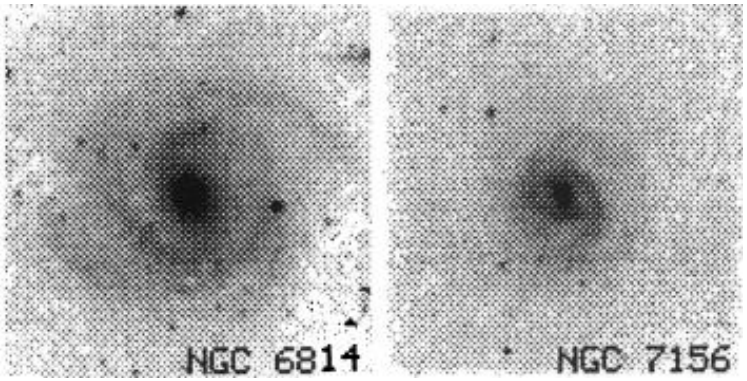
Relevant Material for Lecture 3

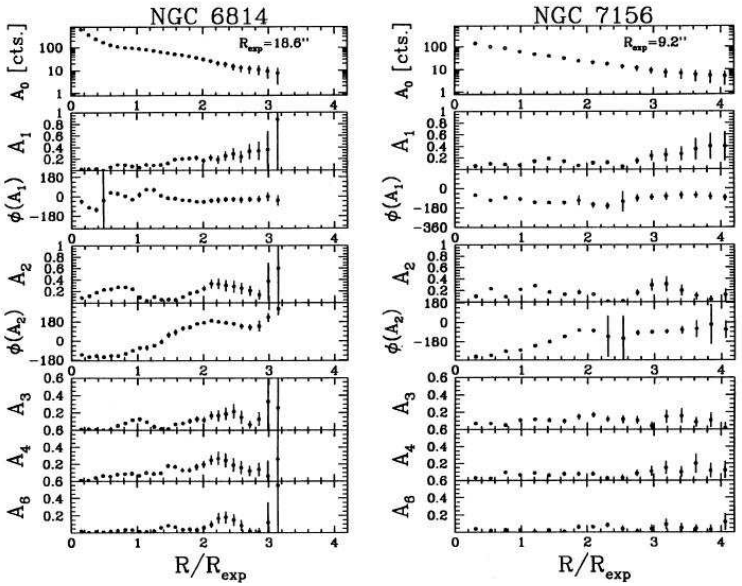
"Galaxies: Structure, Dynamics, and Evolution"

3. **Structure of disc galaxies**

This is the typical image of a spiral

Now analyze in detail





3.1 **How to get spiral arms ?**

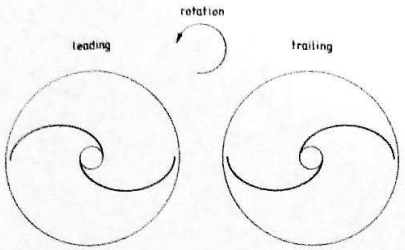


Figure 6.4 Leading and trailing arms.

The strength of the spiral structure can be parametrized by the amplitude of its Fourier components, defined by expressing the surface brightness as a Fourier series (eq. B.66),

$$\frac{I(R,\phi)}{\overline{I}(R)} = 1 + \sum_{m=1}^{\infty} A_m(R)\cos m[\phi - \phi_m(R)] \quad (A_m(R) > 0). \tag{6.1}$$

Here $\overline{I}(R) \equiv (2\pi)^{-1}\int_0^{2\pi} \mathrm{d}\phi\, I(R,\phi)$ is the azimuthally averaged surface brightness at radius R, and A_m and ϕ_m are the amplitude and phase of the mth Fourier component.

If a single Fourier component m dominates the spiral structure, the strength can also be parametrized by the arm-interarm surface-brightness ratio K, which is related to A_m by

$$K = \frac{1 + A_m}{1 - A_m}. \tag{6.2}$$

Most grand-design spiral galaxies have two arms and approximate two-fold rotational symmetry. In near-infrared light, which traces the surface density, the amplitude of the arms lies in the range $0.15 \lesssim A_2 \lesssim 0.6$ (Rix & Zaritsky 1995), corresponding to arm-interarm ratios of $1.4 \lesssim K \lesssim 4$. Grand-design spirals with $m \neq 2$ are rare, although a significant fraction of disk galaxies exhibit lopsided distortions ($A_1 \gtrsim 0.2$) in their outer parts, and careful Fourier decomposition occasionally reveals three-armed spiral patterns (Rix & Zaritsky 1995). The dominance of two-armed patterns in grand-design spirals is a striking observational fact that demands explanation in a successful theory of spiral structure.

(b) Leading and trailing arms Spiral arms can be classified by their orientation relative to the direction of rotation of the galaxy. A **trailing** arm is one whose outer tip points in the direction opposite to galactic rotation, while the outer tip of a **leading** arm points in the direction of rotation (see Figure 6.4).

It is not easy to determine observationally whether the arms of a given galaxy are leading or trailing. In face-on galaxies we cannot determine the

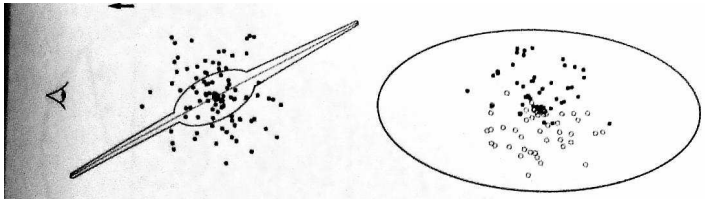


Figure 6.6 Distinguishing near and far sides of a disk galaxy. The dots represent objects such as novae or globular clusters. There is an obscuring dust layer in the central plane of the disk which is shown as a line in the side view at left. In the observer's view, at right, objects behind the dust layer (open circles) are fainter and hence fewer are present in a flux-limited survey.

leading arms (Pasha 1985; Buta, Byrd, & Freeman 2003), and transitory one-armed leading spirals can be produced by plausible dynamical processes, for example encounters with companion galaxies on retrograde orbits. Nevertheless, in the vast majority of cases *spiral arms are trailing*.

(c) The pitch angle and the winding problem The **pitch angle** α at any radius R is the angle between the tangent to the arm and the circle $R = constant$ (see Figure 6.8); by definition $0 < \alpha < 90°$.

It is useful to think of the center of each arm as a mathematical curve in the plane of the galaxy, which we write in the form $\phi + g(R,t) = constant$ where t is the time. Suppose that the galaxy has m-fold rotational symmetry, that is, the arm pattern is unchanged if we rotate the galaxy by $2\pi/m$ radians ($m > 0$). Then a more convenient expression, which defines the locations of all m arms, is

$$m\phi + f(R,t) = \text{constant} \pmod{2\pi}, \tag{6.3}$$

where $f(R,t) \equiv mg(R,t)$ is the **shape function**. It is also useful to introduce the **radial wavenumber**

$$k(R,t) \equiv \frac{\partial f(R,t)}{\partial R}. \tag{6.4}$$

The sign of k determines whether the arms are leading or trailing. If, as we shall always assume, $m > 0$ and the galaxy rotates in the direction of increasing ϕ, then

$$\text{leading arms} \Leftrightarrow k < 0 \quad ; \quad \text{trailing arms} \Leftrightarrow k > 0. \tag{6.5}$$

The pitch angle is given by

$$\cot\alpha = \left| R\frac{\partial\phi}{\partial R} \right|, \tag{6.6}$$

We now conduct a simple thought experiment. At some initial time $t = 0$ we paint a narrow stripe or arm radially outward across the disk of a galaxy. The initial equation of the stripe is $\phi = \phi_0$, where ϕ is the azimuthal angle (Figure 6.8). The disk rotates with an angular speed $\Omega(R)$, where R is the distance from the center of the disk. The disk is said to be in **differential rotation** if $\Omega(R)$ is not independent of R. When the disk is in differential rotation the arm does not remain radial as the disk rotates. The location of the arm $\phi(R, t)$ is described by the equation

$$\phi(R, t) = \phi_0 + \Omega(R)t. \tag{6.8}$$

The pitch angle is given by equation (6.6),

$$\cot \alpha = Rt \left| \frac{\mathrm{d}\Omega}{\mathrm{d}R} \right|. \tag{6.9}$$

For a galaxy with a flat circular-speed curve, $R\Omega(R) = v_c = 200\,\mathrm{km\,s^{-1}}$, $R = 5\,\mathrm{kpc}$, and $t = 10\,\mathrm{Gyr}$, the pitch angle would now be $\alpha = 0.14°$, far

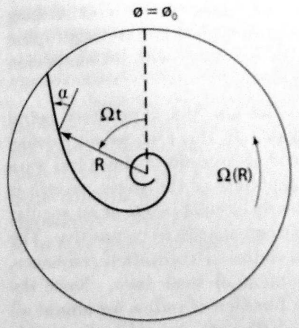


Figure 6.8 How a material arm winds up in a differentially rotating disk. The rotation law is $\Omega(R) \propto R^{-1}$.

smaller than observed pitch angles. This discrepancy is called the **winding problem**: if the material originally making up a spiral arm remains in the arm, the differential rotation of the galaxy winds up the arm in a time short compared with the age of the galaxy. A remarkably clear statement of the winding problem was given over a century ago by Wilczynski (1896).

3.2.3 Nearly circular orbits: epicycles and the velocity ellipsoid

In disk galaxies many stars are on nearly circular orbits, so it is useful to derive approximate solutions to equations (3.68a) that are valid for such orbits. We define

$$x \equiv R - R_\mathrm{g}, \tag{3.75}$$

where $R_\mathrm{g}(L_z)$ is the guiding-center radius for an orbit of angular momentum L_z (eq. 3.72). Thus $(x, z) = (0, 0)$ are the coordinates in the meridional plane of the minimum in Φ_eff. When we expand Φ_eff in a Taylor series about this point, we obtain

$$\Phi_\mathrm{eff} = \Phi_\mathrm{eff}(R_\mathrm{g}, 0) + \tfrac{1}{2}\left(\frac{\partial^2\Phi_\mathrm{eff}}{\partial R^2}\right)_{(R_\mathrm{g},0)} x^2 + \tfrac{1}{2}\left(\frac{\partial^2\Phi_\mathrm{eff}}{\partial z^2}\right)_{(R_\mathrm{g},0)} z^2 + \mathrm{O}(xz^2). \tag{3.76}$$

Note that the term that is proportional to xz vanishes because Φ_eff is assumed to be symmetric about $z = 0$. The equations of motion (3.68a) become very simple in the **epicycle approximation** in which we neglect all terms in Φ_eff of order xz^2 or higher powers of x and z. We define two new quantities by

$$\kappa^2(R_\mathrm{g}) \equiv \left(\frac{\partial^2\Phi_\mathrm{eff}}{\partial R^2}\right)_{(R_\mathrm{g},0)} \quad ; \quad \nu^2(R_\mathrm{g}) \equiv \left(\frac{\partial^2\Phi_\mathrm{eff}}{\partial z^2}\right)_{(R_\mathrm{g},0)}, \tag{3.77}$$

for then equations (3.68a) become

$$\ddot{x} = -\kappa^2 x, \tag{3.78a}$$

$$\ddot{z} = -\nu^2 z. \tag{3.78b}$$

According to these equations, x and z evolve like the displacements of two harmonic oscillators, with frequencies κ and ν, respectively. The two frequencies κ and ν are called the **epicycle** or **radial frequency** and the **vertical frequency**. If we substitute from equation (3.68b) for $\Phi_{\rm eff}$ we obtain[3]

$$\kappa^2(R_{\rm g}) = \left(\frac{\partial^2\Phi}{\partial R^2}\right)_{(R_{\rm g},0)} + \frac{3L_z^2}{R_{\rm g}^4} = \left(\frac{\partial^2\Phi}{\partial R^2}\right)_{(R_{\rm g},0)} + \frac{3}{R_{\rm g}}\left(\frac{\partial\Phi}{\partial R}\right)_{(R_{\rm g},0)}, \tag{3.79a}$$

$$\nu^2(R_{\rm g}) = \left(\frac{\partial^2\Phi}{\partial z^2}\right)_{(R_{\rm g},0)}. \tag{3.79b}$$

Since the circular frequency is given by

$$\Omega^2(R) = \frac{1}{R}\left(\frac{\partial\Phi}{\partial R}\right)_{(R,0)} = \frac{L_z^2}{R^4}, \tag{3.79c}$$

equation (3.79a) may be written

$$\kappa^2(R_{\rm g}) = \left(R\frac{{\rm d}\Omega^2}{{\rm d}R} + 4\Omega^2\right)_{R_{\rm g}}. \tag{3.80}$$

Note that the radial and azimuthal periods (eqs. 3.17 and 3.19) are simply

$$T_r = \frac{2\pi}{\kappa} \quad ; \quad T_\psi = \frac{2\pi}{\Omega}. \tag{3.81}$$

Very near the center of a galaxy, where the circular speed rises approximately linearly with radius, Ω is nearly constant and $\kappa \simeq 2\Omega$. Elsewhere Ω declines with radius, though rarely faster than the Kepler falloff, $\Omega \propto R^{-3/2}$, which yields $\kappa = \Omega$. Thus, in general,

$$\Omega \lesssim \kappa \lesssim 2\Omega. \tag{3.82}$$

Using equations (3.19) and (3.81), it is easy to show that this range is consistent with the range of $\Delta\psi$ given by equation (3.41) for the isochrone potential.

[3] The formula for the ratio κ^2/Ω^2 from equations (3.79) was already known to Newton; see Proposition 45 of his *Principia*.

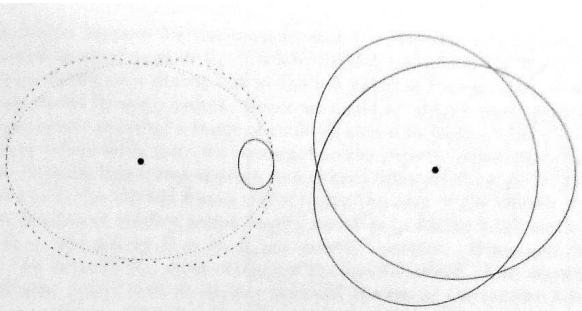


Figure 6.10 The appearance of elliptical orbits in a frame rotating at $\Omega_{\rm p} = \Omega - n\kappa/m$. Left: $(n,m) = (0,1)$, solid line; $(1,2)$, dotted line; $(1,-2)$, dashed line. Right: $(n,m) = (2,3)$.

6.2 Wave mechanics of differentially rotating disks

6.2.1 Preliminaries

(a) Kinematic density waves The galactocentric distance of a particle that orbits in the equatorial plane of an axisymmetric galaxy is a periodic function of time with period T_r (see eq. 3.17). During the interval T_r the azimuthal angle increases by an amount $\Delta\phi$ (eq. 3.18b). These quantities are related to the radial and azimuthal oscillation frequencies $\Omega_r = 2\pi/T_r$ and $\Omega_\phi = \Delta\phi/T_r$. In general, $\Delta\phi/(2\pi)$ is irrational, so the orbit forms a rosette figure such as the one shown in Figure 3.1.

Now suppose that we view the orbit from a frame that rotates at angular speed $\Omega_{\rm p}$. In this frame, the azimuthal angle is $\phi_{\rm p} = \phi - \Omega_{\rm p}t$, which increases in one radial period by $\Delta\phi_{\rm p} = \Delta\phi - \Omega_{\rm p}T_r$. Therefore we can choose $\Omega_{\rm p}$ so that the orbit is closed; in particular, if $\Delta\phi_{\rm p} = 2\pi n/m$, where m and n are integers, the orbit closes after m radial oscillations. In this case

$$\Omega_{\rm p} = \Omega_\phi - \frac{n\Omega_r}{m} \simeq \Omega - \frac{n\kappa}{m}, \tag{6.24}$$

where in the last equality we have approximated Ω_ϕ and Ω_r by their values for nearly circular orbits, the circular frequency Ω and the epicycle frequency κ (see eqs. 3.79). The appearance of the closed orbits in the rotating frame is shown in Figure 6.10.

In general $\Omega(R) - n\kappa(R)/m$ will be a function of radius, so no single choice for $\Omega_{\rm p}$ can ensure that orbits at all radii are closed. In Figure 6.11 we show the behavior of $\Omega - n\kappa/m$ for several values of m and n. The curves are

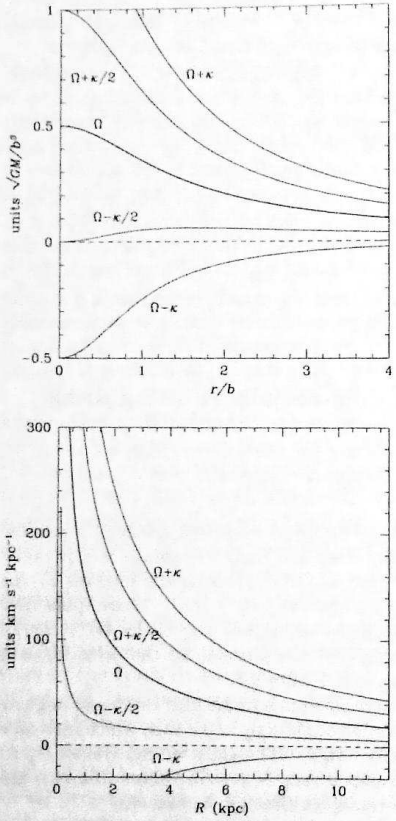


Figure 6.11 Behavior of $\Omega - n\kappa/m$ in: (top) the isochrone potential (eq. 2.47); (bottom) Model I for our Galaxy, described in §2.7.

plotted for two representative galactic circular-speed curves, the isochrone potential (eq. 2.47) and Model I for our Galaxy, as described in Table 2.3.

This diagram exhibits an intriguing fact noticed by Lindblad many decades ago: while most of the $\Omega - n\kappa/m$ curves vary rapidly with radius, the curve for $n = 1, m = 2$ (or $n = 2, m = 4$, etc.) is relatively constant across much of the galaxy.[3] To understand the significance of Lindblad's remark, let us suppose for the moment that $\Omega - \frac{1}{2}\kappa$ were *exactly* constant, and equal to some number Ω_p. Then in a frame rotating at Ω_p the orbits of the type shown as a dotted line in the left panel of Figure 6.10a would be exactly closed at every radius. Hence we could set up a nested, aligned set of these orbits covering a range of radii, as shown in Figure 6.12a. If we fill up these orbits with stars we create a bar-like pattern, which is stationary in the rotating frame and appears as a density wave rotating at the pattern speed Ω_p in the inertial frame. By rotating the axes of the ellipses we can create leading or trailing spiral density waves as in Figure 6.12b and c.

[3] This result is related to the shape of galaxy circular-speed curves in their inner parts. In most galaxies the circular speed rises linearly from the center with a steep slope. Thus, both Ω and κ are large at small radii, so for most values of m and n, $|\Omega - n\kappa/m|$ is much larger near the center than at large radii. However, in the central region where the

In a real galaxy $\Omega - \frac{1}{2}\kappa$ is not exactly constant. Hence, no matter what the value of Ω_p, most orbits are not exactly closed. The orientations of different orbits drift at slightly different speeds, so the pattern tends to twist or wind up. This is a modified version of the winding problem which we have already discussed—but now applied to density waves rather than material arms—and the rate of winding can be calculated in a similar way. Let $\phi_p(R,t)$ be the angle of the major axis of the pattern, as viewed in the frame rotating at the pattern speed. Let us suppose that the major axes are aligned at time $t = 0$; thus $\phi_p(R,0) = \phi_0$. The drift rate is $\partial\phi_p/\partial t = \Omega - \frac{1}{2}\kappa - \Omega_p$; thus

$$\phi_p(R,t) = \phi_0 + [\Omega(R) - \tfrac{1}{2}\kappa(R) - \Omega_p]t \tag{6.25}$$

(cf. eq. 6.8). Equation (6.6) now gives the pitch angle as

$$\cot\alpha = Rt\left|\frac{\mathrm{d}(\Omega - \frac{1}{2}\kappa)}{\mathrm{d}R}\right|. \tag{6.26}$$

In Model I for the Galactic potential of §2.7, the average of $\left|R\,\mathrm{d}(\Omega - \frac{1}{2}\kappa)/\mathrm{d}R\right|$ is about $7\,\mathrm{km\,s^{-1}\,kpc^{-1}}$ between 5 and 10 kpc, and after $t = 10\,\mathrm{Gyr}$ the pitch angle in this region is about $\alpha = 0.8°$. For comparison we computed after equation (6.9) that a material arm would have $\alpha < 0.2°$ in a galaxy with a similar circular-speed curve. Thus, the wave pattern winds up more slowly than the material arm by a factor of five or so. Although the pitch angle is still too small by a factor 10–20, we have come some way towards resolving the winding problem. We conclude that in galaxies with circular-speed curves similar to our own, $n = 1, m = 2$ *density waves can resist the winding process much better than material arms.* This result suggests a natural explanation for the prevalence of two-armed spirals, providing we can find a way to adjust the slow drift rates of all the orbits to a common standard.

Density waves of the type described above are called **kinematic density waves** because they involve only the kinematics of orbits in an axisymmetric potential. In a galaxy, the orbits will deviate from the paths we have assumed

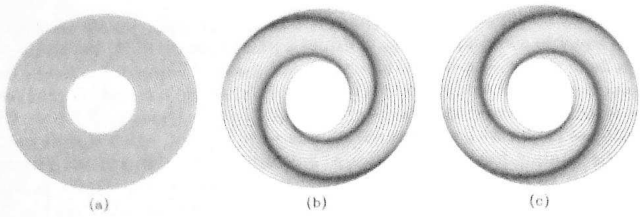


Figure 6.12 Arrangement of closed orbits in a galaxy with $\Omega - \frac{1}{2}\kappa$ independent of radius, to create bars and spiral patterns (after Kalnajs 1973b).

because the spiral pattern itself produces a non-axisymmetric component of the gravitational field. A major goal of spiral-structure theory is to determine whether the non-axisymmetric gravitational field due to the spiral itself can coordinate the drift rates of the orbits in such a way as to produce long-lived spiral patterns.

(b) Resonances Orbits, like springs and drums, have natural resonant frequencies. If the gravitational field generated by spiral structure perturbs an orbit near one of its resonant frequencies, then the response of the orbit is strong, even when the perturbing field is weak. To investigate the response of a stellar disk to non-axisymmetric forces, an essential first step is to locate the resonant orbits.

A gravitational potential that is stationary in a rotating frame can be written in the form $\Phi_1(R,\phi,t) = \Phi(R,\phi - \Omega_p t)$, where Ω_p is the pattern speed of the potential. Examples of systems that generate potentials of this form include the rotating bars seen at the centers of many disk galaxies, a satellite galaxy on a circular orbit in the disk plane, and any stationary spiral structure pattern (i.e., any structure with a well-defined pattern speed). More complicated potentials can be regarded as superpositions of potentials with different pattern speeds.

We now examine the effect of a weak potential of this form on a disk composed of stars on circular or near-circular orbits. Since the potential is periodic in $\phi - \Omega_p t$, it can be decomposed into a series of terms proportional to $\cos[m(\phi - \Omega_p t) + f_m(R)]$, where $m \geq 0$ is an integer. We studied orbits in potentials of this form in §3.3.3, and found that resonances occurred when the circular frequency Ω and the epicycle frequency κ in the unperturbed axisymmetric potential satisfied one of three conditions: $\Omega = \Omega_p$ (corotation resonance), $m(\Omega - \Omega_p) = \kappa$ (inner Lindblad resonance), or $m(\Omega - \Omega_p) = -\kappa$ (outer Lindblad resonance). These resonances occur at specific radii in a differentially rotating disk. The location and even the existence of these radii, called the corotation and Lindblad radii, depend on the circular-speed curve and the pattern speed. For example, inspection of Figure 6.11 show that the isochrone potential has zero or two inner Lindblad radii with $m =$

In reality, the orbits will not be exactly closed in the corotating frame.
Gravity can help: can drive an instability

The simplest type of instability is the Jeans instability for a spherical distribution. Jeans found that systems with a mass larger than the Jeans mass are unstable