Chapter 6

Geometrical Optics

6.1 Overview

Geometrical optics, the study of "rays," is the oldest approach to optics. It is an accurate description of wave propagation whenever the wavelengths and periods of the waves are far smaller than the lengthscales and timescales for variations of the wave amplitude and of the medium through which the waves propagate.

We shall begin our study of geometric optics in Sec. 6.3 by deriving the geometric-optic propagation equations with the aid of the *eikonal* approximation, and we shall elucicate the connection to Hamilton-Jacobi theory. This connection will be made more explicit by demonstrating that a classical, geometric-optics wave can be interpreted as a flux of quanta. In Sec. 6.4 we shall specialize the geometric optics formalism to any situation where a bundle of nearly parallel rays is being guided and manipulated by some sort of apparatus. This is called the *paraxial approximation*, and we shall illustrate it using the problem of magnetically focusing a beam of charged particles and shall show how matrix methods can be used to describe the particle (i.e. ray) trajectories. In Sec. 6.5, we shall turn from scalar waves to the vector waves of electromagnetic radiation. We shall deduce the geometric-optics propagation law for the waves' polarization vector and shall explore the classical version of "geometrical" (or "adiabatic" or "Berry") phase. Finally, In Sec. 6.5, we shall discuss the formation of images in geometrical optics, illustrating our treatment with gravitational lenses. We shall pay special attention to the behavior of images at caustics, and its relationship to catastrophe theory.

6.2 Waves in a Homogeneous Medium

6.2.1 Monochromatic, plane waves

Consider a monochromatic plane wave propagating through a homogeneous medium. Independently of the physical nature of the wave, it can be described mathematically by

$$\psi = Ae^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)} , \tag{6.1}$$

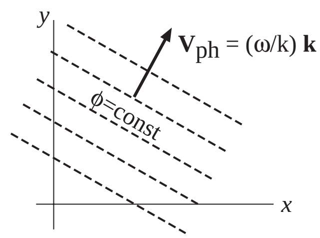

Fig. 6.1: A monochromatic plane wave in a homogeneous medium.

where ψ is any oscillatory physical quantity associated with the wave. If, as is usually the case, the physical quantity is real (not complex), then we must take the real part of Eq. (6.1). In Eq. (6.1) A is the wave's complex amplitude, t and $\mathbf{x}$ are time and location in space, $\omega = 2\pi f$ is the wave's *angular frequency*, and $\mathbf{k}$ is its *wave vector* (with $k \equiv |\mathbf{k}|$ its wave number, $\lambda = 2\pi/k$ its *wavelength*, $\lambda\!\!\!^{-} = \lambda/2\pi$, it reduced wavelengthand $\hat{\mathbf{k}} \equiv \mathbf{k}/k$ its *wave-vector direction*). The quantity in the exponential, $\phi = \mathbf{k}\cdot\mathbf{x} - \omega t$, is the wave's *phase*. Surfaces of constant phase are orthogonal to the propagation direction $\hat{\mathbf{k}}$ and move with the *phase velocity*

$$\mathbf{V}_{\rm ph} \equiv \left(\frac{\partial \mathbf{x}}{\partial t}\right)_\phi = -\frac{\left(\frac{\partial\phi}{\partial t}\right)_\mathbf{x}}{\left(\frac{\partial\phi}{\partial \mathbf{x}}\right)_t} = \frac{\omega}{k}\hat{\mathbf{k}}\ ; \tag{6.2}$$

cf. Fig. 6.1. The frequency ω is determined by the wave vector $\mathbf{k}$ in a manner that depends on the wave's physical nature; the functional relationship $\omega = \Omega(\mathbf{k})$ is called the wave's *dispersion relation*.

Some examples of plane waves that we shall study in this book are: (i) Electromagnetic waves propagating through a dielectric medium with index of refraction n (this chapter), for which ψ could be any Cartesian component of the electric or magnetic field or vector potential and the dispersion relation is

$$\omega = \Omega(\mathbf{k}) = \frac{c}{n}k = \frac{c}{n}|\mathbf{k}|\ , \tag{6.3}$$

with c the speed of light in vacuum. (ii) Sound waves propagating through a solid (Part III) or fluid (liquid or vapor; Part IV), for which ψ could be the pressure or density perturbation produced by the sound wave, and the dispersion relation is the same as for electromagnetic waves (6.3), but with c now the sound speed under some fiducial condition at which (by convention) we set $n = 1$. (iii) Waves on the surface of a deep body of water (depth $\gg \lambda\!\!\!^{-}$; Part IV), for which ψ could be the height of the water above equilibrium and the dispersion relation is

$$\omega = \Omega(\mathbf{k}) = \sqrt{gk} = \sqrt{g|\mathbf{k}|}\ , \tag{6.4}$$

with g the acceleration of gravity. (iv) Flexural waves on a stiff beam or rod (Part III), for which ψ could be the transverse displacement of the beam from equilibrium and the dispersion relation is

$$\omega = \Omega(\mathbf{k}) = \sqrt{\frac{EJ}{\Lambda}}k^2 = \sqrt{\frac{EJ}{\Lambda}}\mathbf{k}\cdot\mathbf{k}\ , \tag{6.5}$$

where EJ is the rod's "flexural rigidity" and Λ is its mass per unit length. (v) Alfvén waves (bending oscillations of plasma-laden magnetic field lines) in a magnetized plasma, for which ψ could be the transverse displacement of a magnetic field line and the dispersion relation is

$$\omega = \Omega(\mathbf{k}) = \mathbf{a} \cdot \mathbf{k}, \tag{6.6}$$

with $\mathbf{a} = \mathbf{B}/\sqrt{\mu_o \rho}$ (SI units), $\mathbf{a} = \mathbf{B}/\sqrt{4\pi\rho}$ the *Alfvén speed*, $\mathbf{B}$ the (homogeneous) magnetic field, μ_o the magnetic permitivity of the vacuum, and ρ the plasma mass density.

6.2.2 Wave packets

Waves in the real world are not precisely monochromatic and planar. Instead, they occupy wave packets that are somewhat localized in space and time. Such wavepackets can be constructed as superpositions of plane waves:

$$\psi(\mathbf{x}, t) = \int \mathrm{A}(\mathbf{k}) e^{i\alpha(\mathbf{k})} e^{i(\mathbf{k}\cdot\mathbf{x} - \omega t)} d^3k \ , \tag{6.7}$$

where A is the modulus and α the phase of the complex amplitude A, and the integration element is $d^3k \equiv dk_x dk_y dk_z$ in terms of components of $\mathbf{k}$ on Cartesian axes x, y, z. Suppose, as is often the case, that $\mathrm{A}(\mathbf{k})$ is sharply concentrated around some specific wave vector $\mathbf{k}_o$. Then in the integral, the contributions from adjacent $\mathbf{k}$'s will tend to cancel each other except in that region of space and time where the oscillatory phase factor changes little with changing $\mathbf{k}$. This is the spacetime region in which the wave packet is concentrated, and its center is where $\boldsymbol{\nabla}_\mathbf{k}(\text{phasefactor}) = 0$:

$$\left(\frac{\partial \alpha}{\partial k_j} + \frac{\partial}{\partial k_j}(\mathbf{k} \cdot \mathbf{x} - \omega t) \right)_{\mathbf{k}=\mathbf{k}_o} = 0 \ . \tag{6.8}$$

Evaluating the derivative with the aid of the wave's dispersion relation, we obtain for the location of the wave packet's center

$$\mathbf{x}_j - \left(\frac{\partial \Omega}{\partial k_j} \right)_{\mathbf{k}=\mathbf{k}_o} t = - \left(\frac{\partial \alpha}{\partial k_j} \right)_{\mathbf{k}=\mathbf{k}_o} = \text{const} \ . \tag{6.9}$$

This tells us that the wave packet moves with the *group velocity*

$$\mathbf{V}_g = \nabla_\mathbf{k} \Omega \ , \quad \text{i.e.} \quad \mathbf{V}_g^j = \left(\frac{\partial \Omega}{\partial k_j} \right)_{\mathbf{k}=\mathbf{k}_o} \ . \tag{6.10}$$

When, as for electromagnetic waves in a dielectric medium or sound waves in a solid or fluid, the dispersion relation has the simple form $\omega = \Omega(\mathbf{k}) \propto k = |\mathbf{k}|$, then the group and phase velocities are the same,

$$\mathbf{V}_g = \mathbf{V}_{\text{ph}} = \frac{c}{n} \hat{\mathbf{k}} \ , \tag{6.11}$$

when the dispersion relation is written in the form of Eq. (6.3), and the waves are said to be *dispersionless*. If the dispersion relation has any other form, then the group and phase

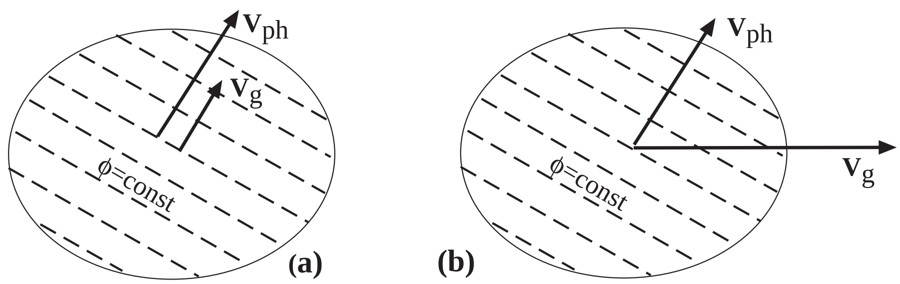

Fig. 6.2: (a) A Wave packet of waves on a deep body of water. The packet is localized in the spatial region bounded by the thin ellipse. Its center moves with the group velocity $\mathbf{V}_g$, and its surfaces of constant phase (the wave's oscillations) move twice as fast and in the same direction, $\mathbf{V}_{\rm ph} = 2\mathbf{V}_g$. This means that the wave's oscillations arise at the back of the packet and move forward through the packet, disappearing at the front. The wavelength of these oscillations is $\lambda = 2\pi/k_o$, where $k_o = |\mathbf{k}_o|$ is the wavenumber about which the wave packet is concentrated [Eq. (6.7)]. (b) An Alfvén wave packet. Its center moves with a group velocity $\mathbf{V}_g$ that points along the homogeneous magnetic field [Eq. (6.14)], and its surfaces of constant phase (the wave's oscillations) move with a phase velocity $\mathbf{V}_{\rm ph}$ that can be in any direction $\hat{\mathbf{k}}$. The phase speed is the projection of the group velocity onto the phase propagation direction, $|\mathbf{V}_{\rm ph}| = \mathbf{V}_g \cdot \hat{\mathbf{k}}$ [Eq. (6.14)], which implies that the wave's oscillations remain fixed inside the packet as the packet moves.

velocities are different, and the wave is said to exhibit *dispersion*; cf. Ex. 6.2. Examples are (see above): (iii) Waves on a deep body of water [dispersion relation (6.4)] for which

$$\mathbf{V}_g = \frac{1}{2}\mathbf{V}_{\rm ph} = \frac{1}{2}\sqrt{\frac{g}{k}}\hat{v}k \; . \tag{6.12}$$

(iv) Flexural waves on a rod or beam [dispersion relation (6.5)] for which

$$\mathbf{V}_g = 2\mathbf{V}_{\rm ph} = 2\sqrt{\frac{EJ}{\Lambda}}k\mathbf{k} \; . \tag{6.13}$$

(v) Alfvén waves in a magnetized plasma [dispersion relation (6.6)] for which

$$\mathbf{V}_g = \mathbf{a} \; , \quad \mathbf{V}_{\rm ph} = (\mathbf{a}\cdot\hat{\mathbf{k}})\hat{\mathbf{k}} \; . \tag{6.14}$$

Notice that the group speed $|\mathbf{V}_g|$ can be less than or greater than the phase speed, and if the homogeneous medium is anisotropic (e.g., for a magnetized plasma), the group velocity can point in a different direction than the phase velocity; see Fig. 6.2

It should be obvious, physically, that the energy contained in a wavepacket must remain always with the packet and cannot move into the region outside the packet where the wave amplitude vanishes. Correspondingly, *the wave packet's energy must propagate with the group velocity* $\mathbf{V}_g$ *and not with the phase velocity* $\mathbf{V}_{\rm ph}$. Similarly, when one examines the wave packet from a quantum mechanical viewpoint, *its quanta must move with the group velocity* $\mathbf{V}_g$. Since we have required that the wave packet have its wave vectors concentrated around $\mathbf{k}_o$, the energy and momentum of each of the packet's quanta are $\tilde{E} = \hbar\Omega(\mathbf{k}_o)$ and $\mathbf{p} = \hbar\mathbf{k}_o$.

****

EXERCISES

Exercise 6.1 *Practice: Group and Phase Velocities*

Derive the group and phase velocities (6.11)–(6.14) from the dispersion relations (6.3)–(6.6).

Exercise 6.2 *Example: Wave-Packet Size and Spreading*

Consider a one-dimensional wave packet, $\psi(x,t) = \int \mathrm{A}(k) e^{i\alpha(k)} e^{i(kx-\omega t)} dk$ with dispersion relation $\omega = \Omega(k)$. Let the amplitude $\mathrm{A}(k)$ be concentrated around $k = k_o$ and contain a range of wave numbers $\Delta k \ll k_o$.

(a) Show that the size of the wave packet is $L \simeq \pi/\Delta k$. Discuss the relationship of this result to the uncertainty principle for the localization of the packet's quanta.

(b) Because the packet contains a finite range of k's, each of which has a different group velocity, the packet's shape and size will change with time. Show that the timescale for the packet to enlarge by a factor 2 is

$$\tau \sim \frac{L^2}{\pi |d^2\Omega/dk^2|_{k=k_o}} .$$

For waves on a deep body of water, how far will a wave packet travel during this spreading time? Express your answer in terms of the packet's wavelength λ_o and its size L.

****

6.3 Waves in an Inhomogeneous, Time-Varying Medium: The Eikonal Approximation

Suppose that the medium in which the waves propagate is spatially inhomogeneous and varies with time. If the lengthscale $\mathcal{L}$ and timescale $\mathcal{T}$ for substantial variations are long compared to the waves' reduced wavelength and period,

$$\mathcal{L} \gg \lambdabar_o = 1/k , \quad \mathcal{T} \gg 1/\omega , \tag{6.15}$$

then the waves can be locally planar and monochromatic. The medium's inhomogeneities and time variations may produce variations in the wave vector $\mathbf{k}$ and frequency ω, but those variations should be substantial only on scales $\gtrsim \mathcal{L} \gg 1/k$ and $\gtrsim \mathcal{T} \gg 1/\omega$. This