Waves in Space Plasmas

An introduction for the course Space Physics II at Uppsala University

Anders I. Eriksson

Swedish Institute of Space Physics
Uppsala Division
and
Department of Astronomy and Space Physics
Uppsala University

Version 0.7.3
April 21, 2004

Contents

Release notes for 0.7.2, 19 August 2003

This compendium have been used in different forms for different courses in Space Physics at Uppsala University since the early 1990s. In this first new version of this compendium for five years, a lot of minor errors have been corrected. Despite being more than a decade old, this text still is in a pre-release stage. It may be time to really release it as 1.0 soon, but there still are a few things I would like to add, and I also need YOUR feedback: please mail any suggestions and corrections to me at Anders.Eriksson@irfu.se.

Release notes for 0.7.3, 21 April 2004

The fact that the only difference to 0.7.2 is the correction of the text to problems 4.6 and 5.4 should not lead you to believe that I think the rest is perfect. Anyway, thanks to Alessandro Retino for spotting these errors. Please send more comments and corrections to Anders.Eriksson@irfu.se.

Chapter 1

Waves and Fourier analysis

1.1 Why study waves?

The first and perhaps most obvious answer is that waves really *exist* in the world around us, and deserve to be studied just because of this. We are surrounded by waves wherever we go and whatever we do: light waves, sound waves, water waves. Who has not been fascinated by the rings from a stone thrown into the water, or by the ocean waves breaking on the shore? Waves are just as common out in space as they are in our everyday surroundings. The so-called "empty space" is filled with a gas of charged particles, a plasma. Because of the plasma, our everyday experience of acoustic and electromagnetic waves is not perfectly applicable in space. The charged particles in the plasma move when they are influenced by the electric and magnetic fields in an electromagnetic wave. Particle motions in a gas is associated with acoustic waves, so in a sense an electromagnetic wave in a plasma is partly acoustic. On the other hand, the motion of the charged particles in an acoustic wave in a plasma causes charge imbalances and current flows, and so create electromagnetic fields. Therefore, a sound wave in a plasma gets a partly electromagnetic character. Because of this connection between electric and mechanical properties, waves in space plasma have a very rich and complicated structure, which is in itself a reason for their study.

Another reason to study waves is that they may be *important.* The practical use we have for acoustic (sound) and electromagnetic (light) waves in our everyday lifes is obvious and can hardly be overestimated. All human communication, with the possible exception of direct physical touch, make use of these waves at some stadium. The plasma waves we study in this course have a direct application to human information exchange by means of radio waves. Waves are also important for large-scale processes in nature. The light waves in the solar radiation heat the earth, but this heating is balanced by cooling due to emission of long-wavelength thermal wave radiation from the earth. Ocean waves erode the coastline. Similarly, plasma waves in space near a planet may "erode" the planetary atmosphere, accelerating ionized particles to speeds above the escape velocity. Recent satellite and radar measurements indicate that the Earth looses oxygen at a rate of some kilograms per second by such processes.

A third reason for studying waves is that for *linear* systems, waves are really the *only* phenomena we need to study for a complete description of the dynamics of the system. A system of linear field equations can always be written as

$$\mathcal{L} f(t, \mathbf{r}) = 0 \tag{1.1}$$

where f denotes the fields (magnetic field, density, temperature, or whatever) and $\mathcal{L}$ is a linear operator, i.e. an expression independent of f[1]. If f_1 and f_2 are two solutions of the linear system (1.1), then the sum $f_1 + f_2$ is another solution to the same system. This well known *principle of superposition* is extremely powerful. Combined with Fourier analysis, where a function is written as a sum (integral) of wavelike quantities (sinusoidal functions), this principle means that if we know the properties of these wavelike quantities in the medium, then all of the system dynamics can be described by just summing over a set of waves. In reality, it turns out that many (perhaps most) systems of interest are non-linear, but we will see that a linear approximation often is very useful, which means that waves are fundamental to our understanding.

[1] We also assume that $\mathcal{L}$ is translationally invariant in time and space. Physically, this implies that we confine our studies to a homogeneous and stationary medium.

1.2 Fourier analysis

From the mathematics courses, we are acquaintained to *Fourier's theorem,* stating that we can write any function $f(t)$ as a superposition of complex sine functions[2]

$$f(t) = \int_{-\infty}^{\infty} f(\omega)\ \exp(-\mathrm{i}\omega t)\,\mathrm{d}\omega, \tag{1.2}$$

where[3]

$$f(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)\ \exp(\mathrm{i}\omega t)\,\mathrm{d}t \tag{1.3}$$

is the *Fourier transform* of $f(t)$. You may have seen other definitions: the factor $1/2\pi$ may be placed in front of any of the integrals, and the signs in the exponentials may also change places. For a physicist, it is suitable to use the definition above for Fourier transforms in time, but to use the opposite sign when transforming a function of a spatial coordinate – we will soon se why. A function of the spatial coordinate x is therefore written

$$f(x) = \int_{-\infty}^{\infty} f(k_x)\ \exp(\mathrm{i}k_x x)\,\mathrm{d}k_x \tag{1.4}$$

where the Fourier transform in space is

$$f(k_x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)\ \exp(-\mathrm{i}k_x x)\,\mathrm{d}x. \tag{1.5}$$

For a function of time and space, we transform one variable at a time, getting

$$\begin{aligned} f(t,x) &= \int_{-\infty}^{\infty} f(\omega,x)\ \exp(-\mathrm{i}\omega t)\,\mathrm{d}\omega \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\omega,k_x)\exp(\mathrm{i}[k_x x-\omega t])\,\mathrm{d}k_x\,\mathrm{d}\omega, \end{aligned} \tag{1.6}$$

where

$$\begin{aligned} f(\omega,k_x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t,k_x)\ \exp(\mathrm{i}\omega t)\,\mathrm{d}t \\ &= \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t,x)\exp(\mathrm{i}[\omega t-k_x x])\,\mathrm{d}x\,\mathrm{d}t. \end{aligned} \tag{1.7}$$

The function $f(t,x)$ is expanded in a sum (integral) of sinusoidal functions of $k_x x - \omega t$, describing plane waves with frequency $\omega/(2\pi)$ and wavelength $2\pi/\omega$ propagating along the x-axis. This is the reason for our choice of different signs in the exponentials of the Fourier integrals in time and space. Fouriers theorem can now be interpreted as stating that all functions of t and x can be written as sums of plane waves, so plane waves are the only things we need bother about. This is a result of fundamental importance.

The extension to three spatial dimensions is straightforward. One easily finds (by transforming one variable at a time) that a function f of four variables (three position coordinates $\mathbf{r}$ and time t) may be written in terms of new variables $\mathbf{k}$ and ω as

$$f(t,\mathbf{r}) = \int\int f(\omega,\mathbf{k})\exp(\mathrm{i}[\mathbf{k}\cdot\mathbf{r}-\omega t])\,\mathrm{d}\omega\,\mathrm{d}^3k \tag{1.8}$$

where

$$f(\omega,\mathbf{k}) = \frac{1}{(2\pi)^4}\int\int f(t,\mathbf{r})\exp(\mathrm{i}[\omega t-\mathbf{k}\cdot\mathbf{r}])\,\mathrm{d}t\,\mathrm{d}^3r. \tag{1.9}$$

[2]A mathematician would here argue that the function must dissapear quickly enough at infinity, satisfy certain continuity conditions etc. We will assume that all functions of physical interest fulfill the relevant requirements.

[3]One may wonder if the definition (1.3) is physically reasonable, as it involves integrating over all time, i.e. not only over the past we at least formally can know something about, but also over all future. In fact, interesting and verifiable physical results turn up when studying waves by integrating not to $+\infty$ but only to present time in (1.3). A nice treatment of these things is found in the book by Brillouin and Sommerfeld. One example is Landau damping, here only briefly and phenomenologically introduced in section 4.3 on page 34; for a better treatment see Swanson page 141.

The integrand of (1.8),

$$f(\omega, \mathbf{k}) \exp(\mathrm{i}[\mathbf{k} \cdot \mathbf{r} - \omega t]) = f(\omega, k_x, k_y, k_z) \exp(\mathrm{i}[k_x x + k_y y + k_z z - \omega t]), \tag{1.10}$$

describes a *planar sinusoidal wave* in three dimensions[4]. The wave has amplitude $f(\omega, k)$ and propagates in the direction of the wave vector $\mathbf{k}$. The modulus $k = |\mathbf{k}|$ is the *wave number*, which is related to the wavelength λ by $k = 2\pi/\lambda$.

The principle of superposition now tells us that if (1.10) is a solution to the field equation (1.1) for all $\mathbf{k}$ and ω, then $f(t, \mathbf{r})$ is a solution as well, as the intergral in (1.8) essentially is a summation over $\mathbf{k}$ and ω. Hence, as all functions can be written in the form (1.8), all solutions to the linear field equation (1.1) can be written as a superposition of plane waves. Thus, for linear systems, *we only have to study plane sinusoidal waves* – everything else can be written as a superposition of such waves.

In general, the operator $\mathcal{L}$ will contain a lot of ∇ and $\partial/\partial t$, so (1.1) will be a partial differential equation. Such equations are hard to solve, but if we only have to look at plane waves, i.e. solutions of the form

$$u(t, \mathbf{r}) = u_0 \exp(\mathrm{i}[\mathbf{k} \cdot \mathbf{r} - \omega t]) \tag{1.11}$$

for scalar quantities and

$$\mathbf{w}(t, \mathbf{r}) = \mathbf{w}_0 \exp(\mathrm{i}[\mathbf{k} \cdot \mathbf{r} - \omega t]) \tag{1.12}$$

for vector fields, we find that

$$\begin{aligned}
\frac{\partial u}{\partial t} &= u_0 \frac{\partial}{\partial t} \exp(\mathrm{i}[\mathbf{k} \cdot \mathbf{r} - \omega t]) = -\mathrm{i}\omega u_0 \exp(\mathrm{i}[\mathbf{k} \cdot \mathbf{r} - \omega t]) = -\mathrm{i}\omega u && (1.13)\\
\frac{\partial \mathbf{w}}{\partial t} &= ... = -\mathrm{i}\omega \mathbf{w} && (1.14)\\
\nabla u &= u_0 (\hat{\mathbf{x}} \frac{\partial}{\partial x} + \hat{\mathbf{y}} \frac{\partial}{\partial y} + \hat{\mathbf{z}} \frac{\partial}{\partial z}) \exp(\mathrm{i}[k_x x + k_y y + k_z z - \omega t]) \\
&= \mathrm{i}(k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}} + k_z \hat{\mathbf{z}}) u_0 \exp(\mathrm{i}[\mathbf{k} \cdot \mathbf{r} - \omega t]) = \mathrm{i}\mathbf{k} u && (1.15)\\
\nabla \cdot \mathbf{w} &= ... = \mathrm{i}\mathbf{k} \cdot \mathbf{w} && (1.16)\\
\nabla \times \mathbf{w} &= ... = \mathrm{i}\mathbf{k} \times \mathbf{w}. && (1.17)
\end{aligned}$$

Thus, for plane sinusoidal waves we can substitute

$$\boxed{\frac{\partial}{\partial t} \longrightarrow -\mathrm{i}\omega} \tag{1.18}$$

and

$$\boxed{\nabla \longrightarrow \mathrm{i}\mathbf{k}} \tag{1.19}$$

Therefore, if we have some partial differential equation, for example

$$\nabla \times \mathbf{E}(t, \mathbf{r}) = -\frac{\partial \mathbf{B}(t, \mathbf{r})}{\partial t} \tag{1.20}$$

we know that the Fourier transforms of the fields will satisfy

$$\mathrm{i}\mathbf{k} \times \mathbf{E}(\omega, \mathbf{k}) = \mathrm{i}\omega \mathbf{B}(\omega, \mathbf{k}). \tag{1.21}$$

When Fourier transforming, we just have to use the substitutions (1.18) and (1.19) to get the transformed equations.

[4]A sinusoidal quantity is here understood to be something which goes like $\exp(i[\mathbf{k} \cdot \mathbf{r} - \omega t])$.

A general procedure for solving linear systems of partial differential equations then is: (1) solve the equations for sinusoidal waves, and (2) build up the solution that satisfies the given initial and boundary conditions by adding plane wave solutions by the Fourier integral (1.8).[5]

The field equations we encounter in physics are partial differential equations. From other courses, we may for example think of the Newtonian gravitational field equation (mechanics), the Schrödinger equation (quantum mechanics), the Navier-Stokes equations (continuum mechanics), Einstein's equations for the gravitational field (general relativity), and, above all, Maxwell's equations for the electromagnetic fields (electromagnetics). Not all of those equations are linear, but all of them can be treated as linear at least for small perturbations from equilibrium[6], and may therefore be studied by the Fourier method. This is the basic reason why plane waves are studied in all these branches of physics: water waves, probability waves, pressure waves, gravitational waves, electromagnetic waves. In a similar manner, the equations that govern the behaviour of the space plasma have wave solutions, so the study of waves is fundamental for the understanding of the space plasma.

1.3 Maxwell's equations

Classical electrodynamics is completely contained in the four *Maxwell equations*[7]. These are Gauss' law for the electric field,

$$\boxed{\nabla \cdot \mathbf{E}(t, \mathbf{r}) = \rho(t, \mathbf{r})/\epsilon_0,} \tag{1.22}$$

Gauss' law for the magnetic field (also known as the condition of no magnetic monopoles),

$$\boxed{\nabla \cdot \mathbf{B}(t, \mathbf{r}) = 0,} \tag{1.23}$$

Faraday-Henry's law,

$$\boxed{\nabla \times \mathbf{E}(t, \mathbf{r}) = -\frac{\partial \mathbf{B}(t, \mathbf{r})}{\partial t},} \tag{1.24}$$

and Ampère-Maxwell's law,

$$\boxed{\nabla \times \mathbf{B}(t, \mathbf{r}) = \mu_0 \mathbf{j}(t, \mathbf{r}) + \frac{1}{c^2}\frac{\partial \mathbf{E}(t, \mathbf{r})}{\partial t}.} \tag{1.25}$$

Here ρ is the charge density (SI unit: C/m^3), $\mathbf{j}$ is the current density (A/m^2), and μ_0, ϵ_0 and c are the usual constants, related by $\mu_0 \epsilon_0 c^2 = 1$.

How do we know that there is any well defined solution to these equations? This is ensured by Helmholtz's theorem, which states that a vector field can be divided into a curl-free part, completely determined by its divergence sources, and a divergence-free part, determined by its curl sources, as long as we have reasonable boundary conditions[8]. The fields are therefore determined by their divergence and curl,

[5] This method is used for solving boundary value problems in the course "Mathematical methods of physics". In the present course, we will only study the plane wave solutions, and not do step (2).

[6] We will have more to say about this process of *linearization* later on (sections 2.3 and 2.4).

[7] The equations are here written in a form in which all information on any material which may be present has to be included in the charge and current densities ρ and $\mathbf{j}$. Another possibility is to use the fields $\mathbf{D}$ and $\mathbf{H}$, and place the description of the medium in the relations $\mathbf{D} = \mathbf{D}(\mathbf{E})$ and $\mathbf{H} = \mathbf{H}(\mathbf{B})$.

[8] See, for example, Panofsky and Phillips, p. 2 – 6.

i.e. by the charge and current densities, and by the boundary conditions[9]. We may thus divide the electric field $\mathbf{E}$ into a curl-free part $\mathbf{E}_\mathrm{S}$ and a divergence-free part $\mathbf{E}_\mathrm{I}$. The curl-free part is known as the *electrostatic field*, satisfying

$$\nabla \cdot \mathbf{E}_\mathrm{S}(t, \mathbf{r}) = \rho(t, \mathbf{r})/\epsilon_0 \tag{1.26}$$

$$\nabla \times \mathbf{E}_\mathrm{S}(t, \mathbf{r}) = 0, \tag{1.27}$$

and may thus be written in terms of a scalar potential as

$$\mathbf{E}_\mathrm{S}(t, \mathbf{r}) = -\nabla\Phi(t, \mathbf{r}). \tag{1.28}$$

Note that even though the field is called "electrostatic" it does not have to be static at all: if the charge density ρ is varying in time, so will $\mathbf{E}_\mathrm{S}$ do. In fact, we will find that a plasma supports electrostatic waves – a phenomenon unknown in vacuum and neutral gases.

The divergence-free part is known as the *induced electric field*, and obeys

$$\nabla \cdot \mathbf{E}_\mathrm{I}(t, \mathbf{r}) = 0 \tag{1.29}$$

$$\nabla \times \mathbf{E}_\mathrm{I}(t, \mathbf{r}) = -\frac{\partial \mathbf{B}(t, \mathbf{r})}{\partial t}. \tag{1.30}$$

The total electric field is

$$\mathbf{E} = \mathbf{E}_\mathrm{S} + \mathbf{E}_\mathrm{I}. \tag{1.31}$$

According to Gauss' law (1.23), there are no divergence sources for the magnetic field, and therefore no magnetic analogy of the electrostatic field.

Problems for Chapter 1

1. *Fourier transforms.* Write down the Fourier transformed Maxwell equations.
2. *Delta function.* Derive the expression

   $$\delta(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega t}\,\mathrm{d}\omega$$

   for the Dirac delta function $\delta(t)$, defined by

   $$f(t) = \int_{-\infty}^{\infty} f(t')\,\delta(t - t')\,\mathrm{d}t',$$

   by use of equations (1.2) and (1.3).
3. *Continuity equation.* Derive the equation of continuity

   $$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0$$

   from Maxwell's equations. What is its physical meaning?
4. *Electrostatic waves.* Show that if a wave field has $\mathbf{E} \parallel \mathbf{k}$, the wave electric field is completely electrostatic. Also show that such a wave cannot exist in vacuum.

[9] In the Maxwell equations, the $\mathbf{B}$-field is a curl source for $\mathbf{E}$ and vice versa, so it is perhaps not obvious that Helmholtz's theorem can be used directly. However, by rewriting the equations in terms the potentials $\Phi(t, \mathbf{r})$ and $\mathbf{A}(t, \mathbf{r})$ the situation becomes clearer. See Wangsness, p. 37, or Jackson, p. 219.

Chapter 2

Linear and linearised wave equations

2.1 Linear or non-linear equations: Why bother?

Modern developments of classical physics have revealed a startling complexity and richness, which almost solely is due to improved insights in non-linear phenomena. Concepts like deterministic chaos originates in non-linear effects. What is it that makes non-linear equations so different from linear equations?

Mathematically, the answer lies in the *principle of superposition*. This tells us that if f_1 and f_2 are solutions to a linear equation, then so is $a\, f_1 + b\, f_2$. For non-linear equations, no such general way of finding a new solution from other, already known, solutions exist. Among other things, this makes it impossible to analyze a nonlinear situation with the Fourier methods outlined in Section 1.2. This may still not sound very exciting. But consider the following two situations, and the difference between linear and nonlinear physics becomes obvious:

First consider the light rays from two torches (flashlights, see Figure 2.1a). Light propagation in air is well described by the Maxwell equations (1.22) – (1.25) with current density $\mathbf{j}$ and charge density ρ both put to zero:

$$\nabla \cdot \mathbf{E}(t, \mathbf{r}) = 0 \tag{2.1}$$

$$\nabla \cdot \mathbf{B}(t, \mathbf{r}) = 0 \tag{2.2}$$

$$\nabla \times \mathbf{E}(t, \mathbf{r}) = -\frac{\partial \mathbf{B}(t, \mathbf{r})}{\partial t} \tag{2.3}$$

$$\nabla \times \mathbf{B}(t, \mathbf{r}) = \frac{1}{c^2}\frac{\partial \mathbf{E}(t, \mathbf{r})}{\partial t}. \tag{2.4}$$

These equations are *linear* in the field variables $\mathbf{E}$ and $\mathbf{B}$. Thus, the rays from the torches are described by linear equations. The rays from one torch is one solution and the rays from the other torch is another:

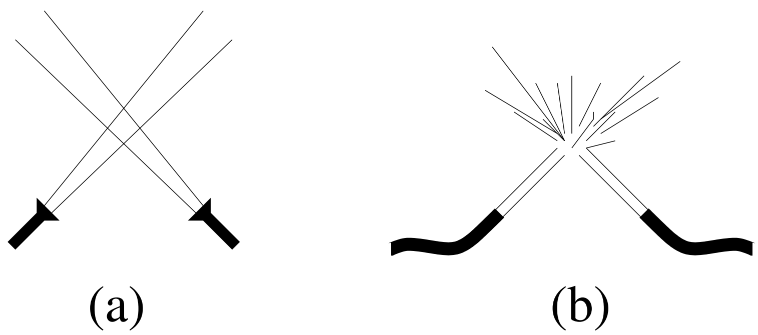


Figure 2.1: The results of crossing (a) two light ray bundles and (b) two water jets are radically different. In (a), the equations are linear, so that the principle of superposition applies, while the equations describing (b) are nonlinear, resulting in turbulent scattering rather than tranquil superposition.

hence the superposition of the two rays should be another solution. A simple experiment shows this to be true: the two ray bundles cross without appreciable scatter, and the overall ray paths are well described as a superposition of the rays from the single torches.

Then consider the very different result you get when crossing the jets from two garden hoses! If having just one hose, we get one well-defined jet, but the resulting pattern from two jets crossing each other certainly does not look like the superposition of the two individual jets (Figure 2.1b). Instead, a lot of spray-producing scattering, with a turbulent and unpredictable fine structure, will occur where the jets cross. This could be expected from a mathematical model of the situation. The main equations governing the motion of a water jet are the Navier-Stokes equations

$$\rho_{\mathrm{m}} \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\nabla p + \rho_{\mathrm{m}} \mathbf{g} + \mu \nabla^2 \mathbf{v} \tag{2.5}$$

and the equation of continuity

$$\frac{\partial \rho_{\mathrm{m}}}{\partial t} + \nabla \cdot (\rho_{\mathrm{m}} \mathbf{v}) = 0. \tag{2.6}$$

In these equations, $\mathbf{v}$ is the velocity field of the fluid, ρ_{m} is its mass density, p the pressure field, $\mathbf{g}$ the acceleration vector of gravity and μ the viscosity. These equations may at first glance not look very nonlinear, but in fact they are. Most important in this respect is often the first term of (2.5), which is nonlinear because

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\, \mathbf{v} \tag{2.7}$$

so that it does contain a term where the field variable $\mathbf{v}$ is multiplied by (a derivative of) itself. Hence, we should not expect the principle of superposition to work in this case. The nonlineraity is the reason for the scattering of the jets where they cross.

2.2 Linear wave equations: Electromagnetic waves in vacuum

Our chief interest is plasma waves, but for a start, we will repeat the theory of electomagnetic (EM) waves in a vacuum. The derivation of their properties is analogous to the plasma wave analyses we will do later, but is simpler, and thus constitutes a good illustration of the method. The equations governing electromagnetic waves in vacuum are (2.1) – (2.4), which as stated in Section 2.1 are *linear* in the field variables $\mathbf{E}$ and $\mathbf{B}$. According to what has been discussed above (page 5), we may concentrate on solutions in terms of plane waves, as any other solution may be written as a superposition (Fourier integral) of plane waves. For plane sinusoidal waves, we can use the substitutions (1.18) and (1.19), so the vacuum Maxwell equations (2.1) – (2.4) become

$$\mathrm{i}\mathbf{k} \cdot \mathbf{E} = 0 \tag{2.8}$$

$$\mathrm{i}\mathbf{k} \cdot \mathbf{B} = 0 \tag{2.9}$$

$$\mathrm{i}\mathbf{k} \times \mathbf{E} = \mathrm{i}\omega \mathbf{B} \tag{2.10}$$

$$\mathrm{i}\mathbf{k} \times \mathbf{B} = -\mathrm{i}\frac{\omega}{c^2}\mathbf{E}. \tag{2.11}$$

Strictly, we should have written $\mathbf{E}(\omega, \mathbf{k})$ and correspondingly for the $\mathbf{B}$-field to emphasize that we are dealing with Fourier amplitudes.

The last two equations above combine to give

$$-\mathrm{i}\frac{\omega}{c^2}\mathbf{E} = \mathrm{i}\mathbf{k} \times \mathbf{B} = \mathrm{i}\mathbf{k} \times (\frac{\mathbf{k}}{\omega} \times \mathbf{E}) \tag{2.12}$$

which with the vector relation[1]

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}) \tag{2.13}$$

[1] See e.g. Physics Handbook chapter M-9

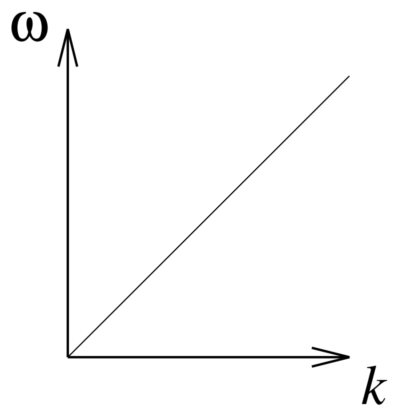


Figure 2.2: Dispersion relation for EM waves in vacuum.

and equation (2.8) may be rewritten as

$$(\omega^2 - k^2c^2)\mathbf{E} = 0. \tag{2.14}$$

This shows that only the Fourier components $\mathbf{E}(\omega, \mathbf{k})$ for which

$$\boxed{\omega^2 = k^2c^2} \tag{2.15}$$

is satisfied can be non-zero. Equation (2.15) is our first example of a *dispersion relation* – a relation between $\mathbf{k}$ och ω that must be satisfied for a non-zero field to exist. The dispersion relation (2.15), plotted in Figure 2.2, does not look too interesting, as its only message is the well known fact that for an EM wave in vacuum, the product of wavelength and frequency is a constant (the speed of light). However, it is important to realize that this is not a general property of all waves, not even of all electromagnetic waves. Waves in general – electromagnetic waves in plasmas or in condensed matter, surface waves in the bathtub, or whatever – most often do not have $\omega \propto k$, and are then known as *dispersive* waves: the EM waves in vacuum thus are non-dispersive.

Dispersion relations may be written in many different ways. To write them on the form $\omega^2 = f(\mathbf{k})$ like (2.15) above is often natural. Another frequently used formulation is to use the *index of refraction* μ, defined by

$$\mu^2 = \frac{k^2c^2}{\omega^2}. \tag{2.16}$$

For the vacuum EM waves, the dispersion relation becomes $\mu = 1$. For a dispersive wave, we get $\mu = \mu(\mathbf{k})$. This situation is well known from optics, where light of different colour have different index of refraction and is refracted in different ways at interfaces between media. Similar phenomena will be encountered in the plasma.

2.3 Linearised wave equations: sound waves

The fundamental equations of motion for a fluid are the Navier-Stokes equations (2.5). For a gas (air, for instance), we can can often neglect the viscosity, and if we also neglect effects of the gravitational field, we get an equation of motion

$$mn(t, \mathbf{r})\frac{\mathrm{d}\mathbf{v}(t, \mathbf{r})}{\mathrm{d}t} = -\nabla p(t, \mathbf{r}) \tag{2.17}$$

where

$$\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla, \tag{2.18}$$

p is the pressure and we have written the mass density $\rho_{\rm m} = mn$, where m is the mass of a molecule and n the molecular density (number of molecules per unit volume). The motion must also satisfy the equation of continuity,

$$\frac{\partial n(t,\mathbf{r})}{\partial t} + \nabla \cdot [n(t,\mathbf{r})\mathbf{v}(t,\mathbf{r})]. \tag{2.19}$$

Pressure p and density n are related by some equation of state. We will here assume that this is the ideal gas law

$$p(t,\mathbf{r}) = n(t,\mathbf{r})KT, \tag{2.20}$$

where the temperature is considered to be constant[2]. We immediately use this to eliminate the pressure from (2.17). The equations we get constitute a complete system describing the evolution of the gas in time and space. If we are interested in waves in the neutral gas, we find that the equations are non-linear, as we have products of the field quantities in the equations. Thus we cannot find any simple wave solutions by the method of section 1.1. However, if we only study small perturbations from an equilibrium, we may *linearise* the equations. We do as follows:

1. **Ansatz**. Rewrite the fields as

$$\mathbf{v}(t,\mathbf{r}) = \mathbf{v}_0 + \mathbf{v}_1(t,\mathbf{r}) \tag{2.21}$$

$$n(t,\mathbf{r}) = n_0 + n_1(t,\mathbf{r}) \tag{2.22}$$

   where

   (a) Terms with index zero are the unperturbed background values of the equilibrium, which is assumed to be constant in time and space.

   (b) Terms with index 1 denotes a small perturbation. Thus,

$$\mathbf{v}_1 \ll \mathbf{v}_0 \tag{2.23}$$

$$n_1 \ll n_0. \tag{2.24}$$

2. **Apply to the field equations**. Put the ansatzes (2.21) and (2.22) into the field equations (2.17) and (2.19)

$$\begin{aligned} m[n_0+n_1]\left(\frac{\partial[\mathbf{v}_0+\mathbf{v}_1]}{\partial t} + ([\mathbf{v}_0+\mathbf{v}_1]\cdot\nabla)[\mathbf{v}_0+\mathbf{v}_1]\right) &= \\ &= -KT\nabla[n_0+n_1] \end{aligned} \tag{2.25}$$

$$\begin{aligned} 0 &= \frac{\partial[n_0+n_1]}{\partial t} + \nabla\cdot([n_0+n_1][\mathbf{v}_0+\mathbf{v}_1]) = \\ &= \frac{\partial[n_0+n_1]}{\partial t} + [n_0+n_1]\nabla\cdot[\mathbf{v}_0+\mathbf{v}_1] + \\ &\quad +[\mathbf{v}_0+\mathbf{v}_1]\cdot\nabla[n_0+n_1] \end{aligned} \tag{2.26}$$

3. **Derivatives of background values disappear**. Use that terms with index 0 are constants

$$m[n_0+n_1]\left(\frac{\partial \mathbf{v}_1}{\partial t} + ([\mathbf{v}_0+\mathbf{v}_1]\cdot\nabla)\mathbf{v}_1\right) = -KT\nabla n_1 \tag{2.27}$$

$$0 = \frac{\partial n_1}{\partial t} + [n_0+n_1]\nabla\cdot\mathbf{v}_1 + [\mathbf{v}_0+\mathbf{v}_1]\cdot\nabla n_1 \tag{2.28}$$

[2] In reality, a better description of sound waves is given by the adiabatic condition p/n^γ = constant, but for purposes of illustration, we use the simpler isothermal approximation.

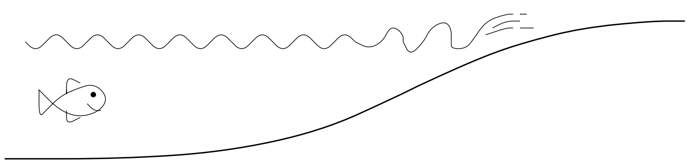


Figure 2.3: The waves we are studying should be small perturbations to a stationary background. One may compare to water surface waves, which can be described as sinusoidal if their amplitude is much less than the depth of the water. However, if the amplitude is comparable to or greater than the average depth, the waves are no longer sinusoidal. Far out at sea, the waves are fairly sinusoidal (at least if the wind is weak), but when they come closer to the shore they grow higher and steeper, loose their sinusoidal shape and break on the shore in a way that definitely not can be described by monochromatic sinusoidal waves. The linearisation is no longer valid, and the linear solutions (sine waves) do not describe the phenomena.

4. **Neglect higher terms**. Because of (2.23) and (2.24), we can neglect terms of form $x_1 y_1$ as compared to terms $x_0 y_1$:

$$mn_0\left(\frac{\partial \mathbf{v}_1(t,\mathbf{r})}{\partial t} + (\mathbf{v}_0\cdot\nabla)\mathbf{v}_1(t,\mathbf{r})\right) = -KT\,\nabla n_1(t,\mathbf{r}) \tag{2.29}$$

$$0 = \frac{\partial n_1(t,\mathbf{r})}{\partial t} + n_0\nabla\cdot\mathbf{v}_1(t,\mathbf{r}) + \mathbf{v}_0\cdot\nabla n_1(t,\mathbf{r}) \tag{2.30}$$

It is this last step which is the linearization. By the procedure above, the nonlinear field equations (2.17) and (2.19) are transformed into the linear equations (2.29) och (2.30) for the perturbation fields n_1 and $\mathbf{v}_1$. We refer to the equations (2.29) and (2.30) as the *linearised equations*. As these equations are linear, the principle of superposition is valid for their solutions, and we may use the Fourier methods from section 1.2.

We now do so, and look for sine wave solutions to the linearised field equations (2.29) and (2.30). By the substitutions $\partial/\partial t \to -\mathrm{i}\omega$ (1.18) and $\nabla \to \mathrm{i}\mathbf{k}$ we get

$$-\mathrm{i}\omega m n_0\mathbf{v}_1 + \mathrm{i}\mathbf{v}_0\cdot\mathbf{k}\mathbf{v}_1 = -\mathrm{i}KT\mathbf{k}n_1 \tag{2.31}$$

$$0 = -\mathrm{i}\omega n_1 + \mathrm{i}n_0\mathbf{k}\cdot\mathbf{v}_1 + \mathrm{i}\mathbf{v}_0\cdot\mathbf{k}n_1. \tag{2.32}$$

In a system moving with the gas, $\mathbf{v}_0 = 0$, and the equations above[3] boil down to

$$\omega\mathbf{v}_1 = \frac{KT}{mn_0}\mathbf{k}n_1 \tag{2.33}$$

$$0 = \omega n_1 - n_0\mathbf{k}\cdot\mathbf{v}_1. \tag{2.34}$$

By using (2.33) in (2.34), we get

$$\begin{aligned} 0 &= \omega^2 n_1 - n_0\mathbf{k}\cdot\frac{KT}{mn_0}\mathbf{k}n_1 = \\ &= (\omega^2 - \frac{KT}{m}k^2)n_1. \end{aligned} \tag{2.35}$$

For any non-zero perturbation, $n_1 \neq 0$ we therefore must have

$$\omega^2 = c_s^2 k^2 \tag{2.36}$$

[3]The alert reader will perhaps protest that $\mathbf{v}_0 = 0$ violates (2.23). The neglect of the $\mathbf{v}_1\cdot\nabla$-terms in comparison to the $\partial/\partial t$-terms in (2.27) and (2.28) now must be motivated by that the amplitude $|\mathbf{v}_1|$ is small compared to the ratio of characteristic dimensions in the problem: $|\mathbf{v}_1| << L/T$, where $1/L = |\nabla|$ and $1/T = |\partial/\partial t|$.

where

$$c_s = \sqrt{\frac{KT}{m}} \tag{2.37}$$

is recognized as the sound speed[4]. Equation (2.36) is our second example of a dispersion relation, a relation between frequency ($\nu = \omega/2\pi$) and wavelength ($\lambda = 2\pi/k$) for the waves we study. The method of deriving this dispersion relation is very general, and we will use it for different types of plasma waves below. However, one must keep in mind that the method works only for small-amplitude perturbations (compare to ocean waves, figure 2.3).

2.4 Linearisation: General method

As linearisation is a fundamental method for studying the response of an equilibrium situation to a small perturbation, and is used in all parts of physics, we summarize the method here.

Assume we have a non-linear system of equations $F(\xi) = 0$ for the field variables ξ. For two-fluid plasma theory we have 14 unknown field components ($\mathbf{E}$, $\mathbf{B}$, $\mathbf{v}_i$, $\mathbf{v}_e$, n_i och n_e), so ξ is a vector in 14 dimensions.

1. **Ansatz**: $\xi(t, \mathbf{r}) = \xi_0 + \xi_1(t, \mathbf{r})$, where $\xi_1 \ll \xi_0$.
2. **Use in field equations**: $F(\xi_0 + \xi_1) = 0$
3. **Derivatives of background values are zero**.
4. **Neglect higher terms**: This gives linearised equations
   $G(\xi_0)\xi_1 = 0$.
5. This homogeneous system of linear equations has non-trivial solutions only if det$(G(\xi_0)) = 0$. This equation is the **dispersion relation**.

A view of what we are doing is that we Taylor expand the non-linear system around ξ_0, where the background fields ξ_0 fulfill the field equations $F(\xi_0) = 0$. We then get

$$0 = F(\xi) = F(\xi_0 + \xi_1) = F(\xi_0) + F'(\xi_0)\xi_1 + ... = F'(\xi_0)\xi_1 + ... \tag{2.38}$$

Neglecting higher terms, we have

$$F'(\xi_0)\xi_1 = 0 \tag{2.39}$$

as our linearised system[5]. Thus, $G(\xi_0)$ above is nothing else than $F'(\xi_0)$. This is a theoretically elegant way of summarizing the linearisation process, but in practize, this Taylor expansion formalism is cumbersome to handle. For the 14-dimensional state vector of two-fluid plasma theory, for instance, we get $F'(\xi_0)$ as a 14×14 matrix, and the dispersion relation det($F'(\xi_0)$) = 0 will thus be rather complex. The more direct approach we used in section 2.3 is ususally the more practical.

Problems for Chapter 2

1. *Sound waves.* Derive the dispersion relation for sound waves assuming adiabatic (p/n^γ =constant) rather than isothermal conditions (compare footnotes on page 12).
2. *Water surface waves.* Derive the dispersion relation $\omega^2 = gk$ (equation 3.57) for long-wavelegth (so that $\nabla p \longrightarrow 0$), small-amplitude surface waves on deep water, neglecting viscosity and surface tension.

---

[4] In reality, there should be a correction due to real sound waves being adiabatic rather than isothermal, so that $c_s = \sqrt{\gamma KT/m}$.

[5] This general method of linearization may remind us of what we once learned about linearisation of non-linear systems around simple critical points in the course Ordinary differential equations (Simmons page 471). Our approach here is similar, just applied to partial rather than ordinary differential equations. We are also interested only in periodic wave solutions, which with the terminology of Simmons means that we only study critical points of vortex type.

Chapter 3

Waves in a cold unmagnetized plasma

3.1 Dispersion relations

A plasma consist of free charges, which we here assume to be of two species: electrons with a number density n_e (unit: m^{-3}), mass m_e and charge $-e$, and ions with number density n_i, mass m_i and charge $+e$. The ions and electrons can in principle move independently of each other, so we may very well have non-zero charge and current densities. In terms of particle motion, these are given by

$$\rho(t,\mathbf{r}) = e[n_\mathrm{i}(t,\mathbf{r}) - n_\mathrm{e}(t,\mathbf{r})] \tag{3.1}$$

and

$$\mathbf{j}(t,\mathbf{r}) = e[n_\mathrm{i}(t,\mathbf{r})\mathbf{v}_\mathrm{i}(t,\mathbf{r}) - n_\mathrm{e}(t,\mathbf{r})\mathbf{v}_\mathrm{e}(t,\mathbf{r})]. \tag{3.2}$$

Maxwell's equations then read

$$\nabla\cdot\mathbf{E}(t,\mathbf{r}) = \frac{e}{\epsilon_0}[n_\mathrm{i}(t,\mathbf{r}) - n_\mathrm{e}(t,\mathbf{r})] \tag{3.3}$$

$$\nabla\cdot\mathbf{B}(t,\mathbf{r}) = 0 \tag{3.4}$$

$$\nabla\times\mathbf{E}(t,\mathbf{r}) = -\frac{\partial\mathbf{B}(t,\mathbf{r})}{\partial t} \tag{3.5}$$

$$\nabla\times\mathbf{B}(t,\mathbf{r}) = \mu_0 e[n_\mathrm{i}(t,\mathbf{r})\mathbf{v}_\mathrm{i}(t,\mathbf{r}) - n_\mathrm{e}(t,\mathbf{r})\mathbf{v}_\mathrm{e}(t,\mathbf{r})] + \frac{1}{c^2}\frac{\partial\mathbf{E}(t,\mathbf{r})}{\partial t}. \tag{3.6}$$

We now introduce a *cold two-fluid model* of the plasma. We consider the electrons and ions to be two different fluids, both assumed to be at zero temperature and hence zero pressure. The equations of motion for ions and electrons are

$$m_\mathrm{i}\frac{\mathrm{d}\mathbf{v}_\mathrm{i}(t,\mathbf{r})}{\mathrm{d}t} = e[\mathbf{E}(t,\mathbf{r}) + \mathbf{v}_\mathrm{i}(t,\mathbf{r})\times\mathbf{B}(t,\mathbf{r})] \tag{3.7}$$

and

$$m_\mathrm{e}\frac{\mathrm{d}\mathbf{v}_\mathrm{e}(t,\mathbf{r})}{\mathrm{d}t} = -e[\mathbf{E}(t,\mathbf{r}) + \mathbf{v}_\mathrm{e}(t,\mathbf{r})\times\mathbf{B}(t,\mathbf{r})], \tag{3.8}$$

respectively. To describe the plasma, we also have the equations of continuity for the two species,

$$\frac{\partial n_\mathrm{i}(t,\mathbf{r})}{\partial t} + \nabla\cdot[n_\mathrm{i}(t,\mathbf{r})\,\mathbf{v}_\mathrm{i}(t,\mathbf{r})] = 0 \tag{3.9}$$

$$\frac{\partial n_\mathrm{e}(t,\mathbf{r})}{\partial t} + \nabla\cdot[n_\mathrm{e}(t,\mathbf{r})\,\mathbf{v}_\mathrm{e}(t,\mathbf{r})] = 0, \tag{3.10}$$

valid as long as there are no ionization or recombination processes. The equations (3.3)–(3.10) form a closed system[1] to which we now will try to find wave solutions.

[1] There are 14 unknowns ($\mathbf{E}, \mathbf{B}, \mathbf{v}_\mathrm{i}, \mathbf{v}_\mathrm{e}, n_\mathrm{i}$, and n_e), but 16 equations. It may seem the system is overdetermined, but that is not the case. The problem is due to the form of Maxwell's equations that we have used. If they are reformulated in terms of the scalar potential Φ and the vector potential $\mathbf{A}$ rather than in terms of $\mathbf{E}$ and $\mathbf{B}$, two equations are trivially fulfilled, and the number of equations and variables becomes identical.

As in the case of pressure waves in a neutral gas presented above in section 2.3, the field equations are non-linear. The equations of continuity are non-linear because they contain the product of two field quantities n and $\mathbf{v}$, while the equations of motion are nonlinear in the $\mathbf{v} \times \mathbf{B}$-terms as well as in the derivative itself: in the convective derivative

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v}, \tag{3.11}$$

the velocity field is multiplied by (a derivative of) itself.

To find wave solutions, we therefore linearise the equations in the same manner as we did with the pressure waves:

1. **Ansatz.**

$$\begin{array}{rcccl} n_{\mathrm{i}}(t,\mathbf{r}) & = & n_0 & + & n_{1\mathrm{i}}(t,\mathbf{r}) \\ n_{\mathrm{e}}(t,\mathbf{r}) & = & n_0 & + & n_{1\mathrm{e}}(t,\mathbf{r}) \\ \mathbf{E}(t,\mathbf{r}) & = & & & \mathbf{E}_1(t,\mathbf{r}) \\ \mathbf{B}(t,\mathbf{r}) & = & & & \mathbf{B}_1(t,\mathbf{r}) \\ \mathbf{v}_{\mathrm{i}}(t,\mathbf{r}) & = & & & \mathbf{v}_{1\mathrm{i}}(t,\mathbf{r}) \\ \mathbf{v}_{\mathrm{e}}(t,\mathbf{r}) & = & & & \mathbf{v}_{1\mathrm{e}}(t,\mathbf{r}) \end{array} \tag{3.12}$$

   where

$$n_{1\mathrm{i}}(t,\mathbf{r}) \ll n_0 \tag{3.13}$$

$$n_{1\mathrm{e}}(t,\mathbf{r}) \ll n_0, \tag{3.14}$$

   while the other fields ($\mathbf{E}, \mathbf{B}, \mathbf{v}_{\mathrm{i}}$ and $\mathbf{v}_{\mathrm{e}}$) are supposed to be zero in the unperturbed equilibrium plasma. The most severe restriction imposed by this assumption is that no effects of a static background magnetic field, like the geomagnetic field, are included.

2. **Put into the field equations** (3.3) – (3.10).

3. **Derivatives of background values disappear.**

4. **Neglect higher terms.** This yields our linearised system:

$$\nabla \cdot \mathbf{E}_1(t,\mathbf{r}) = \frac{e}{\epsilon_0}[n_{1\mathrm{i}}(t,\mathbf{r}) - n_{1\mathrm{e}}(t,\mathbf{r})] \tag{3.15}$$

$$\nabla \cdot \mathbf{B}_1(t,\mathbf{r}) = 0 \tag{3.16}$$

$$\nabla \times \mathbf{E}_1(t,\mathbf{r}) = -\frac{\partial \mathbf{B}_1(t,\mathbf{r})}{\partial t} \tag{3.17}$$

$$\nabla \times \mathbf{B}_1(t,\mathbf{r}) = \mu_0 e[n_0(t,\mathbf{r})\mathbf{v}_{1\mathrm{i}}(t,\mathbf{r}) - n_0(t,\mathbf{r})\mathbf{v}_{1\mathrm{e}}(t,\mathbf{r})] + + \frac{1}{c^2}\frac{\partial \mathbf{E}_1(t,\mathbf{r})}{\partial t} \tag{3.18}$$

$$m_{\mathrm{i}}\frac{\partial \mathbf{v}_{1\mathrm{i}}(t,\mathbf{r})}{\partial t} = e\mathbf{E}_1(t,\mathbf{r}) \tag{3.19}$$

$$m_{\mathrm{e}}\frac{\partial \mathbf{v}_{1\mathrm{e}}(t,\mathbf{r})}{\partial t} = -e\mathbf{E}_1(t,\mathbf{r}) \tag{3.20}$$

$$\frac{\partial n_{1\mathrm{i}}(t,\mathbf{r})}{\partial t} + n_0 \nabla \cdot \mathbf{v}_{1\mathrm{i}}(t,\mathbf{r}) = 0 \tag{3.21}$$

$$\frac{\partial n_{1\mathrm{e}}(t,\mathbf{r})}{\partial t} + n_0 \nabla \cdot \mathbf{v}_{1\mathrm{e}}(t,\mathbf{r}) = 0 \tag{3.22}$$

This procedure is the same as in the case of pressure waves studied in section 2.3.

The equations (3.15) – (3.22) form a system of *linear* equations. We know that the only solutions to this system we have to look for is plane sinusoidal waves, as all other solutions can be built up from sine waves by linear superposition (Fourier integration, see section 1.2). By the substitutions (1.18) and (1.19), the equations can be written in the following form:

$$\begin{aligned}
\mathrm{i}\mathbf{k}\cdot\mathbf{E}_1 &= \frac{e}{\epsilon_0}[n_{1\mathrm{i}}-n_{1\mathrm{e}}] && (3.23)\\
\mathrm{i}\mathbf{k}\cdot\mathbf{B}_1 &= 0 && (3.24)\\
\mathrm{i}\mathbf{k}\times\mathbf{E}_1 &= \mathrm{i}\omega\mathbf{B}_1 && (3.25)\\
\mathrm{i}\mathbf{k}\times\mathbf{B}_1 &= \mu_0 e n_0[\mathbf{v}_{1\mathrm{i}}-\mathbf{v}_{1\mathrm{e}}]-\mathrm{i}\frac{\omega}{c^2}\mathbf{E}_1 && (3.26)\\
-\mathrm{i}\omega m_\mathrm{i}\mathbf{v}_{1\mathrm{i}} &= e\mathbf{E}_1 && (3.27)\\
-\mathrm{i}\omega m_\mathrm{e}\mathbf{v}_{1\mathrm{e}} &= -e\mathbf{E}_1 && (3.28)\\
-\mathrm{i}\omega n_{1\mathrm{i}}+\mathrm{i}n_0\mathbf{k}\cdot\mathbf{v}_{1\mathrm{i}} &= 0 && (3.29)\\
-\mathrm{i}\omega n_{1\mathrm{e}}+\mathrm{i}n_0\mathbf{k}\cdot\mathbf{v}_{1\mathrm{e}} &= 0 && (3.30)
\end{aligned}$$

This system of linear algebraic equations might give a formidable impression because of its size, but is really quite simple to handle. One way is to start by solving the equations of motion (3.27) and (3.28) for the velocities,

$$\mathbf{v}_{1\mathrm{i}} = \mathrm{i}\frac{e}{m_\mathrm{i}\omega}\mathbf{E}_1 \tag{3.31}$$

$$\mathbf{v}_{1\mathrm{e}} = -\mathrm{i}\frac{e}{m_\mathrm{e}\omega}\mathbf{E}_1 \tag{3.32}$$

and solving the equations of continuity (3.29) and (3.30) for the density fluctuations,

$$\begin{aligned}
n_{1\mathrm{i}} &= \frac{n_0}{\omega}\mathbf{k}\cdot\mathbf{v}_{1\mathrm{i}} = \mathrm{i}\frac{n_0 e}{m_\mathrm{i}\omega^2}\mathbf{k}\cdot\mathbf{E}_1 && (3.33)\\
n_{1\mathrm{e}} &= \frac{n_0}{\omega}\mathbf{k}\cdot\mathbf{v}_{1\mathrm{e}} = -\mathrm{i}\frac{n_0 e}{m_\mathrm{e}\omega^2}\mathbf{k}\cdot\mathbf{E}_1. && (3.34)
\end{aligned}$$

Using these expressions, we can eliminate densities and velocities from the Maxwell equations. In particular, the Ampère-Maxwell law (3.26) and the Faraday-Henry law (3.25) yields

$$\begin{aligned}
-\mathrm{i}\frac{\omega}{c^2}\mathbf{E}_1 &= \mathrm{i}\mathbf{k}\times\mathbf{B}_1-\mu_0 e n_0[\mathbf{v}_{1\mathrm{i}}-\mathbf{v}_{1\mathrm{e}}] = \\
&= \mathrm{i}\mathbf{k}\times(\frac{\mathbf{k}}{\omega}\times\mathbf{E}_1)-\mu_0 e n_0[\mathrm{i}\frac{e}{m_\mathrm{i}\omega}\mathbf{E}_1+\mathrm{i}\frac{e}{m_\mathrm{e}\omega}\mathbf{E}_1]. && (3.35)
\end{aligned}$$

As $m_\mathrm{i} \geq 1836\, m_\mathrm{e}$, we may neglect the $1/m_\mathrm{i}$-term. Using the vector relation (2.13), we get

$$-\mathrm{i}\frac{\omega}{c^2}\mathbf{E}_1 = \frac{\mathrm{i}}{\omega}(\mathbf{k}\cdot\mathbf{E}_1)\mathbf{k}-\frac{\mathrm{i}}{\omega}k^2\mathbf{E}_1-\frac{\mathrm{i}\mu_0 n_0 e^2}{m_\mathrm{e}\omega}\mathbf{E}_1. \tag{3.36}$$

Choosing coordinates as

$$\begin{aligned}
\mathbf{k} &= k\hat{\mathbf{x}} && (3.37)\\
\mathbf{E}_1 &= E_x\hat{\mathbf{x}}+E_y\hat{\mathbf{y}} && (3.38)
\end{aligned}$$

makes it possible to write this vector relation as

$$\omega^2(E_x\hat{\mathbf{x}}+E_y\hat{\mathbf{y}}) = -c^2k^2E_x\hat{\mathbf{x}}+c^2k^2(E_x\hat{\mathbf{x}}+E_y\hat{\mathbf{y}})+\frac{n_0e^2}{\epsilon_0 m_\mathrm{e}}(E_x\hat{\mathbf{x}}+E_y\hat{\mathbf{y}}). \tag{3.39}$$

The constant in front of the last term on the right hand side clearly has the dimension of (angular) frequency squared. We therefore introduce a new quantity

$$\boxed{\omega_\mathrm{p} = \sqrt{\frac{n_0e^2}{\epsilon_0 m_\mathrm{e}}}} \tag{3.40}$$

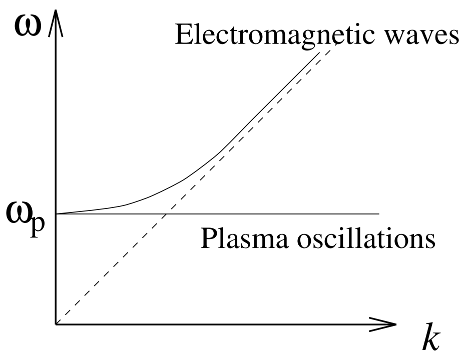


Figure 3.1: Dispersion relations in a cold unmagnetized plasma. The dashed line is the vacuum dispersion relation $\omega = kc$.

which we call *plasma frequency*[2] For the moment, this is just a formal definition: the physical meaning of the plasma frequency will turn up later on (section 3.6). The x-component of (3.39) may then be written

$$\omega^2 E_x = \omega_\mathrm{p}^2 E_x, \tag{3.41}$$

while the y-component becomes

$$\omega^2 E_y = k^2 c^2 E_y + \omega_\mathrm{p}^2 E_y. \tag{3.42}$$

We thus get two independent equations, implying that the x- and y-components of the E-field are independent of each other. When two independent waves exist in this manner, we call them different *wave modes*.

Considering how the coordinates were chosen, the x-component is parallel to $\mathbf{k}$ and is called a *longitudinal* wave mode, while the y-component is perpendicular to $\mathbf{k}$ is a *transversal* mode. From (3.41) we get the dispersion relation

$$\omega^2 = \omega_\mathrm{p}^2 \tag{3.43}$$

for longitudinal waves, and from (3.42) we get the dispersion relation

$$\omega^2 = \omega_\mathrm{p}^2 + c^2 k^2 \tag{3.44}$$

for transversal waves. Both dispersion relations are illustrated in Figure 3.1. In the limit $n_0 \longrightarrow 0$ we have $\omega_\mathrm{p} \longrightarrow 0$, and the dispersion relation for the transverse waves approaches the dispersion relation for EM waves in vacuum, $\omega^2 = k^2 c^2$, so this wave mode is the generalization of light and other EM waves to a plasma. In contrast, the longitudinal mode (3.43) have no counterpart in vacuum or a neutral gas. It is a completely new phenomenon, an electrostatic oscillation, to which we will return in section 3.6. We first concentrate on the transverse waves.

3.2 Electromagnetic waves

Equation (3.42) above told us that the transverse electric wave field, which propagates with $\mathbf{E} \perp \mathbf{k}$, must satisfy the dispersion relation

$$\boxed{\omega^2 = \omega_\mathrm{p}^2 + c^2 k^2} \tag{3.45}$$

As these waves are transverse, $\mathbf{k} \perp \mathbf{E}$ and $\mathbf{k} \cdot \mathbf{E} = 0$, so the electric field is completely induced with no electrostatic component (see equations (1.26) and (1.29) on page 7). A wave of this type is called an

[2]Or, rather, the *plasma angular frequency* (measured in rad/s). Strictly speaking, the plasma frequency, in units of hertz, is $f_\mathrm{p} = \omega_\mathrm{p}/(2\pi)$.

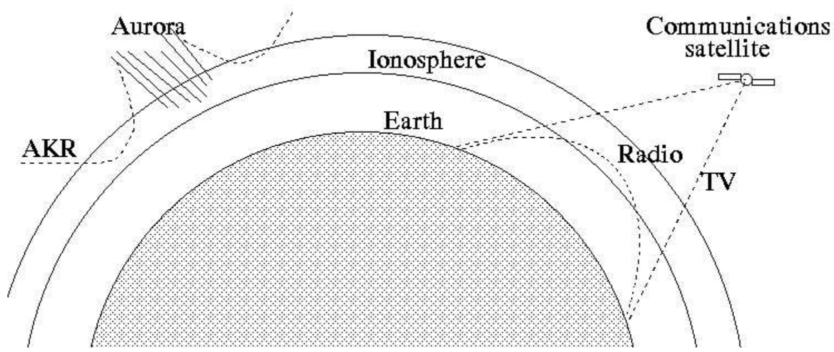


Figure 3.2: The ionosphere prohibits AKR emissions from reaching the ground, and causes radio waves from ground stations to be reflected. Television uses higher frequencies, normally above the maximum plasma frequency in the ionosphere, so TV cannot use the ionosphere for long range communications, but have to rely on cables, line-of-sight propagation, or communication satellites.

electromagnetic wave. As we said above, the dispersion relation (3.45) approaches the dispersion relation (2.15) for electromagnetic waves in a vacuum as n_0 and thus ω_{p} goes to zero. One may also note that for waves with $\omega \gg \omega_{\mathrm{p}}$, the vacuum dispersion relation (2.15) is valid to good accuracy. This implies that visible light is not strongly affected by the passage through the intergalactic or interstellar medium, the plasma in the solar wind, the magnetoshere, or the ionosphere[3].

The most striking consequence of (3.45) is perhaps that only waves with frequencies above the plasma frequency can propagate in the plasma. For $\omega < \omega_{\mathrm{p}}$, (3.45) gives solutions with imaginary k, implying that the wave decreases exponentially in space[4]. A wave of a certain fixed frequency ω thus cannot propagate in a region where the plasma density is so large that[5] $\omega_{\mathrm{p}} > \omega$. This is of great practical importance. Here on the ground, we have $n_0 = 0$, while up in the ionosphere, $n_0 \sim 10^{12}$ m^{-3} or something like that, implying plasma frequencies of up to tens of MHz. Thus waves with lower frequency cannot propagate from the Earth out in space. This is fundamental for radio communications on our planet. A radio wave emitted from the ground bounces in the ionosphere and may return to the surface of the earth far away from the source (section 3.4).

The ionosphere has positive impact on radio communications also in another sense. In the auroral regions, strong wave emissions with frequencies 50 – 500 kHz known as AKR (Auroral Kilometric Radiation) appear. These have a total effect of typically 10 MW, sometimes several GW. If these signals could penetrate down to the ground, they would severly disturb radio communications at least here in the north. But as they cannot penetrate through the ionosphere, they do not disturb us, and in fact they were not discovered until they were measured by satellites.

3.3 Phase and group velocity

If we want to study how fast a certain crest or valley in a sinusoidal wave is moving, we immediately see that it is tied to a particular value of the phase of the wave,

$$\mathbf{k} \cdot \mathbf{r} - \omega t = \text{constant}. \tag{3.46}$$

[3]However, light is affected by the magnetic fields associated with cosmic plasma, which causes the phenomenon of Faraday rotation: the plane of polarization is shifted by the presence of magnetized cosmic plasmas. This provides a means of estimating interstellar and intergalactic magnetic fields. Faraday rotation is treated by Chen, page 133, and its application for cosmic magnetic field estimation is discussed by Longair, page 209.

[4]The solution yielding exponential increase is of course unphysical.

[5]This result is strictly true only for a cold unmagnetized plasma: we will find waves at lower frequencies when we consider thermal effects (Chapter 4) and magnetization of the plasma (Chapter 5. However, even in these more complicated situations, there also exis waves behaving as the analysis in this section shows.

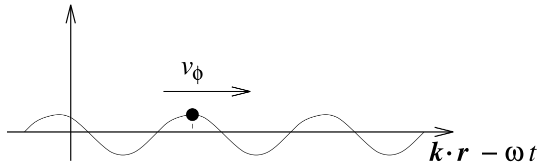


Figure 3.3: A point on a sine wave is identified by its phase $\mathbf{k}\cdot\mathbf{r}-\omega t$. The speed of this point is the phase speed.

This equation defines a relation between the position $\mathbf{r}$ of this particular crest or valley and the time t. We get the velocity by divding by t:

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}[\mathbf{k}\cdot\mathbf{r}-\omega t] = \mathbf{k}\cdot\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} - \omega = kv_\phi - \omega \tag{3.47}$$

$\Longrightarrow$

$$\boxed{v_\phi = \omega/k} \tag{3.48}$$

The speed $v_\phi = \hat{\mathbf{k}}\cdot\frac{d\mathbf{r}}{dt} = \omega/k$ is known as the *phase speed* of the wave, as it is the speed with which the phase is moving. For the dispersion relation (3.45) we get

$$v_\phi^2 = \omega^2/k^2 = \frac{c^2}{1-\omega_\mathrm{p}^2/\omega^2} > c^2. \tag{3.49}$$

The phase speed for an electromagnetic wave in an unmagnetized plasma thus is greater than the speed of light! How does this comply with the demands of the special theory of relativity?

In fact, everything is in order. What special relativity tells us is that information cannot be transported faster than light. But a single plane sine wave conveys no information. Let us assume that we wish to communicate information about when a certain event (dinner, for instance) occurs to some other person P far away. To tell this, we send a short wave packet of a certain frequency f to P. One could think that this means that we only transmit one single frequency; if there is a plasma between us and P, this information would then travel faster than light. But in reality, our signal does not look like $\sin 2\pi f t$ but rather something like $H(t)\sin 2\pi f t$, where

$$H(t) = \begin{cases} 0 & , \quad t<\tau \\ 1 & , \quad \tau<t<\tau+\Delta t \\ 0 & , \quad t>\tau+\Delta t, \end{cases} \tag{3.50}$$

$t=\tau$ is the start time of the pulse, and Δt its length. From Fourier analysis, we know that the Fourier transform of such a wave packet will include all frequencies, not just f. The shorter the length of the pulse, the broader the spectrum of the wave becomes[6]. The only way of transmitting a perfectly monochromatic wave is to keep the transmitter going from $t=-\infty$ till $t=\infty$, in which case it is completely impossible to use it for telling when a certain event happened – whenever P listens to his receiver, he will hear the same tone all the time.

A pulse carrying some information must thus contain all frequencies, but if it has sufficient extent in time, it may be formed so as to have high amplitude only in a small frequency interval. According to (3.49) the frequencies in the pulse will have different phase velocities. This implies that the interference pattern of the different frequencies may travel at a completely different speed. As information is carried by the interference pattern, this unknown speed is of physical interest. So what is it?

Consider a pulse fairly narrow in frequency space, i.e. rather long in time, comprising several wave periods. Only a small interval of frequencies then have Fourier components significantly different from zero. We now look at the simplified case of two frequencies in the pulse. This means that we have infinitely many pulses, as the superposition of two sine waves is a modulated sine wave (see Figure 3.5), but still we have a

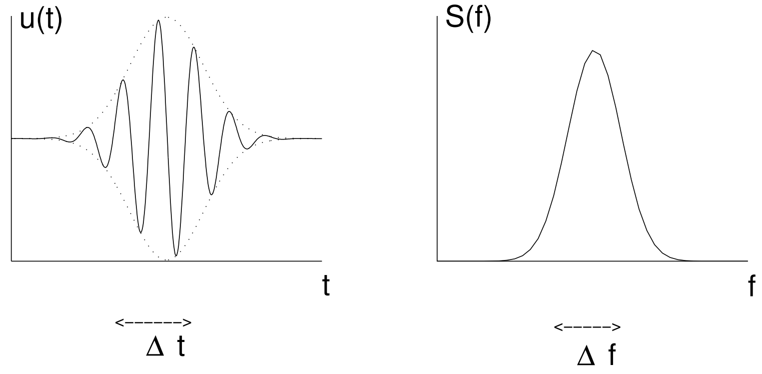


Figure 3.4: A wave pulse $u(t)$ and the magnitude of its complex Fourier spectrum, $S(f) = |u(f)|$. To convey information, we must use waves whoose appearance changes in time – there is no information in a pure sine wave. The duration Δt and the spread in frequency Δf of the wave packet are related by $\Delta f \; \Delta t \geq 1$.

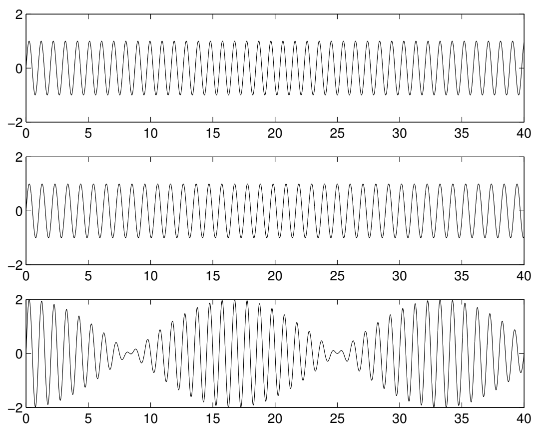


Figure 3.5: Two sine waves with slightly different frequencies (1.03 in the top panel and 0.97 in the center panel) yields an interference pattern known as beats when superposed (lower panel). The beat pattern moves with the group velocity.

model of how to create a localized wave packet. We take the frequencies to be $f - \mathrm{d}f = (\omega - \mathrm{d}\omega)/(2\pi)$ and $f + \mathrm{d}f = (\omega + \mathrm{d}\omega)/(2\pi)$ and to have unit Fourier amplitude. By the dispersion relation, the frequency $\omega - \mathrm{d}\omega$ corresponds to a certain wave number $k - \mathrm{d}k$, and $\omega + \mathrm{d}\omega$ corresponds to $k + \mathrm{d}k$. We may then write the wave as

$$\begin{aligned} u(x,t) = u_0\left(\sin([k-\mathrm{d}k]x - [\omega - \mathrm{d}\omega]t) + \sin([k+\mathrm{d}k]x - [\omega+\mathrm{d}\omega]t)\right) = \\ = 2u_0 \cos(\mathrm{d}k\, x - \mathrm{d}\omega\, t)\sin(kx - \omega t) \end{aligned} \tag{3.51}$$

using well-known trigonometric rules. This is a wave (k, ω) modulated by another wave $(\mathrm{d}k, \mathrm{d}\omega)$. The speed of the carrier wave is $\omega/k = v_\phi$, but what really matters is the speed of the modulation. We call this the *group velocity* v_g, which we can see is

$$v_\mathrm{g} = \frac{\mathrm{d}\omega}{\mathrm{d}k}. \tag{3.52}$$

As this is the speed of the interference pattern, it is the speed at which information is transmitted. The group velocity is therefore physically more important than the phase velocity, as it tells with what speed information and energy is carried with a wave. The phase speed only says something about single Fourier components, which do not exist in real life. Above, we treated a one-dimensional situation[7]. In general, the group velocity is the vector

$$\boxed{\mathbf{v}_\mathrm{g} = \frac{\partial \omega}{\partial \mathbf{k}}} \tag{3.53}$$

where the operator $\partial/\partial\mathbf{k}$ is the gradient in $\mathbf{k}$-space,

$$\frac{\partial}{\partial \mathbf{k}} = \nabla_\mathbf{k} = \hat{\mathbf{x}}\frac{\partial}{\partial k_x} + \hat{\mathbf{y}}\frac{\partial}{\partial k_y} + \hat{\mathbf{z}}\frac{\partial}{\partial k_z}. \tag{3.54}$$

By differentiating the dispersion relation (3.45), we get

$$2\omega\frac{\partial\omega}{\partial k} = 2c^2 k \tag{3.55}$$

$$\Longrightarrow \quad v_\mathrm{g} = \frac{\partial\omega}{\partial k} = c^2\frac{k}{\omega} = \frac{c^2}{v_\phi}. \tag{3.56}$$

As we had $v_\phi > c$ according to (3.49), we obviously get $v_\mathrm{g} < c$. Information is not transmitted faster than light, and the requirements of special relativity are satisfied. Figure 3.6 shows the dispersion relation (3.45), with the dashed lines lines having slopes equal to the phase and group velocities for a certain value of k.

To get some practical experience of a dispersion relation reminiscent of (3.45), one may experiment with water surface waves. The dispersion relation for water waves in deep water[8] is

$$\omega^2 = gk, \tag{3.57}$$

from which follows that

$$2\omega\frac{\partial\omega}{\partial k} = g \tag{3.58}$$

$\Longrightarrow$

$$v_\mathrm{g} = \frac{\partial\omega}{\partial k} = \frac{g}{2\omega} = \frac{gk}{2\omega^2}\frac{\omega}{k} = \frac{1}{2}v_\phi \tag{3.59}$$

as $v_\phi = \omega/k$ per definition, and $gk/\omega^2 = 1$ according to the dispersion relation (3.57). The dispersion relations (3.57) and (3.45) are certainly different (compare Figures 3.6 and 3.7), but they share the property $v_\mathrm{g} < v_\phi$. This can be seen when observing the ring shaped waves from a stone dropped in the water. Each wave crest moves faster than the pattern as a whole (compare Figure 3.8). This phenomenon is similar for radio waves in the ionosphere and for the water waves from the bread crumbs we may observe when feeding the ducks in the Swan Pond.

[6] In quantum mechanics, this is known as *Heisenberg's uncertainty relation*: the shorter the pulse length Δt, the larger the spread in frequency Δf according to $\Delta E \Delta t = h\Delta f \Delta t \geq h$, where h is Planck's constant.

[7] More stringent and complete treatments of the group velocity may be found in the books by Brillouin and by Jackson.

[8] Here "deep" means that the depth should be much larger than the wavelength. We have also neglected surface tension. A more general dispersion relation is $\omega^2 = (g + \frac{T}{\rho_m}k^2)k\tanh(kh)$, where the surface tension $T \approx 0.074$ kg/s^2 in the case of water. More on this is found in, for instance, the nice little book by Lighthill.

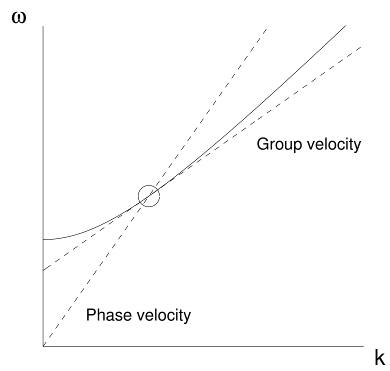


Figure 3.6: The dispersion relation for electromagnetic waves in a cold unmagnetized plasma (solid) and lines with slope equal to v_ϕ and v_g (dashed) at the point on the dispersion curve indicated by the circle.

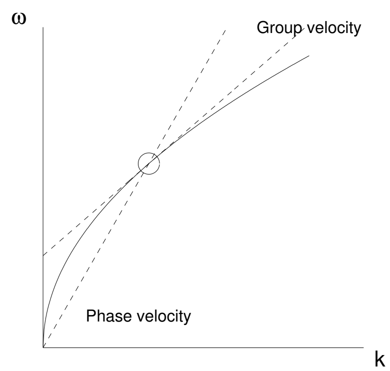


Figure 3.7: The dispersion relation for long wavelength surface waves in deep water (solid) and lines with slope equal to v_ϕ and v_g (dashed) at the point on the dispersion curve indicated by the circle.

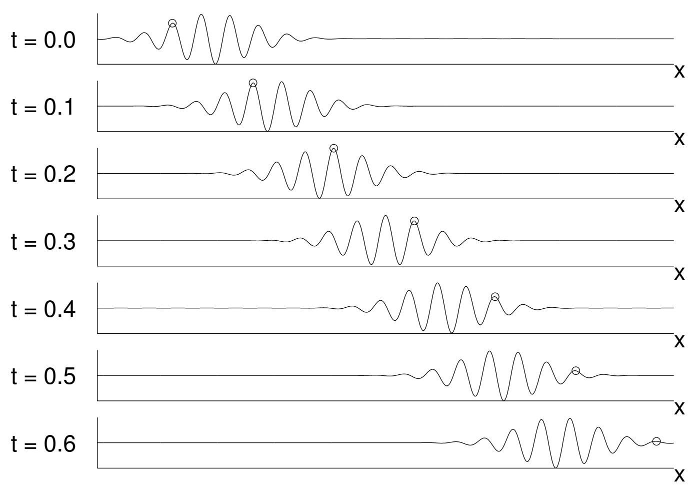


Figure 3.8: If the phase velocity exceeds the group velocity, the wave pattern in a wave packet will change with time. In the example above, $v_\phi = 1.4\, v_\mathrm{g}$. Any particular wave crest, for example the one indicated by a circle in the figure, moves with speed v_ϕ, and is thus seen to move through the packet envelope, which has speed v_g.

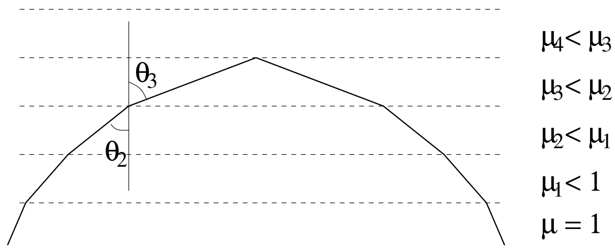


Figure 3.9: Wave propagation in a horizontally stepwise stratified ionosphere.

3.4 Radio wave propagation in the ionosphere

[9] In section 3.2 above, we noted that low frequency EM waves cannot propagate through the ionosphere. A way of studying what happens is to use the well known *Snell's law* from geometrical optics: the angles of incidence θ_1 and θ_2 on different sides of a surface between two media with indices of refraction μ_1 and μ_2 is related by

$$\mu_1 \sin\theta_1 = \mu_2 \sin\theta_2. \tag{3.60}$$

In the terrestrial ionosphere, it will in general be more complicated, as the index of refraction in general depends on the angle of the wave to the magnetic field. We do not consider such complications here, and assume an umnagnetized ionosphere (like on Venus), which means that the dispersion relation (3.45) is applicable. This may be written as

$$\mu^2 = 1 - \omega_{\rm p}^2/\omega^2, \tag{3.61}$$

where $\mu = c/v_\phi$. We consider a horizontally stratified ionosphere, homogeneous in the horizontal direction and with variations in the vertical direction only. If the stratification is stepwise, so that the ionosphere consists of a series of thin layers of different palsma density and index of refraction, we have a situation as in Figure 3.9. If the wave is transmitted from the ground, where $\mu = 1$ as there is no plasma and hence zero plasma frequency, at an angle θ_0 to the vertical, we have

$$\sin\theta_0 = \mu_1 \sin\theta_1 = ... = \mu_j \sin\theta_j. \tag{3.62}$$

In the more general case of a continuously varying plasma density and index of refraction with altitude h, Snell's law reads

$$\mu(h)\,\sin\theta(h) = \sin\theta_0. \tag{3.63}$$

The wave will be reflected at the altitude $h_{\rm r}$ where $\theta(h_{\rm r}) = 90°$, i.e. where

$$\mu(h_{\rm r}) = \sin\theta_0. \tag{3.64}$$

In particular, for a vertical wave, reflection will occur when $\mu = 0$. We define the critical frequency of the ionosphere $\omega_{\rm crit}$ to be the maximum plasma frequency, i.e. the plasma frequency on the altitude where the ionosphere is most dense. A vertically transmitted wave will not be reflected if $\omega > \omega_{\rm crit}$. An obliquely transmitted wave may be reflected even if it has higher frequency. Reflection occurs at the altitude $h_{\rm r}$ where

$$\sin\theta_0 = \mu(h_{\rm r}) = \sqrt{1 - \omega_{\rm p}^2(h_{\rm r})/\omega^2}, \tag{3.65}$$

implying that an oblique wave will be reflected if

$$\omega < \omega_{\rm crit}/\sqrt{1 - \sin^2\theta_0} = \omega_{\rm crit}/\cos\theta_0. \tag{3.66}$$

A consequence of this is illustrated in Figure 3.10.

In what we have done here, we have tacitly assumed that the frequency of the wave stays constant, while the wavelength changes as we go into regions with different refractive index. It may not be evident

[9] This section is based on Bengt Lundborg's lecture notes.

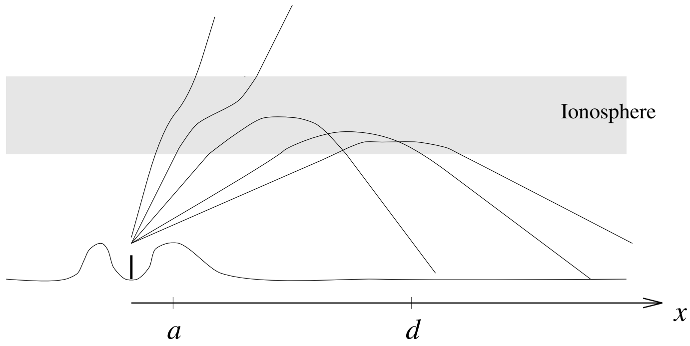


Figure 3.10: A wave with frequency above the critical frequency will propagate out in space if transmitted at a small angle to the vertical, but be reflected for larger angles. This implies that if line-of-sight propagation is prohibited by for instance mountains, there may be a region $a < x < d$ which is neither reached by the direct ray along the line of sight nor by the reflected wave from the ionosphere. This region is called the skip zone.

to everyone why this is so – why isn't it the wavelength that stays constant and the frequency that changes? The reason is that as we assume variations in space in the medium, not variations in time, it is the spatial property of the wave, i.e. the wavelength, that should change, not the temporal quantity of frequency.[10] As long as the horizontal stratification we have assumed is constant in time, the frequency simply cannot change.

3.5 The ionosonde

For a given transmission angle to the vertical, low frequencies will be reflected at lower altitude in the ionosphere than high frequencies. This is used by an instrument called the *ionosonde*, by which the density profile of the ionosphere is studied. If we neglect the influence of the magnetic field, the speed of a pulse with center frequency ω emitted from the ground is the group velocity, which from equation (3.56) is[11]

$$v_{\mathrm{g}} = c^2 k/\omega = c\mu = c\sqrt{1 - \omega_{\mathrm{p}}^2/\omega^2}. \tag{3.67}$$

The time for the pulse to go from the ground ($h = 0$) up to the reflection altitude $h = h_r$ and back again will be

$$T = 2\int_0^{h_r} \frac{dh}{v_{\mathrm{g}}(h)} = \frac{2}{c}\int_0^{h_r} \frac{dh}{\mu(h)}. \tag{3.68}$$

We get different times for different frequencies. By recording $T(\omega)$, one may calculate the density profile $n_{\mathrm{e}}(h)$ up to the altitude of the critical frequency. Above this height, we get no reflection at all. The problem of determining $n_{\mathrm{e}}(h)$ from $T(\omega)$ has a unique solution only if the density is monotonically increasing up to its maximum. A diagram $T(\omega)$ recorded by an ionosonde is called an *ionogramme* (figure 3.11).

[10]Those versed in quantuym theory may note that the frequency and wave number relates to the energy and linear momentum of a quantum of the oscillation as $E = \hbar\omega$ and $\mathbf{p} = \hbar\mathbf{k}$. Energy is conserved in stationary system, while momentum is conserved only in force-free configurations, i.e. systems homeogeneous in space.

[11]Equation (3.67) may be rewritten as $\omega^2 v_{\mathrm{g}}^2(h) = \omega^2 c^2 - \omega_{\mathrm{p}}^2(h) c^2$, which may be compared to the energy expression for a stone you throw into the air with speed v_0 from the ground: $\frac{1}{2}mv^2(h) = \frac{1}{2}mv_0^2 - mgh$. The inertia of the stone corresponds to the square of the frequency, the initial speed of the wave is c, and its "potential energy" is $\omega_{\mathrm{p}}(h)$ in this analogy.

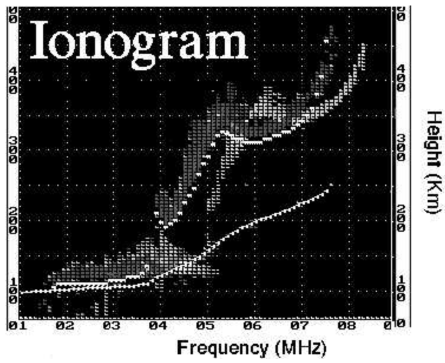


Figure 3.11: Example of an ionogramme, where the propagation time t has been converted to reflection height h.

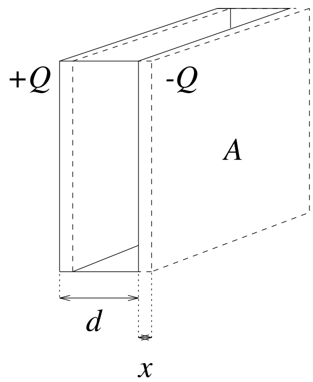


Figure 3.12: Illustration of plasma oscillations. All electrons in the slab of width d are moved a distance x to the right, creating two regions of net charge.

3.6 Plasma oscillations

In section 3.1, we found that two different types of waves could exist in the cold unmagnetized plasma. The longitudinal wave had the remarkable dispersion relation (3.43),

$$\boxed{\omega^2 = \omega_{\mathrm{p}}^2} \tag{3.69}$$

telling us that this oscillation exist only at one single frequency, the plasma frequency, irrespective of wavelength. The phase velocity ω/k simply is ω_{p}/k, i.e., proportional to the wavelength, and the group velocity is zero. What strange wave is this?

In fact it is not really a wave, but rather a stationary eigenoscillation of the plasma, known as the *plasma oscillation*. Assume that we in a plasma move all electrons (but not the ions) within a slab of area A and thickness d, where $d \ll \sqrt{A}$, a distance $x \ll d$ to the right (see Figure 3.12). Net charge will appear in two places: a positive net charge in a slice of thickness x at the left surface of the slab Ad, due to all electrons being removed from here, and a corresponding negative charge at the right surface (Figure 3.12), where the electron density increases. These charges give rise to an electric field, which forces the electrons back to their origonal position. The ions are also affected by the field, but they are much heavier than the electrons and are thus quite immobile compared to the light electrons. When the electrons reach their original positions, there is charge balance and the electric field vanishes. However, the electrons now have kinetic energy, and their inertia makes them pass by the equilibrium point and continue to the left. The charge separation which then appears builds up a new electric field, which stops the electrons and drags them back, and so on. The electrons are oscillating around the equilibrium position with a frequency which we now will determine.

If the electron density is n, the charge in the two slices of thickness x where there is a net charge will be $Q = neAx$. The system may be seen as a plate capacitor with charge Q and distance d between the plates. Such a capacitor has capacitance $C = \epsilon_0 A/d$, so the voltage over it is $U = Q/C = nexd/\epsilon_0$. The electric field between the plates thus is $E = U/d = nex/\epsilon_0$. The restoring force on an electron then will be $F = qE = -ne^2x/\epsilon_0$, so the electron equation of motion is

$$m_\mathrm{e}\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -\frac{ne^2}{\epsilon_0}x. \tag{3.70}$$

The solutions to this differential equation are oscillations at angular frequency

$$\omega = \sqrt{\frac{ne^2}{\epsilon_0 m_\mathrm{e}}} = \omega_\mathrm{p}. \tag{3.71}$$

Hence, the plasma frequency is the oscillation frequency of small charge imbalances in the plasma.

We may note that the plasma oscillation is completely *electrostatic* in the sense defined in section 1.3: the wave vector $\mathbf{k}$ is parallel to the electric field $E_x\hat{\mathbf{x}}$ (equation (3.41)). In chapter 4, we will show that when the pressure is taken into account, the plasma can support propagating *electrostatic waves*, not only the stationary plasma oscillation.

Problems for Chapter 3

1. *Electromagnetic waves.* Derive the dispersion relation for electromagnetic waves (no electrostatic component) in a cold homogeneous unmagnetized plasma consisting of Ca^{2+} and Cl^- with densities $n_{\mathrm{Cl}^-} = 2\,n_{\mathrm{Ca}^{2+}} = 2\,n$.

2. *Interrupted communications.* A spacecraft re-entering the atmosphere can have its radio communication interrupted at frequencies higher than the critical frequency of the ionosphere. Can you think of any explanation?

3. *Skin depth.* A cold unmagnetized slab of plasma occupies the region between $z = 0$ and $z > a$. The plasma density is such as to give $\omega_\mathrm{p} = 2\pi$ rad s^{-1}. Estimate the skin depth, i.e. the distance to which electromagnetic wave fields with frequencies below ω_p will penetrate into the plasma.

4. *Wave energy.* The energy density of an electromagnetic wave in a plasma is contained in the electric wave field ($w_E = \varepsilon_0 E^2/2$), the magnetic wave field ($w_B = B^2/(2\mu_0)$), and the kinetic energy associated with the electron veocity field ($w_K = nm_\mathrm{e}v_\mathrm{e}^2/2$). Calculate the instantaneous and time average (over a wave period) values of these quantities. How do the ratios $< w_B > / < w_E >$ and $< w_K > / < w_E >$ vary with ω?

5. *Wave properties.* An electromagnetic wave with wavelength 100 m and amplitude 10 mV/m is propagating in a plasma of density $n = 10^{10}$ m^{-3}. Calculate the following properties:

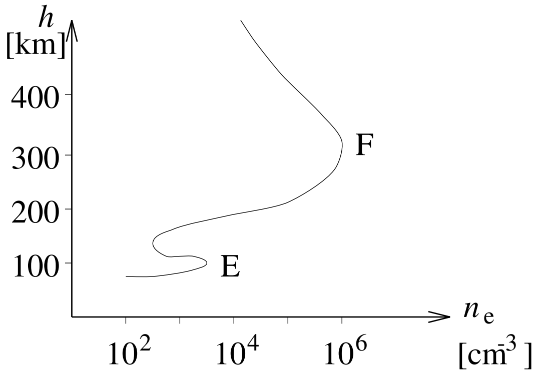
   (a) Frequency, phase velocity and group velocity

   (b) Amplitudes of the magnetic wave field and the electron velocity field

   (c) Average (over a wave period) energy densities in the electric, magnetic and elctron velocity wave fields, average total energy density and average energy flux

   Hint: The average energy flux can be calculated either by considering the average Poynting flux or the average energy density times the group velocity. What is the difference between those methods? Is it practically important?

6. *Critical frequency.* Estimate the critical frequency for the ionospheric profile below.



7. *Pulsar wave dispersion.* As the group velocity is frequency dependant, a pulse containing several frequencies will change its form as it travels through space, so that the frequencies with higher group speed arrive before frequencies with lower group speed. This effect can be seen in the radiation from pulsars, who emit broadband pulses of electromagnetic waves. Show that if $\nu_{\mathrm{p}}^2 \ll \nu^2$, the observed variation of frequency ν with time in the pulsar emission will be

$$\frac{\mathrm{d}\nu}{\mathrm{d}t} \approx -\frac{c}{r}\frac{\nu^3}{\nu_{\mathrm{p}}^2}$$

where r is the distance to the pulsar. If the average interstellar plasma density is 0.1 cm^{-3} and $\mathrm{d}\nu/\mathrm{d}t = -5$ MHz/s is measured on ground for $\nu = 80$ MHz, what is the distance to the pulsar?

8. *Ionospheric wave propagation.* Consider the following model for an unmagnetized ionosphere, on Venus or Titan for instance. The geometry is assumed to be planar (horizontal stratification), and the electron density is given by $n(h) = N\exp(h/L)$. What is the maximum altitude h_0 reached by electromagnetic waves of frequency ω, if they were transmitted from the ground at an angle α to the vertical?

9. *Ionospheric wave propagation.* A radio wave of frequency ν_0 is transmitted vertically from the ground and is reflected at altitude h_0. What frequency should a wave transmitted at an angle ϕ to the vertical have in order to be reflected at the same altitude h_0?

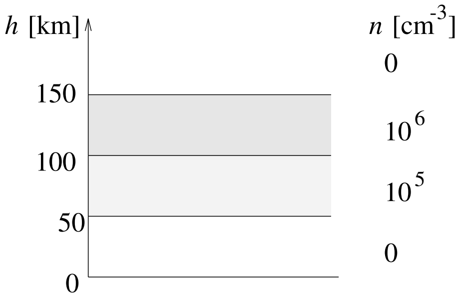
10. *Ray tracing.* The hypothetical planet C16G is so big that its surface can be considered flat and has an ionosphere where the plasma density n varies with altitude as $n = Ny^2/a^2$, where N and a are constants. Let x be a coordinate along the planetary surface. A transmitter at the origin tranmits a radio wave of angular frequency ω in the xy plane at an angle θ_0 to the y axis.

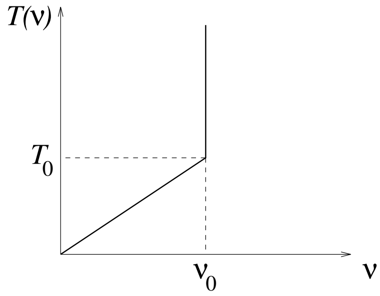
    (a) Show that the ray path is described by the differential equation

    $$\frac{\mathrm{d}y}{\mathrm{d}x} = \pm\frac{\sqrt{\cos^2\theta_0 - \frac{Ne^2y^2}{\varepsilon_0 m_{\mathrm{e}} a^2\omega^2}}}{\sin\theta_0}$$

    (b) Calculate the ray path on the form $y = f(x)$.

11. *Skip zone.* The remarkable planet Qfrxnypladugh-Z, yet to be discovered, lacks intrinsic magnetic field, and has an ionosphere consisting of two plane homogeneous sheaths as in the left figure below. A radio transmitter on the ground operates at 10 MHz. The transmitter is surrounded by mountains, so the direct wave along the line of sight can be neglected. Determine the shortest distance from the transmitter which is reached by the rays. (The region within this distance is known as the skip zone; see also Figure 3.10.)





12. *Ionogramme.* The figure at right above shows an ionogramme trace $T(\nu)$, where ν is the frequency and T the propagation time as defined by equation (3.68). Can you find an ionospheric plasma density profile $n(h)$ resulting in this ionogramme?

Chapter 4

Electrostatic waves

4.1 Langmuir waves

In a cold plasma, there cannot be any waves corresponding to the sound waves in a neutral gas as the term ∇p in the equation of motion is zero when $T = 0$. In a warm plasma, the electron equation of motion is

$$m_{\mathrm{e}} n_{\mathrm{e}}(t, \mathbf{r}) \frac{\mathrm{d}\mathbf{v}_{\mathrm{e}}(t, \mathbf{r})}{\mathrm{d}t} = -\nabla p_{\mathrm{e}}(t, \mathbf{r}) - e\, n_{\mathrm{e}}(t, \mathbf{r})[\mathbf{E}(t, \mathbf{r}) + \mathbf{v}(t, \mathbf{r}) \times \mathbf{B}(t, \mathbf{r})] \tag{4.1}$$

which means that waves corresponding to sound waves may exist. The derivation here is similar to other derivations of dispersion relations in previous chapters: write down the field equations, linearize them, look for sine wave solutions, and eliminate the fields to get the dispersion relation. Let us consider waves with so high frequency that the ions cannot follow the motion of the electrons because of their much higher mass and inertia. The ion density may then be assumed to be constantly $n_{\mathrm{i}} = n_0$. A relation between electron velocity and density is found from the continuity equation,

$$\frac{\partial n_{\mathrm{e}}(t, \mathbf{r})}{\partial t} + \nabla \cdot [n_{\mathrm{e}}(t, \mathbf{r})\mathbf{v}_{\mathrm{e}}(t, \mathbf{r})] = 0. \tag{4.2}$$

For simplicity, the pressure-density relation is assumed to be the equation of state for an ideal gas at isothermal[1] conditions, (2.20):

$$p(t, \mathbf{r}) = n(t, \mathbf{r})KT. \tag{4.3}$$

We confine ourselves to the study of pure *electrostatic waves*, i.e. waves with electric field satisfying $\nabla \times \mathbf{E} = 0$. The electric field may then be described by the electrostatic potential Φ,

$$\mathbf{E}(t, \mathbf{r}) = -\nabla\Phi(t, \mathbf{r}), \tag{4.4}$$

which is given by Gauss' law for the electric field (3.3),

$$\nabla^2\Phi(t, \mathbf{r}) = -\nabla \cdot \mathbf{E}(t, \mathbf{r}) = -\rho/\epsilon_0 = e(n_{\mathrm{e}}(t, \mathbf{r}) - n_0)/\epsilon_0 = e n_{1\mathrm{e}}(t, \mathbf{r})/\epsilon_0 \tag{4.5}$$

where we have used notation as before. The last equation is linear and may be Fourier transformed at once, while the other equations must be treated by the linearization methods of section 2.4. This yields

$$m_{\mathrm{e}} n_0 \frac{\partial \mathbf{v}_{1\mathrm{e}}(t, \mathbf{r})}{\partial t} = n_0 e \nabla\Phi_1(t, \mathbf{r}) - KT\nabla n_{1\mathrm{e}}(t, \mathbf{r}) \tag{4.6}$$

$$\frac{\partial n_{1\mathrm{e}}(t, \mathbf{r})}{\partial t} + n_0 \nabla \cdot \mathbf{v}_{1\mathrm{e}}(t, \mathbf{r}) = 0. \tag{4.7}$$

(4.5) – (4.7) form a system of five linear equations for five unknowns (Φ_1, $n_{1\mathrm{e}}$ och $\mathbf{v}_{1\mathrm{e}}$). We look for plane wave solutions, and get

$$k^2\Phi_1 = -\frac{e}{\epsilon_0} n_{1\mathrm{e}} \tag{4.8}$$

[1] The heat conduction at high frequencies in a collisionless plasma is very small, so a better approximation is the adiabatic condition $p = Cn^{\gamma}$, where $\gamma = 3$ for the one-dimensional case (wave propagation in one given direction) we study here. See also footnote 2.

$$-i\omega m_{\rm e} n_0 \mathbf{v}_{\rm e} = i\mathbf{k}\left[n_0 e\Phi_1 - n_{1\rm e} K T_{\rm e}\right] \tag{4.9}$$

$$-i\omega n_{1\rm e} + i n_0 \mathbf{k}\cdot\mathbf{v}_{\rm e} = 0. \tag{4.10}$$

From these equations, we have

$$\omega n_{1\rm e} = n_0 \mathbf{k}\cdot\mathbf{v}_{\rm e} = \frac{n_0}{\omega m_{\rm e} n_0}\mathbf{k}\cdot\left[\mathbf{k}\left(K T_{\rm e} n_{1\rm e} - n_0 e\Phi_1\right)\right] = \frac{k^2}{\omega m}\left(K T_{\rm e} n_{1\rm e} + \frac{n_0 e^2}{\epsilon_0 k^2} n_{1\rm e}\right), \tag{4.11}$$

so the dispersion relation is[2]

$$\boxed{\omega^2 = \omega_{\rm p}^2 + k^2 v_{\rm e}^2} \tag{4.12}$$

where we used the definition (3.40), and introduced a characteristic electron speed[3]

$$v_{\rm e} = \sqrt{\frac{K T_{\rm e}}{m_{\rm e}}}. \tag{4.13}$$

The waves described by this dispersion relation are called *electron plasma waves* or *Langmuir waves*.

Comparing to the dispersion relation (2.36), we may consider the Langmuir waves as pressure waves in the electron gas. If the dispersion relation is rewritten in terms of the range of the electrostatic field of a particle in the plasma, the Debye length

$$\lambda_{\rm D} = \sqrt{\frac{\epsilon_0 K T_{\rm e}}{n_0 e^2}} = v_{\rm e}/\omega_{\rm p}, \tag{4.14}$$

we get

$$\omega^2 = \omega_{\rm p}^2 (1 + k^2 \lambda_{\rm D}^2). \tag{4.15}$$

For short wavelengths, $2\pi/k \ll \lambda_{\rm D}$, the second term dominates, and the dispersion relation becomes the same as for pressure waves in a neutral gas. This is reasonable, as in this limit the long-range collective effects of the coherent motion of many particles which characterizes a plasma disappears: indeed, the definition of a plasma requires system dimensions to be longer than λ_D. For long wavelengths, $2\pi/k > \lambda_D$, the first term is important, and plasma effects enter the wave behaviour. In the limit of very long wavelengths, we get the plasma oscillation (3.69), which has no counterpart in a neutral gas.

The Langmuir wave (4.12) is a generalization of the plasma oscillation $\omega^2 = \omega_{\rm p}^2$ of cold plasma theory (page 26). The odd feature $v_g = 0$ of the plasma oscillation is not present when we include thermal (pressure) effects, and for the Langmuir waves, we get $v_g = v_{\rm e}\sqrt{1 - \omega_{\rm p}^2/\omega^2}$.

4.2 Ion acoustic waves

The plasma waves we have seen up to now, the electrostatic waves (4.12) as well as the electromagnetic waves (3.45), all propagate only above $\omega_{\rm p}$. For the Langmuir waves, we explicitly assumed high frquency in the derivation. We now do the opposite assumption: assuming waves of so low frequency that the electrons have plenty of time to find an equilibrium. They will then be distributed in space according to the Boltzmann relation

$$n_{\rm e}(t,\mathbf{r}) = n_0 \exp\left(\frac{e\Phi(t,\mathbf{r})}{K T_{\rm e}}\right) \tag{4.16}$$

as their potential energy is $-e\Phi$. As usual, we linearize the equation for waves of small amplitude. Then Φ is assumed small, and we only keep the first term in the Taylor expansion of the exponential,

$$n_{\rm e}(t,\mathbf{r}) \approx n_0\left(1 + \frac{e\Phi(t,\mathbf{r})}{K T_{\rm e}}\right) \tag{4.17}$$

[2]With the definition of $v_{\rm e}$ used here (see footnote 3), a derivation based on adiabatic rather than isothermal conditions yields $\omega^2 = \omega_{\rm p}^2 + 3k^2 v_{\rm e}^2$.

[3]Related to the usually defined electron thermal speed, $v_{\rm th} = \sqrt{2KT_{\rm e}/m_{\rm e}}$.

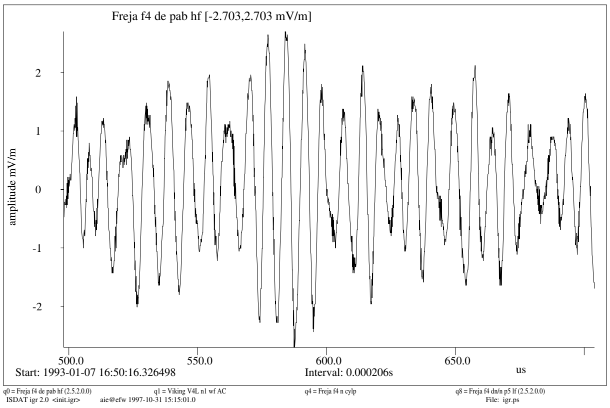


Figure 4.1: Langmuir waves observed by the Freja satellite in auroral regions. Estimate the plasma density!

Hence,

$$n_{1\mathrm{e}}(t,\mathbf{r}) = n_0 \frac{e\Phi(t,\mathbf{r})}{KT_\mathrm{e}} \tag{4.18}$$

where we used the normal notation (see equation (3.12) and following). As the electrons are much lighter than the ions, they will essentially short-circuit any charge imbalance caused by the ion motion, so

$$n_{1\mathrm{i}} = n_{1\mathrm{e}}. \tag{4.19}$$

If we again assume isothermal[4] conditions, the ion equation of motion is

$$m_\mathrm{i} n_0 \frac{\partial \mathbf{v}_{1\mathrm{i}}(t,\mathbf{r})}{\partial t} = -n_0 e \nabla \Phi_1(t,\mathbf{r}) + KT_\mathrm{i} \nabla n_{1\mathrm{i}}(t,\mathbf{r}) \tag{4.20}$$

and their equation of continuity is

$$\frac{\partial n_\mathrm{i}(t,\mathbf{r})}{\partial t} + \nabla \cdot [n_\mathrm{i}(t,\mathbf{r})\mathbf{v}_\mathrm{i}(t,\mathbf{r})] = 0. \tag{4.21}$$

After the usual linearization and Fourier transform procedure, we get

$$-i\omega m_\mathrm{i} n_0 \mathbf{v}_\mathrm{i} = -i\mathbf{k}[n_0 e\Phi + n_{1\mathrm{i}} KT_\mathrm{i}] \tag{4.22}$$

$$-i\omega n_{1\mathrm{i}} + i n_0 \mathbf{k}\cdot\mathbf{v}_\mathrm{i} = 0. \tag{4.23}$$

Combining these two eqations with (4.18) and (4.19) yields

$$\omega^2 n_{1\mathrm{i}} = \omega n_0 \mathbf{k}\cdot\mathbf{v}_\mathrm{i} = \frac{1}{m_\mathrm{i}}\mathbf{k}\cdot[\mathbf{k}(n_{1\mathrm{i}}KT_\mathrm{i} + n_0 e\Phi)] = \frac{k^2}{m_\mathrm{i}}(KT_\mathrm{i} + KT_\mathrm{e})n_{1\mathrm{i}}, \tag{4.24}$$

giving the dispersion relation

$$\boxed{\omega^2 = c_\mathrm{ia}^2 k^2} \tag{4.25}$$

where the ion acoustic speed c_ia is defined by [5]

$$c_\mathrm{ia}^2 = \frac{KT_\mathrm{i} + KT_\mathrm{e}}{m_\mathrm{i}}. \tag{4.26}$$

This wave mode is called the *ion acoustic wave*. We note that in the dispersion relation, we find the mass of the ions, but the temperature for ions as well as electrons. A common situation in space plasmas is to have

[4] As usual, an adiabatic approximation would be more physical, but we stick to the isothermal approximation for simplicity.

[5] Assuming adiabatic ions with $\gamma = 3$ would have resulted in the same dispersion relation (4.25) with $c_\mathrm{ia}^2 = (3KT_\mathrm{i} + KT_\mathrm{e})/m_\mathrm{i}$.

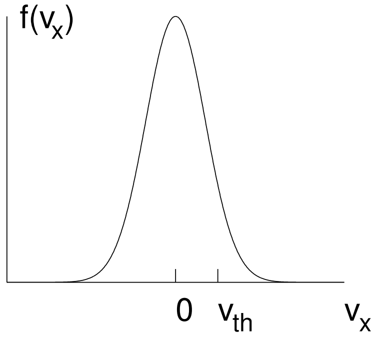


Figure 4.2: Maxwellian velocity distribution in an equilibrium plasma.

$T_\mathrm{i} \ll T_\mathrm{e}$, giving a dispersion relation $\omega^2 = (KT_\mathrm{e}/m_\mathrm{i})k^2$. The properties of the wave are then determined by the ion inertia and the electron temperature. This may be understood as follows. The oscillations are an ongoing energy transformation between particle kinetic energy and potential energy in the wave electrostatic field. As the ions are much heavier than the electrons, their kinetic energy is dominating, and thus the ion mass is important. If the electrons as well as the ions were cold, they would quickly short-circuit the wave electric field, as it has a frequency far below the plasma frequency, which is the frequency where electron inertia becomes important, and there would be no wave. However, as they have a certain temperature, the screening of the electric field is not perfect. Compare to how the electrons cannot completely neutralize a charge imbalance in Debye screening of a stationary charge. As in the Debye case, the screening is less efficient the higher the electron temperature is, making it easier for the wave to propagate, i.e. increasing its speed.

Thus, when considering the ion motion in a thermal plasma, we find an electrostatic mode propagating at frequencies below the plasma frequency.

4.3 Landau damping and kinetic instabilities

A wave may be damped by collisions among the particles in the medium, transfering the ordered energy in the wave to unordered thermal motion in the medium. This means the medium get heated. In a collissionless plasma, like in the magnetosphere or the solar wind, this damping is absent. Still, waves may be damped by what is known as Landau damping.

Consider a plasma in thermodynamic equilibrium with a slight perturbation due to the motion associated with the waves. The electrons in the plasma are then distributed in velocity space as described by the Maxwell-Boltzmann distribution, illustrated in figure 4.2. Consider an electrostatic wave, a Langmuir wave for instance, with phase velocity v_ϕ. Taking the Fourier transform of (4.4), we get

$$\mathbf{E} = -i\mathbf{k}\Phi, \tag{4.27}$$

implying that the wave electric field is parallel to the direction of propagation $\mathbf{k}$. Now consider an electron with speed $v \approx v_\phi$ in the same direction as the wave. This electron will see an electric field which is almost constant in time, as it travels with the wave. We say the particle is in *Landau resonance* with the wave. The consequence of the electron seeing essentially the same field all the time is that it will be accelerated or retarded by the wave electric field, depending on if it is located in a part of the wave field where its velocity $\mathbf{v}$ is parallel or antiparallel to $\mathbf{E}$. After a while, the electron has been accelerated/retarded so much that it no longer is in Landau resonance, so it no longer sees the wave field as constant as it now travels faster or slower than the wave. Thereby further exchange of energy between the wave and the particle is prohibited, but as long as Landau resonance was present, energy was transfered between them. This energy exchange is most efficient when the particle is in perfect Landau resonance, $v = v_\phi$.

Let us assume the velocity of the electron is slightly lower than the wave phase velocity: $v < v_\phi$. If it is retarded, it will never come in perfect Landau resonance, and energy transfer will not be very efficient. If

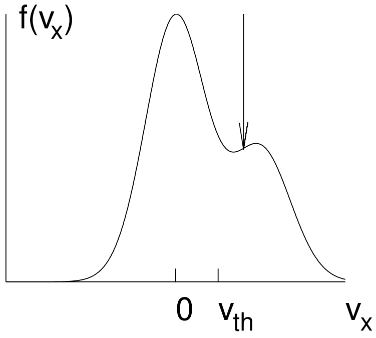


Figure 4.3: Example of an unstable velocity distribution. Waves with phase speed such that $\partial f/\partial v > 0$ (region indicated by arrow) will grow by a process of "inverse Landau damping". Energy goes to the wave from the particles around the maximum in the distribution function. The maximum will therefore be levelled out until there is no positive slope $\partial f/\partial v$.

it is accelerated, then it will reach $v = v_\phi$ where the energy transfer is most efficient, and thus get further accelerated. Energy transfer between the wave field and electrons with speed slightly below the wave phase speed is thus most efficient in the direction from the wave to the particles. If we instead consider particles with slightly higher speed than v_ϕ, a similar argument shows that energy in this case flows most efficient from the particles to the wave. Looking at many particles, the statistical result will be that the wave gives energy to the particles that are slower than the wave, and takes energy from the particles that are faster. Now, from Figure 4.2 it is clear that in an equilibrium plasma, there will always be more of the slower particles than of the faster. Thus, the net result is that energy is converted from electric field energy in the wave to kinetic energy of the particles. This mechanism, known as *Landau damping*, was discovered theoretically by Landau before it could be verified by measurements.

If the electron distribution function looks like Figure 4.3, there is an interval in velocity space where energy will flow from particles to wave: in the region where $\partial f/\partial v > 0$, there are more particles with speed slightly above v_ϕ than slightly below, and here the net energy flow between particles and wave will be reversed, using the same argumentation as above. Thermal fluctuations in the plasma make sure that there always exist some little wave fluctuation of any wavelength, and the fluctuations with v_ϕ in this speed interval will therefore grow. This is an example of a *plasma instability*, where the plasma emits a wave.

4.4 Beam instability

We will now study a simple quantitative model of a situation where waves are generated in a space plasma. From satellite observations, we know that plasma waves are observed together with field-aligned (Birkeland) currents in the auroral regions of the magnetosphere. To understand why, we model the plasma as a population of stationary electrons with density n, through which streams electrons with density Δn and velocity $\mathbf{u}$. For simplicity, we will only study waves of such high frequency that the ions can be considered stationary (compare section 4.1). We assume that the electrons are cold, so that their distribution function is

$$f(\mathbf{v}) = n\delta(\mathbf{v}) + \Delta n\delta(\mathbf{v} - \mathbf{u}). \tag{4.28}$$

A more realistic model would be to assume some spread in velocity space by assigning non-zero temperatures T and T_s for the stationary and streaming electrons, respectively, in which case the distribution function would be a sum of two Maxwell distributions,

$$\begin{aligned} f(\mathbf{v}) \quad &= \quad n\left(\frac{m_\mathrm{e}}{2\pi KT}\right)^{3/2} \exp\left(-\frac{m_\mathrm{e}\mathbf{v}^2}{2KT}\right) + \\ &\quad +\Delta n\left(\frac{m_\mathrm{e}}{2\pi KT_\mathrm{s}}\right)^{3/2} \exp\left(-\frac{m_\mathrm{e}(\mathbf{v}-\mathbf{u})^2}{2KT_\mathrm{s}}\right). \end{aligned} \tag{4.29}$$

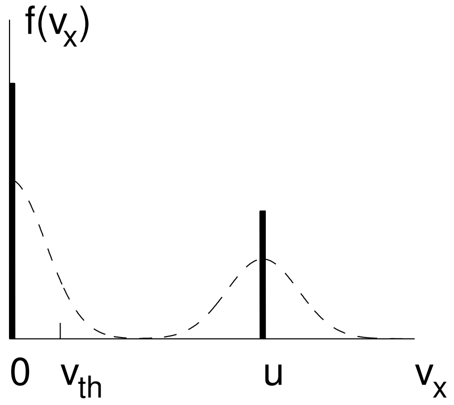


Figure 4.4: Delta velocity distribution used to model a current-carrying plasma (solid bars), and a somewhat more realistic model of two maxwellians (dashed).

In the limit of $T \to 0$ and $T_{\rm s} \to 0$, this collapses to the delta distribution (4.28). The two cases are illustrated in Figure 4.4.

Obviously, the non-zero temperature distribution should be unstable according to the qualitative discussion in the previous section (compare Figures 4.4 and 4.3). The zero-temperature case (cold electrons, delta function distribution) should be a reasonable approximation for the case when $KT \ll \frac{1}{2}m_{\rm e}u^2$ and $KT_{\rm s} \ll \frac{1}{2}m_{\rm e}u^2$, so it is not unreasonable to assume that the basic physics of the instability should be present in the simplified zero-temperature case.

We look for electrostatic waves ($\mathbf{E} = -\nabla\Phi$) in one dimension, so that the wave vector and electron oscillation velocity for any waves which may appear are parallel to the direction of the electron stream $\mathbf{u}$. For a field-aligned current, this direction is along the magnetic field, so there will be no $\mathbf{v} \times \mathbf{B}$ force on the electrons. When the temperatures are zero, so are the pressures. Using index p for the stationary electrons and index s for the streaming population, the equations of motion for the plasma and beam electrons are

$$m_{\rm e}\frac{{\rm d}v_{\rm p}}{{\rm d}t} = e\frac{\partial\Phi}{\partial z} \tag{4.30}$$

and

$$m_{\rm e}\frac{{\rm d}v_{\rm s}}{{\rm d}t} = e\frac{\partial\Phi}{\partial z}, \tag{4.31}$$

where z is a coordinate along the magnetic field. To describe the physics of the situation, we also have the continuity equations for the two electron populations,

$$\frac{\partial n_{\rm p}}{\partial t} + \frac{\partial}{\partial z}(n_{\rm p}v_{\rm p}) = 0 \tag{4.32}$$

$$\frac{\partial n_{\rm s}}{\partial t} + \frac{\partial}{\partial z}(n_{\rm s}v_{\rm s}) = 0, \tag{4.33}$$

and finally, Gauss' law for the electric field,

$$\frac{\partial^2\Phi}{\partial z^2} = \frac{e}{\epsilon_0}(n_{\rm p} + n_{\rm s} - (1+\Delta)n) \tag{4.34}$$

where we have used that the ion density must be $n + \Delta n = (1+\Delta)n$ in order to keep the plasma macroscopically neutral.

The perturbation ansatz for this case is

$$\begin{array}{rcccl} n_{\rm p}(t,\mathbf{r}) & = & n & + & n_{1\rm p}(t,\mathbf{r}) \\ n_{\rm s}(t,\mathbf{r}) & = & \Delta n & + & n_{1\rm s}(t,\mathbf{r}) \\ \Phi(t,\mathbf{r}) & = & & & \Phi_1(t,\mathbf{r}) \\ v_{\rm p}(t,\mathbf{r}) & = & & & v_{1\rm p}(t,\mathbf{r}) \\ v_{\rm s}(t,\mathbf{r}) & = & u & + & v_{1\rm s}(t,\mathbf{r}). \end{array} \tag{4.35}$$

The difference from our earlier treatments of waves in plasmas without streaming particles is the u-term in the ansatz for the velocity for the speed of the beam electrons. This is a very important detail, because it results in new terms in some of the linearized equations. In equation (4.31), we get

$$\begin{aligned} \frac{\mathrm{d}v_{\mathrm{s}}}{\mathrm{d}t} &= \frac{\partial v_{\mathrm{s}}}{\partial t} + v_{\mathrm{s}} \frac{\partial v_{\mathrm{s}}}{\partial z} = \\ &= \frac{\partial v_{1\mathrm{s}}}{\partial t} + (u + v_{1\mathrm{s}}) \frac{\partial v_{1\mathrm{s}}}{\partial z} \approx \\ &\approx \frac{\partial v_{1\mathrm{s}}}{\partial t} + u \frac{\partial v_{1\mathrm{s}}}{\partial z}, \end{aligned} \tag{4.36}$$

and in equation (4.33),

$$\begin{aligned} \frac{\partial}{\partial z}(n_{\mathrm{s}} v_{\mathrm{s}}) &= \frac{\partial}{\partial z}[(\Delta n + n_{1\mathrm{s}})(u + v_{1\mathrm{s}})] \approx \\ &\approx \Delta n \frac{\partial v_{1\mathrm{s}}}{\partial z} + u \frac{\partial n_{1\mathrm{s}}}{\partial z}. \end{aligned} \tag{4.37}$$

Therefore, linearization and Fourier transformation of equations (4.30) –(4.34) yields

$$-i\omega m_{\mathrm{e}} v_{1\mathrm{p}} = ike\Phi_1 \tag{4.38}$$

$$\Rightarrow \quad v_{1\mathrm{p}} = -\frac{e}{m_{\mathrm{e}}\omega} k\Phi_1 \tag{4.39}$$

$$m_{\mathrm{e}}(-i\omega v_{1\mathrm{s}} + ikuv_{1\mathrm{s}}) = ike\Phi_1 \tag{4.40}$$

$$\Rightarrow \quad v_{1\mathrm{s}} = -\frac{e}{m_{\mathrm{e}}(\omega - ku)} k\Phi_1 \tag{4.41}$$

$$-i\omega n_{1\mathrm{p}} + iknv_{1\mathrm{p}} = 0 \tag{4.42}$$

$$\Rightarrow n_{1\mathrm{p}} = \frac{n}{\omega} kv_{\mathrm{p}} = -\frac{ne}{m_{\mathrm{e}}\omega^2} k^2\Phi_1 \tag{4.43}$$

$$-i\omega n_{1\mathrm{s}} + ik\Delta nv_{1\mathrm{s}} + ikun_{1\mathrm{s}} = 0 \tag{4.44}$$

$$\Rightarrow n_{1\mathrm{s}} = \frac{\Delta n}{(\omega - ku)} kv_{1\mathrm{s}} = -\frac{\Delta ne}{m_{\mathrm{e}}(\omega - ku)^2} k^2\Phi_1 \tag{4.45}$$

$$-k^2\Phi_1 = \frac{e}{\epsilon_0}(n_{1\mathrm{p}} + n_{1\mathrm{s}}) \tag{4.46}$$

$$\begin{aligned} \Rightarrow \quad -k^2\Phi_1 &= -\frac{e}{\epsilon_0}\left(\frac{ne}{m_{\mathrm{e}}\omega^2} k^2\Phi_1 + \frac{\Delta ne}{m_{\mathrm{e}}(\omega - ku)^2} k^2\Phi_1\right) = \\ &= -\left(\frac{\omega_{\mathrm{p}}^2}{\omega^2} + \frac{\Delta\omega_{\mathrm{p}}^2}{(\omega - ku)^2}\right) k^2\Phi_1 \end{aligned} \tag{4.47}$$

We thus find the dispersion relation in the plasma with electron stream to be

$$\boxed{\frac{\omega_{\mathrm{p}}^2}{\omega^2} + \frac{\Delta\omega_{\mathrm{p}}^2}{(\omega - ku)^2} = 1,} \tag{4.48}$$

where ω_{p} is the plasma (angular) frequency computed for density n. For $\Delta \to 0$, i.e. when the density of the streaming electrons goes to zero so that the current disappears, this equation becomes $\omega = \omega_{\mathrm{p}}$. The waves we get are therefore generalizations of the plasma oscillation to a plasma with streaming electrons. If we had included thermal effects by putting $T \neq 0$ and including a pressure term, we would have got Langmuir waves.

The dispersion relation (4.48) is a fourth order equation in ω. It may be solved algebraically, but that is a complicated task which yields a rather intransparent solution. Instead, we will study a graphical solution. Let x be the frequency measured in units of the plasma frequency, $x = \omega/\omega_{\mathrm{p}}$, and put $\kappa = ku/\omega_{\mathrm{p}}$. The dispersion relation may then be written

$$\frac{1}{x^2} + \frac{\Delta}{(x - \kappa)^2} = 1 \tag{4.49}$$

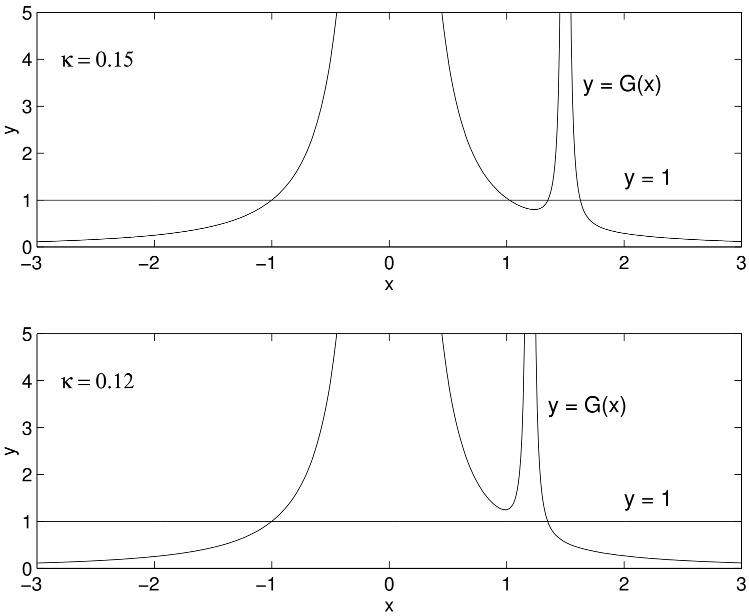


Figure 4.5: Graphical solution of the dispersion relation (4.48).

or simply

$$G(x) = 1, \tag{4.50}$$

where we have defined

$$G(x) = \frac{1}{x^2} + \frac{\Delta}{(x-\kappa)^2}. \tag{4.51}$$

For given values of Δ and κ, the solutions of the dispersion relation are given by the intersection of the curve $y = G(x)$ with the line $y = 1$. If we look at the physics, fixing Δ means fixing the ratio of the densities in the streaming and stationary electron populations. For a given stream velocity u, fixing κ is fixing k, so our graphical solution will tell us about waves with some specified wave length.

Figure 4.5 shows two graphical solutions for $\Delta = 0.01$, i.e. when the 1 % of the electrons are streaming, for two values of κ. In the upper panel, we find four intersections between $y = G(x)$ and $y = 1$. Hence, there are waves with four different frequencies $f = xf_\mathrm{p}$ propagating in the plasma for $\kappa = 0.15$. Two of the roots are close to $x = -1$ and $x = +1$, i.e. to $\omega = \pm\omega_\mathrm{p}$. These corresponds to the usual plasma oscillations. The other two roots are at $x \approx 1.4$ and $x \approx 1.6$. If the rest frame of the beam is denoted by a dash, the transformation of frequency (Doppler shift) between the rest frame of the bulk plasma and the rest frame of the beam is given by $\omega = \omega' + ku$ or $x = x' + \kappa$. The plasma frequency of the beam electrons is $\omega_\mathrm{bp} = \sqrt{\Delta}\omega_\mathrm{p} = \omega_\mathrm{p}/10$, so these two roots represents plasma oscillations of the beam electrons, with wave vector parallel and antiparallel to the beam velocity, respectively.

In the bottom graph, there are only two intersections, and thus only two real solutions to the dispersion relation (4.48), at $x \approx -1$ and $x \approx 1.3$. These corresponds to plasma oscillations in the bulk plasma with wave vector antiparallel to $\mathbf{u}$ and plasma oscillations in the beam with wave vector parallel to $\mathbf{u}$, respectively. However, a fourth degree equation always has four solutions. Thus there must be two complex solutions in addition to the two real roots. Writing a complex solution to (4.48) on the form $\omega = \omega_r + i\gamma$, we find that the wave it describes is of the form

$$e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} = e^{\gamma t}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_r t)}. \tag{4.52}$$

This is a sinusoidal plane wave multiplied by an exponential. If $\gamma < 0$, we have a damped wave, with amplitude decreasing in with time. For $\gamma > 0$, we get growing waves. Will the complex solution to (4.48) describe damped or growing waves? Both! If a fourth order algebraic equation with real coefficients, like

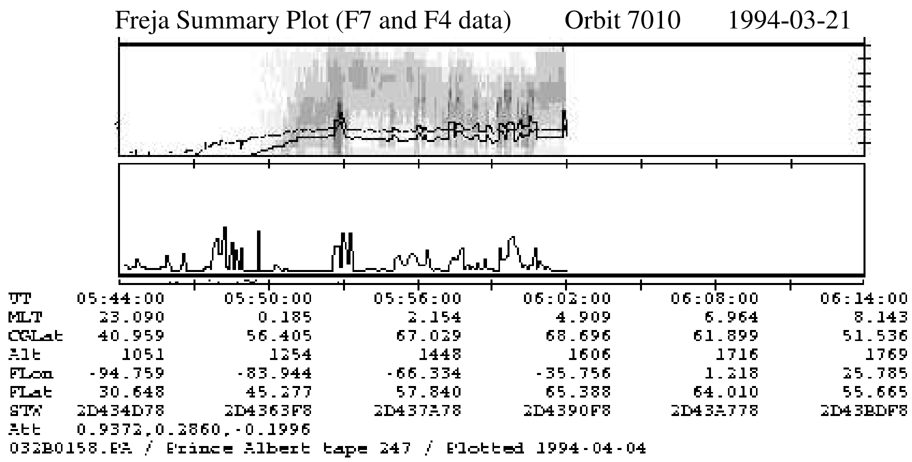


Figure 4.6: Data from the electron spectrometer F7 and the wave instrument F4 on Freja. Upper panel: Downgoing auroral electrons. Vertical scale is logarithmic in energy from 10 eV to 30 keV. Dark shading means high intesity, light shading is low intensity. Lower panel: wave power around the plasma frequency. When electrons around 100 eV appear, for instance at 05:53, the wave power increases.

(4.48), has complex solutions, they come in conjugate pairs. Thus, if $\omega_r + i\gamma$ is one solution, then $\omega_r - i\gamma$ is another. Hence, if the dispersion relation (4.48) has complex roots, then we have growing waves. We say that the plasma is *unstable*: any small perturbation of the right characteristic will grow exponentially until effects not accounted for in our equations (the terms we neglected in our linearization procedure, for instance) inhibits further wave growth. It is interesting to note that while the shorter wavelength ($\kappa = 0.15$) was stable, longer waves ($\kappa = 0.12$) were unstable in this case.

To say something about the details of the unstable plasma waves, for instance, to find the growth rate γ, we have to do an extended analysis of the dispersion relation. For a realistic study, it would also be necessary to take thermal and kinetic effects into account, to find, for example, how Landau damping may stabilize the instability. This is outside the scope of this presentation. However, the derivation above shows some of the quantitative aspects of stability analysis, and indicates the fact that plasmas with streaming electrons often are unstable. This is illustrated in Figure 4.6, which shows the intensity of precipitating auroral electrons and Langmuir waves as measured by the Freja satellite.

Problems for Chapter 4

1. *Wave speeds.* Derive expressions for the phase and group velocity of ion acoustic waves and Langmuir waves.

2. *Collisional damping.* Study how collissions affect Langmuir waves by adding a term $-mn_0\nu\mathbf{v}_\mathrm{e}$ to the right hand side of the equation of motion (4.1) in the blue compendium. Here, ν represents an effective collission frequency. Derive a dispersion relation including the effects of this term. Also derive an explicit expression for Im(ω), and check that the sign of Im(ω) is such as to give damping of the wave.

3. *Langmuir waves.* Show that the time averages of the density of kinetic energy of the electrons $\frac{1}{2}n_0 m_\mathrm{e} v_\mathrm{e}^2$ and the energy density of the electric wave field $\frac{1}{2}\epsilon_0 E^2$ are equal in a Langmuir wave.

4. *Electron-positron plasma.* Determine the dispersion relation for Langmuir waves in an electron-positron plasma.

5. *Ion acoustic wave fields.* An ion acoustic wave of 10 m wavelength is propagating in a plasma with density $n = 10^{10}$ m^{-3}, electron temperature $KT_\mathrm{e} = 1$ eV and $KT_\mathrm{i} = 0.1$ eV. Calculate the amplitude

of the wave density fluctuation field if the wave electric field amplitude is 10 mV/m.

6. *Electrostatic ion cyclotron waves.* Consider electrostatic ($\mathbf{E} = -\nabla\Phi$) waves with $\mathbf{k} \perp \mathbf{B}_0$ in a plasma with electron density n_0, neglegible ion pressure and magnetic field $\mathbf{B}_0$. For the electrons, $n_\mathrm{e} = n_0 \exp(\frac{e\Phi}{KT_\mathrm{e}})$, and the frequency is sufficiently low for them to be able to keep $n_\mathrm{e} = n_\mathrm{i}$. Such waves fulfill the dispersion relation

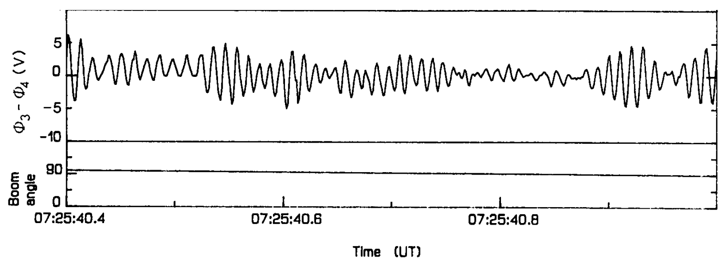
$$\omega^2 = \omega_\mathrm{ci}^2 + \frac{KT_\mathrm{e}}{m_\mathrm{i}} k^2$$

where ω_ci is the ion cyclotron angular frequency. These waves are called electrostatic ion cyclotron waves.

(a) Derive the dispersion relation above! A way of doing this is to write down the equation of motion for the ions (they are only affected by the Lorentz force) and their equation of continuity, linearise these and the electron equation above, and look for plane wave solutions $\exp \mathrm{i}(kx - \omega t)$, where the coordinates has been chosen so as to have$\mathbf{B}_0 = B_0\hat{\mathbf{z}}$ and $\mathbf{k} = k\hat{\mathbf{x}}$.

(b) The figure below shows measurements of electrostatic ion cyclotron waves from the Viking satellite. Estimate the frequency from the figure, and derive the wavelength, phase velocity and group velocity for the waves if $B_0 = 5240$ nT, $KT_\mathrm{e} = 1$ eV and the ions are protons. Also calculate the energy density in the wave electric field. To find the electric field strength from the measured voltage, divide by the antenna length of 80 m.

Viking V4L wave data Orbit 272 1986-04-12



Chapter 5

Low-frequency waves in magnetized plasmas

5.1 Anisotropic plasma

"Magnetized plasma" is a designation for a plasma in which we have a magnetic field. Most of the plasmas in space which are our interest in this course are magnetized: the solar wind, the magnetospheres, ionospheres on magnetized planets. To find plasmas which reasonably can be described as non-magnetized, we have to go to the ionosphere of Venus, or to the inner coma of comets.

The theory for waves in a magnetized plasma is more complicated than the theory for unmagnetized plasmas we have seen the elements of in the preceeding chapters. The basic reason for this is to be found in the equations of motion for ions and electrons,

$$m_{\mathrm{i,e}}\frac{\mathrm{d}\mathbf{v}_{\mathrm{i,e}}(t,\mathbf{r})}{\mathrm{d}t} = \pm e[\mathbf{E}(t,\mathbf{r}) + \mathbf{v}_{\mathrm{i,e}}(t,\mathbf{r}) \times \mathbf{B}(t,\mathbf{r})]. \tag{5.1}$$

In the unmagnetized case, $\mathbf{B}$ as well as $\mathbf{v}_{\mathrm{i,e}}$ are zero in the unperturbed plasma, and only existing as wave fields. In the linearization procedure, the $\mathbf{v} \times \mathbf{B}$ terms therefore disappear altogether. When there is a background magnetic field $\mathbf{B}_0$, we have to rewrite our linearization ansatz as

$$\mathbf{B}(t,\mathbf{r}) = \mathbf{B}_0 + \mathbf{B}_1(t,\mathbf{r}) \tag{5.2}$$

where we assume that the wave field

$$\mathbf{B}_1 \ll \mathbf{B}_0. \tag{5.3}$$

The linearized equations of motion then look like

$$m_{\mathrm{i,e}}\frac{\partial\mathbf{v}_{1i,e}(t,\mathbf{r})}{\partial t} = \pm e[\mathbf{E}_1(t,\mathbf{r}) + \mathbf{v}_{1i,e}(t,\mathbf{r}) \times \mathbf{B}_0] \tag{5.4}$$

for a cold plasma, which we may compare to equations (3.19) and (3.20). The new $\mathbf{v} \times \mathbf{B}$ term introduces a complication, as it will relate $\mathrm{d}v_x/\mathrm{d}t$ to v_y and v_z, etc. The component of the equation of motion which describes the dynamics along the magnetic field is unchanged by this complication, as the $\mathbf{v}_1 \times \mathbf{B}_0$ term does not contribute in this direction. Obviously, the dynamics of particle motion are different in different directions, which means that the plasma is *anisotropic*, in contrast to the isotropic plasmas we have previously considered. It is reasonable to expect that waves also will have different properties for different directions of propagation. This is a not unfamiliar situation: in a crystal, the lattice defines prefered directions, and wave propagation indeed depends on propagation angle to the lattice axes.

We may note that for propagation along the magnetic field, i.e. $\mathbf{k} \parallel \mathbf{B}_0$, the wave modes we have derived previously – radio waves, Langmuir waves, ion acoustic waves – all do exist and have the same dispersion relations as we have derived. In other directions, the properties of the waves may be drastically different. There also are a number of wave modes in a magnetized plasma which have no countepart in the unmagnetized case. Some of these are among the most important wave modes identified in nature, due to their possibility of transporting energy almost unattenuated over long distances. The magnetohydrodynamic approximation, introduced in section 5.2, is sufficient for the derivation of the main properties of these wave modes, which follows in section 5.3.

5.2 Magnetohydrodynamics (MHD)

At the simplest level, a plasma in a magnetic field behaves like any conducting fluid (liquid mercury, for instance) does in the presence of a magnetic field. This is the picture of *magnethydrodynamics (MHD)*: the plasma as a conductive fluid. In such a model, there is no net charge anywhere, so the interaction of the matter with the electromagnetic fields is strictly through the current. Per unit volume, the force on the plasma from the current is

$$n_\mathrm{i} e \mathbf{v}_\mathrm{i} \times \mathbf{B} - n_\mathrm{e} e \mathbf{v}_\mathrm{e} \times \mathbf{B} = \mathbf{j} \times \mathbf{B}, \tag{5.5}$$

so the equation of motion (“Newton’s second law per unit volume) must be

$$\rho_m(t, \mathbf{r}) \frac{\mathrm{d}\mathbf{v}(t, \mathbf{r})}{\mathrm{d}t} = -\nabla p(t, \mathbf{r}) + \mathbf{j}(t, \mathbf{r}) \times \mathbf{B}(t, \mathbf{r}). \tag{5.6}$$

No reference to the electric field is found in the equations, as there is no net charge for it to act on. This restricts the validity of the MHD model to large spatial and long temporal scales, as we know that ion and electron motion may cause charge imbalances on shorter scales[1]. What does “large” and “long” in the sentence above really mean? What quantities are we to compare to? For an unmagnetized plasma, there is one intrinsic length scale, λ_D, and one time scale, $\tau_\mathrm{p} = 2\pi/\omega_\mathrm{p}$. In the magnetized plasma, there are additional parameters to compare to: the gyroperiods of the different particle species s present in the plasma, $\tau_\mathrm{cs} = 2\pi/\omega_\mathrm{cs}$, and the gyroradii of particles of these species with typical thermal velocities, $r_\mathrm{gs} = \sqrt{KT_\mathrm{s}/m_\mathrm{s}}/\omega_\mathrm{cs}$. Thus “large spatial scale” means $L >> r_\mathrm{gs},\ \lambda_\mathrm{D}$, while “long time” means $\tau >> \tau_\mathrm{cs},\ \tau_\mathrm{p}$.

A consequence of that only long temporal and spatial scales are considered is that the perpendicular speed of the plasma will be the $\mathbf{E} \times \mathbf{B}$ drift speed. If the conductivity along the magnetic field lines is almost perfect, no parallel electric fields can develop, so we can write this as[2]

$$\mathbf{E} + \mathbf{v} \times \mathbf{B} = 0. \tag{5.7}$$

In MHD theory, Gauss’ law (1.22) clearly cannot be used to determine the electric field, as perfect neutrality is assumed. Instead, equation (5.7) acts as source equation for the electric field.

For sufficiently slow phenomena, the time derivative in Ampére-Maxwell’s law (1.25) is neglegible compared to the current term, so that

$$\nabla \times \mathbf{B}(t, \mathbf{r}) = \mu_0 \mathbf{j}(t, \mathbf{r}). \tag{5.8}$$

The other two Maxwell equations, (1.23) and (1.24), remain unchanged.

5.3 Hydromagnetic waves

We will focus our interest on cold plasma magnetohydrodynamics, where $p = T = 0$. The waves we will find in this model are known as magnetohydrodynamic waves or hydromagnetic waves. Collecting the equations from the preceding section, we describe the plasma dynamics by the following system:

$$\rho_m(t, \mathbf{r}) \frac{\mathrm{d}\mathbf{v}(t, \mathbf{r})}{\mathrm{d}t} = \mathbf{j}(\mathbf{r}, t) \times \mathbf{B}(t, \mathbf{r}) \tag{5.9}$$

$$\nabla \times \mathbf{E}(t, \mathbf{r}) = -\frac{\partial \mathbf{B}}{\partial t} \tag{5.10}$$

[1] An example are the plasma oscillations and electrostatic waves we considered in the preceeding chapters, for whoose physics charge imbalance is vital.

[2] One may argue for equation (5.7) also from the equations of motion for ions and electrons (3.7, 3.8). When spatial and temporal scales are long, all derivatives with respect to time and space go to zero, leaving only $\mathbf{E} + \mathbf{v} \times \mathbf{B} = 0$. See also problem 1.

$$\nabla \times \mathbf{B}(t, \mathbf{r}) = \mu_0 \mathbf{j}(t, \mathbf{r}) \tag{5.11}$$

$$\mathbf{E}(t, \mathbf{r}) + \mathbf{v}(t, \mathbf{r}) \times \mathbf{B}(t, \mathbf{r}) = 0. \tag{5.12}$$

These equations are non-linear, so we apply our usual method of linearization for small perturbations (section 2.4):

1. **Ansatz.**

$$\begin{array}{rclcl} \rho_m(t, \mathbf{r}) & = & \rho_0 & + & \rho_1(t, \mathbf{r}) \\ \mathbf{B}(t, \mathbf{r}) & = & \mathbf{B}_0 & + & \mathbf{B}_1(t, \mathbf{r}) \\ \mathbf{E}(t, \mathbf{r}) & = & & & \mathbf{E}_1(t, \mathbf{r}) \\ \mathbf{v}(t, \mathbf{r}) & = & & & \mathbf{v}_1(t, \mathbf{r}) \\ \mathbf{j}(t, \mathbf{r}) & = & & & \mathbf{j}_1(t, \mathbf{r}) \end{array} \tag{5.13}$$

   where

$$\rho_1(t, \mathbf{r}) \ll \rho_0 \tag{5.14}$$

$$\mathbf{B}_1(t, \mathbf{r}) \ll \mathbf{B}_0. \tag{5.15}$$

2. **Insert the ansatz into the field equations** (5.9) – (5.12).

3. **Derivatives of background values disappears.**

4. **Neglect quadratic and higher order terms.** This procedure provides us with a set of linearized MHD equations:

$$\rho_{m0} \frac{\partial \mathbf{v}_1(t, \mathbf{r})}{\partial t} = \mathbf{j}_1(\mathbf{r}, t) \times \mathbf{B}_0 \tag{5.16}$$

$$\nabla \times \mathbf{E}_1(t, \mathbf{r}) = -\frac{\partial \mathbf{B}_1(t, \mathbf{r})}{\partial t} \tag{5.17}$$

$$\nabla \times \mathbf{B}_1(t, \mathbf{r}) = \mu_0 \mathbf{j}_1(t, \mathbf{r}) \tag{5.18}$$

$$\mathbf{E}_1(t, \mathbf{r}) + \mathbf{v}_1(t, \mathbf{r}) \times \mathbf{B}_0 = 0. \tag{5.19}$$

As this is a linear system, we may confine our interest to plane waves. The linearized equations transform to

$$-i\omega\rho_0 \mathbf{v}_1 = \mathbf{j}_1 \times \mathbf{B}_0 \tag{5.20}$$

$$i\mathbf{k} \times \mathbf{E}_1 = i\omega \mathbf{B}_1 \tag{5.21}$$

$$i\mathbf{k} \times \mathbf{B}_1 = \mu_0 \mathbf{j}_1. \tag{5.22}$$

$$\mathbf{E}_1 + \mathbf{v}_1 \times \mathbf{B}_0 = 0. \tag{5.23}$$

We choose the z-axis to lie along $\mathbf{B}_0$:

$$\mathbf{B}_0 = B_0 \hat{\mathbf{z}}. \tag{5.24}$$

From now on, we will drop the subscript "1" in the wave fields for $\mathbf{E}$, $\mathbf{v}$ and $\mathbf{j}$, for which there are no background fields that can cause confusion. In component form, the equation of motion then writes

$$-i\omega\rho_{m0} v_x = j_y B_0 \tag{5.25}$$

$$-i\omega\rho_{m0} v_y = -j_x B_0 \tag{5.26}$$

$$v_z = 0, \tag{5.27}$$

while (5.23) appear as

$$E_x = -v_y B_0 \tag{5.28}$$

$$E_y = v_x B_0 \tag{5.29}$$

$$E_z = 0. \tag{5.30}$$

Multiplying (5.25) and (5.26) by $i\omega\mu_0/B_0$ and using the velocity components from (5.28) and (5.29) yields

$$i\omega\mu_0 j_x = \omega^2 E_x/v_\mathrm{A}^2 \tag{5.31}$$

$$i\omega\mu_0 j_y = \omega^2 E_y/v_\mathrm{A}^2 \tag{5.32}$$

where we have introduced *the Alfvén speed*

$$\boxed{c_\mathrm{A} = \frac{B_0}{\sqrt{\mu_0\rho_{m0}}}} \tag{5.33}$$

The physical interpretation of c_A will be unveiled below. The components of Ampère's law (5.22) are

$$\mu_0 j_x = i(k_y B_{1z} - k_z B_{1y}) \tag{5.34}$$

$$\mu_0 j_y = i(k_z B_{1x} - k_x B_{1z}) \tag{5.35}$$

$$\mu_0 j_z = i(k_x B_{1y} - k_y B_{1x}), \tag{5.36}$$

while Faraday-Henry's law (5.21) looks like

$$\omega B_{1x} = -k_z E_y \tag{5.37}$$

$$\omega B_{1y} = k_z E_x \tag{5.38}$$

$$\omega B_{1z} = k_x E_y - k_y E_x \tag{5.39}$$

where we used $E_z = 0$, which follows from (5.30). We now eliminate $\mathbf{B}_1$ from (5.34) - (5.36) by use of (5.37) - (5.39), and insert the resulting expressions for $\mathbf{j}$ in terms of $\mathbf{E}$ in (5.31) and (5.32), whereby we find that

$$(k_y^2 + k_z^2 - \omega^2/c_\mathrm{A}^2)E_x = k_x k_y E_y \tag{5.40}$$

and

$$(k_x^2 + k_z^2 - \omega^2/c_\mathrm{A}^2)E_y = k_x k_y E_x. \tag{5.41}$$

Multiplying these equations, we get

$$\begin{aligned}
0 &= (k_y^2 + k_z^2 - \omega^2/c_\mathrm{A}^2)(k_x^2 + k_z^2 - \omega^2/c_\mathrm{A}^2) - k_x^2 k_y^2 = \\
&= (k_z^2 - \omega^2/c_\mathrm{A}^2)(k_x^2 + k_z^2 - \omega^2/c_\mathrm{A}^2) + \\
&\quad + k_y^2 k_x^2 + k_y^2(k_z^2 - \omega^2/c_\mathrm{A}^2) - k_x^2 k_y^2 = \\
&= (k_z^2 - \omega^2/c_\mathrm{A}^2)(k_x^2 + k_y^2 + k_z^2 - \omega^2/c_\mathrm{A}^2)
\end{aligned} \tag{5.42}$$

which obviously has two solutions. We thus get two dispersion relations representing independent wave modes in the plasma. One of them is known as the *Alfvén mode*,

$$\boxed{\omega^2 = k_z^2 c_\mathrm{A}^2 = k_\parallel^2 c_\mathrm{A}^2} \tag{5.43}$$

while the other mode here will be known as the *compressional mode*[3],

$$\boxed{\omega^2 = (k_x^2 + k_y^2 + k_z^2)c_\mathrm{A}^2 = k^2 c_\mathrm{A}^2} \tag{5.44}$$

[3]These wave modes both have many names, and there is no universal nomenclature. The concept *Alfvén waves* often refers to both the wavemodes, but equally or more often the name Alfén wave is reserved for the wave we call the Alfén mode.

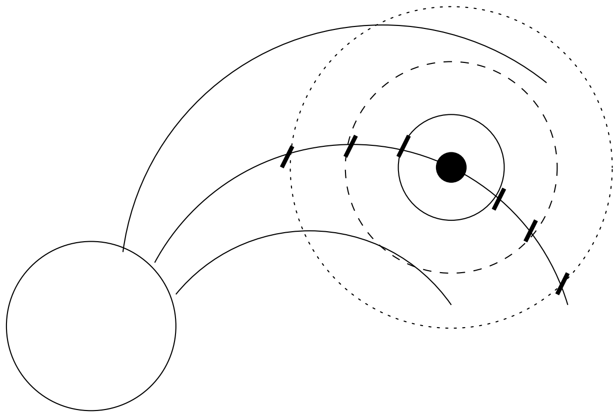


Figure 5.1: The compressional mode propagates isotropically in all directions, as indicated by the concentric circles (which neglects the fact that B and thus c_A varies in a magnetosphere). The wave amplitude therefore decreases as the energy is spread over circles with increasing radius (spheres in three dimension). In contrast, the Alfvén wave, whose wavefronts are denoted by short solid bars in the sketch above, propagates only along the magnetic field lines, with no geometric attenuation with distance. As a result, a source emitting both kinds of MHD waves far out in the magnetosphere will cause a weak compressional wave field far away, while the wave field of the Alfvén mode will have almost the same magnitude even at long distances – provided you sit on the same field line as the wave source; otherwise the Alfvén mode wave field will not be detectable at all.

By the neglect of the displacement current in the Ampère-Maxwell's law, our results are non-relativistic, valid only for $c_\mathrm{A} \ll c$, i.e. for sufficiently ambient weak magnetic field and high plasma density (see the definition (5.33) of the Alfvén speed).

Let us first consider the compressional mode. This mode propagates isotropically in all directions, as only the magnitude k of the wave vector $\mathbf{k}$ is included in the dispersion relation (5.44), not the direction. We can also see that the wave is non-dispersive, as the dispersion relation is a linear relation between k and ω. The propagation characteristics of the compressional wave therefore is reminiscent of a light wave in vacuum or a sound wave in air, but the speed of propagation is c_A rather than c or c_s. As is the case for all non-dispersive waves, the group velocity $\mathbf{v}_\mathrm{g}$ is parallel the wave vector $\mathbf{k}$, and its magnitude is c_A in all directions. A perturbation from a point source thus propagates uniformly in all directions, likes the rings on the water surface when a stone is dropped in a pond. As energy must be conserved, the wave energy transported through any sphere of radius r centered on the point source must be constant, implying that the wave energy density must decrease as $1/r^2$ and the wave fields ($\mathbf{E}_1$, $\mathbf{B}_1$ etc.) like $1/r$.

The Alfvén mode is also characterized by c_A, but otherwise behaves quite differently. From the dispersion relation (5.43),

$$v_\phi = \frac{\omega}{k} = \frac{k_\parallel}{k} c_\mathrm{A} = c_\mathrm{A} \cos\theta, \tag{5.45}$$

where θ is the angle between $\mathbf{B}_0$ and $\mathbf{k}$, and

$$\mathbf{v}_\mathrm{g} = \frac{\partial \omega}{\partial \mathbf{k}} = \pm c_\mathrm{A} \hat{\mathbf{z}} \tag{5.46}$$

where z as usual is chosen along the ambient magnetic field $\mathbf{B}_0$. The result for the group velocity $\mathbf{v}_\mathrm{g}$ is of particular interest: the Alfvén wave transports energy exclusively along the magnetic field! Note that this is true irrespective of size and direction of $\mathbf{k}$ – wherever the wave vector is pointing, the wave will propagate along $\mathbf{B}_0$ only. This means that the Alfvén wave can transport energy over very long distances, without the $1/r^2$ dependence that normally characterizes radiation from a point source (compare Figure 5.1). This makes the Alfvén mode very important in many cosmic contexts.

What do the actual fields from these wave modes look like? Let us choose coordinates so that $\mathbf{B}_0 = B_0 \hat{\mathbf{z}}$ and $\mathbf{k} = k_x \hat{\mathbf{x}} + k_z \hat{\mathbf{z}}$: as there only are two prefered directions in the problem, defined by $\mathbf{B}_0$ and $\mathbf{k}$ this

Alfvén mode	Compressional mode
$E_x \neq 0$ $E_y = 0$ $E_z = 0$	$E_x = 0$ $E_y \neq 0$ $E_z = 0$
$B_{1x} = 0$ $B_{1y} = E_x/c_{\mathrm{A}}$ $B_{1z} = 0$	$B_{1x} = -k_z E_y/\omega$ $B_{1y} = 0$ $B_{1z} = k_x E_y/\omega$
$v_x = 0$ $v_y = -E_x/B_0$ $v_z = 0$	$v_x = E_y/B_0$ $v_y = 0$ $v_z = 0$
$j_x = -\mathrm{i}\omega E_x/\mu_0 c_{\mathrm{A}}^2$ $j_y = 0$ $j_z = \mathrm{i}k_x E_x/\mu_0 c_{\mathrm{A}}$	$j_x = 0$ $j_y = -\mathrm{i}\omega E_y/\mu_0 c_{\mathrm{A}}^2$ $j_z = 0$

Table 5.1: The wave fields of hydomagnetic waves. The coordinates are chosen so as to have $\mathbf{B}_0 = B_0\hat{\mathbf{z}}$ and $\mathbf{k} = k_x\hat{\mathbf{x}} + k_z\hat{\mathbf{z}}$.

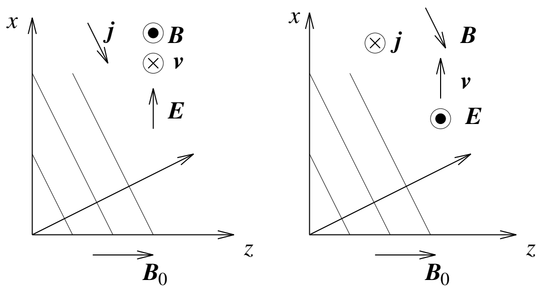


Figure 5.2: Wave fields for the Alfvén wave (left) and the compressional wave (right). Coordinates as described in the text and in the caption to Table 5.3.

implies no loss of generality. We thus have $k_y = 0$, $k_\parallel = k_z$ and $k^2 = k_x^2 + k_z^2$. Using the dispersion relations for Alfvénic (5.43) and compressional (5.44) waves to eliminate ω^2 from equations (5.40) and (5.41), we find that for the compressional wave, $E_x = 0$, while $E_y = 0$ for the Alfvén mode. By going backwards through equations (5.39) to (5.28), we can find all other field components as well. The result is presented in Table 5.3 and in Figure 5.2. Equation (5.12) implies that the magnetic field lines may be considered to be "frozen" in the plasma as has been discussed previously in the course. We may envisage the Alfvén wave mode as in Figure 5.3: it propagates as a ripple on the field lines, very much like waves propagating along a skipping rope or guitar string. The compressional wave behaves more like a sound wave, with compressions and decompressions of the plasma, but unlike the sound wave, the compressional MHD wave does not depend on collisions. For compressional waves, one may envisage a sort of wave motion in a drapery made up of hanging ropes, connected to each other by horizontal springs.

Alfvén waves are very important in many cosmic contexts. As an example, we may consider the terrestrial magnetosphere-ionosphere system. The Alfvén wave is responsible for transport of information from the magnetopause to the ionosphere, because they are the mode transmitting changes in the field-aligned currents flowing through the magnetosphere. This is due to the fact that Alfvén waves have a current component along $\mathbf{B}_0 = B_0\hat{\mathbf{z}}$ (compare Table 5.3). When Fourier transforming a time-dependent field aligned current, the components at non-zero frequency will be Alfén waves[4]. Because of their relation to field-aligned currents, Alfvén waves are important in auroral processes. Looking at Table 5.3, we find that the hydromagnetic waves do not carry any electric field component along $\mathbf{B}_0$, so at first glance they should not be able to accelerate particles in parallel to the magnetic field. However, in auroral regions the perpendicular scale sizes (of for example auroral arcs) often are of the order of the ion gyroradius, where the

[4]Assuming the frequency to be well below ω_{ci}.

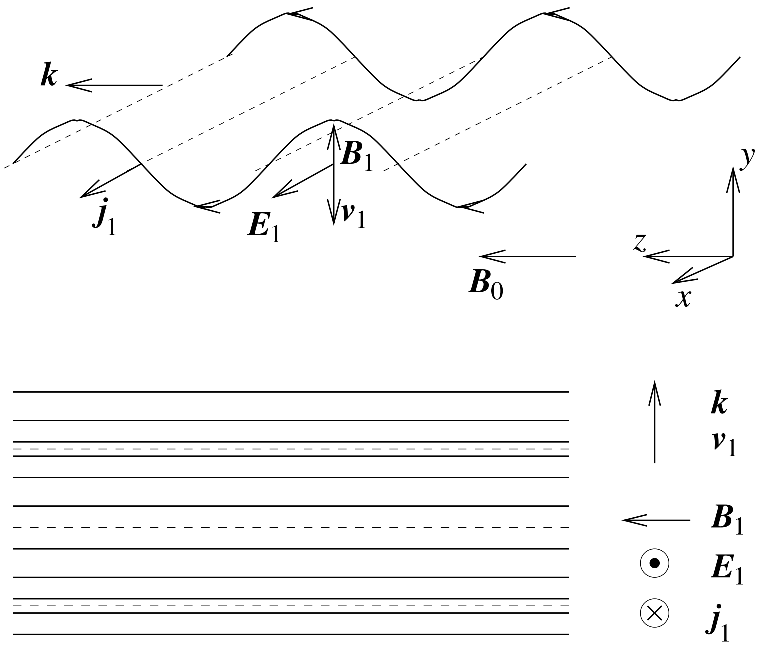


Figure 5.3: Schematic pictures of magnetic field lines for an Alfvén wave propagating parallel to $\mathbf{B}_0$ (upper drawing), and a compressional wave propagating perpendicular to $\mathbf{B}_0$ (lower drawing). Dashed lines are wave fronts.

simple MHD theory we have used here is no longer applicable. More elaborate descriptions in fact show that Alfvén waves in fact carry a parallel electric field if $k_\perp \neq 0$. This field is normally much smaller than the perpendicular electric field, but nevertheless very important.

5.4 Magnetospheric resonances

In the preceding section, we discussed how Alfvén waves can be pictured as ripples on a string, the string in this analogy being a magnetic field line. In the magnetosphere of the Earth or of some other magnetized planet, this "string" has two ends, one in the northern and one in the southern ionosphere of the planet. The "string" thus has finite length, which means that we may expect standing waves, with wavelengths such that an integer number of half wavelengths equal the length of the "string", or field line (Figure 5.4).

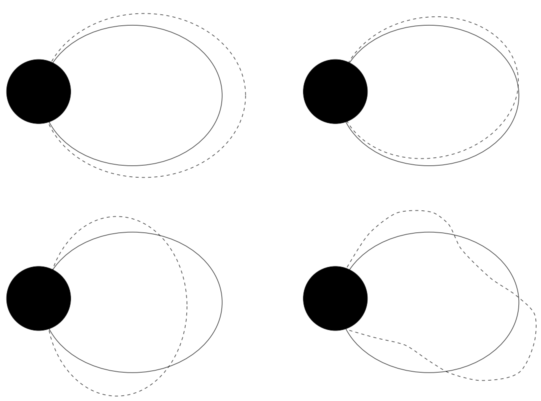


Figure 5.4: Schematic of some of the first standing wave modes on a magnetic field line.

It turns out that this discussion by analogy is basically correct. Figure 5.5 show magnetic field measurements from the geostationary satellite ATS-6. One can descern a fundamental frequency and up to five harmonics in these data. The fundamental frequency changes in time as the satellite follow the rotation of the Earth and the satellite field line thus moves through magnetospheric regions with varying ion composition, plasma density and, to some extent, magnetic field intensity, and therefore varying Alfvén speed. Figure 5.5 also shows a model for the excitation of these resonances. A perturbation in the solar wind hits the magnetopause (the boundary of the magnetosphere) and propagates into the magnetosphere as a compressional wave (here called "fast mode wave"). When the compressional wave passes the field line whose eigenfrequency, or some multiple thereof, is equal to the wave frequency, this field line start oscillating as a standing wave.

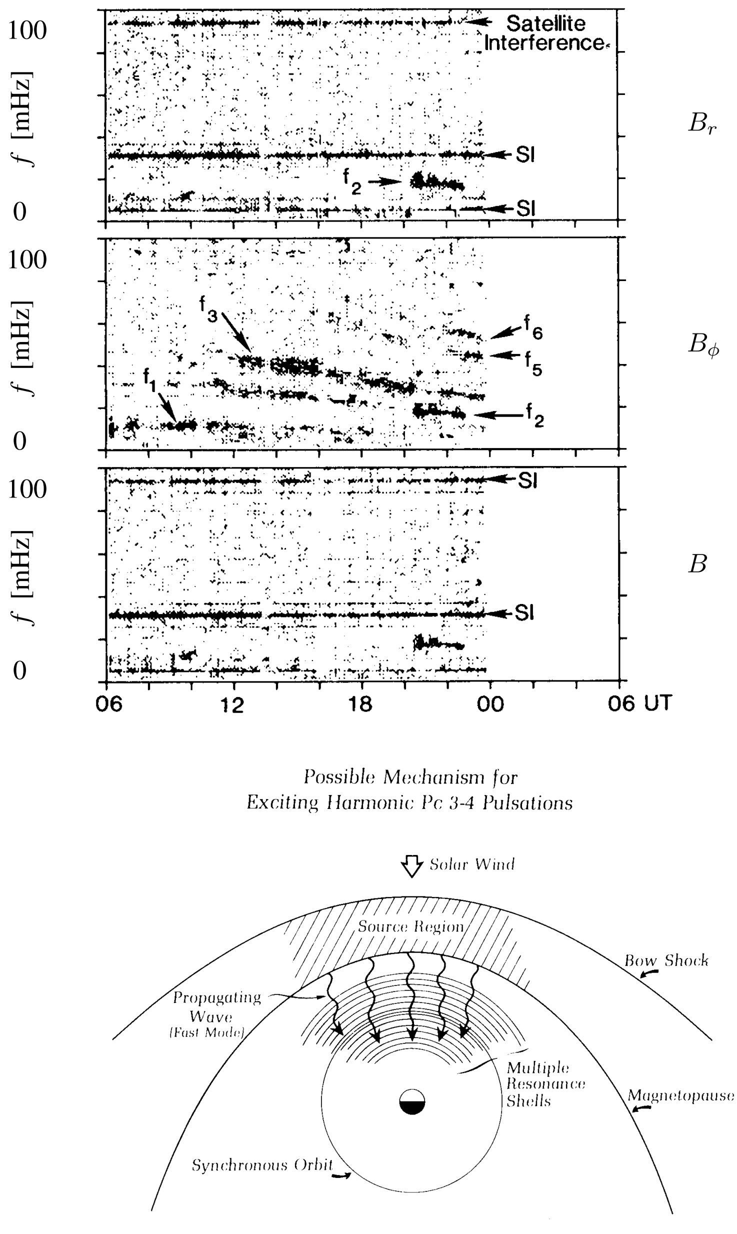


Figure 5.5: Top: Spectra of field line resonances (dark corresponds to high intensity) observed by the magnetospheric satellite ATS-6. Panels show spectra of radial, azimuthal and total magnetic field fluctuation, respectively. Bottom: a model for the excitation of these resonances. UT is universal time. From Takahashi and McPherron.

Problems for Chapter 5

1. *"Frozen-in condition".* Use the fluid equations of motion for a charged particle species s,

$$nm\frac{\mathrm{d}\mathbf{v}_\mathrm{s}}{\mathrm{d}t} = n_\mathrm{s}q_\mathrm{s}(\mathbf{E} + \mathbf{v}_\mathrm{s} \times \mathbf{B}) - \nabla(nKT),$$

to show that

$$\left|\frac{\partial}{\partial t}\right| \ll \omega_\mathrm{c}$$

and

$$|\nabla| \ll 1/r_\mathrm{g} \text{ and } \frac{qvB}{KT}$$

are sufficient conditions for the "frozen-in condition" $\mathbf{E} + \mathbf{v} \times \mathbf{B} = 0$ (equation (5.7)) to be valid.

2. *Energy densities.* Show that the time averages of the wave kinetic energy density $w_\mathrm{K} = \frac{1}{2}\rho_0 v_{1\mathrm{e}}^2$, the wave magnetic field energy density $w_\mathrm{B} = \frac{1}{2\mu_0}B_1^2$ and the wave electric field energy density $w_\mathrm{E} = \frac{1}{2}\epsilon_0 E_1^2$ in an Alfvén wave satisfiy $w_\mathrm{K} = w_\mathrm{B} \gg w_\mathrm{E}$.

3. *Phase relations.* What are the phase relations between the wave fields $\mathbf{E}_1$, $\mathbf{B}_1$, $\mathbf{v}_1$ and $\mathbf{j}_1$ in hydromagnetic waves, i.e., which wave fields are zero simultaneously?

4. *Plasmapause waves (Pc 1).* The figure below shows observations of hydromagnetic waves by the Viking satellite (from Erlandson et al.). Consider the time around 045030 UT, at which time the onboard Langmuir probe indicate an electron density of around 100 cm^{-3}.

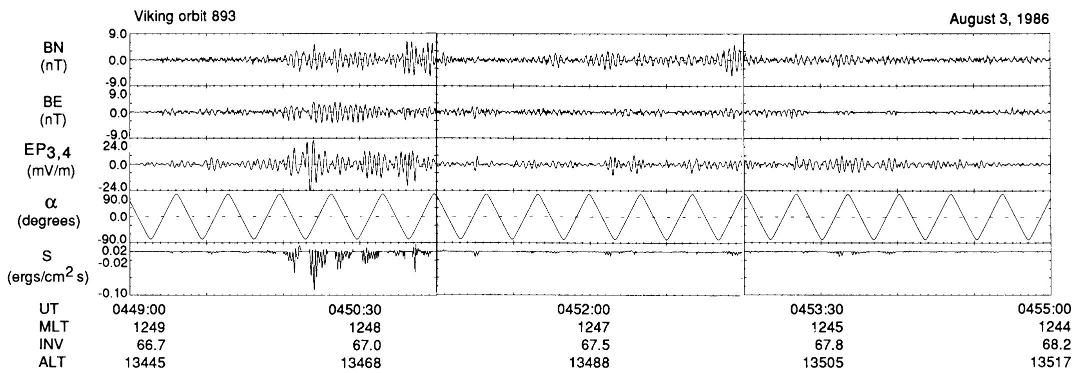
   (a) The magnetometer shows that $B_0 \approx 1600$ nT. Compare this to what would be the case if the Earth had a perfect dipole field with intensity 30 μT on the ground at the equator. The latitude is INV and the altitude is ALT.

   (b) Estimate the Alfvén speed using observed values of $\mathbf{B}_0$ and n_0, using some reasonable assumption for the ion composition.

   (c) Estimate the Alfvén speed using the measured fluctuations in the electric and magnetic fields.

   (d) Estimate the wavelength using the information you have obtained above.

   (e) The energy flux in the wave can be calculated in two ways: by computing the Poynting vector $\mathbf{S} = \mathbf{E}_1 \times \mathbf{B}_1/\mu_0$, which is plotted in the figure, or by multiplying the energy density in the wave fields by the group velocity. Compare the results obtained by these methods.



5. *Shocks.* The bold spaceman Spiff, famous interplanetary explorer, enters the Jovian ionosphere with his spacecraft. His mission is to find out the electron density and temperature of the ionospheric plasma. However, most of his instruments for studying the plasma has been wrecked in a recent fight with hostile aliens, leaving only an electric and a magnetic probe operational. With great skill, he steers the spacecraft parallel to the magnetic field, which has strength 1 nT according to his magnetometer. Spiff knows that an ion acoustic shock wave will form in front of the spacecraft if he travels

faster than the ion acoustic speed, and that an Alfvén shock wave will form if he also exceeds the Alfvén speed. On his instrument for electric field measurements he can see that an electric shock is present, but the magnetometer shows no signs of any magnetic shock wave. What can he say about n_e and T_e from this observation?

References

Works mentioned in the text

- Brillouin, L., and A. Sommerfeld: Wave Propagation and Group Velocity. Academic Press 1960.
- Chen, F. F.: Introduction to Plasma Physics and Controlled Fusion, Volume 1, Plasma Physics. Second edition, Plenum Press 1984.
- Erlandson, RĖ., L. J. Zanetti, T. A. Potemra, L. P. Block and G. Holmgren: Viking Magnetic and Electric Field Observations of Pc 1 Waves at High Latitudes. Journal of Geophysical Research vol. 95, p. 5941, 1990.
- Jackson, J. D.: Classical Electrodynamics. Second edition, Wiley 1975.
- Lighthill, J.: Waves in Fluids. Cambridge University Press 1978.
- Longair, M.: High energy astrophysics, volume 2: Stars, the galaxy and the interstellar medium. Second edition, Cambridge University Press 1994.
- Panofsky, W. K. H., and M. Phillips: Classical electricity and magnetism. Second edition, Addison-Wesley 1962.
- Simmons, G. F.: Differential equations with applications and historical notes. Second edition, McGraw-Hill 1991.
- Swanson, D. G.: Plasma Waves. Academic Press 1989.
- Takahashi, K., and R. L. McPherron: Standing Hydromagnetic Oscillations in the Magnetosphere. Planetary and Space Science, vol. 32, p. 1343, 1984.
- Wangsness, R. K.: Electromagnetic Fields. Second edition, Wiley 1986.

Some other useful books

- Bittencourt, J. A.: Fundamentals of Plasma Physics. Pergamon 1986. (Nice book, at a somewhat higher level than Chen.)
- Kivelson, M. G., and C. T. Rusell (editors): Introduction to Space Physics. Cambridge University Press 1995. (A very useful introduction to space physics.)