

## 2

# Potential Theory

Much of the mass of a galaxy resides in stars. To compute the gravitational potential of a large collection of stars, we should in principle simply add the point-mass potentials of all the stars together. Of course, this is not practicable for the  $\approx 10^{11}$  stars in a typical galaxy, and for most purposes it is sufficient to model the potential by smoothing the mass density in stars on a scale that is small compared to the size of the galaxy, but large compared to the mean distance between stars. In particular, in §1.2.1 we saw that we obtain an excellent approximation to the orbit of a single star in a galaxy by treating the star as a test particle that moves in a smooth potential of this kind. In this chapter we show how the force field of such an idealized galaxy can be calculated.

The chapter is divided into eight sections. We start by reviewing some general results on potential theory, and then in §2.2 we specialize to discuss the simplest potentials, those of spherical bodies. In §2.3 we describe flattened density distributions that also have simple potentials. These special systems give us insight into the potentials of real galaxies and provide useful prototypes, but they are not adequate for accurate modeling. Therefore in §2.4 and §2.5 we describe a variety of general techniques for computing the potentials of aspherical bodies. The potentials of razor-thin disks form an important limiting case, which we discuss in §2.6. In §2.7 we use these results to examine the potential of the Milky Way. In §2.8 and §2.9 we describe how gravitational potentials are found in computer simulations of stellar systems.

The discussion in §§2.4 to 2.6 is rather mathematical, and readers who are willing to take a few results on trust may prefer to move straight from §2.3 to §2.7.

## 2.1 General results

Our goal is to calculate the force  $\mathbf{F}(\mathbf{x})$  on a particle of mass  $m_s$  at position  $\mathbf{x}$  that is generated by the gravitational attraction of a distribution of mass  $\rho(\mathbf{x}')$ . According to Newton's inverse-square law of gravitation, the force  $\mathbf{F}(\mathbf{x})$  may be obtained by summing the small contributions

$$\delta\mathbf{F}(\mathbf{x}) = Gm_s \frac{\mathbf{x}' - \mathbf{x}}{|\mathbf{x}' - \mathbf{x}|^3} \delta m(\mathbf{x}') = Gm_s \frac{\mathbf{x}' - \mathbf{x}}{|\mathbf{x}' - \mathbf{x}|^3} \rho(\mathbf{x}') d^3\mathbf{x}' \quad (2.1)$$

to the overall force from each small element of volume  $d^3\mathbf{x}'$  located at  $\mathbf{x}'$ . Thus

$$\mathbf{F}(\mathbf{x}) = m_s \mathbf{g}(\mathbf{x}) \quad \text{where} \quad \mathbf{g}(\mathbf{x}) \equiv G \int d^3\mathbf{x}' \frac{\mathbf{x}' - \mathbf{x}}{|\mathbf{x}' - \mathbf{x}|^3} \rho(\mathbf{x}') \quad (2.2)$$

is the **gravitational field**, the force per unit mass.

If we define the **gravitational potential**  $\Phi(\mathbf{x})$  by

$$\Phi(\mathbf{x}) \equiv -G \int d^3\mathbf{x}' \frac{\rho(\mathbf{x}')}{|\mathbf{x}' - \mathbf{x}|}, \quad (2.3)$$

and notice that

$$\nabla_{\mathbf{x}} \left( \frac{1}{|\mathbf{x}' - \mathbf{x}|} \right) = \frac{\mathbf{x}' - \mathbf{x}}{|\mathbf{x}' - \mathbf{x}|^3}, \quad (2.4)$$

we find that we may write  $\mathbf{g}$  as

$$\begin{aligned} \mathbf{g}(\mathbf{x}) &= \nabla_{\mathbf{x}} \int d^3\mathbf{x}' \frac{G\rho(\mathbf{x}')}{|\mathbf{x}' - \mathbf{x}|} \\ &= -\nabla\Phi, \end{aligned} \quad (2.5)$$

where for brevity we have dropped the subscript  $\mathbf{x}$  on the gradient operator  $\nabla$ .

The potential is useful because it is a scalar field that is easier to visualize than the vector gravitational field but contains the same information. Also, in many situations the easiest way to obtain  $\mathbf{g}$  is first to calculate the potential and then to take its gradient.

If we take the divergence of equation (2.2), we find

$$\nabla \cdot \mathbf{g}(\mathbf{x}) = G \int d^3\mathbf{x}' \nabla_{\mathbf{x}} \cdot \left( \frac{\mathbf{x}' - \mathbf{x}}{|\mathbf{x}' - \mathbf{x}|^3} \right) \rho(\mathbf{x}'). \quad (2.6)$$

Now

$$\nabla_{\mathbf{x}} \cdot \left( \frac{\mathbf{x}' - \mathbf{x}}{|\mathbf{x}' - \mathbf{x}|^3} \right) = -\frac{3}{|\mathbf{x}' - \mathbf{x}|^3} + \frac{3(\mathbf{x}' - \mathbf{x}) \cdot (\mathbf{x}' - \mathbf{x})}{|\mathbf{x}' - \mathbf{x}|^5}. \quad (2.7)$$

When  $\mathbf{x}' - \mathbf{x} \neq 0$  we may cancel the factor  $|\mathbf{x}' - \mathbf{x}|^2$  from top and bottom of the last term in this equation to conclude that

$$\nabla_{\mathbf{x}} \cdot \left( \frac{\mathbf{x}' - \mathbf{x}}{|\mathbf{x}' - \mathbf{x}|^3} \right) = 0 \quad (\mathbf{x}' \neq \mathbf{x}). \quad (2.8)$$

Therefore, any contribution to the integral of equation (2.6) must come from the point  $\mathbf{x}' = \mathbf{x}$ , and we may restrict the volume of integration to a small sphere of radius  $h$  centered on this point. Since, for sufficiently small  $h$ , the density will be almost constant through this volume, we can take  $\rho(\mathbf{x}')$  out of the integral. The remaining terms of the integrand may then be arranged as follows:

$$\begin{aligned} \nabla \cdot \mathbf{g}(\mathbf{x}) &= G\rho(\mathbf{x}) \int_{|\mathbf{x}' - \mathbf{x}| \leq h} d^3\mathbf{x}' \nabla_{\mathbf{x}} \cdot \left( \frac{\mathbf{x}' - \mathbf{x}}{|\mathbf{x}' - \mathbf{x}|^3} \right) \\ &= -G\rho(\mathbf{x}) \int_{|\mathbf{x}' - \mathbf{x}| \leq h} d^3\mathbf{x}' \nabla_{\mathbf{x}'} \cdot \left( \frac{\mathbf{x}' - \mathbf{x}}{|\mathbf{x}' - \mathbf{x}|^3} \right) \\ &= -G\rho(\mathbf{x}) \int_{|\mathbf{x}' - \mathbf{x}| = h} d^2\mathbf{S}' \cdot \frac{(\mathbf{x}' - \mathbf{x})}{|\mathbf{x}' - \mathbf{x}|^3}. \end{aligned} \quad (2.9a)$$

The last step in this sequence uses the divergence theorem to convert the volume integral into a surface integral (eq. B.43). Now on the sphere  $|\mathbf{x}' - \mathbf{x}| = h$  we have  $d^2\mathbf{S}' = (\mathbf{x}' - \mathbf{x})h d^2\Omega$ , where  $d^2\Omega$  is a small element of solid angle. Hence equation (2.9a) becomes

$$\nabla \cdot \mathbf{g}(\mathbf{x}) = -G\rho(\mathbf{x}) \int d^2\Omega = -4\pi G\rho(\mathbf{x}). \quad (2.9b)$$

If we substitute from equation (2.5) for  $\nabla \cdot \mathbf{g}$ , we obtain **Poisson's equation** relating the potential  $\Phi$  to the density  $\rho$ ;

$$\nabla^2\Phi = 4\pi G\rho. \quad (2.10)$$

This is a differential equation that can be solved for  $\Phi(\mathbf{x})$  given  $\rho(\mathbf{x})$  and an appropriate boundary condition.<sup>1</sup> For an isolated system the boundary condition is  $\Phi \rightarrow 0$  as  $|\mathbf{x}| \rightarrow \infty$ . The potential given by equation (2.3) automatically satisfies this boundary condition. Poisson's equation provides

---

<sup>1</sup> Using the physically correct boundary condition is essential: for example, if  $\Phi(\mathbf{x})$  is a solution, then so is  $\Phi(\mathbf{x}) + \mathbf{k} \cdot \mathbf{x}$ , with  $\mathbf{k}$  an arbitrary constant vector, but the corresponding gravitational fields differ by  $\mathbf{k}$ .

a route to  $\Phi$  and then to  $\mathbf{g}$  that is often more convenient than equations (2.2) or (2.3). In the special case  $\rho = 0$  Poisson's equation becomes **Laplace's equation**,

$$\nabla^2 \Phi = 0. \quad (2.11)$$

If we integrate both sides of equation (2.10) over an arbitrary volume containing total mass  $M$ , and then apply the divergence theorem (eq. B.43), we obtain

$$4\pi G M = 4\pi G \int d^3\mathbf{x} \rho = \int d^3\mathbf{x} \nabla^2 \Phi = \int d^2\mathbf{S} \cdot \nabla \Phi. \quad (2.12)$$

This result is **Gauss's theorem**, which states that *the integral of the normal component of  $\nabla \Phi$  over any closed surface equals  $4\pi G$  times the mass contained within that surface.*

Since  $\mathbf{g}$  is determined by the gradient of a potential, the gravitational field is conservative, that is, the work done against gravitational forces in moving two stars from infinity to a given configuration is independent of the path along which they are moved, and is defined to be the potential energy of the configuration (Appendix D.1). Similarly, the work done against gravitational forces in assembling an arbitrary continuous distribution of mass  $\rho(\mathbf{x})$  is independent of the details of how the mass distribution was assembled, and is defined to be equal to the **potential energy** of the mass distribution. An expression for the potential energy can be obtained by the following argument.

Suppose that some of the mass is already in place so that the density and potential are  $\rho(\mathbf{x})$  and  $\Phi(\mathbf{x})$ . If we now bring in a additional small mass  $\delta m$  from infinity to position  $\mathbf{x}$ , the work done is  $\delta m \Phi(\mathbf{x})$ . Thus, if we add a small increment of density  $\delta \rho(\mathbf{x})$ , the change in potential energy is

$$\delta W = \int d^3\mathbf{x} \delta \rho(\mathbf{x}) \Phi(\mathbf{x}). \quad (2.13)$$

According to Poisson's equation the resulting change in potential  $\delta \Phi(\mathbf{x})$  satisfies  $\nabla^2(\delta \Phi) = 4\pi G(\delta \rho)$ , so

$$\delta W = \frac{1}{4\pi G} \int d^3\mathbf{x} \Phi \nabla^2(\delta \Phi). \quad (2.14)$$

Using the divergence theorem in the form (B.45), we may write this as

$$\delta W = \frac{1}{4\pi G} \int \Phi \nabla(\delta \Phi) \cdot d^2\mathbf{S} - \frac{1}{4\pi G} \int d^3\mathbf{x} \nabla \Phi \cdot \nabla(\delta \Phi), \quad (2.15)$$

where the surface integral vanishes because  $\Phi \propto r^{-1}$ ,  $|\nabla \delta \Phi| \propto r^{-2}$  as  $r \rightarrow \infty$ , so the integrand  $\propto r^{-3}$  while the total surface area  $\propto r^2$ . But  $\nabla \Phi \cdot \nabla(\delta \Phi) = \frac{1}{2} \delta(\nabla \Phi \cdot \nabla \Phi) = \frac{1}{2} \delta(|\nabla \Phi|^2)$ . Hence

$$\delta W = -\frac{1}{8\pi G} \delta \left( \int d^3\mathbf{x} |\nabla \Phi|^2 \right). \quad (2.16)$$

If we now sum up all of the contributions  $\delta W$ , we have a simple expression for the potential energy,

$$W = -\frac{1}{8\pi G} \int d^3\mathbf{x} |\nabla\Phi|^2. \quad (2.17)$$

To obtain an alternative expression for  $W$ , we again apply the divergence theorem and replace  $\nabla^2\Phi$  by  $4\pi G\rho$  to obtain

$$W = \frac{1}{2} \int d^3\mathbf{x} \rho(\mathbf{x})\Phi(\mathbf{x}). \quad (2.18)$$

**The potential-energy tensor** In §4.8.3 we shall encounter the tensor  $\mathbf{W}$  that is defined by

$$W_{jk} \equiv - \int d^3\mathbf{x} \rho(\mathbf{x}) x_j \frac{\partial\Phi}{\partial x_k}, \quad (2.19)$$

where  $\rho$  and  $\Phi$  are the density and potential of some body, and the integral is to be taken over all space. We now deduce some useful properties of  $\mathbf{W}$ , which is known as the **Chandrasekhar potential-energy tensor**.<sup>2</sup>

If we substitute for  $\Phi$  from equation (2.3),  $\mathbf{W}$  becomes

$$W_{jk} = G \int d^3\mathbf{x} \rho(\mathbf{x}) x_j \frac{\partial}{\partial x_k} \int d^3\mathbf{x}' \frac{\rho(\mathbf{x}')}{|\mathbf{x}' - \mathbf{x}|}. \quad (2.20)$$

Since the range of the integration over  $\mathbf{x}'$  does not depend on  $\mathbf{x}$ , we may carry the differentiation inside the integral to find

$$W_{jk} = G \int d^3\mathbf{x} \int d^3\mathbf{x}' \rho(\mathbf{x}) \rho(\mathbf{x}') \frac{x_j (x'_k - x_k)}{|\mathbf{x}' - \mathbf{x}|^3}. \quad (2.21a)$$

Furthermore, since  $\mathbf{x}$  and  $\mathbf{x}'$  are dummy variables of integration, we may relabel them and write

$$W_{jk} = G \int d^3\mathbf{x}' \int d^3\mathbf{x} \rho(\mathbf{x}') \rho(\mathbf{x}) \frac{x'_j (x_k - x'_k)}{|\mathbf{x} - \mathbf{x}'|^3} \quad (2.21b)$$

Finally, on interchanging the order of integration in equation (2.21b) and adding the result to equation (2.21a), we obtain

$$W_{jk} = -\frac{1}{2} G \int d^3\mathbf{x} \int d^3\mathbf{x}' \rho(\mathbf{x}) \rho(\mathbf{x}') \frac{(x'_j - x_j)(x'_k - x_k)}{|\mathbf{x}' - \mathbf{x}|^3}. \quad (2.22)$$

---

<sup>2</sup> Subramanyan Chandrasekhar (1910–1995) was educated in India and England and spent most of his career at the University of Chicago. He discovered the Chandrasekhar limit, the maximum mass of a white dwarf star, and elucidated the concept of dynamical friction in astrophysics (§8.1). He shared the 1983 Nobel Prize in Physics with W. A. Fowler.

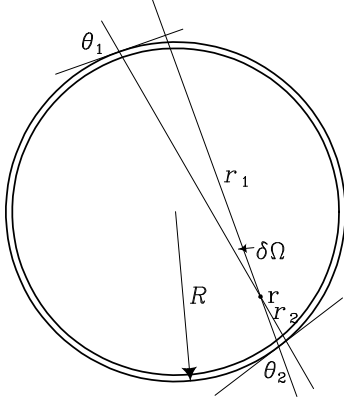


Figure 2.1 Proof of Newton's first theorem.

From this expression we draw the important inference that the tensor  $\mathbf{W}$  is **symmetric**, that is, that  $W_{jk} = W_{kj}$ . If the body is flattened along some axis, say the  $x_3$  axis,  $W_{33}$  will be smaller than the other components because for most pairs of matter elements,  $|x_3 - x'_3| < |x_1 - x'_1|$  or  $|x_2 - x'_2|$ .

When we take the **trace** of both sides of equation (2.22), we find

$$\begin{aligned} \text{trace}(\mathbf{W}) &\equiv \sum_{j=1}^3 W_{jj} = -\frac{1}{2}G \int d^3\mathbf{x} \rho(\mathbf{x}) \int d^3\mathbf{x}' \frac{\rho(\mathbf{x}')}{|\mathbf{x}' - \mathbf{x}|} \\ &= \frac{1}{2} \int d^3\mathbf{x} \rho(\mathbf{x}) \Phi(\mathbf{x}). \end{aligned} \quad (2.23)$$

Comparing this with equation (2.18) we see that  $\text{trace}(\mathbf{W})$  is simply the total gravitational potential energy  $W$ . Taking the trace of (2.19) we have

$$W = - \int d^3\mathbf{x} \rho \mathbf{x} \cdot \nabla \Phi, \quad (2.24)$$

which provides another useful expression for the potential energy of a body.

## 2.2 Spherical systems

### 2.2.1 Newton's theorems

Newton proved two results that enable us to calculate the gravitational potential of any spherically symmetric distribution of matter easily:

**Newton's first theorem** *A body that is inside a spherical shell of matter experiences no net gravitational force from that shell.*

**Newton's second theorem** *The gravitational force on a body that lies outside a spherical shell of matter is the same as it would be if all the shell's matter were concentrated into a point at its center.*

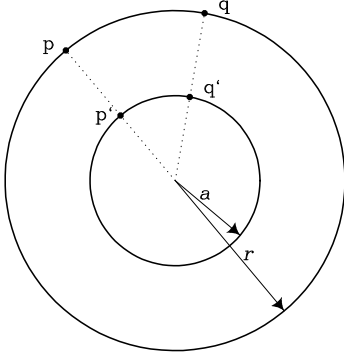
**Figure 2.2** Proof of Newton's second theorem.

Figure 2.1 illustrates the proof of Newton's first theorem. Consider the cone associated with a small solid angle  $\delta\Omega$  centered on the point  $\mathbf{r}$ . This cone intersects the spherical shell of matter at two points, at distances  $r_1$  and  $r_2$  from  $\mathbf{r}$ . Elementary geometrical considerations assure us that the angles  $\theta_1$  and  $\theta_2$  are equal, and therefore that the masses  $\delta m_1$  and  $\delta m_2$  contained within  $\delta\Omega$  where it intersects the shell are in the ratio  $\delta m_1/\delta m_2 = (r_1/r_2)^2$ . Hence  $\delta m_2/r_2^2 = \delta m_1/r_1^2$  and a particle placed at  $\mathbf{r}$  is attracted equally in opposite directions. Summing over all cones centered on  $\mathbf{r}$ , one concludes that the body at  $\mathbf{r}$  experiences no net force from the shell.◁

An important corollary of Newton's first theorem is that the gravitational potential inside an empty spherical shell is constant because  $\nabla\Phi = -\mathbf{g} = 0$ . Thus we may evaluate the potential  $\Phi(\mathbf{r})$  inside the shell by calculating the integral expression (2.3) for  $\mathbf{r}$  located at any interior point. The most convenient place for  $\mathbf{r}$  is the center of the shell, for then all points on the shell are at the same distance  $R$ , and one immediately has

$$\Phi = -\frac{GM}{R}. \quad (2.25)$$

The proof of his second theorem eluded Newton for more than ten years. Yet with hindsight it is easy. The trick (Figure 2.2) is to compare the potential  $\Phi$  at a point  $\mathbf{p}$  located a distance  $r$  from the center of a spherical inner shell of mass  $M$  and radius  $a$  ( $r > a$ ), with the potential  $\Phi'$  at a point  $\mathbf{p}'$  located a distance  $a$  from the center of an outer shell of mass  $M$  and radius  $r$ . Consider the contribution  $\delta\Phi$  to the potential at  $\mathbf{p}$  from the portion of the inner shell with solid angle  $\delta\Omega$  located at  $\mathbf{q}'$ . Evidently

$$\delta\Phi = -\frac{GM}{|\mathbf{p} - \mathbf{q}'|} \frac{\delta\Omega}{4\pi}. \quad (2.26a)$$

But the contribution  $\delta\Phi'$  of the matter in the outer shell near  $\mathbf{q}$  to the potential at  $\mathbf{p}'$  is

$$\delta\Phi' = -\frac{GM}{|\mathbf{p}' - \mathbf{q}|} \frac{\delta\Omega}{4\pi}. \quad (2.26b)$$

Finally, as  $|\mathbf{p} - \mathbf{q}'| = |\mathbf{p}' - \mathbf{q}|$  by symmetry, it follows that  $\delta\Phi = \delta\Phi'$ , and then by summation over all points  $\mathbf{q}$  and  $\mathbf{q}'$  that  $\Phi = \Phi'$ . But we already know that  $\Phi' = -GM/r$ , therefore  $\Phi = -GM/r$ , which is exactly the potential that would be generated by concentrating the entire mass of the inner shell at its center.◁

Alternative proofs of Newton's theorems, using the machinery of spherical harmonics, are given in §2.4.

From Newton's first and second theorems, it follows that the gravitational attraction of a spherical density distribution  $\rho(r')$  on a unit mass at radius  $r$  is entirely determined by the mass interior to  $r$ :

$$\mathbf{F}(r) = -\frac{GM(r)}{r^2}\hat{\mathbf{e}}_r, \quad (2.27a)$$

where

$$M(r) = 4\pi \int_0^r dr' r'^2 \rho(r'). \quad (2.27b)$$

The total gravitational potential may be considered to be the sum of the potentials of spherical shells of mass  $dM(r) = 4\pi\rho(r)r^2dr$ , so we may calculate the gravitational potential at  $\mathbf{r}$  generated by an arbitrary spherically symmetric density distribution  $\rho(r')$  by adding the contributions to the potential produced by shells (i) with  $r' < r$ , and (ii) with  $r' > r$ . In this way we obtain

$$\begin{aligned} \Phi(r) &= -\frac{G}{r} \int_0^r dM(r') - G \int_r^\infty \frac{dM(r')}{r'} \\ &= -4\pi G \left[ \frac{1}{r} \int_0^r dr' r'^2 \rho(r') + \int_r^\infty dr' r' \rho(r') \right]. \end{aligned} \quad (2.28)$$

It is worthwhile to check that the force  $\mathbf{F} = -\nabla\Phi$  obtained from (2.28) recovers the simple expression (2.27a).

An important property of a spherical matter distribution is its **circular speed**  $v_c(r)$ , defined to be the speed of a particle of negligible mass (a **test particle**) in a circular orbit at radius  $r$ . We may readily evaluate  $v_c$  by equating the gravitational attraction  $|\mathbf{F}|$  from equation (2.27a) to the centripetal acceleration  $v_c^2/r$ :

$$v_c^2 = r|\mathbf{F}| = r \frac{d\Phi}{dr} = \frac{GM(r)}{r}. \quad (2.29)$$

The associated angular frequency is called the **circular frequency**

$$\Omega \equiv \frac{v_c}{r} = \sqrt{\frac{GM(r)}{r^3}}. \quad (2.30)$$



The circular speed and frequency measure the mass interior to  $r$ . Another important quantity is the **escape speed**  $v_e$  defined by<sup>3</sup>

$$v_e(r) \equiv \sqrt{2|\Phi(r)|}. \quad (2.31)$$

A star at  $r$  can escape from the gravitational field represented by  $\Phi$  only if it has a speed at least as great as  $v_e(r)$ , for only then does its (positive) kinetic energy  $\frac{1}{2}v^2$  exceed the absolute value of its (negative) potential energy  $\Phi$ . The escape speed at  $r$  depends on the mass both inside and outside  $r$ .

**Potential energy of spherical systems** The simplest expression for the potential energy of a spherical body is obtained from equation (2.24). Substituting equation (2.27a) and integrating over all directions of  $\mathbf{r}$ , we obtain

$$W = -4\pi G \int_0^\infty dr r \rho(r) M(r). \quad (2.32)$$

It is straightforward to show (see Problem 2.2) that the potential-energy tensor of a spherical body is **diagonal**, that is,  $W_{jk} = 0$  for  $j \neq k$ , and has the form

$$W_{jk} = \frac{1}{3} W \delta_{jk}, \quad (2.33)$$

where  $\delta_{ij}$  is unity if  $i = j$  and zero otherwise. Such tensors are said to be **isotropic**.

### 2.2.2 Potentials of some simple systems

It is instructive to discuss the potentials generated by several simple density distributions:

**(a) Point mass** In this case

$$\Phi(r) = -\frac{GM}{r} \quad ; \quad v_c(r) = \sqrt{\frac{GM}{r}} \quad ; \quad v_e(r) = \sqrt{\frac{2GM}{r}}. \quad (2.34)$$

Potentials of this form, and orbits within them, are frequently referred to as **Keplerian** because Kepler first understood that  $v_c \propto r^{-1/2}$  in the solar system.

**(b) Homogeneous sphere** If the density is some constant  $\rho$ , we have  $M(r) = \frac{4}{3}\pi r^3 \rho$  and

$$v_c = \sqrt{\frac{4\pi G \rho}{3}} r. \quad (2.35)$$

---

<sup>3</sup> This result is correct only if the potential  $\Phi(r) \rightarrow 0$  as  $r \rightarrow \infty$ —we have assumed this so far but for systems with very extended mass distributions other zero points may be necessary (cf. eq. 2.62).

Thus in this case the circular speed rises linearly with radius, and the orbital period of a mass on a circular orbit is

$$T = \frac{2\pi r}{v_c} = \sqrt{\frac{3\pi}{G\rho}}, \quad (2.36)$$

independent of the radius of its orbit. The inverse of the angular frequency of a circular orbit is

$$\frac{r}{v_c} = \sqrt{\frac{3}{4\pi G\rho}} = 0.4886(G\rho)^{-1/2}. \quad (2.37)$$

If a small mass is released from rest at radius  $r$  in the gravitational field of a homogeneous sphere, its equation of motion is

$$\frac{d^2 r}{dt^2} = -\frac{GM(r)}{r^2} = -\frac{4\pi G\rho}{3}r, \quad (2.38)$$

which is the equation of motion of a harmonic oscillator of angular frequency  $2\pi/T$ . Therefore no matter what is the initial value of  $r$ , the test mass will reach  $r = 0$  in a quarter of a period, or in a time

$$\sqrt{\frac{3\pi}{16G\rho}} = 0.767(G\rho)^{-1/2}. \quad (2.39)$$

The times in equations (2.37) and (2.39) are rather similar, and this suggests that the time taken for a particle to complete a significant fraction of its orbit is  $\sim (G\rho)^{-1/2}$ , independent of the size and shape of the orbit. This result also holds for inhomogeneous systems, so long as  $\rho$  is replaced by the mean density  $\bar{\rho}$  interior to the particle's current radius. Thus we estimate the crossing time (sometimes also called the **dynamical time**) to be

$$t_{\text{cross}} \simeq t_{\text{dyn}} \simeq (G\bar{\rho})^{-1/2} \quad (2.40)$$

and shall use this as a measure of the characteristic time associated with the orbital motion of a star. Note that a complete orbital period is larger than  $t_{\text{cross}}$  by a factor  $\simeq 2\pi \simeq 6$ .

The potential energy of a homogeneous sphere of radius  $a$  and density  $\rho$  is conveniently obtained from equation (2.32). We have  $M(r) = \frac{4}{3}\pi\rho r^3$ , and therefore

$$W = -\frac{16\pi^2}{3}G\rho^2 \int_0^a dr r^4 = -\frac{16}{15}\pi^2 G\rho^2 a^5 = -\frac{3}{5}\frac{GM^2}{a}. \quad (2.41)$$

Sometimes it is useful to characterize the size of a system that lacks a sharp boundary by quoting the **gravitational radius**, which is defined as

$$r_g \equiv \frac{GM^2}{|W|}. \quad (2.42)$$

For a homogeneous sphere of radius  $a$ ,  $r_g = \frac{5}{3}a$ .

From equation (2.28) the gravitational potential of a homogeneous sphere of radius  $a$  is

$$\Phi(r) = \begin{cases} -2\pi G\rho(a^2 - \frac{1}{3}r^2) & (r < a), \\ -\frac{4\pi G\rho a^3}{3r} & (r > a). \end{cases} \quad (2.43)$$

**(c) Plummer model** We might expect that in many spherical systems the density is roughly constant near the center, and falls to zero at large radii. The potential of a system of this type would be proportional to  $r^2 + \text{constant}$  at small radii and to  $r^{-1}$  at large radii. A simple potential with these properties is the **Plummer model**

$$\Phi = -\frac{GM}{\sqrt{r^2 + b^2}}. \quad (2.44a)$$

The linear scale of the system that generates this potential is set by the **Plummer scale length**  $b$ , while  $M$  is the system's total mass.

From equation (B.53) for  $\nabla^2$  in spherical polar coordinates we have

$$\nabla^2\Phi = \frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{d\Phi}{dr} \right) = \frac{3Gmb^2}{(r^2 + b^2)^{5/2}}. \quad (2.45)$$

Thus from Poisson's equation (2.10) we have that the density corresponding to the potential (2.44a) is

$$\rho(r) = \frac{3M}{4\pi b^3} \left( 1 + \frac{r^2}{b^2} \right)^{-5/2}. \quad (2.44b)$$

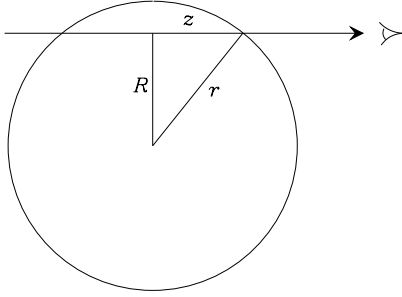
The potential energy of a Plummer model is

$$W = -\frac{3\pi GM^2}{32b}. \quad (2.46)$$

Plummer (1911) used the potential-density pair that is described by equations (2.44) to fit observations of globular clusters. We shall encounter it again in §4.3.3a as a member of the family of stellar systems known as polytropes.

**(d) Isochrone potential** The position of a star orbiting in a Plummer potential cannot be given in terms of elementary functions. However, in Chapter 3 we shall see that all orbits are analytic in the **isochrone potential**

$$\Phi(r) = -\frac{GM}{b + \sqrt{b^2 + r^2}}, \quad (2.47)$$



**Figure 2.3** Projection of a spherical body along the line of sight.

which owes its name to a property of its orbits that we shall derive in §3.1c. By equation (2.29) we have for the circular speed at radius  $r$

$$v_c^2(r) = \frac{GMr^2}{(b+a)^2a}, \quad (2.48a)$$

where

$$a \equiv \sqrt{b^2 + r^2}. \quad (2.48b)$$

When  $r$  is large  $v_c \simeq \sqrt{GM/r}$ , as required for a system of finite mass and extent. By Poisson's equation the density associated with the isochrone potential is

$$\rho(r) = \frac{1}{4\pi G} \frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{d\Phi}{dr} \right) = M \left[ \frac{3(b+a)a^2 - r^2(b+3a)}{4\pi(b+a)^3a^3} \right]. \quad (2.49)$$

Thus the central density is

$$\rho(0) = \frac{3M}{16\pi b^3}, \quad (2.50)$$

and at large radii the density tends to

$$\rho(r) \simeq \frac{bM}{2\pi r^4} \quad (r \gg b). \quad (2.51)$$

**(e) Modified Hubble model** The surface brightnesses of many elliptical galaxies may be approximated over a large range of radii by the **Hubble–Reynolds law**

$$I_H(R) = \frac{I_0}{(1 + R/R_H)^2}, \quad (2.52)$$

or the  $R^{1/4}$  law (eq. 1.17 with  $m = 4$ ). It is possible to solve for the spherical luminosity density  $j(r)$  that generates a given axisymmetric brightness

### Box 2.1: Definitions of surface brightness

Two definitions of surface brightness are used in the astronomical literature. Consider the radiative power that enters a telescope of aperture  $dA$  at normal incidence from directions in the element of solid angle  $d^2\Omega$  due to the distribution of luminosity density  $j(\mathbf{r})$ . We take the coordinate origin to lie at the center of the telescope aperture. Then in the distance range  $(r, r+dr)$  the volume that lies within the given solid-angle element is  $r^2 dr d^2\Omega$ , and the associated luminosity is  $dL = j(\mathbf{r}) r^2 dr d^2\Omega$ . A fraction  $dA/(4\pi r^2)$  of this luminosity enters the telescope, so the power received in the given solid-angle element from the given distance range is  $(4\pi)^{-1} j(\mathbf{r}) dr d^2\Omega dA$ . Summing over all distances and dividing by  $d^2\Omega dA$  we find that the flux per unit solid angle is

$$\hat{I} = \frac{1}{4\pi} \int dr j(\mathbf{r}).$$

The surface brightness  $\hat{I}$  might be reported in units of  $\text{W m}^{-2} \text{sr}^{-1}$ .

The second definition of surface brightness is the integral

$$I = \int dr j(\mathbf{r}).$$

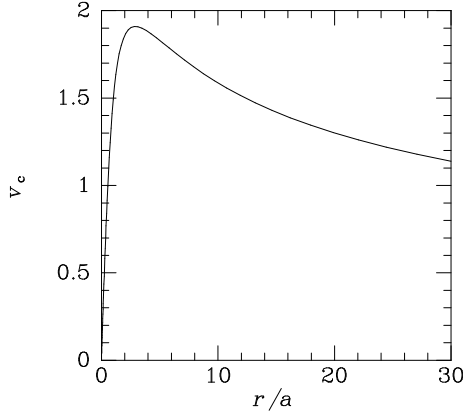
$I$  is the galaxy's luminosity per unit area when viewed from the given direction, and is conveniently given in units of solar luminosities per square parsec. Surface brightnesses are often reported by observers as so many magnitudes per square arcsecond, for example  $20 \text{ mag arcsec}^{-2}$ , meaning that in one square arcsecond of the image as much radiation is received as from a 20th magnitude star. In this book we follow the second convention and call  $I$  the "surface brightness."

distribution  $I(R)$  (see Problem 1.2). However, the resulting formulae for the luminosity distribution of a galaxy that obeys either of these empirical laws are cumbersome (Hubble 1930). Fortunately, the simple luminosity density

$$j_h(r) = j_0 \left( 1 + \frac{r^2}{a^2} \right)^{-3/2}, \quad (2.53)$$

where  $a$  is a constant, gives rise to a surface-brightness distribution that is similar to  $I_H$  (Rood et al. 1972). In fact, in the notation of Figure 2.3 we have that

$$I_h(R) = 2 \int_0^\infty dz j_h(r) = 2j_0 \int_0^\infty dz \left( 1 + \frac{R^2}{a^2} + \frac{z^2}{a^2} \right)^{-3/2}. \quad (2.54)$$



**Figure 2.4** Circular speed versus radius for a body whose projected density follows the modified Hubble model (2.55). The circular speed  $v_c$  is plotted in units of  $\sqrt{Gj_0\Upsilon a^2}$ .

Using the substitution  $y \equiv z/\sqrt{a^2 + R^2}$ , we obtain the **modified Hubble model**

$$I_h(R) = \frac{2j_0a}{1 + R^2/a^2} \int_0^\infty \frac{dy}{(1 + y^2)^{3/2}} = \frac{2j_0a}{1 + R^2/a^2}. \quad (2.55)$$

Thus  $I_h(a) = \frac{1}{2}I_h(0)$ , so  $a$  is the core radius as defined on page 30. At large  $R$ ,  $I_h(R) \propto R^{-2}$ , just as in the Hubble–Reynolds law (2.52). However, the Hubble–Reynolds law and the modified Hubble model behave quite differently near the center: while the luminosity density  $j_h$  is well behaved at the origin, in the Hubble–Reynolds law,  $dI_h/dR \neq 0$  at  $R = 0$ , which implies that the luminosity density  $j_h(r)$  diverges as  $r \rightarrow 0$  (Hubble 1930).

Using equation (2.28) we can calculate the potential that would be generated by the modified Hubble model if its mass were distributed in the same way as its light. If  $\rho(r) = \Upsilon j(r)$ , where  $\Upsilon$  is the constant mass-to-light ratio in the galaxy, one has

$$M_h(r) = 4\pi a^3 \Upsilon j_0 \left[ \ln \left( \frac{r}{a} + \sqrt{1 + \frac{r^2}{a^2}} \right) - \frac{r}{a} \left( 1 + \frac{r^2}{a^2} \right)^{-1/2} \right], \quad (2.56)$$

$$\Phi_h = -\frac{GM_h(r)}{r} - \frac{4\pi G \Upsilon j_0 a^2}{\sqrt{1 + (r/a)^2}}. \quad (2.57)$$

One feature of both the Hubble–Reynolds law and the modified Hubble model is that the mass diverges logarithmically at large  $r$ ; from equation (2.56)  $M_h \simeq 4\pi a^3 \Upsilon j_0 [\ln(2r/a) - 1]$  for  $r \gg a$ . In practice, a galaxy must have a finite mass, so  $j(r)$  must fall below  $j_h(r)$  at sufficiently large  $r$ . Nevertheless, the potential  $\Phi_h$  is finite, and in fact rather nearly equal to  $-GM_h(r)/r$  whenever  $r \gg a$ . This behavior indicates that from the gravitational point of view the density distribution of equation (2.53) behaves much like a point mass at large radii. The circular speed is shown in Figure 2.4. It peaks at  $r = 2.9a$  and then falls nearly as steeply as in the Keplerian case.

**(f) Power-law density model** Many galaxies have luminosity profiles that approximate a power law over a large range in radius. Consider the

structure of a system in which the mass density drops off as some power of the radius:

$$\rho(r) = \rho_0 \left( \frac{r_0}{r} \right)^\alpha. \quad (2.58)$$

The surface density of this system is

$$\Sigma(R) = \frac{\rho_0 r_0^\alpha}{R^{\alpha-1}} \frac{(-\frac{1}{2})! (\frac{\alpha-3}{2})!}{(\frac{\alpha-2}{2})!}. \quad (2.59)$$

We assume that  $\alpha < 3$ , since only in this case is the mass interior to  $r$  finite, namely

$$M(r) = \frac{4\pi\rho_0 r_0^\alpha}{3-\alpha} r^{3-\alpha}. \quad (2.60)$$

From equations (2.60) and (2.29) the circular speed is given by

$$v_c^2(r) = \frac{4\pi G \rho_0 r_0^\alpha}{3-\alpha} r^{2-\alpha}. \quad (2.61)$$

In Chapter 1 we saw that the circular-speed curves of many galaxies are remarkably flat. Equation (2.61) suggests that the mass density in these galaxies is proportional to  $r^{-2}$ , corresponding to  $\alpha = 2$ . In §4.3.3(b) we shall find that this is the density profile characteristic of a stellar-dynamical model called the “singular isothermal sphere.”

The potential difference between radius  $r$  and the reference radius  $r_0$  is

$$\begin{aligned} \Phi(r) - \Phi(r_0) &= G \int_{r_0}^r dr' \frac{M(r')}{r'^2} = \frac{4\pi G \rho_0 r_0^\alpha}{3-\alpha} \int_{r_0}^r dr' r'^{(1-\alpha)} \\ &= \begin{cases} \frac{v_c^2(r_0) - v_c^2(r)}{\alpha - 2} & \text{for } \alpha \neq 2 \\ v_c^2 \ln(r/r_0) & \text{for } \alpha = 2 \end{cases} \end{aligned} \quad (2.62)$$

Equation (2.60) shows that  $M(r)$  diverges at large  $r$  for all  $\alpha < 3$ . However, such models are still useful because by Newton’s first theorem, the mass exterior to any radius  $r$  does not affect the dynamics interior to  $r$ . For  $3 > \alpha > 2$ ,  $v_c^2(r)$  decreases with increasing  $r$  (eq. 2.61), so by (2.62) there is only a finite potential difference between radius  $r$  and infinity. In fact, the escape speed  $v_e(r)$  from radius  $r$  is given by

$$v_e^2(r) = 2[\Phi(\infty) - \Phi(r)] = 2 \frac{v_c^2(r)}{\alpha - 2} \quad (\alpha > 2). \quad (2.63)$$

For  $\alpha < 2$ ,  $\Phi(r)$  grows without limit as  $r \rightarrow \infty$ , so the interpretation of  $\Phi$  as the energy per unit mass required to remove a particle to infinity is no longer valid. The potential is nevertheless a useful concept because the

gravitational field is given by  $\mathbf{g} = -\nabla\Phi$ . Since the light distributions of elliptical galaxies suggest  $\alpha \simeq 3$  at large  $r$  (see eq. 2.53), while the flatness of the circular-speed curves in spiral galaxies suggest  $\alpha \simeq 2$ , it is clear that the escape speeds of galaxies are very uncertain.

**(g) Two-power density models** The luminosity density of many elliptical galaxies can be approximated as a power law in radius at both the largest and smallest observable radii, with a smooth transition between these power laws at intermediate radii (BM §4.3.1). Numerical simulations of the clustering of dark-matter particles suggest that the mass density within a dark halo has a similar structure (§9.3). For these reasons much attention has been devoted to models in which the density is given by

$$\rho(r) = \frac{\rho_0}{(r/a)^\alpha (1 + r/a)^{\beta-\alpha}}. \quad (2.64)$$

With  $\beta = 4$  these models have particularly simple analytic properties, and are known as **Dehnen models** (BM eq. 4.15; Dehnen 1993; Tremaine et al. 1994). BM Table 4.5 gives formulae for the projected surface density of Dehnen models for the cases  $\alpha = 0, 1, \frac{3}{2}$ , and 2. The model with  $\alpha = 1$  and  $\beta = 4$  is called a **Hernquist model** (Hernquist 1990), while that with  $\alpha = 2$  and  $\beta = 4$  is called a **Jaffe model** (Jaffe 1983). Dehnen models with  $\alpha$  in the range  $(0.6, 2)$  provide reasonable models of the centers of elliptical galaxies (BM §4.3.1).

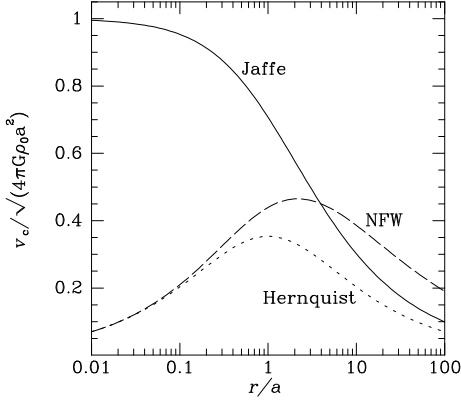
Dark halos can be modeled by equation (2.64) with  $\beta \simeq 3$  and  $\alpha$  in the range  $(1, 1.5)$ . The model with  $(\alpha, \beta) = (1, 3)$  is called the **NFW model** after Navarro, Frenk, & White (1995). The NFW formula contains two free parameters:  $\rho_0$  and  $a$ . Navarro, Frenk, & White (1996) showed that the values taken by these parameters for the halos that formed in their simulations were strongly correlated, so the halos were essentially members of a one-parameter family. The conventional choice of parameter is the distance  $r_{200}$  from the center of the halo at which the mean density is 200 times the cosmological critical density,  $\rho_c$  (eq. 1.52).<sup>4</sup> A more physical choice of parameter is the mass interior to  $r_{200}$ , which is  $M = 200\rho_c \frac{4}{3}\pi r_{200}^3$ . The **concentration** of the halo is  $c \equiv r_{200}/a$ . The central result of Navarro, Frenk, & White (1996) is that at a given value of  $M$  (and therefore  $r_{200}$ ), halos show a relatively small scatter in  $c$ . The mean value of  $c$  falls from  $\simeq 16$  to  $\simeq 6$  as the halo mass increases from  $\sim 3 \times 10^{11} \mathcal{M}_\odot$  to  $\sim 3 \times 10^{15} \mathcal{M}_\odot$ .

According to (2.64) the mass interior to radius  $r$  is

$$M(r) = 4\pi\rho_0 a^3 \int_0^{r/a} ds \frac{s^{2-\alpha}}{(1+s)^{\beta-\alpha}}. \quad (2.65)$$

<sup>4</sup> In §9.2.1 we shall show that  $r_{200}$  approximately divides the interior region in which material has crossed the halo at least once from the exterior one in which matter is still falling in to the halo for the first time.





**Figure 2.5** Circular speed versus radius for the Jaffe, Hernquist, and NFW models.

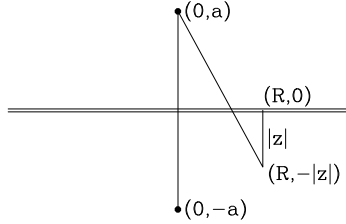
This integral is readily evaluated for integer values of  $\alpha, \beta$ . For the important cases of the Jaffe, Hernquist, and NFW models, we have

$$M(r) = 4\pi\rho_0 a^3 \times \begin{cases} \frac{r/a}{1+r/a} & \text{for a Jaffe model} \\ \frac{(r/a)^2}{2(1+r/a)^2} & \text{for a Hernquist model} \\ \ln(1+r/a) - \frac{r/a}{1+r/a} & \text{for a NFW model.} \end{cases} \quad (2.66)$$

In the Jaffe and Hernquist models the mass asymptotes to a finite value as  $r \rightarrow \infty$ , while in the NFW model the mass diverges logarithmically with  $r$ . The circular speed in each model, which follows immediately from (2.66) and (2.29), is plotted in Figure 2.5. From equation (2.29) the potentials of the three models are

$$\begin{aligned} \Phi &= -G \int_r^\infty dr \frac{M(r)}{r^2} \\ &= -4\pi G \rho_0 a^2 \times \begin{cases} \ln(1+a/r) & \text{for a Jaffe model} \\ \frac{1}{2(1+r/a)} & \text{for a Hernquist model} \\ \frac{\ln(1+r/a)}{r/a} & \text{for a NFW model.} \end{cases} \end{aligned} \quad (2.67)$$

From these formulae and equations (2.32) and (2.42) it is straightforward to show that the Jaffe and Hernquist models have gravitational radii  $r_g = 2a$  and  $6a$ , respectively, while for the NFW model  $r_g$  is undefined.



**Figure 2.6** At the point  $(R, -|z|)$  below Kuzmin's disk, the potential is identical with that of a point mass located distance  $a$  above the disk's center.

## 2.3 Potential-density pairs for flattened systems

Later in this chapter we will show how to obtain the gravitational potentials of systems of arbitrary shape. However, we shall find that the calculation of the gravitational potential and field generated by a given distribution of matter is often an arduous task that leads to cumbersome formulae involving special functions, or numerical calculations. Fortunately, for many purposes it suffices to represent a galaxy by a simple model that has the same gross structure as the galaxy. In this section we describe families of potentials that are generated by fairly simple and realistic axisymmetric density distributions. These potentials help us to understand how the gravitational potential of an initially spherical body is affected by flattening, and in later chapters we shall use several of these potentials to illustrate features of dynamics in axisymmetric galaxies.

### 2.3.1 Kuzmin models and generalizations

Consider the axisymmetric potential

$$\Phi_K(R, z) = -\frac{GM}{\sqrt{R^2 + (a + |z|)^2}} \quad (a \geq 0). \quad (2.68a)$$

As Figure 2.6 indicates, at points with  $z < 0$ ,  $\Phi_K$  is identical with the potential of a point mass  $M$  located at the point  $(R, z) = (0, a)$ ; and when  $z > 0$ ,  $\Phi_K$  coincides with the potential generated by a point mass at  $(0, -a)$ . Hence  $\nabla^2 \Phi_K$  must vanish everywhere except on the plane  $z = 0$ . By applying Gauss's theorem (2.12) to a flat volume that contains a small portion of the plane  $z = 0$ , we conclude that  $\Phi_K$  is generated by the surface density

$$\Sigma_K(R) = \frac{aM}{2\pi(R^2 + a^2)^{3/2}}. \quad (2.68b)$$

The potential-density pair of equations (2.68) was introduced by Kuzmin (1956), but it is often referred to as “Toomre's model 1” because it became widely known in the West only after Toomre (1963) unknowingly re-derived it.

Consider next the potential

$$\Phi_M(R, z) = -\frac{GM}{\sqrt{R^2 + (a + \sqrt{z^2 + b^2})^2}}. \quad (2.69a)$$

When  $a = 0$ ,  $\Phi_M$  reduces to Plummer's spherical potential (2.44a), and when  $b = 0$ ,  $\Phi_M$  reduces to Kuzmin's potential of a razor-thin disk (2.68a). Thus, depending on the choice of the two parameters  $a$  and  $b$ ,  $\Phi_M$  can represent the potential of anything from an infinitesimally thin disk to a spherical system. If we calculate  $\nabla^2 \Phi_M$ , we find that the mass distribution with which it is associated is (Miyamoto & Nagai 1975)

$$\rho_M(R, z) = \left( \frac{b^2 M}{4\pi} \right) \frac{aR^2 + (a + 3\sqrt{z^2 + b^2})(a + \sqrt{z^2 + b^2})^2}{[R^2 + (a + \sqrt{z^2 + b^2})^2]^{5/2} (z^2 + b^2)^{3/2}}. \quad (2.69b)$$

In Figure 2.7 we show contour plots of  $\rho_M(R, z)$  for various values of  $b/a$ . When  $b/a \simeq 0.2$ , these are qualitatively similar to the light distributions of disk galaxies, although there are quantitative differences. For example, from equation (2.69b) we have that  $\rho(R, 0) \propto R^{-3}$  when  $R$  is large, whereas the brightness profiles of disks fall off at least as fast as  $\exp(-R/R_d)$  (eq. 1.7).

Since Poisson's equation is linear in  $\Phi$  and  $\rho$ , the difference between any two potential-density pairs is itself a potential-density pair. Therefore, if we differentiate a potential-density pair with respect to one of its parameters, we obtain a new potential-density pair. For example, Toomre (1963) derived a family of potential-density pairs by differentiating  $\Phi_K(R, z)/a$   $n$  times with respect to  $a^2$ . Similarly, Satoh (1980) obtained a series of spherical potential-density pairs by differentiating  $b^{-2}$  times the Plummer potential and density (eq. 2.44)  $n$  times with respect to  $b^2$ . He flattened these potentials by replacing  $r^2 + b^2$  with  $R^2 + (a + \sqrt{z^2 + b^2})^2$ , and in the limit  $n \rightarrow \infty$  obtained

$$\Phi_S(R, z) = -\frac{GM}{S}, \quad (2.70a)$$

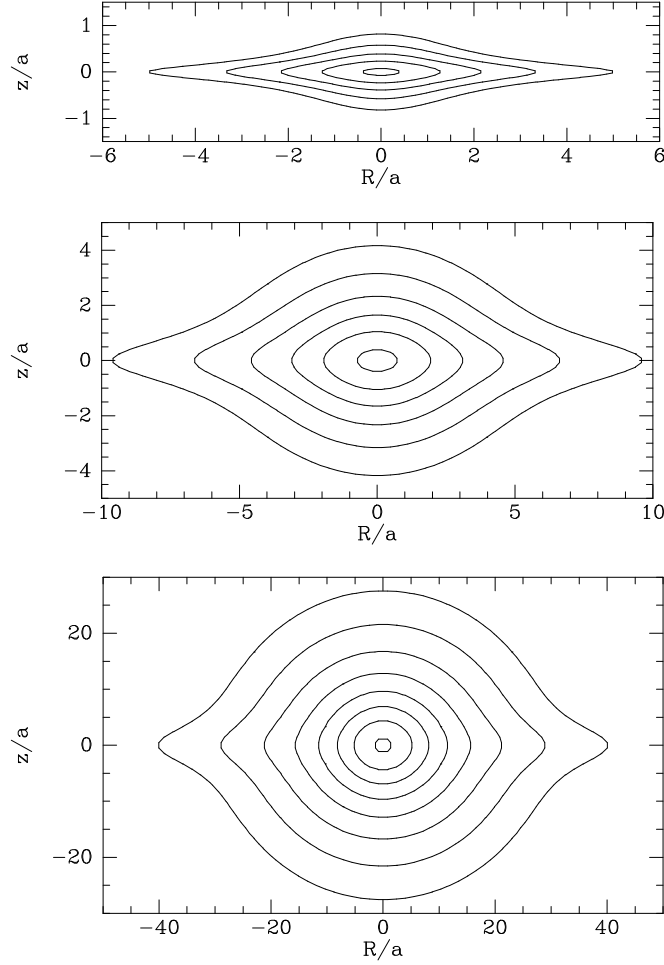
where

$$S^2 \equiv R^2 + z^2 + a \left( a + 2\sqrt{z^2 + b^2} \right). \quad (2.70c)$$

The corresponding density distribution follows from Poisson's equation:

$$\rho_S(R, z) = \frac{ab^2 M}{4\pi S^3(z^2 + b^2)} \left[ \frac{1}{\sqrt{z^2 + b^2}} + \frac{3}{a} \left( 1 - \frac{R^2 + z^2}{S^2} \right) \right]. \quad (2.70b)$$

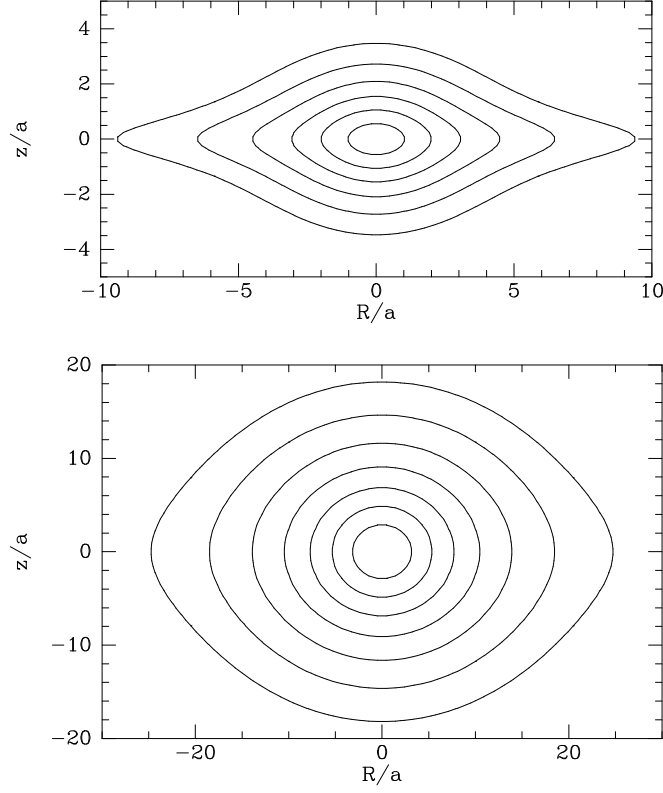
Figure 2.8 shows that at large  $b/a$  the isodensity surfaces  $\Phi_S$  are more nearly elliptical than those of  $\Phi_M$ .



**Figure 2.7** Contours of equal density in the  $(R, z)$  plane for the Miyamoto–Nagai density distribution (2.69b) when:  $b/a = 0.2$  (top);  $b/a = 1$  (middle);  $b/a = 5$  (bottom). There are two contours per decade, and the highest contour levels are  $0.3M/a^3$  (top),  $0.03M/a^3$  (middle), and  $0.001M/a^3$  (bottom).

### 2.3.2 Logarithmic potentials

Since the Kuzmin and other models in the previous subsection all have finite mass, the circular speed associated with these potentials falls off in Keplerian fashion,  $v_c \propto R^{-1/2}$ , at large  $R$ . However, in Chapter 1 it was shown that the circular-speed curves of spiral galaxies tend to be flat at large radii. If at large  $R$ ,  $v_c = v_0$ , a constant, then  $d\Phi/dR = v_0^2/R$ , and hence  $\Phi \propto$



**Figure 2.8** Contours of equal density in the  $(R, z)$  plane for Satoh's density distribution (2.70c) when:  $b/a = 1$  (top);  $b/a = 10$  (bottom). There are two contours per decade, and the highest contour levels are  $0.1M/a^3$  (top), and  $0.001M/a^3$  (bottom).

$v_0^2 \ln R + \text{constant}$  in this region. Therefore, consider the potential

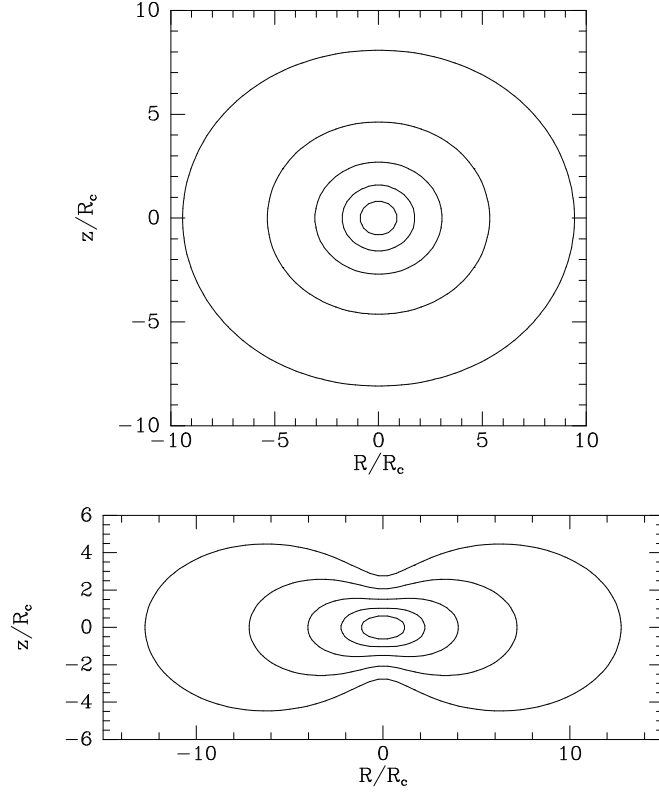
$$\Phi_L = \frac{1}{2}v_0^2 \ln \left( R_c^2 + R^2 + \frac{z^2}{q_\Phi^2} \right) + \text{constant}, \quad (2.71a)$$

where  $R_c$  and  $v_0$  are constants, and  $q_\Phi$  is the axis ratio of the equipotential surfaces. The circular speed at radius  $R$  in the equatorial plane of  $\Phi_L$  is

$$v_c = \frac{v_0 R}{\sqrt{R_c^2 + R^2}}. \quad (2.71b)$$

The density distribution to which  $\Phi_L$  corresponds is

$$\rho_L(R, z) = \frac{v_0^2}{4\pi G q_\Phi^2} \frac{(2q_\Phi^2 + 1)R_c^2 + R^2 + (2 - q_\Phi^{-2})z^2}{(R_c^2 + R^2 + z^2 q_\Phi^{-2})^2}. \quad (2.71c)$$



**Figure 2.9** Contours of equal density in the  $(R, z)$  plane for  $\rho_L$  (eq. 2.71c) when  $q_\Phi = 0.95$  (top),  $q_\Phi = 0.7$  (bottom). There are two contours per decade and the highest contour level is  $0.1v_0^2/(GR_c^2)$ . When  $q_\Phi = 0.7$  the models are unphysical because the density is negative near the  $z$  axis for  $|z| \gtrsim 7R_c$ .

At small  $R$  and  $z$ ,  $\rho_L$  tends to a constant value, and when  $R$  or  $|z|$  is large,  $\rho_L$  falls off as  $R^{-2}$  or  $z^{-2}$ .

The equipotential surfaces of  $\Phi_L$  are spheroids<sup>5</sup> of axial ratio  $q_\Phi$ , but Figure 2.9 shows that the equidensity surfaces are rather flatter and can deviate strongly from spheroids. In fact, if we define the axial ratio  $q_\rho$  of the isodensity surfaces by the ratio  $z_m/R_m$  of the distances down the  $z$  and  $R$  axes at which a given isodensity surface cuts the  $z$  axis and the  $x$  or  $y$  axis, we find

$$q_\rho^2 \simeq \frac{1 + 4q_\Phi^2}{2 + 3/q_\Phi^2} \quad (r \ll R_c) \quad (2.72a)$$

<sup>5</sup> A **spheroid** is the surface generated by rotating an ellipse about one of its principal axes. An **oblate**, or flattened, spheroid is generated if the axis of rotation is the minor axis, and a **prolate**, or elongated, spheroid is generated if this axis is the major axis.

or

$$q_\rho^2 \simeq q_\Phi^4 \left( 2 - \frac{1}{q_\Phi^2} \right) \quad (r \gg R_c). \quad (2.72b)$$

Outside the core, the flattening  $1 - q_\Phi$  of the potential is only about a third that of the density distribution:  $1 - q_\Phi \simeq \frac{1}{3}(1 - q_\rho)$ . The density  $\rho_L$  becomes negative on the  $z$  axis when  $q_\Phi < 1/\sqrt{2} = 0.707$ .

### 2.3.3 Poisson's equation in very flattened systems

In any axisymmetric system with density  $\rho(R, z)$ , Poisson's equation can be written (eq. B.52)

$$\frac{\partial^2 \Phi}{\partial z^2} = 4\pi G \rho(R, z) + \frac{1}{R} \frac{\partial}{\partial R} (R F_R), \quad (2.73)$$

where  $F_R = -\partial\Phi/\partial R$  is the radial force. Now consider, for example, the Miyamoto–Nagai potential-density pair given by equations (2.69). As the parameter  $b \rightarrow 0$ , the density distribution becomes more and more flattened, and at fixed  $R$  the density in the plane  $z = 0$  becomes larger and larger as  $b^{-1}$ . However, the radial force  $F_R$  remains well behaved as  $b \rightarrow 0$ ; indeed, in the limit  $b = 0$ ,  $F_R = -\partial\Phi_K/\partial R$ , where  $\Phi_K(R, z)$  is simply the Kuzmin potential (2.68a). Thus, near  $z = 0$  the first term on the right side of equation (2.73) becomes very large compared to the second, and Poisson's equation simplifies to the form

$$\frac{\partial^2 \Phi(R, z)}{\partial z^2} = 4\pi G \rho(R, z). \quad (2.74)$$

This result applies to almost any thin disk system. It implies that the vertical variation of the potential at a given radius  $R$  depends only on the density distribution at that radius. Effectively, this means that the solution of Poisson's equation in a thin disk can be decomposed into two steps: (i) Approximate the thin disk as a surface density layer of zero thickness and determine the potential in the plane of the disk  $\Phi(R, 0)$  using the models of this section or the more general techniques of §2.6. (ii) At each radius  $R$  solve equation (2.74) to find the vertical variation of  $\Phi(R, z)$ .

Thus we have

$$\Phi(R, z) = \Phi(R, 0) + \Phi_z(R, z) \quad (2.75a)$$

where

$$\Phi_z(R, z) \equiv 4\pi G \int_0^z dz' \int_0^{z'} dz'' \rho(R, z'') + a(R)z. \quad (2.75b)$$

The constant of integration,  $a$ , is zero if the disk is symmetric around the equatorial plane.

## 2.4 Multipole expansion

In the last section we encountered a number of axisymmetric density distributions that give rise to potentials of known form. By adding a few of these distributions together one can obtain quite a wide range of model galaxies that have readily available potentials. However, for many purposes one requires a systematic procedure for calculating the potential of an arbitrary density distribution to whatever accuracy one pleases. The next few sections are devoted to this task.

The first such technique, based on spherical harmonics, works best for systems that are neither very flattened nor very elongated. Hence it is a good method for calculating the potentials of bulges and dark-matter halos (§§1.1.2 and 1.1.3).

Our first step is to obtain the potential of a thin spherical shell of variable surface density. Since the shell has negligible thickness, the task of solving Poisson's equation  $\nabla^2\Phi = 4\pi G\rho$  reduces to that of solving Laplace's equation  $\nabla^2\Phi = 0$  inside and outside the shell, subject to suitable boundary conditions at infinity, at the origin, and on the shell. Now in spherical coordinates Laplace's equation is (eq. B.53)

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \Phi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \Phi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \Phi}{\partial \phi^2} = 0. \quad (2.76)$$

This may be solved by the method of **separation of variables**. We seek special solutions that are the product of functions of one variable only:

$$\Phi(r, \theta, \phi) = R(r)P(\theta)Q(\phi). \quad (2.77a)$$

Substituting equation (2.77a) into (2.76) and rearranging, we obtain

$$\frac{\sin^2 \theta}{R} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) + \frac{\sin \theta}{P} \frac{d}{d\theta} \left( \sin \theta \frac{dP}{d\theta} \right) = -\frac{1}{Q} \frac{d^2 Q}{d\phi^2}. \quad (2.77b)$$

The left side of this equation does not depend on  $\phi$ , and the right side does not depend on  $r$  or  $\theta$ . It follows that both sides are equal to some constant, say  $m^2$ . Hence

$$-\frac{1}{Q} \frac{d^2 Q}{d\phi^2} = m^2, \quad (2.78a)$$

$$\frac{\sin^2 \theta}{R} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) + \frac{\sin \theta}{P} \frac{d}{d\theta} \left( \sin \theta \frac{dP}{d\theta} \right) = m^2. \quad (2.78b)$$

Equation (2.78a) may be immediately integrated to

$$Q(\phi) = Q_m^+ e^{im\phi} + Q_m^- e^{-im\phi}. \quad (2.79a)$$



We require  $\Phi$  to be a periodic function of  $\phi$  with period  $2\pi$ , so  $m$  can take only integer values. Since equations (2.78) depend only on  $m^2$ , we could restrict our attention to non-negative values of  $m$  without loss of generality. However, a simpler convention is to allow  $m$  to take both positive and negative values, so the second exponential in equation (2.79a) becomes redundant, and we may write simply

$$Q = Q_m e^{im\phi} \quad (m = \dots, -1, 0, 1, \dots). \quad (2.79b)$$

Equation (2.78b) can be written

$$\frac{1}{R} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) = \frac{m^2}{\sin^2 \theta} - \frac{1}{P \sin \theta} \frac{d}{d\theta} \left( \sin \theta \frac{dP}{d\theta} \right). \quad (2.80)$$

Since the left side of this equation does not depend on  $\theta$  and the right side does not depend on  $r$ , both sides must equal some constant, which we write as  $l(l+1)$ . Thus equation (2.80) implies

$$\frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) - l(l+1)R = 0, \quad (2.81a)$$

and in terms of  $x \equiv \cos \theta$ ,

$$\frac{d}{dx} \left[ (1-x^2) \frac{dP}{dx} \right] - \frac{m^2}{1-x^2} P + l(l+1)P = 0. \quad (2.81b)$$

Two linearly independent solutions of equation (2.81a) are

$$R(r) = Ar^l \quad \text{and} \quad R(r) = Br^{-(l+1)}. \quad (2.82)$$

The solutions of equation (2.81b) are associated Legendre functions  $P_l^m(x)$  (see Appendix C.5). Physically acceptable solutions exist only when  $l$  is an integer. Without loss of generality we can take  $l$  to be non-negative, and then physically acceptable solutions exist only for  $|m| \leq l$ . When  $m = 0$  the solutions are simply polynomials in  $x$ , called Legendre polynomials  $P_l(x)$ .

Rather than write out the product  $P_l^m(\cos \theta) e^{im\phi}$  again and again, it is helpful to define the spherical harmonic  $Y_l^m(\theta, \phi)$ , which is equal to  $P_l^m(\cos \theta) e^{im\phi}$  times a constant chosen so the  $Y_l^m$  satisfy the orthogonality relation (see eq. C.44)

$$\begin{aligned} \int d^2\Omega Y_l^{m*}(\mathbf{\Omega}) Y_{l'}^{m'}(\mathbf{\Omega}) &\equiv \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi Y_l^{m*}(\theta, \phi) Y_{l'}^{m'}(\theta, \phi) \\ &= \delta_{ll'} \delta_{mm'}, \end{aligned} \quad (2.83)$$

where we have used  $\mathbf{\Omega}$  as a shorthand for  $(\theta, \phi)$  and  $d^2\Omega$  for  $\sin \theta d\theta d\phi$ . The spherical harmonics with  $l \leq 2$  are listed in equation (C.50).

Putting all these results together, we have from equations (2.77a), (2.79b), and (2.82) that

$$\Phi_{lm}(r, \mathbf{\Omega}) = \left( A_{lm} r^l + B_{lm} r^{-(l+1)} \right) Y_l^m(\mathbf{\Omega}) \quad (2.84)$$

is a solution of  $\nabla^2 \Phi = 0$  for all non-negative integers  $l$  and integer  $m$  in the range  $-l \leq m \leq l$ .

Now let us apply these results to the problem of determining the potential of a thin shell of radius  $a$  and surface density  $\sigma(\mathbf{\Omega})$ . We write the potential internal and external to the shell as

$$\Phi_{\text{int}}(r, \mathbf{\Omega}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \left( A_{lm} r^l + B_{lm} r^{-(l+1)} \right) Y_l^m(\mathbf{\Omega}) \quad (r \leq a), \quad (2.85a)$$

and

$$\Phi_{\text{ext}}(r, \mathbf{\Omega}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \left( C_{lm} r^l + D_{lm} r^{-(l+1)} \right) Y_l^m(\mathbf{\Omega}) \quad (r \geq a). \quad (2.85b)$$

The potential at the center must be non-singular, so  $B_{lm} = 0$  for all  $l, m$ . Similarly, the potential at infinity must be zero, so  $C_{lm} = 0$  for all  $l, m$ . Furthermore,  $\Phi_{\text{ext}}(a, \mathbf{\Omega})$  must equal  $\Phi_{\text{int}}(a, \mathbf{\Omega})$  because no work can be done in passing through an infinitesimally thin shell. Hence from equations (2.85) we have

$$\sum_{l=0}^{\infty} \sum_{m=-l}^l A_{lm} a^l Y_l^m(\mathbf{\Omega}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l D_{lm} a^{-(l+1)} Y_l^m(\mathbf{\Omega}). \quad (2.86)$$

The coefficients  $A_{lm} a^l$  etc. of each spherical harmonic  $Y_l^m$  on the two sides of equation (2.86) must be equal, as can be shown by multiplying both sides of the equation by  $Y_{l'}^{m'*}(\mathbf{\Omega})$ , integrating over  $\mathbf{\Omega}$ , and using the orthogonality relation (2.83). Therefore, from equation (2.86) we have

$$D_{lm} = A_{lm} a^{2l+1}. \quad (2.87)$$

Now let us expand the surface density of the thin shell as

$$\sigma(\mathbf{\Omega}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \sigma_{lm} Y_l^m(\mathbf{\Omega}), \quad (2.88)$$

where the  $\sigma_{lm}$  are numbers yet to be determined. To obtain the coefficient  $\sigma_{l'm'}$ , we multiply both sides of equation (2.88) by  $Y_{l'}^{m'*}(\mathbf{\Omega})$  and integrate over  $\mathbf{\Omega}$ . With equation (2.83) we find

$$\int d^2\Omega Y_{l'}^{m'*}(\mathbf{\Omega}) \sigma(\mathbf{\Omega}) = \sigma_{l'm'}. \quad (2.89)$$

Since  $Y_0^0 = 1/\sqrt{4\pi}$ ,  $\sigma_{00} = M/(2a^2\sqrt{\pi})$ , where  $M$  is the mass of the shell.

Gauss's theorem (2.12) applied to a small piece of the shell tells us that

$$\left(\frac{\partial\Phi_{\text{ext}}}{\partial r}\right)_{r=a} - \left(\frac{\partial\Phi_{\text{int}}}{\partial r}\right)_{r=a} = 4\pi G\sigma(\mathbf{\Omega}), \quad (2.90)$$

so inserting equations (2.85) and (2.88) into equation (2.90), we obtain

$$\begin{aligned} -\sum_{l=0}^{\infty} \sum_{m=-l}^l \left( (l+1)D_{lm}a^{-(l+2)} + lA_{lm}a^{l-1} \right) Y_l^m(\mathbf{\Omega}) = \\ 4\pi G \sum_{l=0}^{\infty} \sum_{m=-l}^l \sigma_{lm} Y_l^m(\mathbf{\Omega}). \end{aligned} \quad (2.91)$$

Once again the coefficients of  $Y_l^m$  on each side of the equation must be identical, so with (2.87) we have

$$A_{lm} = -4\pi G a^{-(l-1)} \frac{\sigma_{lm}}{2l+1} \quad ; \quad D_{lm} = -4\pi G a^{l+2} \frac{\sigma_{lm}}{2l+1}. \quad (2.92)$$

Collecting these results together, we have from equations (2.85) that

$$\begin{aligned} \Phi_{\text{int}}(r, \mathbf{\Omega}) &= -4\pi G a \sum_{l=0}^{\infty} \left(\frac{r}{a}\right)^l \sum_{m=-l}^l \frac{\sigma_{lm}}{2l+1} Y_l^m(\mathbf{\Omega}), \\ \Phi_{\text{ext}}(r, \mathbf{\Omega}) &= -4\pi G a \sum_{l=0}^{\infty} \left(\frac{a}{r}\right)^{l+1} \sum_{m=-l}^l \frac{\sigma_{lm}}{2l+1} Y_l^m(\mathbf{\Omega}), \end{aligned} \quad (2.93)$$

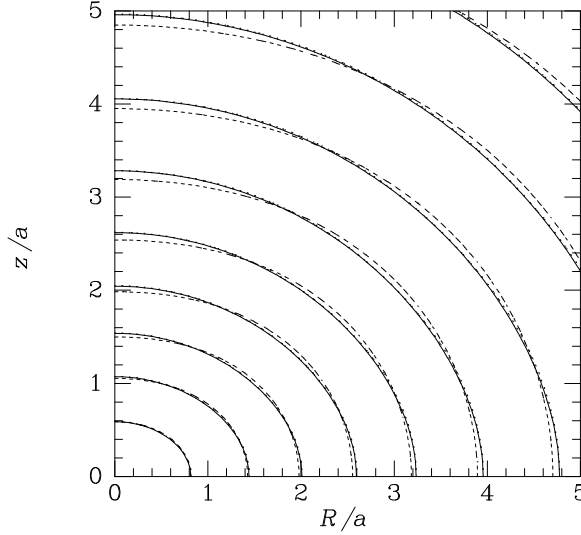
where the  $\sigma_{lm}$  are given by equation (2.89).

Finally we evaluate the potential of a solid body by breaking it down into a series of spherical shells. We let  $\delta\sigma_{lm}(a)$  be the  $\sigma$ -coefficient of the shell lying between  $a$  and  $a + \delta a$ , and  $\delta\Phi(r, \mathbf{\Omega}; a)$  be the corresponding potential at  $r$ . Then we have by equation (2.89)

$$\delta\sigma_{lm}(a) = \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\phi Y_l^{m*}(\mathbf{\Omega}) \rho(a, \mathbf{\Omega}) \delta a \equiv \rho_{lm}(a) \delta a. \quad (2.94)$$

Substituting these values of  $\sigma_{lm}$  into equations (2.93) and integrating over all  $a$ , we obtain the potential at  $r$  generated by the entire collection of shells:

$$\begin{aligned} \Phi(r, \mathbf{\Omega}) &= \sum_{a=0}^r \delta\Phi_{\text{ext}} + \sum_{a=r}^{\infty} \delta\Phi_{\text{int}} \\ &= -4\pi G \sum_{l,m} \frac{Y_l^m(\mathbf{\Omega})}{2l+1} \left( \frac{1}{r^{l+1}} \int_0^r da a^{l+2} \rho_{lm}(a) + r^l \int_r^\infty \frac{da}{a^{l-1}} \rho_{lm}(a) \right). \end{aligned} \quad (2.95)$$

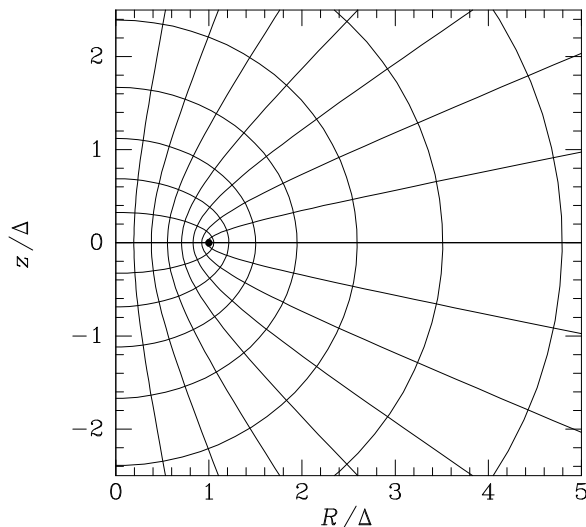


**Figure 2.10** Equipotentials of Satoh's density distribution (2.70c) with  $b/a = 1$ . Full curves show the exact equipotentials computed from equation (2.70a), and dashed curves show the estimate provided by equation (2.95) with the sum over  $l$  extending to  $l = 2$ . Contours based on the sum to  $l = 8$  are also plotted (dotted contours) but almost overlie the full curves.

This equation gives the potential generated by the body as an expansion in **multipoles**: the terms associated with  $l = m = 0$  are the **monopole** terms, those associated with  $l = 1$  are **dipole** terms, those with  $l = 2$  are **quadrupole** terms, and those with larger  $l$  are  $2^l$ -poles. Similar expansions occur in electrostatics (e.g., Jackson 1999). The monopole terms are the same as in equation (2.28) for the potential of a spherical system.<sup>6</sup> Since there is no gravitational analog of negative charge, pure dipole or quadrupole gravitational potentials cannot arise, in contrast to the electrostatic case. In fact, if one places the origin of coordinates at the center of mass of the system, the dipole term vanishes identically outside any matter distribution. While the monopole terms generate a circular-speed curve  $v_c(r) = \sqrt{GM(r)/r}$  that never declines with increasing  $r$  more steeply than in the Keplerian case ( $v_c \propto r^{-1/2}$ ), over a limited range in  $r$  the higher-order multipoles may cause the circular speed to fall more steeply with increasing radius.

As an illustration of the effectiveness of the multipole expansion, we show in Figure 2.10 the contours of Satoh's potential  $\Phi_S(R, z)$  (eq. 2.70a), together with the approximations to this potential that one obtains from equation (2.95) if one includes only terms with  $l \leq 2$  or 8. The flexibility

<sup>6</sup> Thus the spherical-harmonic expansion provides an alternative proof of Newton's first and second theorems.



**Figure 2.11** Curves of constant  $u$  and  $v$  in the  $(R, z)$  plane. Semi-ellipses are curves of constant  $u$ , and hyperbolae are curves of constant  $v$ . The common focus of all curves is marked by a dot. In order to ensure that each point has a unique  $v$ -coordinate, we exclude the disk ( $z = 0$ ,  $R \leq \Delta$ ) from the space to be considered.

of the multipole expansion makes it a powerful tool for numerical work, and it plays a central role in some of the Poisson solvers for N-body codes that will be described in §2.9. However, multipole expansions are poorly suited for modeling the potentials of disks.

## 2.5 The potentials of spheroidal and ellipsoidal systems

Many galaxies have nearly spheroidal or ellipsoidal equidensity surfaces (BM §4.3.3). Moreover, Newton's theorems for spherical bodies can be generalized to include spheroidal and ellipsoidal bodies, so models with isodensity surfaces of this shape are relatively easy to construct. Finally, as the axis ratio shrinks to zero a spheroid becomes a disk, and thus we can obtain the potentials of razor-thin disks as a limiting case of spheroids.

In this section we develop efficient techniques for calculating the potentials of such objects. In §§2.5.1 and 2.5.2 we derive formulae for oblate (i.e., flattened) spheroidal systems, and in §2.5.3 we briefly discuss ellipsoidal systems. Results for prolate spheroidal systems can be obtained either by adapting our derivations for oblate systems, or by specializing the results for ellipsoidal systems. The principal formulae for all three geometries are given in Tables 2.1 and 2.2.

### 2.5.1 Potentials of spheroidal shells

Consider a system in which the isodensity surfaces are similar concentric spheroids. We align our coordinate system such that the  $z$  axis is the minor axis of the spheroids, and designate the principal-axis lengths of the spheroids by  $a$  and  $c$ , with  $a \geq c$ . Spheroidal bodies call for spheroidal coordinates, so we consider the form of Laplace's equation,  $\nabla^2 \Phi = 0$ , in **oblate spheroidal coordinates**. These coordinates employ the usual azimuthal angle  $\phi$  of cylindrical coordinates, but replace the coordinates  $(R, z)$  with new coordinates  $(u, v)$  that are defined by

$$R = \Delta \cosh u \sin v \quad ; \quad z = \Delta \sinh u \cos v \quad (u \geq 0, 0 \leq v \leq \pi), \quad (2.96)$$

where  $\Delta$  is a constant. Figure 2.11 shows the curves of constant  $u$  and  $v$  in the  $(R, z)$  plane. The curves  $u = \text{constant}$  are confocal half-ellipses with foci at  $(R, z) = (\Delta, 0)$ , namely

$$\frac{R^2}{\cosh^2 u} + \frac{z^2}{\sinh^2 u} = \Delta^2. \quad (2.97)$$

Similarly, the curves  $v = \text{constant}$  coincide with the hyperbolae

$$\frac{R^2}{\sin^2 v} - \frac{z^2}{\cos^2 v} = \Delta^2 \quad (2.98)$$

formed by the normals to these ellipses.

From equations (2.96) we see that if we increase one of  $u$ ,  $v$ , and  $\phi$  by a small amount while holding the other two coordinates constant, the point  $(u, v, \phi)$  moves parallel to the three orthogonal unit vectors  $\hat{\mathbf{e}}_u$ ,  $\hat{\mathbf{e}}_v$ ,  $\hat{\mathbf{e}}_\phi$  by the distances  $h_u \delta u$ ,  $h_v \delta v$  and  $h_\phi \delta \phi$ , where the scale factors are

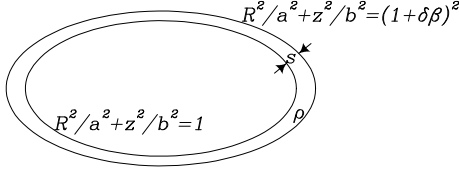
$$h_u = h_v = \Delta \sqrt{\sinh^2 u + \cos^2 v} \quad ; \quad h_\phi = \Delta \cosh u \sin v. \quad (2.99)$$

Hence the gradient of a potential  $\Phi$  may be expressed in these coordinates as (eq. B.39)

$$\nabla \Phi = \frac{1}{\Delta \sqrt{\sinh^2 u + \cos^2 v}} \left[ \frac{\partial \Phi}{\partial u} \hat{\mathbf{e}}_u + \frac{\partial \Phi}{\partial v} \hat{\mathbf{e}}_v \right] + \frac{1}{\Delta \cosh u \sin v} \frac{\partial \Phi}{\partial \phi} \hat{\mathbf{e}}_\phi. \quad (2.100)$$

We further have (eq. B.54)

$$\begin{aligned} \nabla^2 \Phi = & \frac{1}{\Delta^2 (\sinh^2 u + \cos^2 v)} \left[ \frac{1}{\cosh u} \frac{\partial}{\partial u} \left( \cosh u \frac{\partial \Phi}{\partial u} \right) \right. \\ & \left. + \frac{1}{\sin v} \frac{\partial}{\partial v} \left( \sin v \frac{\partial \Phi}{\partial v} \right) \right] + \frac{1}{\Delta^2 \cosh^2 u \sin^2 v} \frac{\partial^2 \Phi}{\partial \phi^2}. \end{aligned} \quad (2.101)$$



**Figure 2.12** A homoeoid of density  $\rho$  is bounded by the surfaces  $R^2/a^2 + z^2/b^2 = 1$  and  $R^2/a^2 + z^2/b^2 = (1 + \delta\beta)^2$ . The perpendicular distance  $s$  between the bounding surfaces varies with position around the homoeoid.

We concentrate on potentials that are functions  $\Phi(u)$  of only the “radial” coordinate  $u$ . For potentials of this class,  $\nabla^2\Phi = 0$  reduces to

$$\frac{d}{du} \left( \cosh u \frac{d\Phi}{du} \right) = 0. \quad (2.102)$$

Hence either

$$\Phi = \Phi_0 \quad (\text{a constant}), \quad (2.103a)$$

or  $d\Phi/du = A \operatorname{sech} u$ , where  $A$  is a constant. Integrating this last equation we find

$$\Phi = -A \sin^{-1}(\operatorname{sech} u) + B, \quad (2.103b)$$

where  $B$  is a constant.

For  $u$  large,  $\operatorname{sech} u \rightarrow \Delta/r \rightarrow 0$ , where  $r$  is the usual spherical radius. So a potential of the form (2.103b) varies as

$$\Phi \simeq -A \operatorname{sech} u + B \rightarrow -\frac{A\Delta}{r} + B, \quad (2.104)$$

Hence, if we set  $B = 0$  and  $A = G\delta M/\Delta$ , the potential given by equation (2.103b) tends to zero at infinity like the gravitational potential of a shell of mass  $\delta M$ . Thus we are led to consider the potential defined by

$$\Phi = -\frac{G\delta M}{\Delta} \times \begin{cases} \sin^{-1}(\operatorname{sech} u_0) & (u < u_0), \\ \sin^{-1}(\operatorname{sech} u) & (u \geq u_0). \end{cases} \quad (2.105)$$

This potential is everywhere continuous and solves  $\nabla^2\Phi = 0$  everywhere except on the spheroid  $u = u_0$  (see eqs. 2.103). Hence it is the gravitational potential of a shell of material on the surface  $u = u_0$ . This shell has principal semi-axes of lengths  $a \equiv \Delta \cosh u_0$  and  $c \equiv \Delta \sinh u_0$ . Hence the shell's **eccentricity**

$$e \equiv \sqrt{1 - \frac{c^2}{a^2}} = \operatorname{sech} u_0, \quad (2.106)$$

and we may rewrite equations (2.105) as

$$\Phi = -\frac{G\delta M}{ae} \times \begin{cases} \sin^{-1}(e) & (u < u_0), \\ \sin^{-1}(\operatorname{sech} u) & (u \geq u_0). \end{cases} \quad (2.107)$$

We can find the surface density of the shell  $u = u_0$  by applying Gauss's theorem (2.12) to the potential (2.107). Since  $\nabla\Phi = 0$  inside the shell, by equation (2.100) the surface density of the shell is

$$\begin{aligned}\Sigma(v) &= \frac{\hat{\mathbf{e}}_u \cdot \nabla\Phi}{4\pi G} = \frac{1}{4\pi G \Delta \sqrt{\sinh^2 u_0 + \cos^2 v}} \left( \frac{d\Phi}{du} \right)_{u=u_0+} \\ &= \frac{\delta M}{4\pi a^2 \sqrt{1 - e^2 \sin^2 v}},\end{aligned}\quad (2.108)$$

where evaluation at  $u_0+$  denotes the limiting value as  $u \rightarrow u_0$  from above.

Equation (2.108) has a simple physical interpretation:  $\Sigma(v)$  is the surface density of the thin shell of uniform density  $\rho$  that is bounded by the two surfaces  $\beta$  and  $\beta + \delta\beta$  of the set of similar spheroids,

$$\frac{R^2}{a^2} + \frac{z^2}{c^2} = \beta^2, \quad (2.109)$$

where  $\delta\beta$  and  $\delta M$  are related by (2.113) below. *Proof:* The small perpendicular  $\mathbf{s}$  in Figure 2.12 runs between the shell's inner and outer skins. Thus  $\mathbf{s} = s\nabla\beta/|\nabla\beta|$ , and at any point on the surface we have  $\delta\beta = (\mathbf{s} \cdot \nabla\beta) = s|\nabla\beta|$ . Hence  $s = \delta\beta/|\nabla\beta|$  and the surface density of the shell is

$$\tilde{\Sigma} = \rho s = \frac{\rho\delta\beta}{|\nabla\beta|} = \left( \frac{R^2}{a^4} + \frac{z^2}{c^4} \right)^{-1/2} \rho\delta\beta. \quad (2.110)$$

Finally, writing  $R = \beta a \sin v$ ,  $z = \beta c \cos v$ , and  $e = \sqrt{1 - c^2/a^2}$ , we find

$$\tilde{\Sigma} = \frac{a\sqrt{1 - e^2} \rho \delta\beta}{\sqrt{1 - e^2 \sin^2 v}}. \quad (2.111)$$

The volume inside an oblate spheroidal shell with semi-axis lengths  $a\beta$  and  $c\beta$  is

$$V = \frac{4}{3}\pi a^2 c \beta^3 = \frac{4}{3}\pi a^3 \beta^3 \sqrt{1 - e^2}, \quad (2.112)$$

so the mass that is bounded by the surfaces  $\beta$  and  $\beta + \delta\beta$  is

$$\delta M = 4\pi\rho a^3 \sqrt{1 - e^2} \beta^2 \delta\beta. \quad (2.113)$$

We can now set  $\beta = 1$  and substitute (2.113) into equation (2.111) to find that  $\tilde{\Sigma} = \Sigma$ , where  $\Sigma$  is given by equation (2.108). ◁

We have shown that the potential (2.107) is generated by a thin shell of uniform density that is bounded by similar spheroids of eccentricity  $e$ . We call such a shell a **thin homoeoid** and have:



**Homoeoid theorem** *The exterior isopotential surfaces of a thin homoeoid are the spheroids that are confocal with the shell itself. Inside the shell the potential is constant.*

The homoeoid theorem applies only to a *thin* homoeoid. But it immediately yields a remarkable property of any homoeoid—that is, of any shell of constant density, no matter how thick, whose inner and outer surfaces are similar (*not* confocal) spheroids:

**Newton’s third theorem** *A mass that is inside a homoeoid experiences no net gravitational force from the homoeoid.*

*Proof:* Break the given homoeoid into a series of thin homoeoids. The interior of the thick homoeoid lies in the interior of each of its component thin homoeoids, and the interior potential of each thin homoeoid is constant. Hence the aggregate interior potential is constant and generates no gravitational force.◁

Newton’s first theorem, for spherical systems, thus emerges as a special case of Newton’s third theorem for spheroidal systems. Newton’s second theorem, for spherical systems, has no analog for spheroidal systems because the potential outside a spheroidal body *does* depend on the distribution of matter within it.

These theorems help us to understand qualitatively the potential of an inhomogeneous spheroidal body. Each shell of the body makes a contribution to the potential that is constant interior to the shell and on the shell, and gradually becomes rounder as one moves outward from the shell. This tendency of isopotentials to become spherical at large radii manifested itself already in §2.4 in the rapid radial decay of the higher multipole components of the gravitational potential. The shape of the isopotential surface at a distance  $r$  from the center of an inhomogeneous spheroidal body represents a compromise between the rather round contributions of the central shells, and the more aspherical contributions of the shells just interior to  $r$ . Thus, if the body is very centrally concentrated, the isopotentials near its edge will be nearly round, while a more homogeneous spheroidal body will have more flattened isopotentials.

### 2.5.2 Potentials of spheroidal systems

We now use equation (2.107) to calculate the gravitational potential of a body whose isodensity surfaces are the similar spheroids

$$\text{constant} = m^2 \equiv R^2 + \frac{z^2}{1 - e^2}, \quad (2.114)$$

i.e., a body in which  $\rho = \rho(m^2)$ . Comparing equations (2.109) and (2.114), we see that  $m = \beta a$ , so the mass of the shell between  $m$  and  $m + \delta m$  is given by equation (2.113) as

$$\delta M = 4\pi\rho(m^2)\sqrt{1 - e^2} m^2 \delta m. \quad (2.115)$$

For a spherical system ( $e = 0$ ) this reduces to the familiar formula  $\delta M = 4\pi\rho r^2\delta r$ .

There is a unique family of confocal spheroids such that one member of the family coincides with the homoeoid labeled by  $m$ . Let  $u_m(R_0, z_0)$  be the label of the member of this family that passes through the point  $(R_0, z_0)$  at which the potential is required (see eq. 2.122 below for an explicit formula). Then if  $(R_0, z_0)$  lies inside the homoeoid  $m$ , we have on setting  $a = m$  in equation (2.107) and substituting for  $\delta M$  from equation (2.115) that the contribution of  $m$  to the potential at  $(R_0, z_0)$  is

$$\delta\Phi_{\text{int}} \equiv \delta\Phi(R_0, z_0) = -4\pi G\rho(m^2)m\delta m \frac{\sqrt{1-e^2}}{e} \sin^{-1}(e). \quad (2.116a)$$

Similarly, if  $(R_0, z_0)$  lies outside the homoeoid,

$$\delta\Phi_{\text{ext}} \equiv \delta\Phi(R_0, z_0) = -4\pi G\rho(m^2)m\delta m \frac{\sqrt{1-e^2}}{e} \sin^{-1}(\text{sech } u_m). \quad (2.116b)$$

The potential of the entire body is the sum of contributions (2.116) from all the homoeoids that make up the body. If we define

$$\psi(m) \equiv \int_0^{m^2} dm^2 \rho(m^2), \quad (2.117)$$

the sum of the  $\delta\Phi_{\text{int}}$  is

$$\sum_{m>m_0} \delta\Phi_{\text{int}} = -2\pi G \frac{\sqrt{1-e^2}}{e} \sin^{-1}(e) [\psi(\infty) - \psi(m_0)], \quad (2.118)$$

where  $m_0$  is the label of the homoeoid that passes through  $(R_0, z_0)$ :

$$m_0^2 \equiv R_0^2 + \frac{z_0^2}{1-e^2}. \quad (2.119)$$

Similarly,

$$\sum_{m<m_0} \delta\Phi_{\text{ext}} = -2\pi G \frac{\sqrt{1-e^2}}{e} \int_0^{m_0^2} dm^2 \rho(m^2) \sin^{-1}(\text{sech } u_m). \quad (2.120)$$

Integrating equation (2.120) by parts,

$$\begin{aligned} \sum_{m<m_0} \delta\Phi_{\text{ext}} &= -2\pi G \frac{\sqrt{1-e^2}}{e} \\ &\times \left\{ [\psi(m) \sin^{-1}(\text{sech } u_m)]_{m=0}^{m_0} - \int_{m=0}^{m_0} \frac{\psi(m) d \text{sech } u_m}{\sqrt{1-\text{sech}^2 u_m}} \right\}. \end{aligned} \quad (2.121)$$

The quantity  $u_m$  appearing in equation (2.121) is a function  $u_m(R_0, z_0)$  by virtue of the condition that  $u_m$  label the spheroid through  $(R_0, z_0)$  that is confocal with the homoeoid  $m = \text{constant}$ . Let the  $\Delta$  parameter of the confocal family of spheroids containing  $m$  be  $\Delta_m$ , and let  $u_\star$  be the label of the homoeoid  $m$  within this family. Then  $m = \Delta_m \cosh u_\star$  and  $\sqrt{1-e^2} m = \Delta_m \sinh u_\star$ , so  $\Delta_m = me$ , and we have from equation (2.97) that

$$\frac{R_0^2}{\Delta_m^2 \cosh^2 u_m} + \frac{z_0^2}{\Delta_m^2 \sinh^2 u_m} = 1, \quad (2.122)$$

which implies  $\frac{R_0^2}{1 + \sinh^2 u_m} + \frac{z_0^2}{\sinh^2 u_m} = m^2 e^2$ .

This is the required equation for  $u_m$ . Thus, in particular,  $m = 0$  implies  $\sinh u_m = \infty$ , and  $m = m_0$  implies  $\sinh u_m = \sqrt{1-e^2}/e$ . Inserting these limits into equation (2.121), and adding the result to equation (2.116), we find

$$\begin{aligned} \Phi(R_0, z_0) = & -2\pi G \frac{\sqrt{1-e^2}}{e} \\ & \times \left( \psi(\infty) \sin^{-1} e - \int_{\sinh u_m = \sqrt{1-e^2}/e}^{\infty} \psi(m) \frac{d \sinh u_m}{1 + \sinh^2 u_m} \right). \end{aligned} \quad (2.123)$$

We can simplify this equation by defining a new variable of integration

$$\tau \equiv a_0^2 e^2 \left[ \sinh^2 u_m - \left( \frac{1}{e^2} - 1 \right) \right], \quad (2.124)$$

where  $a_0$  is any constant. Then equation (2.122) becomes

$$\frac{R_0^2}{\tau + a_0^2} + \frac{z_0^2}{\tau + c_0^2} = \frac{m^2}{a_0^2} \quad (c_0 \equiv \sqrt{1-e^2} a_0), \quad (2.125a)$$

and equation (2.123) becomes

$$\begin{aligned} \Phi(R_0, z_0) = & -2\pi G \frac{\sqrt{1-e^2}}{e} \\ & \times \left( \psi(\infty) \sin^{-1} e - \frac{a_0 e}{2} \int_0^\infty d\tau \frac{\psi(m)}{(\tau + a_0^2) \sqrt{\tau + c_0^2}} \right). \end{aligned} \quad (2.125b)$$

The integral in this equation gives the contributions to  $\Phi$  from homoeoids for which  $(R_0, z_0)$  is an exterior point, with  $\tau = 0$  corresponding to the homoeoid that touches  $(R_0, z_0)$ , and large  $\tau$  corresponding to small homoeoids.

It is instructive to apply equations (2.125) to the determination of the interior potential of a homogeneous spheroid of density  $\rho_0$  and eccentricity  $e$

**Table 2.1** Formulae for the dimensionless quantities  $I \equiv a_2 a_3 a_1^{-1} \int_0^\infty d\tau \Delta^{-1}$  and  $A_i \equiv a_1 a_2 a_3 \int_0^\infty d\tau \Delta^{-1} (a_i^2 + \tau)^{-1}$  that occur in equations (2.128) and Table 2.2.  $\Delta^2(\tau) \equiv \prod_{i=1}^3 (a_i^2 + \tau)$ . The functions  $F(\theta, k)$  and  $E(\theta, k)$  are elliptic integrals (Appendix C.4).

	$a_1 = a_2 > a_3$ (oblate)	$a_1 = a_2 < a_3$ (prolate)	$a_1 > a_2 > a_3$ (triaxial)
	$e \equiv \sqrt{1 - a_3^2/a_1^2}$	$e \equiv \sqrt{1 - a_1^2/a_3^2}$	$k \equiv \sqrt{\frac{a_1^2 - a_2^2}{a_1^2 - a_3^2}}; k'^2 \equiv 1 - k^2; \theta \equiv \cos^{-1} \left( \frac{a_3}{a_1} \right)$
$I$	$2 \frac{\sqrt{1-e^2}}{e} \sin^{-1} e$	$\frac{1}{e} \ln \left( \frac{1+e}{1-e} \right)$	$2 \frac{a_2 a_3}{a_1^2} \frac{F(\theta, k)}{\sin \theta}$
$A_1$	$\frac{\sqrt{1-e^2}}{e^2} \left[ \frac{\sin^{-1} e}{e} - \sqrt{1-e^2} \right]$	$\frac{1-e^2}{e^2} \left[ \frac{1}{1-e^2} - \frac{1}{2e} \ln \left( \frac{1+e}{1-e} \right) \right]$	$2 \frac{a_2 a_3}{a_1^2} \frac{F(\theta, k) - E(\theta, k)}{k^2 \sin^3 \theta}$
$A_2$	$= A_1$	$= A_1$	$2 \frac{a_2 a_3}{a_1^2} \frac{E(\theta, k) - k'^2 F(\theta, k) - (a_3/a_2) k^2 \sin \theta}{k^2 k'^2 \sin^3 \theta}$
$A_3$	$2 \frac{\sqrt{1-e^2}}{e^2} \left[ \frac{1}{\sqrt{1-e^2}} - \frac{\sin^{-1} e}{e} \right]$	$2 \frac{1-e^2}{e^2} \left[ \frac{1}{2e} \ln \left( \frac{1+e}{1-e} \right) - 1 \right]$	$2 \frac{a_2 a_3}{a_1^2} \frac{(a_2/a_3) \sin \theta - E(\theta, k)}{k'^2 \sin^3 \theta}$

**Table 2.2** Potentials and potential-energy tensors of ellipsoidal bodies

Thin shell	$\Phi(\mathbf{x}_{\text{int}}) = -\frac{Ga_1}{2a_2 a_3} I(\mathbf{a}) M_{\text{shell}}$	$\Phi(\mathbf{x}_{\text{ext}}) = -\frac{Ga'_1}{2a'_2 a'_3} I(\mathbf{a}') M_{\text{shell}}$
Homogeneous	$\Phi(\mathbf{x}_{\text{int}}) = -\pi G \rho [I(\mathbf{a}) a_1^2 - \sum_{i=1}^3 A_i(\mathbf{a}) x_i^2]$	$W_{ij} = -\frac{8}{15} \pi^2 G \rho^2 a_1 a_2 a_3 A_i a_i^2 \delta_{ij}$
	$\Phi(\mathbf{x}_{\text{ext}}) = -\pi G \rho \frac{a_1 a_2 a_3}{a'_1 a'_2 a'_3} [I(\mathbf{a}') a_1'^2 - \sum_{i=1}^3 A_i(\mathbf{a}') x_i^2]$	$W = -\frac{8}{15} \pi^2 G \rho^2 a_1^3 a_2 a_3 I$
Inhomogeneous	$\Phi(\mathbf{x}) = -\pi G \frac{a_2 a_3}{a_1} \int_0^\infty \frac{d\tau}{\Delta} \{ \psi(\infty) - \psi[m(\tau, \mathbf{x})] \}$	$W_{ij} = -2\pi^2 G \frac{a_2 a_3}{a_1^4} \mathcal{S} A_i a_i^2 \delta_{ij} \quad W = -2\pi^2 G \frac{a_2 a_3}{a_1^2} \mathcal{S} I \delta_{ij}$

NOTES:  $I$  and  $A_i$  as in Table 2.1.  $\mathbf{x}_{\text{int}}$  and  $\mathbf{x}_{\text{ext}}$  denote points on the interior or exterior of the ellipsoidal shell or body. If  $\sum_{i=1}^3 x_i^2/[a_i^2 + \lambda(\mathbf{x})] = 1$ , then  $a_i'^2 \equiv a_i^2 + \lambda(\mathbf{x})$ ;  $\Delta^2(\tau) \equiv \prod_{i=1}^3 (a_i^2 + \tau)$ ;  $m^2(\tau, \mathbf{x}) \equiv a_1^2 \sum_{i=1}^3 x_i^2/(a_i^2 + \tau)$ ;  $\psi(m) \equiv \int_0^{m^2} \rho(\mathbf{x}) dm^2(0, \mathbf{x})$ ;  $\mathcal{S} \equiv \int_0^\infty dm^2 \rho(m^2) \int_0^{m^2} dm'^2 m' \rho(m'^2) = \frac{1}{2} \int_0^\infty dm [\psi(\infty) - \psi(m)]^2$ .

that has semi-axes of lengths  $a_1$  and  $a_3 = \sqrt{1 - e^2} a_1$ . For this case, equation (2.117) yields

$$\psi(m) = \rho_0 \times \begin{cases} m^2 & (m^2 < a_1^2), \\ a_1^2 & (m^2 \geq a_1^2). \end{cases} \quad (2.126)$$

Equation (2.125b) takes on a particularly simple form if we set the arbitrary constant  $a_0$  equal to  $a_1$ . If  $(R_0, z_0)$  lies inside the spheroid,  $m(\tau)$  is always smaller than  $a_0$ . Hence, we may substitute from equation (2.125a) and (2.126) into (2.125b), to obtain

$$\begin{aligned} \Phi(R_0, z_0) = & -2\pi G \rho_0 a_1^2 \frac{\sqrt{1 - e^2}}{e} \left[ \sin^{-1} e \right. \\ & \left. - \frac{a_1 e}{2} \int_0^\infty \frac{d\tau}{(\tau + a_1^2) \sqrt{\tau + a_3^2}} \left( \frac{R_0^2}{\tau + a_1^2} + \frac{z_0^2}{\tau + a_3^2} \right) \right]. \end{aligned} \quad (2.127)$$

This potential is quadratic in the coordinates and may be written

$$\Phi(\mathbf{x}) = -\pi G \rho_0 (I a_1^2 - A_1 R^2 - A_3 z^2), \quad (2.128)$$

where the dimensionless coefficients  $I$  and  $A_i$  are given in Table 2.1. An expression for the exterior potential of the homogeneous spheroid is given in Table 2.2.

With the help of equations (2.117) and (2.125) we obtain the gravitational field generated by a spheroidal system as

$$\mathbf{g} = -\nabla \Phi = -\pi G \sqrt{1 - e^2} a_0 \int_0^\infty d\tau \frac{\rho(m^2) \nabla m^2}{(\tau + a_0^2) \sqrt{\tau + c_0^2}}, \quad (2.129a)$$

where

$$\nabla m^2 = 2a_0^2 \left( \frac{R}{\tau + a_0^2} \hat{\mathbf{e}}_R + \frac{z}{\tau + c_0^2} \hat{\mathbf{e}}_z \right). \quad (2.129b)$$

We may use equation (2.129) to find the circular speed  $v_c(R)$  in the equatorial plane of an oblate spheroidal galaxy. The radial component of the field (2.129) is

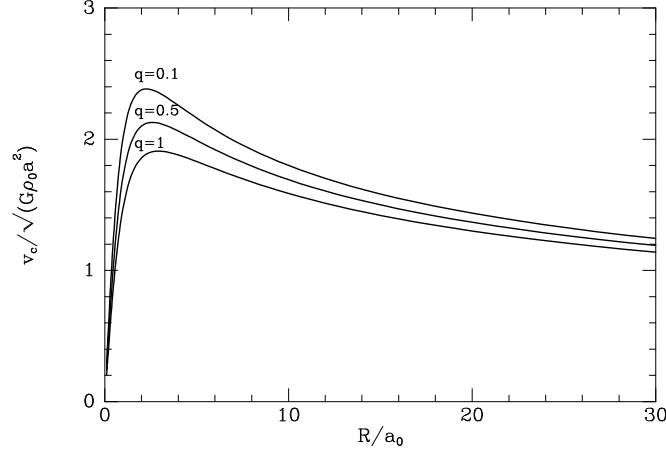
$$g_R(R, z) = -2\pi G \sqrt{1 - e^2} a_0^3 R \int_0^\infty d\tau \frac{\rho(m^2)}{(\tau + a_0^2)^2 \sqrt{\tau + c_0^2}}. \quad (2.130)$$

In the equatorial plane  $z = 0$ , equation (2.125a) yields

$$m = \frac{a_0 R}{\sqrt{\tau + a_0^2}}. \quad (2.131a)$$

Hence

$$\frac{d\tau}{(\tau + a_0^2)^2} = -\frac{2m}{R^2 a_0^2} dm, \quad (2.131b)$$



**Figure 2.13** Circular speed versus radius for three bodies with the same face-on projected density profile (the modified Hubble model, eq. 2.133) but different axis ratios  $q = c/a$ . Though all three bodies have the same mass inside a spheroid of given semi-major axis,  $v_c$  increases with flattening  $1 - q$ .

and equation (2.130) yields

$$v_c^2(R) = -Rg_R(R, 0) = 4\pi G \sqrt{1 - e^2} \int_0^R dm \frac{m^2 \rho(m^2)}{\sqrt{R^2 - m^2 e^2}}. \quad (2.132)$$

Let us see how these formulae work out in a specific case. Consider the oblate spheroidal density distribution

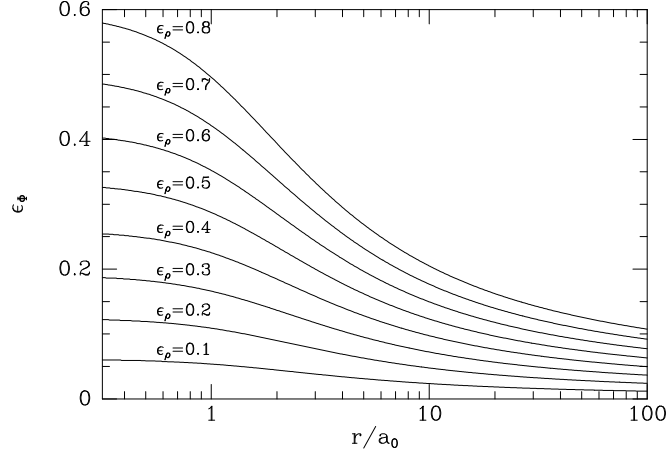
$$\rho(m^2) = \rho_0 \left[ 1 + \left( \frac{m}{a_0} \right)^2 \right]^{-3/2}, \quad (2.133)$$

where  $a_0$  is the core radius and the parameter  $e$  that appears in the definition (2.114) of  $m$  is the eccentricity of the system. In the limit  $e \rightarrow 0$  this reduces to the modified Hubble model (2.53). We substitute for  $\rho$  in equation (2.132) to obtain

$$v_c^2(R) = 4\pi G \rho_0 a_0^3 \frac{\sqrt{1 - e^2}}{e} \int_0^R \frac{m^2 dm}{(a_0^2 + m^2)^{3/2} \sqrt{R^2/e^2 - m^2}}. \quad (2.134)$$

By making the substitution

$$m = \frac{R \sin \theta}{e \sqrt{1 + (R/ea_0)^2 \cos^2 \theta}} \quad (2.135)$$



**Figure 2.14** The ellipticity  $\epsilon_\Phi$  of an equipotential surface versus the surface's semi-major axis length  $r$ . Each curve is labeled by the ellipticity  $\epsilon_\rho = 1 - q$  of the body with density (2.133) that generates the corresponding potential. Notice the rapidity with which the equipotential surfaces become spherical at large  $r/a_0$ .

one may show that the integral of equation (2.134) equals

$$\frac{ek}{R} [F(\theta_m, k) - E(\theta_m, k)], \quad (2.136)$$

where  $F$  and  $E$  are incomplete elliptic integrals (see Appendix C.4),

$$k \equiv \frac{R}{\sqrt{e^2 a_0^2 + R^2}}, \quad \text{and} \quad \theta_m \equiv \sin^{-1} \sqrt{\frac{e^2 a_0^2 + R^2}{a_0^2 + R^2}}. \quad (2.137a)$$

Hence

$$v_c^2(R) = 4\pi G \rho_0 a_0^3 \frac{\sqrt{1-e^2}}{R} k [F(\theta_m, k) - E(\theta_m, k)]. \quad (2.137b)$$

We may use this result to investigate how strongly a galaxy's circular speed is affected by its shape. In Figure 2.13 we plot the circular-speed curves of three galaxies whose density profiles are given by equation (2.133) for axis ratio  $q = \sqrt{1-e^2} = 1$  (spherical system),  $q = 0.5$  (E5 galaxy), and  $q = 0.1$  (the flatness characteristic of disk galaxies). The central density has in each case been adjusted so as to hold constant the mass  $M(a)$  interior to the spheroid of semi-major axis  $a$ . The peak circular speed of the  $q = 0.1$  model is about 20% higher than that of the spherical system because flattening the system increases the radial component of the force between a given mass element and a test mass in the equatorial plane.

For the density distribution defined by equation (2.133), Figure 2.14 shows the ellipticity  $\epsilon_\Phi$  of the isopotential surfaces for several values of the ellipticity  $\epsilon_\rho \equiv 1 - q$  of the density distribution. One sees that in the core  $r < a_0$ ,  $\epsilon_\Phi \gtrsim \frac{1}{2}\epsilon_\rho$ , while at a few core radii  $\epsilon_\Phi \approx \frac{1}{3}\epsilon_\rho$ , and at  $r \gg a$ ,  $\epsilon_\Phi$  rapidly approaches zero. Thus, in the region containing the bulk of the mass, the potential is generally flattened only about a third as much as the density, just as we found for logarithmic potentials in §2.3.2.

### 2.5.3 Potentials of ellipsoidal systems

The problem of calculating the gravitational potential of a body whose iso-density surfaces are similar, coaxial ellipsoids challenged some of the best minds of the eighteenth and nineteenth centuries—see Chandrasekhar (1969) for details. The general problem was solved by George Green (1793–1841), a Nottingham millwright, in 1835 using rather specialized geometrical methods that we shall not present here. Green’s results are natural generalizations of the ones we deduced above by a more accessible technique for axisymmetric systems. We now summarize the results for triaxial systems and refer readers to Kellogg (1953) or Chandrasekhar (1969) for proofs.

On surfaces of constant density, the variable

$$m^2 \equiv a_1^2 \sum_{i=1}^3 \frac{x_i^2}{a_i^2} \quad (2.138)$$

is constant, where  $(x_1, x_2, x_3)$  are Cartesian coordinates and  $a_1, a_2, a_3$  are the semi-axes of the ellipsoid. A thin shell of uniform density, whose inner and outer skins are the surfaces  $m$  and  $m + \delta m$ , generates an exterior potential that is constant on the ellipsoidal surfaces

$$m^2 = a_1^2 \sum_{i=1}^3 \frac{x_i^2}{a_i^2 + \tau}, \quad (2.139)$$

where  $\tau \geq 0$  labels the surfaces. (This is a straightforward extension of the homoeoid theorem proved in §2.5.1.) There is no gravitational field inside such a shell.

We may find the gravitational potential of any body in which  $\rho = \rho(m^2)$  by breaking the body down into thin triaxial homoeoids: the triaxial analog of equation (2.125b) is

$$\Phi(\mathbf{x}) = -\pi G \frac{a_2 a_3}{a_1} \int_0^\infty d\tau \frac{\psi(\infty) - \psi(m)}{\sqrt{(\tau + a_1^2)(\tau + a_2^2)(\tau + a_3^2)}}, \quad (2.140)$$

where  $\psi(m)$  is again defined by (2.117) and  $m = m(\mathbf{x}, \tau)$  through equation (2.139). Merritt & Fridman (1996) give expressions derived from (2.140)



for the gravitational potentials and fields of triaxial generalizations of the Dehnen models of §2.2.2g.

**(a) Ferrers potentials** A particularly simple application of equation (2.140) is to the case in which

$$\rho(m^2) = \begin{cases} \rho_0 (1 - m^2/a_1^2)^n & \text{for } m \leq a_1 \\ 0 & \text{for } m > a_1, \end{cases} \quad (2.141)$$

where  $m = m(\mathbf{x})$  through equation (2.138). By (2.117) we now have

$$\psi(\infty) - \psi(m) = \frac{\rho_0 a_1^2}{n+1} \left(1 - \frac{m^2}{a_1^2}\right)^{n+1} \quad (m \leq a_1). \quad (2.142)$$

Hence the internal potential of a body whose density is of the form (2.141) is

$$\begin{aligned} \Phi(\mathbf{x}) = & -\frac{\pi G \rho_0 a_1 a_2 a_3}{n+1} \int_0^\infty \frac{d\tau}{\sqrt{(\tau + a_1^2)(\tau + a_2^2)(\tau + a_3^2)}} \\ & \times \left(1 - \sum_{i=1}^3 \frac{x_i^2}{\tau + a_i^2}\right)^{n+1}. \end{aligned} \quad (2.143)$$

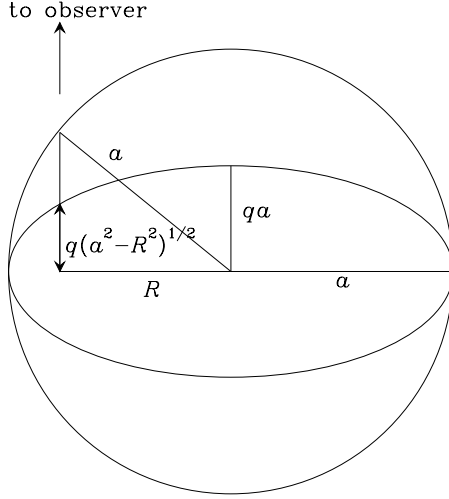
If  $n$  is an integer, the bracket involving  $\mathbf{x}$  in equation (2.143) can be multiplied out, and the potential at any point obtained as a sum of terms of the form  $A_{pqr} x_1^p x_2^q x_3^r$ , where the coefficients  $A_{pqr}$  are independent of  $\mathbf{x}$ . Potentials of this simple form are ideally suited to numerical studies of orbits in triaxial galaxies, such as we shall describe in §3.3. We shall refer to these as **Ferrers potentials**.

The  $n = 0$  Ferrers potential arises from a homogeneous ellipsoid with semi-axes  $a_1, a_2, a_3$ . Expressions for the interior and exterior potentials of such bodies can be derived from equations (2.140) and (2.143) and are given in Tables 2.1 and 2.2.

**(b) Potential-energy tensors of ellipsoidal systems** Roberts (1962) showed that for ellipsoidal bodies equation (2.22) has a remarkably simple form:

$$W_{jk} = -\pi^2 G \frac{a_2 a_3}{a_1^2} \left(\frac{a_j}{a_1}\right)^2 A_j \delta_{jk} \int_0^\infty dm [\psi(\infty) - \psi(m)]^2, \quad (2.144)$$

where the  $A_j$  are given in Table 2.1. Notice that the right side of equation (2.144) comprises a constant times the product of two factors: (i) a factor  $(a_j/a_1)^2 A_j \delta_{jk}$  that depends only on the axial ratios  $(a_2/a_1)$  etc.; and (ii) a factor  $\int dm [\psi(\infty) - \psi(m)]^2$  that is independent of the body's ellipticity and the same for all components of the tensor; this integral can be evaluated from a knowledge of the radial density structure alone. In particular, ratios



**Figure 2.15** A spheroid of axis ratio  $q$  and semi-major axis  $a$  is viewed along a line of sight that cuts the spheroid's equatorial plane perpendicularly at radius  $R$ . This line of sight cuts through the spheroid for a distance  $2q\sqrt{a^2 - R^2}$ .

of potential-energy terms, for example  $W_{11}/W_{33}$ , depend only on the body's ellipticity, and *are entirely independent of the radial density structure* so long as the density is stratified on similar ellipsoids. We shall exploit this useful result in §4.8.3. In Table 2.2 we give expressions for the potential-energy tensors of homogeneous ellipsoids.

## 2.6 The potentials of disks

Most of the light emitted by a typical spiral galaxy comes from a thin disk. Thus we anticipate that a substantial fraction of the galaxy's mass is concentrated in the disk, and it is therefore important to be able to calculate efficiently the gravitational field of a thin disk. We begin by investigating the potential of an idealized axisymmetric disk of zero thickness.

### 2.6.1 Disk potentials from homoeoids

We may consider any axisymmetric disk to be a very flat spheroid and use the formulae of §2.5.1 to obtain the potential. A homogeneous spheroid of density  $\rho$ , semi-axes  $a$  and  $c$ , and axial ratio  $q = c/a$  has mass  $M(a) = \frac{4}{3}\pi\rho qa^3$  (eq. 2.112) and surface density (see Figure 2.15)

$$\Sigma(a, R) = 2\rho q \sqrt{a^2 - R^2}, \quad (2.145)$$

where  $R$  is the usual cylindrical radius. Differentiating these expressions with respect to  $a$ , we obtain the mass  $\delta M(a)$  and the surface density  $\delta \Sigma(a, R)$  of

the thin homoeoid of density  $\rho$ , semi-major axis  $a$ , thickness  $\delta a$ , and axial ratio  $q$ :

$$\delta M(a) = 4\pi\rho qa^2\delta a \quad ; \quad \delta\Sigma(a, R) = \frac{2\rho qa}{\sqrt{a^2 - R^2}}\delta a. \quad (2.146)$$

If we now let  $q$  tend to zero while holding  $2\rho qa \equiv \Sigma_0$  constant, we obtain the mass and surface density of an infinitely flattened homoeoid:

$$\delta M(a) = 2\pi\Sigma_0 a \delta a \quad ; \quad \delta\Sigma(a, R) = \frac{\Sigma_0 \delta a}{\sqrt{a^2 - R^2}}. \quad (2.147)$$

We may construct a razor-thin disk of known surface density  $\Sigma(R)$  by finding the family of homoeoids whose combined surface density equals  $\Sigma(R)$  at all  $R$ . In mathematical language, we have to find the function  $\Sigma_0(a)$  that satisfies the integral equation

$$\Sigma(R) = \sum_{a \geq R} \delta\Sigma(a, R) = \int_R^\infty da \frac{\Sigma_0(a)}{\sqrt{a^2 - R^2}}. \quad (2.148a)$$

This is an Abel integral equation. Its solution is (eq. B.72)

$$\Sigma_0(a) = -\frac{2}{\pi} \frac{d}{da} \int_a^\infty dR \frac{R\Sigma(R)}{\sqrt{R^2 - a^2}}. \quad (2.148b)$$

Note that  $\Sigma_0(a)$  is *not* the same function as  $\Sigma(R)$ . In particular, some of the mass that lies interior to radius  $R$  comes from homoeoids having  $a > R$ . By Newton's third theorem, this portion of matter does not contribute to the gravitational force at  $R$ , because a point in the equatorial plane at radius  $R$  is an interior point of all homoeoids with  $a > R$ . Thus two disks can have identical surface-density distributions for  $R' < R$  and yet have very different force fields at  $R$ . In this respect disks differ from spherical distributions of mass, for which the force at  $r_0$  depends only on the density at  $r < r_0$ . In fact, the surface density of a disk at  $R' > R$  affects the attraction at  $R$  because the annulus of material exterior to  $R$  actually pulls a star placed at radius  $R$  outward, thus partially compensating the inward attraction of the interior matter. For example on the perimeter of a sharp-edged disk, the circular speed can be much higher than at the edge of a spherical body, or a more extended disk, with the same mass interior to this point.

We now calculate the potential of a thin disk by adding the potentials of the thin homoeoids into which we have decomposed it. By Gauss's theorem, the gravitational field is discontinuous across a sheet of finite surface density, but the potential is continuous. Consequently, the potential in the equatorial plane differs infinitesimally from the potential just above or below the disk.

Therefore we need only calculate the potential at points that are external to all homoeoids, and take the limit  $z \rightarrow 0$  to find the potential in the plane.

Equation (2.107) gives the potential outside a homoeoid of mass  $M$ . For a completely flattened homoeoid, we have  $e = 1$  and the mass is given in terms of  $\Sigma_0$  by the first of equations (2.147). Hence, at location  $(R, z)$  the potential of this homoeoid is

$$\delta\Phi(R, z) = -2\pi G \Sigma_0 \delta a \sin^{-1}(\operatorname{sech} u), \quad (2.149)$$

where  $u$  is determined by the equations  $R = a \cosh u \sin v$ ,  $z = a \sinh u \cos v$  (eq. 2.96). Eliminating  $v$  from these equations by using  $1 = \cos^2 v + \sin^2 v$ , we obtain a quadratic equation for  $\cosh^2 u$ :

$$a^2 \cosh^4 u - (R^2 + z^2 + a^2) \cosh^2 u + R^2 = 0. \quad (2.150)$$

The root that we require is the one with  $\cosh^2 u \geq 1$ , which is

$$\begin{aligned} \cosh^2 u &= \frac{1}{2a^2} \left[ R^2 + z^2 + a^2 + \sqrt{(R^2 + z^2 + a^2)^2 - 4a^2 R^2} \right] \\ &= \frac{1}{4a^2} \left[ \sqrt{z^2 + (a+R)^2} + \sqrt{z^2 + (a-R)^2} \right]^2. \end{aligned} \quad (2.151)$$

Taking the square root of both sides and substituting the result into equation (2.149), we obtain

$$\delta\Phi = -2\pi G \Sigma_0 \delta a \sin^{-1} \left( \frac{2a}{\sqrt{z^2 + (a+R)^2} + \sqrt{z^2 + (a-R)^2}} \right). \quad (2.152)$$

Finally, the potential of an axisymmetric disk of arbitrary surface-density profile is obtained by combining this result with equations (2.148), which decompose a disk into homoeoids. We have (Cuddeford 1993)

$$\Phi(R, z) = 4G \int_0^\infty da \sin^{-1} \left( \frac{2a}{\sqrt{+} + \sqrt{-}} \right) \frac{d}{da} \int_a^\infty dR' \frac{R' \Sigma(R')}{\sqrt{R'^2 - a^2}}. \quad (2.153a)$$

where

$$\sqrt{\pm} \equiv \sqrt{z^2 + (a \pm R)^2}. \quad (2.153b)$$

An alternative form is obtained by integrating by parts:

$$\begin{aligned} \Phi(R, z) &= -2\sqrt{2}G \int_0^\infty da \frac{[(a+R)/\sqrt{+}] - [(a-R)/\sqrt{-}]}{\sqrt{R^2 - z^2 - a^2 + \sqrt{+}\sqrt{-}}} \\ &\quad \times \int_a^\infty dR' \frac{R' \Sigma(R')}{\sqrt{R'^2 - a^2}}. \end{aligned} \quad (2.154)$$

This form does not require differentiation of the term that depends on the surface density. Equations (2.153a) and (2.154) are numerically convenient because the inner integrals depend only on  $a$ , and thus can be tabulated on a grid of values of  $a$  at the outset. Then only a single integral is required to evaluate  $\Phi$  at each fresh point  $(R, z)$ . Moreover, the integrands do not oscillate in sign, so numerical integrations converge rapidly.

We are particularly interested in the value of the potential in the equatorial plane. Consider first the case  $a > R$ . For small  $z$ , it is easy to see that  $\sqrt{\pm} \rightarrow (a \pm R) + O(z^2)$ , with the consequence that the numerator of the fraction in the first line in equation (2.154) vanishes like  $z^2$  as  $z \rightarrow 0$ . In the same limit the denominator vanishes too, but more slowly, like  $z$ . Hence, for  $a > R$  the integrand tends to zero with  $z$  and we readily find that

$$\Phi(R, 0) = -4G \int_0^R \frac{da}{\sqrt{R^2 - a^2}} \int_a^\infty dR' \frac{R' \Sigma(R')}{\sqrt{R'^2 - a^2}}. \quad (2.155)$$

To obtain the circular speed at radius  $R$  in the equatorial plane, the natural procedure is to differentiate this expression with respect to  $R$ , which is the upper limit of the outer integral. The usual formula for differentiating such an integral requires us to evaluate the integrand at  $a = R$ , when it diverges. Instead of resolving this problem we return to equation (2.153a), in which we may put  $z$  to zero without any awkwardness, to obtain

$$\Phi(R, 0) = 4G \int_0^\infty da \sin^{-1} \left( \frac{2a}{(a+R) + |a-R|} \right) \frac{d}{da} \int_a^\infty dR' \frac{R' \Sigma(R')}{\sqrt{R'^2 - a^2}}. \quad (2.156)$$

The argument of  $\sin^{-1}$  is unity for  $R < a$  and  $a/R$  otherwise. Consequently, when we differentiate with respect to  $R$  we obtain

$$v_c^2(R) = R \frac{\partial \Phi}{\partial R} = -4G \int_0^R da \frac{a}{\sqrt{R^2 - a^2}} \frac{d}{da} \int_a^\infty dR' \frac{R' \Sigma(R')}{\sqrt{R'^2 - a^2}}. \quad (2.157)$$

**(a) The Mestel disk** As a simple application of equation (2.157) consider a disk in which the surface density is given by

$$\Sigma(R) = \begin{cases} \frac{v_0^2}{2\pi GR} & \text{for } R < R_{\max} \\ 0 & \text{otherwise,} \end{cases} \quad (2.158)$$

where  $v_0$  and  $R_{\max}$  are constants with dimensions of velocity and length, respectively. For this surface density the inner integral in equation (2.157) is proportional to  $\cosh^{-1}(R_{\max}/a)$ , so its derivative with respect to  $a$  follows from

$$\frac{d}{da} \cosh^{-1}(R_{\max}/a) = -\frac{R_{\max}}{a\sqrt{R_{\max}^2 - a^2}}. \quad (2.159)$$

We let the outer radius of the disk,  $R_{\max}$ , tend to infinity. In this limit the disk becomes the **Mestel disk** (Mestel 1963) and the derivative (2.159) becomes  $-1/a$ . Substituting this value into equation (2.157), we find that the circular speed in the Mestel disk is

$$v_c^2 = \frac{2v_0^2}{\pi} \int_0^R \frac{da}{\sqrt{R^2 - a^2}} = v_0^2. \quad (2.160)$$

Hence, the circular speed of a disk in which the surface density is inversely proportional to radius is independent of radius. Moreover, for this surface-density law,  $v_c(R)$  is given by the simple formula

$$v_c^2(R) = \frac{GM(R)}{R}, \quad (2.161a)$$

where

$$M(R) = 2\pi \int_0^R dR' R' \Sigma(R') = \frac{v_0^2 R}{G} \quad (2.161b)$$

is the mass interior to  $R$ . This is precisely analogous to equation (2.29) for a spherical system. Although we have argued that for general disks the circular speed is affected by the mass exterior to  $R$ , for the particular case of the Mestel disk the simple formula (2.161a) happens to give the correct answer—we are aware of no other disks with this property.

**(b) The exponential disk** The surface-brightness profiles of many galactic disks are approximately exponential in form (eq. 1.7). Let us use the results just derived to calculate the potential that such a disk would generate if its surface mass density were also exponential. Setting

$$\Sigma(R) = \Sigma_0 e^{-R/R_d}, \quad (2.162)$$

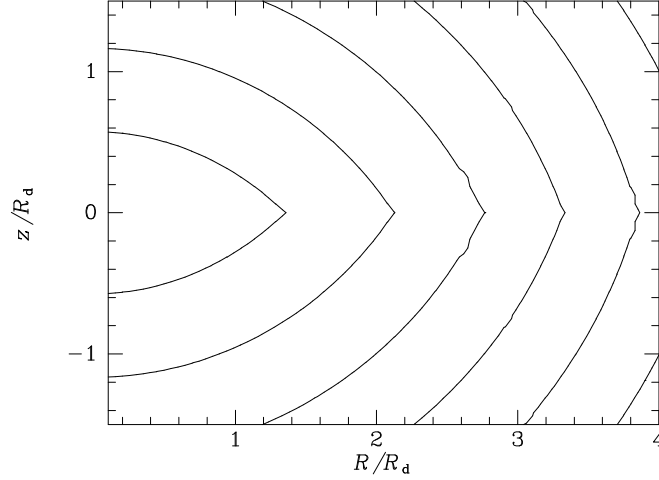
we use equation (C.69) to show that the inner integral in equations (2.153a) to (2.157) is

$$\int_a^\infty dR' \frac{R' \Sigma_0 e^{-R'/R_d}}{\sqrt{R'^2 - a^2}} = \Sigma_0 a K_1(a/R_d), \quad (2.163)$$

where  $K_1$  is a modified Bessel function (Appendix C.7). Figure 2.16 shows contours of the potential that one obtains by substituting this formula into equation (2.154).

The potential in the equatorial plane is given by substituting (2.163) into equation (2.155):

$$\begin{aligned} \Phi(R, 0) &= -4G\Sigma_0 \int_0^R da \frac{a K_1(a/R_d)}{\sqrt{R^2 - a^2}} \\ &= -\pi G\Sigma_0 R [I_0(y)K_1(y) - I_1(y)K_0(y)], \end{aligned} \quad (2.164a)$$



**Figure 2.16** Contours in the  $(R, z)$  plane of constant potential  $\Phi$  for a razor-thin exponential disk. The contour levels are  $GM_d/R_d$  divided by 1.5, 2, 2.5, ..., where  $M_d = 2\pi\Sigma_0 R_d^2$  is the mass of the disk.

where

$$y \equiv \frac{R}{2R_d}, \quad (2.164b)$$

we have used equation (C.70), and the  $I_n$  are modified Bessel functions (Appendix C.7).

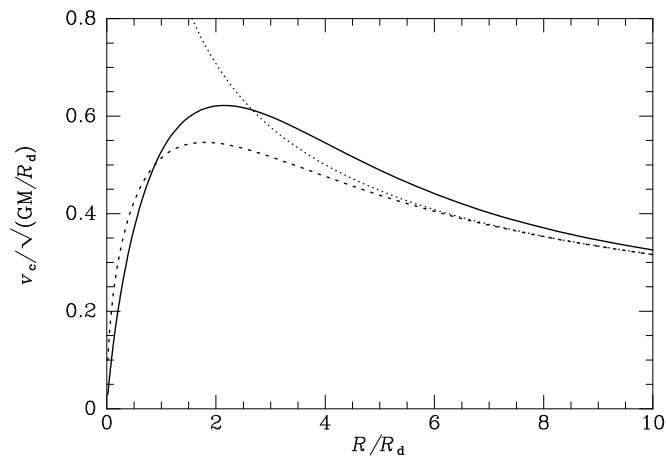
If we differentiate equation (2.164a) with respect to  $R$ , we obtain the circular speed of the exponential disk (Freeman 1970):

$$v_c^2(R) = R \frac{\partial \Phi}{\partial R} = 4\pi G \Sigma_0 R_d y^2 [I_0(y)K_0(y) - I_1(y)K_1(y)]. \quad (2.165)$$

In Figure 2.17 we show this circular speed together with the circular speed of the spherical body that has as much mass  $M_s(r)$  interior to  $r = R$  as the exponential disk, that is,

$$\begin{aligned} M_s(R) = M_d(R) &= 2\pi \int_0^R dR' R' \Sigma_0 e^{-R'/R_d} \\ &= 2\pi \Sigma_0 R_d^2 \left[ 1 - e^{-R/R_d} \left( 1 + \frac{R}{R_d} \right) \right]. \end{aligned} \quad (2.166)$$

The exponential disk achieves a peak circular speed that is about 15% higher than that of the equivalent spherical distribution. The dotted line in Figure 2.17 gives the Keplerian circular speed for a system in which the entire mass of the disk is concentrated at the center. Notice that the disk's circular



**Figure 2.17** The circular-speed curves of: an exponential disk (full curve); a point with the same total mass (dotted curve); the spherical body for which  $M(r)$  is given by equation (2.166) (dashed curve).

speed approaches the Keplerian speed only slowly and from above, whereas the circular speed of the equivalent spherical distribution tends rapidly to the Keplerian speed from below.

**(c) Thick disks** Although galactic disks are thin in the sense that the density falls off much faster perpendicular to the equatorial plane than in the radial direction within the plane, it is frequently essential to take into account the non-zero thickness of the disk in the perpendicular direction. For example, the dynamics of the solar neighborhood, and of the Sun itself, would be very different if the Galactic disk were razor-thin rather than some hundreds of parsecs thick. The techniques for deriving the potential of a spheroidal system that are described in §2.5.1 are not directly applicable to this problem because the density of a galactic disk is generally not constant on spheroids. The multipole expansion described in §2.4 is also not suitable, because it converges slowly for flat systems. Here we describe how to calculate the potential of a disk when the density is of the form

$$\rho(R, z) = \Sigma(R)\zeta(z). \quad (2.167)$$

Physically, this formula implies that a cross-section through the disk always has the same shape, no matter at what radius it is taken. In particular, the characteristic scale height of the disk is independent of  $R$ , an assumption that is in reasonable agreement with observations of edge-on disk galaxies (BM §4.4.3).

It is convenient to normalize the function  $\zeta(z)$  in equation (2.167) such that  $\int dz \zeta(z) = 1$ . With this normalization,  $\Sigma(R)$  is the total surface density



and  $\Sigma(R)\zeta(z) dz$  is the surface density of the layer of material of thickness  $dz$  that lies a distance  $z$  above the equatorial plane. Let  $\Phi_0(R, z)$  be the potential that would be generated by a razor-thin disk with surface density  $\Sigma(R)$  that lay in the plane  $z = 0$ . Then the potential generated at  $(R, z)$  by the layer at distance  $z'$  from the plane is

$$d\Phi(R, z) = dz' \Phi_0(R, z - z')\zeta(z'). \quad (2.168)$$

Adding the contributions to the disk's potential from every layer, we obtain for the overall potential

$$\Phi(R, z) = \int_{-\infty}^{\infty} dz' \zeta(z') \Phi_0(R, z - z'), \quad (2.169)$$

where  $\Phi_0$  can be obtained from either (2.153a) or (2.154).

Consider, for example, the exponential disk (2.162). Then inserting into equation (2.153a) the derivative  $-\Sigma_0(a/R_d)K_0(a/R_d)$  with respect to  $a$  of equation (2.163), and then substituting the resulting value of  $\Phi_0$  into equation (2.169), we find

$$\Phi(R, z) = -\frac{4G\Sigma_0}{R_d} \int_{-\infty}^{\infty} dz' \zeta(z') \int_0^{\infty} da \sin^{-1} \left( \frac{2a}{\sqrt{+} + \sqrt{-}} \right) a K_0(a/R_d), \quad (2.170)$$

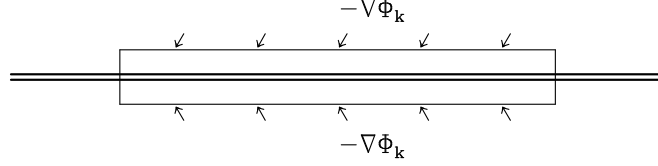
where  $\sqrt{\pm}$  is defined by equation (2.153b) with  $z$  replaced by  $z - z'$ . We shall use this formula in §2.7 below.

### 2.6.2 Disk potentials from Bessel functions

Disk galaxies generally contain non-axisymmetric features such as a bar or spiral arms. Hence it is essential to know how to calculate the potential of a flattened, non-axisymmetric system. Several methods may be employed and none is ideal for every problem. Evans & de Zeeuw (1992) show how the results of §2.5.3 may be used to obtain the potential of razor-thin, elliptical disks in analogy with the work of §2.6.1. However, the technique they present is not easy to apply, so in the rest of this section we present three alternatives, each with its own strengths when applied to particular problems. All these methods can also be applied to axisymmetric disks, and all yield alternatives to the formulae of §2.5.

Above and below an isolated razor-thin disk the gravitational potential satisfies Laplace's equation,  $\nabla^2\Phi = 0$ , with appropriate boundary conditions on the disk and at infinity. In cylindrical coordinates Laplace's equation is (eq. B.52)

$$\frac{1}{R} \frac{\partial}{\partial R} \left( R \frac{\partial \Phi}{\partial R} \right) + \frac{1}{R^2} \frac{\partial^2 \Phi}{\partial \phi^2} + \frac{\partial^2 \Phi}{\partial z^2} = 0. \quad (2.171)$$



**Figure 2.18** The disk mass within the box shown in cross-section equals  $-(4\pi G)^{-1}$  times the integral of the normal component of  $\nabla\Phi_{km}$  over the surface of the box. The horizontal component of  $\nabla\Phi_{km}$  is due to the gravitational attraction from the rest of the galaxy.

Writing

$$\Phi(R, z) = J(R)F(\phi)Z(z), \quad (2.172)$$

we obtain by the method of separation of variables (see §2.4)

$$\frac{1}{J(R)R} \frac{d}{dR} \left( R \frac{dJ}{dR} \right) + \frac{1}{F(\phi)R^2} \frac{d^2 F}{d\phi^2} = -\frac{1}{Z(z)} \frac{d^2 Z}{dz^2} = -k^2, \quad (2.173)$$

where  $k$  is an arbitrary real or complex number. Thus

$$\begin{aligned} 0 &= \frac{d^2 Z}{dz^2} - k^2 Z, \\ 0 &= \frac{1}{J(R)R} \frac{d}{dR} \left( R \frac{dJ}{dR} \right) + \frac{1}{F(\phi)R^2} \frac{d^2 F}{d\phi^2} + k^2. \end{aligned} \quad (2.174)$$

The first of equations (2.174) may be immediately integrated to

$$Z(z) = e^{\pm kz}. \quad (2.175)$$

When multiplied by  $R^2$ , the second of equations (2.174) separates into equations for  $J$  and  $F$ :

$$\begin{aligned} 0 &= \frac{d^2 F}{d\phi^2} + m^2 F \\ 0 &= R \frac{d}{dR} \left( R \frac{dJ}{dR} \right) + k^2 R^2 J(R) - m^2 J. \end{aligned} \quad (2.176)$$

The first of these equations trivially yields  $F \propto e^{im\phi}$ , where  $m$  is an integer. If we make the substitution  $u \equiv kR$ , the second equation simplifies to

$$u \frac{d}{du} \left( u \frac{dJ}{du} \right) + (u^2 - m^2)J(u) = 0. \quad (2.177)$$

We require the solution of this equation to be finite at  $u = 0$  ( $R = 0$ ). This solution is conventionally written  $J_m(u) = J_m(kR)$  and is called the Bessel function of order  $m$  (Appendix C.7). Summarizing these results, we have that the functions

$$\Phi_{\pm}(R, z) = e^{i(m\phi \pm kz)} J_m(kR) \quad (2.178)$$

are solutions of  $\nabla^2 \Phi = 0$ .

Now consider the function

$$\Phi_{km}(R, z) = e^{im\phi - k|z|} J_m(kR), \quad (2.179a)$$

where  $k$  is real and positive.  $\Phi_{km}$  satisfies all the conditions required for it to be the potential generated by an isolated density distribution: it is non-singular at  $R = 0$ , periodic in  $\phi$  and vanishes at large distances from the origin. Furthermore, for  $z > 0$ ,  $\Phi_{km}$  coincides with  $\Phi_-$ , and for  $z < 0$ ,  $\Phi_{km}$  coincides with  $\Phi_+$ . Therefore,  $\Phi_{km}$  solves  $\nabla^2 \Phi = 0$  everywhere except in the plane  $z = 0$ . At  $z = 0$ ,  $\Phi_{km}$  does not satisfy Laplace's equation because its gradient suffers a discontinuity. Figure 2.18 illustrates how we may use Gauss's theorem (eq. 2.12) to evaluate the surface density  $\Sigma_{km}(R, \phi)$  of the sheet that generates this discontinuity. We have that

$$\lim_{z \rightarrow 0+} \frac{\partial \Phi_{km}}{\partial z} = -k e^{im\phi} J_m(kR) \quad \text{and} \quad \lim_{z \rightarrow 0-} \frac{\partial \Phi_{km}}{\partial z} = +k e^{im\phi} J_m(kR). \quad (2.180)$$

The integral of  $\nabla \Phi_{km}$  over the closed unit surface that is shown in the figure must equal  $4\pi G \Sigma_{km}$  from which it follows that

$$\Sigma_{km}(R, \phi) = -\frac{k}{2\pi G} e^{im\phi} J_m(kR). \quad (2.179b)$$

We now use equations (2.179) to find the potential generated by a disk of arbitrary surface density  $\Sigma(R, \phi)$ . If we can find functions  $S_m(k)$  such that

$$\begin{aligned} \Sigma(R, \phi) &= \sum_{m=-\infty}^{\infty} \int_0^{\infty} dk S_m(k) \Sigma_{km}(R, \phi) \\ &= -\frac{1}{2\pi G} \sum_{m=-\infty}^{\infty} \int_0^{\infty} dk k S_m(k) e^{im\phi} J_m(kR), \end{aligned} \quad (2.181)$$

then we will have

$$\begin{aligned} \Phi(R, \phi, z) &= \sum_{m=-\infty}^{\infty} \int_0^{\infty} dk S_m(k) \Phi_{km}(R, \phi, z) \\ &= \sum_{m=-\infty}^{\infty} \int_0^{\infty} dk S_m(k) J_m(kR) e^{im\phi - k|z|}. \end{aligned} \quad (2.182)$$

Multiplying (2.181) through by  $e^{-im'\phi}$  and averaging over  $\phi$ , we obtain

$$\Sigma_{m'}(R) \equiv \frac{1}{2\pi} \int_0^{2\pi} d\phi e^{-im'\phi} \Sigma(R, \phi) = -\frac{1}{2\pi G} \int_0^{\infty} dk k S_{m'}(k) J_{m'}(kR). \quad (2.183)$$

Equation (2.183) states that  $S_m(k)$  is the  $m$ th-order Hankel transform of  $-2\pi G \Sigma_m$  (eq. C.60b). Hankel transforms have properties that are similar to those of the familiar Fourier transforms (Appendix B.4). In particular, they may be inverted by use of equation (C.60a). We find

$$S_m(k) = -2\pi G \int_0^\infty dR R J_m(kR) \Sigma_m(R). \quad (2.184)$$

When we eliminate  $S_m(k)$  between this equation and (2.182), we obtain finally

$$\Phi(R, \phi, z) = -2\pi G \sum_{m=-\infty}^{\infty} \int_0^\infty dk e^{im\phi - k|z|} J_m(kR) \int_0^\infty dR' R' J_m(kR') \Sigma_m(R'). \quad (2.185)$$

**Application to axisymmetric disks** Potential-density pairs for axisymmetric disks are obtained by setting  $z = 0$  in equation (2.182) and restricting the sum to the case  $m = 0$ . We have (Toomre 1963)

$$\Phi(R, 0) = \int_0^\infty dk S_0(k) J_0(kR). \quad (2.186)$$

Differentiating with respect to  $R$  and using the identity  $dJ_0(x)/dx = -J_1(x)$  (eq. C.58), we obtain

$$v_c^2(R) = R \frac{\partial \Phi}{\partial R} = -R \int_0^\infty dk k S_0(k) J_1(kR). \quad (2.187)$$

Substituting for  $S_0(k)$  from equation (2.184) this can be rewritten

$$v_c^2(R) = 2\pi G R \int_0^\infty dk k J_1(kR) \int_0^\infty dR' R' \Sigma(R') J_0(kR'). \quad (2.188)$$

Applying to equation (2.187) the inversion formula for Hankel transforms (eqs. C.60) we find

$$S_0(k) = - \int_0^\infty dR' v_c^2(R') J_1(kR'). \quad (2.189)$$

Substituting this expression for  $S_0$  into equation (2.181), we have

$$\Sigma(R) = \frac{1}{2\pi G} \int_0^\infty dk k J_0(kR) \int_0^\infty dR' v_c^2(R') J_1(kR'). \quad (2.190)$$

Comparison of this formula with our expression (2.188) for the reverse relation reveals complete symmetry between the quantities  $v_c(R)$  and  $2\pi G R \Sigma(R)$

(Kalnajs 1999). Hence our mathematics seems to be saying that it is as easy to determine a disk's surface density from measurements of its circular speed, as to obtain the circular speed from the surface density. Unfortunately, observational constraints destroy this symmetry. The key point is that the left side of either equation (2.188) or (2.190) can be determined at any given value of  $R$  only if the variable on the right side can be measured out to radii at which its value becomes negligible. The surface density declines rapidly with radius, so equation (2.188) can be used to derive accurate values of  $v_c$ . Circular speeds, by contrast, decline little if at all out to the largest observable radii. Consequently, in practice we cannot obtain the data needed to determine  $\Sigma$  accurately from equation (2.190).

### 2.6.3 Disk potentials from logarithmic spirals

An alternative technique for finding non-axisymmetric potential-density pairs was introduced by Kalnajs (1971). The potential  $\Phi(R, \phi)$  at any point in the plane of a disk is

$$\begin{aligned}\Phi(R, \phi) &= -G \int dR' R' \int d\phi' \frac{\Sigma(R', \phi')}{|\mathbf{x} - \mathbf{x}'|} \\ &= -G \int_0^\infty dR' R' \int_0^{2\pi} d\phi' \frac{\Sigma(R', \phi')}{\sqrt{R'^2 + R^2 - 2RR' \cos(\phi' - \phi)}}.\end{aligned}\quad (2.191)$$

The integral in this expression can be simplified if we define a new radial coordinate,

$$u \equiv \ln R, \quad (2.192)$$

and introduce the **reduced potential**  $V$  and the **reduced surface density**  $S$  by

$$\begin{aligned}R^{1/2}\Phi &\equiv V(u, \phi) = e^{u/2}\Phi[R(u), \phi] \\ R^{3/2}\Sigma &\equiv S(u, \phi) = e^{3u/2}\Sigma[R(u), \phi].\end{aligned}\quad (2.193)$$

With these substitutions (2.191) becomes

$$V(u, \phi) = -G \int_{-\infty}^\infty du' \int_0^{2\pi} d\phi' K(u - u', \phi - \phi') S(u', \phi'), \quad (2.194a)$$

where

$$K(u - u', \phi - \phi') \equiv \frac{1}{\sqrt{2}\sqrt{\cosh(u - u') - \cos(\phi - \phi')}}. \quad (2.194b)$$

Now consider the reduced potential  $V_{\alpha m}(u, \phi)$  that is generated by the particular reduced surface density

$$S_{\alpha m}(u, \phi) = e^{i(\alpha u + m\phi)}, \quad (2.195)$$

where  $\alpha$  is a real number and  $m$  is an integer. We have

$$\begin{aligned} V_{\alpha m}(u, \phi) &= -G \int_{-\infty}^{\infty} du' \int_0^{2\pi} d\phi' K(u - u', \phi - \phi') e^{i(\alpha u' + m\phi')} \\ &= -G e^{i(\alpha u + m\phi)} \int_{-\infty}^{\infty} du' \int_0^{2\pi} d\phi' K(u - u', \phi - \phi') e^{i[\alpha(u' - u) + m(\phi' - \phi)]}. \end{aligned} \quad (2.196)$$

If we change to new variables of integration  $u'' \equiv u - u'$  and  $\phi'' \equiv \phi - \phi'$ , equation (2.196) becomes

$$V_{\alpha m} = -GN(\alpha, m) e^{i(\alpha u + m\phi)}, \quad (2.197)$$

where

$$\begin{aligned} N(\alpha, m) &\equiv \int_{-\infty}^{\infty} du'' \int_0^{2\pi} d\phi'' K(u'', \phi'') e^{-i(\alpha u'' + m\phi'')} \\ &= \pi \frac{(\frac{1}{2}m - \frac{3}{4} + \frac{1}{2}i\alpha)! (\frac{1}{2}m - \frac{3}{4} - \frac{1}{2}i\alpha)!}{(\frac{1}{2}m - \frac{1}{4} + \frac{1}{2}i\alpha)! (\frac{1}{2}m - \frac{1}{4} - \frac{1}{2}i\alpha)!}. \end{aligned} \quad (2.198)$$

The kernel  $N(\alpha, m)$  is real and even in both  $\alpha$  and  $m$ .<sup>7</sup> The reduced potential generated by an arbitrary linear combination

$$S(u, \phi) \equiv \sum_{m=-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{d\alpha}{2\pi} A_m(\alpha) e^{i(\alpha u + m\phi)} \quad (2.199a)$$

of surface densities of the form (2.195) is

$$V(u, \phi) = -G \sum_{m=-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{d\alpha}{2\pi} N(\alpha, m) A_m(\alpha) e^{i(\alpha u + m\phi)}. \quad (2.199b)$$

Furthermore, equation (2.199a) states that  $A_m(\alpha)$  is nothing but the Fourier transform of the reduced surface density  $S(u, \phi)$  (eqs. B.64 and B.67). Consequently,

$$A_m(\alpha) = \frac{1}{2\pi} \int_{-\infty}^{\infty} du \int_0^{2\pi} d\phi S(u, \phi) e^{-i(\alpha u + m\phi)}. \quad (2.199c)$$

So we may use equations (2.199) to obtain the potential in the plane  $z = 0$  that is generated by any distribution of surface density. Since  $\alpha u + m\phi = \alpha \ln R + m\phi$  is constant on **logarithmic spirals**, equations (2.199) determine the potential of a disk by decomposing the density into spirals. Note that this derivation does not produce an expression for the value of the potential away from the plane  $z = 0$  (Problem 2.19 remedies this defect).

---

<sup>7</sup> These statements can be proved using equations (C.12) and (C.14).

#### 2.6.4 Disk potentials from oblate spheroidal coordinates

In some situations we require the potential of a disk with a sharp outer edge. If the outer edge is circular, this problem can be efficiently solved using the oblate spheroidal coordinates of §2.5.1.<sup>8</sup> We substitute  $\Phi = U(u)V(v)e^{im\phi}$  into Laplace's equation using (2.101), and separate variables in the usual way. Then we find that  $U$  and  $V$  satisfy

$$\frac{m^2}{\sin^2 v} - \frac{1}{V \sin v} \frac{d}{dv} \left( \sin v \frac{dV}{dv} \right) = l(l+1), \quad (2.200a)$$

$$\frac{m^2}{\cosh^2 u} + \frac{1}{U \cosh u} \frac{d}{du} \left( \cosh u \frac{dU}{du} \right) = l(l+1), \quad (2.200b)$$

where  $l(l+1)$  is the separation constant. The left side of equation (2.200a) is the same as the right side of equation (2.80) with  $V(v)$  substituted for  $P(\theta)$ . Furthermore, the boundary conditions at  $v = 0$  or  $\pi$  (i.e., along the  $z$  axis) that must be satisfied by a physically acceptable function  $V(v)$  are the same as the conditions we imposed on  $P(\theta)$ . Hence

$$V(v, \phi) = V_{lm} Y_l^m(v, \phi), \quad (2.201)$$

where  $V_{lm}$  is a constant and  $Y_l^m$  is the spherical harmonic defined by equation (C.42). Since the potential must be symmetrical about the plane of the disk  $v = \pi/2$ , we must restrict ourselves to values of  $l$  and  $m$  for which  $Y_l^m(v, \phi)$  is an even function of  $\cos v$ . Hence we require  $l - m$  to be even.

If we change the independent variable in equation (2.200b) to  $x = i \sinh u$ , the equation becomes the associated Legendre equation (2.81b) with  $x$  now pure imaginary. If the potential is to vanish at infinity (large  $u$ ), it must be proportional to the solution of equation (2.81b) that vanishes at large  $x$ . This is written  $Q_l^m(x)$  (see Appendix C.5). Hence the functions

$$\Phi_{lm}(u, v, \phi) \equiv \frac{V_{lm}}{Q_l^m(0)} Q_l^m(i \sinh u) Y_l^m(v, \phi) \quad (l - m \text{ even}) \quad (2.202)$$

satisfy Laplace's equation everywhere outside the excluded disk  $u = 0$  ( $z = 0$ ,  $R \leq \Delta$ ) and vanish at infinity. However, there is a discontinuity in the gradient of  $\Phi_{lm}$  on the excluded disk. By Gauss's theorem, this discontinuity is generated by a surface density  $\Sigma_{lm}(v, \phi)$ ; thus from equation (2.100) we have

$$\begin{aligned} 4\pi G \Sigma_{lm}(v, \phi) &= 2 (\hat{\mathbf{e}}_u \cdot \nabla \Phi_{lm})_{u \rightarrow 0+} = \frac{2}{\Delta |\cos v|} \lim_{u \rightarrow 0+} \left( \frac{\partial \Phi_{lm}}{\partial u} \right) \\ &= 2V_{lm} i \lim_{x \rightarrow 0} \left[ \frac{d \ln Q_l^m(x)}{dx} \right] \frac{Y_l^m(v, \phi)}{\Delta |\cos v|}, \end{aligned} \quad (2.203)$$

<sup>8</sup>In the more general case in which the outer edge is an ellipse, one uses ellipsoidal coordinates (Morse & Feshbach 1953; Tremaine 1976b; Evans & de Zeeuw 1992).

where  $x$  approaches zero along the positive imaginary axis. Using equations (C.31) and (C.13) to evaluate the limit in equation (2.203), we then find (Hunter 1963)

$$\Sigma_{lm} = -\frac{2V_{lm}}{\pi^2 G \Delta g_{lm}} \frac{Y_l^m(v, \phi)}{|\cos v|}, \quad (2.204a)$$

where

$$g_{lm} \equiv \frac{(l+m)!(l-m)!}{2^{2l-1} \left[ \left( \frac{l+m}{2} \right)! \left( \frac{l-m}{2} \right)! \right]^2}. \quad (2.204b)$$

A general disk potential, which is a sum over  $l$  and  $m$  of potentials of the form (2.202), is generated by the surface density  $\Sigma(v, \phi)$  that is the sum of surface densities  $\Sigma_{lm}(v, \phi)$ . According to equation (2.204a),  $-2V_{lm}/(\pi^2 G \Delta g_{lm})$  is the coefficient of  $Y_l^m(v, \phi)$  when  $|\cos v| \Sigma(v, \phi)$  is expanded in spherical harmonics. Thus with the orthogonality relation (C.44) we have

$$\frac{2V_{lm}}{\pi^2 G \Delta g_{lm}} = - \int_0^{2\pi} d\phi \int_0^\pi dv \sin v |\cos v| \Sigma(v, \phi) Y_l^{m*}(v, \phi). \quad (2.205)$$

The integrand in equation (2.205) is symmetrical about  $v = \pi/2$  when  $l-m$  is even, so we may restrict the  $v$  integration to the range  $(0, \pi/2)$  and double the result. Hence

$$V_{lm} = -\frac{\pi^2 G g_{lm}}{\Delta} \int_0^{2\pi} d\phi \int_0^\Delta dR R \Sigma(R, \phi) Y_l^{m*}(\sin^{-1}(R/\Delta), \phi). \quad (2.206)$$

All the techniques in this section are special cases of a general method for finding disk potential-density pairs that is described by Qian (1992).

## 2.7 The potential of our Galaxy

In this section we investigate the gravitational field of our own galaxy, the Milky Way. The Galaxy is made of several components, the disk, the bulge, the stellar halo, and the dark halo. The mix of stars, gas and dark matter that makes up a galaxy such as our own varies from component to component and is even likely to depend on location within each component.

Ideally, we should rely solely on dynamical tracers, such as the velocity fields of gas and stars and observations of gravitational lensing, to map out the distribution of mass in the Galaxy. Sadly, at the present time such a project is unfeasible.

Since we are not yet in a position to model the Galactic density and gravitational field in a purely dynamical way, we flesh out the available dynamical constraints with photometric information. In particular, we simply



*assume* that each component has a mass-to-light ratio  $\Upsilon$  that is independent of position. For the reason given above, this procedure is arbitrary and unsatisfactory, but it yields concrete Galactic potentials, which make testable predictions regarding the kinematics of stars and gas. Proceeding in this spirit, we now investigate models of our Galaxy, following Dehnen & Binney (1998a) and BM §10.6.

The brightness distribution of each component is assumed to be similar to those of external galaxies (BM §§4.3 and 4.4), while the size and total luminosity of each component is determined from photometry and star counts, or by fitting to the available dynamical constraints. We do not model the stellar halo here since its contribution to the potential is negligible.

The models are constrained by fitting to the following data (cf. Table 1.2):

- (i) The circular-speed curve  $v_c(R)$  for an assumed value of the solar circular speed,  $v_0 \equiv v_c(R_0)$ . Since this curve is determined from the line-of-sight velocities of tracers such as HI clouds and Cepheid stars, the circular-speed curve depends on  $v_0$ , which must be determined by other methods.
- (ii) The values of the Oort constants (Table 1.2, eq. 3.83, and BM §10.3.3).
- (iii) The total surface density within 1.1 kpc of the Galactic plane near the Sun,  $\Sigma_{1.1}(R_0)$ , and the contribution of the disk to this density (Table 1.1).
- (iv) The velocity dispersion of bulge stars in Baade's window, a line of sight that passes  $\sim 500$  pc from the Galactic center in which absorption by intervening dust is unusually low. We take this dispersion to be  $117 \pm 15 \text{ km s}^{-1}$ .
- (v) The total mass within 100 kpc of the Galactic center (eq. 1.12).
- (vi) The solar radius  $R_0 = 8 \text{ kpc}$ .

The functional forms assumed for each of the Galaxy's components are as follows.

**(a) The bulge** The density of this component is assumed to be

$$\rho_b(R, z) = \rho_{b0} \left( \frac{m}{a_b} \right)^{-\alpha_b} e^{-m^2/r_b^2}, \quad (2.207a)$$

where

$$m = \sqrt{R^2 + z^2/q_b^2}. \quad (2.207b)$$

For  $q_b < 1$  this is an oblate, spheroidal power-law model that is truncated at an outer radius  $r_b$ . Its potential is conveniently calculated from equations (2.125) with

$$e = \sqrt{1 - q_b^2} \quad ; \quad \psi(m) = \rho_{b0} \int_0^m dm^2 \left( \frac{m}{a_b} \right)^{-\alpha_b} e^{-m^2/r_b^2}. \quad (2.208)$$

Near-infrared photometry (BM §10.2.1) suggests values for three of the parameters,  $\alpha_b = 1.8$ ,  $q_b = 0.6$ ,  $r_b = 1.9 \text{ kpc}$ , and without loss of generality,

we can set  $a_b = 1$  kpc. The parameter  $\rho_{b0}$ , and hence the mass of the bulge, are determined by fitting the dynamical constraints.

**(b) The dark halo** By extending the spherical two-power-law models of §2.2.2f to oblate models, the density of this component is taken to have the form

$$\rho_h(R, z) = \rho_{h0} \left( \frac{m}{a_h} \right)^{-\alpha_h} \left( 1 + \frac{m}{a_h} \right)^{\alpha_h - \beta_h}, \quad (2.209)$$

where  $m$  is again given by equation (2.207b) with  $q_b$  replaced by  $q_h$ . The potential of this component, in which the density varies as  $r^{-\alpha_h}$  for  $r \ll a_h$  and  $r^{-\beta_h}$  at large  $r$ , can also be obtained from equation (2.125). Clearly, photometry provides no guidance as to the values of any of the parameters in equation (2.209); all five parameters  $\rho_{h0}$ ,  $a_h$ ,  $\alpha_h$ ,  $\beta_h$ , and  $q_h$  can only be determined by fitting the dynamical constraints. The data we use have little sensitivity to  $q_h$ , and we arbitrarily set it to 0.8.

**(c) The stellar disk** The density of the stellar disk is assumed to fall off exponentially with radius  $R$ , as in equation (1.7), and to depend on distance from the midplane  $z$  through the sum of two exponentials, representing the thin and thick disks described on page 13—this dependence on  $z$  is motivated by observations such as those of Gilmore & Reid (1983), shown in BM Figure 10.25. Mathematically,

$$\rho_d(R, z) = \Sigma_d e^{-R/R_d} \left( \frac{\alpha_0}{2z_0} e^{-|z|/z_0} + \frac{\alpha_1}{2z_1} e^{-|z|/z_1} \right), \quad (2.210)$$

where  $\alpha_0 + \alpha_1 = 1$ ,  $\Sigma_d$  is the central surface density,  $R_d$  is the disk scale length, and  $z_0 = 0.3$  kpc and  $z_1 = 1$  kpc are scale heights for the thin and thick components. The potential generated by this density distribution is given by equation (2.170) with  $\zeta(z)$  replaced by the expression in large brackets in equation (2.210).

**(d) The interstellar medium** The disk formed by a galaxy's interstellar medium (ISM) is thinner and more extended radially than the galaxy's stellar disk (see, for example, BM Figures 8.25 and 9.19). In the case of the Milky Way there is a hole of radius  $\sim 4$  kpc at the center of the disk of the ISM (BM Figure 9.19). These observations are crudely represented by taking the density of the ISM to be

$$\rho_g(R, z) = \frac{\Sigma_g}{2z_g} \exp \left( -\frac{R}{R_g} - \frac{R_m}{R} - \frac{|z|}{z_g} \right), \quad (2.211)$$

with  $R_m = 4$  kpc and  $z_g = 80$  pc. The parameters  $\Sigma_g$  and  $R_g$  are related to the parameters  $\Sigma_d$  and  $R_d$  of equation (2.210) by the assumption that  $R_g = 2R_d$  and that the ISM contributes 25% of the total disk surface density

**Table 2.3** Parameters of Galaxy models

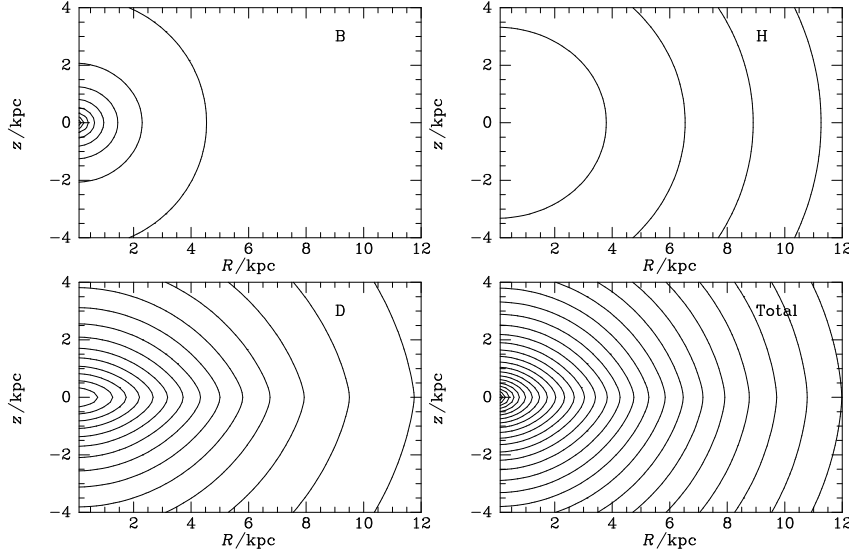
Parameter	Model I	Model II
$R_d/\text{kpc}$	2	3.2
$(\Sigma_d + \Sigma_g)/\mathcal{M}_\odot \text{pc}^{-2}$	1905	536
$\rho_{b0}/\mathcal{M}_\odot \text{pc}^{-3}$	0.427	0.3
$\rho_{h0}/\mathcal{M}_\odot \text{pc}^{-3}$	0.711	0.266
$\alpha_h$	-2	1.63
$\beta_h$	2.96	2.17
$a_h/\text{kpc}$	3.83	1.90
$M_d/10^{10} \mathcal{M}_\odot$	5.13	4.16
$M_b/10^{10} \mathcal{M}_\odot$	0.518	0.364
$M_{h,<10 \text{ kpc}}/10^{10} \mathcal{M}_\odot$	2.81	5.23
$M_{h,<100 \text{ kpc}}/10^{10} \mathcal{M}_\odot$	60.0	55.9
$v_e(R_0)/\text{km s}^{-1}$	520	494
$f_b$	0.05	0.04
$f_d$	0.60	0.33
$f_h$	0.35	0.63

NOTES: In both models  $0.75\Sigma(R_0)$  is contributed by stars, of which  $0.05\Sigma(R_0)$  is in the thick disk. Interstellar gas accounts for the remaining  $0.25\Sigma(R_0)$ . The thin and thick disks have the same scale length  $R_d$ , while the gas disk has scale length  $2R_d$  and a central hole of radius  $R_m = 4 \text{ kpc}$ . The thicknesses of the disks are  $z_0 = 300 \text{ pc}$ ,  $z_1 = 1 \text{ kpc}$ ,  $z_g = 80 \text{ pc}$ . In both models the bulge parameters are  $a_b = 1 \text{ kpc}$ ,  $\alpha_b = 1.8$ ,  $r_b = 1.9 \text{ kpc}$ ,  $q_b = 0.6$ , while the halo axis ratio  $q_h = 0.8$ . The quantity  $v_e(R_0)$  is the escape speed from the solar neighborhood;  $f_b$ ,  $f_d$  and  $f_h$  are the fractions of the radial force supplied by bulge, disk and halo at  $R_0 = 8 \text{ kpc}$ . These are slightly modified forms of Models 1 and 4 of Dehnen & Binney (1998a).

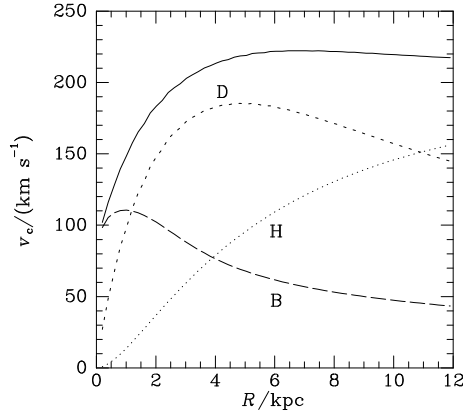
at the solar radius,  $R_0$ . The potential implied by equation (2.211) is best found from equation (2.154).

Dehnen & Binney (1998a) found that fits to the constraints described above could be obtained for a wide range of models made up of the components (a) to (d). The most important single parameter for determining the properties of a model is the scale length of the stellar disk,  $R_d$ . In §1.1.2 we estimated that  $R_d$  lies between 2 and 3 kpc. When  $R_d$  is at the lower end of this range, the disk dominates the gravitational field out to beyond the solar radius, whereas when  $R_d = 3 \text{ kpc}$ , the halo dominates at all radii. It is useful to examine the properties of two extreme models, namely the most and the least halo-dominated models; we designate them Models I and II and list their parameters in Table 2.3.

Model I has a small scale length,  $R_d = 2 \text{ kpc}$ , and gives rise to the isopotential surfaces and circular-speed curves shown in Figures 2.19 and 2.20. At small radii the halo density is  $\rho_h \propto r^{-\alpha_h} = r^2$ , which is the smallest value of  $\alpha_h$  allowed by the fitting program—with this disk scale length, the best fit has the smallest possible halo contribution near the center. Figure 2.20 illus-

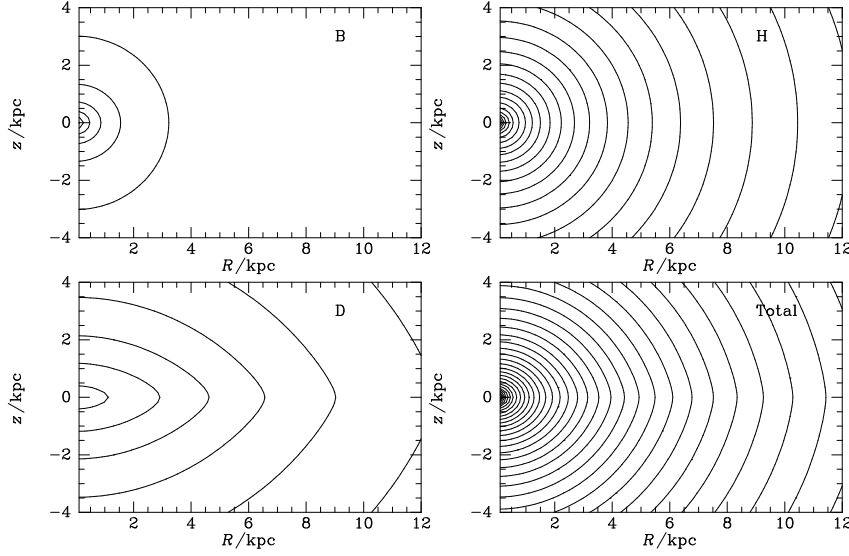


**Figure 2.19** The lower right panel shows equipotential contours of a model of the Galaxy with  $R_d = 2$  kpc (Model I). Contour levels are  $(-0.5, -1, -1.5 \dots) \times (100 \text{ km s}^{-1})^2$ . The top left panel shows the potential of the bulge, while the potentials of halo and disk are shown at top right and lower left, respectively. From top left to lower right the potentials at  $(R, z) = (8 \text{ kpc}, 0)$  are  $-0.28, -10.2, -2.98, -13.46 \times (100 \text{ km s}^{-1})^2$ .



**Figure 2.20** The full curve shows the circular-speed curve of Model I, whose potential is contoured in Figure 2.19. The contributions from the bulge, halo and disk are shown by the long-dashed, dotted and short-dashed curves, respectively. Notice that the total circular speed is given by the sum in quadrature of the circular speeds of the components.

trates the dynamical importance of the disk and bulge interior to the solar radius, showing that at such radii the halo makes only a small contribution to the overall circular speed—since  $v_c^2 \propto g$ , the contribution to the gravitational force is even smaller. This dominance is reflected in the contour plots of Figure 2.19 by the much closer packing of the equipotential contours of the bulge (top left panel) and disk (lower left panel) than those of the halo (top right panel). The equipotential surfaces of the disk are naturally more

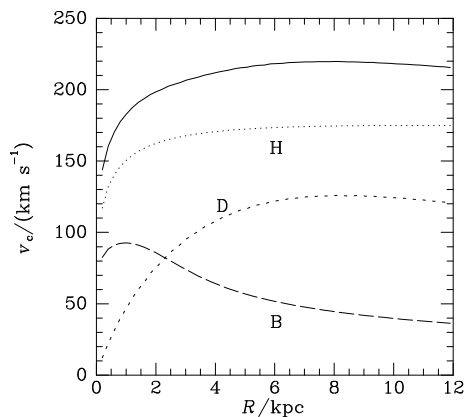


**Figure 2.21** Equipotential contours of the halo-dominated Galaxy model, Model II, which has  $R_d = 3.2$  kpc. The top left panel shows the potential of the bulge, while the potentials of halo and disk are shown at top right and lower left, respectively. The lower right panel shows the overall potential with contour levels  $(-0.5, -1, -1.5 \dots) \times (100 \text{ km s}^{-1})^2$ . From top left to lower right the potentials at  $(R, z) = (8 \text{ kpc}, 0)$  are  $-0.20, -9.83, -2.19, -12.21 \times (100 \text{ km s}^{-1})^2$ .

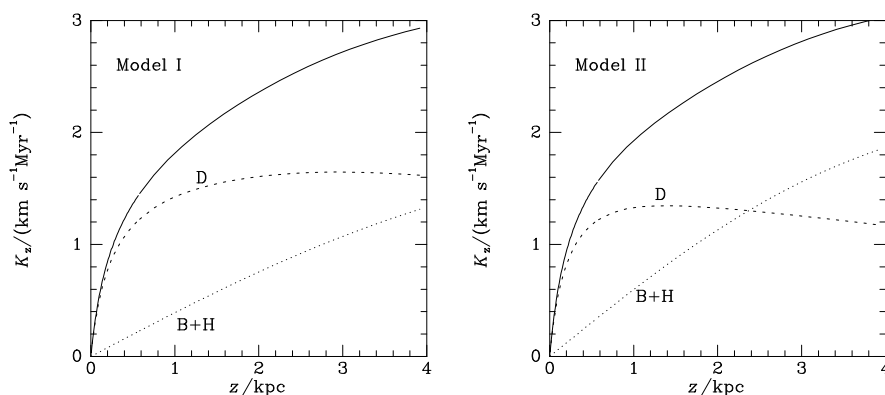
highly flattened than those of either the bulge or the halo, so the equipotential surfaces of the total potential are most flattened at radii  $r \sim 5$  kpc, where the disk's potential dominates.

Although in Model I the disk dominates the gravitational field (potential gradient) at  $R_0$ , the halo makes by far the largest contribution to the total potential at all radii. For example, at the Sun's location the halo contributes  $-10.2 \times (100 \text{ km s}^{-1})^2$  to the overall potential, while the disk and bulge together contribute only  $-3.26 \times (100 \text{ km s}^{-1})^2$ . The large contribution from the halo reflects the its enormous mass, most of it beyond  $R_0$ . Just how much mass the halo contains is ill-determined because the Galaxy's circular speed  $v_c(R)$  is uncertain beyond  $\approx 2R_0$ .

Figures 2.21 and 2.22 analyze the potential of Model II, a model that has a larger disk scale length,  $R_d = 3.2$  kpc. As Figure 2.22 shows, in this model the halo dominates the circular speed at all radii. It does so because it is much more centrally concentrated than the halo of Model I: at small  $r$  its density rises towards the center as  $r^{-1.63}$  rather than falling as in Model I. At the solar position the escape speed in this model is  $v_e(R_0) = 494 \text{ km s}^{-1}$ , which is observationally indistinguishable from  $v_e(R_0) = 520 \text{ km s}^{-1}$  in Model I; both are consistent with the observational estimate  $v_e(R_0) = (550 \pm 50) \text{ km s}^{-1}$  in Table 1.2.



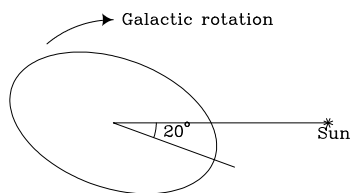
**Figure 2.22** The full curve shows the circular-speed curve of Model II. The contributions from the bulge, halo and disk are shown by the long-dashed, dotted and short-dashed curves, respectively.



**Figure 2.23** In each panel the full curve shows as a function of  $z$  in a Galactic model the force towards the galactic plane,  $K_z = \partial\Phi/\partial z$  at  $R_0 = 8$  kpc. The contributions from the bulge plus halo and disk are shown by the dotted and dashed curves, respectively. The left panel is for disk-dominated model, Model I (Figures 2.19 and 2.20), while the right panel is for Model II (Figures 2.21 and 2.22).

One of the striking conclusions from these models is that the relative contributions of the disk and the halo to the interior mass and the circular speed at  $R_0$  are very uncertain. As  $R_d$  varies from 3.2 kpc to 2 kpc, the mass of the dark halo inside 10 kpc decreases by nearly a factor 2 and the fraction of the gravitational force at  $R_0$  contributed by the halo falls<sup>9</sup> from 0.63 to 0.35. Similar uncertainties are encountered in models of external disk galaxies (van Albada et al. 1985; Sellwood 1999). This degeneracy between the disk and halo parameters has to be resolved by bringing other observational constraints or dynamical arguments to bear, such as those obtained from measurements of the dynamics of galactic bars (§6.5.2e), and

<sup>9</sup> Models in which most of the force at  $\sim 2R_d$  comes from the disk are called “maximum-disk models”—see §6.3.3.



**Figure 2.24** Schematic diagram of the Galactic bar.

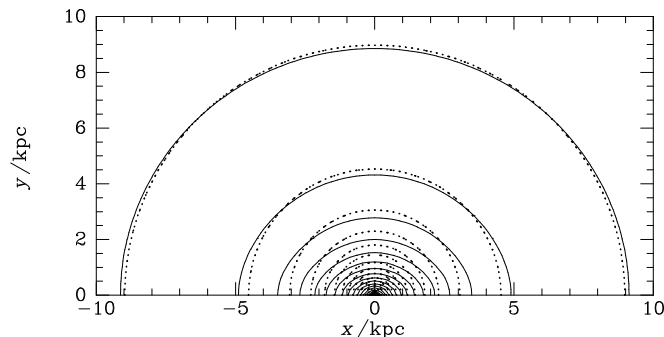
the optical depth to gravitational microlensing towards the Galactic center (Bissantz & Gerhard 2002; Famaey & Binney 2005).

Even though the halo of Model II dominates the circular speed at  $R_0$ , Figure 2.23 shows that the vertical force towards the disk,  $K_z = \partial\Phi/\partial z$ , is dominated by the disk within  $\sim 2$  kpc of the plane. Even a relatively low-mass disk can dominate  $K_z$  in this way because a disk's contribution to  $K_z$  rises extremely quickly near the plane, where the density of disk material is high. Above one scale height,  $\sim 200$  pc, the disk's contribution to  $K_z$  flattens off to the nearly constant value  $2\pi G\Sigma(R)$  (cf. Problem 2.3). By contrast, in both panels of Figure 2.23 the halo's contribution to  $K_z$  (dotted curves) rises nearly linearly with  $z$  out to several kiloparsecs above the Sun. Notice how similar the full curves in Figure 2.23 are: despite the very different contributions to  $v_c$  from disk and halo in the two models, the shape of the observationally measurable quantity  $K_z$  (§4.9.3) is almost the same in the two cases.

**(e) The bulge as a bar** There is both kinematic and photometric evidence that the Milky Way's bulge is in fact a bar, that is, a highly elongated, rapidly rotating stellar system (§6.5 and BM §§9.4 and 10.1). From the vantage point of the Sun, it is hard to determine the precise shape of the bar, but, as sketched in Figure 2.24, the bar is believed to extend to a Galactocentric radius  $\sim 3$  kpc, with its longest axis inclined by about  $20^\circ$  to the line from the Sun to the Galactic center (Bissantz & Gerhard 2002). The lengths of the bar's semi-axes are roughly in the ratios  $1:0.3:0.3$ .

Both photometric studies of the bar itself and comparisons with bars in other galaxies suggest that the isodensity surfaces deviate significantly from ellipsoids (López-Corredoira et al. 2005). Nonetheless, when considering the impact that the bar has on the Galaxy's gravitational potential, it is useful to start by approximating the isodensity surfaces by ellipsoids for in this case we can obtain the potential from equation (2.140)—if a more exact result were required, one could expand the difference  $\rho(\mathbf{x}) - \rho_e(\mathbf{x})$  between the actual density distribution  $\rho$  and the elliptical model  $\rho_e$  in spherical harmonics, and obtain the small correction to the potential from equation (2.95). In this spirit, we estimate the effect of the bar on the Galaxy's potential by fashioning a bar out of the axisymmetric bulge of Model I as follows.

In equations (2.207) we increase the scale radius  $r_b$  from 1.9 kpc to 3 kpc, and redefine  $m$  by  $m^2 = x^2 + (y^2 + z^2)/q^2$ , where  $x$  runs along the bar's long axis. We adopt  $q = 0.35$  and increase  $\rho_0$  so that the bar has roughly the same mass as the original bulge. In Figure 2.25 the full curves show the



**Figure 2.25** Full curves show the intersection with the Galactic plane of the isopotential surfaces of a model of the galactic bar. For comparison the dotted curves show the same curves for the axisymmetric bulge of Model I.

intersection with the Galactic plane of the bar's isopotential surfaces, while the dotted curves show the corresponding curves for the bulge of Model I. As expected, the bar's isopotential surfaces are elongated. The effect is very small near the solar circle but appreciable at  $R \lesssim 5$  kpc. On account of this elongation, the potential now generates tangential forces: Along a radius that makes an angle of  $45^\circ$  with the bar's long axis, the ratio  $F_\phi/F_r$  of the tangential and radial forces from the bar falls from 0.4 at the center to 0.27 at 2 kpc and 0.125 at 4 kpc. From Figure 2.20 we see that the bulge dominates  $F_r$  at  $R \lesssim 1$  kpc, so in this crude model tangential forces are very important at small radii. Conversely, at  $R \simeq 4$  kpc the bulge contributes only 11% of  $F_r$ , so at that radius  $F_\phi$  is only  $\sim 1\%$  of  $F_r$ . Nonetheless, the tangential forces that the bar induces can be dynamically significant for resonant orbits as far out as the solar circle because along such orbits the effect of  $F_\phi$  can accumulate over several periods (§3.7.2 and Dehnen 2000a).

We conclude that accurate models of the triaxial bar are needed to understand the dynamics of the Milky Way at  $R \lesssim 2\text{--}3$  kpc, and possibly beyond.

## 2.8 Potentials from functional expansions

A common theme of many of the methods we have described is the expansion of the gravitational potential and density in a set of functions that are potential-density pairs. We shall encounter such methods again in §2.9.4 as efficient tools for N-body simulation, and in §5.3.2, where we study linear response theory for stellar systems. In this section we re-examine these techniques from a general standpoint.

The basic idea of §2.3 was to approximate a real galactic density distribution by a density for which the potential is known analytically. Only a



small number of different functional forms for the density were presented, so with them the potential of a given galaxy could not be calculated to arbitrary accuracy. In §§2.6.2 to 2.6.4 we described a second approach: representing the system's density as an infinite sum of density distributions of known potential—for example the Bessel-function distributions  $\Sigma_{km}(R, \phi)$  that are defined by equations (2.179). This second approach has the advantage that the potential and density can be approximated to arbitrarily high accuracy. However, a good approximation will require that a large number of terms are taken in the sum for the potential, because none of the density distributions of individual terms resembles real galaxies.

Here we show how to combine the best aspects of these two approaches. In mathematical language, we find pairs of (possibly complex) **basis functions**  $\Phi_\beta(\mathbf{x})$  and  $\rho_\beta(\mathbf{x})$  for  $\beta = 1, 2, \dots$ , that satisfy

$$\nabla^2 \Phi_\beta = 4\pi G \rho_\beta, \quad (2.212)$$

and determine coefficients  $a_\beta$  such that the density of the system under study can be written as the sum

$$\rho(\mathbf{x}) = \sum_{\beta} a_{\beta} \rho_{\beta}(\mathbf{x}); \quad (2.213a)$$

then the system's potential  $\Phi(\mathbf{x})$  is given by

$$\Phi(\mathbf{x}) = \sum_{\beta} a_{\beta} \Phi_{\beta}(\mathbf{x}). \quad (2.213b)$$

We determine the coefficients  $a_\beta$  as follows. We multiply equation (2.213a) by  $-\Phi_\alpha^*$  and integrate over all space to obtain

$$s_\alpha = \sum_{\beta} M_{\alpha\beta} a_\beta, \quad (2.214a)$$

where

$$s_\alpha = - \int d^3\mathbf{x} \Phi_\alpha^*(\mathbf{x}) \rho(\mathbf{x}) \quad ; \quad M_{\alpha\beta} = - \int d^3\mathbf{x} \Phi_\alpha^*(\mathbf{x}) \rho_\beta(\mathbf{x}). \quad (2.214b)$$

The elements of the matrix  $\mathbf{M}$  have a simple physical interpretation:  $M_{\alpha\beta}$  is minus the potential energy of the density distribution  $\rho_\beta$  in the gravitational potential  $\Phi_\alpha^*$ . Using Poisson's equation (2.212) and applying the divergence theorem (B.45), we can show that  $\mathbf{M}$  is a Hermitian matrix:<sup>10</sup>

$$\begin{aligned} M_{\alpha\beta} &= - \frac{1}{4\pi G} \int d^3\mathbf{x} \Phi_\alpha^* \nabla^2 \Phi_\beta \\ &= - \frac{1}{4\pi G} \oint \Phi_\alpha^* \nabla \Phi_\beta \cdot d^2\mathbf{S} + \frac{1}{4\pi G} \int d^3\mathbf{x} \nabla \Phi_\alpha^* \cdot \nabla \Phi_\beta. \end{aligned} \quad (2.215)$$

---

<sup>10</sup> In the language of linear operators, the natural inner product on the space spanned by  $\{\Phi_\alpha\}$  is  $(f, g) = -(4\pi G)^{-1} \int d^3\mathbf{x} f^* \nabla^2 g$ . Then  $(f, g) = (g, f)^*$  is minus the interaction energy of the density distributions associated with the potentials  $f^*$  and  $g$ .

The surface term vanishes when the integral is taken over all space, because  $\Phi_\alpha$  falls off at least as fast as  $r^{-1}$ . This result shows that  $M_{\alpha\beta} = M_{\beta\alpha}^*$  so  $\mathbf{M}$  is Hermitian. Note that  $\mathbf{M}$  does not depend on the galactic mass distribution, so it can be computed once and for all after the basis potentials  $\Phi_\alpha$  have been chosen, and subsequently used time and again to follow the evolution of the potentials of many different galaxies, or of a single dynamically evolving galaxy.

The coefficients  $a_\beta$  can now be found by solving the linear equation (2.214a). We can choose the basis functions to facilitate this process. There are two strategies for exploiting this freedom.

**(a) Bi-orthonormal basis functions** We choose the basis functions such that  $\mathbf{M}$  is the unit matrix, so equation (2.214a) has the trivial solution  $a_\alpha = s_\alpha$ . From equation (2.214b) we see that this requires that the basis functions are **bi-orthonormal**:

$$-\int d^3\mathbf{x} \Phi_\alpha^*(\mathbf{x}) \rho_\beta(\mathbf{x}) = \delta_{\alpha\beta}. \quad (2.216)$$

A straightforward way of ensuring bi-orthonormality is to require that the  $\Phi_\alpha$  are eigenfunctions of the Hermitian operator  $\nabla^2$ —in this case  $\rho_\alpha \propto \Phi_\alpha$ , and the orthogonality of  $\Phi_\alpha$  and  $\rho_\beta$  is assured by the usual theorem that the eigenfunctions of a Hermitian operator are mutually orthogonal. Examples include the three functional expansions for disk potentials described in §§2.6.2 to 2.6.4. In §5.3.2 we shall use bi-orthonormal functional expansions to investigate the stability of stellar systems.

**(b) Designer basis functions** In principle  $\mathbf{a}$  and  $\mathbf{s}$  are vectors of infinite dimension. In practice it is necessary to truncate them and work with finite-dimensional vectors and matrices. The philosophy of **designer basis functions** is to choose the basis functions so that the galaxy can be accurately represented by the smallest possible number of them; the computational savings of having smaller vectors and matrices can more than offset the disadvantage that  $\mathbf{M}$  no longer has a simple form. The approach is most easily described by a concrete example. We examine the important special case in which each basis function is the product of a function of radius  $r$  and a spherical harmonic:

$$\Phi_{\mathbf{n}}(\mathbf{x}) = F_{\mathbf{n}}(r) Y_l^m(\boldsymbol{\Omega}), \quad (2.217)$$

where the potentials are enumerated by the vector  $\mathbf{n} = (n, l, m)$ , which has integer components, rather than the index  $\alpha$  used above; here  $n$  indexes an infinite number of radial functions for each spherical harmonic. Since the spherical harmonics (eq. C.44) are orthogonal,  $\mathbf{M}$  is now block diagonal, that is,  $M_{\mathbf{n}'\mathbf{n}} = 0$  for  $l' \neq l$  or  $m' \neq m$ . Thus the equations (2.214a) can be solved separately for each spherical-harmonic index pair  $(l, m)$ .

Applying the Laplacian operator (C.49) to equation (2.217), we find with Poisson's equation (2.212) that  $\rho_{\mathbf{n}}$  is given by

$$4\pi G \rho_{\mathbf{n}}(\mathbf{x}) = \nabla^2 \Phi_{\mathbf{n}} = Y_l^m(\boldsymbol{\Omega}) \left[ \frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{dF_{\mathbf{n}}}{dr} \right) - \frac{l(l+1)}{r^2} F_{\mathbf{n}} \right]. \quad (2.218)$$

For *any* choice of the radial function  $F_{\mathbf{n}}$ , equations (2.217) and (2.218) yield a potential-density pair. We now exploit this freedom to arrange for the first few functions  $\Phi_{\mathbf{n}}$  in the set to provide a good approximation to the potentials of real galaxies. One way of doing this is to set  $F_{000}(r)$  equal to the potential of one of the two-power-law models described in §2.2.2g. For guidance in choosing  $F_{0lm}$  for  $l > 0$  we turn to the solution of Poisson's equation that we obtained in terms of spherical harmonics in §2.4. Considering the coefficient of  $Y_l^m$  in equation (2.95), we see that the terms proportional to  $r^l$  will be dominant at small  $r$  and the terms proportional to  $r^{-(l+1)}$  will be dominant at large  $r$ . These observations imply that a promising choice for  $F_{0lm}$  is

$$F_{0lm}(r) = F_l(r) \equiv \frac{r^l}{(1 + r/r_0)^{2l+1}} \quad (l > 0), \quad (2.219)$$

where  $r_0$  is a suitable scale radius. If we also apply this equation for  $l = 0$ , then  $F_0(r)$  is proportional to the potential of a Hernquist model (eq. 2.67), consistent with our earlier argument that  $F_{000}$  should be the potential of a two-power-law galaxy model.

We must still expand our basis to include functions  $F_{\mathbf{n}}(r)$  for each given  $(l, m)$  and  $n > 0$  to describe accurately the radial dependence. To this end we write

$$F_{nl}(r) \equiv F_{\mathbf{n}}(r) = U_n(r)F_l(r), \quad (2.220)$$

where  $U_0(r) = 1$  and the set of  $U_n$  is complete. A possible choice would be

$$U_n(r) = u^n \quad \text{where} \quad u \equiv \frac{r/r_0 - 1}{r/r_0 + 1} \quad (-1 \leq u < 1). \quad (2.221)$$

When our choice of basis functions is inserted in equation (2.214b), we find with equation (2.218) that the matrix  $\mathbf{M}$  and the vector  $\mathbf{s}$  take the forms

$$\begin{aligned} M_{\mathbf{n}'\mathbf{n}} &= -(4\pi G)^{-1} \delta_{l'l} \delta_{m'm} \int dr F_{n'l}^*(r) \left[ \frac{d}{dr} \left( r^2 \frac{d}{dr} \right) - l(l+1) \right] F_{nl}(r) \\ s_{\mathbf{n}'} &= - \int d^3\mathbf{x} Y_l^{m*}(\boldsymbol{\Omega}) F_{n'l}^*(r) \rho(\mathbf{x}). \end{aligned} \quad (2.222)$$

It is possible to combine the best features of bi-orthonormal and designer basis functions. For example, if we choose  $F_{nl}(r) = U_n(r)F_l(r)$ , where  $F_l(r)$  is given by (2.219) and  $U_n(r)$  is an appropriate polynomial in the variable  $u$  defined by equation (2.221), the designer basis functions are bi-orthonormal (Hernquist & Ostriker 1992).

## 2.9 Poisson solvers for N-body codes

Our understanding of stellar dynamics has been profoundly advanced by **N-body codes**, computer programs that follow the motion of a large number of masses under their mutual gravitational attraction. Our discussion of these codes is in three parts: here we discuss the algorithms they use to find forces, in §3.4 we discuss the algorithms they use to move particles, and in §4.7.1 we discuss their general principles.

Realistic systems often contain many more particles than it is feasible to follow in a computer—for example, the Milky Way contains in excess of  $10^{11}$  stars and  $10^{69}$  GeV/( $mc^2$ ) dark-matter particles of mass  $m$ . Currently even the largest computers cannot work efficiently with more than  $\sim 10^{10}$  particles. Therefore there are two distinct types of N-body calculation with very different methodologies and problems. **Collisional N-body codes** simulate the evolution of a system with  $N_*$  stars by numerically integrating the equations of motion of exactly  $N = N_*$  particles. **Collisionless N-body codes** simulate the evolution of  $N_*$  stars by following the motion of  $N \ll N_*$  particles.

Collisional N-body codes are used to model systems in which relaxation is important (§1.2 and Chapter 7), in the sense that the relaxation time (1.38) is less than the duration of the numerical integration. They must accurately follow close encounters between stars, the formation and evolution of binary and triple stars, etc. These are challenging tasks because of the wide range of time and length scales involved. For example, a globular cluster may have binary stars with periods as short as hours, and yet evolve on a timescale of 10 Gyr.

Collisionless codes are easier to write but harder to understand and justify. They are used to model systems over times much shorter than the relaxation time. Since the number of particles  $N \ll N_*$ , the relaxation time in the model system is much smaller than that in the system being modeled; the philosophy is that the model is nevertheless accurate because the duration of the integration is much less than the relaxation time of either the real system or the model.

In essence, a collisionless code attempts to mimic the evolution of a system that contains *infinitely* many particles. Consequently, the system's density distribution is to be thought of as a continuum  $\rho(\mathbf{x}, t)$ , and the actual locations of the particles that model the system in the computer should be regarded as Monte-Carlo samplings of the probability-density distribution in position and velocity (§4.1).

N-body simulations—whether collisional or collisionless—use a simple principle: from the current positions of the particles, we derive the gravitational force on each particle. Then we use this force to advance the position and momentum of each particle for a short time, and find new forces. A major challenge is to produce code that will efficiently calculate the gravitational forces on a large number of bodies. We call such code a **Poisson solver**.

Poisson solvers for collisional problems can be designed to be essentially perfect: in §7.4.6d we shall see that the conceptually difficult part of a collisional simulation is integrating the orbits of particles. By contrast, Poisson solvers for collisionless simulations are fundamentally limited by particle noise: since we do not really know  $\rho(\mathbf{r})$  but only the locations of a finite number of particles that sample  $\rho(\mathbf{r})$ , we cannot recover  $\Phi(\mathbf{r})$  accurately no matter how much potential theory we know. On account of this fundamental limitation, Poisson solvers for collisionless simulations inevitably involve a compromise between inadequate resolution and excessive statistical noise. Some solvers are undoubtedly better than others, but the solver that is best suited to a given simulation depends on the scientific problem that is to be addressed. In this section we describe the most important types of Poisson solver, and discuss their strengths and weaknesses.

### 2.9.1 Direct summation

Consider evaluating the force on particle  $\alpha$  by simply summing the contributions from all the other particles in the simulation,

$$\mathbf{F}_\alpha = \sum_{\beta \neq \alpha} Gm_\beta \frac{\mathbf{r}_\beta - \mathbf{r}_\alpha}{|\mathbf{r}_\beta - \mathbf{r}_\alpha|^3}. \quad (2.223)$$

Each such force evaluation involves the calculation of  $N-1$  distances  $|\mathbf{r}_\beta - \mathbf{r}_\alpha|$ . Each distance can be used twice, once for the contribution of particle  $\beta$  to the force on particle  $\alpha$  and once for the force from particle  $\alpha$  on particle  $\beta$ . So, if forces are evaluated by direct summation a minimum of  $\frac{1}{2}N(N-1)$  distances have to be evaluated, where  $N \gg 1$  is the number of particles in the simulation. Thus, the work per timestep increases with  $N$  as  $N^2$ . Below we shall see that there are vastly more efficient ways of evaluating forces, which scale as  $N \ln N$ .

The first  $N$ -body calculations were restricted to small values of  $N$ , for which the difference between  $N^2$  and  $N \ln N$  is not large, so direct summation was a viable option. For the values of  $N > 10\,000$  now current,  $N^2/(N \ln N) > 10^3$ , so direct summation is in principle not an attractive strategy. Notwithstanding this fact it is still in use because its simple formulae can be encoded in hardware, so special-purpose processors can calculate forces between tens of thousands of particles at acceptable speeds (Makino et al. 2003).

**Softening** If two particles, say  $\alpha$  and  $\beta$ , approach each other closely, the term in equation (2.223) that describes the force  $\mathbf{F}_{\alpha\beta}$  between them becomes large. This phenomenon is problematic for both collisional and collisionless calculations, for different reasons.

In the collisional case the divergence of  $\mathbf{F}_{\alpha\beta}$  as  $\mathbf{r}_\alpha \rightarrow \mathbf{r}_\beta$  is a real physical effect, but one that is computationally awkward because it implies that the

equations of motion of particles  $\alpha$  and  $\beta$  have to be integrated with very small timesteps. Unless sophisticated workarounds such as “regularization” (§3.4.7) are employed to handle this situation, close encounters can bring the integration virtually to a halt. “Softening” of the force-law as described below is an expedient that keeps the integration moving along at an acceptable pace at the price of a loss of realism.

In the collisionless case the divergence of  $\mathbf{F}_{\alpha\beta}$  predicted by equation (2.223) is entirely unphysical, for the mass distribution we are trying to model is inherently smooth. The divergence is an artifact of the Monte-Carlo sampling of the density distribution. In this case softening can enhance rather than detract from the realism of the simulation.

Softening involves replacing equation (2.223) with

$$\mathbf{F}_{\alpha} = \sum_{\beta \neq \alpha} Gm_{\beta} S_{\mathbf{F}}(|\mathbf{r}_{\beta} - \mathbf{r}_{\alpha}|) \frac{\mathbf{r}_{\beta} - \mathbf{r}_{\alpha}}{|\mathbf{r}_{\beta} - \mathbf{r}_{\alpha}|}. \quad (2.224)$$

This equation differs from (2.223) only in the replacement of  $r^{-2}$  by  $S_{\mathbf{F}}(r)$ , where  $\mathbf{r} = \mathbf{r}_{\beta} - \mathbf{r}_{\alpha}$ .  $S_{\mathbf{F}}(r)$ , the **force softening kernel**, is some function that tends to  $r^{-2}$  for values of its argument bigger than the **softening length**  $\epsilon$ , and tends smoothly to zero for small values of its argument. Equation (2.224) has the desirable features that: (i) the force exerted by particle  $\beta$  on particle  $\alpha$ ,  $Gm_{\alpha}m_{\beta}S_{\mathbf{F}}(r)\mathbf{r}/r$ , is equal and opposite to the force exerted by  $\alpha$  on  $\beta$ , so Newton’s third law is satisfied; (ii) the force between any two particles acts along the line that joins them; (iii) the force approaches the usual gravitational force at large separations; (iv) the force between two particles at the same location is zero.

$S_{\mathbf{F}}(r)$  is the derivative of another function  $S(r)$ , the **softening kernel**, which appears in the equation for the potential at the location of particle  $\alpha$ :

$$\Phi_{\alpha} \equiv \Phi(\mathbf{r}_{\alpha}) = \sum_{\beta \neq \alpha} Gm_{\beta} S(|\mathbf{r}_{\beta} - \mathbf{r}_{\alpha}|). \quad (2.225)$$

A widely used form of  $S$  is

$$S(\mathbf{r}) = -\frac{1}{\sqrt{r^2 + \epsilon^2}}. \quad (2.226)$$

In this case, the gravitational potential of each particle is just that of a Plummer sphere of scale length  $\epsilon$  (eq. 2.44a). In a collisionless simulation this makes physical sense: the mass that is represented by a particle is in reality distributed through a non-zero volume.<sup>11</sup> The density law of a Plummer sphere (eq. 2.44b) places most of the sphere’s mass inside  $r = \epsilon$ , so the natural choice is to make  $\epsilon$  of order the inter-particle separation.

<sup>11</sup> Note that in this model the softened force between two particles is the force between a point mass and a Plummer sphere, not the force between two Plummer spheres.

In collisionless simulations the choice of softening kernel and softening length is a compromise between maximizing the smoothness of the force-field and minimizing the degradation of the spatial resolution caused by the softening. Despite its popularity, the choice (2.226) of  $S$  is not optimal in this sense, because the density of a Plummer sphere falls off too slowly with radius. A better choice is

$$S(\mathbf{r}) = -\frac{r^2 + \frac{3}{2}\epsilon^2}{(r^2 + \epsilon^2)^{3/2}}, \quad (2.227)$$

which amounts to replacing the potential of each particle with that of a sphere of radius  $\epsilon$  in which the density is proportional to  $(r^2 + \epsilon^2)^{-7/2}$ . See Dehnen (2001) for a discussion of the merits of various softening kernels.

An ideal Poisson solver would allow  $\epsilon$  to vary as a particle moves from a high-density to a low-density region. In practice most contemporary Poisson solvers use fixed softening due to technical difficulties associated with energy conservation.

We now consider algorithms that enable one to evaluate the sum (2.225) to high accuracy in a number of arithmetic operations that scales as  $N \ln N$ , rather as  $N^2$ .

### 2.9.2 Tree codes

Our discussion is based on Dehnen (2000b). The particles are first organized into a structure that was introduced by Barnes & Hut (1986): we place an imaginary cube around the simulation and divide it into eight equal sub-cubes. If any sub-cube contains more than one particle, we divide it in turn into eight equal cubes, and so on until every cube contains at most one particle. This hierarchy of cubes forms an **oct-tree**, and the original cube is called the **root** of the tree. Each cube after the root has a parent cube, and seven sibling cubes. Any cube that contains more than one particle has eight child cubes. There is a clear analogy with a real tree, whose trunk divides into great boughs, which divide into branches, which divide into twigs, which ultimately carry leaves. Particles are the leaves of our oct-tree.

Suppose now that we locate the center of mass of the particles in each cube and evaluate the sums

$$M_0 \equiv \sum_{\alpha} m_{\alpha} \quad ; \quad M_{ij} = \sum_{\alpha} m_{\alpha} x_i^{\alpha} x_j^{\alpha} \quad ; \quad M_{ijk} = \sum_{\alpha} m_{\alpha} x_i^{\alpha} x_j^{\alpha} x_k^{\alpha} \quad (2.228)$$

and so forth, where the sums run over the particles in the cube and  $\mathbf{x}^{\alpha}$  is the position vector of particle  $\alpha$  relative to the cube's center of mass. This hierarchy of sums is closely related to the multipole expansion of §2.4, and these quantities are called the **Cartesian multipole moments**. The three in equation (2.228) are the monopole, quadrupole, and octopole moments.

### Box 2.2: Scaling of tree codes

Consider the work involved in determining the potential at the location of one of  $N$  simulation particles. For simplicity we assume that the system is not very inhomogeneous, although the argument we give can be generalized. The sum over cubes for the potential includes some cubes (“leaves”) that contain only one particle, the rest being “branches” that contain more than one particle. All branch cubes have opening angle  $\leq \theta_o$ , while most of the leaf cubes have opening angle  $> \theta_o$ . If the simulation is representing a system of linear size  $L$ , the leaf cubes have a characteristic linear size  $l = L/N^{1/3}$  and lie within distance  $\sim D = l/\theta_o$  of the given particle; hence there are  $\sim (D/l)^3 = \theta_o^{-3}$  leaf cubes.

Suppose we now increase the number of simulation particles by a factor 8. Then branch cubes will continue to be branch cubes and will not require further subdivision because the angles they subtend are  $< \theta_o$ . Hence the extra work that comes with increased  $N$  arises mainly from the  $\theta_o^{-3}$  leaf cubes that previously had only one particle; most will now need to be subdivided. Thus every time we increase  $N$  by a factor 8 we have to do a fixed amount of additional work. It now follows that the work required to determine the potential at the location of a given particle is proportional to  $\ln N$ , and the work involved in determining the forces on all particles scales as  $N \ln N$ .

The object  $M_i = \sum_{\alpha} m_{\alpha} x_i^{\alpha}$  that might have been listed in equation (2.228) is the cube’s dipole moment, and it vanishes identically by virtue of the fact that  $\mathbf{x}^{\alpha}$  is the position of  $\alpha$  relative to the center of mass. If the cube is one of the tree’s leaves,  $M_0 = m_{\alpha}$  and all the other multipoles vanish. There are of order  $N \log_2 N$  cubes in the tree, so the labor involved in setting it up scales as  $N \ln N$ .

Once the tree has been constructed, we can evaluate the sum (2.225) that yields the potential at any location  $\mathbf{x}$  to good accuracy with only  $\sim \ln N$  operations as follows. We first consider evaluating the force as the sum of the forces due to the monopole, quadrupole, octopole,  $\dots$ ,  $2^k$ -pole etc., of the root cube, up to some predetermined order  $k = K$ . This sum would converge rapidly and could be truncated after a few terms without unacceptable inaccuracy if  $\mathbf{x}$  were remote from the root cube in the sense that the root cube subtended a small **opening angle**  $\theta_o$ , say  $\theta_o \lesssim 1^\circ$ , at  $\mathbf{x}$ . In general,  $\mathbf{x}$  will lie within the root cube, so the desired condition will be not be satisfied. However, we may then instead consider evaluating the potential by summing the multipole contributions from each of the root’s eight child cubes. If any of these subtends an angle at  $\mathbf{x}$  that exceeds  $\theta_o$ , we consider the sum of the multipoles of its children, and so on recursively until the whole potential is obtained as a sum of the multipoles from cubes that



subtend angles at  $\mathbf{x}$  smaller than  $\theta_o$ . Some of these final cubes will be leaves, which are deemed to subtend zero angle, because they contain only a single particle. The key point is that as the number of particles in the simulation is increased, the number of cubes over which we sum to get the force at  $\mathbf{x}$  increases much more slowly—see Box 2.2.

**Cartesian multipole expansion** To derive explicit formulae for the potential in terms of the multipoles of cubes, we now develop some of the theory of Cartesian multipoles. We consider the contribution to the gravitational potential  $\Phi$  at position  $\mathbf{X}$  in cube A that is generated by the matter in cube B. We have

$$\Phi(\mathbf{X}) = G \int_{\text{cube B}} d^3\mathbf{Y} S(\mathbf{X} - \mathbf{Y}) \rho(\mathbf{Y}), \quad (2.229)$$

where  $\rho$  is the matter density and  $S$  is the softening kernel (eq. 2.225). We specialize to the case of well-separated cubes and write  $\mathbf{X} = \hat{\mathbf{X}} + \mathbf{x}$  and  $\mathbf{Y} = \hat{\mathbf{Y}} + \mathbf{y}$ , where  $\hat{\mathbf{X}}$  and  $\hat{\mathbf{Y}}$  are the centers of mass of cubes A and B, respectively. Then we Taylor expand  $S(\mathbf{X} - \mathbf{Y})$  in powers of  $\mathbf{x} - \mathbf{y}$  around  $\hat{\mathbf{X}} - \hat{\mathbf{Y}}$ :

$$\begin{aligned} S(\mathbf{X} - \mathbf{Y}) &= S(\hat{\mathbf{X}} - \hat{\mathbf{Y}}) + \sum_{i=1}^3 (\mathbf{x} - \mathbf{y})_i \frac{\partial S(\mathbf{R})}{\partial R_i} \Big|_{\mathbf{R}=\hat{\mathbf{X}}-\hat{\mathbf{Y}}} \\ &\quad + \frac{1}{2!} \sum_{i,j=1}^3 (\mathbf{x} - \mathbf{y})_i (\mathbf{x} - \mathbf{y})_j \frac{\partial^2 S(\mathbf{R})}{\partial R_i \partial R_j} \Big|_{\mathbf{R}=\hat{\mathbf{X}}-\hat{\mathbf{Y}}} + \dots \quad (2.230) \\ &= \sum_n \frac{1}{n!} (\mathbf{x} - \mathbf{y})^{(n)} \cdot [\partial_{(n)} S(\mathbf{R})]_{\mathbf{R}=\hat{\mathbf{X}}-\hat{\mathbf{Y}}}. \end{aligned}$$

Here the last line employs a notation in which  $\mathbf{r}^{(n)} \equiv r_{i_1} r_{i_2} \dots r_{i_n}$  is a product of  $n$  components of the vector  $\mathbf{r}$  and  $\partial_{(n)}$  is the corresponding product of partial derivatives with respect to these components, while the centered dot implies summation over all the  $n$  indices  $i_k$  (a total of  $3^n$  terms). Substituting (2.230) into (2.229), we obtain

$$\Phi(\hat{\mathbf{X}} + \mathbf{x}) = G \sum_n \frac{1}{n!} \left[ \int d^3\mathbf{y} \rho(\hat{\mathbf{Y}} + \mathbf{y}) (\mathbf{x} - \mathbf{y})^{(n)} \right] \cdot [\partial_{(n)} S(\mathbf{R})]_{\mathbf{R}=\hat{\mathbf{X}}-\hat{\mathbf{Y}}}. \quad (2.231)$$

Using the expansion

$$(\mathbf{x} - \mathbf{y})^{(n)} = \sum_{k=0}^n \frac{n!}{k!(n-k)!} \mathbf{x}^{(k)} \mathbf{y}^{(n-k)}, \quad (2.232)$$

we have

$$\Phi(\hat{\mathbf{X}} + \mathbf{x}) = G \sum_n \sum_{k=0}^n \frac{1}{k!(n-k)!} \mathbf{x}^{(k)} \mathbf{M}^{(n-k)} \cdot [\partial_{(n)} S(\mathbf{R})]_{\mathbf{R}=\hat{\mathbf{X}}-\hat{\mathbf{Y}}}, \quad (2.233)$$

### Box 2.3: Derivatives of a spherically symmetric softening kernel

Consider the case of a spherically symmetric softening kernel  $S(\mathbf{R})$ , i.e., one of the form  $S(|\mathbf{R}|)$ . In this case the evaluation of the first few tensors  $\partial_{(n)}S$  is straightforward:

$$\begin{aligned}
 \partial_{(1)}S(R) &= S' \frac{R_i}{R} \\
 \partial_{(2)}S(R) &= \left( S'' - \frac{S'}{R} \right) \frac{R_i R_j}{R^2} + \frac{S'}{R} \delta_{ij} \\
 \partial_{(3)}S(R) &= \left( S''' - 3 \frac{S''}{R} + 3 \frac{S'}{R^2} \right) \frac{R_i R_j R_k}{R^3} \\
 &\quad + \left( S'' - \frac{S'}{R} \right) \frac{\delta_{ij} R_k + \delta_{jk} R_i + \delta_{ki} R_j}{R^2} \\
 \partial_{(4)}S(R) &= \left( S^{iv} - 6 \frac{S'''}{R} + 9 \frac{S''}{R^2} - 9 \frac{S'}{R^3} \right) \frac{R_i R_j R_k R_l}{R^4} \\
 &\quad + \left( S''' - 3 \frac{S''}{R} + 3 \frac{S'}{R^2} \right) (\delta_{ij} R_k R_l + \delta_{jk} R_i R_l + \delta_{ki} R_j R_l \\
 &\quad + \delta_{il} R_j R_k + \delta_{lj} R_i R_k + \delta_{lk} R_i R_j) / R^3 \\
 &\quad + \left( S'' - \frac{S'}{R} \right) \frac{\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}}{R^2}
 \end{aligned}$$

where

$$\mathbf{M}^{(k)} \equiv \int_{\text{cube B}} d^3 \mathbf{y} \rho(\hat{\mathbf{Y}} + \mathbf{y}) \mathbf{y}^{(k)}. \quad (2.234)$$

When the density distribution is simulated with particles,  $\rho(\mathbf{r}) = \sum_{\alpha} m_{\alpha} \delta(\mathbf{r} - \mathbf{r}^{\alpha})$  is a sum of delta functions centered on the locations of the particles. In this case the integral in (2.234) can be immediately evaluated and yields the objects defined by equations (2.228).

Equation (2.233) expresses the potential generated by cube B at an arbitrary point in cube A as a power series in the components of  $\mathbf{x}$ , the separation between the point of observation and A's center of mass. The coefficients in this power series are made up of (a) the multipole moments of B, which were evaluated during tree construction, and (b) the derivatives  $\partial_{(n)}S$  of the softening kernel evaluated on the vector  $\mathbf{R}$  that joins the centers of mass of the two cubes, which can be evaluated once and for all at the start of the integration. Once these quantities are all to hand, we can quickly evaluate the force on every particle in cube A by summing the series with the appropriate value of  $\mathbf{x}$ . The derivatives  $\partial_{(n)}S$  are explicitly evaluated for a spherically symmetric kernel in Box 2.3.

An important simplification arises because the same derivatives occur in the analogous series for the potential that cube A generates at a point within cube B. Dehnen (2000b) explains how one can exploit this fact, both to accelerate the computation and to achieve exact momentum conservation.

Unlike some of the codes that we are about to present, a tree code employs no grid. This feature enables tree codes to handle problems such as galaxy mergers, in which dense stellar systems move through a large volume of nearly empty space—most grid-based codes cannot handle such problems efficiently because they would waste most of their computing resources covering low-density regions with the same high-resolution grid that they require to handle the dense stellar systems. Another merit of tree codes is that they are readily parallelized by, for example, delegating to a different processor construction of each of the trees that emerge from the root's eight child cubes.

### 2.9.3 Particle-mesh codes

A wide variety of *N*-body codes solve Poisson's equation using estimates of the density at a set of regularly spaced points, the nodes of a mesh or grid. The simplest mesh is a Cartesian grid and we concentrate on this case, although most of the principles we describe carry over to other meshes, such as ones based on spherical polar coordinates. More detail can be found in Hockney & Eastwood (1988).

The process of estimating the density on a grid from the positions of a large number of particles that trace the density is called **mass assignment** because it involves assigning the mass of each particle to one or more nearby nodes. The simplest algorithm is to assign all of the mass of each particle to the nearest node. This procedure, which is known as the **nearest grid point (NGP)** scheme, is rarely used for two reasons. First, the NGP scheme only samples the density distribution crudely. Second, as explained on page 134 below, for technical reasons associated with momentum conservation the NGP scheme leads to solutions of Poisson's equation in which forces change discontinuously midway between nodes.

Better mass-assignment schemes spread the mass of each particle over several nearby nodes. Box 2.4 describes a hierarchy of widely used schemes that spread the mass over  $1, 2^D, 3^D, \dots$  nodes of a  $D$ -dimensional grid. If we decide on one node, we have the NGP scheme. If we want to split a particle's mass over  $2^D$  nodes, the resulting scheme is called the **cloud in cell (CIC)** mass-assignment scheme because it may be visualized by expanding every mass point into a homogeneous cube with side length equal to the grid spacing, and assigning every mass element in this cube to the nearest node (Figure 2.26). If we split a particle's mass over  $3^D$  nodes, we have the **triangular shaped cloud (TSC)** mass-assignment scheme, so called because the scheme may be derived by regarding the particle to be a cloud

### Box 2.4: Mass assignment schemes

A mass-assignment scheme is a function  $W_{\mathbf{p}}(\mathbf{r})$  that gives the fraction of the mass of a particle at  $\mathbf{r}$  that is assigned to node  $\mathbf{p}$ . Here  $\mathbf{p}$  is a triple of integers and the node's location is  $\mathbf{x}_{\mathbf{p}} = (p_1, p_2, p_3)\Delta$ , with  $\Delta$  the grid spacing. The simplest schemes for a three-dimensional Cartesian grid are the product of three functions, each evaluated on one component of  $\mathbf{r}$

$$W_{\mathbf{p}}(\mathbf{r}) = w(x - p_1\Delta)w(y - p_2\Delta)w(z - p_3\Delta).$$

The condition that all the mass is assigned somewhere is  $\sum_{\mathbf{p}} W_{\mathbf{p}}(\mathbf{r}) = 1$ , which is assured if  $M_0 \equiv \sum_i w(x - i\Delta) = 1$ . Additional constraints arise from demanding that the grid-scale fluctuations are as small as possible at large distances. Suppose that we are modeling a one-dimensional potential with softening kernel  $S(x)$ . Then the potential due to a mass  $m$  at  $x$  is  $\Phi(x') = mS(x' - x)$ . The potential computed from a grid is

$$\Phi_g(x') = m \sum_i w(x - i\Delta)S(x' - i\Delta).$$

If  $|x' - x| \gg \Delta$ , we can Taylor expand  $S(x' - i\Delta)$  in powers of  $x - i\Delta$ :

$$\Phi_g(x') = m \sum_i \sum_{n=0}^{\infty} \frac{1}{n!} w(x - i\Delta) S^{(n)}(x' - x) (x - i\Delta)^n,$$

where  $S^{(n)}$  is the  $n$ th derivative of  $S$ . If condition  $M_0 = 1$  is satisfied, the  $n = 0$  term equals the exact potential, so grid-scale fluctuations are minimized if

$$M_n \equiv \sum_i (x - i\Delta)^n w(x - i\Delta) = \text{constant} \quad \text{for } n = 1, 2, \dots$$

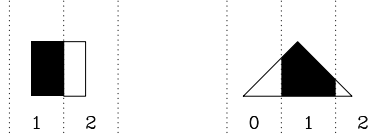
If the mass is distributed between two nodes,

$$w(x) = \begin{cases} 1 - |x|/\Delta & \text{for } |x| < \Delta, \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

satisfies  $M_0 = 1$ ,  $M_1 = 0$ . This is the cloud in cell (CIC) mass-assignment scheme. If the mass is distributed between three nodes,

$$w(x) = \begin{cases} \frac{3}{4} - |x|^2/\Delta^2 & \text{for } |x| < \frac{1}{2}\Delta \\ \frac{1}{2}(\frac{3}{2} - |x|/\Delta)^2 & \text{for } \frac{1}{2}\Delta < |x| < \frac{3}{2}\Delta \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

satisfies  $M_0 = 1$ ,  $M_1 = 0$ ,  $M_2 = \frac{1}{4}\Delta^2$ . This is the triangular shaped cloud (TSC) mass-assignment scheme. There is an alternative three-node scheme in which  $M_2 = 0$ , but in this scheme  $w(x)$  is discontinuous at  $|x|/\Delta = \frac{1}{2}, \frac{3}{2}$ , whereas in the TSC scheme  $w(x)$  is continuous and differentiable.



**Figure 2.26** The CIC (left) and TSC (right) mass-assignment schemes in one dimension. The dotted lines mark the boundaries between cells of the grid, and the rectangle (CIC) and triangle (TSC) represent the mass of the particle. Mass is assigned to cells in proportion to the fraction of the area of the rectangle or triangle that lies in each cell—for example, the shaded portions are assigned to cell 1.

in which the mass density falls linearly with coordinate differences from the location of the particle (Figure 2.26).

In the NGP scheme, the mass assigned to a node changes discontinuously as a particle moves through the grid. In the CIC scheme, the mass changes continuously but with discontinuous first derivative. In the TSC scheme, both the mass and its derivative change continuously. In view of the role a mass-assignment scheme plays in force interpolation—see page 134 below—the progressive increase in smoothness as one proceeds up this hierarchy of mass-assignment schemes is valuable. This advantage must be set against two disadvantages of proceeding up the hierarchy: (i) in three dimensions the computational cost increases rapidly as the number of nodes to which mass is assigned increases from 1 for NGP, to 8 for CIC, 27 for TSC, etc.; (ii) spreading mass over many nodes involves a loss of spatial resolution. In practice the most widely used schemes are CIC and TSC.

Given the density on the mesh, we are ready to solve Poisson’s equation for the potential. To do this we must specify boundary conditions, which in practice are one of two types: **periodic** and **vacuum**. Periodic boundary conditions are used when the grid is imagined to be one cell of an infinite lattice; these are natural for cosmological simulations. With vacuum boundary conditions we require that at large distances  $\Phi \rightarrow -GM/r$ , where  $M$  is the total mass on the grid. These are appropriate for simulations of isolated stellar systems.

**(a) Periodic boundary conditions** We solve Poisson’s equation by representing  $\nabla^2$  as a finite-difference operator. For example, we can make the approximation

$$(\nabla^2 \Phi)_{lmn} \simeq (\Phi_{l+1,m,n} + \Phi_{l-1,m,n} + \Phi_{l,m+1,n} + \Phi_{l,m-1,n} + \Phi_{l,m,n+1} + \Phi_{l,m,n-1} - 6\Phi_{l,m,n})/\Delta^2, \quad (2.235)$$

where  $\Delta$  is the mesh spacing. This substitution converts the Poisson equation into a large system of linear equations for the  $\Phi_{lmn}$  in terms of the known  $\rho_{lmn}$ . The system obtained by equating (2.235) to  $4\pi G\rho_{lmn}$  is most

efficiently solved using discrete Fourier transforms (DFTs; see Appendix G and Problem 2.23).

**(b) Vacuum boundary conditions** Consider now how to solve Poisson's equation with vacuum boundary conditions. We use the integral form of Poisson's equation (2.3), or its smoothed form (2.229), in which the softening kernel  $S$  already contains our choice of boundary conditions. We write

$$\Phi(\mathbf{r}) = G \sum_{\mathbf{p}} S(\mathbf{r} - \mathbf{r}_{\mathbf{p}}) m_{\mathbf{p}}, \quad (2.236)$$

where  $m_{\mathbf{p}}$  is the mass assigned to node  $\mathbf{p}$ . The labor required to evaluate  $\Phi$  at the location of every particle is proportional to the number of particles times the number of nodes. A much more efficient approach is to evaluate  $\Phi$  only at the nodes, and then to interpolate to the locations of particles. At the nodes, equation (2.236) reads

$$\Phi_{\mathbf{q}} = G \sum_{\mathbf{p}} S_{\mathbf{q}-\mathbf{p}} m_{\mathbf{p}}, \quad (2.237)$$

where  $\mathbf{p}$  and  $\mathbf{q}$  are three-vectors with integer components in the range  $(0, K-1)$ , with  $K$  the number of nodes along a side of the computational box. In equation (2.237), each component of  $S$ 's subscript ranges from  $-(K-1)$  (when  $q_1 = 0$  and  $p_1 = K-1$ ) to  $K-1$  (when  $q_1 = K-1$  and  $p_1 = 0$ ). Thus  $S$ 's subscript has components that take on  $2K-1$  values, while the subscript of  $m$  has components that range over only  $K$  possible values. Discrete Fourier transforms provide an extremely efficient way of solving equations of this form, but before we can apply them, we have to make the problem symmetrical in the subscripts, and arrange for them to take  $2^n$  values, for some integer  $n$ . We do this by taking  $K = 2^{n-1}$  and padding  $m$  out to a matrix of size  $(2K)^3$ , whose subscripts cover the range  $(-K, K-1)$  with  $m_{\mathbf{p}}$  set to zero when any component of the subscript lies outside the physically significant range,  $(0, K-1)$ . Finally we eliminate negative indices by making  $S$  and  $m$  periodic functions of each index with period  $2K$ . Thus extended, equation (2.237) takes the form of a discrete convolution. So we take the DFT of both sides of this equation, and, with the aid of the discrete Fourier convolution theorem (eq. G.5), we obtain

$$\hat{\Phi}_{\mathbf{k}} = G \hat{S}_{\mathbf{k}} \hat{m}_{\mathbf{k}}. \quad (2.238)$$

According to this equation, the DFT of  $\Phi$  can be obtained from that of  $m$  simply by multiplying by the DFT of  $S$ . With  $\hat{\Phi}_{\mathbf{k}}$  in hand,  $\Phi$  is easily obtained by doing an inverse DFT. Moreover, the DFT of  $S$  can be calculated once and for all at the start of the simulation and stored for use at each timestep.

### Box 2.5: James's Fourier Poisson solver

We describe the method of James (1977) for imposing vacuum boundary conditions on the system (2.235) without doubling the range of all the indices. This method is both fast and economical with computer memory.

We start from the density values on a cubical grid and imagine the values to represent electric charge density rather than mass density. Next, we find the potential that would be generated by the charges attached to the interior of the grid if the boundary of the grid were grounded—that is, if the potential  $\Phi$  vanished on the boundary. This potential is the solution of Poisson's equation for the given charge density at interior points of the grid, subject to the boundary condition  $\Phi = 0$  on the grid boundary. We obtain the required solution by writing  $\Phi$  in terms of its sine transform (eq. G.7a)

$$\Phi_{lmn} = \sum_{\alpha\beta\gamma=1}^{K-1} \Phi^{\alpha\beta\gamma}(\text{SSS}) \sin\left(\frac{\pi l\alpha}{K}\right) \sin\left(\frac{\pi m\beta}{K}\right) \sin\left(\frac{\pi n\gamma}{K}\right), \quad (1)$$

where  $0 \leq l, m, n \leq K$ . The sine transform automatically ensures that  $\Phi_{lmn}$  vanishes on the boundary. Substituting this expression for  $\Phi_{lmn}$  into the numerical approximation (2.235) for the Laplacian, and then equating this to  $4\pi G\rho_{lmn}$  with  $\rho_{lmn}$  expressed in terms of its sine transform, we discover that each amplitude  $\Phi^{\alpha\beta\gamma}(\text{SSS})$  is simply a multiple of the corresponding amplitude  $\rho^{\alpha\beta\gamma}(\text{SSS})$ . Since the sine transform is its own inverse (eq. G.7b), it is a simple matter to recover the  $\Phi_{lmn}$  from the  $\Phi^{\alpha\beta\gamma}(\text{SSS})$ .

The potential we obtain in this way differs from the one we require because the walls bear charges that differ from those specified in the original problem. (When the walls of the grid are grounded, charges flow along the grounding cables until  $\Phi$  vanishes on the boundary.) We use Gauss's theorem to determine the actual surface density of wall charges  $\Sigma$ —since the potential vanishes outside the box,  $\Sigma = -g_{\perp}/(4\pi G)$ , where  $g_{\perp} = |\mathbf{n} \cdot \nabla\Phi|$  is the magnitude of the force field just inside the walls of the grounded box. Finally, we convolve the resulting charge distribution on the walls with a softening kernel to obtain the potential that it generates when vacuum boundary conditions are applied, and subtract this potential from the potential generated by the mass in the grounded box, to get the potential that the originally specified density distribution generates with vacuum boundary conditions. The softening kernel should be the inverse of the discrete Laplacian used in the calculation of  $\Phi^{\alpha\beta\gamma}(\text{SSS})$ —see James (1977) for details.

The key to the success of this procedure is that the convolution of the wall charge density with the softening kernel is simple because the charge density is hollow—see Appendix G or Magorrian (2007) for details.

This approach to imposing vacuum boundary conditions is very wasteful of computer memory. Box 2.5 describes a clever and much more economical alternative.

Once the potential has been obtained at every node, we use numerical differentiation to obtain the gravitational field  $\mathbf{g}_\mathbf{n}$  at the nodes: for example, the component parallel to the first axis might be

$$\hat{\mathbf{e}}_1 \cdot \mathbf{g}_\mathbf{n} = -\frac{1}{2\Delta}(\Phi_{(n_1+1, n_2, n_3)} - \Phi_{(n_1-1, n_2, n_3)}). \quad (2.239)$$

More generally

$$\mathbf{g}_\mathbf{j} = \sum_{\mathbf{n}} \mathbf{A}_\mathbf{n} \Phi_{\mathbf{j}+\mathbf{n}}, \quad (2.240)$$

where  $\mathbf{A}_\mathbf{n} = -\mathbf{A}_{-\mathbf{n}}$  defines some numerical differentiation scheme. Values of the forces on individual particles are then obtained by interpolation from these values.

In any such scheme, it is desirable that the sum of the forces on all the particles is zero to machine precision, because this is a prerequisite for conservation of the simulation's total momentum; if momentum is not conserved the system is liable to rocket off the grid. Consider therefore the sum of all forces. We let  $W_\mathbf{p}(\mathbf{r})$  denote the fraction of the mass of a particle that lies at  $\mathbf{r}$  that is assigned to node  $\mathbf{p}$ , so

$$m_\mathbf{p} = \sum_{\text{particles } \beta} m_\beta W_\mathbf{p}(\mathbf{r}_\beta). \quad (2.241)$$

We introduce similar functions  $Q_\mathbf{j}(\mathbf{r})$  to describe the interpolation scheme used to calculate forces on particles given the field on the mesh: let the gravitational field at  $\mathbf{r}$  be

$$\mathbf{g}(\mathbf{r}) = \sum_{\text{nodes } \mathbf{j}} Q_\mathbf{j}(\mathbf{r}) \mathbf{g}_\mathbf{j}. \quad (2.242)$$

Then with equation (2.237), (2.240) and (2.241), the sum of all particle forces is

$$\begin{aligned} \mathbf{F}_{\text{tot}} &= \sum_{\text{particles } \alpha} m_\alpha \mathbf{g}(\mathbf{r}_\alpha) = \sum_{\alpha} m_\alpha \sum_{\text{nodes } \mathbf{j}} Q_\mathbf{j}(\mathbf{r}_\alpha) \sum_{\mathbf{n}} \mathbf{A}_\mathbf{n} \Phi_{\mathbf{j}+\mathbf{n}} \\ &= G \sum_{\mathbf{j}, \mathbf{p}} \sum_{\alpha, \beta} m_\alpha m_\beta Q_\mathbf{j}(\mathbf{r}_\alpha) W_\mathbf{p}(\mathbf{r}_\beta) \sum_{\mathbf{n}} \mathbf{A}_\mathbf{n} S_{\mathbf{j}+\mathbf{n}-\mathbf{p}}. \end{aligned} \quad (2.243)$$

The sum over  $\mathbf{n}$ , which is the numerical derivative of the softening kernel, is antisymmetric in  $\mathbf{j}$  and  $\mathbf{p}$ , as can be seen by replacing the dummy variable  $\mathbf{n}$  by  $-\mathbf{n}$  and exploiting the antisymmetry of  $\mathbf{A}_\mathbf{n}$  and the symmetry of  $S_\mathbf{q}$ . In view of this antisymmetry,  $\mathbf{F}_{\text{tot}}$  vanishes if  $Q_\mathbf{p} = W_\mathbf{p}$ , and in fact only this



choice of  $Q_{\mathbf{p}}$  guarantees symmetry of the sum over  $\alpha\beta$  for all possible particle locations. Therefore, we conclude that net momentum can be conserved if and only if we use the same scheme to interpolate forces from the grid as we use to assign mass to the grid. In particular, if we use the NGP scheme, forces change discontinuously midway between nodes.

**(c) Mesh refinement** Stellar systems, whether galaxies, galactic nuclei, star clusters or clusters of galaxies, tend to develop small, tightly bound condensations. These objects are often of considerable astronomical interest. For example, the visible stars of the Milky Way are a dense condensation in a much larger dark-matter halo. The ideal Poisson solver would determine the gravitational field of a structure no matter how small the structure becomes.

In the CIC and TSC mass-assignment schemes, the force between two particles that are separated by no more than the mesh spacing is significantly smaller than the value predicted by Newton's inverse-square law, regardless of the value of  $\epsilon$  that is used in the softening kernel. Hence the dynamics of structures smaller than a few mesh spacings cannot be faithfully followed.

Stellar systems usually have large volumes in which the density is low, and small volumes in which it is high. Hence achieving higher spatial resolution by subdividing the grid (**mesh refinement**) throughout the volume occupied by a system is undesirable: in low-density regions a fine mesh is both computationally expensive and pointless. If higher resolution is to be attained by mesh refinement, we must refine locally and adaptively—at each timestep, high-density regions must be located and equipped with a finer mesh. Handling an adaptive mesh requires intricate code, and if the mesh has a complex geometry, it will not be possible to use DFTs to evaluate the sum (2.237) efficiently. For a discussion of adaptive-mesh codes see Knebe, Green, & Binney (2001).

**(d)  $P^3M$  codes** The **Particle-particle-particle-mesh** or  $P^3M$  technique provides a simpler way of enhancing the spatial resolution of a PM code. The idea behind a  $P^3M$  code is to use the standard PM algorithm to calculate the contribution from distant particles to the force  $\mathbf{F}$  on a particle, while using the direct sum (2.224) to get the contribution to  $\mathbf{F}$  from particles that lie in the same or adjacent cells. The softening kernel  $S_F$  used for the direct sum must be carefully chosen to provide only the difference between the Newtonian force and the force already obtained from the grid.

The efficiency of a  $P^3M$  code depends sensitively on the number of particles,  $N_{\max}$ , in the most densely populated cell, since the cost of evaluating the direct sum scales as  $N_{\max}^2$  rather than as  $N_{\max} \ln(N_{\max})$ .  $P^3M$  codes have been extensively used for cosmological simulations (Efstathiou et al. 1985). In these simulations, galaxy formation leads gradually to higher and higher densities, so  $N_{\max}$  eventually becomes very large and the simulations grind to a halt. This problem can be resolved by using either a tree algorithm (Bode & Ostriker 2003) or a separate  $P^3M$  implementation (Couchman 1991) to handle sums within populous cells.

### 2.9.4 Spherical-harmonic codes

We saw in §2.4 and Figure 2.10 that good approximations to the potentials of stellar systems that are not too far from spherical can be obtained from the first few terms of the spherical-harmonic expansion (2.95).

Two rather different approaches to the numerical implementation of equation (2.95) have been widely used. In the first (McGlynn 1984; Bontekoe 1988), we cover the computational volume with a spherical mesh centered on the estimated location of the center of the system. We identify the particles that lie within each spherical shell around this center, say between radii  $a_i - \Delta/2$  and  $a_i + \Delta/2$ , and then evaluate the integrals (2.94) up to  $l = l_{\max}$ , using the formula

$$m_{lm}(a_i) \simeq \sum_{a_i - \Delta/2 < r_\alpha \leq a_i + \Delta/2} m_\alpha Y_l^{m*}(\theta_\alpha, \phi_\alpha). \quad (2.244)$$

Once the  $m_{lm}(a_i)$  have been evaluated for each shell  $i$ , we use an interpolation algorithm to construct a continuous function  $m_{lm}(a)$ . Then we can numerically estimate at  $r$  the value of the big brackets in equation (2.95):

$$Q_{lm}(r) \equiv \frac{1}{r^{l+1}} \int_0^r da a^{(l+2)} m_{lm}(a) + r^l \int_r^\infty \frac{da}{a^{l-1}} m_{lm}(a). \quad (2.245)$$

The potential at the location of any particle can be calculated from the sum

$$\Phi(\mathbf{r}_\alpha) = -4\pi G \sum_{l=0}^{l_{\max}} \sum_{m=-l}^l \frac{Y_l^m(\theta_\alpha, \phi_\alpha)}{2l+1} Q_{lm}(r_\alpha). \quad (2.246)$$

The second approach uses the functional-expansion technique introduced in §2.8 (Saha 1993). When setting up the simulation we choose a set of functions  $F_{nl}(r)$  in which to do the expansion, and for the first few values of  $l$  evaluate the matrix  $\mathbf{M}$  that is defined by the first of equations (2.222). Finally we store the inverses of these matrices. Then at each timestep we approximate the quantities that are defined by the second of equations (2.222) as sums over particles

$$s_{\mathbf{n}} \simeq \sum_{\alpha} m_\alpha Y_l^{m*}(\theta_\alpha, \phi_\alpha) F_{nl}(r_\alpha), \quad (2.247)$$

where  $\mathbf{n} \equiv (n, l, m)$ . Now we solve the linear equations (2.214) for the  $a_{\mathbf{n}}$  and have that the potential is

$$\Phi(\mathbf{r}) = \sum_{\mathbf{n}} a_{\mathbf{n}} Y_l^m(\boldsymbol{\Omega}) F_{nl}(r). \quad (2.248)$$

When properly constructed—in particular when the code includes all allowed values of  $m$  for each  $l$ —spherical-harmonic codes conserve angular momentum (relative to the chosen center) to machine accuracy, but they do not conserve linear momentum.

### 2.9.5 Simulations of planar systems

Studies of the dynamics of thin disks often confine the particles to a plane. In this case the softening length  $\epsilon$  can be interpreted as the characteristic thickness of the disk: if two particles pass at a fixed vertical separation  $\Delta z$  from one another, their mutual interaction potential is given by equation (2.225) with  $\mathbf{r}_\alpha$  and  $\mathbf{r}_\beta$  two-dimensional vectors, and  $\epsilon$  set equal to  $\Delta z$  in equation (2.226).

A two-dimensional mesh that is  $K$  cells on a side consumes less computer memory and CPU time than the corresponding three-dimensional mesh by a factor  $K$ , so PM codes are much more attractive in two dimensions than in three. Both rectangular (Hohl & Hockney 1969) and polar (Sellwood 1983) meshes have been widely used. On going from three dimensions to two, the functional expansion technique yields a similar efficiency gain because the number of subscripts  $n, l$  etc., that need be summed over decreases from three to two (Earn & Sellwood 1995).

## Problems

**2.1** [1] Show that the gravitational potential energy of a spherical system of finite mass in which the density satisfies  $\lim_{r \rightarrow 0} \rho r^{5/2} = 0$  can be written

$$W = -\frac{G}{2} \int_0^\infty dr \frac{M^2(r)}{r^2}, \quad (2.249)$$

where  $M(r)$  is the mass interior to radius  $r$ .

**2.2** [1] Prove that the Chandrasekhar potential-energy tensor for any spherical body has the form  $W_{jk} = \frac{1}{3}W\delta_{jk}$ , where  $W$  is the potential energy. Hint: start from equation (2.19).

**2.3** [1] Show that the potential of an infinite razor-thin sheet of surface density  $\Sigma$  in the plane  $z = 0$  is  $\Phi = 2\pi G\Sigma|z| + \text{constant}$ , (a) using Gauss's theorem, and (b) from Poisson's equation.

**2.4** [1] (Suggested by A. Toomre) Show that  $\Phi = \ln[r(1 + |\cos(\theta)|)]$  solves Laplace's equation everywhere except when  $r = 0$  or  $\theta = \pi/2$ . By applying Gauss's theorem near  $\theta = \pi/2$ , find the potential of the Mestel disk (2.158) in the limit  $R_{\max} \rightarrow \infty$ .

**2.5** [2] The **finite Mestel disk** is a razor-thin disk with surface density  $\Sigma(R)$  such that (i)  $\Sigma(R) = 0$  for all  $R > R_0$ ; (ii) the circular speed is  $v_c(R) = v_0 = \text{constant}$  for all  $R < R_0$ . The surface density of the finite Mestel disk was first derived by Mestel (1963) but here we describe a short, elegant derivation due to Brada & Milgrom (1995).

(a) Consider a spherical mass distribution with density  $\rho(r) = \frac{1}{2}Ar^{-2}$  for  $r < R_0$  and zero for  $r > R_0$ , where  $A$  is a constant. Argue that the circular speed is independent of radius and independent of  $R_0$  so long as  $r < R_0$ .

(b) Now squash the sphere along one direction, so that its isodensity surfaces are spheroids with axis ratio  $q$ . Argue that the circular speed in the equatorial plane of the squashed system is still independent of radius and independent of  $R_0$  so long as the equatorial radius  $R < R_0$ . Hint: use Newton's third theorem.

(c) By considering the limit  $q \rightarrow 0$ , show that a disk with surface density

$$\Sigma(R) = \begin{cases} (A/R) \cos^{-1}(R/R_0) & (R < R_0) \\ 0 & (R > R_0) \end{cases} \quad (2.250)$$

has a flat circular-speed curve for  $R < R_0$ .

(d) Show that  $v_0^2 = \pi^2 GA$ . Hint: let  $R_0 \rightarrow \infty$  and compare to the infinite Mestel disk.

**2.6** [1] Defining **prolate spheroidal coordinates**  $(u, v)$  by  $R = a \sinh u \sin v$ ,  $z = a \cosh u \cos v$ , where  $a > 0$  is a constant, show that  $R^2 + (a + |z|)^2 = a^2(\cosh u + |\cos v|)^2$ . Hence show that the potential (2.68a) of the Kuzmin disk can be written

$$\Phi_K(u, v) = -\frac{GM}{a} \frac{\cosh u - |\cos v|}{\sinh^2 u + \sin^2 v}. \quad (2.251)$$

In §3.5.3 we show that this potential is an example of a Stäckel potential, in which orbits admit an extra isolating integral.

**2.7** [2] Astronauts orbiting an unexplored planet find that (i) the surface of the planet is precisely spherical and centered on  $r = 0$ ; and (ii) the potential exterior to the planetary surface is  $\Phi = -GM/r$  exactly, that is, there are no non-zero multipole moments other than the monopole. Can they conclude from these observations that the mass distribution in the interior of the planet is spherically symmetric? If not, give a simple example of a non-spherical mass distribution that would reproduce the observations.

**2.8** [1] (Suggested by L. Ciotti) If a transparent, spherical stellar system has constant mass-to-light ratio  $\Upsilon$ , prove that the potential at radius  $r$  is

$$\Phi(r) = -\frac{2G\Upsilon}{r} \int_0^r dx S(x), \quad (2.252)$$

where  $S(x)$  is the strip brightness defined in Problem 1.3.

**2.9** [1] If a transparent, spherical stellar system has constant mass-to-light ratio  $\Upsilon$ , prove that the central potential is (Ciotti 1991)

$$\Phi(0) = -4G\Upsilon \int_0^\infty dR I(R), \quad (2.253)$$

where  $I(R)$  is the surface brightness at projected radius  $R$ .

**2.10** [2] Consider an axisymmetric body whose density distribution is  $\rho(R, z)$  and total mass is  $M = \int d^3\mathbf{r} \rho(R, z)$ . Assume that the body has finite extent,  $\rho(R, z) = 0$  for  $r^2 = R^2 + z^2 > r_{\max}^2$ , and is symmetric about its equator, that is,  $\rho(R, -z) = \rho(R, z)$ .

(a) Show that at distances large compared to  $r_{\max}$ , the potential arising from this body can be written in the form

$$\Phi(R, z) \simeq -\frac{GM}{r} - \frac{G}{4} \frac{(R^2 - 2z^2)}{r^5} \int d^3\mathbf{r}' \rho(R', z') (R'^2 - 2z'^2), \quad (2.254)$$

where the error is of order  $(r_{\max}/r)^2$  smaller than the second term.

(b) Show that at large distances from an exponential disk with surface density  $\Sigma(R) = \Sigma_0 \exp(-R/R_d)$ , the potential has the form

$$\Phi(R, z) \simeq -\frac{GM}{r} \left[ 1 + \frac{3R_d^2(R^2 - 2z^2)}{2r^4} + O(R_d^4/r^4) \right], \quad (2.255)$$

where  $M$  is the mass of the disk.

**2.11** [2] Show that the potential energy of an exponential disk is  $W \simeq -11.627 G \Sigma_0^2 R_d^3$ . Show further that if all stars move on circular orbits, the disk's angular momentum is  $J \simeq 17.462 G^{1/2} \Sigma_0^{3/2} R_d^{7/2}$  and its kinetic energy is  $K \simeq 5.813 G \Sigma_0^2 R_d^3$ . Hence show that for this disk the dimensionless spin parameter  $\lambda \equiv J|E|^{1/2}/GM^{5/2} \simeq 0.4255$ , where  $E = K + W$  is the total energy.

**2.12** [1] The  $r^{-1}$  dependence of the gravitational potential on distance arises because the graviton, which carries the gravitational field, is massless. If the graviton had a mass  $m_g$ , the gravitational potential due to a body of mass  $M$  would be  $\Phi(r) = -GM e^{-\alpha r}/r$ , where  $\alpha = m_g c/\hbar$  (the **Yukawa potential**), which reduces to the Newtonian potential in the limit  $\alpha \rightarrow 0$ . What is the analog of Poisson's equation (2.10) for the Yukawa potential?

**2.13** [2] Prove that the external potentials and gravitational fields of any two confocal spheroids of uniform density and equal mass are everywhere the same.

**2.14** [2] Use equation (2.140) to show that a prolate body with density  $\rho = \rho_0(1 + R^2/a_1^2 + z^2/a_3^2)^{-2}$ , where  $a_3 > a_1$ , generates the potential

$$\Phi(u, v) = -\pi G \rho_0 a_1^2 a_3 \int_0^\infty d\tau \frac{\sqrt{a_3^2 + \tau}}{(\tau + a_3^2 + \lambda)(\tau + a_3^2 + \mu)}, \quad (2.256)$$

where  $(u, v)$  are oblate spheroidal coordinates defined by equation (2.96) with  $\Delta^2 = a_3^2 - a_1^2$ , and we have written  $\lambda \equiv \Delta^2 \sinh^2 u$ ,  $\mu \equiv -\Delta^2 \cos^2 v$ . Decompose the integral in (2.256) into partial fractions to show (without evaluating the integrals) that  $\Phi$  is a Stäckel potential

$$\Phi(\lambda, \mu) = \frac{H(\lambda) - H(\mu)}{\lambda - \mu}, \quad (2.257)$$

where  $H$  is a continuous function (de Zeeuw 1985 and §3.5.3). Finally, show that

$$\Phi(u, v) = -\frac{2\pi G a_1^2 a_3 \rho_0}{\Delta^2} \frac{f(\Delta \sinh u) - f(i\Delta \cos v)}{\sinh^2 u + \cos^2 v}, \quad (2.258a)$$

where

$$f(z) \equiv z \tan^{-1}(z/a_3). \quad (2.258b)$$

Hint: to ensure convergence of the integrals, you may wish to add  $(\tau + a_3^2)^{-\frac{1}{2}}$  to one of the integrands and subtract it from the other. The body with this potential is called the **perfect prolate spheroid**, because it is the only prolate axisymmetric density distribution of constant ellipticity that has a Stäckel potential.

**2.15** [1] Show that the central potential of a thin axisymmetric disk is

$$\Phi(0, 0) = -2\pi G \int_0^\infty dR \Sigma(R). \quad (2.259)$$

Hint: use equation (C.68).

**2.16** [1] Prove that the potential  $\Phi(r)$  is a non-decreasing function of  $r$  in any spherical system. Does the same conclusion hold in an axisymmetric razor-thin disk? If so, prove it; if not, find a counter-example.

**2.17** [2] (Suggested by M. Merrifield) An axisymmetric disk is seen edge-on and has projected mass per unit length  $\mu(X)$ . Show that its surface density is

$$\Sigma(R) = -\frac{1}{\pi} \int_R^\infty \frac{dX}{\sqrt{X^2 - R^2}} \frac{d\mu}{dX}, \quad (2.260)$$

and that its potential is

$$\Phi(R, z) = 2G \int_0^\infty dX \frac{d\mu}{dX} \sin^{-1} \left( \frac{2X}{\sqrt{z^2 + (R+X)^2} + \sqrt{z^2 + (R-X)^2}} \right). \quad (2.261)$$

Hint: use (2.153a).

**2.18** [1] Use equation (2.190) to show that the razor-thin disk for which the circular speed is given by

$$v_c^2 = \frac{v_0^2}{\sqrt{1 + R^2/a^2}} \quad (2.262a)$$

has surface density (Toomre 1963)

$$\Sigma(R) = \frac{v_0^2}{2\pi G R} \left( 1 - \frac{1}{\sqrt{1 + a^2/R^2}} \right). \quad (2.262b)$$

Show that these formulae correspond to the Mestel disk in an appropriate limit. Show that the surface density (2.68b) of the Kuzmin disk is obtained when  $\Sigma(R)$  is differentiated with respect to  $a^2$ , and hence recover the circular speed of this model.

**2.19** [3] We have derived relations between the potential and surface density of non-axisymmetric disks by solving Laplace's equation in cylindrical coordinates (§2.6.2) and oblate spheroidal coordinates (§2.6.4). Derive a relation of this kind by solving Laplace's equation in spherical coordinates, and show that the result is identical with the formula derived using logarithmic spirals (eq. 2.199). Hint: you may need associated Legendre functions  $P_\lambda^m(x)$ , where  $\lambda$  is a real number. See also equations (C.12) and (C.31).

**2.20** [2] Show that the circular speed  $v_c(R)$  in a thin axisymmetric disk of surface density  $\Sigma(R)$  may be written in the form (Mestel 1963)

$$v_c^2(R) = \frac{GM(R)}{R} + 2G \sum_{k=1}^{\infty} \alpha_k \left[ \frac{2k+1}{R^{2k+1}} \int_0^R dR' \Sigma(R') R'^{2k+1} - 2kR^{2k} \int_R^{\infty} dR' \frac{\Sigma(R')}{R'^{2k}} \right], \quad (2.263)$$

where

$$\alpha_k \equiv \int_0^\pi d\theta P_{2k}(\cos \theta) = \pi \left[ \frac{(2k)!}{2^{2k}(k!)^2} \right]^2. \quad (2.264)$$

Hint: start with equation (2.3) and expand  $|\mathbf{x} - \mathbf{x}'|^{-1}$  in Legendre polynomials using equation (C.35).

**2.21** [2] Show that the potential of an axisymmetric disk with surface density  $\Sigma(R)$  is

$$\Phi(R, z) = -\frac{2G}{\sqrt{R}} \int_0^\infty dR' \Sigma(R') kK(k) \sqrt{R'}, \quad (2.265)$$

where  $K(k)$  is a complete elliptic integral (Appendix C.4) and

$$k^2 \equiv \frac{4RR'}{(R+R')^2 + z^2}. \quad (2.266)$$

Hint: start with equation (2.3) and show that  $|\mathbf{x} - \mathbf{x}'|^2 = 4RR'k^{-2}\{1 - k^2 \cos^2[\frac{1}{2}(\phi - \phi')]\}$ . Note that the integral (2.265) has a logarithmic singularity when  $z = 0$  and  $R' \rightarrow R$ , which requires some care when the integral is being evaluated numerically.

**2.22** [3] (Suggested by H. Dejonghe) Prove that the surface density  $\Sigma(x, y)$  and potential  $\Phi(x, y)$  in a disk occupying the plane  $z = 0$  are related by

$$\Sigma(x', y') = \frac{1}{4\pi^2 G} \iint \frac{dx dy}{|\mathbf{x} - \mathbf{x}'|} \left( \frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} \right). \quad (2.267)$$

**2.23** [1] Consider the discrete form of Poisson's equation with periodic boundary conditions that is obtained by using the approximation (2.235) for the value of the Laplacian on a grid with  $K$  mesh cells on a side. Write  $\Phi_{lmn}$  and  $\rho_{lmn}$  in terms of their DFTs,

$$\Phi_{\mathbf{r}} = \sum_{\mathbf{k}} \hat{\Phi}_{\mathbf{k}} e^{2\pi i \mathbf{k} \cdot \mathbf{r} / K} \quad ; \quad \rho_{\mathbf{r}} = \sum_{\mathbf{k}} \hat{\rho}_{\mathbf{k}} e^{2\pi i \mathbf{k} \cdot \mathbf{r} / K}, \quad (2.268)$$

where  $\mathbf{r}$  and  $\mathbf{k}$  are vectors with integer components in the range  $(0, K-1)$ . Show that

$$\Phi_{\mathbf{k}} = \frac{2\pi G \rho_{\mathbf{k}} \Delta^2}{\cos(2\pi k_x / K) + \cos(2\pi k_y / K) + \cos(2\pi k_z / K) - 3}. \quad (2.269)$$

**2.24** [2] In some numerical simulations of spherical stellar systems, spherical symmetry is enforced by treating each of the  $N$  stars as a "superstar," that is, a spherical shell containing a large number of stars with randomly oriented orbits but the same radius, radial velocity, and scalar angular momentum. Let  $m_n$  and  $r_n$  be the mass and radius of the  $n^{\text{th}}$  superstar. Assume that the superstars are sorted in order of increasing radius, that is,  $r_1 < r_2 < \dots < r_N$ .

- (a) Show that the force on superstar  $n$  is

$$\mathbf{F}_n = -\hat{\mathbf{e}}_r \frac{GM_n}{r_n^2} \quad \text{where} \quad M_n \equiv \sum_{j=1}^{n-1} m_j. \quad (2.270)$$

Thus show that the force on every star can be computed in  $O(N)$  steps once the stars are sorted in radius, in contrast to direct summation (§2.9.1), which requires  $O(N^2)$  steps.

- (b) A set of  $N$  superstars can be sorted into increasing order in  $O(N \ln N)$  steps (e.g., Press et al. 1986). Thus, the initial calculation of the forces for an unsorted system requires  $O(N \ln N)$  steps. Show that force calculations at subsequent timesteps require only  $O(N)$  steps.