Kunz Lecture Notes for the 2023 NSF/GPAP School on Plasma Physics for Astrophysicists

Matthew W. Kunz

Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544
Princeton Plasma Physics Laboratory, PO Box 451, Princeton, NJ 08543

(compiled on 2 June 2023)

These are supplementary lecture notes for the 2023 NSF/GPAP Summer School.

CONTENTS

PART I
Introduction to astrophysical plasmas

I.1. What is a plasma?

Astrophysical plasmas are remarkably varied, and so it may appear difficult at first to provide a definition of just what constitutes a "plasma". Is it an ionized, conducting gas? Well, the cold, molecular phase of the interstellar medium has a degree of ionization of $\lesssim 10^{-6}$, and yet is considered a plasma. (Indeed, plenty of researchers still model this phase using ideal magnetohydrodynamics!) Okay, so perhaps a sufficiently ionized, conducting gas (setting aside for now what is meant precisely by "sufficiently")? Well, plasmas don't necessarily have to be good conductors. Indeed, many frontier topics in plasma astrophysics involve situations in which resistivity is fundamentally important.

Clearly, any definition of a plasma must be accompanied by qualifiers, and these qualifiers are often cast in terms of dimensionless parameters that compare length and time scales. Perhaps the most important dimensionless parameter in the definition of a plasma is the *plasma parameter*,

$$\mathit{\Lambda} \doteq n_{\mathrm{e}} \lambda_{\mathrm{D}}^{3}, \tag{I.1}$$

where n_{e} is the electron number density and

$$\lambda_{\mathrm{D}} \doteq \left(\frac{T}{4\pi e^2 n_{\mathrm{e}}} \right)^{1/2} = 7.4 \left(\frac{T_{\mathrm{eV}}}{n_{\mathrm{cm^{-3}}}} \right)^{1/2} \ \mathrm{m} \tag{I.2}$$

is the Debye length. We'll derive this formula for the Debye length and discuss its physics more in § III.1 of these notes, but for now I'll simply state its meaning: it is the characteristic length scale on which the Coulomb potential of an individual charged particle is exponentially attenuated ("screened") by the preferential accumulation (exclusion) of oppositely- (like-) charged particles into (from) its vicinity.[1] Thus, $\mathit{\Lambda}$ reflects the number of electrons in a Debye sphere. Its dependence upon the temperature T suggests an alternative interpretation of $\mathit{\Lambda}$:

$$\mathit{\Lambda} = \frac{T}{4\pi e^2/\lambda_{\mathrm{D}}} \sim \frac{\text{kinetic energy}}{\text{potential energy}}. \tag{I.3}$$

Indeed, if the plasma is in thermodynamic equilibrium with a heat bath at temperature T, then the concentration of discrete charges follows the Boltzmann distribution,

$$n_\alpha(\boldsymbol{r}) = \overline{n}_\alpha \exp\left(-\frac{q_\alpha \phi(\boldsymbol{r})}{T} \right), \tag{I.4}$$

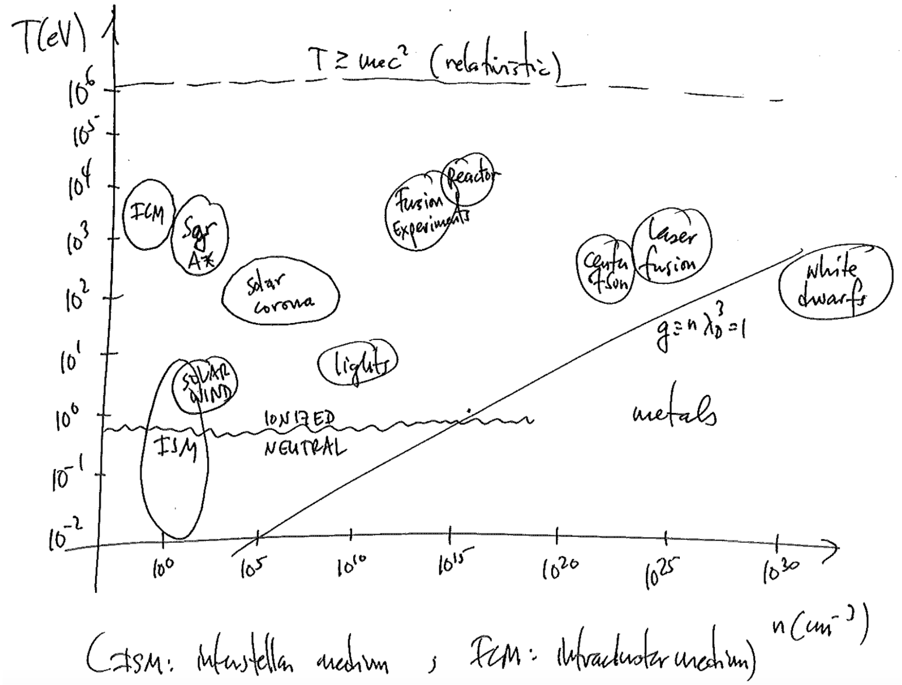
where $\overline{n}_\alpha$ is the mean number density of species α, q_α is its electric charge, and $\phi(\boldsymbol{r})$ is the Coulomb potential. In the limit $\mathit{\Lambda} \to \infty$, the distribution of charges becomes uniform, i.e., the plasma is said to be *quasi-neutral*, with equal amounts of positive and negative charge within a Debye sphere.

Debye shielding is fundamentally due to the polarization of the plasma and the associated redistribution of space charge, and is an example of how a plasma behaves as a dielectric medium. The hotter plasma, the more kinetic energy, the less bound individual electrons are to the protons. When $\mathit{\Lambda} \gg 1$, collective electrostatic interactions are much more important than binary particle–particle collisions, and the plasma is said to be

[1]In this course, sometimes temperature will be measured in Kelvin, and sometimes temperature will be measured in energy units (eV) after a hidden multiplication by Boltzmann's constant k_{B}. An energy of 1 eV corresponds to a temperature of $\sim 10^4$ K (more precisely, $\simeq 1.16 \times 10^4$ K).

weakly coupled. These are the types of plasmas that we will focus on in this course (e.g., the intracluster medium of galaxy clusters has $\Lambda \sim 10^{15}$).

Shown below is a rogue's gallery of astrophysical and space plasmas in the T–n plane, with the $\Lambda = 1$ line indicating a divide between quasi-neutral plasmas (to the left) and metals (to the right):



Clearly, there is a lot of parameter space here and so, to classify these plasmas further, we require additional dimensionless parameters.

I.2. Fundamental length and time scales

Another useful dividing line between different types of astrophysical and space plasmas is whether they are *collisional* or *collisionless.* In other words, is the mean free path between particle–particle collisions, $\lambda_{\rm mfp}$, larger or smaller than the macroscopic length scales of interest, L. If $\lambda_{\rm mfp} \ll L$, then the plasma is said to behave as a *fluid*, and various hydrodynamic and magnetohydrodynamic (MHD) equations can be used to describe its evolution. If, on the other hand, the mean free path is comparable to (or perhaps even larger than) the macroscopic length scales of interest, the plasma cannot be considered to be in local thermodynamic equilibrium, and the full six-dimensional phase space (3 spatial coordinates, 3 velocity coordinates) through which the constituent particles move must be retained in the description. Written in terms of the thermal speed of species α,

$$v_{{\rm th}\alpha} \doteq \left(\frac{2T_\alpha}{m_\alpha} \right)^{1/2}, \tag{I.5}$$

and the collision timescale τ_α, the *collisional mean free path* is

$$\lambda_{{\rm mfp},\alpha} \doteq v_{{\rm th}\alpha} \tau_\alpha . \tag{I.6}$$

For electron–ion collisions,

$$\tau_{\mathrm{ei}} = \frac{3\sqrt{m_{\mathrm{e}}}T_{\mathrm{e}}^{3/2}}{4\sqrt{2\pi}n_{\mathrm{e}}\lambda_{\mathrm{e}}Z^2e^4} \simeq 3.4\times10^5\left(\frac{T_{\mathrm{eV}}^{3/2}}{n_{\mathrm{cm^{-3}}}\lambda_{\mathrm{e}}Z^2}\right)\ \mathrm{s}, \tag{I.7}$$

where Ze is the ion charge and λ_{e} is the electron Coulomb logarithm; for ion–ion collisions,

$$\tau_{\mathrm{ii}} = \frac{3\sqrt{m_{\mathrm{i}}}T_{\mathrm{i}}^{3/2}}{4\sqrt{\pi}n_{\mathrm{i}}\lambda_{\mathrm{i}}Z^4e^4} \simeq 2.1\times10^7\left(\frac{T_{\mathrm{eV}}^{3/2}}{n_{\mathrm{cm^{-3}}}\lambda_{\mathrm{i}}Z^4}\right)\ \mathrm{s}, \tag{I.8}$$

where λ_{i} is the ion Coulomb logarithm. Note that the resulting $\lambda_{\mathrm{mfp,e}}$ and $\lambda_{\mathrm{mfp,i}}$ differ only by a factor of order unity:

$$\lambda_{\mathrm{mfp,e}} = \frac{3}{4\sqrt{\pi}}\frac{T_{\mathrm{e}}^2}{n_{\mathrm{e}}\lambda_{\mathrm{e}}Z^2e^4}, \quad \lambda_{\mathrm{mfp,i}} = \frac{3\sqrt{2}}{4\sqrt{\pi}}\frac{T_{\mathrm{i}}^2}{n_{\mathrm{i}}\lambda_{\mathrm{i}}Z^4e^4},$$

and so one often drops the species subscript on λ_{mfp}. With these definitions, it becomes clear that the plasma parameter (I.1) also reflects the ratio of the mean free path to the Debye length:

$$\Lambda \doteq \frac{n_{\mathrm{e}}\lambda_{\mathrm{D}}^4}{\lambda_{\mathrm{D}}} \sim \frac{T_{\mathrm{e}}^2/n_{\mathrm{e}}/e^4}{\lambda_{\mathrm{D}}} \sim \frac{\lambda_{\mathrm{mfp}}}{\lambda_{\mathrm{D}}}; \tag{I.9}$$

again, a measure of the relative importance of collective effects (λ_{D}) and binary collisions (λ_{mfp}).

Independent of whether a given astrophysical plasma is collisional or collisionless, nearly all such plasmas host magnetic fields, either inherited from the cosmic background in which they reside or produced *in situ* by a dynamo mechanism. There are two ways in which the strength of the magnetic field is quantified. First, the *plasma beta parameter*:

$$\beta_\alpha \doteq \frac{8\pi n_\alpha T_\alpha}{B^2}, \tag{I.10}$$

which reflects the relative energy densities of the thermal motions of the plasma particles and of the magnetic field. Note that

$$\beta_\alpha = \frac{2T_\alpha}{m_\alpha}\times\frac{4\pi m_\alpha n_\alpha}{B^2} = \frac{v_{\mathrm{th}\alpha}^2}{v_{\mathrm{A}\alpha}^2}, \tag{I.11}$$

where

$$v_{\mathrm{A}\alpha} \doteq \frac{B}{\sqrt{4\pi m_\alpha n_\alpha}} \tag{I.12}$$

is the *Alfvén speed* for species α.[2] Second, the *plasma magnetization*, ρ_α/L, where

$$\rho_\alpha \doteq \frac{v_{\mathrm{th}\alpha}}{\Omega_\alpha} \tag{I.13}$$

is the Larmor radius of species α and

$$\Omega_\alpha \equiv \frac{q_\alpha B}{m_\alpha c} \tag{I.14}$$

is the gyro- (or cyclotron, or Larmor) frequency. What distinguishes many astrophysical plasmas from their terrestrial laboratory counterparts is that the former can have $\beta \gg 1$ even though $\rho/L \lll 1$.[3] In other words, a magnetized astrophysical plasma need not have

---

[2] Usually, a single Alfvén speed, $v_{\mathrm{A}} \doteq B/\sqrt{4\pi\varrho}$, is given for a plasma with mass density ϱ.

[3] The $\sim$5 keV intracluster medium of galaxy clusters can be magnetized by a magnetic field as weak as $\sim10^{-18}$ G.

an energetically important magnetic field, and $\beta \gg 1$ does not preclude the magnetic field from having dynamical consequences. You've been warned.

There are two more kinetic scales worth mentioning at this point, which we will come to later in this course: the *plasma frequency*,

$$\omega_{\mathrm{p}\alpha} = \left(\frac{4\pi n_\alpha e^2}{m_\alpha} \right)^{1/2}, \tag{I.15}$$

and the *skin depth* (or *inertial length*),

$$d_\alpha \doteq \frac{c}{\omega_{\mathrm{p}\alpha}} = \left(\frac{m_\alpha c^2}{4\pi n_\alpha e^2} \right)^{1/2}. \tag{I.16}$$

The former is the characteristic frequency at which a plasma oscillates when one sign of charge carriers is displaced from the other sign by a small amount (see § III.2). Indeed, the factor $(4\pi n_\alpha e^2)$ should look familiar from the definition of the Debye length (see (I.2)). The latter is the characteristic scale below which the inertia of species α precludes the propagation of (certain) electromagnetic waves. For example, the ion skin depth is the scale at which the ions decouple from the electrons and any fluctuations in which the electrons are taking part (e.g., whistler waves). The following relationship between the skin depth and the Larmor radius may one day come in handy:

$$d_\alpha = \frac{v_{\mathrm{A},\alpha}}{\Omega_\alpha} = \frac{\rho_\alpha}{\beta_\alpha^{1/2}}. \tag{I.17}$$

I.3. Examples of astrophysical and space plasmas

This part is given as a keynote presentation. Here I simply provide a chart of useful numbers on the next page (ICM = intracluster medium; JET = Joint European Torus, a nuclear fusion experiment; ISM = interstellar medium). For quick reference, the Earth has a ∼0.5 G magnetic field, 1 eV ∼ 10^4 K, 1 au ≈ 1.5×10^{13} cm, 1 pc ≈ 3×10^{18} cm, 1 pc Myr^{-1} ≃ 1 km s^{-1}.

	solar wind @ 1 au (Earth location)	ICM @ ~100 kpc ("cooling radius")	galactic center @ 0.1 pc ("Bondi radius")	JET device (~1 meter)	ISM ("warm")
T	10 eV	8×10^3 eV	2×10^3 eV	10^4 eV	1 eV
n	10 cm^{-3}	5×10^{-3} cm^{-3}	100 cm^{-3}	10^{14} cm^{-3}	1 cm^{-3}
B	100 μG	1 μG	10^3 μG	3×10^4 G	5 μG

v_{thi}	40 km/s	1000 km/s	600 km/s	600 km/s	10 km/s
v_{Ai}	70 km/s	30 km/s	200 km/s	4000 km/s	10 km/s
$\beta_i \equiv \frac{v_{thi}^2}{v_{Ai}^2}$	$\sim 0.3-1$	$\sim 10^3$	~ 10	~ 0.02	~ 1
L	$\lesssim 1$ au	~ 10 kpc $- 100$ kpc	$\lesssim 0.1$ pc	~ 1 m	~ 1 pc $-$ 10's pc
λ_{mfp}	$\sim 0.1-1$ au	$\sim 0.1-10$ kpc	~ 0.01 pc	~ 10 km	$\sim 10^{-7}$ pc
ρ_i	$\sim 10^{-7}$ au	~ 1 npc	~ 1 ppc	~ 0.2 cm	$\sim 10^{-11}$ pc
Ω_i	~ 1 Hz	~ 0.01 Hz	~ 10 Hz	~ 300 MHz	~ 0.05 Hz

PART II

Fundamentals of hydrodynamics

Covered by Dr. Hosking, but here are some supplementary notes…

Unfortunately, fluid dynamics has all but disappeared from the US undergraduate curriculum, as physics departments have made way for quantum mechanics and condensed matter.[4] This is a shame – yes, it's classical physics and thus draws less 'oohs' and 'aahs' from the student (and professorial, for that matter) crowd. But there are many good reasons to study it. First, it forms the bedrock of fascinating and modern topics like non-equilibrium statistical mechanics, including the kinetic theory of gases and particles. Second, it is mathematically rich without being physically opaque. The more you really understand the mathematics, the more you really understand physically what is going on; the same cannot be said for many branches of modern physics. Third, nonlinear dynamics and chaos, burgeoning fields in their own right, are central to arguably the most important unsolved problem in classical physics: fluid turbulence. Solve that, and your solution would have immediate impact and practical benefits to society. Finally,

[4] An excellent textbook from which to learn elementary fluid dynamics is Acheson's *Elementary Fluid Dynamics*. It provides an engaging mix of history, physical insight, and transparent mathematics. I recommend it.

follow in the footsteps of greatness: on Feynman's chalkboard at the time of his death was the remit 'to learn ... nonlinear classical hydro'. With that, let's begin.

II.1. The equations of ideal hydrodynamics

The equations of hydrodynamics and MHD may be obtained rigorously by taking velocity-space moments of the Boltzmann and Vlasov–Landau kinetic equations. *Huh? What?* Okay, we'll get to that soon enough. For now, let's begin with things that you already know: mass is conserved, Newon's second law (force equals mass times acceleration), and the first law of thermodynamics (energy is conserved).

II.1.1. *Mass is conserved: The continuity equation*

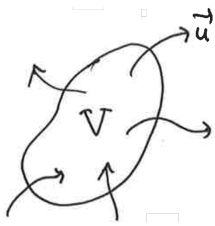
We describe our gaseous fluid by a mass density ρ, which in general is a function of time t and position $\boldsymbol{r}$.[5] Imagine an arbitrary volume $\mathcal{V}$ enclosing some of that fluid. The mass inside of the volume is simply

$$M = \int_{\mathcal{V}} \mathrm{d}V \, \rho. \tag{II.1}$$

Now let's mathematize our intuition: within this fixed volume, the only way the enclosed mass can change is by material flowing in or out of its surface $\mathcal{S}$:

$$\frac{\mathrm{d}M}{\mathrm{d}t} \doteq \int_{\mathcal{V}} \mathrm{d}V \, \frac{\partial \rho}{\partial t} = -\int_{\mathcal{S}} \mathrm{d}\boldsymbol{S} \boldsymbol{\cdot} \rho \boldsymbol{u}, \tag{II.2}$$

where $\boldsymbol{u}$ is the flow velocity.



Gauss' theorem may be applied to rewrite the right-hand side of this equation as follows:

$$\int_{\mathcal{S}} \mathrm{d}\boldsymbol{S} \boldsymbol{\cdot} \rho \boldsymbol{u} = \int_{\mathcal{V}} \mathrm{d}V \, \boldsymbol{\nabla} \boldsymbol{\cdot} (\rho \boldsymbol{u}). \tag{II.3}$$

Because the volume under consideration is arbitrary, the integrands of the volume integrals in (II.2) and (II.3) must be the same. Therefore,

$$\boxed{\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \boldsymbol{\cdot} (\rho \boldsymbol{u}) = 0} \tag{II.4}$$

This is the *continuity equation*; it's the differential form of mass conservation.

Exercise. Go to the bathroom and turn on the sink slowly to get a nice, laminar stream flowing down from the faucet. Go on, I'll wait. If you followed instructions, then you'll see that the stream becomes more narrow as it descends. Knowing that the density of water is very nearly

[5] I sometimes denote the mass density by ϱ to avoid confusion with the Larmor radius ρ. But, given that ρ is standard notation in hydrodynamics for the mass density, and ρ is standard notation in plasma physics for the Larmor radius, you should learn to tell the difference based on the context.

constant, use the continuity equation to show that the cross-sectional area of the stream $A(z)$ as a function of distance from the faucet z is

$$A(z) = \frac{A_0}{\sqrt{1 + 2gz/v_0^2}},$$

where A_0 is the cross-sectional area of the stream upon exiting the faucet with velocity v_0 and g is the gravitational acceleration. If you turn the faucet to make the water flow faster, what happens to the tapering of the stream?

II.1.2. *Newton's second law: The momentum equation*

So far we have an equation for the evolution of the mass density ρ expressed in terms of the fluid velocity $\boldsymbol{u}$. How does the latter evolve? Newton's second law provides the answer: simply add up the accelerations, divide by the mass (density), and you've got the time rate of change of the velocity. But there is a subtlety here: there is a difference between the time rate of change of the velocity in the lab frame and the time rate of change of the velocity in the fluid frame. So which time derivative of $\boldsymbol{u}$ do we take? The key is in how the accelerations are expressed. Are these accelerations acting on a fixed point in space, or are they acting on an element of our fluid? It is much easier (and more physical) to think of these accelerations in the latter sense: given a deformable patch of the fluid – large enough in extent to contain a very large number of atoms but small enough that all the macroscopic variables such as density, velocity, and pressure have a unique value over the dimensions of the patch – what forces are acting on that patch? These are relatively simple to catalog, and we will do so in short order. But first, let's answer our original question: which time derivative of $\boldsymbol{u}$ do we take? Since we have committed to expressing the forces in the frame of the fluid element, the acceleration must likewise be expressed in this frame. The acceleration is *not*

$$\frac{\partial \boldsymbol{u}}{\partial t}. \tag{II.5}$$

Remember what a partial derivative means: something is being fixed! Here, it is the instantaneous position $\boldsymbol{r}$ of the fluid element. Equation (II.5) is the answer to the question, 'how does the fluid velocity evolve at a fixed point in space?' Instead, we wish to fix our sights on the fluid element itself, which is moving. The acceleration we calculate must account for this frame transformation:

$$\boldsymbol{a} = \frac{\partial \boldsymbol{u}}{\partial t} + \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} \boldsymbol{\cdot} \boldsymbol{\nabla} \boldsymbol{u}, \tag{II.6}$$

where $\mathrm{d}\boldsymbol{r}/\mathrm{d}t$ is the rate of change of the position of the fluid element, i.e., the velocity $\boldsymbol{u}(t, \boldsymbol{r})$. This combination of derivatives is so important that it has its own notation:

$$\frac{\mathrm{D}}{\mathrm{D}t} \doteq \frac{\partial}{\partial t} + \boldsymbol{u} \boldsymbol{\cdot} \boldsymbol{\nabla}. \tag{II.7}$$

It is variously referred to as the *Lagrangian derivative*, or comoving derivative, or convective derivative. By contrast, the expression given by (II.5) is the *Eulerian deriative.* Note that the continuity equation (II.4) may be expressed using the Lagrangian derivative as

$$\frac{\mathrm{D} \ln \rho}{\mathrm{D}t} = -\boldsymbol{\nabla} \boldsymbol{\cdot} \boldsymbol{u}, \tag{II.8}$$

which states that incompressible flow corresponds to $\boldsymbol{\nabla} \boldsymbol{\cdot} \boldsymbol{u} = 0$.

So, given some force $\boldsymbol{F}$ per unit volume that is acting on our fluid element, we now know how the fluid velocity evolves: force (per unit volume) equals mass (per unit volume) times acceleration (in the frame of the fluid element):

$$\boldsymbol{F} = \rho \frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t}. \tag{II.9}$$

Now we need only catalog the relevant forces. This could be, say, gravity: $\rho\boldsymbol{g} = -\rho\boldsymbol{\nabla}\Phi$. Or, if the fluid element is conducting, electromagnetic forces (which we'll get to later in the course). But the most deserving of discussion at this point is the pressure force due to the internal thermal motions of the particles comprising the gas. For an ideal gas, the equation of state is

$$P = \frac{\rho k_{\mathrm{B}} T}{m} \doteq \rho C^2, \tag{II.10}$$

where T is the temperature in Kelvin, k_{B} is the Boltzmann constant, m is the mass per particle, and C is the speed of sound in an isothermal gas. Plasma physicists often drop Boltzmann's constant and register temperature in energy units (e.g., eV), and I will henceforth do the same in these notes. How does gas pressure due to microscopic particle motions exert a macroscopic force on a fluid element? First, the pressure must be spatially non-uniform: there must be more or less energetic content in the thermal motions of the particles in one region versus another, whether it be because the gas temperature varies in space or because there are more particles in one location as opposed to another. For example, the pressure force in the x direction in a slab of thickness $\mathrm{d}x$ and cross-sectional area $\mathrm{d}y\,\mathrm{d}z$ is

$$\big[P(t,x-\mathrm{d}x/2,y,z) - P(t,x+\mathrm{d}x/2,y,z)\big]\,\mathrm{d}y\,\mathrm{d}z = -\frac{\partial P}{\partial x}\,\mathrm{d}V. \tag{II.11}$$

Unless the thermal motions of the particles are not sufficiently randomized to be isotropic (e.g., if the collisional mean free path of the plasma is so long that inter-particle collisions cannot drive the system quickly enough towards local thermodynamic equilibrium), there is nothing particularly special about the x direction, and so the pressure force force acting on some differential volume $\mathrm{d}V$ is just $-\boldsymbol{\nabla}P\,\mathrm{d}V$.

Assembling the lessons we've learned here, we have the following force equation for our fluid:

$$\boxed{\rho\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} \doteq \rho\left(\frac{\partial}{\partial t} + \boldsymbol{u}\boldsymbol{\cdot}\boldsymbol{\nabla}\right)\boldsymbol{u} = -\boldsymbol{\nabla}P - \rho\boldsymbol{\nabla}\Phi} \tag{II.12}$$

This equation is colloquially known as the *momentum equation*, even though it evolves the fluid velocity rather than its momentum density. To obtain an equation for the latter, the continuity equation (II.4) may be used to move the mass density into the time and space derivatives:

$$\begin{aligned}
\frac{\partial(\rho\boldsymbol{u})}{\partial t} + \boldsymbol{\nabla}\boldsymbol{\cdot}(\rho\boldsymbol{u}\boldsymbol{u}) &= \frac{\partial\rho}{\partial t}\boldsymbol{u} + \rho\frac{\partial\boldsymbol{u}}{\partial t} + \rho\boldsymbol{u}\boldsymbol{\cdot}\boldsymbol{\nabla}\boldsymbol{u} + \boldsymbol{\nabla}\boldsymbol{\cdot}(\rho\boldsymbol{u})\boldsymbol{u} \\
&= \left[\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\boldsymbol{\cdot}(\rho\boldsymbol{u})\right]\boldsymbol{u} + \rho\left(\frac{\partial}{\partial t} + \boldsymbol{u}\boldsymbol{\cdot}\boldsymbol{\nabla}\right)\boldsymbol{u} \\
&= \left[\qquad 0 \qquad\right]\boldsymbol{u} + \rho\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} = \boldsymbol{F}.
\end{aligned} \tag{II.13}$$

Thus, an equation for the momentum density:

$$\boxed{\frac{\partial(\rho\boldsymbol{u})}{\partial t} + \boldsymbol{\nabla}\boldsymbol{\cdot}(\rho\boldsymbol{u}\boldsymbol{u}) = -\boldsymbol{\nabla}P - \rho\boldsymbol{\nabla}\Phi} \tag{II.14}$$

This form is particularly useful for deriving an evolution equation for the kinetic energy density. Dotting (II.14) with $\boldsymbol{u}$ and grouping terms,

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\rho u^2\right) + \boldsymbol{\nabla}\boldsymbol{\cdot}\left(\frac{1}{2}\rho u^2\boldsymbol{u}\right) = -\boldsymbol{u}\boldsymbol{\cdot}\boldsymbol{\nabla}P - \rho\boldsymbol{u}\boldsymbol{\cdot}\boldsymbol{\nabla}\Phi, \tag{II.15}$$

which is a statement that the kinetic energy of a fluid element changes as work is done by the forces.

Now, how to we know the pressure P? There's an equation for that…

II.1.3. *First law of thermodynamics: The internal energy equation*

There are several ways to go about obtaining an evolution equation for the pressure. One way is to introduce the *internal energy*,

$$e \doteq \frac{P}{\gamma - 1} \tag{II.16}$$

and use the first law of thermodynamics to argue that e is conserved but for $P\,\mathrm{d}V$ work:

$$\boxed{\frac{\partial e}{\partial t} + \boldsymbol{\nabla}\boldsymbol{\cdot}(e\boldsymbol{u}) = -P\boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{u}} \tag{II.17}$$

This is the *internal energy* equation.

Equation (II.17) may be used to derive a total (kinetic + internal + potential) energy equation for the fluid as follows. Do (II.15) + (II.17):

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{1}{2}\rho u^2 + e\right) + \boldsymbol{\nabla}\boldsymbol{\cdot}\left[\left(\frac{1}{2}\rho u^2 + e\right)\boldsymbol{u}\right] &= -\boldsymbol{\nabla}\boldsymbol{\cdot}(P\boldsymbol{u}) - \rho\boldsymbol{u}\boldsymbol{\cdot}\boldsymbol{\nabla}\Phi,\\ &= -(\gamma-1)\boldsymbol{\nabla}\boldsymbol{\cdot}(e\boldsymbol{u}) - \rho\boldsymbol{u}\boldsymbol{\cdot}\boldsymbol{\nabla}\Phi\end{aligned}$$

$$\implies \quad \frac{\partial}{\partial t}\left(\frac{1}{2}\rho u^2 + e\right) + \boldsymbol{\nabla}\boldsymbol{\cdot}\left[\left(\frac{1}{2}\rho u^2 + \gamma e\right)\boldsymbol{u}\right] = -\rho\boldsymbol{u}\boldsymbol{\cdot}\boldsymbol{\nabla}\Phi. \tag{II.18}$$

Now use the continuity equation (II.4) to write

$$\frac{\partial(\rho\Phi)}{\partial t} + \boldsymbol{\nabla}\boldsymbol{\cdot}(\rho\Phi\boldsymbol{u}) = \rho\boldsymbol{u}\boldsymbol{\cdot}\boldsymbol{\nabla}\Phi + \rho\frac{\partial\Phi}{\partial t}. \tag{II.19}$$

Adding this equation to (II.18) yields the desired result:

$$\boxed{\frac{\partial}{\partial t}\left(\frac{1}{2}\rho u^2 + e + \rho\Phi\right) + \boldsymbol{\nabla}\boldsymbol{\cdot}\left[\left(\frac{1}{2}\rho u^2 + \gamma e + \rho\Phi\right)\boldsymbol{u}\right] = \rho\frac{\partial\Phi}{\partial t}} \tag{II.20}$$

The first term in parentheses under the time derivative is sometimes denoted by $\mathcal{E}$.

Yet another way of expressing the internal energy equation (II.17) is to write $e = \rho T/m(\gamma-1)$ and use the continuity equation (II.4) to eliminate the derivatives of the mass density. The result is

$$\frac{\mathrm{D}\ln T}{\mathrm{D}t} = -(\gamma-1)\boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{u}, \tag{II.21}$$

which states that the temperature of a fluid element is constant in an incompressible fluid (*viz.*, one with $\boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{u} = 0$). If this seems intuitively unfamiliar to you, consider this: the hydrodynamic entropy of a fluid element is given by

$$s \doteq \frac{1}{\gamma-1}\ln P\rho^{-\gamma} = \frac{1}{\gamma-1}\ln T\rho^{1-\gamma}. \tag{II.22}$$

Taking the Lagrangian time derivative of the entropy along the path of a fluid element yields

$$\frac{\mathrm{D}s}{\mathrm{D}t} = \frac{\mathrm{D}\ln T}{\mathrm{D}t} - (\gamma - 1)\frac{\mathrm{D}\ln \rho}{\mathrm{D}t}. \tag{II.23}$$

It is then just a short trip back to (II.8) to see that (II.21) is, in fact, the second law of thermodynamics – entropy is conserved in the absence of sources or dissipative sinks:

$$\boxed{\frac{\mathrm{D}s}{\mathrm{D}t} = 0} \tag{II.24}$$

II.2. Summary: Adiabatic equations of hydrodynamics

The adiabatic equations of hydrodynamics, written in conservative form, are:

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla \cdot} (\rho \boldsymbol{u}) = 0, \tag{II.25a}$$

$$\frac{\partial (\rho \boldsymbol{u})}{\partial t} + \boldsymbol{\nabla \cdot} (\rho \boldsymbol{u}\boldsymbol{u}) = -\boldsymbol{\nabla} P - \rho \boldsymbol{\nabla} \varPhi, \tag{II.25b}$$

$$\frac{\partial e}{\partial t} + \boldsymbol{\nabla \cdot} (e \boldsymbol{u}) = -P \boldsymbol{\nabla \cdot} \boldsymbol{u}. \tag{II.25c}$$

The left-hand sides of these equations express advection of, respectively, the mass density, the momentum density, and the internal energy density by the fluid velocity; the right-hand sides represents sources and sinks. If the gravitational potential is due to self-gravity, then one must additionally solve the Poisson equation,

$$\nabla^2 \varPhi = 4\pi G \rho. \tag{II.26}$$

where G is Newton's gravitational constant.

If we instead write these equations in terms of the density, fluid velocity, and entropy and make use of the Lagrangian derivative (II.7), we have

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} = -\rho \boldsymbol{\nabla \cdot} \boldsymbol{u}, \tag{II.27a}$$

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} = -\frac{1}{\rho}\boldsymbol{\nabla} P - \boldsymbol{\nabla} \varPhi, \tag{II.27b}$$

$$\frac{\mathrm{D}s}{\mathrm{D}t} = 0, \tag{II.27c}$$

where $s \doteq (\gamma - 1)^{-1} \ln P\rho^{-\gamma}$. The limit $\gamma \rightarrow \infty$, often of utility for describing liquids, corresponds to $\mathrm{D}\rho/\mathrm{D}t = 0$, i.e., incompressibility.

Exercise. Show that the gravitational force on a self-gravitating fluid element may be written as

$$-\rho \boldsymbol{\nabla} \varPhi = -\boldsymbol{\nabla \cdot} \left(\frac{\boldsymbol{g}\boldsymbol{g}}{4\pi G} - \frac{g^2}{8\pi G}\boldsymbol{I} \right), \tag{II.28}$$

where $\boldsymbol{g} = -\boldsymbol{\nabla}\varPhi$, $g^2 = \boldsymbol{g \cdot g}$, and $\boldsymbol{I}$ is the unit dyadic. The quantity inside the divergence operator is known as the gravitational stress tensor. Because it's written in the form of a divergence, it represents the flux of total momentum through a surface due to gravitational forces.

II.3. Mathematical matters

II.3.1. *Vector identities*

As a start to this section, let me advise you to brush up on your vector calculus…

$$
\begin{aligned}
\boldsymbol{A}\cdot(\boldsymbol{B}\times\boldsymbol{C}) &= \boldsymbol{B}\cdot(\boldsymbol{C}\times\boldsymbol{A}) = \boldsymbol{C}\cdot(\boldsymbol{A}\times\boldsymbol{B}),\\
\boldsymbol{A}\times(\boldsymbol{B}\times\boldsymbol{C}) &= \boldsymbol{B}(\boldsymbol{A}\cdot\boldsymbol{C}) - \boldsymbol{C}(\boldsymbol{A}\cdot\boldsymbol{B}),\\
\boldsymbol{\nabla}\cdot(\boldsymbol{A}\times\boldsymbol{B}) &= \boldsymbol{B}\cdot(\boldsymbol{\nabla}\times\boldsymbol{A}) - \boldsymbol{A}\cdot(\boldsymbol{\nabla}\times\boldsymbol{B}),\\
\boldsymbol{\nabla}\times(\boldsymbol{A}\times\boldsymbol{B}) &= (\boldsymbol{B}\cdot\boldsymbol{\nabla})\boldsymbol{A} - (\boldsymbol{A}\cdot\boldsymbol{\nabla})\boldsymbol{B} - \boldsymbol{B}(\boldsymbol{\nabla}\cdot\boldsymbol{A}) + \boldsymbol{A}(\boldsymbol{\nabla}\cdot\boldsymbol{B}),\\
\boldsymbol{A}\times(\boldsymbol{\nabla}\times\boldsymbol{B}) + \boldsymbol{B}\times(\boldsymbol{\nabla}\times\boldsymbol{A}) &= \boldsymbol{\nabla}(\boldsymbol{A}\cdot\boldsymbol{B}) - (\boldsymbol{A}\cdot\boldsymbol{\nabla})\boldsymbol{B} - (\boldsymbol{B}\cdot\boldsymbol{\nabla})\boldsymbol{A},\\
&\cdots
\end{aligned}
$$

Fluid dynamics is full of these things, and you should either (i) commit them to memory, (ii) carry your NRL formulary with you everywhere, or (iii) know how to quickly derive them using things like

$$\epsilon_{kij}\epsilon_{k\ell m} = \delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{j\ell},$$

where δ_{ij} is the Kronecker delta and ϵ_{ijk} is the Levi–Civita symbol.

II.3.2. *Leibniz's rule and the Lagrangian derivative of integrals*

In the proofs of many conservation laws, a Lagrangian time derivative is taken of a surface or volume integral whose integration limits are time-dependent. In this case, $\mathrm{D}/\mathrm{D}t$ does *not* commute with the integral sign. The trick to dealing with these situations is related to Leibniz's rule:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{a(t)}^{b(t)}\mathrm{d}x\, f(t,x) = \int_{a(t)}^{b(t)}\mathrm{d}x\,\frac{\partial}{\partial t}f(t,x) + f(t,b(t))\frac{\mathrm{d}b}{\mathrm{d}t} - f(t,a(t))\frac{\mathrm{d}a}{\mathrm{d}t}. \tag{II.29}$$

In three dimensions, if we're taking the time derivative of a volume integral whose integration limits $\mathcal{V}(t)$ are time-dependent, the generalization of the above is

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{V}(t)}\mathrm{d}\mathcal{V}\, f(t,\boldsymbol{r}) = \int_{\mathcal{V}(t)}\mathrm{d}\mathcal{V}\,\frac{\partial}{\partial t}f(t,\boldsymbol{r}) + \oint_{\partial\mathcal{V}(t)}\mathrm{d}\boldsymbol{\mathcal{S}}\cdot\bigl[f(t,\boldsymbol{r})\boldsymbol{u}_{\mathrm{b}}(t,\boldsymbol{r})\bigr], \tag{II.30}$$

where $\boldsymbol{u}_{\mathrm{b}}$ is the velocity of the bounding surface $\partial\mathcal{V}(t)$. This is known as the Reynolds transport theorem. In words, the time rate-of-change of a quantity positioned within a moving volume is a combination of the lab-frame rate-of-change of that quantity (i.e., the time derivative at fixed position $\boldsymbol{r}$ – note the partial derivative) and how much of that quantity flowed through the surface. When the velocity of the bounding surface equals the fluid velocity, $\boldsymbol{u}_{\mathrm{b}} = \boldsymbol{u}(t,\boldsymbol{r})$, so that each moving volume corresponds to that of a fluid element, we may replace $\mathrm{d}/\mathrm{d}t$ in (II.30) with the Lagrangian derivative $\mathrm{D}/\mathrm{D}t$:

$$\boxed{\frac{\mathrm{D}}{\mathrm{D}t}\int_{\mathcal{V}(t)}\mathrm{d}\mathcal{V}\, f(t,\boldsymbol{r}) = \int_{\mathcal{V}(t)}\mathrm{d}\mathcal{V}\,\frac{\partial}{\partial t}f(t,\boldsymbol{r}) + \oint_{\partial\mathcal{V}(t)}\mathrm{d}\boldsymbol{\mathcal{S}}\cdot\bigl[f(t,\boldsymbol{r})\boldsymbol{u}(t,\boldsymbol{r})\bigr]} \tag{II.31}$$

You've already encountered an example of this – mass conservation, in which the volume was a "material volume" moving with the fluid element itself:

$$0 = \frac{\mathrm{D}M}{\mathrm{D}t} \doteq \frac{\mathrm{D}}{\mathrm{D}t}\int_{\mathcal{V}(t)}\mathrm{d}\mathcal{V}\,\rho = \int_{\mathcal{V}(t)}\mathrm{d}\mathcal{V}\,\frac{\partial\rho}{\partial t} + \oint_{\partial\mathcal{V}(t)}\mathrm{d}\boldsymbol{\mathcal{S}}\cdot(\rho\boldsymbol{u}).$$

Using the divergence theorem on the final (surface-integral) term gives

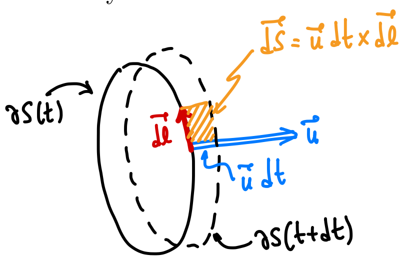
$$0 = \int_{\mathcal{V}(t)} \mathrm{d}\mathcal{V} \left[\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \boldsymbol{\cdot} (\rho \boldsymbol{u}) \right],$$

which provides us with our continuity equation.

A similar rule to (II.31) is needed for time derivatives of surface integrals whose integration limits $\mathcal{S}(t)$ are time-dependent. For a vector field $\boldsymbol{F} = \boldsymbol{F}(t, \boldsymbol{r})$ and a bounding surface $\mathcal{S}(t)$ whose contour $\partial\mathcal{S}(t)$ moves with the fluid velocity $\boldsymbol{u} = \boldsymbol{u}(t, \boldsymbol{r})$, this is given by

$$\boxed{\frac{\mathrm{D}}{\mathrm{D}t} \int_{\mathcal{S}(t)} \mathrm{d}\boldsymbol{\mathcal{S}} \boldsymbol{\cdot} \boldsymbol{F} = \int_{\mathcal{S}(t)} \mathrm{d}\boldsymbol{\mathcal{S}} \boldsymbol{\cdot} \left[\frac{\partial \boldsymbol{F}}{\partial t} + (\boldsymbol{\nabla} \boldsymbol{\cdot} \boldsymbol{F}) \boldsymbol{u} \right] - \oint_{\partial\mathcal{S}(t)} \mathrm{d}\boldsymbol{\ell} \boldsymbol{\cdot} (\boldsymbol{u} \times \boldsymbol{F})} \tag{II.32}$$

(By convention, the contour is taken in the counter-clockwise direction.) Note that $-\mathrm{d}\boldsymbol{\ell} \boldsymbol{\cdot} (\boldsymbol{u} \times \boldsymbol{F}) = \boldsymbol{F} \boldsymbol{\cdot} (\boldsymbol{u} \times \mathrm{d}\boldsymbol{\ell})$. In words, the comoving change of the differential surface element $\mathrm{d}\boldsymbol{\mathcal{S}}$ equals the amount of area swept out in a time $\mathrm{d}t$ via the advection of a differential line element $\mathrm{d}\boldsymbol{\ell}$ on $\partial\mathcal{S}$ by a distance $\boldsymbol{u}\,\mathrm{d}t$:



Equation (II.32) can be used to prove conservation of magnetic flux (§IV.1.1) and conservation of fluid vorticity (§II.4).

II.3.3. $\boldsymbol{u} \boldsymbol{\cdot} \boldsymbol{\nabla} \boldsymbol{u}$ *and curvilinear coordinates*

Finally, the nonlinear combination $\boldsymbol{u} \boldsymbol{\cdot} \boldsymbol{\nabla} \boldsymbol{u}$ that features prominently in the Lagrangian time derivative can be complicated, particularly in curvilinear coordinates where the gradient operator within it acts on the unit vectors within $\boldsymbol{u}$. For example, in cylindrical coordinates (R, φ, z),

$$\begin{aligned}
\boldsymbol{u} \boldsymbol{\cdot} \boldsymbol{\nabla} \boldsymbol{u} &= \boldsymbol{u} \boldsymbol{\cdot} \boldsymbol{\nabla} \big(u_R \hat{\boldsymbol{R}} + u_\varphi \hat{\boldsymbol{\varphi}} + u_z \hat{\boldsymbol{z}} \big) \\
&= (\boldsymbol{u} \boldsymbol{\cdot} \boldsymbol{\nabla} u_R) \hat{\boldsymbol{R}} + (\boldsymbol{u} \boldsymbol{\cdot} \boldsymbol{\nabla} u_\varphi) \hat{\boldsymbol{\varphi}} + (\boldsymbol{u} \boldsymbol{\cdot} \boldsymbol{\nabla} u_z) \hat{\boldsymbol{z}} + \frac{u_\varphi^2}{R} \frac{\partial \hat{\boldsymbol{\varphi}}}{\partial \varphi} + \frac{u_R u_\varphi}{R} \frac{\partial \hat{\boldsymbol{R}}}{\partial \varphi} \\
&= (\boldsymbol{u} \boldsymbol{\cdot} \boldsymbol{\nabla} u_i) \hat{\boldsymbol{e}}_i - \frac{u_\varphi^2}{R} \hat{\boldsymbol{R}} + \frac{u_R u_\varphi}{R} \hat{\boldsymbol{\varphi}},
\end{aligned} \tag{II.33}$$

where, to obtain the final equality, we have used $\partial\hat{\boldsymbol{\varphi}}/\partial\varphi = -\hat{\boldsymbol{R}}$ and $\partial\hat{\boldsymbol{R}}/\partial\varphi = \hat{\boldsymbol{\varphi}}$; summation over the repeated index i is implied in the first term in the final line.

Exercise. Follow a similar procedure to show that, in spherical coordinates (r, θ, φ),

$$\begin{aligned}
\boldsymbol{u} \boldsymbol{\cdot} \boldsymbol{\nabla} \boldsymbol{u} &= \left(u_r \frac{\partial}{\partial r} + \frac{u_\theta}{r} \frac{\partial}{\partial \theta} + \frac{u_\varphi}{r \sin\theta} \frac{\partial}{\partial \varphi} \right) \big(u_r \hat{\boldsymbol{r}} + u_\theta \hat{\boldsymbol{\theta}} + u_\varphi \hat{\boldsymbol{\varphi}} \big) \\
&= (\boldsymbol{u} \boldsymbol{\cdot} \boldsymbol{\nabla} u_i) \hat{\boldsymbol{e}}_i - \frac{u_\theta^2 + u_\varphi^2}{r} \hat{\boldsymbol{r}} + \left(\frac{u_r u_\theta}{r} - \frac{u_\varphi^2 \cot\theta}{r} \right) \hat{\boldsymbol{\theta}} + \left(\frac{u_\theta u_\varphi \cot\theta}{r} + \frac{u_r u_\varphi}{r} \right) \hat{\boldsymbol{\varphi}}.
\end{aligned}$$

The last two terms in the cylindrical $\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u}$, equation (II.33), might look familiar to you from working in rotating frames. Indeed, let us write $\boldsymbol{u} = \boldsymbol{v} + R\Omega(R,z)\hat{\boldsymbol{\varphi}}$, where Ω is an angular velocity, and substitute this decomposition into (II.33):

$$\begin{aligned}\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} &= \bigl[(\boldsymbol{v}+R\Omega\hat{\boldsymbol{\varphi}})\cdot\boldsymbol{\nabla}v_i\bigr]\hat{\boldsymbol{e}}_i + \bigl[(\boldsymbol{v}+R\Omega\hat{\boldsymbol{\varphi}})\cdot\boldsymbol{\nabla}(R\Omega)\bigr]\hat{\boldsymbol{\varphi}} \\ &\quad - \frac{(v_\varphi+R\Omega)^2}{R}\hat{\boldsymbol{R}} + \frac{v_R(v_\varphi+R\Omega)}{R}\hat{\boldsymbol{\varphi}} \\ &= \left[\left(\boldsymbol{v}\cdot\boldsymbol{\nabla}+\Omega\frac{\partial}{\partial\varphi}\right)v_i\right]\hat{\boldsymbol{e}}_i + \left[2\Omega\hat{\boldsymbol{z}}\times\boldsymbol{v} - R\Omega^2\hat{\boldsymbol{R}} + R\hat{\boldsymbol{\varphi}}(\boldsymbol{v}\cdot\boldsymbol{\nabla})\Omega\right] \\ &\quad + \left[\frac{v_R v_\varphi}{R}\hat{\boldsymbol{\varphi}} - \frac{v_\varphi^2}{R}\hat{\boldsymbol{R}}\right]. \end{aligned} \tag{II.34}$$

Each of these terms has a straightforward physical interpretation. The first term in brackets represents advection by the flow and the rotation. The second term in brackets contains the Coriolis force, the centrifugal force, and 'tidal' terms due to the differential rotation, in that order. (The 'tidal' terms can be thought of the fictitious acceleration required for a fluid element to maintain its presence in the local rotating frame as it is displaced radially or vertically. They come from Taylor expanding the angular velocity about a point in the disk.) The third and final term in brackets captures curvature effects due to the cylindrical geometry.

Exercise. Show that the $R\varphi$-component in cylindrical coordinates of the rate-of-strain tensor

$$W_{ij} \doteq \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3}\delta_{ij}\frac{\partial u_k}{\partial x_k}$$

is given by

$$W_{R\varphi} = \frac{1}{R}\frac{\partial u_R}{\partial\varphi} + R\frac{\partial}{\partial R}\frac{u_\varphi}{R}.$$

Hint: $\partial u_i/\partial x_j = [(\hat{\boldsymbol{e}}_j\cdot\boldsymbol{\nabla})\boldsymbol{u}]\cdot\hat{\boldsymbol{e}}_i$ is coordinate invariant.

II.4. Vorticity and Kelvin's circulation theorem

With some vector identities in hand, let's take the curl of the force equation (II.27*b*):

$$\boldsymbol{\nabla}\times\left(\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} = -\frac{1}{\rho}\boldsymbol{\nabla}P - \boldsymbol{\nabla}\Phi\right).$$

The potential term vanishes, since the curl of a gradient is zero. Likewise, the pressure term becomes

$$-\boldsymbol{\nabla}\frac{1}{\rho}\times\boldsymbol{\nabla}P = \frac{1}{\rho^2}\boldsymbol{\nabla}\rho\times\boldsymbol{\nabla}P.$$

As for the left-hand side, the gradient operator commutes with $\partial/\partial t$, but not with $\boldsymbol{u}\cdot\boldsymbol{\nabla}$. Instead,

$$\boldsymbol{\nabla}\times\bigl[(\boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{u}\bigr] = \boldsymbol{\nabla}\times\left[\frac{1}{2}\boldsymbol{\nabla}u^2 - \boldsymbol{u}\times(\boldsymbol{\nabla}\times\boldsymbol{u})\right] = -\boldsymbol{\nabla}\times(\boldsymbol{u}\times\boldsymbol{\omega}),$$

where

$$\boldsymbol{\omega} \doteq \boldsymbol{\nabla}\times\boldsymbol{u} \tag{II.35}$$

is the *fluid vorticity*. The vorticity measures how much rotation a velocity field has (and its direction). Note that it is divergence free, which means that vortex lines cannot end within the fluid – they must either close on themselves (like a smoke ring) or intersect a boundary (like a tornado). Any fresh vortex lines that are made must be created as continuous curves that grow out of points or lines where the vorticity vanishes.

Assembling the above gives the *vorticity equation*,

$$\frac{\partial \boldsymbol{\omega}}{\partial t} - \boldsymbol{\nabla} \times (\boldsymbol{u} \times \boldsymbol{\omega}) = \frac{1}{\rho^2} \boldsymbol{\nabla} \rho \times \boldsymbol{\nabla} P. \tag{II.36}$$

Note that the right-hand side of this equation vanishes if the pressure is *barotropic*, i.e., if $P = P(\rho)$, so that surfaces of constant density and constant pressure coincide. If these surfaces do not coincide, then the fluid is said to have "baroclinicity" or to be "baroclinic". I'll demonstrate below using mathematics what (II.36) means physically, but you already know what the right-hand side means if you pay attention to the weather: areas of high atmospheric baroclinicity have frequent hurricanes and cyclones. In the parlance of fluid dynamics, this is called "baroclinic forcing". Now back to the math...

Dot (II.36) into a differential surface element $\mathrm{d}\boldsymbol{\mathcal{S}}$ normal to the surface $\mathcal{S}$ of a fluid element, integrate over that surface, and use Stokes' theorem to replace the surface integral of a curl with a line integral over the surface boundary $\partial\mathcal{S}$:

$$\int_{\mathcal{S}} \frac{\partial \boldsymbol{\omega}}{\partial t} \boldsymbol{\cdot} \mathrm{d}\boldsymbol{\mathcal{S}} - \oint_{\partial\mathcal{S}} (\boldsymbol{u} \times \boldsymbol{\omega}) \boldsymbol{\cdot} \mathrm{d}\boldsymbol{\ell} = \oint_{\partial\mathcal{S}} \left(-\frac{1}{\rho} \boldsymbol{\nabla} P \right) \boldsymbol{\cdot} \mathrm{d}\boldsymbol{\ell} = -\oint_{\partial\mathcal{S}} \frac{\mathrm{d}P}{\rho}.$$

Using (II.32) to replace the left-hand side by the Lagrangian time derivative of $\boldsymbol{\omega} \boldsymbol{\cdot} \mathrm{d}\boldsymbol{\mathcal{S}}$ yields

$$\frac{\mathrm{D}}{\mathrm{D}t} \int_{\mathcal{S}} \boldsymbol{\omega} \boldsymbol{\cdot} \mathrm{d}\boldsymbol{\mathcal{S}} = -\oint_{\partial\mathcal{S}} \frac{\mathrm{d}P}{\rho}. \tag{II.37}$$

The surface integral on the left-hand side of this equation may be expressed using Stokes' theorem as the *circulation* $\mathit{\Gamma}$:

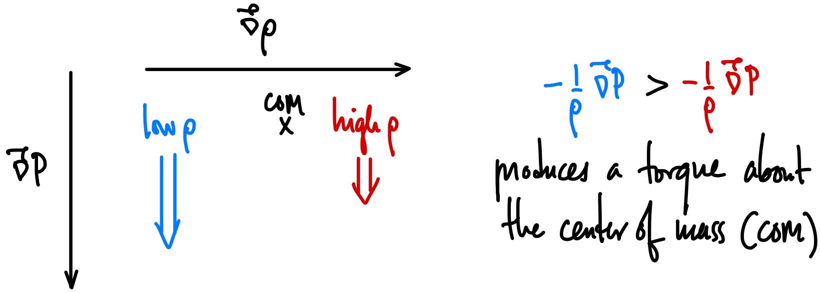
$$\int_{\mathcal{S}} \boldsymbol{\omega} \boldsymbol{\cdot} \mathrm{d}\boldsymbol{\mathcal{S}} = \oint_{\partial\mathcal{S}} \boldsymbol{u} \boldsymbol{\cdot} \mathrm{d}\boldsymbol{\ell} \doteq \mathit{\Gamma}. \tag{II.38}$$

The circulation around the boundary $\partial\mathcal{S}$ can be thought of as the number of vortex lines that thread the enclosed area S. Equation (II.37) then states that the circulation is conserved if the fluid is barotropic – Kelvin's circulation theorem:[6]

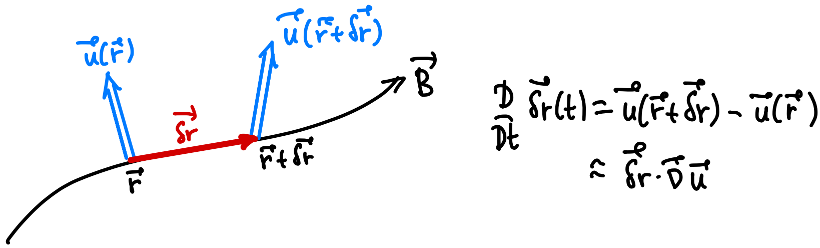
$$\boxed{\frac{\mathrm{D}\mathit{\Gamma}}{\mathrm{D}t} = -\oint_{\partial\mathcal{S}} \frac{\mathrm{d}P}{\rho} = 0 \quad \text{if } P = P(\rho)} \tag{II.39}$$

The figure below illustrates how baroclinic forcing generates vorticity.

[6] The above manipulations require that the surface is simply connected – that is, the region must be such that we can shrink the contour to a point without leaving the region. A region with a hole (like a bathtub drain) is not simply connected.



Another approach to proving (II.39) is to work with $\Gamma = \oint_{\partial S} \boldsymbol{u} \cdot \mathrm{d}\boldsymbol{\ell}$ rather than $\int_S \boldsymbol{\omega} \cdot \mathrm{d}\boldsymbol{S}$ and use the following for how an advected line element of ∂S changes in time:



Exercise. The helicity of a region of fluid is defined to be $\mathcal{H} \doteq \int \boldsymbol{\omega} \cdot \boldsymbol{u} \, \mathrm{d}\mathcal{V}$, where the integral is taken over the volume of that region. Assume that $\Gamma = \text{const}$ and that $\boldsymbol{\omega} \cdot \hat{\boldsymbol{n}}$ vanishes when integrated over the surface bounding $\mathcal{V}$, where $\hat{\boldsymbol{n}}$ is the unit normal to that surface. Prove that the helicity $\mathcal{H}$ is conserved in a frame moving with the fluid, *viz.* $\mathrm{D}\mathcal{H}/\mathrm{D}t = 0$. Note that the fluid need not be incompressible for this property to hold.

The calculation leading to (II.39) can be repeated in a reference frame rotating at a constant angular velocity $\boldsymbol{\Omega}$, in which the fluid velocity is measured to be $\boldsymbol{v} = \boldsymbol{u} - \boldsymbol{\Omega} \times \boldsymbol{r}$ (here, $\boldsymbol{u}$ is the fluid velocity in the inertial frame; see §II.3). The associated vorticity in this rotating frame is

$$\boldsymbol{\omega}_{\mathrm{rot}} = \boldsymbol{\omega} - \boldsymbol{\nabla} \times (\boldsymbol{\Omega} \times \boldsymbol{r}) = \boldsymbol{\omega} - \boldsymbol{\Omega}(\boldsymbol{\nabla} \cdot \boldsymbol{r}) + (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla})\boldsymbol{r} = \boldsymbol{\omega} - 3\boldsymbol{\Omega} + \boldsymbol{\Omega} = \boldsymbol{\omega} - 2\boldsymbol{\Omega}, \quad \text{(II.40)}$$

where $\boldsymbol{\omega} = \boldsymbol{\nabla} \times \boldsymbol{u}$. The circulation in the rotating reference frame is then given by

$$\begin{aligned} \Gamma_{\mathrm{rot}} &= \int_S \boldsymbol{\omega}_{\mathrm{rot}} \cdot \mathrm{d}\boldsymbol{S} = \int_S (\boldsymbol{\omega} - 2\boldsymbol{\Omega}) \cdot \mathrm{d}\boldsymbol{S} \\ &= \oint_{\partial S} \boldsymbol{u} \cdot \mathrm{d}\boldsymbol{\ell} - \int_S 2\boldsymbol{\Omega} \cdot \mathrm{d}\boldsymbol{S} \\ &= \Gamma - \int_S 2\boldsymbol{\Omega} \cdot \mathrm{d}\boldsymbol{S}. \end{aligned} \quad \text{(II.41)}$$

Kelvin's circulation theorem in this rotating frame is therefore

$$\frac{\mathrm{D}\Gamma_{\mathrm{rot}}}{\mathrm{D}t} = -\oint_{\partial S} \frac{\mathrm{d}P}{\rho} - 2\Omega \frac{\mathrm{D}\mathcal{S}_{\mathrm{n}}}{\mathrm{D}t}, \quad \text{(II.42)}$$

where $\mathcal{S}_{\mathrm{n}}$ is component of the surface area oriented normally to $\boldsymbol{\Omega}$. In words, if the projected area of the vortex tube in the plane perpendicular to the rotation vector changes, then the circulation in the rotating frame must change to compensate. This is the origin of Rossby waves, something that will be discussed further in §II.5.2.

II.5. Rotating reference frames

The final calculation in the preceding section provides a natural segue into a discussion of fluid dynamics in rotating reference frames. To begin this discussion, let us first recall equation (II.34), in which the nonlinearity $\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u}$ was written out in cylindrical coordinates for a fluid velocity $\boldsymbol{u}$ consisting of a cylindrical rotation $R\Omega\hat{\boldsymbol{\varphi}}$ with angular velocity $\Omega=\Omega(R,z)$ and a residual velocity $\boldsymbol{v}\doteq\boldsymbol{u}-R\Omega\hat{\boldsymbol{\varphi}}$:

$$\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u}=\left[\left(\boldsymbol{v}\cdot\boldsymbol{\nabla}+\Omega\frac{\partial}{\partial\varphi}\right)v_i\right]\hat{\boldsymbol{e}}_i+\left[2\Omega\hat{\boldsymbol{z}}\times\boldsymbol{v}-R\Omega^2\hat{\boldsymbol{R}}+R\hat{\boldsymbol{\varphi}}(\boldsymbol{v}\cdot\boldsymbol{\nabla})\Omega\right]$$
$$+\left[\frac{v_R v_\varphi}{R}\hat{\boldsymbol{\varphi}}-\frac{v_\varphi^2}{R}\hat{\boldsymbol{R}}\right].$$

When this expansion was introduced in §II.3, each of its components were described physically: 'The first term in brackets represents advection by the flow and the rotation. The second term in brackets contains the Coriolis force, the centrifugal force, and "tidal" terms due to the differential rotation... The third and final term in brackets captures curvature effects due to the cylindrical geometry.' Let's see these terms in action.

Using (II.34), we may express the equations of hydrodynamics (II.27) in cylindrical coordinates in a frame co-moving with the differential rotation. With

$$\frac{\mathrm{D}}{\mathrm{D}t}\to\frac{\partial}{\partial t}+\boldsymbol{v}\cdot\boldsymbol{\nabla}+\Omega\frac{\partial}{\partial\varphi} \tag{II.43}$$

to include advection by the rotation, we have

$$\frac{\mathrm{D}\rho}{\mathrm{D}t}=-\rho\boldsymbol{\nabla}\cdot\boldsymbol{v}, \tag{II.44a}$$
$$\frac{\mathrm{D}v_R}{\mathrm{D}t}=f_R+2\Omega v_\varphi+R\Omega^2+\frac{v_\varphi^2}{R}, \tag{II.44b}$$
$$\frac{\mathrm{D}v_\varphi}{\mathrm{D}t}=f_\varphi-\frac{\kappa^2}{2\Omega}v_R-R\frac{\partial\Omega}{\partial Z}v_z-\frac{v_R v_\varphi}{R}, \tag{II.44c}$$
$$\frac{\mathrm{D}v_z}{\mathrm{D}t}=f_z, \tag{II.44d}$$
$$\frac{\mathrm{D}s}{\mathrm{D}t}=0, \tag{II.44e}$$

where

$$\boldsymbol{f}=-\frac{1}{\rho}\boldsymbol{\nabla}P-\boldsymbol{\nabla}\Phi \tag{II.45}$$

and the combination

$$\boxed{\kappa^2\doteq4\Omega^2+\frac{\partial\Omega^2}{\partial\ln R}=\frac{1}{R^3}\frac{\partial(R^4\Omega^2)}{\partial R}} \tag{II.46}$$

is known as the (square of the) epicyclic frequency. Note that $R^4\Omega^2=\ell^2$, the square of the specific angular momentum ℓ, and so κ^2 measures how much the specific angular momentum associated with the rotation increases or decreases outwards. For Keplerian rotation, $\kappa^2=\Omega^2$.

In §VI.9, these equations will be modified for the presence and evolution of magnetic fields and used to look at linear waves and instabilities that rely on differential rotation. In the meantime, I'll close this portion of the notes by remarking on two useful applications of what you've learned here: the thermal wind equation (§II.5.1) and Rossby waves (§II.5.2).

II.5.1. *Thermal wind equation*

In steady state with $\boldsymbol{v} = 0$, equations (II.44*b*) and (II.44*d*) become

$$0 = -\frac{1}{\rho}\frac{\partial P}{\partial R} - \frac{\partial \varPhi}{\partial R} + R\varOmega^2 \quad \text{and} \quad 0 = -\frac{1}{\rho}\frac{\partial P}{\partial z} - \frac{\partial \varPhi}{\partial z}. \tag{II.47}$$

Taking $\partial/\partial z$ of the first equation, using the second equation, and rearranging yields

$$R\frac{\partial \varOmega^2}{\partial z} = \frac{\hat{\boldsymbol{\varphi}}}{\rho^2} \boldsymbol{\cdot} \bigl(\boldsymbol{\nabla} P \times \boldsymbol{\nabla}\rho\bigr). \tag{II.48}$$

This is the $\hat{\boldsymbol{\varphi}}$ component of the vorticity equation. Note that, if ρ is constant or if $P = P(\rho)$, then the angular velocity $\varOmega$ must be constant on cylinders (this is related to *von Zeippel's theorem*). Now, let us recall the definition of the hydrodynamic entropy, $s = (\gamma - 1)^{-1} \ln P\rho^{-\gamma}$ and use it to replace $\boldsymbol{\nabla} \ln \rho$. The result is

$$R\frac{\partial \varOmega^2}{\partial z} = \frac{\gamma - 1}{\gamma}\, \hat{\boldsymbol{\varphi}} \boldsymbol{\cdot} \left(\boldsymbol{\nabla} s \times \frac{1}{\rho}\boldsymbol{\nabla} P\right) = \hat{\boldsymbol{\varphi}} \boldsymbol{\cdot} \left(\frac{1}{\rho}\boldsymbol{\nabla} P \times \boldsymbol{\nabla} \ln T\right). \tag{II.49}$$

In the Sun, $\boldsymbol{g} = (1/\rho)\boldsymbol{\nabla} P$ is an excellent approximation, with only a tiny angular component due to centrifugal effects. Adopting this simplification and working in spherical coordinates (r, θ, φ), equation (II.49) becomes

$$\boxed{R\frac{\partial \varOmega^2}{\partial z} = \frac{\gamma - 1}{\gamma}\frac{g}{r}\frac{\partial s}{\partial \theta}} \tag{II.50}$$

where $g = GM/r^2$. [The right-hand side of (II.50) can also be written as $-(g/r)\partial \ln T/\partial\theta$.] Equation (II.50) is known as the *thermal wind equation*. It is used often in geophysical applications (e.g., longitudinal entropy gradients driven by temperature differences cause wind shear) and to understand the rotation profile in the convection zone of the Sun.

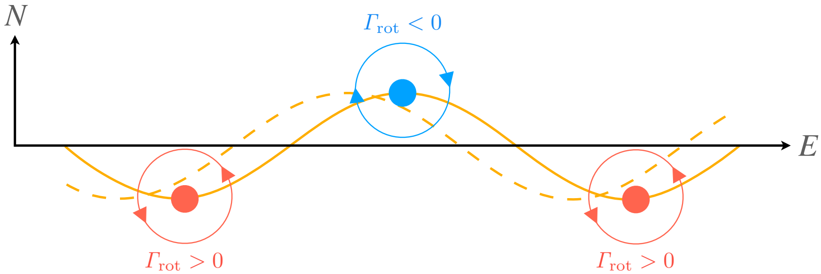
II.5.2. *Rossby waves*

Consider a two-dimensional, incompressible fluid on the surface of uniformly rotating sphere (e.g., a planetary atmosphere). For a constant density or a barotropic equation of state, equation (II.42) becomes

$$\frac{\mathrm{D}}{\mathrm{D}t}\bigl(\varGamma_{\mathrm{rot}} + 2\varOmega\mathcal{S}\cos\theta\bigr) = 0, \tag{II.51}$$

where θ is the angle between the rotation vector and the surface oriented normal to the fluid element. (Note that incompressibility assures $\mathcal{S} = \text{const.}$) This equation states that, as a fluid element makes its way from the equator northwards (*viz.*, from $\theta = \pi/2$ towards $\theta = 0$), its circulation as measured in the rotating frame must decrease. This means that the element must then rotate in the clockwise direction. Likewise, a fluid element that starts at the north pole and moves southwards towards the equator (*viz.*, from $\theta = 0$ towards $\theta = \pi/2$) increases its relative vorticity and thus rotates in the counterclockwise direction.

With this behavior in mind, let's now imagine a small-amplitude, wave-like disturbance at constant latitude (see diagram below). Northward displacements in this wave acquire negative relative vorticity and rotate clockwise; southward displacements acquire positive relative vorticity and rotate counterclockwise. These changes in the velocity of the disturbance actually feed back on the wave itself to make it travel westward; in effect, the wave is advecting itself to the west.



The relationship between the frequency ω and wavevector $\boldsymbol{k}$ for this wave – the *dispersion relation* – is given by

$$\omega = -\frac{k_y}{k_x^2 + k_y^2} \frac{2\Omega \sin\theta}{r}, \tag{II.52}$$

where x denotes the local poloidal direction (pointing southward), y denotes the local azimuthal direction (pointing eastward), and r the spherical radial distance. With $\Omega > 0$ and $k_y > 0$, the phase velocity of the wave $\omega/k_y < 0$, i.e., the wave travels westward. Note that the group velocity, $\partial\omega/\partial k_y$, can be either positive or negative; in general, shorter wavelengths (higher k) have an eastward group velocity and longer wavelengths (smaller k) have a westward group velocity.

These waves are named after the meteorologist Carl Rossby, who derived the mathematics governing this phenomenon in 1939 while at MIT (after which he became assistant director of research at the U.S. Weather Bureau and then moved to University of Chicago as Chair of the Department of Meteorology).[7]

PART III
Fundamentals of plasmas

Covered by Dr. Zhou, but here are some supplementary notes…

Now that we have the fluid equations under our belts, let us discuss why we might expect them to apply to a plasma (instead of the more familiar fluid). There are three concepts to cover in this regard: Debye shielding and quasi-neutrality, plasma oscillations, and collisional relaxation of the plasma to take on a Maxwell–Boltzmann distribution of particle velocities.

III.1. Debye shielding and quasi-neutrality

In § I.1, we mentioned the concept of the *Debye length* and explained its importance in the definition of a plasma. Here we actually derive it from first principles. This derivation starts by recalling that a large plasma parameter $\Lambda \gg 1$ implies that the kinetic energy of the plasma particles is much greater than the potential energy due to Coulomb interactions amongst binary pairs of particles. In this case, the plasma temperature T is much bigger than the Coulomb energy $e\phi \sim e^2/\Delta r \sim e^2 n^{1/3}$, where ϕ is the electrostatic potential, $\Delta r \sim n^{-1/3}$ is the typical interparticle distance, and n is the number density of the particles. Assuming a plasma in local thermodynamic equilibrium, the number density of species α' with charge $q_{\alpha'}$ sitting in the potential ϕ_α of one 'central' particle

[7]See https://images.peabody.yale.edu/publications/jmr/jmr02-01-06.pdf.

of species α ought to satisfy the Boltzmann relation

$$n_{\alpha'}(\boldsymbol{r}) = \overline{n}_{\alpha'} \exp\left(-\frac{q_\alpha \phi_\alpha(\boldsymbol{r})}{T}\right) \approx \overline{n}_{\alpha'}\left(1 - \frac{q_{\alpha'}\phi_\alpha(\boldsymbol{r})}{T}\right), \tag{III.1}$$

where the potential $\phi_\alpha(\boldsymbol{r})$ depends on the distance $\boldsymbol{r}$ from the 'central' particle. To obtain the approximate equality, we have used the assumption $T \gg e\phi_\alpha$ to Taylor expand the Boltzmann factor in its small argument. Inserting (III.1) into the Gauss–Poisson law for the electric field $\boldsymbol{E} = -\boldsymbol{\nabla}\phi_\alpha$, we have

$$\begin{aligned} \boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{E} = -\nabla^2\phi_\alpha &= 4\pi q_\alpha \delta(\boldsymbol{r}) + 4\pi \sum_{\alpha'} q_{\alpha'} n_{\alpha'} \\ &\approx 4\pi q_\alpha \delta(\boldsymbol{r}) + 4\pi \sum_{\alpha'} q_{\alpha'}\overline{n}_{\alpha'} - \underbrace{\left(\sum_{\alpha'} \frac{4\pi \overline{n}_{\alpha'} q_{\alpha'}^2}{T}\right)}_{\doteq \lambda_{\mathrm{D}}^{-2}} \phi_\alpha. \end{aligned} \tag{III.2}$$

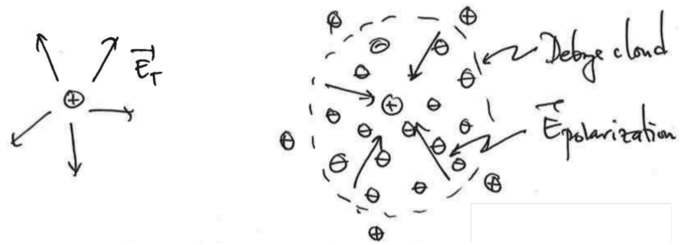
The first term in (III.2) is the point-like charge of the 'central' particle located at $\boldsymbol{r} = \boldsymbol{0}$. The second term is the sum over all charges in the plasma, and equals zero if the plasma is overall charge-neutral (as it should be). The final term introduces the Debye length (see (I.2)), which is the only characteristic scale in (III.2). Note further that this equation has no preferred direction, and so we may exploit its spherical symmetry to recast it as follows:

$$\frac{1}{r^2}\frac{\partial}{\partial r} r^2 \frac{\partial \phi_\alpha}{\partial r} - \frac{1}{\lambda_{\mathrm{D}}^2}\phi_\alpha = 4\pi q_\alpha \delta(\boldsymbol{r}). \tag{III.3}$$

The solution to this equation that asymptotes to the Coulomb potential $\phi_\alpha \to q_\alpha/r$ as $r \to 0$ and to zero as $r \to \infty$ is

$$\boxed{\phi_\alpha = \frac{q_\alpha}{r} \exp\left(-\frac{r}{\lambda_{\mathrm{D}}}\right)} \tag{III.4}$$

This equation states that the bare potential of the 'central' charge is exponentially attenuated ('shielded') on typical distances $\sim\lambda_{\mathrm{D}}$. This is *Debye shielding*, and the sphere of neutralizing charge accompanying the 'central' charge is referred to as the *Debye sphere* (or cloud). Debye shielding of an ion by preferential accumulation of electrons in its vicinity is sketched below:



Note that the electric field due to the polarization of the plasma in response to the ion's bare Coulomb potential acts in the opposite direction to the unshielded electric field.

Now, there was nothing particularly special about the charge that we singled out as our 'central' charge. Indeed, we could have performed the above integration for any charge in the plasma. This leads us to the fundamental tenet in the statistical mechanics of a weakly coupled plasma with $\Lambda \gg 1$: every charge simultaneously hosts its own Debye sphere while being a member of another charge's Debye sphere. The key points are that, by involving a huge number of particles in the small-scale electrostatics of the plasma, these Coulomb-mediated relations (i) make the plasma 'quasi-neutral' on scales $\gg\lambda_{\mathrm{D}}$ and (ii) make collective effects in the plasma much more important than individual binary effects due to particle-particle pairings. The latter is what makes a plasma very different from a neutral gas, in which particle-particle interactions occur through hard-body collisions on scales comparable to the mean particle size.

One consequence of Debye shielding is that the electric fields that act on large scales due to the self-consistent collective interactions between $\sim\Lambda$ Debye clouds are smoothly varying in space and time. As a result, when we write down Maxwell's equations for our quasi-neutral plasma, the fields that appear are these smooth, coarse-grained fields whose spatial structure resides far above the Debye length. Mathematically, we average the Maxwell equations over the microscopic (i.e., Debye) scales, and what remains are the collective macroscopic fields that ultimately make their way into the magnetohydrodynamics of the plasma 'fluid'.

III.2. Plasma oscillations

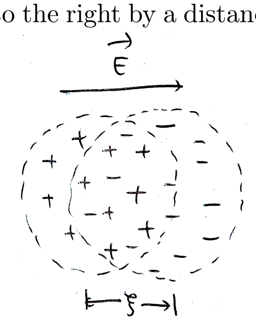
In the previous section, we spoke of a characteristic length scale below which particle-particle interactions are important and above which they are supplanted by collective effects between a large number of quasi-neutral Debye spheres. Is there a corresponding characteristic time scale? The answer is yes, and it may be obtained simply by dimensional analysis: take our Debye length and divide by a velocity to get time. The only velocity in our plasma thus far is the thermal speed, $v_{\mathrm{th}\alpha} = \sqrt{2T/m_\alpha}$, and so that must be it... we have obtained the *plasma frequency* of species α,

$$\omega_{\mathrm{p}\alpha} \doteq \sqrt{\frac{4\pi q_\alpha^2 n_\alpha}{m_\alpha}} \sim \frac{\lambda_{\mathrm{D}}}{v_{\mathrm{th}\alpha}}. \tag{III.5}$$

Of particular importance, given the smallness of the electron mass, is the electron plasma frequency ω_{pe}, which is $\sim\sqrt{m_{\mathrm{i}}/m_{\mathrm{e}}}$ larger than the ion plasma frequency and is generally the largest frequency in a weakly coupled plasma.

Fine. Dimensional analysis works. But what does this frequency actually mean? Go back to our picture of Debye shielding. That was a static picture, in that we waited long enough for the plasma to settle down into charge distributions governed by Boltzmann relations. What if we didn't wait? Surely there was some transient process whereby the particles moved around to configure themselves into these nice equilibrated Debye clouds. There was, and this transient process is referred to as a *plasma oscillation*, and it has a characteristic frequency of (you guessed it) ω_{pe}. Let's show this.

Imagine a spatially uniform, quasi-neutral plasma with well-equilibrated Debye clouds. Shift all of the electrons slightly to the right by a distance ξ, as shown in the figure below:



The offset between the electrons and the ions will cause an electric field pointing from the ions to the displaced electrons, given by $E = 4\pi e n_{\mathrm{e}}\xi$. The equation of motion for the electrons is then

$$m_{\mathrm{e}}\frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} = -eE = -4\pi e^2 n_{\mathrm{e}}\xi = -m_{\mathrm{e}}\omega_{\mathrm{pe}}^2\xi \quad \Longrightarrow \quad \frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} = -\omega_{\mathrm{pe}}^2\xi. \tag{III.6}$$

This is just the equation for a simple harmonic oscillator with frequency ω_{pe}. So, small displacements between oppositely charged species result in *plasma oscillations* (or 'Langmuir oscillations'), a collective process that occurs as the plasma attempts to restore quasi-neutrality in response to some disturbance. Retaining the effects of electron pressure makes these oscillations propagate dispersively with a non-zero group velocity; these *Langmuir waves* have the dispersion relation $\omega^2 \approx \omega_{\mathrm{pe}}^2(1+3k^2\lambda_{\mathrm{De}}^2)$, where k is the wavenumber of the perturbation. More on that later.

III.3. Collisional relaxation and the Maxwell–Boltzmann distribution

In order for the plasma particles to move freely as plasma oscillations attempt to set up equilibrated Debye clouds, the mean free path between particle–particle collisions must be larger than the Debye length. We may estimate the former in term of the collision cross-section σ,

$$\lambda_{\mathrm{mfp}} \sim \frac{1}{n\sigma} \sim \frac{T^2}{ne^4},$$

where the cross-section $\sigma = \pi b^2$ is given by a balance between the Coulomb potential energy, $\sim e^2/b$, across some typical impact parameter b and the kinetic energy of the particles, $\sim T$. Comparing this mean free path to the Debye length (I.2), we find

$$\frac{\lambda_{\mathrm{mfp}}}{\lambda_{\mathrm{D}}} \sim \frac{T^2}{ne^4}\left(\frac{ne^2}{T}\right)^{1/2} \sim n\lambda_{\mathrm{D}}^3 \doteq \Lambda \gg 1.$$

Thus, a particle can travel a long distance and experience the macroscopic fields exerted by the collective electrodynamics of the plasma before being deflected by much the shorter-range, microscopic electric fields generated by another individual particle (recall (I.9)).

The scale separation between the collisional mean free path and the Debye length due to the enormity of the plasma parameter in a weakly coupled plasma says something very important about the statistical mechanics of the plasma. Because $\lambda_{\mathrm{mfp}}/\lambda_{\mathrm{D}} \sim \omega_{\mathrm{pe}}\tau_{\mathrm{ei}} \gg 1$, the particle motions are randomized and the velocity distribution of the plasma particles relaxes to a local Maxwell–Boltzmann distribution on (collisional) timescales that are much longer than the timescale on which particle correlations are established and Coulomb potentials are shielded. As a result, collisions in the plasma occur between partially equilibrated Debye clouds instead of between individual particles, the mathematical result being that the ratio $\lambda_{\mathrm{mfp}}/\lambda_{\mathrm{D}}$ is attenuated by a factor $\sim \ln\Lambda \approx$ 10–40. Thus, the logarithmic factors in the collision times (I.7) and (I.8).

Now, about this collisional relaxation. This school isn't the place to go through all the details of how collision operators are derived, but we need to establish a few facts. First, because of Debye shielding, the vast majority of scatterings that a particle experiences as it moves through a plasma are *small-angle scatterings*, with each event changing the trajectory of a particle by a small amount. These accumulate like a random walk in angle away from the original trajectory of the particle, with an average deflection angle $\langle\theta\rangle = 0$ but with a mean-square deflection angle $\langle\theta^2\rangle$ proportional to the number of scattering events. For a typical electron scattering off a sea of Debye-shielded ions of charge Ze and density n, this angle satisfies

$$\langle\theta^2\rangle \approx \frac{8\pi n L Z^2 e^4}{m_{\rm e}^2 v_{\rm the}^4}\ln\Lambda \tag{III.7}$$

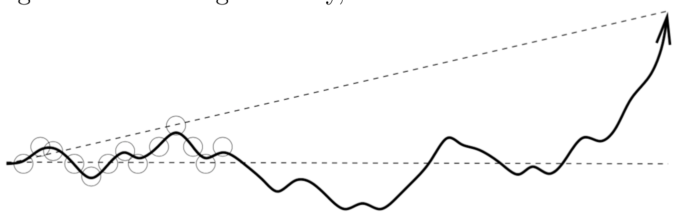
after the electron has traversed a distance L. A large deflection angle, i.e. $\langle\theta^2\rangle \sim 1$, is reached once this distance

$$L \sim \frac{m_{\rm e}^2 v_{\rm the}^4}{8\pi n Z^2 e^4}\frac{1}{\ln\Lambda} \sim v_{\rm the}\tau_{\rm ei} \doteq \lambda_{\rm mfp,e}, \tag{III.8}$$

the collisional mean free path (recall the definition of the electron–ion collision time, equation (I.7)). Noting that the impact parameter for a single 90-degree scattering is $\sim Ze^2/T$, we find the ratio of the cross-section for many small-angle scatterings to accumulate a 90-degree deflection, $\sigma_{\rm multi,90^\circ} \sim 1/nL$ using (III.8), to the cross-section for a single 90-degree scattering, $\sigma_{\rm single,90^\circ} = \pi b^2$ with $b \sim Ze^2/T$, is

$$\frac{\sigma_{\rm multi,90^\circ}}{\sigma_{\rm single,90^\circ}} \sim \ln\Lambda \gg 1. \tag{III.9}$$

Thus, in a weakly coupled plasma, multiple small-angle scatterings are more important than a single large-scale scattering. Visually,



This is the physical origin of the $\ln\Lambda$ reduction in collision time mentioned in the prior paragraph.

So what do these collisions mean for treating our plasma as a fluid? If $\lambda_{\rm mfp}$ is much less than any other macroscopic scale of dynamical interest (i.e., scales on which hydrodynamics occurs), then the *velocity distribution function* $f(\boldsymbol{v})$ of the plasma – that is, the differential number of particles with velocities between $\boldsymbol{v}$ and $\boldsymbol{v}+{\rm d}\boldsymbol{v}$ – is well described by a Maxwell–Boltzmann distribution (often simply called a 'Maxwellian'):

$$f_{\rm M}(v) \doteq \frac{n}{\pi^{3/2}v_{\rm th}^3}\exp\left(-\frac{v^2}{v_{\rm th}^2}\right). \tag{III.10}$$

The factor of $\pi^{3/2}v_{\rm th}^3$ is there for normalization purposes:

$$\int {\rm d}^3\boldsymbol{v}\, f_{\rm M}(\boldsymbol{v}) = 4\pi\int {\rm d}v\, v^2 f_{\rm M}(v) = n \tag{III.11}$$

is the number of particles per unit volume. (Any particle distribution function should satisfy this constraint.) Note that the Maxwellian is isotropic in velocity space, depending only on the speed of the particles (rather than their vector velocity). If these particles are all co-moving with some bulk velocity $\boldsymbol{u}$, then this 'fluid' velocity is subtracted off to ensure an isotropic distribution function in that 'fluid' frame:

$$f_{\rm M}(\boldsymbol{v}) \doteq \frac{n}{\pi^{3/2}v_{\rm th}^3}\exp\left(-\frac{|\boldsymbol{v}-\boldsymbol{u}|^2}{v_{\rm th}^2}\right). \tag{III.12}$$

Note that the first moment of this distribution

$$\int {\rm d}^3\boldsymbol{v}\,\boldsymbol{v} f_{\rm M}(\boldsymbol{v}) = n\boldsymbol{u}; \tag{III.13}$$

and that the (mass-weighted) second moment of this distribution

$$\int \mathrm{d}^3\boldsymbol{v}\, m|\boldsymbol{v}-\boldsymbol{u}|^2 f_{\mathrm{M}}(\boldsymbol{v}) = 3P. \tag{III.14}$$

(Again, any velocity distribution function should satisfy these constraints.)

Different species collisionally relax to a Maxwellian at different rates (e.g., $\tau_{\mathrm{ee}} \sim \tau_{\mathrm{ei}} \sim \sqrt{m_{\mathrm{i}}/m_{\mathrm{e}}}\,\tau_{\mathrm{ii}} \sim (m_{\mathrm{i}}/m_{\mathrm{e}})\tau_{ie}$), and so each species may be described by their own Maxwellians:

$$f_{\mathrm{M},\alpha}(\boldsymbol{v}) \doteq \frac{n_\alpha}{\pi^{3/2} v_{\mathrm{th}\alpha}^3} \exp\left(-\frac{|\boldsymbol{v}-\boldsymbol{u}_\alpha|^2}{v_{\mathrm{th}\alpha}^2}\right). \tag{III.15}$$

But, in the long-time limit, unless some process actively dis-equilibrates the species on a timescale comparable to or smaller than these collision times, all species will take on the *same* $\boldsymbol{u}$ and the *same* T. Their densities are, of course, the same as well, as guaranteed by quasi-neutrality (*viz.*, $\omega_{\mathrm{pe}}\tau \gg 1$ for all collision times τ).

Note then, that when we wrote down our hydrodynamic equations for a scalar pressure (see (II.14) and (II.17)) and didn't affix any species labels to any quantities, we were implicitly assuming that our hydrodynamics occurs on time scales much longer than the collisional equilibration times, so that all species can be well described by local Maxwellians with the same density, fluid velocity, and temperature. Not all astrophysical systems are so cooperative, and anisotropic pressures, velocity drifts between species, and dis-equilibration of species temperatures can often be the norm. Yes, hydrodynamics and MHD are fairly simple, but do not let their simplicity lure you into using them when it's not appropriate to do so – a hard-earned lesson for many astrophysicists.

PART IV
Fundamentals of magnetohydrodynamics

IV.1. The equations of ideal magnetohydrodynamics

Ideal magnetohydrodynamics (MHD) describes the hydrodynamics of a perfectly conducting fluid in the presence of electromagnetic fields. Mass is still conserved, so we still have the continuity equation:

$$\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\boldsymbol{\cdot}(\rho\boldsymbol{u}) = 0. \tag{IV.1}$$

The first law of thermodynamics still holds, so we still have the internal energy equation:

$$\frac{\partial e}{\partial t} + \boldsymbol{\nabla}\boldsymbol{\cdot}(e\boldsymbol{u}) = -P\boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{u}. \tag{IV.2}$$

And Newton's second law still governs the dynamics, so we still have the momentum equation:

$$\frac{\partial(\rho\boldsymbol{u})}{\partial t} + \boldsymbol{\nabla}\boldsymbol{\cdot}(\rho\boldsymbol{u}\boldsymbol{u}) = \boldsymbol{f}. \tag{IV.3}$$

But now we must supplement the force $\boldsymbol{f}$, which was equal to $-\boldsymbol{\nabla}P - \rho\boldsymbol{\nabla}\Phi$ in §II, with the force due to the electromagnetic fields on the conducting fluid elements. To do so, let us view our conducting fluid elements as a coherent collection of ions (with charge $q_{\mathrm{i}} = Ze > 0$) and electrons (with charge $q_{\mathrm{e}} = -e < 0$), and ask how electric and magnetic fields influence each of these species.

The electromagnetic force per unit volume on a collection of charges of species α is given by

$$\boldsymbol{f}_{\mathrm{EM}} = q_\alpha n_\alpha \left(\boldsymbol{E} + \frac{\boldsymbol{u}_\alpha}{c} \times \boldsymbol{B} \right), \tag{IV.4}$$

where n_α is the number density of the species and $\boldsymbol{u}_\alpha$ is that species' bulk velocity. You can think of this simply as the Lorentz force $q_\alpha(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}/c)$ integrated over the ensemble of α charges in each fluid element and divided by the volume of said fluid element. Separating (IV.3) into its charged constituent parts, we then have the momentum equation for species α,

$$\frac{\partial(\rho_\alpha \boldsymbol{u}_\alpha)}{\partial t} + \boldsymbol{\nabla \cdot} (\rho_\alpha \boldsymbol{u}_\alpha \boldsymbol{u}_\alpha) = -\boldsymbol{\nabla} P_\alpha - \rho_\alpha \boldsymbol{\nabla} \varPhi + q_\alpha n_\alpha \left(\boldsymbol{E} + \frac{\boldsymbol{u}_\alpha}{c} \times \boldsymbol{B} \right). \tag{IV.5}$$

At the moment, the trouble is that our continuity equation (IV.1) and internal energy equation (IV.2) make reference to the *total* mass density ρ, the *total* fluid velocity $\boldsymbol{u}$, the *total* pressure P, and the *total* internal energy e. The obvious thing to do, then, is to sum (IV.5) over both species,

$$\sum_\alpha \left[\frac{\partial(\rho_\alpha \boldsymbol{u}_\alpha)}{\partial t} + \boldsymbol{\nabla \cdot} (\rho_\alpha \boldsymbol{u}_\alpha \boldsymbol{u}_\alpha) = -\boldsymbol{\nabla} P_\alpha - \rho_\alpha \boldsymbol{\nabla} \varPhi + q_\alpha n_\alpha \left(\boldsymbol{E} + \frac{\boldsymbol{u}_\alpha}{c} \times \boldsymbol{B} \right) \right], \tag{IV.6}$$

and simplify each of the sums one by one. The first term in (IV.6) becomes familiar after introducing the center-of-mass fluid velocity,

$$\boldsymbol{u} \doteq \frac{1}{\rho} \sum_\alpha \rho_\alpha \boldsymbol{u}_\alpha, \quad \text{where} \quad \rho \doteq \sum_\alpha \rho_\alpha . \tag{IV.7}$$

The second term in (IV.6) requires a bit more work. Write $\boldsymbol{u}_\alpha = \boldsymbol{u} + \Delta \boldsymbol{u}_\alpha$, so that $\Delta \boldsymbol{u}_\alpha$ measures the difference between the bulk flow of species α and the center-of-mass velocity $\boldsymbol{u}$. Then

$$\sum_\alpha \rho_\alpha \boldsymbol{u}_\alpha \boldsymbol{u}_\alpha = \rho \boldsymbol{u}\boldsymbol{u} + \boldsymbol{u} \Big(\cancelto{0}{\sum_\alpha \rho_\alpha \Delta \boldsymbol{u}_\alpha} \Big) + \Big(\cancelto{0}{\sum_\alpha \rho_\alpha \Delta \boldsymbol{u}_\alpha} \Big) \boldsymbol{u} + \sum_\alpha \rho_\alpha \Delta \boldsymbol{u}_\alpha \Delta \boldsymbol{u}_\alpha .$$

The first term here ($\rho \boldsymbol{u}\boldsymbol{u}$) should look familiar: it's the flux of momentum density associated with the total fluid, the same as was seen in §II. Moving the final term of the above expression to the right-hand side of (IV.6) and writing $\sum_\alpha P_\alpha \doteq P$, we have a momentum equation that is starting to look more like (IV.3):

$$\frac{\partial(\rho \boldsymbol{u})}{\partial t} + \boldsymbol{\nabla \cdot} (\rho \boldsymbol{u}\boldsymbol{u}) = -\boldsymbol{\nabla} P - \rho \boldsymbol{\nabla} \varPhi - \sum_\alpha \rho_\alpha \Delta \boldsymbol{u}_\alpha \Delta \boldsymbol{u}_\alpha + \sum_\alpha q_\alpha n_\alpha \left(\boldsymbol{E} + \frac{\boldsymbol{u}_\alpha}{c} \times \boldsymbol{B} \right). \tag{IV.8}$$

Now, this term involving $\Delta \boldsymbol{u}_\alpha$ has nothing really to do with MHD, and was in fact implicitly discarded in §II.1.2, the reason being either that our fluid element is composed of a single species, or that collisions between different species keep their bulk flows very close to the center-of-mass velocity, or that the total mass density and total momentum density are completely dominated by a single species (e.g., the ions). In any of these cases, we may safely drop this term.

Almost there. All that remains to consider is

$$\sum_\alpha q_\alpha n_\alpha \left(\boldsymbol{E} + \frac{\boldsymbol{u}_\alpha}{c} \times \boldsymbol{B} \right).$$

In §III.1, we showed that the densities of the positive and negative charge carriers surrounding a point charge Q in a weakly coupled plasma satisfies

$$\sum_\alpha q_\alpha n_\alpha = \cancelto{0}{\sum_\alpha q_\alpha \bar{n}_\alpha} - \frac{Q}{4\pi\lambda_{\rm D}^3}\,\frac{\exp(-r/\lambda_{\rm D})}{r/\lambda_{\rm D}},$$

and therefore is extremely close to zero well outside of that charge's Debye sphere, i.e., the plasma is *quasi-neutral* on scales $r \gg \lambda_{\rm D}$. MHD concerns itself with just such scales, and so the total electric force on a fluid element in MHD vanishes under quasi-neutrality. This leaves the magnetic term, $(\sum_\alpha q_\alpha n_\alpha \boldsymbol{u}_\alpha)\times\boldsymbol{B}/c$. The sum in parentheses is equivalent to the *current density* of the plasma, $\boldsymbol{j}$, the amount of electric current flowing per unit cross-sectional area. We these principles implemented, our MHD momentum equation is finally here:

$$\boxed{\frac{\partial(\rho\boldsymbol{u})}{\partial t} + \boldsymbol{\nabla}\boldsymbol{\cdot}(\rho\boldsymbol{u}\boldsymbol{u}) = -\boldsymbol{\nabla}P - \rho\boldsymbol{\nabla}\Phi + \frac{\boldsymbol{j}}{c}\times\boldsymbol{B}} \tag{IV.9}$$

Another way to this of this additional term is by analogy with circuits: when a current $\boldsymbol{I}$ flows through a wire of length ℓ in the presence of a magnetic field $\boldsymbol{B}$, there is a force on the wire given by $\boldsymbol{I}\ell\times\boldsymbol{B}/c$. In the fluid context, the 'wire' is the conducting fluid element through which electrons and ions move differentially.

We now have our continuity equation, internal energy equation, and MHD momentum equation. However, in deriving the latter, we have introduced two new variables, $\boldsymbol{j}$ and $\boldsymbol{B}$. The remaining tasks are then to express the current density $\boldsymbol{j}$ in terms of the magnetic field $\boldsymbol{B}$ (since by summing over the momentum equations of each species, we've lost information about each species' bulk flow), and to provide an equation for how the magnetic field evolves. Both of these tasks are solved by Maxwell's equations:

$$\frac{\partial\boldsymbol{B}}{\partial t} = -c\boldsymbol{\nabla}\times\boldsymbol{E}, \quad \boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{B} = 0, \quad \frac{\partial\boldsymbol{E}}{\partial t} = c\boldsymbol{\nabla}\times\boldsymbol{B} - 4\pi\boldsymbol{j}, \quad \boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{E} = 4\pi\sum_\alpha q_\alpha n_\alpha,$$

with the important caveat that the final equation in red (Gauss' law) is rendered completely useless by the quasi-neutrality assumption, $\sum_\alpha q_\alpha n_\alpha \approx 0$. The other equations are (from left to right) Faraday's law of induction, Gauss' law for magnetism (no magnetic monopoles), and Maxwell's version of Ampère's law. No offense to Maxwell, but it turns out that the original Ampère's law,

$$\boldsymbol{j} = \frac{c}{4\pi}\boldsymbol{\nabla}\times\boldsymbol{B}, \tag{IV.10}$$

is just fine our purposes. The displacement current, $(4\pi)^{-1}\partial\boldsymbol{E}/\partial t$, which mathematically and physically connects electromagnetism with the propagation of light, may be rigorously dropped if the fluid velocity satisfies $u^2 \ll c^2$. Why, you ask? Well, this brings us back to the first sentence of this section: we are interested in *perfect conductors*.

A perfect conductor is one that has exactly zero electrical resistance, and so by Ohm's law must have zero electrostatic field. But this doesn't necessarily mean that $\boldsymbol{E} = 0$, because an electric field can be induced by the motion of a conductor through a magnetic field (sometimes called the 'motional emf'). What we mean by a perfect conductor is then that the electric field vanishes *in the frame of the conductor*, or

$$\boldsymbol{E} + \frac{\boldsymbol{u}}{c}\times\boldsymbol{B} = 0. \tag{IV.11}$$

Inserting this equation into the Maxwell–Ampère law and ordering $\partial/\partial t \sim u/\ell$ for some characteristic bulk flow velocity u and gradient lengthscale ℓ, we find that the displacement current

$$\frac{\partial \boldsymbol{E}}{\partial t} \sim \frac{u^2}{c^2}\frac{cB}{\ell} \ll \frac{cB}{\ell} \sim c\boldsymbol{\nabla}\times\boldsymbol{B}$$

if the flow is non-relativistic. As claimed, the original Ampère's law is just fine.

Altogether then, we may close our MHD momentum equation with the following subset of Maxwell's equations:

$$\boxed{\frac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{\nabla}\times(\boldsymbol{u}\times\boldsymbol{B}), \quad \boldsymbol{\nabla}\cdot\boldsymbol{B} = 0, \quad \boldsymbol{j} = \frac{c}{4\pi}\boldsymbol{\nabla}\times\boldsymbol{B}} \tag{IV.12}$$

These equations for the electromagnetic fields $\boldsymbol{B}$ and $\boldsymbol{j}$ – taken alongside (IV.1), (IV.2), and (IV.8) specifying the evolution of the hydrodynamics variables $(\rho,\rho\boldsymbol{u},e)$ – constitute the equations of ideal MHD.

IV.1.1. *Flux freezing: Alfvén's theorem*

Arguably the most important prediction of the ideal MHD equations is that the magnetic flux $\varPhi_B$ through the surface of any fluid element is exactly conserved as that element is advected and deformed by a flow $\boldsymbol{u} = \boldsymbol{u}(t,\boldsymbol{r})$. This is known as 'Alfvén's theorem' or, more colloquially, *flux freezing*. Given Leibniz's rule regarding the time derivatives of surface integrals whose integrations limits $\mathcal{S}(t)$ are time-dependent (eq. (II.32)), the proof itself is trivial:

$$\begin{aligned}
\frac{\mathrm{D}\varPhi_B}{\mathrm{D}t} \doteq \frac{\mathrm{D}}{\mathrm{D}t}\int_{\mathcal{S}(t)} \mathrm{d}\boldsymbol{\mathcal{S}}\cdot\boldsymbol{B} &= \int_{\mathcal{S}(t)} \mathrm{d}\boldsymbol{\mathcal{S}}\cdot\left[\frac{\partial \boldsymbol{B}}{\partial t} + (\boldsymbol{\nabla}\cdot\boldsymbol{B})\boldsymbol{u}\right] - \oint_{\partial\mathcal{S}(t)} \mathrm{d}\boldsymbol{\ell}\cdot(\boldsymbol{u}\times\boldsymbol{B}) \\
\text{(use equation (IV.12))} &= \int_{\mathcal{S}(t)} \mathrm{d}\boldsymbol{\mathcal{S}}\cdot\Big[\boldsymbol{\nabla}\times(\boldsymbol{u}\times\boldsymbol{B})\Big] - \oint_{\partial\mathcal{S}(t)} \mathrm{d}\boldsymbol{\ell}\cdot(\boldsymbol{u}\times\boldsymbol{B}) \\
\text{(use Stokes' theorem)} &= \oint_{\partial\mathcal{S}(t)} \mathrm{d}\boldsymbol{\ell}\cdot(\boldsymbol{u}\times\boldsymbol{B}) - \oint_{\partial\mathcal{S}(t)} \mathrm{d}\boldsymbol{\ell}\cdot(\boldsymbol{u}\times\boldsymbol{B}) \\
&= 0.
\end{aligned} \tag{IV.13}$$

In words, the magnetic flux is conserved in a frame comoving with a fluid element. (This is analogous to Kelvin's circulation theorem governing the circulation; cf. (II.39).)

An alternative description of flux freezing can be stated in terms of line tying: fluid elements that lie on a field line initially will remain on that field line (Lundquist 1951). See Problem 9 in Problem Set 1.

IV.1.2. *Ideal MHD induction equation*

Using a particular vector identity (see §II.3.1), the ideal MHD induction equation may be written in the following form:

$$\frac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{\nabla}\times(\boldsymbol{u}\times\boldsymbol{B}) = -\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{B} + \boldsymbol{B}\cdot\boldsymbol{\nabla}\boldsymbol{u} - \boldsymbol{B}\boldsymbol{\nabla}\cdot\boldsymbol{u}. \tag{IV.14}$$

Each of the terms on the right-hand side has a physical meaning. The first indicates that the magnetic field is advected (carried around by) the fluid flow; when placed on the left-hand side, we obtain the Lagrangian derivative of the magnetic field, $\mathrm{D}\boldsymbol{B}/\mathrm{D}t$. In this Lagrangian frame, the magnetic field can evolve because of two effects. The second term on the right-hand side, $\boldsymbol{B}\cdot\boldsymbol{\nabla}\boldsymbol{u}$, represents stretching of the magnetic field: if the fluid velocity has a gradient along the direction of the magnetic field, different parts of the field line will be carried along at different velocities, causing the field line to stretch. The final term, $-\boldsymbol{B}\boldsymbol{\nabla}\cdot\boldsymbol{u}$, corresponds to compression or rarefaction of the magnetic field.

Indeed, with the continuity equation giving $-\boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{u} = \mathrm{D}\ln\rho/\mathrm{D}t$, we see that co-moving increases (decreases) in the fluid density go hand-in-hand with increases (decreases) in the magnetic-field strength.

A rarely publicized but useful form of the induction equation (IV.14) is obtained by defining the magnetic-field unit vector $\hat{\boldsymbol{b}} \doteq \boldsymbol{B}/B$ and writing separate equations for it and the magnetic-field strength B:

$$\frac{\mathrm{D}\ln B}{\mathrm{D}t} = \big(\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} - \mathsf{I}\big)\boldsymbol{:}\boldsymbol{\nabla}\boldsymbol{u} \quad \text{and} \quad \frac{\mathrm{D}\hat{\boldsymbol{b}}}{\mathrm{D}t} = \big(\mathsf{I} - \hat{\boldsymbol{b}}\hat{\boldsymbol{b}}\big)\boldsymbol{:}\big(\hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla}\boldsymbol{u}\big). \tag{IV.15}$$

These may come in handy one day...

IV.1.3. *Lorentz force: Magnetic pressure and tension*

We now know that perfectly conducting fluids advect, stretch, and compress magnetic fields while conserving magnetic flux. What is the effect of that flux on the dynamics of the fluid element itself? For that, we revisit the Lorentz force in the MHD momentum equation (IV.9), and use Ampère's law to cast the current density in terms of the magnetic field:

$$\boldsymbol{f}_{\rm M} = \frac{\boldsymbol{j}}{c}\times\boldsymbol{B} = \frac{(\boldsymbol{\nabla}\times\boldsymbol{B})\times\boldsymbol{B}}{4\pi} = -\boldsymbol{\nabla}\frac{B^2}{8\pi} + \frac{\boldsymbol{B}\boldsymbol{\cdot}\boldsymbol{\nabla}\boldsymbol{B}}{4\pi}, \tag{IV.16}$$

where to obtain the final equality we have used a well-known vector identity (see §II.3.1). Because $\boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{B} = 0$, this can also be written as

$$\boldsymbol{f}_{\rm M} = -\boldsymbol{\nabla}\boldsymbol{\cdot}\left[\frac{B^2}{8\pi}\mathsf{I} - \frac{\boldsymbol{B}\boldsymbol{B}}{4\pi}\right] = -\boldsymbol{\nabla}\boldsymbol{\cdot}\mathsf{M}, \tag{IV.17}$$

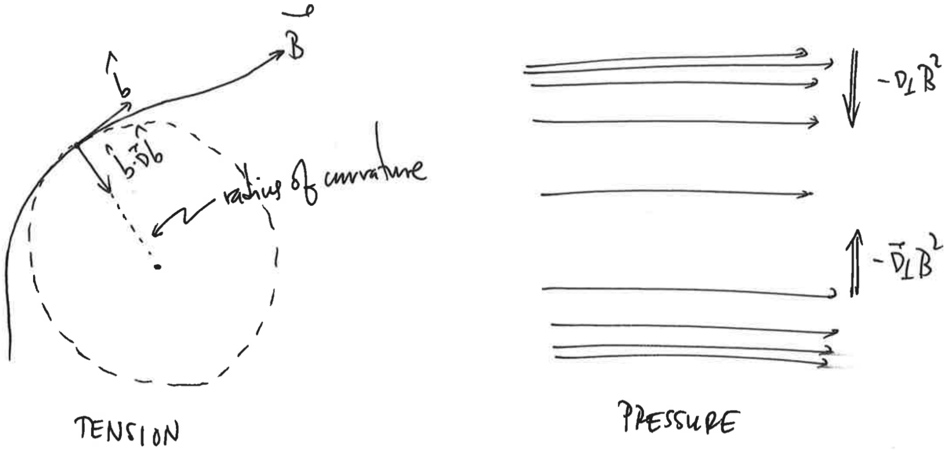
which implicitly defines the 'Maxwell stress', M. This form of the magnetic force suggests a kind of elasticity. To further see this, use the definition of the magnetic unit vector $\hat{\boldsymbol{b}} \doteq \boldsymbol{B}/B$ to write

$$\boldsymbol{B}\boldsymbol{\cdot}\boldsymbol{\nabla}\boldsymbol{B} = B\hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla}(B\hat{\boldsymbol{b}}) = B^2(\hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla}\hat{\boldsymbol{b}}) + \hat{\boldsymbol{b}}\hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla}\frac{B^2}{2}.$$

Using this in (IV.16) and collecting terms yields

$$\boldsymbol{f}_{\rm M} = \frac{B^2}{4\pi}(\hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla}\hat{\boldsymbol{b}}) - \big(\mathsf{I} - \hat{\boldsymbol{b}}\hat{\boldsymbol{b}}\big)\boldsymbol{\cdot}\boldsymbol{\nabla}\frac{B^2}{8\pi}. \tag{IV.18}$$

The first term here corresponds to a curvature force, with $\hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla}\hat{\boldsymbol{b}} \doteq \boldsymbol{\kappa}$ being the curvature of the field lines (see the diagram below). Note that $1/|\boldsymbol{\kappa}|$ is the radius of curvature. When a field line is bent, there is a force pointing towards the local center of curvature that is trying to un-bend the field line and push the plasma towards a lower-energy state in which the magnetic field is straight. The second term in (IV.18) corresponds to a magnetic pressure force acting perpendicular to the field (thus the projection of the gradient onto $\mathsf{I} - \hat{\boldsymbol{b}}\hat{\boldsymbol{b}}$). This term causes the magnetic-field strength to evolve towards being uniform across itself, again seeking a lower-energy state. Magnetic fields like to be straight and evenly spaced, and they will coerce the fluid to adopt motions that drive them towards being straight and evenly spaced.



IV.1.4. *MHD energy equation*

In §II.1.3, we derived an evolution equation for the total energy of a neutral fluid (eq. (II.20)). Here we augment that equation for a perfectly conducting fluid to include the energy of the magnetic field, $B^2/8\pi$. Take the ideal MHD induction equation (IV.14) and dot it with $\boldsymbol{B}/4\pi$:

$$\begin{aligned}
\frac{\partial}{\partial t}\frac{B^2}{8\pi} = \frac{\boldsymbol{B}}{4\pi}\cdot\boldsymbol{\nabla}\times(\boldsymbol{u}\times\boldsymbol{B}) &= \frac{B_i}{4\pi}\,\epsilon_{ijk}\frac{\partial}{\partial x_j}(\boldsymbol{u}\times\boldsymbol{B})_k \\
&= \epsilon_{ijk}\frac{\partial}{\partial x_j}\left[\frac{B_i}{4\pi}(\boldsymbol{u}\times\boldsymbol{B})_k\right] - \epsilon_{ijk}(\boldsymbol{u}\times\boldsymbol{B})_k\frac{\partial}{\partial x_j}\frac{B_i}{4\pi} \\
&= \epsilon_{ijk}\frac{\partial}{\partial x_j}\left[\frac{B_i}{4\pi}(\boldsymbol{u}\times\boldsymbol{B})_k\right] - \epsilon_{ijk}\epsilon_{k\ell m}u_\ell B_m\frac{\partial}{\partial x_j}\frac{B_i}{4\pi} \\
&= \epsilon_{ijk}\frac{\partial}{\partial x_j}\left[\frac{B_i}{4\pi}(\boldsymbol{u}\times\boldsymbol{B})_k\right] - (\delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{j\ell})u_\ell B_m\frac{\partial}{\partial x_j}\frac{B_i}{4\pi} \\
&= -\boldsymbol{\nabla}\cdot\left[\frac{\boldsymbol{B}\times(\boldsymbol{u}\times\boldsymbol{B})}{4\pi}\right] - \frac{\boldsymbol{u}\boldsymbol{B}\,\boldsymbol{:}\,\boldsymbol{\nabla}\boldsymbol{B}}{4\pi} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\frac{B^2}{8\pi}
\end{aligned}$$

$$\Longrightarrow\quad \frac{\partial}{\partial t}\frac{B^2}{8\pi} + \boldsymbol{\nabla}\cdot\left[\frac{\boldsymbol{B}\times(\boldsymbol{u}\times\boldsymbol{B})}{4\pi}\right] = -\frac{\boldsymbol{u}\boldsymbol{B}\,\boldsymbol{:}\,\boldsymbol{\nabla}\boldsymbol{B}}{4\pi} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\frac{B^2}{8\pi}.$$

Note that the quantity inside the divergence on the left-hand side of this equation equals $(c/4\pi)\boldsymbol{E}\times\boldsymbol{B} \doteq \boldsymbol{\mathcal{S}}$... the Poynting flux! In words, magnetic energy (as measured in the lab frame; note the partial time derivative) is transported by the Poynting flux. Those two terms on the right-hand side corresponding to will be cancelled by two equal-and-opposite terms found in the equation for the kinetic energy, obtained by dotting the momentum equation (IV.9) with $\boldsymbol{u}$ and focusing on the Lorentz force:

$$\boldsymbol{u}\cdot\left(-\boldsymbol{\nabla}\frac{B^2}{8\pi} + \frac{\boldsymbol{B}\cdot\boldsymbol{\nabla}\boldsymbol{B}}{4\pi}\right).$$

Yep, they cancel. So, adding the total hydrodynamic energy equation including these Lorentz-force contributions to the magnetic energy equation leads to

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\rho u^2 + e + \rho\varPhi + \frac{B^2}{8\pi}\right) + \boldsymbol{\nabla}\cdot\left[\left(\frac{1}{2}\rho u^2 + \gamma e + \rho\varPhi\right)\boldsymbol{u} + \boldsymbol{\mathcal{S}}\right] = \rho\frac{\partial\varPhi}{\partial t}.$$

But for the impact of a time-varying gravitational potential, the total MHD energy $\mathcal{E} \doteq (1/2)\rho u^2 + e + \rho\Phi + B^2/8\pi$ is conserved.

IV.1.5. *Rotating reference frames*

In §II.3.3, we examined the nonlinear combination $\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u}$ in curvilinear coordinates, finding additional terms that stemmed from differentiating unit vectors and which included Coriolis, centrifugal, and tidal accelerations. Here we take a similar look at the combination $\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{B} - \boldsymbol{B}\cdot\boldsymbol{\nabla}\boldsymbol{u}$ that features in the induction equation (IV.14).

First, use $\partial\hat{\boldsymbol{\varphi}}/\partial\varphi = -\hat{\boldsymbol{R}}$ and $\partial\hat{\boldsymbol{R}}/\partial\varphi = \hat{\boldsymbol{\varphi}}$ in (IV.14) to obtain

$$\frac{\partial\boldsymbol{B}}{\partial t} + \boldsymbol{B}\boldsymbol{\nabla}\cdot\boldsymbol{u} = (-\boldsymbol{u}\cdot\boldsymbol{\nabla}B_i)\hat{\boldsymbol{e}}_i + (\boldsymbol{B}\cdot\boldsymbol{\nabla}u_i)\hat{\boldsymbol{e}}_i + \frac{B_\varphi u_R - B_R u_\varphi}{R}\hat{\boldsymbol{\varphi}}.$$

As in §II.5, if we then decompose the fluid velocity as $\boldsymbol{u} = \boldsymbol{v} + R\Omega(R,z)\hat{\boldsymbol{\varphi}}$, where Ω is an angular velocity, substitute this decomposition into the above equation, and re-group terms, we have

$$\frac{\mathrm{D}B_R}{\mathrm{D}t} = \boldsymbol{B}\cdot\boldsymbol{\nabla}v_R - B_R\boldsymbol{\nabla}\cdot\boldsymbol{v}, \tag{IV.19a}$$

$$\frac{\mathrm{D}B_\varphi}{\mathrm{D}t} = \boldsymbol{B}\cdot\boldsymbol{\nabla}v_\varphi - B_\varphi\boldsymbol{\nabla}\cdot\boldsymbol{v} + B_R\frac{\partial\Omega}{\partial\ln R} + B_z R\frac{\partial\Omega}{\partial z} + \frac{B_\varphi v_R - B_R v_\varphi}{R}, \tag{IV.19b}$$

$$\frac{\mathrm{D}B_z}{\mathrm{D}t} = \boldsymbol{B}\cdot\boldsymbol{\nabla}v_z - B_z\boldsymbol{\nabla}\cdot\boldsymbol{v}, \tag{IV.19c}$$

with $\mathrm{D}/\mathrm{D}t \doteq \partial/\partial t + \boldsymbol{v}\cdot\boldsymbol{\nabla} + \Omega\,\partial/\partial\varphi$. Note that poloidal magnetic fields are sheared into the azimuthal direction by differential rotation.

IV.2. Summary: Adiabatic equations of ideal MHD

The adiabatic equations of MHD, written in conservative form, are:

$$\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot(\rho\boldsymbol{u}) = 0, \tag{IV.20a}$$

$$\frac{\partial(\rho\boldsymbol{u})}{\partial t} + \boldsymbol{\nabla}\cdot(\rho\boldsymbol{u}\boldsymbol{u}) = -\boldsymbol{\nabla}\cdot\left[\left(P + \frac{B^2}{8\pi}\right)\boldsymbol{I} - \frac{\boldsymbol{B}\boldsymbol{B}}{4\pi}\right] - \rho\boldsymbol{\nabla}\Phi, \tag{IV.20b}$$

$$\frac{\partial e}{\partial t} + \boldsymbol{\nabla}\cdot(e\boldsymbol{u}) = -P\boldsymbol{\nabla}\cdot\boldsymbol{u}, \tag{IV.20c}$$

$$\frac{\partial\boldsymbol{B}}{\partial t} - \boldsymbol{\nabla}\times(\boldsymbol{u}\times\boldsymbol{B}) = 0. \tag{IV.20d}$$

The left-hand sides of these equations express advection of, respectively, the mass density, the momentum density, the internal energy density, and the magnetic flux by the fluid velocity; the right-hand sides represents sources and sinks.

If we instead write these equations in terms of the density, fluid velocity, and entropy and make use of the Lagrangian derivative (II.7), we have

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} = -\rho\boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{u}, \tag{IV.21a}$$

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} = -\frac{1}{\rho}\boldsymbol{\nabla}\left(P + \frac{B^2}{8\pi}\right) + \frac{\boldsymbol{B}\boldsymbol{\cdot}\boldsymbol{\nabla}\boldsymbol{B}}{4\pi\rho} - \boldsymbol{\nabla}\varPhi, \tag{IV.21b}$$

$$\frac{\mathrm{D}s}{\mathrm{D}t} = 0, \tag{IV.21c}$$

$$\frac{\mathrm{D}\boldsymbol{B}}{\mathrm{D}t} = \boldsymbol{B}\boldsymbol{\cdot}\boldsymbol{\nabla}\boldsymbol{u} - \boldsymbol{B}\boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{u}, \tag{IV.21d}$$

where $s \doteq (\gamma-1)^{-1}\,\ln P\rho^{-\gamma}$.

PART V
Linear theory of MHD waves

Covered by Dr. Zhou, but here are some supplementary hand-written notes...

MHD waves and linear theory

Let us summarize our dissipationless "ideal MHD" eqns:

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

$$\rho\left(\frac{\partial}{\partial t} + \vec{u}\cdot\vec{\nabla}\right)\vec{u} = -\vec{\nabla}\left(P + \frac{B^2}{8\pi}\right) - \rho\vec{\nabla}\Phi + \frac{\vec{B}\cdot\vec{\nabla}\vec{B}}{4\pi}$$

$$\frac{P}{\gamma-1}\left(\frac{\partial}{\partial t} + \vec{u}\cdot\vec{\nabla}\right)\ln P\rho^{-\gamma} = 0$$

$$\frac{\partial \vec{B}}{\partial t} = \vec{\nabla}\times(\vec{u}\times\vec{B})$$

Now, let us consider a uniform, stationary MHD fluid, threaded by a uniform magnetic field. To orient our coordinate system, we will use $\vec{B} = B_0\hat{z}$, with the directions $\perp$ to the field being x and y. We perturb the fluid with small displacements, which we take (freely) to be sinusoidal:

$$\rho = \rho_0 + \delta\rho\, e^{i\vec{k}\cdot\vec{r} - i\omega t}$$

$$\vec{B} = B_0\hat{z} + \vec{\delta B}\, e^{i\vec{k}\cdot\vec{r} - i\omega t}$$

$$\vec{u} = \phi + \vec{\delta u}\, e^{i\vec{k}\cdot\vec{r} - i\omega t}$$

$$P = P_0 + \delta P\, e^{i\vec{k}\cdot\vec{r} - i\omega t}$$

Small? What's "small"? By "small", I mean that all nonlinearities ($\propto O(\delta^2)$) will be dropped. The result is <u>linear theory</u>.

Before we do this, note that, when computing actual observed quantities, we should take the real part (e.g. $e^{i\theta} \to \cos\theta$, $ie^{i\theta} \to -\sin\theta$, etc.)

First, let's do the simplest thing: $\vec{k} = k\hat{z}$. My notation is usually "$k_\parallel$" in this case, to remind me that k is parallel to the guide field. This notation is used in a lot of plasma physics, but less so in astronomy.
Our linearized MHD eqns. are then

$$-i\omega \frac{\delta\rho}{\rho_0} + ik_\parallel \delta u_\parallel = 0$$

$$-i\omega \vec{\delta u} = -i\frac{k_\parallel \hat{z}}{\rho_0}\left(\delta p + \frac{B_0 \delta B_\parallel}{4\pi}\right) + \frac{ik_\parallel B_0}{4\pi\rho_0}\vec{\delta B}$$

$$-i\omega \frac{\vec{\delta B}}{B_0} = ik_\parallel \vec{\delta u} - \hat{z}\, ik_\parallel \delta u_\parallel \longrightarrow \delta B_\parallel = 0$$ (as is required by $\vec{k}\cdot\vec{\delta B} = 0$)

Note that we don't need to know $\delta\rho$ or δp to solve for the perpendicular ($\perp$) dynamics:

$$\left.\begin{aligned} -i\omega \vec{\delta u_\perp} &= \frac{ik_\parallel B_0}{4\pi\rho_0}\vec{\delta B_\perp} \\ -i\omega \frac{\vec{\delta B_\perp}}{B_0} &= ik_\parallel \vec{\delta u_\perp} \end{aligned}\right\} \left(\omega^2 - k_\parallel^2 V_A^2\right)\frac{\vec{\delta B_\perp}}{B_0} = 0$$

$$\boxed{\begin{gathered}\omega = \pm k_\parallel V_A \\ \text{with } V_A \equiv \frac{B_0}{\sqrt{4\pi\rho_0}}\end{gathered}}$$

These are "Alfvén waves", which are polarized across the guide field and which propagate at speed V_A, the "Alfvén speed". These waves are not associated with any motion along the field nor any changes in density.

Using $\frac{\delta p}{p_0} = \gamma\frac{\delta\rho}{\rho_0}$, the other modes are sound waves: $\boxed{\omega = \pm k_\parallel c_s}$, with $c_s \equiv \left(\frac{\gamma p_0}{\rho_0}\right)^{1/2}$ being the "sound speed".

The fifth mode is $\omega = 0$, and corresponds to a relabeling of fluid elements. It's called the "entropy mode"

Now, let's let $\vec{k} = k_\parallel \hat{z} + \vec{k}_\perp$ — a more general wavevector. Then our linearized equations are

$$\text{(a)} \quad -i\omega \frac{\delta\rho}{\rho_0} + i k_\parallel \delta u_\parallel + i\vec{k}_\perp \cdot \delta\vec{u}_\perp = 0$$

$$-i\omega\,\delta\vec{u} = -\frac{i\vec{k}}{\rho_0}\left(\delta p + \frac{B_0 \delta B_\parallel}{4\pi}\right) + \frac{i k_\parallel B_0}{4\pi\rho_0}\delta\vec{B}$$

$$-i\omega \frac{\delta\vec{B}}{B_0} = i k_\parallel \delta\vec{u} - \hat{z}\left(i k_\parallel \delta u_\parallel + i\vec{k}_\perp \cdot \delta\vec{u}_\perp\right)$$

$$\text{(b)} \quad -i\omega \frac{\delta\vec{B}_\perp}{B_0} = i k_\parallel \delta\vec{u}_\perp \quad \text{and} \quad \text{(c)} \quad -i\omega \frac{\delta B_\parallel}{B_0} = -i\vec{k}_\perp \cdot \delta\vec{u}_\perp$$

$$\text{(d)} \quad -i\omega\,\delta\vec{u}_\perp = -\frac{i\vec{k}_\perp}{\rho_0}\left(\delta p + \frac{B_0\delta B_\parallel}{4\pi}\right) + \frac{i k_\parallel B_0}{4\pi\rho_0}\delta\vec{B}_\perp \quad \text{and}$$

$$\text{(e)} \quad -i\omega\,\delta u_\parallel = -\frac{i k_\parallel}{\rho_0}\left(\delta p + \frac{B_0\delta B_\parallel}{4\pi}\right) + \frac{i k_\parallel B_0}{4\pi\rho_0}\delta B_\parallel.$$

$$\vec{k}_\perp \cdot (d) + k_\parallel (e) \Rightarrow -i\omega\left(\vec{k}_\perp \cdot \delta\vec{u}_\perp + k_\parallel \delta u_\parallel\right) = -\frac{i k^2}{\rho_0}\left(\delta p + \frac{B_0 \delta B_\parallel}{4\pi}\right) + \cancel{0}$$

use (a) :
$$\cancel{+i}\frac{\omega}{\cancel{i}}\left(\frac{\cancel{i}\omega\,\delta\rho}{\rho_0}\right) = \cancel{+i}\frac{k^2}{\rho_0}\left(\delta p + \frac{B_0\delta B_\parallel}{4\pi}\right)$$

* here we've lost the entropy mode *

use $\frac{\delta p}{p_0} = \gamma \frac{\delta\rho}{\rho_0}$:
$$\left(\omega^2 - k^2 c_s^2\right)\frac{\delta\rho}{\rho_0} = k^2 V_A^2 \frac{\delta B_\parallel}{B_0}$$

Now, (d) with (b) gives:
$$\left(\omega^2 - k_\parallel^2 V_A^2\right)\frac{\delta\vec{B}_\perp}{B_0} = -k_\parallel \vec{k}_\perp\left(c_s^2 \frac{\delta\rho}{\rho_0} + V_A^2 \frac{\delta B_\parallel}{B_0}\right)$$

$$\left(\omega^2 - k_\parallel^2 V_A^2\right)\frac{\delta\vec{B}_\perp}{B_0} = -k_\parallel \vec{k}_\perp \frac{\delta B_\parallel}{B_0}\left[c_s^2 \frac{k^2 V_A^2}{\omega^2 - k^2 c_s^2} + V_A^2\right]$$

Note that the parallel and perpendicular components are now coupled!

$$\left(\omega^2 - k_\parallel^2 v_A^2\right) \frac{\delta \vec{B}_\perp}{B_0} = -k_\parallel \vec{k}_\perp v_A^2 \left[\frac{\omega^2}{\omega^2 - k^2 c_s^2}\right] \frac{\delta B_\parallel}{B_0}$$

Before we go any further, note that, if $c_s^2/v_A^2 \gg 1$, then we have $\omega^2 - k_\parallel^2 v_A^2 \simeq 0$, so we get back something like an Alfvén wave in this limit. Proceeding by using $\frac{\delta B_\parallel}{B_0} = -\frac{\vec{k}_\perp \cdot \delta \vec{B}_\perp}{k_\parallel B_0}$, we have

$$\left[\overleftrightarrow{I}\left(\omega^2 - k_\parallel^2 v_A^2\right) - \vec{k}_\perp \vec{k}_\perp v_A^2 \frac{\omega^2}{\omega^2 - k^2 c_s^2}\right] \cdot \frac{\delta \vec{B}_\perp}{B_0} = 0.$$

Taking the determinant and setting it to zero gives the dispersion relation

$$\left(\omega^2 - k_\parallel^2 v_A^2\right)\left[\omega^2 - k_\parallel^2 v_A^2 - k_\perp^2 v_A^2 \frac{\omega^2}{\omega^2 - k^2 c_s^2}\right] = 0.$$

you'll often see this written as

$$\left[\omega^4 - \omega^2 k^2 \left(c_s^2 + v_A^2\right) + k_\parallel^2 v_A^2 k^2 c_s^2\right]$$

But I like it like this because you can take β limits easier.

Note that we recover the Alfvén wave solution $\omega = \pm k_\parallel v_A$. Now we also have $\omega^2 = \frac{k^2 (c_s^2 + v_A^2)}{2} \pm \sqrt{\frac{k^4 (c_s^2 + v_A^2)^2}{4} - k_\parallel^2 v_A^2 k^2 c_s^2}$.

These are the "magnetosonic" modes — the ⊕ solution being the "fast wave" and the ⊖ solution being the "slow wave".

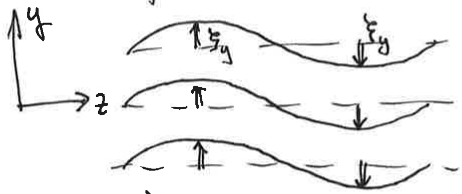
Note that, in the high-β limit, we have

$$\omega_+^2 \simeq \underbrace{k^2 C_s^2}_{\text{sound!}} \quad \text{and} \quad \omega_-^2 \simeq \underbrace{k_\parallel^2 v_A^2}_{\text{Alfvén!}}$$

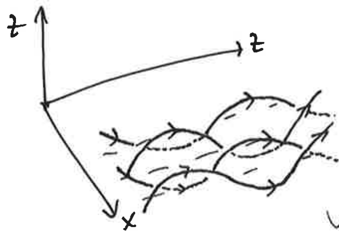
The difference between the slow mode here and an actual shear Alfvén wave is the latter involves no compressive fluctuations, being polarized with $\delta B_\parallel$ exactly $=0$. This is sometimes called a "pseudo-Alfvén" wave.

Here are some pictures of these waves: ($\vec{\xi}$ is displacement)

Alfvén:

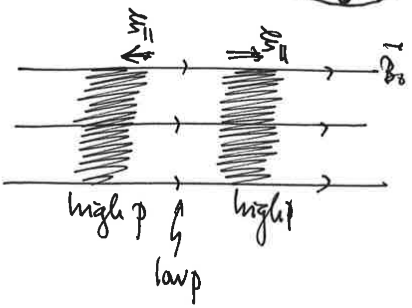


(NB: can have $k_\perp \neq 0$

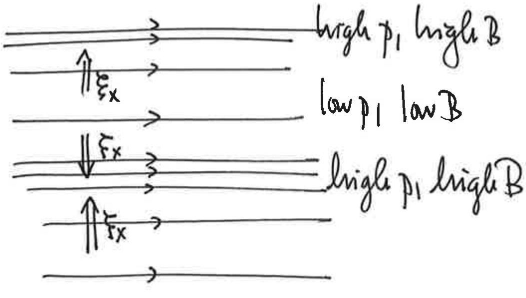


)

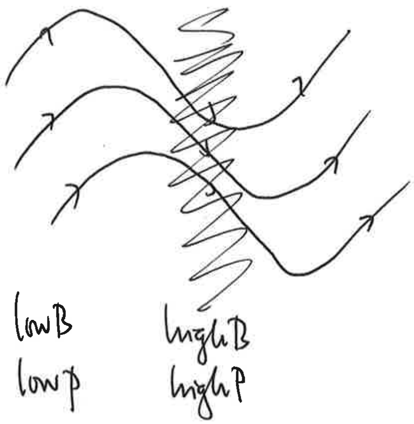
Sound:



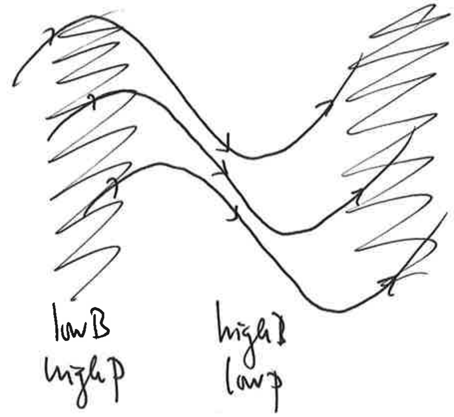
fast w/ $k_\parallel=0$:



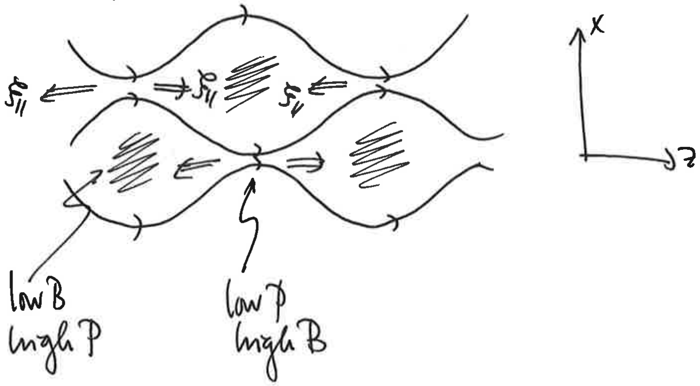
Fast:



Slow:



Slow with $k_\parallel/k_\perp \ll 1$:



Now, this last limit, $k_\parallel/k_\perp \ll 1$, is quite useful for studies of Alfvénic turbulence. What are all waves in this limit? Alfvén is the same: $\pm k_\parallel v_A$. Magnetosonic waves become

$$\omega^2 \approx \frac{k_\perp^2 (c_s^2+v_A^2)}{2}\left[1 \pm \left(1 - \frac{2k_\parallel^2 v_A^2 k_\perp^2 c_s^2}{k_\perp^4 (c_s^2+v_A^2)^2}\right)\right]$$

$$\oplus \text{ FAST} \rightarrow k_\perp^2 (c_s^2+v_A^2) \qquad \ominus \text{ SLOW} \rightarrow k_\parallel^2 v_A^2 \left(\frac{c_s^2}{c_s^2+v_A^2}\right)$$

· Let's look at the slow mode in this limit. Recall from our linear calculation that

$$\delta p = \delta\rho\, c_s^2 = \rho_0 c_s^2 \left(\frac{k^2 V_A^2}{\omega^2 - k^2 c_s^2}\right)\frac{\delta B_{\parallel}}{B_0} \Rightarrow \frac{\delta p}{p_0} + \left(\frac{k^2 V_A^2}{k^2 c_s^2 - \omega^2}\right)\frac{\delta B_{\parallel}}{B_0} = 0.$$

With $k_\perp \gg k_\parallel$ and $\omega^2 \approx k_\parallel^2 V_A^2 \left(\frac{c_s^2}{c_s^2 + V_A^2}\right)$, this becomes

$$\frac{\delta p}{p} + \frac{k_\perp^2 V_A^2 (\delta B_\parallel / B_0)}{k_\perp^2 c_s^2 - \frac{k_\parallel^2 V_A^2 c_s^2}{c_s^2 + V_A^2}} \approx \underbrace{\frac{\delta p}{p} + \frac{V_A^2}{c_s^2}\frac{\delta B_\parallel}{B_0} \approx 0}_{\text{pressure balance!}}$$

PART VI

Linear theory of MHD instabilities

Now let's do some MHD linear instabilities. The program is to set up some equilibria and then subject them to small-amplitude perturbations in the fluid and magnetic field. There are a few different ways of doing this and assessing whether the system is stable or unstable to these perturbations. There's something called the MHD energy principle, which will tell you whether a given set of perturbations about some equilibrium state will bring the system profitably to a lower energy state. There's something called Eulerian perturbation theory, where you subject the equilibrium state to small-amplitude perturbations, formulate those perturbations in the lab frame, and ask whether the perturbations oscillate, grow, or decay. And there's something called Lagrangian perturbation theory, which is same as Eulerian perturbation theory but is formulated in the frame of fluid. Each of these has its advantages depending on the equilibrium state, boundary conditions, and questions being asked. Eulerian perturbation theory is the most straightforward procedure, so we'll start there.

VI.1. A primer on instability

Before attacking the MHD equations, though, let's do something simpler to establish notation and learn the procedure. Consider the following ordinary differential equation:

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + 2\nu \frac{\mathrm{d}x}{\mathrm{d}t} + \varOmega^2 (x - x_0) = 0, \tag{VI.1}$$

where ν and $\varOmega > 0$ are constants. You may recognize this as the equation for a damped simple harmonic oscillator of natural frequency $\varOmega$ whose velocity along the x axis is damped at a rate $\nu > 0$. But let's not yet commit to any particular sign of ν. First, the equilibrium state. This is easy: the oscillator is at rest at $x = x_0$. We now displace the oscillator by a small amount ξ, so that $x(t) = x_0 + \xi(t)$. The equation governing this displacement is

$$\frac{\mathrm{d}^2 \xi}{\mathrm{d}t^2} + 2\nu \frac{\mathrm{d}\xi}{\mathrm{d}t} + \varOmega^2 \xi = 0. \tag{VI.2}$$

This equation admits solutions $\xi \sim \exp(-\mathrm{i}\omega t)$, where ω is a complex frequency that satisfies the *dispersion relation*

$$\omega^2 + 2\mathrm{i}\omega\nu - \varOmega^2 = 0 \quad \Longrightarrow \quad \omega = -\mathrm{i}\nu \pm \sqrt{\varOmega^2 - \nu^2}. \tag{VI.3}$$

How do we assess stability? If the imaginary part of ω is positive, then $-\mathrm{i}\omega$ has a positive real part, and the displacements will grow exponentially in time. If the imaginary part of ω is negative, then $-\mathrm{i}\omega$ has a negative real part, and this corresponds to exponential decay of the perturbation. If ω additionally has a real part, then this represents a growing or decaying oscillator. It's clear from a cursory glance at the dispersion relation (VI.3) that the perturbations oscillate and decay exponentially if $\varOmega > \nu > 0$. If $\nu > \varOmega > 0$, then the perturbations decay without oscillating. But if $\nu < 0$, then there is always an exponentially growing solution. Thus, $\nu > 0$ is the *stability criterion* for this system.

Now, suppose the equation of interest were instead

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + 2\nu \frac{\mathrm{d}x}{\mathrm{d}t} + \varOmega^2 \sin(x - x_0) = 0. \tag{VI.4}$$

The equilibrium is still the same, but if we want simple harmonic oscillator solutions, we're only go to get them if the displacement is small, i.e., $|\xi| \ll x_0$. In that case, we can Taylor expand $\sin(x - x_0) \approx \xi - \xi^3/6 + \dots$. To leading order in ξ, we're back to where we started with (VI.2). This is *linear theory*: identify an equilibrium, perturb the system about that equilibrium, and drop all terms nonlinear in the perturbation amplitude.

Note that we are not solving an initial value problems. We are agnostic about the initial conditions and only ask whether some disturbance will ultimately grow or decay. In some situations (most notably, Landau damping), solving the initial value problem is absolutely essential to obtain the full solution and all the physics involved. But if you just want to calculate the wave-like response of a system to infinitesimally small perturbations and learn whether such a response grows or decays, you need only adopt solutions $\sim \exp(-\mathrm{i}\omega t)$, find the dispersion relation for ω vs $\boldsymbol{k}$, and examine the sign of its imaginary part. (The difference is related to a Laplace vs a Fourier transform in time.)

VI.2. Linearized MHD equations

Take (IV.21) and write

$$\rho = \rho_0(\boldsymbol{r}) + \delta\rho(t,\boldsymbol{r}), \quad \boldsymbol{u} = \delta\boldsymbol{u}(t,\boldsymbol{r}), \quad P = P_0(\boldsymbol{r}) + \delta P(t,\boldsymbol{r}), \quad \boldsymbol{B} = \boldsymbol{B}_0(\boldsymbol{r}) + \delta\boldsymbol{B}(t,\boldsymbol{r});$$

i.e., consider a stratified, stationary equilibrium state threaded by a magnetic field and subject it to perturbations. Never mind how the equilibrium is set up – it is what it is, and we'll perturb it. Neglecting all terms quadratic in δ, equations (IV.21) become

$$\frac{\partial \delta\rho}{\partial t} = -(\delta\boldsymbol{u}\cdot\boldsymbol{\nabla})\rho_0 - \rho_0(\boldsymbol{\nabla}\cdot\delta\boldsymbol{u}), \tag{VI.5}$$

$$\frac{\partial \delta\boldsymbol{u}}{\partial t} = -\frac{1}{\rho_0}\boldsymbol{\nabla}\left(\delta P + \frac{\boldsymbol{B}_0\cdot\delta\boldsymbol{B}}{4\pi}\right) + \frac{\delta\rho}{\rho_0^2}\boldsymbol{\nabla}\left(P_0 + \frac{B_0^2}{8\pi}\right) + \frac{(\boldsymbol{B}_0\cdot\boldsymbol{\nabla})\delta\boldsymbol{B}}{4\pi\rho_0} + \frac{(\delta\boldsymbol{B}\cdot\boldsymbol{\nabla})\boldsymbol{B}_0}{4\pi\rho_0} - \boldsymbol{\nabla}\delta\Phi, \tag{VI.6}$$

$$\frac{\partial \delta\boldsymbol{B}}{\partial t} = -(\delta\boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{B}_0 + (\boldsymbol{B}_0\cdot\boldsymbol{\nabla})\delta\boldsymbol{u} - \boldsymbol{B}_0(\boldsymbol{\nabla}\cdot\delta\boldsymbol{u}), \tag{VI.7}$$

$$\frac{\partial}{\partial t}\left(\frac{\delta P}{P_0} - \gamma\frac{\delta\rho}{\rho_0}\right) = -\delta\boldsymbol{u}\cdot\boldsymbol{\nabla}\ln\frac{P_0}{\rho_0^\gamma}. \tag{VI.8}$$

(A quick way of getting these is to think of δ as a differential operator that commutes with partial differentiation.) Pretty much every gradient of an equilibrium quantity here will give an instability! (Otherwise, you just get back simple linear waves on a homogeneous background.) So let's not analyze this all at once. But I write this system of equations here for two important reasons: (i) it makes clear that we can adopt solutions $\delta \sim \exp(-\mathrm{i}\omega t)$ for the perturbations, since the equations are linear in the fluctuation amplitudes; (ii) we can only adopt full plane-wave solutions $\delta \sim \exp(-\mathrm{i}\omega t + \mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})$ if the fluctuations vary on length scales much smaller than that over which the background varies (the so-called WKB approximation). Otherwise, we have to worry about the exact structure of the background gradients and their boundary conditions.

So these are the themes of most linear stability analyses: a WKB approximation whereby plane-wave solutions are assumed on top of a background state that is slowly varying, and a focus only on whether fluctuations grow or decay rather than their specific spatio-temporal evolution from a set of initial conditions.

VI.3. Lagrangian versus Eulerian perturbations

There is one last thing worth discussing before proceeding with a linear stability analysis of the MHD equations. Just as there is an Eulerian time derivative and a Lagrangian time derivative, there is Eulerian perturbation theory and Lagrangian perturbation theory. The former, in which perturbations are denoted by a 'δ', measures the change in a quantity at a particular point in space. For example, if the equilibrium density at $\boldsymbol{r}$, $\rho(\boldsymbol{r})$, is changed at time t by some disturbance to become $\rho'(t,\boldsymbol{r})$, then we denote the Eulerian perturbation of the density by

$$\rho'(t,\boldsymbol{r}) - \rho(\boldsymbol{r}) \doteq \delta\rho \ll \rho(\boldsymbol{r}). \tag{VI.9}$$

Again, these perturbations are taken *at fixed position*. The latter – Lagrangian perturbation theory – concerns the evolution of small perturbations about a background state *within a particular fluid element* as it undergoes a displacement $\boldsymbol{\xi}$. For example, if a particularly fluid element is displaced from its equilibrium position $\boldsymbol{r}$ to position $\boldsymbol{r}+\boldsymbol{\xi}$, then the density of that fluid element changes by an amount

$$\rho'(t,\boldsymbol{r}+\boldsymbol{\xi}) - \rho(\boldsymbol{r}) \doteq \varDelta\rho. \tag{VI.10}$$

This is a Lagrangian perturbation. To linear order, δ and $\varDelta$ are related by

$$\varDelta\rho \simeq \rho'(t,\boldsymbol{r}) + \boldsymbol{\xi}\boldsymbol{\cdot}\boldsymbol{\nabla}\rho(\boldsymbol{r}) - \rho(\boldsymbol{r}) = \delta\rho + \boldsymbol{\xi}\boldsymbol{\cdot}\boldsymbol{\nabla}\rho. \tag{VI.11}$$

There are many situations in which a Lagrangian approach is easier to use than an Eulerian approach; there are also some situations in which doing so is absolutely necessary (e.g., see §IIIe of Balbus (1988) and §Ic of Balbus & Soker (1989) for discussions of the perils of using Eulerian perturbations in the context of local thermal instability).

Question: It is possible to have zero Eulerian perturbation and yet have finite Lagrangian perturbation. What does this mean physically? Is there a physical change in the system?

The Lagrangian velocity perturbation $\varDelta\boldsymbol{u}$ is given by

$$\varDelta\boldsymbol{u} \doteq \frac{\mathrm{D}\boldsymbol{\xi}}{\mathrm{D}t} = \left(\frac{\partial}{\partial t} + \boldsymbol{u}\boldsymbol{\cdot}\boldsymbol{\nabla}\right)\boldsymbol{\xi}, \tag{VI.12}$$

where $\boldsymbol{u}$ is the background velocity. It is the instantaneous time rate of rate of the displacement of a fluid element, taken relative to the unperturbed flow. Because $\varDelta\boldsymbol{u} = \delta\boldsymbol{u} + \boldsymbol{\xi}\boldsymbol{\cdot}\boldsymbol{\nabla}\boldsymbol{u}$, we have

$$\delta\boldsymbol{u} = \frac{\partial\boldsymbol{\xi}}{\partial t} + \boldsymbol{u}\boldsymbol{\cdot}\boldsymbol{\nabla}\boldsymbol{\xi} - \boldsymbol{\xi}\boldsymbol{\cdot}\boldsymbol{\nabla}\boldsymbol{u}. \tag{VI.13}$$

Note the additional $\boldsymbol{\xi}\boldsymbol{\cdot}\boldsymbol{\nabla}\boldsymbol{u}$ term, representing a measurement of the background fluid gradients by the fluid displacement.

Exercise. Let $\boldsymbol{u} = R\varOmega(R)\hat{\boldsymbol{\varphi}}$, as in a differentially rotating disk in cylindrical coordinates. Consider a displacement $\boldsymbol{\xi}$ with radial and azimuthal components ξ_R and ξ_φ, each depending upon R and φ. Show that

$$\frac{\mathrm{D}\xi_R}{\mathrm{D}t} = \delta u_R \quad \text{and} \quad \frac{\mathrm{D}\xi_\varphi}{\mathrm{D}t} = \delta u_\varphi + \xi_R\frac{\mathrm{d}\varOmega}{\mathrm{d}\ln R}. \tag{VI.14}$$

The second term in the latter equation accounts for the stretching of radial displacements into the azimuthal direction by the differential rotation.

You can think of δ and Δ as difference operators, since we're only working to linear order in the perturbation amplitude: e.g.,

$$\delta\left(\frac{1}{\rho}\right)=\frac{1}{\rho+\delta\rho}-\frac{1}{\rho}\simeq-\frac{\delta\rho}{\rho^2}.$$

But you must be very careful when mixing Eulerian and Lagrangian points of view. Prove the following commutation relations:

$$(i)\quad \left[\delta,\,\frac{\partial}{\partial t}\right]=0;$$

$$(ii)\quad \left[\delta,\,\frac{\partial}{\partial x_i}\right]=0;$$

$$(iii)\quad \left[\Delta,\,\frac{\partial}{\partial t}\right]=-\frac{\partial\xi_j}{\partial t}\frac{\partial}{\partial\xi_j};$$

$$(iv)\quad \left[\Delta,\,\frac{\partial}{\partial x_i}\right]=-\frac{\partial\xi_j}{\partial x_i}\frac{\partial}{\partial\xi_j};$$

$$(v)\quad \left[\Delta,\,\frac{\mathrm{D}}{\mathrm{D}t}\right]=0;$$

$$(vi)\quad \left[\Delta,\,\frac{\mathrm{D}}{\mathrm{D}x_i}\right]=-\xi_j\frac{\partial}{\partial x_j}\frac{\mathrm{D}}{\mathrm{D}t};$$

$$(vii)\quad \left[\frac{\partial}{\partial x_i},\,\frac{\mathrm{D}}{\mathrm{D}t}\right]=\frac{\partial u_j}{\partial x_i}\frac{\partial}{\partial x_j}.$$

You can use these to show that the linearized continuity equation, induction equation, and internal energy equation are

$$\frac{\Delta\rho}{\rho}=-\boldsymbol{\nabla}\cdot\boldsymbol{\xi}, \tag{VI.15}$$

$$\Delta\boldsymbol{B}=\boldsymbol{B}\cdot\boldsymbol{\nabla}\boldsymbol{\xi}-\boldsymbol{B}\boldsymbol{\nabla}\cdot\boldsymbol{\xi}, \tag{VI.16}$$

$$\frac{\Delta T}{T}=-(\gamma-1)\boldsymbol{\nabla}\cdot\boldsymbol{\xi}, \tag{VI.17}$$

respectively. These forms are particularly useful for linear analyses.

Now to calculate something... I'll start with two simple instabilities, the first of which (Jeans instability) will be analyzed using Eulerian perturbation theory, and the second of which (Kelvin–Helmholtz instability) will be analyzed using Lagrangian perturbation theory. Hopefully you'll see why one approach is sometimes easier than the other.

VI.4. Self-gravity: Jeans instability

One of the simplest hydrodynamical waves is a small-amplitude sound wave propagating on an infinite, homogeneous background. Take (IV.21), set $\boldsymbol{B}_0=0$, and assume ρ_0 and P_0 to be constant. The resulting linearized equations are

$$\frac{\partial}{\partial t}\frac{\delta\rho}{\rho_0}=-\boldsymbol{\nabla}\cdot\delta\boldsymbol{u},\qquad \frac{\partial\delta\boldsymbol{u}}{\partial t}=-\frac{1}{\rho_0}\boldsymbol{\nabla}\delta P-\boldsymbol{\nabla}\delta\Phi,\qquad \frac{\partial}{\partial t}\left(\frac{\delta P}{P_0}-\gamma\frac{\delta\rho}{\rho_0}\right)=0. \tag{VI.18a}$$

I've retained the perturbed gravitational potential $\delta\Phi$ in the second equation, because we're going to assume that the fluid is self-gravitating with a potential that obey's

Poisson's equation:[8]

$$\nabla^2 \delta\Phi = 4\pi G \delta\rho. \tag{VI.18b}$$

These equations are linear in δ, and so we may adopt plane-wave solutions, $\delta \sim \exp(-\mathrm{i}\omega t + \mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})$. Substituting this form into (VI.18) gives

$$-\mathrm{i}\omega \frac{\delta\rho}{\rho_0} = -\mathrm{i}\boldsymbol{k}\cdot\delta\boldsymbol{u}, \quad -\mathrm{i}\omega\delta\boldsymbol{u} = -\mathrm{i}\boldsymbol{k}\frac{\delta P}{\rho_0} - \mathrm{i}\boldsymbol{k}\delta\Phi, \quad -\mathrm{i}\omega\left(\frac{\delta P}{P_0} - \gamma\frac{\delta\rho}{\rho_0}\right) = 0, \tag{VI.19a}$$

$$-k^2 \delta\Phi = 4\pi G \delta\rho. \tag{VI.19b}$$

Taking $\boldsymbol{k}\cdot$ the second equation and using the other three yields the dispersion relation

$$\omega\bigl(\omega^2 - k^2 a^2 + 4\pi G \rho_0\bigr) = 0, \tag{VI.20}$$

where $a^2 \doteq \gamma P_0/\rho_0$. The $\omega = 0$ root comes from the perturbed entropy equation, and corresponds to a isentropic relabelling of the fluid elements; its name is the 'entropy mode'. The other two roots correspond to forward- and backward-propagating sound waves under the influence of their own self-gravity:

$$\omega = \pm ka\sqrt{1 - \frac{4\pi G\rho_0}{k^2 a^2}} \tag{VI.21}$$

Self-gravity reduces the speed of the wave for wavenumbers satisfying $ka > (4\pi G\rho_0)^{1/2}$, for which the (expansive) pressure force is greater than the (attractive) gravitational force. At $ka = (4\pi G\rho_0)^{1/2}$, these two forces balance exactly, and the mode is neutrally stable. But for $ka < (4\pi G\rho_0)^{1/2}$, the wavelength is long enough to include a sufficiently large amount of mass in the perturbation to overwhelm the pressure force. Instability ensues, and the mode grows without propagating. This is the *Jeans instability*, named after Sir James Jeans (although Sir Isaac Newton understood the concept over 200 years before the calculation).

The critical wavelength

$$\lambda_{\mathrm{J}} = a\sqrt{\frac{\pi}{G\rho_0}} \tag{VI.22}$$

is referred to as the *Jeans length*. For an isothermal ($\gamma = 1$) molecular cloud of temperature 10 K, number density 200 cm^{-3}, and mean mass per particle $2.33m_{\mathrm{p}}$, the Jeans length is $\simeq$1.5 pc. The corresponding *Jeans mass* enclosed within a spherical volume with λ_{J} as its diameter is

$$M_{\mathrm{J}} = \frac{\pi}{6}\rho_0\lambda_{\mathrm{J}}^3 = 20.3\left(\frac{T_0}{10~\mathrm{K}}\right)^{3/2}\left(\frac{n}{200~\mathrm{cm}^{-3}}\right)^{-1/2}~\mathrm{M}_\odot. \tag{VI.23}$$

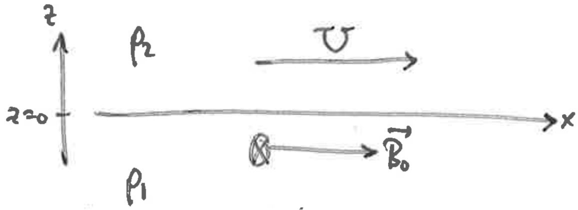
Giant molecular clouds with these parameters have typical masses $\gtrsim 10^4~\mathrm{M}_\odot$, indicating that more must be going on than just thermal pressure support against self-gravity (see: magnetic fields and turbulence). Note that $M_{\mathrm{J}} = M_\odot$ at a density $n \simeq 8.2\times10^4~\mathrm{cm}^{-3}$.

VI.5. Shear: Kelvin–Helmholtz instability

Consider two uniform fluids separated by a discontinuous interface at $z = 0$, as in the figure below:

---

[8] Wouldn't an infinite, homogeneous, self-gravitating fluid collapse under its own weight? Indeed it would. Ignoring this inconvenience is known as the *Jeans swindle*. Following Binney & Tremaine (1987): 'it is a swindle because in general there is no formal justification for discarding the unperturbed gravitational field'.



The fluid above the interface ($z > 0$) has density ρ_2 and equilibrium velocity $\boldsymbol{u}_0 = U\hat{\boldsymbol{x}}$. The fluid below the interface ($z < 0$) has density ρ_1 and is stationary. (We can always transform to a frame in which this fluid is stationary, so why not take advantage of that?) There is a uniform magnetic field $\boldsymbol{B}_0 = B_{0x}\hat{\boldsymbol{x}} + B_{0y}\hat{\boldsymbol{y}}$ oriented parallel to the interface that permeates all of the fluid, which we take to be perfectly conducting. For simplicity, take the fluid to be incompressible, *viz.* $\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0$.

We seek the dispersion relation governing small-amplitude perturbations. It turns out that this problem is most easily analyzed using Lagrangian perturbations rather than Eulerian perturbations – the reason being that the interface and the interfacial pressure between the two fluids must remain continuous as the fluid is perturbed, and it's easier to measure this interface in the frame of the fluid element than in the lab frame.

Take the momentum equation in each of the fluids, above and below, and apply the difference operator $\varDelta \doteq \delta + \boldsymbol{\xi}\cdot\boldsymbol{\nabla}$ while recalling that $[\varDelta, \mathrm{D}/\mathrm{D}t] = 0$ and $\varDelta\boldsymbol{u} = \mathrm{D}\boldsymbol{\xi}/\mathrm{D}t$:

$$\varDelta\left[\rho\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} = -\boldsymbol{\nabla}\left(P + \frac{B^2}{8\pi}\right) + \frac{\boldsymbol{B}\cdot\boldsymbol{\nabla}\boldsymbol{B}}{4\pi}\right]$$

$$\Longrightarrow\quad \rho\frac{\mathrm{D}^2\boldsymbol{\xi}}{\mathrm{D}t^2} = -\boldsymbol{\nabla}\delta\left(P + \frac{B^2}{8\pi}\right) + \frac{\boldsymbol{B}_0\cdot\boldsymbol{\nabla}\delta\boldsymbol{B}}{4\pi}, \tag{VI.24}$$

the form of the right-hand side following because $\boldsymbol{\nabla}B_0 = \boldsymbol{\nabla}P_0 = 0$. Use the linearized induction equation (VI.16) with $\boldsymbol{\nabla}B_0 = 0$, which reads $\delta\boldsymbol{B} = (\boldsymbol{B}_0\cdot\boldsymbol{\nabla})\boldsymbol{\xi}$, and rearrange to obtain

$$\left[\frac{\mathrm{D}^2}{\mathrm{D}t^2} - \frac{(\boldsymbol{B}_0\cdot\boldsymbol{\nabla})^2}{4\pi\rho}\right]\boldsymbol{\xi} = -\frac{1}{\rho}\boldsymbol{\nabla}\delta\left(P + \frac{B^2}{8\pi}\right) \doteq -\frac{1}{\rho}\boldsymbol{\nabla}\delta\varPi. \tag{VI.25}$$

Note that taking the divergence of this equation and using $\boldsymbol{\nabla}\cdot\boldsymbol{\xi} = 0$ (incompressibility) implies that the total perturbed pressure $\varPi$ satisfies

$$\nabla^2\delta\varPi = 0. \tag{VI.26}$$

With the x and y directions being infinite in extent and the background state possessing no structure in those directions, we may write $\delta\varPi = \delta\varPi(z)\exp(\mathrm{i}k_x x + \mathrm{i}k_y y)$ to find

$$\left(-k^2 + \frac{\mathrm{d}^2}{\mathrm{d}z^2}\right)\delta\varPi(z) = 0 \quad\Longrightarrow\quad \delta\varPi(z) \propto \exp(-|kz|),\quad k \equiv \sqrt{k_x^2 + k_y^2}. \tag{VI.27}$$

The absolute value in the argument of the exponential indicates that the perturbation must die off as $z \to \pm\infty$. We may now adopt solutions of the form $\exp(-\mathrm{i}\omega t)$ and evaluate the z component of (VI.25) above and below the interface:

$$\left[(-\mathrm{i}\omega + \mathrm{i}k_x U)^2 + \frac{(\boldsymbol{k}\cdot\boldsymbol{B}_0)^2}{4\pi\rho_2}\right]\xi_{z2} = +\frac{1}{\rho_2}|k|\delta\varPi_2, \tag{VI.28a}$$

$$\left[(-\mathrm{i}\omega \qquad\quad)^2 + \frac{(\boldsymbol{k}\cdot\boldsymbol{B}_0)^2}{4\pi\rho_1}\right]\xi_{z1} = -\frac{1}{\rho_1}|k|\delta\varPi_1, \tag{VI.28b}$$

respectively. At the interface, $\xi_{z1} = \xi_{z2}$ and $\varDelta\varPi_1 = \varDelta\varPi_2$, i.e., the two fluids must move together at the interface and their pressures must hold continuous as they are perturbed.

Because $\boldsymbol{\nabla} B_0 = \boldsymbol{\nabla} P_0 = 0$, the latter implies $\delta\Pi_1 = \delta\Pi_2$. Using this information to match (VI.28*a*) and (VI.28*b*) leads to

$$(\omega - k_x U)^2 \rho_2 + \omega^2 \rho_1 = \frac{(\boldsymbol{k}\boldsymbol{\cdot}\boldsymbol{B}_0)^2}{2\pi} \tag{VI.29}$$

$$\Longrightarrow \quad \boxed{\omega = \frac{k_x U}{2}\frac{\overline{\rho}}{\rho_1}\left\{1 \pm \mathrm{i}\sqrt{\frac{\rho_1}{\rho_2}\left[1 - \frac{(\boldsymbol{k}\boldsymbol{\cdot}\boldsymbol{B}_0)^2}{\pi\overline{\rho}k_x^2U^2}\right]}\right\}} \tag{VI.30}$$

where $\overline{\rho} \doteq 2\rho_1\rho_2/(\rho_1+\rho_2)$ is the reduced mass density. For

$$\frac{(\boldsymbol{k}\boldsymbol{\cdot}\boldsymbol{B}_0)^2}{4\pi\overline{\rho}} < \left(\frac{k_x U}{2}\right)^2, \tag{VI.31}$$

the discriminant is positive and there is a growing (and propagating) mode whose growth rate is proportional to the wavenumber and the velocity shear across the interface. Note that, for $\rho_1 = \rho_2 = \rho$, we have $\overline{\rho} = \rho$, and then (VI.30) becomes

$$\omega = \frac{k_x U}{2}\left[1 \pm \mathrm{i}\sqrt{1 - \frac{(\boldsymbol{k}\boldsymbol{\cdot}\boldsymbol{B}_0)^2}{\pi\rho k_x^2 U^2}}\right];$$

for $U = 0$, this returns a stably propagating shear Alfvén wave, $\omega = \mp(\boldsymbol{k}\boldsymbol{\cdot}\boldsymbol{v}_{\mathrm{A}})$. This indicates that it is the tension in the magnetic-field lines that is responsible for stabilizing the instability. That being said, if the magnetic field is oriented such that $B_{0x} = 0$, then (VI.31) can always be satisfied for small enough $|k_y/k_x|$, no matter how strong is B_{0y}.

The physics is as follows. An upwardly displaced distortion of the interface into region 2 causes a constriction of the velocity there, and the fluid must move faster to conserve its mass. But when it moves faster, the pressure must drop (Bernoulli!). The opposite happens below the interface. Now there is a pressure gradient pushing upwards, reinforcing the displacement, and the process runs away (unless the magnetic tension can stabilize the displacements and propagate them away as Alfvén waves). That's why pressure perturbations were vital in (VI.25).

Question: Does this instability occur in a simple linear shear flow, e.g., $\boldsymbol{u}_0 = Sz\hat{\boldsymbol{x}}$? No! The proof goes as follows. Drop the magnetic field for simplicity. With $\boldsymbol{u}_0 = u_0(z)\hat{\boldsymbol{x}}$, one can show using $\boldsymbol{\nabla\cdot\xi} = 0$ and the momentum equation that

$$\frac{\mathrm{d}^2\xi_z}{\mathrm{d}z^2} - k_x^2\xi_z = \frac{k u_0''}{\omega - k_x u_0}\xi_z.$$

Multiply this by ξ_z^* (the '$*$' denotes the complex conjugate) and integrate between the upper and lower boundaries $z = \pm L$ to obtain

$$\int_{-L}^{L}\mathrm{d}z\left(\xi_z^*\xi_z'' - k_x^2|\xi_z|^2\right) = \int_{-L}^{L}\mathrm{d}z\,\frac{k_x u_0''}{\omega - k_x u_0}|\xi_z|^2.$$

The first term on the left-hand side may be simplified using integration by parts and assuming either periodicity or that ξ_z or ξ_z' vanish at the boundaries. Then

$$\int_{-L}^{L}\mathrm{d}z\left(-|\xi_z'|^2 - k_x^2|\xi_z|^2\right) = \int_{-L}^{L}\mathrm{d}z\,\frac{k_x u_0''}{\omega - k_x u_0}|\xi_z|^2,$$

If the system is unstable, then ω must have an imaginary part, ω_{I}. Writing $\omega = \omega_{\mathrm{R}} + \mathrm{i}\omega_{\mathrm{I}}$, the imaginary part of the above equation is simply

$$\omega_{\rm I} \int_{-L}^{L} {\rm d}z \, \frac{k_x u_0''}{|\omega - k_x u_0|^2} |\xi_z|^2 = 0.$$

This states that u_0'' must be positive over part of the integration range, and negative over the remainder, i.e., u_0'' must pass through zero. Thus, instability requires an *inflection point* (Rayleigh 1880). (Note that the converse is not true: a velocity profile *with* an inflection point is not necessarily unstable.)

VI.6. Buoyancy: Rayleigh–Taylor instability

Using Lagrangian perturbation theory, it is easy to generalize the calculation in the previous section (§VI.5) to include gravity. Again, let the fluid above the interface ($z > 0$) have uniform density ρ_2, and the fluid below the interface ($z < 0$) have uniform density ρ_1. Include the same uniform background magnetic field as before, $\boldsymbol{B}_0 = B_{0x}\hat{\boldsymbol{x}} + B_{0y}\hat{\boldsymbol{y}}$. But now place these fluids in a constant gravitational field $\boldsymbol{g} = -g\hat{\boldsymbol{z}}$, with the gas pressure either side of the interface satisfying hydrostatic equilibrium in the vertical direction:

$$g = -\frac{1}{\rho_1}\frac{{\rm d}P_1}{{\rm d}z} = -\frac{1}{\rho_2}\frac{{\rm d}P_2}{{\rm d}z}.$$

The entire calculation goes through as before, but with the following additions and modifications. First, we must include the perturbed gravitational force in the momentum equation (VI.24), *viz.* $\Delta(\rho\boldsymbol{g}) = -(\Delta\rho)g\hat{\boldsymbol{z}}$. Secondly, because of the background pressure gradient in each of the fluids, we no longer have that $\Delta(\boldsymbol{\nabla}P) = \boldsymbol{\nabla}\delta P$, but rather that $\Delta(\boldsymbol{\nabla}P) = \boldsymbol{\nabla}\delta P + \boldsymbol{\xi}\boldsymbol{\cdot}\boldsymbol{\nabla}(\boldsymbol{\nabla}P)$. Using hydrostatic equilibrium, this may equivalently be written as $\Delta(\boldsymbol{\nabla}P) = \boldsymbol{\nabla}\delta P - \boldsymbol{\xi}\boldsymbol{\cdot}\boldsymbol{\nabla}(\rho g\hat{\boldsymbol{z}})$. Making these two changes in (VI.24) leads to

$$\rho\frac{{\rm D}^2\boldsymbol{\xi}}{{\rm D}t^2} = -\boldsymbol{\nabla}\delta\left(P + \frac{B^2}{8\pi}\right) + \frac{\boldsymbol{B}_0\boldsymbol{\cdot}\boldsymbol{\nabla}\delta\boldsymbol{B}}{4\pi} - (\Delta\rho)g\hat{\boldsymbol{z}} + \boldsymbol{\xi}\boldsymbol{\cdot}\boldsymbol{\nabla}(\rho g). \tag{VI.32}$$

Despite this extra work, however, those two additional terms cancel one another if the fluid is incompressible, since then $\Delta\rho - \boldsymbol{\xi}\boldsymbol{\cdot}\boldsymbol{\nabla}\rho \doteq \delta\rho = 0$. As a result, the *only* difference between this calculation and the Kelvin–Helmholtz calculation in §VI.5 is that the imposition of pressure continuity at the perturbed interface does not imply that $\delta\Pi_1 = \delta\Pi_2$, but rather

$$\Delta\Pi_1 = \Delta\Pi_2 \implies \delta\Pi_1 - \xi_{z1}\rho_1 g = \delta\Pi_2 - \xi_{z2}\rho_2 g \implies \delta\Pi_2 - \delta\Pi_1 = \xi_z(\rho_2 - \rho_1)g.$$

We may then use this in (VI.28) to jump straight to the dispersion relation (cf. (VI.29))

$$(\omega - k_x U)^2\rho_2 + \omega^2\rho_1 = \frac{(\boldsymbol{k}\boldsymbol{\cdot}\boldsymbol{B}_0)^2}{2\pi} + |k|g(\rho_1 - \rho_2), \tag{VI.33}$$

whose solutions are (cf. (VI.30))

$$\omega = \frac{k_x U}{2}\frac{\bar{\rho}}{\rho_1}\left\{1 \pm {\rm i}\sqrt{\frac{\rho_1}{\rho_2}\left[1 + \frac{2|k|g}{k_x^2U^2}\frac{\rho_2 - \rho_1}{\bar{\rho}} - \frac{(\boldsymbol{k}\boldsymbol{\cdot}\boldsymbol{B}_0)^2}{\pi\bar{\rho}k_x^2U^2}\right]}\right\}. \tag{VI.34}$$

By design, this has both Kelvin–Helmholtz and Rayleigh–Taylor in it; let's set $U = 0$ to eliminate the former, in which case

$$\boxed{\omega = \pm{\rm i}\sqrt{|k|g\frac{\rho_2 - \rho_1}{\rho_1 + \rho_2} - \frac{(\boldsymbol{k}\boldsymbol{\cdot}\boldsymbol{B}_0)^2}{2\pi(\rho_1 + \rho_2)}}} \tag{VI.35}$$

This equation states that linear instability requires $\rho_2 > \rho_1$ (heavy on top, light on the bottom), with the difference between the densities being large enough for the destabilizing pressure gradient (Bernoulli!) to overcome the stabilizing magnetic tension. Note that, if $\boldsymbol{B}_0$ is not oriented along the interface, no amount of magnetic field can stabilize the system.

VI.7. Buoyancy: Convective (Schwarzschild) instability

Next up: stratification. Henceforth, ignore self-gravity. Suppose our plasma is immersed in a constant, externally imposed gravitational field $\boldsymbol{g} = -g\hat{\boldsymbol{z}}$ and that its thermal-pressure gradient balances the gravitational acceleration to produce a stationary, equilibrium state. Ignoring for the moment magnetic fields, this hydrostatic equilibrium is described by the equation

$$\frac{1}{\rho_0}\frac{\mathrm{d}P_0}{\mathrm{d}z} = g = \mathrm{const}, \tag{VI.36}$$

where $\rho_0 = \rho_0(z)$. The hydrodynamic equations linearized about this equilibrium are

$$\frac{\partial}{\partial t}\frac{\delta\rho}{\rho_0} + \boldsymbol{\nabla\cdot}\delta\boldsymbol{u} + \delta u_z\frac{\mathrm{d}\ln\rho_0}{\mathrm{d}z} = 0, \tag{VI.37}$$

$$\frac{\partial\delta\boldsymbol{u}}{\partial t} = -\frac{1}{\rho_0}\boldsymbol{\nabla}\delta P - \frac{\delta\rho}{\rho_0}g\hat{\boldsymbol{z}}, \tag{VI.38}$$

$$\frac{\partial}{\partial t}\left(\frac{\delta P}{P_0} - \gamma\frac{\delta\rho}{\rho_0}\right) + \delta u_z\frac{\mathrm{d}}{\mathrm{d}z}\ln\frac{P_0}{\rho_0^\gamma} = 0. \tag{VI.39}$$

Solutions to this set of equations are $\propto\exp(-\mathrm{i}\omega t)$:

$$-\mathrm{i}\omega\frac{\delta\rho}{\rho_0} + \boldsymbol{\nabla\cdot}\delta\boldsymbol{u} + \delta u_z\frac{\mathrm{d}\ln\rho_0}{\mathrm{d}z} = 0, \tag{VI.40}$$

$$-\mathrm{i}\omega\delta\boldsymbol{u} = -\frac{1}{\rho_0}\boldsymbol{\nabla}\delta P - \frac{\delta\rho}{\rho_0}g\hat{\boldsymbol{z}}, \tag{VI.41}$$

$$-\mathrm{i}\omega\left(\frac{\delta P}{P_0} - \gamma\frac{\delta\rho}{\rho_0}\right) + \delta u_z\frac{\mathrm{d}}{\mathrm{d}z}\ln\frac{P_0}{\rho_0^\gamma} = 0. \tag{VI.42}$$

Continued on hand-written notes…

In general, we cannot Fourier transform these eqns. in z, because the coefficients in front of the perturbed quantities are z-dependent. But we can do so in the horizontal (say, x) direction:

$$-i\omega \frac{\delta\rho}{\rho_0} + i k_x \delta u_x + \frac{d\,\delta u_z}{dz} + \delta u_z \frac{d\ln\rho_0}{dz} = 0,$$

$$-i\omega\,\delta u_x = -i k_x \frac{\delta p}{\rho_0},$$

$$-i\omega\,\delta u_z = -\frac{1}{\rho_0}\frac{d\,\delta p}{dz} - \frac{\delta\rho}{\rho_0} g,$$

$$-i\omega\left(\frac{\delta p}{p_0} - \gamma\frac{\delta\rho}{\rho_0}\right) + \delta u_z \frac{d\ln p_0\rho_0^{-\gamma}}{dz} = 0,$$

where now the fluctuations are z-dependent Fourier amplitudes. Denoting $\vec{\delta u} = -i\omega\vec{\xi}$, and dropping the equilibrium "0" subscripts for notational ease, we have

(A) $$\frac{\delta\rho}{\rho} + i k_x \xi_x + \xi_z' + \xi_z \frac{d\ln\rho}{dz} = 0,$$

(B) $$-\omega^2 \xi_x = -i k_x \frac{\delta p}{\rho},$$

(C) $$-\omega^2 \xi_z = -\frac{1}{\rho}\frac{d}{dz}\delta p - \frac{\delta\rho}{\rho} g,$$

(D) $$\frac{\delta p}{p} = \gamma\frac{\delta\rho}{\rho} - \xi_z \frac{d\ln p\rho^{-\gamma}}{dz}.$$

(B) and (D) $\Rightarrow$ $$-\omega^2 \xi_x = -ik_x \frac{P}{\rho}\left[\gamma \frac{\delta\rho}{\rho} - \xi_z \frac{d\ln p\rho^{-\gamma}}{dz}\right]$$

and (A) $\Rightarrow$ $$-\omega^2 \xi_x = +ik_x \frac{P}{\rho}\gamma\left[+ik_x\xi_x + \xi_z' + \xi_z \frac{d\ln\rho}{dz}\right] + ik_x\frac{P}{\rho}\xi_z \frac{d\ln p\rho^{-\gamma}}{dz}$$

$$\Rightarrow (-\omega^2 + k_x^2 a^2)\xi_x = ik_x a^2 \xi_z' + \frac{ik_x a^2}{\gamma}\frac{d\ln p}{dz}\xi_z,$$

where $a^2 \equiv \gamma P/\rho$. Note: $g = -\frac{a^2}{\gamma}\frac{d\ln p}{dz}$, so this is

(E) $$\boxed{(-\omega^2 + k_x^2 a^2)\xi_x = ik_x a^2 \xi_z' - ik_x \xi_z g}$$

(A) $\Rightarrow$ $$\frac{\delta\rho}{\rho} = -\xi_z' - \xi_z\frac{d\ln\rho}{dz} - \frac{ik_x\left[ik_x a^2 \xi_z' - ik_x\xi_z g\right]}{-\omega^2 + k_x^2 a^2}$$

$$\Rightarrow \boxed{\frac{\delta\rho}{\rho} = \frac{\omega^2 \xi_z' + \left[(\omega^2 - k_x^2 a^2)\frac{d\ln\rho}{dz} - k_x^2 g\right]\xi_z}{k_x^2 a^2 - \omega^2}} \quad (a)$$

$$\Rightarrow \frac{\delta p}{p} = \frac{1}{k_x^2 a^2 - \omega^2}\left[\gamma\omega^2\xi_z' + \gamma\left((\omega^2 - k_x^2 a^2)\frac{d\ln\rho}{dz} - k_x^2 g\right)\xi_z + (\omega^2 - k_x^2 a^2)\frac{d\ln p\rho^{-\gamma}}{dz}\xi_z\right]$$

$$= \frac{1}{k_x^2 a^2 - \omega^2}\left[\gamma\omega^2\xi_z' + \omega^2\frac{d\ln p}{dz}\xi_z\right]$$

$$= \boxed{\frac{\omega^2}{k_x^2 a^2 - \omega^2}\left[\gamma\xi_z' + \frac{d\ln p}{dz}\xi_z\right]} \quad (\star)$$

into (c):

$$+\omega^2 \xi_z = +g\left[\frac{\omega^2 \xi_z' + (\omega^2 - k_x^2 a^2)\frac{d\ln p}{dz}\xi_z - k_x^2 g \xi_z}{k_x^2 a^2 - \omega^2}\right]$$

$$+ \frac{1}{\rho}\frac{d}{dz}\left[\frac{\omega^2 p}{k_x^2 a^2 - \omega^2}\left(\xi_z' \gamma + \xi_z \frac{d\ln p}{dz}\right)\right]$$

$$= \frac{\frac{\omega^2}{\rho}\frac{dp}{dz}\left(\xi_z'\gamma + \xi_z \frac{d\ln p}{dz}\right)}{k_x^2 a^2 - \omega^2} + \frac{p}{\rho}\frac{\omega^2}{k_x^2 a^2 - \omega^2}\left(\gamma \xi_z'' + \xi_z' \frac{d\ln p}{dz} + \xi_z \frac{d^2 \ln p}{dz^2}\right)$$

$$- \frac{\omega^2 p}{\rho}\frac{k_x^2}{(k_x^2 a^2 - \omega^2)^2}\frac{da^2}{dz}\left(\xi_z'\gamma + \xi_z \frac{d\ln p}{dz}\right)$$

$$\Rightarrow \omega^2 \xi_z = \frac{g}{k_x^2 a^2 - \omega^2}\left[\omega^2 \xi_z' - k_x^2 g \xi_z + (\omega^2 - k_x^2 a^2)\frac{d\ln p}{dz}\xi_z\right]$$

$$+ \frac{\omega^2}{k_x^2 a^2 - \omega^2}(-g)\left[\xi_z'\gamma + \xi_z \frac{d\ln p}{dz}\right]$$

$$+ \left(\frac{a^2}{\gamma}\right)\frac{\omega^2}{k_x^2 a^2 - \omega^2}\left[\xi_z''\gamma + \xi_z' \frac{d\ln p}{dz} + \xi_z \frac{g\gamma}{a^2}\frac{d\ln T}{dz}\right]$$

$$- \frac{\omega^2}{\gamma}\frac{a^2 k_x^2 a^2}{(k_x^2 a^2 - \omega^2)^2}\frac{d\ln T}{dz}\left[\xi_z'\gamma + \xi_z \frac{d\ln p}{dz}\right]$$

Multiply by $\frac{k_x^2 a^2 - \omega^2}{\omega^2 a^2}$ and group:

ξ_z'': 1.

$$\xi_z': \quad \cancel{\frac{g}{a^2}} - \frac{g\gamma}{a^2} + \cancel{\frac{1}{\gamma}\frac{d\ln p}{dz}} - \frac{k_x^2a^2}{\cancel{\gamma}}\frac{1}{(k_x^2a^2-\omega^2)}\frac{d\ln T}{dz}\cancel{\gamma}$$

$$= \frac{d\ln p}{dz} - \left(\frac{k_x^2a^2}{k_x^2a^2-\omega^2}\right)\frac{d\ln T}{dz} = \frac{d\ln p/dz}{k_x^2a^2-\omega^2}\left[k_x^2a^2-\omega^2 - k_x^2a^2\frac{d\ln T}{d\ln p}\right]$$

$$= \frac{d\ln p/dz}{k_x^2a^2-\omega^2}\left[-\omega^2 + k_x^2a^2\frac{d\ln\rho}{d\ln p}\right] = \frac{\omega^2\frac{d\ln p}{dz} - k_x^2a^2\frac{d\ln\rho}{dz}}{\omega^2-k_x^2a^2}$$

$$\xi_z: \quad -\frac{(k_x^2a^2-\omega^2)}{a^2} - \frac{k_x^2g^2}{\omega^2a^2} - g\frac{d\ln p}{dz}\frac{k_x^2a^2-\cancel{\omega^2}}{\omega^2a^2}$$

$$-\cancel{\frac{g}{a^2}\frac{d\ln p}{dz}} + \cancel{\frac{a^2}{\gamma}\frac{1}{a^2}\frac{g\gamma}{a^2}\frac{d\ln T}{dz}} - \frac{k_x^2a^2}{\gamma(k_x^2a^2-\omega^2)}\frac{d\ln T}{dz}\frac{d\ln p}{dz}$$

$$= \frac{-1}{k_x^2a^2-\omega^2}\left[\frac{k_x^2a^2}{\gamma}\frac{d\ln T}{dz}\frac{d\ln p}{dz} + \frac{gk_x^2}{\omega^2}\frac{d\ln p}{dz}(k_x^2a^2-\omega^2) + \frac{k_x^2g^2}{a^2\omega^2}(k_x^2a^2-\omega^2) + \frac{(k_x^2a^2-\omega^2)^2}{a^2}\right]$$

$$= \frac{1}{\omega^2-k_x^2a^2}\left[\frac{\omega^4}{a^2} - 2\omega^2k_x^2 + k_x^4a^2 - \frac{k_x^2g^2}{a^2} - \cancel{gk_x^2\frac{d\ln p}{dz}} + \frac{k_x^2a^2}{\gamma}\frac{d\ln T}{dz}\frac{d\ln p}{dz} + \frac{gk_x^4a^2}{\omega^2}\frac{d\ln p}{dz} + \frac{k_x^4\cancel{a^2}g^2}{\cancel{a^2}\omega^2}\right]$$

$$-k_x^2g\left(\frac{d\ln p}{dz} - \cancel{\frac{d\ln\rho}{dz}}\right)$$

So,

$$\boxed{\xi_z'' + \xi_z' \left[\frac{\omega^2 \frac{d\ln p}{dz} - k_x^2 a^2 \frac{d\ln \rho}{dz}}{\omega^2 - k_x^2 a^2} \right] + \xi_z \frac{1}{\omega^2 - k_x^2 a^2}\left[\frac{\omega^4}{a^2} - 2\omega^2 k_x^2 + k_x^4 a^2 - k_x^2 g \frac{d\ln p}{dz}\left(1-\frac{1}{\gamma}\right) + g\frac{k_x^4 a^2}{\omega^2}\frac{d\ln p}{dz} + \frac{k_x^4 g^2}{\omega^2}\right] = 0.}$$

This is UGLY!!! And we can't solve it analytically anyhow. It's just a stratified fluid — why is it so complicated?! The reason is twofold: (1) this equation mixes up buoyancy and sound waves — distinct physical effects; and (2) the sound and buoyancy frequencies are functions of height. Let's fix this by adopting an ordering: let

$$\frac{d\xi_z}{dz} \sim i k_z \xi_z = i k_z H \left(\frac{\xi_z}{H}\right) \gg \frac{\xi_z}{H},$$

where $H \equiv \left|\frac{dz}{d\ln p}\right| \sim \left|\frac{dz}{d\ln \rho}\right|$. In other words, we assume that ξ_z varies on a scale $\ll$ the scale of the background. This is a WKB approach. So...

Let $\epsilon \equiv \frac{1}{k_z H} \ll 1$. Also, $k_x \sim k_z$. Now, we must make a decision about the size of ω, by comparing it with

$\dfrac{a}{H} = \dfrac{\gamma g}{a} = \sqrt{\dfrac{\gamma g}{H}}$. There are two choices of interest:

(i) $\omega \sim a/H$

(ii) $\omega \sim ka \sim \underbrace{(a/H)}_{\epsilon} \gg \dfrac{a}{H}$.

First, write $\dfrac{d\xi_z}{dz} = i k_z \xi_z$ with $k_z H \equiv \dfrac{1}{\epsilon}$;  becomes

$$-k_z^2 \xi_z + i k_z \xi_z \left[\frac{\omega^2 \frac{d\ln p}{dz} - k_x^2 a^2 \frac{d\ln \rho}{dz}}{\omega^2 - k_x^2 a^2} \right]$$

$$+ \frac{\xi_z}{\omega^2 - k_x^2 a^2} \left[\frac{\omega^4}{a^2} - 2\omega^2 k_x^2 + k_x^4 a^2 - k_x^2 g \frac{d\ln p}{dz}\left(1 - \frac{1}{\gamma}\right) + \frac{g k_x^4 a^2}{\omega^2} \frac{d\ln p}{dz} + \frac{k_x^4 g^2}{\omega^2} \right] = 0.$$

Now, (i) $\omega \sim a/H$ gives $k_x^2 a^2 \gg \omega^2$ and so the dominant terms are

$$-k_z^2 \xi_z + \xi_z(-k_x^2) - \xi_z \frac{g k_x^2}{\omega^2} \frac{d\ln p}{dz} - \xi_z \frac{k_x^2 g^2}{a^2 \omega^2} = 0$$

$$\Rightarrow \quad k^2 + \frac{g k_x^2}{\omega^2} \underbrace{\left[\frac{d\ln p}{dz} + \frac{g}{a^2} \right]}_{\frac{d\ln p}{dz} - \frac{1}{\gamma}\frac{d\ln p}{dz} = -\frac{1}{\gamma}\frac{d\ln p\rho^{-\gamma}}{dz}} = 0.$$

$$\Rightarrow \boxed{\begin{aligned} \omega^2 &= \frac{k_x^2}{k^2}\frac{g}{\gamma}\frac{d\ln P\rho^{-\gamma}}{dz} \\ &= -\frac{k_x^2}{k^2}\frac{1}{\gamma\rho}\frac{dP}{dz}\frac{d\ln P\rho^{-\gamma}}{dz} \equiv \frac{k_x^2}{k^2}N^2 \end{aligned}}$$

where N^2 is the square of the Brunt - Väisälä frequency.
If $N^2 > 0$, these are called <u>internal waves</u> or <u>g-modes</u>.

Note that different wavenumbers have different velocities (i.e., dispersion) and that ω depends on the direction of $\vec{k}$: $\frac{\partial \omega}{\partial \vec{k}} = \frac{\omega}{k^2}\frac{k_z}{k_x}\left(k_z\hat{x} - k_x\hat{z}\right)$, so that $\vec{k}\cdot\frac{\partial \omega}{\partial \vec{k}} = 0$.

We'll return to the physical cause of these waves later, after the "Boussinesq approximation" is introduced, but, for now, note that $N^2 < 0$ (i.e., upwardly decreasing entropy) gives instability. Go boil some water and think about it.

(ii) $\omega \sim ka \gg a/H$. This gives the following dominant terms:

$$-k_z^2 + \frac{1}{\omega^2 - k_x^2 a^2}\Big[\underbrace{\frac{\omega^4}{a^2} - 2\omega^2 k_x^2 + k_x^4 a^2}_{(\omega^2 - k_x^2 a^2)^2/a^2}\Big] = 0$$

$$\Rightarrow \boxed{\omega^2 = (k_x^2 + k_z^2)\, a^2}$$ <u>Sound waves!</u>

Now, sound waves are often a nuisance in many calculations. They mostly play a rather boring role, and often serve only to make the algebra more tedious. There is something called the Boussinesq approximation, which rigorously filters out sound waves. Let's see how this works in our convection problem...

Return to ⊛ with $\omega \sim \epsilon k a$:

$$\frac{\delta p}{p} \simeq \frac{\omega^2}{k_x^2 a^2}\, i k_z \xi_z \gamma \quad \text{if } k_z \neq 0;$$

$$\simeq \underbrace{\frac{\omega^2}{k_x^2 a^2}}_{\sim \epsilon^2 \ll 1} \frac{d \ln p}{dz} \xi_z \quad \text{otherwise.}$$

So, perhaps we should have dropped perturbations to the gas pressure at some point. Where? In the momentum equation? Hmm... careful. Consider (B) with $\frac{\delta p}{p} \simeq \frac{\omega^2}{k_x^2 a^2} i k_z \xi_z \gamma$

$\Rightarrow \xi_x = -\xi_z \frac{k_z}{k_x}$, or $\vec{k} \cdot \vec{\xi} = 0$. Looks like pressure fluctuations are enforcing (near) incompressibility. Best not to drop them! And (C)?

$$-\omega^2 \xi_z = -i k_z \frac{p}{\rho} \frac{\omega^2}{k_x^2 a^2} i k_z \xi_z \gamma - \frac{\delta \rho}{\rho} g$$

same order!

Okay. So, pressure fluctuations are small, but not so small that they can be dropped from the momentum eqn. What about the entropy eqn?

$$\text{(D)} \Longrightarrow \underbrace{\frac{\delta P}{P}}_{\sim \frac{\xi_z}{H}\frac{\omega^2}{k_x^2 a^2}} = \gamma \frac{\delta\rho}{\rho} - \underbrace{\xi_z}_{\sim \frac{\xi_z}{H},\ \text{required for internal waves}} \frac{d\ln P\rho^{-\gamma}}{dz}$$

or $\sim i k_z \xi_z \gamma \frac{\omega^2}{k_x^2 a^2}$, either way... it's small. So, drop δP from entropy equation! What does that leave us with?

$$\gamma \frac{\delta\rho}{\rho} \simeq \xi_z \frac{d\ln P\rho^{-\gamma}}{dz}$$

$\Longrightarrow \frac{\delta\rho}{\rho} \sim \frac{\xi_z}{H}$. Ah! Look at (A):

$$\underbrace{\frac{\delta\rho}{\rho}}_{\sim \frac{\xi_z}{H}} + \underbrace{i k_z \xi_z}_{\sim k\xi_z} + i k_x \xi_x + \underbrace{\xi_z \frac{d\ln\rho}{dz}}_{\sim \frac{\xi_z}{H}} = 0$$

So, to leading order, we have $\vec{k}\cdot\vec{\xi} = 0$ — incompressibility! Okay, things are consistent, and we have the Boussinesq approx:

$$\frac{\delta P}{P} \sim \frac{1}{kH}\frac{\delta p}{p} \ll \frac{\delta \rho}{\rho} \sim \frac{\delta u}{a} \ll \frac{k\,\delta u}{\omega} \sim k\xi \sim (kH)\frac{\delta u}{a}.$$

Or, defining the Mach number $\mathcal{M}$ and taking it to be small ($\sim \epsilon$),

$$\frac{\delta u}{a} \sim \frac{\delta \rho}{\rho} \sim \frac{\delta T}{T} \sim \frac{1}{\mathcal{M}}\frac{\delta p}{p} \sim \frac{1}{kH} \sim \epsilon \ll 1.$$

In practice, this means:

1) Assume (near) incompressibility ($\vec{\nabla}\cdot\vec{\delta u} = 0$)

2) Drop δp everywhere EXCEPT the momentum eqn. They are enforcing (near) incompressibility.

3) Keep $\delta\rho$ everywhere EXCEPT the continuity eqn. They interact with gravity to give buoyancy.

Watch how much simpler this is...

Ⓐ $\longrightarrow \quad i k_x \xi_x + i k_z \xi_z = 0$

Ⓑ $\longrightarrow \quad -\omega^2 \xi_x = -i k_x \dfrac{\delta p}{\rho}$

$$\longrightarrow \quad \frac{\delta p}{\rho} = \frac{i\omega^2 k_z}{k_x^2}\xi_z$$

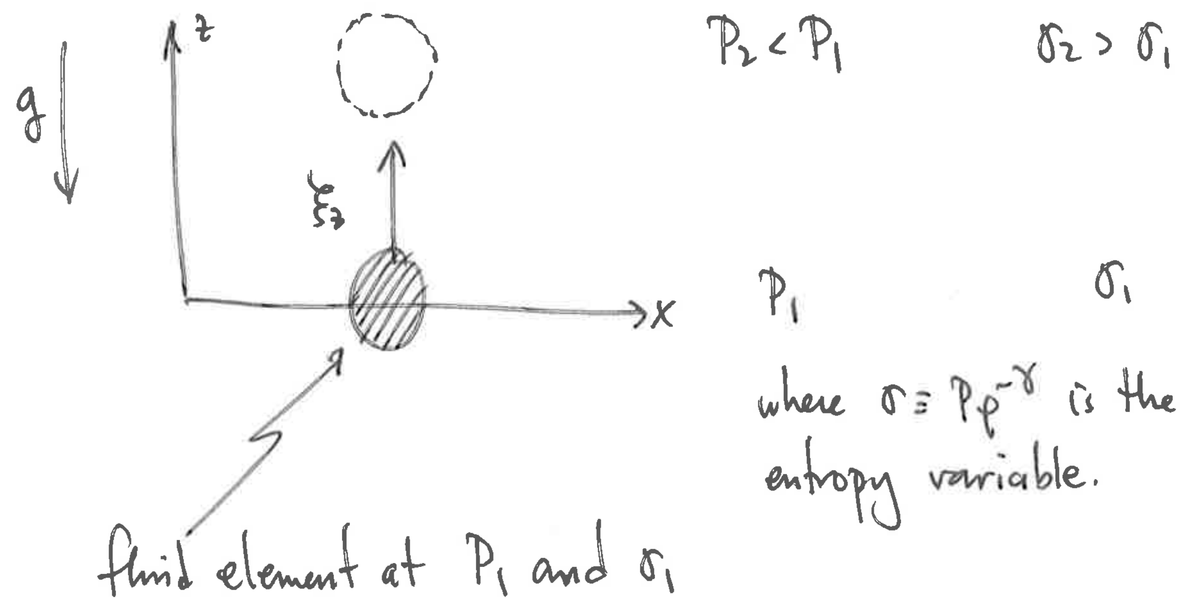
Ⓒ $\longrightarrow \quad -\omega^2 \xi_z = -i k_z \dfrac{\delta p}{\rho} - \dfrac{\delta\rho}{\rho} g$

$$\longrightarrow \quad \omega^2 \xi_z = \frac{k_x^2}{k^2} g \frac{\delta \rho}{\rho}$$

Ⓓ $\longrightarrow \quad 0 = \gamma \dfrac{\delta \rho}{\rho} - \xi_z \dfrac{d\ln p\rho^{-\gamma}}{dz}$

$$\omega^2 = \frac{k_x^2}{k^2} N^2. \quad \text{Done!}$$

What we've done here is eliminated the restoring pressure forces that drive sound waves, essentially by assuming that a^2 is so large that sound waves propagate instantaneously. When the restoring force is purely external (e.g., gravity), the flow behaves as though it were incompressible (nearly). Physically, a slow-moving fluid element remains in pressure balance with its surroundings. This readjustment is what makes buoyancy waves and convection possible. Let us see that explicitly.



where $\sigma \equiv P\rho^{-\gamma}$ is the entropy variable.

Displace fluid element upwards while conserving its entropy. Now it has less entropy than its surroundings. With pressure balance holding, this means that it is also denser than its surroundings. It must fall back to its equilibrium position. Overshooting, it will oscillate at frequency N. (Mathematically, $\Delta\sigma = 0 \Longrightarrow \Delta P/P = \gamma\,\Delta\rho/\rho \Rightarrow \vec{\xi}\cdot\vec{\nabla}\ln P = \gamma\frac{\delta\rho}{\rho} + \gamma\vec{\xi}\cdot\vec{\nabla}\ln\rho \Rightarrow \frac{\delta\rho}{\rho} = \frac{N^2}{g}\xi_z$.)

Now, consider $\sigma_2 < \sigma_1$. Our upwardly displaced fluid element has <u>more</u> entropy than its surroundings, and it will continue to rise $\longrightarrow$ convective instability. The (Karl) <u>Schwarzschild criterion</u> for convective stability is $\boxed{N^2 > 0}$.

Bonus: Exact solution to  for an isothermal atmosphere.

Suppose $\dfrac{d\ln p}{dz} = \dfrac{d\ln \rho}{dz}$ $(T = \text{const})$. Then $\dfrac{d\ln p}{dz} = -\dfrac{\gamma g}{a^2} = \text{const}$.

$\Rightarrow p = p_0 \exp(-z/H)$ with $H \equiv a^2/\gamma g$. Then we have

$$\xi_z'' - \frac{\xi_z'}{H} + \left[\frac{\omega^2 - k_x^2 a^2}{a^2} + \frac{k_x^2 g}{H}\left(1 - \frac{1}{\gamma}\right)\frac{1}{\omega^2}\right]\xi_z = 0.$$

let $\xi_z = f(z)\exp\left(\dfrac{z}{2H}\right)$. Then

$$\xi_z' = f' e^{z/2H} + \frac{f}{2H} e^{z/2H} = f' e^{z/2H} + \xi_z/2H$$

$$\xi_z'' = f'' e^{z/2H} + \frac{f'}{H} e^{z/2H} + \frac{\xi_z}{(2H)^2}$$

$$f'' + \frac{f'}{H} + \frac{f}{(2H)^2} - \frac{f'}{H} - \frac{f}{2H^2} + [\ldots]f = 0.$$

$$f'' + \underbrace{\left[-\frac{1}{4H^2} + \frac{\omega^2 - k_x^2 a^2}{a^2} + \frac{k_x^2 a^2}{\gamma H^2}\left(1 - \frac{1}{\gamma}\right)\frac{1}{\omega^2}\right]}_{= \text{const.}} f = 0.$$

$\Rightarrow f = \exp(\pm i k_z z)$ with $k_z^2 =$ bracket

$$\Rightarrow -k_z^2 - \frac{1}{4H^2} + \frac{\omega^2 - k_x^2 a^2}{a^2} + \frac{k_x^2 a^2}{\omega^2 H^2}\left(\frac{\gamma-1}{\gamma^2}\right) = 0.$$

Mult. by $\omega^2 a^2$ and regroup terms:

$$\omega^4 + \omega^2\left[-k_z^2 a^2 - \frac{a^2}{4H^2} - k_x^2 a^2\right] + k_x^2 a^2\left(\frac{a^2}{H^2}\right)\left(\frac{\gamma-1}{\gamma^2}\right) = 0.$$

$$\Rightarrow \omega^2 = \frac{k^2 a^2 + \frac{a^2}{4H^2}}{2} \pm \frac{1}{2}\left[\left(k^2 a^2 + \frac{a^2}{4H^2}\right)^2 - 4k_x^2 a^2\left(\frac{a^2}{H^2}\right)\left(\frac{\gamma-1}{\gamma^2}\right)\right]^{1/2}.$$

Note that $N^2 \equiv \frac{g}{\gamma}\frac{d\ln P\rho^{-\gamma}}{dz} = \left(\frac{1-\gamma}{\gamma}\right) g\frac{d\ln p}{dz} = \frac{a^2}{H^2}\left(\frac{\gamma-1}{\gamma^2}\right).$

So,

$$\boxed{\omega^2 = \frac{k^2 a^2 + \frac{\gamma^2}{\gamma-1}N^2}{2} \pm \frac{1}{2}\left[\left(k^2 a^2 + \frac{\gamma^2 N^2}{\gamma-1}\right)^2 - 4k_x^2 a^2 N^2\right]^{1/2}}$$

If $(kH)^2 \gg 1$, this becomes $\omega^2 = k^2 a^2$ for the sound wave (plus sign) and $\omega^2 = \frac{k_x^2}{k^2}N^2$ for the g-mode (minus sign). The final term in the square root captures the coupling between these modes.

VI.8. Buoyancy: Parker instability

Continued on hand-written notes...

- A related problem is the Parker Instability, or "magnetic Rayleigh-Taylor instability" (although it is different in detail from RTI and is closer to Schwarzschild Convection). Consider an atmosphere similar to that in our convective instability calculation, but with a magnetic field oriented perpendicularly to gravity with a z-dependence: $\vec{B}_0 = B_0(z)\hat{x}$. The force balance in the equilibrium state now includes a contribution from the magnetic pressure: $g = -\frac{1}{\rho}\left(\frac{dp}{dz} + \frac{dB_0^2}{dz}\right) = \text{const.}$, or

$$\frac{g}{a^2} = -\frac{1}{\gamma}\frac{d\ln P}{dz} - \frac{V_{A0}^2}{a^2}\frac{d\ln B_0}{dz}.$$

Our equations are almost the same:

$$\frac{\delta\rho}{\rho} + ik_x\xi_x + \xi_z' + \xi_z\frac{d\ln\rho}{dz} = 0,$$

$$-\omega^2\xi_x = -ik_x\left(\frac{\delta p}{\rho} + \frac{B_0\,\delta B_x}{4\pi\rho}\right) + \frac{ik_xB_0}{4\pi\rho}\delta B_x + \frac{\delta B_z}{4\pi\rho}\frac{dB_0}{dz},$$

$$-\omega^2\xi_z = -\frac{1}{\rho}\frac{d}{dz}\left(\delta p + \frac{B_0\,\delta B_x}{4\pi}\right) - \frac{\delta\rho}{\rho}g + \frac{ik_xB_0}{4\pi\rho}\delta B_z,$$

$$\frac{\delta p}{p} = \gamma\frac{\delta\rho}{\rho} - \xi_z\frac{d\ln p\rho^{-\gamma}}{dz},$$

but now with magnetic-field perturbations and gradients. The former are given by $\vec{\delta B} = \vec{\nabla}\times(\vec{\xi}\times\vec{B}_0) \Rightarrow \delta B_x = -\frac{d}{dz}(\xi_z B_0)$

$$\delta B_z = ik_xB_0\xi_z$$

First, note that $\dfrac{\delta p + B_0 \delta B_x/4\pi}{\rho}$

$$= a^2 \frac{\delta\rho}{\rho} - \frac{a^2}{\gamma}\frac{d\ln p\rho^{-\gamma}}{dz}\xi_z + \frac{B_0}{4\pi\rho}\left(-\frac{d}{dz}\right)(\xi_z B_0)$$

$$= a^2\left[-ik_x\xi_x - \xi_z' - \xi_z\frac{d\ln\rho}{dz} - \frac{1}{\gamma}\frac{d\ln p\rho^{-\gamma}}{dz}\xi_z - \frac{V_{A0}^2}{a^2}\left(\xi_z' + \xi_z\frac{d\ln B_0}{dz}\right)\right]$$

$$= a^2\left[-ik_x\xi_x - \xi_z'\left(1+\frac{V_{A0}^2}{a^2}\right) + \frac{g}{a^2}\xi_z\right].$$

Then

$$-\omega^2\xi_x = -ik_x a^2\left[-ik_x\xi_x - \xi_z'\left(1+\frac{V_{A0}^2}{a^2}\right) + \frac{g}{a^2}\xi_z\right] + \frac{ik_xB_0}{4\pi\rho}\left[-\xi_z' B_0 - \xi_z B_0\frac{d\ln B_0}{dz}\right] + \frac{ik_xB_0}{4\pi\rho}\xi_z\frac{dB_0}{dz}$$

$$\Rightarrow (-\omega^2 + k_x^2a^2)\xi_x = ik_xa^2\xi_z'\left(1+\cancel{\frac{V_{A0}^2}{a^2}}\right) - ik_xg\xi_z - \cancel{ik_xV_{A0}^2\xi_z'} - \cancel{ik_xV_{A0}^2\frac{d\ln B_0}{dz}\xi_z} + \cancel{ik_xV_{A0}^2\frac{d\ln B_0}{dz}\xi_z}$$

$$\Rightarrow \boxed{(-\omega^2 + k_x^2a^2)\xi_x = ik_xa^2\xi_z' - ik_x\xi_zg}$$

Same as $\#$ w/o B field!

$$\Rightarrow \frac{\delta\rho}{\rho} = -\xi_z' - \xi_z\frac{d\ln\rho}{dz} - ik_x\frac{[ik_xa^2\xi_z' - ik_x\xi_zg]}{(-\omega^2 + k_x^2a^2)}$$

$$\boxed{\frac{\delta\rho}{\rho} = \frac{\omega^2 \xi_z' + \left[(\omega^2 - k_x^2 a^2)\frac{d\ln\rho}{dz} - k_x^2 g\right]\xi_z}{k_x^2 a^2 - \omega^2}}$$

same as ⊚ w/o B field!

Buoyancy is fundamentally the same as in the hydro case. Plugging all of this into the z-component of the momentum eqn. gives $\left(\text{with } \frac{\delta p + \delta B^2/8\pi}{p} = \frac{\omega^2}{\omega^2 - k_x^2 a^2}\left[\frac{\gamma g}{a^2}\xi_z - \gamma\xi_z'\right] - \frac{\gamma v_{A0}^2}{a^2}\xi_z'\right)$

$$-\omega^2 \xi_z = -\frac{1}{\rho}\frac{d}{dz}\left[\frac{p\,\omega^2}{\omega^2 - k_x^2 a^2}\left(\frac{\gamma g}{a^2}\xi_z - \gamma\xi_z'\right) - \frac{B_0^2}{4\pi}\xi_z'\right]$$

$$- g\left\{\frac{\omega^2 \xi_z' + \left[(\omega^2 - k_x^2 a^2)\frac{d\ln\rho}{dz} - k_x^2 g\right]\xi_z}{k_x^2 a^2 - \omega^2}\right\}$$

$$+ \frac{ik_x B_0}{4\pi\rho}\left(ik_x B_0 \xi_z\right)$$

$$\Rightarrow \left(-\omega^2 + k_x^2 v_{A0}^2\right)\xi_z = -\frac{a^2}{\gamma}\frac{d\ln p}{dz}\frac{\omega^2}{\omega^2 - k_x^2 a^2}\left(\frac{\gamma g\,\xi_z}{a^2} - \gamma\xi_z'\right)$$

$$- \frac{a^2}{\gamma}\frac{\omega^2 k_x^2 a^2}{(\omega^2 - k_x^2 a^2)^2}\frac{d\ln T}{dz}\left(\frac{\gamma g\,\xi_z}{a^2} - \gamma\xi_z'\right) + v_{A0}^2\xi_z'' + \xi_z' v_{A0}^2\frac{d\ln B_0^2}{dz}$$

$$- \frac{a^2}{\gamma}\frac{\omega^2}{\omega^2 - k_x^2 a^2}\left(\cancel{\frac{\gamma g}{a^2}\xi_z'} - \frac{\gamma g}{a^2}\xi_z\frac{d\ln T}{dz} - \gamma\xi_z''\right)$$

$$+ \frac{g}{\omega^2 - k_x^2 a^2}\left\{\cancel{\omega^2\xi_z'} + \left[(\omega^2 - k_x^2 a^2)\frac{d\ln\rho}{dz} - k_x^2 g\right]\xi_z\right\}$$

After some straightforward algebra, we find

$$\xi_z'' \left(v_{A0}^2 + \frac{a^2\omega^2}{\omega^2 - k_x^2 a^2} \right)$$

$$+ \xi_z' \left[v_{A0}^2 \frac{d\ln B_0^2}{dz} + \frac{a^2\omega^2}{(\omega^2 - k_x^2 a^2)^2} \left(\omega^2 \frac{d\ln p}{dz} - k_x^2 a^2 \frac{d\ln \rho}{dz} \right) \right]$$

$$+ \xi_z \left[\frac{g}{\gamma} \frac{k_x^2 a^2}{\omega^2 - k_x^2 a^2} \left(\frac{d\ln p\rho^{-\gamma}}{dz} + \frac{\gamma v_{A0}^2}{a^2} \frac{d\ln B_0}{dz} \right) - \frac{g\,\omega^2 k_x^2 a^2}{(\omega^2 - k_x^2 a^2)^2} \frac{d\ln T}{dz} + \omega^2 - k_x^2 v_{A0}^2 \right]$$

All extra terms are $\propto v_{A0}^2$. For $k_z \to 0$, $k_x^2 a^2 \gg 1$, this becomes

$$\boxed{\omega^2 \simeq \underbrace{k_x^2 v_{A0}^2}_{\text{magnetic tension}} + \frac{g}{\gamma}\left(\underbrace{\frac{d\ln p\rho^{-\gamma}}{dz}}_{\text{thermal buoyancy}} + \underbrace{\frac{1}{\beta}\frac{d\ln B_0^2}{dz}}_{\text{magnetic buoyancy}} \right)} \quad \text{w/} \quad \beta = \frac{B_0^2}{8\pi p}.$$

Now, $\frac{d\ln p\rho^{-\gamma}}{dz} + \frac{1}{\beta}\frac{d\ln B_0^2}{dz} > 0$ for stability. But the physics is the same as in Schwarzschild convection — just the pressure balance is different.

VI.9. Rotation

In §II.5, we wrote down the equations of hydrodynamics in a rotating frame – see (II.44). Here we do the same for the equations of MHD. With $\boldsymbol{v} = \boldsymbol{u} - R\mathit{\Omega}(R,z)\hat{\boldsymbol{\varphi}}$ and

$$\frac{\mathrm{D}}{\mathrm{D}t} \doteq \frac{\partial}{\partial t} + \boldsymbol{v}\boldsymbol{\cdot}\boldsymbol{\nabla} + \mathit{\Omega}\frac{\partial}{\partial \varphi},$$

the continuity and force equations are the same,

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} = -\rho\boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{v}, \tag{VI.43}$$

$$\frac{\mathrm{D}v_R}{\mathrm{D}t} = f_R + 2\mathit{\Omega}v_\varphi + R\mathit{\Omega}^2 + \frac{v_\varphi^2}{R}, \tag{VI.44}$$

$$\frac{\mathrm{D}v_\varphi}{\mathrm{D}t} = f_\varphi - \frac{\kappa^2}{2\mathit{\Omega}}v_R - R\frac{\partial\mathit{\Omega}}{\partial z}v_z - \frac{v_R v_\varphi}{R}, \tag{VI.45}$$

$$\frac{\mathrm{D}v_z}{\mathrm{D}t} = f_z, \tag{VI.46}$$

but with the addition of the Lorentz force:

$$\boldsymbol{f} = -\frac{1}{\rho}\boldsymbol{\nabla}\left(P + \frac{B^2}{8\pi}\right) + \frac{\boldsymbol{B}\boldsymbol{\cdot}\boldsymbol{\nabla}B_i}{4\pi\rho}\hat{\boldsymbol{e}}_i + \frac{B_R B_\varphi}{4\pi\rho R}\hat{\boldsymbol{\varphi}} - \frac{B_\varphi^2}{4\pi\rho R}\hat{\boldsymbol{R}} - \boldsymbol{\nabla}\mathit{\Phi}. \tag{VI.47}$$

Note the additional geometric terms $\propto B^2/R$; these are tension forces associated with the bend in the magnetic-field lines as they follow the azimuthal direction. To these equations we must append the induction equation:

$$\frac{\mathrm{D}B_R}{\mathrm{D}t} = -B_R\boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{v} + \boldsymbol{B}\boldsymbol{\cdot}\boldsymbol{\nabla}v_R, \tag{VI.48}$$

$$\frac{\mathrm{D}B_\varphi}{\mathrm{D}t} = -B_\varphi\boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{v} + \boldsymbol{B}\boldsymbol{\cdot}\boldsymbol{\nabla}v_\varphi + \frac{\partial\mathit{\Omega}}{\partial\ln R}B_R + R\frac{\partial\mathit{\Omega}}{\partial z}B_z, \tag{VI.49}$$

$$\frac{\mathrm{D}B_z}{\mathrm{D}t} = -B_z\boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{v} + \boldsymbol{B}\boldsymbol{\cdot}\boldsymbol{\nabla}v_z. \tag{VI.50}$$

With the exception of advection by the differential rotation, the only additions to the induction equation beyond its more customary Cartesian form appear in its azimuthal component: $+R\boldsymbol{B}\boldsymbol{\cdot}\boldsymbol{\nabla}\mathit{\Omega}$ on the right-hand side. This corresponds to stretching of the flux-frozen magnetic field by the differential rotation.

In the hand-written pages that follow, these equations are used to describe the evolution of small fluctuations about a homogeneous, differentially rotating disk with $\mathit{\Omega} = \mathit{\Omega}(R)$, in which the centrifugal acceleration $R\mathit{\Omega}^2$ is balanced by gravity $-\partial\mathit{\Phi}/\partial R$. If the latter is dominated by that of a central point mass M, we have $\mathit{\Phi} = -GM/R$ and so $\mathit{\Omega} = (GM/R^3)^{1/2}$ – i.e., Keplerian rotation.

Before proceeding, I'll write down the linearized MHD equations written in cylindrical coordinates (R,φ,z) in a rotating frame with $\boldsymbol{\mathit{\Omega}} = \mathit{\Omega}(R,z)\hat{\boldsymbol{z}}$. The only assumptions here are that the background magnetic field is uniform, and that the equilibrium state arises from a balance between the centrifugal force and gravity plus thermal-pressure gradients (i.e., we allow for density and pressure stratification in the background state). We also neglect curvature terms of order $\sim(v_\mathrm{A}^2/R)(\delta B/B)$, as these are small compared to the other terms unless the toroidal magnetic field is super-thermal by a factor $\sim(R/H)^{1/2}$, where $H \sim c_\mathrm{s}/\mathit{\Omega}$ is the disk thickness and c_s is the sound speed – an atypical situation.

Without further ado…

$$\left(\frac{\partial}{\partial t}+\mathit{\Omega}\frac{\partial}{\partial\varphi}\right)\delta\rho=-(\delta\boldsymbol{v}\boldsymbol{\cdot}\boldsymbol{\nabla})\rho-\rho(\boldsymbol{\nabla}\boldsymbol{\cdot}\delta\boldsymbol{v}),\tag{VI.51}$$

$$\left(\frac{\partial}{\partial t}+\mathit{\Omega}\frac{\partial}{\partial\varphi}\right)\delta v_R=-\frac{1}{\rho}\frac{\partial}{\partial R}\left(\delta P+\frac{\boldsymbol{B}\boldsymbol{\cdot}\delta\boldsymbol{B}}{4\pi}\right)+\frac{\delta\rho}{\rho^2}\frac{\partial P}{\partial R}+\frac{(\boldsymbol{B}\boldsymbol{\cdot}\boldsymbol{\nabla})\delta B_R}{4\pi\rho}-\frac{\partial\delta\mathit{\Phi}}{\partial R}-2\mathit{\Omega}\delta v_\varphi,\tag{VI.52}$$

$$\left(\frac{\partial}{\partial t}+\mathit{\Omega}\frac{\partial}{\partial\varphi}\right)\delta v_\varphi=-\frac{1}{\rho R}\frac{\partial}{\partial\varphi}\left(\delta P+\frac{\boldsymbol{B}\boldsymbol{\cdot}\delta\boldsymbol{B}}{4\pi}\right)+\frac{\delta\rho}{\rho}\frac{1}{\rho R}\frac{\partial P}{\partial\varphi}+\frac{(\boldsymbol{B}\boldsymbol{\cdot}\boldsymbol{\nabla})\delta B_\varphi}{4\pi\rho}-\frac{1}{R}\frac{\partial\delta\mathit{\Phi}}{\partial\varphi}+\frac{\kappa^2}{2\mathit{\Omega}}\delta v_R+R\frac{\partial\mathit{\Omega}}{\partial z}\delta v_\varphi,\tag{VI.53}$$

$$\left(\frac{\partial}{\partial t}+\mathit{\Omega}\frac{\partial}{\partial\varphi}\right)\delta v_z=-\frac{1}{\rho}\frac{\partial}{\partial z}\left(\delta P+\frac{\boldsymbol{B}\boldsymbol{\cdot}\delta\boldsymbol{B}}{4\pi}\right)+\frac{\delta\rho}{\rho^2}\frac{\partial P}{\partial z}+\frac{(\boldsymbol{B}\boldsymbol{\cdot}\boldsymbol{\nabla})\delta B_z}{4\pi\rho}-\frac{\partial\delta\mathit{\Phi}}{\partial z}\tag{VI.54}$$

$$\left(\frac{\partial}{\partial t}+\mathit{\Omega}\frac{\partial}{\partial\varphi}\right)\delta B_R=(\boldsymbol{B}\boldsymbol{\cdot}\boldsymbol{\nabla})\delta v_R-B_R(\boldsymbol{\nabla}\boldsymbol{\cdot}\delta\boldsymbol{v}),\tag{VI.55}$$

$$\left(\frac{\partial}{\partial t}+\mathit{\Omega}\frac{\partial}{\partial\varphi}\right)\delta B_\varphi=(\boldsymbol{B}\boldsymbol{\cdot}\boldsymbol{\nabla})\delta v_\varphi-B_\varphi(\boldsymbol{\nabla}\boldsymbol{\cdot}\delta\boldsymbol{v})+\frac{\partial\mathit{\Omega}}{\partial\ln R}\delta B_R+R\frac{\partial\mathit{\Omega}}{\partial z}\delta B_z,\tag{VI.56}$$

$$\left(\frac{\partial}{\partial t}+\mathit{\Omega}\frac{\partial}{\partial\varphi}\right)\delta B_z=(\boldsymbol{B}\boldsymbol{\cdot}\boldsymbol{\nabla})\delta v_z-B_z(\boldsymbol{\nabla}\boldsymbol{\cdot}\delta\boldsymbol{v}),\tag{VI.57}$$

$$\left(\frac{\partial}{\partial t}+\mathit{\Omega}\frac{\partial}{\partial\varphi}\right)\delta\sigma=-\delta v_R\frac{\partial\ln P\rho^{-\gamma}}{\partial R}-\delta v_z\frac{\partial\ln P\rho^{-\gamma}}{\partial z},\tag{VI.58}$$

where $\delta\sigma\doteq\delta P/P-\gamma\delta\rho/\rho$.

• <u>Rotational</u> and <u>magnetorotational instability</u>.

Accretion disks are ubiquitous in astrophysics, and they get their namesake by actually facilitating mass accretion onto compact objects like young protostars, neutron stars, black holes, etc. But for this to happen, angular momentum must be redistributed, and it turns out that this is frustratingly difficult in Keplerian disks. The problem is that, hydrodynamically, Keplerian flows are quite stable (we'll show below that they are linearly stable; there is no proof that they are nonlinearly stable, but experimental efforts to find nonlinear instability in hydrodynamic, differentially rotating flows have so far failed). Fluid elements do not like to give up their angular momentum. The culprit is the Coriolis force, a surprisingly strong stabilizing effect. (Indeed, planar shear flows without rotation quite easily disrupt so long as the viscosity is not too large.) Another issue is that the molecular viscosity, which might transport angular momentum purely by frictional means, is absolutely negligible in most all astrophysical fluids. One way out is to posit some anomalous viscosity via (unknown) turbulence. This is the route taken in the classic Shakura & Sunyaev

(1973) paper — assume turbulent transport, characterize it by a scalar viscosity, and take that viscous stress to be proportional to the gas pressure:

$$T_{R\varphi} = \alpha_{ss}\, p$$

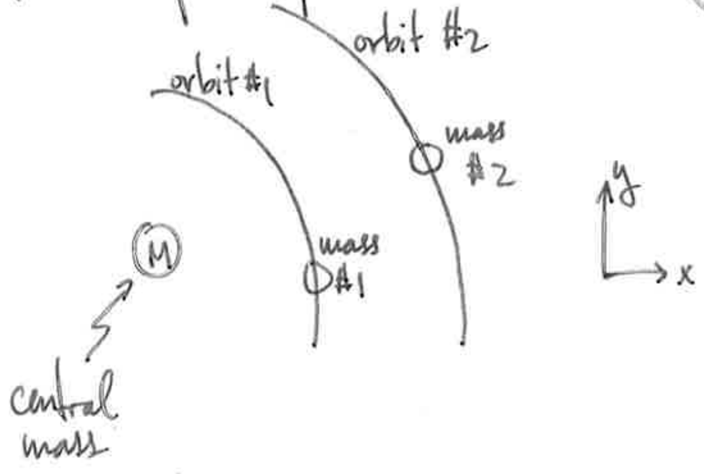
gas pressure

proportionality constant

R-φ component of the stress tensor, responsible for transporting φ momentum in the R direction.

This led to the "α-disk" framework of accretion disks, which has been extremely profitable, but woefully unsatisfying. This changed in 1991.

Let's pause here and explore the above claims a bit further. I said a Keplerian disk is hydrodynamically stable to small disturbances. Let's prove it. There are two ways to do this — using point masses in orbits, and using the full hydro eqns. in a rotating frame. Here's the first:



The eqns. of motion for these masses are

$$\ddot{x} - 2\Omega\dot{y} = -\frac{d\Omega^2}{d\ln R}\,x$$

$$\ddot{y} + 2\Omega\dot{x} = 0$$

These are called the "Hill equations" (Hill 1878).
They include the Coriolis force and an extra term in the "radial" equation for the x displacement that accounts for the "tidal" force (the local difference between the centrifugal force and gravity). And they are local — note the Cartesian coordinate system with x pointing locally radial and y pointing locally azimuthal. Take solutions $x, y \sim e^{i\omega t}$ to compute the normal modes of this system:

$$\begin{bmatrix} -\omega^2 + \frac{d\Omega^2}{d\ln R} & 2\Omega i\omega \\ -2\Omega i\omega & -\omega^2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0 \Rightarrow \omega^2\left(\omega^2 - \frac{d\Omega^2}{d\ln R}\right) = 4\Omega^2\omega^2$$

$$\Rightarrow \omega^2 - \underbrace{\left(4\Omega^2 + \frac{d\Omega^2}{d\ln R}\right)}_{\equiv \kappa^2} = 0 \Rightarrow \boxed{\omega = \pm\kappa}$$

κ^2 ← "epicyclic frequency"

These are epicyclic oscillations when $\kappa^2 > 0$, and exponentially growing disturbances when $\kappa^2 < 0$.

Note that $4\Omega^2 + \frac{d\Omega^2}{d\ln R} = \frac{1}{R^3}\frac{d\ell^2}{dR}$, where $\ell = \Omega R^2$ is the (specific) angular momentum. Thus,

$$\boxed{\frac{d\ell^2}{d\ln R} > 0 \Leftrightarrow \text{linear stability}}$$

"Rayleigh criterion"

The fluid way: Let's assume incompressibility for simplicity. Going back to our hydrodynamic equs., with gravity from a central point mass and $\vec{u} = \vec{v} + R\Omega\hat{\varphi}$, we have

$$\left(\frac{\partial}{\partial t} + \Omega\frac{\partial}{\partial\varphi}\right)v_R + \vec{v}\cdot\vec{\nabla}v_R - 2\Omega v_\varphi - R\Omega^2 - \frac{v_\varphi^2}{R}$$

$$= -\frac{\partial p}{\partial R}\frac{1}{\rho} + g_R,$$

$$\left(\frac{\partial}{\partial t} + \Omega\frac{\partial}{\partial\varphi}\right)v_\varphi + \vec{v}\cdot\vec{\nabla}v_\varphi + 2\Omega v_R + v_R\frac{d\Omega}{d\ln R} + \frac{v_R v_\varphi}{R}$$

$$= -\frac{1}{R}\frac{\partial p}{\partial\varphi}\frac{1}{\rho},$$

where $g_R = -\frac{GM}{R^2}$. Our equilibrium state is $\vec{v}=0$, p = constant, and $g_R = -R\Omega^2 \Rightarrow \Omega^2 = \frac{GM}{R^3}$, a Keplerian orbit. Writing $\vec{v} = 0 + \vec{\delta v}$ and $p = p_0 + \delta p$, our equations to linear order in δ are

$$\left(\frac{\partial}{\partial t} + \Omega\frac{\partial}{\partial\varphi}\right)\delta v_R - 2\Omega\delta v_\varphi = -\frac{1}{\rho}\frac{\partial}{\partial R}\delta p,$$

$$\left(\frac{\partial}{\partial t} + \Omega\frac{\partial}{\partial\varphi}\right)\delta v_\varphi + 2\Omega\delta v_R + \frac{d\Omega}{d\ln R}\delta v_R = -\frac{1}{\rho}\frac{1}{R}\frac{\partial\delta p}{\partial\varphi}.$$

For simplicity, let us neglect the $\partial/\partial\varphi$ derivatives and let $\delta \sim \exp(-i\omega t + i k_R R + i k_z z)$. Then, with

$$\vec{\nabla}\cdot\vec{\delta v} = 0 \Rightarrow k_R \delta v_R = -k_z \delta v_z \quad \text{and} \quad \frac{\partial}{\partial t}\delta v_z = -\frac{1}{\rho}\frac{\partial}{\partial z}\delta p,$$

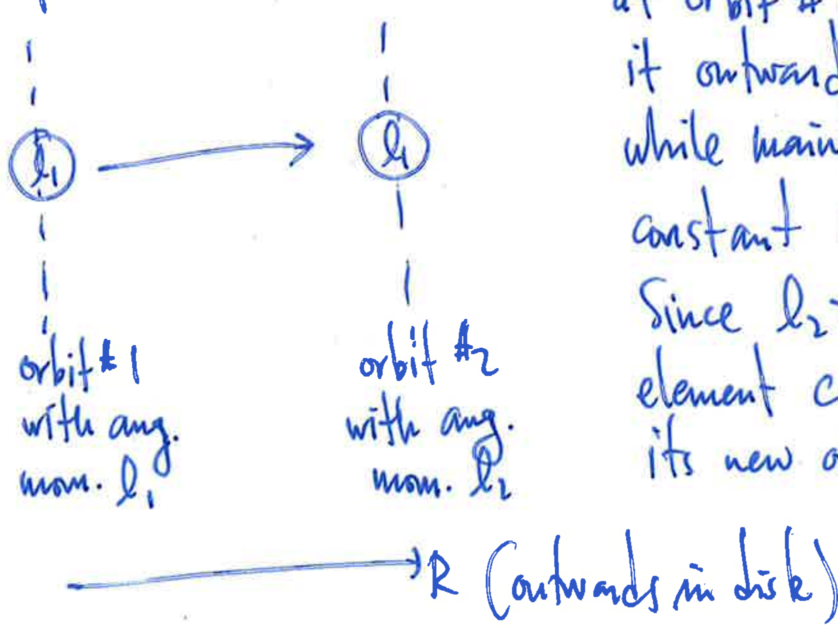
we find $\frac{\delta p}{\rho} = -\frac{k_R}{k_z^2}\omega\, \delta v_R$ and so

$$\begin{bmatrix} -i\omega \frac{k^2}{k_z^2} & -2\Omega \\ \underbrace{\frac{\kappa^2}{2\Omega}}_{=2\Omega + \frac{d\Omega}{d\ln R}} & -i\omega \end{bmatrix} \begin{bmatrix} \delta v_R \\ \delta v_\varphi \end{bmatrix} = 0 \Rightarrow \boxed{\omega^2 = \frac{k_z^2}{k^2}\kappa^2}$$

Same stability criterion as before, $\kappa^2 > 0$.

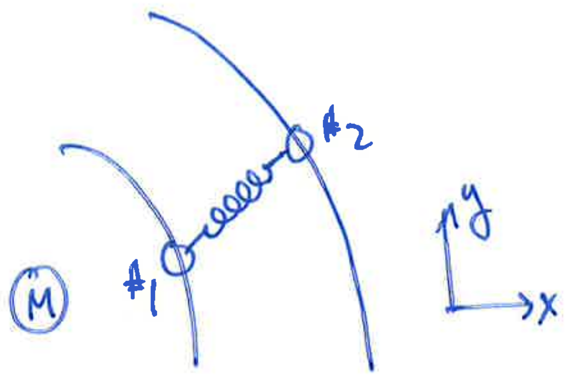
Physically, what's going on? Take a fluid element at orbit #1 and displace it outwards to orbit #2 while maintaining a constant ang. momentum. Since $l_2 > l_1$, the fluid element cannot stay in its new orbit and must return back to orbit #1.

$\Rightarrow$ STABLE.



Now, back to 1991...

Steve Balbus and John Hawley, then both at Univ. of Virginia, found by a straightforward linear analysis and clever use of 90's supercomputers, that a small but finite magnetic field is all that is required to linearly destabilize Keplerian rotation. How could this be missed? The answer is complicated. The instability — at first known as the "Balbus-Hawley instability" but now goes by the moniker "magnetorotational instability" (MRI) — appeared in a little-known Russian paper by Velikhov in 1959, and 2 years later made its way into Chandrasekhar's classic text on "Hydrodynamic and hydromagnetic stability". But there it appeared in a rather odd guise, at least to an astronomer thinking about accretion disks — Couette flow, i.e., rotational flow excited by placing a (conducting) fluid between two cylindrical walls rotating at different speeds. It wasn't until B&H rediscovered it and placed it in the astrophysical context that the instability became appreciated as a possible solution to the accretion problem. What followed was an industry of linear analysis and nonlinear numerical simulations aiming to characterize the MRI in a wide variety of disk systems. But let's go back to the beginning:



Take the Hill system but attach a spring between the two masses — magnetic fields act as springs, so you can imagine this being a field line threading two fluid elements. Add Hooke's law to the eqns. of motion:

$$\ddot{x} - 2\Omega\dot{y} = -\frac{d\Omega^2}{d\ln R}x - Kx$$

w/ K = spring constant

$$\ddot{y} + 2\Omega\dot{x} = -Ky$$

$x, y \sim e^{-i\omega t}$

$$\Rightarrow \begin{bmatrix} -\omega^2 + \frac{d\Omega^2}{d\ln R} + K & 2\Omega i\omega \\ -2\Omega i\omega & -\omega^2 + K \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

$$\Rightarrow \left(\omega^2 - \frac{d\Omega^2}{d\ln R} - K\right)\left(\omega^2 - K\right) = 4\Omega^2\omega^2$$

$$\Rightarrow (\omega^2 - K)\left(\omega^2 - K - 4\Omega^2 - \frac{d\Omega^2}{d\ln R}\right) = 4\Omega^2 K$$

extra extra extra

Solutions are $\omega^2 - K = \frac{K^2}{2} \pm \sqrt{\left(\frac{K^2}{2}\right)^2 + 4\Omega^2 K}$,

whose (−) solution is unstable if $\boxed{K + \frac{d\Omega^2}{d\ln R} < 0}$

This is important, because Keplerian disks have $\frac{d\Omega^2}{d\ln R} < 0$! Note that the spring cannot be too strong here. Interestingly,

$$\boxed{\frac{d l^2}{d\ln R} > 0 \text{ for hydro stability} \longrightarrow \frac{d\Omega^2}{d\ln R} > 0 \text{ for MHD stability}}$$

One can show that the Lagrangian change in the ang. mom. of a fluid element as it is displaced is given by

$$\frac{\Delta l}{l} = \frac{x}{R}\left(\frac{\kappa^2}{2\Omega^2} - \frac{i\omega}{\Omega}\frac{y}{x}\right)$$

$$= -\frac{x}{R}\left(\frac{2K}{\omega^2 - K}\right)$$

the spring broke conservation of angular momentum!

If $K \ll |\omega^2| \sim \Omega^2$, then outward displacements $(x>0)$ gain angular momentum as they are torqued by the spring (NB: $\omega^2 < 0$ corresponds to growth, so $\Delta l \propto (x/R)$).

At max. growth (take $\frac{\partial}{\partial K}$ of disp. relation and find extrema of ω^2), the growth rate $-i\omega = \frac{1}{2}\left|\frac{d\Omega}{d\ln R}\right|$ and

$$\left.\frac{\Delta l}{l}\right|_{max} = \frac{2x}{R}\left(1 - \frac{1}{4}\left|\frac{d\Omega}{d\ln R}\right|\right).$$

Now the fluid picture: (let $\vec{B}_0 = B_0\hat{z}$ for simplicity)

$$\left(\frac{\partial}{\partial t} + \Omega\frac{\partial}{\partial\varphi}\right)\delta v_R - 2\Omega\,\delta v_\varphi = -\frac{1}{\rho}\frac{\partial}{\partial R}\left(\delta p + \frac{\vec{B}_0\cdot\delta\vec{B}}{4\pi}\right)$$
$$+ \frac{\vec{B}_0\cdot\vec{\nabla}\,\delta B_R}{4\pi\rho},$$

$$\left(\frac{\partial}{\partial t} + \Omega\frac{\partial}{\partial\varphi}\right)\delta v_\varphi + \frac{\kappa^2}{2\Omega}\,\delta v_R = -\frac{1}{\rho}\frac{1}{R}\frac{\partial}{\partial\varphi}\left(\delta p + \frac{\vec{B}_0\cdot\delta\vec{B}}{4\pi}\right)$$
$$+ \frac{\vec{B}_0\cdot\vec{\nabla}\,\delta B_\varphi}{4\pi\rho}.$$

We need equs. for δB_R and δB_φ (δB_z is determined from $\vec{\nabla}\cdot\delta\vec{B} = 0$). So, take the induction equ. and write it in cylindrical coordinates w/ $\vec{u} = \vec{v} + R\Omega\hat{\varphi}$:

$$\left(\vec{B}\cdot\vec{\nabla}\vec{u} = \hat{e}_i\,\vec{B}\cdot\vec{\nabla}v_i + \frac{v_R B_\varphi}{R}\hat{\varphi} - \frac{v_\varphi B_\varphi}{R}\hat{R} - \Omega B_\varphi\hat{R} + \hat{\varphi}B_R\left(\Omega + \frac{d\Omega}{d\ln R}\right)\right)$$

$$\Rightarrow \left(\frac{\partial}{\partial t} + \Omega\frac{\partial}{\partial\varphi}\right)\delta B_R = \vec{B}_0\cdot\vec{\nabla}\,\delta v_R,$$

$$\left(\frac{\partial}{\partial t} + \Omega\frac{\partial}{\partial\varphi}\right)\delta B_\varphi = \vec{B}_0\cdot\vec{\nabla}\,\delta v_\varphi + \delta B_R\,\frac{d\Omega}{d\ln R}.$$

Again, take $\partial/\partial\varphi = 0$ and let $\delta \sim \exp(-i\omega t + ik_R R + ik_z z)$.

$$\vec{\nabla}\cdot\vec{\delta v} = 0 \longrightarrow \delta v_z = -\frac{k_R}{k_z}\,\delta v_R$$
$$\vec{\nabla}\cdot\vec{\delta B} = 0 \longrightarrow \delta B_z = -\frac{k_R}{k_z}\,\delta B_R$$

$$\frac{\partial}{\partial t}\delta v_z = -\frac{1}{\rho}\frac{\partial}{\partial z}\left(\delta p + \frac{\vec{B}_0\cdot\vec{\delta B}}{4\pi}\right) + \frac{\vec{B}_0\cdot\vec{\nabla}\,\delta B_z}{4\pi\rho}$$

$$\Rightarrow \frac{\delta p}{\rho} + \frac{\vec{B}_0\cdot\vec{\delta B}}{4\pi\rho} = \frac{-\omega k_R}{k_z^2}\delta v_R - \frac{\vec{k}\cdot\vec{B}_0}{4\pi\rho}\frac{k_R}{k_z^2}\delta B_R$$

$$\Rightarrow \begin{cases} -i\omega\,\delta v_R - 2\Omega\,\delta v_\varphi = -ik_R\left[-\dfrac{\omega k_R}{k_z^2}\delta v_R - \dfrac{\vec{k}\cdot\vec{B}_0}{4\pi\rho}\dfrac{k_R}{k_z^2}\delta B_R\right] + \dfrac{i\vec{k}\cdot\vec{B}_0}{4\pi\rho}\delta B_R \\[2ex] -i\omega\,\delta v_\varphi + \dfrac{\kappa^2}{2\Omega}\delta v_R = \dfrac{i\vec{k}\cdot\vec{B}_0}{4\pi\rho}\delta B_\varphi \end{cases}$$

$$\Rightarrow \begin{cases} -i\omega\dfrac{k^2}{k_z^2}\delta v_R - 2\Omega\,\delta v_\varphi = \dfrac{i\vec{k}\cdot\vec{B}_0}{4\pi\rho}\dfrac{k^2}{k_z^2}\delta B_R \\[2ex] \dfrac{\kappa^2}{2\Omega}\delta v_R - i\omega\,\delta v_\varphi = \dfrac{i\vec{k}\cdot\vec{B}_0}{4\pi\rho}\delta B_\varphi \end{cases}$$

and
$$\begin{cases} -i\omega\,\delta B_R = i\vec{k}\cdot\vec{B}_0\,\delta v_R \\ -i\omega\,\delta B_\varphi = i\vec{k}\cdot\vec{B}_0\,\delta v_\varphi + \delta B_R\dfrac{d\Omega}{d\ln R} \end{cases}$$

solve these for $\vec{\delta v}$ and plug into these

$$
\left[\begin{array}{l}
-i\omega \dfrac{k^2}{k_z^2}\left(\dfrac{-i\omega\,\delta B_R}{i\mathbf{k}\cdot\mathbf{B}_0}\right) - 2\Omega\left(\dfrac{-i\omega\,\delta B_\varphi - \delta B_R\, d\Omega/d\ln R}{i\mathbf{k}\cdot\mathbf{B}_0}\right) = \dfrac{i\mathbf{k}\cdot\mathbf{B}_0}{4\pi\rho}\dfrac{k^2}{k_z^2}\,\delta B_R \\[2ex]
\dfrac{\kappa^2}{2\Omega}\left(\dfrac{-i\omega\,\delta B_R}{i\mathbf{k}\cdot\mathbf{B}_0}\right) - i\omega\left(\dfrac{-i\omega\,\delta B_\varphi - \delta B_R\, d\Omega/d\ln R}{i\mathbf{k}\cdot\mathbf{B}_0}\right) = \dfrac{i\mathbf{k}\cdot\mathbf{B}_0}{4\pi\rho}\,\delta B_\varphi
\end{array}\right.
$$

Cleaning up...

$$
\begin{bmatrix}
-\omega^2 + (\mathbf{k}\cdot\mathbf{v}_A)^2 - \dfrac{k_z^2}{k^2}\dfrac{d\Omega^2}{d\ln R} & 2\Omega i\omega \dfrac{k_z^2}{k^2} \\[2ex]
-2\Omega i\omega & -\omega^2 + (\mathbf{k}\cdot\mathbf{v}_A)^2
\end{bmatrix}
\begin{bmatrix} \delta B_R \\[2ex] \delta B_\varphi \end{bmatrix} = 0.
$$

Look familiar? $K \to (\mathbf{k}\cdot\mathbf{v}_A)^2$! Magnetic tension is a spring. Dispersion relation:

$$
\left[\omega^2 - (\mathbf{k}\cdot\mathbf{v}_A)^2\right]\left[\omega^2 - (\mathbf{k}\cdot\mathbf{v}_A)^2 - \kappa^2\frac{k_z^2}{k^2}\right] = 4\Omega^2(\mathbf{k}\cdot\mathbf{v}_A)^2
$$

$$
\Rightarrow \omega^2 - (\mathbf{k}\cdot\mathbf{v}_A)^2 = \frac{\kappa^2}{2} \pm \sqrt{\left(\frac{\kappa^2}{2}\right)^2 + 4\Omega^2(\mathbf{k}\cdot\mathbf{v}_A)^2}
$$

unstable if $\boxed{(\mathbf{k}\cdot\mathbf{v}_A)^2 + \dfrac{d\Omega^2}{d\ln R} < 0}$

Note: Can write discriminant as $\left[\frac{\kappa^2}{2} + (\mathbf{k}\cdot\mathbf{v}_A)^2\right]^2 - (\mathbf{k}\cdot\mathbf{v}_A)^2\left[(\mathbf{k}\cdot\mathbf{v}_A)^2 + \kappa^2 - 4\Omega^2\right]$

$$
= \left[\frac{\kappa^2}{2} + (\mathbf{k}\cdot\mathbf{v}_A)^2\right]^2 - (\mathbf{k}\cdot\mathbf{v}_A)^2\left[(\mathbf{k}\cdot\mathbf{v}_A)^2 + \frac{d\Omega^2}{d\ln R}\right]
$$

This means Keplerian disks are linearly unstable, provided the magnetic field isn't so strong that all the wavenumbers $k_z = 2\pi/\lambda_z$ that can fit within the height of the disk satisfy $k_z^2 v_{Az}^2 > \left|\frac{d\Omega^2}{d\ln R}\right|$ — then tension stabilizes all relevant modes.

See Balbus & Hawley 1998 Rev. Mod. Phys. for more.

PART VII
Charged particle motion

Covered by Dr. Tolman, but here are some supplementary notes...

So far, we have concerned ourselves with the response of fluid elements to both imposed and self-consistently generated electromagnetic and gravitational fields. But those fluid elements are composed of charged (and neutral) particles; it would be good to know how those particles move through phase space. Now, we all know Newton's equations of motion for a particle in the presence of electric and magnetic fields:

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \boldsymbol{v}, \quad \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = \frac{q}{m}\Big[\boldsymbol{E}(t,\boldsymbol{r}) + \frac{\boldsymbol{v}}{c}\times\boldsymbol{B}(t,\boldsymbol{r})\Big]. \tag{VII.1}$$

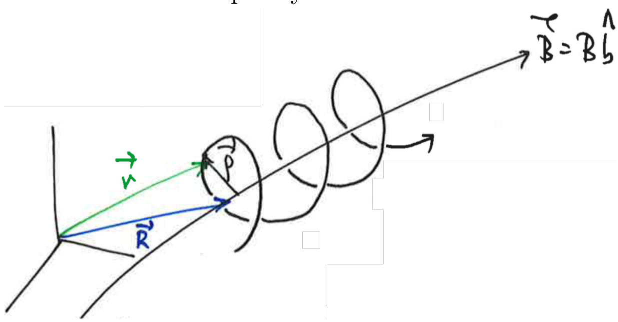
But solutions to (VII.1) are surprisingly subtle, even in seemingly simple situations...

VII.1. Particle motion in uniform electric and magnetic fields

Consider the motion of a single, charged particle. Start by decomposing the particle's position into a Larmor position $\boldsymbol{\rho}$ and a guiding-center position $\boldsymbol{R}$, *viz.*,

$$\boldsymbol{r} = \boldsymbol{\rho} + \boldsymbol{R} = -\frac{\boldsymbol{v}\times\hat{\boldsymbol{b}}}{\Omega} + \boldsymbol{R}, \tag{VII.2}$$

where $\Omega \doteq qB/mc$ is the Larmor frequency:

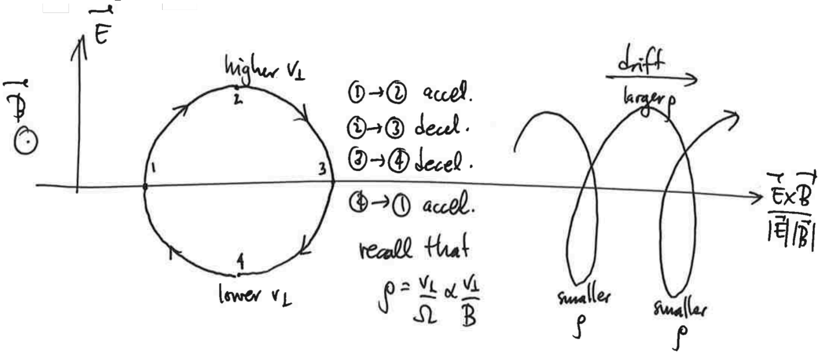


The Larmor position just oscillates around the guiding center at a rate $\dot{\vartheta} \simeq -\Omega$ (more on this later). Using this decomposition, let's begin with something relatively simple: particle motion in constant electric and magnetic fields.

Rearranging (VII.2) and taking the time derivative,

$$\begin{aligned}
\dot{\boldsymbol{R}} &= \dot{\boldsymbol{r}} - \dot{\boldsymbol{\rho}} \\
&= \boldsymbol{v} + \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}\times\frac{\hat{\boldsymbol{b}}}{\Omega} \\
&= \boldsymbol{v} + \frac{q}{m}\Big(\boldsymbol{E} + \frac{\boldsymbol{v}}{c}\times\boldsymbol{B}\Big)\times\frac{\hat{\boldsymbol{b}}}{\Omega} \qquad \text{(using (VII.1))} \\
&= \boldsymbol{v} + \frac{q\boldsymbol{E}\times\hat{\boldsymbol{b}}}{m\Omega} - \boldsymbol{v}_\perp \\
&= v_\parallel\hat{\boldsymbol{b}} + \frac{c\boldsymbol{E}\times\boldsymbol{B}}{B^2} \doteq v_\parallel\hat{\boldsymbol{b}} + \boldsymbol{v}_E \qquad \text{(VII.3)} \\
&= \text{parallel streaming of the guiding center + `E cross B drift'}
\end{aligned}$$

Note that the perpendicular drift is charge independent; ions and electrons drift in the same direction with the same speed. Thus, no currents are generated by this type of guiding-center drift. The physical origin of the $\boldsymbol{E} \times \boldsymbol{B}$ drift is the dependence of the gyroradius of a particle on $v_\perp$, which periodically changes due to acceleration by the perpendicular component of the electric field:



You'll see when we study ideal MHD that particles $\boldsymbol{E} \times \boldsymbol{B}$ drift in order to stay on a given magnetic-field line.

For more a general force $\boldsymbol{F}$, the perpendicular drift is

$$\boldsymbol{v}_F \doteq \frac{\boldsymbol{F} \times \hat{\boldsymbol{b}}}{m\Omega}, \tag{VII.4}$$

which is generally charge dependent and thus results in currents.

VII.2. Particle motion in a non-uniform magnetic field

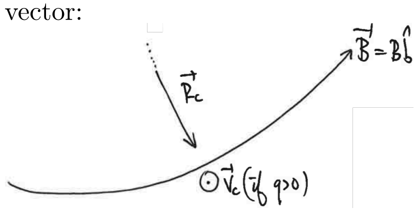
Next, let's keep the uniform electric field, but allow the magnetic field to vary in space. Equation (VII.3) acquires an additional term due to gradients in the magnetic field along the particle orbit:

$$\dot{\boldsymbol{R}} = v_\parallel \hat{\boldsymbol{b}} + \boldsymbol{v}_E + \boldsymbol{v} \times \frac{\mathrm{d}}{\mathrm{d}t} \frac{\hat{\boldsymbol{b}}}{\Omega}. \tag{VII.5}$$

The final term in (VII.5) includes two new drifts, which can be obtained rigorously using 'guiding-center theory' (and we will, in §VII.4). But they can also be obtained quite readily if you already know their names: 'curvature drift' and 'grad-B drift'. The former suggests we look at the centrifugal force on a particle as it follows a curved magnetic-field line:

$$\boldsymbol{F}_\mathrm{c} = \frac{m v_\parallel^2}{r_\mathrm{c}} \hat{\boldsymbol{r}}_\mathrm{c}, \quad \text{where } \hat{\boldsymbol{r}}_\mathrm{c} = -r_\mathrm{c} \hat{\boldsymbol{b}} \cdot \boldsymbol{\nabla} \hat{\boldsymbol{b}} \tag{VII.6}$$

with r_c being the radius of curvature of the field line. The unit vector $\hat{\boldsymbol{r}}_\mathrm{c}$ points in the direction of the curvature vector:



Feeding (VII.6) into (VII.4), we obtain the curvature drift,

$$\boldsymbol{v}_\mathrm{c} \doteq \frac{\boldsymbol{F}_\mathrm{c} \times \hat{\boldsymbol{b}}}{m\Omega} = -\frac{v_\parallel^2}{\Omega} (\hat{\boldsymbol{b}} \cdot \boldsymbol{\nabla} \hat{\boldsymbol{b}}) \times \hat{\boldsymbol{b}}. \tag{VII.7}$$

Note that it is charge dependent.

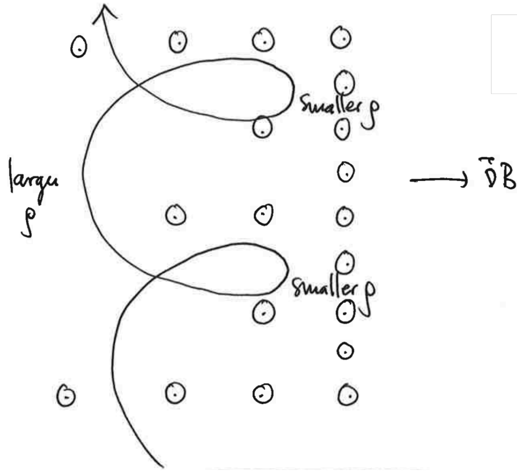
As for the 'grad-B drift', imagine a magnetic dipole with moment

$$\boldsymbol{\mu} = \frac{1}{2} q\boldsymbol{r} \times \frac{\boldsymbol{v}}{c} = -\frac{1}{2} q\rho \frac{v_\perp}{c} \hat{\boldsymbol{b}} = -\frac{mv_\perp^2}{2B} \doteq -\mu\hat{\boldsymbol{b}}, \tag{VII.8}$$

exposed to an inhomogeneous magnetic field. The force on a dipole is equal to $\boldsymbol{\nabla}(\boldsymbol{\mu}\boldsymbol{\cdot}\boldsymbol{B}) = -\mu\boldsymbol{\nabla}B$, and so (using (VII.4)), there is a drift given by

$$\boldsymbol{v}_{\nabla B} \doteq \frac{-\mu\boldsymbol{\nabla}B \times \hat{\boldsymbol{b}}}{m\Omega} = \frac{v_\perp^2}{2\Omega} \hat{\boldsymbol{b}} \times \boldsymbol{\nabla} \ln B. \tag{VII.9}$$

This drift results from the increase (decrease) in the gyroradius of a particle as the particle enters a region of decreased (increased) magnetic-field strength:



The grad-B drift is also charge dependent.

Note that, in a force-free field configuration with $\boldsymbol{\nabla} \times \boldsymbol{B} \parallel \boldsymbol{B}$, we have $\hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla}\hat{\boldsymbol{b}} = \boldsymbol{\nabla}_\perp \ln B$. Thus, from (VII.7) and (VII.9),

$$\boldsymbol{v}_{\rm curv} + \boldsymbol{v}_{\nabla B} = \frac{v_\parallel^2 + v_\perp^2/2}{\Omega} \hat{\boldsymbol{b}} \times \boldsymbol{\nabla} \ln B.$$

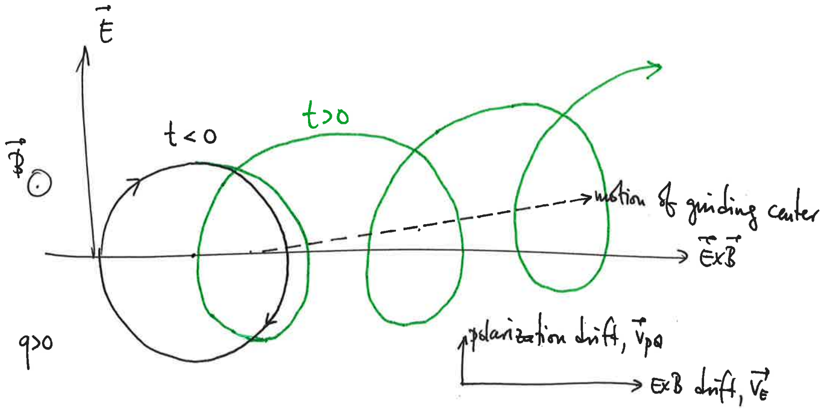
Averaged over all particles, these drifts are $\sim v_{\rm th}(\rho/\ell_B)$, which is typically (very!) sub-thermal.

VII.3. Particle motion in a time-dependent electric field

If $\boldsymbol{E}$ has some explicit time dependence, then there is yet another drift called 'polarization drift', which can be thought of as being due to an inertial force $-m\,{\rm d}\boldsymbol{v}_E/{\rm d}t$ on the guiding center:

$$\boldsymbol{v}_{\rm pol} \doteq -\frac{{\rm d}\boldsymbol{v}_E}{{\rm d}t} \times \frac{\hat{\boldsymbol{b}}}{\Omega} = \frac{1}{\Omega}\frac{c}{B}\frac{\partial \boldsymbol{E}_\perp}{\partial t}. \tag{VII.10}$$

If an electric field is suddenly switched on in a plasma, the ions will drift faster than the electrons (!), thus polarizing the plasma. The idea here is that, if the electric field varies as the particle navigates its gyro-orbit and does not average to zero, the result is a net shift of the guiding center in the direction of $\partial\boldsymbol{E}_\perp/\partial t$ for positive charges and in the opposite direction for negative charges. The simplest way to picture this is to consider switching on a linearly increasing perpendicular electric field at $t = 0$:



Because the ions and electrons have different signs of polarization drift, there is a current produced:

$$\boldsymbol{j}_{\mathrm{pol}} = \rho \left(\frac{c}{B} \right)^2 \frac{\partial \boldsymbol{E}_{\perp}}{\partial t}, \tag{VII.11}$$

where $\rho \doteq m_{\mathrm{i}} n_{\mathrm{i}} + m_{\mathrm{e}} n_{\mathrm{e}}$ is the mass density. This current is dominated by the heavier species (ions), since that species has a larger gyro-period and thus is displaced by a much larger distance by the changing electric field during each orbit. By analogy with standard electrodynamics in dielectric media, in which[9]

$$\boldsymbol{j}_{\mathrm{pol}} = \frac{\varepsilon}{4\pi} \frac{\partial \boldsymbol{E}}{\partial t},$$

we see that the effective permittivity $\varepsilon = (c/v_{\mathrm{A}})^2$, where $v_{\mathrm{A}} \doteq B/\sqrt{4\pi\rho}$ is the Alfvén speed. (Polarization current is tied to the propagation of Alfvén waves.) Since we often have $c/v_{\mathrm{A}} \gg 1$, most plasmas have $\varepsilon \gg 1$, i.e., they behave as strongly polarizable media.

VII.4. Guiding-center theory

This is more advanced material. It is a beautiful exercise in asymptotic expansion, but there isn't enough time in this school for go through the procedure in detail. I encourage to you read it on your own and work through the calculations.

There is a systematic way of deriving drifts that are due to the non-constantly of forces along a particle's orbit, so long as these forces vary slowly. By that, we mean that the length scales (ℓ) and time scales (τ) over which the forces vary are long compared to ρ and Ω^{-1}, respectively:

$$\frac{\rho}{\ell} \ll 1, \quad (\Omega\tau)^{-1} \ll 1.$$

To enact this scale hierarchy, we introduce a small parameter,

$$\epsilon \doteq \frac{\rho}{\ell} \sim (\Omega\tau)^{-1},$$

and expand (VII.1) in powers of ϵ. Not surprisingly, we will find a fast gyromotion and a slow guiding-center motion.

---

[9]Because of the standard undergraduate training in electromagnetism, you may not be familiar with dielectrics in Gaussian units. If that's true, then note the following conventions: $\boldsymbol{D} = \boldsymbol{E} + 4\pi\boldsymbol{P} \doteq \varepsilon\boldsymbol{E} \doteq (1 + 4\pi\chi_{\mathrm{e}})\boldsymbol{E}$.

Start by writing $\boldsymbol{R} \doteq \boldsymbol{r} - \boldsymbol{\rho}$ as before, but now with the Larmor vector defined by

$$\boldsymbol{\rho} = -\frac{(\boldsymbol{v} - \boldsymbol{v}_E) \times \hat{\boldsymbol{b}}}{\Omega}. \tag{VII.12}$$

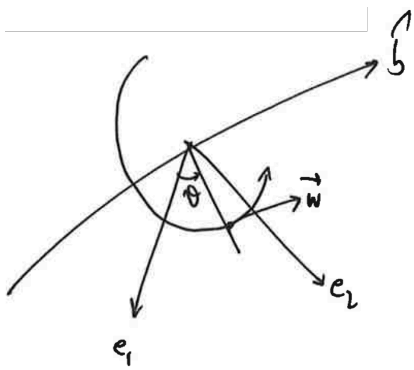
The reason for separating out $\boldsymbol{v}_E$ from the other drifts is that the $\boldsymbol{E} \times \boldsymbol{B}$ is *not* small in ϵ. (Indeed, this is why $\boldsymbol{E} \times \boldsymbol{B}$ motion plays such a prominent role in MHD.) For ease of notation, write

$$\boldsymbol{w} \doteq \boldsymbol{v} - \boldsymbol{v}_E, \tag{VII.13}$$

so that $\boldsymbol{\rho} = -\boldsymbol{w} \times \hat{\boldsymbol{b}}/\Omega$. Now, we know that the directions parallel ($\|$) and perpendicular ($\perp$) to the magnetic field behave differently (certainly in the $\epsilon \ll 1$ limit), so write

$$\boldsymbol{w} = v_\| \hat{\boldsymbol{b}} + \boldsymbol{w}_\perp = v_\| \hat{\boldsymbol{b}} + w_\perp \big(\hat{\boldsymbol{e}}_1 \cos\vartheta + \hat{\boldsymbol{e}}_2 \sin\vartheta\big), \tag{VII.14}$$

where ϑ is the gyrophase:



The coordinates $(\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2, \hat{\boldsymbol{b}})$ are functions of $(t, \boldsymbol{r})$ as the particle sweeps around the changing, inhomogeneous magnetic field. What follows is a gradual shift of the particle coordinates from $(\boldsymbol{r}, \boldsymbol{v})$ to $(\boldsymbol{R}, v_\|, w_\perp, \vartheta)$.

Let us first examine the motion of the guiding-center position, which follows from (VII.1) and (VII.12):

$$\dot{\boldsymbol{R}} = \dot{\boldsymbol{r}} - \dot{\boldsymbol{\rho}} = \underbrace{v_\| \hat{\boldsymbol{b}}}_{0} + \underbrace{\boldsymbol{v}_E}_{0} - \underbrace{\frac{\mathrm{d}\boldsymbol{v}_E}{\mathrm{d}t} \times \frac{\hat{\boldsymbol{b}}}{\Omega}}_{1} + \underbrace{\boldsymbol{w} \times \frac{\mathrm{d}}{\mathrm{d}t}\frac{\hat{\boldsymbol{b}}}{\Omega}}_{1}. \tag{VII.15}$$

The order in ϵ of each term (relative to $v_{\rm th}$) has been noted. To leading order, there is parallel streaming and the $\boldsymbol{E} \times \boldsymbol{B}$ drift. The next-order terms are those dependent upon spatiotemporal changes in the electromagnetic fields along the particle's trajectory.

Next, the evolution of the parallel velocity:

$$\dot{v}_\| = \frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{v} \cdot \hat{\boldsymbol{b}}) = \underbrace{\frac{q}{m} E_\|}_{-1} + \underbrace{(\boldsymbol{v}_E + \boldsymbol{w}) \cdot \frac{\mathrm{d}\hat{\boldsymbol{b}}}{\mathrm{d}t}}_{0},$$

where the ordering is given relative to $v_{\rm th}/\tau$. That $\mathcal{O}(\epsilon^{-1})$ term is a problem; it says that $E_\|$ accelerates particles along field lines on the timescale of a Larmor gyration. Since ions and electrons are accelerated in opposite directions, this would lead to a rapid charge separation, ultimately violating our assumption of slowly varying fields. $E_\parallel$ must be $\mathcal{O}(\epsilon)$:

$$\dot{v}_\parallel = \frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{v}\boldsymbol{\cdot}\hat{\boldsymbol{b}}) = \underbrace{\frac{q}{m}E_\parallel}_{0} + \underbrace{(\boldsymbol{v}_E+\boldsymbol{w})\boldsymbol{\cdot}\frac{\mathrm{d}\hat{\boldsymbol{b}}}{\mathrm{d}t}}_{0}\,. \tag{VII.16}$$

Following similar steps, one can also show that

$$\dot{w}_\perp = -\hat{\boldsymbol{e}}_\perp\boldsymbol{\cdot}\left(\underbrace{v_\parallel\frac{\mathrm{d}\hat{\boldsymbol{b}}}{\mathrm{d}t}}_{0} + \underbrace{\frac{\mathrm{d}\boldsymbol{v}_E}{\mathrm{d}t}}_{0}\right). \tag{VII.17}$$

All these terms have clean physical interpretations. Parallel electric fields accelerate particles along field lines; the plane of the perpendicular drifts tilts as the particles stream along a varying $\hat{\boldsymbol{b}}$; and parallel motion can become perpendicular motion if $\hat{\boldsymbol{b}}$ changes along the orbit.

It's a bit more work to show that

$$\dot{\vartheta} = \underbrace{-\varOmega}_{-1} \underbrace{-\,\hat{\boldsymbol{e}}_2\boldsymbol{\cdot}\frac{\mathrm{d}\hat{\boldsymbol{e}}_1}{\mathrm{d}t}}_{0} \underbrace{-\,\frac{\boldsymbol{w}_\perp\times\hat{\boldsymbol{b}}}{w_\perp^2}\boldsymbol{\cdot}\left(v_\parallel\frac{\mathrm{d}\hat{\boldsymbol{b}}}{\mathrm{d}t}+\frac{\mathrm{d}\boldsymbol{v}_E}{\mathrm{d}t}\right)}_{0}, \tag{VII.18}$$

and so I'll show you the steps. (It should be obvious that the dominant term is $-\varOmega$, i.e., $\dot{\vartheta} = -\varOmega + \mathcal{O}(\epsilon^0) + \dots$) Here are those steps:

$$\begin{aligned}
\frac{\mathrm{d}\boldsymbol{w}_\perp}{\mathrm{d}t} &= \frac{\mathrm{d}w_\perp}{\mathrm{d}t}\hat{\boldsymbol{e}}_\perp + w_\perp\left(\frac{\mathrm{d}\hat{\boldsymbol{e}}_1}{\mathrm{d}t}\cos\vartheta + \frac{\mathrm{d}\hat{\boldsymbol{e}}_2}{\mathrm{d}t}\sin\vartheta\right) + \underbrace{w_\perp(-\hat{\boldsymbol{e}}_1\sin\vartheta+\hat{\boldsymbol{e}}_2\cos\vartheta)\,\dot{\vartheta}}_{=\,-\boldsymbol{w}_\perp\times\hat{\boldsymbol{b}}} \\
&= -\hat{\boldsymbol{e}}_\perp\hat{\boldsymbol{e}}_\perp\boldsymbol{\cdot}\left(v_\parallel\frac{\mathrm{d}\hat{\boldsymbol{b}}}{\mathrm{d}t}+\frac{\mathrm{d}\boldsymbol{v}_E}{\mathrm{d}t}\right) + w_\perp\left(\frac{\mathrm{d}\hat{\boldsymbol{e}}_1}{\mathrm{d}t}\cos\vartheta + \frac{\mathrm{d}\hat{\boldsymbol{e}}_2}{\mathrm{d}t}\sin\vartheta\right) - (\boldsymbol{w}_\perp\times\hat{\boldsymbol{b}})\,\dot{\vartheta}
\end{aligned}$$

$$\begin{aligned}
\Longrightarrow\quad -\frac{\boldsymbol{w}_\perp\times\hat{\boldsymbol{b}}}{w_\perp^2}\boldsymbol{\cdot}\frac{\mathrm{d}\boldsymbol{w}_\perp}{\mathrm{d}t} &= \cancelto{0}{\frac{(\boldsymbol{w}_\perp\times\hat{\boldsymbol{b}})\boldsymbol{\cdot}\hat{\boldsymbol{e}}_\perp\hat{\boldsymbol{e}}_\perp}{w_\perp^2}\boldsymbol{\cdot}\left(v_\parallel\frac{\mathrm{d}\hat{\boldsymbol{b}}}{\mathrm{d}t}+\frac{\mathrm{d}\boldsymbol{v}_E}{\mathrm{d}t}\right)} \\
&\quad + \underbrace{\left(-\,\hat{\boldsymbol{e}}_1\boldsymbol{\cdot}\frac{\mathrm{d}\hat{\boldsymbol{e}}_2}{\mathrm{d}t}\sin^2\vartheta + \hat{\boldsymbol{e}}_2\boldsymbol{\cdot}\frac{\mathrm{d}\hat{\boldsymbol{e}}_1}{\mathrm{d}t}\cos^2\vartheta\right)}_{\text{since }\hat{\mathbf{e}}_1\boldsymbol{\cdot}\mathrm{d}\hat{\mathbf{e}}_1/\mathrm{d}t\,=\,\hat{\mathbf{e}}_2\boldsymbol{\cdot}\mathrm{d}\hat{\mathbf{e}}_2/\mathrm{d}t\,=\,0} + \dot{\vartheta}
\end{aligned}$$

$$\begin{aligned}
\Longrightarrow\quad \dot{\vartheta} &= -\frac{\boldsymbol{w}_\perp\times\hat{\boldsymbol{b}}}{w_\perp^2}\boldsymbol{\cdot}\frac{\mathrm{d}\boldsymbol{w}_\perp}{\mathrm{d}t} - \hat{\boldsymbol{e}}_2\boldsymbol{\cdot}\frac{\mathrm{d}\hat{\boldsymbol{e}}_1}{\mathrm{d}t} \qquad (\text{since }\; -\hat{\boldsymbol{e}}_1\boldsymbol{\cdot}\mathrm{d}\hat{\boldsymbol{e}}_2/\mathrm{d}t = \hat{\boldsymbol{e}}_2\boldsymbol{\cdot}\mathrm{d}\hat{\boldsymbol{e}}_1/\mathrm{d}t) \\
&= -\varOmega - \hat{\boldsymbol{e}}_2\boldsymbol{\cdot}\frac{\mathrm{d}\hat{\boldsymbol{e}}_1}{\mathrm{d}t} - \frac{\boldsymbol{w}_\perp\times\hat{\boldsymbol{b}}}{w_\perp^2}\boldsymbol{\cdot}\left(v_\parallel\frac{\mathrm{d}\hat{\boldsymbol{b}}}{\mathrm{d}t}+\frac{\mathrm{d}\boldsymbol{v}_E}{\mathrm{d}t}\right).
\end{aligned}$$

So, we now have the evolution of $(\boldsymbol{R}, v_\parallel, w_\perp, \vartheta)$, but it's given in terms of $(\boldsymbol{r}, \boldsymbol{v})$. To proceed, we must write the latter in terms of the former.

To do that, we Taylor expand about the guiding-center position; e.g.,

$$\hat{\boldsymbol{b}}(t,\boldsymbol{r}) = \hat{\boldsymbol{b}}(t,\boldsymbol{R}) - \frac{\boldsymbol{w}_\perp \times \hat{\boldsymbol{b}}}{\Omega} \boldsymbol{\cdot} \boldsymbol{\nabla} \hat{\boldsymbol{b}}(t,\boldsymbol{R}) + \dots \tag{VII.19}$$

Also,

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{\boldsymbol{r},\boldsymbol{v}} = \left.\frac{\partial}{\partial t}\right|_{\boldsymbol{R},v_\parallel,w_\perp,\vartheta} + \dot{\boldsymbol{R}}\boldsymbol{\cdot}\left.\frac{\partial}{\partial \boldsymbol{R}}\right|_{t,v_\parallel,w_\perp,\vartheta} + \dot{v}_\parallel \left.\frac{\partial}{\partial v_\parallel}\right|_{t,\boldsymbol{R},w_\perp,\vartheta} + \dot{w}_\perp \left.\frac{\partial}{\partial w_\perp}\right|_{t,\boldsymbol{R},v_\parallel,\vartheta} + \dot{\vartheta} \left.\frac{\partial}{\partial \vartheta}\right|_{t,\boldsymbol{R},v_\parallel,w_\perp} . \tag{VII.20}$$

Henceforth, I'll be suppressing the argument $(t,\boldsymbol{R})$ on $\hat{\boldsymbol{b}}$ and $\boldsymbol{v}_E$ and the what's-held-fixed labels on the partial derivatives. Using (VII.19) and (VII.20), we must evaluate our $(\mathrm{d}/\mathrm{d}t)(\boldsymbol{R}, v_\parallel, w_\perp, \vartheta)$ order by order in ϵ.

At $\mathcal{O}(\epsilon^{-1})$, we have $\dot{\vartheta} = -\Omega$, i.e, Larmor gyration. At $\mathcal{O}(\epsilon^0)$, $\dot{\boldsymbol{R}} = v_\parallel \hat{\boldsymbol{b}} + \boldsymbol{v}_E$, which is the same as guiding-center motion in constant fields. Next, work on $\dot{v}_\parallel$ and $\dot{w}_\perp$. Begin by noticing that

$$\dot{\vartheta}\frac{\partial}{\partial\vartheta} = -\Omega\frac{\partial}{\partial\vartheta} + \mathcal{O}(\epsilon^0)$$

is the biggest term in $\mathrm{d}/\mathrm{d}t$ (see (VII.20)). Thus,

$$\frac{\mathrm{d}\hat{\boldsymbol{b}}}{\mathrm{d}t} = \left(\frac{\partial}{\partial t} + v_\parallel \hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla} + \boldsymbol{v}_E\boldsymbol{\cdot}\boldsymbol{\nabla}\right)\hat{\boldsymbol{b}} + \Omega\frac{\partial}{\partial\vartheta}\left(\frac{\boldsymbol{w}_\perp\times\hat{\boldsymbol{b}}}{\Omega}\right)\boldsymbol{\cdot}\boldsymbol{\nabla}\hat{\boldsymbol{b}} + \mathcal{O}(\epsilon), \tag{VII.21}$$

$$\frac{\mathrm{d}\boldsymbol{v}_E}{\mathrm{d}t} = \left(\frac{\partial}{\partial t} + v_\parallel \hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla} + \boldsymbol{v}_E\boldsymbol{\cdot}\boldsymbol{\nabla}\right)\boldsymbol{v}_E + \underbrace{\Omega\frac{\partial}{\partial\vartheta}\left(\frac{\boldsymbol{w}_\perp\times\hat{\boldsymbol{b}}}{\Omega}\right)}_{=\,\boldsymbol{w}_\perp}\boldsymbol{\cdot}\boldsymbol{\nabla}\boldsymbol{v}_E + \mathcal{O}(\epsilon), \tag{VII.22}$$

where (to remind you) $\hat{\boldsymbol{b}}$ and $\boldsymbol{v}_E$ are functions of $(t,\boldsymbol{R})$. (The difference between, say, $\hat{\boldsymbol{b}}(t,\boldsymbol{r})$ and $\hat{\boldsymbol{b}}(t,\boldsymbol{R})$ can be packed into the omitted $\mathcal{O}(\epsilon)$ terms.) Using (VII.21) and (VII.22) in the evolution equations (VII.16) and (VII.17) for $v_\parallel$ and $w_\perp$, respectively, gives

$$\frac{\mathrm{d}v_\parallel}{\mathrm{d}t} = \frac{qE_\parallel}{m} + (\boldsymbol{v}_E + \boldsymbol{w})\boldsymbol{\cdot}\left(\frac{\mathrm{D}\hat{\boldsymbol{b}}}{\mathrm{D}t} + \boldsymbol{w}_\perp\boldsymbol{\cdot}\boldsymbol{\nabla}_\perp\hat{\boldsymbol{b}}\right), \tag{VII.23}$$

$$\frac{\mathrm{d}w_\perp}{\mathrm{d}t} = -\hat{\boldsymbol{e}}_\perp\boldsymbol{\cdot}\left[\left(\frac{\mathrm{D}}{\mathrm{D}t} + \boldsymbol{w}_\perp\boldsymbol{\cdot}\boldsymbol{\nabla}_\perp\right)(v_\parallel\hat{\boldsymbol{b}} + \boldsymbol{v}_E)\right], \tag{VII.24}$$

where

$$\frac{\mathrm{D}}{\mathrm{D}t} \doteq \frac{\partial}{\partial t} + (v_\parallel\hat{\boldsymbol{b}} + \boldsymbol{v}_E)\boldsymbol{\cdot}\boldsymbol{\nabla} \tag{VII.25}$$

is the Lagrangian time derivative in the parallel-streaming and $\boldsymbol{E}\times\boldsymbol{B}$-drifting frame. In (VII.23) and (VII.24) we find a mix of terms that are independent of ϑ and dependent upon ϑ. For example, grouping such terms in (VII.23),

$$\frac{\mathrm{d}v_\parallel}{\mathrm{d}t} = \left\{\frac{qE_\parallel}{m} + \boldsymbol{v}_E\boldsymbol{\cdot}\frac{\mathrm{D}\hat{\boldsymbol{b}}}{\mathrm{D}t}\right\} + \left\{\boldsymbol{w}\boldsymbol{\cdot}\left(\frac{\mathrm{D}\hat{\boldsymbol{b}}}{\mathrm{D}t} + \boldsymbol{w}_\perp\boldsymbol{\cdot}\boldsymbol{\nabla}_\perp\hat{\boldsymbol{b}}\right) + \boldsymbol{w}_\perp\boldsymbol{v}_E\boldsymbol{:}\boldsymbol{\nabla}_\perp\hat{\boldsymbol{b}}\right\}. \tag{VII.26}$$

To separate the two groups, we introduce the gyro-averaging procedure

$$\langle\dots\rangle_{\boldsymbol{R}} \doteq \frac{1}{2\pi}\oint \mathrm{d}\vartheta\,(\dots), \tag{VII.27}$$

where the gyrophase integral is taken at fixed $\boldsymbol{R}$. The following identities are useful:

$$\langle \boldsymbol{w} \rangle_{\boldsymbol{R}} = w_\parallel \hat{\boldsymbol{b}}, \quad \langle \boldsymbol{w}\boldsymbol{w} \rangle_{\boldsymbol{R}} = w_\parallel^2 \hat{\boldsymbol{b}}\hat{\boldsymbol{b}} + \frac{w_\perp^2}{2}(\mathsf{I} - \hat{\boldsymbol{b}}\hat{\boldsymbol{b}}). \tag{VII.28}$$

Applying the gyro-average to (VII.26) and using these identities yields

$$\langle \dot{v}_\parallel \rangle_{\boldsymbol{R}} = \left\{ \frac{qE_\parallel}{m} + \boldsymbol{v}_E \boldsymbol{\cdot} \frac{\mathrm{D}\hat{\boldsymbol{b}}}{\mathrm{D}t} \right\} + \left\{ \frac{w_\perp^2}{2} \underbrace{(\mathsf{I} - \hat{\boldsymbol{b}}\hat{\boldsymbol{b}}) \boldsymbol{:} \boldsymbol{\nabla}_\perp \hat{\boldsymbol{b}}}_{= -\hat{\boldsymbol{b}} \boldsymbol{\cdot} \boldsymbol{\nabla} \ln B} \right\}$$

$$\boxed{\langle \dot{v}_\parallel \rangle_{\boldsymbol{R}} = \frac{qE_\parallel}{m} + \boldsymbol{v}_E \boldsymbol{\cdot} \frac{\mathrm{D}\hat{\boldsymbol{b}}}{\mathrm{D}t} - \frac{w_\perp^2}{2B} \hat{\boldsymbol{b}} \boldsymbol{\cdot} \boldsymbol{\nabla} B} \tag{VII.29}$$

So, guiding-center acceleration along field lines is driven by (1) parallel electric fields, (2) a fictitious force that accounts for boosting to the non-inertial frame of a varying $\boldsymbol{v}_E$, and (3) mirroring forces by parallel gradients in the magnetic-field strength. The interpretation of the second term is aided by noting that

$$\boldsymbol{v}_E \boldsymbol{\cdot} \frac{\mathrm{D}\hat{\boldsymbol{b}}}{\mathrm{D}t} = -\frac{\mathrm{D}\boldsymbol{v}_E}{\mathrm{D}t} \boldsymbol{\cdot} \hat{\boldsymbol{b}},$$

since $\boldsymbol{v}_E \boldsymbol{\cdot} \hat{\boldsymbol{b}} = 0$. In the third term, you should recognize the combination $w_\perp^2/2B$.

Doing the same for $w_\perp \dots$

$$\begin{aligned} \frac{\mathrm{d}w_\perp}{\mathrm{d}t} &= -(\hat{\boldsymbol{e}}_1 \cos\vartheta + \hat{\boldsymbol{e}}_2 \sin\vartheta) \boldsymbol{\cdot} \left(v_\parallel \frac{\mathrm{D}\hat{\boldsymbol{b}}}{\mathrm{D}t} + \frac{\mathrm{D}\boldsymbol{v}_E}{\mathrm{D}t} \right) \\ &\quad - (\hat{\boldsymbol{e}}_1 \cos\vartheta + \hat{\boldsymbol{e}}_2 \sin\vartheta) \boldsymbol{\cdot} \left(v_\parallel \boldsymbol{w}_\perp \boldsymbol{\cdot} \boldsymbol{\nabla}_\perp \hat{\boldsymbol{b}} + \boldsymbol{w}_\perp \boldsymbol{\cdot} \boldsymbol{\nabla}_\perp \boldsymbol{v}_E \right) \end{aligned}$$

$$\begin{aligned} \Longrightarrow \quad \langle \dot{w}_\perp \rangle_{\boldsymbol{R}} &= -\big(\langle \cos\vartheta\, \boldsymbol{w}_\perp \rangle_{\boldsymbol{R}} \boldsymbol{\cdot} \boldsymbol{\nabla}_\perp v_\parallel \hat{\boldsymbol{b}} \big) \boldsymbol{\cdot} \hat{\boldsymbol{e}}_1 - \big(\langle \cos\vartheta\, \boldsymbol{w}_\perp \rangle_{\boldsymbol{R}} \boldsymbol{\cdot} \boldsymbol{\nabla}_\perp \boldsymbol{v}_E \big) \boldsymbol{\cdot} \hat{\boldsymbol{e}}_1 \\ &\quad - \big(\langle \sin\vartheta\, \boldsymbol{w}_\perp \rangle_{\boldsymbol{R}} \boldsymbol{\cdot} \boldsymbol{\nabla}_\perp v_\parallel \hat{\boldsymbol{b}} \big) \boldsymbol{\cdot} \hat{\boldsymbol{e}}_2 - \big(\langle \sin\vartheta\, \boldsymbol{w}_\perp \rangle_{\boldsymbol{R}} \boldsymbol{\cdot} \boldsymbol{\nabla}_\perp \boldsymbol{v}_E \big) \boldsymbol{\cdot} \hat{\boldsymbol{e}}_2 \\ &= -\frac{w_\perp}{2} \hat{\boldsymbol{e}}_1 \hat{\boldsymbol{e}}_1 \boldsymbol{:} \boldsymbol{\nabla}_\perp v_\parallel \hat{\boldsymbol{b}} - \frac{w_\perp}{2} \hat{\boldsymbol{e}}_1 \hat{\boldsymbol{e}}_1 \boldsymbol{:} \boldsymbol{\nabla}_\perp \boldsymbol{v}_E \\ &\quad - \frac{w_\perp}{2} \hat{\boldsymbol{e}}_2 \hat{\boldsymbol{e}}_2 \boldsymbol{:} \boldsymbol{\nabla}_\perp v_\parallel \hat{\boldsymbol{b}} - \frac{w_\perp}{2} \hat{\boldsymbol{e}}_2 \hat{\boldsymbol{e}}_2 \boldsymbol{:} \boldsymbol{\nabla}_\perp \boldsymbol{v}_E \\ &= -\frac{w_\perp}{2} (\mathsf{I} - \hat{\boldsymbol{b}}\hat{\boldsymbol{b}}) \boldsymbol{:} \boldsymbol{\nabla} v_\parallel \hat{\boldsymbol{b}} - \frac{w_\perp}{2} (\mathsf{I} - \hat{\boldsymbol{b}}\hat{\boldsymbol{b}}) \boldsymbol{:} \boldsymbol{\nabla} \boldsymbol{v}_E \end{aligned}$$

$$\boxed{\langle \dot{w}_\perp \rangle_{\boldsymbol{R}} = \frac{v_\parallel w_\perp}{2B} \hat{\boldsymbol{b}} \boldsymbol{\cdot} \boldsymbol{\nabla} B - \frac{w_\perp}{2} (\mathsf{I} - \hat{\boldsymbol{b}}\hat{\boldsymbol{b}}) \boldsymbol{:} \boldsymbol{\nabla} \boldsymbol{v}_E} \tag{VII.30}$$

And, in a similar manner,

$$\boxed{\langle \dot{\vartheta} \rangle_{\boldsymbol{R}} = -\mathit{\Omega} - \hat{\boldsymbol{e}}_2 \boldsymbol{\cdot} \frac{\mathrm{D}\hat{\boldsymbol{e}}_1}{\mathrm{D}t} - \frac{v_\parallel}{2} \hat{\boldsymbol{b}} \boldsymbol{\cdot} \boldsymbol{\nabla} \times (v_\parallel \hat{\boldsymbol{b}} + \boldsymbol{v}_E)} \tag{VII.31}$$

But this one doesn't really matter – we'll only ever need the leading-order $\dot{\vartheta} = -\mathit{\Omega}$.

We can also go back and compute the $\mathcal{O}(\epsilon)$ terms in $\dot{\boldsymbol{R}}$ (see (VII.15)), in order to see the appearance of inhomogeneities in the evolution of the guiding center:

$$
\begin{aligned}
\langle \dot{\boldsymbol{R}} \rangle_{\boldsymbol{R}} &= v_{\parallel}\hat{\boldsymbol{b}} + \boldsymbol{v}_E - \left\langle \frac{\mathrm{d}\boldsymbol{v}_E}{\mathrm{d}t} \times \frac{\hat{\boldsymbol{b}}}{\Omega} \right\rangle_{\boldsymbol{R}} + \left\langle \boldsymbol{w} \times \frac{\mathrm{d}}{\mathrm{d}t}\frac{\hat{\boldsymbol{b}}}{\Omega} \right\rangle_{\boldsymbol{R}} \\
&= v_{\parallel}\hat{\boldsymbol{b}} + \boldsymbol{v}_E - \frac{\mathrm{D}\boldsymbol{v}_E}{\mathrm{D}t} \times \frac{\hat{\boldsymbol{b}}}{\Omega} + v_{\parallel}\hat{\boldsymbol{b}} \times \frac{\mathrm{D}\hat{\boldsymbol{b}}}{\mathrm{D}t} + \left\langle \boldsymbol{w}_{\perp} \times \left(\boldsymbol{w}_{\perp} \times \boldsymbol{\nabla}_{\perp} \frac{\hat{\boldsymbol{b}}}{\Omega} \right) \right\rangle_{\boldsymbol{R}} \\
&= v_{\parallel}\hat{\boldsymbol{b}} + \boldsymbol{v}_E - \frac{\mathrm{D}\boldsymbol{v}_E}{\mathrm{D}t} \times \frac{\hat{\boldsymbol{b}}}{\Omega} + v_{\parallel}\hat{\boldsymbol{b}} \times \frac{\mathrm{D}\hat{\boldsymbol{b}}}{\mathrm{D}t} \\
&\quad + \frac{w_{\perp}^2}{2}\left[\hat{\boldsymbol{e}}_1 \times \left(\hat{\boldsymbol{e}}_1 \times \boldsymbol{\nabla}\frac{\hat{\boldsymbol{b}}}{\Omega} \right) + \hat{\boldsymbol{e}}_2 \times \left(\hat{\boldsymbol{e}}_2 \times \boldsymbol{\nabla}\frac{\hat{\boldsymbol{b}}}{\Omega} \right) \right]
\end{aligned}
$$

$$
\boxed{\langle \dot{\boldsymbol{R}} \rangle_{\boldsymbol{R}} = \left[v_{\parallel} + \frac{w_{\perp}^2}{2\Omega}\hat{\boldsymbol{b}}\boldsymbol{\cdot}(\boldsymbol{\nabla}\times\hat{\boldsymbol{b}}) \right]\hat{\boldsymbol{b}} + \boldsymbol{v}_E + \frac{w_{\perp}^2}{2\Omega}\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\ln B + \hat{\boldsymbol{b}}\times\left(v_{\parallel}\frac{\mathrm{D}\hat{\boldsymbol{b}}}{\mathrm{D}t} + \frac{1}{\Omega}\frac{\mathrm{D}\boldsymbol{v}_E}{\mathrm{D}t} \right)} \tag{VII.32}
$$

Again, a reminder: every $\hat{\boldsymbol{b}}$ and $\boldsymbol{v}_E$ in these formulae are evaluated at $(t, \boldsymbol{R})$. From left to right, we have (1) parallel streaming (including an $\mathcal{O}(\epsilon)$ correction to the parallel velocity), (2) $\boldsymbol{E}\times\boldsymbol{B}$ drift, (3) grad-B drift, (4) curvature drift, and (5) polarization drift.

VII.5. First adiabatic invariant

The equation for $\langle \dot{w}_{\perp} \rangle_{\boldsymbol{R}}$, (VII.30), implies something special. Note that

$$
\begin{aligned}
(\mathsf{I} - \hat{\boldsymbol{b}}\hat{\boldsymbol{b}})\boldsymbol{:}\boldsymbol{\nabla}\boldsymbol{v}_E &= \underbrace{\boldsymbol{\nabla}\boldsymbol{\cdot}\left(\frac{c\boldsymbol{E}\times\hat{\boldsymbol{b}}}{B} \right)}_{\text{use Faraday's law}} - \hat{\boldsymbol{b}}\boldsymbol{\cdot}\left(\hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla}\frac{c\boldsymbol{E}\times\hat{\boldsymbol{b}}}{B} \right) \\
&= -\frac{\partial \ln B}{\partial t} \underbrace{- c\boldsymbol{E}\boldsymbol{\cdot}\left(\boldsymbol{\nabla}\times\frac{\hat{\boldsymbol{b}}}{B} \right) - \hat{\boldsymbol{b}}\boldsymbol{\cdot}\left(\hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla}\frac{c\boldsymbol{E}\times\hat{\boldsymbol{b}}}{B} \right)}_{\text{user vector identities to expand}} \\
&= -\frac{\partial \ln B}{\partial t} - \frac{c\boldsymbol{E}\boldsymbol{\cdot}(\boldsymbol{\nabla}\times\hat{\boldsymbol{b}})}{B} \underbrace{- \frac{c\boldsymbol{E}\boldsymbol{\cdot}(\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\ln B)}{B} - \frac{c\hat{\boldsymbol{b}}\hat{\boldsymbol{b}}\boldsymbol{:}\boldsymbol{\nabla}(\boldsymbol{E}\times\hat{\boldsymbol{b}})}{B}}_{\text{use vector identities to rearrange}} \\
&= -\frac{\partial \ln B}{\partial t} - \underbrace{\frac{c\boldsymbol{E}\boldsymbol{\cdot}(\boldsymbol{\nabla}\times\hat{\boldsymbol{b}})}{B}}_{\substack{\text{write} \\ \boldsymbol{E} = \boldsymbol{E}_{\parallel}\hat{\boldsymbol{b}} + \boldsymbol{E}_{\perp}}} - \underbrace{\frac{c\boldsymbol{E}\times\hat{\boldsymbol{b}}}{B}}_{=\boldsymbol{v}_E}\boldsymbol{\cdot}\boldsymbol{\nabla}\ln B + \frac{c\boldsymbol{E}_{\perp}\boldsymbol{\cdot}(\boldsymbol{\nabla}\times\hat{\boldsymbol{b}})}{B} \\
&= -\frac{\partial \ln B}{\partial t} - \underbrace{\frac{cE_{\parallel}\hat{\boldsymbol{b}}\boldsymbol{\cdot}(\boldsymbol{\nabla}\times\hat{\boldsymbol{b}})}{B}}_{\substack{\text{is } \mathcal{O}(\epsilon) \text{ relative} \\ \text{to other terms}}} - \boldsymbol{v}_E\boldsymbol{\cdot}\boldsymbol{\nabla}\ln B \\
&= -\frac{\partial \ln B}{\partial t} - \boldsymbol{v}_E\boldsymbol{\cdot}\boldsymbol{\nabla}\ln B + \mathcal{O}(\epsilon)
\end{aligned}
$$

And so (VII.30) becomes

$$\begin{aligned}\langle \dot{w}_\perp \rangle_{\boldsymbol{R}} &= \frac{v_\parallel w_\perp}{2}\hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla}\ln B - \frac{w_\perp}{2}\left(-\frac{\partial \ln B}{\partial t} - \boldsymbol{v}_E\boldsymbol{\cdot}\boldsymbol{\nabla}\ln B\right) + \mathcal{O}(\epsilon) \\ &= \frac{w_\perp}{2}\left(\frac{\partial}{\partial t} + v_\parallel \hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla} + \boldsymbol{v}_E\boldsymbol{\cdot}\boldsymbol{\nabla}\right)\ln B + \mathcal{O}(\epsilon) \\ &= \frac{w_\perp}{2}\frac{\mathrm{D}\ln B}{\mathrm{D}t} + \mathcal{O}(\epsilon),\end{aligned}$$

which implies

$$\boxed{\langle \dot{\mu} \rangle_{\boldsymbol{R}} = \mathcal{O}(\epsilon)} \tag{VII.33}$$

where $\mu \doteq w_\perp^2/2B(t,\boldsymbol{R})$. In words, the magnetic moment μ is constant on the time and length scales of the field variation. Its constantly is telling us that, on these time and length scales, ϑ is an ignorable coordinate. (This property forms the basis of gyrokinetics.) More fundamentally, μ conservation is telling us that plasmas are 'diamagnetic', that is, all particle-generated fluxes add to reduce the ambient field. The total change in B is proportional to the change in the perpendicular kinetic energy of the particle. The greater the plasma thermal energy, the more it excludes the magnetic field.

For a fluid element containing an ensemble of magnetized particles, μ conservation implies that the thermal pressure perpendicular to the local magnetic field of that fluid element $P_\perp \doteq \langle m w_\perp^2/2 \rangle \propto nB$, where the angle brackets $\langle\,\cdot\,\rangle$ denote the ensemble average. We'll return to this important point later in the course.

VII.6. Adiabatic invariance

μ is one of several adiabatic invariants, which are related to the exactly conserved Poincaré invariants of classical mechanics. Adiabatic invariance is one of the most important concepts in the plasma physics of weakly collisional plasmas. The invariants emerge from the periodic motion induced by the magnetic field, and derive from the Hamiltonian action $\oint \boldsymbol{\wp}\boldsymbol{\cdot}\mathrm{d}\boldsymbol{q}$ around a loop representing nearly periodic motion. μ is the 'first adiabatic invariant' of plasma physics; the corresponding periodic motion is obviously the gyromotion of a particle about a magnetic field. The canonical momentum $\wp$ in this case is the particle' angular momentum, $mv_\perp\rho$; the angular variable ϑ is the q, the conjugate coordinate. If the particle's orbit changes slowly, either because $\partial_t \ln B \ll \varOmega$ or because the particle is drifting slowly into a region of varying field strength and/or geometry, then the action changes very little.[10] You might see a 'simple' derivation of μ conservation in some textbooks, rather different from the guiding-center-theory approach

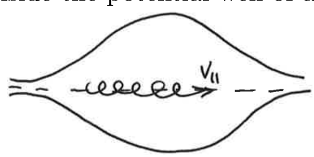
[10]How little? Kruskal (1958, 1962) and Northrop (1963*b*) showed that μ is conserved 'to all orders', meaning that, if μ can be written as an expansion in the small parameter ϵ, $\mu = \mu_0 + \epsilon\mu_1 + \epsilon^2\mu_2 + \dots$, then $\varDelta\mu \doteq \mu - \mu_0 = c_1 \exp(-c_2\epsilon)$, where c_1 and c_2 are positive constants of order unity.

we've taken above. It runs something like this:

$$
\begin{aligned}
(\Delta\mu \text{ in one orbit}) &= \frac{\Delta(mw_\perp^2/2)}{B} - \mu\frac{\Delta B}{B} = \frac{1}{B}\int_0^{2\pi/\Omega} \mathrm{d}t\, \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}mw_\perp^2\right) - \mu\frac{\Delta B}{B} \\
&= \frac{1}{B}\int_0^{2\pi/\Omega} \mathrm{d}t\, q\boldsymbol{w}_\perp \boldsymbol{\cdot} \boldsymbol{E}_\perp - \mu\frac{\Delta B}{B} \\
&= \frac{q}{B}\oint \mathrm{d}\boldsymbol{\ell}_\perp \boldsymbol{\cdot} \boldsymbol{E}_\perp - \mu\frac{\Delta B}{B} \\
&= \mu\frac{\Delta B}{B} - \mu\frac{\Delta B}{B} = 0.
\end{aligned}
$$

The idea is that the electric field associated with the change in the magnetic field accelerates the particle, increasing its perpendicular energy in such a way that μ is conserved.

A nice example of adiabatic invariance at work is magnetic mirroring. Imagine a magnetized particle trapped inside the potential well of a static magnetic bottle:



The energy of the particle is conserved,

$$\varepsilon = \frac{1}{2}mv_\parallel^2 + \frac{1}{2}mv_\perp^2 = \text{const},$$

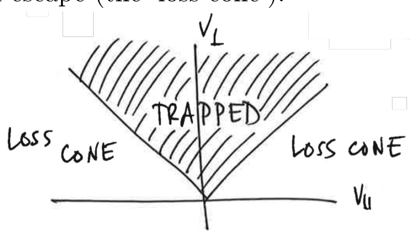
as is its magnetic moment, $\mu = \text{const}$. Thus, as the particle moves from its initial position where the magnetic-field strength is B_0 into a region where the field strength is B, its parallel velocity, initially $v_{\parallel 0}$, must adjust according to these constraints:

$$\frac{1}{2}mv_{\parallel 0}^2 + \mu B_0 = \frac{1}{2}mv_\parallel^2 + \mu B = \varepsilon \quad \Longrightarrow \quad v_\parallel = \pm\sqrt{\frac{2}{m}(\varepsilon - \mu B)} \tag{VII.34}$$

With ε and μ constant, this establishes a relationship between the parallel velocity of the particle and the local magnetic-field strength (at the particle's gyro-center): if B increases in the particle's frame, $v_\perp$ must increase by μ conservation, and $v_\parallel$ must then decrease by energy conservation. If the particle encounters a strong enough magnetic field that $v_\parallel \to 0$, the particle is said to 'reflect' off of the strong-field region. The criterion for reflection is (VII.34) with $v_\parallel = 0$:

$$\frac{1}{2}mv_{\parallel 0}^2 + \mu(B_0 - B) = 0 \quad \Longrightarrow \quad \frac{v_{\parallel 0}}{v_{\perp 0}} \leqslant \sqrt{\frac{B}{B_0} - 1} \text{ for confinement} \tag{VII.35}$$

This defines a critical pitch angle separated particles that are trapped inside the magnetic bottle from those that can escape (the 'loss cone'):

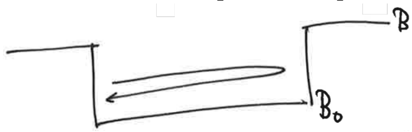


Collisions, which break μ, would of course promote the leakage of particles out of the trapped region.

Now, what if the ends of the mirror were to move slowly?

VII.7. Second adiabatic invariant

Imagine a charged particle confined in a square-well potential:



Assume that the bounce time (i.e., the time required for the particle to transit the mirror, bounce, and return to its starting point) is much less than the time over which the ends of the mirror move. There will be an approximately conserved quantity,

$$\mathcal{J} \doteq \oint \mathrm{d}s\, m v_{\parallel}, \tag{VII.36}$$

associated with the periodic bounce motion of the guiding center in the evolving mirror. This integral – the second adiabatic invariant – is taken over the 'bounce orbit' of the guiding center, with the differential $\mathrm{d}s$ oriented along the local magnetic-field direction and the limits being the turning points of the bounce orbit.[11] For example, if the mirror shrinks adiabatically, then $v_{\parallel}$ increases.[12] Proving this is more involved, and $\mathcal{J}$ is typically a less robust invariant than μ (although it is of crucial importance for the persistence of the van Allen belts, by ensuring that precessing particles trapped in the Earth's magnetic field return to their native field line after circumnavigating the Earth). If you're interested in the finer details, consult Northrop (1963*a*, pg. 294).

VII.8. Third adiabatic invariant

There is yet another adiabatic invariant associated with the periodic motion of charged particles in a magnetic field, but it often receives much less attention than the first two because of its lesser utility. The reason is because the associated periodic motion is not as general as, say, a particle gyrating about a field line. In this case, the approximately conserved quantity

$$\mathcal{K} \doteq \oint \mathrm{d}\ell\, \wp_{\phi} \simeq e \oint \mathrm{d}\ell\, A_{\phi} = e\Phi \tag{VII.37}$$

is the magnetic flux enclosed within a periodic orbit caused by cross-field drifts. (The drift velocity v_{ϕ} is typically small compared to eA_{ϕ}, thus the '$\simeq$' in (VII.37).) If the particle orbit also involves bouncing between two turning points in a magnetic mirror, then the periodic orbit associated with the drift motion is to be evaluated at the 'bounce center' (just as $\mathcal{J}$ is to be evaluated using an orbit of the guiding center). As with all adiabatic invariants, there is a comparison of time scales that must be done; here, it is

[11] The canonical momentum here is technically $mv_{\parallel} + eA_{\parallel}/c$, but the latter (vector-potential) term representing the momentum associated with the electromagnetic field, once integrated over the bounce orbit, equals the total amount of magnetic flux enclosed by the orbit ($=0$).

[12] Note that both μ and $\mathcal{J}$ are of the form (energy)/(frequency). This is the general form of an adiabatic invariant. Think of $E/\omega = \hbar$ (Einstein) or $\oint p\, \mathrm{d}q = nh$ (Sommerfeld). Einstein, at the Solray conference in 1911, said that this is the general form of an adiabatic invariant, and that this is what ought to be quantized.

between the time scale on which the magnetic field varies and the period of the drift orbit. In the Earth's inner magnetosphere, the time for trapped particles with energies of $\sim$MeV to circumnavigate the Earth via their cross-field drifts is $\sim$1 hr, and so any geomagnetic storms would interfere with $\mathcal{K}$ conservation. Again, the separation of time scales and field geometry required for $\mathcal{K}$ conservation is not particularly general, but it is important to bear in mind that particles like to keep the total magnetic flux constant within both their gyro-orbits (μ conservation) and their drift orbits ($\mathcal{K}$ conservation).

VII.9. Summary of single-particle drifts and their currents

Here is a summary of the single-particle guiding-center drifts that we've discussed:

$$\boldsymbol{v}_E = \frac{c}{B}\boldsymbol{E}\times\hat{\boldsymbol{b}}, \quad \boldsymbol{v}_{\nabla B} = \frac{w_\perp^2}{2\Omega}\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\ln B, \quad \boldsymbol{v}_{\rm c} = \frac{v_\parallel^2}{\Omega}\hat{\boldsymbol{b}}\times(\hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla}\hat{\boldsymbol{b}}), \quad \boldsymbol{v}_{\rm pol} = \frac{1}{\Omega}\frac{c}{B}\frac{\partial\boldsymbol{E}_\perp}{\partial t},$$

where the latter two are part of the more general "acceleration" drift,

$$\frac{\hat{\boldsymbol{b}}}{\Omega}\times\frac{\rm D}{{\rm D}t}(v_\parallel\hat{\boldsymbol{b}}+\boldsymbol{v}_E).$$

Note that these are all perpendicular to the magnetic field. Later in these notes, we will make a connection between the single-particle drifts and the magnetohydrodynamic equations. For that, we actually need one more "drift" – the *diamagnetic flow* – and something called the *magnetization current*. These contributions, neither of which are associated with true particle drifts, are discussed in the next two sections. But first it will help to compute the currents associated with the particle drifts and include them here. As we have already emphasized, the $\boldsymbol{E}\times\boldsymbol{B}$ drift is species independent, and thus contributes no current in a quasi-neutral plasma. What about the others?

To compute the perpendicular currents associated with the particle drifts, $\boldsymbol{j}_{\perp,{\rm dr}}$, we imagine a plasma whose particles' velocities are distributed according to a distribution function $f_\alpha(\boldsymbol{v})$ for each species α. The perpendicular current is then obtained by affixing a species label α to the drifts we computed, multiplying each of them by q_α, summing over species, and integrating over the velocity space after weighting each drift by f_α, *viz.*

$$\begin{aligned}
\boldsymbol{j}_{\perp,{\rm dr}} &= \sum_\alpha q_\alpha\int{\rm d}\boldsymbol{v}\,\boldsymbol{v}_{{\rm dr},\alpha}f_\alpha \\
&= \sum_\alpha q_\alpha\int{\rm d}\boldsymbol{v}\left[\frac{w_\perp^2}{2\Omega_\alpha}\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\ln B+\frac{\hat{\boldsymbol{b}}}{\Omega_\alpha}\times\frac{\rm D}{{\rm D}t}(v_\parallel\hat{\boldsymbol{b}}+\boldsymbol{v}_{\rm E})\right]f_\alpha \\
&= \frac{c}{B}\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\ln B\sum_\alpha\int{\rm d}\boldsymbol{v}\,\frac{1}{2}m_\alpha w_\perp^2 f_\alpha+\frac{c}{B}\hat{\boldsymbol{b}}\times(\hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla}\hat{\boldsymbol{b}})\sum_\alpha\int{\rm d}\boldsymbol{v}\,m_\alpha v_\parallel^2 f_\alpha \\
&\quad+\frac{c}{B}\hat{\boldsymbol{b}}\times\left(\frac{\partial\hat{\boldsymbol{b}}}{\partial t}+\boldsymbol{v}_E\boldsymbol{\cdot}\boldsymbol{\nabla}\hat{\boldsymbol{b}}+\hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla}\boldsymbol{v}_E\right)\sum_\alpha\int{\rm d}\boldsymbol{v}\,m_\alpha v_\parallel f_\alpha \\
&\quad+\frac{c}{B}\hat{\boldsymbol{b}}\times\left(\frac{\partial\boldsymbol{v}_E}{\partial t}+\boldsymbol{v}_E\boldsymbol{\cdot}\boldsymbol{\nabla}\boldsymbol{v}_E\right)\sum_\alpha\int{\rm d}\boldsymbol{v}\,m_\alpha f_\alpha.
\end{aligned}\tag{VII.38}$$

Each of the above integrals over the distribution function f_α have a name: $\int{\rm d}\boldsymbol{v}\,m_\alpha f_\alpha = m_\alpha n_\alpha$ is the mass density, $\int{\rm d}\boldsymbol{v}\,m_\alpha v_\parallel f_\alpha = m_\alpha n_\alpha u_{\parallel\alpha}$ is the parallel component of the bulk momentum density, and

$$\int{\rm d}\boldsymbol{v}\,\frac{1}{2}m_\alpha w_\perp^2 f_\alpha = p_{\perp\alpha}, \quad \int{\rm d}\boldsymbol{v}\,m_\alpha v_\parallel^2 f_\alpha = p_{\parallel\alpha}+m_\alpha n_\alpha u_{\parallel\alpha}^2$$

are measures of the particle energies perpendicular and parallel to the magnetic-field direction. Namely, $p_{\perp\alpha}$ ($p_{\parallel\alpha}$) measures the energetic content of the random ("thermal") motions of the particles of species α in the direction perpendicular (parallel) to the local magnetic field. Substituting these expressions into (VII.38) yields

$$\begin{aligned} \frac{B}{c}\boldsymbol{j}_{\perp,\mathrm{dr}} &= \sum_\alpha p_{\perp\alpha}\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\ln B + \sum_\alpha \Big(p_{\parallel\alpha} + m_\alpha n_\alpha u_{\parallel\alpha}^2\Big)\hat{\boldsymbol{b}}\times(\hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla}\hat{\boldsymbol{b}}) \\ &\quad + \sum_\alpha m_\alpha n_\alpha \hat{\boldsymbol{b}}\times\left[u_{\parallel\alpha}\frac{\partial\hat{\boldsymbol{b}}}{\partial t} + u_{\parallel\alpha}\boldsymbol{v}_E\boldsymbol{\cdot}\boldsymbol{\nabla}\hat{\boldsymbol{b}} + u_{\parallel\alpha}\hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla}\boldsymbol{v}_E + \frac{\partial \boldsymbol{v}_E}{\partial t} + \boldsymbol{v}_E\boldsymbol{\cdot}\boldsymbol{\nabla}\boldsymbol{v}_E\right]. \end{aligned} \tag{VII.39}$$

This may be further simplified using $\hat{\boldsymbol{b}}\times\hat{\boldsymbol{b}} = 0$ and writing

$$\frac{\mathrm{d}}{\mathrm{d}t_\alpha} \doteq \frac{\partial}{\partial t} + \boldsymbol{u}_\alpha\boldsymbol{\cdot}\boldsymbol{\nabla}, \quad \text{where } \boldsymbol{u}_\alpha \doteq u_{\parallel\alpha}\hat{\boldsymbol{b}} + \boldsymbol{v}_{\mathrm{E}}. \tag{VII.40}$$

The result is that

$$\frac{B}{c}\boldsymbol{j}_{\perp,\mathrm{dr}} = \sum_\alpha p_{\perp\alpha}\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\ln B + \sum_\alpha p_{\parallel\alpha}\hat{\boldsymbol{b}}\times(\hat{\boldsymbol{b}}\boldsymbol{\cdot}\boldsymbol{\nabla}\hat{\boldsymbol{b}}) + \sum_\alpha m_\alpha n_\alpha\hat{\boldsymbol{b}}\times\frac{\mathrm{d}\boldsymbol{u}_\alpha}{\mathrm{d}t_\alpha}. \tag{VII.41}$$

So there are currents associated with the grad-B drift, the curvature drift, and the acceleration drifts (which include the polarization drift). We'll return to this formula in §VII.11 after discussing the magnetization current.

VII.10. Magnetization current

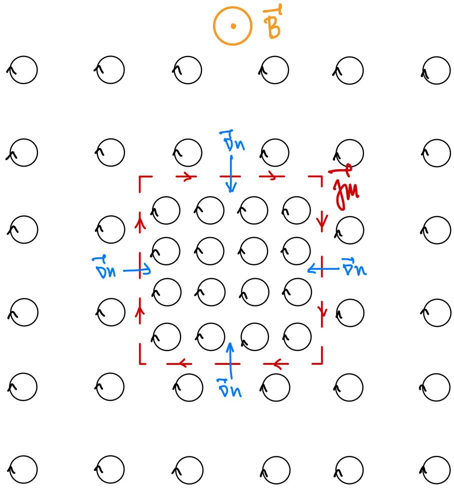
Plasmas are diamagnetic, a fact we pointed out when discussing μ conservation: the greater the plasma (perpendicular) thermal energy, the more it excludes the magnetic field. There is a macroscopic current, not caused by single-particle drifts, associated with this property. Essentially, because a magnetized plasma may be thought of as being composed of magnetic dipoles, each of which being associated with a gyro-orbiting particle, a plasma may be considered as a magnetic material. From basic electromagnetism, the current of a magnetic material in which the magnetization is non-uniform is given by $\boldsymbol{j}_M = c\boldsymbol{\nabla}\times\boldsymbol{M}$, where $\boldsymbol{M}$ is the magnetization per unit volume due to these magnetic dipoles. The latter may be obtained by integrating up all the magnetic moments of each of the particles, $-\mu\hat{\boldsymbol{b}}$ (see (VII.8)), weighted by the particle distribution function:

$$\boldsymbol{M} = -\hat{\boldsymbol{b}}\sum_\alpha\int\mathrm{d}\boldsymbol{v}\,\mu_\alpha f_\alpha = -\frac{\hat{\boldsymbol{b}}}{B}\sum_\alpha\int\mathrm{d}\boldsymbol{v}\,\frac{1}{2}m_\alpha w_\perp^2 f_\alpha = -\frac{\hat{\boldsymbol{b}}}{B}\sum_\alpha p_{\perp\alpha}. \tag{VII.42}$$

The resulting current is

$$\boldsymbol{j}_M = c\boldsymbol{\nabla}\times\boldsymbol{M} = -c\boldsymbol{\nabla}\times\left(\frac{\hat{\boldsymbol{b}}}{B}\sum_\alpha p_{\perp\alpha}\right). \tag{VII.43}$$

The figure below illustrates the origin of this current. In this example, there are more ions gyrating about the (uniform) magnetic field in the center of the plasma than near the edge, and so there is a density (and thus pressure) gradient pointing inwards (indicated by the blue arrows). Therefore, there are more particles whose field-perpendicular velocities are oriented clockwise along the red dashed line than there are particles whose velocities are oriented counter-clockwise. The difference results in a current that flows as indicated, in the $\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}n$ direction. A similar effect occurs if the density of guiding centers is uniform but the particles' perpendicular velocities are larger in some region of space than they are elsewhere. Alternatively, one may think of the magnetization current in terms of diamagnetism: if the perpendicular thermal energy of the particles is larger in one region than in another, the ability of the plasma to exclude magnetic fields is inhomogeneous. This produces a current.



VII.11. Total plasma current and the diamagnetic flow

Let us add up all the perpendicular currents we have discussed thus far:

$$\begin{aligned}\boldsymbol{j}_\perp &= \boldsymbol{j}_M + \boldsymbol{j}_{\perp,\mathrm{dr}} \\ &= -c\boldsymbol{\nabla}\times\left(\frac{\hat{\boldsymbol{b}}}{B}\sum_\alpha p_{\perp\alpha}\right) + \frac{c}{B}\sum_\alpha p_{\perp\alpha}\hat{\boldsymbol{b}}\times\boldsymbol{\nabla}\ln B + \frac{c}{B}\sum_\alpha p_{\parallel\alpha}\hat{\boldsymbol{b}}\times(\hat{\boldsymbol{b}}\cdot\boldsymbol{\nabla}\hat{\boldsymbol{b}}) \\ &\quad + \frac{c}{B}\sum_\alpha m_\alpha n_\alpha \hat{\boldsymbol{b}}\times\frac{\mathrm{d}\boldsymbol{u}_\alpha}{\mathrm{d}t_\alpha}.\end{aligned} \tag{VII.44}$$

The first two terms may be combined to yield

$$\frac{B}{c}\boldsymbol{j}_\perp = -\boldsymbol{\nabla}\times\left(\hat{\boldsymbol{b}}\sum_\alpha p_{\perp\alpha}\right) + \sum_\alpha p_{\parallel\alpha}\hat{\boldsymbol{b}}\times(\hat{\boldsymbol{b}}\cdot\boldsymbol{\nabla}\hat{\boldsymbol{b}}) + \sum_\alpha m_\alpha n_\alpha\hat{\boldsymbol{b}}\times\frac{\mathrm{d}\boldsymbol{u}_\alpha}{\mathrm{d}t_\alpha}. \tag{VII.45}$$

This next step is unlikely to be clear to you, but let's introduce the tensor

$$\boldsymbol{P}_\alpha = p_{\perp\alpha}\big(\boldsymbol{I}-\hat{\boldsymbol{b}}\hat{\boldsymbol{b}}\big) + p_{\parallel\alpha}\hat{\boldsymbol{b}}\hat{\boldsymbol{b}}, \tag{VII.46}$$

where $\boldsymbol{I}$ is the unit dyadic. (This is the form of the thermal pressure tensor in a magnetized plasma.) Noting that $\boldsymbol{\nabla}\times\hat{\boldsymbol{b}} - \hat{\boldsymbol{b}}\cdot(\hat{\boldsymbol{b}}\cdot\boldsymbol{\nabla}\hat{\boldsymbol{b}}) = \hat{\boldsymbol{b}}\hat{\boldsymbol{b}}\cdot(\boldsymbol{\nabla}\times\hat{\boldsymbol{b}})$ is parallel to $\hat{\boldsymbol{b}}$, equation (VII.45) becomes simply

$$\frac{B}{c}\boldsymbol{j}_\perp = \hat{\boldsymbol{b}}\times\sum_\alpha\left(\boldsymbol{\nabla}\cdot\boldsymbol{P}_\alpha + m_\alpha n_\alpha\frac{\mathrm{d}\boldsymbol{u}_\alpha}{\mathrm{d}t_\alpha}\right). \tag{VII.47}$$

Taking $\hat{\boldsymbol{b}}\times$ (VII.47) and using the vector identity $\hat{\boldsymbol{b}}\times(\hat{\boldsymbol{b}}\times\boldsymbol{A}) = -\boldsymbol{A}_\perp$, we obtain

$$\sum_\alpha\left(m_\alpha n_\alpha\frac{\mathrm{d}\boldsymbol{u}_\alpha}{\mathrm{d}t_\alpha} + \boldsymbol{\nabla}\cdot\boldsymbol{P}_\alpha\right)_\perp = \frac{\boldsymbol{j}\times\boldsymbol{B}}{c}. \tag{VII.48}$$

If this looks familiar to you, congratulations! You know MHD, and now you see how single-particle drifts and magnetization current fit into the MHD description.

One last thing. The quantity

$$\boldsymbol{u}_{\mathrm{dia},\alpha} \doteq \frac{\hat{\boldsymbol{b}}}{\Omega_\alpha} \times \frac{\boldsymbol{\nabla} \cdot \boldsymbol{P}_\alpha}{m_\alpha n_\alpha} \tag{VII.49}$$

that appears implicitly in (VII.47) has a name – it is referred to as the *diamagnetic flow velocity* of species α.[13] It is *not* a particle drift, but rather refers to net flux of gyrating particles passing through a reference surface due to an inhomogeneous distribution of guiding centers.

PART VIII
Kinetic theory of plasmas

If you read Part VII, then you might think plasma physics is solved. At any time, compute the electromagnetic fields produced by the charged particles, and then use these fields to evolve the particle phase-space positions using (VII.1), or perhaps to evolve the guiding-center phase-space positions using (VII.29)–(VII.32). Done. Well, not quite. First, there are simply too many particles to follow in a real plasma. There are $\sim 10^{28}$ particles in this room alone. One data dump of $\boldsymbol{r}$ and $\boldsymbol{v}$ for all of these particles would be $\sim 5 \times 10^{17}$ TB (!!!) Secondly, we're not really all that interested in every single particle; we usually want bulk information, like density, momentum, pressure, heat flux, etc. Thirdly, and perhaps most importantly, a many-body system like a plasma is chaotic. Even if we could solve all phase-space trajectories of all the particles given some initial conditions, we would have to admit that those initial conditions are completely arbitrary, and that infinitesimal changes to those initial conditions will yield microscopically different results. It'd be a shame if that mattered, wouldn't it?

This is where a statistical approach comes in handy: What is the probability that a particle will have position $\boldsymbol{r}$ and velocity $\boldsymbol{v}$ in some six-dimensional phase-space interval $\mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{v}$? How does this probability evolve? Under what conditions is this probabilistic evolution accurate enough to yield meaningful predictions for a single realization of the system? Answering these questions is the job of *kinetic theory*. Think about flipping a coin. We all know that the odds of getting heads is 50%. But those are the *odds* – not very useful is you're betting your career on a single coin toss. Or even two coin tosses, or ten. But perhaps very useful if you flip the coin 10^{28} times and make a wager on the percentage of tosses that came up heads. Or, if you're rich and have 10^{28} coins lying around, you could flip them all at the same time and count the number of heads as a fraction of the total. Why kinetic theory works as a predictive theory is that, statistically, there is no difference between running, say, a particle-in-cell code with 100 particles per cell many, many, many times each with different initial conditions randomly sampled from some prescribed distribution, and running the same code just once with 10^{28} particles per cell and a single realization of the initial conditions. The statistics you learn from the

[13] I am deliberately *not* calling it the "diamagnetic drift velocity", as some are wont to do. Nothing is actually drifting, so this moniker makes no sense! Later in this course, we will show that the diamagnetic flow is what one obtains when Taylor-expanding an equilibrium distribution $f(\mathcal{E}, \mu, \boldsymbol{R})$ that is a function of the particle energy $\mathcal{E}$, magnetic moment μ, and guiding-position $\boldsymbol{R}$ about the particle position $\boldsymbol{r} \doteq \boldsymbol{R} + \boldsymbol{\rho}$ and computing its first velocity-space moment.

first experiment should be an accurate representation of the actual results in the second experiment.

The trick is actually building a rigorous, predictive kinetic theory of plasmas. Here, I'll sketch how one is built, highlighting its assumptions. The finer details are quite advanced, but can be found in my lecture notes for AST554 at Princeton University.

VIII.1. The Klimontovich equation as a microscopic description of a plasma

A complete description of a plasma would emerge if one were to have knowledge of all the coordinates and momenta of all of the constituent particles, as well as the electromagnetic fields in which they move and which they self-consistently produce. We've already discussed why such a description would be untenable with which to work, but let us nevertheless adopt this microscopic standpoint and see where it leads.

Start by defining the function

$$F_\alpha(t,\boldsymbol{r},\boldsymbol{v}) = \sum_{i=1}^{N_\alpha} \delta(\boldsymbol{r}-\boldsymbol{R}_{\alpha i}(t))\delta(\boldsymbol{v}-\boldsymbol{V}_{\alpha i}(t)), \tag{VIII.1}$$

which completely specifies the positions $\boldsymbol{R}_{\alpha i}(t)$ and velocities $\boldsymbol{V}_{\alpha i}(t)$ of N_α particles of species α as functions of time. Note that

$$\lim_{\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}\to 0} \int \mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}\, F_\alpha(t,\boldsymbol{r},\boldsymbol{v})$$

is either unity or zero, depending upon whether there is a particle at $(\boldsymbol{r},\boldsymbol{v})$ at time t, so that

$$\int \mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}\, F_\alpha(t,\boldsymbol{r},\boldsymbol{v}) = N_\alpha. \tag{VIII.2}$$

Thus, the microscopic state of the plasma at any time t would be known if one were to know $\boldsymbol{R}_{\alpha i}$ and $\boldsymbol{V}_{\alpha i}$ at $t=0$ and their temporal evolution. Hamilton's equations of motion provide us with the latter:

$$\frac{\mathrm{d}\boldsymbol{R}_{\alpha i}}{\mathrm{d}t} = \boldsymbol{V}_{\alpha i} \qquad \text{and} \qquad \frac{\mathrm{d}\boldsymbol{V}_{\alpha i}}{\mathrm{d}t} = \frac{q_\alpha}{m_\alpha}\left(\boldsymbol{E}_{\mathrm{m}} + \frac{\boldsymbol{V}_{\alpha i}}{c}\times\boldsymbol{B}_{\mathrm{m}}\right), \tag{VIII.3}$$

where q_α and m_α are the charge and mass of species α, and

$$\boldsymbol{E}_{\mathrm{m}} = \boldsymbol{E}_{\mathrm{m}}(t,\boldsymbol{R}_{\alpha i}(t)) \qquad \text{and} \qquad \boldsymbol{B}_{\mathrm{m}} = \boldsymbol{B}_{\mathrm{m}}(t,\boldsymbol{R}_{\alpha i}(t)) \tag{VIII.4}$$

are the "microphysical" electric and magnetic fields evaluated at the particle position $\boldsymbol{R}_{\alpha i}$ at time t. The adjective "microphysical" here is meant to indicate that $\boldsymbol{E}_{\mathrm{m}}$ and $\boldsymbol{B}_{\mathrm{m}}$ are the fields self-consistently generated by the particles themselves. These satisfy Maxwell's equations:

$$\boldsymbol{\nabla}\times\boldsymbol{E}_{\mathrm{m}} = -\frac{1}{c}\frac{\partial \boldsymbol{B}_{\mathrm{m}}}{\partial t}, \tag{VIII.5}$$

$$\boldsymbol{\nabla}\times\boldsymbol{B}_{\mathrm{m}} = \frac{1}{c}\frac{\partial \boldsymbol{E}_{\mathrm{m}}}{\partial t} + \frac{4\pi}{c}\sum_\alpha q_\alpha \int \mathrm{d}\boldsymbol{v}\, \boldsymbol{v}F_\alpha(t,\boldsymbol{r},\boldsymbol{v}), \tag{VIII.6}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{E}_{\mathrm{m}} = 4\pi\sum_\alpha q_\alpha \int \mathrm{d}\boldsymbol{v}\, F_\alpha(t,\boldsymbol{r},\boldsymbol{v}), \tag{VIII.7}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{B}_{\mathrm{m}} = 0. \tag{VIII.8}$$

Because Maxwell's equations are linear, we can add to these fields any that may be externally imposed: $\boldsymbol{E}_{\rm m} \to \boldsymbol{E}_{\rm m} + \boldsymbol{E}_{\rm ext}$ and $\boldsymbol{B}_{\rm m} \to \boldsymbol{B}_{\rm m} + \boldsymbol{B}_{\rm ext}$. This will be useful for describing magnetized plasmas threaded by an external magnetic field. Before we proceed any further, two things are worth noting:

(1) The electric and magnetic fields in (VIII.3) omit the contribution from particle (αi). In other words, a particle does not interact electromagnetically with itself.

(2) Writing $(\boldsymbol{r}, \boldsymbol{v})$ and $\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}$ all the time is exhausting. Denote $\boldsymbol{x} = (\boldsymbol{r}, \boldsymbol{v})$ and $\mathrm{d}\boldsymbol{x} = \mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{v}$, i.e., $\boldsymbol{x}$ is the phase-space coordinate and $\mathrm{d}\boldsymbol{x}$ is a small volume of phase space. Likewise, $\boldsymbol{X}_{\alpha i} = (\boldsymbol{R}_{\alpha i}, \boldsymbol{V}_{\alpha i})$.

Now, let us consider how $F_\alpha(t, \boldsymbol{x}) \doteq F_\alpha(t, \boldsymbol{r}, \boldsymbol{v})$ evolves:

$$
\begin{aligned}
\frac{\partial F_\alpha}{\partial t} &= \frac{\partial}{\partial t} \sum_{i=1}^{N_\alpha} \delta(\boldsymbol{x} - \boldsymbol{X}_{\alpha i}(t)) \\
&= \sum_{i=1}^{N_\alpha} \frac{\mathrm{d}\boldsymbol{X}_{\alpha i}}{\mathrm{d}t} \cdot \frac{\partial}{\partial \boldsymbol{X}_{\alpha i}} \delta(\boldsymbol{x} - \boldsymbol{X}_{\alpha i}(t)) \\
&= -\sum_{i=1}^{N_\alpha} \frac{\mathrm{d}\boldsymbol{X}_{\alpha i}}{\mathrm{d}t} \cdot \frac{\partial}{\partial \boldsymbol{x}} \delta(\boldsymbol{x} - \boldsymbol{X}_{\alpha i}(t)) \\
&= -\sum_{i=1}^{N_\alpha} \left\{ \boldsymbol{V}_{\alpha i} \cdot \boldsymbol{\nabla} + \frac{q_\alpha}{m_\alpha} \left[\boldsymbol{E}_{\rm m}(t, \boldsymbol{R}_{\alpha i}(t)) + \frac{\boldsymbol{V}_{\alpha i}}{c} \times \boldsymbol{B}_{\rm m}(t, \boldsymbol{R}_{\alpha i}(t)) \right] \cdot \frac{\partial}{\partial \boldsymbol{v}} \right\} \delta(\boldsymbol{x} - \boldsymbol{X}_{\alpha i}(t)) \\
&= -\sum_{i=1}^{N_\alpha} \left\{ \boldsymbol{v} \cdot \boldsymbol{\nabla} + \frac{q_\alpha}{m_\alpha} \left[\boldsymbol{E}_{\rm m}(t, \boldsymbol{r}) + \frac{\boldsymbol{v}}{c} \times \boldsymbol{B}_{\rm m}(t, \boldsymbol{r}) \right] \cdot \frac{\partial}{\partial \boldsymbol{v}} \right\} \delta(\boldsymbol{x} - \boldsymbol{X}_{\alpha i}(t)) \\
&= -\left\{ \boldsymbol{v} \cdot \boldsymbol{\nabla} + \frac{q_\alpha}{m_\alpha} \left[\boldsymbol{E}_{\rm m}(t, \boldsymbol{r}) + \frac{\boldsymbol{v}}{c} \times \boldsymbol{B}_{\rm m}(t, \boldsymbol{r}) \right] \cdot \frac{\partial}{\partial \boldsymbol{v}} \right\} \sum_{i=1}^{N_\alpha} \delta(\boldsymbol{x} - \boldsymbol{X}_{\alpha i}(t)) \\
&= -\left\{ \boldsymbol{v} \cdot \boldsymbol{\nabla} + \frac{q_\alpha}{m_\alpha} \left[\boldsymbol{E}_{\rm m}(t, \boldsymbol{r}) + \frac{\boldsymbol{v}}{c} \times \boldsymbol{B}_{\rm m}(t, \boldsymbol{r}) \right] \cdot \frac{\partial}{\partial \boldsymbol{v}} \right\} F_\alpha(\boldsymbol{t}, \boldsymbol{x})
\end{aligned}
$$

$$
\Longrightarrow \quad \boxed{\left[\frac{\partial}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} + \frac{q_\alpha}{m_\alpha} \left(\boldsymbol{E}_{\rm m} + \frac{\boldsymbol{v}}{c} \times \boldsymbol{B}_{\rm m} \right) \cdot \frac{\partial}{\partial \boldsymbol{v}} \right] F_\alpha(t, \boldsymbol{x}) = 0} \qquad \text{(VIII.9)}
$$

Equation (VIII.9) is called the *Klimontovich equation*. While it is equivalent to phase-space conservation, it is *not* a statistical equation. With proper initial conditions, it is completely deterministic. Together with Maxwell's equations (VIII.5)–(VIII.8), the densities and fields are determined for all time.

The Klimontovich equation (VIII.9) can be thought of as expressing the incompressibility of the substance $F_\alpha(t, \boldsymbol{x})$ as it moves in phase space: $\mathrm{D}F_\alpha/\mathrm{D}t = 0$, where $\mathrm{D}/\mathrm{D}t$ is the phase-space Lagrangian (i.e., comoving) derivative. Nicholson (1983) writes, "is it any wonder that a point particle is incompressible?" Phase-space trajectories that follow the characteristics of (VIII.9) and start from a region where $F_\alpha = 0$ will carry that null information along with them. Likewise with regions where $F_\alpha = 1$. Thus, the phase space is populated in a very choppy way. For that reason, as well as the simple fact that, despite some mathematics, we haven't actually simplified anything, the Klimontovich equation as a description of the plasma is not worth much practical use. It does, however, form the basis of a *statistical* description of the plasma. But, for that, we need some kind of averaging process…

VIII.2. The Liouville ("Leé-ooo-ville") distribution

Just as the microscopic state of a plasma is completely specified by the coordinates and momenta of its constituent particles, the statistical properties of the plasma are completely determined by the probabilistic distribution of said particles. Thus, we introduce the distribution function P_N of the coordinates and momenta of all of the $N \doteq \sum_\alpha N_\alpha$ particles in the system. Specifically,

$$P_N \prod_\alpha \mathrm{d}\boldsymbol{X}_{\alpha 1}\mathrm{d}\boldsymbol{X}_{\alpha 2}\dots\mathrm{d}\boldsymbol{X}_{\alpha N_\alpha}$$

gives the probability that, at time t, the phase-space coordinates of the particles of species α have the values $\boldsymbol{X}_{\alpha 1}$, $\boldsymbol{X}_{\alpha 2}$, …, $\boldsymbol{X}_{\alpha N_\alpha}$ in the range $\mathrm{d}\boldsymbol{X}_{\alpha 1}\mathrm{d}\boldsymbol{X}_{\alpha 2}\dots\mathrm{d}\boldsymbol{X}_{\alpha N_\alpha}$. This $6N$-dimensional phase space is called the "$\mathit{\Gamma}$ space". The microscopic state of the plasma is expressed in the $\mathit{\Gamma}$ space by a point $\{\boldsymbol{X}_{\alpha i}\}$. A few important points:

(1) The system points $\{\boldsymbol{X}_{\alpha i}\}$ do not interact with one another and so P_N satisfies a continuity equation of the Liouville kind:

$$\frac{\mathrm{D}P_N}{\mathrm{D}t} \doteq \frac{\partial P_N}{\partial t} + \sum_\alpha \sum_{i=1}^{N_\alpha} \frac{\mathrm{d}\boldsymbol{X}_{\alpha i}}{\mathrm{d}t}\cdot\frac{\partial P_N}{\partial \boldsymbol{X}_{\alpha i}} = 0; \tag{VIII.10}$$

i.e., the probability density is conserved along a characteristic trajectory in phase space.

(2) Because P_N is a probability, we have

$$\int \prod_\alpha \mathrm{d}\boldsymbol{X}_{\alpha 1}\mathrm{d}\boldsymbol{X}_{\alpha 2}\dots\mathrm{d}\boldsymbol{X}_{\alpha N_\alpha}\, P_N \doteq \int \mathrm{d}\boldsymbol{X}_{\rm all}\, P_N = 1,$$

where I've introduced the shorthand $\mathrm{d}\boldsymbol{X}_{\rm all}$ to indicate integration over all of the $\mathit{\Gamma}$ space (including all species).

(3) In thermodynamic equilibrium, P_N equals the Gibbs distribution

$$D_N \doteq \frac{1}{\mathcal{Z}}\exp\left(-\frac{\mathcal{H}}{T}\right), \tag{VIII.11}$$

where $\mathcal{H} = \mathcal{H}(\mathit{\Gamma})$ is the Hamiltonian (kinetic plus potential energy), T is the (species-independent!) equilibrium temperature (in energy units), and

$$\mathcal{Z} \doteq \int \prod_\alpha \mathrm{d}\boldsymbol{X}_{\alpha 1}\dots\mathrm{d}\boldsymbol{X}_{\alpha N_\alpha}\,\exp\left(-\frac{\mathcal{H}}{T}\right) \tag{VIII.12}$$

is the partition function. Plasmas are usually non-equilibrium systems, and so we will need to know how P_N evolves in time from a given starting distribution $P_N(0)$.

(4) It is profitable to think of P_N in the statistical-mechanics ensemble sense: imagine $\mathcal{N}$ replicas of our plasma, all macroscopically identical but microscopically different, with the system points $\{\boldsymbol{X}_{\alpha i}\}$ scattered over the $\mathit{\Gamma}$ space. Then P_N can be defined from

$$P_N \prod_\alpha \mathrm{d}\boldsymbol{X}_{\alpha 1}\mathrm{d}\boldsymbol{X}_{\alpha 2}\dots\mathrm{d}\boldsymbol{X}_{\alpha N_\alpha} \doteq \lim_{\mathcal{N}\to\infty}\frac{\mathcal{N}_s}{\mathcal{N}}, \tag{VIII.13}$$

where $\mathcal{N}_s$ is the number of those system points contained in an infinitesimal volume $\prod_\alpha \mathrm{d}\boldsymbol{X}_{\alpha 1}\dots\mathrm{d}\boldsymbol{X}_{\alpha N_\alpha}$ in the $\mathit{\Gamma}$ space around $\{\boldsymbol{X}_{\alpha i}\}$. (Why can we do this for a plasma? Hint: think about the accuracy of using a statistical description of an N-body system to describe any one realization of the system. What happens to the model's predictive power when N is not very large?) This codifies mathematically the experiment mentioned in the introduction to this Part: draw some initial conditions from a bucket, time advance them, record the outcome; draw another set of initial conditions from a bucket, time advance them, record the outcome; draw yet another set... do this procedure many, many, many times; then realize that calculating the mean outcome of all these experiments is statistical the same as evolving forward in time the probability distribution describing the likelihood of all possible initial conditions.

VIII.3. Reduced distribution functions

With a probability distribution in hand, we can perform an ensemble average over all these realizations of the plasma. Each one of these realizations is deterministic, but the system is stochastic between different realizations. Averaging amongst all these different realizations will turn our spiky "fine-grained" F_α into the smooth "coarse-grained" distribution. For example,

$$\int \mathrm{d}\boldsymbol{X}_{\alpha 2} \dots \mathrm{d}\boldsymbol{X}_{\alpha N_\alpha} \prod_\beta \mathrm{d}\boldsymbol{X}_{\beta 1} \mathrm{d}\boldsymbol{X}_{\beta 2} \dots \mathrm{d}\boldsymbol{X}_{\beta N_\beta}\, P_N$$

is the joint probability that particle $\alpha 1$ has coordinates in $(\boldsymbol{X}_{\alpha 1})$ to $(\boldsymbol{X}_{\alpha 1} + \mathrm{d}\boldsymbol{X}_{\alpha 1})$ *irrespective* of the coordinates of particles $\alpha 2$, ..., αN_α, $\beta 1$, $\beta 2$, ..., βN_β, etc. This *reduced distribution function* is called the *one-particle distribution function*. It can be normalized to one's tastes. I choose the following:[14]

$$f_\alpha(t, \boldsymbol{x}) \doteq N_\alpha \int \mathrm{d}\boldsymbol{X}_{\alpha 2} \dots \mathrm{d}\boldsymbol{X}_{\alpha N_\alpha} \prod_\beta \mathrm{d}\boldsymbol{X}_{\beta 1} \mathrm{d}\boldsymbol{X}_{\beta 2} \dots \mathrm{d}\boldsymbol{X}_{\beta N_\beta}\, P_N, \tag{VIII.14}$$

The operative word here is "irrespective". Of course the probability of, say, an electron being at some phase-space position $\boldsymbol{x}$ is impacted by an ion being nearby at $\boldsymbol{x}' \approx \boldsymbol{x}$, but this information is not is f_α. The influence of a near neighbor on the distribution of a particle is contained in a less reduced description, e.g., the *two-particle distribution function*:

$$f_{\alpha\beta}(t, \boldsymbol{x}, \boldsymbol{x}') \doteq N_\alpha N_\beta \int \mathrm{d}\boldsymbol{X}_{\alpha 2} \dots \mathrm{d}\boldsymbol{X}_{\alpha N_\alpha} \mathrm{d}\boldsymbol{X}_{\beta 2} \dots \mathrm{d}\boldsymbol{X}_{\beta N_\beta} \prod_\gamma \mathrm{d}\boldsymbol{X}_{\gamma 1} \mathrm{d}\boldsymbol{X}_{\gamma 2} \dots \mathrm{d}\boldsymbol{X}_{\gamma N_\gamma}\, P_N. \tag{VIII.15}$$

Then $f_{\alpha\beta}(t, \boldsymbol{x}, \boldsymbol{x}')\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{x}'/N_\alpha N_\beta$ is the joint probability that particle $\alpha 1$ is at $\boldsymbol{x}$ in interval $\mathrm{d}\boldsymbol{x}$ *and* particle $\beta 1$ is at $\boldsymbol{x}'$ in interval $\mathrm{d}\boldsymbol{x}'$, irrespective of all other particles.

Note three things:

(1) The species labels α and β could refer to the same type of particle ($\alpha = \beta$), in which case $\beta 1 \to \alpha 2$. (Particles are indistinguishable amongst a species, and so the exact numerical indices do not matter.) In this case, $N_\beta \to N_\alpha - 1$.

(2) The two-particle distribution function $f_{\alpha\beta}$ is still a reduced distribution, but, as opposed to the one-particle distribution function f_α, it contains some information

[14] The reason for the N_α is so that $\int \mathrm{d}\boldsymbol{v}\, f_\alpha(t, \boldsymbol{x})$ is the number density n_α, a customary normalization for the one-particle distribution function. Others might introduce a prefactor $\mathcal{V}$ for volume, which makes $\int \mathrm{d}\boldsymbol{v}\, f_\alpha(t, \boldsymbol{x})$ equal to the fraction of the mean number density $\overline{n}_\alpha \doteq N_\alpha/\mathcal{V}$ in that volume.

about two-body interactions. If the particles do not interact, then $f_{\alpha\beta} = f_\alpha f_\beta$, the product of one-particle distribution functions... simple.

(3) One could of course generalize this process. For example, the three-particle distribution function is

$$f_{\alpha\beta\gamma} \doteq N_\alpha N_\beta N_\gamma \int \frac{\mathrm{d}\boldsymbol{X}_{\rm all}}{\mathrm{d}\boldsymbol{X}_{\alpha 1}\mathrm{d}\boldsymbol{X}_{\beta 1}\mathrm{d}\boldsymbol{X}_{\gamma 1}}\, P_N; \tag{VIII.16}$$

the four-particle distribution function is

$$f_{\alpha\beta\gamma\delta} \doteq N_\alpha N_\beta N_\gamma N_\delta \int \frac{\mathrm{d}\boldsymbol{X}_{\rm all}}{\mathrm{d}\boldsymbol{X}_{\alpha 1}\mathrm{d}\boldsymbol{X}_{\beta 1}\mathrm{d}\boldsymbol{X}_{\gamma 1}\mathrm{d}\boldsymbol{X}_{\delta 1}}\, P_N;$$

and so on.

We combine this machinery with the Klimontovich distribution (VIII.1) as follows.

Each term in $F_\alpha = \sum_i \delta(\boldsymbol{x} - \boldsymbol{X}_{\alpha i})$ describes the location of a particle in terms of its initial conditions, and P_N describes the probability of a particle having a certain set of initial conditions, and so the reduced descriptions of P_N can be expressed in terms of the averages of products of F_α over all possible initial conditions. These averages are defined by

$$\langle G(F_\alpha, F_\beta, \dots, F_\gamma)\rangle \doteq \int \mathrm{d}\boldsymbol{X}_{\rm all}\, P_N\, G(F_\alpha, F_\beta, \dots, F_\gamma). \tag{VIII.17}$$

Let's put this to work.

VIII.4. Towards the Vlasov equation

Integrate the Klimontovich distribution (VIII.1) over the Liouville distribution (see (VIII.17)):

$$\begin{aligned}
\langle F_\alpha(t,\boldsymbol{x})\rangle &\doteq \sum_{i=1}^{N_\alpha} \int \mathrm{d}\boldsymbol{X}_{\rm all}\, P_N\, \delta(\boldsymbol{x} - \boldsymbol{X}_{\alpha i}(t)) \\
&= N_\alpha \int \mathrm{d}\boldsymbol{X}_{\rm all}\, P_N\, \delta(\boldsymbol{x} - \boldsymbol{X}_{\alpha 1}(t)) \quad \text{(particles are indistinguishable)} \\
&= N_\alpha \int \mathrm{d}\boldsymbol{X}_{\alpha 2}\dots\mathrm{d}\boldsymbol{X}_{\alpha N_\alpha} \prod_\beta \mathrm{d}\boldsymbol{X}_{\beta 1}\mathrm{d}\boldsymbol{X}_{\beta 2}\dots\mathrm{d}\boldsymbol{X}_{\beta N_\beta}\, P_N \\
&\doteq f_\alpha(t,\boldsymbol{x}) \quad \text{(def'n of one-particle distribution function, (VIII.14)).}
\end{aligned} \tag{VIII.18}$$

Similarly, the average electromagnetic fields are obtained by averaging the microscopic fields $\boldsymbol{E}_{\rm m}$ and $\boldsymbol{B}_{\rm m}$, which depend upon the positions of the (point-like) particles, over the probable locations of all of the particles:

$$\boldsymbol{E} \doteq \langle \boldsymbol{E}_{\rm m}\rangle = \int \mathrm{d}\boldsymbol{X}_{\rm all}\, P_N\, \boldsymbol{E}_{\rm m} \qquad \text{and} \qquad \boldsymbol{B} \doteq \langle \boldsymbol{B}_{\rm m}\rangle = \int \mathrm{d}\boldsymbol{X}_{\rm all}\, P_N\, \boldsymbol{B}_{\rm m}. \tag{VIII.19}$$

Using (VIII.18) and (VIII.19) in the Maxwell equations (VIII.5)–(VIII.8) gives

$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t}, \tag{VIII.20}$$

$$\nabla \times \boldsymbol{B} = \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} + \frac{4\pi}{c}\sum_\alpha q_\alpha \int \mathrm{d}\boldsymbol{v}\, \boldsymbol{v} f_\alpha(t,\boldsymbol{r},\boldsymbol{v}), \tag{VIII.21}$$

$$\nabla \cdot \boldsymbol{E} = 4\pi \sum_\alpha q_\alpha \int \mathrm{d}\boldsymbol{v}\, f_\alpha(t,\boldsymbol{r},\boldsymbol{v}), \tag{VIII.22}$$

$$\nabla \cdot \boldsymbol{B} = 0. \tag{VIII.23}$$

Simple. This is because Maxwell's equations are linear.

The difficulty is that the Klimontovich equation (VIII.9) is *not*. It has a quadratic nonlinearity, which it what makes it so hard to solve. Let's see that. The integral of (VIII.9) over the Liouville distribution is

$$\frac{\partial}{\partial t}\langle F_\alpha\rangle + \boldsymbol{v}\cdot\nabla\langle F_\alpha\rangle + \left\langle \frac{q_\alpha}{m_\alpha}\left(\boldsymbol{E}_\mathrm{m} + \frac{\boldsymbol{v}}{c}\times\boldsymbol{B}_\mathrm{m}\right)\cdot\frac{\partial F_\alpha}{\partial \boldsymbol{v}}\right\rangle = 0. \tag{VIII.24}$$

The first two terms in (VIII.24) involve only the one-particle distribution function f_α (see (VIII.18)). The third and final term can be manipulated further by decomposing the microscopic electromagnetic fields into their statistical means and deviations:

$$\boldsymbol{E}_\mathrm{m} = \langle \boldsymbol{E}_\mathrm{m}\rangle + \delta\boldsymbol{E} \doteq \boldsymbol{E} + \delta\boldsymbol{E} \qquad \text{and} \qquad \boldsymbol{B}_\mathrm{m} = \langle \boldsymbol{B}_\mathrm{m}\rangle + \delta\boldsymbol{B} \doteq \boldsymbol{B} + \delta\boldsymbol{B}. \tag{VIII.25}$$

The fields $\boldsymbol{E}$ and $\boldsymbol{B}$ are smooth and coarse-grained; they are the "macroscopic" fields obtained by averaging the microscopic fields over all possible positions of the plasma particles, weighted by the Liouville distribution. The remainders, $\delta\boldsymbol{E}$ and $\delta\boldsymbol{B}$, are spiky and fine-grained; they capture the influence of the discrete nature of the particles on the electromagnetic fields. Using (VIII.25) in the Klimontovich equation (VIII.24) and likewise writing $F_\alpha = f_\alpha + \delta F_\alpha$, we obtain

$$\boxed{\left[\frac{\partial}{\partial t} + \boldsymbol{v}\cdot\nabla + \frac{q_\alpha}{m_\alpha}\left(\boldsymbol{E} + \frac{\boldsymbol{v}}{c}\times\boldsymbol{B}\right)\cdot\frac{\partial}{\partial \boldsymbol{v}}\right] f_\alpha(t,\boldsymbol{x}) = -\left\langle \frac{q_\alpha}{m_\alpha}\left(\delta\boldsymbol{E} + \frac{\boldsymbol{v}}{c}\times\delta\boldsymbol{B}\right)\cdot\frac{\partial \delta F_\alpha}{\partial \boldsymbol{v}}\right\rangle} \tag{VIII.26}$$

(If there are externally imposed electric and magnetic fields, they can be added to $\boldsymbol{E}$ and $\boldsymbol{B}$, respectively.) The one-particle distribution function evolves because particles move around in configuration space and accelerate in velocity space (the left-hand side), and because of correlations between discrete particles and spiky electromagnetic fields (the right-hand side).

Before deciding what to do with the right-side of (VIII.26), recall that, on scales $L \gtrsim \lambda_\mathrm{D}$, individual particle particles are shielded and what remains are fields due to the collective action of a large number of particles. Also recall that the Coulomb potential is long-range, and so the fields decay on distances long compared to the interparticle spacing ($\lambda_\mathrm{D} \gg \delta r$). This gives collective behavior: interaction of individual particles with the mean ("macroscopic") fields generated by all other particles. This means that the entire left-hand side of (VIII.26) consists of terms that vary smoothly in phase space, since it's entirely insensitive to the discrete nature of the plasma. The right-hand side, by contrast, is very sensitive, and is ultimately responsible for collisional effects. And with $\delta\boldsymbol{E}$ and $\delta\boldsymbol{B}$ depending on δF_α through the Coulomb and Biot-Savart laws, this right-hand side is *quadratic* in δF. Yuck.

VIII.5. Solving the BBGKY hierarchy

The problem is that this quadratic right-hand side of (VIII.26) cannot be expressed solely in terms of f_α, and so we have here what is known as a *closure problem*. Through many manipulations (see §§II.4, II.5 of these notes if you're interested), one can actually write the right-hand side in terms of the two-particle distribution function $f_{\alpha\beta}$. That may seem like progress, but the two-particle distribution depends on the three-particle distribution $f_{\alpha\beta\gamma}$, and the three-particle distribution depends on the four-particle distribution $f_{\alpha\beta\gamma\delta}$, etc. It never ends. This is called the *BBGKY hierarchy*, named after Bogoliubov, Born, Green, Kirkwood, & Yvon (1935–1949).

At this point, it's worth reiterating the definitions of f_α and $f_{\alpha\beta}$. f_α is the one-particle distribution function – the probability that a particle of species α has phase-space position $\boldsymbol{x}$ at time t in the interval $\mathrm{d}\boldsymbol{x}$ *regardless* of all other particles. No particle–particle interactions are encoded in f_α. $f_{\alpha\beta}$, on the other hand, is the joint probability that a particle of species α has phase-space position $\boldsymbol{x}$ at time t *and* a particle of species β has phase-space position $\boldsymbol{x}'$ at time t, *regardless* of all other particles. Now, suppose all particles were truly uncorrelated (i.e., no collisions). Then $f_{\alpha\beta}(t,\boldsymbol{x},\boldsymbol{x}') = f_\alpha(t,\boldsymbol{x})f_\beta(t,\boldsymbol{x}')$, and the right-hand side of (VIII.26) would vanish. This would return the *Vlasov equation*,

$$\boxed{\dot{f}_\alpha \doteq \left[\frac{\partial}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla} + \frac{q_\alpha}{m_\alpha}\left(\boldsymbol{E} + \frac{\boldsymbol{v}}{c}\times\boldsymbol{B}\right)\cdot\frac{\partial}{\partial \boldsymbol{v}}\right] f_\alpha(t,\boldsymbol{x}) = 0 \quad \text{(Vlasov equation)}} \tag{VIII.27}$$

Again, the one-particle distribution function evolves because particles move around in configuration space and accelerate in velocity space, but this time the only interactions that each particle has with other particles is indirect, through the coarse-grained, smooth, collective electromagnetic fields.

This suggests that we introduce some function, say, $g_{\alpha\beta}(t,\boldsymbol{x},\boldsymbol{x}')$, which captures the difference between $f_{\alpha\beta}$ and $f_\alpha f_\beta$:

$$f_{\alpha\beta} = f_\alpha f_\beta + g_{\alpha\beta}. \tag{VIII.28}$$

This is the first step in what is known as the *Mayer cluster (or cumulant) expansion*. It splits the statistically independent pieces of $f_{\alpha\beta}$, which have multiplicative probabilities, apart from the statistically dependent piece. The difference is the *two-particle correlation function*. It's almost always useful to split off the piece of a joint probability distribution that corresponds to uncorrelated events. Nicholson (1983) on page 54 of his textbook has a cute analogy concerning correlated and uncorrelated coin tosses and die rolls. I prefer Yahtzee: the difference between rolling each die separately versus putting them all in the can and shaking them all and rolling them all out at the same time, so that their mutual collisions influence which side of each die faces up when the system comes to rest. Whatever you prefer, the fact that the right-hand side of (VIII.26) can be written succinctly in terms of the two-particle correlation function, *viz.*

$$\boxed{\dot{f}_\alpha = \sum_\beta \frac{q_\alpha q_\beta}{m_\alpha}\int \mathrm{d}\boldsymbol{x}'\,\frac{\partial}{\partial \boldsymbol{r}}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\cdot\frac{\partial}{\partial \boldsymbol{v}} g_{\alpha\beta}(t,\boldsymbol{x},\boldsymbol{x}')} \tag{VIII.29}$$

attests to the fact that "collisions" are, in fact, correlations established by Coulomb interactions in the probabilities that two particles will be found near one another.

There is a natural small parameter in a weakly coupled plasma:

$$\Lambda^{-1} \doteq (n\lambda_{\mathrm{D}}^3)^{-1} \lll 1; \tag{VIII.30}$$

i.e., there are many particles in a Debye sphere. Recall that this also means that the average potential energy of the plasma is small compared to the average kinetic energy. To the extent that the potential energy due to interactions can be neglected, the plasma behaves like an ideal gas; thus, Λ^{-1} measures the size of departures of the thermodynamic properties of the plasma from those of an ideal gas.

Before explaining what this means for our BBGKY hierarchy, let us compare this situation with that of a gas of neutral particles. In that situation, the range of the interaction force r_0 is much smaller than the mean spacing δr of the particles $\sim n^{-1/3}$. Then it makes sense to expand particle correlations in the small parameter nr_0^3, and thus neglect the triple correlation. In other words, particle–particle collisions are sufficiently rare due to the small cross section that three-body collisions are much rarer than two-body collisions, with the presence of a third body affecting the collision between two bodies at an asymptotically small level. In a plasma, by contrast, $r_0 \approx \lambda_{\rm D} \gg n^{-1/3}$ implies $nr_0^3 \gg 1$. This is because Debye screening limits the range of the interaction potential, but to a value that is still large compared to the average interparticle separation (i.e., the Coulomb force is long range compared to the scattering force of direct two-body collisions, but has its long range attenuated by Debye screening). However, this does not mean that three-body interactions are more important than two-body interactions, despite $nr_0^3 \gg 1$ for a plasma. This is because, even though a charged particle is interacting with all the particles in its Debye sphere and thus undergoes $\sim\Lambda$ simultaneous Coulomb collisions, such collisions are *weak*, in the sense that the effect of, say, particle A on particle B's orbit is small enough that the collision between particle B and another particle C is practically unaffected. This is because collisions in an ionized plasma result in small-angle (rather than large-angle) deflections. Another way of saying this is that the joint distribution $f_{\alpha\beta}$ of two particles in a small volume ($n^{-1} \ll \mathcal{V} \ll \lambda_{\rm D}^3$) is determined by the many particles outside of the volume rather than by the separation of the two particles from one another; i.e., $f_{\alpha\beta} \approx f_\alpha f_\beta$.

What this means for our kinetic theory is that, ultimately, three-particle (and higher) correlations can be dropped without much consequence, and $g_{\alpha\beta}$ can be written in terms of f_α and f_β. This closes the hierarchy, and allows one to write the right-hand side of (VIII.29) in a relatively simple form that is referred to as the *collision operator*. This operator encodes the impact of the discrete nature of charged particles, and the electric noise generated by this discreteness, on the otherwise smooth trajectories of the particles. The next section offers a few words on such collisions and what they imply for the irreversibility of a weakly coupled plasma.

VIII.6. A brief primer on collisions and irreversibility

First, take another look at the Vlasov equation (VIII.27), rewritten here:

$$\dot{f}_\alpha(t,\boldsymbol{x}) \doteq \left[\frac{\partial}{\partial t} + \boldsymbol{v}\boldsymbol{\cdot}\boldsymbol{\nabla} + \frac{q_\alpha}{m_\alpha}\left(\boldsymbol{E} + \frac{\boldsymbol{v}}{c}\times\boldsymbol{B}\right)\boldsymbol{\cdot}\frac{\partial}{\partial\boldsymbol{v}}\right] f_\alpha(t,\boldsymbol{x}) = 0. \tag{VIII.31}$$

Note that setting $t \rightarrow t' \doteq -t$, $\boldsymbol{r} \rightarrow \boldsymbol{r}' \doteq \boldsymbol{r}$, $\boldsymbol{v} \rightarrow \boldsymbol{v}' \doteq -\boldsymbol{v}$, $f_\alpha \rightarrow f'_\alpha \doteq f_\alpha(-t,\boldsymbol{r},-\boldsymbol{v})$, $\boldsymbol{E} \rightarrow \boldsymbol{E}' \doteq \boldsymbol{E}(-t,\boldsymbol{r})$, and $\boldsymbol{B} \rightarrow \boldsymbol{B}' \doteq -\boldsymbol{B}(-t,\boldsymbol{r})$ in (VIII.31) changes nothing:

$$\left[\frac{\partial}{\partial t'} + \boldsymbol{v}'\boldsymbol{\cdot}\boldsymbol{\nabla}' + \frac{q_\alpha}{m_\alpha}\left(\boldsymbol{E}' + \frac{\boldsymbol{v}'}{c}\times\boldsymbol{B}'\right)\boldsymbol{\cdot}\frac{\partial}{\partial\boldsymbol{v}'}\right] f'_\alpha(t',\boldsymbol{x}') = 0.$$

Thus, the Vlasov–Maxwell set of equations is *time-reversible*. All information about the phase-space fluid elements is preserved for all time.

Next, calculate the evolution of the entropy,

$$\mathcal{S} \doteq -\sum_\alpha \int \mathrm{d}\boldsymbol{x}\, f_\alpha(t,\boldsymbol{x}) \ln f_\alpha(t,\boldsymbol{x}), \tag{VIII.32}$$

in a Vlasov plasma:

$$\dot{\mathcal{S}} = -\sum_\alpha \int \mathrm{d}\boldsymbol{x}\, \dot{f}_\alpha \bigl(1 + \ln f_\alpha\bigr) = 0. \tag{VIII.33}$$

Entropy is constant. What a comforting thought.

Of course, these things aren't generally true. The world is not time-reversible, no matter how much we wish it to be so. In dropping

$$-\left\langle \frac{q_\alpha}{m_\alpha}\left(\delta\boldsymbol{E} + \frac{\boldsymbol{v}}{c}\times\delta\boldsymbol{B}\right)\boldsymbol{\cdot}\frac{\partial\delta F_\alpha}{\partial\boldsymbol{v}}\right\rangle$$

from the right-hand side of the Liouville-averaged Klimontovich equation (see (VIII.26)), we have lost entropy-increasing collisional dissipation and irreversibility. It's actually a lot of work to (rigorously or not) derive the appropriate collision operator, Balescu–Lenard (Balescu 1960; Lenard 1960) or Landau (Landau 1937), and so I'll simply write down the latter, as if it's entirely obvious from whence it came:

$$C[f_\alpha] = \sum_\beta \frac{2\pi q_\alpha^2 q_\beta^2 \ln\lambda_{\alpha\beta}}{m_\alpha}\frac{\partial}{\partial\boldsymbol{v}}\boldsymbol{\cdot}\int \mathrm{d}\boldsymbol{v}'\,\boldsymbol{U}(\boldsymbol{v}-\boldsymbol{v}')\boldsymbol{\cdot}\left(\frac{1}{m_\alpha}\frac{\partial}{\partial\boldsymbol{v}} - \frac{1}{m_\beta}\frac{\partial}{\partial\boldsymbol{v}'}\right) f_\alpha(\boldsymbol{v}) f_\beta(\boldsymbol{v}'), \tag{VIII.34}$$

where $\boldsymbol{U}(\boldsymbol{u}) \doteq (u^2\boldsymbol{I} - \boldsymbol{u}\boldsymbol{u})/u^3$ and $\ln\lambda_{\alpha\beta}$ is the Coulomb logarithm (see the NRL plasma formulary). The details of all this can be found here. In the meantime, here are some properties that any rigorously derived collision operator ought to satisfy:

- If $f_\alpha \geqslant 0$ at $t=0$, then $f_\alpha \geqslant 0$ for all time t.
- Particle number is conserved:

  $$\int \mathrm{d}\boldsymbol{v}\, C[f_\alpha] = 0 \quad \text{for each } \alpha.$$

- Total momentum is conserved:

  $$\sum_\alpha \int \mathrm{d}\boldsymbol{v}\, m_\alpha \boldsymbol{v}\, C[f_\alpha] = 0.$$

  (NB: momentum of each individual species is *not* conserved. Newton would have a problem with that.)
- Total kinetic energy is conserved:

  $$\sum_\alpha \int \mathrm{d}\boldsymbol{v}\, \frac{1}{2} m_\alpha v^2\, C[f_\alpha] = 0.$$

  (NB: again, this holds only for the entire plasma, not each species by itself.)
- The entropy $\mathcal{S}$ (see (VIII.32)) can either increase or remain constant under the action of the collision operator. It *cannot* decrease!
- Maxwell distributions for all species with equal temperatures and mean velocities are a time-independent solution:

  $$f_{\mathrm{M},\alpha} = \frac{n_\alpha}{\pi^{3/2} v_{\mathrm{th}\alpha}^3}\exp\left(-\frac{|\boldsymbol{v}-\boldsymbol{u}|^2}{v_{\mathrm{th}\alpha}^2}\right), \quad v_{\mathrm{th}\alpha}^2 \doteq \frac{2T}{m_\alpha}$$

with the same temperature T and mean velocity $\boldsymbol{u}$ for all α.

- As $t \to \infty$, any f_α satisfying $f_\alpha \geqslant 0$ approaches a Maxwell distribution with equal temperatures and mean velocities for all species.

It's left as an exercise to the reader to show that (VIII.34) satisfies all of these properties.

In practice, the Balescu–Lenard and Landau collision operators are rarely used. When collision aren't simply thrown out altogether, various simplified operators are chosen, mostly based on their analytical and/or numerical tractability, but also because some of them can be obtained rigorously as certain limits of the full Landau operator. Here are just a few, which are provided mainly so that you recognize them if they ever cross into your future light cone:

Bhatnagar–Gross–Krook (BGK): $C[f_\alpha] = -\nu(f_\alpha - f_{\mathrm{M},\alpha})$

Lenard–Bernstein:
$$C[f_\alpha] = \nu \frac{\partial}{\partial \boldsymbol{v}} \cdot \Bigg[\underbrace{(\boldsymbol{v} - \boldsymbol{u}) f_\alpha}_{\text{drag}} + \underbrace{\frac{v_{\mathrm{th}\alpha}^2}{2} \frac{\partial f_\alpha}{\partial \boldsymbol{v}}}_{\text{diffusion}} \Bigg]$$

Lorentz:
$$C[f_\alpha] = \nu \mathcal{L}[f_\alpha]$$
$$\text{where } \mathcal{L} \doteq \underbrace{\frac{1}{2}\left[\frac{\partial}{\partial \xi}(1-\xi^2)\frac{\partial}{\partial \xi} + \frac{1}{1-\xi^2}\frac{\partial}{\partial \phi_v^2} \right]}_{\text{velocity pitch-angle diffusion; } \xi \doteq \cos\theta_v}$$

Fokker–Planck:
$$C[f_\alpha] = \underbrace{-\frac{\partial}{\partial \boldsymbol{v}} \cdot \big[\boldsymbol{A}_\alpha(\boldsymbol{v}) f_\alpha \big]}_{\text{drag}} + \underbrace{\frac{1}{2} \frac{\partial}{\partial \boldsymbol{v}} \frac{\partial}{\partial \boldsymbol{v}} : \big[\boldsymbol{B}_\alpha(\boldsymbol{v}) f_\alpha \big]}_{\text{diffusion}} .$$

The velocity-dependent Fokker–Planck coefficients $\boldsymbol{A}_\alpha$ and $\boldsymbol{B}_\alpha$ are related to the "jump moments" $\langle \Delta \boldsymbol{v} \rangle$ and $\langle \Delta \boldsymbol{v} \Delta \boldsymbol{v} \rangle$, which are expectation values for changes and correlations in particle velocities over a short (but not too short – there is a Markov assumption involved) interval of time.

VIII.7. Moments of the kinetic equation

Accepting that there *is* a collision operator – whatever it is – one may proceed to take moments of the kinetic equation to obtain "fluid" equations. Start with the Vlasov–Landau equation, repeated here for convenience:

$$\dot{f}_\alpha \doteq \frac{\partial f_\alpha}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} f_\alpha + \frac{q_\alpha}{m_\alpha}\Big(\boldsymbol{E} + \frac{\boldsymbol{v}}{c} \times \boldsymbol{B} \Big) \cdot \frac{\partial f_\alpha}{\partial \boldsymbol{v}} = C[f_\alpha].$$

We could of course add additional forces on the charged particles, such as that due to gravity, $m_\alpha \boldsymbol{g}$. Since we'll use quasi-neutrality to eliminate $\boldsymbol{E}$ at some point, let's do that:

$$\dot{f}_\alpha \doteq \frac{\partial f_\alpha}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} f_\alpha + \left[\frac{q_\alpha}{m_\alpha}\Big(\boldsymbol{E} + \frac{\boldsymbol{v}}{c} \times \boldsymbol{B} \Big) + \boldsymbol{g} \right] \cdot \frac{\partial f_\alpha}{\partial \boldsymbol{v}} = C[f_\alpha]. \tag{VIII.35}$$

Now, $\boldsymbol{v}$ contains both thermal and mean velocities. It is useful to split them apart (e.g., because they might have very different magnitudes):

$$\boldsymbol{w} \doteq \boldsymbol{v} - \boldsymbol{u}_\alpha(t, \boldsymbol{r}), \tag{VIII.36}$$

where

$$\boldsymbol{u}_\alpha(t,\boldsymbol{r}) \doteq \frac{1}{n_\alpha}\int \mathrm{d}\boldsymbol{v}\,\boldsymbol{v} f_\alpha, \quad n_\alpha(t,\boldsymbol{r}) \doteq \int \mathrm{d}\boldsymbol{v}\, f_\alpha. \tag{VIII.37}$$

Enacting this transformation of variables, $f_\alpha(t,\boldsymbol{r},\boldsymbol{v}) \to f_\alpha(t,\boldsymbol{r},\boldsymbol{w})$, through the use of

$$\left.\frac{\partial}{\partial t}\right|_{\boldsymbol{v}} = \left.\frac{\partial}{\partial t}\right|_{\boldsymbol{w}} + \left.\frac{\partial \boldsymbol{w}}{\partial t}\right|_{\boldsymbol{v}} \cdot \frac{\partial}{\partial \boldsymbol{w}} = \left.\frac{\partial}{\partial t}\right|_{\boldsymbol{w}} - \frac{\partial \boldsymbol{u}_\alpha}{\partial t}\cdot\frac{\partial}{\partial \boldsymbol{w}}, \tag{VIII.38}$$

$$\left.\frac{\partial}{\partial \boldsymbol{r}}\right|_{\boldsymbol{v}} = \left.\frac{\partial}{\partial \boldsymbol{r}}\right|_{\boldsymbol{w}} + \left.\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{r}}\right|_{\boldsymbol{v}} \cdot \frac{\partial}{\partial \boldsymbol{w}} = \left.\frac{\partial}{\partial \boldsymbol{r}}\right|_{\boldsymbol{w}} - \frac{\partial \boldsymbol{u}_\alpha}{\partial \boldsymbol{r}}\cdot\frac{\partial}{\partial \boldsymbol{w}}, \tag{VIII.39}$$

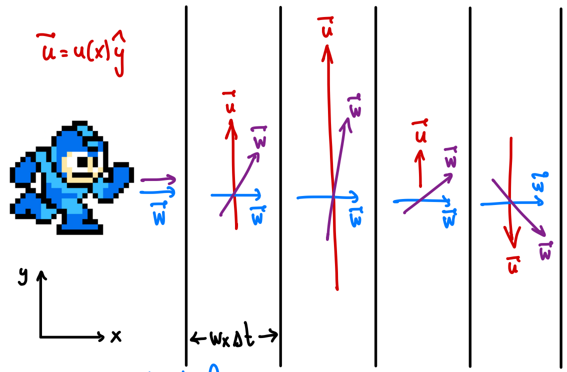
equation (VIII.35) becomes

$$\frac{\mathrm{D}f_\alpha}{\mathrm{D}t_\alpha} + \boldsymbol{w}\cdot\boldsymbol{\nabla} f_\alpha + \Bigg[\underbrace{\frac{q_\alpha}{m_\alpha}\left(\boldsymbol{E} + \frac{\boldsymbol{u}_\alpha}{c}\times\boldsymbol{B}\right) + \boldsymbol{g} - \frac{\mathrm{D}\boldsymbol{u}_\alpha}{\mathrm{D}t_\alpha}}_{\doteq\, \boldsymbol{a}_\alpha(t,\boldsymbol{r})} + \frac{q_\alpha}{m_\alpha}\left(\frac{\boldsymbol{w}}{c}\times\boldsymbol{B}\right) - \boldsymbol{w}\cdot\boldsymbol{\nabla}\boldsymbol{u}_\alpha\Bigg]\cdot\frac{\partial f_\alpha}{\partial \boldsymbol{w}}$$
$$= C[f_\alpha], \tag{VIII.40}$$

where

$$\frac{\mathrm{D}}{\mathrm{D}t_\alpha} \doteq \frac{\partial}{\partial t} + \boldsymbol{u}_\alpha\cdot\boldsymbol{\nabla} \tag{VIII.41}$$

is the Lagrangian time derivative taken in the frame comoving with the mean velocity $\boldsymbol{u}_\alpha$ of species α. The additional acceleration terms in (VIII.40) that result from the frame transformation, *viz.* $\mathrm{D}\boldsymbol{u}_\alpha/\mathrm{D}t_\alpha$ and $\boldsymbol{w}\cdot\boldsymbol{\nabla}\boldsymbol{u}_\alpha$, are the result of boosting to a time- and space-dependent frame. The former term is fairly self-explanatory – particles must be accelerated so as to continue residing in the "fluid element" they comprise, which is itself being accelerated by various (magneto)hydrodynamic forces that result in $\mathrm{D}\boldsymbol{u}_\alpha/\mathrm{D}t_\alpha$– but the latter deserves some discussion. Imagine you are trying to walk at constant velocity $\boldsymbol{w} = w\hat{\boldsymbol{x}}$ across several layers of differentially moving conveyor belts with velocities $\boldsymbol{u} = u(x)\hat{\boldsymbol{y}}$, as in the figure below. In your frame (and the frame of the conveyor belts), your velocity will always be $w\hat{\boldsymbol{x}}$. But, in the lab frame, your velocity will include the velocity of the conveyor belts. This means that, extra time you step onto a new conveyor belt that has some velocity oriented in the y direction that is different from that of the last conveyor belt, you will be accelerating in the lab frame. That is, your velocity in the lab frame will change over an interval of time from one conveyor belt to the next. Mathematically, the figure below corresponds to an acceleration $w\Delta u_y/\Delta x$ every time you step from one conveyor belt at position x with velocity $u\hat{\boldsymbol{y}}$ to another conveyor belt at position $x+\Delta x$ with velocity $(u+\Delta u)\hat{\boldsymbol{y}}$. The difference between these two points of view is enacted by adding $-\boldsymbol{w}\cdot\boldsymbol{\nabla}\boldsymbol{u}_\alpha$ to the acceleration term of (VIII.40).



$\vec{w}$ in conveyor-belt frame is $\longrightarrow \vec{w}$, always
$\vec{w}$ in lab frame changes by $\Delta\vec{w} = (\vec{w}\cdot\vec{\nabla}\vec{u})\Delta t = \left(w_x \frac{du}{dx}\hat{y}\right)\Delta t$
with every timestep Δt.

Next, take those moments:

$$\int \mathrm{d}\boldsymbol{w}\,(\text{VIII.40}):\quad \frac{\mathrm{D}}{\mathrm{D}t_\alpha}\cancelto{n_\alpha}{\int \mathrm{d}\boldsymbol{w}\, f_\alpha} + \cancelto{0}{\int \mathrm{d}\boldsymbol{w}\,\boldsymbol{w}\cdot\boldsymbol{\nabla} f_\alpha} + \boldsymbol{a}_\alpha\cdot\cancelto{0}{\int \mathrm{d}\boldsymbol{w}\,\frac{\partial f_\alpha}{\partial \boldsymbol{w}}}$$

$$+ \frac{q_\alpha}{m_\alpha}\cancelto{0}{\int \mathrm{d}\boldsymbol{w}\left(\frac{\boldsymbol{w}}{c}\times\boldsymbol{B}\right)\cdot\frac{\partial f_\alpha}{\partial\boldsymbol{w}}} - \underbrace{\int \mathrm{d}\boldsymbol{w}\,(\boldsymbol{w}\cdot\boldsymbol{\nabla}\boldsymbol{u}_\alpha)\cdot\frac{\partial f_\alpha}{\partial \boldsymbol{w}}}_{\overset{\text{bp}}{=}\, -(\boldsymbol{\nabla}\cdot\boldsymbol{u}_\alpha)\cancelto{n_\alpha}{\int \mathrm{d}\boldsymbol{w}\, f_\alpha}}$$

$$= \cancelto{0}{\int \mathrm{d}\boldsymbol{w}\, C[f_\alpha]}$$

$$\Longrightarrow \quad \boxed{\frac{\mathrm{D}n_\alpha}{\mathrm{D}t_\alpha} + n_\alpha\boldsymbol{\nabla}\cdot\boldsymbol{u}_\alpha = 0} \quad \text{(continuity equation for species } \alpha) \tag{VIII.42}$$

$$\int \mathrm{d}\boldsymbol{w}\, m_\alpha\boldsymbol{w}(\text{VIII.40}):\quad \frac{\mathrm{D}}{\mathrm{D}t_\alpha}\cancelto{0}{\int \mathrm{d}\boldsymbol{w}\, m_\alpha\boldsymbol{w} f_\alpha} + \int \mathrm{d}\boldsymbol{w}\, m_\alpha\boldsymbol{w}\boldsymbol{w}\cdot\boldsymbol{\nabla} f_\alpha + m_\alpha\boldsymbol{a}_\alpha\cdot\underbrace{\int \mathrm{d}\boldsymbol{w}\,\boldsymbol{w}\frac{\partial f_\alpha}{\partial \boldsymbol{w}}}_{\overset{\text{bp}}{=}\, -n_\alpha\mathsf{I}}$$

$$+ q_\alpha\cancelto{0}{\int \mathrm{d}\boldsymbol{w}\,\boldsymbol{w}\left(\frac{\boldsymbol{w}}{c}\times\boldsymbol{B}\right)\cdot\frac{\partial f_\alpha}{\partial\boldsymbol{w}}} - \cancelto{0}{\int \mathrm{d}\boldsymbol{w}\, m_\alpha\boldsymbol{w}(\boldsymbol{w}\cdot\boldsymbol{\nabla}\boldsymbol{u}_\alpha)\cdot\frac{\partial f_\alpha}{\partial\boldsymbol{w}}}$$

$$= \int \mathrm{d}\boldsymbol{w}\, m_\alpha\boldsymbol{w}C[f_\alpha]$$

$$\Longrightarrow \quad \boxed{\boldsymbol{\nabla}\cdot\mathsf{P}_\alpha - m_\alpha n_\alpha\boldsymbol{a}_\alpha = \int \mathrm{d}\boldsymbol{w}\, m_\alpha\boldsymbol{w}C[f_\alpha] \doteq \boldsymbol{R}_\alpha} \quad \text{(force equation for species } \alpha) \tag{VIII.43}$$

where

$$\boxed{\mathsf{P}_\alpha \doteq \int \mathrm{d}\boldsymbol{w}\, m_\alpha\boldsymbol{w}\boldsymbol{w} f_\alpha} \tag{VIII.44}$$

is the thermal pressure tensor of species α and $\boldsymbol{R}_\alpha$ is the friction force on species α (recall Newton's third law, $\sum_\alpha \boldsymbol{R}_\alpha = 0$). Equation (VIII.43) may of course be rewritten in the following, perhaps more familiar, form:

$$m_\alpha n_\alpha \frac{\mathrm{D}\boldsymbol{u}_\alpha}{\mathrm{D}t_\alpha} = q_\alpha n_\alpha \Big(\boldsymbol{E} + \frac{\boldsymbol{u}_\alpha}{c} \times \boldsymbol{B}\Big) + m_\alpha n_\alpha \boldsymbol{g} - \boldsymbol{\nabla \cdot} \boldsymbol{\mathsf{P}}_\alpha + \boldsymbol{R}_\alpha. \tag{VIII.45}$$

If we sum (VIII.45) over species, the electric-field term vanishes by quasineutrality, $\sum_\alpha q_\alpha n_\alpha = 0$. Then, defining the total mass density $\varrho \doteq \sum_\alpha m_\alpha n_\alpha$ and the mean center-of-mass velocity $\boldsymbol{u} \doteq \varrho^{-1} \sum_\alpha m_\alpha n_\alpha \boldsymbol{u}_\alpha$, equation (VIII.45) implies

$$\varrho\Big(\frac{\partial}{\partial t} + \boldsymbol{u \cdot \nabla}\Big)\boldsymbol{u} = \frac{\boldsymbol{j}}{c} \times \boldsymbol{B} + \varrho \boldsymbol{g} - \boldsymbol{\nabla \cdot} (\boldsymbol{\mathsf{P}} + \boldsymbol{\mathsf{D}}), \tag{VIII.46}$$

where

$$\boldsymbol{j} = \sum_\alpha q_\alpha n_\alpha \boldsymbol{u}_\alpha \tag{VIII.47}$$

is the current density, $\boldsymbol{\mathsf{P}} \doteq \sum_\alpha \boldsymbol{\mathsf{P}}_\alpha$ is the total pressure tensor, and

$$\boldsymbol{\mathsf{D}} \doteq \sum_\alpha m_\alpha n_\alpha \Delta\boldsymbol{u}_\alpha \Delta\boldsymbol{u}_\alpha \tag{VIII.48}$$

is a tensor composed of species drifts relative to the center-of-mass velocity,

$$\Delta\boldsymbol{u}_\alpha \doteq \boldsymbol{u}_\alpha - \boldsymbol{u}. \tag{VIII.49}$$

(Note that $\sum_\alpha m_\alpha n_\alpha \Delta\boldsymbol{u}_\alpha = 0$, by definition.) Returning to those moments...

$$\begin{aligned}
\int \mathrm{d}\boldsymbol{w}\, m_\alpha w_i w_j (\text{VIII.40}) : \quad & \frac{\mathrm{D}}{\mathrm{D}t_\alpha} \underbrace{\int \mathrm{d}\boldsymbol{w}\, m_\alpha w_i w_j}_{= \mathsf{P}_{\alpha,ij}} + \int \mathrm{d}\boldsymbol{w}\, m_\alpha w_i w_j \boldsymbol{w \cdot \nabla} f_\alpha \\
& + m_\alpha a_{\alpha,k} \cancelto{0}{\int \mathrm{d}\boldsymbol{w}\, w_i w_j \frac{\partial f_\alpha}{\partial w_k}} \\
& + \underbrace{\int \mathrm{d}\boldsymbol{w}\, q_\alpha w_i w_j \Big(\frac{\boldsymbol{w}}{c} \times \boldsymbol{B}\Big) \boldsymbol{\cdot} \frac{\partial f_\alpha}{\partial \boldsymbol{w}}}_{\substack{\overset{\mathrm{bp}}{=} - \int \mathrm{d}\boldsymbol{w}\, q_\alpha \left[w_i \left(\frac{\boldsymbol{w}}{c} \times \boldsymbol{B}\right)_j + \left(\frac{\boldsymbol{w}}{c} \times \boldsymbol{B}\right)_i w_j\right] f_\alpha \\ = -\frac{q_\alpha}{m_\alpha}\left(\frac{\boldsymbol{\mathsf{P}}_\alpha}{c} \times \boldsymbol{B}\right)_{ij} - \frac{q_\alpha}{m_\alpha}\left(\frac{\boldsymbol{\mathsf{P}}_\alpha}{c} \times \boldsymbol{B}\right)_{ji}}} \\
& - \underbrace{\int \mathrm{d}\boldsymbol{w}\, m_\alpha w_i w_j (\boldsymbol{w \cdot \nabla} u_{\alpha,\ell}) \frac{\partial f_\alpha}{\partial w_\ell}}_{\substack{\overset{\mathrm{bp}}{=} - \int \mathrm{d}\boldsymbol{w}\, m_\alpha \left[w_i (\boldsymbol{w \cdot \nabla} u_{\alpha,j}) + (\boldsymbol{w \cdot \nabla} u_{\alpha,i}) w_j + w_i w_j (\boldsymbol{\nabla \cdot u}_\alpha)\right] f_\alpha \\ = -(\boldsymbol{\mathsf{P}}_\alpha \boldsymbol{\cdot \nabla u}_\alpha)_{ij} - (\boldsymbol{\mathsf{P}}_\alpha \boldsymbol{\cdot \nabla u}_\alpha)_{ji} - \boldsymbol{\mathsf{P}}_\alpha (\boldsymbol{\nabla \cdot u}_\alpha)}} \\
& = \int \mathrm{d}\boldsymbol{w}\, m_\alpha w_i w_j C[f_\alpha].
\end{aligned} \tag{VIII.50}$$

Define the heat flux tensor for species α:

$$\boldsymbol{\mathsf{Q}}_\alpha \doteq \int \mathrm{d}\boldsymbol{w}\, m_\alpha \boldsymbol{w}\boldsymbol{w}\boldsymbol{w} f_\alpha. \tag{VIII.51}$$

Then, equation (VIII.50) becomes

$$\frac{\mathrm{D}\boldsymbol{P}_\alpha}{\mathrm{D}t_\alpha} + \boldsymbol{\nabla}\cdot\boldsymbol{Q}_\alpha + \frac{q_\alpha}{m_\alpha}\left[\left(\frac{\boldsymbol{P}_\alpha}{c}\times\boldsymbol{B}\right) + \left(\frac{\boldsymbol{P}_\alpha}{c}\times\boldsymbol{B}\right)^T\right] + \left[(\boldsymbol{P}_\alpha\cdot\boldsymbol{\nabla}\boldsymbol{u}_\alpha) + (\boldsymbol{P}_\alpha\cdot\boldsymbol{\nabla}\boldsymbol{u}_\alpha)^T\right] + \boldsymbol{P}_\alpha(\boldsymbol{\nabla}\cdot\boldsymbol{u}_\alpha) = \int \mathrm{d}\boldsymbol{w}\, m_\alpha \boldsymbol{w}\boldsymbol{w} C[f_\alpha], \tag{VIII.52}$$

where the superscript T denotes the transpose. In component form, (VIII.52) reads

$$\frac{\mathrm{D}P_{\alpha,ij}}{\mathrm{D}t_\alpha} + (\boldsymbol{\nabla}\cdot\boldsymbol{Q}_\alpha)_{ij} + \frac{q_\alpha}{m_\alpha}\big(\epsilon_{jk\ell}P_{\alpha,ik}B_\ell + \epsilon_{ik\ell}P_{\alpha,jk}B_\ell\big) + \big(\delta_{i\ell}P_{\alpha,jk} + \delta_{j\ell}P_{\alpha,ik} + \delta_{k\ell}P_{\alpha,ij}\big)\frac{\partial u_{\alpha,\ell}}{\partial r_k} = \int \mathrm{d}\boldsymbol{w}\, m_\alpha w_i w_j C[f_\alpha]. \tag{VIII.53}$$

Usually the trace of this equation is taken, with

$$p_\alpha \doteq \frac{1}{3}\mathrm{tr}\,\boldsymbol{P}_\alpha. \tag{VIII.54}$$

Then (VIII.52) provides an evolutionary equation for the internal energy:

$$\frac{3}{2}\frac{\mathrm{D}p_\alpha}{\mathrm{D}t_\alpha} + \boldsymbol{\nabla}\cdot\boldsymbol{q}_\alpha + \boldsymbol{P}_\alpha\,\boldsymbol{:}\,\boldsymbol{\nabla}\boldsymbol{u}_\alpha + \frac{3}{2}p_\alpha\boldsymbol{\nabla}\cdot\boldsymbol{u}_\alpha = Q_\alpha, \tag{VIII.55}$$

where

$$\boldsymbol{q}_\alpha \doteq \int \mathrm{d}\boldsymbol{w}\,\frac{1}{2}m_\alpha w^2\boldsymbol{w} f_\alpha \tag{VIII.56}$$

is the conductive heat flux of species α and

$$Q_\alpha \doteq \int \mathrm{d}\boldsymbol{w}\,\frac{1}{2}m_\alpha w^2 C[f_\alpha] \tag{VIII.57}$$

is the collisional energy exchange. Further writing

$$\boxed{\boldsymbol{P}_\alpha \doteq p_\alpha\boldsymbol{I} + \boldsymbol{\Pi}_\alpha,} \tag{VIII.58}$$

where $\boldsymbol{\Pi}_\alpha$ is the viscous stress tensor of species α and using (VIII.42) to replace $\boldsymbol{\nabla}\cdot\boldsymbol{u}_\alpha$ in (VIII.55) by $\mathrm{d}\ln n_\alpha/\mathrm{d}t$, the internal energy equation (VIII.55) provides an equation for the hydrodynamic entropy:

$$\boxed{\frac{3}{2}p_\alpha\frac{\mathrm{D}}{\mathrm{D}t_\alpha}\ln\frac{p_\alpha}{n_\alpha^{5/3}} = -\boldsymbol{\nabla}\cdot\boldsymbol{q}_\alpha - \boldsymbol{\Pi}_\alpha\,\boldsymbol{:}\,\boldsymbol{\nabla}\boldsymbol{u}_\alpha + Q_\alpha} \tag{VIII.59}$$

Finally, using (VIII.58), the force equation (VIII.45) becomes

$$\boxed{m_\alpha n_\alpha\frac{\mathrm{D}\boldsymbol{u}_\alpha}{\mathrm{D}t_\alpha} = q_\alpha n_\alpha\left(\boldsymbol{E} + \frac{\boldsymbol{u}_\alpha}{c}\times\boldsymbol{B}\right) + m_\alpha n_\alpha\boldsymbol{g} - \boldsymbol{\nabla}p_\alpha - \boldsymbol{\nabla}\cdot\boldsymbol{\Pi}_\alpha + \boldsymbol{R}_\alpha} \tag{VIII.60}$$

Clearly, to close the system of hydrodynamic equations (viz., (VIII.42), (VIII.59), and (VIII.60)), we require $(\boldsymbol{\Pi}_\alpha, \boldsymbol{q}_\alpha, \boldsymbol{R}_\alpha, Q_\alpha)$ expressed in terms of the lower "fluid" moments $(n_\alpha, \boldsymbol{u}_\alpha, p_\alpha)$. This is the purpose of what is called the *Chapman–Enskog expansion*, which is only possible when the collisional mean free path is much smaller than the lengthscales of interest (e.g., gradient scales) so that the distribution function f_α is nearly Maxwellian. This will give a tractable kinetic equation, without time variation, which will close the moment equations and allow evolution on a slow timescale. The result is *magnetohydrodynamics*.

Note that, if we were simply to drop $(\boldsymbol{\Pi}_\alpha, \boldsymbol{q}_\alpha, \boldsymbol{R}_\alpha, Q_\alpha)$ from the fluid equations above, we'd have the set

$$\frac{\mathrm{D}n_\alpha}{\mathrm{D}t_\alpha} + n_\alpha \boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{u}_\alpha = 0, \tag{VIII.61}$$

$$m_\alpha n_\alpha \frac{\mathrm{D}\boldsymbol{u}_\alpha}{\mathrm{D}t_\alpha} = q_\alpha n_\alpha\left(\boldsymbol{E} + \frac{\boldsymbol{u}_\alpha}{c}\times\boldsymbol{B}\right) + m_\alpha n_\alpha \boldsymbol{g} - \boldsymbol{\nabla}p_\alpha, \tag{VIII.62}$$

$$\frac{3}{2}p_\alpha \frac{\mathrm{D}}{\mathrm{D}t_\alpha}\ln\frac{p_\alpha}{n_\alpha^{5/3}} = 0, \tag{VIII.63}$$

which are just the equations of ideal MHD written out for each species α.

VIII.8. Landau damping via Newton's 2nd law

Imagine an electron moving along the z axis with speed v_0. Slowly turn on a wave-like electric field: $\boldsymbol{E}(t,z) = E_0\cos(\omega t - kz)\,\mathrm{e}^{\epsilon t}\hat{\boldsymbol{z}}$, where ω is the frequency as k is the wavenumber of the wave. The adverb "slowly" is captured by the $\mathrm{e}^{\epsilon t}$ factor with $\epsilon \ll 1$. We'll take $\epsilon \to +0$ at the end of the calculation; its only purpose is to establish an arrow of time. The goal is to solve perturbatively for the motion of the electron by assuming that E_0 is so small that it changes the electron's trajectory only a little bit over several wave periods. The solution illustrates the physical mechanism of Landau damping (Lifshitz & Pitaevskii 1981).

The equations of motion are

$$\frac{\mathrm{d}z}{\mathrm{d}t} = v_z, \tag{VIII.64}$$

$$\frac{\mathrm{d}v_z}{\mathrm{d}t} = -\frac{e}{m_e}E_0\cos(\omega t - kz)\mathrm{e}^{\epsilon t}. \tag{VIII.65}$$

The solution to lowest order in E_0 is trivial: $z(t) = v_0 t$ and $v_z(t) = v_0 = \text{const}$. Write $z(t) = v_0 t + \delta z(t)$ and $v_z(t) = v_0 + \delta v_z(t)$ and calculate the first-order changes δz and δv_z. Equation (VIII.64) becomes

$$\frac{\mathrm{d}\delta v_z}{\mathrm{d}t} = -\frac{e}{m_e}E(t,z(t)) \approx -\frac{e}{m_e}E(t,v_0 t) = -\frac{eE_0}{m_e}\,\mathrm{Re}\,\mathrm{e}^{[\mathrm{i}(\omega - kv_0)+\epsilon]t}. \tag{VIII.66}$$

Integrating this gives

$$\begin{aligned}\delta v_z(t) &= -\frac{eE_0}{m_e}\int_0^t \mathrm{d}t'\,\mathrm{Re}\,\mathrm{e}^{[\mathrm{i}(\omega-kv_0)+\epsilon]t'} \\ &= -\frac{eE_0}{m_e}\,\mathrm{Re}\,\frac{\mathrm{e}^{[\mathrm{i}(\omega-kv_0)+\epsilon]t} - 1}{\mathrm{i}(\omega - kv_0)+\epsilon} \\ &= -\frac{eE_0}{m_e}\frac{\epsilon\,\mathrm{e}^{\epsilon t}\cos[(\omega-kv_0)t] - \epsilon + (\omega-kv_0)\mathrm{e}^{\epsilon t}\sin[(\omega-kv_0)t]}{(\omega-kv_0)^2+\epsilon^2}.\end{aligned} \tag{VIII.67}$$

Integrating again,

$$\begin{aligned}\delta z(t) &= \int_0^t \mathrm{d}t'\,\delta v_z(t') = -\frac{eE_0}{m_e}\int_0^t \mathrm{d}t'\,\mathrm{Re}\,\frac{\mathrm{e}^{[\mathrm{i}(\omega-kv_0)+\epsilon]t}-1}{\mathrm{i}(\omega-kv_0)+\epsilon} \\ &= -\frac{eE_0}{m_e}\left\{\mathrm{Re}\,\frac{\mathrm{e}^{[\mathrm{i}(\omega-kv_0)+\epsilon]t}-1}{[\mathrm{i}(\omega-kv_0)+\epsilon]^2} - \frac{\epsilon t}{(\omega-kv_0)^2+\epsilon^2}\right\} \\ &= -\frac{eE_0}{m_e}\left\{\frac{[\epsilon^2-(\omega-kv_0)^2][\mathrm{e}^{\epsilon t}\cos[(\omega-kv_0)t]-1]+2\epsilon\,(\omega-kv_0)\,\mathrm{e}^{\epsilon t}\sin[(\omega-kv_0)t]}{[(\omega-kv_0)^2+\epsilon^2]^2}\right. \\ &\qquad \left. -\ \frac{\epsilon t}{(\omega-kv_0)^2+\epsilon^2}\right\}. \end{aligned} \tag{VIII.68}$$

The first-order correction to the electric field evaluated at the particle position is

$$\delta E(t,z(t)) = E(t,z(t)) - E(t,v_0t) = \delta z(t)\frac{\partial E(t,v_0t)}{\partial z} = \delta z(t)\,k\sin[(\omega-kv_0)t]E_0\mathrm{e}^{\epsilon t}, \tag{VIII.69}$$

with $\delta z(t)$ given by (VIII.68). The work done by the field on the electron per unit time is the power gained by the electron (and thus lost by the wave). Denoting an average over timescales satisfying $\omega^{-1} \ll t \ll \epsilon^{-1}$ by $\langle\,\cdot\,\rangle$, this power is

$$\begin{aligned} P(v_0) &= -e\langle E(t,z(t))v_z(t)\rangle = -e\langle[E(t,v_0t)+\delta E(t,z)][v_0+\delta v_z(t)]\rangle \\ &= -e\Big\langle \underbrace{E(t,v_0t)v_0}_{\substack{\text{vanishes}\\\text{under}\\\text{averaging}}} + \underbrace{E(t,v_0t)\delta v_z(t)}_{\substack{\text{only }\cos^2\\\text{term survives}\\\text{averaging}}} + \underbrace{\delta E(t,z(t))v_0}_{\substack{\text{only }\sin^2\\\text{term survives}\\\text{averaging}}}\Big\rangle + \mathcal{O}(\delta^2) \\ &\approx \frac{e^2E_0^2}{m_e}\,\mathrm{e}^{2\epsilon t}\left\langle\frac{\epsilon}{(\omega-kv_0)^2+\epsilon^2}\cos^2[(\omega-kv_0)t] + \frac{2kv_0\epsilon(\omega-kv_0)}{[(\omega-kv_0)^2+\epsilon^2]^2}\sin^2[(\omega-kv_0)t]\right\rangle \\ &= \frac{e^2E_0^2}{2m_e}\mathrm{e}^{2\epsilon t}\left[\frac{\epsilon}{(\omega-kv_0)^2+\epsilon^2} + \frac{2kv_0\epsilon(\omega-kv_0)}{[(\omega-kv_0)^2+\epsilon^2]^2}\right] \end{aligned}$$

$$\Longrightarrow \quad \boxed{P(v_0) = \frac{e^2E_0^2}{2m_e}\,\mathrm{e}^{2\epsilon t}\frac{\mathrm{d}\chi}{\mathrm{d}v_0} \quad \text{with} \quad \chi \doteq \frac{\epsilon v_0}{(\omega-kv_0)^2+\epsilon^2}} \tag{VIII.70}$$

If $v_0 \lesssim \omega/k$ (particle lagging the wave), then $\mathrm{d}\chi/\mathrm{d}v_0 > 0$ and so $P(v_0) > 0$, indicating that energy is being transferred from the field to the electron. The wave damps. If $v_0 \gtrsim \omega/k$ (particle leading the wave), then $\mathrm{d}\chi/\mathrm{d}v_0 < 0$ and so $P(v_0) < 0$, indicating that energy is being transferred from the electron to the field. The wave grows.

Suppose there is now a distribution of these electrons, $F(v_0)$. The total power per unit volume going into (or out of) this distribution is

$$P = \int \mathrm{d}v_z\,F(v_z)P(v_z) = \frac{e^2E_0^2}{2m_e}\mathrm{e}^{2\epsilon t}\int \mathrm{d}v_z\,F(v_z)\frac{\mathrm{d}\chi}{\mathrm{d}v_z} \stackrel{\text{bp}}{=} -\frac{e^2E_0^2}{2m_e}\mathrm{e}^{2\epsilon t}\int \mathrm{d}v_z\,F'(v_z)\chi(v_z). \tag{VIII.71}$$

Take $\epsilon \to +0$ and use Plemelj's formula,

$$\lim_{\epsilon\to+0}\frac{1}{x-\zeta\mp\mathrm{i}\epsilon} = \mathrm{PV}\frac{1}{x-\zeta} \pm \mathrm{i}\pi\delta(x-\zeta),$$

where PV denotes the principal value and $\delta(x)$ is the Dirac delta function, to show that

$$\chi(v_z) = \frac{\epsilon v_z}{(\omega - kv_z)^2 + \epsilon^2} = -\frac{\mathrm{i}}{2}\left(\frac{v_z}{kv_z - \omega - \mathrm{i}\epsilon} - \frac{v_z}{kv_z - \omega + \mathrm{i}\epsilon}\right) \to \pi \frac{\omega}{k^2}\delta(v_z - \omega/k). \tag{VIII.72}$$

Using this limit in (VIII.71) leads to

$$\boxed{P = -\frac{e^2 E_0^2}{2m_e k^2}\pi\omega F'(\omega/k)} \tag{VIII.73}$$

If $\omega F'(\omega/k) < 0$ (> 0), there are more resonant particles lagging (leading) the wave than there are leading (lagging) the wave, resulting in a net transfer of energy to the electrons (wave). It's left as an exercise to the reader to show that, for $F'(\omega/k) < 0$, this power comes at the expense of the electric energy (i.e., damping), and that, for $F(\omega/k) > 0$, the energy loss from the electrons goes into growing the electric energy (i.e., instability).

REFERENCES

Balbus, Steven A. 1988 On Thermal Instability and Hydrostatic Equilibrium in Cooling Flows. *Astrophys. J.* **328**, 395.

Balbus, Steven A. & Soker, Noam 1989 Theory of Local Thermal Instability in Spherical Systems. *Astrophys. J.* **341**, 611.

Balescu, R. 1960 Irreversible Processes in Ionized Gases. *Physics of Fluids* **3**, 52–63.

Kruskal, M. D. 1958 The Gyration of a Charged Particle. *Project Matterhorn Publications and Reports* .

Kruskal, M. D. 1962 Asymptotic Theory of Hamiltonian and other Systems with all Solutions Nearly Periodic. *Journal of Mathematical Physics* **3**, 806–828.

Landau, L. 1937 Transport equation in the case of Coulomb interaction. *Zh. Eksper. i Teoret. Fiz.* **7**, 203.

Lenard, A. 1960 On Bogoliubov's kinetic equation for a spatially homogeneous plasma. *Annals of Physics* **10**, 390–400.

Lifshitz, E. M. & Pitaevskii, L. P. 1981 *Physical kinetics*.

Lundquist, S. 1951 On the stability of magneto-hydrostatic fields. *Phys. Rev.* **83**, 307.

Nicholson, D. 1983 *Introduction to plasma theory*. New York: John Wiley & Sons, Inc.

Northrop, T. G. 1963*a* Adiabatic Charged-Particle Motion. *Reviews of Geophysics and Space Physics* **1**, 283–304.

Northrop, T. G. 1963*b* *The adiabatic motion of charged particles*. New York: Interscience Publishers.

Rayleigh, J. W. S. 1880 On the stability, or instability, of certain fluid motions. *Proc. London Math. Soc.* **11**, 57.