4

Equilibria of Collisionless Systems

In §1.2 we introduced the idea that stellar systems may be considered to be collisionless: we obtain a good approximation to the orbit of any star by calculating the orbit that it would have if the system's mass were smoothly distributed in space rather than concentrated into nearly point-like stars. Eventually, the true orbit deviates significantly from this model orbit, but in systems with more than a few thousand stars, the deviation is small for a time $\lesssim t_{\rm relax}$ that is much larger than the crossing time $t_{\rm cross}$. In fact, for a galaxy $t_{\rm relax}$ is usually much larger even than the age of the universe, so the approximation that the potential is smooth provides a complete description of the dynamics.

In this chapter we consider model stellar systems that would be perfect equilibria if $t_{\rm relax}$ were arbitrarily large. Such models are the primary tool for comparisons of observations and theory of galaxy dynamics. In Chapter 7 we shall see that they are also applicable to globular clusters, even though $t_{\rm relax}$ is significantly smaller than the cluster's age, so long as it is recognized that the equilibrium evolves slowly, on a timescale of order $t_{\rm relax}$.

We assume throughout that the stellar systems we examine consist of N identical point masses, which might be stars or dark-matter particles. Although unrealistic, this assumption greatly facilitates our work and has no impact on the validity of our results.

In §4.1 we derive the equation that allows us to find equilibria, and discuss its connection to observational data. In §4.2 we show that solutions of the equation can be readily found if integrals of motion in the galactic potential are known, and in §§4.3 to 4.5 we use such solutions to study models with a variety of symmetries. In §4.6 we show that it is advantageous to express solutions in terms of action integrals. Unfortunately, in many practical cases insufficient integrals are known to obtain relevant solutions, so in §§4.7 and 4.8 we discuss alternative strategies, starting with heavily numerical approaches and moving on to approximate techniques that are based on moments of the fundamental equation. In §4.9 we draw on techniques developed throughout the chapter to hunt for massive black holes and dark halos in galaxies using observations of the kinematics of their stars. In §4.10 we address the question "what determines the distribution of stars in a galaxy?" This is a difficult question to which we shall have to return in Chapter 9.

4.1 The collisionless Boltzmann equation

When modeling a collisionless system such as an elliptical galaxy, it is neither practical nor worthwhile to follow the orbits of each of the galaxy's billions of stars. Most testable predictions depend on the probability of finding a star in the six-dimensional phase-space volume $\mathrm{d}^3\mathbf{x}\mathrm{d}^3\mathbf{v}$ around the position $\mathbf{x}$ and velocity $\mathbf{v}$. Therefore we define the **distribution function** (or DF for short) f such that $f(\mathbf{x},\mathbf{v},t)\,\mathrm{d}^3\mathbf{x}\mathrm{d}^3\mathbf{v}$ is the probability that at time t a randomly chosen star, say star 1, has phase-space coordinates in the given range. Since by assumption all stars are identical, this probability is the same for stars $2,3,\ldots,N$. By virtue of its definition f is normalized such that

$$\int \mathrm{d}^3\mathbf{x}\mathrm{d}^3\mathbf{v}\, f(\mathbf{x},\mathbf{v},t) = 1, \tag{4.1}$$

where the integral is over all phase space.

Let $\mathbf{w} = (\mathbf{x},\mathbf{v})$ be the usual Cartesian coordinates, and consider an arbitrary region $\mathcal{V}$ of phase space. The probability of finding star 1 in $\mathcal{V}$ is $P = \int_{\mathcal{V}} \mathrm{d}^6\mathbf{w}\, f(\mathbf{w})$. Let $\mathbf{W}$ represent some arbitrary set of phase-space coordinates, and let $F(\mathbf{W})$ be the corresponding DF; that is the probability of finding star 1 in $\mathcal{V}$ is $P = \int_{\mathcal{V}} \mathrm{d}^6\mathbf{W}\, F(\mathbf{W})$. If $\mathcal{V}$ is small enough, f and F will be approximately constant throughout it, and we can take them outside the integrals for P. Thus

$$P = f(\mathbf{w}) \int_{\mathcal{V}} \mathrm{d}^6\mathbf{w} = F(\mathbf{W}) \int_{\mathcal{V}} \mathrm{d}^6\mathbf{W}. \tag{4.2}$$

If the coordinates $\mathbf{W}$ are canonical, equation (D.81) implies that $\int_{\mathcal{V}} \mathrm{d}^6\mathbf{w} = \int_{\mathcal{V}} \mathrm{d}^6\mathbf{W}$. Substituting this relation into (4.2), we conclude that $F(\mathbf{W}) = f(\mathbf{w})$. Therefore, the DF has the same numerical value at a given phase-space point in *any* canonical coordinate system. This invariance enables us henceforth to treat $\mathbf{w} = (\mathbf{q}, \mathbf{p})$ as an arbitrary system of canonical coordinates.

Any given star moves through phase space, so the probability of finding it at any given phase-space location evolves with time. We now derive the differential equation that is satisfied by f as a consequence of this evolution. As f evolves, probability must be conserved, in the same way that mass is conserved in a fluid flow. The conservation of fluid mass is described by the continuity equation (F.3)

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial \mathbf{x}} \cdot (\rho \dot{\mathbf{x}}) = 0, \tag{4.3}$$

where ρ and $\dot{\mathbf{x}} = \mathbf{v}$ are the density and velocity of the fluid. The analogous equation for the conservation of probability in phase space is

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial \mathbf{w}} \cdot (f \dot{\mathbf{w}}) = 0. \tag{4.4}$$

We now use Hamilton's equations (D.54) to eliminate $\dot{\mathbf{w}} = (\dot{\mathbf{q}}, \dot{\mathbf{p}})$. The second term in equation (4.4) becomes

$$\begin{aligned} \frac{\partial}{\partial \mathbf{q}} \cdot (f \dot{\mathbf{q}}) + \frac{\partial}{\partial \mathbf{p}} \cdot (f \dot{\mathbf{p}}) &= \frac{\partial}{\partial \mathbf{q}} \cdot \left(f \frac{\partial H}{\partial \mathbf{p}} \right) - \frac{\partial}{\partial \mathbf{p}} \cdot \left(f \frac{\partial H}{\partial \mathbf{q}} \right) \\ &= \frac{\partial f}{\partial \mathbf{q}} \cdot \frac{\partial H}{\partial \mathbf{p}} - \frac{\partial f}{\partial \mathbf{p}} \cdot \frac{\partial H}{\partial \mathbf{q}} \\ &= \dot{\mathbf{q}} \cdot \frac{\partial f}{\partial \mathbf{q}} + \dot{\mathbf{p}} \cdot \frac{\partial f}{\partial \mathbf{p}}, \end{aligned} \tag{4.5}$$

where we have used the fact that $\partial^2 H/\partial \mathbf{q} \partial \mathbf{p} = \partial^2 H/\partial \mathbf{p} \partial \mathbf{q}$. Substituting this result into equation (4.4) we obtain the **collisionless Boltzmann equation**[1]

$$\frac{\partial f}{\partial t} + \dot{\mathbf{q}} \cdot \frac{\partial f}{\partial \mathbf{q}} + \dot{\mathbf{p}} \cdot \frac{\partial f}{\partial \mathbf{p}} = 0, \tag{4.6}$$

which is a partial differential equation for f as a function of six phase-space coordinates and time.

Equation (4.6) can be rewritten in a number of forms, each of which is useful in different contexts. Equation (4.5) enables us to write

$$\begin{aligned} 0 &= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{q}} \cdot \frac{\partial H}{\partial \mathbf{p}} - \frac{\partial f}{\partial \mathbf{p}} \cdot \frac{\partial H}{\partial \mathbf{q}} \\ &= \frac{\partial f}{\partial t} + [f, H], \end{aligned} \tag{4.7}$$

[1] Often also called the Vlasov equation, although it is a simplified version of an equation derived by L. Boltzmann in 1872. See Hénon (1982).

where the square bracket is a Poisson bracket (eq. D.65).

An alternative form of the collisionless Boltzmann equation can be derived by extending to six dimensions the concept of the convective or Lagrangian derivative (see eq. F.8). We define

$$\frac{\mathrm{d}f}{\mathrm{d}t} \equiv \frac{\partial f}{\partial t} + \dot{\mathbf{w}} \cdot \frac{\partial f}{\partial \mathbf{w}}; \tag{4.8}$$

$\mathrm{d}f/\mathrm{d}t$ represents the rate of change of the local probability density as seen by an observer who moves through phase space with a star. Comparison of equations (4.6) and (4.7) shows that $\dot{\mathbf{w}} \cdot (\partial f/\partial \mathbf{w}) = [f, H]$, so the convective derivative can also be written

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + [f, H], \tag{4.9}$$

and the collisionless Boltzmann equation (4.6) is simply

$$\frac{\mathrm{d}f}{\mathrm{d}t} = 0. \tag{4.10}$$

In words, the flow through phase space of the probability fluid is incompressible; the phase-space density f of the fluid around a given star always remains the same.[2] In contrast to flows of incompressible fluids such as water, the density will generally vary greatly from point to point in phase space; the density is constant as one follows the flow around a particular star but the density around different stars can be quite different.

In terms of inertial Cartesian coordinates, in which $H = \frac{1}{2}v^2 + \Phi(\mathbf{x}, t)$ with Φ the gravitational potential, the collisionless Boltzmann equation reads

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} - \frac{\partial \Phi}{\partial \mathbf{x}} \cdot \frac{\partial f}{\partial \mathbf{v}} = 0. \tag{4.11}$$

In cylindrical coordinates we have (eq. 3.66) $H = \frac{1}{2}(p_R^2 + p_\phi^2/R^2 + p_z^2) + \Phi$ so with (4.7) the collisionless Boltzmann equation becomes[3]

$$\begin{aligned} \frac{\partial f}{\partial t} + p_R \frac{\partial f}{\partial R} + \frac{p_\phi}{R^2}\frac{\partial f}{\partial \phi} + p_z \frac{\partial f}{\partial z} - \left(\frac{\partial \Phi}{\partial R} - \frac{p_\phi^2}{R^3}\right)\frac{\partial f}{\partial p_R} \\ - \frac{\partial \Phi}{\partial \phi}\frac{\partial f}{\partial p_\phi} - \frac{\partial \Phi}{\partial z}\frac{\partial f}{\partial p_z} = 0. \end{aligned} \tag{4.12}$$

[2] A simple example of an incompressible flow in phase space is provided by an idealized marathon race in which all runners travel at constant speeds: at the start of the course, the spatial density of runners is large but they travel at a wide variety of speeds; at the finish, the density is low, but at any given time all runners passing the finish line have nearly the same speed.

[3] A reader unconvinced of the usefulness of the Hamiltonian formalism should try deriving either (4.12) or (4.14) directly from (4.11).

To obtain the Hamiltonian for motion in spherical polar coordinates we replace in (3.218) $\partial S/\partial r$ by p_r, $\partial S/\partial\theta$ by p_θ and $\partial S/\partial\phi$ by p_ϕ and find

$$H = \tfrac{1}{2}\Big(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\phi^2}{r^2\sin^2\theta}\Big) + \Phi. \tag{4.13}$$

Using this expression in (4.7) we find

$$\begin{aligned}\frac{\partial f}{\partial t} + p_r\frac{\partial f}{\partial r} + \frac{p_\theta}{r^2}\frac{\partial f}{\partial\theta} + \frac{p_\phi}{r^2\sin^2\theta}\frac{\partial f}{\partial\phi} - \Big(\frac{\partial\Phi}{\partial r} - \frac{p_\theta^2}{r^3} - \frac{p_\phi^2}{r^3\sin^2\theta}\Big)\frac{\partial f}{\partial p_r}\\ - \Big(\frac{\partial\Phi}{\partial\theta} - \frac{p_\phi^2\cos\theta}{r^2\sin^3\theta}\Big)\frac{\partial f}{\partial p_\theta} - \frac{\partial\Phi}{\partial\phi}\frac{\partial f}{\partial p_\phi} = 0.\end{aligned} \tag{4.14}$$

Conversion to rotating coordinates is discussed in Problem 4.1.

4.1.1 Limitations of the collisionless Boltzmann equation

(a) Finite stellar lifetimes The physical basis of the collisionless Boltzmann equation is conservation of the objects that are described by the DF. Stars are not really conserved because they are born and die, so their flow through phase space would be more accurately described by an equation of the type

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \mathbf{v}\cdot\frac{\partial f}{\partial\mathbf{x}} - \frac{\partial\Phi}{\partial\mathbf{x}}\cdot\frac{\partial f}{\partial\mathbf{v}} = B - D, \tag{4.15}$$

where $B(\mathbf{x},\mathbf{v},t)$ and $D(\mathbf{x},\mathbf{v},t)$ are the rates per unit phase-space volume at which stars are born and die. In the collisionless Boltzmann equation, $B-D$ is set to zero. This is a useful approximation to the truth if and only if $B-D$ is smaller in magnitude than terms on the left of equation (4.15). The term $\mathbf{v}\cdot\partial f/\partial\mathbf{x}$ is of order vf/R, where v and R are the characteristic speed and radius in the galaxy. The ratio R/v is simply the crossing time $t_{\rm cross}$ (§1.2). Similarly, $\partial\Phi/\partial\mathbf{x}$ is of order the characteristic acceleration a, so the term $(\partial\Phi/\partial\mathbf{x})\cdot(\partial f/\partial\mathbf{v})$ is of order af/v. Since $a\approx v/t_{\rm cross}$, the two last terms in the middle section of equation (4.15) are of order $f/t_{\rm cross}$. Thus consider the ratio

$$\gamma = \left|\frac{B-D}{f/t_{\rm cross}}\right|. \tag{4.16}$$

The collisionless Boltzmann equation is valid if $\gamma \ll 1$, which requires that the fractional change in the number of stars per crossing time is small.

The significance of this criterion can be clarified by some concrete examples. We consider two contrasting stellar types: M dwarfs, which have masses $\lesssim 0.5\,\mathcal{M}_\odot$ and live longer than the age of the universe; and O stars, which have masses $\gtrsim 20\,\mathcal{M}_\odot$ and lifetimes $\lesssim 10\,$Myr (BM Tables 3.13 and 5.3). In an elliptical galaxy the rate of formation of M dwarfs is negligible, and even the oldest M dwarfs have not had time to evolve significantly. Hence, the collisionless Boltzmann equation will apply accurately to the DF of M dwarfs ($\gamma \leq 0.01$). Now consider the contrasting case of O stars in the Milky Way. These stars have lifetimes significantly shorter than a crossing time $\sim 100\,\text{Myr}$. In fact, an O star will scarcely move from its birthplace before it dies, and the phase-space distribution of such stars will depend entirely on the processes that govern star formation, and not at all on the collisionless Boltzmann equation ($\gamma \simeq 10$). In between these extremes, the collisionless Boltzmann equation will apply quite accurately to main-sequence populations in the Milky Way less massive than $\sim 1.5\,\mathcal{M}_\odot$, since these stars live for $\gtrsim 1\,\text{Gyr}$, which will generally be some tens of crossing times. In certain circumstances the collisionless Boltzmann equation may even be applied to a population of short-lived objects, for example planetary nebulae in an elliptical galaxy, because the phase-space distributions of the objects' births and deaths are to a good approximation identical, so $B - D \simeq 0$.

(b) Correlations between stars The average number density of stars in an infinitesimal volume of phase space is Nf. However, in practice all we can hope to measure is the number density in some volume of phase space large enough to contain many stars. The natural assumption to make is that the density in such a volume is simply $N\overline{f}$, where $\overline{f}$ is the average of f within this volume.[4] However, this assumption will only be correct if the positions of stars in phase space are uncorrelated: that is, knowing that star 1 is at $\mathbf{w}$ makes it neither more nor less likely that another star, say star 2, is at an adjacent phase-space location $\mathbf{w}'$. Mathematically, we assume that the probability of finding star 1 in the volume $\mathrm{d}^6\mathbf{w}$ at $\mathbf{w}$ and star 2 in $\mathrm{d}^6\mathbf{w}'$ at $\mathbf{w}'$ is simply the product $f(\mathbf{w})\mathrm{d}^6\mathbf{w}\, f(\mathbf{w}')\mathrm{d}^6\mathbf{w}'$ of the probabilities of finding star 1 at $\mathbf{w}$ and star 2 at $\mathbf{w}'$—in §7.2.4 we shall call such distributions "separable". When the assumption of separability holds, the probability $P_\mathcal{V}(k)$ that we will find k stars in a given volume $\mathcal{V}$ of phase space is given by the Poisson distribution (Appendix B.8)

$$P_\mathcal{V}(k) = \frac{\mu^k}{k!}\mathrm{e}^{-\mu} \quad \text{where} \quad \mu \equiv N\overline{f}\mathcal{V}. \tag{4.17}$$

It is easy to show that the mean number of stars predicted by this probability distribution is $\langle k\rangle = N\overline{f}\mathcal{V}$. Thus $N\overline{f}$ is indeed the expectation value of the stellar number density, if the DF is separable. Two obvious corollaries are that the mean mass within $\mathcal{V}$ is

$$\langle m\rangle = M\overline{f}(\mathbf{w})\mathcal{V}, \tag{4.18}$$

where M is the total mass of the stellar system, and the mean luminosity emitted within $\mathcal{V}$ is

$$\langle l\rangle = L\overline{f}(\mathbf{w})\mathcal{V}, \tag{4.19}$$

[4] This function is sometimes called the **coarse-grained** DF. The standard DF is then called the **fine-grained** DF to eliminate any danger of confusion with $\overline{f}$.

where L is the system's luminosity.

In reality, the presence of star 1 at $\mathbf{x}$ always increases the probability that star 2 will be found at some nearby position $\mathbf{x}'$ because stars attract one another. Hence, the assumption that the probability distribution of individual stars is separable is never strictly valid. In Chapter 7 we shall explore the effect of such correlations on the evolution of stellar systems. However, in this chapter we assume that separability holds, as it very nearly does for many stellar systems, because the force on a star from its neighbors is very much smaller than the force from the rest of the system.

4.1.2 Relation between the DF and observables

At any fixed position $\mathbf{x}$, the integral

$$\nu(\mathbf{x}) \equiv \int \mathrm{d}^3\mathbf{v}\, f(\mathbf{x},\mathbf{v}) \tag{4.20}$$

gives the probability per unit volume of finding a particular star at $\mathbf{x}$, regardless of its velocity. Multiplying by the total number N of stars in the population, we obtain the real-space number density of stars

$$n(\mathbf{x}) \equiv N\nu(\mathbf{x}). \tag{4.21}$$

In the Galaxy $n(\mathbf{x})$ can in principle be determined from star counts, and thus $\nu(\mathbf{x})$ can be derived from $n(\mathbf{x})$. In other galaxies it is not usually possible to count stars, but we can derive $\nu(\mathbf{x})$ from the luminosity density $j(\mathbf{x}) = L\nu(\mathbf{x})$, where L is the luminosity of the stellar population (BM §4.2.3).

It is often convenient to modify the definition of the DF so that $f\mathrm{d}^6\mathbf{w}$ represents not the probability of finding a given star in the phase-space volume $\mathrm{d}^6\mathbf{w}$, but rather the expected number, total mass, or total luminosity of the stars in $\mathrm{d}^6\mathbf{w}$. These modifications correspond to multiplying f by N, M, or L, respectively. Ideally these different definitions would be reflected in different notations for the DF. In practice the definition is usually clear from the context, and f is conventionally used to denote all of these quantities.

Dividing f by ν we obtain the probability distribution of stellar velocities at $\mathbf{x}$

$$P_{\mathbf{x}}(\mathbf{v}) = \frac{f(\mathbf{x},\mathbf{v})}{\nu(\mathbf{x})}, \tag{4.22}$$

which can be directly measured near the Sun (BM §10.3). In external galaxies $P_{\mathbf{x}}$ can be probed through the **line-of-sight velocity distribution** (LOSVD; BM §11.1), which gives for a particular line of sight through the galaxy the fraction $F(v_\parallel)\mathrm{d}v_\parallel$ of the stars that have line-of-sight velocity within $\mathrm{d}v_\parallel$ of $v_\parallel$. Almost all galaxies are sufficiently far away that all vectors from the observer to a point $\mathbf{x}$ in the galaxy are very nearly parallel to the fixed unit vector $\hat{\mathbf{s}}$ from the observer to the center of the galaxy. Then $x_\parallel \equiv \hat{\mathbf{s}} \cdot \mathbf{x}$ and $v_\parallel \equiv \hat{\mathbf{s}} \cdot \mathbf{v}$ are the components of $\mathbf{x}$ and $\mathbf{v}$ parallel to the line of sight. We also define $\mathbf{x}_\perp \equiv \mathbf{x} - x_\parallel \hat{\mathbf{s}}$ and $\mathbf{v}_\perp \equiv \mathbf{v} - v_\parallel \hat{\mathbf{s}}$ to be the components of $\mathbf{x}$ and $\mathbf{v}$ in the plane of the sky. The relation between $P_\mathbf{x}(\mathbf{v})$ and $F(\mathbf{x}_\perp, v_\parallel)$ is

$$\begin{aligned} F(\mathbf{x}_\perp, v_\parallel) &= \frac{\int \mathrm{d}x_\parallel\, \nu(\mathbf{x}) \int \mathrm{d}^2\mathbf{v}_\perp\, P_\mathbf{x}(v_\parallel \hat{\mathbf{s}} + \mathbf{v}_\perp)}{\int \mathrm{d}x_\parallel\, \nu(\mathbf{x})} \\ &= \frac{\int \mathrm{d}x_\parallel \mathrm{d}^2\mathbf{v}_\perp\, f(\mathbf{x}, \mathbf{v})}{\int \mathrm{d}x_\parallel \mathrm{d}^3\mathbf{v}\, f(\mathbf{x}, \mathbf{v})}. \end{aligned} \tag{4.23}$$

The LOSVD is frequently quantified by two numbers, the mean line-of-sight velocity $\overline{v}_\parallel$ and the dispersion $\sigma_\parallel$ about this mean. We have

$$\begin{aligned} \overline{v}_\parallel(\mathbf{x}_\perp) \equiv \int \mathrm{d}v_\parallel\, v_\parallel F(\mathbf{x}_\perp, v_\parallel) &= \frac{\int \mathrm{d}x_\parallel \mathrm{d}^3\mathbf{v}\, v_\parallel f(\mathbf{x}, \mathbf{v})}{\int \mathrm{d}x_\parallel \mathrm{d}^3\mathbf{v}\, f(\mathbf{x}, \mathbf{v})} \\ &= \frac{\int \mathrm{d}x_\parallel\, \nu(\mathbf{x})\, \hat{\mathbf{s}} \cdot \overline{\mathbf{v}}}{\int \mathrm{d}x_\parallel\, \nu(\mathbf{x})}, \end{aligned} \tag{4.24a}$$

where we have defined the **mean velocity** at location $\mathbf{x}$

$$\overline{\mathbf{v}}(\mathbf{x}) \equiv \int \mathrm{d}^3\mathbf{v}\, \mathbf{v} P_\mathbf{x}(\mathbf{v}) = \frac{1}{\nu(\mathbf{x})} \int \mathrm{d}^3\mathbf{v}\, \mathbf{v} f(\mathbf{x}, \mathbf{v}). \tag{4.24b}$$

The **line-of-sight velocity dispersion** is defined to be

$$\begin{aligned} \sigma_\parallel^2(\mathbf{x}_\perp) &\equiv \int \mathrm{d}v_\parallel\, (v_\parallel - \overline{v}_\parallel)^2 F(\mathbf{x}_\perp, v_\parallel) \\ &= \frac{\int \mathrm{d}x_\parallel \mathrm{d}^3\mathbf{v}\, (\hat{\mathbf{s}} \cdot \mathbf{v} - \overline{v}_\parallel)^2 f(\mathbf{x}, \mathbf{v})}{\int \mathrm{d}x_\parallel \mathrm{d}^3\mathbf{v}\, f(\mathbf{x}, \mathbf{v})}. \end{aligned} \tag{4.25}$$

The line-of-sight velocity dispersion is determined both by the variation in the mean velocity $\overline{v}_\parallel(\mathbf{x})$ along the line of sight, and the spread in stellar velocities at each point in the galaxy around $\overline{\mathbf{v}}(\mathbf{x})$. This spread is characterized by the **velocity-dispersion tensor**

$$\begin{aligned} \sigma_{ij}^2(\mathbf{x}) &\equiv \frac{1}{\nu(\mathbf{x})} \int \mathrm{d}^3\mathbf{v}\, (v_i - \overline{v}_i)(v_j - \overline{v}_j) f(\mathbf{x}, \mathbf{v}) \\ &= \overline{v_i v_j} - \overline{v}_i \overline{v}_j. \end{aligned} \tag{4.26}$$

The velocity-dispersion tensor is manifestly symmetric, so we know from matrix algebra that at any point $\mathbf{x}$ we may choose a set of orthogonal axes $\hat{\mathbf{e}}_i(\mathbf{x})$ in which $\boldsymbol{\sigma}^2$ is diagonal, that is, $\sigma_{ij}^2 = \sigma_{ii}^2 \delta_{ij}$ (no summation over i, and $\delta_{ij} = 1$ for $i = j$ and zero otherwise). The ellipsoid that has the

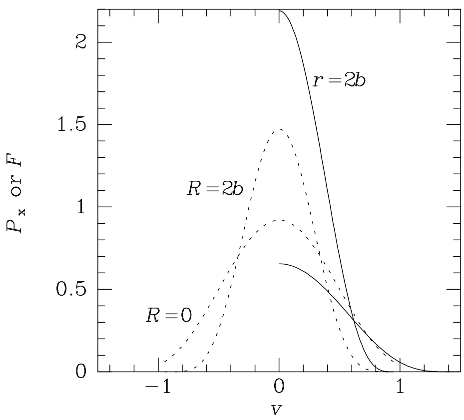

Figure 4.1 The full curves show the velocity distributions $P_{\mathbf{x}}(v)$ at the center of a Plummer model (lower curve) and at $r = 2b$ (upper curve). The dashed curves show the LOSVD $F(v_\parallel)$ along two lines of sight, $R = 0, 2b$. The DF of the Plummer model is given by equation (4.91b).

diagonalizing coordinate axes $\hat{\mathbf{e}}_i(\mathbf{x})$ for its principal axes and σ_{11}, σ_{22} and σ_{33} for its semi-axis lengths is called the **velocity ellipsoid** at $\mathbf{x}$.

To determine the relation between the velocity-dispersion tensor and the line-of-sight velocity dispersion, we let $u(\mathbf{x}) \equiv \hat{\mathbf{s}} \cdot \overline{\mathbf{v}}(\mathbf{x}) - \overline{v}_\parallel$ be the difference between the mean velocity parallel to the line of sight at $\mathbf{x}$ and the mean velocity for the entire line of sight. Then we can rewrite (4.25) as follows:

$$\begin{aligned}\sigma_\parallel^2(\mathbf{x}_\perp) &= \frac{\int \mathrm{d}x_\parallel \mathrm{d}^3\mathbf{v}\, [\hat{\mathbf{s}} \cdot (\mathbf{v} - \overline{\mathbf{v}}) + u]^2 f(\mathbf{x}, \mathbf{v})}{\int \mathrm{d}x_\parallel \mathrm{d}^3\mathbf{v}\, f(\mathbf{x}, \mathbf{v})} \\ &= \frac{\int \mathrm{d}x_\parallel\, \nu(\mathbf{x})\big(\hat{\mathbf{s}} \cdot \boldsymbol{\sigma}^2 \cdot \hat{\mathbf{s}} + u^2\big)}{\int \mathrm{d}x_\parallel\, \nu(\mathbf{x})},\end{aligned} \tag{4.27}$$

where we introduce the notation $\hat{\mathbf{s}} \cdot \boldsymbol{\sigma}^2 \cdot \hat{\mathbf{s}} \equiv \sum_{ij} \hat{s}_i \sigma_{ij}^2 \hat{s}_j$.

These results show that once ν, $\overline{\mathbf{v}}$ and $\boldsymbol{\sigma}^2$ are known at each point in a model, the observable quantities $v_\parallel$ and $\sigma_\parallel^2$ can be determined for that model. This fact makes ν, $\overline{\mathbf{v}}$ and σ_{ij}^2, all functions of $\mathbf{x}$, vital links between observations and theoretical models. Moreover, we shall see in §4.8 that in equilibrium stellar systems there are simple relations between these quantities and the gravitational field (the Jeans equations).

Notice that while $v_\parallel$ depends only on the mean velocity field $\overline{\mathbf{v}}(\mathbf{x})$, there are contributions to $\sigma_\parallel^2$ from both $\boldsymbol{\sigma}^2$ and $\overline{\mathbf{v}}$. Moreover, both contributions are inherently positive, so $\sigma_\parallel^2$ is in general larger than the average of the intrinsic squared velocity dispersion $\hat{\mathbf{s}} \cdot \boldsymbol{\sigma}^2 \cdot \hat{\mathbf{s}}$ along the line of sight.

One shortcoming of $\overline{v}_\parallel$ and $\sigma_\parallel^2$ as probes of the dynamics of a galaxy is that they are hard to measure accurately because they are sensitive to the contributions of the small number of high-velocity stars.

An example In §4.3.3a we shall encounter an exceptionally simple model system called the Plummer model. This is a non-rotating spherical system in which the velocity distribution $P_{\mathbf{x}}$ depends only on $v \equiv |\mathbf{v}|$, and the gravitational potential is given by equation (2.44a). The full curves in Figure 4.1 show $P_{\mathbf{x}}(v)$ at the center of the system and at $r = 2b$, where b is the Plummer scale length. Notice that $P_{\mathbf{x}}$ vanishes for speeds larger than the escape speed $\sqrt{2|\Phi(\mathbf{x})|}$ (eq. 2.31). At small radii, where $|\Phi|$ is relatively large, a graph of $P_{\mathbf{x}}$ versus v is wide and gently peaked, while at large radii, where $|\Phi|$ is much smaller, a plot of $P_{\mathbf{x}}(v)$ shows a high, narrow peak.

The LOSVD $F(v_{\|})$ along a line of sight through a Plummer model depends on the projected distance $R = |\mathbf{x}_{\perp}|$ between the line of sight and the center of the model because it is a weighted mean of the velocity distributions $P_{\mathbf{x}}(v)$ for different points along the line of sight (eq. 4.23). The dashed curves in Figure 4.1 show the LOSVD for the line of sight through the center, and one further out. Notice that the LOSVD at each radius is more centrally peaked than the velocity distribution at that radius. There are two reasons for this phenomenon. First, $F(v_{\|})$ is depressed at large values of $v_{\|}$ by the integral over $\mathbf{v}_{\perp}$ in (4.23) because the range of allowed values of $\mathbf{v}_{\perp}$ diminishes rapidly as $v_{\|}$ approaches the escape speed. A subsidiary effect is that the line of sight through the center samples points that are physically far from the cluster center, where the velocity distribution $P_{\mathbf{x}}$ is narrowly peaked around $v = 0$.

4.2 Jeans theorems

In §3.1.1 we introduced the concept of an integral of motion in a given stationary potential $\Phi(\mathbf{x})$. According to equation (3.56), a function of the phase-space coordinates $I(\mathbf{x}, \mathbf{v})$ is an integral if and only if

$$\frac{\mathrm{d}}{\mathrm{d}t} I[\mathbf{x}(t), \mathbf{v}(t)] = 0 \tag{4.28}$$

along any orbit. With the equations of motion this becomes

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \frac{\partial I}{\partial \mathbf{x}} \cdot \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} + \frac{\partial I}{\partial \mathbf{v}} \cdot \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = 0, \quad \text{or} \quad \mathbf{v} \cdot \frac{\partial I}{\partial \mathbf{x}} - \frac{\partial \Phi}{\partial \mathbf{x}} \cdot \frac{\partial I}{\partial \mathbf{v}} = 0. \tag{4.29}$$

Comparing this with equation (4.11), we see that the condition for I to be an integral is identical with the condition for I to be a steady-state solution of the collisionless Boltzmann equation. This leads to the following theorem, first stated by Jeans (1915).

Jeans theorem *Any steady-state solution of the collisionless Boltzmann equation depends on the phase-space coordinates only through integrals of motion in the given potential, and any function of the integrals yields a steady-state solution of the collisionless Boltzmann equation.*