STRUCTURE AND DYNAMICS OF GALAXIES

1. Distribution of stars in the Milky Way Galaxy

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/∼vdkruit

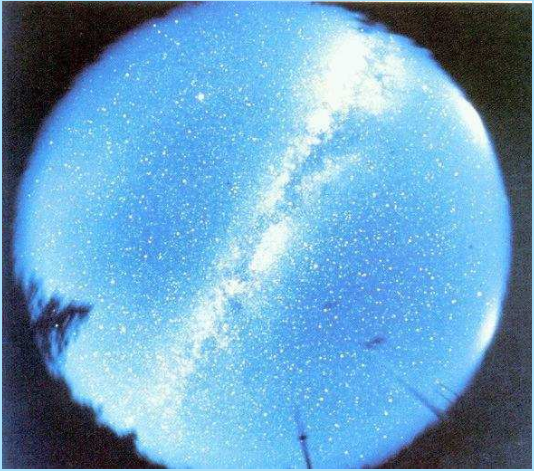
Beijing, September 2011

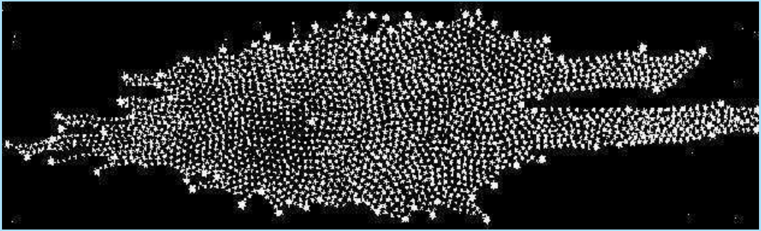
Outline

Historical introduction

Herschel and Kapteyn

Our **Galaxy** can be seen on the sky as the Milky Way, a band of faint light.



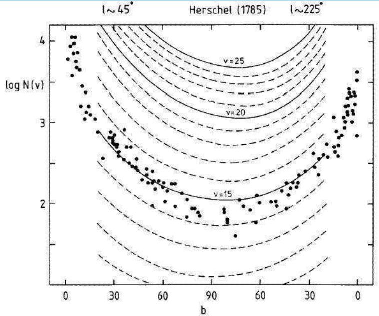


The earliest attempts to study the structure of the Milky Way Galaxy (the Sidereal System; really the whole universe) on a global scale were based on star counts.

William Herschel (1738 – 1822) performed such “star gauges” and assumed that (1) all stars have equal intrinsic luminostities and (2) he could see stars out ot the edges of the system.

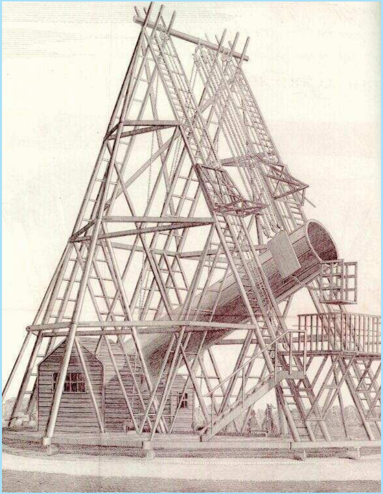
Then the distance to the edge of the system in any direction is proportional to the square-root of the number of stars per square degree.

It can be shown by comparing to current star counts that Herschel counted stars down to about visual magnitude 14.5[1].



⇒ Counted down to ≈ 15 V-mag.

[1] P.C. van der Kruit, A.&A. 157, 244 (1986)



From "Equalisation of starlight"-experiments Herschel estimated his "Space-penetrating powers":

Unaided eye: 12 times Sirius

20-ft telescope: 75 times unaided eye

⇒ 14.8 mag fainter than brightest stars.

Jacobus C. Kapteyn (1851 – 1922) improved upon this by determining locally the luminosity function $\Phi(M)$, that is the frequency distribution of stars as a function of their absolute magnitudes.

The observed distribution of stars $N_{\rm m}$ in a given direction as a function of apparent magnitude m relates to the space density of stars $\Delta(\rho)$ at distance ρ as

$$\frac{dN_{\rm m}}{dm} = 0.9696 \int_0^\infty \rho^2 \Delta(\rho) \Phi(m - 5 \log \rho) d\rho$$

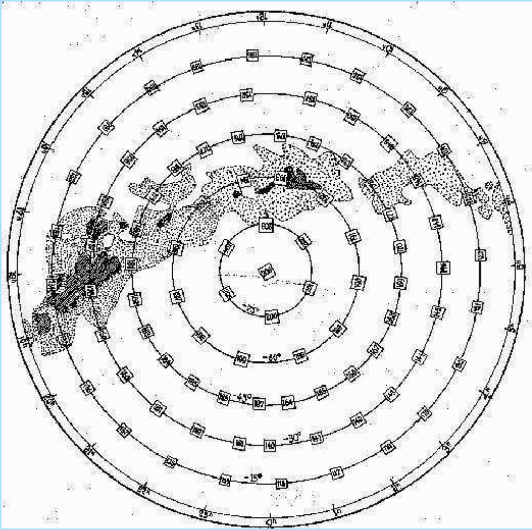
Kapteyn proceeded to investigate (numerical) methods to invert this integral equation in order to solve it.

Kapteyn suspected that interstellar absorption was present and even predicted that it would give rise to reddening[2].

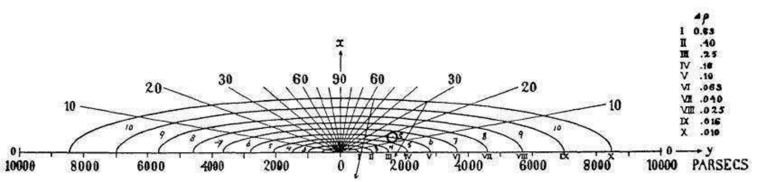
But he found that the reddening was small (0.031 ± 0.006 mag per kpc in modern units) and chose to ignore it.

Under Kapteyn's leadership an international project on Selected Areas over the whole sky to determine star counts (and eventually spectral types and velocities) in a systematic way was started.

[2] J.C. Kapteyn, Ap.J. 29, 46 & 30, 284/398 (1909)



Towards the end of his life he used star counts to construct what became known as the Kapteyn Universe[3]:



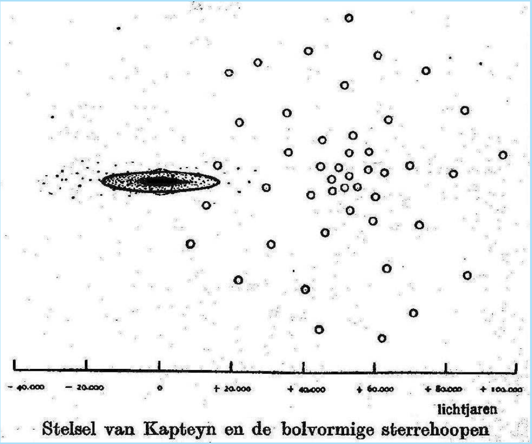
The Sun is near the center.

That was suspicious and later was found to result from the neglect of insterstellar absorption.

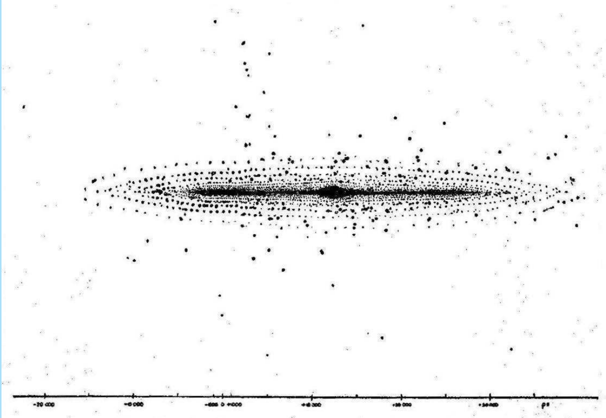
[3] J.C Kapteyn & P.J. van Rhijn, Ap.J. 52, 23 (1920); J.C. Kapteyn, o.J. 55, 302 (1922)

Shapley and Hubble

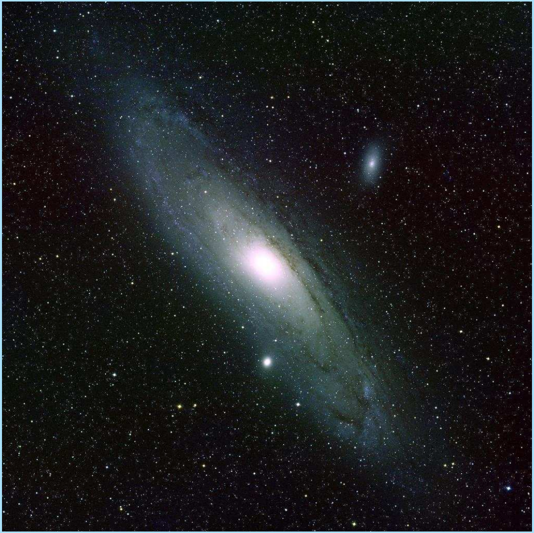
Indeed the work of Harlow Shapley (1885 – 1972) on the distances of Globular Clusters showed that the Sidereal System really was much larger.



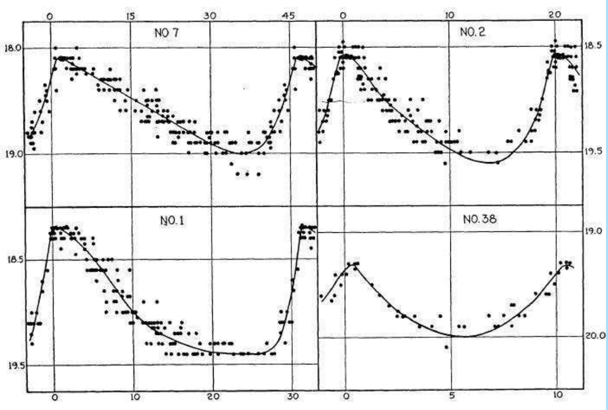
Astronomers like Jan H. Oort (1900 – 1992) found that absorption reconciled the two models.



Dr. Oort's diagram van het melkwegstelsel



An important step was made by Edwin Hubble (1889–1953), who showed, using Cepheids, that the Andromeda Nebula is an 'Island Universe', a separate stellar system outside the Galaxy.



Hubble[4] found a distance of 275 kpc. The current value is 780 kpc.

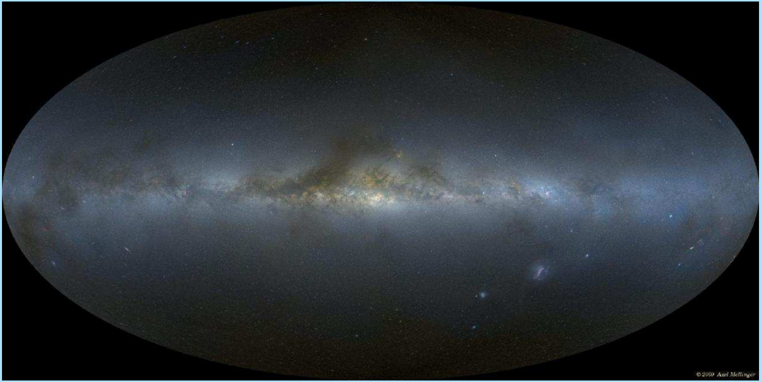
[4] E. Hubble, Ap.J. 69, 103 (1929)

So the Galaxy is one of very many, seen edge-on.

Luminosity distribution in the Galaxy

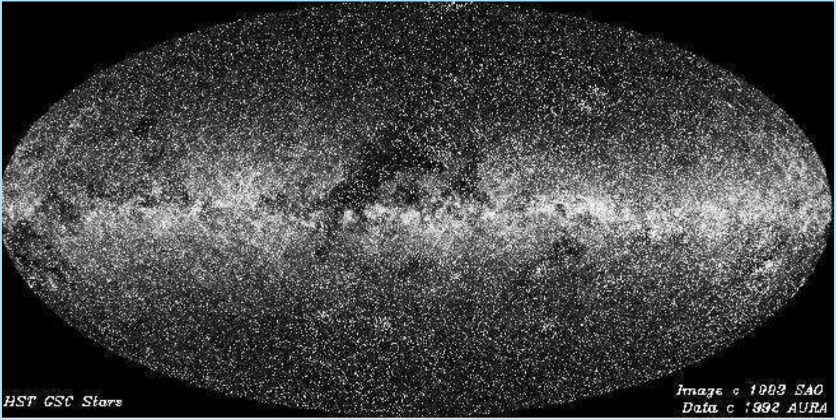
Modern views of the Milky Way

Here is a composite picture[5] covering the full sky at $36''\mathrm{pixel}^{-1}$.

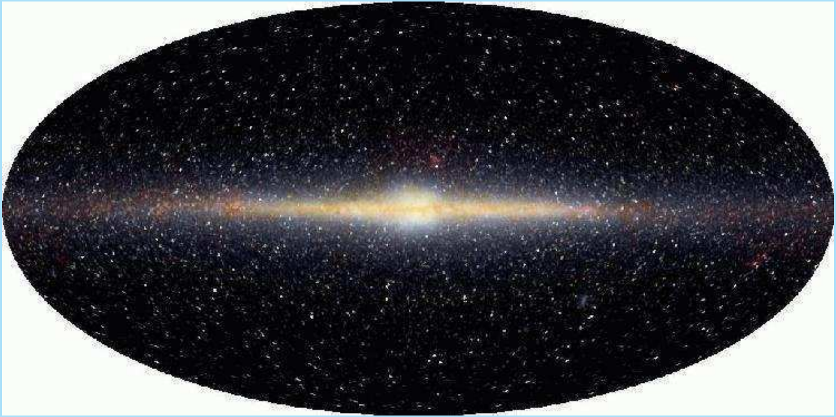


[5] A. Mellinger, P.A.S.P. 121, 1180 (2009); also Astronomy Picture of the Day for 2009 November 25: antwrp.gsfc.nasa.gov/apod/ap091125.html

Here is a plot of all stars in the Guide Star Catalogue of the Hubble Space Telescope down to about magnitude 16.



The Cosmic Background Explorer (COBE) satellite did see the Milky Way in the near-infrared as follows:



Direct measurements of the surface brightness of the Galaxy are difficult due to other contributions:

The sky contributions in the visual with some comparisons are as follows:

	$S_{10}(V)_{\mathrm{G2V,V}}$	V-mag arcsec^{-2}
Disk of sun	$\sim 10^{17}$	$\sim$ -15
Daylight	$\sim 3 \times 10^{11}$	$\sim$ -1
Full moon	$\sim 10^{11}$	0.5
Airglow	50	23.5
Zodiacal light (ecliptic)	180	22.0
Zodiacal light (pole)	80	23.0
Bright stars ($m_{\mathrm{V}} < 6$)	20	24.5
Integrated starlight (plane)	300	21.5
Integrated starlight (pole)	30	24.0
Diffuse Galactic light (plane)	50	23.5
Diffuse Galactic light (pole)	2	27.0
Cosmic background	~ 1	~ 28.0

The property $S_{10}(V)_{\mathrm{G2V},\lambda}$ denotes the equivalent number of G2V-stars in the λ-band per square degree that have magnitude 10 in the V-band.

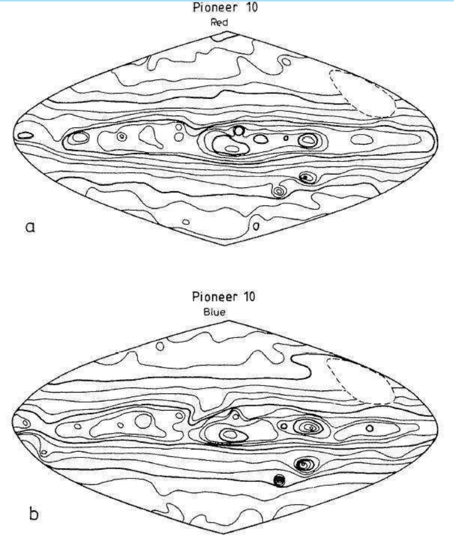
Pioneer 10 photometry

The zodiacal light is the biggest problem when studying the background distribution of starlight.

The problem is the reverse for people interested in studying zodiacal light.

The satellite Pioneer 10 was launched in March 1972 and reached Jupiter in December 1973.

During its trip in the asteroid belt and beyond it swept the skies and made a map of the background starlight free of zodiacal light.


Pioneer 10
Red
a
Pioneer 10
Blue
b

STRUCTURE AND DYNAMICS OF GALAXIES

2. Kinematics of the Milky Way

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www. astro.rug.nl/∼vdkruit

Beijing. September 2011

Outline

Differential rotation

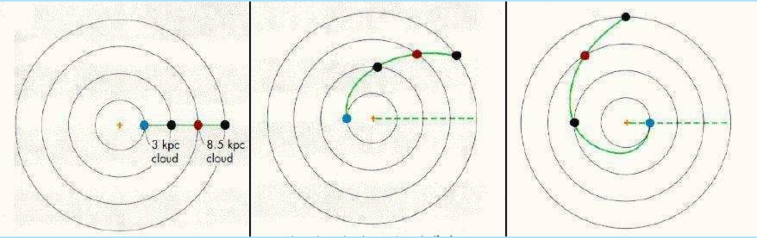
- Relative motions
- Local approximations and Oort constants

Rotation curves and mass distributions

Differential rotation

Relative motions

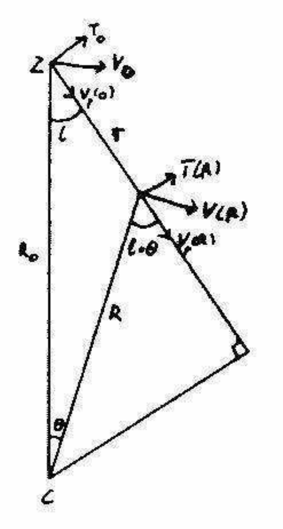
The Galaxy does not rotate like a solid wheel. The period of revolution varies with distance from the center. This is called differential rotation.

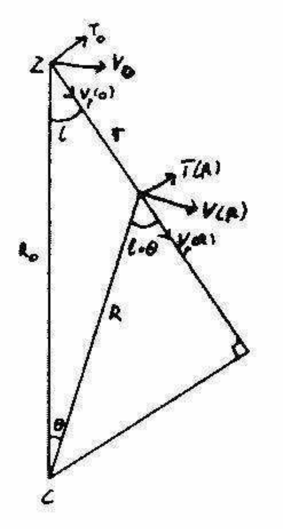


Each part moves with respect to those parts that do not happen to be at the same galactocentric distance.

Say, the rotation speed is $V(R)$ and in the solar neighborhood it is $V_\circ$.

If the Sun Z is at a distance $R_\circ$ from the center C, then an object at distance r from the Sun at Galactic longitude l has a radial velocity w.r.t. the Sun $V_{\rm rad}$ and a tangential velocity T.





$$V_{\rm rad} = V_{\rm r}(R) - V_{\rm r}(0) = V(R)\sin(l+\theta) - V_\circ \sin l$$

$$T = T(R) - T(0) = V(R)\cos(l+\theta) - V_\circ \cos l$$

$$R\sin(l+\theta) = R_\circ \sin l$$

$$R\cos(l+\theta) = R_\circ \cos l - r$$

Substitute this and we get

$$V_{\rm rad} = R_\circ \left(\frac{V(R)}{R} - \frac{V_\circ}{R_\circ} \right) \sin l \tag{1}$$

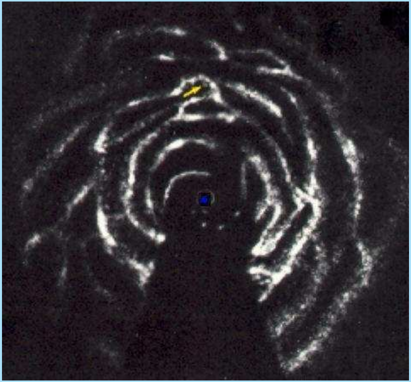
$$T = R_\circ \left(\frac{V(R)}{R} - \frac{V_\circ}{R_\circ} \right) \cos l - \frac{r}{R} V(R) \tag{2}$$

So, if we would know the rotation curve $V(R)$ we can calculate the distance R from observations of $V_{\rm rad}$. From this follows r with an ambiguity symmetric with the sub-central point.

The latter is that point along the line-of-sight that is closest to the Galactic Center.

$V(R)$ can be deduced in each direction l by taking the largest observed radial velocity. This will be the rotation velocity at the sub-central point.

With the 21-cm line of HI, the distribution of hydrogen in the Galaxy has been mapped[1]. This was the first indication that the Galaxy is a spiral galaxy.



[1] K.K. Kwee, C.A. Muller & G. Westerhout, Bull. Astron. Inst. Neth. 12, 211 (1954); J.H. Oort, F.J. Kerr & G. Westerhout, Mon.Not.R.A.S. 118, 379 (1958) and J.H. Oort, I.A.U. Symp. 8, 409 (1959)

Local approximations and Oort constants

We now make local approximations; that is $r \ll R_\circ$.

Change to angular velocities $\omega(R) = V(R)/R$ and $\omega_\circ = V_\circ/R_\circ$ and make a Tayler expansion

$$f(a+x) = f(a) + x\frac{df(a)}{da} + \frac{1}{2}x^2\frac{d^2f(a)}{d^2a} + \ldots.$$

for the angular rotation velocity

$$\omega(R) = \omega_\circ + (R - R_\circ)\left(\frac{d\omega}{dR}\right)_{R_\circ} + \tfrac{1}{2}(R - R_\circ)^2\left(\frac{d^2\omega}{dR^2}\right)_{R_\circ}$$

The cosine-rule gives

$$R = R_\circ \left[1 + \left(\frac{r}{R_\circ} \right)^2 - \frac{2r}{R_\circ} \cos l \right]^{1/2}$$

Make a Tayler expansion for this expression and ignore terms of higher order than $(r/R_\circ)^3$.

$$R = R_\circ \left[1 - \frac{r}{R_\circ} \cos l + \frac{1}{2} \left(\frac{r}{R_\circ} \right)^2 (1 - \cos^2 l) \right]$$

$$R - R_\circ = -r \cos l + \frac{1}{2} \frac{r^2}{R_\circ} (1 - \cos^2 l)$$

$$(R - R_\circ)^2 = r^2 \cos^2 l$$

Substitute this in the equation for ω

$$\begin{aligned}\omega(R) &= \omega_\circ + \left(\frac{d\omega}{dR}\right)_{R_\circ} R_\circ \left[-\frac{r}{R_\circ}\cos l + \frac{1}{2}\left(\frac{r}{R_\circ}\right)^2 (1-\cos^2 l)\right] \\ &+ \frac{1}{2}\left(\frac{d^2\omega}{dR^2}\right)_{R_\circ} R_\circ^2 \left(\frac{r}{R_\circ}\right)^2 \cos^2 l\end{aligned}$$

or in linear velocity

$$\begin{aligned}V_{\mathrm{rad}} &= \left(\frac{r}{R_\circ}\right)^2 \left(\frac{d\omega}{dR}\right)_{R_\circ} \frac{R_\circ^2}{2}\sin l - \frac{r}{R_\circ}\left(\frac{d\omega}{dR}\right)_{R_\circ} R_\circ^2 \sin l \cos l \\ &+ \frac{1}{2}\left(\frac{r}{R_\circ}\right)^2 \left[-\left(\frac{d\omega}{dR}\right)_{R_\circ} R_\circ^2 + \left(\frac{d^2\omega}{dR^2}\right)_{R_\circ} R_\circ^3\right] \sin l \cos^2 l\end{aligned}$$

Use $2\sin l \cos l = \sin 2l$ and ignore terms with $(r/R_\circ)^2$ and higher orders. Then

$$V_{\rm rad} = -\tfrac{1}{2} R_\circ \left(\frac{d\omega}{dR}\right)_{R_\circ} r \sin 2l \equiv A r \sin 2l$$

So, stars at the same distance r will show a systematic pattern in the magnitude of their radial velocities accross the sky with Galactic longitude.

For stars at Galactic latitude b we have to use the projection of the velocities onto the Galactic plane:

$$V_{\rm rad} = A r \sin 2l \cos b$$

For the tangential velocities we make a change to proper motions μ. In equivalent way we then find

$$\frac{T}{r} = 4.74\mu = -\omega_\circ + \frac{3}{2}\left(\frac{d\omega}{dR}\right)_{R_\circ} r\cos l - \left(\frac{d\omega}{dR}\right)_{R_\circ} R_\circ \cos^2 l + \frac{r}{2R}\left[-\left(\frac{d\omega}{dR}\right)_{R_\circ} + \left(\frac{d^2\omega}{dR^2}\right)_{R_\circ} R_\circ^2\right]\cos^3 l$$

Now use $\cos^2 l = \frac{1}{2} + \frac{1}{2}\cos 2l$ and ignore all terms $(r/R_\circ)$ and higher order.

$$\begin{aligned} 4.74\mu &= -\omega_\circ - \tfrac{1}{2}\left(\frac{d\omega}{dR}\right)_{R_\circ} R_\circ - \tfrac{1}{2}R_\circ\left(\frac{d\omega}{dR}\right)_{R_\circ}\cos 2l \\ &\equiv B + A\cos 2l \end{aligned}$$

Now the distance dependence has of course disappeared. Agian for higher Galactic latitude the right-hand side will have to be multiplied by $\cos b$.

The constants A and B are the Oort constants. Oort first made the derivation above (in 1927) and used this to deduce the rotation of the Galaxy from observations of the proper motions of stars.

The Oort constanten can also be written as

$$A = \frac{1}{2}\left[\frac{V_\circ}{R_\circ} - \left(\frac{dV}{dR}\right)_{R_\circ}\right]$$

$$B = -\frac{1}{2}\left[\frac{V_\circ}{R_\circ} + \left(\frac{dV}{dR}\right)_{R_\circ}\right]$$

Furthermore

$$A + B = -\left(\frac{dV}{dR}\right)_{R_\circ} \quad ; \quad A - B = \frac{V_\circ}{R_\circ}$$

Current best values are

$$\begin{array}{ll} R_\circ \sim 8.5 \text{ kpc} & A \sim 13 \text{ km s}^{-1} \text{ kpc}^{-1} \\ V_\circ \sim 220 \text{ km s}^{-1} & B \sim -13 \text{ km s}^{-1} \text{ kpc}^{-1} \end{array}$$

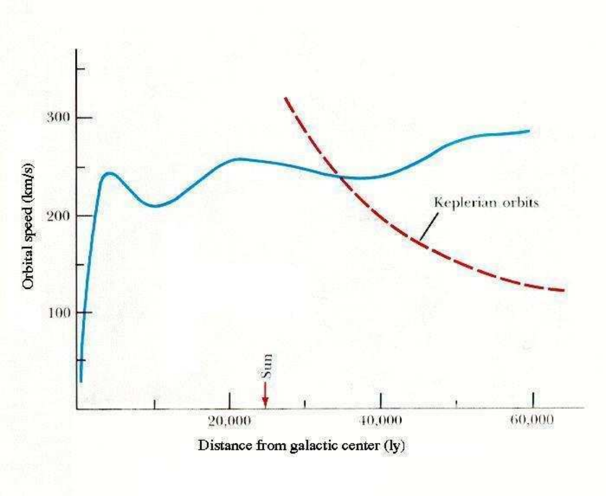
Rotation curves and mass distributions

The rotation curve $V(R)$ is difficult to derive beyond $R_\circ$ and this can only be done with objects of known distance such as HII regions).

In a circular orbit around a point mass M we have $M = V^2R/G$ (as in the Solar System). This is called a Keplerian rotation curve.

One expects that the rotation curve of the Galaxy tends to such a behavior as one moves beyond the boundaries of the disk. However, we do see a flat rotation curve.

One determination of the Galactic rotation curve:



We see that up to large distances from the center the rotation velocity does not drop.

We also see this in other galaxies. It shows that more matter must be present than what we observe in stars, gas and dust and this is called dark matter.

With the formula estimate the mass within $R_\circ$ as $\sim 9.6 \times 10^{10} M_\odot$.

At the end of the measured rotation curve this enclosed mass becomes $\sim 10^{12} M_\odot$.

STRUCTURE AND DYNAMICS OF GALAXIES

3. Stellar Populations, classification of galaxies

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/~vdkruit

Beijing, September 2011

Outline

Stellar Populations

Origin of the concept
Vatican Symposium
The current situation

Classification

Definition by Hubble and later extensions
Correlations along the Hubble sequence

Origin of the concept

Origin of the concept

Lindblad[1] in 1925 argued that the Galaxy is made up of a set of components with a continuous range of flattening.

Baade[2] in 1944 resolved red stars in the central regions of M32 and the elliptical companions and introduces the concept of two stellar populations, mainly based on the characteristics of their H-R diagrams. Population I is in the disk and has blue stars and

Population II in the halo with globular cluster type H-R diagrams with red stars the brightest.

[1] B. Lindblad, Arkiv. Mat. Astron. Fysik 19A, No. 21 (1925)
[2] W. Baade, Ap.J. 100, 137 and 147 (1944)

THE RESOLUTION OF MESSIER 32, NGC 205, AND THE CENTRAL REGION OF THE ANDROMEDA NEBULA*

W. BAADE
Mount Wilson Observatory

Received April 27, 1944

ABSTRACT

Recent photographs on red-sensitive plates, taken with the 100-inch telescope, have for the first time resolved into stars the two companions of the Andromeda nebula—Messier 32 and NGC 205—and the central region of the Andromeda nebula itself. The brightest stars in all three systems have the photographic magnitude 21.3 and the mean color index +1.3 mag. Since the revised distance-modulus of the group is $m - M = 22.4$, the absolute photographic magnitude of the brightest stars in these systems is $M_{pg} = -1.1$.

The Hertzsprung-Russell diagram of the stars in the early-type nebulae is shown to be closely related to, if not identical with, that of the globular clusters. This leads to the further conclusion that the stellar populations of the galaxies fall into two distinct groups, one represented by the well-known H-R diagram of the stars in our solar neighborhood (the slow-moving stars), the other by that of the globular clusters. Characteristic of the first group (type I) are highly luminous O- and B-type stars and open clusters; of the second (type II), short-period Cepheids and globular clusters. Early-type nebulae (E–Sa) seem to have populations of the pure type II. Both types seem to coexist in the intermediate and late-type nebulae.

The two types of stellar populations had been recognized among the stars of our own galaxy by Oort as early as 1926.

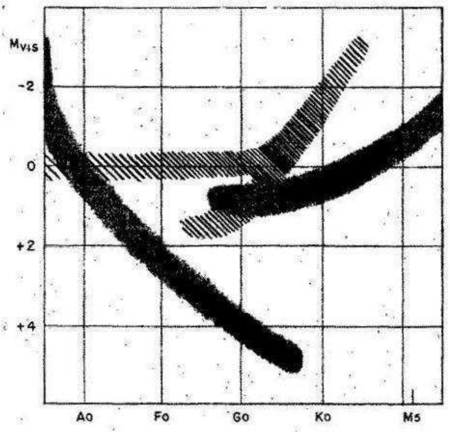
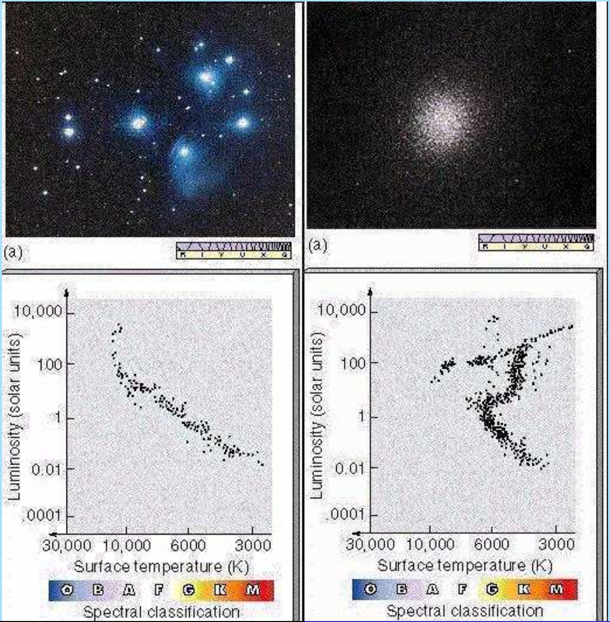
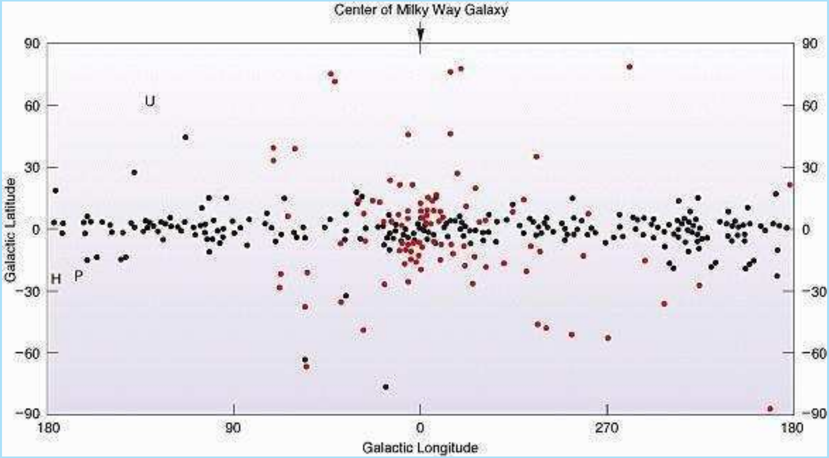


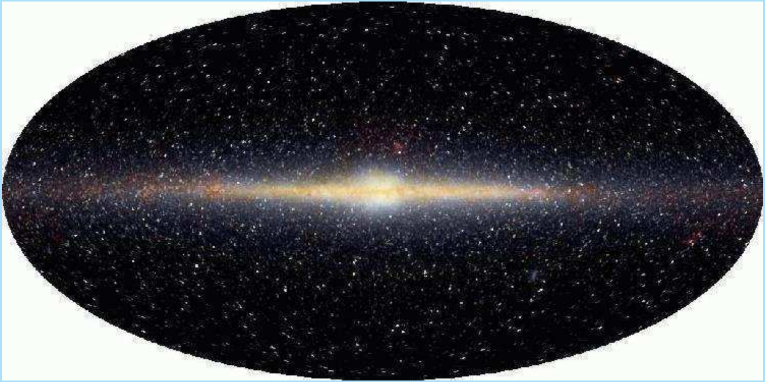
FIG. 1.—Shaded areas: ordinary H-R diagram (type I). Hatched area: H-R diagram of stars in globular clusters (type II).


(a)
10,000
100
1
0.01
.0001
Luminosity (solar units)
30,000 10,000 6000 3000
Surface temperature (K)
O B A F G K M
Spectral classification
(a)
10,000
100
1
0.01
.0001
Luminosity (solar units)
30,000 10,000 6000 3000
Surface temperature (K)
O B A F G K M
Spectral classification

The Galaxy as consisting of two basic populations can be seen in the distribution on the sky of globular (red) versus galactic clusters.



and in the near-infrared image of the Galaxy with the DIRBE experiment on board the Cosmic Background Explorer COBE.



Vatican Symposium

In 1957 the Vatican Symposium on stellar populations defined five stellar populations with a decreasing age, increasing flattening and increasing metal abundance.

Population	$\lvert z \rvert$ (pc)	$\lvert V_z \rvert$ (km/s)	Typical members
Extreme Pop. I	120	8	Gas, Young stars associated with spiral structure, Supergiants, Cepheids, T Tauri stars, Galactic Clusters of Trumpler's Class I
Older Pop. I	160	10	A-Type stars, Strong-line stars
Disk Population	400	17	Stars of galactic nucleus, Planetary Nebulae, novae, RR Lyrae stars with periods below 0.4 days, Weak-line stars
Interm. Pop. II	700	25	"High-velocity stars" with z-velocities exceeding 30 km/sec, Long-period variables <M5e with periods below 250 days
Halo Pop. II	2000	75	Subdwarfs, Globular clusters with high z-velocity, RR Lyrae stars with periods longer than 0.4 days

The current situation.

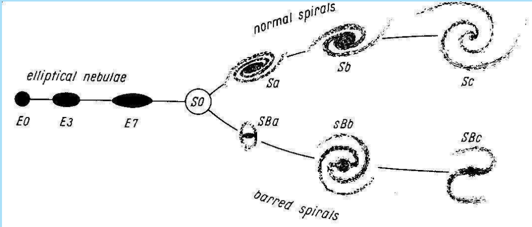
- Dark halo, presumably non-baryonic.
- Population II.
- Thick disk.
- Old disk, sometimes called thin disk.
- Population I.

Classification

Definition by Hubble and later extensions

Classification systems have been described in detail by Allan Sandage in Volume IX of “Stars and Stellar Systems”[3].

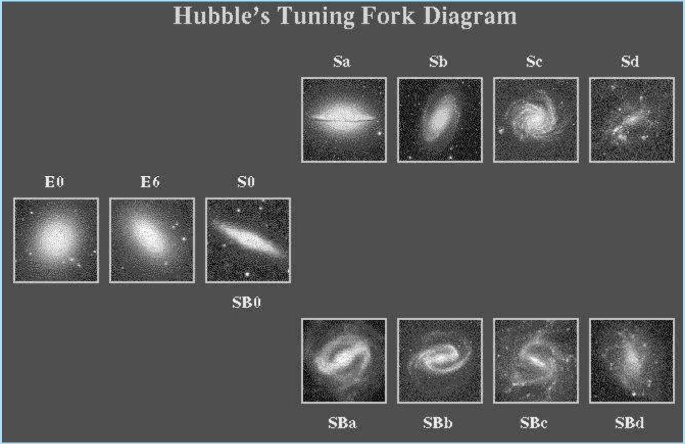
The Hubble classification scheme starts with Hubble’s scheme of the 1920’s (his well-known tuning fork).



[3] Available at http://nedwww.ipac.caltech.edu/level5/Sandage/frames. html

Originally the S0 class was not included. Hubble introduced it in the 1930's.

Here is a modern WWW-version of the Tuning Fork.

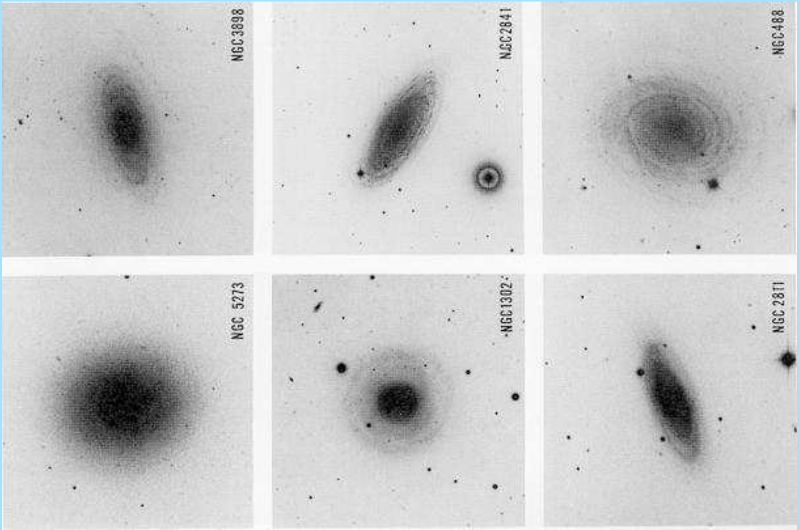


The Hubble Classification System has the following criteria:

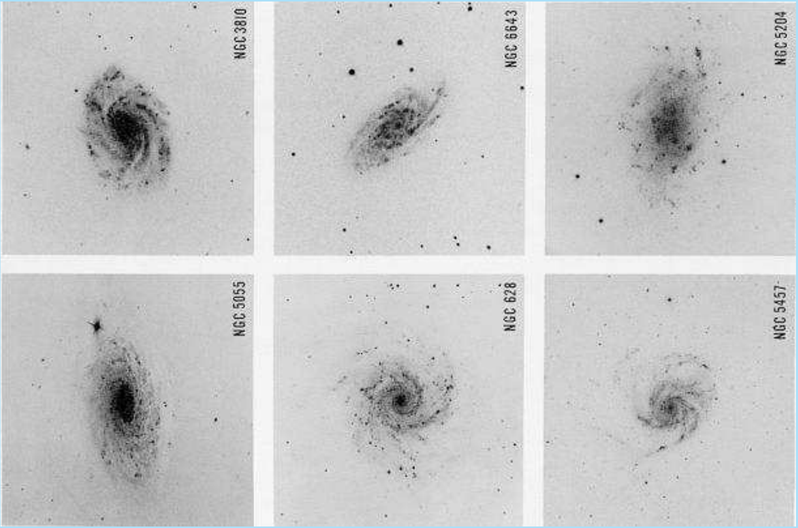
- Ellipticals – E0 to E7 depending on the apparent flattening (En with $n = 10 \times (a - b)/a$).
- Spirals either with or without a bar (S or SB) and subclasses a to c depending on
  - Bulge-to-disk ratio
  - Pitch angle of spiral arms
  - Development of arms ("strength" of HII regions)
- Irregulars Irr

The following figures from Sandage's paper illustrate the system.

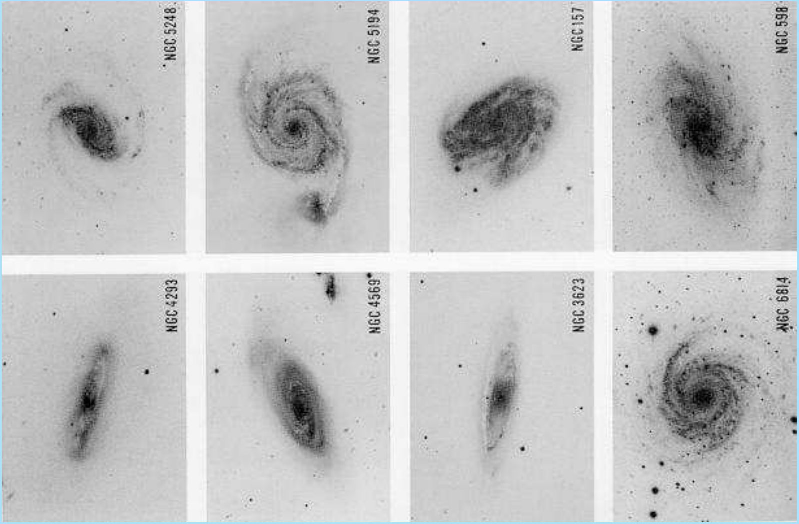
S0 and Sa with thin arms.



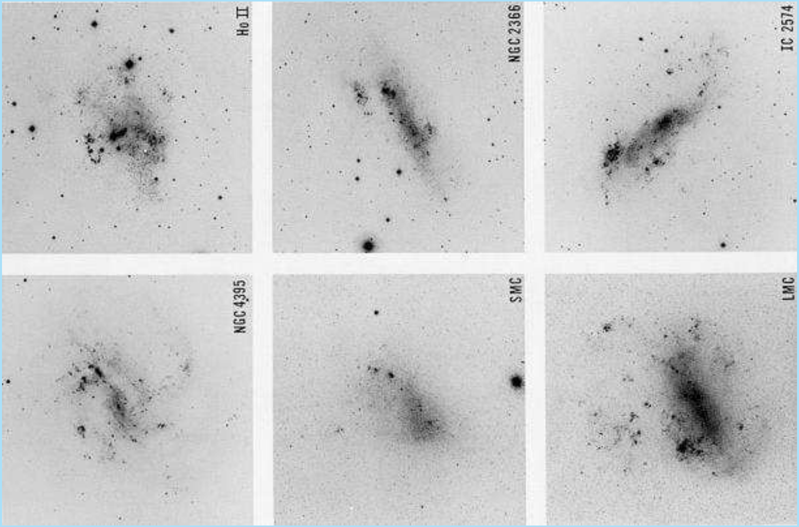
Sb and Sc with thin arms.



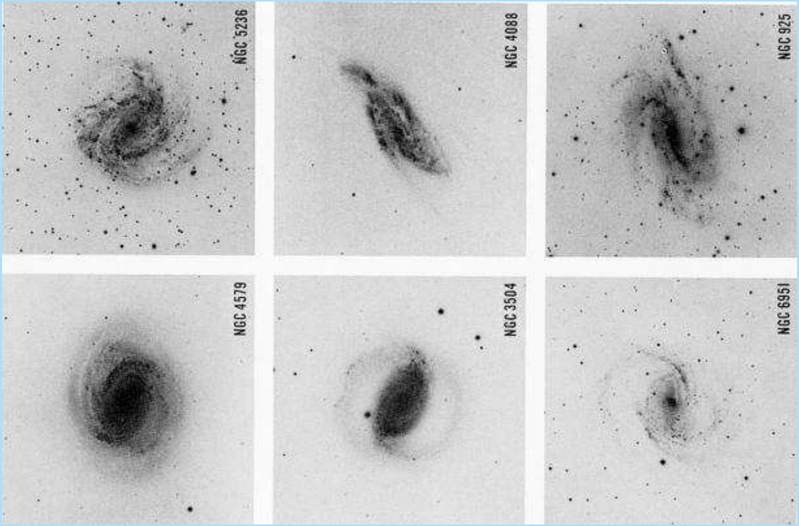
Sa to Sc with heavy arms.



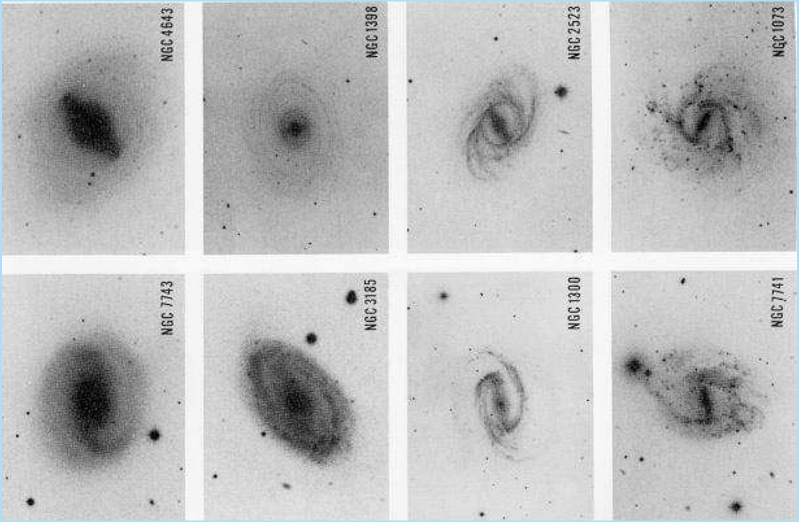
Irregulars Irr , later called Sd and Sm.



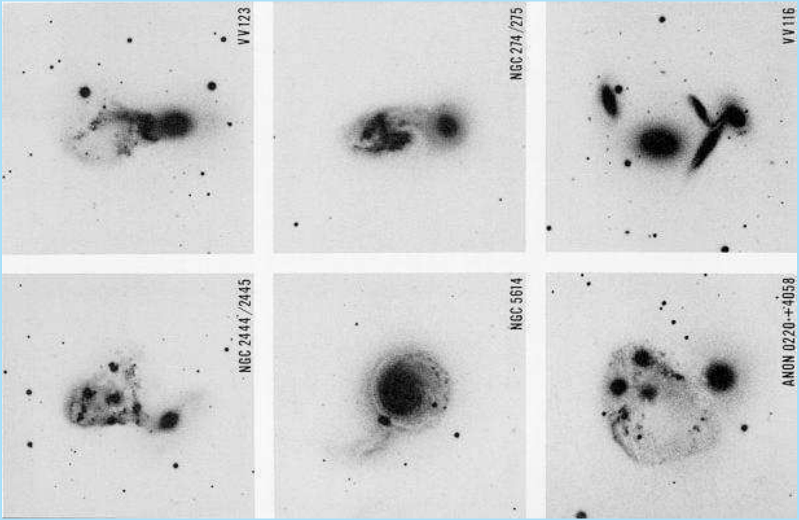
Spirals with small bars (SAB).



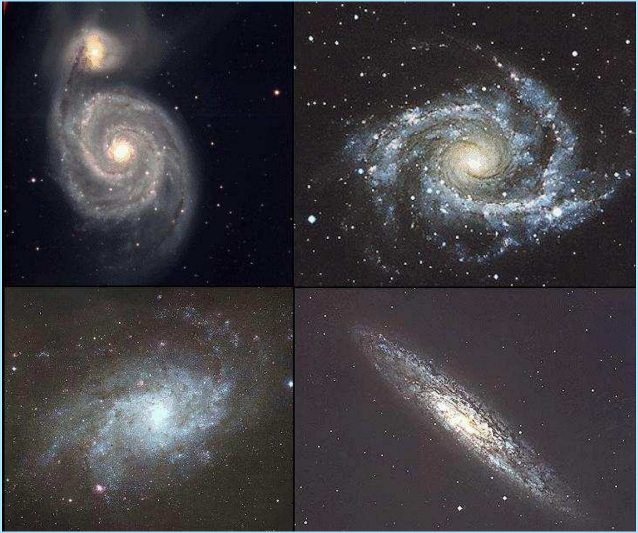
Spirals with heavy bars (SB).



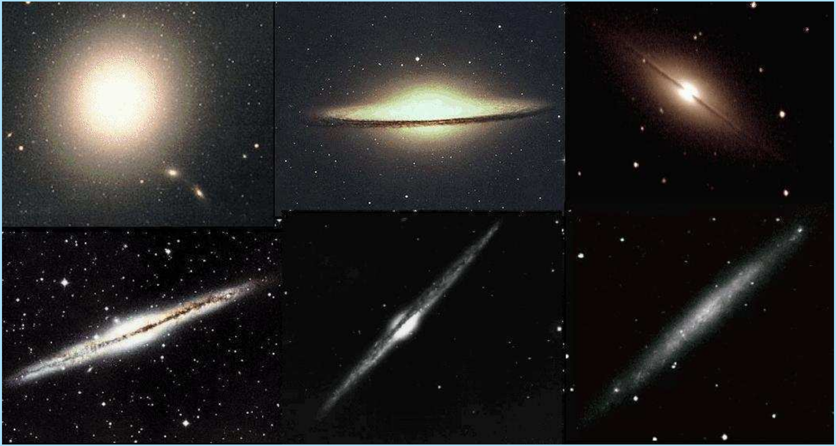
It is not possible to classify interacting galaxies.



Some pictures of galaxies with modern telescopes.



A set of pictures of edge-on galaxies along the Hubble sequence.



De Vaucouleurs later introduced Sd and Im ("Magellanic irregulars) to replace many of the Irr. Also he used the intermediate classification SAB.

He also introduced the varieties r (arms begin from an internal ring, often at the end of a bar) and s (no internal ring).

Correlations along the Hubble sequence

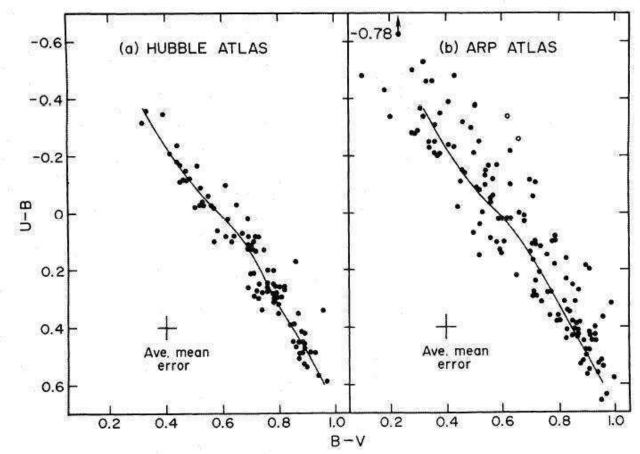
Hubble classification correlates with integrated colors[4] and relative HI content[5], so is apparently related to the history of star formation.

The colors of E-galaxies are about $(B - V) \sim 0.9$, $(U - B) \sim 0.6$ and those for late type galaxies $(B - V) \sim 0.4$, $(U - B) \sim -0.3$.

The HI content is expressed as the hydrogen mass to luminosity ratio

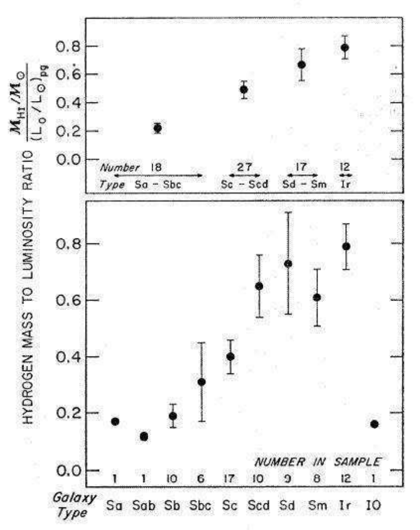
[4] R.B. Larson & B.M. Tinsley, Ap.J. 219, 46 (1978)
[5] M.S. Roberts, A.J. 74, 859 (1969)

The Hubble Atlas has normal galaxies; the Arp Atlas has disturbed and interacting galaxies.



Note that both the colors and these HI/L ratios are distance independent, since both are ratios of fluxes.

It follows that the Hubble sequence is one according to the relative importance of the two fundamental populations.



STRUCTURE AND DYNAMICS OF GALAXIES

4. Galactic dynamics: Fundamental equations

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/∼vdkruit

Beijing, September 20011

The collisionless Boltzmann equation

Studies of galactic dynamics start with two fundamental equations. The first is the *continuity equation*, also called the *Liouville* or *collisionless Boltzmann equation*.

It states that in any element of phase space the time derivative of the distribution function equals the number of stars entering it minus that leaving it, if no stars are created or destroyed.

Write the distribution function in phase space as $f(x, y, z, u, v, w, t)$ and the potential as $\Phi(x, y, z, t)$.

Now look first for the one-dimensional case at a position x, u. After a time interval dt the stars at $x - dx$ have taken the place of the stars at x, where $dx = udt$.

So the change in the distribution function is

$$df(x,u) = f(x - udt, u) - f(x,u)$$

$$\frac{df}{dt} = \frac{f(x - udt, u) - f(x,u)}{dt} = \frac{f(x - dx, u) - f(x,u)}{dx} u = -\frac{df(x,u)}{dx} u$$

For the velocity replace the positional coordinate with the velocity x with u and the velocity u with the acceleration du/dt. But according to Newton's law we can relate that to the force or the potential. So we get

$$\frac{df}{dt} = -\frac{df(x,u)}{du}\frac{du}{dt} = \frac{df(u,x)}{du}\frac{d\Phi}{dx}$$

The total derivative of the distribution function then is

$$\frac{\partial f(x,u)}{\partial t} + \frac{\partial f(x,u)}{\partial x} u - \frac{\partial f(x,u)}{\partial u} \frac{\partial \Phi}{\partial x} = 0$$

In three dimensions this becomes

$$\frac{\partial f}{\partial t} + u \frac{\partial f}{\partial x} + v \frac{\partial f}{\partial y} + w \frac{\partial f}{\partial z} - \frac{\partial \Phi}{\partial x} \frac{\partial f}{\partial u} - \frac{\partial \Phi}{\partial y} \frac{\partial f}{\partial v} - \frac{\partial \Phi}{\partial z} \frac{\partial f}{\partial w} = 0$$

Usually dynamical systems are assumed to be in equilibrium so that we have

$$u\frac{\partial f}{\partial x} + v\frac{\partial f}{\partial y} + w\frac{\partial f}{\partial z} - \frac{\partial \Phi}{\partial x}\frac{\partial f}{\partial u} - \frac{\partial \Phi}{\partial y}\frac{\partial f}{\partial v} - \frac{\partial \Phi}{\partial z}\frac{\partial f}{\partial w} = 0.$$

This is the collisionless Boltzmann equation

Usually (especially in disk galaxies) we work in cylindrical coordinates.

The distribution function then is $f(R, \theta, z, V_{\rm R}, V_\theta, V_{\rm z}, t)$ and the collisonless Boltzmann equation becomes

$$\begin{aligned} V_{\rm R}\frac{\partial f}{\partial R} + \frac{V_\theta}{R}\frac{\partial f}{\partial \theta} + V_{\rm z}\frac{\partial f}{\partial z} + \left(\frac{V_\theta^2}{R} - \frac{\partial \Phi}{\partial R}\right)\frac{\partial f}{\partial V_{\rm R}} & \\ - \left(\frac{V_{\rm R} V_\theta}{R} + \frac{1}{R}\frac{\partial \Phi}{\partial \theta}\right)\frac{\partial f}{\partial V_\theta} - \frac{\partial \Phi}{\partial z}\frac{\partial f}{\partial V_{\rm z}} & = 0. \end{aligned}$$

For axial symmetry this reduces to

$$V_{\mathrm{R}}\frac{\partial f}{\partial R}+V_{\mathrm{z}}\frac{\partial f}{\partial z}-\left(\frac{\partial \Phi}{\partial R}-\frac{V_\theta^2}{R}\right)\frac{\partial f}{\partial V_{\mathrm{R}}}-\frac{V_{\mathrm{R}}V_\theta}{R}\frac{\partial f}{\partial V_\theta}-\frac{\partial \Phi}{\partial z}\frac{\partial f}{\partial V_{\mathrm{z}}}=0.$$

For spherical symmetry this reduces further to

$$V_{\mathrm{R}}\frac{\partial f}{\partial R}-\left(\frac{\partial \Phi}{\partial R}-\frac{V_\theta^2}{R}\right)\frac{\partial f}{\partial V_{\mathrm{R}}}=0.$$

Here the velocity V_θ corresponds to the angular momentum of the system.

Poisson's equation

The second fundamental equation is Poisson's equation, which says that the gravitational potential derives from the combined gravitational forces of all the matter.

It can be written as

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} \equiv \nabla^2 \Phi = 4\pi G \rho(x, y, z)$$

In cylindrical coordinates

$$\frac{\partial^2 \Phi}{\partial R^2} + \frac{1}{R}\frac{\partial \Phi}{\partial R} + \frac{1}{R^2}\frac{\partial^2 \Phi}{\partial \theta^2} + \frac{\partial^2 \Phi}{\partial z^2} = 4\pi G \rho(R, \theta, z).$$

For the axisymmetric case

$$\frac{\partial K_{\mathrm{R}}}{\partial R} + \frac{K_{\mathrm{R}}}{R} + \frac{\partial K_{\mathrm{z}}}{\partial z} = -4\pi G\rho(R,z)$$

the spherical case

$$\frac{1}{R^2}\frac{\partial}{\partial R}\left(R^2\frac{\partial\Phi}{\partial R}\right) = 4\pi G\rho(R).$$

and the plane-parallel case

$$\frac{dK_{\mathrm{z}}}{dz} = -4\pi G\rho(z).$$

The collissionless Boltzmann and Poisson equations together completely describe the dynamics of a system.

The Poisson equation always refers to the total mass density distribution ρ. In the Boltzmann equation we may be looking at the distribution function of a sub-component, for which the mass density then is denoted by ν.

In a self-gravitating system of course ρ and ν are the same.

Hydrodynamic equations

In practice we never observe full distribution functions, but only the first three moments of it in the form of density, systematic motion and amount of random motion of velocity dispersion.

The hydrodynamic, moment or Jeans equations are obtained from the collissionless Boltzmann equation by multiplication by a velocity to some power followed by integration over all velocities (as in calculating moments for a distribution).

The Boltzman equation was

$$u\frac{\partial f}{\partial x} + v\frac{\partial f}{\partial y} + w\frac{\partial f}{\partial z} - \frac{\partial \Phi}{\partial x}\frac{\partial f}{\partial u} - \frac{\partial \Phi}{\partial y}\frac{\partial f}{\partial v} - \frac{\partial \Phi}{\partial z}\frac{\partial f}{\partial w} = 0.$$

First we change to the often used notation to write this as

$$\frac{\partial f}{\partial t} + v_i\frac{\partial f(\vec{x},\vec{v})}{\partial x_i} - \frac{\partial \Phi}{\partial x_i}\frac{\partial f(\vec{x},\vec{v})}{\partial v_i} = 0$$

Implicit is that we sum over all the values for $i = 1, 2, 3$.

Next we take the convention $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dv_1 dv_2 dv_3 \equiv \int d^3v$

Then the zeroth, first and second order moments in velocity become

$$\int f \; d^3v = \nu$$

$$\frac{1}{\nu} \int v_i f \; d^3v = \langle v_i \rangle$$

$$\frac{1}{\nu} \int v_i v_j f \; d^3v = \langle v_i v_j \rangle$$

From now on I write $f = f(\vec{x}, \vec{v})$.

Zeroth order moment of the Boltzmann equation

$$\int \frac{\partial f}{\partial t} d^3v + \int v_i \frac{\partial f}{\partial x_i} d^3v - \frac{\partial \Phi}{\partial x_i} \int \frac{\partial f}{\partial v_i} d^3v = 0$$

This can be rewritten as

$$\frac{\partial \nu}{\partial t} + \frac{\partial}{\partial x_i} \int v_i f \; d^3v - \frac{\partial \Phi}{\partial x_i} \int f(v_i)]_{-\infty}^{\infty} \, d^2v_{\neq i} = 0$$

Then

$$f(v_i)]_{-\infty}^{\infty} = 0 \quad \Longrightarrow \quad \frac{\partial \nu}{\partial t} + \frac{\partial}{\partial x_i}(\nu \langle v_i \rangle) = 0$$

First order moment of the Boltzmann equation

$$\int v_j \frac{\partial f}{\partial t} d^3v + \int v_i v_j \frac{\partial f}{\partial x_i} d^3v - \frac{\partial \Phi}{\partial x_i} \int v_j \frac{\partial f}{\partial v_i} d^3v = 0$$

Now

$$\int v_j \frac{\partial f}{\partial v_i} d^3v = \int f(v_i)]_{-\infty}^{\infty} d^2 v_{\neq i} - \int \left(\frac{\partial v_j}{\partial v_i} \right) f \; d^3v = 0 - \delta_{ij}\nu$$

so

$$\frac{\partial}{\partial t}\left(\nu \langle v_j \rangle\right) + \frac{\partial}{\partial x_i}\left(\nu \langle v_i v_j \rangle\right) + \nu \frac{\partial \Phi}{\partial x_j} = 0$$

Second order moment of the Boltzmann equation

Similarly we find

$$\nu \frac{\partial \langle v_j \rangle}{\partial t} - \langle v_i \rangle \frac{\partial (\nu v_j)}{\partial x_i} + \frac{\partial (\nu \langle v_i v_j \rangle)}{\partial x_i} + \nu \frac{\partial \Phi}{\partial x_j} = 0$$

This equation is often rewritten using the velocity dispersion tensor:

$$\sigma_{ij}^2 = \langle (v_i - \langle v_i \rangle) \times (v_j - \langle v_j \rangle) \rangle = \langle v_i v_j \rangle - \langle v_i \rangle \langle v_j \rangle = \overline{v_i v_j} - \overline{v_i}.\overline{v_j}$$

Then

$$\frac{\partial (\nu \sigma_{ij}^2)}{\partial x_i} = \frac{\partial (\nu \langle v_i v_j \rangle)}{\partial x_i} - \langle v_j \rangle \frac{\partial (\nu \langle v_i \rangle)}{\partial x_i} - \nu \langle v_i \rangle \frac{\partial \langle v_j \rangle}{\partial x_i}$$

So we can write the second order Bolzmann equation as

$$\nu \frac{\partial \langle v_j \rangle}{\partial t} + \nu \langle v_i \rangle \frac{\partial \langle v_j \rangle}{\partial x_i} + \frac{\partial (\nu \sigma_{ij}^2)}{\partial x_i} + \nu \frac{\partial \Phi}{\partial x_j} = 0$$

So we see that the zeroth, first and second order Boltzmann equations describe relations between the density distribution of a component ν, the mean motions $\langle v_i \rangle$ and the random motions $\langle v_i v_j \rangle$ or σ_{ij} with the potential Φ.

Densities, mean and random motions are *in principle* observables.

Jeans equations

These moment equations are also called Jeans equations and are usually applied in equilibrium when f is not a function of time.

In the practical case the velocity dispersion tensor is assumed to have a diagonal form, i.e. there is a velocity ellipsoid with semi-major axes $\sigma_{11}, \sigma_{22}, \sigma_{33}$ and all cross-terms equal to zero.

In general the Jeans equation cannot be solved without additional assumptions.

And in practice we measure only surface density distributions projected onto the plane of the sky and velocities and velocity dispersion projected onto the line-of-sight.

In the axi-symmetric case the Jeans equations are derived in the same way.

For the radial direction we find:

$$\frac{\partial}{\partial R}(\nu\langle V_{\rm R}^2\rangle)+\frac{\nu}{R}\{\langle V_{\rm R}^2\rangle - V_{\rm t}^2 - \langle (V_\theta - V_{\rm t})^2\rangle\} + \frac{\partial}{\partial z}(\nu\langle V_{\rm R} V_{\rm z}\rangle) = \nu K_{\rm R}$$

By assumption we have taken here $V_{\rm t} = \langle V_\theta\rangle$ and $\langle V_{\rm R}\rangle = \langle V_{\rm z}\rangle = 0$.

This can be rewritten as:

$$-K_{\rm R} = \frac{V_{\rm t}^2}{R} - \langle V_{\rm R}^2\rangle \left[\frac{\partial}{\partial R}(\ln \nu\langle V_{\rm R}^2\rangle) + \frac{1}{R}\left\{1 - \frac{\langle (V_\theta - V_{\rm t})^2\rangle}{\langle V_{\rm R}^2\rangle}\right\}\right] + \langle V_{\rm R} V_{\rm z}\rangle \frac{\partial}{\partial z}(\ln \nu\langle V_{\rm R} V_{\rm z}\rangle)$$

The last term reduces in the symmetry plane to

$$\langle V_{\rm R} V_{\rm z}\rangle \frac{\partial}{\partial z}(\ln \nu\langle V_{\rm R} V_{\rm z}\rangle) = \frac{\partial}{\partial z}\langle V_{\rm R} V_{\rm z}\rangle$$

and may then be assumed zero.

For the azimuthal direction the moment equation is seldom used, because it only contains cross-terms of the velocity tensor. It reads

$$\frac{2\nu}{R}\langle V_{\rm R} V_{\theta}\rangle + \frac{\partial}{\partial R}(\nu\langle v_{\rm R} V_{\theta}\rangle) + \frac{\partial}{\partial z}(\nu\langle V_{\theta} V_{\rm z}\rangle) = 0$$

In the vertical direction the moment equation becomes

$$\frac{\partial}{\partial z}(\nu\langle V_{\rm z}^2\rangle) + \frac{\nu\langle V_{\rm R} V_{\rm z}\rangle}{R} + \frac{\partial}{\partial R}(\nu\langle V_{\rm R} V_{\rm z}\rangle) = \nu K_{\rm z}$$

For spherical symmetry we have velocities $V_{\rm R}$, V_θ and V_ϕ

$$\frac{\partial}{\partial R}(\nu\langle V_{\rm R}^2\rangle) + \frac{\nu}{R}\{2\langle V_{\rm R}^2\rangle - V_{\rm t}^2 - \langle (V_\theta - V_{\rm t})^2\rangle - \langle V_\phi^2\rangle\} = \nu K_{\rm R}$$

In plane-parallel layers the Jeans equation reduces to

$$\frac{d}{dz}\left\{\nu\langle V_{\rm z}^2\rangle\right\} = \nu K_{\rm z}$$

STRUCTURE AND DYNAMICS OF GALAXIES

5. Galactic dynamics: Virial equations, integrals of motion

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/~vdkruit

Beijing, September 2011

Virial equations

The virial equations are derived from the first-order moment Jeans equation for a self-gravitating system (so $\nu = \rho$) by taking its first order moment over spatial coordinates.

$$\frac{\partial}{\partial t}\left(\rho \bar{v}_j\right) + \frac{\partial}{\partial x_i}\left(\rho \overline{v_i v_j}\right) + \rho \frac{\partial \Phi}{\partial x_j} = 0$$

So we get

$$\int x_k \frac{\partial\left(\rho \bar{v}_j\right)}{\partial t} d^3x = -\int x_k \frac{\partial}{\partial x_i}\left(\rho \overline{v_i v_j}\right) d^3x - \int x_k \rho \frac{\partial \Phi}{\partial x_j} d^3x$$

Moment of inertia tensor

Look at the term on the left

$$\int x_k \frac{\partial \left(\rho \bar{v}_j\right)}{\partial t} d^3x$$

and define the moment of inertia tensor

$$I_{jk} = \int \rho x_j x_k d^3x$$

Take the first derivative of this tensor.

$$\frac{d}{dt} I_{jk} = \int \frac{\partial \rho}{\partial t} x_j x_k d^3 x$$

Now recall the zeroth-order moment Jeans equation:

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i} (\rho \langle v_i \rangle) = 0$$

Then we can write

$$\frac{d}{dt} I_{jk} = - \int \frac{\partial \left(\rho \bar{v}_i \right)}{\partial x_i} x_j x_k d^3 x$$

This reduces to

$$\frac{d}{dt} I_{jk} = \int \rho \bar{v}_i \left(\delta_{ij} x_k + \delta_{ik} x_j \right) d^3x = \int \rho \left(\bar{v}_j x_k + \bar{v}_k x_j \right) d^3x$$

and

$$\frac{d^2}{dt^2} I_{jk} = \int \left[x_k \frac{\partial}{\partial t} \left(\rho \bar{v}_j \right) + x_j \frac{\partial}{\partial t} \left(\rho \bar{x}_k \right) \right] d^3x$$

The moment of inertia tensor should be symmetric with respect to the coordinates, so

$$\frac{d^2}{dt^2} \left(\tfrac{1}{2} I_{jk} \right) = \int x_k \frac{\partial}{\partial t} \left(\rho \bar{v}_j \right) d^3x$$

Kinetic energy tensor

Now take the first term on the right and use integration by parts.

$$
\begin{aligned}
-\int x_k \frac{\partial}{\partial x_i}\left(\rho \overline{v_i v_j}\right) d^3x &= \int \rho \overline{v_i v_j} \frac{\partial x_k}{\partial x_i} d^3x - \int \frac{\partial}{\partial x_i}\left(x_k \rho \overline{v_i v_j}\right) d^3x \\
&= \int \delta_{ik} \rho \overline{v_i v_j} \frac{\partial x_k}{\partial x_i} d^3x - \int \delta_{ik} \frac{\partial}{\partial x_i}\left(x_k \rho \overline{v_i v_j}\right) d^3x \\
&= \int \rho \overline{v_k v_j} d^3x - 0 = 2K_{kj}
\end{aligned}
$$

where we have defined the kinetic energy tensor.

We can distinguish between the ordered and random motions using

$$\overline{v_k v_j} = \bar{v}_k . \bar{v}_j + \sigma^2_{kj}$$

This gives rize to a motions tensor T_{jk} and a velocity dispersion tensor Π_{jk}

$$\begin{aligned} K_{ij} &= \int \rho \bar{v}_i . \bar{v}_j d^3x + \frac{1}{2} \int \rho \sigma^2_{ij} d^3x \\ & T_{ij} + \frac{1}{2} \Pi_{ij} \end{aligned}$$

Potential energy tensor

Finally the second term on the right. This we define as the potential energy tensor.

$$W_{jk} = - \int x_j \frac{\partial \Phi}{\partial x_k} d^3 x$$

This finally gives

$$\tfrac{1}{2} \frac{d^2}{dt^2} I_{ij} = 2T_{ij} + \Pi_{ij} + W_{ij}$$

The trace of the tensors give the total energies, so the trace of the last equation reduces for the static case to

$$2T + \Pi = 2K = -W$$

Integrals of motion

Recall the collisonless Boltzmanmn equation

$$\frac{\partial f}{\partial t} + u\frac{\partial f}{\partial x} + v\frac{\partial f}{\partial y} + w\frac{\partial f}{\partial z} - \frac{\partial \Phi}{\partial x}\frac{\partial f}{\partial u} - \frac{\partial \Phi}{\partial y}\frac{\partial f}{\partial v} - \frac{\partial \Phi}{\partial z}\frac{\partial f}{\partial w} = 0.$$

Now consider the equations of motion of an individual star:

$$\frac{dx}{dt} = u,\ \frac{dy}{dt} = v,\ \frac{dz}{dt} = w,\ \frac{du}{dt} = -\frac{\partial \Phi}{\partial x},\ \frac{dv}{dt} = -\frac{\partial \Phi}{\partial y},\ \frac{dw}{dt} = -\frac{\partial \Phi}{\partial z}$$

Fill this in and we get

$$\frac{\partial f}{\partial t} + \frac{dx}{dt}\frac{\partial f}{\partial x} + \frac{dy}{dt}\frac{\partial f}{\partial y} + \frac{dz}{dt}\frac{\partial f}{\partial z} + \frac{du}{dt}\frac{\partial f}{\partial u} + \frac{dv}{dt}\frac{\partial f}{\partial v} + \frac{dw}{dt}\frac{\partial f}{\partial w} \equiv \frac{Df}{Dt} = 0.$$

So along the path of any star in phase space the total derivative of the distribution function Df/Dt is zero.

The density in phase space is constant along the path of any star and the flow of stars in phase space is incompressible.

The equations of motion of a star can be rearranged as:

$$dt = \frac{dx}{u} = \frac{dy}{v} = \frac{dz}{w} = \frac{du}{-\partial\Phi/\partial x} = \frac{dv}{-\partial\Phi/\partial y} = \frac{dw}{-\partial\Phi/\partial z}$$

These are 6 independent ordinary differential equations which yield 6 integration constants for each orbit.

These integration constants thus correspond to a set of 6 independent properties with each combination of values related to a particular stellar orbits.

The distribution function f then simply tells which of these orbits are actually populated, so the general solution of the Boltzmann equation can be written as

$$f(x, y, z, u, v, w) = F(I_1, I_2, ..., I_6)$$

The I's are called the integrals of motion.

The question is then to what physical properties (if any!) these integrals of motion correspond.

Summarizing we have:

- Integrals of motion are functions $I_i(\vec{r}, \vec{v}, t)$ that are constant along an orbit (or $DI/Dt = 0$).
- In phase space there are surfaces $I_i(\vec{r}, \vec{v}, t) = \text{constant}$ and the orbit is the intersection of these surfaces.
- There cannot be more than 6 integrals of motion.

Isolating integrals of motion

We see that the distribution function depends *only* on the integrals of motion. So what are these?

One can be identified as the energy, which is always conserved along an orbit:

$$I_1 = E = \frac{1}{2}(u^2 + v^2 + w^2) + \Phi(x, y, z) = \mathrm{constant}$$

This is called an *isolating* integral of motion, because for particular values it isolates hyper-surfaces in phase space.

The others in general are non-isolating and are only implicit in the numerical integration of an orbit.

In an axisymmetric potential there is a second isolating integral: the angular momentum in the direction of the symmetry axis z is also conserved along an orbit.

$$I_2 = J = RV_\theta$$

Then we have

$$f(R, z, V_{\mathrm{R}}, V_\theta, V_{\mathrm{z}}) = F(E, J)$$

Actually, in a spherically symmetric potential all three components of the angular momentum are isolating intergals.

In the case of the Galaxy near the plane (at small z) the potential is separable and the R- and z-motions will then be decoupled

$$\Phi(R,z) = \Phi_1(R) + \Phi_2(z)$$

Then the decoupled z-energy is a third integral of motion:

$$I_3 = \frac{1}{2} V_z^2 + \Phi_2(z)$$

I will have much more to say later about the so-called third integral problem, which is related to this.

In general any symmetry in the potential or any coordinate system in which the potential can be separated gives rise to integrals of motion.

The integrals that I mentioned for these specific cases restrict the orbit of a star to certain regions of 6-dimensional phase space.

That is why they are called isolating integrals of motion.

But not all integrals of motion have this property and they are called non-isolating integrals and are not of much use.

The concept isolating versus non-isolating will be illustrated next with a simple example.

Non-isolating integrals of motion

Consider the two-dimensional harmonic oscillator with different periods. The equations of motion are

$$x = X \sin \alpha(t - t_x) \quad ; \quad y = Y \sin \beta(t - t_y)$$

Obviously when α/β is rational the orbit is periodic and has a single path.

What are the integrals of motion? First realise that

$$\frac{dx}{dt} = X\alpha \cos \alpha(t - t_x)$$

From x and dx/dt we can form a time-independent parameter:

$$I_1 = \left(\frac{dx}{dt}\right)^2 + \alpha^2 x^2 = X^2(\alpha^2+1) = \mathrm{constant}$$

This then is an integral of motion and confines x to the interval $(-X < x < X)$.

Similarly we have

$$I_2 = \left(\frac{dy}{dt}\right)^2 + \beta^2 y^2 = Y^2(\beta^2+1)$$

Together these integrals then confine the orbit to the area $(-X < x < X, -Y < y < Y)$.

There is a third time-independent quantity that we can derive as follows.

Eliminate t from the two equations of motion; then we get

$$I_3 = \frac{1}{\alpha} \arcsin\left(\frac{x}{X}\right) + \frac{1}{\beta} \arcsin\left(\frac{y}{Y}\right) = t_x - t_y$$

This can be re-arranged as

$$x = X \sin\left[\alpha I_3 - \frac{\alpha}{\beta} \arcsin\left(\frac{y}{Y}\right)\right]$$

Now $\arcsin(y/Y)$ repeats every interval 2π and therefore the second term repeats every interval $2\pi\alpha/\beta$.

If α/β is rational we then get for any value of y a finite number of values for x between $-X$ and X and therefore the orbit is periodic. Then I_3 can also assume a finite number of values and therefore is an isolating integral of motion.

But if α/β is irrational, the second term can assume an infinity of values and x also is not constrained and I_3 can have an infinite number of values and does not constrain the orbit within the area $(-X < x < X, -Y < y < Y)$.

Then I_3 is a non-isolating integral of motion and of no practicle value.

So we see that:

- The number of isolating integrals of motion depend on *both* the potential and the particular orbit and
- For a particular potential some orbits can have more isolating integrals than others.

A further illustration of non-isolating integrals of motions is phase mixing[1].

Assume that stars move in a potential $\Phi(\vec{r})$ and have closed orbits on $(\vec{r}, \vec{v})$. One integral of motion is the total energy of a star

$$E = \frac{1}{2} v^2 + \phi(\vec{r})$$

The orbital period $T(E)$ depends on E. Take for the starting position $\vec{r_\circ}$.

Then the orbital phase angle ψ of the star at time t is

$$\psi(E, \vec{r}) = \psi(E, \vec{r_\circ}) + 2\pi \frac{t}{T(E)}$$

[1] K.C. Freeman, Stars & Stellar Sytems IX, 409 (1975)

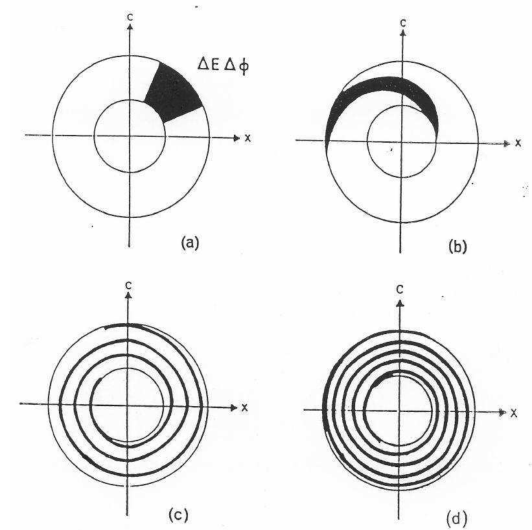
Therefore

$$\psi(E, \vec{r_\circ}) = \psi(E, \vec{r}) - 2\pi \frac{t}{T(E)}$$

is another integral of motion.

So the distribution function can be written as $f(E, \psi - 2\pi t/T)$ and we can follow f in the (E, ψ)-plane.

Say, it intially starts as a distribution limited by values of E and ψ. Then since T is a function of E we find a development as in the following schematic figure.


c
ΔE Δϕ
x
(a)
c
x
(b)
c
x
(c)
c
x
(d)

Although initially confined to a small range $\Delta E \Delta \psi$, the distribution function evolves to a distribution over all phases.

So the distribution function looses its dependence on phase angle and the second integral is non-isolating.

The only isolating integral is the energy.

In general, it may be stated that the non-isolating integrals do define surfaces in phase space, they come close in phase space to any point allowed by the isolating integrals and therefore provide no further constraints on the properties of the orbits.

Jeans' theorem

Jeans' theorem is:
Any arbitrary function of the integrals of motions satisfies the collisionless Boltzmann equation

This is so because the distribution function is constant along the path of an orbit, $Df/Dt = 0$. If f is any function of $I_1 .... I_n$.

$$\frac{Df}{Dt} = \sum_{i=1}^{n} \frac{\partial f}{\partial I_i} \frac{dI_i}{\partial t} = 0$$

However, in order to make a self-consistent system as a solution that resembles a real galaxy, we also need to satisfy the Poisson equation. This is referred to as the self-consistency problem.

Now, the integral of $f(I_i)$ over all integrals I_i at any position is the local density and this must be single valued.

But in general we only know the (single valued) isolating integrals.

Lynden-Bell[2] inferred from this that the distribution function can be completely defined by the isolating integrals only.

E.g. in a system that is spherical in all its properties (so it must depend on the magnitude of the angular momentum, but not its direction) the distribution function is $f = f(E, L^2)$.

Lynden-Bell[3] showed that it is possible for rotating systems to be spherical, while intuitively one expects it to be always oblate.

[2] D. Lynden-Bell, MNRAS 123, 1 (1962)
[3] D. Lynden-Bell, MNRAS 120, 240 (1960)

STRUCTURE AND DYNAMICS OF GALAXIES

6. Galactic dynamics: Timescales

Piet van der Kruit
Kapteyn Astronomical Institute, University of Groningen, the Netherlands
www.astro.rug.nl/∼vdkruit

Beijing, September 2011

Range of timescales

There are a few timescales that are important.

- Crossing time, which is simply the radius divided by the velocity R/V.
  For a galaxy we take some charactristic radius and typical velocity.
  Note that for a uniform sphere with mass M and radius R we have for the typical velocity the circular speed and then

$$V = \sqrt{\frac{GM}{R}} \quad \rho = \frac{3M}{4\pi R^3} \quad t_{\rm cross} = \sqrt{\frac{3}{4\pi G\rho}}$$

- For a galaxy the crossing time is of the order of 10^8 years.
- Hubble time, which is an estimate of the age of the Universe and therefore of galaxies. It is of the order of 10^{10} years.
- The fact that the crossing time is much less than the Hubble time the suggests that we may take the system in dynamical equilibrium.
- Two-body relaxation. This is important for two reasons:
  - Collisions between stars are extremely rare, so collisional pressure is unimportant (contrary to a gas), and
  - Two-body encounters are able to virialize a galaxy so that the kinetic energy of the stars acts as a pressure to stabilize the system, balancing the potential energy.

Two-body relaxation time

Two-body encounters provide processes for a galaxy to come into equilibrium and "virialize", which means that the stellar velocity distribution randomizes.

We will now estimate this relaxation time.

Suppose that we have a cluster of radius R and mass M, made up of N stars with mass m, moving with a mean velocity V.

If two stars pass at a distance r, the acceleration is about Gm/r^2.

Say, that it lasts for the period when the stars are less than the distance r from the closest approach and therefore for a time $2r/V$.

The total change in V^2 is then (acceleration times time)

$$\Delta V^2 \sim \left(\frac{2Gm}{rV}\right)^2$$

The largest possible value of r is obviously R.

For the smallest, we may take $r = r_{\min}$, where ΔV^2 is equal to V^2 itself, since then the approximation breaks down. It is not critical, since we will need the logarithm of the ratio $R/r_{\min}$.

So we have

$$r_{\min} = \frac{2Gm}{V^2}$$

The density of stars is $3N/4\pi R^3$ and the surface density $N/\pi R^2$.

The number of stars with impact parameter r is then the surface density times $2\pi r dr$.

After crossing the cluster once the star has encountered all others. We can calculate the total change in V^2 by integrating over all r

$$(\Delta V^2)_{\rm tot} = \int_{r_{\rm min}}^{R} \left(\frac{2Gm}{rV}\right)^2 \frac{2Nr}{R^2} dr = \left(\frac{2Gm}{RV}\right)^2 2N \ln \Lambda$$

where $\Lambda = R/r_{\rm min}$.

The relaxation time is equal to the number of crossing times it takes for $(\Delta V^2)_{\rm tot}$ to be become equal to V^2.

Since a crossing time is of order R/V and since the virial theorem tells us that $V^2 \sim GNm/R$, we find

$$t_{\rm relax} \sim \frac{RN}{8V \ln \Lambda} \sim \left(\frac{R^3 N}{Gm}\right)^{1/2} \frac{1}{8 \ln \Lambda}$$

With the expression above for $r_{\rm min}$ we find

$$\Lambda = \frac{R}{r_{\rm min}} = \frac{RV^2}{2Gm} \sim \frac{GNm}{2GRm} \sim \frac{N}{2} \sim N$$

The final expression for the two-body relaxation time then is

$$t_{\mathrm{relax}} \sim \left(\frac{R^3}{GM}\right)^{1/2} \frac{N}{8 \ln N}$$

This ranges from about 10^9 years for globular clusters to 10^{12} years for clusters of galaxies.

Within galaxies encounters are unimportant and they can be treated as collisionless systems.

Violent relaxation

If galaxies are relaxed systems another mechanism must be at work. This is violent relaxation[1].

This occurs when the potential changes on timescales comparable to the dynamical timescale.

If $E(\vec{v},t) = \frac{1}{2}v^2 + \Phi(\vec{x},t)$ then

$$\begin{aligned}\frac{dE}{dt} &= \frac{dE}{d\vec{v}}\frac{d\vec{v}}{dt} + \frac{d\Phi}{dt} = \vec{v}\frac{d\vec{v}}{dt} + \frac{d\Phi}{dt} \\ &= -\frac{\partial\vec{r}}{\partial t}\frac{\partial\Phi}{\partial\vec{r}} + \frac{\partial\Phi}{\partial t} + \frac{\partial\Phi}{\partial\vec{r}}\frac{d\vec{r}}{dt} \\ &= \frac{\partial\Phi}{\partial t}\end{aligned}$$

[1] D. Lynden-Bell, MNRAS 136,101 (1967)

Thus a star can change its energy in a collisonless system by a time-dependent potential, such as during the collapse of a galaxy.

The timescale associated with violent relaxation is, according to Lynden-Bell

$$t_{\rm vr} \sim \langle \frac{\dot{\Phi}^2}{\Phi^2} \rangle$$

So the timescale of violent relaxation is of the order of that of the change of the potential.

A very important aspect is that the change in a star's energy is independent of its mass, contrary to other relaxaton mechanisms, such as two-body encounters, which give rise to mass segregation.

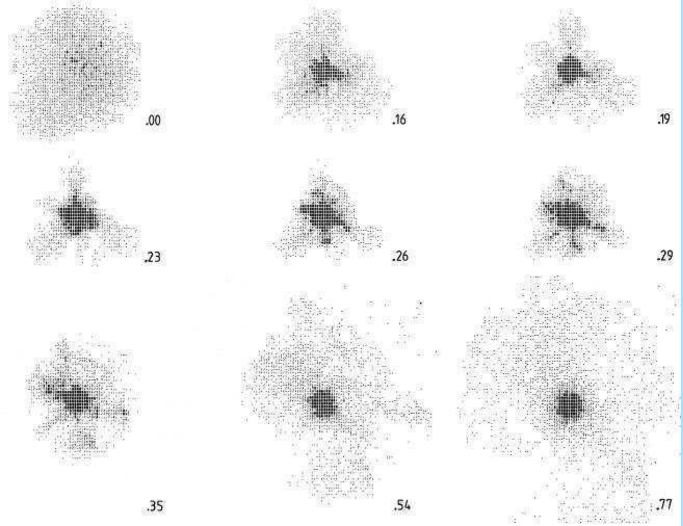
Also some of the information on the initial condition will get lost.

Van Albada[2] was the first to numerically simulate violent relaxation.

He found some remarkable things:

- If the collapse factor was large, irregular initial conditions gave rise to an $R^{1/4}$-law[3] surface density distribution, as observed in elliptical galaxies over a range of up to 12 magnitudes.
- The binding energy of particles before and after collapse correlate, showing that some information on the initial state is not wiped out.

[2] T.S. van Albada, MNRAS 201, 939 (1982)
[3] $\log I(r) = \log I_\circ - 3.33(r/r_e)^{1/4}$.


.00
.16
.19
.23
.26
.29
.35
.54
.77

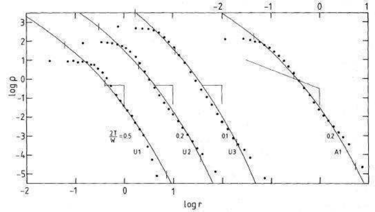
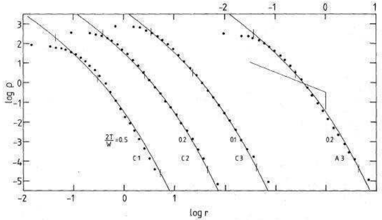
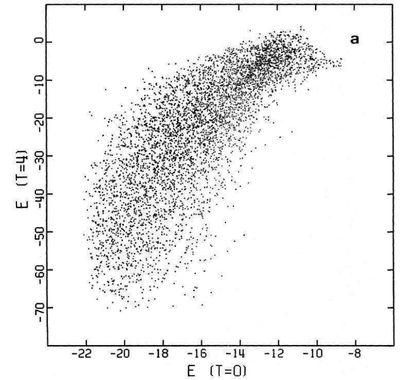
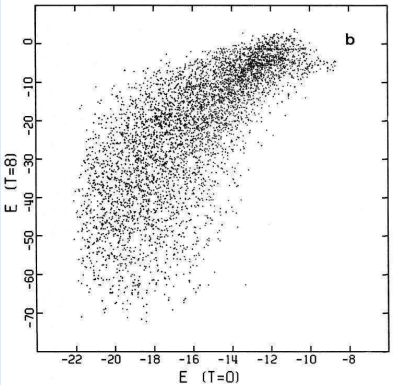


Figure 4. Density distribution in final equilibrium model compared with that for the $r^{1/4}$ law (solid line; Young 1976), for models U and A1. Scaling: equilibrium models and $r^{1/4}$ law model have same half-mass radius r_h and same total mass. Short vertical dashes along $\rho(r)$ for $r^{1/4}$ law model indicate radii containing 10, 50 and 99 per cent of the total mass. Density and radius of starting model are indicated by short straight lines.




a
E (T=4)
E (T=0)
0
-10
-20
-30
-40
-50
-60
-70
-22
-20
-18
-16
-14
-12
-10
-8


b
E (T=8)
E (T=0)
0
-10
-20
-30
-40
-50
-60
-70
-22
-20
-18
-16
-14
-12
-10
-8

Dynamical friction

As a star moves through a background of other stars, the small deflections will give a small overdensity behind the star and consequently induce a drag.

Suppose that a body of mass m moves in a circular orbit with radius R through a background of bodies with mass M at a speed $V_{\rm c}$ and assume that the background is an isothermal sphere[4] with $V_{\rm c}$ the circular speed (and $V_{\rm c}/2$ the velocity dispersion).

[4]An isothermal sphere is a distribution where everywhere the velocity dispersion is constant and isotropic and that is in equilibirium with its own gravity; see later.

Then the loss of angular momentum is about

$$\frac{dJ}{dR} \sim -0.4 \frac{Gm^2}{R} \ln \Lambda$$

where

$$\Lambda = \frac{R_c V_c^2}{G(m+M)}$$

R_c is the core radius of the isothermal sphere (the typical lengthscale of the background density distribution).

The timescale of dynamical friction for the body to spiral into the center is then

$$t_{df} \sim \frac{R^2 V_c}{Gm \ln \Lambda}$$

This timescale is large and only relevant for globular clusters in the inner halo or for galaxies in the central parts of clusters.

STRUCTURE AND DYNAMICS OF GALAXIES

7. Galactic dynamics: Stellar orbits

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/∼vdkruit

Beijing, September 2011

Outline

Orbits in symmetric potentials

Spherical potentials

The equation of motion in a spherical potential is in vector notation

$$\ddot{\mathbf{R}} = -\frac{d\Phi}{d\mathrm{R}}\hat{\mathbf{e}}_{\mathrm{R}}$$

The angular momentum is

$$\mathbf{R} \times \dot{\mathbf{R}} = \mathbf{L}$$

This is constant and the orbit therefore is in a plane.

We then use polar coordinates in this plane these two equations become

$$\ddot{R} - R\dot{\theta}^2 = -\frac{d\Phi}{dR}$$

$$R^2\dot{\theta} = L$$

Integrating this we get

$$\tfrac{1}{2}\dot{R}^2 + \frac{1}{2}\frac{L^2}{R^2} + \Phi(R) = E$$

The energy E is constant.

If $E < 0$ then the star is bound between radii $R_{\rm max}$ and $R_{\rm min}$, which are the roots of

$$\frac{1}{2}\frac{L^2}{R^2} + \Phi(R) = E$$

The radial period is the interval between the times the star is at $R_{\rm min}$ and $R_{\rm max}$ and back.

$$\begin{aligned} T_{\rm R} &= 2\int_{R_{\rm min}}^{R_{\rm max}} dt = 2\int_{R_{\rm min}}^{R_{\rm max}} \frac{dR}{\dot{R}} = \\ & 2\int_{R_{\rm min}}^{R_{\rm max}} \frac{dR}{\{2[E-\Phi(R)] - L^2/R^2\}^{1/2}} \end{aligned}$$

In the azimuthal direction the angle θ changes in the time $T_{\rm R}$ by

$$\Delta\theta = \int_0^{T_{\rm R}} \frac{d\theta}{dR} dR = 2 \int_0^{T_{\rm R}} \left(\frac{L}{R^2}\right) \frac{dR}{\dot{R}}$$

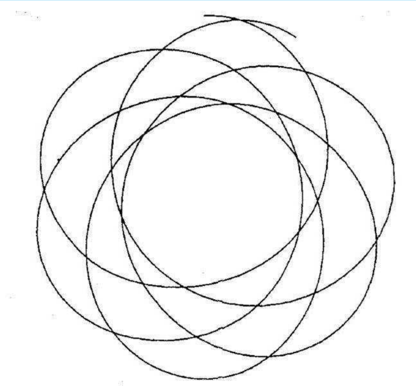
This can be evaluated further in terms of $T_{\rm R}$, which depends upon the particular potential.

The orbit is closed if

$$\Delta\theta = 2\pi \frac{m}{n}$$

with m and n integers.

This is not generally true and the orbit then has the form of a rosette and can the star visit every point within $(R_{\rm min}, R_{\rm max})$.



Even in the simple case of a spherical potential, the equation of motion of the orbit must be integrated numerically.

The Rosette orbit can be closed by observing it from a rotating frame (see below under resonances), when it is rotating at an angular velocity of

$$\Omega_\mathrm{p} = \frac{(\Delta\theta - 2\pi)}{T_\mathrm{R}}$$

We will treat two special cases which can be solved analytically.

The harmonic oscillator

This concerns the potential of a uniform sphere

$$\Phi = \tfrac{1}{2}\Omega^2 R^2.$$

Then we take cartesian coordinates $x = r\cos\theta$, $y = r\sin\theta$ and then

$$\frac{d^2x}{dt^2} = -\Omega^2 x \quad ; \quad \frac{d^2y}{dt^2} = -\Omega^2 y$$

Then

$$x = X\cos(\Omega t + a_{\mathrm{x},\circ}) \quad ; \quad y = Y\cos(\Omega t + a_{\mathrm{y},\circ})$$

The orbits are closed ellipses centered on the origin and $\Delta\theta$ is equal to π in T_{R}.

The Keplerian potential

The potential now is that of a point source in the center and this is the well-known two-body problem[1]:

$$\Phi = -\frac{GM}{R}$$

The orbits are closed ellipses with one focus at the origin:

$$R = \frac{a(1-e^2)}{\{1+\cos(\theta-\theta_\circ)\}}$$

[1]There is a complete derivation of the two-body problem available at (http://www.astro.rug.nl/∼vdkruit/jea3/homepage/two-body.pdf).

Here semi-major axis a and excentricity e are related to E and L by

$$a = \frac{L^2}{GM(1-e^2)} \quad ; \quad E = -\frac{GM}{2a}$$

$$R_{\mathrm{max}}, R_{\mathrm{min}} = a(1 \pm e)$$

$$T_{\mathrm{R}} = T_{\theta} = 2\pi\sqrt{\frac{a^3}{GM}} = T_{\mathrm{R}}(E)$$

Now $\Delta\theta = 2\pi$ in T_{R}.

Galaxies have mass distributions somewhere between these two extremes, so we may expect that $\Delta\theta$ is in the range π to 2π in T_{R}.

Axisymmetric potentials

We now have a potential $\Phi = \Phi(R, z)$, that may be applicable to disk galaxies. The equations of motion are

$$\ddot{R} - R\dot{\theta}^2 = -\frac{\partial \Phi}{\partial R}$$

$$\frac{d}{dt}(R^2\dot{\theta}) = 0$$

$$\ddot{z} = \frac{d^2 z}{dt^2} = -\frac{\partial \Phi}{\partial z}$$

Integration of middle one of these equations gives

$$L_{\mathrm{z}} = R^2\dot{\theta}$$

The motion in the meridional plane then can be described by an effective potential

$$\ddot{R} = -\frac{\partial \Phi_{\rm eff}}{\partial R}$$

$$\ddot{z} = -\frac{\partial \Phi_{\rm eff}}{\partial z}$$

where

$$\Phi_{\rm eff} = \Phi(R,z) + \frac{L_{\rm z}^2}{2R^2}$$

The energy of the orbit is

$$E = \tfrac{1}{2}\dot{R}^2 + \tfrac{1}{2}\dot{z}^2 + \Phi_{\rm eff}(R,z)$$

The orbit is trapped inside the appropriate contour $E = \Phi_{\rm eff}$, which is called the zero-velocity curve.

Only orbits with low $L_{\rm z}$ can approach the z-axis.

The minimum in $\Phi_{\rm eff}$ occurs for $\nabla \Phi_{\rm eff} = 0$, or at $z = 0$ and where

$$\frac{\partial \Phi}{\partial R} = \frac{L_{\rm z}^2}{R^3}$$

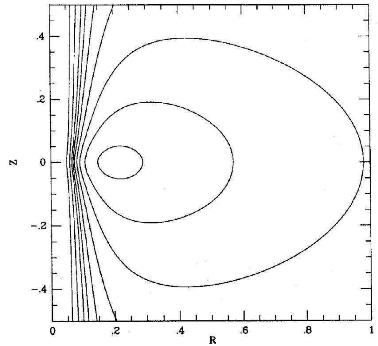
This corresponds to the circular orbit with $L = L_{\rm z}$.

It is the highest angular momentum orbit that is possible for a given E, or in other words, it has all its kinetic energy in θ-motion.

As an example we take the logarithmic potential

$$\Phi(R,z) = \tfrac{1}{2} V_\circ^2 \ln\left(R^2 + \frac{z^2}{q^2}\right)$$

Here are countours of Φ_{eff} for the case $q = 0.5$ and $L_z = 0.2$.

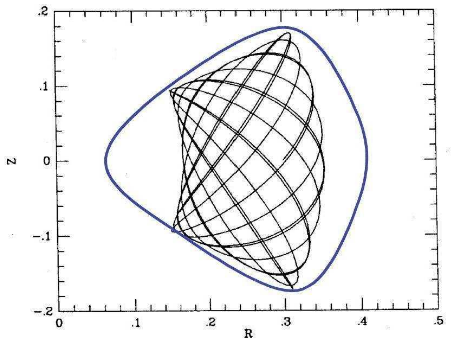


The minimum in Φ_{eff} occurs where $\nabla\Phi_{\mathrm{eff}} = 0$ that is in the plane ($z = 0$).

If E and L_z were the only two isolating integrals the orbits would be able to visit all points within their zero-velocity curves.
In simulations this is often not the case and there must be a third integral.

Here is the case of actual simulated orbits in a slightly flattened logarithmic potential. We show the motion in the meridional plane, rotating along with the angular momentum of the orbit.

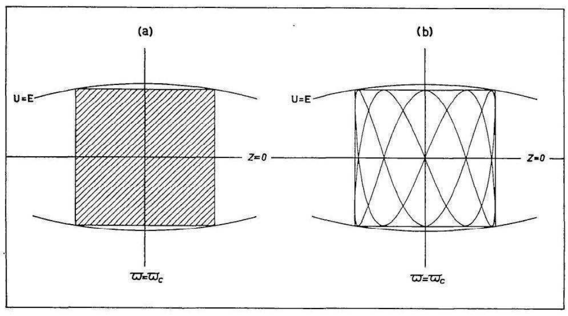
The blue line is the zero-velocity curve corresponding to this orbit.



Third integral

Recall that for small deviations from the symmetry plane the energy in the z-direction was a third isolating integral.

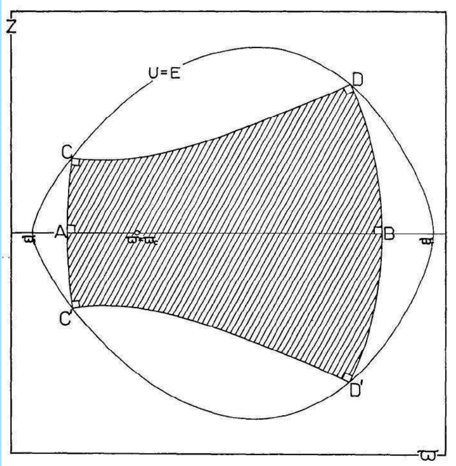
Here are two diagrams from an early study by Ollongren[2]. We have either periodic or non-periodic orbits.



[2] A. Ollongren, B.A.N. 16, 241 (1962)

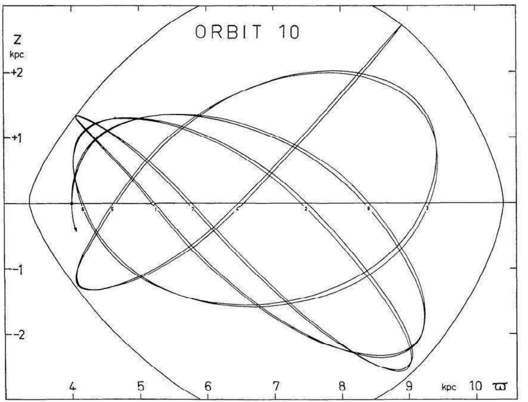
Ollongren did numerical integrations using the potential of a recent model of the mass distribution in the Galaxy by Schmidt[a].

He found that there was a distortion of the box that was covered by the orbit.



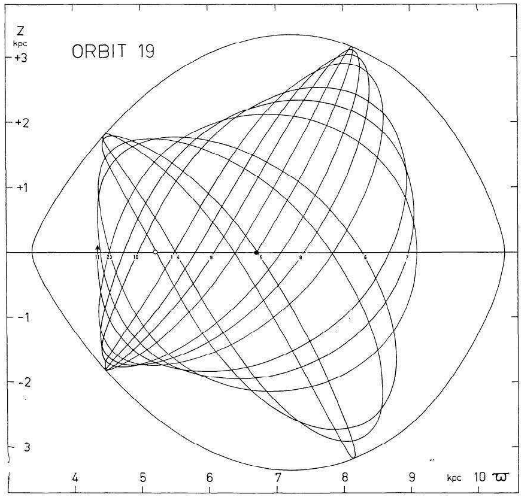
[a] M. Schmidt, B.A.N. 13, 15 (1956)

He also found that the most general separable case was in elliptical coordinates, in which a third integral is quadratic in the velocities[3].

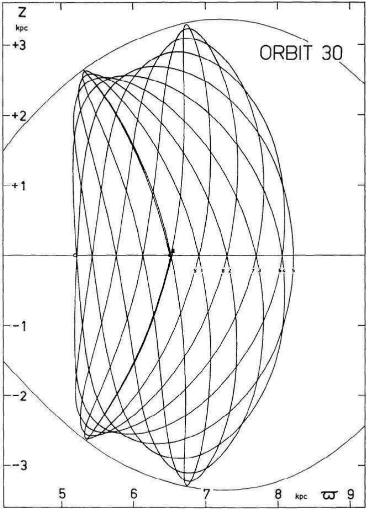


Orbit 10, *low*, *special*, frequency ratio near 14 : 17. In revolution 5 the fundamental point D is approached closely, after which the previous path is nearly retraced back to the starting point ($\varpi = \varpi_A$, $z = 0$).

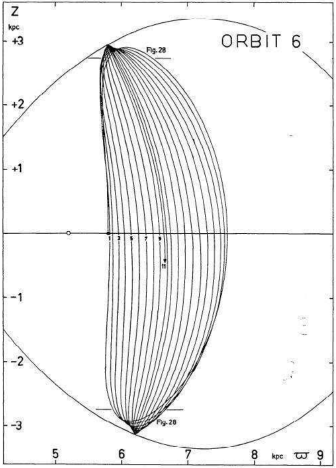
[3]See also H.C. van de Hulst, B.A.N. 16, 235 (1962)



Orbit 19, *intermediate*. In revolution 11 the fundamental point A is approached closely. Motion of orbital plane is shown in Figure 33.



Orbit 30, *intermediate*, close to orbit 3. The trajectory departs only slowly from the trajectory of orbit 3.



Orbit 6, *high*, *special*. Sense of rotation of all revolutions drawn is clockwise, except for revolution 11, which is of the switching type. Region within the boundary for $|z| > 2.75$ enlarged in Figure 28.

Summarizing:

- If E and L_z are the only two isolating integrals, the orbit would visit all points within the zero-velocity curves.
- In practice it was found that there are limiting surfaces that seem to forbid the orbit to fill the whole volume within the zero-velocity curves.
- This behaviour is very common for orbits in axisymmetric potentials, when the combination (E, L_z) is not too far from that of a circular orbit. A third integral is present, although in general its form cannot be explicitly written down.

Surface of section

For each orbit the energy $E(R, z, \dot{R}, \dot{z})$ is an integral, so only three of the four coordinates can be independent, say R, z and $\dot{R}$.

The orbit can visit every point in $(R, z, \dot{R})$-space as far as allowed by E.

Now take a slice through $(R, z, \dot{R})$-space, e.g. at $z = 0$. This is called a *surface of section*.

The orbits' successive crossings of $z = 0$ generate a set of points inside the region $E = \frac{1}{2}\dot{R}^2 + \Phi_{\rm eff}(R, 0)$.

Hénon & Heiles[4] did a famous study of third integrals and surfaces of section. They used a convenient analytical potential in coordinates (x, y):

$$\Psi(x, y) = \frac{1}{2}(x^2 + y^2 + 2x^2 y - \frac{2}{3}y^3)$$

The figure shows consecutive crossings of the surface of section $(y, \dot{y})$.

After an infinite time the full curve will be filled.

This is a signature of a third isolating integral; the orbit is constrained inside the zero-velocity curve.

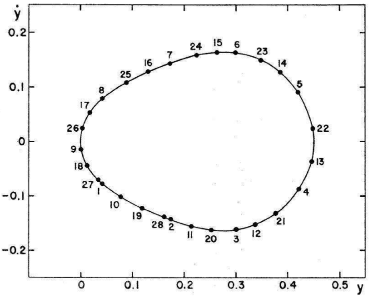
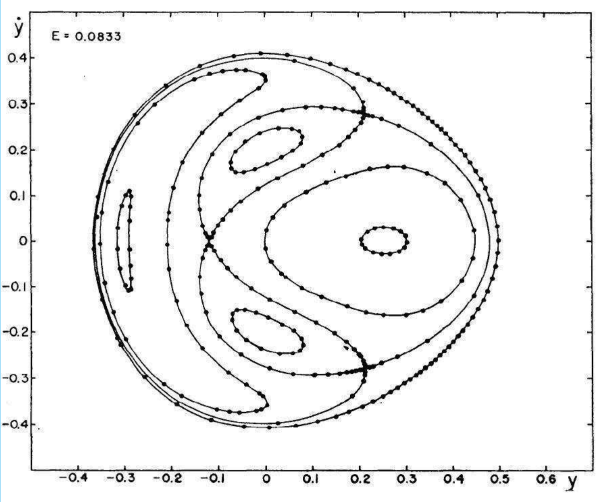


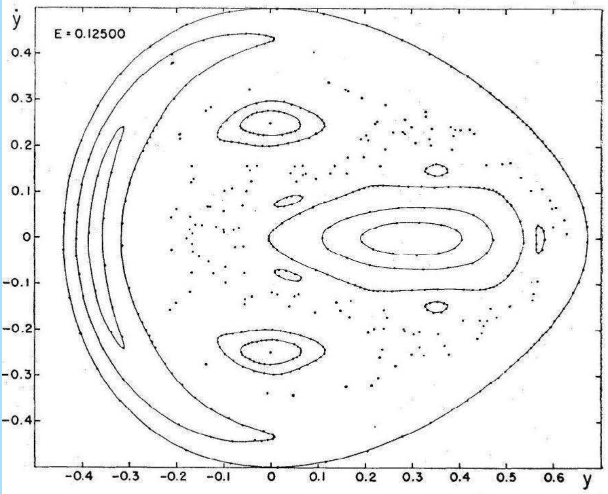
FIG. 3. A typical set of points P_i; $E=0.08333$.

[4] M. Hénon & C. Heiles, A.J. 69, 73 (1964)

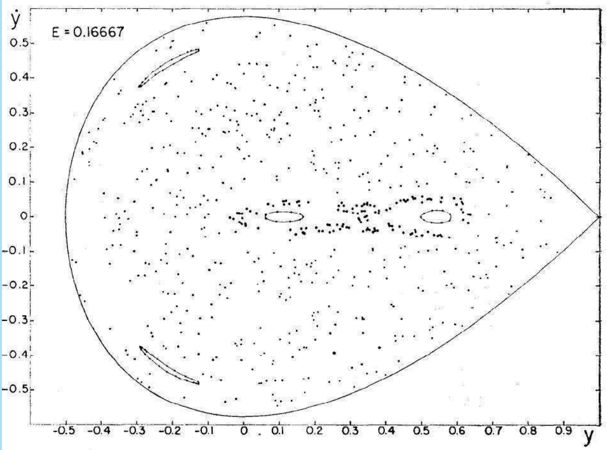
Here are some orbits for $E = 0.08333$. All have a third integral.



Here are orbits for $E = 0.125$. Now some orbits have no third integral.

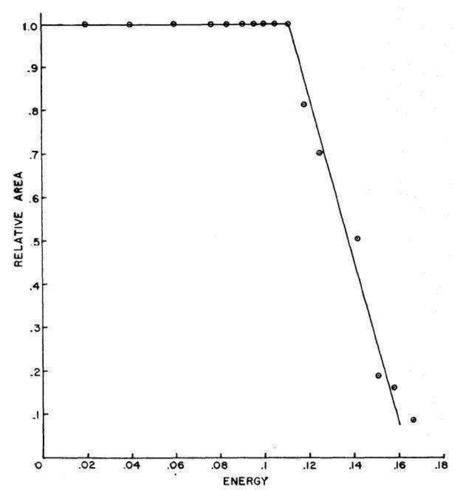


For $E = 0.16667$ almost no orbits have a third integral.



Hénon & Heiles divised a method to derive the fraction of orbits that have a thrid integral for each energy.

For $E < 0.11$ all orbits have a third integral, but for $E > 0.17$ almost none do.



If there is no other integral then these points fill the whole region.

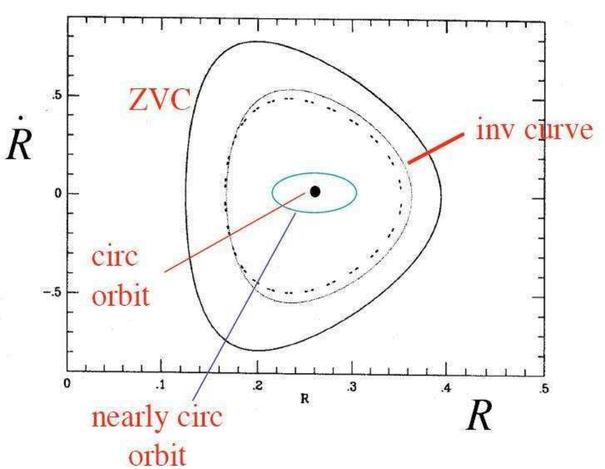
If there is another integral, then its surface $I_{\mathrm{R}}(R, z, \dot{R})$ cuts the plane in a *curve* $I_{\mathrm{R}}(R, 0, \dot{R}) = \mathrm{constant}$.

A *periodic* orbit is a *point* or a set of points on the $(R, \dot{R})$ surface of section.

Such curves and points are called *invariant*, because they are invariant under the mapping of the surface of section onto itself generated by the orbit.

Invariant points often have closed invariant curves around them on the surface of section. These represent *stable* periodic orbits. Ones where invariant curves cross are *unstable* periodic orbits.

This diagram (taken from Ken Freeman) summarizes the points.



Rotating non-axisymmetric potentials

In cases of bars or some elliptical galaxies we may consider a potential that rotates with a rigid angular velocity Ω.

Then the equation of motion is

$$\ddot{\mathbf{r}} = -\nabla\Psi - 2(\boldsymbol{\Omega} \times \mathbf{r}) - \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r})$$

The second term on the right is the Coriolis force and the third one the centrifugal force.

Then we can define an effective potential, so that

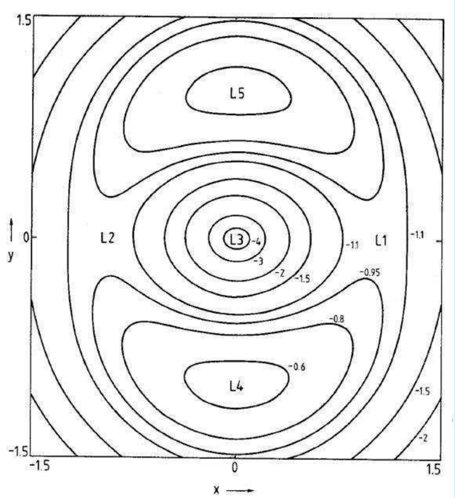
$$\ddot{\mathbf{r}} = -\nabla\Psi_{\mathrm{eff}} - 2(\boldsymbol{\Omega} \times \mathbf{r})$$

Such a potential has equipotential curves in the $z = 0$ plane that show neutral points.

L_1 and L_2 are saddle points and are unstable.

L_3 is a minimum and is stable.

L_4 and L_5 are maxima that can either be stable or unstable.

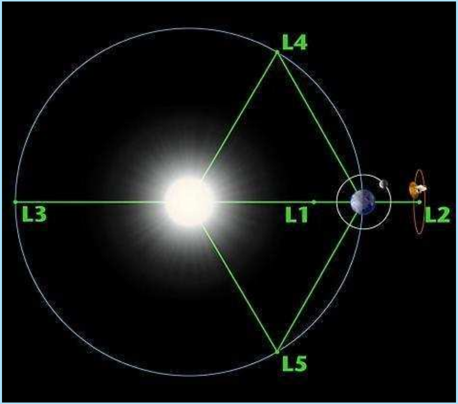


These point should in spite of their notation not be confused with Lagrange points in the restricted three-body problem, although there is some similarity.

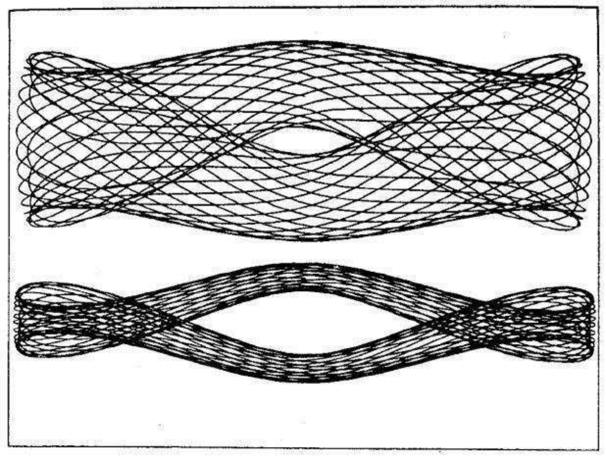
There are two bodies (here Sun and Earth) in circular orbits.

The Lagrange points L_1, L_2 and L_3 are saddle points and unstable.

L_4 and L_5 are stable.



Stars describe orbits that reinforce the bar potential.



STRUCTURE AND DYNAMICS OF GALAXIES

8. Galactic dynamics: Epicycle orbits, instabilities

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/~vdkruit

Beijing, September 2011

Outline

Epicycle orbits

- Epicycle theory
- Vertical motion
- Resonances

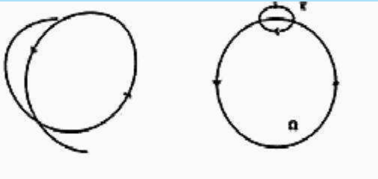
Instabilities

- Jeans instability
- Toomre criterion for local stability
- Goldreich–Lynden-Bell criterion
- Global stability
- Tidal radius

Epicycle orbits

Epicycle theory

For small deviation from the circular rotation, the orbits of stars can be described as epicyclic orbits.



If $R_\circ$ is a fudicial distance from the center and if the deviation $R - R_\circ$ is small compared to $R_\circ$, then we have in the radial direction

$$\frac{d^2}{dt^2}(R-R_\circ) = \frac{V^2(R)}{R} - \frac{V_\circ^2}{R_\circ} = 4B(A-B)(R-R_\circ) = -\kappa^2(R-R_\circ),$$

where the last approximation results from making a Taylor expansion of $V(R)$ at $R_\circ$ and ignoring higher order terms.

This equation is of the form $\ddot{x} = -\kappa^2 x$ and is easily integrated

$$R - R_\circ = \frac{V_{\mathrm{R},\circ}}{\kappa} \sin \kappa t,$$

In the tangential direction we have

$$\frac{d\theta}{dt} = \frac{V(R)}{R} - \frac{V_\circ}{R_\circ} = -2\frac{A-B}{R_\circ}(R - R_\circ),$$

where θ is the angular tangential deviation seen from the Galactic center. Then

$$\theta R_\circ = -\frac{V_{\mathrm{R},\circ}}{2B} \cos \kappa t$$

The orbital velocities are

$$V_{\mathrm{R}} = V_{\mathrm{R},\circ} \cos \kappa t,$$

$$V_{\theta} - V_{\theta,\circ} = \frac{V_{\mathrm{R},\circ}\kappa}{-2B} \sin \kappa t.$$

The period in the epicycle equals $2\pi/\kappa$ and κ is the epicyclic frequency

$$\kappa = 2\{-B(A-B)\}^{1/2}.$$

In the solar neighborhood κ $\sim$36 km s^{-1} kpc^{-1}.

For a flat rotation curve we have

$$\kappa = \sqrt{2}\frac{V_\circ(R)}{R}.$$

Through the Oort constants and the epicyclic frequency, the parameters of the epicycle depend on the local forcefield, because these are all derived from the rotation velocity and its radial derivative.

The direction of motion in the epicycle is opposite to that of galactic rotation.

The ratio of the velocity dispersions or the axis ratio of the velocity ellipsoid in the plane for the stars can be calculated as

$$\frac{\langle V_{\mathrm{R}}^2\rangle^{1/2}}{\langle V_{\theta}^2\rangle^{1/2}} = \sqrt{\frac{-B}{A-B}}.$$

For a flat rotation curve this equals 0.71.

With this result the hydrodynamic equation can then be reduced to the so-called asymmetric drift equation. Recall

$$-K_{\mathrm{R}} = \frac{V_{\mathrm{t}}^2}{R} - \langle V_{\mathrm{R}}^2\rangle\left[\frac{\partial}{\partial R}(\ln \nu\langle V_{\mathrm{R}}^2\rangle) + \frac{1}{R}\left\{1 - \frac{\langle (V_\theta - V_{\mathrm{t}})^2\rangle}{\langle V_{\mathrm{R}}^2\rangle}\right\}\right] + \langle V_{\mathrm{R}} V_{\mathrm{z}}\rangle \frac{\partial}{\partial z}(\ln \nu\langle V_{\mathrm{R}} V_{\mathrm{z}}\rangle)$$

For the case the cross-dispersion in the last term is zero, we can now write

$$V_{\rm rot}^2 - V_{\rm t}^2 =$$

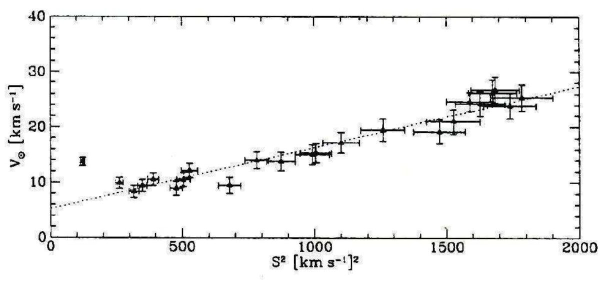
$$-\langle V_{\rm R}^2 \rangle \left\{ R\frac{\partial}{\partial R} \ln \nu + R\frac{\partial}{\partial R} \ln \langle V_{\rm R}^2 \rangle + \left[1 - \frac{B}{B-A} \right] \right\}.$$

Here $V_{\rm rot}$ is the 'circular' velocity that corresponds directly to a centrifigal force $V_{\rm rot}^2/R$ equal to the gravitational force $K_{\rm R}$.

If the asymmetric drift $(V_{\rm rot} - V_{\rm t})$ is small, the left-hand term can be approximated by

$$V_{\rm rot}^2 - V_{\rm t}^2 \sim 2V_{\rm rot}(V_{\rm rot} - V_{\rm t}).$$

The term asymmetric drift comes from the observation that objects in the Galaxy with larger and larger velocity dispersion lag more and more behind in the direction of Galactic rotation.



Vertical motion

For the vertical motion the equivalent approximation is also that of a harmonic oscillator.

For a constant density the hydrodynamic equation reduces to

$$K_{\rm z} = \frac{d^2 z}{dt^2} = -4\pi G \rho_0 z.$$

Integration gives

$$z = \frac{V_{\rm z,\circ}}{\lambda} \sin \lambda t \quad ; \quad V_{\rm z} = V_{\rm z,\circ} \cos \lambda t.$$

The period equals $2\pi/\lambda$ and the vertical frequency λ is

$$\lambda = (4\pi G \rho_0)^{1/2}.$$

For the solar neighbourhood we have ρ_0 ~0.1 $M_\odot$ pc^{-3}.

With the values above for $R_\circ$, $V_\circ$, A and B, the epicyclic period $\kappa^{-1} \sim 1.7 \times 10^8$ yrs and the vertical period $\lambda^{-1} \sim 8 \times 10^7$ yrs.

This should be compared to a period of rotation of 2.4 $\times 10^8$ yrs.

The Sun moves with ~20 km s^{-1} towards the Solar Apex at Galactic longitude $\sim 57^\circ$ and latitude $\sim +27^\circ$.

From the curvature of the ridge of the Milky Way the distance of the Sun from the Galactic Plane is estimated as 12 pc.

The axes of the solar epicycle are about ~0.34 kpc in the radial direction and ~0.48 kpc in the tangential direction.

The amplitude of the vertical motion is ~85 pc.

Resonances

The most important ones are between epicyclic frequency and some other frequency that we will call *pattern speed* $\Omega_{\rm p}$.

The *inner Lindblad resonance* occurs for

$$\Omega_{\rm p} = \Omega_{\rm rot}(R) - \frac{\kappa}{2}$$

where $\Omega_{\rm rot}(R)$ is the angular rotation speed.

This resonance occurs at the radius, where –in a rotating frame with angular velocity $\Omega_{\rm p}$– the particle goes through 2 epicycles in the same time is it goes once around the centre. The resulting orbit in that frame then is closed and has an oval shape.

It goes back to Lindblad's discovery that the property $\Omega_{\rm rot}(R) - \kappa/2$ in the inner Galaxy is roughly constant with R.

The pattern speed may be identified with that of the rotating frame in which the spiral *pattern* (not the spiral arms as physical structures themselves) is stationary or with the body rotation of a bar or oval distortion.

Equivalently we have the *outer Lindblad resonance*

$$\Omega_{\rm p} = \Omega_{\rm rot}(R) + \frac{\kappa}{2}$$

and *co-rotation*

$$\Omega_{\rm p} = \Omega_{\rm rot}(R)$$

Higher order Lindblad resonances (involving κ/n) sometimes also play a role.

Instabilities

Jeans instability

We then start with the Jeans instability in a homogeneous medium.

There are various ways of describing it to within an order of magnitude.

The first is to make use of the virial theorem

$$2\ T_{\rm kin} + \Omega = 0$$

for stability against gravitational contraction.

In a uniform, isothermal sphere the kinetic energy is

$$T_{\rm kin} = 1/2\, M\langle V^2\rangle$$

and the potential energy

$$\Omega = -\frac{3}{5}\frac{GM^2}{R}$$

So the sphere will contract when its mass M is larger than the value required by the virial theorem.

This is called the Jeans mass $M_{\rm Jeans}$, which then comes out as

$$M_{\rm Jeans} = \left(\frac{5}{3G}\right)^{3/2}\left(\frac{3}{4\pi}\right)^{1/2}\left(\frac{\langle V^2\rangle^3}{\rho}\right)^{1/2}$$

A method that gives roughly the same result starts by calculating the free-fall time of a homogeneous sphere.
Anywhere the equation of motion is

$$\frac{d^2 r}{dt^2} = -\frac{G\ M(r)}{r^2} = -\frac{4\pi}{3} G\rho r$$

Solve this and apply for $r = 0$, then

$$t_{\rm ff} = \left(\frac{3\pi}{32 G\rho} \right)^{1/2}$$

The free-fall time is independent of the initial radius and depends only on the density. Now, if there were no gravity a star will move out to the radius of the sphere R in a time

$$t = \frac{R}{\langle V^2 \rangle^{1/2}}$$

For marginal stability the two have to be equal and it follows that the Jeans length is

$$R_{\mathrm{Jeans}} = \left(\frac{3\pi}{32} \frac{\langle V^2 \rangle}{G\rho} \right)^{1/2}$$

Sometimes in the literature the Jeans length is taken as the *diameter* of the sphere.

Toomre criterion for local stability

Next we need to consider Toomre's[1] criterion for local stability:

$$Q = \frac{\langle V_{\mathrm{R}}^2 \rangle^{1/2} \kappa}{3.36 G \sigma}$$

$\langle V_{\mathrm{R}}^2 \rangle^{1/2}$ is the stellar velocity dispersion in the R-direction, σ is the local disk surface density and κ is the epicyclic frequency.

An approximate derivation of Toomre's criterion can be made for an infinitesimally thin disk.

1. At small scales the Jeans instability needs to be considered.

[1] A. Toome, Ap.J. 139, 1217 (1964)

Take an area with radius R and surface density σ. The equation of motion is

$$\frac{d^2R}{dt^2} = -\pi G\sigma$$

Solve this and apply for $R = 0$; this gives the free-fall time

$$t_{\rm ff} = \left(\frac{2R}{\pi G\sigma}\right)^{1/2}$$

A star moves out to radius R in a time

$$t = \frac{R}{\langle V^2\rangle^{1/2}}$$

and this must for marginal stability be equal to the free-fall time.

This then gives the Jeans length

$$R_{\rm Jeans} = \frac{2\langle V^2 \rangle}{\pi G \sigma}$$

2. At large scale we need to consider stability resulting from differential rotation.

Take an area with radius $R_\circ$; the angular velocity from differential rotation is

$$\Omega = B$$

The centrifugal force is then

$$F_{\rm cf} = R_\circ \Omega^2$$

Let it contract to radius R, then the angular velocity becomes

$$\Omega = \frac{R_\circ^2 B}{R^2}$$

and the centrifugal force

$$F_{\rm cf} = R\Omega^2 = \frac{R_\circ^4 B^2}{R^3}$$

If the contraction is dR then

$$\frac{dF_{\rm cf}}{dR} = -\frac{3R_\circ^4 B^2}{R^4}$$

Now look at the gravitational force

$$F_{\rm grav} = -\frac{G\pi R_\circ^2 \sigma}{R^2}$$

This is correct to within a factor 2 for a flat distribution. Then

$$\frac{dF_{\rm grav}}{dR} = \frac{2\pi G R_\circ^2 \sigma}{R^3}$$

At $R = R_\circ$ these two must compensate each other, so

$$R_{\rm crit} = \frac{2\pi G \sigma}{3B^2}$$

and the disk is stable for all $R > R_{\rm crit}$.

3. Toomre's stability criterion then follows by considering that the disk is stable at all scales if the minimum radius for stability by differential rotation is equal to or smaller than the maximum radius for stability by random motions (the Jeans radius).

Thus

$$\langle V^2 \rangle^{1/2}_{\rm crit} = \frac{\pi}{\sqrt{3}} \frac{G\sigma}{B}$$

In practice $B \approx -A$ (for flat rotation curves), so we can write

$$\langle V^2 \rangle^{1/2}_{\rm crit} \sim 2\pi \left(\frac{2}{3} \right)^{1/2} \frac{G\sigma}{\kappa} = 5.13 \frac{G\sigma}{\kappa}$$

Toomre in his precise treatment found a constant of 3.36.

Goldreich–Lynden-Bell criterion

This can be extended to the criterion, that Goldreich and Lynden-Bell[2] derived for stability of gaseous disks of finite thickness against sheared instabilities:

$$\frac{\pi G\overline{\rho}}{4B(B-A)} \lesssim 1$$

This follows from the result for the Toomre criterion above as follows.

From the vertical oscillation above we find that the maximum distance from the plane is

$$z_{\circ} = \frac{V_{\mathrm{z},\circ}}{(4\pi G\rho_{\circ})^{1/2}}$$

[2] R. Goldreich & D. Lynden-Bell, MNRAS 193, 189 (1965)

Equate the critical velocity dispersion in our derivation of the Toomre criterion to $V_{\mathrm{z},\circ}$, then

$$\frac{12}{\pi} G^{-1} z_\circ^2 \rho_\circ \frac{B^2}{\sigma^2}$$

Now take a mean density $\overline{\rho}$ equal to $\sigma / z_\circ$ and to $\frac{1}{2}\rho_0$ and using $(B - A) \approx 2B$, we get

$$\frac{\pi}{3} G \frac{\overline{\rho}}{B(B - A)} \sim 1$$

These sheared instabilities were proposed by Goldreich & Lynden-Bell as a possible mechanism for the formation of spiral structure.

More recently, Toomre[3] has studied the process in stellar disks and finds an instability based on shear due to differential rotation, that he called *swing amplification*. This process is prevented when

$$X = \frac{R\kappa^2}{2\pi m G \sigma} \gtrsim 3$$

where m is the number of arms. For $-B \approx A$ (a flat rotation curve) this can be written as

$$\frac{Q V_{\rm rot}}{\langle V_{\rm R}^2 \rangle^{1/2}} \gtrsim 3.97\ m$$

This is Toomre's local stability citerion if the velocity dispersion is replaced by $0.22\ V_{\rm rot}/m$.

[3] A. Toomre, Normal Galaxies, ed. S.M. Fall & D. Lynden-Bell, 111 (1981)

Global stability

For global stability there is a global condition due to Efstathiou, Lake & Negroponte[4] from numerical experiments, which reads

$$Y = V_{\rm rot} \left(\frac{h}{GM_{\rm disk}} \right)^{1/2} \gtrsim 1.1$$

For a pure exponential disk with surface density $\sigma(R) = \exp(-R/h)$ without any dark halo $Y = 0.59$.

For a flat rotation curve it is then easy to show that the condition implies that within the disk radius of 4 to 5 scalelengths h the mass in the halo should exceed that of the disk by a factor of about 3.5.

[4] G. Efstathiou, G. Lake & J. Negroponte, MNRAS 199, 1069 (1982)

For a flat rotation curve and an exponential disk Y can be rewritten as

$$Y = 0.615 \left\{ \frac{Q R V_{\rm rot}}{h \langle V_{\rm R}^2 \rangle^{1/2}} \right\}^{1/2} \exp\left(\frac{R}{2h}\right)$$

and this gives

$$\frac{Q V_{\rm rot}}{\langle V_{\rm R}^2 \rangle^{1/2}} \gtrsim 7.91$$

Comparing this to the equation for swing amplification we see that for spirals that are stable against global modes, swing amplification is possible for all modes with $m \geq 2$, at least at those radii where the rotation curve is flat.

Ostriker & Peebles[5] have also found from numerical experiments a general condition for global stability.

Stability occurs only when the ratio of kinetic energy in rotation S to the potential energy Ω

$$t = \frac{S}{|\Omega|} \lesssim 0.14$$

The virial theorem says that $2S + 2R + \Omega = 0$, where S is the kinetic energy in random motions.

Since $R/S > 0$, we would have expected t to have the range $0 - 0.5$ available.

[5] J.P. Ostriker & P.J.E. Peebles, Ap.J. 186, 467 (1973)

The criterion translates into $R/S \gtrsim 2.5$, while for the local Galactic disk it is about 0.15.

So disk galaxies require additional material with high random motion in order to conform to the criterion, either in the disk itself (e.g. the stars in the central region) or in the dark halo.

Tidal radius

Globular clusters have tidal radii due to the force field of the Galaxy. These radii can be estimated as follows.

Assume two point masses M (the Galaxy) and m (the cluster) and a separation R in a circular orbit (the following can be adapted to elliptical orbits as well with R the smallest separation).

Kepler's third law says

$$\frac{T^2}{a^3} = \frac{4\pi^2}{G(M+m)}$$

For a circular orbit we can find the angular velocity of the globular cluster around the center of gravity is

$$\Omega = \left[\frac{G(M+m)}{R^3} \right]^{1/2}$$

The center of gravity is at a distance $MR/(M+m)$ from the cluster.

Take a star at distance r from the center of the cluster in the direction of M and calculate where the total force on that star is zero. Thus in terms of accelleration (after dividing by G)

$$\frac{M}{(R-r)^2} - \frac{m}{r^2} - \frac{M+m}{R^3} \left(\frac{MR}{M+m} - r \right) = 0$$

Since r is much less than R we may expand the first term

$$\frac{M}{(R-r)^2} \approx \frac{M}{R^2}\left(1+2\frac{r}{R}\right)$$

Since m is small compared to M the third term can be reduced to

$$\frac{M+m}{R^3}\left(\frac{MR}{M+m}-r\right) = \frac{M}{R^2} - \frac{mr}{R^3}$$

Then the equation reduces to

$$\frac{3Mr}{R^3} - \frac{m}{r^2} = 0$$

The tidal radius then is the solution for r of this equation:

$$r_{\rm tidal} \sim R \left(\frac{m}{3M} \right)^{1/3}$$

For $M = 10^{12}$ M$_\odot$, $m = 10^5$ M$_\odot$ and $R = 10$ kpc we get $r_{\rm tidal} \approx 30$ pc.

Observed tidal radii can be used to constrain the mass distribution in the Galaxy.

STRUCTURE AND DYNAMICS OF GALAXIES

9. Galactic dynamics: the velocity ellipsoid

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/∼vdkruit

Beijing, September 2011

The Schwarzschild distribution

The distribution of space velocities of the local stars can be described with the so-called ellipsoidal distribution.

This was first introduced by Karl Schwarzschild and is therefore also called the Schwarzschild distribution.

The distribution is Gaussian along the principal axes, but has different dispersions. This anisotropy was Schwarzschild's explanation of the "star-streams" that were discovered by Kapteyn.

The general equation for the Schwarzschild distribution is

$$
\begin{aligned}
f(R, z, V_{\mathrm{R}}, V_\theta, V_{\mathrm{z}}) \quad = \quad & \frac{8\langle V_{\mathrm{R}}^2\rangle\langle V_\theta^2\rangle\langle V_{\mathrm{z}}^2\rangle}{\pi^{3/2}}\nu \\
& \exp\left[-\frac{V_{\mathrm{R}}^2}{2\langle V_{\mathrm{R}}^2\rangle} - \frac{(V_\theta - V_{\mathrm{t}})^2}{2\langle V_\theta^2\rangle} - \frac{V_{\mathrm{z}}^2}{2\langle V_{\mathrm{z}}^2\rangle} - \right. \\
& \left. \frac{V_{\mathrm{R}} V_\theta}{2\langle V_{\mathrm{R}} V_\theta\rangle} - \frac{V_{\mathrm{R}} V_{\mathrm{z}}}{2\langle V_{\mathrm{R}} V_{\mathrm{z}}\rangle} - \frac{(V_\theta - V_{\mathrm{t}}) V_{\mathrm{z}}}{2\langle V_\theta V_{\mathrm{z}}\rangle}\right]
\end{aligned}
$$

There is an interesting deduction that can be made from this ellipsoidal velocity distribution, which was done by Oort in the same paper in which he discovered differential rotation, defined the Oort constants and laid the foundations for "stellar dynamics"[1].

Take the asymmetric drift equation, insert this distribution and add the condition that $z = 0$ is a plane of symmetry.

Then you get an equation in terms of velocities and multiplications thereof that has to be identical, so that all terms need to be zero.

This is a lot of algebra (see Oort's paper).

[1] J.H.Oort, B.A.N. 4, 269 (1928), see also his chapter in Stars & Stellar Systems V, Galactic Structure, ed. Adriaan Blaauw & Maarten Schmidt, 455 (1965)

$$\Pi\frac{\partial f}{\partial\varpi}+\frac{\Theta}{\varpi}\left(\Theta\frac{\partial f}{\partial\Pi}-\Pi\frac{\partial f}{\partial\Theta}\right)+Z\frac{\partial f}{\partial z}+K_{\varpi}\frac{\partial f}{\partial\Pi}+K_z\frac{\partial f}{\partial Z}=0 \quad (6)$$

This equation is generally solvable *), but at present I shall only consider some particular solutions, which take account of the fact that the distribution of the peculiar motions of the stars has been found to approximate very closely to a function of the following form.

We shall assume, then, that the velocity distribution is of the ellipsoidal type and that the centre of symmetry of this distribution has a velocity Θ_0 with respect to the stationary co-ordinate system. The directions of the axes of the Schwarzschild ellipsoid will be left undetermined for the present, so that we find a distribution function of the following form:

$$f=f_0 e^{-h^2\Pi^2-k^2(\Theta-\Theta_0)^2-l^2Z^2-m\Pi(\Theta-\Theta_0)-n\Pi Z-p(\Theta-\Theta_0)Z} \quad (8)$$

in which h, k, l, m, n, p, f_0 and Θ_0 are functions of ϖ and z. Inserting (8) in equation (6) we get after dividing by $-f$ and arranging according to powers of Π, Θ, Z:

$$\begin{aligned}
&\Pi^3\frac{\partial h^2}{\partial\varpi}+\Pi^2\Theta\left(\frac{\partial m}{\partial\varpi}-\frac{m}{\varpi}\right)+\Pi^2 Z\left(\frac{\partial h^2}{\partial z}+\frac{\partial n}{\partial\varpi}\right)+\Pi\Theta^2\left(\frac{\partial k^2}{\partial\varpi}+\frac{2h^2-2k^2}{\varpi}\right)+\Pi\Theta Z\left(\frac{\partial m}{\partial z}+\frac{\partial p}{\partial\varpi}-\frac{p}{\varpi}\right)+\\
&+\Pi Z^2\left(\frac{\partial l^2}{\partial\varpi}+\frac{\partial n}{\partial z}\right)+\Theta^3\frac{m}{\varpi}+\Theta^2 Z\left(\frac{\partial k^2}{\partial z}+\frac{n}{\varpi}\right)+\Theta Z^2\frac{\partial p}{\partial z}+Z^3\frac{\partial l^2}{\partial z}-\Pi^2\frac{\partial(m\Theta_0)}{\partial\varpi}-\\
&2\Pi\Theta\left\{\frac{\partial(k^2\Theta_0)}{\partial\varpi}-\frac{k^2\Theta_0}{\varpi}\right\}-\Pi Z\left\{\frac{\partial(m\Theta_0)}{\partial z}+\frac{\partial(p\Theta_0)}{\partial\varpi}\right\}-\Theta^2\frac{m\Theta_0}{\varpi}-2\Theta Z\frac{\partial(k^2\Theta_0)}{\partial z}-Z^2\frac{\partial(p\Theta_0)}{\partial z}+\\
&+\Pi\left\{\frac{\partial(k^2\Theta_0^2)}{\partial\varpi}+2h^2K_{\varpi}+nK_z-\frac{1}{f_0}\frac{\partial f_0}{\partial\varpi}\right\}+\Theta(mK_{\varpi}+pK_z)+Z\left\{\frac{\partial(k^2\Theta_0^2)}{\partial z}+2l^2K_z+nK_{\varpi}-\frac{1}{f_0}\frac{\partial f_0}{\partial z}\right\}-\\
&m\Theta_0K_{\varpi}-p\Theta_0K_z=0
\end{aligned} \quad (9)$$

As this equation must hold for all values of Π, Θ and Z, the co-efficients of the different powers must vanish separately. We thus get the following conditions:

$$m = p = 0 \tag{10}$$

$$\frac{\partial h^2}{\partial \varpi} = \frac{\partial l^2}{\partial z} = 0 \tag{11}$$

$$\frac{\partial h^2}{\partial z} + \frac{\partial n}{\partial \varpi} = 0; \quad \frac{\partial l^2}{\partial \varpi} + \frac{\partial n}{\partial z} = 0; \quad \frac{\partial k^2}{\partial z} + \frac{n}{\varpi} = 0 \tag{12}$$

$$\frac{\partial k^2}{\partial \varpi} = \frac{2(k^2 - h^2)}{\varpi} \tag{13}$$

$$\frac{\partial (k^2 \Theta_\circ)}{\partial \varpi} = \frac{k^2 \Theta_\circ}{\varpi} \tag{14}$$

$$\frac{\partial (k^2 \Theta_\circ)}{\partial z} = 0 \tag{15}$$

$$\frac{1}{f_\circ}\frac{\partial f_\circ}{\partial \varpi} = \frac{\partial (k^2 \Theta_\circ^2)}{\partial \varpi} + nK_z + 2h^2 K_\varpi \tag{16}$$

$$\frac{1}{f_\circ}\frac{\partial f_\circ}{\partial z} = \frac{\partial (k^2 \Theta_\circ^2)}{\partial z} + nK_\varpi + 2l^2 K_z \tag{17}$$

stars considered. In our present notation we have thus:

$$A = \frac{1}{2}\left(\frac{\Theta_\circ}{\varpi} - \frac{\partial \Theta_\circ}{\partial \varpi}\right)$$

and similarly for the quantity derived from proper motions:

$$B = \frac{1}{2}\left(-\frac{\Theta_\circ}{\varpi} - \frac{\partial \Theta_\circ}{\partial \varpi}\right)$$

Thus, inserting these in (19):

$$h^2/k^2 = -B/(A-B) \tag{26}$$

The result is

$$\begin{aligned}
2\langle V_{\rm R}^2\rangle &= C_1 + \tfrac{1}{2}C_5 z^2 \\
2\langle V_{\theta}^2\rangle &= C_1 + C_2 R^2 + \tfrac{1}{2}C_5 z^2 \\
2\langle V_{\rm z}^2\rangle &= C_4 + \tfrac{1}{2}C_5 z^2 \\
2\langle V_{\rm R} V_{\rm z}\rangle &= -C_5 R z \\
\langle V_{\rm R} V_{\theta}\rangle &= \langle V_{\theta} V_{\rm z}\rangle = 0 \\
V_{\rm t} &= \frac{C_3 R}{C_1 + C_2 R^2 + \tfrac{1}{2}C_5 z^2}
\end{aligned}$$

The constants C_1 to C_5 are positive constants.

The density distribution at $z = 0$ follows from

$$\frac{\partial \ln \nu}{\partial R} = 2C_1 K_\mathrm{R} + \frac{C_2^2 R^2 + (2C_1 C_3^2 - C_1 C_2)R}{(C_2 R^2 + C_1)^2} - \frac{C_1 R}{C_5 R^2 + 2C_4}$$

and the vertical gradient from

$$\begin{aligned} \frac{\partial \ln \nu}{\partial z} \quad = \quad & (C_5 R^2 + 2C_4)K_\mathrm{z} - C_5 z \left[R K_\mathrm{R} + \right. \\ & \left. \frac{2(C_2 + 2C_3^2)R^2 + C_5 z^2 + 2C_1}{(2C_2 R^2 + C_5 z^2 + 2C_1)^2} + \frac{1}{C_5 z^2 + 2C_1} \right] \end{aligned}$$

Oort's derivation only holds if the stellar velocity distribution is exactly Gaussian.

It is too restrictive (e.g. it does not allow high-velocity stars) and therefore, it cannot be used for a description of galactic dynamics.

In reality, the velocity distributions are not precisely Gaussian and are better seen as a superposition of Gaussians (such as of groups of stars with similar ages).

So, these equations are of historical interest only. However, it is interesting to see that Oort assumed that $C_5 = 0$. This uncoupled the radial and vertical motion (as for a third integral).

Properties of the velocity ellipsoid

For the solar neighbourhood, but probably anywhere in galactic disks, the velocity distribution of the stars is very anisotropic.

- The *ratio of the radial versus tangential velocity dispersions* is determined by the local differential rotation and can be derived using the epicycle approximation.

  The axis ratio of the epicycles depend on the local Oort constants and therefore axis ratio of the velocity ellipsoid is

$$\frac{\langle V_\theta^2 \rangle}{\langle V_\mathrm{R}^2 \rangle} = \frac{-B}{(A-B)}$$

- The *ratio of the vertical to radial velocity dispersion* is unconstrained, as a result of the third integral.

  However, the existence of a third integral does not necessarily imply that the velocity distribution has to be anisotropic.

  If no third integral would exist, the velocity distribution would have to be isotropic, according to Jeans.

- The *long axis of the velocity ellipsoid in the plane* should point to the center.

  However, it does not in practice. This is called the "deviation of the vertex" and presumably is due to local irregularities in the Galactic gravitational field.
- The *long axis of the velocity ellipsoid outside the plane* has an unknown orientation.

  This has been a longstanding problem, also sometimes referred to as the "tilt" of the velocity ellipsoid.

  Oort assumed the long axis to be parallel to the Galactic plane ($C_5 = 0$), but later assumed it to be pointing always towards the Galactic center.

There is an interesting consequence in this respect of flat rotation curves[2].

Take the Poisson equation for the axisymmetric case

$$\frac{\partial K_{\mathrm{R}}}{\partial R} + \frac{K_{\mathrm{R}}}{R} + \frac{\partial K_{\mathrm{z}}}{\partial z} = -4\pi G\rho(R, z)$$

For a flattened disk, it can be shown that the first two terms in or near the plane $z = 0$ are

$$\frac{\partial K_{\mathrm{R}}}{\partial R} + \frac{K_{\mathrm{R}}}{R} \approx 2(A - B)(A + B)$$

[2] P.C. van der Kruit & K.C. Freeman, Ap.J. 303, 556 (1986)

In 1965, Oort[3] estimated that the first two terms are in the solar neighborhood and in the plane of the Galaxy about 34 times smaller than the third term.

For a flat rotation curve we have $A = -B$, so the equation reduces to that for a plane-parallel case.

On this basis one may expect for small distances from the plane that the long axis is parallel to the plane.

So with flat rotation curves the plane-parallel case turns out to be a much better description of reality than may expected on the basis of the form of the Poisson equation.

[3] J.H. Oort, Stars & Stellar Systems V, Galactic Structure, ed. Adriaan Blaauw & Maarten Schmidt, p. 455 (1965)

The closure problem

The hydrodynamical equations were obtained by multiplication of the Liouville equation with velocities and then integrating over all velocity space.

This system is not complete (there is a "closure problem"): there are only three equations for eight unknowns (the density, rotation velocity, three velocity dispersions and three "cross-dispersions" as a function of position).

In principle one could take higher order moments (by multiplying the Jeans equations with velocities once more and again integrating over all velocities), but this produces more extra unknowns than extra equations.

However, with reasonable assumptions[4] there has been some progress.

It works as follows. In analogy to the second moment

$$\sigma_{\rm ab}(R,z) = \langle V_{\rm a} V_{\rm b}\rangle = \frac{1}{\nu}\int (V_{\rm a} - \langle V_{\rm a}\rangle)(V_{\rm b} - \langle V_{\rm b}\rangle) f d^3 V$$

one defines the third and fourth moments as

$$S_{\rm abc}(R,z) = \langle V_{\rm a} V_{\rm b} V_{\rm c}\rangle = \frac{1}{\nu}\int (V_{\rm a} - \langle V_{\rm a}\rangle)(V_{\rm b} - \langle V_{\rm b}\rangle)(V_{\rm c} - \langle V_{\rm c}\rangle) f d^3 V$$

$$\begin{aligned} T_{\rm abcd}(R,z) &= \langle V_{\rm a} V_{\rm b} V_{\rm c} V_{\rm d}\rangle \\ &= \frac{1}{\nu}\int (V_{\rm a} - \langle V_{\rm a}\rangle)(V_{\rm b} - \langle V_{\rm b}\rangle)(V_{\rm c} - \langle V_{\rm c}\rangle)(V_{\rm d} - \langle V_{\rm d}\rangle) f d^3 V \end{aligned}$$

---

[4] P.O. Vandervoort, Ap.J. 195, 333 (1975); and in particular P. Amendt & P. Cuddeford, Ap.J. 368, 79 (1991); P. Cuddeford & P. Amendt, MNRAS 256, 166 (1992)

The third moment correponds to the "skewness" (e.g. $S_{\rm RRR}/(\sigma_{\rm RR})^{3/2}$). It is zero for a Gaussian, since this is completely symmetric.

The fourth moment corresponds to the "kurtosis" (e.g. $T_{\rm RRRR}/(\sigma_{\rm RR})^{2}$), which decribes how peaked the distribution is; a Gaussian has a kurtosis of 3.

The assumptions of Amendt & Cuddeford were

- All parameters can be expanded in terms of a small parameter ϵ, which is the ratio of the radial velocity dispersion to the rotation velocity.
- The ordering scheme of these remains such that only terms in the leading order have to be taken. Thus e.g. in

$$S_{\rm abc} = \sum_{n=0}^{\infty} \epsilon^{n+3} S_{\rm abc}^{n+3}$$

  the higher order components of $S_{\rm abc}$ become smaller with n.
- The velocity distributions are Gaussian (Schwarzschild) up to one more order than required by the equations. This happens to translate e.g. for the kurtosis into

$$\frac{T_{\rm Rzzz}}{\sigma_{\rm RZ}^2 \sigma_{\rm zz}^2} = 3 + O(\epsilon^3)$$

These assumptions mean that we have to do with a cool, highly flattened and quasi-isothermal system.

Then the system can be closed and four more equations result after a lot of algebra. Here they are from the publication

Since it is rare for $\overline{v_\phi}(R, z)$ to be known *a priori*, particularly its z-behavior, we use equations (A1) and (A4) to eliminate $\overline{v_\phi}$ and v, and after some tedious algebra arrive at a system of four partial differential equations for the four components of the velocity dispersion tensor, in terms of the potential Φ. The first three of these equations are

$$\Phi_{,z}(\sigma_{zz}^2\,\partial_z\,\sigma_{Rz}^2 + \sigma_{Rz}^2\,\partial_R\,\sigma_{Rz}^2) = \frac{1}{3}\left[\frac{(R^3\Phi_{,R})_{,R}}{R^3} - \Phi_{,zz}\right]\sigma_{Rz}^2\,\sigma_{zz}^2 - \frac{\Phi_{,z}}{R}\,\sigma_{Rz}^4 + \frac{\Phi_{,Rz}}{3}\,\sigma_{zz}^2(\sigma_{zz}^2 - \sigma_{RR}^2)\,, \tag{79}$$

$$\sigma_{zz}^2\,\partial_z\,\sigma_{zz}^2 + \sigma_{Rz}^2\,\partial_R\,\sigma_{zz}^2 = 0\,, \tag{80}$$

$$\begin{aligned}\Phi_{,z}\left[4\,\sigma_{Rz}^2\,\partial_z\,\sigma_{Rz}^2 + \sigma_{zz}^2\,\partial_z\,\sigma_{RR}^2 + 4\,\sigma_{RR}^2\,\partial_R\,\sigma_{Rz}^2 + \frac{4}{\sigma_{zz}^2}(\sigma_{RR}^2\,\sigma_{zz}^2 - \sigma_{Rz}^4)\partial_z\,\sigma_{zz}^2 + \sigma_{Rz}^2\,\partial_R\,\sigma_{RR}^2\right] &= 2\Phi_{,Rz}\,\sigma_{Rz}^2\,\sigma_{zz}^2 - \frac{1}{R}(6\Phi_{,z} + 4R\Phi_{,Rz})\sigma_{Rz}^2\,\sigma_{RR}^2 \\ &+ \frac{1}{R}(6\Phi_{,R} + 2R\Phi_{,RR})\sigma_{zz}^2\,\sigma_{RR}^2 - 2\Phi_{,zz}\,\sigma_{Rz}^4 + 2\Phi_{,Rz}\,\frac{\sigma_{Rz}^6}{\sigma_{zz}^2} - \frac{8}{R}\,\Phi_{,R}\,\sigma_{zz}^2\,\sigma_{\phi\phi}^2 + \frac{8}{R}\,\Phi_{,z}\,\sigma_{Rz}^2\,\sigma_{\phi\phi}^2\,.\end{aligned} \tag{81}$$

The new form of the fourth equation, equation (C1), is cumbersome and is included in Appendix C.

For ease of notation we make the substitutions $\sigma^2_{Rz} = t$; $\sigma^2_{RR} = u$; $\sigma^2_{\phi\phi} = v$; $\sigma^2_{zz} = w$, and use subscript comma notation to denote partial differentiation on the potential. When we eliminate the mean velocity and number density from equation (78) we obtain the final equation of our closed system of partial differential equations:

$$
\begin{aligned}
&\Phi_{,z}\Big\{(2t^4 + u^2w^2 - 3uwt^2)(\partial_z w)^2 + (2t^3w - utw^2)(\partial_z w\,\partial_z t) + (4t^4 - 3uwt^2)(\partial_z w\,\partial_R t) \\
&\quad - 2t^2w^2(\partial_z t)^2 - 3t^3w(\partial_R t\,\partial_z t) - \frac{1}{2}w^2t^2(\partial_R t\,\partial_z u) - uwt^2(\partial_R t)^2 - \frac{1}{2}t^3w(\partial_R t\,\partial_R u) - \frac{1}{2}w^2t^2(\partial_z w\,\partial_z u) - \frac{1}{2}t^3w(\partial_z w\,\partial_R u)\Big\} \\
&\quad + [\Phi_{,Rz}(t^3 - uwt)(w^2 - uw + t^2) - \Phi_{,zz}t^2w(t^2 - uw)]\partial_z w \\
&\quad + \left\{\Phi_{,Rz}(2w^2 - 2uw + t^2) + \left[\frac{1}{R^3}(R^3\Phi_{,R})_{,R} - 2\Phi_{,zz}\right]tw - \frac{3}{R}\Phi_{,z}t^2\right\}t^2w\,\partial_z t \\
&\quad + \left[2\Phi_{,Rz}t^3(w^2 + t^2) - \frac{3}{R}(R\Phi_{,z})_{,R}t^3uw - 2\Phi_{,zz}t^4w + \frac{1}{R^3}(R^3\Phi_{,R})_{,R}uw^2t^2\right]\partial_R t \\
&\quad + \left[\frac{1}{2R^3}(R^3\Phi_{,R})_{,R}w^2 - \frac{1}{2R^3}(R^3\Phi_{,z})_{,R}tw\right]t^2w\,\partial_z u + \left[\frac{1}{2R^3}(R^3\Phi_{,R})_{,R}tw - \frac{1}{2R^3}(R^3\Phi_{,z})_{,R}t^2\right]t^2w\,\partial_R u \\
&\quad - \frac{2}{R}(\Phi_{,z}t - \Phi_{,R}w)t^2w^2\,\partial_z v - \frac{2}{R}(\Phi_{,z}t - \Phi_{,R}w)t^3w\,\partial_R v = \frac{4}{R^2}(\Phi_{,z}t - \Phi_{,R}w)t^3wv\,. \qquad \text{(C1)}
\end{aligned}
$$

These equations can be used to derive further information on the velocity ellipsoid in cool, flattened galaxies (i.e. in disks).

There are a few applications.

The tilt of the velocity ellipsoid.

From the equations it can be found that

$$\frac{\partial \langle V_{\rm R} V_{\rm z} \rangle}{\partial z}(R,0) = \lambda(R) \left(\frac{\langle V_{\rm R}^2 \rangle - \langle V_{\rm z}^2 \rangle}{R} \right)(R,0)$$

with

$$\lambda(R) = \left[R^2 \frac{\partial^3 \Phi}{\partial R \partial z^2} \left(3 \frac{\partial \Phi}{\partial R} + R \frac{\partial^2 \Phi}{\partial R^2} - 4R \frac{\partial^2 \Phi}{\partial z^2} \right)^{-1} \right](R,0)$$

For a flat rotation curve this gives

$$\lambda(R,0) = \left(\frac{2\pi G R^3}{V_{\rm t}^2 - 8\pi G R^2 \rho} \frac{\partial \rho}{\partial R} \right)(R,0)$$

The radial dependence velocity dispersions.

A solution of the equations has the following form

$$f_1(R)\left(\frac{\partial\langle V_{\mathrm{R}}^2\rangle}{\partial R}\right)(R,0)+f_2(R)\langle V_{\mathrm{R}}^2\rangle(R,0)=f_3(R)$$

The functions f have complicated forms and are related to the local potential and kinematics through parameters α, β and γ.

$$\alpha=-\left(\frac{\partial^2\Phi}{\partial z^2}\right)(R,0)=-\lambda^2$$

where λ is the vertical frequency.

$$\begin{aligned}\beta &= \left(\frac{\partial^2 \Phi}{\partial R^2}\right)(R,0) + \frac{3}{R}\left(\frac{\partial \Phi}{\partial R}\right)(R,0) \\ &= \frac{1}{R^3}\left(\frac{\partial (R^2 V_{\mathrm{t}}^2)}{\partial R}\right)(R,0) = -\kappa^2\end{aligned}$$

with κ the epicyclic frequency.

$$\begin{aligned}\gamma &= \frac{1}{4}\left\{R\left(\frac{\partial^2 \Phi}{\partial R^2}\right)\left(\frac{\partial \Phi}{\partial z}\right)^{-1} + 3\right\}(R,0) \\ &= \left(\frac{\langle V_\theta^2\rangle}{\langle V_{\mathrm{z}}^2\rangle}\right)(R,0)\end{aligned}$$

which is the anisotropy in the velocity distribution.

This can be solved for a given potential; the most realistic solution is with a logarithmic-exponential potential

$$\Phi(R, z) = A \ln R - BR - Cz^2 \exp\left(-\frac{R}{h}\right),$$

which has

$$\left(\frac{\partial^2 \Phi}{\partial z^2}\right)(R, 0) = 2C \exp\left(-\frac{R}{h}\right)$$

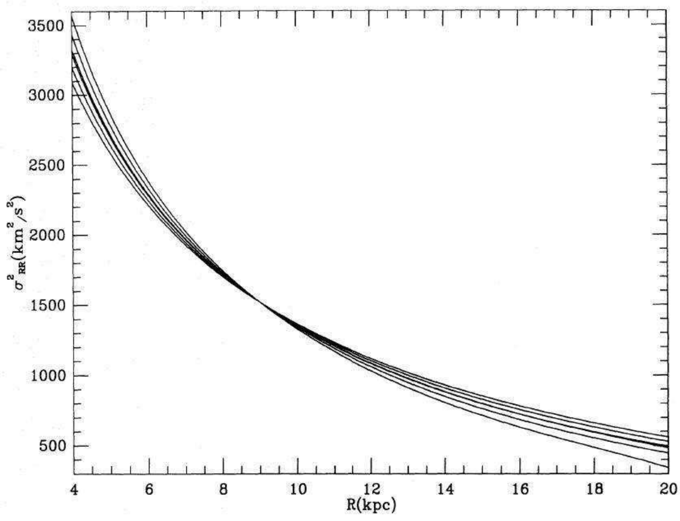
and thus an exponential density profile (as has been observed for the surface brightness distribution).

The resulting distributions show

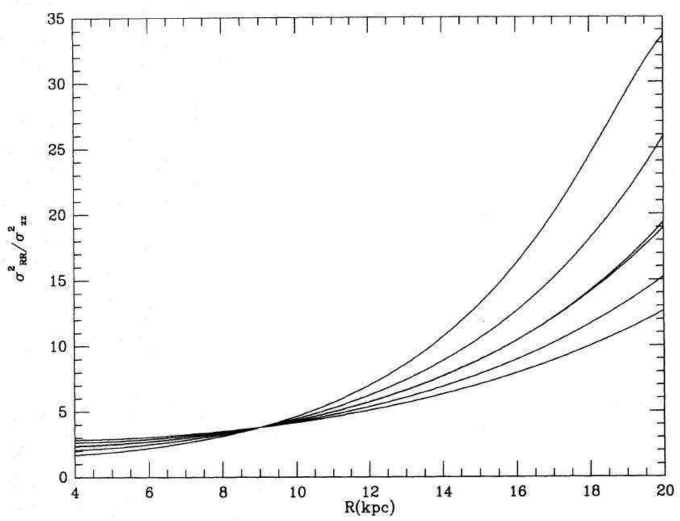
- The radial velocity dispersion $\langle V_{\mathrm{R}}^2 \rangle$ decreases more or less exponentially with radius
- The velocity anisotropy $\langle V_{\mathrm{R}}^2 \rangle / \langle V_{\mathrm{z}}^2 \rangle$ is roughly constant (in the inner regions at least)
- Toomre Q is constant with radius, except near the center.

The following graphs show this for a number of combinations of values for C and h.

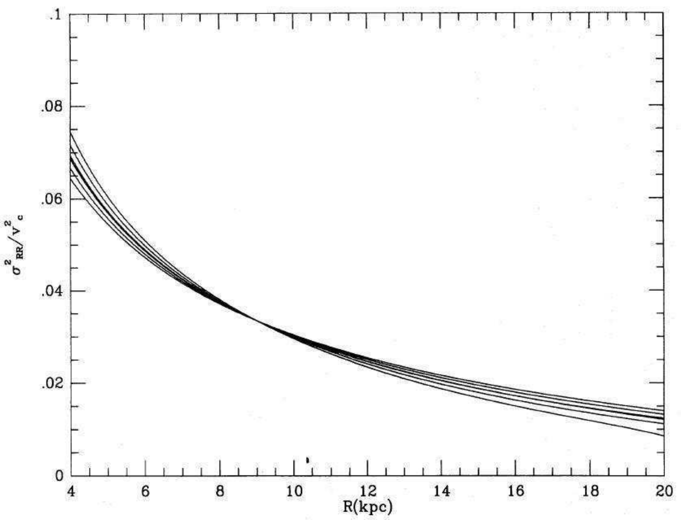
The (square of the) radial velocity dispersion $\langle V_{\mathrm{R}}^2 \rangle$.



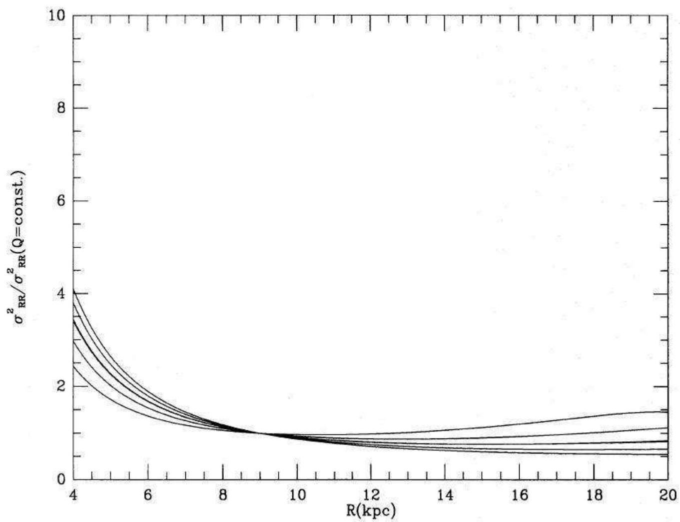
The axis ratio of the velocity ellipsoid $\langle V_{\mathrm{R}}^2 \rangle / \langle V_{\mathrm{z}}^2 \rangle$.



The temperature parameter $\langle V_{\mathrm{R}}^2 \rangle / V_{\mathrm{t}}^2$.



The axis ratio of the velocity ellipsoid w.r.t. Q = constant.



A further application is the following new equation

$$\left(\frac{\partial^2 \langle V_{\rm z}^2 \rangle}{\partial z^2}\right)(R,0) = -\lambda(R)\left[\left(\frac{\langle V_{\rm R}^2\rangle - \langle V_{\rm z}^2\rangle}{R}\right)\frac{\partial \ln \langle V_{\rm z}^2\rangle}{\partial R}\right](R,0)$$

Since $\lambda(R) > 0$, $\langle V_{\rm R}^2\rangle > \langle V_{\rm z}^2\rangle$ and $\langle V_{\rm z}^2\rangle$ decreasing with R, the righthand side of the equation has to be positive.

That means that $\langle V_{\rm z}^2\rangle$ has a minimum in the plane.

So disks are not strictly isothermal in z and numerical values suggest less peaked in density than the exponential function.

The final application gives a more accurate estimate of the velocity anisotropy in the plane through

$$\frac{\langle V_\theta^2\rangle}{\langle V_\mathrm{R}^2\rangle} = \frac{1}{2}\left\{1+\frac{\partial\ln V_\mathrm{t}}{\partial\ln R}-\frac{S_{\theta\theta\theta}}{V_\mathrm{t}\langle V_\mathrm{R}^2\rangle}+\frac{1}{\nu R V_\mathrm{t}\langle V_\mathrm{R}^2\rangle}\frac{\partial R^2\nu S_{\mathrm{RR}\theta}}{\partial R}+\right.$$
$$\left.\frac{R}{V_\mathrm{t}\langle V_\mathrm{R}^2\rangle}\frac{\partial S_{\mathrm{R}\theta\mathrm{z}}}{\partial z}+\frac{V_\mathrm{t}^2-V_\mathrm{rot}^2}{V_\mathrm{t}\langle V_\mathrm{R}^2\rangle^2}S_{\mathrm{RR}\theta}+\frac{T_{\mathrm{RR}\theta\theta}}{\langle V_\mathrm{R}^2\rangle^2}\right\}$$

In practice this can be approximated as

$$\frac{\langle V_\mathrm{R}^2\rangle}{\langle V_\theta^2\rangle}=\frac{1}{2}\left(1+\frac{\partial\ln V_\mathrm{t}}{\partial\ln R}+\frac{T_{\mathrm{RR}\theta\theta}}{\langle V_\mathrm{R}^2\rangle^2}\right)$$

This constitutes a small correction to the classical result

$$\frac{\langle V_\mathrm{R}^2\rangle}{\langle V_\theta^2\rangle}=\frac{1}{2}\left(1+\frac{\partial\ln V_\mathrm{t}}{\partial\ln R}\right)=\frac{-B}{A-B}$$

STRUCTURE AND DYNAMICS OF GALAXIES

10. Galactic dynamics: The self-consistency problem and potential theory

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/$\sim$vdkruit

Beijing, September 2011

The self-consistency problem

Ideally, one would like to construct self-consistent, self-gravitating models for galaxies, by solving the two coupled, fundamental equations:

$$u\frac{\partial f}{\partial x} + v\frac{\partial f}{\partial y} + w\frac{\partial f}{\partial z} - \frac{\partial \Phi}{\partial x}\frac{\partial f}{\partial u} - \frac{\partial \Phi}{\partial y}\frac{\partial f}{\partial v} - \frac{\partial \Phi}{\partial z}\frac{\partial f}{\partial w} = 0.$$

and

$$\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} \equiv \nabla^2 \Phi = 4\pi G \rho(x, y, z)$$

Unfortunately, in general this is not possible.

There are two possible apporaches:

- The direct method. Assume a potential Φ on the basis of the density distribution, inferred from observations. Then use the observed kinematics to derive further properties of the distribution function.
- The inverse method. Make a guess for the dependence of the distribution function on the isolating integrals and calculate the density, potential, motions and velocity distrubutions.

The direct approach is straightforward in e.g. the case of the vertical distributions in a galactic disk (where it reduces to a one-dimensional treatment).

The inverse method makes use of functional solutions of well-defined cases, such a isothermal models.

First we turn to the direct method.

The Schwarzschild method

Schwarzschild[1] proceeds as follows:

- Choose a density distribution for the system you want to model.
- Solve Poission's equation (usually numerically).
- Compute a library (many hundreds) of orbits in this potential and calculate the density distribution that each orbit generates.
- Add these with appropriate weights to recover the density distribution started from (usually this involves "linear or quadratic programming").

[1] M. Schwarzschild, Ap.J. 232, 236 (1979)

Often it is possible to use constraints as the observations of the kinematics of the stars, i.e. their motions and velocity dispersions.

There is uncertainty whether any outcome is unique.

But it is an extremely powerful approach.

Isothermal solutions and related results

For simple geometries full semi-analytical solutions for the distribution function to the set of two fundamental equations can be obtained.

These solutions refer to self-gravitating systems, which means that ρ and ν are the same.

Examples are spherical density distributions or density distributions on stratified layers with isothermal velocity distributions (equal velocity dispersions at all positions),

Isothermal sphere and King models

The Poisson equations for spherical symmetry was

$$\frac{1}{R^2}\frac{\partial}{\partial R}(R^2 K_{\rm R}) = -4\pi G\rho(R)$$

and the Jeans equation

$$\frac{\partial}{\partial R}(\nu\langle V_{\rm R}^2\rangle) + \frac{\nu}{R}\{2\langle V_{\rm R}^2\rangle - V_{\rm t}^2 - \langle (V_\theta - V_{\rm t})^2\rangle - \langle V_\phi^2\rangle\} = \nu K_{\rm R}$$

If the velocity distribution is isotropic and if there is no rotation this reduces to

$$\langle V^2\rangle\frac{\partial\rho}{\partial R} = \rho K_{\rm R}$$

Here V is the radial velocity.

The equations can be combined to give

$$\frac{\langle V^2 \rangle}{R^2} \frac{\partial}{\partial R} \left(R^2 \frac{\partial \ln \rho}{\partial R} \right) = -4\pi G \rho$$

The solution is

$$\rho(R) = \frac{\langle V^2 \rangle}{2\pi G} R^{-2}$$

This is called the singular isothermal sphere, since the density is infinite at the center.

Note that we have not constrained the functional form of the velocity distribution.

A well-behaved solution is obtained by assuming that the velocity distribution is Gaussian.

There is in this spherical, non-rotating case only one isolating integral of motion, namely the energy E.

According to Jeans' theorem then the distribution function is only a function E.

So take the distribution function to be

$$f(E) = \mathrm{const.} \times \mathrm{e}^{-E/\langle V^2 \rangle}$$

With $E = -\Phi + \frac{1}{2}V^2$ integration over all V gives

$$\rho(R) = \rho(0)\mathrm{e}^{-\Phi(R)/\langle V^2\rangle}$$

Now set the boundary conditions $\rho(0) = \rho_\circ$ and $(d\rho/dR)_{\mathrm{z}=0} = 0$.
Then the solution

$$\rho(R) = \rho_0 \mathrm{e}^{-\Phi}$$

can be found from a numerical integration where Φ follows from

$$\mathrm{e}^{-\Phi} = \frac{1}{\chi^2}\frac{d}{d\chi}\left(\chi^2\frac{d\Phi}{d\chi}\right) \quad ; \quad \chi = \left(\frac{\langle V^2\rangle}{4\pi G\rho_0}\right)^{1/2} R$$

For large R this becomes

$$\rho(R) = \frac{\langle V^2 \rangle}{2\pi G} R^{-2}$$

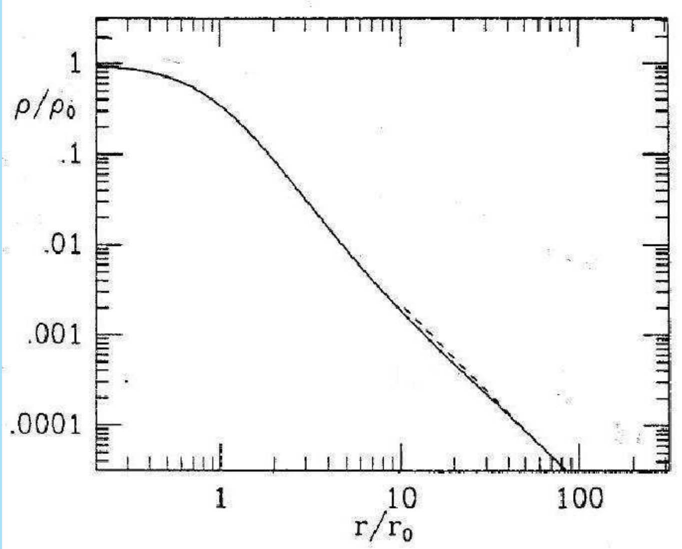
and thus approaches the singular isothermal sphere.

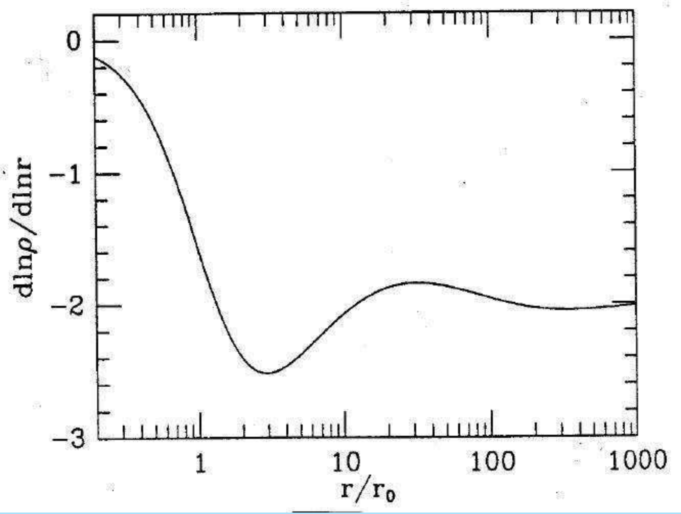
This solution has a natural length-scale that is called the core radius (also King radius)

$$R_0 = \left(\frac{4\pi G \rho_0}{9 \langle V^2 \rangle} \right)^{-1/2}$$

At this core radius the projected surface density is roughly half the central one.

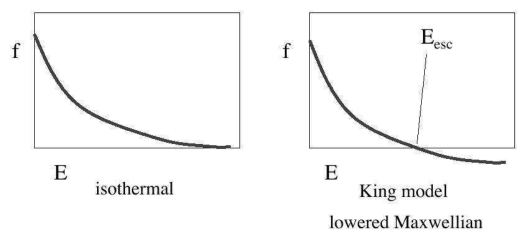
The next slides show the density distribution and the logarithmic density slope.


1
.1
.01
.001
.0001
ρ/ρ_0
1
10
100
r/r_0


0
−1
−2
−3
dlnρ/dlnr
1
10
100
1000
r/r₀

King models

King models are adapted isothermal spheres with a tidal radius R_t and a corresponding upper boundary in the velocity distribution.



The distribution function is

$$f(E) = \mathrm{const.} \left[\mathrm{e}^{-E/\langle V^2 \rangle} - \mathrm{e}^{-E_\mathrm{esc}/\langle V^2 \rangle} \right] \quad \text{for } E < E_\mathrm{esc}$$

$$0 \quad \text{for } E > E_\mathrm{esc}$$

Using again $E = -\Phi + \frac{1}{2}V^2$ and defining the zero-point of Φ such that $E_{\rm esc} = 0$ we may write this as

$$f(E) = \mathrm{const.}\left[\mathrm{e}^{-E/\langle V^2\rangle} - 1\right] \quad \text{for } E > 0$$

Integrating over all velocities then gives

$$\rho(R) = \rho_\circ \left[e^{\Phi(R)/\langle V^2\rangle} \mathrm{erf}\left(\sqrt{\frac{\Phi}{\langle V^2\rangle}}\right) - \sqrt{\frac{4\phi}{\pi\langle V^2\rangle}}\left(1 + \frac{2\Phi}{3\langle V^2\rangle}\right)\right]$$

Here erf is the Error Function.

Then we get

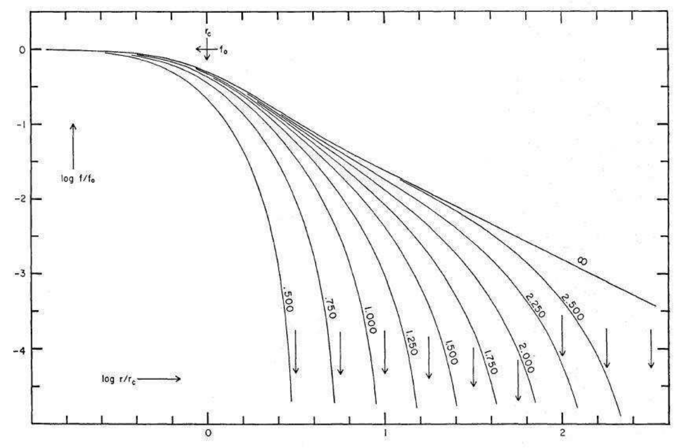
$$\frac{d}{dR}\left(R^2\frac{d\Phi}{dR}\right) = -4\pi G\rho_\circ R^2\left[e^{\Phi(R)/\langle V^2\rangle}\mathrm{erf}\left(\sqrt{\frac{\Phi}{\langle V^2\rangle}}\right) - \sqrt{\frac{4\phi}{\pi\langle V^2\rangle}}\left(1+\frac{2\Phi}{3\langle V^2\rangle}\right)\right]$$

This again has to be numerically integrated from the center outwards.

At the tidal radius R_t the density drops to zero.

The ratio $c = \log(R_\mathrm{t}/R_\circ)$ is called the concentration.

Here are some models in projected surface density[2].



[2] I.R. King, A.J. 71, 64 (1966)

The total mass is

$$M(R_\mathrm{t}) = \frac{2}{G} \langle V^2 \rangle r_\circ \, f\left(\frac{R_\mathrm{t}}{R_\circ}\right)$$

and the central surface density

$$\sigma_\circ = \rho_0 r_0 \, g\left(\frac{R_\mathrm{t}}{R_\circ}\right)$$

The functions f and g can only be calculated numerically and are given in the literature. The velocity dispersion is

$$\langle V^2 \rangle^{1/2} \propto \frac{\rho_\circ M(R_\mathrm{t})}{f\left(R_\mathrm{t}/R_\circ\right) g\left(R_\mathrm{t}/R_\circ\right)}$$

King models are useful to describe globular clusters and to some extent elliptical galaxies.

Isothermal sheet and other vertical distributions

For a self-gravitating isothermal sheet the basic equations become

$$\frac{\partial K_{\mathrm{z}}}{\partial z} = -4\pi G\rho(z)$$

and

$$\langle W^2\rangle \frac{\partial \nu}{\partial z} = \nu K_{\mathrm{z}}$$

The two basic equations can be combined into

$$-4\pi G\rho(z) = \langle W^2\rangle \frac{d^2}{dz^2}\left\{\ln\frac{\rho(z)}{\rho(0)}\right\}$$

The solution is

$$\rho(z) = \frac{\langle W^2 \rangle}{2\pi G z_0^2} \operatorname{sech}^2 \left(\frac{z}{z_\circ} \right)$$

The corresponding surface density is

$$\sigma = 2 z_0 \rho_\circ$$

and the relation to the velocity dispersion

$$\langle W^2 \rangle = \pi G \sigma z_\circ$$

The vertical force results from integration of Poisson's equation as

$$K_{\mathrm{z}} = -2 \frac{\langle W^2 \rangle}{z_\circ} \tanh \left(\frac{z}{z_\circ} \right)$$

Usefull approximations are

$$\mathrm{sech}^2\left(\frac{z}{z_0}\right) = \exp\left(-\frac{z^2}{z_0^2}\right) \qquad \text{for} \quad z \ll z_0$$

$$\mathrm{sech}^2\left(\frac{z}{z_0}\right) = 4\exp\left(-\frac{2z}{z_0}\right) \qquad \text{for} \quad z \gg z_0$$

The isothermal sheet is used to describe vertical distributions in stellar disks.[3]

[3] P.C. van der Kruit & L. Searle, A.&A. 95, 105 (1981)

For a second isothermal component of negligible mass and different velocity dispersion in this force-field we find

$$\rho_{\mathrm{II}}(z) = \rho_{\mathrm{II}}(0)\mathrm{sech}^{2p}\left(\frac{z}{z_\circ}\right)$$

where

$$p = \frac{\langle W^2 \rangle}{\langle W^2 \rangle_{\mathrm{II}}}$$

An application of this is for example the HI-gas layer inside a stellar disk that contains most of the surface density.

Exponential and sech-distributions

The isothermal sheet is only an approximate description of the vertical distribution of stars in disks of galaxies. There is a range of generations of stars, each with their own velocity dispersion.

Often used is the exponential distribution, since it is a convenient fitting function.

Since the velocity dispersion now varies with z we have to write the equation in terms of the velocity dispersion in he plane $\langle W^2 \rangle_\circ^{1/2}$. The equations corresponding to this case are[4]:

$$\rho(z) = \frac{\langle W^2 \rangle_\circ}{2\pi G Z_\mathrm{e}^2} \exp\left(-\frac{z}{z_\mathrm{e}}\right)$$

[4] P.C. van der Kruit, A.&A., 192, 117 (1988)

$$\sigma = 2 z_{\rm e} \rho_\circ$$

$$\langle W^2 \rangle_\circ = \pi G \sigma z_{\rm e}$$

$$K_{\rm z} = -2\pi G\sigma \left\{ 1 - \exp\left(-\frac{z}{z_{\rm e}} \right) \right\}$$

If an isothermal component of negligible mass moves in this force field, then

$$\rho_{\rm II}(z) = \rho_{\rm II}(0) \exp\left[-\frac{2pz}{z_{\rm e}} + 2p \left\{ 1 - \exp\left(-\frac{z}{z_{\rm e}} \right) \right\} \right]$$

where now

$$p = \frac{\langle W^2 \rangle_0}{\langle W^2 \rangle_{\rm II}}$$

As an intermediate case between the isothermal solution and the exponetial it is also possible to use the sech-distribution[5].

This corresponds probably closest to reality. The equations then are:

$$\rho(z) = \frac{2\langle W^2\rangle_{\mathrm{II}}}{\pi^3 G z_{\mathrm{e}}^2} \mathrm{sech}\left(\frac{z}{z_{\mathrm{e}}}\right)$$

$$\sigma = \pi \rho_\circ z_{\mathrm{e}}$$

$$\langle W^2\rangle_{0\circ} = \frac{\pi^2}{2} G \sigma z_{\mathrm{e}}$$

$$K_{\mathrm{z}} = -4 G \sigma \arctan\left\{\sinh\left(\frac{z}{z_{\mathrm{e}}}\right)\right\}$$

[5] P.C. van der Kruit, A.&A. 192, 127 (1988)

For the second isothermal component we now get

$$\rho_{\mathrm{II}}(z) = \rho_{\mathrm{II}}(0)\exp\left\{-\frac{8}{\pi^2}pI\left(\frac{z}{z_{\mathrm{e}}}\right)\right\}$$

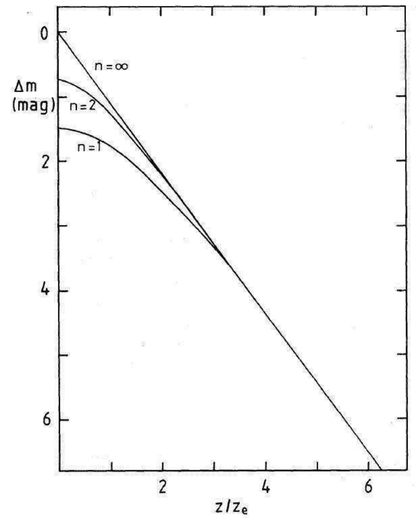
where

$$I(y) = \int_0^y \arctan(\sinh x)dx$$

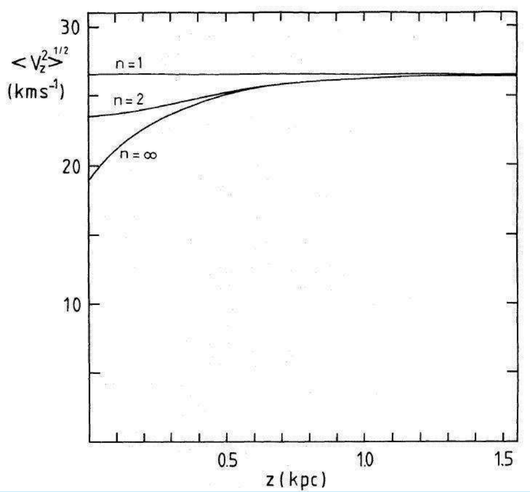
This integral can be evaluated easily by numerical methods or through a series expansion.

The properties are illustrated in the following figures, where properties appropriate for the Solar Neighborhood have been chosen.

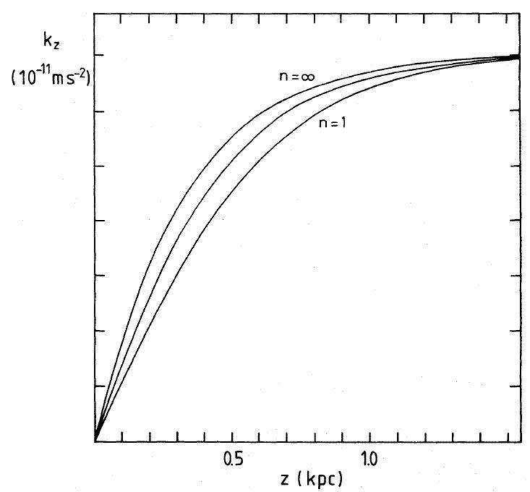
The density distributions as a function of z expressed in magnitudes.



The velocity dispersions as a function of z.



The vertical force K_z as a function of z.



Potential theory

General axisymmetric theory

Much attention has been paid to inverting Poisson's equation. For the axisymmetric case:

$$\frac{\partial^2\Phi}{\partial R^2} + \frac{1}{R}\frac{\partial\Phi}{\partial R} + \frac{\partial^2\Phi}{\partial z^2} = 4\pi G\rho(R,z)$$

so that the potential (and the forces) can be calculated when the density distribution is given.

This is a limited problem in that it does not involve the continuity equation and the distribution function and therefore is not a general solution for a dynamical system, such as the isothermal solutions above.

At the basis lies the Hankel (or Fourier-Bessel) transform, which in the radial direction for the density is

$$\tilde{\rho}(k,z) = \int_0^\infty u J_0(ku)\rho(u,z)du$$

J_0 is the Bessel function of the first kind.

The important property, why this is useful, is that the transform can be inverted:

$$\rho(R,z) = \int_0^\infty k J_0(kR)\tilde{\rho}(k,z)dk$$

Now, if we take this transform in the radial direction for both sides of the Poisson equation we get[6]

$$-k^2\tilde{\Phi}(k,z) + \frac{\partial^2}{\partial z^2}\tilde{\Phi}(k,z) = 4\pi G\tilde{\rho}(k,z)$$

This linear non-homogeneous ordinary differential equation can be solved to give

$$\tilde{\Phi}(k,z) = -\frac{2\pi G}{k}\int_{-\infty}^{\infty} \exp\left(-k|z-v|\right)\tilde{\rho}(k,v)dv$$

[6] S. Casertano, MNRAS 203, 735 (1983)

Using this, Poisson's equation can then be inverted to

$$\Phi(R,z) = -2\pi G \int_0^\infty \int_{-\infty}^\infty J_0(kR)\tilde{\rho}(k,v)e^{-k|z-v|}dv\,dk$$

Then

$$\Phi(R,z) = -2\pi G \int_0^\infty \int_0^\infty \int_{-\infty}^\infty uJ_0(kR)J_0(ku)\rho(u,v)e^{-k|z-v|}dv\,du\,dk$$

The integrations are simpler when the density is separable

$$\rho(R,z) = \sigma_{\mathrm{R}}(R)\rho_{\mathrm{z}}(z)$$

The forces follow by taking the negative derivatives of the potential in the radial and vertical directions.

$$K_{\mathrm{R}}(R,z) = -\frac{\partial\Phi(R,z)}{\partial R} = -2\pi G \int_0^\infty \int_0^\infty \int_{-\infty}^\infty ukJ_1(kR)J_0(ku)\rho(u,v)e^{-k|z-v|}dv\,du\,dk$$

and

$$K_{\mathrm{z}}(R,z) = -\frac{\partial\Phi(R,z)}{\partial z} = -2\pi G \int_0^\infty \int_0^\infty \int_{-\infty}^\infty uJ_0(kR)J_0(ku)\rho(u,v)\mathrm{sign}(z-v)e^{-k|z-v|}dv\,du\,dk$$

STRUCTURE AND DYNAMICS OF GALAXIES

11. Galactic dynamics: Various potentials

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/∼vdkruit

Beijing, September 2011

Outline

The exponential disk

Forces and potential

We saw that Poisson's equation can then be inverted to

$$\Phi(R,z) = -2\pi G \int_0^\infty \int_{-\infty}^\infty J_0(kR)\tilde{\rho}(k,v)e^{-k|z-v|}dv\,dk$$

Then

$$\Phi(R,z) = -2\pi G \int_0^\infty \int_0^\infty \int_{-\infty}^\infty uJ_0(kR)J_0(ku)\rho(u,v)e^{-k|z-v|}dv\,du\,dk$$

The integrations are simpler when the density is separable

$$\rho(R,z) = \sigma_{\mathrm{R}}(R)\rho_{\mathrm{z}}(z)$$

The forces follow by taking the negative derivatives of the potential in the radial and vertical directions.

$$K_{\mathrm{R}}(R,z) = -\frac{\partial \Phi(R,z)}{\partial R} = -2\pi G \int_0^\infty \int_0^\infty \int_{-\infty}^\infty uk J_1(kR) J_0(ku) \rho(u,v) e^{-k|z-v|} dv\, du\, dk$$

and

$$K_{\mathrm{z}}(R,z) = -\frac{\partial \Phi(R,z)}{\partial z} = -2\pi G \int_0^\infty \int_0^\infty \int_{-\infty}^\infty u J_0(kR) J_0(ku) \rho(u,v) \mathrm{sign}(z-v) e^{-k|z-v|} dv\, du\, dk$$

There are various ways of proceeding from here. The first is by taking an analytical form for the density distribution.

Kuijken and Gilmore[1] have done this for exponential disks.

If the radial density distribution is exponential

$$\sigma_{\rm R}(R) = \sigma_0 \exp\left(-R/h\right)$$

then the Hankel transform becomes

$$\int_0^\infty \sigma_0 J_0(ku) u e^{-u/h} du = \frac{\sigma_0 h^2}{(k^2 h^2 + 1)^{3/2}}$$

[1] K. Kuijken & G. Gilmore, MNRAS vol. 239, 571 (1989)

The potential can then be written as

$$\Phi(R,z) = -2\pi G h^2 \int_0^\infty \int_{-\infty}^\infty \frac{J_0(kR)}{(k^2h^2+1)^{3/2}} \rho_z(v) e^{-k|z-v|} dv\, dk$$

First note that if $\rho_z(z)$ is symmetric around $z=0$, then

$$I_z(k,z) = \int_{-\infty}^\infty \rho_z(v) e^{-k|z-v|} dv$$

$$= 2e^{k|z|} \int_0^{|z|} \rho_z(v) \cosh(kv) dv + 2\cosh(kz) \int_{|z|}^\infty \rho_z(v) e^{-kv} dv$$

$$= e^{-k|z|} \int_0^{|z|} \rho_z(v) e^{kv} dv + e^{k|z|} \int_{|z|}^\infty \rho_z(v) e^{-kv} dv + e^{-k|z|} \int_0^\infty \rho_z(v) e^{-kv} dv$$

Kuijken and Gilmore first solve for an exponential z-distribution:

$$\rho_{\mathrm{z}} = \exp\left(-|z|/z_{\mathrm{e}}\right)$$

Solving for this gives

$$\Phi(R,z) = -4\pi G\sigma_0 h^2 z_{\mathrm{e}} \int_0^\infty \frac{J_0(kR)}{(k^2h^2+1)^{3/2}} \frac{e^{-k|z|} - z_{\mathrm{e}} k e^{-|z|/z_{\mathrm{e}}}}{1-k^2z_{\mathrm{e}}^2} dk$$

The possible term for which the denominator is zero ($kz_{\mathrm{e}} = 1$) is still finite; the last quotient is in that case

$$\frac{1}{2z_{\mathrm{e}}k}(1+k|z|)e^{-k|z|}$$

The forces are

$$K_{\rm R}(R,z) = -4\pi G\sigma_0 h^2 z_{\rm e} \int_0^\infty k \frac{J_1(kR)}{(k^2h^2+1)^{3/2}} \frac{e^{-k|z|} - z_{\rm e} k e^{-|z|/z_{\rm e}}}{1-k^2z_{\rm e}^2} dk$$

and

$$K_{\rm z}(R,z) = -4\pi G\sigma_0 h^2 z_{\rm e} \int_0^\infty k \frac{J_0(kR)}{(k^2h^2+1)^{3/2}} \, \mathrm{sign}(z) \frac{e^{-k|z|} - e^{-|z|/z_{\rm e}}}{1-k^2z_{\rm e}^2} dk$$

Next they assume that the density distribution is given by

$$\rho(R,z) = \rho_0 \exp\left(-R/h\right) \mathrm{sech}^n(z/nz_\mathrm{e})$$

For $n = 0$ we have again the exponential z-distribution with vertical, exponential scaleheight z_e.
For $n = 2$ we have the locally isothermal disk[2] and for $n = 1$ the "sech-disk"[3].

Kuijken and Gilmore show that the potential can be written as

$$\Phi(R,z) = -4\pi G \rho_0 h^2 z_\mathrm{e} 2^n \int_0^\infty J_0(kR)(k^2h^2+1)^{-3/2} \times$$

$$\sum_{m=0}^{\infty} \binom{-n}{m} \frac{(1+2m/n)\exp\left(-k|z|\right) - z_\mathrm{e} k \exp\left[-(1+2m/n)|z|/z_\mathrm{e}\right]}{(1+2m/n)^2 - k^2 z_\mathrm{e}^2} dk$$

[2] P.C. van der Kruit & L. Searle, A.&A. 95, 105 (1981)
[3] P.C. van der Kruit, A.&A. 192, 117 (1988)

The possible term, for which $m = n(kz_{\rm e} - 1)/2$, has a zero denominator and must be written as

$$\frac{1}{2z_{\rm e}k}\binom{-n}{m}(1+k|z|)e^{-k|z|}$$

The binomial with the upper coefficient negative can be written as follows

$$\binom{-n}{m} = \frac{(-n)(-n-1)........(-n-m+1)}{m!}$$

$$= (-1)^m \binom{m+n-1}{n-1} = (-1)^m \frac{(m+n-1)!}{(n-1)!m!}$$

So the potential is in this case expressed as a sum of those for exponential z-distributions.

This is essentially related to the fact that the sech is written as a sum of exponentials:

$$\mathrm{sech}\, x = 2 \sum_{j=0}^{\infty} (-1)^j e^{-(2j+1)|x|}$$

This well-known expansion suffers from the fact that it does not work for $x = 0$, because the terms are alternatingly $+1$ and -1.

This does not necessarily make it unsuitable, because after integration each term gets divided by $-(2j+1)$ and the series will converge even for $x = 0$.

However, it may remain slow for small x. For example the sum for $x = 0$

$$2\sum_{j=0}^{\infty}\frac{(-1)^j}{2j+1} = \frac{\pi}{2}$$

takes 32 steps to reach an accuracy of 1%.

Similar expressions as above can be found for the forces, but this will not be fully written out here.

Rotation curves

Casertano[4] has derived an expression for the potential in the plane for an arbitrary density distribution in order to find the rotation curve of a disk with a density distribution derived from surface photometry.

He uses the radial force in the plane and performs the integration over k first (rather than over u).

The equation for the radial force in the plane for a symmetrical z-distribution is

$$K_{\mathrm{R}}(R,0) = -4\pi G \int_0^\infty \int_0^\infty \int_0^\infty ukJ_1(kR)J_0(ku)\rho(u,v)e^{-kv}\,dv\,du\,dk$$

[4] S. Casertano, MNRAS 203, 735 (1983)

It helps to have the same order Bessel functions and get rid of the linear factor k by integrating in parts

$$\int_0^\infty uJ_0(ku)\rho(u,v)du = \frac{u}{k}J_1(uk)\rho(u,v)\Big|_0^\infty - \frac{1}{k}\int_0^\infty uJ_1(uk)\frac{\partial\rho(u,v)}{\partial u}du$$

Then

$$K_{\mathrm{R}}(R,0) = -4\pi G\int_0^\infty\int_0^\infty\int_0^\infty uJ_1(kR)J_1(uk)\frac{\partial\rho(u,v)}{\partial u}e^{-kv}dv\,dk\,du$$

and this can be solved to give

$$K_{\mathrm{R}}(R,0) = 8G\int_0^\infty\int_0^\infty\sqrt{\frac{u}{Rp}}\frac{\partial\rho(u,v)}{\partial u}[K(p)-E(p)]du\,dv$$

where

$$p = x - \sqrt{x^2-1}, \quad x = \frac{R^2+u^2+v^2}{2Ru}$$

K and E are the complete elliptic integrals of the second and first kind respectively for which good approximations are known. For the z-dependence of the density one can take an exponential or the isothermal distribution.

Casertano's work can be extended to the potential, vertical force and the radial force out of the plane. First start with $K_{\rm R}$ at arbitrary z.

At a general position we had

$$K_{\rm R}(R,z) = -2\pi G \int_0^\infty \int_0^\infty \int_{-\infty}^\infty ukJ_1(kR)J_0(ku)\rho(u,v)e^{-k|z-v|}dv\,du\,dk$$

As Casertano we can do the integration over k (after integration by parts) and obtain

$$\int_0^\infty J_1(kR)J_1(uk)e^{-k|z-v|}dk = \frac{(2-p^2)K(p)-2E(p)}{\pi p\sqrt{Ru}}$$

where

$$p = 2\frac{\sqrt{Ru}}{\sqrt{(z-v)^2+(R+u)^2}}$$

This is the same as Casertano found (except that he had $z=0$), but he chose to rework it further to the form above.

The formula for p has a singularity at $R=u=z=0$. Note however that for $R=u=0$ we already have $p=0$ for all z, so that we should take $p=0$ also for $z=0$. Of course this only occurs when evaluating the force in the center.

The radial force now becomes

$$K_{\mathrm{R}}(R,z) = 2G \int_0^{\infty} \int_{-\infty}^{\infty} \frac{(2-p^2)K(p) - 2E(p)}{p\sqrt{Ru}} \frac{\partial \rho(u,v)}{\partial u} \, du \, dv$$

For the vertical force and the potential itself we have a product of Bessel functions of equal order before the integration by parts, but this of different order after that.

When then the integration over k is done, we get expressions which contain the Heuman Lambda function. This can be rewritten only in forms that involve incomplete elliptic integrals of the first and second kind or the elliptic integral of the third kind, but these are much more difficult to evaluate numerically.

Also the integrals over u must then be written as the sum of two different integrals, one from 0 to R and one from R to ∞. So it is better to start with the forms before the integration by parts.

For the vertical force we start with

$$K_z(R,z) = -2\pi G \int_0^\infty \int_0^\infty \int_{-\infty}^\infty uJ_0(kR)J_0(ku)\rho(u,v)\mathrm{sign}(z-v)e^{-k|z-v|}dv\,du\,dk.$$

The integration over k yields

$$\int_0^\infty kJ_0(kR)J_0(ku)e^{-k|z-v|}dk = \frac{|z-v|p^3}{4\pi(1-p^2)\sqrt{(uR)^3}}E(p)$$

and we get

$$K_z(R,z) = -\frac{G}{2}\int_0^\infty \int_{-\infty}^\infty \mathrm{sign}(z-v)\frac{u|z-v|p^3E(p)}{(1-p^2)\sqrt{(uR)^3}}\rho(u,v)dv\,du$$

For the potential we start with

$$\Phi(R,z) = -2\pi G \int_0^\infty \int_0^\infty \int_{-\infty}^\infty u J_0(kR) J_0(ku) \rho(u,v) e^{-k|z-v|} dv\, du\, dk$$

The integration over k now yields

$$\int_0^\infty J_0(kR) J_0(ku) e^{-k|z-v|} dk = \frac{p}{\pi\sqrt{uR}} K(p)$$

The potential then is given by

$$\Phi(R,z) = -2G \int_0^\infty \int_{-\infty}^\infty \frac{upK(p)}{\sqrt{uR}} \rho(u,v) dv\, du$$

Various potentials

There are in the literature many particular potentials that can be used to describe galaxies, but are not isothermal.

The most important ones will be summarized here.

These are **not** solutions of the Liouville and Poisson equation. Rather they are convenient expressions for the potential or density distribution that can be inserted analytically in Poisson's equation.

Plummer, Kuzmin and Toomre models

The **Plummer model** was originally used to describe globular clusters.

The potential has the simple spherical form

$$\Phi(R) = -\frac{GM}{\sqrt{R^2 + a^2}}$$

The corresponding density distribution is

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

The **Kuzmin model** derives from the potential

$$\Phi(R,z) = -\frac{GM}{\sqrt{R^2 + (a + |z|)^2}}$$

This is an axisymmetric potential that can be used to describe very flat disks.

The corresponding surface density is

$$\sigma(R) = \frac{aM}{2\pi(R^2 + a^2)^{3/2}}$$

The **Toomre models** derive from the Kuzmin model by differentiating with respect to a^2.

The n-th model follows after $(n-1)$ differentiations:

$$\sigma_{\rm n}(R) = \sigma(0)\left(1 + \frac{R^2}{4n^2a^2}\right)$$

The corresponding potential can be derived by differentiating the potential an equal number of times.

It can be seen that Toomre's model 1 (which has $n = 1$) is Kuzmin's model.

The limiting case of $n \rightarrow \infty$ becomes a Gaussian surface density model.

Logarithmic potentials

These are made to provide rotation curves that are not Keplerian for large R.

Since these can be flattened they provide an alternative to the simple isothermal sphere. The potential is

$$\Phi(R,z) = \frac{V_\circ^2}{2} \ln\left(r_\circ^2 + R^2 + \frac{z^2}{c^2}\right)$$

$V_\circ$ is the rotation velocity for large radii and c controls the flattening of the isopotential surfaces ($c \leq 1$).

The density distribution is

$$\rho(R,z) = \frac{V_\circ^2}{4\pi G c^2} \frac{(2c^2+1)r_\circ^2 + R^2 + 2z^2[1-1/(2c^2)]}{(r_\circ^2 + R^2 + z^2/c^2)^2}$$

At large radii $R \gg r_\circ$ the isodensity surfaces have a flattening

$$\left(\frac{b}{a}\right)^2 = c^4(2-c^{-2})$$

In the inner regions $R \ll r_\circ$ it is

$$\left(\frac{b}{a}\right)^2 = \frac{1+4c^2}{2+3c^{-2}}$$

The rotation curve is

$$V_{\rm rot} = \frac{V_\circ R}{\sqrt{r_\circ^2 + R^2}}$$

Oblate spheroids

Assume that all iso-density surfaces are confocal ellipsoids with axis ratio c/a and therefore excentricity

$$e = \sqrt{1 - \frac{c^2}{a^2}}$$

Let the density along the major axis be $\rho(R)$. Define

$$\alpha(R, z) = R^2 + \frac{z^2}{1 - e^2}$$

The forces and the potential can then be calculated. I will not treat the full derivation[5], but simply list the equations.

[5] See Binney & Tremaine, section 2.5

Inside the spheroid the forces and potential are

$$K_{\mathrm{R}} = -\frac{4\pi G\sqrt{1-e^2}}{e^3} R \int_0^{\sin^{-1} e} \rho(\alpha) \sin^2\beta d\beta$$

$$K_{\mathrm{z}} = -\frac{4\pi G\sqrt{1-e^2}}{e^3} z \int_0^{\sin^{-1} e} \rho(\alpha) \tan^2\beta d\beta$$

$$\Phi(R,z) = \frac{4\pi G\sqrt{1-e^2}}{e} \left[\int_0^{\delta} \rho(\alpha)\alpha\beta d\alpha + \sin^{-1} e \int_{\delta}^{a} \rho(\alpha)\alpha d\alpha\right]$$

Here

$$\delta^2 = R^2 + \frac{z^2}{1-e^2}$$

$$\alpha^2 = \frac{R^2\sin^2\beta + z^2\tan^2\beta}{e^2}$$

Outside the spheroid ($\alpha > a$) we have

$$K_{\mathrm{R}} = -\frac{4\pi G\sqrt{1-e^2}}{e^3} R \int_0^{\gamma} \rho(\alpha) \sin^2 \beta d\beta$$

$$K_{\mathrm{z}} = -\frac{4\pi G\sqrt{1-e^2}}{e^3} z \int_0^{\gamma} \rho(\alpha) \tan^2 \beta d\beta$$

$$\Phi(R,z) = \frac{4\pi G\sqrt{1-e^2}}{e} \int_0^{a} \rho(\alpha)\alpha\beta d\alpha$$

Here γ follows from

$$R^2 \sin^2 \gamma + z^2 \tan^2 \gamma = a^2 e^2$$

Infinitesimally thin disks

This is analogous to the treatment of general disk potentials above, but now the vertical distribution is a δ-function.

The equation we had before based on the Hankel-transform was

$$\Phi(R,z) = -2\pi G \int_0^\infty \int_0^\infty \int_{-\infty}^\infty u J_0(kR) J_0(ku) \rho(u,v) e^{-k|z-v|} dv\, du\, dk$$

The potential can be written for the infinitesimally thin disk as

$$\Phi(R,z) = -2\pi G \int_0^\infty \exp\,(-k|z|) J_0(kR) \int_0^\infty \sigma(r) J_0(kr) r\, dr\, dk$$

The rotation velocity then becomes

$$V_{\rm c}^2(R) = -R \int_0^\infty S(k) J_1(kR)\; k\; dk$$

where

$$S(k) = -2\pi G \int_0^\infty J_0(kR)\sigma(R) dR$$

It may be useful to calculate the surface density corresponding to a known rotation curve $V_{\rm c}(R)$.

Using the inversion of the first equation above it can be shown that

$$\sigma(R) = \frac{1}{\pi^2 G}\left[\frac{1}{R}\int_0^R \frac{dV_c^2}{dr} K\left(\frac{r}{R}\right) dr \; + \; \int_R^\infty \frac{1}{r}\frac{dV_c^2}{dr} K\left(\frac{R}{r}\right) dr\right]$$

where K is the complete elliptic integral.

There is a contribution from the part of the disk beyond R.

This also holds for disks with finite thickness as long as the density distribution is not described by spheroids.

In general the rotation curve of a disk depends on the surface density at all radii.

Mestel disk

This has the surface density distribution

$$\sigma(R) = \sigma_\circ \frac{R_\circ}{R}$$

The corresponding rotation curve is flat and has

$$V_\mathrm{c}^2(R) = 2\pi G \sigma_\circ R_\circ = \frac{GM(R)}{R}$$

where $M(R)$ is the mass interior to R.

This is treated in a famous paper by Freeman[6]. The surface density is

$$\sigma(R) = \sigma_\circ \exp\left(-\frac{R}{h}\right)$$

The corresponding potential from the equation above for a infinitessimaly thin disk is

$$\Phi(R,0) = -\pi G \sigma_\circ R \left[I_\circ\left(\frac{R}{2h}\right) K_1\left(\frac{R}{2h}\right) - I_1\left(\frac{R}{2h}\right) K_0\left(\frac{R}{2h}\right) \right]$$

Here I and K are the modified Bessel functions.

[6] K.C. Freeman, Ap.J. 160, 811 (1970)

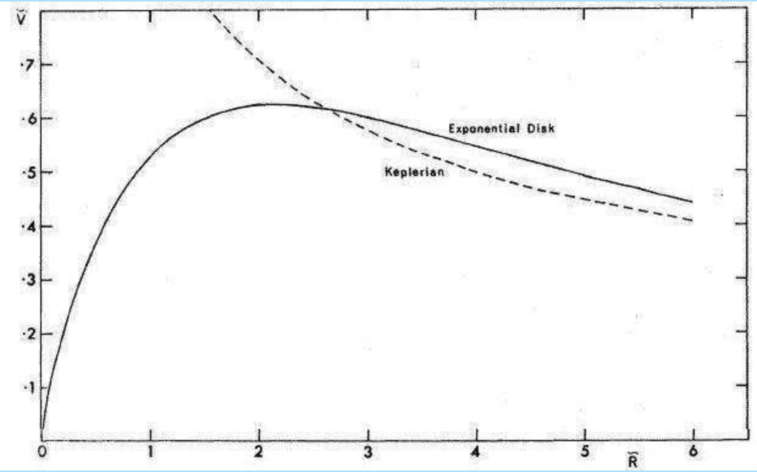
The rotation curve is (again with the equation above for infinitessimally thin disks)

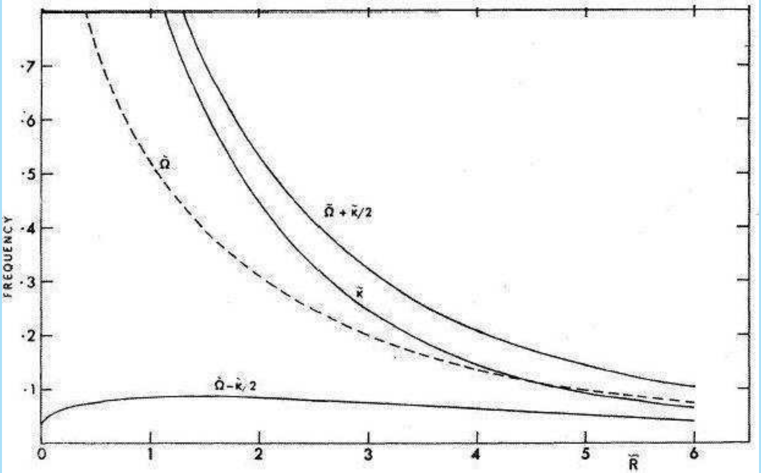
$$V_c^2(R) = 4\pi G\sigma_\circ h \left(\frac{R}{2h}\right)^2 \left[I_0\left(\frac{R}{2h}\right) K_0\left(\frac{R}{2h}\right) - I_1\left(\frac{R}{2h}\right) K_1\left(\frac{R}{2h}\right)\right]$$

The total potential energy of the disk is

$$\Omega \approx -11.6 G\sigma_\circ^2 h^3$$

The rotation curve and the corresponding resonances are shown in the next figures. Note the approximate constancy of $\Omega - \kappa/2$ with radius.


Ṽ
·7
·6
·5
·4
·3
·2
·1
0
1
2
3
4
5
R̃
6
Exponential Disk
Keplerian


FREQUENCY
·7
·6
·5
·4
·3
·2
·1
0
1
2
3
4
5
6
R̃
Ω̃
Ω̃ + κ̃/2
κ̃
Ω̃ − κ̃/2

Stäckel potentials

Stäckel potentials are potentials that can be written as separable functions in ellipsoidal coordinate systems.

They are defined as follows[7].

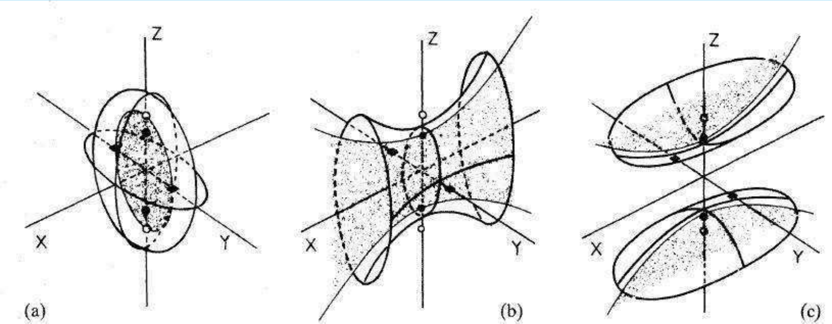
If (x, y, z) is a cartesian coordinate system, then the ellipsoidal coordinates (λ, μ, ν) are the three roots for τ of

$$\frac{x^2}{\tau + \alpha} + \frac{y^2}{\tau + \beta} + \frac{z^2}{\tau + \gamma} = 1$$

where $\alpha < \beta < \gamma$ are three constants.

[7] P.T. de Zeeuw, MNRAS 236, 273 (1985)

The coordinate system is illustrated in the picture below.



I will here only treat the axisymmetric case with oblate density distributions (which means a prolate potential distribution), which applies to disk galaxies[8].

In that case the coordinate system is spheroidal and it can be seen as a further generalisation of the axisymmetric, plane-parallel case, where the potential is separable in R and z.

[8] See also H. Dejonghe & P.T. de Zeeuw, Ap.J. 333, 90 (1988); S.M. Kent & P.T. de Zeeuw, A.J. 102, 1994 (1991)

Coordinate system

The new coordinate system is (λ, ϕ, ν). The relation with the axisymmetric system (r, ϕ, z) is, that λ and ν are the two roots for τ of

$$\frac{r^2}{\tau + \alpha} + \frac{z^2}{\tau + \gamma} = 1$$

with

$$0 \leq \nu \leq \lambda$$

The constants α and γ are sometimes also given in the form

$$\alpha = -a^2, \quad \gamma = -c^2$$

These correspond to a focal distance

$$\Delta = (|\gamma - \alpha|)^{1/2} = (|a^2 - c^2|)^{1/2}$$

Note that λ and ν have a dimension of length^2.

The coordinate surfaces are spheroids for constant λ and hyperboloids for constant ν with the z-axis as rotation axis.

The case for flattened disks obtains, when $-\alpha > -\gamma$, so that $-\gamma = c^2 \leq \nu \leq -\alpha = a^2 \leq \lambda$.

Spheroids of constant λ then are prolate, while the hyperboloids of constant ν have two sheets.

On each meridional plane of constant ϕ we then have elliptical coordinates (λ, ν) with foci on the z-axis at $z = \pm\Delta$.

Note that the mass distribution is oblate, although the coordinate system is prolate.

Other relations between the two coordinate systems are

$$r^2 = \frac{(\lambda+\alpha)(\nu+\alpha)}{\alpha-\gamma} \quad ; \quad z^2 = \frac{(\lambda+\gamma)(\nu+\gamma)}{\gamma-\alpha}$$

and

$$\lambda, \nu = \frac{1}{2}(r^2 + z^2 - \gamma - \alpha) \pm \frac{1}{2}\sqrt{(r^2 - z^2 + \gamma - \alpha)^2 + 4r^2 z^2}$$

Also

$$\lambda + \nu = r^2 + z^2 - \alpha - \gamma \quad ; \quad \lambda\nu = \alpha\gamma - \gamma r^2 - \alpha z^2$$

Note that ν and λ occupy different, but contiguous parts of the positive real line.

- In the plane we have $\nu = -\gamma$, $\lambda = r^2 - \alpha$
- On the z-axis
  - $\nu = z^2 - \gamma$, $\lambda = -\alpha$ for $0 \leq |z| \leq \Delta$
  - $\nu = -\alpha$, $\lambda = z^2 - \gamma$ for $|z| \geq \Delta$.

The potential and the density distribution

Suppose that the potential Φ, which is minus the usual potential Φ and therefore always positive, can be separated as follows

$$\Phi(\lambda,\nu) = \frac{(\lambda+\gamma)G(\lambda) - (\nu+\gamma)G(\nu)}{\lambda-\nu}$$

Such potentials are called (axi-symmetric) Stäckel potentials.

For models with a finite mass M the potential should tend to zero for large radii, which means that for $\lambda \to \infty$ we get

$$G(\lambda) \sim \frac{GM}{\lambda^{1/2}}$$

The density ρ, which is defined such that $\rho\ dx\ dy\ dz$ is the mass in the volume element $dx\ dy\ dz$, can be calculated from Poisson's equation, which has the complicated form

$$\pi G\rho(\lambda,\nu)(\nu-\lambda) = (\lambda+\alpha)(\lambda+\gamma)\frac{\partial^2\Phi}{\partial\lambda^2} + \left(\tfrac{3}{2}\lambda+\tfrac{1}{2}\alpha+\gamma\right)\frac{\partial\Phi}{\partial\lambda} -$$

$$(\nu+\alpha)(\nu+\gamma)\frac{\partial^2\Phi}{\partial\nu^2} - \left(\tfrac{3}{2}\nu+\tfrac{1}{2}\alpha+\gamma\right)\frac{\partial\Phi}{\partial\nu}$$

The Kuzmin equation gives the properties, when the density on the z-axis are given:
Assume that this density is $\varphi(\tau)$, where $\tau=\lambda,\nu$ and note from above that on the z-axis we always have $\tau = z^2-\gamma$ for all z.

Then the density is

$$\rho(z) = \varphi(z^2 - \gamma) = \varphi(\tau)$$

Define the primitive function of $\varphi(\tau)$ as

$$\psi(\tau) = \int_{-\gamma}^{\tau} \varphi(\sigma)\, d\sigma$$

Then the density follows from

$$\rho(\lambda, \nu) = \left(\frac{\lambda + \alpha}{\lambda - \nu}\right)^2 \varphi(\lambda) -$$

$$2\frac{(\lambda + \alpha)(\nu + \alpha)}{(\lambda - \nu)^2}\frac{\psi(\lambda) - \psi(\nu)}{\lambda - \nu} + \left(\frac{\nu + \alpha}{\lambda - \nu}\right)^2 \varphi(\nu)$$

The total mass is

$$M = 2\pi \int_{-\gamma}^{\infty} \frac{\sigma + 2\gamma - \alpha}{\sqrt{\sigma + \gamma}} \varphi(\sigma)\ d\sigma = 4\pi \int_{0}^{\infty} (z^2 + \Delta^2) \varphi(z)\ dz$$

The potential follows from

$$G(\tau) = 2\pi G \psi(\infty) - \frac{2\pi G}{\sqrt{\tau + \gamma}} \int_{-\gamma}^{\tau} \frac{\sigma + \alpha}{2(\sigma + \gamma)^{3/2}} \psi(\sigma)\ d\sigma$$

Velocities, angular momentum and integrals of motion

In order to convert velocities we write

$$\cos\Theta = \left[\frac{(\nu+\alpha)(\lambda+\gamma)}{(\alpha-\gamma)(\lambda-\nu)}\right]^{1/2} \quad ; \quad \sin\Theta = \left[\frac{(\lambda+\alpha)(\nu+\gamma)}{(\gamma-\alpha)(\lambda-\nu)}\right]^{1/2}$$

Velocities are related for the oblate mass models $(\gamma-\alpha>0)$ as

$$V_{\rm r} = V_\lambda \cos\Theta - V_\nu \sin\Theta \quad ; \quad \mathrm{sign}\,(z)\;V_{\rm z} = V_\lambda \sin\Theta + V_\nu \cos\Theta$$

and

$$V_\lambda = V_{\rm r}\cos\Theta + \mathrm{sign}\,(z)\;V_{\rm z}\sin\Theta \quad ; \quad V_\nu = -V_{\rm r}\sin\Theta + \mathrm{sign}\,(z)\;V_{\rm z}\cos\Theta$$

Note that V_λ and V_ν are velocities in the local Cartesian system and do *not* describe the changes in λ and ν.

For the momenta we need the coefficients of the coordinate system

$$P^2 = \frac{\lambda - \nu}{4(\lambda + \alpha)(\lambda + \gamma)} \quad ; \quad R^2 = \frac{\nu - \lambda}{4(\nu + \alpha)(\nu + \gamma)}$$

The momenta then are

$$p_\lambda = PV_\lambda, \quad p_\phi = rV_\phi, \quad p_\nu = RV_\nu.$$

The angular momenta are

$$L_{\rm x} = y\dot{z} - z\dot{y} = rV_{\rm z}\sin\phi - z(V_{\rm r}\sin\phi + V_\phi\cos\phi)$$
$$L_{\rm y} = z\dot{x} - x\dot{z} = -rV_{\rm z}\cos\phi + z(V_{\rm r}\cos\phi - V_\phi\sin\phi)$$
$$L_{\rm z} = x\dot{y} - y\dot{x} = rV_\phi$$

The total angular momentum L is

$$L^2 = (r^2 + z^2)V_\phi^2 + (rV_{\rm z} - zV_{\rm r})^2$$

Integrals of motion

It can then be shown that there are three integrals of motion, namely

$$I_1 = E = -\left(\frac{p_\lambda^2}{2P^2} + \frac{p_\phi^2}{2r^2} + \frac{p_\nu^2}{2R^2} \right) + \Phi(\lambda, \nu)$$

$$I_2 = \tfrac{1}{2} L_{\mathrm{z}}^2$$

$$I_3 = \tfrac{1}{2}(L_{\mathrm{x}}^2 + L_{\mathrm{y}}^2) + (\gamma - \alpha) \left[\tfrac{1}{2} V_{\mathrm{z}}^2 - z^2 \frac{G(\lambda) - G(\nu)}{\lambda - \nu} \right]$$

The equations of motion then are

$$p_\lambda^2 = \frac{1}{2(\lambda+\alpha)}\left[G(\lambda) - \frac{I_2}{\lambda+\alpha} - \frac{I_3}{\lambda+\gamma} - E\right]$$

$$p_\phi^2 = 2I_2$$

$$p_\nu^2 = \frac{1}{2(\nu+\alpha)}\left[G(\nu) - \frac{I_2}{\nu+\alpha} - \frac{I_3}{\nu+\gamma} - E\right]$$

In the meridional plane the orbits are restricted to the area defined by

$$-\gamma \leq \nu \leq \nu_0, \quad \lambda_1 \leq \lambda \leq \lambda_2$$

where the turning points ν_0, λ_1 and λ_2 are the values for ν and λ for which respectively V_ν and V_λ are zero.

The case $\nu = -\gamma$ corresponds to $z = 0$.

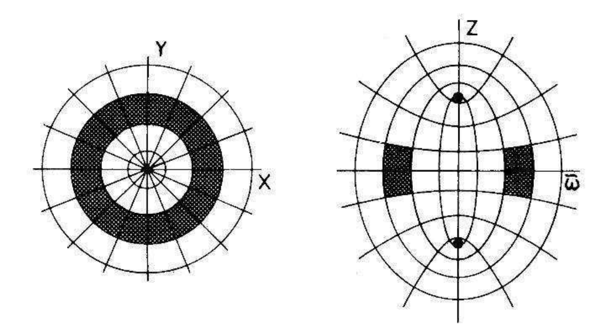
The turning points are the three solutions $\tau_1 \leq \tau_2 \leq \tau_3$ of

$$G(\tau) - \frac{I_2}{\tau + \alpha} - \frac{I_3}{\tau + \gamma} - E = 0$$

where in general there should be

- one solution $\tau_1 \leq -\alpha$, which is ν_0, and
- two solutions $-\alpha \leq \tau_2 \leq \tau_3$, which are λ_1 and λ_2.

In the case of an oblate mass distribution (prolate coordinate system) all orbits are "short axis tubes", bounded by two prolate spheroids and one hyperboloid of one sheet.



STRUCTURE AND DYNAMICS OF GALAXIES

12. Luminosity distributions: Bulges and disks

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/∼vdkruit

Beijing, September 2011

Outline

Luminosity distributions

- Bulge luminosity laws
- Luminosity distributions in disks

Component separation

- Moderately inclined spirals
- Edge-on spirals

Luminosity distributions

Bulge luminosity laws

Reynolds[1] made the first fit to the M31-bulge.

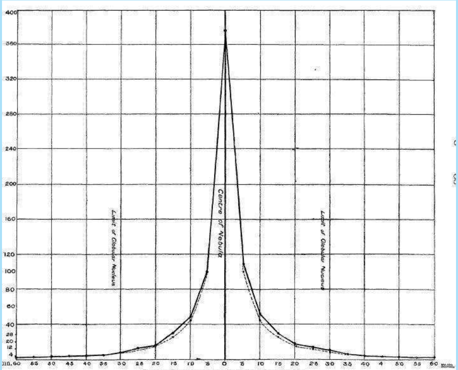
He used the function:

$$(x+1)^2 y = \mathrm{constant}$$

with x the radial distance and y the "light ratio" (relative surface brightness on a linear scale).

He went out to only 6.9 arcmin ($\sim$ 1.4 kpc). At this radius the surface brightness is 21 B-mag arcsec^{-2}.

[1] H.H.Reynolds, MNRAS. 74, 132 (1913)



Light Curve of the Andromeda Nebula.

Hubble used this later in the form:

$$I(R) = I_{\circ}(R + a)^{-2}$$

The most commonly used fitting function is the so-called $R^{1/4}$-law found empirically by de Vaucouleurs[2].

$$\log \frac{I(R)}{I_\mathrm{e}} = -3.3307 \left[\left(\frac{R}{R_\mathrm{e}} \right)^{1/4} - 1 \right]$$

$$R_\mathrm{e} = \mathrm{Effective\ radius}$$

$$\mu(0) = \mu_\mathrm{e} + 8.3268$$

$$L = 7.215 \pi I_\mathrm{e} R_\mathrm{e}^2 (b/a)$$

[2] G. de Vaucouleurs, Ann. d'Astrophys. 11, 247 (1948)

For this there is a numerical deprojection formula from Young[3], which has an approximation for large R (in $L_\odot$ pc^{-3}):

$$L(R) = 52.19 \left(\frac{L}{R_\mathrm{e}}\right)^3 \left(\frac{R}{R_\mathrm{e}}\right)^{-7/8} \cdot \exp\left[-7.67\left(\frac{R}{R_\mathrm{e}}\right)^{1/4}\right]$$

If flattened $R \rightarrow \alpha = \sqrt{R^2(b/a)^2 + z^2}$.

More physical rather than empirical are the King models[4], which work best for globular clusters and also better for elliptical galaxies than bulges.

[3] P.J. Young, A.J. 81, 807 (1976)
[4] I. King, A.J. 71, 64 (1966)

They are based on isothermal distributions with upper limits on the energy of the particles and are therefore isothermal spheres with a tidal radius.

Jarvis & Freeman[5] introduce also rotation and study the effects of the gravitational effects of the disk.

The starting point is a distribution function, which is a truncated Maxwellian:

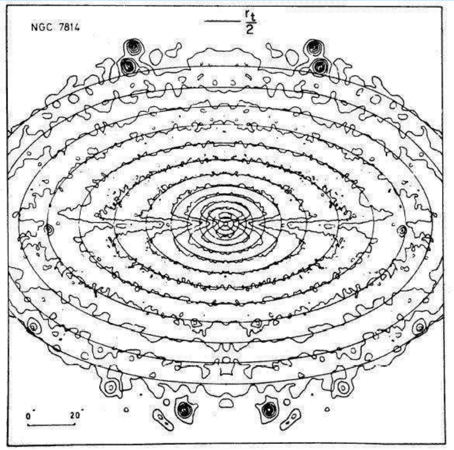
$$f(E, J) = \alpha[\exp(-\beta E) - \exp(\beta E_\circ)] \exp(\gamma J)$$

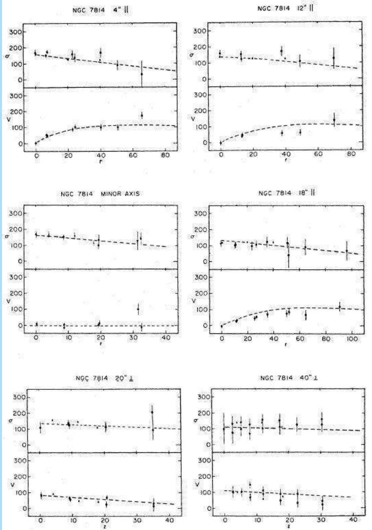
$E \leq E_\circ$ is the energy per unit mass and J the angular momentum parallel to the symmetry axis.

For $\gamma = 0$ we get the King models.

[5] B. Jarvis & K.C. Freeman, Ap.J. 295, 314 and 324 (1986)

Jarvis and Freeman take a constant M/L and include effects of disk potential, and are able to reproduce observations of both isophotes and (stellar) kinematics.

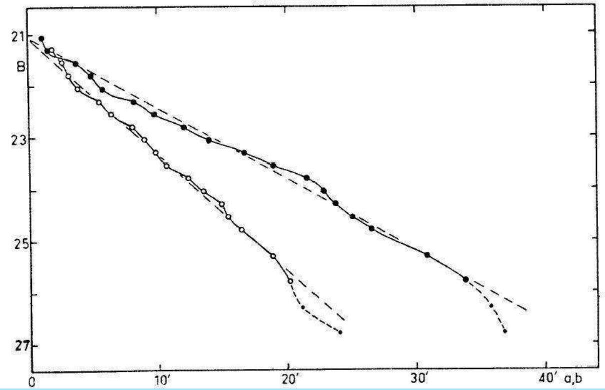




The conclusion is that bulges are consistent with isotropic, oblate spheroids, flattened mostly by rotation.

Luminosity distributions in disks

De Vaucouleurs[6] discovered that radial surface brightness profiles of disks are exponential.



[6] G. de Vaucouleurs, Ap.J. 130, 728 (1959)

A famous paper on exponential disks and the corresponding dynamics is by Freeman[7].
The surface brightness is

$$I(R) = I_{\circ} \exp\left(-R/h\right)$$

in linear units ($L_{\odot}\ \mathrm{pc}^{-2}$).

In magnitudes arcsec^{-2} it is a straight line.

The total luminosity is

$$L = 2\pi h^2 I_{\circ}$$

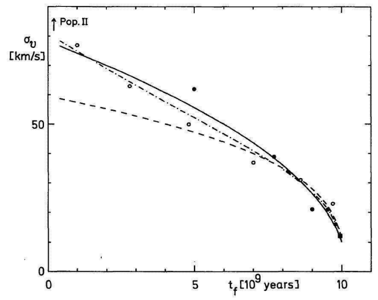
[7] K.C. Freeman, Ap.J. 160, 811 (1970)

Vertical distributions can (away from the dust lane) of the old disk population be approximated with an isothermal sheet.

This is not unreasonable in view of the Age - Velocity dispersion relation[a] of stars in the solar neighborhood.

Star older then a few Gyr have dispersions of the order 50 km $\sec^{-1}$.

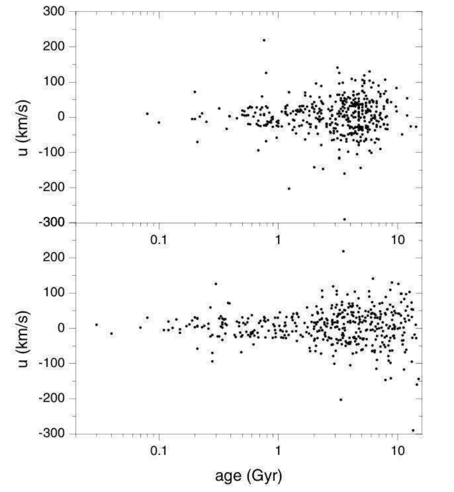
[a] R. Wielen, A.&A. 60, 263 (1977)



With the HIPPARCOS astrometric satellite better data are possible.

Here is a more recent version of the relation.[a]

[a] H. J. Rocha-Pinto et al. A.&A. 423, 517 (2004)



Increase of the u peculiar velocity with age, for uncorrected and corrected chromospheric ages.

The three-dimensional distribution of stars in disks was therefore proposed[8] (with the inclusion of a cut-off radius, so that $R < R_{\rm max}$) as

$$L(R,z) = L(0,0)\,\exp\left(-R/h\right)\,\mathrm{sech}^{\,2}(z/z_\circ)$$

$$I(R) = 2z_\circ L(0,0)\,\exp\left(-R/h\right)$$

$$\langle V_z^2\rangle = \pi G I(R) z_\circ (M/L)$$

[8] P.C. van der Kruit & L. Searle, A.&A. 95, 105 (1981)

For large z-distances:

$$z/z_\circ \gg 1 \;\; \text{then} \;\; \mathrm{sech}\,^2(z/z_\circ) = 4\,\exp\,(-2z/z_\circ)$$

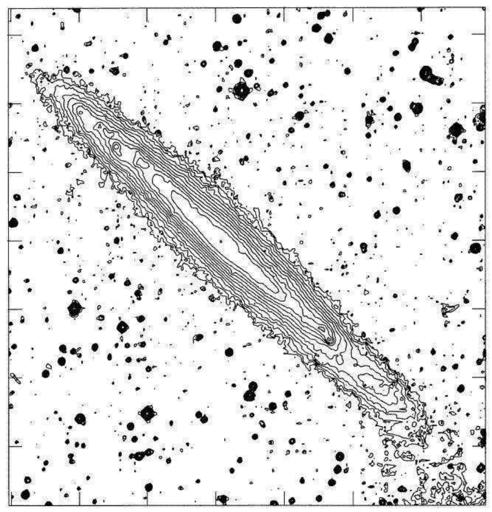
Near the plane:

$$z/z_\circ \ll 1 \;\; \text{then} \;\; \mathrm{sech}\,^2(z/z_\circ) = \; \exp\,(-z^2/z_\circ^2)$$

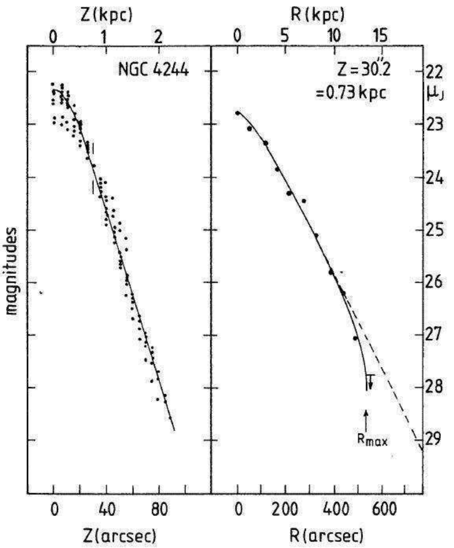
For $R_{\mathrm{max}} \rightarrow \infty$:

$$I(R,z) = 2hL(0,0)(R/h)K_1(R/h)\,\mathrm{sech}\,^2(z/z_\circ)$$

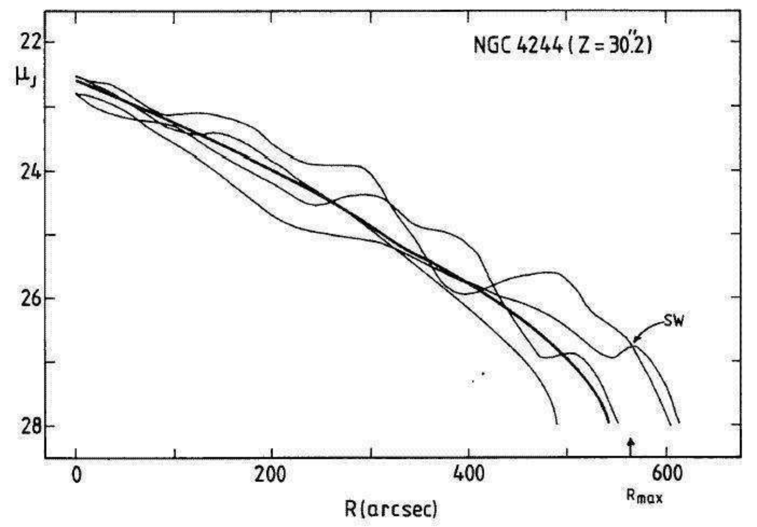
Here is an isophote map of the pure disk, edge-on galaxy NGC 4244.



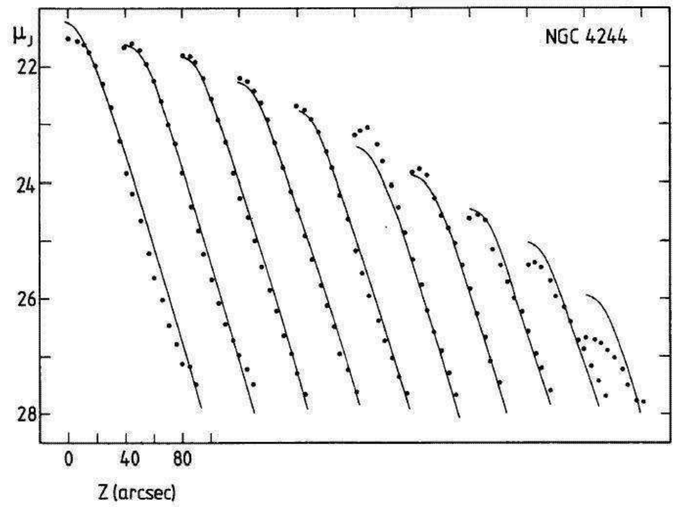
We fit profiles, averaged symmetrically, in z at various R and shifted in coincidence (left) and at a radial profile at a suitable z above the dustlane.



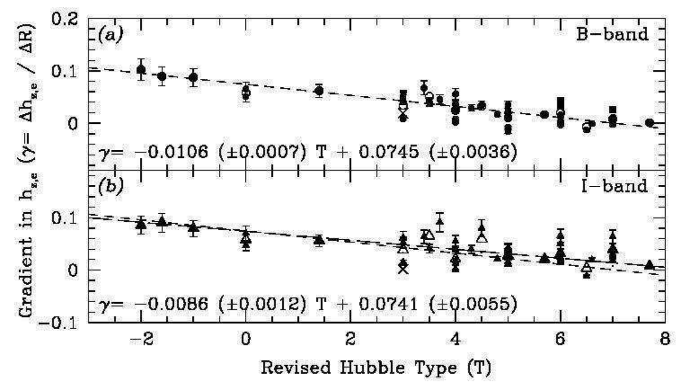
Here is the fit in directions parallel to the major axis.



And parallel to the minor axis.



A closer look at a larger set of edge-on galaxies[9] shows that the constancy of the vertical scaleheight $z_\circ$ does not hold for early type galaxies.

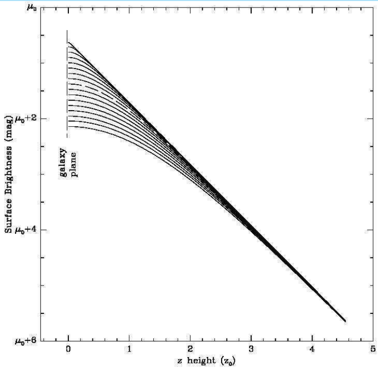


[9]R. de Grijs & R.F. Peletier, A.&A. 320, L21 (1997)

It is unlikely that at moderate and small distances above the plane the stellar population is isothermal.

Therefore a set of functions was proposed to allow for this[a]

$$L(z) = L(0)2^{-2/n} \,\mathrm{sech}^{\,2/n}\left(\frac{nz}{2z_\circ}\right)$$



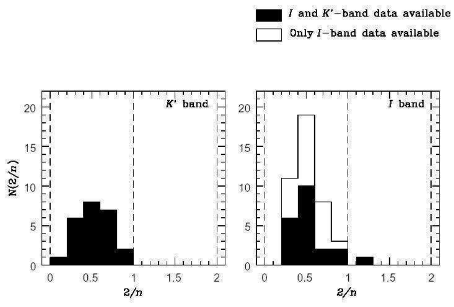
[a] P.C. van der Kruit, A.&A. 192, 117 (1988)

This ranges from the isothermal distribution for $n = 1$ to an exponential for $n = \infty$.

Fits[a] give

$$2/n = 0.54 \pm 0.20$$

in the K-band (2.2 μ).



[a] R. de Grijs, R.F. Peletier & P.C. van der Kruit, A.&A. 327, 966 (1997)

Component separation

Moderately inclined spirals

The usual assumption is to view the galaxy as built up of an exponential disk and an $R^{1/4}$-bulge.

Parameters of the fit then are:

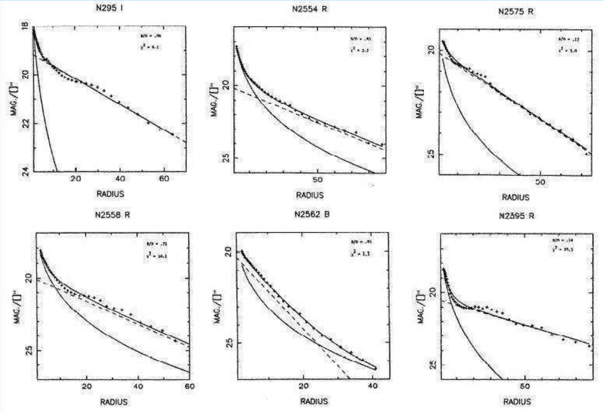
- $\mu_{\rm e}$ and $R_{\rm e}$ for the bulge
- $\mu_\circ$ and h for the disk

This is usually done with some least-squares procedure after a first guess at parameters for the dominant component.

Test on artificial images[10] show that this usually works well.

[10] J.M. Schombert & G.D. Bothun, A.J. 92, 60 (1987)

Here are some actial component separations form Schombert & Bothun.



Edge-on spirals

We now fit to a projected exponential, locally isothermal disk and an $R^{1/4}$-bulge.

Parameters of the fit now are:

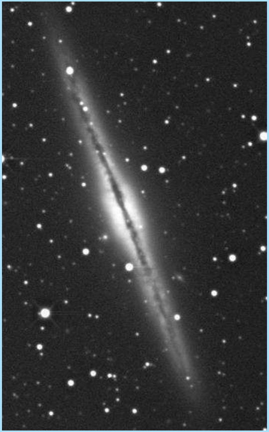
- $\mu_{\circ}$, h and $z_{\circ}$ for the disk
- $\mu_{\rm e}$, $R_{\rm e}$ and b/a for the bulge

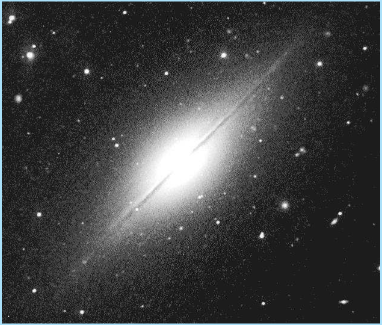
The fit is made first for the dominant component and this is subtracted from the observed distribution.

We look at two examples:

NGC 891[11]. This is an Sb in which the disk dominates the light.

NGC 7814[12]. This is an Sa and the bulge dominates the light.

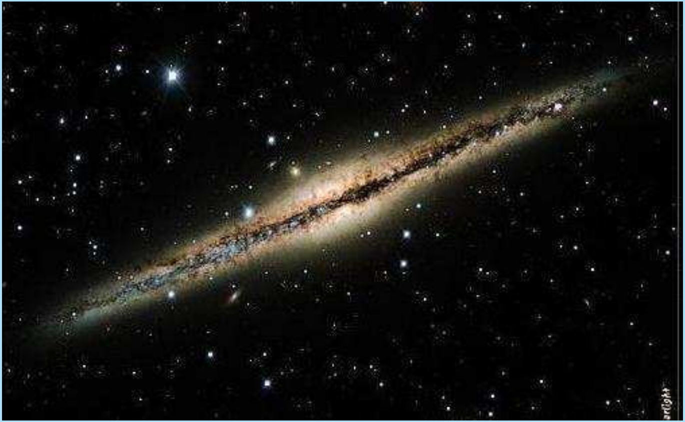




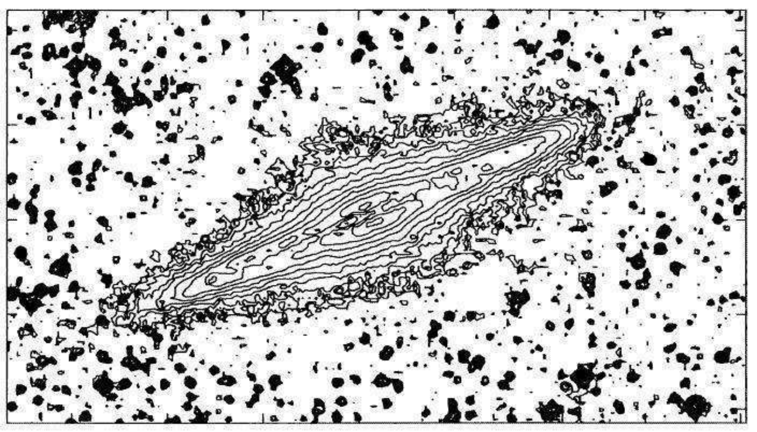
[11] van der Kruit & Searle, A.&A. 95, 116 (1981)

[12] van der Kruit & Searle, A.&A. 110, 79 (1982)

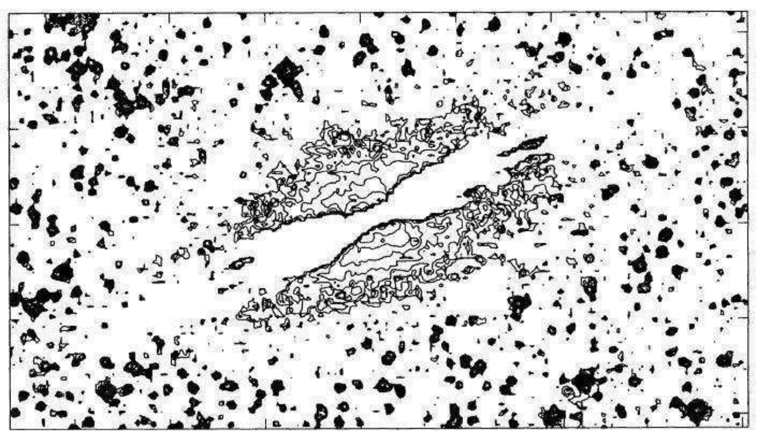
NGC 891 (D = 9.5 Mpc)



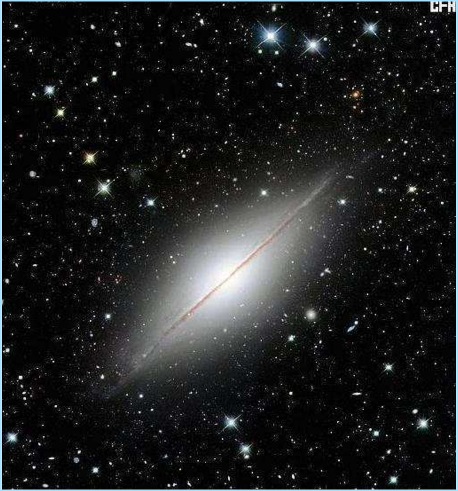
We start with the original image (here the $J \approx B$ band) after "subtraction" foreground stars by interpolation.



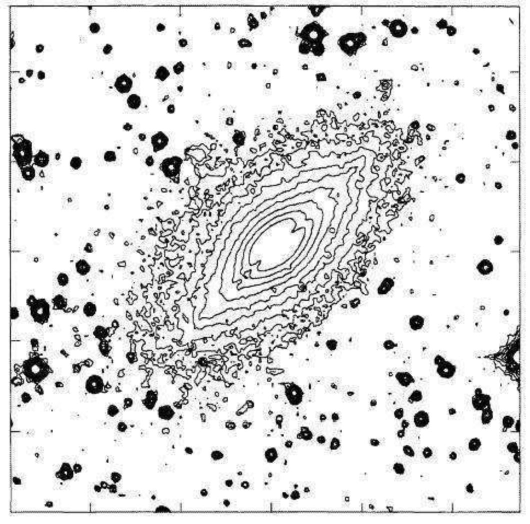
Then we make a fit for the disk from composite R- and z-profiles and subtract this from the data. We then find the bulge brightness distribution.



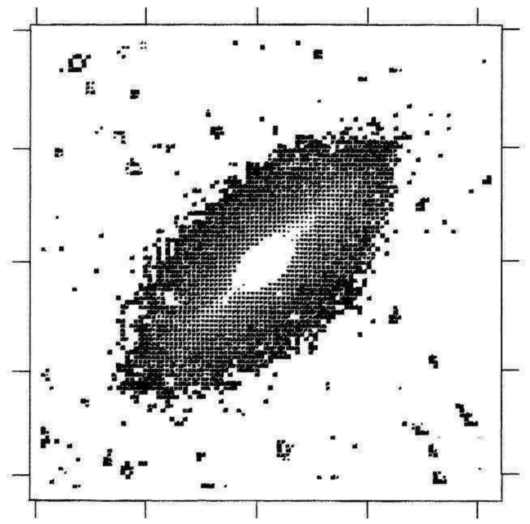
NGC 7814 (D = 15 Mpc)

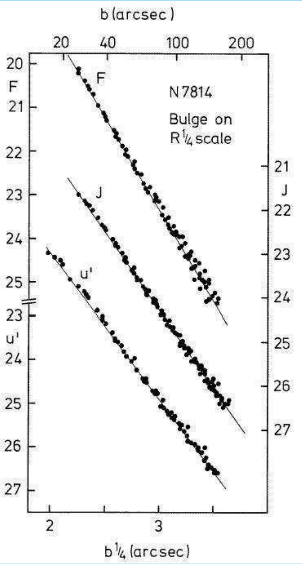


The procedure now is to find a bulge model and subtract that from the observations to reveal the disk.



Note the color change in the bulge (again bluer in the outer parts)[13] .





$$\mu_{U'} = 14.87 + 3.32 b^{1/4}$$

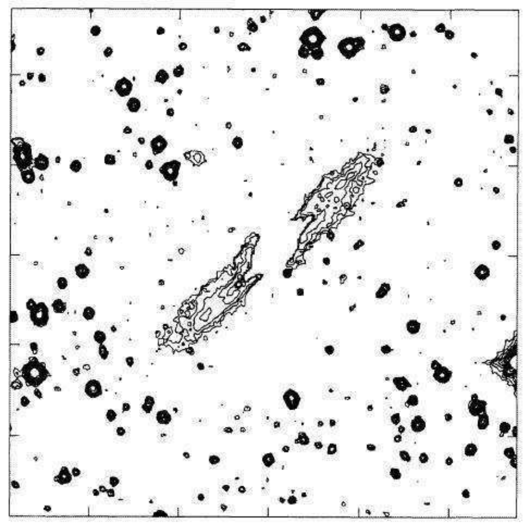
$$\mu_{J\,\mathrm{opt}} = 13.72 + 3.55 b^{1/4}$$

$$\mu_{V} = 13.08 + 3.75 b^{1/4}$$

$$\mu_{F} = 10.70 + 4.20 b^{1/4}$$

$$\mu_{J} = 9.19 + 4.36 b^{1/4}$$

$$\mu_{K} = 8.07 + 4.43 b^{1/4}$$



STRUCTURE AND DYNAMICS OF GALAXIES

13. Luminosity distributions: Parameters

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/∼vdkruit

Beijing, September 2011

Outline

Disk galaxies

Distribution of parameters

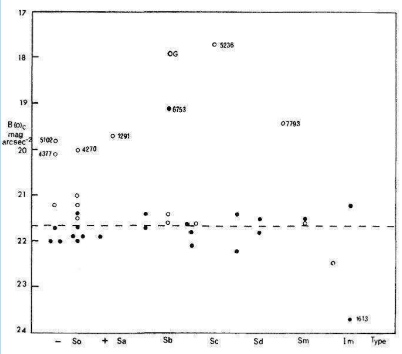
Ken Freeman[1] was the first to study the distribution of properties of exponential disks.

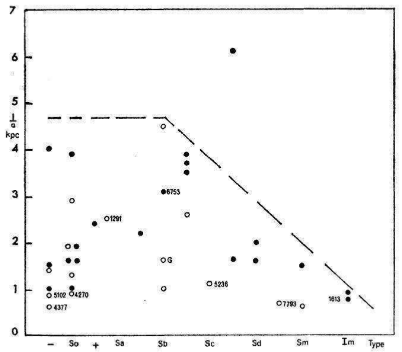
His results are in the following two figures; the small range of (extrapolated) face-on, central surface brightness is known as "Freeman's Law":

$$\mu_\circ = 21.67 \pm 0.30 \ \ \mathrm{B-mag\ arcsec}^{-2}$$

This has generated considerable discussion. The problem is that samples need to be statistically complete and Freeman's sample had serious selection effects.

[1] K.C. Freeman, Ap.J. 160, 811 (1970)


B(0)c mag arcsec-2
17
18
19
20
21
22
23
24
5102
4377
4270
1291
G
6753
5236
7793
1613
–
So
+
Sa
Sb
Sc
Sd
Sm
Im
Type

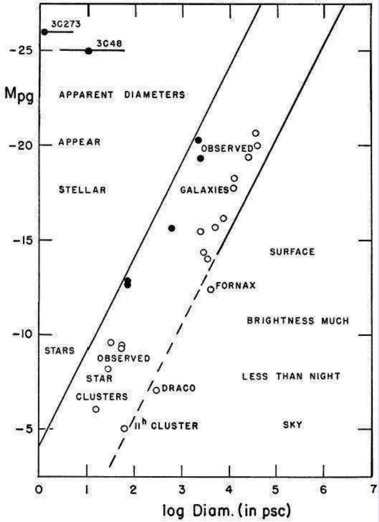

1/α kpc
0
1
2
3
4
5
6
7
6753
1291
G
5236
5102
4270
4377
7793
1613
–
So
+
Sa
Sb
Sc
Sd
Sm
Im
Type

Selection effects

The selection was discussed first by Arp [a].

We see that there is a narrow band in this diagram,excluding objects that either are have surface brightnesses that are too faint or that appear stellar.

[a] H.C. Ap.J. 142, 402 (1965).



The selection effects operating here are:

- For a particular luminosity and a faint $\mu_\circ$ we get a large h, but for the most part the object is fainter than sky.
- For the same luminosity and a bright $\mu_\circ$ we get small h and the object will appear starlike.

We will quantify this below.

First we will consider the $V/V_{\rm max}$-test for completeness.

For this we need to know the selection criteria of the sample. These could be for example all objects down to a certain angular diameter (at some isophotal level) or integrated apparent magnitude.

Suppose that an object has a distance R. Now shift it in distance untill it drops out of the sample due to the completeness limit and call this distance $R_{\rm max}$.

Then we have V as the volume corresponding to R and $V_{\rm max}$ as the volume relating to $R_{\rm max}$.

Now, in case of a uniform space distribution each object has an uniform chance to be actually located throughout the volume $V_{\rm max}$.

In otherwords, the property $V/V_{\rm max}$ calculated for all objects in the sample should be distributed uniformly over the interval 0 to 1.

Note that $V/V_{\rm max}$ can usually be calculated without knowing the actual distance.

In practice the test is to calculate $\langle V/V_{\rm max}\rangle$. For a compete sample it is required that

$$\langle V/V_{\rm max}\rangle = 0.5.$$

The error in $\langle V/V_{\rm max}\rangle$ is $(12\ n)^{-1/2}$.

This is so, because all numbers between 0 and 1 have an average of 0.5 and a dispersion of $\sqrt{12}$.

Selection and Freeman's law

Mike Disney[2] suggested that Freeman's law is the result of sample selection (and not only of incompleteness).

In the process he also addressed the equivalent for elliptical galaxies, called Fish's law.

The analysis was later extended as in the following[3].

Assume luminosity-law (in linear units)

$$\sigma(R) = \sigma_\circ \exp \; - (R/h)^{1/b}$$

$b = 1$: exponential disk
$b = 4$: $R^{1/4}$ bulge or elliptical galaxy.

[2] M.Disney, Nature 263, 573 (1975)

[3] M. Disney & S. Phillipps, Mon.Not.R.A.S. 205, 1253 (1983); see also J.I. Davies, Mon.Not.R.A.S. 244, 8 (1990)

We then have for the integrated luminosity:

$$L_{\mathrm{tot}} = \int_0^\infty 2\pi R\sigma(R)dR = (2b)!\pi\sigma_\circ h^2$$

a. Diameter selection.

Suppose that a sample is complete for a radius larger than θ_{lim} arcsec at an isophote of μ_{lim} magnitudes arcsec^{-2}. For a radius R and a distance d the angular diameter is $\theta = R/d$ radians.

For clarity we now do the derivation only for an exponential disk.

The disk has an apparent radius

$$R_{\mathrm{app}} = h \ln\left(\frac{\sigma_\circ}{\sigma_{\mathrm{lim}}}\right)$$

In magnitudes arcsec^{-2} this is

$$R_{\rm app} = 0.4 \ln 10 \; h \, (\mu_{\rm lim} - \mu_\circ)$$

With $L = 2\pi\sigma_\circ h^2$ this becomes

$$R_{\rm app} = \frac{0.4 \ln 10}{\sqrt{2\pi}} \left(\frac{L}{\sigma_\circ} \right)^{-1/2} (\mu_{\rm lim} - \mu_\circ)$$

This can be rewritten as

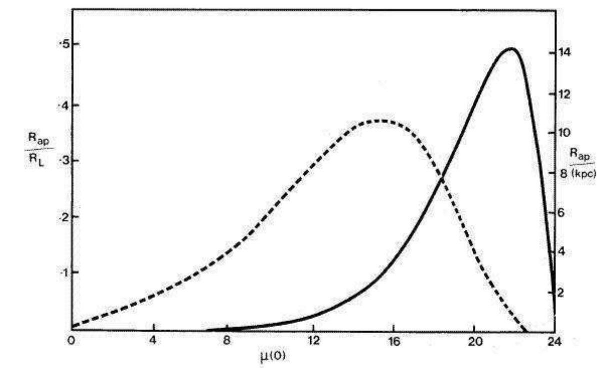
$$R_{\rm app} \sqrt{\frac{\pi\sigma_{\rm lim}}{L}} = \frac{0.4 \ln 10}{\sqrt{2}} 10^{-0.2(\mu_{\rm lim} - \mu_\circ)} (\mu_{\rm lim} - \mu_\circ)$$

The square-root term on the lefthand side is a kind of fiducial radius, that Disney and Phillipps write as $R_{\rm L}$.

The case with $\beta = 4$ for elliptical galaxies is

$$\frac{R_{\rm app}}{R_{\rm L}} = \frac{(0.4 \ln 10)^4}{\sqrt{8!}} 10^{-0.2(\mu_{\rm lim} - \mu_\circ)} (\mu_{\rm lim} - \mu_\circ)^4$$

In the following figure we see the behavior of $R_{\rm app}/R_{\rm L}$ as a function of the central surface brightness $\mu_\circ$ for the case of a diameter selection at an isophote of 24 (B-)magnitudes arcsec^{-2}.



The apparent diameter for exponential disks (full line) peaks at a central surface brightness of $(\mu_{\rm lim} - \mu_\circ) = 2.171$; for elliptical galaxies (dashed line) this occurs at $(\mu_{\rm lim} - \mu_\circ) = 8.686$.

Now when we express surface brightness μ in magnitudes arcsec^{-2} and distances (such as $\sqrt{\sigma/L}$) in parsec we can derive

$$\frac{L}{\sigma_{\rm lim}} = 10^{0.4(\mu_{\rm lim} - M + 5)}$$

Then for the maximum distance for a galaxy to remain in the sample d in parsec and angular radius limit $\theta_{\rm lim}$ in arcsec we get

$$d_{\rm size} = \frac{0.4 \ln 10}{\sqrt{2\pi}} \frac{\mu_{\rm lim} - \mu_\circ}{\theta_{\rm lim}} 10^{0.2(\mu_\circ - M + 5)}.$$

For the general case the result is

$$d_{\rm size} = \frac{(0.4 \ln 10)^b}{\sqrt{\pi (2b)!}} \frac{(\mu_{\rm lim} - \mu_\circ)^b}{\theta_{\rm lim}} 10^{0.2(\mu_\circ - M + 5)}$$

The maximum of d occurs at

$$\mu_{\circ,\max} = \mu_{\rm lim} - \frac{b}{0.2 \ln 10}$$

b. Integrated magnitude selection

Now the sample is supposed complete up to a limiting integrated apparent magnitude $m_{\rm lim}$ within an isophote $\mu_{\rm lim}$.

Assume that the image is overexposed at isophote $\mu_{\rm M}$ to allow for photographic surveys and define

$$s = 0.4 \ln 10(\mu_{\rm M} - \mu_\circ) \;\; ; \;\; p = 0.4 \ln 10(\mu_{\rm lim} - \mu_\circ)$$

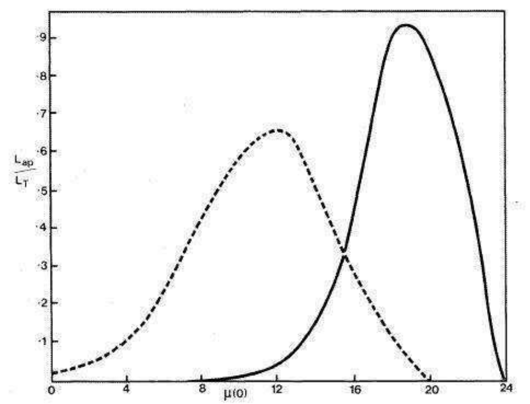
The maximum distance then comes out as

$$d_{\rm magn} = \left[A_{\rm s}{\rm e}^{-s} - A_{\rm p}{\rm e}^{-p} \right]^{1/2} 10^{0.2(m_{\rm lim} - M + 5)}$$

with

$$A_{\rm s} = \sum_{n=0}^{n=2b} \frac{s^n}{n!} \quad ; \quad A_{\rm p} = \sum_{n=0}^{n=2b-1} \frac{p^n}{n!}$$

The following figure below is for a limiting isophote of 24 magnitues arcsec^{-2} and a saturation isophote of 19 magnitudes arcsec^{-2}.



Again we see maxima as for diameter selection.

Note that both diameter and magnitude selection works in favor of disks around Freeman's surface brightness and elliptical systems near Fish's value.

Some actual values: For Palomar Sky Survey:

$\mu_{\rm lim} \approx$ 24 B-mag arcsec^{-2}
$\mu_{\rm M} \approx$ 19 B-mag arcsec^{-2}

Diameter selection: d^3 peaks at:
– 21.8 B-mag arcsec^{-2} for $b = 1$
– 15.3 B-mag arcsec^{-2} for $b = 4$

Magnitude selection: d^3 peaks at:
– 18.5 B-mag arcsec^{-2} for $b = 1$
– 12.0 B-mag arcsec^{-2} for $b = 4$

Observed:
$b = 1$: 21.6 ± 0.3 B-mag arcsec^{-2} (Freeman's law)
$b = 4$: 14.8 ± 0.9 B-mag arcsec^{-2} (Fish's law)

In any catalogue each galaxies has a value for d according to the selection criteria.

If both diameter and magnitude selection play a role the smalles of the two values is the appropriate one.

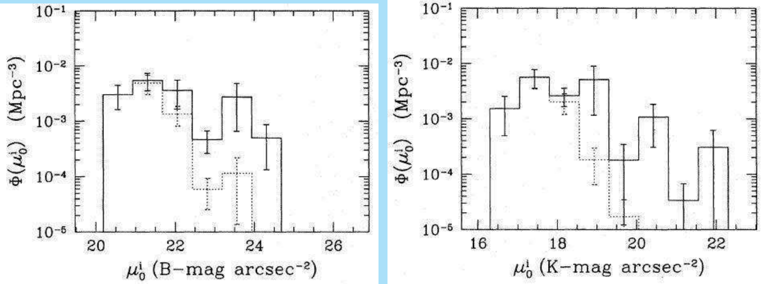
We can then define the visibility as the value for d^3 for each galaxy: in an unbiased sample and a uniform distribution a value of $\mu_\circ$ will occur at a frequency $\propto d^3$.

The equations for the visibility can of course also be used to correct complete sample for the volumes over which galaxies are sampled as a function of their properties in order te obtain space densities as a function of parameters.

This can be used to study the question of the origin of Freeman's law and whether it results from selection effects.

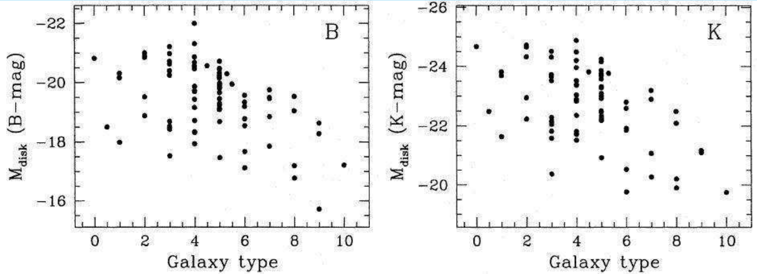
Allen & Shu[4] were the first to suggest that the selection only works at the faint level and that there is only a real upper limit to the central surface brightnesses.

This is confirmed by Roelof de Jong[5], who also confirmed that the faint surface brightness disks are all of late type[6].

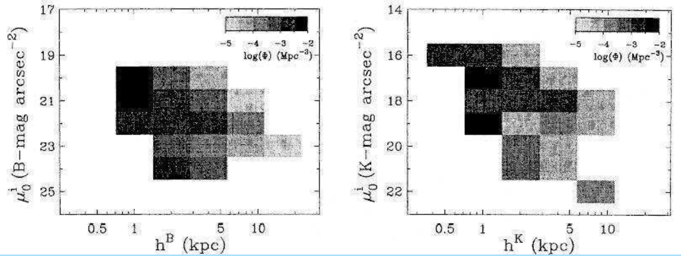


[4]R.J. Allen & F.H. Shu, Ap.J. 227, 67, (1979)
[5]R.S. de Jong, A.&A. 313, 45 (1996)
[6]P.C. van der Kruit, A.&A. 173, 59 (1987)

This is related to the fact that late type galaxies generally have fainter disks.



Data can be combined in bi-variate distribution functions.



From a weighing with the total luminosity it can be estimated that high surface brightness galaxies probably provide the majority of the luminosity density in the universe.

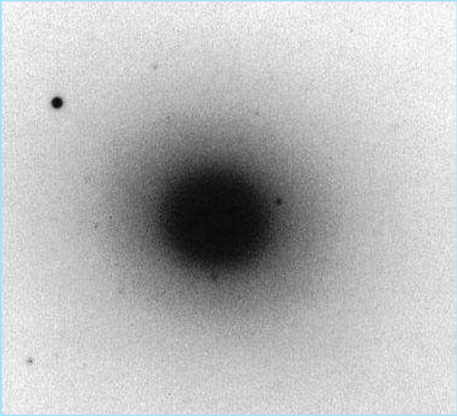
Elliptical galaxies

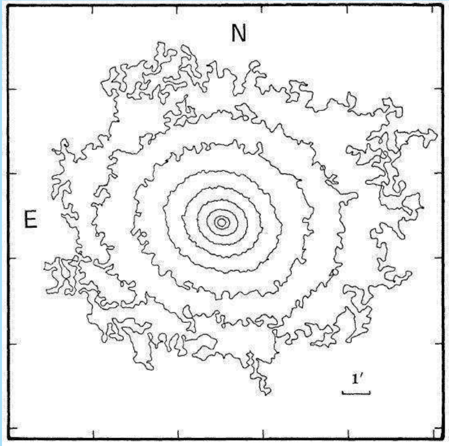
Luminosty distributions

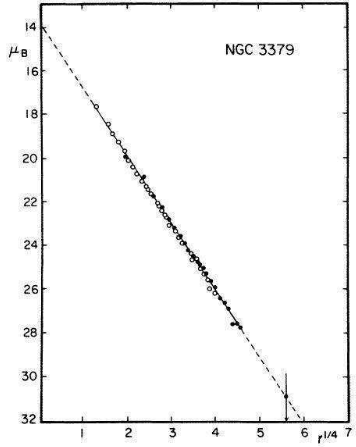
Elliptical galaxies usually conform to the $R^{1/4}$-law and look smooth and regular.

NGC 3379 has been used as a prototype and standard for surface photometry[a].

[a] G. de Vaucouleurs & M. Capaccioli, Ap.J.Suppl. 40, 699 (1979)




N
E
1′


NGC 3379
μ_B
14
16
18
20
22
24
26
28
30
32
1
2
3
4
5
6
7
$r^{1/4}$

Detailed study shows that the isophotal structure of ellipticals is usually much more complicated.

In particular there are isophote twists and deviations from ellipticity.

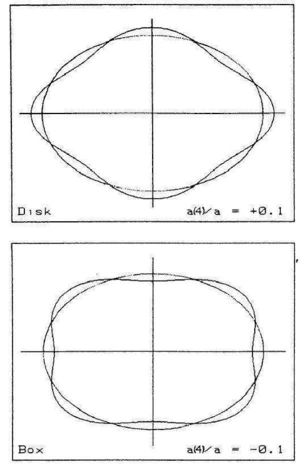
The latter are described by parameters $a(i)$.
These describe the deviations from pure ellipses in multiplicity i[7].
These are derived from Fourier analysis of the isophote shapes relative to the best fitting ellipse.

By definition (because of the ellipse fit) $a(i) = 0$ for $i = 0, 1, 2$.

[7] R. Bender, S. Döbereiner & C. Möllenhoff, A.&A.Suppl. 74, 385 (1988)

The most interesting is $a(4)$, which is negative for "boxy" isophotes and positive for "disky" isophotes.

Here are some examples of non-zero parameters $a(4)$.



We will now look at fits in a boxy galaxy.

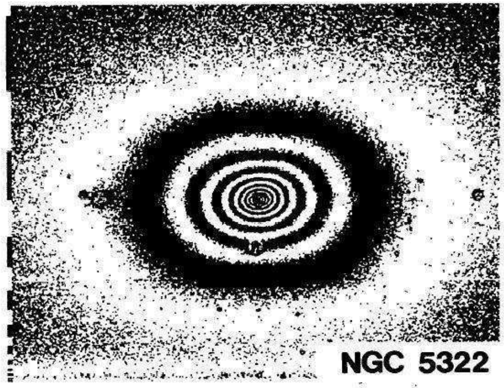
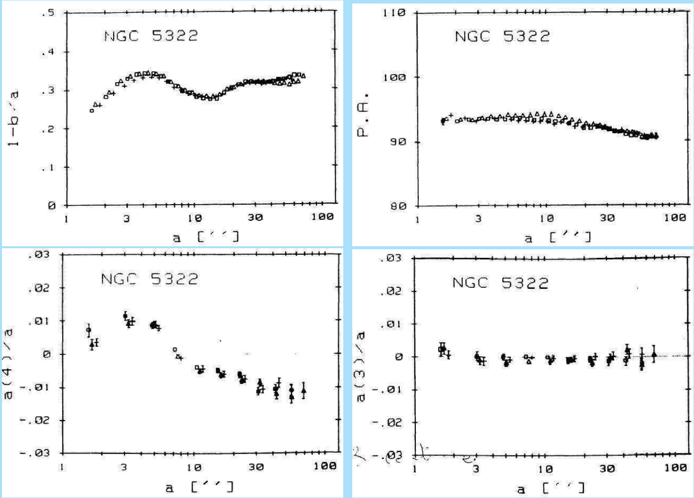


FIGURE 7. — R-image of NGC 5322, an elliptical galaxy with box-shaped isophotes ($a(4)/a \sim -0.01$).


NGC 5322
1-b/a
a ['']
NGC 5322
P.A.
a ['']
NGC 5322
a(4)/a
a ['']
NGC 5322
a(3)/a
a ['']

And here are fits a disky galaxy.

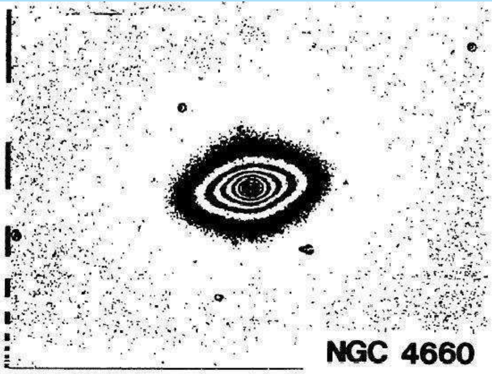
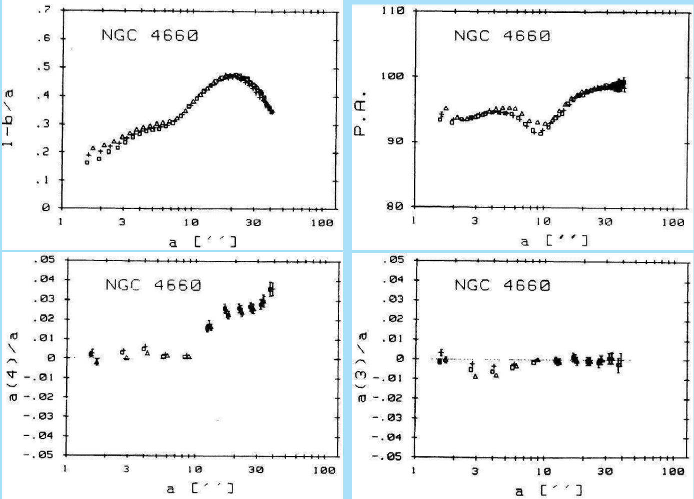


FIGURE 6. — R-image of NGC 4660, an elliptical galaxy with a disk-component in the isophotes ($a(4)/a \sim +0.03$).

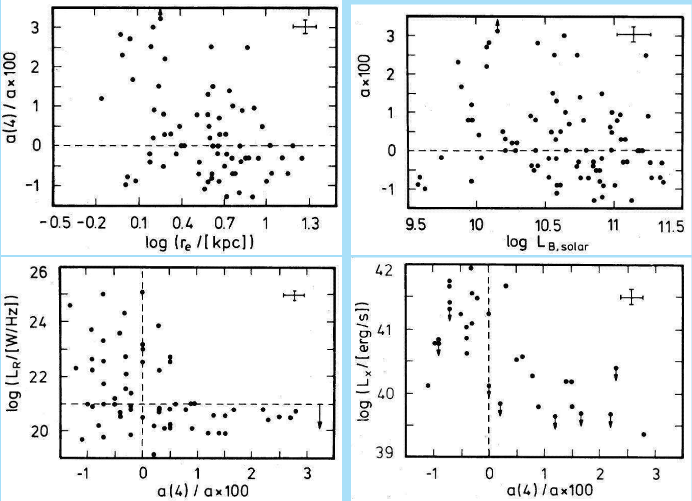

NGC 4660
1-b/a
a [″]
NGC 4660
P.A.
a [″]
NGC 4660
a(4)/a
a [″]
NGC 4660
a(3)/a
a [″]

The global $a(4)$ parameter for a sample of galaxies does not correlate with effective radius or integrated luminosity[8].

However, galaxies with strong radio emission or X-ray halo's are almost always boxy.

It has been suggested that "boxyness" results from interactions.

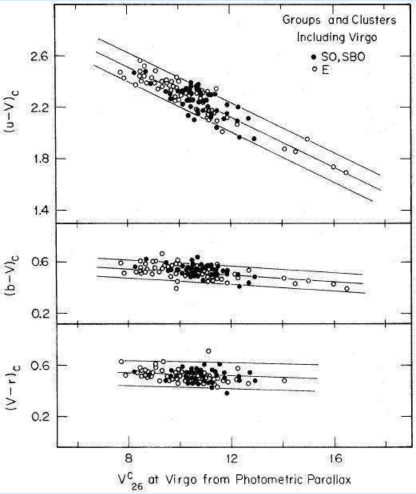
[8] R. Bender, P. Surma, S. Döbereiner, C. Möllenhoff & R. Madejsky, A.&A. 217, 35 (1989)

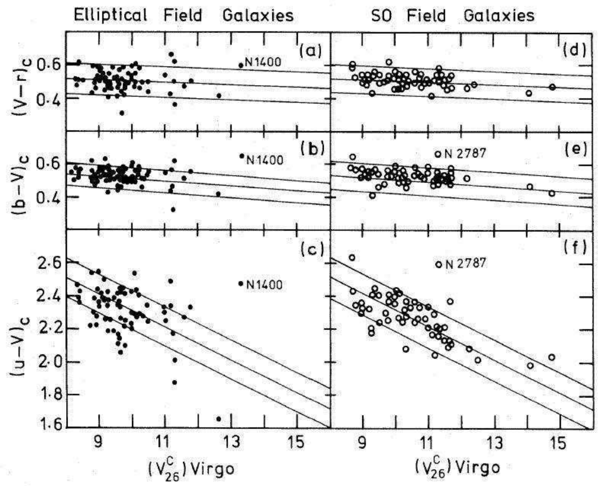

a(4) / a×100
log (r_e /[kpc])
a×100
log L_B,solar
log (L_R/[W/Hz])
a(4) / a×100
log (L_x/[erg/s])
a(4) / a×100

There is a well-defined color – magnitude relation for early-type galaxies[9].
The relation is the same in clusters and in the field.

It is actually one between metallicity and mass (or escape velocity).

[9] A. Sandage, Ap.J. 176, 21 (1972)
N, Visvanathan & A. Sandage, Ap.J. 216, 214 (1977)
A. Sandage & N. Visvanathan, Ap.J. 223, 707 and 225, 742 (1978)


Groups and Clusters
Including Virgo
• SO, SBO
∘ E
$(u-V)_c$
2.6
2.2
1.8
1.4
$(b-V)_c$
0.6
0.2
$(V-r)_c$
0.6
0.2
8
12
16
V^c_{26} at Virgo from Photometric Parallax

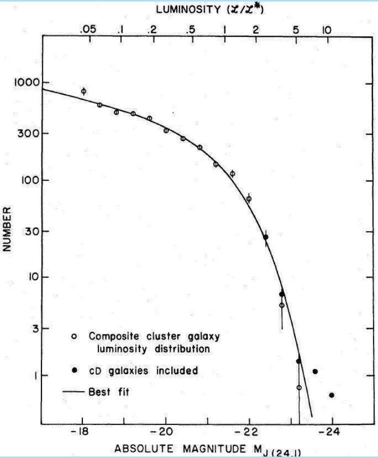

Elliptical Field Galaxies
S0 Field Galaxies
(a)
(b)
(c)
(d)
(e)
(f)
N1400
N 2787
$(V-r)_c$
$(b-V)_c$
$(u-V)_c$
0.6
0.4
2.6
2.4
2.2
2.0
1.8
1.6
9
11
13
15
(V_{26}^C) Virgo

The luminosity function of galaxies is fitted with the Schechter-function[a]

$$\phi(L)dL \propto (L/L^\star)^\alpha \exp\left(-L/L^\star\right) d(L/L^\star)$$

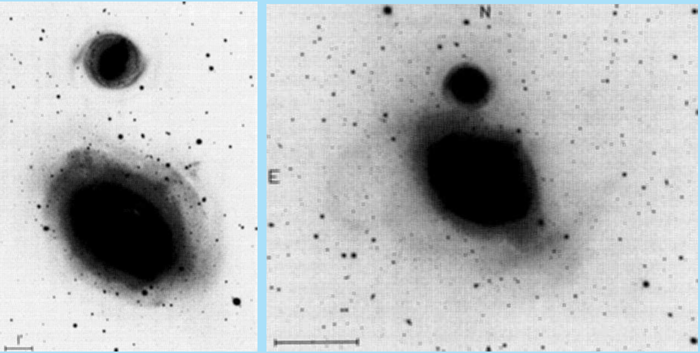
The best fits have $\alpha \sim 1.2$ and $L^\star$ corresponding to $M_{\rm B}^\star \sim -20.6$.

[a] P. Schechter, Ap.J. 203, 297 (1976)

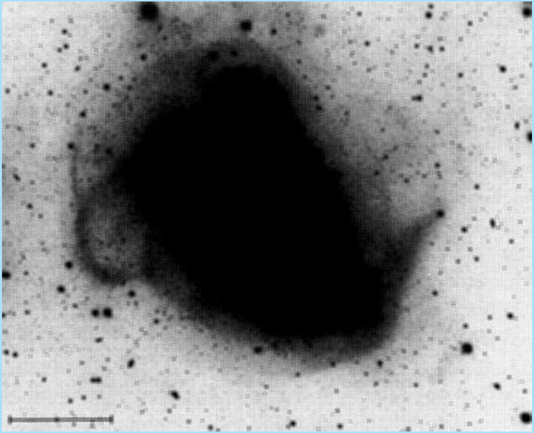


Shells and ripples

In the outer parts faint "shells and ripples" are seen, such as in NGC 1316 = Fornax A[10].



[10] F. Schweizer, Ap.J. 237, 303 (1980)



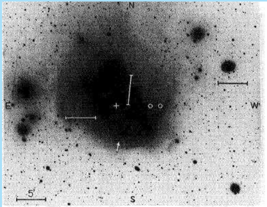
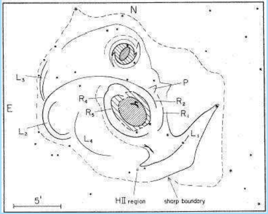


FIG. 1

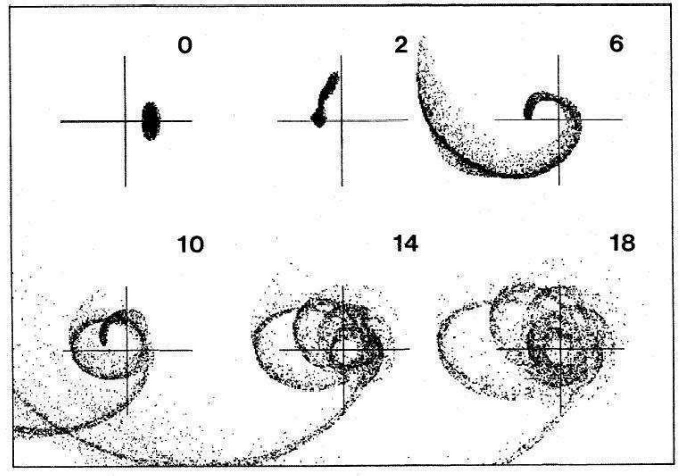


Numerical experiments[11] show that these can be the result of a collission with a disk galaxy.

In the figure on the next frame we see how the disk evolves in the potential of a 100 times more massive elliptical galaxy in a typical encounter.

The unit of time is the circular period at a characteristic radius in the potential.

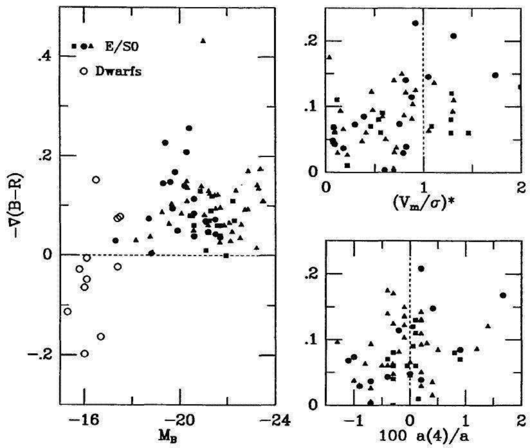
[11] P.J. Quinn, Ap.J. 279, 596 (1984)


0
2
6
10
14
18

Color gradients

Important for formation models is the correlation of color gradients with structural and dynamical properties.

Color gradients usually are defined as the change in color index in magnitudes per decade in radius or
$\bigtriangledown(B-V) = \Delta(B-V)/\Delta(\log r)$.


E/S0
Dwarfs
−∇(B−R)
.4
.2
0
−.2
−16
−20
−24
M_B
.2
.1
0
0
1
2
(V_m/σ)*
.2
.1
0
−1
0
1
2
100 a(4)/a

The property $(V_{\rm m}/\sigma)^*$ is normalised to unity for an isotropic oblate rotator.

- Ellipticals have significant color gradients. The light becomes redder towards the center.
- However, dwarf spheroidals have inverse gradients. This may be due to recent star formation.
- Anisotropic galaxies have smaller gradients.
- Also boxy galaxies tend to have smaller gradients.
- There is no strong correlation between the strength of the color gradient and the luminosity or velocity dispersion.

STRUCTURE AND DYNAMICS OF GALAXIES

14. Photometric evolution

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/~vdkruit

Beijing, September 2011

Outline

Photometric evolution

- Fundamentals
- Analytical models
- Detailed studies
- Schmidt's law for star formation

Population synthesis

Photometric evolution

Fundamentals

The fundamental discussion is by Tinsley[1].

The Initial Mass Function (IMF) is the distribution over stellar masses during star formation.

It is determined in the solar neighborhood independently for low and high mass stars:

- Low masses ($M < 1 M_\odot$) from general distribution of masses of older stars in the disk, since these are all still present.
- High masses ($M > 1 M_\odot$) from distribution of stellar masses in actual clusters and associations.

[1] B.M. Tinsley, Fund. Cosmic Physics 5, 287 (1980)

Normalisation is done such that the two parts join smoothly at $\approx 1 M_\odot$ (continuity constraint).

An usefull analytic form of IMF:

$$\phi(M) = x M_{\mathrm{L}}^{x} M^{-(1+x)} dM$$

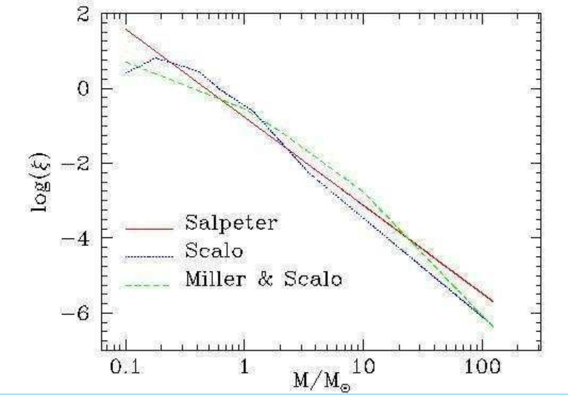
for

$$M_{\mathrm{L}} < M < M_{\mathrm{U}}$$

Usually $M_{\mathrm{L}} = 0.1 M_\odot$ and $M_{\mathrm{U}} = 50 M_\odot$.

The "Salpeter-function" has $x = 1.35$.

Here are some forms of the IMF often used.



Bi-model star formation was proposed by Larson[2]. It says that the two modes of star formation of high- and low-mass stars are independent and normalisation of the IMF must be done separately.

[2] R.B. Larson, Mon. Not. R.A.S. 218, 409 (1986)

The Star Formation Rate(SFR) is the total mass in newly formed stars as a function of time.

In the solar neighborhood it has been roughly constant with time.

It may vary between galaxies, but is usually taken independent of position in a galaxy.

With an IMF and a SFR it is possible to calculate the luminosity and colors of galaxies as a function of time.

This is done by first calculating the photometric evolution of a star clusters by assuming an IMF and using stellar evolution tracks.

In principle this needs to be done for different metal abundances.

These clusters can then be added according to the SFR (and the evolution of metal abundance with time).

Analytical models

Single burst.

First look at the Main Sequence; we have approximately:

$$L \propto M^{\alpha}$$

Rough values for α are 4.9 in U, 4.5 in B and 4.1 in V.

The main-sequence life-time is:

$$t_{\mathrm{MS}} = M^{-\gamma}$$

With M in $M_{\odot}$ the unit of time is $\approx 10^{10}$ years. A good value for γ is 3.

Assume that stars formed all at $t = 0$ and that the total mass is $\psi_\circ$. Then

$$\begin{aligned} L_{\mathrm{MS}}(t) &= \int_{M_{\mathrm{L}}}^{M_{\mathrm{t}}} \psi_\circ M^\alpha \phi(M) dM \\ &= \frac{x}{\alpha - x} M_{\mathrm{L}}^x \psi_\circ M_{\mathrm{t}}^{\alpha - x}, \end{aligned}$$

where

$$M_{\mathrm{t}} = t^{1/\gamma}$$

Now look at the giants. Assume all giants have a luminosity L_{G} and are in that stage for a time t_{G}.

Reasonable values for L_{G} are 35 in U, 60 in B and 90 $L_\odot$ in V and 0.03 for t_{G}.

The number of giants at time t is then

$$\begin{aligned} N_{\rm G}(t) &= \psi_\circ \phi(M_{\rm t}) \left| \frac{dM}{dt_{\rm MS}} \right|_{\rm M=M_t} t_{\rm G} \\ &= \psi_\circ \frac{x}{\gamma} M_{\rm L}^{x} M_{\rm t}^{\gamma - x} t_{\rm G} \end{aligned}$$

Now we can derive the Single Burst luminosity at time t:

$$L_{\rm SB}(t) = L_{\rm MS}(t) + N_{\rm G}(t) L_{\rm G}$$

Using $U_\odot = 5.40$, $B_\odot = 5.25$ and $V_\odot = 4.70$, and $M_{\rm L} = 0.1 M_\odot$, the following table can be calculated.

t	$(U-B)$	$(B-V)$	$(M/L)_{\rm B}$
0.01	-0.34	0.12	0.15
0.03	-0.06	0.45	0.38
0.1	0.18	0.64	1.12
0.3	0.38	0.79	2.79
1	0.56	0.90	6.95
3	0.66	0.96	14.9

Ongoing star formation.

Write the SFR as $\psi(t)$. Then

$$L(t)=\int_0^t \psi(t-t')L_{\rm SB}(t')dt'$$

For two extreme cases we get at $t = 1$:

Model	(U-B)	(B-V)	$(M/L)_{\mathrm{B}}$
Single burst	0.56	0.90	7.0
Constant SFR	-0.25	0.24	1.0

This spans the range of the observed two-color diagram with the single burst corresponding to elliptical and S0 galaxies and the constant SFR for Sc and later types.

Now let us look at some more detailed studies.

Detailed studies

Searle, Sargent & Bagnuolo[3] find the following luministies and colors for sinble bursts a number of slopes of the IMF.

TABLE 1

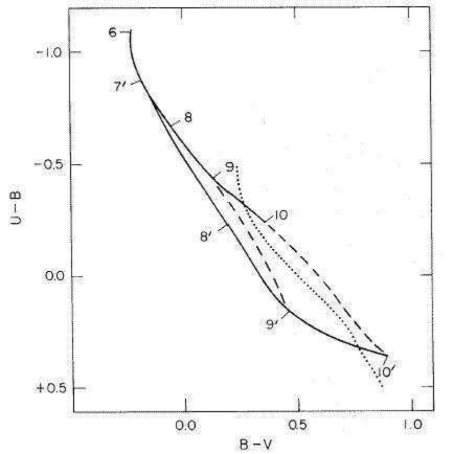
THE BRIGHTNESS AND COLORS OF MODEL STAR CLUSTERS AS A FUNCTION OF AGE

AGE	$\alpha = 2.1$			$\alpha = 2.45$			$\alpha = 3.2$		
(10^7 yrs)	M_V	$B-V$	$U-B$	M_V	$B-V$	$U-B$	M_V	$B-V$	$U-B$
0.1....	−7.2	−0.23	−1.18	−6.6	−0.22	−1.10	−5.6	−0.22	−1.10
0.3....	−7.6	−0.19	−0.96	−6.9	−0.21	−0.96	−5.9	−0.18	−0.96
1.0....	−7.4	−0.15	−0.83	−6.9	−0.18	−0.78	−6.0	−0.14	−0.78
3.0....	−6.2	−0.05	−0.60	−6.0	−0.03	−0.58	−5.6	−0.05	−0.58
10.0....	−5.0	+0.19	−0.22	−5.0	+0.19	−0.23	−5.0	+0.11	−0.23
30.0....	−3.8	+0.21	+0.03	−3.9	+0.34	0.00	−4.4	+0.27	+0.06
100.0....	−2.5	+0.44	+0.12	−2.8	+0.46	+0.16	−3.6	+0.47	+0.22
300.0....	−1.6	+0.66	+0.26	−1.9	+0.67	+0.28	−2.9	+0.68	+0.29
1000.0....	−0.9	+0.89	+0.38	−1.2	+0.90	+0.36	−2.3	+0.89	+0.36

Using this they get a predicted two-color diagram with the Salpeter IMF as in the following figure.

[3] L. Searle, W.L.W. Sargent & W.G. Bagnuolo, Ap.J. 179,427 (1973)

Here numbers x show the location of models of ages 10^x years old; with primes for SB models and unprimed for constant SF. All normal galaxies lie to the right of the dotted line.



Searle *et al.* conclude that the models and observations are consistent with:

- All galaxies $\approx 10^{10}$ years old.
- IMF everywhere similar to local IMF.
- Mean SFR averaged over sufficiently large area's and long times generally declines with time.
- Decay times vary among late-type galaxies; some show bursts, some show uniform SFR.

Larson & Tinsley[4] add:

- Precise form of SFR is not important. Important is only SFR over the last $\approx 10^8$ years to mean SFR over the life of the galaxy.
- Effects of different ages, metallicities and upper stellar masses are small.
- Interacting galaxies show more scatter in two-color diagram. This can be explained with bursts of 5% (fraction of mass to total stellar mass at time of burst; $b \sim 0.05$) and duration $\tau \approx 2 \times 10^7$ years.

[4] R.B. Larson & B.M. Tinsley, Ap.J. 219, 46 (1978)

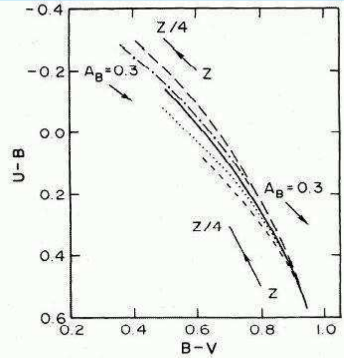
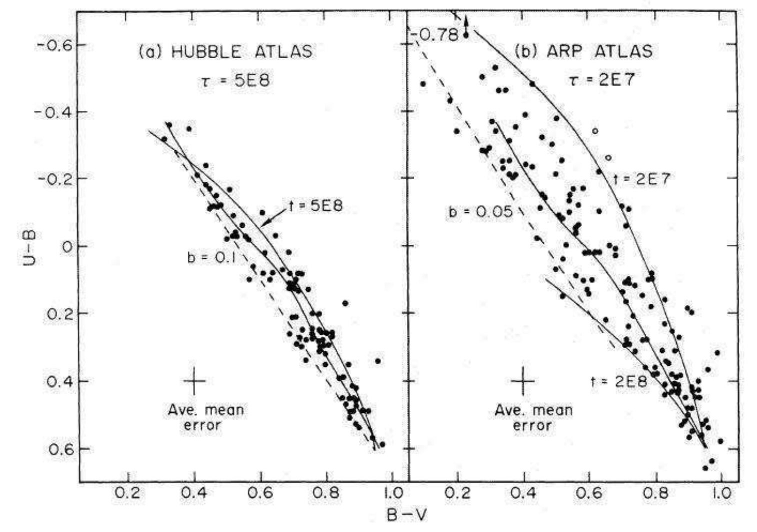


FIG. 7.—Colors of models with monotonic SFRs and age 10^{10} yr. *Heavy line*, local IMF. *Long dashes*, IMF with slope $x = 1$. *Short dashes*, $x = 2$. The foregoing use case T supergiant colors and have an upper mass limit $m_U = 30\ M_\odot$. *Dot-dashes*, $x = 1$, $m_U = 30\ M_\odot$, and case C supergiant colors. *Dots*, $x = 1$, case T, and $m_U = 10\ M_\odot$. The reddening vectors for $A_B = 0.3$ show the RC2 formula for galactic reddening which depends on $B - V$. The other vectors indicate schematically how colors of red and blue galaxies, respectively, may change with a factor 4 reduction in metallicity.


(a) HUBBLE ATLAS
τ = 5E8
t = 5E8
b = 0.1
Ave. mean
error
(b) ARP ATLAS
τ = 2E7
-0.78
t = 2E7
b = 0.05
t = 2E8
Ave. mean
error
U−B
B−V
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.2
0.4
0.6
0.8
1.0

Rob Kennicutt[5] adds the integrated H_α fluxes (in the form of an equivalent width, providing independent information on recent formation of heavy stars.

Equivalent width is the wavelength interval in the continuum that corresponds to as much flux as the line.

His most important results are the following slides:

[5] R.C. Kennicutt, Ap.J. 272, 54 (1983)

- The slope of upper IMF is roughly that of the Salpeter function.

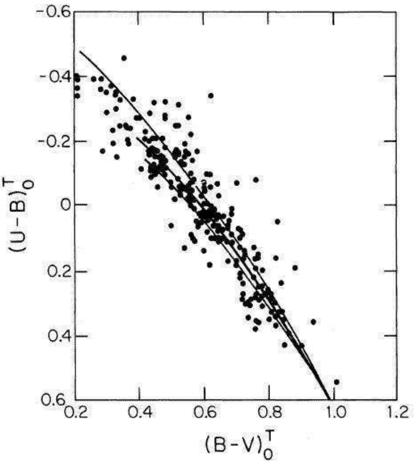


FIG. 3.—Two-color diagram from Shapley-Ames spiral galaxies, along with the model galaxy disk colors described in the text. The three curves correspond to the different mass functions adopted, the Miller and Scalo function (*lowest curve*), the extended Miller-Scalo (i.e., "Salpeter") function, and the shallow m^{-2} IMF (*top curve*).

- Galaxies have same upper mass limit in the IMF ($\approx 50 M_\odot$).

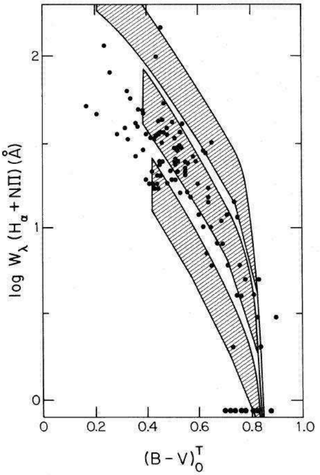


FIG. 4.—Observed emission line equivalent widths and corrected RC2 colors for observed galaxies, along with the evolutionary models. The effect of dust has been shown by plotting each model as an area, as described in the text. The IMFs corresponding to each model are the same as in Fig. 3.

Roelof de Jong[6] derives models to study the color gradients in disks and among different disks.

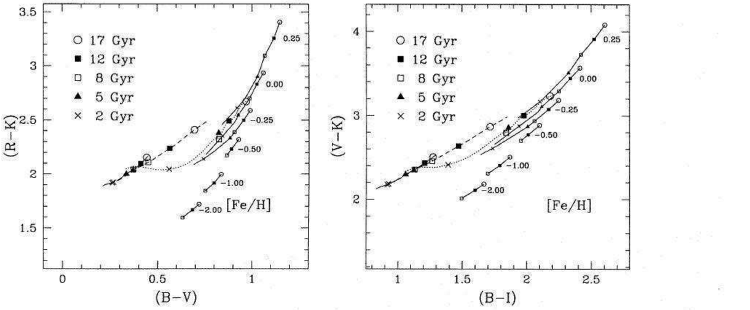


Fig. 6. Evolutionary color–color plots of stellar synthesis models. The symbols indicate the number of years after creation of this population. To the right in each panel, the different ages connected by solid lines, are the single burst models of Worthey (1994) for different metallicities. The corresponding [Fe/H] values are indicated next to them. To the left in each panel are the solar metallicity models of Bruzual & Charlot (1996). The dotted line indicates the single burst evolution. The dashed line is a model with an exponentially declining star formation rate. The leftmost dot-dashed line, overlapping the blue part of the exponentially declining SFR model, indicates a model with constant star formation. Bruzual & Charlot used the Johnson R and I passbands which were here converted to Kron-Cousins R and I passbands using the equations of Bessell (1979).

[6] R.S. de Joing, A.&A. 313, 377 (1996)

His conclusions are:

- Dust reddening plays a minor role.
- Outer parts have lower average ages and are more metal poor than inner parts of disks.
- Late type galaxies ($T \geq 6$)[7] have lower metallicites and younger average ages.

[7] Following de Vaucouleurs himself the de Vaucouleurs types are given numerical values, e.g. T=1 $\rightarrow$ Sa, T=3 $\rightarrow$ Sb, etc. So here is meant types later than Sc.

Schmidt's law for star formation

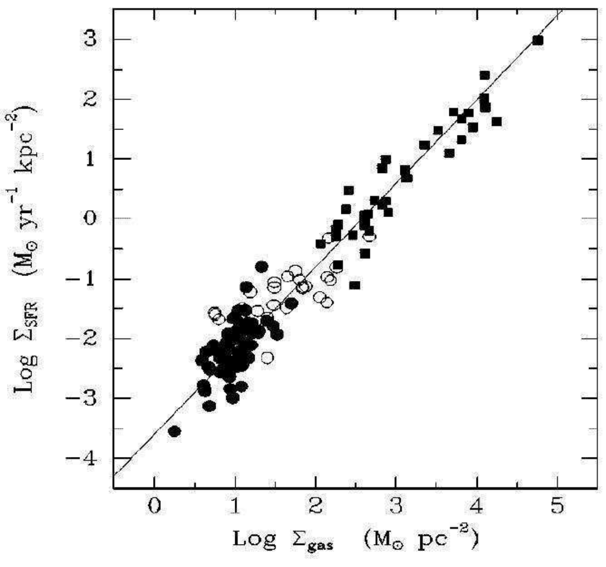
Maarten Schmidt[8] proposed that the star formation rate relates to the gas density as

$$\mathrm{SFR} \propto \rho^2$$

Often this was immediately translated in (observable) surface properties.
The latest result[9] gives

$$\Sigma_{\mathrm{SFR}} = (2.5 \pm 0.7) \times 1^{-4} \left(\frac{\Sigma_{\mathrm{gas}}}{1 M_\odot \mathrm{pc}^{-2}} \right)^{1.4 \pm 0.15} \quad M_\odot \mathrm{year}^{-1} \mathrm{kpc}^{-2}$$

[8] M. Schmidt, Ap.J. 129, 243 (1959)
[9] R.C. Kennicutt, Ann.Rev.Astron.Astrophys. 36, 189 (1998)


3
2
1
0
−1
−2
−3
−4
0
1
2
3
4
5
Log Σ_{SFR} ($M_\odot$ yr^{-1} kpc^{-2})
Log Σ_{gas} ($M_\odot$ pc^{-2})

Population synthesis

Her one attempts to fit the intermediate resolution spectra with those of observed stars.

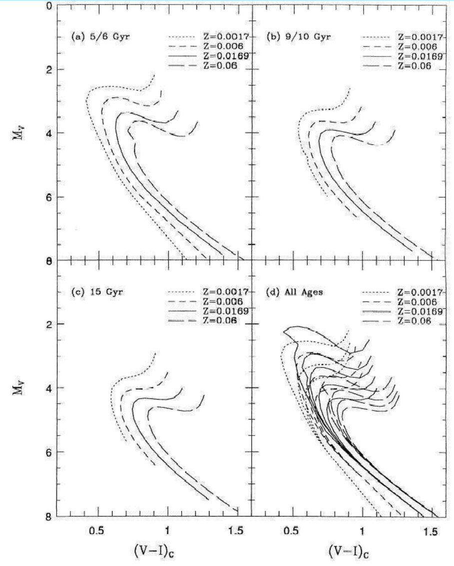
Best method now is by fitting integrated spectra of generations of particular age and metallicity[10].

The steps are the following.:

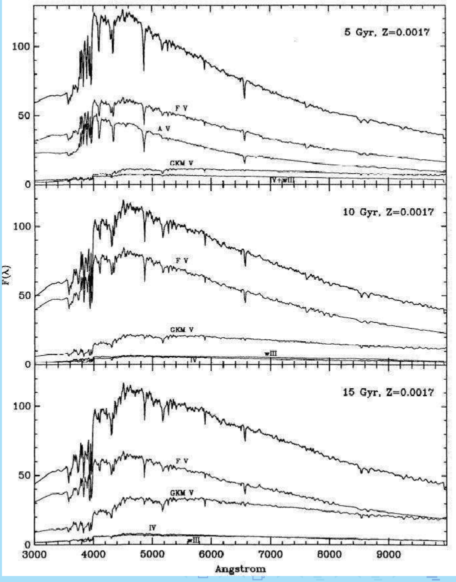
- Measure spectra of stars of various ages and metallicity.
- Synthesize integrated spectra of generations from a set of isochrones.
- Fit using least-squares techniques to galaxy spectra.

[10] For example A.J. Pickles, Ap.J. 296, 340 (1985); Ap.J.Suppl. 59, 33 (1985) or A.J. Pickles & P.C. van der Kruit A.&A.Suppl, 84, 421 (1990) and 91, 1 (1991)

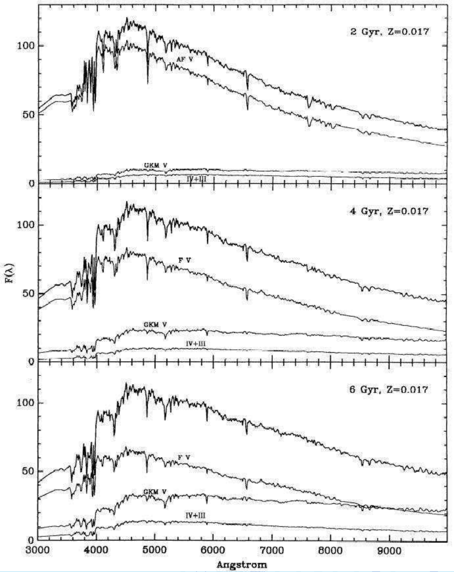
Here is a set of isochrones used by Pickles & van der Kruit.



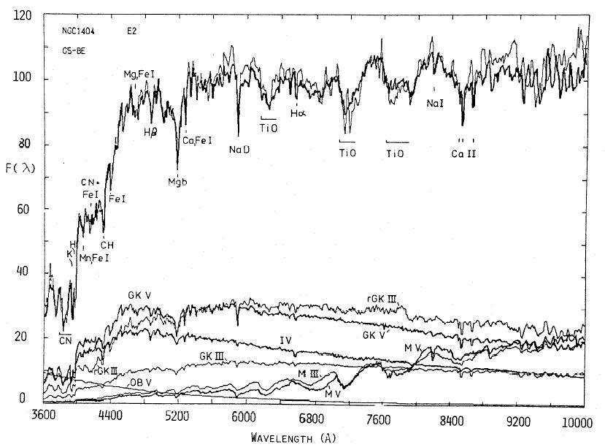
These are synthesized spectra of a metal poor cluster at three ages.



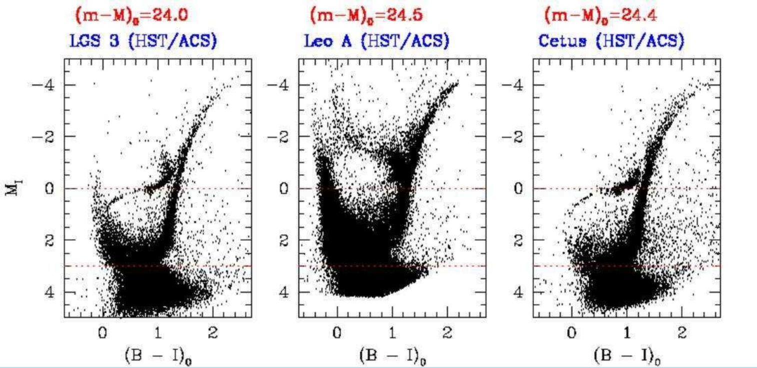
These are synthesized spectra of a metal rich cluster at three ages.



This a an example of a spectrum of an elliptical galaxy fitted by a set of stellar spectra.

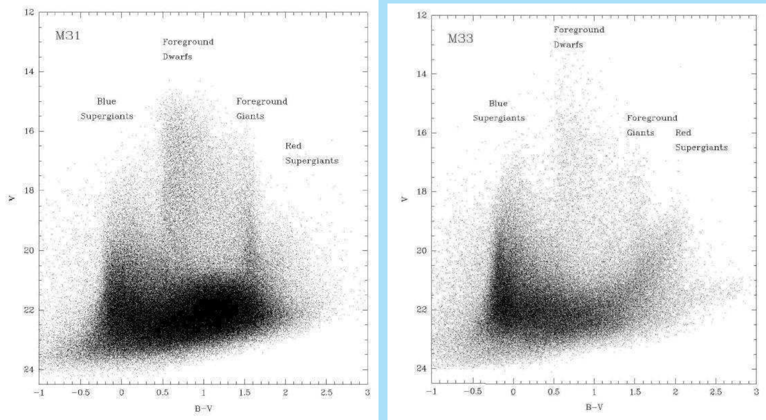


It is now possible to directly observe colour-magnitude diagrams in dwarfs galaxies in the Local Group[11].



[11] See E. Tolstoy, V. Hill & M. Tosi, Ann.Rev.A.&A. 47, 371 (2009) for a review.

Even in M31 and M33 it has been possible now[12].

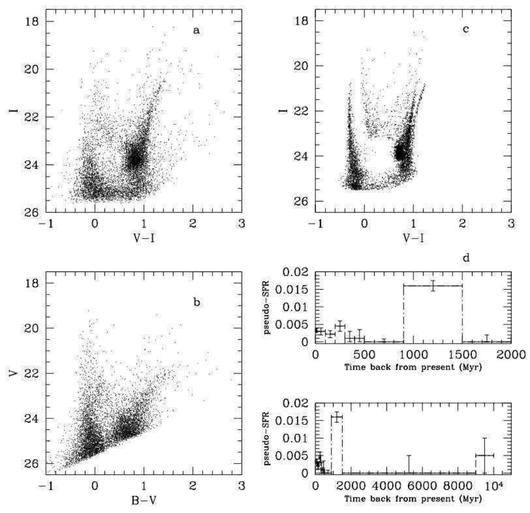


[12] P. Massay et al., A.J. 131, 2486 (2006)

One can use such data to derive star formation histories (SFH) and for studies of abundance distributions.

Here is the SFH for Leo I based on HST data[a] (left data, right convolved model and SFH).

[a] E. Tolstoy et al., A.J., 116, 1244 (1998)



STRUCTURE AND DYNAMICS OF GALAXIES

15. Kinematics of spiral galaxies

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/∼vdkruit

Beijing, September 2011

Outline

Analysis of HI observations
- Moment analysis
- Tilted rings

Examples of HI observations
- Example of an inclined galaxy: NGC 5055
- Example of an edge-on galaxy: NGC 891

HI velocity dispersions

CO and H_2

HI observations

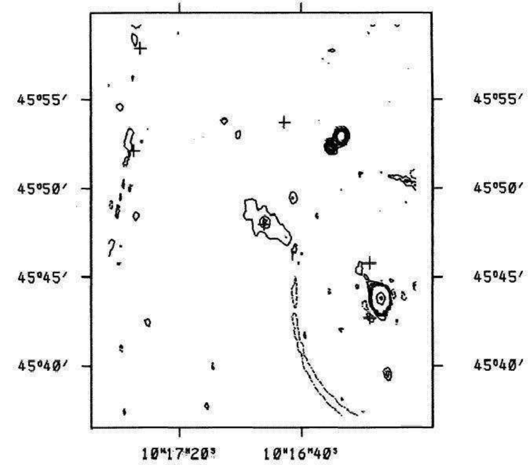
As an example I take the observations of NGC 3198 with the Westerbork Synthesis Radio Telescope.

These observations are part of the Palomar-Westerbork Survey of northern spiral galaxies[1].

This survey combined 21-cm observations of the neutral hydrogen with three-color optical surface photometry from photographic plates with the Palomar 48-inch Schmidt-telescope.

[1] B.M.H.R. Wevers, Ph.D. Thesis, 1984, B.M.H.R. Wevers, P.C. van der Kruit & R.J Allen, A.&A.Suppl. 66, 505 (1986)

The first thing to do is add up the channels at which no HI is present to find the continuum map.



The continuum radiation is mostly non-thermal synchrotron emission from relativistic electrons moving in the galactic magnetic field.

At the position of the HII-regions there also is thermal free-free emission from interaction between free electrons and ionized hydrogen (protons).

This particular galaxy has radio emission from the center and some extended faint emission from the disk.

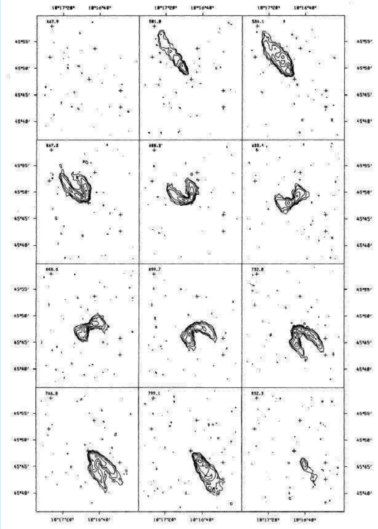
This continuum map is then subtracted from all channel maps to reveal the distribution of HI at various velocities.

The continuum map should be produced from as many channel maps as possible, so that the noise in it is low compared to that in the channel maps themselves.

Here are the channel maps of NGC 3198 as far as they contain neutral hydrogen emission.

The radial velocity increases from left-top (468 km sec^{-1}) to right-bottom (832 km sec^{-1}) in steps of 33 km sec^{-1}.

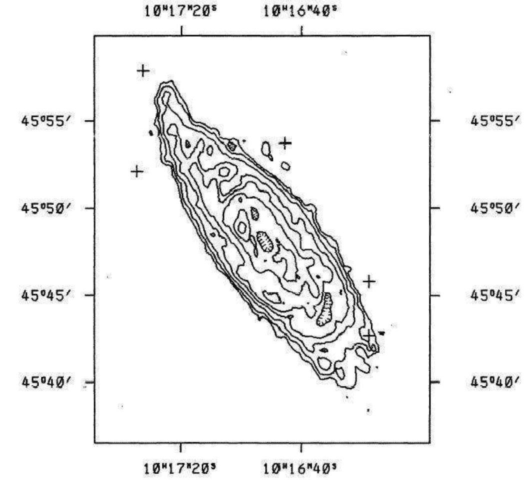
Obviously the northern (top) part is approaching us with respect to the systemic velocity and the southern part is receeding.



Analysis of HI observations

These channel maps can be added to produce the map with the distribution of neutral hydrogen, the total HI-map.

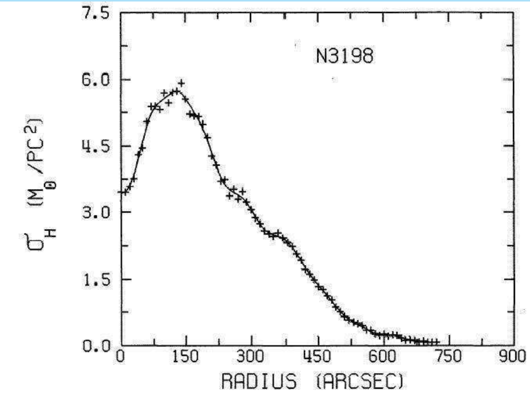
To suppress noise usually this is preceded by blocking out the areas in each of the channel maps that appear to have no HI-signal and thus contain only noise.



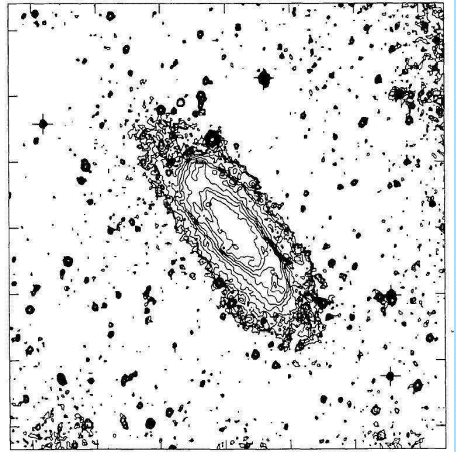
Moment analysis

From this map the radial HI profile can be produced by averaging in azimuthal annuli.

In practice this is done after analysis of the velocity field in order to find the position of the center and the orientation parameters (direction of major axis and inclination).



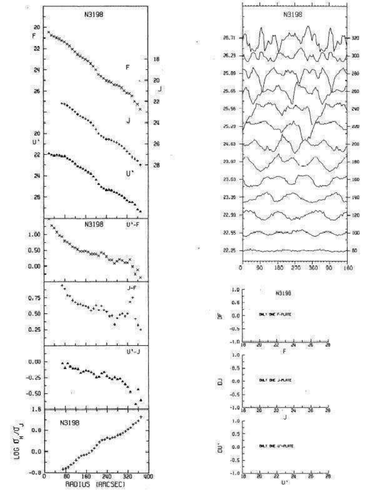
One can then take the optical map(s) and derive the radial luminosity profiles.



These can be further extended with radial color profiles and radial profile of the HI-surface density versus optical surface luminosity.

Here we have on the left from top to bottom the surface brightness profiles in three color bands, the radial profiles of three color indices and ratio of the (face-on) surface density if HI over the surface brightnes.

On the right are azimuthal color profiles and at the bottom differences of surface brightnesses from independent meassurements (not applicable here).



The profiles are tabulated here.

The units are magnitudes per arcsec^2 for surface brightness (or equivalently in solar luminosities per pc^2), magnitudes for color, solar masses per pc^2 for HI-surface densities and solar masses over solar luminosities for the density–brightness ratio.

NGC 3198 surface brightness

radius arcsec	U'	J	F	U'-J	J-F	U'-F
0.	-	-	-	-	-	-
10.	-	-	-	-	-	-
20.	21.91	-	20.43	-	-	1.48
30.	22.03	-	20.75	-	-	1.28
40.	22.08	-	20.88	-	-	1.20
50.	22.04	-	20.97	-	-	1.07
60.	22.07	-	21.10	-	-	0.97
70.	22.16	22.19	21.24	-0.03	0.95	0.92
80.	22.15	22.25	21.36	-0.10	0.89	0.79
90.	22.35	22.38	21.59	-0.03	0.79	0.76
100.	22.46	22.54	21.76	-0.08	0.78	0.70
110.	22.65	22.76	22.05	-0.11	0.71	0.60
120.	22.88	22.99	22.31	-0.11	0.68	0.57
130.	23.08	23.20	22.55	-0.12	0.65	0.53
140.	23.18	23.36	22.71	-0.18	0.65	0.47
150.	23.33	23.49	22.86	-0.16	0.63	0.47
160.	23.49	23.63	23.01	-0.14	0.62	0.48
170.	23.64	23.79	23.19	-0.15	0.60	0.45
180.	23.80	23.97	23.42	-0.17	0.55	0.38
190.	24.05	24.25	23.67	-0.20	0.58	0.38
200.	24.39	24.64	24.01	-0.25	0.63	0.38
210.	24.76	25.00	24.44	-0.24	0.56	0.32
220.	25.05	25.24	24.62	-0.19	0.62	0.43
230.	25.22	25.39	24.83	-0.17	0.56	0.39
240.	25.33	25.56	25.03	-0.23	0.53	0.30
250.	25.31	25.57	25.11	-0.26	0.46	0.20
260.	25.37	25.65	25.18	-0.28	0.47	0.19
270.	25.51	25.76	25.43	-0.25	0.33	0.08
280.	25.60	25.88	25.45	-0.28	0.43	0.15
290.	25.77	26.04	25.56	-0.27	0.48	0.21
300.	25.89	26.21	25.71	-0.32	0.50	0.18
310.	26.08	26.44	25.98	-0.36	0.46	0.10
320.	26.30	26.70	26.20	-0.40	0.50	0.10
330.	26.46	26.94	26.25	-0.48	0.69	0.21
340.	26.47	27.24	26.49	-0.77	0.75	-0.02
350.	26.74	27.41	26.99	-0.67	0.42	-0.25
360.	27.15	27.59	27.27	-0.44	0.32	-0.12
370.	27.37	27.98	27.73	-0.61	0.25	-0.36

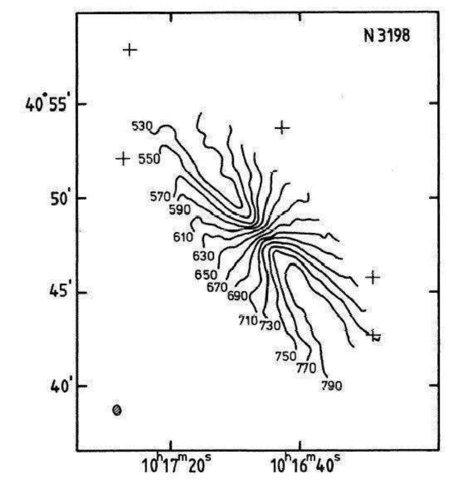
NGC 3198 HI surface density

radius arcsec	σ(HI) $M_\odot/pc^2$	σ(J) $L_\odot/pc^2$	LOG(σ(HI)/σ(J)) $M_\odot/L_\odot$
0.	3.453	-	-
30.	3.758	-	-
60.	5.039	-	-
90.	5.312	20.800	-.593
120.	5.706	11.859	-.318
150.	5.553	7.483	-.129
180.	5.154	4.809	.030
210.	4.262	5.862	.360
240.	3.741	1.112	.527
270.	3.299	.925	.552
300.	3.057	.611	.699
330.	2.586	.312	.919
360.	2.546	.171	1.172
390.	2.238	-	-
420.	1.723	-	-
450.	1.331	-	-
480.	1.024	-	-
510.	.656	-	-
540.	.499	-	-
570.	.355	-	-
603.	.271	-	-
630.	.236	-	-
660.	.132	-	-
690.	.105	-	-
720.	.081	-	-

The velocity field follows from deriving at each position the radial velocity.

This can be done either by moment analysis of the HI-profile or a fit with a Gaussian.

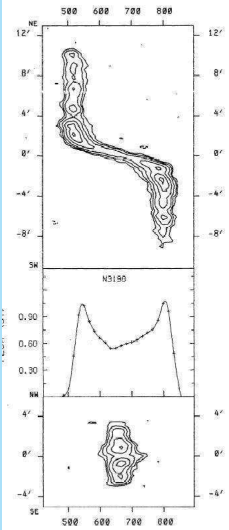
This is called a spider diagram.



Helpful for further analysis are also position-velocity diagrams (or x,V-diagrams), which have position along a line (or curve) on one axis and radial velocity on the other.

The figure shows the x,V-diagrams along the major and minor axis.

Also useful is the integrated profile.



Tilted rings

The next step is to analyse the velocity field in terms of the orientation of the plane of the disk and the rotation curve.

A first guess for the major axis direction and the inclination can be obtained from the distribution of HI and/or the optical image.

Assume we have a disk galaxy with a rotation curve $V_{\rm rot}(R)$.
The position angle of the major axis is $\Phi_\circ$ and the inclination is i (defined as zero for face-on).
Take the coordinates on the sky as (r, Φ) and in the plane of the galaxy (R, θ). Then

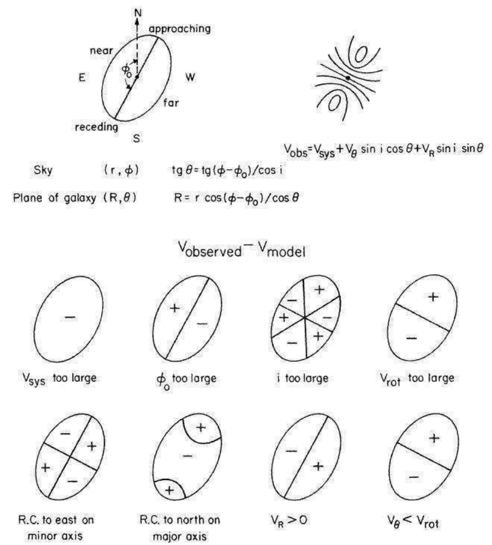
$$R = r\frac{\cos(\Phi - \Phi_\circ)}{\cos\theta} \qquad \tan\theta = \frac{\tan(\Phi - \Phi_\circ)}{\cos i}$$

$$V_{\rm obs} = V_{\rm sys} + V_{\rm rot}(R)\sin i\cos\theta$$

We can calculate the pattern of the residual velocity field after subtraction of a model.

We then see that errors in each parameter produce different patterns and therefore in principle these parameters can be determined independently[a].

[a] See P.C. van der Kruit & R.J Allen, Ann.Rev.Astron.Astrophys. 16, 103 (1978)



The usual procedure to determine the velocity field is as follows.

From the optical maps the position of the center, the position angle of the major axis and the inclination are estimated.

Then in rings in the galaxy plane (which corresponds to ellipses on the sky) the observed velocities are converted into "rotation velocities" along the ring.

Then changes in the parameters are introduced; this changes the run of deduced rotation velocity along the ring.

The parameters are optimized until these variations along the ring are minimal.

In practice it turns out that in particular in the outer regions the planes of the rings change.

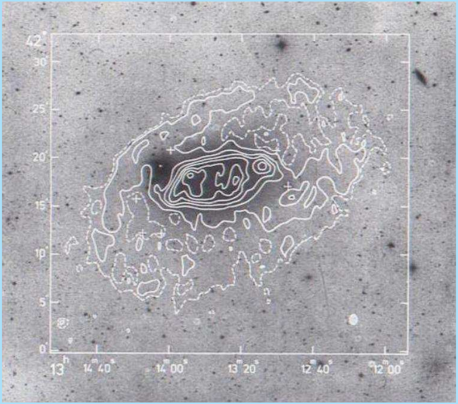
Examples of HI observations

Example of an inclined galaxy: NGC 5055

This is illustrated with the observations of NGC 5055[2].

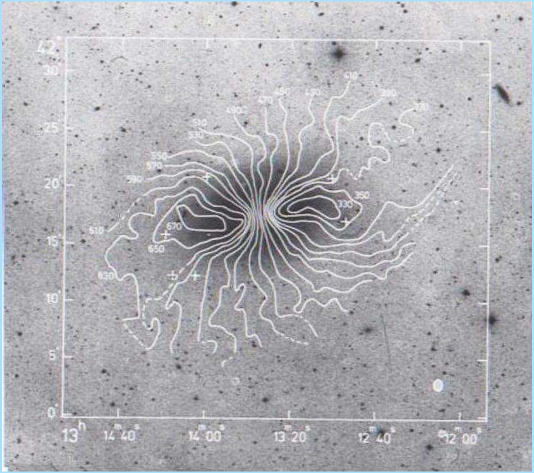
[2] A. Bosma, Ph.D. thesis, 1978; A.J. 86, 1791 (1981)

Here is the distribution of HI.

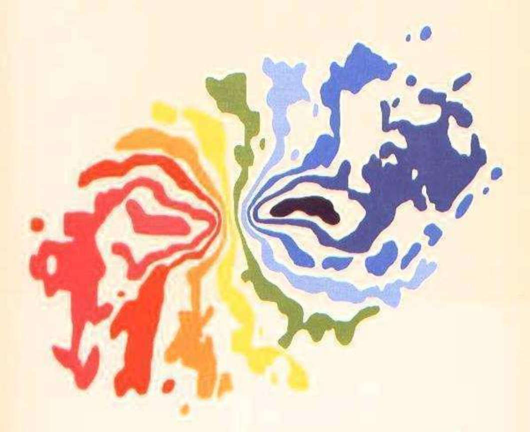


The distribution of the HI in the outer parts suggests that the plane of the disk changes. This is called a "warp".

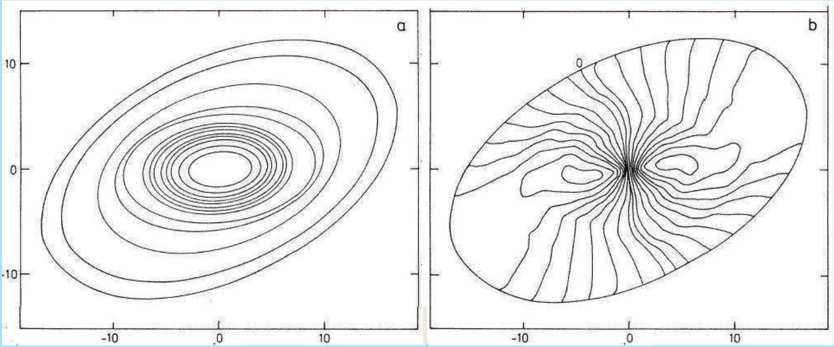
We also see distortions in the velocity field.



The velocity field is conveniently represented in color (from Albert Bosma's thesis):



The distribution and velocity field of the HI can be fitted with "inclined rings" with pure rotation in a changing plane.



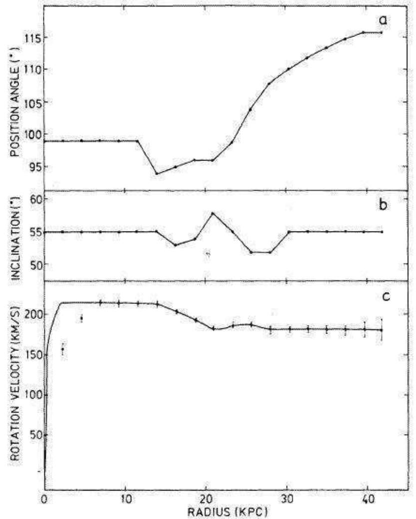
The figure shows from top to bottom:

Position angle of the major axis

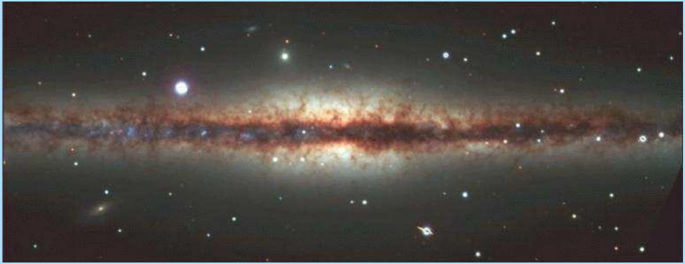
Inclination

Rotation velocity

We return to the matter of warps later.

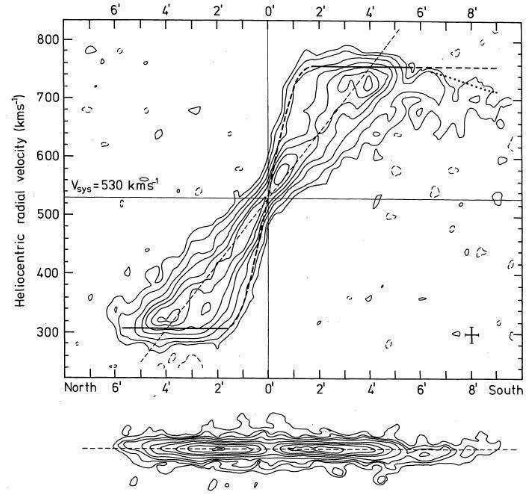


Example of an edge-on galaxy: NGC 891



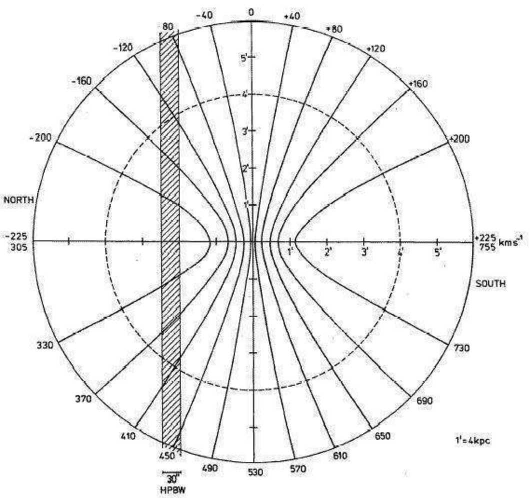
The observations are from Sancisi & Allen[3].

[3] R. Sancisi & R.J. Allen, A.&A. 74, 73 (1979)


6'
4'
2'
0'
2'
4'
6'
8'
800
700
600
500
400
300
Heliocentric radial velocity (kms⁻¹)
V_sys = 530 kms⁻¹
North
South

The position-velocity diagram (l, V-diagram) now is a projection of the plane of the galaxy with only a ambiguity around the "line of nodes".

This can be seen when we draw lines of equal line of sight velocity on the plane of the galaxy.



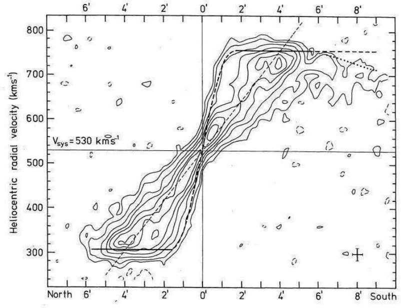
It is possible to model the I, V-diagram in terms of a distribution of the HI and a rotation curve.

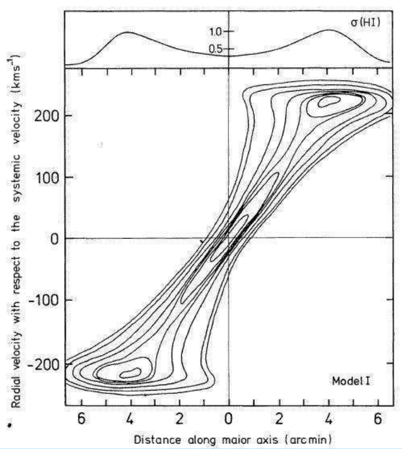
The radial HI distribution can be estimated by "decomposing" the observed HI on the sky under the assumption of circular symmetry.

The "extreme" or "high" velocities give a first estimate of the rotation curve.

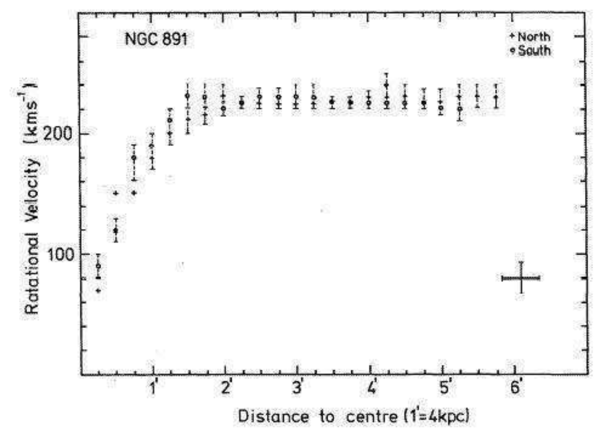
To properly model the I, V-diagram one needs to assume an HI velocity dispersion.

NGC 891 does not have an extended HI disk beyond the stellar disk and the HI layer appears very flat.


Heliocentric radial velocity (kms^{-1})
V_{sys} = 530 kms^{-1}
North
South


σ(HI)
Radial velocity with respect to the systemic velocity (kms^{-1})
Model I
Distance along major axis (arcmin)

The resulting rotation cuve is typical with a sharp rise and then remaining constant.

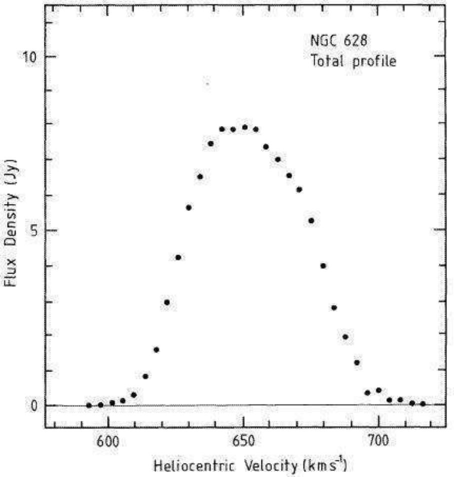


HI velocity dispersions

NGC 628 is very close to face-on and can therefore be used to measure the velocity dispersion of the HI[4].

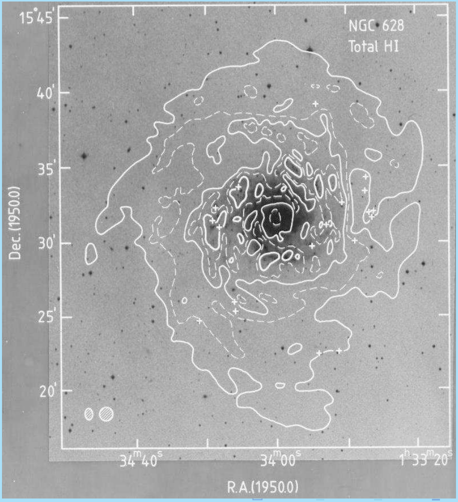
[4] G.S. Shostak & P.C. van der Kruit, A.&A. 132, 20 (1984)

The fact the NGC 628 is close to face-on is visible in the width of the integrated HI profile.



The HI is much more extended than the optical image.

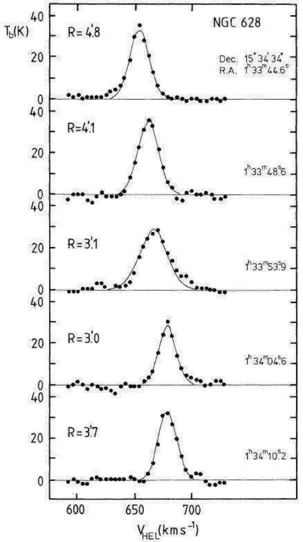
Also the spiral structure continues in the HI beyond the stellar disk and the optical spiral arms.



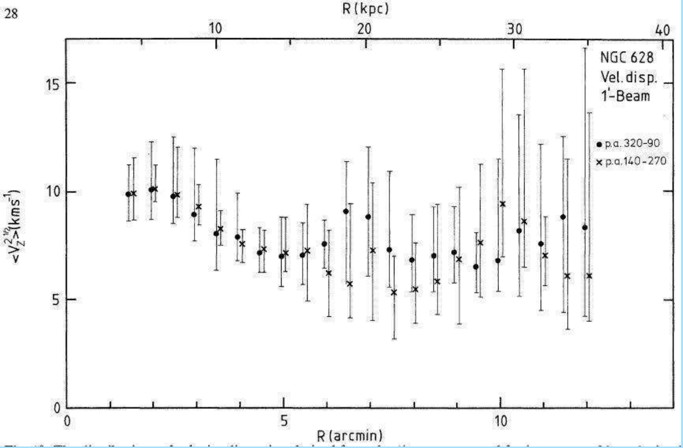
The next thing we can do is determine the velocity dispersion of the HI.

For this we need a face-on galaxy, because the gradient of systematic motion should be small accross a beam.
Here are some individual profiles at various distances from the center.

It can be seen that Gaussians can be fit very well to these profiles.



The HI velocity dispersion is between 7 and 10 km/s at all radii.



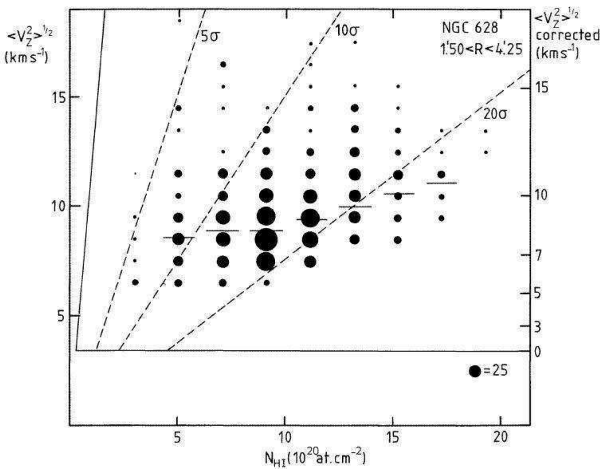
The velocity dispersion of the HI is expected to be isotropic due to cloud collisions.

This is confirmed by observations of more inclined (and large angular size) galaxies.

The value of 10 km/s corresponds roughly to a kinetic temperature of 10^4 K.

This is the temperature where cooling of the interstellar medium gets very effective due to ionisation of hydrogen.

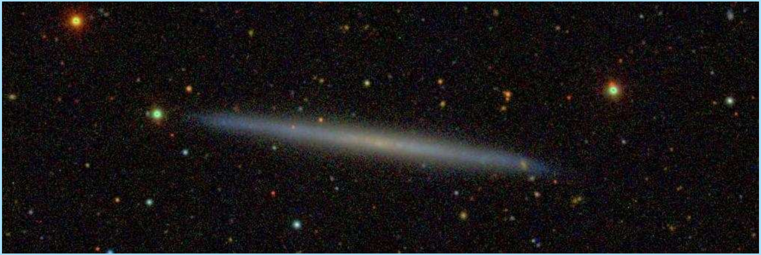
Closer analysis shows that within the optical image the velocity dispersion is systematically higher in areas of higher surface density (the spiral arms).



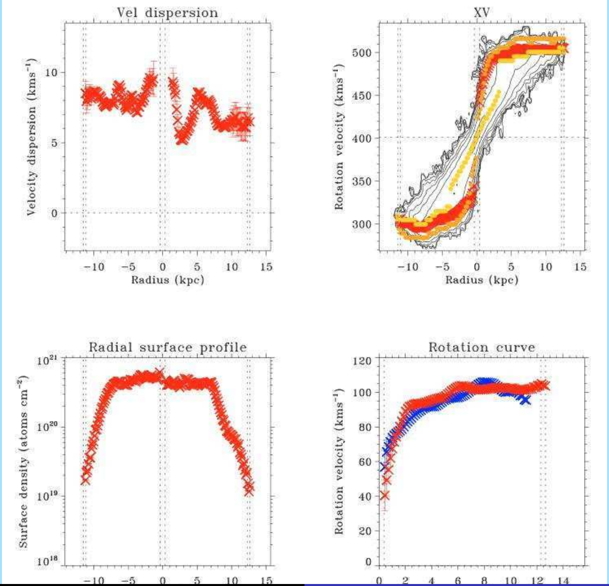
This is probably related to heating of the gas by star formation.

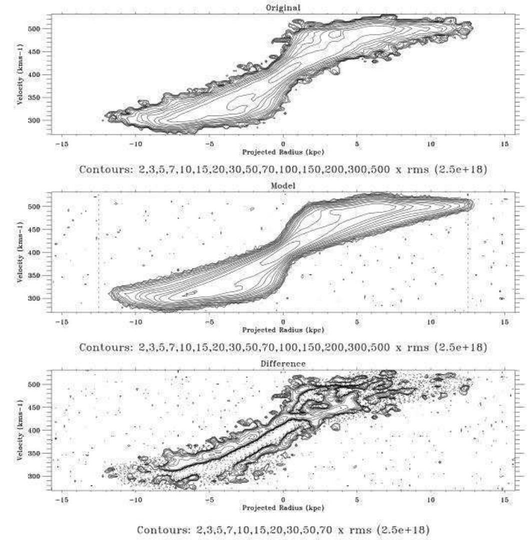
If the signal-to-noise of the data is good enough, we now can in edge-on systems als fit the radial HI distribution, the rotation curve and the velocity dispersion at the same time.[5]

Here are fits tot the superthin galaxy UGC7321.



[5]J.C. O'Brien, K.C. Freeman & P.C. van der Kruit, A.&A. 151, A62 & A63 (2010).


Vel dispersion
Velocity dispersion (kms^{-1})
Radius (kpc)
XV
Rotation velocity (kms^{-1})
Radius (kpc)
Radial surface profile
Surface density (atoms cm^{-2})
Rotation curve
Rotation velocity (kms^{-1})


Original
Velocity (kms-1)
Projected Radius (kpc)
Contours: 2,3,5,7,10,15,20,30,50,70,100,150,200,300,500 x rms (2.5e+18)
Model
Velocity (kms-1)
Projected Radius (kpc)
Contours: 2,3,5,7,10,15,20,30,50,70,100,150,200,300,500 x rms (2.5e+18)
Difference
Velocity (kms-1)
Projected Radius (kpc)
Contours: 2,3,5,7,10,15,20,30,50,70 x rms (2.5e+18)

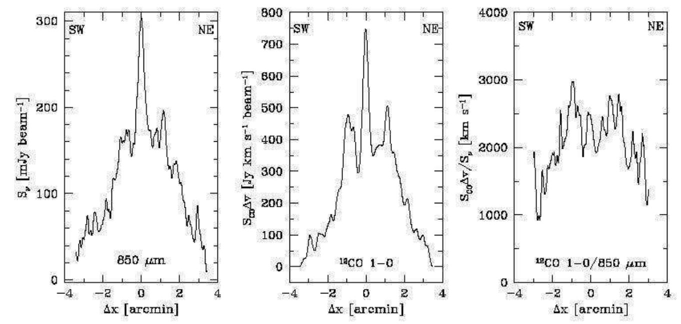
CO and H_2

The distribution of molecular hydrogen is often inferred from observations of CO at (sub-)millimeter wavelengths.

The assumption is that everywhere the ratio between these two molecules is the same.

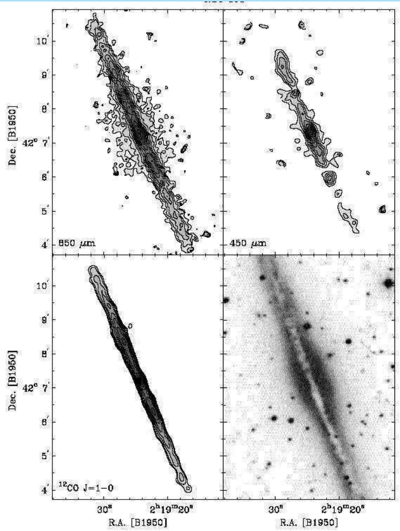
This is a dubious assumption, as this ratio is very likely dependent upon metallicity and physical conditions.

Here are some observations of NGC 891[6].



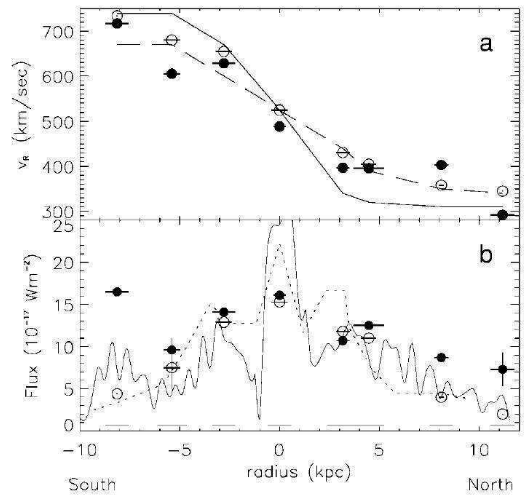
Here the near-infrared observations are also shown (these should show the distribution of the dust).

[6] F.R. Israel, P.P. van der Werf & R.J.P. Tilanus, A.&A. 334, L83 (1999)


10′
9′
8′
42° 7′
6′
5′
4′
Dec. [B1950]
850 μm
450 μm
^{12}CO J=1–0
30^s
2^{h}19^{m}20^s
R.A. [B1950]

Only recently has it been possible to directly measure lines of H_2 with the Infrared Space Observatory (ISO)[a].

We see here observations of the S(0) (28.2 μ) (filled) and S(1) (17.0 μ) (open) lines, compared with CO-observations.



[a] E.A. Valentijn & P.P. van der Werf, Ap.J. 522, L29 (1999)

STRUCTURE AND DYNAMICS OF GALAXIES

16. Rotation curves and dark matter

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/∼vdkruit

Beijing, September 2011

Outline

Tully-Fisher relation

Rotation curves and mass distribution

- Exponential disk
- Dark matter halo
- Maximum disk hypothesis
- Independent checks on the maximum disk hypothesis
- Modified dynamics

Tully-Fisher relation

For exponential disks:

$$M \propto \sigma_\circ h^2 \qquad V_{\max} \propto (\sigma_\circ h)^{1/2}$$

Then

$$M \propto V_{\max}^4 \sigma_\circ^{-1}$$

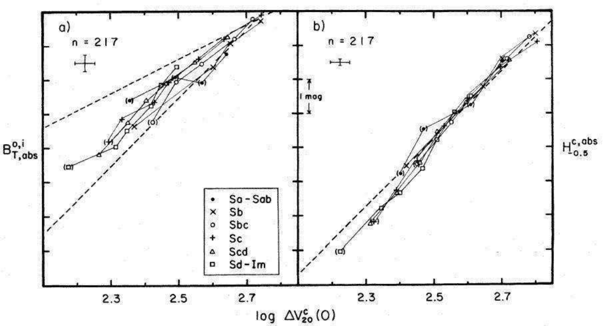
With Freeman's law and constant mass to light ratio M/L:

$$L \propto V_{\max}^4$$

This is the Tully-Fisher relation which has indeed been observed[1]. In practice $V_{\max}$ is measured from the total width of the HI-profile, corrected for inclination, at a level 20 or 50% of the peak.

[1] R.B. Tully & J.R. Fisher, A.&A. 54, 661 (1977)

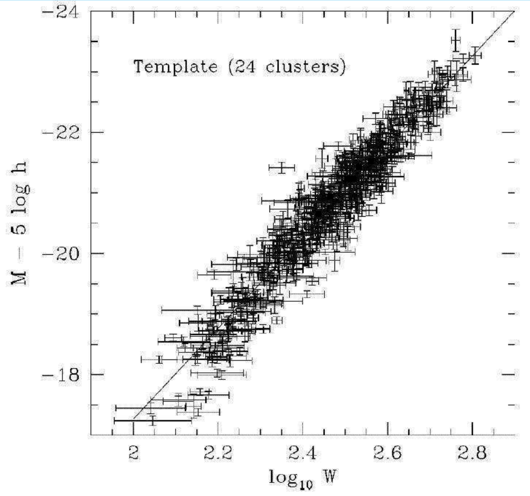
Aaronson & Mould[2] find exponents of 3.5 in B and 4.3 in H (1.6μ).



There is debate about the slope in observed relations.

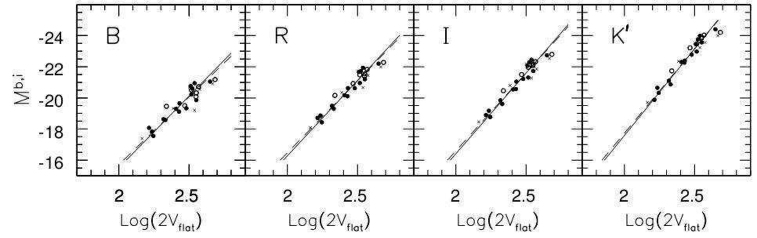
[2] M. Aaronson & J.R. Mould, Ap.J. 265, 1 (1983)

In the I-band Giovanelli *et al.*[a] find from 555 galaxies in 24 clusters a slope of 7.68 ± 0.13 (in magnitudes, which corresponds to 3.07 ± 0.05).



[a] R. Giovanelli & 6 other authors, Ap.J. 477, L1 (1997)

A recent study of the Ursa Major Cluster[3] shows that the relation is tightest in the K′-band and there the slope is 11.3 ± 0.5 (exponent 4.5 ± 0.2).



[3] M.A.M. Verheijen, Ph.D. thesis (1997) and Ap.J. 563, 694 (2001)

Rotation curves and mass distribution

Exponential disk

The exponential disk has a surface density distribution

$$\sigma(R) = \sigma_\circ e^{-(R/h)}$$

where $\sigma_\circ$ is the central surface density and h the scalelength. The total mass of the disk out to infinity is $M = 2\pi\sigma_\circ h^2$.

When it is self-gravitating and infinitessimally thin, the corresponding rotation curve has the analytic form[4]:

$$V^2_{\rm rot}(R) = \pi G h \sigma_\circ \left(\frac{R}{h}\right)^2 [I_0 K_0 - I_1 K_1]$$

I and K are modified Bessel functions *evaluated at $R/2h$*.

[4] K.C. Freeman, Ap.J. 160, 811 (1970)

This rotation curve has the properties

- that it rizes from the center to a maximum at $R = 2.2h$ with

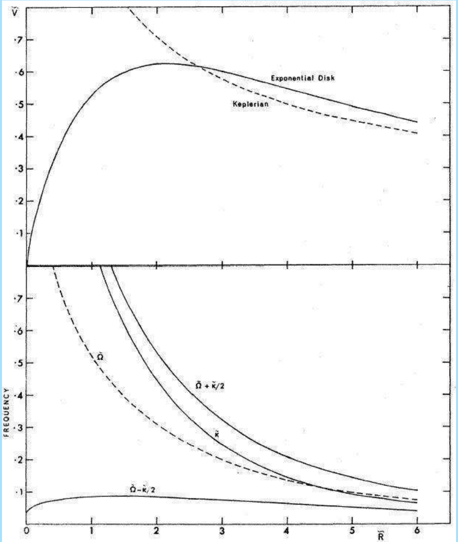
$$V_{\max} = 0.8796(\pi G h \sigma_0)^{1/2}$$

- and becomes Keplerian at large R.

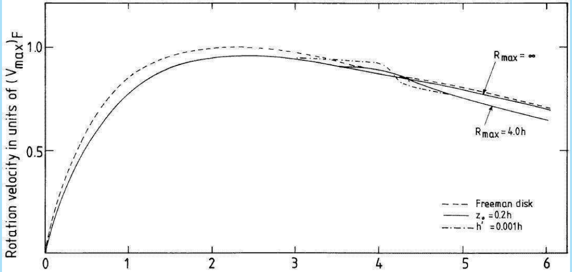
In the next figure the axes are dimensionless, such that $\tilde{R} = R/h$ and $\tilde{V} = V\sqrt{h/GM}$.

The lower half of the figure has the angular frequency Ω, the epicyclic frequency κ and the Lindblad resonance frequencies $\Omega \pm \kappa/2$.

These frequencies are in dimensionaless units of $\sqrt{GMh^3}$.



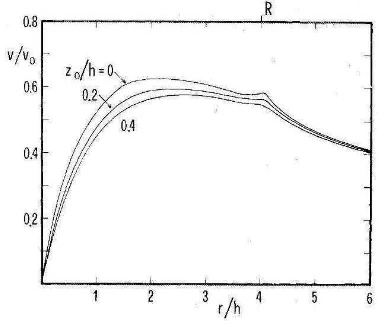
The rotation curve changes slightly when allowance is made for the finite thickness and the truncation[5].

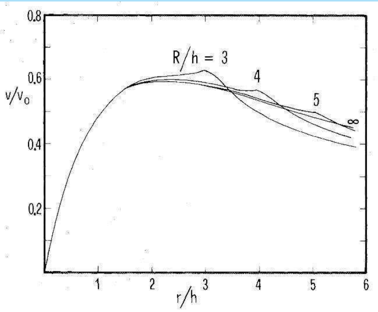


The dashed line has a infinitely thin disk, the full-drawn line has a finite thickness ($z_\circ = 0.2h$) without and with a shallow truncation (the scalelength changes by a factor 5 at $R_{\max}$). The dot-dashed curve has a very sharp edge.

[5]P.C. van der Kruit & L. Searle, A.&A. 110, 61 (1982)

Here are similar figures from another study[6] with a truncation as a linear drop in surface density over a radial range $\delta = 0.2h$.





On the left the thickness of the disk is varied and on the right the radius of the truncation.

[6] S. Casertano, Mon.Not.R.A.S. 203, 735 (1983)

Dark matter halo

Observations of spiral galaxies show flat rotation curves that do not show the Keplerian decline beyond the optical edge.

So add a dark halo with $\rho \propto R^{-2}$ at large R.

This can be an isothermal sphere[7] or some other analytical function[8].

In practice one may also directly infer a predicted rotation curve from the disk by calculated from the observed surface brightness profile.

[7] e.g. C. Carignan & K.C. Freeman, Ap.J. 294, 494 (1985)
[8] K. Begeman, Ph.D. thesis (1987)

In the general case that the disk density distribution is $\rho(R,z)$, the rotation curve from the corresponding self-gravitating disk is

$$V_{\rm c}^2(R) = -8GR \int_0^\infty r \int_0^\infty \frac{\partial \rho(r,z)}{\partial r} \, \frac{K(p)-E(p)}{(Rrp)^{1/2}} \, dz \, dr$$

with

$$p = x - (x^2-1)^{1/2} \ \text{ and } \ x = \frac{R^2+r^2+z^2}{2Rr}$$

When the density distribution is separable in $\sigma(R)$ and $Z(z)$ this becomes

$$V_{\rm c}^2 = -8GR \int_0^\infty r\sigma(r) \int_0^\infty \frac{\partial Z(z)}{\partial z} \frac{K(p)-E(p)}{(Rrp)^{1/2}} \, dz \, dr$$

The vertical distribution can for example be assumed to be the isothermal sheet.

We may in addition have a bulge with observed surface density $\sigma(r)$; then for the self-gravitating case we have

$$V_c^2(R) = \frac{2\pi G}{R} \int_0^R r\sigma(r)\, dr +$$

$$\frac{4G}{R} \int_R^\infty \left[\arcsin \left(\frac{R}{r} \right) - \frac{R}{(r^2 - R^2)^{1/2}} \right] r\sigma(r)\, dr$$

For the dark halo the assumed the density law

$$\rho(R) = \rho_\circ \left[1 + \left(\frac{R}{R_c} \right)^2 \right]^{-1}$$

results in

$$V_c^2(R) = 4\pi G \rho_\circ R_c^2 \left[1 - \frac{R_c}{R} \arctan \left(\frac{R}{R_c} \right) \right]$$

To get the total rotation curve for a system consisting of three components add these circular velocities in quadrature:

$$V_{\rm circ}(R) = \left[V^2_{\rm disk}(R) + V^2_{\rm bulge}(R) + V^2_{\rm halo}(R)\right]^{1/2}$$

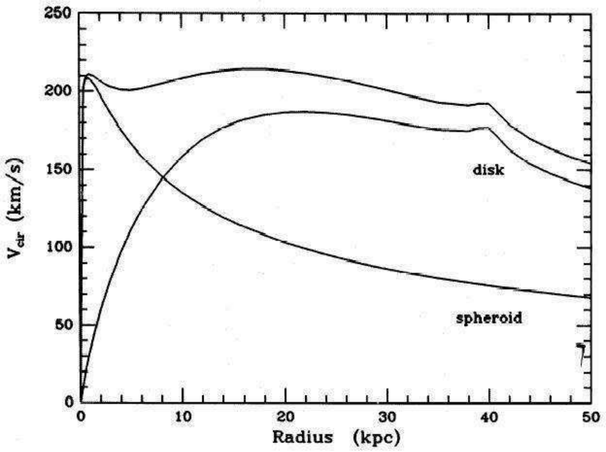
One can make things easier by fitting an exponential disk to the observations and use the analytic form of the corresponding rotation curve.

If in addition there is gas, this should be treated in the same way.

In practice we have for the stars only surface *brightness* distributions, so we need an undetermined mass-to-light ratio M/L in order to turn this into a surface *density* distribution.

From the solar neighborhood we can only find that M/L is of order a few in solar units.

In principle one can make an approximately flat rotation curve by a careful tuning of the disk and bulge contributions, as here for the Galaxy.

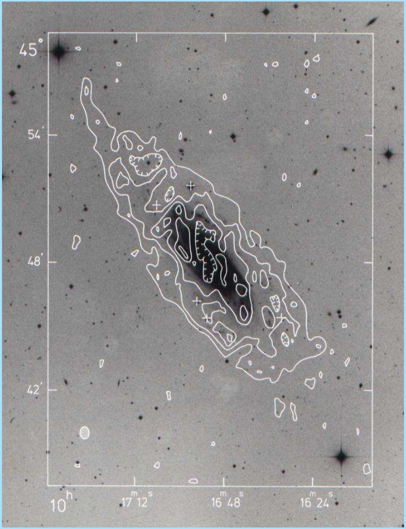


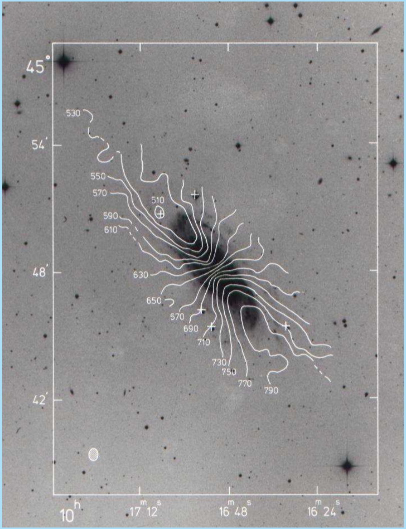
Maximum disk hypothesis

The following is from an analysis of the rotation curve of NGC 3198[a], which has essentially no bulge.

The HI extends out to 11 scalelengths.

[a] T.S. van Albada, J.N. Bahcall, K. Begeman & R. Sancisi, Ap.J. 295, 305 (1985)


45°
54′
48′
42′
10h
17m12s
16m48s
16m24s

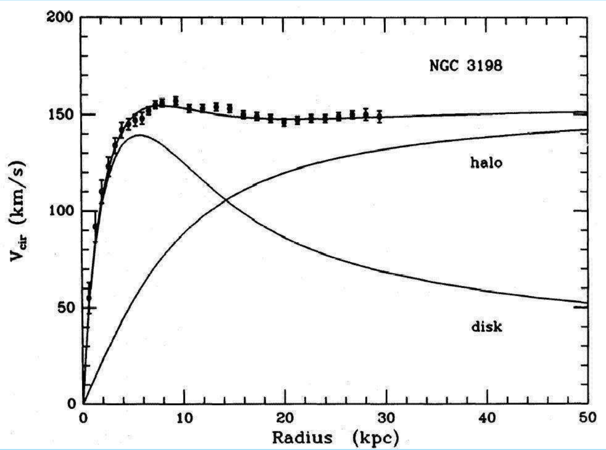

45°
54′
48′
42′
530
550
570
590
610
510
630
650
670
690
710
730
750
770
790
10h
17m12s
16m48s
16m24s

The procedure then is to choose an M/L of the disk that gives the maximum amplitude of the disk rotation curve that is allowed by the observations.

The two free parameters of the dark halo, core radius $R_{\rm c}$ and central density $\rho_\circ$ are then used to fit the rotation curve.

This is called the "maximum disk hypothesis", since it is a fit to the rotation curve with the largest amount of mass possible in the disk (and the largest M/L).

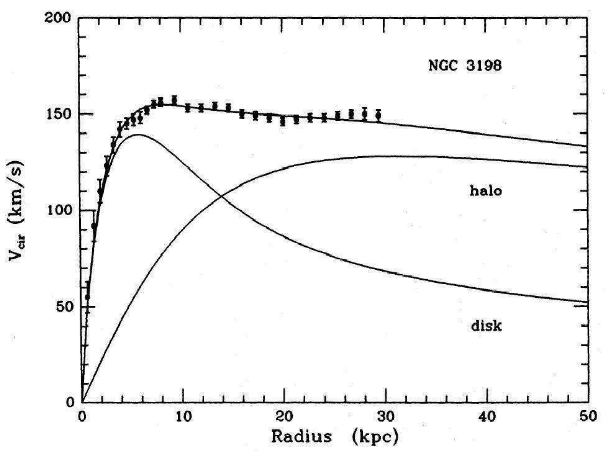
The maximum disk solution to the rotation curve of NGC 3198 looks as follows.



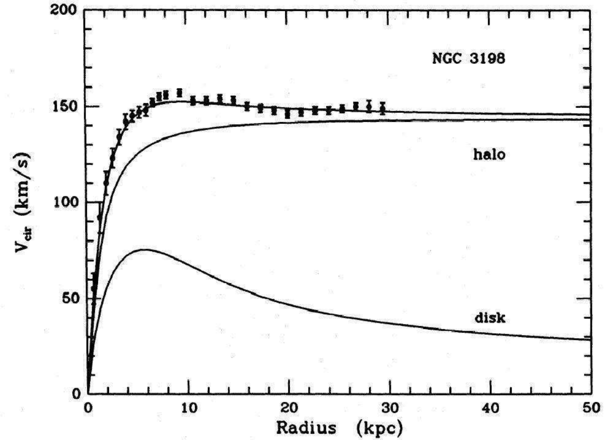
This particular model for NGC 3198 has a total mass of 15×10^{10} $M_\odot$ within 30 kpc.

Within this radius the ratio of dark to visible matter is 3.9. At the optical edge this ratio is 1.5.

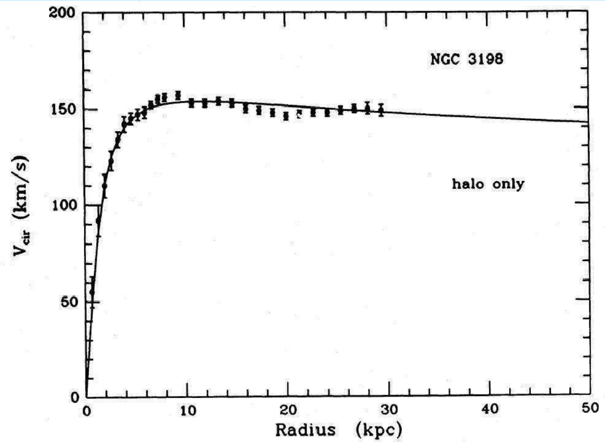
By adjusting the halo parameters one can minimize the dark halo mass by assuming that the rotation curve falls beyond the last measured point.


NGC 3198
halo
disk
V_{cir} (km/s)
Radius (kpc)
0
50
100
150
200
10
20
30
40
50

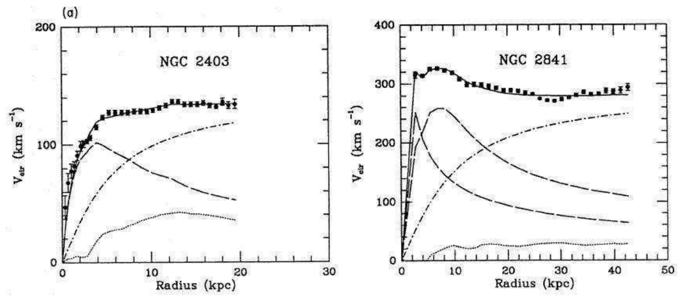
The difficulty with the maximum disk hypothesis is that it is possible to make similar good fits with lower disk masses ...



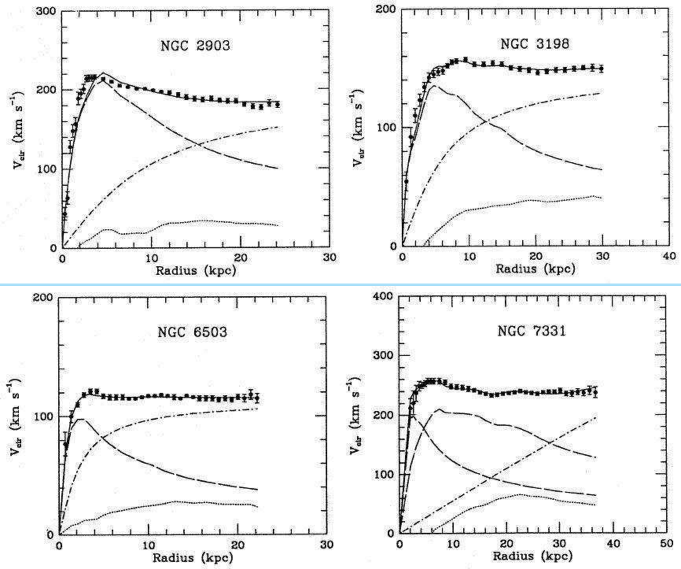
... and even no disk mass at all!



Begeman[9] observed 8 spirals, of which HI in NGC 2841 goes out to 17.8 h (43 kpc).



[9] K. Begeman, Ph.D. thesis (1987); K. Begeman, A.H. Broeils & R.H. Sanders, Mon.Not.R.A.S. 249, 523 (1991)


NGC 2903
NGC 3198
NGC 6503
NGC 7331
V_{cir} (km s^{-1})
Radius (kpc)

Independent checks on the maximum disk hypothesis

There are independent ways in which the maximum disk hypothesis can be checked by independent measurement of M/L.

a. The truncation feature in the rotation curve:

The truncation feature in the rotation curve can in principle be used to estimate the mass of the disk. It has been done in two cases where the mass of the halo within the truncation radisu has been estimated:

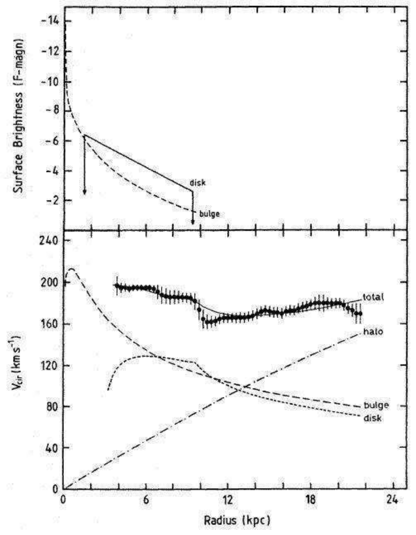
- NGC 5907[10]: $(M_{\rm halo})_{\rm R_{opt}} \approx 60\%$ (so not maximum disk)
- NGC 4013[11]: $(M_{\rm halo})_{\rm R_{opt}} \approx 25\%$

[10] S. Casertano, Mon.Not.R.A.S. 203, 735 (1983)
[11] R. Bottema, A.&A. 306, 345 (1996)

In NGC 4013 the disk and bulge must dominate dynamically in the inner regions.

The truncation feature is clearly visible.

However, the fit to the rotation curve is not maximum disk.



b. Maximum rotation versus scalelength

Another interesting argument is the following[12].

For a pure exponential disk the maximum in the rotation curve occurs at $R = 2.2h$ with an amplitude of

$$V_{\rm max} \propto \sqrt{h\sigma_\circ} \propto \sqrt{\frac{M_{\rm disk}}{h}}$$

For fixed disk-mass $M_{\rm disk}$ this gives

$$\frac{\partial \log V_{\rm max}}{\partial \log h} = -0.5$$

[12] S. Courteau & H.-W. Rix, Ap.J. 513, 561 (1999)

Remember that the Tully-Fisher relation is a tight correlation between maximum rotation and total luminosity of disk galaxies.

The total luminosity of an exponential disk is $L = 2\pi\mu_\circ h^2$.

Then at a given absolute magnitude (or mass) lower scalelength disks should have higher rotation.

So, if disk-dominated galaxies are maximum disk (in practice $V_{\rm disk} \sim 0.85 V_{\rm total}$) this should be seen in *scatter* in the Tully-Fisher relation

This is not observed and the estimate is that on average $V_{\rm disk} \sim 0.6 V_{\rm total}$.

c. Thickness of the HI-layer.

The thickness of the gas layer can be used to measure the surface density of the disk independent of the rotation curve.

The density distribution of the exponential, locally isothermal disk was:

$$\rho_*(R,z) = \rho_*(0,0)\,\exp\left(-R/h\right)\,\mathrm{sech}^{\,2}(z/z_\circ)$$

If the HI has a velocity dispersion $\langle V_{\mathrm{z}}^2\rangle_{\mathrm{HI}}^{1/2}$, and if the stars dominate the gravitational field

$$\rho_{\mathrm{HI}}(R,z) = \rho_{\mathrm{HI}}(R,0)\,\mathrm{sech}^{\,2p}(z/z_\circ)$$

$$p = \frac{\langle V_{\mathrm{z}}^2\rangle_*}{\langle V_{\mathrm{z}}^2\rangle_{\mathrm{HI}}}$$

The full width at half maximum of this distribution is:

$$W_{\rm HI} = 1.663 p^{-1/2} z_\circ \ \ \text{for } p \gg 1$$

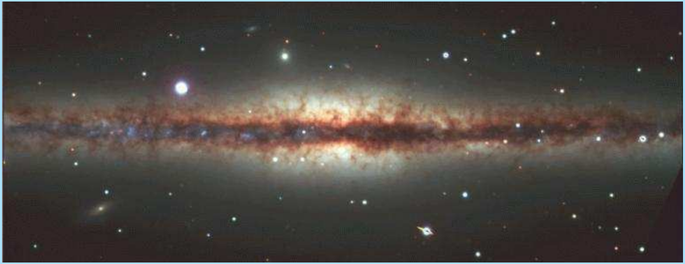
$$W_{\rm HI} = 1.763 p^{-1/2} z_\circ \ \ \text{for } p = 1$$

Then to within 3%

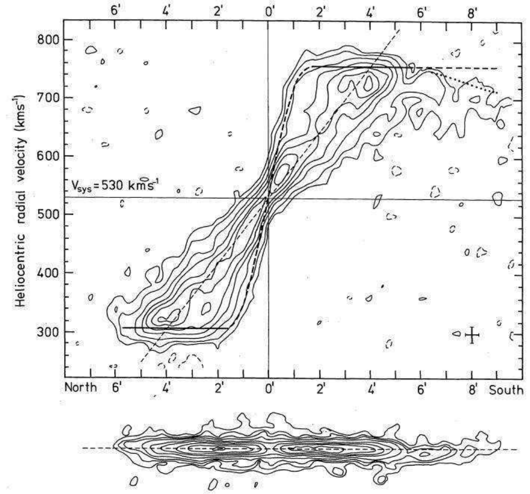
$$W_{\rm HI} = 1.7 \langle V_z^2 \rangle_{\rm HI}^{1/2} \left[\frac{\pi G (M/L) \mu_\circ}{z_\circ} \right]^{-1/2} \exp\left(R/2h\right)$$

So the gas layer increases exponentially in thickness with an e-folding of $2h$.

We now look at an analysis of the HI-layer in NGC 891[13] from measurements by Sancisi & Allen[14].

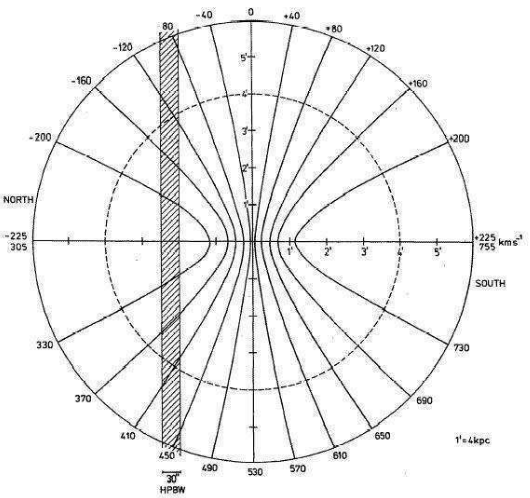


[13] P.C. van der Kruit, A.&A. 99, 298 (1981)
[14] R. Sancisi & R.J. Allen, A.&A. 74, 73 (1979)

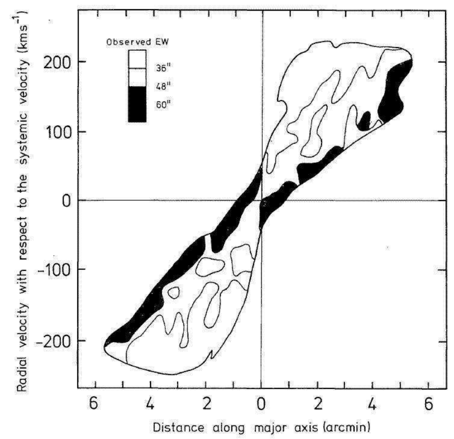

6' 4' 2' 0' 2' 4' 6' 8'
800
700
600
500
400
300
Heliocentric radial velocity (kms^{-1})
V_{sys} = 530 kms^{-1}
North 6' 4' 2' 0' 2' 4' 6' 8' South

The position-velocity diagram (l, V-diagram) is a projection of the plane of the galaxy with only a ambiguity around the "line of nodes".

This can be seen when we draw lines of equal line of sight velocity on the plane of the galaxy.



Here is a measure of the thickness.

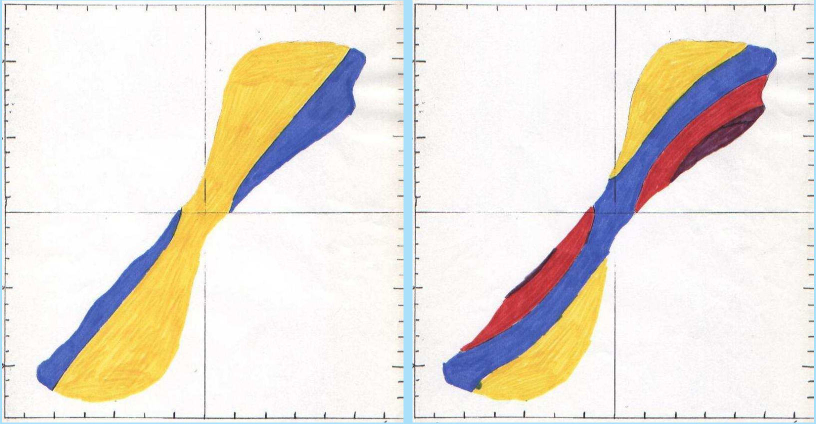


Three particular models were then calculated:

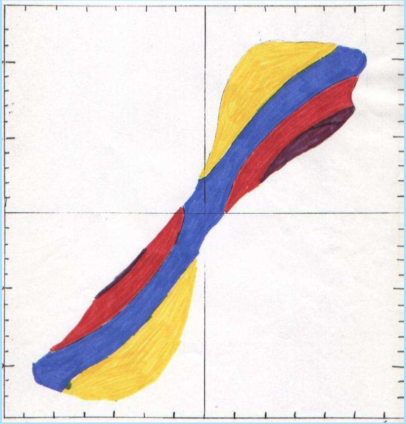
- Model I, which has 40% of the mass within the optical radius in the disk,
- Model II with all the mass (including the dark mass) in the disk,
- Model III with a constant thickness of the HI-layer.

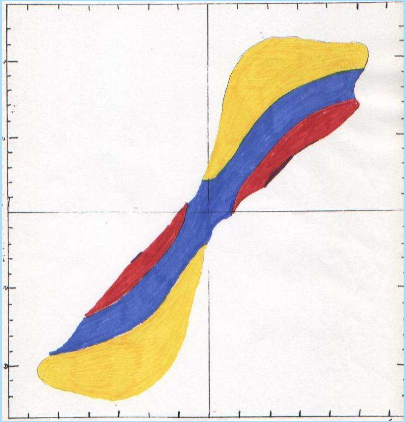
The W_{HI} in the observations were then calculated for disks with inclinations of 87.5 and 90°.

Here is the equivalent width in the (x, V)-diagram for Model I with inclinations of 90° (left) and $87\overset{\circ}{.}5$ (right).

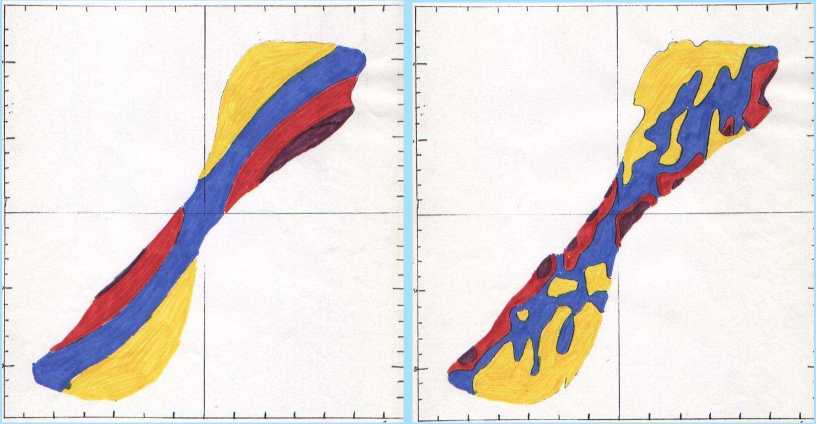


Here is the equivalent width in the (x, V)-diagram for Model I (left) and Model II (right) both at an inclination of 87°5.

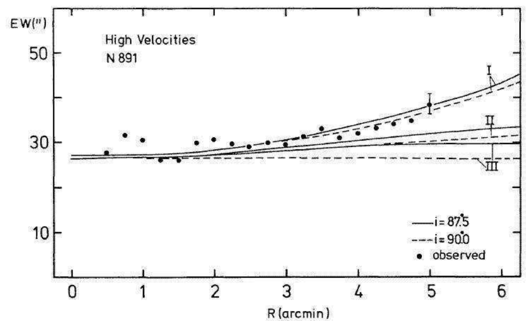




Here is the equivalent width in the observed (x, V)-diagram (right) and that for Model I with an inclination of 87°.5 (left).



Here we see the thickness over the “high” velocities only (190 to 230 km/s), compared to observations.



NGC 891 is not maximum disk. Also this analysis shows that the dark matter cannot be in the disk.

d. Thickness of the stellar disk

The vertical motions of the stars can be combined with the thickness of stellar disks to estimate of the disk surface densities σ.

For the isothermal sheet with space density

$$\rho(z) = \rho(0)\ \mathrm{sech}^{\,2}(z/z_\circ)$$

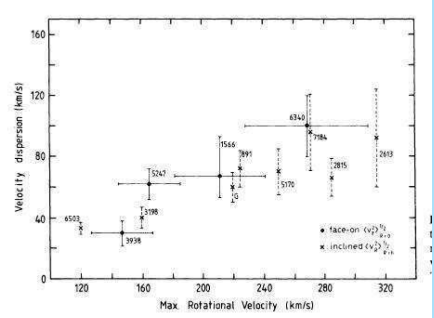
we had for the stellar velocity dispersion

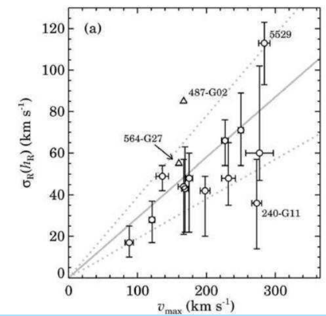
$$\langle V_{\mathrm{z}}^2\rangle^{1/2} = \sqrt{2\pi G\rho(0)}z_\circ = \sqrt{\pi G\sigma}z_\circ$$

Roelof Bottema[15] found that the stellar velocity dispersion at a fiducial radius correlates maximum in the rotation curve.

[15] R. Bottema, A.&A. 275, 16 (1993)

On the left Bottema's original correlation and on the right the same from a more recent study[16].





[16] M. Kregel, P.C. van der Kruit & K.C. Freeman, Mon.Not.R.A.S. 358, 503 (2005)

Using this relation we can estimate the disk surface density if we know $z_\circ$ and the rotation curve.

Statistical analysis of samples of galaxies gives[17] then is

$$\frac{V_{\rm rot,disk}}{V_{\rm rot,obs}} = 0.56 \pm 0.06.$$

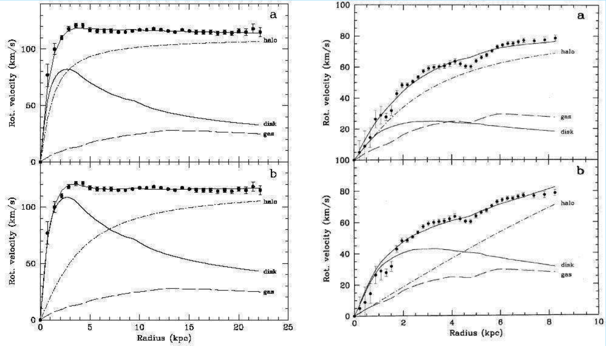
A working definition[18] of this ratio for a maximum disk is

$$\frac{V_{\rm rot,disk}}{V_{\rm rot,obs}} = 0.85 \pm 0.10.$$

So, in general galaxy disk appear to be NOT maximum disk.

[17] R. Bottema, A.&A. 275, 16 (1993); M. Kregel, P.C. van der Kruit & K.C. Freeman, Mon.Not.R.A.S. 358, 5003 (2004)
[18] P.D. Sackett, Ap.J. 483, 103 (1997)

Bottema's analysis[19] on a high surface brightness and a low-surface brightness galaxy gives a model according to the stellar velocity dispersion as at the top and the maximum disk hypothesis as at the bottom.



[19] R. Bottema, A.&A. 328, 517 (1997)

e. Our Galaxy

The hydrodynamical equation describes how the distribution and kinematics of a tracer population relates to the vertical gravitational force.

$$-K_{\mathrm{z}} = \frac{1}{\nu}\frac{\partial}{\partial z}(\nu\sigma_{\mathrm{zz}}^2) + \frac{1}{\nu R}\frac{\partial}{\partial R}(\nu R\sigma_{\mathrm{Rz}}^2)$$

The second term can usually be neglected and if the tracer population is isothermal then

$$K_{\mathrm{z}} = \sigma_{\mathrm{zz}}^2 \frac{\partial}{\partial z}\ln \nu(z)$$

The Poisson equation relates the gravitational field to the total density distribution ρ.

At small distances z from the plane these equations can be combined to give

$$4\pi G\rho_\circ = \frac{\partial}{\partial z}\left[\frac{1}{\nu}\frac{\partial}{\partial z}(\nu\sigma_{\mathrm{zz}}^2)\right]$$

One can use samples of for example K giants or (older) F dwarfs to this. This idea goes back to Kapteyn[20] and Oort[21].
Modern analyses of this kind have been done by Bahcall[22] and Kuijken & Gilmore[23].

Bahcall finds for the space density in the solar neighborhood $0.21 \pm 0.04\ \mathrm{M_\odot pc^{-3}}$.

[20] J.C. Kapteyn, Ap.J. 55, 302 (1922)
[21] J.H. Oort, Bull.Astron.Inst.Neth. 6, 249 (1932)
[22] J.N. Bahcall, Ap.J. 276, 156 and 169, Ap.J. 287, 926 (1984)
[23] K.H. Kuijken & G. Gilmore, Mon.Not.R.A.S. 239, 571, 605 and 651 (1989)

Observed are the following contributions.

Component	density
Main sequence stars	0.044
Subgiants and giants	0.002
White dwarfs	0.005
ISM (atomic & molec. gas, dust)	0.045
Population II	0.0001
Total	0.096

So in this case a total of about $0.1\ \mathrm{M_\odot pc^{-3}}$ is unaccounted for. This problem has been known for many years and is known as the "Oort limit".

Large numbers of brown dwarfs or stellar remnants cannot completely be ruled out.

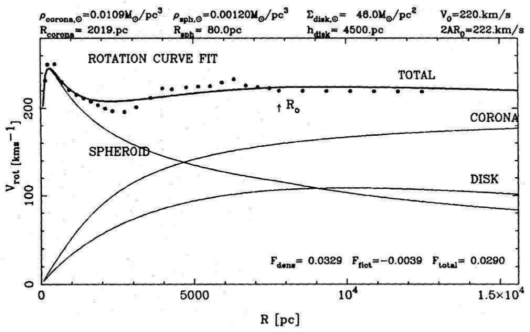
Kuijken & Gilmore on the other hand find that the local density is about 0.10 $\mathrm{M_{\odot}pc^{-3}}$ and that there is no convincing evidence for missing matter.

In terms of surface density of the Galactic disk, Bahcall finds a value of $66 \pm 8\ \mathrm{M_{\odot}pc^{-2}}$. This is distributed as follows:

Component	mass $\mathrm{M_{\odot}pc^{-2}}$	luminosity $\mathrm{L_{\odot}pc^{-2}}$
Main sequence stars	23.9	9.7
Subgiants and giants	1.0	13.3
White dwarfs	3.6	0.0
Interstellar medium	4.5	0.0
Unseen matter	33.0	0.0
(Population II)	(3.0)	(1.5)
Total	66.0	23.0

Kuijken & Gilmore find a total surface density of $46 \pm 9\ \mathrm{M_\odot pc^{-2}}$, of which $35 \pm 5\ \mathrm{M_\odot pc^{-2}}$ is in stars and $13 \pm 3\ \mathrm{M_\odot pc^{-2}}$ in gas and dust.

They also propose the following fit to the rotation curve of the Galaxy.



It follows that the Galaxy is NOT maximum disk.

With $\kappa \sim 31$ km sec^{-1} kpc^{-1} and $\sigma_{\mathrm{RR}} \sim 40$ km sec^{-1} the Toomre parameter can be determined as

$$Q \sim 2.1.$$

Disk stars have varying vertical distributions, according to the velocity dispersion – age relation.

This is also reflected in the (exponential) scaleheight derived from counts as a function of absolute magnitude.

Modified dynamics

Flat rotation curves may show that classic Newtonian gravity does not work at large distances[24]. For this purpose Modified Newtonian Dynamics (MOND)[25] was developed.

This has an accelaration $\vec{g}$, which is related to Newtonian acceleration $\vec{g}_{\mathrm{N}}$ as

$$\vec{g}\left(\frac{g}{a_\circ}\right)\left[1+\left(\frac{g}{a_\circ}\right)^2\right]^{-1/2} = \vec{g}_{\mathrm{N}}$$

with $a_\circ \sim 1.2 \times 10^{-8}$ cm sec^{-2}.

[24] e.g. R.H. Sanders, Mon.Not.R.A.S. 223, 539 (1986); K. Begeman, A.H. Broeils & R.H. Sanders, Mon.Not.R.A.S. 249,523 (1991)

[25] e.g. M. Milgrom, Ap.J. 270, 365 (1983)

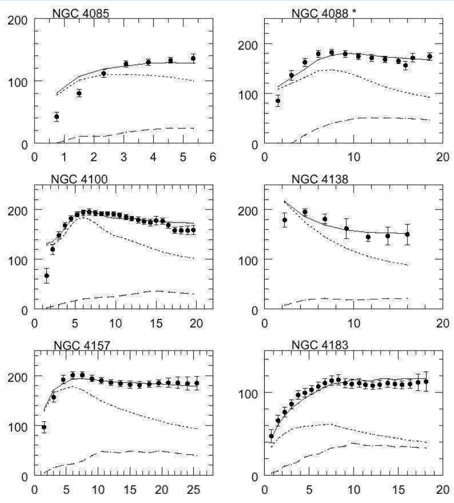
- For large accelerations $g/a_\circ$ this reduces to Newtonian gravity. So on small scales (in the solar system or the inner parts of galaxies) we have $g = g_{\rm N} \propto R^{-2}$ and Keplerian rotation with $V_{\rm rot}^2 \propto R^{-1}$.
- But at low accelerations is becomes $g = (g_{\rm N} a_\circ)^{1/2}$. Since now $g \propto R^{-1}$ this gives rise to $V_{\rm rot}^2 \propto R^0 =$ constant.

The result is that flat rotation curves can be produced without introducing a dark halo .

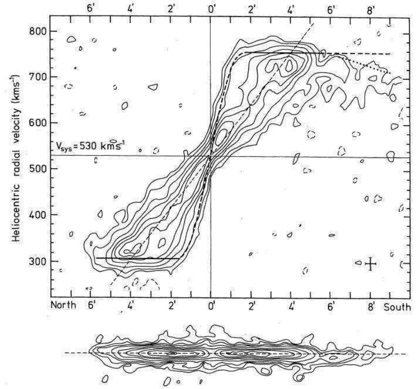
Here are some fits to actual rotation curves[a].

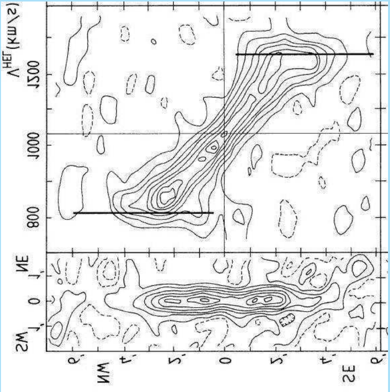
The full lines are the MOND-fits and the other lines show Newtonian curves for the stars and gas.

[a] R.H. Sanders & M.A.M. Verheijen, Ap.J. 503, 97 (1998)

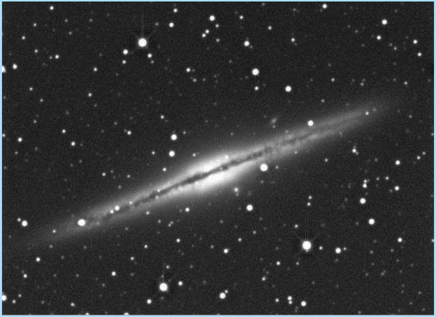


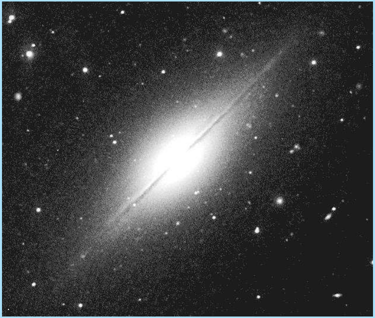
NGC 891 and NGC 7814 have the very similar rotation curves.





but completely different light distributions.[26]





This is inconsistent with MOND.

[26] See also F. Fraternali, R. Sancisi & P. Kamphuis, Astron.Astrophys. 531, A64, 2011

STRUCTURE AND DYNAMICS OF GALAXIES

17. Warps and dust.

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/∼vdkruit

Beijing, September 2011

Outline

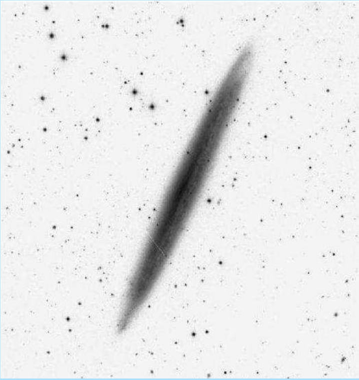
Warps in HI
- Warps: observations
- Warps: origin

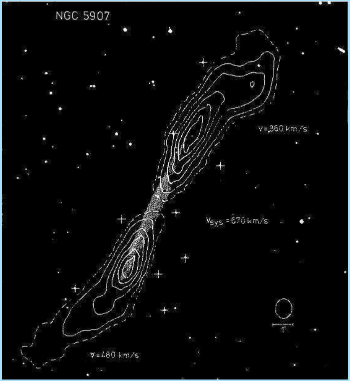
Dust and absorption
- Holmberg's analysis
- Analysis of Disney et al.
- Background galaxies

Warps in HI

Warps: observations

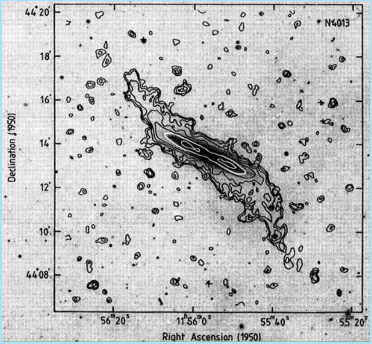
- Warps in the HI in external galaxies are most readily observed in edge-on systems as NGC 5907[1].

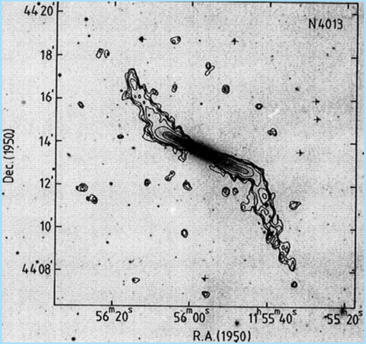




[1] R. Sancisi, A.&A. 74, 73 (1976)

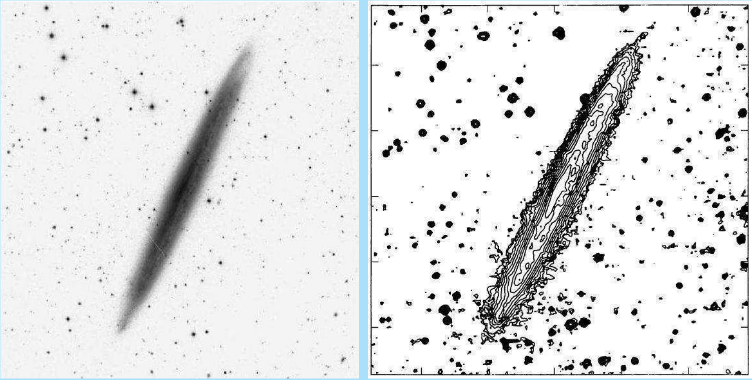
- An extreme example is "prodigious warp" in NGC 4013[2].
- The warp is very symmetric and starts suddenly near the end of the optical disk (see the extreme channel maps on the left).





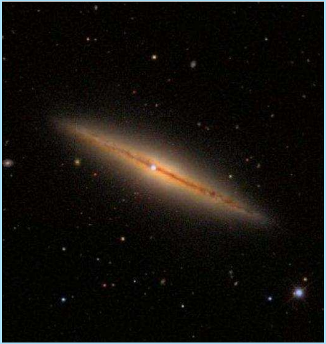
[2] R. Bottema, G.S.Shostak & P.C. van der Kruit, Nature 328, 401 (1987); R. Bottema, A.&A. 295, 605 (1995) and 306, 345 (1996)

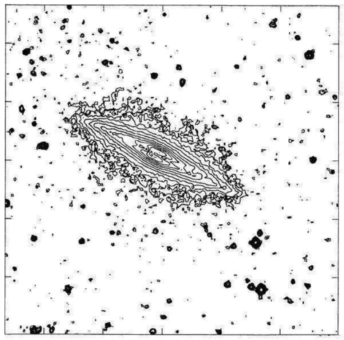
- It is interesting to note that the NGC 5907 has a clear and sharp truncation[3] in its stellar disk, where also the warp starts.



[3] P. C. van der Kruit & L. Searle, A.&A. 110, 61

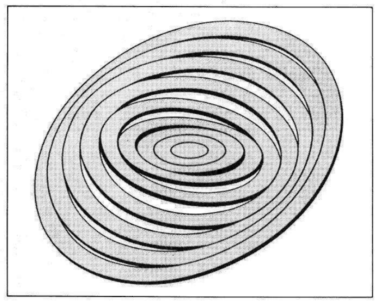
- NGC 4013 also has a clear truncation[4] in its stellar disk. The three-dimensional analysis[5] does confirm that in deprojection the warp strats very close to the truncation radius.

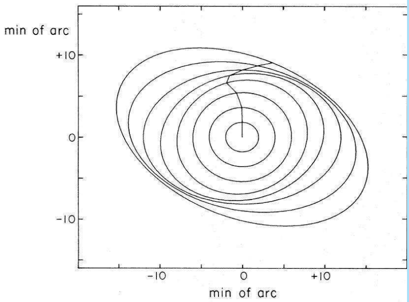




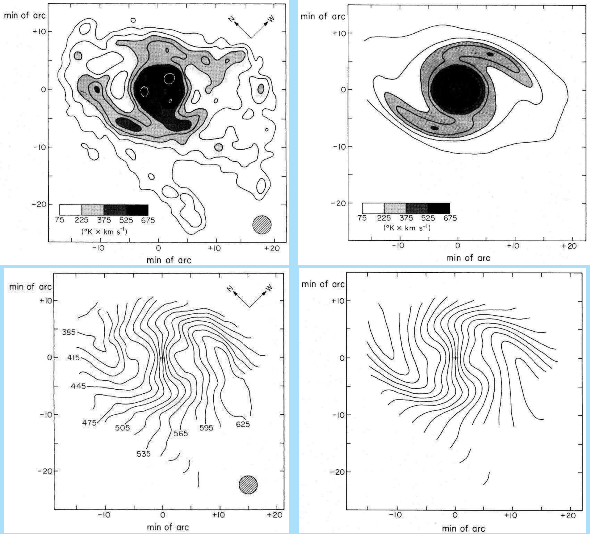
[4] P. C. van der Kruit & L. Searle, *op. cit.*
[5] R. Bottema, *op. cit.*

- Warps were already seen in less inclined systems, such as M83[6].
- These "kinematic warps" were fitted with so-called "tilted-ring models".





[6] D.H. Rogstad, I.A. Lockhart & M.C.H. Wright, Ap.J. 193, 309 (1974)


min of arc
+10
0
-10
-20
75 225 375 525 675
(°K × km s⁻¹)
-10
0
+10
+20
min of arc
N
W
385
415
445
475
505
535
565
595
625

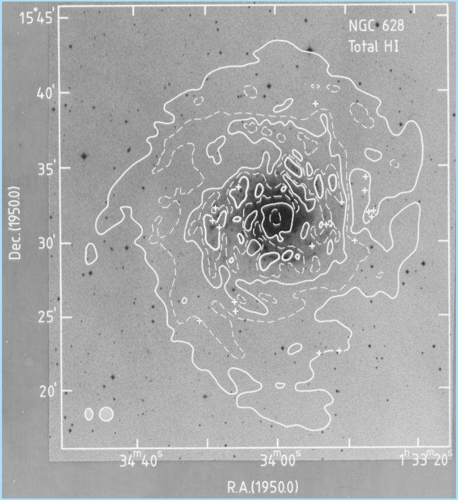
Velocity dispersions

NGC 628 is very close to face-on and can therefore be used to measure the velocity dispersion of the HI[7].

[7] G.S. Shostak & P.C. van der Kruit, A.&A. 132, 20 (1984)

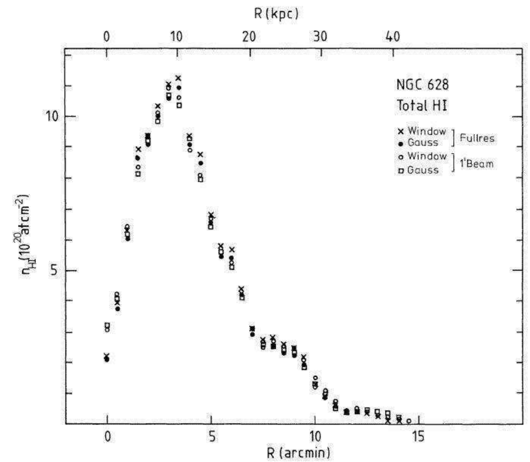
The HI is much more extended than the optical image.

Also the spiral structure continues in the HI beyond the stellar disk and the optical spiral arms.



Since the disk is so close to face-on we can derive the radial distribution of the HI from simple averaging in circular annuli on the sky.

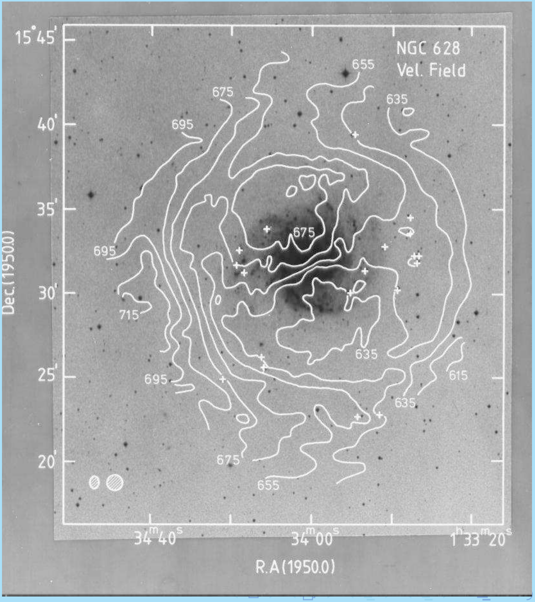
There is a feature in the profile at the edge of the stellar dsisk ($\sim$ 6 arcmin).



The velocity field looks regular in the central part, but has clear deviations in the outer part.

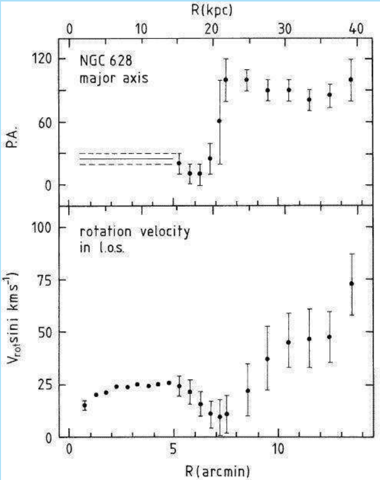
The disk is warped and the HI-plane moves actually through the plane of the sky.

At a radius of about 7 arcmin the observed velocity is about the systemic velocity.



The parameters of the tilted-ring model show this also.

At about 7 arcmin the position angle moves through a large angle and the observed rotation drops to zero and then increases again.



The rotation curve has an amplitude of ~25 km/s. For a galaxy of this type and absolute magnitude (using the Tully-Fisher relation; see later) the rotation velocity should be 200 to 250 km/s.

The inclination is then only 5 to 7°.

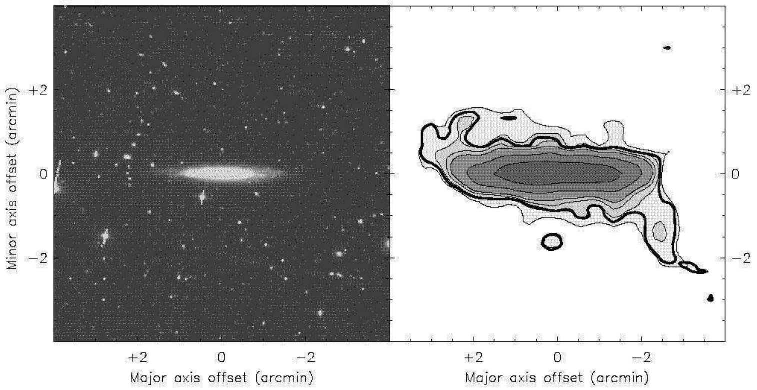
Over the optical part we can derive the residual velocity field when that from rotation is subtracted from the observations.

This shows no systematic pattern and has an r.m.s. value of only 3.9 km/s.

Any systematic pattern of vertical motion is small (or mimic that of rotation) and the disk is therefore be extremely flat.

For comparison, in the solar neighborhood a vertical velocity of 4 km/s corresponds to an amplitude of only 45 pc.

García Ruiz[8] has done a survey of edge-on galaxies.



[8] I. García-Ruis, Ph.D. thesis (2001); I. García-Ruiz, R. Sancisi & K.H. Kuijken, A.&A. 394, 796 (2002)

His major findings are;

- All galaxies, in which the HI is more extended than the stellar disk have warps.
- The warp usually starts near the edge of the stellar disk.
- Galaxies in rich environments tend to have larger and more asymmetric warps.

Warps: origin

Briggs[9] formulated a set of rules of behaviour for HI- warps.

RULES OF BEHAVIOR FOR GALACTIC WARPS

F. H. Briggs
Kapetyn Astronomical Institute, University of Groningen, and Department of Physics and Astronomy, University of Pittsburgh

Received 1989 July 21; accepted 1989 September 19

ABSTRACT

A sample of galaxies is now available for which H I 21 cm line observations allow the development of detailed kinematic models based on concentric, circular rings with adjustable inclinations and orbital velocity. By examining these warped systems in a variety of reference frames, clear empirically determined "rules" for the behavior of galactic warps have emerged.

Analysis of 12 galaxies with extended, warped H I disks show the following:

1. The H I layer typically is planar within R_{25}, but warping becomes detectable within $R_{\rm Ho} = R_{26.5}$. Warping within $R_{\rm Ho}$ appears consistent with a common (i.e., straight) line of the nodes (LON) measured in the plane defined by the innermost regions of the galaxies.

2. Warps change character at a transition radius near $R_{\rm Ho}$.

3. For radii larger than $R_{\rm Ho}$, the LON measured in the plane of the inner galaxy advances in the direction of galaxy rotation for successively larger radii. Thus, the nodes lie along leading spirals in this frame of reference.

4. The galaxy kinematics uniquely specify a new reference frame in which there is a common LON for oribits within the transition radius and also a *differently oriented* straight LON for the gas outside the transition radius. This new reference frame is typically inclined by less than 10° to the plane of the inner galaxy.

The lack of a common LON throughout the entire warped disk argues against models that rely on normal bending modes to maintain warp coherence at all radii. Instead, the emerging picture may require galaxy models with two distinct regimes. Behavior in the outer regime is consistent with models that have the LON regressing most rapidly for oribits that are in closest proximity to the flat, stellar disk. In the inner regime, the disk may be settling into a warped mode.

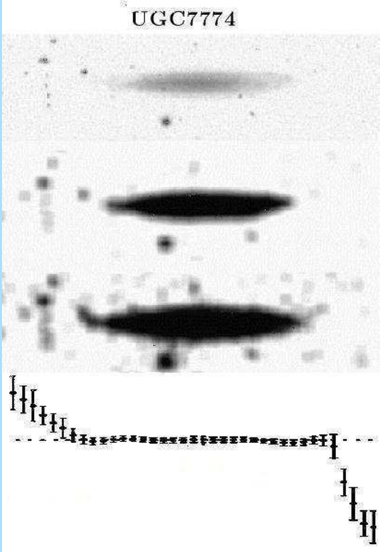
[9] F.H. Briggs, Ap.J. 352, 15 (1990)

The most important aspects of Brigg's rules for the present discussion are:

- The HI layer typically is planar within R_{25}, but warping becomes detectable near $R_{Ho} = R_{26.5}$.
- Warps change character at a transition radius near R_{Ho}.
- The outer warp defines a reference frame.

A recent finding[a] indicates that warps start just beyond the truncation radius.

[a] P.C. van der Kruit, A.&A. 466, 883 (2007)



Properties of warps can be summarized as follows:

- All galaxies with extended HI disks have warps.
- Many galaxies have relatively sharp truncations.
- In edge-on galaxies the HI warps sets in just beyond the truncation radius, for less inclined systems it sets in near the Holmberg radius.
- In many cases the rotation curve shows a feature that indicates that there is at the truncation radius also a sharp drop in mass surface density.
- The onset of the warp is abrupt and discontinuous. and there is a steep slope in HI-surface density at this point.
- Inner disks are extremely flat and the warps define a single "new reference frame".

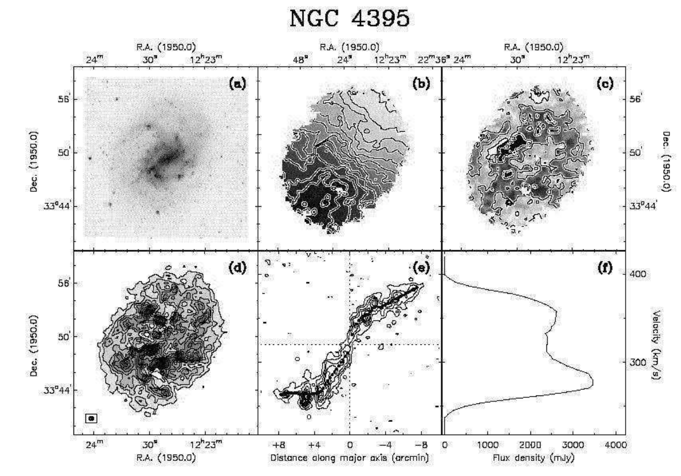
This may mean that the inner stellar disk formed first with a truncation and that the HI in the warp fell in later with another orientation of its angular momentum.

Often spiral galaxies are "lob-sided"[10] in their outer HI, such as NGC4395.

This has been explained as disks that are lying off-center in a dark halo[11].

[10] R.H.M. Schoenmakers, Ph.D. thesis (1999), R.S. Swaters, R.H.M. Schoenmakers, R. Sancisi & T.S. van Albada, Mon.Not.R.A.S. 304, 330 (1999)

[11] S.E. Levine & L.S. Sparke, Ap.J. 496, L13 (1998); E. Noordermeer, L.S. Sparke & S.E. Levine, Mon.Not.R.A.S. 328, 1064 (2001)



(panel c has residual velocities)

Dust and absorption

Holmberg's analysis

The earliest study is by Holmberg[12].

He defined an apparent face-on surface brightness from the apparent magnitude m and and the angular major-axis diameter a

$$\mu'_{\rm obs} = m + 5 \log a$$

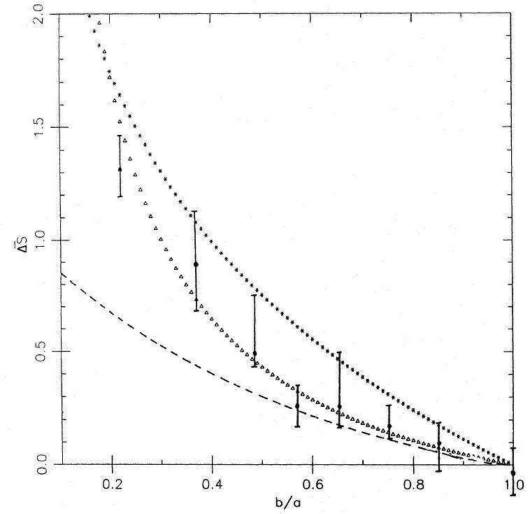
He then plotted this against the axis ratio b/a on the sky.

The inclination i is related to the axis ratio as

$$\sec i = a/b$$

for a not too edge-on disk ($a/b < 3$).

[12] E. Holmberg. Medd. Lund Obs. Ser. 2, No. 136 (1958)


2.0
1.5
1.0
0.5
0.0
$\bar{\Delta S}$
0.2
0.4
0.6
0.8
1.0
b/a

Holmberg's fit to the data (triangles) then is

$$\mu'_{\mathrm{obs}}(i) = \mu'(0) + A_{\mathrm{B}}\{\sec i - 1\}$$

$A_{\mathrm{B}} = 0.40$ mag for Sa-Sb
$A_{\mathrm{B}} = 0.28$ mag for Sc

So his conclusion was that disk of galaxies are not optically thick.

However, it should be realised that Holmberg's fit is not physical, since it is actually that of a dust sheet in front of a stellar disk.

Later with the IRAS satellite is was found that often for galaxies $L_{\mathrm{FIR}}/L_{\mathrm{opt}} \sim 1$ or more.

Realise that for a thin, opaque dust layer in the central plane of stellar disk we expect:

- $A_{\mathrm{B}} = 0.75$ mag.
- No change in color index
- $L_{\mathrm{FIR}}/L_{\mathrm{opt}} \sim 1$

In the Galaxy we are *not* in an optically thick part of the disk.

Extinctions towards the poles are estimated between 0 and 0.2 mag in B.

But there may be denser parts and towards the center absorption may in general increase in galaxies.

Analysis of Disney et al.

Disney *et al.*[13] collected information from various sources, parametrizing it as

$$\mu_{\rm obs}(i) = \mu_\circ - 2.5 n_{\rm eff} \log(a/b)$$

In a completely optically thin disk one expects $n_{\rm eff} = 1$ and in an optically thick disk $n_{\rm eff} \leq 1$.

Then for samples in the Second Reference Catologue (RC2) and the Revised Shapley-Ames Catalogue (RSA) the following values are found for $n_{\rm eff}$:

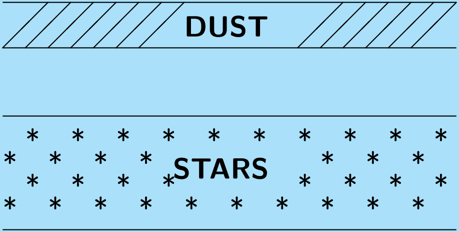
[13] M. Disney, J. Davies & S. Phillipps, Mon.Not.R.A.S. 239, 939 (1989)

Type	Holmberg	RC2	RSA
Sa-Sb	0.46	0.72	0.46
Sbc	0.46	0.68	0.65
Sc	0.65	0.68	0.65
Sd	-	0.68	0.65
Sdm-Im	-	0.96	0.82

So there is certainly evidence for some absorption.

Now look at some simple models of Disney *et al.*

♠ SCREEN MODEL



The dust layer has optical thickness τ, the stellar disk emissivity E^* and thickness T.

The observed surface brightness the becomes

$$L(i) = E^* T \sec i \exp\{-\tau \sec i\}$$

Note that Holmberg's μ' is $L'(i) = L(i)\cos i$, so

$$\mu'(i) = \mu'_\circ + A^\circ_{\mathrm{B}} \sec i = \mu'(0) + A^\circ_{\mathrm{B}}(\sec i - 1)$$

The total face-on absorption becomes

$$A^\circ_{\mathrm{B}} = 1.086\tau$$

For $\tau \ll 1$

$$L(i) = E^* T \sec i$$

The observed surface brightness is $L(\tau, i)$ and the bolometric surface brightness is $L(0,0)\sec i$.

Consider a circular area πa^2, then total luminosity is

$$L_{\rm bol} = \pi a^2 L(0,0) = \pi a^2 E^* T$$

The observed face-on luminosity is

$$L_{\rm opt} = \pi a^2 L(\tau, 0)$$

If the dust re-radiates isotropically

$$L_{\rm FIR} = L_{\rm bol} - L_{\rm opt} = \pi a^2 \{L(0,0) - L(\tau, 0)\}$$

The FIR surface brightness at inclination i then is

$$L_{\rm FIR}(i) = \sec i \{ L(0,0) - L(\tau,0) \}$$

and we can calculate (drop the τ's)

$$\frac{L_{\rm FIR}}{L(i)} = \sec i \frac{E^* T - L(0)}{L(i)}$$

So we get for the Screen Model

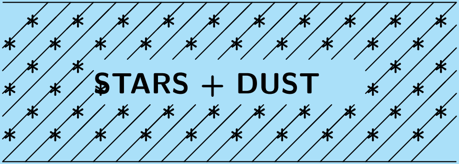
$$\frac{L_{\rm FIR}}{L(i)} = \exp\{\tau \sec i\} - 1$$

For the optically thin case $\tau \ll 1$ this reduces to

$$\frac{L_{\rm FIR}}{L(i)} = \tau \sec i$$

♠ SLAB MODEL

Now make the model more realistic.



The results then become:

$$L(i) = \frac{E^* T}{\tau} \left[1 - \exp\{-\tau \sec i\}\right]$$

$$A_{\mathrm{B}}^{\circ} = -2.5 \log \left\{ \frac{1 - \exp(-\tau)}{\tau} \right\}$$

$$\frac{L_{\mathrm{FIR}}}{L(i)} = \sec i \, \frac{\tau - 1 + \exp\{-\tau\}}{1 - \exp\{-\tau \sec i\}}$$

For the optically thick case $\tau \gg 1$

$$L(i) = \frac{E^* T}{\tau} = \text{ constant}$$

$$A_{\mathrm{B}} = 2.5 \log \tau$$

$$\frac{L_{\mathrm{FIR}}}{L(i)} = (\tau - 1) \sec i$$

For the optically thin case $\tau \ll 1$

$$L(i) = E^* T \sec i$$

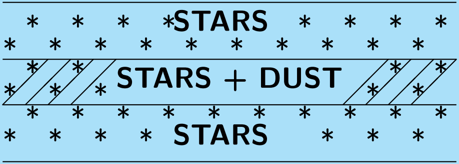
So L' is independent of i.

$$A_{\mathrm{B}} = -2.5 \log \left(1 - \frac{\tau}{2}\right)$$

$$\frac{L_{\mathrm{FIR}}}{L(i)} = \frac{\tau}{2}$$

♠ SANDWICH MODEL

In real galaxies the dust layer is thinner than the stellar disk.



Let the thickness of dust layer be pT. Then

$$L(i) = E^* T \sec i \left[\frac{1-p}{2} \{1 + \exp\left(-\tau \sec i\right)\} + \frac{p}{\tau \sec i} \{1 - \exp\left(-\tau \sec i\right)\} \right]$$

The optically thick case $\tau \gg 1$ now becomes

$$L(i) = E^* T \sec i \, \frac{1-p}{2}$$

$$A_{\rm B} = -2.5 \log \left\{ \frac{1-p}{2} \right\}$$

$$\frac{l_{\rm FIR}}{L(i)} = \sec i \frac{(1+p)\tau - 2p}{(1-p)\tau + 2p}$$

If $p \ll 1$

$$L(i) = \frac{E^* T}{2} \sec i$$

$$A_{\rm B} = 0.753$$

$$\frac{L_{\rm FIR}}{L(i)} = \sec i$$

The optically thin case $\tau \ll 1$ gives

$$L(i) = E^* T \sec i \left\{ 1 - \frac{1-p}{2} \tau \sec i \right\}$$

$$A_{\rm B} = -2.5 \log \left\{ 1 - \frac{1-p}{2} \tau \right\}$$

$$\frac{L_{\rm FIR}}{L(i)} = \frac{\tau}{2}$$

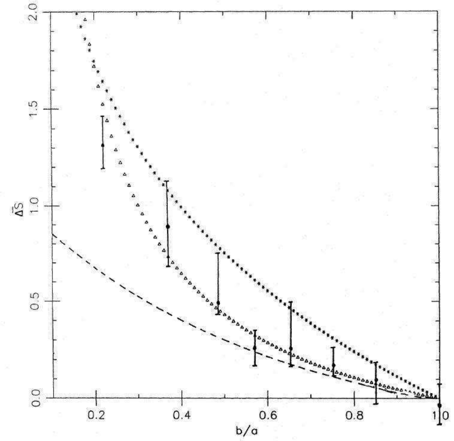
If $p \ll 1$

$$L(i) = E^* T \sec i \left(1 - \frac{\tau}{2} \sec i \right)$$

$$A_{\rm B} = -2.5 \log \left(1 - \frac{\tau}{2} \right)$$

$$\frac{L_{\rm FIR}}{L(i)} = \frac{\tau}{2}$$

- triangles: $\tau < 1$ Screen Model
- stars: $\tau \gg 1$ Slab Model
- dashes: $\tau \gg 1$ Sandwich Model ($p = 0.5$).



The optically thin Slab and Sandwich Models predict no dependence of Holmberg surface brightness on inclination.

So observations are consistent with optically thick models, but the results are very geometry dependent and therefore not yet conclusive.

The near-IR data are also not entirely conclusive. $L_{\rm FIR}$ can be very large compared to $L_{\rm opt}$ if star-formation occurs extensively in very thick, obscured, but localized area's (GMC's)

Disney *et al.* also calculate triplex models as above, which give similar results as these simple models.

We can still extend the analysis by looking at the colors.

In all models we had:

$$L(i) = E^* \, TF(p, \tau, i) \sec i$$

Take

$$\tau_{\rm V} = 0.75 \tau_{\rm B}$$

$$\frac{L_{\rm B}(i)}{L_{\rm V}(i)} = \frac{E^*(B)}{E^*(V)} \frac{F(p, \tau_{\rm B}, i)}{F(p, \tau_{\rm V}, i)}$$

The color change between inclination 0° and 70° then is:

$$\Delta(B - V) = -2.5 \log \left\{ \frac{F(p, \tau_{\rm B}, 70) F(p, \tau_{\rm V}, 0)}{F(p, \tau_{\rm V}, 70) F(p, \tau_{\rm B}, 0)} \right\}$$

For the Sandwich Model we have:

- Optically thin ($\tau \ll 1$):

$$F(p, \tau, i) = 0.5 \Rightarrow \Delta(B - V) = 0$$

- Optically thick ($\tau \gg 1$):

$$F(p, \tau, i) = \frac{1 - p}{2} \Rightarrow \Delta(B - V) = 0$$

Here are some values for $\Delta(B - V)$ as a function of optical thickness.

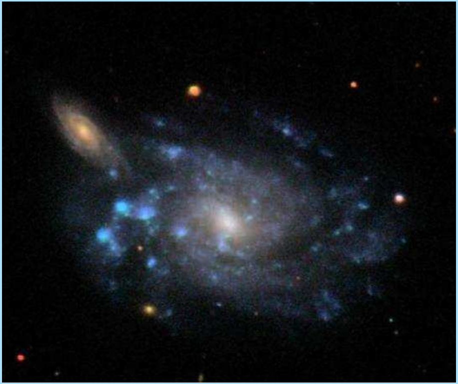
τ	Screen	p = 1	p = 0.5	p = 0.1
0.1	0.05	0.02	0.02	0.02
0.5	0.26	0.09	0.06	0.04
1.0	0.52	0.13	0.04	-0.01
2.0	1.04	0.11	-0.04	-0.07
5.0	2.61	0.02	-0.04	-0.01
10.	5.22	0.02	0.02	0.02

- For small τ B is always more affected than V, so redder with inclination.

- For large τ at high inclination we see only up to the dust, so we have unreddened colors.
However at face-on there is still reddening and disks become bluer with inclination.

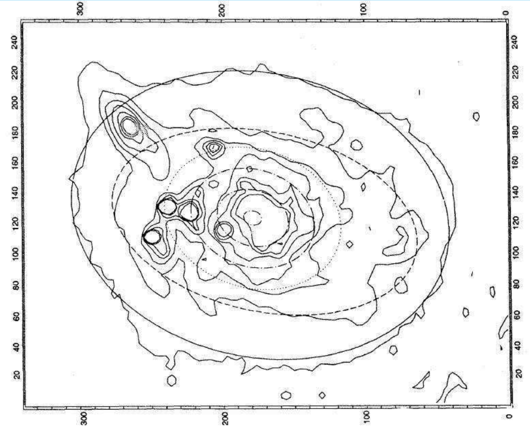
Background galaxies

A very effective test in principle is to look for galaxies seen through disks as in the pair NGC450/UGC807[14].

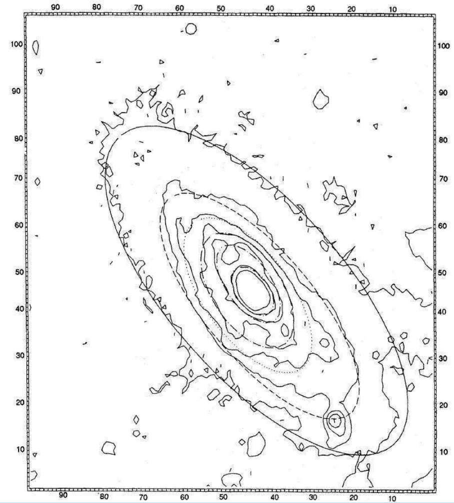


[14] Y. Andredakis & P.C. van der Kruit, A.&A. 265, 396 (1992)

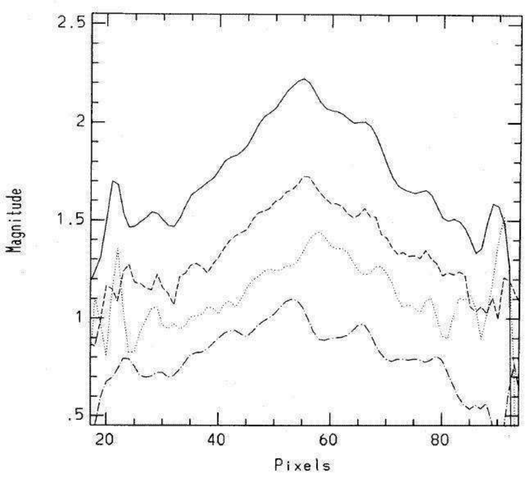
In the photometry we can deduce the surface brightness distribution of NGC 450 in the area of overlap.



Subtraction then gives the "uneffected" image of UGC807.

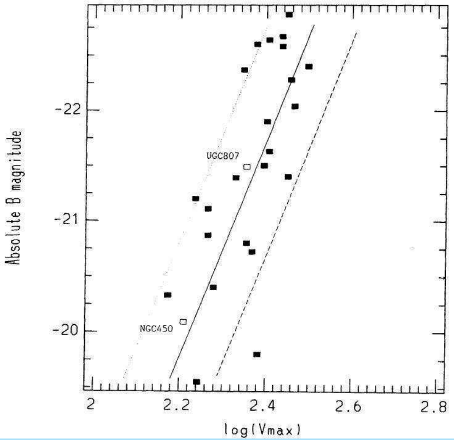


This shows no color changes, so there is no significant gradient in absorption.



Both galaxies conform to the Tully-Fisher relation.

The maximum absorption allowed is 0.3 magnitudes in the V-band.



More sophisticated is to study images of galaxies with the Hubble Space Telescope and identify background galaxies.

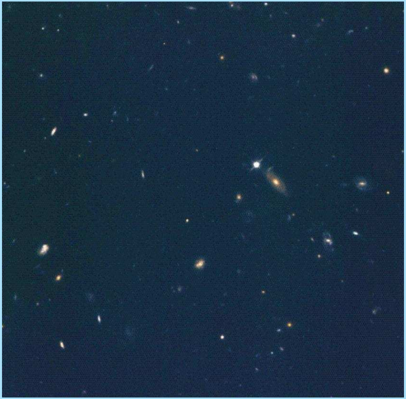
For this one takes images from the HST archive, essentially from the key-project to derive the distance scale through Cepheids and calibration of the TF-relation[15].

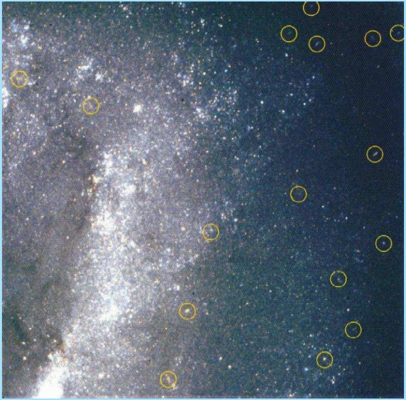
Then the test can be done by adding the Hubble Deep Field (HDF) with the appropriate noise and background level and see what fraction of these galaxies are recovered.

With this synthetic field method[16] evidence for some absorption has been found.

[15] see www.ipac.caltech.edu/H0kp/.

[16] R.A. González, R.J. Allen, B. Dirch, J.C. Ferguson, D. Calzetta & N. Panagia, Ap.J. 506, 152 (1998)

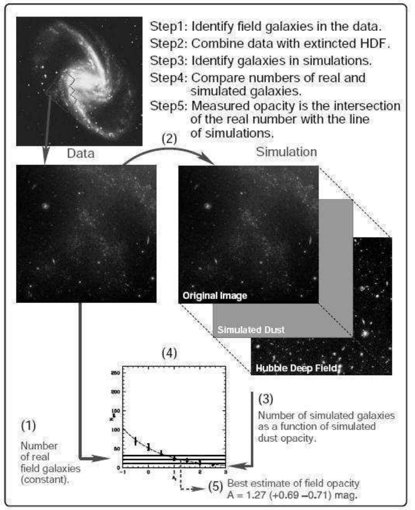




The synthetic field method works as follows. One starts with a set of HST images of nearby galaxies.

This is then compared to images where the HDF has been superposed with various amount of dimming.

The dimming where the same number of galaxies per unit solid angle is found then shows the amount of absorption.

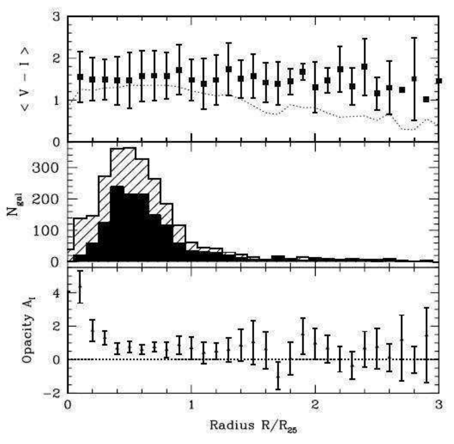


Here is the final result from this project[17]:

Top: Average color of background galaxies in observed fields and in HDF (dotted line) .

Middle: Number of observed galaxies (filled histogram) and in synthessized fields.

Bottom: Inferred extinction.



[17] B.W. Holwerda, Ph.D. Thesis; B.W. Holwerda, R.A. Gonzalez, R.J. Allen & P.C. van der Kruit, Ap.J. 129, 1381 (2004)

STRUCTURE AND DYNAMICS OF GALAXIES

18. Stellar kinematics and spiral structure

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/~vdkruit

Beijing, September 2011

Outline

Spiral structure

Density wave theory

Stochastic star formation model

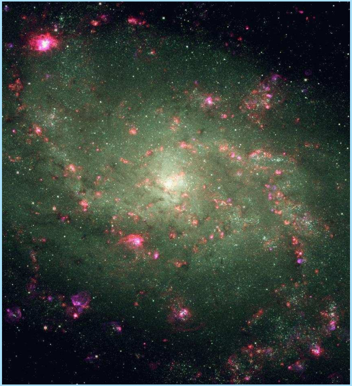
Stellar kinematics

Spiral structure

Density wave theory

We distinguish two types of spiral structure, grand design ...

and flocculent.



A comparative study of these two classes[1] suggests that in grand-design spiral strcuture there seems to be a strong underlying spiral wave in the stellar disk, while not in flocculent ones.

The density wave theory[2] was a response to the "winding dilemma", where material arms would wind up in a matter of 10^8 years or less.

The density wave is a spiral pattern, whose shape does not change with time, and which moves through the stellar and insterstellar disk.

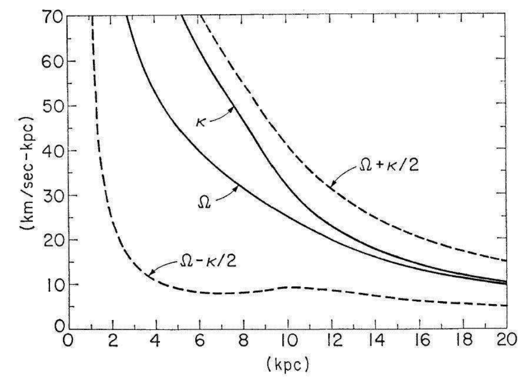
---

[1] B.G. Elmegreen & D.B. Elmegreen, Ap.J.Suppl. 54, 127 (1984)

[2] C.C. Lin & F.H. Shu, Ap.J. 140,646 (1964), C.C. Lin, C. Yuan & F.H. Shu, Ap.J. 155, 721 (1969)

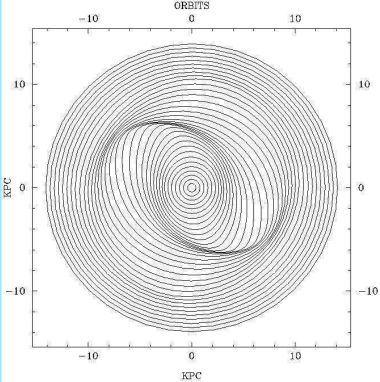
At the basis of a good description we can take the deduction that in the disk of our Galaxy (and in many others) the inner Lindblad resonance $\Omega - \kappa/2$ is fairly constant.

In this resonance a star goes through two epicycles during one revolution around the center. That means it describes a closed oval orbit in a rotating coordinate system with $\Omega - \kappa/2$.



In a disk where this property is constant over most radii we can get the following situation, where the stars are forced in orbits that line up as a spiral pattern.

In a coordinate frame, rotating with the pattern speed $\Omega_{\rm p} = \Omega - \kappa/2$, the spiral pattern remains unchanged.



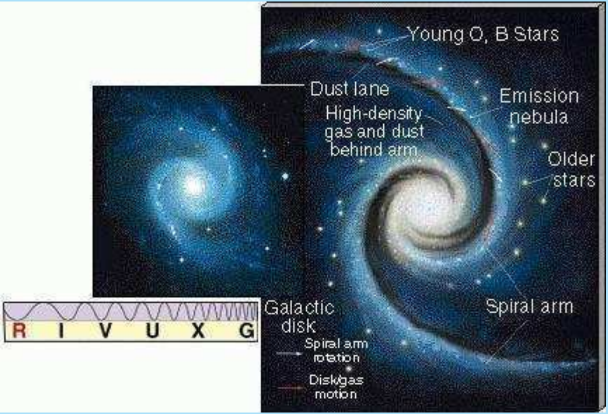
In the original density wave theory the density perturbations maintain themselves. The response of the stars to the perturbed gravitational field by the density concentrations in the arms results in a self-sustaining pattern of density perturbations.

It was realized later by Toomre and others that the dissipation of energy in the waves is quick enough ($\sim 10^8$ years) that rejuvenation is required regularly.

It took until the first part of the seventies, before the underlying wave in the stellar disk was discovered in surface photometry[3].

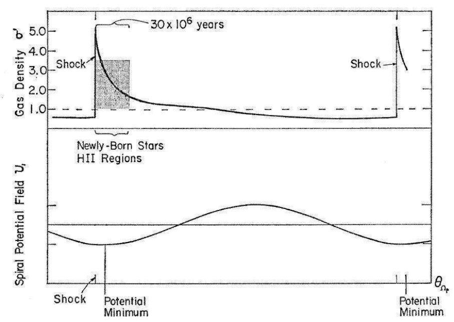
[3] F. Schweizer, Ap.J.Suppl. 31, 313 (1976)

The strongest confirmation came from studies of the interstellar medium.



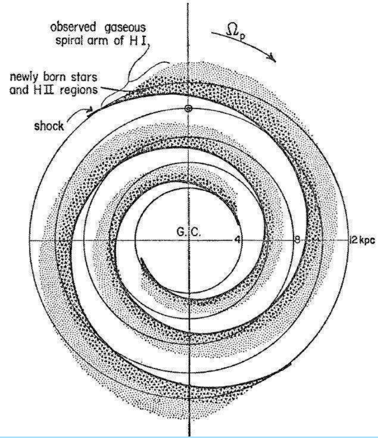
The reponse of the gas and dust is a-linear, since the relative velocities involved are supersonic[4].

This gives shocks at the inner sides of the spiral arms and associated dustlanes and star formation.



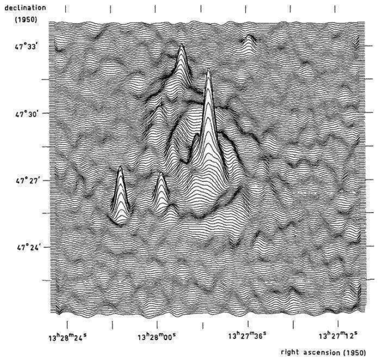
[4]W.W. Roberts, Ap.J. 158, 123 (1969)

The "delay" between dustlanes and HII-regions concerns the time between onset of gravitational instability and birth of MS-stars.

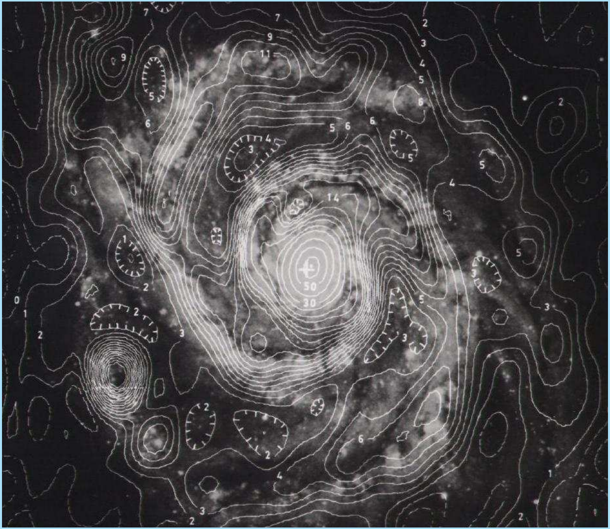


It was also confirmed by radio continuum studies with the new WSRT[5] in M51.

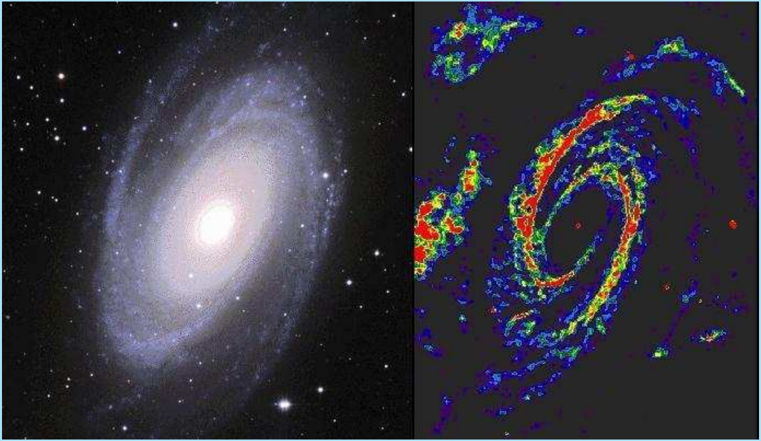
The compression holds at least for the magnetic field and possibly the relativistic electrons, so the synchrotron radiation will be enhanced at the inside of the arms and at the dustlanes.



[5]D.S. Mathewson, P.C. van der Kruit & W.N. Brouw, A.&A. 17, 468 (1972)



The next thing was to try and measure the streaming motions due to the density wave. This was tried in M81 using HI.



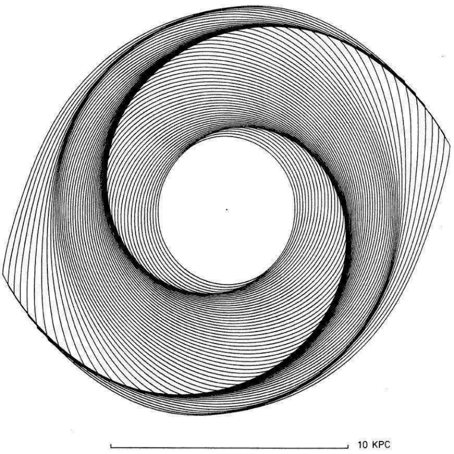
The Ph.D. thesis of H.C.D. Visser[6] analysed this in detail.

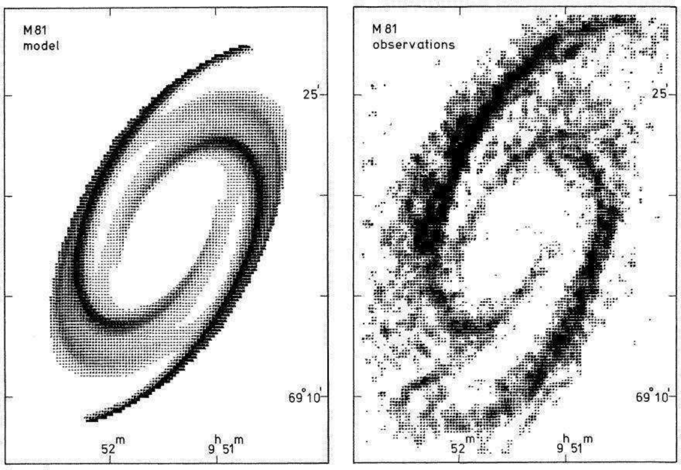
He used the surface photometry of Scheizer and HI-measurements at Westerbork.

With that he was able to find an internally consistent representation of the observations of at the same time both the HI surface density distribution and the HI velocity field.

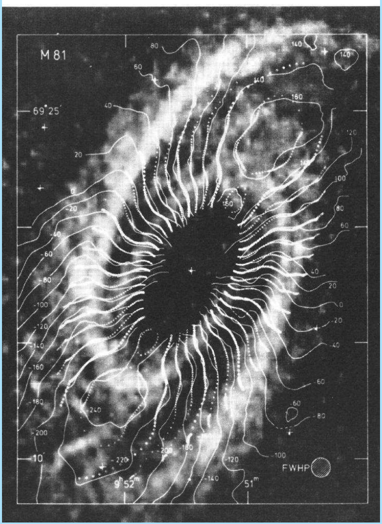
Here are the (non-linear) streamlines of the gas.

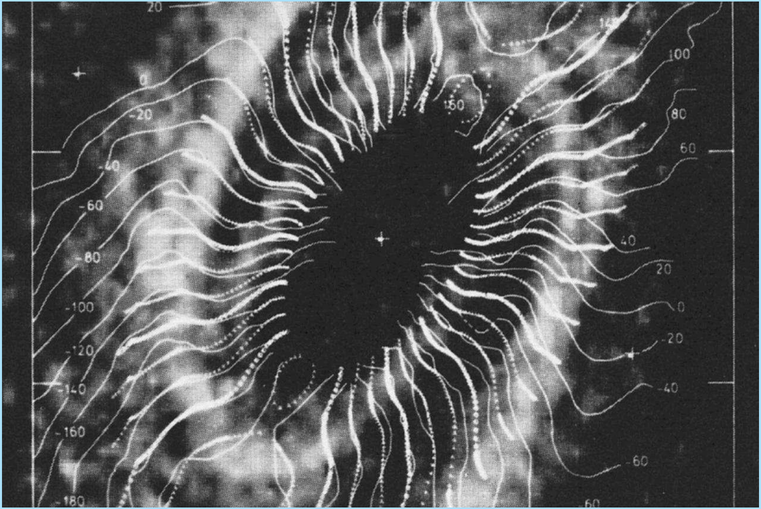
[6] 1978; see also A.&A. 88, 159 (1980)


10 KPC

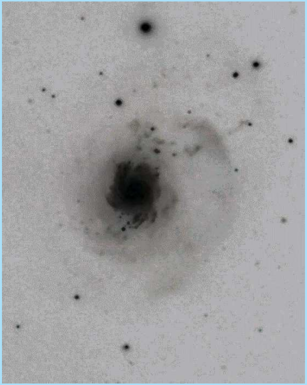

M 81
model
25'
69°10'
52m
9h51m
M 81
observations
25'
69°10'
52m
9h51m

The streaming motions are of the order of 10 km s^{-1}.



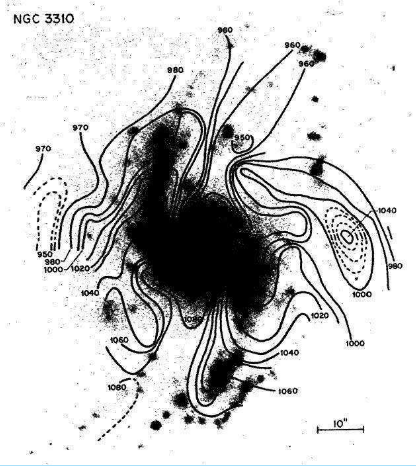

20
0
-20
-40
-60
-80
-100
-120
-140
-160
-180
150
100
80
60
40
20
0
-20
-40
-60
-60

A very exceptional case is the disturbed, star burst galaxy NGC 3310, which is probably an example of a recent merger[7].

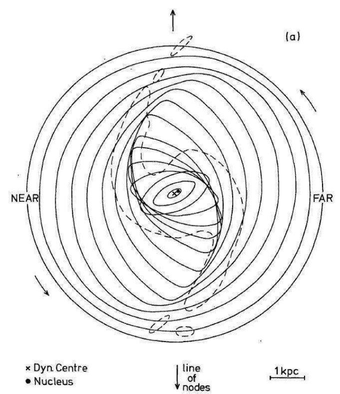


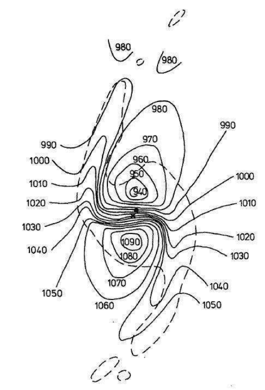
[7] P.C. van der Kruit & A.G. de Bruyn, A.&A. 48, 373 (1976); P.C. van der Kruit, A.&A. 49, 161 (1976)

The velocity field shows strong signs of streaming motions related to the spiral arms.



The streaming motions are here up to a third or so of the rotation velocity.





Stochastic star formation model

Density waves may be generated by tidal interactions, such as in M51 or in NGC 3310, or through Toomre's swing amplification.

The flocculent spiral stucture is probably the result of stochastic self-propagating star formation[8].

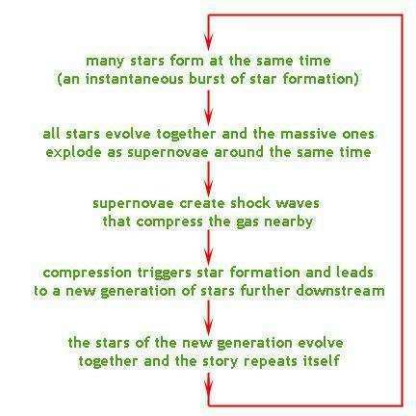
Since the propagation and induced star formation is never 100%, also this will die out unless there is also spontaneous star formation.

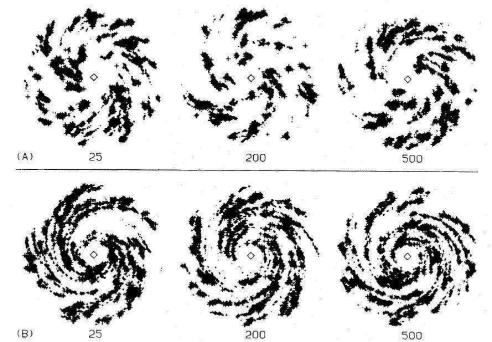
[8] H. Gerola & P.E. Seiden, Ap.J. 223, 129 (1978)

In this model star formation through supernova explosions is postulated to stimulate star formation in the neighborhood.

Such structures are then drawn out by differential rotation into arm-like features.

On the next page some simulations.





It has been suggested[9] that grand-design spiral structure is produced by bars or tidal encounters, while flocculent spiral structure results if the disk is left by itself.

[9] J. Kormendy & C.A. Norman, Ap.J. 233, 539 (1979)

Stellar kinematics

To measure stellar kinematics one needs to analyse absorption line spectra.

The assumption is that the galaxy spectrum is essentially that of a late-G to early K-giant (the "template"), shifted by a radial velocity and broadened by the velocity distribution.

This is based on the fact that the integrated light from an old population is dominated by the stars in the upper part of the Giant Branch.

The fundamental equation is

$$G(\log\lambda) = \alpha\, S(\log\lambda - \delta)\ *\ B$$

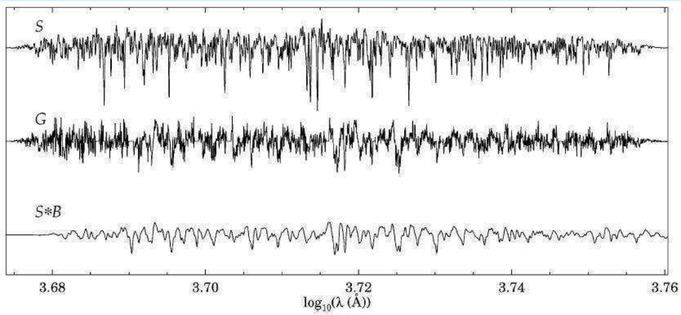
$$\begin{array}{lcl} G(\log\lambda) & = & \text{galaxy spectrum} \\ S(\log\lambda) & = & \text{template spectrum} \\ B & = & \text{broadening function} \\ \delta & = & \text{radial velocity} \\ \langle V^2\rangle^{1/2} & = & \text{velocity dispersion (the second moment of } B\text{)} \end{array}$$

Analysis is therefore exclusively based on Fourier methods[10], using:

$$\tilde{G}(k) = \gamma\ \tilde{T}(k)\cdot\tilde{B}$$

[10] Following the fundamental discussion by S.M. Simkin, A.&A. 31, 129 (1971)

Here is an example[11]

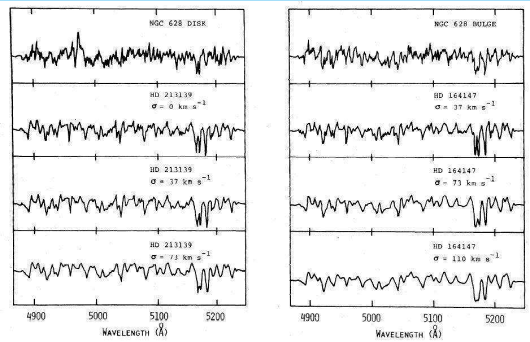


[11] From M. Kregel, P.C. van der Kruit & K.C. Freeman, Mon.Not.R.A.S. 351, 1247 (2004)

An often used part of the spectrum is around 5000Å, where one finds the Mg b triplet and many Fe I lines.

The figure[a] shows at the top galaxy exposures and below broadened spectra of template K-giants.

[a] From van der Kruit & Freeman, Ap.J. 278, 81 (1984)

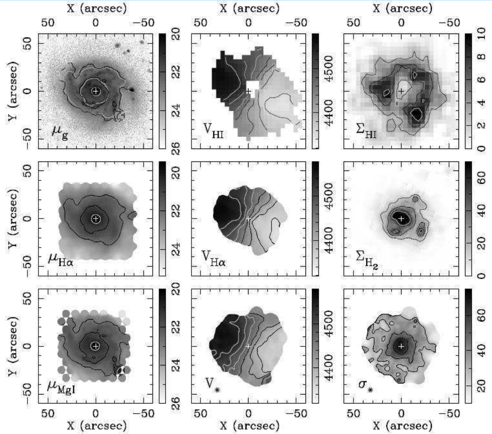


There are three general methods.

- Power spectrum method[12].
  - δ from cross-correlation peak
  - $\langle V^2 \rangle^{1/2}$ from slope of power spectrum
- Fourier quotient method[13].
  - Assume B is a Gaussian
  - Then $\tilde{B}$ is also a Gaussian (but complex)
  - Fit a Gaussian to $\tilde{G}(k)/\tilde{T}(k)$
- Cross-correlation method[14].
  - δ from cross-correlation peak
  - $\langle V^2 \rangle^{1/2}$ from width of cross-correlation peak

[12] G.D. Illingworth & K.C. Freeman, Ap. J. 188, L83 (1974)
[13] due to Paul Schechter; W.L.W. Sargent, P.L. Schechter, A. Boksenberg & K. Shortridge, Ap.J. 212, 326 (1977)
[14] J. Tonry & M. Davis, A.J. 84, 1511 (1979)

The major progress in this area is the use of integral-field units, as in the DiskMass project[15] so that large areas can be observed at once (and compared to other data).



[15] M.A. Bershady, M.A.W. Verheijen, et al., Ap.J. 716, 198 & 234 (2010)

STRUCTURE AND DYNAMICS OF GALAXIES

19. Dynamics of spiral galaxies: Stars

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/∼vdkruit

Beijing, September 2011

Outline

Stellar velocity dispersions

Z-velocity dispersion

If disks have constant mass-to-light ratios M/L, the density can be described by

$$\rho(R,z) = \rho(0,0)\,\exp\left(-R/h\right)\,\mathrm{sech}^2(z/z_\circ)$$

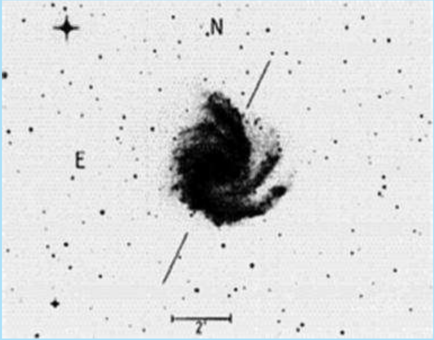
The vertical velocity dispersion then is

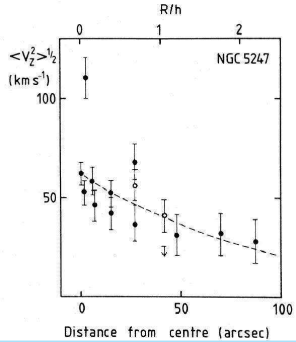
$$\langle V_{\mathrm{z}}^2\rangle^{1/2} = \sqrt{2\pi G\rho(R,0)}z_\circ$$

and it is expected that

$$\langle V_{\mathrm{z}}^2\rangle^{1/2} \propto \exp\left(-R/2h\right)$$

This can be tested by observations in face-on systems, e.g. NGC 5247[1].





[1]P.C. van der Kruit & K.C. Freeman 1986, Ap.J. 303, 556 (1968)

The fit is

$$\langle V_z^2 \rangle^{1/2} = (62 \pm 7) \exp\left[-(0.42 \pm 0.10)\ R/h\right] \mathrm{km\ s}^{-1}$$

This is consistent with M/L about constant.

It has been confirmed in various studies since then.[2]

[2] See recent review by P.C. van der Kruit & K.C. Freeman, Ann. Rev. A.&A. 49, 301, 2011

R- and θ -velocity dispersions

From fundamental kinematics we have

$$\frac{\langle (V_\theta - V_{\mathrm{t}})^2 \rangle}{\langle V_{\mathrm{R}}^2 \rangle} = \frac{B}{B - A}$$

So, if we know the rotation curve we know the ratio of the radial and tangential velocity dispersion.

The other property to consider is the asymmetric drift.

The hydronamic equation can be written as

$$-K_{\mathrm{R}} = V_{\mathrm{t}}^2 - \langle V_{\mathrm{R}}^2 \rangle \frac{\partial}{\partial R} \ln(\nu \langle V_{\mathrm{R}}^2 \rangle) + \frac{1}{R} \left\{ \langle V_{\mathrm{R}}^2 \rangle - \langle (V_\theta - V_{\mathrm{t}})^2 \rangle + \langle V_{\mathrm{z}} V_{\mathrm{R}} \rangle \frac{\partial}{\partial z} (\ln \nu \langle V_{\mathrm{z}} V_{\mathrm{R}} \rangle) \right\}$$

Poisson's equation is

$$\frac{\partial K_{\mathrm{R}}}{\partial R} + \frac{K_{\mathrm{R}}}{R} + \frac{\partial K_{\mathrm{z}}}{\partial z} = -4\pi G \rho$$

For small z it can be shown that

$$\frac{\partial K_{\mathrm{R}}}{\partial R} + \frac{K_{\mathrm{R}}}{R} = 2(A - B)(A + B)$$

and for a flat rotation curve $A = -B$, so that

$$\frac{\partial K_{\mathrm{z}}}{\partial z} = -4\pi G \rho$$

Then

$$\langle V_{\mathrm{z}} V_{\mathrm{R}} \rangle = 0$$

Obviously we have

$$K_{\mathrm{R}} = V_{\mathrm{rot}}^2 / R$$

For an exponential disk with constant M/L

$$\frac{\partial}{\partial R} \ln \nu = -\frac{1}{h}$$

The asymmetric drift equation then becomes

$$V_{\rm rot}^2 - V_{\rm t}^2 = \langle V_{\rm R}^2 \rangle \left[\frac{R}{h} - R \frac{\partial}{\partial R} \ln \langle V_{\rm R}^2 \rangle - \left\{ 1 - \frac{B}{B-A} \right\} \right]$$

There are now two possibilities for observing. The first is to measure $\langle V_{\rm R}^2 \rangle^{1/2}$ directly from spectra.

The difficulty is the line-of-sight integration. This has to be treated by modeling as was done in the edge-on galaxy NGC 5170[3].

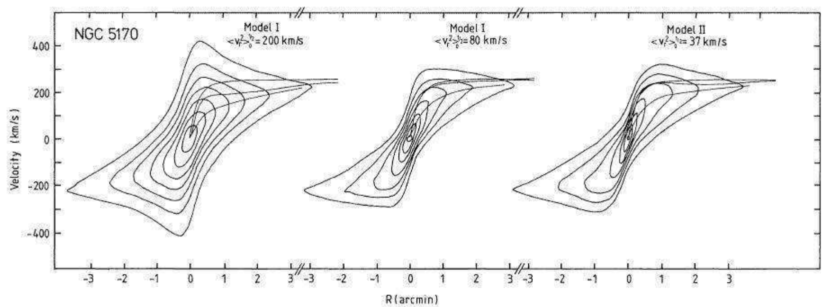
The profiles now have become asymmetric.

[3] R. Bottema, P.C. van der Kruit & K.C. Freeman, Ap.J. 178. 77 (1987)

Using an estimate of the circular motion from the HI-rotation curve one can calculate the profiles in a stellar "l,V-diagram".

To do this one needs an assumed radial variation of the velocity dispersion, the rotation curve (and from that the Oort constants) and the density distribution of the stars.

In the figure here we see a few such simulations. The three lines in each panel are form top to bottom: the circular motion from HI-observations, the stellar rotation velocity and peaks of Gaussians fitted to the resulting profiles.


NGC 5170
Model I
$<v_r^2>_0^{1/2} = 200$ km/s
Model I
$<v_r^2>_0^{1/2} = 80$ km/s
Model II
$<v_r^2>_0^{1/2} = 37$ km/s
Velocity (km/s)
400
200
0
-200
-400
-3 -2 -1 0 1 2 3
R (arcmin)

The second option is to measure the asymmetric drift.

The relevant equation was

$$V_{\mathrm{rot}}^2 - V_{\mathrm{t}}^2 = \langle V_{\mathrm{R}}^2 \rangle \left[\frac{R}{h} - R \frac{\partial}{\partial R} \ln \langle V_{\mathrm{R}}^2 \rangle - \left\{ 1 - \frac{B}{B-A} \right\} \right]$$

So we see that we need to measure:

- V_{rot}, A and B from HI-synthesis or emission line spectroscopy.
- V_{t} from absorption line spectroscopy.
- h from surface photometry.

For a flat rotation curve:

$$\frac{B}{B-A} = 0.5 \text{ and } \kappa^2 = \frac{2V_{\rm rot}^2}{R^2}$$

For small asymmetric drift:

$$V_{\rm rot}^2 - V_{\rm t}^2 \approx 2V_{\rm rot}(V_{\rm rot} - V_{\rm t})$$

Now consider two possibilities:

- Model I with $\langle V_{\rm R}^2\rangle/\langle V_{\rm z}^2\rangle$ constant. Then

$$\langle V_{\rm R}^2\rangle^{1/2} \propto \exp\left(-R/2h\right)$$

$$V_{\rm rot} - V_{\rm t} = \frac{\langle V_{\rm R}^2\rangle}{2V_{\rm rot}}\left(\frac{2R}{h} - 0.5\right)$$

- Model II with Q constant. Then

$$\langle V_{\rm R}^2\rangle^{1/2} \propto R\exp\left(-R/h\right)$$

$$V_{\rm rot} - V_{\rm t} = \frac{\langle V_{\rm R}^2\rangle}{2V_{\rm rot}}\left(\frac{3R}{h} - 2.5\right)$$

How different are these models? For comparison calculate a Q (arbitrarily set to unity at one scalelength) for the first model:

R/h	Q
1.0	1.17
1.5	**1.00**
2.0	0.96
3.0	1.06
4.0	1.31
5.0	1.73

We see that the models are really not different up to four h.

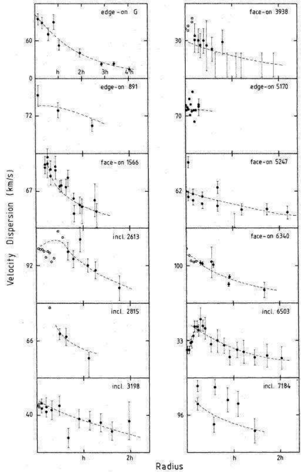
Numerical experiments on dynamics of stellar disks give $Q \sim 1.5 - 2.0$ at all radii.

The Bottema relations

R. Bottema[a] observed stellar velocity dispersions in a set of 12 galaxies.

He then defined as fiducial values the radial velocity dispersion at one scalelength for inclined systems and the vertical velocity dispersion in the center for face-on systems.

This difference should roughly correct for the ratio between these dispersions.



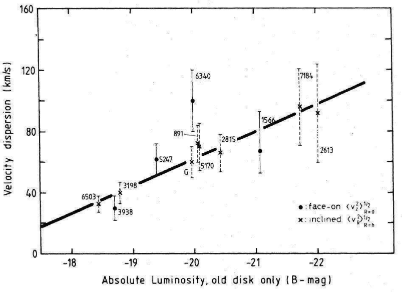
---

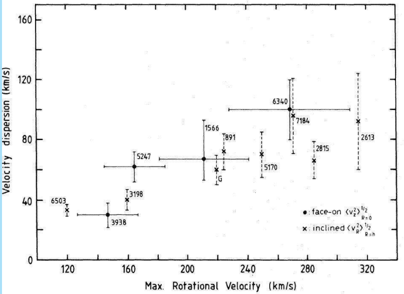
[a] Ph.D. thesis (1995); Bottema, A.&A. 275, 16 (1993)

He then found the following relations

$$\langle V_{\rm R}^2 \rangle_{\rm R=h}^{1/2} = \langle V_{\rm z}^2 \rangle_{\rm R=0}^{1/2} = -17 \times M_{\rm B} - 279 \ \ {\rm km/s}$$

$$\langle V_{\rm R}^2 \rangle_{\rm R=h}^{1/2} = \langle V_{\rm z}^2 \rangle_{\rm R=0}^{1/2} = 0.29 V_{\rm rot} \ \ {\rm km/s}$$





Can we understand these relations?

From the definition of Q we have

$$Q \propto \langle V_{\rm R}^2 \rangle^{1/2} \kappa \sigma^{-1}$$

For a flat rotation curve

$$\kappa \propto V_{\rm rot} R^{-1}$$

An exponential disk has

$$\sigma \propto \mu_\circ (M/L) \exp\left(-R/h\right)$$

Combining these equations gives

$$\langle V_{\rm R}^2 \rangle_{\rm h}^{1/2} \propto \mu_\circ (M/L) Q h V_{\rm rot}^{-1}$$

Now $L \propto \mu_\circ h^2$ and the Tully-Fisher relation gives $L \propto V_{\rm rot}^n$ with $n \approx 4$, so

$$\langle V_{\rm R}^2 \rangle_{\rm h}^{1/2} \propto \mu_\circ (M/L) Q V_{\rm rot} \propto \mu_\circ (M/L) Q L^{1/4}$$

So we expect that $\mu_\circ$, M/L and Q or at least their product are constant between disks.

Implications for maximum disk and stability

We had for hydrostatic equilibrium at the center

$$\langle V_{\mathrm{z}}^2 \rangle_{\mathrm{R=0}}^{1/2} = (2.3 \pm 0.1)\sqrt{G\sigma_{\circ} z_{\mathrm{e}}}$$

$\sigma_{\circ}$ is the central surface density and the range in the constant results from the choice of n.
The maximum rotation velocity of the exponential disk then is

$$_{\mathrm{disk}} = 0.88\sqrt{\pi G \sigma_{\circ} h} = (0.69 \pm 0.03)\langle V_{\mathrm{z}}^2 \rangle_{\mathrm{R=0}}^{1/2} \sqrt{\frac{h}{z_{\mathrm{e}}}}$$

With the Bottema relation between this central velocity dispersion and the maximum observed rotation velocity we get

$$\frac{V_{\rm disk}}{V_{\rm rot}} = (0.21 \pm 0.08)\sqrt{\frac{h}{z_{\rm e}}}$$

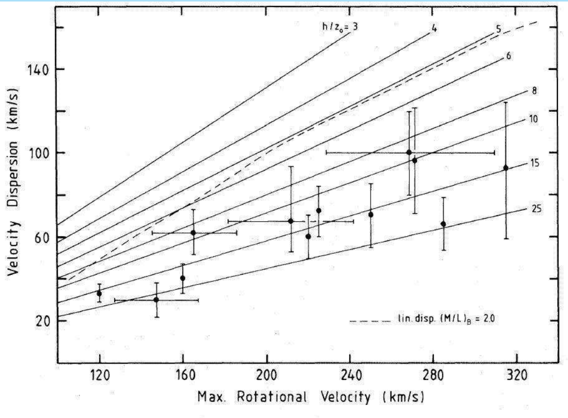
Analysis of a sample of edge-on galaxies gives for the ratio of scaleparameters 7.3 ± 2.2[4], so that

$$\frac{V_{\rm disk}}{V_{\rm rot}} = (0.57 \pm 0.22)$$

So disks in general are not maximum disk.

[4] M. Kregel, P.C. van der Kruit & R. de Grijs, Mon.Not.R.A.S. 334, 646 (2002)

Bottema[5] first showed with this argument that his relations implied that for maximum disk situations the stellar disks should be much flatter than observed.



[5] R. Bottema, A.&A. 275, 16 (1993)

For a flat rotation curve we have

$$\kappa = 2\sqrt{B(B-A)} = \sqrt{2}\frac{V_{\rm rot}}{R}$$

From the definition of Q and applying at $R = h$ we get

$$\langle V_{\rm R}^2 \rangle_{\rm R=h}^{1/2} = \frac{3.36\,G}{\sqrt{2}}\,Q\,\frac{\sigma(R=h)h}{V_{\rm rot}}$$

Using hydrostatic equilibrium (also at $R = h$) gives[6]

$$\frac{\langle V_{\rm z}^2 \rangle^{1/2}}{\langle V_{\rm R}^2 \rangle^{1/2}} = \sqrt{\frac{(7.2 \pm 2.5)}{Q}\,\frac{z_{\rm e}}{h}}$$

[6] P.C. van der Kruit & R. de Grijs, A.&A. 352, 129 (1999)

In the solar neighborhoud this axis ratio of the velocity ellisoid is ~ 0.5[7] and for the Galaxy we have $z_{\rm e} \sim 0.35$ kpc and $h \sim 4$ kpc, so that

$$Q \sim 2.5.$$

Taking all data and methods together it is found that this applies in all galaxies; disks are locally stable according to the Toomre criterion.
Numerical studies give such values for Q when disks are marginaly stable.

[7] W. Dehnen & J. Binney, Mon.Not.R.A.S. 298, 387 (1998)

Global stability

Swing amplification[8] of disturbances occurs as a result of the shear in rotating disks and turns these disturbances into growing trailing spiral waves.

It can be formulated in a criterion for prevention of this instability[9]

$$X = \frac{R\kappa^2}{2\pi G m \sigma(R)} \gtrsim 3$$

Here m is the number of spiral arms.

[8] A. Toomre, in a Cambridge conference on Structure and Evolution of Galaxies (1981)

[9] J.R. Sellwood, IAU Symp. 100, 197 (1983)

For a flat rotation curve this can be rewritten as

$$\frac{QV_{\rm rot}}{\langle V_{\rm R}^2\rangle^{1/2}} \gtrsim 3.97m$$

and with Bottema's relation it translates into

$$Q \gtrsim 1.1m$$

To prevent strong asymmetric $m = 1$ or bar-like $m = 2$ instabilities we require $Q \gtrsim 2$.

Numerical studies have indicated that disks with velocity dispersions as observed show global instabilities when evolving by themselves.

Disks can be stabilised by massive halos and therefore global stability requires that the disk mass has to be less than a certain fraction of the total mass, according to the criterion[10]

$$Y = V_{\rm rot} \left(\frac{h}{GM_{\rm disk}} \right)^{1/2} \gtrsim 1.1$$

This implies that within $R_{\rm max}$ the mass in the halo $M_{\rm halo} > 75\%$. This is also not true for maximum disk.

[10] G. Efstathiou, G. Lake & J. Negroponte, Mon.Not.R.A.S. 199, 1069 (1982)

The criterion can be rewritten as

$$Y = 0.615 \left[\frac{QRV_{\rm rot}}{h\langle V_{\rm R}^2 \rangle^{1/2}} \right]^{1/2} \exp \left(-\frac{R}{2h} \right) \gtrsim 1.1$$

Evaluating this at $R = h$ and using the Bottema relation gives

$$Q \gtrsim 2$$

STRUCTURE AND DYNAMICS OF GALAXIES

20. Elliptical galaxies: Global dynamics

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/∼vdkruit

Beijing, September 2011

Outline

Fundamental Plane

Rotation and shapes

- Flattening of oblate spheroids
- $V_{\rm m}/\bar{\sigma} - \epsilon$ relation and triaxiality
- Detailed kinematics

Fundamental Plane

With Fish's law (constant central surface brightness) and constant M/L then follows the Faber-Jackson relation[1] between luminosity L and stellar velocity dispersion σ:

$$L \propto \sigma^4$$

This is equivalent to the Tully-Fisher relation for spirals.

There is also a relation between diameter D_Σ (the radius at which the mean surface brightness is 20.75 mag arcsec^{-2}) and the velocity dispersion[2]:

$$D_\Sigma \propto \sigma^{4/3}$$

[1] S.M. Faber & R.E. Jackson, Ap.J. 204, 668 (1976)
[2] A. Dressler *et al.*, Ap.J. 313, 42 (1987)

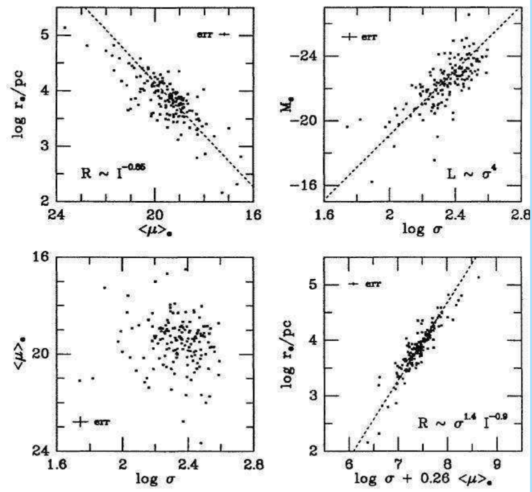
This can be used to decrease the scatter in the FJ-relation by including surface brightness ($\langle SB_{\rm e} \rangle$ = mean surface brightness within the effective radius) as a second parameter

$$L \propto \sigma^{2.65} \langle SB_{\rm e} \rangle^{-0.65}.$$

The "fundamental plane" of elliptical galaxies is a relation between some consistently defined radius (e.g. core radius) R, the observed central velocity dispersion σ and a consistently defined surface brightness I[3]:

$$R \propto \sigma^{1.4 \pm 0.15} I^{-0.9 \pm 0.1}$$

[3] see J. Kormendy & G. Djorgovski, Ann.Rev.Astron.Astrophys. 27, 235 (1989)


log r_e/pc
5
4
3
2
24
20
16
$\langle\mu\rangle_e$
err
R ~ $I^{-0.85}$
M_e
−24
−20
−16
1.6
2
2.4
2.8
log σ
err
L ~ σ^4
16
20
24
$\langle\mu\rangle_e$
1.6
2
2.4
2.8
log σ
err
log r_e/pc
5
4
3
2
6
7
8
9
log σ + 0.26 $\langle\mu\rangle_e$
err
R ~ $\sigma^{1.4}$ $I^{-0.9}$

In broad terms the Fundamental Plane can be understood as follows.

For equilibrium the Virial Theorem states that

$$2T_{\rm k} + \Omega = 0$$

where $T_{\rm k}$ is the total kinetic energy and Ω the potential energy.

The kinetic energy is proportional to MV^2 and the potential energy to M^2/R. Here M is the total mass, V a typical internal velocity and R some characteristic radius.

All the information on the detailed density and velocity structure is in the proportionality constants.

Thus we have

$$M \propto RV^2$$

For elliptical galaxies the kinetic energy is dominated by that in random motions rather then rotation. So for V we will take the mean velocity dispersion[4] σ.

With the mass-to-light ratio M/L, we replace M with $L(M/L)$ with L the total luminosity. For R we take a typical radius such as the effective radius; then we get

$$R \propto L\left(\frac{M}{L}\right)\sigma^2$$

[4] If σ is the observed line-of-sight velocity dispersion, the typical velocity is actually the three-dimensional velocity dispersion 3σ.

If I is the mean surface brightness within R we have $I \propto LR^{-2}$ and

$$R \propto \sigma^2 I^{-1} \left(\frac{M}{L}\right)^{-1}$$

The observed FP was

$$R \propto \sigma^{1.4 \pm 0.15} I^{-0.9 \pm 0.1}$$

The coefficients are close to the observed ones. Differences arise because of variations in actual structural parameters and possible dependence of M/L on M and/or σ.

Rotation and shapes

Flattening of oblate spheroids

If we consider elliptical galaxies to be oblate spheroids, flattened by rotation we can estimate how much rotation is needed using the virial equation.

Let the spheroid be flattened along the z-axis. Then the symmetry with respect to this axis requires

$$\langle V_{\rm R} \rangle = \langle V_{\rm z} \rangle = \langle V_{\rm R} V_\theta \rangle = \langle V_{\rm z} V_\theta \rangle = 0$$

The rotational velocity is $\langle V_\theta \rangle$.

Start with the motions tensor

$$T_{ij} = \frac{1}{2} \int \bar{v}_i . \bar{v}_j d^3 x$$

We have

$$\langle v_{\rm x}\rangle = \langle V_\theta\rangle \sin\theta \quad ; \quad \langle v_{\rm y}\rangle = \langle V_\theta\rangle \cos\theta$$

Then

$$\begin{aligned} T_{\rm xy} &= \tfrac{1}{2}\int \rho\langle v_{\rm x}\rangle\langle v_{\rm y}\rangle d^3x \\ &= \tfrac{1}{2}\int_0^{2\pi}\int_0^{\infty}\int_{-\infty}^{\infty} \rho(R,z)\langle V_\theta\rangle^2 \sin\theta\cos\theta \; dz \; dR \; d\theta \\ &= 0 \end{aligned}$$

since

$$\int_0^{2\pi} \sin\theta\cos\theta d\theta = \tfrac{1}{2}\int_0^{2\pi} \sin(2\theta)d\theta = \tfrac{1}{2}\sin^2(\theta)\,\Big|_0^{2\pi} = 0$$

Similarly, *all* non-diagonal elements of the tensors T_{ij}, Π_{ij} and W_{ij} can be shown to be equal to zero.

Them because of symmetry in the system we must also have

$$T_{xx} = T_{yy} \quad ; \quad \Pi_{xx} = \Pi_{yy} \quad ; \quad W_{xx} = W_{yy}$$

So the only non-trivial virial equations are

$$2T_{xx} + \Pi_{xx} + W_{xx} = 0 \quad ; \quad 2T_{zz} + \Pi_{zz} + W_{zz} = 0$$

So

$$\frac{2T_{xx} + \Pi_{xx}}{2T_{zz} + \Pi_{zz}} = \frac{W_{xx}}{W_{zz}}$$

The ratio W_{xx}/W_{zz} for density distributions with surfaces of equal density being confocal ellipsoids can be shown to be independent of the actual radial dependence of the density. I illustrate that now.

Assume that the axis ratio is c/a and therefore the excentricity

$$e = \sqrt{1 - \frac{c^2}{a^2}}$$

Let the density along the major axis be $\rho(R)$. Define

$$\alpha(R, z) = R^2 + \frac{z^2}{1 - e^2}$$

Then inside the spheroid with radius a the forces and potential are

$$K_{\rm R} = -\frac{4\pi G\sqrt{1 - e^2}}{e^3} R \int_0^{\sin^{-1} e} \rho(\alpha) \sin^2 \beta d\beta$$

$$K_z = -\frac{4\pi G\sqrt{1-e^2}}{e^3} z \int_0^{\sin^{-1} e} \rho(\alpha) \tan^2\beta d\beta$$

$$\Phi(R,z) = \frac{4\pi G\sqrt{1-e^2}}{e}\left[\int_0^{\delta} \rho(\alpha)\alpha\beta d\alpha + \sin^{-1} e \int_{\delta}^{a} \rho(\alpha)\alpha d\alpha\right]$$

Here

$$\delta^2 = R^2 + \frac{z^2}{1-e^2}$$

and

$$\alpha^2 = \frac{R^2\sin^2\beta + z^2\tan^2\beta}{e^2}$$

With partial integration we may write in the equation for $K_{\rm R}$

$$\int_0^{\sin^{-1} e} \rho \sin^2 \beta d\beta = \rho B_1 - \int_0^{\sin^{-1} e} \frac{\partial \rho}{\partial \beta} d\beta$$

with

$$B_1 = \int_0^{\sin^{-1} e} \sin^2 \beta d\beta = \tfrac{1}{2}(\beta - \sin \beta \cos \beta)\Big|_0^{sin^{-1} e}$$

This is a constant and then

$$K_{\rm R} = -\frac{4\pi G \sqrt{1-e^2}}{e^3} R B_1 \left[\rho - \int_0^{\sin^{-1} e} \frac{\partial \rho}{\partial \beta} d\beta \right]$$

Similarly

$$K_{\mathrm{z}} = -\frac{4\pi G\sqrt{1-e^2}}{e^3} z B_2 \left[\rho - \int_0^{\sin^{-1} e} \frac{\partial \rho}{\partial \beta} d\beta \right]$$

with

$$B_2 = \int_0^{\sin^{-1} e} \tan^2 \beta d\beta = (-\beta + \tan\beta)|_0^{sin^{-1} e}$$

Now remember that

$$W_{RR} = -\int R \frac{\partial \Phi}{\partial R} d^3x = \int R K_{\mathrm{R}} d^3x$$

$$W_{zz} = -\int z \frac{\partial \Phi}{\partial z} d^3x = \int z K_{\mathrm{z}} d^3x$$

So in the ratio W_{xx}/W_{zz} the dependence on the functional form of ρ disappears[5].

In fact, to a good approximation, for oblate bodies we have then

$$\frac{2T_{xx} + \Pi_{xx}}{2T_{zz} + \Pi_{zz}} = \frac{W_{xx}}{W_{zz}} \propto \left(\frac{c}{a}\right)^{-0.9}$$

Now consider the cases where the system is either rotating or not or has an isotropic or anisotropic velocity distribution.

[5] The actual ratio is related to parameters in Table 2-1 of Binney & Tremaine.

A. Isotropic and rotating.

Then the velocity dispersion σ is independent of direction. But it may vary with the ellipsoidal surface it is on and therefore we use a density-weighted rms (one-dimensional) velocity dispersion $\bar{\sigma}$. So, if the total mass is M

$$\Pi_{xx} = \int \rho \sigma_{xx}^2 d^3x = M\bar{\sigma}^2 = \Pi_{zz}$$

Say, the density-weighted rotation velocity (around the z-axis) is $\bar{V}$; then $v_x^2 = \frac{1}{2}\bar{V}^2$, and we get

$$T_{zz} = 0$$

$$T_{xx} = \frac{1}{2} \int \rho v_x^2 d^3x = \frac{1}{4} M\bar{V}^2 = T_{yy}$$

Therefore

$$\frac{\frac{1}{2}M\bar{V}^2 + M\bar{\sigma}^2}{M\bar{\sigma}^2} = \left(\frac{c}{a}\right)^{-0.9}$$

This can be reduced to

$$\frac{\bar{V}}{\bar{\sigma}} = \sqrt{2\left[\left(\frac{c}{a}\right)^{-0.9} - 1\right]}$$

This is interesting, since it shows that a large amount of rotation is necessary to give rise to flatterning. E.g. for a rather modest flattening of $c/a = 0.7$ one needs $\bar{V} \sim 0.9\bar{\sigma}$.

B. Anisotropic and non-rotating

Then $T_{xx} = 0$ and $\Pi_{xx} = M\bar{\sigma}_{xx}^2$, $\Pi_{zz} = M\bar{\sigma}_{zz}^2$

This gives

$$\frac{\bar{\sigma}_{zz}}{\bar{\sigma}_{xx}} \sim \left(\frac{c}{a}\right)^{-0.9}$$

For the same modest flattening of $c/a = 0.7$ one now needs only a small anisotropy $\bar{\sigma}_{zz}/\bar{\sigma}_{xx} \sim 0.85$.

C. Anisotropic and rotating

Write

$$\Pi_{zz} = (1-\delta)\Pi_{xx} = (1-\delta)M\bar{\sigma}^2$$

We have again $T_{zz} = 0$ and $2T_{xx} = \frac{1}{2}M\bar{V}^2$.

Then

$$\frac{\bar{V}}{\bar{\sigma}} = \sqrt{2\left[(1-\delta)\left(\frac{c}{a}\right)^{-0.9} - 1\right]}$$

This would mean that we can expect a relation between $\bar{V}/\bar{\sigma}$ and the ellipticity $\epsilon = 1 - (c/a)$ in elliptical galaxies.

However, we observe these systems from random orientations and see an apparent flattening, a projected rotation and the integrated velocity dispersion along the line-of-sight.

It turns out that this only shifts the galaxies that are oblate, isotropic rotators in the *apparent* $(V_{\rm m}/\bar{\sigma} - \epsilon)$-plane roughly along the line of the correlation[6].

So we can compare the observations with the predictions from the anisotropic, rotating case.

[6] See Binney & Tremaine, section 4.3 (page 217)

$V_{\rm m}/\bar{\sigma} - \epsilon$ relation and triaxiality

Originally elliptical galaxies were thought to be simple systems, mainly supported by random motions and flattened by rotation.

The rotation turned out to be too small to provide the flattening so this had to be due to anisotropic velocity distributions.

A parameter used is the ratio of the observed (projected) maximum rotation velocity $V_{\rm m}$ and the observed line-of-sight velocity dispersion at the center $\bar{\sigma}$.

This is a measure of the relative importance of rotation and random motions.

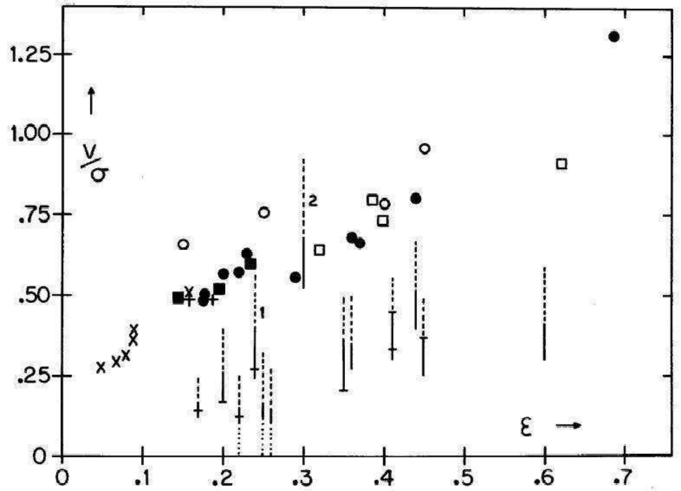
It can be compared to the observed flattening $\epsilon = 1 - b/a$ with a and b the (projected) major and minor axis[7].

The symbols in the next graph indicate models with isotropic velocity dispersions that are flattened by rotation and seen under various inclinations.

The bars are data and rotate less than expected for the observed flattening.

Note that the models lie on a well-defined line where the intrinsic relation roughly coincides with the projected one.

[7] G. Illingworth, Ap.J. 218, L43 (1977)


1.25
1.00
.75
.50
.25
0
V/σ
2
1
ε
0
.1
.2
.3
.4
.5
.6
.7

Further work[8] showed that spiral bulges and faint ellipticals are fast rotators.

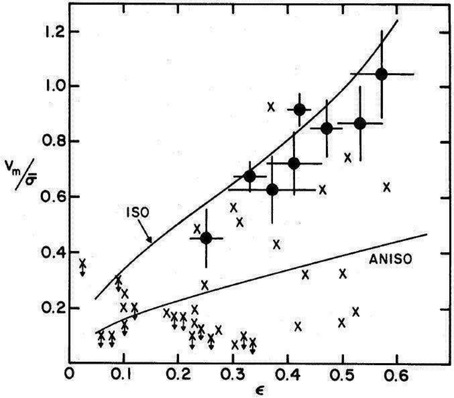
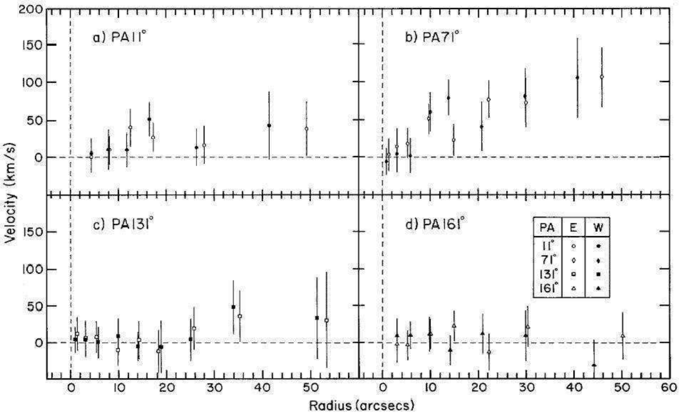


FIG. 3.—Comparison of bulge data (*filled circles*) with all available elliptical galaxy data (*crosses*, *arrows* indicate upper limits) in the dimensionless rotation-ellipticity plane. Derivation of V_m, $\bar{\sigma}$, and ε is discussed in the text. The line labeled ISO represents projected models of oblate spheroids with isotropic residual velocities and rotational flattening. The line labeled ANISO describes a typical anisotropic oblate model with σ_z smaller than σ_r and σ_θ.

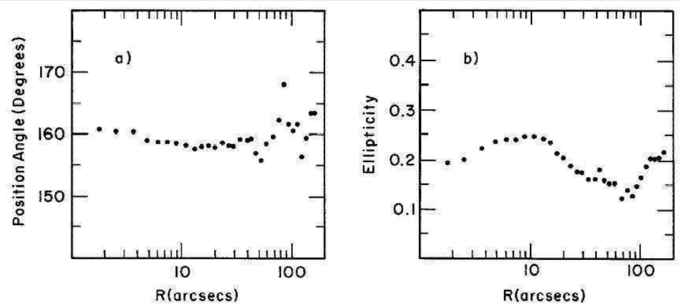
[8] e.g. J. Kormendy & G. Illingworth, Ap.J. 256, 460 (1982)

Minor axis rotation was first discovered in NGC 4261[9].



[9]R.L. Davies & M. Birkinshaw, Ap.J. 303, L45 (1986)

The maximum rotation is in p.a.$\sim 70^\circ$, while the isophotes have major axis at $\sim 160^\circ$.



The suggestion was made that this galaxy is prolate.

It turned out that elliptical galaxies are triaxial[10].

This explains the $(V_{\rm m}/\sigma - \epsilon)$-relation, the isophote twists and the minor axis rotation.

Minor axis rotation can result from[11]:

- projection effects in triaxial systems or
- misalingment of the angular momentum and the shortest axis.

[10] J. Binney, Mon.Not.R.A.S. 183, 779 (1978)

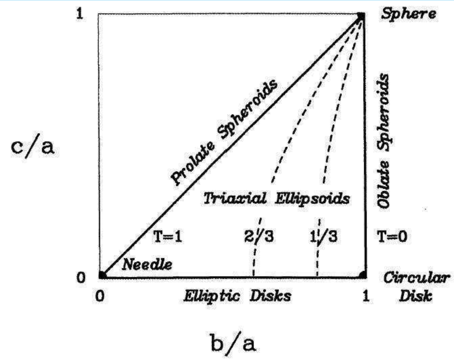
[11] M. Franx, G. Illingworth & P.T. de Zeeuw, Ap.J. 383, 112 (1991)

Define the misalignment $\psi_{\rm int}$ as the angle between the intrinsic short axis and the angular momentum.

Define for axes $a \geq b \geq c$ the triaxiality

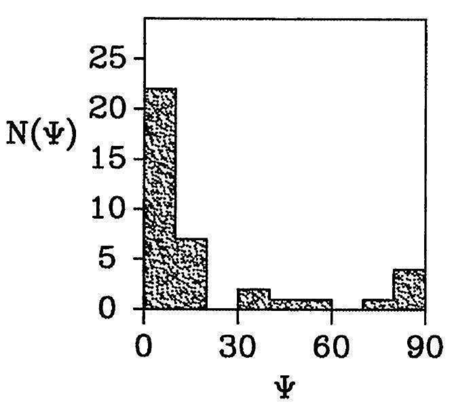
$$T = \frac{a^2 - b^2}{a^2 - c^2} = \frac{1 - b^2/a^2}{1 - c^2/a^2}$$

Thus T=0: oblate;
T=1: prolate.



We can *measure* the apparent ellipticity ϵ and the apparent misalignment ψ (the ratio of maximum observed velocity on the apparent axes)

$$\tan\psi = \frac{v_{\rm min}}{v_{\rm maj}}.$$



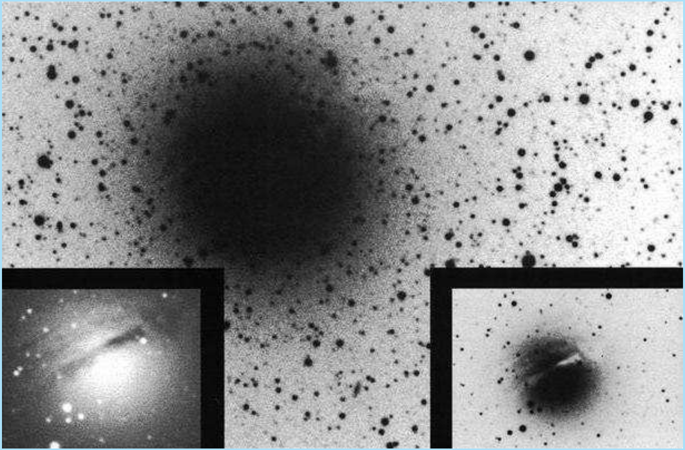
The distributions observed give the following rough indications:

- Most (at least 50%) ellipticals have a small $\psi_{\rm int}$ ($\lesssim 10^\circ$), but some ($\approx$ 10%) rotate along their major axis.
- $\langle T \rangle \approx 0.3$ and T has a wide distribution with possibly as much as 40% of the galaxies prolate.
- The ratio c/a has a peak at about 0.6-0.7.

Dust lanes are often seen[12] and occur usually along the apparent minor axis, but also sometimes along the major axis.

[12] F. Bertola & G. Galletta, Ap.J. 226, L115 (1978)

Here is NGC 1947.



In triaxial potentials stable orbits are possible, but the detailed kinematics depends on the galaxy shape and body rotation.

In principle dust lanes can be used to determine the intrinsic shape of an individual galaxy [13].

[13] R.L. Merritt & P.T. de Zeeuw, Ap.J. 267, L19 (1983); J. Kormendy & G. Djorkovski, Ann.Rev.A&A. 27, 235 (1989)

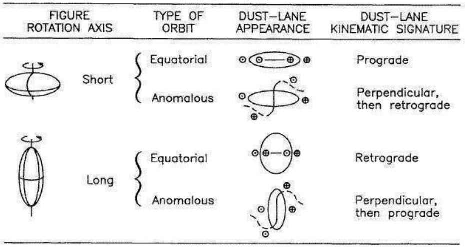


Figure 1 Stable orbits of gas in a rotating triaxial galaxy (adapted from Merritt & de Zeeuw 1983). As illustrated, the figure tumbles in the direction of stellar rotation ($\Omega_p > 0$); if $\Omega_p < 0$, the sense of gas rotation is reversed. Assume that the figure rotates about its shortest or longest axis (*left*). The second column gives the kind of orbit, and the third sketches resulting dust lanes seen edge-on. Anomalous orbits have different orientations at different radii (van Albada et al. 1982). They are the analogues of polar orbits in a stationary potential; at small radii, where Ω_p is unimportant, they are polar. At large radii, the figure rotates several times during an orbit and so is effectively oblate-spheroidal; then the orbit is equatorial (Simonson 1982). In between, the orbits have skew orientations determined by the Coriolis force. The schematic illustrations of dust lanes show the directions of stellar and gas motion; $\odot$ indicates approach, and $\oplus$ indicates recession. The right column states the kinematic signature, i.e. the sense of rotation of the dust lane with respect to the stars.

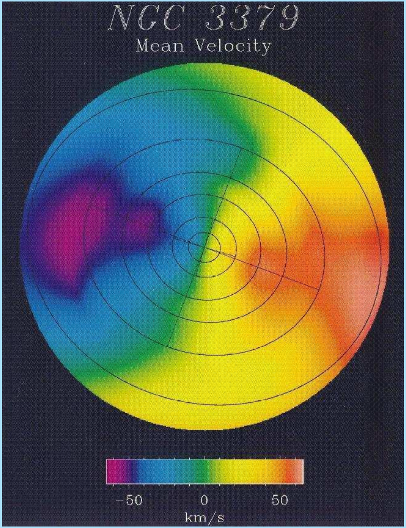
Detailed kinematics

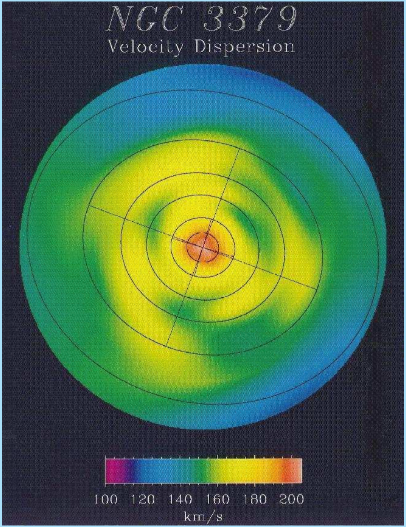
Detailed kinematics, including higher order moments of the velocity distyribution, of the velocity distributions can now be observed very well.

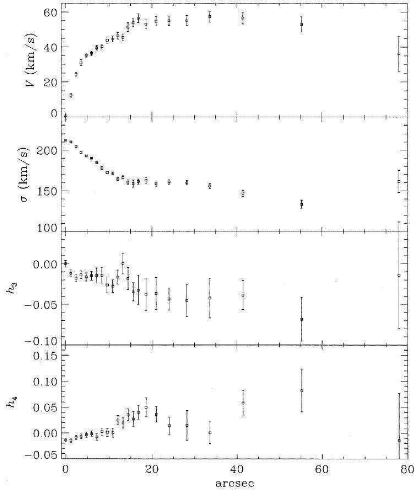
An example is a study of NGC3379[14].

Dynamical modeling shows that NGC 3379 may be a flattened, weakly triaxial system seen in an orientation that makes it appear round.

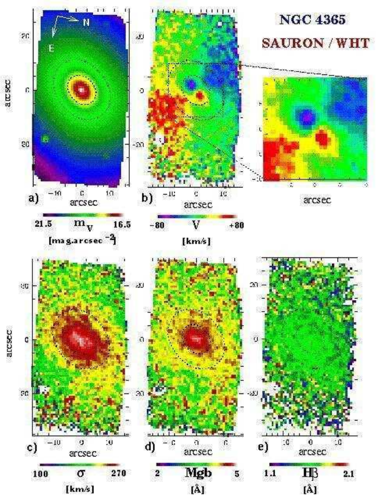
[14] T.S. Statler & T. Smecker-Hane, A.J. 117, 839 (1999)


NGC 3379
Mean Velocity
−50
0
50
km/s


NGC 3379
Velocity Dispersion
100
120
140
160
180
200
km/s


60
40
20
0
V (km/s)
200
150
100
σ (km/s)
0.00
−0.05
−0.10
h_3
0.15
0.10
0.05
0.00
−0.05
h_4
0
20
40
60
80
arcsec

Recently the SAURON integral field spectrograph has been built and used to survey kinematics and structure of elliptical galaxies[a].



---

[a] P.T. de Zeeuw et al., Mon.Not.R.A.S. 329, 513 (2002)

STRUCTURE AND DYNAMICS OF GALAXIES

21. Elliptical galaxies: Dynamical structure

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/∼vdkruit

Beijing, September 2011

Central kinematics and black holes

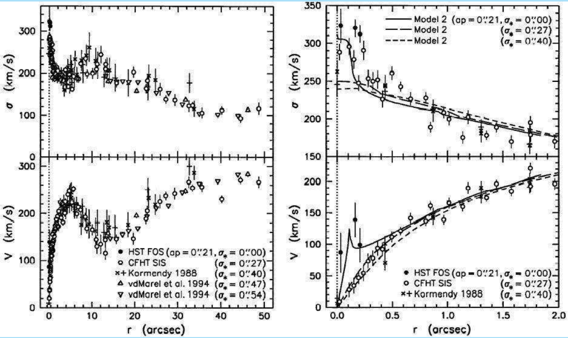
The central regions often show kinematics deviating from the outer parts.

These distinct cores may show:

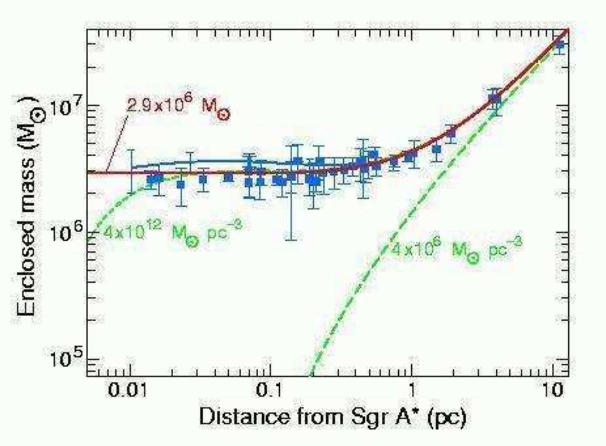
- Rapid rotation in the core but slow rotation in the main body
- Opposite rotation in the core relative to that in the main body
- Core rotation along the minor axis.

The distinct cores usually show small velocity dispersions, which suggest a two-component galaxy consisting of an elliptical with a small central disk.

Evidence for black holes comes from rapid rotation and high velocity dispersions in the inner regions, such as in NGC 4594[1] or our own Galaxy.



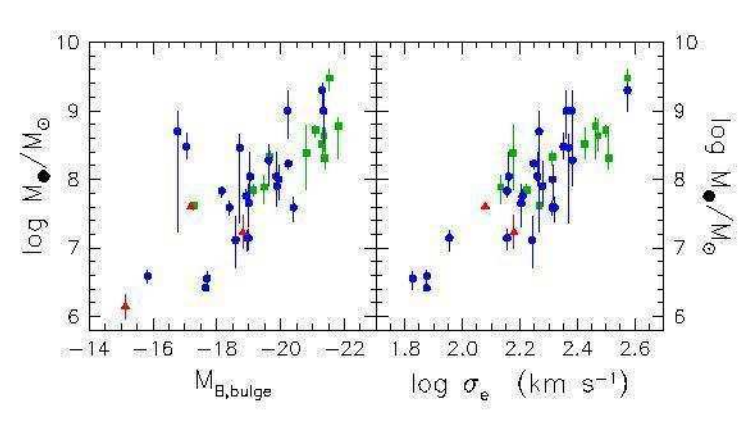
[1] J. Kormendy et al., Ap.J. 473, L91 (1996)


Enclosed mass ($M_\odot$)
10^7
10^6
10^5
$2.9 \times 10^6\ M_\odot$
$4 \times 10^{12}\ M_\odot\ pc^{-3}$
$4 \times 10^6\ M_\odot\ pc^{-3}$
0.01
0.1
1
10
Distance from Sgr A* (pc)

A compilation of all available data[2] shows a tight correlation between the mass of the black hole and the luminosity or velocity dispersion in the main body of the elliptical galaxy or bulge.

Probably this means no more than that larger galaxies have more material to feed into the center.

[2] S. Tremaine et al., Ap.J. 574, 740 (2002)


log M● /M⊙
10
9
8
7
6
−14
−16
−18
−20
−22
M B,bulge
1.8
2.0
2.2
2.4
2.6
log σe (km s−1)

Dynamical models

Stäckel potentials

The most simple description of an elliptical is that of King models, which are isothermal spheres with tidal radii and truncations in the velocity distributions. For these we have can estimate the total mass from

$$\frac{M}{L} = \frac{9\sigma^2}{2\pi G I_0 r_{\rm c}}.$$

However, we have seen that ellipticals have anisotropic velocity distribrutions and are in general triaxial.

A describtion then is with Stäckel potentials, which are potentials that are separable in ellipsoidal coordinates.

These are coordinates (λ, μ, ν) that are the three roots of τ for

$$\frac{x^2}{\tau+\alpha} + \frac{y^2}{\tau+\beta} + \frac{z^2}{\tau+\gamma} = 1$$

with $\alpha \leq \beta \leq \gamma$ three constants. It then turns out that

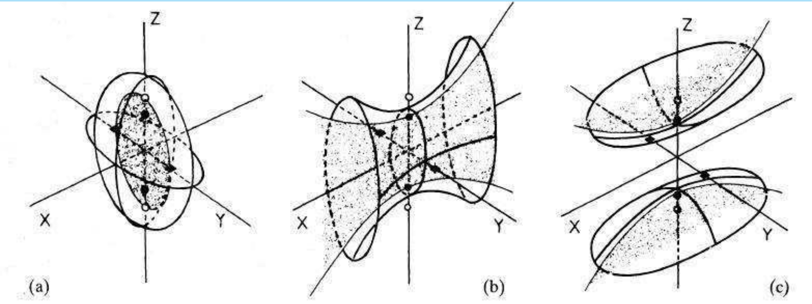
$$-\gamma \leq \nu \leq -\beta \leq \mu \leq -\alpha \leq \lambda$$

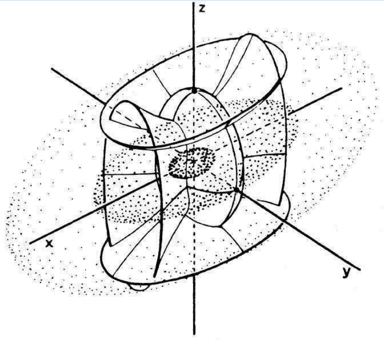
The line element is $ds^2 = P^2 d\lambda^2 + Q^2 d\mu^2 + R^2 d\nu^2$ with

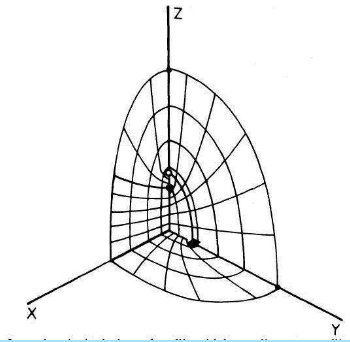
$$P^2 = \frac{(\lambda-\mu)(\lambda-\nu)}{4(\lambda+\alpha)(\lambda+\beta)(\lambda+\gamma)} \quad ; \quad Q^2 = \frac{(\mu-\nu)(\mu-\lambda)}{4(\mu+\alpha)(\mu+\beta)(\mu+\gamma)}$$

$$R^2 = \frac{(\nu-\lambda)(\nu-\mu)}{4(\nu+\alpha)(\nu+\beta)(\nu+\gamma)}$$

In such coordinate systems surfaces of constant λ are ellipsoids, of constant μ hyperboloids of one sheet and of constant ν hyperboloids of two sheets.




z
x
y


Z
X
Y

Stäckel potentials are of the form

$$\Phi(\lambda,\mu,\nu) = -\frac{F(\lambda)}{(\lambda-\mu)(\lambda-\nu)} - \frac{F(\mu)}{(\mu-\nu)(\mu-\lambda)} - \frac{F(\nu)}{(\nu-\lambda)(\nu-\mu)}$$

This can be used to describe triaxial galaxies[3].

Many density distributions can be locally approximated with a Stäckel potential.

For example, it is possible to derive a local approximation to the the potential in a disk with a flat rotation curve by a Stäckel potential[4].

[3] P.T. de Zeeuw & D. Lynden-Bell, Mon.Not.R.A.S. 215, 713 (1985); P.T. de Zeeuw, Mon.Not.R.A.S. 216, 273 (1985)

[4] T.S. Statler, Ap. J. 344, 217 (1989)

If the density is specified on the z-axis and if the potential is of the Stäckel-form in a specified ellipsoidal coordinate system, then the density at any point can be calculated with the so-called generalized Kuzmin formula[5].

A set of models with simple density profiles has been calculated[6] to illustrate the usefulness.

A nice example is the modified Hubble model, which has

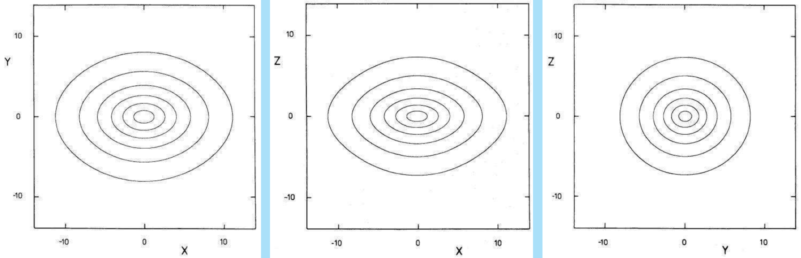
$$\rho(z) = \rho_\circ(1 + z^2)^{-3/2}$$

Then the coordinate system determines what the axis ratio's are in the density distributions and these change with radius.

[5] P.T. de Zeeuw, Mon.Not. R.A.S. 216, 599 (1985)

[6] P.T. de Zeeuw, R. Peletier & M. Franx, Mon.Not.R.A.S. 221, 1001 (1986)

Here are isodensity curves for a typical triaxial modified Hubble model (contour interval log 3).



So, this density distribution has smooth isodensity surfaces <u>and</u> has in a potential of Stäckel form!

The perfect ellipsoid

Every orbit in a Stäckel potential is the sum of three motions, one in each coordinate.

As a result motion is bounded by coordinate surfaces.

It is of use to study the types of orbits that arise in triaxial potentials.

A beautiful illustration is the case of the perfect ellipsoid[7], which is both stratified on concentric (triaxial) ellipsoids and produces exactly a Stäckel potential.

[7] P.T. de Zeeuw, Mon.Not.R.A.S. 216, 273 (1985)

The perfect ellipsoid has the density distribution

$$\rho = \frac{\rho_\circ}{(1+\tilde{m}^2)^2} \quad ; \quad \tilde{m}^2 = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \quad ; \quad a \geq b \geq c$$

This has semi-axes $\tilde{m}a$, $\tilde{m}b$ and $\tilde{m}c$ and falls off as $\tilde{m}^{-4}$ at large distances.

The function $F(\tau)$ in the equation for the potential then is

$$F(\tau) = \pi G \rho_\circ abc(\tau+\alpha)(\tau+\gamma) \int_0^\infty \frac{\sqrt{u-\beta}}{\sqrt{(u-\alpha)(u-\gamma)}} \frac{du}{u+\tau}$$

There are exact solutions for the (isolating) integrals of motion:

$$H = X + Y + Z$$

$$J = (\mu + \nu)X + (\nu + \lambda)Y + (\lambda + \mu)Z$$

$$K = \mu\nu X + \nu\lambda Y + \lambda\mu Z$$

where

$$X = \frac{P^2\dot{\lambda}^2}{2} - \frac{F(\lambda)}{(\lambda - \mu)(\lambda - \nu)} \;\; ; \;\; Y = \frac{Q^2\dot{\mu}^2}{2} - \frac{F(\mu)}{(\mu - \nu)(\mu - \lambda)}$$

$$Z = \frac{R^2\dot{\nu}^2}{2} - \frac{F(\nu)}{(\nu - \lambda)(\nu - \mu)}$$

These integrals are all quadratic in velocity and have the dimension of an energy.

It is more insightfull to write the integrals as the energy (as usual) and two non-classical integrals:

$$I_1 = H$$

$$I_2 = \frac{\alpha^2 H + \alpha J + K}{\alpha - \gamma}$$

$$I_3 = \frac{\gamma^2 H + \gamma J + K}{\gamma - \alpha}$$

Special case I: the prefect prolate spheroid. Here $\gamma = \beta$ (so the long axis is the x-axis). Since

$$-\gamma \leq \nu \leq -\beta \leq \mu \leq -\alpha \leq \lambda$$

we have

$$\nu = -\gamma = \beta$$

The third integral then becomes the (classical) angular momemtum along the x-axis

$$I_3 = \frac{1}{2}(y\dot{z} - z\dot{y})^2 = \frac{1}{2}L_{\mathrm{x}}^2$$

The integral I_2 remains a non-classical one.

Special case II: the perfect oblate spheroid. Then we have $\mu = -\beta = -\alpha$.

In this case the angular momentum around the z-axis is an isolating integral:

$$I_2 = \frac{1}{2}(x\dot{y} - y\dot{x})^2 = \frac{1}{2}L_z^2$$

I_3 is the well-know third integral of Galactic dynamics.

I_3 remains a non-classical integral.

Special case III: If we collapse the perfect oblate spheroid along the symmetry axis we get the Kuzmin disk.

With $\mu = -\beta = -\alpha$ and $\gamma = 0$ we get the same I_2 as above and in addition

$$I_3 = \frac{1}{2} L_{\rm x}^2 + \frac{1}{2} L_{\rm y}^2 + \frac{1}{2} a \dot{z}^2 - a|z|\Phi$$

(a is the coordinate system focal distance above and below the plane)

I_3 has the property of an energy associated with the z-axis.

In this case we then have three isolating integrals E, I_2 and I_3.

Special case IV: the perfect sphere. Then
$\mu = \nu = -\gamma = -\beta = -\alpha$. So

$$J = \frac{1}{2}L^2 - 2\alpha H \quad : \quad K = \alpha^2 - \frac{1}{2}\alpha L^2 \quad ; \quad I_2 + I_3 = \frac{1}{2}L^2$$

with $\vec{L}$ the total angular momentum vector $(L_\mathrm{x}, L_\mathrm{y}, L_\mathrm{z})$.

Then there are four isolating integrals of motion , namely the total energy E and the three components of the angular momentum L_x, L_y and L_z.

Types of orbits

For dynamical studies it is important to investigate the possible general types of orbits in the kind of potential considered. Here we look at orbits in triaxial potentials using the perfect ellipsoid.

It can be shown that the equations of motion become

$$E = 2(\tau + \beta)p_\tau^2 + \Phi_{\rm eff}(\tau)$$

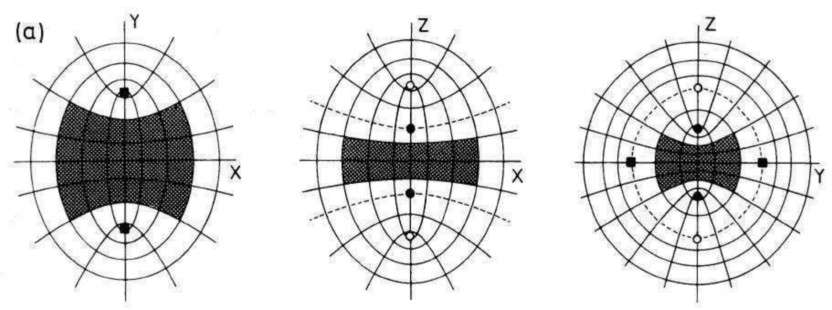
with

$$p_\lambda = P^2 \dot{\lambda} \quad ; \quad p_\mu = Q^2 \dot{\mu} \quad ; \quad p_\nu = R^2 \dot{\nu}$$

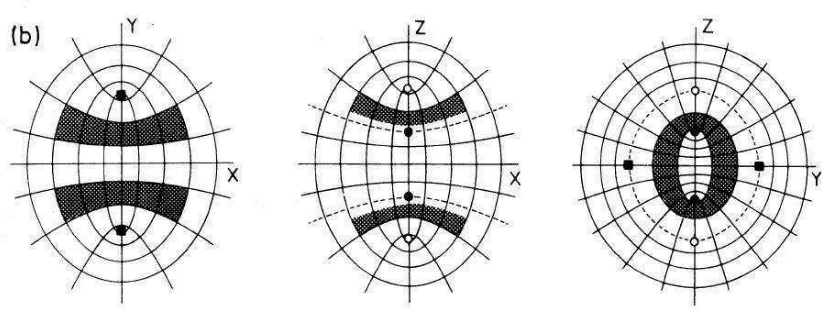
$$\Phi_{\rm eff} = \frac{I_2}{\tau + \alpha} + \frac{I_3}{\tau + \gamma} - G(\tau)$$

Depending on the values of the integrals there are four general types of orbits.

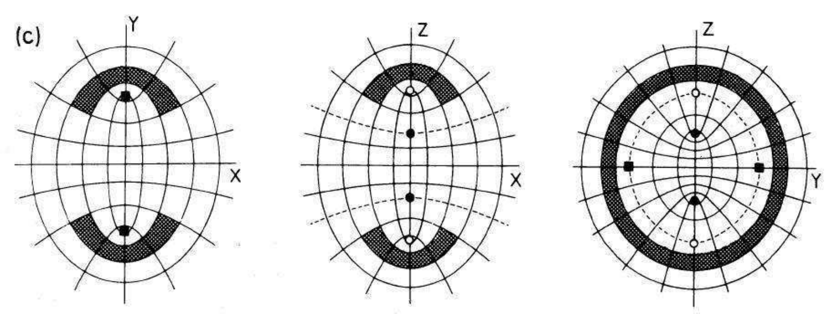
Box orbits



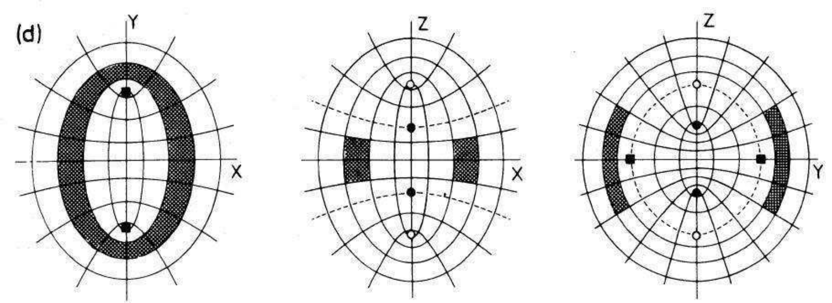
Inner long axis tube orbits



Outer long axis tube orbits



Short axis tube orbits



Next consider orbits in the (x, y)-plane.

This is for $\nu = -\gamma$ and $p_\nu^2/2R^2 = 0$.

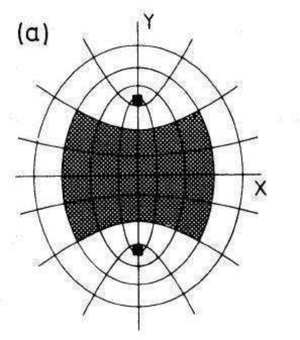
It can be shown that for orbits in this plane we have

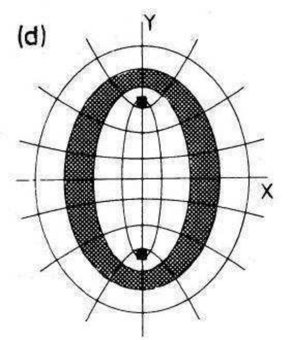
$$I_3 = 0$$

Then two types of orbits remain, which are versions of the orbits earlier, but now collapsed onto the (x, y)-axis.

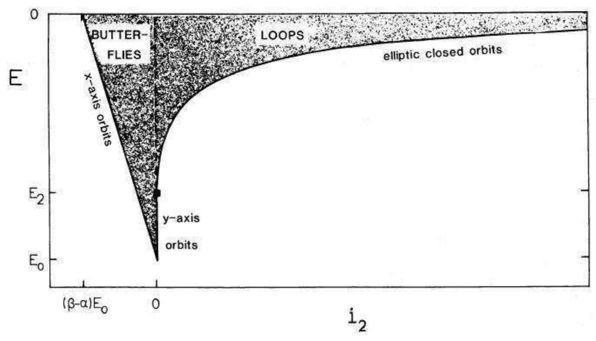
These orbits turn out to be stable for perturbations perpendicular to this plane.

The two types of orbits that remain are butterflies (collapsed box orbits with $I_2 < 0$; left) and loops (collapsed short axis tubes with $I_2 > 0$; right), resp. inside or outside the foci.





The orbits can be distinguished according to the integrals.



The limiting cases are *x*-axis orbits, *y*-axis orbits (which are unstable for *x*-perturbations) and elliptic closed orbits.

Then orbits in the (x, z)-plane.

Since $\mu = -\beta$ or $\nu = -\beta$

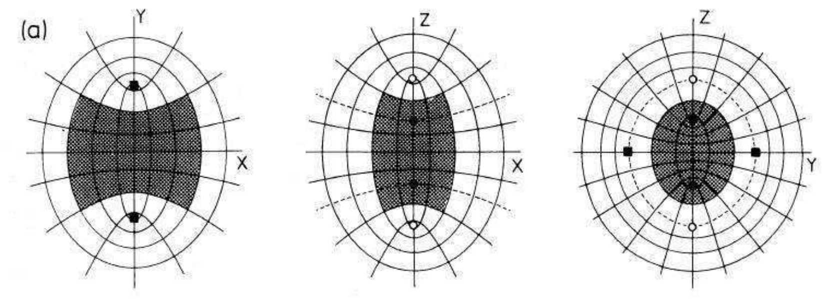
$$E - E_\circ = \frac{I_2}{\alpha - \beta} + \frac{I_3}{\gamma - \beta}$$

The fundamental orbits are again butterflies and loops.

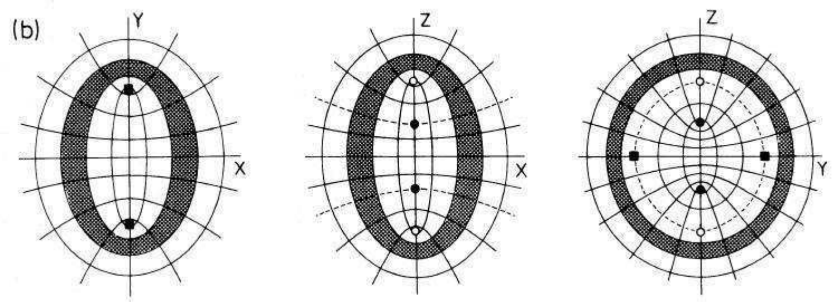
The butterflies can either be stable (and then are collapsed box orbits) or unstable for perturbations in the y-direction. When stable they are collapsed box orbits.

The loops are all unstable.

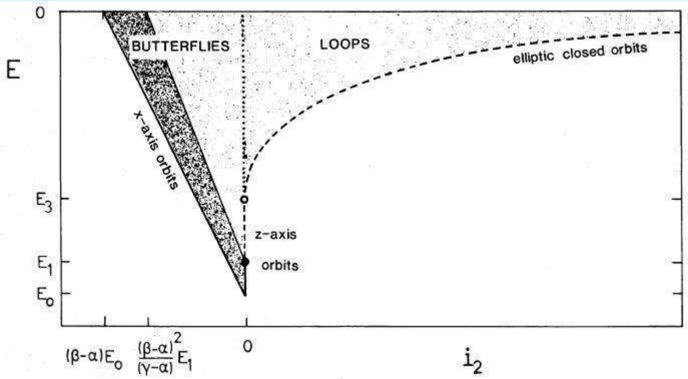
Unstable butterfly



Unstable loop



Classification of (x, z)-orbits (shaded is stable, dashed is unstable periodic orbits).



Orbits in the (y, z)-plane.

Now $\lambda = -\alpha$ or $\mu = -\alpha$.

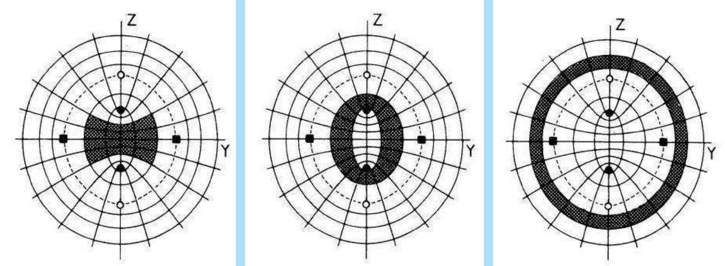
Now we have

$$I_2 = 0$$

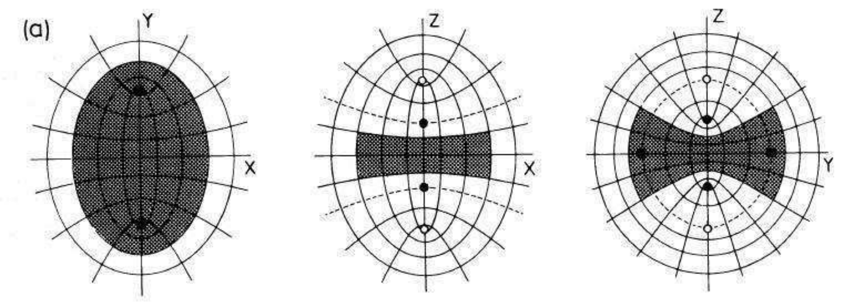
.
We have again butterflies and loops, but these can now be both stable and unstable.

The stable butterfly is a collapsed box orbits. There are two types of stable loops, either collapsed inner or outer long axis tubes.

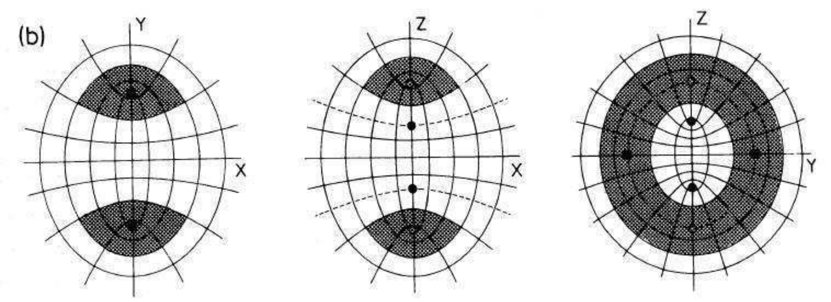
Left the stable butterfly and on the right the two stable loops.



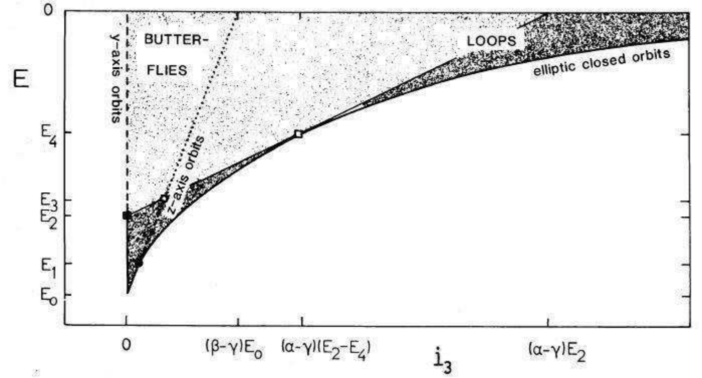
Unstable butterfly



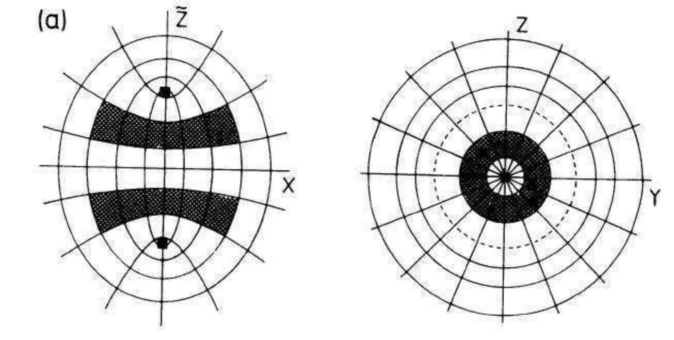
Unstable loop



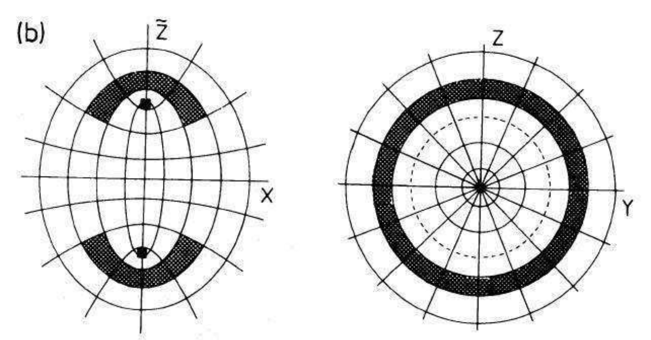
Classification of orbits



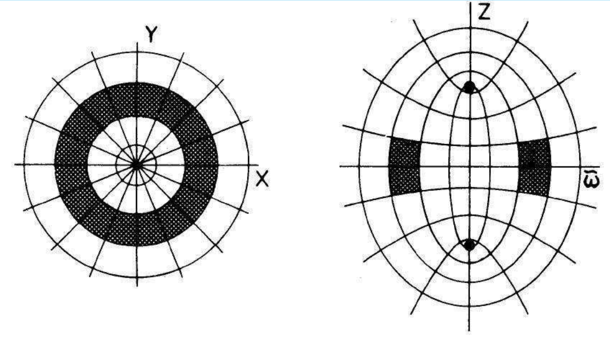
In the case of a prolate spheroid only two types of orbits are possible.
Here is the inner long axis tube.



The vertical axis indicates that this is any meridional plane perpendicular to x.
The other possibility is the outer long axis tube



In the case of the oblate spheroid only short axis tube orbits are possible.



Dark matter

Solutions for isotropic models usually have gradients in M/L, while for triaxial models solutions with constant M/L are usually possible.

The manner to proceed and make progress then is to consider higher order moments of the observed velocity profiles.

For example Carollo et al.[8] show that at least three out of their four ellipticals must have dark haloes.

[8] C.M. Carollo, P.T. de Zeeuw, R.P. van der Marel, I.J. Danziger & E.E. Qian, 441, L25 (1995)

X-ray halos

X-ray emission at large radii can also be used to measure masses of large ellipticals and clusters.

Measure the X-ray emissivity distribution $\epsilon(r)$ from the distribution on the sky and the X-ray energy distribution.

Infer from the distribution of ϵ the density distribution of the gas $\rho_{\rm gas}(R)$ and the distribution of temperature $T(r)$.

Then the hydrostatic equation gives for the pressure P

$$\frac{dP}{dR} = -\frac{GM(<R)}{R^2}\rho_{\rm gas}(R)$$

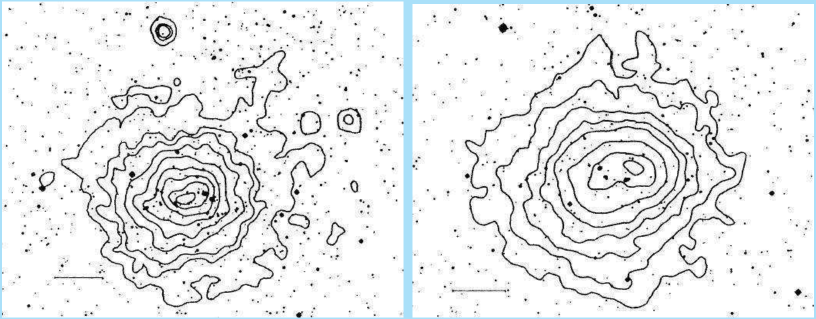
The ideal gas equation gives

$$P = \rho_{\rm gas}\frac{kT}{\mu m_{\rm p}}$$

Then

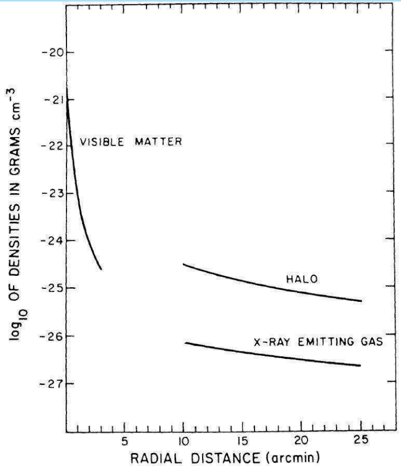
$$M(<R) = -\frac{kT(R)R}{G\mu m_{\rm p}}\left[\frac{d\log\rho_{\rm gas}}{d\log R} + \frac{d\log T}{d\log R}\right].$$

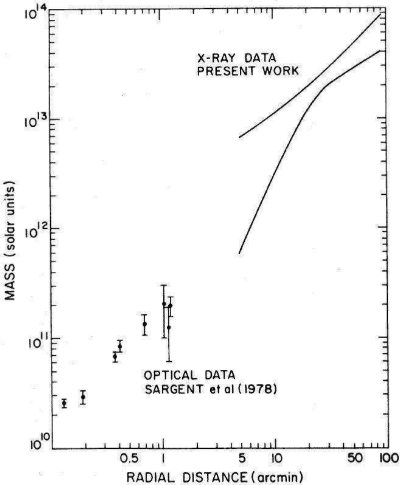
Here are X-ray distributions in two clusters of galaxies.



The next two graphs show the analysis of the giant elliptical M 87 in the center of the Virgo cluster[9].

[9] Fabricant & Gorenstein, Ap.J. 267, 535 (1983)


log10 OF DENSITIES IN GRAMS cm-3
-20
-21
-22
-23
-24
-25
-26
-27
VISIBLE MATTER
HALO
X-RAY EMITTING GAS
5
10
15
20
25
RADIAL DISTANCE (arcmin)


MASS (solar units)
10^14
10^13
10^12
10^11
10^10
X-RAY DATA
PRESENT WORK
OPTICAL DATA
SARGENT et al (1978)
0.5
1
5
10
50
100
RADIAL DISTANCE (arcmin)

Shells can also be used. Simulations show that their spacing depends on the mass profile.

Finally we can measure masses of whole clusters of galaxies.

The Virial Theorem $2T + \Omega \sim 0$ for equilibrium for a uniform, spherical distribution gives

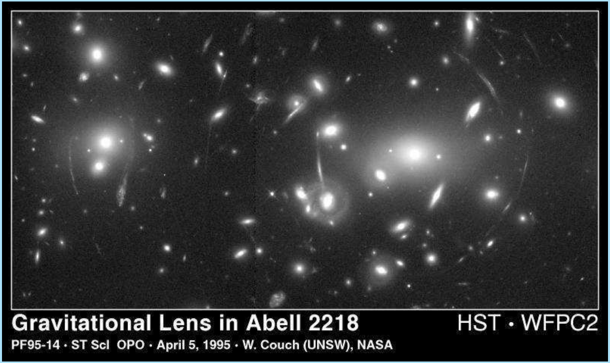
$$2T = \sum mV^2 \sim M\langle V^2\rangle \sim -\Omega \sim \frac{3GM}{5R}$$

Thus

$$M \sim \frac{R\sigma_{\rm v}^2}{G} \sim \left(\frac{R}{1\ \rm Mpc}\right)\left(\frac{\sigma_{\rm v}}{10^3\ \rm km\ s^{-1}}\right)^2 10^{15}\ \rm M_{\odot}$$

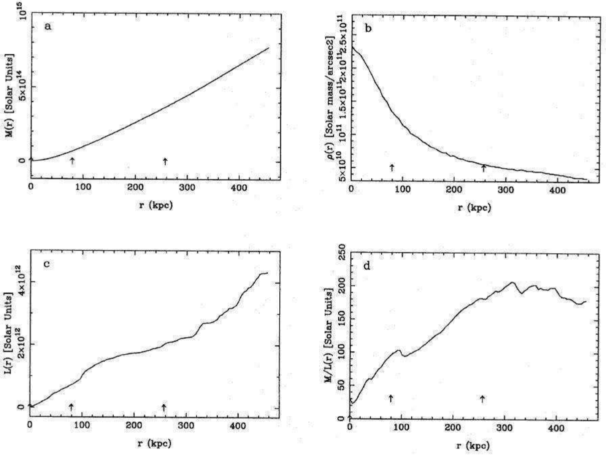
This indicates masses of up to 10^{15} $M_{\odot}$.

Nowadays also gravitational arcs can be used (e.g. in Abell 2218[10]).



[10] J.P. Kneib et al., A.&A. 303, 27 (1995)

Here are the inferred distributions.



STRUCTURE AND DYNAMICS OF GALAXIES

22. Chemical evolution

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/∼vdkruit

Beijing, September 2011

Abundance gradients

Bulges have color gradients (become bluer with radius).

This is due to metallicity changes.

For a low [Fe/H] in an old population:

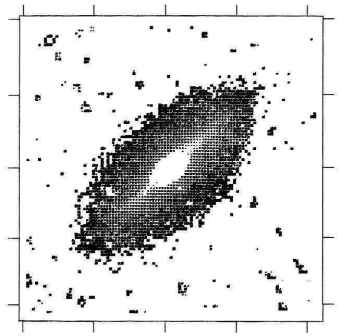
- The effective temperature of the giant branch is higher
- There is less line-blanketing
- The horizontal branch is more extended towards the blue.

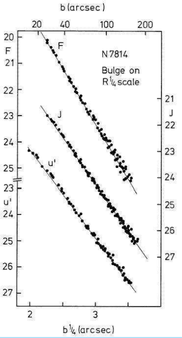
The relation between color and metallicity can be calibrated using the integrated light of Galactic globular clusters.

The range in (U-B),(B-V) in bulges is roughly that in globular clusters.

So the range in metallicity in bulges is 1 - 2 dex in [Fe/H]. }

There is such a pronounced color gradient in the bulge of NGC 7814[1]





$$\mu_{U'} = 14.87 + 3.32 b^{1/4}$$

$$\mu_{J\,\mathrm{opt}} = 13.72 + 3.55 b^{1/4}$$

$$\mu_{V} = 13.08 + 3.75 b^{1/4}$$

$$\mu_{F} = 10.70 + 4.20 b^{1/4}$$

Disks have gradients in emission line ratio's in HII regions.

Some prominent emission lines in spectra of HII-regions are the following:

Ion	Wavelength
[OII]	3726/3729
Hδ	4101
Hγ	4340
Hβ	4861
[OIII]	4959/5007
Hα	6562
[NII]	6548/6583
[SII]	6716/6731

An often used parameter is the "excitation", which is the ratio of the strengths of the [OIII] and Hβ lines.

These are at about the same wavelength, so this ratio is not sensitive to extinction corrections.

The exitation could change due to a number of effects:

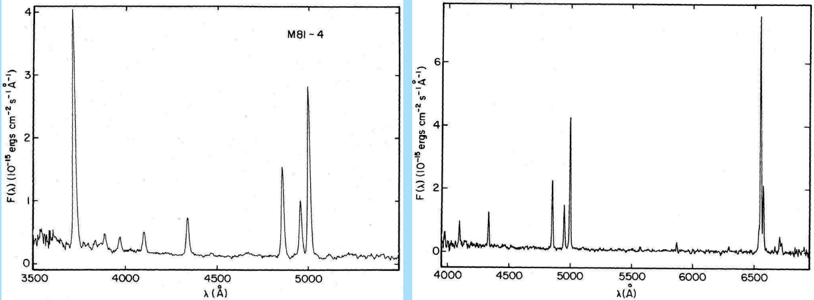
- Changing dust content and therefore radiation field
- Changing stellar temperatures; increasing $T_{\rm eff}$ gives increasing excitation
- Changing abundance because of cooling through O- and N-ions:
  A lower oxygen abundance gives an increased $T_{\rm e}$ and then we get stronger O-lines; thus [OIII]/Hβ *increases with decreasing metallicity*.

Detailed studies[a] have shown that the effect of abundance gradients is probably the most important.

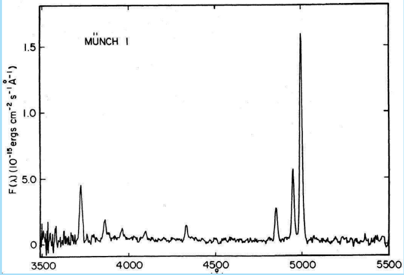
As an example we have a detailed look at measurement in M81[b] between 3 and 15 kpc.

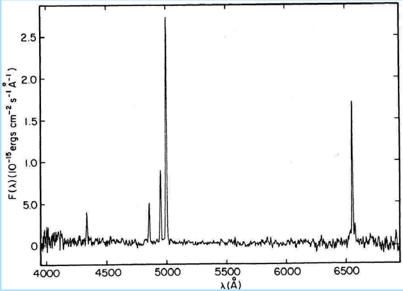
[a] L. Searle, Ap.J. 168, 327 (1973)
[b] D.R. Garnett & G.A. Shields, Ap.J. 317, 82 (1987)

This is the spectrum of an HII-region at R = 7 kpc.

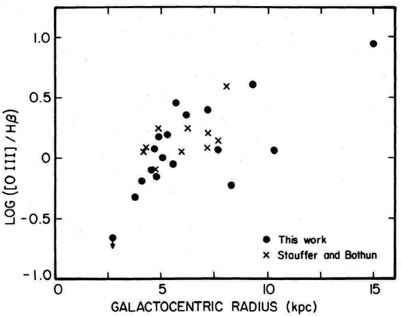


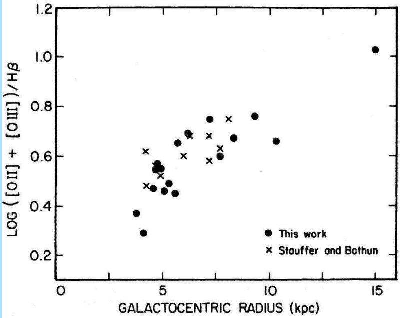
This is the spectrum of an HII-region at $R = 15$ kpc.





Here we see the gradients in [OIII]/Hβ ratio and the ([OIII]+[OII])/Hβ ratio.



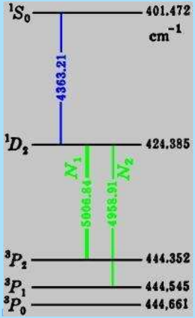


The use of [OIII] and [OII]] has the advantage that two levels of ionisation of the oxygen are taken.

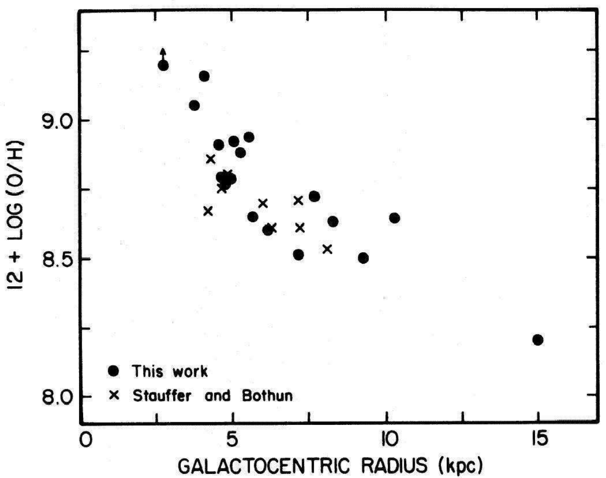
The disadvantage is that the extinction corrections are important.

The line ratio's must be transformed into abundances. The calibration of excitation into abundance can be done in two ways:

- Measure the weak [OIII] line at λ 4363 in addition to the lines at λ 4959 and 5007. Then lines are measured involving the same level and from this the electron temperature $T_{\rm e}$ can be calculated. This allows the determination of the oxygen-hydrogen ratio.



- The second possibility is to calculate full sets of photoionization models of HII regions.



The result for these measurements in M81 is a gradient of -0.08 dex kpc^{-1} in [O/H].

This is a typical value for spiral disks, including our own.

Theory of chemical evolution

Take a volume (either a whole galaxy or a part of it) and define within that volume:

$M_{\rm g}$ = Mass in gas

M_* = Mass in stars

$M_{\rm Z}$ = Mass in heavy elements

$Z(t) = M_{\rm Z}(t)/M_{\rm g}(t)$ = Abundance

$$y = \frac{\rm Mass\ injected\ in\ new\ metals}{\rm Mass\ locked\ in\ long-lived\ stars} = \rm Yield^{2}$$

[2] Searle & Sargent, Ap.J. 173,25 (1972)

The Instantaneous Recycling Approximation (IRA) says that star evolution of heavy stars is instantaneous and that the products are mixed instaneously into the interstellar medium.

Assume the system is closed (no inflow or outflow of gas).

Then the fundamental equations are:

$$\frac{dM_{\rm Z}}{dt} = y\frac{dM_*}{dt} - Z(t)\frac{dM_*}{dt}$$

$$\frac{dM_{\rm g}}{dt} = -\frac{dM_*}{dt}$$

The Simple Model

This assumes that $Z(t = 0) = Z_\circ = 0$.

Define

$$x = \frac{M_{\rm g}(t)}{M_{\rm tot}}$$

The fundamental equations can then be solved to give

$$Z(t) = y \ln\left(\frac{1}{x}\right)$$

The metal abundance of the gas is an increasing function of the gas fraction x and time.

Stars have the abundance of the gas at the time of their birth.

The fraction of stars at time t with abundance $Z \leq Z_1 (\leq Z(t))$ is:

$$F(Z) = \frac{1 - x_1}{1 - x}$$

$$x_1 = \exp - \left(\frac{Z_1}{y} \right)$$

So

$$\langle Z \rangle = y \frac{1 - x(1 - \ln x)}{1 - x}$$

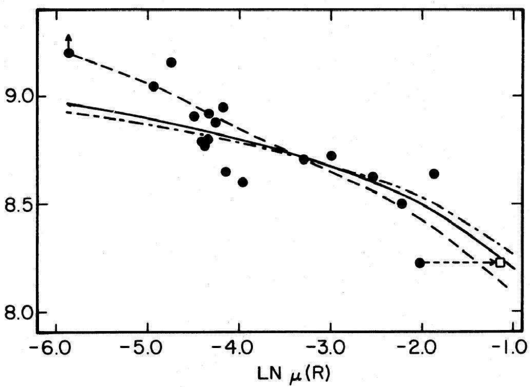
Use up all the gas ($x \rightarrow 0$), then $\langle Z \rangle \rightarrow y$.

So: Abundance of gas $\rightarrow \infty$.

The *mean* abundance of stars $\rightarrow y$.

The observations in M81 can be used to test this model.

For that purpose the radius has been replaced by the gas fraction (from the HI and the photometry) $\mu(R)$.



The thick line shows the observed distribution and Simple Model is the full-drawn line.

The Simple Model suffers from the G-dwarf problem:
It predicts far too many stars of low metallity.

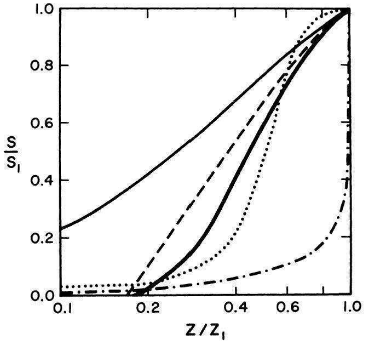


Figure 5 Metallicity distributions. S/S_1 is the fraction of G-K dwarfs in the solar neighborhood with metal abundance less than Z, where Z_1 is the present interstellar abundance (except as noted below). *Heavy line:* schematic representation of the data after removing an estimated dispersion due to observational errors (after Pagel & Patchett 1975). *Light solid line:* the "simple model" [Equation (3)]. *Dashed line:* effect of a finite initial abundance, $Z_0 = 0.17\,Z_1$. *Dash-dotted line:* an infall model [Equation (4)]. *Dotted line:* the infall model with a log gaussian distribution of Z at all times, with $\sigma(\log Z) = 0.2$. In this case, Z_1 is the value at which $S/S_1 \approx 1$ (cf Tinsley 1975a).

The simple model predicts that of the G-dwarfs in the solar neighborhood more than 40% should have a metallicity less than 0.2 of solar.

This fundamental problem was first noted by Maarten Schmidt[3].

There are two general ways to cure this; namely a non-zero abundance in the gas at the beginning or an extended inflow of unenriched material.

We will now explore these two options.

[3] M. Schmidt, Ap.J. 137, 758 (1963)

The Extended Simple Model.

The assumptions are the same as in the simple model, except that $Z_\circ \neq 0$.

This is also known as Prompt Initial Enrichment (PIE).

Then everywhere replace y with $y + Z_\circ$ and the equations look the same.

The solution then is

$$Z(t) = Z_\circ + y \ln \left(\frac{1}{x} \right)$$

So, now when we use up all gas, we get

Abundance of gas $\rightarrow \infty$.

Mean abundance of stars $\rightarrow y + Z_\circ$.

Because the metallicity of the gas is initially finite, there are (much) fewer metal-poor stars.

The Prompt Initial Enrichment Model now is the dashed line.

This is a better representation of the observed distribution.

(The thick line was the observed distribution and the full-drawn line the simple model) .

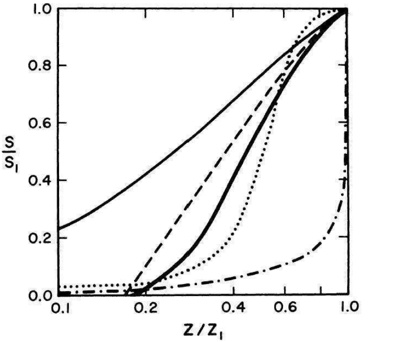


Figure 5 Metallicity distributions. S/S_1 is the fraction of G-K dwarfs in the solar neighborhood with metal abundance less than Z, where Z_1 is the present interstellar abundance (except as noted below). *Heavy line*: schematic representation of the data after removing an estimated dispersion due to observational errors (after Pagel & Patchett 1975). *Light solid line*: the "simple model" [Equation (3)]. *Dashed line*: effect of a finite initial abundance, $Z_0 = 0.17\,Z_1$. *Dash-dotted line*: an infall model [Equation (4)]. *Dotted line*: the infall model with a log gaussian distribution of Z at all times, with $\sigma(\log Z) = 0.2$. In this case, Z_1 is the value at which $S/S_1 \approx 1$ (cf Tinsley 1975a).

The Inflow Model.

Assume an inflow $f(t)$ of unprocessed material.

This means that there is less gas in the beginning compared to the simple model and the enrichtment then proceeds much faster and therefore decreases the predicted number of G-dwarfs.

The second fundamental equation becomes

$$\frac{dM_{\rm g}}{dt} = -\frac{dM_*}{dt} + f(t)$$

This model cannot be solved analytically in the general case, but it can be done fore the extreme inflow model, where $M_{\rm g}$ = constant.

Define

$$\mu = \frac{M_*(t)}{M_{\rm g}}$$

Then

$$Z(t) = y \left\{ 1 - \exp\left(-\mu\right) \right\}$$

It can then be found that

$$F(Z) = \frac{\mu}{\mu_1}$$

$$\langle Z \rangle = y - \frac{y}{\mu} + \frac{y}{\mu} \exp\left(-\mu\right)$$

If we now use up all gas, we get

$\mu \rightarrow \infty$ and $\langle Z \rangle \rightarrow y$.

Abundance of gas $\rightarrow y$.

Mean abundance of stars $\rightarrow y$.

The extreme infall model is the dash-dotted line.

(The thick line wass the observed distribution, the full-drawn line simple model and the dashed line the Prompt Initial Enrichment Model.)

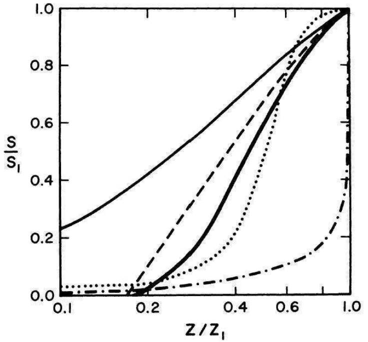


Figure 5 Metallicity distributions. S/S_1 is the fraction of G-K dwarfs in the solar neighborhood with metal abundance less than Z, where Z_1 is the present interstellar abundance (except as noted below). *Heavy line:* schematic representation of the data after removing an estimated dispersion due to observational errors (after Pagel & Patchett 1975). *Light solid line:* the "simple model" [Equation (3)]. *Dashed line:* effect of a finite initial abundance, $Z_0 = 0.17\,Z_1$. *Dash-dotted line:* an infall model [Equation (4)]. *Dotted line:* the infall model with a log gaussian distribution of Z at all times, with $\sigma(\log Z) = 0.2$. In this case, Z_1 is the value at which $S/S_1 \approx 1$ (cf Tinsley 1975a).

The extreme inflow model is much too extreme in that it now predicts too few metal-poor stars.

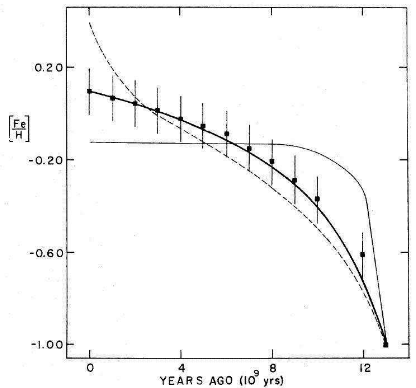
So, $M_{\rm gas}$ must have decreased with time.

The dotted line in the previous figure is an example of an adapted infall model.

The inflow is possibly seen in our Galaxy as the high-velocity clouds.

The best fit can be found for the Solar Neighborhood with a combination of prompt initial enrichement and inflow.

When the time is made explicit (e.g. by assuming that the SFR is constant) this model can reproduce the metallicity - age relation[4].



[4] B.A. Twarog, Ap.J. 242, 242 (1980)

The Simple Model with Bells and Whistles.[5]

This term is now used for any model that relaxes the assumptions of the simple model, but was used originally for models with outflow of processed material.

Let there be an outflow of processed material $g(t)$.

Then the fundamental equations become

$$\frac{dM_{\rm Z}}{dt} = y\frac{dM_*}{dt} - Z(t)\frac{dM_*}{dt} - Z(t)g(t)$$

$$\frac{dM_{\rm g}}{dt} = -\frac{dM_*}{dt} - g(t)$$

[5] J.R. Mould, P.A.S.P. 96, 773 (1984)

For an illustrative case that can be solved analytically, take

$$g(t) = \alpha \frac{dM_*}{dt}$$

Then we have the fundamental equations back with y replaced with an *effective* yield

$$y' = \frac{y}{1+\alpha}$$

The solution is then

$$Z(t) = \frac{y}{1+\alpha} ln \left(\frac{1}{x} \right)$$

Use up all gas, then:

Abundance of gas $\rightarrow \infty$.

Mean abundance of stars $\rightarrow y' = y/(1+\alpha)$.

For elliptical galaxies there is a mass - metallicity relation[6]. This can be explained if elliptical galaxies have (had) outflow of processed material, which must haven been more pronounced in smaller systems.

[6] J.R. Mould, P.A.S.P. 96, 773 (1984)

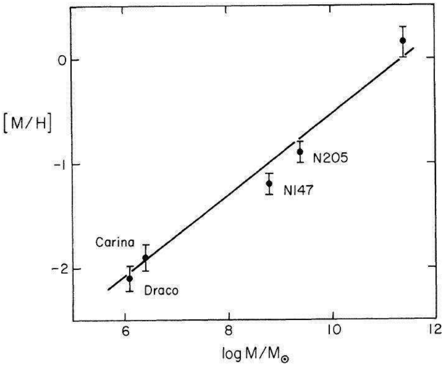


FIG. 2—A mass-metallicity relation for elliptical galaxies. The unlabeled point shows the metallicity inferred for the brightest ellipticals from integrated light models.

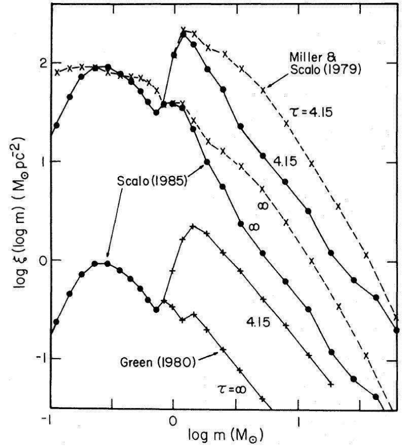
Bi-modal star formation

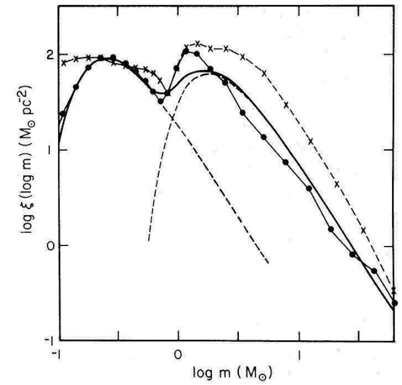
It is possible to relax the continuity constraint in the determination of the IMF and assume bi-modal star formation[7].

This is based on the idea of two modes of star formation, that are independent.

This continuity constraint can be relaxed and that also is a possible solution of the G-dwarf problem, since it uncouples the formation of lighter stars from the enrichment by the massive stars.

[7] R.B. Larson, Mon.Not.R.A.S. 218, 409 (1986)


Miller & Scalo (1979)
τ=4.15
4.15
∞
Scalo (1985)
4.15
Green (1980)
τ=∞
log ξ (log m) (M⊙ pc⁻²)
log m (M⊙)


log ξ (log m) (M⊙ pc⁻²)
log m (M⊙)

If C is the number of stars formed $(\log M)^{-1}$ pc^{-2} Gyr^{-1}:

$$C(\log M, t) = SFR_1(t) \cdot IMF_1(\log M) + SFR_2(t) \cdot IMF_2(\log M)$$

$$IMF_{\mathrm{k}}(\log M) = 2.55 M_{\mathrm{k}} M^{-2} \exp \left[- \left(\frac{M_{\mathrm{k}}}{M} \right)^{3/2} \right]$$

$$SFR_{\mathrm{k}}(t) = A_{\mathrm{k}} \exp \left(- \frac{t}{\tau_{\mathrm{k}}} \right)$$

- Low mass: $\tau_1 = \infty$, $M_1 = 0.30 M_\odot$, $A_1 = 1.85 M_\odot \mathrm{pc}^{-2} \mathrm{Gyr}^{-1}$
- High mass: $\tau_2 = 3.4$ Gyr, $M_2 = 2.2 M_\odot$, $A_2 = 41 M_\odot \mathrm{pc}^{-2} \mathrm{Gyr}^{-1}$

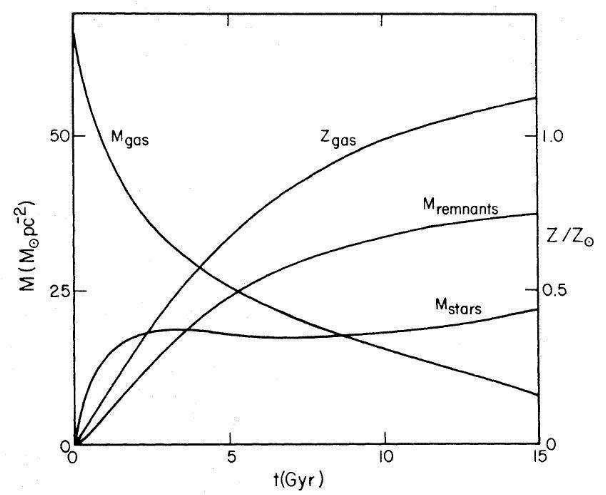
Effects of bi-modal star formation:

• This explains in a natural way the occurrence of two types of associations; the O-associations having OB-stars and the T-associations having only T Tauri stars.

• A smaller amount of mass has gone into long-lived stars per unit luminosity of newly formed stars during the whole history. This solves the problem of the gas consumption time-scale (why do all galaxies use their gas in another Hubble time or less?).

• More mass is in invisible remnants of massive stars (white dwarfs, etc.).
For $M_{\rm remnant} = 0.38 + 0.15 M_*$ this adds up to about 3/4 of the mass density.
This solves the local missing mass problem (Oort limit), but is only compatible with observations if the fading time is less than 10 Gyr.

• Rapid early increase in [Fe/H] combined with low relative SR in low-mass stars. This solves the G-dwarf problem of the simple model for chemical evolution.


M$_{gas}$
Z$_{gas}$
M$_{remnants}$
M$_{stars}$
M (M$_\odot$ pc^{-2})
Z / Z$_\odot$
t(Gyr)
0
5
10
15
25
50
0.5
1.0

Comparison to observations

- Abundance gradients in bulges:
  This results from a change in the effective yield with radius due to changing escape rates of processed gas.
- Overall abundances of ellipticals:
  There is a correlation of [Fe/H] with $M_{\rm V}$, which follows if for more massive systems the gas has more difficulty to escape.
- Disk abundance differences between galaxies:
  Earlier types have higher metallicities, because more gas has been used in star formation.

- Gas abundance gradients in disks:
  This results from radial gradient in relative gas consumption and content.
- Stellar abundance gradients in disks:
  No gradients should result if most of the gas is used up (the mean stellar abundance is then equal to the yield); at least it should be smaller than in the gas.

STRUCTURE AND DYNAMICS OF GALAXIES

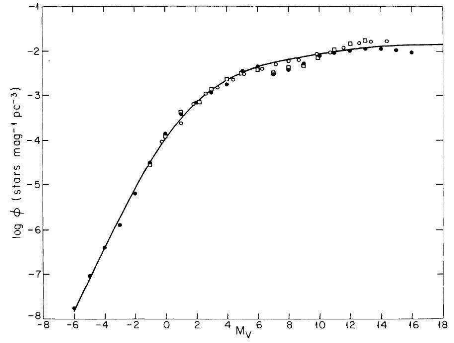
23. Formation of galaxies: The Milky Way Galaxy

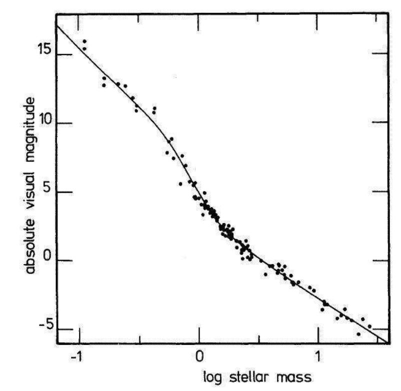
Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/∼vdkruit

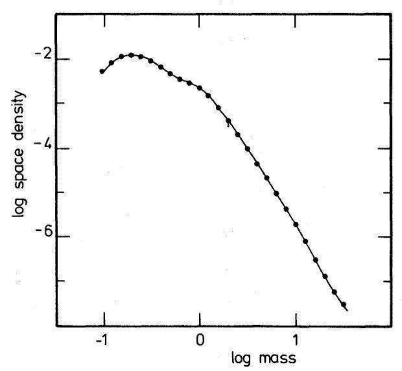
Beijing, September 2011

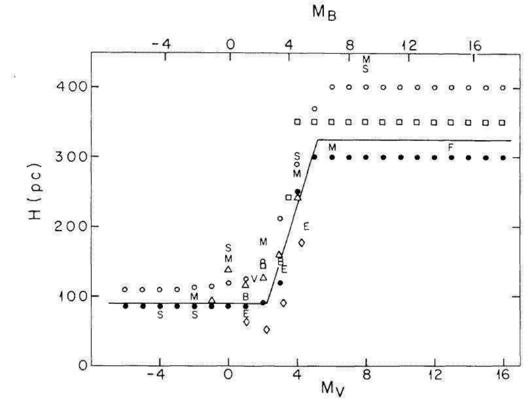
The disk population

The local Mass Function (density in $M_\odot$ pc^{-3} per mass interval of 0.1 log $M_\odot$ as a function of stellar mass) derives from the observed Luminosity Function (number of stars per magnitude interval per pc^3) using the Mass-Luminosity Relation.




absolute visual magnitude
15
10
5
0
-5
-1
0
1
log stellar mass


log space density
-2
-4
-6
-1
0
1
log mass



The old disk scaleheight is about 325 pc.

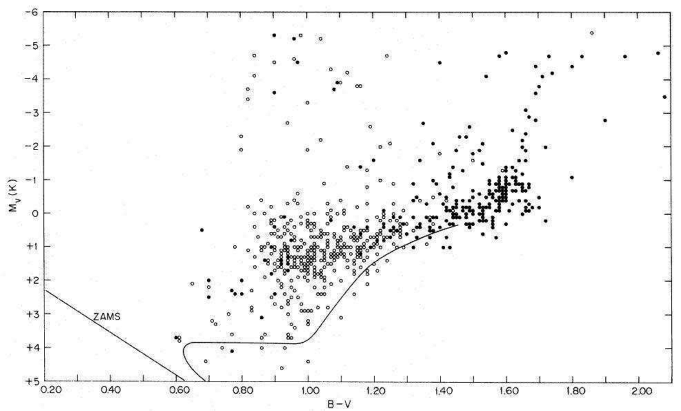
Subgiants and giants have emission ("chromospheric") components in the CaII K-line. The strength of this component gives the absolute magnitude and hence the distance.

This has been done for a sample of about 700 bright stars[1].

The line in the figure (next frame) is the (sub-)giant branch of NGC 188 (age $\approx 10 \times 10^9$ years).

This shows that the old disk population contains stars with ages at least up to that age.

[1] O.C. Wilson, Ap.J. 205, 823 (1976)


-6
-5
-4
-3
-2
-1
0
+1
+2
+3
+4
+5
M_V(K)
ZAMS
0.20
0.40
0.60
0.80
1.00
1.20
1.40
1.60
1.80
2.00
B-V

Eggen, Lynden-Bell and Sandage collapse model

This classical paper ELS[2] contains a study of properties of samples of high- and low-velocity dwarfs.

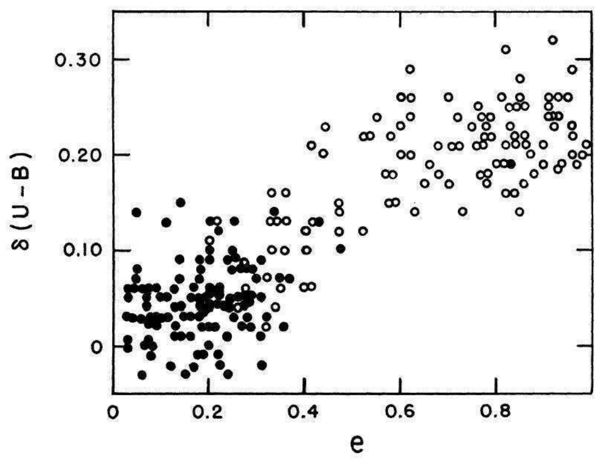
These samples have determinations of parallax, proper motion, radial velocity and photometry and spectral type.

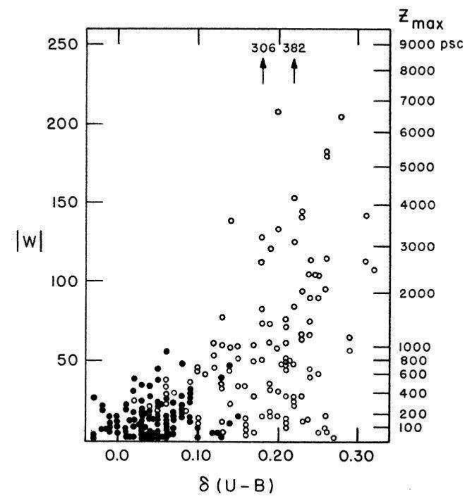
They determined the three components of the space velocity and computed from that the "excentricity" (from the radial excusion in the plane) and the angular momentum.

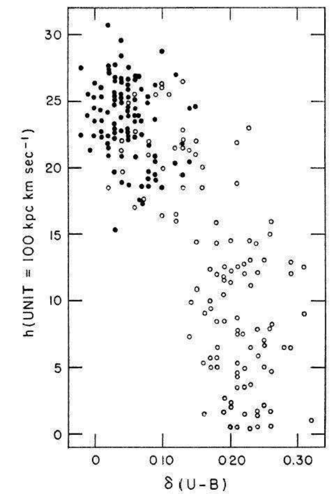
The ultraviolet excess $\delta(U-B)$[3] is an indication of the metallicity, since for these stars it results from line blanketing (more absorption lines in the UV than in the visual).

[2] O.J. Eggen, D. Lynden-Bell & A. Sandage, Ap.J. 136, 748 (1962)
[3] Difference in color observed from that expected from the spectral type

- The vertical velocity, orbital excentricity and angular momentum correlate with the UV-excess.



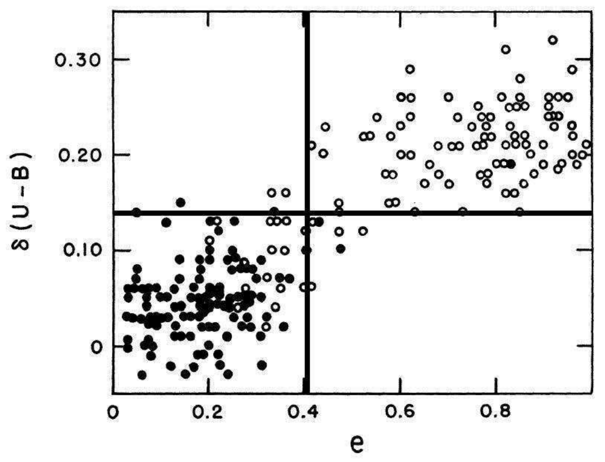

z_{max}
9000 psc
8000
7000
6000
5000
4000
3000
2000
1000
800
600
400
200
100
250
200
150
100
50
|W|
306 382
0.0
0.10
0.20
0.30
δ (U − B)

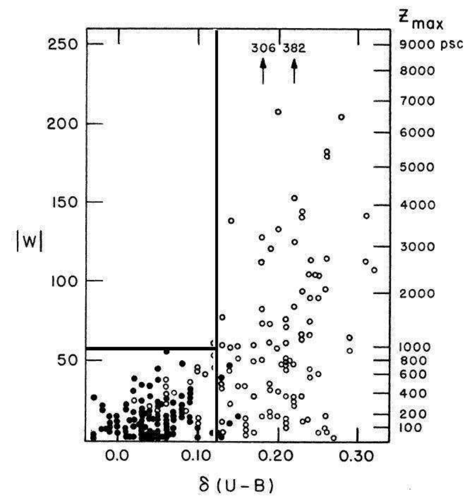

h (UNIT = 100 kpc km sec⁻¹)
30
25
20
15
10
5
0
0
0.10
0.20
0.30
δ (U − B)

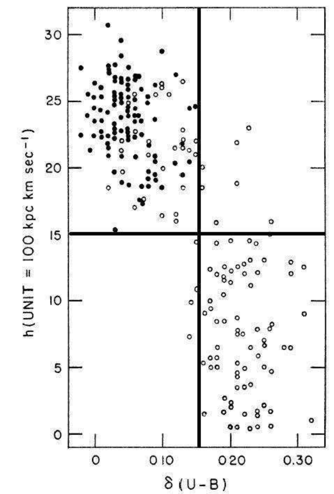
- The continuous progression of metal content from halo to disk stars provides evidence that the Galaxy collapsed.
- Metal-poor stars go up to $z \approx 10$ kpc while the old disk only goes up to $z \approx 400$ pc. The vertical collapse is thus about a factor 25.
- The occurence of very high excentricities among halo stars indicates rapid disk collapse. A strong increase in gravitation will elongate circular orbits when the collapse proceeds on timescale less than the orbital period ($\approx 10^8$ years).
- From the observed angular momentum the estimated radial collapse factor is about 10.

ELS described the process as a continuous one, but even their figures can be interpreted as showing two discrete components.

However, their graphs may be interpreted differently.

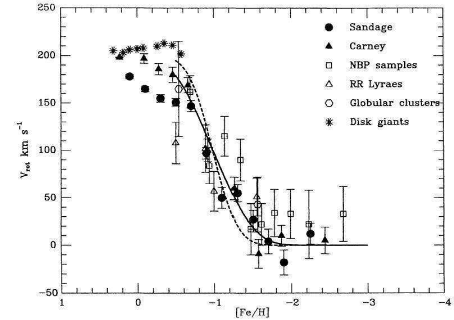



Z max
306
382
9000 psc
8000
7000
6000
5000
4000
3000
2000
1000
800
600
400
200
100
250
200
150
100
50
|W|
0.0
0.10
0.20
0.30
δ (U−B)


30
25
20
15
10
5
0
h (UNIT = 100 kpc km sec⁻¹)
0
0.10
0.20
0.30
δ (U−B)

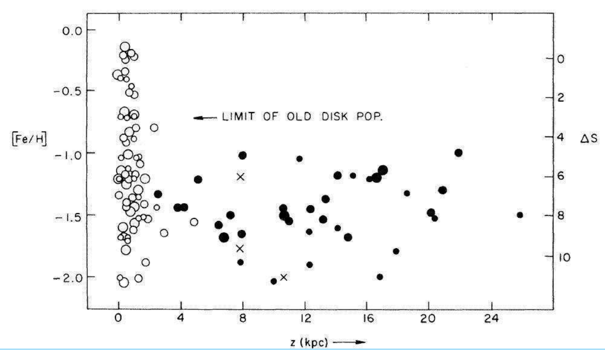
There is more evidence for the basic discreteness of Galactic structure.

One example is the asymmetric drift (the lagging behind in rotation of components with higher velocity dispersion) as a function of metallicity [Fe/H].



Here we see the rotation velocity with respect to an inertial frame.

Also the upper limit of the distribution of metallicity of disk and halo RR Lyrae stars[4] does not show a gradual decline with height above the plane.



[4] T.D. Butler, T.D. Kinman & R.P. Kraft, A.J. 84, 993 (1979)

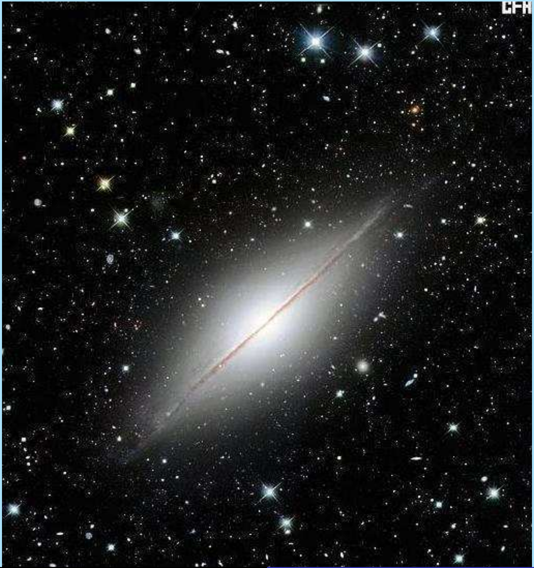
The surface photometry of NGC 7814[5] reveals some important information.

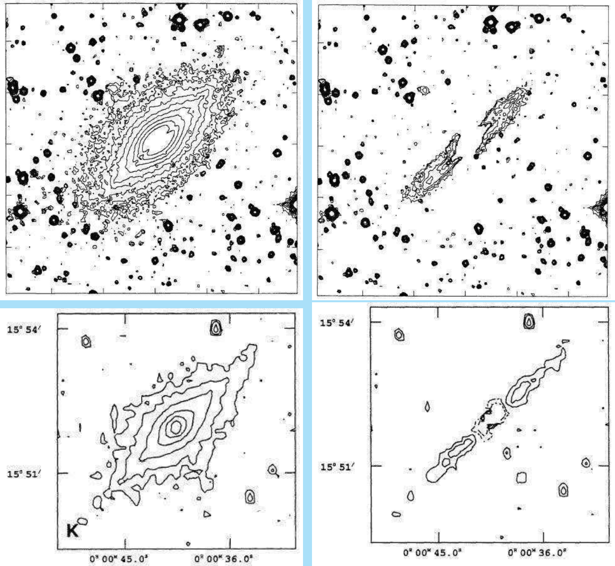
This galaxy is bulge-dominated, but the photometry showed bulge isophotes with all identical axis ratios.

Analysis of the data then showed that it is possible to separate the surface brightness distribution into two distinct components (spheroid and disk) with discretely different flattenings.

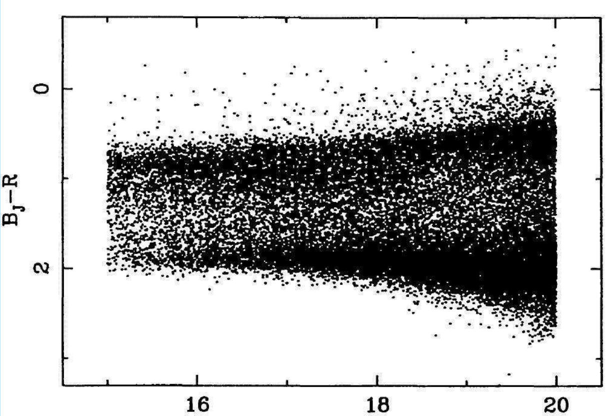
This seemed to indicate that star formation occured in two discrete epochs, one before and one after disk collapse.

[5] P.C. van der Kruit & L. Searle, A.&A. 110, 79 (1982)




15° 54′
15° 51′
K
0ʰ 00ᵐ 45.0ˢ
0ʰ 00ᵐ 36.0ˢ
15° 54′
15° 51′
0ʰ 00ᵐ 45.0ˢ
0ʰ 00ᵐ 36.0ˢ

We see also the basic two-component structure in the colors at faint star counts.



- Few stars bluer than $(B-V) \sim 0.4$. This corresponds to the MS turn-off of the extremely metal-poor halo population.
- The peak at $(B-V) \sim 0.6$. This is the MS turn-off of the halo population.
- The peak at $(B-V) \sim 1.5$. This is the cool MS of the disk population.
- The absence of stars redder than $(B-V) \sim 2.0$. This indicates the absence of large amounts of M-dwarfs to provide the missing local mass.

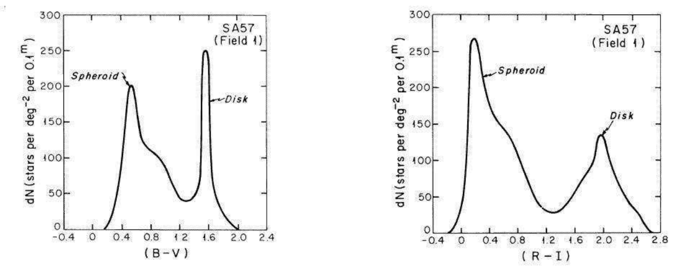
The thick disk

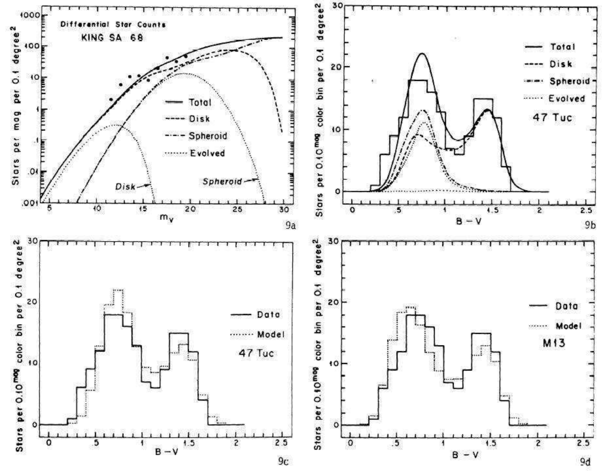
The old situation (before about 1980) was that there was no clear evidence for a substabtial Intermediate Population II and there were basically two discrete components (halo and disk).

Bahcall & Soneira[6] built a Galaxy model with distinct disk and halo components. This was later improved as the Standard Galaxy Model[7].

[6] Ap.J. Suppl. 44, 73 (1980)
[7] Bahcall & Soneira, Ap.J.Suppl. 55.67 (1984)

They showed that it could very well reproduce faint star counts and color distributions in two "Selected Areas", for which deep data were available, namely SA 57 $(l, b) = (65{,}86)$ and SA 68 (111,-46).




Differential Star Counts
KING SA 68
Total
Disk
Spheroid
Evolved
Disk
Spheroid
Stars per mag per 0.1 degree^2
m_V
9a
47 Tuc
Stars per 0.10^{mag} color bin per 0.1 degree^2
B − V
9b
Data
Model
9c
M13
9d

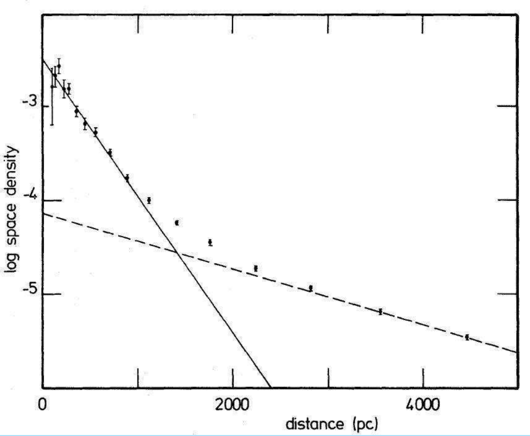
Gilmore & Reid[8] did deep star counts in South Galactic Pole.

They selected only those stars near the MS turn-off on the basis of their colors. For these turn-off and subgiants they determined photometric parallaxes.

They found two components in the disk:

- The "thin disk" (really the old disk) with exponential scaleheight $h_z \approx 300$ pc
- A new component that they called the "thick disk" with $h_z \approx 1350$ pc.
- The local normalisation was such that the thick disk has in the plane $\approx 2\%$ of the stars and this corresponds to $\approx 9\%$ of the face-on surface brightness.

[8] G. Gilmore & N. Reid, Mon.Not.R.A.S. 202, 1025 (1983)


-3
-4
-5
log space density
0
2000
4000
distance (pc)

Bahcall and collaborators[9] conclude that the Standard Galaxy Model BS84 is consistent with counts in all fields available and inconsistent with a model including a thick disk and inconclusive when a metal-rich Luminosity Function (LF) is used for the thick disk.

Gilmore and collaborators[10] present a model with a thick disk (G84) and claims consistency with the count: for the thick disk they use the LF of the globular cluster 47 Tuc ($[Fe/H] \approx -0.7$).

---

[9] J.N. Bahcall & R.M Soneira, Ap.J.Suppl. 55, 67 (1984) (BS84); J.N. Bahcall *et al.*, 299, 616 (1985)

[10] G. Gilmore, Mon.Not.R.A.S. 207, 223 (1984) (G84); G. Gilmore *et al.*, Mon.Not.R.A.S. 213, 257 (1085)

So the earlier disagreement due to choice of the LF of the intermediate component; BS84 and G84 reproduce star counts only if the LF of metal-rich globular cluster is used for it.

The conclusion then is that star counts by themselves are not conclusive evidence for a thick disk or Intermediate Population II.

We first look at external galaxies.

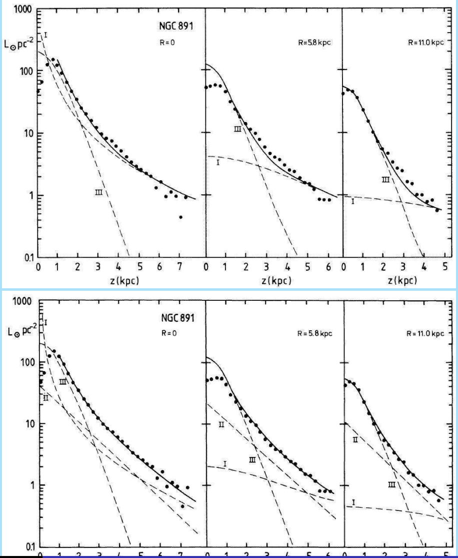
Thick disks in external galaxies

We can look at edge-on external galaxies, such as NGC 891, which is very similar to our Galaxy[11] and construct equivalent "BS84" and "G84" models.

	Galaxy	NGC 891	Galaxy	NGC 891
	"BS84" old disk		"G84" old disk	
h (kpc)	4.5 - 5	4.9	4.5 - 5	4.9
z_o (kpc)	0.6 - 0.7	0.99	0.6 - 0.7	0.99
$R_{\rm max}$ (kpc)	22	21	22	21
$L_{\rm tot}$ ($L_\odot$)	$\sim 1.1 \times 10^{10}$	6.7×10^{10}	$\sim 1.1 \times 10^{10}$	6.7×10^{10}
	"BS84" thick disk		"G84" thick disk	
$h_{\rm R}$ (kpc)	no thick disk	no thick disk	~ 4.5	5
$h_{\rm z}$ (kpc)	no thick disk	no thick disk	~ 1.3	1.5
$L_{\rm tot}$ ($L_\odot$)	no thick disk	no thick disk	$\sim 2 \times 10^8$	2×10^8
	"BS84" spheroid		"G84" spheroid	
$R_{\rm e}$ (kpc)	~ 2.7	2.3	~ 2.7	2.3
$(1-e^2)^{1/2}$	~ 0.7	~ 0.6	~ 0.7	~ 0.6
$L_{\rm tot}$ ($L_\odot$)	$\sim 1.5 \times 10^9$	1.2×10^9	$\sim 1.0 \times 10^9$	4.9×10^8

[11] P.C. van der Kruit, A&A, 140, 470 (1984)

Both these models fit the surface photometry well.



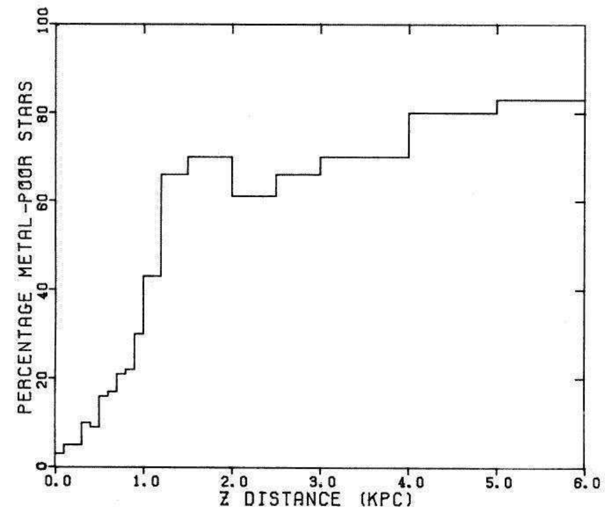
Kinematical evidence for a thick disk

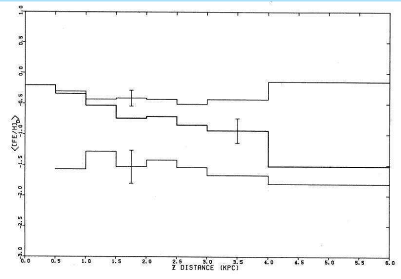
Hartkopf & Yoss[12] compiled DDO photometry and vertical velocities of G & K giants.

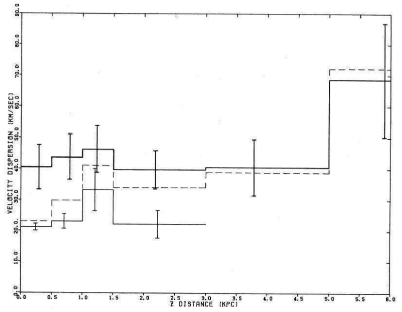
The distribution is separable into two components, each about isothermal:

- $\langle[\mathrm{Fe/H}]\rangle \approx -0.4$; $\langle W^2\rangle^{1/2} \approx 20$ km s^{-1}
- $\langle[\mathrm{Fe/H}]\rangle \approx -1.5$; $\langle W^2\rangle^{1/2} \approx 40$ km s^{-1}

[12] W.I. Hartkopf & K.M. Yoss, 87, 1679 (1982)


PERCENTAGE METAL-POOR STARS
0
20
40
60
80
100
0.0
1.0
2.0
3.0
4.0
5.0
6.0
Z DISTANCE (KPC)





Rose[13] found Red Horizontal Branch stars in the North Galactic Pole field similar to the ones in globular cluster M71 ([Fe/H] ~ -0.6).

These are too metal poor to be old disk stars.

[13] J.A. Rose, A.J. 90, 787 (1985)

They constitute 5% of all non-halo giants in the field and have $h_z \lesssim 0.5$ kpc and $\langle W^2 \rangle^{1/2} \sim 40$ km s^{-1}.

This is fully consistent with the thick disk of Gilmore & Reid.

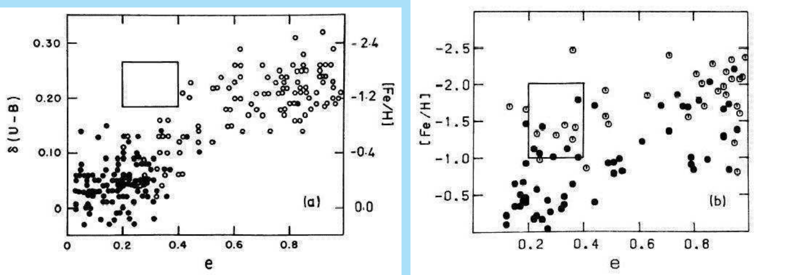
Norris *et al.*[14] found stars with $[\mathrm{Fe/H}] \leq -1.0$; $e \leq 0.4$.

This area is empty in the ELS study and would correspond to positions of stars in an Intermediate Population II.

These stars have $\langle W^2 \rangle^{1/2} = 61 \pm 9$ km s^{-1}.

---

[14] J. Norris, M.S. Bessel & A.J. Pickles, Ap.J.Suppl. 58, 463 (1985)



At the left the ELS diagram. The rectangle gives corresponding areas in both diagrams.

The thick disk is real and could be an Intermediate Population II.

It is probably discrete from the Old Disk Population and possibly also from the Halo Population II.

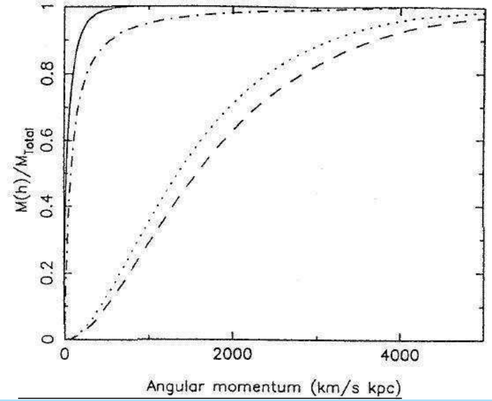
In face-on surface brightness is only of the order of 10% compared to the disk in the solar neighborhood.

So we may distinguish the following components[15] (lengths in kpc and velocities in km/s).

Component	Pop I	Old Disk	Thick disk	Halo
h_z	0.1	0.3	$\sim$ 1.5	$\sim$ 4
$\langle[\mathrm{Fe/H}]\rangle$	$\sim$ 0.0	-0.3	-0.6	-1.5
$\sigma_{[\mathrm{Fe/H}]}$	$\sim$ 0.15	$\sim$ 0.2	$\sim$ 0.3	$\sim$ 0.5
Asym. Drift	small	$\sim$ 10	$\sim$ 40	$\sim$ 150
$\langle W^2\rangle^{1/2}$	$\sim$ 10	25	45	100

[15] See also G. Gilmore, R.F.G. Wyse & K.H. Kuijken, Ann.Rev.A.&A. 27, 555 (1989)

It is possible to make an estimate of the cumulative distribution $M(h)/M_{\rm total}$ of specific angular momentum[16] h in each of these components.



The solid line is the bulge, the dashed-dotted line the halo, the dotted curve the thick disk and the dashed curve the old (thin) disk.

The bulge is related to the halo, but the thick disk to the disk.

[16] Angular momentum per unit mass

Globular clusters

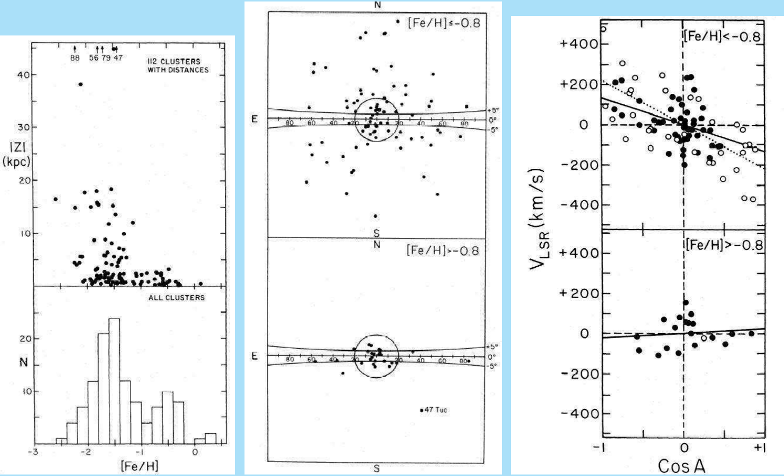
Globular clusters have long been known to be made up of two sub-systems, one following the traditional halo and with metal-poor clusters and one flattened and with less metal-poor systems.

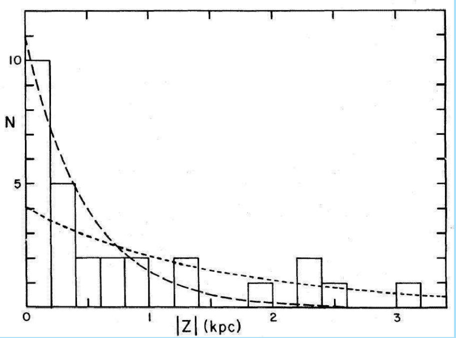
These have been called G- and F-clusters or disk- and halo-clusters.

They also display a bi-modal metallicity distribution with a division at $[\mathrm{Fe/H}] \approx -0.8$.

Also there is a clear difference in asymmetric drift (or rotation velocity of the group as a whole) and velocity dispersion.

This is seen in the radial velocity with respect to the Local Standard of Rest (LSR) as a function of the angle A with the apex of the LSR.


88 56 79 47
112 CLUSTERS WITH DISTANCES
|Z| (kpc)
ALL CLUSTERS
N
[Fe/H]
[Fe/H]≤-0.8
[Fe/H]>-0.8
N
S
E
47 Tuc
[Fe/H]<-0.8
[Fe/H]>-0.8
V_{LSR} (km/s)
Cos A

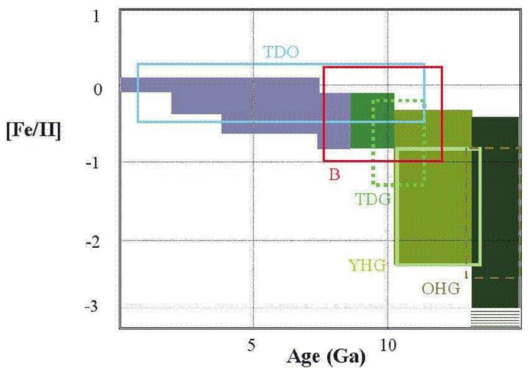


The metal-rich clusters form a disk-system with properties much like the thick disk[17].

[17] R. Zinn, Ap.J. 293, 424 (1985)

	disk-clusters	halo-clusters
[Fe/H]	> -0.8	< -0.8
h_{z} (kpc)	0.5-1.5	–
V_{rot} (km/s)	152 ± 29	50 ± 23
σ_{los} (km/s)	72 ± 11	116 ± 9

A summary picture of the structure of the Galaxy is given in this age-metallicity relation[18].

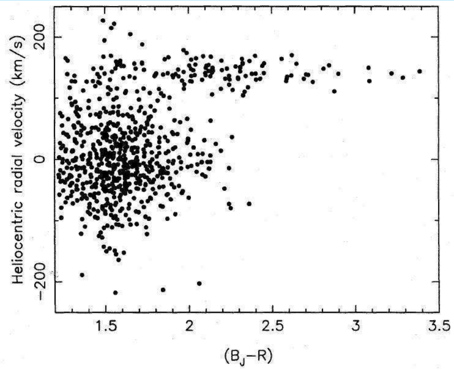


TDO = thin disk open clusters
TDG = thick disk globular clusters
B = bulge
YHG = young halo globular clusters
OHG = old halo globular clusters

[18] K.C. Freeman & J. Bland-Hawthorn, Ann.Rev.A.&A. 40, 487 (2002)

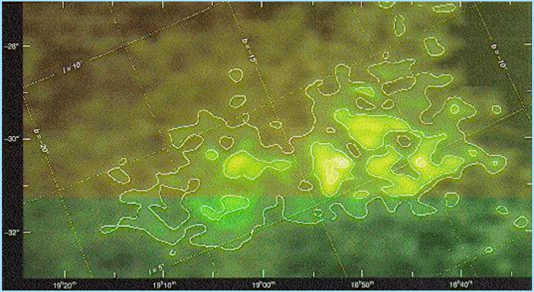
The Sagittarius dwarf

In the course of a study of the kinematics of a sample of stars in the Galactic bulge[19] a curious feature in the distribution was found.



[19] R.O. Ibata, G. Gilmore & M.J. Irwin, Mon.Not.R.A.S. 277, 781 (1995)

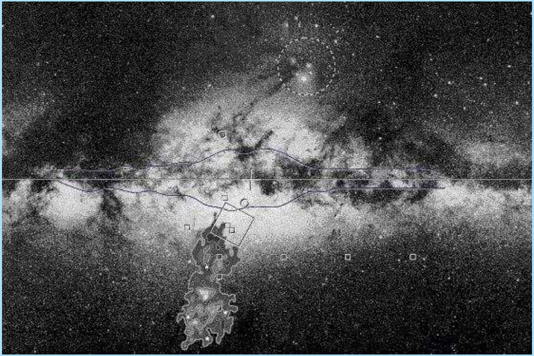
Tracing it accross the sky mapped out the Sagittarius Dwarf.



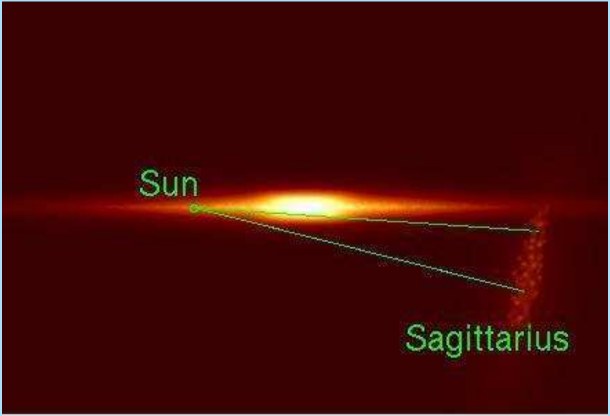
The distance is about 24 kpc and it is comparable in size and luminosity to a large dwarf spheroidal galaxy.

It apparently is approaching the disk of the Galaxy.

Detailed follow-up studies[20] indicate that it is on an orbit with a period of about 1 Gyr and it must have gone through the disk a few times before.



[20] R.A. Ibata, R.F.G. Wyse, G. Gilmore, M.J. Irwin & N.B. Suntzeff, A.J. 113, 634 (1997)


Sun
Sagittarius

STRUCTURE AND DYNAMICS OF GALAXIES

24. The formation of galaxies

Piet van der Kruit
Kapteyn Astronomical Institute
University of Groningen, the Netherlands
www.astro.rug.nl/~vdkruit

Beijing, September 2011

Galaxies at high redshift

First we look at some results of the Sloan Digital Sky Survey (SDSS)[1], that surveyed a large part of the northern sky outside the Galactic Plane in five optical wavelenght bands.

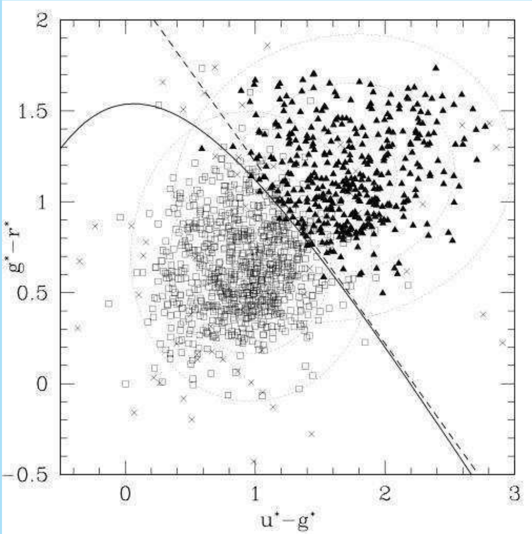
In the SDSS there is a routine to identify galaxies and do photometry and of many objects low-resolution spectra are taken.

The following is from a study[2] of the colors a sample of almost 150,000 galaxies at high Galactic latitude, of which 287 have been studied for morphology and 500 have spectra.

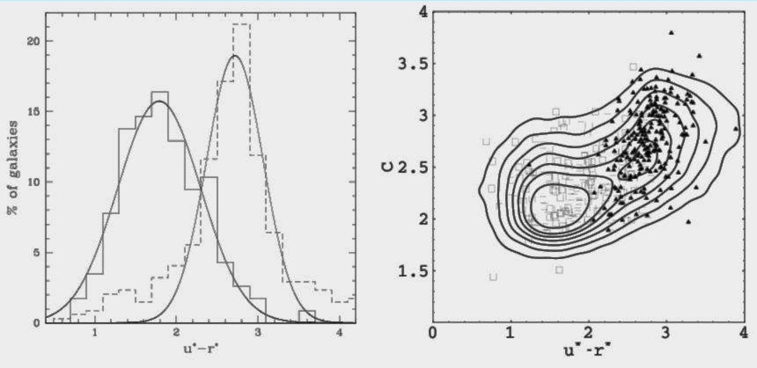
[1] D.G. York et al.. A.J. 120, 1579 (2000)
[2] Strateva et al., Ap.J. 122, 1861 (2001)

Fron statistical studies it turns out that the distribution of the colors is bi-modal.

The separator is at $(u^* - r^*) = 2.22$ (dashed line).



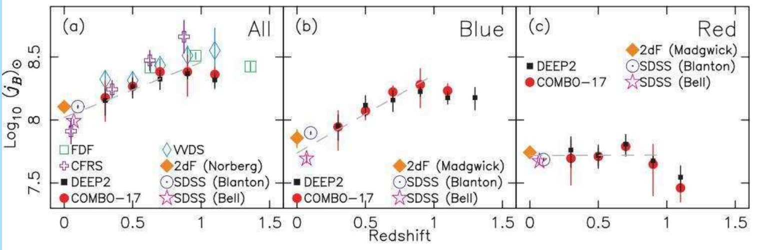
The important result is that the red peaks correspond to early types (E, S0, Sa) and the blue peak to late types (Sb, Sc, Irr).



c is a concentration index; triangles early, squares late types.

It is possible to study the time evolution of both groups from samples at different redshifts.

An extensive study[3] shows that the luminosity density of blue galaxies has decreased by 0.6 dex since $z \sim 1$, while that for the red galaxies has remained constant.



[3] S.M. Faber et al., Ap.J.665, 265 (2007)

When one looks in detail[4] there is little change in the mass function of massive galaxies with redshift out to $z \sim 1$.

However, the morphological mix changes, with more early-types at later times.

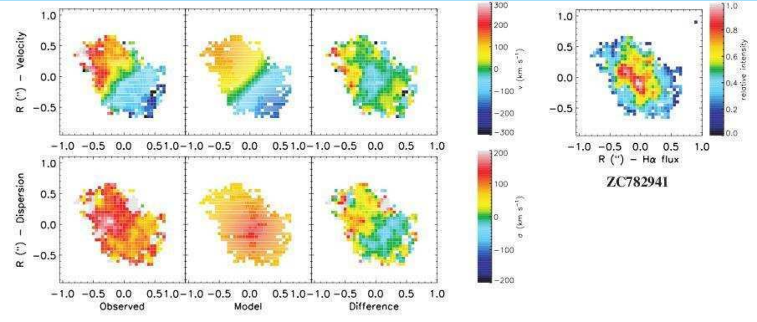
E/S0 galaxies dominate the higher mass population, spirals that at lower masses. The transition changes from $(1-2) \times 10^{11}$ M$_\odot$ at $z \sim 1$ to 3×10^{10} M$_\odot$ at $z = 0$

This "downsizing" phenomenon means that the most massive galaxies stop forming stars first and lower mass galaxies later.

[4] K. Bundy, R.S. Ellis & C.J. Conselice, Ap.J. 625, 621 (2005)

More massive galaxies then evolve into spheroidal systems at earlier times, and this morphological transformation may be completed 1-2 Gyr after star formation ceases.

It is possible now to derive velocity fields of star-forming galaxies at large redshift through emission lines.[5]

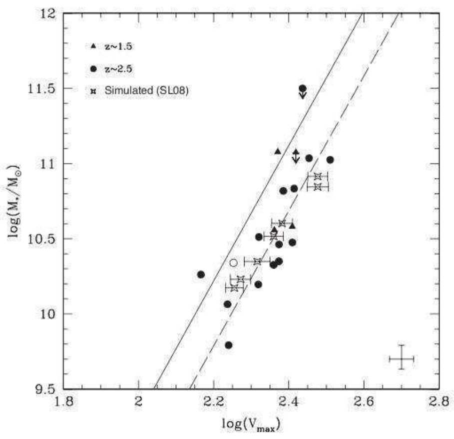


[5] Cresci et al., Ap.J. 697, 115 (2009)

A large fraction of star-forming galaxies at $z \sim 2$ have large, rotating disks.

The dynamical mass correlates with the stellar mass from stellar population models

They show a (stellar mass) Tully-Fisher relation with the same slope as at present, but with different zero-point.



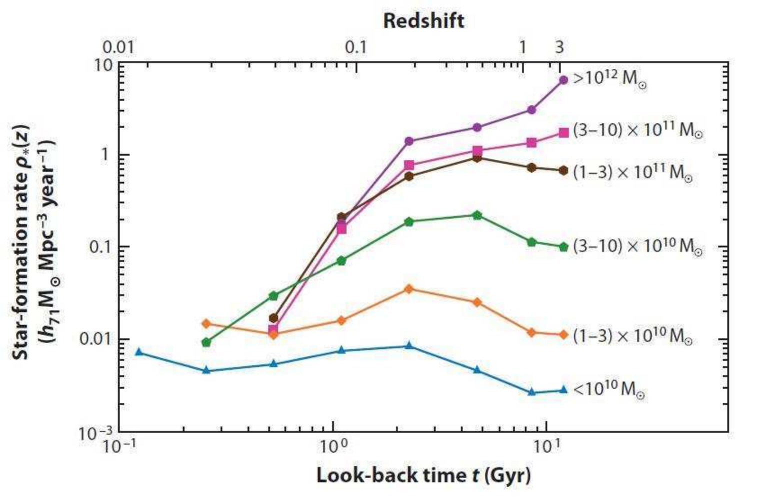
The disks at high redshift have regular velocity fields and result most likely not from mergers, but rather from smooth, but rapid gas inflow.

- In most massive galaxies star-formation ceased early and the result was elliptical galaxies.
- Many of the current disks in large galaxies are basically already in place at redshifts of about 2, when the Universe was only $1-2$ Gyr old.
- Bulges probably formed later and are still forming from merging and capturing of satellites.

Important is the concept of down-sizing.

It says that star formation in the early Universe took predominantly place in larger systems; currently in smaller galaxies.

This may be connected to merging, which was extensive and early in massive galaxies.


Redshift
0.01
0.1
1
3
10
1
0.1
0.01
10^-3
10^-1
10^0
10^1
Star-formation rate ρ*(z)
(h71 M⊙ Mpc^-3 year^-1)
Look-back time t (Gyr)
>10^12 M⊙
(3–10) × 10^11 M⊙
(1–3) × 10^11 M⊙
(3–10) × 10^10 M⊙
(1–3) × 10^10 M⊙
<10^10 M⊙

Galaxy formation

Background

a. Two paradigms

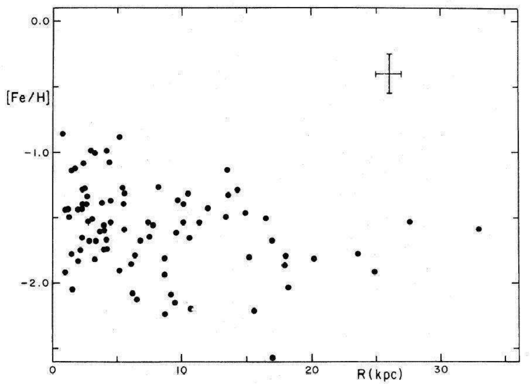
ELS[6] studied the motions of stars in the solar neighborhood and found correlations between metal abundance and the kinematics.

They concluded that the Galaxy was formed during a relatively rapid collapse.

SZ[7] studied the abundance distributions of globular clusters.

[6] O.J. Eggen, D. Lynden-Bell & A.R. Sandage, Ap.J. 136, 748 (1962)
[7] L. Searle & R. Zinn, Ap.J. 225, 357 (1978)

Beyond 8 kpc from the center the distribution over abundance is fairly wide, but does not change with galactocentric distance.

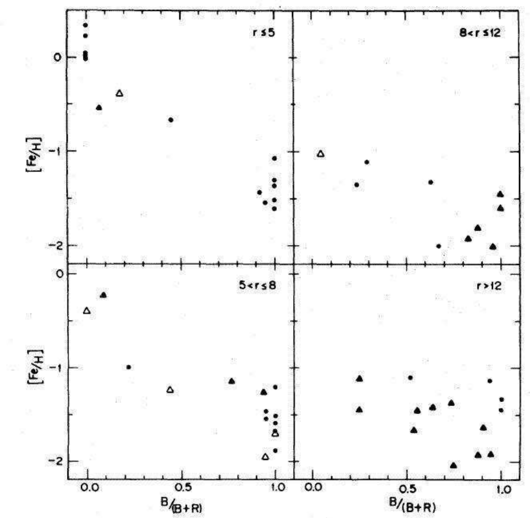


The extent of the Horizontal Branch depends in first instance on metallicity.

However, it has been known that the HB-morphology varies also among clusters of the same metallicity.

This is called the second parameter.

It is characterized by the parameter $B/(B+R)$. B is the number of HB-stars to the blue of the RR-Lyrae gap and R the number to the red.


r ≤ 5
8 < r ≤ 12
5 < r ≤ 8
r > 12
[Fe/H]
[Fe/H]
0
-1
-2
0.0
0.5
1.0
B/(B+R)
B/(B+R)

SZ found that in the tightly bound inner regions $B/(B+R)$ correlates well with abundance, but in the outer halo there is a great diversity of HB-morphology at a given abundance.

They suggested that the second parameter is age.

They concluded that all the above is consistent with a picture in which the build-up of the halo occurs over an extended period during which small fragments (of up to $\sim 10^8$ M$_\odot$ or so) continue to fall in.

These fragments loose gas after a while (due to supernova explosions) and will have a mean metal abundance equal to the effective yield.

The effective yield will have a range and distribution that is stochastic and should show no correlation with galactocentric distance.

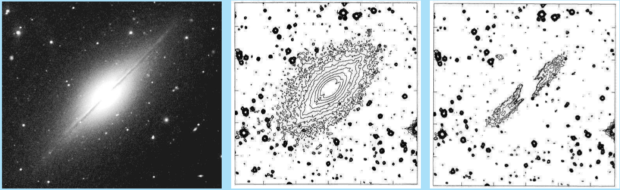
b. Basic two-component structure.

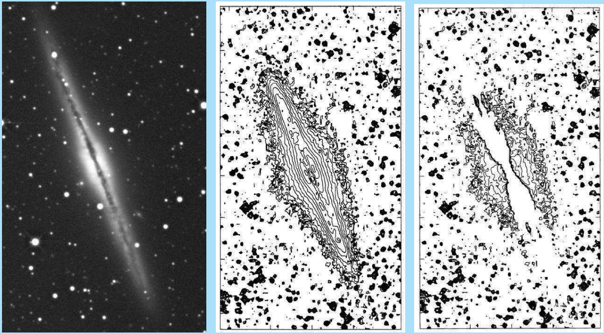
In spite of the possible presence of a thick disk (which has of order 10% of the disk mass), spiral galaxies are basically consist of two distinct components with discrete flattening[8].
This seems to point to two discrete epochs of star formation:

Before collapse – dissipationless – Population II

After collapse – dissipational – Thick and thin disk; Population I

[8]P.C. van der Kruit & L. Searle, A.&A. 110, 79 (1982)





Bulge formation

The observed properties of bulges are:

- $R^{1/4}$-law.
- Generally color (=abundance) gradients
- Isochromes have the same shape as isophotes (in NGC 7814)

Color gradients are often interpreted as evidence for dissipational collapse.

However, then the more metal-rich parts should be more flattened than the metal-poor parts.

Further numerical experiments[9] of dissipationless collapse with violent relaxation shows:

- From irregular initial conditions follows an $R^{1/4}$ distribution
- Statistical conservation of binding energy and thus gradients.

The properties of bulges are consistent with them forming early on in a dissipationless collapse over a longer timescale with fragments falling in for a few Gyr.

[9] T.S. van Albada, Mon.Not.R.A.S. 201, 939 (1982)

Disk formation

Disk formation is of course dissipational.

First we have to look into the question of the origin of angular momentum.

The angular momentum in disks is due to tidal torques between (proto-)galaxies in the early universe[10].

It can be described by a dimensionless parameter

$$\lambda = J|E|^{1/2}G^{-1}M^{-5/2} \approx 0.08$$

where J is the total angular momentum, E the total energy and M the total mass.

[10] P.J.E. Peebles, A.&A. 11, 377 (1969)

Numerical experiments give

$$\lambda = 0.07 \pm 0.03$$

This predicts insufficient angular momentum to explain rotation of disk galaxies in traditional models without dark matter.

The canonical working model has the following characteristics [11]:

- Disk and dark halo have the same distribution of specific angular momentum (= angular momentum per unit mass).
- Disks collapse with detailed conservation of angular momentum.

For tidal torques to work one needs $\sim$ 10 times as much mass in dark halo as in the disk.

[11] M. Fall & G. Efstathiou, Mon.Not.R.A.S. 193, 189 (1980)

Finally we need Mestel's hypothesis[12].

Mestel noted that the rotation and mass distribution in the disk of the Galaxy gave a distribution of specific angular momentum similar to that of a uniformly rotating, uniform sphere.

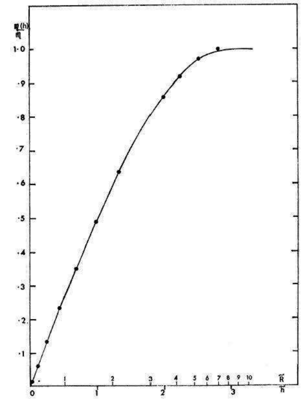
The hypothesis then is that disks form from such a Mestel-sphere with detailed conservation of angular momentum.

The normalized distribution of specific angular momentum $h_{\rm s}$ in the Mestel sphere is

$$\frac{M(h_{\rm s})}{M} = 1 - \left(1 - \frac{h_{\rm s}}{h_{\rm max}}\right)^{3/2}$$

[12] L. Mestel, Mon.Not.R.A.S. 126, 553 (1963)

Freeman[13] has noted already that the self-gravitating exponential disk also has roughly this distribution.



The curve is for the exponential disk and the points for the Mestel sphere.

[13] K.C. Freeman, Ap.J. 160, 811 (1970)

Gunn[14] noted that in a flat rotation curve the Mestel distribution would in centrifugal equilibrium give an approximately exponential radial surface density distribution.

On this basis we can consider the following scenario for disk galaxy formation[15].

We make the following assumptions based on the discussion above:

- The protogalaxy is a Mestel sphere.
- The angular momentum results from tidal torques and $\lambda \sim 0.07$.
- There is a uniform mix of dark and luminous matter (so they have the same specific angular momentum distribution).

[14] J.E. Gunn, in "Astrophysical Cosmology", ed. Brück, Coyne & Longair, Pont. Acad. Scient, Vatican, p. 233 (1982)

[15] P.C. van der Kruit, A.&A. 173, 59 (1987)

For the protogalaxy the total mass is M, and at maximum expansion the density is $\rho_\circ$ and the radius $R_\mathrm{m} = (3M/4\pi\rho_\circ)^{1/3}$.

At maximum expansion then the potential energy is

$$\Omega = -\frac{3GM^2}{5R_\mathrm{m}}$$

and the total angular momentum

$$J = \frac{2}{5} M h_\mathrm{max}$$

At maximum expansion the energy is essentially gravitational ($|E| = |\Omega|$; in virial equilibrium it is a factor 2 smaller). Then

$$h_\mathrm{max} = \frac{5}{2}\left(\frac{5}{3}\right)^{1/2} G^{1/2}\lambda M^{1/2} R_\mathrm{m}^{1/2}$$

Now first consider the halo formation.

There is some star formation in the inner regions to form the Population II stars. These settle dissipationlessly in the bulge.

So we get an $R^{1/4}$-bulge with an abundance gradient.

The dark matter settles dissipationlessly in something like an isothermal sphere.

Assume the amount of dark matter to be

$$M_{\mathrm{H}} = (1 - \Gamma)M$$

Let this settle in an isothermal sphere with radius R_{H}. Then the potential energy can be calculated as

$$\Omega_{\mathrm{H}} = -G(1 - \Gamma)^2 \frac{M^2}{R_{\mathrm{H}}}$$

The viral theorem requires (after completion of the collapse of the dark halo) that

$$E_{\rm H} = \frac{\Omega_{\rm H}}{2} = -G(1-\Gamma)^2 \frac{M^2}{2R_{\rm H}}$$

But originally the energy was

$$E_{\rm H} = -G(1-\Gamma) \frac{3M^2}{5R_{\rm m}}$$

Energy is conserved during dissipationless collapse, so

$$R_{\rm H} = \frac{5}{6}(1-\Gamma)R_{\rm m}$$

$$V_{\rm m}^2 = \frac{6}{5}\left(\frac{G}{1-\Gamma}\right)\frac{M}{R_{\rm m}}$$

Now look at the disk formation.

The remaining gas has a mass ΓM (minus bulge stars, but assume this to be small).

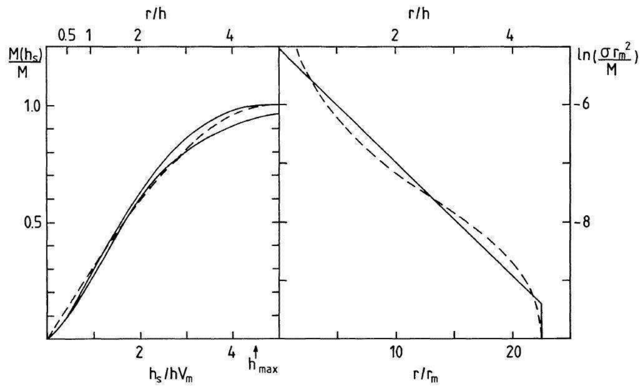
This then settles in a disk *with* dissipation, but conserves the specific angular momentum distribution.

The force field in which this happens is that of the dark halo. Parametrize the final (flat) rotation curve is as

$$V_{\rm rot}^2 = V_{\rm m}^2 \frac{R^2}{R_{\rm m}^2 + R^2} \left[1 - \gamma \ln \left(\frac{R^2}{R_{\rm m}^2 + R^2} \right) \right]$$

From real galaxies we know that the precise value of γ is not important (but $\gamma \approx 0.1$) and $R_{\rm m} \approx (0.1 - 0.5)h$.

Then calculate the surface density distribution of the disk that results; this is a roughly exponential disk with an edge at $\approx 4.5 (\equiv \beta) h$.



On the left we have the specific angular momentum distribution (dashed line is the Mestel distribution; fulldrawn lines are exponential disks with a flat rotation curve for an edge at infinity and $4.5h$).

On the right we see the surface density distribution from the Mestel distribution in the flat rotation curve (dashed) and a pure exponential truncated at $4.5h$.

Assume for simplicity that $\Gamma \ll 1$, so the disk does not seriously affect the force field.

The figure shows an inner excess; this may in reality be the bulge.

How does the thick disk originate? Is at a relic of the violent processes at the moment of disk collapse?

The outer HI beyond the optical truncation and the observed warps may be the result of later infall. Is that why warps start at the optical edge?

From an examination of the figure we deduce the resulting scalelength

$$h = \frac{h_{\mathrm{max}}}{\beta V_{\mathrm{m}}} = \frac{25\lambda}{6\sqrt{2}\beta} \frac{R_{\mathrm{m}}}{(1-\Gamma)^{1/2}}$$

and the central surface density

$$\sigma_{\circ} = \frac{36}{625} \left(\frac{4}{3\sqrt{\pi}} \right)^{2/3} \left(\frac{\beta}{\lambda} \right)^{2} \frac{\Gamma}{1-\Gamma} \rho_{\circ}^{2/3} M^{1/3}$$

In models of hierarchical clustering, galaxies form at about the same time and

$$\frac{\delta\rho}{\rho} \propto M^{-(3+n)/6}$$

with $n =$ -1.5 to 0.

So $\rho_\circ$ is about constant and has only a small dependence on M. Then we get $\sigma_\circ$ about constant for Γ constant.

For $\lambda = 0.07$ and $\beta = 4.5$ we get (V in km s^{-1}, M in M$_\odot$, R in kpc, ρ in M$_\odot$ pc^{-3}, etc.):

$$\frac{\Gamma}{(1-\Gamma)^{1/2}} = 1.5\frac{\sigma_\circ h}{V_{\rm m}^2} \qquad R_{\rm m} = 22\frac{h}{(1-\Gamma)^{1/2}} \qquad R_{\rm H} = 18(1-\Gamma)h$$

$$M = 4.2\times 10^6 (1-\Gamma)^{1/2} V_{\rm m}^2 h \qquad \rho_\circ = 9.7\times 10^8 (1-\Gamma)^2 \frac{V_{\rm m}^2}{h^2}$$

Now apply this to our Galaxy, which has $h = 5$ kpc, $V_{\rm m} = 220$ km s^{-1}, $\sigma_\circ = 400\ M_\odot$ pc^{-2}.

$$\Gamma = 0.06 \qquad R_{\rm m} = 115\,{\rm kpc} \qquad R_{\rm H} = 90\,{\rm kpc} = 18h$$

$$M = 1.0 \times 10^{12}\,{\rm M}_\odot \qquad \rho_\circ = 2 \times 10^{-4}\,{\rm M}_\odot\,{\rm pc}^{-3}$$

For other galaxies we find $\Gamma = 0.04$ - 0.11 and $\rho_\circ \approx 10^{-4}$ M$_\odot$ pc^{-3}.

For $\Omega = 1$ it has been deduced that

$$\frac{\delta\rho}{\rho} = \frac{9\pi^2}{16}$$

For $H = 75$ km s^{-1} Mpc^{-1} this then implies a redshift of galaxy formation of about 3.5.

Finally calculate the disk luminosity

$$L_{\rm disk} = \frac{L}{M}\Gamma^2(1-\Gamma)V_{\rm m}^4\mu_\circ^{-1}$$

So with Freeman's law, constant (M/L) and Γ we get the Tully-Fisher relation.

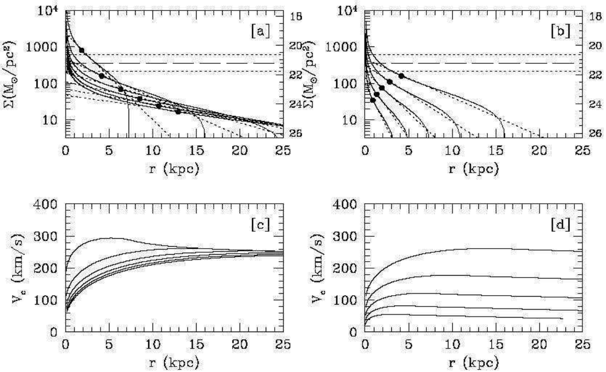
This schematic model has been greatly improved by Dalcanton et al.[16].

They do more realistic calculations, taking all gravitation into account, take a range in λ, etc.

The assumption of a range in λ now translates in a range of predicted central surface densities.

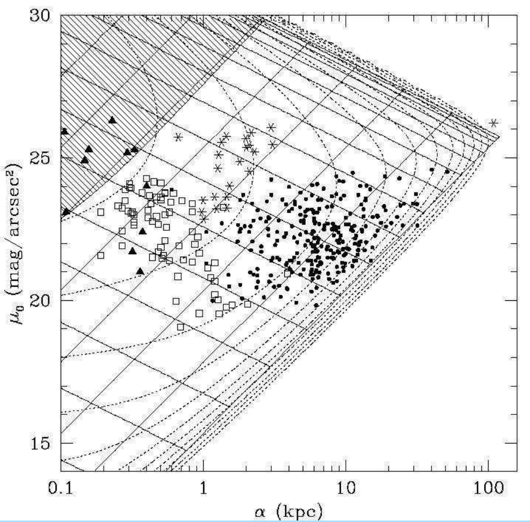
The resulting disk density profiles and rotation curves are in the following figure.

[16] J.J. Dalcanton, D.N. Spergel & F.J. Summers, Ap.J. 482, 659 (1997)



On the left we have models for λ= 0.03 - 0.18; $M = 10^{12}$ $M_\odot$. On the right we have λ= 0.06; $M = 10^{10} - 10^{12}$ $M_\odot$.

The models project in the (surface brightness - scalelength) plane as follows.



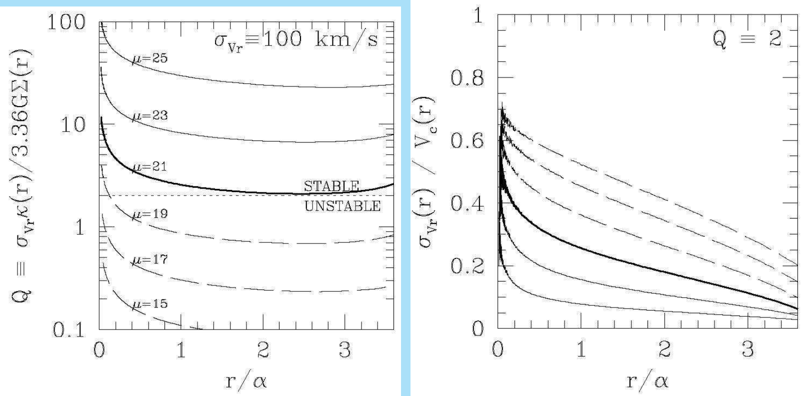
The dashed line are lines of equal expected density in this plane for $M/L = 3$ in B, $\Gamma = 0.05$ and $H = 50$ km s^{-1} kpc^{-1}. This is based on an assumed mass distribution as a Schechter function.

The solid lines with positive slope are of equal mass and those of negative slope of constant angular momentum.

In the hatched region gas pressure is expected to prevend the galaxies from collapsing.

The data are various not statistically complete samples (the filled triangles are Local Group spheroidals).

The disk stability and stellar velocity dispersion as a function of radius gives the following results.



Although in broad terms probably still applicable, this model will have to be augmented to incorporate the effects of infall of companions, such as the Sagittarius Dwarf into our own Galaxy.