Action-based distribution functions for spheroidal galaxy components

Lorenzo Posti,[1][*] James Binney,[2] Carlo Nipoti[1] and Luca Ciotti[1]

[1]*Dipartimento di Fisica e Astronomia, Università di Bologna, viale Berti-Pichat 6/2, I-40127 Bologna, Italy*
[2]*Rudolf Peierls Centre for Theoretical Physics, Keble Road, Oxford OX1 3NP, UK*

ABSTRACT

We present an approach to the design of distribution functions that depend on the phase-space coordinates through the action integrals. The approach makes it easy to construct a dynamical model of a given stellar component. We illustrate the approach by deriving distribution functions that self-consistently generate several popular stellar systems, including the Hernquist, Jaffe, and Navarro, Frenk and White models. We focus on non-rotating spherical systems, but extension to flattened and rotating systems is trivial. Our distribution functions are easily added to each other and to previously published distribution functions for discs to create self-consistent multicomponent galaxies. The models this approach makes possible should prove valuable both for the interpretation of observational data and for exploring the non-equilibrium dynamics of galaxies via *N*-body simulations.

Key words: galaxies: kinematics and dynamics – galaxies: structure – dark matter.

1 INTRODUCTION

Axisymmetric equilibrium models are extremely useful tools for the study of galaxies. A real galaxy will never be in perfect dynamical equilibrium – it might be accreting dwarf satellites, or being tidally disturbed by the gravitational field of the group or cluster to which it belongs, or displaying spiral structure – but an axisymmetric equilibrium model will usually provide a useful basis from which a more realistic model can be constructed by perturbation theory.

By Jeans (1915) theorem, every equilibrium model can be described by a distribution function (DF) that depends on the phase-space coordinates $(\boldsymbol{x}, \boldsymbol{v})$ only through isolating integrals of motion. In an axisymmetric potential, most orbits prove to be quasi-periodic, with the consequence that they admit three isolating integrals (Arnold 1978). Consequently, a generic DF for an axisymmetric equilibrium galaxy is a function of three variables.

The major obstacle to exploiting this insight is that we have analytic expressions for only two isolating integrals of motion in a general axisymmetric potential, namely the energy $E = \frac{1}{2}v^2 + \Phi(\boldsymbol{x})$ and the component of the angular momentum about the symmetry axis, $J_\phi = (\boldsymbol{x} \times \boldsymbol{v})_z$. Several authors have examined model galaxies with DFs of the two-integral form $f(E, J_\phi)$ (Prendergast & Tomer 1970; Wilson 1975; Rowley 1988; Evans 1994), but in such models the velocity dispersions σ_R and σ_z in the radial and vertical directions are inevitably equal. This condition is seriously violated in our Galaxy and we have no reason to suppose that the condition is better satisfied in any external galaxy. Hence, it is mandatory to

extend the DF's argument list to include a 'non-classical' integral, I_3, for which we do not have a convenient expression.

Since any function $J(E, J_\phi, I_3)$ of three isolating integrals is itself an isolating integral, we actually have an enormous amount of freedom as to what integrals to use as arguments of the DF. Given that we must use at least one integral for which we lack an expression for its dependence on $(\boldsymbol{x}, \boldsymbol{v})$, there is a powerful case for making the DF's arguments action integrals. These integrals are alone capable as serving as the three momenta J_i of a canonical coordinate system – this property makes them the bedrock of perturbation theory. Their canonically conjugate variables, the angles θ_i, have two remarkable properties: (i) along any orbit they increase linearly with time at rates $\Omega_i(\boldsymbol{J})$, so

$$\theta_i(t) = \theta_i(0) + \Omega_i(\boldsymbol{J})\,t, \tag{1}$$

and (ii) they make the ordinary phase-space coordinates periodic functions

$$\boldsymbol{x}(\boldsymbol{\theta} + 2\pi\boldsymbol{m}, \boldsymbol{J}) = \boldsymbol{x}(\boldsymbol{\theta}, \boldsymbol{J}) \quad \text{(integer } m_i\text{).} \tag{2}$$

The actions J_i also have nice properties. In particular, (i) any triple of finite numbers (J_r, J_ϕ, J_z) with $J_r, J_z \geq 0$ corresponds to a bound orbit with the orbit $\boldsymbol{J} = 0$ being that on which a star is stationary at the middle of the galaxy, and (ii) the volume of phase space occupied by orbits with actions in $\mathrm{d}^3\boldsymbol{J}$ is $(2\pi)^3\mathrm{d}^3\boldsymbol{J}$. Consequently, any non-negative function $f(\boldsymbol{J})$ that tends to zero as $|\boldsymbol{J}| \to \infty$ and has a finite integral $\int \mathrm{d}^3\boldsymbol{J}\, f(\boldsymbol{J})$ specifies a valid galaxy model of mass

$$M = (2\pi)^3 \int \mathrm{d}^3\boldsymbol{J}\, f(\boldsymbol{J}). \tag{3}$$

[*] E-mail: lorenzo.posti@unibo.it

The actions are defined by integrals

$$J_i = \frac{1}{2\pi} \oint_{\gamma_i} d\boldsymbol{x} \cdot \boldsymbol{v}, \tag{4}$$

where γ_i is a closed path in phase space. If we require that the first action J_r quantifies the extent of a star's radial excursions and the third action J_z quantifies the extent of its excursions either side of the potential's equatorial plane, then the actions are unambiguously defined. What we here call J_r is sometimes called J_R or J_u, and what we call J_z is sometimes called J_ϑ or J_v, but no significance attaches to these different notations. In a spherical potential $J_z = L - |J_\phi|$, where L is the magnitude of the angular momentum vector.

To obtain the observable properties of a model defined by $f(\boldsymbol{J})$, for example its density distribution $\rho(\boldsymbol{x}) = \int d^3\boldsymbol{v} \, f(\boldsymbol{J})$ and its velocity dispersion tensor $\sigma_{ij}^2(\boldsymbol{x})$, one has to be able to evaluate $\boldsymbol{J}(\boldsymbol{x}, \boldsymbol{v})$ in an arbitrary gravitational potential. Recently, a number of techniques have been developed for doing this (Binney 2012a; Sanders & Binney 2014, 2015). Consequently, while the last word on action evaluation has likely not yet been written, we now have algorithms that enable one to extract the observables from a DF $f(\boldsymbol{J})$ with reasonable accuracy.

DFs $f(\boldsymbol{J})$ that depend on the phase-space coordinates only through the actions were first used to model the disc of our Galaxy in an assumed gravitational potential (Binney 2010, 2012b). Recently, Binney (2014, hereafter B14) showed how to derive the self-consistent gravitational potential that is implied by a given $f(\boldsymbol{J})$ by exploring a family of flattened, rotating models that he derived from the 'ergodic' DF of the isochrone model: that is the DF $f(H)$ that depends on the phase-space coordinates only through the Hamiltonian $H = \frac{1}{2}v^2 + \Phi(\boldsymbol{x})$. Hénon (1960) derived the isochrone's ergodic DF, and in the case of the isochrone potential explicit expressions are available for $\boldsymbol{J}(\boldsymbol{x}, \boldsymbol{v})$ and $H(\boldsymbol{J})$ (Gerhard & Saha 1991). Substituting $H(\boldsymbol{J})$ in $f(H)$ B14 obtained the DF $f(\boldsymbol{J})$ of the isotropic isochrone model. In this paper, we present simple analytic functions $f(\boldsymbol{J})$ that generate nearly isotropic models of other widely used models, such as the Hernquist (1990), Jaffe (1983), and Navarro, Frenk & White (1996, hereafter NFW) models.

Once a DF of the form $f(\boldsymbol{J})$ is available for a spherical, non-rotating model, the procedure B14 used to flatten the isochrone sphere and to set it rotating can be used to flatten and/or set rotating one's chosen model. So DFs for spherical models in the form $f(\boldsymbol{J})$ are valuable starting points from which quite general axisymmetric models are readily constructed.

Galaxies are generally considered to consist of a number of components, such as a disc, a bulge, and a dark halo, that cohabit a single gravitational potential. If we represent each component by a DF of the form $f(\boldsymbol{J})$, it is straightforward to find the gravitational potential in which they are all in equilibrium (e.g. Piffl et al. 2014; Piffl, Penoyre & Binney in preparation). An analogous composition using DFs of the form $f(E, J_\phi, I_3)$ has never been achieved and may be impossible, because when components are added, their potentials must be added, and the energies of physically similar orbits in a given component are quite different before and after we add in the potential of another component. For example, the orbit on which a star sits at the centre of the galaxy will have different energies before and after addition. If E is used as an argument of the DF, the change in E will change the density of stars on the given orbit, which is contrary to the fundamental idea of building up the galaxy by adding components. By contrast, the actions of the orbit on which a star sits at the galactic centre vanish in any potential, and if a component is defined by $f(\boldsymbol{J})$, it contributes the same density of stars to this orbit regardless of the external potential in which that component finds

itself. This fact is a major motivation for discovering what DF of the form $f(\boldsymbol{J})$ is required to generate each component of a galaxy.

The DF of an isotropic spherical model must depend on the actions only via the Hamiltonian $H(\boldsymbol{J})$. The dependence of f on H is readily obtained from the inversion formula of Eddington (1916), but an exact expression for $H(\boldsymbol{J})$ is only available for the isochrone potential and its limiting cases, the harmonic oscillator and Kepler potentials. Our ignorance of $H(\boldsymbol{J})$ for potentials other than the isochrone amounts to a barrier to the extension of B14's approach to model building. One way to break through this barrier is to devise numerical approximations to $H(\boldsymbol{J})$ and some success has been had in this direction by Fermani (2013) and Williams, Evans & Bowden (2014). In this paper, we pursue a slightly different strategy, which is to develop simple algebraic expressions for DFs $f(\boldsymbol{J})$ that generate self-consistent models that closely resemble popular spherical systems. We also show that a very simple form of $f(\boldsymbol{J})$ generates a model that is almost identical to the isochrone sphere and we give a useful analytic expression for the radial action as a function of energy and angular momentum for a Hernquist sphere.

The paper is organized as follows. In Section 2, we use analytic arguments to infer $f(\boldsymbol{J})$ for scale-free models. These models are not physically realizable as they stand, so in Section 3 we consider models that consist of two power-law sections joined at a break radius. In Section 4, we extract realizable models from scale-free models by the alternative strategy of adding a core to the system and/or tidally truncating the model. Section 5 sums up.

2 POWER-LAW MODELS

Consider a gravitational potential that scales as a power of the distance from the galactic centre, i.e. $\Phi(\xi\boldsymbol{x}) \propto \xi^a \Phi(\boldsymbol{x})$ with $a \neq 0$: in the limit $a \to 0$ the gravitational potential tends to a logarithmic potential, which is an interesting special case that we will treat in Section 2.1.

An orbit in a power-law potential has time-averaged kinetic and potential energies, K and W, respectively, that are related by the virial theorem: $2K = aW$. The instantaneous total energy, given by the sum of the instantaneous kinetic and potential energies, is conserved along the orbit and consequently is given by

$$E = K + W = \left(\frac{a}{2} + 1\right) W. \tag{5}$$

In any power-law potential, we need only to study orbits of one arbitrarily chosen energy E because each of these orbits can be rescaled to a similar orbit at any given energy E'. Indeed, if an orbit is rescaled by a spatial factor, i.e. $\boldsymbol{x} \to \boldsymbol{x}' = \xi\boldsymbol{x}$, then the orbit's total energy scales as

$$E \to E' = \xi^a E, \tag{6}$$

since obviously $W \to W' = \xi^a W$. Further $v^2 \propto K = \frac{1}{2}aW$, so under rescaling $\boldsymbol{v} \to \boldsymbol{v}' = \xi^{a/2}\boldsymbol{v}$.

Given the scalings derived above for $\boldsymbol{x}$ and $\boldsymbol{v}$ it follows that

$$\boldsymbol{J} \to \boldsymbol{J}' = \xi^{1+a/2}\boldsymbol{J}. \tag{7}$$

Thus, both the energy and the actions of an orbit that is rescaled by the spatial factor ξ are rescaled by powers of this factor.

From equations (6) and (7) we deduce that the Hamiltonian is of the form

$$H(\boldsymbol{J}) = [h(\boldsymbol{J})]^{a/(1+a/2)}, \tag{8}$$

where $h(\boldsymbol{J})$ is a homogeneous function of degree one, i.e. $h(\zeta\boldsymbol{J}) = \zeta h(\boldsymbol{J})$ for every constant ζ. In particular, H is itself a homogeneous

function of the three actions of degree $a/(1 + a/2)$. It is easy to check that equation (8) gives the correct scalings $H \propto |\boldsymbol{J}|$ and $H \propto |\boldsymbol{J}|^{-2}$ for the harmonic oscillator ($a = 2$) and Kepler ($a = -1$) potentials. Williams et al. (2014) derive a closely related result in which a specific form is proposed for $h(\boldsymbol{J})$.

The homogeneous function h is strongly constrained by the orbital frequencies. Indeed

$$\frac{\Omega_i}{\Omega_j} = \frac{\partial H / \partial J_i}{\partial H / \partial J_j} = \frac{\partial h / \partial J_i}{\partial h / \partial J_j}. \tag{9}$$

In a scale-free model, the frequency ratio on the left is a homogeneous function of degree zero, i.e. scale independent, in agreement with the right-hand side. A natural choice for h that we will use extensively is

$$h(\boldsymbol{J}) = J_r + \frac{\Omega_\phi(\boldsymbol{J})}{\Omega_r(\boldsymbol{J})} |J_\phi| + \frac{\Omega_z(\boldsymbol{J})}{\Omega_r(\boldsymbol{J})} J_z. \tag{10}$$

In a scale-free model, this is homogeneous of degree one, as required. Moreover, so long as the frequency ratios do not change rapidly within a surface of constant energy in action space, the derivatives of h satisfy equation (9) to good precision.

In the definition (10) of $h(\boldsymbol{J})$, the modulus of the angular momentum J_ϕ appears because we are concerned with the construction of the part of the DF that is even in J_ϕ. If we wish to set the model rotating, we will add to this even part an odd part as discussed by B14.

Consider now the density distribution that generates a power-law potential. In the spherical case,[1] we have

$$\frac{\mathrm{d}\Phi(r)}{\mathrm{d}r} = \left.\frac{\partial \Phi(\xi r)}{r \partial \xi}\right|_{\xi=1} = \frac{a}{r} \Phi(r). \tag{12}$$

Hence

$$4\pi G \rho = \frac{1}{r^2} \frac{\mathrm{d}}{\mathrm{d}r} \left(r^2 \frac{\mathrm{d}\Phi}{\mathrm{d}r} \right) = \frac{1}{r^2} \frac{\mathrm{d}}{\mathrm{d}r} (ar\Phi) = \frac{a + a^2}{r^2} \Phi. \tag{13}$$

If $a = -1$, we recover the expected result $\rho = 0$, but for $a \neq 0$ we obtain the polytropic relation for index $n = 1 - 2/a$ (e.g. Binney & Tremaine 2008, section 4.3.3a):

$$\rho \propto |\Phi|^{1-2/a}. \tag{14}$$

From this relation, it is easy to derive the ergodic DF

$$f(E) \propto E^{-(4+a)/2a} \tag{15}$$

from Eddington's formula (e.g. Evans 1994). From equations (8) and (15) it follows that the DF of a power-law model is

$$f(\boldsymbol{J}) = [h(\boldsymbol{J})]^{-(4+a)/(2+a)}. \tag{16}$$

The DF of a power-law model is itself a power law of the three actions and the exponent is completely determined by that of $\Phi(\boldsymbol{x})$.

---

[1] In the non-spherical case,

$$\begin{aligned} 4\pi G \rho(\xi \boldsymbol{r}) &= \frac{a + a^2}{\xi^2 r^2} \Phi(\xi \boldsymbol{r}) + \frac{1}{\xi^2 r^2 \sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial \Phi(\xi \boldsymbol{r})}{\partial\theta} \right) \\ &\quad + \frac{1}{\xi^2 r^2 \sin^2\theta} \frac{\partial^2 \Phi(\xi \boldsymbol{r})}{\partial\phi^2} \\ &= 4\pi G \xi^{a-2} \rho(\boldsymbol{r}). \end{aligned} \tag{11}$$

Consequently, ρ and Φ have simple scalings with r but they are not necessarily functions of each other.

2.1 Logarithmic potentials

Now consider the limit $a \to 0$ when the scaling of Φ becomes additive

$$\Phi(\xi \boldsymbol{x}) = \Phi(\boldsymbol{x}) + v_c^2 \log(\xi), \tag{17}$$

where v_c is a constant that one can easily show is the circular speed. Since galaxies have quite flat circular speed curves, potentials of this form are very useful.

The kinetic energy K does not change on rescaling, while the potential energy $W \to W' = W + v_c^2 \log(\xi)$, so

$$E \to E' = E + v_c^2 \log(\xi). \tag{18}$$

The invariance of K implies invariance of $\boldsymbol{v}$ under orbit rescaling, so the scaling of the actions is

$$\boldsymbol{J} \to \boldsymbol{J}' = \xi \boldsymbol{J} = \exp\left(\frac{E' - E}{v_c^2} \right) \boldsymbol{J}. \tag{19}$$

We now use each side of this equation as the argument of a homogeneous function of degree one, $h(\boldsymbol{J})$, and obtain

$$h(\boldsymbol{J}') = \exp\left(\frac{E' - E}{v_c^2} \right) h(\boldsymbol{J}), \tag{20}$$

or on rearrangement

$$E' = E + v_c^2 \log[h(\boldsymbol{J}')/h(\boldsymbol{J})]. \tag{21}$$

Here E' and E are the energies of any two orbits whose actions $\boldsymbol{J}'$ and $\boldsymbol{J}$ are proportional to each other. We can choose to make $\boldsymbol{J}$ an orbit with vanishing energy, and we can choose h to be the homogeneous function that satisfies $h(\boldsymbol{J}) = 1$ as $\boldsymbol{J}$ moves over the surface $E = 0$ in action space. With these choices, we have

$$H(\boldsymbol{J}') = v_c^2 \log[h(\boldsymbol{J}')]. \tag{22}$$

The ergodic DF that self-consistently generates the spherical logarithmic potential is well known to be

$$f(H) = \exp\left(\frac{E_0 - H}{\sigma^2} \right), \tag{23}$$

where $\sigma^2 = v_c^2/2$ and E_0 is a constant (e.g. Binney & Tremaine 2008, section 4.3.3b). Using equation (22) it follows that the ergodic DF is

$$f(\boldsymbol{J}) = \text{constant} \times [h(\boldsymbol{J})]^{-2}. \tag{24}$$

This result is consistent with the limit $a \to 0$ of equation (16) for a power-law model.

Note that equation (24) implies that the phase-space density diverges as $\boldsymbol{J} \to 0$. It follows that this DF unambiguously specifies the singular isothermal sphere, in contrast to the DF (23), from which one can derive both cored and singular isothermal spheres (e.g. Binney & Tremaine 2008, section 4.3.3b). It is characteristic of DFs of the form $f(\boldsymbol{J})$ that they uniquely and transparently specify the phase-space density both at the centre of the model ($\boldsymbol{J} = 0$) and for marginally bound orbits ($\boldsymbol{J} \to \infty$). From a DF that depends on energy, by contrast, the phase-space density at the centre of the model is implicitly specified by the boundary condition adopted at $r = 0$ when solving Poisson's equation for the self-consistent potential.

The considerations of the last paragraph apply equally to the power-law DFs (16): although we used the standard form (15) of the energy-based DF of the polytropes to derive this DF, it implies infinite phase-space density at the system's centre, so it is inconsistent with familiar cored polytropes, such as the Plummer model.

3 TWO-POWER MODELS

Any power-law model is problematic in the sense that the mass interior to radius r diverges as $r \to \infty$ if the density declines as r^{-b} with $b \leq 3$, and the mass outside radius r diverges as $r \to 0$ when $b \geq 3$. Hence, there is no value of b for which the model is physically reasonable at both large and small r. One way we can address this problem is to assume that ρ scales as different powers of radius at small and large radii. A widely used family of models of this type is given by the density profile

$$\rho(r) = \frac{\rho_0}{(r/r_b)^\alpha (1 + r/r_b)^{\beta-\alpha}}, \tag{25}$$

where r_b is the break radius (e.g. Binney & Tremaine 2008). Three particular cases of importance are the Jaffe (1983) model $(\alpha, \beta) = (2, 4)$, the Hernquist (1990) model $(\alpha, \beta) = (1, 4)$, which belong to the family of Dehnen (1993) models $(\beta = 4)$, and the NFW model $(\alpha, \beta) = (1, 3)$ (Navarro et al. 1996). The ergodic DFs of the Jaffe and Hernquist models are known analytic function, but that of the NFW model is not. Our goal in this section is to find analytic functions $f(\boldsymbol{J})$ that generate models that closely resemble these three classic models.

In the regime $r \ll r_b$ the mass $M(r)$ enclosed by the sphere of radius r is $M \propto r^{3-\alpha}$, so the gravitational acceleration is $d\Phi/dr \propto r^{1-\alpha}$ and thus the potential drop between radius r and the centre is

$$\Phi(r) - \Phi(0) \propto r^{2-\alpha} \quad \text{or} \quad \log(r) \text{ when } \alpha = 2. \tag{26}$$

Setting $a = 2 - \alpha$ we can now employ the results we derived above for power-law potentials to conclude that

$$f(\boldsymbol{J}) = [h(\boldsymbol{J})]^{-(6-\alpha)/(4-\alpha)}. \tag{27}$$

The Hernquist and NFW models both have $\alpha = 1$ so we expect their DFs to have asymptotic behaviour

$$f(\boldsymbol{J}) = [h(\boldsymbol{J})]^{-5/3} \text{ as } |\boldsymbol{J}| \to 0. \tag{28}$$

A Jaffe model has $\alpha = 2$, so the asymptotic behaviour of the Jaffe model's DF as $\boldsymbol{J} \to 0$ is given by equation (24).

Consider now the asymptotic behaviour of a two-power model as $r \to \infty$. If the model has finite mass, the potential will asymptote to the Kepler potential, $\Phi \propto r^{-1}$, so $\rho \propto |\Phi|^\beta$. In the Kepler regime, the dependence of the Hamiltonian on the actions is (e.g. Binney & Tremaine 2008, equation 3.226a)

$$H(\boldsymbol{J}) = [g(\boldsymbol{J})]^{-2}, \tag{29}$$

where $g(\boldsymbol{J})$ is a homogeneous function of degree one. Although ρ is a simple power of $|\Phi|$ we cannot employ the polytropic formula (15), because that rests on Poisson's equation, which does not apply in this case: the model's envelope is a collection of test particles that move in the Kepler potential generated by its core. We instead go back to Eddington's formula

$$f(\mathcal{E}) \propto \frac{d}{d\mathcal{E}} \int_0^{\mathcal{E}} \frac{d\Psi}{\sqrt{\mathcal{E} - \Psi}} \frac{d\rho}{d\Psi}, \tag{30}$$

where $\mathcal{E} = -E$ and $\Psi = -\Phi$. From this formula, it is easy to show that $\rho \propto \Psi^\beta$ implies

$$f(\mathcal{E}) \propto \mathcal{E}^{\beta-3/2}. \tag{31}$$

Combining this with equation (29) we conclude that for $\beta > 3$ the asymptotic behaviour of a double-power DF is

$$f(\boldsymbol{J}) = [g(\boldsymbol{J})]^{-2\beta+3} \text{ as } |\boldsymbol{J}| \to \infty. \tag{32}$$

For the Jaffe and Hernquist models $\beta = 4$, so for these models

$$f(\boldsymbol{J}) = [g(\boldsymbol{J})]^{-5} \text{ as } |\boldsymbol{J}| \to \infty. \tag{33}$$

Now that we have the asymptotic behaviour of f in the limits of both small and large $\boldsymbol{J}$, it is straightforward to devise a suitable form of the DF

$$f(\boldsymbol{J}) = \frac{M_0}{J_0^3} \frac{[1 + J_0/h(\boldsymbol{J})]^{(6-\alpha)/(4-\alpha)}}{[1 + g(\boldsymbol{J})/J_0]^{2\beta-3}}. \tag{34}$$

Here M_0 is a constant that has the dimensions of a mass and J_0 is a characteristic action. If the two homogeneous functions are normalized such that $h(\boldsymbol{J}) \simeq g(\boldsymbol{J}) \simeq |\boldsymbol{J}|$, orbits that linger near the break radius r_b have $|\boldsymbol{J}| \simeq J_0$. These conditions ensure that f tends to the required powers of h and g when $|\boldsymbol{J}| \ll J_0$ and $|\boldsymbol{J}| \gg J_0$, respectively.

We use different homogeneous functions for the regimes of small and large $\boldsymbol{J}$ because the frequency ratios in these two regimes will differ. In the Kepler regime, which is handled by g, all frequencies are equal, so if we require an isotropic model we choose

$$g(\boldsymbol{J}) = J_r + |J_\phi| + J_z. \tag{35}$$

In the regime of small $\boldsymbol{J}$, $\Omega_r > \Omega_\phi = \Omega_z$, and we take h to be of the form (10) with a frequency ratio that is less than unity. Unfortunately, in this regime the frequency ratio does vary over a surface of constant energy and an exactly isotropic model cannot be constructed using constant ratios. We simply use $\Omega_\phi/\Omega_r = \Omega_z/\Omega_r = 1/2$, which are the frequency ratios of a harmonic oscillator.

The DF (34) is infinite on the orbit $\boldsymbol{J} = 0$ of a star that is stationary at the model's centre. *Cuspy* models such as the Hernquist, Jaffe and NFW models do have such centrally divergent DFs, while in other *cored* systems the phase-space density reaches a finite maximum. Cored systems will be treated in Section 4.

3.1 Technicalities

Here, we touch on some technical issues that arise when one sets out to recover the observable properties of a model from the DF that defines it. The first step is to normalize the DF to the desired total mass by evaluating the integral (3). When the DF depends only on the function $h(\boldsymbol{J})$ defined by equation (10) [i.e. the case $g(\boldsymbol{J}) = h(\boldsymbol{J})$] with the frequency ratios $\omega \equiv \Omega_\phi/\Omega_r = \Omega_z/\Omega_r$ taken to be constant, it is convenient to change coordinates from (J_r, J_ϕ, J_z) to (J_r, L, J_z) and integrate out J_z, and then to change coordinates to (h, L) and integrate out L. Then one finds

$$\frac{M}{(2\pi)^3} = \int dh\, f(h) \int_0^{h/\omega} dL\, L = \frac{1}{2\omega^2} \int_0^\infty dh\, h^2 f(h). \tag{36}$$

In the more general case, when $h(\boldsymbol{J}) \neq g(\boldsymbol{J})$, the integral (3) cannot be reduced to one dimension. Equation (36) can be written as

$$\frac{M}{(2\pi)^3} = M_0 \int d\boldsymbol{y} \frac{[1 + 1/h(\boldsymbol{y})]^{(6-\alpha)/(4-\alpha)}}{[1 + g(\boldsymbol{y})]^{2\beta-3}}, \tag{37}$$

where $\boldsymbol{y} \equiv \boldsymbol{J}/J_0$. The integral in equation (37) is dimensionless and depends only on the model's parameters α, β and on the forms of the homogeneous functions h and g. It can therefore be computed at the outset. Then the value of M_0 can be set that ensures that the model has whatever mass is required.

The physical scales of the models are determined by the action scale J_0 and by the mass scale M_0, so the natural length-scale is

$$r_0 \equiv \frac{J_0^2}{GM_0}. \tag{38}$$

In following sections, we will present $f(\boldsymbol{J})$ analogues of three classic models that have a finite mass: the Hernquist, Jaffe and

Table 1. The ratio of the half-mass radius r_h to the scale radius r_0, defined by equation (38), for the $f(J)$ isochrone, $f(J)$ Hernquist and $f(J)$ Jaffe models. For comparison, we list also the ratio r_h/r_b, where r_b is the break radius, of the corresponding classical models.

	Isochrone	Hernquist	Jaffe
r_h/r_0	3.4	2.42	0.76
r_h/r_b	3.06	2.41	1

isochrone models. For our analogue models the top row of Table 1 gives the ratio r_h/r_0 of half-mass radius to the scale radius defined by equation (38). The second row gives for the classical models the ratio of r_h to the break radius, and we see that for the Hernquist model $r_0 = r_b$ to good precision, while in the other two cases the difference between r_0 and r_b is less than 25 per cent.

Once $f(J)$ has been normalized, we are able to determine the potential $\Phi(x)$ that the model self-consistently generates by the iterative procedure described by B14.

## 3.2 Worked examples

3.2.1 The Hernquist model

The Hernquist (1990) model is an interesting example both because it is a widely used model and because we can derive its ergodic DF as a function of the actions for comparison with the $f(J)$ model given by equation (34) with $(\alpha, \beta) = (1, 4)$, which hereafter we refer to as $f(J)$ Hernquist model.

In Appendix A, we derive an analytic expression for $J_r = J_r(H, L)$ in the spherical Hernquist potential. By numerically inverting this expression, we arrive at $H = H(J_r, L)$ for the Hernquist sphere. Combining this with the sphere's ergodic DF, which was given already by Hernquist (1990), we have the exact $f = f[H(J)]$. In Fig. 1, we show surfaces in action space on which this DF is constant together with surfaces on which DF of the $f(J)$ Hernquist model is constant. The differences are small but apparent and arise because the surfaces of constant energy are not exactly planar.

Fig. 2 compares the radial profiles of density, circular speed and radial component of velocity dispersion in the exact isotropic model and in the $f(J)$ Hernquist model. The largest discrepancy is in the velocity dispersion and reflects the fact that the model is significantly radially biased around r_0. The long-dashed curve in Fig. 3 shows that the $f(J)$ Hernquist model has a slight radial bias
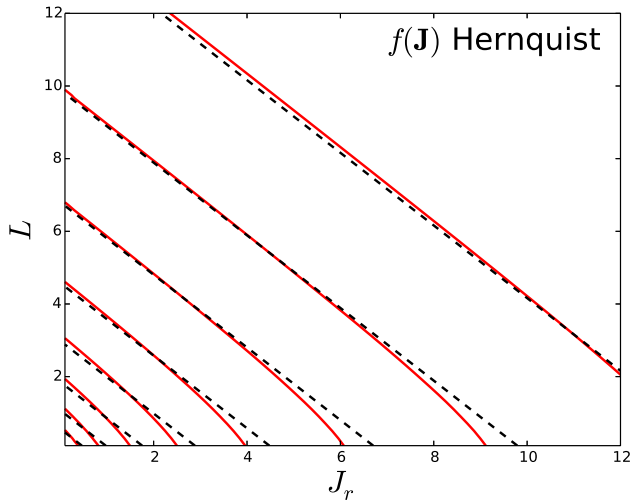
Figure 1. The red full curves show surfaces on which the DF of the classical isotropic Hernquist sphere is constant in the (J_r, L) plane of action space, while the black dashed curves show surfaces on which the corresponding $f(J)$ DF is constant.
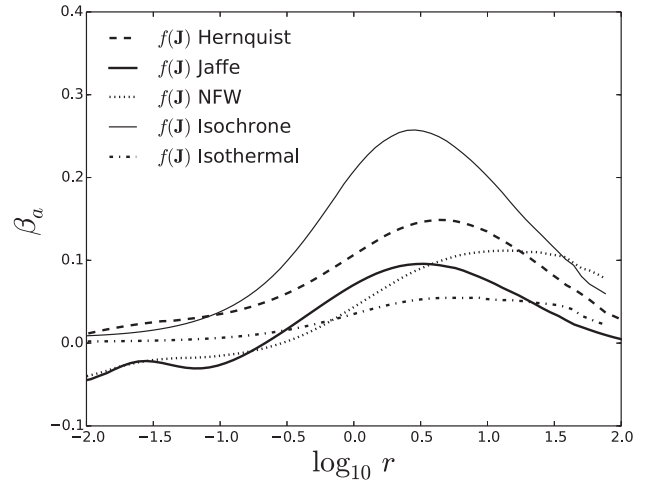
Figure 3. Anisotropy profiles for $f(J)$ Hernquist, $f(J)$ Jaffe, $f(J)$ NFW, $f(J)$ isochrone and $f(J)$ isothermal models. The profiles are normalized to r_0 (equation 38).
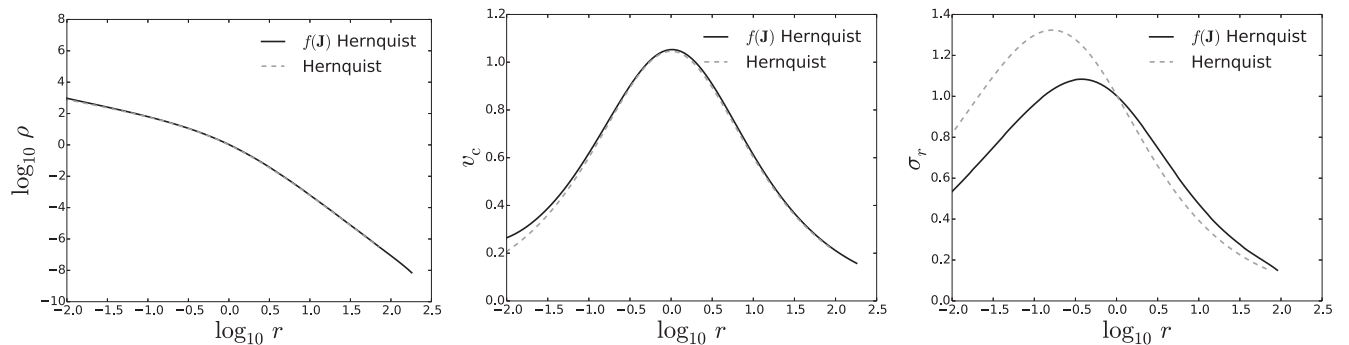
Figure 2. Density (left-hand panel), circular velocity (central panel), and radial velocity dispersion (right-hand panel) profiles for the classical isotropic Hernquist sphere (normalized to r_b) and for the $f(J)$ Hernquist model (normalized to r_0).

at all radii by plotting the anisotropy parameter

$$\beta_{\rm a} = 1 - \frac{\sigma_\phi^2 + \sigma_z^2}{2\sigma_r^2}. \tag{39}$$

By virtue of the adopted form of g (equation 35), $\beta_{\rm a} \to 0$ in the Keplerian regime. Even though the potential is not harmonic at the centre, still the model tends to isotropy also at small radii, which justifies our simple choice for $h(\boldsymbol{J})$.

3.2.2 The Jaffe model

The Jaffe (1983) model behaves as Hernquist's at large radii, while tending to $\rho \propto r^{-2}$ close to the centre. Fig. 4 shows the radial profiles of the $f(\boldsymbol{J})$ Jaffe model defined by setting $(\alpha, \beta) = (2, 4)$ in the DF (34), and compares them with the classical isotropic model. The discrepancies in σ_r are due to the slight radial bias of the $f(\boldsymbol{J})$ model around r_0. The full curve in Fig. 3 shows that this bias is actually quite mild – $|\beta_{\rm a}| < 0.1$.

3.2.3 NFW halo

The NFW model has $\beta = 3$ with the consequence that its mass diverges logarithmically as $r \to \infty$ and its potential is never Keplerian. Consequently, the reasoning used to construct a DF above equation (34) does not apply. If we nevertheless adopt equation (34) with $(\alpha, \beta) = (1, 3)$, we obtain a DF that implies that

as $J \to \infty$ the mass with actions less than J diverges like $\log J$. Asymptotically, the circular speed of the standard NFW model is

$$v_{\rm c} \sim \sqrt{\frac{\log(1 + r/r_0)}{r}}, \tag{40}$$

so in this model the action of a circular orbit is $J_\phi \sim \sqrt{r \log r}$. This shows that mass diverging like $\log J$ in action space corresponds, to leading order, to divergence of the mass in real space like $\log r$. Hence, it is plausible that the DF (34) with $(\alpha, \beta) = (1, 3)$ generates a model similar to the NFW model.

Computation of $\rho(r)$ for the $f(\boldsymbol{J})$ model with $(\alpha, \beta) = (1, 3)$ bears out this expectation. However, the slope of the model's density profile at large r is slightly steeper than desired, and a better fit to the classical NFW profile is obtained by adopting

$$f(\boldsymbol{J}) = \frac{M_0}{J_0^3} \frac{[1 + J_0/h(\boldsymbol{J})]^{5/3}}{[1 + g(\boldsymbol{J})/J_0]^{2.9}}. \tag{41}$$

Fig. 5 shows the radial profiles of the classical NFW model and those of the model generated by the DF (41), which we shall call $f(\boldsymbol{J})$ NFW model. The dotted curve in Fig. 3 shows that this model is mildly radially biased at radii larger than r_0 and it becomes very slightly tangentially biased for $r < r_0$. These anisotropies account for the difference between the σ_r profiles of the $f(\boldsymbol{J})$ and classical NFW models.

4 CORES AND CUTS

In the last section, we addressed the problematic nature of power-law models – that their mass diverges at either small or large radii – by introducing separate slopes of the dependence of f on $\boldsymbol{J}$ at small
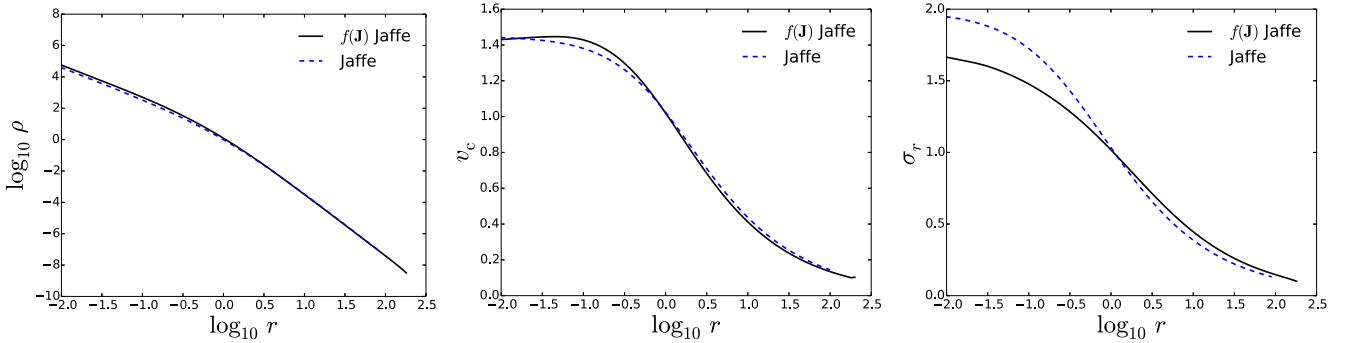

Figure 4. Same as Fig. 2, but for the classical isotropic Jaffe sphere and for the $f(\boldsymbol{J})$ Jaffe model.
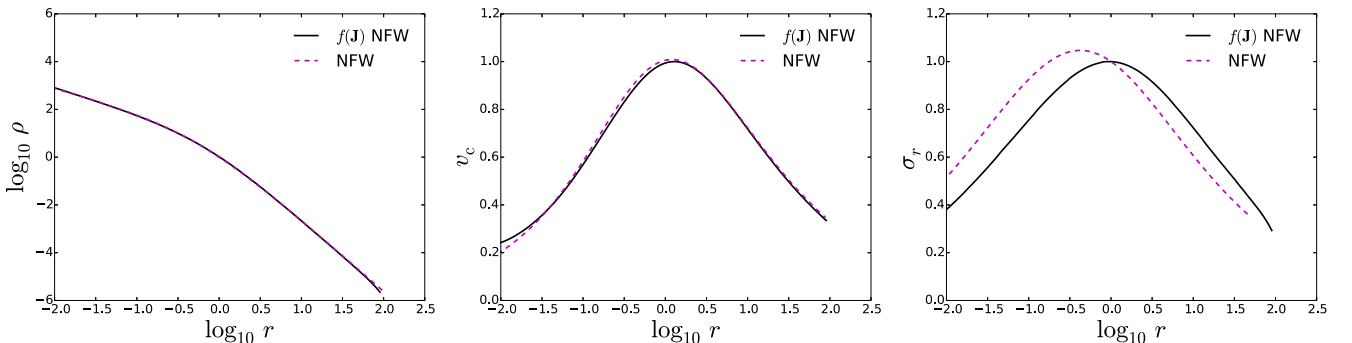

Figure 5. Same as Fig. 2, but for the classical isotropic NFW sphere and for the $f(\boldsymbol{J})$ NFW model defined by equation (41).

and large $\boldsymbol{J}$. The recovered models had central density cusps similar to those of the Hernquist, Jaffe and NFW models. If a homogeneous core is required, the natural DF to adopt is

$$f(\boldsymbol{J}) = \frac{M_0}{J_0^3} \frac{1}{[1 + g(\boldsymbol{J})/J_0]^{2\beta-3}}, \tag{42}$$

for then the phase-space density has the finite value M_0/J_0^3 at the centre of the model, and the asymptotic density profile is expected to be $\rho \propto r^{-\beta}$. For $\beta \leq 3$ the system has infinite mass, so for these models we taper the DF by subtracting a constant from the value given by equation (42)

$$f(\boldsymbol{J}) \mapsto f'(\boldsymbol{J}) = \max\left[0, f(\boldsymbol{J}) - f(\boldsymbol{J}_{\mathrm{t}})\right], \tag{43}$$

where $\boldsymbol{J}_{\mathrm{t}}$ is some large action, which defines a truncation radius

$$r_{\mathrm{t}} = \frac{|\boldsymbol{J}_{\mathrm{t}}|^2}{GM}. \tag{44}$$
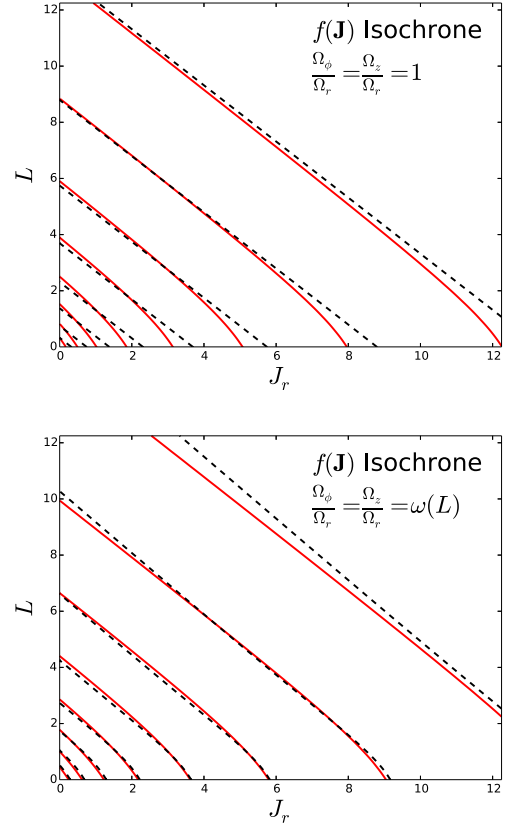
4.1 Isochrone model

Fig. 6 compares the density profiles of the model equation (42) generates for $\beta = 4$ (black curves) with those of the isochrone (Hénon 1960). The two models are extremely similar, so we shall refer to the model generated by the DF (42) when $\beta = 4$ as the $f(\boldsymbol{J})$ isochrone model. The density profiles of the two models are essentially identical, but at $r \simeq r_0$ σ_r is slightly smaller in the $f(\boldsymbol{J})$ isochrone than in the classical isochrone because the $f(\boldsymbol{J})$ isochrone is mildly radially biased near r_0 – the thin full curve in Fig. 3 shows $\beta_{\mathrm{a}}(r)$ for this model. It is non-zero because action space surfaces of $f(\boldsymbol{J})$ do not quite coincide with surfaces of constant $H(\boldsymbol{J})$, as the upper panel of Fig. 7 shows by plotting contours of f and H. For the isochrone potential, we have an analytic expression for the frequency ratio Ω_ϕ/Ω_r as a function of L. The lower panel of Fig. 7 shows that the constant-energy and constant-DF contours are more closely aligned when the argument of the homogeneous function uses the exact frequency ratio.

Given that the exact DF of the isochrone is a complicated function of $\boldsymbol{J}$, it is astonishing that the trivial DF (42) provides such a good approximation to it.

4.2 Cored isothermal sphere

In Section 2.1, we derived an approximation (24) to the DF of the singular isothermal sphere. Here, we modify this model into one that has a core and (ii) tapering


Figure 7. Surfaces of constant $H(\boldsymbol{J})$ (red) and of constant $f(\boldsymbol{J})$ black in action space for two $f(\boldsymbol{J})$ isochrone models with different choice of the function g appearing in the DF (42). In the upper panel, g is $J_r + L$ whereas in the lower panel it is $J_r + (\Omega_\phi/\Omega_r)L$ where the frequency ratio is a function of L.

its density at large radii so the model's mass becomes finite. Then the DF is

$$f(\boldsymbol{J}) = \frac{M_0}{J_0^3} \max\left(0, [1 + J_0/h(\boldsymbol{J})]^2 - [1 + J_0/h(\boldsymbol{J}_{\mathrm{t}})]^2\right), \tag{45}$$

where $h(\boldsymbol{J})$ is given by equation (10) with both frequency ratios set to $1/\sqrt{2}$ and $\boldsymbol{J}_{\mathrm{t}} = (0, v_{\mathrm{c}}r_{\mathrm{t}}, 0)$. As full curves in Fig. 8 show this DF generates a model that has a core that extends to r_0 and a density profile that plunges to zero near the truncation radius r_{t}. The short-dashed curve in Fig. 3 shows the model's anisotropy parameter β_{a}, which is always small ($|\beta_{\mathrm{a}}| < 0.04$).
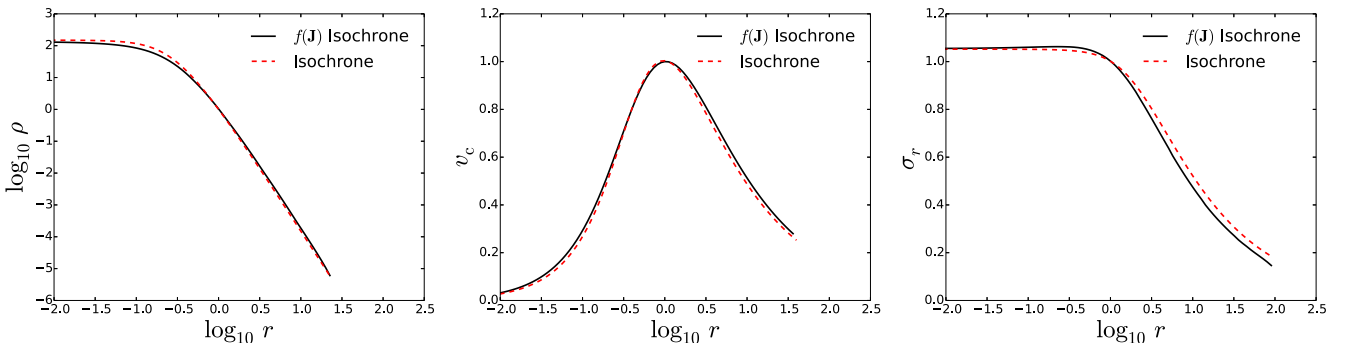
Figure 6. Same as Fig. 2, but for the classical isotropic isochrone sphere and for the $f(\boldsymbol{J})$ isochrone model.
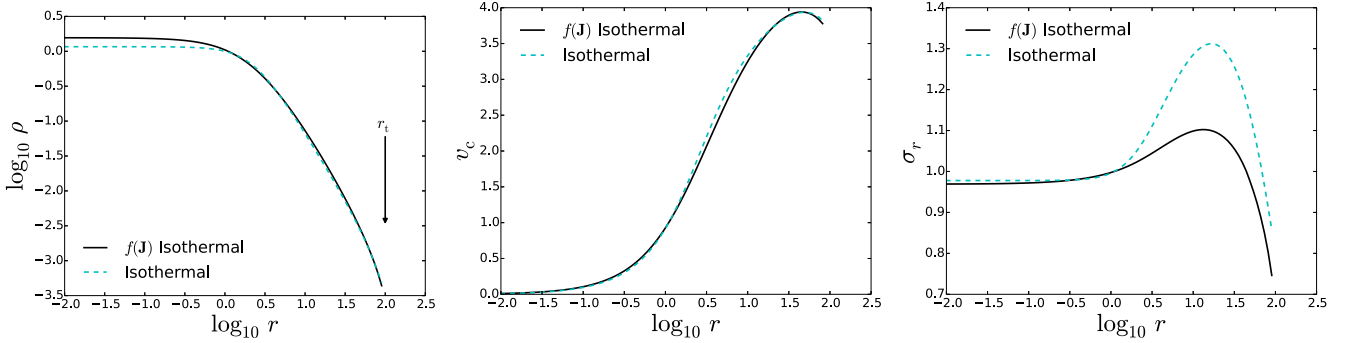
Figure 8. Same as Fig. 2, but for the truncated isotropic cored isothermal sphere (equation 46) and for the $f(\boldsymbol{J})$ truncated isothermal model. We show the location of the truncation radius defined by equation (44).

An ergodic model with a simple functional form of $\rho(r)$ to which we can compare our $f(\boldsymbol{J})$ model has

$$\rho(r) = \frac{v_c^2}{2\pi G \left(r^2 + r_b^2\right)} e^{-r^2/r_t^2}. \tag{46}$$

The dashed curve in the left-hand panel of Fig. 8 shows that the model defined by the DF (46) provides an excellent fit to the density profile of our $f(\boldsymbol{J})$ model. Curiously, in the $f(\boldsymbol{J})$ model $\sigma_r(r)$ is more nearly constant within r_t than in either of the models with analytic density profiles. The dashed curve in the right-hand panel of Fig. 8 shows that the model defined by the DF (46) has a significantly deeper central depression in σ_r than the $f(\boldsymbol{J})$ model.

5 CONCLUSIONS

Studies of both our own and external galaxies will benefit from the availability of a flexible array of *dynamical* models of galactic components such as disc, bulge and dark halo. The construction of general models of this type is rather straightforward when one decides to start from an expression for the component's DF as a function of the action integrals J_i. In this paper, we have illustrated this fact by deriving simple analytic forms for DFs that self-consistently generate models that closely resemble the isochrone, Hernquist, Jaffe, NFW and truncated isothermal models. In previous papers, Binney (2010, 2012b) has given simple analytic DFs that provide excellent fits to the structure of the Galactic disc, so now DFs are available for all commonly occurring galactic components.

Our models are tailored to minimize velocity anisotropy at both small and large radii. In all of them, the anisotropy parameter β_a peaks at intermediate radii. The peak is by far sharpest in the $f(\boldsymbol{J})$ isochrone, but even in this model β_a stays below 0.25.

Our presentation has been elementary in the sense that we have confined ourselves to spherical, almost isotropic components that live in isolation. However, B14 showed that given a near-ergodic DF $f(\boldsymbol{J})$ of a component such as those presented here, it is trivial to modify it so it generates a system that is flattened by velocity anisotropy, or by rotation, or by a combination of the two. Equally important, when the DF of an individual component is given as $f(\boldsymbol{J})$, it is straightforward to add components. Such addition was exploited by Piffl et al. (2014) in a study of the contribution of dark matter to the gravitational force on the Sun: in that study the models fitted to data comprised a sum of DFs $f(\boldsymbol{J})$ for the disc and the stellar halo. The dark halo was assigned a density distribution rather than a DF, but Piffl et al. (in preparation) represent the dark

halo by the $f(\boldsymbol{J})$ NFW model, making the Galaxy a completely self-consistent object. A key point for such work is that the mass of each component can be specified at the outset.

Our approach has several points of contact with that of Williams et al. (2014) and Evans & Williams (2014), who derive approximations to $H(\boldsymbol{J})$ for models that are defined by DFs of the form $f(E, L)$. In particular, they show that for their models better approximations to the iso-energy surfaces in action space can be obtained if one's homogeneous function has as its argument the sum of a linear function of the actions, as used here, and a small term $\epsilon \sqrt{L J_r}$. We expect that the anisotropy of our models could be enhanced by adding such a term.

In addition to assisting in the dynamical interpretation of observations of galaxies, the models that this work makes possible could provide useful initial conditions for *N*-body simulations. The first step would be the construction of a self-consistent galaxy model from a judiciously chosen DF. Then one could Monte Carlo sample the action space using the DF as the sampling density, and torus mapping (e.g. Binney & McMillan 2011) could be used to generate an orbital torus at each of the selected actions. Finally, some number *n* of initial conditions $(\boldsymbol{x}, \boldsymbol{v})$ would be selected on each torus, uniformly space in the angles θ_i. The resulting simulation would be in equilibrium to whatever precision had been used in the solution of Poisson's equation, and it would experience a 'cold start' (Sellwood 1987). Moreover, given that it would be possible to evaluate the original DF at any phase-space point, the model would lend itself to the method of perturbation particles (Leeuwin, Combes & Binney 1993) in which the simulation particles represent the difference between a dynamically evolving model and an underlying equilibrium rather than the whole model. This method has been little used in the past on account of the lack of interesting models with known DFs, which is precisely the need that we have here supplied.

REFERENCES

Arnold V. I., 1978, The Mathematical Methods of Classical Mechanics. Springer, New York
Binney J., 2010, MNRAS, 401, 231
Binney J., 2012a, MNRAS, 426, 1324
Binney J., 2012b, MNRAS, 426, 1328
Binney J., 2014, MNRAS, 440, 787 (B14)
Binney J., McMillan P. J., 2011, MNRAS, 413, 1889
Binney J., Tremaine S., 2008, Galactic Dynamics, 2nd edn. Princeton Univ. Press, Princeton, NJ
Byrd P. F., Friedman M. D., 1971, Handbook of Elliptic Integrals for Engineers and Scientists. Springer-Verlag, Berlin
Ciotti L., 1996, ApJ, 471, 68
Dehnen W., 1993, MNRAS, 265, 250
Dickson L. E., 1914, Elementary Theory of Equations. Wiley Inc., Hoboken, NJ
Eddington A. S., 1916, MNRAS, 76, 572
Evans N. W., 1994, MNRAS, 267, 333
Evans N. W., Williams A. A., 2014, MNRAS, 443, 791
Fermani F., 2013, DPhil thesis, Univ. Oxford
Gantmacher F. R., 1959, The Theory of Matrices Vol. 2. Am. Math. Soc., Providence, Rhode Island
Gerhard O. E., Saha P., 1991, MNRAS, 251, 449
Hénon M., 1960, Ann. Astrophys., 23, 474
Hernquist L., 1990, ApJ, 356, 359
Jaffe W., 1983, MNRAS, 202, 995
Jeans J. H., 1915, MNRAS, 76, 70
Leeuwin F., Combes F., Binney J., 1993, MNRAS, 262, 1013
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Piffl T. et al., 2014, MNRAS, 445, 3133
Prendergast K. H., Tomer E., 1970, AJ, 75, 647
Rowley G., 1988, ApJ, 331, 124
Sanders J. L., Binney J., 2014, MNRAS, 441, 3284
Sanders J. L., Binney J., 2015, MNRAS, in press
Sellwood J. A., ARA&A, 25, 151
Williams A. A., Evans N. W., Bowden A. D., 2014, MNRAS, 442, 1405
Wilson C. P., 1975, ApJ, 80, 175

APPENDIX A: ANALYTICAL EXPRESSION FOR THE RADIAL ACTION IN THE HERNQUIST SPHERE

The radial action is defined as

$$J_\mathrm{r} = \frac{1}{2\pi} \oint p_r \mathrm{d}r = \frac{1}{\pi} \int_{r_1}^{r_2} \mathrm{d}r \sqrt{2E - 2\Phi(r) - \frac{L^2}{r^2}}, \tag{A1}$$

where $\Phi(r) = -GM/(r + r_\mathrm{b})$ and r_1, r_2 are the pericentric and apocentric radii for the given energy E and angular momentum L, i.e. the two roots of the integrand in equation (A1). Introducing the dimensionless quantities $s \equiv r/r_\mathrm{b}$, $\mathcal{E} \equiv -Er_\mathrm{b}/GM$ and $l = L/\sqrt{2GMr_\mathrm{b}}$, equation (A1) can be rewritten as

$$J_\mathrm{r} = \frac{\sqrt{2GMr_\mathrm{b}}}{\pi} \int_{s_1}^{s_2} \mathrm{d}s \sqrt{-\mathcal{E} + \Psi(s) - \frac{l^2}{s^2}}, \tag{A2}$$

where $\Psi(s) \equiv 1/(1 + s)$ is the relative dimensionless potential. We now change the integration variable from $s = (1 - \Psi)/\Psi$ to Ψ

(Ciotti 1996), and have

$$J_\mathrm{r} = \frac{\sqrt{2GMr_\mathrm{b}}}{\pi} \int_{\Psi_2}^{\Psi_1} \mathrm{d}\Psi \, \frac{\sqrt{\mathcal{P}(\Psi)}}{(1 - \Psi)\Psi^2}, \tag{A3}$$

where $\Psi_1 \equiv \Psi(s_1)$, $\Psi_2 \equiv \Psi(s_2)$ and

$$\mathcal{P}(\Psi) = -\mathcal{E}(1 - \Psi)^2 + \Psi(1 - \Psi)^2 - l^2\Psi^2 \tag{A4}$$

is a cubic in Ψ, the roots of which can be found by standard methods (e.g. Dickson 1914). Ψ_1, Ψ_2 are two roots in the physical range $0 \leq \Psi \leq 1$. Let A be the third real root, so

$$\mathcal{P}(\Psi) = (\Psi_1 - \Psi)(\Psi - \Psi_2)(A - \Psi). \tag{A5}$$

While it is physically obvious that two of the three real solutions of equation (A4) are in the range (0, 1) and the remaining one is outside ($A > 1$), we remark that the same conclusion can be reached by purely algebraic arguments by using the Routh–Hurwitz theorem (see e.g. Gantmacher 1959). By evaluating equations (A4) and (A5) at $\Psi = 0$ one gets $A = \mathcal{E}/\Psi_1\Psi_2 > 0$. By splitting into its partial fractions the integrand in equation (A3), it is possible to express the integral for J_r in terms of complete elliptic integrals (see Byrd & Friedman 1971):

$$J_r = \frac{\sqrt{2GMa}}{\pi} D_5 [D_1 \Pi(\alpha_1, k^2) + D_2 \mathrm{E}(k^2)$$
$$+ D_3 \mathrm{K}(k^2) + D_4 \Pi(\alpha_2, k^2)], \tag{A6}$$

where K, E, Π are, respectively, the complete elliptic integral of the first, second and third kind,

$$\alpha_1 \equiv \frac{\Psi_1 - \Psi_2}{\Psi_2}, \quad \alpha_2 \equiv \frac{\Psi_1 - \Psi_2}{1 - \Psi_2}, \quad k^2 \equiv \frac{\Psi_1 - \Psi_2}{A - \Psi_2} \tag{A7}$$

and finally

$$\begin{aligned}
D_1 &= [(1 - 2\Psi_1)\Psi_2 + \Psi_1]A + \Psi_1\Psi_2, \\
D_2 &= \Psi_2(\Psi_2 - A), \\
D_3 &= \Psi_2(A - 2), \\
D_4 &= 2\Psi_2(A - 1)(\Psi_1 - 1), \\
D_5 &= -\sqrt{A - \Psi_2}/D_2.
\end{aligned} \tag{A8}$$

We have tested the formula (A6) for consistency by numerically integrating equation (A1) for a large set of orbits at different (E, L) and the numerical and analytical results agree within the error of the employed routine.

This paper has been typeset from a TeX/LaTeX file prepared by the author.