QUEEN MARY
UNIVERSITY OF LONDON

Department of Physics

ASTM-052: Extragalactic Astrophysics
by
Peter Clegg

Note 3. Dynamics of Galaxies

Table of Contents

Chapter 3 : Dynamics of Galaxies 1

1. Introduction 1

2. Stellar Dynamics 1
  2.1 Independence of Stellar Motion 1
  2.2 Stellar Collisions 1
    2.2.1 Mean-Free Path and Collision time 1
    2.2.2 Direct Physical Collisions 1
    2.2.3 Gravitational “Collisions” 2
    2.2.4 Strong Gravitational Interactions 3

3. Gas Dynamics 4
  3.1 Simple Treatment 4
    3.1.1 Collective Motion 4
    3.1.2 The Free-Fall Time 4
    3.1.3 The Jeans Criterion 5
  3.2 Hydrodynamics of the Gas 5
    3.2.1 The Equation of Continuity 5
    3.2.2 The Euler Equation 6
    3.2.3 The Poisson Equation 7
    3.2.4 The Equation of State 7
    3.2.5 Perturbations and Linearisation of the Equations 7
    3.2.6 The Dispersion Relation 8
    3.2.7 Jeans Condensation Revisited 9
    3.2.8 The Time-Scale of Collapse 9
    3.2.9 Shock Waves 10

4. Dynamics of Spiral Discs 12
  4.1 Overall Description 12
  4.2 Co-ordinate System 12
  4.3 Motion Perpendicular to the Plane of the Disc 13
    4.3.1 Equations of Motion 13
  4.4 Motion of Stars in the Plane of the Disc 13
    4.4.1 Circular Motion 13
    4.4.2 Epicyclic Motion 14
    4.4.3 Resonant Orbits 15
  4.5 Spiral Structure 15
    4.5.1 Introduction 15
    4.5.2 The Winding Problem 15
    4.5.3 The Physical Mechanism Sustaining Spiral Arms 16
    4.5.4 Density-Wave Theory 17
    4.5.5 The Linearised 18
    4.5.6 The Dispersion Relation 18
    4.5.7 The Effects of and on the Stars 20
    4.5.8 The Speed of a Spiral Pattern 20
    4.5.9 Testing the Theory 20

Bibliography for Chapter 3 21

CHAPTER 3: DYNAMICS OF GALAXIES

1. Introduction

In this chapter, we shall explore the way that the material contents of galaxies move around and influence each other. In particular, I shall show that stars and gas behave very differently. I shall also discuss ideas about the origin of spiral structure in disc galaxies. For a more detailed discussion, see [1].

2. Stellar Dynamics

2.1 Independence of Stellar Motion

The two main constituents of galaxies – the stars and the gas – behave very differently. This is one of the reasons why spiral galaxies are so different from ellipticals. We shall see in a moment that most stars in galaxies ignore the presence of other *individual* stars, responding only to their overall gravitational field[1]. The gas, on the other hand, behaves as an entity: individual clouds of gas behave as continuous fluids and the clouds interact violently with each other. We shall return to the behaviour of the gas later when considering the formation of stars and the existence of spiral structure. First, let us tackle the motions of stars.

2.2 Stellar Collisions

2.2.1 MEAN-FREE PATH AND COLLISION TIME

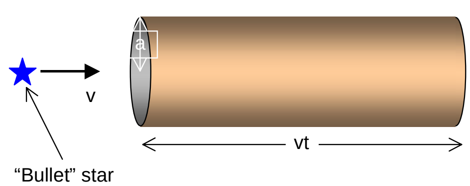

Figure 3-1. Volume swept out by circle of influence.

Suppose there are n stars per unit volume of space, with average velocity v. Consider a "bullet" star, travelling at v through these "target" stars. Let us assume that the bullet star interacts with the targets if their centres lie with a circle of influence (which I shall discuss in more detail below) of the centre of the bullet, which has radius a. The *cross-section* σ for collisions is therefore given by

$$\sigma = \pi a^2 \tag{2.1}$$

In time t, the circle of influence will sweep out a cylindrical volume V given by

$$V = \sigma \times vt\ , \tag{2.2}$$

as shown in Figure 3-1. This cylinder contains N target stars, where

$$N = V \times n = (\sigma vt) \times n \tag{2.3}$$

and there will therefore be N collisions in the time t. The average time τ_c between *individual* collisions is therefore given by

$$\tau_c \equiv \frac{t}{N} = \frac{t}{(\sigma vt)n} = \frac{1}{\sigma vn}\ . \tag{2.4}$$

In this time, the bullet star travels its *mean-free path* λ_c given by

$$\lambda_c := v\tau_c = \frac{1}{\sigma n}\ . \tag{2.5}$$

You should note that the mean-free path is independent of the velocity of the stars. Let us now investigate the frequency of different types of stellar collision.

2.2.2 DIRECT PHYSICAL COLLISIONS

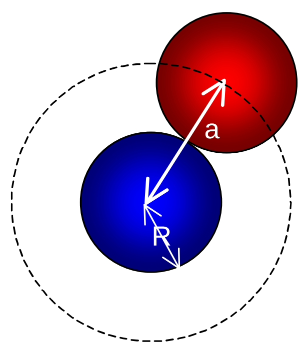

Figure 3-2. Direct physical collision between two stars

This is the most obvious and familiar type of collision: two stars physically collide with each other with severe consequences for their continued existence as individuals. If the average radius of the stars is R then, as shown in

Figure 3-2, the radius of the sphere of influence is *twice* the radius of the stars:

$$a = 2R\ . \tag{2.6}$$

The mean-free path λ_{coll} for direct collisions between stars is therefore given by

$$\lambda_{coll} = \frac{1}{4\pi R^2 n} \tag{2.7}$$

and the corresponding collision time τ_{coll} by

$$\tau_{coll} = \frac{1}{4\pi R^2 nv}\ . \tag{2.8}$$

[1]This applies only to so-called field stars and not to stars in clusters. Nor is it true of the centres of galaxies where the density of stars can be very high.

Putting in numerical values, we get

$$\lambda_{\text{coll}}(\text{kpc}) \sim \frac{1.6\times10^{11}}{(R/R_{\text{sun}})^2\, n\left(\text{pc}^{-3}\right)}; \qquad (2.9)$$

$$\tau_{\text{coll}}(\text{y}) \sim \frac{1.5\times10^{20}}{(R/R_{\text{sun}})^2\, n\left(\text{pc}^{-3}\right) v\left(\text{km s}^{-1}\right)} \qquad (2.10)$$

What values do λ_{coll} and τ_{coll} have in normal circumstances? Let us take the solar neighbourhood as an example. The typical radii and masses of stars are those of the sun, R_{sun} and M_{sun} respectively[2]. The density ρ of material – mostly stars – in the solar neighbourhood is of the order of 0.1 $\text{M}_{\text{sun}}\,\text{pc}^{-3}$ so that the number-density n of stars is given by

$$n \sim \frac{\rho}{M_{\text{sun}}} \sim 0.1\,\text{pc}^{-3}. \qquad (2.11)$$

Finally, typical velocities[3] of stars in the solar neighbourhood are about 20 km s^{-1}. Putting in these numbers together, we get

$$\lambda_{\text{coll}} \sim 10^9 \text{ Mpc} \quad \text{(solar neighbourhood)}; \qquad (2.12)$$

$$\tau_{\text{coll}} \sim 10^{20} \text{ y} \quad \text{(solar neighbourhood)}. \qquad (2.13)$$

This value of λ_{coll} is some million times the radius of the visible universe and τ_{coll} is some ten billion times the age of the universe. Since the solar neighbourhood is fairly typical, we can say confidently that solid-body collisions between stars almost never happen throughout most of the Galaxy! Even in the centre of a globular where densities are of order 10^4 pc^{-3} and velocities of order 15 km s^{-1},

$$\lambda_{\text{coll}} \sim 100\,\text{Mpc} \quad \text{(globular cluster)}; \qquad (2.14)$$

$$\tau_{\text{coll}} \sim 10^{15} \text{ y} \quad \text{(globular cluster)}. \qquad (2.15)$$

2.2.3 Gravitational "Collisions"

Although stars may physically collide only very infrequently, they do of course interact gravitationally and this interaction affects their motion. I am going to make a "back-of-the-envelope" study of this gravitational interaction; I shall use this sort approach a lot in the course because it gives a quick insight to problems whilst usually giving an answer within about an order of magnitude of more detailed treatment.

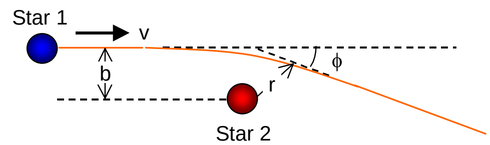

Figure 3-3. Gravitational deflection.

Figure 3-3 shows the gravitational interaction between two stars. Star 1 approaches star 2 at relative velocity v and with *impact parameter* b, defined as the closest approach the two stars would make to each other if they did *not* interact gravitationally. Suppose that star 1 is deflected through an angle ϕ (the angle between the asymptotes to its path). The precise calculation of ϕ is tedious but an estimate of its value can be obtained very simply. The gravitational force F between the two stars is given by

$$F(r) = \frac{Gm_1m_2}{r^2}, \qquad (2.16)$$

where m_1 and m_2 are the masses of stars 1 and 2 respectively, and r is their instantaneous separation. Rather than follow the interaction given by equation (2.16) over the infinite time for which it applies, I shall approximate it as follows. First, I shall use the *fixed* value F of the force between the two stars when they are separated by the impact parameter b. Secondly, I shall allow this force to operate only for a *finite* time. We then have

$$F \equiv F(b) = \frac{Gm_1m_2}{b^2}. \qquad (2.17)$$

How long should we allow this force to act for? The two stars are separated by a distance comparable to b for a time τ_b of order:

$$\frac{\text{Distance travelled by star 1 while force has value } F}{\text{Velocity of star 1}}$$

From Figure 3-3, we can see that

$$\tau_b \sim \frac{2b}{v}. \qquad (2.18)$$

From this, we can estimate the *impulse* I – equal to the force multiplied by the time during which it acts – given by star 2 to star 1:

$$I \sim F \times \tau_b \sim \frac{Gm_1m_2}{b^2} \times \frac{2b}{v} = \frac{2Gm_1m_2}{bv}. \qquad (2.19)$$

[2]Note that, when I use solar quantities such as R_{sun} as variables, they are italicised. When used as units, as in 0.1 M_{sun}, they are not.

[3]I am speaking here of the random, or *peculiar*, motions of stars over and above their systematic orbital velocity of about 220 km s^{-1} about the Galactic centre.

Deleted: R_o
Formatted
Deleted: M_o
Formatted
Formatted

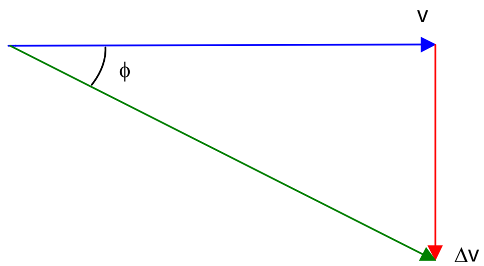

Figure 3-4. Vector diagram of velocities.

Suppose that, as a result of the interaction, star 1 suffers a change Δv in its velocity, perpendicular to its original velocity, as shown in the vector diagram of Figure 3-4. Since its change of momentum must equal the impulse given to it, we obtain

$$m_1 \Delta v \sim \frac{2Gm_1 m_2}{bv} \qquad (2.20)$$

or

$$\Delta v \sim \frac{2Gm_2}{bv} . \qquad (2.21)$$

Finally, using Figure 3-4, we can estimate the angle ϕ through which star 1 is deflected during the encounter. We find that

$$\phi \approx \frac{\Delta v}{v} \sim \frac{2Gm_2}{bv^2} . \qquad (2.22)$$

To get an expression for b, let us note that a volume V of space contains N stars, where

$$N = V \times n , \qquad (2.23)$$

n being the number-density of stars as before. An individual star therefor occupies, on average, a volume $V_{\rm star}$ given by

$$V_{\rm star} = \frac{V}{N} = \frac{1}{n} \qquad (2.24)$$

so that the typical value of b – the separation between stars at their closest approach – is given by

$$b \sim V_{\rm star}^{1/3} = n^{-1/3} . \qquad (2.25)$$

From relations (2.22) and (2.25), we get

$$\phi \sim \frac{2Gm_2 n^{1/2}}{v^2} . \qquad (2.26)$$

Again using solar-neighbourhood values to estimate a typical value for ϕ, we find

$$\phi \sim 10^{-5} \text{ radian} \approx 2 \text{ arcsec} . \qquad (2.27)$$

In other words, a typical gravitational interaction between stars produces only a tiny deviation of the stars from their original path[4].

2.2.4 STRONG GRAVITATIONAL INTERACTIONS[5]

From relationship (2.22), we can see that only close encounters will cause significant angular deflection of stars. I shall define a *strong gravitational interaction* as one that deflects a star through one radian[6]; as we shall see shortly, this is not quite a arbitrary as it may seem. From (2.22), we see that the impact parameter $b_{\rm strong}$ needed for this is given by

$$b_{\rm strong} \sim \frac{2Gm}{v^2} , \qquad (2.28)$$

where I have dropped the suffix 2 on m as being no longer necessary. Substituting $b_{\rm strong}$ for the radius of the sphere of influence in equation (2.1), we get from equation (2.5) for the mean-free path $\lambda_{\rm relaxl}$[7] between strong interactions,

$$\lambda_{\rm relax} = \frac{1}{\pi b_{\rm strong}^2 n} = \frac{v^4}{4\pi G^2 m^2 n} , \qquad (2.29)$$

with corresponding collision time $\tau_{\rm relax}$ given by

$$\tau_{\rm relax} = \frac{\lambda_{\rm grav}}{\rm v} = \frac{v^3}{4\pi G^2 m^2 n} . \qquad (2.30)$$

Putting in numerical values, we get

$$\lambda_{\rm relax}(\text{kpc}) = 4.3 \times \frac{v^4(\text{km s}^{-1})}{\left(\frac{m}{M_{\rm sun}}\right)^2 n(pc^{-3})} ; \qquad (2.31)$$

$$\tau_{\rm relax}(\text{y}) = 4.2 \times 10^9 \, \frac{v^3(\text{km s}^{-1})}{\left(\frac{m}{M_{\rm sun}}\right)^2 n(pc^{-3})} . \qquad (2.32)$$

Now one radian, which is about 60° is a very appreciable deviation of a star from its original path. Therefore, $\tau_{\rm relax}$ is a measure of the time needed by a group of stars to influence each other significantly and hence to come into dynamical equilibrium with each other. It is known as the *gravitational relaxation time*.

[4]This conclusion does not hold for very dense regions, such as the centres of globular clusters or the centres of galaxies, where the densities and their random velocities are high.

[5]I deliberately adopt a very simple approach here. A more detailed treatment gives somewhat lower values of relaxation times but the overall conclusions still hold.

[6] An alternative, and more realistic, approach is to calculate the number of collisions necessary to produce a total deviation of one radian. This is a random walk problem and gives a similar result.

[7] The reason for the choice of the suffix "relax" will be clear shortly.

Although the choice of one radian as the criterion for a significant deflection may appear rather arbitrary, there is another reason for choosing this value. It is easy to show that the impact parameter b_{strong} is that which makes the gravitational potential energy of a star – in the field of another – equal to its kinetic energy. Under such condition, we should expect the gravitational interaction to have a significant effect on the motion of the star.

Using solar-neighbourhood value once again, we find that

$$\lambda_{\text{relax}} \sim 7{,}000\ \text{Mpc} \quad \text{(solar neighbourhood);} \tag{2.33}$$

$$\tau_{\text{relax}} \sim 3\times10^{14}\ \text{y} \quad \text{(solar neighbourhood).} \tag{2.34}$$

so that gravitational collisions again have very little effect. For the centres of globular clusters. however,

$$\lambda_{\text{relax}} \sim 20\ \text{kpc} \quad \text{(globular cluster);} \tag{2.35}$$

$$\tau_{\text{relax}} \sim 1.5\times10^{9}\ \text{y} \quad \text{(globular cluster)} \tag{2.36}$$

so that these have had time to become gravitationally relaxed, as we might expect from their smooth, spherical appearance. Note, though, that the mean-free path is much bigger than the size of the cluster so that stars have to cross the cluster many times for the relaxation to take place.

We can conclude that, except in particularly dense regions such as the centres of globular clusters and the nuclear regions of spiral galaxies, stars move more or less independently of other individual stars. This fact is important in tracing the evolution of galaxies, as I shall show later.

3. Gas Dynamics

3.1 Simple Treatment

3.1.1 Collective Motion

As I said above, gas exerts pressure, which can be transmitted – at the speed of sound – over large distances within a cloud of gas. It also exhibits viscosity. Both these properties complicate the study of the (hydro-)dynamics of gas in galaxies and I shall not attempt more than an outline of most of the effects. I shall first give a very simple treatment of the collapse of a gas cloud before going into more detail in section 3.2.

3.1.2 The Free-Fall Time

Consider a system, of mass M, in which the *only* forces acting are gravitational: there are no pressure forces in this idealised system. For simplicity I shall take the system to be spherically symmetric and initially at rest, as shown in Figure 3-5. The equation of motion of a particle of mass m, situated at distance r from the centre of the system, is

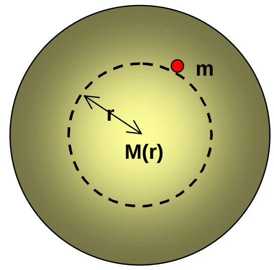

Figure 3-5. Free-fall of gas cloud.

$$m\frac{d^2r}{dt^2} = -\frac{GM(r)}{r^2}m\,, \tag{3.1}$$

where, as usual, $M(r)$ is the mass interior to radius r. It is easy to show that, as they accelerate inwards, no particle overtakes a particle initially closer to the centre of the system. For any chosen particle initially at r_o, therefore, $M(r)$ has the constant value $M(r_o)$. Let τ_{ff} be the time for the chosen particle to reach the centre of the system, remembering that I am not allowing any pressure gradients to build up to stop the motion. τ_{ff} is called the *free-fall time*. A back-of-the-envelope approach allows us to estimate τ_{ff} as follows. For the particle to reach the centre, it has to travel a distance r_o. Since the time it takes to do this is τ_{ff}, a crude estimate of its average velocity $\langle v\rangle$ is given by

$$\langle v\rangle \sim -\frac{r_o}{\tau_{\text{ff}}}\,. \tag{3.2}$$

An equally crude estimate of its average acceleration $\langle a\rangle$ is given by

$$\langle a\rangle \equiv \left\langle\frac{d^2r}{dt^2}\right\rangle \sim \frac{\langle v\rangle}{\tau_{\text{ff}}} \sim -\frac{r_o}{\tau_{\text{ff}}^2}\,. \tag{3.3}$$

The negative signs occur in relationships (3.2) and (3.3) because the motion is in the direction of decreasing r. Finally, a crude average value $\langle F\rangle$ of the gravitational force on the particle is given by

$$\langle F\rangle \equiv \left\langle\frac{GM(r_o)}{r^2}\right\rangle \sim \frac{GM(r_o)}{(r_o/2)^2} = 4\frac{GM(r_o)m}{r_o^2}\,. \tag{3.4}$$

From relationships (3.1), (3.3) and (3.4), we get the approximate relationship

$$\frac{r_o}{\tau_{\text{ff}}^2} \sim 4\frac{GM(r_o)}{r_o^2} \tag{3.5}$$

or

$$\tau_{\mathrm{ff}} \sim \left[\frac{4\,GM(r_{\mathrm{o}})}{r_{\mathrm{o}}^{3}}\right]^{-1/2} . \qquad (3.6)$$

We can re-write equation (3.6) as

$$\tau_{\mathrm{ff}} \sim \left[\frac{(16\pi/3)\,GM(r_{\mathrm{o}})}{(4\pi/3)r_{\mathrm{o}}^{3}}\right]^{-1/2} . \qquad (3.7)$$

But $(4\pi/3)r_{\mathrm{o}}^3$ is the volume of the sphere of radius r_{o} so that

$$\frac{M(r_{\mathrm{o}})}{(4\pi/3)r_{\mathrm{o}}^{3}} \sim \rho . \qquad (3.8)$$

where ρ is a measure of the average density of the system. Hence

$$\tau_{\mathrm{ff}} \sim 2\sqrt{\frac{\pi}{3}}(G\rho)^{-1/2} \sim (G\rho)^{-1/2} . \qquad (3.9)$$

since we are only considering orders of magnitude, and are not concerned with small numerical factors.

Notice that the free-fall time is independent of both the size and the total mass of the system. Although derived for the special case of a spherical cloud, this expression gives the order of magnitude of the time taken by *any* self-gravitating system to collapse if no other forces act upon it.

3.1.3 The Jeans Criterion

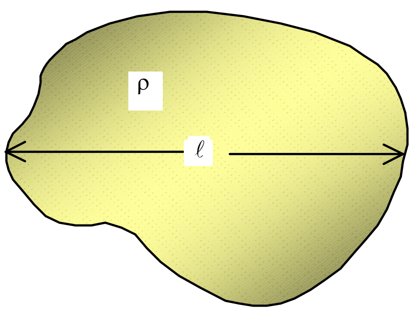

Figure 3-6. Jeans Collapse

In a real cloud of gas, we shall to have pressure gradients[8] that may help to support the cloud against gravitational collapse. Consider the cloud shown in Figure 3-6. If no pressure gradient existed within it, the cloud would start to collapse on a time-scale given by relationship (3.9). If the cloud wants to stop itself collapsing, it must set up a pressure gradient before it is too late. Now changes in pressure travel at the speed of sound so, to distribute this pressure-gradient correctly, the cloud will need at least as long as the time τ_{sound} that it takes for a sound-wave to cross the cloud. This is given by

$$\tau_{\mathrm{sound}} \sim \frac{\ell}{v_{\mathrm{sound}}} , \qquad (3.10)$$

ℓ being the characteristic size of the cloud and v_{sound} being the speed of sound in the gas[9]. The cloud will collapse if it free-falls faster than it can set up the pressure gradient, that is if

$$\tau_{\mathrm{ff}} < \tau_{\mathrm{sound}} . \qquad (3.11)$$

Using relationships (3.9) and (3.10), we see that the cloud will collapse if

$$\ell \gtrsim \ell_{\mathrm{J}} := \left(\frac{v_{\mathrm{sound}}^{2}}{\mathrm{G}\rho}\right)^{1/2} \qquad (3.12)$$

The quantity ℓ_{J} defined by relation (3.12) is called the *Jeans length* after Sir James Jeans, who first advanced this idea.

Corresponding to the Jeans length is the Jeans mass M_{J} which is the mass of the material contained within a region the size of the Jeans length. Clearly this mass is of the order of the density ρ of the material multiplied by ℓ_{J}^{3}:

$$M_{\mathrm{J}} := \ell_{\mathrm{J}}^{3}\rho = \left(\frac{v_{\mathrm{sound}}^{2}}{\mathrm{G}}\right)^{3/2} \rho^{-1/2} \qquad (3.13)$$

and the cloud will collapse if

$$M \gtrsim M_{\mathrm{J}} . \qquad (3.14)$$

Note that, the *denser* the cloud, the *smaller* the mass required to cause collapse.

3.2 Hydrodynamics of the Gas

3.2.1 The Equation of Continuity

I shall not discuss the theory of hydrodynamics in any detail but merely want to give the fundamental principles involved. There are four basic equations which convey these principles. The first, the equation of continuity, expresses the commonplace that matter cannot be created or destroyed. Consider a fluid in which a flow is taking place and select the volume V shown in Figure 3-7. Any increase in the mass M of material contained within the volume V must be provided by material flowing across the boundary S of the volume. Hence, we can write

[8] It is important to remember that it is the pressure *gradient* that transmits a net force between neighbouring parts of a fluid. A uniform pressure exerts the same force on both "sides" of any element of the fluid and therefore exerts no net force.

[9] I shall consider more precisely what is meant by the "speed of sound" later.

$$\begin{aligned} &\textit{Rate of increase of mass within } V \\ &= \textit{Rate of flow of mass across } S \end{aligned} \qquad (3.15)$$

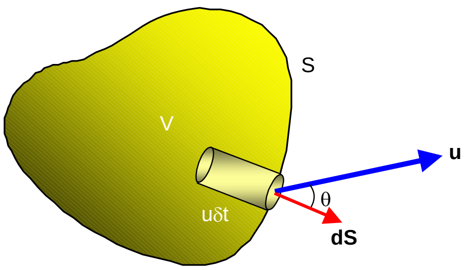

Figure 3-7, Flow of liquid.

It is easy to write down an expression for the left-hand side of this equation. We have

$$\frac{dM}{dt} = \frac{d}{dt}\int_V \rho \, dV = \int_V \frac{\partial \rho}{\partial t} dV \,, \qquad (3.16)$$

where we can change the order of differentiation and integration because the volume V is fixed.

Obtaining an expression for the right hand side of the equation (3.15) is a little trickier. Consider the infinitesimal element ***dS*** of the bounding surface, shown in the diagram, and suppose the velocity of the fluid at this point is ***u***, in the direction shown. In time δt, the volume of fluid flowing *out of* the volume is simply the volume δV of the slanted cylinder shown. It is clear from the figure that

$$\delta V = \mathbf{u}\,\delta t \times \mathbf{dS} \cos\theta = \delta t(\mathbf{u.dS}) . \qquad (3.17)$$

If the density of the fluid is ρ at this point, then the *total* mass δM of material flowing *into* the volume in time δt is given by

$$\delta M = -\rho\,\delta V = -\delta t \int_S \rho\,\mathbf{u.dS} \qquad (3.18)$$

so that

$$\frac{dM}{dt} = -\int_S \rho\,\mathbf{u.dS} \qquad (3.19)$$

Equating the expressions (3.16) and (3.19) for *dM/dt*, we get

$$\int_V \frac{\partial \rho}{\partial t} dV = -\int_S \rho\,\mathbf{u.dS} = -\int_V \nabla.(\rho\,\mathbf{u}) dV \,, \qquad (3.20)$$

where I have used Gauss' theorem to get the second equality. We can re-write (3.20) as

$$\int_V \left[\frac{\partial \rho}{\partial t} - \nabla.(\rho\,\mathbf{u})\right] dV = 0 \,, \qquad (3.21)$$

or, since the we chose the volume V quite arbitratrarily,

$$\frac{\partial \rho}{\partial t} - \nabla.(\rho\,\mathbf{u}) = 0 \,. \qquad (3.22)$$

which is the equation of continuity[10].

3.2.2 The Euler Equation

The Euler equation is common-sense application of Newton's second law although I first need to discuss the concept of the hydrodynamic derivative. Consider any function f of the four variables (x,y,z,t): f might be density or pressure, for example. In general, taking the differential of f, we have

$$df = \frac{\partial f}{\partial t} dt + \sum_{i=1}^{3} \frac{\partial f}{\partial x_i} dx_i \qquad (3.23)$$

where $x_i = x$, etc. Now fix my attention on the rate of change of the function f which we would measure, in some given element of the fluid, *while moving with that element*; denote this rate of change of f by Df/Dt. Dividing equation 2.14 by δt throughout, we obtain

$$\begin{aligned} \frac{\mathrm{D}f}{\mathrm{D}t} &= \frac{\partial f}{\partial t} + \sum_{i=1}^{3} \frac{\partial f}{\partial x_i} \frac{dx_i}{dt} \\ &= \frac{\partial f}{\partial t} + \sum_{i=1}^{3} \frac{\partial f}{\partial x_i} u_i \,, \end{aligned} \qquad (3.24)$$

where u_i is the ith component of the velocity of the fluid. The second equality in equation (3.24) follows from the fact that we are moving with the fluid so that the rate of change dx_i/dt of our x_i-co-ordinate is u_i. The quantity denoted by D/Dt is called the *hydrodynamic derivative*.

Equation (3.24) can be written in compact form as

$$\frac{\mathrm{D}f}{\mathrm{D}t} = \frac{\partial f}{\partial t} + (\mathbf{u}.\nabla) f \,. \qquad (3.25)$$

We now use equation (3.25) to find the acceleration D**u**/Dt of the fluid. Newton's second law says that force is equal to mass times acceleration. Consider a small mass m of the fluid. Then

$$m\frac{\mathrm{D}\mathbf{u}}{\mathrm{d}t} = \mathbf{f} \,, \qquad (3.26)$$

where **f** is the force acting on the mass m. If we define **F** to be the force *per unit mass* of fluid, so that

[10] Equation 2.13 is, of course, similar to the equation of conservation of charge in electromagnetism, the "mass current-density" $\rho\,\mathbf{u}$ replacing the electrical current-density ***j***.

$$\mathbf{F} = \frac{\mathbf{f}}{m}, \tag{3.27}$$

then equation (3.26) becomes

$$\frac{\mathrm{D}\mathbf{u}}{\mathrm{d}t} = \mathbf{F}. \tag{3.28}$$

There are two[11] forces acting, the gravitational force and the pressure-gradient force. By definition of the gravitational potential Φ, the gravitational force $\mathbf{f}_G$ acting on mass m of the fluid is given by

$$\mathbf{f}_G = -m\nabla\Phi \tag{3.29}$$

so that the gravitational force $\mathbf{F}_G$ per unit mass is given by

$$\mathbf{F}_G = \frac{1}{m}\mathbf{f}_G = -\nabla\Phi. \tag{3.30}$$

By considering the pressure forces acting on opposite sides of a small cube of fluid, it is easy to show that the force $\mathbf{F}_p$ per unit mass of the fluid, arising from a pressure gradient ∇p, is given by

$$\mathbf{F}_p = -\frac{1}{\rho}\nabla p. \tag{3.31}$$

Finally, from equations (3.28), (3.30) and (3.31), we get the *Euler equation* for the velocity $\mathbf{u}$ of the fluid:

$$\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} = -\nabla\Phi - \frac{1}{\rho}\nabla p. \tag{3.32}$$

3.2.3 The Poisson Equation

The third equation we need is the *Poisson equation* for the gravitational potential Φ:

$$\nabla^2\Phi = 4\pi G\rho, \tag{3.33}$$

which is easily obtained by analogy with the electrostatic potential and charge-density in electromagnetism..

3.2.4 The Equation of State

So far, we have three equations relating the four variables $\mathbf{u}$, ρ, p and Φ. We need one more equation in order to get a solution for any one of the variables; this is the equation of state relating p and ρ:

$$p \equiv p(\rho). \tag{3.34}$$

In fact, if we assume that all motion is adiabatic, it turns out that all we need is the *differential* relation[12]

$$\left(\frac{\partial p}{\partial \rho}\right)_S = u^2. \tag{3.35}$$

Then,

$$\nabla p = \frac{dp}{d\rho}\nabla\rho = u^2\nabla\rho. \tag{3.36}$$

We shall see later that u is related to the velocity of sound in the gas. Using (3.36), we can eliminate p from the Euler equation and find

$$\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} = -\nabla\Phi - \frac{u^2}{\rho}\nabla\rho \tag{3.37}$$

We are now poised to explore the behaviour of the gas in galaxies!

3.2.5 Perturbations and Linearisation of the Equations

We now have three (differential) equations for the three unknowns ρ, $\mathbf{u}$ and Φ. Given appropriate boundary conditions we could, in principle, solve them. In practice the solution is difficult because the equations are non-linear. If we are content to explore small *perturbations* from the equilibrium state, it is possible to obtain approximate equations which are linear and which can therefore be solved more easily. Moreover, if we have any two solutions of a linear equation, the sum of these solutions is also a solution. This allows us to use Fourier techniques in seeking solutions. Using this perturbation approach, the Jeans criterion of section 3.1.3 can be derived in a more rigorous fashion. The main reason for introducing the equations, though, is that they form the foundation for the theory of spiral structure, which I shall describe later.

Consider the density ρ of the fluid as an example. Suppose it can be represented by a known part ρ_0 together with a perturbation ρ_1, which is a small fraction of ρ_0:

$$\rho = \rho_0 + \rho_1. \tag{3.38}$$

Doing the same for the pressure p, the velocity $\mathbf{u}$ and the gravitational potential Φ, we arrive at the equations

$$\begin{aligned} p &= p_0 + p_1; \\ \mathbf{u} &= \mathbf{u}_0 + \mathbf{u}_1; \\ \Phi &= \Phi_0 + \Phi_1. \end{aligned} \tag{3.39}$$

We assume that the quantities q_0 themselves solutions of the basic equations (3.22), (3.32) and (3.33). If we

[11] For simplicity, I shall ignore electromagnetic forces although, in practice, these undoubtedly play an important rôle.

[12] Do not confuse this scalar quantitiy u with the vector velocity $\mathbf{u}$ of the fluid.

substitute for ρ and $\mathbf{u}$, from equations (3.38) and (3.39), in the equation of continuity (3.22), we get

$$\frac{\partial}{\partial t}(\rho_0 + \rho_1) + \nabla.[(\rho_0 + \rho_1)(\mathbf{u}_0 + \mathbf{u}_1)] = 0 . \quad (3.40)$$

Because the quantities q_1 are small compared with the q_0, we can neglect second order terms in the q_1s and so reduce equation (3.40) to the form

$$\left[\frac{\partial \rho_0}{\partial t} + \nabla.(\rho_0 \mathbf{u}_0)\right] + \left[\frac{\partial \rho_1}{\partial t} + \nabla.(\rho_0 \mathbf{u}_1 + \rho_1 \mathbf{u}_0)\right] = 0 . \quad (3.41)$$

The first term in square brackets in equation (3.41) vanishes identically because ρ_0 and $\mathbf{u}_0$ are chosen to satisfy the equation of continuity (3.22). We are therefore left with

$$\frac{\partial \rho_1}{\partial t} + \nabla.(\rho_0 \mathbf{u}_1 + \rho_1 \mathbf{u}_0) = 0 , \quad (3.42)$$

which is *linear* in the perturbations ρ_1 and $\mathbf{u}_1$.

We can linearise the Euler equation (3.32) in a similar way to get

$$\frac{\partial \mathbf{u}_1}{\partial t} = -\nabla \Phi_1 - \frac{u^2}{\rho_0} \nabla \rho_1 . \quad (3.43)$$

The Poisson equation (3.33), which is already linear, reduces to

$$\nabla^2 \Phi_1 = 4\pi G \rho_1 , \quad (3.44)$$

3.2.6 The Dispersion Relation

I now want to show that the above equations can have wave-like solutions; I shall show this only for ρ_1 but the result also holds for p_1, $\mathbf{u}_1$ and Φ_1. Suppose that the background quantities ρ_0, p_0, and Φ_0 are homogenous and that the background fluid is at rest. Then

$$\rho_0 = \text{constant};$$
$$p_0 = \text{constant};$$
$$\mathbf{u}_0 = 0;$$
$$\Phi_0 = \text{constant}.$$

Under these conditions, equation (3.42) reduces to

$$\frac{\partial \rho_1}{\partial t} + \rho_0 \nabla.\mathbf{u}_1 = 0 . \quad (3.45)$$

Taking the derivative of this equation with respect to time, we get

$$\frac{\partial^2 \rho_1}{\partial t^2} + \rho_0 \frac{\partial}{\partial t}(\nabla.\mathbf{u}_1) = 0 . \quad (3.46)$$

Take the divergence of the Euler equation (3.43) to get

$$\nabla.\frac{\partial \mathbf{u}_1}{\partial t} = -\nabla^2 \Phi_1 - \frac{u^2}{\rho_0} \nabla^2 \rho_1 , \quad (3.47)$$

or

$$\frac{\partial}{\partial t} \nabla.\mathbf{u}_1 = -4\pi G \rho_1 - \frac{u^2}{\rho_0} \nabla^2 \rho_1 , \quad (3.48)$$

where I have interchanged the order of differentiation in the first equality and used the Poisson equation (3.44). If we now multiply equation (3.48) by ρ_0 and subtract it from equation (3.46), we get

$$\frac{\partial^2 \rho_1}{\partial t^2} = u^2 \frac{\partial^2 \rho_1}{\partial x^2} + (4\pi G \rho_o)\rho_1 . \quad (3.49)$$

If the second term on the right hand side of this equation were absent, we should have an ordinary wave equation for the perturbation ρ_1 showing it to propagate with the speed of sound u.

Any wave can be decomposed into its Fourier components, so let us try a plane-wave solution of the form

$$\rho_1(t, x) = \rho_{1o} \exp\left[i2\pi\left(\nu t - \frac{x}{\lambda}\right)\right] = \rho_{1o} e^{i(\omega t - kx)} \quad (3.50)$$

where the angular frequency ω and the wave-number k are given by

$$\omega = 2\pi\nu ; \quad (3.51)$$

$$k = \frac{2\pi}{\lambda} . \quad (3.52)$$

Substituting this solution into equation (3.49), we get the *dispersion relation*

$$\omega^2 = u^2 k^2 - 4\pi G \rho_o . \quad (3.53)$$

If the second term on the right hand side of equation (3.53) were absent – if we could "switch off" gravity – we should have the usual terrestrial dispersion relation between the frequency and wavelength of a wave:

$$\omega^2 = u^2 k^2 . \quad (3.54)$$

The term $4\pi G\rho_0$ modifies the equation in a profound way, however. Neither the phase nor the group velocity is any longer independent of the frequency of the wave: we find that

$$u_{\text{phase}} \equiv \frac{\omega}{k} = u\sqrt{1 - \frac{4\pi G \rho_o}{u^2 k^2}} = u\sqrt{1 - \frac{G\rho_o}{\pi u^2}\lambda^2} \quad (3.55)$$

whereas

$$u_{\text{group}} \equiv \frac{d\omega}{dk} = \frac{u}{\sqrt{1 - \dfrac{4\pi G\rho_o}{u^2 k^2}}} = \frac{u}{\sqrt{1 - \dfrac{G\rho_o}{\pi u^2}\lambda^2}} . \quad (3.56)$$

Information therefore travels faster at longer wavelengths.

The real importance of the gravitational term, though, is that, if it is sufficiently large, it can make ω^2 *negative*; that is, it can make the frequency imaginary! For

$$4\pi G\rho_o > u^2 k^2 , \quad (3.57)$$

we have

$$\omega^2 < 0 . \quad (3.58)$$

Put

$$\omega = \pm i\alpha , \quad (3.59)$$

where α is real. The solution of the wave equation then becomes

$$\rho_1 = \rho_{1o} e^{\mp\alpha t} e^{-ikx} . \quad (3.60)$$

This is no longer a travelling wave but a *standing-wave* disturbance which is unstable: its amplitude either grows or decays exponentially with time[13], depending on the sign of α.

3.2.7 JEANS CONDENSATION REVISITED

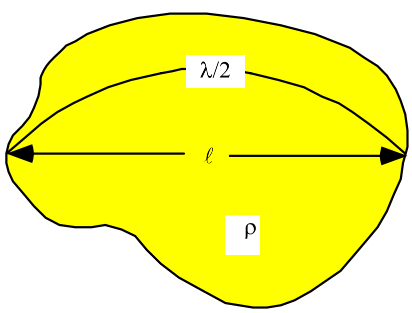

Figure 3-8. Maximum wavelength in a cloud.

Equation (3.57) shows that ω is imaginary, and therefore that exponential growth of the perturbation can occur, if[14]

$$k < k_J := \left(\frac{4\pi G\rho}{u^2}\right)^{1/2} \quad (3.61)$$

or

$$\lambda > \lambda_J := \left(\frac{\pi u^2}{G\rho}\right)^{1/2} . \quad (3.62)$$

Consider the possible modes of oscillation of the cloud of gas, with characteristic dimension ℓ, shown in Figure 3-8. Since the empty space outside the cloud cannot support any density fluctuations[15], the edges of the cloud must be nodes of the wave. The lowest-frequency mode is therefore such that one half-wavelength just spans the cloud so that the wavelength λ of this lowest mode of oscillation is given by

$$\frac{\lambda}{2} \sim \ell . \quad (3.63)$$

We see from equation (3.62), therefore, that the density of the cloud can grow exponentially, that is the cloud can collapse exponentially, if

$$\ell > \ell_J = \frac{\lambda_J}{2} = \left(\frac{\pi u^2}{4G\rho}\right)^{1/2} . \quad (3.64)$$

The mass M contained within a cloud of diameter ℓ is given by

$$M \sim \frac{4\pi}{3}\left(\frac{\ell}{2}\right)^3 \rho \quad (3.65)$$

so the cloud can collapse if its mass satisfies the relation

$$M > M_J := \frac{\pi^{5/2}}{48}\left(\frac{u^2}{G}\right)\rho^{-1/2} . \quad (3.66)$$

The Jeans length ℓ_J defined by equation (3.64) differs form that obtained by the less sophisticated method of section 3.1.3 [equation (3.12)] by only a factor of $\pi^{1/2}/2$=0.89 whilst the Jeans mass M_J differs from the previous estimate [equation (3.13)] by $\pi^{5/2}/48$=0.36.

3.2.8 THE TIME-SCALE OF COLLAPSE

Equation (3.9) shows that a cloud whose mass exceeds the Jeans mass collapses – at least initially – as exp$(-\alpha t)$. It therefore increases its density by a factor e in a time τ_{collapse} given by

$$\tau_{\text{collapse}} = \frac{1}{\alpha} . \quad (3.67)$$

[13]The derivation of the wave equation required that the disturbance was small. You will realise that, if the exponential growth occurs, the perturbation will not *remain* small and the above treatment will cease be valid after a time of the order of 1/α.

[14] It is no longer necessary to distinguish between the background density ρ_o and the perturbation ρ_1 so we shall now drop the suffix zero.

[15] There is, of course, no such thing as completely empty space but the approximation is useful here.

Now the dispersion relation (3.53), with ω replaced by $i\alpha$, shows that

$$\begin{aligned}\alpha^2 &= -u^2k^2 + 4\pi G\rho_o \\ &= -u^2\left(\frac{4\pi}{\lambda}\right)^2 + 4\pi G\rho_o \\ &\approx -u^2\left(\frac{2\pi}{\ell}\right)^2 + 4\pi G\rho_o .\end{aligned} \tag{3.68}$$

where I have used relations (3.63). As the cloud collapses, the first term on the right hand side of equation (3.68) varies inversely as the square of the size of the cloud, whereas the density in the second term varies inversely as the cube of the size of the cloud. As the cloud shrinks, therefore, the second term eventually dominates the right hand side of the equation and we have

$$\alpha \approx \sqrt{4\pi G\rho} . \tag{3.69}$$

From relation (3.67) we have, therefore,

$$\tau_{\text{collapse}} \sim (4\pi G\rho)^{-1/2} = \frac{1}{\sqrt{4\pi}}\tau_{\text{ff}} , \tag{3.70}$$

the last relation coming from equation (3.15). In other words, the cloud collapses on a time-scale comparable with the free-fall time. We can see what this means physically by remembering the discussion of section 3.1.3: if the pressure gradient is insufficient to prevent it, collapse takes place and the pressure gradient becomes increasingly ineffectual against the effects of gravity.

3.2.9 Shock Waves[16]

Let us now temporarily ignore gravitational effects. The derivation of equation (3.49) for the propagation of density perturbations in a gas depended on these perturbations being small. I now want to consider what happens if we consider large disturbances. We need first to consider how the speed of sound, given by equation (3.35), depends upon the density ρ of the gas. For a perfect gas – which is a good approximation to interstellar gas – undergoing adiabatic compression or rarefaction, the relationship between the pressure p and the density ρ is

$$p = p_o\left(\frac{\rho}{\rho_o}\right)^\gamma , \tag{3.71}$$

where p_o and ρ_o are constants and γ is the ratio of the specific heat at constant pressure to that at constant volume. Hence

$$u^2 = \frac{d}{d\rho}\left[p_o\left(\frac{\rho}{\rho_o}\right)^\gamma\right] = \gamma\frac{p_o}{\rho_o}\rho^{\gamma-1} \propto \rho^{\gamma-1} \tag{3.72}$$

Since γ is greater than unity, the adiabatic sound speed *increases* as the density increases

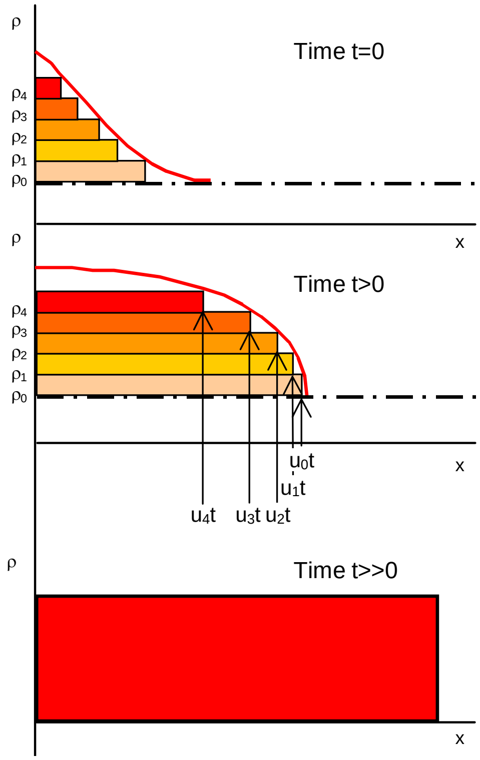

Figure 3-9. Development of a shock front.

Consider the density perturbation shown by the full line in the top diagram of Figure 3-9 propagating into a gas of undisturbed density ρ_0. Let us approximate this disturbance by the sum of the *small* rectangular disturbances shown as dashed curves in the figure. After time t, the lowest rectangular disturbance will have travelled a distance $u_0 t$, where

$$u_0 = u(\rho_0) \tag{3.73}$$

is the velocity of sound in the undisturbed gas. The second small disturbance, however, will be travelling in the gas which has *already been disturbed* by the first disturbance and whose density has risen to ρ_1. The second disturbance will therefore travel at the higher speed u_1 and its leading edge will tend to catch up with the leading edge of the first disturbance, as shown in the middle diagram. The same applies to succeeding disturbances so that the front of the overall disturbance will steepen as shown. Obviously, the succeeding small disturbances can never actually overtake one another: the density cannot be multi-valued at any point! After sufficient time, therefore, the overall disturbance will become the *shock front*, shown in the lower diagram,

[16] This material will not be examined in detail.

which will be travelling faster than the speed of sound in the undisturbed gas.

Let us investigate the relationships between the conditions in the material behind a shock front and those in the un-shocked material. It is most convenient to work in a frame of reference in which the shock front itself is at rest, as shown in Figure 3-10.

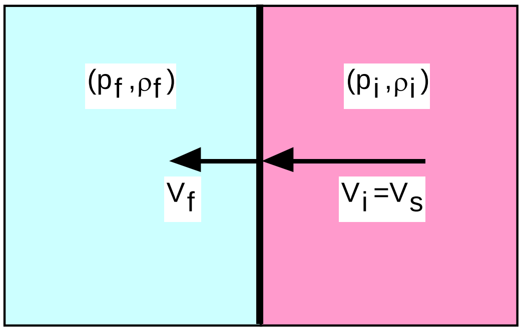

Figure 3-10. Frame of shock front.

In this case, the material upstream of the shock – the un-shocked material – has pressure p_i and density ρ_i and is flowing into the shock front with velocity V_s, the velocity of the shock front in the gas. Downstream of the front, the material, with pressure p_f and density ρ_f, is flowing away from the front with velocity V_f. To relate the upstream and downstream quantities, we need to remember that mass, momentum and energy must be conserved as the material crosses the front.

Consider the material flowing across the front within a cylinder of cross-sectional area A, as shown in Figure 3-11. In time t, a volume $Av_s t$ of material will flow into the front and a volume $AV_S t$ will flow out of it. Conservation of mass then tells us that

$$\rho_i \times AV_S t = \rho_f \times AV_f t \tag{3.74}$$

or

$$\rho_i V_S = \rho_f V_f \, . \tag{3.75}$$

The momentum flowing into the front in time t is $(\rho_i AV_S t)V_S$ whilst that flowing out is $(\rho_f AV_f t)V_f$.,. The change in momentum must be caused by the impulse provided by the difference in pressure across the front. We have, therefore,

$$\rho_f V_f^2 At - \rho_i V_i^2 At = (p_i - p_f)At \tag{3.76}$$

or

$$p_i + \rho_i V_S^2 = p_f + \rho_f V_f^2 \, . \tag{3.77}$$

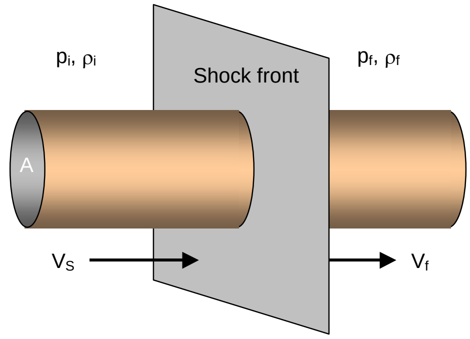

Figure 3-11. Material crossing shock front.

Finally, the energy flowing into the front is the sum of the internal energy $u_i At$, where u_i is un-shocked internal energy density, the and the kinetic energy $(\rho_i At)V_S^2/2$. In addition, the pressure does work $p_i At$ pushing the gas into the front. Similarly, the gas emerging from the front carries internal and kinetic energy away and does work *on* the surrounding gas. The conservation of energy therefore requires that[17]

$$\begin{aligned} & p_i At + u_i At + \frac{1}{2}\rho_i At V_S^2 \\ & = p_f At + u_f At + \frac{1}{2}\rho_f At V_f^2 \end{aligned} \tag{3.78}$$

or, using equation (3.75),

$$\frac{p_i + u_i}{\rho_i} + \frac{1}{2}V_S^2 = \frac{p_f + u_f}{\rho_f} + \frac{1}{2}V_f^2 \, . \tag{3.79}$$

The solution of the three simultaneous equations (3.75), (3.77) and (3.79) is straightforward but tedious. The result is

$$\begin{aligned} & \left(\frac{p_f + u_f}{\rho_f}\right) - \left(\frac{p_i + u_i}{\rho_i}\right) \\ & = \frac{1}{2}\left(\frac{\rho_i + \rho_f}{\rho_i \rho_f}\right)(p_f - p_i). \end{aligned} \tag{3.80}$$

This is known as the *Hugoniot*. For a perfect gas at temperature T, we have

$$u = c_V T; \qquad p = \frac{\rho}{\mu \, m_H} kT, \tag{3.81}$$

where c_V is the specific heat at constant volume, μ is the mean molecular weight of the gas and m_H is the mass of the hydrogen atom.

But

[17] I am not considering situations in which energy is input to the gas from other sources such as by photons in an ionisation front.

$$c_p = c_V + \frac{\rho}{\mu \, m_H} k; \qquad \gamma := \frac{c_p}{c_V}, \tag{3.82}$$

where c_p is the specific heat at constant pressure. From equations (3.81) and (3.82), we have

$$u + p = \frac{\gamma}{\gamma - 1} p \tag{3.83}$$

so that equation (3.80) can be written in the form

$$\frac{\rho_f}{\rho_i} = \frac{(\gamma_f + 1)p_f + (\gamma_f - 1)p_i}{(\gamma_i - 1)p_f + (\gamma_i + 1)p_i}. \tag{3.84}$$

If the downstream pressure is very much greater than the upstream pressure, it is called a *strong shock*, in which case.

$$\frac{\rho_f}{\rho_i} \approx \frac{\gamma_f + 1}{\gamma_i - 1}. \tag{3.85}$$

The diffuse interstellar medium is mainly monatomic, with $\gamma = 5/3$, before and after the shock. In these regions, therefore, the density downstream is four times the upstream density in a strong shock. At the other extreme, a strong shock in a dense molecular cloud, with $\gamma_i = 7/3$, leaves material dissociated, with $\gamma_f = 5/3$. In this case, the density jumps by a factor of five.

Concentrating on the latter case, we see from equation (3.75), it is clear that the gas leaves the shock front at one fifth of the shock velocity. From equations (3.81) and (3.84), we have for the ratio of the downstream temperature T_f of the gas to the upstream temperature T_i,

$$\frac{T_f}{T_i} \approx \frac{p_f}{p_i} \frac{\gamma_i - 1}{\gamma_f - 1}. \tag{3.86}$$

Equation (3.86) shows that the temperature of the gas increases in direct proportion to the jump in pressure. There is a simple physical explanation for this: the bulk kinetic energy upstream of the shock is converted into disordered internal energy at the shock, raising the temperature of the gas.

Finally, from equations (3.75), (3.77) and (3.84), it can be shown that the shock velocity V_S is related to the initial density and the final pressure in a strong shock by

$$V_S^2 \approx \frac{(\gamma_f + 1)p_f}{2\rho_i}. \tag{3.87}$$

We shall see that shocks play an important rôle in the behaviour of interstellar gas.

4. Dynamics of Spiral Discs

4.1 Overall Description

We saw in chapter 2 that spiral galaxies consist of several components, the halo, the nucleus and the disc. We have also seen that the nucleus is similar to spheroidal galaxies. I shall here concentrate upon the disc component; in particular I want to discuss the spiral structure, which is such a striking feature of the discs of these galaxies.

When we looked at the masses of spiral galaxies in chapter 2, we found that the disc material in spirals undergoes roughly plane-circular motion about the centre of the galaxy. This is only an approximation, though. If we take truly circular motion in a plane as the zeroth-order approximation, then we must add higher order motions to get an accurate picture. In practice, we shall consider only a first-order approximation as follows. First, we shall assume that deviations of the disc material from circular motion are very small. Secondly, we shall assume that the motion perpendicular to the plane is completely "decoupled" from that in the plane. By this I mean that, when analysing the motion perpendicular to the plane, we can completely forget about the motion in the plane, and vice versa.

Why can we do this? First, as we saw in section 2, the motion of stars is essentially governed by the overall gravitational field of the galaxy and not by the fields of *individual* stars. Secondly, the motion of the material at any point is determined by the gravitational field at that point. If motion *perpendicular* to the plane does not take the star into regions where the gravitational forces *parallel* to the plane are significantly different from those acting in the plane, then the perpendicular motion will not significantly affect the motion in the plane. And *vice versa*; if the motion in the plane does not change the gravitational forces acting perpendicular to the plane, then this perpendicular motion will be unaffected of the motion in the plane. Because the motions are relatively small, it turns out that both assumptions are good and this makes the problem much easier. I shall deal with the plane and perpendicular motions in turn.

4.2 Co-ordinate System

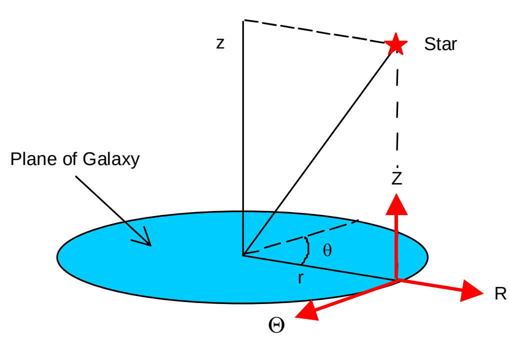

Figure 3-12. Co-ordinate system.

Because the disc component of the galaxy exhibits cylindrical symmetry to a zeroth-order approximation, it is sensible to use cylindrical polar co-ordinates to describe it, as shown in Figure 3-12. Relative to the centre of the galaxy, we shall use the co-ordinates *(r,θ)* in the plane and the co-ordinate z perpendicular to the plane[18]. We will denote by the capitals (R,Θ,Z) the linear velocities corresponding to motion in the (r,θ,z) directions respectively. The symbol Ω will be used to denote the *angular* velocity $d\theta/dt$.

4.3 Motion Perpendicular to the Plane of the Disc[19]

4.3.1 Equations of Motion

If we assume that the z-component of the motion of disc stars is independent of the motion in the plane, we can write for the equation of motion of a star of mass m perpendicular to the plane

$$m\ddot{z} = -m\frac{\partial\Phi(r,\theta,z)}{\partial z} \approx -m\frac{\partial\Phi(z)}{\partial z}, \tag{4.1}$$

where $\Phi(\rho,\theta,z)$ is the gravitational potential and the approximation follows from the argument of section 4.1 A dot denotes differentiation with respect to time. Assuming that we are only considering small deviations from the plane, we can expand $\partial\Phi(z)/\partial z$ as a Taylor series and keep only the first order terms:

$$\begin{aligned}\frac{\partial\Phi(z)}{\partial z} &\approx \left.\frac{\partial\Phi(z)}{\partial z}\right|_{z=0} + z\left.\frac{\partial^2\Phi(z)}{\partial z^2}\right|_{z=0} \\ &= \Phi_0' + z\Phi_0'',\end{aligned} \tag{4.2}$$

where a prime denotes differentiation with respect to z. We shall also assume, reasonably I think, that the disc is a plane of symmetry of the galaxy so that the first term on the right hand side of equation (4.2) vanishes. We can get the second term from Poisson's equation and, again following the discussion of section 4.1, we find that

$$\Phi_0'' \equiv \left.\frac{\partial^2\Phi}{\partial z^2}\right|_{z=0} \approx \left.\nabla^2\Phi\right|_{z=0} = 4\pi G\rho_o, \tag{4.3}$$

where ρ_o is the density in the plane. Using equations (4.1), (4.2) and (4.3), we get as the equation of motion in the z-direction

$$\ddot{z} = -(4\pi G\rho_o)z. \tag{4.4}$$

Equation (4.4) shows that, to this approximation, the motion perpendicular to the plane is simple harmonic and is determined by the density in the disc. If we were able to trace out the motion of a disc star in the z-direction, we could determine the density in the disc. This is impossible for external galaxies but it can be applied to the Galactic disc in the solar neighbourhood. Even here, we cannot, follow the motion of an individual star because the oscillatory period is far too long. We have once again to resort to statistical arguments. The result is that the density in the solar neighbourhood comes out at about $0.15\ M_{sun}\ pc^{-3}$. Remember that this is derived dynamically. If we try to account for this by adding up the contributions of directly observed matter, we find a total of about $0.11\ M_{sun}\ pc^{-3}$, made up of about $0.08\ M_{sun}\ pc^{-3}$ of stars and $0.03\ M_{sun}\ pc^{-3}$ of gas. Even locally, in the solar neighbourhood, we come across the problem of dark matter!

4.4 Motion of Stars in the Plane of the Disc

4.4.1 Circular Motion[20]

If the motion were truly circular about the centre of the galaxy, r and z would be constant and we should have for the velocity Θ of any star at distance r from the centre of the galaxy

$$-\frac{m\Theta^2}{r} = -m\frac{\partial\Phi}{\partial r}, \tag{4.5}$$

which says that the centripetal force necessary for circular motion is supplied by the radial gradient of the gravitational potential. [The minus sign appears on the left-hand side of the equation because the centripetal force is directed towards the centre.] What should we use for Φ? I have already discussed the existence of massive, spherically symmetric halos about spiral galaxies when talking about rotation curves in chapter 2. Let us therefore make the assumption of spherical symmetry of the mass distribution, justifying it by the agreement of prediction with observation.

For a spherically symmetric distribution of matter, the gravitational field at a distance r from the centre of the distribution depends only upon the material contained within r and is directed towards the centre. We can therefore write

$$\frac{\partial\Phi(r)}{\partial r} = \frac{GM(r)}{r^2} \tag{4.6}$$

where $M(r)$ is the total mass contained within radius r. From equations (4.2) and (4.6) we find that

$$\Theta(r) = \left[\frac{GM(r)}{r}\right]^{1/2}. \tag{4.7}$$

We have already seen in chapter 2 that rotation curves are very flat so a reasonable approximation, for regions other than the central parts of the galaxy, is to assume that Θ is independent of r, i.e.

$$\Theta = \Theta_o = \text{constant}. \tag{4.8}$$

[18] Note that this differs from the historical use of ϖ – a script π – rather than r.

[19] This is not formally part of the course and will not be examined. It is here for completeness.

[20] For convenience, I reproduce some of the discussion of Chapter 2 here.

As we have seen in chapter 2, the consequence of this is that the mass $M(r)$ increases linearly with r and that the density $\rho(r)$ falls off as the inverse square of r.

4.4.2 Epicyclic Motion

We have so far considered stars in perfectly circular orbits although we know that this is only an approximation. As well as the z-component of motion, which we have already considerrd, disc stars have small peculiar velocities *within* the plane, superimposed upon their circular motion.

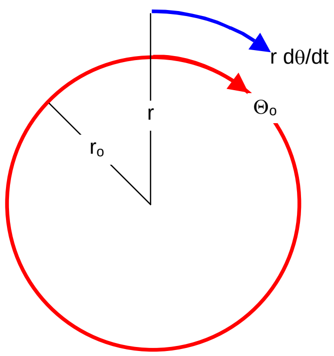

Figure 3-13. Perturbation of circular motion.

Suppose we take an individual star with circular velocity Θ_o, at radius r_o from the centre of the galaxy, and displace it slightly in the radial direction to distance r, as shown in Figure 3-13, *whilst keeping its angular momentum the same*. What will its subsequent motion be? Because the only force acting is a central force, the star's angular momentum must *remain* constant after the displacement. Now the angular momentum L is given by

$$L = mr\Theta = mr^2\Omega \,, \tag{4.9}$$

where m is the mass of the star and Ω is the angular velocity of the star about the centre of the galaxy:

$$\Omega = \dot{\theta} \,. \tag{4.10}$$

The constancy of the angular momentum requires that

$$mr\Theta = mr^2\Omega = mr_o^2\Omega_o = mr_o\Theta_o \,. \tag{4.11}$$

The star's circular velocity Θ is therefore given by

$$\Theta = \Theta_o \times \frac{r_o}{r} \propto \frac{1}{r} \,. \tag{4.12}$$

As the star moves out, therefore, its circular velocity will decrease from the unperturbed value of Θ_o. It will therefore be moving around the centre of the galaxy *slower* than the other stars which are in circular orbit at this slightly larger radius and which therefore have velocity Θ_o. Relative to these stars, therefore, the perturbed star will appear to be moving backwards. Furthermore, with its reduced velocity, it will not have sufficient centrifugal force to overcome the gravitational force and it will tend to drop in towards the centre of the galaxy. As it drops down below its original circular orbit at r_o, it will continue to conserve angular momentum and this means that the it will now be going *too fast* with respect to stars in circular orbit at this reduced radius and will overtake them. It will also now having too much centrifugal force, will move back out again … .

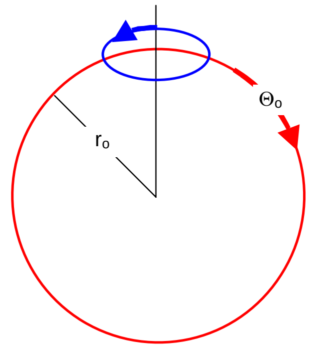

Figure 3-14. Rotating frame of reference.

This is best illustrated by using a frame of reference rotating with angular velocity Ω_o such that

$$r_o\Omega_o = \Theta_o \,. \tag{4.13}$$

In this rotating frame of reference, an unperturbed star at radius r_o appears be at rest. By the arguments above, the perturbed star will appear to be moving backwards in this frame when it is at radii $r > r_o$ and moving forwards when $r < r_o$, as shown in Figure 3-14. This motion of a perturbed star is called epicyclic motion, from the Greek $\varepsilon\pi\iota$ meaning upon and $\kappa\upsilon\kappa\lambda o\varsigma$ meaning circle.

The detailed treatment of this epicyclic motion is tedious and I simply quote the results here. In the frame of Figure 3-14 a star at radius r traces out an ellipse[21] with angular frequency $\kappa(r)$, given by

$$\kappa^2(r) = -4B(r)[A(r) - B(r)] \,, \tag{4.14}$$

where the Oort parameters $A(r)$ and $B(r)$ are given by

$$\begin{aligned} A(r) &= +\frac{1}{2}\left[\frac{\Theta(r)}{r} - \frac{d\,\Theta(r)}{dr}\right]; \\ B(r) &= -\frac{1}{2}\left[\frac{\Theta(r)}{r} + \frac{d\,\Theta(r)}{dr}\right]. \end{aligned} \tag{4.15}$$

It is easy to see that A is a measure of the *differential rotation* of the galaxy: if the galaxy were to rotate like a solid body – that is with constant angular velocity – A

[21]Note that the frame in which an epicyclic orbit appears as an ellipse depends upon the radius r from the centre; different frames must be used at each radius.

would be zero and B would be equal to the minus angular velocity Ω.

4.4.3 RESONANT ORBITS

In general, the orbit of a star undergoing epicyclic motion will not be closed in the inertial (non-rotating frame). Consider the radial component of epicyclic motion for a star at radius r. [The angular component of the motion will look after itself.] The frequency of radial oscillation is $\kappa/2\pi$ so that the time $T_{\text{epicycle}}(q)$ taken for the star to complete q radial epicyclic oscillations is given by

$$T_{\text{epicycle}}(q) = q \times \frac{2\pi}{\kappa} . \quad (4.16)$$

The frequency of rotation about the centre of the galaxy is $\Omega/2\pi$ so that the time $T_{\text{orbit}}(p)$ needed for p complete orbital rotations about the centre of the galaxy is given by

$$T_{\text{orbit}}(p) = p \times \frac{2\pi}{\Omega} . \quad (4.17)$$

The orbit of a star would therefore be closed after p orbits of the galaxy if

$$T_{\text{epicycle}}(q) = T_{\text{orbit}}(p) \quad (4.18)$$

or, from equations (4.4) and (4.17), if

$$\frac{p}{\Omega} = \frac{q}{\kappa} . \quad (4.19)$$

Condition (4.19) will not normally be met in the inertial (non-rotating) frame which we have used previously. It is always possible, however, to choose a *rotating* frame in which the epicyclic orbits of stars at any particular radius from the centre of a galaxy *are* closed. Consider a frame rotating with velocity Ω_p (the reason for the choice of subscript will become clear later). In this frame, the orbital angular velocity Ω' of the star is given by

$$\Omega' = \Omega - \Omega_p . \quad (4.20)$$

The condition for closure in *this* frame is, therefore,

$$\pm \frac{p}{\Omega'} = \frac{q}{\kappa} . \quad (4.21)$$

The ± sign occurs because Ω' may be positive or negative; in other words, Ω_p may be smaller or larger that Ω (p and q are positive by definition). From equations (4.20) and (4.21), we get for closure of orbits

$$\Omega_p = \Omega \pm \frac{p}{q}\kappa . \quad (4.22)$$

We shall see later that, for our Galaxy, the values $p = 1$, $q = 2$ have special significance.

I have spent some time on this topic because it is relevant to the study of spiral structure. The next section develops another aspect of this that is particularly important.

4.5 Spiral Structure

4.5.1 INTRODUCTION

The spiral pattern seen in many galaxies is one of the most beautiful and spectacular sights in astronomy. We see no other regular pattern on such an enormous scale. It is therefore natural that a lot of effort has gone into trying to explain this structure. In spite of this, the *origin* of spiral patterns is not entirely clear although we do seem to have a satisfactory explanation of how they may maintain themselves once they have formed.

4.5.2 THE WINDING PROBLEM

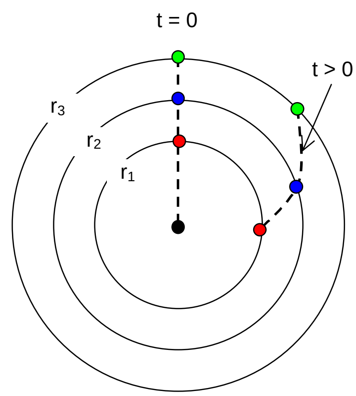

Figure 3-15. The origin of the winding problem.

We can see very easily that the material in each spiral arm of a galaxy must constantly be changing. Figure 3-15 shows three stars or clouds of gas at different radial distances r_1, r_2 and r_3 from the centre of the galaxy, where

$$r_2 = \frac{3}{2} r_1 = \frac{1}{2} r_3 . \quad (4.23)$$

Initially, at time t = 0 say, these clouds happen to be aligned along a radius. If they were to remain so aligned the galaxy would have to rotate as a rigid body and the angular velocity $\Omega(r)$ at any radius r would actually be independent of r. We have seen, however, that far from Ω being constant, it is the *circular* velocity $\Theta(r)$ which is approximately independent of r. Let us assume for the moment that Θ is truly constant with value Θ_o so that

$$\Omega(r) = \frac{\Theta(r)}{r} = \frac{\Theta_o}{r} . \quad (4.24)$$

The, after a time t, a star at radius r will have turned through an angle $\theta(r,t)$ given by

$$\theta(r,t) = \Omega(r)t = \frac{\Theta_o}{r}t \ . \qquad (4.25)$$

Consider two clouds at radii r_1 and r_2. It is clear that the cloud at radius r_2 will have moved through only half the angle that the cloud at r_1 has moved in any given time. This is shown in Figure 3-15; at time *t*, the third cloud from the centre has moved through an angle of only $\pi/4$ while the second cloud out has moved through $\pi/2$.

It might be argued that this differential rotation gives just the spiral pattern we are looking for. It is easy to see that this cannot be so. The time for a cloud to complete an orbit of a galaxy, although obviously dependent on its distance from the centre, is of the order of 10^8 years. On the other hand, galaxies are some 10^{10} years old, as we shall see later. In the lifetime of the galaxy, therefore, there have been some hundred or so orbits. If, for the sake of argument, I assume that the innermost cloud out in Figure 3-15 had completed 100 orbits, then the outermost cloud out would have completed 200 and the line joining them in the figure would be wound up one hundred times. Because we do not see galaxies with arms wound this tightly, we must seek some mechanism which preserves the relatively loosely wound spirals.

Although the tightness of winding in spirals changes as one goes along the Hubble sequence, the arms being less tightly wound in later types than in earlier, the total range is not large. There is no evidence, moreover, that the sequence is an evolutionary one, and that arms are either winding or unwinding. We are therefore justified in assuming that it is the spiral *pattern* that remains more or less unchanged with time. This can only be achieved if the pattern rotates like a rigid body, with angular velocity Ω_p, independent of radius from the centre of the galaxy.

4.5.3 The Physical Mechanism Sustaining Spiral Arms

If we reject the idea that the stars and clouds of gas making up an arm at any given time *remain* within that arm, the only alternative is that new material must be moving through the arms, flowing in at one side and out the other. In this section, I give a qualitative description; some more quantitative detail is given in the next in the next section.

Figure 3-16 shows the basic scheme for a two-armed spiral. We assume that the spiral pattern, represented by the thick lines, rotates like a rigid body with angular velocity Ω_p and that the stars and gas move in essentially circular orbits, represented by the thin lines, around the centre of the galaxy. I shall also assume that this material is always *overtaking* the arms. In this model, the arms represent *concentrations* of matter so that the density in the arms is higher than in the rest of the plane. The gravitational potential in the disc, therefore, no longer has axial symmetry. Because of this distortion in the potential, the stars and gas will be perturbed slightly from their circular orbits. If the system is to be self-perpetuating, this perturbation must be just such as to maintain the distortion of the potential.

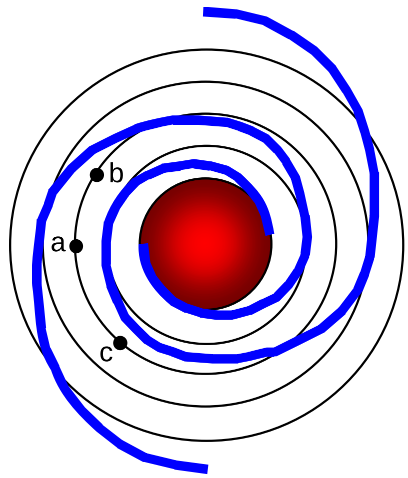

Figure 3-16. Maintenance of spiral structure.

The behaviour can be seen qualitatively from the figure. Consider first a star or cloud of gas at position *a*. It is roughly equidistant from both arms and will therefore suffer no net radial force. At *b*, on the other hand, it will be attracted outward by the extra material in the nearer (outside) arm and will move to a slightly larger radius. Conserving its angular momentum, it will slow down a little in its orbit and will tend to linger in the vicinity of the arm. The material bunched in this way gives rise to the additional potential that caused the bunching in the first place. When a star eventually leaves a spiral arm, as at *c*, it is attracted inwards and speeds up, so regaining its initial speed. In this way, the spiral perturbation is preserved although the material which gives rise to it is constantly changing[22].

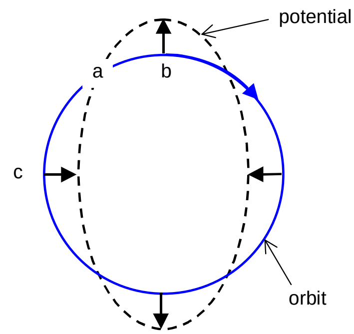

Figure 3-17. Spiral perturbations in a rotating frame.

[22] The effect is similar to a bottleneck on a motor way. Assume a steady state of motion, not a complete jam, so that cars leave the jam at the same rate as those joining it. If one carriageway is closed off, there are more cars per unit length of the motor way there than at other places. Note that, although the increased density of traffic is always at the same place, it is made up of a constantly changing population of cars.

The effect is shown schematically in Figure 3-17, which is drawn in a frame of reference rotating at the angular velocity Ω_p of the two-armed spiral pattern. The force acting on the object – a star or cloud of gas – is represented by the arrows between the circular orbit of the object and the ellipse representing the departure from circular symmetry of the gravitational field. Clearly this distorted gravitational field subjects the object to a perturbing force of frequency Ω' given by

$$\Omega' = \Omega - \Omega_p . \tag{4.26}$$

4.5.4 Density-Wave Theory[23]

The accepted description of spiral structure is that a *density wave* propagates through the material of the disc (see [2], for example). The full treatment of density waves in complicated and I shall outline the procedure, giving results that are valid only for *tightly-wound* spirals.

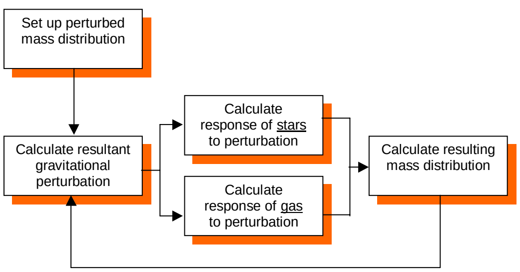

Figure 3-18. Self-consistent solution for density waves.

The overall procedure is shown in Figure 3-18. Starting at the top left-hand corner, we set up an initial perturbed mass-distribution and calculate the resultant gravitational potential. We then calculate separately[24] the dynamical responses of the stars and the gas to this distribution. Finally, we calculate the mass distribution resulting from these responses and the corresponding potential. This potential is fed back into the start of the process and we continue going around the loop until the result is self-consistent.

The behaviour of the gas is well described by the equations of hydrodynamics. As we saw in section 2, however, stars are virtually collisionless in their motion and this makes stellar dynamics difficult. Fortunately we can obtain rather good insight to the behaviour of density waves by ignoring completely the presence of the stars. This is at first sight surprising when you consider that the stars dominate the mass of the disc. I shall explain the reason later.

Rather than deal with the *volume* density – the mass per unit volume ρ – it is convenient to use the *surface* density σ, the mass per unit surface area of the disc. Obviously,

$$\sigma(r,\theta) = \int_{-\infty}^{+\infty} \rho(r,\theta,z)\,dz . \tag{4.27}$$

As when we dealt with sound waves, let us consider a first-order perturbation $\sigma_1(r,\theta,t)$ to the unperturbed, cylindrically symmetric time-independent distribution $\sigma_0(r)$:

$$\sigma(r,\theta,t) = \sigma_0(r) + \sigma_1(r,\theta,t) . \tag{4.28}$$

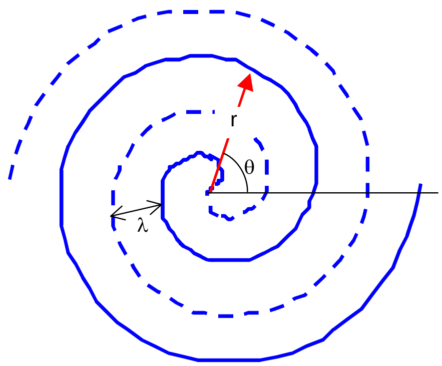

Figure 3-19. Spiral pattern.

Let us try solutions of the form[25]

$$\begin{aligned}\sigma_1(r,\theta,t) &= \sigma_{1o} \exp[i(\alpha t - m\theta - kr)] \\ &= \sigma_{1o} \exp[i\phi]\end{aligned} \tag{4.29}$$

Why? Because it represents a rotating, spiral wave. To see this, consider the pattern generated by following the locus of constant values of σ_1. From equation (4.29) we have

$$\alpha t - m\theta - kr + n2\pi = \phi_o = \text{constant} , \tag{4.30}$$

where the term $n2\pi$ appears because σ_1, given by equation (4.29), is periodic in ϕ with period 2π. Adding $n2\pi$ to ϕ does not therefore change σ_1. Let us first trace the pattern at a given *time* which, for convenience, we take as $t = 0$; then we have

$$m\theta + kr - n2\pi = -\phi_o = \text{constant} . \tag{4.31}$$

From equation (4.31), we have

$$r = \frac{1}{k}(n2\pi - m\theta - \phi_o) , \tag{4.32}$$

which is an m-armed spiral, as shown in Figure 3-19 for $m = 2$. [For clarity, the second arm is shown dashed.] To

[23] The analytical details of this section will not be examined.

[24] Remember that the stars and gas behave very differently.

[25] In fact, we should represent the density wave as a Fourier series of similar terms but the first-term is adequate for demonstrating the principles of the theory.

see this, consider the moving an angle $\Delta\theta$ around the pattern at a constant value of r. If we are get to a point where the phase – and therefore σ_1 – has the same value ϕ_o, we must have from equation (4.31).

$$m(\theta+\Delta\theta)+kr-(n+1)2\pi=-\phi_o \quad (4.33)$$

so that

$$\Delta\theta=\frac{2\pi}{m} \quad (4.34)$$

and the pattern repeats every $2\pi/m$.

The radial distance λ between adjacent arms of the spiral – the wavelength of the spiral wave – is got by keeping the *angle* θ fixed and putting

$$r+\lambda=\frac{1}{k}\left[\phi_o+(n+1)2\pi-m\theta\right]. \quad (4.35)$$

Taking equation (4.32) from equation (4.35), we get

$$\lambda=\frac{2\pi}{k} \quad (4.36)$$

so that k has its usual meaning of wavenumber.

If we differentiate equation (4.30) with respect to t at constant r, we get

$$\left.\frac{\partial\theta}{\partial t}\right|_r=\frac{\omega}{m} \quad (4.37)$$

so that the spiral wave rotates at a speed Ω_p given by

$$\Omega_p=\frac{\omega}{m}. \quad (4.38)$$

Finally, differentiating equation (4.30) with respect to t at constant θ we get

$$\left.\frac{\partial r}{\partial t}\right|_\theta=\frac{\omega}{k} \quad (4.39)$$

so that, at a given angular position, a wave with wave-vector k advances radially into the interstellar gas with velocity v_{spiral} given by

$$v_{spiral}=\frac{\omega}{k}=\frac{1}{m}\frac{\Omega_p}{k}. \quad (4.40)$$

4.5.5 The Linearised

Because we are taking σ_1 to be small compared with σ_0, we seek to linearise the equations of hydrodynamics, again as we did for sound waves. I shall not go into details, because these are messy, but simply quote the results. The only difference between this and the earlier treatment is that we now have *two-dimensional*, rather than three-dimensional, hydrodynamics because we are considering the surface density of gas in the plane. Because the equations involve pressure, we need the two-dimensional analogue of the three-dimensional pressure p.

In a classical gas at temperature T, we have

$$p=nkT\ , \quad (4.41)$$

where n is the number-density of molecules. We also have that

$$\frac{1}{2}nm\langle v^2\rangle\left[=\frac{1}{2}\rho\langle v^2\rangle\right]=\frac{3}{2}kT\ , \quad (4.42)$$

where m is the mass of the molecules. Hence

$$p=\frac{1}{2}\rho\langle v^2\rangle. \quad (4.43)$$

It turns out that, in interstellar gas, the overall pressure is dominated by the mean-squared *turbulent* velocity $\langle a^2\rangle$ of the gas, rather than the mean-squared velocity of the molecules. Equation (4.43) is therefor replaced with

$$p=\frac{1}{2}\rho\langle a^2\rangle. \quad (4.44)$$

Finally, it is easy to show that the two-dimensional pressure[26] P is given, in terms of the surface-density σ, by

$$P=\sigma\langle a^2\rangle. \quad (4.45)$$

After solving the linearised equations, we get for σ_1,

$$\frac{\sigma_{1o}}{\sigma_0}=-i\frac{kg_{radial}}{\kappa^2-m^2(\Omega_p-\Omega)^2+k^2\langle a^2\rangle}\ ,$$

where κ is the epicyclic frequency defined in equation (4.14) and g_{radial} is the radial component of the perturbed gravitational force, given by

$$g_{radial}=2\pi iG\sigma_{1o}\frac{k}{|k|}. \quad (4.46)$$

4.5.6 The Dispersion Relation

If I substitute for g_{radial} from equation (4.46) into equation (4.49), we get after some manipulation

$$m^2(\Omega_p-\Omega)^2=\langle a^2\rangle k^2+\kappa^2-2\pi G\sigma_0|k|, \quad (4.47)$$

which, using equation (4.38), can also be written as

[26] The dimensions of P are force per unit length, rather than the force per unit area of p.

$$(\omega - m\Omega)^2 = \left\langle a^2 \right\rangle k^2 + \kappa^2 - 2\pi G\sigma_0 |\, k\,| . \qquad (4.48)$$

You should compare this with the dispersion relation (3.54) for sound waves in a gravitational field:

$$\omega^2 = u^2 k^2 - 4\pi G \rho_o .$$

On the left-hand side of both equations appears the frequency of the perturbation, albeit modified by the pattern frequency in the case of the density wave. On the right-hand side in both cases, the wave-vector appears multiplied by a velocity and gravity appears in a destabilising term, albeit involving the wave vector *k* in the density-wave case. [Remember that it is this term that gave rise to gravitational collapse in section 3.2.7.] In the case of the density wave, the gravitational term is off-set by the essentially positive term involving the epicyclic frequency.

Let us re-write equation (4.47) again, now in the form

$$\frac{|\, k\,|}{k_o} = 1 + x^2 - \nu^2 , \qquad (4.49)$$

where:

- $$k_o = \frac{\kappa^2}{2\pi G\sigma_0} ; \qquad (4.50)$$
- Toomre's stability number *x*, is given by
$$x = \frac{\left\langle a^2 \right\rangle k^2}{\kappa^2} ; \qquad (4.51)$$
- the dimensionless, normalised frequency ν is given by
$$\nu = \frac{m(\Omega_p - \Omega)}{\kappa} . \qquad (4.52)$$

Before we analyse the dispersion relation (4.49), let us think for a moment about the normalised frequency ν. At any distance from the centre of the galaxy, $(\Omega_p - \Omega)$ is difference in angular velocity between the spiral pattern and the gas. Because there are *m* arms in the pattern, the gas therefore meets a spiral arm with frequency $m(\Omega_p - \Omega)/2\pi$. In time *T*, therefore, the gas has $m(\Omega_p - \Omega)T/2\pi$ encounters with an arm. In the same time, an individual cloud will make $\kappa T/2\pi$ epicyclic oscillations. Hence ν, which can be re-written as,

$$\nu = \frac{\lfloor m\,(\Omega_p - \Omega)T/2\pi \rfloor}{[\kappa T/2\pi]} , \qquad (4.53)$$

measures the number of encounters of a gas cloud with a spiral arm *per* epicyclic oscillation. If ν is an integer, then there is *resonance* between epicycles and arm encounters and we might expect this to cause havoc with the motion (see below).

Returning to (4.49), we see that the left-hand side of the equation is essentially non-negative. In fact, if we are to have a wave at all, so that | *k* | is not zero, then the left-hand side must be positive so we must have

$$\nu^2 > 1 + x^2 , \qquad (4.54)$$

This relation must be satisfied if the gravitational perturbation, produced by the perturbation in surface density σ_1, actually acts to *bring about* that perturbation in density. In other words, (4.54) is a condition for self-consistency. From (4.54), we have

$$-\sqrt{1+x^2} < \nu < +\sqrt{1+x^2} \qquad (4.55)$$

or, from equation (4.52),

$$\Omega(r) - \frac{\kappa(r)}{m}\sqrt{1+x^2} < \Omega_p < \Omega(r) - \frac{\kappa(r)}{m}\sqrt{1+x^2} \; , (4.56)$$

where I have restored the explicit dependence of the frequencies $\Omega(r)$ and $\kappa(r)$ from the centre of the galaxy[27]. Relation (3.32) tells us that the pattern speed Ω_p – which must be constant if we are to have an enduring density-wave – is constrained to lie between values that are a function of *r*.

If the root-mean-squared turbulent velocity $<a^2>^{1/2}$ were zero, then *x* Toomre's stability number would also be zero and we should have

$$\Omega(r) - \frac{\kappa(r)}{m} < \Omega_p < \Omega(r) + \frac{\kappa(r)}{m} . \qquad (4.57)$$

$(\Omega - \kappa/m)$ and $(\Omega + \kappa/m)$ are called the inner an outer *Lindblad resonances*, respectively. The condition (4.57) is easy to understand physically. In terms of the dimensionless frequency ν, it is

$$-1 < \nu < +1 . \qquad (4.58)$$

Remembering the discussion following the definition of ν (4.52), we see that inequality (4.58) says that we must avoid resonances between the frequency with which a gas cloud meets perturbing effects of the arms and its epicyclic frequency[28]. We know that, if we repeatedly "hit" something at its resonant frequency, the oscillations tend to build up and the object may destroy itself. A dramatic example of this in "everyday" life is collapse of the Tacoma Narrows suspension bridge in Washington State, USA, recorded in a famous film sequence. In the absence of any turbulent velocity,

---

[27] $x(r)$ is also a function of this distance.

[28] This is just a special case of the condition for resonant orbits discussed previously.

therefore, spiral density waves can therefore only exist between the Lindblad resonances.

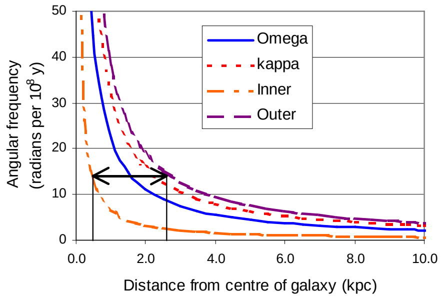

Figure 3-20. The Lindblad resonances.

Figure 3-20 shows $\Omega(r)$, $\kappa(r)$ and the two Lindblad resonances for a two-armed spiral, plotted as a function of distance r from the centre of a galaxy. It is clear from the diagram that a spiral wave whose angular velocity is about 13 radians per 10^8 years, for example, can exist in this galaxy only between about 0.5 and 1.5 kpc from the centre. Note that the *circle of co-rotation*, where $\Omega_p = \Omega$, lies between the two resonances.

In practice, the turbulent velocity is not zero and inequality (4.55) shows that this turbulence *increase* the range over which the spiral pattern can exist. Hence the name *stability number* for x.

4.5.7 The Effects of and on the Stars

You will remember that we have totally neglected the stars in the above discussion. This is not entirely justified but we can see qualitatively that only a fraction of them respond to the perturbations. The stars will be undergoing epicyclic motion and they will be doing so almost unaffected by the other stars (the collisionless approximation) in sharp contrast to the gas. A star will therefore freely wander in and out about its mean radius. Those stars whose radial epicyclic excursions take them more that the spacing between arms will naturally become confused about which arm they are supposed to be reacting to. They will not therefore respond to the perturbations. Only a fraction of the stars, therefore, those that have radial amplitudes small compared with the inter-arm spacing, will respond to the density waves. That is why we can get a good approximation using the gas alone.

4.5.8 The Speed of a Spiral Pattern

Although we have come up with a plausible theory of spiral structure, we have not yet compared this with observation. Let us first try to fit the pattern speed Ω_p of a given spiral.

Equation (4.29) can be re-arrange to give θ as a function of r:

$$\theta = -\frac{1}{m}\left\{\frac{1}{i}\ln\left[\frac{\sigma_1(r,\theta,t)}{\sigma_{1o}}\right] - \alpha t + kr\right\}. \qquad (4.59)$$

If we follow a maximum value of σ_1 – that is, if we follow and arm – at some fixed value of t we can write equation (4.59) as

$$\theta = -\frac{1}{m}\{\text{constant} + kr\} = \Psi(r). \qquad (4.60)$$

say[29]. Knowing now that spiral structure can exist only between the limits given by (3.40), I can re-write this as the integral of the derivative of Ψ:

$$\theta(r) - \theta(r_{\text{in}}) = \frac{1}{m}\int_{r_{\text{in}}}^{r} \frac{d\Psi(r')}{\mathrm{d}r'}dr' = \frac{1}{m}\int_{r_{\text{in}}}^{r} k(r')dr', \qquad (4.61)$$

where r_{in} is the inner limit given by (4.56). It is clear that the solution of (4.48) for k depends only on Ω_p, $\Omega(r)$, $\kappa(r)$, $<a^2(r)>^{1/2}$ and $\sigma_0(r)$, all of which are observable quantities in principle. We can therefore write

$$\theta(r) - \theta(r_{\text{in}}) = \frac{1}{m}\int_{r_{\text{in}}}^{r} k\left[\Omega_p, \Omega(r'), \kappa(r'), \left\langle a^2(r)\right\rangle, \sigma_0(r')\right]dr' \qquad (4.62)$$

and adjust the parameter Ω_p so that the model fits the observed run of $\theta(r)$ with r. Typical values of Ω_p are found to be in the range 10 to 40 km s^{-1} kpc^{-1}, corresponding to a rotational period for the pattern of about 100 to 500 million years.

4.5.9 Testing the Theory

I have described a model of spiral structure that consists of linear sinusoidal waves of density in the material of the galactic disc. Why should these produce the *visible* spiral structure, which consists of mainly massive young stars and their associated HII regions? In fact linear theory is inadequate for this and we have to introduce some non-linear element. When we do this, we find that shock-fronts exist in the spiral arms. We shall also see that only trailing spiral arms are likely to exist.

If we take the azimuthal angle θ to increase in the clockwise direction, equation (4.29):

$$\begin{aligned}\sigma_1(r,\theta,t) &= \sigma_{1o}\exp[i(\alpha t - m\theta - kr)] \\ &= \sigma_{1o}\exp[i\phi]\end{aligned}$$

descibes a *trailing* spiral, as shown on the left of **Error! Reference source not found.**. To see this, consider following with t a locus of constant phase at a given radius. Then we have

[29] Remember that, in general, $k(r)$ will be a function of r.

$$0 = \alpha - m\left(\frac{\partial\theta}{\partial t}\right)_r \quad (4.63)$$

so that

$$\left(\frac{\partial\theta}{\partial t}\right)_r = \frac{\alpha}{m} > 0 \,. \quad (4.64)$$

This shows that the spiral rotates clockwise. Now consider following with *r* a locus of constant phase at a given time. We have

$$0 = -m\left(\frac{\partial\theta}{\partial r}\right)_t - k \quad (4.65)$$

so that

$$\left(\frac{\partial\theta}{\partial r}\right)_t = -\frac{k}{m} < 0 \,. \quad (4.66)$$

so that the angle decreases with *r*.

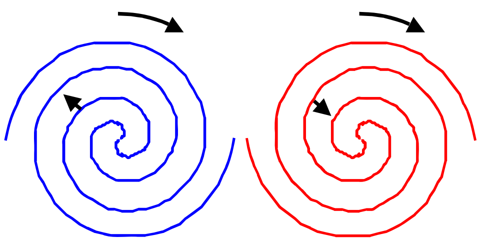

Figure 3-21. Trailing and leading spirals.

For a *leading* spiral, also rotating clockwise and shown on the right of Figure 3-21, the equation corresponding to (4.29) is

$$\sigma_1(r,\theta,t) = \sigma_{1o}\exp[i(\alpha t - m\theta + kr)] \,. \quad (4.67)$$

In a trailing spiral, the wave advances outward into the disc: at constant θ,

$$0 = \alpha - k\left(\frac{\partial r}{\partial t}\right)_\theta \,, \quad (4.68)$$

whereas it advances inward in a leading spiral.

In a leading spiral, therefore, the wave front moves toward generally less dense material in which the speed of sound is lower. It therefore tends to pile up on itself until it produces a shock front, as describes in section 3.2.9. As I shall show shortly, this tends to produce stars. The leading spiral, on the other hand, is moving into a region of increasing density and higher speed of sound. This allows the wave to "get away from itself" and reduces the effect of the perturbation. Since the spiral arms are dominated by hot young stars, therefore, we may conclude that only trailing waves are likely to exist[30].

Remember the discussion of Jeans collapse in sections 3.1.3 and 3.2.7. We saw there that only clouds of gas greater than a critical size could collapse to form stars. This critical size was inversely proportional to the square root of the density, as shown by equation (3.66). Suppose now that a cloud a little above the critical size enters a shock front. This will compress the cloud, increasing its density. If the increase in density is sufficient, the cloud will now be *less* than the critical size and will collapse in the free-fall time given by equation (3.70). Hence shock waves are capable of causing gas clouds to collapse to form stars in times of the order of 10^6 y[31]. Since the orbital period of the material about the centre of the galaxy is some 10^8 years, this means that stars form within a few degrees of the shock front. The massive O and B stars, which distinguish the arms, stay on the main sequence for only about 10^6 years so that we should expect the arms themselves to be a few degrees wide, as indeed we observe. All this is quite well illustrated in the image of the Whirlpool galaxy, shown in Figure 3-22.

Figure 3-22. Star-formation in the Whirlpool galaxy.

Bibliography for Chapter 3

[1] Binney, J and Merrifield, M. *Galactic Astronomy*, Princeton University Press, 1998. ISBN 0-691-02565-7

[2] Rohlfs, K. *Lectures on Density Wave Theory*, Springer-Verlag, Berlin (1977).

---

[30] In a few cases, it is possible to measure the sense of the spiral arms of a galaxy directly (cf. [1]). In all such cases, the wave are found to be trailing.

[31] This is an over-simplified description of star-formation in spiral arms. For more detail see the course PHY-410 *The Interstellar Medium*.