



**Action-based
self-consistent models**

Eugene Vasiliev

Oxford University

Gaia Challenge, Stockholm, October 2016

Fundamental equations

1. Collisionless Boltzmann equation:

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

gravitational potential

distribution function

(Assumption: a galaxy is a collisionless system in a steady state)

2. Poisson equation:

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

density

(Assumption: Newtonian gravity)

3. The link:

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

(Assumption: self-consistency)

Ways of solving the equations

1. Collisionless Boltzmann equation:

Jeans theorem states that the DF may depend only on the integrals of motion

$$f = f(\mathcal{I}(\mathbf{x}, \mathbf{v})), \quad \mathcal{I} = \{E, L, \dots\}.$$

← depend on the potential Φ

2. Poisson equation:

$$\Phi(\mathbf{x}) = - \iiint d^3x' G \rho(\mathbf{x}') \times \frac{1}{|\mathbf{x} - \mathbf{x}'|}.$$

3. The link:

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

Actions as integrals of motion

- ▶ One may use any set of integrals of motion, **but** actions are special:

- ▶ For bounded multiperiodic motion, actions are defined as

$$J = \frac{1}{2\pi} \oint \mathbf{p} \, d\mathbf{x}, \text{ where } \mathbf{p} \text{ are canonically conjugate momenta for } \mathbf{x}$$

- ▶ Action/angle variables $\{\mathbf{J}, \boldsymbol{\theta}\}$ are the most natural way of describing the motion: from Hamilton's equations we have

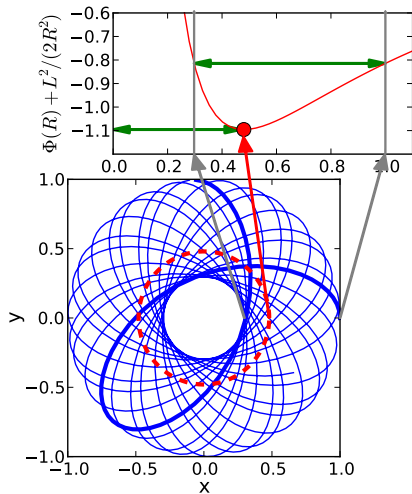
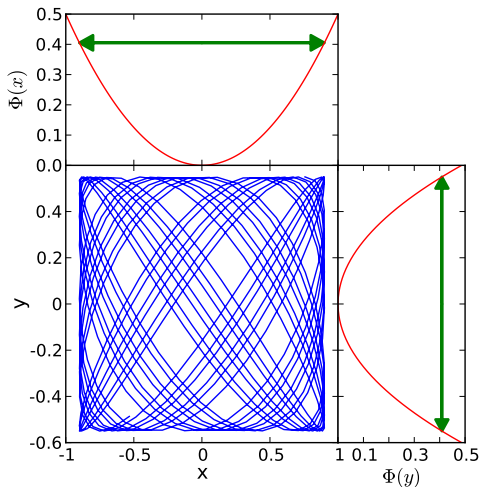
$$\frac{dJ_i}{dt} = -\frac{\partial H}{\partial \theta_i} = 0 \text{ (actions are integrals of motion), and}$$

$$\frac{d\theta_i}{dt} = \frac{\partial H}{\partial J_i} \equiv \Omega_i \text{ (angles increase linearly with time);}$$

here $H(\mathbf{J})$ is the Hamiltonian and $\boldsymbol{\Omega}(\mathbf{J})$ are the frequencies.

Examples of action/angle variables

The meaning of the action/angle variables may vary for different classes of orbits, but generally describes the extent of oscillation in a particular direction.



Pros and cons of action/angle variables

- + Most natural description of motion (angles change linearly with time); once \mathbf{J} and $\mathbf{\Omega}$ have been found, orbit computation is trivial.
- + Possible range for each action variable is $[0..\infty)$ or $(-\infty..\infty)$, independently of the other ones (unlike E and L , say).
- + Canonical coordinates: the volume of phase space $d^3x d^3v = d^3J d^3\theta$.
- + Actions are adiabatic invariants (are conserved under slow variation of potential).
- + Serve as a good starting point in perturbation theory.
- No general way of expressing the Hamiltonian $H \equiv \Phi(\mathbf{x}) + \frac{1}{2}\mathbf{v}^2$ in terms of actions (i.e., solving the Hamilton–Jacobi equation).
- Not easy to compute them in a general case.
- + Efficient methods for conversion between $\{\mathbf{x}, \mathbf{v}\}$ and $\{\mathbf{J}, \boldsymbol{\theta}\}$ have been developed in the last few years.

“Classical” methods

- ▶ Spherical systems:

two of the actions can be taken to be the *azimuthal action*

$J_\phi \equiv L_z$ and the *latitudinal action* $J_\vartheta \equiv L - |L_z|$;

the third one (the *radial action*) is given by a 1d quadrature:

$$J_r = \frac{1}{\pi} \int_{r_{\min}}^{r_{\max}} dr \sqrt{2[E - \Phi(r)] - L^2/r^2},$$

where r_{\min} , r_{\max} are the peri- and apocentre radii.

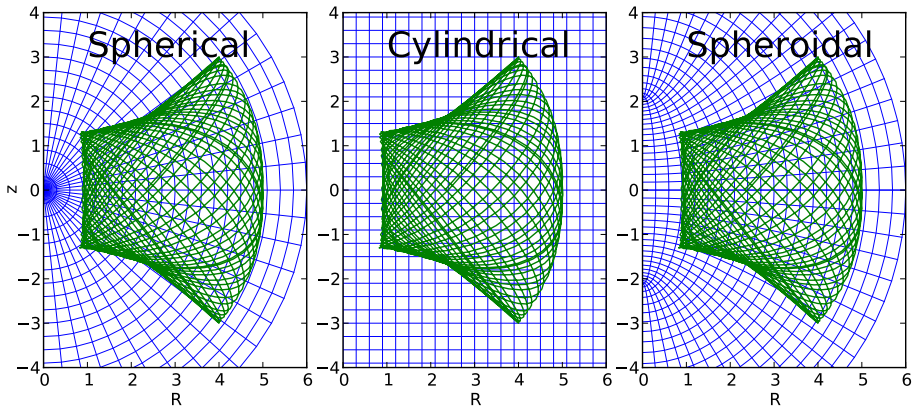
Angles are given by 1d quadratures. For special cases (the isochrone potential, and its limiting cases – Kepler and harmonic potentials), these integrals are computed analytically.

Note: a related concept in celestial mechanics are the Delaunay variables.

- ▶ Flattened axisymmetric systems – the **epicyclic approximation**: motion close to the disc plane is nearly separable into the in-plane motion (J_ϕ and J_r as in the spherical case) and the vertical oscillation with a fixed energy E_z in a nearly harmonic potential (J_z).

State of the art: Stäckel fudge

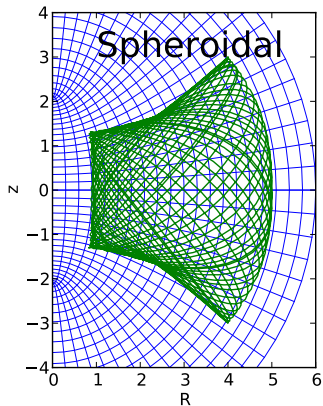
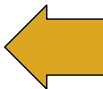
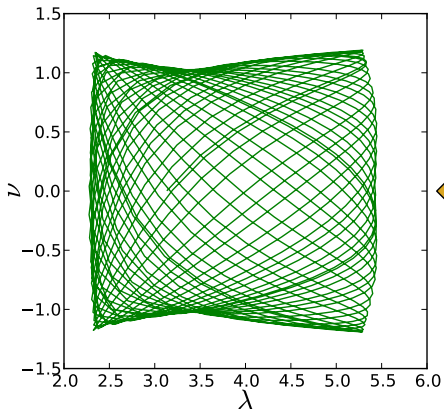
Fact: orbits in realistic axisymmetric galactic potentials are much better aligned with prolate spheroidal coordinates.



State of the art: Stäckel fudge

Fact: orbits in realistic axisymmetric galactic potentials are much better aligned with prolate spheroidal coordinates.

One may explore the assumption that the motion is separable in these coordinates (λ, ν) .



Stäckel fudge (Binney 2012)

The most general form of potential that satisfies the separability condition is the Stäckel potential¹: $\Phi(\lambda, \nu) = -\frac{f_1(\lambda) - f_2(\nu)}{\lambda - \nu}$.

The motion in λ and ν directions, with canonical momenta p_λ, p_ν , is governed by two separate equations:

$$2(\lambda - \Delta^2) \lambda p_\lambda^2 = \left[E - \frac{L_z^2}{2(\lambda - \Delta^2)} \right] \lambda - [I_3 + (\lambda - \nu)\Phi(\lambda, \nu)],$$

$$2(\nu - \Delta^2) \nu p_\nu^2 = \left[E - \frac{L_z^2}{2(\nu - \Delta^2)} \right] \nu - [I_3 + (\nu - \lambda)\Phi(\lambda, \nu)].$$

Under the approximation that the separation constant I_3 is indeed conserved along the orbit, this allows to compute the actions:

$$J_\lambda = \frac{1}{\pi} \int_{\lambda_{\min}}^{\lambda_{\max}} p_\lambda d\lambda, \quad J_\nu = \frac{1}{\pi} \int_{\nu_{\min}}^{\nu_{\max}} p_\nu d\nu.$$

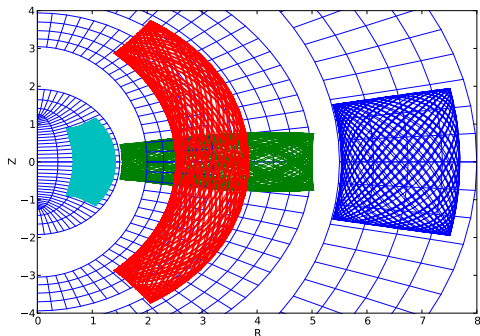
¹Note that the potential of the Perfect Ellipsoid (de Zeeuw 1985) is of the Stäckel form, but it is only one example of a much wider class of potentials.

Stäckel fudge in practice

A rather flexible approximation: for each orbit, we have the freedom of using two functions $f_1(\lambda)$, $f_2(\nu)$ (directly evaluated from the actual potential $\Phi(R, z)$) to describe the motion in two independent directions.

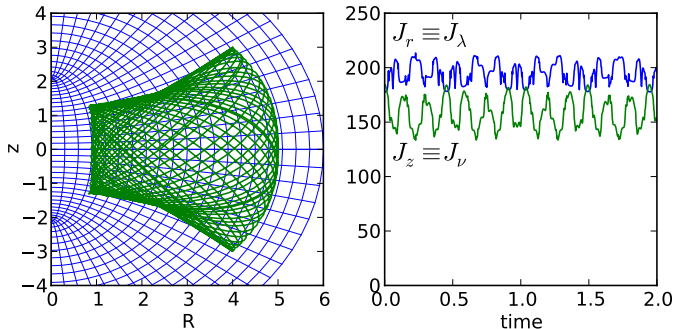
These functions are different for each orbit (implicitly depend on E, L_z, l_3).

Moreover, we may choose the interfocal distance Δ of the auxiliary prolate spheroidal coordinate system for each orbit independently.



Accuracy of Stäckel fudge

Accuracy of action conservation using the Stäckel fudge:
 $\lesssim 1\%$ for most disc orbits, $\lesssim 10\%$ even for high-eccentricity orbits.



Self-consistent models of spherical systems

- ▶ Integrals of motion:

energy $E = \Phi(r) + v^2/2$, angular momentum $L = |\mathbf{x} \times \mathbf{v}|$.

- ▶ Gravitational potential:

$$\Phi(\mathbf{x}) \equiv \Phi(r) = -4\pi G \left[r^{-1} \int_0^r dr' \rho(r') r'^2 + \int_r^\infty dr' \rho(r') r'^{-1} \right].$$

- ▶ Distribution function:

$f(E)$ (isotropic) or $f(E, L)$ (general anisotropic);
in the isotropic case

$$\begin{aligned} \rho(r) &= \int_0^{v_{\text{esc}}} 4\pi v^2 dv f(\Phi(r) + v^2/2) \\ &= \int_{\Phi(r)}^0 dE 4\pi f(E) \sqrt{2[E - \Phi(r)]}. \end{aligned}$$

Two possible approaches

1. From f to ρ :

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\Phi(r)}{dr} \right) = 4\pi G \iiint d^3v f(\Phi(r) + v^2/2, |\mathbf{x} \times \mathbf{v}|)$$

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

Examples: Plummer, King, lowered isothermal models, etc.

2. From ρ to f : first compute $\Phi(r)$ and its inverse $r(\Phi)$;

$$\rho(r) = \int_{\Phi(r)}^0 dE 4\pi f(E) \sqrt{2[E - \Phi(r)]} \implies$$
$$f(E) = \frac{d}{dE} \int_E^0 d\Phi \frac{d\rho(r(\Phi))}{d\Phi} \frac{1}{4\pi \sqrt{2[E - \Phi(r)]}}$$

(Eddington inversion), or its generalizations for anisotropic $f(E, L)$ (Ossipkov–Merritt, Cuddeford, etc.)

Examples: isochrone, Hernquist model, etc.

Axisymmetric systems

- ▶ Integrals of motion:

$$E = \Phi(R, z) + v^2/2, \quad L_z = R v_\phi,$$

third integral $I_3 = ?$ (known analytically in special cases).

- ▶ Gravitational potential:

$\Phi(R, z) = ?$ (can be reduced to a 1d integral over $\rho(R, z)$ in special cases, e.g., for $\rho(R^2 + z^2/q^2)$ or $\rho_R(R)\rho_z(z)$, etc.)

- ▶ Distribution function:

$f(E, L_z)$ (two-integral) or $f(E, L_z, I_3)$ (general three-integral).

Two approaches are now entirely different from the spherical case.

1. From f to ρ :

Iterative approach [e.g., Prendergast&Tomer 1970];

2. From ρ to f :

– Contour integral method for $f(E, L_z)$ [Hunter&Qian 1993];

– Non-parametric discrete representation of distribution function.

Models with a prescribed distribution function

Goal: $f(\mathbf{J}) \implies \rho(\mathbf{x})$

given a particular expression for the distribution function $f(\mathbf{J})$,
construct the corresponding self-consistent potential/density pair.

- ▶ Assume an initial guess for $\Phi(\mathbf{x})$;
- ▶ Construct the mapping $\{\mathbf{x}, \mathbf{v}\} \implies \{\mathbf{J}, \boldsymbol{\theta}\}$ in this potential;
- ▶ Compute $\rho(\mathbf{x}) = \iiint d^3v f(\mathbf{J}(\mathbf{x}, \mathbf{v}))$;
- ▶ Solve the Poisson equation to find new $\Phi(\mathbf{x})$;
- ▶ Repeat until convergence.

[Binney 2014; Piffl et al.2015; Cole & Binney 2016; this work

Advantages of using actions

1. Action/angle variables are canonical \implies

the total mass of the model is computed trivially

$$M = \int f(\mathbf{x}, \mathbf{v}) d^3x d^3v = \int f(\mathbf{J}) d^3J (2\pi)^3,$$

does not depend on Φ , does not change between iterations.

2. Multicomponent models:

trivial superposition of separate $f_k(\mathbf{J})$ without changing the functional form of each component;

addition of a new component \implies

adiabatic modification of existing density profiles

(e.g., dark matter halo response to the formation of a baryonic disc).

3. Faster and more robust convergence ($\sim 5 - 10$ iterations).

How to compute the potential in a general case

1. Direct integration:

$$\Phi(\mathbf{x}) = - \iiint d^3x' \rho(\mathbf{x}') \times \frac{G}{|\mathbf{x} - \mathbf{x}'|}.$$

3. Spherical-harmonic expansion:

$$\Phi(r, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \Phi_{lm}(r) Y_l^m(\theta, \phi),$$

$$\Phi_{lm}(r) = -\frac{4\pi G}{2l+1} \times \\ \times \left[r^{-1-l} \int_0^r dr' \rho_{lm}(r') r'^{l+2} + r^l \int_r^{\infty} dr' \rho_{lm}(r') r'^{1-l} \right],$$

$$\rho_{lm}(r) = \int_0^{\pi} d\theta \int_0^{2\pi} d\phi \rho(r, \theta, \phi) Y_l^{m*}(\theta, \phi).$$

How to compute the potential in a general case

2. Azimuthal-harmonic (Fourier) expansion:

$$\Phi(R, z, \phi) = \sum_{m=-\infty}^{\infty} \Phi_m(R, z) e^{im\phi},$$

$$\rho_m(R, z) = \frac{1}{2\pi} \int_0^{2\pi} d\phi \rho(R, z, \phi) e^{-im\phi},$$

$$\Phi_m(R, z) = - \iint dR' dz' \rho_m(R', z') \times \Xi_m(R, z, R', z'),$$

analytic expr. for Green's function:

$$\begin{aligned} \Xi_m(R, z, R', z') &\equiv \int_0^{\infty} dk 2\pi G J_m(kR) J_m(kR') \exp(-k|z - z'|) = \\ &= \frac{2\sqrt{\pi} \Gamma(m + \frac{1}{2}) {}_2F_1(\frac{3}{4} + \frac{m}{2}, \frac{1}{4} + \frac{m}{2}; m + 1; \xi^{-2})}{\sqrt{RR'} (2\xi)^{m+1/2} \Gamma(m + 1)} \end{aligned}$$

$$\text{where } \xi \equiv \frac{R^2 + R'^2 + (z - z')^2}{2RR'}.$$

How to compute the potential in a general case

1. Direct integration:

$$\Phi(\mathbf{x}) = - \iiint d^3x' \rho(\mathbf{x}') \times \frac{G}{|\mathbf{x} - \mathbf{x}'|}.$$

2. Azimuthal harmonic expansion:

$$\Phi(R, z, \phi) = \sum_{m=-\infty}^{\infty} \Phi_m(R, z) e^{im\phi}.$$

3. Spherical harmonic expansion:

$$\Phi(r, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \Phi_{lm}(r) Y_l^m(\theta, \phi).$$

interpolated functions

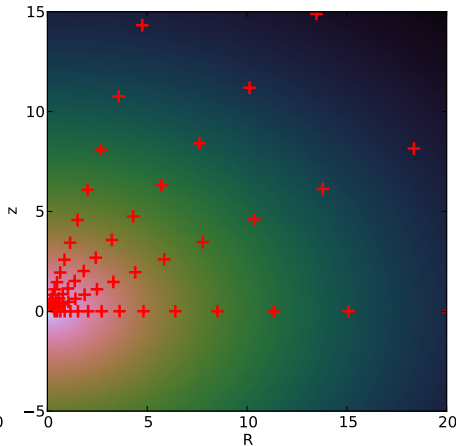
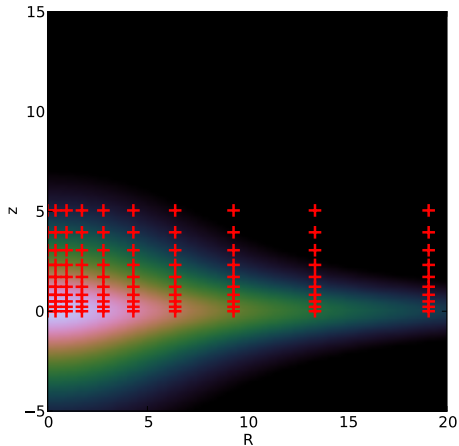
4. Basis-set expansion:

$$\Phi(r, \theta, \phi) = \sum_{n=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=-l}^l \Phi_{nlm} A_{nl}(r) Y_l^m(\theta, \phi).$$

(example: self-consistent field method of Hernquist&Ostriker 1992)

Two types of potential approximations used in models

- ▶ for disc-like components – azimuthal-harmonic expansion;
- ▶ for spheroidal components – spherical-harmonic expansion.



Gravitational potential extracted from N-body models

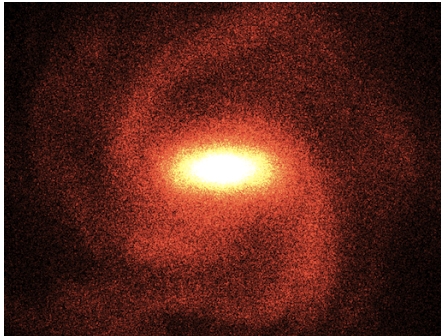
The spherical-harmonic and azimuthal-harmonic potential approximations can also be constructed from N -body models.

Advantages:

fast evaluation, smooth forces, suitable for orbit analysis.

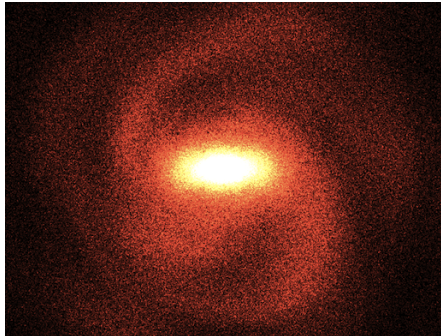
Real N -body model

(from Roca-Fabrega et al. 2013, 2014)



Potential approximation

(suitable for test-particle integrations,
e.g. Romero-Gomez et al. 2011)



Models with prescribed density profile

Goal: $\rho(\mathbf{x}) \implies f(\mathcal{I})$

construct a self-consistent model for the given density profile (possibly with additional kinematic constraints).

Need to somehow invert the integral equation

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

Variants of methods:

- ▶ Based on N -body models: made-to-measure and alike...
- ▶ Based on orbits: Schwarzschild's orbit-superposition method.
- ▶ Based on distribution-function "building blocks": this work.

Schwarzschild's orbit-superposition method

Discretize both the density profile and the distribution function:

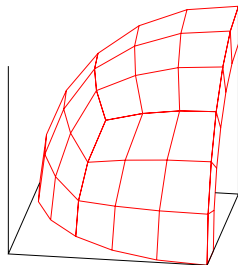
$\rho(\mathbf{x}) \implies$ cells of a spatial grid;
mass of each cell is

$$m_c = \iiint_{\mathbf{x} \in V_c} \rho(\mathbf{x}) d^3x$$

$f(\mathcal{I}) \implies$ collection of orbits:

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

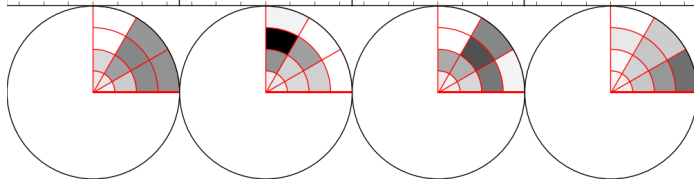
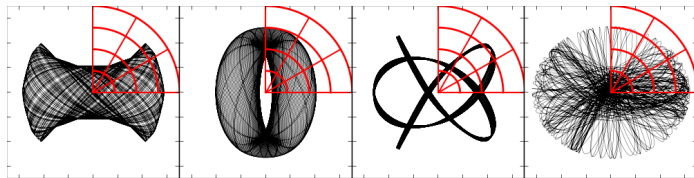
adjustable weight of each orbit



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

Schwarzschild's orbit-superposition method

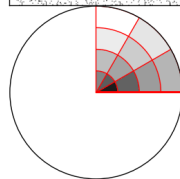
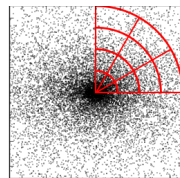
orbits in the model



discretized orbit density

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

target density



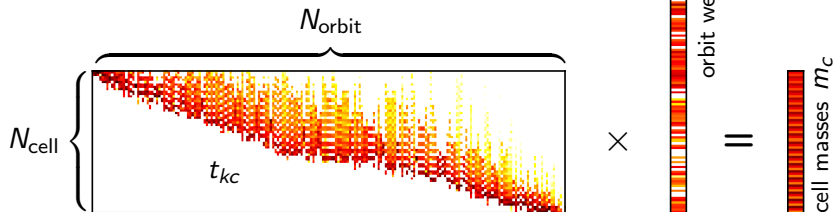
discretized density

(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

Solve the linear system with constraints $w_k \geq 0$
(linear or non-linear optimization problem)



Importance of regularization:



non-regularized



regularized

Building blocks for distribution function

Same idea: invert the integral equation

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

by decomposing f into a sum of building blocks (basis functions):

$$f(\mathcal{I}) = \sum_{k=1}^{N_{\text{basis}}} w_k f_k(\mathcal{I}),$$

computing the projections of all basis functions at a grid of points \mathbf{x}_c

$$\rho_{kc} \equiv \iiint d^3v f_k(\mathcal{I}(\mathbf{x}_c, \mathbf{v})),$$

and solving the optimization problem

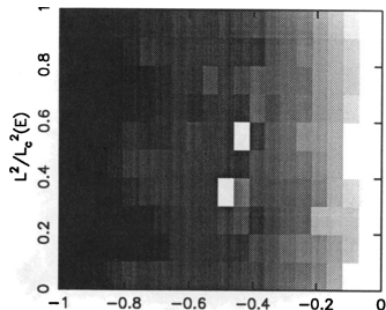
$$\rho_c \equiv \rho(\mathbf{x}_c) = \sum_{k=1}^{N_{\text{basis}}} w_k \rho_{kc}, \quad w_k \geq 0, \quad c = 1..N_{\text{constraint}},$$

to find the weights of basis functions w_k .

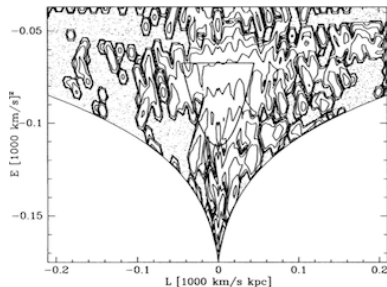
Building blocks for distribution function

Similar approaches suggested previously:

- ▶ 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&Thiébaud(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 $f(E, L)$.

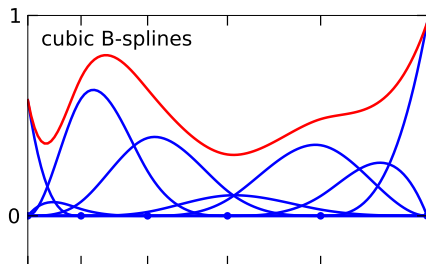
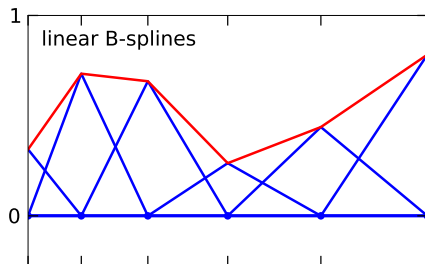
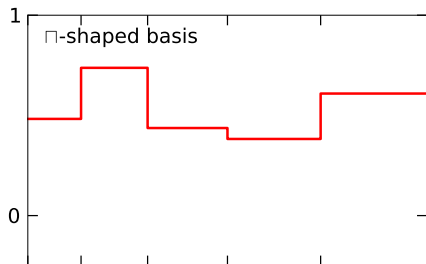
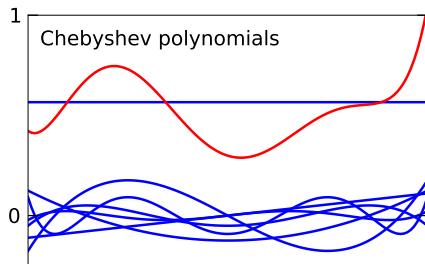


E [Merritt 1993]



[Kuijken 1994]

Building blocks for interpolation



B-splines of degree N :

flexible choice of grid points, locality, smoothness (increases with N), nonnegativity.

Models with non-parametric distribution function

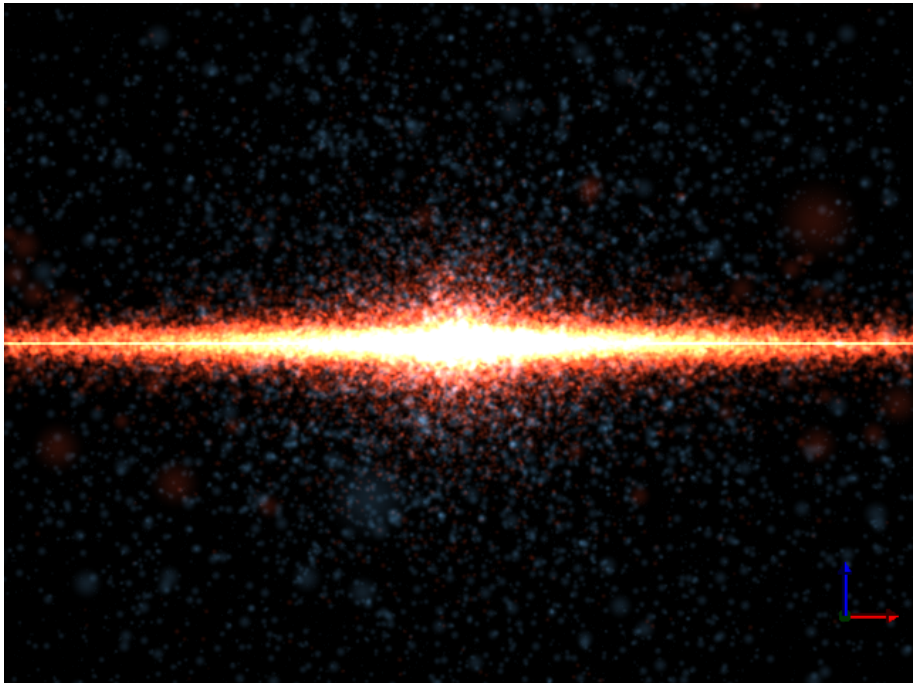
- ▶ $f(\mathbf{J})$ represented as an interpolated 3d function in action space (tensor product of 1d B-splines);
- ▶ weights of basis functions found by solving a linear or quadratic optimization problem (constraints: values of density at a 3d grid in space);
- ▶ smoothness: choice of degree of B-splines;
- ▶ regularization: minimum-curvature condition for the 3d interpolant (roughness penalty);
- ▶ possibility of determining $f(\mathbf{J})$ from an N -body model or from discrete observational points, using maximum penalized likelihood method (work in progress..).

Advantages of models based on distribution function

- ▶ Clear physical meaning
(localized structures in the space of integrals of motion);
- ▶ Easy to compare different models
(how to compare two N -body or N -orbit models?);
- ▶ Easy to compare models to discrete observational data;
- ▶ Easy to sample particles from the distribution function
(convert to an N -body model);
- ▶ Stability analysis
(perturbation theory most naturally formulated in terms of actions);

Caveats:

- ▶ Implicitly rely on the integrability of the potential, ignore the presence of resonant orbit families;
- ▶ So far implemented only for axisymmetric models
(not a fundamental limitation).



AGAMA library – Action-based galaxy modeling architecture

- ▶ Extensive collection of gravitational potential models (analytic profiles, azimuthal- and spherical-harmonic expansions);
- ▶ Conversion to/from action/angle variables (fast and accurate method for spherical potentials, Stäckel fudge for axisymmetric potentials, torus mapping);
- ▶ Action-based distribution functions; generation of N -body models and determination of best-fit parameters of DF and potential;
- ▶ Self-consistent multicomponent models with action-based DFs: (iterative method for $f(\mathbf{J}) \Rightarrow \rho(\mathbf{x})$, non-parametric DF recovery $\rho(\mathbf{x}) \Rightarrow f(\mathbf{J})$);
- ▶ Efficient and carefully designed C++ implementation, examples, Python and Fortran interfaces, plugins for galpy, NEMO, AMUSE;



<https://github.com/GalacticDynamics-Oxford/Agama>

Conclusions

- ▶ Advantage of models based on distribution functions;
- ▶ Advantage of actions as arguments of distribution functions;
- ▶ Two approaches for construction of self-consistent models;
- ▶ Software available for the community.

THANK YOU!