Kinetic Theory of Stellar Systems

Prof. J.J. Binney

Oxford University
Michaelmas Term 2014

URL

These notes can be downloaded from

`www-thphys.physics.ox.ac.uk/people/JamesBinney/lectures.html`

Books

Anyone interested in stellar dynamics will benefit from studying V.I. Arnold's beautiful *Mathematical Methods of Classical Mechanics* (Springer)

The standard text on stellar dynamics is *Galactic Dynamics* by J. Binney and S. Tremaine (Princeton). We'll refer to this as "BT08".

A useful summary of stellar discs by J.A.Sellwod is Chapter 18 of "Galactic structure and Stellar Populations", Vol 5 of "Planets stars and stelar systems", eds T.D.Oswalt & G Gilmore, Springer 2013

Some relevant material that's not covered in *Galactic Dynamics* can be found in Chapter 3 of *Secular Evolution of Galaxies*, eds J. Falcon-Barroso & J.H. Knapen (CUP)(arXiv:1202.3403)

Part I: Spheroidal Systems

Part II: Stellar Discs

1 Gravity: unshielded long-range force

We have to deal with particles that may be stars, planets, or dark-matter particles of mass $< 1\,\mathrm{TeV}$. We shall speak of "stars" regardless of whether the particles really are stars.

In a gas the particles interact by a short-range force, 2 at a time; in a plasma they interact by a long-range force, but shielding suppresses system-wide accumulation of force. In a stellar system the force is long-range and unshielded.

In this regard a stellar system is more complex than a plasma, but it is simpler because relativistic effects can be neglected – in a plasma the cancelling electrostatic forces are so huge that miniscule relativistic corrections to them become crucial – magnetism. In a stellar system one can neglect relativist corrections to the "electrostatic" force Gm_1m_2/r^2.

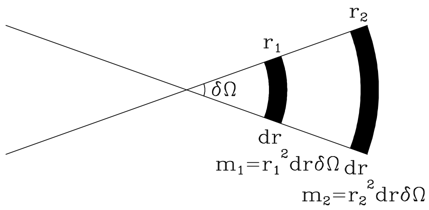

Fig. 1. Each shaded portio of the cone contributes equally to the force on the star at its apex.

The contribution to the force on a test star of mass m from stars at distance r in a cone is $\delta\mathbf{F} = [Gm\rho(r)r^2\delta\Omega/r^2]\delta r$, so over distances within which $\rho \sim$ const, equal forces come from equal intervals in distance. That is the net force is sensitive to the bulk of the system and the contribution of closest neighbours is very small. Hence, the force can be accurately estimated by the **mean-field model** in which the mass of each star is spread through a sphere comparable in size to the inter-star distance; i.e., we compute the potential $\Phi(\mathbf{x})$ and force $-m\nabla\Phi$ from a smooth mass distribution $\rho(\mathbf{x})$.

Our zeroth-order approximation to the motion of a star is the trajectory that follows from its initial $(\mathbf{x}, \mathbf{v})$ and $\Phi(\mathbf{x})$.

1.1 The virial theorem

We now prove a very useful result that follows from the scale-free nature of the gravitational interaction. Suppose we differentiate a kind of moment of inertial $I = \sum_\alpha m_\alpha |\mathbf{x}_\alpha|^2$:

$$\frac{\mathrm{d}^2 I}{\mathrm{d}t^2} = 2\sum_\alpha m_\alpha \left(\mathbf{x}_\alpha \cdot \ddot{\mathbf{x}}_\alpha + |\dot{\mathbf{x}}_\alpha|^2\right). \tag{1.1}$$

Now

$$\ddot{\mathbf{x}}_\alpha = G\sum_\beta \frac{m_\alpha m_\beta}{|\mathbf{x}_\alpha - \mathbf{x}_\beta|^3}(\mathbf{x}_\beta - \mathbf{x}_\alpha) \tag{1.2}$$

so

$$\sum_\alpha \mathbf{x}_\alpha \cdot \ddot{\mathbf{x}}_\alpha = G\sum_{\alpha\beta} \frac{m_\alpha m_\beta}{|\mathbf{x}_\alpha - \mathbf{x}_\beta|^3}\mathbf{x}_\alpha \cdot (\mathbf{x}_\beta - \mathbf{x}_\alpha). \tag{1.3}$$

Interchanging α, β on the right side and adding to the original equation, we find

$$2\sum_\alpha \mathbf{x}_\alpha \cdot \ddot{\mathbf{x}}_\alpha = G\sum_{\alpha\beta} \frac{m_\alpha m_\beta}{|\mathbf{x}_\alpha - \mathbf{x}_\beta|^3}(\mathbf{x}_\alpha - \mathbf{x}_\beta)\cdot(\mathbf{x}_\beta - \mathbf{x}_\alpha) = -G\sum_{\alpha\beta} \frac{m_\alpha m_\beta}{|\mathbf{x}_\alpha - \mathbf{x}_\beta|} = 2W, \tag{1.4}$$

where W is the system's potential energy. Inserting this into (1.1)

$$\frac{\mathrm{d}^2 I}{\mathrm{d}t^2} = 2W + 4K, \tag{1.5}$$

where K is the system's kinetic energy. If the system is statistically in a steady state, $I =$ const to within Poisson noise, and we have the **virial theorem**

$$2K + W = 0. \tag{1.6}$$

2 CBE, Jeans thm, constants of motion

By Liouville's thm, in this approximation the probability density of stars $f(\mathbf{x}, \mathbf{v})$ obeys the Collisionless Boltzmann Equation

$$\begin{aligned}\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \dot{\mathbf{q}} \cdot \frac{\partial f}{\partial \mathbf{q}} + \dot{\mathbf{p}} \cdot \frac{\partial f}{\partial \mathbf{p}} &= 0 \\ \frac{\partial f}{\partial t} + \frac{\partial H}{\partial \mathbf{p}} \cdot \frac{\partial f}{\partial \mathbf{q}} - \frac{\partial H}{\partial \mathbf{q}} \cdot \frac{\partial f}{\partial \mathbf{p}} &= 0 \\ \frac{\partial f}{\partial t} &= \{H, f\}\end{aligned} \tag{2.1}$$

The CBE states that the distribution function (DF) f of an equilibrium system is a constant of motion. So any non-negative function $f(I_1, \ldots, I_n)$ of constants of motion with $\int \mathrm{d}^3\mathbf{x}\mathrm{d}^3\mathbf{v}\, f < \infty$ provides a model of an equilibrium system (**Jeans' thm**).

2.1 Angle-action variables

In a (homogeneous) plasma we have $\mathbf{v} = \text{const}$, so $f(\mathbf{v})$ solves CBE. In a stellar system $\mathbf{v} \neq \text{const}$ but we can write $f(H)$ or if spherical $f(H, \mathbf{L})$ or if axisymmetric $f(H, L_z)$.

But in axisymmetric $\Phi(R, z)$, $f(H, L_z)$ is not generic. Numerical integration of orbits shows they are **quasi-perioidic**; this means that the coordinates (and momenta) along an orbit can be expanded as

$$\mathbf{x}(t) = \sum_{\mathbf{n}} \mathbf{X_n} \mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\boldsymbol{\Omega}t}$$

where $\mathbf{n} = (n_1, n_2, n_3)$ has integer components and $\boldsymbol{\Omega} = (\Omega_1, \Omega_2, \Omega_3)$ is a triple of **characteristic frequencies**. (If the orbit were periodic, the Fourier expansion would contain only one characteristic frequency; it's quasiperiodic because there are more frequencies than arguments, here t.) The quasi-periodic nature of orbits implies (see Arnold's book) $\exists$ 3 constants of motion I_i.

Any $F(I_1, I_2, I_3)$ is a const of motion, so we have infinite freedom in what we use for arguments of the DF. By far the best choice is a set (J_1, J_2, J_3) st $\exists$ canonoically conjugate variables θ_i. Then $\{\theta_i, J_j\} = \delta_{ij}$, etc. Because $0 = \dot{\mathbf{J}} = \{\mathbf{J}, H\} = -\partial H/\partial\boldsymbol{\theta}$, we have $H(\mathbf{J})$. Also

$$\dot{\boldsymbol{\theta}} = \frac{\partial H}{\partial \mathbf{J}} = \boldsymbol{\Omega}(\mathbf{J}) = \text{const}, \tag{2.2}$$

so

$$\boldsymbol{\theta}(t) = \boldsymbol{\theta}(0) + \boldsymbol{\Omega}t, \tag{2.3}$$

i.e., the **angle variables** θ_i increase linearly with time. We scale the **action integrals** J_i such that $\mathbf{x}, \mathbf{v}$) along an orbit become periodic with period 2π:

$$x(\boldsymbol{\theta} + 2\pi\mathbf{m}, \mathbf{J}) = x(\boldsymbol{\theta}, \mathbf{J}) \text{ with integer } m_i. \tag{2.4}$$

The **action integrals** J_i are defined up to a set of discrete canonical transformations (generating function $S(\boldsymbol{\theta}, \mathbf{J}') = \boldsymbol{\theta} \cdot \mathbf{M} \cdot \mathbf{J}'$ where the matrix $\mathbf{M}$ has integer elements). For an axisymmetric system the actions are uniquely defined by requiring that

$$\begin{aligned}&J_r \text{ quantifies radial excursions} \\ &J_\phi = L_z \\ &J_z \text{ quantifies oscilations perpendicular to the equatorial plane.}\end{aligned} \tag{2.5}$$

In the spherical limit $J_z = L - |L_z|$ is the angular momentum in the (x, y) plane.

A generic equilibrium system has $f(\mathbf{J})$.

3 Encounters

We can consider that the difference $\delta\Phi$ between the real Φ and the mean-field Φ perturbs the motion of stars in the mean-field Φ. $\delta\Phi$ comprises a series of Kepler potentials minus the potential obtained by smoothing out a particle's mass. Surely the dominant partner is the Kepler potential, which diverges at the location of the scattering particle. So we compute an encounter between two stars, a scatterer of mass m and a scattered star of mass M. The encounter is completely described by the steady motion of the system's centre of mass and the scattering off a fixed Kepler potential $-G(M+m)/r$ of the **reduced particle**, which has mass

$$\mu = \frac{mM}{m+M} \text{ and velocity } \mathbf{V} = \mathbf{v}_M - \mathbf{v}_m. \tag{3.1}$$

Since the centre-of-mass velocity is unchanged by the encounter

$$m\Delta\mathbf{v}_m + M\Delta\mathbf{v}_M = 0 \tag{3.2}$$

and a little algebra yields

$$\Delta\mathbf{v}_M = \frac{m}{m+M}\Delta\mathbf{V}. \tag{3.3}$$

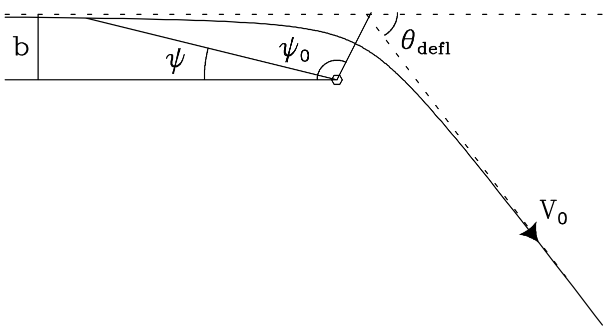

Fig. 2. Definition of the angles describing scattering of the reduced particle

If the encounter has impact parameter b and the reduced particle's initial velocity is $\mathbf{V}_0$, its angular-momentum per unit mass is $L = bV_0$. Then its polar coordinates (r, ψ) in the invariant plane satisfy (BT08 3.24)

$$\frac{1}{r} = C\cos(\psi - \psi_0) + \frac{G(M+m)}{b^2V_0^2}, \tag{3.4}$$

where C and ψ_0 are constants determined by the initial conditions. Differentiating (3.4) and using $r^2\dot{\psi} = L = bV_0$

$$\frac{\mathrm{d}r}{\mathrm{d}t} = Cr^2\dot{\psi}\sin(\psi - \psi_0) = CbV_0\sin(\psi - \psi_0). \tag{3.5}$$

We take the origin of ψ to be the direction to the particle at $t = -\infty$ and then evaluating (3.5) at $t = -\infty$

$$-V_0 = CbV_0\sin(-\psi_0) \tag{3.6}$$

Evaluating (3.4) at $t = -\infty$ gives

$$0 = C\cos\psi_0 + \frac{G(M+m)}{b^2V_0^2}, \tag{3.7}$$

so

$$\tan\psi_0 = -\frac{bV_0^2}{G(M+m)}. \tag{3.8}$$

The point of closest approach is $\psi = \psi_0$ and the orbit is symmetrical about this point, so the deflection angle is $\theta_{\rm defl} = 2\psi_0 - \pi$. By conservation of energy the particle's speed returns to V_0 as $t \to +\infty$. It follows that the changes in $\mathbf{V}$ parallel and perpendicular to the original velocity are

$$\begin{aligned}
|\Delta\mathbf{V}_\perp| &= V_0 \sin\theta_{\rm defl} = V_0|\sin 2\psi_0| = \frac{2V_0|\tan\psi_0|}{1+\tan^2\psi_0} \\
&= \frac{2bV_0^3}{G(M+m)}\left(1+\frac{b^2V_0^4}{G^2(M+m)^2}\right)^{-1} \\
|\Delta\mathbf{V}_\parallel| &= V_0(1-\cos\theta_{\rm defl}) = V_0(1+\cos 2\psi_0) = \frac{2V_0}{1+\tan^2\psi_0} \\
&= 2V_0\left(1+\frac{b^2V_0^4}{G^2(M+m)^2}\right)^{-1}.
\end{aligned}$$

Using (3.3) to extract the components of $\Delta\mathbf{v}_M$

$$\begin{aligned}
|\Delta\mathbf{v}_{M\perp}| &= \frac{2mbV_0^3}{G(M+m)^2}\left(1+\frac{b^2V_0^4}{G^2(M+m)^2}\right)^{-1} \\
|\Delta\mathbf{v}_{M\parallel}| &= \frac{2mV_0}{M+m}\left(1+\frac{b^2V_0^4}{G^2(M+m)^2}\right)^{-1}.
\end{aligned}$$

We have now to add the velocity changes caused by a succession of encounters as M moves through a sea of stars m and common velocity $\mathbf{v}_m$. Each encounter produces a $\Delta\mathbf{v}_{M\perp}$ in a different direction, and in a uniform sea these will sum to zero. But since $|\Delta\mathbf{v}_{M\parallel}|$ is always directed opposite to $\mathbf{V}_0$ these contributions accumulate as a net drag on M.

3.1 Dynamical friction

Let's focus on the fraction of the field stars that have velocities within $\mathrm{d}^3\mathbf{v}_m$ of $\mathbf{v}_m$. The spatial density of such stars is $f(\mathbf{v})\mathrm{d}^3\mathbf{v}_m$ so the rate of encounters with them at impact parameter b is

$$2\pi b\,\mathrm{d}b\,V_0\,f(\mathbf{v}_m)\mathrm{d}^3\mathbf{v}_m$$

so they change $\mathbf{v}_M$ at a rate

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{v}_M}{\mathrm{d}t} &= \mathbf{V}_0 f(\mathbf{v}_m)\mathrm{d}^3\mathbf{v}_m 2\pi\int_0^{b_{\rm max}}\mathrm{d}b\,b\frac{2mV_0}{M+m}\left(1+\frac{b^2V_0^4}{G^2(M+m)^2}\right)^{-1} \\
&= 2\pi\ln(1+\Lambda^2)G^2m(M+m)f(\mathbf{v}_m)\mathrm{d}^3\mathbf{v}_m\frac{\mathbf{v}_m-\mathbf{v}_M}{|\mathbf{v}_m-\mathbf{v}_M|^3}
\end{aligned}\tag{3.9a}$$

where

$$\Lambda \equiv \frac{b_{\rm max}V_0^2}{G(M+m)}.\tag{3.9b}$$

Typically Λ is large > 1000 and its logarithm is insensitive to V_0. If we approximate $\ln(1+\Lambda^2) \simeq 2\ln\Lambda$ by a constant, the **Coulomb logarithm**, equation (3.9) states that each point in velocity space contributes to the drag on M as if the density $f(\mathbf{v}_m)$ in velocity space were a mass density in real space that is attracting gravitationally. It follows that if the velocity distribution is isotropic (i.e., $f(\mathbf{v}_m) = f(|\mathbf{v}_m|)$), the net drag comes entirely from stars that move more slowly than M. In fact it's easy to show that in this case when we sum over all velocities $\mathbf{v}_m$ the drag is

$$\frac{\mathrm{d}\mathbf{v}_M}{\mathrm{d}t} = -16\pi^2\ln\Lambda G^2m(M+m)\int_0^{v_M}\mathrm{d}v_m\,v_m^2 f(v_m)\frac{\mathbf{v}_M}{v_M^3}.\tag{3.10}$$

This is **Chandrasekhar's dynamical friction formula**. We'll discuss it in the case that $M \gg m$.

Then the *deceleration* of M is proportional to M, so the force on M is proportional to M^2. the physical origin of the deceleration is the gravitational pull on M from the region of enhanced density of field stars that it creates behind it, as a ship leaves a wake on the ocean. The mass of the wake is proportional to M so the force it imposes on M is proportional to M^2

The deceleration is also proportional to the density nm of the stars being scattered but independent of the masses of individual stars.

3.2 Stochastic acceleration

Equation (3.10) states that M experiences a frictional drag even when $M = m$ and we are in fact dealing with a field star. This observation suggests the conclusion that the whole population of field stars will slow down until they are at rest. This conclusion is absurd because it would violate conservation of energy, and is inconsistent with our expectation that the field-star population should have a Maxwellian distribution as a steady state. The reason the stars do not slow to rest is that in thermal equilibrium balance is achieved between dynamical friction and stochastic acceleration associated with the random jolts $\Delta\mathbf{v}_{M\perp}$. The jolts drive a random walk in velocity space, and acting alone they would cause the whole population to diffuse off to infinity. The job of dynamical friction is to pull them back to $v_m = 0$. Thermal equilibrium is characterised by balance between outward diffusion and inward drag.

We could simply add in quadrature the contributions $\Delta\mathbf{v}_{M\perp}$ given by equation (3.9) but it is easier and more instructive to resort to a crude computation. We compute the impulse the reduced particle would receive if it sped past the origin at speed V_0.

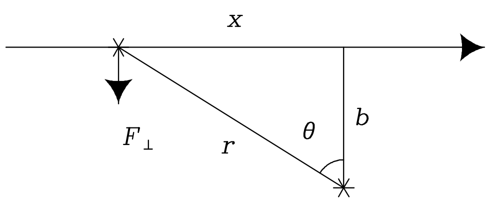

Fig. 3. Computing reduced-particle scattering in the impulse approximation

$$\Delta\mathbf{V}_\perp = \int_{-\infty}^{\infty} \mathrm{d}t\, \mathbf{g}_\perp = \int \mathrm{d}t \frac{G(M+m)}{r^2}\frac{b}{r} \simeq bG(M+m)\int_{-\infty}^{\infty} \mathrm{d}t \frac{1}{(b^2+V_0^2t^2)^{3/2}} = \frac{2G(M+m)}{bV_0} \tag{3.11}$$

For $bV_0/G(M+m) \gg 1$ this is in perfect agreement with our exact formula (3.9). However, as $b \to 0$ it diverges whereas the exact formula tends to zero. So we should not use our approximate formula for b smaller than

$$b_{90} \equiv \frac{G(M+m)}{V_0^2}, \tag{3.12}$$

which corresponds to $\psi_0 = 3\pi/4$ and $\theta_{\mathrm{defl}} = \pi/2$. Note that $\ln(b_{\max}/b_{90}) \simeq \ln\Lambda$.

For forward scattering $\Delta\mathbf{v}_{M\perp} > \Delta\mathbf{v}_{M\parallel}$ and we know that scattering at large b through small angles dominates. So we sum in quadrature $\Delta\mathbf{v}_{M\perp} \simeq 2Gm/bV_0$ [eq. (3.3)], and by analogy with eq. (3.9) find

$$\frac{\mathrm{d}\mathbf{v}_M^2}{\mathrm{d}t} = V_0 f(\mathbf{v}_m)\mathrm{d}^3\mathbf{v}_m 2\pi\int_{b_{90}}^{b_{\max}} \mathrm{d}b\, b\frac{4G^2m^2}{b^2V_0^4} = 8\pi G^2m^2\frac{f(\mathbf{v}_m)\mathrm{d}^3\mathbf{v}_m}{|\mathbf{v}_m - \mathbf{v}_M|}\ln(b_{\max}/b_{90}) \tag{3.13}$$

Now specialising to an isotropic $f(v_m)$ we integrate over all velocities

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{v}_M^2}{\mathrm{d}t} &= 8\pi G^2 m^2 \ln\Lambda \int \mathrm{d}v_m\, v_m^2 f(v_m) \int \mathrm{d}\theta\, \frac{\sin\theta}{\sqrt{v_m^2+v_M^2-2v_m v_M\cos\theta}} \int \mathrm{d}\phi \\
&= 16\pi^2 G^2 m^2 \ln\Lambda \int \mathrm{d}v_m\, v_m^2 f(v_m) \left[\frac{\sqrt{v_m^2+v_M^2-2v_m v_M\cos\theta}}{2v_m v_M}\right]_0^\pi \\
&= 16\pi^2 G^2 m^2 \ln\Lambda \int \mathrm{d}v_m\, v_m^2 f(v_m) \frac{1}{\max(v_m, v_M)} \\
&= 16\pi^2 G^2 m^2 \ln\Lambda \left(\frac{1}{v_M}\int_0^{v_M} \mathrm{d}v_m\, v_m^2 f(v_m) + \int_{v_M}^\infty \mathrm{d}v_m\, v_m f(v_m)\right).
\end{aligned} \tag{3.14}$$

Let's apply this formula in the case $M = m$ so we are actually studying the acceleration of a typical field star. Then a typical value of v_M is σ, the velocity dispersion of the field stars, and the big bracket is of order $n/8\pi\sigma$ (we know that $\int_0^\infty \mathrm{d}v\, v^2 f(v) = n/4\pi$), where $n = \rho/m$ is the number density of field stars, so the time for $\mathbf{v}_M$ to diffuse to a different part of velocity space is

$$t_{\mathrm{relax}} = \frac{v_M^2}{\mathrm{d}\mathbf{v}_M^2/\mathrm{d}t} \simeq \frac{\sigma^3}{2\pi G^2 m\rho \ln\Lambda}. \tag{3.15}$$

To understand the structure of this formula, note that $\sigma^2 \sim G\mathcal{M}/R$, where $\mathcal{M}$ and R are the total mass and linear scale of the system, and $(G\rho)^{-1/2} = t_{\mathrm{dyn}}$ is its dynamical time. So

$$t_{\mathrm{relax}} \sim 0.1\frac{\mathcal{M}}{m}\frac{\sigma}{R}t_{\mathrm{dyn}}^2 \sim 0.1 N t_{\mathrm{dyn}} \tag{3.16}$$

where $N \equiv \mathcal{M}/m$ is the number of particles in the system.

3.3 Encounters at large impact parameters

The Coulomb logarithm $\ln\Lambda$ is a relic of our finding that in a strictly homogeneous star field, the integrals for the drag and stochastic acceleration effects would diverge as we pushed to large b. Real systems are inhomogeneous, so the integrals terminate naturally with b of order the system size, but the conclusion is inescapable that the dominant fluctuations are associated with encounters at the largest conceivable impact parameters. This conclusion is worrying, but a back-of-envelope calculation explains what's really happening.

In any volume $\mathcal{V}$, which in the mean contains N stars of mass m and thus a total mass $\mathcal{M} = Nm$ the actual mass fluctuates around this mean by $\sqrt{N}m = \sqrt{m\mathcal{M}}$. Consequently the gravitational acceleration at some point due to $\mathcal{V}$ fluctuates by $\delta g \sim G\sqrt{m\mathcal{M}}/R^2$, where R is the characteristic system size. If the force is anomalously high, it will remain so for a time $t \sim \mathcal{V}^{1/3}/\sigma \sim (\mathcal{M}/M)^{1/3}R/\sigma$, where M and σ are the system's total mass and random velocity. Multiplying the magnitude of the anomalous acceleration by its duration we obtain an estimate of the contribution from $\mathcal{V}$ to the fluctuating velocity of our particle

$$\delta v \sim \delta g\, t \sim \frac{G\sqrt{m\mathcal{M}}}{R^2}\frac{(\mathcal{M}/M)^{1/3}R}{\sigma} \sim \frac{G\sqrt{m\mathcal{M}}}{R}\frac{(\mathcal{M}/M)^{1/3}}{\sqrt{GM/R}} \sim \sqrt{\frac{Gm}{R}}(\mathcal{M}/M)^{5/6}. \tag{3.17}$$

Thus Poisson fluctuations in the content of the largest sub-volumes make the largest contributions to the jitter in $\mathbf{v}$. Of course there are more small sub-volumes than large ones, and their fluctuations refresh more rapidly than those of the large sub-volumes, but these effects don't effectively compensate for the factor $(\mathcal{M}/M)^{5/6}$ in (3.17) by which a bigger volume has a bigger impact. We conclude that the logarithmic divergences we encountered signal that fluctuations in the gravitational potential are dominated by Poisson noise in the number of particles on one side of the system rather than another. We have idealised the fluctuations as arising from many simultaneous encounters, but this idealisation is not well-founded. It would be sounder to work in terms of thermally excited normal modes of oscillation that have a characteristic scale that is comparable to the system's size.

In an ideal gas the number of molecules in a given volume experiences Poisson fluctuations as was assumed above, and these fluctuations can be considered to arise from thermally excited sound waves. The self-gravity of a stellar system makes the system more compressible on large scales than an ideal gas, with the consequence that the fluctuations should have a larger amplitude than simple Poisson fluctuations.

3.4 Equipartition

For M/m large, dynamical friction is proportional to mM while stochastic acceleration is proportional to m^2. Hence the speed at which the two effects come into balance decreases with increasing M:

$$\frac{\mathrm{d}\mathbf{v}_M^2/\mathrm{d}t}{|\mathrm{d}\mathbf{v}_M/\mathrm{d}t|} = \frac{m}{M+m} \frac{\frac{1}{v_M}\int_0^{v_M} \mathrm{d}v_m\, v_m^2 f(v_m) + \int_{v_M}^{\infty} \mathrm{d}v_m\, v_m f(v_m)}{\frac{1}{v_M^2}\int_0^{v_M} \mathrm{d}v_m\, v_m^2 f(v_m)}. \tag{3.18}$$

The denominator grows steadily with v_M while the numerator falls very slowly with v_M, so the second fraction declines with increasing v_M. Thus balance is achieved in the sense that the second fraction is $\sim (M+m)/m$ at a speed that declines as M increases.

4 Fokker-Planck equation

Let $P(\mathbf{w}, \boldsymbol{\Delta})$ be the probability per unit time that a star with phase-space location $(\mathbf{x}, \mathbf{v})$ suffers an encounter that increments its velocity by $\boldsymbol{\Delta}$. Then we can write the **master equation**

$$f(\mathbf{v}, t+\delta t) = f(\mathbf{v}, t) - \delta t \int \mathrm{d}^3\boldsymbol{\Delta}\, f(\mathbf{v}, t) P(\mathbf{v}, \boldsymbol{\Delta}) + \delta t \int \mathrm{d}^3\boldsymbol{\Delta}\, f(\mathbf{v}-\boldsymbol{\Delta}, t) P(\mathbf{v}-\boldsymbol{\Delta}, \boldsymbol{\Delta}), \tag{4.1}$$

where for brevity we have suppressed dependencies on $\mathbf{x}$. The first integral counts stars that are scattered away from $\mathbf{v}$ to some other velocity in the interval δt, while the second integral counts stars that are scattered to $\mathbf{v}$. We have seen that the dominant encounters are distant ones that cause small deflections, so we may argue that $P(\mathbf{v}, \boldsymbol{\Delta})$ is non-negligible only for small $\boldsymbol{\Delta}$. By contrast, neither $f(\mathbf{v}, t)$ nor $P(\mathbf{v}, \boldsymbol{\Delta})$ is a rapidly-varying function of their first argument. This being so, we may Taylor expand the product $f(\mathbf{v}-\boldsymbol{\Delta}, t)P(\mathbf{v}-\boldsymbol{\Delta}, \boldsymbol{\Delta})$ in its first argument and retain only three terms

$$f(\mathbf{v}-\boldsymbol{\Delta}, t)P(\mathbf{v}-\boldsymbol{\Delta}, \boldsymbol{\Delta}) \simeq f(\mathbf{v}, t)P(\mathbf{v}, \boldsymbol{\Delta}) - \boldsymbol{\Delta}\cdot\frac{\partial}{\partial \mathbf{v}}[f(\mathbf{v}, t)P(\mathbf{v}, \boldsymbol{\Delta})] + \tfrac{1}{2}\Delta_i\Delta_j \frac{\partial^2}{\partial v_i \partial v_j}[f(\mathbf{v}, t)P(\mathbf{v}, \boldsymbol{\Delta})] \tag{4.2}$$

When we substitute this approximation into the second integral of the master equation, we obtain

$$\frac{\delta f}{\delta t} = -\int \mathrm{d}^3\boldsymbol{\Delta}\, \boldsymbol{\Delta}\cdot\frac{\partial}{\partial \mathbf{v}}[f(\mathbf{v}, t)P(\mathbf{v}, \boldsymbol{\Delta})] + \tfrac{1}{2}\int \mathrm{d}^3\boldsymbol{\Delta}\, \Delta_i\Delta_j \frac{\partial^2}{\partial v_i \partial v_j}[f(\mathbf{v}, t)P(\mathbf{v}, \boldsymbol{\Delta})]. \tag{4.3}$$

The derivatives can be taken out of the integrals, and then we have

$$\frac{\mathrm{d}f}{\mathrm{d}t} = -\frac{\partial}{\partial \mathbf{v}}\cdot[\mathbf{D}(\mathbf{v})f(\mathbf{v}, t)] + \tfrac{1}{2}\frac{\partial^2}{\partial v_i \partial v_j}[D_{ij}(\mathbf{v})f(\mathbf{v}, t)], \tag{4.4a}$$

where the first and second-order **diffusion coefficients** are

$$\begin{aligned} D_i(\mathbf{v}) &\equiv \int \mathrm{d}^3\boldsymbol{\Delta}\, \Delta_i P(\mathbf{v}, \boldsymbol{\Delta}), \\ D_{ij}(\mathbf{v}) &\equiv \int \mathrm{d}^3\boldsymbol{\Delta}\, \Delta_i\Delta_j P(\mathbf{v}, \boldsymbol{\Delta}). \end{aligned} \tag{4.4b}$$

The first-order coefficient $\mathbf{D}$ simply describes dynamical friction: in an isotropic system, $\mathbf{D} \propto -\mathbf{v}$ and has a magnitude that is readily deduced from Chandrasekhar's friction formula (3.10). The second-order coefficient D_{ij} is simply an elaboration of the stochastic heating term (3.14). The **Fokker-Planck equation** (4.4a)

4.1 Orbit-averaged Fokker-Planck equation

The derivation of the Fokker-Planck equation above is based on the conceit that particles suffer changes of velocity while at a particular location $\mathbf{x}$. But since distant encounters are the strongest drivers of deviation from mean-field orbits, this conceit is indefensible. It is better to start from the conception that at any time the system is in statistical equilibrium, so by the Jeans theorem its distribution function is $f(\mathbf{J})$. Encounters cause f to evolve according to the **orbit-averaged Fokker-Planck equation**

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial \mathbf{J}} \cdot [\mathbf{D}(\mathbf{J}) f(\mathbf{J}, t)] + \tfrac{1}{2} \frac{\partial^2}{\partial J_i \partial J_j} [D_{ij}(\mathbf{J}) f(\mathbf{J}, t)], \tag{4.5a}$$

where

$$\begin{aligned} D_i(\mathbf{J}) &\equiv \int \mathrm{d}^3\boldsymbol{\Delta}\, \Delta_i P(\mathbf{J}, \boldsymbol{\Delta}), \\ D_{ij}(\mathbf{v}) &\equiv \int \mathrm{d}^3\boldsymbol{\Delta}\, \Delta_i \Delta_j P(\mathbf{J}, \boldsymbol{\Delta}). \end{aligned} \tag{4.5b}$$

The derivation of this equation precisely parallels the derivation of (4.4a) with $\mathbf{v} \to \mathbf{J}$ and the suppression of the conjugate variable, now $\boldsymbol{\theta}$ unnecessary because f has no dependence on $\boldsymbol{\theta}$. $P(\mathbf{J}, \boldsymbol{\Delta})$ is now the probability per unit time that a star on the orbit $\mathbf{J}$ suffers a perturbation that increments its action integrals by $\boldsymbol{\Delta}$.

Note that the Fokker-Planck equation can be written

$$\frac{\partial f}{\partial t} = -\mathrm{div}\, \mathbf{F}, \tag{4.6a}$$

where the diffusive flux is

$$F_i \equiv D_i f - \tfrac{1}{2} \frac{\partial}{\partial J_j} [D_{ij} f]. \tag{4.6b}$$

This form of the equation manifests the conservation of stars as they diffuse. The second term usually produces a flux that is in the opposite direction to the gradient of f, just as the flux of heat in a bar is in the opposite direction to the temperature gradient. The vector $\mathbf{D}$ is usually directed towards the origin of action space.

The orbit-averaged diffusion equation was criticised by Hénon on the grounds that it does not encompass the important process of evaporation of stars from a cluster (BT08 §7.5.2). This is because no finite action corresponds to an unbound orbit.

5 Binary stars

To this point our strategy has been to ignore correlations between stars beyond those implied by the system's large-scale inhomogeneity – we have characterised the system by its one-particle distribution function. However, a significant fraction of stars are members of a binary system and binary stars can have a big impact on the long-term evolution of star clusters.

When a field star has a close encounter with a binary star, energy is exchanged between the binary's internal energy $E_b < 0$ and the translational KE of the field and binary stars. If $|E_b|$ is decreased too much, the binary is disrupted (ionised). If $|E_b|$ is increased, the cluster is heated. The field star may exchange places with one of the binary stars.

5.1 Soft binaries

If $E_b| < \frac{1}{2} m\sigma^2$ the binary is **soft**. The internal speed of the binary is slower than the random motion in the cluster. On average, interactions with field stars make soft binaries softer, until they are disrupted.

5.2 Hard binaries

If $|E_b| > \frac{1}{2}m\sigma_2$ the binary is **hard**. When a field star comes close to a hard binary, an unstable triple star is formed. Eventually one of the three stars is ejected with a velocity comparable to the original binary velocity, so faster than σ. By conservation of energy the final binary is harder than the original binary: **Heggie's law**: hard binaries become harder, soft binaries softer.

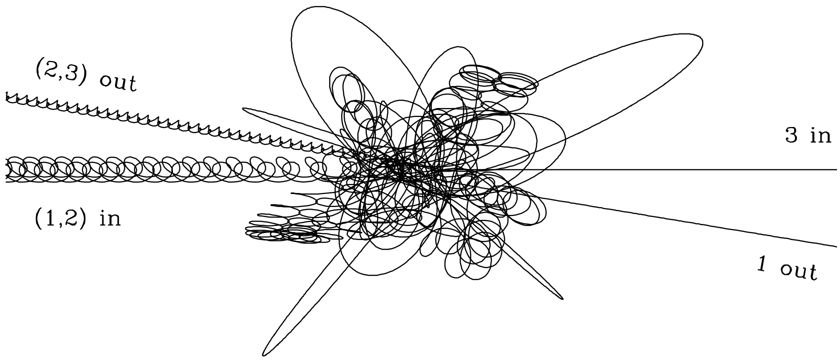

Fig. 4. A close encounter between a binary star and a single star

The bottom line: hard binaries are an energy source for a cluster just as nuclear fusion is an energy source for stars. Soft binaries come and go without significant impact.

5.2.1 Formation of hard binaries A binary can form in a cluster of point masses only through the interaction of 3 bodies: one is there to carry away the energy released as the other two form a binary. If a hard binary is to form the relative velocity of the future binary members has to change by of order itself, so the impact parameter has to be not much larger than b_{90}, which yields a $\pi/2$ deflection. The characteristic time t_{90} interval between a given star experiencing a large deflection is given by $nb_{90}^2vt_{90} \simeq 1$, and the probability that when this encounter occurs there is a third particle within b_{90} is $P_3 \simeq nb_{90}^3$. Hence the time required for a given star to have a non-negligible probability of entering a tight binary is

$$t_2 \simeq \frac{t_{90}}{P_3} \simeq \frac{1}{n^2 b_{90}^5 v} \simeq \frac{1}{n^2 v}\left(\frac{v^2}{Gm}\right)^5 \tag{5.1}$$

But by the virial thm (1.6) $r_0 \sim GNm/v^2$ so with (3.15)

$$\frac{t_2}{t_{\rm relax}} \simeq \frac{1}{n^2 v}\left(\frac{v^2}{Gm}\right)^5 \frac{2\pi G^2 m^2 n \ln\Lambda}{v^3} \simeq \frac{1}{nr_0^3}N^3 2\pi\ln\Lambda \simeq N^2 2\pi \ln\Lambda. \tag{5.2}$$

The time required for one of the N stars in the core to get into a binary is smaller than t_2 by a factor N and the first tight binary will form after

$$t_{\rm fst} \sim t_2/N \simeq 2\pi N \ln\Lambda t_{\rm relax}. \tag{5.3}$$

6 Thermal equilibrium?

6.1 Negative specific heat & core collapse

The virial thm (1.6) requires the potential energy W and the kinetic energy K to satisfy $W = -2K$. Also $E = K + W$ so $E = -K$. If follows that $\delta K/\delta E = -1$. But $\delta K \sim \delta T$, temperature, and $\delta E \sim \delta Q$, heat. So by the virial thm a stellar system has negative specific heat $C = \delta Q/\delta T$.

Let's put our stellar system in a sphere & drop it into a heat bath.

If $T_{\rm bath} < T_{\rm star}$, heat will flow out of system into bath. $T_{\rm star}$ will rise accelerating heat loss: runaway (Lynden-Bell & Wood, 38, 495, 1968)

A star cluster can be conceptually divided into self-gravitating body and a low-mass envelope trapped in the body's potential well. Envelope = bath, body = system and we expect **core collapse**

6.2 Entropy

The Shannon entropy of a system's probability distribution $p(\mathbf{w}_1, \ldots, \mathbf{w}_N) \equiv p(\tau)$ is $S = -\int \mathrm{d}^{6N}\tau\, p \ln p$. In so far as correlations between stars can be neglected, $p(\tau) = f(\mathbf{w}_1)f(\mathbf{w}_2)\cdots$, and $S = -N\int \mathrm{d}^6\mathbf{w}\, f \ln f$. Divide system into body mass M_1 and halo mass $M_2 \ll M_1$ with characteristic radius r_2. The halo moves in the gravitational field of the body, so its velocity dispersion satisfies $\sigma_2^2 \sim GM_1/r_2$, and most of its stars are located where

$$f_2 \sim \frac{1}{r_2^3\sigma_2^3} \sim \frac{1}{(GM_1 r_2)^{3/2}} \tag{6.1}$$

Now $|E_2| \sim GM_2M_1/r_2$ so

$$S_2 \sim \tfrac{3}{2}N_2 \ln(GM_1r_2) \sim -\tfrac{3}{2}N_2 \ln|E_2| + \text{const} \tag{6.2}$$

By shrinking the body slightly we can obtain energy and feed it to the halo and thus make $|E_2|$ as small as we please, driving $S_2 \to \infty$. Hence there is no upper bound on the system's entropy.

In ordinary statistical mechanics states of thermal equilibrium are states of maximum entropy given E and M. We have shown that these states do not exist for a stellar system, so such systems are incapable of achieving thermal equilibrium. They must be constantly increasing their entropy by moving to states of higher central concentration. This is the physical principle that drives the evolution of both stars and galaxies.

6.3 Evaporation

Encounters in a stellar system will drive $n(\mathbf{v}) = f(\mathbf{x},\mathbf{v})$ towards a Maxwellian. But stars that reach $v > v_{\rm esc}(\mathbf{x}) = \sqrt{2\Phi(\mathbf{x})}$ will escape. By the virial thm (1.6)

$$\tfrac{1}{2}M\overline{v_{\rm esc}}^2 \equiv -\int \mathrm{d}^6\mathbf{w}\, f\Phi = -\int \mathrm{d}^3\mathbf{x}\rho\Phi == -2\mathrm{PE} = 2\int \mathrm{d}^6\mathbf{w}\, fv^2 \equiv 2M\overline{\sigma}^2 \tag{6.3}$$

so

$$\overline{v_{\rm esc}} = 2\overline{\sigma} \tag{6.4}$$

In a Maxwellian a fraction $\sim 1/135$ of stars have speeds greater than 2σ. Hence the system will lose 1/135 of its mass each relaxation time by **evaporation** of stars

7 What determines dynamical equilibrium?

7.1 Violent relaxation

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(\tfrac{1}{2}v^2 + \Phi) = \mathbf{v}\cdot\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} + \frac{\partial\Phi}{\partial t} + \mathbf{v}\cdot\nabla\Phi = \frac{\partial\Phi}{\partial t}. \tag{7.1}$$

Thus in a collisionless system the energies of stars can change only in so far as the mean gravitational field is fluctuating. If the system starts far from dynamical equilibrium, for a few dynamical time it experiences $\sim$ critically damped oscillation, during which time Φ is strongly time-dependent and energy is redistributed between stars (**violent relaxation**. This redistribution is monopoly-capitalistic: the rich get richer and the poor poorer. The gravitational energy released in the initial collapse is first fed into large-scale modes of oscillation, and these rapidly Landau damp and in this way increase the kinetic energy of the particles to the value required by the virial theorem for an equilibrium configuration.

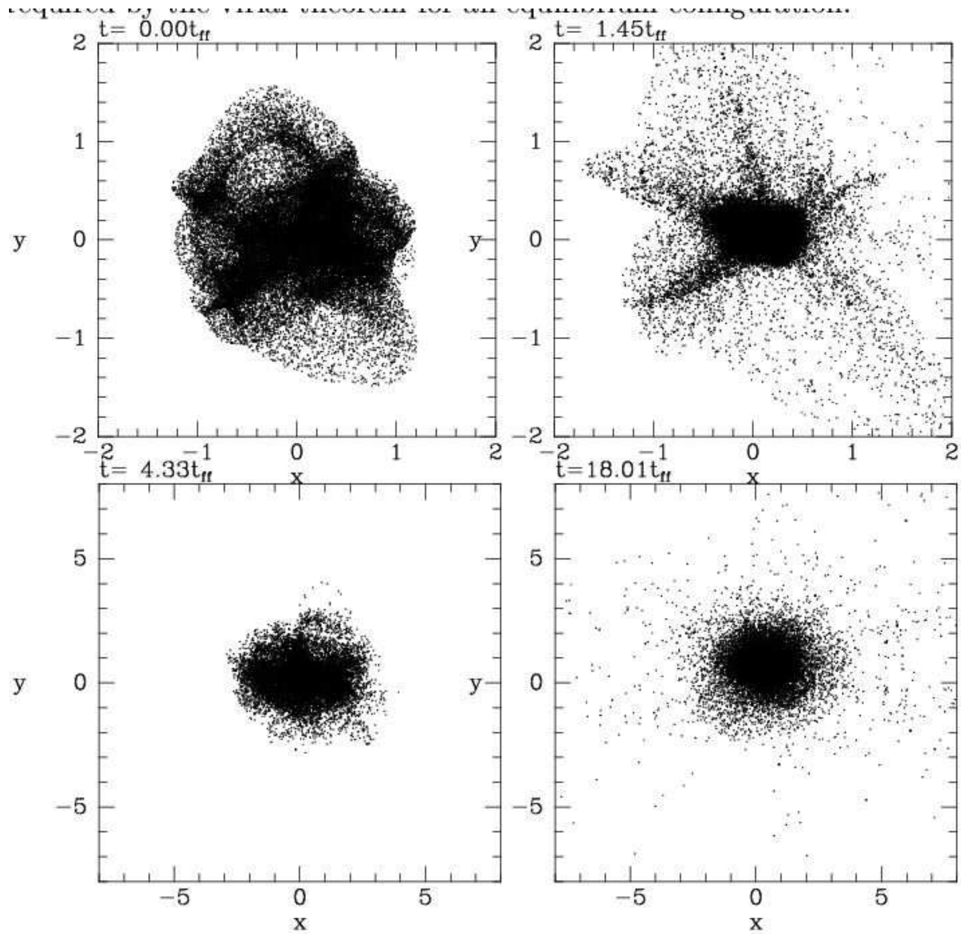

Fig. 5. Violent relaxation of a system of point masses that starts far from equilibrium (BT08)

7.2 Cluster Evolution

The evolution of stellar clusters under the influence of 2-body scattering has been computed in at least four ways:

1. Direct N-body modelling. This is expensive because one has to handle close encounters and short-period binary orbits in parallel with following the evolution of tens of thousands of stars for a Hubble time.

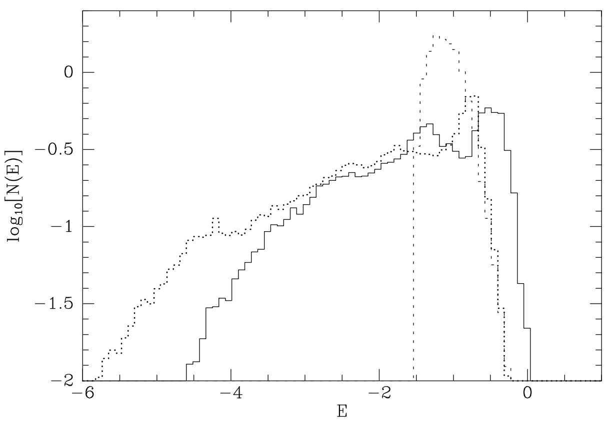

Fig. 6. The distribution over energy of the particles at three times (0, $1.45t_{\rm ff}$, $18t_{\rm ff}$) during the relaxation of Fig. 5. (BT08)

2. N-body integration of the CBE with stochastically applied velocity changes that generate the diffusion coefficients $\overline{\Delta\mathbf{v}}$ and $\overline{\Delta v_i \Delta v_j}$ computed from 2-body scattering.
3. Following the cluster's evolution when $f(\mathbf{J})$ is subjected to stochastic changes that generate orbit-averaged diffusion coefficients from 2-body scattering.
4. Evolution of a self-gravitating gas sphere with the thermal conductivity is computed from the diffusion coefficients (Lynden-Bell & Eggleton, MNRAS, 191, 483, 1980).

Fig. 7 shows the evolution of an initial Plummer sphere computed by solving the orbit-averaged F-P eqn (Takahashi, K, PASJ, 47, 561 1995). The cluster undergoes self-similar evolution towards infinite central density; outside a core radius r_0 the density becomes a power law $\rho \sim r^{-2.23}$. In this zone the velocity distribution becomes slightly radially biased, but in the original envelope ($r > 1$) it becomes almost completely radial. If τ is the time remaining to the singularity, $r_0 \propto \tau^{0.53}$, the core mass $M_0 \propto \tau^{0.41}$, and the central relaxation time $\propto \tau$. In fact $\tau \simeq 300 t_{\rm relax}(0)$.

Unfortunately, reality is less exciting because core collapse will be averted by the formation of a hard binary (and quite likely there were some present all along). Indeed, eq. (5.3) gives the time $t_{\rm fst}$ required for the first hard binary to form, and this becomes a smaller multiple of $t_{\rm relax}$ as the number $N_0 = M_0/m$ of stars in the core dwindles. Once $t_{\rm fst}$ drops below $\sim 300 t_{\rm relax}$, the appearance of a hard binary becomes likely. Acording to (5.3), $t_{\rm fst} < 300 t_{\rm relax}$ for $N_0 \lesssim 30$, so we expect a hard binary to appear not later than when $N_0 \simeq 30$. The energy released by the binary causes the core to expand and cool, arresting the outward flow of heat through the power-law segment. If the number of stars in the cluster is large enough, the process of core collapse followed by expansion can repeat (Fig. 8 from Breeden & Cohn, ApJ, 448, 672, 1995)

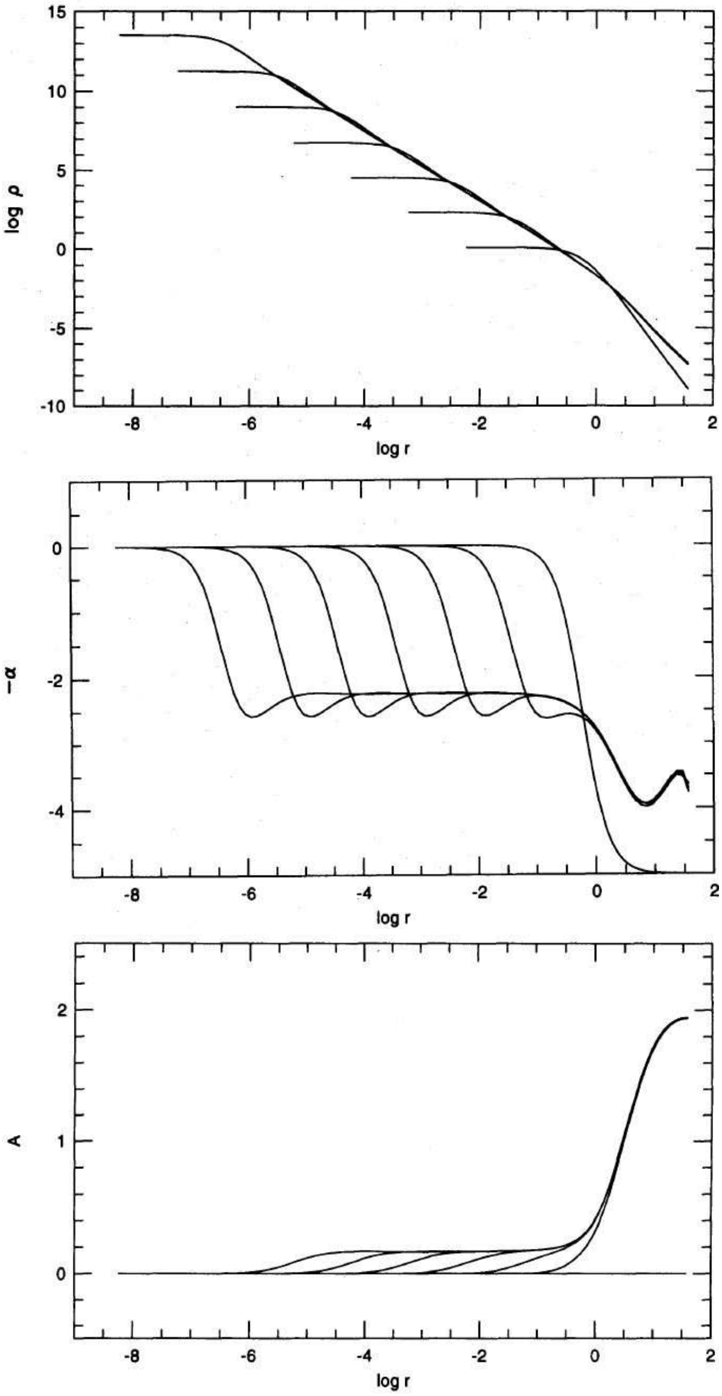

Fig. 7. Evolution of a star cluster that starts as a Plummer model. Top: density, middle: slope of log-density profile, bottom: velocity anisotropy profile $1-(\sigma_\phi^2+\sigma_\theta^2)/\sigma_r^2$ (Takahashi 1995)

Part II: Discs

8 How to heat a stellar disc

Gas clouds radiate energy and slump into near-circular orbits in galactic potential wells – a circular orbit is the orbit that has least energy for a given angular momentum L_z. Hence stars are born on nearly circular orbits, crowded into a small part of phase space. In action space (a true 3d projection of 6d phase space) stars are born near the $J_\phi = L_z$ axis. The kinetic theory of discs is all about the diffusion of stars away from this axis to increase entropy.

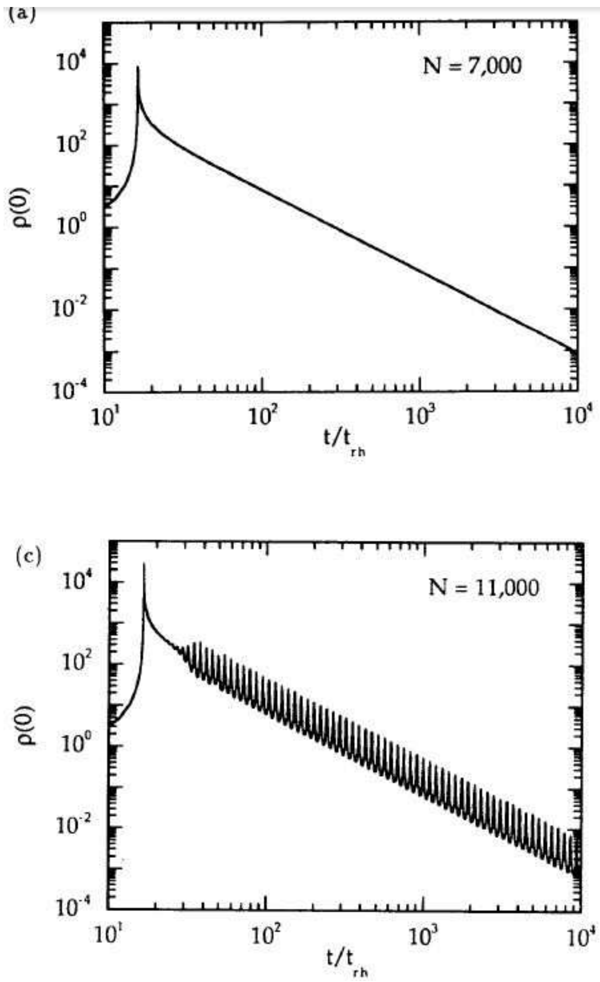

Fig. 8. The approach to cluster core collapse and its aftermath for clusters with 7000 and 11,000 stars (Breeden & Cohn 1995)

The diffusion is associated with stars moving to orbits of greater eccentricity and higher inclination – the one motion increases the radial velocity dispersion in the disc, σ_r, the other increases the vertical velocity dispersion σ_z. Thus the disc "heats".

Disc heating does not require an external source of E:

$$\frac{\partial H}{\partial L_z} = \frac{\partial H}{\partial J_\phi} = \Omega_\phi = \frac{v_c}{R}$$

In real discs the circular speed v_c never rises as fast as R (in galaxies like ours $v_c \sim$ const) so Ω_ϕ is a decreasing fn of R. Hence if a quantity of L_z can be transported from R_1 to $R_2 > R_1$ a quantity of energy

$$\delta E = [\Omega_\phi(R_1) - \Omega_\phi(R_2)]\delta L_z \tag{8.1}$$

is made available for random as opposed to circular motion. Thus entropy increase is all about shifting L_z outwards.

8.1 Molecular viscosity

Just as collisions of molecules in a gas give rise to viscosity, stellar encounters will cause viscous transport of momentum – a shear flow distorts the local velocity distribution from a Maxwellian, and encounters try to drive it back to a Maxwellian. The relaxation time (3.16) is the time required to restore a disturbed Maxwellian and the timescale on which molecular viscosity redistributes angular momentum cannot be shorter than $t_{\rm relax}$, so let's evaluate the latter near the Sun.

$$t_{\rm relax}\frac{\sigma^3}{2\pi G^2 m\rho\ln\Lambda}\simeq\frac{\sigma_z t_z^2}{2\pi\ln\Lambda}\frac{\sigma_z^2}{Gm},\tag{8.2}$$

where we have used that $t_z=(G\rho)^{-1/2}\simeq 100\,{\rm Myr}$ is the dynamical time for oscillations perpendicular to the Galactic plane. Now $\sigma_z\simeq 20\,{\rm km\,s^{-1}}$, which is about 70% of the Earth's orbital speed, so $Gm/\sigma_z^2\simeq 1.5R_\oplus\sim 10^{-5}\,{\rm pc}$ and hence

$$t_{\rm relax}\simeq 0.1\frac{200\,{\rm km\,s^{-1}}}{10^{-5}\,{\rm pc}}(100\,{\rm Myr})^2.\tag{8.3}$$

Finally $1\,{\rm km\,s^{-1}}$ moves you $\sim 1\,{\rm pc}$ in $1\,{\rm Myr}$ so $t_{\rm relax}\sim 10^9\,{\rm Myr}$, way longer than the age of the Universe. That is, molecular viscosity is far too weak to redistribute significant angular momentum.

8.2 Empirical evidence for disc heating

Blue stars are massive and short-lived, while red stars have small masses & long lives. So if you bin stars near the Sun by colour the mean age of stars in the blue bins is smaller than the mean age of stars in the red bins. A plot of random velocities in the radial (U) and tangential (V) directions shows a steady increase of σ with redness ($B-V$) & thus age (Fig. 9 from Aumer & Binney, MNRAS, 397, 1286, 2009). Part of this increase is due to encounters of stars with massive gas clouds, but most arises from particle-wave interactions.

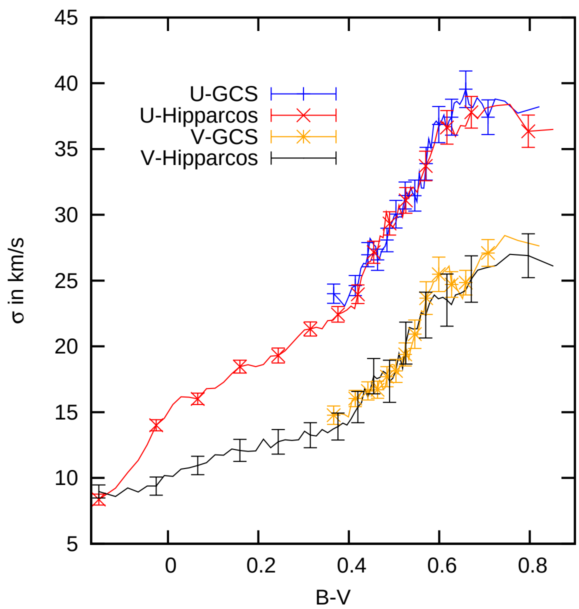

Fig. 9. Random velocities of stars near the Sun when binned by colour $B-V$. The curves show components in the radial (upper curves) and azimuthal (lower curves) directions (Aumer & Binney 2009)

9 Diffusion tensor from angle-action variables

Let's compute the second-order diffusion tensor D_{ij} for the orbit-averaged FPE. The potential Φ has a dominant time-independent part ϕ_0 and a fluctuating part, and we expand the latter in the angle-action variables proper to Φ_0:

$$\Phi(\mathbf{x},t)=\Phi_0(\mathbf{x})+\Phi_1(\mathbf{x},t)=\Phi_0+\sum_{\mathbf{n}}\Phi_{\mathbf{n}}(\mathbf{J},t)\cos(\mathbf{n}\cdot\boldsymbol{\theta}+\psi_{\mathbf{n}})\tag{9.1}$$

Hamilton's equation for $\mathbf{J}$ is

$$\dot{\mathbf{J}} = -\frac{\partial H}{\partial \boldsymbol{\theta}} = \sum_{\mathbf{n}} \mathbf{n}\Phi_{\mathbf{n}}(\mathbf{J},t)\cos(\mathbf{n}\cdot\boldsymbol{\theta}+\psi_{\mathbf{n}}). \tag{9.2}$$

To get a random change $\boldsymbol{\Delta}$ in $\mathbf{J}$ we need to integrate this equation for longer than the autocorrelation time of Φ_1. We expand the variables in powers of Φ_1/Φ_0:

$$\mathbf{J}(t) = \mathbf{J}_0 + \boldsymbol{\Delta}_1(t) + \boldsymbol{\Delta}_2(t) + \cdots \quad \text{and} \quad \boldsymbol{\theta}(t) = \boldsymbol{\theta}_0 + \boldsymbol{\Omega}_0 t + \boldsymbol{\theta}_1(t) + \cdots \tag{9.3}$$

$\boldsymbol{\Delta}_t$ is obtained by integrating (9.2) along an unperturbed orbit

$$\begin{aligned}\boldsymbol{\Delta}_1(t) &= \sum_{\mathbf{n}} \int_0^T \mathrm{d}t\, \Phi_{\mathbf{n}}(\mathbf{J},t)\sin(\mathbf{n}\cdot\boldsymbol{\theta}+\psi_{\mathbf{n}}) \\ &= \sum_{\mathbf{n}} \int_0^T \mathrm{d}t\, \Phi_{\mathbf{n}}(\mathbf{J},t)\sin[\mathbf{n}\cdot)\boldsymbol{\theta}_0 + \boldsymbol{\Omega}t) + \psi_{\mathbf{n}}].\end{aligned} \tag{9.4}$$

We multiply this equation by itself and average over all phases $\boldsymbol{\theta}_0$. After reordering the integrals so the integral over $\boldsymbol{\theta}_0$ is done first, the innermost integral is

$$(2\pi)^{-3}\int \mathrm{d}^3\boldsymbol{\theta}_0\ \sin[\mathbf{n}\cdot(\boldsymbol{\theta}_0+\boldsymbol{\Omega}_0 t)+\psi_{\mathbf{n}}]\sin[\mathbf{n}'\cdot(\boldsymbol{\theta}_0+\boldsymbol{\Omega}_0 t')+\psi_{\mathbf{n}'}] \tag{9.5}$$

Using $2\sin A\sin B = \cos(A-B)-\cos(A+B)$ and that the integral of any cosine that depends on $\boldsymbol{\theta}_0$ will vanish, we conclude that the innermost integral is

$$\tfrac{1}{2}\cos[\mathbf{n}\cdot\boldsymbol{\Omega}_0(t-t')]\,\delta_{\mathbf{n}\mathbf{n}'}. \tag{9.6}$$

Next we take the ensemble-average of the random variable $\Phi_{\mathbf{n}}$ under the assumption that it's a stationary random process so its autocorrelation depends only on the lag $t-t'$:

$$\langle\Phi_{\mathbf{n}}(\mathbf{J},t)\Phi_{\mathbf{n}}(\mathbf{J},t')\rangle = c_{\mathbf{n}}(\mathbf{J},t-t'). \tag{9.7}$$

Now

$$\begin{aligned}\langle\Delta_i\Delta_j\rangle &= \tfrac{1}{2}\sum_{\mathbf{n}} n_i n_j \int_0^T \mathrm{d}t \int_0^T \mathrm{d}t'\, c_{\mathbf{n}}(\mathbf{J},t-t')\cos[\mathbf{n}\cdot\boldsymbol{\Omega}_0(t-t')] \\ &= \tfrac{1}{4}\sum_{\mathbf{n}} n_i n_j \int_{-T}^T \mathrm{d}v\, c_{\mathbf{n}}(\mathbf{J},v)\cos(\mathbf{n}\cdot\boldsymbol{\Omega}_0 v)\int_{|v|}^{`2T-|v|} \mathrm{d}u \\ &= \tfrac{1}{2}\sum_{\mathbf{n}} n_i n_j \int_{-T}^T \mathrm{d}v\, c_{\mathbf{n}}(\mathbf{J},v)\cos(\mathbf{n}\cdot\boldsymbol{\Omega}_0 v)(T-|v|).\end{aligned} \tag{9.8}$$

We are choosing T to be bigger than the autocorrelation time, so $c(\mathbf{J},v)$ is non-negligible only for $v \ll T$. Hence we can approximate $T-|v| \simeq T$. With this approximation $\langle\Delta_i\Delta_j\rangle$ becomes proportional to T, and the diffusion coefficient is the constant of proportionality, i.e.

$$D_{ij} = \tfrac{1}{2}\sum_{\mathbf{n}} n_i n_j\, \widetilde{c}_{\mathbf{n}}(\mathbf{J},\mathbf{n}\cdot\boldsymbol{\Omega}), \tag{9.9}$$

where

$$\widetilde{c}_{\mathbf{n}}(\mathbf{J},\omega) = \int_{-T}^T \mathrm{d}v\, c_{\mathbf{n}}(\mathbf{J},v)\cos(\omega v) = \int_{-T}^T \mathrm{d}v\ \langle\Phi_{\mathbf{n}}(\mathbf{J},t)\Phi_{\mathbf{n}}(\mathbf{J},t-v)\rangle\cos(\omega v) \tag{9.10}$$

is the power spectrum of the fluctuations.

Equation (9.9) implies that a star is able to diffuse through action space to the extent that the fluctuations contain power at one of its natural frequencies $\mathbf{n}\cdot\boldsymbol{\Omega}$. If the fluctuations are narrow-band because they are associated with normal modes of the system, only a minority of stars will diffuse – those that resonate with the oscillations.

9.1 Wave-particle scattering

Consider a perturbing potential

$$\Phi_1(R,\phi,t) \propto \mathrm{e}^{\mathrm{i}(kR+m\phi-\omega t)} \tag{9.11}$$

Depending on the values taken by k and m, this describes a steadily rotating pattern that can range from axisymmetry ($m=0$) to barred ($k=0$) through spirals of varying pitch angle. In a frame that rotates at angular velocity

$$\omega_\mathrm{p} = \omega/m \tag{9.12}$$

the potential is stationary. Hence the in this frame the Hamiltonian is a constant of motion. This Hamiltonian is **Jacobi's invariant**

$$H = E - \omega_\mathrm{p} J_\phi. \tag{9.13}$$

Hence changes in the energy E (numerical value of the time-dependent Hamiltonian in an inertial frame) and changes in angular momentum J_ϕ are related by

$$\delta E = \omega_\mathrm{p} \delta J_\phi. \tag{9.14}$$

The energy associated with random motion is the difference

$$E_\mathrm{rand} = E - E_\mathrm{circ}(J_\phi)$$

between a star's energy and the energy of a circular orbit with its angular momentum. Differencing this equation and using equation (9.13) we get

$$\delta E_\mathrm{rand} = \left(\omega_\mathrm{p} - \frac{\partial E_\mathrm{circ}}{\partial J_\phi}\right)\delta J_\phi = (\omega_\mathrm{p} - \omega_\mathrm{circ})\delta J_\phi. \tag{9.15}$$

Hence inside the **corotation** resonance (where stars orbit at the same angular velocity as the pattern), E_rand is increased if stars donate angular momentum to the wave, and outside corotation E_rand is increased when stars absorb angular momentum from the wave. At corotation absorption or emission of angular momentum leaves E_rand unchanged.

The key resonances are the **inner Lindblad resonance** (**ILR**) where $\Omega_r = m(\Omega_\phi - \omega_\mathrm{p})$ (stars overtake the pattern at the frequency of their radial oscillations) and the **outer Lindblad resonance** (**OLR**) where $\Omega_r = m(\omega_\mathrm{p} - \Omega_\phi)$ (the pattern overtakes stars at their radial frequency). A disc would heat if waves gathered J_ϕ at the ILR and deposited it at the OLR.

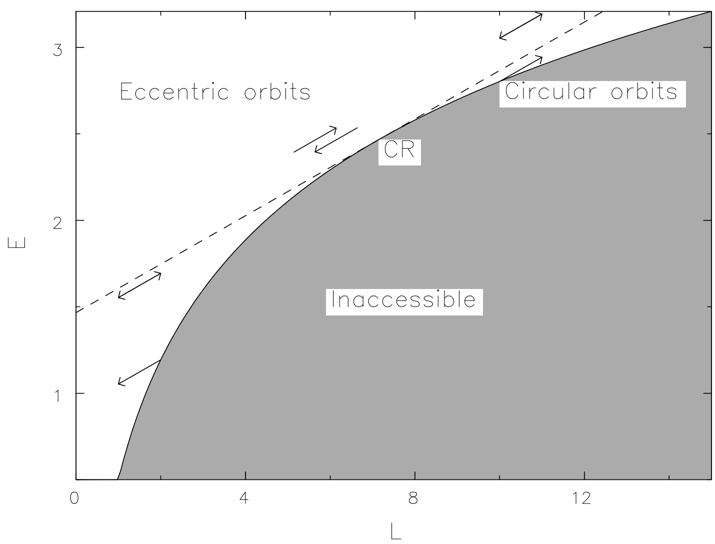

Fig. 10. A Lindblad diagram [(E, L_z) plane] with arrows showing motion of stars when resonantly scattered (Sellwood & Binney 2002)

10 Wave mechanics of discs

The stars of a disc are coupled to one another by gravity, so it is to be expected that waves can propagate through the disc. Perhaps the easiest waves to imagine are bending waves: suppose we displace downward a patch of the disc. Then the attraction of the undisplaced stars will pull our patch back up, and our patch will pull the other stars down. So we expect a ripple of vertical displacement to propagate out from our patch like a ripple propagating along a chain. Differential rotation within the disc will make the dynamics of the ripples complex, but it's intuitively clear that stars throughout the disc will eventually be affected by our original displacement.

Galactic discs do show systematic distortions from planarity, but probably more important are waves of displacement within the disc. A systematic variation of the direction of the major axes of the elliptical orbits of stars generates a spiral pattern of stellar density

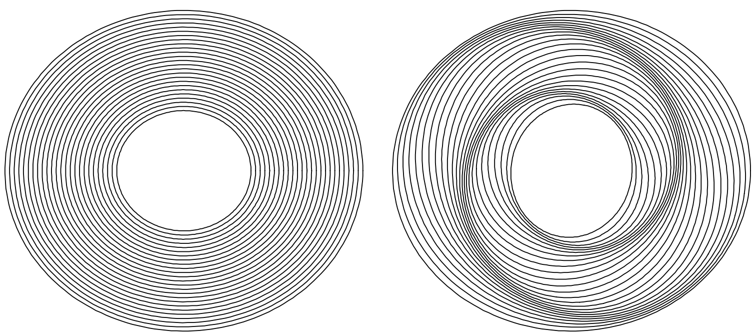

Fig. 11. Left: a series of nearly circular ellipses. Right: the same ellipses with systematically rotated principal axes (BT08)

This pattern gives rise to a spiral perturbation to the gravitational field, which perturbs orbits, thus over time modifying the density distribution. It is physically plausible that as a consequence a spiral pattern will propagate through the disc.

We can get insight into wave propagation with a local (WKB) analysis by positing that the wavelength of the waves is small compared to R. There are 2 important cases: (i) tightly wound spirals and loosely wound waves at corotation.

10.1 Tightly wound spirals

$$\Phi_1(R,\phi,t) = \epsilon \mathrm{e}^{\mathrm{i}(kR+m\phi-\omega t)} \tag{10.1}$$

where $kR \gg 1$. By solving Poisson's equation in cylindrical coordinates (in which it separates), it's not hard to show that this perturbation is generated by a thin sheet in which the surface density of stars is

$$\Sigma_1 = -\frac{|k|}{2\pi G}\Phi_1. \tag{10.2}$$

Φ_1 perturbs the DF to $f_0(\mathbf{J})+f_1(\boldsymbol{\theta},\mathbf{J},t)$. We compute f_1 from the linearised CBE, integrate over $\mathbf{v}$ to obtain the response density $\Sigma_{\rm resp}(R,\phi,t)$ and obtain the **Lin-Shu-Kalnajs dispersion relation** by equating $\Sigma_{\rm resp}$ to Σ_1:

$$\begin{aligned} m(\omega_{\rm p}-\omega_{\rm circ})^2 &= \Omega_r^2 - 2\pi G\Sigma_0|k|\mathcal{F}(s,\chi) \qquad \text{stars} \\ m(\omega_{\rm p}-\omega_{\rm circ})^2 &= \Omega_r^2 - 2\pi G\Sigma_0|k| + c_{\rm s}^2k^2 \qquad \text{gas.} \end{aligned} \tag{10.3a}$$

Here

$$s \equiv \frac{\omega - m\Omega_\phi}{\Omega_r} \tag{10.3b}$$

is the ratio of driving frequency a star experiences to its natural radial frequency, while

$$\chi \equiv \left(\frac{\sigma_R k}{\Omega_r}\right)^2 \tag{10.3c}$$

is the square of the ratio of typical radial excursions $\sigma_R/(\Omega_r/2\pi)$ to the wavelength $2\pi/k$, and $\mathcal{F}$ is the **reduction factor**

$$\mathcal{F}(s,\chi) \equiv 2(1-s^2)\frac{\mathrm{e}^{-\chi}}{\chi}\sum_{l=1}^{\infty}\frac{\mathcal{I}_l(\chi)}{1-s^2/l^2}, \tag{10.3d}$$

where $\mathcal{I}_l$ is a modified Bessel function. $\mathcal{F}$ is a decreasing function of increasing χ and $\mathcal{F}\to 1$ as $\chi \to 0$.

The second line in (10.3a) is the dispersion relation for tightly-wound spiral waves in an isothermal gas disc. It states that the frequency at which a star is shaken $m(\omega_\mathrm{p}-\omega_\mathrm{circ})$ should be equal to the quadrature sum of the radial frequency and the sound-wave frequency kc_s less the amount $2\pi G\Sigma_0|k|$ by which gravity makes the disc "squishy". This squishiness factor is least effective at long wavelengths because at these wavelengths the field lines can leak most easily into empty space. At the very shortest wavelengths the kc_s dominates: waves become sound waves.

The Lin-Shu-Kalnajs relation for a stellar disc in the first line of (10.3a) differs from the gas relation in trading the kc_s term for the presence of the reduction factor $\mathcal{F}$, which reduces the effectiveness of gravity when the stars have a large velocity dispersion σ_R. It's to be expected that increased velocity dispersion will affect the stellar dynamics much as increased sound speed will the gas dynamics.

For $\sigma_R > 0$ or $c_\mathrm{s} > 0$ the right sides of the dispersion relations (10.3a) are > 0 for very small and very large k with a minimum in between. If the right side dips below zero, the disc is unstable. When $\sigma_R = 0$, the stellar disc is unstable for

$$k > k_\mathrm{crit} \equiv \frac{\Omega_r^2}{2\pi G\Sigma_0}. \tag{10.4}$$

The disc is stable for all k provided

$$1 < Q \equiv \frac{\Omega_r \sigma_R}{3.36 G\Sigma_0}, \tag{10.5}$$

which is called **Toomre's Q**.

When the disc is stable, the dispersion relation cannot be satisfied with ω_p very close to ω_circ: the region around corotation doesn't support waves.

10.2 Swing amplifier

Although the region around corotation does not support steady, sinusoidally oscillating tightly-wound waves, solutions can be found (Goldreich & Lynden-Bell 65 for gas, Julian & Toomre 1966 for stars) for short-wavelength waves and these have an interesting temporal evolution. Differential rotation shears an initially leading-spiral group of waves into a trailing group and in the process amplifies it. The factor by which the packet is amplified depends sensitively on Q and

$$X \equiv \frac{k_\mathrm{crit}}{k_\phi}. \tag{10.6}$$

and for $Q \simeq 1.5$ can exceed 10. This process of amplification of waves as they are sheared from leading to trailing is the **swing amplifier**.

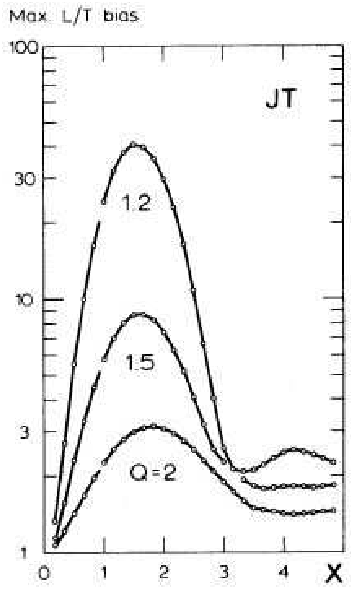

Fig. 12. Swing-amplification factor for three values of Toomre's Q as a function of X (eq. (10.6) (Toomre 1981)

10.3 Discs as lasing cavities

Combining tightly-wound running-wave solutions with the swing amplifier yields the following picture of disc dynamics (Toomre in "The structure and evolution of normal galaxies, eds S.M. Fall and D. Lynden-Bell, CUP, 1981). Noise-generated leading waves propagate towards CR, are there swing amplified and reflected to trailing waves that propagate inwards. If these are somehow reflected to leading waves, they will again be amplified and reflected at CR and the disc will be unstable. If they reach a Lindblad resonance, they will be Landau damped and the disc will be stable. This picture is supported by (i) the empirical fact that discs with slowly-rising circular speed v_c tend to be unstable while discs with steep central rises in v_c to an approximately constant value are stable, and (ii) the morphology of the normal modes of the small number of discs for which these have been computed.

N-body simulations of formally stable discs were until recently puzzling. If a simulation of such a disc is continued long enough, spiral structure eventually grows to order unity and a strong bar forms from the disc. The more particles there are in the disc, the longer it takes for the bar to form (Fig. 13). Power spectra at three stages in the evolution of a 50 M particle simulation show activity near the ILR (Fig. 14 from Sellwood, ApJ 751, 44, 2012)

Sellwood (2012) and Sellwood & Carlberg (2014) interpret this as follows. Swing amplified noise heats the disc at its ILR. The heated region behaves like a half-silvered mirror and reflects some of the swing-amplified power en route to an ILR that lies inside the first ILR. So some once-amplified noise is amplified again, and is again part-reflected. Eventually all waves of this second episode get through to their ILR, and heat the disc, creating a new and more reflective half-silvered mirror. This cycle repeats until the disc has a very irregular DF and swing amplification becomes highly effective.

The disc is more strongly heated at ILR than at OLR because at ILR stars stay on resonance as their actions ar shifted whereas at OLR they quickly go off resonance (Fig. 15). Activity starts inside & spreads outwards as the inside heats.

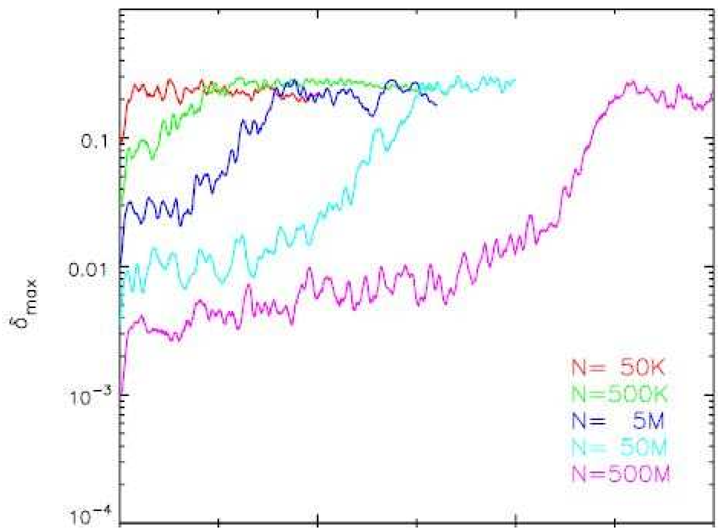

Fig. 13. The amplitude of fluctuations as a function of time in 5 simulations of the same galactic disc with increasing particle number. The t scale runs 0 to 3000 units. (Sellwood 2012)

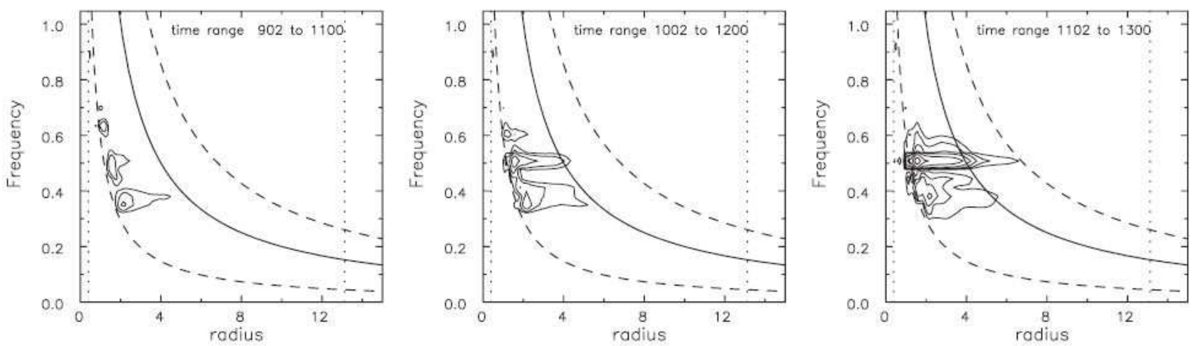

Fig. 14. The contours show the amplitude of the temporal FT of the $m = 2$ component of the density over 3 time intervals in the simulation of Fig. 13 that has 50M particles. The frequency $\omega = m\omega_{\rm p}$ is plotted vertically. The full curve marks $m\omega_{\rm circ}$ at each radius and the dashed curves show $m\omega_{\rm circ} \pm \Omega_r$, i.e., the frequencies that put an ILR or OLR at the radius. (Sellwood 2012)

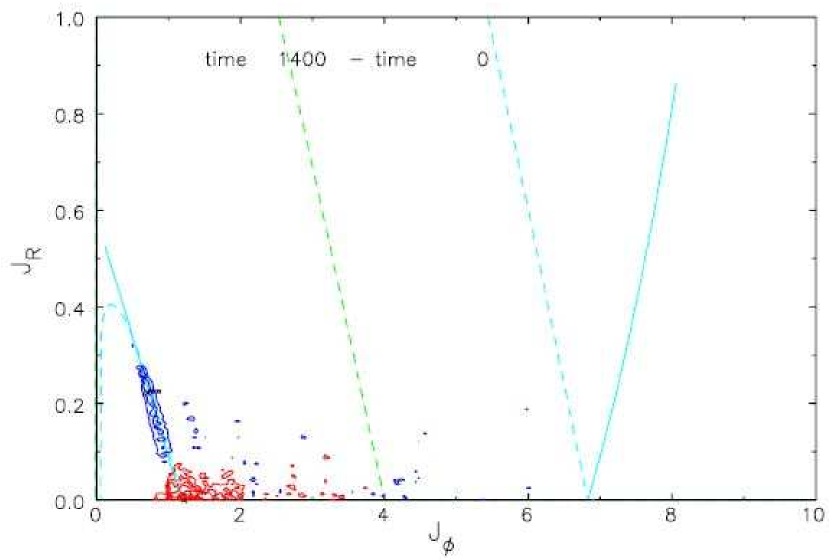

Fig. 15. Changes to the DF in the 50 M particle simulation of Fig. 13. Blue contours mark increased density, red contours decreased density. The dashed lines indicate ILR, CR and OLR for $\omega_{\rm p} = 0.25$. Solid cyan lines show trajectories of particle at constant Jacobi invariant.

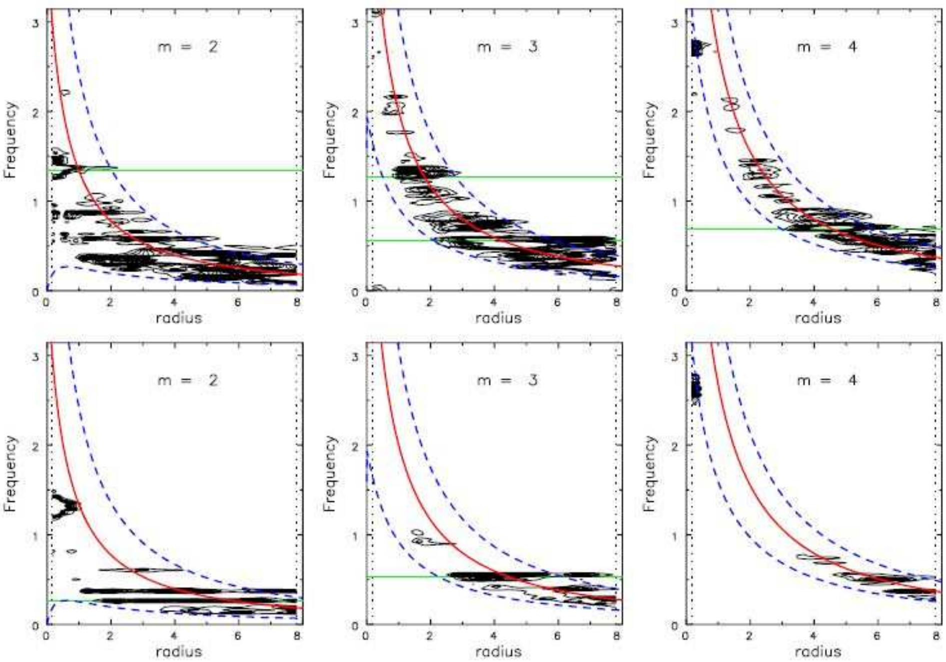

Fig. 16. Contours show the amplitude of the temporal Fourier transform of the $m = 2, 3, 4$ component of the density at each radius in a three-dimensional, 20M particle simulation of a disc. The top row covers the 1st half of the simulation, the bottom row the 2nd half. The full curve shows $m\omega_{\rm circ}$ while the dashed curves show $m\omega_{\rm circ} \pm \Omega_r$. (Sellwood & Carlberg 2014)

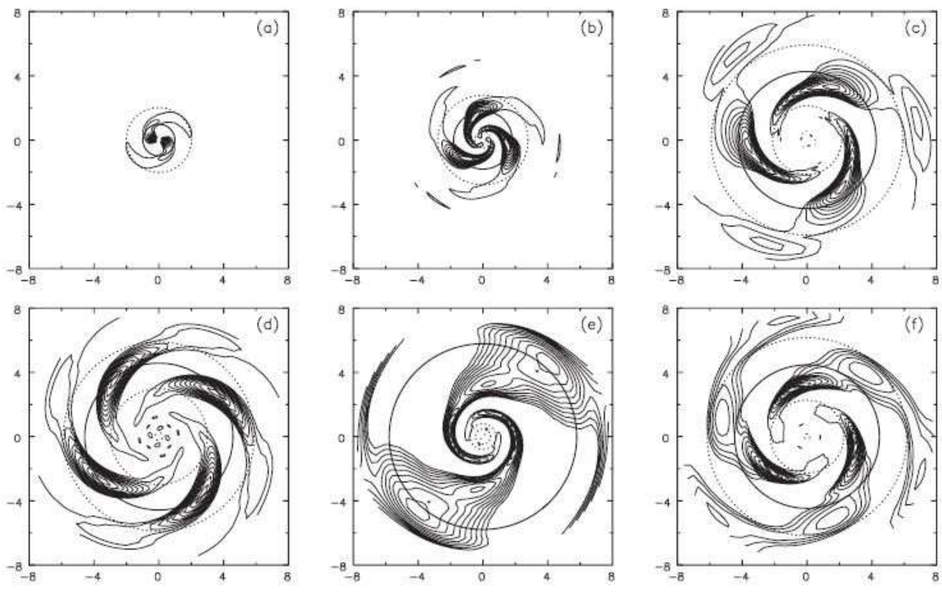

Fig. 17. Spatial structure of the modes marked in green on Fig. 16. (Sellwood & Carlberg 2014)

10.4 Derivation of the Lin-Shu-Kalnajs dispersion relation

We write $f(\boldsymbol{\theta}, \mathbf{J}, t) = f_0(\mathbf{J}) + f_1(\boldsymbol{\theta}, \mathbf{J}, t)$, where $(\boldsymbol{\theta}, \mathbf{J})$ are the angle-action variables of the equilibrium system defined by $f_0(\mathbf{J})$. Since the Poisson bracket $\{,\}$ is linear in each slot the linearised CBE $\partial f/\partial t = \{H, f\}$ becomes

$$\frac{\partial f_1}{\partial t} = \{H_0, f_1\} + \{H_1, f_0\} = -\boldsymbol{\Omega}_0 \cdot \frac{\partial f_1}{\partial \boldsymbol{\theta}} + \frac{\partial \Phi_1}{\partial \boldsymbol{\theta}} \cdot \frac{\partial f_0}{\partial \mathbf{J}} \tag{10.7}$$

This linear equation for f_1 is time-translation invariant so we may seek solutions in which f_1 and Φ_1 are $\propto \mathrm{e}^{-\mathrm{i}\omega t}$, i.e.,

$$f_1(\boldsymbol{\theta}, \mathbf{J}, t) = \sum_{\mathbf{n}} \delta f_{\mathbf{n}}(\mathbf{J}) \mathrm{e}^{\mathrm{i}(\mathbf{n}\cdot\boldsymbol{\theta} - \omega t)}$$
$$\Phi_1(\boldsymbol{\theta}, \mathbf{J}, t) = \sum_{\mathbf{n}} \delta \Phi_{\mathbf{n}}(\mathbf{J}) \mathrm{e}^{\mathrm{i}(\mathbf{n}\cdot\boldsymbol{\theta} - \omega t)}$$

When these expansions are inserted into (10.6) we can equate coefficients of $\mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}}$ to obtain

$$\delta f_{\mathbf{n}}(\mathbf{J}) = \frac{\mathbf{n}\cdot\partial f_0/\partial\mathbf{J}}{\mathbf{n}\cdot\boldsymbol{\Omega}_0 - \omega} \delta\Phi_{\mathbf{n}}(\mathbf{J})$$

This equation is very similar to one that occurs in Landau's analysis of a plasma, with **J** replacing **n**.

Our next step is to expand Φ_1 from (10.1) in angle-action coordinates. We can do this if we adopt the **epicycle approximation**. In this approximation we linearise the equations of motion of a star of angular momentum L_z around the circular orbit with this angular momentum. This circular orbit has radius R_{g}, and provides the **guiding centre**. Our star moves on an elliptical orbit around the guiding centre as the latter moves on a circular orbit around the galaxy

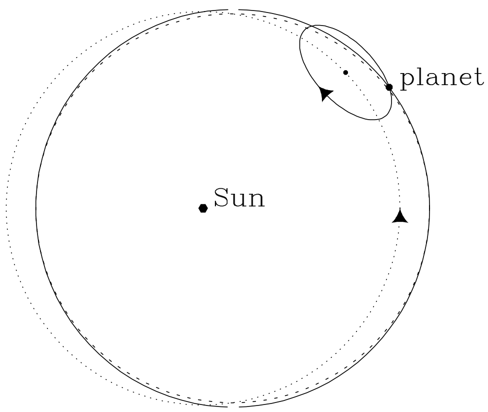

Fig. 18. An elliptical Kepler orbit as an elliptical epicycle on a circular guiding-centre orbit

In this approximation

$$R(\mathbf{J}, \boldsymbol{\theta}) = R_{\mathrm{g}}(J_\phi) + a(\mathbf{J})\cos\theta_r, \qquad \phi(\mathbf{J}, \boldsymbol{\theta}) = \theta_\phi + \frac{\gamma a}{R_{\mathrm{g}}}\sin\theta_r \tag{10.8}$$

where

$$a(\mathbf{J}) = \sqrt{\frac{2J_r}{\Omega_r}} \text{ and } \gamma = 2\Omega_\phi/\Omega_r. \tag{10.8b}$$

In the epicycle approximation, we neglect the J_r dependency of $\boldsymbol{\Omega}$, so $\boldsymbol{\Omega}(\mathbf{J}) \simeq \boldsymbol{\Omega}(0, J_\phi)$, the components of this vector being the **epicycle frequencies**. It follows that the Fourier decomposition of Φ_1 is

$$\begin{aligned}\Phi_1(R, \phi, t) &= \epsilon \mathrm{e}^{\mathrm{i}(kR + m\phi - \omega t)} \\ &= \epsilon \mathrm{e}^{\mathrm{i}kR_{\mathrm{g}}} \exp(\mathrm{i}ka\cos\theta_r) \mathrm{e}^{\mathrm{i}m\theta_\phi} \exp\left(\mathrm{i}\frac{m\gamma a}{R_{\mathrm{g}}}\sin\theta_r\right) \mathrm{e}^{-\mathrm{i}\omega t}.\end{aligned} \tag{10.9}$$

Now we use

$$\begin{aligned}\exp\left[\mathrm{i}\left(ka\cos\theta_r + \frac{m\gamma a}{R_{\mathrm{g}}}\sin\theta_r\right)\right] &= \exp[\mathrm{i}a\mathcal{K}\sin(\theta_r + \alpha)] \\ &= \sum_l \mathcal{J}_l(\mathcal{K}a)\mathrm{e}^{\mathrm{i}l(\theta_r + \alpha)},\end{aligned} \tag{10.10a}$$

where

$$\alpha(J_\phi) \equiv \arctan\left(\frac{m\gamma}{kR_{\rm g}}\right) \qquad \mathcal{K}(J_\phi) \equiv \sqrt{k^2 + \frac{m^2\gamma^2}{R_{\rm g}^2}} \tag{10.10b}$$

are, respectively, the pitch and the total wavenumber of the spirals. We can now write Φ_1 in the form

$$\Phi_1(\boldsymbol{\theta}, \mathbf{J}, t) = \epsilon \sum_{l=-\infty}^{\infty} \mathrm{e}^{\mathrm{i}(kR_{\rm g}+l\alpha)} \mathcal{J}_l(\mathcal{K}a) \mathrm{e}^{\mathrm{i}(l\theta_r + m\theta_\phi - \omega t)}, \tag{10.11}$$

which implies that

$$\delta\Phi_{(l,m,0)}(\mathbf{J}) = \epsilon \mathrm{e}^{\mathrm{i}(kR_{\rm g}+l\alpha)} \mathcal{J}_l(\mathcal{K}a), \tag{10.12}$$

so

$$\delta f_{(l,m,0)}(\mathbf{J}) = \epsilon \frac{\mathbf{n}\cdot\partial f_0/\partial\mathbf{J}}{\mathbf{n}\cdot\mathbf{\Omega}_0 - \omega} \mathrm{e}^{\mathrm{i}(kR_{\rm g}+l\alpha)} \mathcal{J}_l(\mathcal{K}a) \qquad \mathbf{n} = (l, m, 0). \tag{10.13}$$

Now we need to integrate this perturbed distribution function over velocities to determine the perturbed density and compare it with (10.2), the density required to generate the assumed Φ_1. We exploit the fact that $(\mathbf{x}, \mathbf{v})$ and $(\boldsymbol{\theta}, \mathbf{J})$ are both systems of canonical coordinates, so $\mathrm{d}^2\mathbf{x}\,\mathrm{d}^2\mathbf{v} = \mathrm{d}^2\boldsymbol{\theta}\,\mathrm{d}^2\mathbf{J}$:

$$\begin{aligned}\Sigma_{\rm resp}(R,\phi,t) = {}& \frac{\epsilon}{R} \int \mathrm{d}^2\mathbf{J}\mathrm{d}^2\boldsymbol{\theta}\, \delta\left(\phi - \theta_\phi - \frac{\gamma a}{R_{\rm g}}\sin\theta_r\right) \delta(R - R_{\rm g} - a\cos\theta_r) \\ & \times \sum_{\mathbf{n}=(l,m)} \frac{\mathbf{n}\cdot\partial f_0/\partial\mathbf{J}}{\mathbf{n}\cdot\mathbf{\Omega}_0 - \omega} \mathrm{e}^{\mathrm{i}(kR_{\rm g}+l\alpha)} \mathcal{J}_l(\mathcal{K}a)\, \mathrm{e}^{\mathrm{i}(\mathbf{n}\cdot\boldsymbol{\theta} - \omega t)}.\end{aligned} \tag{10.14}$$

We use the Dirac functions to do the integrals over θ_ϕ and J_ϕ. Subsequently every occurrence of $R_{\rm g}$ should be replaced by $R - a\cos\theta_r$ but by virtue of the tight-winding approximation we neglect the difference between $R_{\rm g}$ and R except when it occurs multiplied by k in an exponential. Then

$$\begin{aligned}\Sigma_{\rm resp}(R,\phi,t) = {}& \frac{\epsilon}{R} \mathrm{e}^{\mathrm{i}(kR+m\phi-\omega t)} \frac{\mathrm{d}J_\phi}{\mathrm{d}R_{\rm g}}\bigg|_{R_{\rm g}=R} \sum_{l=-\infty} \mathrm{e}^{\mathrm{i}l\alpha} \int \mathrm{d}J_r\, \mathcal{J}_l(\mathcal{K}a) \frac{\mathbf{n}\cdot\partial f_0/\partial\mathbf{J}}{\mathbf{n}\cdot\mathbf{\Omega}_0 - \omega} \\ & \times \int \mathrm{d}\theta_r\, \exp\left[\mathrm{i}\left(l\theta_r - m\frac{\gamma a}{R}\sin\theta_r - ka\cos\theta_r\right)\right] \qquad \mathbf{n} = (l, m).\end{aligned} \tag{10.15}$$

We have $\mathrm{d}J_\phi/\mathrm{d}R_{\rm g} = R_{\rm g}\Omega_\phi/\gamma$ and for f_0 we use

$$f_0(\mathbf{J}) = \frac{\gamma\Sigma_0}{2\pi\sigma^2}\, \mathrm{e}^{-\Omega_r J_r/\sigma^2} \tag{10.16}$$

which generates a biaxial Gaussian velocity distribution in the epicycle approximation. We again use (10.10) to express the exponential of sinusoids as a sum of Bessel functions times exponentials. That done the integral over θ_r can be done, leaving

$$\begin{aligned}\Sigma_{\rm resp}(R,\phi,t) &= \frac{\epsilon\Omega_r^2\Sigma_0}{\sigma^4} \mathrm{e}^{\mathrm{i}(kR+m\phi-\omega t)} \sum_{l=-\infty} \frac{-l}{l\Omega_r + m\Omega_\phi - \omega} \int \mathrm{d}J_r\, |\mathcal{J}_l(\mathcal{K}a)|^2 \mathrm{e}^{-\Omega_r J_r/\sigma^2} \\ &= \frac{\epsilon\Omega_r\Sigma_0}{\sigma^2} \mathrm{e}^{\mathrm{i}(kR+m\phi-\omega t)} \sum_{l=-\infty} \frac{-l\mathcal{I}_l(\chi)\mathrm{e}^{-\chi}}{l\Omega_r + m\Omega_\phi - \omega},\end{aligned} \tag{10.17}$$

where $\mathcal{I}_l$ is a modified Bessel function and χ is defined by (10.3c). The Lin-Shu-Kalnajs dispersion relation (10.3a) follows by equating $\Sigma_{\rm resp}$ to Σ_1 given by (10.2).