2001

A Study of Stellar Orbits in a Rotating, Gaseous Bar.

Eric Ian Barnes
Louisiana State University and Agricultural & Mechanical College

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

A STUDY OF STELLAR ORBITS IN A ROTATING, GASEOUS BAR

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Physics and Astronomy

by
Eric Barnes
B.S., Northeast Missouri State University, Kirksville, 1995
M.S., Louisiana State University and A&M College, Baton Rouge, 1998
May 2001

ACKNOWLEDGEMENTS

There is one person to whom I owe a large debt of gratitude. Dr. Joel Tohline has been my advisor and mentor for the last three years and it was he who had the notion to investigate the system presented here. Somehow, he could extract useful information from my often wildly rambling reports. I have to agree with a recent referee's report Joel received that stated, "This Tohline guy is impressive." Along with Joel, Dr. Dana Browne and Dr. Juhan Frank have given me many helpful hints and suggestions for which I am grateful.

I must also thank my family: Mom, Dad, Somer, Corey, Uncle Fred, and Auntie Bonnie Darling Dear. Special thanks to my grandparents. Without their support and encouragement I never would have made it this far.

Graduate school is bearable if you have friends like I have had. The Nation 75 (Officials); Zach Allen, Karl Misselt (The Sweat), and Jason Elmore, have my eternal gratitude for keeping me from doing work. On the other hand, I would also like to thank the rest of the "Hydro" group: John Cazes, Howie Cohl, and Patrick Motl, because they actually *helped* me do my work. And last, but not least, muchas gracias to the rest of the physics crew: Reggie Madjoe, Erik Blaufuss, Phil Walsh, and Siong Heng. Drinkies for all my mates!

This work has been supported, in part, by funding from the U.S. National Science Foundation through grant AST 99-87344, from the Louisiana Board of Regents, LEQSF, through grant (1996-01)-GF-08, and from NASA/LaSPACE

through grant NGT5-40035. This work was also supported, in part, by grants of high-performance computing time through NPACI machines at the San Diego Supercomputing Center.

The Digitized Sky Surveys were produced at the Space Telescope Science Institute under U.S. Government grant NAG W-2166. The images taken from these surveys (Figures 1.1 and 1.2) are based on photographic data obtained using the Oschin Schmidt Telescope on Palomar Mountain and the UK Schmidt Telescope. The plates were processed into the present compressed digital form with the permission of these institutions.

The National Geographic Society – Palomar Observatory Sky Atlas (POSS-I) was made by the California Institute of Technology with grants from the National Geographic Society.

The Oschin Schmidt Telescope is operated by the California Institute of Technology and Palomar Observatory.

The UK Schmidt Telescope was operated by the Royal Observatory Edinburgh, with funding from the UK Science and Engineering Research Council (later the UK Particle Physics and Astronomy Research Council), until 1988 June, and thereafter by the Anglo-Australian Observatory. The blue plates of the southern Sky Atlas and its Equatorial Extension (together known as the SERC-J), as well as the Equatorial Red (ER), and the Second Epoch [red] survey (SES) were all taken with the UK Schmidt.

Supplemental funding for sky-survey work at the STScI is provided by the European Southern Observatory.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

In an effort to better understand the formation and evolution of barred galaxies, the properties of orbits in the effective potential of one specific model of a rapidly rotating, steady-state gas-dynamical bar that has been constructed via a self-consistent hydrodynamical simulation have been examined. This bar is used to test the following idea. If primordial galaxies evolve to a rapidly rotating, bar-like configuration before a significant amount of star formation has taken place, and then stars form from the gas that makes up the bar, the initial stellar distribution function should be much different than those used in previous bar formation studies.

As a first step towards understanding such a distribution function, orbits in the two-dimensional, equatorial slice of the above mentioned bar are studied. Orbits that result from a systematic search of initial conditions are compared to orbits that have initial conditions determined by the Restriction Hypothesis. The Restriction Hypothesis is the implementation of the idea that stars are forming from the gas that makes up the bar. Specifically, the initial velocities of Restriction Hypothesis orbits are set equal to the known gas velocities at the points of formation. It is found that Restriction Hypothesis orbits are a subset of all possible orbits and that the most important regular orbit family has a "bow tie" shape. These orbits are vastly different than the main family of orbits previously thought to sustain bar shaped distributions.

Extending the Restriction Hypothesis to the fully three-dimensional bar potential, a method of characterizing the resulting orbits based on the number of conserved quantities respected by the orbits has been utilized. These conserved quantities are known as integrals of motion and are related to the number of dimensions in which a phase space orbit exists. This technique is found to be robust and provides a straightforward way of categorizing orbits. Using this technique, it has been determined that a large percentage of examined three-dimensional Restriction Hypothesis orbits respect at least two integrals of motion.

1. INTRODUCTION

1.1 Galactic Astronomy 101

Prior to the early twentieth century, many astronomers, most notably Messier, observed what were called *nebulae*, meaning "clouds". These nebulae were diffuse sources of light that were cataloged so that they would not be confused with transient objects like comets. The nature of these nebulae was argued about for many years, reaching a head in the Shapley-Curtis debate of 1920. Shapley argued that the nebulae are associated with our own Milky Way galaxy, while Curtis took the position that nebulae are actually galaxies themselves and are extragalactic entities. In a sense, both were correct. A subset of the nebulae are associated with the Milky Way. These are now known to be either star clusters or clouds of dust and gas surrounding stars. The extragalactic nature of the remaining nebulae was finally settled when Hubble measured the distances to several nebulae using variable stars contained within them (Hubble 1925). The distances Hubble measured provided strong proof that these nebulae exist outside the Milky Way.

With the realization that there are galaxies external to our own, astronomers began to study these objects (as well as the Milky Way) in earnest and discovered that galaxies are complex systems consisting of stars, gas, and dust. They come in a variety of forms but are usually categorized into four basic morphologies defined by Hubble: elliptical, lenticular, spiral, and irregular (Hubble 1936). Ellipticals (E) range in projected shape (that is, the

two-dimensional outline of their true three-dimensional shape) from circular (denoted by E0) to elongated, cigar-like, forms (denoted by E7). It is commonly believed that many E galaxies are truly triaxial objects. Galaxies classified as 'spiral' display patterns of stars that lie along spiral "arms". A representative "normal" spiral is NGC 2997, shown here in Fig. 1.1. Between the E and spiral galaxy types lie the lenticular (S0) galaxies. These galaxies appear to have a basic disk shape, but do not show evidence of spiral structure. Spirals are divided into "normal" (like NGC 2997) and "barred" galaxies. Figure 1.2 is a picture of a typical barred spiral galaxy, NGC 1300. The original Hubble "tuning fork" diagram (Fig. 1.3[1]) illustrates the bifurcation of normal (S) and barred (SB) spirals.

Both barred and normal spirals are further subdivided into what are called Hubble types. The Hubble types (*a*, *b*, *c*) describe two aspects of a galaxy's image; bulge-to-disk ratio and openness of the spiral arms. Hubble type *a* means that a galaxy has a significant amount of stars in its central region, called the bulge, as well as having tightly wrapped spiral arms. Type *c* signifies a galaxy with a small bulge and open, or unwrapped, spiral arms. Hubble type *b* is an intermediate class. For the above examples, NGC 2997 is an Sc galaxy, while NGC 1300 is classified as SB*b*. Irregular (I) is a catch-all category for galaxies that do not fit any of the above molds. In a few cases, irregular galaxies also contain aspects of other morphologies, for example, the bar-like stellar distribution of the Small Magellanic Cloud that orbits the Milky Way.

[1]Reprinted from Hartmann & Impey (1994) by permission of Brooks/Cole Publishing. See Appendix B for the letter of consent.

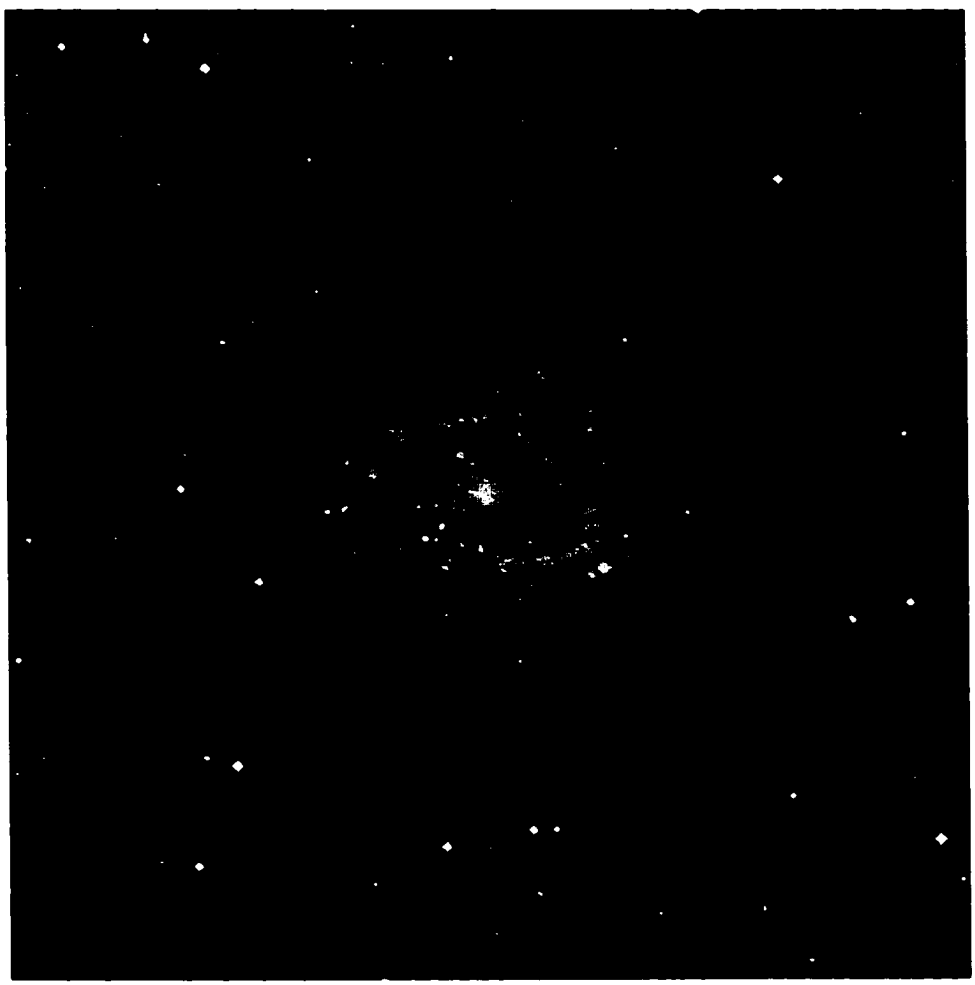

Figure 1.1: A prototypical spiral galaxy, NGC 2997. Taken from the STScI Digitized Sky Survey.

Figure 1.2: A good example of a SBb barred spiral galaxy, NGC 1300. Taken from the STScI Digitized Sky Survey.

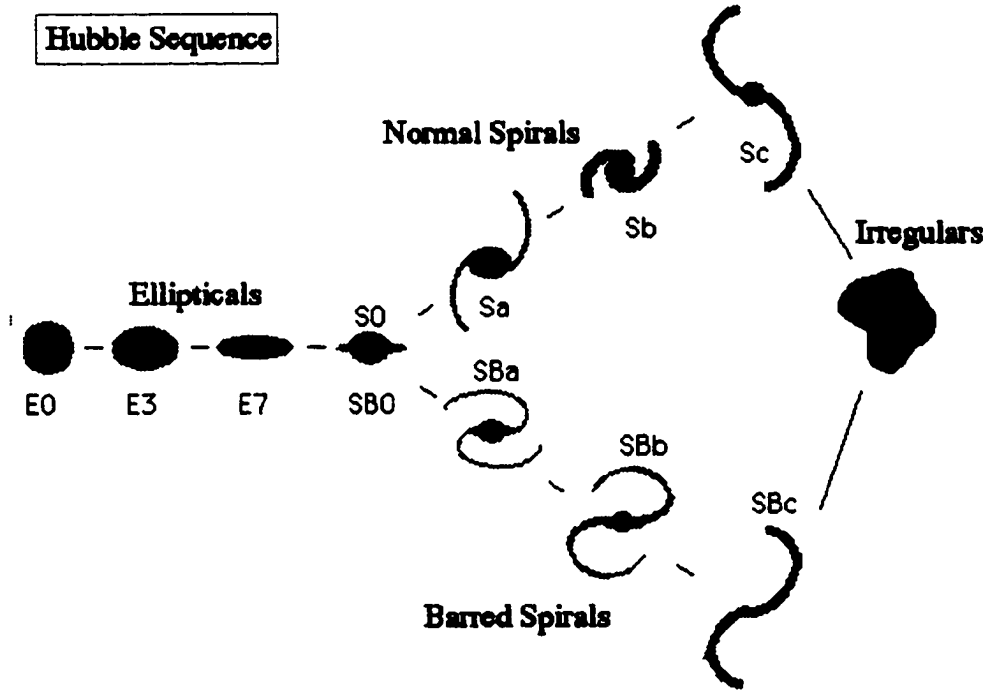

Figure 1.3: The Hubble sequence of galaxy morphologies. Taken from Hartmann & Impey (1994).

1.2 Galaxy Formation

With the understanding that galaxies are distant systems much like our own Milky Way, astronomers and astrophysicists set out to explain how they formed. The modern understanding of galaxy formation begins with the discovery of the Cosmic Microwave Background (CMB) radiation (Penzias & Wilson 1965). This was the evidence that gave the "big bang" theory of the evolution of the universe widespread support. The CMB is an extremely isotropic distribution of radiation created when the universe first cooled to a temperature at which atoms could exist, about 500,000 years after the big bang. The expansion of the universe Doppler shifts this light to longer wavelengths that we call the microwave portion of the electromagnetic spectrum. The isotropic nature of the CMB (no one line of sight is more intense than another to one part in 10^5) stands in stark contrast to the present state of the local universe in which galaxies show no sign of being uniformly distributed. It is generally agreed that the tiny (one part in 10^5) fluctuations in the CMB correspond to the seeds of the universe's large scale structure that is seen at the current epoch (see e.g., Longair 1998).

The other piece of the modern galaxy formation puzzle is what has become known as "dark matter." Dark matter is mass which only interacts with the universe gravitationally. There is no direct evidence for dark matter, for example, no one has isolated a particle of dark matter, but there is a great deal of indirect evidence for its existence. Before the discovery of the CMB, Zwicky was studying clusters of galaxies. He found that the kinematics of the cluster galaxies implied a greater mass for the cluster than could be obtained

by summing the individual masses of visible galaxies (Zwicky 1933). Later, it would be found that individual galaxies show evidence of being influenced by dark matter as well. In particular, curves of circular velocity versus radial distance appear to not follow a Newtonian gravitational law. Specifically, both neutral and ionized hydrogen gas in many spiral galaxies have circular velocities larger than expected (e.g., Rubin & Ford 1970; Krumm & Salpeter 1979). The simplest way to reconcile Newtonian gravity with these findings is to introduce distributions of dark matter around galaxies. These distributions are known as dark matter haloes.

The anisotropies in the CMB and dark matter are the two basic ingredients of modern theories of galaxy formation. The basic scenario holds that the anisotropies in the CMB reflect the distribution of dark matter seeds of large scale structure. Simulations of galaxy formation show that as these seeds grow and interact gravitationally, many aspects of the large scale structure of the universe can be reproduced (e.g., Pearce *et al.* 1999). However, there has not yet been an attempt to create individual galaxies (complete with stellar populations) in these simulations. To do so would involve several additional complex steps. For example, first, a dark matter halo and the galaxy's constituent gas would have to be combined. Next, the gas would have to form stars. Finally, the dynamics of the ensuing stellar population would have to be followed to determine the type of galaxy formed. Of the preceeding steps, the most uncertain is the star formation piece. Issues related to the formation of the first generation of stars and, in particular, the dynamical systems (i.e., galaxies) with which these stars are associated is a

central component of this study. Specifically, the focus here will be on the formation of barred spiral galaxies.

1.3 Barred Systems

Barred spiral galaxies are common in the universe at the current epoch. At least 50% of face-on spiral galaxies have been determined to be barred (e.g., Knapen *et al.* 2000). It is more difficult to assign a barred fraction to edge-on galaxies simply because the obvious barred signature is no longer present. One determination of the fraction of barred galaxies in edge-on galaxies has been made by measuring the percentage of edge-on spirals with boxy or peanut-shaped bulges (Lütticke *et al.* 2000a). The facts that the observed percentage of these bulges ($\approx$ 45%) is similar to the percentage of face-on bars and that three-dimensional (3D) N-body simulations make bulges with similar shapes (Pfenniger & Friedli 1991) is, at least, circumstantial evidence that roughly half of all nearby spiral galaxies are barred (Lütticke *et al.* 2000b). The slightly lower percentage for the edge-on galaxies can be explained by the idea that bars seen end-on (along their major axes) project an elliptical bulge and therefore have not been included in the Lütticke *et al.* (2000a) count.

Despite the ubiquity of barred systems, it has been unclear whether or not normal (S) and barred (SB) spirals are simply opposite ends of a spectrum of bar shapes. Intermediate, weakly barred galaxies do exist (category SA in modified Hubble sequences); so what is the connection? Recently, Abraham & Merrifield (2000) have developed a quantified Hubble sequence. Based on a standard sample of galaxies, they have proposed an answer to this

question. Their study suggests that instead of normal and barred spirals forming opposite ends of a single, normal distribution of galaxy types, the true distribution is bimodal, as suggested by Hubble's original tuning fork sequence (Fig. 1.3). They further suggest that this bimodality could be evidence for a difference in the formation mechanism for the two types of galaxies.

It is commonly believed that normal spiral galaxies are the result of axisymmetric gaseous disks forming stars (e.g., Larson 1976) and that the spiral structure is the result of density waves (Lin & Shu 1964, 1966). The formation of barred galaxies is a less well understood phenomenon. The simplest way to form a bar is through the much discussed "bar instability" that occurs for both collisionless (stellar) and gaseous sytems (however, see Sellwood 2000 for an opposing viewpoint). An excellent review of the bar instability is given in Binney & Tremaine (1987) and is summarized here. Numerical simulations of N-body systems have shown that under certain conditions, an initially axisymmetric (stellar-dynamical) system will deform to a new configuration where most of the particles form a bar shaped object (the small amount of remaining mass forms an extended axisymmetric system) (e.g., Hohl 1971; Hohl & Zang 1979; Sellwood 1981). Ostriker & Peebles (1973), following work by Hohl (1971) and Kalnajs (1972), empirically found that the bar instability occurs in collisionless systems when the ratio of rotational kinetic energy ($T_{\rm rot}$) to the absolute value of gravitational potential energy ($|W|$) is greater than ≈ 0.14. The analogous value for gaseous systems is $T_{\rm rot}/|W| \approx 0.25$. Using data from local regions of the Milky Way, Ostriker &

Peebles (1973) determined a $T_{\rm rot}/|W|$ value that implies our Galaxy should be strongly barred. Their work suggests that there should be dark matter halo components of normal spiral galaxies, in this case to stabilize the disks against bar formation.

Systems that can succumb to the bar instability are called "cool." This terminology can be made clear by introducing the scalar virial theorem. The virial theorem states that (Binney & Tremaine 1987),

$$\frac{1}{2}\frac{d^2I}{dt^2} = 2T_{\rm rot} + \Pi + W, \tag{1.1}$$

where I is the moment of inertia of the system; Π is a measure of the random, or thermal, kinetic energy; and the other terms are defined as before. If the system is considered to be in a steady state, then the left-hand side of eq. (1.1) is zero, and with some rearranging eq. (1.1), can be written as,

$$\frac{T_{\rm rot}}{|W|} + \frac{1.5\Pi}{|W|} = \frac{1}{2}. \tag{1.2}$$

When the kinetic energy of the system is dominated by rotational energy, as is necessary for the bar instability, the thermal energy suffers and hence the system is "cool". While it appears that the bar instability is best explained in terms of the "swing amplifier" (Goldreich & Lynden-Bell 1965; Julian & Toomre 1966; Toomre 1981), the critical value of $T_{\rm rot}/|W|$ identified by Ostriker & Peebles (1973) remains a useful rule-of-thumb for whether or not systems will undergo a global bar instability.

1.4 Stellar Dynamics and Orbit Terminology

This study will focus on stellar orbits in models of time-invariant galaxy potentials. With this in mind, it is useful to present here a brief review of

some basics of stellar dynamics. Stars in galaxies are presumed to form a collisionless system, that is, the forces felt by any one star are due only to the mean forces produced by all other stars. The distribution function, f, of stars must therefore satisfy the collisionless Boltzmann equation (Binney & Tremaine 1987),

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_i \frac{\partial f}{\partial w_i}\dot{w}_i = 0, \tag{1.3}$$

where the independent variables w_i are the phase space coordinates, e.g., $(x, y, z, \dot{x}, \dot{y}, \dot{z})$. The distribution function can be thought of as the probability that a given system will have any given set of phase space coordinates. Alternatively, f can be viewed as a system's phase space density, thereby making eq.(1.3) equivalent to Liouville's theorem. A by-product of this equation is that any time-independent function of phase space coordinates, known as an integral of motion, must be a valid solution of eq. (1.3). This fact is incorporated into the Jeans Theorem: any steady-state solution of eq. (1.3) must be a function of integrals of motion and any function of integrals of motion must likewise solve eq. (1.3) (Binney & Tremaine 1987). A stronger, qualified statement can be made. If all orbits in a given 3D potential are regular (see next paragraph), then any solution to eq. (1.3) is a function of three integrals of motion. This is known as the Strong Jeans Theorem (Binney & Tremaine 1987). Indeed, in potentials that have analytically prescriptible integrals of motion, such as spherical or axisymmetric systems, a variety of distribution functions have been investigated (e.g., King 1966; Kalnajs 1976). In part, this study will be examining whether or not

the gravitational potential of one realistic barred galaxy model can support a stellar distribution function whose stellar orbits are all regular.

The term "regular" is commonly used to describe orbits that respect a number of isolating integrals of motion that is greater than or equal to the degrees of freedom of the orbit (Binney & Tremaine 1987). For example, an orbit in a 2D potential must have at least two isolating integrals in order to be considered regular. In contrast, irregular orbits are fully ergodic, i.e., these orbits will sample every energetically available region of phase space. A closely related type of orbit is called "quasi-ergodic" in this study. Quasi-ergodic orbits are neither regular nor entirely ergodic; they sample large portions of phase space without covering all of the energetically allowed region. Despite this minor difference, both ergodic and quasi-ergodic orbits respect only one complete isolating integral of motion, either the specific energy, ϵ, or the Jacobi constant, ϵ_J (the specific effective energy in a rotating frame, see eq. 3.2).

1.5 This Study

The main motivation of the research presented in this dissertation is to investigate barred galaxy formation from a different perspective than has been taken in the past. Previous numerical work has concentrated on purely stellar-dynamical (collisionless) initial states that lead to barred structures (see §1.3). As alluded to earlier, the overall galaxy formation scenario behind these simulations is that an axisymmetric disk of gas forms stars which then undergo a bar instability as discussed in §1.3. However, it seems equally plausible that the initially axisymmetric gas disk could undergo the bar insta-

bility before a large number of stars have formed. In this case, the resulting stellar distribution function may be significantly different from that in the previously discussed "stars first" example. The goal of this study is to understand the stellar distribution function that can result from an initially gaseous, barred galaxy system.

Cazes (1999) has created a self-consistent, gaseous bar during an investigation of binary star formation. This bar model, referred to here as the Cazes bar, is scaled to a galaxy-sized object in this study. Within the overall attempt to understand stellar orbits in the Cazes bar, the first step towards understanding an emerging stellar distribution function is to investigate equatorial plane orbits (see Chapter 4). There are two reasons for beginning with equatorial orbits. First, there exists a straightforward procedure for documenting two-dimensional orbits (see §3.3). Second, a substantial literature of two-dimensional orbit studies has developed over the past several decades, providing comparison studies. Equatorial orbits with initial conditions unconstrained by Cazes bar gasdynamics are investigated first. In an attempt to gain some insight into the variety of orbits seen, orbits in two analytical approximations to the equatorial Cazes bar potential are also studied. Having cataloged the types of orbits possible in the Cazes bar equatorial plane, a simple transition from gas to stars is studied next. The initial velocities of stars are restricted to be consistent with that of the gas that makes up the Cazes bar. The impact of shocks in the Cazes bar on such orbits is also investigated.

In order to analyze orbits in the fully 3D Cazes bar in a manner similar to the equatorial orbits, an accurate 3D orbital characterization tool is needed. Whereas 2D orbits can be easily characterized by standard techniques, such as surfaces of section (see §3.3), 3D orbits do not yet have a standard analysis tool. Techniques like the spectral dynamics method (Binney & Spergel 1982; Carpintero & Aguilar 1998) and the Lyapunov exponent method (see Udry & Pfenniger 1988; Merritt & Valluri 1996 for good discussions) are both commonly cited, but, they are not particularly useful for this study (see Chapter 5). Instead, the correlation integral characterization method of Grassberger & Procaccia (1983) and Carnevali & Santangelo (1984) has been developed in considerable detail. The reliable and quantitative nature of this method is demonstrated here for a variety of potentials.

With the correlation integral characterization method validated, it is then applied to a large sample of orbits in the 3D analytical approximation to the Cazes bar potential. These orbits are restricted to specific values of effective energy in an effort to understand the nature of possible stellar distribution functions. For example, this type of analysis can determine whether or not quasi-ergodic orbits are likely to play significant roles in real galaxies.

The remaining chapters of this dissertation are organized in the following manner. Chapter 2 contains descriptions of the various models of galaxy potentials (analytical and numerical) that are examined in this study. The tools that are used to create and characterize the orbits are discussed in Chapter 3. The investigation of equatorial orbits is the subject of Chapter 4. Chapter 5 details the implementation of the correlation integral characterization

method, while Chapter 6 describes its usefulness in analyzing orbits in the three-dimensional analytical Cazes bar. Finally, the conclusions of this study are given in Chapter 7. Note that the results summarized in Chapter 4 are drawn from Barnes & Tohline (2001), and the Chapter 5 material appears in Barnes (2001).

2. MODELS

2.1 Numerical Model

2.1.1 Cazes Bar Formation

Self-gravitating, triaxial configurations that either are stationary in inertial space or are spinning about their shortest axis are of broad astrophysical interest. Aside from their relevance to the global properties of spiral and elliptical galaxies, spinning triaxial configurations are thought to be a stage through which dense cores of molecular clouds must evolve in order to produce binary stars (Lebovitz 1987; Cazes & Tohline 2000). Such configurations also can arise in the context of the late stages of stellar evolution (Lai, Rasio, & Shapiro 1993; New, Centrella, & Tohline 2000). In recent years interest in triaxial compact stellar objects has been renewed because they are potentially detectable sources for the gravitational-wave detectors that are being constructed worldwide.

The theoretical understanding of such structures has grown out of the general class of incompressible, ellipsoidal figures of equilibrium originally identified over 100 years ago by Maclaurin, Jacobi, Dedekind, and Riemann, and recently studied in detail by Chandrasekhar (1969). The Riemann S-type ellipsoids, in particular, are an extremely useful family of equilibrium fluid configurations because they have analytically prescriptible properties that span a broad range of geometric parameters. Unfortunately, Riemann ellipsoids are not completely satisfactory models of galaxies, protostellar clouds,

or compact stellar objects because they are uniform-density configurations with very simple internal flows, whereas most astrophysically interesting systems are centrally condensed objects that exhibit a wide assortment of different angular momentum profiles.

In an effort to study the rotational fission instability in more realistic models of protostellar gas clouds, Cazes (1999) recently has utilized numerical hydrodynamic techniques to construct two different steady-state models of rapidly rotating, triaxial gas clouds having a compressible (specifically, an $n = 3/2$ polytropic) equation of state. These models have been described in detail by Cazes (1999). It appears that these are the only fully self-consistent models of self-gravitating, compressible gas bars with nontrivial internal flows that have been presented or discussed in the literature. Because these models provide structures that are more realistic than Riemann ellipsoids, the properties of one of them — specifically the one referred to as "Model B" in Cazes (1999) — will be examined here in the context of the formation and evolution of barred galaxies. Hereafter, this model will be referred to as the "Cazes bar."

The gas-dynamical simulation described by Cazes (1999) that ultimately produced the Cazes bar began from a rotationally flattened, axisymmetric, $n = 3/2$ polytropic gas cloud that was in equilibrium and dynamically stable against axisymmetric disturbances. The initial model was constructed with an angular velocity profile such that, in equatorial projection, the model had uniform vortensity, where vortensity is defined as the ratio of vorticity ($\vec{\nabla} \times \vec{v}$, where $\vec{v}$ is a velocity field) to surface density. The model had a ratio of ro-

tational to gravitational potential energy $t = T_{\rm rot}/|W| = 0.282$ and therefore was sufficiently rapidly rotating that it was unstable toward the development of a bisymmetric, nonaxisymmetric distortion. Although primarily bar-like in structure, the eigenfunction of the unstable bisymmetric mode had a slight, loosely wound, two-armed spiral character. Some redistribution of angular momentum occurred via gravitational torques as the mode grew to nonlinear amplitude. After approximately 30 dynamical times, the system settled down into a new, dynamically stable, spinning bar-like structure containing 98% of the initial cloud mass and 95% of the cloud's original total angular momentum. At this point in the system's evolution, Cazes reconfigured the hydrodynamical code so that the evolution could be continued in a frame of reference that was rotating at a constant angular frequency, the pattern frequency of the bar, and then he followed the system's evolution through an additional 30 dynamical times. This extended evolution showed that, to a high degree of accuracy, the Cazes bar had settled into a steady-state configuration and was dynamically stable.

2.1.2 Cazes Bar Properties

As detailed in the last column of Table 3 of Cazes & Tohline (2000; see also the bottom panels of their Figs. 8 and 9), the bar extends along the major (x) axis to a dimensionless[1] radius of $x_{\rm max} = 1.07$, has an intermediate (y)-to-

[1] As discussed in §3.1 of Cazes (1999), the hydrodynamical simulation that created the "Model B" Cazes bar was performed using a set of dimensionless units so that the model could be straightforwardly scaled to a variety of different types of astrophysically interesting systems. A so-called "polytropic" system of units was adopted in which $M_0 = G = K = 1$, where G is the gravitational constant, K is the polytropic constant in the ($n = 3/2$) polytropic equation of state, and M_0 is the total mass of the initial, axisymmetric, equilibrium configuration from which the Cazes bar formed. As is tabulated in Table 3 of Cazes (1999), in these units, the Cazes bar has a mass $M = 0.958$, a total angular

major (x) axis ratio of approximately 0.52, and possesses two shallow off-axis density maxima at $|x| = 0.31$. Slices of the Cazes bar density distribution along the three principal planes are shown here in Fig. 2.1. The two spiral "kinks" that are immediately apparent in the second and fourth quadrants of the isodensity contours of Fig. 2.1c identify the location, in the equatorial plane, of the two weak standing shocks that accompany the bar's internal flow, as described more fully below.

As discussed by Cazes (1999), the bar is spinning about its shortest (z) axis in a counter-clockwise direction with respect to Fig. 2.1c, with a well-defined pattern frequency,[2] $\Omega = 0.522$, and exhibits a global ratio of rotational to gravitational potential energy, $t = 0.235$. The bar appears to be spinning as a solid object but, in reality, it is not. Instead, as viewed from a frame spinning with the bar's pattern frequency, each Lagrangian fluid element in the bar moves along a well-defined streamline in a periodic, prograde orbit (counter-clockwise in Fig. 2.1c) with a frequency that varies with position along the streamline. The nested fluid streamlines (see the bottom panel of Fig. 9 in Cazes & Tohline 2000) do not cross one another, but streamlines

momentum $J = 0.941$, a semi-major axis length $R_{\rm max} = 8.47$, a pattern frequency $\Omega = 0.522$, and a maximum density $\rho_{\rm max} = 6.69 \times 10^{-3}$ Note that all of the figures in this manuscript show lengths that have been additionally scaled to the equatorial radius ($R_{\rm eq} = 7.95$) of the initial axisymmetric model from which the "Model B" Cazes bar was formed; hence, $x_{\rm max} \equiv R_{\rm max}/R_{\rm eq} = 1.07$. The appendix in Williams & Tohline (1987), for example, shows in detail how any physical variable can be converted from this "polytropic" system of units to more familiar dimensional units. By way of illustration, when the Cazes bar is scaled to $M_0 = 10^{10} M_\odot$ and $x_{\rm max} = 2$kpc, it has a pattern period $P_{\rm pat} = 2\pi/\Omega \approx 1 \times 10^7$ yr and a maximum density of $\approx 3 \times 10^{-22}$ g cm^{-3} (see also Cazes 1999).

[2]At the start of its "steady-state" evolution (time $\tau = 32\tau_{\rm dyn}$), the Cazes bar had a dimensionless pattern frequency of $\Omega = 0.488$, in accord with the value of the frame rotation frequency Ω_0 that is listed in Table 3 of Cazes (1999). At the end of their simulation ($\tau = 59\tau_{\rm dyn}$), however, the pattern frequency had shifted slightly, to the value $\Omega = 0.522$ that will be used here.

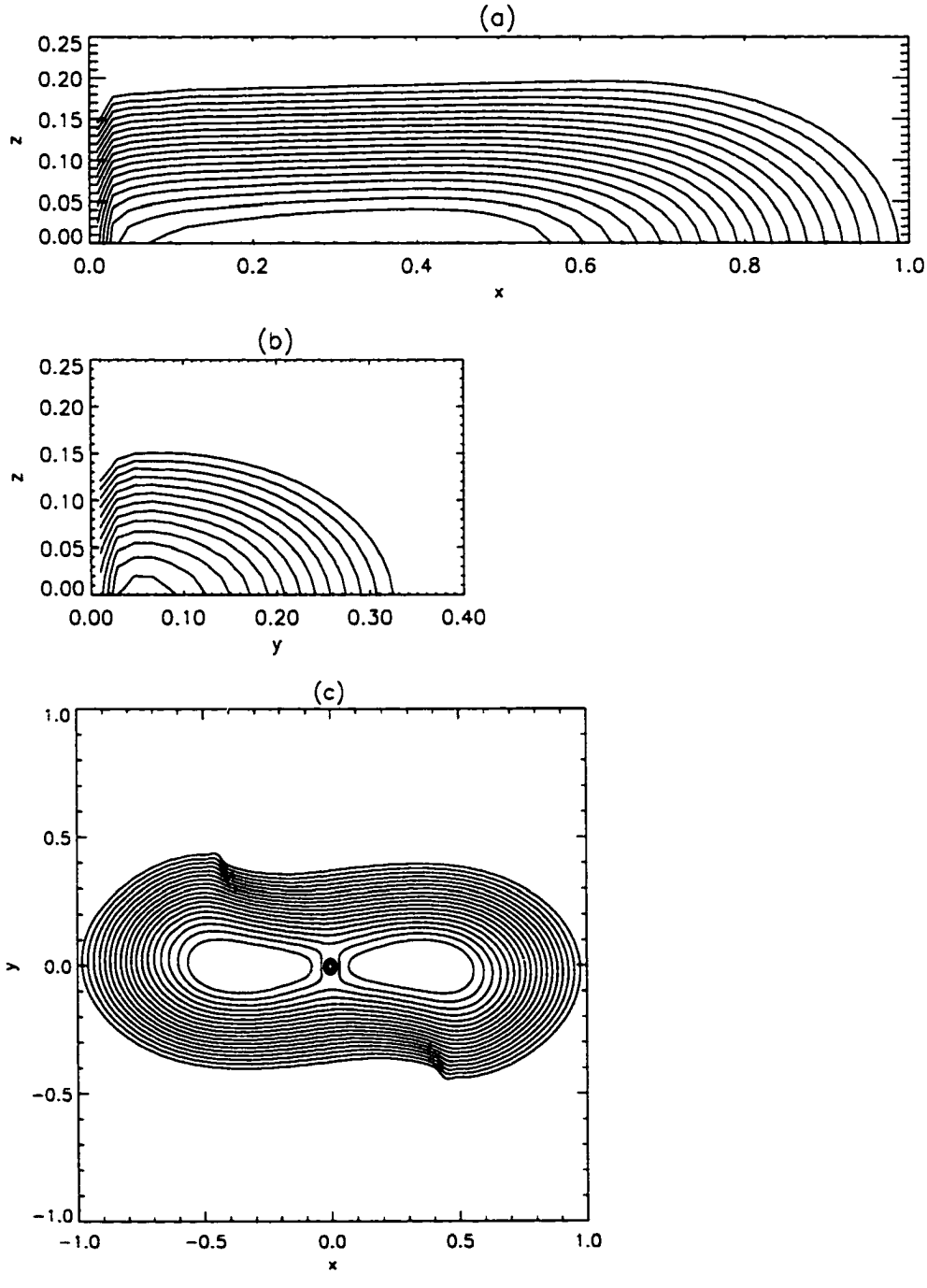

Figure 2.1: Cazes bar isodensity contours in the three principal planes. (a) The $x-z$ plane. (b) The $y-z$ plane. (c) The equatorial $(x-y)$ plane. Note the presence of shocks in (c).

associated with the lowest density (outermost) regions of the bar contain a pair of standing shock fronts. The velocity of fluid elements that follow these outermost streamlines becomes supersonic (in the frame rotating with the bar) as they "fall" along the length of the bar then, with the aid of the shock, the flow becomes subsonic in order to bend around the end of the bar. The two standing shocks are evidenced by the kinks in the isodensity contours displayed in Fig. 2.1c; see also the related violin Mach surface in the bottom panel of Fig. 8 in Cazes (1999). Moving radially outward along the shock, the flow exhibits Mach numbers that vary smoothly from 1.0 to roughly 2.0. Hence, along its entire length, the standing shock is relatively weak. The equatorial velocity field of the Cazes bar is illustrated in Fig. 2.2.

Figure 2.3 displays equipotential contours of the effective potential,

$$\Phi_{\rm eff}(x,y,z) \equiv \Phi(x,y,z) - \frac{1}{2}\Omega^2(x^2+y^2), \tag{2.1}$$

that is generated by the rotating Cazes bar for projections along the three principal axes. Hereafter, the numerically determined effective potential of the Cazes bar will be referred to as $\Phi_{\rm CB}$. Notice that, as with simpler models of rotating bars or oval distortions, e.g., Binney & Tremaine §3.3.2 (1987), $\Phi_{\rm CB}(x,y)$ (Fig. 2.3c) displays four prominent extrema outside of the central, elongated potential well. Two relative maxima appear above and below the bar (these are associated with the traditional $L4$ and $L5$ Lagrange points), and two saddle points (associated with the $L1$ and $L2$ Lagrange points) are marked by asterisks to the left and right of the bar. The $L1$ and $L2$ points are located at a dimensionless distance $R_{L2} = 1.36$ from the origin and, for all practical purposes, define the maximum extent of the bar along the major

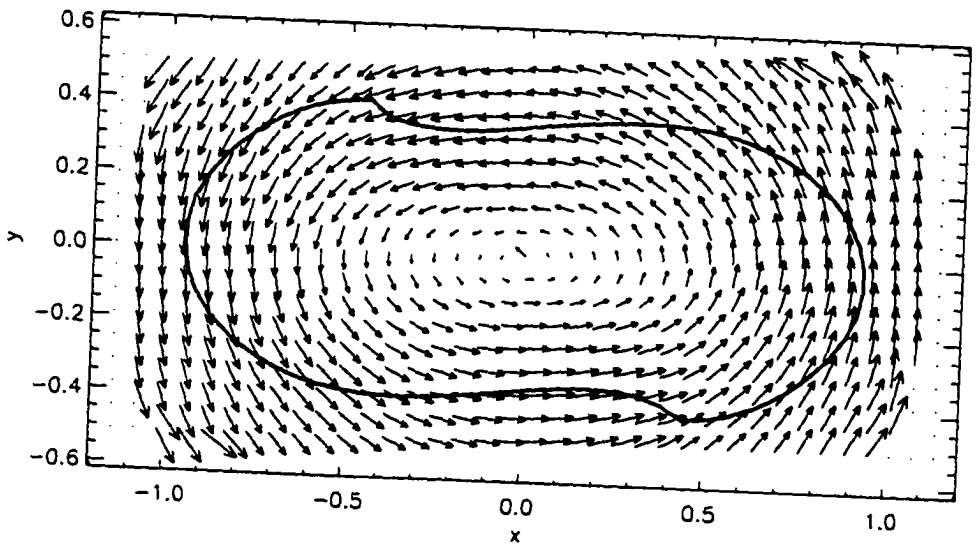

Figure 2.2: The equatorial velocity field (in the rotating frame) of the Cazes bar. The solid line shows the outermost density contour from Fig. 2.1c. Note that the gas flows in a prograde direction.

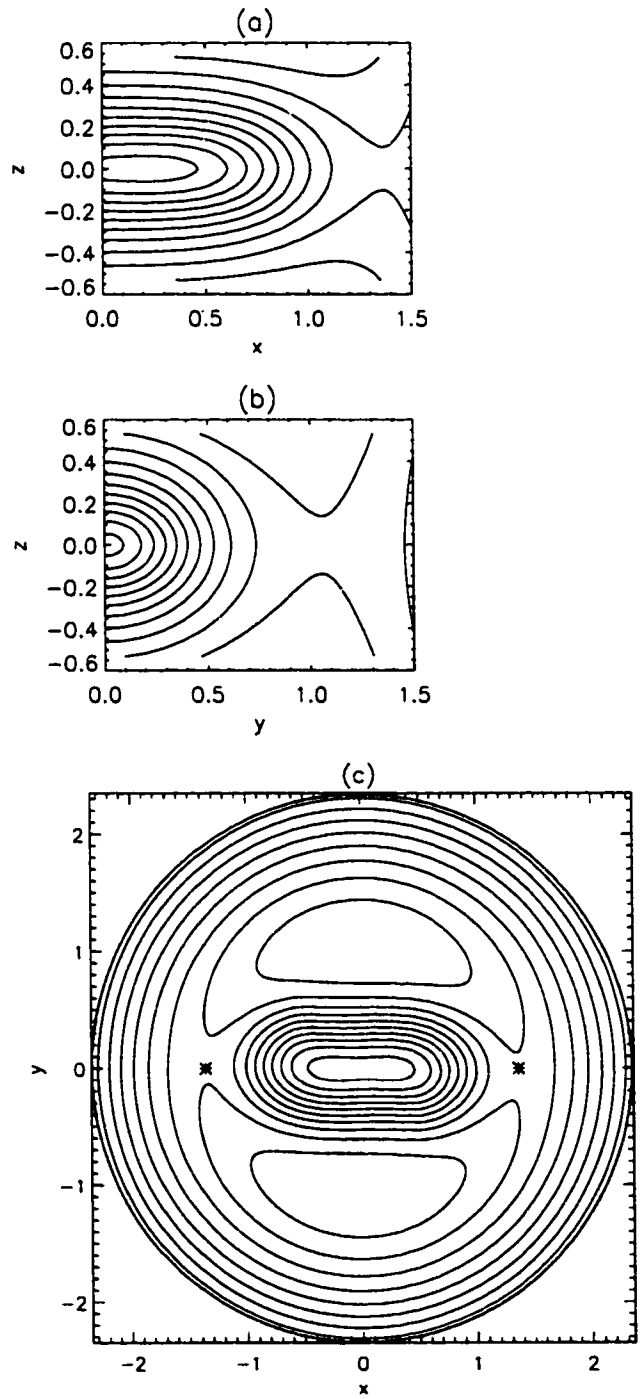

Figure 2.3: Cazes bar equipotential surfaces in the three principal planes. (a) The $x - z$ plane. (b) The $y - z$ plane. (c) The equatorial $(x - y)$ plane. The asterisks mark the $L1$ and $L2$ points.

axis. The solid curves in Figs. 2.4a and 2.4b show the quantitative variation in $\Phi_{CB}(x, y)$ along the major and intermediate axes, respectively, of the bar. Along the intermediate axis, for example, the effective potential varies from a value $\Phi_{min} = -1.018$ at $y = 0$ to a value associated with the $L4$ and $L5$ maxima of $\Phi_{L4,L5} = -0.503$. Along the major axis the effective potential climbs to a somewhat lower value, $\Phi_{L1,L2} = -0.603$, before dropping again at positions $|x| > R_{L2}$. As Fig. 2.1c illustrates, the Cazes bar has two mild off-axis density maxima. These density maxima help support a corresponding pair of slight off-axis minima in the effective potential. The minima are not evident from the contour levels used in Fig. 2.3c, but they can be seen in Fig. 2.4a. Note that the equipotential contours do not trace out simple quadratic surfaces — they have, instead, an overall "peanut" shape — and the contours exhibit a slight spiral twist. When the numerically generated Φ_{CB} is replaced with an analytical "fit" (see §2.2 and §4.3), an attempt will be made to mimic these characteristic features.

2.2 Analytical Models

In the course of this study, various analytical potentials are used in addition to Φ_{CB}. The most important of these is an analytical function that displays the characteristics of the Cazes bar potential and has been created to facilitate a comparison of Φ_{CB} with other previously investigated analytical potentials. Additionally, the validation of the implementation of the correlation integral method presented in Chapter 5 utilizes two previously studied analytical, nonrotating potentials, namely the Richstone and Hénon-Heiles potentials as well as an iterative map known as the Hénon map.

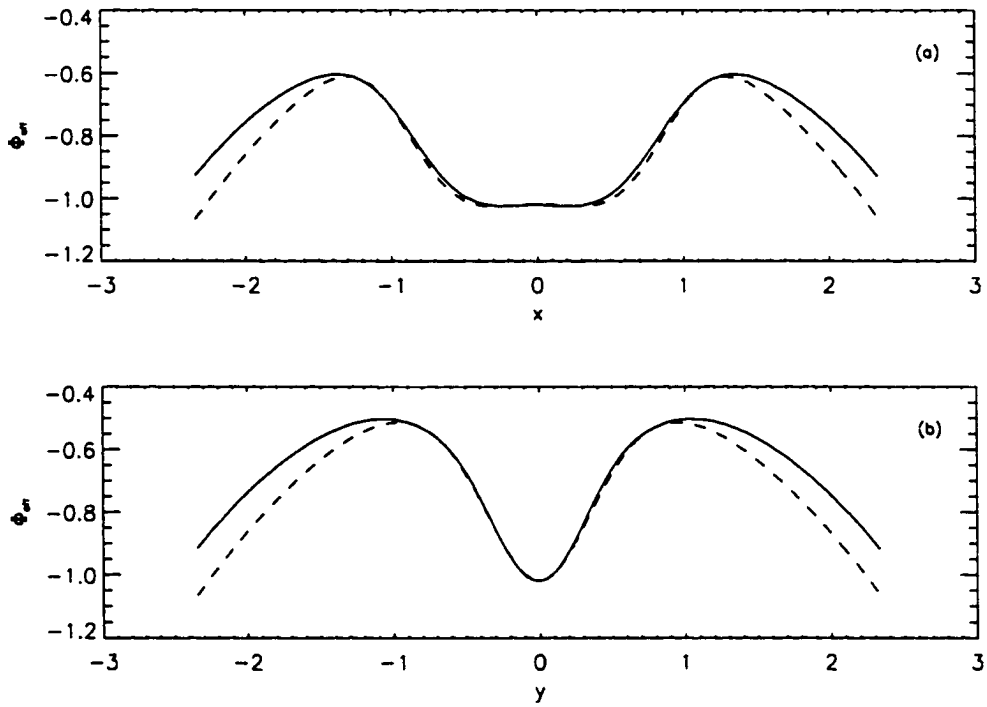

Figure 2.4: Comparison of $\Phi_{CB}(x, y)$ (solid curves) and $\Phi_{aCB}(x, y)$ (dashed curves). (a) Variations along the major (x) axis. (b) Variations along the intermediate (y) axis. Note the presence of two shallow, off-axis minima in (a).

2.2.1 Analytical Cazes Bar

The analytical form that has been created to mimic Φ_{CB} has the form,

$$\Phi_{aCB}(x,y,z) = N\left\{1-\left(1+\left(\frac{x}{R_{L2}}\right)^{\alpha}+\left(\frac{y}{qR_{L2}}\right)^{2}+\left(\frac{z}{q_z z_{lim}}\right)^{\gamma}\right)^{-\beta}\right\} - \frac{1}{2}\Omega^2(x^2+y^2)+\Phi_{min}, \quad (2.2)$$

where N is a normalization factor; q determines the strength of the bar-like distortion in the equatorial plane; q_z determines the strength of the bar distortion in the $x-z$ plane; and α, β, and γ are exponents whose values are to be determined. R_{L2} is the distance along the major axis from the center of the bar to its L2 Lagrange point, and z_{lim} is a vertical scale height. We will only consider values of q, q_z, z_{lim}, and R_{L2} for which the x-axis coincides with the major axis of the bar; the y-axis is the intermediate axis; and the z-axis is along the bar's minor axis. The angular velocity of the bar and the value of the potential at the center are taken from the Cazes bar, that is $\Omega = 0.5218$ and $\Phi_{min} = -1.018$, respectively. When z is set to zero, this potential effectively mimics the structure displayed in Fig. 2.3c. It is the variation of this analytical, equatorial potential along the major and minor axes that is superimposed on the variation of Φ_{CB} in Fig. 2.4.

In an effort to illustrate how well this analytical approximation to the Cazes bar potential, Φ_{aCB}, matches the potential that was derived numerically in the Cazes & Tohline (2000) hydrodynamical simulation, equipotential contours from the analytical Cazes bar are presented in Fig. 2.5 for comparison with Fig. 2.3. Figure 2.5a shows a slice of the analytical Cazes bar along the positive half of the major axis in the meridional plane. The pa-

rameters for the analytical potential pictured here are: $N = 0.7$, $q = 0.8$, $q_z = 1.5$, $R_{L2} = 1.36$, $z_{lim} = 0.65$, $\alpha = 4$, $\beta = 4$, and $\gamma = 1.7$. Equipotential contours in the $y - z$ plane of the analytical Cazes bar are illustrated in Fig. 2.5b. Figure 2.5c displays equipotential contours in the equatorial planes of the analytical Cazes bar. As a check on the applicability of the analytical Cazes bar as a substitute for the Cazes bar, 10 particle orbit integrations have been performed with identical initial conditions in both potentials. In every case, the resulting orbital projections had very similar (although not exact) morphologies.

2.2.2 Richstone Potential

Another analytical potential that has been used in this study is the scale-free, logarithmic potential referred to as the Richstone potential (Richstone 1982). One nice feature of this potential is that it can be used to study either 2D or 3D orbits. To investigate fully 3D orbits, the following form of the potential is used,

$$\Phi_R(x, y, z) = \frac{v_0^2}{2} \ln \left(x^2 + y^2 + \frac{z^2}{q^2} + R_c^2 \right), \tag{2.3}$$

where v_0 is the constant circular speed for the potential, q is a measure of the flattening of the potential, and R_c is a core radius. There are three possible potential shapes:

- $q > 1$ - the potential is prolate spheroidal;
- $q = 1$ - the potential is spherical;
- $q < 1$ - the potential is oblate spheroidal.

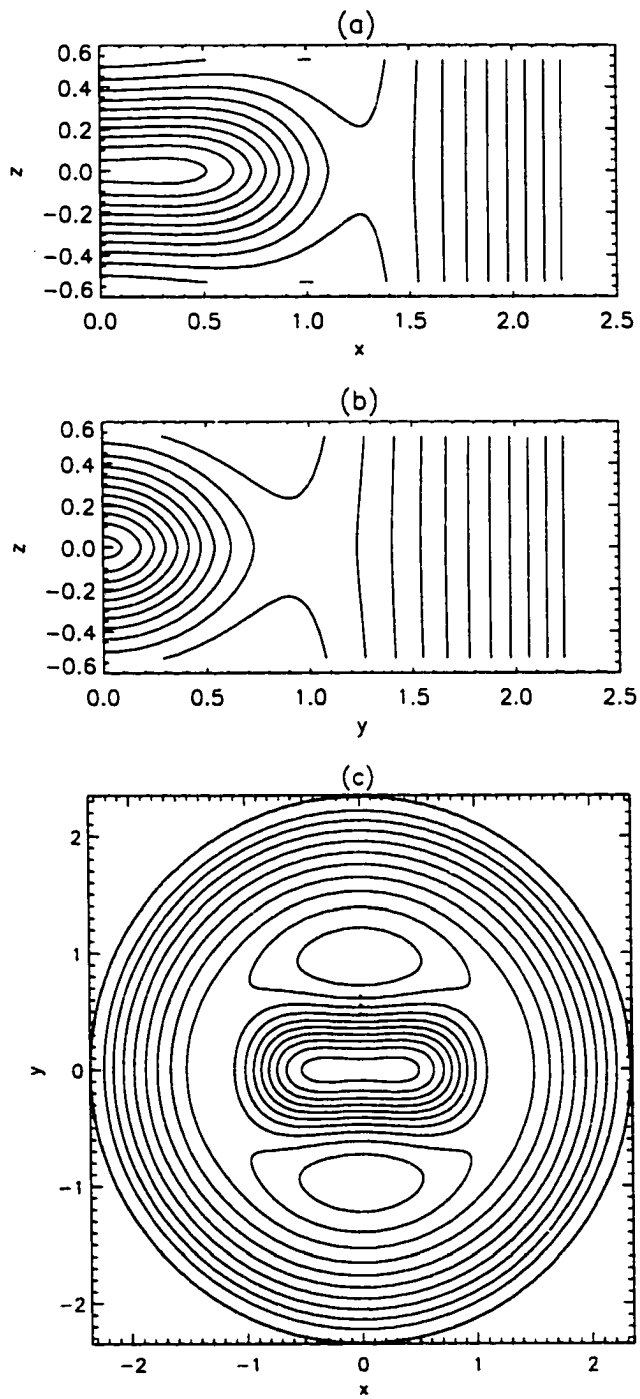

Figure 2.5: Analytical Cazes bar equipotential surfaces in the three principal planes. (a) The $x-z$ plane. (b) The $y-z$ plane. (c) The equatorial $(x-y)$ plane. This figure should be compared with Fig. 2.3.

As in Richstone (1982), the parameters are given the following values, $q = 0.75$ and $R_c = 0.1$.

In order to examine 2D orbital motion in the Richstone potential, first transform eq.(2.3) to cylindrical coordinates (R, ϕ, z). It is clear that Φ_R is axisymmetric, so the Lagrangian is cyclic in ϕ, which means that the z-component of angular momentum is constant. With this restriction, orbits in the 2D Richstone potential move under the influence of the following effective potential,

$$\Phi_{R,\text{eff}}(R, z) = \frac{v_0^2}{2} \ln\left(R^2 + \frac{z^2}{q^2} + R_c^2\right) + \frac{L_z^2}{2R^2}. \tag{2.4}$$

2.2.3 Hénon-Heiles Potential

The Hénon-Heiles potential (Hénon & Heiles 1964) is a well-known potential that supports both regular and quasi-ergodic orbits. This potential has the form,

$$\Phi_{HH}(R, z) = \frac{1}{2}(R^2 + z^2 + 2R^2 z - \frac{2}{3}z^3). \tag{2.5}$$

For energies $\epsilon < 0.01$, almost all orbits in this potential are regular. As the energy is increased, more and more of phase space is occupied by quasi-ergodic orbits. At an energy $\epsilon = 1/6$, the situation is reversed in that most orbits are quasi-ergodic.

2.2.4 The Hénon Map

Unlike the previous potentials, the Hénon map cannot be directly related to a physical system. Instead, it is prescribed by the following set of iterative equations (Hénon 1976):

$$x_{i+1} = y_i + 1 - a x_i^2, \tag{2.6}$$

and

$$y_{i+1} = bx_i, \tag{2.7}$$

where a and b are constants. For the specific values $a = 1.4$ and $b = 0.3$ used in this study, the Hénon map displays a variety of interesting characteristics (see Chapter 5).

3. TOOLS

3.1 Numerical Integration Scheme

Over the past few decades, numerical integration of the equations of motion has become the standard way to investigate stellar orbits in galaxies. In general, the equation of motion has the following vector form,

$$\ddot{\vec{x}} = -\vec{\nabla}\Phi - \vec{\Omega} \times (\vec{\Omega} \times \vec{x}) - 2\vec{\Omega} \times \dot{\vec{x}}, \tag{3.1}$$

where $\vec{\Omega}$ is the angular velocity of a rotating frame of reference, $\vec{x}$ is the position vector in that frame, and Φ is the gravitational potential of the system. When a rotating potential is investigated in this study, the angular velocity vector always points in the positive z-direction. When Ω is nonzero, calculations are performed in the rotating frame. In such a frame, the Hamiltonian for a test mass may be written as,

$$\begin{aligned} H = \epsilon_J &\equiv \epsilon - \vec{\Omega} \cdot \vec{L} \\ &= \frac{1}{2}(p_x^2 + p_y^2 + p_z^2) - \Omega(xp_y - yp_x) + \Phi \\ &= \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + \Phi_{\text{eff}}, \end{aligned} \tag{3.2}$$

where the canonical momenta are,

$$\begin{aligned} p_x &= \dot{x} - \Omega y, \\ p_y &= \dot{y} + \Omega x, \\ p_z &= \dot{z}, \end{aligned}$$

ϵ_J is the specific Jacobi constant (or effective specific energy), ϵ is the total specific energy, $\vec{\Omega}$ is the angular velocity of the bar, $\vec{L}$ is the angular momentum of the particle in the rotating frame, and $\Phi_{\rm eff}$ is defined as in eq. (2.1). The N spatial coordinates form what is called a configuration space, while the inclusion of the canonical momenta give rise to a $2N$-dimensional phase space. For example, the spatial Cartesian coordinates (x, y, z) and canonical momenta (p_x, p_y, p_z) form a six-dimensional phase space.

Time sequences of phase space points are calculated with a Verlet integration scheme (Verlet 1967). This is straightforward to implement for non-rotating potentials, such as the Hénon-Heiles or 3D Richstone potentials. When a rotating potential is examined, the Verlet scheme is modified so that two Verlet steps are performed per fixed timestep. This is done because velocity dependent Coriolis terms must be included in the accelerations. In order to provide optimum values of velocities for evaluation of the Coriolis terms, a first Verlet step is used to obtain a first estimate of the velocities, but particle positions and velocities are not permanently updated at this step. Then, for the same timestep, the second Verlet step returns and updates a more accurate subsequent position and velocity.

When equatorial orbits in the Cazes bar are investigated in this study (see §4.2), $\Phi_{\rm CB}(x, y)$ values are used to determine accelerations. One difficulty with this approach is that $\Phi_{\rm CB}(x, y)$ has been interpolated to an 800×800 Cartesian grid and, hence, a finite-differencing scheme must be used to evaluate derivatives of the potential. For consistency, the analytical, equatorial potentials are also evaluated on the same size grid and finite-differencing is

also applied. A 5-point finite difference stencil is used in each direction to represent derivatives. As a check on the error introduced with the finite differences, the analytically derived gradient of the equatorial, analytical Cazes bar potential has been compared to the finite difference gradient of the same potential defined on an 800×800 grid. The fractional difference is on the order of 10^{-5}. Another difficulty is that particle initial positions are chosen to be on grid lines, but as the orbit integration progresses, each particle position moves continuously. When the particle's position does not fall precisely on the intersection of two grid lines, both components of the acceleration are evaluated at the four grid points that surround the particle's position, then linearly interpolate these to the particle's position. With this scheme, the Hamiltonian of individual orbits is conserved to better than 0.5% and the average value of the Hamiltonian is within 0.1% of the specified Jacobi constant over 10^7 timesteps. That is, each timestep the value of $1/2(v_x^2+v_y^2)+\Phi_{\rm eff}(x,y)$ changes by at most 0.5% and over a long integration, the timestep variations tend to cancel each other out.

While the above procedure provides orbits that can be analyzed with simple techniques (see §3.3), the methods that will be discussed at the end of this chapter (§3.4 and 3.5) require higher accuracy. For this reason, the orbit integrations that are discussed in Chapter 5 are performed with analytically specified potentials and accelerations. Using analytical forms with the modified Verlet algorithm provides a level of energy conservation ($\Delta\epsilon/\epsilon < 10^{-6}$ over 10^9 timesteps) adequate for these high accuracy methods.

3.2 Orbit Population Techniques

3.2.1 Equatorial Shooting Technique

The shooting technique that has been used to investigate unrestricted orbits in the equatorial potentials (see Chapter 4) is as follows. First, a value of ϵ_J is chosen. Second, an initial position x_i is selected along the major axis of the bar. At this position, start with $\dot{x}_i = 0.0$. From ϵ_J, x_i, and $|\dot{x}_i|$, the corresponding value of $|\dot{y}_i|$ is uniquely determined. Then a particle trajectory is integrated with each of the four combinations of these initial velocity components: $(\dot{x}_i, \dot{y}_i)$,$(-\dot{x}_i, \dot{y}_i)$,$(\dot{x}_i, -\dot{y}_i)$, $(-\dot{x}_i, -\dot{y}_i)$. Without changing the initial position, another value of $|\dot{x}_i|$ is chosen, typically proceeding in steps of 0.2. A new $|\dot{y}_i|$ is thereby determined and the integration is repeated for each of the four velocity combinations. This cycle continues until the maximum allowed value of $\dot{x}_i$ has been reached for that initial position. At that point, the initial position is changed and the entire procedure is repeated until the energetically limiting position along the major axis is reached. As a result, roughly 200 unique orbits were examined for each selected value of ϵ_J. Finally, it is emphasized that, throughout this study, equatorial-plane coordinate axes are always oriented such that the x-axis coincides with the major axis of the bar, as in Fig. 2.1c.

3.2.2 Restriction Hypothesis

One of the integral parts of the Cazes bar model is the self-consistent flow field, the equatorial slice of which is shown in Fig. 2.2. Knowledge of how the gas moves in the Cazes bar potential is what allows the Restriction Hypothesis (hereafter, RH) to be implemented. The RH is the idea that

forms the basis of how the transition from gas to stars will be handled in this study. According to the RH, when stars form from a gaseous system, the initial velocities of the stars must be consistent with the velocity field of the gas. For example, when all gas flows on prograde orbits, as in the Cazes bar, stars cannot have retrograde initial velocities. However, once the stars have formed, they are completely decoupled from the gas (i.e., they no longer react to gas pressure). Using the physical quantities in footnote 2, it can be shown that the decoupling of stellar orbits in the Cazes bar is an accurate description since in a Hubble time (10^{10} years) stellar orbits with typical velocities lose only $\approx 10^{-7}$ of their kinetic energy to drag forces. Expanding upon this decoupling idea, if a gaseous system is very cold (low pressure), then the resulting stellar orbits will be more similar to gas orbits than if the parent gas system is hot (high pressure). This approach to creating stellar distribution functions has interesting consequences, as will be discussed in Chapters 4 & 5.

3.3 Surfaces of Section

A surface of section is a convenient and simple way to visualize the structure of a two-dimensional conservative system's phase space. Here, conservative means that either energy or ϵ_J is constant. This study deals only with conservative systems. The existence of this constant decreases the accessible dimensionality of phase space by one. For example, a conservative system with a 2D configuration space has an associated 4D phase space, but any phase space trajectory can be at most three-dimensional. A surface of section is a 2D slice of this restricted 3D phase space. The choice of which

slice to use is arbitrary, but usually one spatial coordinate is set to zero and the surface of section is formed from the other spatial coordinate and its conjugate momentum. The conventions used in this study are discussed below.

There is some ambiguity in the astrophysics literature over the question of which surfaces of section to use when characterizing a two-dimensional potential: (x, p_x) or (y, p_y). For example, Binney (1982a) used (x, p_x) while Sparke & Sellwood (1987) and Teuben & Sanders (1985) examined (y, p_y). In order to glean as much information as possible about the equatorial orbits in $\Phi_{\rm CB}$, we have decided to look at both the (x, p_x) and (y, p_y) surfaces of section. We will base our primary categorizing criteria on the (x, p_x) surfaces of section but, as is illustrated below, the (y, p_y) surfaces of section also can convey some important information, so we use them accordingly. Each (x, p_x) surface of section is obtained by plotting the x-component of the position and canonical momentum every time the particle crosses the x-axis with $p_y > 0$. Alternatively, by plotting the y-component of the position and canonical momentum each time the particle crosses the y-axis with $p_x < 0$, a (y, p_y) surface of section is created.

Orbital characterization with surfaces of section is not quantitative, but is fairly easy to judge by eye. A periodic (closed) orbit will create a surface of section that is composed of a finite number of points. A regular (quasi-periodic) orbit appears as a collection of smooth curves in a surface of section diagram. Ergodic orbits create surfaces of section that are completely filled with points within the energetically allowed region. Surfaces of section for

quasi-ergodic orbits are composed of points that do not lie on smooth curves and yet do not completely fill energetically allowed areas. Many examples of these diagrams are given in Chapter 4.

3.4 Lyapunov Exponents

The technique for measuring the largest Lyapunov exponent (hereafter referred to simply as the Lyapunov exponent) follows from the prescription given in Benettin *et al.* (1976). Lichtenberg & Lieberman (1983) give a good introduction to the method, starting from the definition of the Lyapunov exponent, which is summarized here. The assumption is made that two nearby trajectories diverge exponentially with time, i.e.,

$$d(t) = d(0)e^{kt}, \tag{3.3}$$

where d denotes a phase space distance. The Lyapunov exponent can then be defined to be,

$$k \equiv \lim_{t\to\infty, d(0)\to 0} \left(\frac{1}{t}\right) \ln\left[\frac{d(t)}{d(0)}\right]. \tag{3.4}$$

If a chosen orbit is quasi-ergodic, so that two nearby trajectories diverge as $e^{\sigma t}$, then $k = \sigma =$ constant. If, instead, the chosen orbit is regular and nearby trajectories diverge only as a power law in time, $d(t) \sim d(0)t^{\alpha}$, then

$$k = \frac{\alpha}{t} \ln t. \tag{3.5}$$

In this case, a plot of $\ln k$ vs. $\ln t$ should have a slope of -1 for large values of t, independent of the precise value of the exponent α.

The definition of k as given in eq.(3.4) is computationally unsatisfactory. Exponential growth can quickly lead to numbers that a computer cannot

represent. Benettin *et al.* (1976) suggest, instead, that an orbit be broken into n finite time lengths, τ (see their Fig. 1 or Fig. 5.6 in Lichtenberg & Lieberman 1983 for a pictoral representation of this idea). Then, every τ time units, the distance between neighboring trajectories should be re-normalized to the distance between the two at the beginning of the orbit. With this technique, the Lyapunov exponent is defined to be (Benettin *et al.* 1976),

$$k_n = \frac{1}{n\tau} \sum_{i=1}^{n} \ln \frac{d(i\tau)}{d(0)}, \tag{3.6}$$

where, in the limit of $n \to \infty$, $k_n \to k$.

The task that remains is to calculate whether or not nearby trajectories exponentially diverge. To do this, it is necessary to follow a particular phase space trajectory, $\vec{\xi}(t)$, as well as a neighboring trajectory, $\vec{\xi}(t) + \vec{\Delta\xi}(t)$. $\vec{\xi}$ is a phase space position vector and $\vec{\Delta\xi}$ is initially small. The vector of interest is $\vec{\Delta\xi}$, that is, the vector difference between the orbit that is integrated and the nearby trajectory. Following Lichtenberg & Lieberman (1983), the following definition is made, $\vec{w} \equiv \vec{\Delta\xi}$. From a linear stability analysis (as in Binney & Tremaine 1987, §3.5.3), then,

$$\frac{d\vec{w}}{dt} = \mathcal{M} \cdot \vec{w}. \tag{3.7}$$

$\mathcal{M}$ is a tensor that has components defined by,

$$\mathcal{M} \equiv \frac{\partial \vec{F}}{\partial \vec{\xi}}, \tag{3.8}$$

where $\vec{F} = \left(-\frac{\partial H}{\partial \vec{x}}, \frac{\partial H}{\partial \vec{p}}\right)$, H is the system's Hamiltonian, $\vec{x}$ is the generalized coordinate vector, and $\vec{p}$ is the conjugate momentum vector. In a system

with N degrees of freedom, there are now $3N$ equations that must be solved – N for the particle trajectory and $2N$ for the phase space difference vector.

The $2N$ differential equations given by eq.(3.7) are translated into finite-difference equations that are solved alongside the equations of motion. For orbits that will be analyzed here, k_n is determined for $n\tau = 10^2$, 5×10^2, 10^3, 5×10^3, and 10^4. In addition, the value of τ has been chosen to be the same as Δt, so the difference vector is renormalized every timestep. The presentation of Lyapunov exponents is in the form of plots showing $\ln k_n$ versus $\ln n\tau$. This form has been chosen for these plots because regular orbits (those insensitive to small changes in initial conditions) have slopes ≈ -1, while quasi-ergodic orbits have slopes ≈ 0 because they are sensitive to initial conditions.

3.5 Correlation Integral Method

The dimensionality of a phase space orbit can be determined by calculating a correlation integral and observing its dependence on phase space distance. Operationally, the steps in determining the dimensionality of a phase space orbit are:

1. Integrate an orbit for a sufficient number of timesteps ("sufficient" will be clarified below).
2. Choose a set of sampling points from the orbit.
3. Calculate the correlation integral, $C(r)$, as defined by eq. (3.9).
4. Measure a slope from a plot of $\ln C(r)$ vs. $\ln r$.

Because topics 1 and 2 are somewhat connected, they are discussed together. From various trials, investigating phase spaces of different sizes, it has been

determined that approximately 50 orbital periods must be completed in order for an accurate correlation integral to be found. This means that more timesteps are necessary for each 3D orbit than for each 2D orbit. With the timestep that has been used throughout the investigation of the correlation integral method (see Chapter 5), the ratio between the number of required timesteps is $\approx$ 100. (It is likely that these numbers would change if a different integration scheme was used.) The sampling points are chosen following the suggestion in Carnevali & Santangelo (1984); specifically, random points are selected from a subset of the orbital phase space points. The subset consists of 10^5 points taken at equal timestep intervals. Normally, 10% of these points are randomly chosen to be sampling points. While this number of sampling points is usually adequate for the purposes of this study, it generally is best to use as many sampling points as is computationally practical.

The correlation integral is determined numerically using the following formula (Grassberger & Procaccia 1983):

$$C(r) = \lim_{N\to\infty} \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1,\neq i}^{N} \Theta(r - |\vec{\xi}_i - \vec{\xi}_j|), \tag{3.9}$$

where Θ is the Heaviside step function, the $\vec{\xi}_i$ are phase space position vectors of points on a phase space orbit, r is a phase space distance, and N is the number of sampling points from the orbit. It is now clear why it is advisable to use as large a value of N as practical, as the correlation integral is defined more accurately when N is larger. A naive evaluation of eq.(3.9) is very slow for large N, even if one takes advantage of the symmetry of the Heaviside function.

The procedure for speed-up used in this study divides phase space into "bins". Once the sampling points have been chosen, the minimum and maximum values for each dimension of phase space are found. Actually, when dealing with conservative systems, one dimension of phase space may be neglected because there is a conserved quantity (either energy or Jacobi's constant). In this study, this shortcut is used only when analyzing a 6D phase space orbit (the z coordinate is ignored). A certain number of bins per dimension (the same for all dimensions) is chosen. The sizes of each dimension's bins are then given by (maxval − minval)/(# of bins). The sampling points are then assigned to bins, each of which has an index assigned to it. For example, in a s dimensional phase space with b bins per dimension, the bin indices run from $1 \rightarrow b^s$. Now, the list of N random points is sorted (smallest to largest) according to their corresponding bin index. A loop over phase space distances, r in eq.(3.9), is begun. Each r value is compared to the smallest bin length. If the ratio is less than one, only points in the same or first neighboring bins are included in the sum. As the ratio increases, more and more neighboring bins (and hence neighboring points) are taken into the sum. When the ratio is greater than or equal to the number of bins per dimension, all points are included. This nearest-neighbor hunt is performed for each of the b^s bins. Finally, a new r value is chosen and the cycle begins again. This method of calculating correlation integrals is most effective when the r value is small enough that only nearest neighbor cells need be included. Basically, this binning technique allows the regions of phase space that contribute nothing to the sum in eq. (3.9) to be ignored.

The final step in determining the dimensionality of a phase space orbit is measuring a slope from the plot of $\ln C(r)$ vs. $\ln r$. A log-log plot is used because of the behavior of the correlation integral. Grassberger & Procaccia (1983) state that, for $r \ll 1$, $C(r) \propto r^{\nu}$, where ν is the dimensionality of the phase space orbit. This is good news since the calculation of $C(r)$ is most efficient for small r. So, when $\ln C(r)$ is plotted against $\ln r$, the slope of any linear section can be interpreted as the dimensionality ν of the phase space orbit within that range of r. In order to ascertain the reliability of the values of ν that we have measured for each orbit, a minimized χ^2 linear fit to the $\ln C(r) - \ln(r)$ data for five independent sets of sampling points has also been calculated. Also, at least one order of magnitude in r must be covered by the linear section to be considered. It is the average slopes ($D = \langle \nu \rangle$) and standard deviations (σ) of these five linear fits that are reported in the legends of figures such as Fig. 5.10b.

4. ANALYSIS OF EQUATORIAL ORBITS

4.1 Background

Following the lead of Contopoulos *et al.* (1989), it is useful to group previously published theoretical studies of the structure and stability of barred galaxies into the following four broad categories:

1. Orbit calculations for various two-dimensional (2D) analytical, bar potentials.

2. N-body simulations of stellar bar formation.

3. N-body simulations of interstellar cloud collisions in bars.

4. Studies of gasdynamical flows in externally prescribed "stellar" bar potentials.

This work does not fit naturally into any of these categories because it involves a detailed analysis of the properties of a self-gravitating gasdynamical, rather than stellar-dynamical, bar-like configuration. However, there are strong parallels between this study and the analysis presented by Sparke & Sellwood (1987) of a purely stellar-dynamical bar, so a reminder of the key features of this earlier work are presented for the reader.

In an effort to ascertain not only what kinds of orbits are allowed in the potential well of rapidly rotating, barred galaxies but also which orbit families are likely to be populated by stars in such galaxies, Sparke & Sellwood (1987,

hereafter SS) combined the tools and analysis techniques that previously had been associated with the separate categories of investigation listed as items 1 and 2, above. First, they used a 2D N-body code to construct a steady-state model of a rapidly rotating, infinitesimally thin, bar (initially including a small axisymmetric bulge component, a surrounding axisymmetric disk, and a "hot" component). Then, using a standard "shooting technique" along with surfaces of section and a characteristic diagram (e.g., see Fig.4.8), they mapped out the properties of available (stable and unstable) orbits in the effective potential well of this numerically generated, 2D steady-state bar. Finally, they identified which of the numerous possible orbit families were actually being populated by particles in their N-body simulation. Among other things, the SS work provided strong evidence in support of a "restriction hypothesis"[1] first alluded to by Teuben & Sanders (1985), namely, that a real barred galaxy contains stars that largely follow a favorable subset of all possible orbit families. More specifically, SS found that the stellar distribution function that was associated with their steady-state bar, DF_{SS}, was dominated by particles whose trajectories were associated with, and generally trapped around the x_1 family of orbits as defined by Contopoulos & Papayannopoulos (1980).

This restriction hypothesis has received additional support from Pfenniger & Friedli (1991), who extended the SS analysis to fully three-dimensional (3D) N-body models of steady-state bars, as well as from the study of

[1]The phrase "restriction hypothesis" has been adopted here in an effort to encapsulate the essence of the first sentence of §4.3 in SS as well as the similar implication that appears in Teuben & Sanders (1985). This phrase does not appear in either the SS or the Teuben & Sanders (1985) reference.

Berentzen *et al.* (1998), in which a small amount of gas (8% of the total galaxy mass) was included in a self-consistent fashion along with a 3D N-body simulation. It is not obvious how a nonlinear dynamical simulation - and by inference a real galaxy - that starts from a nearly axisymmetric distribution function, DF_{axisym}, is able to preferentially select this restricted set of orbits (primarily related to the x_1 family) while evolving to a bar-like configuration, but the outcome makes sense. Indeed, other generally available orbits, such as orbits associated with the x_2 or x_4 families (Contopoulos & Papayannopoulos 1980), have trajectories that generally do not support the overall shape of the bar.

Here, an analysis that is very similar to the one presented by SS is performed, but for a 2D equatorial slice of the 3D Cazes bar. This analysis has been motivated, in part, by the facts that the steady-state gasdynamical Cazes bar is the first detailed model of its kind and that it was unknown to what degree its global attributes resemble the properties of previous N-body counterparts. By using a shooting technique to inject test particles (stars) into the potential well of the equatorial Cazes bar and then following the motion of the stars through many orbital periods, surface of section diagrams have been produced to facilitate such a comparison.

4.2 Cazes Bar Potential

4.2.1 Composite Surfaces of Section

Figure 4.1 shows a composite (x, p_x) surface of section diagram for six separate regular orbits that arise in the equatorial Cazes bar potential when the Jacobi constant $\epsilon_J = -0.75$. (This value of the Jacobi constant has

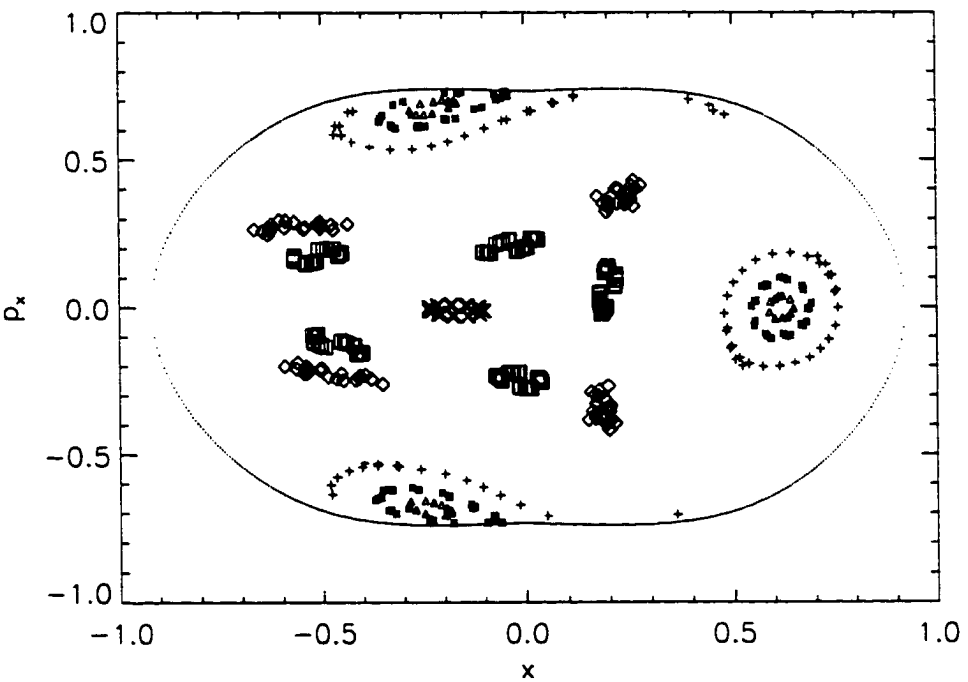

Figure 4.1: The (x, p_x) composite surface of section diagram for 6 selected regular orbits with $\epsilon_J = -0.75$ that are supported by $\Phi_{\mathrm{CB}}(x, y)$. The dotted line surrounding the invariant curves is the zero velocity curve.

been selected for illustrative purposes only; it has no special significance other than it lies between $\Phi_{\min}$ and $\Phi_{L1,L2}$). The contour that corresponds to this value of ϵ_J is identified by the dashed-dotted solid line in Fig. 4.5a. Additionally, the surface of section diagrams in this chapter also display zero-velocity curves, i.e. the locus of points in the surface of section diagram at which the potential equals the value of ϵ_J. Figure 4.1 contains:

- one elongated region (marked by $\times$ symbols) confined to a narrow, short segment of the negative x axis;
- five disconnected regions (marked by squares) that lie mostly at negative values of x and surround the narrow elongated region;
- four disconnected regions (represented by diamonds) having $|p_x|$ values that are generally larger than that of the five regions marked by squares;
- three islands identified by two separate, but nested surfaces of section (marked by asterisks and triangles);
- a set of three curves (shown as $+$ symbols) that appear to define a boundary between the three islands and the region of the diagram occupied by the other surfaces of section.

Figure 4.2 is a (y, p_y) composite surface of section that complements the (x, p_x) surface of section shown in Fig. 4.1. The orbits that create each of the surfaces of section in Fig. 4.2 are marked by the corresponding symbols in Fig. 4.1. For example, the orbit that forms the smallest three-island surface of section in Fig. 4.1 (marked with triangles) creates the skewed

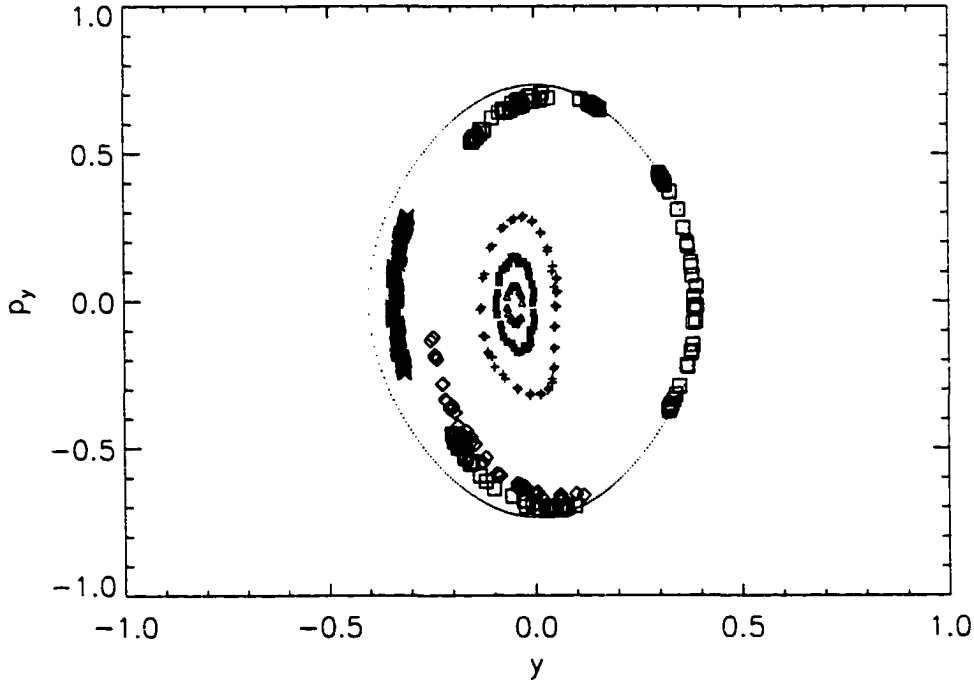

Figure 4.2: The (y, p_y) composite surface of section diagram for the same 6 orbits represented in Fig. 4.1. As in Fig. 4.1, the dotted line is the zero velocity curve.

ellipse that is centered on $y \approx -0.05$ in Fig. 4.2. Note that the majority of points that make up the (y, p_y) surfaces of section fall at negative values of y, suggesting that the orbits from which they are derived are retrograde. It will be demonstrated that the surface of section marked by crosses is derived from an orbit which appears to be related to the x_4 family of retrograde orbits. The surfaces of section marked by squares and diamonds belong to retrograde orbits with higher order resonances. However, the orbits associated with the three islands in Fig. 4.1 are, in fact, prograde.

One striking feature of all of the surfaces of section that make up Figs. 4.1 and 4.2 is the lack of symmetry. In nonrotating, bisymmetric potentials, surfaces of section show reflection symmetry about both the $x = 0$ $(y = 0)$ axis and the $p_x = 0$ $(p_y = 0)$ axis. A variety of such symmetric surfaces of section may be found in Binney & Tremaine §3.3 (1987). Rotating potentials lose the reflection symmetry about the $x = 0$ $(y = 0)$ axis, but generally retain it across the $p_x = 0$ $(p_y = 0)$ axis. Some examples of surfaces of section with this symmetry intact may be seen in SS as well as in Teuben & Sanders (1985). Figures 4.1 and 4.2 exhibit the expected rotational based asymmetry with respect to the $x = 0$ $(y = 0)$ axis, but they also display a slight asymmetry about the $p_x = 0$ $(p_y = 0)$ axis. As the analytical approximation to Φ_{CB} is developed in §4.3, the primary goal is to reproduce the primary features seen in Figs. 4.1 and 4.2, including this asymmetry.

4.2.2 Individual Orbits

While the surface of section diagram provides a useful tool for categorizing orbits, the orbits themselves are of primary importance. To begin, a

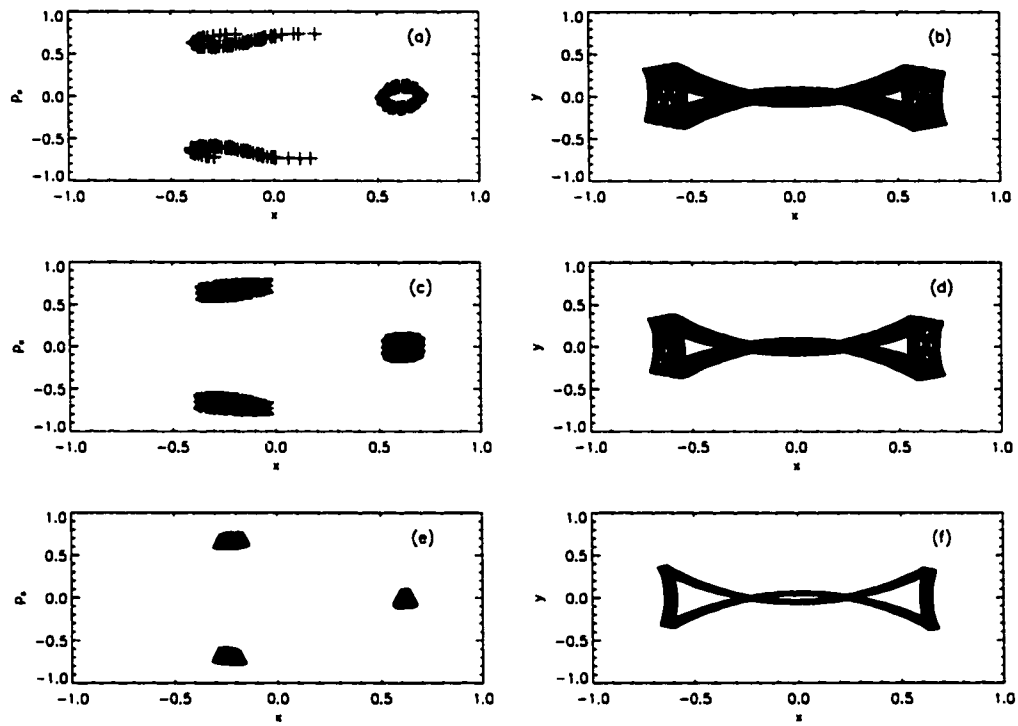

Figure 4.3: Plots of the 6 individual surfaces of section taken from the Fig. 4.1 composite diagram and their corresponding orbits. Surfaces of section are shown on the left, (a,c,e,g,i,k); orbits are on the right, (b,d,f,h,j,l). The symbols used for each surface of section are the same as in Fig. 4.1.

description of the regular orbits that have just been identified in connection with the Cazes bar potential is given. The frames in the left column of Fig. 4.3 isolate individual (x, p_x) surfaces of section from the Fig. 4.1 composite diagram, while the frames in the right column of Fig. 4.3 illustrate the $x - y$ orbital trajectories from which each corresponding surface of section was derived. Figure 4.3f shows the nearly closed orbit that leads to the smallest, three-island surface of section (marked by triangles) illustrated in Fig. 4.1. This orbit, as well as each of the two closely related orbits depicted

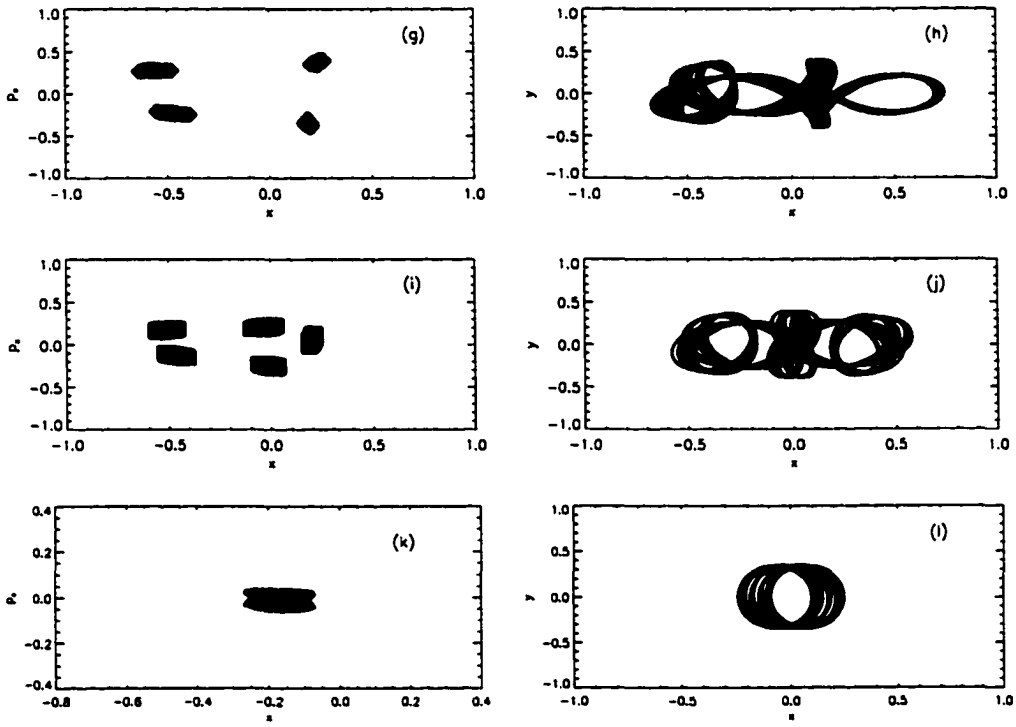

Figure 4.3: Continued

in Fig. 4.3d and Fig. 4.3b, has the shape of a bow tie. Hence the regions of the surface of section diagrams that are occupied by these orbits — the three-islands in (x, p_x) and the skewed ellipses near the origin in (y, p_y) — will be referred to as the bow tie regions. Particles travel on bow tie orbits in a counter-clockwise direction (i.e., the overall motion is prograde) and make four radial oscillations before completing one full orbit cycle. Hence, the orbits illustrated in frames $a - f$ of Fig. 4.3 are almost certainly related to the 4/1 family of orbits discussed by Contopoulos (1988; see especially his Fig. 1a). However, during two of the radial oscillations in a bow tie orbit, the particle passes very close to and, indeed, around the origin in such a way that its direction of motion formally becomes retrograde. This is why the (y, p_y) surface of section for these orbits generally resides at negative values of y. It is noted as well that the bow tie orbits do not exhibit perfect reflection symmetry about the x-axis. For example, the top and bottom sections of the orbit shown in Fig. 4.3d seem to be tilted with respect to the intermediate (y) axis; and the bottom of the "v" shape that is formed on the top of the orbit shown in Fig. 4.3b does not lie directly above the inverted "v" that is formed on the bottom of that orbit.

The relatively simple orbit shown in Fig. 4.3l is a retrograde orbit. This is clear from the (x, p_x) surface of section (Fig. 4.3k), which shows that each time a particle on this orbit crosses the x-axis with a positive p_y it is to the left of the origin (i.e., at negative x), as well as from the (y, p_y) surface of section (marked with crosses in Fig. 4.2), which shows that each time the particle crosses the y-axis with a negative p_x it is below the origin (i.e., at

negative y). This orbit is almost certainly a member of the x_4 family of orbits, as defined by Contopoulos & Papayannopoulos (1980).

The regular orbits shown in Figs. 4.3h and 4.3j are also largely retrograde. However, these orbits are much more complex than the one illustrated in Fig. 4.3l. Using the terminology of Contopoulos (1988), Fig. 4.3h displays a 5/1 orbit; that is, the orbit makes five radial oscillations for every complete orbit cycle. Similarly, Fig. 4.3j displays a 6/1 orbit. It is easier to understand why these two nearly closed orbits display, respectively, four and five disconnected regions in the (x, p_x) surface of section diagram if, rather than counting radial oscillations, the number of y-oscillations the orbit undergoes before completing one full (horizontal) excursion along the bar is counted. In this sense, Fig. 4.3h displays a $4:1$ orbit while Fig. 4.3j displays a $5:1$ orbit, exactly matching the number of disconnected regions that arise in the (x, p_x) surface of section diagram.

Like a number of other previously investigated, nonaxisymmetric potentials, the Cazes bar potential supports a rich variety of regular orbits that have a recognizable $n:m$ oscillatory pattern, in the sense just discussed. Particles following these trajectories complete n oscillations perpendicular to the major (x) axis in the time that it takes them to complete m circuits along the major axis. In the case of a closed orbit in which the oscillations perpendicular to the major axis actually cross the major axis, such an orbit would be represented by n distinct points in a (x, p_x) surface of section diagram. However, it is also possible that not every oscillation will cross the major axis. As an example of this variety, Fig. 4.4a illustrates a nearly closed, regular

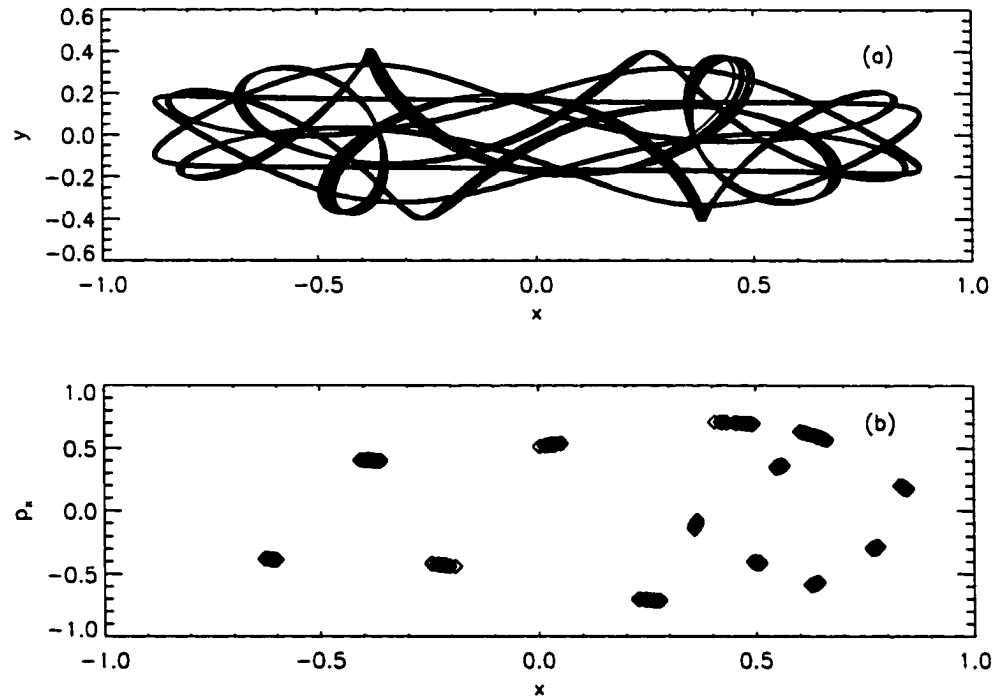

Figure 4.4: Plots illustrating the behavior of a 15:5 orbit that is supported in $\Phi_{CB}(x, y)$. (a) Orbit showing 15 vertical oscillations for every 5 horizontal oscillations. (b) The corresponding (x, p_x) surface of section diagram.

15:5 orbit. However, the (x, p_x) surface of section diagram in Fig. 4.4b displays only 13 islands. The difference between what is expected (15 islands) and what is observed (13 islands) is due to the interesting behavior of this particular orbit. When the two apparently straight sections of the orbit (at $y \approx 0.2$ and $y \approx -0.2$) are closely scrutinized, they each definitely exhibit a small y oscillation. Since neither of these nearly horizontal segments crosses the x-axis, neither generates a corresponding island in Fig. 4.4b.

The Cazes bar potential also allows quasi-ergodic orbits to develop. In fact, $40 \pm 5\%$ of the ≈ 200 orbits studied at this Jacobi constant are quasi-

ergodic (as judged from surface of section diagrams). This statistical measure has been determined as follows. First, the sample is divided into two random subsamples, each containing half the total number of orbits. The average number of each orbit type (in this case, regular or quasi-ergodic) is calculated for the two subsamples. This process is then repeated ten times, and the mean and the standard deviation of the separate averages are calculated for the ten trials.

As mentioned earlier, quasi-ergodic orbits wander through the bar without an overall shape. For this reason, they are difficult to discuss individually, but collectively they have characteristics of interest. These orbits cross the major axis of the bar many times as they move along the length of the bar. Most importantly, they support the shape of the bar. A sample quasi-ergodic orbit is shown superimposed on equipotential contours of the Cazes bar in Fig. 4.5a. The corresponding surface of section is shown in Fig. 4.5b.

4.2.3 Composite Surfaces of Section for Varying ϵ_J

While the previous sections have dealt with orbits at a single effective energy, it is interesting to see phase space structure at various Jacobi constant levels. Figure 4.6 contains composite (x, p_x) surface of section diagrams for four separate values of ϵ_J: (a) $\epsilon_J = -0.96$ (near the bottom of the potential well); (b) $\epsilon_J = -0.85$; (c) $\epsilon_J = -0.75$ (this is the same as Fig. 4.1); and (d) $\epsilon_J = -0.63$ (almost at the $L1, L2$ energy level). Figure 4.7 shows the corresponding (y, p_y) surfaces of section. As before, quasi-ergodic surfaces of section are not shown, but do exist at each of these energies. The most striking aspect of these diagrams is their similarity to one another. The

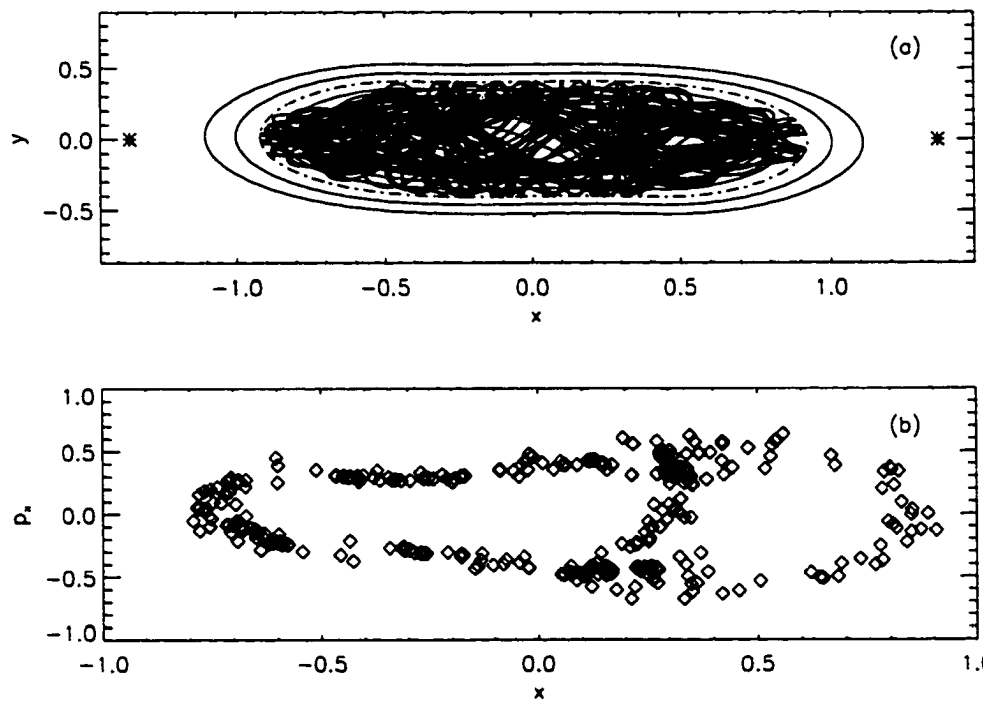

Figure 4.5: (a) A quasi-ergodic orbit with $\epsilon_J = -0.75$ that is supported by $\Phi_{CB}(x, y)$ is shown superimposed on equipotential contours of that potential. The dashed-dotted contour drawn at $\Phi_{CB} = -0.75$ also serves as a boundary of the area inside which this orbit is confined. As in Fig. 2.3c, asterisks mark the positions of the $L1$ and $L2$ Lagrange points. (b) The (x, p_x) surface of section for this orbit.

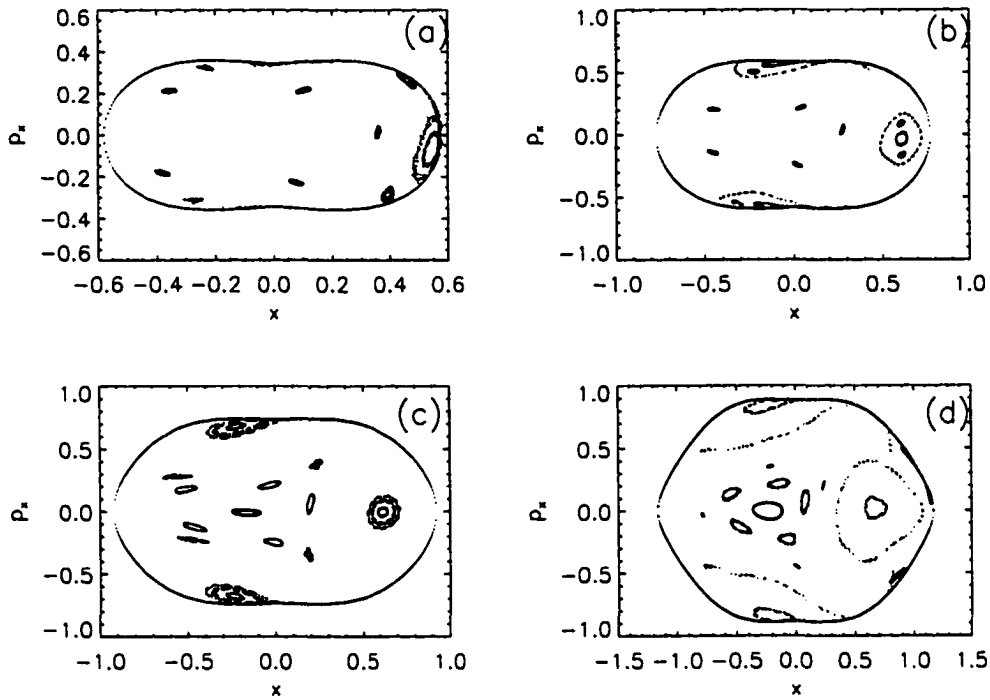

Figure 4.6: Composite (x, p_x) surfaces of section for four different values of ϵ_J. (a) $\epsilon_J = -0.96$; this value of the Jacobi constant traps particles near the bottom of the potential well. (b) $\epsilon_J = -0.85$. (c) $\epsilon_J = -0.75$. (d) $\epsilon_J = -0.63$; this value of the Jacobi constant allows particles to move throughout the entire bar. Note the presence of bow tie and $5:1$ orbital surfaces of section in each frame. Also, x_4 orbits appear only in (c) and (d), while x_1 orbits appear only in (d) (see also Fig. 4.8).

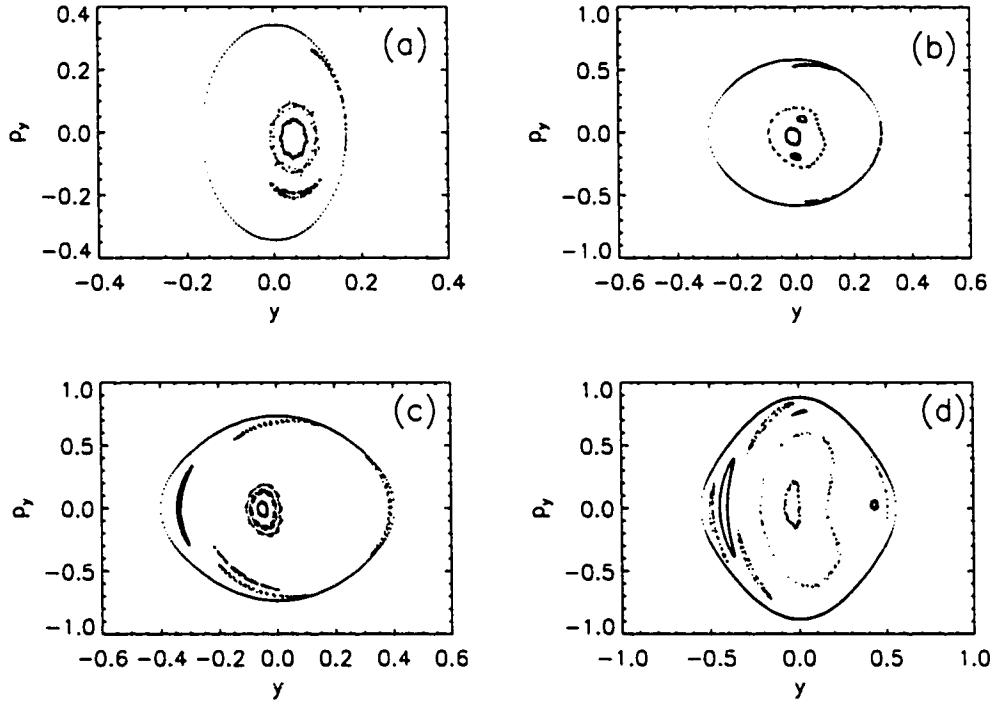

Figure 4.7: Composite (y, p_y) surfaces of section for 4 different values of ϵ_J. The energy levels for (a), (b), (c), and (d) are the same as in Fig. 4.6.

bow tie and 5 : 1 orbits appear in all diagrams. One difference among these diagrams is the small loop that appears at ($x \approx -0.1, p_x \approx 0.0$) in Figs. 4.6c and 4.6d. As mentioned in §3.2, these loops are created by x_4 orbits. More difficult to distinguish is the presence of an x_1 surface of section in Fig. 4.6d. This surface of section is composed of two pieces that lie close to the zero velocity curve (outer dotted line). A better view of the x_1 surface of section is shown in Fig. 4.7d; it is the small loop located at ($y \approx 0.4, p_y = 0.0$). The relationship between these (x_1, x_4, and bow tie) orbits and the energy range over which they exist is best illustrated by a characteristic diagram, as shown here in Fig. 4.8. This diagram displays the location at which each periodic orbit crosses the y-axis as a function of the energy (Jacobi constant in this case) of that orbit. Figure 4.8 demonstrates that the bow tie orbits are the dominant regular orbital family in the Cazes bar potential.

4.3 Analytical Potentials

4.3.1 Analytical Cazes Bar Potential

The analytical potential described in §2.2.1 is used (with $z = 0$) to better understand the particular variety of orbits supported by the equatorial Cazes bar potential. Numerous surfaces of section for this analytically prescribed effective potential have been generated from orbits with $\epsilon_J = -0.75$. In this potential, approximately 50% of the orbits studied are quasi-ergodic. The composite (x, p_x) surface of section for some of the regular orbits is shown in Fig. 4.9; Fig. 4.10 shows the corresponding composite (y, p_y) surface of section diagram. Figures 4.9 and 4.10 closely resemble Figs. 4.1 and 4.2, respectively. It appears that the analytical potential that has been used to

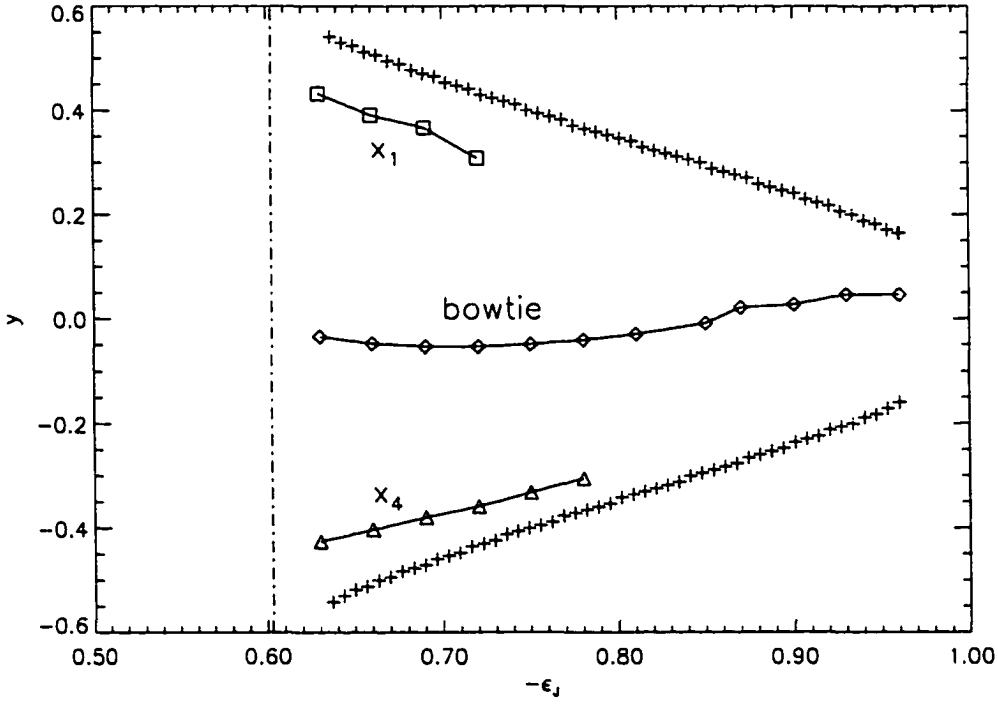

Figure 4.8: A characteristic diagram for three families of orbits in the equatorial Cazes bar potential. The dashed-dotted line at $-\epsilon_J = 0.603$ marks the value of the potential at the $L1, L2$ points. The + symbols represent the position along the y-axis where $\epsilon_J = \Phi_{CB}$. Bowtie orbits exist over the entire energy range that exists inside the Cazes bar. Near the bottom of the potential well ($-\epsilon_J \approx 0.85$), the periodic bow tie orbits become fully prograde. The x_4 and x_1 families exist over only a limited (higher energy) range.

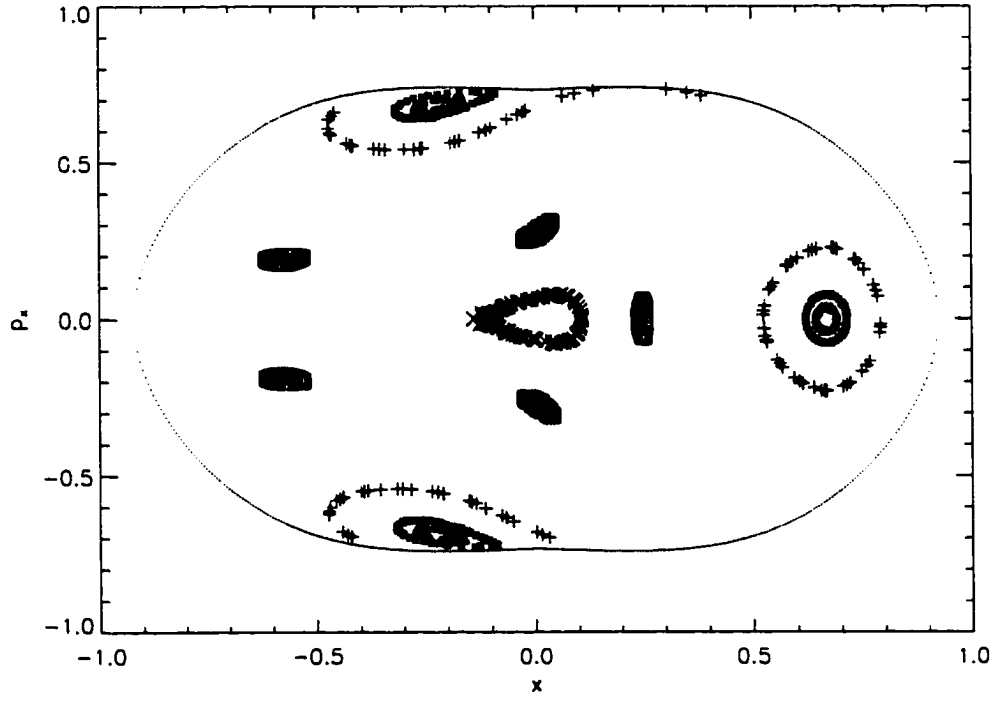

Figure 4.9: The (x, p_x) composite surface of section diagram for 5 selected regular orbits with $\epsilon_J = -0.75$ that are supported by the analytical Cazes bar potential described in §2.2.1. As before, the dotted line surrounding the invariant curves denotes the zero velocity boundary. This diagram should be compared with Fig. 4.1.

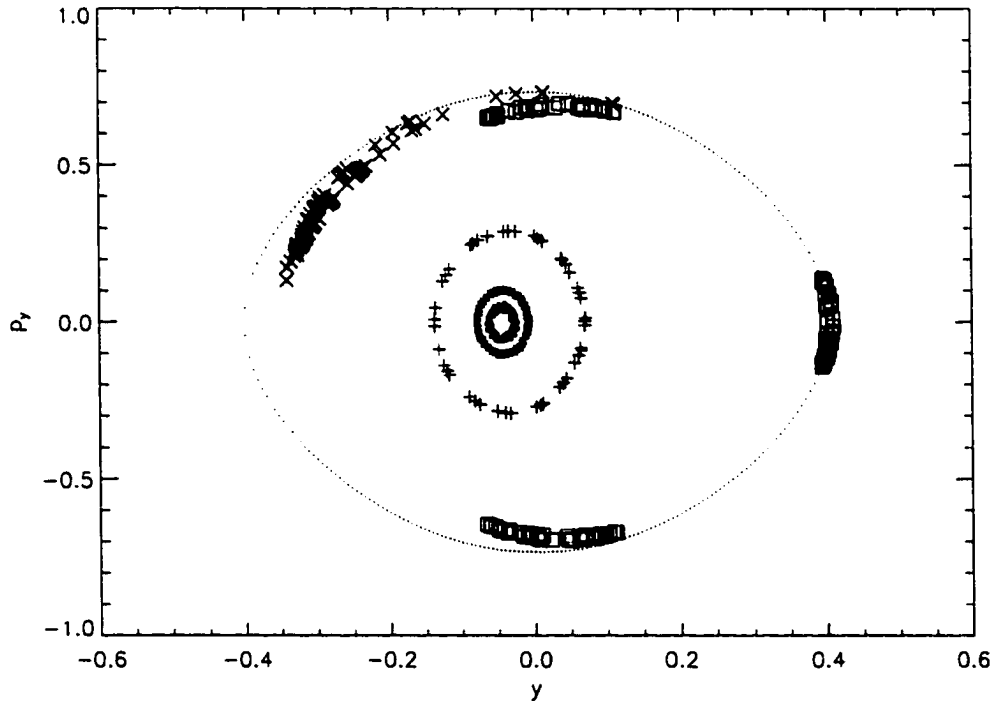

Figure 4.10: The (y, p_y) composite surface of section diagram for the same orbits represented in Fig. 4.9. Again, the dotted line is the zero velocity curve. This figure should be compared with Fig. 4.2.

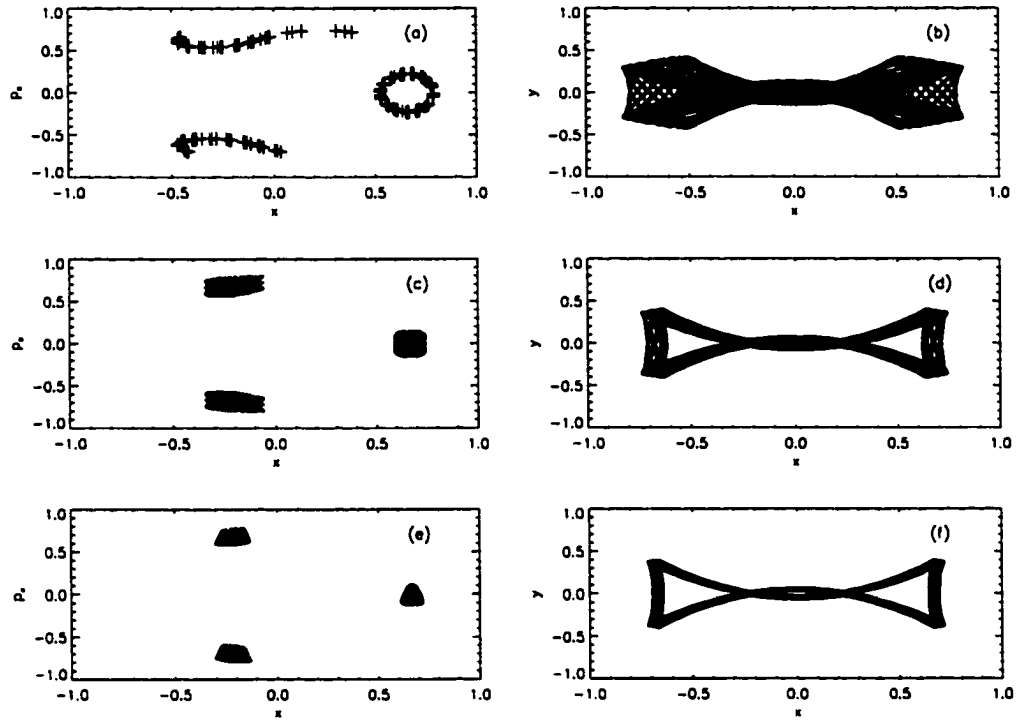

Figure 4.11: Plots of the 5 individual surfaces of section taken from the Fig. 4.9 composite diagram and their corresponding orbits. Surfaces of section are shown on the left, (a,c,e,g,i); orbits are on the right, (b,d,f,h,j). A comparison between this figure and Fig. 4.3 illustrates the degree to which the analytically specified effective potential discussed in §2.2.1 supports orbits that are like the orbits supported by $\Phi_{CB}(x, y)$.

generate Figs. 4.9 and 4.10 is indeed an appropriate model for the Cazes bar. Figures 4.11a-j show individual surfaces of section from the Fig. 4.9 composite diagram, along with the orbits that created them. The orbits shown in Figs. 4.11b, 4.11d, and 4.11f are very reminiscent of the orbits shown in Figs. 4.3b, 4.3d, and 4.3f, respectively. (The orbit equivalent to the one shown in Fig. 4.3h is not pictured.) Note, however, that the bow tie orbits now exhibit a reflection symmetry about the $y = 0$ axis. Also,

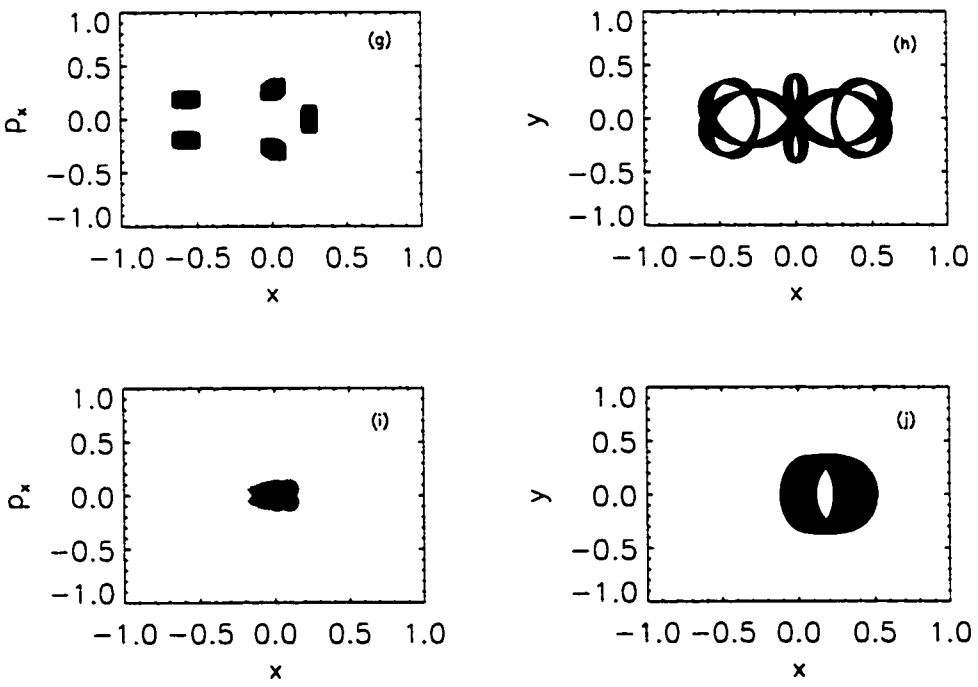

Figure 4.11: Continued.

attempts to find x_4 orbits in the rotating, analytical potential at this energy were unsuccessful. The small lobe marked by $\times$ symbols in Figs. 4.9 and 4.11i corresponds to a retrograde orbit that appears to be trapped in the shallow, off-axis minimum that lies along the positive x axis (see Fig. 2.4a).

Figure 4.11h shows a 5:1 orbit (in the $y:x$ oscillation notation introduced in §4.2.2) that resembles the 5 : 1 orbit that was found in the Cazes bar (Fig. 4.3j). The major difference between Fig. 4.11g (4.11h) and Fig. 4.3i (4.3j) is the reflection symmetry exhibited by the former and the lack of symmetry in the latter. However, the fact that these orbits correspond in overall appearance with those in Fig. 4.3 is further evidence (along with the similar appearances of Figs. 2.3c and 2.5c and the major and minor axis fits shown in Fig. 2.4) that the analytical potential closely matches the Cazes bar potential.

4.3.2 Twisted Analytical Cazes Bar Potential

In an effort to construct an analytical effective potential that supports orbits having all of the asymmetries seen in the equatorial Cazes bar orbits, a slight spiral twist has been added to the potential function given in eq.(2.2). Specifically, the chosen twisted potential has the form,

$$\Phi_{\text{eff}}(x',y') = N\left\{1-\left(1+\left(\frac{x'}{R_c}\right)^{\alpha}+\left(\frac{y'}{qR_c}\right)^{2}\right)^{-n/2}\right\}-\frac{1}{2}\Omega^2(x'^2+y'^2)+\Phi_{\text{min}}, \tag{4.1}$$

where

$$x' \equiv x\cos(a\sqrt{x^2+y^2}) - y\sin(a\sqrt{x^2+y^2})$$

and

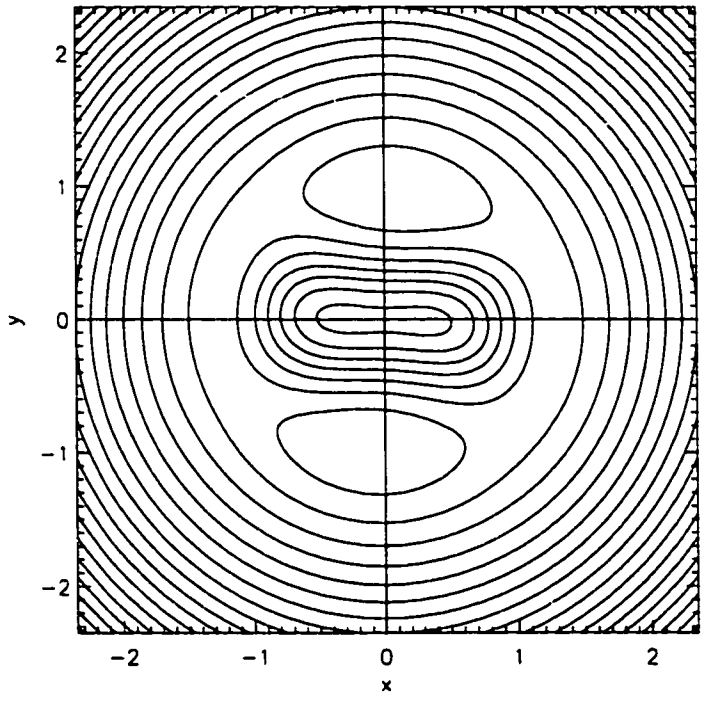

Figure 4.12: Equipotential contours of the twisted analytical potential that is defined by eq.(4.1). A comparison between this diagram and the one shown in Fig. 2.3c illustrates the degree to which the analytical "fit" matches $\Phi_{CB}(x, y)$.

$$y' \equiv x \sin(a\sqrt{x^2 + y^2}) + y \cos(a\sqrt{x^2 + y^2}).$$

For this potential, shown in Fig. 4.12, $a = 0.1$; otherwise the values of the parameters are the same as for the analytical Cazes bar potential discussed in §2.2.1.

The composite (x, p_x) surface of section for regular orbits with $\epsilon_J = -0.75$ that are supported by this twisted potential is shown in Fig. 4.13; Fig. 4.14 shows the corresponding composite (y, p_y) surface of section.

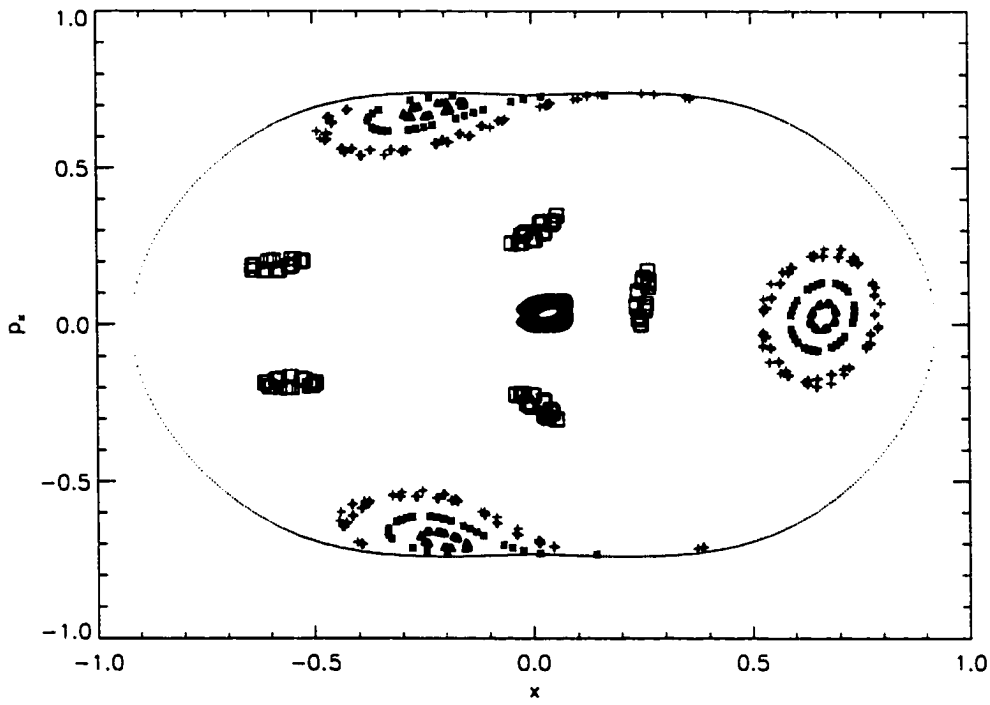

Figure 4.13: The (x, p_x) composite surface of section diagram for 5 selected regular orbits with $\epsilon_J = -0.75$ that are supported by the twisted analytical Cazes bar potential described in §4.2.2. As before, the dotted line marks the zero velocity curve. This diagram should be compared with Fig. 4.1.

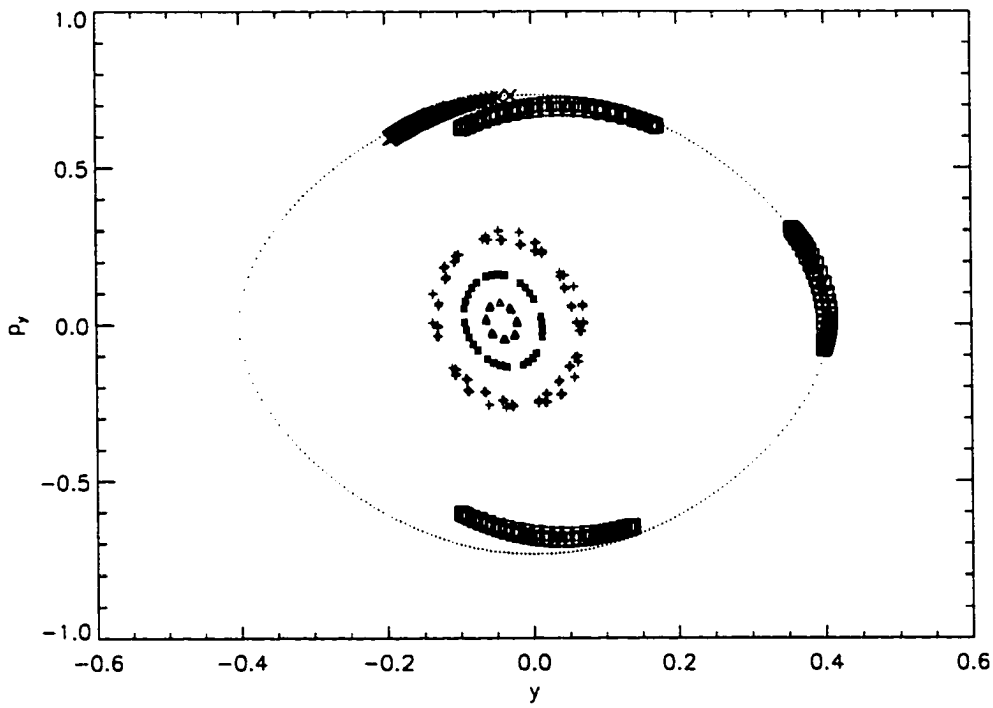

Figure 4.14: The (y, p_y) composite surface of section diagram for the same orbits represented in Fig. 4.13. The dotted line is the zero velocity curve. This figure should be compared with Fig. 4.2.

About 50% of orbits studied in this potential are quasi-ergodic. Notice that the overall composite surface of section bears a strong resemblance to that of the untwisted potential. The major difference lies in the symmetry of the surface of section. The reflection symmetry about the $p_x = 0$ (and $p_y = 0$) axis is now gone. An example of this feature is that the 3 islands (marked by triangles and asterisks) that are positioned symmetrically in the Fig. 4.9 surface of section diagram are twisted slightly from those positions in Fig. 4.13. The effect of the twisting of the potential on the orbits can be seen in Figs. 4.15a-j. The orbits in Fig. 4.15 resemble those from §3.2 even more closely than the orbits shown in Fig. 4.11 in that there is now an asymmetry due to the twisting. The effective potential given by eq. (4.1) appears to provide an excellent approximation to $\Phi_{CB}(x, y)$.

4.4 Restriction Hypothesis

Up to this point, many stellar orbits have been identified that, in principle, could be supported by the Cazes bar potential. In the context of the RH, the question becomes, "Which of these orbits would be populated by stars that form from the gas and, therefore, have initial velocities determined by the gas in the bar?"

4.4.1 Restriction Hypothesis Orbits

In order to maintain consistency between the discussion of the RH orbits and the previously discussed orbits, the focus is turned to orbits of stars that are created with a Jacobi constant $\epsilon_J = -0.75$. However, the method of choosing initial conditions as outlined in §3.2 must be abandoned. From the gas motions that are an integral part of the Cazes bar structure, there

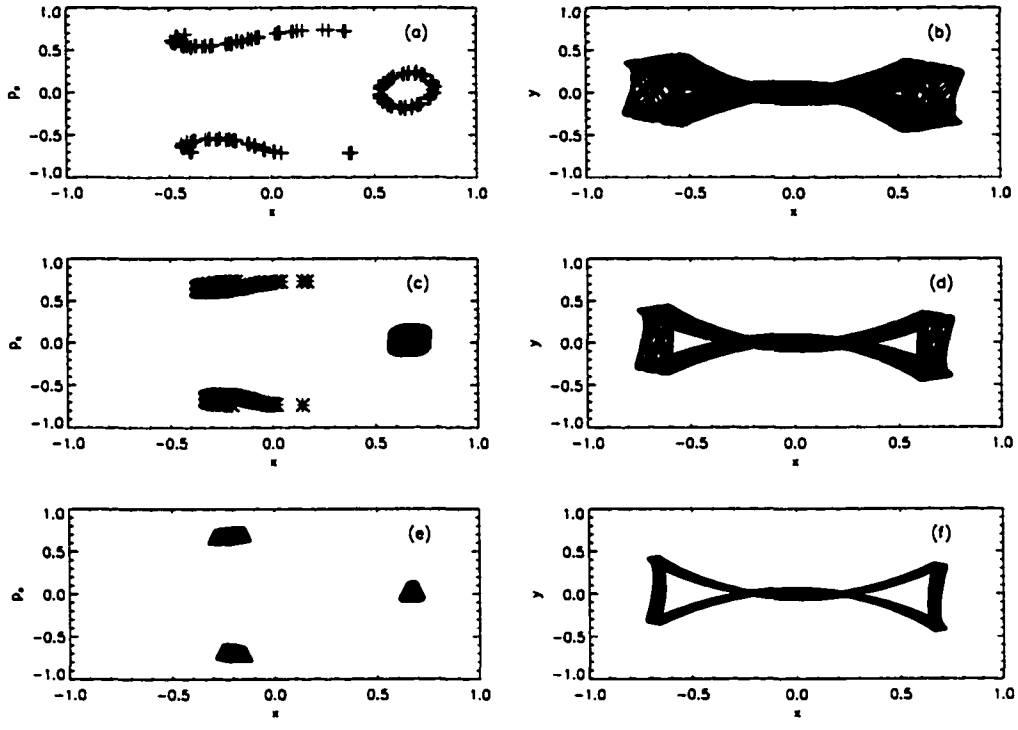

Figure 4.15: Plots of the 5 individual surfaces of section taken from the Fig. 4.13 composite diagram and their corresponding orbits. Surfaces of section are shown on the left, (a,c,e,g,i); orbits are on the right, (b,d,f,h,j). A careful comparison between these plots and the corresponding ones displayed in Fig. 4.11 shows that the slight spiral "twist" that has been added to eq.(2.2) in order to generate eq.(4.1) produces a "north-south" asymmetry like the one that arises in $\Phi_{CB}(x, y)$ (see Fig. 4.3).

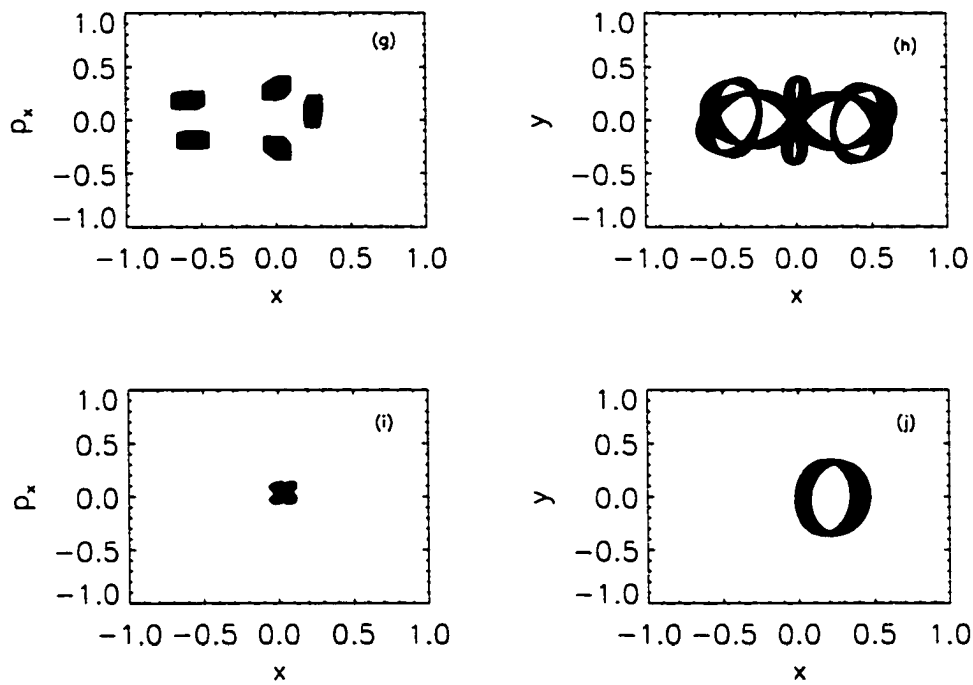

Figure 4.15: Continued.

are specific values of the velocity associated with each coordinate position in the bar. With this in mind, Fig. 4.16 shows contours of constant ϵ_J, where the known velocity of the gas has been used in the determination of ϵ_J at each (x, y) location. The dashed-dotted contour underlying the large assortment of symbols identifies at what locations in the bar stars could form with $\epsilon_J = -0.75$. Because these contours are dependent on the velocity field of the gas, the shocks mentioned in §2.1 become noticeable in Fig. 4.16, whereas they were not readily identifiable in the earlier plots of Φ_{CB} (Fig. 2.3c).

In order to investigate the behavior of an assortment of RH orbits, 30 particles have been positioned along the $\epsilon_J = -0.75$ contour, as shown in Fig. 4.16, and each has been assigned the velocity of the Cazes bar gas at that location. For discussion purposes, these particles are divided into two groups: one that begins on the positive side of the major axis and terminates where the $\epsilon_J = -0.75$ contour crosses the intermediate (y) axis; and one that begins near the intermediate axis and ends where this energy contour crosses the negative side of the major axis. The composite (x, p_x) surface of section diagrams that have been derived from the first and second groups are shown in Figs. 4.17 and 4.19, respectively. Corresponding (y, p_y) composite surface of section diagrams for the first and second groups are displayed in Figs. 4.18 and 4.20, respectively. The symbols marking the initial positions of the particles in Fig. 4.16 are the same symbols used to make the corresponding surfaces of section in Figs. 4.17, 4.18, 4.19, and 4.20.

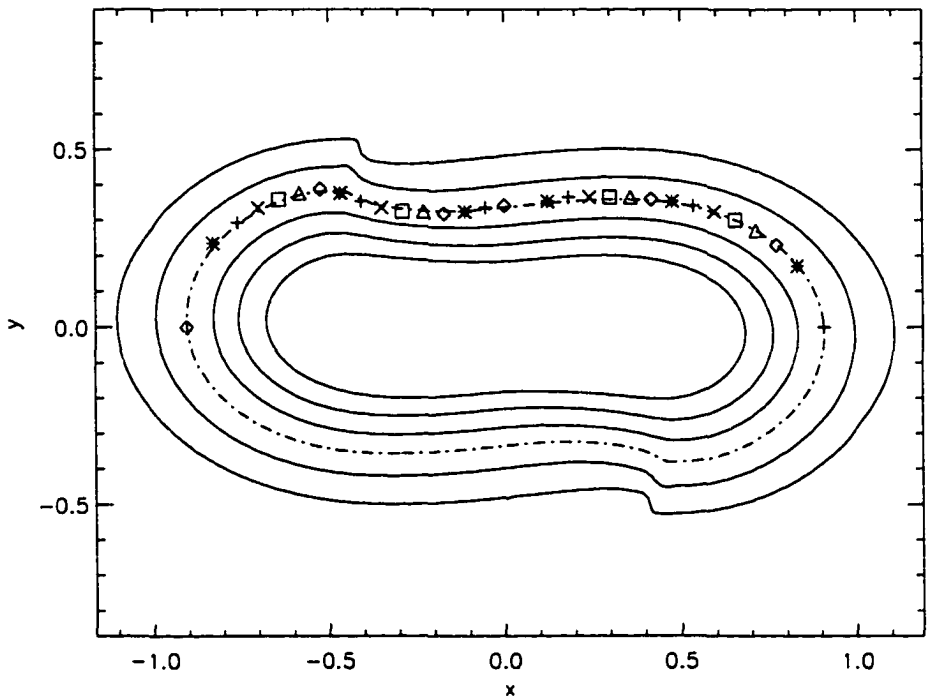

Figure 4.16: Contours of constant ϵ_J where, as discussed in §3.1, the value of ϵ_J at each coordinate position is given by eq.(3.2) with $\Phi_{\rm eff}(x,y) = \Phi_{\rm CB}(x,y)$ and the velocity components $(\dot{x},\dot{y})$ are specified by the velocity of the gas at each position in the steady-state Cazes bar. The dashed-dotted contour is for $\epsilon_J = -0.75$; the spacing between contours is 0.05. Also shown are the initial positions of 30 particles representing stars that form from the gas according to the "Restriction Hypothesis" with $\epsilon_J = -0.75$.

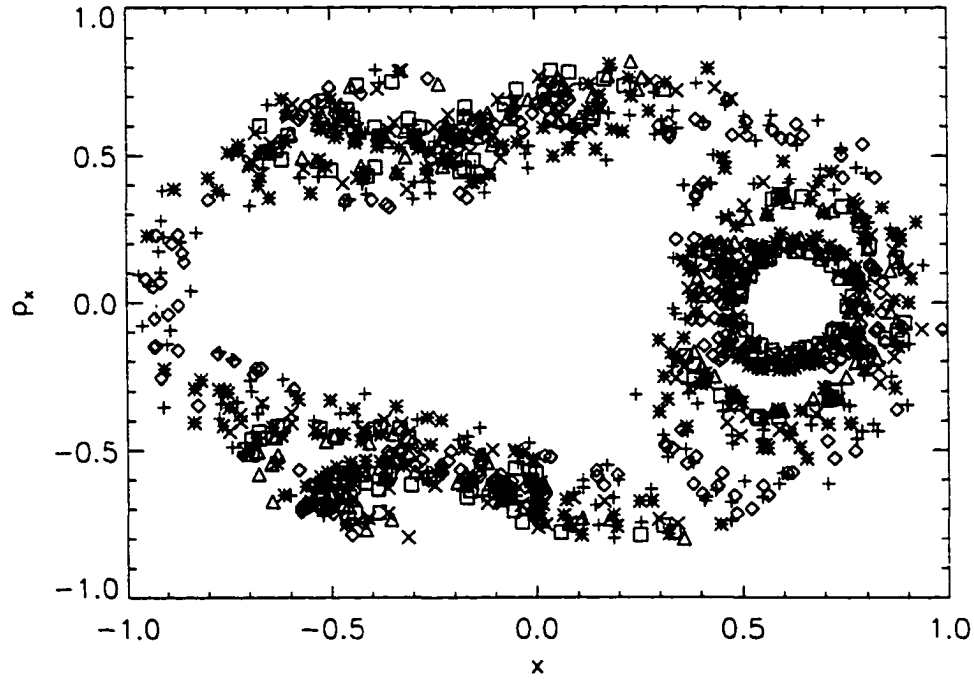

Figure 4.17: The (x, p_x) composite surface of section diagram that results from following the orbits of the "first" group of 15 particles in $\Phi_{CB}(x, y)$, as discussed in §4.4.1. For each "section", the initial particle position is identified by the corresponding symbol in the first quadrant of Fig. 4.16; the initial velocity is specified by the Cazes bar gas velocity at each location. This figure should be compared with Fig. 4.1 keeping in mind that quasi-ergodic orbits have been included here, whereas for clarity they were omitted in Fig. 4.1. Note that none of the retrograde orbits shown in Figs. 4.1 and 4.3 are populated under the RH.

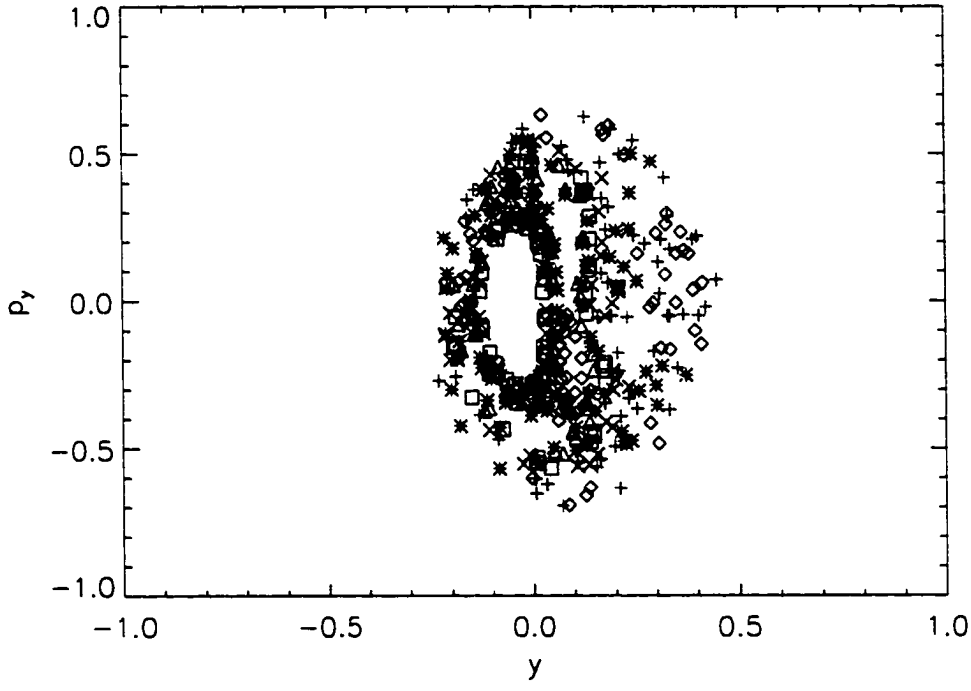

Figure 4.18: Complementing Fig. 4.17, this shows the (y, p_y) composite surface of section diagram generated by the orbits of the group of 15 particles identified in the first quadrant of Fig. 4.16. (See the caption to Fig. 4.17 for relevant details.) This diagram should be compared with Fig. 4.2.

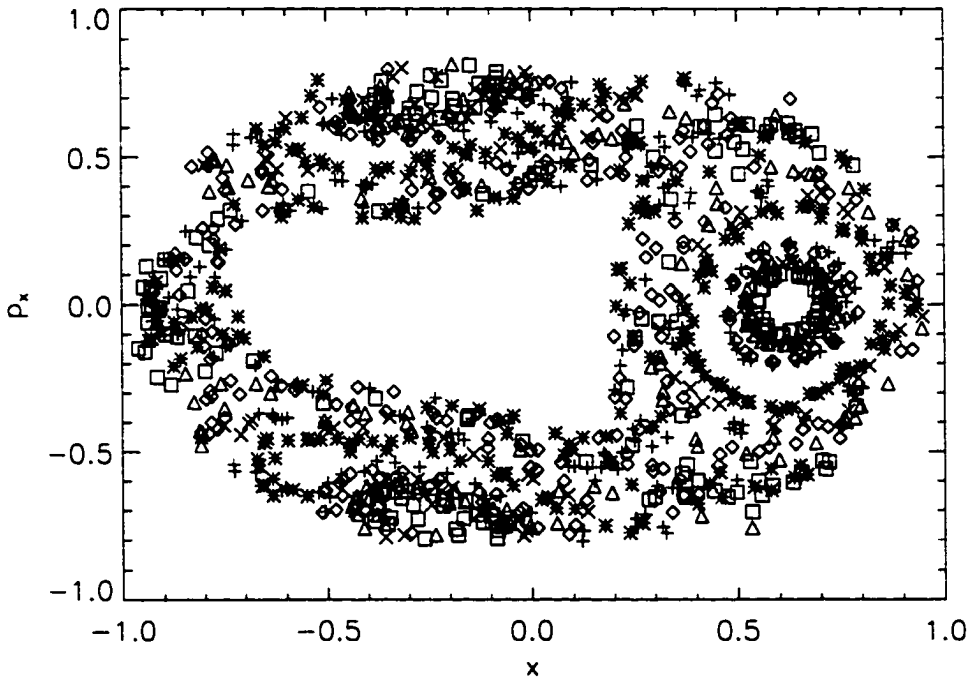

Figure 4.19: The same as Fig. 4.17, but for the "second" group of 15 particles identified in the second quadrant of Fig. 4.16.

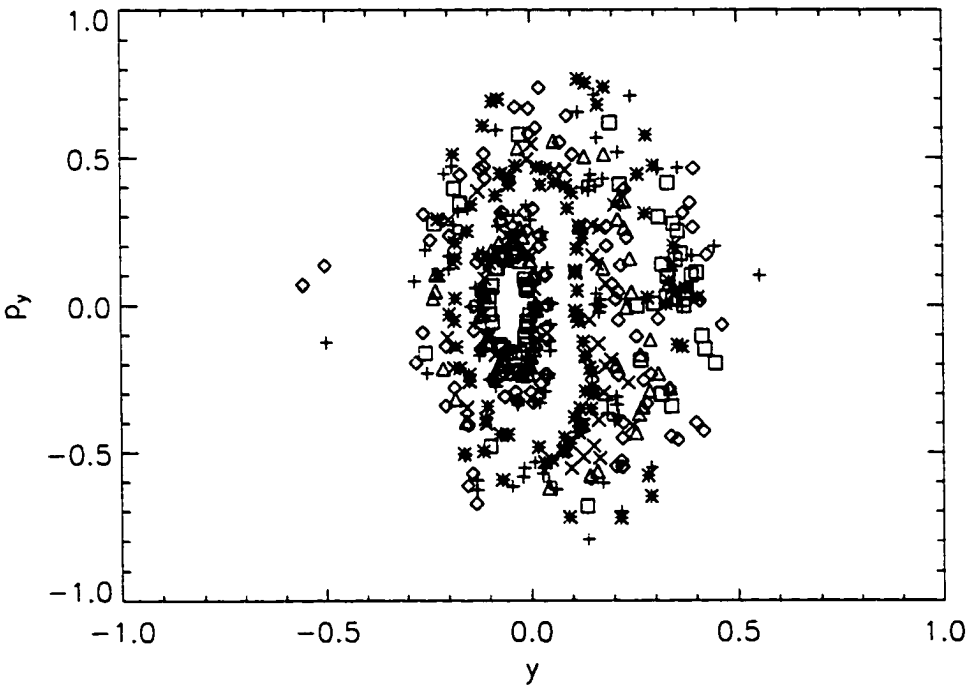

Figure 4.20: The same as Fig. 4.18, but for the orbits of particles identified in the second quadrant of Fig. 4.16.

By comparing these new figures with Figs. 4.1 and 4.2, an assessment can be made of the general impact of the RH. One of the most striking differences between these figures results simply from the fact that the surfaces of section of quasi-ergodic orbits have been included in Figs. 4.17 through 4.20, whereas the equivalent orbits were not displayed in Figs. 4.1 or 4.2. In this way it is clear at a glance that the RH permits a mixture of both regular and quasi-ergodic orbits to be populated. In particular, 9 of the 30 starting positions shown in Fig. 4.16, that is, $31 \pm 8\%$ using the statistics discussed in §4.2.2, of the RH particles produce quasi-ergodic orbits. It is also clear that the bow tie region of phase space is well-populated within the constraints of the RH. The holes that appear at the centers of the bow tie orbit "islands" in Figs. 4.17 through 4.20 falsely suggest that a strictly periodic bow tie orbit does not arise under the RH. Instead, these empty regions — and the analogous gaps that appear between some of the other regular orbit surfaces of section — arise because the spacing that has been chosen between initial particle positions in Fig. 4.16 was relatively coarse. With a finer spacing, these regions would have been filled by regular bow tie orbits, and two points, one for each group, would have been identified on the precisely periodic bow tie orbit. Most significantly, the composite surface of section diagrams (Figs. 4.17 and 4.19) that result from the RH present a large region of phase space that is completely unoccupied. This is the region that previously had been occupied by retrograde orbits. It is obvious, therefore, that under the constraints of the RH, no true retrograde

orbits are produced. This is perhaps not surprising, given that all of the gas in the Cazes bar is moving along prograde streamlines.

It is informative to study the sequence of orbits that appears as one moves to different starting positions along the $\epsilon_J = -0.75$ contour of Fig. 4.16, in a counter-clockwise fashion starting from the position marked by the plus symbol on the positive x-axis. This point on the major axis is the initial position for an orbit that is similar to that shown in Fig. 4.3b. Moving along the contour of constant ϵ_J, each successive initial position gives rise to bow tie orbits that are more and more closed. The fifth and sixth points (a square and a $\times$) have nearly identical orbits; they form the innermost three-island surfaces of section in Fig. 4.17 and the corresponding innermost curves of the bow tie region in Fig. 4.18. A particle starting from a position somewhere between these two points would probably trace the periodic bow tie orbit. Proceeding toward the minor axis, the sequence reverses and the orbits become less closed. The twelfth point (marked by a $\times$) is the origin for an orbit that is basically the same as that for the first point. The last three initial positions in this group, up to and including the point on the minor axis, produce quasi-ergodic orbits. These are the orbits that, for example, create the swarm of points that surround the three bow tie region islands in Fig. 4.17.

The second group of initial positions whose (x, p_x) surfaces of section are displayed in Fig. 4.19 and (y, p_y) surfaces of section are shown in Fig. 4.20 are now discussed. In the absence of the shock, it would be reasonable to assume that the progression seen in the first group would simply be reversed.

However, whereas only three positions nearest the minor axis gave rise to quasi-ergodic orbits in the first group, five positions nearest the minor axis lead to quasi-ergodic orbits in group two. The sixth position (marked by a $\times$) produces an orbit similar to the one shown in Fig. 4.3b and leads to a large three-island surface of section in Fig. 4.19. The next four points (ending with the triangle just after the shock) mirror the sequence in the first group by beginning bow tie orbits that become more and more closed. The eleventh point in the second group (marked by a square) is the initial position for the most closed orbit in both groups. This orbit forms the smallest three-island surface of section in Fig. 4.19 and the smallest curve in the bow tie region of Fig. 4.20. Continuing towards the major axis, the orbits become less closed. The point on the major axis (marked by a diamond) gives rise to an orbit that is the same as that orbit associated with the first point among the first group of points. Since the number of quasi-ergodic RH orbits is greater for the quadrant containing a shock, it seems that the presence of a shock (even one as mild as this) can influence orbital structure. The effects of shocks are examined more directly in the next section.

Observations in the solar neighborhood indicate that stars are born with velocities that have some dispersion about the mean motion of the gas. With this in mind, the properties of stellar orbits that begin with the velocity of the gas plus a modest random component have been examined. The results vary little from what has already been presented here under the strict RH, even if a random velocity of up to 30% of the magnitude of the initial velocity is added. For larger (but $\leq 50\%$) random velocities, orbits that were

bow tie-shaped with no perturbation maintain their basic morphology, but unperturbed quasi-ergodic orbits can be "kicked" into regular orbits.

4.4.2 Orbits Originating Near Shocks

If stars were to form from the gas with equal probability at all locations throughout the Cazes bar, then the distribution function of stars that would be created at a Jacobi constant $\epsilon_J = -0.75$ would contain a uniform mixture of all the orbits discussed in §4.4.1. For example, most would be quasi-periodic and 30%, by number, would be quasi-ergodic. In real barred galaxies, however, one usually does not find that star formation occurs at a uniform rate throughout the entire volume of the bar. In particular, the star formation rate usually is higher in the vicinity of a shock (Binney & Merrifield 1998, §5.1.8). Since the Cazes bar model contains shocks, this is an ideal opportunity to examine how such a process would impact the resulting distribution function of newly formed stars.

In order to model this scenario, four groups of 15 particles are placed in the vicinity of the shock structure that is evident in the fourth quadrant of Fig. 2.1c; see Fig. 4.21 for details regarding the distribution of these particles. These particles were given the gas velocity corresponding to their initial positions, according to the RH. Note that, as described in §2.1.2, the shock becomes stronger as the distance from the major axis increases, but along most of its length, the shock is relatively weak. More specifically, using the initial particle positions in Fig. 4.21 as a guide: the diamond located at $y \approx -0.31$ identifies the contour level at which the shock front officially begins (Mach number 1.0); the flow reaches Mach 1.5 between the square and

triangle at $y \approx -0.45$; and at the lowest density contour shown, the Mach number is approximately 2.0. These particles no longer share a common value of ϵ_J. Hence, individual surfaces of section for regular orbits are likely to overlap and it becomes much less useful to produce composite surface of section diagrams. For this reason, only individual surfaces of section will be discussed here.

From this entire group of 60 initial particle positions, only the three particles in the second column and farthest from the x-axis (marked by a plus symbol, asterisk, and diamond) follow quasi-ergodic orbits. All other particles follow regular (quasi-periodic) orbits. Focusing, then, on this second column of particles, the particle that began farthest from the x-axis (a plus symbol in Fig. 4.21) creates the (x, p_x) surface of section shown in Fig. 4.22a, and the corresponding orbit shown in Fig. 4.22b. Moving progressively closer to the $x-$axis, most of the particles trace orbits that have the general bow tie shape. For example, particles starting from the positions marked by the square ($y_i \approx -0.43$) and the asterisk ($y_i \approx -0.34$) generate the orbits shown in Figs. 4.22d and 4.22f. The orbits shown in Figs. 4.22h and 4.22j are followed by particles that are deep in the central region of the potential well [initial positions marked by the plus ($\epsilon_J = -0.899$) and diamond ($\epsilon_J = -0.953$) symbols that appear closest to the major axis in Fig. 4.21]. These orbits are quite thin and have a strong overall bar shape. (Note that these orbits do not appear to be thin in the figures because the vertical axis has been expanded in order to reveal more orbit details.)

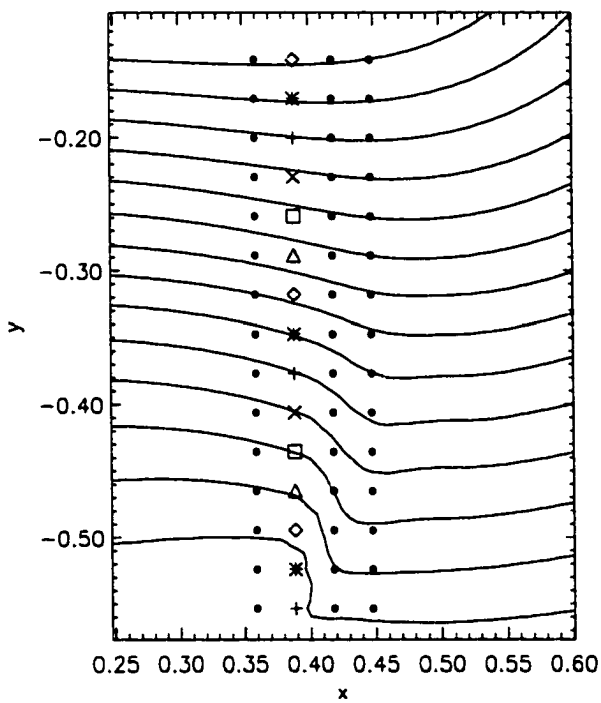

Figure 4.21: Initial positions of the four groups of 15 particles placed in the vicinity of the shock that is present in the fourth quadrant of the equatorial Cazes bar, superimposed on isodensity contours showing the fourth quadrant structure of the bar. The Jacobi constants for the particles that are marked by the vertical column containing a variety of symbols are as follows (from most negative y to least negative y): -0.617, -0.613, -0.631, -0.653, -0.676, -0.699, -0.724, -0.751, -0.782, -0.814, -0.843, -0.870, -0.899, -0.927, -0.953. See Fig. 2.1c for a less magnified view of this region.

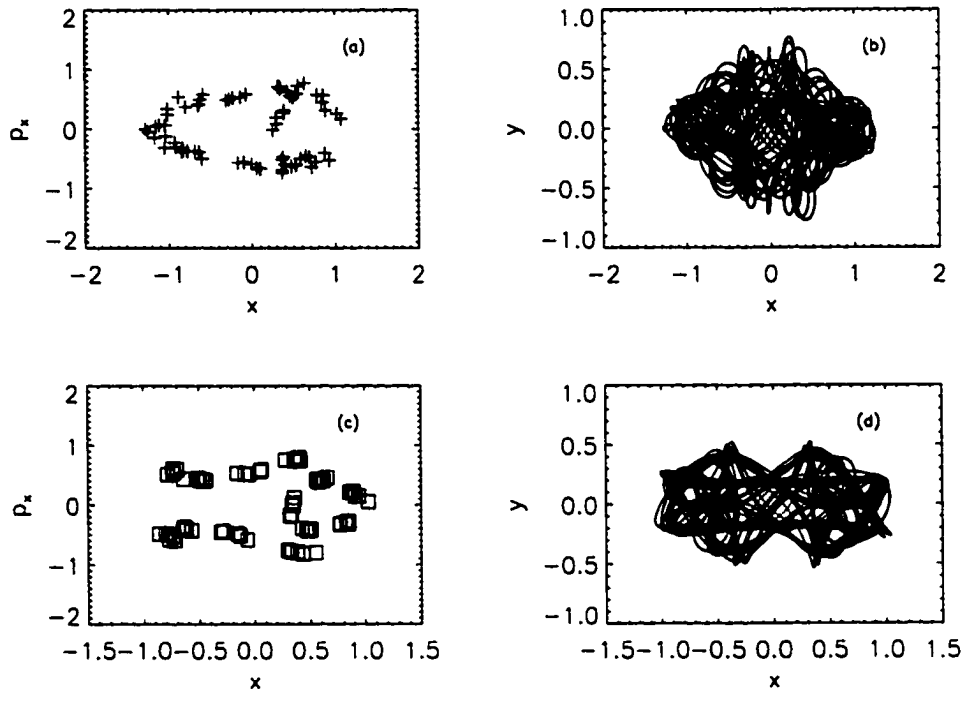

Figure 4.22: Plots of the (x, p_x) surfaces of section and corresponding orbits produced in $\Phi_{CB}(x, y)$ by 5 of the 15 particles whose initial positions are shown in Fig. 4.21. Surfaces of section are on the left, (a,c,e,g,i); orbits are on the right, (b,d,f,h,j). In each case, the symbol used in the surface of section matches the symbol used to mark the corresponding particle's initial position in Fig. 4.21. As discussed in §5.2, the initial particle velocity is specified by the Cazes bar gas velocity at the particle's initial position, as prescribed by the RH.

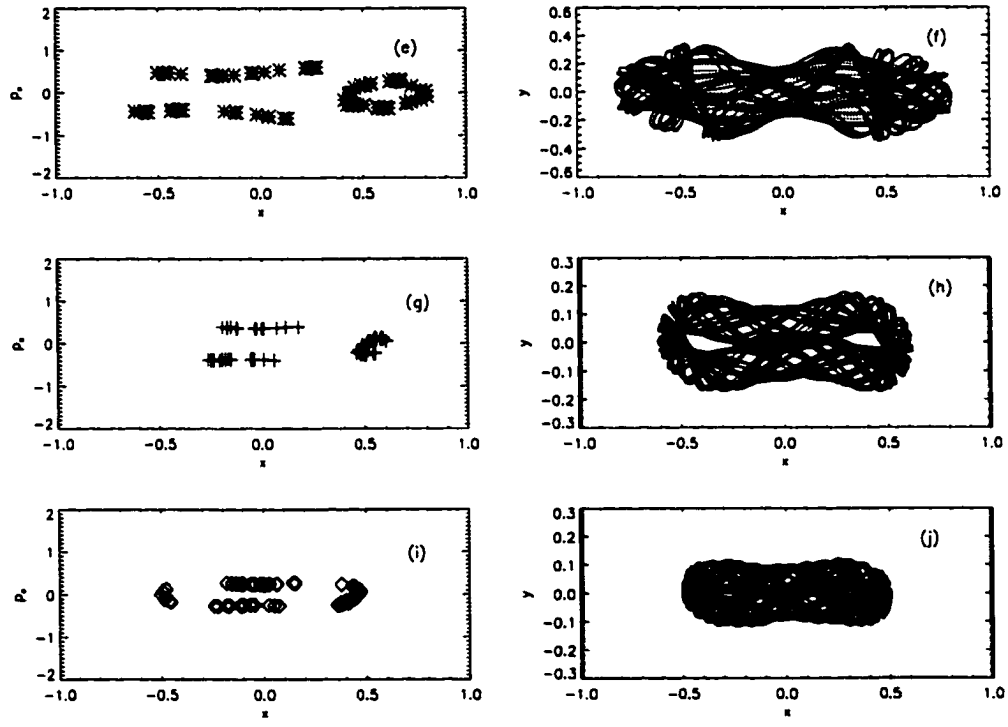

Figure 4.22: Continued.

It seems that the presence of the quasi-ergodic orbits is connected to the large velocities that are present in the gas that is located immediately before the shock. Since bow tie orbits have turning points in the vicinity of this shock, stars that are created with small x-velocities in this region (i.e., from the post-shock gas) have a better chance to fall onto such an orbit than do stars that are created with a large x-velocity (i.e., from pre-shock gas). Basically, these high velocity stars are shot through the region occupied by bow tie orbits and onto the only other available trajectories, that is, quasi-ergodic orbits. Hence, the presence of the shock does influence the trajectories onto which stars will be injected according to the RH, but in such a way that stars which form from the post-shock gas are unlikely to end up on quasi-ergodic orbits.

4.4.3 Active Galactic Nuclei Fueling

One interesting side note to the results presented in this chapter involves active galactic nuclei (AGN). One of the most exciting areas in current astronomy is the development of an accurate theoretical model to describe AGN. An incredibly large range of AGN lie at the centers of many different types of galaxies. For example, AGN known as Seyferts appear in spiral galaxies (Seyfert 1943), while radio galaxies are most commonly associated with elliptical galaxies (Spinrad *et al.* 1985). The different species of AGN are differentiated in a variety of ways; qualities of spectral lines, total luminosity, and/or luminosity variability. The fact that many AGN are variable on very short astrophysical timescales, $\tau_{\rm var} < 1\,{\rm yr}$, is evidence that the luminosity must be produced in astrophysically compact regions,

$r \approx c\tau_{\rm var} \approx 10^{-2}$ parsecs. Currently, the most likely candidate to be the engine of AGN is a supermassive ($10^{7-8} M_\odot$) black hole that is being fed by an accretion disk. In what has become known as the unified model of AGN, the mass of the black hole regulates the total luminosity of the AGN while the geometry and orientation of the accretion disk determines the AGN's observed properties (Binney & Merrifield 1998, §4.6.4).

The energy output of AGN is related to the amount of mass that is accreted by the black hole in the following way;

$$\dot{E} = \alpha \dot{M} c^2, \tag{4.2}$$

where $\dot{E}$ is the AGN luminosity, α is an efficiency coefficient, $\dot{M}$ is the mass transfer rate, and c is the speed of light. The quantity of interest here is $\dot{M}$. In the equatorial Cazes bar, bow tie orbits can transport mass very close to the center of the potential. In a very rough calculation, a mass transfer rate can be determined by calculating the amount of mass per unit time on bow tie orbits that comes within the tidal radius of a black hole that is placed at the center of the Cazes bar. Once tidally disrupted, all of the mass is considered to be transferred to the black hole. In this simple picture, the mass transfer rate is determined by two factors: one, the fraction of bow tie orbits; and two, the formation rate of pre-stellar clouds on bow tie orbits. With the fraction of RH bow tie orbits ($\approx 50\%$) and a representative AGN luminosity, $\dot{E} = 10^{39}$ J/s, the necessary cloud formation rate is $\dot{M} = 1\, M_\odot$/yr. The fact that this number is on the order of the star formation rate for galaxies is suggestive. More detailed work may show this mechanism to be important for AGN fueling.

5. ORBITAL CHARACTERIZATION

5.1 Background

The study of orbits in 3D gravitational potentials is of broad astrophysical interest. Whether the orbits under investigation are those of asteroids in the solar system (e.g., Pilat-Lohinger *et al.* 1999), stars in a globular cluster or a galaxy (e.g., Carpintero *et al.* 1999), or galaxies in a cosmological simulation (e.g., Colpi *et al.* 1999), methods of characterizing the orbits can be extremely useful, especially when an attempt is made to understand what the "typical" behavior is of a very large collection of orbits in a particular physical system. These methods are most useful when they are quantitative, reliable, and applicable to a wide variety of systems.

While the importance of regular orbits (and therefore integrals of motion) has been the focus of many stellar dynamics research projects, the importance and likely existence of quasi-ergodic orbits in realistic galaxy potentials also has been discussed (Goodman & Schwarzschild 1981; Binney 1982b; Habib *et al.* 1997; Valluri & Merritt 1998). While the previously mentioned studies involved analytical potentials, Barnes & Tohline (2001) have also found that quasi-ergodic stellar orbits are quite common in a numerically-created, gaseous bar. Quasi-ergodic orbits also appear in N-body simulations of galactic bars (see, for example, Sparke & Sellwood 1987). A specific example of a situation in which quasi-ergodic orbits play an important role is in galaxy models with massive central objects (e.g., Udry & Pfenniger 1988, Valluri

& Merritt 1998). In a detailed analysis of the influence that massive central objects have on stellar orbits, Gerhard & Binney (1985) suggest that sufficiently massive central objects will scatter into quasi-ergodic orbits trajectories that initially: 1) are regular; 2) pass near the central object; and 3) support nonaxisymmetric distributions, such as bars. With the current interest in understanding active galactic nuclei as well as the recent discovery of a link between the masses of central objects and central stellar velocity dispersions (Gebhardt *et al.* 2000; Ferrarese & Merritt 2000), more accurate modeling of the distribution functions of such systems could provide some insight into the observational evidence.

In an effort to better characterize orbits (both regular and quasi-ergodic) in astrophysically interesting potentials, the relative utility of the correlation integral method has been examined. As pointed out by Grassberger & Procaccia (1983) and Carnevali & Santangelo (1984), the correlation integral method is a flexible and accurate characterization technique that has been widely utilized by physicists but has been virtually ignored by the stellar dynamical community. This technique distinguishes orbits based on the number of isolating integrals of motion that are respected by any given orbit. This makes it a useful, quantitative tool for examining stellar distribution functions that would arise in potentials such as the ones mentioned above. Regular and nonregular orbits can be differentiated. Within the regular class, periodic and quasi-periodic orbits can also be distinguished from one another. Additionally, this technique can distinguish between 3D orbits that respect two isolating integrals and those that respect only one. This ability to dis-

tinguish in a clear and quantitative fashion between various types of orbits makes the correlation integral method particularly well suited to involvement in studies of 3D potentials where quasi-ergodic orbits are likely to play an important role. The primary objective of this chapter is to demonstrate the quantitative reliability and broad utility of the correlation integral method by applying it as a characterization tool to orbits in a wide variety of potentials – most of which are familiar to the stellar dynamics community – and by comparing and contrasting it to other techniques that have been broadly used to characterize orbits. This will place the community in a position to effectively utilize this tool in a wide variety of problems; hopefully, others will be motivated as well to adopt it as a useful tool.

5.2 Selected Orbits

The characterizations of phase space orbits discussed in this chapter are given in Table 5.1. As indicated in the first column of Table 5.1, the orbits are identified by the potential (or mapping) in which they exist. Listed in column 2 are the corresponding figure numbers that contain visual summaries of the orbital analysis. The last two columns of Table 5.1 are provided as a quick reference to the Lyapunov exponent and correlation integral characterizations that have been determined for each orbit. If the next to last column (labeled Lyapunov) contains an 'I' (denoting an orbit insensitive to small changes in initial conditions), that orbit is regular; an 'S' (for sensitive) denotes an orbit that is not regular. The last column holds the measured dimensionality D of each phase space orbit. The initial conditions (positions, velocities, and energies) for each of the orbits are listed in Appendix A. Before undertaking a

Table 5.1. Orbit Characterizaton Information

Orbit	Fig. #	ϵ or ϵ_J	Lyapunov	D
Hénon map	5.1	...	1.21	
2D Richstone #1	5.2	-0.574	I	2
2D Richstone #2	5.3	-0.614	I	2
2D Richstone #3	5.4	0.0	I	2
2D Richstone #4	5.5	-4.56(-2)	I	1
Hénon-Heiles #1	5.6	0.125	I	1
Hénon-Heiles #2	5.7	0.125	I	2
Hénon-Heiles #3	5.8	0.125	S	2.4
Hénon-Heiles #4	5.9	0.125	S	2.3
2D Cazes bar #1	5.10	-0.75	I	1
2D Cazes bar #2	5.11	-0.75	I	2
2D Cazes bar #3	5.12	-0.75	I	2
2D Cazes bar #4	5.13	-0.75	S	2.5
3D Richstone #1	5.14	-0.269	I	3
3D Richstone #2	5.15	-0.263	I	3
3D Richstone #3	5.16	-0.229	I	3
3D Richstone #4	5.17	-0.389	I	2.9
3D Richstone #5	5.18	-0.461	I	2.9
3D Cazes bar #1	5.19	-0.922	I	3
3D Cazes bar #2	5.20	-0.909	I	3
3D Cazes bar #3	5.21	-0.658	I	3
3D Cazes bar #4	5.22	-0.674	I	2.7
3D Cazes bar #5	5.23	-0.681	S	4
3D Cazes bar #6	5.24	-0.695	S	4.3
3D Cazes bar #7	5.25	-0.645	S	4.3

discussion of the characterization of these selected orbits, let us examine the expectations for the various models. The 2D phase space of the Hénon map is not derived from a Hamiltonian system, so there are no integrals of motion for this system *per se*. However, each of the 4D phase spaces (2D Richstone, Hénon-Heiles, and 2D Cazes bar) is a Hamiltonian system. In these cases, the potentials are time-invariant so the energy (or ϵ_J) must be an integral, and we should find that $D \leq 3$ for all orbits. Fully ergodic orbits will have $D = 3$, while quasi-ergodic orbits will have $2 < D < 3$. Quasi-ergodic orbits respect only one full integral of motion, but they also have some (unknown) restriction in addition to that of their ergodic cousins. Both ergodic and quasi-ergodic orbits are sensitive (S) to initial conditions, so the Lyapunov exponent should remain nearly constant with time, as discussed in §3.4. On the other hand, a regular orbit must have a total number of isolating integrals that is equal to or greater than the number of spatial degrees of freedom (in this case ≥ 2, (i.e., $D \leq 2$); specific energy (or ϵ_J) plus 1 or 2 unknown integrals). So, the allowed phase space for each regular (but not periodic) orbit should be 2D (4 phase space dimensions minus the 2 isolating integral dimensions). The behavior of the Lyapunov exponent should reflect the fact that a regular orbit is insensitive (I) to initial conditions, that is, the slope of $\ln k_n$ vs. $\ln n\tau$ should be ≈ -1.

Orbits in 3D Hamiltonian systems occupy 6D phase spaces. However, since the potentials are still time-invariant, the measured dimensionality of the phase space orbit should be, at most, five. Regular orbits in 3D potentials must respect at least three integrals of motion, $D \leq 3$, and as before, the

Lyapunov exponent should decrease with time for a regular orbit. Returning to the nonregular orbits, fully ergodic orbits should display a dimensionality $D = 5$, while quasi-ergodic orbits should have a noninteger value $4 < D < 5$. Orbits with $3 < D \leq 4$ are neither regular nor quasi-ergodic. They respect (at least) two full integrals of motion and will be referred to as semiregular orbits in this work. All nonregular orbits should have Lyapunov exponent plots that have $\ln k_n$ vs. $\ln n\tau$ slopes that are ≈ 0.

5.3 2D Phase Space

As mentioned earlier, the Hénon map phase space is fractal (see §2.2.4). The accepted value for the dimensionality of this phase space orbit is $D = 1.25 \pm 0.02$ (Grassberger & Procaccia 1983). Figure 5.1a shows the (x, y) phase space structure of the Hénon map; Fig. 5.1b is a magnified view illustrating the fractal nature of this phase space orbit. Figure 5.1c shows the $\ln C(r)$ vs. $\ln r$ plot (hereafter, the $C - r$ plot) that was obtained in this study for the Hénon map using the correlation integral method. The result is $D = 1.21 \pm 0.01$. This is also the value that Grassberger & Procaccia (1983) obtained using the correlation integral method. The difference between this result and the accepted value is discussed in their paper. They claim that there is a systematic error in the determination of the Hénon map correlation integral. However, their suggested remedy does not work well when phase space is populated by numerically integrated orbits. In the following sections it will be shown that this slight disagreement does not diminish the quantitative usefulness of the correlation integral method as a tool for characterizing orbits.

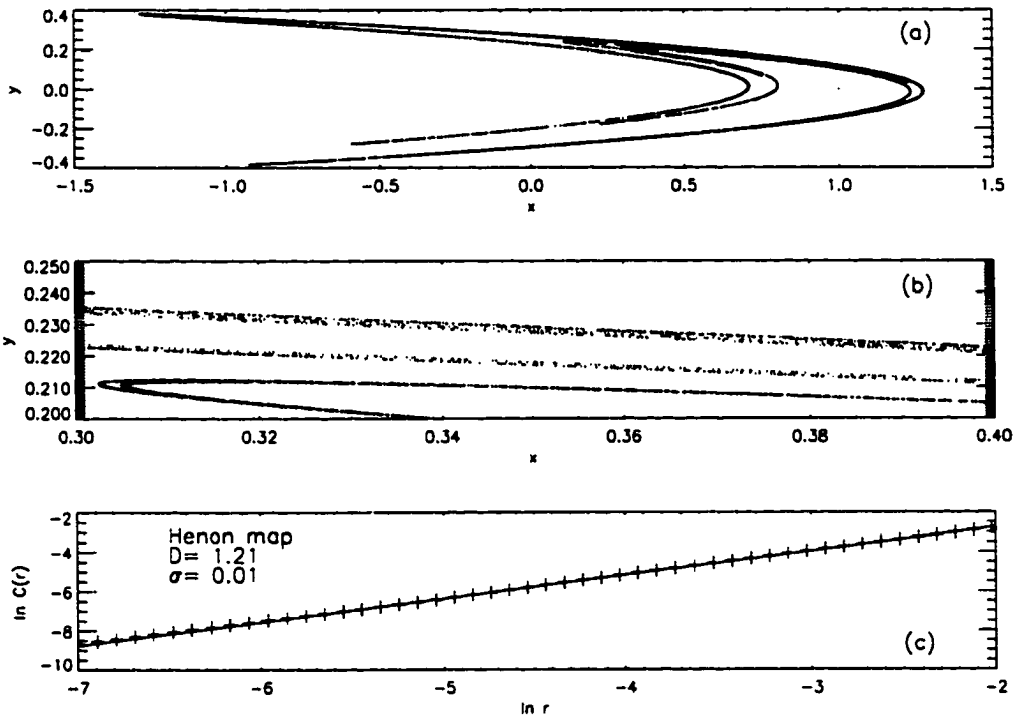

Figure 5.1: Hénon map. (a) Plot of phase space points obtained after 100,000 iterations of eqs. (2.6) and (2.7). (b) Magnification of the plot in (a). The small scale substructure is a telltale sign of the fractal nature of the phase space orbit. (c) A plot of the correlation integral ($\ln C(r)$) versus phase space distance ($\ln r$) for this mapping. The value of D shown in the legend has been measured from the slope of the line; σ is the standard deviation measured for D, as described in §3.5.

5.4 4D Phase Space

For orbits in 2D potentials, the presentation of results will adhere to the following form: In each figure, the frame labeled (a) contains the orbital trajectory; frame (b) displays the $C - r$ plot obtained using the correlation integral method; frame (c) shows the Lyapunov exponent plot as $\ln k_n$ vs. $\ln n\tau$; and frame (d) contains either the $(x - p_x)$ or $(R - p_R)$ surface of section diagram for the orbit. The legends of frames labeled (b) also will provide quantitative information derived from the $C - r$ plot as discussed in §3.5, namely the slope D and standard deviation σ. In most Lyapunov exponent plots, a dot-dashed line with a slope of -1 also is included because, as discussed in §3.4, the behavior of k_n should approach this slope as $n\tau \to \infty$ if an orbit is regular.

The results for four orbits in the 2D Richstone potential are shown in Figs. 5.2 through 5.5. In every case, all three techniques for characterizing the orbits indicate that the orbits are regular: the surface of section plots are composed of invariant curves; the Lyapunov exponent drops with time with a slope of minus one; and from the correlation integral, the measured dimensionality has a value ≤ 2. Note that the Lyapunov exponent plots in Figs. 5.2c and 5.5c are basically identical and therefore make no distinction between the closed and unclosed regular orbits. However, the difference is clearly illustrated by the differing slopes of the $C - r$ plots in Figs. 5.2b and 5.5b: the periodic orbit displays $D = 1$, instead of $D = 2$.

At this point, it is worth noting some general features of $C - r$ plots derived from orbits (of any dimensionality). As the value of r increases, the

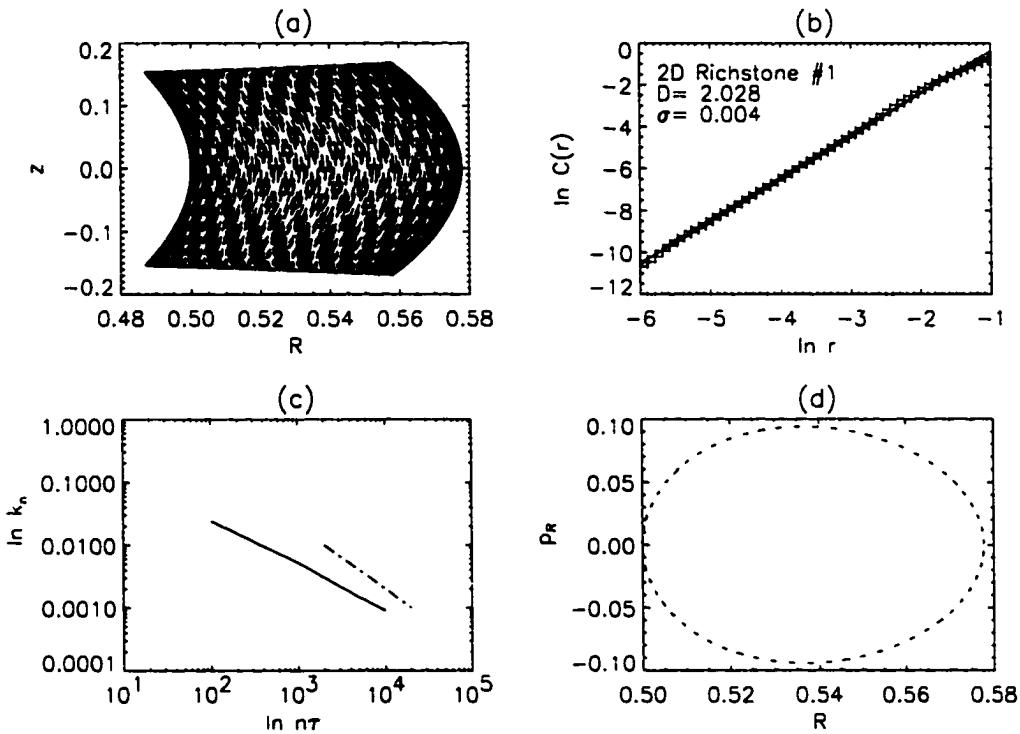

Figure 5.2: Orbit #1 in the 2D Richstone potential; see the Appendix for initial conditions. (a) The $R-z$ trajectory of this orbit. (b) The associated $C-r$ plot listing the average slope (D) and standard deviation (σ). (c) The plot of the Lyapunov exponent ($\ln k_n$) versus integration time ($\ln n\tau$) for this orbit. (d) The $(R-p_R)$ surface of section for this orbit.

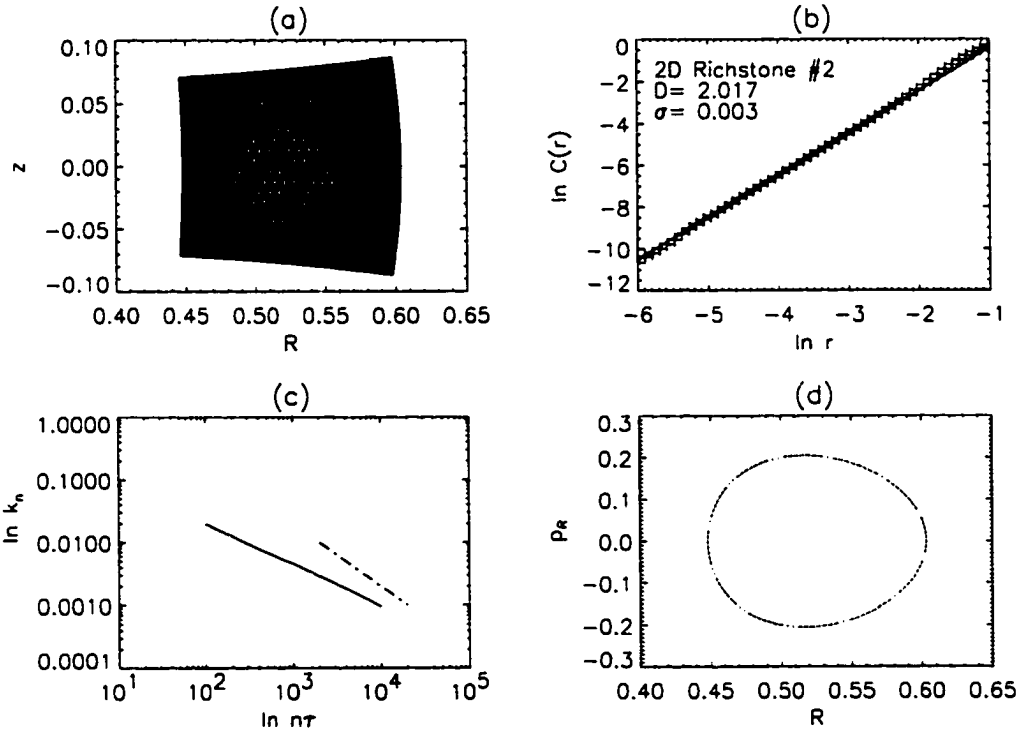

Figure 5.3: Orbit #2 in the 2D Richstone potential; each frame contains information as described in the caption to Fig. 5.2.

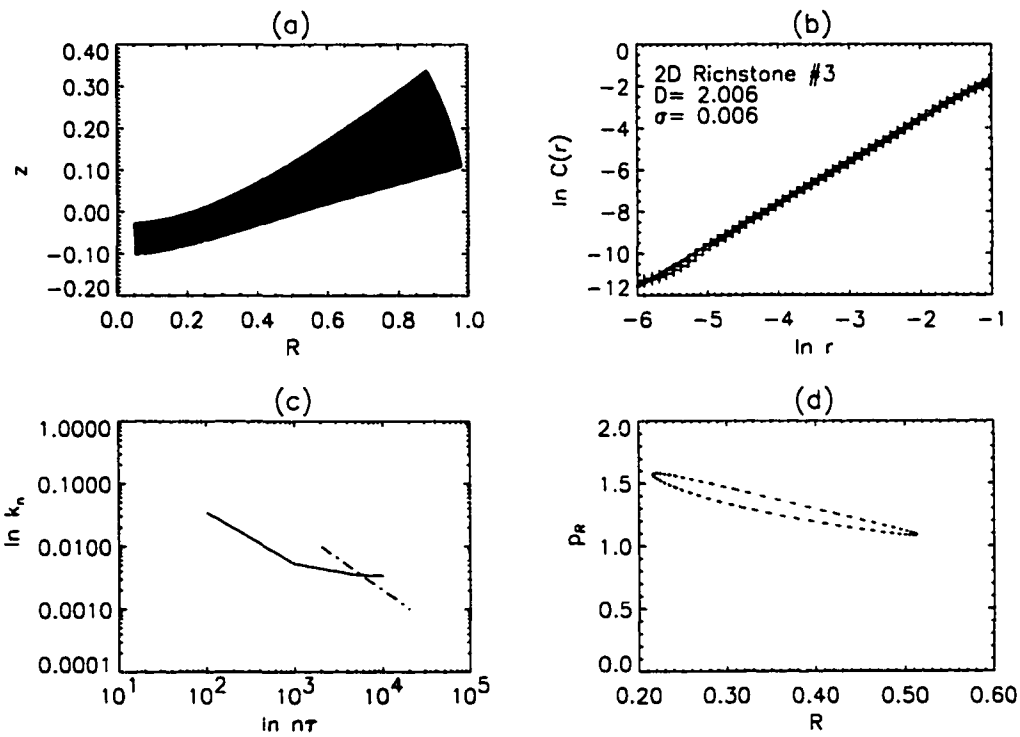

Figure 5.4: Orbit #3 in the 2D Richstone potential; each frame contains information as described in the caption to Fig. 5.2.

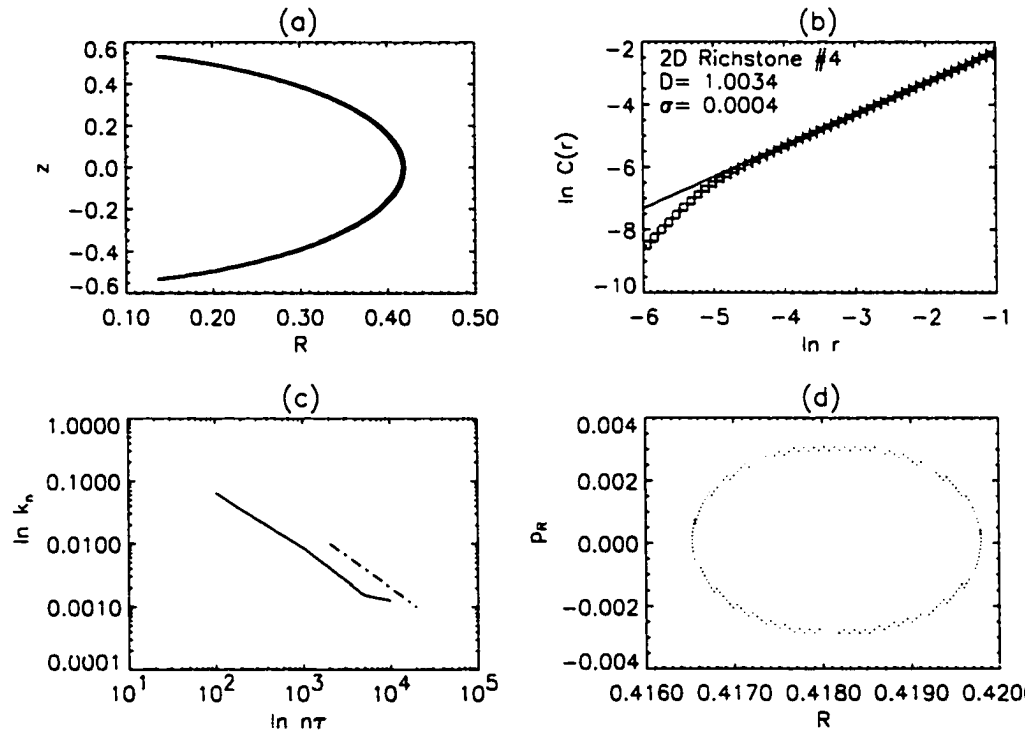

Figure 5.5: Orbit #4 in the 2D Richstone potential; each frame contains information as described in the caption to Fig. 5.2.

plotted points will often deviate from the line of slope D. This is because a clean linear relation is expected only for $r \ll 1$. Another way to think about this is that, for a sufficiently large value of phase space distance r_0, the entire phase space orbit will be enclosed. Then, for $r \geq r_0$, $C(r) = 1.0$, so the $C - r$ plot must reach a slope of zero at large enough values of r. Deviations from a linear slope of D may also occur at the smallest values of r, but for a different reason. Since only a finite number of sampling points is used, there is necessarily a lower limit to the smallest distance that can be measured between any two points. Basically, this is a small-number statistics problem. As a larger number of sampling points is used, the linear fit generally becomes tighter and extends to smaller values of r.

Four orbits with $\epsilon = 1/8$ in the Hénon-Heiles potential are displayed and analyzed in Figs. 5.6 through 5.9. Unlike the 2D Richstone potential, the Hénon-Heiles potential supports some quasi-ergodic orbits at this energy. All three methods of characterization indicate that the orbits shown in Figs. 5.6 and 5.7 are regular, while the ones shown in Figs. 5.8 and 5.9 are quasi-ergodic. Focusing on the two quasi-ergodic orbits, notice that, (1) the Lyapunov exponent is approximately constant, signaling an exponential departure of two initially neighboring trajectories; (2) the surface of section is no longer composed of a smooth invariant curve; and (3) the $C - r$ plot identifies a non-integer, fractal dimension greater than two. We know that the orbits shown in Figs. 5.8 and 5.9 are quasi-ergodic rather than fully ergodic because of the measured dimensionality. This is a distinction that can be made clearly from the $C - r$ plot, but it is not possible from a mea-

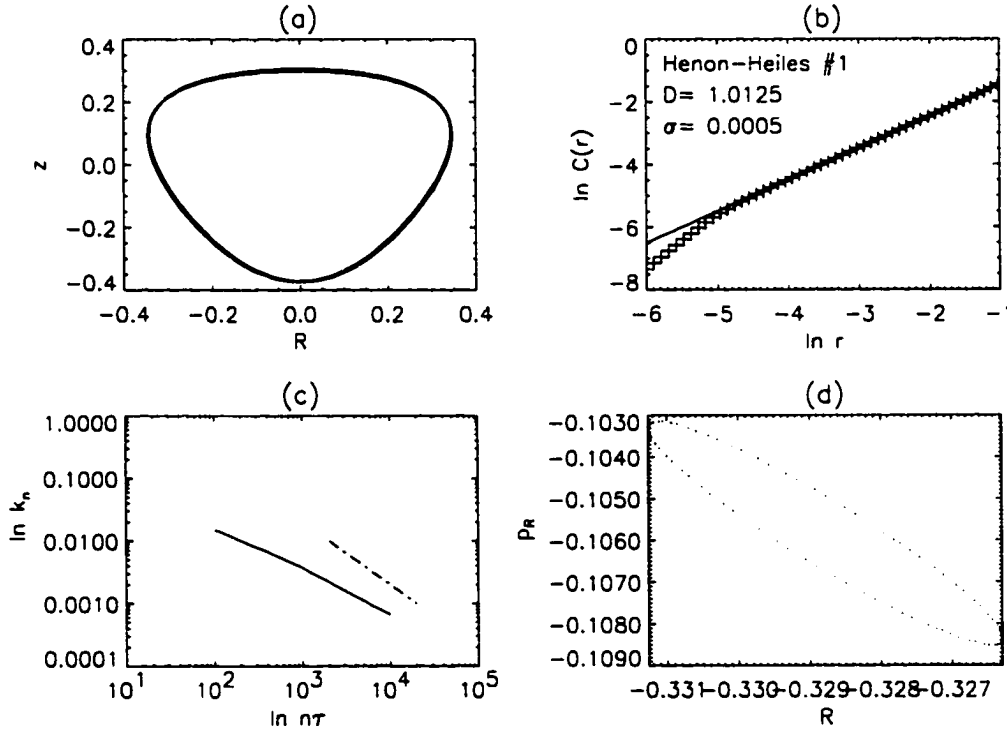

Figure 5.6: Orbit #1 in the Hénon-Heiles potential; see the Appendix for initial conditions. (a) The $R-z$ trajectory of this orbit. (b) The associated $C-r$ plot listing the average slope (D) and standard deviation (σ). (c) The plot of the Lyapunov exponent ($\ln k_n$) versus integration time ($\ln n\tau$) for this orbit. (d) The ($R-p_R$) surface of section for this orbit.

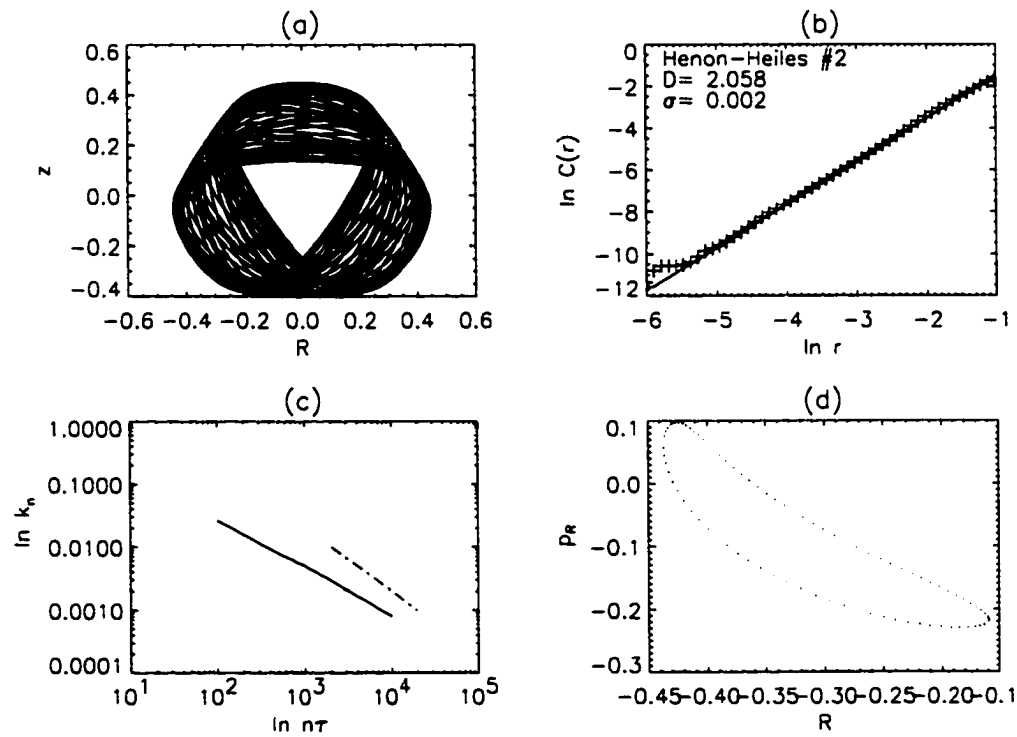

Figure 5.7: Orbit #2 in the Hénon-Heiles potential; each frame contains information as described in the caption to Fig. 5.6.

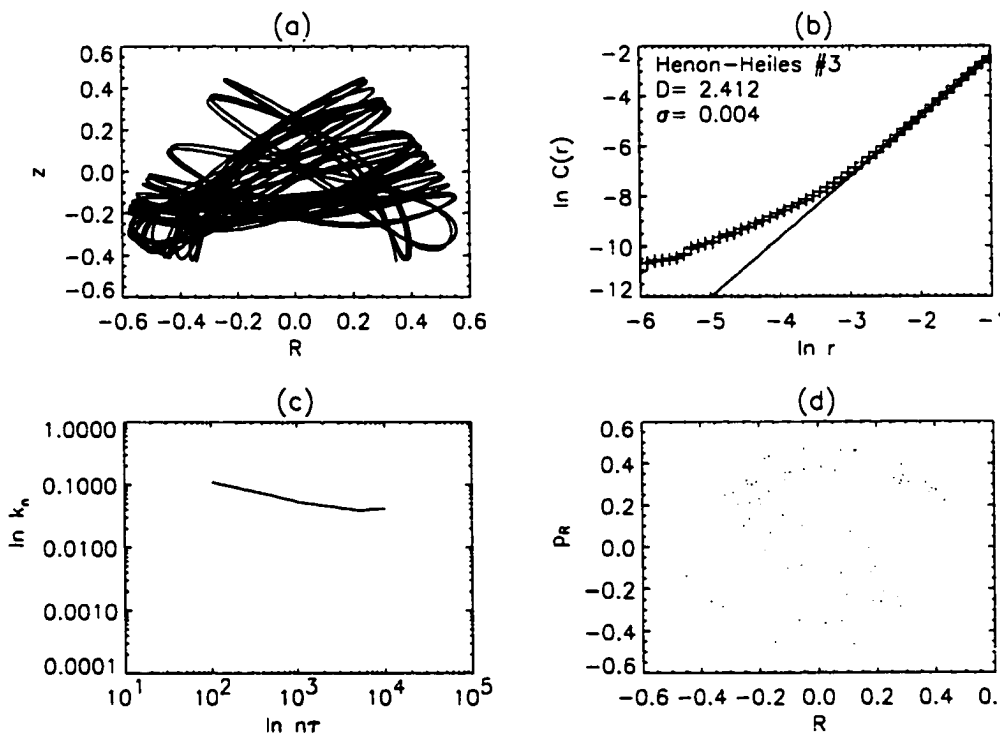

Figure 5.8: Orbit #3 in the Hénon-Heiles potential; each frame contains information as described in the caption to Fig. 5.6.

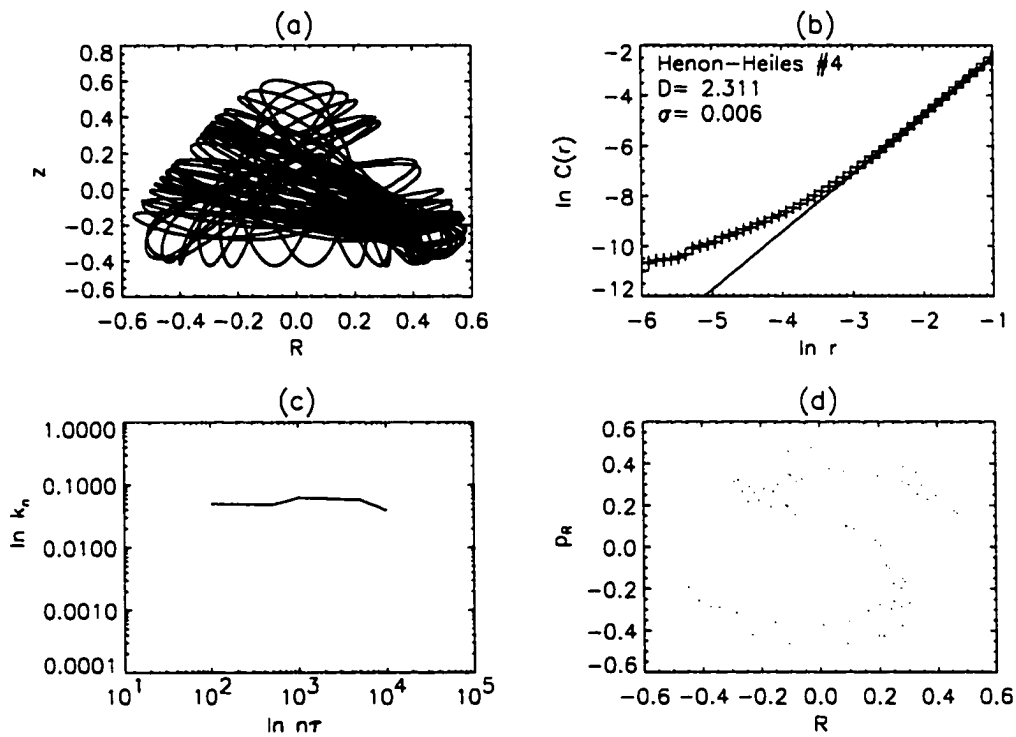

Figure 5.9: Orbit #4 in the Hénon-Heiles potential; each frame contains information as described in the caption to Fig. 5.6.

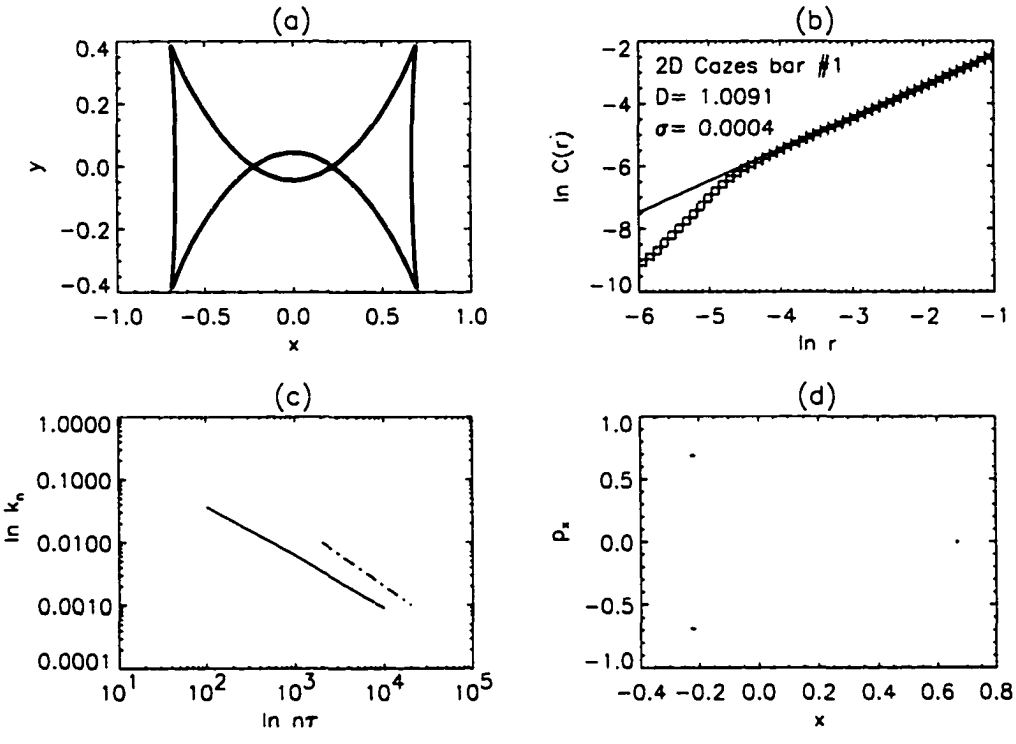

Figure 5.10: Orbit #1 in the 2D Cazes bar potential; see the Appendix for initial conditions. (a) The $x - y$ trajectory of this orbit. (b) The associated $C - r$ plot listing the average slope (D) and standard deviation (σ). (c) The plot of the Lyapunov exponent ($\ln k_n$) versus integration time ($\ln n\tau$) for this orbit. (d) The $(x - p_x)$ surface of section for this orbit.

surement of the Lyapunov exponent alone. The deviations from D at small values of r in Figs. 5.8b and 5.9b are simply artifacts of the limited number of orbital timesteps. This is similar to the small number statistics problem that was discussed above. When the number of timesteps is increased, these deviations disappear and the line with slope D extends to smaller r values.

Four different orbits that are supported by the 2D analytical Cazes bar potential are presented here in Figs. 5.10 through 5.13. Once again, all three methods of characterizing these 2D orbits agree: the orbits shown in

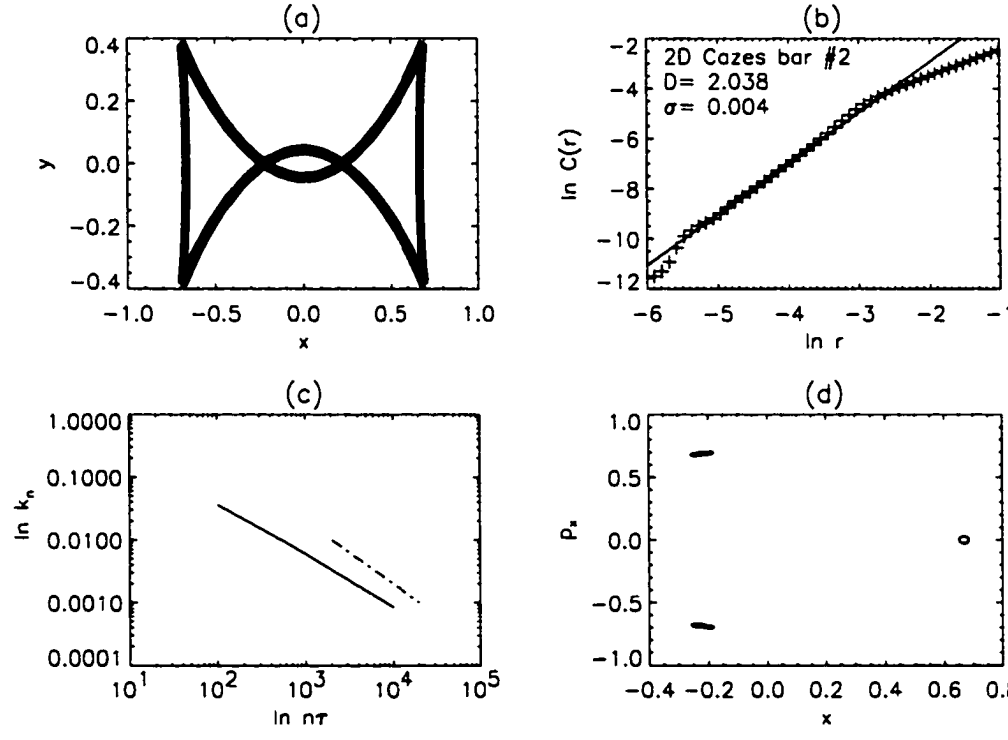

Figure 5.11: Orbit #2 in the 2D Cazes bar potential; each frame contains information as described in the caption to Fig. 5.10.

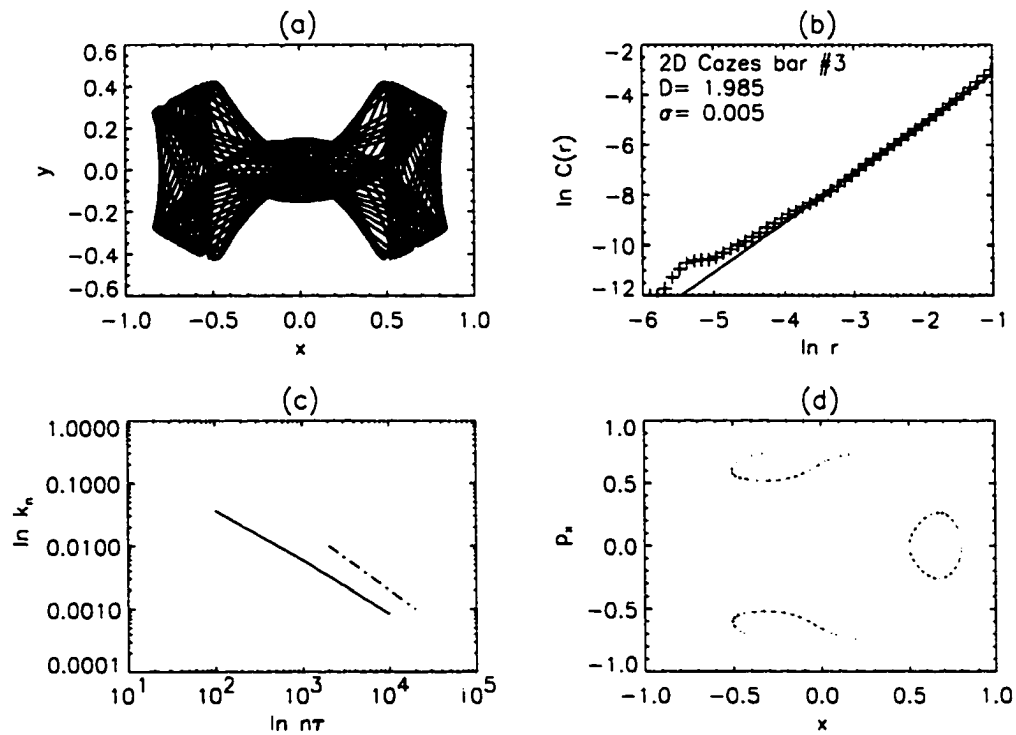

Figure 5.12: Orbit #3 in the 2D Cazes bar potential; each frame contains information as described in the caption to Fig. 5.10.

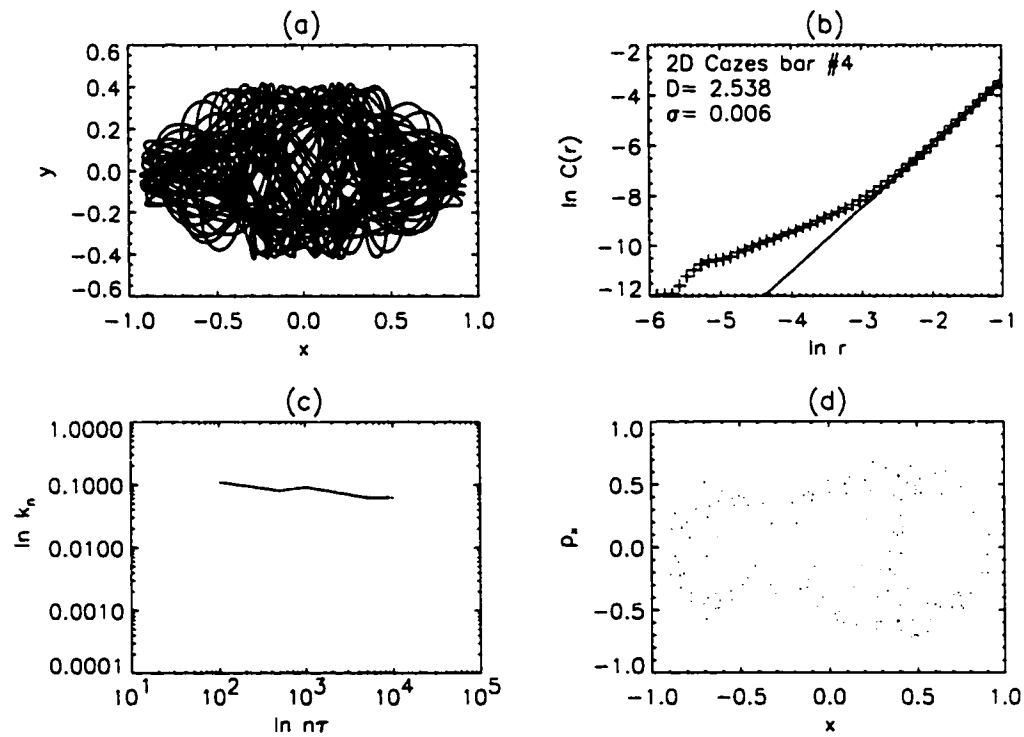

Figure 5.13: Orbit #4 in the 2D Cazes bar potential; each frame contains information as described in the caption to Fig. 5.10.

Figs. 5.10, 5.11, and 5.12 are regular, while the one shown in Fig. 5.13 is quasi-ergodic. The orbit shown in Fig. 5.10a has $D \approx 1$, rather than $D = 2$; hence, it is periodic. This orbit is the parent of the 'bow tie' family of orbits that were discussed extensively in Chapter 4. The orbit shown in Fig. 5.11a appears to be trapped near the periodic bow tie orbit. The most interesting aspect of this orbit is visible in the $C - r$ plot, Fig. 5.11b. There are two linear sections in the $C - r$ plot. The one that exists for small r has a slope of $D \approx 2$. At larger r, however, the linear section has a slope of $D \approx 1$. This slope discontinuity persists even when this orbit is followed through 100 times as many timesteps. This suggests that the discontinuity and the r location of the discontinuity are physically relevant (as opposed to the cases in the previous paragraph where differing slopes were numerical artifacts). We interpret this difference using the following analogy presented by Gleick (1987). Imagine viewing a ball of string from very far away. If asked to describe the dimensionality of the ball of string, which appears to be a point, the answer would be zero. Moving closer, the ball can be seen to be extended with dimensionality equal to 2. Moving even closer, the ball now appears to have a finite extent in a direction perpendicular to the dimensions already present as well, and is thus a 3D object. However, observing the ball of string at very close range, the one-dimensional nature of the string becomes apparent. Likewise, the dimensionality of a phase space orbit can change depending on what length scale is observed. Indeed, for small enough values of r, all phase space orbits derived from particle trajectories are really 1D objects. With this in mind, the segment with slope $D \approx 1$ is interpreted

as demonstrating that the orbit is not far from being closed. The r value at which the change in slope occurs identifies a critical length scale for this orbit that cannot be determined via the Lyapunov exponent method or from surfaces of section.

The quasi-ergodic orbit shown in Fig. 5.13 has a dimensionality $D \approx 2.54$. Unlike the discontinuity present in the regular orbit $C - r$ plot (Fig. 5.11b), the apparently linear slope at smaller r values is dependent on the number of timesteps taken for the orbit, as with the orbits shown in Figs. 5.8 and 5.9.

5.5 6D Phase Space

As with the 2D orbits previously discussed, the figures containing results for 3D orbits all have a similar form. Each figure will consist of: a frame labeled (a) illustrating the $x - y$ projection of the orbit; a frame labeled (b) illustrating the $x - z$ orbital projection; a frame labeled (c) illustrating the $y - z$ projection of the orbit; a frame (d) showing the $C - r$ plot for the orbit; and finally, a frame labeled (e) that displays the Lyapunov exponent plot derived from the orbit.

Figures 5.14 through 5.18 display results for five different orbits in the 3D Richstone potential. As expected, these results support the characterization of these orbits as regular. Notice that the slopes in Figs. 5.17d and 5.18d have noninteger values. This does not mean that the orbits are quasi-ergodic. The behavior of the curves in Figs. 5.17e and 5.18e shows that these orbits are regular and thus respect 3 full isolating integrals. However, they also appear to have ≈ 0.1 additional integrals. Another interpretation of this noninteger

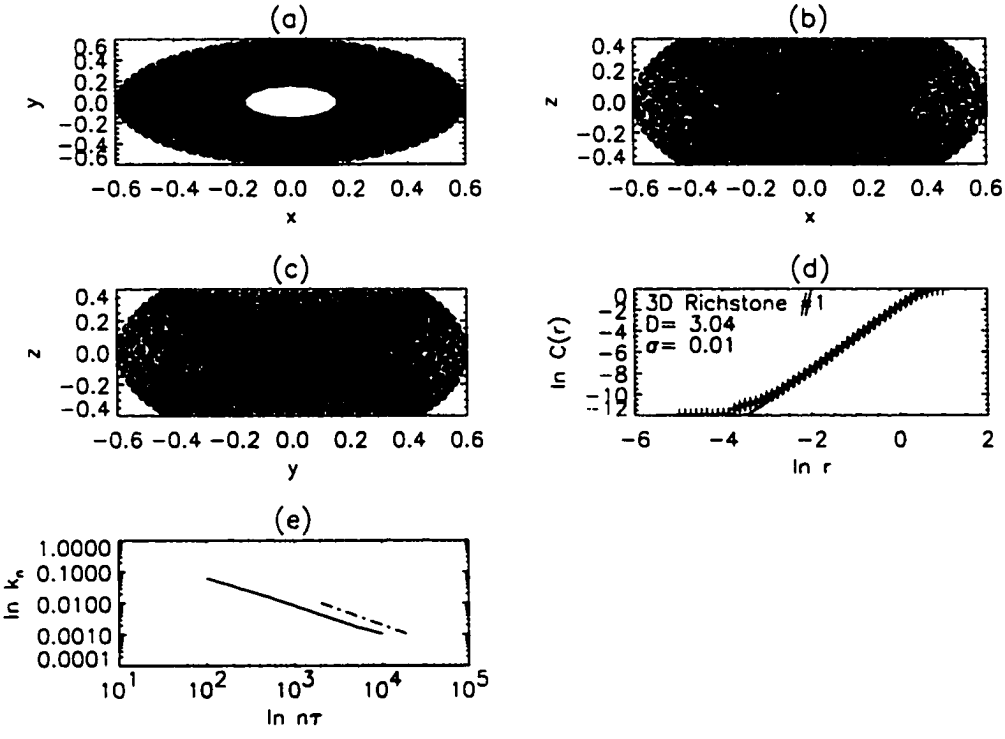

Figure 5.14: Orbit #1 in the 3D Richstone potential; see the Appendix for initial conditions. (a) The $x-y$ plane projection of this orbit. (b) The $x-z$ plane projection of the same orbit. (c) The $y-z$ plane projection of the same orbit. (d) The associated $C-r$ plot listing the average slope (D) and standard deviation (σ). (e) The plot of the Lyapunov exponent ($\ln k_n$) versus integration time ($\ln n\tau$) for this orbit.

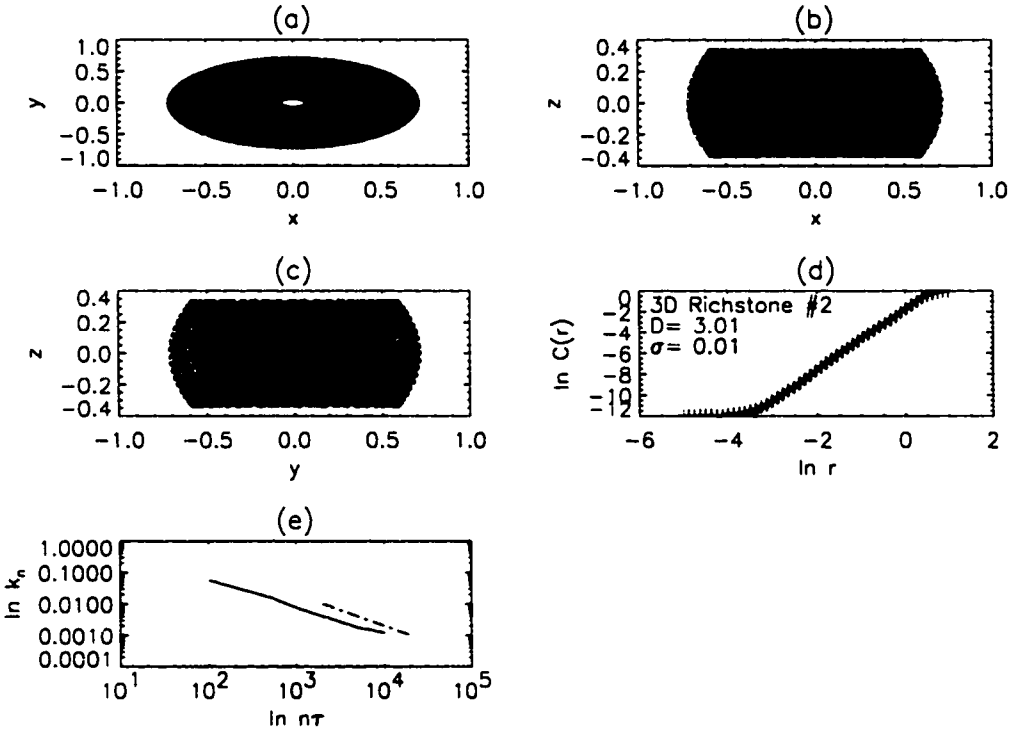

Figure 5.15: Orbit #2 in the 3D Richstone potential; each frame contains information as described in the caption to Fig. 5.14.

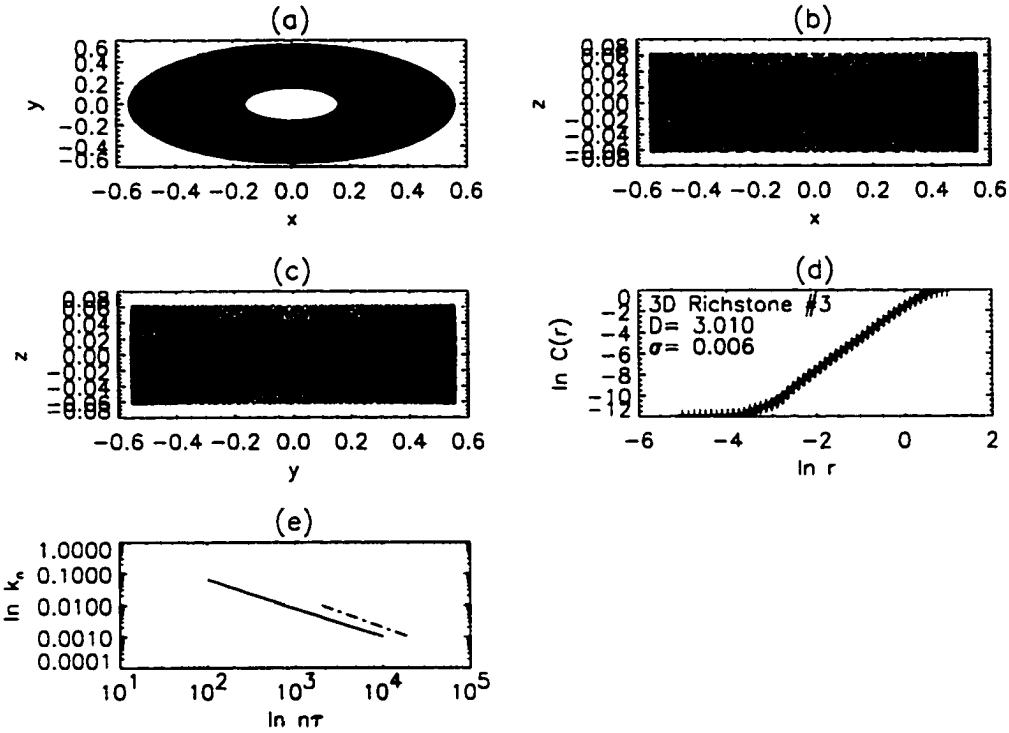

Figure 5.16: Orbit #3 in the 3D Richstone potential; each frame contains information as described in the caption to Fig. 5.14.

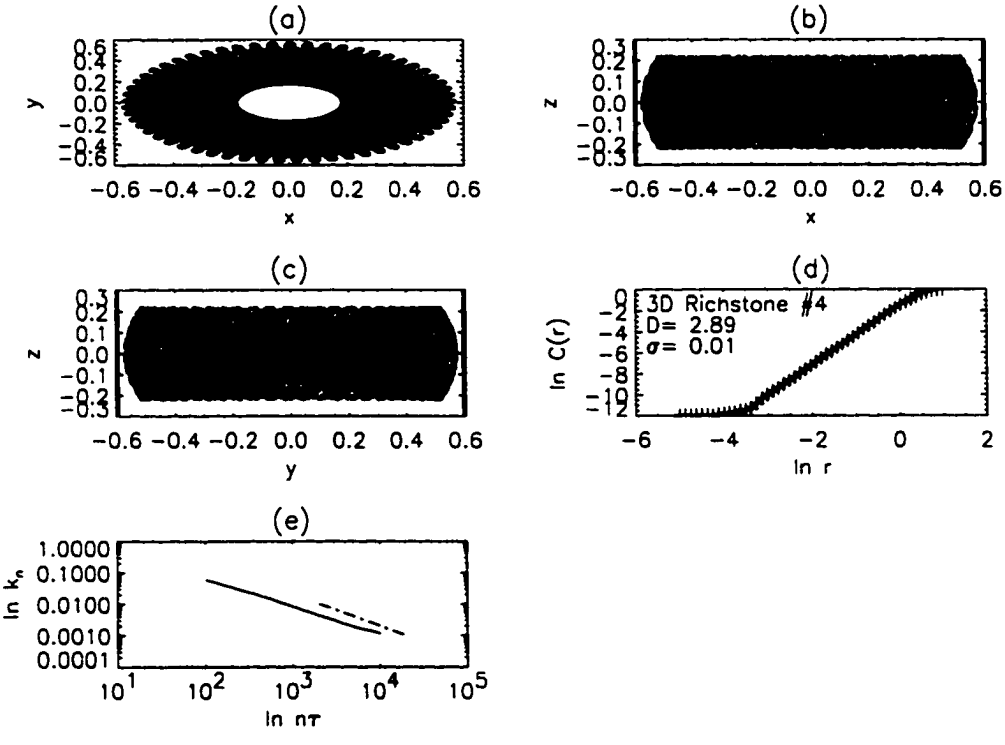

Figure 5.17: Orbit #4 in the 3D Richstone potential; each frame contains information as described in the caption to Fig. 5.14.

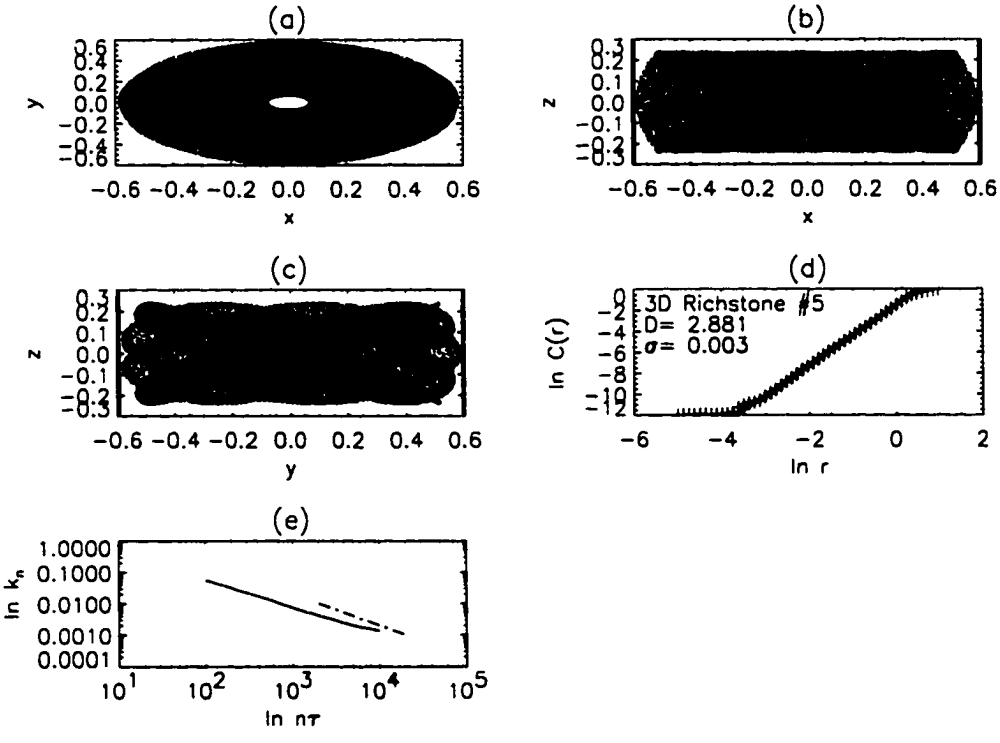

Figure 5.18: Orbit #5 in the 3D Richstone potential; each frame contains information as described in the caption to Fig. 5.14.

value is that there is another integral that is isolating only during 10% of the orbital period (Carnevali & Santangelo 1984). It is also possible that these results are due to a problem with our implementation of the correlation integral method. One possible explanation of these departures from the expected value of 3 is due to a problem with the sampling of phase space. It is possible that the orbits were not integrated long enough for the orbit to completely cover a full three dimensions. However, runs that extended five and ten times as long did not result in measured dimensionalities vastly different from the original orbit. It is also possible that not enough sampling points were used. However, using up to $N = 100,000$ made little difference to the result.

In the context of this discussion, the following question needs to be addressed: What is the size of any systematic error in the correlation integral method? There is no simple analytical or theoretical answer to this question. However, it can be said empirically that in the 2D cases discussed in §5.4 above, the D values of regular orbits differed from the nearest integer values by $< 3\%$. The quasi-ergodic orbits differed from the nearest integer values by more than 13%. In the cases of Figs. 5.17 and 5.18, the difference between D and the nearest integer are $\approx 4\%$ for both, whereas the differences for the rest of the 3D Richstone orbits are $< 1\%$. It could be argued that the 4% difference is really within the systematic error for an orbit in a 3D potential. While this may be the case, it is believed that the difference is physically relevant. The reason for this belief can be seen when Fig. 5.17 is compared to Fig. 5.14. Note that the orbital projections shown in Fig. 5.14 seem to

fill complete areas, for example, the points in Fig. 5.14a appear to fill an annulus. However, if an annulus were drawn on top of Fig. 5.17a, the points would not completely fill it. Instead, there appear to be "spokes" near the edge of the annulus with empty spaces between them. It is possible that this unfilled area is a signature of a regular, but noninteger dimensional, phase space orbit. For the purpose of characterizing orbits, this apparently extra fractional integral of motion does not present a problem. Orbits like this do respect 3 full isolating integrals ($D \leq 3$) and are regular.

5.6 The 3D Analytical Cazes Bar

In the previous sections, phase space orbits in well understood systems were analyzed using the correlation integral method. In this section, the focus is on orbits supported by the 3D analytical Cazes bar potential. This unique potential supports regular orbits, quasi-ergodic orbits, and orbits in-between those two classes, providing the opportunity to use the correlation integral method to analyze orbits of a type not previously studied. Unlike the transition from the 2D to 3D Richstone potentials, it is difficult to predict what behaviors will be seen in $C-r$ and Lyapunov exponent plots for orbits in the 3D analytical Cazes bar potential simply by knowing the results from the 2D case. The 3D Richstone potential has an obvious symmetry, and so it was quite certain that the 3D orbits would respect three integrals of motion; the two observed in the 2D case, ϵ and the unknown integral, as well as L_z. The 3D analytical Cazes bar has no such symmetry. There is no *a priori* reason to believe that any orbits in the 3D analytical Cazes bar potential should be regular. Some orbits could respect two integrals, ϵ_J and

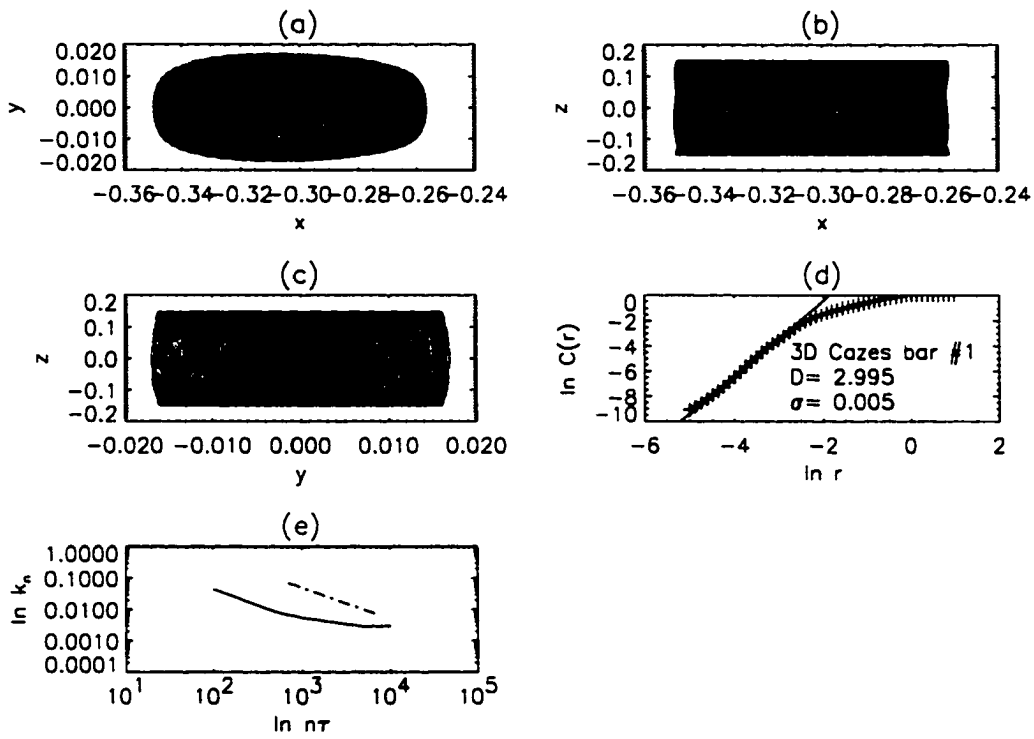

Figure 5.19: Orbit #1 in the 3D analytical Cazes bar potential; see the Appendix for initial conditions. (a) The $x-y$ plane projection of this orbit. (b) The $x-z$ plane projection of the same orbit. (c) The $y-z$ plane projection of the same orbit. (d) The associated $C-r$ plot listing the average slope (D) and standard deviation (σ). (e) The plot of the Lyapunov exponent ($\ln k_n$) versus integration time ($\ln n\tau$) for this orbit.

the unknown integral present in the 2D case, but there is no obvious third integral. Thus, the expectation is that some $C-r$ plots for 3D analytical Cazes bar orbits should have slopes $D \approx 4$ (respecting the integrals of motion seen in the 2D case) and the rest, quasi-ergodic orbits, should have slopes $4 < D < 5$.

Surprisingly, Figures 5.19 through 5.22 show four regular orbits that are supported by the 3D analytical Cazes bar. From the various orbital projec-

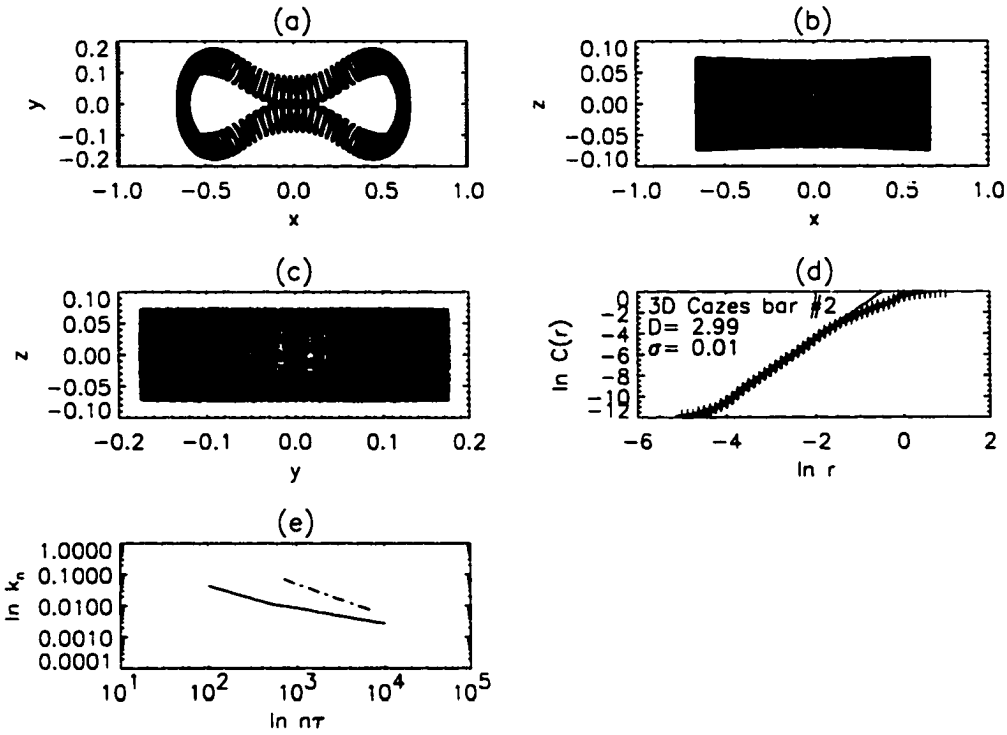

Figure 5.20: Orbit #2 in the 3D analytical Cazes bar potential; each frame contains information as described in the caption to Fig. 5.19.

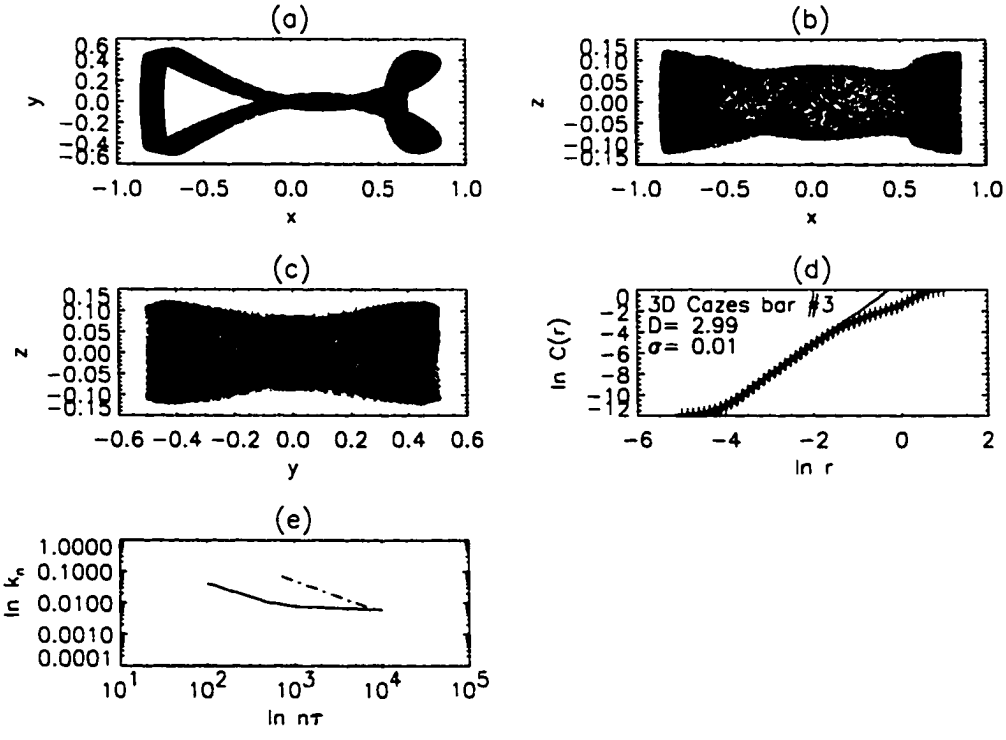

Figure 5.21: Orbit #3 in the 3D analytical Cazes bar potential; each frame contains information as described in the caption to Fig. 5.19.

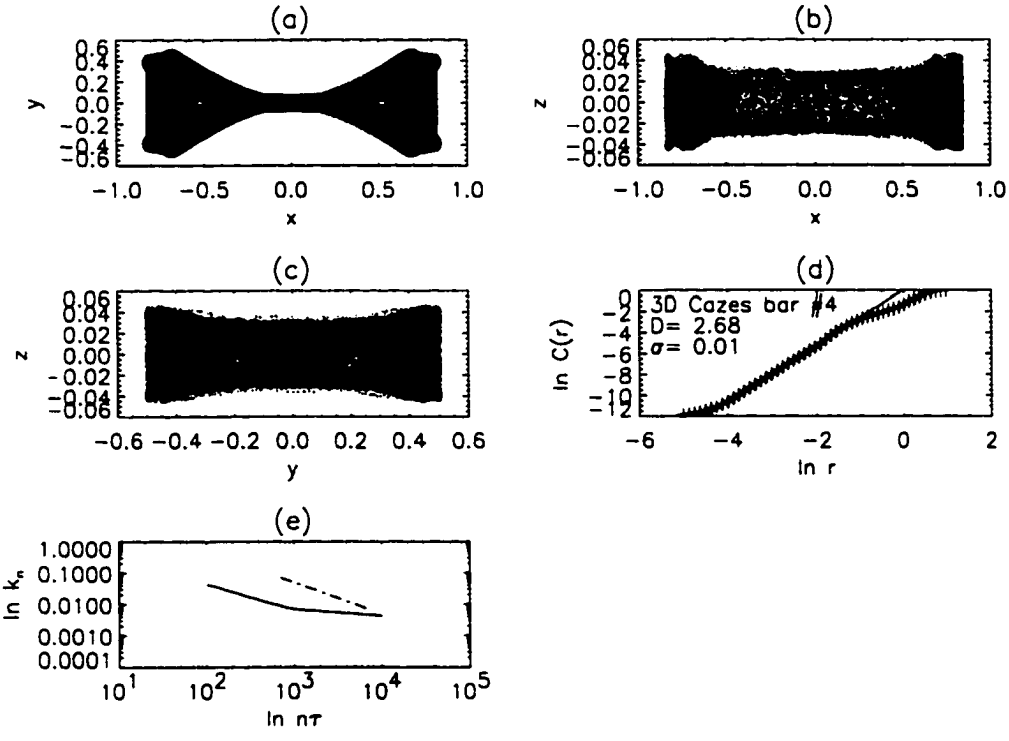

Figure 5.22: Orbit #4 in the 3D analytical Cazes bar potential; each frame contains information as described in the caption to Fig. 5.19.

tions, it seems that the regularity is most evident in the $x - y$ projections. Notice that the orbit shown in Fig. 5.19 is actually trapped in a slight potential minimum located "off-axis" along the major (x) axis. The three orbits detailed in Figs. 5.20 through 5.22 all seem to be 3D extensions of the 2D bow tie orbits. Note that the frames labeled (b) in Figs. 5.21 and 5.22 show that particles on these orbits will spend most of their time at the ends of the bow tie. It is also interesting to note that the measured slope is much shallower for the orbit in Fig. 5.22 than that in Fig. 5.21. Again, this fractional dimensionality does not mean that the orbit is quasi-ergodic. It appears to agree with the results found in the 3D Richstone case; regular orbits can have some additional (fractional) constraint on their phase space occupation. Despite this, the correlation integral characterization of these orbits as regular is supported by the Lyapunov exponent plots shown in each figure.

None of the three orbits shown in Figs. 5.23 through 5.25 is regular. However, whereas the orbits in Figs. 5.24 and 5.25 display quasi-ergodic behavior in both the correlation integral method and the Lyapunov exponent plot, the orbit in Fig. 5.23 is not quasi-ergodic. According to the correlation integral method, this orbit respects 2 integrals of motion and therefore, cannot be quasi-ergodic. However, the Lyapunov exponent plot makes no distinction between this orbit and those in Figs. 5.24 and 5.25. It is also difficult to differentiate this orbit from those in Figs. 5.24 and 5.25 by looking at the orbital projections. The orbit shown in Fig. 5.23 is the in-between class of orbit that was referred to above. It seems unlikely that systematic

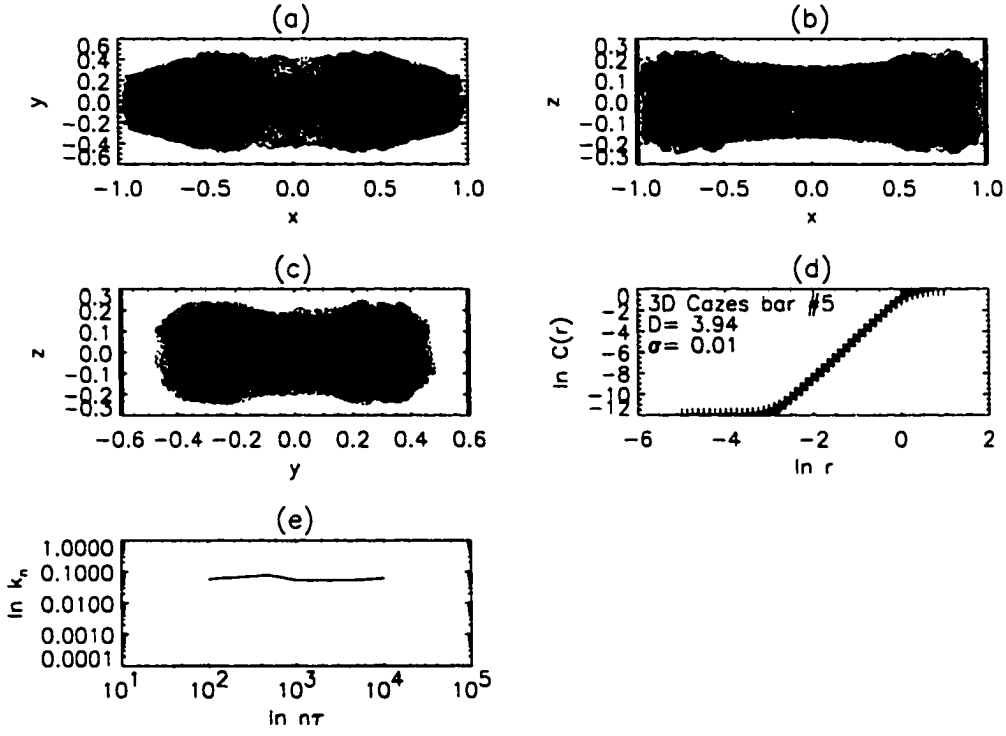

Figure 5.23: Orbit #5 in the 3D analytical Cazes bar potential; each frame contains information as described in the caption to Fig. 5.19.

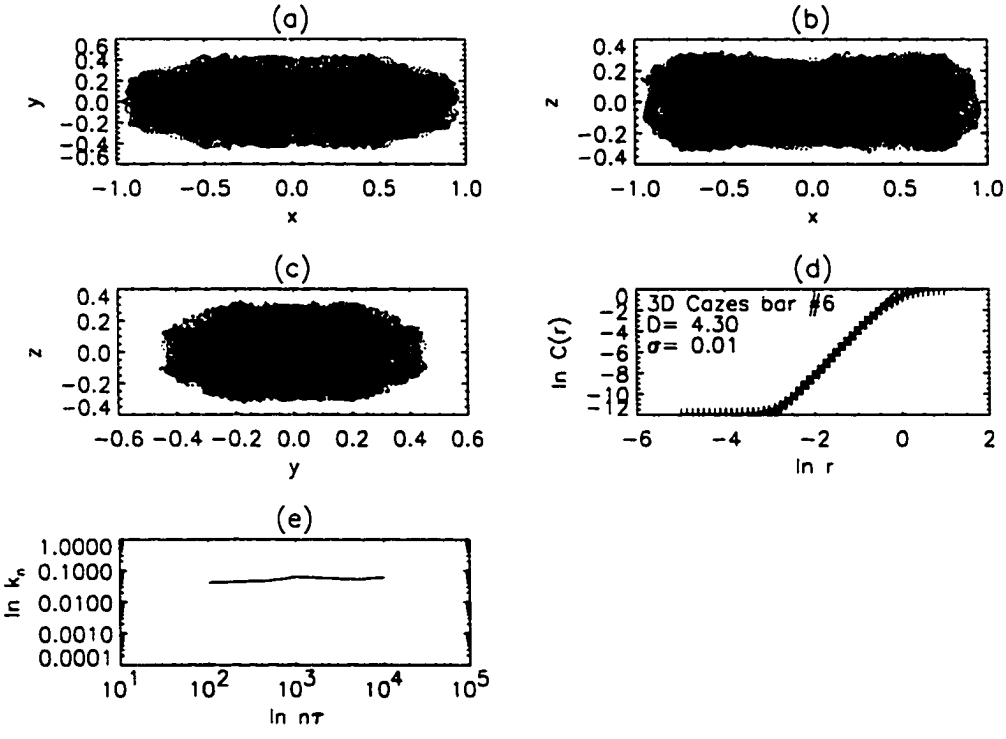

Figure 5.24: Orbit #6 in the 3D analytical Cazes bar potential; each frame contains information as described in the caption to Fig. 5.19.

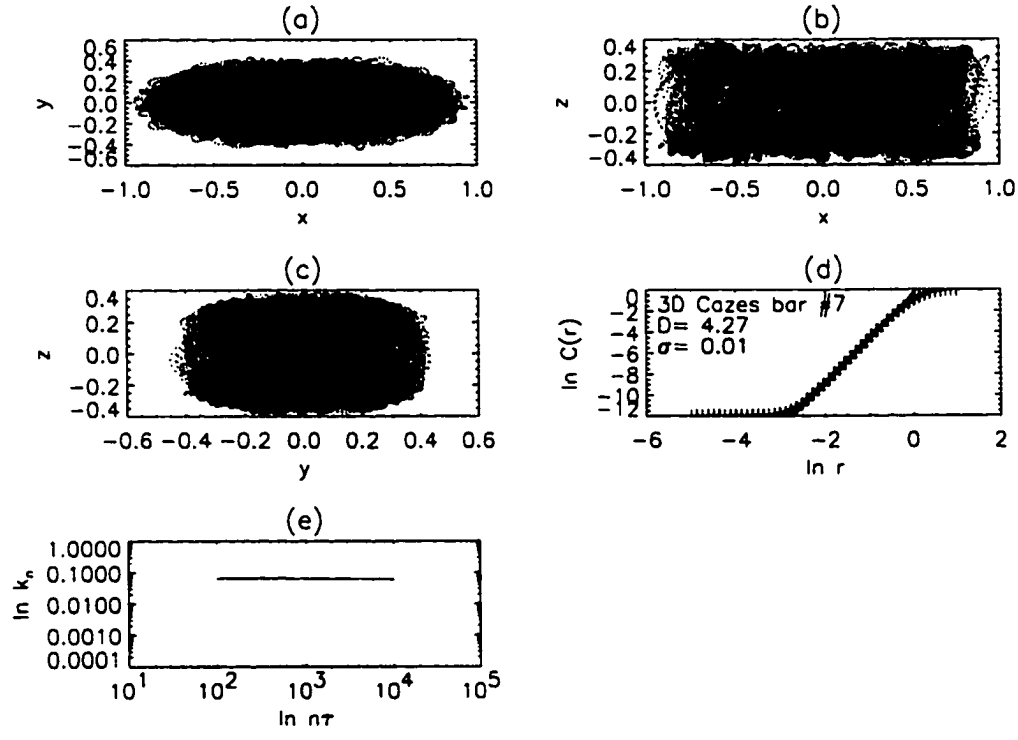

Figure 5.25: Orbit #7 in the 3D analytical Cazes bar potential; each frame contains information as described in the caption to Fig. 5.19.

error is to blame for the correlation integral result. The two undisputed quasi-ergodic orbits have slopes $\approx$ 7% greater than the nearest integer. So, even with the 3% systematic error, a similar orbit could, at best, display a slope 4% greater than the nearest integer slope. In fact, the measured slope in Fig. 5.23d is only 1.5% smaller than the nearest integer slope.

The reader may have noticed that some of the Lyapunov exponent plots shown in the preceeding figures display a tendency to "flatten out" at large $n\tau$ values, as in Figs. 5.4 and 5.19. This behavior is due to numerical inaccuracy and can be remedied by reducing the timestep used in the integration scheme. However, even with the current timestep the Lyapunov exponent calculations take a considerable amount of time. Since there is a clear difference between the Lyapunov exponent behaviors of regular and quasi-ergodic orbits despite the fact that the slopes are not precisely equal to -1, there has been no attempt to improve this situation.

6. ANALYSIS OF ANALYTICAL 3D CAZES BAR ORBITS

With the correlation integral method validated, it is time to use it as a tool to fully analyze an astrophysically interesting potential, the 3D analytical Cazes bar. This chapter contains a calculated study of one hundred orbits with specific values of ϵ_J. Twenty-five RH orbits are chosen at each of the following ϵ_J levels: -0.96, -0.85, -0.75, -0.63. These values have been chosen to match the values used in connection with the investigation of equatorial Cazes bar orbits (see Chapter 4). The initial positions of these orbits have been chosen to cover surfaces of constant ϵ_J. In practice, this coverage is determined by first identifying constant z planes that lie inside the constant ϵ_J surfaces. Then, in each of these planes, (x, y) positions are determined for grid points with the specific ϵ_J value under investigation. If more than twenty-five points are identified for any ϵ_J level, then a subsample of the points is used. The final sample of points for each ϵ_J value covers one quadrant of each surface. With the symmetries of the 3D analytical Cazes bar, this accurately represents fully covering the surface.

The results of the investigation of these orbits are summarized in the figures below. Figure 6.1 displays plots of the number of integrals of motion respected by each investigated orbit. The number of integrals of motion (I) has been obtained in each case by measuring the slope D in a $C - r$ plot and then subtracting D from the total dimensionality of phase space, in this case $I = 6 - D$. Each frame shows the measured results for all 25 orbits

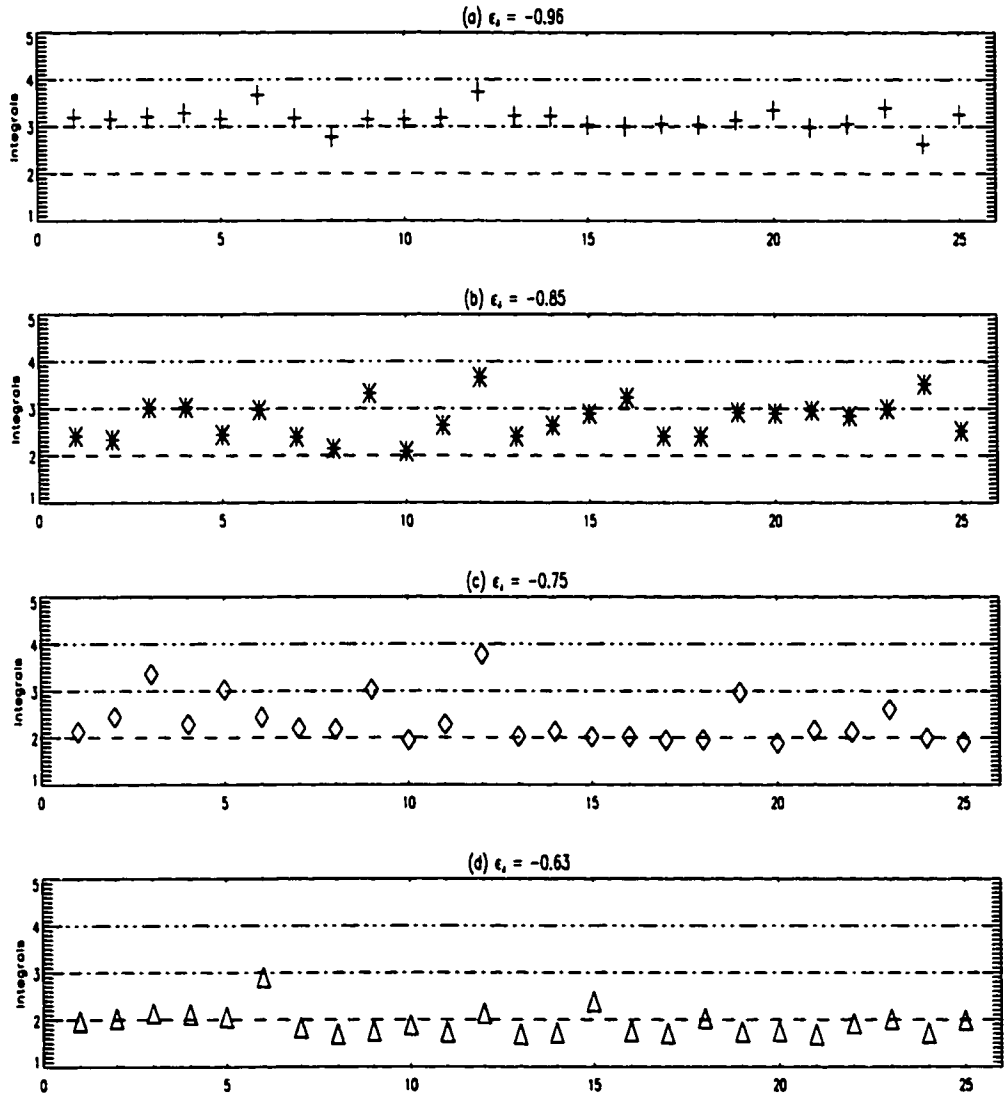

Figure 6.1: Plots showing the number of respected integrals for each of 100 different 3D analytical Cazes bar RH orbits. (a) The plus symbols mark orbits with $\epsilon_J = -0.96$. (b) The asterisks mark orbits with $\epsilon_J = -0.85$. (c) $\epsilon_J = -0.75$ orbits are represented by diamonds. (d) $\epsilon_J = -0.63$ orbits are marked with triangles.

with the same ϵ_J value. In each frame, the initial z position for the orbits increases to the right along the abscissa (however, the maximum value of z at the right edge is not the same for all frames). Figure 6.2 displays a histogram of the numbers of integrals of motion respected by each of the one hundred orbits studied here. The columns are centered on the values listed below them and extend ± 0.125 to either side. There seems to be evidence of a bimodal distribution around the integer values 2 and 3. Figure 6.3 is a two-dimensional histogram that shows the number of integrals of motion respected by the orbits as a function of their ϵ_J values.

A number of observations can be made from these figures. First, $39 \pm 6\%$ of the orbits investigated are regular ($I \geq 3$), $47 \pm 4\%$ are semiregular ($2 \leq I < 3$), and $15 \pm 4\%$ are quasi-ergodic ($I < 2$). (These quoted fractions have been determined in the same way as those discussed in §4.2.2.) These numbers would most likely change if, for example, a small random velocity component were added to the RH velocities (as discussed in §4.4.1) or if all of the orbits originated near the shocks present in the 3D Cazes bar (as discussed for the 2D case in §4.4.2). Second, there is a clear trend for orbits with more negative ϵ_J values to be regular. Third, there seems to be no correlation between the initial z position of the orbits and the corresponding value of I. The main inference to be drawn is that the effective energy of a star plays a large role in determining what type of orbit it will follow. It is interesting to note that this same general behavior has also been observed for orbits in the Hénon-Heiles potential (Hénon & Heiles 1964). In both cases, the relative fraction of nonregular orbits increases with increasing energy.

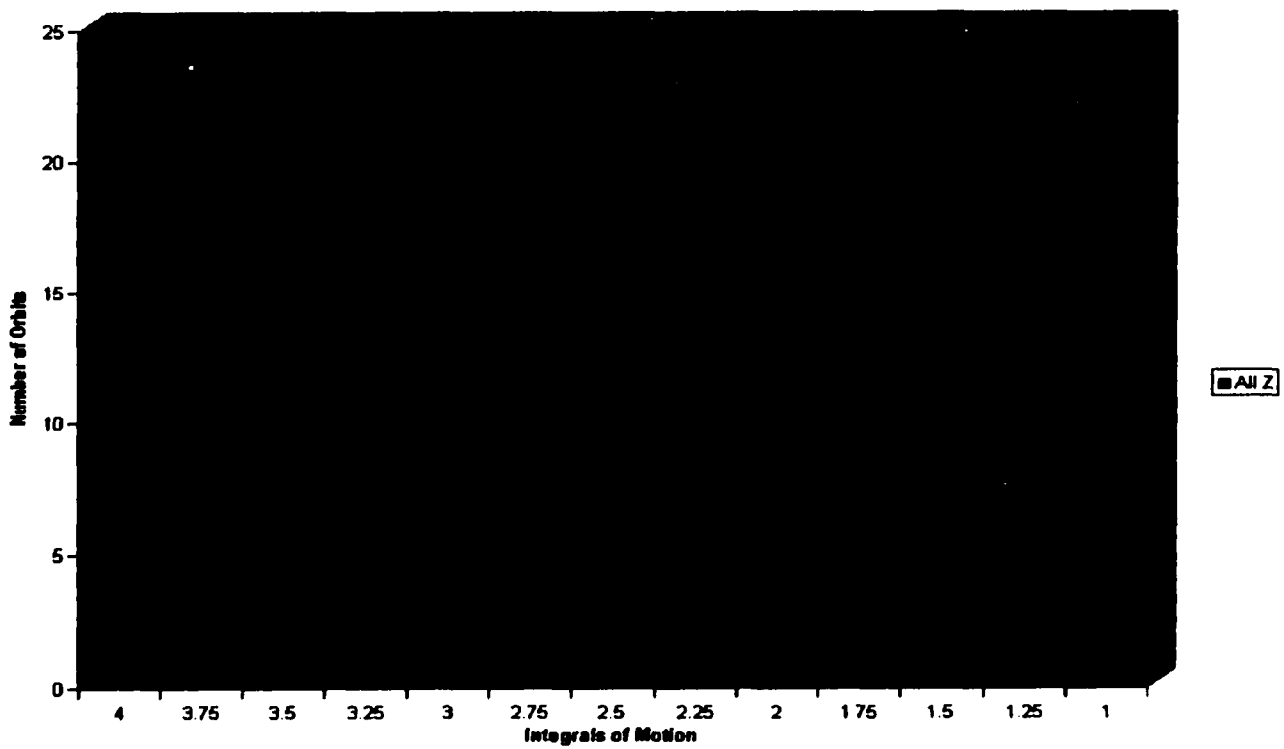

Figure 6.2: Histogram of the number of isolating integrals respected by the 3D analytical Cazes bar RH orbits. Note the clustering of values near 2 and 3. The columns are centered on the numbers below them and extend ± 0.125 to either side.

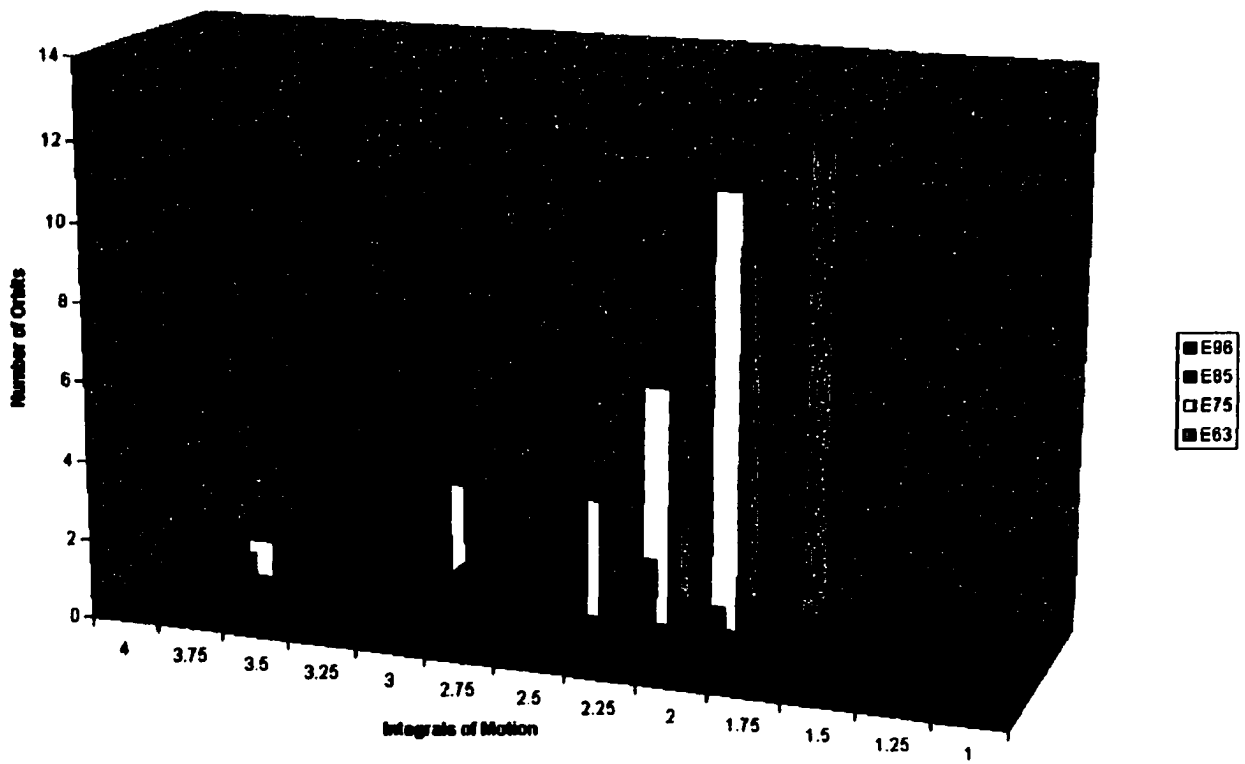

Figure 6.3: The 2D histogram of the number of isolating integrals respected by the 3D analytical Cazes bar RH orbits as a function of ϵ_J. The integral of motion columns are the same as in Fig. 6.2. The ϵ_J values increase from -0.96 in the front to -0.63 in the back.

Since these are simply test particle orbits, the behavior of a realistic stellar distribution function (i.e., one in which the density of stars is related to the potential) is unknown. However, these results are the first step towards understanding the types of orbits that can be populated during the transition from gas to stars in a nonaxisymmetric galaxy. The next step will involve combining an N-body simulation with a hydrodynamics simulation in order to follow the proper evolution of the entire galaxy system. This is a large task that is only beginning to be tackled. Despite the preliminary nature of the results presented here, it seems fair to say that this study adds to the previously discussed evidence that nonregular orbits can play important roles in galaxies (sec §5.1).

7. CONCLUSIONS

7.1 Equatorial-Plane Orbits in Barred Galaxies

In an effort to more fully understand the formation and evolution of barred galaxies, the properties of stellar orbits in a rotating, gaseous bar-shaped system have been studied. The system that has been examined in detail is one of the steady-state models developed by Cazes (1999, see also Cazes & Tohline 2000) in a recent three-dimensional hydrodynamical simulation. Initially, a shooting technique along with surface of section diagrams have been used to probe the orbits that are supported in the equatorial plane of this "Cazes bar". This analysis shows that the Cazes bar potential, $\Phi_{\rm CB}(x,y)$, supports a roughly equal mixture of regular and quasi-ergodic orbits. There also are indications that the presence of a shock front increases the ratio of quasi-ergodic to regular orbits. Virtually all of the regular prograde orbits appear to belong to a single family that have a bow tie shape. These orbits are almost certainly related to the 4/1 family of orbits described by Contopoulos (1988) because particles on bow tie orbits make four radial oscillations for each complete azimuthal cycle. But they differ from the 4/1 orbit illustrated in, for example, Fig. 1a of Contopoulos (1988) in that they pass very close to, and around, the center of the potential well twice each orbit cycle (see, for example, Fig. 4.3f). The equatorial Cazes bar potential also supports a variety of regular retrograde orbits, including some that appear to be members of the x_4 orbit family.

The analysis indicates that, over a large range of energies, the Cazes bar does not support the family of x_1 orbits (see the characteristic diagram, Fig. 4.8). This is perhaps the most striking difference between $\Phi_{CB}(x,y)$ and the potential wells that have been generated through self-consistent N-body simulations. N-body simulations tend to produce bars with stellar distribution functions, such as that in Sparke & Sellwood (1987), that are dominated by x_1 orbits. One possible reason for this is that the Cazes bar has a higher ratio of rotational to gravitational potential energy $T_{rot}/|W|$ than typical N-body bars and as a result, along its major axis, the Cazes bar potential is very shallow. In order to approximate this behavior, an analytical effective potential was designed that, while exhibiting a traditional quadratic dependence — *i.e.*, changing as $(y/R_{L2})^2$ — along the intermediate axis, changes as $(x/R_{L2})^4$ from the center along the major axis.

There are several interesting points to be made about the bow tie orbit family and about stars that might be injected into bow tie orbits. Although bow tie orbits should certainly be classified as a prograde orbit family, stars that move along bow tie orbits will *appear* to be moving in a retrograde sense on the portions of their orbits that are nearest the center of the bar. It is interesting to note that such counterrotation has been seen in at least one barred galaxy (Prada & Gutierrez 1999). Also, any star that moves along a bow tie orbit will (a) spend most of its time near the "four corners" of the orbit and (b) pass very close to the center of the potential well twice each orbit. When coupled with the discovery that a significant fraction of stars that form from gas in the Cazes bar will be injected into bow tie orbits, the

first of these points suggests that a gaseous bar should produce a distribution function, DF_{bar}, that is rather boxy or peanut-shaped. This is in contrast to distribution functions like DF_{SS} that are dominated by the x_1 family of orbits and are therefore more elliptical in shape.

The second of these points has led to the suggestion that star formation in a primarily gaseous bar may provide a mechanism for funneling matter in toward the center of a galaxy in situations where gas dissipation alone does not work efficiently. As noted by Norman & Silk (1983), triaxial potentials can provide a means of transporting stellar mass to a central black hole. Stars (or the dense cores of molecular clouds) that travel close to a central black hole can become tidally disrupted, and the resulting gas can form an accretion disk that fuels an active galactic nucleus (AGN) (c.f., Evans & Kochanek 1989; Ho, Filippenko, & Sargent 1997). Admittedly, the present model has not examined to what extent a central point mass will scatter and, thereby, disrupt the regular bow tie orbit (Gerhard & Binney 1985). However, the existence of orbits that travel near the center of the potential over such a large range of energies ($-0.96 < \epsilon_J < -0.63$) is intriguing. Rough calculations show that it may be possible for objects on bow tie orbits to transport enough mass to the central mass to power an AGN, but more detailed work must be done to solidify this idea.

7.2 Fully Three-Dimensional Orbits

Building on the work of Grassberger & Procaccia (1983) and Carnevali & Santangelo (1984), a flexible technique known as the correlation integral method also has been developed for characterizing orbits in 3D potentials.

This method analyzes phase space orbits and returns a single number, the dimensionality of the phase space orbit. From this number and the dimensionality of the underlying phase space, the number of isolating integrals of motion respected by an orbit can be determined. This number can then be used as a quantitative characterization attribute.

The implementation of the correlation integral method has been tested for a variety of previously studied systems. In all cases, more familiar characterization tools, such as surfaces of section and Lyapunov exponents, support the results obtained with the correlation integral method. The advantages of the correlation integral are most apparent when used to characterize orbits in 3D potentials because more traditional characterization techniques become much less useful when applied to 6D phase spaces associated with realistic models. Despite the fact that the analytical Cazes bar has no obvious geometrical symmetries that give rise to analytical integrals of motion, the correlation integral method demonstrates that it does support regular orbits. Additionally, the correlation integral method distinguishes between orbits that respect two integrals of motion and those that respect only one integral.

The simple fact that the correlation integral method can reproduce the results of other characterization methods is not enough to warrant its adoption. Here the various categorizing methods used in Chapters 4 & 5 are compared and contrasted.

- When analyzing orbits in a 4D phase space, surface of section diagrams are the simplest and clearest way to characterize orbits. However, there

is no simple quantitative measure that describes quasi-ergodic orbits in surface of section diagrams. Also, these diagrams are not easily translated to 6D phase spaces. The correlation integral method addresses both of these problems.

- Lyapunov exponents provide quantitative measures of orbital regularity in arbitrary 2D and 3D potentials. Unfortunately, all regular orbits, closed and unclosed, share the same signature in Lyapunov exponent plots. For orbits in 3D potentials, the behavior of the Lyapunov exponent is also the same for all nonregular orbits. That is, no distinction is made between orbits that respect only one integral of motion and those that respect two. The correlation integral method distinguishes between these types of regular (periodic and unclosed) and nonregular (respecting one or two integrals of motion) orbits.

Applying the correlation integral method analysis to a large sample of Restriction Hypothesis orbits in the 3D analytical Cazes bar potential has revealed that semiregular orbits dominate (by number) the types of orbits that are supported by the Cazes bar. Also, the value of ϵ_J appears to be the major factor in determining whether or not a given orbit will be regular. Specifically, orbits near the bottom of the potential well (lower energy) are much more likely to be regular than orbits with higher energy. The initial z position of an orbit apparently has no significant influence on the orbit's regularity.

The correlation integral method should prove useful as a tool for characterizing the properties of orbits in a wide variety of Hamiltonian dynamical

systems. However, there are three specific cases of astrophysical interest to which the correlation integral method seems particularly well suited: analyzing stellar distribution functions in analytically and numerically specified models of steady-state galactic potentials (especially those with Hamiltonian chaos); investigating the orbits that are supported by galactic potentials formed in cosmological simulations; and quantifying the response of stellar systems to potentials that contain central point masses.

7.3 Gasdynamical Versus Stellar-Dynamical Bars

The possibility that galaxies form central bar-like structures while still in a predominantly gaseous state is the idea underlying this study. Because it has been constructed in a self-consistent manner, the Cazes bar presents a reasonable representation of such a newly formed, gaseous galaxy configuration. If stars form from the gas in such a barred galaxy, the proposed Restriction Hypothesis illustrates the orbits into which the stars would be injected at the time of their formation. The analysis presented herein indicates that the distribution function DF_{bar} of such a system of stars would contain no retrograde orbits, but it would consist of a reasonable mixture of quasi-ergodic orbits and regular prograde orbits predominately related to the bow tie (4/1) orbit family. It is important to emphasize that these stellar orbits are distinctly different from the orbits that gas particles follow in the Cazes bar. Elements of gas are accelerated by local pressure gradients as well as by gradients in the underlying gravitational potential; also, unlike stellar orbits, gas particle orbits do not cross one another. As illustrated by Cazes (1999), within the steady-state Cazes bar the gas moves along closed stream-

lines that are approximately elliptical in shape. It is safe to say that no stars that form from such a gas flow will have similarly elliptical orbits. Searching many different initial conditions for particles in Φ_{CB}, no orbits were found that even approximated the gas streamlines.

Finally then, the following question must be raised: If a purely gaseous galaxy were to initially evolve into the form of a steady-state Cazes bar, then slowly create stars from the gas, injecting them according to the Restriction Hypothesis into the orbits that make up DF_{bar}, could a smooth evolutionary transition be made between the purely gaseous bar and one that is entirely made up of stars but that otherwise exactly resembles the Cazes bar? Using a technique similar to Schwarzschild's method (Schwarzschild 1979,1982) or that of Contopoulos & Grosbøl (1988), it is conceivable that the right combination of regular, semiregular, and quasi-ergodic orbits could be assembled to produce a steady-state stellar-dynamical bar[1]. And this configuration may even closely resemble the Cazes bar. (Given that an analytical function Φ_{eff} has been found that closely approximates Φ_{CB}, it should be relatively straightforward to conduct such a study.) However it seems unlikely that a system of stars that forms according to the Restriction Hypothesis from the Cazes bar could lead to such a configuration because the specific distribution of gas in the Cazes bar is unlikely to produce the required proportion of bow tie and quasi-ergodic orbits. For example, if in order to create a steady-state

[1]These methods have been designed to produce self-consistent potential-density pairs for N-body systems. First, the types of orbits supported by a given potential are investigated. Then, these orbits are populated to produce a density distribution. This density distribution is exactly that needed to provide the original potential. These methods are designed to provide model stellar distribution functions for any given potential.

stellar bar one needs N_{ϵ_J} bow tie orbits with energy ϵ_J, then there must be the right proportion of gas with energy ϵ_J at the proper positions within the Cazes bar to form stars for these orbits. With this additional constraint, it seems unlikely that there would be a clean transformation between a gaseous and a stellar system. It is suspected, instead, that after more than half of the gas has been converted into stars, the entire configuration would dynamically relax to a new configuration that is dominated by the collective dynamics of the stars. Since such an evolution would begin from a relatively high $T_{\rm rot}/|W|$ configuration that contains a large number of stars in bow tie orbits, it would be interesting to know whether this final state has a more boxy or peanut shape than the stellar dynamical configurations that have been created via N-body simulations from initially axisymmetric distribution functions. It may be necessary to answer this question before being able to state with any certainty whether barred galaxies form from initially axisymmetric ($\mathrm{DF_{axisym}}$) or nonaxisymmetric ($\mathrm{DF_{bar}}$) stellar distributions.

REFERENCES

Abraham, R., & Merrifield, M. 2000, AJ, 120, 2835

Barnes, E., & Tohline, J. 2001, ApJ, April 10, in press

Barnes, E. 2001, ApJ, submitted

Benettin, G., Galgani, L., & Strelcyn, J. 1976, Phys. Rev. A, 14, 2338

Berentzen, I., Heller, C., Shlosman, I., & Fricke, K. 1998, MNRAS, 300, 49

Binney, J. 1982a, MNRAS, 201, 1

Binney, J. 1982b, MNRAS, 201, 15

Binney, J., & Spergel, D. 1982, ApJ, 252, 308

Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton: Princeton University Press)

Binney, J., & Merrifield, M. 1998, Galactic Astronomy (Princeton: Princeton University Press)

Cazes, J. 1999, Ph.D. Thesis, Louisiana State University, Baton Rouge

Cazes, J., & Tohline, J. 2000, ApJ, 532, 1051

Carnevali, P., & Santangelo, P. 1984, ApJ, 281, 473

Carpintero, D., & Aguilar, L. 1998, MNRAS, 298, 1

Carpintero, D., Muzzio, J., & Wachlin, F. 1999, CeMDA, 73, 159

Chandrasekhar, S. 1969, Ellipsoidal Figures of Equilibrium (New Haven: Yale Univ. Press)

Colpi, M., Mayer, L., & Governato, F. 1999, ApJ, 525, 720

Contopoulos, G., & Papayannopoulos, Th. 1980, A&A, 92, 33

Contopoulos, G. 1988, A&A, 201, 44

Contopoulos, G., & Grosbøl, P. 1989, A&A Rev., 1, 261

Contopoulos, G., Gottesman, S., Hunter, J., & England, M. 1989, ApJ, 343, 608

Evans, C., & Kochanek, C. 1989, ApJ, 346, L13

Ferrarese, L., & Merritt, D. 2000, ApJ, 539, L9

Gebhardt, K., *et al.* 2000, ApJ, 539, L13

Gerhard, O., & Binney, J. 1985, MNRAS, 216, 467

Gleik, J. 1987, Chaos: Making a New Science (New York:Penguin)

Goldreich, P., & Lynden-Bell, D. 1965, MNRAS, 130, 125

Goodman, J., & Schwarzschild, M. 1981, ApJ, 245, 1087

Grassberger, P., & Procaccia, I. 1983, Phys. Rev. Lett., 50, 346

Habib, S., Kandrup, H., & Mahon, M. 1997, ApJ, 480, 155

Hartmann, W., & Impey, C. 1994, Astronomy: The Cosmic Journey (Pacific Grove:Brooks/Cole), http://astronomica.org/textbook

Hénon, M. 1976, Comm. Math. Phys., 50, 69

Hénon, M., & Heiles, C. 1964, AJ, 69, 73

Ho, L., Filippenko, A., & Sargent, W. 1997, ApJ, 487, 591

Hohl, F. 1971, ApJ, 168, 343

Hohl, F., & Zang, T. 1979, AJ, 84, 585

Hubble, E. 1925, ApJ, 62, 409

Hubble, E. 1936, The Realm of the Nebulae (New Haven: Yale University Press)

Julian, W., & Toomre, A. 1966, ApJ, 146, 810

Kalnajs, A. 1972, ApJ, 175, 63

Kalnajs, A. 1976, ApJ, 205, 751

King, I. 1966, AJ, 71, 64

Knapen, J., Shlosman, I., & Peletier, R. 2000, ApJ, 529, 93

Krumm, N., & Salpeter, E. 1979, AJ, 84, 1138

Lai, D., Rasio, F., & Shapiro, S. 1993, ApJS, 88, 205

Larson, R. 1976, MNRAS, 176, 31

Lebovitz, N. 1987, in Highlights of Astronomy Vol. 8, ed. D. McNally (Boston:Kluwer), 129

Lichtenberg, A., & Lieberman, M. 1983, Regular and Stochastic Motion (New York:Springer-Verlag)

Lin, C., & Shu, F. 1964, ApJ, 132, 286

Lin, C., & Shu, F. 1966, Proc. Nat. Acad. Sci., 55, 229

Longair, M. 1998, Galaxy Formation (New York:Springer-Verlag)

Lütticke, R., Dettmar, R.-J., & Pohlen, M. 2000a, A&AS, 145, 405

Lütticke, R., Dettmar, R.-J., & Pohlen, M. 2000b, A&A, 362, 435

Merritt, D., & Valluri, M. 1996, ApJ, 471, 82

New, K., Centrella, J., & Tohline, J. 2000, Phys. Rev. D, 62, 064019

Norman, C., & Silk, J. 1983, ApJ, 266, 502

Ostriker, J., & Peebles 1973, ApJ, 202, 353

Pearce, F., Jenkins, A., Frenk, C., Colberg, J., White, S., Thomas, P., Couchman, H., Peacock, J., & Efstathiou, G. 1999, ApJ, 521, L99

Penzias, A., & Wilson, R. 1965, ApJ, 142, 419

Pfenniger, D., & Friedli, D. 1991, A&A, 252, 75

Pilat-Lohinger, E., Dvorak, R., & Burger, Ch. 1999, CeMDA, 73, 117

Prada, F., & Gutierrez, C. 1999, ApJ, 517, 123

Richstone, D. 1982, ApJ, 252, 496

Rubin, V., & Ford, W. 1970, ApJ, 159, 379

Schwarzschild, M. 1979, ApJ, 232, 236

Schwarzschild, M. 1982, ApJ, 263, 599

Sellwood, J. 1981, A&A, 99, 362

Sellwood, J. 2000, in Dynamics of Galaxies: from the Early Universe to the Present, ed. F. Combes, G. Mamon, & V. Charmandaris, ASP Conference Series, Vol. 197

Seyfert, C. 1943, ApJ, 97, 28

Sparke, L., & Sellwood, J. 1987, MNRAS, 225, 653

Spinrad, H., Djorgovski, S., Marr, J., & Aguilar, L. 1985, PASP, 97, 932

Teuben, P., & Sanders, R. 1985, MNRAS, 212, 257

Toomre, A. 1981, in The Structure and Evolution of Normal Galaxies, ed. S.M. Fall & D Lynden-Bell (Cambridge: Cambridge Univ. Press)

Udry, S., & Pfenniger, D. 1988, A&A, 198, 135

Valluri, M., & Merritt, D. 1998, ApJ, 506, 686

Verlet, L. 1967, Phys. Rev., 159, 98

Williams, H., & Tohline, J. 1987, ApJ, 315, 594

Zwicky, F. 1933, Helvetica Physica Acta, 6, 110

APPENDIX A: ORBIT INITIAL CONDITIONS

Table A.1. Orbit Initial Conditions

Orbit	x_0	y_0	z_0	$\dot{x}_0$	$\dot{y}_0$	$\dot{z}_0$
Hénon map	0.0	0.0	...	...	...	...
2D Richstone #1	0.5	0.0	...	0.0	0.4	...
2D Richstone #2	0.5	0.0	...	0.2	0.2	...
2D Richstone #3	0.5	0.0	...	1.12	0.23	...
2D Richstone #4	0.136	-0.532	...	0.0	0.0	...
Hénon-Heiles #1	0.0	0.3	...	0.422	0.0	...
Hénon-Heiles #2	0.0	0.14	...	0.481	0.0	...
Hénon-Heiles #3	0.0	0.24	...	0.402	0.2	...
Hénon-Heiles #4	0.0	0.24	...	0.281	0.35	...
2D Cazes bar #1	0.67	0.0	...	0.0	0.611	...
2D Cazes bar #2	0.65	0.0	...	0.0	0.628	...
2D Cazes bar #3	0.5	0.0	...	0.0	0.712	...
2D Cazes bar #4	0.5	-0.5	...	0.0	0.507	...
3D Richstone #1	0.5	0.0	0.0	0.4	0.4	0.7
3D Richstone #2	0.6	0.0	0.2	0.5	0.2	-0.05
3D Richstone #3	0.5	0.0	0.3	0.4	0.5	0.01
3D Richstone #4	0.5	0.0	0.0	0.4	0.5	0.4
3D Richstone #5	0.3	0.3	-0.1	0.2	0.6	-0.5
3D Cazes bar #1	-0.282	-6.92(-2)	-0.147	2.34(-2)	-1.85(-2)	-4.95(-4)
3D Cazes bar #2	0.658	0.0	-7.35(-2)	-1.23(-2)	0.126	4.68(-4)
3D Cazes bar #3	-0.679	-0.504	-0.11	5.88(-2)	-1.67(-2)	3.01(-3)
3D Cazes bar #4	-0.679	-0.504	-3.68(-2)	5.7(-2)	-1.97(-2)	1.14(-3)
3D Cazes bar #5	-0.271	0.384	0.11	-0.363	1.97(-3)	-2.73(-3)
3D Cazes bar #6	0.408	0.389	-0.147	-0.254	3.23(-2)	-2.87(-4)
3D Cazes bar #7	0.222	-0.173	0.331	0.276	-1.6(-2)	3.3(-3)

APPENDIX B: LETTER OF CONSENT

From: keith.dodson@brookscole.com

To: Eric Barnes ¡barnes@baton.phys.lsu.edu¿

Date: Thu, 1 Mar 2001 18:14:21 -0800

Subject: Re: permission to reproduce figure

Dear Eric,

Please consider this e-mail permission to use the figure in your thesis. However, please indicate the source either in the legend or immediately beneath the figure. Thank you.

Keith Dodson Editor

From: Eric Barnes ¡barnes@baton.phys.lsu.edu¿ on 03/01/2001 12:38:20 PM

To: Keith Dodson/BCP/International Thomson Publishing@ITP

Subject: permission to reproduce figure

My name is Eric Barnes and I am trying to get permission to reproduce Figure 24-20 from the on-line textbook "Astronomy: The Cosmic Journey" in my Ph.D. dissertation. I contacted Dr. Impey (one of the authors) and he told me that only the publisher could give consent. This will not be published in any journal. Please let me know if there is anything that I can/need to do. Thank you for your time.

VITA

Eric Barnes was born on August 14th, 1973, in Joliet, Illinois. He graduated from Lebanon Community High School in 1991 and began his undergraduate education at Northeast Missouri State University later the same year. In the spring of 1995, Eric graduated from Northeast. After a restful year, he began graduate studies at Louisiana State University (LSU) in the fall of 1996. A master of science degree in physics was obtained in 1998 after successfully completing a qualifying exam. During the 1999/2000 academic year he was president of the Graduate Student Association at LSU. He expects to receive the degree of Doctor of Philosophy in physics in May 2001 and to take a postdoctoral research position at Rutgers University shortly thereafter.

DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Eric Barnes

Major Field: Physics

Title of Dissertation: A Study of Stellar Orbits in a Rotating, Gaseous Bar

Approved:

Major Professor and Chairman

Dean of the Graduate School

EXAMINING COMMITTEE:

Date of Examination:

March 20, 2001