The Astrophysical Journal, 391:531–549, 1992 June 1

INFERRING THE MASS OF SPHERICAL STELLAR SYSTEMS FROM VELOCITY MOMENTS

Herwig Dejonghe
Rijksuniversiteit Gent, Sterrenkundig Observatorium, Krijgslaan 281-S9, B-9000 Gent, Belgium

AND

David Merritt
Department of Physics and Astronomy, Serin Physics Laboratories, Rutgers University, Piscataway, NJ 08855
Received 1991 June 10; accepted 1991 December 4

ABSTRACT

We discuss the usefulness of line-of-sight velocity distributions for constraining the potential and kinematics of a nonrotating spherical system when nothing is known a priori about its radial mass distribution. The virial theorem, because it requires an assumed radial dependence for the potential, yields only very broad constraints on the total mass; the velocity dispersion profile is also insufficient to determine the potential, because of the unknown radial dependence of the anisotropy. To do better, one must make use of the additional information contained within the distribution of line-of-sight velocities at every projected radius. We develop a formalism, based on velocity moments, for doing so. First, we show that—if the potential $\Phi(r)$ is known, and the distribution function of the observed sample is of the form $f(E, L^2)$—the joint distribution of positions and velocities $v_p(r_p, v_p)$ uniquely determines that distribution function. We prove this by deriving expressions for the intrinsic velocity moments in terms of the projected moments, for all orders. As a by-product, we obtain an alternative set of Jeans equations in a very simple form. Second, we show that most assumed potentials are inconsistent with a given $v_p(r_p, v_p)$, since they imply intrinsic velocity moments that are negative at some radii, and hence a distribution function that is negative in some regions of phase space. The potential can therefore be constrained by the requirement that the inferred moments be nonnegative. We find no reason to believe that $\Phi(r)$ is *uniquely* determined by $v_p(r_p, v_p)$; however, a number of arguments suggest that it is very strongly constrained.

The expressions relating the observed velocity moments of order $2n$ to the intrinsic moments involve the nth power of $d\Phi/dr$. If we treat $\Phi(r)$ as unknown, an estimate of its most extreme functional forms can be determined through a process of constrained optimization, where the constraints (which include positivity of the inferred moments) are of order n in the coefficients defining the potential. When only the velocity dispersion profile is known, these constraints are linear; we solve the corresponding linear programming problem to find the most extreme potentials consistent with a family of velocity dispersion profiles derived from a generalized Plummer model. We find that, when the velocity dispersion profile is mildly centrally peaked, the range of allowed total masses and central densities is almost as great as the range allowed by the virial theorem alone; thus, the velocity dispersion profile typically contains little additional information about the radial form of the potential. When higher order velocity moments are specified, the problem becomes a more difficult one of nonlinearly constrained optimization. We show, for the same family of Plummer models, that the requirement of positive fourth-order moments can substantially reduce the range of allowed potentials. This leads us to suggest that methods that explicitly or implicitly account for positive distribution functions will constrain the potential to the fullest. However, given the difficulty of evaluating high-order velocity moments from imperfect data, we suggest an alternative scheme that uses the observed positions and velocities directly, without binning or computation of moments. Such a scheme should be especially efficient in cases where the velocity data are limited and discrete, as in dwarf spheroidal galaxies.

Subject headings: celestial mechanics, stellar dynamics — galaxies: kinematics and dynamics

1. INTRODUCTION

Broadly speaking, kinematical data may be used in two ways to constrain the dynamical state of a stellar system. If the underlying potential is thought to be known, and if an approximately steady state may be assumed, the data can be used to estimate the phase-space distribution function f of the observed sample, where f depends only on the isolating integrals of motion in the assumed potential. This is essentially the classical "Jeans problem," much studied in the context of our Galaxy. In the original, self-consistent formulation of the Jeans problem, the potential was calculated from the luminous matter. In modern astronomy, this prescription is much less justified, since we now know that much, even most, of the matter that determines the potential of a stellar system can be distributed in a very different manner from the light. In this more uncertain situation, greater demands are placed on the kinematical data: they must somehow yield both the phase-space distribution of luminous objects, as well as the functional form of the potential in which the objects move. Often these two problems are viewed as separate: for instance, the virial theorem appears to make a definite prediction about the total mass of a system that depends only on the mean square velocity of a set of test objects, with no assumptions about their detailed velocity distribution. However the virial theorem can only be applied if the spatial dependence of the potential is specified in advance; in the absence of such knowledge, it imposes only order-of-magnitude constraints on the total mass. The same is true of the usual modifications to the virial theorem, such as the "projected mass method" of Heisler, Tremaine, & Bahcall (1985): none of these is appropriate for estimating the total mass unless the form of the matter distribution is known.

An estimate of the spatial dependence of the potential can obviously only be made if the data themselves contain some information about the variation of kinematical quantities with position. In a few important cases, the form of the potential follows directly from the Jeans equation, using observable, low-order velocity moments of the phase-space distribution function of some set of test objects. For instance, the distribution of mass in a circular galactic disk follows immediately from the rotation curve (assuming a geometry for the potential, e.g., that all the matter lies in the disk); the distribution of mass perpendicular to a disk is given uniquely by the variation of number density and velocity dispersion of some test population with height (assuming a plane-stratified geometry). However in most astrophysically interesting cases, the potential is not uniquely specified by the observable moments. This is true, for instance, of spherical galaxies, in which the observed moments contain contributions from both the radial and tangential components of the intrinsic moments. In nonspherical, pressure-supported systems, the indeterminacy is even greater, since the intrinsic shape and orientation may be unknown, and the intrinsic velocity dispersion may have three independent components at every position. Most stellar systems apart from disk galaxies can be assigned to this category, including open and globular clusters, dwarf spheroidal galaxies, elliptical galaxies, and clusters of galaxies. The few studies of these systems that allow for different luminous and dark matter distributions generally conclude that a discouragingly large range of potentials can satisfy the data equally well, even under the restrictive assumption of spherical symmetry (e.g., Katz & Richstone 1985; Merritt 1987; Pryor & Kormendy 1990).

In spite of this indeterminacy, a number of observational programs are currently underway to gather large samples of velocity data for stellar systems of various types, with the goal of determining their masses and kinematics. Examples are studies of globular clusters, dwarf elliptical galaxies, and the Galactic bulge, based on velocities of individual stars (e.g., Meylan & Mayor 1991; Mateo et al. 1991; te Lintel Hekkert & Dejonghe 1989); plans to map the external potentials of early-type galaxies using velocities of the globular clusters or planetary nebulae that surround them (e.g., Mould et al. 1990; Ford et al. 1989); and the continuing accumulation of radial velocity data for galaxies in clusters (e.g., Teague, Carter, & Gray 1990). Much progress is also being made on the derivation of kinematical information from integrated spectra of elliptical galaxies: line-of-sight velocity distributions near the projected center, where surface brightnesses are high (e.g., Bender 1990); or more limited information, such as mean velocities and velocity dispersions, in the faint outskirts of these galaxies (e.g., Bertin et al. 1989). These programs have been motivated in part by the fact that the distribution and density of dark matter as a function of scale size is an important datum in constraining models for the formation of structure in the universe. However, in spite of the considerable observational effort being devoted to these programs, there is at present little concensus on how the new data can best be used.

The most complete kinematical information that one can hope to obtain in a spherical stellar system, based on distant observations at a single epoch, is the distribution of line-of-sight velocities as a function of projected radius $\nu_p(r_p, v_p)$, where $\nu_p(r_p, v_p)dv_p$ is the surface density at r_p of objects (e.g., stars) with line-of-sight velocities in the interval v_p to $v_p + dv_p$. (Throughout we neglect absorption, so that observed quantities are always simple projections along the line of sight.) Henceforth we will call $\nu_p(r_p, v_p)$ the "projected distribution function." The projected number density and velocity dispersion profiles are moments over v_p of $\nu_p(r_p, v_p)$, that is (assuming no rotation),

$$\nu_p(r_p) = \int_{-\infty}^{\infty} \nu_p(r_p, v_p)dv_p \, , \tag{1a}$$

$$\nu_p(r_p)\sigma_p^2(r_p) = \int_{-\infty}^{\infty} \nu_p(r_p, v_p)v_p^2 \, dv_p \, . \tag{1b}$$

Similarly, the "velocity moment profile" of nth order is

$$\mu_p^n(r_p) = \int_{-\infty}^{\infty} \nu_p(r_p, v_p)v_p^n \, dv_p \, . \tag{1c}$$

Mathematically, specification of the infinite set of moments $\mu_p^n(r_p)$ is equivalent to specification of $\nu_p(r_p, v_p)$; in practice, of course, only the lowest order moments are accessible to observation. The connection between the projected distribution function $\nu_p(r_p, v_p)$ and the intrinsic distribution function $f(r, v_r, v_T)$ is the rather daunting integral equation

$$\nu_p(r_p, v_p) = 2 \int_{r_p}^{\infty} \frac{r\,dr}{\sqrt{r^2 - r_p^2}} \iint dv_{r_p} dv_\theta \, f(r, v_r, v_T) \, , \tag{2}$$

where (v_{r_p}, v_θ) are Cartesian components along the polar coordinates (r_p, θ) in the plane of the sky, and (v_r, v_T) are the intrinsic, radial, and tangential velocity components. If $f(r, v_r, v_T)$ could be determined uniquely from $\nu_p(r_p, v_p)$, the problem would be solved, since both the phase-space density as well as the potential (from the Boltzmann—not the Poisson—equation) would be known. However, there are likely to be many $f(r, v_r, v_T)$ consistent with a given $\nu_p(r_p, v_p)$; nor is there any reason to expect that, among the infinite set of possible f, all imply the *same* potential.

To see how $\nu_p(r_p, v_p)$ might constrain $\Phi(r)$, it is helpful to make explicit use of the fact that f reflects a steady state, that is, that it depends on the phase-space variables only through the energy $E = v^2/2 + \Phi(r)$ and the squared angular momentum $L^2 = r^2 v_T^2$. This allows us to rewrite equation (2) as

$$\nu_p(r_p, v_p) = 2 \int_{r_p}^{\infty} \frac{r\,dr}{\sqrt{r^2 - r_p^2}} \iint dv_{r_p} dv_\theta \, f(E, L^2) \, . \tag{3}$$

In effect, we have imposed the additional requirement that the inferred f be constant on surfaces of constant E and L^2. Equation (3) is a complete statement of the problem: it contains not only the relation between observed and intrinsic quantities, but also the most general expression of the fact that the intrinsic distribution of positions and velocities must reflect a steady state. Now since both the (unknown) $f(E, L^2)$ and the (observed) $\nu_p(r_p, v_p)$ are functions of two variables, it is tempting to conclude from a naive application of Fredholm theory (integrals as continuous limits of systems of linear equations) that equation (3) has a solution $f(E, L^2)$ for *any* specified potential $\Phi(r)$, that is, that the kinematical data impose no constraints whatsoever on $\Phi(r)$. As it happens, this is indeed the case, at least in a narrow mathematical sense; the proof is given in § 2 below. However, it turns out that many assumed potentials imply, through equation (3), a distribution function that is negative in some parts of phase space, and that can therefore be rejected as unphysical. The problem of inferring the potential of a spherical stellar system from kinematical data can therefore be stated very generally as follows: given the joint distribution of projected radii and line-of-sight velocities, one searches for potentials that imply—through equation (3)—a distribution function that is everywhere nonnegative.

Stated in this way, it is clear that the problem of mass estimation in spherical systems is fundamentally different from the classical problems of determining the mass in and around spiral galaxies: the equation that determines $\Phi(r)$ is an inequality, that is, $f > 0$, in which f is defined implicitly, through a triple integral; and, expressed in terms of Φ, this equation is highly nonlinear. Thus we are faced with a "nonlinear implicit integral inequality." We do not know of any theory about such equations, let alone a general method of solution. Certainly such a problem must be very difficult, or impossible, to solve in the general case. There is no obvious reason to believe, for instance, that $\Phi(r)$ is *uniquely* determined by $\nu_p(r_p, v_p)$, even if the latter function were known exactly: it might be underdetermined in some cases, overdetermined in others.

This conclusion seems at first puzzling: we are accustomed to thinking of mass estimation in terms of equations like the virial theorem, which are linear, explicit relations (not inequalities) between known quantities (positions, velocities), and unknown ones (masses, densities). For instance, the classical virial theorem for a spherical stellar system states

$$\langle v^2 \rangle = \langle \boldsymbol{r} \cdot \nabla\Phi \rangle , \qquad (4)$$

where angle brackets denote number-weighted averages over all space for some sample of objects (e.g., Goldstein 1980). Comparing equation (4) to equation (3), and noting that $\langle v^2 \rangle$ is determined uniquely by $\nu_p(r_p, v_p)$, we might conclude that only *certain* $\Phi(r)$ (those satisfying eq. [4]) could be made consistent with a given projected distribution function. In fact, we demonstrate below that any choice of $\Phi(r)$ that is *not* consistent with the virial theorem necessarily implies a distribution function f whose second velocity moments are negative at large radii. More specifically, the virial theorem may be interpreted (at least in this context) as a condition on the potential such that the intrinsic second moments $\nu\sigma_r^2(r)$, $\nu\sigma_\vartheta^2(r)$ corresponding to an observed velocity dispersion profile $\sigma_p(r_p)$ do not become negative at large r. As it turns out, the equations relating the intrinsic second moments to the observed dispersions are linear in Φ; thus the virial theorem is a linear relation. Furthermore, the virial theorem does not guarantee positivity of f, even at the level of its second moments, at *all* radii; the equations which do so are—unlike the virial theorem—inequalities. In the same way, we show that a potential that violates the higher order virial constraints discussed by Kent (1991) implies a distribution function whose higher order velocity moments are negative at large radii. However, these virial constraints, important as they are, do not tell the whole story: there may well exist potentials that satisfy a set of virial constraints, but that, through equation (3), still imply a phase-space distribution function that is negative in *some* parts of phase space. It is this more general requirement of a positive phase-space density, and not the less restrictive virial theorems, that fundamentally constrains the potential.

In the present paper, then, we address the following two questions:

1. In an *assumed* spherical potential $\Phi(r)$, is $f(E, L^2)$ uniquely determined by $\nu_p(r_p, v_p)$?
2. To what extent is $\Phi(r)$ itself determined by $\nu_p(r_p, v_p)$?

The answer to the first question is "yes": we prove this (§ 2) through consideration of the projected velocity moments as functions of r_p, by deriving the equations that relate the intrinsic moments to the observed moments, for every order. Thus, in principle, we are able to "invert" the data to derive the complete distribution function in the assumed potential. The inversion equations for the second and fourth velocity moments have been derived previously (Binney & Mamon 1982; Merrifield & Kent 1990). However, our expressions, even for these low orders, are substantially simpler: we show that the intrinsic velocity moments can always be expressed as a single integral over known quantities. We give explicit expressions for the inversion equations up to order four, and a general formula from which the inversion equations of all orders can be computed.

The second question is trickier. It is clear that, for *certain* $\nu_p(r_p, v_p)$, the potential is uniquely determined. For example, consider a spherical galaxy consisting of only circular orbits. The projected velocity dispersion at $r_p = 0$ will be zero in such a galaxy. Clearly the only orbital distribution that can produce a vanishing central dispersion in projection is one containing only circular orbits; therefore the velocity anisotropy is completely specified, and the radial variation of the circular velocity, and hence the potential, can be inferred uniquely from $\sigma_p^2(r_p)$. Note that, in this case, even the *lowest* moments of $\nu_p(r_p, v_p)$ are sufficient to completely solve the problem: the extra information contained within the full line-of-sight velocity distribution is not needed. As we show below, this is not generally the case, but we might nevertheless expect that the lowest order projected moments will sometimes strongly, perhaps even uniquely, constrain the potential.

Equally clearly, precise knowledge of the form of $\nu_p(r_p, v_p)$ at large v_p must place rather strong constraints on the potential. It is obvious that

$$\nu_p(r_p, v_p) = 0 \quad \text{for} \quad |v_p| \geq \sqrt{-2\Phi(r_p)} , \qquad (5)$$

since stars with higher velocities would escape. This condition is merely the expression of our faith that no significant number of unbound stars can be present in a relaxed stellar system. Hence Dejonghe (1987, hereafter Paper II) suggested that—if one in addition assumes that bound stars fill all of available phase space—the tails of ν_p would directly yield

$$\Phi(r) = -\tfrac{1}{2} v_{p,\max}^2(r) , \qquad (6)$$

with $|v_{p,\max}(r)|$ the extent of the "line profile" at $r_p = r$. This argument suggests the intrinsically unstable nature of a potential function determination: clearly, very high quality data are required to constrain $\Phi(r)$ exactly. Other methods may conceal this unstable nature, but can never fully remedy it.

These fairly general considerations suggest that the projected distribution function contains much information about the potential. We are not able to demonstrate here, however, that $\nu_p(r_p, v_p)$ *uniquely* specifies $\Phi(r)$, nor can we ascertain that such a theorem must exist in the general case. In the absence of such a proof, we pose, and show how to answer, a simpler question:

3. How wide a range of potentials are consistent with a given, *finite* set of velocity moment profiles?

In other words: given the dependence of, say, the projected number density, velocity dispersion, and fourth velocity moment on projected radius, how strongly is the potential constrained? We show (§ 5) that this question can be naturally posed in the language of optimization theory. When the only constraints are the number density and velocity dispersion profiles, the problem reduces to one of linear programming, for which efficient and general algorithms exist. We use such an algorithm to derive the most extreme potentials consistent with a family of velocity dispersion profiles derived from a generalized Plummer model. We find—in agreement with previous studies—that the range of potentials is generally very wide, and in fact that the velocity dispersion profile often contains little more information about the potential than is contained within the virial theorem alone. When the constraints include the fourth-order moments as a function of radius, the problem of constraining the potential becomes a more difficult one of nonlinear optimization, for which completely general algorithms do not exist. Nevertheless we are able to show that even limited information about the fourth moments can substantially reduce the range of allowed potentials.

Our results suggest that rather high moments of the line-of-sight velocity distribution are required to usefully constrain the radial form of the potential in spherical stellar systems. However it is well known that, as a practical matter, sample moments past the second or fourth generally have extremely high variances; furthermore, by converting a two-dimensional data set $\{r_p, v_p\}$ into binned moments, one effectively throws away a great deal of information. It would clearly be useful to develop an algorithm for potential estimation that makes full use of discrete data. The results presented here suggest the form that such an algorithm should take. Since, in an assumed potential, the distribution function $f(E, L^2)$ is uniquely determined by $v_p(r_p, v_p)$, the algorithm should begin by finding the nonnegative f from which the data are "most likely" to have been drawn, in some assumed potential $\Phi(r)$. Next, it should test whether other potentials permit distribution functions with even greater "likelihoods." To the extent that $\Phi(r)$ is uniquely determined (or overdetermined) by $v_p(r_p, v_p)$, the algorithm should be able to find a single, "most likely" potential and distribution function, and this "most likely" solution should approach the correct one as the data set increases in size. One such algorithm is described in a forthcoming paper (Merritt & Saha 1992).

One final point is in order before presenting the calculations. Our conclusion that low-order velocity moments place only very weak constraints on the form of the gravitational potential of a hot stellar system seems, at first sight, to be at odds with much observational work. The ratios of dynamical to luminous mass densities derived from line-of-sight velocity dispersions, using the standard, parametric mass estimators (e.g., the virial theorem, or the core-fitting formula), are remarkably consistent: inferred mass-to-light ratios almost always lie between about 1 and 10, and they depend on metallicity and age roughly as expected. Furthermore, the observed scatter in relations such as the Faber-Jackson law seems no bigger than would be expected on the basis of measurement errors, plus uncertainties about intrinsic shapes. This evidence suggests that velocity dispersions do, indeed, contain considerable information about galactic masses, and that stellar systems are much more uniform in their stellar populations and orbital makeups than they could be. At the same time, it is also true that the detailed form of the distribution function is unknown for any class of hot stellar system, aside from, perhaps, globular clusters (and even here, the preferred form of f is a reflection mostly of theoretical preconceptions, not observational constraints). Furthermore, in disk galaxies, where the orbital geometry is thought to be well understood, the dark matter always has a very different distribution than the luminous matter, and the same may well be true in the outer parts of elliptical galaxies. Thus, our feeling is that the impressive regularity in the kinematical properties of hot stellar systems is not sufficient justification for ignoring the finer details of their phase-space structure and mass distributions—especially since, as this paper attempts to make plausible, those finer details are in principle accessible with the kinds of kinematical data that are now becoming available. Although our conclusions about the usefulness of velocity moments may seem pessimistic, our larger point is an optimistic one: that is should soon be possible to *infer*, rather than simply *model*, the dynamical structure of elliptical galaxies and other hot stellar systems.

2. SOLUTION OF THE JEANS PROBLEM FOR A SPHERICAL STELLAR SYSTEM

2.1. *Preliminaries*

2.1.1. *Augmented Moments*

In order to understand more fully how observed velocity moments may constrain our knowledge of the distribution function $f(E, L^2)$, in a *given* potential $\Phi(r) = -\psi(r)$, it is useful to first look closely at the definition of an *intrinsic* moment

$$\mu_{n,i,j}(r) = \iiint f(E, L^2) v_r^n v_\varphi^i v_\vartheta^j \, dv_r \, dv_\varphi \, dv_\vartheta \tag{7}$$

with the triple integral ranging over all velocities. When this integral is written more explicitly (Dejonghe 1986, hereafter Paper I), one sees that the integral does not really yield the intrinsic moment $\mu_{n,i,j}(r)$, but rather a function of two variables $\tilde{\mu}_{n,i,j}(\psi, r)$. Since we can only know $\mu_{n,i,j}(r) = \tilde{\mu}_{n,i,j}[\psi(r), r]$, the integral equation (7) must be highly degenerate, that is, there must exist many f consistent with a given $\mu_{n,i,j}(r)$. This is indeed the case, and many explicit examples of this degeneracy have been given (e.g., Paper I; Paper II). We call the functions $\tilde{\mu}_{n,i,j}(\psi, r)$ "augmented moments," and the space (ψ, r) in which they assume values we call "augmented configuration space."

It is convenient to define

$$x = r^2 , \tag{8a}$$

$$y = r_p^2 . \tag{8b}$$

The spherical coordinates (r, φ, ϑ) are seen, after projection, as polar coordinates (r_p, θ); because of circular symmetry in the projection, the polar angle θ will nowhere appear explicitly. Since augmented configuration space is a two-dimensional space, we define the partial derivatives

$$\partial_\psi = \frac{\partial}{\partial \psi} ,$$

$$\partial_x = \frac{\partial}{\partial x} = \frac{1}{2r} \frac{\partial}{\partial r} . \tag{9}$$

In configuration space, however, only the ordinary derivative

$$D_x = \frac{d}{dx} = (\partial_x + \psi' \partial_\psi)_{[\psi(x), x]} \tag{10}$$

is defined, with $\psi'(x) = (D_x \psi)(x)$.

It can be shown (Paper I) that

$$\mu_{2n,2i,2j} = \frac{1}{\pi} B\left(i + \frac{1}{2}, j + \frac{1}{2}\right) \mu_{2n,2(i+j)} , \tag{11}$$

with

$$\mu_{2n,2m} = 2\pi \iint f(E, L^2) v_r^{2n} v_T^{2m+1}\, dv_r\, dv_T \tag{12}$$

and $v_T = \sqrt{v_\vartheta^2 + v_\varphi^2}$. We will call these moments the "anisotropic moments," in contrast to the true moments defined in equation (7), and the even simpler "isotropic moments" which one can define in the isotropic case $f(E)$ (Paper I). In our case, the anisotropic moments are necessary and sufficient, and there is no real need to deal with the true moments (7), not even for computing the line-of-sight moments, as we will see. One may also show that

$$\tilde{\mu}_{2n,2m}(\psi, x) = \frac{2^{m+n}}{\sqrt{\pi}} \frac{\Gamma(n+1/2)}{\Gamma(n+m)} \times \int_0^\psi (\psi - \psi')^{m+n-1}\, \partial_x^m (x^m \tilde{\mu}_{0,0})(\psi', x) d\psi' , \tag{13}$$

valid for $m + n \geq 1$. Using equation (13), one proves easily the recursion relations

$$\tilde{\mu}_{2n,2m}(\psi, r) = \frac{\Gamma(n+1/2)}{\Gamma(n+m+1/2)}\, \partial_x^m [x^m \tilde{\mu}_{2(n+m),0}] \tag{14a}$$

and

$$\tilde{\mu}_{2n-2,0}(\psi, r) = \frac{1}{2n-1}\, \partial_\psi(\tilde{\mu}_{2n,0}) . \tag{14b}$$

Any augmented moment in principle suffices to determine f (Paper I). In particular, this is true for the augmented number density $\tilde{\nu} = \tilde{\mu}_{0,0}$.

Equations (14) give the following useful relations between the second-order moments

$$\nu\sigma_r^2 = \mu_{2,0} , \tag{15a}$$

$$\nu\sigma_\varphi^2 = \nu\sigma_\vartheta^2 = \tfrac{1}{2}\mu_{0,2} = \partial_x(x\tilde{\mu}_{2,0}) , \tag{15b}$$

and the fourth-order moments

$$\nu\langle v_r^4 \rangle = \mu_{4,0} , \tag{16a}$$

$$\nu\langle v_r^2 v_\varphi^2 \rangle = \nu\langle v_r^2 v_\vartheta^2 \rangle = \tfrac{1}{2}\mu_{2,2} = \tfrac{1}{3}\,\partial_x(x\tilde{\mu}_{4,0}) , \tag{16b}$$

$$\nu\langle v_\varphi^4 \rangle = \nu\langle v_\vartheta^4 \rangle = 3\nu\langle v_\vartheta^2 v_\varphi^2 \rangle = \tfrac{3}{8}\mu_{0,4} = \tfrac{1}{2}\,\partial_x^2(x^2\tilde{\mu}_{4,0}) . \tag{16c}$$

2.1.2. *The Jeans Equations*

Equations (14) give expressions for all moments of order $2(n + m)$ in terms of the augmented moment $\tilde{\mu}_{2n,0}$; since one augmented moment determines the distribution function completely (Paper I), one should be able to derive the Jeans equations from equations (14).

We will now show that, indeed, equations (14) *are* the Jeans equations. To this end we only need to remember that the Jeans equations relate moments in configuration space, while equations (14) relate moments in augmented configuration space. For example, if we substitute

$$\partial_x = D_x - \psi'\,\partial_\psi$$

in equation (15b),

$$\tfrac{1}{2}\mu_{0,2} = \partial_x(x\tilde{\mu}_{2,0}) = \mu_{2,0} + x\,\partial_x(\tilde{\mu}_{2,0}) ,$$

we obtain

$$\tfrac{1}{2}\mu_{0,2} = \mu_{2,0} + xD_x\mu_{2,0} - x\psi'\nu(x) , \tag{17}$$

where we made use of equation (13),

$$\partial_\psi(\tilde{\mu}_{2,0}) = \tilde{\nu} .$$

After transforming equation (17) to more conventional notations (see eq. [15]), we recover

$$D_r(\nu\sigma_r^2) + \frac{2}{r}\,\nu\left(\sigma_r^2 - \frac{1}{2}\,\sigma_T^2\right) + \nu D_r(\Phi) = 0 , \tag{18}$$

which is the familiar second-order Jeans equation.

As a second example we consider the three nontrivial fourth-order moments. We expect to recover the two nontrivial fourth-order Jeans equations as expressions (16b) and (16c) which give $\mu_{2,2}$ and $\mu_{0,4}$ as a function of $\mu_{4,0}$. Indeed, using equation (16b) we have

$$\mu_{2,2} = \tfrac{2}{3}\,\partial_x(x\tilde{\mu}_{4,0}) ,$$

and, using the same procedure as above, we get

$$\tfrac{2}{3}[D_x(x\mu_{4,0}) - x\psi'\mu_{2,0}] = \mu_{2,2} , \tag{19}$$

or

$$D_r(\nu\langle v_r^4 \rangle) + \frac{\nu}{r}\,(2\langle v_r^4 \rangle - 3\langle v_r^2 v_T^2 \rangle) + 3\nu\sigma_r^2 D_r\Phi = 0 , \tag{20}$$

which is one of the sought equations.

To recover the other Jeans equation we start from equation (16c),

$$\mu_{0,4} = \tfrac{4}{3}\,\partial_x^2(x^2\tilde{\mu}_{4,0})$$

which equals

$$\mu_{0,4} = \tfrac{4}{3}(D_x - \psi'\,\partial_\psi)[x\,\partial_x(x\tilde{\mu}_{4,0}) + x\tilde{\mu}_{4,0}] ,$$

or, with equation (16b),

$$\mu_{0,4} = D_x(2x\mu_{2,2} + \tfrac{4}{3}x\mu_{4,0}) - 4\psi' x(\tfrac{1}{2}\mu_{0,2} + \mu_{2,0}) . \tag{21}$$

Taking account of equation (19), this equation becomes

$$\mu_{0,4} = 4\mu_{2,2} + 2xD_x\mu_{2,2} - 2\psi' x\mu_{0,2} ,$$

or, in more conventional notations,

$$\nu\langle v_T^4 \rangle = 4\nu\langle v_r^2 v_T^2 \rangle + rD_r(\nu\langle v_r^2 v_T^2 \rangle) + r\nu\langle v_T^2 \rangle D_r\Phi . \tag{22}$$

2.1.3. *The Augmented Moments: Alternate Definition*

The augmented moments defined in equation (13) were very close to the ordinary moments (7), in that $\tilde{\mu}_{2n,m}[\psi(r), r] = \mu_{2n,m}(r)$. While we have a clear intuition for the significance of moments up to the second and, possibly, the fourth orders, beyond that, the concept of moments becomes abstract and essentially of technical interest. We can therefore freely redefine the moments in a way that is more suitable for our purposes, without compromising too much of accepted practice. Therefore, we define new augmented moments as

$$\tilde{\omega}_m^\ell(\psi, r) = \partial_x(\tilde{\omega}_{m-1}^\ell) = \partial_x^m(\tilde{\omega}_0^\ell) = \partial_x^m(\tilde{\mu}_{2\ell,0}) , \quad 0 \leq m \leq \ell , \tag{23}$$

and, of course,

$$\omega_m^\ell(r) = \tilde{\omega}_m^\ell[\psi(r), r] .$$

From equation (14) we obtain easily

$$\mu_{2(\ell-m),m} = \frac{\Gamma(\ell - m + 1/2)}{\Gamma(\ell + 1/2)}\, m! \sum_{i=0}^{m} \binom{m}{i} \frac{x^i}{i!}\, \omega_i^\ell . \tag{24}$$

Conversely, (24) is a set of m linear equations in m variables, which is most easily solved by recursion:

$$\omega^{\ell}_m = \frac{1}{x^m}\,\frac{\Gamma(\ell+1/2)}{\Gamma(\ell-m+1/2)}\,\mu_{2(\ell-m),2m} - m!\sum_{i=0}^{m-1}\binom{m}{i}\frac{x^{i-m}}{i!}\,\omega^{\ell}_i\,. \tag{25}$$

In particular

$$\omega^{\ell}_0 = \mu_{2\ell,0}\,, \tag{26}$$

and thus $\omega^0_0 = \mu_{0,0} = \nu$, and

$$\begin{aligned}\omega^1_0 &= \mu_{2,0} = \nu\sigma_r^2\,,\\ \omega^1_1 &= \frac{1}{x}\left(\frac{1}{2}\,\mu_{0,2} - \mu_{2,0}\right) = \frac{1}{x}\,\nu(\sigma_\vartheta^2 - \sigma_r^2)\,,\\ &= -\frac{1}{x}\,\nu\sigma_r^2\beta\,,\end{aligned} \tag{27}$$

which shows a clear connection with Binney's anisotropy parameter $\beta = 1 - \sigma_\vartheta^2/\sigma_r^2$. Also,

$$\begin{aligned}\omega^2_0 &= \mu_{4,0} = \nu\langle v_r^4\rangle\,,\\ \omega^2_1 &= \frac{1}{x}\left(\frac{3}{2}\,\mu_{2,2} - \mu_{4,0}\right) = \frac{1}{x}\,\nu(3\langle v_r^2 v_\vartheta^2\rangle - \langle v_r^4\rangle)\,,\\ \omega^2_2 &= \frac{2}{x^2}\left(\frac{3}{8}\,\mu_{0,4} - 3\mu_{2,2} + \mu_{4,0}\right) = \frac{2}{x^2}\,\nu(\langle v_r^4\rangle\\ &\qquad - 6\langle v_r^2 v_\vartheta^2\rangle + \langle v_\varphi^4\rangle)\,.\end{aligned}$$

We obtained the Jeans equations from the definitions of the moments essentially by replacing a partial derivative in augmented configuration space by a (total) derivative in configuration space. This suggests an alternate set of Jeans equations, in terms of the alternate moments. To obtain these, we simply convert the equations (23) to configuration space:

$$\omega^{\ell}_m(x) = D_x[\omega^{\ell}_{m-1}(x)] - (2\ell-1)\psi'\omega^{\ell-1}_{m-1}(x)\,. \tag{28}$$

This is a simple set of $(\ell - 1)$ equations that (recursively) express $\omega^{\ell}_m(x)$, $m > 0$, in terms of derivatives of ω^{ℓ}_0 and moments of order less than ℓ.

2.2. *Solution of the Jeans Problem*

We now show how to retrieve the distribution function, given the projected moments at all radii.

We start with equation (A10) of Paper II, which gives an expression for an arbitrary moment of the line-of-sight velocity distribution at projected radius r_p and position r along that radius in terms of the anisotropic moments (13):

$$\begin{aligned}\mu_{2n}(r_p, r) &= \frac{1}{\sqrt{\pi}}\,\frac{2^n\Gamma(n+1/2)}{\Gamma(n)}\\ &\quad\times\sum_{i=0}^{n}\binom{n}{i}\frac{r_p^i}{i!}\int_0^{\psi} d\psi'(\psi-\psi')^{n-1}\,\partial^i_{r^2}\,\tilde{\nu}(\psi', r)\,,\end{aligned} \tag{29}$$

which we write, with equations (13) and (8),

$$\mu_{2n}(x, y) = \sum_{i=0}^{n}\binom{n}{i}\frac{y^i}{i!}\,\partial^i_x\tilde{\mu}_{2n,0}(\psi, x)\,. \tag{30}$$

In practice, however, we can only know $\mu_{2n,0}(x)$, and therefore we introduce in equation (30) the total derivative (10). In order to express ∂_x in terms of ordinary derivatives with respect to x, we need an explicit expression for these derivatives. One finds

$$\begin{aligned}\partial^i_x\tilde{\mu}_{2n,0}(\psi, x) &= \partial^i_x\tilde{\omega}^n_0(\psi, x)\\ &= D^i_x\omega^n_0 - \sum_{j=0}^{i-1} D^j_x(\psi'\,\partial_\psi\tilde{\omega}^n_{i-j-1})\,,\end{aligned}$$

or, using equations (14b) and (26):

$$\partial^i_x\tilde{\mu}_{2n,0}(\psi, x) = D^i_x\omega^n_0 - (2n-1)\sum_{j=0}^{i-1} D^j_x(\psi'\omega^{n-1}_{i-j-1})\,. \tag{31}$$

Substituting the last equation into equation (30), we get

$$\begin{aligned}\mu_{2n}(x, y) &= \sum_{i=0}^{n}\binom{n}{i}\frac{y^i}{i!}\,D^i_x\,\mu_{2n,0}(x)\\ &\quad - (2n-1)\sum_{i=0}^{n}\binom{n}{i}\frac{y^i}{i!}\sum_{j=0}^{i-1} D^j_x(\psi'\omega^{n-1}_{i-j-1})\,.\end{aligned}$$

Using

$$\mu^{2n}_p = 2\int_r^{+\infty}\mu_{2n}(r, r_p)\,\frac{r\,dr}{\sqrt{r^2 - r_p^2}} \tag{32a}$$

$$= \int_y^{+\infty}\mu_{2n}(x, y)\,\frac{dx}{\sqrt{x-y}}\,, \tag{32b}$$

we obtain

$$\sum_{i=0}^{n}\binom{n}{i}\frac{y^i}{i!}\int_y^{+\infty} D^i_x\,\mu_{2n,0}(x)\,\frac{dx}{\sqrt{x-y}} = \mu^{2n}_p(y) + Q^{2n}(y) \tag{33}$$

with

$$Q^{2n}(y) = (2n-1)\sum_{k=1}^{n}\sum_{i=k}^{n}\binom{n}{i}\frac{y^i}{i!}\int_y^{+\infty} D^{i-k}_x(\psi'\omega^{n-1}_{k-1})\,\frac{dx}{\sqrt{x-y}}\,. \tag{34}$$

This is an equation for $\mu_{2n,0}$. The right-hand side of equation (33) can be considered completely known, since $Q^{2n}(y)$ is written in terms of all intrinsic moments of order $n - 1$, which are supposed to be the result of a previous iteration. One can easily verify that the left-hand side of equation (33) equals

$$\frac{1}{n!}\,D^n_y\left[y^n\int_y^{+\infty}\omega^n_0(x)\,\frac{dx}{\sqrt{x-y}}\right],$$

while equation (34) can be recast as (see Appendix A)

$$\begin{aligned}Q^{2n}(y) &= (2n-1)D^n_y\sum_{k=1}^{n}\left\{y^{n+k}\int_0^1 du\,\frac{u^k(1-u)^{k-1}}{(k-1)!}\right.\\ &\quad\left.\times\left[\sum_{i=0}^{n-k}\frac{u^i(1-u)^{n-k-i}}{(n-k-i)!(i+k)!}\right]\int_{yu}^{+\infty}(\psi'\omega^{n-1}_{k-1})(x)\,\frac{dx}{\sqrt{x-yu}}\right\}.\end{aligned} \tag{35}$$

Hence equation (33) becomes

$$\begin{aligned}\int_y^{+\infty}\omega^n_0(x)\,\frac{dx}{\sqrt{x-y}} &= n\int_0^1\mu^{2n}_p(uy)(1-u)^{n-1}\,du\\ &\quad + (2n-1)\sum_{k=1}^{n} y^k\int_0^1 du\,\frac{u^k(1-u)^{k-1}}{(k-1)!}\\ &\quad\times\left[\sum_{i=1}^{n-k}\binom{n}{i+k}u^i(1-u)^{n-k-i}\right]\\ &\quad\times\int_{yu}^{+\infty}\psi'\omega^{n-1}_{k-1}\,\frac{dx}{\sqrt{x-yu}}\,.\end{aligned} \tag{36}$$

Finally, we complete the inversion

$$\omega_0^n(x) = -\frac{n}{\pi}\int_0^1 du(1-u)^{n-1}D_x\int_x^{+\infty}\mu_p^{2n}(uy)\frac{dy}{\sqrt{y-x}}$$
$$-\frac{2n-1}{\pi}\sum_{k=1}^{n}\int_0^1 du\,\frac{u^{k-1/2}(1-u)^{k-1}}{(k-1)!}$$
$$\times\left[\sum_{i=0}^{n-k}\binom{n}{i+k}u^i(1-u)^{n-k-i}\right]D_x\int_{xu}^{+\infty}dx'$$
$$\times(\psi'\omega_{k-1}^{n-1})(x')\int_0^1\frac{dv}{\sqrt{v}\sqrt{1-v}}\left[\frac{x'}{u}v+x(1-v)\right]^k, \quad (37)$$

which, upon substituting $v=\sin^2 t$, can also be written as

$$\omega_0^n(x) = D_x\int_0^1 du\left[a_1(u)\int_{xu}^{+\infty}\mu_p^{2n}(y)\frac{dy}{\sqrt{y-xu}}\right.$$
$$+\sum_{k=1}^{n}a_{2,k}(u)\int_{xu}^{+\infty}dx'(\psi'\omega_{k-1}^{n-1})(x')$$
$$\left.\times(xux')^{k/2}P_k\left(\frac{xu+x'}{2\sqrt{xux'}}\right)\right], \quad (38a)$$

with

$$a_1(u) = -\frac{n}{\pi}\frac{(1-u)^{n-1}}{\sqrt{u}},$$
$$a_{2,k}(u) = -(2n-1)\frac{u^{-1/2}(1-u)^{k-1}}{(k-1)!}$$
$$\times\left[\sum_{i=0}^{n-k}\binom{n}{i+k}u^i(1-u)^{n-k-i}\right], \quad (38b)$$

and $P_k(z)$ the Legendre polynomial of degree k. A computationally more convenient representation is

$$\omega_0^n(x) = \int_0^{+\infty}dy\mu_p^{2n}(y)b_1(x,y)$$
$$+\int_0^{+\infty}\psi'(x')dx'\sum_{k=1}^{n}\omega_{k-1}^{n-1}(x')b_{2,k}(x,x'), \quad (39a)$$

with

$$b_1(x,y) = D_x\int_0^{\min(1,y/x)}\frac{a_1(u)du}{\sqrt{y-xu}},$$
$$b_{2,k}(x,x') = D_x\int_0^{\min(1,x'/x)}a_{2,k}(u)(xx'u)^{k/2}P_k\left(\frac{xu+x'}{2\sqrt{xx'u}}\right)du \quad (39b)$$

functions that are independent of the data. Note that the intrinsic moments of arbitrary order are given by *single* integrations of known functions.

Once ω_0^n is determined, we can calculate all order n moments with equation (28), and proceed to the determination of ω_0^{n+1}. The algorithm guarantees that all unnormalized moments will be regular, that is, finite. We are thus led to the theorem:

THEOREM. In a nonrotating, spherical stellar system with known potential, the line-of-sight velocity distribution $v_p(r_p, v_p)$ determines the distribution function $f(E, L^2)$ completely.

Proof. Knowledge of $v_p(r_p, v_p)$ is equivalent to knowledge of all projected moments $\mu_p^{2n}(r_p)$. We have shown that then, in principle, all moments $\omega_0^n(r) = \mu_{2n,0}(r)$ follow. For a fixed r,

$$\mu_{2n,0}(r) = \frac{2^n\Gamma(n+1/2)}{\sqrt{\pi}\Gamma(n)}\int_0^{\psi(r)}(\psi-\psi')^{n-1}\tilde{v}(\psi', r)d\psi',$$

and these functions are known for all n. These are moments with respect to ψ of $\tilde{v}(\psi, r)$ for $0\le\psi\le\psi(r)$. Define the Laplace transform

$$\mathscr{L}(s, r) = \int_0^{\psi(r)}e^{s[\psi(r)-\psi']}\tilde{v}(\psi', r)d\psi'.$$

Then we will have recovered $\tilde{v}(\psi, r)$ if we know $\mathscr{L}(s, r)$. Now

$$D_s^n\mathscr{L}(s, r)|_{s=0} = \mu_{2n,0}(r),$$

and thus

$$\mathscr{L}(s, r) = \sum_{n\ge 0}\mu_{2n,0}(r)\frac{s^n}{n!}.$$

Once we know $\tilde{v}(\psi, r)$, $f(E, L^2)$ follows from formulae such as equations (1.4.28) and (1.4.33) in Paper I.

3. INTEGRAL CONSTRAINTS

Until now, we have allowed complete freedom in the specification of the potential. Based on the arguments given in the Introduction, however, we might expect that certain potentials will lead, by way of the inversion equations just derived, to nonphysical distribution functions. Here we show that the requirement that f be well behaved at large radii imposes an infinite set of integral constraints on the potential, the lowest order of which is the usual virial theorem. The higher order integral constraints are nonlinear functions of Φ; thus, even though they are infinite in number, they need not constrain the potential uniquely. We stress, however, that these integral constraints do not reflect *all* the information about $\Phi(r)$ that is contained within the projected distribution function, since a potential that satisfies the integral constraints may still imply velocity moments that are negative at some radii, and therefore an f that is negative somewhere in phase space. Just as the velocity dispersion profile contains more information about the potential than the virial theorem alone, so the higher order velocity moment profiles contain information that is lost in the higher order integral constraints. Furthermore, as we show, exact compliance with the virial theorems is required only when attempting to infer the intrinsic properties of a stellar system at large radii; the central properties are only weakly affected by imposition of the virial constraints.

3.1. *The Second-Order Equations*

Binney & Mamon (1982) first showed how to obtain the two functions $\sigma_r^2(r)$ and $\sigma_\vartheta^2(r)$ from the single function $\sigma_p^2(r_p)$ by requiring that the solution so obtained satisfy the Jeans equation in an assumed potential $\Phi(r)$. Their deconvolution equation is equivalent to equation (39) with $n=1$, with one important difference: while the formula derived here is guaranteed—because of a judicious choice of integration constant—to give a well-behaved solution at $r = 0$, their equations imply a divergent central pressure unless the assumed potential is in precise virial equilibrium with the line-of-sight velocity dispersions. Binney & Mamon (and later Merrifield & Kent 1990, in their discussion of the fourth moments) interpreted this pathological behavior as a *consequence* of noncompliance with the virial theorem. Their interpretation is disturbing, since it implies that the inferred, central properties of a stellar system are strongly dependent on the behavior of Φ and $\nu_p \sigma_p^2$ at points far from the center. We show here that a different interpretation is possible. If the adopted potential is "virially inconsistent" with the observed velocity dispersion profile, our formulae imply that the intrinsic velocity moments $\sigma_r^2(r)$ and $\sigma_\vartheta^2(r)$—while remaining well behaved at the origin—are ill behaved at *large* radii. Thus, noncompliance with the virial theorem means only that the inferred solution will be poorly behaved at radii where the data are themselves, typically ill determined; the inferred, central values of $\sigma_r^2(r)$ and $\sigma_\vartheta^2(r)$ will be nearly the same as their values in a model that is in precise virial equilibrium. This distinction may seem a trivial one when dealing with the second moments, since compliance with the second-order virial theorem can always be assured by adjusting a single normalizing constant in the potential, for example, the total mass. The advantage of our formulation becomes clear when dealing with the higher order moments: imposition of the "virial" constraints corresponding to velocity moments of order $2n$ would require that $\Phi(r)$ simultaneously satisfy $n(n+1)/2$ integral conditions—an awkward state of affairs unless the potential is characterized by a very large number of free parameters. Our treatment demonstrates that knowledge of the projected properties of a stellar system at large radii is only necessary when attempting to constrain the potential or distribution function at large radii.

We begin by rederiving the second-order equations. We first relate the surface density to the intrinsic number density:

$$\nu_p(r_p) = 2 \int_r^{+\infty} \nu(r) \frac{r\,dr}{\sqrt{r^2 - r_p^2}} . \tag{40a}$$

Introducing equation (8), we get

$$\nu_p(y) = \int_y^{+\infty} \nu(x) \frac{dx}{\sqrt{x - y}} , \tag{40b}$$

enabling us to obtain $\nu(x)$ with

$$\nu(x) = -\frac{1}{\pi} \int_x^{+\infty} D_y(\nu_p) \frac{dy}{\sqrt{y - x}} . \tag{41}$$

Second, the observed velocity dispersion is related to ν and the spatial velocity dispersion σ_r^2 and $\sigma_\varphi^2 = \sigma_\vartheta^2 = [1 - \beta(r)]\sigma_r^2$ through

$$\frac{1}{2} \nu_p \sigma_p^2(r_p) = \int_{r_p}^{+\infty} \nu(r)\sigma_r^2(r) \frac{r\,dr}{\sqrt{r^2 - r_p^2}} - r_p^2 \int_{r_p}^{+\infty} \nu(r)\sigma_r^2(r)\beta(r) \frac{dr}{r\sqrt{r^2 - r_p^2}} . \tag{42}$$

This equation has a known left-hand side, but the right-hand side still contains two unknown functions. This indeterminacy can be eliminated by the Jeans equation (18), which is our third equation. Combining equations (42) and (18), one finds

$$\frac{1}{2} \nu_p \sigma_p^2 - \frac{1}{2} r_p^2 \int_{r_p}^{+\infty} \nu D_r \psi \frac{dr}{\sqrt{r^2 - r_p^2}} = \int_{r_p}^{+\infty} \nu\sigma_r^2 \frac{r\,dr}{\sqrt{r^2 - r_p^2}} + \frac{1}{2} r_p^2 \int_{r_p}^{+\infty} D_r(\nu\sigma_r^2) \frac{dr}{\sqrt{r^2 - r_p^2}} , \tag{43}$$

which must be solved for $\sigma_r^2(r)$.

We denote

$$g(y) = (\nu_p \sigma_p^2)(y) - y \int_{\sqrt{y}}^{+\infty} \nu D_r \psi \frac{dr}{\sqrt{r^2 - y}} , \tag{44}$$

a known function. Equation (43) then becomes

$$g(y) = \int_y^{+\infty} \nu\sigma_r^2 \frac{dx}{\sqrt{x - y}} + y \int_y^{+\infty} D_z(\nu\sigma_r^2) \frac{dx}{\sqrt{x - y}} . \tag{45}$$

The key point here is to note that this can be written simply as

$$g(y) = D_y\left(y \int_y^{+\infty} \nu\sigma_r^2 \frac{dx}{\sqrt{x - y}} \right) , \tag{46}$$

which can be readily integrated:

$$\int_{y_0}^{y} g(y')dy' = y \int_y^{+\infty} \nu\sigma_r^2 \frac{dx}{\sqrt{x - y}} , \tag{47}$$

where y_0 is a constant of integration. Regularity of the solution $\nu\sigma_r^2$ at $r = 0$ dictates $y_0 = 0$, and we define $G(y) = \int_0^y g(y')dy'$. Inverting equation (47) then gives

$$(\nu\sigma_r^2)(x) = -\frac{1}{\pi} \int_x^{+\infty} D_y\left[\frac{G(y)}{y}\right] \frac{dy}{\sqrt{y - x}} . \tag{48}$$

A more common way to write this is

$$\nu(r)\sigma_r^2(r) = \frac{G(\infty)}{2r^3} - \frac{2}{\pi r^3} \int_r^{\infty} \left[\frac{r}{\sqrt{r_p^2 - r^2}} + \cos^{-1}\left(\frac{r}{r_p}\right)\right] g(r_p) r_p\,dr_p$$

or, substituting equation (44) for $g(r_p)$ and simplifying further,

$$\nu(r)\sigma_r^2(r) = \frac{G(\infty)}{2r^3} - \frac{2}{\pi r^3} \int_r^{\infty} \left[\frac{r}{\sqrt{r_p^2 - r^2}} + \cos^{-1}\left(\frac{r}{r_p}\right)\right] \times \nu_p(r_p)\sigma_p^2(r_p) r_p\,dr_p + \frac{2}{3r^3} \int_r^{\infty} \left(r'^3 + \frac{r^3}{2}\right) \nu(r') \frac{d\Phi}{dr'}\,dr' . \tag{49a}$$

Using the Jeans equation (18), we find the following for the tangential dispersion:

$$\nu(r)\sigma_\vartheta^2(r) = -\frac{G(\infty)}{4r^3} - \frac{1}{2\pi r^3} \int_r^{\infty} \left[\frac{r(r_p^2 + r^2)}{r_p^2\sqrt{r_p^2 - r^2}} + \cos^{-1}\left(\frac{r}{r_p}\right)\right] \times \frac{d(\nu_p \sigma_p^2)}{dr_p} r_p^2\,dr_p - \frac{1}{3r^3} \int_r^{\infty} (r'^3 - r^3)\nu(r') \frac{d\Phi}{dr'}\,dr' . \tag{49b}$$

Equations (49) are *mathematically* acceptable solutions for $\nu\sigma_r^2$ and $\nu\sigma_\vartheta^2$. because, by construction, the integrand in equation (48) is always well behaved: although it may not be obvious from equations (49), there is no further condition to be imposed in order that $\nu\sigma_r^2$ and $\nu\sigma_\vartheta^2$ be finite for $r = 0$. This is a consequence of the nature of equation (46) featuring a differential operator, which introduces an arbitrary constant in the solution; this constant was chosen such that the integrand in equation (48) always behaves. Hence, from a purely mathematical point of view, the solution for the "pressures" $\nu\sigma_r^2$ and $\nu\sigma_\vartheta^2$ is satisfactory since it is nonsingular everywhere.

However as $r \to \infty$ (or equivalently, at the edge of a finite system), these equations imply

$$\nu\sigma_r^2 \to \frac{G(\infty)}{2r^3}\,, \quad \nu\sigma_\vartheta^2 \to -\frac{G(\infty)}{4r^3}\,.$$

In other words, unless $G(\infty) = 0$, then either σ_r^2 or σ_ϑ^2 must become negative at large radii. Furthermore, for $r \to \infty$, the "pressure" $\nu\sigma_r^2$ behaves as r^{-3}. Therefore, if ν tends to zero sufficiently fast, σ_r^2 will blow up for $r \to \infty$, in spite of a behaving $\nu\sigma_r^2$. On the other hand, at any point along the line of sight, the line profile has locally, for physical reasons, a finite variance. That variance is a linear combination of σ_r^2 and σ_ϑ^2, and thus σ_ϑ^2 must tend to $-\infty$ at large radii—as indeed implied by equation (49b). Thus a physically permissible solution must have

$$G(\infty) = \int_0^\infty g(y)dy = 0\,, \tag{50}$$

or, using equation (44),

$$3\int_0^{+\infty} \nu_p\sigma_p^2 r_p\, dr_p = 2\int_0^{+\infty} \nu(r)\frac{d\Phi}{dr} r^3\, dr\,, \tag{51}$$

which is the virial theorem (4). We conclude that singular dispersions (not to be confused with pressures) at large radii are sufficient to reject the solution on the basis of negativity of the distribution function. The virial theorem may therefore be interpreted—at least in this context—as a consequence of the positivity and regularity of f. This is consistent with the fact that the virial equation can be obtained by integrating the distribution function, which is only a valid operation if that distribution function is sufficiently regular. By the same token, a similar argument holds for any singular higher order moment. We will use this result in the next subsection to derive integral constraints from the higher order inversion equations.

Our recovery of the virial theorem is hardly unexpected: it is well known that one may derive the virial theorem by integrating the spherical Jeans equation (18), multiplied by $4\pi r^3$, from zero to infinity,

$$\langle v^2\rangle = (\boldsymbol{r}\cdot\boldsymbol{\nabla}\Phi) + 4\pi(r^3\nu\sigma_r^2)_0^\infty\,, \tag{52}$$

and requiring that $r^3\nu\sigma_r^2$ vanish at large and small radii. The new point to be made here is that, even if the adopted potential does not satisfy the virial theorem exactly, the inferred kinematical solution can still be well behaved at the origin. In other words, the *central* structure of a model need not be strongly dependent on the behavior of $\nu_p\sigma_p^2$ or Φ at large radii—a fortunate state of affairs, since the behavior of these quantities at large radii is typically poorly determined. By contrast, Binney & Mamon (1982), in their equation (12a), effectively chose for their constant of integration $y_0 = \infty$, rather than $y_0 = 0$, leading them to an expression for $\nu\sigma_r^2$ that is equivalent to our equation (49a) without the term $G(\infty)/2r^3$. Thus, their algorithm gives a divergent $\nu\sigma_r^2$ at $r = 0$ unless the velocity dispersion profile is precisely "consistent," in a virial sense, with the assumed potential. In our more general formulae, a potential that is not "virially consistent" with a given velocity dispersion profile will lead to a solution that begins to behave badly only outside of some radius, beyond which—hopefully—the data are poorly determined as well. In spite of this slight (but important) difference, our inversion equations are identical to theirs after imposition of the virial constraint.

A final remark is in order on the interpretation of the condition (51). Generally, the virial theorem is seen as a scaling condition on Φ. However, it might equally well be seen as a condition on the *assumption* that the distribution function depends on E and L only. To see this, it is instructive to look at equation (42) again, which is the integral equation containing all the observational constraints. It is clear that, if we had not set $\sigma_\varphi^2 = \sigma_\vartheta^2$, even the Jeans equations

$$D_r\nu\sigma_r^2 + \frac{\nu}{r}(2\sigma_r^2 - \sigma_\varphi^2 - \sigma_\vartheta^2) - \nu D_r\psi = 0$$

would not have been sufficient to close the equations and determine the dispersions. Yet, a scaling condition such as equation (51) would always have to be imposed. This follows from noting that, even if $\sigma_\varphi^2 \neq \sigma_\vartheta^2$, there always will exist a function $f(x)$ such that $\sigma_\varphi^2 = f\sigma_\vartheta^2$. If we now first assume such a $f(x)$, the analysis of Binney & Mamon (1982) could be repeated and would again lead to a condition serving the same purpose as equation (51). We conclude that, although a condition such as equation (51) will always exist in any algorithm that produces a unique solution for the intrinsic pressures, its form may depend on the a priori conditions and symmetries that we assume or impose. Hence, equation (51) is a condition on the two-integral assumption for the distribution function *as well as* a scaling condition on Φ, which, in this case, depends on the assumed phase space structure very weakly, through the density $\nu(r)$. In the next subsection, we will see that the dependence on phase-space structure will grow stronger as we go to higher orders.

3.2. *The Higher Order Constraints*

In spite of the fact that the solution (39) for $\omega_0^n(x)$ will be regular everywhere, the intrinsic normalized nth order moment will not, in general, have an acceptable behavior for $x \to +\infty$. This can be most easily seen by considering equation (36), the left-hand side of which tends to zero as fast as $r\nu\psi$. The right-hand side however, that we write as $R + S$, with

$$R = \frac{n}{y}\int_0^y \mu_p^{2n}(t)\left(1 - \frac{t}{y}\right)^{n-1} dt \tag{53a}$$

$$S = \frac{(2n-1)}{y}\sum_{k=1}^{n}\int_0^y dt\,\frac{t^k(1-t/y)^{k-1}}{(k-1)!} \times\left[\sum_{i=0}^{n-k}\binom{n}{i+k}\left(\frac{t}{y}\right)^i\left(1-\frac{t}{y}\right)^{n-k-i}\right] \times\int_t^{+\infty}\psi'\omega_{k-1}^{n-1}\frac{dx}{\sqrt{x-t}}\,, \tag{53b}$$

exhibits, for $y \to +\infty$, polynomial behavior. In particular, we can easily see that R and S are, in the limit, polynomials of order n in $1/y$ of the form $R + S = \sum_{i=1}^n (c_i + d_i)y^{-i}$. Hence, the polynomials $R + S$ must be identically zero in general, thereby imposing n additional constraints for every determined intrinsic moment ω_0^n. In Appendix B we show that these constraints are

$$\lim_{y\to+\infty}\left[\int_0^1 \mu_p^{2n}(yu)u^m\,du + (2n-1)\sqrt{\pi}\sum_{k=1}^{n} a_{k,m}^n \times \int_0^1 \psi'(yu)\omega_{k-1}^{n-1}(yu)u^m(yu)^{k+1/2}\,du\right] = 0\,, \quad 0 \le m \le n-1\,, \quad (54a)$$

with

$$a_{k,m}^n = \binom{n-1}{k-1}\frac{(k+m)!}{k!\Gamma(k+m+3/2)} \times {}_3F_2(k-n, -m, 1; k+1, 1-n; 1)\,. \quad (54b)$$

The integrals in equation (54a) converge under quite generous conditions for the projected and intrinsic moments. Stronger yet, the limits will produce multiple zeros for both terms, and the conditions must be seen as conditions on the asymptotic behavior as $y \to +\infty$. We can get rid of the multiple zeros by multiplying equation (54a) by y^{m+1}, thus obtaining a stronger version of these constraints, only valid in case the integrals exist:

$$\int_0^{+\infty}\mu_p^{2n}(y)y^m\,dy + (2n-1)\sqrt{\pi} \times \sum_{k=1}^{n} a_{k,m}^n \int_0^{+\infty}\psi'\omega_{k-1}^{n-1}x^{k+m+1/2}\,dx = 0\,, \quad 0 \le m \le n-1\,. \quad (55)$$

As Kent (1991) has noted, the constraints (55) can also be obtained by taking spatial moments of the Jeans equations. Within our formalism, that derivation is straightforward. We begin by calculating the integral on the sky of a projected $2n$th order moment

$$\int \mu^{2n}(y)y^m\,dA = \pi\int_0^{+\infty}\mu^{2n}(y)y^m\,dy = \pi\int_0^{+\infty}y^m\,dy\int_y^{+\infty}\mu_{2n}(x, y)\frac{dx}{\sqrt{x-y}}\,,$$

according to equation (32). After substituting equation (30) and integrating over y, we obtain

$$\int_0^{+\infty}\mu^{2n}(y)y^m\,dy = \sqrt{\pi}\sum_{i=0}^{n}\binom{n}{i} \times \frac{(m+i)!}{i!\Gamma(m+i+3/2)}\int_0^{+\infty}\omega_i^n(x)x^{m+k+1/2}\,dx\,.$$

We now substitute equation (31) and perform a number of partial integrations, with the result

$$\int_0^{+\infty}\mu^{2n}(y)y^m\,dy = \sqrt{\pi}b_{0,m}^n\int_0^{+\infty}x^{m+1/2}\omega_0^n\,dx - (2n+1)\sqrt{\pi}\sum_{k=1}^{n} b_{k,m}^n\int_0^{+\infty}\psi'(x)\omega_{k-1}^{n-1}(x)x^{k+m+1/2}\,dx\,,$$

and

$$b_{k,m}^n = \frac{(-)^k}{\Gamma(m+k+3/2)}\sum_{i=k}^{n}\binom{n}{i}\frac{(-)^i(m+i)!}{i!}\,.$$

It turns out that

$$b_{0,m}^n = \frac{m!}{\Gamma(m+3/2)}\,{}_2F_1(-n, 1+m; 1; 1) = 0\,, \quad m < n\,,$$

while

$$b_{k,m}^n = \frac{(k+m)!}{k!\Gamma(m+k+3/2)}\binom{n}{k} \times {}_3F_2(k-n, k+m+1, 1; k+1, k+1; 1)\,, \quad 1 \le l \le n\,.$$

We will therefore have recovered equation (55) if we can prove that $a_{k,m}^n = b_{k,m}^n$, or

$$k\,{}_3F_2(k-n, -m, 1; k+1, 1-n; 1) = n\,{}_3F_2(k-n, k+m+1, 1; k+1, k+1; 1)\,.$$

This is indeed an apparent identity upon writing it as

$$k\int_0^1 u^{k-1}\,{}_2F_1(k-n, -m; 1-n; 1-u)du = n\int_0^1 u^{k-1}\,{}_2F_1(k-n, k+m+1; k+1; 1-u)du = n\int_0^1 u^{n-1}\,{}_2F_1(k-n, -m; k+1; 1-1/u)du$$

and comparing the term in $(-m)_i$.

Since we will need them below, we give here the integral constraints for orders two and four:

$$\int_0^\infty \mu_p^2 r_p\,dr_p = \frac{2}{3}\int_0^\infty \nu(r)\frac{d\Phi}{dr}r^3\,dr\,, \quad (56a)$$

$$\int_0^\infty \mu_p^4 r_p\,dr_p = \frac{2}{5}\int_0^\infty \nu(r)(3\sigma_r^2 + 2\sigma_\vartheta^2)\frac{d\Phi}{dr}r^3\,dr\,, \quad (56b)$$

$$\int_0^\infty \mu_p^4 r_p^3\,dr_p = \frac{4}{35}\int_0^\infty \nu(r)(\sigma_r^2 + 6\sigma_\vartheta^2)\frac{d\Phi}{dr}r^5\,dr \quad (56c)$$

3.3. *Significance of the Virial Constraints*

We have argued that the virial constraints are best interpreted as conditions on the potential such that the inferred kinematics of a stellar system are physically reasonable at large radii. In practice, the large-radius structure of a stellar system is generally poorly constrained observationally, because of finite sample sizes or uncertain background corrections; thus, noncompliance with the virial constraints leads to uncertainties in the kinematics only at radii where the data are also uncertain. Nevertheless, the virial constraints—to the degree that the terms within them depending on line-of-sight velocities can be accurately computed from limited data—do, themselves, constrain the form of the potential. Merrifield & Kent (1990) went so far as to speculate that the infinite set of virial constraints would be formally sufficient to fully determine $\Phi(r)$, even in the absence of any additional information about the projected distribution function. We can neither prove nor disprove their conjecture here, since the virial constraints are nonlinear functions of Φ for $n > 1$. However, as we have shown, these constraints are simply one piece of the larger problem: while they imply good behavior of f at large radii, they say nothing about its positivity at general points in phase space. Furthermore, converting all available velocity information to global moments is an inefficient way to deal with limited data, since, in the process, one loses considerable information about the dependence of projected distribution functions on r_p and v_p. In a practical sense, then, we imagine that there must exist more efficient ways of inferring gravitational potentials from limited data than via the virial constraints.

4. INVERSION EQUATIONS FOR ORDERS 2 AND 4

Here we present explicit expressions for the intrinsic moments in terms of the observed moments, for orders two and four. We begin by assuming that the virial constraints of corresponding order are exactly satisfied. The two second-order equations are then

$$\nu(r)\sigma_r^2(r) = -\frac{2}{\pi r^3}\int_r^\infty \left[\frac{r}{\sqrt{r_p^2 - r^2}} + \cos^{-1}\left(\frac{r}{r_p}\right)\right]\nu_p(r_p)\sigma_p^2(r_p)r_p\,dr_p + \frac{2}{3r^3}\int_r^\infty \left(r'^3 + \frac{1}{2}r^3\right)\nu(r')\frac{d\Phi}{dr'}dr' , \quad (57a)$$

$$\nu(r)\sigma_\vartheta^2(r) = -\frac{1}{2\pi r^3}\int_r^\infty \left[\frac{r(r_p^2 + r^2)}{r_p^2\sqrt{r_p^2 - r^2}} + \cos^{-1}\left(\frac{r}{r_p}\right)\right] \times \frac{d(\nu_p\sigma_p^2)}{dr_p}r_p^2\,dr_p - \frac{1}{3r^3}\int_r^\infty (r'^3 - r^3)\nu(r')\frac{d\Phi}{dr'}dr' , \quad (57b)$$

and the three fourth-order equations are

$$\nu(r)\langle v_r^4(r)\rangle = \frac{2}{\pi r^5} \times \int_r^\infty \left[3r\sqrt{r_p^2 - r^2} + (3r_p^2 - 2r^2)\cos^{-1}\left(\frac{r}{r_p}\right)\right] \times \mu_p^4(r_p)r_p\,dr_p + \frac{1}{35r^5} \times \int_r^\infty \left[\sigma_r^2(r')\left(-9\frac{r^7}{r'^2} + 42r^5 + 84r^2r'^3 - 12r'^5\right) + \sigma_\vartheta^2(r')\left(9\frac{r^7}{r'^2} + 7r^5 + 56r^2r'^3 - 72r'^5\right)\right]\nu(r')\frac{d\Phi}{dr'}dr' , \quad (58a)$$

$$\nu(r)\langle v_r^2(r)v_\vartheta^2(r)\rangle = -\frac{1}{3\pi r^5} \times \int_r^\infty \left[\frac{r(9r_p^2 - 5r^2)}{\sqrt{r_p^2 - r^2}} + (9r_p^2 - 2r^2)\cos^{-1}\left(\frac{r}{r_p}\right)\right] \times \mu_p^4(r_p)r_p\,dr_p + \frac{1}{105r^5} \times \int_r^\infty \left[\sigma_r^2(r')\left(-18\frac{r^7}{r'^2} + 42r^5 - 42r^2r'^3 + 18r'^5\right) + \sigma_\vartheta^2(r')\left(18\frac{r^7}{r'^2} + 7r^5 - 28r^2r'^3 + 108r'^5\right)\right] \times \nu(r')\frac{d\Phi}{dr'}dr' , \quad (58b)$$

$$\nu(r)\langle v_\vartheta^4(r)\rangle = -\frac{1}{16\pi r^5} \times \int_r^\infty \left[\frac{r(9r_p^4 + 6r^4 + r^2r_p^2)}{r_p^2\sqrt{r_p^2 - r^2}} + (9r_p^2 + 4r^2)\cos^{-1}\left(\frac{r}{r_p}\right)\right] \times \frac{d\mu_p^4}{dr_p}r_p^2\,dr_p + \frac{1}{70r^5} \times \int_r^\infty \left[\sigma_r^2(r')\left(-54\frac{r^7}{r'^2} + 84r^5 - 21r^2r'^3 - 9r'^5\right) + \sigma_\vartheta^2(r') \times \left(54\frac{r^7}{r'^2} + 14r^5 - 14r^2r'^3 - 54r'^5\right)\right]\nu(r')\frac{d\Phi}{dr'}dr' . \quad (58c)$$

An equation equivalent to equation (57a) was first derived by Binney & Mamon (1982); as previously noted, our expression is simpler, involving only a single integration. Equation (58a) for $\langle v_r^4(r)\rangle$ was first given, in a different form, by Merrifield & Kent (1990); again, our expression contains one fewer integration. Equations (57b), (58b), and (58c) are presented here for the first time. We checked that each expression gives the correct, intrinsic moments for the family of anisotropic Plummer models of Paper II.

If the adopted potential does not satisfy the second- or fourth-order virial constraints, then the solutions given above are not correct. To the two, second-order equations must be added

$$\nu(r)\sigma_r^2(r) \to \nu(r)\sigma_r^2(r) + \frac{G(\infty)}{2r^3} , \quad (59a)$$

$$\nu(r)\sigma_\vartheta^2(r) \to \nu(r)\sigma_\vartheta^2(r) - \frac{G(\infty)}{4r^3} , \quad (59b)$$

where

$$G(\infty) = 2\int_0^\infty \nu_p(r_p)\sigma_p^2(r_p)r_p\,dr_p - \frac{4}{3}\int_0^\infty \nu(r')\frac{d\Phi}{dr'}r'^3\,dr' . \quad (60)$$

To the three fourth-order equations must be added

$$\nu(r)\langle v_r^4(r)\rangle \to \nu(r)\langle v_r^4(r)\rangle + \frac{H_1(\infty)}{r^3} - \frac{H_2(\infty)}{r^5} , \quad (61a)$$

$$\nu(r)\langle v_r^2(r)v_\vartheta^2(r)\rangle \to \nu(r)\langle v_r^2(r)v_\vartheta^2(r)\rangle - \frac{H_1(\infty)}{6r^3} + \frac{H_2(\infty)}{2r^5} , \quad (61b)$$

$$\nu(r)\langle v_\vartheta^4(r)\rangle \to \nu(r)\langle v_\vartheta^4(r)\rangle - \frac{H_1(\infty)}{8r^3} - \frac{3H_2(\infty)}{8r^5} , \quad (61c)$$

where

$$H_1(\infty) = 2\int_0^\infty \mu_p^4(r_p)r_p\,dr_p - \frac{4}{5}\int_0^\infty (3\sigma_r^2 + 2\sigma_\vartheta^2)\nu(r')\frac{d\Phi}{dr}r'^3\,dr' , \quad (62a)$$

$$H_2(\infty) = 3\int_0^\infty \mu_p^4(r_p)r_p^3\,dr_p - \frac{12}{35}\int_0^\infty (\sigma_r^2 + 6\sigma_\vartheta^2)\nu(r')\frac{d\Phi}{dr'}r'^5\,dr' . \quad (62b)$$

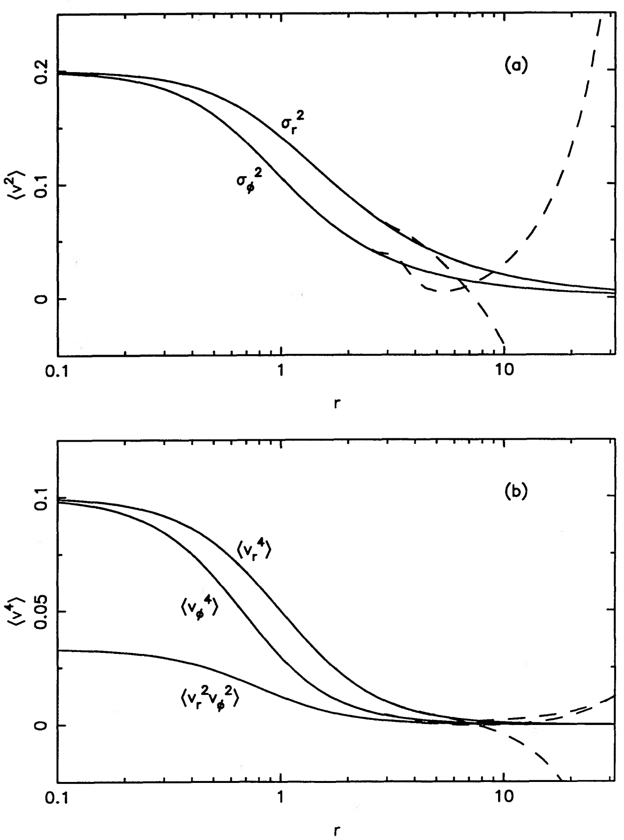

FIG. 1.—Intrinsic velocity moments derived from modified, Plummer-model profiles, as described in the text.

As noted above, these formula yield solutions that are regular at $r = 0$, but that behave unphysically beyond some radius r whose value depends on the degree to which the virial constraints are violated.

As an illustration of the application of these equations, we computed the intrinsic moments corresponding to a set of line-of-sight moment profiles, with an assumed potential that did not satisfy the virial constraints. For $\sigma_p^2(r_p)$ we took the anisotropic Plummer model profile of equation (73), with anisotropy $q = 1$, modified by the factor $[1 + (r_p - r_0)^2]^{-1}$ for $r_p > r_0 = 3$. The potential was taken to be the Plummer potential (74). Figure 1*a* shows that the inferred intrinsic moments have nearly the correct values at radii $r \lesssim r_0$ and behave unacceptably only at much larger radii. Similarly, Figure 1*b* shows the behavior of the intrinsic, fourth-moment profiles when $\mu_p^4(r_p)$ is modified by the same factor. Again, the unphysical behavior is limited to radii well outside of the region where the line-of-sight moments deviate from the "correct" values. We conclude that—as a practical matter—no very severe penalty is to be paid for noncompliance with the virial theorems.

5. USING VELOCITY MOMENTS TO CONSTRAIN POTENTIALS

We have seen that the problem of determining the potential of a spherical stellar system from the lower velocity moments of a set of test particles is generally an underconstrained one. This means that—in the absence of preconceptions about the proper form of Φ or f—it almost never makes sense to search for a single, "best-fit" potential; rather, one should seek an *optimal* potential, that is, a potential that maximizes or minimizes (subject to the observational constraints) some "objective function" chosen to represent a physically interesting quantity, for example, the total mass or the central density. In the past, the standard approach to this problem has been to note that equations like equation (3), which relate line-of-sight velocity moments to the underlying distribution function f, are linear in f. One can therefore test the suitability of an assumed potential $\Phi(r)$ by using a linear algorithm to construct an $f(E, L^2)$ that is nonnegative and that reproduces the observed moments (within the observational errors) in the assumed Φ (e.g., Richstone & Tremaine 1984; Newton & Binney 1984; Dejonghe 1989). By trying many different forms for $\Phi(r)$, one can eventually put limits on the mass distribution. The major difficulty with this approach is obvious: because it focuses on the construction of f rather than Φ, it does not lend itself to an efficient search for the full range of potentials consistent with the data.

An alternative approach is suggested by the inversion equations (57) and (58) derived above. Note, for instance, that the intrinsic velocity dispersions implied by a given $\sigma_p^2(r)$ are related to the assumed potential through a linear equation. If we express this potential as a sum of basis functions with unspecified coefficients,

$$\Phi(r) = \sum_i a_i \Phi_i(r) , \tag{63}$$

the a_i can be determined through a linear programming algorithm, where the constraints include positivity of the intrinsic moments $\nu\sigma_r^2$ and $\nu\sigma_\vartheta^2$ at every radius. The objective function need only be linear in the a_i; such functions include the total mass, the central or mean density, the central value of the potential, etc. If we require the potential to be consistent with higher order observed moments, say of order $2n$, then the equations relating the intrinsic moments to the potential are of order n in the a_i. Linear programming will not work in this case; however, in the last few years, a number of reasonably general and efficient algorithms for nonlinear optimization have been published (e.g., Schittkowski 1980). Of course, any technique based only on velocity moments will not yield a complete distribution function, and there is no guarantee that the derived intrinsic moments, even if everywhere positive, could be derived from a fully nonnegative f. On the other hand, it is our purpose here to explore the power of the constraints imposed by positive intrinsic moments, and it is hard to see how any of the existing methods that yield f as well can be made to fit logically in the scheme worked out in the previous paragraphs.

In the absence of any other constraints, the "optimized" potential that results from such an algorithm is often strongly unphysical, in the sense that the corresponding mass density,

$$\rho = \frac{1}{4\pi G} \nabla^2 \Phi , \tag{64}$$

is negative, or oscillates with radius. It therefore makes sense to add to the algorithm the (linear) constraints

$$\sum_i a_i \rho_i(r) > 0 , \quad \sum_i a_i \frac{d\rho_i}{dr} < 0 \tag{65}$$

at all r, where $\rho_i = \nabla^2 \Phi_i / 4\pi G$.

The behavior of the algorithm also depends strongly on the choice for the objective function to be maximized or minimized. For instance, telling the routine to find the solution with the maximum total mass usually produces a poor result, since arbitrarily large amounts of matter at large radii have almost no effect on the internal dynamics. Similarly, the maximum central mass density is poorly defined, since a central singularity of insignificant mass affects the dynamics only very close to the center. Two well-defined, and astrophysically interesting, choices for the objective function are the *minimum* total mass, and the *minimum* central density. Minimizing either quantity gives (as we show below) a well-defined solution. Furthermore, both quantities can be derived simply if the only constraint on the potential is the virial theorem (Merritt 1987). The minimum mass in this case is

$$GM_{\min} = \frac{3}{4}\frac{\int_0^\infty \nu_p(r_p)\sigma_p^2(r_p)r_p\,dr_p}{\int_0^\infty \nu(r)r\,dr}, \tag{66}$$

and corresponds to a model in which all the matter is concentrated in a central point. The minimum density (under the assumption $d\rho/dr < 0$) is

$$G\rho_{\min} = \frac{9}{16\pi}\frac{\int_0^\infty \nu_p(r_p)\sigma_p^2(r_p)r_p\,dr_p}{\int_0^\infty \nu(r)r^4\,dr}, \tag{67}$$

and corresponds to a model in which the matter is uniformly distributed. Adding additional constraints beyond the virial theorem usually gives "minimum mass" and "minimum density" solutions that are similar in character (though less extreme) than these.

5.1. *Second Moments*

Suppose that the data comprise only the projected number density and velocity dispersion profiles, with nothing known about the higher moments. If we assume that the potential satisfies the virial theorem, that is, that

$$\frac{3}{2}\int_0^\infty \nu_p\sigma_p^2 r_p\,dr_p = \int_0^\infty \nu\frac{d\Phi}{dr}r^3\,dr = \sum_i a_i\int_0^\infty \nu\frac{d\Phi_i}{dr}r^3\,dr, \tag{68}$$

then the radial and tangential components of the velocity dispersion are given by equations (57), which may be written

$$\nu(r)\sigma_r^2(r) = \frac{1}{r^3}\left[-I(r) + \sum_i a_i A_i(r)\right], \tag{69a}$$

$$\nu(r)\sigma_t^2(r) = \frac{1}{r^3}\left[-J(r) + \sum_i a_i B_i(r)\right]. \tag{69b}$$

The functions $I(r)$ and $J(r)$ contain the kinematical information

$$I(r) = \frac{2}{\pi}\int_r^\infty\left[\frac{r}{\sqrt{r_p^2 - r^2}} + \cos^{-1}\left(\frac{r}{r_p}\right)\right]\nu_p(r_p)\sigma_p^2(r_p)r_p\,dr_p, \tag{70a}$$

$$J(r) = -\frac{1}{2}I(r) + \frac{r^3}{\pi}\int_r^\infty \frac{d}{dr_p}(\nu_p\sigma_p^2)\frac{dr_p}{\sqrt{r_p^2 - r^2}}, \tag{70b}$$

while $A_i(r)$ and $B_i(r)$ depend on the potential basis functions

$$A_i(r) = \frac{2}{3}\int_r^\infty\left(r'^3 + \frac{1}{2}r^3\right)\nu(r')\frac{d\Phi_i}{dr'}\,dr', \tag{71a}$$

$$B_i(r) = -\frac{1}{3}\int_r^\infty(r'^3 - r^3)\nu(r')\frac{d\Phi_i}{dr'}\,dr'. \tag{71b}$$

The additional constraints on the a_i are therefore

$$\sum_i a_i A_i(r) - I(r) > 0, \tag{72a}$$

$$\sum_i a_i B_i(r) - J(r) > 0; \tag{72b}$$

these conditions must, of course, be satisfied at every r. In practice, it is sufficient to satisfy the constraints (72) on a discrete radial grid, with the number of grid points N_{grid} comparable to or somewhat greater than the number of potential basis functions N_{basis}. In most of what follows, we took $N_{\text{basis}} = 30$, $N_{\text{grid}} = 100$.

We used a linear programming algorithm (routine DLPRS of IMSL) to derive optimal potentials for a stellar system with the "observed" number density and velocity dispersion profiles

$$\nu_p(r_p) = \frac{1}{\pi}\frac{1}{(1 + r_p^2)^2}, \tag{73a}$$

$$\nu_p(r_p)\sigma_p^2(r_p) = \frac{3}{32}\frac{1}{6 - q}\frac{1}{(1 + r_p^2)^{5/2}}\left(3 - \frac{5q}{4}\frac{r_p^2}{1 + r_p^2}\right). \tag{73b}$$

Equation (73a) is the projected density profile of a Plummer (1911) model, with unit total number of stars; equation (73b) defines a family of velocity dispersion profiles, parametrized by the variable q, corresponding to the anisotropic distribution functions derived in Paper II. These distribution functions were designed to reproduce the Plummer number density profile in the Plummer potential,

$$\Phi_{\text{Plummer}}(r) = \frac{1}{\sqrt{1 + r^2}}, \tag{74}$$

with unit total mass. The variable q, which ranges from $-\infty$ to 2, specifies the distribution function through equation (19) of Paper II. For our purposes, however, q may be thought of purely as a parameter that defines the shape of the velocity dispersion profile. When $q = -\infty$, the underlying distribution function contains only circular orbits, so that the projected velocity dispersion falls to zero at the center. When $q = 2$, the distribution function is strongly weighted toward highly radial orbits, and the dispersion profile is centrally peaked. Henceforth we will replace q by $Q \equiv q/(8 - q)$, which ranges from -1 (circular orbits) to 1/3 (strongly radial orbits). For the models of Paper II, Q is equal to the "global anisotropy" $(T_r - T_T/2)/(T_r + T_T/2)$, where T_r and T_T are the kinetic energies in radial and tangential motions. Figure 2 shows a set of projected velocity dispersion profiles from this family.

A useful set of potential basis functions, rooted in the Plummer model, has been defined by Clutton-Brock (1973):

$$\Phi_i(r) = -\frac{U_i(\xi)}{\sqrt{1 + r^2}}, \quad \xi = \frac{r^2 - 1}{r^2 + 1}, \tag{75}$$

where the U_i are Chebyshev polynomials of the second kind:

$$U_i(\cos\theta) = \frac{\sin[(i + 1)\theta]}{\sin\theta}. \tag{76}$$

The corresponding mass densities are

$$\rho_i(r) = \frac{K_i}{4\pi G}\frac{U_i(\xi)}{(1 + r^2)^{5/2}}, \quad K_i = 4i(i + 2) + 3. \tag{77}$$

Note that $\rho_0(r) = 3/(4\pi G)(1 + r^2)^{-5/2}$, the Plummer (1911) density law with unit total mass; similarly $\Phi_0(r) = \Phi_{\text{Plummer}}(r)$.

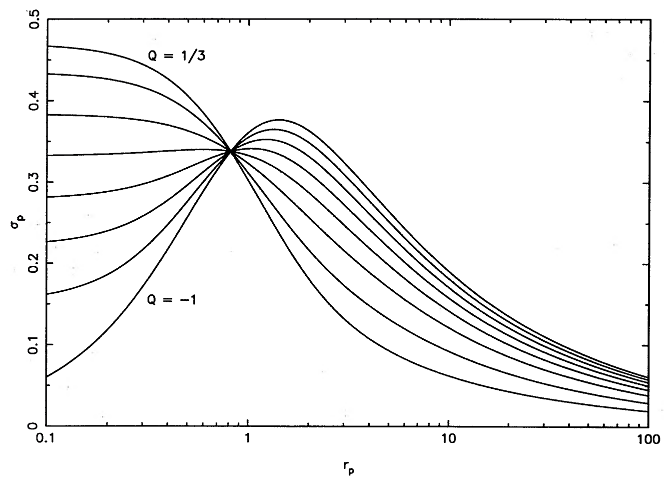

FIG. 2.—Velocity dispersion profiles, for $Q = -1, -0.8. -0.6, -0.4, -0.2$, 0., 0.2, and 1/3.

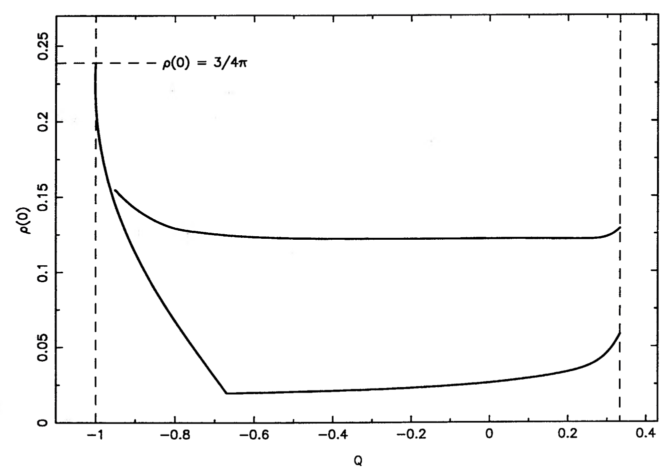

FIG. 3.—Minimum central mass densities, as a function of Q. *Lower solid line:* lower limit imposed by the constraint that the intrinsic velocity dispersions be positive; *upper solid line:* lower limit when the fourth-order virial constraints are added.

If the matter generating the potential is assumed to be distributed in the same way as the test particles, then the potential corresponding to these profiles is simply the Plummer potential, for which

$$M = 1\ , \quad \rho(0) = \frac{3}{4\pi} = 0.2387\ldots .$$

If we relax the assumption that "mass follows light," the virial theorem (in the form of eq. [66]) constrains the total mass to be greater than

$$M_{\rm min} = \frac{3\pi}{32} = 0.2945\ldots ;$$

the lower limit on the central density implied by the virial theorem (eq. [67]) is zero in this case, which may be traced to the fact that the projected number density of a Plummer model falls off more rapidly than r^{-3}. Notice that the minimum mass is independent of Q, a consequence of the fact that all of the velocity dispersion profiles correspond to models with fixed kinetic energy.

Figures 3 and 4 illustrate the lower limits on the total mass and central density when the velocity moments $\nu\sigma_r^2$ and $\nu\sigma_t^2$ inferred from $\nu_p \sigma_p^2$ are constrained to be nonnegative at all radii. For $Q = -1$, only a single potential satisfies the constraints, the Plummer potential (74). This is because the $Q = -1$ velocity dispersion profile is that of a model constructed from circular orbits, and, as argued in the Introduction, such a profile uniquely defines the potential. However as Q increases above -1, the minimum central density drops rapidly to a value of ~ 0.02, about a factor of 10 below its value in the mass-follows-light model, after which it remains roughly fixed. Similarly, the minimum mass drops smoothly from 1 at $Q = -1$ to ~ 0.3—only slightly above the lower limit set by the virial theorem—at $Q \approx 0.2$. Figure 5 shows that the minimum-density solutions come in two types, separated by $Q \approx -0.67$ (where a "kink" occurs in the plot of Fig. 3). When $Q \lesssim -0.67$, the minimum-density solution is strongly tangentially anisotropic close to the center; the central density is apparently constrained by the requirement that $\sigma_r^2 > 0$ at small r. When $Q \gtrsim -0.67$ (to the right of the "kink" in Fig. 4) the minimum density solutions are radially anisotropic at large radii, and their central densities are limited by the requirement that $\sigma_t^2 > 0$ at large r. By contrast, all of the minimum mass solutions have the same qualitative character: the anisotropy becomes strongly tangential at large radii. The behavior of these solutions is easy to understand in a qualitative sense. When the matter defining the potential is centrally concentrated, orbital velocities fall off rapidly with radius, so that trajectories at large radii must be nearly tangential in order to project the maximum fraction of their velocity along the line of sight. Similarly, when the matter is very extended, orbits must become relatively more radial at large radii, so that a small fraction of their velocity is projected along the line of sight.

These figures demonstrate—rather discouragingly— just how poorly velocity dispersion profiles constrain the underlying matter distribution. This is especially true when (as often occurs in nature) the velocity dispersion profile is centrally peaked: Figures 3 and 4 show that, for $Q \gtrsim -0.2$, the constraints imposed by the complete velocity dispersion profile are only marginally better than those imposed by the virial theorem alone. For the most strongly peaked profiles, that is, $Q \gtrsim 0.1$, even a model in which all of the mass is within about 1/10 of the Plummer model scale length can be made to fit the data. The velocity dispersion profile is therefore primarily

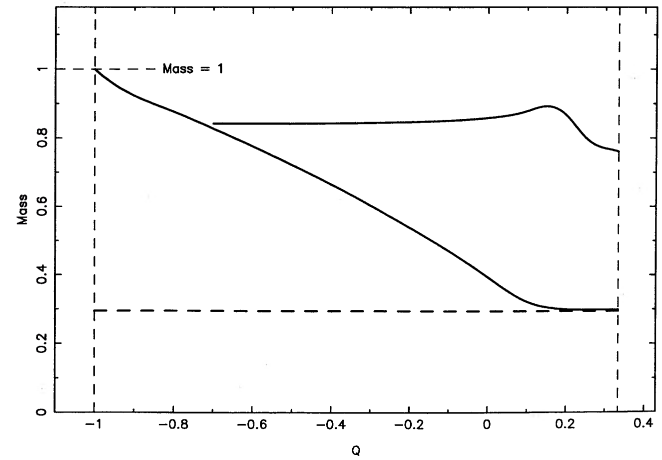

FIG. 4.—Minimum total masses, as a function of Q. *Dashed line:* lower limit imposed by the virial theorem; *lower solid line:* lower limit imposed by the constraint that the intrinsic velocity dispersions be positive; *upper solid line:* lower limit when the fourth-order virial constraints are added.

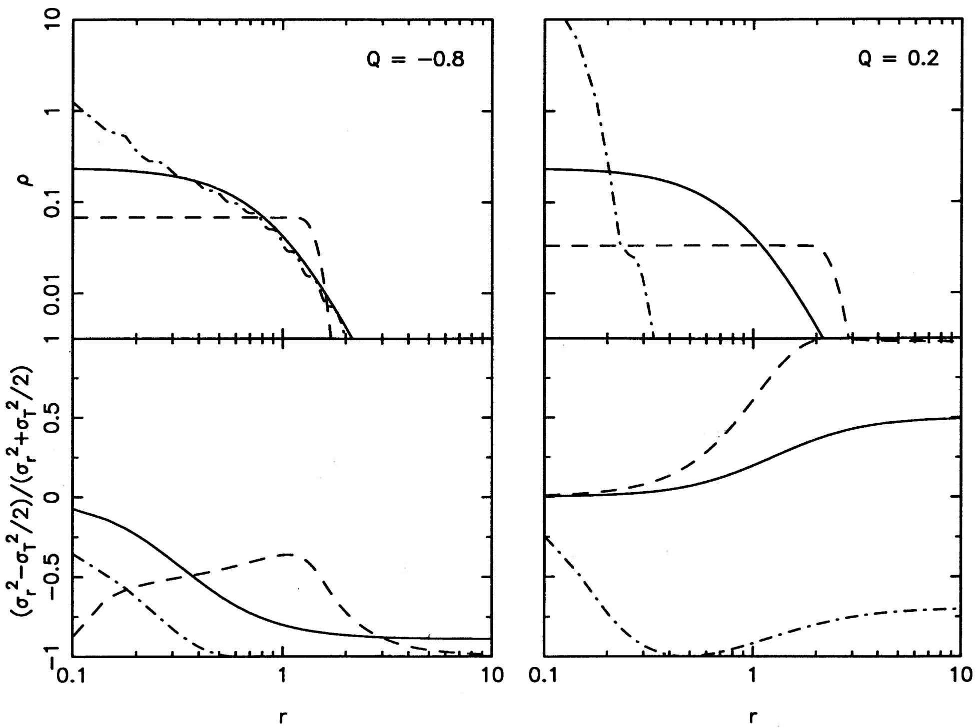

FIG. 5.—Character of the velocity-dispersion-constrained optimal solutions. *Left:* $Q = -0.8$; *right:* $Q = 0.2$. *Solid lines:* mass follows light; *dashed lines:* minimum central density; *dash-dot lines:* minimum total mass.

useful, not for constraining the potential, but rather for indicating the nature of the *kinematical* solution corresponding to an *assumed* potential.

Of course, if one knew enough about galaxy formation to cast doubt on certain kinematical states, then some of these solutions could be rejected. For instance, it is often argued that pressure-supported stellar systems should become radially anisotropic at intermediate to large radii, whether they formed via a slow or a rapid collapse. However, these arguments give no clue about where to draw the line between "acceptable" and "unacceptable" solutions. Should mildly tangentially anisotropic solutions be permitted, or only solutions that are substantially radially anisotropic? Should galaxies that show evidence for past interactions or mergers be expected to have qualitatively different distribution functions from isolated galaxies? Given the current, dismal state of knowledge about galaxy origins and evolution, we believe that it is dangerous to exclude models on the basis of "plausibility" alone. Instead, one should make use of the additional information contained within the full line-of-sight velocity distribution.

5.2. *Fourth Moments*

Next we consider how the additional information contained within the fourth-order velocity moments further constrains the potential. We begin by adding, to the constraints (65) and (72), the requirement that the solution satisfy the two fourth-order "virial theorems,"

$$\int_0^\infty \mu_p^4 r_p\, dr_p = \frac{2}{5}\int_0^\infty \nu(r)(3\sigma_r^2 + 2\sigma_\vartheta^2)\frac{d\Phi}{dr} r^3\, dr\,,$$

$$\int_0^\infty \mu_p^4 r_p^3\, dr_p = \frac{4}{35}\int_0^\infty \nu(r)(\sigma_r^2 + 6\sigma_\vartheta^2)\frac{d\Phi}{dr} r^5\, dr\,.$$

In other words, we imagine that the data are sufficiently good to yield the velocity dispersion profile accurately at all radii, but that only the radially averaged fourth moments are known. In terms of the a_i, these constraints may be written

$$\tfrac{5}{2}\int_0^\infty \mu_p^4 r_p\, dr_p = \sum_i a_i C_i^I + \sum_{i,j} a_i a_j D_{ij}^I\,, \tag{78a}$$

$$\tfrac{35}{4}\int_0^\infty \mu_p^4 r_p^3\, dr_p = \sum_i a_i C_i^{II} + \sum_{i,j} a_i a_j D_{ij}^{II}\,, \tag{78b}$$

with

$$C_i^I = -\int_0^\infty [3I(r) + 2J(r)]\frac{d\Phi_i}{dr}\, dr\,, \tag{79a}$$

$$C_i^{II} = -\int_0^\infty [I(r) + 6J(r)]\frac{d\Phi_i}{dr} r^2\, dr\,, \tag{79b}$$

$$D_{ij}^I = \int_0^\infty [3A_i(r) + 2B_i(r)]\frac{d\Phi_j}{dr}\, dr\,, \tag{79c}$$

$$D_{ij}^{II} = \int_0^\infty [A_i(r) + 6B_i(r)]\frac{d\Phi_j}{dr} r^2\, dr\,. \tag{79d}$$

By satisfying these constraints, we guarantee that the distribution function wil be positive and regular at large radii, at least to the extent that the fourth-order moments are able to do so.

The additional constraints are quadratic in the a_i, which means that we can no longer use a linear programming routine to find the optimal potentials. Instead we used the "nonlinearly constrained minimization" routine NCONG of IMSL. This routine—like virtually all nonlinear optimization routines—does not guarantee that the solution it finds is *glob-*

ally optimum; in general, the routine will select a solution that depends on the starting point, that is, on the initial guess for the a_i. However, a number of tests suggested that the solution in our case is insensitive to the starting point; results below were calculated from the "feasible" starting vector $a_0 = 1$, $a_i > 0$, $i > 0$, corresponding to the Plummer potential. Although the nonlinear routine ran quite quickly (more quickly, in fact, than the linear programming routine), it often failed to give a solution when $N_{\rm basis}$ was chosen larger than about 10. (A different routine, E04VCF of NAG, was found to fail in the same way.) The failures seemed to be due to certain coefficients in the constraint equations which, for large $N_{\rm basis}$, can be very small (see, e.g. Gill, Murray, & Wright 1981). Results presented here were obtained with $N_{\rm basis} = 10$; they are therefore probably not as accurate as the values presented in the previous subsection, calculated with $N_{\rm basis} = 30$.

Figures 3 and 4 show that the addition of the fourth-order virial constraints does, indeed, limit the range of central densities and total masses more strongly than the velocity dispersions alone. The minimum central density is typically $\sim 50\%$ of the constant-M/L value, while the total allowed mass is always greater than $\sim 75\%$ of the Plummer model mass. Figure 6 shows, however, that the range in mass density profiles allowed by the fourth-order virial constraints is still rather great: it includes matter distributions that are both significantly more and less centrally concentrated than the "luminous" objects.

The numerical difficulties that we encountered with the nonlinear optimization routine deterred us from investigating how much more strongly one could constrain the potential using the full, projected fourth-moment profile. However, Figure 4 suggests that $\mu_p^4(r_p)$ would probably impose rather stringent lower limits on at least the total mass for this family of Plummer models.

6. DISCUSSION

In a general way, the points made here—that the potential $\Phi(r)$ implied by a set of kinematical data is often strongly underdetermined; that the only fundamental constraints on Φ arise from the necessity that the inferred distribution function f be positive; that these constraints are nonlinear functions of Φ; that the virial theorems, both the usual second-order expression as well as the higher order integral constraints, may be interpreted (at least in this context) as consequences of the positivity and regularity of f; and that limited data often place a more stringent lower limit on the total mass of a stellar system than on its central density— are relevant to any attempt to infer the potential of a spherical stellar system from line-of-sight velocities. Put simply, the determination of the matter distribution of hot stellar systems from kinematical data is a difficult problem, and qualitatively different from the classical problems of determining the mass in and around rotating disks.

When the kinematical data are limited to a projected velocity dispersion profile, we have shown that the most extreme (or otherwise "optimal") potentials can be inferred through a straightforward process of linear programming, and that the range of allowed potentials is generally very great. This technique can, and probably should, be applied to real data: it is at least as simple as other, linear techniques based on the construction of f (e.g., Richstone & Tremaine 1984; Dejonghe 1989), and has the advantage that the form of $\Phi(r)$ need not be specified a priori. Perhaps its greatest disadvantage is the necessity to specify v_p and σ_p at all projected radii. In practice, this requirement would typically mandate some sort of smooth extrapolation of the observed profiles, certainly at large radii (where surface brightnesses are low), and sometimes at small

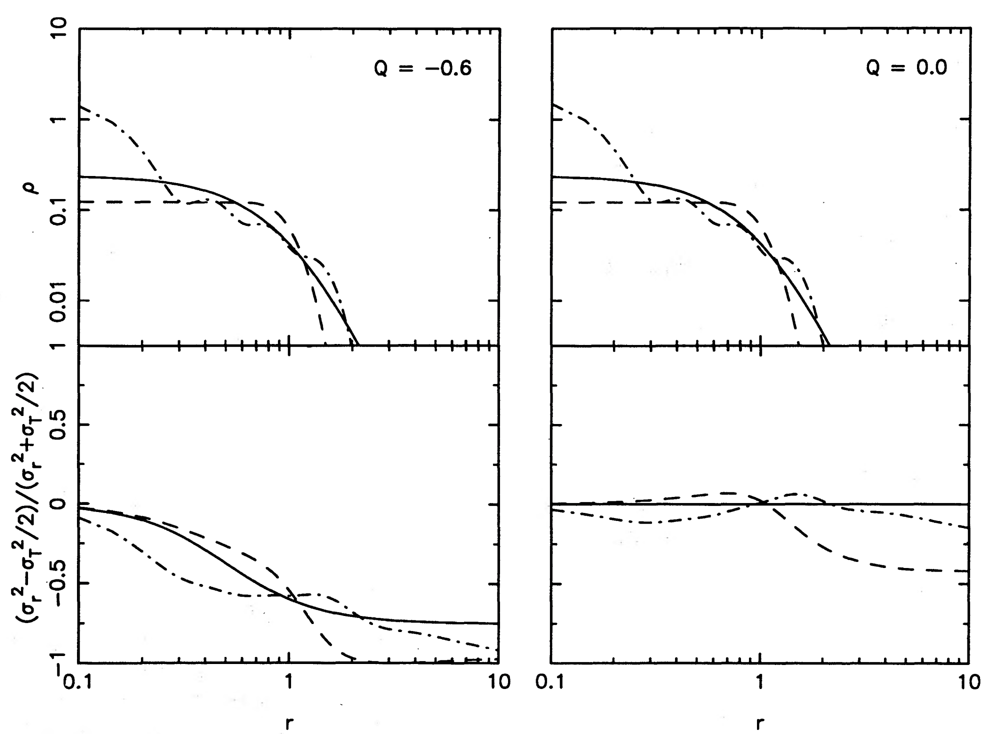

FIG. 6.—Character of the optimal solutions constrained by velocity dispersions and by the fourth-order integral constraints. *Left*: $Q = -0.6$; *right*: $Q = 0.0$. *Solid lines*: mass follows light; *dashed lines*: minimum central density; *dash-dot lines*: minimum total mass.

radii as well, as when searching for dynamical signatures of black holes. We note also that the application of an algorithm like ours to real data would almost certainly result in greater formal uncertainties in $\Phi(r)$ than previous authors have quoted, since, in the past, masses have almost always been inferred from formulae or algorithms that contain implicit restrictions (generally motivated by mathematical simplicity, not physical plausibility) on the form of Φ or f. It should also be possible to extend this technique to spherical systems with rotation, and perhaps even to axisymmetric subsystems in spherical potentials, after making an assumption about the angle between the symmetry axis and the line of sight.

In principle, one can do better with higher order moments. We do not believe that this approach is likely to be very fruitful in practice: first, because the numerical algorithms become much more complicated when moments above the second are included as constraints (i.e., nonlinear optimization vs. linear programming); second, because the accuracy with which moments of order n can be determined from limited data falls off sharply with n; and third, because the higher order moments are likely to be strongly affected by contamination, substructure, slight departures from sphericity, or other effects which are difficult to treat with any generality in an algorithm such as ours. Furthermore, when the kinematical data are *discrete*—as they are in most cases, excluding distant, unresolved elliptical galaxies—it is often the case that even the projected velocity dispersion profile cannot be meaningfully approximated, much less the higher order profiles. For instance, the largest data sets currently available for dwarf spheroidal galaxies contain only about 30 velocities (e.g., Mateo et al. 1991).

There is, however, no fundamental reason why the determination of potentials should be based on velocity *moments*. One could equally well interpret a set of observed positions and velocities as an approximation to the joint distribution $\nu_p(r_p, v_p)$ and ask (using eq. [3]) what potentials are capable of reproducing this two-dimensional data set (within the observational uncertainties) with a nonnegative f. Of course, a very large number of positions and velocities, perhaps thousands, would be required to yield a reasonably exact estimate of $\nu_p(r_p, v_p)$. However, there is some reason to believe, as argued in the Introduction, that even less-than-complete knowledge of $\nu_p(r_p, v_p)$ might yield usefully strong constraints on $\Phi(r)$. For instance, Merritt (1987) used the roughly 300 radial velocities of galaxies in the Coma cluster to construct $N(v_p)$, the overall distribution of line-of-sight velocities, and concluded that models in which the dark matter was less centrally concentrated than the galaxies were most likely. In the case of Coma, the velocity histogram was noticeably skew, a probable consequence of contamination by foreground galaxies. However, individual galaxies and star clusters are typically much more isolated and relaxed than galaxy clusters, and it seems reasonable to hope that as few as ~ 100 radial velocities in such a system would usefully constrain its potential, even in the absence of any a priori knowledge about the radial distribution of matter. Probably the simplest way to proceed is to assume a potential, then to find an f from which the data—in the form of discrete, unbinned positions and velocities—are "most likely" to have been drawn; repeating the procedure with different $\Phi(r)$ would then yield a "most probable" potential (or, more likely, a range of potentials, all of which are equally probable). The behavior of one such algorithm is discussed by Merritt & Saha (1992). We are optimistic that the application of such algorithms to real data will soon remove the indeterminacy that has heretofore always plagued the observational study of elliptical galaxies and other hot stellar systems.

This work was supported in part by NSF grant AST 90-16515 to D. M. We benefited from conversations with S. Kent, J. Kormendy, M. Merrifield, D. Richstone, and P. Saha. This work was begun while both authors were visiting the Aspen Center for Physics in 1990.

APPENDIX A

In this appendix we calculate the function

$$D_y^{-n}\left(\frac{Q^{2n}}{(2n-1)}\right).$$

The symbol D_y^{-n} is shorthand for n successive integrations, which introduce additional constants. The regularity of the intrinsic moments at $y = 0$ dictates that we choose

$$D_y^{-n}[\phi(y)] = \frac{1}{(n-1)!}\int_0^y dt(y-t)^{n-1}\phi(t) ,$$

and thus we consider

$$\frac{1}{(n-1)!}\int_0^y dt(y-t)^{n-1}\sum_{i=k}^{n}\binom{n}{i}\frac{t^i}{i!}D_t^{i-k}[f(t)] , \tag{A1}$$

with

$$f(t) = \int_t^{+\infty}(\psi'\omega_{k-1}^{n-1})(x)\frac{dx}{\sqrt{x-t}} . \tag{A2}$$

Successive partial integration transforms equation (A1) to

$$\frac{1}{(n-1)!}\sum_{i=k}^{n}\binom{n}{i}\frac{(-)^{i-k}}{i!}\int_0^y D_t^{i-k}[t^i(y-t)^{n-1}]f(t)dt ,$$

or

$$\sum_{i=k}^{n} \binom{n}{i} \frac{1}{i!} \int_0^y dt f(t) \sum_{j=0}^{i-k} \frac{(-)^{i-k}(i-k)!}{j!(i-k-j)!(i-j)!(n-1-i+j+k)!} t^{i-j}(y-t)^{n-1-i+j+k} .$$

We introduce $p = i - j$, and rewrite the above equation as

$$\sum_{p=k}^{n} \int_0^y dt \frac{f(t)t^p(y-t)^{n-1+k-p}}{p!(p-k)!(n-1+k-p)!} \sum_{i=p}^{n} \binom{n}{i} \frac{(-)^{i-p}(i-k)!}{(i-p)!i!} . \tag{A3}$$

The sum can be evaluated by noting that

$$\sum_{i=p}^{n} \binom{n}{i} \frac{(-)^{i-p}(i-k)!}{(i-p)!i!} = \frac{n!(p-k)!}{(n-p)!p!} \, {}_2F_1(p-n, p-k+1; p+1; 1)$$

$$= \frac{(p-k)!(n+k-1-p)!}{(n-p)!(k-1)!} ,$$

which in turn reduces equation (79) further to

$$\frac{1}{(k-1)!} \int_0^y dt f(t) t^k \sum_{i=0}^{n-k} \frac{t^i(y-t)^{n-1-i}}{(n-k-i)!(i+k)!} , \tag{A4}$$

or

$$y^{n+k} \frac{1}{(k-1)!} \int_0^1 du f(yu) u^k \sum_{i=0}^{n-k} \frac{u^i(1-u)^{n-1-i}}{(n-k-i)!(i+k)!} , \tag{A5}$$

together of course with equation (A1).

APPENDIX B

CALCULATION OF THE INTEGRAL CONSTRAINTS

We now determine more explicitly the functions R and S, defined in equation (53). This we do by expanding the polynomials $(1 - t/y)^i$ that appear in their definitions. This is particularly easy for R, and we get

$$R = \sum_{m=0}^{n-1} c_m y^{-m-1} , \tag{B1a}$$

with

$$c_m = n\binom{n-1}{m}(-)^m \int_0^y \mu_p^{2n}(t) t^m \, dt . \tag{B1b}$$

The calculation of

$$S = \sum_{m=0}^{n-1} d_m y^{-m-1} , \tag{B2a}$$

is more complicated. First we determine the coefficient in y^{-m-1} by simultaneously expanding $(1 - t/y)^{k-1}$ and $(1 - t/y)^{n-k-i}$ in equation (53b). We get after some manipulations

$$d_m = (-)^m(2n-1)\binom{n-1}{m} \sum_{k=1}^{n} q_{k,m}^n \int_0^y dt t^{k+m} \int_t^{+\infty} \psi' \omega_{k-1}^{n-1} \frac{dx}{\sqrt{x-t}} , \tag{B2b}$$

with

$$q_{k,m}^n = \frac{m!(n-m-1)!}{(k-1)!(n-1)!} \left[\sum_{j=0}^{k-1} \binom{k-1}{j} (-)^j t^j\right] \left[\sum_{i=0}^{\max(n-k,\, m-j)} \binom{n}{i+k} t^i \binom{n-k-i}{m-i-j} (-)^{-i-j} t^{-i-j}\right] .$$

Interchanging summations in $q_{k,m}^n$ yields

$$q_{k,m}^n = \frac{m!(n-m-1)!}{(k-1)!(n-1)!} \sum_{i=0}^{\min(n-k,m)} \binom{n}{i+k} (-)^i \sum_{j=\max(m-n+k,0)}^{\min(m-i,k-1)} \binom{k-1}{j} \binom{n-k-i}{m-i-j} . \tag{B3}$$

The inner sum in this expression can be written as

$$\frac{(n-k-i)!}{(m-i)!(n-k-m)!} \sum_{j=\max(m-n+k,0)}^{\min(m-i,k-1)} \frac{(-k+1)_j(-m+i)_j}{(n-k-m+1)_j j!} . \tag{B4}$$

When $m - n + k \leq 0$, equation (79) reduces to

$$\frac{(n-k-i)!}{(m-i)!(n-k-m)!}\, {}_2F_1(-k+1,\ -m+i;\ n-k-m+1;\ 1) = \frac{(n-i-1)!}{(n-m-1)!(m-i)!} ,$$

while, when $m - n + k > 0$, expression (79) can be transformed to

$$\frac{(k-1)!}{(n-m-1)!(m-n+k)!}\, {}_2F_1(m-n+1,\ k-n+i;\ m-n+k+1;\ 1) = \frac{(n-i-1)!}{(n-m-1)!(m-i)!} ,$$

yielding again the same result. Thus $q^n_{k,m}$ now becomes

$$q^n_{k,m} = \frac{m!}{(k-1)!(n-1)!} \sum_{i=0}^{\min(n-k,m)} \binom{n}{i+k} (-)^i \frac{(n-i-1)!}{(m-i)!} ,$$

which is, after some manipulations, fairly concisely written as

$$q^n_{k,m} = \frac{1}{(k-1)!} \binom{n}{k}\, {}_3F_2(k-n,\ -m,\ 1;\ k+1,\ 1-n;\ 1) . \tag{B5}$$

In the limit $y \to +\infty$, the coefficient $(c_m + d_m)y^{-m-1}$ must be zero, and thus follow the n constraints

$$\lim_{y\to+\infty} \left[n \int_0^y \mu_p^{2n}(y') \left(\frac{y'}{y}\right)^m \frac{dy'}{y} + (2n-1) \sum_{k=1}^{n} q^n_{k,m} \int_0^y \frac{dy'}{y} \left(\frac{y'}{y}\right)^m y^k \int_{y'}^{+\infty} \psi' \omega^{n-1}_{k-1} \frac{dx}{\sqrt{x-t}} \right] = 0 \qquad 0 \leq m \leq n-1 . \tag{B6}$$

After interchanging of the integrations, we get

$$\lim_{y\to+\infty} \left[\int_0^y \mu_p^{2n}(y') \left(\frac{y'}{y}\right)^m \frac{dy'}{y} + (2n-1)\sqrt{\pi} \sum_{k=1}^{n} a^n_{k,m} \int_0^y \psi'(x) \omega^{n-1}_{k-1}(x) \left(\frac{x}{y}\right)^m x^{k+1/2} \frac{dx}{y} \right] = 0 , \qquad 0 \leq m \leq n-1 , \tag{B7a}$$

with

$$a^n_{k,m} = \binom{n-1}{k-1} \frac{(k+m)!}{k!\Gamma(k+m+3/2)}\, {}_3F_2(k-n,\ -m,\ 1;\ k+1,\ 1-n;\ 1) , \tag{B7b}$$

or, if we put $u = y'/y$, respectively $u = x/y$,

$$\lim_{y\to+\infty} \left[\int_0^1 \mu_p^{2n}(yu) u^m\, du + (2n-1)\sqrt{\pi} \sum_{k=1}^{n} a^n_{k,m} \int_0^1 \psi'(yu) \omega^{n-1}_{k-1}(yu) u^m (yu)^{k+1/2}\, du \right] = 0 , \qquad 0 \leq m \leq n-1 . \tag{B7c}$$

REFERENCES

Bender, R. 1990, A&A, 229, 441
Bertin, G., et al. 1989, ESO Messenger, 56, 19
Binney, J., & Mamon, G. A. 1982, MNRAS, 200, 361
Clutton-Brock, M. 1973, Ap&SS, 23, 55
Dejonghe, H. 1986, Phys. Rep., 133, No. 3-4 (Paper I)
———. 1987, MNRAS, 224, 13 (Paper II)
———. 1989, ApJ, 343, 113
Ford, H. C., Ciardullo, R., Jacoby, G. H. & Hui, X. 1989, in IAU Symp. 131, Planetary Nebula, ed. S. Torres-Peimbert (Dordrecht: Kluwer), 335
Gill, P. E., Murray, W., & Wright, M. H. 1981, Practical Optimization (London: Academic)
Goldstein, J. 1980, Classical Mechanics, 2d ed. (Reading: Addison-Wesley)
Heisler, J., Tremaine, S., & Bahcall, J. N. 1985, ApJ, 298, 8
Katz, N., & Richstone, D. O. 1985, ApJ, 296, 331
Kent, S. M. 1991, MNRAS, 247, 702
te Linkel Hekkert, P., & Dejonghe, H. 1989, in Dynamics of Astrophysical Disks, ed. J. Sellwood (Cambridge: Cambridge Univ. Press), 139
Mateo, M., Olszewski, E., Welch, D. L., Fischer, P., & Kunkel, W. 1991, AJ, 102, 914
Merrifield, M. R., & Kent, S. M. 1990, AJ, 99, 1548
Merritt, D. 1987, ApJ, 313, 121
Merritt, D., & Saha, P. 1992, Rutgers Astrophysics Preprint, no. 125
Meylan, G., & Mayor, M. 1991. STScI, Preprint, no. 542
Mould, J. R., Oke, J. B., de Zeeuw, P. T., Nemec, J. M. 1990, AJ, 99, 1823
Newton, A. J., & Binney, J. 1984, MNRAS, 210, 711
Plummer, H. C. 1911, MNRAS, 71, 460
Pryor, C., & Kormendy, J. 1990, AJ, 100, 127
Richstone, D. O., & Tremaine, S. 1984, ApJ, 286, 27
Schittkowski, K. 1980, Lecture Notes in Economics and Mathematical Systems, 183
Teague, P. F., Carter, D., & Gray, P. M. 1990, ApJS, 72, 715