A fast algorithm for estimating actions in triaxial potentials

Jason L. Sanders\textsuperscript{1,2\textdagger} and James Binney\textsuperscript{1}

\textsuperscript{1}Radolf Peierls Centre for Theoretical Physics, Keble Road, Oxford OX1 3NP, UK
\textsuperscript{2}Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK

Accepted 2014 December 5. Received 2014 December 4; in original form 2014 September 5

ABSTRACT

We present an approach to approximating rapidly the actions in a general triaxial potential. The method is an extension of the axisymmetric approach presented by Binney, and operates by assuming that the true potential is locally sufficiently close to some Stäckel potential. The choice of Stäckel potential and associated ellipsoidal coordinates is tailored to each individual input phase-space point. We investigate the accuracy of the method when computing actions in a triaxial Navarro–Frenk–White potential. The speed of the algorithm comes at the expense of large errors in the actions, particularly for the box orbits. However, we show that the method can be used to recover the observables of triaxial systems from given distribution functions to sufficient accuracy for the Jeans equations to be satisfied. Consequently, such models could be used to build models of external galaxies as well as triaxial components of our own Galaxy. When more accurate actions are required, this procedure can be combined with torus mapping to produce a fast convergent scheme for action estimation.

Key words: methods: numerical – Galaxy: kinematics and dynamics – galaxies: kinematics and dynamics.

1 INTRODUCTION

The haloes that form in baryon-free cosmological simulations almost always have triaxial shapes (Jing & Suto 2002; Allgood et al. 2006; Vera-Ciro et al. 2011). When baryons are added to the simulations, many dark haloes become more spherical (Kazantzidis et al. 2004; Bailin et al. 2005; Valluri et al. 2010), but the most successful current models suggest our Galaxy’s dark halo is triaxial (Law & Majewski 2010; Vera-Ciro & Helmi 2013). Moreover, there is considerable observational evidence that the so-called cored, slowly-rotating elliptical galaxies are generally triaxial (Cappellari et al. 2011). Hence dynamical models of triaxial stellar systems are of considerable astronomical interest.

The first triaxial models were made by violent relaxation of an $N$-body model (Aarseth & Binney 1978), and these models prompted Schwarzschild (1979) to develop the technique of orbit superposition so triaxial models with prescribed density profiles could be constructed. Schwarzschild’s work gave significant insight into how triaxial systems work for the first time, and this insight was enhanced by de Zeeuw (1985), who showed that Stäckel potentials provided analytic models of orbits in a very interesting class of triaxial systems. The most important subsequent development in the study of triaxial systems was the demonstration by Merritt & Valluri (1999) that when a triaxial system lacks a homogeneous core, as real galaxies do, box orbits tend to become centrophobic resonant box orbits.

Work on axisymmetric models in the context of our Galaxy has increased awareness of the value in stellar dynamics of the intimately related concepts of the Jeans’ theorem and action integrals. Ollongren (1962) established that the space of quasi-periodic orbits in galactic potentials is three-dimensional. Jeans’ theorem tells us that any non-negative function $f$ on this space provides an equilibrium stellar system. The key to gaining access to the observable properties of this tantalizing array of stellar systems, is finding a practical coordinate system for orbit space. A coordinate system for orbit space comprises a set of three functions $I_1(x, v)$ that are constant along any orbit in the gravitational potential $\Phi(x)$ of the equilibrium system. A major difficulty is that in the case of a self-consistent system $\Phi$ has to be determined from $f(I)$ by computing the model’s density, and the latter can be computed only when $\Phi(x)$ is known. Hence the computation of $\Phi(x)$ has to be done iteratively, and expressions are needed for the $I_1$ that are valid in any reasonable potential $\Phi(x)$, not merely the potential of the equilibrium model.

If the system is axisymmetric, the energy $E$ and component of angular momentum $L_z$ are integrals that are defined for any axisymmetric potential $\Phi(R, z)$, and equilibrium models of axisymmetric systems have been constructed from distribution functions (DFs) of the form $f(E, L_z)$ (Prendergast & Tomer 1970; Wilson 1975; Rowley 1988). However, these two-integral models are not generic, and they are much harder to construct than generic models when the DF is specified as a function $f(I)$ of the actions (Binney 2014).

\textdagger E-mail: jls@ast.cam.ac.uk

© 2015 The Authors
Published by Oxford University Press on behalf of the Royal Astronomical Society
Moreover, knowledge of the DF as a function of the actions is the key to Hamiltonian perturbation theory, and the ability to perturb models is crucial if we are to really understand how galaxies work, and evolve over time. Actions are also the key to modelling stellar streams, which are themselves promising probes of our Galaxy’s distribution of dark matter (Helmh & White 1999; Tremaine 1999; Eyre & Binney 2011; Sanders & Binney 2013).

Action integrals constitute uniquely advantageous coordinates for orbit space, which is often called action space because its natural Cartesian coordinates are the actions. Actions can be defined for any quasi-periodic orbit and, uniquely among isolating integrals, they can be complemented by canonically conjugate coordinates, the angle variables \( \theta_i \). These have the convenient properties of (i) increasing linearly in time, so

\[
\theta_i(t) = \theta_i(0) + \Omega_i(J)t,
\]

and (ii) being periodic such that any ordinary phase-space coordinate such as \( x \) satisfies \( x(\theta, J) = x(\theta + 2\pi m, J) \), where \( m \) is any triple of integers.

The discussion above amply motivates the quest for algorithms that yield angle-action coordinates \( (\theta, J) \) given ordinary phase-space coordinates \( (x, v) \), and vice versa. These algorithms are usefully divided into convergent and non-convergent algorithms. Convergent algorithms yield approximations to the desired quantity that can achieve any desired accuracy given sufficient computational resource, whereas non-convergent algorithms provide, more cheaply, an approximation of uncontrolled accuracy. Torus mapping (Kaasalainen & Binney 1994; McMillan & Binney 2008) is a convergent algorithm that yields \( x(\theta, J) \) and \( v(\theta, J) \), while Sanders & Binney (2014) introduced a convergent algorithm for \( \theta(x, v) \) and \( J(x, v) \). Both algorithms work by constructing the generating function for the canonical mapping of some ‘toy’ analytic system of angle-action variables into the real phase space. Torus mapping has been demonstrated only in two dimensions, but both axisymmetric and two-dimensional static barred potentials have been successfully handled, and there is no evident obstacle to generalizing to the three-dimensional case. Sanders & Binney (2014) treated the triaxial case, but the restriction to lower dimensions and axisymmetry is trivial.

These convergent algorithms are numerically costly, and, in the axisymmetric case, non-convergent algorithms have been used extensively, especially for extracting observables from a DF \( f(J) \). These extractions require \( J \) to be evaluated at very many phase-space points, and speed is more important than accuracy. The adiabatic approximation (Binney 2010) has been extensively used in modelling the solar neighbourhood (e.g. Schönrich & Binney 2009) but its validity is restricted to orbits that keep close to the Galactic plane. Stäckel fitting (Sanders 2012) has been successfully used to model stellar streams (Sanders 2014a). This method estimates the actions as those in the best-fitting Stäckel potential for the local region a give orbit probes. Sanders (2014b) shows that it is less cost-effective than the ‘Stäckel fudge’ that was introduced by Binney (2012a). Binney (2012b) used the ‘Stäckel fudge’ to model the solar neighbourhood and to explore the first family of self-consistent stellar systems with specified \( f(J) \) (Binney 2014). The Stäckel fudge was recently used by Pfiff et al. (2014) to place by far the strongest available constraints on the Galaxy’s dark halo. In this paper, we extend the Stäckel fudge to triaxial systems.

We begin in Section 2 by showing how to find the actions in a triaxial Stäckel potential. In Section 3, we extend the Stäckel fudge to general triaxial potentials. In Section 4, we apply this algorithm to a series of orbits in a triaxial Navarro–Frenk–White (NFW) potential, and in Section 5 we construct the first triaxial stellar systems with specified DFs \( f(J) \), and demonstrate that, notwithstanding the uncontrolled nature of the fudge as an approximation, the models satisfy the Jeans equations to good accuracy. In Section 6, we describe a new convergent algorithm for obtaining \( (J, \theta) \) from \( (x, v) \). Finally we conclude in Section 7.

## 2 TRIAXIAL STÄCKEL POTENTIALS

In this section, we show how actions can be found in a triaxial Stäckel potential. The presentation here follows that given by de Zeeuw (1985).

### 2.1 Ellipsoidal coordinates

Triaxial Stäckel potentials are expressed in terms of ellipsoidal coordinates \((\lambda, \mu, \nu)\). These coordinates are related to the Cartesian coordinates \((x, y, z)\) as the three roots of the cubic in \( \tau \)

\[
\frac{x^2}{(\tau + \alpha)} + \frac{y^2}{(\tau + \beta)} + \frac{z^2}{(\tau + \gamma)} = 1, \tag{1}
\]

where \(\alpha, \beta, \gamma\) are constants defining the coordinate system. For the potential explored later, we choose to set \( x \) as the major axis, \( y \) as the intermediate axis and \( z \) as the minor axis, such that \(-\gamma \leq v \leq -\beta \leq \mu \leq -\alpha \leq \lambda\). Surfaces of constant \( \lambda \) are ellipsoids, surfaces of constant \( \mu \) are hyperboloids of one sheet (flared tubes of elliptical cross-section that surround the \( x \) axis), and surfaces of constant \( \nu \) are hyperboloids of two sheets that have their extremal point on the \( z \)-axis. In the plane \( z = 0 \), lines of constant \( \lambda \) are ellipses with foci at \( y = \pm\Delta_1 \equiv \pm\sqrt{\beta - \alpha} \), whilst, in the plane \( x = 0 \), lines of constant \( \mu \) are ellipses with foci at \( z = \pm\Delta_2 \equiv \pm\sqrt{\gamma - \beta} \).

The expressions for the Cartesian coordinates as a function of the ellipsoidal coordinates are

\[
x = \frac{(\lambda + \alpha)(\mu + \alpha)(v + \alpha)}{(\alpha - \beta)(\alpha - \gamma)(\lambda + \alpha)}, \tag{2}
\]

\[
y = \frac{(\lambda + \beta)(\mu + \beta)(v + \beta)}{(\beta - \alpha)(\beta - \gamma)(\lambda + \beta)}, \tag{3}
\]

\[
z = \frac{(\lambda + \gamma)(\mu + \gamma)(v + \gamma)}{(\gamma - \alpha)(\gamma - \beta)(\lambda + \gamma)}. \tag{4}
\]

Note that a Cartesian coordinate \((x, y, z)\) gives a unique \((\lambda, \mu, \nu)\), whilst the point \((\lambda, \mu, \nu)\) corresponds to eight points in \((x, y, z)\). Therefore, we will only consider potentials with this symmetry, i.e. triaxial potentials with axes aligned with the Cartesian axes.

The generating function, \( S \), to take us between Cartesian, \((x, y, z, p_x, p_y, p_z)\), and ellipsoidal coordinates, \((\lambda, \mu, \nu, p_\lambda, p_\mu, p_\nu)\), is

\[
S(p_x, p_y, p_z, \lambda, \mu, \nu) = p_x(x(\lambda, \mu, \nu) + p_y(y(\lambda, \mu, \nu) + p_z(z(\lambda, \mu, \nu), \tag{5}
\]

Using \(p_i = \partial S / \partial x_i\) we find, for instance,

\[
p_x = \frac{p_x}{2} \sqrt{\frac{(\mu + \alpha)(v + \alpha)}{(\alpha - \beta)(\alpha - \gamma)(\lambda + \alpha)}} \tag{6}
\]

\[
+ \frac{p_y}{2} \sqrt{\frac{(\mu + \beta)(v + \beta)}{(\beta - \alpha)(\beta - \gamma)(\lambda + \beta)}} \tag{7}
\]

\[
+ \frac{p_z}{2} \sqrt{\frac{(\mu + \gamma)(v + \gamma)}{(\gamma - \alpha)(\gamma - \beta)(\lambda + \gamma)}} \tag{8}
\]
There are similar equations for $p_z$ and $p_\nu$. Inversion of these three equations gives us expressions for $p_z$, $p_x$, and $p_\nu$ as functions of $p$, and $\tau$. For a general triaxial potential, $\Phi$, we can express the Hamiltonian, $H$, in terms of the ellipsoidal coordinates as

$$H = \frac{1}{2}(p_x^2 + p_y^2 + p_z^2) + \Phi(x, y, z).$$

where

$$P^2 = \frac{(\lambda - \mu)(\lambda - v)}{4(\lambda + \alpha)(\lambda + \beta)(\lambda + \gamma)},$$

$$P^2_\mu = \frac{(\mu - v)(\mu - \lambda)}{4(\mu + \alpha)(\mu + \beta)(\mu + \gamma)},$$

$$P^2_\nu = \frac{(v - \mu)(v - \lambda)}{4(v + \alpha)(v + \beta)(v + \gamma)}.$$

2.2 Stöckel potentials

The most general triaxial Stöckel potential, $\Phi_3$, can be written as

$$\Phi_3(\lambda, \mu, v) = \frac{f(\lambda)}{(\lambda - \mu)(v - \lambda)} + \frac{f(\mu)}{(\mu - v)(\mu - \lambda)} + \frac{f(v)}{(v - \mu)(\mu - v)}.$$  

$\Phi_3$ is composed of three functions of one variable. Here, we denote the three functions with the same letter, $f$, as their domains are distinct. Additionally, $f(\tau)$ must be differentiable everywhere and continuous at $\tau = -\alpha$ and $\tau = -\beta$ for $\Phi_3$ to be finite at $\lambda = \mu = -\alpha$ and $\mu = v = -\beta$. With this form for the potential we can solve the Hamilton–Jacobi equation (see de Zeeuw 1985). We write $p_\tau = \partial W/\partial \tau$ and equate the Hamiltonian to the total energy, $E$, in equation (5). We then multiply through by $(\lambda - \mu)(\mu - v)$ to find

$$(v - \mu)\left(2(\lambda + \alpha)(\lambda + \beta)(\lambda + \gamma)\frac{\partial W}{\partial \lambda}\right)^2 - f(\lambda) - \lambda^2 E$$

$$+ (\lambda - v)\left(2(\mu + \alpha)(\mu + \beta)(\mu + \gamma)\frac{\partial W}{\partial \mu}\right)^2 - f(\mu) - \mu^2 E$$

$$+ (\mu - \lambda)\left(2(v + \alpha)(v + \beta)(v + \gamma)\frac{\partial W}{\partial v}\right)^2 - f(v) - v^2 E$$

$$= 0.$$  

(8)

We make the Ansatz $W = \sum W_i(\tau)$ and define

$$U(\tau) = 2(\tau + \alpha)(\tau + \beta)(\tau + \gamma)\frac{\partial W}{\partial \tau} - f(\tau) - \tau^2 E,$$

such that the Hamilton–Jacobi equation becomes

$$(v - \mu)U(\lambda) + (\lambda - v)U(\mu) + (\mu - \lambda)U(v) = 0.$$  

(10)

Taking the second derivative of this expression with respect to $\tau = \{\lambda, \mu, v\}$ we find that

$$U(\tau) = \alpha \tau - b,$$

where $a$ and $b$ are constants. Therefore, the equations for the momenta can be written as

$$2(\tau + \alpha)(\tau + \beta)(\tau + \gamma)p^2 = \tau^2 E - \tau a + b + f(\tau).$$

(12)

For an initial phase-space point, $(x_0, v_0)$, we find $r_0(x_0, v_0)$ and $p_\tau(x_0, v_0)$ using the coordinate transformations and can then find the integrals $a$ and $b$ by solving equation (12) (see de Zeeuw 1985, for more details). These integrals are related to the classical integrals $I_2$ and $I_3$ in a simple way. As $p_\tau$ is only a function of $\tau$, the actions are then given by the 1D integrals

$$J_\tau = \frac{2}{\pi} \int_{\tau_-}^{\tau_+} d\tau \left[p_\tau(\tau)\right].$$

(13)

where $(\tau_-, \tau_+)$ are the roots of $p_\tau(\tau) = 0$, which we find by using Brent’s method to find points where the right-hand side of equation (12) vanishes. Note that for loop orbits, we must divide the ‘radial’ action by two ($J_\tau$ for the short-axis loops and outer long-axis loops, $J_\nu$ for the inner long-axis loops). In Table 1, we give the limits $(\tau_-, \tau_+)$ of the action integrals and the physical meaning of each of the actions for each of the four orbit classes.

The approach to finding the actions presented here requires an explicit form for $f$. In the next section, we will show how we can circumnavigate the need for this explicit form, which allows us to use the same equations for a general potential.

3 The Triaxial Stöckel Fudge

We now show how we can use the insights from Stöckel potentials to estimate actions in a more general potential. For a general triaxial potential, $\Phi$, we can attempt to find the actions by assuming that the general potential is close to a Stöckel potential. Given a general potential we can solve the Hamilton–Jacobi equation for a particular coordinate system, $(\alpha, \beta, \gamma)$ (see Section 4.1). If $\Phi$ were a Stöckel potential, these quantities would be given by, for instance,

$$\chi_\lambda(\lambda, \mu, v) = \frac{f(\mu) - f(v)}{\mu - v} + \frac{v f(\mu) - \mu f(v)}{\mu - v}.$$  

Therefore, for a general potential, we can write

$$f(\tau) \approx \chi(\lambda, \mu, v) + C_\tau + D_\tau,$$

(16)

where $C_\tau$ and $D_\tau$ are constants provided we always evaluate $\chi$, with two of the ellipsoidal coordinates fixed. For instance, we always evaluate $\chi_\lambda$ at fixed $\mu$ and $v$. 

Table 1. Actions in a triaxial Stöckel potential. We give the limits of the action integrals and the physical meaning of each of the actions for each of the four orbit classes. The numbers in brackets after the orbit class are the orbit classification numbers used in Section 4.2.
When we substitute these expressions into equation (12) we find
\[ 2(\tau + \alpha)(\tau + \beta)(\tau + \gamma)p_1^2 = \tau^2 E - \tau A_t + B_t + \chi_t(\lambda, \mu, v). \]
\[ (17) \]
For each \( \tau \) coordinate there are two new integrals of motion given by \( A_t = a - C_t \) and \( B_t = b + D_t \).

Given an initial phase-space point, \((x_0, v_0)\), and a coordinate system, \((\alpha, \beta, \gamma)\), we can calculate the ellipsoidal coordinates \((\lambda_0, \mu_0, v_0)\) using equation (18) to write
\[
B_t = 2(\tau_0 + \alpha)(\tau_0 + \beta)(\tau_0 + \gamma)p_1^2 - \tau_0^2 E + \tau_0 A_t - \chi(\lambda_0, \mu_0, v_0).
\]
\[ (18) \]
It remains to find an expression for \( A_t \) as a function of the initial phase-space point. To proceed, we consider the derivative of the Hamiltonian with respect to \( \tau \). In a Stäckel potential, we can stay on the orbit while changing \( \tau \) and \( p_1(\tau) \) with all the other phase-space variables held constant. Therefore, in a Stäckel potential \( \partial H/\partial \tau = 0 \). Here, we consider \( \partial H/\partial \lambda \) and will give the results for \( \mu \) and \( v \) afterwards. Using equation (5) we write
\[
0 = \left( \frac{\partial H}{\partial \lambda} \right)_{\mu, v} = \frac{1}{2} \frac{\partial}{\partial \lambda} \left[ \frac{p_1^2}{p_2^2} + \frac{1}{2} \frac{p_2^2}{(\mu - \lambda)p_2^2} + \frac{1}{2} \frac{p_3^2}{(\nu - \lambda)p_2^2} + \frac{\partial \Phi}{\partial \lambda} \right].
\]
\[ (19) \]
To evaluate \( \partial [p_1^2/p_2^2] / \partial \lambda \) we use equation (18) to write
\[
2(\lambda + \alpha)(\lambda + \beta)(\lambda + \gamma)p_2^2 = 2(\lambda_0 + \alpha)(\lambda_0 + \beta)(\lambda_0 + \gamma)p_{1,0}^2 + (\lambda^2 - \lambda_0^2)E - \lambda_0 A_t - \chi(\lambda_0, \mu_0, v_0) + \chi_0(\lambda_0, \mu_0, v_0),
\]
\[ (20) \]
that such that
\[
1 \frac{p_1^2}{2} = \frac{Q + (\lambda^2 - \lambda_0^2)E - (\lambda - \lambda_0)A_t}{(\lambda - \mu)(\lambda - \nu)} - \Phi(\lambda, \mu_0, v_0),
\]
\[ (21) \]
where
\[
Q = 2(\lambda_0 + \alpha)(\lambda_0 + \beta)(\lambda_0 + \gamma)p_{1,0}^2 + \chi_0(\lambda_0, \mu_0, v_0).
\]
\[ (22) \]
Upon substituting equation (21) into equation (19), we note that the derivatives of \( \Phi \) cancel. Therefore, evaluating \( \partial H/\partial \lambda \) at the initial phase-space point we find
\[
A_t = 2\lambda_0 E - (2\lambda_0 - \mu_0 - v_0) \left( \Phi(\lambda_0, \mu_0, v_0) + \frac{1}{2} \frac{p_{1,0}^2}{p_2^2} \right) - \frac{1}{2} \frac{p_{1,0}^2(\lambda_0 - v_0)}{p_2^2} - \frac{1}{2} \frac{p_{1,0}^2(\lambda_0 - \mu_0)}{p_2^2}.
\]
\[ (23) \]
This can be simplified further to
\[
A_t = (\mu_0 + v_0)E + \frac{1}{2} \frac{p_{1,0}^2(\mu_0 - \mu_0)}{p_2^2} + \frac{1}{2} \frac{p_{1,0}^2(\mu_0 - v_0)}{p_2^2}.
\]
\[ (24) \]
Note that \( A_t \) is independent of \( \lambda_0 \) and \( p_{1,0} \) (except implicitly in the energy, \( E \)) as \( p_{1,0} \) contains cancelling factors of \((\lambda_0 - \tau_0)\). Similarly
\[
A_{\mu} = (\mu_0 + v_0)E + \frac{1}{2} \frac{p_{1,0}^2(\mu_0 - \mu_0)}{p_2^2} + \frac{1}{2} \frac{p_{1,0}^2(\mu_0 - v_0)}{p_2^2}.
\]
\[ (25) \]
For a true Stäckel potential, given an initial phase-space point we can find six integrals of motion, \((A_1, A_2, A_3, B_1, B_2, B_3)\) from equations (18), (24), and (25). Note that a general Stäckel potential only admits three integrals of motion so the six derived integrals of motion are not independent. This procedure gives identical results to evaluating the integrals as in de Zeeuw (1985). Note that the expressions for these integrals do not explicitly involve the function \( f(\tau) \) - they only involve the potential, \( \Phi \). With the integrals of motion calculated, we are in a position to find \( p_1(\tau) \) and hence the actions from equation (13).

For a general potential we may find six approximate integrals of motion using the same equations, and hence estimate the actions. In this case, although the potential may admit only three true integrals of motion, the six approximate integrals of motion are independent estimates of true integrals of motion. Again, as the expressions do not require \( f(\tau) \) they can be evaluated for a general potential. In Appendix A, we show how the angles and frequencies can be estimated using the same approach.

3.1 Relation to axisymmetric case

The above procedure extends the work of Binney (2012a). Binney (2012a) constructed the ‘Stäckel fudge’ algorithm for estimating actions in a general axisymmetric potential \( \Phi(R, z) \), where \( R \) and \( z \) are the usual cylindrical polar coordinates. We now relate the procedure to that of Binney (2012a) to develop further understanding.

Oblate axisymmetric Stäckel potentials are associated with prolate elliptic coordinates \((\lambda, \nu)\) given by the roots for \( \tau \) of
\[
\frac{R^2}{\tau + \alpha} + \frac{\nu^2}{\tau + \gamma} = 1,
\]
\[ (26) \]
where \(-\nu \leq v \leq -\alpha \leq \lambda\). Binney (2012a) uses the coordinates \((u, v)\) which are related to \((\lambda, \nu)\) via
\[
sinh^2 u = \frac{\lambda + \alpha}{\gamma - \alpha},
\]
\[ (27) \]
\[
\cosh v = \frac{\nu + \gamma}{\gamma - \alpha},
\]
\[ (28) \]
such that
\[
R = \sqrt{\gamma - \alpha} \sinh u \sin v,
\]
\[ (29) \]
\[
z = \sqrt{\gamma - \alpha} \cosh u \cos v.
\]
An oblate axisymmetric Stäckel potential can be written as
\[
\Phi_3(\lambda, \nu) = -\frac{f(\lambda) - f(v)}{\lambda - v},
\]
and the equations for the momenta are given by (de Zeeuw 1985)
\[
2(\tau + \alpha)(\tau + \gamma)p_1^2 = E(\tau + \gamma) - \left( \frac{\tau + \gamma}{\tau + \alpha} \right) I_2 = I_3 + f(\tau).
\]
\[ (30) \]
For axisymmetric potentials \( I_3 = \frac{1}{2} L_z^2 \), where \( L_z \) is the \( z \)-component of the angular momentum. For a general oblate axisymmetric potential, \( \Phi \), we define
\[
\chi(\lambda, \nu) = -(\nu - \lambda)\Phi,
\]
\[ (31) \]
\[
\chi(\lambda, \nu) = -(\nu - \lambda)\Phi.
\]
\[ (32) \]
Therefore, for a general potential, we can write,
\[ f(\tau) = \chi_\alpha(\lambda, v) + D_\tau, \]
where \(D_\tau\) are constants provided we evaluate \(\chi_\alpha\) at constant \(v\) and vice versa. We can write the equations for the momenta as
\[ 2(\tau + \alpha)(\tau + \gamma) p_\tau^2 = E(\tau + \gamma) - \left(\frac{\tau + \gamma}{\tau + \alpha}\right) I_2 - B_\tau + \chi_\alpha(\lambda, v), \]
where we have defined the integral of motion \(B_\tau = I_1 - D_\tau\). \(B_\tau\) may be found given an initial phase-space point and we then integrate the equations for the momenta to find the actions. Note that in this case only two integrals of the motion, \(B_\tau\), need to be found, as, in the axisymmetric case, we can find two exact integrals of motion, \(E\) and \(L_1\). This is the procedure followed in Binney (2012a) and, despite the differing conventions and presentation, this method gives identical results to that of Binney (2012a).

4 TESTS

For the purposes of testing the above algorithm, we use a triaxial NFW halo (Navarro, Frenk & White 1997; Jing & Suto 2002):
\[ \Phi(x, y, z) = \frac{GM_s}{m} \log \left(1 + \frac{m}{m_0}\right) \]
where \(m = \sqrt{x^2 + y^2 + z^2} \).
\[ \Phi(x, y, z) = \frac{GM_s}{m} \log \left(1 + \frac{m}{m_0}\right) \]
We set \(y_s = 0.95\), \(z_s = 0.85\), \(m_0 = 10\) kpc, and \(GM_s = (1109 \text{ km s}^{-1})^2\) kpc. In Fig. 1, we show the equipotential contours in the \(z = 0\) and \(y = 0\) planes. It is perhaps more conventional to include the triaxiality in the density (e.g. Jing & Suto 2002), but, for simplicity, we have chosen to include triaxiality in the potential. For our choice of parameters this leads to negative densities along the \(z\)-axis for \(z \gtrsim 130\) kpc. This is well outside the region we will probe in our experiments so we are not concerned that our model is unphysical at large \(z\).

4.1 Selection of coordinate system

The accuracy of the above routine for a general potential will depend upon our choice of coordinate system, \((\alpha, \beta, \gamma)\). Note that the potential is fixed and this coordinate system acts only as a set of parameters in the algorithm to find the actions. We can freely set \(\gamma = -1\) kpc$^2$ as the coordinate system only depends on \(\Delta_1 = \sqrt{\beta - \alpha}\) and \(\Delta_2 = \sqrt{\gamma - \beta}\). For each orbit, we consider we are in a position to choose different \(\Delta_i\). Here, we consider how we can choose suitable \(\Delta_i\) given an initial phase-space point.

In Sanders (2012), the mixed derivative \(\partial_\alpha \partial_\beta [\ell(\lambda - \nu)\Phi]\) was used to select an appropriate coordinate system in an axisymmetric potential. For the triaxial case, we could construct a similar quantity:
\[ \partial_\alpha \partial_\beta [\ell(\lambda - \mu)(\mu - \nu)(\nu - \lambda)\Phi]. \]
However, this expression would involve third derivatives of the potential so is undesirable. Binney (2014) selected a coordinate system by fitting ellipses to shell orbits at each energy, \(E\). We follow a similar procedure: we assume that the best choice of coordinate system is solely a function of \(E\).

In a Stäckel potential the short-axis closed loops are ellipses confined to the plane \(z = 0\) with foci at \(y = \pm \Delta_1 = \pm \sqrt{\beta - \alpha}\), whilst the long-axis closed loops are confined to the plane \(x = 0\) with foci at \(z = \pm \Delta_2 = \pm \sqrt{\gamma - \beta}\). Additionally, for these closed loop orbits only one of the actions is non-zero (\(J_\beta\) for the short-axis closed loop and \(J_\gamma\) for the long-axis closed loop).

For a general potential, we use these facts to select appropriate values for \(\Delta_\gamma\) using a two-step procedure: given a value for \(E\) we find the two closed loop orbits – one around the short-axis and one around the long-axis, and with these closed orbits found we alter the position of the foci to optimize the action estimates from our algorithm. Note that the structure of the closed orbits is independent of any choice of the foci positions such that the two steps of the procedure are distinct.

First, to find the closed orbits with energy \(E\), we select a point along the intermediate axis, \(y = y_f\), and launch an orbit with speed \(v = \sqrt{2(\Phi(0, y_f, 0))}\) in either the \(x\) (for the short-axis loop) or \(z\) direction (for the long-axis loop). The next time the orbit crosses the \(y\)-axis, we note the \(y\)-intercept, \(y = y_f\) and calculate \([-y_f - y_f]\). We repeat this procedure with a new \(y_f\) until we have minimized \([-y_f - y_f]\) using Brent’s method. We only integrate half of the orbit and assume that the other half can be obtained by symmetry to avoid misidentifying fish-tail resonant orbits as closed loop orbits.

With the closed orbits with energy \(E\) in our potential found, we turn to estimating the location of the foci. Using the long-axis closed loop orbit integration we find an estimate of \(\Delta_2\) by minimizing the standard deviation of the \(J_\gamma\) estimates from each time-step with respect to \(\beta\) using Brent’s method. The action estimates are found using the algorithm outlined in Section 3. This procedure is not sensitive to the choice of \(\alpha\). Once we have found \(\beta\), we perform a similar procedure for the short-axis loop: vary \(\alpha\) until we have minimized the standard deviation of \(J_\beta\).

We perform the above procedure for a range of energies from \(E_{\min} = \Phi(0, y_{\min}, 0)\) to \(E_{\max} = \Phi(0, y_{\max}, 0)\), tabulating the found values of \(\alpha\) and \(\beta\) for interpolation. For the NFW potential, we adopt \(y_{\min} = 0.05\) kpc and \(y_{\max} = 60\) kpc. In Figs 2 and 3, we plot the standard deviation of the actions of the closed loop orbits against \(\Delta_2\) and \(\Delta_1\) for the constant energy surface with \(E = \Phi(0, m_0, 0) = -(290 \text{ km s}^{-1})^2\). In both cases there is a clear minimum in the standard deviation. In Fig. 2, we show the standard deviation in \(J_\gamma\), as a function of \(\Delta_2\), \(\Delta_2 = \sqrt{\gamma - \beta}\) using two different values for \(\alpha\). The results are indistinguishable. Provided we initially choose a sufficiently negative value of \(\alpha\) that the optimal \(\beta\) satisfies \(\beta > \alpha\), we are free to first set \(\Delta_2\) and then choose \(\Delta_1\).

4.1.1 Coordinate system procedure

For clarity, we now summarize the above procedure as follows.

(i) Given a general potential, create a regularly-spaced grid in energy, \(E\), between some minimum and maximum energy.
J. L. Sanders and J. Binney

Figure 2. Standard deviation in $J_\nu$ as a function of $/Delta_2$ for the closed long-axis loop orbit shown in the inset. The solid line shows the results if we set $\alpha = -80$ kpc$^2$ whilst the red crosses show the results if we set $\alpha = -20$ kpc$^2$. The choice of $/Delta_2$ is insensitive to $\alpha$. In the inset, the two red arrows show the initial position vector for the orbit and that position vector rotated by 90° anticlockwise. The black squares show the chosen location of the foci $z = \pm /Delta_2$.

Figure 3. Standard deviation in $J_\mu$ as a function of $/Delta_1$ for the closed short-axis loop orbit shown in the inset. In the inset, the two red arrows show the initial position vector for the orbit and that position vector rotated by 90° clockwise. The black squares show the chosen location of the foci $y = \pm /Delta_1$.

Figure 4. Closed-loop choice of $/Delta_1$ (solid black) and $/Delta_2$ (dashed red) as a function of energy, $E$, for the NFW potential described in Section 4. The range of energies covered corresponds to the energies of particles dropped from 0.5 to 30 kpc along the intermediate axis. The vertical blue dotted line gives the energy of the surface explored in Section 4.2.

(ii) At each grid-point, $E_i$, find the short-axis and long-axis closed loops by integrating orbits launched at $(0, y_k, 0)$ with velocity $\sqrt{2(E_i - \Phi(0, y_k, 0))}$ in the direction of the long-axis or short-axis, respectively. The closed loops will cross the $y$-axis for the first time at $(0, -y_k, 0)$. We store the phase-space points $(x_j, v_j)$ at each time sample $t_j$.

(iii) Minimize the standard deviation of the $J_\mu(x_j, v_j)$ from the long-axis closed loop orbit integration with respect to $\beta$ to find $/Delta_2$.

(iv) Minimize the standard deviation of the $J_\nu(x_j, v_j)$ from the short-axis closed loop orbit integration with respect to $\alpha$ to find $/Delta_1$.

We call the $/Delta_1$ and $/Delta_2$ found using this procedure the closed-loop estimates.

4.1.2 Coordinate system results

In Fig. 4, we have plotted the closed-loop choice of $/Delta_1$ and $/Delta_2$ as a function of the energy. We see that for low energies (very centrally confined orbits) $/Delta_i$ tends to zero. Due to the cusp at the centre of the NFW potential, loop orbits exist right down to the centre of the potential. The foci must lie within these loop orbits so $/Delta_i$ must decrease as we go to lower energy. As we increase the energy $/Delta_i$ increases with $/Delta_1 < /Delta_2$.

To check the closed-loop estimates, we launch a series of orbits of constant energy $E = \Phi(0, m_0, 0) = -(290 \text{ km s}^{-1})^2$ at linearly-spaced intervals along the $y$-axis with velocity vectors in the $(x, z)$ plane oriented at differing linearly-spaced angles, $\theta$, to the $x$-axis and integrate the orbits for approximately 10.3 Gyr storing phase-space points every 0.1 Gyr. Note again that the orbit integration is in the fixed NFW potential and so the structure of an orbit is independent of any choice of $\alpha$ and $\beta$. The choice of $\alpha$ and $\beta$ only affects the recovery of the actions and we wish to find the optimal choice of $\alpha$ and $\beta$ for each orbit, i.e. the choice that makes the actions as constant in time as possible. Therefore, we minimize the sum of the variances of the actions with respect to $\alpha$ and $\beta$. The results of this procedure are shown in Fig. 5. We see that the majority of orbits yield optimal $/Delta_i$ similar to the closed-loop estimates. At the extremes of $y$, $/Delta_i$ deviates from this choice. These are the box
Action estimation in triaxial potentials

Figure 5. Choice of $\Delta_1$ and $\Delta_2$ which minimizes the variation in the actions for a range of orbits confined to a constant energy surface. Each orbit was launched at $y$ on the intermediate axis with angle $\theta$ from the long-axis. The dashed black line gives the values chosen by only inspecting the closed loop orbits as specified in Section 4.1.

orbits and they seem to favour lower $\Delta_i$. At fixed $y$ the choice of $\Delta_i$ is not so sensitive to $\theta$.

We could improve our choice of $\Delta_1$ and $\Delta_2$ by making the choice a function of an additional variable. For instance, we could make the choice a function of the total angular momentum, which is not an integral of motion. However, we will see that we cannot significantly improve the action recovery with a better choice of $\Delta_i$.

4.2 Accuracy

We now briefly inspect the accuracy of the action recovery using the triaxial St"ackel fudge. We take three orbits from the surface of constant energy explored in the previous section. The three orbits are a box orbit with $y = 1.8234$ kpc, $\theta = 0.6$ rad (shown in Fig. 6), a short-axis loop orbit with $y = 4.8234$ kpc, $\theta = 0.4$ rad (shown in Fig. 7), and a long-axis loop orbit with $y = 3.8234$ kpc, $\theta = 1.2$ rad (shown in Fig. 8). The top row of each figure shows three projections of the orbit, while the three lower panels show the action estimates calculated at each point along the orbit using the closed-loop choice of $\Delta_i$ in blue, and in green those obtained with the choice of $\Delta_i$ that minimizes the spread in the action estimates. Clearly no procedure for determining $\Delta_i$ will give superior performance to that obtained with the latter, which is expensive to compute because it requires orbit integration. The intersection of the black lines in the bottom panels of Figs 6 to 8 show the 'true' actions calculated with the method of Sanders & Binney (2014). The distributions of coloured points from the St"ackel Fudge scatter around the true actions, as one would hope. The extent of the green distributions, obtained with the computationally costly values of $\Delta_i$, are at best a factor 2 smaller than the distributions of blue points, obtained with the cheap value of $\Delta_i$. From this experiment, we conclude that there is not a great deal to be gained by devising a better way to evaluate the $\Delta_i$.

In Appendix A, we show how well the angle coordinates are recovered for these orbits.

The actions of the box orbit are $(J_\lambda, J_\mu, J_\nu) = (686, 192, 137)$ kpc km s$^{-1}$ and our method yields errors of $(\Delta J_\lambda, \Delta J_\mu, \Delta J_\nu) = (56, 39, 22)$ kpc km s$^{-1}$ so approximately 10–20 per cent. If we adjust $\Delta_i$ to minimize the spread in the action estimates along the orbit, we find errors of $(\Delta J_\lambda, \Delta J_\mu, \Delta J_\nu) = (17, 19, 16)$ kpc km s$^{-1}$ so approximately $\lesssim$ 10 per cent. We can achieve a factor of 2 improvement for $J_\lambda$ and $J_\mu$.

The actions of the short-axis loop orbit are $(J_\lambda, J_\mu, J_\nu) = (55, 752, 78)$ kpc km s$^{-1}$ and our method yields errors of $(\Delta J_\lambda, \Delta J_\mu, \Delta J_\nu) =$ 56.38, 38.74, 21.75) kpc km s$^{-1}$, $(\Delta J_\lambda / J_\lambda, \Delta J_\mu / J_\mu, \Delta J_\nu / J_\nu) =$ (0.08, 0.219, 0.16)
\( \Delta J_x = (2, 3, 1) \text{ kpc km s}^{-1} \) so \( \lesssim 4 \) per cent. If we adjust \( \Delta_i \) to minimize the spread in the action estimates along the orbit, we find errors of \((\Delta J_x, \Delta J_y, \Delta J_z) = (0.8, 2.0, 0.9) \text{ kpc km s}^{-1}\).

The actions of the long-axis loop orbit are \((J_x, J_y, J_z) = (50, 102, 680) \text{ kpc km s}^{-1}\) and our method yields errors of \((\Delta J_x, \Delta J_y, \Delta J_z) = (4, 5, 6) \text{ kpc km s}^{-1} \) so \( \lesssim 8 \) per cent. If we adjust \( \Delta_i \) to minimize the spread in the action estimates along the orbit, we yield errors of \((\Delta J_x, \Delta J_y, \Delta J_z) = (2.0, 2.5, 4.2) \text{ kpc km s}^{-1}\).

For all the orbits shown in Fig. 5 (sampled from the constant energy surface \( E = \Phi(0, m_0, 0) = -(290 \text{ km s}^{-1})^2 \)), we have plotted

Figure 7. Action estimates for example short-axis loop orbit using the triaxial Stäckel fudge. See Fig. 6 for more information on each panel.

Figure 8. Action estimates for example long-axis loop orbit using the triaxial Stäckel fudge. See Fig. 6 for more information on each panel.
If we wish to draw the consequent defined by \( \sqrt{p} \alpha x = \alpha / \Delta^1 \), \( \beta \mid \) and \( = \mu \mu \sqrt{y} \) we have that \( \gamma \sqrt{z} \) and \( \gamma = -x z \sqrt{\lambda} \), \( \lambda \) shows that the St"uckel approximation (see Table 1): \( \lambda_\pm = -\alpha \), \( \mu_\pm = -\beta \) and \( v_\pm = -\gamma \) correspond to a box orbit (classification number 0), \( \mu_+ = \beta \), \( \mu_- = -\alpha \) to a short-axis loop orbit (1), \( \lambda_\pm = -\alpha \), \( v_\pm = -\beta \) to an inner long-axis loop (2), and \( \mu_+ = -\alpha \), \( v_\pm = -\beta \) to an outer long-axis loop (3). The orbit classification number is calculated as an average of these classifications along the orbit. With this scheme, orbits near the boundaries of the orbit classes that are chaotic or resonant are allocated non-integer orbit classification numbers. We see that the largest action errors occur at the interfaces between the orbit classes. In particular, \( \Delta J_x \) and \( \Delta J_y \) are largest along the box-short-axis-loop interface, whilst \( \Delta J_\mu \) is largest at the box-long-axis-loop interface. It is at these boundaries that the orbits pass close to the foci so clearly our choice of foci affects the action recovery for these orbits.

In general, we find that the action recovery for loop orbits is good, as these orbits probe a small radial range of the potential. For box orbits the recovery deteriorates as these orbits probe a larger central region of the potential. Additionally, we have seen that by altering \( \Delta \), we can achieve up to a factor of 2 improvement in the accuracy of the actions for both the loop and box orbits.

### 4.3 Surfaces of section

For understanding the behaviour of dynamical systems, Poincaré (1892) introduced the concept of a surface of section. These diagrams simplify the motion of a high-dimensional dynamical system. A regular orbit in an integrable triaxial potential permits three constants of the motion, thus confining the motion to a 3-torus. If we choose to only plot the series of points where the orbit passes through a 4-surface in phase space, e.g. defined by \( y = 0 \) and \( z = 0 \), the phase-space points will be confined to a line, or a consequent, which may be visualized clearly.

We can test the St"uckel approach outlined here by seeing how well it reproduces the surfaces of section. To produce the true surface of section, we must integrate the orbit in the true potential and find the phase-space points where the orbit crosses our chosen 4-surface. Here, we use 4-surfaces defined by one of the spatial axes. The orbit will never pass through a given spatial axis in a finite time so we can only produce points arbitrarily close to the given axis. If we require the points where the orbit crosses the \( x \)-axis, we integrate until the \( y \) and \( z \) steps have bracketed \( y = 0 \) and \( z = 0 \). We then bisect the integration step \( N_{\text{max}} \) times choosing the interval that brackets \( y = 0 \) and if the interval still brackets \( z = 0 \) after the \( N_{\text{max}} \) bisections we store the final point. We integrate over our chosen step-size with a Dortmund–Prince eighth-order adaptive integration scheme with an absolute accuracy of \( \epsilon = 1 \times 10^{-10} \). We choose a step-size of 0.005 kpc and set \( N_{\text{max}} = 10 \). This scheme produces points that are \( \leq 0.001 \) kpc away from the \( x \)-axis for the box orbit considered below.

To produce the corresponding surface of section from the St"uckel method, we determine \( r \) along our chosen spatial axis using equation (2) between the determined limits in \( \tau \), and use equation (17) to find the corresponding \( p_r \). From \( p_r \), we can use expressions such as equation (4) to calculate \( \mu_r \), \( p_{\alpha} \), and \( p_\beta \), if we wish to draw the consequent defined by \( y = 0 \), \( z = 0 \) we have that \( \mu = -\beta \) and \( v = -\gamma \) such that \( x = \sqrt{\lambda + \alpha} \), and \( p_\lambda = \sqrt{4(\lambda + \alpha)p_\mu} \). If we wish to draw the consequent defined by \( y = 0 \), \( z = 0 \) we have that for \( |y| > \Delta_1 \), \( \mu = -\alpha \) and \( v = -\gamma \) such that \( y = \sqrt{\lambda + \beta} \), and \( p_\lambda = \sqrt{4(\lambda + \beta)p_\mu} \), whilst for \( |y| < \Delta_1 \), \( \lambda = -\alpha \) and \( v = -\gamma \) such that \( y = \sqrt{\mu + \beta} \) and \( p_\lambda = \sqrt{4(\mu + \beta)p_\mu} \).

Fig. 11 shows that the St"uckel approximation consequents for the short-axis loop lie close to the true consequent. Those of the long-axis loop are slightly worse. The box orbit seems problematic. For

\[ \text{Fig. 9. Logarithm of the fractional error in the actions for a selection of orbits in the constant energy surface } E = \Phi(0, m_\alpha, 0) = -(290 \text{ km s}^{-1})^2 \text{ for the triaxial NFW potential. The } x \text{-axis shows the position along the intermediate axis at which the orbits were launched (y), and the colour-coding in the right-hand panel shows the angle, } \theta, \text{ in the } x-z \text{ plane at which the orbits were launched.} \]
controls whether the $f$ is smaller than its maximum value. This behaviour is $z = 0.7$ for large $z$. However, we see $+|\mu|$, $+|\nu|$, and $+|\lambda|$ of the model. We do not construct a self-consistent model but in $z$ of the model. We do not construct a self-consistent model but instead consider $f$ to be a tracer population in the externally applied

$$
\int d^3v \rho(x, v),
$$

Note that, as the potential is time-independent, the Hamiltonian is time-reversible and we need only integrate over half the velocity space and multiply the result by two. We integrate up to the maximum velocity, $v_{\text{max}}$ at $x$, given by $v_{\text{max}} = \sqrt{2\Phi(x)}$. We will later calculate these moments extensively to demonstrate that the action-based distribution functions obey the Jeans’ equations. In Figs 12 and 13, we plot the density of the radially biased ($\zeta = 3.28$) and tangential-biased ($\zeta = 0.7$) models. We display contours of constant

$$
\Delta \mu / \text{kpc km s}^{-1}
$$

Figure 11. Surfaces of section for the three test orbits in the triaxial NFW potential. In the left-hand panel we show the box orbit, the central panel shows the short-axis loop orbit and the right-hand panel shows the long-axis loop orbit. In each panel, the solid black line gives the true curve of consequents from orbit integration. The narrower coloured lines give the consequents from the Stäckel approximation coloured by $|x_1|$ of the initial phase-space point where $x_1 = x$ for the short-axis loop and box orbit, and $x_1 = y$ for the long-axis loop orbit. The text above each plot gives the plane that defines the surface of section.

5 A TRIAXIAL MODEL WITH SPECIFIED DF

The main purpose of the algorithm presented here is to calculate efficiently the moments of triaxial distribution functions. We have seen that the errors in the actions reported by the scheme can be large. However, when calculating moments of a distribution function, many action evaluations are required and there is scope for errors to substantially cancel, leaving the final value of the moment quite accurate. In this section, we demonstrate this phenomenon by for the first time constructing triaxial models from an analytic DF $f(J)$. We adopt a simple distribution function (Posti et al. 2014)

$$
f(x, v) = f(J(x, v)) = (J_0 + |J_1| + |\zeta J_2| + |\eta J_3|)^p,
$$

where $J_0 = 10 \text{ km s}^{-1} \text{kpc}$ is a scale action, $\zeta$ controls whether the model is tangentially/radially biased and $\eta$ controls the flattening
density in two planes along with the density along a line parallel to the x-axis decomposed into its contributions from each orbit class. The density is calculated using the adaptive Monte Carlo Divonne routine in the \textsc{CUBA} package of Hahn (2005). The class of each orbit is determined by the limits of the motion in $\tau$: $\lambda_\pm = -\alpha$, $\mu_\pm = -\beta$ and $\nu_- = -\gamma$ correspond to a box orbit, $\mu_\pm = \beta$, $\mu_\pm = -\alpha$ to a short-axis loop orbit, and $\nu_- = -\gamma$, $\nu_+ = -\beta$ to a long-axis loop. As we are calculating the density close to the x-axis, the long-axis loop orbits, which loop the x-axis, do not contribute significantly to the density integral. We see that for the radially biased model the box orbits are the dominant contributors whilst for the tangentially biased model the short-axis loop orbits are the major contributors.

We will now perform some checks to see whether our distribution functions are accurate.

5.1 Normalization

One quick check of our action estimation scheme is how accurately it recovers the normalization. To keep the normalization finite we set $p = -3.5$ for this section. We are able to calculate the normalization of our DF in two distinct ways. First, we calculate the normalization analytically from the DF as

$$M_{\text{true}} = (2\pi)^3 \int d^3J f(J)$$

$$= (2\pi)^3 \int_0^\infty \int_0^\infty \int_0^\infty dJ_\lambda dJ_\mu dJ_\nu f(J)$$

$$= -\eta \zeta (p+1)(p+2)(p+3).$$

Note that for each $J$ in the appropriate range there are two loop orbits – one circulating clockwise and one anticlockwise. Therefore, we must multiply the normalization by two for these orbits. However, we have defined the ‘radial’ action to be four times the integral from $\tau_-$ to $\tau_+$ for these orbits so these factors cancel (Binney & Spergel 1984; de Zeeuw 1985). Additionally, we can calculate the mass as

$$M_{\text{est}} = 8 \int_{(x,y,z)>0} d^3x \int d^3v f(J(x,v)).$$
Figure 14. Accuracy of Jeans’ equation calculation for the radially biased model ($\zeta = 3.28$). In the top half of each panel, we show $-\rho \partial \Phi / \partial x_j$ as a series of black dots and $\partial (\rho \sigma_{ij}^2) / \partial x_i$ as a red line. In the bottom half, we show the percentage error difference between these quantities. Each panel shows a single component, i.e. a single $j$, along the line parametrized by the coordinate $x_k$ and given above the top-right corner of each panel. The bottom three panels all correspond to the same line. The grey dashed line is the zero-error line.

For each spatial coordinate, we make the transformation $u_i = 1/(1 + x_i)$ to make the integrand flatter. The limits of the integral are now $u_i = [0, 1]$. To reduce numerical noise, we split the integral such that we calculate the contribution near the axes separately. We perform the integral using the Monte Carlo Divonne routine. For the tangentially biased model ($\zeta = 0.7$), we find $M_{\text{est}} \approx 1.006M_{\text{true}}$ and for the radially biased model ($\zeta = 3.28$) $M_{\text{est}} \approx 1.007M_{\text{true}}$ so despite the often large errors in the actions the normalization of the model is well recovered.

**5.2 The Jeans equation**

Our distribution function must satisfy the collisionless Boltzmann equation

$$\frac{df}{dt} = 0.$$ \hfill (38)

In turn this means the distribution function must satisfy the Jeans equations (see equation 4.209 of Binney & Tremaine 2008)

$$\frac{\partial (\rho \sigma_{ij}^2)}{\partial x_i} = -\rho \frac{\partial \Phi}{\partial x_j}.$$ \hfill (39)

A simple test of our action-based distribution functions is checking whether they satisfy these equations. The right hand side is calculated from analytic differentiation of the potential and multiplying by the density. The left-hand side is found by numerically differentiating the three-dimensional integrals $\rho \sigma_{ij}^2$ and summing the appropriate contributions. Numerical differentiation of an integral leads to significant noise. To combat this, we use an adaptive vectorized integration-rule cubature scheme implemented in the **CUBATURE** package from Steven Johnson (http://ab-initio.mit.edu/wiki/index.php/Cubature). Using a fixed-rule adaptive routine means the noise in the integrals is controlled such that the numerical derivatives are less noisy. In Figs 14 and 15, we show how accurately the Jeans equations are satisfied along several lines through the potential for our two models. We plot each side of each Jeans equation for a choice of $j$ along a range of lines, along with the percentage error difference between the two sides of the equation. We avoid calculating the derivatives of the moments along the axes as the numerical differentiation is awkward there. In general, we find $\lesssim 10$ per cent error for nearly all tested points with the majority having $\lesssim 4$ per cent over a range of $\sim 8$ orders of magnitude.

Despite the large errors introduced by the action estimation scheme, we have produced a distribution function that satisfies the Jeans equations to reasonable accuracy. Even for the heavily radially biased model, which has large contributions from the box orbits, the Jeans equations are well satisfied. This gives us confidence that
models based on triaxial distribution functions can be constructed using the scheme we have presented.

5.3 Error in the moments

When comparing smooth models to data, we are primarily interested in whether the action-estimation scheme produces accurate enough actions to reproduce the features of the data. If the broad features of the data are on scales larger than the errors in the individual actions, our method should be sufficiently accurate to recover these features from an appropriate \( f(J, \theta) \). In this case, the error in the moments is more important than the error in the individual actions. Larger errors in the actions are expected to lead to larger errors in the moments but the relationship between the two is unclear.

We have seen that, despite the presented method introducing large errors in some actions, the moments of the DF are well recovered such that the normalization is accurate to 0.6 per cent and the Jeans’ equations are accurate to \( \lesssim 4 \) per cent. In this section, we develop some understanding of how errors in the actions translate into errors in the moments. Initially, we can make some progress by considering the normalization of the DF: \( \int d^3x \, d^3v \, f(x, v) = \int d^3J \, d^3\theta \, f(J) \).

Suppose we have a set of angle-action variables \( (J', \theta') \) that are not the true angle-action variables \( (J, \theta) \), we can relate the two sets via the generating function, \( S(J, \theta) \), such that

\[
S(J, \theta) = J \cdot \theta + \sum_n S_n(J) \sin n \cdot \theta.
\]

Now suppose we evaluate the normalization using \( J'(x, v) \). We transform the volume element \( d^3J \, d^3\theta \) to \( d^3J' \, d^3\theta' \) via the Jacobian

\[
\det \left( \frac{\partial \theta'}{\partial \theta} \right)_J = \det \left( 1 + \sum_n n \otimes \partial J_n \cos n \cdot \theta' \right)
\]

\[
\simeq |1 + \sum_n n \cdot \partial J_n \cos n \cdot \theta' |.
\]

We assume that the approximate angle-action variables are sufficiently close to the true angle-action variables that the second term on the right-hand side is much less than unity. Therefore, the
difference in the normalization is given by
\[
\int d^3 J d\theta' \left[ f(J) - f(J') \right] = \int d^3 J d\theta \partial_J f \cdot (J - J')
\]
\[
\cong \int d^3 J d\theta' \left( 1 + \sum_n n \cdot \partial_J S_n \cos n \cdot \theta' \right)
\]
\[
\times \left[ -\sum_m (m \cdot \partial_J J_m) \cos m \cdot \theta' + \ldots \right].
\]
(42)

We see that, after integrating over \(\theta'\), all terms with odd powers of trigonometric functions vanish, and the leading order error in the normalization is
\[
-4\pi^3 \int d^3 J \sum_n (n \cdot \partial_J f)(n \cdot \partial_J S_n) S_n.
\]
This term is second order in the Fourier components of the generating function. Note that the sign of this term is unclear as both the \(S_n\) and \(\partial_J S_n\) can be positive and negative. We anticipate \(\partial_J f\) is negative such that the density falls with radius. From equation (40), we find that to leading order the errors in the angle-action variables averaged over an orbit are
\[
\Delta J_i \leq \sqrt{\frac{2}{3} \sum_n |S_n|} \quad \text{and} \quad \Delta \theta_i \leq \sqrt{\frac{2}{3} \sum_n |\partial_J S_n|}.
\]

Therefore, we can see that the error in the normalization is approximately first order in the error in the actions, \(\Delta J_i\), the error in the angles, \(\Delta \theta_i\), and the gradient of the distribution function, \(\partial_J f\). Therefore, we anticipate that the relative error in the normalization will be small when \(\Delta J \ll (\partial_J f)^{-1}\) for all points in action space. This is essentially the expected result. Consider the distribution functions of Binney (2014); these are of the form \(f(J) \sim \exp(-v J_i/\sigma_i^2)\) such that \(\Delta J_i \ll \sigma_i^2/v\) for a good estimate of the normalization, where \(v\) is the vertical epicycle frequency. Near the Sun \(v \approx 0.1\) Myr\(^{-1}\) and \(\sigma_z \approx 30\) km s\(^{-1}\) so \(\Delta J_i \ll 10\) kpc km s\(^{-1}\). For the distribution function considered in the previous section we require \(\Delta J \ll \max(J, J_0)\).

For the moments of the distribution function, we expect similar results but we are not able to explicitly calculate the leading order errors in these quantities. Instead, we briefly show how the error in the density changes with the errors in the actions. We begin by calculating the density from a triaxial DF for which we know the true density – the perfect ellipsoid (de Zeeuw 1985), which has density profile
\[
\rho(x, y, z) = \frac{\rho_0}{(1 + m^2)^2},
\]
where
\[
m^2 \equiv \frac{x^2}{x_p^2} + \frac{y^2}{y_p^2} + \frac{z^2}{z_p^2}, \quad x_p \geq y_p \geq z_p \geq 0.
\]
(44)

We set \(\rho_0 = 7.2 \times 10^6 M_\odot\) kpc\(^{-3}\), \(x_p = 5.5\) kpc, \(y_p = 4.5\) kpc and \(z_p = 1\) kpc. The actions in this potential can be found exactly using the scheme presented above by setting \(\alpha = -x_0^2\) and \(\beta = -y_0^2\). We use the tangentially biased DF of the previous section. For four different Cartesian positions we calculate the true density, and then proceed to calculate the density when the logarithm of the actions are scattered normally by some fixed amount \(\sqrt{\langle (\Delta J)^2 \rangle}\). We plot the error in the density as a function of \(\Delta J\) in Fig. 16. We find that the relative error in the density goes as \(\sim \Delta J^{3/2}\) for the three densities near the axes, and is flatter for the density at \((x, y, z) = (4, 4, 4)\) kpc.

This procedure is artificial as we have used non-canonical coordinates to evaluate the density. Therefore, for a fuller test we instead choose to calculate the density using the St"uckel fudge scheme but changing \(\alpha\) away from the truth. In this case, the error in the density is systematic. In the lower panel of Fig. 16, we plot the density error from this procedure as a function of the distribution-function-weighted rms action error. Again we find the density goes as \(\sim \sqrt{\langle (\Delta J)^2 \rangle^{1.2}}\) for all but the density on the \(z\)-axis. As we are changing \(\alpha\) this does not significantly affect the actions of the long-axis loop orbits, which dominate the density budget along the \(z\)-axis.

6 EXTENDING THE SCOPE OF TORUS MAPPING

We have seen that the triaxial St"uckel fudge can produce large errors in the actions for some orbits, but that the moments of an action-based DF are nevertheless well recovered using the method. However, even if one requires accurate actions, the St"uckel fudge can be valuable because it enables one to construct a torus through a given point \((x, y)\) by torus mapping rather than orbit integration (Sanders & Binney 2014). This option may be essential because torus mapping works even in a chaotic portion of phase space (Kaasalainen 1995), while the approach based on orbit integration is already
Problematic when resonantly trapped orbits take up a significant portion of phase space, and it breaks down with the onset of chaos.

We proceed as follows. First, we use the Stäckel fudge to obtain approximate actions

\[ J = J_0(x, v). \]

Then by torus mapping we obtain the torus with actions \( J \). On account of errors, the given point \((x, v)\) will not lie on the constructed torus, but one can identify the nearest point \((x', v')\) that does lie on the torus. We find this point by minimizing the tolerance

\[ \eta = |\Omega|^2|\mathbf{x}(\theta) - \mathbf{x}|^2 + |\mathbf{v}(\theta) - \mathbf{v}|^2, \]

with respect to the angle, \( \theta \), on the torus. \( \Omega \) is the frequency vector of the constructed torus. We use the Stäckel fudge estimate of the angles as an initial guess for the minimization. For the point \((x', v')\) we know the true actions:

\[ J = J_0(x', v'). \]

Now we use the Stäckel fudge to obtain approximate actions for this point

\[ J' = J_0(x', v'). \]

If, as we expect, the errors in the Stäckel fudge are systematic rather than random, then

\[ J_0(x, v) = J_0(x, v) + \Delta \]

with \( \Delta \) a slowly varying function of phase-space position. So a better estimate of the true actions of the original point \((x, v)\) is

\[ J'' = J + \Delta = J + (J - J') = 2J - J'. \]

If one is of a nervous disposition, one now uses torus mapping to construct the torus with actions \( J'' \) and seeks the point on this torus that is closest to the given point and applies the Stäckel fudge there, and so on. This cycle can be repeated until the nearest point on the constructed torus satisfies some tolerance \( \eta = \eta_t \).

In Fig. 17, we show an illustration of this procedure for the axisymmetric case. We use the axisymmetric Stäckel fudge as given in Section 3.1 and the torus construction code as presented in McMillan & Binney (2008). For the axisymmetric Stäckel fudge we set \( \gamma = -1 \) kpc\(^2\) and \( \alpha = -20 \) kpc\(^2\), such that the foci are at \( z = \pm \sqrt{\gamma} = \pm 4.4 \) kpc. We construct a torus of actions \((J, L_z, J_z)\) = (244.444, 3422.213, 488.887) kpc km s\(^{-1}\) in the ‘best’ potential from McMillan (2011). This potential is an axisymmetric multicomponent Galactic potential consisting of two exponential discs representing the thin and thick discs, an axisymmetric bulge model from Bissantz & Gerhard (2002) and an NFW dark halo. The parameters of the mass model were chosen to satisfy recent observational constraints. We produce a series of \((x, v)\) points on the constructed torus. The axisymmetric Stäckel fudge gives errors of \( \Delta J \approx 13 \) kpc km s\(^{-1}\) in the recovery of the actions for these points. After one iteration of the above procedure we find actions accurate to \( \Delta J \approx 0.3 \) kpc km s\(^{-1}\), and after the procedure has converged to a tolerance \( \eta_t = (0.1 \) kpc km s\(^{-1}\))^2 the actions are accurate to \( \Delta J \approx 0.07 \) kpc km s\(^{-1}\). The majority of \((x, v)\) points converge in less than five iterations. We limited the number of iterations to 20 and a few \((x, v)\) did not converge within 20 iterations. This is due to non-linear behaviour of both the torus construction and Stäckel code in small phase-space volumes. It is clear, however, that only one iteration is required for a substantial improvement in the actions, and it is hard to make a case for more iterations.

Finally, we demonstrate how the above method operates for a range of high-action orbits. We integrate a series of orbits in the

**Figure 17.** Illustration of the iterative torus scheme presented in Section 6. The black points show the initial estimate of the actions for a series of points along an orbit calculated using the axisymmetric Stäckel fudge of Binney (2012a). The dark blue points show the improved estimate from a single iteration of the torus scheme, and the light green points show the estimate after the scheme was deemed to converge to an accuracy of \( \eta_t = (0.1 \) kpc km s\(^{-1}\))^2. The top-right inset shows a histogram of the number of iterations required to reach this accuracy. Above the plot we show the standard deviation of the action estimates for each set of points. The black dashed lines give the true action estimate. The bottom-left inset shows a zoom-in of the central region of the main plot.

‘best’ potential from McMillan (2011) launched at five linearly spaced points \( x_0 \) along the x-axis such that \( 4 \) kpc \( \leq \) \( x_0 \) \( \leq \) 12 kpc. We launch the orbits with velocity \( v = (v_1 \cos \theta, v_1, v_1 \sin \theta) \) where \( v_1 = \sqrt{v_{10}^2 + \Phi(x_0, 0, 0)} \) and we choose four linearly spaced values of \( v_1 \) such that \( 0.5v_{10} \leq v_1 \leq 0.8v_{10} \) and 5 linearly spaced angles \( \theta \) such that \( 0.2 \) rad \( \leq \theta \leq \frac{\pi}{2} \) rad. The range of radial and vertical actions for this collection of orbits is shown in the top panel of Fig. 18 and is approximately 1 kpc km s\(^{-1}\) \( \leq J \leq \) 800 kpc km s\(^{-1}\) and 1 kpc km s\(^{-1}\) \( \leq J_z \leq \) 800 kpc km s\(^{-1}\) (these are calculated as the mean of the fudge estimates along the orbit). For each orbit, we find the standard deviation of the action estimates from the axisymmetric Stäckel fudge method and the iterative torus method for 10 widely time-separated phase-space points along the orbit. We use \( \eta_t = (0.1 \) kpc km s\(^{-1}\))^2, limit the maximum number of iterations to five and construct the tori with a relative error of \( 1 \times 10^{-4} \). We plot the results in Fig. 18. We see the majority of orbits have lower iterative torus errors than fudge errors and follow a broad line that lies approximately two to three orders of magnitude beneath the 1:1 line. However, there are several orbits that lie close to the 1:1 line indicating the procedure has not converged to a greater accuracy than the initial accuracy produced by the Stäckel fudge. These orbits are either near-resonant so require more careful action assignment (Kaasalainen 1995), or have one action very much greater than the other (near-radial or shell orbits) so require more accurate torus
actions of a selection of orbits shown in the top panel, along with the absolute error in the action for these orbits from the axisymmetric Stäckel fudge of Binney (2012a) plotted against the absolute error from the iterative torus procedure for a range of orbits in the bottom panel. The black dots show the results for the radial action, whilst the red crosses show the results for the vertical action. The blue dotted line is the 1:1 line.

Figure 18. Actions of a selection of orbits shown in the top panel, along with the absolute error in the action for these orbits from the axisymmetric Stäckel fudge of Binney (2012a) plotted against the absolute error from the iterative torus procedure for a range of orbits in the bottom panel. The black dots show the results for the radial action, whilst the red crosses show the results for the vertical action. The blue dotted line is the 1:1 line.

7 CONCLUSIONS

We have presented a method for estimating the actions in a general triaxial potential using a Stäckel approximation. The method is an extension of the Stäckel fudge introduced by Binney (2012a) for the axisymmetric case. We have investigated the accuracy of the method for a range of orbits in an astrophysically relevant triaxial potential. We have seen that the recovery of the actions is poorest for the box orbits, which probe a large radial range of the potential, and much better for the loop orbits, which are confined to a more limited radial range. The only parameters in the method are the choice of the focal positions \( \Delta_x \), which are selected for each input phase-space point. We have detailed a procedure for selecting these based on the energy of the input phase-space point. This choice is not optimal but, by adjusting \( \Delta_x \), we can, at best, increase the accuracy of the actions of a factor of 2 for the triaxial NFW potential considered. However, to achieve this accuracy requires additional computation for each input phase-space point (e.g. orbit integration). For general potentials the best action estimates will be achieved when locally (over the region a given orbit probes) the potential is well approximated by some Stäckel potential. Many potentials of interest are not well fitted globally by Stäckel potentials so the accuracy of the action estimates will deteriorate for orbits with large radial actions.

The advantage of this method over other methods for estimating the actions in a triaxial potential is speed. Unlike the convergent method introduced by Sanders & Binney (2014), we obtain the actions without integrating an orbit – we only use the initial phase-space point. We have only to evaluate several algebraic expressions, find the limits of the orbits in the \( \tau \) coordinate and perform Gaussian quadrature. These are all fast calculations. However, this speed comes at the expense of sometimes disappointing accuracy. If accurate results are required, the Stäckel fudge can be combined with torus mapping to form a rapidly convergent scheme for the determination of \( \mathbf{J}(x, v) \). We demonstrated how such a scheme performed in the axisymmetric case and found a single torus construction provided a high level of accuracy that is not significantly improved by further torus constructions.

We went on to construct, for the first time, triaxial stellar systems from a specified DFs \( f(\mathbf{J}) \) in Section 5. We demonstrated the mass of these models is well recovered using the Stäckel fudge, and we showed how the error in the density of these models varies as a function of the action error. Notwithstanding the errors in individual actions, both a radially biased model and a tangentially biased model satisfy the Jeans equations to good accuracy. This is because individual errors largely cancel during integration over velocities when computing moments such as the density \( \rho(x) \) and the pressure tensor \( \rho \sigma_{ij}^2(x) \).

The results presented in this paper have focused on a limited range of astrophysically relevant models: we have used a single specific NFW potential and two simple distribution functions that depend on a linear sum of the actions. However, we anticipate that our results will extend to more general distribution functions. We have investigated analytically how the normalization of the distribution function varies with the error in the action estimates and shown that the normalization is well recovered provided the error in the actions is smaller than the action scale over which the distribution function varies significantly, i.e. \( \Delta J \ll f(\mathbf{J})^{-1} \). Therefore, the recovery of the moments is expected to be most accurate for distribution functions with shallow radial density profiles and to deteriorate with the steepness of the required profile.

Whilst the scheme presented here does not give accurate enough actions for working with streams (Sanders 2014a), we have shown...
that it is an appropriate and powerful tool for constructing models from specified DFs \( f(J) \). A key property of DFs of the form \( f(J) \) is that they can be trivially added to build up a multicomponent system. Hence the ability to extract observables from DFs of the form \( f(J) \) is likely to prove extremely useful for interpreting data on both external galaxies (Cappellari et al. 2011) and our Galaxy, in which components such as the stellar and dark haloes may be triaxial, and the bulge certainly is.

**ACKNOWLEDGEMENTS**

We thank Paul McMillan for provision of the torus construction machinery used in Section 6 and the Oxford Galactic Dynamics group for helpful comments. JLS acknowledges the support of STFC. JB was supported by STFC by grants R22138/GA001 and ST/K00106X/1. The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 321067.

**REFERENCES**

Aarseth S. J., Binney J., 1978, MNRAS, 185, 227
Allgood B., Flores R. A., Primack J. R., Kravtsov A. V., Wechsler R. H., Faltenbacher A., Bullock J. S., 2006, MNRAS, 367, 1781
Bailin J. et al., 2005, ApJ, 627, L17
Binney J., 2010, MNRAS, 401, 2318
Binney J., 2012a, MNRAS, 426, 1324
Binney J., 2012b, MNRAS, 426, 1328
Binney J., 2014, MNRAS, 440, 787
Binney J., Spergel D., 1984, MNRAS, 206, 159
Binney J., Tremaine S., 2008, Galactic Dynamics, 2nd edn. Princeton Univ.
Press, Princeton, NJ
Bissantz N., Gerhard O., 2002, MNRAS, 330, 591
Cappellari M. et al., 2011, MNRAS, 413, 813
de Zeeuw T., 1985, MNRAS, 216, 273
Eyre A., Binney J., 2011, MNRAS, 413, 1852
Hahn T., 2005, Comput. Phys. Commun., 168, 78
Helmi A., White S. D. M., 1999, MNRAS, 307, 495
Jing Y. P., Suto Y., 2002, ApJ, 547, 538
Kaasalainen M., 1995, Phys. Rev. E., 52, 1193
Kaasalainen M., Binney J., 1994, MNRAS, 268, 1033
Kazantzidis S., Kravtsov A. V., Zennner A. R., Allgood B., Nagai D., Moore B., 2004, ApJ, 611, L73
Law D. R., Majewski S. R., 2010, ApJ, 714, 229
McMillan P. J., 2011, MNRAS, 414, 2446
McMillan P. J., Binney J., 2008, MNRAS, 390, 429
Merritt D., Valluri M., 1999, AJ, 118, 1777
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Ollongren A., 1962, Bull. Astron. Inst. Neth., 16, 241
Pfiff T. et al., 2014, MNRAS, 445, 3133
Poincaré H., 1892, Les méthodes nouvelles de la mécanique céleste, Gauthier-Villars, Paris
Posti L., Binney J., Nicotra C., Ciotti L., 2014, preprint (arXiv:1407.1020)
Prendergast K. H., Tomer E., 1970, AJ, 75, 674
Rowley G., 1988, ApJ, 331, 124
Sanders J., 2012, MNRAS, 426, 128
Sanders J. L., 2014a, MNRAS, 443, 423
Sanders J. L., 2014b, DPhil thesis, Univ. Oxford
Sanders J. L., Binney J., 2013, MNRAS, 433, 1826
Sanders J. L., Binney J., 2014, MNRAS, 441, 3284
Schönrich R., Binney J., 2009, MNRAS, 396, 203
Schwarzschild M., 1979, ApJ, 232, 236
Tremaine S., 1999, MNRAS, 307, 877
Valluri M., Debattista V. P., Quinn T., Moore B., 2010, MNRAS, 403, 525
Vera-Ciro C., Helmi A., 2013, ApJ, 773, L4
Vera-Ciro C. A., Sales L. V., Helmi A., Frenk C. S., Navarro J. F., Springel V., Vogelsberger M., White S. D. M., 2011, MNRAS, 416, 1377
Wilson C. P., 1975, AJ, 80, 175

**APPENDIX A: ANGLES AND FREQUENCIES**

With the framework presented in Section 3 we are also in a position to find the angles, \( \theta \), and frequencies, \( \Omega \). Following de Zeeuw (1985) we write

\[
\frac{\partial E}{\partial \tau} = \frac{1}{\Omega_{1}} \frac{\partial J_{1}}{\partial \tau},
\]

\[
\frac{\partial E}{\partial a} = 0 + \sum_{\mu=2}^{\mu} \Omega_{\mu} \frac{\partial J_{\mu}}{\partial a},
\]

\[
\frac{\partial E}{\partial b} = 0 + \sum_{\mu=2}^{\mu} \Omega_{\mu} \frac{\partial J_{\mu}}{\partial b}.
\]

(A1)

Inversion of these equations gives, for instance,

\[
\Omega_{\lambda} = \frac{1}{\Gamma} \frac{\partial (J_{\mu}, J_{\nu})}{\partial (a, b)},
\]

where \( \Gamma = \frac{\partial (J_{\mu}, J_{\nu}, J_{\lambda})}{\partial (E, a, b)} \).

(A2)

\( \Omega_{\mu} \) and \( \Omega_{\nu} \) are given by cyclic permutation of \( \{\lambda, \mu, \nu\} \). To find the derivatives of \( J_{\mu} \) with respect to the integrals we differentiate equation (13) under the integral sign at constant \( \tau \). From equation (17), we know \( p_{\tau}(\tau, E, A, B_{\tau}) \). We note that

\[
\frac{\partial}{\partial \tau} \left|_{\tau} \right. = \frac{\partial A_{\tau}}{\partial a} \frac{\partial}{\partial a} + \frac{\partial A_{\tau}}{\partial b} \frac{\partial}{\partial b},
\]

\[
\frac{\partial}{\partial \tau} \left|_{\tau} \right. = \frac{\partial B_{\tau}}{\partial a} \frac{\partial}{\partial a} + \frac{\partial B_{\tau}}{\partial b} \frac{\partial}{\partial b}.
\]

(A3)

as \( A_{\tau} = a - C_{\tau} \) and \( B_{\tau} = b + D_{\tau} \), where \( C_{\tau} \) and \( D_{\tau} \) are independent of \( \tau \). The required derivatives are

\[
\frac{\partial p_{\tau}}{\partial \tau} \left|_{\tau} \right. = \frac{\tau^{2}}{4p_{\tau}(\tau + \alpha)(\tau + \beta)(\tau + \gamma)},
\]

\[
\frac{\partial p_{\tau}}{\partial a} \left|_{\tau} \right. = - \frac{1}{\tau} \frac{\partial p_{\tau}}{\partial E} \left|_{\tau} \right.,
\]

\[
\frac{\partial p_{\tau}}{\partial b} \left|_{\tau} \right. = \frac{1}{\tau^{2}} \frac{\partial}{\partial E} \left|_{\tau} \right.,
\]

(A4)

Note that \( p_{\tau} \) can vanish at the limits of integration. The change of variables

\[
\tau = \tau \sin \theta + \tilde{\tau}; \; \tilde{\tau} = \frac{1}{2}(\tau_{+} + \tau_{-}); \; \tau = \frac{1}{2}(\tau_{+} - \tau_{-})
\]

causes the integrand to go smoothly to zero at the limits. To find the angles, we use the generating function, \( W(\lambda, \mu, \nu, J_{\mu}, J_{\mu}, J_{\nu}) \), given by

\[
W = \sum_{\tau=1}^{\tau} W_{\tau} = \sum_{\tau=1}^{\tau} \int_{\tau_{-}}^{\tau_{+}} d\tau' p_{\tau}' + F_{\tau}(p_{\tau}, x) \int_{\tau_{-}}^{\tau_{+}} d\tau' |p_{\tau}'|.
\]

(A5)

\( F_{\tau} \) are factors included to remove the degeneracy in the \( \tau \) coordinates such that \( \theta_{\tau} \) covers the full range 0 to \( 2\pi \) over one oscillation in the Cartesian coordinates. These factors can be written in the form

\[
F_{\tau}(p_{\tau}, x) = \Pi(\lambda_{\tau} + \alpha)\Theta(-x) + \Theta(-p_{\tau}),
\]

\[
F_{\mu}(p_{\mu}, x) = \Pi(\mu_{\tau} + \beta)\Theta(-y) + \Theta(-p_{\mu}).
\]
The angles calculated using the triaxial St"ackel fudge presented in this paper for three different orbits in the triaxial NFW potential. The solid red lines show the angles calculated from the initial angle estimate and the frequency estimates for approximately one period.

\[ + \Pi(v_{\pm} + \beta) \Pi(\mu_{\pm} + \alpha) \left[ \frac{\tau}{2\pi} + \Theta(-x) \right] \]
\[ + \Pi(v_{\pm} + \beta) \Pi(\lambda_{\pm} + \alpha) \Theta(-p_{\mu}) \]
\[ \mathcal{F}_{i}(p_{\nu}, x) = \Theta(-z) + \Theta(-p_{\nu}) \tag{A6} \]

where \( \Theta \) is the Heaviside step function and \( \Pi \) is one when its argument is zero and zero otherwise. The \( \Pi \)-function in \( \mathcal{F}_{i} \) takes care of the cases when the orbit is a box or inner long-axis loop.

The \( \Pi \)-functions in \( \mathcal{F}_{\mu} \) take care of the cases when the orbit is a short-axis loop or a box, an outer long-axis loop, and an inner long-axis loop, respectively. The angles are given by

\[ \theta_{i} = \frac{\partial W}{\partial I_{i}} = \sum_{l=E,A,B} \frac{\partial W}{\partial I_{l}} \frac{\partial I_{l}}{\partial J_{i}} \tag{A7} \]

The first term on the right is, up to factors, the indefinite integral of the derivatives of \( J_{l} \) with respect to the integrals found previously, whilst the second term is found from inverting these derivatives. We have chosen the zero-point of \( \theta_{i} \) to correspond to \( \tau = \tau_{0}, p_{\tau} > 0 \) and \( \dot{x}_{i} \geq 0 \) for all \( i \), except for the outer long-axis loop orbits which have \( \theta_{\mu} = 0 \) at \( \mu = -\alpha, p_{\tau} > 0 \) and \( \dot{x}_{i} \geq 0 \). Note that the angles are the \( 2\pi \) modulus of the \( \theta_{i} \) found from the above scheme.

In Fig. A1, we show the angles calculated from the St"ackel fudge for the three orbits investigated in Section 4.2. We use the automatic choice of \( \Delta_{1} \) for the box and short-axis loop orbit, and the choice that minimizes the spread in actions for the long-axis loop orbit. The short-axis loop orbit shows the expected straight-line structure in the angle coordinates, whilst for the long-axis loop and box orbits there is clear deviation from this expected straight line. We also show the angles calculated using the initial angle estimate and the average of the frequency estimates along the orbit. We see that they are well recovered but after approximately one period the error in the frequencies is sufficient for these angles to deviate from the angle estimates.

The standard deviations in the frequencies are reasonably large. For the box orbit, the mean frequencies are given by \( \Omega = (18.1, 20.3, 24.3) \text{ kpc}^{-1} \text{ km s}^{-1} \) with errors \( \Delta \Omega = (0.2, 0.8, 1.3) \text{ kpc}^{-1} \text{ km s}^{-1} \). For the short-axis loop the mean frequencies are given by \( \Omega = (34.9, 21.9, 25.0) \text{ kpc}^{-1} \text{ km s}^{-1} \) with errors \( \Delta \Omega = (0.6, 0.1, 0.2) \text{ kpc}^{-1} \text{ km s}^{-1} \). For the long-axis loop the mean frequencies are given by \( \Omega = (36.9, 22.4, 24.0) \text{ kpc}^{-1} \text{ km s}^{-1} \) with errors \( \Delta \Omega = (1.6, 1.5, 0.8) \text{ kpc}^{-1} \text{ km s}^{-1} \). We note that the frequency errors are largest at the turning points of the orbits for the loop orbits or near the centre of the potential for the box orbit.