Theoretical Astrophysics I: Stellar Structure and Evolution
Lecture 4: Convective energy transport

Petri Käpylä & Ivan Milić
pkapyla,milic@leibniz-kis.de

Institut für Sonnenphysik (KIS), Freiburg

May 22, 2025

Energy transport mechanisms in stars

- Radiation: photons transport the energy. Mean free path is very short and this can be modeled as a diffusion process down the temperature gradient, i.e.,

$$\mathbf{F}_{\mathrm{rad}} = -K_{\mathrm{rad}} \boldsymbol{\nabla} T, \tag{1}$$

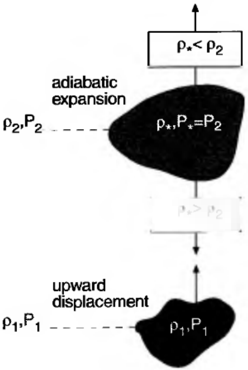
  where K_{rad} is the radiative conductivity.
- Conduction: heat is transported because of collisions of particles. Analogous to 1, this is written as

$$\mathbf{F}_{\mathrm{cd}} = -K_{\mathrm{cd}} \boldsymbol{\nabla} T. \tag{2}$$

  Typically $K_{\mathrm{rad}} \gg K_{\mathrm{cd}}$ (Except where?).
- Convection: gas is opaque to radiation, becomes *unstable*, and fluid motions transport the energy. This leads to very complicated dynamics and cannot in general be represented in equally simple terms as radiation or conduction.

Intuitive picture of convective instability

- Consider a fluid element with density ρ_1 and pressure P_1 displaced upward to a level where $\rho = \rho_2$ and $P = P_2$.
- If the density inside the element ($\rho_\star$) is larger (smaller) than the ambient density, it is pulled back (continues to accelerate).
- Assumption: no heat exchange between fluid element and the surroundings.



Source: Prialnik

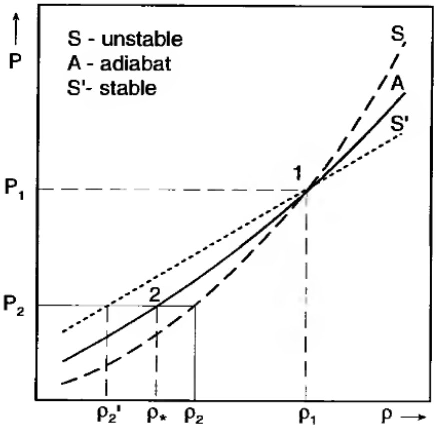
Schwarzschild criterion

- Consider the atmospheres A, S, and S': the stably stratified case corresponds to S', where

$$\frac{\partial P}{\partial \rho} < \left(\frac{\partial P}{\partial \rho}\right)_{\rm a} \qquad (3)$$

- This form is not particularly useful so we will recast it in terms of a temperature gradient.
- Recall the 1st law of thermodynamics:

$$dQ = du + PdV. \qquad (4)$$



Source: Prialnik

Schwarzschild criterion

- The ideal gas equation can be written as:

$$P = \mathcal{R}_{\rm spec}\rho T = (c_P - c_V)\rho T, \tag{5}$$

where $c_P = \frac{dQ}{dT}\big|_P$ and $c_V = \frac{dQ}{dT}\big|_V$ are specific heat capacities at constant pressure and at constant volume (unit J K^{-1} kg^{-1}). Their ratio is $\gamma = c_P/c_V = 5/3$ (aka adiabatic index).

- Furthermore, with the specific internal energy $u = c_V T$ this becomes

$$P = \mathcal{R}_{\rm spec}\rho T = (\gamma - 1)\rho u, \quad \text{and} \quad u = \frac{1}{\gamma - 1}\frac{P}{\rho}. \tag{6}$$

Schwarzschild criterion

- For an adiabatic process $dQ = 0$. Furthermore, the specific volume is $V = \rho^{-1}$, and $dV = -d\rho/\rho^2$. Thus,

$$dQ = du + PdV. \tag{7}$$

$$\frac{1}{\gamma - 1}\left(\frac{dP}{\rho} - P\frac{d\rho}{\rho^2}\right) - P\frac{d\rho}{\rho^2} = 0. \tag{8}$$

$$\frac{1}{\gamma - 1}\left(\frac{dP}{P} - \frac{d\rho}{\rho}\right) - \frac{d\rho}{\rho} = 0. \tag{9}$$

$$\frac{1}{\gamma - 1}\left(\frac{\rho}{P}\frac{dP}{d\rho} - 1\right) - 1 = 0. \tag{10}$$

$$\frac{\rho}{P}\frac{dP}{d\rho} = \left(\frac{\rho}{P}\frac{dP}{d\rho}\right)_{\mathrm{a}} = 1 + (\gamma - 1) = \gamma. \tag{11}$$

Schwarzschild criterion

- Going back to Eq. (3) we can write:

$$\frac{\rho}{P}\frac{dP}{d\rho} < \left(\frac{\rho}{P}\frac{dP}{d\rho}\right)_{\mathrm{a}} = \gamma. \tag{12}$$

- For an ideal gas

$$\frac{dP}{P} = \frac{d\rho}{\rho} + \frac{dT}{T}. \tag{13}$$

- Multiply by P/dP, make use of Eq. (12) and define ∇_{a}:

$$1 = \left(\frac{P}{\rho}\frac{d\rho}{dP} + \frac{P}{T}\frac{dT}{dP}\right)_{\mathrm{a}} \equiv \frac{1}{\gamma} + \nabla_{\mathrm{a}}, \tag{14}$$

or:

$$\nabla_{\mathrm{a}} = 1 - \frac{1}{\gamma} = \frac{\gamma - 1}{\gamma}. \tag{15}$$

Schwarzschild criterion

- Extending the definition of ∇ to the general case

$$\nabla \equiv \frac{P}{T}\frac{dT}{dP}, \tag{16}$$

  we can rewrite the stability condition (12) as

$$\nabla < \nabla_{\mathrm{a}}. \tag{17}$$

  This criterion was derived by Karl Schwarzschild in 1906.
- Sometimes the quantity $\Delta\nabla = \nabla - \nabla_{\mathrm{a}}$ (superadiabaticity) is used to denoted whether a layer is convectively stable or not.
- In stellar convection zones (apart from the near-surface layers) $\Delta\nabla$ is very small, e.g., $\mathcal{O}(10^{-8} \ldots 10^{-3})$ in the Sun.

Schwarzschild criterion – reflection

- Violation of Schwarzschild criterion is a necessary but not sufficient condition for convection to occur (Why?)

Schwarzschild criterion – reflection

- Violation of Schwarzschild criterion is a necessary but not sufficient condition for convection to occur (Why?)
- Internal friction in the gas was been neglected $\rightarrow$ in reality $\Delta\nabla$ has to exceed a finite value $\Delta\nabla_{\mathrm{min}}$ (which depends on T, ρ, rotation, magnetic fields, etc.) for convection to ensue.
- Typically this is represented by a Rayleigh number which is a ratio of convective transport to diffusive transport:

$$\mathrm{Ra} = \frac{gd^4}{\nu\chi H_p}\Delta\nabla, \tag{18}$$

  where $g = GM/r^2$, ν is the kinematic viscosity, $\chi = K_{\mathrm{rad}}/\rho c_P$ is the radiative diffusivity, and H_P is the pressure scale height.
- Critical value for free convection is $\mathrm{Ra_c} \approx 1707$. In the Sun, $\mathrm{Ra} \approx 10^{20}$.

When and why does convection happen?

The Schwarzschild criterion is very general and we would like to have something more spectific to stars.

- Consider a radiative star in hydrostatic equilibrium where the whole luminosity is transported by radiative diffusion:

$$F = -4\pi r^2 K \frac{\partial T}{\partial r}, \tag{19}$$

  where K is the radiative conductivity:

$$K = \frac{16\sigma T^3}{3\kappa\rho} = \frac{4cT^3}{3a\kappa\rho}. \tag{20}$$

- With this the temperature gradient is

$$\frac{\partial T}{\partial r} = -\frac{3}{4ac}\frac{\kappa\rho}{T^3}\frac{F}{4\pi r^2}. \tag{21}$$

When and why does convection happen?

- Recast this in terms of radiative pressure:

$$P_{\mathrm{rad}} = \tfrac{1}{3} a T^4, \quad \text{or} \quad dP_{\mathrm{rad}} = \tfrac{4}{3} T^3 dT. \tag{22}$$

- Now we can write the equation in terms of P_{rad}:

$$\frac{dP_{\mathrm{rad}}}{dr} = -\kappa\rho \frac{F}{4\pi r^2}. \tag{23}$$

- Finally, we use the hydrostatic equation:

$$\frac{dP}{dr} = -\rho \frac{Gm}{r^2}, \tag{24}$$

to obtain:

$$\frac{dP_{\mathrm{rad}}}{dP} = \frac{\kappa F}{4\pi Gm}. \tag{25}$$

When and why does convection happen?

- The sought relation is:

$$\frac{dP_{\mathrm{rad}}}{dP} = \frac{\kappa F}{4\pi G m}. \tag{26}$$

- What can you say based on this equation?

When and why does convection happen?

- The sought relation is:

$$\frac{dP_{\mathrm{rad}}}{dP} = \frac{\kappa F}{4\pi G m}. \tag{27}$$

- What can you say based on this equation?
- Because $P = P_{\mathrm{gas}} + P_{\mathrm{rad}}$, $dP > dP_{\mathrm{rad}}$, or $\frac{dP_{\mathrm{rad}}}{dP} < 1$.
- This means that for a *radiative* star

$$\kappa F < 4\pi G m. \tag{28}$$

- This condition can be violated if κF is sufficiently big and at least a part of the energy must be transported by means other than radiation.
- Increase κ due to partial ionization: outer layers of low mass stars such as the Sun.
- Increase F due to strong temperature dependence of energy production: CNO-cycle and triple-α in massive stars.

Eddington luminosity

- Eq. (28) can be rewritten in terms of L and M

$$L < \frac{4\pi GM}{\kappa}. \tag{29}$$

- This is a fundamental limit for radiative luminosity (Eddington luminosity). If L exceeds this, a radiation-driven *stellar wind* will commence.

- For a radiative star the total flow of energy (unit: W) is transported by radiation and given by

$$F = F_{\mathrm{rad}} = 4\pi r^2 \frac{4ac}{3} \frac{T^4}{\kappa \rho H_p} \nabla_{\mathrm{rad}}, \tag{30}$$

where ∇_{rad} is the radiative temperature gradient; see, Eq. (21).
- When convection is present $F = F_{\mathrm{rad}} + F_{\mathrm{conv}} \neq F_{\mathrm{rad}}$, and therefore $\nabla \neq \nabla_{\mathrm{rad}}$.
- Therefore we need to compute F_{conv} in order to calculate ∇.

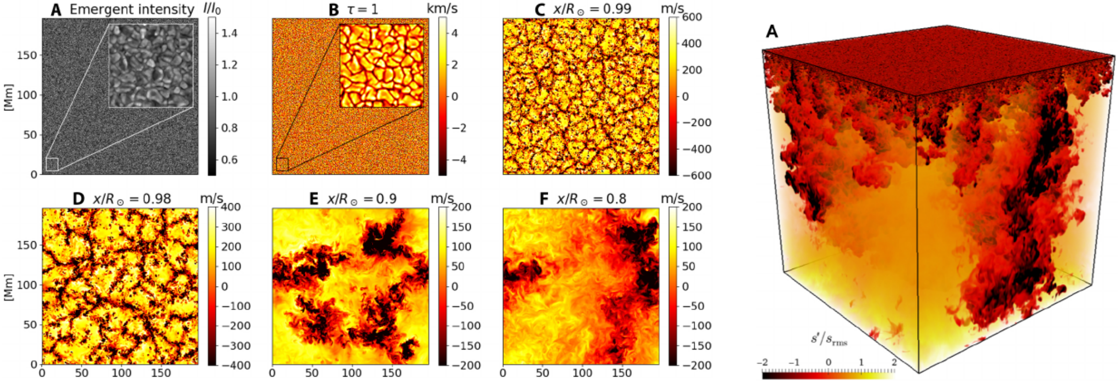
Mixing length theory

Convection is highly turbulent and chaotic, and occurs in 3D. This in itself is already a nearly unsolvable problem (possibly next lecture). Stellar evolution models require a simple presciption of convection in 1D (although there is no rigorous way to do this!).

- Assume discrete gas elements that move a vertical distance ℓ before dissolving.
- This distance is called the *mixing length*, and is given by $\ell = \alpha_{\mathrm{MLT}} H_P$, where $\alpha_{\mathrm{MLT}} = \mathcal{O}(1)$. The functional form is a choice and cannot be derived rigorously from first principles!
- The mixing length theory is analogous to the concept of mean-free path in thermodynamics. It originates from Ludwig Prandtl (1925) who used it to describe turbulent motions in the laboratory.
- It was adapted to astrophysics by Ludwig Biermann in the 1930s and the most widely used version is due to Erika Vitense (1953); see also Böhm-Vitense (1958).
- What can you conclude from the basic assumption of the mixing length theory?

Mixing length theory

Pressure scale height increases with depth → convection cells get bigger.



Source: Hotta et al. (2019), *Science Advances*, 5, 2307

- Consider the convective energy (more precisely enthalpy) flux (unit: W/m^2) is:

$$F_{\rm conv} = c_P \rho u \delta T, \tag{31}$$

where u is the convective velocity, and $\delta T = T - \overline{T}$.

- Assume that the "average" convective blob has traveled a distance $d = \ell/2$.
- Then the temperature fluctuation is:

$$\frac{\delta T}{T} = \frac{d}{T}\left(\frac{\partial T}{\partial r} - \left.\frac{\partial T}{\partial r}\right|_{\rm ad}\right) = (\nabla - \nabla_{\rm ad})\frac{\ell}{2H_P}, \tag{32}$$

- Assume that half of the work done by the buoyancy force when travelling a radial distance $\ell/2$ goes to kinetic energy:

$$-\frac{1}{2}g\frac{\delta\rho}{\rho}\frac{\ell}{2} = \frac{1}{2}g\frac{\delta T}{T}\frac{\ell}{2} = (\nabla - \nabla_{\rm ad})\frac{g\ell^2}{8H_P}. \tag{33}$$

Mixing length theory

- Assume that half of this goes to the kinetic energy of the fluid element:

$$u^2 = (\nabla - \nabla_{\rm ad}) \frac{g\ell^2}{8H_P}. \tag{34}$$

- Combining Eqs. (32) and (34) gives the convective flux:

$$F_{\rm conv} = c_P \rho T g^{1/2} \frac{\ell^2}{4\sqrt{2} H_P^{3/2}} (\nabla - \nabla_{\rm ad})^{3/2}. \tag{35}$$

- What do these equations imply? Hint: recall the Schwarzschild criterion.
- We made a shortcut in the analysis here. Any idea what this was?

Mixing length theory

- What do these equations imply? Hint: recall the Schwarzschild criterion.
- The factor $\nabla - \nabla_{\rm ad}$ means that when the stratification changes from unstable to stable the convective flow abruptly stops. Therefore mixing length theory is *local* and does not allow *overshooting*.
- We made a shortcut in the analysis here. Any idea what this was?
- There are still radiative losses from the convective fluid elements and therefore the temperature gradient inside the element ($\nabla_{\rm e}$) is not equal to $\nabla_{\rm ad}$.
- This leads to an additional equation (known as the cubic equation) that needs to be solved for ∇.
- However, this is only important very near the surfaces of stars and many stellar evolution codes simply set $\nabla = \nabla_{\rm ad}$ when the Schwarzschild criterion is violated.

Mixing length theory

- Mixing length model of the solar convection zone.
- Huge differences in pressure, density, superadiabaticity, and velocity.
- Also the convective turnover time $\tau_c = \ell/u$ varies several orders of magnitude.

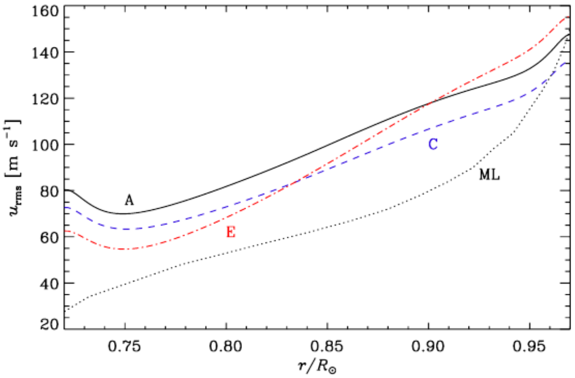
Table 6.1. Convection zone of a standard solar model ($Z+n$ means $Z \times 10^n$)

$r/r_\odot$	P [Pa]	T [K]	$\rho\left[\frac{\mathrm{kg}}{\mathrm{m}^3}\right]$	η_H	η_He	η_{He^+}	$\Delta\nabla$	ΔT [K]	$v\left[\frac{\mathrm{m}}{\mathrm{s}}\right]$	F_C/F
1.000	9.55+03	5.78+3	2.51−4	.00	.00	.00	−1.1−1	0.0+0	0	0.00
1.000	1.18+04	6.23+3	2.87−4	.00	.00	.00	9.7−2	7.7+0	152	0.00
1.000	1.34+04	6.74+3	3.01−4	.00	.00	.00	4.1−1	4.6+2	1181	0.05
1.000	1.45+04	7.21+3	3.05−4	.00	.00	.00	5.4−1	1.5+3	2172	0.32
1.000	2.19+04	9.26+3	3.52−4	.02	.00	.00	2.0−1	1.6+3	2427	0.96
1.000	3.33+04	1.05+4	4.57−4	.05	.00	.00	9.7−2	9.2+2	2042	0.99
.999	5.04+04	1.14+4	6.16−4	.09	.00	.00	5.7−2	5.9+2	1783	1.00
.999	7.64+04	1.22+4	8.44−4	.13	.00	.00	3.7−2	4.1+2	1581	1.00
.999	1.16+05	1.30+4	1.17−3	.17	.00	.00	2.5−2	2.9+2	1412	1.00
.999	1.76+05	1.37+4	1.63−3	.21	.00	.00	1.7−2	2.2+2	1266	1.00
.999	2.66+05	1.45+4	2.28−3	.25	.00	.00	1.3−2	1.7+2	1139	1.00
.998	4.04+05	1.53+4	3.18−3	.28	.00	.00	9.2−3	1.3+2	1026	1.00
.998	6.12+05	1.62+4	4.46−3	.32	.00	.00	6.9−3	1.0+2	925	1.00
.998	9.28+05	1.71+4	6.24−3	.36	.00	.00	5.2−3	8.0+1	837	1.00
.997	1.41+06	1.80+4	8.72−3	.40	.00	.00	3.9−3	6.4+1	758	1.00
.997	2.13+06	1.91+4	1.22−2	.44	.00	.00	3.0−3	5.3+1	688	1.00
.997	3.23+06	2.03+4	1.70−2	.48	.01	.00	2.4−3	4.3+1	626	1.00
.996	4.90+06	2.16+4	2.37−2	.52	.01	.00	1.9−3	3.6+1	571	1.00
.996	7.42+06	2.31+4	3.28−2	.56	.02	.00	1.5−3	3.1+1	524	1.00
.995	1.13+07	2.48+4	4.51−2	.61	.04	.00	1.2−3	2.7+1	481	1.00
.995	1.71+07	2.68+4	6.21−2	.65	.07	.00	9.3−4	2.3+1	441	1.00
.994	3.18+07	3.03+4	9.89−2	.72	.15	.00	6.8−4	1.9+1	393	1.00
.993	5.94+07	3.47+4	1.56−1	.78	.30	.00	4.9−4	1.5+1	350	1.00
.992	9.01+07	3.84+4	2.09−1	.81	.43	.00	3.9−4	1.4+1	326	1.00
.991	1.37+08	4.28+4	2.79−1	.84	.57	.00	3.2−4	1.2+1	305	1.00
.990	2.07+08	4.81+4	3.68−1	.87	.70	.00	2.6−4	1.1+1	286	1.00
.988	3.14+08	5.47+4	4.82−1	.89	.80	.00	2.1−4	1.0+1	269	1.00
.987	4.76+08	6.28+4	6.27−1	.91	.86	.01	1.6−4	9.3+0	251	1.00
.985	7.21+08	7.25+4	8.13−1	.93	.88	.03	1.2−4	8.1+0	232	1.00
.983	1.09+09	8.37+4	1.05+0	.94	.85	.09	9.0−5	6.8+0	212	1.00
.981	1.66+09	9.65+4	1.37+0	.94	.77	.19	6.5−5	5.6+0	194	1.00
.976	3.09+09	1.19+5	2.04+0	.95	.57	.40	4.1−5	4.4+0	171	1.00
.971	5.77+09	1.49+5	3.00+0	.96	.38	.61	2.6−5	3.5+0	153	1.00
.967	8.75+09	1.73+5	3.88+0	.96	.28	.72	2.0−5	3.1+0	142	1.00
.962	1.33+10	2.02+5	4.99+0	.96	.21	.79	1.5−5	2.7+0	132	1.00
.956	2.01+10	2.37+5	6.42+0	.96	.16	.84	1.1−5	2.3+0	123	1.00
.949	3.05+10	2.78+5	8.24+0	.97	.13	.87	8.1−6	2.0+0	114	1.00
.942	4.62+10	3.27+5	1.06+1	.97	.11	.89	5.9−6	1.8+0	105	1.00
.932	7.00+10	3.85+5	1.36+1	.97	.09	.91	4.3−6	1.5+0	98	1.00
.922	1.06+11	4.53+5	1.74+1	.97	.08	.92	3.2−6	1.3+0	90	.99
.910	1.61+11	5.34+5	2.24+1	.97	.07	.93	2.3−6	1.1+0	84	.99
.896	2.44+11	6.29+5	2.87+1	.97	.07	.93	1.7−6	9.7−1	78	.98
.880	3.70+11	7.42+5	3.68+1	.97	.06	.94	1.2−6	8.4−1	72	.97
.862	5.61+11	8.75+5	4.73+1	.98	.06	.94	9.1−7	7.2−1	67	.95
.841	8.50+11	1.03+6	6.07+1	.98	.05	.95	6.6−7	6.2−1	62	.91
.818	1.29+12	1.22+6	7.79+1	.98	.05	.95	4.7−7	5.2−1	57	.86
.778	2.41+12	1.56+6	1.14+2	.98	.04	.96	2.7−7	3.8−1	48	.71
.732	4.56+12	2.01+6	1.66+2	.98	.04	.96	1.1−7	1.9−1	34	.34
.717	5.44+12	2.15+6	1.85+2	.98	.04	.96	5.6−8	1.1−1	26	.15
.710	5.96+12	2.23+6	1.95+2	.98	.04	.96	1.9−8	3.8−2	15	.03

Stix (2002), The Sun: An Introduction

Mixing length theory vs. simulations

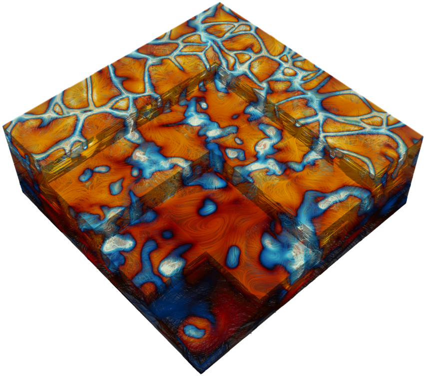
- Convective velocities in 3D simulations are of the same order of magnitude as from MLT.
- Exact match cannot be expected beacuse MLT is a very coarse approximation + simulations have their own issues (next lecture).

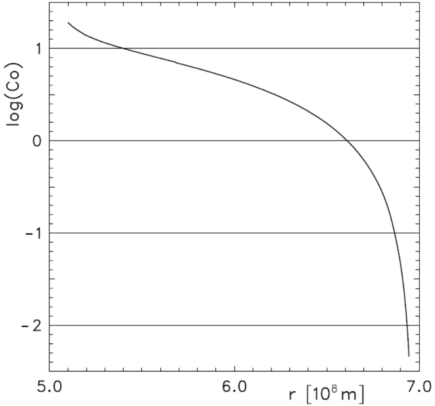


Average convective velocity in simulations (A, C, E) and from MLT

Mixing length theory – issues?

- Non-rotating convection looks similar to what mixing length theory predicts.

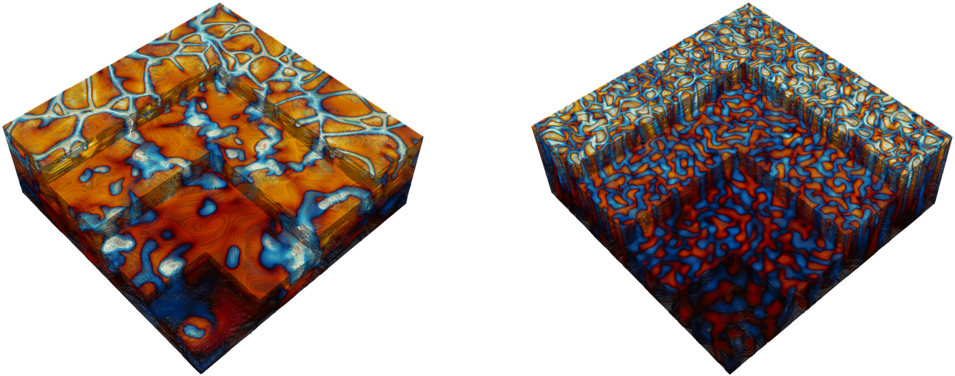




Left: Non-rotating convection. Right: logarithm of the Coriolis number $\mathrm{Co} = 2\Omega_\odot/\tau_c$ using mixing length model data.

Mixing length theory – issues?

- Non-rotating convection looks similar to what mixing length theory predicts. But not if rotation is sufficiently strong.



Left: Non-rotating convection. Right: rapidly rotating convection.

Perfect is the worst enemy of good enough?

- Mixing length theory (+ some additional tweaks to account for overshooting) is generally good enough to capture the big picture.
- However, the interplay of convection, rotation and magnetism leads to another set of interesting problems that can also have repercussions to everyday life.