Equilibrium models of stellar systems

Eugene Vasiliev

Institute of Astronomy, Cambridge

45th Heidelberg Physics Graduate Days, October 2020

Distribution function in stellar dynamics

A stellar system composed of a large number $N \gg 1$ of "identical" stars can be described by a distribution function (DF):

$f(\mathbf{x}, \mathbf{v};\ t)$ – probability density in the phase space (at time t),

$\int f(\mathbf{x}, \mathbf{v})\, d^3x\, d^3v = 1$ (unit-normalized), or alternatively

$\int f(\mathbf{x}, \mathbf{v})\, d^3x\, d^3v = M_{\rm total}$ (mass-normalized).

For a multicomponent system, define a separate DF f_k for each species k, or an "extended distribution function" (EDF) with additional arguments:

$f(\mathbf{x}, \mathbf{v};\ \boldsymbol{\eta};\ t)$ (η may be age, mass, chemical composition, etc.)

Properties of the distribution function

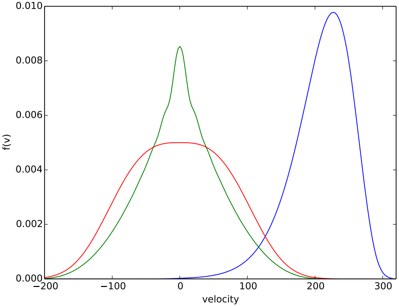
DF offers a complete description of the stellar population:

- density $\rho(\mathbf{x}) = \int f(\mathbf{x},\mathbf{v})\, d^3v$,
- mean velocity $\overline{\mathbf{v}}(\mathbf{x}) = \dfrac{1}{\rho(\mathbf{x})} \int \mathbf{v}\, f(\mathbf{x},\mathbf{v})\, d^3v$,
- second moment of velocity $\overline{v_{ij}^2}(\mathbf{x}) = \dfrac{1}{\rho(\mathbf{x})} \int v_i\, v_j\, f(\mathbf{x},\mathbf{v})\, d^3v$,

  velocity dispersion tensor $\sigma_{ij}^2 \equiv \overline{v_{ij}^2} - \overline{v_i}\,\overline{v_j}$,
- more generally, velocity distribution at a given point

  $\mathfrak{f}(v_1;\, \mathbf{x}) = \dfrac{1}{\rho(\mathbf{x})} \int f(\mathbf{x},\mathbf{v})\, dv_2\, dv_3$

  (note how non-Gaussian it can be!),
- etc.



Evolution of the distribution function

As the system evolves, stars move along their trajectories, but the DF at the location of any star is conserved (Liouville's theorem):

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{x}}\frac{d\mathbf{x}}{d\mathbf{t}} + \frac{\partial f}{\partial \mathbf{v}}\frac{d\mathbf{v}}{d\mathbf{t}}$$

$$= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{x}}\,\mathbf{v} - \frac{\partial f}{\partial \mathbf{v}}\frac{\partial \Phi}{\partial x} = 0$$

collisionless Boltzmann (or Vlasov) equation

generally valid for large-N systems (e.g., entire galaxies), in which the relaxation time is long compared to their age (even if the system is not in a steady state!)

Evolution of the distribution function

As the system evolves, stars move along their trajectories, but the DF at the location of any star is ~~conserved~~ (Liouville's theorem):

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{x}} \frac{d\mathbf{x}}{d\mathbf{t}} + \frac{\partial f}{\partial \mathbf{v}} \frac{d\mathbf{v}}{d\mathbf{t}}$$

$$= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{x}}\, \mathbf{v}\; - \frac{\partial f}{\partial \mathbf{v}} \frac{\partial \Phi}{\partial x} = \Gamma[f]$$

collision operator

(generalized Landau equation or its simplifications, including various forms of Fokker–Planck equation) – applies to thermodynamically old systems such as globular clusters

Phase mixing and the Jeans theorem

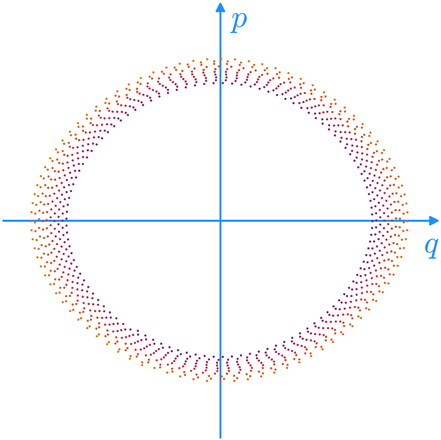
In realistic galactic potentials, orbital frequencies $\boldsymbol{\Omega} \equiv \dfrac{\partial H}{\partial \mathbf{J}} \neq \text{const}$, hence an initially localized ensemble of points eventually mixes in orbital phase and fills the entire accessible region for its integrals of motion.

Jeans theorem

In a steady state, the DF may depend only on the integrals of motion:

$$f(\mathbf{x}, \mathbf{v}) = f\big(\mathcal{I}(\mathbf{x}, \mathbf{v};\, \Phi)\big)$$

Here $\mathcal{I}$ may be energy E, or actions $\mathbf{J}$ – anything that is conserved by the given potential.

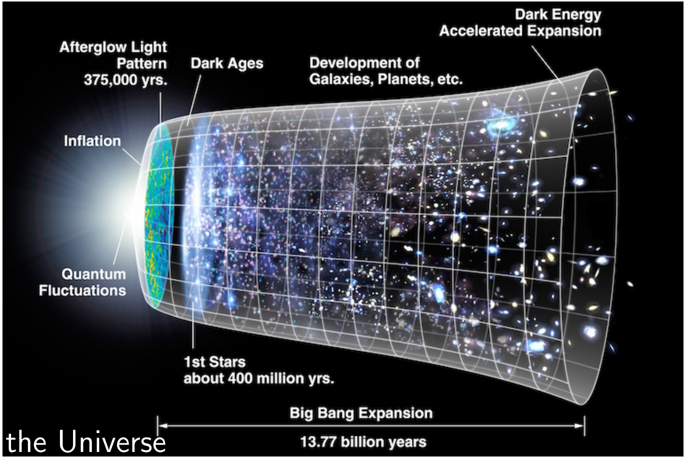


Equilibrium models

A lot of effort in stellar dynamics is devoted to construction of models of stellar systems in a dynamical equilibrium (used interchangeably with "steady state"), also known as "dynamical modelling" (not to be confused with an N-body simulation of the evolution of a self-gravitating stellar system!)

Fred Hoyle

vs.



the Universe

Why steady state?

Distribution function of stars $f(\mathbf{x}, \mathbf{v}, t)$
satisfies [sometimes] the collisionless Boltzmann equation:

$$\frac{\partial f(\mathbf{x}, \mathbf{v}, t)}{\partial t} + \mathbf{v} \frac{\partial f(\mathbf{x}, \mathbf{v}, t)}{\partial \mathbf{x}} - \frac{\partial \Phi(\mathbf{x}, t)}{\partial \mathbf{x}} \frac{\partial f(\mathbf{x}, \mathbf{v}, t)}{\partial \mathbf{v}} = 0.$$

Potential $\Leftrightarrow$ mass distribution

Why steady state?

Distribution function of stars $f(\mathbf{x}, \mathbf{v}, t)$
satisfies [sometimes] the collisionless Boltzmann equation:

$$\mathbf{v}\,\frac{\partial f(\mathbf{x}, \mathbf{v}\quad)}{\partial \mathbf{x}} - \frac{\partial \Phi(\mathbf{x}\quad)}{\partial \mathbf{x}}\,\frac{\partial f(\mathbf{x}, \mathbf{v}\quad)}{\partial \mathbf{v}} = 0.$$

3D

Steady-state assumption $\Longrightarrow$ Jeans theorem:

$$f(\mathbf{x}, \mathbf{v}) = f\left(\mathcal{I}(\mathbf{x}, \mathbf{v};\ \Phi)\right)$$

3D – 6D

integrals of motion ($\leq$ 3D?), e.g., $\mathcal{I} = \{E, L, \ldots\}$

Two aspects of dynamical modelling

1. Construction of ad-hoc self-consistent equilibrium models,
   in which the DF $f(\mathcal{I})$ and the potential Φ are consistent with each other.

   Often these models would be "as simple as possible (but no simpler)".

2. Construction of models of real stellar systems based on some observations.

   These measurements would provide at least one component of velocity (usually the line-of-sight velocity distribution, or some of its moments) at some number of locations on the sky (ideally, a densely sampled 2d map, but often only a few fibers or slits).

   The goal is to infer the distribution of the **total** mass (stars + dark matter + central black hole + ...) from the observed kinematics of some tracers.

   Sometimes these models could be taken from a family of well-studied theoretical models of the first kind, and in other cases, constructed to match the observations as closely as possible, without enforcing any specific form.

Methods: overview

0. Virial theorem: $2\,K + W = 0$

kinetic energy → K; W ← potential energy

virial mass estimators: $G\,M \propto r\,\sigma^2$ [e.g., Wolf+ 2010; Churazov+ 2010]

1. Jeans equations

2. Distribution functions

3. Schwarzschild's orbit superposition

4. Made-to-measure

Jeans equations

Multiply the CBE by various powers of velocity and integrate over velocities.

A simple 1d example:

CBE for $f(z, v_z;\ t)$: $$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial z}\, v_z - \frac{\partial f}{\partial v_z}\,\frac{\partial \Phi}{\partial z} = 0$$

Jeans equations

Multiply the CBE by various powers of velocity and integrate over velocities.

A simple 1d example:

$$\text{multiply by } v_z^0 : \qquad \int_{-\infty}^{\infty} \left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial z} v_z - \frac{\partial f}{\partial v_z} \frac{\partial \Phi}{\partial z} \right] dv_z = 0$$

$$\frac{\partial}{\partial t} \int_{-\infty}^{\infty} f(z, v_z)\, dv_z + \frac{\partial}{\partial z} \int_{-\infty}^{\infty} f(z, v_z)\, v_z\, dv_z - \frac{\partial \Phi(z)}{\partial z} \cancel{\Big(f(z, v_z) \Big)}\Big|_{v_z=-\infty}^{+\infty} = 0$$

vanishes because

$f \to 0$ as $|v_z| \to \infty$

$$\frac{\partial \rho}{\partial t} + \frac{\partial \left(\rho\, \overline{v_z} \right)}{\partial z} = 0$$

(continuity equation)

Jeans equations

Multiply the CBE by various powers of velocity and integrate over velocities.

A simple 1d example:

multiply by v_z : $$\int_{-\infty}^{\infty} \left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial z} v_z - \frac{\partial f}{\partial v_z} \frac{\partial \Phi}{\partial z} \right] v_z \, dv_z = 0$$

$$\frac{\partial(\rho \, \overline{v_z})}{\partial t} + \frac{\partial\left(\rho \, \overline{v_z^2}\right)}{\partial z} - \frac{\partial \Phi}{\partial z} \left[\cancel{\left(f \, v_z \right)\Big|_{v_z=-\infty}^{+\infty}} - \int_{-\infty}^{\infty} f \, \frac{\partial v_z}{\partial v_z} \, dv_z \right] = 0$$

Jeans equations

Multiply the CBE by various powers of velocity and integrate over velocities.

A simple 1d example:

multiply by v_z :
$$\int_{-\infty}^{\infty} \left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial z}\, v_z - \frac{\partial f}{\partial v_z}\, \frac{\partial \Phi}{\partial z} \right] v_z \, dv_z = 0$$

$$\frac{\partial (\rho\, \overline{v_z})}{\partial t} + \frac{\partial \left(\rho\, \overline{v_z^2}\right)}{\partial z} + \frac{\partial \Phi}{\partial z}\, \rho = 0$$

or

$$\rho\, \frac{\partial \overline{v_z}}{\partial t} + \rho\, \overline{v_z}\, \frac{\partial \overline{v_z}}{\partial z} + \frac{\partial \left(\rho\, \sigma_z^2\right)}{\partial z} + \frac{\partial \Phi}{\partial z}\, \rho = 0$$

analog of Euler equation

Jeans equations

Multiply the CBE by various powers of velocity and integrate over velocities.

A simple 1d example:

multiply by v_z :
$$\int_{-\infty}^{\infty}\left[\frac{\partial f}{\partial t}+\frac{\partial f}{\partial z}\,v_z-\frac{\partial f}{\partial v_z}\,\frac{\partial \Phi}{\partial z}\right] v_z\,dv_z=0$$

$$\frac{\partial\left(\rho\,\overline{v_z^2}\right)}{\partial z}+\frac{\partial\Phi}{\partial z}\,\rho=0$$

(steady state, no bulk motion)

measure from observations

or

want to infer

$$\frac{\partial\left(\rho\,\sigma_z^2\right)}{\partial z}+\frac{\partial\Phi}{\partial z}\,\rho=0$$

e.g., vertical force above an infinite, uniform galactic disc –
toy example, real life is not so simple!

Jeans equations (general form, steady state)

Multiply the CBE by velocity v_i and integrate over the 3d velocity space:

$$0 = \int d^3\mathbf{v}\ v_i \sum_j \left[v_j \frac{\partial f(\mathbf{x},\mathbf{v})}{\partial x_j} - \frac{\partial \Phi(\mathbf{x})}{\partial x_j} \frac{\partial f(\mathbf{x},\mathbf{v})}{\partial v_j} \right]$$

$$= \sum_j \left[\int d^3\mathbf{v}\ v_i\, v_j \frac{\partial f}{\partial x_j} - \frac{\partial \Phi}{\partial x_j} \int d^3\mathbf{v}\ v_i \frac{\partial f}{\partial v_j} \right]$$

$$= \sum_j \left[\frac{\partial \left(\int d^3\mathbf{v}\ v_i\, v_j\, f \right)}{\partial x_j} + \frac{\partial \Phi}{\partial x_j} \int d^3\mathbf{v} \underbrace{\frac{\partial v_i}{\partial v_j}}_{=\delta_{ij}} f \right]$$

pressure gradient

$$= \sum_j \frac{\partial \left(\rho\, \overline{v_i\, v_j} \right)}{\partial x_j} + \frac{\partial \Phi}{\partial x_i} \rho,$$

gravitational force

hydrostatic equilibrium

where $\rho = \int d^3\mathbf{v}\, f, \quad \overline{v_i\, v_j} = \frac{1}{\rho} \int d^3\mathbf{v}\ v_i\, v_j\, f.$

3 equations, 6 components of $\overline{v_i\, v_j}$ – underdetermined system!

Jeans equations: spherical case

In the spherical non-rotating case, only one nontrivial equation remains:

$$0 = \frac{d\left(\rho\, \sigma_r^2\right)}{dr} + \frac{d\Phi}{dr}\,\rho + \frac{\rho}{r}\left(2\sigma_r^2 - \sigma_\theta^2 - \sigma_\phi^2\right)$$

$$= \frac{d\left(\rho\, \sigma_r^2\right)}{dr} + \frac{d\Phi}{dr}\,\rho + \frac{2\beta}{r}\,\rho\,\sigma_r^2$$

where $\beta(r) \equiv 1 - \dfrac{\sigma_\theta^2(r) + \sigma_\phi^2(r)}{2\sigma_r^2(r)}$ is the anisotropy coefficient :

$\beta = 1$ – purely radial orbits,

$\beta > 0$ – radially anisotropic case,

$\beta = 0$ – isotropic case,

$\beta < 0$ – tangentially anisotropic case,

$\beta = -\infty$ – purely circular orbits.

Jeans equations: spherical case

In the spherical non-rotating case, only one nontrivial equation remains:

$$
\begin{aligned}
0 &= \frac{d\left(\rho\,\sigma_r^2\right)}{dr} + \frac{d\Phi}{dr}\,\rho + \frac{\rho}{r}\left(2\sigma_r^2 - \sigma_\theta^2 - \sigma_\phi^2\right) \\
&= \frac{d\left(\rho\,\sigma_r^2\right)}{dr} + \frac{d\Phi}{dr}\,\rho + \frac{2\beta}{r}\,\rho\,\sigma_r^2
\end{aligned}
$$

The relation between σ_r and β is given by

$$
\begin{aligned}
\sigma_r^2(r) &= \frac{1}{\rho(r)\,g(r)} \int_r^\infty \frac{G\,M(r')\,\rho(r')\,g(r')}{r'^2}\,dr' \\
g(r) &\equiv \exp\left[2\int_0^r \frac{\beta(r')}{r'}\,dr'\right]
\end{aligned}
$$

[van der Marel 1994; Mamon & Łokas 2005]

Jeans equations: spherical case in projection

We usually measure only the projected (surface) density $\Sigma(R)$ and the line-of-sight velocity dispersion $\sigma^2_{\rm los}(R)$:

$$\Sigma(R) = 2\int_R^\infty \frac{\rho(r)\,r}{\sqrt{r^2-R^2}}\,dr$$

$$\sigma^2_{\rm los}(R) = \frac{2}{\Sigma(R)}\int_R^\infty \left(1-\beta(r)\,\frac{R^2}{r^2}\right)\,\frac{\sigma_r^2(r)\,\rho(r)\,r}{\sqrt{r^2-R^2}}\,dr$$

[Binney & Mamon 1982]

Unknown functions:

$\rho(r)$ — can be obtained by deprojecting $\Sigma(r)$

$\beta(r)$ — parameters of the tracer population (stars);

$\sigma_r(r)$ — related by the Jeans equation

$\Phi(r)$ or $M(r)$ or $v_{\rm circ}(r)$ – total potential (stars, dark matter, SMBH, etc.)

Jeans equations: spherical mass inversion

In the isotropic case ($\beta = 0$):

$$\rho(r) = -\frac{1}{\pi} \int_r^\infty \frac{d\Sigma(R)}{dR} \frac{dR}{\sqrt{R^2 - r^2}}$$

$$v_{\rm circ}^2(r) \equiv \frac{G\,M(r)}{r} = \frac{1}{\pi\,\rho(r)} \int_r^\infty \frac{d^2\left[\Sigma(R)\,\sigma_{\rm los}^2(R)\right]}{dR^2} \frac{R\,dR}{\sqrt{R^2 - r^2}}$$

For a general anisotropic case (when $\beta(r)$ is assumed to be known), one may express $v_{\rm circ}^2(r)$ using double integrals over $\beta(r)$, which can be computed analytically for several common functional forms of $\beta(r)$ [Mamon & Boué 2009; Wolf+ 2009].

These expressions involve 1st and 2nd derivatives of [noisy] observables...

Jeans equations: mass–anisotropy degeneracy

Jeans equations are not closed and do not allow one to determine σ, β and Φ simultaneously without making further assumptions.

There are several possible ways to lift this degeneracy:

- Use of higher-order moments (e.g., kurtosis) or virial shape parameters [Merrifield & Kent 1990; Richardson & Fairbairn 2013, 2014; Read & Steger 2017].
- Use of multiple independent tracer populations * [Walker & Peñarrubia 2011; Amorisco & Evans 2011].
- Use of extra information from proper motions [e.g., Wilkinson+ 2002; Strigari+ 2007; Massari+ 2019].

* Typically, one may determine the mass inside some specifically chosen radius ($\simeq r_{\text{half-mass}}$) nearly independently of β [e.g., Wolf+ 2010; Lyskova+ 2012]. With multiple spatially-distinct stellar populations, one may constrain the enclosed mass profile at several radii.

Jeans equations: axisymmetric case

$\Phi(R,z)$ – total gravitational potential

$\rho(R,z)$ – density of tracers

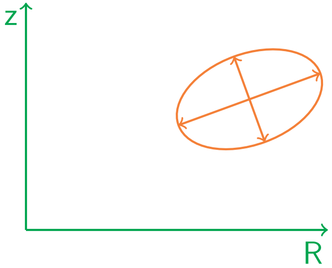
$$0 = \rho\,\frac{\partial\Phi}{\partial z} + \frac{\partial(\rho\,\sigma_z^2)}{\partial z} + \frac{\partial(\rho\,\overline{v_R v_z})}{\partial R} + \frac{\rho\,\overline{v_R v_z}}{R}$$

$$0 = \rho\,\frac{\partial\Phi}{\partial R} + \frac{\partial(\rho\,\sigma_R^2)}{\partial R} + \frac{\partial(\rho\,\overline{v_R v_z})}{\partial z} + \frac{\rho\,\left(\sigma_R^2 - \overline{v_\phi^2}\right)}{R}$$

Two equations for four unknown functions (components of the velocity ellipsoid tensor):

$\sigma_R^2,\;\; \sigma_z^2,\;\; \overline{v_R v_z},\;\; \overline{v_\phi^2} = \overline{v_\phi}^2 + \sigma_\phi^2.$

Need further assumptions about the orientation of the velocity ellipsoid in the meridional plane.



Jeans equations: axisymmetric case – semi-isotropic

Assume $\overline{v_R v_z} = 0$ and $\sigma_R^2 = \sigma_z^2$:

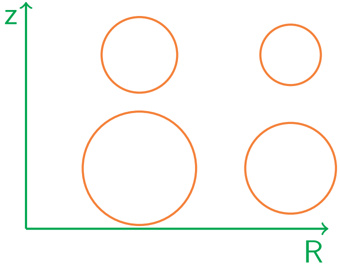
$$0 = \rho \frac{\partial \Phi}{\partial z} + \frac{\partial (\rho\, \sigma_R^2)}{\partial z}$$

$$0 = \rho \frac{\partial \Phi}{\partial R} + \frac{\partial (\rho\, \sigma_R^2)}{\partial R} + \frac{\rho \left(\sigma_R^2 - \overline{v_\phi^2}\right)}{R}$$

Used in many papers throughout 1980s – 2000s

Still need to decide* how to split $\overline{v_\phi^2} = \overline{v_\phi}^2 + \sigma_\phi^2$
e.g., assume full isotropy $\sigma_\phi^2 = \sigma_R^2$ (unrealistic!)

* this is true for all variants of Jeans equations



Jeans equations: axisymmetric case – spherical alignment

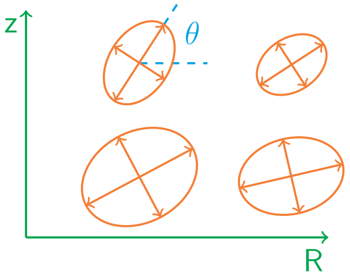
Assume orientation of the velocity ellipsoid towards the galactic center:

$$\tan 2\theta = \frac{2\,\overline{v_R v_z}}{\sigma_R^2 - \sigma_z^2} = \frac{2\,Rz}{R^2 - z^2}$$

$$0 = \rho\,\frac{\partial\Phi}{\partial z} + \frac{\partial(\rho\,\sigma_z^2)}{\partial z} + \frac{\partial(\rho\,\overline{v_R v_z})}{\partial R} + \frac{\rho\,\overline{v_R v_z}}{R}$$

$$0 = \rho\,\frac{\partial\Phi}{\partial R} + \frac{\partial(\rho\,\sigma_R^2)}{\partial R} + \frac{\partial(\rho\,\overline{v_R v_z})}{\partial z} + \frac{\rho\,\left(\sigma_R^2 - \overline{v_\phi^2}\right)}{R}$$

A good approximation for realistic galaxies;
advocated by Binney 2014; Evans+ 2016
but more complicated and rarely used
(although see Cappellari 2019);
need further assumptions about the shape
of the velocity ellipsoid



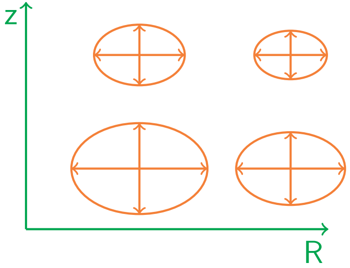
Jeans equations: axisymmetric case – cylindrical alignment

Assume $\overline{v_R v_z} = 0$ and $\sigma_R^2/\sigma_z^2 = b = \text{const}$:

$$0 = \rho \frac{\partial \Phi}{\partial z} + \frac{\partial(\rho\, \sigma_R^2)}{\partial z} \frac{1}{b}$$

$$0 = \rho \frac{\partial \Phi}{\partial R} + \frac{\partial(\rho\, \sigma_R^2)}{\partial R} + \frac{\rho\left(\sigma_R^2 - \overline{v_\phi^2}\right)}{R}$$

Jeans Anisotropic Method (JAM)
[Cappellari 2008; Watkins+ 2013; Zhu+ 2016]



Jeans equations: axisymmetric case – one-dimensional

Assume $\overline{v_R v_z} = 0$ and consider only one equation at a fixed R:

$$0 = \rho \frac{\partial \Phi}{\partial z} + \frac{\partial(\rho\, \sigma_z^2)}{\partial z}$$

Used in various studies to infer the vertical profile of the potential in the Solar neighborhood.

Jeans models of observed stellar systems

- Choose your approximation (spherical / axisymmetric, velocity alignment, etc.)
- Choose the parameters of the model:
  density profile, potential model, anisotropy coefficient, etc.
- Compute the velocity dispersion tensor $\overline{v_i\, v_j}\,(\mathbf{x})$
- Compute observable quantities ($\sigma_{\rm los}(X, Y)$, etc.)
- Compare with the data and evaluate the quality of fit
- Repeat for many different choices of model parameters,
  find the best ones and determine their uncertainties

2. Distribution function-based models

1. Collisionless Boltzmann equation + Jeans theorem:

$$f(\mathbf{x},\mathbf{v}) = f\big(\mathcal{I}(\mathbf{x},\mathbf{v};\ \Phi)\big), \qquad \mathcal{I} = \left\{E \equiv \Phi(\mathbf{x}) + \tfrac{1}{2}|\mathbf{v}|^2, \ldots\right\}$$

integrals of motion – depend on Φ

2. Poisson equation:

$$\nabla^2\Phi(\mathbf{x}) = 4\pi\, G\, \rho(\mathbf{x})$$

3. The link:

$$\rho(\mathbf{x}) = \iiint d^3\mathbf{v}\ f(\mathbf{x},\mathbf{v})$$

Two alternative approaches: $f(\mathcal{I}) \Longrightarrow \rho,\ \Phi$ or $\rho,\ \Phi \Longrightarrow f$.

Distribution function-based models, spherical case

1. $f(E,L) \Longrightarrow \Phi(r),\ \rho(r)$:

$$\nabla^2\Phi = \frac{1}{r^2}\frac{d}{dr}\left(r^2\,\frac{d\Phi(r)}{dr}\right) = 4\pi\,G\iiint d^3\,\mathbf{v}\ \underbrace{f\left(\overset{\substack{E\\ \parallel}}{\Phi(r)+\tfrac{1}{2}|\mathbf{v}|^2},\ \overset{\substack{L\\ \parallel}}{|\mathbf{x}\times\mathbf{v}|}\right)}_{\text{assumed functional form}}$$

second-order integro-differential equation for $\Phi(r)$.

Examples:

- Polytropes: $f(E) \propto |E|^{n-3/2}$ (e.g. $n=5$ is the Plummer(1911) model)
- Lowered isothermal models:

$$f(E,L) \propto \left[\exp\left(-\tfrac{E}{\sigma^2}\right) - \text{const}\right]\exp\left(-\tfrac{L^2}{2\sigma^2\,r_a^2}\right)$$

[King 1962; Michie 1963; Wilson 1975; Gieles & Zocchi 2015]

Distribution function-based models, spherical case

2. $\Phi(r),\ \rho(r) \Longrightarrow f(E,L)$: several choices for factorizations of $f(E,L)$.

- Eddington inversion formula for isotropic $f(E)$: [Eddington 1916]

$$f(E) = \frac{1}{2\pi^2} \int_E^0 \frac{d^2\rho}{d\Phi^2} \frac{d\Phi}{\sqrt{2(\Phi - E)}}, \qquad \rho(\Phi) = \rho(r)\big|_{\Phi(r)=\Phi}$$

- Cuddeford–Osipkov–Merritt inversion: [Osipkov 1979; Merritt 1985; Cuddeford 1991]

  $f(E,L) = f_Q(Q)\, L^{-2\beta_0}, \quad Q \equiv E + L^2/(2\, r_a^2),$

  $f_Q(Q)$ is given by a similar integral expression.

  Anisotropy coefficient β ranges from β_0 at small r to $\beta_\infty = 1$ at $r \gg r_a$.

- (Quasi-)separable models:

  $f(E,L) = f_E(E)\, h(x), \quad x \equiv L^\alpha \big/ \big[L_0^\alpha + L_{\mathrm{circ}}^\alpha(E)\big]$ [Gerhard 1991; Saha 1992]

  $f(E,L) = f_E(E)\, f_L(L), \quad f_L = (1 + L^2/L_0^2)^{\beta_0-\beta_\infty}\, L^{-2\beta_0}$ [Wojtak+ 2008]

  $f_E(E)$ is determined numerically from a Volterra integral equation.

Distribution function-based models, axisymmetric case

1. $f\left(E, L_z\,[, I_3]\right) \Longrightarrow \Phi(R,z),\ \rho(R,z)$: iterative approach

   [Prendergast & Tomer 1975; Rowley 1988; Kuijken & Dubinski 1995; Widrow+ 2005; Binney 2014; Piffl+ 2015; Sanders & Evans 2016; Vasiliev 2019]

   - Assume a functional form for $f(\mathcal{I})$ and a starting guess for $\Phi(\mathbf{x})$;
   - Establish the integrals of motion $\mathcal{I}(\mathbf{x}, \mathbf{v};\ \Phi)$;
   - Compute $\rho(\mathbf{x}) = \iiint d^3\mathbf{v}\, f\left(\mathcal{I}(\mathbf{x}, \mathbf{v};\ \Phi)\right)$;
   - Compute the new potential from the Poisson equation: $\nabla^2\Phi = 4\pi\, G\, \rho$.
   - Repeat until convergence.

2. $\Phi(R,z),\ \rho(R,z) \Longrightarrow f(E, L_z)$ expressed as a contour integral

   [Hunter & Qian 1993] – cumbersome and rarely used in practice;

   two-integral DF is not very realistic – has $\sigma_R = \sigma_z$.

Distribution function-based models, nonparametric

Goal: determine the DF from the observed kinematics.

Spherical case: projected DF $\mathcal{F}(R, v_{\rm los}) \implies f(E, L)$ [Dejonghe & Merritt 1992].
Axisymmetric edge-on case: $\overline{v_{\rm los}}(X, Y), \sigma_{\rm los}(X, Y) \implies f(E, L_z)$ [Merritt 1996].

The solution is given by inverting the integral equation

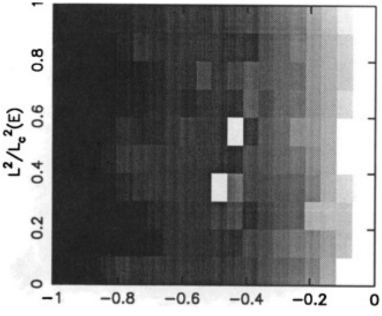
$$\mathcal{F}(X, Y, v_{\rm los}) = \int dZ \iint dv_X\, dv_Y\; f\big(\mathcal{I}[\mathbf{x}, \mathbf{v}]\big)$$

A practical approach:

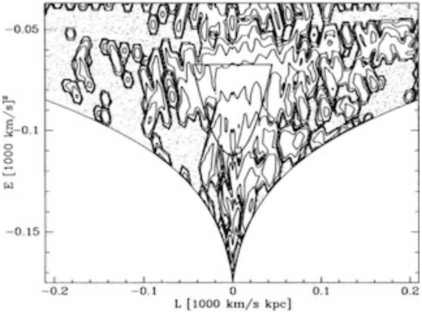
- discretize the projected DF into $\mathcal{F}^{(n)} \equiv \mathcal{F}(X^{(n)}, Y^{(n)}, v_{\rm los}^{(n)})$
- represent $f(\mathcal{I})$ as a sum of basis functions with unknown amplitudes: $f = \sum_k A_k f_k$;
- compute the projection of each basis function $\mathcal{F}_k(X, Y, v_{\rm los})$;
- find the best-fit amplitudes A_k satisfying $\sum_k A_k \mathcal{F}_k^{(n)} = \mathcal{F}^{(n)}$.

Distribution function-based models, nonparametric

- Dejonghe 1989; Merritt & Saha 1993: $f_k(E, L)$ as Fricke components $|E|^{\alpha} L^{-2\beta}$;
- Merritt 1993, 1996: histograms (Π-shaped blocks) for $f(E, L)$ or $f(E, L_z)$;
- Kuijken 1994; Pichon & Thiébaut 1998: bilinear interpolation for $f(E, L_z)$;
- Dehnen & Gerhard 1994: Chebyshev polynomial basis for $f(E, L_z)$;
- Magorrian 2014: superposition of multivariate Gaussian 'blobs' for $f(E, L)$.
- Magorrian 2019: rectangular blocks for $f(E, L, L_z)$.



E [Merritt 1993]



[Kuijken 1994]

Distribution function models of observed stellar systems

- Choose your approximation (spherical / axisymmetric, DF class, etc.)

$f \Rightarrow \Phi$ $\Phi \Rightarrow f$

Parametric DF | Fixed-form DF | Non-parametric DF

assume f | assume Φ

compute Φ | compute f | compute observables for f_k

compute observables | compute weights of f_k

- Compare with the data and evaluate the quality of fit
- Repeat for many different choices of model parameters, find the best ones and determine their uncertainties

3. Schwarzschild's orbit-superposition method

Introduced by Schwarzschild (1979) as a practical approach
for constructing self-consistent triaxial models with prescribed $\rho(\mathbf{x}) \Leftrightarrow \Phi(\mathbf{x})$.

integrals of motion

To invert the equation $\rho(\mathbf{x}) = \iiint f\left(\mathcal{I}\left[\mathbf{x}, \mathbf{v} \mid \Phi\right]\right) \, d^3\mathbf{v}$,

discretize both the density profile and the distribution function:

$\rho(\mathbf{x}) \implies$ cells of a spatial grid; mass of each cell is $M_c = \iiint\limits_{\mathbf{x}\in V_c} \rho(\mathbf{x}) \, d^3x$;

$f(\mathcal{I}) \implies$ collection of orbits with unknown weights [to be determined]:

$$f(\mathcal{I}) = \sum_{k=1}^{N_{\mathrm{orb}}} w_k \, \delta(\mathcal{I} - \mathcal{I}_k)$$

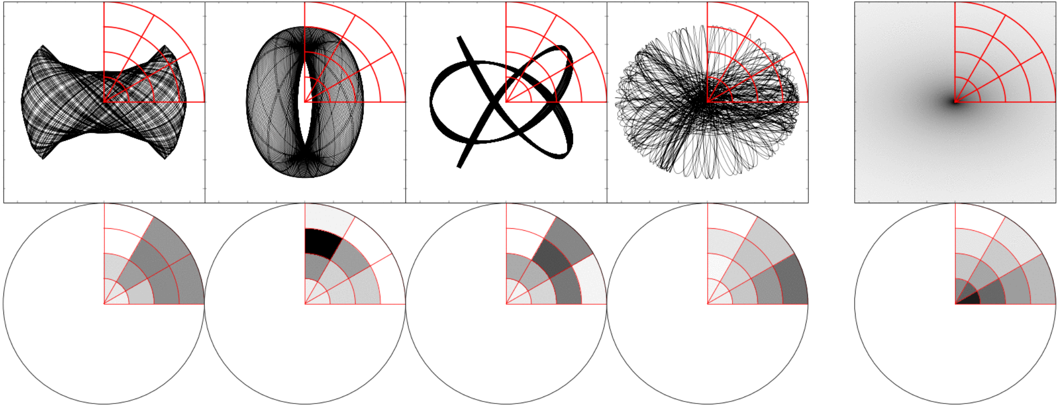
each orbit is a delta-function in the space of integrals of motion

adjustable weight of each orbit

Schwarzschild's orbit-superposition method: self-consistency

orbits in the model

target density



discretized orbit density

(fraction of time t_{kc} that k-th orbit spends in c-th cell)

discretized density

(mass M_c in grid cells)

For each c-th cell we require $\sum_k w_k \, t_{kc} = M_c$, where $w_k \geq 0$ is orbit weight

Schwarzschild's orbit-superposition method: fitting procedure

- Assume some potential $\Phi(\mathbf{x})$
  (e.g., from the deprojected luminosity profile plus parametric DM halo or SMBH)
- Construct the orbit library in this potential:
  for each k-th orbit, store its contribution to the discretized density profile t_{kc}, $c = 1..N_{\rm cell}$ and to the kinematic observables u_{kn}, $n = 1..N_{\rm obs}$
- Solve the constrained optimization problem to find orbit weights w_k:

$$\text{minimize } \chi^2 + \mathcal{S} \equiv \sum_{n=1}^{N_{\rm obs}} \left(\frac{\sum_{k=1}^{N_{\rm orb}} w_k\, u_{kn} - U_n}{\delta U_n} \right)^2 + \mathcal{S}(\{w_k\})$$

regularization term

observational constraints

their uncertainties

$$\text{subject to } w_k \ge 0, \quad k = 1..N_{\rm orb},$$

$$\sum_{k=1}^{N_{\rm orb}} w_k\, t_{kc} = M_c, \quad c = 1..N_{\rm cell}$$

density constraints (cell masses)

- Repeat for different choices of potential and find the one that has lowest χ^2

Schwarzschild's orbit-superposition method: implementations

Several commonly used independent implementations of the method:

- theoretical studies in triaxial geometry: Schwarzschild 1979, 1993; Pfenniger 1984; Statler 1987; Merritt & Fridman 1996; Siopis & Kandrup 2000; Vasiliev 2013
- spherical codes: Richstone & Tremaine 1984; Rix+ 1997; Jalali & Tremaine 2010; Breddels & Helmi 2013; Kowalczyk+ 2017
- axisymmetric: "Leiden" [van der Marel, Cretton, Cappellari, . . . – since 1998]
- axisymmetric: "Nukers" [Gebhardt, Richstone, Kormendy, . . . – since 2000]
- axisymmetric: "MasMod" [Valluri, Merritt, Emsellem – since 2004]
- triaxial/Milky Way bar: Zhao, Wang, Mao 1996, 2012
- triaxial: van den Bosch, van de Ven, de Zeeuw, Zhu, . . . – since 2008 $\Rightarrow$ "Dynamite"
- triaxial: "Forstand" [Vasiliev & Valluri 2020]

4. Made-to-measure (M2M) N-body models

Introduced by Syer & Tremaine 1996 as a way of constructing "tailored" N-body models satisfying some observational constraints.

Ingredients:

- N-particle system with time-dependent phase-space coordinates and weights $\{\mathbf{x}_k, \mathbf{v}_k, w_k\}\Big|_{k=1}^{N_{\rm body}}$ moving in a potential $\Phi(\mathbf{x})$
- Observational constraints U_n and their uncertainties δU_n, $n = 1..N_{\rm obs}$
- Model predictions for these observations: $V_n = \sum_{k=1}^{N_{\rm body}} w_k \, \underbrace{K_n(\mathbf{x}_k, \mathbf{v}_k)}_{\text{some predefined kernels}}$

Objective:

- minimize $\Omega \equiv \frac{1}{2}\sum_{n=1}^{N_{\rm obs}} \Delta_n^2 + \mathcal{S}\big(\{w_k\}\big)$,

  where $\Delta_n \equiv (V_n - U_n)/\delta U_n$ is the deviation in n-th constraint,

  $\mathcal{S}\big(\{w_k\}\big)$ is some measure of smoothness (regularization term),

  by varying the particle weights w_k.

Made-to-measure models

Objective is satisfied when $\dfrac{\partial \Omega}{\partial w_k} \equiv \sum_{n=1}^{N_{\rm obs}} \dfrac{\Delta_n\, K_n(\mathbf{x}_k, \mathbf{v}_k)}{\delta U_n} + \dfrac{\partial \mathcal{S}}{\partial w_k}$ is 0 for all k

Procedure:

repeat until $\Delta_n \approx 0$

- Evolve the N-body system in time: $\dot{\mathbf{x}}_k = \mathbf{v}_k,\ \dot{\mathbf{v}}_k = -\left.\dfrac{\partial \Phi}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_k}$
- Adjust the particle weights: $\dot{w}_k = -\dfrac{w_k}{\tau_{\rm ch}} \dfrac{\partial \Omega}{\partial w_k}$ (force-of-change)
- To reduce fluctuations, replace $\Delta_n(t)$ by a time-smoothed $\tilde{\Delta}_n(t) \equiv \dfrac{1}{\tau_{\rm sm}} \int_0^\infty \Delta_n(t-\tau)\, \exp\left(-\dfrac{\tau}{\tau_{\rm sm}}\right)\, d\tau$ in the above expression

* remove particles with too small w_k, split particles with too large w_k

* recompute the potential $\Phi(\mathbf{x})$ from particle positions and weights

Made-to-measure vs. Schwarzschild method

Both represent the DF as a large ensemble of δ-functions
with weights as free parameters in the model:

- N-body particles ($\sim 10^5 - 10^6$)
- time-average during evolution
- iteratively adjust weights (handmade gradient descent method)
- may adjust the potential during the fitting procedure
- live N-body system – easy to test the stability
- more expensive in CPU time

- orbits ($\sim 10^3 - 10^5$)
- compute entire orbits beforehand
- solve a large-scale constrained optimization problem by black-box routines
- potential fixed in advance (need to construct a new orbit library each time a new potential is chosen)
- need to convert orbit library into an N-body model first

Made-to-measure / tailored N-body models in practice

Several independent implementations of the method:

- *NMAGIC* (MPI/Garching group, Gerhard et al.):
  Milky Way bar/bulge [Bissantz+ 2014; Portail+ 2015, 2017],
  Andromeda bar/bulge [Blaña Díaz+ 2018],
  external galaxies [de Lorenzi, Morganti, Das, . . . 2007+]
- Milky Way bar/bulge; M87 halo [Long, Mao, Shen, Zhu, . . . 2010+]
- *PRIMAL*: Milky Way disk and bar [Hunt & Kawata 2013+]
- "theoretical" (no obs.applications) code of Dehnen 2009
- Deg 2010 (thesis, unpublished)
- Malvido & Sellwood 2014
- [non-M2M] iterative method of Rodionov & Athanassoula 2009
- [non-M2M] iterative method of Yurin & Springel 2014

Summary of modelling methods

method	ensures $f \geq 0$	smooth DF	assumptions on functional form	geometry, rotation	cost
Jeans	$-$	n/a	parametric β, vel.alignment	Sph, Axi	low
DF	$+$	$+$	func. form of DF	S,A,Tri	varies
		$\pm$	nonparametric	S,A	high (?)
Schwarzschild	$+$	$-$	–"–	S,A,T,Ω	high
Made-to-measure	$+$	$-$	–"–	S,A,T,Ω	high