Jeans Instability and Gravitational Collapse

Gravity introduces a new element into the physics of sound waves (aka stable perturbations) in a gas. Simply put, a gas cloud's self gravity can cause perturbations on sufficiently large scales to become *unstable*. The goal here is to present a brief overview of fluid dynamics, and to give both rigorous and non-rigorous derivations of the Jeans criterion (named after Sir James Jeans, who first published this analysis in 1902) for gravitational instability.

We set the stage by first considering sound waves in a non-gravitating gas, establishing some basic methodology and results. Then we include gravity, and ask how it changes the picture.

Fluid Dynamics 101

We consider a fluid in which quantities such as the number and mass densities $n(\mathbf{x},t)$ and $\rho(\mathbf{x},t)$, the pressure $P(\mathbf{x},t)$, the temperature $T(\mathbf{x},t)$, and the velocity field $\mathbf{v}(\mathbf{x},t)$ can all sensibly be defined at any point $\mathbf{x}$. As a practical matter, this means that the fluid must be such that the local interparticle spacing $\ell = n^{-1/3}$ is much less than the scale of interest λ. Many (but *not* all) astrophysical systems fall into this category.

The behavior of the fluid is governed by three fundamental equations. First, the **continuity equation** is a statement of conservation of mass at any point:

$$\frac{\partial \rho}{\partial t} \;+\; \nabla\cdot(\rho\mathbf{v}) \;=\; 0. \tag{1}$$

Integrating this equation over any small volume V surrounding the point $\mathbf{x}$ and applying the divergence theorem to the second term, this equation just says that the rate of change of $\int_V \rho dV$, the total mass within V, is equal to $\oint_S \rho\mathbf{v}\cdot .d\mathbf{S}$, the rate at which mass flows outward across the surface of V.

The second fundamental equation is the **Euler equation**, which is just the fluid equivalent of Newton's second law:

$$\frac{\partial \mathbf{v}}{\partial t} \;+\; \mathbf{v}\cdot\nabla\mathbf{v} \;=\; -\,\frac{\nabla P}{\rho}. \tag{2}$$

The right-hand side is just the pressure gradient, the force per unit volume on the fluid, divided by the density, the mass per unit volume. The left-hand side is the *acceleration* of the fluid. We can understand the latter statement by considering the rate of change of any quantity $Q(\mathbf{x},t)$ that describes the fluid. As we move with a fluid element, Q changes for two reasons. First, the value of Q at any point $\mathbf{x}$ is changing if $\partial Q/\partial t \neq 0$ there. Second, as the fluid moves, the element samples different regions of space, which will have different values of Q if ∇Q is nonzero. Applying the chain rule for partial differentiation, we have, writing $\mathbf{x} = (x,y,z)$,

$$Q(x+\delta x, y+\delta y, z+\delta z, t+\delta t) - Q(x,y,z,t) \;\approx\; \frac{\partial Q}{\partial x}\delta x + \frac{\partial Q}{\partial y}\delta y + \frac{\partial Q}{\partial z}\delta z + \frac{\partial Q}{\partial t}\delta t,$$

so

$$\begin{aligned}
\frac{dQ}{dt} \quad &\equiv \quad \lim_{\delta\to 0}\frac{Q(x+\delta x, y+\delta y, z+\delta z, t+\delta t) - Q(x,y,z,t)}{\delta t} \\
&= \quad \frac{\partial Q}{\partial x}v_x + \frac{\partial Q}{\partial y}v_y + \frac{\partial Q}{\partial z}v_z + \frac{\partial Q}{\partial t} \\
&= \quad \mathbf{v}\cdot\nabla Q + \frac{\partial Q}{\partial t},
\end{aligned}$$

where $\mathbf{v} = (v_x, v_y, v_z) = \lim_{\delta t \to 0} \delta\mathbf{x}/\delta t$. The quantity dQ/dt, the rate of change of Q moving with the fluid, is sometimes called the *convective derivative* of Q. Setting $Q = \mathbf{v}$, we see that the right hand side of Equation (2) is simply $d\mathbf{v}/dt$. Note that the right-hand side of Equation (2) as written contains only the pressure force. Other forces, such as magnetic fields or internal viscous terms, could also appear there, but we will neglect them here.

Equations (1) and (2) represent 4 equations (1 scalar, 1 vector) in 5 unknowns (ρ, P, and $\mathbf{v}$). We need another equation to find a solution. That equation is the equation of state of the fluid

$$P = P(\rho, S) \tag{3}$$

where S is entropy. We will confine ourselves here to isentropic flows, with $S =$ constant.

These equations have many interesting solutions, but our interest here is more limited. To start, let's consider a uniform fluid and look at the behavior of small perturbations to that background solution. Specifically, the unperturbed fluid motion is steady, $\partial/\partial t \equiv 0$, with uniform density and pressure ρ_0 and P_0, and zero velocity $\mathbf{v}_0 = \mathbf{0}$. You can easily verify that this solution trivially satisfies Equations (1) and (2).

Now let's consider a small perturbation to this baseline solution. We write

$$\begin{aligned} \rho &= \rho_0 + \rho_1 \\ P &= P_0 + P_1 \\ \mathbf{v} &= \mathbf{v}_1 \end{aligned}$$

and substitute these into Equations (1) and (2):

$$\frac{\partial \rho_0}{\partial t} + \frac{\partial \rho_1}{\partial t} + \nabla \cdot [(\rho_0 + \rho_1)\mathbf{v}_1] = 0 \tag{4}$$

$$\frac{\partial \mathbf{v}_1}{\partial t} + \mathbf{v}_1 \cdot \nabla \mathbf{v}_1 = -\frac{\nabla(P_0 + P_1)}{\rho_0 + \rho_1}. \tag{5}$$

We then *linearize* these equations by expanding them to first order, throwing away all products of small terms (having subscript 1). This is the basic mathematical technique employed in all studies of the stability of nonlinear dynamical systems—find an equilibrium solution, then look at the first-order equations describing small perturbations to that solution. (Why? Because we know how to solve linear equations and they give us valuable insight into the physics of the system.)

In this case, discarding all products involving more than one small term, and imposing the properties of the unperturbed solution, Equations (4) and (5) become

$$\frac{\partial \rho_1}{\partial t} + \rho_0 \nabla \cdot \mathbf{v}_1 = 0 \tag{6}$$

$$\rho_0 \frac{\partial \mathbf{v}_1}{\partial t} = -\nabla P_1. \tag{7}$$

Note that the tricky nonlinear $\mathbf{v} \cdot \nabla \mathbf{v}$ term in the Euler equation, which is the source of a host of complex physical phenomena in real fluids, ranging from turbulent dissipation to shock waves, has disappeared.

The key point is that Equations (6) and (7) are soluble. Taking the partial derivative of Equation (6) with respect to time and taking the divergence of Equation (7), and noting that $\partial(\nabla \cdot \mathbf{v}_1)/\partial t = \nabla \cdot (\partial \mathbf{v}_1/\partial t)$, we can eliminate the $\mathbf{v}_1$ terms and find

$$\frac{\partial^2 \rho_1}{\partial t^2} - \nabla^2 P_1 = 0. \tag{8}$$

We now use the equation of state, Equation (3), to obtain

$$P_1 \equiv \delta P = \left.\frac{\partial P}{\partial \rho}\right|_{S,0} \delta\rho = \left.\frac{\partial P}{\partial \rho}\right|_{S,0} \rho_1 \,,$$

so Equation (8) becomes

$$\frac{\partial^2 \rho_1}{\partial t^2} - \left.\frac{\partial P}{\partial \rho}\right|_{S,0} \nabla^2 \rho_1 = 0. \tag{9}$$

We recognize Equation (9) as the *wave equation*, the solutions to which are traveling waves with wave speed c_s, where

$$c_s^2 = \left.\frac{\partial P}{\partial \rho}\right|_{S,0}. \tag{10}$$

For an ideal gas with $P = \rho kT/m = A\rho^\gamma$, the adiabatic sound speed is

$$\begin{aligned} c_s &= \sqrt{A\gamma\rho^{\gamma-1}} = \sqrt{\frac{\gamma P}{\rho}} \\ &= \sqrt{\frac{\gamma kT}{m}} \\ &= 0.91\,\mathrm{km/s}\ \gamma^{1/2} \left(\frac{T}{100\,\mathrm{K}}\right)^{1/2} \left(\frac{m}{m_H}\right)^{-1/2}, \end{aligned} \tag{11}$$

where m is the mean molecular weight.

Linearity means that any linear combination of solutions is also a solution. In that case, since any solution can be decomposed into a Fourier series (for a finite region) or a Fourier transform (for an infinite domain), either of which is just a sum of plane-wave solutions of the form

$$\rho_1 \propto e^{i\mathbf{k}\cdot\mathbf{x} - i\omega t} \tag{12}$$

(and similarly for other quantities), we can understand the physics of the system by considering the behavior of plane-wave modes. Here $\omega = 2\pi f$ is the angular frequency and $\mathbf{k}$ is the *wave vector*, whose direction is the direction of motion of the plane wave and whose magnitude (the *wavenumber*) is $k = 2\pi/\lambda$, where λ is the wavelength.

Substituting Equation(12) into Equation (9) we obtain the familiar relation between wavenumber and frequency:

$$\omega = c_s\, k \tag{13}$$

or

$$f\lambda = c_s. \tag{14}$$

Self-gravitating fluids

Now let's consider how the above analysis changes when gravity is included. Gravity leaves the continuity equation (1) unchanged, and introduces an additional force into the right-hand side of the Euler equation (2), which now becomes

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v} = -\frac{\nabla P}{\rho} - \nabla\phi, \tag{15}$$

where the gravitational acceleration is $\mathbf{a} = -\nabla\phi$ and the gravitational potential ϕ satisfies Poisson's equation

$$\nabla^2\phi \;=\; 4\pi G\rho. \tag{16}$$

In the spirit of the earlier approach, we seek a steady ($\partial/\partial t = 0$) unperturbed solution with uniform density ρ_0 and pressure P_0, and zero velocity. We note in passing that this formally leads to a contradiction for the potential: Substituting the "0" quantities into the Euler equation implies $\nabla\phi_0 = 0$, so $\phi_0 =$ constant, but Poisson's equation (16) says $\nabla^2\phi > 0$. This contradiction is sometimes called the *Jeans swindle*. Binney & Tremaine (p. 402) discuss some arguments to get around it in practice; see also the article by Kiessling (2003, Adv.Appl.Math. 31, 132; arXiv:astro-ph/9910247v1).

In any case, this technical problem with the unperturbed state is not regarded as a serious flaw in the perturbation analysis. We might imagine that other forces, such as rotation or magnetism, are also at play in the initial state, allowing a self-consistent solution to exist. For example, consider a homogeneous self-gravitating fluid of density ρ contained within a rotating cylinder of radius R. The cylinder and the fluid rotate at angular speed Ω about the axis of the cylinder. It is readily shown from Equations (15) and (16) that the fluid can be in true equilibrium, with no pressure gradient (i.e. the gravitational and centrifugal forces balance), if $\Omega^2 = 2\pi G\rho$.

Ignoring the details of the Jeans swindle and writing

$$\begin{aligned} \rho &= \rho_0 + \rho_1 \\ P &= P_0 + P_1 \\ \mathbf{v} &= \mathbf{v}_1 \\ \phi &= \phi_0 + \phi_1, \end{aligned}$$

substituting into Equations (1), (15), and (16), and again using Equation (10) to eliminate P_1, we obtain the linearized equations

$$\frac{\partial\rho_1}{\partial t} + \rho_0\nabla\cdot\mathbf{v}_1 \;=\; 0 \tag{17}$$

$$\rho_0\frac{\partial\mathbf{v}_1}{\partial t} \;=\; -c_s^2\nabla\rho_1 - \nabla\phi_1. \tag{18}$$

$$\nabla^2\phi_1 \;=\; 4\pi G\rho_1 \tag{19}$$

Combining, as before, the time derivative of Equation (17) with the divergence of Equation (18) and employing Equation (19) to eliminate $\nabla^2\phi_1$, we arrive at

$$\frac{\partial^2\rho_1}{\partial t^2} - c_s^2\nabla^2\rho_1 \;=\; -\,4\pi G\rho_0\rho_1, \tag{20}$$

which is identical to Equation (9) apart from the additional term on the right-hand side.

Again looking for plane-wave solutions, we substitute Equation (12) into Equation (20) to obtain the new relation between wavenumber and frequency:

$$\omega^2 \;=\; c_s^2\,k^2 - 4\pi G\rho_0\,. \tag{21}$$

The difference between Equations (13) and (21) is profound. In the former case, ω^2 is positive (and so ω is real) for any k. In the latter, since the second term on the right is negative, it is possible for ω^2 to be *negative* for some k. What does this mean? Setting $\omega^2 = -\alpha^2$, we have $\omega = \pm i\alpha$ and $e^{i\omega t} = e^{\mp\alpha t}$. In other words, the oscillatory behavior found for real ω is replaced by exponential growth for imaginary ω—the disturbance is *unstable*. Any wavenumber k leading to $\omega^2 < 0$ represents a perturbation whose amplitude grows exponentially in time. An overdense region becomes denser and denser, leading to gravitational collapse.

From Equation (21), it is clear that all plane-wave perturbations having

$$k \ < \ k_J \ \equiv \ \frac{\sqrt{4\pi G\rho_0}}{c_s} \tag{22}$$

are unstable. The critical wavenumber k_J is called the **Jeans wavenumber**. In terms of wavelength $\lambda = 2\pi/k$, this means that all perturbations having

$$\lambda \ > \ \lambda_J \ \equiv \ c_s\sqrt{\frac{\pi}{G\rho_0}} \tag{23}$$

are unstable. The critical wavelength λ_J is the **Jeans length**. All scales larger than λ_J are gravitationally unstable.

It is common to recast the above criteria in terms of the total mass involved. The mass contained within a spherical volume having a diameter equal to the Jeans length is called the **Jeans mass**, M_J, where

$$M_J \ \equiv \ \frac{4\pi}{3}\rho_0\left(\tfrac{1}{2}\lambda_J\right)^3 \ = \ \frac{\pi^{5/2}}{6}\,\frac{c_s^3}{G^{3/2}\rho_0^{1/2}} \tag{24}$$

Some Simpler Arguments

Armed with the exact solution, we can explore some simpler, "back of the envelope" arguments to understand the result.

First, as discussed in Binney & Tremaine, let's imagine compressing a spherical region of radius R in a homogeneous gas cloud. If $R \to (1-\alpha)R$, where α is small and positive, and the mass M remains constant, then $\rho_0 \propto M/R^3 \to (1+3\alpha)\rho_0$. From Equation (10), the pressure change is $\delta P = c_s^2\delta\rho = 3\alpha c_s^3\rho_0$ and so—to order of magnitude at least—the pressure gradient between the center of the region and the edge is $\sim \delta P/R = 3\alpha c_s^3\rho_0/R$ and the (outward) pressure force per unit mass is

$$a_p \ \approx \ \frac{\delta P}{\rho_0 R} \ \approx \ \frac{3\alpha c_s^2}{R}. \tag{25}$$

As the region is compressed, the gravitational acceleration at its surface, GM/R^2, also increases. The extra (inward) gravitational acceleration due to the compression is

$$a_g \ \approx \ \frac{2GM\alpha}{R^2} \ = \ \frac{8\pi G\rho_0 R\alpha}{3}. \tag{26}$$

where we have used the fact that $M = \frac{4}{3}\pi R^3\rho_0$. Comparing Equations (25) and (26), we see that the gravitational force dominates (and continues to dominate) if $a_p < a_g$, or

$$R^2 \ > \ \frac{9}{8\pi}\frac{c_s^2}{G\rho_0}. \tag{27}$$

Apart from the numerical coefficient, we see that the critical value of R is just the Jeans length (Equation 23).

Following Sparke & Gallagher (p. 355) we can reach the same conclusion by considering the thermal and potential energies of the region. The total gravitational potential energy of a homogeneous sphere of radius R and mass $M = \frac{4\pi}{3}R^3\rho_0$ is easily shown to be $U = -3GM^2/5R \propto R^5$, while the thermal energy (for an ideal gas) is $K = \frac{3}{2}MkT/m = (3/\gamma)\frac{1}{2}Mc_s^2 \propto R^3$, from Equation (11). Comparing the two, we see that the total energy $U + K$ becomes negative (i.e. the region's gravity dominates its thermal energy) for

$$\frac{3}{2\gamma}Mc_s^2 \quad < \quad \frac{3}{5}\frac{GM^2}{R}$$

so

$$c_s^2 \quad < \quad \frac{2\gamma}{5}\frac{GM}{R} \quad = \quad \frac{8\pi\gamma}{15}G\rho_0 R^2$$

or

$$R^2 \quad > \quad \frac{15}{8\pi\gamma}\frac{c_s^2}{G\rho_0}, \tag{28}$$

again essentially the same result as Equations (23) and (27).

Finally, we note that the instability criterion (Equation 23) can be rewritten (again neglecting the numerical factors) as

$$\frac{R}{c_s} \;\gtrsim\; (G\rho_0)^{-1/2}. \tag{29}$$

The right-hand side is the *free-fall* time for the clump—the time scale for it to collapse under gravity if pressure were negligible—while the left-hand side is the time taken for a sound wave to cross the system. The Jeans criterion can thus be interpreted as meaning that sound cannot traverse the region (and hence pressure cannot operate) in time to prevent the collapse.

All of the above examples confirm that the Jeans instability is the result of competition between thermal (pressure/sound speed) forces and gravitational forces. In each case, however the gravity is quantified, gravitational collapse ensues when gravity dominates over the thermal term.

Consequences of the Jeans Instability

Let's evaluate the Jeans length and mass, Equations (23) and (24), for parameters of astrophysical interest. Plugging in numbers typical of dense molecular cores (with particle mass $m = 3.3 \times 10^{-24}$ g), we obtain

$$\lambda_J \quad = \quad 1.0\,\mathrm{pc}\,\left(\frac{T}{10\,\mathrm{K}}\right)^{1/2}\left(\frac{n}{10^3\,\mathrm{cm}^{-3}}\right)^{-1/2} \tag{30}$$

$$M_J \quad = \quad 26\,M_\odot\,\left(\frac{T}{10\,\mathrm{K}}\right)^{3/2}\left(\frac{n}{10^3\,\mathrm{cm}^{-3}}\right)^{-1/2}, \tag{31}$$

where $c_s = 260$ m/s for $T = 10$ K and $\gamma = 5/3$, although given the effectiveness of cooling in maintaining constant temperature, a better approximation might be the isothermal $\gamma = 1$, as assumed in S&G, in which case $c_s \approx 200$ m/s. Thus any dense molecular core containing more than a few tens of solar masses of gas is unstable, and will collapse in roughly a free-fall time

$$t_{ff} \;=\; (G\rho)^{-1/2} \;=\; 2.1\,\mathrm{Myr}\,\left(\frac{n}{10^3\,\mathrm{cm}^{-3}}\right)^{-1/2}. \tag{32}$$

As the unstable region collapses its density rises, but the atomic and molecular processes cooling it remain efficient so long as the cloud is optically thin to the cooling radiation, and the temperature remains roughly constant. As a result, the Jean's mass steadily decreases as the collapse proceeds, and the collapsing cloud *fragments* into lower and lower mass pieces, each collapsing on its own free-fall time scale. Our initial collapsing cloud is on its way to becoming a star cluster.

The process ends when the fragments become so dense that they are optically thick to the radiation cooling them, so the radiation can no longer escape and the temperature begins to rise, stabilizing the collapse and forming a *protostar*. Experts and simulations disagree on precisely when this point is reached, and on the subsequent evolution of the protostars, but this is the subject of active research. At stake is our understanding of the stellar mass function and the physical state of young star clusters.