Global dynamics in galactic triaxial systems. I

P. M. Cincotta,1,2 C. M. Giordano1,2, and M. J. Pérez1,2

1 Facultad de Ciencias Astronómicas y Geofísicas, Universidad Nacional de La Plata, Paseo del Bosque S/N, La Plata (1900), Argentina
e-mail: jperez@fcaglp.unlp.edu.ar
2 Instituto de Astrofísica de La Plata (IALP), Consejo Nacional de Investigaciones Científicas y Técnicas de la República Argentina

Received 19 December 2005 / Accepted 18 April 2006

ABSTRACT

We present a theoretical analysis of the global dynamics in a triaxial galactic system using a 3D integrable Hamiltonian as a simple representation. We include a thorough discussion on the effect of adding a generic non-integrable perturbation to the global dynamics of the system. We adopt the triaxial Stäckel Hamiltonian as the integrable model and compute its resonance structure in order to understand its global dynamics when a perturbation is introduced. We provide a theoretical discussion about diffusive processes taking place in phase space.

Key words. methods: numerical – galaxies: structure – chaos – diffuson – galaxies: kinematics and dynamics

1. Introduction

Observations with HST revealed the presence of very high stellar densities at the centers of early-type galaxies (Crane et al. 1993; Ferrarese et al. 1994), suggesting a power law ($r^{-p}$) fit for them. The evidence for large central masses was also reinforced from high-resolution kinematic studies of nuclear stars and gas, which revealed the presence of compact dark objects with masses in the range of $10^{5.3} - 10^{6.5} \ M_{\odot}$, presumably super-massive black holes (Ford et al. 1998). These observational results have produced a substantial change in the classic thinking on dynamics in triaxial galaxies.

Results obtained from numerical simulations show that the addition of a central mass to an integrable triaxial potential has strong effects on its dynamics, at least for the boxlike orbits that cover the central region of triaxial galaxies. Black holes and central density cusps scatter these particular orbits during each close passage inducing chaos in the system. The sensitivity of boxlike orbits to deflections also leads to a rounder central distribution of mass (see for instance Gerhard & Binney 1985, and Udry & Pfenniger 1988). This slow evolution towards axisymmetry suggests that stationary triaxial configurations cannot exist for a central density cusp. Valluri & Merritt (1998) find that in most of early-type galaxies, chaotic evolution is determined by the mass of the central black hole, ($M_{bh}$), rather than by the slope of the density profile. They show that when the central mass contains 2\% of the galaxy mass, a transition to global stochasticity occurs. For such large values of $M_{bh}$, the box-orbit phase space is almost completely stochastic and diffusive processes could take place over very short timescales.

This result is substantially attractive because this critical black hole mass was close to the observed one (Kormendy & Ritchie 1995) and also close to the mass that induced a sudden evolution toward axisymmetry in N-body simulations (Merritt & Quinlan 1998). However, this is no longer true, since from the works of, for instance, Ferrarese & Merritt (2000) and Merritt & Ferrarese (2001) it is known that the mass of black holes in galaxies, derived from the black hole demographic relationships, are 0.1–0.2% of the ellipsoid mass in which they are embedded.

Merritt & Fridman (1996) arrive at the same conclusion analyzing two triaxial power law models: the steep ($p = 2$) and the weak ($p = 1$) cusps. They find, in agreement with Gerhard & Binney (1985) and Schwarzschild (1993), that triaxial galaxies with such high concentration of mass would evolve toward a central axisymmetry, as box orbits gradually lose their distinguishability. For these models, in which a large fraction of phase space is dominated by chaotic dynamics, the construction of self-consistent solutions requires the inclusion of stochastic orbits as well as the regular ones. A system thus built evolves, mainly close to its center, as stochastic orbits mix, through phase space. In order to obtain stationary solutions, they build the “fully-mixed models” keeping stochastic orbits out of the central part of the system where chaotic orbits mix ergodically driving to a rounder distribution and destroying the triaxial self-consistency. They find that although it is possible to build this kind of solution for a weak cusp model, this is not the case for a strongly concentrated model. This would imply that triaxiality is not consistent with a high central density. The discussion of the existence of stationary non-axisymmetric stellar systems is still open. Poon & Merritt (2002) were able to construct triaxial equilibrium with central black holes which were both regular and stable, but under a very strong assumption. They need the chaotic building blocks to be fully mixed for the triaxial equilibrium to co-exist with central singularities, and there is no evidence yet that in 3D systems with divided phase space, a completely connected chaotic component actually exists. Moreover, it seems that this could happen only when the chaotic component has a large measure ($\approx 1$) and $t \to \infty$, which, from a physical point of view, would not be possible in galactic systems (see for instance Giordano & Cincotta 2004 and references therein and also Sect. 5).

All these conclusions rest on the strong hypothesis that diffusive processes drive stochastic orbits to mix, covering all chaotic
regions of phase space on a timescale of the order of the Hubble time. Taking into account the fundamental role that diffusive processes should play in galactic dynamics (as well as in asteroidal and planetary dynamics), we will discuss the relevance of diffusive processes in phase space theoretically as well as numerically. In this first paper we address the theoretical discussion; numerical studies will be presented in a forthcoming paper.

Global dynamics of triaxial galactic systems have been studied in previous works, for example by Wachlin & Ferraz-Mello (1998) and Papaphilippou & Laskar (1998). They represented elliptical galaxies using, respectively, a generalization of a double-power-law spherical mass model and the axisymmetric softened version of the 3D logarithmic potential. For these models, they applied frequency map analysis (Laskar 1993) in order to investigate the orbital structure of the system. Both analyses revealed stochastic motion in both models. Although they show the main resonances of these systems and investigate the relevance of chaotic motion in these models, they did not focus on the study of the diffusive mechanisms that may present in these systems. Moreover, the frequency map analysis is not an efficient tool for such of studies since it does not provide a good measure of chaos (see Cincotta & Simó 2000; Cincotta et al. 2003).

Here, we compute the theoretical resonant structure of an integrable triaxial system. Particularly, we focus our attention on the Stäckel triaxial model, which we assume to be a rough approximation of an elliptical galaxy. Then we add a generic non-integrable perturbation in order to show how the resonance structure is distorted by the effect of this perturbation.

In Sect. 2, we briefly recall the main characteristics of the integrable Stäckel model. Numerical details related to the computation of the resonances and the analysis of the resonance structure are presented in Sect. 3. In Sect. 4, we review and provide a theoretical discussion about the transition to chaos as a consequence of resonance interaction. Finally, we discuss the different processes that could lead to chaotic diffusion in phase space.

2. The integrable model

Motion in a smooth gravitational field becomes quite simple if the number of isolating integrals equals the number of degrees of freedom, and much work in galactic dynamics has focused on finding integrable models for galactic potentials. Kuzmin (1956, 1975) showed that there is a unique, ellipsoidal stratified mass model for which the corresponding potential has three global integrals of the motion, quadratic in the velocities. Kuzmin’s model, explored in detail by de Zeeuw (1985) who called it the “Perfect Ellipsoid”, has a large, constant-density core in which the orbital motion is similar to that of a three-dimensional linear oscillator. Every orbit in the core of the Perfect Ellipsoid fills a region close to a rectangular parallelepiped, or box. These trajectories were called box orbits. At higher energies in the Perfect Ellipsoid, box orbits persist and three new orbital families appear: the tubes (inner and outer long axis tubes and short axis tubes), orbits that preserve the direction of their circulation around either the long or short axis of the figure. Tube orbits respect an integral of motion analogous to the angular momentum, and hence – unlike box orbits – avoid the center. As pointed out by Schwarzschild (1981), these four families, box, inner and outer long axis tubes and short axis tubes, reproduce the general form of triaxial galaxies.

The Perfect Ellipsoid model has been used by several authors as a first approximation to represent an elliptical galaxy.

In this section, we address some relevant aspects of the integrable Stäckel model – separable in ellipsoidal coordinates – whose three global integrals of motion admit analytical expressions.

The Perfect Ellipsoid is represented by stratified density function distributed over concentric ellipsoids of semi-axes $ma$, $mb$ and $mc$, given by

$$\rho = \frac{\rho_0}{(1 + m^2)^{3/2}},$$

where $\rho_0$ represents the central density and $m$.

$$m^2 = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}, \quad a \geq b \geq c \geq 0,$$

is constant on an ellipsoidal shell. Following Chandrasekhar (1969), it is possible to obtain the potential for the Perfect Ellipsoid, which in ellipsoidal coordinates $(\lambda, \mu, \nu)$ adopts the Stäckel form:

$$V(\lambda, \mu, \nu) = -\frac{1}{4h_2^2} G(\lambda) - \frac{1}{4h_2^2} G(\mu) - \frac{1}{4h_2^2} G(\nu),$$

where $\alpha = -a^2$, $\beta = -b^2$, $\gamma = -c^2$ and $h_2^2$, $h_3^2$, $h_4^2$ are the metric coefficients of the ellipsoidal coordinates and $G(\tau)$ is a function given in terms of an elliptical integral of the third kind (see de Zeeuw 1985).

The Stäckel model has three explicit global analytic integrals, namely the adelphic integrals $I_2$ and $I_3$ and the total energy $H$, and indeed all orbits in it can be determined in terms of simple quadratures.

The adelphic integrals can be considered as generalizations of the angular momentum integrals that exist in axisymmetric and spherical potentials, but also as generalizations of the energy integrals always present in separable potentials in Cartesian coordinates. Every orbit in Stäckel potential is the sum of three motions, one in each coordinate. As a result, motion is bounded by coordinate surfaces.

The adelphic integrals $I_2$ and $I_3$ are, in fact, linear combinations of other integrals $J$ and $K$:

$$I_2 = \frac{\omega^2 H + \alpha J + K}{\alpha - \gamma}, \quad I_3 = \frac{\gamma^2 H + \gamma J + K}{\gamma - \alpha},$$

where the energy $H$, $J$ and $K$ are functions of the ellipsoidal coordinates and conjugate momenta (see de Zeeuw 1985).

As a consequence of the separability of the Stäckel potential, the equations of motion can be recast as:

$$p_\tau^2 = \frac{H - V_{eff}(\tau)}{2(\tau + \beta)},$$

where $\tau = \lambda, \mu, \nu$, and $V_{eff}(\tau)$ denotes the effective potential,

$$V_{eff}(\tau) = \frac{I_2}{(\tau + \alpha)} + \frac{I_3}{(\tau + \gamma)} - G(\tau).$$

This characteristic of the system allows us to analyze the orbital structure according to the values of $p_\tau^2$ for each set of chosen values of the integrals. Since the motion is only possible for $p_\tau^2 \geq 0$, $\tau = \lambda, \mu, \nu$, the motion in each coordinate is either an oscillation between turning points defined by $p_\tau^2 = 0$, or a rotation whenever $p_\tau^2 > 0$ for all $\tau$. The combination of these two kinds of motion will determine the class of orbit. The ranges where $p_\tau^2$ are non-negative define a volume space, determined by the integrals, where the motion takes place. Thus, the orbital structure of the system may well be studied by analyzing the values of (5) for
different combinations of chosen values for its integrals. Such
a systematic study can be performed by fixing one of the
integrals to compute the momenta in the space determined by
the two remaining ones.

de Zeeuw (1985) carried out a detailed classification of or-
bits for this Stäckel model by analyzing the tridimensional
space of integrals. A projection of this space on the energy
surface is shown in Fig. 1, where the two inner curves in the figure
separate the four different orbital families: box, inner and outer long
axis tubes and short axis tubes. These curves are separatrices of dif-
ferent kinds of motion and they play a central role in the dynamics
of the system when it is perturbed (see Sect. 4).

Although the properties of this model have already been
thoroughly studied (de Zeeuw & Lynden-Bell 1985), we will
complement them by computing and analyzing the resonance
structure of the system in its integrals space. The knowledge of
this resonance structure will allow us to predict the global dy-
namical behavior of the system when a non-integrable perturba-
tion is introduced.

3. Resonance structure of the integrable model

In order to understand how an integrable system is modified by
the effect of a non-integrable perturbation, we require a picture
of its resonance structure, since it yields a preliminary knowl-
edge of the way in which resonances in the integrable system
may be distorted when a perturbation is turned on, avoiding
the integration of the equations of motion and the variational ones.
Further, this gives us information about the mechanisms that
could give rise to the transition from regularity to chaos by dif-
ferent processes.

In the present section, we compute the resonance structure
for the integrable Stäckel model in the non-perturbed integral
space \( I = (H, I_3, I_2) \). At this stage we introduce action-angle
variables, which in many respects are the natural variables for
the description of the motion. Indeed, from the geometrical point
of view the actions (which are functions of the global integrals)
define the tori structure that foliate the phase space. For given
values of the action, the angle variables describe the orbit on the
torus.

Thus, we briefly outline the procedure for the case of motion
separable in ellipsoidal coordinates, whose results apply not
only to the particular case of the Perfect Ellipsoid, but to every
Stäckel potential. The action variables \( J_i \) associated with each
ellipsoidal coordinate are defined by

\[
J_t = \frac{1}{2\pi} \int p_t d\tau = \frac{2}{\pi} \int_{\tau_{\max}}^{\tau_{\min}} p_t d\tau, \quad \tau = \lambda, \mu, \nu, \quad (7)
\]

where, from Eqs. (5) and (6), the momenta \( p_t \) can be written
(de Zeeuw 1985)

\[
p_t = \frac{H - I_2/(\tau + \alpha) - I_3/(\tau + \beta) + G(\tau)}{2(\tau + \beta)}, \quad (8)
\]

and the integration is performed over all values of \( \tau \) for which
\( p_t^2 \geq 0 \).

From Eqs. (7) and (8) we obtain \( J_t = J_t(H, I_2, I_3) \) and de-
noting the action vector by \( \mathbf{J} = (J_1, J_2, J_3) \), we can formally
invert this relation in order to get \( H(\mathbf{J}) \). For given values of
the integrals, \( H = E, I_2 = i_2 \) and \( I_3 = i_3 \), we get the values of
the actions \( J_t \) that fix the torus where the motion takes place.
The associated angle variables \( \theta_t \) canonically conjugate to the
actions \( J_t \), are linear functions of the time \( t \),

\[
\theta_t = \omega_t(\mathbf{J}) \cdot t + \theta_t(0),
\]

where \( \theta_t(0) \) are constants determined by the three remnant initial
conditions and the quantities (integrals) \( \omega_t(\mathbf{J}) \) given by

\[
\omega_t(\mathbf{J}) = \frac{\partial H(\mathbf{J})}{\partial \mathbf{J}}
\]

are the frequencies of the motion in each degree of freedom.

If the three frequencies are incommensurable, the orbit
densely fills the volume allowed to it by the values of the inte-
grals of motion, i.e. the region where \( p_t^2 \) are non-negative. Thus,
for these particular values of the actions (or the integrals) we
have a strong non-resonant 3D torus on which the motion is er-
godic. This occurs when the frequencies satisfy the so-called
diophantine condition (see, for instance, Giorgilli 1990) which
states that for a given integer vector \( \mathbf{m} \) the frequency vector
satisfies

\[
|\mathbf{m} \cdot \omega| \geq \gamma |\mathbf{m}|^{-\alpha}, \quad |\mathbf{m}| = |m_1| + |m_2| + |m_3|,
\]

where \( \gamma > 0 \) and \( \alpha > 2 \) in case of a 3D system. Whenever
a relation of the form

\[
\mathbf{m} \cdot \omega(\mathbf{J}) = 0, \quad \mathbf{m} \in \mathbb{Z}^3/0
\]

is satisfied, we obtain a resonance condition for the actions or the
frequencies. Those values of the actions that fulfill (12) lead to
a resonant torus, that is, orbits are not ergodic on a 3D torus but
on a 2-dimensional one. It may be interpreted that the resonance
condition provides a relation between the actions that lead to
a new local integral which confines the motion to a manifold of
lower dimensionality.

In order to obtain the resonance structure on an energy sur-
face, a particular value of the energy \( H \) has to be fixed and the
resonance condition (12) solved for each value of \( \mathbf{J} \).

The frequencies for ellipsoidal coordinates are given by (10)
and, since actions do not admit explicit analytical expressions in
terms of the integrals, frequencies (10) have to be calculated by

![Fig. 1. Orbital Classification for the Perfect Ellipsoid in the I_2-I_3-plane, for a fixed value of the energy E = -0.3, de Zeeuw (1985). Curves drawn in this plane are separatrices of different families of orbits.](image-url)
numerical means. Indeed, since $H = H(J, J_2, J_3)$, we may write (de Zeeuw 1985)

\[
\frac{\partial H}{\partial J} = \omega_1 \frac{\partial J_1}{\partial J} + \omega_2 \frac{\partial J_2}{\partial J} + \omega_3 \frac{\partial J_3}{\partial J} = 1,
\]

\[
\frac{\partial H}{\partial J_2} = \omega_1 \frac{\partial J_1}{\partial J_2} + \omega_2 \frac{\partial J_2}{\partial J_2} + \omega_3 \frac{\partial J_3}{\partial J_2} = 0,
\]

\[
\frac{\partial H}{\partial J_3} = \omega_1 \frac{\partial J_1}{\partial J_3} + \omega_2 \frac{\partial J_2}{\partial J_3} + \omega_3 \frac{\partial J_3}{\partial J_3} = 0.
\]

Then,

\[
\omega_1 = \frac{1}{\Delta} \frac{\partial (J_1, J_2, J_3)}{\partial (J_2, J_3)}, \quad \omega_2 = \frac{1}{\Delta} \frac{\partial (J_1, J_2)}{\partial (J_2, J_3)} \omega_3 = \frac{1}{\Delta} \frac{\partial (J_1, J_3)}{\partial (J_2, J_3)},
\]

where we have written

\[
\Delta = \frac{\partial (J_2, J_3)}{\partial (J_2, J_3)}.
\]

The partial derivatives that occur in the above expressions can be evaluated by differentiating under the integrals, (7)

\[
\frac{\partial J_2}{\partial \tau} = 2 \pi \int_{\tau_{\text{min}}}^{\tau_{\text{max}}} \frac{\text{d}\tau}{(\tau + \alpha)(\tau + \beta)p_\tau},
\]

\[
\frac{\partial J_2}{\partial \tau} = 2 \pi \int_{\tau_{\text{min}}}^{\tau_{\text{max}}} \frac{\text{d}\tau}{(\tau + \alpha)(\tau + \beta)p_\tau},
\]

\[
\frac{\partial J_3}{\partial \tau} = 2 \pi \int_{\tau_{\text{min}}}^{\tau_{\text{max}}} \frac{\text{d}\tau}{(\tau + \alpha)(\tau + \beta)p_\tau}.
\]

The values $\tau_{\text{min}}$ and $\tau_{\text{max}}$ define the range in each ellipsoidal coordinate where the motion is allowed, that is, the coordinate space regions where the condition $p_\tau^2 \geq 0$ is satisfied. The above formulas permit the calculation of the actions $J_i$ and the frequencies $\omega_i$ without integrating the equations of motion. For a Stäckel potential, and given values $E, I_2, I_3$ of the integrals of motion, the actions follow from (7), and the frequencies can be computed from (14) and (17).

Thus, in order to compute the frequencies for particular values of $H$, $I_2$, and $I_3$, we have to integrate (17) by numerical means and obtain the roots $p_\tau^2 = 0$, which provides the values of $\tau_{\text{min}}$ and $\tau_{\text{max}}$. This numerical procedure restricts the accuracy with which the frequencies are obtained. We have calculated the integration limits in (17) by means of the zbrent subroutine (Press et al. 1994), setting the tolerance parameter equal to $10^{-10}$. Further, we have performed the numerical integration for (17) by means of the Romberg method -implemented in the midpoint and gromo subroutines (Press et al. 1994) – with an accuracy of the order of $10^{-10}$. Also, for particular values of the integrals $H$, $I_2$, and $I_3$, the functions involved in (17) are not well defined for $\tau_{\text{min}}$ and $\tau_{\text{max}}$. This shortcoming has been avoided by shrinking the integration ranges to $(\tau_{\text{min}}^+, \tau_{\text{max}}^-, \text{d}\tau)$, $\text{d}\tau \sim 10^{-10}$ being a suitable value as a compromise between the need to minimize the final errors in the frequencies and the computational time required to the whole numerical procedure.

From the analysis of all the above factors that affect the accuracy with which the frequencies can be obtained, in practice we take as the resonance condition the relation

\[
|m \cdot \omega(J)| < 10^{-4},
\]

which determines the different resonant surfaces in the integrals space $(H, I_2, I_3)$.

In order to visualize the resonance structure of the Stäckel model, we fix a value for the energy $H$, which defines the energy surface in the integrals space. Figure 2 shows the resonance structure for the Perfect Ellipsoid in the $I_2I_3$-plane, corresponding to $H = -0.3$, this particular value of the energy being in the range $-0.9074 \leq E < 0$ corresponding to the Stäckel model, representative of the dynamical behavior of the system. The curves shown in this figure arise from the intersections between the energy surface and several resonant surfaces calculated from (17) in terms of $I_2$ and $I_3$, for different resonant vectors $m$ satisfying $|m| = |m_1| + |m_2| + |m_3| < 8$ (see text).

In order to understand how the resonance structure of the Stäckel model determines the orbital structure of the system, in Fig. 3 we present three particular resonances, namely, $(1, -3, 2), (1, -2, 1)$, and $(3, -1, -1)$, as well as three tori located in different regions on the $I_2I_3$-plane. Over these tori, orbits will proceed with different dynamical features depending on their location with respect to the resonant curves. Thus, the nonresonant torus depicted in the figure lies in a region of the phase space where the resonant condition does not hold for any $m$ (or where the diophantine condition (11) holds), and any orbit lying on such a 3D torus will cover it densely and uniformly as $t \rightarrow \infty$. As mentioned above, in the 3D torus the motion is ergodic. Meanwhile, motion on a torus located on a resonant curve is such that any trajectory inhabits a submanifold of dimensionality two as a consequence of the resonance condition. Therefore, an orbit on such a torus will densely cover a 2D (elliptic) torus, the resonant torus. Closure in configuration space requires an additional, independent resonance condition. Any torus located on the intersection of two resonances will lead to a periodic orbit since the two independent resonance conditions restrict the motion to one dimension (a 1D torus).

Figures 1 and 2 show the orbital structure in the Stäckel model, the internal curves in Fig. 1, given by $I_2 = 0$ and $E = V_{\text{eff}}(-\beta)$, play an important role as they bound as well as separate the four regions corresponding to the four families of orbits of the Perfect Ellipsoid. The separatrices are very unstable, even a very small perturbation dramatically modifies the dynamics on them. In the integrable model the smooth
separatrices are present because the stable and unstable manifolds of a given object (in this case a 2D hyperbolic torus) exactly match. However, in the surroundings of the separatrices, the resonance structure is rather intricate.

The resonance structure of the integrable Stäckel model allows us to forecast what the dynamical behavior of the system could be if a non-integrable perturbation were added. As it will be discussed in the next section, the interaction of resonances are responsible for the start of chaotic motion when a perturbation is turned on.

4. Theoretical discussion of the perturbative effects

Though a deep insight into the dynamics of the perturbed Stäckel system gained by numerical means will be addressed in a forthcoming paper, several of its main features can be inferred from the resonance structure of the integrable model.

We review a theoretical discussion of how a perturbation would modify the dynamics of the integrable Stäckel model, following the general description given by Chirikov (1979) and Cincotta (2002). We consider a system represented by the Hamiltonian

$$H(J, \theta) = H_0(J) + \epsilon V(J, \theta),$$

where

$$\epsilon V(J, \theta) = \epsilon \sum_{m} V_m(J) \cos(m \cdot \theta).$$

Here \((J, \theta)\) denote the tridimensional action-angle coordinates for the unperturbed Hamiltonian \(H_0\), \(m\) is a tridimensional integer vector, \(V_m\) certain real functions and \(\epsilon\), the perturbation parameter, is a real number.

Elliptical galaxies might be represented by a smooth triaxial potential \(H_0\), like the Stäckel one, plus a perturbation given by a multipolar expansion, \(\epsilon V\). We therefore assume that the perturbation is analytic, which may not be the case if we add a central singularity.

For \(\epsilon = 0\) we have \(H = H_0(J)\) and the motion is stable\(^1\) for any initial condition, since we have the complete set of the three integrals of motion \(J_i\) (\(H_0\) being cyclic in \(\theta\)), i.e. the Hamiltonian is completely integrable. The phase vector evolves linearly with time with a frequency vector \(\omega(J) = \partial H_0/\partial J\), with \(\det(\partial \omega(J)/\partial J) \neq 0\) so that \(\omega(J)\) is a one-to-one function. This condition guarantees the nonlinearity of the system.

The presence of the perturbation in general disrupts the integrability of the Hamiltonian \(H_0\) (due to the phase dependence of \(V\)), leading to a variation of the unperturbed integrals \(J_i\). The stability of the motion breaks down when a large change in the actions takes place, i.e., when a “gross” instability sets up.

To describe the motion of a star in the Hamiltonian (18) and (19), we take advantage of a perturbative technique to obtain approximate solutions, assuming that the perturbation parameter is small, \(\epsilon \ll 1\), that is, we consider a near-integrable Hamiltonian system. The perturbative approach given by the so-called asymptotic series implies, roughly, that the variation of the unperturbed actions is computed via a power series in the perturbation parameter.

It is well known that the effect of a perturbative Fourier component (as in (19)) is stronger as long as the time variation of its phase, \(\phi_m = m \cdot \theta\), is slow. In the limit case of constant phase we reach the resonance condition for the unperturbed frequencies. If we are far from a resonance (i.e., the initial conditions correspond to values of the actions or the frequencies that satisfy the diphase condition (11)), it can be shown that the motion is stable. On taking advantage of the so-called averaging method, we neglect the oscillating part of the perturbation, retaining only its average value.

When we are near to a resonance condition, the asymptotic series technique no longer work due to the appearance of the so-called small denominators in the coefficients of the asymptotic series. These small denominators are the resonant values \(\omega_m = m \cdot \omega\) that may lead to divergent series. It can be shown that the set \(\omega_m\) for all integer vectors is, in general, dense everywhere in phase space. Therefore, to find initial conditions “far from a resonance” is not straightforward.

The geometric features of resonances in the action or integral space are represented in Figs. 2 and 3. Any point on this plane is a torus since the “position vector” is given by the three values of the actions (or the integrals). In action space, the resonance equation \(m \cdot \omega(J) = 0\) leads to another 2-dimensional surface, whose local normal at the resonant point \(J = J\) is

$$n' = (\partial (m \cdot \omega(J))/\partial J)_{J'}.$$

Further, the conservation of the unperturbed energy provides the 2-dimensional energy surface \(H_0(J) = E\). In what follows we will consider only convex Hamiltonians, a condition that is verified by the Stäckel Hamiltonian. The subspace defined by the intersection of both the resonant and the energy surfaces has dimension 1.

By definition, the frequency vector \(\omega\) is normal to the energy surface in the action space. The latter condition, together with the resonance equation, shows that the resonant vector \(m\) lies in the plane tangent to the energy surface at \(J = J\). Furthermore, a simple inspection of the equations of motion for only one resonant perturbing term, that is \(\epsilon V = \epsilon V_m(J) \cos(m \cdot \theta)\), shows that \(J\) is parallel to the resonant vector \(m\). This picture of the dynamics allows us to conclude that the motion under a single resonant perturbation proceeds on the tangent plane to the energy surface at the point \(J = J\) along the direction of the resonant vector. As \(\epsilon \to 0\), the resonant perturbation preserves the unperturbed energy.

The above discussion considering only one resonant term shows that the perturbation (19) only depends on a single phase,
namely the resonant phase $m_t \cdot \theta$. If we perform a canonical transformation in such a way that one of the three new phases, say $\psi_1$, is such that $\dot{\psi}_1 = m_1 \cdot \theta$, then the resulting Hamiltonian will be cyclic in the remaining two phases $\psi_2, \psi_3$ and the new momenta $p_2, p_3$ will be integrals of the motion.

This canonical transformation may be done through

$$\psi_1 = \mu_2 \theta; \quad \dot{J}_1 = J'_1 + p_2 \mu_1,$$

where the sum over repeated indexes is understood and $\mu_k$ is a $3 \times 3$ matrix. This linear transformation could be thought of as a local change of basis where the new basis vectors are:

$$\mu_1 = m_y, \quad \mu_2 = \left( \frac{\omega(J')}{|\omega(J')|} \right) \mu_3 \equiv e$$

the being latter normal to both $\mu_2$ and to $n^\prime$. In general $e$ will not be normal to $m_y$.

According to the discussion given by Chirikov (1979) and Cincotta (2002), the Hamiltonian (18) and (19) in the vicinity of a single resonant perturbation reduces to a simple pendulum model:

$$H_r(p_1, \psi_1) \approx H_0(J') + \frac{p_1^2}{2M} + eV_m(J') \cos \psi_1,$$

where $H_r$ stands for the resonant Hamiltonian and $M$ is the “non-linear mass” given by: $1/M = m_\mu (\partial \omega / \partial J_\lambda) - m_k$. Thus $p_1$ changes with time following the pendulum model for a given value of the integral $H_r$ in the $m_\mu$ direction. The remaining momenta $p_2, k = 2, 3$ are set to zero so that $J'$ is a point of the trajectory. Therefore, since $H_r$ is independent of time, we again have the full set of three (local) integrals of motion for the problem of a single resonant perturbation: $H_r, p_2, p_3$.

The variations of the unperturbed actions are determined by the strength of the perturbation and the non-linear mass, in the fashion (see Cincotta 2002 for further details)

$$(\Delta J)_i \equiv |J - J'| \sim \sqrt{V_m(J')} |\epsilon|.$$

Thus, the resonant surface of the unperturbed system in action space becomes a resonant layer, whose width is given by $(\Delta J)_i$.

Since $J - J' = p$ and in the new basis $p = p_1 \mu_1 + p_2 \mu_2 + p_3 \mu_3$, we note that $p_2$ and $p_3$ being local integrals, the motion proceeds from the resonant point in the direction of $\mu_1$, i.e. across the resonance.

Cincotta (2002), following Chirikov (1979), discusses how different terms in the perturbation modify the dynamics of the integrable system $H_0$. For clarity we recast (18) and (19) as

$$H(J, \theta) = H_0(J) + eV_m(J) \cos(m_3 \cdot \theta) + eV_m(J) \cos(m_2 \cdot \theta) + e \sum_{m=m_0}^{m_\mu} V_m(J) \cos(m \cdot \theta),$$

If we take into account all terms in (25), all momenta change with time and then variations in $p_2$ and $p_3$ could occur, since now the Hamiltonian depends on $\theta_2$ and $\theta_3$. These variations may proceed along the $\mu_2$ and $\mu_3$ directions, and recalling that $p_2$ is normal to the unperturbed energy surface, the variations in $p_2$ take into account small changes in the unperturbed energy when the perturbation is switched on. These variations are bounded and of order $\epsilon$. However, $\mu_3$ is tangent to the energy surface and by definition it is directed along the resonance, so one could expect variations in $p_3$ that, at first sight, will be unbounded (see next section).

Let us now discuss the effects of the different terms in (25). As shown above, on taking $V_m = 0$ for all $m \neq m_\mu$ and picking up initial conditions such that $m_\mu \cdot \omega(J') = 0$, the torus structure in the vicinity of $J'$ is preserved because of the existence of the three local integrals $H_r, p_2$ and $p_3, H_r(p_1, \psi_1)$ being the pendulum Hamiltonian. Depending on the value of $H_r$, the momenta $p_1$ will oscillate or rotate around $J'$. The resonant layer corresponds to the oscillation domain of the pendulum energy. As it is well-known, oscillations and rotations in a pendulum are separated by a smooth separatrix $p_1(\psi_1')$.

If we switch on $V_m$, which we assume to be the dominant non-resonant perturbing term, its main effect on the dynamics of the system in the vicinity of $J'$ would be to split the stable and unstable manifolds associated with the unstable points of the pendulum leading to a splitting of separatrices and the smooth curve $p_1'(\psi_1')$ becomes a stochastic layer of finite width. Thus, under the effect of this perturbative term, oscillations and rotations are actually separated by a layer of chaotic motion; $p_1(\psi_1')$, $p_1'(\psi_1')$ shows a bounded unstable, chaotic behavior confined to this thin layer. Figure 4 illustrates this fact for two given resonances in the Stäckel model in the unperturbed integral space $(H_0, I_2, I_3)$ for a fixed value of $H_0$. There, the local basis $\mu_1', \mu_2', \mu_3'$, at a given resonant torus $J'$, has been included, the $\mu'_i$ the $\mu_i$ at $J'$ being mapped to the integral space.

If we turn on the next relevant term in (25), say $V_m$, it is possible to show that this third term might cause unbounded chaotic variations in $p_3$ while $p_1$ lies in the stochastic layer; there could be motion in the $\mu_3$ direction, that is, along the stochastic layer of the resonance.

Figure 5 illustrates in a schematic way how the whole resonant structure of the Stäckel model (given in Fig. 2) could be distorted by the effect of a generic perturbation. All resonances would become layer, whose widths $(\Delta J)$ strongly depend on $J'$, $m, V_m$ but all of them have order $\epsilon$.

Geometrically, when the system is exactly at some resonance, we have a 2D elliptical torus, while when the system is on the border of the resonance we get a 2D hyperbolic torus. Thus the center of a resonance layer over the integral space form a chain of elliptic torus while the border of the resonance layer corresponds to a chain of hyperbolic torus. This is schematically shown in Fig. 6 for the same resonances of the Stäckel potential depicted in Fig. 5.

All the orbits trapped in resonances constitute subfamilies of the Stäckel orbital families. As we discussed at the
beginning of this section, if the system is far from any resonance, the motion is stable. The presence of the perturbation only produces small changes of order \( \epsilon \) in the unperturbed actions. Thus a non-resonant value of the unperturbed actions or integrals leads to a 3D non-resonant torus that support quasi-periodic orbits. Accordingly to Eqs. (7) and (8), the three local integrals are \( H = H_0 + \epsilon V, I_2 = I_2 + \delta \delta I_2, I_3 = I_3 + \delta \delta I_3, \delta I_2 \) being bounded constants. Therefore, in the neighbourhood of the values of the unperturbed integrals \( (I_2, I_3) \) the orbital structure is preserved.

On the other hand, orbits trapped in a single resonance at \( J' \) also respect three local integrals that in action space are the pendulum Hamiltonian \( H_p \) and two independent linear combinations of the unperturbed actions at the resonant value, for instance, \( K_1 = a_1 J'_1 + a_2 J'_2, K_3 = a_3 J'_1 + a_4 J'_2 \), where the coefficients \( a_i \) depend only on the \( \mu_k \). For \( H_p = 0 \) we are exactly on a 2D elliptical torus that supports a stable resonant orbit that is the same as in the unperturbed system (since \( p_1 \equiv 0 \) and \( J = J' \)). For positive values of \( H_p \) new 3D tori appear that support nearly resonant orbits that respect the above mentioned local integrals. The nature of these new orbits are different to that of the unperturbed system at \( J' \). When \( H_p \) takes the value that corresponds to the pendulum separatrix, we arrive at to the 2D hyperbolic tori that support unstable orbits.

5. Discussion on the origin of chaos and diffusion

The dynamical description becomes highly intricate when we consider the interaction among all the perturbative terms. Indeed, as mentioned above, the effect of a perturbation on resonance is to form a thin stochastic layer at the border of the resonance layer. Since the resonance layer has a width \( \sim \sqrt{\epsilon} \), as long as the perturbation increases its strength, the resonances become wider and the overlap of the stochastic layers of different resonances could take place when \( \epsilon \) reaches some critical value \( \epsilon_c \). When this occurs, the resonances overlap. The dynamics in this case become unstable due to the intersections of the stable and unstable manifolds of different resonances. The resulting motion is completely stochastic (see Cincotta 2002) whose main characteristic is its local instability. Figure 5 provides a schematic illustration of the effect of a perturbation on the resonant structure. From this figure, we can clearly see that if we increase the perturbation, a massive overlap of resonances will occur. Due to the dimensionality of the system (3D), resonance intersections take place (which is not the case in 2D systems). Around the point of intersection, the stochastic layers of different resonances overlap, as Fig. 5 shows schematically.

For the time being, an analytical description of the dynamics at the resonance intersection is still lacking, because if we retain two resonant terms in the expansion (25), for instance \( V_{m_p} \) and \( V_{m_m} \), the Hamiltonian depends then on two resonant phases \( (m_p - \theta) \) and \( (m_m - \theta) \), and as can be shown, this system is not integrable. A picture of the dynamics at a resonance crossing is given in Cincotta et al. (2003), where a zoom of the dynamical behaviour at the region close to the intersection of two resonances with their corresponding stochastic layers is presented. For this case, the intersection of these two resonances generates a stable periodic orbit and a central region of regular or mild chaotic motion bounded by a thin layer of unstable motion that presents at the same time several stable domains separated by extremely thin chaotic filaments. However, it is possible to infer that at the intersection of the unperturbed resonances a periodic orbit appears whose stability cannot be determined a priori and that the zone of stochastic motion should be bounded (see Chirikov 1979).

The known mechanisms that lead to a transition from regularity to chaos are overlap of resonances (including resonance crossings) and Arnold diffusion-like processes (see for instance, Chirikov 1979; Cincotta 2002; Giordano & Cincotta 2004).

As we have already stated, chaos means variation of the unperturbed integrals. This is usually called chaotic diffusion in the literature. Unfortunately no theory exists that describes global diffusion in phase space. In other words, it is not possible to estimate either its rate or its direction or route. Although one can obtain accurate values of the Lyapunov exponents, the KS entropy, the MEGNO (see Cincotta et al. 2003; Cincotta & Simó 2000) or any other indicator of the stability of the motion, they provide only local values of the diffusion rate. A given orbit in a chaotic component of the phase space could have two positive and large values for the Lyapunov exponents which does not mean that the unperturbed integrals would change much. This is a

Fig. 5. Schematic representation of the effect of a perturbation on the resonant structure of the St"ackel model (given in Fig. 2). All the resonances become layers of width \( \Delta J \).

Fig. 6. Geometric structure of the perturbed resonances shown in Fig. 5. The central curve, that corresponds to the exact resonance is a chain of 2D elliptical tori, while the borders of the resonance correspond to a chain of 2D hyperbolic tori.
natural consequence of the structure of the phase space of almost all dynamical systems like galaxies.

Thus what is actually relevant is the extent of the domain and the time-scale over which the diffusion may occur. In a recent work Giordano & Cincotta (2004) showed numerically that in models similar to that of an elliptical galaxy, the time-scale over which diffusion becomes relevant is several orders of magnitude of the Hubble time. On the other hand, in models corresponding to planetary or asteroid dynamics, diffusion may occur over physical time-scales.

One can make predictions about the global dynamical behaviour of a perturbed system by studying its resonance structure. For example, in Fig. 2 we observe a rather intricate resonance structure in the box orbits domain, with several crossings of resonance as well as the fact that they are very close to each other. Thus, we expect that this region would be dominated by chaotic dynamics when the perturbation is activated, for instance by adding a mass concentration in the center of the system, the box domain in the $I_1 I_2$-plane will be the first to be dominated by chaos. This is also true for any other kind of integrable model and perturbation, whatever its strength at the center may be (see, for instance Papapliopoulou & Laskar 1998).

We could come to a similar conclusion for those regions close to the separatrices which separate different families of orbits, $I_2 = 0$ and $E = V_{e ff}(\beta)$. Several resonance intersections along these separatrices can be observed, then, when a non-integrable perturbation is added, the dynamics in this region would become rather intricate due to the fact that the stable and unstable manifolds of different 2D hyperbolic tori will tangle in a very complicated fashion (see, for instance Cincotta et al. 2003). This explains the instabilities found by de Zeeuw (1985) along the $I_1$-axis; in fact, in the St"{u}ckel model all 2D tori with $I_2 = 0$ are unstable (hyperbolic), so that the dynamics in this region becomes rather chaotic under the effect of a non-integrable perturbation.

An important fact to be stated is that when chaos sets up, the unperturbed global integrals (or actions) have a discontinuous dependence on phase space variables. Indeed, close to the resonant torus, despite the existence of three local integrals, the unperturbed orbital structure is not preserved and the topology of the phase space changes. Moreover, on the stochastic layer at least one integral does not exist.

Thus, in the St"{u}ckel model, for $\epsilon = 0$ we get Fig. 2, while for $\epsilon \ll \epsilon_c$ the dynamic picture will be similar to that given in Fig. 5. Resonances do not present a significant overlap and chaotic motion will be confined to the narrow stochastic layers and around the resonance crossings. Thus we cannot expect large variations of the unperturbed integrals and therefore chaotic diffusion should be irrelevant. However, theoretically, it might be possible that chaotic diffusion takes place even for very small values of the perturbation ($\epsilon \ll \epsilon_c$). Since for very small $\epsilon$ values the resonance structure is still present, one could expect that the perturbative terms drive diffusion along the stochastic layers of a given resonance, i.e. in the $\mu_1$ direction.

Indeed, Arnol’d (1964) proved for a very specific Hamiltonian, that there exists a mechanism that could connect two distant tori in the chain of 2D hyperbolic tori located on the border of the resonance. This means that variations of the unperturbed actions along the resonance occur. Since the actual motion around any hyperbolic tori is stochastic, chaotic diffusion along the stochastic layer of a given resonance could take place. The mechanism that allows motion in the $\mu_1$ direction is the so-called Arnol’d Mechanism (see Giorgilli 1990). However, the existence of a transition chain of hyperbolic tori does not mean that stochastic motion would spread over the whole web of resonances. Moreover, in general, it is not possible to find this transition chain in most non-linear Hamiltonian systems, even in the case of the simpler ones. Nevertheless, it is often assumed that this is the case and that chaotic motion will take place over the whole resonance web. This implies that a global instability exists and it is usually called Arnol’d diffusion. This conjecture supports, for instance, concepts like mixing or ergodicity, among others. Let us recall that the global instability properties of near integrable Hamiltonian systems are far from being well-understood after the pioneering work of Arnol’d (see Lochak 1999).

For $\epsilon \gg \epsilon_c$, the overlap of resonances begins to operate destroying the torus structure of phase space and the resonance web disappears, and chaotic diffusion begins. In this case, we expect that diffusion could operate through overlap of resonances, but the extent of the diffusive zone will be confined, at most, to the overlapping domain. There is no guarantee that the whole chaotic component of phase space is connected, even in the case of massive overlap (Giordano & Cincotta 2004). But also in this case, it is usually stated that the chaotic component of phase space is fully connected through Arnol’d diffusion. Arguments based on these yet unproved statements have given rise to many of the current ideas taken for granted on dynamics in elliptical galaxies, as can be verified for instance in the works of Merritt & Valluri (1996), Merritt (1999) and Udy & Penniger (1988) among others.

Therefore, in a theoretical approach, while it is possible to get much information about the dynamical behaviour of the system as long as the perturbation increases its strength, nothing can be concluded about the diffusive processes in phase space of near-integrable systems. Numerical integrations become the only way to investigate if chaotic diffusion could play a significant role in particular models of galaxies. This will be addressed in a forthcoming paper using the St"{u}ckel potential discussed here.

We clearly discriminate the dynamical behaviour of the system depending on the values of the integrals. Close to strong non-resonant tori, the orbital structure of the unperturbed St"{u}ckel potential is preserved and the local integrals are corrections of order $\epsilon$ of the unperturbed global integrals. On the other hand, when the system is close to a resonant tori, the unperturbed orbital structure is not preserved, new local integrals appear and the topology of phase space changes. These new local integrals are the pendulum Hamiltonian and linear combinations of the unperturbed actions at the resonant point.

Therefore, in such a system, it is not possible to assume that the distribution function of the galaxy, in the whole regular component, has the form $f(I_2, I_3)$. This could be true only for strong non-resonant tori, but since resonances are dense in phase space, the distribution function locally exists and has the form $f_r(H, I_2, I_3) + \epsilon g(H, I_2, I_3)$ in neighbourhood of non-resonant tori that support quasi-periodic orbits and $f_r(H, K_2, K_3)$ in vicinity of resonant tori.

On the other hand, nothing can be said about the dependence of $f$ in the chaotic region. In a similar fashion, since there is no theoretical support to argue that the whole chaotic region is fully connected, one should define a local distribution function only in those regions of phase space that could be visited by a single orbit. It is very likely that in some regions two local quasi-invariants exist (for example in the stochastic layers of the resonances) and in some other regions only the energy remains constant. Clearly, discontinuous dependence of $f$ on the integrals, and consequently on phase space variables, is expected.
On the other hand, much progress can be made in the understanding of diffusive processes in phase space of near-integrable Hamiltonian systems, including the dynamical behaviour around resonance crossings. Indeed, as all the evidence shows, galaxies should exhibit a divided phase space, with both regular and chaotic components, which is the phase space structure of near-integrable Hamiltonians for low-to-moderate perturbations.

Taking into account these arguments that are supported by theoretical considerations, we suggest that the direction to take in constructing equilibrium models of elliptical galaxies.

The selected representation of an elliptical galaxy does not encompass the case of a central singularity. In such a case, one deals with a non-analytic perturbation which admits no Fourier expansion as (18) and the perturbative approach is no longer suitable. Nevertheless, it is clear that such a perturbation $\sim 1/r$ mainly affects the dynamics in the box domain due to the fact that these orbits pass arbitrarily close to the center: box orbits are strongly scattered producing in general an almost spherical configuration near the center (see however Poon & Merritt 2002). Therefore, the perturbative approach addressed here should be reformulated so that a singularity at the origin can be taken into account. This is not a simple task. Although there is observational evidence of central mass concentrations in elliptical galaxies that could be black holes (see for instance Ferrarese & Ford 2005), the question of whether they actually are supermassive black holes is still an open matter.

Due to the fact that the described perturbative approach succeeds in studying actual physical problems such as planetary and asteroidal systems or dynamics in particle accelerators as well as the motion of charged particles in magnetic bottles, we argue that the theoretical formulation discussed here, which the only one developed so far, is a fairly useful tool when dealing with galactic systems.

Although one may argue that the action-angular variables do not constitute an adequate set to yield a description of galactic systems, there exists numerical evidence of galactic models exhibiting large regions of phase space corresponding to regular motion. Therefore, the Hamiltonian of a star moving in a galactic potential can always be written in the fashion $H(J) = H_0(J) + \epsilon V(J, \Theta)$, the key being what should be considered as $V(J, \Theta)$. Since all these mathematical theories have served well to provide a thorough description of chaotic systems, they may be of use to yield a good first order approximation of the motion of a star in a smooth stationary potential. Moreover, despite the fact that the actual dynamics of a galactic system comes from numerical simulations, theory that permits us to elucidate the matter, at least in an heuristic way is still useful.

Acknowledgments. This work was partially supported by grants of the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) and Fundación Antorchas. The authors acknowledge the valuable comments and suggestions of the referee.

References

Arnold, V. I. 1964, Sov. Math.-Dokl., 5, 581
Chirikov, B. 1979, Phys. Rep., 52, 63
Cincotta, P. M. 2002, New AR, 46, 13
Cincotta, P. M., & Simó, C. 2000, A&AS, 147, 205
Cincotta, P. M., Giordano, C. M., & Simó, C. 2003, Phys. D, 182, 151
Giorgilli, A. 1990, in Les Méthodes Modernes de la Mécanique Céleste, ed. C. Froeschlé, & C. Benest, 249
Kuzmin, G. G. 1956, AZh, 33, 27
Kuzmin, G. G. 1973, in The Dynamics of Galaxies and Star Clusters, ed. T. B. Oinarov, 71
Louchay, P., in Hamiltonian Systems with Three or More Degrees of Freedom, ed. C. Simo, NATO ASI (Dordrecht: Kluwer), 168