Mon. Not. R. Astron. Soc. **277**, 1341–1353 (1995)

Nearly self-consistent disc–bulge–halo models for galaxies

Konrad Kuijken[1] and John Dubinski[2]
[1] *Kapteyn Institute, PO Box 800, 9700 AV Groningen, The Netherlands*
[2] *Lick Observatory, University of California, Santa Cruz, CA 95064, USA*

Accepted 1995 July 11. Received 1995 June 19; in original form 1995 February 13

ABSTRACT

We describe methods for setting up self-consistent disc–bulge–halo galaxy models. The bulge and halo distribution functions (DFs) are functions of E and L_z only. The halo's flattening and rotation can be specified. The disc DF is a function of E and L_z and a third 'integral', E_z, the vertical energy, which is approximately conserved in a warm disc with vertical extent. The models also have finite extent, making them suitable for N-body simulation. A simulation of a sample model shows that in practice the models are very close to equilibrium, making them ideal for experiments on instabilities in galactic discs. We also present a sequence of models closely resembling the Milky Way mass distribution with 5 exponential scale radii and varying halo mass and radial extent.

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

1 INTRODUCTION

This paper describes a set of semi-analytic models for the phase-space distribution functions of axisymmetric disc galaxies. They contain three components, corresponding to a disc, bulge and halo. The distribution functions are simple functions of three integrals of motion: the two analytic ones (energy and angular momentum about the axis of symmetry) and one approximate integral which describes the vertical motions in the disc component. These distribution functions yield a unique density for each component in any given potential: since all three components affect each other gravitationally, a numerical solution of the Poisson equation is required for self-consistent models.

Setting up a stable three-dimensional disc for N-body experiments is a difficult task. Strategies to handle this problem have included adiabatically growing the disc mass distribution in a self-consistent halo/bulge model (e.g. Barnes 1988), treating the halo and bulge as a static background (e.g. Sellwood & Merritt 1994; Quinn, Hernquist & Fullagar 1993), or directly solving the Jeans equations (under a suitable ansatz) for the complete system of the disc, bulge and halo to find the velocity dispersions (Hernquist 1993), and realizing these dispersions with (typically) Gaussian distributions. In a recent study, Sellwood & Merritt (1994) used the DF calculated by Kalnajs (1976) of the Kuzmin–Toomre disc to set up the radial and azimuthal velocities, but this method is restricted to that class of models, and they did use the Jeans equations for the vertical velocities.

While these methods are useful for discs with small particle number N where the discreteness noise is dominant, as N increases the approximate initial conditions introduce subtle transient behaviour which can interfere with the interpretation of experiments. Adiabatic growth of the disc, or relaxation from non-equilibrium initial conditions, has the further disadvantage that the relaxed initial conditions are not under perfect control of the experimenter. It is for example not uncommon to observe outward propagating rings of overdensity from the warmer disc centre (where the Gaussian approximation is the poorest), and transients in the disc velocity dispersion which can change the initial values by ~ 20 per cent before the particles settle into equilibrium (Hernquist, Mihos & Walker, private communication). As a result, though these models relax quickly to equilibrium, the velocity dispersion profiles change from the input values. The density profiles usually do not change significantly. While such subtle effects do not matter in experiments in which a disc will be strongly perturbed (say by a merger), they can be significant in investigations of the rather subtle instabilities which arise from the inherent fragility of a cool disc. Examples of these problems are spiral structure formation (Sellwood & Carlberg 1984), bar instabilities (e.g. Hernquist & Weinberg 1992), bending instabilities in counter-streaming discs (e.g. Sellwood & Merritt 1994; Merritt & Sellwood 1994), disc warping (Dubinski & Kuijken 1995), and disc heating by satellite accretion or other tidal perturbations (e.g. Quinn et al. 1993). Simulations with N in the millions are now becoming possible with greater computing power. The subtle transients at start-up, due to various approximations which were previously drowned out by discreteness, become correspondingly more bothersome. A cleaner set of initial conditions is desirable.

In this paper we describe the construction of our models. In Section 2, we present the distribution functions (DFs) of the disc, bulge and halo and describe a method for calculating the potential of the self-consistent model. The bulge and halo have the DF of King (1966) models and their flattened generalization, the lowered Evans distributions (Kuijken & Dubinski 1994). These DFs depend only on the classical integrals of motion, and can be combined straightforwardly. The disc DF is a generalization to three dimensions of the planar model devised by Shu (1969) (see also Kuijken & Tremaine 1992), similar to the construction by Binney (1987), and allows almost arbitrary specification of the radial variation of density and velocity dispersion. Since two-integral distribution functions for discs are not realistic (Oort 1965), we employ an approximate third integral. Section 3 contains N-body demonstrations and tests of equilibrium for a specific model. In Section 4, we present a sequence of models with different halo extent suitable for describing the Milky Way. Finally, we discuss the advantages and limitations of these models in Section 5 and summarize our results.

2 DISTRIBUTION FUNCTION

Our strategy is an extension of the technique we employed in the construction of the lowered Evans models (Kuijken & Dubinski 1994), and earlier applied by, e.g., Prendergast & Tomer (1970) and Rowley (1988). The starting point is a chosen analytic form for the DF, written in terms of known integrals of motion. Any such DF represents an equilibrium model in any gravitational potential that respects these integrals, with space density explicitly determined by this potential. We can therefore construct self-gravitating models by requiring that the potential and the density be related by Poisson's equation. This last step must generally be done numerically.

In order of increasing complexity, we now detail the DFs of the three galactic components.

2.1 The bulge distribution function

For the bulge DF we take a King model (King 1966). This DF has the form

$$f_{\rm bulge}(E) = \begin{cases} \rho_{\rm b}(2\pi\sigma_{\rm b}^2)^{-3/2}\exp[(\Psi_0-\Psi_{\rm c})/\sigma_{\rm b}^2]\{\exp[-(E-\Psi_{\rm c})/\sigma_{\rm b}^2]-1\} & \text{if } E<\Psi_{\rm c}, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

It depends on the three parameters: $\Psi_{\rm c}$ (the cutoff potential of the bulge), $\rho_{\rm b}$ (approximately the central bulge density, ignoring the effects of the DF truncation) and $\sigma_{\rm b}$, which governs the velocity dispersion of the bulge component. Ψ_0 is the gravitational potential at the centre of the model.

The density of the bulge component in a potential Ψ is obtained by integrating its DF over all velocities, resulting in

$$\rho_{\rm bulge}(\Psi) = \rho_{\rm b}\left[e^{(\Psi_0-\Psi)/\sigma_{\rm b}^2}\,{\rm erf}(\sqrt{(\Psi_{\rm c}-\Psi)/\sigma_{\rm b}^2}) - \pi^{-1/2}e^{(\Psi_0-\Psi_{\rm c})/\sigma_{\rm b}^2}\left(2\sqrt{(\Psi_{\rm c}-\Psi)/\sigma_{\rm b}^2}-\frac{4}{3}[(\Psi_{\rm c}-\Psi)/\sigma_{\rm b}^2]^{3/2}\right)\right] \tag{2}$$

where $\Psi<\Psi_{\rm c}$, and zero density elsewhere. ${\rm erf}(x)\equiv 2\pi^{-1/2}\int_0^x \exp(-t^2)\,{\rm d}t$ is the usual error function.

In what follows, we will normally choose $\sigma_{\rm b}<\sigma_0$ and $\Psi_{\rm c}<0$ to make the bulge more centrally condensed, and more radially confined, than the halo (the latter has a cutoff at zero energy).

2.2 The halo distribution function

We use the DF of a lowered Evans model (Kuijken & Dubinski 1994) for the halo. This DF is a truncation at zero energy of the models discovered by Evans (1993) for the flattened logarithmic potential. It takes the form

$$f_{\rm halo}(E,L_z^2) = \begin{cases} [(AL_z^2+B)\exp(-E/\sigma_0^2)+C][\exp(-E/\sigma_0^2)-1] & \text{if } E<0, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

The density corresponding to this DF is given by equation (9) of Kuijken & Dubinski (1994), and is repeated here for completeness:

$$\begin{aligned} \rho_{\rm halo}(R,\Psi) \;=\; & \tfrac{1}{2}\pi^{3/2}\sigma_0^3(AR^2\sigma_0^2+2B)\,{\rm erf}(\sqrt{-2\Psi}/\sigma_0)\exp(-2\Psi/\sigma_0^2) \\ & +(2\pi)^{3/2}\sigma_0^3(C-B-AR^2\sigma_0^2)\,{\rm erf}(\sqrt{-\Psi}/\sigma_0)\exp(-\Psi/\sigma_0^2) \\ & +\pi\sqrt{-2\Psi}[\sigma_0^2(3A\sigma_0^2R^2+2B-4C)+\frac{4}{3}\Psi(2C-A\sigma_0^2R^2)]. \end{aligned} \tag{4}$$

The halo DF has five free parameters: the potential well depth Ψ_0, the velocity and density scales σ_0 and ρ_1, the halo core radius $R_{\rm c}$ and the flattening parameter q (the last three of these contained within the parameters A, B, and C). For convenience, we have defined a characteristic halo radius $R_{\rm a}$ which replaces the density scale ρ_1 of Kuijken & Dubinski (1994):

$$R_{\rm a} = \left(\frac{3}{2\pi G\rho_1}\right)^{1/2}\sigma_0 e^{\Psi_0/2\sigma_0^2}; \tag{5}$$

it is roughly the radius at which the halo rotation curve, if continued at its $R=0$ slope, would reach the value $2^{1/2}\sigma_0$.

Arbitrary amounts of rotation can be added to the halo model by splitting the DF into parts with positive and negative L_z.

2.3 The disc distribution function

In the construction of a realistic three-integral disc distribution function, the issue of a third integral cannot be evaded (as it was for the bulge and halo components) for it is an observed fact that in the solar neighbourhood (e.g., Wielen 1974) and in the discs of other galaxies (Bottema 1993) the vertical and radial dispersions are different, which is not possible in any DF that depends only on energy and angular momentum. The simplest approximate third integral in an axisymmetric disc system is the energy in the vertical oscillations, $E_z \equiv \Psi(R,z) - \Psi(R,0) + \frac{1}{2}v_z^2$. It is quite well conserved along nearly circular orbits which have no large radial or vertical excursions. We will use this quantity as third integral for the disc DF in our models. More sophisticated possible integrals are discussed by Kent & de Zeeuw (1991).

Armed with a third integral, the disc distribution function can be constructed by extending the planar DF discussed by Shu (1969) and Kuijken & Tremaine (1992) into the vertical dimension, similar to the DF constructed by Binney (1987). Thus, we have

$$f_{\rm disc}(E_{\rm p},L_z,E_z) = \frac{\Omega(R_{\rm c})}{(2\pi^3)^{1/2}\kappa(R_{\rm c})}\frac{\widetilde{\rho}_{\rm d}(R_{\rm c})}{\widetilde{\sigma_R}^2(R_{\rm c})\widetilde{\sigma_z}(R_{\rm c})}\exp\left[-\frac{E_{\rm p}-E_{\rm c}(R_{\rm c})}{\widetilde{\sigma_R}^2(R_{\rm c})}-\frac{E_z}{\widetilde{\sigma_z}^2(R_{\rm c})}\right]. \tag{6}$$

Here, $E_{\rm p} \equiv E - E_z$ is the energy in planar motions, L_z is the specific angular momentum about the axis of symmetry, $R_{\rm c}$ and $E_{\rm c}$ are the radius and energy of a circular orbit with angular momentum L_z, and Ω and κ are the circular and epicyclic frequencies at radius $R_{\rm c}$. The density corresponding to this DF is obtained by integrating over the three velocity components. The v_R- and v_z-integrals are straightforward, leaving the v_ϕ-integral:

$$\begin{aligned}\rho_{\rm disc}(R,z) \;=\; & \int_0^\infty \left\{\left[{\rm d}v_\phi \equiv {\rm d}R_{\rm c}\left(\frac{R_{\rm c}\kappa(R_{\rm c})^2}{2R\Omega(R_{\rm c})}\right)\right]\frac{2\widetilde{\rho}_{\rm d}(R_{\rm c})\Omega(R_{\rm c})}{(2\pi)^{1/2}\widetilde{\sigma_R}(R_{\rm c})\kappa(R_{\rm c})}\times\right.\\ & \left.\exp\left[-\frac{\Psi(R,0)-\Psi(R_{\rm c},0)}{\widetilde{\sigma_R}^2(R_{\rm c})}-\left(\frac{R_{\rm c}^2}{R^2}-1\right)\frac{v_{\rm c}^2(R_{\rm c})}{2\widetilde{\sigma_R}^2(R_{\rm c})}-\frac{\Psi(R,z)-\Psi(R,0)}{\widetilde{\sigma_z}^2(R_{\rm c})}\right]\right\}.\end{aligned} \tag{7}$$

In the $z = 0$ plane, this expression reduces to $\widetilde{\rho}_{\rm d}(R)$ with fractional error $O(\widetilde{\sigma_R}^2/v_{\rm c}^2)$, and to the same order the radial velocity distribution is Gaussian with dispersion $\widetilde{\sigma_R}(R)$ (see Kuijken & Tremaine 1992). The essence of the construction in equation (6) is the replacement of the radius R (which is not an integral of motion) by the epicyclic radius $R_{\rm c}$ (which is a function of angular momentum, and therefore is conserved along orbits). In warm discs, in which excursions from circular orbits are small but not negligible, this parametrization still provides a good starting point for constructing a DF with given radial density and velocity dispersion profiles. The vertical structure of this disc is approximately isothermal, with the scale height set by the vertical velocity dispersion $\widetilde{\sigma_z}(R_{\rm c})$ and the vertical potential gradient.

Observed large disc galaxies have vertical light profiles similar to the $\mathrm{sech}^2(z)$-dependence expected for a vertically isothermal, self-gravitating sheet (van der Kruit & Searle 1981): therefore, as long as the disc is mainly confined in z by its own gravity, the DF of equation (6) will be a reasonable model.

In any gravitational potential, we can adjust the 'tilde' functions $\widetilde{\rho}$, $\widetilde{\sigma_R}$ and $\widetilde{\sigma_z}$ to the desired disc characteristics. In this paper, we arrange for the disc density to be approximately radially exponential with scale length $R_{\rm d}$ and truncated at radius $R_{\rm out}$:

$$\rho_{\rm disc}(R,z) = \frac{M_{\rm d}}{8\pi R_{\rm d}^2 z_{\rm d}}{\rm e}^{-R/R_{\rm d}}\,{\rm erfc}\left(\frac{r-R_{\rm out}}{2^{1/2}\delta R_{\rm out}}\right)\exp\left[-0.8676\frac{\Psi_z(R,z)}{\Psi_z(R,z_{\rm d})}\right]. \tag{8}$$

Here $M_{\rm d}$ is a parameter which is close to the mass of the disc unless the disc is severely truncated or the vertical structure is far from $\mathrm{sech}^2(z/z_{\rm d})$. $\delta R_{\rm out}$ governs the sharpness of the truncation. The vertical density of these discs is constructed to depend exponentially on the vertical potential $\Psi_z(R,z) \equiv \Psi(R,z) - \Psi(R,0)$, and to drop from the mid-plane value by a factor $\mathrm{sech}^2(1) \simeq {\rm e}^{-0.8676}$ at a height of $z_{\rm d}$, similar to the behaviour of a constant-thickness isothermal sheet.

Given a total potential for the model, we then set the disc tilde functions in the disc DF as follows. In the limit of very small velocity dispersions these functions are the actual mid-plane density and velocity dispersions. We first choose the function $\widetilde{\sigma_R}(R_{\rm c})$, approximately determining the radial velocity dispersion in the disc. $\widetilde{\rho}$ and $\widetilde{\sigma_z}$ are then iteratively adjusted so that the densities on the mid-plane and at height $z = z_{\rm d}$ agree with those of equation (8). It turns out that, at least for the models described in the remainder of this paper, this recipe yields a DF which has a space density close to that given by equation (8).

2.4 Calculation of the combined potential

The distribution functions for the various galaxy components all imply a unique volume density in a given potential. To construct a self-gravitating model, we need to find the potential in which the combined density is also the one implied by Poisson's equation, i.e.

$$\nabla^2\Psi(R,z) = 4\pi G[\rho_{\rm disc}(R,\Psi,\Psi_z) + \rho_{\rm bulge}(\Psi) + \rho_{\rm halo}(R,\Psi)]. \tag{9}$$

We solve equation (9) using a spherical harmonic expansion, following Prendergast & Tomer (1970), with two significant modifications. First, we have found that the disc density obtained by integrating the disc DF (equation 6) over all velocities is close

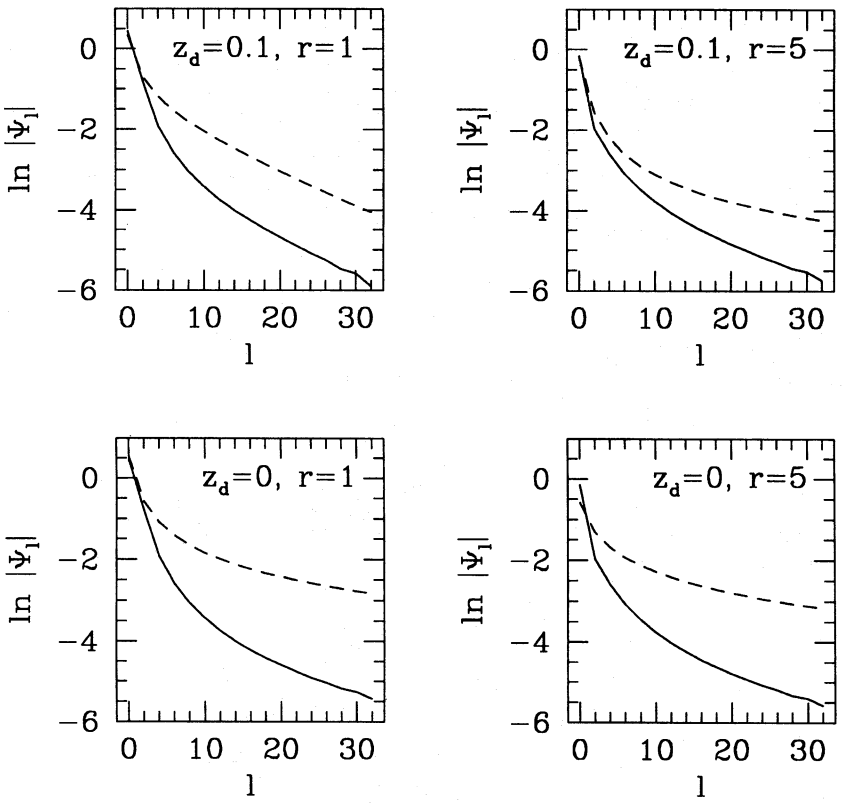

Figure 1. Comparison between the harmonic coefficients required to describe the potential for a radially exponential, vertically isothermal disc with (solid lines) and without (dashed lines) the $\Psi^{\dagger}_{\rm disc}$ term. In all cases shown, the disc scale length is 1. Top panels: disc scale height 0.1 at $r = 1$ (left) and $r = 5$ (right). Bottom panels: as the top ones, except for a disc with zero thickness.

to the value given by equation (8) (this was, after all, what the disc DF was designed to do). We therefore use this latter expression in the solution of Poisson's equation: this change avoids the single integral that would have to be calculated numerically each time the disc density was needed. Secondly, and more importantly, for realistically thin discs a spherical harmonic expansion is not very efficient, since high-order terms must be taken before a good approximation can be obtained. Moreover, zero-thickness discs, which form a very regular limit physically, require a prohibitive number of terms in the series. We therefore construct an analytic potential which represents the high-frequency terms correctly, and only fit harmonics to the residue.

A possible analytic 'high harmonics' disc potential $\Psi^{\dagger}_{\rm disc}$ is obtained by vertically integrating the disc density twice (i.e. by solving Poisson's equation for the disc component ignoring the radial gradient terms). Assuming that the vertical profile of the disc is that of a self-gravitating, plane-parallel sheet, this approach yields

$$\Psi^{\dagger\prime}_{\rm disc} = 4\pi G\rho_{\rm d}(R) z_{\rm d}^2 \ln\cosh(z/z_{\rm d}). \tag{10}$$

This potential, as written, is not suitable for removing the high harmonics because the density corresponding to this potential, $\nabla^2\Psi^{\dagger\prime}_{\rm disc}/4\pi G$, does not converge to 0 at large distances from the origin. The residual density can therefore not be fitted conveniently with spherical harmonic coefficients. We resolve this problem by replacing cylindrical radius R by spherical radius r in equation (10): since the radial density profile has a cutoff, the resulting potential and its corresponding density will indeed converge at large radii, while at small z the potential still models well the higher frequency terms of the disc.

The high-frequency disc potential we use is then

$$\Psi^{\dagger}_{\rm disc}(r,z) = \frac{GM_{\rm d}z_{\rm d}}{2R_{\rm d}^2} \ln\cosh(z/z_{\rm d})\ \mathrm{erfc}\left(\frac{r-R_{\rm out}}{2^{1/2}\delta R_{\rm out}}\right) \mathrm{e}^{-r/R_{\rm d}}. \tag{11}$$

The corresponding volume density is rather involved, but can be obtained using the result

$$\nabla^2 f(r) \ln\cosh z = f''(r)\ln\cosh z + 2\frac{f'(r)}{r}(z\tanh z + \ln\cosh z) + f(r)\,\mathrm{sech}^2 z, \tag{12}$$

where once again it should be remembered that r is the spherical radius. Note that the final term reproduces the disc density of a $\mathrm{sech}^2(z)$ disc to $O(z/R)^2$, and that the other terms have traded a vertical derivative for a radial one, and so are subdominant in thin discs. Moreover, as long as f'/r and f'' decay faster than $1/r^2$ for small r, these correction terms tend to zero at the origin.

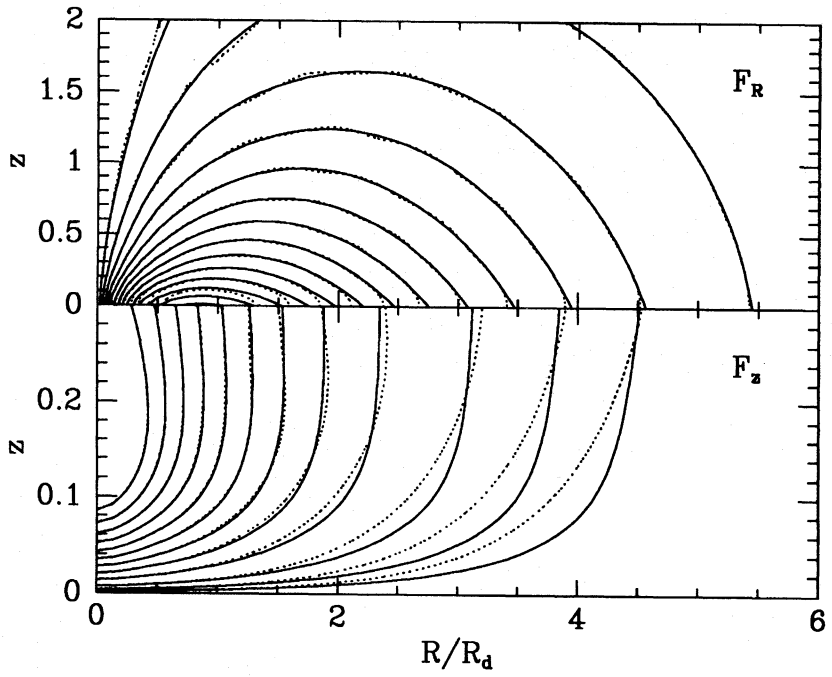

Figure 2. Representative contours of the resulting radial and vertical forces for an $l = 32$ pure multipole expansion (dotted line) and $l = 8$ expansion of the residue plus $\Psi^{\dagger}_{\rm disc}$ (solid line). Higher order expansions of the residue plus $\Psi^{\dagger}_{\rm disc}$ give the same result as the $l = 8$ expansion. The forces derived from the $l = 32$ pure multipole expansion still have significant errors at small z and intermediate R.

The power of this approach is illustrated in Fig. 1, where the results of a bare Legendre series for the potential of exponential discs are compared with the harmonics required to fit only the residual $\Psi_{\rm disc} - \Psi^{\dagger}_{\rm disc}$.

Note that it is important to have the resolution to follow accurately the *radial* variation in this density near the truncation radius, where the f'' term now magnifies the change in disc density.

Experience shows that, for a disc scale height of $0.15R_{\rm d}$ (i.e., exponential scale height $R_{\rm d}/13.3$ far from the disc plane), the residual density can be well-fitted with a series truncated at $l = 8$, whereas even a series expansion to $l = 32$ for the disc density alone does not give satisfactory results (Fig. 2).

2.5 Softening

In practice, N-body simulations employ a softened form of gravity in order to suppress two-body relaxation. Our models, as formulated above, will therefore not be in equilibrium under these modified forces. In principle, it is possible to solve a suitably modified Poisson equation to allow for the softening: a simply way to do this would be to smooth the density with the appropriate kernel before solving Poisson's equation. While possible, this extra smoothing step can be computationally expensive, and we have not implemented it in what follows. As will be seen, effects of this deficiency are small provided the smoothing length is smaller than relevant length scales in the model.

3 A SAMPLE MODEL

The advantage of galaxy models built from distribution functions is that the kinematics are fully specified. In this section we describe how this kinematic information can be used to set up equilibrium initial conditions for N-body simulations of disc–bulge–halo galaxies, and provide some sample results.

3.1 Setting up an N-body realization

We can generate an N-body realization of a galaxy by randomly sampling from the DFs for each component. The bulge and the halo are straightforward to generate since the systems are nearly spherical and the velocity ellipsoids are nearly isotropic. The methods are described in detail in Kuijken & Dubinski (1994). A particle's position is first determined by sampling from the density distribution. With this position, one can find the local maximum of the DF (at $(v_x, v_y, v_z) = (0,0,0)$) and then use the acceptance–rejection technique to find a velocity. This involves selecting the three components of the velocity at random from a

Table 1. Galaxy model parameters.

	Disc					Bulge			Halo				
Model	M_d	R_d	R_t	z_d	δR_{out}	Ψ_c	σ_b	ρ_b	Ψ_0	σ_0	q	C	R_a
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Sample	1.00	1.0	5.0	0.15	0.3	-2.0	0.50	10.0	-4.0	1.00	0.9	0.1	0.5
MW–A	0.87	1.0	5.0	0.10	0.5	-2.3	0.71	14.5	-4.6	1.00	1.0	0.1	0.8
MW–B	0.87	1.0	5.0	0.10	0.5	-2.9	0.71	14.5	-5.2	0.96	1.0	0.1	0.8
MW–C	0.87	1.0	5.0	0.10	0.5	-3.7	0.71	14.5	-6.0	0.93	1.0	0.1	0.8
MW–D	0.87	1.0	5.0	0.10	0.5	-4.7	0.71	14.5	-7.0	0.92	1.0	0.1	0.8

(1) disc mass, (2) disc scale radius, (3) disc truncation radius, (4) disc scale height, (5) disc truncation width, (6) bulge cutoff potential, (7) bulge velocity dispersion, (8) bulge central density, (9) halo central potential, (10) halo velocity dispersion, (11) halo potential flattening, (12) halo concentration, $C = R_c^2/R_K^2$ (Kuijken & Dubinski 1994), (13) characteristic halo radius.

Table 2. Galaxy model properties.

	Disc			Bulge		Halo	
Model	M	$\sigma_{R,0}$	R_e/R_d	M	R_e/R_d	M	R_e/R_d
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample	0.94	0.50	5.6	0.29	1.7	9.6	44.9
MW-A	0.82	0.47	6.0	0.42	1.0	5.2	21.8
MW-B	0.82	0.47	6.0	0.43	1.0	9.6	30.1
MW-C	0.82	0.47	6.0	0.43	1.0	19.8	44.0
MW-D	0.82	0.47	6.0	0.43	1.0	37.0	72.8

(1) disc mass, (2) disc central radial velocity dispersion, $\sigma_{R,0}$, (3) disc radial extent (radius where density drops to zero) in disc scale lengths, (4) bulge mass, (5) bulge radial extent in disc scale lengths, (6) halo mass, (7) halo radial extent in disc scale lengths.

velocity sphere with radius equal to the escape velocity. A random value, f_{ran}, of the DF is also chosen between 0 and the local maximum. If f_{ran} is less than the value of the DF at the chosen velocity then the velocity is accepted, otherwise it is rejected and another attempt is made.

Sampling from the disc DF is slightly more troublesome, since the disc is thin and the local velocity maximum of the disc DF must be found at each point (it is not at $\boldsymbol{v} = 0$ as for the bulge and halo). As before, we sample first from the density distribution to find particle positions and then from the DF to find velocities. Since the discs are generally warm, the maximum of the velocity distribution occurs at a point where the azimuthal component of velocity v_ϕ is less than the local circular velocity, v_{circ}. v_R and v_z are zero at the local maximum. We find the local velocity maximum for each particle position using standard methods (Press et al. 1993) and then use the acceptance–rejection technique as before to select a velocity.

3.2 An N-body realization

We set up an N-body realization of a sample model to test the validity of the equilibrium. This galaxy model has a disc:bulge mass ratio of 4:1 and a halo:disc mass ratio of 1.5:1 within 5 disc scale lengths. The total mass of the halo is about 10 times the disc mass with the halo extending to 40 disc scale lengths (see Tables 1 and 2 for the model parameters and resulting properties). The rotation curve is fairly flat out to a radius of 10 disc scale lengths (Fig. 3).

The disc is warm with a Toomre $Q = 1.7$ at the disc half-mass radius. The value of Q is fairly constant throughout the disc, though rising both in the centre within 1 scale length and near the edge where the surface density tapers off to zero (Fig. 4).

We performed various simulations to assess the validity of our approximate disc DF and the stability of the overall system. We used a tree code for all of the simulations (Barnes & Hut 1986; Hernquist 1987; Dubinski 1988). The critical opening angle was set to $\theta = 0.9$ with forces between cells and particles calculated to quadrupole order. The particle softening radius was set to 0.025 disc scale lengths (or 0.17 disc scale heights). The orbital time of the model at the disc half-mass radius was 13 units. We set the leapfrog timestep to 0.1 units and ran the models to $t = 96$ or approximately 7.5 orbital times. We set up four simulations:

(i) a disc of 40 000 test particles orbiting in the derived system potential;
(ii) a gravitating disc of 40 000 particles with a static bulge and halo potential;
(iii) a disc, bulge, and halo with 40 000, 10 000, and 50 000 particles respectively; and
(iv) a disc, bulge and halo with 80 000, 20 000, and 200 000 particles.

For all of the models, we followed the time evolution of the disc surface density profile, $\Sigma(R)$, the disc velocity dispersion profile

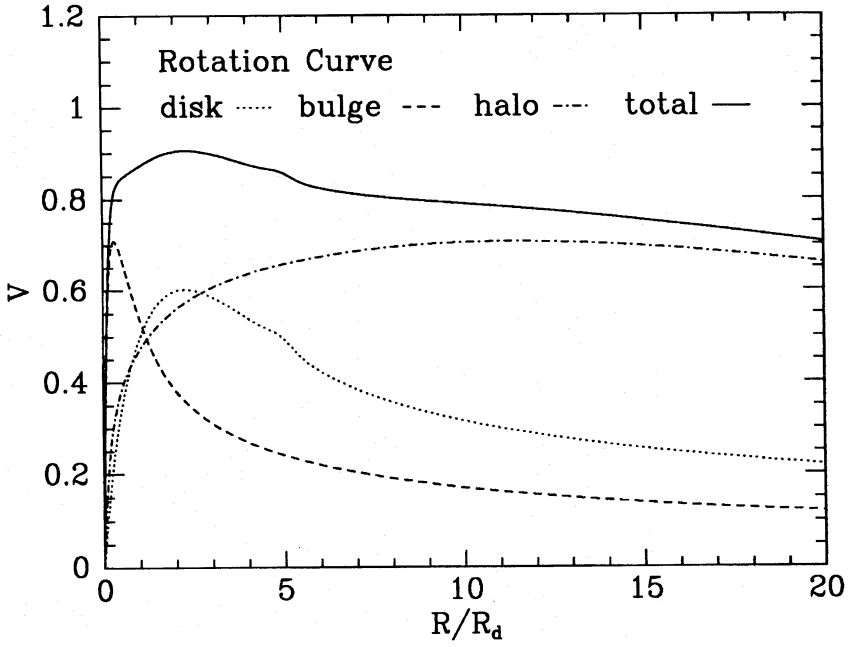

Figure 3. Rotation curve of a sample model.

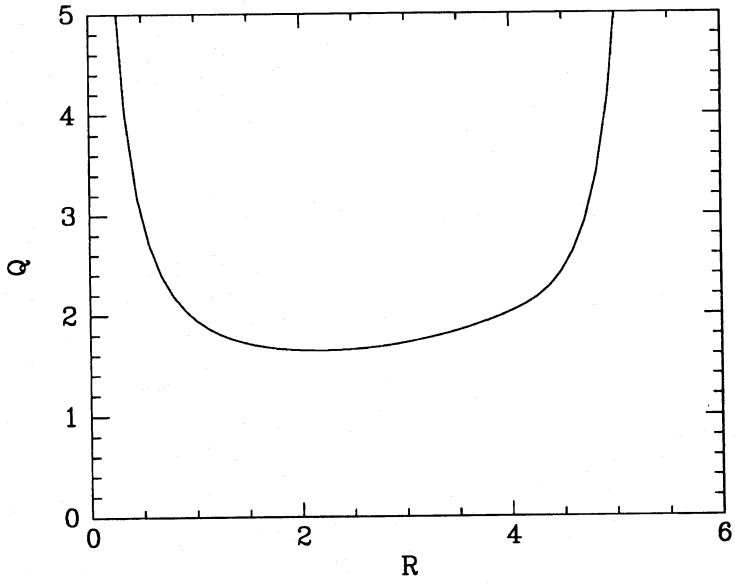

Figure 4. Toomre's stability parameter $Q = \sigma_r \kappa / 3.36 G\Sigma$ versus the radius.

averaged in rings, and the disc scale height versus radius. For the bulge and halo, we calculated the spherically averaged density profiles.

The first simulation provides a simple check of our DF, since the test particles should maintain a constant density and velocity dispersion profile in the total potential of the model. This test is most critical for the disc since we have approximated a third integral with the vertical energy E_z, and it is not guaranteed that this quantity is conserved sufficiently well to maintain a decent equilibrium. For our sample model, we find that the approximation works very well. There is essentially no change in the density profile (Fig. 5) and velocity dispersion profiles (Fig. 6).

In the second simulation, we made the disc 'live' while maintaining a static bulge and halo potential. We thereby avoid external perturbations from a noisy bulge and halo potential (which we will see below can significantly heat the disc). The surface density profile is nearly constant, once again deviating only slightly from the initial conditions. The disc is initially cool enough to suffer from spiral arm instabilities and this manifests itself as a gradual rise in σ_R and σ_ϕ throughout the disc (Fig. 6) (e.g. Sellwood & Carlberg 1984; Toomre & Kalnajs 1991). This self-heating is unavoidable in cool discs, even when set up from a DF

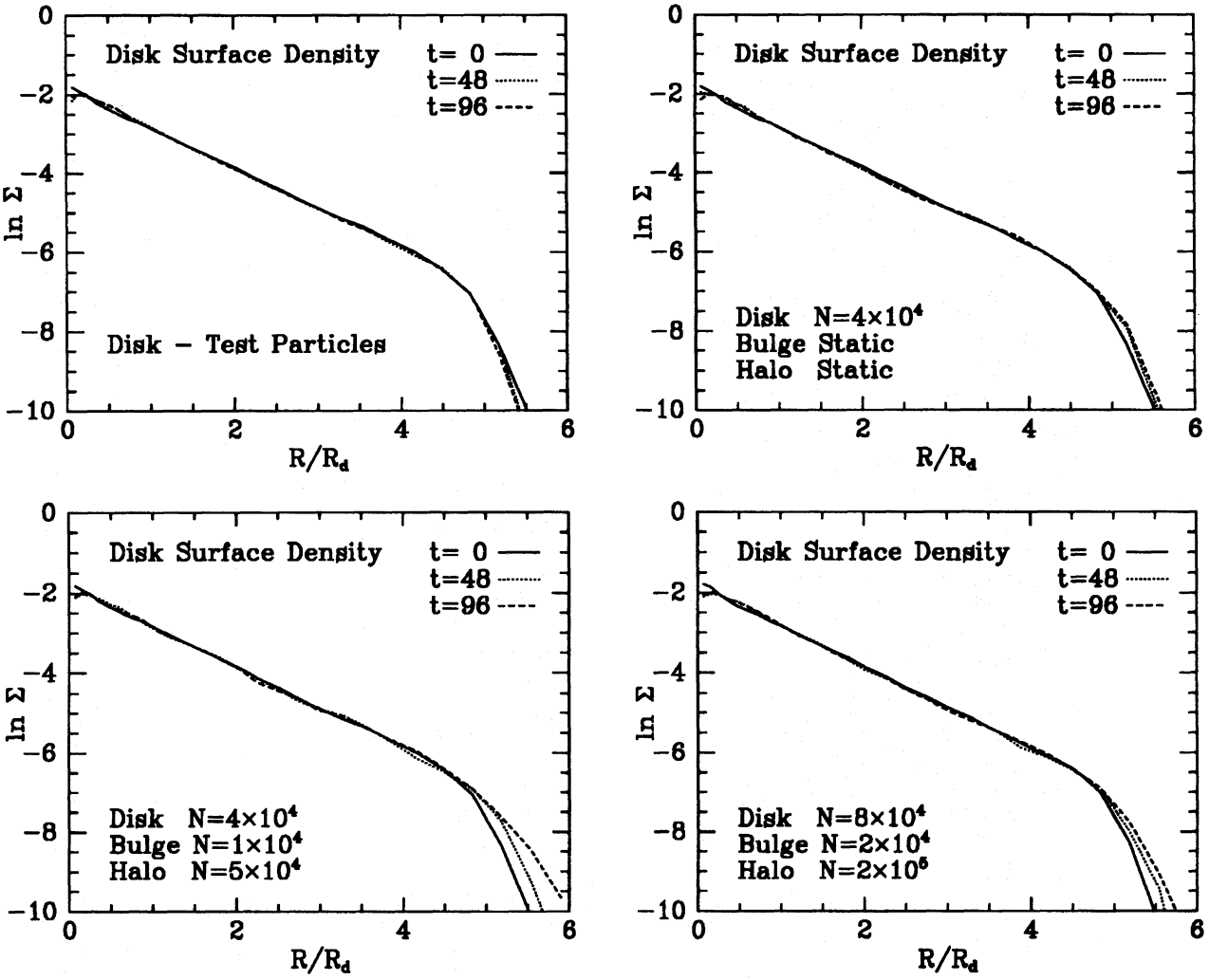

Figure 5. The disc surface density profile as a function of time for the four simulations as labelled.

that is formally in equilibrium. The vertical velocity dispersion, σ_z, remains fairly constant except at the edge of the disc where it is poorly sampled by particles and discreteness noise is large, and near the centre where the disc is hot and the assumed constancy of E_z along orbits built into the DF breaks down. Nevertheless, σ_z does not show any significant initial transients over most of the disc, suggesting that the disc DF represents a starting point very close to equilibrium. We do not expect spiral arm formation to heat the disc vertically, since the modes grow in the disc plane and couple weakly to vertical oscillations (e.g., Jenkins & Binney 1990). We also plotted the rms height of disc particles averaged in rings (Fig. 7). There is a slight increase with time in the average scale height which may arise either from the particle softening, which would weaken the local disc self-gravity and allow it to puff up, or from gradual heating by the disc's own discreteness.

The final two simulations, in which the halo and bulge are also live, show the strong effect of a coarse-grained halo on the disc evolution. A particle halo adds a significant source of disc heating (Hernquist 1993) when the number of halo particles is too small. The halo is effectively composed of massive black holes that bombard the disc as in the model of Lacey & Ostriker (1985). The continuous bombardment of halo particles heats the disc at a rate $\sigma^2 = (\sigma_0^2 + Dt)$, where D is proportional to the black hole mass and local halo density. The disc velocity ellipsoid therefore changes and the vertical scale height increases. This effect is strongest at the disc centre where the halo density is the greatest, and is weakened as the number of halo particles is increased, in agreement with the Lacey–Ostriker results. The graininess of the halo can also excite bending waves, especially in the outer disc where the disc density is low. The discs in simulation (iii) do indeed exhibit bending movements that resemble a flag waving in the breeze. This behaviour was less apparent in the larger simulation (iv). Again, as in simulation (ii), gravitational softening can also cause the disc to puff up.

Fig. 6 shows the velocity dispersion profiles for the simulations with particle bulges and haloes. We can see that there is significant heating for the smaller-N model over and above the heating of the disc in the static halo potential. The disc heating for the large-N halo is significantly smaller with the σ_z increasing only at disc radii with $R > 3R_d$. The disc scale height also does not

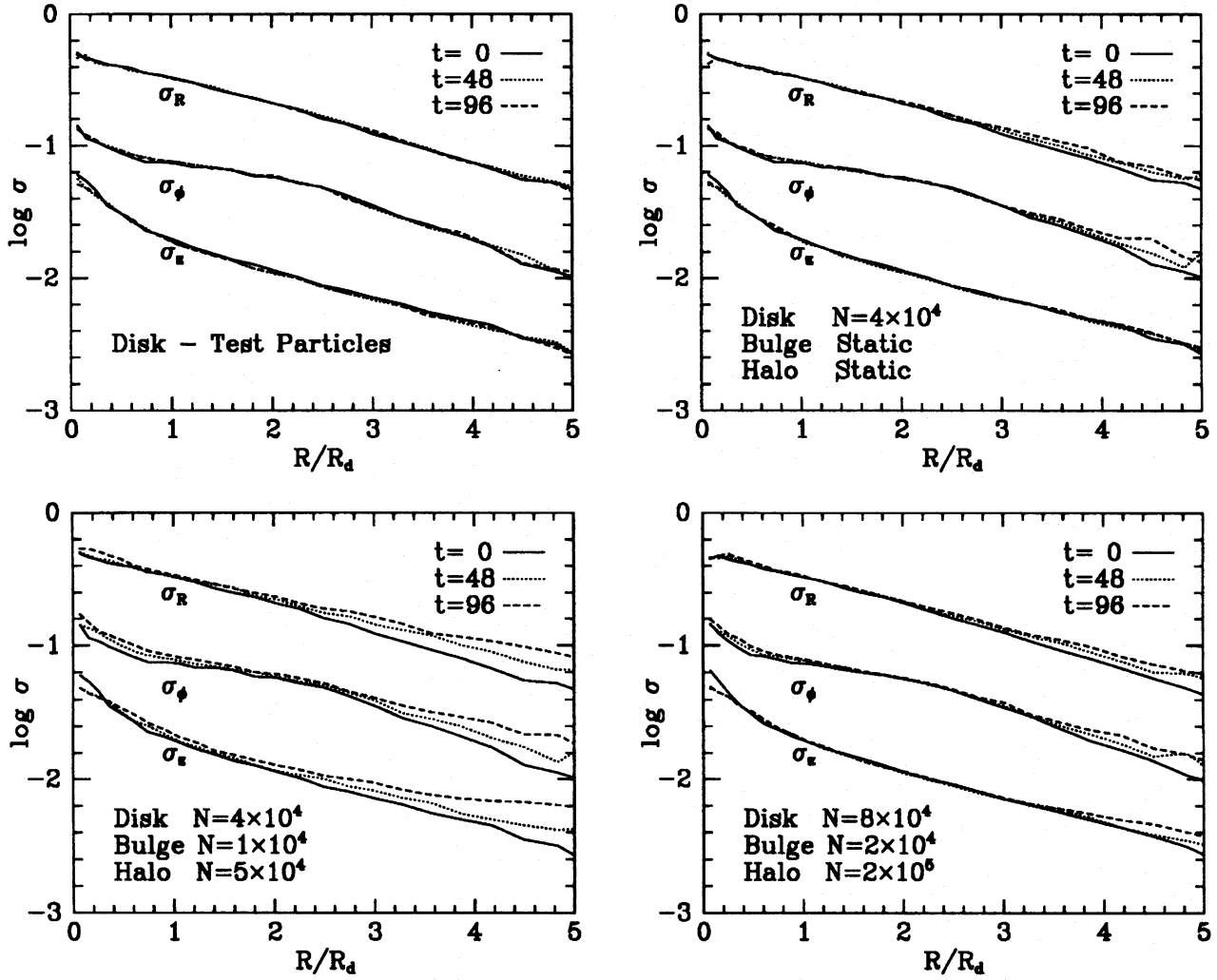

Figure 6. The velocity dispersion profiles as a function of time for the four simulations as labelled. To avoid overlap, σ_ϕ and σ_z are offset by 0.5 and 1.0 dex.

grow as much for the larger-N model in accord with expected heating rates (Fig. 7). The models are apparently converging to a smooth equilibrium limit as we increase N.

Unlike the disc, the bulge and halo maintain their integrity during the simulations. Fig. 8 shows the constancy of the density profiles of the models at three times. The slight drop in the density in the bulge within $r < 0.1$ can be attributed to the particle softening radius $\epsilon = 0.025$. The reduced gravity from the softening allows the bulge to expand a little at the centre and therefore drop in density. Despite this minor defect, the bulge and halo stay in equilibrium.

A more sensitive test of stability is to examine the distribution of the particles in various phase-space projections. In Fig. 9 we compare the density in three such projections at the beginning and end ($t = 96$) of the largest simulation. Only very minor changes are seen.

At late times (5–8 orbital times), a modest bar begins to develop within 1 scale radius in the completely live models. The bar forms sooner in the smaller-N models. The seeds of these bar instabilities probably originate in the discreteness noise of the halo also observed by Hernquist (private communication). The onset of bar formation can only be delayed by increasing N.

In summary, a disc DF with E_z as an approximate third integral results in a disc–bulge–halo galaxy model very close to formal equilibrium. We can therefore follow subtle changes in the disc structure and kinematics resulting from a variety of instabilities.

In the next section, we find a set of model parameters which produce a mass model closely resembling the Galaxy.

4 MODELS OF THE MILKY WAY

In Table 1, we present the parameters for generating a sequence of four models, MW-A, B, C, and D, which have mass distributions and rotation curves closely resembling those of the Milky Way within 5 scale radii. The disc and bulge mass distributions are the

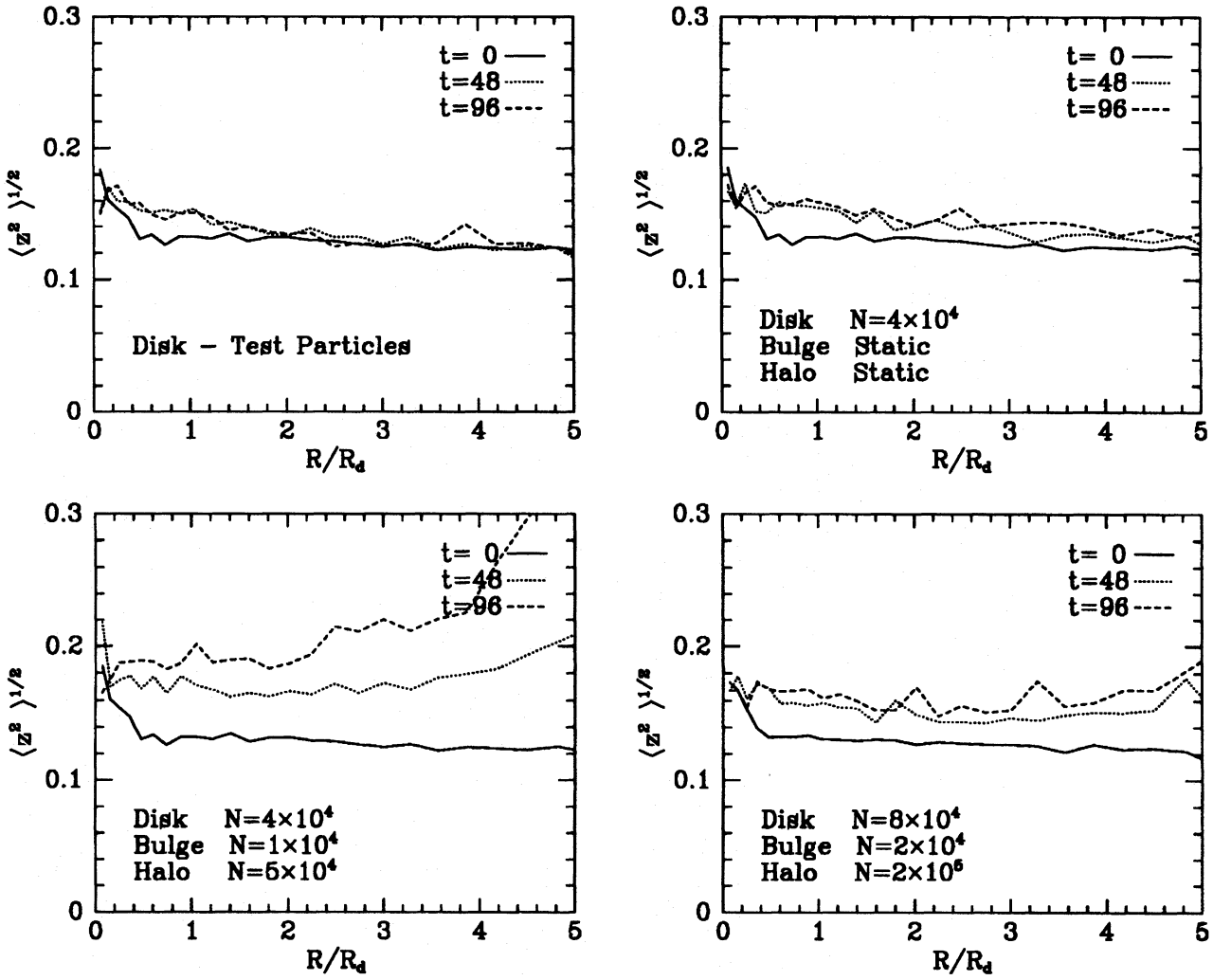

Figure 7. The disc scale height $\langle z^2 \rangle^{1/2}$ as a function of time for the four simulations as labelled.

same for each model with mass and extent of the halo increasing through the sequence (Table 2). The haloes are all chosen with $q = 1.0$, though they are slightly squashed in the self-consistent galaxy models. Model MW-D has the halo with the largest mass and has the most realistic representation of the outer Galaxy. These models were found by trial and error and renormalized so that the flat portion of the rotation curve had $V_c \approx 1.0$. The contributions to the radial acceleration in the solar neighbourhood ($R = 1.8R_d$) from disc, bulge and halo are comparable in these models, as found by Kuijken & Gilmore (1989) in their study of the local disc surface density. The natural units for length, velocity, and mass for these dimensionless models are $R_d = 4.5$ kpc, $V = 220$ km s^{-1}, and $M = 5.1 \times 10^{10}$ M$_\odot$. The central velocity dispersion was chosen so that the observed radial velocity dispersion of 42 km s^{-1} at the solar radius ($R = 1.8R_d$) would be reproduced in the model. Figs 10 and 11 show the rotation curves out to 5 scale radii and 50 scale radii respectively.

We tested the stability of the models against bar formation and found that they were not immediately unstable. A small bar within 1 scale radius did form after four orbital times, though increasing the number of particles delayed the formation to later times. It seems that the observed instability was due to the discreteness of the simulation rather than the intrinsic mass distribution. In the limit of large N, these particular models are probably stable against bar formation. The fact that we observe a bar in the Galactic Centre (e.g. Weinberg 1992) might imply a larger disc:halo mass ratio within the solar circle than these models assume.

5 SUMMARY

We have described a method for setting up self-consistent galaxy models with a disc, bulge and halo where each component is described by a separate phase-space distribution function (DF). The halo DF depends on particle energy E and angular momentum L_z, and has the form of a lowered Evans model (Kuijken & Dubinski 1994), resulting in a flattened, finite-radius model. The bulge DF is a King model (King 1966) which is a function only of E. We introduce a new DF for the disc, a version of the planar DF of

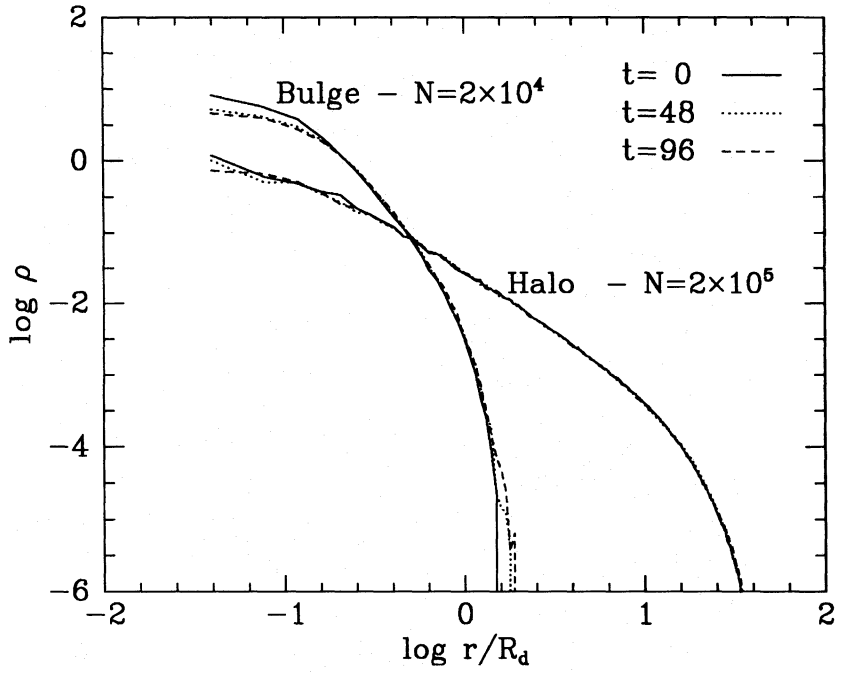

Figure 8. The bulge and halo density profiles for the largest simulation as a function of time. The bulge relaxes to a slightly lower density for $r < 0.1$ probably as a result of the gravitational softening, but overall the density profiles remain constant.

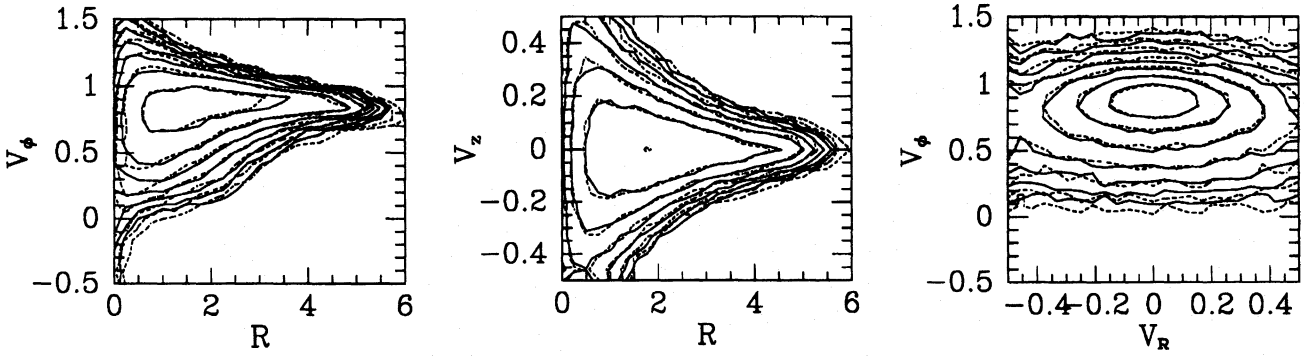

Figure 9. Various phase-space projections of the largest simulation of the sample model, at the beginning (solid contours) and end (dashed contours) of the run. For clarity, only the disc particles are shown. Contours are representative number densities declining by factors of 2 from the central maxima. The phase-space density does not change significantly over the course of the run, demonstrating the good quality of the initial equilibrium.

Shu (1969) modified to include vertical structure. It is a function of E, L_z and E_z, the vertical energy. The disc DF is approximate in that the third 'integral', E_z, introduced to describe the vertical motion, is not conserved exactly. Nevertheless, in practice the disc DF yields a model very close to equilibrium. The three components respond to the total gravitational potential of the model. The main advantages of these models over previous ones are the following.

(i) The flattening and rotation of the halo can be specified in the lowered Evans model DF, which allows examination of new problems where these parameters are important, such as the dynamics of galactic warps and satellite orbital evolution.

(ii) The disc does not suffer from additional transient adjustments at start-up beyond the expected (and unavoidable) spiral and bar instabilities. The potential of the system and velocity ellipsoids (particularly those of the disc) will therefore not change significantly from the initial state (provided the softening employed in the N-body code is sufficiently small). These models are therefore useful for studying problems involving the effects of small perturbations on the disc, such as disc heating by sinking satellites or external tidal fields.

(iii) Since the DF is known explicitly, many quantities (line profiles, for example) can be calculated directly without the need for simulation.

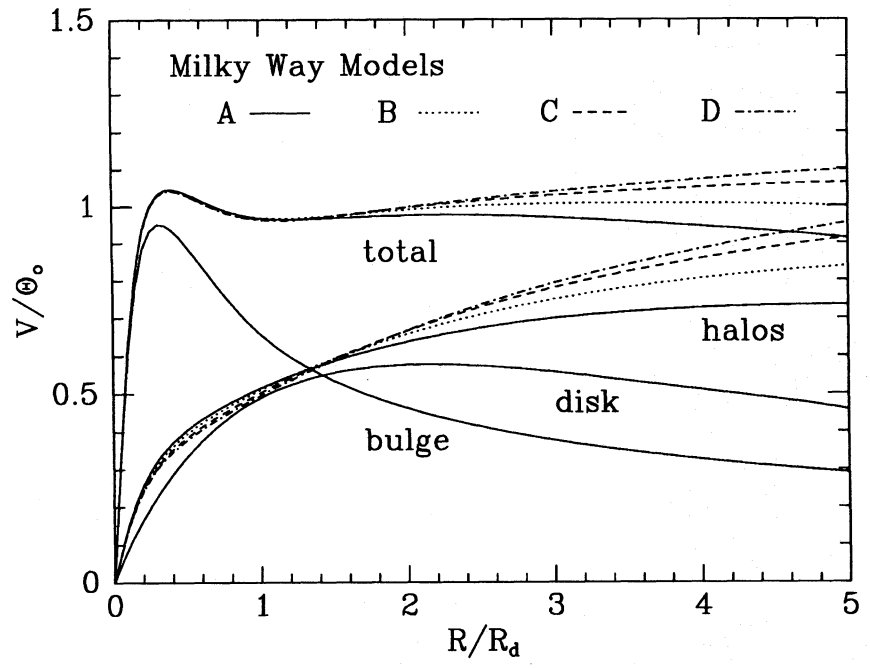

Figure 10. The rotation curves for the Milky Way models, MW-A, B, C, and D, showing contributions from the disc, bulge and halo in the inner regions within $R < 5R_d$.

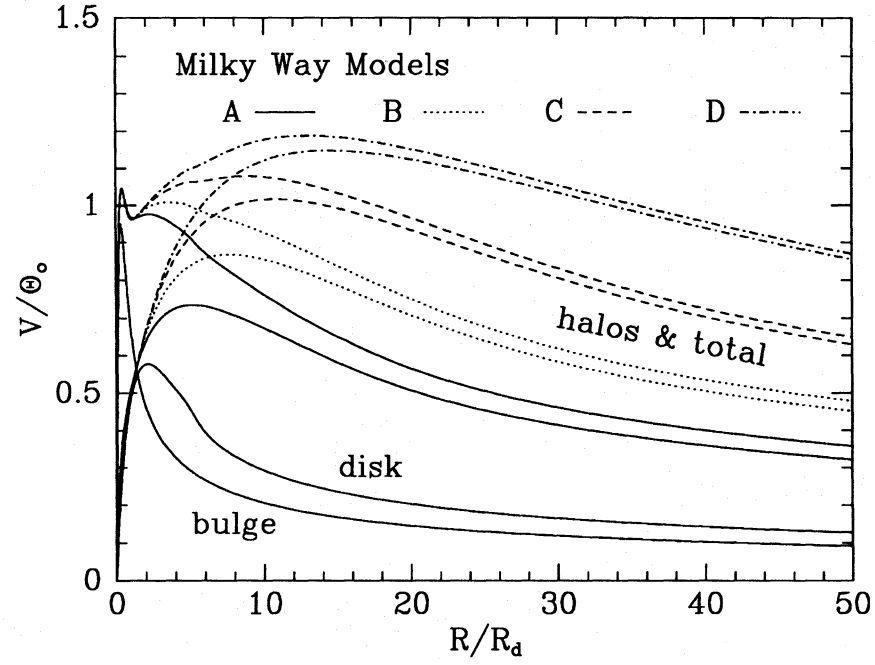

Figure 11. The rotation curves for the Milky Way models, MW-A, B, C, and D, showing contributions from the disc, bulge and halo in the outer regions out to $R = 50R_d$.

These same techniques can be generalized to models with different bulge and halo DFs and discs with different density distributions. Of course, equilibrium does not imply stability, and it is by no means guaranteed that these models are free from spiral and bar instabilities. In practice, though, models can be found that are apparently stable over many rotation periods. Disc heating by the halo is troublesome, and ultimately can only be avoided by going to very large N ($\sim 10^6$) if we wish to evolve a system for a Hubble time.

REFERENCES

Barnes J., 1988, ApJ, 331, 699
Barnes J., Hut P., 1986, Nat, 324, 446
Binney J., 1987, in Gilmore G., Carswell B., eds, The Galaxy. Reidel, Dordrecht, p. 399
Bottema R., 1993, A&A, 275, 16
Dubinski J., 1988, M.Sc. Thesis, University of Toronto
Dubinski J., Kuijken K., 1995, ApJ, 442, 492
Evans N.W., 1993, MNRAS, 260, 191
Hernquist L., 1987, ApJS, 64, 715
Hernquist L., 1993, ApJS, 86, 389
Hernquist L., Weinberg M., 1992, ApJ, 400, 80
Jenkins A., Binney J., 1990, MNRAS, 245, 305
Kalnajs A.J., 1976, ApJ, 175, 63
Kent S., de Zeeuw T., 1991, AJ, 102, 1994
King I.R., 1966, AJ, 67, 471
Kuijken K., Dubinski J., 1994, MNRAS, 269, 13
Kuijken K., Gilmore G., 1989, MNRAS, 239, 605
Kuijken K., Tremaine S., 1992, in Sundelius B., ed., Dynamics of Disc Galaxies. Göteborg University Press, Göteborg, p. 71
Lacey C.G., Ostriker J.P., 1985, ApJ, 299, 633
Merritt D., Sellwood J., 1994, ApJ, 425, 551
Oort J.H., 1965, in Blaauw A., Schmidt A., eds, Stars and Stellar Systems, vol. 5. Univ. Chicago Press, Chicago, chapter 21
Prendergast K.H., Tomer E., 1970, AJ, 75, 674
Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P., 1993, Numerical Recipes in C. Cambridge Univ. Press, Cambridge, section 7
Quinn P.J., Hernquist L., Fullagar D.P., 1993, ApJ, 403, 74
Rowley G., 1988, ApJ, 331, 124
Sellwood J.A., Carlberg R.G., 1984, MNRAS, 282, 61
Sellwood J.A., Merritt D., 1994, ApJ, 425, 530
Shu F.H., 1969, ApJ, 158, 505
Toomre A., Kalnajs A., 1991, in Sundelius B., ed., Dynamics of Disc Galaxies. Göteborg University Press, Göteborg, p. 341
Weinberg M.D., 1992, ApJ, 384, 81
Wielen R., 1974, Highlights of Astronomy, 3, 395
van der Kruit P.C., Searle L., 1981, A&A, 95, 101

This paper has been produced using the Royal Astronomical Society/Blackwell Science LaTeX style file.