Astr 511: Galaxies as galaxies

University of Washington

Mario Jurić & Željko Ivezić

Lecture 9:

Dynamics III: Building Galaxy Models

Using the CBE to Build Galaxy Models

- Binney & Tremaine, §4.

- Barnes, Galaxies, https://www.ifa.hawaii.edu/~barnes/ast626_09/scbe.pdf

Integrals of Motion

An **integral of motion** is any function $I(\mathbf{r}, \mathbf{v})$ of the phase space coordinates $(\mathbf{r}, \mathbf{v})$ that satisfies:

$$\frac{d}{dt} I(\mathbf{r(t)}, \mathbf{v(t)}) = 0 \tag{1}$$

along *all* orbits $(\mathbf{r(t)}, \mathbf{v(t)})$. We're already familiar with a few integrals of motion – e.g., the total energy, $E = 1/2|\mathbf{v}|^2 + \Phi(r)$ is an integral in time-independent potentials, the angular momentum, $\mathbf{L}$, is an integral in spherical systems, and the $\hat{z}$ component of the angular momentum, L_z is an integral in axisymmetric systems.

Jeans Theorem

Any integral of motion is the solution of the time-independent CBE. Proof:

$$\frac{dI}{dt} = 0 = \frac{\partial I}{\partial \mathbf{r}} \cdot \frac{d\mathbf{r}}{dt} + \frac{\partial I}{\partial \mathbf{v}} \cdot \frac{d\mathbf{v}}{dt} = \ \mathbf{v} \cdot \frac{\partial I}{\partial \mathbf{r}} - \nabla\Phi \cdot \frac{\partial I}{\partial \mathbf{v}} = 0 \qquad (2)$$

and this is the same condition satisfied by the steady state ($\frac{\partial f}{\partial t} = 0$) solution of the CBE:

$$\frac{\partial f}{\partial t} + \mathbf{v}\nabla f - \nabla\Phi \frac{\partial f}{\partial \mathbf{v}} = 0 \qquad (3)$$

when we substitute $f \to I$. Q.E.D.

Integrals of Motion and the CBE

Any steady-state solution of the CBE depends on the phase space coordinates only through integrals of motion in the given potential, and any function of the integrals yields a steady state solution of the CBE. Proof:

- $\rightarrow$ Suppose f is a steady state solution of the CBE. Then, as we've seen on the previous slide, it will satisfy the condition to be an integral of motion.

- $\leftarrow$ Conversely, if I_1 through I_n are n integrals, and if f is any function of n variables, then:

$$\frac{d}{dt} f[I_1(\mathbf{x},\mathbf{v}),...,I_n(\mathbf{x},\mathbf{v})] = \sum_{m=1}^{n} \frac{\partial f}{\partial I_m} \frac{dI_m}{dt} = 0 \tag{4}$$

  so f satisfies the CBE (hint: write out df/dt to see that another form of the CBE is $df/dt = 0$).

A recipe to build galaxies!

Why is this interesting? Because it gives us a **recipe to build (idealized) models galaxies**, and collisionless stellar systems in general, without resorting to N-body techniques!

These models can frequently capture many of the qualitative properties we observe in real galaxies (density profiles, kinematical properties).

They may also be starting points for N-body simulations.

Example: Isotropic models of spherical galaxies

In the case of isotropic (i.e., no preferred direction) models of spherical galaxies, the distribution function $f(\mathbf{x}, \mathbf{v}) = f(E)$ can only be a function of specific energy, $E = v^2/2 + \Phi(\mathbf{r})$.

Recipe ingredients:

#1 The Poisson equation for a spherically symmetric system:

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = 4\pi G\rho(r) \tag{5}$$

(note: we often take the boundary condition that $\Phi \to 0$ as $r \to \infty$)

Example: Isotropic models of spherical galaxies

In the case of isotropic (i.e., no preferred direction) models of spherical galaxies, the distribution function $f(\mathbf{x}, \mathbf{v}) = f(E)$ can only be a function of specific energy, $E = v^2/2 + \Phi(\mathbf{r})$.

Recipe ingredients:

#2 The expression for density (an integral of the DF over the (isotropic) velocity field):

$$\rho = 4\pi \int_0^{v_e} dv v^2 f(v^2/2 + \Phi(r)) \tag{6}$$

where $v_e = \sqrt{-2\Phi(r)}$ is the escape velocity at radius r. Alternatively, we can switch to energy as the independent variable and we have:

$$\rho = 4\pi \int_{\Phi}^{0} dE \sqrt{2E - 2\Phi} f(E) \tag{7}$$

From f to ρ: Plummer Model

Given any functional form for $f(E)$ which is non-negative for all $E < 0$, use either (6) or (7) to calculate the function $\rho(\Phi)$, and insert the result into (5).

Example: **Plummer model**:

$$f(E) = \begin{cases} F \cdot (-E)^{7/2}, & E < 0, \\ 0, & E \geq 0, \end{cases}$$

where F is a constant. We use (7) to obtain $\rho(\Phi)$ and plug the result into (5):

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = K(-\Phi)^5 \tag{8}$$

where K is another constant. The solution gives us a model with the density profile:

$$\rho(r) = \frac{3M}{4\pi a^3}\left(1 + \frac{r^2}{a^2}\right)^{-5/2} \tag{9}$$

where M is the total mass and a is the characteristic scale (related to the constant F).

From f to ρ: King Model

The Plummer model was originally devised to describe observations of star cluster.

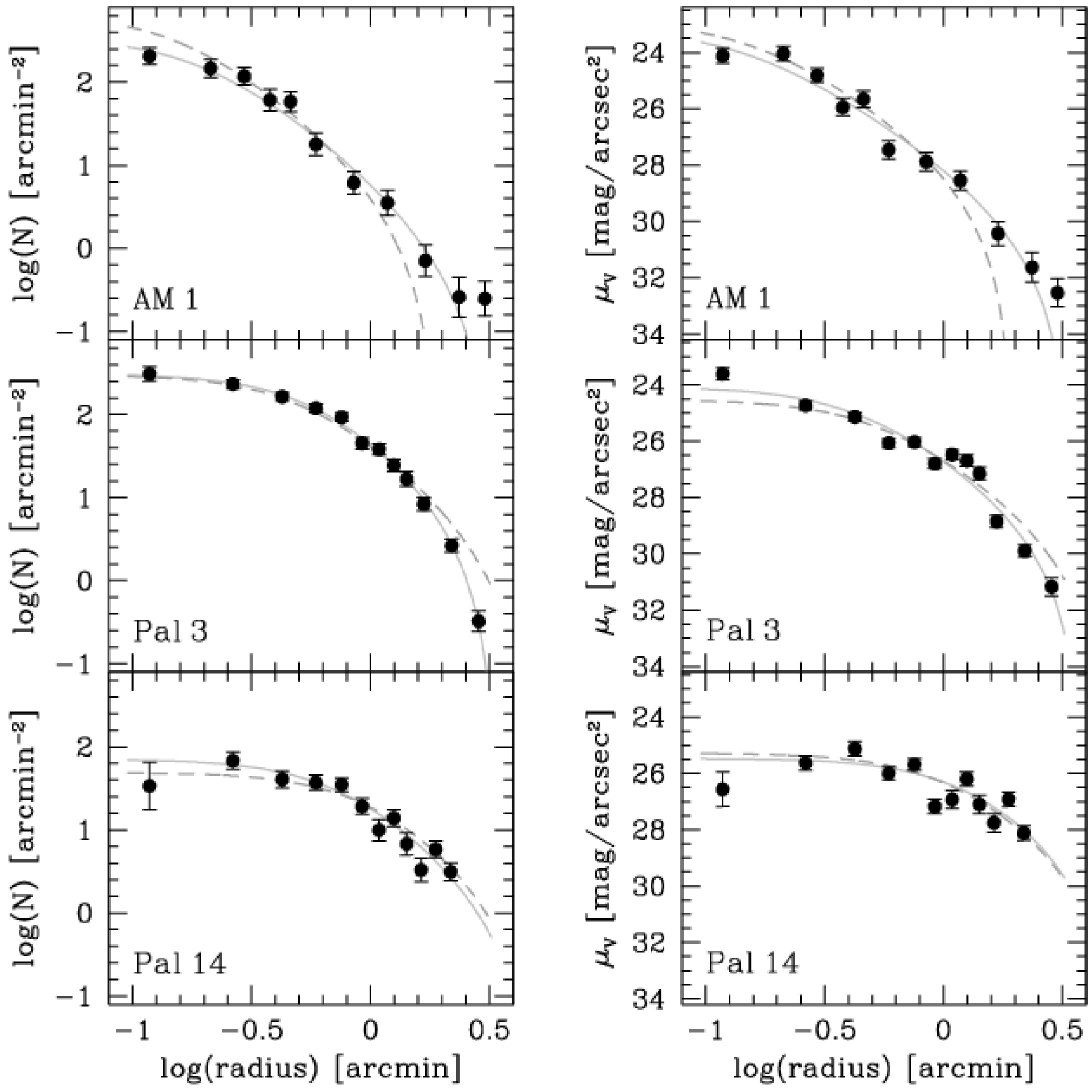
Another popular model, originally introduced by Ivan King (King 1966) to explain the observations of globular clusters is the **King model**:

$$f(E) = \begin{cases} \rho_1 (2\pi\sigma)^{-3/2} \exp(-(E - E_0)/\sigma^2 - 1), & E < E_0, \\ 0, & E \geq E_0, \end{cases}$$

where ρ_1, σ, and E_0 are parameters of the model. Just like with the Plummer model, the Poisson equation can be solved to derive $\rho(r)$ (see Eq. 4.111 in BT).

The three parameters above can be related to the central brightness, Σ_0, the **core radius** r_c, at which the brightnes drops to 50% of central, and **tidal radius**, r_t, at which the brightness vanishes.

Example: King Profiles



Hilker (2005)

From ρ to f

Given any functional form for $\rho(r)$ which is non-negative everywhere, it can be shown that:

$$f(E) = \frac{1}{\sqrt{8}\pi^2}\frac{d}{dE}\int_E^0 d\Phi \frac{\rho'(\Phi)}{\sqrt{\Phi - E}} \tag{10}$$

where $\rho'(Phi) = d\rho(\Phi)/d\Phi$ and $\rho(\Phi)$ can be obtained by solving the Poisson equation and inverting (see BT § 4, for details).

The above equation is useful in constructing isotropic models of spherical systems with known density profiles (e.g., inferred from observations). In some cases, these can be derived analytically (e.g., Jaffe 1983, or Henrquist 1990).

Generally, it can be solved numerically – it's frequently used to set up initial conditions for spherical systems in N-body calculations.

Comparison of various profiles

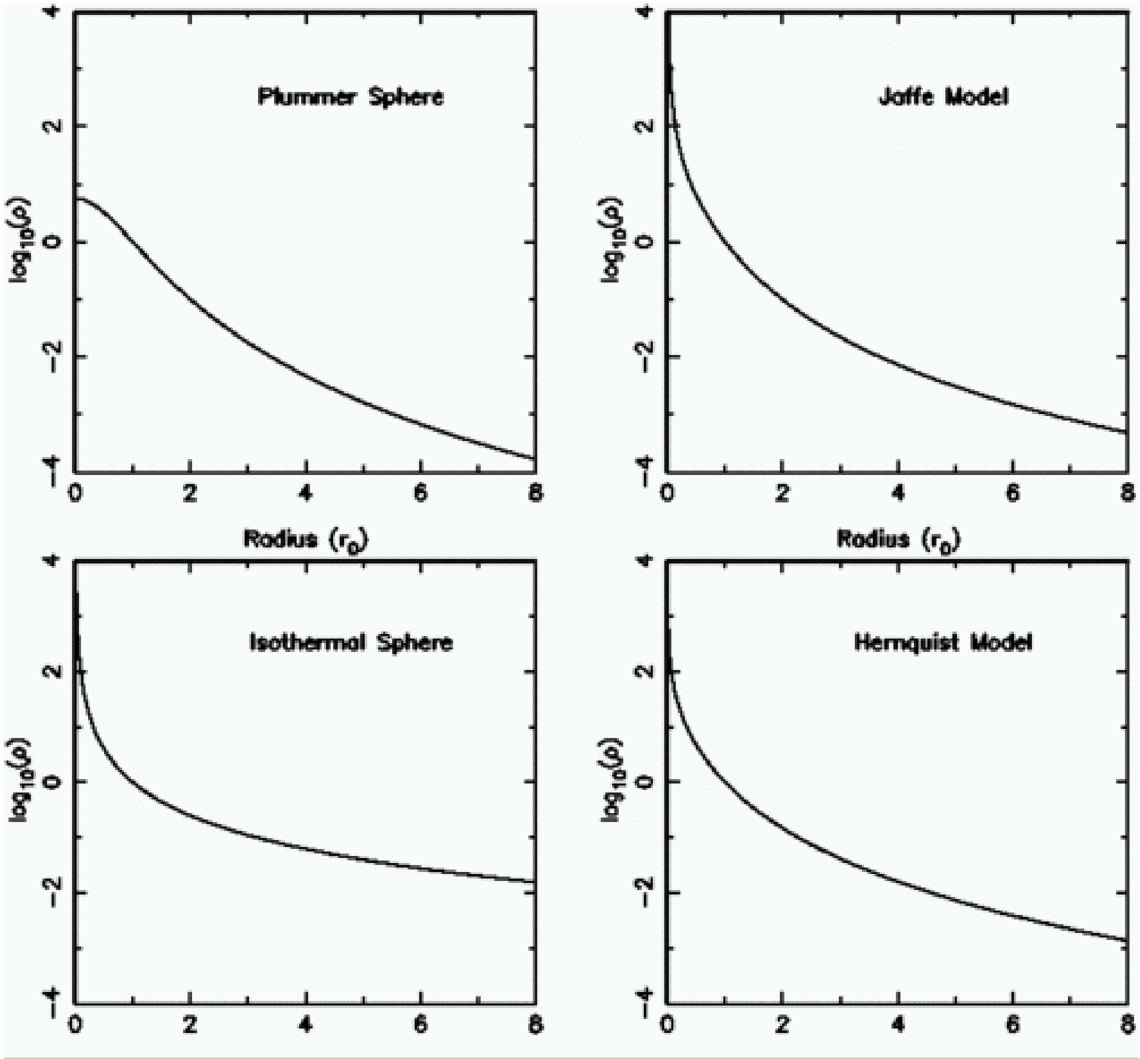
The profiles derived above supply us with convenient starting points for building models of stellar systems (galaxies and globular clusters).

- **Plummer** model has a finite density in the core, and falls of as r^{-5}. It matches globular clusters reasonably well.

- **King** model has a finite core density and outer radius, and is a good description of globular clusters.

- **Hernquist & Jaffe** models both fall of as r^{-4} at large radii, with different slopes in the core (Jaffe is steeper). This fall-off at large radii is observed in elliptical galaxies, and theoretically well motivated (violent relaxation; BT §4.10.2).

Comparison of various profiles



plot by Chris Flynn, Swinburne

More about spherically symmetric profiles

Plummer, Jaffe and Hernquist are all realizations of a broader class of **double power-law** density profiles described by:

$$\rho(r) \propto \frac{1}{r^{\gamma}(1 + r^{1/\alpha})^{(\beta-\gamma)\alpha}} \tag{11}$$

These behave as $\rho \propto r^{-\gamma}$ at $r \ll 1$ and $\rho \propto r^{-\beta}$ at $r \gg 1$. Many of the well known density profiles fall into this category:

(α, β, γ)	Name	Reference
$(1, 3, 1)$	NFW Profile	Navarro, Frenk & White, 1997, ApJ, 490, 493
$(1, 4, 1)$	Hernquist Profile	Hernquist, 1990, ApJ, 356, 359
$(1, 4, 2)$	Jaffe Profile	Jaffe, 1983, MNRAS, 202, 995
$(1, 4, \frac{3}{2})$	Moore Profile	Moore et al., 1999, MNRAS, 310, 1147
$(\frac{1}{2}, 2, 0)$	Modified Isothermal Sphere	Sacket & Sparke, 1990, ApJ, 361, 409
$(\frac{1}{2}, 3, 0)$	Modified Hubble Profile	Binney & Tremaine, p. 39
$(\frac{1}{2}, 4, 0)$	Perfect Sphere	de Zeeuw, 1985, MNRAS, 216, 273
$(\frac{1}{2}, 5, 0)$	Plummer Model	Plummer, 1911, MNRAS, 71, 460

Table from Frank van den Bosch, Yale

Building Arbitrary Equilibrium Galaxies: Schwarzschild's Method

The examples given previously rely on our ability to derive closed-form solutions to equations such as (10). Some of these solutions lead to models that are useful in that they correspond to profiles seen in some galaxies and globular clusters.

However useful that is, the analysis of the previous section does not give us a general, *practical*, method to derive a steady-state distribution function for an *arbitrary* distribution of density, $\rho(\mathbf{x})$.

We'll now describe a semi-numerical technique due to Schwarzschild that promises to do just that. It is commonly referred to as **Schwarzschild's method** (Schwarzschild 1979).

Galaxy as a superposition of stars on orbits

A galaxy, as observed (as *imaged*) is nothing more than a superposition of light coming from $\sim 10^{10}$ stars moving on orbits within its potential.

If we imagine the orbits as *paths* along which stars traverse the galactic potential, we quickly see that some may be more frequented than others. For example an orbit that takes a star around the center of the Milky Way that's parallel to the plane of the disk is (roughly) ~ 1000 times more likely to have a star than an orbit perpendicular to it.

Schwarzschild's key insight was that if we developed a library of all (or approximately all) orbits that an observed density distribution admits, we could try to construct a *superposition* that reproduces the observed density $\rho(\mathbf{x})$.

Schwarzschild's Method in a Nutshell

1. Given $\rho(\mathbf{x})$ find the potential $\Phi(\mathbf{x})$.

2. Construct a large library of orbits admitted by $\Phi(\mathbf{x})$

   - Choose a large number (N) of initial conditions, each assigned to a particle.

   - Integrate (numerically) each one of these over a long time $t \gg t_{\text{cross}}$, tracing out its orbit.

   - *Pixelize* each orbit onto a grid of K cells, where the contribution of the orbit to each pixel will be proportional to the *time* a particle on the orbit spends in the pixel.

3. Find a linear combination of orbits from the orbit library, each with non-negative weights, that reproduces $\rho(\mathbf{x})$ when added together.

Schwarzschild's Method: #1. Density and potential

As the density distribution is known, the potential is obtained by solving the Poisson equation:

$$\nabla^2 \Phi = -4\pi G \rho(\mathbf{x}) \tag{12}$$

No closed form solutions of this equation exist in general, but we can solve it numerically on a finite grid. For more, see the discussion of *Poisson solvers* in BT §2.9).

This gives us the potential $\Phi(\mathbf{x})$ generated by the matter distribution $\rho(\mathbf{x})$, within which we can trace out the motions of test particles (revealing their orbits).

Schw. Method: #2. Construct the orbit library

Our next goal is to construct a *library* of all (at least in principle) orbits that are possible in the derived potential. We don't have to store the actual orbits (i.e., $\mathbf{x}(t)$ paths) into our library; what matters is their *contribution to the observables*, typically, the density.

We therefore begin by **gridding** the space occupied by the galaxy, subdividing it into K cells (voxels). As our particle traverses the galaxy, it will spend a varying amount of time δt_j in each of our K cells. Note that this includes $\delta t_j = 0$, i.e., some cells may never be visited).

The contribution to the observed density in each cell will be proportional to the time the particle spends in the cell.. We store these contributions (in the most naive implementation, as a 3D data cube for each orbit). The contribution to other observables (e.g., the LOSVD) could be computed and stored as well.

Example: A regular non-resonant orbit in an axisymmetric potential.

1.2 Catastrophes in Orbital Observables **3**

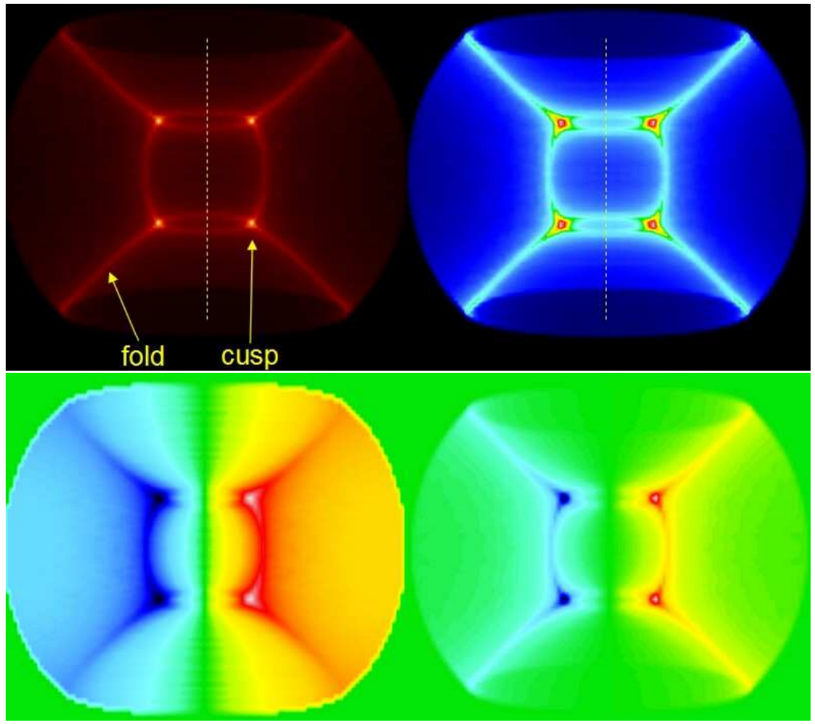


Fig. 1.2. *Top Panels:* projected surface brightness of a regular non-resonant orbit in an axisymmetric potential. On the left a single-color linear colormap was used, while in all the other plots the same linear colormap of Fig. 1.4 was adopted. *Bottom Panels:* on the left the mean velocity of the above orbit is shown. On the right the velocity was weighted with the surface brightness: this is what enters into the computation of the observed mean velocity.

Cappellari et al (2003); astro-ph:0302274v1

Example: A regular non-resonant orbit in a triaxial potential.

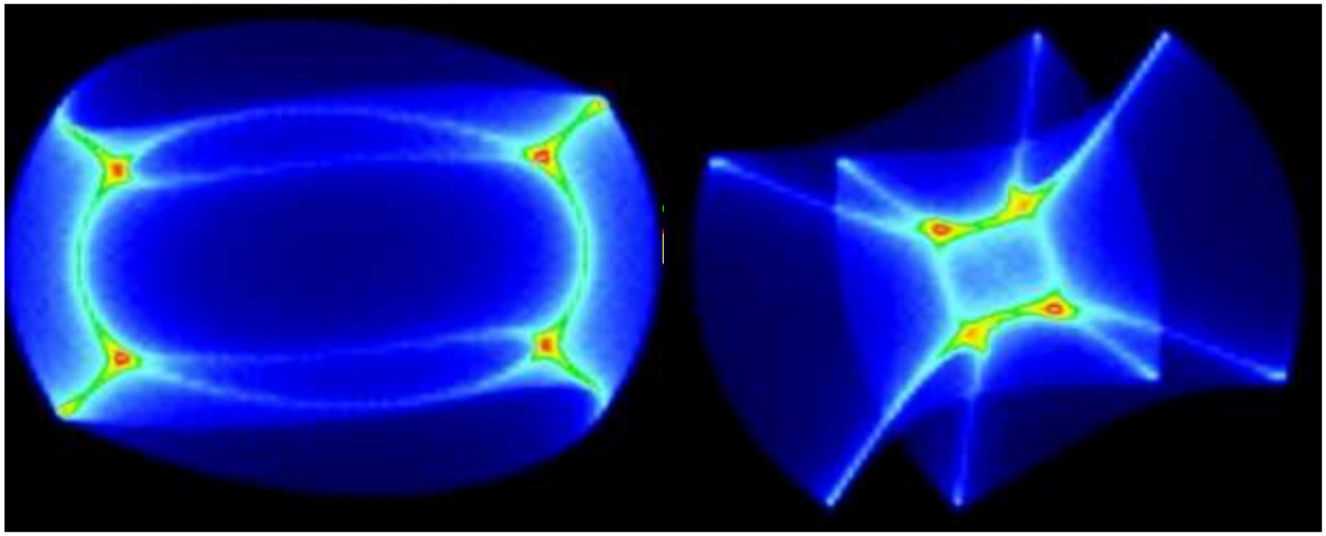


Fig. 1.3. Projected surface brightness of a regular non-resonant tube orbit (left) and box orbit (right) in a triaxial potential.

Cappellari et al (2003); astro-ph:0302274v1

Schwarzschild's Method: #3. Find a superposition that corresponds to the observations

Let's assume the only observable we're trying to reproduce is the density, $\rho(\mathbf{x})$. The contribution to cell j of volume V_j is equal to:

$$m_j = \rho(\mathbf{x_j})V_j \tag{13}$$

. Let us assume that each orbit in our library is populated by a large number of stars, uniformly distributed in orbital phase, and that the total mass of stars on orbit i is $w_i M$, were w_i is the weight, and M is the total mass of the galaxy. Therefore, the mass in cell j will be equal to the sum of the contributions from orbits that cross it, or:

$$m_{j,\text{constructed}} = M \sum_{i=1}^{N} w_i p_{ij} \tag{14}$$

Schwarzschild's Method: #3. Find a superposition that corresponds to the observations

Therefore, we're looking for a solution of:

$$0 = \Delta_j \equiv m_j - M \sum_{i=1}^{N} w_i p_{ij} \tag{15}$$

.

This is a set of K linear equations for the N unknown weights w_i. Note that the condition that $\sum_j m_j = M$ implies that $\sum_j w_j = 1$.

In principle, by choosing $N = K$ would make it possible to trivially solve this sytem; however, for the model to be physical, all w_i have to be positive. In practice, this choice leads to some values of w_i being negative.

Schwarzschild's Method: #3. Find a superposition that corresponds to the observations

The way we work around this problem is to take $N \gg K$, i.e., we observe many more orbits than spatial cells. In that case, the points that satisfy the equations form an $(N - K)$ dimensional subspace of the N dimensional space of weight vectors $\mathbf{w}$.

A physically meaningful solution exists if this subspace passes through the region where all $w_i > 0$. It's possible there are no solutions – i.e., the input $\rho(x)$ cannot be explained as a steady-state self-gravitating system.

But note that if a solution exists, there will **generally be infinitely many solutions** $\mathbf{w}$. Every one corresponds to an acceptable galaxy model.

The Objective Function

That is, **we either find no solution, or an infinite set of possible solutions**.

We must bring additional information to the table, introduce *additional constraints* to narrow down the set of acceptable solutions, perhaps even down to a single one. This is normally done by choosing the solution that maximizes some **objective function** $\phi(\mathbf{w})$.

A simple choice is one where $\phi(\mathbf{w})$ is linear:

$$\phi(\mathbf{w}) = \sum_i \phi_i w_i \tag{16}$$

Solving the system (15) under these constraints is an exercise in a mathematical optimization technique called **linear programming**. See `http://www.math.ucla.edu/~tom/LP.pdf` for a consise introduction.

Choosing the Objective Function

The **objective function** is how we impose additional constraints (that is, **import additional knowledge about the system**) into the modeling procedure.

While linear objective function is straightforward to write and apply, it is not optimal. For one, it's not obvious how to physically motivate the maximization of some linear function of the weights. Schwarzschild himself pretty much picked his ϕ_i at random, though his goal was to prove that there *exist* self-consistent models of triaxial galaxies. There are also other, more technical reasons, why a linear $\phi(\mathbf{w})$ is not optimal (see BT for details).

Other options: maximizing the entropy, $\phi(\mathbf{w}) = S \equiv -\sum_i w_i \ln w_i$ (see BT §4.10.1), or a quadratic objective function of the form $\phi(\mathbf{w}) = -\sum_i w_i^2/W_i$ where $W_i > 0$. Solving with the latter is an exercize in **quadratic programming**; the physical meaning is to find the values of w_i that are closest to W_i while satisfying $\rho(\mathbf{x})$. Into W_i we can encode any priors we have on the expected orbital structure of the galaxy being studied.

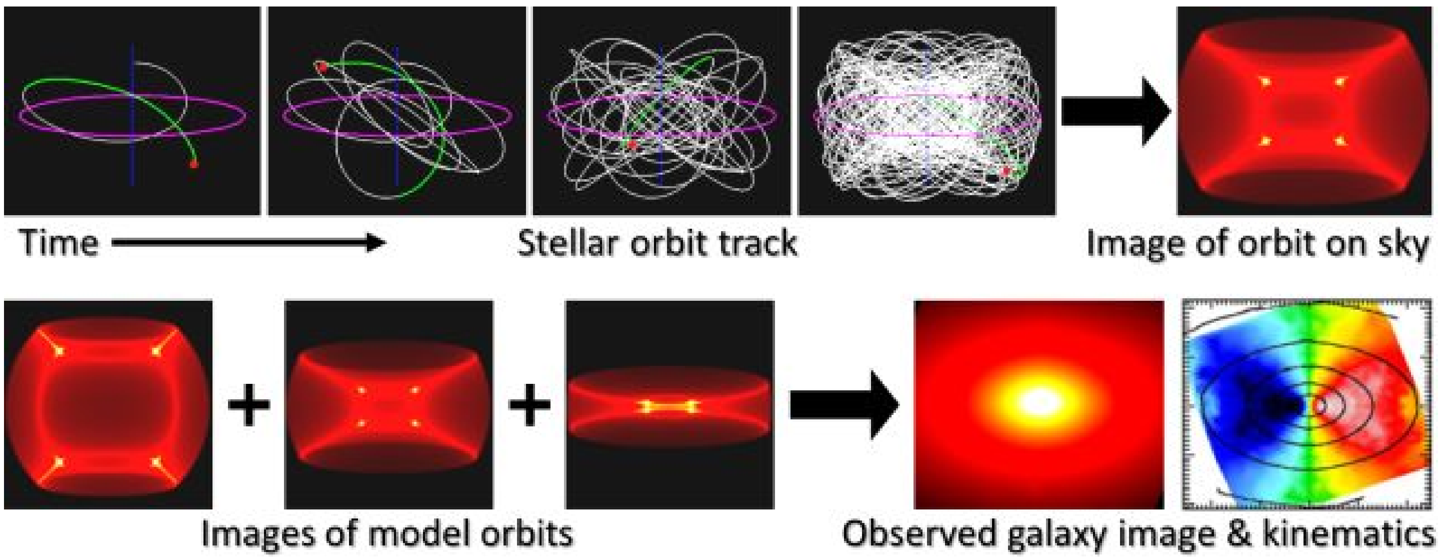
Extensions of Schwarzschild's Method

A particularly fruitful extension of Schwarzschild's method is to model kinematic data (e.g., such as those coming from integral field units).

These observations result in measurements of LSOVD at pixels within the galaxies' footprints, and the LOSVD at each point is a linear function of the orbit weights, w_i. So the function to *minimize* to is simply χ^2.

In other words, $\phi(\mathbf{w}) = -\chi^2$. This can be solved using quadratic programming techniques. It has been applied to the understanding of dynamics of early type galaxies, as well as the search for central black holes.

Schwarzschild's Method: Building an ETG constrained by IFU observations



Cappellari (2015); http://adsabs.harvard.edu/abs/2014arXiv1410.7329C

Applications of Schwarzschild Modelling

Schwarzschild modelling has been successfuly applied to a variety of topics. Two more recent ones:

- Search for massive black holes at the centers of galaxies (e.g., Richtone & Tremaine 1985; van der Marel et al. 1998; Gebhardt et al. 2003), leading to the discovery of $M - \sigma$ relation.

- Modeling of the large-scale dynamics of early type galaxies (e.g., Capellari et al. 2006), constraining the mass densities and orbital structure.

The Importance of a Good Orbit Library

It's important that the orbit library is fairly complete, i.e., that it combines a sufficiently wide variety of types of orbits, with a reasonably dense sampling of the phase space:

- If this is not the case, there may not be even a single solution to the constraint equaton (15) with non-negative weights and secondly,

- we often wish to explore a full set of allowable models, to understand better understand the limitations and the constraining power of our dataset.

In practice, generating a good orbit library requires some intuition and experience. This part is as much art as science.