Measuring the global properties of disk galaxies

May 8, 2022

This problem is about characterizing the properties of a galaxy produced by a cosmological simulations. The physical processes responsible for the formation and evolution of galaxies influence their shapes, their masses and their sizes among other properties. In this exercise, we use a characteristic scale, the exponential scale-length, as a scale for the size of a disk galaxy.

1 Loading Cosmological Simulation Data

1. Register an account on https://www.tng-project.org/

2. Get familiar with the data specifications (some of these are described in this page).

3. Download and load particles from Subhalo 117258 **in the TNG50** simulation run. Note there is additional information about the subhalos that might be used in the subsequent questions, from https://www.tng-project.org/data/search/?sim=TNG100-1snap=99. `h5py` python package may be useful. A single halo should be ∼100MB worth data.

4. The coordinates of the particles are given in units as specified on the project webpage. Much of the units are expressed in comoving kiloparsecs per h. At redshift 0 (i.e. present-day), comoving units and physical units are the same, so that only the reduced Hubble factor h is relevant in this case. Convert the units to physical kiloparsecs. Some relevant background literature on distance measurements in cosmology can be found at the following link.

5. Visualize the distribution of the stellar particles with a 2D color plot showing the log of the surface density (in units of solar masses per squared kpc). One can address this question with the three following steps:

a) Make a 2d histogram of the count of stars in bins of x and y
b) Weight that histogram by the masses of each particles (in solar mass!)
c) Divide the values obtained at **b)** by the surface element dxdy (set by the x and y bins sizes) ... with logarithmic scaling.

2 Defining reference frame

The particles coordinates and velocities are given with respect to the simulation box. To address galactic dynamics questions, it is often more handy to work in a coordinate centered on the center of the galaxy and that reflects well its symmetry properties.

6. It is relevant to define a new coordinate system, centered on the center of the galaxy and where the disk lies in the x-y plane. One way to do this, for disk galaxies, is to compute the angular momentum of the disk (the stars rotate about the galactic center, in the plane of the disk), and align the z coordinates with it.

a) Express the coordinates and the velocities with respect to the center of mass of the star particles, and with respect to the velocity of the center of mass.

b) Write a function that calculates the angular momentum vector of the inner disk (say the inner 3-4 kpc) with respect to the center of the subhalo, where

$$\vec{L} = \sum_i m_i \vec{r}_i \times \vec{v}_i \tag{1}$$

c) Define a new *orthonormal* coordinate system where the vertical axis aligns with the angular momentum calculated at **2)**.

d) Write a function that transform the old coordinates of the stellar particles (w.r.t the simulation box) to the new coordinates (w.r.t the newly defined basis). To do this, write the transformation matrix from the old base to the new base (at least in python, vectorized operations are much faster).

7. In the new coordinate system,

a) visualize the distribution of the stellar particles with a 2D histogram showing the log of the surface density (in units of solar masses per squared kpc)

b) plot the the log surface density profile (= log Sigma as a function of radius)

c) Measure the slope by eye (dlog Sigma / dR). Assuming that the profiles of disk galaxies are well desribed by an exponential Sigma propto exp(-R/Rd), what would be the scale-length Rd of that galaxy?

3 Measuring Physical Properties of Galaxies (Optional)

Now we would like to have a quantitative estimate for the exponential scale-length of the disk of the galaxy. We can write a model for the mass surface density

$$\Sigma(R, R_d, \Sigma_0) = \Sigma_0 \exp(-R/R_d) \tag{2}$$

Assuming that stellar particles are drawn from such a distribution, with masses and positions $\{m_i, x_i, y_i\}$, we can write their probability density as

$$p(x, y) = \exp(-R/R_d)/R_d^2, \tag{3}$$

which would be the spatial distribution of the stellar particles had they all equal masses. Since stellar particles have different masses, we will construct a weighted likelihood function accounting for that more massive stellar particles are more representative of the surface density distribution.

If all data are drawn independently, their joint probability given the scale-length R_d is (where we leave out irrelevant terms)[1]

$$\mathcal{L}(R_d; x, y, m) = p(\{x_i, y_i\}|R_d, \{m_i\}) \propto \prod p(x_i, y_i)^{m_i}. \tag{4}$$

This probability should be maximized for the (true) parameter R_d^{true} from which the stellar particles were sampled. The goal is to estimate it.

8. Write functions that compute the output of eq. 3 and the log of eq. 4.

9. Plot the log likelihood evaluated on the star particles, as a function of R_d for R_d values going from 0.5 kpc to 10 kpc. Where is the minimum?

10. Optimize the log likelihood by finding numerically the R_d value that maximizes the log likelihood. The function `scipy.optimize.minimize` used on the *minus* log likelihood may be useful.

11. How different is that result from the original measurement in question **7.c)**? Comment on what could be responsible from these differences (use the figures produced during the problem).

[1] More accurately, this can be modelled with a Poisson point process that can also account for the normalizing density Σ_0. This exercise simplifies to the most relevant terms only, and can be simplified further setting $m_i = 1$ (i.e. assuming all star particles have the same mass).

4 Tips

Help for 6.c)
Let's call the old coordinate system (the one of the simulation box) base vectors $\{\vec{e}_x, \vec{e}_y, \vec{e}_z\}$ and the new ones the same with '.
i. Take $\vec{e}'_z = \vec{L}/|\vec{L}|$.
ii. The cross product operation produces naturally orthogonal vectors. One can take a new 'un-normalized' vector $\vec{u}'_x = \vec{e_x} \times \vec{e}'_z$, and normalize it with $\vec{e}'_x = \vec{u}'_x/|\vec{u}'_x|$.
iii. Now there is left to define an $\vec{e}'_y$ that is orthogonal to both $\vec{e}'_x$ and $\vec{e}'_z$, and to normalize it.

Help for 6.d)
Construct the matrix $\mathbf{T}$ of the change of basis from the simulation box to the new basis. By definition $\vec{x}_{\text{old}} = T\vec{x}_{\text{new}}$.Since the two bases are orthogonal, obtaining the new coordinates is just multiplying with the transpose $\vec{x}_{\text{new}} = \mathbf{T}^T\vec{x}_{\text{old}}$. The elements of T are obtained by taking the dot products of $\vec{e_i}$ and $\vec{e}'_j$ (where i, j are x, y, z)

Help for 7.b)
The surface mass density profile as a function of radius is not a histogram of radius. The former has units of $M_\odot kpc^{-2}$ whereas the latter has units of kpc^{-1}. For a tip: the difference between a surface number density profile $p(x, y) = f(R)$ (a function of radius only if the galaxy is axisymetric) and that of a radial density profile $p(R)$ is related by a factor R, as $p(x, y) \propto \frac{1}{R}p(R)$. In other words: the surface element $dS = Rd\phi dR$.