1.12   Stability: Jeans mass and spiral structure

Until this point we have been concerned primarily with building equilibrium systems. We have not yet addressed the question of whether these equilibrium configurations are stable or unstable. The standard method for determining the stability of a system in equilibrium is to perturb it slightly. Either it will return to its previous configuration and oscillate around it (stable) or it will continue in the direction of the perturbation (unstable). The primary forces at work are gravity, which drives collapse, and pressure, which resists collapse. Since we have been treating the collection of stars as a collisionless fluid, it is not quite correct to think of a conventional fluid pressure, but for our purposes we take it that the stellar velocities produce a kind of effective pressure.

In the best of all possible worlds, one would begin with a time-independent distribution function and corresponding potential that solve the collisionless Boltzmann and Poisson equations. One then would consider small deviations of order $\epsilon \ll 1$ that can change as a function of time.

$$
\begin{aligned}
f(\vec{\mathbf{x}}, \vec{\mathbf{v}}, t) &= f_o(\vec{\mathbf{x}}, \vec{\mathbf{v}}) + \epsilon f_1(\vec{\mathbf{x}}, \vec{\mathbf{v}}, t) \\
\Phi(\vec{\mathbf{x}}, t) &= \Phi_o(\vec{\mathbf{x}}) + \epsilon \Phi_1(\vec{\mathbf{x}}, t).
\end{aligned}
\tag{1.169}
$$

Substituting (1.169) into the CBE and Poisson equation and equating terms of $\mathcal{O}(\epsilon)$, we get

$$
\frac{\partial f_1}{\partial t} + \vec{\mathbf{v}} \cdot \frac{\partial f_1}{\partial \vec{\mathbf{x}}} - \vec{\nabla}\Phi_0 \cdot \frac{\partial f_1}{\partial \vec{\mathbf{v}}} - \vec{\nabla}\Phi_1 \cdot \frac{\partial f_0}{\partial \vec{\mathbf{v}}} = 0
\tag{1.170}
$$

and

$$
\nabla^2 \Phi_1(\vec{\mathbf{x}}, t) = 4\pi G \int f_1(\vec{\mathbf{x}}, \vec{\mathbf{v}}, t) d^3\vec{\mathbf{v}} \quad .
\tag{1.171}
$$

But as with our study of equilbria, solution for, $f_1(\vec{\mathbf{x}}, \vec{\mathbf{v}}, t)$, a function of 7 variables, is impractical.

We resort, instead, to perturbing Jeans' equations, and thus get the perturbations to low order moments of the distribution function, which ought to work reasonably well for large scale perturbations. But even Jeans' equation, the first velocity moment of the CBE,

$$
\nu\frac{\partial \bar{v}_j}{\partial t} + \bar{v}_i \nu \frac{\partial \bar{v}_j}{\partial x_i} = -\nu\frac{\partial \Phi}{\partial x_j} - \frac{\partial}{\partial x_j}(\nu\sigma_{ij}^2).
$$

presents difficulties. Six distinct components of the velocity dispersion tensor must be separately perturbed. Such a treatment is "beyond the scope of the present treatment." We therefore retreat to a fluid (as opposed to collisionless) treatment of the perturbations to Jeans' equations, assuming an isotropic velocity dispersion tensor. We multiply the number density ν by a particle mass, giving a density ρ. Dividing through by density and associating the $(\rho\sigma^2)$ term with pressure P gives the (collisional) fluid analog, the Euler equation:

$$
\frac{\partial \vec{\mathbf{v}}}{\partial t} + (\vec{\mathbf{v}} \cdot \vec{\nabla})\vec{\mathbf{v}} = -\vec{\nabla}\Phi - \frac{1}{\rho}\vec{\nabla}P.
\tag{1.172}
$$

The zeroth moment of the CBE gives the mass continuity equation

$$
\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho\vec{\mathbf{v}}) = 0.
\tag{1.173}
$$

Just as with the distribution function above, we consider small deviations from the steady-state equilibrium solutions to the fluid equations (1.172) and (1.173). For small ϵ, we have the following expressions for the density, velocity, pressure, and potential of the fluid:

$$
\begin{aligned}
\rho(\vec{\mathbf{x}}, t) &= \rho_0(\vec{\mathbf{x}}) + \epsilon\rho_1(\vec{\mathbf{x}}, t) \\
\vec{\mathbf{v}}(\vec{\mathbf{x}}, t) &= \vec{\mathbf{v}}_0(\vec{\mathbf{x}}) + \epsilon\vec{\mathbf{v}}_1(\vec{\mathbf{x}}, t) \\
P(\vec{\mathbf{x}}, t) &= P_0(\vec{\mathbf{x}}) + \epsilon P_1(\vec{\mathbf{x}}, t) \\
\Phi(\vec{\mathbf{x}}, t) &= \Phi_0(\vec{\mathbf{x}}) + \epsilon\Phi_1(\vec{\mathbf{x}}, t)
\end{aligned}
\tag{1.174}
$$

The $\mathcal{O}(\epsilon)$ terms give

$$
\frac{\partial\vec{\mathbf{v}}_1}{\partial t} + (\vec{\mathbf{v}}_0 \cdot \vec{\nabla})\vec{\mathbf{v}}_1 + (\vec{\mathbf{v}}_1 \cdot \vec{\nabla})\vec{\mathbf{v}}_0 = -\vec{\nabla}\Phi_1 - \vec{\nabla}h_1,
\tag{1.175}
$$

$$
\frac{\partial\rho_1}{\partial t} + \vec{\nabla} \cdot (\rho_0\vec{\mathbf{v}}_1) + \vec{\nabla} \cdot (\rho_1\vec{\mathbf{v}}_0) = 0,
\tag{1.176}
$$

and

$$
\nabla^2\Phi_1 = 4\pi G\rho_1,
\tag{1.177}
$$

where the term h_1 in equation (1.175) is the first-order perturbation to the *enthalpy*, defined as

$$
h \equiv \int_0^\rho \frac{dP(\rho)}{\rho},
\tag{1.178}
$$

giving the perturbation expression for the enthalpy as

$$
h_1 = \left(\frac{dP}{d\rho}\right)_{\rho_0} \frac{\rho_1}{\rho_0} = v_s^2\frac{\rho_1}{\rho_0},
\tag{1.179}
$$

where v_s is the sound speed in the fluid.

To get Eqns. (1.175-1.177) into a form which can be solved analytically, Jeans made an additional simplifying assumption, that of a static, infinite background density ($\rho_0 = \mathrm{const}, \vec{\mathbf{v}}_0 = 0, \Phi_0 = 0$). This assumption is *neither* physically realistic nor self-consistent, but it allows us to write the fluid equations as two coupled first-order partial differential equations:

$$
\frac{\partial\rho_1}{\partial t} + \rho_0\vec{\nabla} \cdot \vec{\mathbf{v}}_1 = 0
\tag{1.180}
$$

and

$$
\frac{\partial\vec{\mathbf{v}}_1}{\partial t} = -\vec{\nabla}h_1 - \nabla\Phi_1.
\tag{1.181}
$$

These can be combined into a single second-order equation in the density perturbation ρ_1

$$
\frac{\partial^2\rho_1}{\partial t^2} - v_s^2\nabla^2\rho_1 - 4\pi G\rho_0\rho_1 = 0.
\tag{1.182}
$$

This is wave equation, so we try solutions of the form

$$
\rho_1(\vec{\mathbf{x}}, t) = C\exp[i(\vec{\mathbf{k}} \cdot \vec{\mathbf{x}} - \omega t)].
\tag{1.183}
$$

Substituting (1.183) into (1.182) gives a dispersion relation

$$\omega^2 = v_s^2 k^2 - 4\pi G \rho_0. \tag{1.184}$$

If $\omega^2 > 0$, the perturbations propagate like sound waves through the fluid, i.e. the perturbations are *stable*. If $\omega^2 < 0$, we get a real, positive exponent in the equation (1.183), giving unstable exponential growth. Setting $\omega^2 = 0$ gives the critical wavenumber, or equivalently the critical length scale ($\lambda = 2\pi/k$), for which initial density perturbations will grow exponentially. This wavelength is called the *Jeans length* and is given by

$$\lambda_J^2 = \frac{\pi v_s^2}{G \rho_0}. \tag{1.185}$$

The amount of matter within such a critical volume is easily calculated assuming a spherical density perturbation of radius λ_J and average density ρ_0. This *Jeans mass* is typical of the structure we can expect to be formed due to unstable gravitational collapse:

$$M_J = \frac{4\pi}{3} \rho_0 \lambda_J^3. \tag{1.186}$$

Our derivation thus far has been for *collisional* fluids. How are these results used to understand collisionless stellar systems? The speed of sound v_s^2 must be replaced by the velocity dispersion σ^2. This gives a dispersion relation that is slightly different from equation (1.184) but the Jeans mass remains unchanged.

For an isolated spherical system of point masses, the crossing time r/σ is approximately the same as the free-fall collapse time, $t \sim 1/\sqrt{G\rho}$. Nothing smaller than the whole system can be unstable to gravitational collapse.

Going back to the linear analysis, we can use the same methods with modifications appropriate to other geometries. Cylindrical coordinates would seem a good choice for fluid disks. We assume perturbations of the form $e^{im\phi}$ for integer values of m. The simplest case is that of $m = 0$, which gives the dispersion relation

$$\omega^2 = v_s^2 k^2 - 2\pi G \Sigma |k| + \kappa^2, \tag{1.187}$$

where Σ is the integrated surface density of the disk and κ is the same epicyclic frequency introduced in section (1.7). A dimensionless factor, *Toomre's Q*, defined by

$$Q \equiv \frac{v_s \kappa}{\pi G \Sigma} \tag{1.188}$$

allows the dispersion relation to be written as

$$\omega^2 = Q^2 k^2 \left(\frac{\pi G \Sigma}{\kappa} \right)^2 - 2\pi G \Sigma |k| + \kappa^2. \tag{1.189}$$

In this form the stability of the system depends solely on Q:

$$
\begin{aligned}
Q &= 1 && \text{gives perfect square, } \omega^2 \geq 0 \\
Q &> 1 && \text{stable for all } k \\
Q &< 1 && \text{unstable for some } k
\end{aligned}
\tag{1.190}
$$
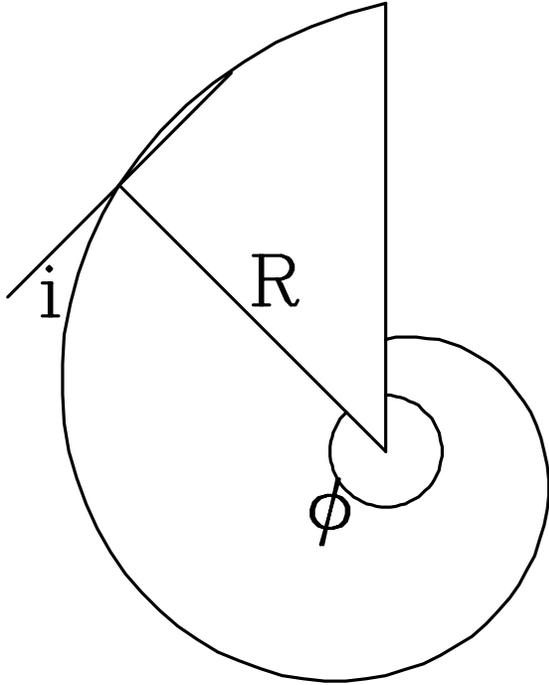
Figure 1.25: The crest of an $m = 1$ perturbation makes angle i with the tangential direction.

A similar treatment of a collisionless stellar disk yields a similar dispersion relation, with the qualification that the dimensionless constant π used in defining Toomre's Q, (eqn. 1.188), is replaced by the dimensionless constant 3.36, a change of less than 10%.

For values of $m = 1, 2, ...$, the picture gets somewhat more complicated. The first complication is that the notation changes – we switch from ω to $\Omega_p \equiv \omega/m$ where the subscript p denotes pattern speed. Full derivation of the dispersion relation is again "beyond the scope of the present treatment." There are nonetheless some important points that are relatively easy to follow.

First, there is a problem. Suppose that the stars in a disk are in circular orbits, with some orbits overpopulated so as to cause an apparent spiral. The position of the "crest" in polar coordinates is described by

$$\phi(R, t) = \phi_0 + \Omega(R)t \quad . \tag{1.191}$$

The the crest will in general make an angle i with the tangential direction. As illustrated in Figure 1.25 we find

$$\cot i = \left| R\frac{\partial \phi}{\partial R} \right| = R \left| \frac{d\Omega}{dR} \right| t \quad \text{whence} \tag{1.192}$$

$$\cot i = \left| \frac{d\ell n\Omega}{d\ell nR} \right| \Omega t \quad . \tag{1.193}$$

The relevant time t is of order the Hubble time, in which case the product Ωt is of order 100π. If $\Omega(R)$ is not constant (solid body rotation) any spiral that is more than a few galactic "years" old will have a very small inclination i – it will have wound up.

Lindblad came up with a scheme that partially mitigates the wind-up problem. Recall that with a few special exceptions, the orbits in a circularly symmetric potential do not close – the radial and aziumthal periods, T_r and T_a are different. There is a phase $\Delta\phi \equiv \omega_a T_r$ which is not in general a multiple of 2π, nor is it in general an integer fraction thereof. But in a frame of reference rotating
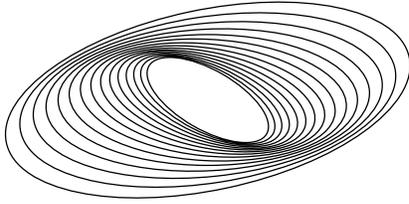
Figure 1.26: A series of nested ellipses produce a spiral arm. Such a "kinematic" spiral can only persist if there is a rotating frame in which all of the orbits are closed ellipses. This is *not* possible for a logarithmic potential.
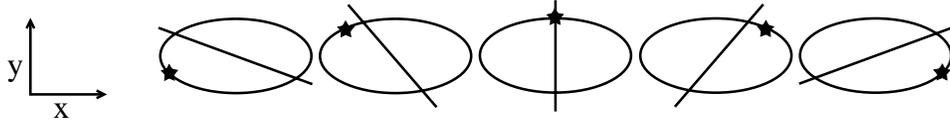
Figure 1.27: A star executes an elliptical epicycle as a spiral arm winds up. The coordinate system is in the frame of guiding center, which is moving toward the left. Notice that the star spends a large part of the epicycle close to the spiral arm.

with orbital frequency Ω_p, the phase of the orbit is given by

$$\Delta\phi_p = \Delta\phi - \omega_p T_r \quad . \tag{1.194}$$

If $\Delta\phi_p = 2\pi n/m$ the orbit closes afer m radial oscillations and

$$\Omega_p = \frac{\Delta\phi}{T_r} - \frac{\Delta\phi_p}{T_r} = \omega_a - \frac{n}{m}\omega_r \approx \Omega - \frac{n}{m}\kappa \quad , \tag{1.195}$$

where κ is the epicyclic frequency. Taking $n = 1$ and $m = 2$ gives closed ellipses. Let us suppose that there is a range of radii for which $\Omega - \kappa/2$ is constant. If one could then somehow arrange to populate at set of perturbed orbits with the correct phase, one might produce a pattern like that in Figure 1.26 that would persist for some time.

Lindblad's construction may point in the right direction, producing roughly the correct pattern speed, but it does not work well for our flat rotation curve Mestel disk model, for which $\kappa = \sqrt{2}\Omega$ giving $\Omega_p = 0.29\Omega$. Spirals wind up more slowly, but they still wind up.

But Lindblad's construction is kinematic, in that the spiral is an accident of initial conditions rather than a dynamical entity. Various investigators have attempted to constuct persistent spiral patterns for which self-gravity in some way overcomes the wind-up problem. Chief among these have been Lin and his collaborators, who have advanced the scenario of quasi-stationary *density waves*. For many galaxies the wind-up problem appears to preclude quasi-stationary density waves, but they may be possible in galaxies in which a bar forces them.

An alternative approach is to give up on the quest for immortality to embrace short-lived spirals. Toomre in particular has emphasized that close encounters between galaxies will cause tides that produce magnificent transient spiral patterns. The *swing amplification* of leading spiral arms, illustrated in Figure 1.27, plays a major role in this model. Measuring the inclination i of the spiral arm with respect to the tangential (horizontal) direction we have

$$\cot i = -R\frac{d\Omega}{dR}t \quad , \tag{1.196}$$

where $\frac{d\Omega}{dR} < 0$. Differentiation with respect to time and using the definitions of Oort's constants,

Equations 1.114 and 1.115, we that the time rate of change of the inclination angle is given by

$$\frac{di}{dt} = \frac{2A}{1 + A^2 t^2} \quad . \tag{1.197}$$

For a Mestel disk the epicyclic frequency is $\kappa = \sqrt{2}A$, roughly equally to the time rate of change of the inclination for most of an epicycle. Stars whose motions are in phase with the spiral arm stay in phase and self-gravitate, increasing the amplitude of the disturbance. Eventually the arm winds up.

Tides play a role, in particular, in the "grand design" spirals, starting with M51 for which the companion responsible for the tide is clearly visible, and M81, whose spiral arms appear to result from an encouneter with M82, which itself is undergoing a starburst.

Yet another scheme for producing spiral arms is *self-propagating star formation.* The idea is that star formation propagates along an interface much like a detonation front, which is sheared out by differential rotation. This produces patterns more like those seen in the "flocculent" spirals. The bottom line on spiral structure is that different mechanisms may be at work in barred, grand design and flocculent spirals.