Derive the onset of instability for Rayleigh–Bénard convection

This document is a walk-through of each step in deriving the onset of Rayleigh–Bénard instability. The main purpose of this note is to introduce a general procedure to solving instability problems, and how to use nondimensionalization and linear stability theory. Some of the calculations may be too technical to fluid dynamics; please let the TF know if you find some steps difficult to understand.

Step 1: Set up the problem

When solving a fluid dynamics problem, it is always advisable to illustrate the setup and write down all the relevant parameters and variables, as well as any assumptions. The same idea applies to solid mechanics problems, or basically any continuum mechanics problem.

Let us first illustrate a Rayleigh–Bénard convection setup. We have a layer of fluid confined between two plates, where the bottom plate is being uniformly heated up, $T_0 > T_h$. Due to temperature difference, the density variation leads to fluid motion generated by buoyancy. This motion is balanced by the viscous forces in fluid. The ratio between these two forces determines whether the fluid motion is stable, and we will look at this dimensionless ratio closely in later steps.

![Diagram showing Rayleigh–Bénard convection](image)

Note that although we illustrate the problem in 3D, we are mostly interested in the 2D case because it is easier to study. Therefore, we will only be looking at fluid motion in the $x$ and $y$ direction, which is effectively the same as looking at slices of the fluid motion in the $z$ direction.

We then list all the variables and parameters of interest. Usually we obtain them by writing down the governing equations first; here we are doing things in a reversed order for pedagogy reasons:

<table>
<thead>
<tr>
<th>variables</th>
<th>parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>fluid density $\rho$</td>
<td>volume coefficient of thermal expansion $\alpha$</td>
</tr>
<tr>
<td>fluid velocity $u$</td>
<td>gravitation $g$</td>
</tr>
<tr>
<td>temperature $T$</td>
<td>dynamic viscosity $\mu$, kinematic viscosity $v = \mu/\rho$</td>
</tr>
<tr>
<td>pressure $p$</td>
<td>thermal diffusivity $\kappa$</td>
</tr>
</tbody>
</table>
Step 2: Write down the governing equations

Since the density is no longer constant, we cannot directly use the incompressible Navier–Stokes equations as the governing equations. Before we go back to the mass conservation and momentum conservation equations, let us first consider an equation of state (EOS) for density, i.e. how \( \rho \) depends on temperature \( T \) and pressure \( p \). A standard EOS is based on the Boussinesq approximation:\(^1\)

\[
\rho = \rho_0 \left( 1 + \alpha ( T - T_0 ) \right)
\]

(1)

where \( \alpha < 0 \) is the volume coefficient of thermal expansion. Usually, fluid expands and becomes less dense as it is heated. \( \rho \) is assumed to have a fixed part\(^2\) \( \rho_0 \) and the other part that has a linear dependence on temperature \( T \). The Boussinesq approximation states the density variation is only important in the buoyancy term.

Next let us look at mass conservation. Since the variation of \( \rho \) is very small, the general form still reduces essentially to the continuity equation (or incompressibility constraint):

\[
\partial_t \rho + \nabla \cdot ( \rho \mathbf{u} ) = 0 \quad \implies \quad \nabla \cdot \mathbf{u} = 0
\]

(2)

Then we move on to momentum conservation, i.e. Navier–Stokes equation. Since the density variation is only important in the buoyancy term, we have

\[
\rho_0 \frac{D \mathbf{u}}{Dt} = -\nabla p + \mu \nabla^2 \mathbf{u} - \rho g \mathbf{\hat{y}}
\]

\[
\implies \rho_0 \left( \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} \right) = -\nabla p + \mu \nabla^2 \mathbf{u} - \rho_0 (1 + \alpha (T - T_0)) g \mathbf{\hat{y}}
\]

(3)

Lastly, we need to close the system because we have four variables but only three equations so far. There are two ways to think about the origin of the last equation: (1) an energy equation with respect to (w.r.t.) temperature; (2) the temperature variable in the Boussinesq approximation. Since temperature \( T \) is a scalar field, we can use the advection–diffusion equation to evolve it:

\[
\partial_t T + (\mathbf{u} \cdot \nabla) T = \kappa \nabla^2 T
\]

(4)

In summary, the system of governing equations to solve is

\[
\begin{align*}
\text{momentum:} & \quad \rho_0 \left( \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} \right) = -\nabla p + \mu \nabla^2 \mathbf{u} - \rho_0 (1 + \alpha (T - T_0)) g \mathbf{\hat{y}} \\
\text{mass:} & \quad \nabla \cdot \mathbf{u} = 0 \\
\text{energy:} & \quad \partial_t T + (\mathbf{u} \cdot \nabla) T = \kappa \nabla^2 T \\
\text{equation of state:} & \quad \rho = \rho_0 (1 + \alpha (T - T_0))
\end{align*}
\]

(5)

\(^1\)Boussinesq approximation: https://en.wikipedia.org/wiki/Boussinesq_approximation_(buoyancy)

\(^2\)\( \rho_0 \) is the fluid density at temperature \( T_0 \).
Step 3: Nondimensionalize the governing equations

We could just go ahead and solve the system of governing equations. However, our result would depend on the physical and computational units we put in. To make our result more general (and not to worry about units in the calculation), we nondimensionalize the equations by removing physical dimensions. From personal experience, some benefits of nondimensionalization are:

- reduce the number of parameters (e.g. with Buckingham $\pi$ theorem$^3$)
- working with more general and much simpler governing equations
- describe the problem only with dimensionless number
- extract dominant scaling or characteristic feature of a system

There are many approaches to nondimensionalizing a system. Here, we first find the existing scales in the problem. Since the distance between the two plates is $h$, a natural choice of length scale of this problem is $h$. For time scale $\tau$, the convention is to use thermal scaling $\tau = h^2/\kappa$. We have temperature scale determined by the two plates, $\Delta T = T_h - T_0$. We can also derive two other scales:

<table>
<thead>
<tr>
<th>length scale</th>
<th>time scale</th>
<th>temperature scale</th>
<th>velocity scale</th>
<th>pressure scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h$</td>
<td>$\tau = \frac{h^2}{\kappa}$</td>
<td>$\Delta T = T_h - T_0$</td>
<td>$U = \frac{h}{\tau} = \frac{\kappa}{h}$</td>
<td>$P = \rho_0 U^2 \frac{\nu}{\kappa}$</td>
</tr>
</tbody>
</table>

One way to check that the units indeed match is to explicitly write down the SI units of each variable and operator. Use $M$ for mass, $L$ for length and $T$ for time, we have

$$[u] = \left[\frac{L}{T}\right], \quad [\rho_0] = \left[\frac{M}{L^3}\right], \quad [P] = \left[\frac{M}{LT^2}\right], \quad [\nu] = \left[\frac{L^2}{T}\right], \quad [\nabla] = \left[\frac{1}{L}\right], \quad [\nabla] = \left[\frac{1}{T}\right]$$

We then decompose the variables and operators into a dimensionless parameters with dimensional scales, where all terms with an underline are dimensionless:

$$\underline{u} = \frac{\kappa}{h} \underline{u}, \quad \underline{p} = \underline{P} \underline{p} = \rho_0 \kappa^2 \underline{v} \frac{\nu h^2}{\kappa} p, \quad \underline{T} = \Delta T \underline{T} + T_0, \quad \underline{\nu} = \frac{1}{h^2} \underline{\nabla}, \quad \underline{\nabla} = \underline{u} \cdot \nabla = \frac{\kappa}{h^2} \underline{\nabla}$$

Substituting into the momentum equation, we have

$$\rho_0 \left( \frac{\partial \underline{u}}{\partial t} + (\underline{u} \cdot \nabla) \underline{u} \right) = -\nabla \underline{p} + \mu \nabla^2 \underline{u} - \rho_0 (1 + \alpha (T - T_0)) \underline{g} \hat{y}$$

$$\Rightarrow \quad \left( \frac{\partial \underline{u}}{\partial t} + (\underline{u} \cdot \nabla) \underline{u} \right) = -\frac{1}{\rho_0} \nabla \underline{p} + \nu \nabla^2 \underline{u} - (1 + \alpha (T - T_0)) \underline{g} \hat{y}$$

$$\Rightarrow \quad \frac{\kappa}{h} \frac{\kappa}{h^2} \left( \frac{\partial \underline{u}}{\partial t} + (\underline{u} \cdot \nabla) \underline{u} \right) = -\frac{1}{\rho_0} \frac{\kappa}{h^2} \frac{\kappa}{h^2} \nabla \underline{p} + \nu \frac{1}{h^2} \frac{\kappa}{h^2} \nabla \underline{u} - \underline{g} \hat{y} - \alpha \Delta T \underline{T} g \hat{y}$$

$$\Rightarrow \quad \frac{\kappa^2}{h^4} \left( \frac{\partial \underline{u}}{\partial t} + (\underline{u} \cdot \nabla) \underline{u} \right) = -\frac{v \kappa}{h^3} \nabla \underline{p} + \frac{v \kappa}{h^3} \nabla^2 \underline{u} - \underline{g} \hat{y} - \alpha \Delta T \underline{T} g \hat{y}$$

$^3$Buckingham $\pi$ theorem: https://en.wikipedia.org/wiki/Buckingham_%CF%80_theorem

$^4$Characteristic dynamic pressure is $\rho_0 U^2$. $\nu/\kappa$ shows the ratio between momentum diffusion and thermal diffusion. We will see at the end of Step 3 that this dimensionless ratio is just the Prandtl number. This number will also pop up if we directly do the algebra without defining a pressure scale.
Let \( \tilde{p} \) be the modified pressure to account for the reference temperature term, i.e. we can absorb \( g \hat{y} \) into the dimensionless pressure \( \hat{p} \): 
\[
\hat{p} = p + \frac{h^2}{v\kappa} g \hat{y} \implies \nabla \hat{p} = \nabla p + \frac{h^3}{v\kappa} g \hat{y}
\]  
(7)

Further simplify the momentum equation, we have 
\[
\implies \frac{\kappa^2}{h^3} \left( \frac{\partial u}{\partial t} + (u \cdot \nabla) u \right) = -\frac{v\kappa}{h^3} \nabla \hat{p} + \frac{v\kappa}{h^3} \nabla^2 u - \alpha \Delta T g T \hat{y}
\]
\[
\implies \frac{\kappa}{v} \left( \frac{\partial u}{\partial t} + (u \cdot \nabla) u \right) = -\nabla \hat{p} + \nabla^2 u - \frac{\alpha \Delta T g h^3}{v\kappa} T \hat{y}
\]  
(8)

where we drop the underlines and tilde for notational simplicity. We have two dimensionless number:

- **Prandtl number**\(^5\): \( Pr = \frac{v}{\kappa} \) = momentum diffusivity / thermal diffusivity
- **Rayleigh number**\(^6\): \( Ra = \frac{\alpha \Delta T g h^3}{v\kappa} \) = convection / conduction = \( \frac{\rho g}{v\kappa/h^3} = \frac{\rho U}{h^2} \) = buoyant force / viscous force

Many dimensionless numbers in fluid dynamics are named after famous fluid dynamicists. In Rayleigh–Bénard convection, the most important dimensionless number is the Rayleigh number, which quantifies the ratio between buoyancy (i.e. temperature-driven inertial forces) and viscous forces. A high \( Ra \) implies turbulent flow fields.

We can perform similar nondimensionalization on the other two governing equations [HW] (not really needing EOS). Thus, our system of nondimensionalized governing equations becomes

| \text{momentum:} | \frac{1}{Pr} \left( \frac{\partial u}{\partial t} + (u \cdot \nabla) u \right) = -\nabla p + \nabla^2 u - Ra \ T \hat{y} | \\ 
| \text{mass:} | \nabla \cdot u = 0 | \\ 
| \text{energy:} | \partial_t T + (u \cdot \nabla) T = \nabla^2 T |  
(9)

with nondimensionalized boundary conditions:

- at \( y = 0 \): \( u = 0 \) and \( T = 0 \)
- at \( y = 1 \): \( u = 0 \) and \( T = 1 \)

Nondimensionalization has greatly simplified our governing equations and showed us the dominant terms in the system. However, there is still one more operation on the governing equations before we can start solving it. We will linearize them to find the critical temperature difference where an organized cellular motion sets in. We will also look at the B.C.s in more detail in the next step.

\(^5\)Ludwig Prandtl: https://en.wikipedia.org/wiki/Ludwig_Prandtl
Prandtl number: https://en.wikipedia.org/wiki/Prandtl_number

\(^6\)Lord Rayleigh: https://en.wikipedia.org/wiki/John_William_Strutt,_3rd_Baron_Rayleigh
Rayleigh number: https://en.wikipedia.org/wiki/Rayleigh_number
Step 4: Find the steady state solutions

For linear stability theory, we need to first find the steady state solution. One obvious steady state solution for the fluid velocity \( \mathbf{u} = (u, v) \) is \( u^{SS} = v^{SS} = 0 \), i.e. the fluid is not moving. Substituting in:

\[
\begin{align*}
\text{momentum } \dot{x} : & \quad \frac{1}{Pr} \left( \partial_t u^{SS} + (u^{SS} \cdot \nabla) u^{SS} \right) = -\partial_x p^{SS} + \nabla^2 u^{SS} \\
\text{momentum } \dot{y} : & \quad \frac{1}{Pr} \left( \partial_t v^{SS} + (u^{SS} \cdot \nabla) v^{SS} \right) = -\partial_y p^{SS} + \nabla^2 v^{SS} - Ra T^{SS} g
\end{align*}
\]

\[
\rightarrow \begin{cases}
\text{momentum } \dot{x} : & \partial_x p^{SS} = 0 \\
\text{momentum } \dot{y} : & \partial_y p^{SS} = -Ra T^{SS} g
\end{cases} \quad \rightarrow \quad \begin{cases}
\partial_y p^{SS}(y) = -Ra T^{SS} g
\end{cases}
\]  (11)

We can also substitute the steady state solution into the temperature equation. At steady state, temperature is not changing w.r.t. time, i.e. \( \partial_t T = 0 \). What’s more, temperature is only varying in the \( y \) direction, i.e. \( \partial_x T = 0 \). Applying boundary conditions \( T(0) = 0 \) and \( T(1) = 1 \), we have:

\[
\begin{align*}
\text{energy: } & \quad \partial_t T^{SS} + (u^{SS} \cdot \nabla) T^{SS} = \kappa \left( \partial_x^2 T + \partial_y^2 T \right) \\
\rightarrow \quad \text{energy: } & \quad \partial_y^2 T^{SS} = 0 \quad \rightarrow \quad T^{SS}(y) = y
\end{align*}
\]  (12)

Step 5: Apply linear stability theory and linearize about the steady state

Now we disturb the steady state system slightly by introducing a small perturbation \( \epsilon \ll 1 \):

\[
\begin{align*}
\mathbf{u} &= \underbrace{u^{SS}}_{\text{steady state}} + \underbrace{\epsilon \hat{u}}_{\text{perturbation}} , \quad v = v^{SS} + \epsilon \hat{v}, \quad p = p^{SS} + \epsilon \hat{p}, \quad T = T^{SS} + \epsilon \hat{T}
\end{align*}
\]  (13)

Let us first apply linear stability theory on the momentum \( \dot{x} \) equation. Substitute these linearized solutions and collect terms based on \( O(\epsilon^n) \) orders, we have:

\[
\begin{align*}
\text{momentum } \dot{x} : & \quad \frac{1}{Pr} \left( \partial_t (u^{SS} + \epsilon \hat{u}) + (u^{SS} + \epsilon \hat{u}) \partial_x (u^{SS} + \epsilon \hat{u}) + (v^{SS} + \epsilon \hat{v}) \partial_y (u^{SS} + \epsilon \hat{u}) \right) \\
& = -\partial_x (p^{SS} + \epsilon \hat{p}) + \nabla^2 (u^{SS} + \epsilon \hat{u}) \\
\mathcal{O}(\epsilon^0) : & \quad \frac{1}{Pr} \left( \partial_t u^{SS} + (u^{SS} \cdot \nabla) u^{SS} \right) = -\partial_x p^{SS} + \nabla^2 u^{SS} \quad \text{steady state solutions!} \\
\mathcal{O}(\epsilon^1) : & \quad \frac{1}{Pr} \partial_t \hat{u} = -\partial_x \hat{p} + \nabla^2 \hat{u}
\end{align*}
\]

We are only interested in perturbations to \( O(\epsilon^1) \). Perform the same algebra to other equations [HW]:

\[
\begin{align*}
\text{momentum } \dot{x} : & \quad \frac{1}{Pr} \partial_t \hat{u} = -\partial_x \hat{p} + \nabla^2 \hat{u} \quad [1] \\
\text{momentum } \dot{y} : & \quad \frac{1}{Pr} \partial_t \hat{v} = -\partial_y \hat{p} + \nabla^2 \hat{v} - Ra \hat{T} g \quad [2] \\
\text{energy : } & \quad \partial_t \hat{T} + \hat{v} \partial_y T^{SS} = \nabla^2 \hat{T} \\
\text{mass : } & \quad \partial_x \hat{u} + \partial_y \hat{v} = 0 \quad [4]
\end{align*}
\]

We have finally found the linearized nondimensional governing equations for this problem!
Step 6: Simplify the governing equations to ODE

We have four coupled PDEs, and our goal is to convert them into a single ODE\(^7\). First, we remove the pressure term by \(\{\partial_y[1] - \partial_x[2]\}\) to find:

\[
\frac{1}{Pr} \partial_t \left( \partial_y \hat{u} - \partial_x \hat{v} \right) = \nabla^2 \left( \partial_y \hat{u} - \partial_x \hat{v} \right) + Ra \partial_x \hat{T} \quad [5]
\]

Then we eliminate \(\hat{u}\) by \(\{\partial_x[5] \text{ and } [4]\}\) to get:

\[
- \frac{1}{Pr} \partial_t \nabla^2 \hat{v} = -\nabla^4 \hat{v} + Ra \partial_x^2 \hat{T} \quad [6]
\]

Equations [3] and [6] have 2 unknowns together, \(\hat{v}\) and \(\hat{T}\). Since they are heuristically periodic in the \(x\) direction, and the instability is growing exponential in time, suitable guesses for \(\hat{v}\) and \(\hat{T}\) are

\[
\hat{v} = v^* \exp(\omega t) \sin(kx) \\
\hat{T} = T^* \exp(\omega t) \sin(kx)
\]

where \(\omega\) is the growth rate of instability. Therefore, the onset of instability of Rayleigh–Bénard convection is when \(\omega\) is imaginary, i.e. \(\Re(\omega) = 0^+\).

Plugging the guesses into equations [3] and [6] and doing algebra, we have

\[
[3'] : \quad \omega T^* + v^* = (D^2 - k^2) T^* \\
[6'] : \quad - \frac{1}{Pr} \omega(D^2 - k^2) v^* = -(D^2 - k^2)^2 v^* - Ra k^2 T^*
\]

where for notational convenience, \(D = \partial/\partial y\). The onset of instability is when \(\omega \sim 0\), thus

\[
[3''] : \quad v^* = (D^2 - k^2) T^* \\
[6''] : \quad (D^2 - k^2)^2 v^* = -Ra k^2 T^*
\]

Finally, combine equations [3''] and [6''] into a single ODE via substitution, and we have:

\[
(D^2 - k^2)^3 v^* = -Ra k^2 T^* \quad [7]
\]

This is a 6th-order ODE, thus we need 6 B.C.s (transformed into linearized nondimensional space):

at \(y = 0\): \( v = 0 \) (no flux) \( \partial_y v = 0 \) (no-slip) \( \hat{T} = 0 \)

at \(y = 1\): \( v = 0 \) (no flux) \( \partial_y v = 0 \) (no-slip) \( \hat{T} = 0 \) (small perturbation)

\[
\rightarrow \quad \text{at } y = 0, 1: \quad v^* = Dv^* = (D^2 - k^2)^2 v^* = 0 \quad [23]
\]

This is an eigenvalue problem with eigenvalue\(^8\) \(Ra\)! There will not be an eigenfunction \(v^*\) satisfying our ODE for any combination of \(k\) and \(Ra\): the solutions lie on different contours representing different modes of instability. The smallest value of \(Ra\) on each contour is the critical point for convection to form.

\(^7\) Not necessary if doing everything numerically. But our goal here is to find the onset of instability analytical equation.

\(^8\) You can think of the Rayleigh number as a dimensionless temperature difference.
Step 7: Solve the eigenvalue problem

Let \( D_k = -\frac{(D^2 - k^2)^3}{k^2} = -\frac{(\partial_y^2 - k^2)^3}{k^2} \) and \( f = v^* \), then the ODE becomes

\[
D_k f = Ra f
\]  
(24)

There are multiple ways of solving this matrix equation:

- analytics/algebra (with Mathematica)
- Newton's method, shooting method, and other root-finding methods
- represent \( D_k \) as a matrix and find the eigenvalues

The critical Rayleigh number is \( Ra_c = 1708 \). When \( Ra > Ra_c \), instability between two rigid plates occurs. The Rayleigh number shows how each parameter determines the stability of the base/steady state. For example, fluid viscosity \( \nu \) plays a stabilizing role; the larger the value of \( \nu \), the larger the temperature difference \( \Delta T \) needed before convection happens.

(Optional: Analytically solve the eigenvalue problem by hand)

We cannot analytically solve the governing equation [7] by hand. That is why we should use numerical methods to computationally solve for phase diagram of each \( k \) and \( Ra \). However, we can solve the alternative problem when the two plates are assumed stress-free. In this artificial case, we have

\[
\partial_y u = 0, \quad \text{continuity:} \quad \partial_y (\partial_x u + \partial_y v) = 0 \implies D^2 v = 0
\]  
(25)

thus the B.C.s simplify to

at \( y = 0, 1 \): \( v^* = Dv^* = D^4 v^* = 0 \)  
(26)

which means the even derivatives are zero at \( y = 0, 1 \). Assume the solution take the form

\[
v^* = A_n \sin(n \pi y)
\]  
(27)

which satisfies the ODE and the modified B.C.s. Plugging into equation [7], we have

\[
(-n^2 \pi^2 - k^2)^3 A_n \sin(n \pi y) = -Ra k^2 A_n \sin(n \pi y) \implies Ra = \frac{(n^2 \pi^2 + k^2)^3}{k^2} > 0
\]  
(28)

The minimum \( Ra_c \) is found when \( n = 1 \), where

\[
\partial_k \left( \frac{(n^2 + k^2)^3}{k^2} \right) = 0 \implies k_c = \frac{\pi}{\sqrt{2}} \implies Ra_c = \frac{27 \pi^4}{4} \approx 657
\]  
(29)

This stress-free boundary was in fact the case solved by Lord Rayleigh in 1916, according to whom captured the essentials of the problem. Here we see the power of nondimensionalization, linear stability theory, and proper simplifications in studying fluid instability problems.
References


[3] AM205 (Fall 2020): Group activity “Applying class concepts to describe convection cells”