

Chapter 2

Angle-Action Estimation in a General Axisymmetric Potential

2.1 Introduction

In the study of dynamical systems, it is becoming increasingly important to be able to process and understand large multi-dimensional data sets efficiently. The stars in our own Galaxy, the Milky Way, are being increasingly observed in full six-dimensional phase-space through the combination of astrometry and radial velocity measurements. Full 6D phase-space information is currently available for stars in the solar neighbourhood from the Geneva-Copenhagen and RAVE surveys (Nordström et al. 2004; Zwitter et al. 2008) and this is to be greatly expanded on by the space mission *Gaia* (Perryman et al. 2001). Beyond our Galaxy, the advent of integral-field spectroscopy has led to projects such as SAURON (Bacon et al. 2001 and subsequent papers), which mapped the kinematics of a representative sample of 72 nearby elliptical and spiral galaxies, and subsequently ATLAS3D (Cappellari et al. 2011 and subsequent papers), which combined SAURON observations with CO and HI observations to study the kinematics of a complete volume-limited sample of 260 local early-type galaxies.

Although no galaxy is ever in perfect dynamical equilibrium, equilibrium dynamical models are central to the interpretation of observations of both our Galaxy and external galaxies. N -body simulations of cosmological clustering yield a picture in which dark-matter haloes are far from dynamical equilibrium only during short-lived and quite rare major mergers. In general a dark-matter halo can be well approximated by a dynamical equilibrium that is mildly perturbed by accretion. A major reason for the importance of equilibrium models is that we can infer a galaxy's gravitational potential, and thus its dark-matter distribution, only to the extent that the galaxy is in equilibrium. Moreover, equilibrium models are the simplest models and more complex configurations, involving spiral structure or an on-going minor merger for example, are best modelled as perturbations of an equilibrium model. Schwarzschild modelling (Schwarzschild 1979) constructs equilibrium models by describing the configuration of a model by a weighted set of orbits. However, this approach is not the most natural as each orbit is characterised by its initial phase-space coordinates.

Additionally, observational data is often understood by performing large N -body simulations. Whilst such models are straightforward to produce, the configurations of the models are difficult to control and characterise. It is necessary that techniques are developed which can simplify both observational and simulation data without losing the richness of the phase-space information.

2.1.1 Angle-Action Variables

The natural way to model a dynamical equilibrium is via Jeans theorem, which assures us that the system's distribution function (DF) can be assumed to be a non-negative function of isolating integrals. Since one expects a smooth time-independent gravitational potential to admit up to three functionally independent isolating integrals, Jeans theorem states that we should be able to represent an equilibrium stellar system by the density of stars in a three-dimensional space of integrals rather than in full six-dimensional phase space. This reduction in dimensionality makes the system very much easier to comprehend and model.

Since any function of integrals is itself an integral, infinitely many different integrals may be used as arguments of the DF. However, the action integrals J_i stand out as uniquely suited to be used as arguments of the DF. Along with the angles, the actions form a set of canonical coordinates that can be used to express the equations of motion in a trivial form: the actions are integrals of the motion whilst the angles increase linearly with time. Such a formulation instantly reduces the complexity of any dynamical data set. Along an orbit, the six phase-space dimensions are reduced to three angle coordinates. Angle-action variables can be defined for any quasi-periodic orbit. Initially introduced to study celestial mechanics, angle-action variables now have great potential for galaxy dynamics due to their attractive properties. It is particularly convenient to use actions as arguments of the DF as (i) they are adiabatic invariants, (ii) the zero-point of an action is well defined and (iii) the range of values an action may take is independent of the other actions. The angle-action variables also provide a basis for the development of a perturbative solution to the equations of motion (see Binney and Tremaine 2008 for a much fuller discussion of the merits of angle-action variables). Finally, we mention that angle-action variables can be used to hunt for substructure in the Galaxy (e.g. McMillan and Binney 2008), and will be used in Chap. 6 to model tidal streams.

Despite the aforementioned advantages, angle-action variables remain awkward to work with in practical applications due to the difficulty of their calculation in a general potential. They are easily calculated when the potential is spherical and with more work can be numerically calculated when the potential is of Stäckel form. However, neither of these approaches is satisfactory when working with realistic galaxy potentials, because such potentials do not satisfy these conditions. The development of methods to estimate angle-action variables in a general potential is crucial if we are to benefit from the advantages of angle-action variables and the wealth of techniques that utilise them. For some applications, such as stream modelling, we require

accurate actions for a few data points, such that we can afford to spend a long time evaluating the actions. However, for other applications, such as disc modelling, we require a fast algorithm for estimating the actions of many data points. Therefore, a variety of algorithms are required.

In this chapter, we present methods for estimating angle-action variables in a general axisymmetric potential. The methods presented in this chapter fall into two categories: slower convergent algorithms that produce very accurate actions, and faster non-convergent algorithms that produce less accurate actions. The majority of these methods are based on the analytic calculation of the actions in an axisymmetric Stäckel potential, so, in Sect. 2.2, we give a brief overview of the determination of angle-action variables in this class of potential. We continue by presenting a series of non-convergent methods based on the Stäckel potential. In Sect. 2.3, we present a method for estimating the angle-actions by locally fitting a Stäckel potential to the region a given orbit explores in an axisymmetric potential. The actions and angles in the general axisymmetric potential are estimated as those in the fitted Stäckel potential. The results of the method are examined by analysing artificial data in Sect. 2.3.3, and we demonstrate the practical application of the method by inspecting the Geneva-Copenhagen Survey in angle-action space in Sect. 2.3.4. In Sect. 2.4, we present the adiabatic approximation from Schönrich and Binney (2012) that is based on separability in cylindrical polar coordinates. In Sect. 2.5, we present a new method that improves on the adiabatic approximation by assuming separability in a general prolate spheroidal coordinate system. In Sect. 2.6, we present the method from Binney (2012) that assumes the potential is close to a Stäckel potential. Finally, we show how these non-convergent methods can be combined with torus construction (McMillan and Binney 2008) to produce a convergent method for finding the actions from a phase-space point $(\mathbf{x}, \mathbf{v})$. The results of all of these methods are compared in Sect. 2.8. The work on which this chapter is based was published by Sanders (2012).

2.2 Actions and Angles in a Stäckel Potential

The most general class of potentials in which we are able to calculate the angle-action variables analytically is that of Stäckel potentials. In a confocal ellipsoidal coordinate system, these potentials produce separable Hamilton-Jacobi equations. A full discussion of Stäckel potentials is given in de Zeeuw (1985). Here we limit the discussion to oblate axisymmetric Stäckel potentials which are associated with prolate spheroidal coordinates (λ, ϕ, ν) . A specific prolate spheroidal coordinate system is defined by two constants (a, c) . These coordinates are related to cylindrical polar coordinates (R, ϕ, z) by

$$\frac{R^2}{\tau - a^2} + \frac{z^2}{\tau - c^2} = 1, \quad (2.1)$$

where λ and ν are the roots of τ such that $c^2 \leq \nu \leq a^2 \leq \lambda$. Surfaces of constant λ are prolate spheroids and surfaces of constant ν are two-sheeted hyperboloids of revolution that intersect the spheroids orthogonally. Given λ and ν we can find expressions for R and z as a function of (λ, ν) ,

$$\begin{aligned} R^2 &= \frac{(\lambda - a^2)(\nu - a^2)}{c^2 - a^2}, \\ z^2 &= \frac{(\lambda - c^2)(\nu - c^2)}{a^2 - c^2}. \end{aligned} \quad (2.2)$$

To convert from the canonical coordinates, (R, z, p_R, p_z) , to the coordinates (τ, p_τ) , we introduce the generating function $S(\lambda, \nu, p_R, p_z)$ given by

$$S(\lambda, \nu, p_R, p_z) = p_R R(\lambda, \nu) + p_z z(\lambda, \nu). \quad (2.3)$$

We can then find $p_\tau = \partial S / \partial \tau$ as

$$\begin{aligned} p_\lambda &= \frac{1}{2} p_R \sqrt{\frac{\nu - a^2}{(c^2 - a^2)(\lambda - a^2)}} + \frac{1}{2} p_z \sqrt{\frac{\nu - c^2}{(a^2 - c^2)(\lambda - c^2)}}, \\ p_\nu &= \frac{1}{2} p_R \sqrt{\frac{\lambda - a^2}{(c^2 - a^2)(\nu - a^2)}} + \frac{1}{2} p_z \sqrt{\frac{\lambda - c^2}{(a^2 - c^2)(\nu - c^2)}}. \end{aligned} \quad (2.4)$$

Similarly, differentiating Eq. (2.2) we find $p_\lambda = P_\lambda^2 \dot{\lambda}$ and $p_\nu = P_\nu^2 \dot{\nu}$ where the dot denotes differentiation with respect to time and

$$\begin{aligned} P_\lambda^2 &= \frac{\lambda - \nu}{4(\lambda - a^2)(\lambda - c^2)}, \\ P_\nu^2 &= \frac{\nu - \lambda}{4(\nu - a^2)(\nu - c^2)}. \end{aligned} \quad (2.5)$$

Inversion of Eq. (2.4) allows us to write the Hamiltonian as

$$\begin{aligned} H &= \frac{1}{2}(p_R^2 + p_z^2) + \frac{L_z^2}{2R^2} + \Phi(R, z) \\ &= \frac{1}{2} \left(P_\lambda^2 \dot{\lambda}^2 + P_\nu^2 \dot{\nu}^2 + \frac{L_z^2}{R^2(\lambda, \nu)} \right) + \Phi(\lambda, \nu) \\ &= \frac{1}{2} \left(\frac{p_\lambda^2}{P_\lambda^2} + \frac{p_\nu^2}{P_\nu^2} + \frac{L_z^2}{R^2(\lambda, \nu)} \right) + \Phi(\lambda, \nu), \end{aligned} \quad (2.6)$$

where L_z is the z -component of the angular momentum. A potential, Φ_S , is of Stäckel form in a particular prolate spheroidal coordinate system if

$$\Phi_S = -\frac{f(\lambda) - f(\nu)}{\lambda - \nu}. \quad (2.7)$$

Φ_S is fully defined by a single function $f(\tau)$. A single function may be used as λ and ν take different ranges of values except at $\lambda = a^2, \nu = a^2$, where we require f to be continuous so the potential remains finite. As in all axisymmetric potentials, the energy, E , and z -component of the angular momentum, L_z are isolating integrals. In a Stäckel potential we are in the fortunate position of being able to find analytically a third isolating integral, I_3 . To do this we solve the Hamilton-Jacobi equation. We introduce the generating function, $W(\lambda, \nu, \mathbf{J})$ for the transformation between the prolate spheroidal coordinates and the as yet unknown actions and corresponding angles. With this generating function we write $p_\tau = \partial W / \partial \tau$ and set the Hamiltonian at fixed $\mathbf{J}$ to the energy E . We make the Ansatz that $W = \sum_\tau W_\tau(\tau)$ and find that

$$\begin{aligned} \lambda E + f(\lambda) - 2(\lambda - a^2)(\lambda - c^2) \left(\frac{\partial W_\lambda}{\partial \lambda} \right)^2 + \left(\frac{c^2 - a^2}{\lambda - a^2} \right) \frac{L_z^2}{2} \\ = \nu E + f(\nu) - 2(\nu - a^2)(\nu - c^2) \left(\frac{\partial W_\nu}{\partial \nu} \right)^2 + \left(\frac{c^2 - a^2}{\nu - a^2} \right) \frac{L_z^2}{2}. \end{aligned} \quad (2.8)$$

We introduce the separation constant

$$K = c^2 E - \frac{L_z^2}{2} + I_3, \quad (2.9)$$

where I_3 is the third integral. We now have expressions for the momenta as a function of τ and the classical integrals i.e.

$$2(\tau - a^2)(\tau - c^2)p_\tau^2 = (\tau - c^2)E - \left(\frac{\tau - c^2}{\tau - a^2} \right) \frac{L_z^2}{2} - I_3 + f(\tau). \quad (2.10)$$

This expression allows us to find the third isolating integral as

$$I_3 = (\tau - c^2)E - \left(\frac{\tau - c^2}{\tau - a^2} \right) \frac{L_z^2}{2} + f(\tau) - 2(\tau - a^2)(\tau - c^2)p_\tau^2. \quad (2.11)$$

We define the action variables, J_λ and J_ν , as

$$J_\tau = \frac{1}{2\pi} \oint p_\tau d\tau, \quad (2.12)$$

where the integration is over all values of τ for which $p_\tau^2 \geq 0$. As $p_\tau = p_\tau(\tau, E, I_2, I_3)$, the actions are solely functions of the isolating integrals and thus constants of the motion. The third action, J_ϕ , is simply L_z . Therefore, given a Cartesian phase-space point $(\mathbf{x}_0, \mathbf{v}_0)$, we can find the three isolating integrals,

$\mathbf{I} = (E, L_z, I_3)$, using the coordinate transformation and perform the 1D integrals in Eq. (2.12) to find the actions.

The actions give an absolute measure of the extent of the oscillations of the orbit in each of the coordinates. At large radii, the prolate spheroidal coordinate system becomes spherical such that $\lambda \approx R^2 + z^2$. Therefore, we can think of J_λ as a measure of the radial oscillations. The ν coordinate increases as we move away from the $z = 0$ plane, so we may think of J_ν as a measure of the vertical oscillations.

The corresponding angle coordinates, θ_τ and θ_ϕ , are calculated using the generating function, $W(\lambda, \phi, \nu, J_\lambda, L_z, J_\nu)$ for the canonical transformation from $(\lambda, \phi, \nu, p_\lambda, L_z, p_\nu)$ to $(\theta_\lambda, \theta_\phi, \theta_\nu, J_\lambda, L_z, J_\nu)$. The angles are found by differentiating the generating function with respect to the respective action such that

$$\theta_\tau = \frac{\partial W}{\partial J_\tau} \text{ for } \tau = \{\lambda, \nu\}; \quad \theta_\phi = \frac{\partial W}{\partial L_z}. \quad (2.13)$$

A full list of formulae, as well as a discussion of how to perform the quadratures numerically, is given in Appendix A. In this appendix, we also give expressions for the frequencies, $\boldsymbol{\Omega} = \partial H / \partial \mathbf{J}$, which will be useful in later chapters.

2.3 Estimating Actions in a Fitted Stäckel Potential

We have seen how to find actions in a Stäckel potential. We now go on to show how we can use these insights to estimate actions in a general potential. Given the ease with which we can calculate actions and angles in a Stäckel potential, it seems sensible to investigate how well a Stäckel potential can fit a Galaxy model so that we may estimate the actions and angles as those calculated in this best-fitting potential. It has been known for some time that Stäckel potentials do not give a good fit to the potential of the Galaxy globally due to the rigid conditions they must fulfil. Dejonghe and de Zeeuw (1988) outline a method for fitting a general axisymmetric potential with a Stäckel potential, which can be applied both globally and locally. These authors produced global fits for the Bahcall-Schmidt-Soneira Galaxy model (Bahcall et al. 1982) with errors in the potential nowhere exceeding 3 %, and Jasevicius (1994) carried out a similar analysis on a broader range of Milky Way potential models with similar results. As expected, the fits are worst in the central 0.5 kpc of the Galaxy. De Bruyne et al. (2000) sought to fit axisymmetric potentials locally using a set of Stäckel potentials in order to calculate the third isolating integral, I_3 . When applied to a Miyamoto-Nagai potential, I_3 was found to vary by approximately 10 % along an orbit. Here we follow the method presented by Dejonghe and de Zeeuw (1988).

Suppose we have an axisymmetric potential $\Phi(R, z)$ that we wish to fit by a Stäckel potential Φ_{fit} . We begin by choosing a prolate spheroidal coordinate system by specifying (a, c) . The coordinate system is fully specified by the combination $(a^2 - c^2)$, so we are free to set $c^2 = 1$, which reduces numerical difficulties. We determine a by using a property of an axisymmetric Stäckel potential (de Zeeuw 1984). It follows from Eq. (2.7) that for a Stäckel potential, Φ_S ,

$$\frac{\partial^2}{\partial \lambda \partial \nu} [(\lambda - \nu) \Phi_S] = 0. \quad (2.14)$$

Therefore, for a general potential Φ we use this equation as a definition for the coordinate system. We find this gives an estimate for a at a point (R, z)

$$a^2 - c^2 = R^2 - z^2 - \left[3z \frac{\partial \Phi}{\partial R} - 3R \frac{\partial \Phi}{\partial z} + Rz \left(\frac{\partial^2 \Phi}{\partial R^2} - \frac{\partial^2 \Phi}{\partial z^2} \right) \right] \bigg/ \frac{\partial^2 \Phi}{\partial R \partial z}. \quad (2.15)$$

Later, we calculate a sufficiently accurate value of a by evaluating this expression at multiple positions along an orbit and averaging.¹ With this choice of a we transform $\Phi(R, z)$ to $\Phi(\lambda, \nu)$ and specify the fitting region: $\lambda_- \leq \lambda \leq \lambda_+$, $\nu_- \leq \nu \leq \nu_+$. A global fit corresponds to $\nu_- = c^2$, $\nu_+ = \lambda_- = a^2$, $\lambda_+ = \infty$. We define the *auxiliary function*

$$\chi(\lambda, \nu) \equiv -(\lambda - \nu) \Phi(\lambda, \nu). \quad (2.16)$$

If the potential Φ is of Stäckel form, this auxiliary function is simply $\chi(\lambda, \nu) = f(\lambda) - f(\nu)$. We seek the function f that makes Φ_{fit} most like Φ by minimising the square difference of the potential auxiliary function and the fitting potential auxiliary function, χ_{fit} , over the fit region. Therefore, we minimise the functional

$$F[f] = \int_{\lambda_-}^{\lambda_+} d\lambda \int_{\nu_-}^{\nu_+} d\nu \Lambda(\lambda) N(\nu) (\chi(\lambda, \nu) - f(\lambda) + f(\nu))^2, \quad (2.17)$$

where $\Lambda(\lambda)$ and $N(\nu)$ are weighting functions allowing us to acquire a better fit in certain areas. These functions must be finite when integrated over the fitting region. We choose the normalised weighting functions

$$\Lambda(\lambda) = 4\lambda^{-5}(\lambda_-^{-4} - \lambda_+^{-4})^{-1}, \quad N(\nu) = (\nu_+ - \nu_-)^{-1}. \quad (2.18)$$

This choice of weighting functions gives preferential weight to smaller values of λ where the potential is harder to fit. Analytic minimisation of the functional F results in a best fit function

$$f(\lambda) = \bar{\chi}(\lambda) - \frac{1}{2} \bar{\bar{\chi}}, \quad f(\nu) = -\bar{\chi}(\nu) + \frac{1}{2} \bar{\bar{\chi}}, \quad (2.19)$$

¹For an oblate Stäckel potential, this expression gives the exact $a^2 - c^2$ at every point. However, for a general oblate potential, this expression can give negative values such that $a^2 < 0$. This implies that the best-fitting coordinate system is oblate (see Sect. 2.3.5). However, this only seems to occur when the equipotentials are near spherical (e.g. at high angular momentum in McMillan potential). For the logarithmic potential explored later this formula selects prolate coordinate systems for $q < 1$ and oblate for $q > 1$.

where

$$\begin{aligned}
 \bar{\chi}(\lambda) &= \int_{\nu_-}^{\nu_+} d\nu \chi(\lambda, \nu) N(\nu), \\
 \bar{\chi}(\nu) &= \int_{\lambda_-}^{\lambda_+} d\lambda \chi(\lambda, \nu) \Lambda(\lambda), \\
 \bar{\bar{\chi}} &= \int_{\nu_-}^{\nu_+} \int_{\lambda_-}^{\lambda_+} d\lambda d\nu \chi(\lambda, \nu) \Lambda(\lambda) N(\nu).
 \end{aligned} \tag{2.20}$$

The derivation of these equations is given in Appendix B. The quality of the fit we have achieved is then measured by $F[f]$.

2.3.1 Procedure

Combining the above two sections, we can estimate the actions and angles of a phase point $(\mathbf{x}, \mathbf{v})$ by first fitting a Stäckel potential to the given potential over the region the orbit probes, and then calculating the angle-action variables in this fitted potential. Therefore, given a point $(\mathbf{x}, \mathbf{v})$, we follow this procedure:

1. We begin by calculating the z -component of the angular momentum, L_z , and the energy, E , in the ‘true’ potential.
2. We then integrate the orbit in the ‘true’ potential. We use the initial time-steps of the orbit integration to find the best-fitting coordinate system: at several points along the orbit we evaluate Eq. (2.15) and average to find a sufficiently accurate value for a . With the coordinate system found, we continue integrating to find the edges of the orbit, λ_+ , λ_- and ν_+ , that define the fitting region. The edges of the orbit are given approximately by the points where $\dot{\tau} = 0$ in the best-fitting prolate spheroidal coordinate system. The minimum and maximum τ edges are distinguished by inspecting the sign of $\ddot{\tau}$. For all realistic potentials every orbit crosses the $z = 0$ plane so we set $\nu_- = c^2$.
3. We can now find a best-fitting Stäckel potential over this region. Using Eq. (2.19) we tabulate $f(\lambda)$ and $f(\nu)$ for 40 points in (λ_-, λ_+) and (ν_-, ν_+) respectively so that we may interpolate these smooth functions. Any call outside the ranges is calculated fully using Eq. (2.19) with a full re-computation of $\bar{\chi}(\tau)$.
4. With the best fit potential now calculated, we find I_3 using Eq. (2.11) for three points on the boundary of the orbit (on the minimum λ edge, the maximum λ edge and the maximum ν edge) and take an average. We have already found these three points when determining the edges of the orbit, so this choice involves minimum additional computational effort and provides a fair estimate for I_3 over a large region of the orbit. However, this choice of I_3 can lead to the initial phase-space point $(\mathbf{x}, \mathbf{v})$ being forbidden. Therefore, with this choice of I_3 , we check whether $p^2(\nu) > 0$ and $p^2(\lambda) > 0$ for the initial phase-space point using Eq. (2.11), and,

if not, then we calculate I_3 from Eq. (2.10) using only the initial phase-space point. This procedure² reduces the numerical noise around the turning points, particularly in R .

5. With the three isolating integrals calculated, we are in a position to estimate the actions and angles using the method outlined in Sect. 2.2. The limits of the orbit are redetermined by finding from Eq. (2.10) the points where $p_\tau^2 = 0$ using Brent's method, and are not given by $\tau_\pm$. We do this because using $\tau_\pm$ may result in including points in the integration where $p_\tau^2 < 0$.

Table 2.3 quantifies the efficiency of this procedure.

2.3.2 Discussion

Kent and de Zeeuw (1991) propose several methods for estimating I_3 by assuming the potential is close to separable. For general disc orbits in realistic Galactic potentials, they found that the most accurate of the techniques was the 'least-squares method'. This method sought to minimise

$$\int dt \left[(\tau(t) - c^2)E - \left(\frac{\tau(t) - c^2}{\tau(t) - a^2} \right) \frac{L_z^2}{2} + f(\tau(t)) - 2(\tau(t) - a^2)(\tau(t) - c^2)p_\tau^2(t) - I_3 \right]^2$$

with respect to the coordinate system and I_3 , where the integral is in reality a sum over phase-space points from an orbit integration. In a general potential, we require some approximation for $f(\tau)$. The method of Kent and de Zeeuw (1991) does not require an explicit fit of a Stäckel potential to the true potential. The authors chose to approximate $f(\lambda) \approx -(\lambda - c^2)\Phi(R(\lambda, c^2), 0)$. In Sect. 2.6 and Chap. 4 we employ a similar approach to estimating $f(\tau)$ from a general potential. As we have already performed an orbit integration as part of our procedure, we could choose to estimate the best-fitting coordinate system and I_3 in the manner presented by Kent and de Zeeuw. With a best-fitting coordinate system chosen we could then find the best-fitting Stäckel potential hopefully to further refine our estimate of I_3 . Then with estimates of both $f(\lambda)$, $f(\nu)$, I_3 and the coordinate system we could use Eq. (2.12) to find the actions.

The calculation of the actions is sensitive to the value of the momenta throughout the orbit. Therefore, it is important that we find a good approximation to both the coordinate system and $f(\tau)$. Later we will show that, from the proposed methods, the estimate of $f(\tau)$ that produces the smallest spread in the action estimates is obtained by fitting a Stäckel potential as done here. The method of Kent and de Zeeuw (1991), whilst appropriate for finding a good choice of coordinate system and estimate for I_3 , is probably not optimal for action estimation.

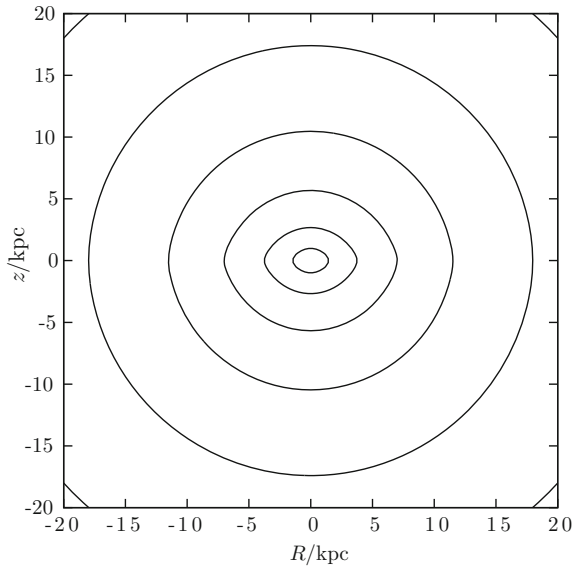
²Later we will adopt a slightly different procedure (see Chap. 6 and Appendix G for details).

2.3.3 Application

We now investigate how successful the above routine is for estimating the angle-action variables in a general axisymmetric potential. To demonstrate the applicability of the method to data, we choose a realistic Milky Way potential from McMillan (2011a). This potential consists of two exponential discs for the thick and thin discs of the Galaxy and two spheroids for the bulge and dark matter halo. We select the ‘best’ model from this paper, which had parameters for the components of the potential set by current constraints. The equipotential contours for this model are plotted in Fig. 2.1. It is clear that, as we move out from the centre, the contours become more circular so we anticipate that they are better fit by surfaces of constant λ and ν . Therefore, we expect more accurate estimates of the angle-action variables for orbits at larger radii. Also, orbits that probe a large range of R and/or z should have less accurate angle-action variable estimates as these orbits probe a large range of curvature of the equipotential contours. Therefore, we expect the method to work best for small J_λ and J_ν but large L_z .

de Zeeuw et al. (1986) show that for the corresponding density of a Stäckel potential to be everywhere non-negative, the density cannot fall off faster than r^{-4} , where r is the spherical radius. The best potential from McMillan (2011a) has a flat rotation curve, and an Navarro-Frenk-White halo (Navarro et al. 1996). Therefore, the density distribution towards the centre goes approximately as r^{-2} and falls to r^{-3} in the outskirts. We, therefore, expect the best-fitting Stäckel potentials to be suitable. However, as we are only locally fitting Stäckel potentials, we can produce

Fig. 2.1 Contours of $\ln(\Phi(R, z)/\Phi(0, 0))$ for McMillan’s best-fitting Milky Way potential. The contours increase from the centre in equally space units of 0.15 with the central contour at -0.15



unphysical global Stäckel potentials i.e. $\rho(\mathbf{x}) < 0$ for some $\mathbf{x}$, that are still suitable locally.

We assess the validity of the method by comparing the results with the ‘exact’ angle-action variables calculated using the ‘torus machine’ (McMillan and Binney 2008). Orbital tori are three-dimensional surfaces characterised by the three actions $\mathbf{J} = (J_R, L_z, J_z)$ obtained as the images of analytic tori under a canonical transformation. The strength of the torus machine lies in constructing a torus given a set of actions, $\mathbf{J}$, such that the phase-space coordinates, $(\mathbf{x}, \mathbf{v})$, may be obtained as functions of the angles, $\boldsymbol{\theta}$, over the surface of the torus. Therefore, a simple test for the Stäckel potential fitting procedure is first to produce a list of phase space coordinates with fixed actions but randomly chosen angles using the torus machine. The success of the method is then measured by how accurately the angle-action variables can be reproduced. We note that the canonical transformation produced by the torus machine maps J_R into J_λ and J_z into J_ν . From now on, we will use the more intuitive notation for the actions, J_R and J_z , and similarly for the angles, θ_R and θ_z .

The errors in the actions of a given torus from the torus machine may be estimated from the residuals of the Hamiltonian over the surface of the torus. The error in the Hamiltonian, ΔH , is related directly to the error in the actions by

$$\Delta H = \frac{\partial H}{\partial \mathbf{J}} \Delta \mathbf{J} = \boldsymbol{\Omega} \cdot \Delta \mathbf{J} \quad (2.21)$$

where we find the frequency $\boldsymbol{\Omega}$ directly from the torus machine. Assuming the errors in J_R and J_z are approximately equal and uncorrelated, the error in the actions may be estimated as

$$\Delta J \approx \frac{\Delta H}{\sqrt{\Omega_R^2 + \Omega_z^2}}. \quad (2.22)$$

The true angles of an orbit in a potential increase linearly with time. The errors in the torus angles are estimated by the residuals of the angles away from this expected straight line. Clearly, we require the angle and action errors from the torus machine to be smaller than the errors from the Stäckel-fitting procedure in order to state anything meaningful about the systematic errors from our method.

2.3.3.1 Single Torus

Here we discuss the results of applying the procedure to 10,000 randomly generated points from the torus³ $\mathbf{J} = (J_R, L_z, J_z) = (0.078, 1.9, 0.097) \text{ kpc}^2 \text{ Myr}^{-1}$. This torus was chosen to be representative of the actions of a disc star in the solar neighbourhood. For this torus, the errors in the actions and angles are $\Delta J/J = 0.01\%$ and $(\Delta \theta_R, \Delta \theta_\phi, \Delta \theta_z) = (1.0, 0.2, 1.0) \times 10^{-5} \text{ rad}$. The orbit in the (R, z) plane is

³Throughout this chapter, the actions are stated in units of $\text{kpc}^2 \text{ Myr}^{-1} = 977.8 \text{ kpc km s}^{-1}$.

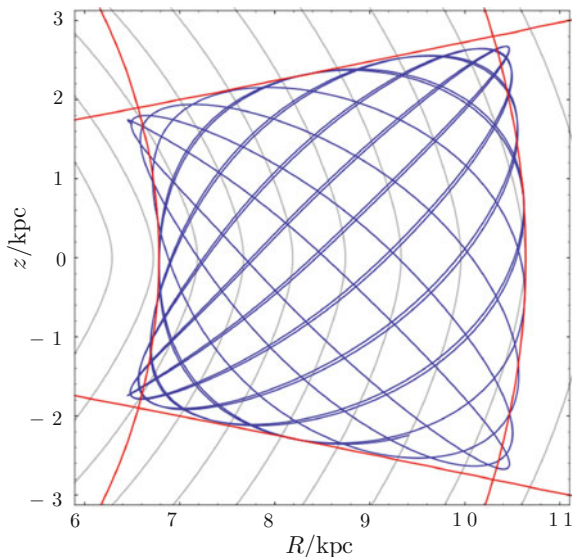


Fig. 2.2 Fit region for a single orbit—the blue line shows the orbit with actions $\mathbf{J} = (J_R, L_z, J_z) = (0.078, 1.9, 0.097) \text{ kpc}^2 \text{ Myr}^{-1}$. The black lines are the equipotential contours of McMillan’s best-fitting Milky Way potential. The red lines show the lines of constant λ and ν that define the region over which the potential is fitted

shown in Fig. 2.2. This orbit has apses at $R \approx (6.5, 10.5) \text{ kpc}$ and $z_{\text{max}} \approx 2.8 \text{ kpc}$. Also shown in the figure are the curves defining the fit region and equipotential contours for McMillan’s best-fitting potential.

The residuals in the fitted potential over the fitting region defined in Fig. 2.2 are plotted in Fig. 2.3. Everywhere within the fitting region, the error in the potential is less than 0.2 % of the maximum difference in the potential across the fitting region. We note here that a good fit for the potential does not necessarily correlate with an accurate calculation of the actions. Small changes in the potential can cause large changes in the motion of a particle, so, whilst a good fit for the potential is necessary, we don’t expect the errors in the actions to be of similar order.

The 10,000 phase-space points are shown in scatter plots of (R, θ_R) and (z, θ_z) in Fig. 2.4. We can see that R and z are periodic in the angles. We define the zero-point of θ_R such that the radial periapsis and apoapsis correspond to $\theta_R = 0$ and $\theta_R = \pi$ respectively. θ_z is defined such that $z = 0$ corresponds to $\theta_z = 0, \pi$ and $z = \pm z_{\text{max}}$ corresponds to $\theta_z = \pi/2, 3\pi/2$. The zero-point of θ_ϕ is defined such that $\theta_\phi = \phi$ at periapsis. The spread of the z coordinates of the points at a given angle is much larger than the spread in the R coordinates.

When the Stäckel fitting method is applied to this set of phase-space points, we find that the root-mean-square (RMS) deviations of the actions, ΔJ , are given by $\Delta J_R/J_R \approx 4.9\%$ and $\Delta J_z/J_z \approx 4.2\%$. We also find that there is a very tight anticorrelation between ΔJ_R and ΔJ_z . All phase-space points have the same energy

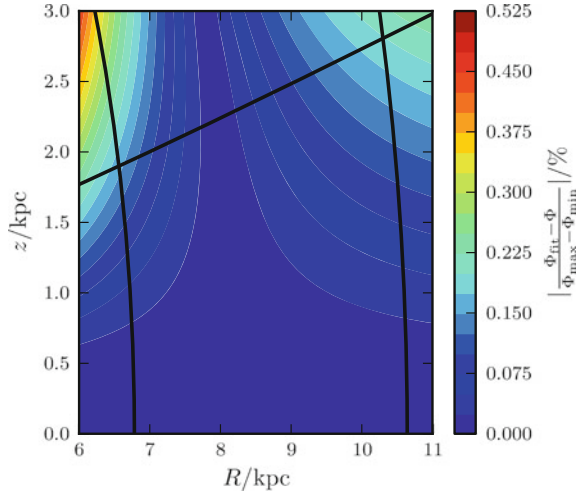


Fig. 2.3 Filled contour plot of the percentage difference between the best-fitting Stäckel potential and McMillan's best-fitting Milky Way potential. $\Phi_{\min}$ and $\Phi_{\max}$ give the values of the potential on the minimum and maximum λ edges respectively. Also plotted in black are the curves of constant λ and ν which define the region over which the potential is fitted. This is the fit region corresponding to the orbit shown in Fig. 2.2, with actions $\mathbf{J} = (J_R, L_z, J_z) = (0.078, 1.9, 0.097) \text{ kpc}^2 \text{ Myr}^{-1}$. We see that, within the fitting region, the difference between the fitted potential and the potential we are attempting to fit is less than 0.2 % of the maximum potential difference across the fitting region

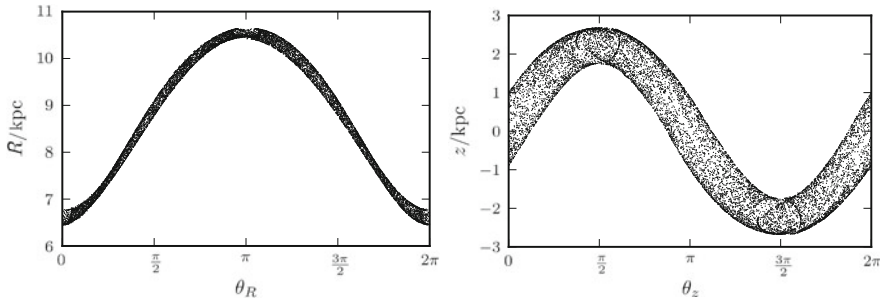


Fig. 2.4 Scatter plot of R against θ_R and z against θ_z for 10,000 randomly selected phase-space points from the torus detailed in Sect. 2.3.3.1

as we are using the potential that was used to integrate the orbit to find the energy. Therefore, all the points lie along the intersection of the surface of constant energy with the (J_R, J_z) plane. If we overestimate J_R , we must underestimate J_z in order to have the correct energy. The RMS deviations in the angles for the 10,000 phase-space points are $(\Delta\theta_R, \Delta\theta_\phi, \Delta\theta_z) = (4.1, 1.1, 5.1) \times 10^{-2} \text{ rad}$.

Errors as a function of angle: It is informative to investigate how the errors in the derived actions and angles vary with true angle around the torus. The derived actions are approximately independent of the true angles as they depend only on

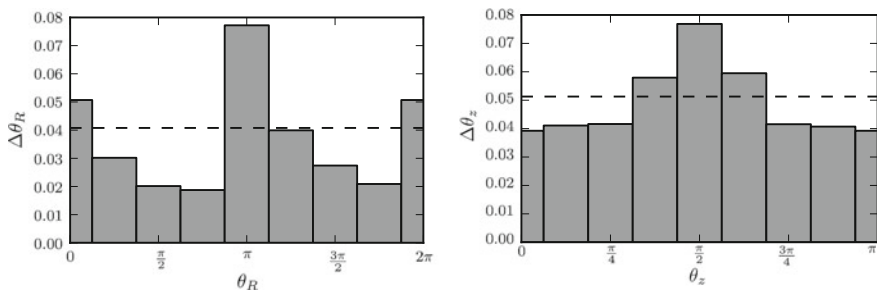


Fig. 2.5 RMS error in the angles binned as a function of angle for the torus in Sect. 2.3.3.1. The *dashed line* shows the total RMS error from all the points on the torus. For the bottom panel, we have taken advantage of the symmetry in the $z = 0$ plane and mapped the $\theta_z = (0, 2\pi)$ interval onto $\theta_z = (0, \pi)$ such that $\theta_z = \pi/2$ corresponds to $\pm z_{\max}$ etc. The largest error occurs at the apsides for both cases

the path of the orbit, which is determined by the fitted potential and not the initial point on the orbit. Any small variation is due to the choice of prolate spheroidal coordinate system and variations in the fitted potential. However, we find that the error in the derived angle varies with true angle. In Fig. 2.5 we plot the RMS errors in the angles binned as a function of true angle for both θ_R and θ_z . Maximum errors occur at the turning points in the (R, θ_R) and (z, θ_z) plots shown in Fig. 2.4. For the radial and vertical angle, the largest error occurs at apoapsis. In a Stäckel potential, the momenta, p_τ , depend on τ and the isolating integrals, which, once determined, are taken to be constant. Therefore, at a given location in the orbit, the angle is solely a function of the position coordinates and the velocity information is essentially ignored. Around turning points in the orbit, the velocity coordinates contain the majority of the information, whilst the position coordinates are changing very slowly. Therefore at turning points the errors in the angles are large as the angle coordinates are estimated using this reduced phase-space information. In general, the errors in θ_z are larger than the errors in θ_R .

2.3.3.2 Multiple Tori

We have seen that the method gives reasonable estimates for the actions for a particular torus, but, in order to use the method with confidence, we need to see how the errors depend on the torus. Here we repeat the above procedure for a range of different tori that probe the different regions of the potential. We work with two groups of tori: those with low actions and torus machine errors less than $\Delta J/J = 0.01\%$ and $(\Delta\theta_R, \Delta\theta_\phi, \Delta\theta_z) = (27.0, 5.1, 990) \times 10^{-6}$ rad and those with high actions and torus machine errors less than $\Delta J/J = 1\%$ and $(\Delta\theta_R, \Delta\theta_\phi, \Delta\theta_z) = (2.0, 1.7, 1.7) \times 10^{-2}$ rad. The low-action group consists of 100 tori with actions $J_R = (0.001, 0.005, 0.01, 0.05, 0.1) \text{ kpc}^2 \text{ Myr}^{-1}$, $J_z = (0.001, 0.005, 0.01, 0.05, 0.1) \text{ kpc}^2 \text{ Myr}^{-1}$ and $L_z = (1.0, 2.0, 3.0, 4.0) \text{ kpc}^2 \text{ Myr}^{-1}$.

These tori probe the region $3 \text{ kpc} < R < 22 \text{ kpc}$, $|z| < 5 \text{ kpc}$ and are chosen to be representative of disc-type tori. The high-action group consists of 36 tori with actions $J_R = (0.5, 1.0, 5.0) \text{ kpc}^2 \text{ Myr}^{-1}$, $J_z = (0.5, 1.0, 5.0) \text{ kpc}^2 \text{ Myr}^{-1}$ and $L_z = (1.0, 2.0, 3.0, 4.0) \text{ kpc}^2 \text{ Myr}^{-1}$. These tori probe the region $2 \text{ kpc} < R < 120 \text{ kpc}$, $|z| < 100 \text{ kpc}$. We include the second group to demonstrate that the method can deal with orbits that deviate very far from the plane and probe a very large region of the potential. We would like to be able to apply the method to halo stars and tidal streams, so it is important to understand the errors for these high-action tori.

Actions: As mentioned previously, we expect the errors in J_R and J_z will be large when $J_R/|L_z|$ and/or $J_z/|L_z|$ are large. In this regime, the orbit probes a large central region of the potential so we anticipate the potential fit will be poorer. In Fig. 2.6, the RMS deviations in the actions for the complete orbit sample are plotted against the combination of the actions $(J_R + J_z)/|L_z|$. We can see that, as anticipated, the absolute errors correlate with this action combination. In fact, the correlation is much tighter than the individual correlations with $J_R/|L_z|$ and $J_z/|L_z|$, so the errors in the method are dependent on the sum of the actions ($J_R + J_z$). It is this measure that tells us how much an orbit strays from a circular orbit and thus how much of the potential it explores.

We also note from Fig. 2.6, that, at a given value of $(J_R + J_z)/|L_z|$, the errors in J_R and J_z are of similar magnitudes. As explained above, the errors in J_R and J_z compensate for each other to recover the correct energy. In Fig. 2.7, we plot this correlation between the RMS errors in J_R and J_z . A consequence of this tight correlation is that when one action is much greater than the other, the relative error in the smaller action will be much greater than the relative error in the larger action. However, it is worth noting that the absolute error is far more important than the relative error. Given a distribution function for a steady-state galaxy, $f(\mathbf{J})$, the absolute error in f is given by

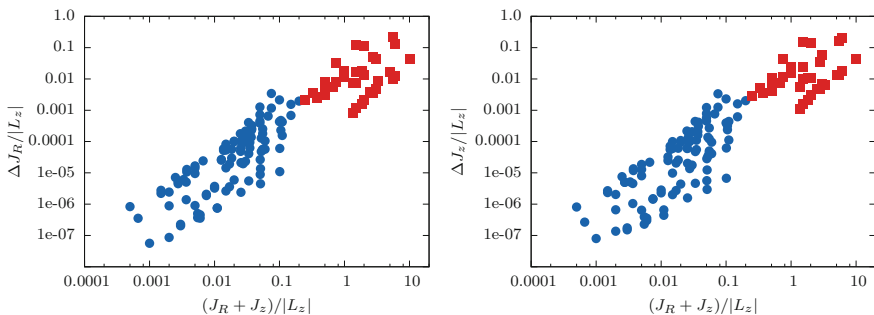
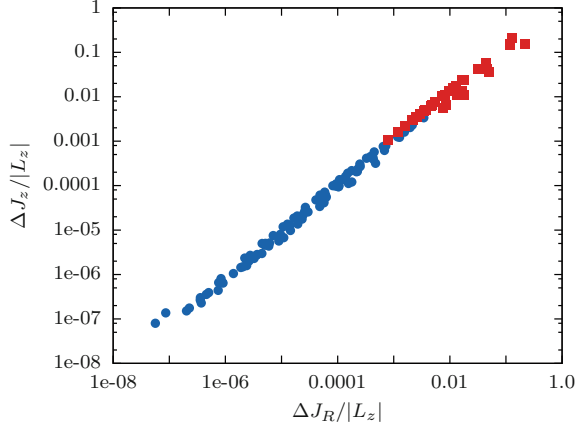


Fig. 2.6 Absolute RMS deviations of the actions for the sample of tori detailed in Sect. 2.3.3.2. The blue circles are data for which the relative errors in the actions from the torus machine are less than 0.01 % and the red squares are those with errors less than 1 %. The errors correlate loosely with both J_R and J_z separately, but there is a much tighter correlation between the errors and $(J_R + J_z)$

Fig. 2.7 RMS deviations in J_R and J_z for the sample of tori detailed in Sect. 2.3.3.2. The *blue circles* are data for which the relative errors in the actions from the torus machine are less than 0.01 % and the *red squares* are those with errors less than 1 %



$$(\Delta f)^2 = \sum_{i,j} \frac{\partial f}{\partial J_i} \frac{\partial f}{\partial J_j} \text{cov}(J_i, J_j), \quad (2.23)$$

where $\text{cov}(X, Y)$ is the covariance between variables X and Y . In the case of uncorrelated errors between the actions, this simply becomes

$$(\Delta f)^2 = \sum_i \left(\frac{\partial f}{\partial J_i} \Delta J_i \right)^2. \quad (2.24)$$

The distribution function for the Milky Way is approximately exponential in the actions (Binney 2010 and Chap. 8):

$$f(\mathbf{J}) \sim \prod_i e^{a_i J_i}, \quad (2.25)$$

where a_i is independent of the action J_i . Therefore, the absolute error in f is given by

$$(\Delta f)^2 = \sum_i (f a_i \Delta J_i)^2. \quad (2.26)$$

Similarly, the relative error in the distribution function is given by

$$\left(\frac{\Delta f}{f} \right)^2 = \sum_i \left(\frac{\partial \ln f}{\partial \ln J_i} \frac{\Delta J_i}{J_i} \right)^2 = \sum_i (a_i \Delta J_i)^2. \quad (2.27)$$

Both the absolute and relative error in the distribution function are determined by the absolute errors in the actions, so we need not be overly concerned that the relative error in one action is much larger than the relative error in another.

From the relationship illustrated in Fig. 2.6, we can estimate the error in a given estimate of J_R and J_z . Performing a linear fit to both sets of data points independently, we find that, for $(J_R + J_z)/|L_z| \lesssim 10$, a good fit for the RMS errors in both J_R and J_z is given by

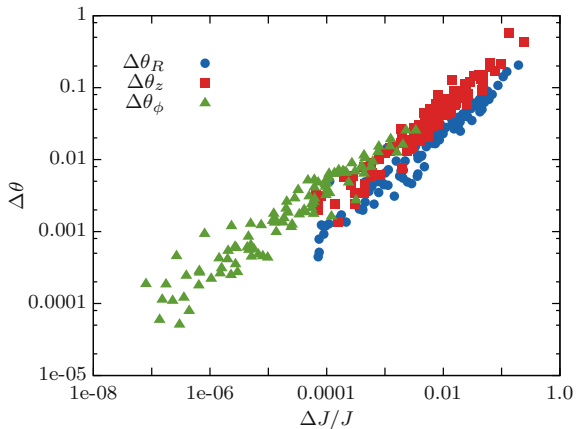
$$\Delta J \approx 0.01 \frac{(J_R + J_z)^{\frac{3}{2}}}{|L_z|^{\frac{1}{2}}}. \quad (2.28)$$

The errors in the actions have a weak dependence on L_z as orbits with higher L_z explore regions of the potential that are more spherical and hence easier to fit with a Stäckel potential (see Fig. 2.1).

Angles: We present the RMS deviations in the angles for the 100 tori in the low-action group in Fig. 2.8. In general, the relative errors in the angles are larger than the errors in the actions. The calculation of an action involves only a single integral, whereas the corresponding angle calculation involves nine integrals. Each integral folds in more error from the fitting method, so we expect the relative errors in the angle variables to be significantly larger than the errors in the actions. From Fig. 2.8, we see that the errors in the angles correlate with the relative error in the actions. The error in θ_ϕ has been plotted against $\Delta J_R/L_z$, whilst the other two angles have been plotted against the relative error in their respective action. As the errors in J_R and J_z are approximately equal (Fig. 2.7), the three sets of points are essentially θ_i against $\Delta J/J_i$. As with the errors in the actions, we may estimate the error in a given calculation of the angles by fitting the data in Fig. 2.8. We find the errors are approximately given by

$$\Delta\theta_R \approx \left(\frac{\Delta J}{J_R}\right)^{0.75}, \quad \Delta\theta_z \approx \left(\frac{\Delta J}{J_z}\right)^{0.75}, \quad \Delta\theta_\phi \approx \left(\frac{\Delta J}{L_z}\right)^{0.5}. \quad (2.29)$$

Fig. 2.8 RMS deviations in the angles for the 100 low-action tori as a function of the relative error in their respective action. For θ_ϕ , we have plotted the error against $\Delta J_R/L_z$



2.3.4 Geneva-Copenhagen Survey

Now that we have understood the systematic errors of the method, we apply it to real data. The Geneva-Copenhagen Survey (GCS) (Nordström et al. 2004) is a sample of 16682 nearby F and G stars, and is perhaps the best data set with full 6D phase-space information. It provides us with a platform to motivate a discussion of errors involved in a practical calculation. The GCS has been analysed by many authors and specifically looked at in angle-action space by Sellwood (2010) and McMillan (2011b). From the table produced by Holmberg et al. (2009), we select the 13,518 objects that have full 6D phase-space information. We correct the data for the solar velocity with respect to the local standard of rest as calculated by Schönrich et al. (2010) i.e. $(U, V, W)_\odot = (11.1, 12.24, 7.25) \text{ km s}^{-1}$.⁴ Using the ‘best’ model Galactic potential from McMillan (2011a) sets the solar radius as $R_0 = 8.29 \text{ kpc}$ and the velocity of the local standard of rest as $v_{\text{LSR}} = 239.1 \text{ km s}^{-1}$. The results of estimating the actions and angles using the Stäckel-fitting procedure are shown in Fig. 2.9. These are very similar to the equivalent plots from Sellwood (2010) and McMillan (2011b). We see that the plot of J_R against L_z has a markedly parabolic shape. The minimum of this parabola corresponds to the circular orbit at the solar radius. The shaded red region is inaccessible by stars that are at the solar position and have $J_z = 0$. We can understand this using the epicyclic approximation (Binney and Tremaine 2008): a star with a given L_z has some guiding-centre radius R_c . For $J_R = 0$, the star follows a circular orbit of radius R_c at an angular rate L_z/R_c^2 , such that $\phi = \theta_\phi = t L_z/R_c^2$. Non-zero J_R produces radial oscillations about this circular orbit, such that a star circulates the guiding centre of the $J_R = 0$ orbit. A star requires enough radial action to produce a radial oscillation large enough to pass through the solar neighbourhood, hence the parabolic shape. The angles θ_R and $\theta_\phi - \phi$ describe the location of the star with respect to the guiding centre. They are correlated as, whilst a star leads the guiding centre, $\theta_\phi < \phi$ and $\theta_R < \pi$, whilst, when trailing, $\theta_\phi > \phi$ and $\theta_R > \pi$. At the solar position $\phi = \pi$, hence the observed selection effect in Fig. 2.9. From the plot of θ_ϕ against θ_R , we see that the majority of stars are at $\theta_R = 0, \pi$ corresponding to the apses of their radial motion. However, there is still a lot of structure in between these extrema: the peak at $\theta_R/\pi \approx 0.54$ corresponds to the Hyades moving group.

2.3.4.1 Structure in the GCS

Since Dehnen (1998) investigated the kinematics of the solar neighbourhood using data from the *Hipparcos* satellite, it has been known that, when viewed in the (U, V) velocity plane, the local distribution of stars consists of a series of groups and clusters. These structures were classified by Famaey et al. (2005) and are all thought to have a dynamical origin (De Simone et al. 2004; Antoja et al. 2010). From the GCS sample, we can identify several of the larger structures from the peaks in the (U, V)

⁴We work in a right-handed Galactocentric Cartesian coordinate system with the positive x direction pointing towards the Galactic centre.

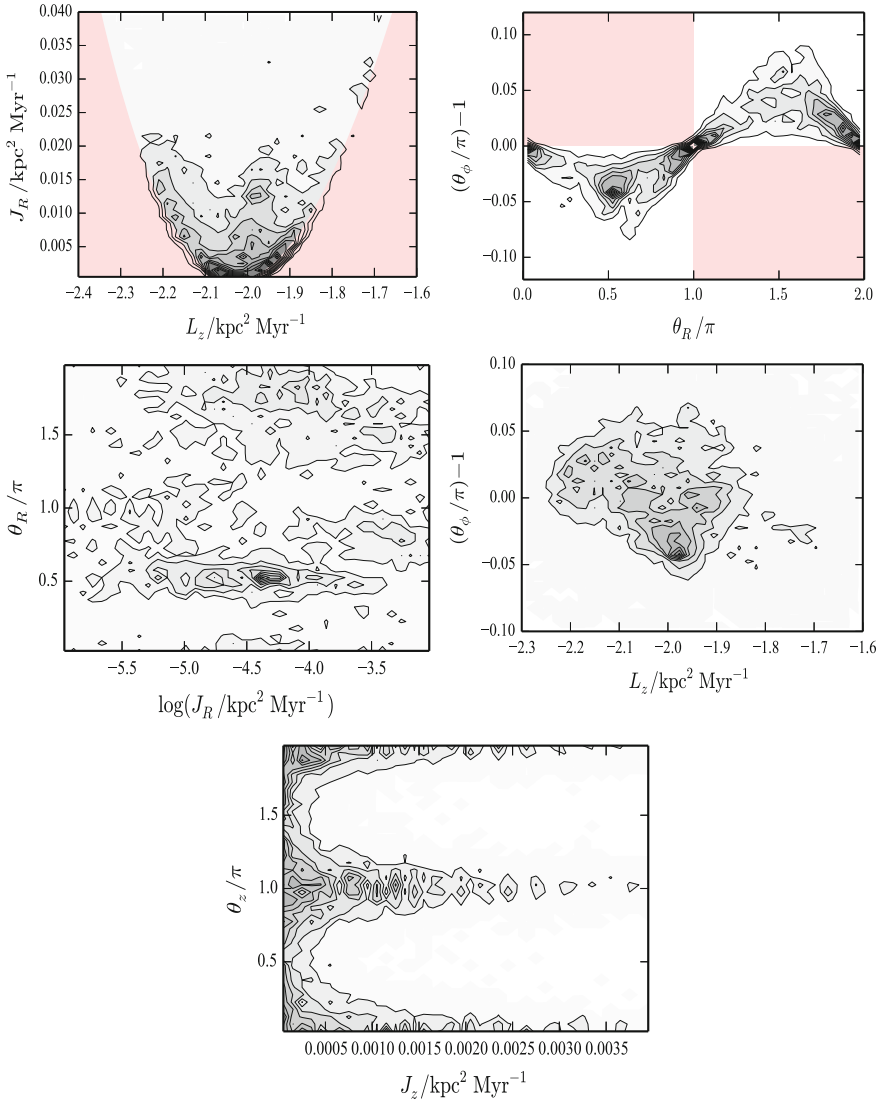
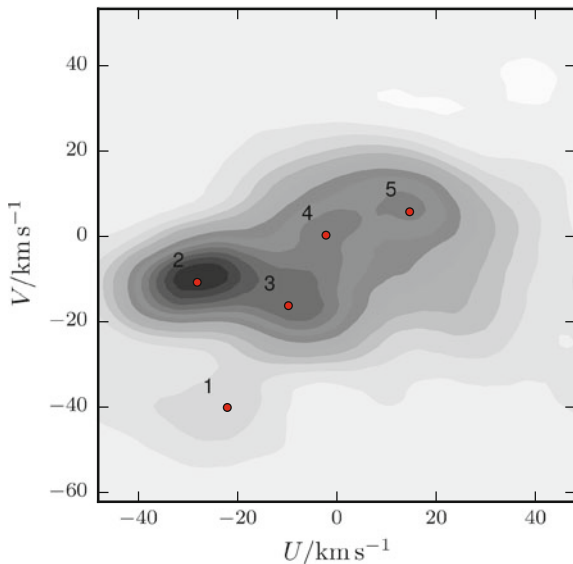


Fig. 2.9 Density contours for the actions and angles of 13,518 objects from the Geneva-Copenhagen Survey. The contours are linearly spaced. The *red shaded region* is inaccessible by stars that are at the solar position and have $J_z = 0$. The strong peak at $\theta_R/\pi \approx 0.54$ is due to the Hyades. We also see the Hercules stream at $\log J_R/\text{kpc}^2\text{Myr}^{-1} \approx -3.3$ and $\theta_R/\pi \approx 0.85$

distribution. This is done by binning the stars in U and V and then performing a wavelet transform. In Fig. 2.10, we show the results of the wavelet transform for structure on scales $\sim 12\text{ km s}^{-1}$. The peaks in this plot are identified with the known

Fig. 2.10 Density contours for the (U, V) plane of 13,518 objects from the Geneva-Copenhagen Survey smoothed to a scale of $\sim 12 \text{ km s}^{-1}$ via a stationary wavelet transform. The five major structures are identified by red points: 1 Hercules, 2 Hyades, 3 Pleiades, 4 Coma Berenice, 5 Sirius



groups, specifically the Hercules, Hyades, Pleiades, Coma Berenice and Sirius. As discussed by McMillan (2011b), for a star located at the solar position, each point in the (U, V) plane represents a unique point in the $(J_R, L_z, \theta_R, \theta_\phi)$ space. Lines of constant L_z and θ_ϕ form an approximately Cartesian grid in the (U, V) plane, whilst lines of constant J_R and θ_R form an approximately polar grid. Therefore, we can find a reduced set of angle-actions (excluding any vertical action and angle) for each of the peaks identified in Fig. 2.10. From inspecting the (U, V) plane, we can see that the distribution deviates from axisymmetric equilibrium, so estimating the actions and angles assuming an axisymmetric potential is clearly an oversimplification. However, this is a necessary first step to create a basis on which we can include non-axisymmetric perturbations.

We represent each group by a 6D phase-space point placed at the solar position with zero vertical velocity but U and V determined by the identification from Fig. 2.10. We then estimate the actions and angles corresponding to this point, again using McMillan’s ‘best’ potential. The results are shown in Table 2.1. This gives us an opportunity to discuss the various sources of errors in a realistic use of the method. There are two sources of error in this calculation—the systematic errors introduced by the Stäckel fitting method and the errors in the input coordinates. Holmberg et al. (2009) estimated the space-velocity errors for each star as 1.5 km s^{-1} , which when combined with the errors estimated in Schönrich et al. (2010) for the solar motion gives velocity errors of $(\Delta U, \Delta V) = (1.7, 1.6) \text{ km s}^{-1}$. The majority of this error arises from the uncertainty in the distances. We convolve each peak position by these errors to calculate 10,000 Gaussianly distributed points in (U, V) space about these peaks, and then estimate the errors in the output actions and angles by the RMS scatter in the resulting (J, θ) coordinates. These errors can then be compared and combined with the known systematic errors from Sect. 2.3.3.2 and are shown in Table 2.1.

Table 2.1 Velocities with respect to the local standard of rest and angle-action coordinates for density peaks of known structures in the Geneva-Copenhagen Survey

	$U / \text{km s}^{-1}$	$V / \text{km s}^{-1}$		
Hercules	−22.1	−40.1		
Hyades	−28.2	−10.8		
Pleiades	−9.8	−16.3		
Coma Berenice	−2.2	0.3		
Sirius	14.7	5.7		
	$J_R / \text{kpc}^2 \text{Myr}^{-1}$	$L_z / \text{kpc}^2 \text{Myr}^{-1}$	θ_R	$\theta_\phi - \pi$
Hercules	$(4.0 \pm 0.3^{0.006}_{0.3}) \times 10^{-2}$	-1.69 ± 0.01	$2.65 \pm 0.04^{0.008}_{0.036}$	$-(9.6 \pm 0.7^{0.05}_{0.68}) \times 10^{-2}$
Hyades	$(1.2 \pm 0.1^{0.001}_{0.1}) \times 10^{-2}$	-1.94 ± 0.01	$1.91 \pm 0.07^{0.005}_{0.07}$	$-0.118 \pm 0.007^{0.0001}_{0.0066}$
Pleiades	$(7.3 \pm 0.1^{0.005}_{0.1}) \times 10^{-3}$	-1.89 ± 0.01	$2.68 \pm 0.08^{0.004}_{0.077}$	$-(4.2 \pm 0.7^{0.006}_{0.7}) \times 10^{-2}$
Coma Berenice	$(1.7 \pm 1.4^{0.0002}_{1.4}) \times 10^{-4}$	-2.03 ± 0.01	$1.5 \pm 1.1^{0.0009}_{1.1}$	$-(9.3 \pm 6.4^{0.0008}_{6.4}) \times 10^{-3}$
Sirius	$(3.5 \pm 0.8^{0.001}_{0.8}) \times 10^{-3}$	-2.07 ± 0.01	$5.28 \pm 0.13^{0.003}_{0.13}$	$0.06 \pm 0.01^{0.00002}_{0.006}$

The errors in U and V are 1.7 km s^{-1} and 1.6 km s^{-1} respectively. The errors in the angles and actions are presented as a_c^b where a is the total error and b and c are the contributions from the systematic errors of the Stäckel fitting method and the errors in the space-velocities respectively, such that $a^2 = b^2 + c^2$. In all cases, the observational error dominates the systematic error

We find that, for all the coordinates, the error in the data dominates the systematic error introduced by the method. As noted by McMillan (2011b) the relationship between errors in $(\mathbf{x}, \mathbf{v})$ and $(\mathbf{J}, \boldsymbol{\theta})$ is non-trivial. Specifically, at very low radial actions, any error in the velocity can introduce a 2π error in the θ_R coordinate. We see this occurring for the Coma Berenice peak—the error in the θ_R coordinate is large as the peak is positioned very close to the origin of the (U, V) plane. We also see that the errors in J_R and θ_ϕ are of order one for Coma Berenice.

We note that we have not included any error for the size of the structures in phase-space, nor any error due to the choice of smoothing from the wavelet transform, nor any error for the assumption that all the stars are situated at the solar position, nor any error in the choice of gravitational potential. Investigating the error in the actions due to the range of viable potentials for the Milky Way is beyond the scope of this investigation. Even with this underestimated error, the errors from the data dominate the systematic errors. We conclude that, given the accuracy of the current data, the determination of the angle-action coordinates using the Stäckel-fitting method is not limited by the well-understood systematic errors.

Table 2.2 Deviations between actions and angles for stars in the Geneva-Copenhagen Survey sample estimated with the Stäckel fitting method and the data from McMillan (2011b)

	RMS difference	Expected RMS error
$\Delta J_R / 10^{-4} \text{ kpc}^2 \text{ Myr}^{-1}$	5.4	4.9
$\Delta J_z / 10^{-4} \text{ kpc}^2 \text{ Myr}^{-1}$	7.6	4.9
$\Delta \theta_R / 0.01 \text{ rad}$	2.6	0.9
$\Delta \theta_\phi / 0.01 \text{ rad}$	1.2	0.6
$\Delta \theta_z / 0.01 \text{ rad}$	7.7	40.1

The second column gives the RMS of the expected systematic errors from the Stäckel fitting method

2.3.4.2 Comparison with McMillan (2011b)

McMillan (2011b) calculated the angles and actions of the GCS sample by using the torus machine iteratively. Here we compare the results to our own estimates of the angle-actions. The potential used by McMillan (2011b) was the ‘convenient’ potential detailed in McMillan (2011a), which places the Sun at a Galactocentric radius $R_0 = 8.5 \text{ kpc}$ with a velocity of the local standard of rest $v_{\text{LSR}} = 244.5 \text{ km s}^{-1}$. We calculate the RMS deviations between McMillan’s data and ours, and present the results in Table 2.2. We also show the expected RMS errors in the angles and actions estimated using the Stäckel fitting method. The largest discrepancy between our data and McMillan’s occurs for the θ_z coordinate. The expected error for this coordinate is also large, and very much larger than the actual error. The largest error in θ_z is produced by stars near their turning points. At low z the majority of stars are well away from their turning points, and the resulting errors in the angles are smaller than if we observed a sample of stars that were uniformly sampled in θ_z . Only those stars with very low vertical action have corresponding large errors in θ_z . For all the other variables, the discrepancy between our data and McMillan’s data seems to be in agreement with the expected systematic errors of our method.

2.3.5 Prolate Axisymmetric Potentials

Later we will need to use the above algorithm on prolate axisymmetric potentials, so here we extend the algorithm for this case. Oblate spheroidal coordinates are associated with prolate axisymmetric potentials. These are defined by the roots

$$\frac{R^2}{\tau - a^2} + \frac{z^2}{\tau - c^2} = 1, \quad (2.30)$$

where λ and ν are the roots of τ such that $a^2 \leq \lambda \leq c^2 \leq \nu$. Note that now surfaces of constant ν are oblate spheroids and surfaces of constant λ are two-sheeted hyperboloids of revolution. A general prolate axisymmetric potential accommodates two classes of orbit: inner and outer long-axis loops. An inner long-axis loop crosses $z = 0$ inside the focus $\sqrt{c^2 - a^2}$ and has $\nu_0 = c^2$. These orbits are approximately radially bound by surfaces of constant λ and vertically by surfaces of constant ν . An outer long-axis loop crosses outside this focus and has $\lambda_1 = c^2$. These orbits are approximately radially bound by surfaces of constant ν and vertically by surfaces of constant λ . We choose to set $c^2 = 1$ and set a^2 using Eq. (2.15).⁵ The equations for the actions, angles and frequencies are identical to the oblate case. We, therefore, adopt the same procedure as outlined in Sect. 2.3: we begin by integrating the orbit to find a best choice of coordinate system, find the boundaries of the orbit, fit a Stäckel potential to this region using Eqs. (2.19) and (2.21), and then estimate the actions, angles and frequencies as those in this best-fitting potential. We must ensure that, when a coordinate bounces instead of oscillating, we include an appropriate factor of 2 and we must add an additional π to the appropriate angle when $z < 0$ to remove the degeneracy in the choice of coordinates. Note that, for an outer long-axis loop orbit, $J_R = J_\nu$ and $J_z = J_\lambda$, whereas for an inner long-axis loop orbit $J_R = J_\lambda$ and $J_z = J_\nu$.

As an example, we look at loop orbits in the axisymmetric logarithmic potential described by

$$\Phi(R, z) = \frac{V_c^2}{2} \ln \left(R^2 + \frac{z^2}{q^2} \right), \quad (2.31)$$

where $V_c = 220 \text{ km s}^{-1}$ and q controls the flattening. We generate tori with the actions $(J_R, L_z, J_z) = (0.29, 3.8, 0.45) \text{ kpc}^2 \text{ Myr}^{-1}$ using the torus machine for $q = (0.6, 0.9, 1.1, 1.9)$. On each torus we generate an initial condition $(\mathbf{x}, \mathbf{v})$ with $(\theta_R, \theta_\phi, \theta_z) = (0.1, 0.0, 0.0) \text{ rad}$, and integrate this initial condition for 500 adaptive time-steps. In Fig. 2.11 we show the recovered actions, angles and frequencies for this orbit, along with the input actions and frequencies from the torus machine. For $q = 0.6, 0.9$ the orbit is a short-axis loop, for $q = 1.1$ the orbit is an outer long-axis loop, and for $q = 1.9$ the orbit is an inner long-axis loop. The actions for the orbits in the near-spherical potentials ($q = 0.9, 1.1$) are accurate to $\sim 2\%$, whilst for the flatter potentials ($q = 0.6, 1.9$) the errors increase to $\sim 5\%$. We see the anti-correlation between the radial and vertical actions. As the flattening is increased, we see the expected decrease in the vertical frequency, Ω_z . The frequency recovery has small biases.

⁵Note that a^2 can be less than zero, making a imaginary. However, all physical quantities depend on $\sqrt{c^2 - a^2}$, so this is irrelevant.

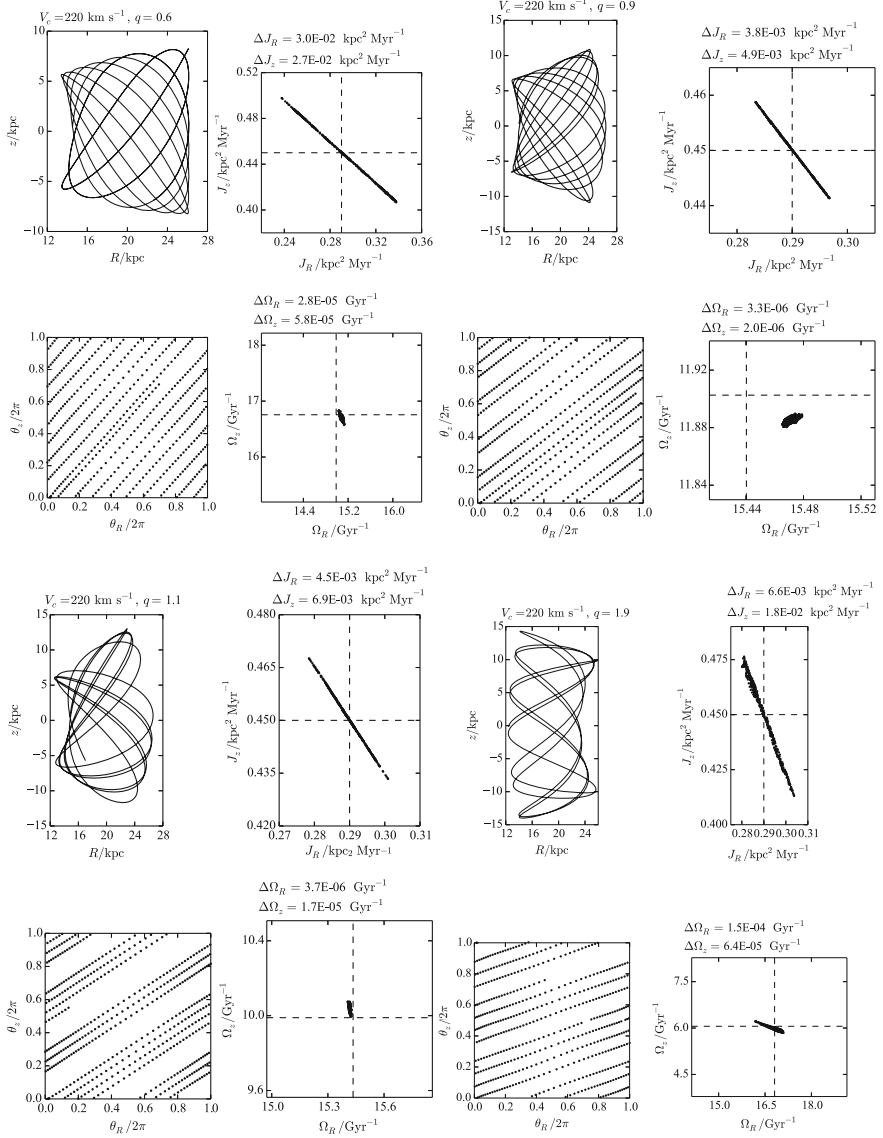


Fig. 2.11 Actions, angles and frequencies for orbits integrated in the logarithmic potential with differing flattening, q . Each set of four panels corresponds to a different flattening. The *top left* is $q = 0.6$, *top right* $q = 0.9$, *bottom left* $q = 1.1$, and *bottom right* $q = 1.9$. Within each set of four panels, we show the orbit in the meridional plane in the *top left* panel, the J_R and J_z estimates in the *top right* (with the respective standard deviations given above each plot), the angle recovery in the *bottom left*, and Ω_R and Ω_z estimates in the *bottom right*. The dashed lines are the true values from the torus machine

2.4 Polar Adiabatic Approximation

We now move on to discuss other methods for estimating the actions in an axisymmetric potential. The adiabatic approximation provides an alternative method for estimating the actions. In its simplest form (Binney 2010), the adiabatic approximation is based on the observation that the oscillation of a star perpendicular to the Galactic plane is much more rapid than the oscillation in the plane, so we may consider the vertical motion to be determined by a slowly varying potential. This approximation assumes that the motion in the plane is unaffected by the motion perpendicular to the plane. The absence of energy transfer between the radial and vertical motion leads to an underestimate of the centrifugal potential for the radial motion and hence an underestimate of the maximum radius of an orbit. Binney and McMillan (2011) attempted to resolve this issue by replacing L_z with $(|L_z| + J_z)$ in the effective radial potential. Schönrich and Binney (2012) improved on this by including a correction to the radial energy due to the changes in the vertical energy along an orbit. It is this final approach that we present here.

Following Schönrich and Binney (2012) we assume that the vertical motion at a given radius, R_0 , is governed by the potential $\Psi_z(z) = \Phi(R_0, z) - \Phi(R_0, 0)$ such that the vertical energy, E_z , is

$$E_z = \frac{1}{2}v_z^2 + \Psi_z(z). \quad (2.32)$$

Then the vertical action is estimated to be

$$J_z = \frac{2}{\pi} \int_0^{z_{\max}} dz v_z, \quad (2.33)$$

where $z_{\max}$ is the vertical height where the vertical velocity, v_z , is zero. By linear interpolation we may reverse this calculation such that, for a given pair of J_z and R_0 , we may calculate $E_z(J_z, R_0)$. Over the course of an orbit we take J_z to be constant but the vertical energy will be changing as the orbit explores different radii. For overall energy conservation, this energy must be transferred from the vertical motion into the radial motion. Therefore, the radial motion is governed by the one-dimensional potential

$$\Psi_R(R) = \Phi(R, 0) + \frac{L_z^2}{2R^2} + E_z(J_z, R) - E_z(J_z, R_c), \quad (2.34)$$

where R_c is the guiding-centre radius (the radius of a circular orbit with z -component of angular momentum L_z). Using this potential, we estimate the radial action as

$$J_R = \frac{1}{\pi} \int_{R_p}^{R_a} dR v_R, \quad (2.35)$$

where R_p and R_a are the radii where the radial velocity, v_R , is zero.

2.5 Ellipsoidal Adiabatic Approximation

The adiabatic approximation of the previous section assumed the motion could be considered separable in cylindrical polar coordinates. We call this method the polar adiabatic approximation (PAA) to differentiate it from the method presented here. It is not obvious that the coordinates in which the motions are ‘most separable’ are the polar coordinates, (R, z) . Prolate spheroidal coordinates (λ, ϕ, ν) provide an alternative set of coordinates for describing motion in an axisymmetric potential, which are important as they are linked to Stäckel potentials. Here we present an alternative method based on the polar adiabatic approximation, but instead considering the motion in the prolate spheroidal coordinates to be separable. We call this method the ellipsoidal adiabatic approximation (EAA).

For a general axisymmetric potential, $\Phi(R, z)$, we recall from Eq. (2.6) that the Hamiltonian in prolate spheroidal coordinates is given by

$$H = \frac{1}{2} \left(\frac{p_\lambda^2}{P_\lambda^2} + \frac{p_\nu^2}{P_\nu^2} + \frac{L_z^2}{R^2(\lambda, \nu)} \right) + \Phi(\lambda, \nu) \quad (2.36)$$

where we have chosen a specific coordinate system, $p_{\lambda, \nu}$ are the conjugate momenta to λ, ν and $P_\lambda^2 = \frac{\lambda - \nu}{(\lambda - a^2)(\lambda - c^2)}$ and $P_\nu^2 = \frac{\nu - \lambda}{(\nu - a^2)(\nu - c^2)}$.

Now we assume the ‘vertical’ motion follows an ellipse of constant $\lambda = \lambda_0$ such that the ν coordinate is determined by the potential

$$\Psi_\nu(\nu) = \frac{L_z^2}{2R^2(\lambda_0, \nu)} - \frac{L_z^2}{2R^2(\lambda_0, c^2)} + \Phi(\lambda_0, \nu) - \Phi(\lambda_0, c^2). \quad (2.37)$$

The energy of the ν oscillations is given by

$$E_\nu = \frac{1}{2} P_\nu^2 \dot{\nu}^2 + \Psi_\nu(\nu) \quad (2.38)$$

such that the vertical action is found as

$$J_z = \frac{2}{\pi} \int_{c^2}^{\nu_+} d\nu p_\nu = \frac{2}{\pi} \int_{c^2}^{\nu_+} d\nu \sqrt{2P_\nu^2(\lambda, \nu)} \sqrt{E_\nu - \Psi_\nu(\lambda, \nu)}, \quad (2.39)$$

where ν_+ is the root of $(E_\nu - \Psi_\nu(\lambda, \nu))$. Therefore, given a 6D phase-space point $(\mathbf{x}, \mathbf{v})$ we may find the best prolate spheroidal coordinate system using Eq. (2.15), evaluate E_ν at this point via the coordinate transformation and then carry out the integration along the curve of constant λ that passes through the phase-space point.

Given a value of λ , L_z and the vertical energy E_ν we are able to find the vertical action. The vertical energy varies with λ whilst the vertical action should, by definition, remain constant. Therefore, this calculation may be reversed such that at a given λ we may determine the vertical energy $E_\nu = E_\nu(\lambda, L_z, J_z)$. This is done by tabulating E_ν for a range of values of λ , L_z and J_z . Armed with this tabulation, we

calculate the radial action in much the same way as when using the polar adiabatic approximation. The radial motion is governed by the effective potential along the $z = 0$ axis given by

$$\Psi_\lambda(\lambda) = \Phi(\lambda, c^2) + \frac{L_z^2}{2R^2(\lambda, c^2)} + E_\nu(\lambda, L_z, J_z). \quad (2.40)$$

The total energy of the orbit is then given by

$$E_{\text{tot}} = \frac{1}{2} P_\lambda^2 \dot{\lambda}^2 + \Psi_\lambda(\lambda) \quad (2.41)$$

and the radial action is calculated as

$$J_R = \frac{1}{\pi} \int_{\lambda_-}^{\lambda_+} d\lambda p_\lambda = \frac{1}{\pi} \int_{\lambda_-}^{\lambda_+} d\lambda \sqrt{2P_\lambda^2(\lambda, c^2) \sqrt{E_{\text{tot}} - \Psi_\lambda(\lambda)}} \quad (2.42)$$

where λ_+ and λ_- are the roots of $(E_{\text{tot}} - \Psi_\lambda(\lambda))$.

An illustration of how this method differs from the polar adiabatic approximation of the previous section is shown in Fig. 2.12. The blue line shows the line along which the PAA vertical-action integration is performed. The red line shows the line of constant λ along which the EAA vertical-action integration is performed. The points at which they intersect gives the $(R, \pm|z|)$ coordinates of the input phase-space point. The EAA clearly captures the shape of the boundaries of the orbit.

As we will see in Sect. 2.8, the EAA offers a considerable improvement over the PAA but unfortunately, due to the energy correction now being a function of three

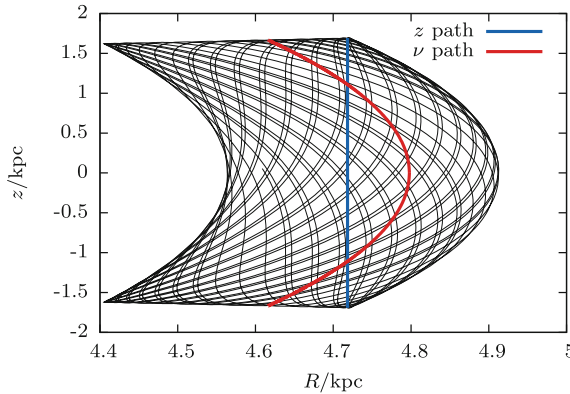


Fig. 2.12 Illustration of the difference between the PAA and EAA. An example orbit is shown by the black line. The blue line shows the line along which the PAA vertical-action integration is performed. The red line shows the line of constant λ along which the EAA vertical-action integration is performed. The points at which they intersect gives the $(R, \pm|z|)$ coordinates of the input phase-space point

variables (λ, L_z, J_z) , this approach takes slightly longer. However, the E_ν tabulation requires a very small number of L_z values (here we use five) over the required L_z range ($0.5 \text{ kpc}^2 \text{ Myr}^{-1} \leq L_z \leq 4.5 \text{ kpc}^2 \text{ Myr}^{-1}$) for a sufficient level of accuracy. Note that the PAA is contained within the EAA. In the limit of very large focal distance, the surfaces of constant λ and ν tend to those of constant R and z . In this limit, the vertical action becomes independent of L_z and the EAA tends to the PAA.

2.6 Axisymmetric Stäckel Fudge

Binney (2012) presented a method for estimating the actions in a general axisymmetric potential by assuming it is close to a Stäckel potential. The presentation here differs from that in Binney (2012), but is chosen to match better to the triaxial extension presented in Chap. 4. For a general oblate axisymmetric potential, Φ , we define

$$\begin{aligned}\chi_\lambda(\lambda, \nu) &\equiv -(\lambda - \nu)\Phi, \\ \chi_\nu(\lambda, \nu) &\equiv (\lambda - \nu)\Phi.\end{aligned}\tag{2.43}$$

If Φ were a Stäckel potential, these quantities would be given by

$$\begin{aligned}\chi_\lambda(\lambda, \nu) &= f(\lambda) - f(\nu), \\ \chi_\nu(\lambda, \nu) &= f(\nu) - f(\lambda).\end{aligned}\tag{2.44}$$

Therefore, for a general potential we can write,

$$f(\tau) \approx \chi_\tau(\lambda, \nu) + D_\tau,\tag{2.45}$$

where D_τ are constants provided we evaluate χ_λ at constant ν and vice versa. We can write the equations of motion (Eq. 2.10) as

$$2(\tau - a^2)(\tau - c^2)p_\tau^2 = E(\tau - c^2) - \left(\frac{\tau - c^2}{\tau - a^2}\right)\frac{L_z^2}{2} - B_\tau + \chi_\tau(\lambda, \nu),\tag{2.46}$$

where we have defined the integrals of motion $B_\tau = I_3 - D_\tau$. Given an initial phase-space point, we use Eq. (2.15) to find a suitable coordinate system, calculate λ , ν , p_λ and p_ν , and use Eq. (2.46) to find the integrals B_τ . Equation (2.46) is then integrated over an oscillation in τ to find the actions as in Eq. (2.12). We note that for the λ integral we keep ν fixed at the input value, and vice versa. This is the procedure followed in Binney (2012).

2.7 Iterative Torus Machine

The methods presented so far suffer from the disadvantage that the error in the action is not controlled. As the methods are not convergent, we are unable to find the exact action regardless of how much computation we perform. McMillan and Binney (2008) used the torus machine iteratively to find the actions and angles of a given phase-space point. This typically involved around 20 torus fits per phase-space point and relied on a good initial guess for the actions and angles for fast convergence. Such a method is potentially the most accurate way to determine the angle-action variables but also the most costly. The authors report that it takes around 15 seconds to perform the iterative procedure on a single phase-space point. Improved methods for the initial estimate will naturally improve the usability of an iterative approach. We can use the results from the Stäckel-fitting approach presented in this chapter as an input to an iterative torus scheme. Here we detail such an approach.

1. Given an initial phase-space point $(\mathbf{x}_i, \mathbf{v}_i)$, we use the Stäckel-fitting method to find an approximate angle-actions $(\boldsymbol{\theta}_S, \mathbf{J}_S)$.
2. We construct a torus of action $\mathbf{J}_S$.
3. We find the closest phase-space point to $(\mathbf{x}_i, \mathbf{v}_i)$ on the torus by minimizing the tolerance

$$\eta = |\boldsymbol{\Omega}|^2 |\mathbf{x}(\boldsymbol{\theta}) - \mathbf{x}_i|^2 + |\mathbf{v}(\boldsymbol{\theta}) - \mathbf{v}_i|^2 \quad (2.47)$$

with respect to the angle coordinates of the torus, $\boldsymbol{\theta}$, using $\boldsymbol{\theta}_S$ as an initial guess. Here $\boldsymbol{\Omega}$ is the frequency of the torus and is used to equate position and velocity separations. We perform the minimisation using the Nelder-Mead simplex algorithm (Nelder and Mead 1965). This minimisation procedure produces a new phase-space point $(\mathbf{x}_S, \mathbf{v}_S)$.

4. For the phase-space point $(\mathbf{x}_S, \mathbf{v}_S)$, we use the Stäckel-fitting method to find an estimate of the angle-actions $(\boldsymbol{\theta}_P, \mathbf{J}_P)$.
5. The error in the action reported by the Stäckel approximation is given by $\Delta \mathbf{J} = \mathbf{J}_P - \mathbf{J}_S$. Therefore, if we assume that the error made by the Stäckel approximation is approximately constant over this small region of phase space, we expect that our initial estimate, $\mathbf{J}_S$, is also in error by $\Delta \mathbf{J}$. A better approximation to the action is

$$\mathbf{J}_{\text{new}} = \mathbf{J}_S - \Delta \mathbf{J} = \mathbf{J}_S + (\mathbf{J}_S - \mathbf{J}_P). \quad (2.48)$$

6. We can then construct a torus of action $\mathbf{J}_{\text{new}}$ and repeat the procedure to refine our estimate of the mis-estimate $\Delta \mathbf{J}$.

This method should converge provided the errors in the Stäckel algorithm and the errors in the torus construction vary smoothly with position and velocity.

This procedure is preferable to the Stäckel-fitting method as the accuracy of the procedure can be assessed by evaluating the tolerance, η . However, this increased accuracy will necessarily come with an increase in computation time. To test the

procedure, we produce a series of phase-space points on the torus with actions $(J_R, L_z, J_z) = (0.25, 3.5, 0.5) \text{ kpc}^2 \text{ Myr}^{-1}$ using the torus machine with an accuracy $\Delta J/J = 10^{-7}$. In Fig. 2.13, we plot the action estimates for this series of phase-space points from the Stäckel-fitting algorithm along with the recovered actions after a single iteration of the above procedure, and the action estimates after the algorithm has reached a tolerance of $\eta = (0.1 \text{ km s}^{-1})^2$, or the number of iterations exceeds 20. We find that the Stäckel algorithm initially produces a spread in the action estimates of approximately 4 %. After one iteration of the above routine, we yield actions accurate to ~ 0.1 %. Further iterations yield slightly more accurate actions (~ 0.05 %). It appears a single iteration is satisfactory and the benefit of further iterations is small. From the inset in Fig. 2.13, we see that the majority of initial conditions only took a few iterations to converge to the required tolerance. Note that some phase-space points took up to 20 iterations of the procedure and still did not converge to the required tolerance. This is not to say that the iterative procedure wandered away from the target actions, but just that the iterative procedure oscillated near the tar-

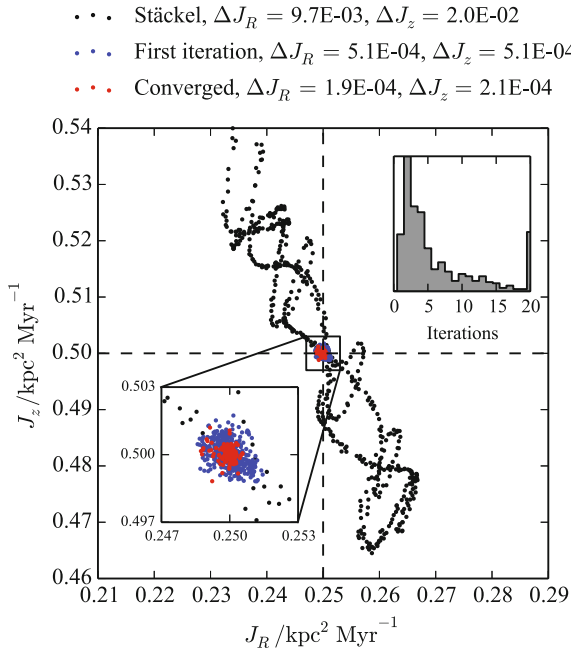


Fig. 2.13 Actions recovered using the iterative torus machine approach for a series of phase-space points on the torus with actions $(J_R, L_z, J_z) = (0.25, 3.5, 0.5) \text{ kpc}^2 \text{ Myr}^{-1}$. The *black points* show the actions from the first estimate using the Stäckel-fitting method from this chapter. The *red points* show the recovered actions after one iteration of the procedure, and the *blue points* show the actions after the procedure was deemed to converge to a tolerance of $\eta = (0.1 \text{ km s}^{-1})^2$. The standard deviations of the three sets of points are shown above the plot. The *bottom-left* inset shows a zoom-in of the highlighted rectangle in the main plot, and the *top-right* inset shows a histogram of the number of iterations performed

get actions without ever reaching the required tolerance. This is probably due to non-linearities in both the Stäckel-fitting method and the torus construction. If one chooses to construct more accurate tori i.e. lower $\Delta J/J$, the number that fail to converge is reduced. Finally, we note that the above scheme takes approximately 10 seconds to perform three iterations.

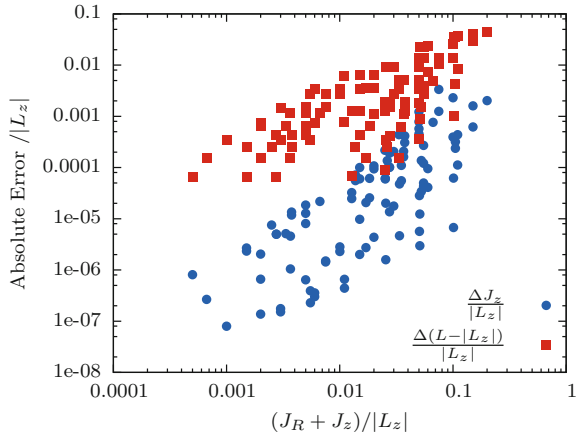
2.8 Method Comparison

When estimating actions, we are aiming for a balance between accuracy of the action recovery and speed. The most accurate method is the convergent iterative torus approach we presented in the previous section. However, this accuracy comes at the expense of additional computational time required to construct tori. Here we critically compare the methods presented in the chapter by investigating the accuracy of the action estimates and time taken to perform many action evaluations. We do not include the iterative torus approach in the following comparisons as the torus machinery is used to generate phase-space points of known actions. Therefore, we expect the approach to converge to the same accuracy as that used to generate the tori.

2.8.1 Total Angular Momentum

Before investigating the methods presented in this chapter, we show how the Stäckel-fitting method compares to the crudest estimate of the vertical action. Some authors have hunted for structure in the distribution of stars in spaces defined by phase space functions other than the actions. For example, Helmi and de Zeeuw (2000) use the set of variables (E, L_z, L) , where $L = |\mathbf{x} \times \mathbf{v}|$ is the total angular momentum, to attempt to find substructure within numerical simulations of disrupted satellite galaxies, whereas Helmi et al. (2006) considered the ‘APL space’ of apocentre, pericentre and z -component of the angular momentum in order to identify signatures of past accretion events in the Geneva-Copenhagen Survey of the solar neighbourhood (Nordström et al. 2004). The total angular momentum is only conserved when we are considering spherical potentials. In a spherical potential, the vertical action is simply $J_z = L - |L_z|$. Here we investigate how much better we are doing when we estimate the vertical action using a Stäckel fit than if we simply use L . Figure 2.14 shows the absolute RMS error in the vertical action for the 100 low-action tori taken from the lower panel of Fig. 2.6 along with the RMS error in the spherical vertical action, $(L - |L_z|)$. The Stäckel-fitting method gives approximately two orders of magnitude improvement in the vertical-action error compared to simply using L .

Fig. 2.14 RMS deviations of J_z for the Stäckel fitting method (blue circles) and the RMS deviations in the spherical vertical action ($L - |L_z|$) (red squares) for the 100 low-action tori detailed in Sect. 2.3.3.2



2.8.2 Single Torus

We now compare the action recovery for two tori using four methods: the polar adiabatic approximation (PAA), the ellipsoidal adiabatic approximation (EAA), the Stäckel fudge (SF), and locally fitting Stäckel potentials (FIT). We do not compare the accuracy of the convergent iterative torus approach as this approach can be made arbitrarily accurate. We use the torus machine to generate 10,000 phase-space points on the tori $(J_R, L_z, J_z) = (0.001, 2.0, 0.001)\text{kpc}^2\text{Myr}^{-1}$, and $(J_R, L_z, J_z) = (0.1, 2.0, 0.1)\text{kpc}^2\text{Myr}^{-1}$ in the ‘best’ potential from McMillan (2011a), and estimate the actions using the four methods. In Fig. 2.15 we show the two tori in the meridional plane along with the action estimates for each method. The most accurate method for both tori is the FIT method, whilst the least accurate is the PAA. For the low-action torus, the PAA produces actions accurate to $\sim 1\%$, whilst the methods based on Stäckel potentials all produce actions accurate to $\sim 0.01\%$. The main difference between the three methods based on Stäckel potentials is the recovery of the radial action which is most accurate for the FIT method, and least accurate for the EAA. For the high-action torus, the PAA produces actions accurate to $\sim 10\%$, whilst the other three methods produce actions accurate to $\sim 4\%$. The SF method is the least accurate of these three and the FIT method is the most accurate, although the differences are very subtle.

2.8.3 Multiple Tori

We now inspect the accuracy of the action recovery for a selection of tori. We use the ‘low-action’ tori from Sect. 2.3.3.2. These 100 tori probe the region $3\text{ kpc} < R < 22\text{ kpc}$, $|z| < 5\text{ kpc}$ and are chosen to be representative of disc-type tori. For each

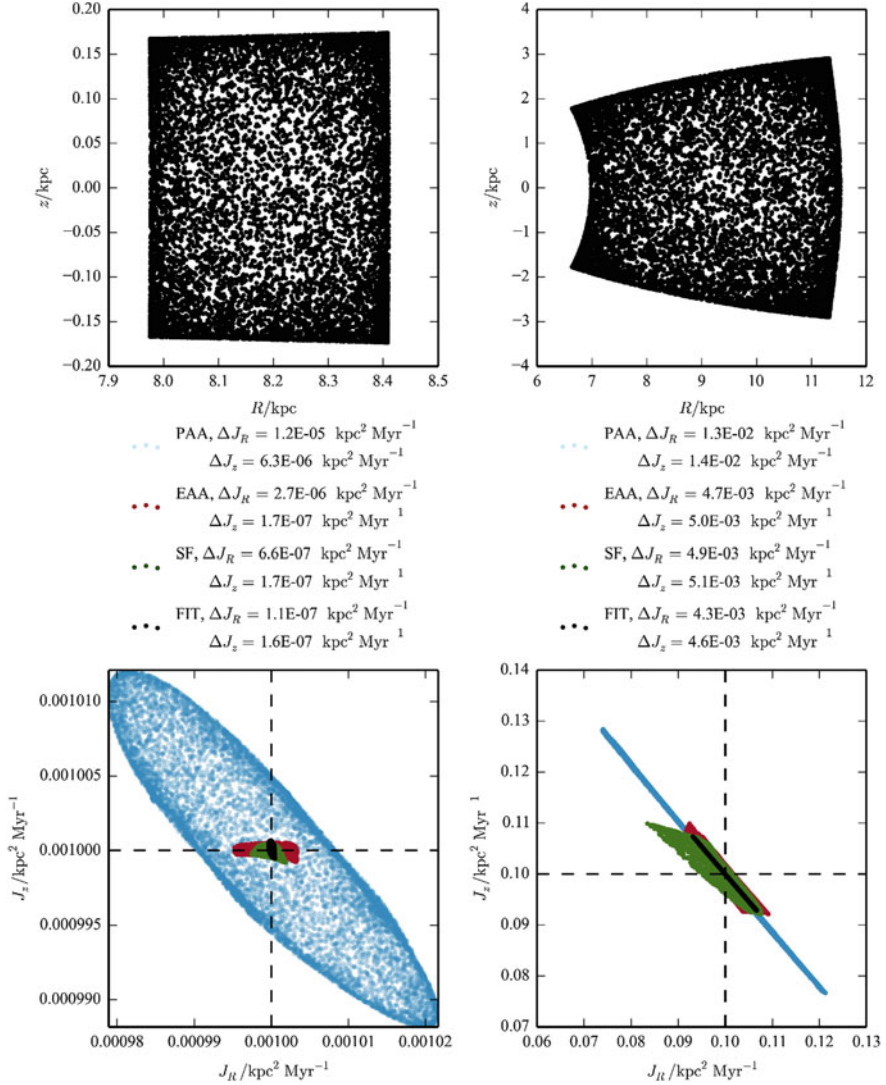


Fig. 2.15 Accuracy of actions using the polar adiabatic approximation (PAA), ellipsoidal adiabatic approximation (EAA), Stäckel fudge (SF) and locally fitting Stäckel potentials. The *top two panels* show 10,000 (R, z) points on two tori in McMillan (2011a) ‘best’ potential with actions $(J_R, L_z, J_z) = (0.001, 2.0, 0.001) \text{ kpc}^2 \text{ Myr}^{-1}$ (left), and $(J_R, L_z, J_z) = (0.1, 2.0, 0.1) \text{ kpc}^2 \text{ Myr}^{-1}$ (right). The *lower two panels* show the radial and vertical action estimates for these phase-space points using four methods: the PAA in blue, the EAA in red, the Stäckel fudge method (SF) from Binney (2012) in green, and locally fitting Stäckel potentials (FIT) in black. We show the standard deviations of the action estimates for each method between the plots

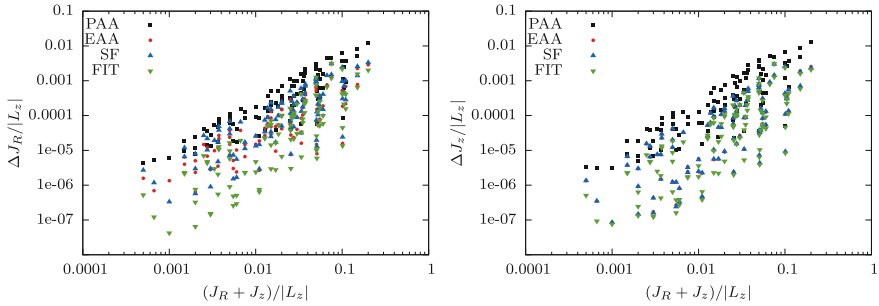


Fig. 2.16 Absolute RMS deviations of the radial (*left*) and vertical (*right*) actions for the polar adiabatic approximation (PAA, *black squares*), the ellipsoidal adiabatic approximation (EAA, *red diamonds*), the Stäckel fudge method (SF, *blue upwards-pointing triangles*), and the Stäckel fitting method (FIT, *green downwards-pointing triangles*). In the *right panel*, the *red diamonds* are not visible as the error in the vertical action reported by the EAA is near identical to that reported by the SF

Table 2.3 Comparison of the accuracy of the action recovery for three tori along with the time taken to evaluate the actions of 10,000 phase-space points for each of the four methods presented in this chapter

$J_R/ L_z $	$J_z/ L_z $	$\Delta J_R^{\text{PAA}}/\Delta J_R^{\text{FIT}}$	$\Delta J_R^{\text{EAA}}/\Delta J_R^{\text{FIT}}$		$\Delta J_R^{\text{SF}}/\Delta J_R^{\text{FIT}}$
0.001	0.001	91	7.4		6.5
0.01	0.01	46	7.0		8.4
0.1	0.1	6.5	1.5		1.7
$J_R/ L_z $	$J_z/ L_z $	Time _{PAA} /s	Time _{EAA} /s	Time _{SF} /s	Time _{FIT} /s
0.001	0.001	4.5	7.2	2.5	10.1
0.01	0.01	6.4	8.3	2.6	15.1
0.1	0.1	8.3	10.7	2.6	44.6

torus, we plot the standard deviation in the action estimates for 10,000 randomly selected phase-space points from each of the four methods in Fig. 2.16.

For all methods, we find that the error in the actions correlates approximately with $(J_R + J_z)/|L_z|$. The PAA is clearly the least accurate of the four methods in both the recovery of J_R and J_z . The SF and EAA give near identical errors in J_z such that we cannot see the EAA points in the plot as they are all behind the SF points. We are using identical choices of a^2 for the two methods so the accuracy of the J_z recovery seems limited by this choice. The error in the J_R recovery is, in general, larger for the SF method than the EAA method, but not significantly. The most accurate method for both J_R and J_z is the FIT method. In Table 2.3 we show the ratio of the RMS spreads in the J_R recovery for each method. We see for low-action tori, $J_i/|L_z| \lesssim 0.1$, where $i = R, z$, the FIT method is 2 orders of magnitude more accurate than the PAA, whilst for $J_i/|L_z| \approx 0.1$ this reduces to a single order of magnitude. For the low-action tori $J_i/|L_z| \lesssim 0.1$ the FIT method is 2 orders of

magnitude more accurate than both the EAA and SF methods, but this reduces to only a 1.5 increase in accuracy for $J/|L_z| \lesssim 0.1$. We expect that the FIT method is the most accurate of the Stäckel methods as it uses the most information to calculate an appropriate value of a^2 , whilst the EAA and SF use only the initial phase-space point to estimate a^2 .

It is also worth noting that the relative errors for the PAA can be as much as order one for orbits with large J_z —the PAA performs best for orbits that stay near the plane. The accuracy of the actions estimated by the EAA are approximately an order of magnitude better than those determined using the PAA.

2.8.4 Computational Cost

In Table 2.3, we also give the time taken to evaluate 10,000 actions on a 2.5 GHz i5 processor. For all tori, the SF method is the fastest, followed by the PAA, then the EAA, and finally the FIT method takes the longest time. We can understand this hierarchy as the SF method requires no interpolation, the PAA uses a 2D interpolation, the EAA uses a 3D interpolation and finally the FIT method requires an orbit integration and interpolation of the resulting fitted potential. For $J_i/|L_z| = 0.001$ ($i = R, z$) the FIT method requires 4 times the computational expense for a factor of 6.5 increase in the accuracy over the SF method, whilst this increases to 20 times the computational expense for only a 1.7 increase in accuracy for $J_i/|L_z| = 0.1$. The EAA takes $\gtrsim 4$ times longer than the SF for only a marginal increase in the action accuracy.

In conclusion, the full Stäckel-fitting method presented in this chapter is only marginally more accurate than both the ellipsoidal adiabatic approximation and the Stäckel fudge method at the expense of more computation. Both the ellipsoidal adiabatic approximation and the Stäckel fudge method provide a clear route to tabulating the actions on a grid which then may be interpolated allowing significant speed-ups (see Binney (2012) and Chap. 8). However, the adiabatic approximations require an initial tabulation of the vertical energies and any call outside this grid is problematic. The Stäckel fudge does not suffer from this disadvantage as any action calls outside the grids may be found on the fly. For low-action applications, such as disc modelling, the ellipsoidal adiabatic approximation and the Stäckel fudge are superior methods, but for more accurate high-action applications, such as stream modelling, the Stäckel-fitting method is slightly superior.

2.9 Conclusions

We have detailed methods for estimating the angle-action variables in a general axisymmetric potential given a 6D phase-space point, $(\mathbf{x}, \mathbf{v})$. The first and most thoroughly explored method is based on locally fitting a Stäckel potential to the region of the potential an orbit probes and then estimating the true actions and angles

as those in the fitted Stäckel potential. We have investigated the systematic errors of this method by producing phase-space points of known actions and angles using the torus machine (McMillan and Binney 2008) and then assessing how well the method can reproduce these variables. For a single representative disc-type torus, the errors in the angles are largest for phase-space points near apsis and the errors in the actions are of order a few percent. For a collection of tori chosen to be representative of both disc and halo-type tori, the absolute error in the actions is found to scale with the sum of the vertical and radial actions. The errors in the angles scale with the relative error in their corresponding action. We demonstrated the use of the method by application to the Geneva-Copenhagen Survey (GCS). As this survey is only local, the angle-action space does not reveal much more information than velocity space. However, we present angle-action coordinates for the peaks of the clumps and streams present in the survey and use them to study the relative impact on estimated angles and actions of observational errors and the known systematic errors of the method. We show that the observational errors are dominant. In Chap. 8, we will inspect the GCS data again in the context of extended distribution function modelling of the Galactic disc.

In addition to the Stäckel-fitting method, we have also presented two new methods for finding angles, actions and frequencies in an axisymmetric potential. The first builds on the adiabatic approximation presented in Schönrich and Binney (2012), but, instead of assuming the motion is separable in cylindrical polar coordinates, assumes the motion is separable in prolate spheroidal coordinates. We demonstrated the improvement in the accuracy of the actions calculated for a range of orbits, and showed that the accuracy rivalled the full Stäckel-fitting method for a much smaller computational cost. One shortcoming of the methods based on fitting Stäckel potentials is that they can not be made arbitrarily accurate. The accuracy is limited by how close the true potential is to a Stäckel potential over the region of interest. The torus machinery (McMillan and Binney 2008) constructs a series expansion for the generating function from a toy to target torus, and can be made arbitrarily accurate by using more terms in the generating function. Such a procedure is designed to find $(\mathbf{x}, \mathbf{v})$ given $(\mathbf{J}, \boldsymbol{\theta})$. We presented an iterative method for finding $(\mathbf{J}, \boldsymbol{\theta})$ from $(\mathbf{x}, \mathbf{v})$ by using the Stäckel-fitting procedure as a first estimate, and refining the estimate using the torus machinery. We showed that, for a single torus, only a single iteration was required to improve the error in the action by a factor of 20, and further iterations did little to improve this.

We closed the chapter by comparing the presented methods for estimating the actions in an axisymmetric potential. The Stäckel-fitting method gives results approximately two orders of magnitude more accurate than assuming the potential is spherical. The Stäckel fudge method from Binney (2012) provides the best compromise between speed and accuracy for both low- and high-action disc tori. However, it is not as accurate as the full Stäckel-fitting method, particularly at low actions. The polar adiabatic approximation is improved on by the ellipsoidal adiabatic approximation which offers an accuracy comparable to the Stäckel fudge method but at greater computational expense.

We have demonstrated that the methods presented here are suitable for many disc and halo-type orbits. The procedure will not work for resonant orbits or chaotic orbits. However, the occurrence of these orbits in realistic galaxy axisymmetric potentials is rare and the great majority of stars are on quasi-periodic non-resonant orbits (Ollongren 1962; Martinet and Mayer 1975).

It is hoped that these methods will lead to more widespread use of angle-action variables when analysing data. Whilst the GCS can be easily analysed in velocity space, angle-action variables should enable us to reveal structures, which are more dispersed in phase-space, in larger surveys.

2.9.1 *Future Work*

The potential used for testing the methods in this chapter consists of two stellar discs, bulge and halo. The gas disc is also a crucial component of any Galactic model (Dehnen and Binney 1998), and is confined to the plane with a scale-height of 40 pc. This component will affect the action recovery for low vertical action tori and some methods presented in this chapter may be more appropriate than others for dealing with this.

We have limited the discussion in this chapter to axisymmetric potentials. Whilst for our own spiral galaxy the axisymmetric approach may suffice, for analysing elliptical galaxies a triaxial approach must be developed. There are also triaxial Stäckel potentials (see de Zeeuw 1985 and Chap. 4), so it should be possible to expand the approach outlined in this chapter to triaxial potentials. The extension of the angle-action estimation is simple, but the fitting procedure is more complex when the potential is triaxial. de Zeeuw and Lynden-Bell (1985) discuss how a general triaxial potential may be fitted both locally and globally by a Stäckel potential. The method for global fitting is the three-dimensional generalisation of the method used in this chapter so involves multiple multi-dimensional integrals. Also, the best choice of coordinate system involves minimising the least-square difference with respect to two coordinate parameters, so a more computationally expensive procedure than the simple method used in this chapter may be required for finding the best coordinate system. In the next two chapters, we will go on to discuss other approaches to estimating the actions in a general triaxial potential.

References

- Antoja T, Figueras F, Torra J, Valenzuela O, Pichardo B (2010) *Lect Notes Essays Astrophys* 4:13
 Bacon R et al (2001) *MNRAS* 326:23
 Bahcall JN, Schmidt M, Soneira RM (1982) *ApJL* 258:L23
 Binney J (2010) *MNRAS* 401:2318
 Binney J (2012) *MNRAS* 426:1324
 Binney J, McMillan P (2011) *MNRAS* 413:1889

- Binney J, Tremaine S (2008) Galactic dynamics, 2nd edn. Princeton University Press, Princeton
- Cappellari M et al (2011) MNRAS 413:813
- De Bruyne V, Leeuw F, Dejonghe H (2000) MNRAS 311:297
- De Simone R, Wu X, Tremaine S (2004) MNRAS 350:627
- de Zeeuw PT, Lynden-Bell D (1985) MNRAS 215:713
- de Zeeuw T (1984) Ph.D. thesis, Leiden University
- de Zeeuw T (1985) MNRAS 216:273
- de Zeeuw T, Peletier R, Franx M (1986) MNRAS 221:1001
- Dehnen W (1998) AJ 115:2384
- Dehnen W, Binney J (1998) MNRAS 294:429
- Dejonghe H, de Zeeuw T (1988) ApJ 329:720
- Famaey B, Jorissen A, Luri X, Mayor M, Udry S, Dejonghe H, Turon C (2005) A & A 430:165
- Helmi A, de Zeeuw PT (2000) MNRAS 319:657
- Helmi A, Navarro JF, Nordström B, Holmberg J, Abadi MG, Steinmetz M (2006) MNRAS 365:1309
- Holmberg J, Nordström B, Andersen J (2009) A & A 501:941
- Jasevicius V (1994) Baltic Astron 3:232
- Kent SM, de Zeeuw T (1991) AJ 102:1994
- Martinet L, Mayer F (1975) A & A 44:45
- McMillan PJ (2011a) MNRAS 414:2446
- McMillan PJ (2011b) MNRAS 418:1565
- McMillan PJ, Binney JJ (2008) MNRAS 390:429
- Navarro JF, Frenk CS, White SDM (1996) ApJ 462:563
- Nelder JA, Mead R (1965) Comput J 7:308
- Nordström B et al (2004) A & A 418:989
- Ollongren A (1962) Bull Astron Inst Neth 16:241
- Perryman MAC et al (2001) A & A 369:339
- Sanders J (2012) MNRAS 426:128
- Schönrich R, Binney J (2012) MNRAS 419:1546
- Schönrich R, Binney J, Dehnen W (2010) MNRAS 403:1829
- Schwarzschild M (1979) ApJ 232:236
- Sellwood JA (2010) MNRAS 409:145
- Zwitter T et al (2008) AJ 136:421

Dynamics of the Milky Way

Tidal Streams and Extended Distribution Functions for
the Galactic Disc

Sanders, J.

2015, XXII, 264 p. 85 illus., 35 illus. in color., Hardcover

ISBN: 978-3-319-18771-6