Modern stellar dynamics, lecture 8: dynamical modelling in practical applications

Eugene Vasiliev

Institute of Astronomy, Cambridge

Part III / MAst course, Winter 2022

Purpose of dynamical modelling

Determine the mass distribution of a stellar system from the kinematics of some tracer population(s) under the assumption of dynamical equilibrium.

Methods

- **0.** Virial theorem: 2K + W = 0kinetic energy \sim potential energy virial mass estimators: $GM \propto r \sigma^2$
 - 1. Jeans equations
 - 2. Distribution functions
 - 3. Orbit superposition
 - 4. Made-to-measure

Forward modelling approach:

- assume a functional form for the model;
- adopt some plausible model parameters; 5
- predict the observable kinematics;
- vary the model parameters to maximize the likelihood of observed values.

3. Schwarzschild's orbit-superposition method

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

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:

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



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

 $f(\mathcal{I}) = \sum_{k=1}^{N_{orb}} w_k \, \delta(\mathcal{I} - \mathcal{I}_k)$ each orbit is a delta-function in the space of integrals of motion
adjustable weight of each orbit [to be determined]

Schwarzschild's orbit-superposition method: self-consistency



For each *c*-th cell we require $\sum_{k} w_k t_{kc} = M_c$, where $w_k \ge 0$ 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_{cell} and to the kinematic observables u_{kn}, n = 1..N_{obs}

Solve the constrained optimization problem to find orbit weights w_k :

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

subject to $w_k \ge 0$, $k = 1..N_{orb}$, observational constraints
 $\sum_{k=1}^{N_{orb}} w_k t_{kc} = M_c$, $c = 1..N_{cell}$ density constraints (cell masses)

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

Schwarzschild's orbit-superposition method: fitting procedure

Solve the linear system with non-negativity constraints on the solution vector $w_k \ge 0$ (linear or non-linear optimization problem)



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 {x_k, v_k, w_k} |<sup>N_{body}_{k=1} moving in a potential Φ(x)
 </sup>
- ▶ Observational constraints U_n and their uncertainties δU_n , $n = 1..N_{obs}$

• Model predictions for these observations: $V_n = \sum_{k=1}^{N_{body}} w_k K_n(\mathbf{x}_k, \mathbf{v}_k)$

Objective:

• minimize $\Omega \equiv \frac{1}{2} \sum_{n=1}^{N_{obs}} \Delta_n^2 + S(\{w_k\})$, where $\Delta_n \equiv (V_n - U_n)/\delta U_n$ is the deviation in *n*-th constraint, $S(\{w_k\})$ is some measure of smoothness (regularization term), by varying the particle weights w_k .

some predefined kernels

Made-to-measure models

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

Procedure:

- Evolve the *N*-body system in time: $\dot{\mathbf{x}}_k = \mathbf{v}_k, \ \dot{\mathbf{v}}_k = -\frac{\partial \Phi}{\partial \mathbf{x}}\Big|_{\mathbf{x}=\mathbf{x}_k}$
- Adjust the particle weights: $\dot{w_k} = -\frac{w_k}{\tau_{ch}} \frac{\partial \Omega}{\partial w_k}$ (force-of-change)
- To reduce fluctuations, replace $\Delta_n(t)$ by a time-smoothed

 $\tilde{\Delta}_n(t) \equiv \frac{1}{\tau_{\rm sm}} \int_0^\infty \Delta_n(t-\tau) \, \exp\left(-\frac{\tau}{\tau_{\rm sm}}\right) \, d\tau \text{ in the above expression}$

* 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

Schwarzschild method

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

VS.

- 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

Applications of dynamical modelling

Milky Way:

- Nuclear star cluster
- Central region (bar/bulge)
- Disc in the Solar neighbourhood
- Outskirts (halo)

External galaxies:

- Satellites of the Milky Way
- Galactic nuclei and supermassive black holes
- Large samples of galaxies in the local Universe

unresolved

Nuclear star cluster

A central compact cluster of stars, with radius of a few pc and mass of a few $\times 10^7 M_{\odot}$, containing the central supermassive black hole (Sgr A^{*}). Only visible in the [near-]infrared.



Nuclear star cluster and the central supermassive black hole

Individual stellar orbits around the central SMBH Sgr A* are traced over $\gtrsim 20$ years, with some stars completing more than orbit. The 2020 Nobel prize was awarded for the discovery of a *supermassive compact object*.





Nuclear star cluster and the central supermassive black hole

R/R_(4.3e6 Msun BH)

4.4

4.3

R. = 8277 pc fixed

+0.5 % precisio

The mass of the black hole measured from stellar orbits is $4.3 \times 10^6 M_{\odot}$. The latest interferometric observations with the GRAVITY instrument at VLT achieve $\lesssim 1\%$ error on M_{\bullet} , firmly detect relativistic effects (pericentre precession and gravitational redshift), and begin to constrain the extended mass distribution.



Nuclear star cluster and the central supermassive black hole

Black hole mass growth



Nuclear star cluster models

Ingredients: star counts and/or diffuse light profile; individual V_{los} and proper motion measurements, or integrated-light kinematic maps

Methods: Jeans analysis, distribution functions, Schwarzschild models

Outputs: M_{\bullet} , stellar M/L, enclosed mass profile, velocity anisotropy, ...

A testbed for M_{\bullet} measurement in external galaxies, for which we do not have a "true answer" from time-resolved stellar orbits!





It has been argued that the non-negativity of the distribution function imposes the constraint $\gamma \ge \beta_0 + 1/2$ (An & Evans 2006). This relation is violated in large parts of the $\beta_0 - \gamma$ preferred region (Figure 1(e)) and this issue deserves a separate investigation. Including this limit will likely result in slightly steeper γ and increased tangential anisotropy. A distribution function analysis similar to that of Wu & Tremaine (2006) will be useful to confirm the present results.

[Do+ 2013]

Galactic bar (formerly known as bulge)

Ingredients: star counts, v_{los} and PM from ground-based infrared surveys (2MASS, VVV, BRAVA, APOGEE) in selected low-extinction fields; microlensing depth Methods: Schwarzschild, N-body and M2M models (triaxial, rotating) Outputs: mass distribution, orbit structure, pattern speed



Galactic disc

Ingredients: *millions* of stars with 6d phase-space measurements, but mostly within 1–2 kpc from the Sun; ages and chemical properties: [Fe/H], [α /Fe], ... Methods: axisymmetric Jeans analysis, distribution functions Outputs: circular-velocity curve, vertical structure, local DM density Challenges: extinction, selection effects, perturbations & disequilibrium



Gaia 6d sample [credit: ESA]



Terminological challenge: definitions of thin & thick discs.

Thin: scale height ~ 0.3 kpc, metal-rich, α -weak, young and intermediate age, velocity dispersion increases with age. Thick: scale height ~ 1 kpc, more metal-poor, α -enhanced, old.



Juric et al. 2008

0.8

0.6

0.4

0.2

Juric et al 2008.

model notential

 $P_*(R_0, z)$

Gilmore & Reid 1983

 $p_{dm} = 0.012 \text{ M}_{\odot} \text{ pc}^{-3}$

One-dimensional vertical hydrostatic equilibrium and mass distribution:

Jeans eqn in z:

$$\frac{1}{R} \frac{\partial (R \rho_* \overline{v_R v_z})}{\partial R} + \frac{\partial (\rho_* \overline{v_z}^2)}{\partial z} + \rho_* \frac{\partial \Phi}{\partial z} = 0,$$
tracer density
Poisson eqn:

$$\frac{1}{R} \frac{\partial}{\partial R} \left(R \frac{\partial \Phi}{\partial R} \right) + \frac{\partial^2 \Phi}{\partial z^2} = 4\pi G \rho.$$
total density
total density
rotation curve term $= \frac{1}{R} \frac{\partial v_{circ}^2}{\partial R} \approx 0$
Neglecting tilt and rotation curve terms, we have

$$\rho = -\frac{1}{4\pi G} \frac{\partial}{\partial z} \left(-\frac{\partial \Phi}{\partial z} \right) = -\frac{1}{4\pi G} \frac{\partial}{\partial z} \left(\frac{1}{\rho_*} \frac{\partial (\rho_* \sigma_z^2)}{\partial z} \right),$$
vertical force K_z
or, integrating the Poisson equation over z up to some height h,

$$\Sigma(h) \equiv \int_{-h}^{h} \mathrm{d}z \ \rho(z) = \frac{|K_z(h)|}{2\pi \ G} = -\left[\frac{1}{2\pi \ G \ \rho_\star} \frac{\partial(\rho_\star \sigma_z^2)}{\partial z}\right]_{z=h} \text{ (using tracers at } z=h\text{).}$$

1d Jeans equations: use only the tracer density and velocity dispersion profiles



Alternatively, one may invoke the Jeans theorem and define the 1d DF of tracers: $f(z, v_z) = f(E_z), \quad E_z \equiv \Phi(z) + \frac{1}{2}v_z^2.$

Then the tracer density is

$$\rho_{\star}(z) = \int_{-\infty}^{\infty} \mathrm{d}v_z \ f(z, v_z) = 2 \int_{\Phi(z)}^{\infty} \mathrm{d}E_z \ \frac{f(E_z)}{\sqrt{2[E_z - \Phi(z)]}}.$$

On the other hand, we may construct $f(E_z)$ from the full velocity distribution $f(0, u_z)$ at z=0, i.e., using only the local stellar kinematics.

A star with velocity u_z at z = 0 will have a velocity $v_z \equiv \sqrt{u_z^2 - 2\Phi(h)}$ at z = h, and the integration over velocity can be written as

$$\rho_{\star}(h) = 2 \int_{0}^{\infty} \mathrm{d}v_{z} \ f(z, v_{z}) = 2 \int_{\sqrt{2\Phi(h)}}^{\infty} \mathrm{d}u_{z} \ f(0, u_{z}) \ \frac{u_{z}}{\sqrt{u_{z}^{2} - 2\Phi(h)}}$$

Although this equation cannot be easily inverted to get $\Phi(h)$ explicitly, one can try various parametric forms of $\Phi(h)$ and choose the one that best reproduces the link between the density profile $\rho_*(h)$ and the local velocity distribution $f(0, u_z)$ [Kuijken & Gilmore 1989, 1991; Holmberg & Flynn 2000].

1d DF: use only the tracer density profile and the full velocity distribution at z = 0



Different components (stars, cold and hot gas, dark matter) have different vertical distributions, so to break the degeneracy between them, one needs to measure the vertical force $K_z(h)$ or equivalently the surface density $\Sigma(h)$ at several values of h, preferrably above 1 kpc. OTOH for larger h the 1d approximation becomes increasingly less accurate. The local density of *dark* matter (what remains after subtracting all other components) is ~ 10% of the total in-plane density – challenging to determine accurately!





local DM density estimates [de Salas & Widmark 2021]

Galactic disc: global structure and the circular-velocity curve

ilers et al. 2018 (this work)

Lopez-Corredoira et al. 2014

luang et al. 2016

Kafle et al. 2012

350

300

1-S 1200

n.: linear fit

all stellar components -

halo: NFW-profile fit

all stellar components

thin disk

thick dis

More sophisticated Jeans models use the $\overline{v_{\phi}}$ and $\sigma_{R,\phi,z}$ profiles in the 2d meridional plane, while DF-based models compare the predicted velocity distributions $f(v_R), f(v_{\phi}), f(v_z)$ with observed histograms in a large range of R, z.



Galactic outskirts and the global Milky Way mass profile

Ingredients: stars with well-measured distances (RR Lyrae, BHB, red giants, ...), globular clusters and satellite galaxies as test particles; v_{los} and more recently PM from Gaia and HST

Methods: spherical Jeans analysis, distribution functions (usually simple power-law "tracer mass estimators"), stellar streams, cosmological *N*-body sims, . . .

Outputs: mass profile, shape of DM halo, substructures in stellar halo, ... Challenges: substructure, multiple components, disequilibrium



Galactic outskirts and the global Milky Way mass profile

Stellar halo: total mass $\sim 10^9 M_{\odot}$; double-power-law density profile with a break radius at $\sim 20 - 30$ kpc; highly radial velocity anisotropy in metal-rich component (remnant of an early major accretion event).



Galactic outskirts and the global Milky Way mass profile

Globular clusters: \sim 150 objects with full 6d phase-space information, reaching out to \gtrsim 100 kpc (though most of them are in the inner part of the Galaxy). Satellite galaxies: \sim 30 objects at considerably larger radii (50 – 200 kpc), also with position/velocity measurements, but not necessarily in equilibrium...



Milky Way satellites

name	distance	luminosity	
LMC	50	1e9	
SMC	64	4e8	
Sgr	27	5e7	
Fornax	147	2e7	
Leo I	250	4e6	
Sculptor	86	2e6	
Leo II	230	6e5	
Sextans	105	5e5	
Carina	86	4e5	
Draco	76	3e5	
Ursa Mi	nor 76	3e5	
Antlia 2	130	2e5	
ultrafaint galaxies			



Sculptor dSph

Fornax dSph

Milky Way satellites

Number of stars with measured velocities or PM: $\gtrsim 10^7$ for LMC, $\sim 2\times 10^5$ for Sgr, few $\times 10^3$ for classical dSph, $\mathcal{O}(10)$ for ultrafaints.

PM from Gaia DR3 are helpful for cleaning up the membership selection, but are usable for modelling (have uncertainties smaller than velocity dispersion) only for LMC, SMC and Sgr.

Line-of-sight velocity dispersions ranging from ~ 25 km/s for LMC down to 1-3 km/s for ultrafaint galaxies (comparable to measurement uncertainties and affected by binary stars at the level of a few km/s).

Potential likely dominated by dark matter even in the central regions.



Mass modelling of dwarf galaxies

Ingredients: density profiles from star counts; v_{los} of 100s–1000s stars Methods: Jeans analysis, distribution functions, Schwarzschild models Outputs: DM density profile, stellar velocity anisotropy Challenges: breaking the mass–anisotropy degeneracy, e.g., by using higher-order velocity moments [Merrifield&Kent 1990; Łokas&Mamon 2003; Richardson & Fairbairn 2013; Read & Steger 2017] or PM data [Wilkinson+ 2002; Strigari+ 2007; Massari+ 2019]



Multiple populations in dwarf galaxies

Two-population modelling: chemically distinct stellar populations with different kinematics and spatial distribution, living in the same potential \implies probe the dark matter profile at different sweet-spot radii.

> (108 7.5 M)

 $\log_{10} [M(<R_{half})]$

65





Cusp-to-core transformation in dwarf galaxies

Standard cosmological model (ACDM) predicts cuspy density profiles of dark matter haloes (e.g., Navarro– Frenk–White, $\rho \propto r^{-1}$ at small r), but dynamical models of *some* dwarf galaxies prefer cored profiles. One possible way of transforming cusps into cores is via bary-onic feedback – rapid, non-adiabatic expansion caused by gas outflows caused by starbursts, followed by slow, adiabatic accumulation of gas, repeated many times.





Cusp-to-core transformation in dwarf galaxies

In this scenario, we may expect that galaxies with low stellar mass and short duration of star formation history (such as Draco) should retain cusps, while galaxies with more extended star formation history (such as Fornax) should acquire cores, and this seems to be supported by observations.

Draco

Fornax

Radius (kpc)

夏 10-

[Read+ 2019]

SFR (M_☉/yr)

10

SFR (M_☉/yr)

10

Age (Gyr)



Dynamical modelling in galaxies outside the Local Group

Ingredients: surface brightness profiles, integrated-field spectroscopy (2d maps of V_{los} distribution, and sometimes also stellar age and metallicity maps) Methods: axisymmetric Jeans models, Schwarzschild and M2M models

Outputs: supermassive black hole mass, stellar M/L, dark halo properties, ...

Challenges: deprojection degeneracies (usually ignored!), finite spatial resolution, limited spatial coverage, excessive flexibility in models (!)



Dynamical modelling in galaxies outside the Local Group

Usually the total potential is generated by the [deprojected] stellar light distribution multiplied by the mass-to-light ratio M/L, the central black hole M_{\bullet} , the dark halo, etc. In Schwarzschild models, a new orbit library is built for each choice of the potential, and orbit weights are assigned to reproduce the kinematic maps as closely as possible while retaining dynamical self-consistency. The difference χ^2 between the model and the data is plotted as a function of model parameters, and the quantities of interest (such as M_{\bullet}) and their uncertainties are determined from $\Delta \chi^2 \equiv \chi^2 - \chi^2_{min} \leq$ some threshold.





Dynamical modelling in galaxies outside the Local Group

In Jeans models, the procedure is quite similar, but instead of the entire orbit library, the kinematics properties of the model are predicted for the given choice of potential parameters and the single anisotropy parameter *b* or β_z (which determines the shape of the velocity ellipsoid) and compared with the observed maps of the full second moment of velocity $\overline{v^2} \equiv \overline{v}^2 + \sigma^2$ (not the entire set of GH moments!).

Jeans models are much faster than Schwarzschild models and use less data, but are still believed to produce reliable constraints on the potential. $\underbrace{_{MGC4578 V_{MGC}}}_{V_{MGC} JMM \ model}$



Scaling relations for large galaxy samples

Dynamical determination of black hole masses in many galaxies has led to a discovery of relatively tight relations between M_{\bullet} and other galaxy properties (mass or luminosity, velocity dispersion, etc.) [Ferrarese & Merritt 2000; Gebhardt+ 2000], which have been further refined with time (notably, not all galaxy types follow the same relations).

OTOH the dynamical analysis of galaxies at larger scales ($\gtrsim r_{half-light}$) unveiled a dichotomy between fast and slow rotators and its dependence on the morphology and the

total mass [Emsellem+ 2007; Cappellari+ 2007]. $M_{\bullet} - \sigma$ relation 1010 morphology (ellipticity) vs. kinematics relation ATLAS^{3D} + SAMI Pilot 0.8 Isotropic Rotator 0.7 109 $\delta = 0.7 \times \varepsilon_{intr}$ 0.6 $\lambda_R = 0.08 + \varepsilon/4$ M_{\bullet}/M_{c} No Rotation 0.5 **Complex Velocity** 1.08 (⁹) 0.4 KDC **Counter Rotating Regular Rotator** 0.3 107 0.2 0.1 [Kormendy & Ho 2013] [Cappellari 2016] 0.3 0.4 0.5 0.6 0.7 0.8 0.9 60 80 100 200 300 400 $\sigma_{\rm c}$ (km s⁻¹)

Uncertainties and degeneracies in determining the potential

$$\mathcal{F}(X, Y, V_{\text{los}}) \implies \begin{cases} f(E, I_2, I_3) \\ \Phi(x, y, z) \end{cases}$$

Expect a large range of possible potentials consistent with observed kinematics!

Constraints get tighter when increasing spatial coverage (by excluding unrealistic orbit distributions) *or* having imperfect/noisy data

(see discussions in Dejonghe&Merritt 1992; Valluri+ 2004; Magorrian 2006, 2013).



Mock data: spherical Hernquist model with a black hole, two variants of spatial coverage

Summary

- Dynamical modelling is the art of determining the intrinsic structure and mass distribution of a stellar system from the limited observable kinematic information.
- Different methods perform better in particular situations (e.g., when having discrete or integrated-light kinematics, V_{los} only or all three velocity components, etc.), and detailed comparisons of several methods on identical input data are scarce and inconclusive.
- The dynamical modelling problem is *intrinsically degenerate* in most cases, even when using the equilibrium assumption!
- Nevertheless, it remains the primary source of knowledge about unseen mass in the Universe (dark matter, massive black holes).