Chapter 15

Basic concepts of small-amplitude waves in anisotropic dispersive media

Systems of linear differential equations can often be studied conveniently using Fourier analysis. If any one quantity oscillates sinusoidally at a particular frequency, ω, then all the others must oscillate at the same frequency (or not at all), and the problem becomes one of finding the relative amplitudes and phases of the various oscillating quantities. The fluid plasma equations do *not* constitute a set of linear differential equations, so we cannot in general assume that nonlinear coupling between frequencies will be absent. However, if we consider only situations where the oscillations are small enough, then the equations can be 'linearized'. This means that the fluid equations are solved to zeroth order with no waves present. In the simplest case, considered here, that solution is the trivial one—a uniform isotropic plasma immersed in a steady (or even zero) magnetic field. Next we consider a first-order expansion of the equations in terms of small wave-like perturbations, neglecting second- and higher-order terms. This means that whenever we see two oscillating quantities multiplied together, since they are both small, we consider this to be a higher-order term and we neglect it. For any real situation, we then have to go back and verify that this neglect is justified: are the amplitudes we calculate in our real situation small enough that the nonlinear terms are actually negligible compared to the linear ones? For now, however, we will consider just the idealized small-amplitude limit.

15.1 EXPONENTIAL NOTATION

In the linear regime, all oscillating quantities can be represented with 'exponential notation'. For example, the density perturbation could be

$$n_1 = \bar{n}_1 \exp[\mathrm{i}(\mathbf{k} \cdot \mathbf{x} - \omega t + \delta_n)] \tag{15.1}$$

where the overbar on the $\bar{n}_1$ indicates that it is serving as a real wave amplitude, rather than an oscillating quantity (note that the overbar does *not* indicate a time average.) The quantity $\mathbf{k}$ is the vector wave-number, or 'wave-vector', and λ, the wavelength, is $2\pi/k$. The vector $\mathbf{k}$ can have components in all directions. In an anisotropic medium like a magnetized plasma, the direction as well as the magnitude of $\mathbf{k}$ plays a crucial role in the wave dynamics. Along directions in which the component of $\mathbf{k}$ is large, the wavelength is short, so quantities vary rapidly in space; along directions in which the component of $\mathbf{k}$ is small, the wavelength is long, and so quantities vary slowly in space. Of course, the fact that we have small-amplitude perturbations does not imply that this plane-wave spatial variation necessarily gives the best description of the oscillations. Indeed, planar geometry is too simple to treat a cylindrical or otherwise specially shaped real situation, if the size of the plasma is not much greater than a wavelength. Then only the $\exp(-\mathrm{i}\omega t + \mathrm{i}\delta_n)$ time dependence applies, and a different spatial dependence is appropriate.

For now, we will deal with idealized plane waves only. In the particularly simple case where the plane wave-fronts align with surfaces of constant x, we can write

$$n_1 = \bar{n}_1 \exp[\mathrm{i}(k_x x - \omega t + \delta_n)]. \tag{15.2}$$

For definiteness, we can take δ_n to be 0 (i.e. no phase shift, an assumption that does not sacrifice generality since we can choose to measure the phase shift of *everything else* relative to n_1). If we choose the standard convention that the measurable part of n_1 is its real part, we have

$$n_1 = \bar{n}_1 \cos[(k_x x - \omega t)]. \tag{15.3}$$

This represents a wave traveling with a phase velocity $v_\mathrm{p} \equiv \omega/k_x$.

In the case of a vector wave-number, we define a vector phase velocity

$$\mathbf{v}_\mathrm{p} \equiv \omega\mathbf{k}/k^2 = (\omega k_x/k^2)\hat{\mathbf{x}} + (\omega k_y/k^2)\hat{\mathbf{y}} + (\omega k_z/k^2)\hat{\mathbf{z}}.$$

An observer traveling at speed ω/k in the direction of propagation of the wave, $(\mathbf{k}/k)$, stays at a constant wave phase. We can see this by supposing that $\mathbf{x}$ varies as $\mathbf{v}_\mathrm{p}t$, in which case the argument of the exponential, $\mathrm{i}(\mathbf{k}\cdot\mathbf{x} - \omega t + \delta_n)$, is independent of time. In this Unit we will *always* consider $\mathrm{Re}(\omega)$ to be positive, since a negative $\mathrm{Re}(\omega)$ corresponds to a wave propagating in the opposite direction from $\mathbf{k}$; we will handle such a case with $\mathbf{k} \to -\mathbf{k}$. The quantity $\mathrm{Im}(\omega)$ represents damping ($\mathrm{Im}(\omega) < 0$) or growth ($\mathrm{Im}(\omega) > 0$) of the wave in time. Similarly, $\mathrm{Im}(\mathbf{k})$ represents growth or damping in space.

Other quantities such as flow velocities and electric and magnetic fields will have the same character of spatial and temporal variation, i.e. $\exp[\mathrm{i}(\mathbf{k}\cdot\mathbf{x} - \omega t)]$, but will have different phases and amplitudes. Indeed, each vector component of each quantity has its own phase and amplitude. For example, we can write the electric field as

$$
\begin{aligned}
\mathbf{E}_1 &= \bar{E}_{x1}\hat{\mathbf{x}}\cos[(\mathbf{k}\cdot\mathbf{x}-\omega t+\delta_{Ex})] + \bar{E}_{y1}\hat{\mathbf{y}}\cos[(\mathbf{k}\cdot\mathbf{x}-\omega t+\delta_{Ey})] \\
&\quad + \bar{E}_{z1}\hat{\mathbf{z}}\cos[(\mathbf{k}\cdot\mathbf{x}-\omega t+\delta_{Ez})] \\
&= \mathrm{Re}\{\bar{E}_{x1}\hat{\mathbf{x}}\exp[\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t+\delta_{Ex})] + \bar{E}_{y1}\hat{\mathbf{y}}\exp[\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t+\delta_{Ey})] \\
&\quad + \bar{E}_{z1}\hat{\mathbf{z}}\exp[\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t+\delta_{Ez})]\} \\
&= \mathrm{Re}\{\bar{E}_{x1}\exp(\mathrm{i}\delta_{Ex})\hat{\mathbf{x}}\exp[\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t)] + \bar{E}_{y1}\exp(\mathrm{i}\delta_{Ey})\hat{\mathbf{y}} \\
&\quad \times\exp[\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t)] + \bar{E}_{z1}\exp(\mathrm{i}\delta_{Ez})\hat{\mathbf{z}}\exp[\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t)]\} \\
&= \mathrm{Re}\{[\bar{E}_{x1}\exp(\mathrm{i}\delta_{Ex})\hat{\mathbf{x}} + \bar{E}_{y1}\exp(\mathrm{i}\delta_{Ey})\hat{\mathbf{y}} + \bar{E}_{z1}\exp(\mathrm{i}\delta_{Ez})\hat{\mathbf{z}}] \\
&\quad \times\exp[\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t)]\}
\end{aligned}
\tag{15.4}
$$

where δ_{Ex}, δ_{Ey} and δ_{Ez} are real phase delays between E_{x1}, E_{y1}, E_{z1} and n_1, and all the amplitude factors (the quantities with the overbars) are again taken to be real. This is a painfully non-compact form for $\mathbf{E}_1$. The same information can be written as

$$
\mathbf{E}_1 = \mathrm{Re}\{\underline{E}_1\exp[\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t)]\}
\tag{15.5}
$$

where the *underlined italic* $\underline{E}_1$ is now a complex vector (i.e. it has six scalars associated with it), but it is independent of time and space. To translate between these two notations recognize that, for example,

$$
\tan\delta_{Ex} = \mathrm{Im}(\underline{E}_1\cdot\hat{\mathbf{x}})/\mathrm{Re}(\underline{E}_1\cdot\hat{\mathbf{x}})
\tag{15.6}
$$

and

$$
\bar{E}_{x1} = |\underline{E}_1\cdot\hat{\mathbf{x}}| = [(\underline{E}_1\cdot\hat{\mathbf{x}})(\underline{E}_1\cdot\hat{\mathbf{x}})^*]^{1/2}
\tag{15.7}
$$

where the asterisk indicates a complex conjugate. In equations (15.6) and (15.7), the terms on the far left-hand side are the real phase delay and the real amplitude, while the other terms are built from the complex wave amplitudes.

As we proceed to use this notation, we will take even more advantage of its compactness. All of the first-order terms in our equations (and therefore one multiplier in every additive term in the first-order equations) will contain the same exponential factor. Therefore we can simply drop the exponential factor without difficulty, so long as we are always clear about which are the first-order multiplicative terms. (For example, we will often find terms like $\underline{E}_1 \times \mathbf{B}_0$, and it is important to remember which one is the perturbed quantity.) Finally, in the interest of further conciseness of notation, we will drop the underlined italics which indicates a complex wave amplitude: *all the first-order terms will be complex wave amplitudes*, so that we may return to using a simple bold-faced vector such as $\mathbf{E}_1$, *with the understanding that the exponential factor is implicit and that the physical vector quantity is the real part.* We will, however, retain the subscripts indicating order everywhere in this Unit, as well as the distinction of boldface versus plain to show vector versus scalar quantities.

There is one pitfall in this more-or-less standard approach. Sometimes we find ourselves multiplying together two first-order quantities to evaluate some second-order quantity, and often then time-averaging this second-order quantity. For example, suppose we want the time average of $\mathbf{A}_1 \cdot \mathbf{B}_1$; the proper answer is $\frac{1}{2}\mathrm{Re}[\mathbf{A}_1 \cdot \mathbf{B}_1^*]$.

Problem 15.1: Show that the time average of the dot product of two physical vector fields, $\mathbf{A}_1$ and $\mathbf{B}_1$, is $\langle \mathbf{A}_1 \cdot \mathbf{B}_1 \rangle = \frac{1}{2}\mathrm{Re}[\mathbf{A}_1 \cdot \mathbf{B}_1^*]$. The left-hand side of this equation represents the time-average of the *physical* fields, while the right-hand side evaluates this time-average in terms of the complex wave amplitudes. Allow arbitrary phase differences between $\mathbf{A}_1$ and $\mathbf{B}_1$.

15.2 GROUP VELOCITIES

We have already discussed the phase velocity of a wave—the speed at which a point of constant phase propagates forward along $\mathbf{k}/k$. If we make up a wave-packet of fast oscillations grouped together in time and space, as shown in Figure 15.1, this is the speed at which individual crests within the packet travel. However, these crests need not travel at the speed that the overall packet moves; the crests within the packet can slide forward or backward relative to the bundle of energy and information that constitutes the wave-packet. Indeed this frequently must be the case, since we will find that phase velocities in a plasma often exceed the speed of light, but the velocity of the group of waves (the 'group velocity') must be less than this, from fundamental considerations of special relativity.

Figure 15.1 shows a packet of oscillations with a Gaussian envelope. The amplitude $A(x)$ is given by

$$A(x) = \mathrm{Re}[\exp(-x^2/2\sigma^2)\exp(\mathrm{i}k_0 x)] \tag{15.8}$$

where we have chosen $k_0\sigma \gg 1$, so that there are many oscillations within the packet. The question we would like to investigate is: how does this wave-packet propagate in a dispersive medium where ω depends on k? Without deriving the principles of Fourier analysis, let us assert and later prove that the same $A(x)$ given in equation (15.8) can also be written

$$A(x) = \mathrm{Re}\left(\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp(\mathrm{i}kx)\exp[-\sigma^2(k-k_0)^2/2]\mathrm{d}k\right). \tag{15.9}$$

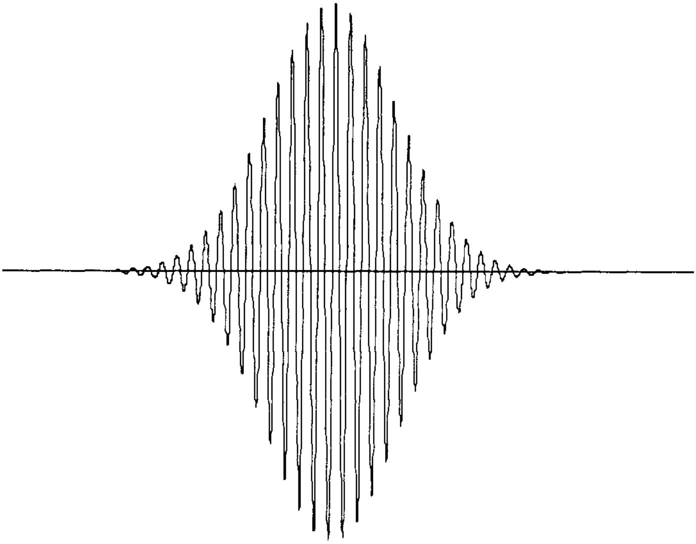

Figure 15.1. Wave-packet with a Gaussian envelope, constructed such that $k_0\sigma \ll 1$.

Equation (15.9) says that a wave-packet localized in space, x, can be considered to have been constructed of an integral over plane waves localized in wave-number, k.

Problem 15.2: Prove that the two forms of $A(x)$ given in equations (15.8) and (15.9) are equivalent. (A few tricks: transform $k' = k - k_0$; use the technique of completing the square in the exponent to transform the integral into an integral over a simple Gaussian; finally, use the facts that there are no poles in the complex plane for the resulting integrand, and that it goes to zero exponentially as Re $k \to \pm\infty$, so that any integral along a contour parallel to the real axis will give the same result.)

Equation (15.9) (and Figure 15.1) can be viewed as $t = 0$ freeze-frames of a set of propagating waves. The time evolution of this system is then just

$$A(x,t) = \mathrm{Re}\left(\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\{\mathrm{i}[kx - \omega(k)t]\}\exp[-\sigma^2(k-k_0)^2/2]\mathrm{d}k\right) \tag{15.10}$$

where we have explicitly denoted the k dependence of ω by using $\omega(k)$. For a narrow enough wave-packet in k space (which means a large σ, i.e. wide in physical space), we can approximate $\omega(k) \approx \omega(k_0) + (\partial\omega/\partial k)_{k_0}(k - k_0)$.

We further assume that the medium is dispersive, but not too dispersive, by neglecting quadratic terms in the expansion of ω in $(k-k_0)$. So, proceeding for our moderately dispersive medium, we obtain

$$A(x,t) = \mathrm{Re}\Bigg(\exp\{\mathrm{i}[k_0(\partial\omega/\partial k)_{k_0} - \omega(k_0)]t\} \times \frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\{\mathrm{i}[kx - k(\partial\omega/\partial k)_{k_0}t]\}\exp[-\sigma^2(k-k_0)^2/2]\,\mathrm{d}k\Bigg). \tag{15.11}$$

Now the factor beginning with $\sigma/\sqrt{2\pi}$ is *exactly* $A(x-(\partial\omega/\partial k)_{k_0}t, 0)$—in other words, the original $t=0$ freeze-frame, but translating at velocity $(\partial\omega/\partial k)_{k_0}$. This is just what we were looking for: the velocity of our wave-packet. So what is the factor on the first line? It is an overall space-independent time oscillation corresponding to the fact that the wave fronts are propagating at the phase velocity, ω/k, while the wave-packet moves at the group velocity, $\partial\omega/\partial k$, not equal to ω/k.

15.3 RAY-TRACING EQUATIONS

In an inhomogeneous plasma the trajectory of a wave-packet will be curved, responding to gradients in the plasma properties. We can derive the ray-tracing equations for the propagation of localized wave energy in a plasma simply from the considerations above. Consider a wave-packet localized not only in the longitudinal direction (parallel to $\mathbf{k}_0$), but also in the transverse direction (perpendicular to $\mathbf{k}_0$). For simplicity (but without loss of generality) let us assume $\mathbf{k}_0 \parallel \hat{\mathbf{x}}$, giving $\mathbf{k}_0 = k_0\hat{\mathbf{x}}$. Then the wave amplitude we desire can be expressed as

$$A(\mathbf{x}) = \mathrm{Re}[\exp(-x^2/2\sigma_x^2 - y^2/2\sigma_y^2 - z^2/2\sigma_z^2)\exp(\mathrm{i}k_0x)]. \tag{15.12}$$

By analogy with equation (15.9), we can re-express $A(\mathbf{x})$ in terms of its Fourier transform:

$$A(\mathbf{x}) = \mathrm{Re}\Bigg(\frac{\sigma_x\sigma_y\sigma_z}{(2\pi)^{3/2}}\int_{-\infty}^{\infty}\exp(\mathrm{i}\mathbf{k}\cdot\mathbf{x}) \times \exp[-\sigma_x^2(k_x-k_0)^2/2 - \sigma_y^2k_y^2/2 - \sigma_z^2k_z^2/2]\,\mathrm{d}^3\mathbf{k}\Bigg). \tag{15.13}$$

As before, we now consider this as a ‘freeze-frame’ picture at $t=0$, and include a factor $\exp(-\mathrm{i}\omega t)$, acknowledging that $\omega=\omega(\mathbf{k})$, where $\mathbf{k}$ is a vector quantity in our anisotropic medium. Carrying through a Taylor expansion as before, we approximate

$$\omega \simeq \omega(\mathbf{k}_0) + (\mathbf{k} - \mathbf{k}_0) \cdot \nabla_{\mathbf{k}}\omega|_{k_0} \tag{15.14}$$

where the meaning of $\nabla_{\mathbf{k}}\omega|_{k_0}$ is given by

$$\nabla_{\mathbf{k}}\omega \equiv \hat{\mathbf{x}}\frac{\partial\omega}{\partial k_x} + \hat{\mathbf{y}}\frac{\partial\omega}{\partial k_y} + \hat{\mathbf{z}}\frac{\partial\omega}{\partial k_z} = \frac{\partial\omega}{\partial\mathbf{k}} \tag{15.15}$$

evaluated at $\mathbf{k} = \mathbf{k}_0$. If we carry through the same analysis as equations (15.9)–(15.11), but in three dimensions, we will find our 'freeze-frame' $A(\mathbf{x})$ translating at a vector group velocity given by

$$\mathbf{v}_g = \frac{\partial\omega}{\partial\mathbf{k}} \tag{15.16}$$

with an overall time-dependent oscillation superimposed, as before. Note that $\mathbf{v}_g$ may not only have a different magnitude from $\mathbf{v}_p$, but even a different *direction*.

Problem 15.3: Prove Equation (15.16), following the derivation given in one dimension in equations (15.9)–(15.11).

We are assuming that the plasma medium is inhomogeneous so, based on our experience with light rays and lenses, there is no reason to expect the location of the peak of the $\mathbf{k}$ spectrum, $\mathbf{k}_0$, to be preserved. On the other hand, since the background medium is by hypothesis linear and time-independent, $\omega(\mathbf{k}_0)$ should be constant. This means that the total derivative of ω, moving with the wave-packet, must vanish. Assuming we know $\omega = \omega(\mathbf{x}, \mathbf{k})$ for our medium, the total derivative of ω can be expressed in terms of its partial derivatives by

$$\frac{\mathrm{d}\omega}{\mathrm{d}t} = \mathbf{v}_g \cdot \left.\frac{\partial\omega}{\partial\mathbf{x}}\right|_{\mathbf{k}} + \frac{\mathrm{d}\mathbf{k}_0}{\mathrm{d}t} \cdot \left.\frac{\partial\omega}{\partial\mathbf{k}}\right|_{\mathbf{x}} = 0. \tag{15.17}$$

The partial derivative with respect to $\mathbf{x}$ is at fixed $\mathbf{k}$, and *vice versa*. Thus we have, in general, 'equations of motion' or 'ray-tracing equations' for our wave-packet:

$$\frac{\mathrm{d}\mathbf{k}_0}{\mathrm{d}t} = -\left.\frac{\partial\omega}{\partial\mathbf{x}}\right|_{\mathbf{k}} \qquad \frac{\mathrm{d}\mathbf{x}_0}{\mathrm{d}t} = \left.\frac{\partial\omega}{\partial\mathbf{k}}\right|_{\mathbf{x}}. \tag{15.18}$$

As the wave-packet propagates it maintains the peak of its frequency spectrum, but its wave-number spectrum transforms. To trace out a 'ray' one must integrate forward in time the packet's position in both $\mathbf{x}$- and $\mathbf{k}$-space, since the future propagation depends on *both* $\mathbf{x}_0$ and $\mathbf{k}_0$.

The analogy to Hamiltonian mechanics is evident, as is the parallel with quantum mechanics, where $\hbar\omega$ is identified as the energy of a photon and $\hbar\mathbf{k}$ as its momentum. The ray-tracing equations are only valid in the limit of so-called 'geometrical optics', where the wave-packet is also well localized in physical space such that $\delta\mathbf{x} \cdot \partial\omega/\partial\mathbf{x} \ll \omega$, where $\delta\mathbf{x} = \sigma_x\hat{\mathbf{x}} + \sigma_y\hat{\mathbf{y}} + \sigma_z\hat{\mathbf{z}}$, and is well localized in $\mathbf{k}$-space such that $\delta\mathbf{k} \cdot \partial\omega/\partial\mathbf{k} \ll \omega$, where $\delta\mathbf{k} = \hat{\mathbf{x}}/\sigma_x + \hat{\mathbf{y}}/\sigma_y + \hat{\mathbf{z}}/\sigma_z$.

In this same limit of geometrical optics, we can use the Wentzel–Kramers–Brillouin (WKB) approximation to determine the wave phase at any location along the ray trajectory. In this approach we note that $\mathbf{k}_0(t)$ is implicitly a function of $\mathbf{x}_0(t)$ along the ray, since both are explicitly functions of t. If we imagine sending out a steady beam of radiation, rather than a wave-packet, the energy will still propagate along the group velocity vector. Along this ray-trajectory, now, the continuous spatial derivative of the wave phase will be $\mathbf{k}_0$, while the time-derivative of the phase will continue to be $-\omega_0$ (which does not vary in time or space). Thus the phase difference at fixed time between two points $\mathbf{x}_0$ and $\mathbf{x}_1$ along the ray path, $\mathbf{l}$, is given by

$$\Delta\phi = \int_{\mathbf{x}_0}^{\mathbf{x}_1} \mathbf{k}_0 \cdot d\mathbf{l}.$$