Galactic Thin Disk

Of the components of our MILKY WAY GALAXY, the thin disk is the most prominent part to our eyes. It manifests itself as the band of faint light that we see encircling the whole sky. Except for the bulge in the direction of the center of our Galaxy, the stars that make up the Milky Way as we see it are part of the thin disk, just as our Sun is part of that same population of stars. A remarkable fact is that when we look at the Milky Way from the southern hemisphere in winter—when the center of the Galaxy is almost at the zenith at midnight—the appearance is strikingly similar to somewhat later Hubble class, edge-on galaxies such as NGC 891 (of Hubble type Sb).

The young disk

As we will see below, the thin disk is made up of stars of a large range of ages. The youngest have only just formed out of the interstellar medium; the gas and dust between the stars out of which stars are being formed is generally considered to be part of the young disk. The gas can be mapped easily using the 21 cm line of neutral hydrogen (H I) and this has been done extensively for our Galaxy since RADIOASTRONOMY started to be practised in a systematic manner in the 1950s. It took until the 1970s, with the advent of aperture synthesis radio telescopes, until it could be done with sufficient angular resolution in external galaxies. The neutral hydrogen extends usually much further than the stellar disk, often exhibiting a deviation from the plane of the inner galaxy. This can be seen in edge-on galaxies as a curving upward on one side and a symmetric curving downward at the opposite. In moderately inclined galaxies there are kinematic signatures of these so-called warps.

The measurements of the neutral gas have to be supplemented by observations of the distribution of the molecular gas, most abundant in molecular hydrogen H_2. This, however, cannot be observed directly and its distribution is inferred from radio measurements of spectral lines of carbon monoxide, CO. This is transferred into distributions of H_2 by a still rather dubious 'X factor' which descibes the ratio of H_2 versus CO. This X factor is predicted on theoretical grounds and the dubious assumption is made that it is the same from galaxy to galaxy and with galactocentric radius in spite of known variations in heavy-element abundance and radiation field (the latter being important for photodissociation (see PHOTODISSOCIATION REGIONS)). The molecular gas is found to be concentrated to a large extent in giant molecular clouds (GMCs) that contain masses up to $10^6 M_\odot$ (see INTERSTELLAR MOLECULAR CLOUDS). In these GMCs star formation appears to be occurring.

The interstellar gas then is distributed in a very thin layer—much thinner than the older disk stars—with a thickness of order 100 pc. The molecular gas is more concentrated toward the galaxy's center. The velocity distribution at any position in the disk is largely isotropic with (one-dimensional) velocity dispersions of order 7 km s^{-1}.

STAR FORMATION occurs in groups. This is because for the conditions in the interstellar medium a gas cloud can only be unstable with respect to its own gravity if its mass exceeds about 10^5 or so $M_\odot$. This so-called Jeans' mass can be derived crudely as follows. Consider a spherical region of radius R and density ρ, or mass $M = \frac{4}{3}\pi R^3\rho$. The gravitational (potential) energy can be calculated to be

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

The kinetic energy due to the random motions with dispersion $\langle V^2\rangle^{1/2}$ is roughly

$$T = \tfrac{1}{2}M\langle V^2\rangle.$$

The virial theorem then tells us that this region will not contract under its own gravity as long as $\Omega + 2T$ is positive, or in other words as long as there is sufficient kinetic energy to overcome the gravitational force. This then translates into a maximum radius for stability that is called the Jeans' length,

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

and the mass contained within this radius is the Jeans' mass.

In the interstellar medium near the Sun the gas density is of order 1 H atom cm^3 and the velocity dispersion is about 7 km^{-1}. Then $R_{\rm Jeans} \sim 400$ pc. This is more than the thickness of the gas layer. A circular area of the disk with this radius contains about $10^6 M_\odot$ in the form of gas.

Clusters of newly formed stars are called associations. There are two types, depending on the kind of stars seen in them. The most prominent are the OB ASSOCIATIONS; their brightest stars are heavy O and B stars and they are usually accompanied by H II REGIONS, where the interstellar gas is ionized by the UV radiation from these hot stars. The second type is called T associations; these are made up of T TAURI STARS, which are stars of solar mass or somewhat less contracting on their way to the main sequence. The existence of these two types of associations is one basis for a proposed concept of bimodal star formation, in which massive and less massive stars are formed independently.

The distribution of OB associations can be studied in much detail. In agreement with the fact that associations are formed out of the interstellar medium recently, they are also found in our Galaxy in a thin layer, comparable in thickness with that of the gas. The local system of OB stars deviates from that of the plane of the Galaxy; it is confined to a narrow, tilted layer; this is called GOULD'S BELT, since it is reflected in the distribution of OB stars on the sky and this is also generally along a great circle inclined to that of the Milky Way. The velocity distribution of OB stars also deviates from being isotropic, although the velocity dispersion is comparable with that of the gas. The long axis of the velocity distribution does not point toward the Galactic center (has a so-called vertex deviation); this is probably the result of local disturbances in the smooth density distribution in the disk due to the spiral structure of the Galaxy.

Open clusters

Stars are being born in groups or clusters, each representing of the order of $10^5 M_\odot$. Most clusters dissolve in the course of time as a result of the gravitational interactions with the surroundings. Irregularities, such as those associated with spiral structure, are believed to play a major role in this, but also condensations in the interstellar medium, such as the GMCs, are important. It is not known which stars belong to the same cluster in which the Sun was one of the stars being born. Finding out about this is almost certainly impossible. Yet, some clusters or at least their more solidly bound central parts survive, in some cases as long as the age of the disk.

So, we have a whole range of clusters of varying ages. Their ages can be derived from their HERTZSPRUNG–RUSSELL DIAGRAMS. The youngest are the associations discussed above, the oldest—such as the cluster NGC 188—have ages up to about 10^{10} yr. Local examples are the Hyades (consisting of relatively bright stars over a fairly large area of sky and having an age of almost 10^9 yr) and the Pleiades (the brightest stars being visible to the naked eye and having an age of about 10^8 yr).

Star clusters are entities that have been able to retain their structure and to form a relatively securely bound group of stars. Clusters that have dissolved by gravitational tidal effects still may survive as recognizable clusterings in phase space. Even if the prominent structure of clusters in coordinate space has disappeared, their signatures in velocity space may survive on a longer timescale. Structure in the distribution of space velocities of common stars in the neighborhood of the Sun does show clear records of the dynamical history of the disk in the form of the so-called moving groups: collections of stars that are isolated in phase space and appear to share other properties such as age and/or abundance of heavier chemical elements.

The old disk

The old disk contains most of the mass in stars in the disks of our and other galaxies. In spite of the prominence of spiral structure in pictures of DISK GALAXIES, the underlying smooth component contains most of the light (and mass). The fact that spiral structure is so evident in atlases of galaxies derives from the fact that photographic print emphasize the gradients (photography essentially being a logarithmic process). Digital surface photometry shows that even the bright H II regions in the SPIRAL ARMS and the dustlanes on the inside of the arms are only relatively minor disturbances on the underlying brightness distribution. In a vertical column through the disk of our Galaxy in the solar neighborhood, the OB stars contribute only a third or so to the integrated surface brightness. For what follows we can take in external galaxies the surface brightness of the disk to be mainly representing the distribution of older disk stars, even at relatively blue optical wavelengths. The light from the older disk stars is of course dominated by their brightests members, i.e. the giant branch stars among them.

From studies of the surface brightness distributions in spiral galaxies—starting with those by de Vaucouleurs in the 1940s and 1950s—it was found that the radial fall-off from the center onwards is exponential. That is to say, the surface brightness can be fitted to a function of the form

$$L(R) = L_0 \exp\left(-\frac{R}{h}\right)$$

where h is called the scale length. The surface brightness $L(R)$ is expressed here in $L_\odot$ pc^{-2}, but in practice is measured in magnitudes per arcsec^2 and indicated by μ. For the majority of the bright spiral galaxies, Freeman in 1970 found that the central surface brightness was remarkably constant at 21.6 B magnitudes arcsec^{-2}. In the meantime, many galaxies with fainter central surface brightnesses have been found, but these are mostly small dwarf galaxies or a minority of so-called low-surface-brightness galaxies (constituting a small fraction of the mass in stars in the universe), but the result stands for the common disk galaxies such as our own.

The cause of this exponential nature of the distribution of stars appears to be the distribution of angular momentum in the gas from which galaxies form. The angular momentum probably results from tidal interactions between protogalaxies in the early universe, when they were much closer together. Irregularities in their mass distribution made them susceptible to gravitational tidal fields from their neighbors, resulting in a spinning-up of these gas clouds. It is commonly accepted that this, when the protogalactic gas clouds detached themselves from their surroundings in the expanding universe, resulted in entities closely resembling a uniform-density, uniformly rotating sphere. That is to say, the more or less spherical density distribution had a roughly uniform density and rotated everywhere with the same angular velocity. If this distribution of angular momentum is conserved during the collapse of the gas (after some of it was taken away as the stars that form population II in the halo) into a disk, then a roughly exponential distribution as observed will follow. This assumes that the gravitational field is mostly determined by that of a dark halo with a density distribution falling off as R^{-2} as implied by the rotation curves of spiral galaxies. The exponential nature of the light (and presumably mass distributions in galaxy disks) seems to reflect the effect of how galaxies acquired their angular momentum in the early universe.

The older disk stars form a system that is clearly thicker than that of the young population (see also GALACTIC THICK DISK). From the common stars in the solar neighborhood this follows both from counts that can be used to derive their spatial distribution as well as from their kinematics. The velocity dispersion of stars such as our Sun locally is of the order of 20 km s^{-1} in the direction perpendicular to the Galactic plane. This is much higher than that of the young population and consequently the thickness of the thin disk of older stars must be much larger than that of the gas and very young stars. Since this system of old disk stars is almost certainly effectively self-gravitating (that is to say, these stars themselves provide essentially all of the gravitational force in the direction perpendicular to the plane; the z-direction) it can be expected that the spatial density distribution should approximate that of an isothermal sheet. This is a stratified layer in which the velocity dispersion is constant with distance from the plane.

It is well known that such an isothermal, stratified distribution, where stars move under the collective gravitational force of themselves, can be written as

$$\rho(z) = \rho(0)\,\mathrm{sech}^2\left(\frac{z}{z_0}\right)$$

where the vertical scale parameter z_0 is related to the velocity dispersion $\langle V_z^2\rangle^{1/2}$ by

$$z_0 = \frac{\langle V_z^2\rangle}{\pi G\sigma}$$

with σ the surface density (integrated over a vertical column):

$$\sigma = 2\rho(0)z_0.$$

Comparison with the distribution of stars in the solar neighborhood and their kinematics shows that this is a remarkably good approximation to the actual situation. The inferred vertical density distribution at larger distances from the plane approximates an exponential:

$$\mathrm{sech}^2\left(\frac{z}{z_0}\right) \approx 4\exp\left(-2\frac{z}{z_0}\right).$$

This has indeed been observed in the surface brightness distribution in external disk galaxies when seen edge on. What is more remarkable is that the scale parameter z_0 is to an excellent approximation independent of distance from the center. Using the exponential radial distribution mentioned above, it follows that the space distribution of old disk stars in a spiral galaxy can be described very well by

$$\rho(R, z) = \rho(0, 0)\exp\left(-\frac{R}{h}\right)\mathrm{sech}^2\left(\frac{z}{z_0}\right).$$

Dynamical heating

When we look at more detail into the distribution of velocities of common stars in the solar neighborhood we find that their velocity dispersions increase with age. Apparently, stars are being born with the velocity distribution of the interstellar medium from which they form and subsequently gain velocity. Although the interstellar medium appears—as a result of collisions between gas clouds—to be essentially isotropic, the process by which stars on average gain in space motion quickly makes their velocity distribution anisotropic. The gain in velocity is at first rather rapid, but then seems to level off when the age is a few billion years. In the radial direction the gain in velocity is largest, leveling off at about 40 km s^{-1}, while in the vertical direction it reaches only about 20 km s^{-1}. This difference is possible because the motion in the vertical direction is independent of the motion in the Galactic plane. The increase of space motion of disk stars with age is usually referred to as dynamical heating or as secular evolution.

There are two processes that may be responsible for this dynamical heating. The first, proposed by Spitzer and Schwarzschild in the 1950s, is by scattering of stellar orbits by encounters with large mass concentrations in the interstellar medium in the disks; these were later identified with the GMCs mentioned above. This will statistically speed up the motions of the stars, but will quickly become less efficient and the space motions of the stars will level off as observed. This is so because, when the motions increase, stars will spend a larger fraction of the time out of the thin layer where the GMCs reside and the scattering will quickly become less effective.

The second mechanism was proposed in the 1960s by Barbanis and Woltjer and is the scattering of the stars by encounters with spiral arms. These also constitute variations in the gravitational field. This mechanism is less effective in the vertical direction, because the period of vertical motion of stars is very different from that from encountering the spiral structure. In the radial direction there is much more of a resonance between these encounters and radial oscillations and the effect is more efficient. Theoretical studies indicate that the first mechanism would predict the ratio of the vertical to radial velocity dispersions to be larger than observed, so it is likely that (at least in our Galaxy in the neighborhood of the Sun) the second mechanism has been more effective. Still, at various positions in different galaxies the actual dynamical heating is probably accomplished by both processes in varying relative importance.

Random motion of the old disk stars in galaxies is vital for matters of stability. If there were no relative motion of the stars, the disks of old stars would not be able to be as smooth as they are; any statistical density fluctuation would quickly result in gravitational collapse and the disks would break up in many dense concentrations. This is a dilemma that goes back to the days of Isaac Newton: if every star attracts every other one then why would they not have collapsed into a dense concentration? Newton's incorrect answer was that stars are distributed evenly in space, such that every star is pulled equally in all directions by the other stars. We now know that the correct answer is that stability derives from the relative motions of the stars, providing in essence a kind of pressure that prevents these gravitational instabilities from occurring.

Now return to galactic disks. We have seen that the random motions of the stars can stabilize any distribution up to a certain radius. For flat disks we can rewrite this Jeans' criterion as follows. In an infinitesimally thin disk we have a surface density σ and a velocity dispersion $\langle V^2\rangle^{1/2}$. Then the kinetic energy in a region with radius R is

$$T = \tfrac{1}{2}\pi R^2\sigma\langle V^2\rangle$$

and the potential energy is roughly

$$\Omega = \frac{2\pi^3}{3}G\sigma^2 R^3.$$

For this case of a very thin disk then the Jeans' length is

$$R_{\mathrm{Jeans}} = \frac{3\langle V^2\rangle}{2\pi^2 G\sigma}.$$

For the old stellar disk in our Galaxy in the neighborhood of the Sun, the surface density is about $80M_\odot$ pc^{-2} and the velocity dispersion roughly 40 km s^{-1}. Then this radius comes out as about 5 kpc. So we find that random stellar motions can stabilize the old stellar disk only up to moderate scales.

So something else has to be responsible for stabilizing the disk over larger length scales. This stability can be provided by differential rotation, which means that the angular rotation velocity varies with galactocentric distance. Parts further out have a lower angular velocity and as a result of this, seen from any point in the disk, these parts lag behind in the motion around the center, while the parts further in go ahead. This appears for not too large distances as a rotation around this reference position, and the angular velocity at which this occurs is described by the so-called Oort constant B. If the rotation velocity around the center of the galaxy is V_{rot} then

$$B = -\frac{1}{2}\left(\frac{V_{\mathrm{rot}}}{R} + \frac{\mathrm{d}V_{\mathrm{rot}}}{\mathrm{d}R}\right).$$

This provides a centrifugal force that can counteract the gravity and in this manner can provide stability. Consider again a circular area in the disk with radius R and surface density σ. For a particle at the edge this centrifigul force is B^2R. On the other hand, the gravitational force from the matter within R is approximately that due to the mass M inside R, or $GM/R^2 = \pi G\sigma$. Now, contrary to the Jeans' stability mechanism there is a *minimum* radius for stability, because as R increases the centrifugal force due to differential rotation increases. This minimum radius occurs where this centrifugal force equals the attraction due to gravity and is

$$R_{\min} = \frac{\pi G\sigma}{B^2}.$$

In our Galaxy in the solar neighborhood the Oort constant $B \sim -13$ km s^{-1} kpc^{-1}. This then gives $R_{\min} \sim 5$ kpc.

As Alar Toomre first described in 1964, a galactic disk is thus stable at all length scales, if this minimum radius is equal to or smaller than the Jeans' length. From the numbers given, we see that the disk of our Galaxy is indeed stable, but only just. So, disks can be stabilized at small scales by random motions and at large scales by differential rotation. Putting the two together then gives for a disk with given surface density and differential rotation a minimum velocity dispersion required for stability:

$$\langle V^2\rangle^{1/2}_{\min} = \left(\frac{2\pi}{3}\right)^{1/2}\frac{G\sigma}{|B|}.$$

Stars that have a small velocity with respect to the the general rotation will, in a coordinate system that rotates along with this general rotation, descibe a small elliptical orbit (an epicycle) in which it goes around in a period $2\pi/\kappa$ and

$$\kappa^2 = -4B\frac{V_{\mathrm{rot}}}{R}.$$

κ is the so-called epicyclic frequency. In actual galaxies the rotation velocity V_{rot} is constant with radius over most of the disk and then $\kappa = -2\sqrt{2}B$.

Then we can rewrite the equation above as

$$\langle V^2\rangle^{1/2}_{\min} = 4.09\frac{G\sigma}{\kappa}.$$

In a more detailed analysis and for arbitrary rotation curves, Toomre derived this minimum velocity dispersion actually as

$$\langle V^2\rangle^{1/2}_{\min} = 3.36\frac{G\sigma}{\kappa}.$$

Disks of spiral galaxies, including our own, do not appear to be violently unstable. However, there is spiral structure and star formation, so some instability must occur. As it turns out, as best as we can measure the disks are close to the border between stability and instability. This presumably is the cause of the fact that star formation is regulated as a moderately slow process, whereby the conversion of the interstellar gas into stars proceeds at a reasonable rate, as observed from the distribution of stellar ages in the solar neigborhood. As we can derive from the Toomre minimum velocity dispersion criterion above, the minimum dispersion decreases with galactocentric radius, the epicycle frequency decreasing as R^{-1}, but the surface density σ exponentially. This is indeed roughly as observed. Furthermore, it implies also a decrease in the velocity dispersion in the direction perpendicular to the plane in such a way that it roughly produces the constant thickness of the thin disk (the radial constancy of the scale height parameter z_0).

Bibliography

The material presented above (and many more related matters) is discussed in more detail in

Gilmore G, King I R and van der Kruit P C 1990 *The Milky Way as a Galaxy* (University Science Books)

Piet van der Kruit