

DOE/ET-53088-392

IFSR #392

**Resonances, the Devil's Staircase and Transport  
in Area-Preserving Maps**

*Qi Chen*

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

**August 1989**

# **Resonances, the Devil's Staircase and Transport in Area-Preserving Maps**

Publication No. \_\_\_\_\_

Qi Keith Chen, Ph.D.

The University of Texas at Austin, 1989

Supervising Professor: Philip J. Morrison

Chaotic transport in few degree of freedom Hamiltonian systems is of considerable importance in various fields, including plasma confinement, accelerator physics, intramolecular dynamics, celestial mechanics, and condensed matter physics. In two degree of freedom systems represented by area-preserving maps, we introduce a Markov transport model to describe various statistical properties in the irregular components where there are no invariant tori to prevent global transport. States of the Markov chain are regions delineated by the stable and unstable manifolds of the hyperbolic periodic orbits. We show that resonances give a complete partition of the phase space in the supercritical regime so that almost all points in the phase space are identified with particular resonance states. We then apply the Markov model to a purely chaotic system — the sawtooth map, and derive exact analytic results for transport rates. These are compared with the

numerical rates. We conclude that in the chaotic regime, the Markov model gives reasonably good predictions for transport properties in the irregular components. In order to calculate transport rates for real systems in the chaotic regime, we need to locate highly unstable orbits. We develop a numerical method, the orbit extension method, for finding both unstable ordered periodic orbits and the principal heteroclinic orbits between two resonances. This method actually takes advantage of the instability in the chaotic regime and gives both a stable and an efficient prescription for finding unstable orbits.

**Resonances, the Devil's Staircase and Transport  
in Area-Preserving Maps**

APPROVED BY  
SUPERVISORY COMMITTEE:

Philip A. Morrison

James Z. Meiss

Harry L. Swanson

Robert W. Hyman

Wendell Horton

DEDICATED TO MY PARENTS

## Acknowledgments

First and foremost, I want to express my deep gratitude to my dissertation advisor James Meiss. It has been a great privilege and a rewarding experience to collaborate with and learn from him. He taught me how to think physics and show me the way in the maze of dynamics. Throughout my years of study under him, I have enjoyed his infinite accessibility, his deep knowledge, and his constant support and encouragement.

Secondly, I want to thank Phillip Morrison, for serving as chairman of my dissertation committee. Phil has broadened my knowledge of dynamics and physics and made me see a fuller picture of the universe.

Thirdly, I want to thank Ian Percival, whose insight has been a constant inspiration. It was during my visit at Queen Mary College in the summer of 86 and in the lively research atmosphere of School of Mathematics that I was christened as a dynamicist.

I would like to thank other members of my dissertation committee. I thank Wendell Horton for his continued interests in my research and its application to plasma problems; Harry Swinney and Robert Wyatt for advices and support.

There are other people I have benefitted discussions with and learned from. In particular, I want to thank Robert MacKay, Françoise Argoul, Alain Arnéodo, Dwight Barkley, Roger Broucke, Izthack Dana, Dominique Escande, Hans Koch, Norman Murray, Razi Tavokol, Charles Radin, Franco Vivaldi, Bob Williams, Xiaojin Wang, and Xuehua Yu.

Discussions with fellow graduate students are invaluable for my education. In particular, I want to thank Hyungtae Kook for dynamics, Chris Kueny on various physics applications, and Yongwu Rong for mathematics. Other people come to mind are, Raul Acevedo, Nigel Bird, Changye Chen, Guoyong Fu, and Yangzhong Zhang.

I am grateful to the Institute for Fusion Studies for providing a wonderful independent research environment. I thank the staff at the Institute for various conveniences they provide me, in particular, Dawn East, Buff Miner, Michael Sternberg, Sarolyn Stewart, Carolyn Valentine, and Lynn Williams.

I acknowledge financial support from the Department of Physics from 1984 to 1986; the Institute for Fusion Studies from 1986 until the present; and U.K. SERC and NATO during my visit at Queen Mary College in 1986 and Warwick University in 1987.

## Table of Contents

<b>Acknowledgments</b>	<b>v</b>
<b>Abstract</b>	<b>vii</b>
<b>Table of Contents</b>	<b>ix</b>
<b>List of Figures</b>	<b>xii</b>
<b>1. Introduction</b>	<b>1</b>
<b>2. Hamiltonian Dynamical Systems</b>	<b>8</b>
2.1 Hamiltonian Dynamical Systems	9
2.1.1. Dynamical Systems	9
2.1.2. Hamiltonian Dynamics	12
2.1.3. Variational Principle and Vortex Flows	16
2.1.4. Autonomous and Nonautonomous Systems	22
2.2. Discrete Hamiltonian Systems	25
2.2.1. Derivation of the Return Map from flow	25
2.2.2. Generating Functions	29
2.2.3. Symplectic Algorithms	34
2.3. Invariant Structures and KAM Theorem	38
2.3.1. Integrable Systems	38
2.3.2. KAM Theorem and Its Applications	41
2.3.3. Variational Principle for Invariant Tori and Cantori	47



2.4. Applications	50
2.4.1. Magnetic Field Flow as a Vortex Flow	50
2.4.2. Passive Scalars and Drift Wave Instability	55
2.4.3. Cyclotron	57
<b>3. Resonances and the Devil's Staircase in Area-Preserving Twist Maps</b>	<b>59</b>
3.1. Twist Maps of the Cylinder	61
3.1.1. Twist Maps	61
3.1.2. Action Principle	65
3.1.3. Example	67
3.2. Monotone Sets and Minimizing Orbits	68
3.2.1. Monotone Sets and Frequencies	68
3.2.2. Continued Fraction and Farey Tree	69
3.2.3. Minimizing Orbits	72
3.2.4. Minimax Orbits	75
3.3. Partial Barriers, Turnstiles and Resonances	77
3.3.1. Irregular Components	77
3.3.2. Partial Barriers, Turnstiles and Resonances	79
3.3.3. Flux, Areas and Actions	85
3.4. Relation to Average Lagrangian	94
3.4.1. Average Lagrangian	94
3.4.2. Area Devil's Staircase and Average Lagrangian	97
3.4.3. Implications for $\Delta W(v \rightarrow v')$	102
3.4.4. Application to Calculating the Area under an Invariant Circle	105
3.5. The Markov Transport Model	110
<b>4. Symbolic Dynamics of the Sawtooth Map</b>	<b>115</b>
4.1. Introduction	116
4.2. The Sawtooth Map and Its Orbital Structures	120

4.3. Minimizing and Minimax Orbits	125
4.3.1. Minimizing Periodic Orbits	125
4.3.2. Cantori and Minimizing Homoclinic Orbits	128
4.3.3. Minimax Periodic Orbits	132
4.4. Partial Barriers, Turnstiles and Resonances	136
4.4.1. Partial Barriers and Turnstiles	136
4.4.2. Resonances	138
4.4.3. Generalized Turnstiles	142
4.4.4. Application to Piecewise Linear Standard Map	
4.5. Invariant Sets within Resonances	150
4.6. Escape Dynamics	163
<b>5. Orbit Extension Method for Finding Unstable Orbits</b>	<b>173</b>
5.1. Introduction	174
5.2. Newton's Method	177
5.2.1. Twist Maps	177
5.2.2. Newton's Method	179
5.3. Orbits Extension Method	185
5.3.1. Bunching	186
5.3.2. Ordered Periodic Orbits	190
5.3.3. Principal Heteroclinic Orbits	197
<b>6. Summaries and Conclusions</b>	<b>203</b>
<b>References</b>	<b>207</b>
<b>Vita</b>	

## List of Figures

2.1	Hamiltonian flow preserves the symplectic area.	14
2.2	Conserved circulation under the vortex flow.	20
2.3	Poincaré's integral invariant.	21
2.4	Mapping at a period.	26
2.5	Surface of Section.	27
2.6	Cylindrical Coordinate of the field line flow.	53
2.7	Cyclotron.	57
3.1	Twist condition.	62
3.2	Net Flux.	63
3.3	A rotational invariant circle.	64
3.4	Flux.	79
3.5	Partial barrier of a pair of periodic orbits.	80
3.6	Partial separatrix.	81
3.7	(1,3) resonance of the standard map.	82
3.8	Partial barrier of a cantorus.	84
3.9	Illustration of the fundamental area formula.	85
3.10	Flux between two resonances.	92
3.11	The flux function $\Delta W(v \rightarrow v')$ .	103
4.1	a) Resonances of the sawtooth map.	118
	b) Some resonances of the standard map.	119
4.2	Flux across the (3,8) periodic orbit for the sawtooth map.	137
4.3	Illustration of the calculation of the resonance area.	139
4.4	The area devil's staircase function.	141

4.5	An orbit wandering around the (1,2) resonance.	143
4.6	Schematic illustration of the generalized turnstile.	146
4.7	Flux across the (3,8) periodic orbit for the piecewise linear standard map.	149
4.8	Illustration of the $\mathcal{L}$ - $\mathcal{R}$ coding.	151
4.9	Construction of the horseshoe for trapped orbits.	154
4.10	Trapped periodic orbits in a resonance.	156-159
4.11	Trapped rotational periodic orbits as a function of $\lambda$ .	160
4.12	Number of periodic points trapped in a resonance as a function of the period.	162
4.13	a) Typical survival probability.	166
	b) Typical numerical average escape rate.	167
4.14	Escape rate as a function of the Lyapunov number.	170
4.15	Escape rate as a function of the residue for small residues.	171
<hr/>		
5.1	Configuration points of the (8,21) minimizing orbit as the parameter varies.	187
5.2	Four levels of the orbit extension method.	193
5.3	Phase space plot of the final configuration in Fig. 5.2.	194
5.4	Sketch of points on orbits homoclinic to and heteroclinic between two minimizing periodic orbits.	198
Table 3.1, 3.2 Area under the golden mean invariant circle of the standard map.		106-109

**Resonances, the Devil's Staircase and Transport  
in Area-Preserving Maps**

by

QI KEITH CHEN, B.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

Copyright  
by  
Qi Keith Chen  
1989

# **Chapter 1**

## **Introduction**

Generically, area-preserving maps have a phase space which is divided between regular and irregular components. In the regular components, the dynamics is mostly quasiperiodic rotation on invariant circles called Kolmogorov, Arnol'd and Moser tori; while in the irregular components, the motion is chaotic with positive Lyapunov exponents. The task for transport theory is to study the statistical properties and behaviors of motions within and transitions between regions in the irregular components.

Statistical properties of chaotic motions in an irregular component can be described by a Markov partition, a partition into regions with a list of allowable transitions between them, such that for each allowable transition there is a unique orbit. Indeed Pesin has shown that such a partition exists for components on which the Lyapunov exponents are nonzero almost everywhere [Pesin, 1977]. While we can always assign transition rates between different regions, which may depend on the past history of the orbit. If the transition rates are independent of the past history, such a description is a Markov process.

The key concept proposed by MacKay, Meiss and Percival in their 1983 pioneering work is the concept of partial barriers formed from hyperbolic invariant sets, which impede transport in chaotic regions. Transport from one region to another takes place through turnstiles, the flux regions of partial barriers. Overlaps of a turnstile with various regions gives the allowable transitions. Assuming rapid mixing in the irregular components, the flux from one region to another divided by the area of that region gives the transition rate. Regions correspond to states in the Markov chain. A natural choice of regions is through a delineation of those partial barriers with minimum flux. It turns



out that in the chaotic regime, the minimum flux partial barriers are characterized by the frequencies of ordered hyperbolic invariant sets. Partial barriers are formed from pieces of the stable and unstable manifolds of the corresponding ordered hyperbolic orbits. For an irrational frequencies, such an orbit is a cantor set, an ordered invariant set which is the remnants of the integrable invariant torus. It has a cantor set structure. For a rational frequency, there are two partial barriers with local minimum flux, the upper and lower partial separatrices; these are built upon the upper and lower minimizing homoclinic orbits.

To facilitate practical calculations of transport rates, discrete Markov models are proposed. There are basically two choice of states: Regions bounded by the partial barriers of cantori, or regions bounded by the rational partial separatrices. In the former case, since irrational numbers are uncountable, there are an uncountably many states in the model. This makes the model intractable in practice. A discrete approximation can be introduced by observing that "noble" cantori usually have local minimum flux across them, so states can be chosen as regions bounded by partial barriers with noble frequencies (the noble numbers are countable).

Near criticality, flux through cantori obey scaling relations, this is reflected in the actual transport properties, for instance, the diffusion coefficient. However, far from criticality, the choice of the noble partial barriers does not seem to be useful to describe the actual transport processes. This is because, unlike the critical case, we can no longer restrict only to the noble cantori; in fact, all cantori have a large flux, and usually these are all of the same order of magnitude. Therefore, all cantori should be treated

"democratically"; the choice of noble cantori leaves out too much of the details and structure in the actual transport process. However, it is unclear how to choose a suitable, but countable set of cantori to partition the phase space in this case.

A more natural partition of phase space is to use partial separatrices. They are formed from pieces of the stable and unstable manifolds of the hyperbolic periodic orbits. Each rational number defines two partial separatrices, upper and lower. They have common end points at the corresponding ordered hyperbolic periodic orbit. The region delineated by the upper and lower partial separatrices is called the resonance for its kinship with a resonance in an integrable system. A resonance is uniquely characterized by its rational frequency. Resonances are countable, and according to both perturbation theory and numerical results, the size of a resonance usually decreases exponentially with the level of the rational frequency in the Farey tree. So a finite state approximation of the Markov chain is obtained upon neglecting exponentially small regions in the phase space.

In the supercritical regime where there are no rotational invariant tori to prevent "global" transport, resonances must take over the whole phase space so that almost all points in the phase space are identified with some state in the Markov chain and a reasonably good statistical description of the real transport process is given by the Markov model. A more precise reformulation of the idea of the complete partition of phase space by resonances is that the area under an invariant set as a function of its frequency is a complete devil's staircase.

In this dissertation, we show first that the area staircase is indeed complete in the supercritical regime. We then apply the Markov transport model to a chaotic area-preserving map — the sawtooth map, which allows exact calculation of transport rates. Through comparison of numerical studies and predictions of the Markov model, we show that for large Lyapunov exponents, the Markov model gives a good and sometimes an exact description of the transport process. However, as one approaches the integrable system, the discrepancies become larger and the transport process is no longer completely accounted for by the model. We also develop a numerical technique, the orbit extension method, for finding highly unstable orbits. This is necessary for the calculation of the actual transport rates.

One topic not investigated in this dissertation is the effect of boundary circles on the Markov model. Boundary circles are invariant circles at the threshold of destruction. Renormalisation theory predicts there is a critical scaling for boundary circles; however, a complete understanding of these renormalisation properties is still lacking. Scaling properties near hierarchical boundary circles are believed to be the prime reason for the long time correlations in the correlation functions in the chaotic regimes.

This dissertation is organized as follows. Chapter 2 is a brief review of Hamiltonian dynamical systems. I include some basic concepts and techniques, and a few applications of Hamiltonian dynamics. The material covered is not complete. Some of the neglected topics are: Poisson dynamics, perturbation techniques, bifurcation theory, linear stability, and the Melnikov method.

In chapter 3, we review some basic facts of area-preserving twist maps of the cylinder. Based on Aubry-Mather theory, we discuss the recently developed concepts of partial barriers, turnstiles and resonances and define the area devil's staircase function. We relate the area function to the devil's staircase discovered by Aubry [1983a] in the study of the commensurate-incommensurate transition in the classical Frenkel-Kontorova model. As a corollary, we show that resonances give a complete partition of phase space in the supercritical regime. This result also gives a practical way to calculate the area under an invariant circle, a quantity which is useful in applications.

Chapter 4 is a case study of a purely hyperbolic system, the sawtooth map. For this system, we analytically construct partial barriers and resonances, and calculate transport rates. The motivation is to pin down the effectiveness of the Markov transport model in irregular components, and to get around technical difficulties of mixed systems which complicate the testing of the assumptions, especially the determination of the locations and properties of boundary circles. We discuss symbolic characterizations of orbits, construct partial barriers and resonances, derive analytic formulas for flux and resonance areas, and study escape dynamics from a single resonance. We compare predictions of the Markov model with numerical results, and conclude that the Markov model gives good predictions for transport rates in the chaotic regime. However, the theory no longer completely accounts for the transport when the system is nearly integrable.

Chapter 5 is a discussion of numerical techniques for finding various orbits important in the calculation of transport rates. These orbits are typically highly unstable, indeed the

positivity of the Lyapunov exponent implies sensitive dependence on initial conditions and is a key impediment to devising stable numerical prescriptions. We develop the orbit extension method, for finding ordered unstable periodic orbits and the principal heteroclinic orbits between two resonances. This method actually takes advantage of the instability and gives both a stable and an efficient prescription for finding unstable orbits.

**Chapter 2**  
**Hamiltonian Dynamical Systems**

## 2.1 Hamiltonian Dynamical Systems

Hamiltonian dynamics has a paramount importance to the study of Nature. As Newton says in his Principia, "the laws which we have explained abundantly serve to account for all the motions of the celestial bodies and of our sea".

### 2.1.1. Dynamical Systems

Dynamics is concerned with the motion of systems, that is, their change of state with time  $t$ . The past and future of a deterministic dynamical system is completely determined by its present state. The state of the system is described by a set of first order differential equations

$$\frac{d\mathbf{x}}{dt} = \mathbf{v}(\mathbf{x}) \tag{2.1.1}$$

where  $\mathbf{x}$  is a point on an  $n$ -dimensional  $C^{r+1}$  differentiable manifold  $\mathcal{M}$  which we call the phase space of the system (A  $C^{r+1}$  manifold is a topological space where each point has a neighborhood which is  $C^{r+1}$  diffeomorphic to  $\mathbb{R}^n$ , here we require  $r \geq 1$ ),  $\mathbf{v}(\mathbf{x})$  is a  $C^r$  smooth vector field on the manifold. Equation (2.1.1) are the equations of motion of the system. The vector field can also be time dependent; in the extended phase space  $\mathcal{M} \times \mathcal{R}$ , it can be treated in a similar fashion as the time independent vector field.

For example, the equations describing the earth orbiting the sun in polar coordinates are

$$\frac{d^2x}{dt^2} - r \frac{d\theta}{dt} = \frac{GM_s}{r^2}$$

$$\frac{d}{dt} (r^2 \frac{d\theta}{dt}) = 0 \tag{2.1.2}$$

where  $G$  is the gravitational constant and  $M_s$  is the mass of the sun. Though (2.1.2) is a set of two second order equations, it is equivalent to four first order equations. In fact any set of higher order differential equations is equivalent to a set of first order differential equations, so there is no loss of generality in restricting to first order equations.

Let  $x$  be the initial condition at time 0. The solution  $g^t x$  of equation (2.1.1) defines a map  $g^t$  from  $\mathcal{M}$  into itself. The basic theorem of the theory of ordinary differential equations states that locally the vector field  $v$  is diffeomorphic to a constant vector field at any nonsingular point  $v \neq 0$ . Therefore, locally this map exists and is unique, and is a diffeomorphism, since constant vector field has a unique integral curve passing through a given point. The  $C^1$  smooth condition for the vector field also guarantees the uniqueness of the integral curve at singular points.

The local phase flow  $g^t$  forms a local one parameter group of diffeomorphisms

$$\begin{aligned} 1) & g^0 = \text{identity}, \\ 2) & g^{t+s} = g^t g^s \text{ for } s, t \text{ small enough.} \end{aligned} \tag{2.1.3}$$

By patching solutions in small neighborhoods in the phase space, the solution can be extended uniquely to a finite or an infinite time interval. In the finite extension case, the motion is undefined at some finite time, and is referred to as terminating.



For vector fields with compact support, the solution can be extended bi-infinitely. Such solution  $g^t$  is called a phase flow and forms a one parameter group of diffeomorphisms.

### 2.1.2. Hamiltonian Dynamics

Hamilton's equations describe motion in a even dimensional symplectic phase space  $\mathcal{M}^{2n}$ , where  $n$  is the number of degree of freedom (d.o.f) of the system. The coordinate chart is usually parametrized by  $\mathbf{z} = (\mathbf{q}, \mathbf{p}) = (q^1, q^2, \dots, q^n, p^1, p^2, \dots, p^n)$ , called the symplectic (or canonical) coordinates. Hamilton's equations are

$$\begin{aligned} \frac{d\mathbf{q}}{dt} &= \frac{\partial H(\mathbf{q}, \mathbf{p})}{\partial \mathbf{p}} \\ \frac{d\mathbf{p}}{dt} &= -\frac{\partial H(\mathbf{q}, \mathbf{p})}{\partial \mathbf{q}} \end{aligned} \quad (2.1.4)$$

or

$$\frac{d\mathbf{z}}{dt} = \mathbf{J}dH(\mathbf{z}) \quad (2.1.5)$$

where  $\mathbf{J}$  is called the cosymplectic tensor; it maps a differential one form to a vector field. In symplectic coordinates

$$\mathbf{J} = \begin{pmatrix} 0 & \mathbf{I} \\ -\mathbf{I} & 0 \end{pmatrix} \quad (2.1.6)$$

$H$  is the Hamiltonian function. The vector field  $\mathbf{J}dH$  is called the Hamiltonian vector field.

A symplectic manifold  $(\mathcal{M}^{2n}, \omega)$  is an even dimensional manifold equipped with a closed nondegenerate differential two form  $\omega$

$$\omega = \sum_{ij} \omega_{ij} dz^i \wedge dz^j \quad (2.1.7)$$

A two form is nondegenerate if  $\omega(\xi, \eta) = 0 \forall \eta \Leftrightarrow \xi = 0$  for  $\xi, \eta$  in the tangent space  $T\mathcal{M}_{\mathbf{z}}$ . A nontrivial example of a symplectic manifold is a two dimensional torus, where the symplectic form is the area form. In symplectic coordinates,  $\omega$  is given by

$$\omega = \sum_{i=1}^n dp^i \wedge dq^i \equiv \mathbf{dp} \wedge \mathbf{dq} \quad (2.1.8)$$

which implies that the matrix

$$\omega = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

The two form  $\omega$  associates each vector  $\xi$  in the tangent space  $T\mathcal{M}_{\mathbf{z}}$  at point  $\mathbf{z}$  to a one form in the cotangent space  $T^*\mathcal{M}_{\mathbf{z}}$  by

$$\gamma_{\xi} = \omega(\cdot, \xi) = \sum_{i,j} \omega_{ij} \xi^j dz^i \quad (2.1.9)$$

Since  $\omega$  is nondegenerate, this map in fact is an isomorphism. Therefore, every one form is also associated uniquely with a vector. We shall denote the isomorphism by  $J: T\mathcal{M}_{\mathbf{z}} \rightarrow T^*\mathcal{M}_{\mathbf{z}}$ .  $J$  is the cosymplectic tensor and is nondegenerate

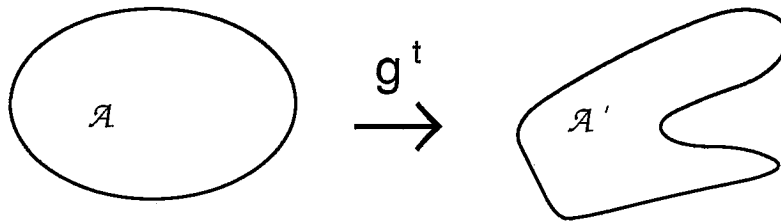
$$J_{ij} = (\omega^{-1})_{ij} \quad (2.1.10)$$

The most salient feature of the Hamiltonian flow  $g^t$  is the preservation of the symplectic area (Fig. 2.1)

$$\int g^t_{\mathcal{A}} \omega = \int_{\mathcal{A}} \omega \quad (2.1.11)$$

Where  $\mathcal{A}$  is a two chain (a two dimensional surface), or equivalently

$$(g^t)^* \omega = \omega \quad (2.1.12)$$



$$\text{Area}(\mathcal{A}) = \text{Area}(\mathcal{A}')$$

Figure 2.1. Hamiltonian flow preserves the symplectic area

In particular, the phase space volume  $\omega^n$  is conserved under the flow; this is Liouville's Theorem. In fact, the existence of a Hamiltonian function is not necessary for this property: any Locally Hamiltonian vector field (the vector field  $J\gamma$  generated by a close one form  $\gamma$ ) preserves the symplectic form. This is easily verified algebraically by noting that the Lie derivative of the symplectic form  $\omega$  under the flow vanishes. We give another proof in section 2.1.3.

Darboux has shown that locally a symplectic manifold has a normal form: in a neighborhood of a points  $\mathbf{z}$  in the symplectic manifold  $\mathcal{M}^{2n}$ , one can always choose a symplectic coordinate

$(\mathbf{q}, \mathbf{p}) = (q^1, q^2, \dots, q^n, p^1, p^2, \dots, p^n)$  such that the symplectic form  $\omega$  has the standard form (2.1.8).

### 2.1.3. Variational Principles and Vortex Flows

The Hamiltonian function in (2.1.4) can also be time dependent. In this case the equations are essentially unchanged

$$\begin{aligned}\frac{d\mathbf{q}}{dt} &= \frac{\partial H(\mathbf{q}, \mathbf{p}, t)}{\partial \mathbf{p}} \\ \frac{d\mathbf{p}}{dt} &= - \frac{\partial H(\mathbf{q}, \mathbf{p}, t)}{\partial \mathbf{q}}\end{aligned}\tag{2.1.13}$$

Equations (2.1.13) can be cast into a variational form, the principle of least action in phase space

$$\delta \int \gamma = 0\tag{2.1.14}$$

where

$$\gamma = \sum_{i=1}^n p^i dq^i - H dt \equiv \mathbf{p} d\mathbf{q} - H dt\tag{2.1.15}$$

under the free boundary condition

$$\gamma = 0 \text{ at end points}$$

Suppose we make a coordinate transformation from  $(\mathbf{q}, \mathbf{p}, t)$  to mixed variables  $\mathbf{z} = (z^0, z^1, \dots, z^{2n})$

$$\begin{aligned}p^i &= p^i(z^0, z^1, \dots, z^{2n}) \\ q^i &= q^i(z^0, z^1, \dots, z^{2n}) \\ t &= t(z^0, z^1, \dots, z^{2n})\end{aligned}\tag{2.1.16}$$

Then the one form (2.1.5) becomes

$$\gamma = \sum_{\mu=0}^{2n} \gamma_{\mu} dz^{\mu} \quad (2.1.17)$$

Now all coordinates are equivalent: formally there is no distinct time coordinate. Thus it is natural to work in an odd-dimensional extended phase space  $\mathcal{M}^{2n \times \mathcal{R}}$  in a covariant fashion. Here  $\gamma$  is called the fundamental one form of the Hamiltonian flow. The fundamental one form is defined up to a gauge transformation

$$\sum_{\mu=0}^{2n} \gamma'_{\mu} dz^{\mu} = \sum_{\mu=0}^{2n} \gamma_{\mu} dz^{\mu} + dS \quad (2.1.18)$$

since the variation vanishes for an exact one form,  $dS$ .

The equations of motion are given by taking the variation of (2.1.15). This is facilitated by introducing a parameter  $\lambda$ , which acts as a formal "time" along the flow, so that the variational integral can be written in Lagrange form

$$\delta \int L d\lambda = 0 \quad (2.1.19)$$

where

$$L = \sum_{\mu} \gamma_{\mu} \frac{dz^{\mu}}{d\lambda} \quad (2.1.20)$$

The Euler-Lagrange equations obtained from (2.1.19) and (2.1.20) are

$$\sum_{\nu} \omega_{\mu\nu} \frac{dz^{\nu}}{d\lambda} = 0 \quad (2.1.21)$$

where

$$\omega_{\mu\nu} = \gamma_{\nu,\mu} - \gamma_{\mu,\nu} = (d\gamma)_{\mu\nu} \quad (2.1.22)$$

This is called the phase space Lagrangian formulation. Notice that our Lagrangian has twice as many independent variables as the number of coordinates in the extended phase space, and is always linear in "time" derivative. Although this formulation is more general than the usual Lagrangian formulation of Hamiltonian flow, such as Hamilton's principle of least action, when the fundamental form is given by (2.1.15), it is equivalent to Hamilton's principle (Arnol'd,1978).

In order that equations (2.1.21) define a unique directional flow, the two form  $d\gamma$  must have a unique null direction at every point in space  $\mathbf{v}$ :  $d\gamma(\cdot, \mathbf{v}) = 0$ , this requires that the rank of  $d\gamma$  is  $2n$  at every point. Such a two form is called nonsingular. The unique flow defined by the null direction of  $d\gamma$  is called the vortex flow. There are  $2n+1$  equations in (2.1.21), of which  $2n$  are independent. This is easily seen if we let  $\lambda = z^0$  and write down the spatial and "temporal" components of (2.1.21): the "temporal" component is a consequence of spatial components. This is a manifestation of the fact that the rank of  $d\gamma$  is  $2n$ .

For the Hamiltonian flow in canonical coordinates, the fundamental one form is given by (2.1.15), so

$$\omega = dp \wedge dq - dH \wedge dt \quad (2.1.23)$$



which is obviously nonsingular. The unique null direction in extended phase space is given by  $\mathbf{v} = (H_{\mathbf{p}}, -H_{\mathbf{q}}, 1)$ , which is the usual Hamiltonian vector field.

Symmetries relate to conservation laws in the vortex flow formulation since it is in essence a Lagrangian system. The simplest version of Noether's theorem states that if the components of the fundamental one form are independent of the coordinate, say  $z^\alpha$ , then  $\gamma_\alpha$  is an invariant of the flow. This is easy to verify from the equations of motion

$$\frac{d\gamma^\alpha}{d\lambda} = \sum_{\mu} \gamma_{\alpha\mu} \frac{dz^\mu}{d\lambda} = - \sum_{\mu} \omega_{\alpha\mu} \frac{dz^\mu}{d\lambda} = 0 \quad (2.1.24)$$

There is a conserved circulation under the flow (Fig. 2.2), the integral invariant of Poincaré-Cartan. Consider a closed curve  $c^1$ , everywhere transverse to the vortex flow direction  $\mathbf{v}$ . Under the flow, this curve forms a tube called the vortex tube. Any curve  $c^2$  which encircles the same vortex tube to  $c^1$  under the flow, since  $c^1 - c^2 = \partial s$ , where  $s$  is a two cycle representing a section of the vortex tube. Then

$$\int_{c^1} \gamma = \int_{c^2} \gamma \quad (2.1.25)$$

This follows from Stokes' theorem:

$$\int_{c^1} \gamma - \int_{c^2} \gamma = \int_{\partial s} \gamma = \int_s d\gamma$$

However by definition of the flow, the value of  $d\gamma$  vanishes on any infinitesimal two plane containing the flow vector  $\mathbf{v}$ , therefore the integral on the right hand side is equal to zero.

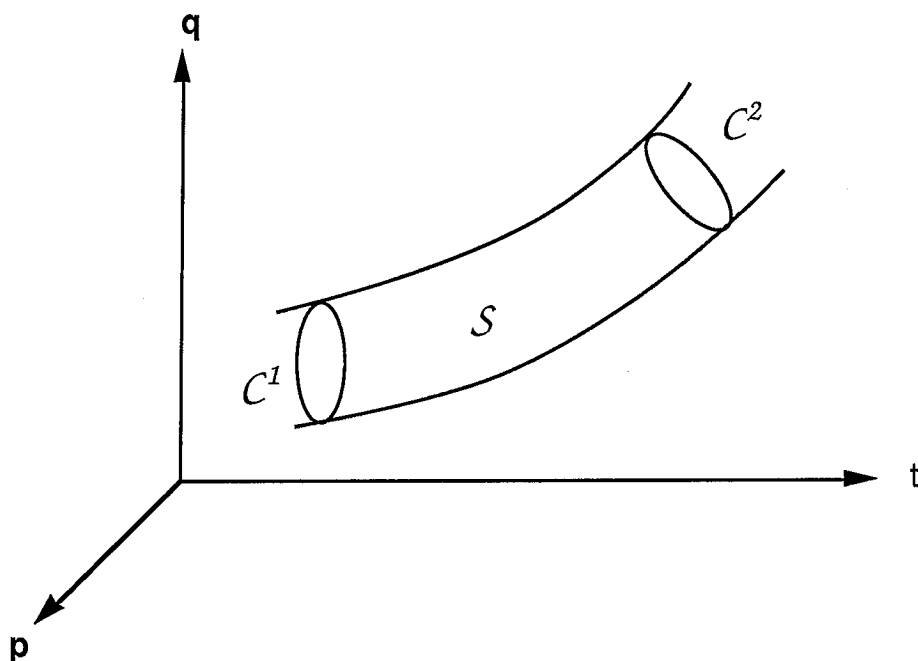


Figure 2.2. Conserved circulation under the vortex flow.

For Hamiltonian flow in canonical coordinates, the fundamental one form is (2.1.15). If we choose  $C_1$  to be any closed curve in the constant time phase space  $\mathcal{M}^{2n}$ , and  $C_2$  its time evolution after time  $t$  (Fig. 2.3), we get the Poincaré invariant for time dependent Hamiltonian flow (2.1.11) or (2.1.12).

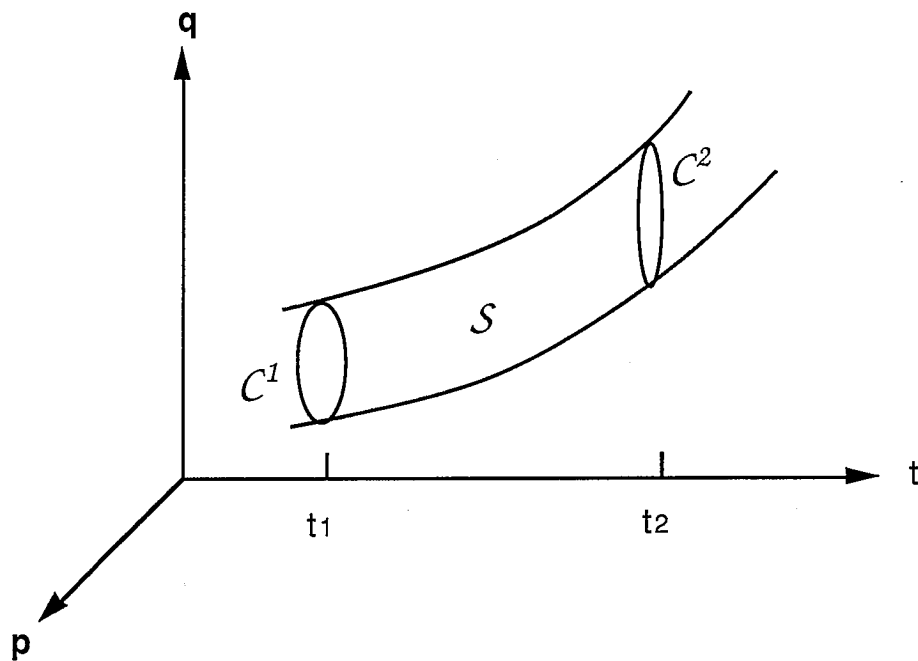


Figure 2.3. Poincaré's integral invariant

### 2.1.4. Autonomous and Nonautonomous Hamiltonian Systems

If the Hamiltonian function is independent of the time, the system is called autonomous; otherwise, the system is nonautonomous. In the following, we show that in canonical coordinates, an  $n$  d.o.f nonautonomous system is equivalent to an  $n+1$  d.o.f autonomous system with and vice versa. This fact allows us to apply theorems on autonomous systems to nonautonomous systems.

Consider a Hamiltonian flow given by the Hamiltonian function  $H(\mathbf{q}, \mathbf{p}, t)$ , where  $(\mathbf{q}, \mathbf{p}) = (q^1, q^2, \dots, q^n, p^1, p^2, \dots, p^n)$ . Define new canonical coordinates  $(\mathbf{q}', \mathbf{p}') = (q'^1, q'^2, \dots, q'^n, p'^1, p'^2, \dots, p'^{n+1}) = (q^1, q^2, \dots, q^n, t, p^1, p^2, \dots, p^n, -E)$ , and the new Hamiltonian function

$$H'(\mathbf{q}', \mathbf{p}') = H(\mathbf{q}, \mathbf{p}, t) - E \quad (2.1.26)$$

It is easily verified that equation (2.1.13) is now equivalent to

$$\begin{aligned} \frac{dq'}{dt} &= \frac{\partial H(\mathbf{q}', \mathbf{p}')}{\partial p'} \\ \frac{dp'}{dt} &= -\frac{\partial H(\mathbf{q}', \mathbf{p}')}{\partial q'} \end{aligned} \quad (2.1.27)$$

The above result is not surprising since any  $2n$  dimensional time dependent flow is equivalent to a  $2n+1$  dimensional time independent flow. An additional dimension is introduced in order to preserve the symplectic structure.

The converse result is also of interest and useful. Given a time independent  $n$  d.o.f. Hamiltonian function  $H(\mathbf{q}, \mathbf{p})$ , then  $H$  is conserved. We can choose any generalized coordinate as the new "time" variable and the conjugate coordinate represents the new "time" dependent Hamiltonian. For example, given

$$H(\mathbf{q}, \mathbf{p}) = E \quad (2.1.28)$$

Let  $(\mathbf{q}', \mathbf{p}') = (q^1, q^2, \dots, q^{n-1}, p^1, p^2, \dots, p^{n-1}) = (q^1, q^2, \dots, q^{n-1}, p^1, p^2, \dots, p^{n-1})$ ,  $\tau = q^n$  and  $H'(\mathbf{q}', \mathbf{p}', \tau) = -p^n(q^1, q^2, \dots, q^{n-1}, p^1, p^2, \dots, p^{n-1}, q^n, E)$ , where we solve  $p^n$  from Eq(2.1.28). Then

$$\begin{aligned} \frac{d\mathbf{q}'}{d\tau} &= \frac{\partial H'(\mathbf{q}', \mathbf{p}', \tau, E)}{\partial \mathbf{p}'} \\ \frac{d\mathbf{p}'}{d\tau} &= -\frac{\partial H'(\mathbf{q}', \mathbf{p}', \tau, E)}{\partial \mathbf{q}'} \end{aligned} \quad (2.1.29)$$

where  $E$  serves as a parameter.

In  $(\mathbf{q}, \mathbf{p}) = (q^1, q^2, \dots, q^n, p^1, p^2, \dots, p^n)$  coordinates, the fundamental one form is

$$\gamma = \sum_{i=1}^n p^i dq^i - H(\mathbf{p}, \mathbf{q}) dt$$

In the new coordinates  $(q^1, q^2, \dots, q^{n-1}, p^1, p^2, \dots, p^{n-1}, q^n, E)$ , the fundamental form becomes

$$\gamma = \sum_{i=1}^{n-1} p^i dq^i - E dt - H'(\mathbf{p}', \mathbf{q}', \tau, E) d\tau$$

The equations of motion are

$$\frac{dq^i}{d\tau} = \frac{\partial H'(\mathbf{q}', \mathbf{p}', \tau, E)}{\partial p^i}$$

$$\frac{dp^i}{d\tau} = - \frac{\partial H'(\mathbf{q}', \mathbf{p}', \tau, E)}{\partial q^i} \quad (i=1, \dots, n-1) \quad (2.1.30)$$

$$\frac{dE}{d\tau} = 0$$

$$\frac{dt}{d\tau} = - \frac{\partial H'(\mathbf{q}', \mathbf{p}', \tau, E)}{\partial E} \quad (2.1.31)$$

The last pair of equations are exactly conservation of energy and the equation of motion for the coordinate  $q^n$ .

## 2.2. Discrete Hamiltonian Systems

A discrete Hamiltonian system is a symplectic (or canonical) map which maps a symplectic phase space into itself and preserves the symplectic two form. A symplectic map arises naturally from the successive intersections of a Hamiltonian flow trajectory with a surface of section. Reducing a flow to a map is a great conceptual simplification: the dynamics of the map contains all the qualitative features of the flow and some quantitative features near the surface of section. Furthermore, the dimensionality of the phase space is reduced by one, so that both mathematical proofs and numerical calculations of properties of trajectories can be approached more conveniently and efficiently.

### 2.2.1. Derivation of the Return Map from Flow

Consider a Hamiltonian phase flow  $g^t$  and introduce a formal partition of the time interval  $\{ \dots, t_{-n}, \dots, t_{-1}, t_0, t_1, \dots, t_n, \dots \}$ . Then the map  $T_i = g_{t_i}^{t_{i+1}}$ , where  $g_{t_i}^{t_{i+1}}$  denotes the flow from time  $t_i$  to  $t_{i+1}$ , is a symplectic map, since the flow preserves the symplectic form. It is important that the solution can be extended to all real time, so that successive applications of the maps  $T_i$  are defined for all integer  $i$ .

For time periodic flow with period  $\tau$ , choose the crossing time as  $\{ \dots, -n\tau, \dots, -\tau, 0, \tau, \dots, n\tau, \dots \}$ . The return map  $T_i$  is then time independent (independent of  $i$ ); therefore, we reduce the dynamics to a lower dimension. Since the flow is periodic, it is time translational invariant under  $\tau$ , hence satisfies  $g_{t_1+\tau}^{t_2+\tau} =$

$g_{t_1}^{t_2}$ . therefore  $g_0^{n\tau} = (g_0^\tau)^n$ . At any time  $i\tau$ ,  $i \in \mathbb{Z}$ , the return map  $T_i$  is given uniquely by  $g_0^\tau$  (Figure 2.4).

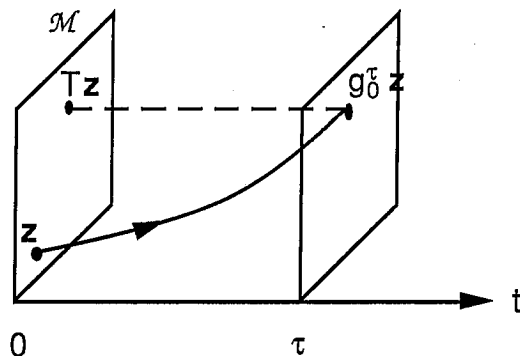


Figure 2.4. Mapping at a period.

For example, if the flow is a free motion except at time  $i\tau$ ,  $i \in \mathbb{Z}$ , when it receives a kick  $\mathbf{f}(\mathbf{q})$ , then the map derived from the flow has the form

$$\begin{aligned} \mathbf{p}' &= \mathbf{p} + \mathbf{f}(\mathbf{q}) \\ \mathbf{q}' &= \mathbf{q} + \mathbf{p}' \end{aligned} \tag{2.2.1}$$

where  $\mathbf{q}'$ ,  $\mathbf{p}'$  are the position and momentum immediately before the next kick. If  $\mathbf{f}(\mathbf{q})$  is the gradient of some potential function,  $\mathbf{f}(\mathbf{q}) = -\nabla V(\mathbf{q})$ , then this gives a class of symplectic maps and techniques used for continuous Hamiltonian flow, e.g. perturbation theory can be applied to this class of systems.

For an  $n$  d.o.f. autonomous system with a periodic solution, we can reduce the dynamics in the neighborhood of the periodic solution from a  $2n-1$  dimensional energy surface to a local map around that solution in a  $2n-2$  dimensional subspace. By "local"



we mean the map is defined only in a neighborhood of the periodic solution. The symplectic form in this subspace is again preserved. We prove this in the following.

Consider a point  $\mathbf{z}_0$  of the periodic solution on the  $2n-1$  dimensional energy surface and define any  $2n-2$  dimensional surface transverse to the velocity vector  $\mathbf{v}(\mathbf{z}_0)$  in (2.1.10) as a cross section. Since  $\mathbf{z}_0$  is periodic, from any point  $\mathbf{z}$  sufficiently near to  $\mathbf{z}_0$  on a cross section, the flow will return to the section at  $T\mathbf{z}$  after some finite time. This defines a local map on the cross section, the technique is called Poincaré surface of section. The return map is obviously independent of the crossing time  $n$ , and we show it also preserves the symplectic area on the cross section.

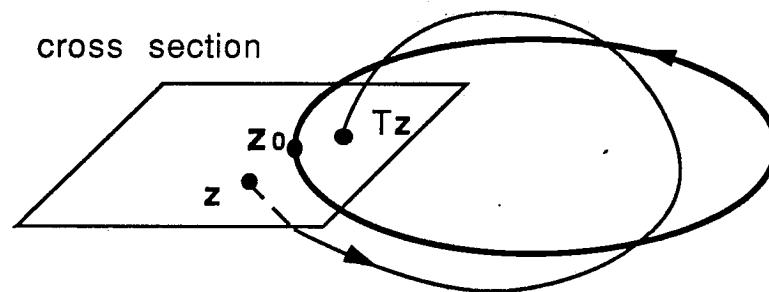


Figure 2.5. Surface of section.

Since the flow is identical to an  $n-1$  d.o.f. nonautonomous flow, near the periodic solution, we can always choose  $q^n$  as the new time coordinate. The fundamental one form is

$$\gamma = \sum_{i=1}^{n-1} p^{i,j} dq^i - H' dt' = \sum_{i=1}^n p^i dq^i$$

By Poincaré-Cartan invariance, the one form is preserved for any cross section transverse to the flow, hence

$$\int_{C^1} \gamma = \int_{C^2} \gamma$$

or

$$\int_{T\mathcal{A}} \omega^2 = \int_{\mathcal{A}} \omega^2$$

This shows the symplectic area is preserved.

### 2.2.2. Generating Functions

A symplectic map  $T$  is a map from a symplectic manifold onto itself which preserves the symplectic form

$$(q', p') = T(q, p)$$

$$T^* \omega = \omega \quad (2.2.2)$$

In canonical coordinates Eq(2.2.2) is equivalent

$$(DT)^t J DT = J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \quad (2.2.3)$$

where  $( )^t$  denotes the transpose of a matrix. This means that the matrix corresponding to the derivative of the map  $DT$  is a symplectic matrix

If the map satisfies the nondegeneracy condition

$$\det \left( \frac{\partial q'}{\partial p} \right) \neq 0$$

then there is an action formulation for the symplectic map, i.e., there exists a local generating function such that

$$\begin{aligned} p &= - \frac{\partial F(q, q')}{\partial q} \\ p' &= \frac{\partial F(q, q')}{\partial p_{q'}} \end{aligned} \quad (2.2.4)$$

or

$$\mathbf{p}'d\mathbf{q}' - \mathbf{p}d\mathbf{q} = dF(\mathbf{q},\mathbf{q}') \quad (2.2.5)$$

The generating function  $F(\mathbf{q},\mathbf{q}')$  is defined up to an arbitrary constant by

$$F(\mathbf{q},\mathbf{q}') = \int_{\mathbf{Q},\mathbf{Q}'}^{\mathbf{q},\mathbf{q}'} \mathbf{p}'d\mathbf{q}' - \mathbf{p}d\mathbf{q}$$

Where  $\mathbf{p}(\mathbf{q},\mathbf{q}')$  is defined locally by the nondegeneracy condition on  $\partial\mathbf{q}'/\partial\mathbf{p}$ , and  $\mathbf{p}'(\mathbf{q},\mathbf{q}')$  by using the map. The point  $(\mathbf{Q},\mathbf{Q}')$  is arbitrary and the integral is path independent because the map is symplectic. The nondegeneracy condition for the action function is

$$\det \left( \frac{\partial^2 F}{\partial \mathbf{q} \partial \mathbf{q}'} \right) \neq 0$$

In this case, an orbit of the map is represented completely by its configuration  $\{ \mathbf{q}_i, i \in \mathbb{Z} \}$ . If there exists a global generating function, the equations of motion of the map can be reformulated by setting the variation of the action

$$W(\{\mathbf{q}_i\}) = \sum_{i=M}^N F(\mathbf{q}_i, \mathbf{q}_{i+1}) \quad (2.2.6)$$

to zero with fixed ends  $\mathbf{q}_M$  and  $\mathbf{q}_N$ . The resulting Euler-Lagrangian equations are

$$F_1(\mathbf{q}_i, \mathbf{q}_{i+1}) + F_2(\mathbf{q}_{i-1}, \mathbf{q}_i) = 0 \quad (2.2.7)$$

where the subscript 1 and 2 denote derivatives with respect to the first and second argument of the generating function, respectively. Therefore, by Eq(2.2.4) the two momentum agree at each point of the orbit. The action formulation plays a central role in the analysis of symplectic maps.

For symplectic return maps derived from flows given by the variational principle (2.1.14), there is a relation between the generating function of the discrete map and the action of the flow. Take the variation of the integral (2.1.15) with time at both ends fixed and allowing configuration points to vary, Then the map  $T_i$  is generated by

$$\mathbf{p}'d\mathbf{q}' - \mathbf{p}d\mathbf{q} = dF_i(\mathbf{q}', \mathbf{q})$$

where

$$F_i(\mathbf{q}', \mathbf{q}) = \text{sta} \int_{\substack{\mathbf{q}(t_{i+1}) = \mathbf{q}' \\ \mathbf{q}(t_i) = \mathbf{q}}} \mathbf{p}d\mathbf{q} - Hdt \quad (2.2.8)$$

sta means the integration is taken over the stationary trajectory from  $\mathbf{q}$  to  $\mathbf{q}'$ . If the flow is periodic and the strobing time is taken as multiples of the period, the generating function is independent of  $i$ .

For the local return map derived from an autonomous flow, the generating function is again independent of  $i$ . Now the end points are chosen to be at the intersection of the energy surface

and the cross section, which is independent of the choice of the crossing  $i$ . One simple case is when we have an angular coordinate  $q^n$  for the flow, the generating function is

$$F(\mathbf{x}', \mathbf{x}) = \int_{\mathbf{x}, q^n}^{\mathbf{x}', q^{n+1}} \mathbf{p} d\mathbf{q} - H dt \quad (2.2.9)$$

where  $\mathbf{x}$  denotes the  $n-1$  components of  $\mathbf{q}$  excluding  $q^n$ . For integrable and near integrable systems, there exists such an angle coordinate.

Therefore theoretical results of symplectic maps can be applied to the flow, in principle. However, in practice, this requires us to do the integral in (2.2.8) or (2.2.9), which is usually difficult unless it is done numerically. A symplectic numerical algorithm is discussed in section 2.2.3.

For example, for one d.o.f. time periodic flows, flux and area have a straight forward meaning and can be obtained by applying the translation formula (2.2.8). However, for autonomous flow, the reduction to a symplectic map relies on the existence of periodic solution, and the map only characterizes local structures near the periodic motion. The existence of such a periodic orbit is not known a priori. In fact, even for two d.o.f. flow, the orbital structure is not completely understood. The motion depends on the topology of the energy surface, which is a three dimensional manifold (there are basically eight distinct topological structures of three dimensional manifolds). The notion of flux and area in two dimensional map need to be generalized in this case. MacKay and Meiss have shown how to generalize flux to two d.o.f. flow

[MacKay and Meiss 87]; however, the three dimensional resonance regions and their relation to the resonances of two dimensional twist maps remain to be clarified.

### 2.2.3. Symplectic Algorithms

A symplectic integration algorithm (SIA) is a numerical scheme which preserves the symplectic nature of the flow. Symplectic integration is important for practical applications; for instance, in tokamak experiments, it is important to know the location of the magnetic islands. This is usually done by numerically integrating the magnetic field line using the Runge-Kutta method; however, This scheme does not preserve the Hamiltonian nature of the field line flow. Therefore in the long run, the magnetic field is no longer divergence free, and no longer reflects the qualitative features of the magnetic field in the tokamak — a monopole field is quite different from a sourceless field.

Symplectic integration was first studied by De Vogelaere. A systematic approach was achieved through the work of Feng [Feng 1986] and Channell [Channel and Scoval 1988]. We shall follow the latter.

Consider the Hamiltonian flow given by (2.1.13), starting from  $(\mathbf{q}_0, \mathbf{p}_0)$  at time 0. We want to find the phase space position after a time step  $\delta$ . The initial position is mapped to the current position by a symplectic map  $g^\delta$ . To devise an integrator which is symplectic and approximates  $g^\delta$  order by order in  $\delta$ , we seek an order by order canonical transformation.

Consider for example, a generating function  $K(\mathbf{q}, \mathbf{p}_0, t)$ , the transformation is



$$\mathbf{q}_0 = -\frac{\partial H(\mathbf{q}, \mathbf{p}_0, t)}{\partial \mathbf{p}_0}$$

$$\mathbf{p} = -\frac{\partial H(\mathbf{q}, \mathbf{p}_0, t)}{\partial \mathbf{q}} \quad (2.2.8)$$

Expand the generating function order by order in  $\delta$ ; the zeroth order is the identical transformation

$$K(\mathbf{q}, \mathbf{p}_0, t) = -\mathbf{p}_0 \mathbf{q} + \delta K_1(\mathbf{q}, \mathbf{p}_0, t) + \delta^2 K_2(\mathbf{q}, \mathbf{p}_0, t) + \dots$$

Expanding  $(\mathbf{q}, \mathbf{p})$  in power series of  $\delta$  in terms of  $(\mathbf{q}_0, \mathbf{p}_0)$  gives

$$\mathbf{q} = \mathbf{q}_0 + \mathbf{q}_1 \delta + \mathbf{q}_2 \frac{\delta^2}{2} + \dots$$

$$\mathbf{p} = \mathbf{p}_0 + \mathbf{p}_1 \delta + \mathbf{p}_2 \frac{\delta^2}{2} + \dots \quad (2.2.10)$$

then using Hamilton's equations yields

$$\mathbf{q}_1 = K_{1,p}$$

$$\mathbf{p}_1 = -K_{1,q}$$

$$\mathbf{q}_2 = 2(K_{1,pq}\mathbf{q}_1 + K_{1,pt} + K_{2,p})$$

$$\mathbf{p}_2 = -2(K_{1,qq}\mathbf{q}_1 - K_{1,qt} - K_{2,q})$$

where the indices after a comma denote differentiation. By integrating (2.2.13) order by order in  $\delta$ , we have

$$\begin{aligned}q_1 &= H_{,p} \\ p_1 &= -H_{,q}\end{aligned}$$

$$\begin{aligned}q_2 &= H_{,pq}q_1 + H_{,pp}p_1 + H_{,pt} \\ p_2 &= -H_{,qq}q_1 - H_{,qp}p_1 - H_{,qt}\end{aligned}$$

comparing this with the canonical transformation yields

$$K_1 = H$$

$$K_2 = -(H_{,p} \cdot H_{,q} + H_{,t})/2 \quad (2.2.11)$$

Sometimes it is convenient to use other generating functions, for example, the generating function of the second type  $K'(q_0, p)$  [Feng 1986]. It is possible to find  $K'$  by going through the same procedure as above; however, it is easier to find it directly by transforming between different types of generating functions. For example, the transformation between  $K'$  and  $K$  is

$$K'(p, q_0) = K(p, q_0) + p \cdot q + p_0 \cdot q_0 \quad (2.2.12)$$

or

$$\begin{aligned}p q_0 + K'_1(p, q_0) \delta + K'_2(p, q_0) \delta^2 \\ = -p_0 q + p q + p_0 q_0 + K_1(p_0, q) \delta + K_2(p_0, q) \delta^2\end{aligned}$$

move the term  $p_0 q_0$  to the left hand side

$$p_1 q_0 \delta + p_2 q_0 \frac{\delta^2}{2} + K'_1(p, q_0) \delta + K'_2(p, q_0) \delta^2$$

$$= p_1 q \delta + p_2 q \frac{\delta^2}{2} + K_1(p_0, q) \delta + K_2(p_0, q) \delta^2$$

From the first order equation, we have

$$p_1 q_0 + K'_1(p_0, q_0) = p_1 q_0 + K_1(p_0, q_0)$$

hence

$$K'_1 = K_1$$

the second order equation gives

$$p_2/2 + p_1 K'_{1,p} + K'_2 = p_1 q_1 + p_2/2 + q_1 K_{1,q} + K_2$$

hence

$$\begin{aligned} K'_2 &= p_1 q_1 + q_1 K_{1,q} - p_1 K_{1,p} + K_2 \\ &= K_{1,p} K_{1,q} + K_2 = (H_{,p} H_{,q} - H_{,t})/2 \end{aligned} \tag{2.2.13}$$

Similarly, we obtain expansions for other generating functions.

## 2.3. Invariant Structures and KAM Theorem

### 2.3.1. Integrable Systems

The symplectic structure imposes strong conditions on integrable systems. In fact it turns out that, for an  $n$  d.o.f. Hamiltonian system, we only need to know  $n$  first integrals to integrate the equations instead of the  $2n$  first integrals required for general ordinary differential equations. If the motion lies on a connected and compact manifold, then the latter is topologically an  $n$  dimensional torus in a  $2n$  dimensional symplectic manifold. Furthermore, this torus is geometrically a Lagrangian manifold (The symplectic form  $\omega$  vanishes on the torus) [Arnol'd 1978].

The above result is formalized in Liouville's theorem. We first introduce the Poisson bracket. On a symplectic manifold, the Poisson bracket  $\{F_1, F_2\}$  of two functions  $F_1, F_2$  is  $\omega(JdF_1, JdF_2)$ . Two functions are in involution with each other if their Poisson bracket is zero. For example, if a canonical system can be integrated by the method of Hamilton-Jacobi, then it has  $n$  first integrals in involution. Since the new canonical momenta  $P_i$  are constant and clearly in involution.

Theorem: Given  $n$  functions in involution on a  $2n$ -dimensional symplectic manifold

$$F_1, \dots, F_n \quad \{F_i, F_j\} = 0 \quad i, j = 1, 2, \dots, n$$

consider a level set of the functions  $F_i$

$$\mathcal{M}_f = \{x: F_i(x) = f_i, i=1, \dots, n\}$$

Assume all  $F_i$  are independent on  $\mathcal{M}_f$  (i.e.,  $dF_i$  are linearly independent at each point on  $\mathcal{M}_f$ ). Then

1.  $\mathcal{M}_f$  is a smooth manifold, invariant under the phase flow with Hamiltonian function  $H = F_i$  ( $i = 1, \dots, n$ ).
2. If  $\mathcal{M}_f$  is compact and connect, then it is diffeomorphic to the  $n$ -dimensional torus .
3. The phase flow with Hamiltonian function  $H$  determines a conditionally periodic motion on  $\mathcal{M}_f$ , i.e., in angular coordinates  $\phi = (\phi_1, \dots, \phi_n)$ , we have
 
$$d\phi/dt = v(f) \tag{2.3.1}$$
4. The canonical equations with Hamiltonian function  $H$  can be integrated by quadrature.

Since  $dF_i$  are linearly independent at each point of  $\mathcal{M}_f$ , by the implicit function theorem,  $\mathcal{M}_f$  is an  $n$ -dimensional submanifold of the  $2n$ -dimensional phase space. The functions  $F_1, \dots, F_n$  give  $n$  tangent vector fields  $JdF_i$  on  $\mathcal{M}_f$ . These vector fields commute with each other: the commutator of two vector fields is their Poisson bracket, and hence is zero. They are independent since  $dF_i$  are independent and  $J$  is nonsingular.  $JdF_i$  are tangent (parallel) to  $\mathcal{M}_f$  since the derivative of the functions  $F_j$  in the direction of the field  $JdF_j$  is the Poisson bracket  $(F_j, F_j)$  hence vanishes. The only  $n$ -dimensional connected and compact manifold parallelizable by  $n$  independent vector fields is the  $n$ -torus. Hence the motion occurs on an  $n$ -torus which is called an invariant torus.

The invariant torus  $\mathcal{M}_f$  is a Lagrangian manifold. Since the vector fields  $\{ JdF_i \}$  form a basis of the tangent space at each point, and  $\omega(JdF_i, JdF_j)$  is the Poisson brackets hence vanishes.

The action angle coordinates can be constructed by integrating the one form  $\mathbf{p}d\mathbf{q}$  over primitive cycles. Let  $C_1, C_2, \dots, C_n$ , be a basis for the one dimensional cycles on the  $n$  torus  $\mathcal{M}_f$ , then

$$I_i(\mathbf{f}) = \frac{1}{2\pi} \int_{C_i} \mathbf{p}d\mathbf{q} \quad (2.3.2)$$

are the action variables.  $\phi$  are angular coordinates on the torus  $\mathcal{M}_{\mathbf{f}}$ . The transformation from  $(\mathbf{p}, \mathbf{q})$  to  $(\mathbf{I}, \phi)$  is canonical. The action angle variables are not unique, since any transformation  $\phi' = \phi + \mathbf{c}(\mathbf{I})$  will give a new set of angle variables.

When the  $F_j$  cease to be independent for some values of  $\mathbf{f}$ , then  $\mathcal{M}_{\mathbf{f}}$  ceases to be a manifold. Such critical values of  $\mathbf{f}$  correspond to separatrices dividing the phase space into different integrable regions. Some of these regions can have unbounded motion similar to the unbounded motions in the pendulum system.

### 2.3.2. KAM Theorem and Its Applications

Integrable systems are extremely rare, however they can be considered as first approximations to real systems. For example, the motion of the earth can be regarded as an integrable Keplerian motion around the sun with only small perturbations due to other planets.

A near integrable system is given by the Hamiltonian in action angle variables

$$H = H_0(\mathbf{I}) + \varepsilon H_1(\mathbf{I}, \theta) \quad (2.3.3)$$

Poincaré called the study of motion in this system the fundamental problem of dynamics.

The fate of the motion described by (2.3.3) depends on the frequencies of the unperturbed system and is answered by the celebrated Kolmogorov-Arnold-Moser theorem (KAM theorem). It was first suggested by Kolmogorov from a non-rigorous perturbation theory (method of averaging), later proved by Arnol'd for analytic Hamiltonians and Moser for two dimensional, sufficiently smooth, twist maps from an  $2n$  dimensional annulus to itself of the form

$$\begin{aligned} \mathbf{I}' &= \mathbf{I} + \varepsilon \mathbf{f}(\mathbf{I}, \phi, \varepsilon) \\ \phi' &= \phi + \omega(\mathbf{I}) + \varepsilon \mathbf{g}(\mathbf{I}, \phi, \varepsilon) \quad \mathbf{I} \in \mathcal{B}, \phi \in \mathcal{T}^n \end{aligned} \quad (2.3.4)$$

where  $\mathcal{B}$  is a solid ball in  $\mathcal{R}^n$ . Obviously, under the nondegenerate condition to be defined below, the Poincaré map derived from (2.3.3) is a twist map (the technique of surface of section is always possible since we have angle coordinates). Conversely,

Douady [1982] has shown that any twist map of the form (2.3.4) can be derived from Poincaré surface of section of some Hamiltonian of the form (2.3.3). Therefore, the theorem for twist maps can be deduced from Hamiltonian flows and vice versa.

For the integrable system given by  $H_0(I)$ , the motion is a rotation on an invariant torus with frequency vector  $\omega(I) = \partial H_0(I)/\partial I$ . A torus is said to be nonresonant (incommensurate) if the frequencies are rationally independent: if  $k \in \mathbb{Z}^n$ , and  $(k, \omega) = 0$ , then  $k = 0$ . The unperturbed motion system is said to be nondegenerate if the frequencies are functionally independent

$$\det \left( \frac{\partial \omega}{\partial I} \right) = \det \left( \frac{\partial^2 H_0}{\partial I^2} \right) \neq 0 \quad (2.3.5)$$

The unperturbed system is called isoenergetically nondegenerate if one of the frequencies does not vanish and the ratios of the remaining  $n-1$  frequencies to it are functionally independent on the energy level  $H_0 = \text{const}$

$$\det \begin{pmatrix} \frac{\partial^2 H_0}{\partial I^2} & \frac{\partial H_0}{\partial I} \\ \frac{\partial H_0}{\partial I} & 0 \end{pmatrix} \neq 0 \quad (2.3.6)$$

In a nondegenerate or isoenergetic system the resonant tori form a dense set of full measure. The sets of resonant tori with any number of independent frequencies greater than one are dense and have measure zero.



For the map case, the unperturbed mapping is said to be nondegenerate if

$$\det \left( \frac{\partial \omega}{\partial \mathbf{I}} \right) \neq 0 \quad (2.3.7)$$

Theorem (KAM for flows, [Arnol'd 1988]): Consider Hamiltonian (2.3.3) such that  $H_0(\mathbf{I})$  is nondegenerate or isoenergetically nondegenerate, then for a sufficiently small  $\varepsilon$  most nonresonant invariant tori are only slightly deformed, so that in the phase space of the perturbed system, there are invariant tori densely filled with conditionally-periodic phase curves, with the number of independent frequencies equal to the number of degrees of freedom. These invariant tori fill most of phase space in the sense that the measure of the complement of their union is small, in the case of isoenergetic nondegeneracy the invariant tori fill most of each level manifold of the energy.

Theorem (KAM for maps, [Arnol'd 1988]): Consider the twist map (2.3.4). Suppose the unperturbed map is analytic and nondegenerate. Then for any sufficiently small perturbation of class  $C^r$  with  $r > 2n+1$  in the annulus  $\mathcal{B} \cup \mathcal{T}^n$  there exist invariant tori close to the tori  $I = \text{const}$ , and the measure of the complement of their union is small when the perturbation is small. The iterates of each point on an invariant torus densely fill the torus.

The invariant tori in KAM theorem are called KAM tori or the Kolmogorov set. The measure of the complement of the Kolmogorov set does not exceed a quantity of the order  $\sqrt{\varepsilon}$ , which is the typical size of the resonances which arise in perturbation theory. The deformation of the preserved torus depends on the

arithmetic properties of the frequencies and typically is of the order  $\sqrt{\epsilon}$ .

We can say more about the perturbed motions of (2.3.3) in the case of two d.o.f. systems. In fact the fate of almost all invariant tori can be decoded. First, for an isoenergetically nondegenerate system with two d.o.f., for all initial conditions the action variables remain forever near their initial values; since the energy level sets are three dimensional, and the KAM tori are two dimensional and fill a large part of an energy level set. Each such two torus divides the three dimensional energy level manifold. A phase curve starting in a gap between two invariant tori of the perturbed system remains forever trapped between these two tori. The oscillations of the action variables do not exceed a quantity of order  $\sqrt{\epsilon}$  from perturbation theory, because the measure of the difference between a torus and its unperturbed counterpart are bounded by quantities of this order.

The perturbation treatment fails for resonant tori. To give an approximate description of the fate of resonant tori, we average nonresonant terms and retain only one resonant term in the Hamiltonian [Chirikov, 1979]. This Hamiltonian, generically, has the form of a pendulum, and the size of the separatrix is of order  $\sqrt{\epsilon}$ . However, the neglected terms through averaging have the effect of splitting the separatrices: the periodic solutions are preserved by the Poincaré-Birkhoff theorem [see, for instance, Arnol'd and Avez 1968], the surfaces asymptotic (homoclinic or heteroclinic) for  $t \rightarrow +\infty$ , and  $t \rightarrow -\infty$  to these periodic solutions are no longer necessarily identical, in fact they typically cross each other transversely, which is the main reason for chaotic motions in two d.o.f. Hamiltonian systems. Far from the separatrices one finds all the tori obtained through averaging, except for an  $O(\exp(-c/\sqrt{\epsilon}))$  region, where  $c$  is a

constant. Near separatrices a special analysis shows that there exist tori exponentially close to the separatrices, so the separatrices are confined in a zone of exponentially small width and the angle between the split separatrices is exponentially small. It is possible to estimate the splitting between separatrices, which gives a way to prove the nonintegrability of the system. This involves perturbation near the separatrix and is called the Melnikov integral method. The validity of the perturbation treatment near separatrices is not obvious and is proved through a deep mathematical theorem.

An important application of the KAM theorem is the nonlinear stability of motion near periodic solutions. The tangent motion near fixed points or periodic solutions is mostly understood now, the latter corresponds to fixed points of the return map on the surface of section. There is a whole well developed machinery [MacKay 1986a, Howard and MacKay 87] to analyze the linear stability of a fixed point or periodic motion. For a finite time interval, the nonlinear motion near fixed points can be studied by applying Birkhoff's normal form. This, in fact, gives proofs for stability of the solar system for a finite but long time. However, the analysis of nonlinear stability of a fixed point or a periodic orbit requires more mathematical sophistication and is resolved only through application of the KAM theorem.

Assume the equilibrium position is stable in the linear approximation so that the characteristic frequencies around it are defined. Arnol'd proved that if there is no resonance relation of order less than or equal to four among characteristic frequencies, then almost all motions around the equilibrium point are quasiperiodic rotations on invariant tori [Arnol'd 1978]. This holds for autonomous flows, time periodic flows, as

well as symplectic maps. Therefore the equilibrium point is stable.

Notice, when  $n > 2$ , the stability is different from the usual Lyapunov stability. The periodic orbit is stable only in the measure theoretical sense. There exists the so called "Arnol'd Diffusion": points arbitrary near the elliptic periodic orbits can diffuse out, since the  $n$ -dimensional invariant tori do not divide the  $(2n-1)$ -dimensional energy level manifold. In this case, the "gaps" corresponding to different resonances are connected to one another, so that the tori do not prevent a phase curve originating near a resonance from wandering far away. This is called the "topological instability", since the KAM theorem basically proves the metric stability. So the Arnol'd diffusion is analogous to a drift in a fibered (by the invariant tori) sponge through resonance holes of the order  $\sqrt{\epsilon}$ .

### 2.3.3. Variational Principle for Invariant Tori and Cantori

Integrable and near integrable systems have been our key focus up till now. Far from integrability, the invariant tori and elliptic islands are destroyed; the regions with no invariant tori or elliptic islands are called the zones of instability. There are other more exotic invariant structures in the zones of instability. In the case of two dimensional area-preserving maps, some of the structures are known; they are cantori and horseshoes. We discuss cantori here.

An invariant torus on which the motion is incommensurate is the extremal of a variational principle. In the case of a symplectic twist map given by action principle (2.2.3), an incommensurate motion on a torus  $\Sigma$  is parametrized by a family of angle variables  $\theta: \mathbf{x} = \mathbf{x}_\Sigma(\theta), \mathbf{p} = \mathbf{p}_\Sigma(\theta)$ , such that  $\{\mathbf{x}(\theta_0+vt), \mathbf{p}(\theta_0+vt)\}$  is an orbit. The motion is a stationary state of the following action

$$\int_0^1 F(\mathbf{x}(\theta), \mathbf{x}(\theta+v)) d\theta \quad (2.3.8)$$

where  $\int d\theta$  means integration over  $(\theta_1, \theta_2, \dots, \theta_n)$ .

Now if  $\{\mathbf{x}(\theta_0+vt), \mathbf{p}(\theta_0+vt)\}$  is an orbit, the variation of the action is

$$\int_0^1 -\mathbf{p}(\theta)\delta\mathbf{x}(\theta) + \mathbf{p}(\theta+\nu)\delta\mathbf{x}(\theta+\nu)d\theta = 0$$

since  $\mathbf{x}(\theta)$ ,  $\mathbf{p}(\theta)$  are periodic.

On the other hand, the variation of (2.3.8) gives

$$F_1(\mathbf{x}(\theta), \mathbf{x}(\theta+\nu)) + F_2(\mathbf{x}(\theta-\nu), \mathbf{x}(\theta)) = 0$$

where the subscript of  $F$  denotes taking derivative with respect to the first or second argument in the generating function. These are equations of motion (2.2.5) for a discrete Hamiltonian system. Therefore  $\{\mathbf{x}_\Sigma(\theta), \mathbf{p}_\Sigma(\theta)\}$  represents an invariant torus if and only if the variation of the action (2.3.8) is zero.

Aubry [Aubry and Le Daeron, 1983] and Mather [1982] showed that for twist maps on the cylinder, there are always minima of the variational principle (2.3.8) in the space of monotone functions for every irrational frequency  $\nu$ . If the solution  $\mathbf{x}(\theta)$  is continuous, then it represents an invariant curve homeomorphic to a circle, and its restriction to this curve is topologically conjugate to the rotation through an irrational angle  $\nu$ . If  $\mathbf{x}(\theta)$  is discontinuous, then, the discontinuity points are dense on the circle, so the set given by the closure of  $\{\mathbf{x}(\theta+\nu t)\}$  is totally disconnected. It is also closed and perfect, therefore it is a Cantor set. These invariant Cantor sets were named cantori by Percival. Motion on the cantori is characterized by an irrational frequency  $\nu$ . Cantori can be thought of as remains of destroyed invariant tori. They occur generically in area-preserving maps.

Although cantori do not form absolute barriers to transport as invariant tori do, they can form partial barriers and restrict motion to one side or the other for a long time period, though trajectories finally leak through gaps in the Cantor set. Further details of partial barriers formed from cantori are considered in the next chapter.

We do not know if there are Cantor set solutions generally in higher dimensional Hamiltonian systems. However, there are examples of nearly hyperbolic four dimensional symplectic maps where there exist stationary Cantor set like solutions. It is not known if they are minimizing configurations [Chen, MacKay, and Meiss 1989].

The generic structure of two dimensional area-preserving maps is clear. When the system is nearly integrable, there are invariant tori and elliptic island chains in phase space. The island chains are hierarchical, so there are islands around islands. As the nonlinearity increases, some of the invariant tori are destroyed, and they are replaced by invariant cantori. The elliptic island chains are also destroyed, and they are usually replaced by other Cantor set like structures called horseshoes.

## 2.4. Applications

In this section, we give some few degree of freedom Hamiltonian systems arising from practical applications. The dynamics we discuss can all be reduced to two dimensional area-preserving maps. Other applications can be found in a reprint collection compiled by MacKay and Meiss [MacKay and Meiss, 1988].

### 2.4.1 Magnetic Field Line Flow as a Vortex Flow

For purpose of plasma confinement, if a static equilibrium with surfaces of constant pressure exists, the associated magnetic field must have nested toroidal flux surfaces, i.e. the magnetic field is exactly integrable. This is because if a magnetically confined plasma is in mechanical equilibrium, and if the plasma is sufficiently isolated so that it is characterized by a nearly scalar pressure  $P$ , then the force balance gives

$$\nabla P = \mathbf{J} \times \mathbf{B} \quad (2.4.1)$$

where  $\mathbf{B}$  is the magnetic field which satisfies

$$\nabla \cdot \mathbf{B} = 0 \quad (2.4.2)$$

and  $\mathbf{J}$  is the current density

$$\mathbf{J} = \nabla \times \mathbf{B} \quad (2.4.3)$$

The above set of equations are called the magnetostatic equations.



As a consequence of (2.4.1)

$$\mathbf{B} \cdot \nabla P = 0 \quad (2.4.4)$$

This implies the field lines lie in the constant pressure surfaces. For confinement, such a surface must be closed and bounded (compact), and the magnetic field can have no null points on it (any such point breaks the guiding center approximation and leads to eventual leakage of the plasma out of the surface). The only such physical realizable surface is topologically a torus. These surfaces are called flux surfaces [Hazeltine and Meiss 1985].

To study the feasibility of forming these nested toroidal surfaces, we need to investigate the magnetic field line flow, i.e. find the integral curve whose tangent vector is parallel to the given static magnetic field [Cary and Littlejohn, 1983]. This corresponds to the lowest order motion of the confined plasma (neglecting all gyromotion and guiding center drifts). The equations for the field line flow are

$$\mathbf{B} \times d\mathbf{x}/d\lambda = 0 \quad (2.4.5)$$

where  $\lambda$  is a parameter of the integral curve. Eq(2.4.5) can be put in a variational form

$$\delta \int \mathbf{A} = 0$$

where the fundamental one form is the vector potential

$$\mathbf{A} = A_1 dx^1 + A_2 dx^2 + A_3 dx^3 \quad (2.4.6)$$

As we discussed in section 2.1.3, this defines a Hamiltonian flow if the two form

$$d\mathbf{A} = B_3 dx^1 \wedge dx^2 + B_1 dx^2 \wedge dx^3 + B_2 dx^3 \wedge dx^1 \quad (2.4.7)$$

is nondegenerate. Here the  $B_i$  are the components of the magnetic field. This implies the magnetic field is non-null (nonsingular). This condition is satisfied for the toroidal configuration. However, singular magnetic fields are important, for instance, in space plasma problems.

Therefore the generic magnetic field line flow is equivalent to a one and half degree of freedom Hamiltonian flow, hence flux surfaces exist for all values of the rotational transform (to be defined below) if the system possesses a perfect symmetry (Noether's theorem). For instance, azimuthally symmetric geometry gives a system with nested toroidal flux surfaces. The problem with azimuthal symmetric geometry is that the rotational transform can only be produced by an internal toroidal current. This is the mechanism of Tokmak confinement. Toroidal devices with the rotational transform produced by external coils are called stellarators; as far as is known, they can only be achieved without perfect symmetry.

It is instructive to consider the field configuration with azimuthal symmetry

$$\gamma = A_z(z,R)dz + A_R(z,R)dR + A_\phi(z,R)d\phi \quad (2.4.8)$$

where  $A_z$ ,  $A_R$ ,  $A_\phi$  are the components of the vector potential in cylindrical coordinates  $(z, R, \phi)$  (Figure 2.5). From Noether's theorem,  $A_\phi$  is conserved and we write

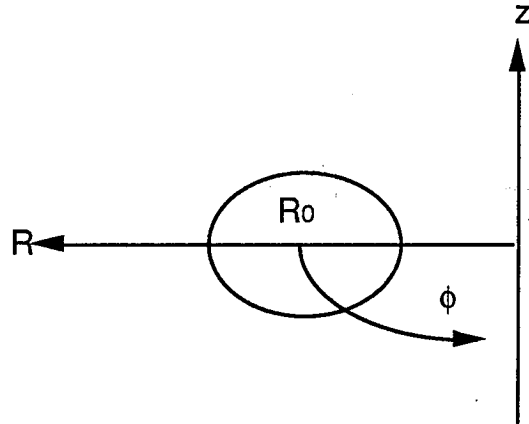


Figure 2.6. Cylindrical coordinates of the field line flow.

$$\psi(z, R) = -A_\phi(z, R) \quad (2.4.9)$$

$\psi$  plays the role of the Hamiltonian function. The surfaces  $\psi = \text{const.}$  are the flux surfaces. These surfaces are azimuthally invariant, hence if they cut a closed curve in the  $z$ - $R$  plane, they generate a torus by rotation about the  $z$ -axis. The innermost torus, which is an extremal of  $\psi$ , is degenerate and is a closed curve. The poloidal flux between two flux surfaces  $C_a$  and  $C_b$  is

$$\int_{C_a} \gamma - \int_{C_b} \gamma = 2\pi (\psi_a - \psi_b) \quad (2.4.10)$$

Hence is basically the poloidal flux.

The action variable is

$$u = 1/2\pi \int A_z(z,R)dz + A_R(z,R)dR \quad (2.4.11)$$

The contour is given by  $\psi(z,R) = \text{const}$ .  $u$  is a function only of the energy  $\psi$ . By Stokes' theorem,  $u$  is the toroidal flux. The quantity  $du(\psi)/d\psi = q(\psi)$  is called the safety factor.

Hence in action angle variables, the fundamental one form has the canonical form

$$\gamma = u d\theta - \psi d\phi \quad (2.4.12)$$

The angle variable  $\theta$  is not unique and can be chosen in different ways depending on the problem. The field line equations in these coordinates are

$$\begin{aligned} d\theta/d\phi &= 1/q \\ du/d\phi &= 0 \end{aligned} \quad (2.4.13)$$

So the safety factor is the inverse of the rotation transform  $\tau$  of the field line flow.

Perturbations of the perfect symmetric field result in the breakup of the pressure surfaces with rational rotational transform and magnetic islands (resonances) emerge. In this case, it will be more fruitful to investigate the return map of the field line flow.

### 2.4.2. Passive Scalars and Drift Wave Instability

In hydrodynamic systems, it is useful to use small particles to visualize the fluid flow. Ideally these particles are massless; they are driven by the velocity field of the flow  $\mathbf{u}(\mathbf{x},t)$ , hence are called passive scalars (PS). Some local physical quantities such as temperature or the density of a second fluid can be regarded under some conditions as PS. Therefore it is important to understand the dynamics of PS in the theory of mixing of fluids.

For two dimensional flows, the motion of a passive scalar is described by Hamilton's equations. Let  $\Psi$  be the stream function of the flow, then

$$\begin{aligned}\frac{dx}{dt} &= -\frac{\partial\Psi}{\partial y} \\ \frac{dy}{dt} &= \frac{\partial\Psi}{\partial x}\end{aligned}\tag{2.4.14}$$

Hence  $x$  serves as the canonical momentum,  $y$  the canonical coordinate, and  $\Psi$  the Hamiltonian function. Therefore steady two dimensional flows are integrable and mixing of fluids is highly inefficient. However, if the stream function is time dependent, PS typically undergo the same kind of chaotic motion arising in two d.o.f. Hamiltonian systems and the transport is strongly enhanced. This enhancement of transport is known as chaotic advection.

A similar problem arising in Plasma physics is the study of the drift wave instability. This instability occurs in a plasma

with density gradients and produces anomalous transport of particles across the confining magnetic field. Drift waves are low frequency electrostatic waves in which the cross field particle motion is given by the  $\mathbf{E} \times \mathbf{B}$  drift of the guiding center. The motion of the test particle is Hamiltonian

$$\begin{aligned} \frac{dx}{dt} &= -\frac{c}{B} \frac{\partial \phi}{\partial y} \\ \frac{dy}{dt} &= \frac{c}{B} \frac{\partial \phi}{\partial x} \end{aligned} \quad (2.4.15)$$

where  $\phi(x,y,t)$  is the electrostatic potential. This is a one and half d.o.f. system.

In the presence of  $N$  drift waves of a plasma bounded in  $x$  and periodic in  $y$ , the electrostatic potential is

$$\phi(x,y,t) = -Ex + \sum_{k=1}^N A_k \sin(k_x x) \cos(k_y y - \omega t + \beta_k) \quad (2.4.16)$$

where  $\mathbf{k} = (\pi n/L_x, 2\pi m/L_y)$  are the eigenmodes determined by the boundary conditions. For a single drift wave, the system is integrable, and it can only produce a localised convection of the plasma with no net transport. However, the presence of a small secondary drift wave produces stochastic motion along the boundaries of the convective motion and give rise to a net plasma transport. The onset of global transport can be treated by the Melnikov function method [Horton 1986].

### 2.4.3. Cyclotron

The map (2.2.1) can also describe the motion of particles in a cyclotron (Fig. 2.6). Suppose there is a time dependent voltage drop  $V(t)$  across a narrow azimuthal gap in a magnetic field  $\mathbf{B} = B_0 \mathbf{e}_z$ . The time for an electron to go one cycle is

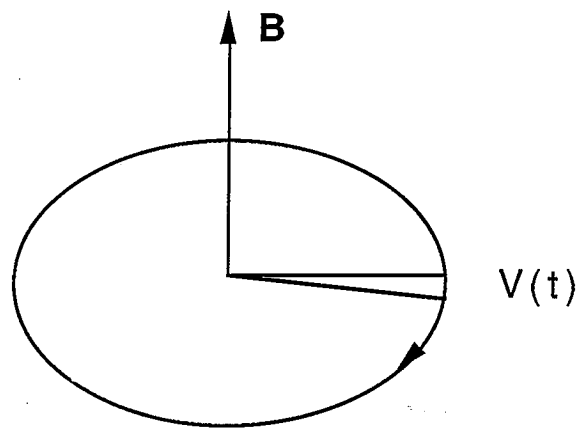


Figure 2.7. The cyclotron.

$$T = 2\pi/\Omega = 2\pi m\gamma c/eB = 2\pi E/eBc$$

where  $E$  is the particle energy  $m\gamma c^2$ . The change in energy upon traversing the gap is

$$\Delta E = -eV(t)$$

Let  $(E, t)$  be the energy and time just before the gap; then after one cycle, their new values are

$$\begin{aligned} E' &= E - eV(t) \\ t' &= t + 2\pi E'/eBc \end{aligned} \tag{2.4.17}$$

Let  $y = E/e$ ,  $x = tBc/2\pi$ ,  $f(x) = V(2\pi x/Bc)$ , we recover the map  
(2.2.1)

$$y' = y + f(x)$$

$$x' = x + y'$$



## **Chapter 3**

### **Resonances and the Devil's Staircase in Area-Preserving Twist Maps**

In this chapter, we introduce a Markov transport model based on the flux and resonance theory of transport for twist maps of the cylinder. The main result of this chapter is to show that in the supercritical regime when there are no rotational invariant circles, resonances give a complete partition of the phase space. Section 3.1 to 3.3 is a review of twist maps of the cylinder: minimizing and minimax orbits, the construction of partial barriers and resonances, and the definition of flux and area devil's staircase function. In section 3.4, we show that there is a simple relation between the average Lagrangian of a minimizing orbit and the area under this orbit, i.e., the area function is the derivative of the average Lagrangian with respect to frequency. This implies area function is a complete devil's staircase in the supercritical regime. In section 3.5, we introduce the Markov model.

### 3.1. Twist Maps of the Cylinder

#### 3.1.1 Twist Maps

Consider the cylinder  $M = \mathcal{T}^1 \times \mathcal{R}$  parametrized by canonical coordinates  $(\theta, y)$ , where  $\theta \in [0, 1)$  is the angle coordinate,  $y \in \mathcal{R}$  can be considered as the action variable. The area form is

$$\omega = dy \wedge d\theta \quad (3.1.1)$$

An area-preserving twist map of the cylinder  $\mathcal{M} = \mathcal{T}^1 \times \mathcal{R}$  is an orientation preserving  $C^1$  diffeomorphism  $T: \mathcal{M} \rightarrow \mathcal{M}$

$$(\theta', y') = T(\theta, y) \quad (3.1.2)$$

such that  $T$  preserves the topological ends of the cylinder and the area form (3.1.1), and there exists a constant  $K$  such that

$$\partial\theta'/\partial y \geq K > 0 \quad (3.1.3)$$

If the map is derived from a flow, it necessarily preserves the topological ends since the flow provides a smooth connection of the mapping to the identity mapping. Twist is also a common condition. For example, near a typical elliptic point, there exists a coordinate system such that the return map is given by a Birkhoff normal form which satisfies the twist condition

$$\begin{aligned} r' &= r + h(r, \theta) \\ \theta' &= \theta + 2\pi\omega + p_2 r^2 + \dots + p_{2k} r^{2k} + g(r, \theta) \end{aligned} \quad (3.1.4)$$

Here  $h$  and  $g$  are  $\mathcal{O}(r^{2k})$ , and thus as small as one likes.

It is often useful to lift the map to the plane  $\mathbb{R}^2$ , which is the universal covering space of the cylinder. The lift of a map  $f$  is a diffeomorphism  $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , such that  $\pi F = f\pi$ , where  $\pi$  is the covering map  $\pi(x,y) = (x \bmod 1, y)$ . We shall denote the lift coordinates by  $(x,y)$ . Let  $R$  be the translation  $R(x,y) = (x+1,y)$ , then  $F$  commutes with  $R$ , hence  $x'(x+1,y) = x'(x,y) + 1$ . If  $F'$  is another lift then  $F' = FR^k$ , for some  $k \in \mathbb{Z}$ .

If  $F$  is the lift of a twist map, then  $\partial x'/\partial y \geq K > 0$ , therefore, the first iterate of a vertical line ( $x = \text{const}$ ) tilts to the right.

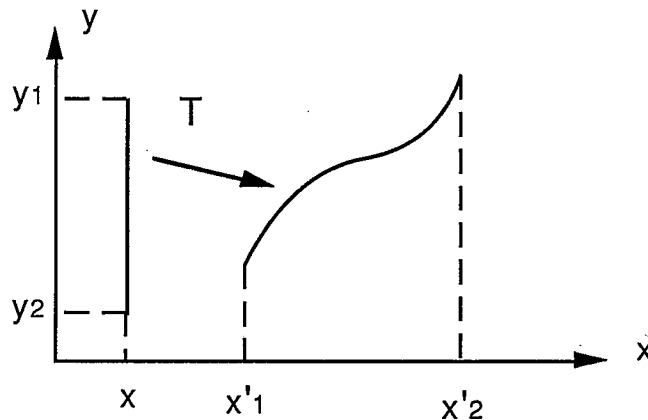


Figure 3.1. Twist Condition.

Hence the map  $(x,y) \rightarrow (x,x')$  is invertible,  $(x,x')$  can be chosen as good coordinates on  $\mathbb{R}^2$ .

We shall also use the concept of a reversible area-preserving map. An area-preserving map  $T$  is reversible if there is a reversor  $S$  with  $S^2 = \text{Id}$  which conjugates  $T$  to its inverse  $T^{-1} = STS$ . The fixed points of  $S$  form curves called symmetric lines [MacKay, 1982]. Near any fixed points of a reversor, there exist normal symmetry coordinates  $(x,y)$  such that  $S(x,y) = (x,-y)$ .

For reversible maps, there always appear to be symmetric periodic orbits of all rotation numbers on each of the symmetry lines. Furthermore, in some cases, there appears to be one symmetry line on which occur periodic orbits of nonnegative residue (see section 3.2.3) for all rotation numbers. Such reflection lines are called dominant. Maps with dominant lines are said to possess dominant symmetry.

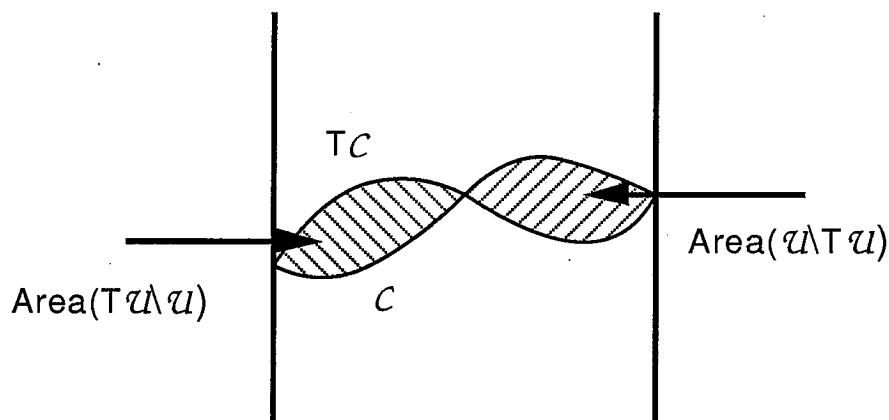


Figure 3.2. The net flux.

Let  $c$  be a loop encircle the cylinder once and  $u$  be the region under  $c$ , then  $u$  is an open set homeomorphic to the cylinder. The net flux of  $T$  is

$$\mathcal{F} = \text{Area}(T(u) \setminus u) - \text{Area}(u \setminus T(u)) \quad (3.1.4)$$

The net flux is independent of the choice of  $c$ . It can be thought of as the flux of area across a rotational circle. Therefore, a map which has a rotational invariant circle must have zero net flux.

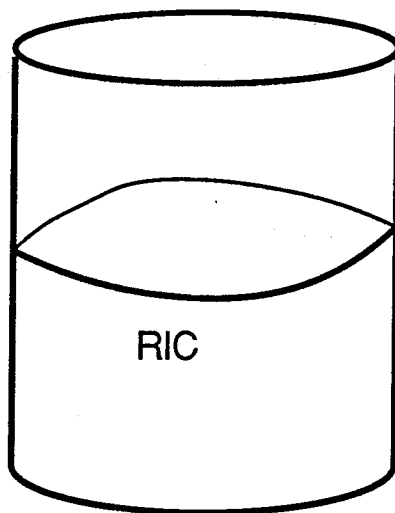


Figure 3.3. A rotational invariant circle (RIC).

We shall be interested in twist maps with zero net flux.

### 3.1.2. Action Principle

Area-preserving twist maps have an action representation, and orbits of the map can be obtained from a variational principle.

Theorem: Let  $T: \mathcal{R}^2 \rightarrow \mathcal{R}^2$  be a lift of  $C^r$  area-preserving twist map with  $(x',y') = T(x,y)$ . Then there exists a  $C^{r+1}$  function  $F: \mathcal{R}^2 \rightarrow \mathcal{R}$  such that

$$y'dx' - ydx = dF(x,x') \quad (3.1.5)$$

The function  $F$  is called the generating function of the map  $T$ . It is unique up to addition of constant and satisfies the twist

condition  $\frac{\partial^2 F}{\partial x \partial x'} \leq C < 0$ .

If  $F$  is the generating function of  $T$ , so is  $F'(x,x') = F(x-1,x'-1)$ , since  $T$  commutes with  $R(x,y) = (x-1,y)$ . Thus  $F(x,x') - F(x-1,x'-1)$  is a constant. This is in fact the net flux of the map  $T$ , as we will see from the fundamental area formula (3.3.1).

The generating function  $F_n$  of  $n$  iterates of the map  $T$  is given by the sum of the generating function  $F$ .

$$F_n(x_0, x_n) = \text{sta} \left[ \sum_{t=0}^n F(x_t, x_{t+1}) \right] \quad (3.1.6)$$

where  $\text{sta}$  means all  $x_t$  are taken as the stationary values of the total action with fixed ends  $x_0$  and  $x_n$ .

A state is a bi-infinite sequence  $\{ \{ x_t \} : x_t \in \mathcal{R}, t \in \mathcal{Z} \}$ . A segment  $\{ x_{MN} \}$  of the state is a finite subsequence  $\{ x_t \in \mathcal{R}, M \leq t \leq N \}$ . An orbit is a state whose arbitrary finite segments  $x_{MN}$  are stationary configurations of the action

$$W(\{x_t\}) = \sum_{t=M}^N F(x_t, x_{t+1}) \quad (3.1.7)$$

In this sense, we refer an orbit as a stationary state.

Equation (3.1.7) defines the classical Frenkel-Kontorova model (F-K model) as a one dimensional chain of atoms with nearest neighbor interaction  $F(x_i, x_{i+1})$ . A large class of generating functions have the form

$$F(x, x') = K(x-x') - V(x) \quad (3.1.8)$$

where  $K$  is the spring function, which is convex to satisfy the twist condition, and  $V$  is a periodic potential function so the map is of the cylinder and has zero net flux.

A periodic orbit  $\{ x_t \}$  is of type  $(m, n)$  if for all  $t$ ,  $x_{t+n} = x_t + m$ . Let  $x_{m,n} = \{ \{ x_t \} : x_{t+n} = x_t + m, t \in \mathcal{Z} \}$ . Define the action  $W_{m,n}$  by

$$W_{m,n}(\{x_t\}) = \sum_{t=0}^{n-1} F(x_t, x_{t+1}) \quad x_n = x_0 + m \quad (3.1.9)$$

It is obvious that orbits of type  $(m, n)$  corresponds to stationary states of  $W_{m,n}$  with respect to variations preserving  $x_n = x_0 + m$ .



### 3.1.3 Example

An example of the twist map with zero flux is the much-studied standard map where

$$F(x,x') = \frac{(x' - x)^2}{2} - \frac{k}{(2\pi)^2} \cos(2\pi x) \quad (3.1.10)$$

or

$$\begin{aligned} y' &= y - \frac{k}{2\pi} \sin(2\pi x) \\ x' &= x + y' \end{aligned} \quad (3.1.11)$$

The symmetry reversor is

$$S(x,y) = (-x+y,y) \quad (3.1.12)$$

The standard map has zero net flux since  $F(x+1,x'+1) = F(x,x')$ . The  $y$  coordinate of the standard map is also periodic. So the phase space can be wrapped to a torus.

## 3.2. Monotone Sets and Minimizing Orbits

### 3.2.1. Monotone Sets and Frequencies

Let  $T$  be the lift of an area-preserving twist map and let  $\pi$  be the projection  $\pi(x,y) = x$ . If for a given point  $x \in \mathcal{R}^2$  the limit

$$v(x) = \lim_{t \rightarrow \pm\infty} (\pi(T^t(x)) - \pi(x))/t \quad (3.2.1)$$

exists, then  $v$  is called the frequency or rotation number of  $x$  for  $T$ . The limit is obviously independent of the choice of point on an orbit and hence it is the frequency for the orbit. A type  $(m,n)$  orbit has frequency  $m/n$ .

A nonempty set  $\mathcal{M}$  which is both invariant under  $T$  and  $\mathcal{R}$  (the translation) is monotone if  $\pi|_{\mathcal{M}}$  is injective and

$$x, x' \in \mathcal{M}, \pi(x) < \pi(x') \Rightarrow \pi(T(x)) < \pi(T(x')) \quad (3.2.2)$$

We say an orbit of  $\{x_t\}$  is monotone if all of its translates are monotone

$$x_t + j < x_{t'} + j' \Rightarrow x_{t+1} + j < x_{t'+1} + j' \quad (3.2.3)$$

Monotone sets are called ordered in this sense.

All monotone orbits have frequencies [MacKay and Stark 1985].

### 3.2.2. Continued Fraction and Farey Tree

The structure and organization of monotone orbits depends on the number theoretical properties of their frequencies.

The continued fraction expansion [Hardy and Wright, 1979, Khinchin, 1964] of  $v$  is

$$v = [a_0, a_1, \dots, a_i, \dots] = a_0 + 1/(a_1 + 1/(\dots + 1/(a_i + \dots))) \quad (3.2.3)$$

where the  $a_i$ 's are positive integers. If  $v$  is rational, this expansion terminates and has two representations:  $[a_0, a_1, \dots, a_n]$  and  $[a_0, a_1, \dots, a_n - 1, 1]$  with  $a_n > 1$ ; Otherwise if  $v$  is irrational, the expansion is infinite.

If the continued fraction is periodic, it is a solution of a quadratic equation with integer coefficients, therefore is called a quadratic irrational. If  $a_i = 1$ , for all integer  $i$  larger than some integer  $j$ , then the number is called noble. The noblest number is the golden mean  $\gamma = (\sqrt{5}+1)/2 = [1, 1, 1, \dots]$ . If all  $a_i$  are bounded, the number is of constant type.

Truncations at some finite stage give the convergents of a continued fraction

$$m_i/n_i = [a_0, a_1, \dots, a_i]$$

These are the best rational approximants to  $v$  in the sense that  $|v - m/n| > |v - m_i/n_i|$  for all other  $m/n$  with  $n < n_{i+1}$ . Note that successive convergents to  $v$  approximate  $v$  from opposite sides. Every convergent is close to the number it approximates in the sense that

$$|nv - m| < C/n$$

where  $C = 1$ . However, if  $C$  is too small ( $C < 1/\sqrt{5}$ ), there exists  $v$  such that only finitely many convergents satisfy the inequality.

A Farey tree is a method for organizing all rationals between a given pair in a binary tree. Begin with a pair of neighboring rationals  $m/n$  and  $m'/n'$ , which satisfy the relation  $mn' - nm' = \pm 1$  (we assume  $+1$  here so that  $m/n < m'/n'$ ), level one of the tree is generated from these two by adding their numerators and denominators:  $m''/n'' = (m+m')/(n+n')$ . This rational is called the mediant of  $m/n$  and  $m'/n'$ . It is easy to see that  $m''$  and  $n''$  have no common divisor, and  $m''/n''$  is a neighbor to both its parents. To construct the second level, find the mediant of  $m''/n''$  and each of its parents. Continuing this construction leads to a binary tree, which gives all the rational numbers between  $m'/n'$  and  $m/n$ . For example, starting from  $0/1$  and  $1/0$ , we get the master tree which gives all nonnegative rational numbers.

The sum of the continued fraction elements of  $m/n$  gives the level which it occurs in the master tree

$$\text{Level}([a_0, a_1, \dots, a_i]) = \sum_{j=0}^i a_j$$

The parents of a rational can easily be obtained from its continued fraction expansion:

$$m/n = [a_0, a_1, \dots, a_i]$$

With the convention that the final element  $a_j > 1$ , the continued fraction expansion for their parents are  $[a_0, a_1, \dots, a_{j-1}, a_{j-1}]$  and  $[a_0, a_1, \dots, a_{j-1}]$

### 3.2.3. Minimizing Orbits

A segment  $x_{MN}$  is minimizing if  $W$  is globally minimizing with respect to fixing the end points  $x_M$  and  $x_N$ . A state  $\{x_t\}$  is minimizing if every finite segment is minimizing. A minimizing state is called a minimizing orbit.

Minimizing orbits have the following important properties:

a) Aubry's fundamental lemma: If  $\{u\}$ ,  $\{v\}$  are two minimizing states, then  $u_n - v_n$  has at most one change of sign or zero in  $\mathbb{Z}$ . Furthermore, if  $u_n - v_n$  does have a change of sign or zero, then  $u_n - v_n$  is uniformly bounded away from zero everywhere else, and in particular  $\{u\}$  and  $\{v\}$  cannot be asymptotic to each other.

b) Every minimizing orbit is monotone, and hence in particular has a rotation number.

Therefore minimizing orbits are characterized by their frequencies. There are recurrent minimizing orbits for each frequency.

c) All recurrent orbits with frequency  $m/n$  are periodic orbits of type  $(m,n)$ . There exists at least one periodic minimizing orbit of type  $(m,n)$ . The set of all periodic orbits of type  $(m,n)$  is denoted as  $\mathcal{M}_{m/n}$ . This is a monotone set.

We can define periodic orbits in  $\mathcal{M}_{m/n}$  to be adjacent if there is no other orbit in  $\mathcal{M}_{m/n}$  between them. Note if  $\{x_t\}$  is a minimum of  $W_{m,n}$ , so is  $x'_t = x_{t+r} + s$ , for  $r,s \in \mathbb{Z}$ . These correspond to the orbit on cylinder. If there is only one equivalent minimizing orbit in  $\mathcal{M}_{m/n}$ , the right and left adjacent

(or neighboring) orbits of  $\{x_t\}$  are  $\{x_t^r = x_{t+r} - s\}$  and  $\{x_t^l = x_{t+l} - j\}$ , where  $r, s, l, j$  are given uniquely by

$$mr - ns = 1, ml - nj = -1, r, l < n$$

d) For every irrational number  $\nu$ , there exists a monotone set  $\mathcal{M}_\nu$  of recurrent minimizing states of this rotation number.  $\mathcal{M}_\nu$  is either a rotational invariant circle or an invariant cantor set. Furthermore every orbit in  $\mathcal{M}_\nu$  is dense in  $\mathcal{M}_\nu$ .  $\mathcal{M}_\nu$  can be parametrized by two hull functions  $f^+(\theta)$  and  $f^-(\theta)$ , which are precisely the minimizing solution of Percival's action (2.3.8) in the space of monotone functions.  $f^+(\theta)$  and  $f^-(\theta)$  satisfy

- 1)  $f^\pm(\theta)$  are strictly increasing.
- 2)  $f^\pm(\theta+1) = f^\pm(\theta) + 1$ .
- 3)  $f^-(\theta)$  is left continuous
- 4)  $f^+(\theta)$  is right continuous.
- 5)  $f^-(\theta) \leq f^+(\theta)$

If  $f^-(\theta) = f^+(\theta) = f(\theta)$  is continuous, then  $\mathcal{M}_\nu$  is a rotational invariant circle; otherwise, if  $f^+(\theta)$  and  $f^-(\theta)$  are discontinuous, then  $\mathcal{M}_\nu$  is an invariant cantor set called a cantorus. The discontinuities come in orbits under  $\theta \rightarrow \theta + \nu$ . The discontinuities of  $f^+(\theta)$  and  $f^-(\theta)$  correspond to endpoints of gaps in the cantorus. Since the orbit is monotone, these endpoints do not cross upon iteration, therefore, we can define the left end and right end orbits  $\{x_t^l\}, \{x_t^r\}$  for the cantorus.

There are also nonrecurrent minimizing orbits for each frequency  $\nu$ . They are orbits heteroclinic to the gap endpoints. When  $\nu$  is rational, these orbits are crossing points of the stable and unstable minimizing orbits (minimizing orbits are usually

hyperbolic). When  $\nu$  is irrational, these orbits fall in the gaps of the cantor. Let  $\text{Min}_\nu$  be the set of all minimizing states of frequency  $\nu$ , we find the following classification of the minimizing orbits.

**Classification for  $\nu$  irrational** The set  $\mathcal{M}_\nu$  is unique and all recurrent minimizing states of rotation number  $\nu$  belong to it. If  $\mathcal{M}_\nu$  is a circle then every minimizing state of rotation number  $\nu$  is recurrent, and lies in  $\mathcal{M}_\nu$ . If  $\mathcal{M}_\nu$  is a cantor set then every state in  $\text{Min}_\nu$  is either in  $\mathcal{M}_\nu$  or lies in some gap of  $\mathcal{M}_\nu$  and is asymptotic to the orbits of the end points define that gap.  $\text{Min}_\nu$  is thus a monotone set.

**Classification for  $\nu = m/n$  rational** All recurrent orbits in  $\text{Min}_{m/n}$  are periodic orbits of type  $p, q$  and thus are in  $\mathcal{M}_{m/n}$ . Any other orbit in  $\text{Min}_{m/n}$  must either satisfy  $x_{n+t} > x_t + m$  or  $x_{n+t} < x_t + m$ . The former are called advancing orbits and the latter retreating orbits. Any advancing or retreating orbit is heteroclinic to a pair of adjacent periodic orbits, and is confined to the gap defined by these periodic orbits. An advancing orbit with frequency  $m/n$  is denoted by  $(m, n)^+$  orbit, and the corresponding retreating orbit is denoted by  $(m, n)^-$  orbit. Conversely for each neighboring pair of periodic orbits there are at least two minimizing heteroclinic orbits, one advancing and one retreating.



### 3.2.4. Minimax Orbits

Let us define the residue of a period  $n$  orbit as

$$R = (2 - \text{Tr}(DT^n))/4 = (2 - \lambda - \lambda^{-1})/4 \quad (3.2.3)$$

Where  $\lambda, \lambda^{-1}$  are the eigenvalues of the linear map of the periodic orbit. The residue is related to the linear stability of the orbit as follows

$$\begin{aligned} R < 0, & \text{ regular hyperbolic} \\ R = 0, & \text{ regular parabolic} \\ 0 < R < 1, & \text{ elliptic} \\ R = 1, & \text{ inversion parabolic} \\ R > 1, & \text{ inversion hyperbolic} \end{aligned} \quad (3.2.4)$$

**Theorem (Poincaré-Birkhoff):** Let  $T$  be an area-preserving twist map of the cylinder with zero net flux. Then for every rational number  $p/q$ ,  $T$  has at least two periodic orbits of type  $p,q$ : one of non-positive residue and one of non-negative residue.

Given a minimizing  $(m,n)$  periodic orbit  $\{ M_t \}$ , let  $\{ M_t^r \}$  be the same orbit shifted to the right of  $\{ M_t \}$ .  $\{ M_t^r \}$  has the same action as  $\{ M_t \}$  since the net flux is zero. Since the space is compact, there must be a saddle in between. This gives another periodic orbit, the minimax orbit  $\{ S_t \}$ . In a similar fashion, one can show the existence of minimax  $(m,n)^+$  and  $(m,n)^-$  homoclinic orbits.

If the rotation number of the minimizing set is irrational, and  $\mathcal{M}_\nu$  is an invariant cantor set, let  $\{l_t\}$ ,  $\{r_t\}$  be the orbits of the endpoints of some gap in  $\mathcal{M}_\nu$ , the sum

$$\Delta W(x_t) = \sum_{t=-\infty}^{\infty} [F(x_t, x_{t+1}) - F(l_t, l_t)]$$

converges for sequences lying in the orbit of the gap,

$$l_t \leq x_t \leq r_t$$

It can be shown that  $\Delta W(x_t)$  is differentiable and nonnegative and is zero for  $\{l_t\}$  and  $\{r_t\}$ . Mather [1982] proved that there is a minimax orbit  $\{S_t\}$  in between. It is homoclinic to the Cantor set. The action difference between the minimax orbit and the minimizing orbit is denoted as  $\Delta W_\nu$ . Mather proved that this is a continuous function of  $\nu$  at the irrationals, and it is zero if and only if there exists a rotational invariant circle with rotation number  $\nu$ .

The difference of the action between a pair of minimizing and minimax orbit has a physical interpretation as the flux across the partial barriers. This is discussed in the next section.

A result of MacKay and Meiss [1983] states that a minimizing orbit has nonpositive residue, and a minimax periodic orbit has nonnegative residue. Therefore, minimizing orbits are usually hyperbolic.

### 3.3. Partial Barriers, Turnstiles and Resonances

#### 3.3.1. Irregular components

An irregular component is by definition the complement of invariant circles, elliptic periodic orbits and invariant circles around them.

For a hyperbolic period  $n$  orbit, the linear mapping has two vectors corresponding to the stable and unstable directions. The stable manifold theorem implies the stable and unstable directions can be continued nonlinearly into stable and unstable manifolds:  $W^S = \{ z, T^n z \rightarrow z_0 \text{ as } j \rightarrow \infty \}$ , where  $z_0$  is some point on the orbit; similarly  $W^u = \{ z, T^{-n} z \rightarrow z_0 \text{ as } j \rightarrow \infty \}$ . Stable and unstable manifolds typically intersect with each other transversely, and such intersections are called heteroclinic points if the manifolds correspond to different periodic orbits, and homoclinic if they are the same periodic orbit. The advancing  $(m,n)^+$  and retreating  $(m,n)^-$  minimizing orbits we discussed in section 3.3.3. are homoclinic orbits to the  $(m,n)$  minimizing periodic orbit.

Whenever there is a transverse crossing of the stable and unstable manifolds, there is an irregular component. In fact, by following the lobes of the stable and unstable manifold, one can construct an invariant set called a "horseshoe", whose points are labeled by a doubly infinite sequence of 0s and 1s, such that iterating the map is equivalent to shift the symbol sequence one place to the left.

Numerical experiments show that an irregular component has a dense orbit, has a positive measure [Umberger and Farmer, 1985]. Furthermore, it appears to be ergodic and has the same positive Lyapunov exponent

$$\lambda = \lim_{n \rightarrow \infty} \ln(\text{Tr}(DT^n(z)))/n > 0$$

almost everywhere.

For twist maps of the cylinder, irregular components are usually bounded by two rotational invariant circles. These invariant circles act as absolute barriers between irregular components. We are interested in global transport where all rotational invariant circles are destroyed. In the following, we assume that all minimizing orbits are hyperbolic. More precisely, we assume that the lower bound of the average Lyapunov exponents for all minimizing orbits is greater than zero. This implies there is no rotational invariant circle [MacKay and Percival 1985].

### 3.3.2. Flux, Partial Barriers, and Resonances

A useful concept is the flux from one region to another. Consider a closed loop  $C$  encircling its interior  $I$ , and their iterates  $C'$  and  $I'$  (Fig 3.4), the flux across  $C$  is

$$F = \text{Area}(I - I') = \text{Area}(I' - I)$$

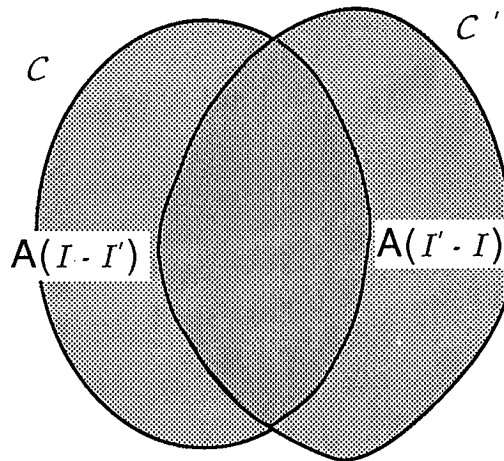


Figure 3.4. Flux.

If we choose  $C$  as a rotational circle, since the net flux is zero, the upward flux is equal to the downward flux.

A rotational circle acts as a "partial barrier" for the map. The simplest way to construct a partial barrier is to use the minimizing and minimax periodic  $(m,n)$  orbits. Choose a gap in the minimizing orbit, called the principal gap, if the map possesses a dominant symmetry, the gap is chosen to be the one

across the dominant line. Fill the gap with an arbitrary line  $\mathcal{L}_0$  passing through the minimax orbit. Take  $n-1$  preimages of this curve to fill the remaining gaps. This results a rotational circle, a partial barrier. Take one more preimage of  $\mathcal{L}_0$ , we get  $\mathcal{L}_{-n}$ . It must get back to the principle gap with the same end points as  $\mathcal{L}_0$  and must cross  $\mathcal{L}_0$  at the point in the  $(m,n)$  minimax orbit. The region below  $\mathcal{L}_0$  and above  $\mathcal{L}_{-n}$  is the region which crosses the partial barrier from below to above on one iteration of the map. Similarly the region below  $\mathcal{L}_{-n}$  but above  $\mathcal{L}_0$  crosses the partial barrier from above to below upon one iteration of the map. The union of these two regions has a figure-of-eight structure called a turnstile.

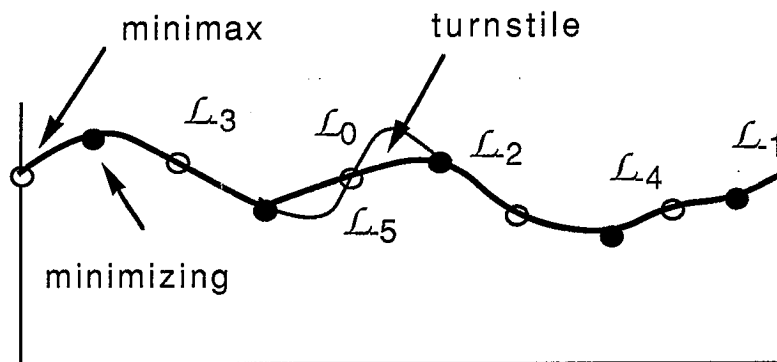


Figure 3.5. Partial barrier of a  $(2,5)$  orbit.

A more naturally constructed partial barrier is a partial separatrix. Consider the minimizing  $(m,n)$  periodic orbit; it has two minimizing homoclinic orbit, the advancing (upper)  $(m,n)^+$  orbit and the retreating (lower)  $(m,n)^-$  orbits. Consider the advancing  $(m,n)^+$  orbit (Fig. 3.6); choose a principal gap in this orbit. Connect it with a segment of the unstable manifold  $u_0$ ; this segment naturally goes through the minimax homoclinic orbit. Take preimages of the unstable segment  $u_0$  and they

converge to the minimizing periodic orbit and together they form part of the unstable manifold of the minimizing periodic orbit. We also take images of the stable segment  $s_0$  in the principal gap. The set  $\{ u_t, t \leq 0 \} \cup \{ s_t, t > 0 \}$ , which is the union of the unstable segments for  $t \leq 0$  and the stable segments for  $t > 0$ , forms the upper partial separatrix. If the stable and unstable manifolds join together smoothly, this gives the usual upper separatrix for an integrable system. Again, the turnstile is formed by taking the preimage of the partial barrier. Each segment has a preimage on the partial barrier, except for  $s_1$ , which becomes  $s_0$  in the principal gap. The left lobe of the turnstile is the set of points which will cross the partial separatrix from below to above in the next iteration and the right lobe from above to below.

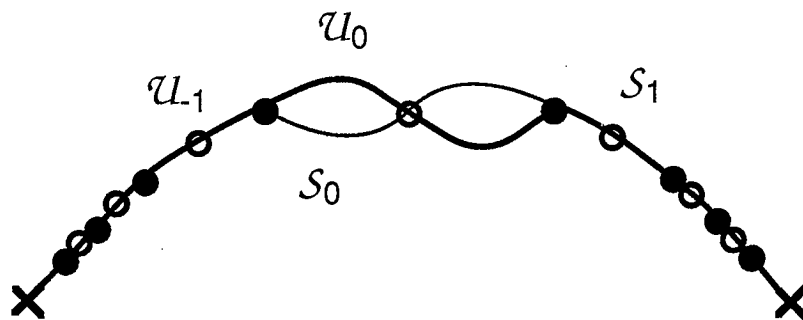


Figure 3.6. The upper partial separatrix.

The lower partial separatrix is formed using the same procedure, which is the union of the unstable segments for  $t \leq 0$  and the stable segments for  $t > 0$ . The region bounded by the upper and lower partial separatrices is the  $(m,n)$  resonance [Channon and Lebowitz 1980, MacKay, Meiss and Percival 1987]. A resonance has an island chain structure. We call the island formed in the principal gap the central island. Fig. 3.7 shows the

upper and lower partial separatrices for the (1,3) resonance of the standard map. The upper turnstile area of an  $(m,n)$  resonance gives the total flux from the  $(m,n)$  resonance to all other resonances above it. Similarly, the lower turnstile represents the flux to all resonances below the  $(m,n)$  resonance.

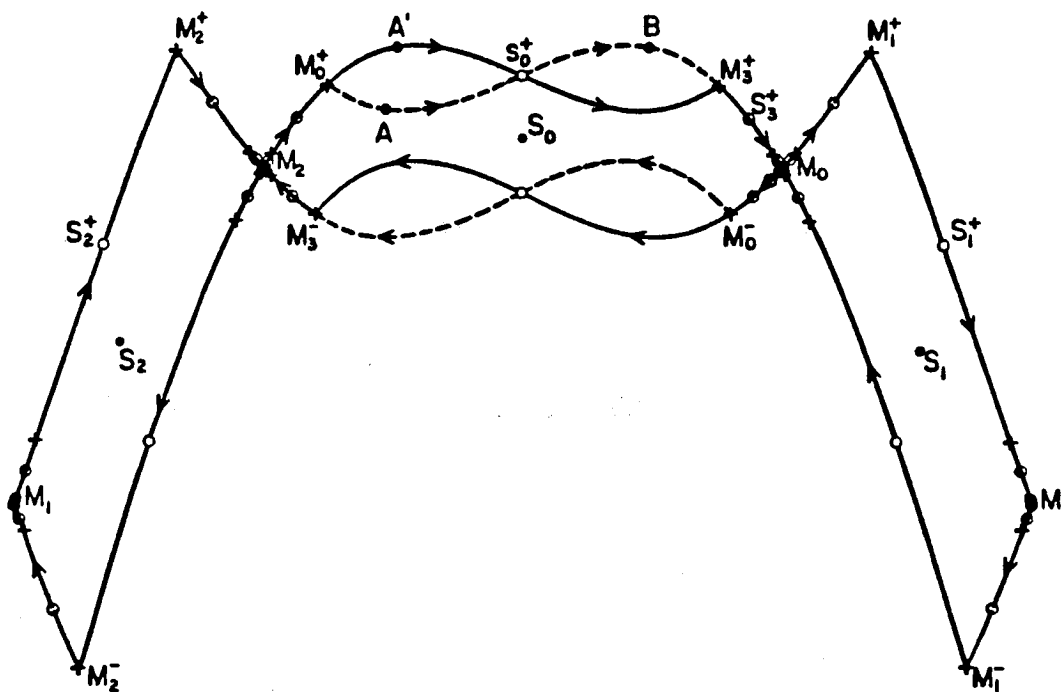


Figure 3.7. (1,3) resonance for the standard map in symmetry coordinates.  $M_t$  are points on the period 3 minimizing orbit,  $S_t$  are points on the period three minimax orbit;  $M_t^+$  and  $M_t^-$  are points on the minimizing homoclinic orbits,  $S_t^+$  and  $S_t^-$  are points on the minimax homoclinic orbit.



In general, a different choice of homoclinic point at which to switch from unstable to stable manifold gives different shape to the resonance; however, the area of the resonance is independent of the choice of the switching point. A consistent choice is important in order that resonances of different rotation numbers do not intersect with each other and fit together to give a complete partition of phase space.

For maps with dominant symmetry, there is a natural choice of the principal gap, which is the dominant gap for all the minimizing orbits. The minimizing orbits never have points on the symmetry line; therefore, the two points nearest to, and on different sides of the symmetry line define the dominant gap of a minimizing orbit. This fits all resonances nicely in phase space, and they never overlap. Since all turnstiles are in the dominant gap, their union forms a "chimney" on the dominant line. All flux between resonances takes place in this region.

We can also form partial barriers from cantori. The idea is again to use the stable and unstable sets. Define the stable set  $s(x,y)$  of  $(x,y)$  to be the set of points  $(x',y')$  such that the distance between  $T^t(x',y')$  and  $T^t(x,y)$  goes to zero as  $t \rightarrow \infty$ ; and the unstable set  $u(x,y)$  of  $(x,y)$  to be the set of points  $(x',y')$  such that the distance between  $T^t(x',y')$  and  $T^t(x,y)$  goes to zero as  $t \rightarrow -\infty$ . Choose a principal gap in the cantorus; the end points have the same stable set and unstable set, since they converge together in both directions of time. Furthermore, the minimax orbit belongs to both sets. We denote the stable and unstable sets in the principal gap by  $s_0$  and  $u_0$ . Let  $\mathcal{L}_t = T^t(s_0)$  for  $t > 0$ , and  $\mathcal{L}_t = T^t(u_0)$  for  $t \leq 0$ . If there is only one family of gaps, then the set  $\{ \mathcal{L}_t, t \in \mathbb{Z} \}$  forms a rotational circle (A Cantor set always has a countable set of gaps, so there is at most a

countable number of families, so we can repeat the construction for each family). The preimage of  $\mathcal{L}_1$  is  $\mathcal{S}$ .  $\mathcal{S}$  and  $\mathcal{U}$  must cross each other at the minimax orbit, and they form a turnstile.

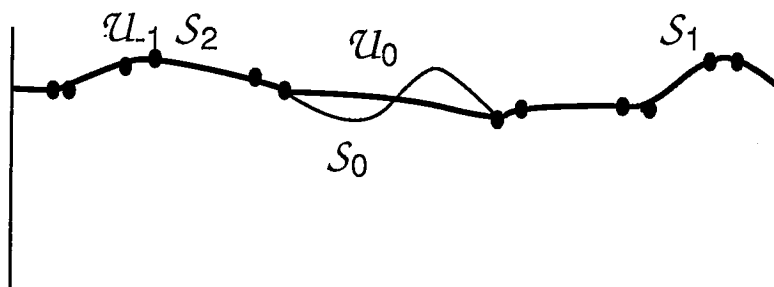


Figure 3.8. Partial barrier of a cantorus.

### 3.3.3. Flux, Areas and Actions

#### A. The Fundamental Area Formula

Starting from the action principle for twist maps, we can derive formulas for the flux across and the area under partial barriers constructed from both cantori and homoclinic orbits

Let  $c$  be a directed curve in the plane. Parametrize it by  $\lambda$  ranging over  $[0,1]$  so  $c(\lambda) = \{ x(\lambda), y(\lambda) \}$ . Define the algebraic area,  $A$ , under  $c$  to be the signed area bounded by the loop formed from  $c$ , the vertical lines  $x = x(0)$  and  $x = x(1)$  and the horizontal line  $y = 0$ . The direction of loop is set as that of increasing  $\lambda$  over  $c$ . Let  $c'$  be the image of  $c$  under the map  $T$ . The area under  $c'$  is denoted  $A'$ . Integrate (3.1.5) yields

$$\Delta F = F(x(1), x'(1)) - F(x(0), x'(0)) = A' - A \quad (3.3.1)$$

Thus the difference between the areas is the difference between the generating functions evaluated at the end points. Note the curve  $c$  does not enter explicitly in (3.3.1).

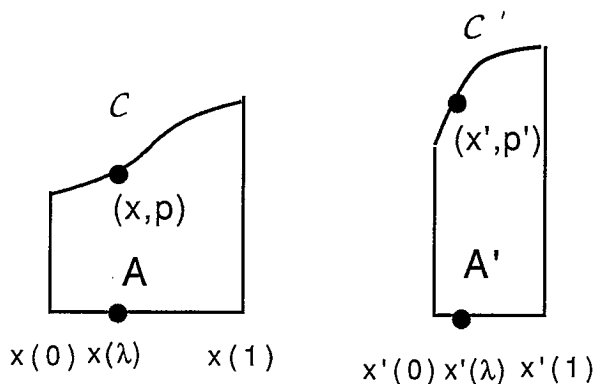


Figure 3.9. Illustration of the fundamental area formula.

## B. Flux through of a Pair of Periodic Orbits

As a simple application of the fundamental formula, let us calculate the flux through the turnstile in a pair of periodic orbits. Let  $A_0$  be the area under the segment  $\mathcal{L}_0$  connecting the point  $M_0$  on the minimizing orbit with  $S_0$  on the maximax orbit. Similarly,  $A_t$  are the iterates of this area. The fundamental formula reads

$$A_t - A_{t-1} = F(S_{t-1}, S_t) - F(M_{t-1}, M_t)$$

The area of the turnstile is  $A_0 - A_{-n}$

$$\begin{aligned} \text{Flux} &= A_0 - A_{-n} = \sum_{t=-n+1}^0 [ F(S_{t-1}, S_t) - F(M_{t-1}, M_t) ] \\ &= \sum_{t=0}^{n-1} [ F(S_{t-1}, S_t) - F(M_{t-1}, M_t) ] = \Delta W(m, n) \end{aligned} \quad (3.3.2)$$

The flux is simply the difference in action between the maximax and the minimizing orbits; therefore it does not depend on the choice of the curve  $\mathcal{L}_0$  connecting  $M_0$  and  $S_0$ .

## C. Stable and Unstable Segments

Consider an orbit  $\{ M_t \}$  which is hyperbolic (this includes the case of a cantorus whose Lyapunov exponent is positive). Any two points on the stable or the unstable manifold of  $\{ M_t \}$  are called future or past asymptotic; if they are on both manifolds, they are homoclinic. The area under a stable or unstable segment between two homoclinic points can be calculated by the fundamental formula (3.3.1).

Parametrize the stable segment  $S_t$  of a future asymptotic pair  $\{Y_t\}$  and  $\{Z_t\}$  by  $\lambda$ , such that  $S_t(0) = Y_t$ ,  $S_t(1) = Z_t$  and  $S_{t+1}(\lambda) = TS_t(\lambda)$ . Then the area under  $S_t$  is

$$A^{S_t} = A^{S_{t+1}} - \Delta F_t = A^{S_{t+2}} - \Delta F_{t+1} - \Delta F_t = \dots = -\sum_{j=t}^{-\infty} \Delta F_j = -\Delta W^f_t$$

Similarly if  $\{Y_t\}$  and  $\{Z_t\}$  are past asymptotic, and  $A^{U_t}$  the area under the unstable segment, then

$$A^{U_t} = \Delta W^p_t$$

If  $\{Y_t\}$  and  $\{Z_t\}$  are homoclinic, the signed area between the unstable and stable segments is independent of  $t$

$$A^{U_t} - A^{S_t} = \Delta W = \sum_{j=-\infty}^{\infty} \Delta F_j \quad (3.3.3)$$

Once the total area under a single segment is known, the total area under all the segments  $S_t$  after time  $t$  is

$$A^f_t = \sum_{k=1}^{\infty} A^{S_{t+k}} = -\sum_{k=1}^{\infty} \Delta W^f_{t+k} = -\sum_{k=1}^{\infty} k \Delta F_{t+k}$$

Similarly the area under all unstable segments at time  $t$  and before is

$$A^p_t = \sum_{k=-\infty}^0 A^{U_{t+k}} = \sum_{k=-\infty}^0 \Delta W^p_{t+k} = -\sum_{k=-\infty}^0 k \Delta F_{t+k}$$

Finally if  $\{Y_t\}$  and  $\{Z_t\}$  are homoclinic, the total area under all the future stable segments and the past unstable segments is

$$A_t = A^p_t + A^f_t = - \sum_{k=-\infty}^0 k \Delta F_{t+k} \quad (3.3.4)$$

In general, (3.3.4) is not independent of  $t$ , in fact,  $A_{t+1} = A_t + \Delta W$ , where  $\Delta W$  is given by (3.3.3). Our applications involve pair of orbits whose action difference is zero; therefore, the area is independent of  $t$ .

#### D. Area Devil's Staircase Function

The area under the gaps of a cantorus is given by

$$A = - \sum_{t=-\infty}^{\infty} t \left[ F(x^r_t, x^r_{t+1}) - F(x^l_t, x^l_{t+1}) \right] \quad (3.3.6)$$

where  $\{ M^l_t \}$  is the orbit of the left endpoints of a gap in the cantorus and  $\{ M^r_t \}$  the orbit of the right end points. The action difference between  $\{ M^l_t \}$  and  $\{ M^r_t \}$  vanishes, hence the area is independent of the choice of crossing points. The total area under the partial barrier is the sum of (3.3.6) and the area under the cantorus. Since hyperbolic cantori have zero length [MacKay, Meiss and Percival 1987], the area under the partial barrier is given by (3.3.6).

We can also find the area under the upper and lower partial separatrices of a resonance. It suffices to consider the period one case, since any period  $n$  resonance is period one under  $n$  iteration of the map. The upper area is given by (3.3.4) where  $\{ Y_t \} = \{ M^+_t \}$  is the minimizing homoclinic orbit, and  $\{ Z_t \} = \{ M^+_{t+1} \}$ , its right shift. Subtracting the action of the fixed point  $\{ x_F \}$  from the first term in the bracket while simultaneously adding it to the second term, we have

$$A^+(0/1) = \sum_{t=-\infty}^{\infty} \left[ F(x^+_t, x^+_{t+1}) - F(x_F, x_F) \right]$$

Now it is easy to find the area under the upper partial separatrix of a period  $n$  resonance

$$A^+(m/n) = n \sum_{t=-\infty}^{\infty} \left\{ \sum_{i=0}^{n-1} \left[ F(x^+_{tn+i}, x^+_{tn+i+1}) - F(x_i, x_{i+1}) \right] \right\} \quad (3.3.7a)$$

where  $\{ x^+_t \}$  is the  $(m,n)^+$  minimizing orbit. The term in the curly bracket can be interpreted as the action sum of the map  $T^n$ .

Similarly, the area under the lower partial separatrix is

$$A^-(m/n) = n \sum_{t=-\infty}^{\infty} \left\{ \sum_{i=0}^{n-1} \left[ F(x_i, x_{i+1}) - F(x^-_{tn+i}, x^-_{tn+i+1}) \right] \right\} \quad (3.3.7b)$$

where  $\{ x^-_t \}$  is the  $(m,n)^-$  minimizing orbit.

Now, we can define area as a function of frequency: for rational frequency, it has two values, the upper area is given by the area under the upper (partial) separatrix, and the lower area is the area under the lower (partial) separatrix; For irrational frequency, it is either given by the area under the rotational invariant torus or the area under the invariant cantorus with that frequency. So the area as a function of frequency has a jump at each rational number and is monotonically increasing; therefore, it is a devil's staircase function. A devil's staircase function is called complete if the sum of its jumps equals to its total variation. An area devil's staircase function is plotted in Figure 4.4.



### E. Flux Through Partial Separatrices and Cantori

Formula (3.3.3) can be used to calculate the total flux through a partial separatrix or a partial barrier formed from a cantorus. In the former case, we choose  $\{ Y_t \}$  to be the minimizing homoclinic orbit  $\{ M_t^\pm \}$ , and  $\{ Z_t \}$  the minimax homoclinic orbit  $\{ S_t^\pm \}$ . The total flux through the upper partial separatrix is called the upper flux of the resonance,  $\Delta W^+(m,n)$ ; and the total flux through the lower partial separatrix is called the lower flux,  $\Delta W^-(m,n)$ .

For a partial barrier formed from a cantorus,  $\{ Y_t \}$  is an orbit  $\{ M_t \}$  on the cantorus,  $\{ Z_t \}$  is the minimax homoclinic orbit  $\{ S_t \}$  to the cantorus. This is precisely Mather's  $\Delta W_V$  (section 3.2.4).

We can also find the flux from one resonance to another. In this case, we choose orbits which are heteroclinic between the resonances: Take  $\{ Y_t \}$  the left heteroclinic orbit  $\{ H_t^l \}$ ,  $\{ Z_t \}$  the right heteroclinic orbit  $\{ H_t^r \}$  (see Fig.3.5). Numerically, we find  $\{ H_t^l \}$  is locally minimizing, and  $\{ H_t^r \}$  is locally minimax (see section 5.3.3 for more detail).

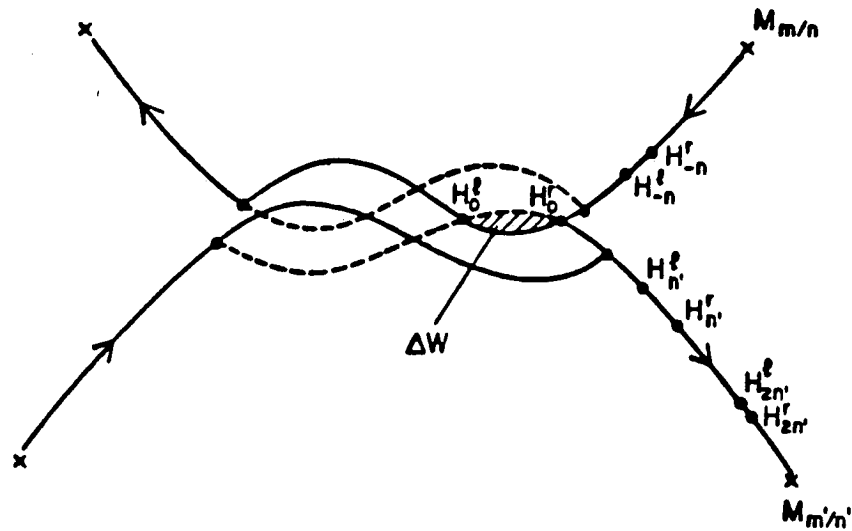


Figure 3.10. Flux between two resonances.

In fact, there are possibly four such pairs of heteroclinic orbits. For a given pair of resonances  $(m,n)$  and  $(m',n')$  ( $m/n < m'/n'$ ), the heteroclinic orbit from upper partial separatrix of  $(m,n)$  to the upper partial separatrix of  $(m'/n')$  is denoted by  $(m/n^+; m'/n'^+)$ , and the corresponding flux is  $\Delta W(m/n^+ \rightarrow m'/n'^+)$ . Similarly, we define  $\Delta W(m/n^+ \rightarrow m'/n'^-)$  etc.. The function  $\Delta W(m/n^+ \rightarrow v)$  can be interpreted as the flux of the upper partial separatrix of  $(m,n)$  resonance to all resonances beyond the frequency  $v$ . Therefore for rational  $v = m'/n'$ , there is a jump in  $\Delta W(m/n^+ \rightarrow v)$ , one corresponds to the flux from the  $(m,n)$  resonance to all other resonances beyond frequency  $m'/n'$  including the  $(m',n')$  resonance, which is  $\Delta W(m/n^+ \rightarrow m'/n'^+)$ ; the other excluding the resonance, which is  $\Delta W(m/n^+ \rightarrow m'/n'^-)$ . For irrational frequency  $v$ , there is only one  $\Delta W(m/n^+ \rightarrow v)$ . This defines a devil's staircase function, which is discontinuous at

each rational frequency, and continuous at each irrational frequency. We shall see that this function is complete.

The flux from one resonance to another is

$$F(m/n \rightarrow m'/n') = [ \Delta W(m/n^+ \rightarrow m'/n'^+) - \Delta W(m/n^+ \rightarrow m'/n'^-) ] \\ - [ \Delta W(m/n^- \rightarrow m'/n'^+) - \Delta W(m/n^- \rightarrow m'/n'^-) ] \quad (3.3.8)$$

Flux satisfies the balance equation by zero net flux

$$\sum_{m'/n'} F(m/n \rightarrow m'/n') = \sum_{m'/n'} F(m'/n' \rightarrow m/n) \quad (3.3.9)$$

which says that the total flux out of a resonance is equal to the total flux into that resonance.

### 3.4. Relation to Average Lagrangian

#### 3.4.1. Average Lagrangian

Of all the minimizing configurations in the F-K model, the one with the lowest action is the ground state. It has trivially frequency zero. Consider a chain of atoms in an external field of strength  $E$ , we want to find the ground state frequency  $\nu$  as a function of  $E$ . The action for a finite segment of state  $\{x_t\}$  is

$$W(\{x_t\}) = \sum_{t=M}^N F(x_t, x_{t+1}) - E(x_{t+1} - x_t) \quad (3.4.1)$$

Therefore, the ground state is given by minimizing the action per site,  $L(\nu) - E\nu$ , where  $L(\nu)$  is the average Lagrangian for a minimizing orbit of frequency  $\nu$

$$L(\nu) = \lim_{N' - N \rightarrow \infty} \frac{\sum_{t=N}^{N'-1} F(x_t, x_{t+1})}{N' - N} \quad (3.4.2)$$

We show  $L(\nu)$  is a well defined function of frequency  $\nu$ .

For a periodic orbit, the average Lagrangian is obviously equal to the action of the orbit divided by the period. For a homoclinic orbit  $\{x^+_j\}$  with frequency  $m/n$ ; the average Lagrangian is

$$L\left(\frac{m}{n}\right) = \lim_{k' - k \rightarrow \infty} \frac{\sum_{i=k}^{k'-1} \sum_{l=0}^{n-1} F(x^+_{in+l}, x^+_{in+l+1})}{(k' - k)n}$$

Since the homoclinic orbit converges exponentially to the minimizing  $(m,n)$  orbit  $\{x_j\}$ , the limit in  $k'-k$  exists, and is

$$L\left(\frac{m}{n}\right) = \frac{\sum_{t=0}^{n-1} F(x_t, x_{t+1})}{n} = L\left(\frac{m}{n}\right)$$

This shows  $L(v)$  exists for each rational frequency.

To show that  $L(v)$  exists for irrational frequencies, we need the following ergodic theorem:

Theorem [Arnol'd and Avez 1968]: If  $R$  is a rotation of the unit circle represented by  $[0,1)$  through an irrational angle, and  $f$  is a Riemann integrable function, then the time average of the function  $f$  is equal to its spatial average, i.e.

$$\lim_{N \rightarrow \infty} \frac{\sum_{t=0}^{N-1} f(x_t, x_{t+1})}{N} = \int_0^1 f(\theta) d\theta \quad (3.4.3)$$

For an irrational frequency  $v$ , the minimizing orbit is either an invariant circle or a cantorus. In either case there exists a monotonically increasing function  $f$  such that:

$$x_n = f(nv + \alpha) \quad (3.4.4)$$

Besides its explicit dependence on the frequency  $v$  in its argument, the functional form of  $f$  also depends on the frequency  $v$ . For a subcritical invariant circle,  $f$  is analytic, and for a hyperbolic cantorus,  $f$  can be written as the sum of step functions:

$$f(x) = \sum_{t=-\infty}^{\infty} f_t H(x - x_t) \quad (3.4.5)$$

where  $H(x) = 1, x \geq 0; H(x) = 0, x < 0$  is the Heaviside function and  $f_t$  is the discontinuity across a gap of the cantorus. For convenience, we shall let  $x(\theta) = f(\theta)$  in the following.

It follows from the ergodic theorem that for an irrational frequency  $\nu$ :

$$L(\nu) = \int_0^1 F(x(\theta), x(\theta+\nu)) d\theta \quad (3.4.6)$$

Indeed, this average action is the action for an invariant torus or cantorus as formulated by Percival (see section 2.3.3).

$L(\nu)$  is a convex function of  $\nu$ . Given  $\nu_1, \nu_2$ , we construct a state by choosing a fraction  $\lambda$  of points from the minimizing orbit of frequency  $\nu_1$ , and a fraction of  $(1-\lambda)$  from the minimizing orbit of frequency  $\nu_2$ . This state has frequency  $\lambda\nu_1+(1-\lambda)\nu_2$  and average Lagrangian  $\lambda L(\nu_1)+(1-\lambda)L(\nu_2)$ . However, it is not the minimizing orbit of frequency  $\lambda\nu_1+(1-\lambda)\nu_2$  since it is not ordered, hence

$$\lambda L(\nu_1) + (1-\lambda)L(\nu_2) \leq L(\lambda\nu_1+(1-\lambda)\nu_2) \quad (3.4.7)$$

i.e.,  $L(\nu)$  is a convex function. Any convex function  $L(\nu)$  has monotonically increasing left and right derivatives  $L^-(\nu)$  and  $L^+(\nu)$  which are equal almost everywhere [Rudin 1987].

The function  $\nu(E)$  now is given implicitly by minimizing  $L(\nu)-E\nu$  with respect to  $\nu$ , giving

$$L^-(\nu) \leq E \leq L^+(\nu) \quad (3.4.8)$$

Hence  $\nu(E)$  is continuous, monotonically increasing, and has a flat plateau at each rational frequency. This is a devil's staircase function.

### 3.4.2. Area Devil's Staircase and Average Lagrangian

Now we show that the derivative of the average action with respect to frequency gives the area devil's staircase:

$$A^{\pm}\left(\frac{m}{n}\right) = \left. \frac{d^{\pm}L(v)}{dv} \right|_{v = m/n}$$

$$A(v) = \left. \frac{dL(v)}{dv} \right|_{v \text{ irrational}} \quad (3.4.9)$$

Let us calculate  $L^{+}(v)$  and  $L^{-}(v)$  for a rational frequency  $m/n$ . To approach the minimizing  $(m,n)$  orbit from above and below, we use the minimizing  $(mk+m_1, nk+n_1)$  and  $(mk+m_2, nk+n_2)$  periodic orbits, respectively, and let the integer  $k$  go to infinity [MacKay, Meiss and Percival 1987]. The integers  $m_1, n_1, m_2, n_2$  are uniquely determined by:

$$m_1n - n_1m = 1, \quad m_2n - n_2m = -1, \quad n_1, n_2 < n$$

As  $k$  goes to infinity, these two orbits limit to the upper and lower minimizing orbits homoclinic to the  $(m,n)$  orbit, respectively. Let  $\{x^{k+}_t\}$  denote the minimizing  $(mk+m_1, nk+n_1)$  orbit, then the derivative of  $L(m/n)$  from above is given by:

$$\begin{aligned}
L^{+} \left( \frac{m}{n} \right) &= \lim_{k \rightarrow \infty} \frac{1}{mk+m_1 / nk+n_1 - m/n} \left[ \frac{\sum_{t=0}^{nk+n_1-1} F(x_t, x_{t+1})}{nk+n_1} - \frac{\sum_{t=0}^{n-1} F(x_t, x_{t+1})}{n} \right] \\
&= \lim_{k \rightarrow \infty} \left[ n \sum_{t=0}^{nk+n_1-1} F(x_t, x_{t+1}) - nk+n_1 \sum_{t=0}^{n-1} F(x_t, x_{t+1}) \right] \\
&= n \sum_{i=-\infty}^{\infty} \sum_{t=0}^{n-1} [F(x_{in+t}^+, x_{in+t+1}^+) - F(x_t, x_{t+1})]
\end{aligned}$$

Similar result holds for  $L^{-}(v)$ . Since these equations are identical to Eq(3.3.7), this implies Eq(3.4.9) for rational frequency.

For a subcritical circle with irrational frequency,  $x(\theta)$  in (3.4.6) is analytic, so we can interchange the order of differentiation and integration; therefore,

$$\begin{aligned}
\frac{dL(v)}{dv} &= \int_0^1 F_2(x(\theta), x(\theta+v)) \frac{dx(\theta+v)}{d\theta} d\theta + \\
&\int_0^1 F_2(x(\theta), x(\theta+v)) \frac{\partial x(\theta+v)}{\partial v} d\theta + \int_0^1 F_1(x(\theta), x(\theta+v)) \frac{\partial x(\theta)}{\partial v} d\theta
\end{aligned}$$

Where again the subscripts 1 and 2 of  $F$  denote derivatives with respect to the first and second argument of  $F$ . The partial derivative  $\partial x(\theta)/\partial v$  is due to the dependence of the orbit  $x(\theta)$  on the frequency  $v$ . Using eq(3.1.4), this becomes:



$$\begin{aligned} \frac{dL(v)}{dv} &= \int_0^1 p(\theta+v) dx(\theta+v) + \int_0^1 p(\theta+v) \frac{\partial x(\theta+v)}{\partial v} d\theta - \int_0^1 p(\theta) \frac{\partial x(\theta)}{\partial v} d\theta \\ &= \int_0^1 p(\theta) dx(\theta) = A(v) \end{aligned}$$

In fact, the condition that  $x(\theta)$  be analytic can be relaxed. As long as  $x(\theta)$  is monotone and transitive, the above equation holds so that it is applicable to the critical invariant circle.

For a hyperbolic cantorus, since  $x(\theta)$  is a sum of step functions, the integrand  $F(x(\theta), x(\theta+v))$  in (3.4.6) is also a sum of step functions:

$$F(x(\theta), x(\theta+v)) = \sum_{t=-\infty}^{\infty} F_t H(\theta - \theta_t) \quad (3.4.10)$$

where

$$\theta_t \equiv vt + \alpha \pmod{1} \quad (3.4.11)$$

$F_t$  is nothing but the discontinuity of the generating function at the gap in the cantorus. Let  $\{x_t^l\}$  be the orbit of the left endpoints of a gap in the cantorus,  $\{x_t^r\}$  the orbit of the right endpoints, thus:

$$F_t \equiv F(x_t^r, x_{t+1}^r) - F(x_t^l, x_{t+1}^l) \quad (3.4.12)$$

The function on the right hand side of (3.4.10) is periodically extended to the whole real line in  $\theta$ .

Taking the derivative of the average action  $L(v)$  yields:

$$\begin{aligned} \frac{dL(v)}{dv} &= \int_0^1 \sum_t \left[ F_1(x_t^r, x_{t+1}^r) \frac{\partial x_t^r}{\partial v} + F_2(x_t^r, x_{t+1}^r) \frac{\partial x_{t+1}^r}{\partial v} \right. \\ &\quad \left. - F_1(x_t^l, x_{t+1}^l) \frac{\partial x_t^l}{\partial v} - F_2(x_t^l, x_{t+1}^l) \frac{\partial x_{t+1}^l}{\partial v} \right] H(\theta - \theta_t) d\theta - \sum_t F_t \frac{d(vt + \alpha)}{dt} \\ &= \int_0^1 \sum_t \left( p_{t+1}^r \frac{\partial x_{t+1}^r}{\partial v} - p_{t+1}^l \frac{\partial x_{t+1}^l}{\partial v} \right) H(\theta - \theta_{t+1} + v) d\theta \\ &\quad - \int_0^1 \sum_t \left( p_t^r \frac{\partial x_t^r}{\partial v} - p_t^l \frac{\partial x_t^l}{\partial v} \right) H(\theta - \theta_t) d\theta - \sum_t F_t \end{aligned}$$

where in the last equality, the integrand of the first term is the shift in the  $\theta$  variable by an angle  $v$  of that of the second term. These two terms cancel due to the periodicity of the integrand. Thus:

$$dL(v)/dv = - \sum_{t=-\infty}^{\infty} t \left[ F(x_t^r, x_{t+1}^r) - F(x_t^l, x_{t+1}^l) \right]$$

This is exactly the same formula as Eq(3.3.6).

Aubry [1982] showed that the derivative of the average Lagrangian with respect to frequency is a complete devil's

staircase when the lower bound of the Lyapunov exponents of all the minimizing orbits is positive; this implies all invariant circles are destroyed and all cantori are hyperbolic; Since we have shown that the area devil's staircase is identical to the derivative of the average action, resonances fill the entire phase space.

### 3.4.3. Implications for $\Delta W(v \rightarrow v')$

As an immediate consequence of the completeness of the area devil's staircase function, the function  $\Delta W(m/n^+ \rightarrow v)$  and  $\Delta W(m/n^- \rightarrow v)$  are complete devil's staircase functions: since the total area of partial barriers of cantori is zero, there is no flux to these partial barriers.

In Figure 3.11, we plot the function  $\Delta W(1/3 \rightarrow v)$ . Figure 3.11a is for the standard map at  $k = 1.40$ . In fact, it is an overlay of two functions,  $\Delta W(1/3^+ \rightarrow v)$  for  $v > 1/3$ , and  $\Delta W(1/3^- \rightarrow v)$  for  $v < 1/3$ . The heteroclinic orbits are found using the orbit extension method (chapter 5). The behavior of the function  $\Delta W(v \rightarrow v')$  is more clearly illustrated in Figure 3.11b. We plot the function  $\Delta W(1/3^+ \rightarrow v)$  for  $v > 1/3$  for the sawtooth map at  $k = 0.08$  (the sawtooth map will be discussed in detail in chapter 4). The upper turnstile of (1,3) resonance overlaps with the lower turnstile of (2,5) resonance but not with its upper turnstile. Therefore, upon one single step, a point trapped in (1,3) resonance can go no farther up than the (2,5) resonance. In the figure,  $v$  is varied over all the rationals from  $1/3$  to  $2/5$  up to period 28. The point shown at  $v = 1/3$  is the upper turnstile area of the (1,3) resonance, i.e.,  $\Delta W(m, n^+)$ . The function  $\Delta W(1/3^+ \rightarrow v)$  is necessarily decreasing as  $|v - 1/3|$  increases because the overlap area must be smaller than the area of the turnstile itself. Furthermore  $\Delta W(1/3^+ \rightarrow v)$  has discontinuities across every rational since the overlap with the  $(m, n)^+$  turnstile is significantly different from that with the  $(m, n)^-$  turnstile, because the area corresponding to the  $(m, n)$  resonance itself intervenes. For example, the area of overlap of the  $(1, 3)^+$  with the  $(3, 8)^-$  turnstile is  $\Delta W(1/3^+ \rightarrow 3/8^-) = 0.95 \times 10^{-3}$ , while the overlap of the  $(1, 3)^+$  turnstile with the  $(3, 8)^+$  turnstile is

considerably less ( $0.49 \times 10^{-3}$ ). The function  $\Delta W(1/3^+ \rightarrow \nu')$  is a complete devil's staircase. In fact, the sum of the total jumps up to period 27 is  $1.392 \times 10^{-3}$ , while the total area of the upper turnstile is  $1.400 \times 10^{-3}$ , so the difference is less than  $0.8 \times 10^{-5}$ .

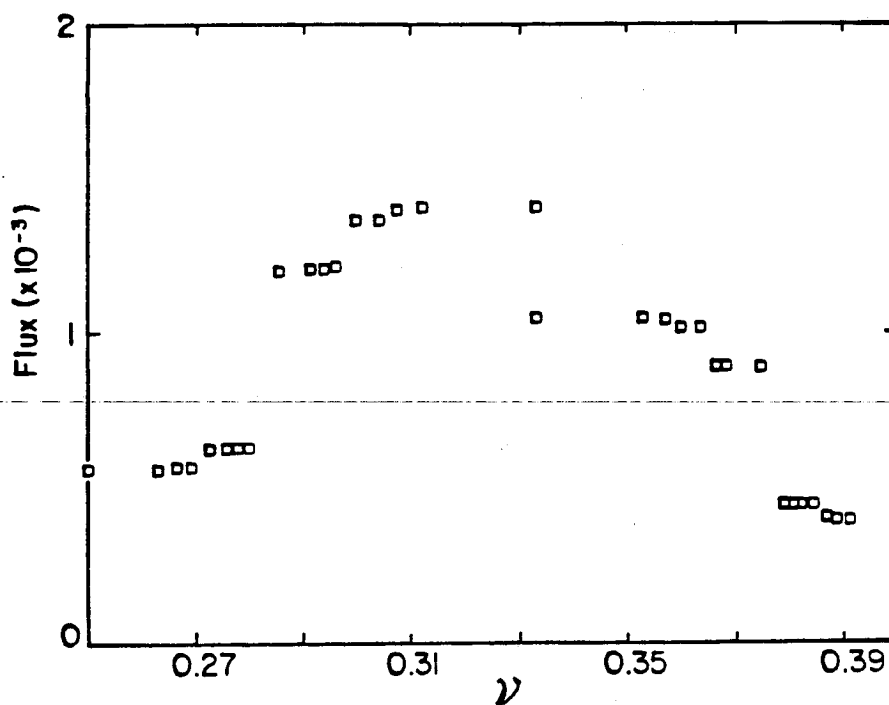


Figure 3.11a. The function  $\Delta W(1/3^+ \rightarrow \nu)$  for the standard map at  $k = 1.4$ .

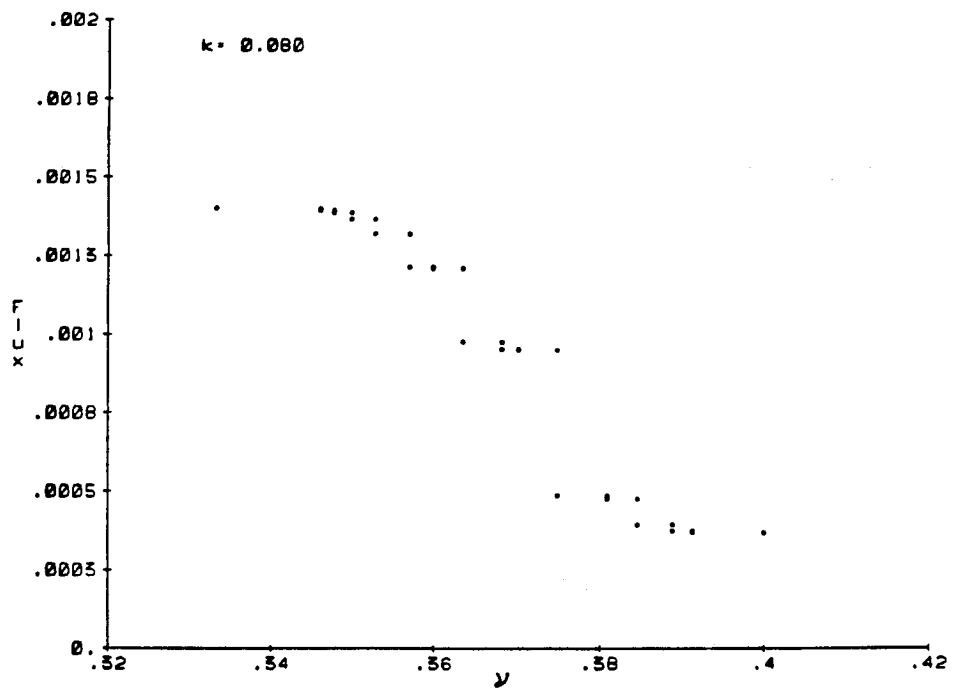


Figure 3.11b. The function  $\Delta W(1/3^+ \rightarrow \nu)$  for the sawtooth map at  $k = 0.08$ .

### 3.4.4. Application to Calculating the Area Under an Invariant Circle

Equation (3.4.7) gives a convenient formula for computing the area under an invariant circle, which is useful in practical applications [see, eg. Dana and Reinardt, 1987]. We use periodic orbits whose frequencies are the successive rational approximants of the irrational frequency of the invariant circle to obtain the derivative of the average action.

The results for the golden mean invariant circle at two parameter values for the standard map are shown in table 1.  $A(l)$  is the area calculated using rational approximants at level  $l$ , the convergence rate  $\delta(l)$  is defined as:

$$\delta(l) = (A(l+1) - A(l))/(A(l) - A(l-1)) \quad (3.4.10)$$

Notice this method converges considerably faster than the linear interpolation method (see table 2). In the latter case, we compute the area under a periodic orbit by connecting neighboring points with straight lines. The convergence rate is so slow that it is virtually impossible to find the area without the cost of extremely long approximating periodic orbits.

Table 1

Area under the golden mean invariant circle calculated by the method mentioned in this paper at different parameter values. Orbits are calculated to precision  $10^{-12}$ .  $A(l)$  is the area calculated from the  $l$ (th) rational approximating periodic orbit of the golden mean circle,  $\delta(l)$  is the convergence rate at the level  $l$  defined by eq(3.4.10).

$k = 0.9$

Level	M/N	A(l)	$\delta(l)$
4	5/8	0.6104229	—
5	8/13	0.6164991	-0.39350
6	13/21	0.6141081	-0.39701
7	21/34	0.6150573	-0.39712
8	34/55	0.6146804	-0.39303
9	55/89	0.6148285	-0.38682
10	89/144	0.6147712	-0.38299
11	144/233	0.6147932	-0.38205
12	233/377	0.6147848	-0.38165
13	377/610	0.6147880	-0.38197
14	610/987	0.6147868	—



$$k = k_c = 0.971635406$$

Level	M/N	A(l)	$\delta(l)$
4	5/8	0.6103039	—
5	8/13	0.6157595	-0.37235
6	13/21	0.6137281	-0.37417
7	21/34	0.6144882	-0.37174
8	34/55	0.6142056	-0.37373
9	55/89	0.6143112	-0.37233
10	89/144	0.6142719	-0.37326
11	144/233	0.6142866	-0.37266
12	233/377	0.6142811	-0.37304
13	377/610	0.6142831	-0.37280
14	610/987	0.6142824	-0.37295
15	987/1597	0.6142827	-0.37286
16	1597/2584	0.6142826	—

Table 2

Area under the golden mean invariant circle calculated by linearly interpolating the neighboring points of the periodic orbits. All Orbits are calculated to precision  $10^{-12}$ .  $A(l)$  is the area calculated from the  $l$ (th) rational approximating periodic orbit of the golden mean circle,  $\delta(l)$  is the convergence rate at the level  $l$  defined by eq(3.4.10).

$k = 0.9$

Level	M/N	A(l)	$\delta(l)$
4	5/8	0.7081045	—
5	8/13	0.6708096	0.38843
6	13/21	0.6563233	0.96231
7	21/34	0.6423830	0.60258
8	34/55	0.6339828	0.78309
9	55/89	0.6274048	0.67063
10	89/144	0.6229934	0.69026
11	144/233	0.6199483	0.63581
12	233/377	0.6180122	0.63471
13	377/610	0.6167834	0.61881
14	610/987	0.6102303	—

$$k = k_c = 0.971635406$$

Level	M/N	A(l)	$\delta(l)$
4	5/8	0.7058759	—
5	8/13	0.6688289	0.38648
6	13/21	0.6545110	0.93595
7	21/34	0.6411104	0.58059
8	34/55	0.6333304	0.75352
9	55/89	0.6274674	0.66256
10	89/144	0.6235831	0.71557
11	144/233	0.6208036	0.69140
12	233/377	0.6188818	0.70734
13	377/610	0.6175225	0.70114
14	610/987	0.6165694	0.70601
15	987/1597	0.6158965	0.70456
16	1597/2584	0.6154224	—

### 3.5. The Markov Transport Model

The numerical observation that an irregular component has a positive Lyapunov exponent allows a stochastic description of the transport process: since any uncertainty in the initial positions in the phase space is amplified exponentially under iterations of the map till it is of the same order as the size of the system under study. So we are forced to "coarse grain" the actual underlying dynamics, and give a probabilistic description of the transport.

In the supercritical regime partitioning of the phase space by resonances gives a natural discrete coarse grain procedure. First, resonances are nonintersecting and cover the whole phase space area, so that almost all positions in the phase space are inside some resonance state. Secondly, the stable and unstable manifolds of a minimizing periodic orbit typically cross each other transversely, so there is a stochastic layer inside any resonance in which the motion can be regarded as mixing, therefore rapidly losing memory of its history. Lastly, the overlaps of turnstiles of partial separatrices with other resonances give all the allowable transitions between resonances. All these suggest a Markov description of transport where we assume there is an immediate loss of memory within resonances.

Consider a random motion of a particle through a discrete set of states. A Markov chain is a stochastic process where the transition probability from one state to another at time  $t$  is independent of  $t$ , i.e., if  $\mathcal{T}(i,t \rightarrow j,t+1)$  is the probability of starting the particle in state  $i$  at time  $t$  and finding it in state  $j$  at time  $t+1$ , then

$$\mathcal{T}(i,t \rightarrow j,t+1) = \mathcal{T}(i,0 \rightarrow j,1) \equiv \mathcal{T}_{ij} \quad (3.5.1)$$

The transition probability  $\mathcal{T}_{ij}$  (or transition matrix) is nonnegative and satisfies

$$\sum_j \mathcal{T}_{ij} = 1 \quad (3.5.2)$$

Let  $P_i(t)$  be the probability in state  $i$  at time  $t$ , the evolution equation of the probability distribution is

$$P_i(t+1) = \sum_j P_j(t) \mathcal{T}_{ji} \quad (3.5.3)$$

This is called the master equation of the Markov chain.

In the simplest version of the Markov model, states are resonance islands labeled by  $(r,s)$ , where  $r$  denotes the resonance type  $(m_r, n_r)$ , in fact we can use the frequency of the resonance  $m_r/n_r$  as  $r$ ; and  $s = 0, 1, \dots, n_r-1$  denotes the island. By convention, we choose  $s = 0$  to be the central island in the resonance chain, and  $(r,t)$  to be the inverse iterate of  $(r,t+1 \bmod n_r)$  for  $t > 1$ .

The transition matrix elements are specified as follows:

- 1) For  $s \neq 0$ , each iteration maps  $(r,s)$  to  $(r,s+1)$ , the next island in the same resonance chain. Hence the transition matrix from  $(r,s)$  island to  $(r,s+1 \bmod n_r)$  island is 1 and to any other is 0.
- 2) Transitions between different resonances take place in the  $(r,0)$  island. Let the flux between resonance  $r$  and  $r'$  ( $r \neq r'$ ) be

$F(r,r')$ , the total flux out of resonance  $r$  be  $\mathcal{F}lux(r) (= \sum_{r'} F(r,r'))$ , and the stochastic area in the central island of the resonance  $r$  be  $A(r)$ , which is the area of the central island minus the area of regular motions in it, then

$$\begin{aligned} \mathcal{T}(r,0 \rightarrow r',1) &= F(r,r')/A(r) \\ \mathcal{T}(r,0 \rightarrow r,1) &= 1 - \mathcal{F}lux(r)/A(r) \end{aligned} \quad (3.5.4)$$

All other transition matrix elements are zero.

With the above specification, the transition matrix obviously satisfies

$$\sum_{r',s'} \mathcal{T}(r,s \rightarrow r',s') = 1. \quad (3.5.5)$$

We discuss a few consequences of this model.

a) The simplest prediction of the Markov model is that for escape from a single resonance the survival probability decays exponentially. We may restrict consideration to the escape from the central island of a period  $n$  resonance, since the other islands are iterates of each other. Hence

$$P(nt) = (1 - \mathcal{F}lux/A)^t$$

so the escape rate is

$$\alpha = (\log(1 - \mathcal{F}lux/A))/n \quad (3.5.6)$$

Numerical studies show exponential decay of escape rate, and in some cases, the escape rates compare favorably with the prediction. A model system is studied in detail in chapter 4.

b) Another consequence of the Markov model is that if the phase space is compact, then any initial probability distribution will approach to a uniform distribution equilibrium state

$$\lim_{t \rightarrow \infty} P_{r,s}(t) \rightarrow cA(r) \quad (3.5.7)$$

where  $c$  is the normalization constant.

A steady state  $P_{r,s}$  satisfies

$$P_{r,s} = \sum_{r',s'} P_{r',s'} \pi(r',s' \rightarrow r,s) \quad (3.5.8)$$

Using (5.3.7), and write  $\pi(r',s' \rightarrow r,s) = F(r',s';r,s)/A(r')$ , we get

$$A(r) = \sum_{r',s'} F(r',s';r,s) = \sum_{r',s'} F(r,s;r',s')$$

where in the last equation, we used the balance condition for the flux. This is exactly Eq(3.5.5).

c) Dana, Murray, and Percival [1989] have investigated the diffusion coefficient for periodic maps of the cylinder using the Markov model. Since the map is periodic in the vertical direction, we can use another integer index  $l$  to label the repeated cell structure in the vertical direction. They start with an arbitrary initial distribution, and compute the quantity

$$D = \lim_{t \rightarrow \infty} |2P(r,s,l;t)/2t| \quad (3.5.9)$$

which in some sense is a coarse grained diffusion coefficient. They show that  $D$  is independent of the initial probability distribution and is related to the largest eigenvalue  $\beta(k)$  of the transition matrix in the Fourier representation as follows

$$D = -d^2\beta(k)/dk^2|_{k=0} \quad (3.5.10)$$

This result shows some agreement with the actual diffusion coefficient. An interesting topic along this direction is to calculate the correlation function of the Markov model.

We can refine our Markov model to incorporate the effects of stable elliptic regions within a resonance. This can be achieved by subdivision of the resonance states into a hierarchical tree of states. In fact, we can subdivide a resonance by higher class resonances, this is possible as long as there are elliptic islands inside so that the map has twist around the elliptic orbit. However, for a hyperbolic system, we do not know any natural subdivision.



## **Chapter 4**

# **Symbolic Dynamics of the Sawtooth Map**

#### 4.1. Introduction

The Markov model discussed in the last chapter allows rapid computation of transport rates and is being used in applications [Meiss et. al. 1984, Davis 1985], but it depends on some statistical assumptions that need to be checked carefully for their accuracy.

In this chapter, we use the sawtooth map as a test case for the Markov transport model. It has several advantages:

1. Cantori, partial barriers, turnstiles and resonance boundaries can all be obtained analytically, as we show in this chapter.
2. These structures resemble those of smooth maps like the standard map (unless the perturbations are close to critical, so the sawtooth structures are useful first approximations for smooth systems under large perturbations. This is shown in Figure 4.1. They are also identical to those of certain piecewise linear maps in some parameter ranges.
3. For all non-zero values of the parameter, the sawtooth map is a completely chaotic K-system. However, this property is also a disadvantage, because the map cannot represent the effect near critical boundary circles, where the invariant tori have just been destroyed. The behavior near boundary circles is important for smooth systems like the standard map, and is believed to be the mechanism which gives long time tails in correlation functions in area preserving maps [Meiss and Ott, 1986].

As we shall see, a purely chaotic system has more structures than its name suggests. The skeleton of the chaos is a bounded invariant set. This invariant set has intriguing properties: its dynamics is characterized by the shift map on a restricted symbol sequence space.

The sawtooth map has been studied extensively by various authors [Percival 1980; Cary and Meiss 1981; Aubry 1983b; Percival and Vivaldi 1987a, 1987b; Bird and Vivaldi 1988]. In particular, the minimizing orbits were first obtained by Aubry in 1983.

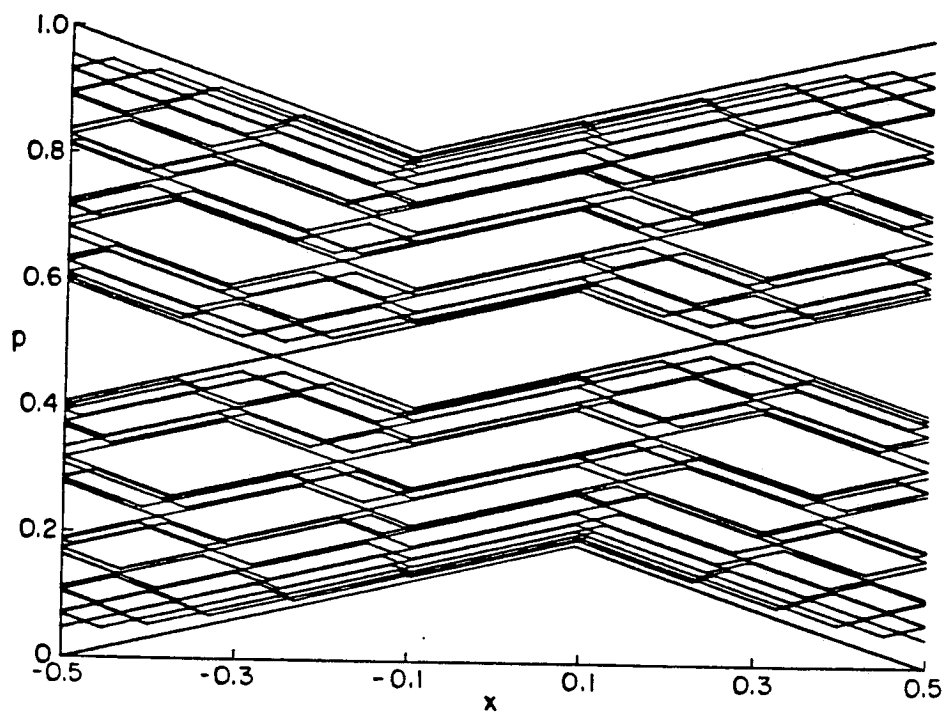


Figure 4.1a. Resonances of the sawtooth map for  $K = 0.1$  up to period 10.

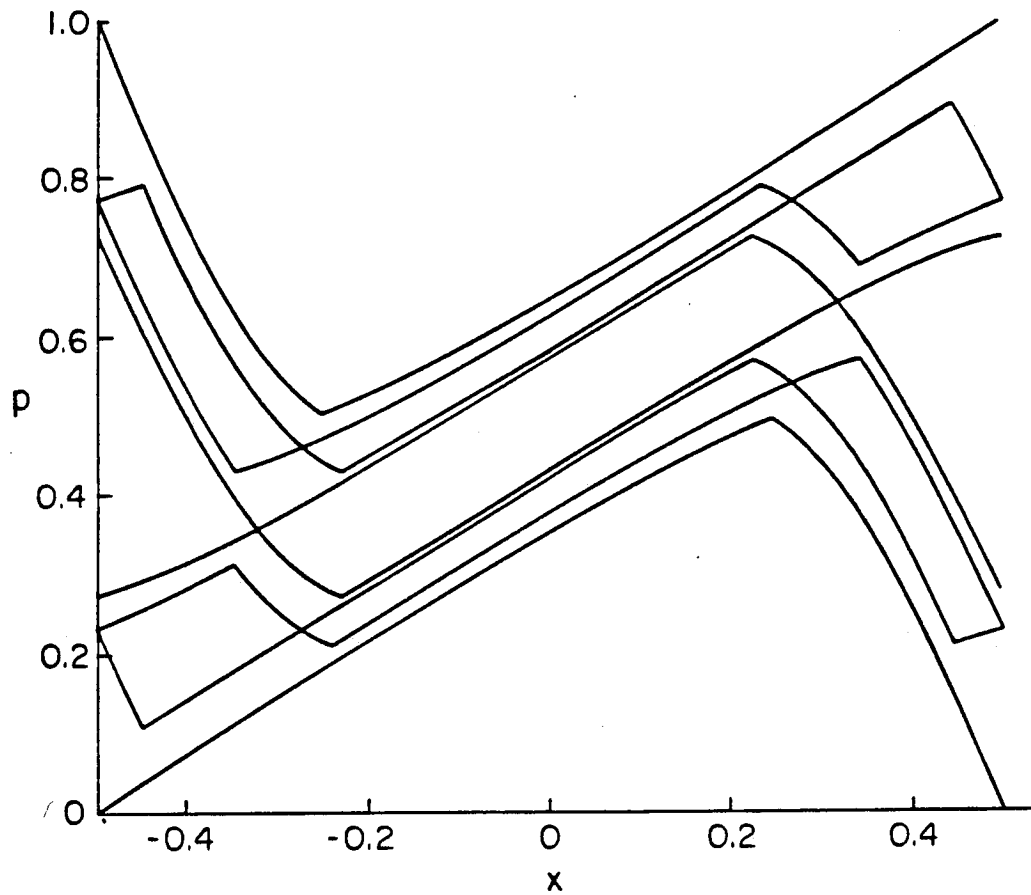


Figure 4.1b. Some resonances of the standard map at  $K = 1.972$ . Note that away from the principal gap, around  $x = 0$ , the resonances have a rectangular shape similar to the sawtooth map.

## 4.2. The Sawtooth Map and Its Orbital Structures

The sawtooth map is derived by linearizing the standard map about the latter's fixed points. It is defined by

$$\begin{aligned} p_{t+1} &= p_t + Kf(x_t) \\ x_{t+1} &= x_t + p_{t+1} \text{ mod } 1 \end{aligned} \quad (4.2.1)$$

where  $x$  is the configuration coordinate,  $x \in \mathcal{T}^1$ , and  $p$  the momentum coordinate,  $p \in \mathcal{R}$ ; we shall take  $\mathcal{T}^1$  as the unit interval  $[0,1)$ .  $t$  is the discrete time variable,  $K$  is the "stochasticity parameter", and the force function is

$$f(x) = x - 1/2 \text{ for } x \in (0,1) \quad (4.2.2)$$

The force is temporarily undefined at  $x = 0, 1$ . Note the force is discontinuous on the line  $x = 0$ . This discontinuity is the basic source of "nonlinearity" in the sawtooth map. Notice that the map is also periodic in  $p$ ; define a unit cell  $0 \leq p < 1, 0 \leq x < 1$ , the orbit structure is repeated in unit cells that are one unit of momentum apart. We shall consider the parameter range  $K > 0$  where the map is hyperbolic.

The lift of the sawtooth map to its universal covering space, the plane  $\mathcal{R}^2$  is

$$\begin{aligned} p_{t+1} &= p_t + KS(X_t) \\ X_{t+1} &= X_t + p_{t+1} \end{aligned} \quad (4.2.3)$$

or its Lagrangian or Newtonian form is

$$X_{t+1} - 2X_t + X_{t-1} = KS(X_t) \quad (4.2.4)$$

where  $X \in \mathcal{R}$  is the lifted angle coordinate.  $S(X)$  is the periodic sawtooth function defined by

$$\begin{aligned} S(X) &= X - 1/2 \text{ for } X \in (0,1) \\ S(X+k) &= S(X) \text{ for } k \in \mathbb{Z}. \end{aligned} \quad (4.2.5)$$

The sawtooth map has a reversor  $R$  (section 3.1.1)

$$R(x,p) = (-x+p,p) \quad (4.2.6)$$

In the following, we shall use the Newtonian form of the map; however the  $p$  coordinates are easily retrieved by using the original map (4.2.3).

Orbits of the lift map correspond to stationary configurations of a one dimensional Frenkel-Kontorova chain of particles connected by harmonic springs with action given by

$$W(\{X_t\}) = \frac{(X_{t+1} - X_t)^2}{2} - \frac{K}{2} (S(X_t))^2 \quad (4.2.7)$$

The potential is periodic, and within each well it is quadratic.

It is convenient to decompose  $X$  into a fractional part and an integral part

$$X_t = x_t + M_t \quad (4.2.8)$$

Then Newton's equation becomes

$$\begin{aligned} x_{t+1} - 2x_t + x_{t-1} - K(x_t - 1/2) &= -(M_{t+1} - 2M_t + M_{t-1}) \\ &\equiv -(c_t - c_{t-1}) \equiv -b_t \end{aligned} \quad (4.2.9)$$

where we have introduced the so called linear codes

$$c_t = M_{t+1} - M_t$$

$$b_t = c_t - c_{t-1} \quad (4.2.10)$$

Therefore if  $x_t$  crosses the discontinuity line  $k$  time(s) upon iteration, the  $c$  code is  $k$ . If the crossing is from left to right,  $k$  is positive; if the crossing is from right to left,  $k$  is negative. The integer  $M_t$  labels the potential well where the  $t$ (th) particle sits. The code  $c_t$  represents the shift in the well number between successive particles, hence gives a crude integer measure of the stretching in the string between these particles, and  $b_t$  represents the integer part of the shear in the stretching.

The equation of motion (4.2.8) is a second order inhomogeneous difference equation. To solve it, we need to invert the infinite tridiagonal matrix. This can be achieved by a Green function which satisfies the equation

$$G_{t+1,s} - 2G_{ts} + G_{t-1,s} - KG_{ts} = \delta_{ts} \quad (4.2.11)$$

Since  $G_{t+k,s+k}$  satisfies the same equation as  $G_{ts}$ , the solution only depends on the difference  $t-s$

$$G_{ts} = G_{t-s,0}$$

The unique bounded solution is given by

$$G_{ts} = -\lambda^{-|t-s|/\sqrt{D}} \quad (4.2.12)$$



where  $\lambda$  is the greater root of the homogeneous difference equation, which is also the greater eigenvalue of the linear map

$$\begin{aligned}\lambda &= 1 + (K + \sqrt{D})/2 \\ D &= K^2 + 4K\end{aligned}\tag{4.2.13}$$

therefore

$$\begin{aligned}x_t - 1/2 &= \sum_{s=-\infty}^{\infty} G_{ts} (-b_s) \\ x_t &= \sum_{s=-\infty}^{\infty} \frac{\lambda^{-|t-s|}}{\sqrt{D}} b_s + 1/2\end{aligned}\tag{4.2.14}$$

Hence, the angular coordinates of an orbit are uniquely characterized by the b-code, and vice versa. The relation between the angular configurations and the b-code is linear. The b-codes characterize equivalent classes of orbits on the torus, we call it the "linear code". In the same way, the c-codes characterize equivalent classes of orbits on the cylinder, and the M-codes characterize orbits on the plane.

Not all linear codes give physical orbits. There are an infinite number of constraints for the code corresponding to the requirement that

$$0 < x_t < 1$$

or

$$-1/2 < \sum_{s=-\infty}^{\infty} \frac{\lambda^{-|t-s|}}{\sqrt{D}} b_s < 1/2 \quad (4.2.15)$$

We shall refer orbits that do not satisfy the above constraint as "ghost orbits". Note that as  $K$  increases, the coefficients of the  $b$ 's decrease. Hence ghost orbits may become real orbits as  $K$  increases. This occurs when one or more points of a ghost orbit hits the discontinuity lines at  $x=0$  and  $x=1$ , and move inward from there.

### 4.3. Minimizing and Minimax Orbits

Symbol sequences for minimizing orbits are given by their monotone property. Although the potential here is not differentiable, Aubry-Mather theory (section 3.2) still applies to this model.

**Theorem** (Aubry 1983b) For a minimizing orbit with rotation frequency  $\nu$ , there exists a phase constant  $\alpha$  such that the configuration points  $X_t$  belong to the interval  $[M_t, M_{t+1}]$ , where  $M_t = \text{int}(\nu t + \alpha)$ .

Hence minimizing orbits are monotone. This follows immediately from the above theorem

$$X_t + j < X_{t'} + j' \Leftrightarrow X_{t+1} + j < X_{t'+1} + j' \quad (4.3.1)$$

#### 4.3.1 Minimizing Periodic Orbits

A periodic orbit of type  $(m,n)$  satisfies  $X_{t+n} = X_t + m$ . Note that the symbol sequence of a periodic orbit is also periodic. Summing up the series in Eq(4.2.12) gives

$$x_t = \sum_{s=0}^{n-1} \frac{(\lambda^{-s} + \lambda^{n-s})}{\sqrt{D} (1 - \lambda^{-n})} b_{s+t} + \frac{1}{2} \quad (4.3.2)$$

The symbol sequence for the  $(m,n)$  minimizing orbit is

$$M_t = \text{int}\left(\frac{1}{2} + \frac{mt}{n}\right) \quad (4.3.3)$$

A few simple relations for the symbol sequence of periodic minimizing configuration are useful. Letting  $\ell$  be the unique solution of the Diophantine equation  $m\ell = -1 \pmod{n}$ ,  $\ell < n$ . This is equivalent to letting  $j/\ell$  the left convergent of  $m/n$

$$m\ell - nj = -1, \ell < n \quad (4.3.4)$$

A simple calculation using Eq(4.3.4) gives

$$\begin{aligned} b_{t+\ell} &= b_t & 1 < t < n-1 \\ b_{1+\ell} &= b_1 - 1 \\ b_{n-1+\ell} &= b_{n-1} - 1 \\ b_\ell &= b_0 + 2 \end{aligned} \quad (4.3.5)$$

An important characterization of these orbits is the "gap" function, which we now define. In the following discussion, it is convenient to let  $x_0$  be the leftmost point of the orbit in the unit interval  $[0,1)$ . The neighboring point to the left of  $x_t$  is  $x_{t+\ell}$ . The gap function at time  $t$  for the minimizing periodic configuration  $\{ \xi_t \}$  is the distance between  $x_t$  and  $x_{t+\ell}$

$$\xi_t = x_t - x_{t+\ell} + \delta_{0,t} \quad (4.3.6)$$

For minimizing periodic orbit, the equation satisfied by  $\xi_t$  is

$$\begin{aligned} \xi_{t+1} &= (2+k) \xi_t + \xi_{t-1} \\ &= (-b_t + b_{t+\ell} + \delta_{0,t+1} + \delta_{0,t-1} - 2\delta_{0,t}) - k\delta_{0,t} \\ &= -k\delta_{0,t} \end{aligned} \quad (4.3.7)$$

hence for a period  $n$  orbit

$$\xi_t^{(n)} = K \frac{(\lambda^{-s} + \lambda^{n-s})}{\sqrt{D} (1 - \lambda^{-n})} \quad (4.3.8)$$

where we use the superscript  $n$  to emphasize that the gap function depends only on the period of the orbit, and not its frequency. Eq(4.3.8) was first obtained by Percival and Vivaldi [1987b] by a continuity argument.

An immediate consequence of Eq(4.3.8) is that the leftmost point of the minimizing orbit is given by

$$x_0^{(n)} = \frac{\xi_0^{(n)}}{2} = \frac{K(1 + \lambda^n)}{2\sqrt{D}(1 - \lambda^{-n})} \quad (4.3.9)$$

### 4.3.2 Cantori and Minimizing Homoclinic Orbits

Let  $\sigma \in \{+,-\}$ , and define  $\text{int}^\sigma(x)$  and  $\{x\}^\sigma$  to be the right and left continuous integral and fractional parts of  $x$ . The M-code which gives the orbit corresponding to the gap end points of a cantorus with frequency  $\nu$  is

$$M^\sigma_t = \text{int}^\sigma(\nu t) \quad (4.3.10)$$

The orbit of right or left gap end points (for  $\sigma = +$  or  $-$  respectively) is

$$x^\sigma_t = \sum_{n=-\infty}^{\infty} \frac{\lambda^{-|n|}}{\sqrt{D}} b^\sigma_{t+n} + 1/2$$

$$x^\sigma_t = \sum_{n=-\infty}^{\infty} \frac{\lambda^{-|n|}}{\sqrt{D}} [\{v(t+1+n)\}^\sigma - 2\{v(t+n)\}^\sigma + \{v(t-1+n)\}^\sigma] + 1/2$$

$$= -\frac{(\lambda^{-1} - \lambda^{-1})}{\sqrt{D}} \{vt\}^\sigma + (\lambda + \lambda^{-1} - 2) \sum_{n=-\infty}^{\infty} \frac{\lambda^{-|n|}}{\sqrt{D}} \{v(t+n)\}^\sigma + 1/2$$

$$= \{vt\}^\sigma - K \sum_{n=-\infty}^{\infty} \frac{\lambda^{-|n|}}{\sqrt{D}} \{v(t+n)\}^\sigma + 1/2$$

Introducing the sawtooth function

$$s^\sigma(\theta) = \{\theta + 1/2\}^\sigma - 1/2$$

and replacing  $\nu t - 1/2$  by  $\theta$ , we obtain a continuum of orbits  $\{x^\sigma(\theta + \nu n) \mid n \in \mathbb{Z}\}$

$$x^\sigma(\theta) = \{\theta + 1/2\}^\sigma - K \sum_{n=-\infty}^{\infty} \frac{\lambda^{-|n|}}{\sqrt{D}} s^\sigma(\theta + \nu n) \quad (4.3.11)$$

As we shall show,  $0 < x^\sigma(\theta) < 1$ , Hence

$$X^\sigma(\theta) = \theta + 1/2 - K \sum_{n=-\infty}^{\infty} \frac{\lambda^{-|n|}}{\sqrt{D}} s^\sigma(\theta + \nu n) \quad (4.3.12)$$

We note the following properties of  $X^\sigma(\theta)$

a)  $X^\sigma(\theta)$  is  $\sigma$ -continuous. This follows immediately since  $\{\theta\}^\sigma$  is  $\sigma$ -continuous.

b) The derivative of  $X^\sigma(\theta)$  vanishes at all points of continuity. Since the series for  $X^\sigma(\theta)$  is uniformly convergent, its derivative can be computed term by term. Evaluating the derivative at a point of continuity

$$\partial X^\sigma(\theta) / \partial \theta = 1 - K \sum_{n=-\infty}^{\infty} \frac{\lambda^{-|n|}}{\sqrt{D}} = 0$$

c)  $X^\sigma(\theta)$  is a monotone function of  $\theta$ . The jump across a point  $\theta_{mn} = m - \nu n - 1/2$  is

$$\Delta X_n = K \lambda^{-|n|} / \sqrt{D}$$

This in fact is the gap function of the cantor. The total variation of  $X^\sigma(\theta)$  from  $-1/2$  to  $1/2$  is the sum of all jumps. Hence  $X^\sigma(\theta)$  is a sum of step functions

$$X^\sigma(\theta) = \sum \Delta X_n H^\sigma(\theta - \theta_{mn}) + \text{const.}$$

where  $H^\sigma(\theta - \theta_{mn})$  is the left or right continuous step function.

d) A bound of the norm of  $x^\sigma(\theta)$  can be obtained by noticing  $x^\sigma(0) = 1/2$  and  $x^\sigma(\theta)$  is odd almost everywhere. The value of  $x^\sigma(\theta)$  is the sum of the jumps corresponding to discontinuities  $\theta_{mn} \in [0, \theta]$ . Providing  $|\theta| < 1/2$ , the term  $\Delta X_0$  does not contribute, hence

$$|x^\sigma(\theta) - 1/2| < 1/2 (\sum \Delta X_n) < 1/2$$

Therefore

$$0 < x^\sigma(\theta) < 1$$

The invariant set indeed is inside the fundamental unit interval.

The closure of the set  $\mathcal{M}_v = \{ x^\sigma(\theta), x^\sigma(\theta+v) \mid \theta \in \mathcal{R}, \sigma \in \{+, -\} \}$  is a Cantor set. Since  $x^\sigma(\theta)$  is  $\sigma$ -continuous,  $\mathcal{M}_v$  has no isolated points, hence is perfect and closed. Now between any two distinct points  $x_1, x_2$  of  $\mathcal{M}_v$  corresponding to parameter values  $\theta_1, \theta_2$  of  $x^\sigma(\theta)$ , there is a jump at  $\theta_{mn} = m - nv - 1/2 \in (\theta_1, \theta_2)$ . Hence  $\mathcal{M}_v$  is totally disconnected. Therefore  $\mathcal{M}_v$  is a Cantor set.

The gap function for a cantor is defined as the gaps between the end points of the Cantor set. It is easily obtained from formulas for the right and left end points or by taking the limit  $n$  to infinity for the periodic gap function



$$\xi_t = K\lambda^{-|t|/\sqrt{D}} \quad (4.3.13)$$

Note the cantorus gap function for the sawtooth map does not depend on the frequency. This also applies to the gap function for the minimizing homoclinic orbits, the latter can be viewed as a limiting cantorus when the frequency tends to a rational number.

Equation (4.3.13) has interesting implications, e.g. the fractal dimension of any cantorus for the sawtooth map is zero. In fact, for a hyperbolic cantorus, it can be shown that the projected measure is zero and the gaps go to zero exponentially, and there are only finitely many orbits of gaps. Therefore the projected measure after removing the  $n$  largest gaps is at most  $C\alpha^{-n}$ , for some  $C$ . So one can cover  $\mathcal{M}_v$  with  $n$  intervals  $I_j$  of length  $C\alpha^{-n}$ . Thus  $K_n(s) = \sum |I_j|^s \leq nC\alpha^{-n}$ . The Hausdorff dimension  $D_H(\mathcal{M}_v) = \inf \{ s \geq 0: K_n(s) \rightarrow 0 \text{ as } n \rightarrow \infty \} = 0$ . This fact was first discovered for the golden mean cantorus in the standard map [Li and Bak, 1986], and later was proved for any hyperbolic cantorus [MacKay 1987].

When the frequency  $v$  is a rational number and is equal to  $m/n$ , equation (4.3.10) gives the symbol sequence for the upper or lower  $(m,n)$  homoclinic orbits. Eq(4.3.13) also gives the gap function for the upper or lower  $(m,n)$  homoclinic orbit.

### 4.3.3. Minimax Periodic Orbits

The ordered periodic orbit given by (4.4.2) and (4.4.3) is the ordinary hyperbolic one, which minimizes the action sum. For maps with sufficiently smooth functions  $f(x)$ , the Poincare-Birkhoff theorem guarantees the existence of a second ordered periodic orbit with the same winding number. This orbit is elliptic or hyperbolic with reflection, and minimaximizes the action sum. For the sawtooth map,  $f(x)$  is discontinuous, and the existence of such an orbit is not self evident. We show here that a second ordered periodic orbit indeed exists if  $f(x)$  is defined on the discontinuity line as follows:

$$f(x=0) = 0 \tag{4.3.14}$$

This definition is a natural one, since it leads to an antisymmetric function  $f(x) = -f(-x)$ . The map is then invariant under inversion  $(x,p) \rightarrow (-x,-p)$ . The new orbit has a point on the discontinuity line  $x=0$ , so that its linear stability is undefined (the tangent space does not exist). However, the orbit is obviously unstable, since the map is hyperbolic everywhere else.

We shall now show, by explicit construction, that there are ordered minimax periodic orbits for the sawtooth map. First we note that any minimax periodic configuration, if it exists, must have at least one point at the maximum of the potential,  $x=0$ . If this were not true then the action on the orbit would be differentiable, with a second variation given by

$$\delta^2 W(\{X_t\}) = \sum_{s=0}^{n-1} \sum_{t=0}^{n-1} H_{s,t} \delta X_s \delta X_t$$

$$H_{s,t} = -\delta_{s,t-1} - \delta_{s,t+1} + (2+K)\delta_{s,t} \quad (4.3.15)$$

where the indices  $s, t$  are assumed to be cyclic modulo  $n$ , the period of the orbit. Thus  $\delta^2 W > 0$ , for  $K > 0$ , i.e. this configuration is always a local minimum of the action. On the other hand, a minimax  $(m, n)$  configuration can have at most one point at  $x = 0$  since each point of an ordered  $(m, n)$  configuration lies between the two points on the corresponding  $(m, n)$  minimizing configuration.

Now we claim that the minimax orbit  $\{x'_t\}$  of rotation number  $m/n$  has the same symbol sequence as the minimizing  $(m, n)$  orbit  $\{x_t\}$  and is given by

$$\begin{aligned} x'_0 &= (x_0 + x_{\ell-1})/2 = 0 \\ x'_t &= (x_t + x_{t+\ell})/2 \quad t \neq 0 \pmod{n} \end{aligned} \quad (4.3.16)$$

The geometrical meaning of (4.3.16) is simple: the points on the minimax periodic orbit are exactly halfway between the neighboring points of the corresponding minimizing periodic orbit. For  $1 < t < n-1$ , Eq(4.2.9) becomes

$$\begin{aligned} x'_{t+1} - 2x'_t + x'_{t-1} - K(x'_t - 1/2) &= -(b_t + b_{t+\ell})/2 \\ &= -b_t \end{aligned}$$

We need also to verify Eq(4.2.9) for  $t = 1$  and  $t = n-1$ . It suffices to consider the case  $t = 1$

$$x'_2 - 2x'_1 + x'_0 - K(x'_1 - 1/2) = -(b_1 + b_{1+\ell} + 1)/2$$

$$= -b_1$$

The equation for the minimax orbit at  $t = 0$  is automatically satisfied

$$x'_{-1} + x'_1 = 1 = -b_0$$

Finally it is not difficult to see that the action, Eq(4.3.15), has precisely one downward direction, corresponding to the variation of the coordinate  $X_0$ .

We can also start directly from the gap function to find the minimax orbit

$$\xi'_t = \xi'_t - \xi'_{t+l} + \delta_{0,t}$$

Assuming that  $\{ x'_t \}$  has a unique point at  $x=0$  and has the same linear code as the minimizing periodic orbit, we have

$$x'_{t+1} - 2x'_t + x'_{t-1} - K(x'_{t-1}/2) = -b_t + K\delta_{0,t}/2$$

The term  $K\delta_{0,t}/2$  comes from the correction at  $t=0$  where the force is zero. The equation for  $\xi'_t$  is

$$\xi'_{t+1} - 2\xi'_t + \xi'_{t-1} - K\xi'_t = -K\delta_{0,t} + K(\delta_{0,t} - \delta_{0,t+l})/2$$

where the second term is due to the definition that the force term is zero at  $x = 0$ . therefore

$$\xi'_{t+1} - 2\xi'_t + \xi'_{t-1} - K\xi'_t = -K(\delta_{0,t} + \delta_{0,t+l})/2$$

or

$$\xi'_t = (\xi_t + \xi_{t+l})/2 \quad (4.3.17)$$

which again says the minimax orbit sits exactly halfway between the minimizing orbit. In fact, this also applies to the minimax homoclinic orbits and orbits homoclinic to cantori.

## 4.4. Partial Barriers, Turnstiles and Resonances

### 4.4.1. Partial Barriers and Turnstiles

The flux across the  $(m,n)$  minimizing orbits is obtained by constructing the turnstile of the orbit. We connect the neighboring points of the  $(m,n)$  minimizing orbit by straight line segments, then iterate these line segments one step backwards. The construction is shown in Figure 4.2. Note that these line segments pass through the  $(m,n)$  minimax orbit. Due to the discontinuity of the sawtooth map on its dominant symmetry line, one line segment is broken to two pieces. The region bounded by the dominant symmetry line, the broken pieces and the original line segments is the flux region, or turnstile. It is shown as the two shaded triangles in Figure 4.2. Hence the area of either one of the two triangles gives us the flux  $\Delta W$  across the  $(m,n)$  minimizing orbit

$$\Delta W(m,n) = \frac{1}{4} \Delta p x_0$$

where  $\Delta p$  is the discontinuity in  $p$  at  $x = 0$ . By Eq(4.2.1)

$$\Delta p' = \Delta p - K$$

Since the discontinuity maps to a continuous segment after one iteration,  $\Delta p' = 0$ , so

$$\Delta p = K$$

thus using Eq(4.3.9) we obtain

$$\Delta W(m,n) = \frac{K^2 (1 + \lambda^n)}{8\sqrt{D} (1 - \lambda^{-n})} \quad (4.4.1)$$

The flux across a cantor or a homoclinic orbit is given by letting  $n$  tend to infinity

$$\Delta W_c = K^2/8\sqrt{D} \quad (4.4.2)$$

One can also find the flux by taking the action difference of the corresponding minimax and minimizing orbits

$$\begin{aligned} \Delta W(m,n) &= W(\{X'_t\}) - W(\{X_t\}) \\ &= \frac{1}{8} \left\{ 2K\xi_0 - \sum_{t=0}^{n-1} [ (\xi_{t+1} - \xi_t)^2 + k \xi_t^2 ] \right\} \end{aligned}$$

which, by Eq(4.3.8), yields Eq(4.4.1) again.

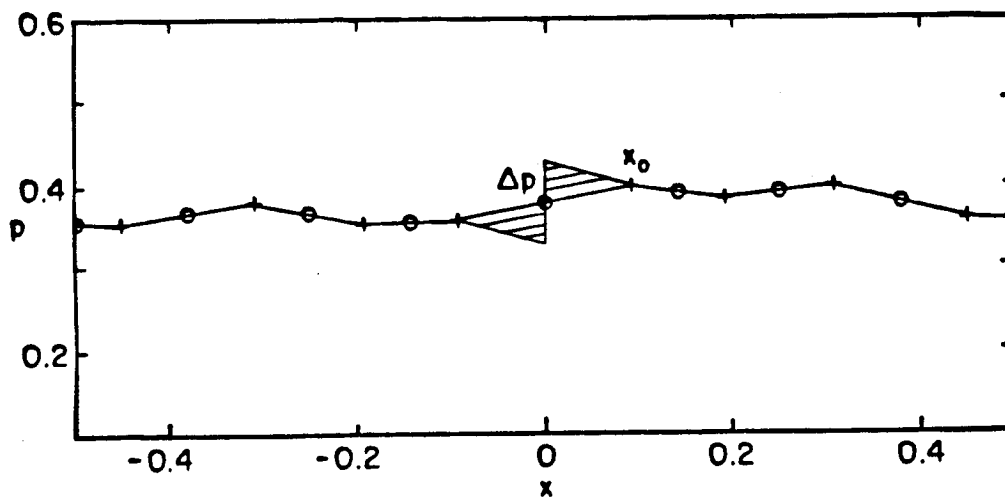


Figure 4.2. Flux across the (3,8) periodic orbit at  $K = 0.1$ . The crosses are the points on the minimizing orbit. The dots are on the minimax orbit. The shaded region is the turnstile.

#### 4.4.2. Resonances

An  $(m,n)$  resonance is the region bounded by the upper and lower partial separatrices for a minimizing  $(m,n)$  orbit. For the sawtooth map, the stable and the unstable manifolds of any periodic orbit are straight line segments: the slope does not depend on the period of the orbit. Thus the resonances of the sawtooth map are constructed simply by drawing the unstable (stable) manifolds of the points on the periodic orbit till they intersect the stable (unstable) manifolds of the neighboring points; these define the upper and lower partial separatrices of the resonances. Some of the resonances of the sawtooth map are shown in Figure 4.1a. Note they are all parallelograms.

It is straight forward to calculate the area of an  $(m,n)$  resonance, This is illustrated in Figure 4.3 for the  $(1,2)$  resonance. For an  $(m,n)$  minimizing orbit, the leftmost point is given by Eq(4.3.9), that of the homoclinic orbit is obtained by taking  $n$  to infinity; thus  $h_1 = K/\sqrt{D}$  and  $h_2 = x^{(n)}_0 - x^{(\infty)}_0 = K/\sqrt{D} (\lambda^n - 1)$ . Simple geometry gives

$$h = h_2(\tan\alpha + \tan\beta)$$

where  $\tan\alpha$  and  $-\tan\beta$  are the slopes of the unstable and the stable manifolds

$$\tan\alpha = K/(\lambda - 1) \quad \tan\beta = K/(1 - \lambda^{-1})$$

$$h(m,n) = h_2(\tan\alpha + \tan\beta) = k/(\lambda^n - 1) \quad (4.4.3)$$

$$A(m,n) = nh(h_1 + h_2) = \frac{nK^2}{\sqrt{D} (\lambda^n + \lambda^{-n} - 2)} \quad (4.4.4)$$



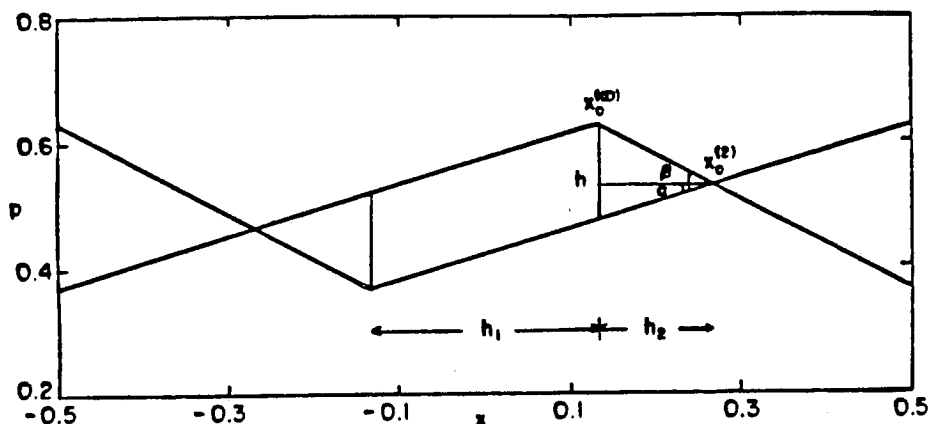


Figure 4.3. Illustration of the calculation of resonance area for the (1,2) resonance.

One can verify that formulas (4.4.3), (4.4.4) also apply to (0,1) and (1,1) resonances if we take them as the same resonance.

Note the area is inversely proportional to the absolute value of the residue. This seems to hold generically. The numerical computations of MacKay, Meiss and Percival [1987] show that for smooth maps the resonance area is, within numerical errors, inversely proportional to the absolute value of the residue for large residue minimizing orbits.

There are two complete staircases in the sawtooth map, one is given by the height function (4.4.3), the other by the area function (4.4.4) (Figure 4.4). Since these two quantities depend only on  $n$ , we denote them by  $h(n)$  and  $A(n)$ . It is not difficult to

prove these two devil's staircases are complete. For height function, we sum over all the steps

$$\sum_{n=1}^{\infty} \phi(n) h(n) = 1$$

Where  $\phi(n)$  is the Euler  $\phi$ -function [Hardy and Wright 1979], i.e. the number of positive integers not greater than and relatively prime to  $n$ . The above sum is a standard result

$$k \sum_{n=1}^{\infty} \phi(n) / (\lambda^n - 1) = K / (\lambda + \lambda^{-1} - 2) = 1 \quad (4.4.5)$$

Introducing a new variable  $\theta$  defined by  $\lambda = e^\theta$ , we can rewrite Eq(4.4.5) as

$$\sum_{n=1}^{\infty} \phi(n) / (e^{n\theta} - 1) = [4 \sinh^2(\theta/2)]^{-1}$$

Taking the derivative with respect to  $\theta$ , we get

$$\sum_{n=1}^{\infty} \phi(n) n / \sinh^2(n\theta/2) = \cosh(\theta/2) / \sinh^3(\theta/2)$$

which is exactly the completeness condition for the area staircase

$$\sum_{n=1}^{\infty} \phi(n)A(n) = 1 \quad (4.4.6)$$

The complement set of the resonances cuts a vertical line in a cantor set. From Eq(4.4.5), we see that this set again has zero fractal dimension. We expect this to hold generically for supercritical area-preserving maps.

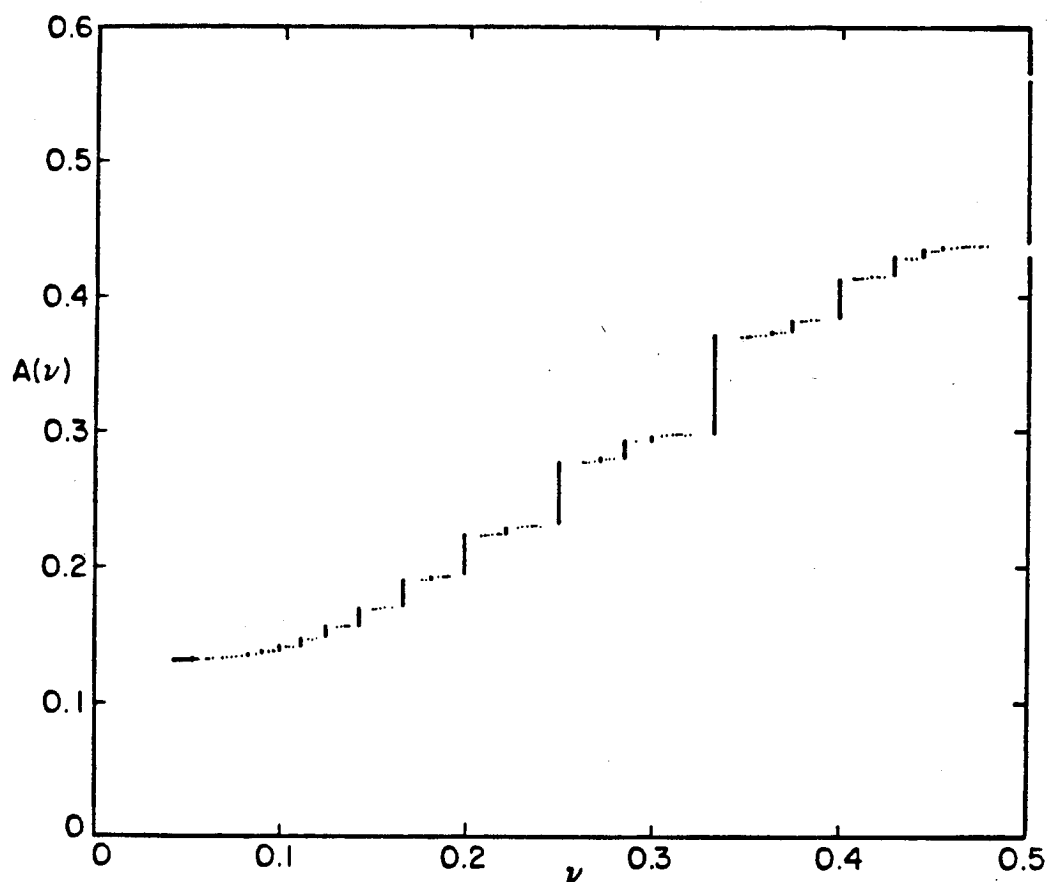


Figure 4.4. The area devil's staircase function in the sawtooth map at  $K = 0.3$ . Only half the staircase is plotted. Another half can be obtained by the reflection symmetry:  $A(1-v) = 1 - A(v)$

#### 4.4.3. Generalized Turnstiles

The geometrical form of the total flux region defined as the union of the upper and lower turnstiles undergoes qualitative changes which distinguish dynamics inside a resonances. There are basically three distinct cases,  $1 < \lambda^n < 2$ ,  $2 < \lambda^n < 3$ , and  $\lambda^n > 3$ . For the last case, the shape of the escape region changes, and we call the new structures "generalized turnstiles".

Let us restrict to the central island of the  $(m,n)$  resonance (the one surrounding  $x = 0$ ). For small  $\lambda^n$ , motions near the center of the island are quite "regular": Consider a point on the left side of the discontinuity, near the center of the island. The orbit is determined by the stable and unstable manifolds of the left unstable periodic point. So it moves on a hyperbola till it crosses the discontinuity line from left to right; Then it moves on a right hyperbola till it crosses the discontinuity line from right to left; and so on. Therefore, the orbit rotates around the center of the island, in a similar fashion to a stable elliptic island for the smooth map case. However, this analogy is not complete: almost all points in the central island will eventually rotate out of the island and escape since the map is hyperbolic. For small  $\lambda^n$ , the escape time can be very long, and the island structure is prominently seen. Figure 4.5 Shows such a behavior for the  $(1,2)$  resonance.

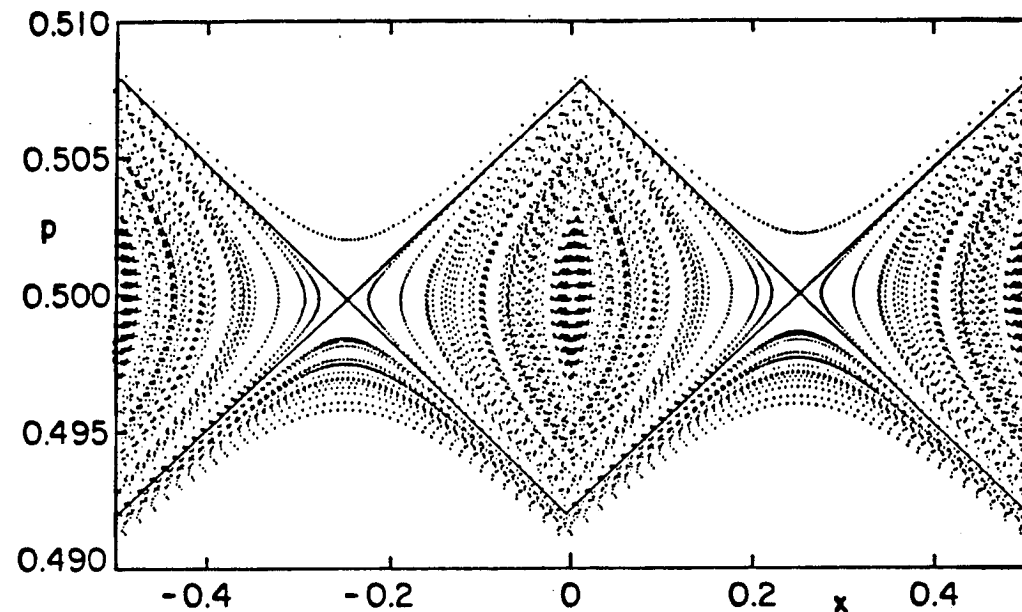


Figure 4.5. An orbit wandering around the (1,2) resonance at  $K = 0.001$  or  $\lambda^2 = 1.056$ .

This island structure is "modified" if the upper and lower turnstiles touch each other. This occurs when

$$h(m,n) = 2\Delta p = K$$

or

$$\lambda^n = 2 \tag{4.4.7}$$

For  $\lambda^n > 2$ , the escape region (the union of the turnstiles) is a connected component, and it cuts the resonance into two disjoint pieces. This has interesting implications for the dynamics inside the resonance. In particular, the trapped invariant set becomes a Cantor set.

The form of the escape region undergoes a further qualitative change when the upper turnstile cuts across the lower boundary of the resonance and the lower turnstile cuts across the upper boundary (see Figure 4.6.). This occurs when

$$h(m,n) = \Delta p = K/2$$

or

$$\lambda^n = 3 \tag{4.4.8}$$

For  $\lambda^n > 3$ , only parts of the upper and lower turnstiles remain inside the resonance (the trapezoidal regions CEMJ and DFNI in Figure 4.6). The regions outside the resonance can no longer be regarded as the flux regions. We call these trapezoidal regions the upper and lower "generalized turnstiles".

In this case, the total flux is no longer  $2\Delta W$  given in section 4.4.1, but the sum of the areas of the shaded parallelogram regions

$$\text{Flux} = \frac{K^2}{4\sqrt{D}} \left[ 1 - \frac{(\lambda^n - 3)^2}{(\lambda^n - 1)^2} \right] \tag{4.4.9}$$

It is easy to verify that at  $\lambda^n = 3$ , the total flux is one third of the area of an island.

According to the Markov model, the probability of escaping a period  $n$  resonance per iteration is

$$P_e(\lambda^n) = \frac{\text{Flux}(m,n)}{\text{Area}(m,n)/n}$$

Hence

$$\begin{aligned} P_e(\lambda^n) &= (\lambda^n + \lambda^{-n} - 2)/4 & \lambda^n < 3 \\ &= 1 - 2\lambda^{-n} & \lambda^n \geq 3 \end{aligned} \tag{4.4.10}$$

It is easy to verify that  $P_e$  increases monotonically from 0 to 1 as  $\lambda^n$  varies from 1 to  $\infty$ .

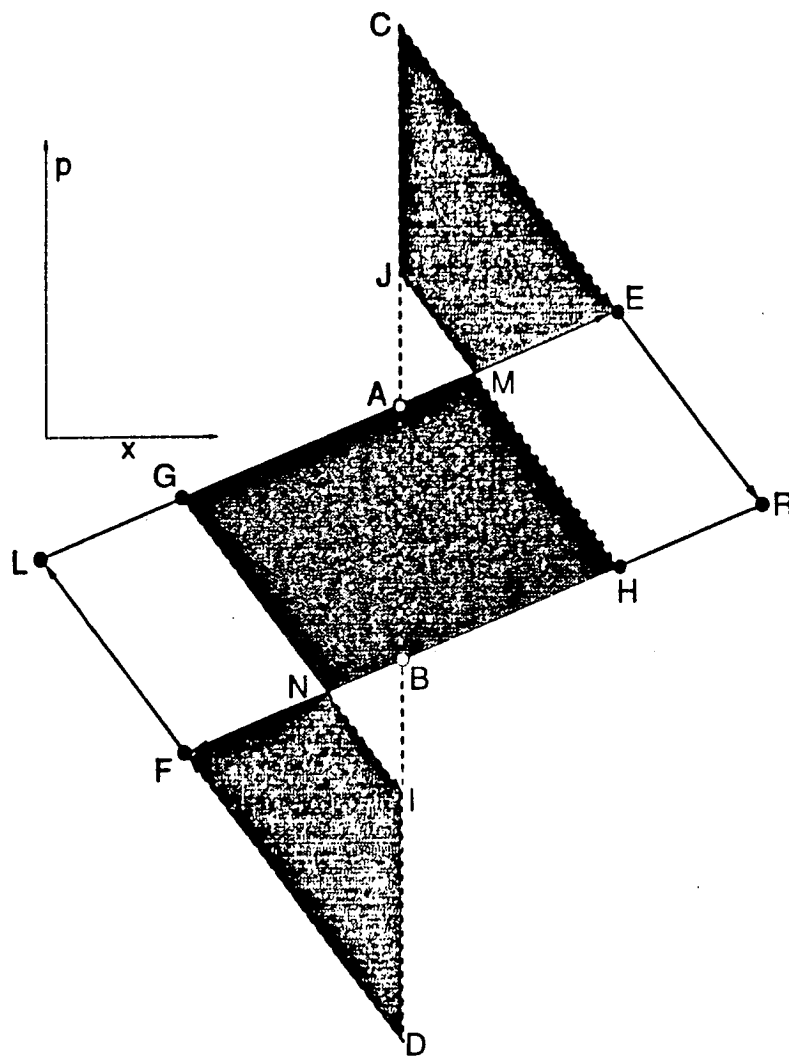


Figure 4.6. Schematic illustration of the central island of a general resonance, in the case  $\lambda > 3$ , with its generalized turnstiles (shaded regions).



#### 4.4.4. Application to Piecewise Linear Standard map

We remark here that the sawtooth map is intimately related to various piecewise linear, continuous area-preserving maps studied by Bullet [1986]. Define a new force function in (4.2.2) by

$$f(x) = \begin{cases} K(\varepsilon-1)x/\varepsilon & 0 < x \leq \varepsilon/2 \\ K(x-1/2) & \varepsilon/2 < x \leq 1 - \varepsilon/2 \\ K(\varepsilon-1)(x-1)/\varepsilon & 1 - \varepsilon/2 < x \leq 1 \end{cases} \quad (4.4.11)$$

where  $0 < \varepsilon < 1$ . If the size of the main gap  $\xi^{(\infty)}_0$  is larger than  $\varepsilon$ , then all the minimizing and minimax configurations for Eq(4.4.11) are identical to those for the sawtooth map with the same parameter value. This is because all points of a minimizing orbit fall inside the hyperbolic region  $\{ x: |x-1/2| < (1-\varepsilon)/2 \}$ , where the maps are identical. Furthermore a minimax orbit will have only one point outside this region (at  $x=0$ ), but both maps have  $F(0) = 0$ . Thus the minimizing or minimax orbits of the sawtooth map are also the minimizing or minimax periodic orbits of the piecewise linear map. The critical parameter value for which  $\xi^{(\infty)}_0 = \varepsilon$

$$K_{Cr} = 4\varepsilon^2 / (1 - \varepsilon^2) \quad (4.4.12)$$

As shown by Bullet, above this critical value there are no invariant tori. For  $K = K_{Cr}$  invariant tori of all frequencies can be constructed by connecting the neighboring points of the corresponding minimizing homoclinic orbits or cantori by straight line segments.

The resonance structure of these piecewise linear maps at or above this critical value is therefore identical to that of the sawtooth map. For instance the resonances of the piecewise linear standard map ( $\varepsilon = 1/2$ ) studied by Bullett at  $K = 4/3$  correspond to those of the sawtooth map at  $K = 4/3$  (where  $\xi^{(\infty)}_0 = 1/2$ ), except in the region of the principal gap. The turnstile of an  $(m,n)$  periodic orbit above the critical parameter value is constructed in Figure 4.7, and it is easy to see the flux is given by

$$\Delta W(m,n) = K(\xi^{(n)}_0 - \varepsilon)/8 \quad (4.4.13)$$

Similarly the flux through a cantorus is

$$\Delta W_C = K(\xi^{(\infty)}_0 - \varepsilon)/8 \quad (4.4.14)$$

Near the critical parameter value, Eq(4.4.14) implies that the flux through any cantorus grows as

$$\Delta W_C = \varepsilon(1-\varepsilon^2)(K-K_{Cr})/16 + O(K-K_{Cr})^2 \quad (4.4.15)$$

That  $\Delta W$  grows linearly with  $K-K_{Cr}$  is unusual since for a smooth map renormalization theory predicts an exponent of 3.01 for any noble cantorus [MacKay, Meiss and Percival 1984].

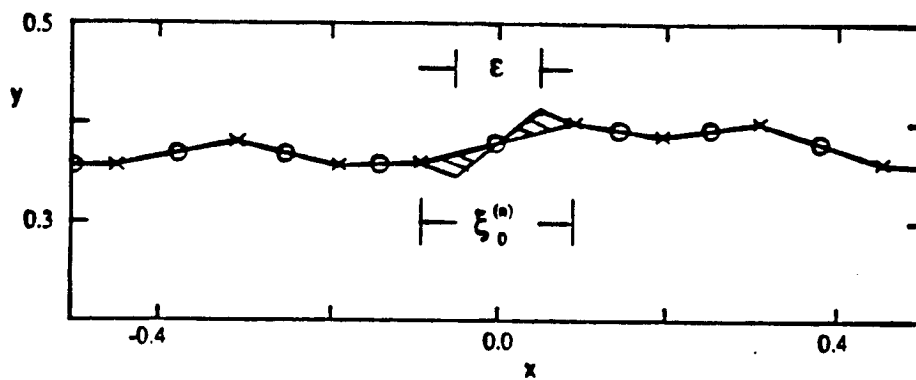


Figure 4.7. Flux across the (3,8) periodic orbit for the piecewise linear map for  $\varepsilon = 0.1$  at  $K = 0.1$ . The shaded region is the turnstile.

#### 4.5. Invariant Sets within Resonances

Orbits which are trapped within a resonance give rise to the islands around islands structure for smooth systems [Meiss 1987, Meiss and Ott 1986]. This structure impedes the escape of orbits from a resonance because outside the last vibrational invariant circle around the minimax periodic point there are vibrational cantori. Each cantorus has a turnstile which impedes the transport of orbits. This structure is repeated hierarchically for each stable island within the resonance. It is natural to ask if such a structure exists in the purely hyperbolic sawtooth map and how it affects the escape dynamics.

In the smooth case some of the trapped orbits can be characterized by their rotation frequency about the central elliptic point, or their rotation frequency about another trapped orbit. For the sawtooth map, it is more convenient to characterize the set of trapped orbits in terms of an extension of the linear code. This gives a complete characterization of the trapped set.

An orbit is uniquely determined by the linear code  $\{ b_t \}$ . For a trapped orbit, however, it is convenient to use a modified code which more naturally describes its properties. The new code is constructed from two characteristics: the resonance within which the periodic orbit is trapped,  $(m,n)$ , and an  $\mathcal{L}$ - $\mathcal{R}$  symbol sequence  $\{ a_j \}$  of the orbit. To construct the latter, consider the central island of the  $(m,n)$  resonance. The discontinuity line divides it into two halves, call them  $\mathcal{L}$  and  $\mathcal{R}$  for left and right, respectively (Figure 4.8). A periodic orbit within the  $(m,n)$  resonance has a period which is multiple of  $n$ , say  $qn$ ; it has  $q$  points in each island. To construct the code, pick a starting

point  $x_0$  in the central island and let  $a_0=0$  if  $x_0$  is in  $\mathcal{L}$ , otherwise  $a_0=1$ ; continue for the remaining  $q$  points, letting  $a_j=0$  if  $x_{jn}$  is in  $\mathcal{L}$  and so on. Therefore  $\{a_j\}$  has period  $q$ . For example, the vibrating ordered orbit with rotation frequency  $3/8$  (rotation number relative to  $T^n$ ) has a  $\mathcal{L}$ - $\mathcal{R}$  symbol sequence  $[1,1,0,1,0,0,1,0]$ , where we use the square brackets to denote the periodicity. There are two vibrating ordered orbits for odd period, one has an extra point on the left side of the discontinuity, and the other has one more point on the right. They are reflections of each other under the reversor  $R$  given by (4.2.6). For example,  $\mathcal{L}$ - $\mathcal{R}$  symbol sequence for the  $2/7$  orbits are:  $[1100110]$  and  $[1100100]$ . This symbol sequence is natural and corresponds to the usual symbol sequence for a hyperbolic horseshoe. In fact, this coding scheme, i.e., resonance type plus  $\mathcal{L}$ - $\mathcal{R}$  symbol, gives a general code for almost all orbits of the sawtooth map, since the resonances fill the full measure of phase space.

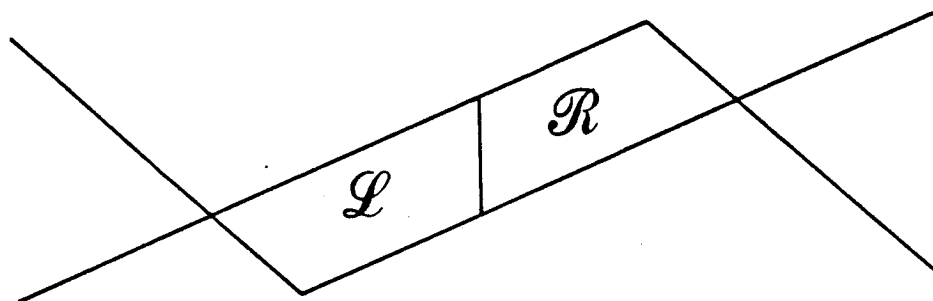


Figure 4.8. Illustration of the  $\mathcal{L}$ - $\mathcal{R}$  coding for the  $(1,2)$  resonance.

The linear code  $\{ b_i \}$  can be easily obtained from the above two characteristics. Suppose the c-code for the  $(m,n)$  orbit is  $[c_0, c_1, c_2, \dots, c_{n-1}]$ , where time 0 corresponds to the leftmost point of the  $(m,n)$  orbit. The c-code for the island orbits is then given by  $[1-a_0, c_1, c_2, \dots, c_{n-2}, a_1, 1-a_1, c_1, c_2, \dots, c_{n-2}, a_2, \dots, 1-a_{k-1}, c_1, \dots, a_0]$ . There is no change in the code for points other than those in the central island. Since this orbit follows the rotation order of the  $(m,n)$  orbit, if  $a_j=0$ , the crossing is deferred to the next step, while when  $a_j=1$ , the crossing occurs at the same step. This explains the occurrence of the pair  $(a_j, 1-a_j)$  at times  $(i_{n-1}, i_n)$ . Using Eq(4.2.10), the linear b-code for times  $(i_{n-1}, i_n, i_{n+1})$  is given by  $(1, -1, 0)$  when  $a_j = 1$ , and is  $(0, 1, -1)$  when  $a_j = 0$ .

The above analysis assumes that the resonance period  $n > 2$ . The cases  $n = 1$  and  $n = 2$  are special. For  $(m,n) = (1,2)$ , using the same argument, the c-code is given by:  $[1-a_0, a_1, 1-a_1, \dots, 1-a_{n-1}, a_0]$ , therefore the linear code at time  $(2i-1, 2i)$  is  $(1, -1)$ , if  $a_j=1$ ; or  $(-1, 1)$ , if  $a_j=0$ . The c-code for  $(m,n) = (0,1)$  is:  $[a_1 - a_0, a_2 - a_1, \dots, a_0 - a_{n-1}]$ , therefore  $b_i = a_{i+1} - 2a_i + a_{i-1}$ .

Although we only discussed periodic orbits within a given resonance, it is obvious that the same analysis applies to any orbits which are forever trapped inside a resonance. Therefore the  $\mathcal{L}\text{-}\mathcal{R}$  symbol sequence characterizes the entire invariant set of the resonance.

If the  $\mathcal{L}\text{-}\mathcal{R}$  symbol is of period 1, and is  $[1]$ , the linear code is the same as that of  $(m,n)$  orbit. If the  $\mathcal{L}\text{-}\mathcal{R}$  symbol is  $[0]$ , the linear code is that of  $(m,n)$  orbit shifted to the left once. These two cases therefore do not give trapped orbits, but the usual  $(m,n)$  orbit. Heteroclinic orbits from left to right are given by

(...,0,0,1,1,...). In a similar fashion, one can construct codes for heteroclinic orbits approaching higher period orbits in different time directions.

Orbits corresponding to a given linear code do not necessarily exist: one has to check that all configuration points satisfy  $0 < x < 1$ . However, when  $\lambda^n > 3$ , all  $\mathcal{L}\text{-}\mathcal{R}$  codes indeed occur. This criterion is exactly that for the formation of "generalized turnstiles" (see Eq(4.4.8)).

To show that all codes occur, consider the central island of the resonance under the map  $T^n$  (see Figures 4.9). Using the boundary of the upper and lower turnstiles, we divide the central island into three regions (Figure 4.9a). Upon one iteration, the upper homoclinic point  $G$  is mapped to  $E$ , and the lower homoclinic point  $H$  is mapped to  $F$ ; therefore, the vertical strip  $\mathcal{V}_1$  is mapped to  $\mathcal{H}_1$ , and  $\mathcal{V}_2$  to  $\mathcal{H}_2$ , Figure 4.9b. The middle strip  $\mathcal{M}$ , which is the union of the generalized turnstiles, is mapped out of the island. This defines an orientation preserving horseshoe map (except that points mapped out of the resonance may return at some later time; however, this does not affect the structure of invariant set inside the resonance). The invariant set is given by the nonescaping region of the central island under repeated application of the map in both the forward and the backward time directions. The result is a Cantor set which is naturally hyperbolic, thus the dynamics on this invariant set is a full two shift, i.e., is topologically conjugate to a shift map in the space of bi-infinite sequences with two symbols. This justifies our choice of  $\mathcal{L}\text{-}\mathcal{R}$  symbols to code the dynamics. The periodic horseshoe orbits are shown in Figure 4.10 for illustration.

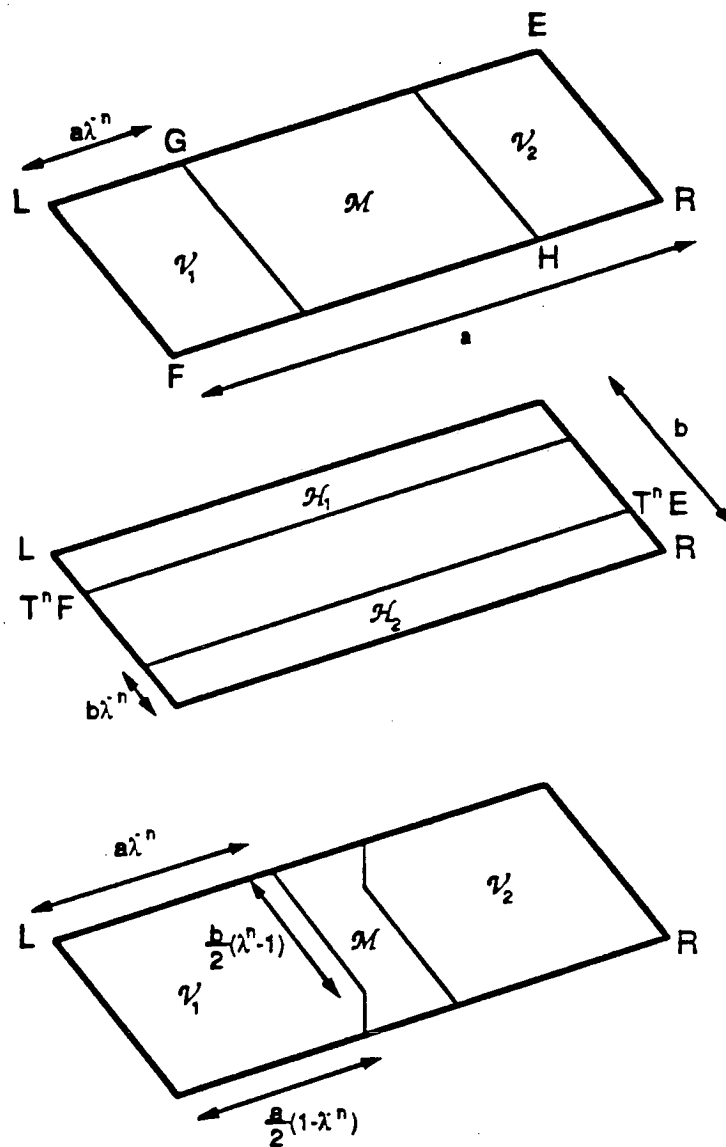


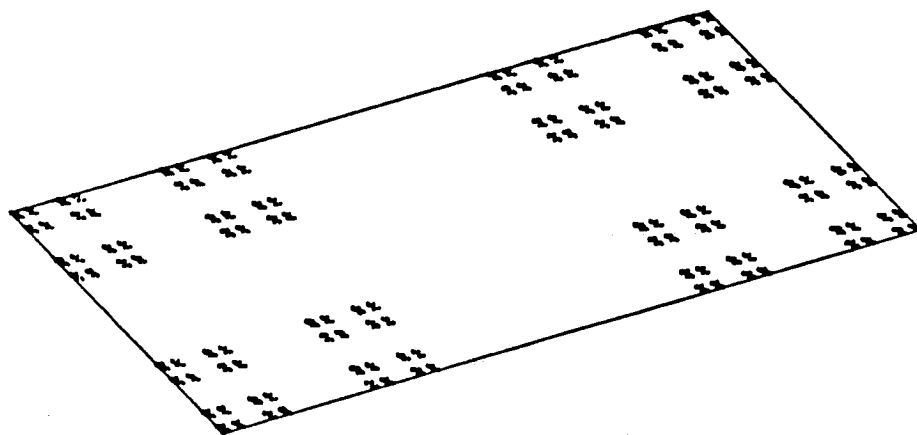
Figure 4.9 Construction of the horseshoe for trapped orbits. For  $\lambda > 3$ , a) the two trapped regions  $\nu_1$ ,  $\nu_2$  and the union of the turnstiles in the central island; b) upon iteration by  $T^m$ ,  $\nu_1$  becomes  $\mathcal{H}_1$  and  $\nu_2$  becomes  $\mathcal{H}_2$ . c) When  $2 < \lambda < 3$ , the union of the turnstiles gives  $\mathcal{M}$ , and the trapped regions are  $\nu_1$  and  $\nu_2$ .



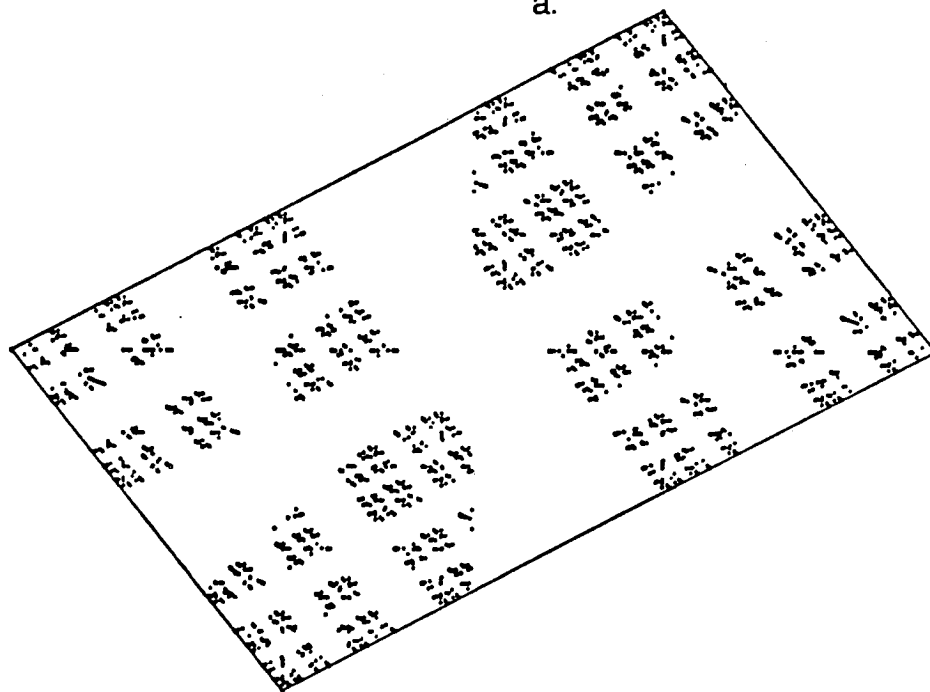
Now it is easy to see why the [0] and the [1] orbits correspond to the same periodic orbit. Under  $T^n$ , the coding implies that the orbit stays in the left or the right rectangle in each level, therefore it is identical to the  $(m,n)$  orbit.

As  $\lambda^n$  decreases from 3, some of the trapped orbits hit the discontinuity line and disappear; the horseshoe is replaced by a fractal structure. When  $\lambda^n$  is near 1, we see an obvious island structure which is reminiscent of a stable, elliptic island. These are illustrated in Figure 4.10 and Figure 4.5.

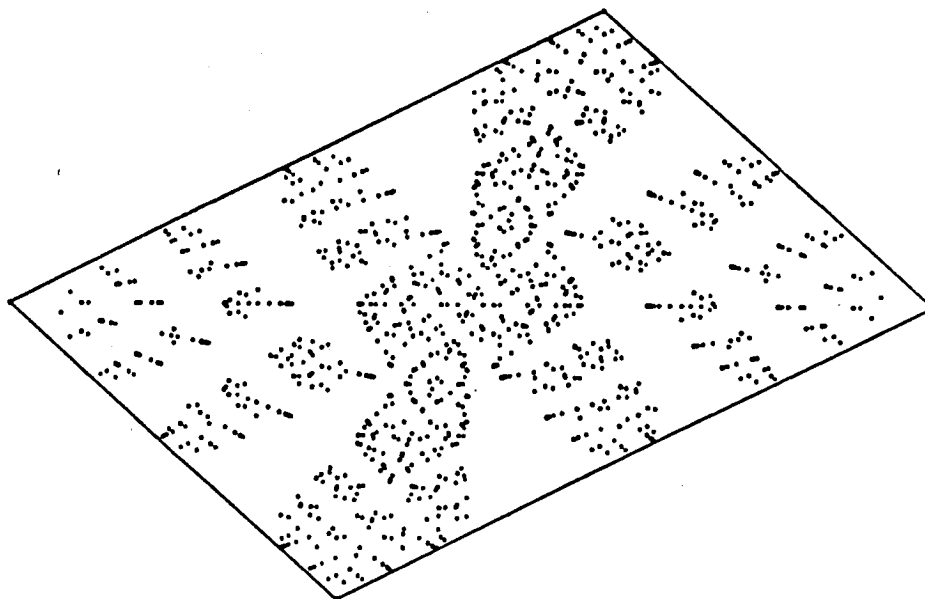
There are horseshoes in the sawtooth map even for very small  $K$ , since for large enough  $n$  there will be always a resonance with  $\lambda^n > 3$ . Furthermore, island orbits do not always correspond to collapsing orbits (A periodic orbit which collapses on to an orbit of lower period at some parameter value. [Bird and Vivaldi, 1988]) — the latter bifurcate only from integer parameter values (cat maps and the twist map). In fact, even the orbits which are purely rotating about the minimax point (ordered orbits under  $T^n$ ) are not all collapsing orbits. This is illustrated in Figure 4.11 for the ordered orbit with rotation frequency  $3/10$ , where we follow the orbit as the parameter  $K$  increases. For small parameter values two of the points fall outside the interval  $0 < x < 1$ , and thus the orbit is a "ghost". It appears only at some finite parameter value.



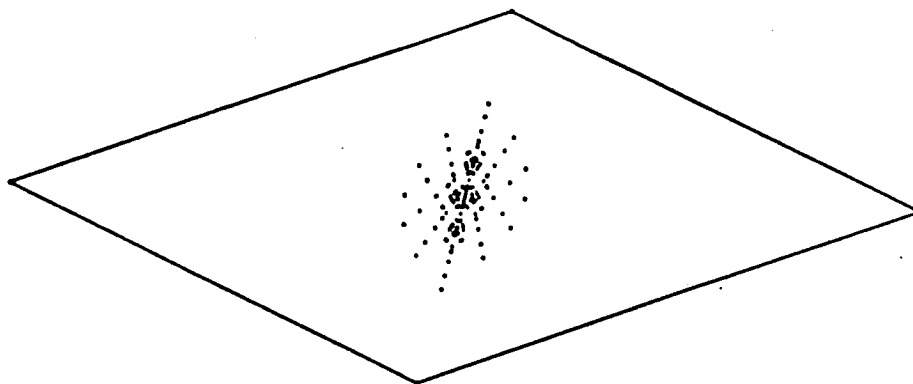
a.



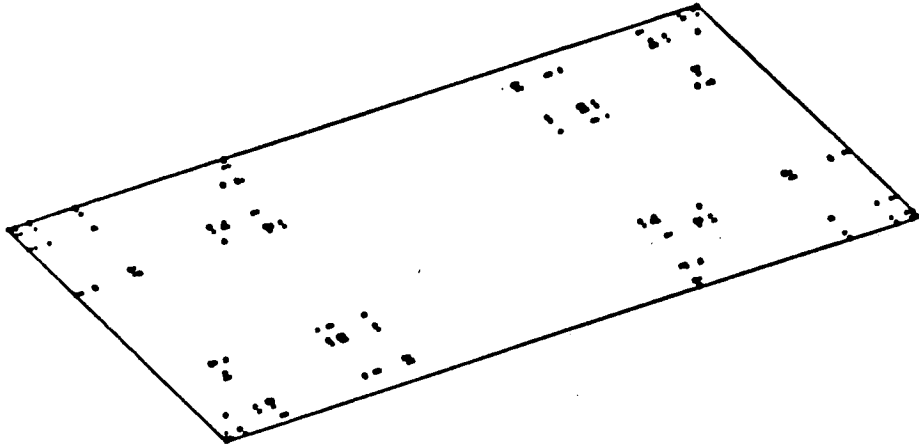
b.



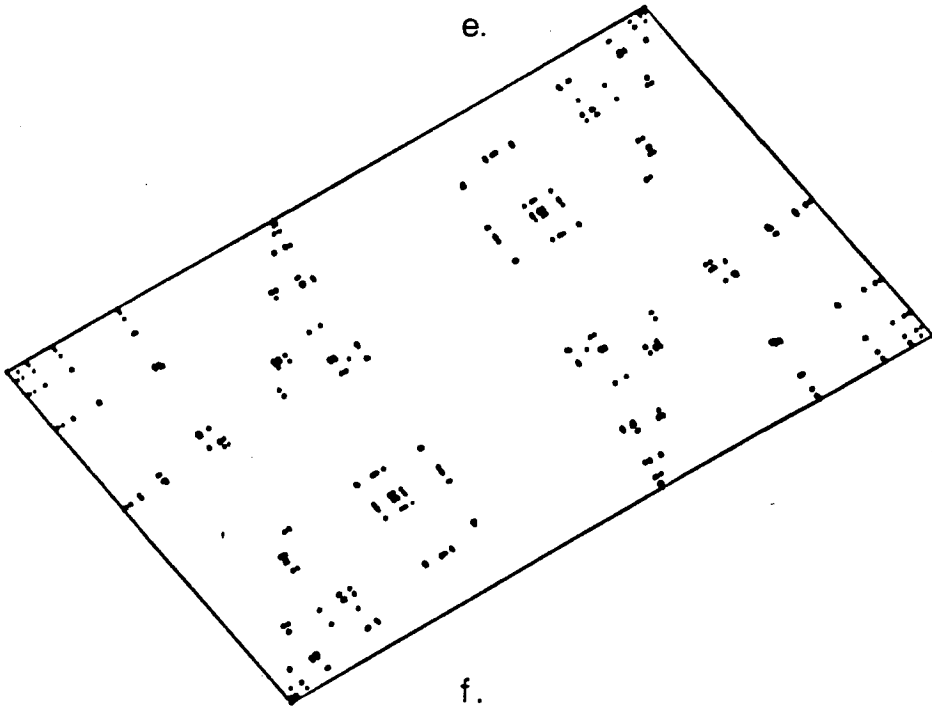
c.



d.



e.



f.

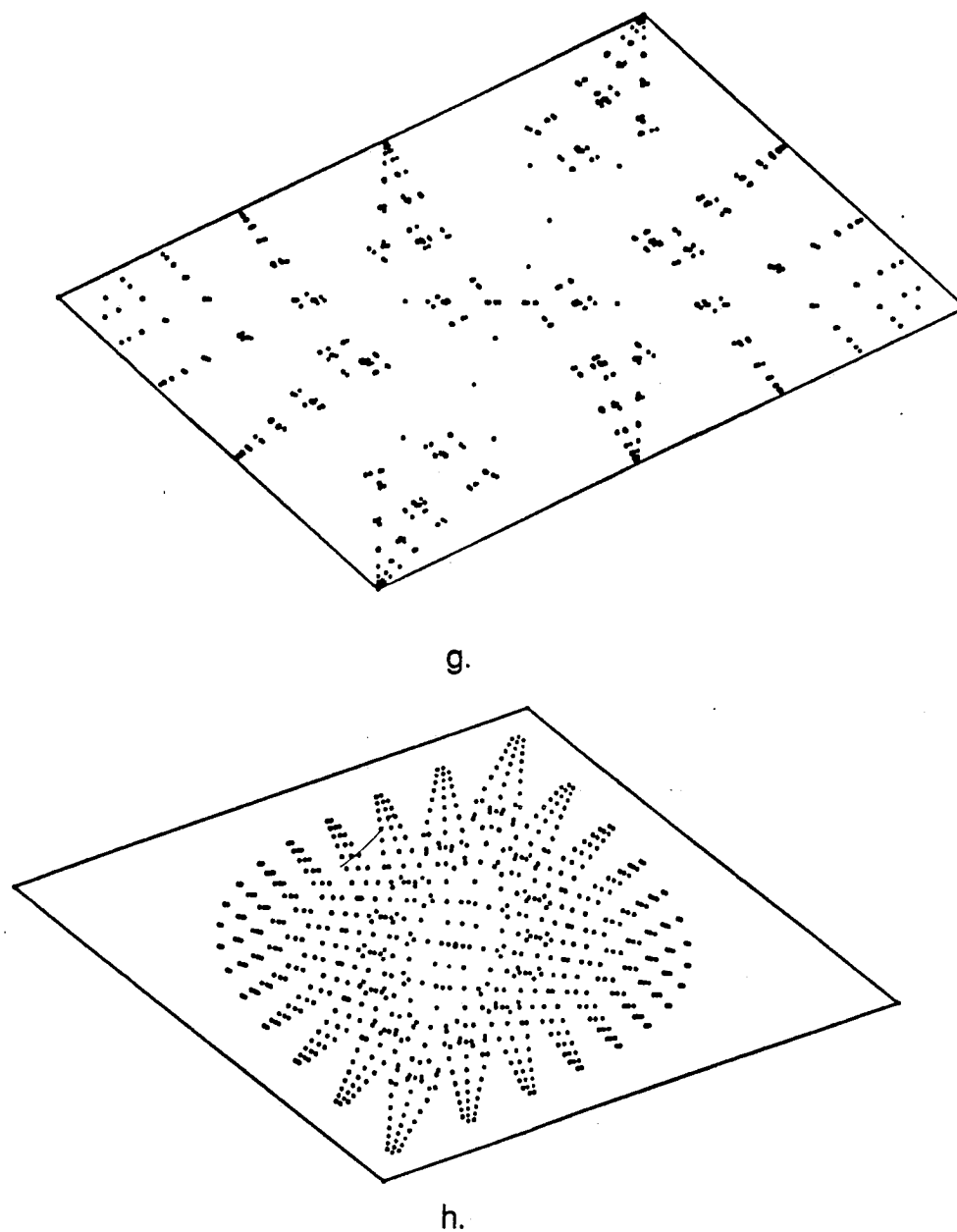


Figure 4.10. Trapped periodic orbits in a resonance, computed using the L-R coding up to a maximum period. a)  $\lambda = 3.2$ , period 8; b)  $\lambda = 2.4$ , period 12; c)  $\lambda = 1.8$ , period 12; d)  $\lambda = 1.2$ , period 12. Trapped rotational (ordered) periodic orbits in a resonance up to a maximum period. e)  $\lambda = 3.2$ , period 20; f)  $\lambda = 2.4$ , period 20; g)  $\lambda = 1.8$ , period 20; h)  $\lambda = 1.2$ , period 30.

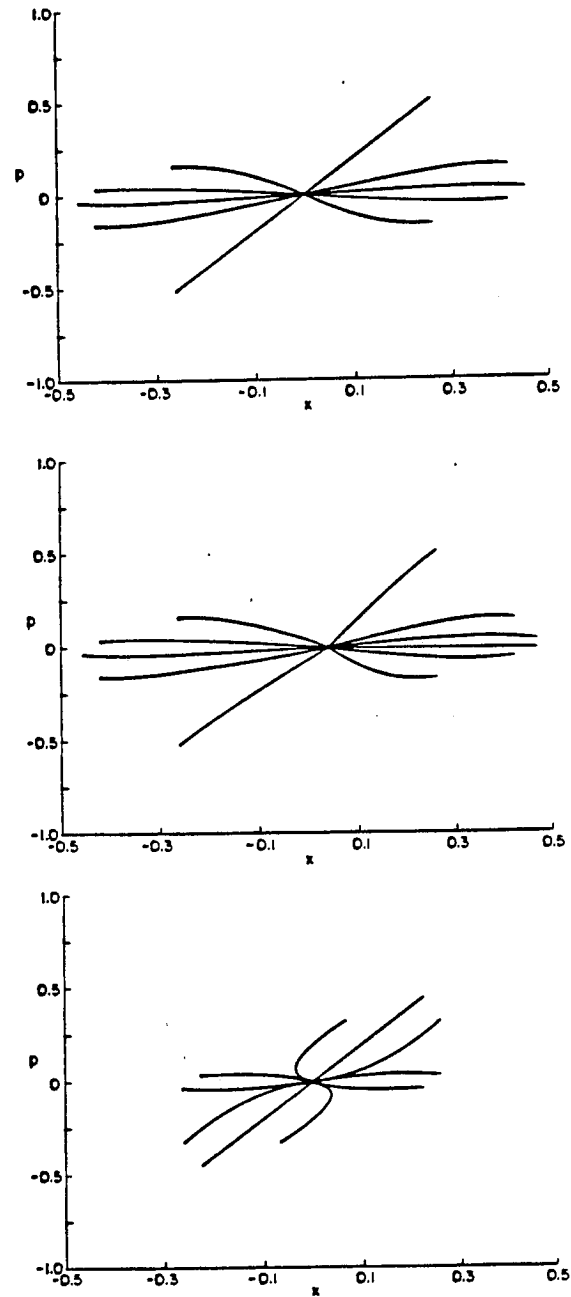


Figure 4.11. Trapped rotational periodic orbits. Each figure shows the points in a single orbit as a function of  $\lambda$  as  $\lambda$  varies from 1 to 3. a) 1/10 orbit; b) 1/11 orbit; c) 3/10 orbit.

The number of periodic points trapped in a resonance grows exponentially with the period. For a period  $n$  resonance the exponent is given by the topological entropy of  $T^n$  restricted to the central island of the resonance. Let  $N_n(q, \lambda)$  be the number of periodic points in the  $(m, n)$  central island with period  $nq$  at Lyapunov exponent  $\log(\lambda)$ . It has the universal form (see below)

$$N_n(q, \lambda) = N_1(q, \lambda^n)$$

The topological entropy of  $T^n$ , restricted to the central island, is defined as

$$\text{ent}(\lambda^n) = \lim_{q \rightarrow \infty} \log(N(q, \lambda)) / q \quad (4.5.1)$$

In Figure 4.12, we plot  $\log(N(q, \lambda))$  as a function of  $q$ ; the straight line is the entropy multiplied by the period.

When  $\lambda^n > 3$ , the dynamics is a full two shift; therefore, there are  $2^q$  period  $q$  points, and the entropy is  $\log 2$ .

The fact that  $N_n(q, \lambda)$  has a universal form follows from the fact that the dynamics in the central island can be described geometrically. As is clear from Figures 4.9 and 4.10, there are basically three cases:  $\lambda^n > 3$ ,  $2 < \lambda^n < 3$ ,  $\lambda^n < 2$ . In each of these cases we can divide the central island into three regions. The central strip  $\mathcal{M}$  is given by the union of the turnstiles or the generalized turnstiles (Figure 4.9); therefore, in studying the invariant set, we can neglect this strip. The remaining two strips, labelled  $\mathcal{L}$  and  $\mathcal{R}$ , are defined as the regions which stay in the resonance upon one iteration, and do not include the points

on the discontinuity line. For these strips, the map is geometrically equivalent to squeezing by a factor  $\lambda^n$  in the stable direction and stretching by a factor  $\lambda^n$  in the unstable direction; the left and right  $(m,n)$  periodic points remain fixed, respectively. As shown in Figure 4.9, the detailed shape of the  $\mathcal{L}$  and  $\mathcal{R}$  stripes depends only on  $\lambda^n$ ; therefore, the number of periodic points under the map  $T^n$  depends only on  $\lambda^n$ .

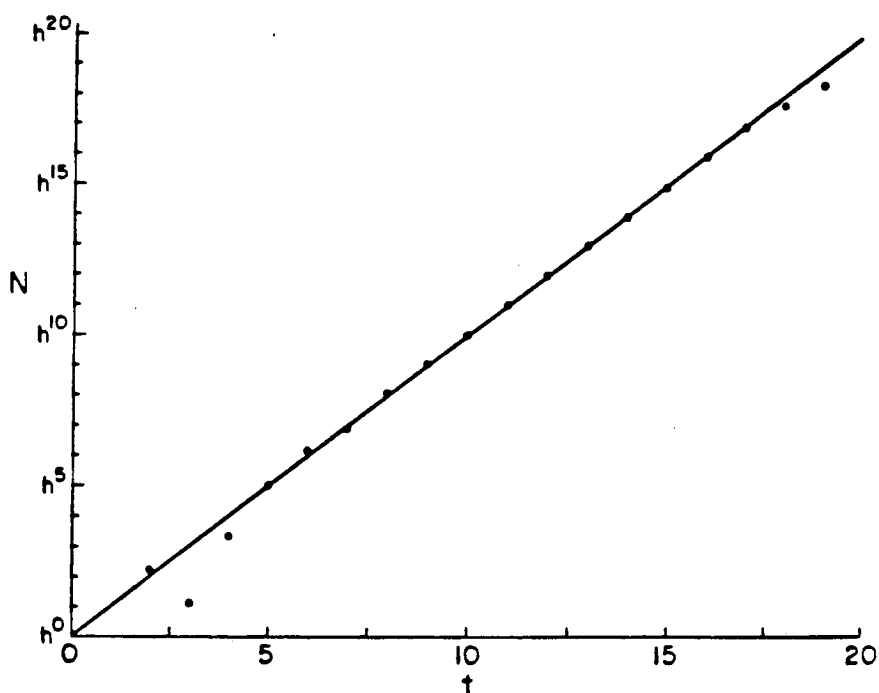


Figure 4.12. Number of periodic points trapped in a resonance as a function of period at  $\lambda = 2.4$ . The number is shown on a logarithmic axis with a base 1.859 which is the exponential of the entropy. For period 18 and 19,  $N(t)$  deviates from the exponential due to numerical error in checking the existence condition (4.2.15).



#### 4.6. Escape Dynamics

The simplest transport problem is the escape dynamics from a single resonance: iterate the  $(m,n)$  resonance  $t$  times and determine the area  $A_n(t)$  which remains in the resonance at time  $t$ . The survival probability at time  $t$  is

$$P_n(t) = A_n(t)/A_n(0) \quad (4.6.1)$$

We expect a priori that this probability decays exponentially, since the map is hyperbolic, and so define the escape rate

$$\alpha_n(\lambda) = - \lim_{t \rightarrow \infty} \ln(P_n(t))/t \quad (4.6.2)$$

Since escape takes place only in the central island, we may restrict consideration to it under the map  $T^n$ . As discussed in Sec. 4.5, the dynamics only depends on  $\lambda^n$ ; therefore the escape rate has a universal form

$$\alpha_n(\lambda) = \alpha_1(\lambda^n)/n \quad (4.6.3)$$

Hence we need only study escape from the  $(0,1)$  resonance, which facilitates the numerical simulation considerably. For simplicity, we drop the subscript 1 in the following discussion.

When  $\lambda > 3$ , we can obtain a closed form for the escape rate, because the dynamics restricted to the central island is a horseshoe map. Upon each iteration, the central strip is mapped out of the resonance; the  $\mathcal{L}$  and the  $\mathcal{R}$  strips are squeezed uniformly by a factor  $\lambda$  in the stable direction and stretched

uniformly by a factor  $\lambda$  in the unstable direction,  $2/\lambda$  of the original area remains either in the  $\mathcal{L}$  strip or the  $\mathcal{R}$  strip. Therefore the survival probability at time  $t$  is given by

$$P(t) = (2/\lambda)^t$$

hence the escape rate is

$$\alpha(\lambda) = \ln(\lambda) - \ln(2) \quad (4.6.4)$$

In this case, the dynamics is purely mixing: under each iteration the remaining area is stretched uniformly in the unstable direction, covering uniformly the width of the resonance. Thus the escaping probability in each iteration is the same for all time. This means the escape is a Markov process when  $\lambda > 3$ .

The Markov transport model predicts the probability of escape per step is Flux/Area; therefore,

$$P(t) = (1 - \text{Flux}/\text{Area})^t \quad (4.6.5)$$

Using the flux and area formulas (4.4.2), (4.4.9) and (4.4.4), the Markov rate is

$$\begin{aligned} \alpha_M(\lambda) &= -\log(1 - [\lambda + \lambda^{-1} - 2]/4) & \lambda < 3 \\ &= \log(\lambda) - \log(2) & \lambda > 3 \end{aligned} \quad (4.6.6)$$

Thus this rate is exact for  $\lambda > 3$ .

Equation (4.6.4) can also be interpreted in the following way. Let  $\mathcal{T}$  be the trapped, invariant set of the resonance. Then,

according to the ergodic theory (the so called "thermodynamic formalism"), the escape rate is the difference between the Lyapunov exponent and the topological entropy of  $\mathcal{T}$ :

$$\alpha_L(\mathcal{T}) = L(\mathcal{T}) - \text{ent}(\mathcal{T}) \quad (4.6.7)$$

This is generally true for hyperbolic systems [Bohr and Rand 1987; Grebogi, Ott, and Yorke 1988; Hsu, Ott, and Grebogi 1988]. For the sawtooth resonance, when  $\lambda > 3$ , the topological entropy is  $\ln 2$ , therefore, the escape rate is given by (4.6.4).

When  $\lambda < 3$ , the entropy, and hence  $\alpha_L$ , can be determined numerically by computing the growth rate of the number of periodic orbits. This number decreases with  $1/\lambda$  as the various  $\mathcal{L}$ - $\mathcal{R}$  coded orbits cross the discontinuity, and cease to exist (see Sec 4.5). Checking the existence of the periodic orbits, involves large computations which we find are prohibitive when  $\lambda < 1.5$ .

Numerical calculation of the escape rate can be done by a Monte Carlo experiment: begin with a large number of points (typically  $10^6$ - $10^7$ ) chosen on an equally spaced grid within the resonance. Iterate each point, computing the number of points still trapped at time  $t$ . Figure 4.13a shows the log plot of the survival probability. An average escape rate at time  $t$  can be defined as

$$\beta(t) = - \log(P(t)) / t \quad (4.6.8)$$

At  $t = 1$ , this rate is exactly given by (4.6.6), since this rate refers to the fraction of area lost after one step. The numerical

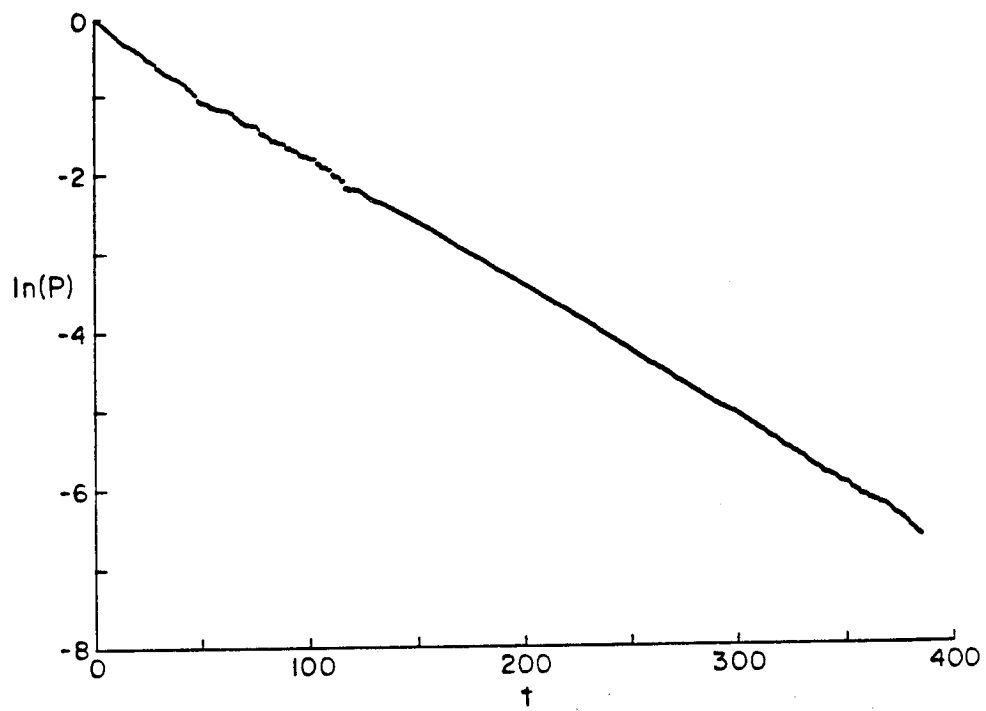
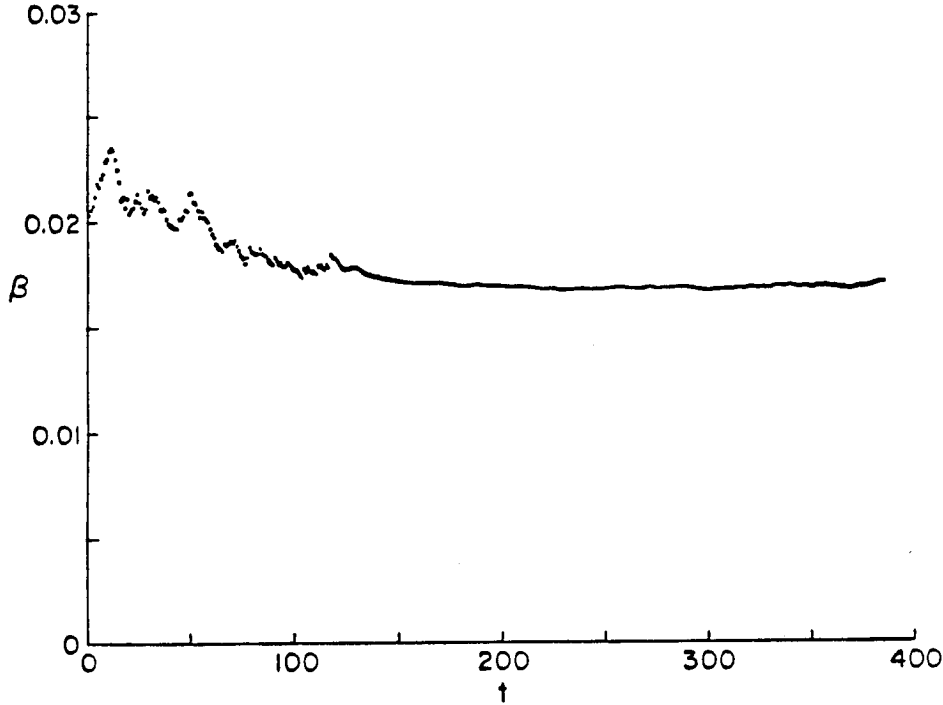


Figure 4.13. a) Typical survival probability.



b) Typical numerical average escape rate.

computations show that the rate oscillates wildly for small time, and then relaxes to a fixed value (Figure 4.13b). This oscillation is more prominent for small Lyapunov number  $\lambda$ . Beyond the oscillation region, the survival probability decays exponentially until there are too few trapped points remaining (around  $10^3$ ) to obtain good statistics. We derive a numerical escape rate by a least square fit to the exponential regime.

Figure 4.14 shows escape rate as a function of  $\lambda$ . The smooth curve is the Markov rate (4.6.6). The second curve is the Lyapunov rate (4.6.7); this is computed by finding all the trapped periodic orbits up to a period such that the relative error in the rate is 2%. We are unable to calculate the entropy for small  $\lambda$  because the convergence with period is extremely slow and excessive computations are needed. The numerical rates, from a fit to (4.6.2), are plotted as triangles. At moderate values of  $\lambda$  ( $1.5 < \lambda < 3.0$ ), the numerical rates tend to fall close to  $\alpha_L$ . However,  $\alpha_M$  is not far off, the largest discrepancy is about 10%; therefore, we conclude  $\alpha_M$  is a reasonable estimate of the exact rate, especially because computing the entropy requires considerable CRAY time.

For small  $\lambda$ , the deviation between  $\alpha_M$  and numerical rates is larger; in fact, the two rates approach zero with different functional forms. This is shown in Figure 4.15a, which plots the escape rates as a function of absolute value of the residue

For small residue, from Eq(4.6.6),  $\alpha_M$  scales linearly with R

$$\alpha_M = |R| \quad (4.6.9)$$

Fig 4.15b shows the numerical rate scales with a different exponent. The least square fitting gives an exponent of  $1.4 \pm 0.1$ .

This error estimate is not reliable, due to the difficulty of fitting the numerical escape rate data to an exponential. It seems reasonable that the exponent for the escape rate will be the same as the exponent for the flux at small residue, which is 1.5. The deviation from the Markov model for small nonlinearity is not surprising: to properly describe escape from a single resonance, we need to compute the area which has entered the resonance from the incoming turnstiles. This area corresponds to empty regions which are seen in the Monte Carlo simulations. When the Lyapunov exponent is sufficiently small, these regions are not very well mixed, and the Markov description is no longer valid.

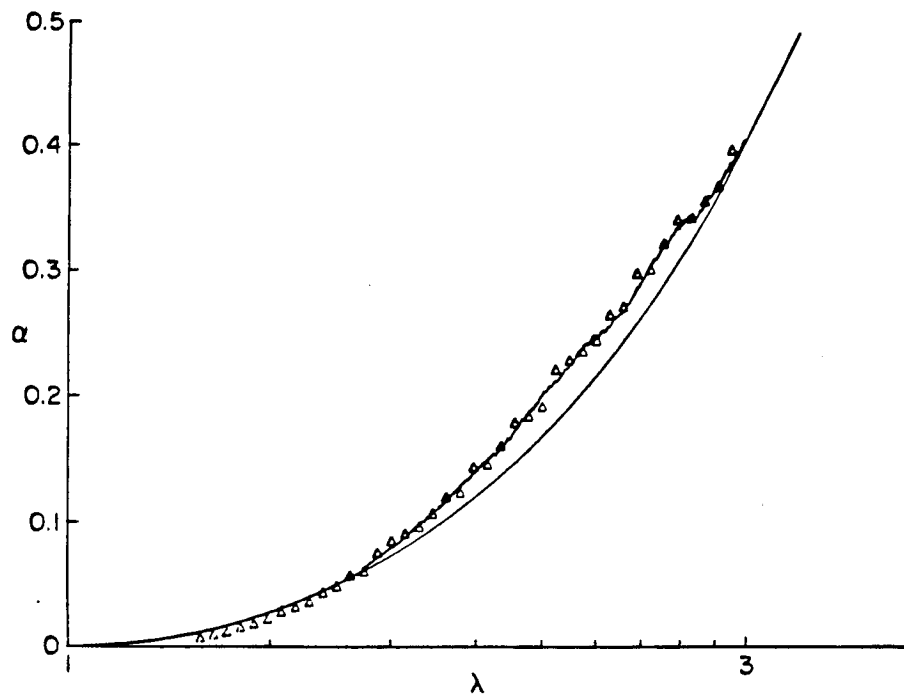


Figure 4.14. Escape rate as a function of the Lyapunov number. The smooth line is the Markov rate. The dark line is the Lyapunov rate, which is determined numerically within a 2% relative error. The Triangles are numerical rates.



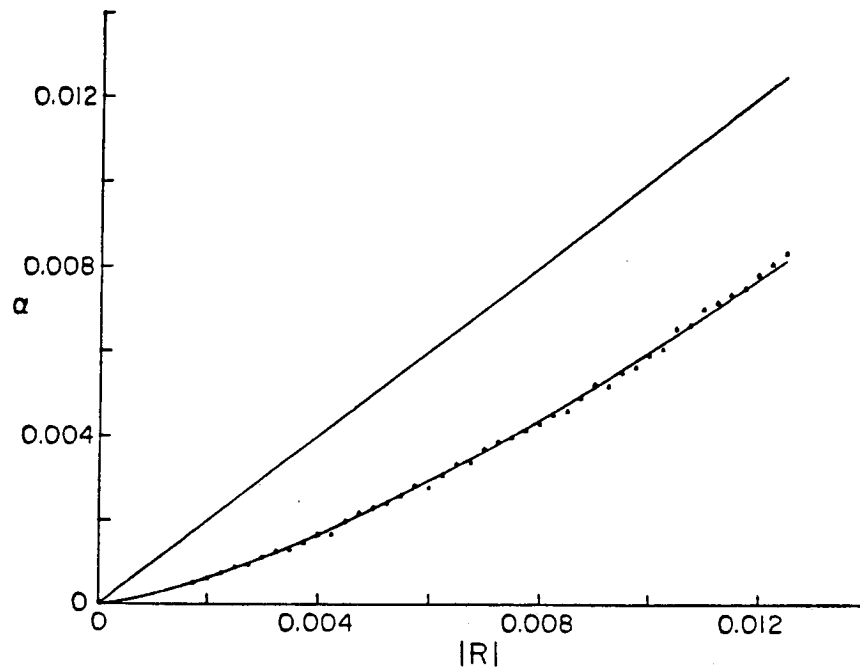
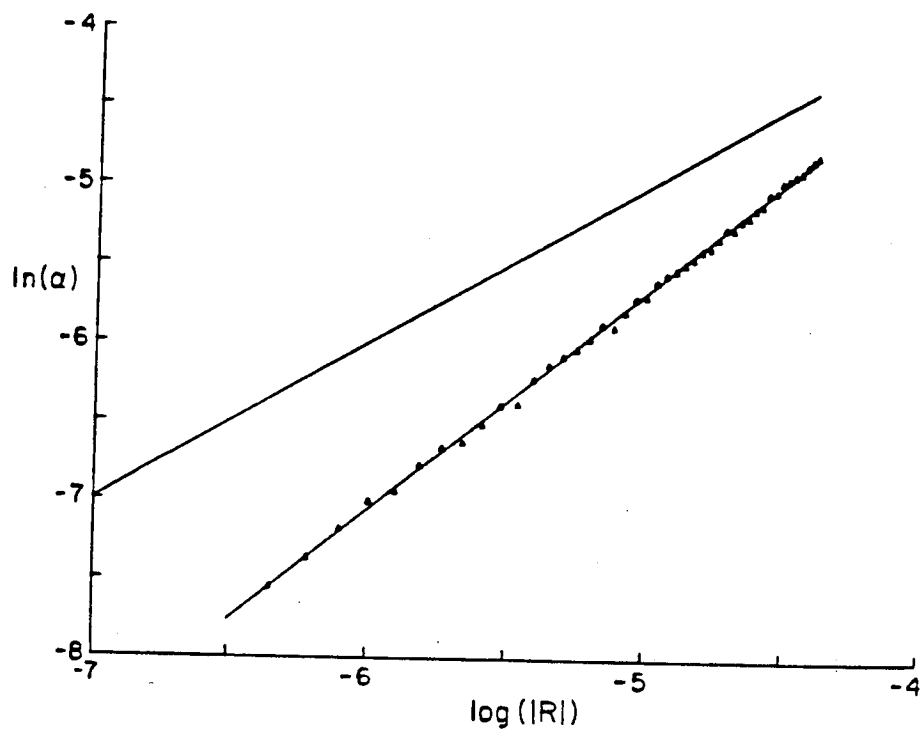


Figure 4.15a. Escape rate as a function of the residue at small residue. The straight line is the Markov rate. The triangles are numerical rates. The smooth curve is the least square fit to a scaling behavior.



b) Escape rate in log-log plot.

---

## **Chapter 5**

# **Orbit Extension Methods for Finding Unstable Orbits**

## 5.1. Introduction

Calculation of flux and resonance area requires the computation of various orbits. These orbits are typically unstable; indeed, the stochastic region is characterized by positive Lyapunov exponents, which implies sensitive dependence on initial conditions. It is a delicate procedure to locate highly unstable orbits for Hamiltonian systems.

As we have seen, periodic orbits play important roles in the dynamics in the chaotic regime. The set of periodic orbits are usually dense in the invariant set of the dynamics. Quasiperiodic orbits and homoclinic orbits are well approximated by periodic orbits, the error of the approximation decreases exponentially with the increase of the period, the exponent is negative of the Lyapunov exponent [MacKay, Meiss and Percival 1987]. Hence in studying chaotic dynamics, small period orbits usually give good approximations to the real orbits. We shall mostly concentrate on periodic orbits; however, methods to locate heteroclinic and homoclinic orbits are also discussed.

The easiest method to find periodic orbits is by iterating the map, searching for a point which returns to itself. This is particularly useful when the map possesses a reflection symmetry (involution). The method is well documented [MacKay 82], and is successful in locating periodic orbits with small residues. However, this method is doomed to failure due to the exponential growth of numerical errors when the orbit is highly unstable (even if the orbit is followed by continuation of parameter from a stable region). A variation of the method is to implement the usual two dimensional Newton method to root

search the equation  $T^n(x,p) = (x,p)$ . However this suffers the same disease that plagues the iteration method.

The fact that twist maps are Lagrangian systems makes it possible to design stable numerical methods to find highly unstable orbits. Orbits are stationary configurations of the action

$$W(\{X_t\}) = \sum_{t=M}^N F(X_t, X_{t+1}) \quad (5.1.1)$$

For type  $(m,n)$  period orbits, the variation is under the constraint  $X_n = X_0 + m$ .

Aubry has introduced the gradient flow method [Aubry 1983a; Peyrard and Aubry 1983]. This consists of integrating the set of  $n$  coupled differential equations to find the minimum of  $W$

$$\frac{dX_t}{d\tau} = -\nabla W \quad (5.1.2)$$

where  $\tau$  is an integral parameter. Starting with an initial condition of uniform rotation with frequency  $m/n$ :  $X_t(0) = \text{int}(\frac{m}{n}t + \alpha)$ , the limiting configuration  $\{X_t(\infty)\}$  gives the minimizing  $(m,n)$  orbit [Angenent 1984]. This method converges well for unstable orbits but requires the integration of a large number of coupled differential equations, one for each point on the orbit, and is often too time consuming.

Implementing Newton's method for the Lagrangian system gives a more stable iteration scheme [Schnellnuber et. al. 1986]. Its main disadvantage is that the computation time grows

rapidly with the increase of the period. With the usual brute force LU decomposition to invert the matrix, the computation time grows as  $n^3$  [Press et. al. 1986], where  $n$  is the period of the orbit. However, for Lagrangian systems, the computation can be arranged so that the time to find an orbit is proportional to its period [Mestel and Percival 1987, Kook and Meiss 1989b]. Another problem with this method is that it is so sensitive to the initial configuration that one often ends up with a different orbit than that desired: in chaotic regions there are many close periodic orbits with the same period. Thus careful choice of initial configuration is necessary for success, and most of the effort in finding an orbit is often spent searching for good starting points. This slows the method considerably (especially for long orbits) and diminishes its advantage.

In this chapter, we address the above two problems. We first review the Green function method for Newton iteration scheme proposed by Mestel and Percival [1987], then a few relevant methods for determining good initial configurations. A detailed account of the orbit extension method for finding both ordered periodic orbits and principal heteroclinic orbits is presented in section 3. The latter can be viewed as both a stable and an economic prescription for finding unstable orbits and resolves the initial configuration problem with Newton's method. We have applied the method to the determination of flux from one resonance to another, this requires the calculation of highly unstable orbits heteroclinic between two resonances. The Newton iteration program for periodic orbits is published elsewhere [Chen and Mestel 1988].

## 5.2 Newton's Method

### 2.1 Twist Maps

We will consider a twist map acting on a phase space  $(X,p)$  given by

$$\begin{aligned} p_{t+1} &= p_t - V'_k(X_t) \\ X_{t+1} &= X_t + p_{t+1} \end{aligned} \quad (5.2.1)$$

or equivalently in Lagrangian form by

$$X_{t+1} - 2X_t + X_{t-1} = - V'_k(X_t) \quad (5.2.2)$$

Here  $V'_k$  is the derivative of a potential,  $V_k$ , which is assumed to be an even, periodic function of  $X$ , with period unity:

$$V_k(X+1) = V_k(X) = V_k(-X) \quad (5.2.3)$$

The periodicity condition implies that  $X$  can be folded into the unit interval (e.g. modulus unity) so the phase space can be considered as a cylinder. We shall denote the angle variable in the half open unit interval  $[0,1)$  by lower case letter  $x$ .

The evenness condition (5.2.3) implies that the map has a reflection symmetry about  $x=0$ , and is not necessary for our method, but convenient for the exposition. Without loss of generality, we also assume that the extremum of  $V$  at  $x=0$  is a minimum. For simplicity, we also suppose that this is its only minimum. Equation (5.2.3) then implies that its maximum occurs at  $x=1/2$ . Finally, the potential depends on a parameter  $k$ , which governs its height, and thus the strength of the nonlinearity.

An orbit is denoted by its configuration  $\{ X_t, t = \dots, -1, 0, 1, \dots \}$ .  
Orbits are stationary points of the action

$$W(\{X_t\}) = \sum_{t=M}^N F(X_t, X_{t+1}) \quad (5.2.4)$$

where the generating function is

$$F(X, X') = (X - X')^2 / 2 - V_k(X) \quad (5.2.5)$$

If the orbit has period  $n$ , we denote it by  $\{X_1, X_2, \dots, X_{n-1}\}$ , with the additional conditions  $X_{t+n} = X_t$  for some integer  $m$ . Of course, knowledge of the configurations also gives the momenta through (5.2.1).



### 5.2.2 Newton's Method

There are two steps to find a periodic orbit with Newton's method: i) choose a trial configuration, and ii) invert the tangent map. For a given trial configuration  $\{X_t\}$  (on the real line), the tangent map or Hessian matrix is

$$H_{i,j} = \delta_{i+1,j} - 2\delta_{i,j} + \delta_{i-1,j} + \delta_{i,j}V''(X_j) , \quad (5.2.6)$$

where  $\delta$  is the Kronecker delta, and we have omitted the "k" subscript on  $V$  for clarity. To obtain the correction,  $\{\Delta X_t\}$ , to  $\{X_t\}$  one must solve the linear problem

$$\sum_j H_{i,j} \Delta X_j = R_i , \quad (5.2.7)$$

where  $R_i$  is the residue vector of the trial configuration:

$$R_i = X_{i+1} - 2X_i + X_{i-1} + V'(X_i) . \quad (5.2.8)$$

Solution of (5.2.7) requires proper consideration of the boundary conditions of the desired orbit. For example, suppose we are interested in an orbit which begins at  $X_{-1}$  and reaches  $X_n$  after  $n+1$  iterations. In this case the  $0^{\text{th}}$  and  $(n-1)^{\text{st}}$  rows of  $H$  must be changed since  $\Delta X_{-1}$  and  $\Delta X_n$  are zero. In contrast, to find a periodic orbit, we set  $\Delta X_n = \Delta X_0$  and  $\Delta X_{-1} = \Delta X_{n-1}$ , but allow  $\Delta X_{-1}$  and  $\Delta X_{n-1}$  to vary. In this case the Hessian becomes a periodic matrix.

Consider first the case of periodic boundary conditions. Define a periodic  $n \times n$  Hessian matrix  $h$  by

$$\begin{aligned}
h_{0,j} &= \delta_{1,j} - 2\delta_{0,j} + \delta_{n-1,j} + \delta_{0,j}V''(X_0) \\
h_{i,j} &= \delta_{i+1,j} - 2\delta_{i,j} + \delta_{i-1,j} + \delta_{i,j}V''(X_j) \quad 0 < i < n-1 \\
h_{n-1,j} &= \delta_{0,j} - 2\delta_{n-1,j} + \delta_{n-2,j} + \delta_{n-1,j}V''(X_{n-1}) \\
h_{i+n,j+n} &= h_{i,j}
\end{aligned} \tag{5.2.9}$$

The inverse matrix  $g$  which we call the periodic Green's function satisfies

$$\sum_{j=0}^{n-1} h_{i,j}g_{j,k} = \delta_{i,k}$$

Since  $h$  is periodic and symmetric, the inverse matrix is also periodic and symmetric.

$$\sum_{j=0}^{n-1} h_{i,j}g_{k,j} = \delta_{i,k} \tag{5.2.10}$$

Let  $G_{k,j}$  be the Green function of  $H_{i,j}$  in (5.2.6), it satisfies

$$\sum_{j=-\infty}^{\infty} H_{i,j}G_{k,j} = \delta_{i,k} \tag{5.2.11}$$

Following Mestel and Percival [1987], We represent the Green function  $g_{k,j}$  by  $G_{k,j}$  in the fundamental interval  $k \leq j < k+n$  for fixed  $k$  and periodically extend  $g_{k,j}$

$$g_{k+n,j} = g_{k,j+n} = g_{k,j} = G_{k,j} \quad k \leq j < n+k \tag{5.2.12}$$

Hence

$$\Delta X_k = \sum_{j=k}^{k+n-1} g_{k,j}R_j \tag{5.2.13}$$

Let  $w^+{}_j, w^-{}_j$  be two linearly independent solutions of the homogeneous equation

$$\sum_{j=-\infty}^{\infty} H_{i,j} w_j = w_{i+1} - (2-V''(X_i))w_i + w_{i-1} = 0 \quad (5.2.14)$$

Then we can write

$$G_{k,j} = A_k w^+{}_j + B_k w^-{}_j \quad j \geq k \quad (5.2.15)$$

$G_{k,j}$  satisfy the inhomogeneous equation (5.2.11)

$$\sum_{j=-\infty}^{\infty} H_{i,j} G_{k,j} = G_{k,i+1} - (2-V''(X_i))G_{k,i} + G_{k,i-1} = \delta_{i,k} \quad (5.2.16)$$

Note that all components of (5.2.16) for fixed  $k$  are satisfied by  $g_{k,j}$  except for the  $j = k$  and  $k+n-1$  components

$$g_{k,k+1} - (2-V''(X_k))g_{k,k} + g_{k,k+n-1} = 1 \quad (5.2.17a)$$

$$g_{k,k} - (2-V''(X_{k-1}))g_{k,k+n-1} + g_{k,k+n-2} = 0 \quad (5.2.17b)$$

The second equation will be identical to the  $k+n-1$  component of (5.2.16) if we choose

$$G_{k,k} = G_{k,k+n} \quad (5.2.18a)$$

Now consider the  $k+n$  component of (5.2.16)

$$G_{k,k+1+n} - (2-V''(X_k))G_{k,k+n} + G_{k,k+n-1} = 0$$

This will be identical to (5.2.17a) if

$$G_{k,k+1} - G_{k,k+n+1} = 1 \quad (5.2.18b)$$

Equations (5.2.18a) and (5.2.18b) determines  $g_{k,j}$  completely through (5.2.12) and (5.2.15).

The Wronskian of two solutions  $u_t, v_t$  of the homogeneous second order difference equation (5.2.14) is

$$Wr(u_t, v_t) = u_t v_{t+1} - u_{t+1} v_t \quad (5.2.19)$$

It can be verified that it is a constant.

Solving (5.2.18) for  $A_k$  and  $B_k$ , we get

$$g_{k,j} = [(w^{-}_{k+n} - w^{-}_k) w^{+}_j - (w^{+}_{k+n} - w^{+}_k) w^{-}_j] / \Delta \quad (5.2.20)$$

where  $\Delta$  is determined by the equation

$$\Delta = Wr(w^{n-}, w^{n+}) + Wr(w^{+}, w^{n-}) + Wr(w^{n+}, w^{-}) + Wr(w^{-}, w^{+}) \quad (5.2.21)$$

and

$$\begin{aligned} w^{n+}_t &= w^{+}_{n+t} \\ w^{n-}_t &= w^{-}_{n+t} \end{aligned}$$

Hence (5.2.7) becomes

$$\Delta X_k = [(w^{-}_{k+n} - w^{-}_k) c^{+}_k - (w^{+}_{k+n} - w^{+}_k) c^{-}_k] / \Delta \quad (5.2.22)$$

$$c^{+}_k = \sum_{j=k}^{k+n-1} w^{+}_j R_j \quad (5.2.23)$$

and similarly for  $c_k^-$ . The time for computing  $\Delta X_k$  can now be arranged to be linear in  $n$  [Mestel and Percival 1987].

To find an orbit with given initial and final points  $x_{-1}$  and  $x_n$ , we again use the Green function, with a slight modification in boundary conditions. The Hessian matrix  $h$  at the endpoints is now given by

$$\begin{aligned} h_{0,j} &= \delta_{1,j} - 2\delta_{0,j} + \delta_{0,j}V''(X_0) \\ h_{i,j} &= \delta_{i+1,j} - 2\delta_{i,j} + \delta_{i-1,j} + \delta_{i,j}V''(X_j) \\ h_{n-1,j} &= -2\delta_{n-1,j} + \delta_{n-2,j} + \delta_{n-1,j}V''(X_{n-1}) \end{aligned} \quad (5.2.24)$$

For  $k = 0, n-1$ , we can again write the solution  $g_{k,j}$  in terms of  $G_{k,j}$ .  $G_{0,j}, G_{n-1,j}$  will satisfy the same equation by  $g_{0,j}, g_{n-1,j}$  if

$$\begin{aligned} G_{0,-1} &= 1, \quad G_{0,n} = 0 \\ G_{n-1,-1} &= 0, \quad G_{n-1,n} = 1 \end{aligned} \quad (5.2.25)$$

For  $k \neq 0, n-1$ , we write

$$\begin{aligned} g_{k,j} &= G_{k,j} & 0 \leq j \leq k \\ g_{k,j} &= G'_{k,j} & k \leq j \leq n-1 \end{aligned} \quad (5.2.26)$$

where

$$\begin{aligned} G_{k,j} &= A_k w^{+j} + B_k w^{-j} \\ G'_{k,j} &= A'_k w^{+j} + B'_k w^{-j} \end{aligned}$$

The boundary conditions are now given by

$$\begin{aligned} G_{k,-1} = 1, \quad G'_{k,k+1} - G_{k,k+1} = 1 \\ G'_{k,n} = 1, \quad G_{k,k-1} - G'_{k,k-1} = 1 \end{aligned} \tag{5.2.23}$$

Asymptotic boundary conditions can be treated in exactly the same way as the fixed boundary condition so long as we know the asymptotic behavior.

The linear problems for all three boundary conditions are identical for those points of the orbits far from the boundaries. Thus, if the corrections to the boundary points are small, the method used in the periodic case can even be used for the other boundary conditions without much error.

### 5.3. Orbit Extension Method

The success of Newton's method depends critically on the choice of initial configuration. Unlike the gradient flow method, for which one is able to find the basin of attraction (in fact, order is preserved by this method [Angenent 1984]), The basin of attraction for Newton's method is a very complicated fractal set in  $n$  dimensional complex space for a period  $n$  orbit. This makes the choice of the initial configuration an extremely delicate matter. There are two cases for which a simple choice suffices for ordered orbits: when the orbit is stable or has small residue, uniform rotation gives a sufficiently good initial guess; when the residue is extremely large, all points of the orbit fall into the minimum of the potential, and a sensible initial configuration can be perturbatively obtained to first order in  $1/k$ .

While these techniques work reasonably well in the two limiting cases, they still suffer from the problem that many iterations are required to find the correct orbit; this makes the method very time-consuming for long orbits. Furthermore they are applicable only to ordered orbits. Our purpose is to develop a method which will be quicker and more flexible.

The orbit extension method is motivated by the fact that often a short orbit segment already contains a lot of information about a related longer orbit. This is familiar in the case of quasi-periodic orbits, which can be approximated by periodic orbits with nearby frequencies. The orbit extension method is a technique for systematically exploiting the bunching of points on an unstable orbit.

### 5.3.1. Bunching

Our method for finding orbits of twist maps is based on the observation that the configuration points of an orbit tend to cluster into groups as the nonlinearity is increased. Examples of this are shown in Figure 5.1a for the  $(m,n) = (8,21)$  minimizing orbit of the sawtooth map, and in Figure 5.1b for the standard map. These figures show the configuration points  $\{x(j), j = 0, 1, \dots, 20\}$ , where  $j$  denotes the order in the unit interval, as  $k$  is varied from 0 to 3.5.

At  $k = 0$  all orbits are simple rotations. For a minimizing orbit, as  $k \rightarrow 0^+$ ,  $x_t \rightarrow \{mt/n + 1/2n\}$  ( $\{X\}$  is the fractional part of  $X$ ). The rotation frequency determines the relationship between the spatial and temporal ordering of the points: upon iteration each point shifts  $m$  to the right, thus if  $x(j) = x_t$ , then  $x(j+m) = x_{t+1}$ . The ordering is preserved as  $k$  varies since the curves do not cross in Figure 5.1.

However, as  $k$  increases they tend to bunch together in groups. For  $k \approx 0.2$  for the sawtooth map, and  $k \approx 1.5$  for the standard map, some of the points cluster together in pairs, leaving thirteen well separated points or pairs of points. As  $k$  increases these clusters successively approach each other in pairs, until near  $k \approx 3$  there are five distinguishable groups, and for  $k \approx 10$  there are only three. In fact, as  $k \rightarrow \infty$  all the points approach  $x = 1/2$ , where the action is minimum. The way in which this bunching occurs is determined by the Farey sequence for the frequency  $m/n$ .



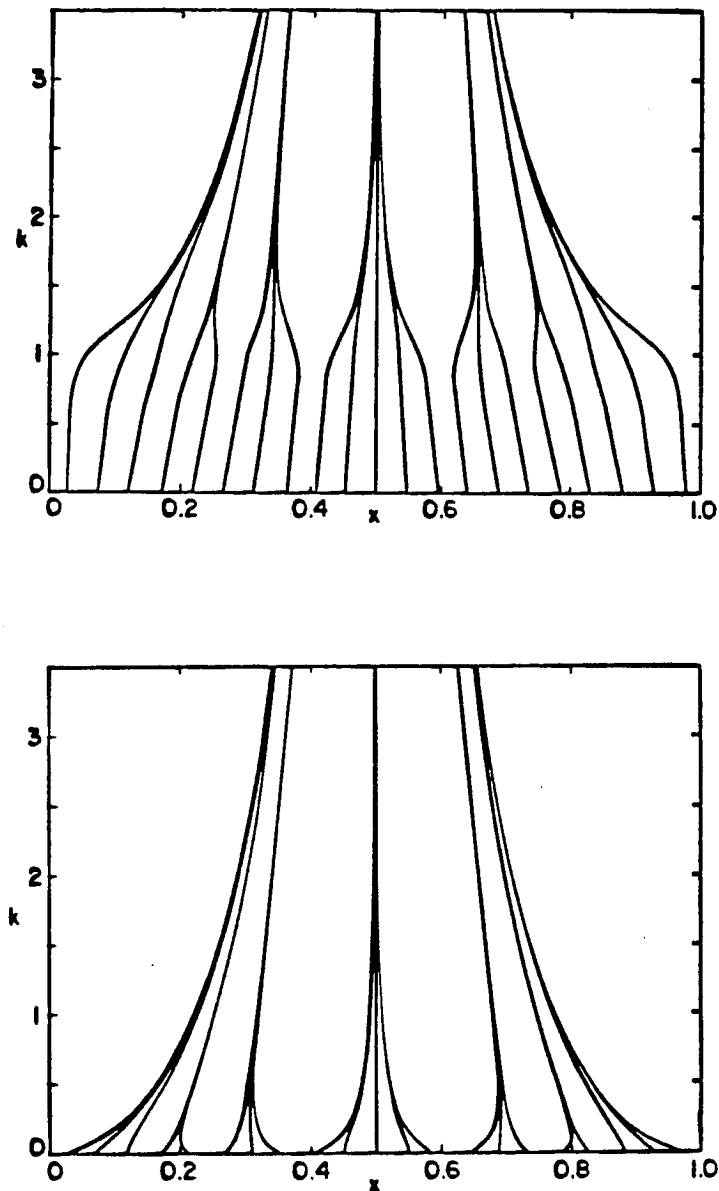


Figure 5.1. Configuration points of the (8,21) minimizing orbit as the parameter  $k$  varies. (1a) is for the sawtooth map and (1b) for the standard map. At  $k=0$  the points are equally spaced, but as  $k$  increases bunching towards the minimum of  $-V(x)$  occurs. These orbits were obtained by the extension method.

For a rational number on the Farey tree, the "mother" is defined to be the rational parent with the larger denominator; the "father" is defined to be the parent with the smaller denominator. Note that the mother is closer to its child than the father is. The parents can easily be obtained from the continued fraction expansion of their child:

$$m/n = [a_1, a_2, \dots, a_j] = a_1 + 1/(a_2 + 1/(\dots + 1/a_j)) \quad (5.3.1)$$

where the  $a_i$ 's are positive integers. With the convention that the final element  $a_j > 1$ , the continued fraction expansion for the mother is  $[a_1, a_2, \dots, a_{j-1}, a_{j-1}]$  and for the father is  $[a_1, a_2, \dots, a_{j-1}]$ .

The Farey sequence for a rational is its maternal lineage: thus for  $8/21$  one obtains the sequence  $\{1/2, 1/3, 2/5, 3/8, 5/13\}$ . In this case the Farey sequence is identical to the sequence of continued fraction convergents; more generally when the  $a_i \neq 1$ , "intermediates" as well as convergents are included in the Farey sequence. Each convergent in the Farey sequence gives a best approximation to  $m/n$ . For example,  $2/5$  is the nearest rational to  $8/21$  for any with denominator less than 8.

As is seen in Figure 5.1 the bunching of points on an orbit occurs with groups defined by successive levels of the Farey sequence (for even periods the cluster around  $x = 0.5$  must be arbitrarily separated into two groups).

The bunching leaves gaps in configuration space, and as  $k$  increases, the largest of these increases in size. For a minimizing orbit, this "principal" gap occurs around  $x = 0$  where the generating function has a maximum. Because the points on

the orbit are ordered, we know that the iterates of the endpoints of a gap are also gap endpoints. Thus we can follow the orbit of a gap, delineated by its endpoints. For the orbit in Figure 5.1, the image of the principal gap is the gap around  $x \approx 0.4$ , and its preimage is the gap around  $x \approx 0.6$ . These are both narrower than the principal gap. In fact the size of a gap decreases approximately exponentially with time away from the principal gap; the mean rate is the Lyapunov exponent of the orbit. This implies that strong bunching is only present when the orbit is strongly unstable, and is the reason that there is no bunching when  $k$  is small for the standard map. For the case of the sawtooth map, the local exponentiation rate is spatially constant, and the bunching can be analyzed exactly.

For one period of a Farey parent, the orbit of the principal gap is ordered according to the parent's frequency. For example, when  $k \approx 3$  in Figure 5.1b, the orbit of the principal gap follows the ordering of the  $2/5$  parent for  $-3 < t < 3$ , but for larger  $|t|$  the gap collapses to a small spatial scale, and the  $2/5$  ordering is not sufficient to describe the motion. Similarly, when  $k \approx 2$  the principal gap follows the  $3/8$  ordering for  $-5 < t < 4$ .

The bunching properties of unstable orbits will be exploited extensively below.

### 5.3.2. Ordered Periodic Orbits

The spatial ordering of points for an  $(m,n)$  orbit is defined by their positions in the unit interval  $[0,1)$  (in this section, configuration points are taken modulus one). We denote the configuration points in the unit interval by  $x(j)$  where  $j$  designates the spatial ordering:

$$0 \leq x(0) < x(1) < \dots < x(n-1) < 1 \quad (5.3.2)$$

It is observed that minimizing orbits have all points nonzero, and a minimax orbit has  $x(0) = 0$ . The principal gap in the orbit is the gap around  $x = 0$ . For a minimizing orbit, we focus on the endpoints of the principal gap which, since our convention is to put points in the interval  $[0,1)$ , are  $x(0)$  and  $x(n-1)$ . On the other hand, a minimax orbit has a point in the center of each gap, so we focus on the center point of the principal gap, which is  $x(0)$ . Choose the origin of time such that  $x_0 = x(0)$ .

There is an easy translation between time ordering and spatial ordering for an ordered orbit; it is based on the Farey sequence for the orbit frequency. Let  $p/r$  and  $q/l$  ( $p/r < q/l$ ) be the parents of  $m/n$ . Suppose  $x(j) = x_t$ , then its spatial neighbors are given by:

$$x(j-1) = x_{t+l}, \quad x(j+1) = x_{t+r}, \quad (5.3.3)$$

i.e. the right neighbor is obtained upon  $r$  iterations, and the left neighbor upon  $l$  iterations. With the above formula, we are able to find the endpoints of the largest gap by successive iteration from any point on the orbit. For example, for a minimizing orbit,

since  $x_0$  is the right endpoint of the principal gap,  $x_{\ell}$  is its left endpoint.

The extension method is based on iteration on the levels of the Farey sequence, and is given by the following algorithm:

- 1) Find the Farey sequence of  $m/n$ . Begin at a level which has a sufficiently small period, say  $\leq 5$ . Denote the frequency at this level by  $p/q$ .
- 2) Find an approximate configuration  $\{x_t, t = 0, 1, \dots, q-1\}$  at the lowest level. It is sufficient to choose a simple trial as discussed in section 2.3, and iterate Newton's method twice.
- 3) Proceed to the next Farey level with daughter rotation number  $p'/q'$ ; denote the denominators of the parents by  $r'$  and  $\ell'$ .
- 4) Obtain the the initial guess for the new configuration at this level. For a minimizing orbit,  $\{x'_t\}$ , is obtained from  $\{x_t\}$  first by setting the new endpoints of the principal gap equal to the old. This means that the largest gap in the new trial orbit is approximated by the largest gap in the old one. We then set the second largest gaps equal, and so forth. Since the gaps decrease approximately monotonically away from the largest, this is done by identifying temporal neighbors of the principal gap. Continue to do this until all  $q'$  points on the new orbit are assigned:

$$\begin{aligned}
 x'_t &= x_t \\
 x'_{\ell'-t} &= x_{\ell-t} & t = 0, 1, 2, \dots, [\ell'/2], \\
 \\
 x'_{q'-t} &= x_{q-t} \\
 x'_{\ell'+t} &= x_{\ell+t} & t = 1, 2, \dots, [(q'-\ell')/2], \quad (5.3.4)
 \end{aligned}$$

where  $[ ]$  indicates integer part of a number. Note there is overlap in the assignments for  $\ell'$  or  $q'-\ell'$  even. If this is the

case, the best strategy is to use the average of the two assignments. For a minimax orbit, the situation is simpler, since there is only one point to assign per gap:

$$\begin{aligned} x'_t &= x_t & t &= 0,1,2,\dots [q'/2] , \\ x'_{q'-t} &= x_{q-t} & t &= 1,2, \dots [q'/2] , \end{aligned} \quad (5.3.5)$$

- 5) Apply Newton's method with periodic boundary conditions to the current configuration.
- 6) If  $q' < n$ , remove the primes on the symbols to indicate a transition to the next Farey level, and go to 3).
- 7) Otherwise, this final configuration is used as the initial guess of the  $(m,n)$  orbit. It generally takes only one or two more iterations of Newton's method to obtain the  $(m,n)$  orbit to a given precision.

Several levels of orbit extension are shown for the minimizing  $(13,34)$  orbit in Figure 5.2 The top configuration is that of the  $2/5$  orbit; the assignment of the points for the next level  $(3/8)$  gives the second configuration. Upon iteration the points split apart, resulting in the third configuration. Below that is the result of the next two iterations, giving approximations to the  $(5,13)$  and  $(8,21)$  orbits. The final configuration shown is indistinguishable from the actual  $(8,21)$  orbit, and on the scale of the graph, from the  $(13,34)$  orbit as well.

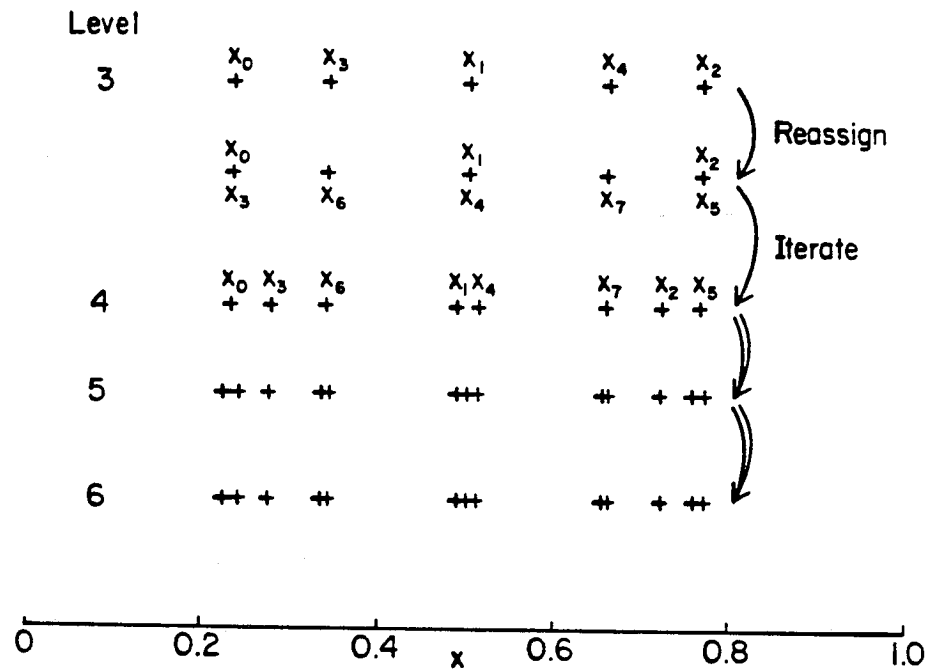


Figure 5.2. Four levels of the orbit extension method for an orbit of the standard map at  $k = 2.0$ . The uppermost configuration is the exact configuration of the (2,5) orbit. The next shows the reassignment of the points to give the (3,8) ordering. The third is the result of a single Newton iteration on this trial. The last two show the results of iterating the (5,13) and the (8,21) orderings. The final orbit has a residue of order  $10^{10}$ , and higher levels cannot be distinguished on this scale.

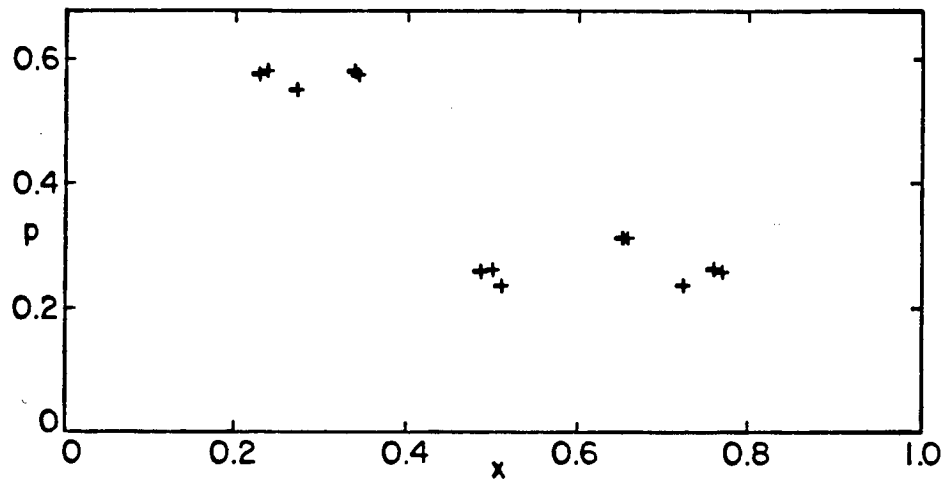


Figure 5.3. Phase space plot of the final configuration in Fig. 5.2. Note that configuration space bunching implies phase space bunching.



The orbit extension method has the advantage that it only needs the approximating configurations of small period orbits. Newton iteration is applied only once to the intermediate configurations because the general structure of the bunching is only qualitative. It is quite efficient considering that most of the intermediate configurations have very short periods. So the effective time cost of all the intermediate steps is just that for two or three Newton iterations of the (m,n) orbit.

Theoretically, there is no limitation to the period of the orbit one wants to find. Numerically, because of the exponential divergence of the elements of the Hessian matrix with the period of the orbit, very long orbits cannot be found due to computer overflow. We have been able to find periodic orbits up to residue  $10^{25}$  using a double precision (128 bit) algorithm on the ICL 2988 computer. This is far better than that needed for physically motivated studies because the smallest gap on an orbit scales as the inverse of the residue; so orbits with residue larger than the inverse of the precision are numerically indistinguishable. Similarly, quantities such as the Lyapunov exponent, and the flux also have errors of order the inverse of the residue.

Typically the orbit extension method converges with less than half the computational time than that required for Newton's method on simple trial configurations of the full orbit. However, the extension method can be slower when the orbit has small residue: the analysis of bunching is based on the existence of a positive Lyapunov exponent, which certainly breaks down for small residue orbits. Thus, the orbit extension is useful when the absolute value of the residue of the orbit is larger than one. By contrast, the most efficient method for smaller

residue is to iterate the map for one period, forcing the trajectory to return to its origin by a two dimensional secant method.

### 5.3.3. Principal Heteroclinic Orbits

When we study transport, we are primarily interested in quantities such as area and flux. The flux out of a resonance is determined by the orbits homoclinic to the resonance. The flux from one resonance to another is determined by the corresponding heteroclinic orbits.

A type  $(m,n)$  principal homoclinic orbit is an ordered orbit which approaches a periodic  $(m,n)$  orbit in both directions of time. Let  $h_0$  represent a point on such an orbit in the gap  $(x(0),x(1))$  of the  $(m,n)$  orbit. Ordering implies that subsequent points on the homoclinic orbit are in the corresponding gaps of the periodic orbit, and furthermore that  $h_{j_n} \rightarrow x(1)$  and  $h_{-j_n} \rightarrow x(0)$  as  $j \rightarrow \infty$ . Such an orbit is called right (left)-going and denoted with superscript "+" ("-") if the limit is approached as  $j \rightarrow \infty$  ( $j \rightarrow -\infty$ ). The existence of such orbits is proved by Aubry [Aubry and Le Daeron, 1983]. There are actually two such orbits of each type: they are obtained at minimum and minimax values of the action, and will be denoted by  $M$  and  $S$ , respectively (see Figure 5.4). Choose  $t=0$  such that  $S^\pm_0 = 0$  and  $M^\pm_0$  is the leftmost point on the minimizing orbit. Ordering then implies

$$\begin{aligned} M^+_{t-n} &< S^+_t < M^+_t, \\ M^-_{t+n} &< S^-_t < M^-_t. \end{aligned} \tag{5.3.6}$$

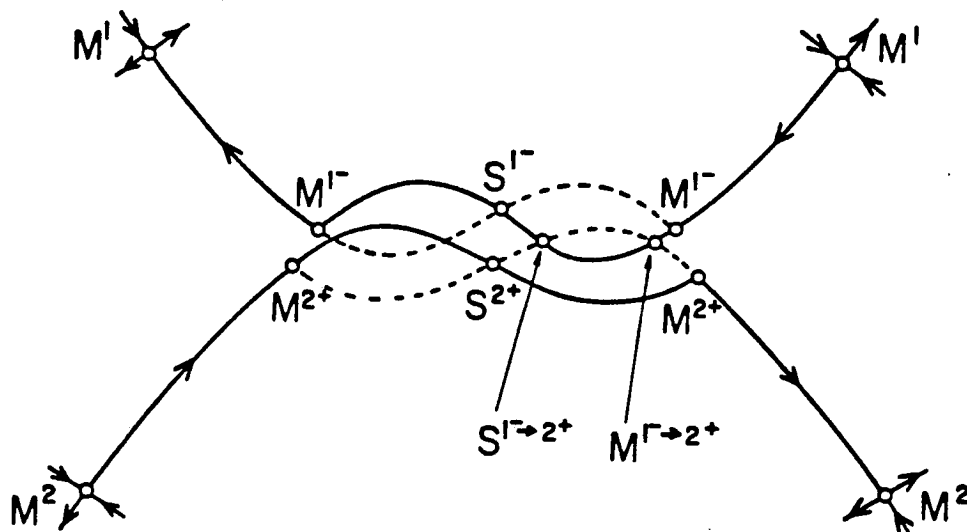


Figure 5.4. Sketch of points on orbits homoclinic to and heteroclinic between minimizing periodic orbits  $M^1$  and  $M^2$ . The solid lines are segments of stable and unstable manifolds forming the lower and upper boundaries of the  $M^1$  and  $M^2$  resonances, respectively. The dashed lines are extensions of the right-going manifolds forming the turnstiles. Points on principal heteroclinic orbits,  $S^{1^- \rightarrow 2^+}$  and  $M^{1^- \rightarrow 2^+}$ , occur at the intersection of turnstile boundaries. Past points of these orbits fall on the right-going unstable manifold of  $M^1$ , and future points on the right-going stable manifold of  $M^2$ .

A heteroclinic orbit denoted by  $(m_1, n_1) \rightarrow (m_2, n_2)$ , approaches an  $(m_1, n_1)$  orbit in the forward time direction and an  $(m_2, n_2)$  orbit in the backward time direction. Such orbits clearly do not exist when there is an invariant circle between the two periodic orbits; however, Birkhoff has shown that non-existence of an invariant circle is sufficient to guarantee the existence of an orbit which goes from some neighborhood of  $(m_1, n_1)$  to some neighborhood of  $(m_2, n_2)$  [Birkhoff 1950, p.111]. Furthermore, Mather has proved that heteroclinic orbits do exist in this case. Since all absolutely minimizing configurations are either periodic or quasiperiodic, a principal heteroclinic orbit cannot be absolutely minimizing; however, Mather's result implies the existence of locally minimizing and minimax orbits [Mather, personal communication, 1985; Hall 1987].

There are possibly four types of  $(m_1, n_1) \rightarrow (m_2, n_2)$  orbits, depending on whether the orbit leaves  $(m_1, n_1)$  on a right- or left-going unstable manifold and whether it approaches  $(m_2, n_2)$  on a right- or left-going stable manifold. Using the superscripts + and - to indicate right- or left-going, respectively, then the orbit denoted, for example, by  $(m_1, n_1)^- \rightarrow (m_2, n_2)^+$  leaves  $(m_1, n_1)$  on a left-going unstable manifold, and approaches  $(m_2, n_2)$  on a right-going stable manifold. At each point on a heteroclinic orbit the unstable manifold of  $(m_2, n_2)$  intersects the stable manifold of  $(m_2, n_2)$  as shown in Figure 5.4. Corresponding to each  $(m_1, n_1) \rightarrow (m_2, n_2)$  orbit there is also a  $(m_2, n_2) \rightarrow (m_1, n_1)$  orbit; the latter occur at the unlabeled stable-unstable manifold intersections in Figure 5.4. By Eq. (5.2.3) these two orbits are equivalent upon reflection about the origin: they are not symmetric themselves. As we will discuss below,

these heteroclinic orbits characterize the transport from  $(m_1, n_1)$  to  $(m_2, n_2)$  and vice versa.

A principal heteroclinic orbit is a heteroclinic orbit which is partially ordered in the following sense. Consider for example  $(m_1, n_1)^- \rightarrow (m_2, n_2)^+$  and suppose  $m_1/n_1 > m_2/n_2$ . Let  $M^{1-}$  and  $S^{1-}$  be the minimizing and minimax  $(m_1, n_1)^-$  orbit, etc. Let  $x_0$  be the leftmost point on the heteroclinic orbit:  $S^{2+}_0, S^{1-}_0 < x_0 < M^{2+}_0, M^{1-}_0$  (by convention  $S^{2+}_0 = S^{1-}_0 = 0$ ). The partial ordering condition is:

$$\begin{aligned} S^{2+t} < x_t < M^{2+t}, t > 0 \\ S^{1-t} < x_t < M^{1-t}, t < 0 \end{aligned} \quad (5.3.7)$$

Similarly, the ordering conditions can be easily given for the other three types of principal heteroclinic orbits.

The non-existence of invariant circles is not sufficient for the existence of principal heteroclinic orbits. These exist only when the turnstiles of the two resonances overlap.

Since a homoclinic orbit lies on the unstable and stable manifolds of the corresponding minimizing periodic orbit, its behavior in the neighborhood of the periodic orbit is governed by the latter's characteristic exponent. Indeed, most points of a homoclinic orbit typically lie very near the corresponding periodic orbit. Furthermore, the ordering condition (5.3.7) implies that a heteroclinic orbit is well approximated by the corresponding homoclinic orbits near the periodic points. This suggests the use of segments of homoclinic orbits as the initial guess for the corresponding heteroclinic orbit. This idea corresponds to the method used by Mather in his existence proof.

The method for finding the principal homoclinic orbits is reviewed in [MacKay, Meiss and Percival 1987]. Suppose we want to find the homoclinic orbits for resonance  $(m,n)$ . We first find the parents  $(p,r)$  and  $(q,l)$  of  $m/n$  ( $p/r < q/l$ ). The left-going principal homoclinic orbit is approximated by the  $(mj+p, nj+r)$  orbit and the right going one by the  $(mj+q, nj+l)$  orbit for  $j$  large enough. These periodic orbits can be found by the extension method, as before.

The orbit extension method to find a minimizing  $(m_1, n_1)^- \rightarrow (m_2, n_2)^+$  orbit consists of the following steps:

- 1) Find the minimizing  $(m_2, n_2)^+$  orbit  $\{x^{2+}\}$  and the minimizing  $(m_1, n_1)^-$  orbit  $\{x^{1-}\}$ .
- 2) Position the time origin of the homoclinic orbits at their leftmost points, take a segment of the  $(m_1, n_1)^-$  orbit from some time  $-T$ , up to time zero, and a segment of  $(m_2, n_2)^+$  orbit from time zero up to  $T$ . The initial guess for the  $(m_1, n_1)^- \rightarrow (m_2, n_2)^+$  heteroclinic orbit consists of the union of these two segments, with the choice of one of the two points at  $t = 0$ :

$$\{x^{1-T}, \dots, x^{1-1}, \max(x^{1-0}, x^{2+0}), x^{2+1}, \dots, x^{2+T}\} \quad (5.3.8)$$

- 3) Apply Newton's method with asymptotically periodic boundary conditions once.
- 4) Append another segment of the minimizing  $(m_1, n_1)^-$  orbit from  $-2T$  to  $-T-1$ , to the negative time end and the minimizing  $(m_2, n_2)^+$  orbit from  $T+1$  to  $2T$  to the positive time end of the resulting orbit.

- 5) Repeat 3) and 4) until the orbit is sufficiently long to achieve the desired accuracy.
- 6) This final configuration is used as the initial guess for the  $(m_1, n_1)^- \rightarrow (m_2, n_2)^+$  orbit; however, usually only one or two Newton iterations are required for convergence at this stage.

Similar prescriptions can be developed to find the other principal heteroclinic orbits.

We find numerically that the principal heteroclinic orbits are locally minimizing or minimax. These two orbits are labeled as  $M^{1^- \rightarrow 2^+}$  and  $S^{1^- \rightarrow 2^+}$  respectively in Figure 5.4. The time used to find a heteroclinic orbit is of the same order as the time used to find a periodic orbit of the same length. In cases where the principal heteroclinic orbits do not exist, Newton's method does not converge.

We have used this method to calculate the flux function from one resonance to another. This is discussed in section 3.4.3.



**Chapter 6**  
**Summaries and Conclusions**

In summary, chaotic transport from one part of phase space to another can be described as successive transitions between different resonance islands. Resonances are regions delineated by pieces of the stable and unstable manifolds of the hyperbolic periodic orbits. Resonances give a complete partition of phase space in the supercritical regime where there are no rotational invariant circles: they cover the whole phase space except for a measure zero set. Resonance boundaries are called partial separatrices. When an orbit enters a resonance island, it follows the rotational order of that resonance till it reaches the upper or lower exiting turnstile, where it makes a transition to another resonance. If the turnstiles of two resonances overlap, then there is a direct transition between these resonances; otherwise, a transition requires successive entering and exiting intermediate resonances. Resonances, partial separatrices and turnstiles provides a framework of our picture of chaotic motions.

Resonances form a natural partition of the phase space in a statistical description of transport. They are chosen as coarse grained states in the Markov chain. The overlap of turnstiles of resonances specify all the transitions between these resonances. Assuming a complete "randomization" under the dynamics, the transition rate from one resonance to another is given by the overlapping turnstile area, the flux, divided by the area of that resonance. Resonances form a countable set, their areas decrease exponentially with their period. These properties make them particularly suitable for a finite state approximation of the model in an actual calculation of transport rates. One expects rapid convergence in any reasonable truncation scheme.

The statistical assumptions made in the Markov model require careful verification. Numerical studies show that there are basically three regimes: In the strongly chaotic regime, the mixing rate described by the Lyapunov exponent is fast enough to smooth out any nonuniformities in the distribution, this corresponds to the occurrence of horseshoes (or generalized turnstiles) in the resonance islands. When the mixing rate is slower, transport rates deviate from the predictions of the Markov model. However, they are not very far off and can be used to give an economic estimate of the actual transport rates. The model breaks down for nearly integrable systems.

In order to estimate transport rates for real systems from the Markov model, we need to calculate flux and resonance areas. The action formulation allows these calculation through determination of periodic, homoclinic, and heteroclinic orbits. However, these orbits are typically highly unstable; the sensitive dependence on initial conditions is a key impediment to devising stable numerical prescriptions. A stable and efficient method — the orbit extension method, is developed for finding both unstable ordered periodic orbits and the principal heteroclinic orbits between two resonances.

A major challenge in the future is to generalize the idea of resonances, partial barriers and turnstiles to higher dimensional Hamiltonian systems. Some results in this direction are discussed in Kook and Meiss [1989a] and MacKay, Meiss and Stark [1989]. There is evidence for resonance zones in four dimensional symplectic maps; however, we do not know how to form the resonance boundary. Certainly, the stable and unstable sets of the hyperbolic periodic orbits play an important role. Similarly it is unclear how to construct higher dimensional partial barriers and whether the symplectic flux, the escaping

symplectic area, bears any relation to the actual escaping dynamics.

Another problem raised by this dissertation is to study the transport in nearly integrable Hamiltonian systems. As suggested by our study, the Markov model fails to account for the slow drift in the tiny channels of the stochastic web, since the mixing assumption is not easily satisfied there. However, it does appear that the characteristic time in the web scales inversely with the flux. What is the mechanism for this?

## References

R. Abraham and J.E. Marsden, 1976, Foundations of Mechanics, 2nd Edition, Benjamin-Cummings.

S.B. Angenent, 1984, The Periodic Orbits of an Area-Preserving Twist Map, Leiden preprint.

H. Aref, 1984, Stirring by Chaotic Advection, J. Fluid Mech. **143**, 1-21.

V.I. Arnol'd, 1973, Ordinary Differential Equations, MIT Press.

V.I. Arnol'd, 1978, Mathematical Methods of Classical Mechanics, Graduate Texts in Mathematics 60, Springer-Verlag.

Vi.I. Arnol'd, 1983, Geometrical Methods in the Theory of Ordinary Differential Equations, Springer-Verlag.

V.I. Arnol'd (ed), 1986, Encyclopaedia of Mathematical Sciences vol. 3, Dynamical Systems III, Spinger-Verlag.

V.I. Arnol'd and A. Avez, 1968, Ergodic Problems of Classical Mechanics, Benjamin.

S. Aubry, 1982, The Devil's Staircase Transformation in Incommensurate Lattices, Lect. Notes in Math. **925**, 221-245.

S. Aubry, 1983a, The Twist Map, the Extended Frenkel-Kontorova Model and the Devil's Staircase, Physica **7D**, 240-258.

S. Aubry, 1983b, Exact Models with A Complete Devil's Staircase, *J. Phys. C*, **16**, 2497-2508.

S. Aubry and P.Y. Le Daeron, 1983, The Discrete Frenkel-Kontorova Model and Its Extensions, *Physica* **8D**, 381-422.

S. Aubry, P.Y. Le Daeron and G. Andre, 1982, Classical Ground States of a One-Dimensional Model for Incommensurate Structures, Preprint, CEN Saclay.

D. Bensimon and L.P. Kadanoff, 1984, Extended Chaos and Disappearance of KAM trajectories, *Physica* **13D**, 82-89.

M.V. Berry, 1978, Regular and Irregular Motion, in Topics in Nonlinear Dynamics, (S. Jorna, ed), AIP Conf. Proc. **45**, 16-120.

N. Bird and F. Vivaldi, 1988, Periodic Orbits of the Sawtooth Maps, *Physica* **30D**, 164.

G.D. Birkhoff, 1950, Collected Mathematical Papers, Vols. 1,2, and 3, AM. Math. Soc., Providence.

T. Bohr and D. Rand, 1987, The Entropy Function for the Characteristic Exponents, *Physica* **25D**, 387-389.

S. Bullet, 1986, Invariant Circles for the Piecewise Linear Standard map, *Commun. Math. Phys.* **107**, 241.

J.R. Cary and R.G. Littlejohn, 1983, Noncanonical Hamiltonian Mechanics and Its application to Magnetic Field Line Flow, *Ann. Phys.* **151**, 1-34.

J.R. Cary and J.D. Meiss, 1981, Rigorously Diffusive, Deterministic Map, Phys. Rev. **A24**, 2664-2668.

P.J. Channell and C. Scovel, 1988, Symplectic Integration of Hamiltonian Systems, submitted to Nonlinearity.

S.R. Channon and L.J. Lebowitz, 1980, Numerical Experiments in Stochasticity and Heteroclinic Oscillation, in Nonlinear Dynamics, Ann. New York Acad. Sci. **357**, 108-118, R.H.G. Helleman, ed.

Q. Chen, 1987, Area as a Devil's Staircase in Twist Maps, Phys. Lett. **123**, 444-450.

Q. Chen, J.D. Meiss, and I.C. Percival, 1987, Orbit Extension Method for Finding Unstable Orbits, Physica **29D**, 143-154.

Q. Chen and B. Mestel, 1988, Taylor-Chirikov Map Package, Computer Physics Communications **51**, 463-476.

Q. Chen and J.D. Meiss, 1989, Flux, Resonances and the Devil's Staircase for the Sawtooth Map, Nonlinearity **2**, 347-356.

Qi Chen, R.S. MacKay, and J.D. Meiss, 1989, Cantori for Symplectic Maps, submitted to Nonlinearity.

B.V. Chirikov, 1979, A Universal Instability of Many-Dimensional Oscillator Systems, Phys. Reports **52**, 262-379.

B.V. Chirikov and D.L. Shepelyanski, 1984, Correlation Properties of Dynamical Chaos in Hamiltonian Systems, Physica **13D**, 395-400.

I. Dana and S. Fishman, 1985, Diffusion in the Standard Map, *Physica* **17D**, 63-74.

I. Dana and W.P. Reinhardt, 1987, Adiabatic Invariance in the Standard Map, *Physica* **28D**, 115-142.

I. Dana, N. Murray, and I.C. Percival, 1989, Resonances and Diffusion in Periodic Hamiltonian Maps, *Phys. Rev. Lett.* **62**, 233-235.

R. Devaney, 1986, *An Introduction to Chaotic Dynamical Systems*, Benjamin/Cummings.

M.J. Davis, 1985, Bottlenecks to Intramolecular Energy Transfer and the Calculation of Relaxation Rates, *J. Chem. Phys.* **83**, 1016-1035.

R. Douady, 1982, Une Démonstration Directe de L'équivalence des Théorèmes de Tores Invariants pour Difféomorphismes et Champs des Vecteurs, *C.R. Acad. Sci., Paris, Sér. I* **295**, 201-204.

Feng Kang, 1986, Different Schemes for Hamiltonian Formalism and Symplectic Geometry, *J. of Computational Math*, **4**, 279.

C. Grebogi, E. Ott and J. Yorke, 1988, Unstable Periodic Orbits and the Dimensions of Multifractal Chaotic Attractor, *Phys. Rev.* **A37**, 1711-1724.

J. Greene, 1979, A Method for Computing the Stochastic Transition, *J. Math. Phys.* **20**, 1183-1201.



J. Greene, R.S. MacKay and J. Stark, 1986, Boundary Circles for Area-preserving Maps, *Physica* **21D**, 267-295.

J. Guckenheimer and P. Holmes, 1983, Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields, *Applied Mathematical Sciences* 42, Springer-Verlag.

G.R. Hall, 1987, A Topological Version of a Theorem of Mather's on Shadowing in Monotone Twist Maps," preprint.

J.D. Hanson, J.R. Cary and J.D. Meiss, 1985, Algebraic Decay in Self-Similar Markov Chains, *J. Stat. Phys.* **39**, 327-345.

G.H. Hardy and E.M. Wright, 1979, *Introduction to the Theory of Numbers*, 5th Edition, Clarendon Press.

R.D. Hazeltine and J.D. Meiss, 1985, Shear Alfvén Dynamics of Toroidally Confined Plasmas, *Phys. Reports* **121**, 1-164.

W. Horton, 1986, Onset of Stochasticity and the Diffusion Approximation in Drift Waves, *Plas. Phys. and Cont. Fusion*, **27**, 937-956.

J.E. Howard and R.S. MacKay, 1987, Linear Stability of Symplectic Maps, *J. Math. Phys.* **28**, 1036-1051.

G. Hsu, E. Ott and C. Grebogi, 1988, Strange Saddles and the Dimensions of Their Invariant Manifold, *Phys. Lett.* **127A**, 199-204.

H. Kantz and P. Grassberg, 1985, Repellers, Semi-Attractors, and Long Lived Chaotic Transients, *Physica* **17D**, 75-86.

C.F.F. Karney, 1983, Long-time Correlations in the Stochastic Regime, *Physica* **8D**, 360-380.

A. Katok, 1983, Periodic Orbits and Quasiperiodic Orbits for Twist Maps, in *Dynamical Systems and Chaos*, L. Darrido, ed. Springer Lect. Notes in Phys. **179**, 47-65.

A.Y. Khinchin, 1964, *Continued Fractions*, University of Chicago Press, Chicago.

H.T. Kook and J.D. Meiss, 1989a, Periodic Orbits for Reversible, Symplectic Mappings, *Physica* **35D**, 65-89.

H.T. Kook and J.D. Meiss, 1989b, Applications of Newton's Method to Lagrangian Dynamical Systems, *Physica* **D**, in Press.

L.D. Landau and E.M. Lefshitz, 1976, *Mechanics*, 3rd ed., Course of Theoretical Physics, vol.1, Pergamon Press, Oxford.

O.E. Langford, 1973, Introduction to the Mathematical Theory of Dynamical Systems, in *Chaotic Behavior of Deterministic Systems*, North Holland, Amsterdam.

W. Li and P. Bak, 1986, Fractal Dimension of Cantori, *Phys. Rev. Lett.* **57**, 655-658.

A.J. Lichtenberg and M.A. Lieberman, 1983, *Regular and Stochastic Motion*, Applied Mathematical Sciences 38, Springer-Verlag.

R.G. Littlejohn, 1983, Variational Principle of Guiding Center Motion, *J. Plasma Phys.* **29**, 121-125.

R.S. MacKay, 1982, Renormalization in Area-Preserving Maps, Ph.D. Thesis, Princeton.

R.S. MacKay, 1983, A Renormalization Approach to Invariant Circles in Area-Preserving Maps, *Physica*, **7D**, 283-300.

R.S. MacKay, 1986a, Stability of Equilibria of Hamiltonian Systems, in *Nonlinear Phenomena and Chaos*, S. Sarkar (ed.), Adam Hilger, Bristol, 254-270.

R.S. MacKay, 1986b, Transition to Chaos for Area-Preserving Maps, in *Nonlinear Dynamics Aspects of Particle Accelerators*, J.M. Jowett, M. Month and S. Turner (eds.), Springer-Verlag Lecture Notes in Physics **247**, 390-454.

R.S. MacKay, 1987, Hyperbolic Cantori have dimension zero, *J. Phys. A: Math Gen.* **20**, L559-561.

R.S. MacKay and J.D. Meiss, 1983, Linear Stability of Periodic Orbits in Lagrangian Systems, *Phys. Lett. A*, **98**, 92-94.

R.S. MacKay and J.D. Meiss, 1986, Flux and Differences in Action for Continuous Time Hamiltonian Systems, *J. Phys. A Letters* **19**, 225-229.

R.S. MacKay and J.D. Meiss (eds.), 1987, *Hamiltonian Dynamical Systems*, Adam Hilgar (Bristol).

R.S. MacKay, J.D. Meiss, and I.C. Percival, 1984, Transport in Hamiltonian Systems, *Physica* **13D**, 55-81.

R.S. MacKay, J.D. Meiss and I.C. Percival, 1987, Resonances in area-preserving maps, *Physica* **27D**, 1-20.

R.S. MacKay, J.D. Meiss, and J. Stark, 1989, Converse KAM Theory for Symplectic Twist Maps, preprint.

R.S. MacKay and I.C. Percival, 1985, Converse KAM: Theory and Practice, Commun. Math. Phys. **98**, 469-512.

R.S. MacKay and J. Stark, 1985, Lectures on Orbits of Minimal Action for Area-Preserving Maps, University of Warwick Preprint.

J.N. Mather, 1982, Existence of Quasi-Periodic Orbits for Twist Homeomorphisms of Annulus, Topology, **21**, 457-467.

J.D. Meiss, 1986a, Class Renormalization: Islands around Islands, Phys. Rev. A **34**, 2375-2383.

J.D. Meiss, 1986b, Transport Near the Onset of Stochasticity, Part. Accel. **19**, 9-24.

J.D. Meiss, J.R. Cary, D.F. Escande, R.S. MacKay, I.C. Percival and J.L. Tennyson, 1985, Dynamical Theory of Anomalous Particle Transport, in Plasma Physics and Controlled Nuclear Fusion Research 1984, IAEA, Vienna, **3**, 441-448.

J.D. Meiss and E.Ott, 1986, Markov Tree Model of Transport in Area Preserving Maps, Physica **20D**, 387-402.

B. Mestel and I.C.Percival, 1987, Newton method for highly unstable orbits, Physica **24D**, 172.

P.J. Morrison, 1982, Poisson Brackets for Fluids and Plasmas, in Mathematical Methods in Hydrodynamics and Integrability in

Dynamical Systems, M. Tabor and Y.M. Treve, eds. Am. Inst. Phys. Conf. Proc. **88**, 13-46.

P.J. Morrison and J.M. Greene, 1980, Non-Canonical Hamiltonian Density Formulation of Hydrodynamics and Ideal Magnetohydrodynamics, Phys. Rev. Lett. **45**, 790-794.

P.J. Olver, 1988, Applications of Lie Groups to Differential Equations, Springer-Verlag, New York.

I.C. Percival, 1980, Variational Principles for Invariant Tori and Cantori, in Non-Linear Dynamics and Beam-Beam Interaction, M. Month and J.C. Herrera, eds. , Am. Inst. Phys. Conf. Proc. **57**, 310-320.

I.C. Percival and F. Vivaldi, 1987a, Arithmetical Properties of Strongly Chaotic Motions, Physica **25D**, 105.

I.C. Percival and F. Vivaldi, 1987b, A Linear Code for the Sawtooth and Cat Maps, Physica **27D**, 373.

Ya.B. Pesin, 1977, Lyapunov Characteristic Exponents and the Smooth Ergodic Theory, Russ. Math. Surveys **32:4**, 55-144.

M. Peyrard and S. Aubry, 1983, Critical Behavior at the Transition in the Discrete Frenkel-Kontrova Model, J. Phys. C, **16**, 1593-1608.

W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, 1986, Numerical Recipes, Cambridge University Press.

W. Rudin, 1987, Real and Complex Analysis, 3rd edition, McGraw Hill Book Company.

H.J. Schnellhuber, H. Urbschat and A. Block, 1986, Calculation of Cantori, Phys. Rev. **33A**, 2856.

S.J. Shenker and L.P.Kadanoff, 1982, Critical Behavior of a KAM Surface: I Empirical Results, J. Stat. Phys. **27**, 631-656.

D.K. Umberger and J.D. Farmer, 1985, Fat Fractals on the Energy Surface, Phys. Rev. Lett. **55**, 661-664.

## Vita

Qi Keith Chen, was born in Shanghai, China on October 7, 1963, the son of Zhu Ren Chen and Xiao Er Xu. After completing his work at Fudan High School affiliated to Fudan University, in 1980, he entered Fudan University, Shanghai, China. He received the degree of Bachelor of Science in Physics in July, 1984. In the Fall of 1984, he entered The Graduate School of The University of Texas at Austin. He was employed as a Teaching Assistant in Physics from September 1984 to January 1986. He has been employed as a Research Assistant at the Institute for Fusion Studies since January 1986.

Permanent Address: 1670-12 Kongjiang Road, Suite 504  
Shanghai, People's Republic of China

This dissertation was typeset by the author.