ASTR 333/433 — Dark Matter — Midterm Review — Spring 2024

OBSERVATIONAL EVIDENCE FOR MASS DISCREPANCIES

There are many lines of evidence for mass discrepancies in the universe. Historically, some of the more important include the

- Oort discrepancy in the solar neighborhood (Oort)
- Stability of galactic disks (e.g., Peebles & Ostriker)
- Flat rotation curves of spiral galaxies (Rubin, Bosma)
- Large velocity dispersions of stars in dwarf spheroidals (Aaronson; many others)
- Clusters of Galaxies
  - Large velocity dispersions of cluster member galaxies (Zwicky)
  - Gravitational lensing
  - Hydrostatic equilibrium of hot intracluster gas (as seen in X-rays and the SZ effect)
- Large Scale Structure (needs a boost to get it to grow: Peebles)
- Cosmic gravitating mass density exceeds baryon density from Big Band Nucleosynthesis: $\Omega_m > \Omega_b$

Together, the last two imply the need for some form of non-baryonic, dynamically cold dark matter that does not interact via electromagnetism (so zero interaction cross-section with photons). This implies the existence of a new kind of particle outside those known in the Standard Model of Particle Physics.

DARK MATTER CANDIDATES

- Baryonic
  - Conventional
    * Brown dwarfs, Jupiters, very faint stars
    * Very cold molecular gas (in thermal equilibrium with cosmic background radiation)
    * Warm-hot gas (diffuse gas that has a temperature near to 10^5 K is very hard to detect)
  - Exotic
    * White dwarfs, neutron stars, black holes (stellar remnants)
    * Strange nuggets (including 'macros'), primordial black holes
- Non-baryonic
  - Hot Dark Matter
    * Massive neutrinos
  - Cold Dark Matter
    * WIMPs (Weakly Interacting Mass Particles) and extensions
    * Primordial black holes, strange nuggets (act as non-baryonic CDM if made before BBN)
  - Other
    * Warm dark matter (WDM), Self-interacting dark matter SIDM, Superfluid dark matter, Axions, etc.

The existence of dark matter is an inference based on the assumption that gravity is normal; the observed discrepancies could also indicate a change in the force law.

VIRIAL THEOREM

Can be derived be requiring that the moment of inertia tensor be stationary for an object in equilibrium. Boils down to

$$2\langle \mathbf{K} \rangle + \langle \mathbf{W} \rangle = 0$$

where $\mathbf{K}$ is the kinetic energy and $\mathbf{W}$ is the gravitational potential energy. For N particles of mass m such that $M = Nm$,

$$M = 2\frac{\sigma^2}{G} R_{\rm hms}$$

where σ is the velocity dispersion representing the kinetic energy and $R_{\rm hms}$ is the harmonic radius

$$\frac{1}{R_{\rm hms}} = \sum \frac{1}{r_{i,j}}$$

which appears because the gravitational potential depends on the separation between each pair of masses i & j. It is commonly approximated as

$$R_{\rm hms} \approx 1.25 R_e$$

where R_e is the radius containing half the total light. This is widely applicable in many systems (King), but not perfect in all. It also assumes that the distribution of mass traces the distribution of light. This is not obviously a good assumption if most of the mass is dark: the half-mass radius could be rather different from the half-light radius. For completeness, it is good to know that for a spherical system, the half mass radius in 3D is related to the projected half mass radius as

$$R_{1/2} = \frac{4}{3} R_e.$$

For anisotropic orbits in spherical systems, the mass estimator becomes

$$M(r) = \frac{r\sigma_r^2}{G}\left(-\frac{\partial \ln n_*}{\partial \ln r} - \frac{\partial \ln \sigma_r^2}{\partial \ln r} - 2\beta\right)$$

where σ_r is the radial velocity dispersion, n_* is the distribution of tracers, and β is the anisotropy parameter:

$$\beta = 1 - \frac{\sigma_t^2}{\sigma_r^2}$$

and σ_t is the tangential velocity dispersion. Note that for
circular orbits, $\sigma_r = 0$; $\beta \to -\infty$;
isotropic orbits, $\sigma_r = \sigma_t$; $\beta = 0$;
radial orbits, $\sigma_t = 0$; $\beta = 1$.
Note also that the anisotropy parameter quantifies how the kinetic energy is partitioned between radial and tangential motion.
Remember: All this boils down to Newton's equation for circular velocity around a *dynamical* mass $M(r)$

$$V^2(r) = \frac{GM(r)}{r}$$

which, more generally, follows from the Poisson equation

$$\nabla^2 \Phi = 4\pi G \rho(r, \theta, \phi).$$

Solving the latter leads to constants of order unity in $V^2 = GM/r$ for non-spherical mass distributions, but $V^2 = GM/r$ always gets you in the right ballpark. We measure V and r, and find from this that extragalactic systems have dynamical $M(r) > M_b(r)$, the observed baryonic mass — hence the need for dark matter.

DISK STABILITY

Ostriker & Peebles (1973) showed that a cold, thin, rotating disk of point masses (stars) was unstable to the unchecked growth of the bar instability. A bare disk operating under Newtonian gravity self-shreds after only a few dynamical times. This contradicts the observation that dynamically cold spiral galaxies exist and are ubiquitous throughout the universe. Observations to high redshift show that this situation is sustained over at least half a Hubble time. Their solution was to impose an external potential; i.e., that of an invisible dark matter halo. Numerically, stability is achieved for

$$t_k = \frac{T}{|W|} \lesssim 0.14$$

where T is the kinetic energy in rotation. The total kinetic energy $K = T + \Pi/2$ where Π is the kinetic energy in random motions. For spiral galaxies, $T \gg \Pi$, so $K \approx T$ and according to the virial relation we should have $T/W \approx 1/2$. To lower this to the requisite 0.14, we increase the gravitational potential energy W by adding an invisible component in the form of a static dark matter halo.

Athanassoula et al. (1987) showed that the reverse was also true. Some disk instability is needed to drive the observed bars and spiral structure. If the potential energy is increased with out limit, disks become too stable and never form the observed structures. This leads to a lower limit on disk mass.

The modern situation is much more complicated, and does not boil down to a simple equation for the fractional kinetic energy t_k in rotation. Nevertheless, there is a clear need to stabilize disks, but also a tension between that and the need to not over-stabilize disks to the point that bars and spiral arms never form.

OORT DISCREPANCY

The restoring force K_z for vertical motions to a plane-parallel disk is

$$K_z = -\frac{\partial \Phi}{\partial z} = \frac{1}{n(z)} \frac{\partial (n\sigma_z^2)}{\partial z}$$

where $n(z)$ is the vertical profile of the tracer population, often taken to be exponential: $n(z) = n_0 e^{-z/z_0}$. Locally, this boils down to a simple relation between the vertical velocity dispersion σ_z, the surface density Σ, and the scale height z_0:

$$\sigma_z^2 = 2\pi G \Sigma z_0.$$

More generally,

$$K_z(R) = 2\pi G\Sigma(R) + 2z(A^2 - B^2) = 2\pi G\Sigma(R) - 2z\left(\frac{V}{R}\right)\left(\frac{dV}{dR}\right)$$

where A and B are the geenralized Oort parameters, and are themselved a fnction of radius. Note that in addition to the term directly attributable to the surface density $\Sigma(R)$, there is also a term for the shape of the gravitational potential in the radial direction that manifests as the $A^2 - B^2$ term. This latter term will cause a discrepancy between the dynamical surface density and the baryonic surface density even if there is no extra dark matter in the disk just because there is a quasi-spherical halo that causes the gradient in the rotation curve to deviate from the Keplerian case. This causes a lot of confusion in the literature, with seemingly contradictory claims that often boil down to differing assumptions about this term. (Of course, for a flat rotation curve, $dV/dR = 0$).

Definitions of Galactic Quantities

Distance to Galactic Center: R_0. Circular speed of LSR: Θ_0 (distinct from the solar motion relative to the LSR).

Orbital frequency: $\Omega = \frac{V}{R} = A - B$; Orbital Period $P = \frac{2\pi}{\Omega} = \frac{2\pi R}{V}$

Oort Constant $A = \frac{1}{2}\left(\frac{V}{R} - \frac{dV}{dR}\right)_{R_0}$; Oort Constant $B = -\frac{1}{2}\left(\frac{V}{R} + \frac{dV}{dR}\right)_{R_0}$

Epicyclic frequency: $\kappa^2 = -4B\Omega = -4B(A - B)$.

EXPONENTIAL DISKS

The surface brightness profile of spiral galaxies is frequently approximated with the *exponential disk*:

$$\Sigma(R) = \Sigma_0 e^{-R/R_d}$$

where Σ_0 is the central surface brightness (in $L_\odot\,\mathrm{pc}^{-2}$) and R_d is the scale length (in pc or kpc). Integrating this profile gives a cumulative enclosed luminosity

$$L(< x) = L_{tot}[1 - (1 + x)e^{-x}]$$

where $x = R/R_d$ and $L_{tot} = 2\pi\Sigma_0 R_d^2$. For a mass-to-light ratio Υ, the total stellar mass is $M_* = \Upsilon L_{tot}$. The rotation curve of a "spherical disk" is then

$$V^2(R) = \frac{GM_*}{R}[1 - (1 + x)e^{-x}]$$

while the more accurate thin disk rotation curve is

$$V^2(R) = \frac{2GM_*}{R_d} y^2[I_0(y)K_0(y) - I_1(y)K_1(y)]$$

where $y = R/(2r_d)$ and I and K are modified Bessel functions of the first and second kind. These (or a more accurate numerical representation of the observed light distribution) can be combined with dark matter halo models to fit the total rotation: $V_{tot}^2(R) = V_{disk}^2(R) + V_{DM}^2(R)$.

DARK MATTER HALO MODELS

Psuedo-Isothermal

Empirically motivated halo form with density

$$\rho_{iso}(r) = \frac{\rho_0}{1 + (r/R_c)^2}$$

where ρ_0 is the core density and R_c the core radius. These combine to define the asymptotic flat rotation velocity $V_\infty = \sqrt{4\pi G\rho_0 R_c^2}$ that appears in the rotation curve

$$V_{iso}(r) = V_\infty \sqrt{1 - \frac{R_c}{r}\tan^{-1}\left(\frac{r}{R_c}\right)}.$$

The limiting behaviors are $\rho \to$ constant and $V \propto r$ as $r \to 0$ and $\rho \propto r^{-2}$ and $V \to$ constant as $r \to \infty$.

NFW

NFW halos are found to results in dark matter-only simulations with density profile

$$\rho_{NFW}(r) = \frac{\rho_s r_s^3}{r(r + r_s)^2}$$

where ρ_s is the density at the break radius r_s where the density profile rolls over from inner to outer behavior. The rotation curve is

$$V_{NFW}(r) = V_{200}\sqrt{\frac{\ln(1 + cx) - cx/(1 + cx)}{x[\ln(1 + c) - c/(1 + c)}}$$

where $c = R_{200}/r_s$ and

$$V_{200}\sqrt{\frac{GM_{200}}{R_{200}}}$$

where R_{200} is the radius enclosing a density 200 times the critical density of the universe.

The limiting behaviors are $\rho \propto r^{-1}$ and $V \propto r^{1/2}$ as $r \to 0$ and $\rho \propto r^{-3}$ as $r \to \infty$.

LAWS OF GALACTIC ROTATION

The dynamics of disk galaxies are well-organized, and can be summarized with a few simple rules:

1. **Flat Rotation Curves** (the Rubin-Bosma Law)

   The rotation curves of galaxies tend towards an approximately constant rotation speed that persists to indefinitely large radii.

2. **Renzo's Rule** (Sancisi's Law)

   For any feature in the luminosity profile there is a corresponding feature in the rotation curve, and vice versa.

3. **The Baryonic Tully-Fisher Relation (BTFR)**

   The flat rotation speed of a galaxy correlates with its baryonic mass (the sum of stars and gas).

   $M_b = M_* + M_g = A\,V_f^4$ with $A = 47\ \mathrm{M}_\odot\,\mathrm{km}^{-4}\,\mathrm{s}^4$.

4. **The Central Density Relation (CDR)**

   The dynamically measured central mass surface density of a galaxy correlates with its photometrically measured central surface brightness.

   $\Sigma_{\mathrm{dyn},0} = \mathcal{S}(\Sigma_{\mathrm{bar},0}/\Sigma_\dagger)$

5. **The Radial Acceleration Relation (RAR)** (Milgrom's Law)

   The observed centripetal acceleration correlates with that predicted by the distribution of baryonic mass.

   $\mathrm{g}_{\mathrm{obs}} = \mathcal{F}(\mathrm{g}_{\mathrm{bar}}/\mathrm{g}_\dagger)$

Note that flat rotation curves (1) follow from the low acceleration limit of the RAR (5). Indeed, each of the first four laws can be derived from or are implied by the RAR. The three quantitative relations (3), (4), and (5) all involve the same acceleration scale: $\mathrm{g}_\dagger$ directly in (5),
through $\mathrm{g}_\dagger = G\,\Sigma_\dagger$ in (4),
and in (3),

$$\mathrm{g}_\dagger = \frac{\zeta V_f^4}{G M_b}$$

where $\zeta \approx 0.8$ accounts for the cylindrical geometry of disk galaxies. This order-unity factor is a mild function of radius that can be derived from the equation for the rotation curve of a razor thin exponential disk.

Cosmological Framework

Dark matter halos are thought to form by gravitational collapse of over-dense regions in an otherwise expanding universe.

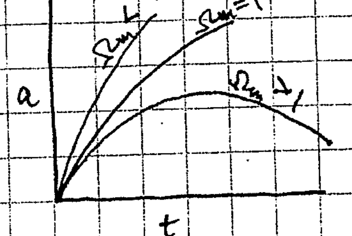
$$a = \frac{1}{1+z}$$

A little necessary context:

Friedmann eqn:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3}$$

scale size, mass density, curvature, cosmological constant



a is the scale size of the universe — a dimensionless quantity that encodes the physical separation between comoving coordinate tracers (e.g., galaxies w/ zero peculiar motion)

a(t) is the expansion history of the universe following from the solution of Friedmann's eqn.

Note that the Hubble parameter $H = \frac{\dot{a}}{a}$ is the expansion rate.

H must vary with time; it's current measured value is the misnamed Hubble "constant" $H_0 = \left(\frac{\dot{a}}{a}\right)_0$ measured now at $t = t_0$, the age of the U.

It is also convenient to define the density parameter $\Omega_m = \frac{\rho}{\rho_{crit}}$

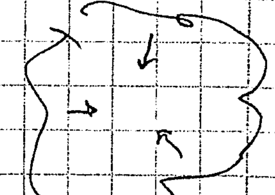
which is the ratio of the actual mass density to the critical density

$$\rho_{crit} = \frac{3H^2}{8\pi G}$$ that defines the over/under between eternal expansion and eventual recollapse.

Note that ρ_{crit} evolves with H, as do ρ & Ω_m.
Only if $\Omega_m = 1$ exactly does it remain 1 for eternity.

Monolithic Galaxy Formation — Tophat overdensity

1. Collapse of giant primordial gas clouds



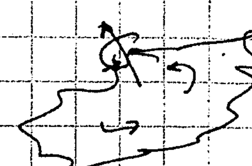
2. Stars start to form during collapse
   - low metallicity
   - eccentric orbits
   - oldest stars present

   } → Globular clusters + Galactic stellar halo

   orbits retain memory of collapse

stars don't dissipate!

3. Gas settles into disk
   - gas dissipates
   - settles into plane, specified by net initial angular momentum



   - collapses & contracts until rotationally supported

     $\lambda_i = 0.05 \rightarrow 20\times$ collapse

     where $\lambda = \frac{J|E|^{1/2}}{G M^{5/2}}$ is the dimensionless angular momentum "spin parameter" from primordial tidal torques
   - halo adiabatically compressed by gas infall

4. Stars form in disk

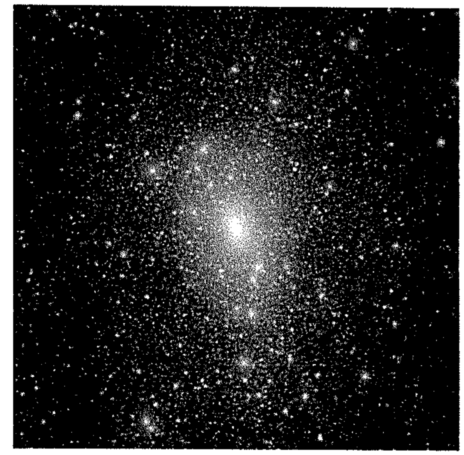
   might evolve as closed box,
   or gas accretion may continue at a low rate
   (helps with G-dwarf problem)

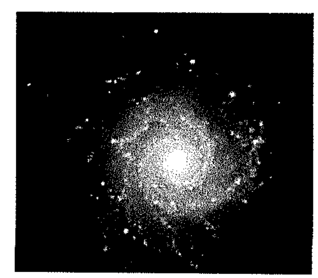
Has to happen in this sequence:
Only gas can dissipate to form a disk
Once formed, stars retain memory of their initial conditions
Hence getting all the stars to orbit in the same direction in the same plane requires the gas disk to settle first.

Hierarchical galaxy formation with Cold Dark Matter

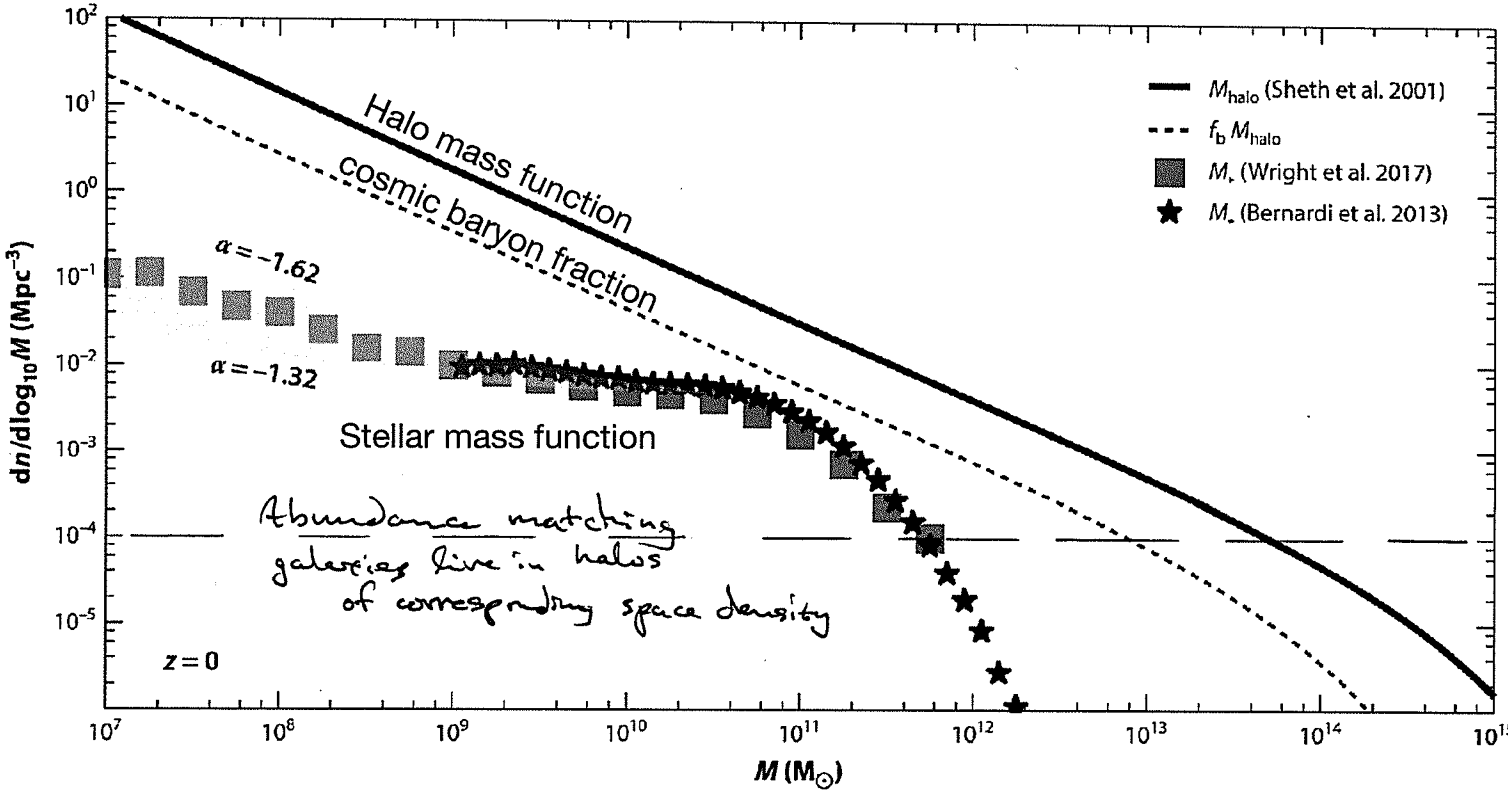
Sequence of events in galaxy formation

1. Dark matter halos form; merge into ever larger masses
2. Baryons fall into the potential wells of DM halos
3. Gas cools & dissipates, sinks to centers of DM halos
    - Halos compressed by sinking baryons
    - gas forms rotating disks at centers of DM halos
4. Stars form in disks
    - Feedback heats gas, dissuading further gas accretion
    - might rearrange dark matter
5. Mergers transform some disks into ellipticals
    - star formation enhanced then truncated by mergers
6. Renewed gas accretion may re-form disks around spheroids
    - thus becoming the bulges of S0s and early type (Sa, Sb) spirals
7. Merging slows; more gradual accretion of dark matter and gas may continue
8. Galaxies


dark matter from simulation


spiral galaxy

Halo and stellar mass function
(Bullock & Boylan-Kolchin)



Bullock JS, Boylan-Kolchin M. 2017.
Annu. Rev. Astron. Astrophys. 55:343–87

$f_* = \frac{M_*}{M_{200}}$ varies with M_*

Schechter fn: $\Phi(M_*) = \Phi^* e^{-\left(\frac{M}{M^*}\right)} \left(\frac{M}{M^*}\right)^{\alpha}$

Φ^* characteristic density
M^* characteristic mass
α faint end slope

integrated mass density $\rho = \int M \Phi(M) dM = M^* \Phi^* \Gamma(\alpha + 2)$ incomple Γ fn ~ 1 for $\alpha \sim -1$

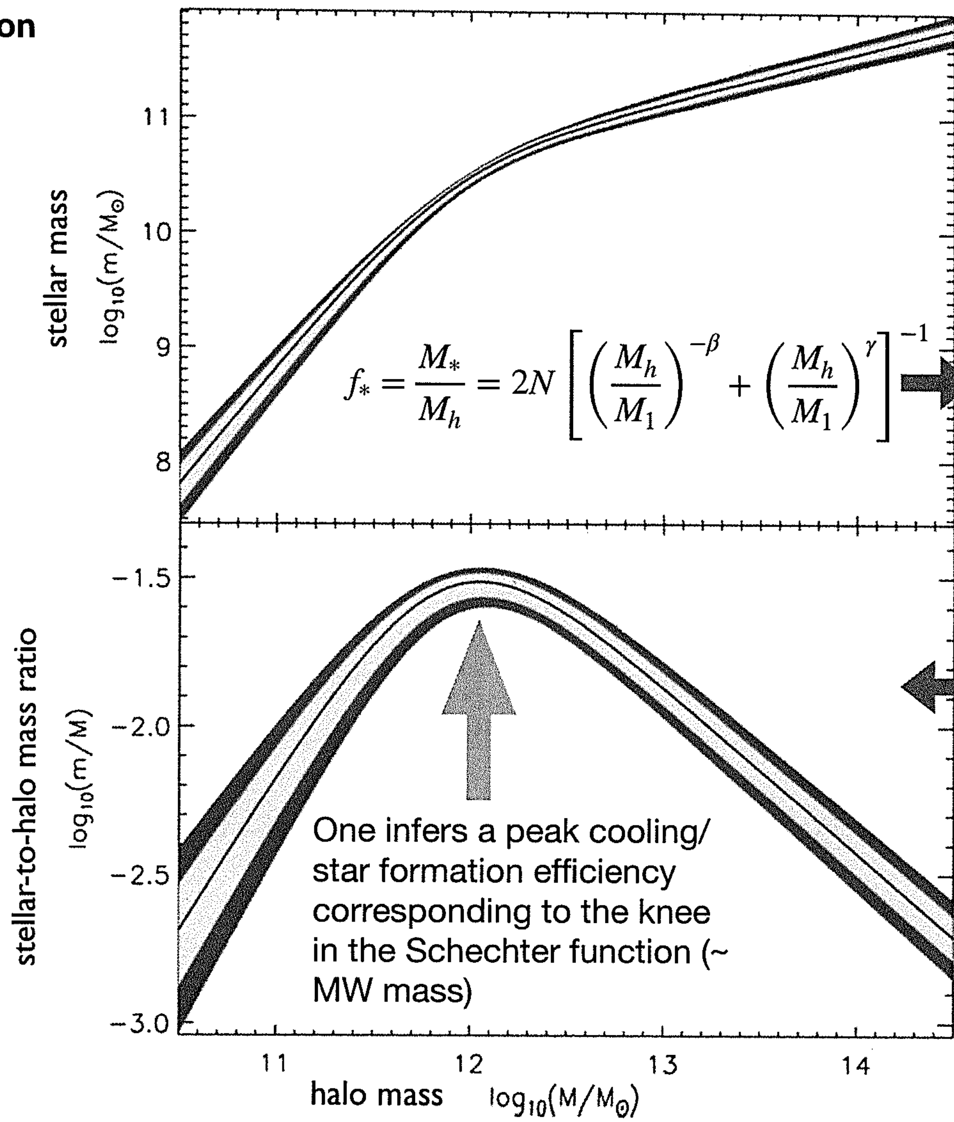
Stellar Mass-Halo Mass Relation from Abundance Matching

Behroozi et al (2013)
Moster et al (2013)
Kravstov et al (2019)

From
"abundance matching"

Match the number density of simulated dark matter halos to that observed for galaxies as a fcn of mass

Note that a large range in stellar mass gets wedged into a narrow range in halo mass.



Common fitting function used to describe stellar mass-halo mass relation where M_1, β, γ, N are parameters to fit the shape seen at bottom. There's a bend point, a low- and high-mass slope, and a normalization. This is useful for relating the halo mass to the observable stellar mass.