

5 Instabilities

5.1 Gravitational instability

- In our study of sound waves, we considered what happens when you introduce small perturbations to a uniform fluid that is initially at rest, but deliberately ignored the effects of the self-gravity of the gas. We now relax this assumption, and ask what happens to these perturbations when gravitational effects are included.
- With the gravitational term included, our perturbed version of Euler's equation becomes

$$(\rho_0 + \rho_1) \left(\frac{\partial \vec{v}_1}{\partial t} + \vec{v}_1 \cdot \nabla \vec{v}_1 \right) = -\nabla(p_0 + p_1) - (\rho_0 + \rho_1)\nabla(\phi_0 + \phi_1), \quad (481)$$

where ϕ_0 is the gravitational potential of the unperturbed fluid, and ϕ_1 is the perturbation to the potential arising from the density perturbation in the gas. We determine the potential from the density via Poisson's equation:

$$\nabla^2(\phi_0 + \phi_1) = 4\pi G(\rho_0 + \rho_1), \quad (482)$$

and so

$$\nabla^2\phi_0 = 4\pi G\rho_0, \quad (483)$$

$$\nabla^2\phi_1 = 4\pi G\rho_1. \quad (484)$$

- We now encounter our first problem: what to do about ϕ_0 ? In a truly infinite, uniform medium, ϕ_0 is also uniform everywhere and we can therefore set it to zero. However, we have known since the time of Newton that this situation is unstable, with the slightest perturbation away from uniformity leading to gravitational collapse. In addition, ϕ_0 is not zero if we consider a uniform but finite system.
- In his original analysis of this problem, Sir James Jeans assumed that we could nevertheless disregard the ϕ_0 term, a procedure that has subsequently become known as the **Jeans swindle**. Although it is not strictly justified in the present case, there are many physical situations in which this approach is justified: e.g. in a rotating body that is in equilibrium, such as a disk, the force from the unperturbed potential is simply the centripetal force required to balance the centrifugal force arising from the rotation, and hence can be neglected in our perturbation analysis.
- If we apply the Jeans swindle, then our perturbed versions of the continuity and Euler equations can be reduced to the following form

$$\frac{\partial \rho_1}{\partial t} = -\rho_0 \nabla \cdot \vec{v}_1, \quad (485)$$

$$\rho_0 \frac{\partial \vec{v}_1}{\partial t} = -\frac{\gamma p_0}{\rho_0} \nabla \rho_1 - \rho_0 \nabla \phi_1. \quad (486)$$

These are the same equations as we saw previously in our study of sound waves, with the addition of one extra term on the right-hand side of the momentum equation, corresponding to the force due to the self-gravity of the perturbations.

- Taking the time derivative of the first of these equations and the spatial derivative of the second, and then combining them, we obtain

$$\frac{\partial^2 \rho_1}{\partial t} = c_s^2 \nabla^2 \rho_1 + \nabla^2 \phi_1. \quad (487)$$

We can then use the perturbed form of Poisson's equation to write this as

$$\frac{\partial^2 \rho_1}{\partial t} = c_s^2 \nabla^2 \rho_1 + 4\pi G \rho_0 \rho_1. \quad (488)$$

We see that once again, we have a wave equation, but that the self-gravity of the perturbation provides an additional driving term.

- If we now consider a plane wave solution of the form $e^{-i(\vec{k}\cdot\vec{x}+\omega t)}$, then we obtain the dispersion relation

$$\omega^2 = c_s^2 k^2 - 4\pi G \rho_0. \quad (489)$$

- When k is very large (i.e. our perturbations are physically small), the gravitational term is unimportant and this expression reduces to the dispersion relation for a simple, undamped sound wave. On the other hand, when the perturbations are very large, so that $c_s^2 k^2 \ll 4\pi G \rho_0$, we can write the dispersion relation approximately as

$$\omega^2 = -4\pi G \rho_0, \quad (490)$$

which implies that

$$\omega = (4\pi G \rho_0)^{1/2} i, \quad (491)$$

and hence that the perturbation grows exponentially as

$$\rho_1 \propto \exp [(4\pi G \rho_0)^{1/2} t]. \quad (492)$$

We see therefore that in this limit, the gas is **gravitationally unstable**: any tiny perturbation away from uniformity will amplify exponentially on a timescale $t_{\text{grav}} = (4\pi G \rho_0)^{-1/2}$, comparable to the gravitational free-fall time of the gas.

- The critical wavenumber at which we move from the oscillatory regime to the gravitationally unstable regime is given by

$$k_J^2 = \frac{4\pi G \rho_0}{c_s^2}. \quad (493)$$

Using this, we can define a corresponding critical length scale, known as the **Jeans length**

$$\lambda_J = \frac{2\pi}{k_J} = \left(\frac{\pi c_s^2}{G \rho_0} \right)^{1/2}, \quad (494)$$

and a critical mass scale, known as the **Jeans mass**

$$M_J = \frac{4\pi}{3} \rho_0 \left(\frac{\lambda_J}{2} \right)^3 = \frac{\pi^{5/2}}{6} \frac{c_s^3}{G^{3/2} \rho^{1/2}}. \quad (495)$$

(Note that this value for M_J is the same, to within a small numerical factor, as we would get if we asked what mass we need in order for the magnitude of the gravitational energy of a perturbation to exceed its thermal energy).

- Perturbations that have sizes larger than the Jeans length or masses larger than the Jeans mass are gravitationally unstable, and hence will undergo gravitational collapse. For $\lambda \sim \lambda_J$ and $M \sim M_J$, pressure forces remain important and collapse proceeds slowly, while for $\lambda \gg \lambda_J$ and $M \gg M_J$, pressure forces are unimportant and collapse proceeds as if the gas were in free-fall.
- This scale-dependence of the collapse rate can be seen more clearly if we rewrite the dispersion relation in terms of a dimensionless growth rate $\Omega = i\omega t_{\text{grav}}$ and a dimensionless wave-number $\nu = k/k_J$. In these dimensionless units, we have

$$\Omega^2 = 1 - \nu^2, \quad (496)$$

demonstrating very clearly that Ω increases for decreasing ν (i.e. increasing physical size of the perturbation) and is a maximum for $\nu = 0$ (i.e. an infinitely large perturbation). In practical terms, this means that in three dimensions, the fastest growing mode of the gravitational instability is the one which corresponds to the overall collapse of the medium.

- An important consequence of this result is that in an approximately spherically symmetrical collapse, it is difficult for gravitational instability to cause **fragmentation** of the gas unless it is seeded on small scales by large density perturbations, since otherwise the small-scale perturbations grow slowly in comparison to the collapse timescale of the gas as a whole.
- Let us now examine what happens if, instead of an infinite uniform medium, we have an infinite thin sheet of matter with initially uniform surface density Σ_0 .
- For a thin sheet, the fluid equations take the form

$$\frac{\partial \Sigma}{\partial t} + \nabla \cdot (\Sigma \vec{v}) = 0, \quad (497)$$

$$\Sigma \left(\frac{\partial \vec{v}}{\partial t} + \vec{v} \cdot \nabla \vec{v} \right) = -\nabla p - \Sigma \nabla \Phi, \quad (498)$$

where $\Sigma = \int_{-\infty}^{+\infty} \rho dz$ is the surface density of the sheet and the gravitational potential Φ satisfies the thin-disk version of Poisson's equation:

$$\nabla^2 \Phi = 4\pi G \Sigma \delta(z), \quad (499)$$

where δ is the Dirac delta function.

- If we apply first-order perturbation theory to this set of equations, we obtain the following equations for our perturbed surface density Σ_1 , velocity \vec{v}_1 and potential Φ_1 :

$$\frac{\partial \Sigma_1}{\partial t} = -\Sigma_0 \nabla \cdot \vec{v}_1, \quad (500)$$

$$\Sigma_0 \frac{\partial \vec{v}_1}{\partial t} = -c_s^2 \nabla \Sigma_1 - \Sigma_0 \nabla \Phi_1, \quad (501)$$

$$\nabla^2 \Phi_1 = 4\pi G \Sigma_1 \delta(z). \quad (502)$$

Note that to derive the last of these equations, we have again used the Jeans swindle to allow us to disregard the unperturbed potential.

- The continuity and momentum equations have a very similar form to their 3D analogues, and thus it is natural to adopt similar trial solutions for Σ_1 and \vec{v}_1 , namely

$$\Sigma_1 = \Sigma_a e^{-i(\vec{k} \cdot \vec{x} + \omega t)}, \quad (503)$$

$$\vec{v}_1 = \vec{v}_a e^{-i(\vec{k} \cdot \vec{x} + \omega t)}, \quad (504)$$

where Σ_a and \vec{v}_a are constants.

- The appropriate solution to use for the potential perturbation requires a little more thought. Our solution must satisfy the perturbed version of Poisson's equation, which motivates the functional form

$$\Phi_1 = \Phi_a e^{-|kz|} e^{-i(\vec{k} \cdot \vec{x} + \omega t)}, \quad (505)$$

where $\Phi_a = 2\pi G \Sigma_a / |k|$.

- Substituting these solutions into the perturbation equations and solving for the dispersion relation linking ω and k yields

$$\omega^2 = c_s^2 k^2 - 2\pi G \Sigma k. \quad (506)$$

Again, we see that when k is small, gravity dominates and the sheet becomes gravitationally unstable, while if k is large, thermal pressure dominates and the sheet is stable. However, the presence of k in the gravitational term does lead to a qualitative difference in the behavior of the instability in this case.

- This change in behavior can be seen more clearly if we once again write the dispersion relation in dimensionless units. If we define a characteristic length scale

$$H = \frac{c_s^2}{\pi G \Sigma}, \quad (507)$$

and use this to define a dimensionless growth rate $\Omega = i\omega H / c_s$ and a dimensionless wavenumber $\nu = kH$, then we can show that

$$\Omega^2 = 2\nu - \nu^2. \quad (508)$$

In order for the perturbation to grow, we require that $\Omega^2 > 0$, which implies that $\nu < 2$. The critical wavenumber at which we first see the instability is therefore $k_{\text{crit}} = 2/H$. If we now consider what happens as we increase the size of the perturbations (i.e. decrease k and ν), we see that while initially the grow rate increases, it soon reaches a maximum value, when $\nu = 1$, and thereafter decreases. We see therefore that, unlike the 3D case, the fastest growing mode in the 2D case occurs when $k = 1/H = k_{\text{crit}}/2$, and that this mode grows much more rapidly than the collapse timescale of the sheet as a whole.

- Consequently, fragmentation in a 2D thin sheet tends to produce fragments with a single characteristic mass $M_{\text{char}} = 4M_{\text{crit}}$, where M_{crit} is the critical mass required for fragmentation.
- Note also that although in our analysis here, we have considered an infinitely thin sheet, a similar result can be shown to apply for the more realistic case of a self-gravitating isothermal sheet.

5.2 Toomre instability

- The original version of the type of instability that has become known as the **Toomre instability** was derived by Toomre in 1964 for a stellar disk, i.e. a rotating disk of (collisionless) stars. However, a very similar instability exists in the gas of a rotating disk of gas, and it is the latter case that we examine here.
- Consider a rotating thin disk, with constant surface density Σ and constant angular velocity Ω_0 . How does the disk respond to the presence of small perturbations?
- A rotating thin disk is a relatively simple generalization of the case of a non-rotating thin sheet that we considered in the previous section. In the rotating case, and in the frame that is co-rotating with the fluid, the fluid equations become

$$\frac{\partial \Sigma}{\partial t} + \nabla \cdot (\Sigma \vec{v}) = 0, \quad (509)$$

$$\frac{\partial \vec{v}}{\partial t} + \vec{v} \cdot \nabla \vec{v} = -\frac{\nabla p}{\Sigma} - \nabla \Phi - 2\vec{\Omega} \times \vec{v} + \Omega^2(x\hat{e}_x + y\hat{e}_y), \quad (510)$$

$$\nabla^2 \Phi = 4\pi G \Sigma \delta(z). \quad (511)$$

The effect of the rotation is to introduce two additional terms in the momentum equation, corresponding to the Coriolis and centrifugal forces, respectively.

- In the co-rotating frame, and in the absence of perturbations, we know that the potential gradient in the x - y plane must exactly balance the centrifugal force, i.e.

$$\nabla \Phi_0 = \Omega^2(x\hat{e}_x + y\hat{e}_y). \quad (512)$$

- Therefore, in this case when we apply first-order perturbation theory to this set of equations, we do not need to rely on the Jeans swindle in order to neglect the $\nabla \Phi_0$

term. Instead, we simply note that it balances the centrifugal force term, and hence eliminate both from the resulting perturbation equations.

- To first order, we therefore have

$$\frac{\partial \Sigma_1}{\partial t} = -\Sigma_0 \nabla \cdot \vec{v}_1, \quad (513)$$

$$\frac{\partial \vec{v}_1}{\partial t} = -\frac{c_s^2}{\Sigma_0} \nabla \Sigma_1 - \nabla \Phi_1 - 2\vec{\Omega} \times \vec{v}_1, \quad (514)$$

$$\nabla^2 \Phi_1 = 4\pi G \Sigma_1 \delta(z). \quad (515)$$

and using the same trial solutions as in the thin-sheet case, we can show that

$$\omega^2 = c_s^2 k^2 + 4\Omega^2 - 2\pi G \Sigma_0 |k|. \quad (516)$$

- In order for the disk to remain stable in the presence of small perturbations, we must have $\omega^2 > 0$. We see therefore that thermal pressure and rotation both act to stabilize the disk, while the self-gravity of the perturbation destabilizes it.
- In addition, we see that the thermal pressure term and the rotation term have very different dependences on scale. The contribution from rotation is independent of scale, while the thermal pressure contribution scales as k^2 , and so is much larger on small scales than on large scales. In comparison, the self-gravity term scales as k , and so grows on small scales, although not as rapidly as the thermal pressure term.
- By comparing the thermal pressure and self-gravity terms, we can show that perturbations with wave-numbers larger than

$$k_J = \frac{2\pi G \Sigma_0}{c_s^2} \quad (517)$$

are stable. (Note that this is the same result as in the non-rotating thin sheet case). Similarly, by comparing the rotation and self-gravity terms, we can show that all perturbations with wave-numbers smaller than

$$k_{\text{rot}} = \frac{2\Omega^2}{\pi G \Sigma_0} \quad (518)$$

are also stable.

- We therefore see that if $k_{\text{rot}} > k_J$, the disk will be stable on all scales. On the other hand, if $k_{\text{rot}} < k_J$, then there is a range of wave-numbers $k_{\text{rot}} < k < k_J$ for which the disk is unstable to the growth of small perturbations.
- The condition that $k_{\text{rot}} > k_J$ can also be written as

$$\frac{2\Omega^2}{\pi G \Sigma_0} > \frac{2\pi G \Sigma_0}{c_s^2}. \quad (519)$$

We can rearrange this inequality to give

$$\left(\frac{\Omega c_s}{\pi G \Sigma_0}\right)^2 > 1, \quad (520)$$

which simplifies to

$$\frac{\Omega c_s}{\pi G \Sigma_0} > 1. \quad (521)$$

Therefore, if we define the **Toomre stability parameter** as

$$Q \equiv \frac{\Omega c_s}{\pi G \Sigma_0}, \quad (522)$$

we see that the disk will be stable if $Q \geq 1$ and unstable if $Q < 1$.

- In our analysis here, we have assumed a fixed angular velocity Ω , but a similar instability holds locally when Ω varies with radius in the disk. However, in this case, the condition for instability becomes

$$Q \equiv \frac{\kappa c_s}{\pi G \Sigma_0} < 1, \quad (523)$$

where κ is the epicyclic frequency of the disk. Similarly, Σ_0 and c_s may also vary with radius, and so in general, Q is a function of radius, meaning that at some radii, the disk may be unstable, while at others it is stable.

- As an example, consider the formation of giant planets in a protoplanetary disk. This is believed to be a consequence of gravitational instability, but close to the star, the disk will be strongly heated by the star, raising c_s and Q and hence suppressing the instability. The fact that we nevertheless see many giant planets at radii that are very close to their host stars, the so-called **hot Jupiters**, is therefore good evidence for some form of migration occurring in the disk.

5.3 Parker instability

- Another important instability that can occur on large scales within galactic disks is the so-called **Parker instability**. For this to operate, we require the gas to be magnetized and that the magnetic field is oriented such that the magnetic field lines are initially parallel to the mid-plane of the galactic disk.
- The first question to address is what is the equilibrium configuration of the field and the gas in this case? In order to analyze this, we make a couple of simplifying assumptions. We assume that the pressure due to non-thermal particles (cosmic rays) is zero, and that the gas is at rest. Neither of these assumptions is strictly true in the real ISM, but they allow us to capture the essence of the instability without unnecessary complications.

- We assume, without loss of generality, that the magnetic field is oriented parallel to the y axis, so that

$$\vec{B}_0 = B_0(z)\hat{e}_y, \quad (524)$$

and that the pressure and density are uniform in the x - y plane. We also assume, following Parker, that the ratio of the thermal to the magnetic pressure, the so-called **plasma β parameter**, is constant:

$$\beta \equiv \frac{8\pi P_0}{B_0^2} = \text{constant}. \quad (525)$$

- If the gas is in equilibrium, then it follows that

$$(1 + \beta^{-1})\frac{dP_0}{dz} = -\rho_0\frac{d\Phi}{dz}, \quad (526)$$

where $\rho_0(z)$ is the density of the unperturbed gas and $\Phi(z)$ is the gravitational potential. We can eliminate the density by using the relationship $\rho_0 = P_0/c_s^2$, and so

$$(1 + \beta^{-1})\frac{dP_0}{dz} = -\frac{P_0}{c_s^2}\frac{d\Phi}{dz}. \quad (527)$$

- Solving this equation yields the following expressions for the thermal pressure and the magnetic field strength

$$P_0(z) = P_0(0) \exp\left[-\frac{\Phi(z)}{(1 + \beta^{-1})c_s^2}\right], \quad (528)$$

$$B_0(z) = B_0(0) \exp\left[-\frac{\Phi(z)}{2(1 + \beta^{-1})c_s^2}\right], \quad (529)$$

where

$$\frac{8\pi P_0(0)}{B_0^2(0)} = \beta. \quad (530)$$

When β is large, meaning that thermal pressure dominates, we therefore have the standard exponential atmosphere, while when β is small, the scale-height of the gas is greatly increased, thanks to the substantial magnetic pressure support.

- In this equilibrium configuration, only magnetic pressure is important; magnetic tension plays no role, because the field lines are straight. However, if we now perturb the gas away from this equilibrium state, then the magnetic tension plays an important role in determining its subsequent evolution.
- We can understand qualitatively what happens if we think of the magnetic field as a “light” fluid held down by a “heavy fluid” (the gas). If we perturb the field upwards, then the magnetic pressure remains the same (initially), but the weight of the gas acting downwards on the field is reduced. The net force therefore acts upwards, further displacing the field – in other words, the field is buoyant.

- As the field starts to move upwards, magnetic tension forces come into play. They suppress the growth of short-wavelength perturbations, since in this case the field curvature is large, and so is the tension force. For very long wavelengths, however, the tension forces are small and are insufficient to stop the growth of the instability.
- As the field starts to buckle, the force on the gas due to gravity is no longer perpendicular to the field lines. It therefore drives a flow of gas along the field lines, allowing the gas to drain from the “peaks” of the magnetic field into the troughs. This reduces the weight of the gas acting to anchor the magnetic field in the peaks, allowing them move even further upwards.
- A detailed analysis of this instability with the tools of perturbation theory shows that the maximum growth rate is roughly the free-fall rate and occurs for perturbations with a wavelength

$$\lambda \simeq \frac{2\pi(1 + \beta^{-1})c_s^2}{g_0}, \quad (531)$$

where g_0 is the acceleration due to gravity at the midplane of the disk. Shorter wavelength perturbations are stabilized by magnetic tension, while longer wavelength perturbations are unstable, but grow more slowly than this mode.

- If we approximate the gravitational potential close to the midplane as $\Phi(z) = g_0|z|$, then the scale-height of the gas distribution implied by Equation 528 is

$$H = \frac{(1 + \beta^{-1})c_s^2}{g_0}, \quad (532)$$

and so the fastest growing mode of the Parker instability has a wavelength that is roughly $\lambda \simeq 2\pi H$.

5.4 Kelvin-Helmholtz instability

- The **Kelvin-Helmholtz instability** is an example of an instability occurring at the interface between two fluids.
- Consider a flow in the x direction made up of two separate fluids. Fluid 1, located in the half-plane $z < 0$ has velocity U_1 and density ρ_1 . Fluid 2, located at $z > 0$, has velocity U_2 and density ρ_2 .
- For simplicity, we assume that the effects of gravity are negligible and that the fluid is incompressible. The latter assumption means that the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) = 0 \quad (533)$$

reduces to the constraint

$$\nabla \cdot \vec{v} = 0. \quad (534)$$

We also assume that the flow is **irrotational**, i.e. that its vorticity $\vec{\omega} = \nabla \times \vec{v} = 0$.

- If $\nabla \cdot \vec{v} = 0$ and $\nabla \times \vec{v} = 0$, then we can write \vec{v} as the gradient of a scalar potential Φ , provided that this potential satisfies

$$\nabla^2 \Phi = 0. \quad (535)$$

We therefore have what is known as a **potential flow**.

- In the present case, we therefore have

$$\vec{v}_2 = \nabla \Phi_2 \quad z > 0, \quad (536)$$

$$\vec{v}_1 = \nabla \Phi_1 \quad z < 0. \quad (537)$$

- We next note that the Euler equation for an incompressible, constant density gas can be written as

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla) \vec{v} = -\nabla \left(\frac{P}{\rho} \right). \quad (538)$$

Using the identity

$$(\vec{v} \cdot \nabla) \vec{v} = \frac{1}{2} \nabla v^2 - \nabla \times \vec{v}, \quad (539)$$

and the assumption that \vec{v} is a potential flow (which implies that $\nabla \times v = 0$), we can rewrite this as

$$\nabla \frac{\partial \Phi}{\partial t} + \frac{1}{2} \nabla v^2 + \nabla \left(\frac{P}{\rho} \right) = 0, \quad (540)$$

which implies that

$$\frac{\partial \Phi}{\partial t} + \frac{1}{2} v^2 + \frac{P}{\rho} = \text{constant}. \quad (541)$$

(This is basically just Bernoulli's equation in slightly different guise).

- In the unperturbed flow, the velocity potentials Φ_1 and Φ_2 are simply $U_1 x$ and $U_2 x$, respectively. However, we want to understand what happens when we perturb the system, so let us consider perturbing the potentials to

$$\Phi_1 = U_1 x + \phi_1, \quad (542)$$

$$\Phi_2 = U_2 x + \phi_2, \quad (543)$$

where $\phi_i \ll \Phi_i$ for $i = 1, 2$. Note that the perturbing potentials must also satisfy $\nabla^2 \phi = 0$.

- We next introduce a function $\xi(x, t)$ that describes the z position of the interface between the two fluids.
- At $t = 0$, we have $\xi(x, 0) = 0$ for all x . At $t > 0$, however, the interface can move if the fluids have a non-zero velocity in the z direction. Since the flow must remain continuous across the interface, the motion of the interface in the z direction is just the same as the motion of the fluid immediately surrounding it, i.e. $\partial \Phi_1 / \partial z$ evaluated at $z = \xi$, or a similar expression for fluid 2.

- In a frame comoving with the fluid in the x direction, we therefore have

$$\frac{D\xi}{Dt} = \frac{\partial\Phi_1}{\partial z}, \quad (544)$$

while in a frame fixed in space, we instead have

$$\frac{\partial\xi}{\partial t} + v_{1,x} \frac{\partial\xi}{\partial x} = \frac{\partial\Phi_1}{\partial z}, \quad (545)$$

where $v_{1,x}$ is the velocity of the flow in the x direction.

- We know that

$$v_{1,x} = \frac{\partial\Phi_1}{\partial x}. \quad (546)$$

Therefore, the evolution of ξ satisfies

$$\frac{\partial\xi}{\partial t} + \frac{\partial\Phi_1}{\partial x} \frac{\partial\xi}{\partial x} = \frac{\partial\Phi_1}{\partial z}. \quad (547)$$

- Now, since

$$\frac{\partial\Phi_1}{\partial x} = U_1 + \frac{\partial\phi_1}{\partial x}, \quad (548)$$

we can also write this as

$$\frac{\partial\xi}{\partial t} + \left(U_1 + \frac{\partial\phi_1}{\partial x} \right) \frac{\partial\xi}{\partial x} = \frac{\partial\Phi_1}{\partial z}. \quad (549)$$

Moreover, since U_1 has no dependence on z , we also know that $\partial\Phi_1/\partial z = \partial\phi_1/\partial z$, implying that

$$\frac{\partial\xi}{\partial t} + \left(U_1 + \frac{\partial\phi_1}{\partial x} \right) \frac{\partial\xi}{\partial x} = \frac{\partial\phi_1}{\partial z}. \quad (550)$$

- In equilibrium, the x derivative of ξ is zero. It only becomes non-zero once we perturb the potential, and if the perturbation is small, the resulting derivative will also be small. Therefore, we can consider the term

$$\frac{\partial\phi_1}{\partial x} \frac{\partial\xi}{\partial x} \quad (551)$$

to be of second order in our perturbed quantities. Therefore, if we carry out our analysis only to linear order, we can neglect this term, leaving us with

$$\frac{\partial\xi}{\partial t} + U_1 \frac{\partial\xi}{\partial x} = \frac{\partial\phi_1}{\partial z}. \quad (552)$$

We can apply the same chain of reasoning in fluid 2, which gives us a second equation for ξ :

$$\frac{\partial\xi}{\partial t} + U_2 \frac{\partial\xi}{\partial x} = \frac{\partial\phi_2}{\partial z}. \quad (553)$$

- Returning to the Bernoulli equation, we note that we can write it as

$$\frac{\partial \Phi_1}{\partial t} + \frac{1}{2} \left[\left(U_1 + \frac{\partial \phi_1}{\partial x} \right)^2 + \left(\frac{\partial \phi_1}{\partial z} \right)^2 \right] + \frac{P_1}{\rho_1} = \frac{1}{2} U_1^2 + \frac{\bar{P}}{\rho_1}, \quad (554)$$

where \bar{P} is the initial pressure. Expanding this to first order in our small perturbation, we get

$$\frac{\partial \phi_1}{\partial t} + \frac{1}{2} U_1^2 + U_1 \frac{\partial \phi_1}{\partial x} + \frac{P_1}{\rho_1} = \frac{1}{2} U_1^2 + \frac{\bar{P}}{\rho_1}, \quad (555)$$

Rearranging this then gives an expression for P_1 :

$$P_1 = \bar{P} - \rho_1 \left(\frac{\partial \phi_1}{\partial t} + U_1 \frac{\partial \phi_1}{\partial x} \right) \quad (556)$$

Similarly, in fluid 2 we have

$$P_2 = \bar{P} - \rho_2 \left(\frac{\partial \phi_2}{\partial t} + U_1 \frac{\partial \phi_2}{\partial x} \right) \quad (557)$$

Finally, since $P_1 = P_2$ (i.e. the system is in pressure equilibrium), we find that

$$\rho_1 \left(\frac{\partial \phi_1}{\partial t} + U_1 \frac{\partial \phi_1}{\partial x} \right) = \rho_2 \left(\frac{\partial \phi_2}{\partial t} + U_1 \frac{\partial \phi_2}{\partial x} \right) \quad (558)$$

- Now that we have the governing equations for ϕ_1 , ϕ_2 and ξ , we can solve for the dispersion relation describing the evolution of the perturbation. We consider the ansatz

$$\phi_1 = \phi_1(z) \exp [i(kx - \omega t)]. \quad (559)$$

Since $\nabla^2 \phi_1 = 0$, we know that

$$\frac{\partial^2 \phi_1(z)}{\partial z^2} \exp [i(kx - \omega t)] + i^2 k^2 \phi_1(z) \exp [i(kx - \omega t)] = 0, \quad (560)$$

and hence that

$$\frac{\partial^2 \phi_1(z)}{\partial z^2} = k^2 \phi_1(z). \quad (561)$$

- This equation has solutions $\phi_1(z) \propto \exp(kz)$ and $\phi_1(z) \propto \exp(-kz)$. However, recall that fluid 1 fills the half-plane with $z < 0$. A long way away from the interface, we expect the fluid to be undisturbed, and hence require that $\phi_1 \rightarrow 0$ as $z \rightarrow -\infty$. This behaviour is satisfied by the first solution, but not the second, and so we can discard the latter.
- A similar line of argument tells us that $\phi_2 \propto \exp(-kz)$. Therefore, for a single Fourier mode:

$$\phi_1 = \hat{\phi}_1 \exp(kz) \exp [i(kx - \omega t)], \quad (562)$$

$$\phi_2 = \hat{\phi}_2 \exp(-kz) \exp [i(kx - \omega t)], \quad (563)$$

$$\xi = \hat{\xi} \exp [i(kx - \omega t)], \quad (564)$$

where $\hat{\phi}_1$, $\hat{\phi}_2$, and $\hat{\xi}$ are the relevant mode amplitudes.

- Substituting these into Equations 552 and 553 gives us

$$-i\omega\hat{\xi} + ikU_1\hat{\xi} = k\hat{\phi}_1 \exp(kz), \quad (565)$$

$$-i\omega\hat{\xi} + ikU_2\hat{\xi} = -k\hat{\phi}_2 \exp(kz). \quad (566)$$

Similarly, from Equation 558, we find that

$$i\rho_1 [kU_1 - \omega] \hat{\phi}_1 \exp(kz) = i\rho_2 [kU_2 - \omega] \hat{\phi}_2 \exp(-kz). \quad (567)$$

- Combining these equations, we find that

$$\frac{i\rho_1}{k} (kU_1 - \omega) (ikU_1 - i\omega)\hat{\xi} = \frac{i\rho_2}{-k} (kU_2 - \omega) (ikU_2 - i\omega)\hat{\xi}. \quad (568)$$

This equation has a trivial solution $\hat{\xi} = 0$, but also has a non-trivial solution if ω and k satisfy

$$\rho_1 (kU_1 - \omega)^2 = -\rho_2 (kU_2 - \omega)^2, \quad (569)$$

$$\rho_1 (k^2U_1^2 - 2kU_1\omega + \omega^2) = -\rho_2 (k^2U_2^2 - 2kU_2\omega + \omega^2). \quad (570)$$

Rearranging this gives us the following quadratic equation for ω :

$$\omega^2 (\rho_1 + \rho_2) - 2k (\rho_1U_1 + \rho_2U_2) \omega + k^2 (\rho_1U_1^2 + \rho_2U_2^2) = 0. \quad (571)$$

- Solving for ω , we find that

$$\omega = \frac{k (\rho_1U_1 + \rho_2U_2)}{\rho_1 + \rho_2} \pm \frac{2k}{2 (\rho_1 + \rho_2)} \sqrt{(\rho_1U_1 + \rho_2U_2)^2 - (\rho_1 + \rho_2) (\rho_1U_1^2 + \rho_2U_2^2)}. \quad (572)$$

Consider the term inside the square root. We can expand this as

$$\rho_1^2U_1^2 + 2\rho_1\rho_2U_1U_2 + \rho_2^2U_2^2 - \rho_1^2U_1^2 - \rho_2^2U_2^2 - \rho_1\rho_2U_2^2 - \rho_2\rho_1U_1^2 = -\rho_1\rho_2 (U_1 - U_2)^2. \quad (573)$$

Therefore, the term inside the square root is always negative if $U_1 \neq U_2$, implying that ω always has an imaginary part.

- Specifically, we find that

$$\omega = \frac{k (\rho_1U_1 + \rho_2U_2)}{\rho_1 + \rho_2} \pm i \frac{\sqrt{\rho_1\rho_2}k}{(\rho_1 + \rho_2)} |U_1 - U_2|. \quad (574)$$

Therefore, the interface between the fluids always displays oscillatory behaviour (unless $U_1 = U_2 = 0$, i.e. the fluid is at rest), but also has both growing and decaying modes whenever $|U_1 - U_2| \neq 0$, i.e. whenever one of the fluids is moving relative to the other one.

- The presence of a growing mode means that in any shear flow of this type, small wave-like perturbations of the interface between the fluids will grow into large “billows”. The subsequent non-linear evolution of these billows (which is not captured by our linear analysis above) leads to them rolling up into vortex-like structures. This instability is the **Kelvin-Helmholtz instability**.

- The Kelvin-Helmholtz instability grows on a characteristic timescale

$$t_{\text{grow}} = \frac{\rho_1 + \rho_2}{\sqrt{\rho_1 \rho_2}} \frac{1}{k |U_1 - U_2|}. \quad (575)$$

In the special case when $\rho_1 = \rho_2$, we have

$$t_{\text{grow}} = \frac{2}{k |U_1 - U_2|}. \quad (576)$$

On the other hand, if $\rho_1 \gg \rho_2$, we have instead

$$t_{\text{grow}} \simeq \frac{\sqrt{\rho_1/\rho_2}}{k |U_1 - U_2|}, \quad (577)$$

while if $\rho_2 \gg \rho_1$, we have

$$t_{\text{grow}} \simeq \frac{\sqrt{\rho_2/\rho_1}}{k |U_1 - U_2|}, \quad (578)$$

- We therefore see that the instability grows most rapidly for fluids that have very similar densities, while if the fluids have very different densities, the growth rate is slower by approximately a factor of the square-root of the density contrast. We also note that $t_{\text{grow}} \propto k^{-1}$, indicating that small modes grow faster than large ones. Consequently, the size of the non-linear vortices created by the Kelvin-Helmholtz instability tend to grow with time.
- Another important point to note is that

$$t_{\text{grow}} \propto \frac{1}{|U_1 - U_2|}. \quad (579)$$

This means that Kelvin-Helmholtz instability occurs far more rapidly when there are sudden changes in the velocity of the shear flow. This also allows us to reason about the effects of viscosity (which was neglected in the above analysis) on the instability. Viscosity will act to reduce the difference between U_1 and U_2 , and we would therefore expect the growth of the Kelvin-Helmholtz instability to occur more slowly in very viscous fluids than in fluids with negligible viscosity.

5.5 Rayleigh-Taylor instability

- In our discussion of the Kelvin-Helmholtz instability, we assumed that the effects of gravity are negligible. Now let's relax this assumption and see what happens.
- We consider the same initial setup as before, with two fluids with densities ρ_1 and ρ_2 and velocities U_1 and U_2 separated by an interface whose position is specified by $\xi(x, t)$. However, this time we assume that there is a homogeneous gravitational field \vec{g} with strength g pointing in the negative z direction.

- In the presence of this field, Euler's equation becomes

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla) \vec{v} = -\nabla \left(\frac{P}{\rho} \right) + \vec{g}, \quad (580)$$

and the Bernoulli equation becomes

$$\frac{\partial \Phi}{\partial t} + \frac{1}{2} v^2 + \frac{P}{\rho} + gz = \text{constant}. \quad (581)$$

- Our derivation of the equations linking ξ , ϕ_1 and ϕ_2 relies only on the velocity across the interface being constant and hence is the same regardless of whether or not a gravitational field is present. Therefore, in this case the equations

$$\frac{\partial \xi}{\partial t} + U_1 \frac{\partial \xi}{\partial x} = \frac{\partial \phi_1}{\partial z}, \quad (582)$$

$$\frac{\partial \xi}{\partial t} + U_2 \frac{\partial \xi}{\partial x} = \frac{\partial \phi_2}{\partial z} \quad (583)$$

remain valid.

- When $g \neq 0$, however, the third of our equations, derived from the Bernoulli equation, takes the form

$$\frac{\partial \Phi_1}{\partial t} + \frac{1}{2} \left[\left(U_1 + \frac{\partial \phi_1}{\partial x} \right)^2 + \left(\frac{\partial \phi_1}{\partial z} \right)^2 \right] + \frac{P_1}{\rho_1} + g\xi = \frac{1}{2} U_1^2 + \frac{\bar{P}}{\rho_1}, \quad (584)$$

where the absence of the gravitational term on the right-hand side comes from the fact that the interface starts at $z = 0$ in the unperturbed state. Expanding this to first order as before, we get

$$\frac{\partial \phi_1}{\partial t} + U_1 \frac{\partial \phi_1}{\partial x} + \frac{P_1}{\rho_1} + g\xi = \frac{\bar{P}}{\rho_1}, \quad (585)$$

and hence

$$P_1 = \bar{P} - \rho_1 \left(\frac{\partial \phi_1}{\partial t} + U_1 \frac{\partial \phi_1}{\partial x} + g\xi \right). \quad (586)$$

Similarly,

$$P_2 = \bar{P} - \rho_2 \left(\frac{\partial \phi_2}{\partial t} + U_2 \frac{\partial \phi_2}{\partial x} + g\xi \right). \quad (587)$$

- Previously, we assumed that the fluid was in pressure equilibrium and hence equated P_1 and P_2 . Is this assumption still valid in the presence of a gravitational field? Clearly, on large scales there must be a pressure gradient in the vertical direction when the flow is in equilibrium, to balance the gravitational attraction. However, on small scales (i.e. scales much smaller than the scale height), close to the interface, we can still assume that P_1 and P_2 are approximately constant, and hence it is still valid to equate them. Doing so yields

$$\rho_1 \left(\frac{\partial \phi_1}{\partial t} + U_1 \frac{\partial \phi_1}{\partial x} + g\xi \right) = \rho_2 \left(\frac{\partial \phi_2}{\partial t} + U_2 \frac{\partial \phi_2}{\partial x} + g\xi \right). \quad (588)$$

- If we now consider the same trial solutions as before, we find that

$$-i\omega\hat{\xi} + ikU_1\hat{\xi} = k\hat{\phi}_1 \exp(kz), \quad (589)$$

$$-i\omega\hat{\xi} + ikU_2\hat{\xi} = -k\hat{\phi}_2 \exp(kz), \quad (590)$$

and

$$i\rho_1 [kU_1 - \omega] \hat{\phi}_1 \exp(kz) + \rho_1 g \hat{\xi} = i\rho_2 [kU_2 - \omega] \hat{\phi}_2 \exp(-kz) + \rho_2 g \hat{\xi}. \quad (591)$$

- Combining these equations yields the dispersion relation valid in the case $g \neq 0$:

$$\omega^2 (\rho_1 + \rho_2) - 2k (\rho_1 U_1 + \rho_2 U_2) \omega + k^2 (\rho_1 U_1^2 + \rho_2 U_2^2) + (\rho_2 - \rho_1)kg = 0. \quad (592)$$

- Now consider the special case where the fluid is at rest, i.e. where $U_1 = U_2 = 0$. In this case, the dispersion relation reduces to the form

$$\omega^2 (\rho_1 + \rho_2) + (\rho_2 - \rho_1)kg = 0. \quad (593)$$

Therefore,

$$\omega^2 = \frac{(\rho_1 - \rho_2)kg}{\rho_1 + \rho_2}. \quad (594)$$

If $\rho_1 > \rho_2$ (i.e. the lower fluid is denser than the upper fluid), then $\omega^2 > 0$ and so this arrangement is stable for all k . On the other hand, if $\rho_2 > \rho_1$ (i.e. the denser fluid lies on top), then $\omega^2 < 0$ and the arrangement is **unstable** for all k . This instability, which is driven by the buoyancy of the lighter fluid, is known as the **Rayleigh-Taylor instability**.

5.6 Thermal instability

- The final instability that we consider here has nothing to do with the gravity of the gas, but instead is driven entirely by the interplay between thermal pressure and radiative heating and cooling. This is the **thermal instability**.
- Suppose that the net cooling rate of a parcel of gas is given by

$$\dot{Q} = n^2\Lambda - n\Gamma, \quad (595)$$

where Λ is the cooling rate per unit volume and Γ is the heating rate. [NB. We use this form for \dot{Q} because most heating processes are proportional to a single power of the density (e.g. cosmic ray heating or the photoelectric effect), while most cooling processes are proportional to the density squared].

- In thermal equilibrium, heating balances cooling and so we simply have $\dot{Q} = 0$.
- Now consider what happens when we perturb this equilibrium state. We will consider two types of perturbations: **isochoric** perturbations, in which we hold the density constant while changing the temperature, and **isobaric** perturbations in which we hold the pressure constant while changing the temperature.

- Suppose we increase the temperature. For our equilibrium to be thermally stable, this must lead to $\dot{Q} > 0$, so that the gas cools back to the equilibrium state. If, on the other hand, it leads to $\dot{Q} < 0$, the gas will continue to heat up, moving further and further from equilibrium – it will be **thermally unstable**.
- Similarly, if we decrease the temperature, then we need $\dot{Q} < 0$ in order for the gas to return to equilibrium. Our condition for thermal instability is hence that:

$$\frac{\partial \dot{Q}}{\partial T} < 0. \quad (596)$$

For an isochoric perturbation, we have the **Parker criterion**:

$$\left[\frac{\partial \dot{Q}}{\partial T} \right]_{\rho} < 0. \quad (597)$$

On the other hand, for isobaric perturbations we have the **Field criterion**:

$$\left[\frac{\partial \dot{Q}}{\partial T} \right]_P < 0. \quad (598)$$

- To see what this implies in terms of the temperature dependence of the cooling function, let us suppose that we can approximate Λ and Γ as power-laws:

$$\Lambda = \Lambda_0 T^{\alpha}; \quad \Gamma = \Gamma_0 T^{\beta}. \quad (599)$$

- The Parker criterion then becomes:

$$\alpha \frac{n^2 \Lambda}{T} - \beta \frac{n \Gamma}{T} < 0. \quad (600)$$

We can rearrange this to obtain:

$$\alpha (\Lambda n^2 - \Gamma n) + (\alpha - \beta) \Gamma n < 0. \quad (601)$$

- In equilibrium, the term in the first set of brackets is zero, and hence the condition for isochoric thermal instability becomes:

$$(\alpha - \beta) \Gamma n < 0. \quad (602)$$

Since $\Gamma n > 0$, this implies that the medium will be isochorically unstable whenever $\alpha < \beta$.

- In the local ISM, $\beta \simeq 0$, and our condition for isochoric thermal instability becomes a constraint on the temperature dependence of our cooling rate. If we assume that at temperatures $T \ll 10^4$ K this is dominated by fine structure emission from C^+ , then $\Lambda \propto T^{0.1} e^{-92/T}$ when C^+ -H collisions dominate and $\Lambda \propto T^{-1/2} e^{-92/T}$ when C^+ -electron collisions dominate. Below a few hundred K, we therefore always have $\alpha > 0$ and the medium is isochorically stable. However, at higher temperatures, it can be isochorically unstable if collisions with electrons dominate the C^+ cooling rate.

- What about isobaric instability? If we ignore any changes in composition, then $P = \text{const}$ implies that $n \propto T^{-1}$. Therefore, the Field criterion is:

$$(\alpha - 2)\frac{n^2\Lambda}{T} - (\beta - 1)\frac{n\Gamma}{T} < 0 \quad (603)$$

Hence we have:

$$(\alpha - 2)(\Lambda n^2 - \Gamma n) + (\alpha - 2 - \beta + 1)n\Gamma < 0, \quad (604)$$

which for a gas initially in thermal equilibrium reduces to the condition:

$$\alpha - \beta - 1 < 0. \quad (605)$$

For $\beta = 0$, we therefore have thermal instability if $\alpha < 1$, i.e. if the cooling function increases with increasing temperature more slowly than $\Lambda \propto T$.

- In the local ISM, this condition is generally satisfied for all temperatures in the range $100 < T < 6000$ K, regardless of whether electron collisions or hydrogen collisions dominate. Therefore, the gas is always isobarically unstable in this regime.
- What is the size of the structures created by isobaric thermal instability? For the gas to be isobaric, it must be able to maintain pressure equilibrium internally. This means that these structures have a maximum size that is approximately:

$$L_{\text{iso}} = c_s t_{\text{cool}}, \quad (606)$$

where t_{cool} is the cooling time of the gas. For a gas with $T = 100$ K, we have $\Lambda_{\text{C}^+} = 1.1 \times 10^{-27} \text{ erg cm}^3 \text{ s}^{-1}$ (where we have assumed solar metallicity and that hydrogen collisions dominate), and hence the cooling time is approximately

$$t_{\text{cool}} = \frac{1.5nkT}{\Lambda_{\text{C}^+}n^2}, \quad (607)$$

$$= \frac{2.1 \times 10^{-14}}{1.1 \times 10^{-27}} n^{-1}, \quad (608)$$

$$= 0.6n^{-1} \text{ Myr}. \quad (609)$$

The sound speed for $T = 100$ K is approximately 1 km s^{-1} , and the isobaric length scale is thus

$$L_{\text{iso}} \simeq 0.6 n^{-1} \text{ pc}. \quad (610)$$

- The minimum size of the isobaric perturbations is set by another physical effect, thermal conduction. As the perturbation cools, it becomes colder than the surrounding gas. Heat will therefore flow into it via thermal conduction. If the heat flow into the perturbation balances the radiative losses then the perturbation will not cool and the instability will be suppressed.
- To estimate the length scale on which this becomes important, we start by recalling the heat flow equation:

$$\vec{q} = -\kappa \vec{\nabla} T. \quad (611)$$

Here, \vec{q} is the flow of heat through the surface of the perturbation and κ is the thermal conductivity of the gas.

- We can write this equation in integral form as:

$$\dot{Q} = -\kappa \oint \vec{\nabla}T \cdot d\vec{S}. \quad (612)$$

We can get a crude estimate of the value of \dot{Q} by noting that $\nabla T \sim T/L$, where L is the size of the perturbation. Hence:

$$|\dot{Q}| \sim \kappa \times \frac{T}{L} \times L^2 \sim \kappa T L. \quad (613)$$

- For thermal conduction to balance cooling, we must have $n^2 \Lambda L^3 \sim \kappa T L$. From this, we can derive an estimate for the minimum length scale that our perturbation can have:

$$L_{\min} = \left(\frac{\kappa T}{n^2 \Lambda} \right)^{1/2}. \quad (614)$$

This minimum length scale is known as the **Field length**, and is often written as L_F .

- When the ionization of the gas is low (or the magnetic field is strong), the dominant contribution to the thermal conductivity comes from atomic hydrogen. In this case, we have

$$\kappa = 2.5 \times 10^3 T^{1/2} \text{ cm}^{-1} \text{ K}^{-1} \text{ s}^{-1}. \quad (615)$$

If the gas temperature is 100 K, as before, then this implies that $L_F \sim 0.02 n^{-1}$ pc, significantly smaller than L_{iso} .

- The isobaric thermal instability can therefore create dense structures in the gas with a wide range of scales, $L_F < L < L_{\text{iso}}$. However, these structures are generally not self-gravitating. To see this, note that in order for a density perturbation to be gravitationally unstable, its size must be larger than the Jeans length L_J . Therefore, in order for the initial small isobaric perturbation to be gravitationally unstable, it must have $L > L_J$, which is possible only if $L_{\text{iso}} > L_J$. To within a small numerical factor, we have $L_{\text{iso}} = c_s t_{\text{cool}}$ and $L_J = c_s t_{\text{ff}}$, where t_{ff} is the gravitational free-fall timescale. Therefore, the initial perturbation is unstable only if $t_{\text{cool}} > t_{\text{ff}}$.
- Typically, in conditions where thermal instability is important, cooling is efficient and $t_{\text{cool}} \ll t_{\text{ff}}$. Therefore, in these conditions, the initial isobaric perturbations are not gravitationally unstable.
- As the perturbations evolve, their temperatures decrease and their densities increase. As $L_J \propto T^{1/2} n^{-1/2}$, this leads to a decrease in the Jeans length. However, their physical dimensions also decrease, scaling as $L \propto n^{-1/3}$ if the perturbation is roughly spherical. The ratio of the Jeans length to the size of the perturbation therefore scales as

$$\frac{L_J}{L} \propto T^{1/2} n^{-1/6}. \quad (616)$$

Since $n \propto T^{-1}$ for an isobaric perturbation, this implies that

$$\frac{L_J}{L} \propto T^{2/3}. \quad (617)$$

Therefore, if $L_J \ll L$ initially, the perturbation will become unstable only if the change in temperature is very large. However, in most circumstances in which gas is susceptible to thermal instability, the temperature changes by no more than a factor of 100, and often by much less. Since $100^{2/3} \sim 20$, this implies that if $t_{\text{cool}} < 0.05t_{\text{ff}}$ initially, then $L > L_J$ throughout the evolution of the isobaric perturbation. To give some concrete numbers, for the local ISM, we have $t_{\text{cool}} < 0.01t_{\text{ff}}$ in the CNM, and an even smaller ratio in warmer gas, and so in this case, the perturbations remain stable.

- On the other hand, isochoric perturbations can become gravitationally unstable, since they are not constrained to have sizes $L < L_{\text{iso}}$. In addition, the strong pressure gradients that are created during the initial cooling phase can have a large effect on the dynamics of the gas, and in particular can be an efficient way of generating turbulence in the cold gas.