Fundamentals of Nonlinear Optics

ECED 6400 Lecture Notes

February 15, 2014

Contents

Chapter 1

Introduction

In this course, we will be describing all optical phenomena classically within the framework of *macroscopic Maxwell's equations* written in terms of *macroscopic* electromagnetic fields. The latter are obtained by averaging rapidly varying microscopic fields over spatial scales much larger than characteristic material microstructure scales (atomic size, lattice scale, etc). The averaging procedure is examined in detail in standard electrodynamics textbooks[1]. Within the framework of such a phenomenological approach, which circumvents a detailed microscopic light-matter interaction description, *external* or *driving* volume charge and current densities, $\rho_{\rm ex}$ and $\mathbf{J}_{\rm ex}$, give rise to the electromagnetic fields obeying the Maxwell equations in the form

$$\nabla \cdot \mathbf{D} = \rho_{\rm ex}, \tag{1.1}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{1.2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{1.3}$$

and

$$\nabla \times \mathbf{H} = \mathbf{J}_{\rm ex} + \frac{\partial \mathbf{D}}{\partial t}. \tag{1.4}$$

The set of equations (1.1) through (1.4) is not closed, however, until we provide any information about the material media. Such information is furnished by supplying phenomenological constitutive relations among the four fields, $\mathbf{E}$, $\mathbf{D}$, $\mathbf{B}$, and $\mathbf{H}$. Without much loss of generality we will assume hereafter that all material media are nonmagnetic, which holds true for virtually all natural media at optical frequencies [2]. We can then represent the magnetic constitutive relation in its simplest form as

$$\mathbf{B} = \mu_0 \mathbf{H}, \tag{1.5}$$

[1] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, NY, 1999) 3rd edition.

[2] This criterion, however, breaks down for some artificial materials, the so-called metamaterials, which we will not consider in this course.

with μ_0 being the free space permeability in the SI units we will be employing hereafter. A general electric constitutive relation states

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}; \qquad \mathbf{P} = \mathbf{P}(\mathbf{E}), \tag{1.6}$$

where ϵ_0 is the free space permittivity and $\mathbf{P}$ is a macroscopic polarization field. The latter is in turn a function of the applied electric field. For sufficiently weak applied fields, $\mathbf{P}$ is a linear function of $\mathbf{E}$; this is a regime of *linear optics*. However, even in the linear optics regime, the dependence of the polarization on the applied electric field can be rather complicated to account for possible medium anisotropy and–temporal and sometimes even spatial–dispersion. While the former implies that the medium response in a particular direction can be affected by the electric field components orthogonal to this direction, the latter acknowledges the fact that the medium response at a given space-time point can depend on the applied electric field in the past (temporal dispersion) and/or on the fields in the neighborhood of the spatial point (spatial dispersion). We will study all these cases in detail in the subsequent chapters.

As the magnitude of the applied electric field increases, the linear relationship between $\mathbf{P}$ and $\mathbf{E}$ breaks down and we enter the realm of *nonlinear optics*. If the electric field intensity is far below a critical value, $E_{\rm cr} \sim 10^9$ V/cm needed to ionize a material atom, the resulting polarization can be expressed as a series in increasing powers of the electric field. Schematically, such a series can be expressed as

$$P = \chi^{(1)} E + \chi^{(2)} E^2 + \chi^{(3)} E^3 + \ldots, \tag{1.7}$$

where we ignored the vector nature of the fields as well as dispersion, for simplicity. The expansion coefficients, $\chi^{(1)}$ and $\chi^{(2)}$, etc., are identified as linear and nonlinear susceptibilities, respectively. The linear and nonlinear susceptibilities should be treated as phenomenological constants in our classical description. The condition $E \ll E_{\rm cr}$ is typically met with a vast majority of laser sources which rarely generate fields in excess of 10^6 V/cm. However, even if the applied field does not exceed $E_{\rm cr}$, the power series expansion can fail, provided the carrier frequency of the field lies close to any internal resonance of the medium. In the latter case, the material response tends to saturate at high enough field intensities. The proper quantitative description of such nonlinear saturation phenomena calls for a quantum mechanical treatment of the medium. Whenever, the power expansion of $\mathbf{P}$ is valid, though, we shall refer to the lowest-order term in the expansion as a linear contribution and designate the rest to be *nonlinear polarization* such that

$$\mathbf{P} = \mathbf{P}_{\rm L} + \mathbf{P}_{\rm NL}. \tag{1.8}$$

Next, the external charge and current densities are not independent from each other. Rather they are related by another fundamental law, the charge conservation law, which takes the form of a well-known continuity equation viz.,

$$\frac{\partial \rho_{\rm ex}}{\partial t} + \nabla \cdot \mathbf{J}_{\rm ex} = 0. \tag{1.9}$$

The external $\rho_{\rm ex}$ and $\mathbf{J}_{\rm ex}$ drive the electromagnetic fields which, in turn, induce *internal* charge and current densities, ρ and $\mathbf{J}$, inside a medium. The induced charges and

currents can be of either free (conduction) or bound (polarized) type and they also obey the continuity equation,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0. \tag{1.10}$$

We stress that charge conservation (1.9) amounts to a fundamental law which does not follow from Maxwell's equations. The electromagnetic energy conservation law, however, does follow from the Maxwell equations by the same token as the mechanical energy conservation follows from Newton's laws.

To derive the electromagnetic energy conservation law or the Poynting theorem, we take dot products of the both sides of Eqs. (1.3) and (1.4) with $\mathbf{H}$ and $\mathbf{E}$, respectively, and use the constitutive relations (1.5) and (1.6), yielding

$$\mathbf{H} \cdot (\nabla \times \mathbf{E}) = -\mu_0 \mathbf{H} \cdot \frac{\partial \mathbf{H}}{\partial t} \tag{1.11}$$

and

$$\mathbf{E} \cdot (\nabla \times \mathbf{H}) = \mathbf{J}_{\rm ex} \cdot \mathbf{E} + \epsilon_0 \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} + \mathbf{E} \cdot \frac{\partial \mathbf{P}}{\partial t}. \tag{1.12}$$

On subtracting Eq. (1.11) from Eq. (1.12) term by term, we obtain

$$\frac{\epsilon_0}{2} \frac{\partial E^2}{\partial t} + \frac{\mu_0}{2} \frac{\partial H^2}{\partial t} + \mathbf{J}_{\rm ex} \cdot \mathbf{E} = \mathbf{E} \cdot (\nabla \times \mathbf{H}) - \mathbf{H} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot \frac{\partial \mathbf{P}}{\partial t}. \tag{1.13}$$

Further, using the vector identity

$$\nabla \cdot (\mathbf{E} \times \mathbf{H}) = \mathbf{H} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{H}) \tag{1.14}$$

we arrive, after minor algebra, at a differential form of the electromagnetic energy conservation equation

$$\frac{\partial w_{em}}{\partial t} + \nabla \cdot \mathbf{S} = -\mathbf{J}_{\rm ex} \cdot \mathbf{E} - \mathbf{E} \cdot \frac{\partial \mathbf{P}}{\partial t}. \tag{1.15}$$

Here the electromagnetic energy density w_{em} is defined in the same way as in free space,

$$w_{em} = \tfrac{1}{2}\epsilon_0 E^2 + \tfrac{1}{2}\mu_0 H^2, \tag{1.16}$$

and we introduced the electromagnetic energy flux density, the so-called Poynting vector, by the expression

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}. \tag{1.17}$$

Equation (1.15) is often referred to as Poynting's theorem. In essence, it implies that the time rate of change of the electromagnetic energy density is determined by the energy flux density minus losses associated with external as well as internal currents. The second term on the r.h.s of Eq. (1.15) describes Ohmic losses associated with external currents and the third one is identified with the energy loss caused by induced polarization currents, including the ones associated with the generation of nonlinear polarizations. To reexpress the right-hand side of Eq. (1.15) in a more symmetric from, we can explicitly define the induced polarization currents as

$$\mathbf{J} = \frac{\partial \mathbf{P}}{\partial t}, \tag{1.18}$$

and introduce the corresponding induced charge densities as

$$\rho = -\nabla \cdot \mathbf{P}, \tag{1.19}$$

such that the continuity equation (1.10) is respected. We emphasize that our separation of the charges into external (driving) and internal (induced) is much more natural at optical frequencies than the standard separation into free and bound charges. The latter is a rather arbitrary division[3] which can be confusing at optical frequencies, especially for metals[4]. We will then unify free and bound induced charges and currents under the umbrella of ρ and $\mathbf{J}$.

To better understand Eq. (1.15), we transform it into the integral form

$$\frac{dW_{em}}{dt} = -\oint_{\sigma} d\boldsymbol{\sigma} \cdot \mathbf{S} - \int_{v} dv(\mathbf{J}_{\rm ex} + \mathbf{J}) \cdot \mathbf{E}. \tag{1.20}$$

where

$$W_{em} = \int_{v} dv w_{em}, \tag{1.21}$$

is the total energy of electromagnetic field inside a given volume v, and we used a divergence theorem to convert a volume integral on the the r.h.s of (1.20) into the surface one. Equation (1.20) then implies that the total energy change inside a finite region of the medium can occur as a result of the energy outflow through the boundary surface of the region as well as via energy losses inside the region associated with driving and induced currents. This situation is schematically illustrated in Fig. 1.

In many practical situations in nonlinear optics one deals with pulse or beam fields with their carriers oscillating at optical frequencies. Such fast oscillations can never be detected by even the fastest modern detectors whose response time is much larger that an optical period. Consequently, it makes sense to talk about the field quantities averaged over many optical cycles – it is those quantities that can actually be registered in optical measurements anyway. In the absence of external currents, $\mathbf{J}_{\rm ex} = 0$, we can rewrite the time-averaged Poynting theorem as

$$\left\langle \frac{\partial w_{em}}{\partial t} \right\rangle + \nabla \cdot \langle \mathbf{S} \rangle = -\left\langle \mathbf{E} \cdot \frac{\partial \mathbf{P}}{\partial t} \right\rangle. \tag{1.22}$$

Here we define time-averaged quantities such as the average Poynting vector by the expression

$$\langle \mathbf{S} \rangle = \langle \mathbf{E} \times \mathbf{H} \rangle \equiv \frac{1}{T} \int_{t-T/2}^{t+T/2} dt\, (\mathbf{E} \times \mathbf{H}), \tag{1.23}$$

where T is an optical period. In case of monochromatic fields, which can be conveniently represented via complex amplitudes as

$$\mathbf{E} = \tfrac{1}{2}(\boldsymbol{\mathcal{E}} e^{-i\omega t} + c.\, c.), \tag{1.24}$$

[3] Yu. A. Illinskii, L. V. Keldysh, *Electromagnetic response of material media* (Plenum Press, New York, NY, 1994).

[4] S. A. Maier, *Plasmonics, Fundamentals and Applications* (Springer, Berlin, 2007).

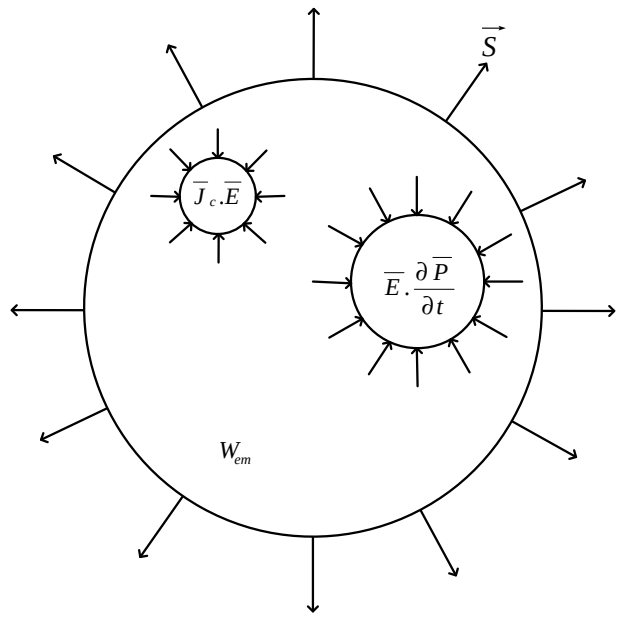

Figure 1.1: Schematic illustration of energy conservation in nonlinear media.

and

$$\mathbf{H} = \tfrac{1}{2}(\boldsymbol{\mathcal{H}} e^{-i\omega t} + c.\ c.), \tag{1.25}$$

Eq. (1.23) can be shown to reduce to

$$\langle \mathbf{S} \rangle = \tfrac{1}{2}\mathrm{Re}(\boldsymbol{\mathcal{E}} \times \boldsymbol{\mathcal{H}}^*). \tag{1.26}$$

Chapter 2

Plane electromagnetic waves in linear media

2.1 Plane waves in free space

In the absence of external charges and currents, Maxwell's equations in free space take the form

$$\nabla \cdot \mathbf{E} = 0, \tag{2.1}$$

$$\nabla \cdot \mathbf{H} = 0, \tag{2.2}$$

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t}, \tag{2.3}$$

and

$$\nabla \times \mathbf{H} = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \tag{2.4}$$

Linearity, stationarity, and homogeneity of Maxwell's equations in free space point to the existence of plane-wave solutions in the form

$$\mathbf{E}(\mathbf{r}, t) = \mathrm{Re}\{\boldsymbol{\mathcal{E}} e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}\}, \qquad \mathbf{H}(\mathbf{r}, t) = \mathrm{Re}\{\boldsymbol{\mathcal{H}} e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}\}. \tag{2.5}$$

By linearity of Maxwell's equations in free space, we can drop the real part and deal with complex phasors describing the waves directly. The real part can be taken at the end of all calculations to yield physical (real) electric and magnetic fields of a plane wave.

The Maxwell equations in the plane-wave form can be rewritten as

$$\mathbf{k} \cdot \boldsymbol{\mathcal{E}} = 0, \tag{2.6}$$

$$\mathbf{k} \cdot \boldsymbol{\mathcal{H}} = 0, \tag{2.7}$$

$$\mathbf{k} \times \boldsymbol{\mathcal{E}} = \omega \mu_0 \boldsymbol{\mathcal{H}}, \tag{2.8}$$

and

$$\mathbf{k} \times \boldsymbol{\mathcal{H}} = -\omega \epsilon_0 \boldsymbol{\mathcal{E}}. \tag{2.9}$$

In Eqs. (2.6) – (2.9) we dropped plane-wave phasors on both sides.

Next, we can exclude the magnetic field from the fourth Maxwell equation leading to

$$\mathbf{k} \times (\mathbf{k} \times \boldsymbol{\mathcal{E}}) = -\epsilon_0 \mu_0 \omega^2 \boldsymbol{\mathcal{E}}. \tag{2.10}$$

Rearranging the double cross-product on the left-hand side of Eq. (2.10), we arrive at

$$\mathbf{k}(\mathbf{k} \cdot \boldsymbol{\mathcal{E}}) - k^2 \boldsymbol{\mathcal{E}} = -\epsilon_0 \mu_0 \omega^2 \boldsymbol{\mathcal{E}}. \tag{2.11}$$

With the aid of Eq. (2.6), we obtain

$$(k^2 - \mu_0 \epsilon_0 \omega^2) \boldsymbol{\mathcal{E}} = 0, \tag{2.12}$$

implying that

$$k = \omega \sqrt{\epsilon_0 \mu_0} = \omega / c \tag{2.13}$$

where we introduced the speed of light in vacuum

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} = 3 \times 10^8 \text{ m/s}. \tag{2.14}$$

Equation (2.13) is a dispersion relation for plane electromagnetic waves in free space; it relates the wave number to the wave frequency. The complex amplitudes $\boldsymbol{\mathcal{E}}$ and $\boldsymbol{\mathcal{H}}$–which determine the directions of $\mathbf{E}$ and $\mathbf{H}$–are not independent, but are related by the Maxwell equations (2.8) or (2.9). For instance, from the knowledge of $\boldsymbol{\mathcal{E}}$ one can determine $\boldsymbol{\mathcal{H}}$ using Eq. (2.8),

$$\boldsymbol{\mathcal{H}} = \frac{(\mathbf{e}_k \times \boldsymbol{\mathcal{E}})}{\eta_0}, \tag{2.15}$$

where $\mathbf{e}_k = \mathbf{k}/k$ and η_0 is the **free space impedance** defined as

$$\eta_0 = \sqrt{\frac{\mu_0}{\epsilon_0}} \simeq 377\,\Omega. \tag{2.16}$$

By the same token, $\mathbf{E}_0$ can be inferred from $\mathbf{H}_0$ with the help of Eq. (2.9):

$$\boldsymbol{\mathcal{E}} = -\eta_0 (\mathbf{e}_k \times \boldsymbol{\mathcal{H}}). \tag{2.17}$$

It follows at once from Eqs. (2.15) and (2.17) that $\boldsymbol{\mathcal{E}}$, $\mathbf{k}$ and $\boldsymbol{\mathcal{H}}$ are mutually orthogonal for a plane wave in free space.

By convention, the wave polarization is associated with the time evolution of the electric field vector. Let us consider a plane wave propagating along the z-axis in free space. As, $\mathbf{k} = k\mathbf{e}_z$, and $\mathbf{E} \perp \mathbf{k}$, the electric field in the phasor form reads

$$\mathbf{E}(z,t) = \text{Re}\{(\mathbf{e}_x |\mathcal{E}_x| e^{i\phi_{0x}} + \mathbf{e}_y |\mathcal{E}_y| e^{i\phi_{0y}}) e^{i(kz - \omega t)}\}, \tag{2.18}$$

We will now show that, in general, the tip of the electric field vector moves around an ellipse as the time evolves. This general polarization is called **elliptic**. To proceed,

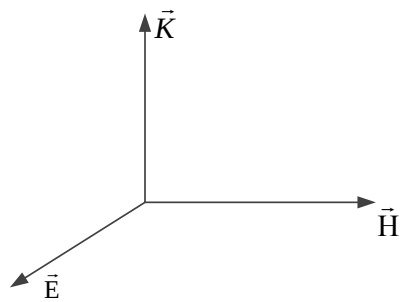

Figure 2.1: Mutual orientation of $\mathbf{E}$, $\mathbf{H}$ and $\mathbf{k}$ of a plane wave propagating in free space.

we rewrite the complex amplitude in the rectangular form as

$$\begin{aligned}\mathcal{E}_x\mathbf{e}_x + \mathcal{E}_y\mathbf{e}_y &= \underbrace{(\mathbf{e}_x|\mathcal{E}_x|\cos\phi_{0x} + \mathbf{e}_y|\mathcal{E}_y|\cos\phi_{0y})}_{\mathbf{U}} \\ &\quad + i\underbrace{(\mathbf{e}_x|\mathcal{E}_x|\sin\phi_{0x} + \mathbf{e}_y|\mathcal{E}_y|\sin\phi_{0y})}_{\mathbf{V}}. \end{aligned} \tag{2.19}$$

Note that $\mathbf{U}$ and $\mathbf{V}$ are not orthogonal which makes the situation tricky. We can however introduce a transformation from $\mathbf{U}$ and $\mathbf{V}$ to $\mathbf{u}$, $\mathbf{v}$ involving an auxiliary parameter θ such that

$$\mathbf{U} + i\mathbf{V} = (\mathbf{u} + i\mathbf{v})e^{i\theta}, \tag{2.20}$$

It follows at once from Eq. (2.20) that

$$\mathbf{U} = \mathbf{u}\cos\theta - \mathbf{v}\sin\theta, \qquad \mathbf{V} = \mathbf{u}\sin\theta + \mathbf{v}\cos\theta. \tag{2.21}$$

Inverting Eqs. (2.21), we obtain

$$\mathbf{u} = \mathbf{U}\cos\theta + \mathbf{V}\sin\theta, \qquad \mathbf{v} = \mathbf{U}\sin\theta - \mathbf{V}\cos\theta. \tag{2.22}$$

We can now use our freedom to choose θ wisely. In particular, choosing it such that $\mathbf{u}\cdot\mathbf{v} = 0$ (orthogonal axes), we obtain by taking the dot product of $\mathbf{u}$ and $\mathbf{v}$,

$$\tan 2\theta = \frac{2\mathbf{U}\cdot\mathbf{V}}{U^2 - V^2} \implies \theta = \frac{1}{2}\tan^{-1}\left(\frac{2\mathbf{U}\cdot\mathbf{V}}{U^2 - V^2}\right). \tag{2.23}$$

Here we made use of the trigonometric identities, $\sin 2\theta = 2\sin\theta\cos\theta$ and $\cos 2\theta = \cos^2\theta - \sin^2\theta$. By combining Eqs. (2.19) and (2.20), we can rewrite our field as

$$\mathbf{E}(z,t) = \mathrm{Re}\{(\mathbf{u} + i\mathbf{v})e^{i(kz-\omega t+\theta)}\}. \tag{2.24}$$

Using the orthogonality of $\mathbf{u}$ and $\mathbf{v}$, we can write the two orthogonal components of the field, E_u and E_v as

$$E_u = u\cos(kz - \omega t + \theta), \qquad E_v = v\sin(kz - \omega t + \theta). \tag{2.25}$$

It follows from Eq. (2.25) that

$$\frac{E_u^2}{u^2} + \frac{E_v^2}{v^2} = 1, \tag{2.26}$$

where u and v are given by Eq. (2.22) and θ by Eq. (2.23). Eq. (2.26) manifestly represents an ellipse with the semi-major axis making the angle θ with the x-axis as is shown in Fig. 3.6. The tip of $\mathbf{E}$ can move either clockwise or counterclockwise along

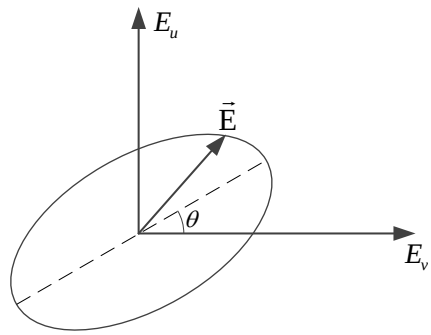

Figure 2.2: Illustrating elliptic polarization.

the ellipse; depending on the direction of motion of $\mathbf{E}$, the polarization is left-hand or right-hand elliptical. In the left-hand (right-hand) elliptical polarization, the fingers of your left (right) hand follow the direction of rotation and the thumb points to the wave propagation direction. Thus, for a general **elliptic** polarization, the electric field

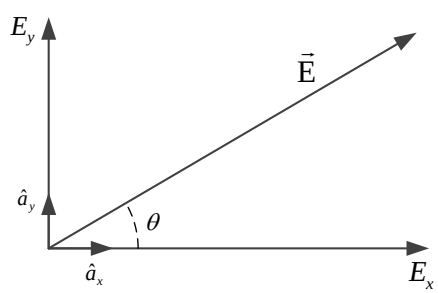

Figure 2.3: Illustrating linear polarization.

amplitude takes the form

$$\mathbf{E}(z,t) = \mathbf{e}_x|\mathcal{E}_x|\cos(kz - \omega t + \phi_{0x}) + \mathbf{e}_y|\mathcal{E}_y|\cos(kz - \omega t + \phi_{0y}). \tag{2.27}$$

Although, in general, the electric field is elliptically polarized, there are two important particular cases. The electric field is said to be **linearly** polarized if the phases of

two orthogonal components of the field in Eq. (2.18) are the same, $\phi_{0x} = \phi_{0y}$.
In this case,

$$\mathbf{E}(z,t) = (\mathbf{e}_x|\mathcal{E}_x| + \mathbf{e}_y|\mathcal{E}_y|)\cos(kz - \omega t + \phi_0), \tag{2.28}$$

and the electric field is always directed along the line making the angle

$$\alpha = \tan^{-1}(|\mathcal{E}_y|/|\mathcal{E}_x|) \tag{2.29}$$

with the x-axis as is shown in Fig. 3.7.
If the phases of the two orthogonal components in Eq. (2.19) differ by $\pi/2$, and $|E_{0x}| = |E_{0y}|$, the wave is said to be **circularly** polarized. In this case

$$\mathbf{E}(z,t) = |\mathcal{E}|[\mathbf{e}_x \cos(kz - \omega t + \phi_0) \mp \mathbf{e}_y \sin(kz - \omega t + \phi_0)]. \tag{2.30}$$

In a circularly polarized wave, the $\mathbf{E}$ has the same magnitude but is moving along

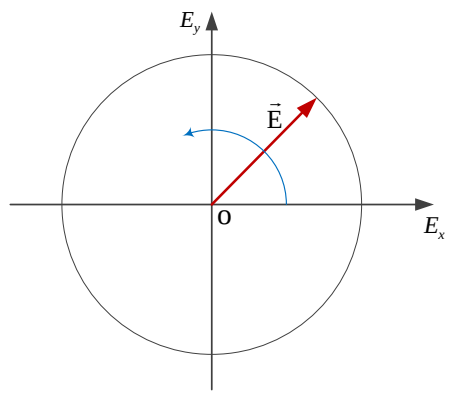

Figure 2.4: Illustrating circular polarization.

the circle. In the case of "-" sign in Eq. (2.30), $\mathbf{E}$ moves counterclockwise around the circle and the wave is **left circularly** polarized; for the "+" sign it is **right circularly** polarized.

2.2 Plane waves in homogeneous dielectrics

We now consider general phenomenological electric constitutive relations for stationary, homogeneous linear media. As a medium can be anisotropic and dispersive, we can introduce the relative permittivity and conductivity tensors, ϵ_{ij} and σ_{ij} and express $\mathbf{D}$ and $\mathbf{J}$ in terms of $\mathbf{E}$ as

$$D_i(\mathbf{r},t) = \epsilon_0 \sum_{j=x,y,z} \int_{-\infty}^{\infty} dt' \int d\mathbf{r}'\, \epsilon_{ij}(\mathbf{r}-\mathbf{r}', t-t') E_j(\mathbf{r}', t'), \tag{2.31}$$

and

$$J_i(\mathbf{r},t) = \sum_{j=x,y,z} \int_{-\infty}^{\infty} dt' \int d\mathbf{r}'\, \sigma_{ij}(\mathbf{r}-\mathbf{r}', t-t') E_j(\mathbf{r}', t'). \tag{2.32}$$

In Eqs. (2.31) and (2.32), the permittivity and conductivity tensors depend only on coordinate and time differences because of homogeneity and stationarity of the medium: All properties of such media are invariant with respect to translations in time and displacements in space.

The translational invariance of the system prompts the use of plane-wave expansions via Fourier transforms, i. e.,

$$\mathbf{D}(\mathbf{r},t) = \int d\omega \int d\mathbf{k}\, \boldsymbol{\mathcal{D}}(\mathbf{k},\omega) e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \tag{2.33}$$

with similar expressions for the other fields. In physical terms, Fourier expansions give all possible plane waves allowed to propagate in such media; the Fourier coefficients specify field amplitudes of these plane waves. Introducing also Fourier expansions of the permittivity and conductivity tensors viz.,

$$\epsilon_{ij}(\mathbf{r},t) = \int d\omega \int d\mathbf{k}\, \epsilon_{ij}(\mathbf{k},\omega) e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \tag{2.34}$$

and

$$\sigma_{ij}(\mathbf{r},t) = \int d\omega \int d\mathbf{k}\, \sigma_{ij}(\mathbf{k},\omega) e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \tag{2.35}$$

we can use convolution properties of Fourier transforms to cast Eqs. (2.31) and (2.32) to

$$\mathcal{D}_i(\mathbf{k},\omega) = \epsilon_0 \sum_{j=x,y,z} \epsilon_{ij}(\mathbf{k},\omega) \mathcal{E}_j(\mathbf{k},\omega), \tag{2.36}$$

and

$$\mathcal{J}_i(\mathbf{k},\omega) = \sum_{j=x,y,z} \sigma_{ij}(\mathbf{k},\omega) \mathcal{E}_j(\mathbf{k},\omega). \tag{2.37}$$

Next, on taking Fourier transforms of Eqs. (1.6), and (1.18) and combining Eqs. (2.36) as well as (2.37), we can establish a relation between the permittivity and conductivity tensors in the Fourier space,

$$\epsilon_{ij}(\mathbf{k},\omega) = \delta_{ij} + \frac{i}{\epsilon_0 \omega} \sigma_{ij}(\mathbf{k},\omega). \tag{2.38}$$

Exercise 2.1. *Derive Eq. (2.38).*

Thus, we conclude that the permittivity and conductivity tensors are actually related and one can be eliminated in favor of the other. In condensed-matter calculations, it is the conductivity tensor that is typically employed. On the other hand, optical wave propagation in the media is more conveniently examined in terms of the permittivity tensor. In the absence of external charges and currents, the Maxwell equations (1.1) through (1.4) can be greatly simplified in the Fourier space to read

$$\mathbf{k} \cdot \boldsymbol{\mathcal{D}} = 0, \tag{2.39}$$

$$\mathbf{k} \cdot \boldsymbol{\mathcal{H}} = 0, \tag{2.40}$$

$$\mathbf{k} \times \boldsymbol{\mathcal{E}} = \mu_0 \omega \boldsymbol{\mathcal{H}}, \tag{2.41}$$

and

$$\mathbf{k} \times \boldsymbol{\mathcal{H}} = -\omega \boldsymbol{\mathcal{D}}. \tag{2.42}$$

Next, eliminating the magnetic field from Eqs. (2.39) - (2.42), and using the constitutive relation (2.36), we can express Eqs. (1.1) and (2.41) in the component form as

$$\sum_{i,j=x,y,z} k_i \epsilon_{ij}(\mathbf{k}, \omega) \mathcal{E}_j(\mathbf{k}, \omega) = 0, \tag{2.43}$$

and

$$\sum_{j=x,y,z} \left[k^2 \delta_{ij} - k_i k_j - \tfrac{\omega^2}{c^2} \epsilon_{ij}(\mathbf{k}, \omega) \right] \mathcal{E}_j(\mathbf{k}, \omega) = 0. \tag{2.44}$$

Eqs. (2.43) and (2.44) determine all possible plane electromagnetic waves supported by a given medium.

Exercise 2.2. *Show that Eqs. (2.43) and (2.44) are always compatible.*

Eq. (2.43 is called a generalized transversality condition, whereas Eq. (2.44) is a dispersion relation for the waves. The existence of nontrivial plane-wave solutions to Eq. (2.44) can be expressed in terms of a determinant condition as

$$\text{Det} \left[k^2 \delta_{ij} - k_i k_j - \tfrac{\omega^2}{c^2} \epsilon_{ij}(\mathbf{k}, \omega) \right] = 0. \tag{2.45}$$

Let us now consider the important limiting case of an isotropic dielectric. It can be inferred by inspection that the dielectric permittivity tensor of an isotropic medium can only be composed of δ_{ij} and $k_i k_j$ implying that

$$\epsilon_{ij}(\mathbf{k}, \omega) = \delta_{ij} A(k, \omega) + k_i k_j B(k, \omega), \tag{2.46}$$

where $A(k, \omega)$ and $B(k, \omega)$ are scalar functions. Instead of using A and B, however, it will prove convenient to divide ϵ_{ij} into a part transverse to the $\mathbf{e}_k = \mathbf{k}/k$ direction, and that longitudinal to $\mathbf{e}_k$. Such a decomposition can be accomplished via

$$\epsilon_{ij}(\mathbf{k}, \omega) = \epsilon_\perp(k, \omega) \left(\delta_{ij} - \tfrac{k_i k_j}{k^2} \right) + \epsilon_\parallel(k, \omega) \tfrac{k_i k_j}{k^2}. \tag{2.47}$$

On substituting from Eq. (2.47) into Eqs. (2.43) and (2.44), the latter can be transformed to

$$\epsilon_\parallel(k, \omega)(\mathbf{k} \cdot \boldsymbol{\mathcal{E}}) = 0, \tag{2.48}$$

and

$$\left[k^2 - \tfrac{\omega^2}{c^2} \epsilon_\perp(k, \omega) \right] \left[\boldsymbol{\mathcal{E}} - \tfrac{\mathbf{k}(\mathbf{k} \cdot \boldsymbol{\mathcal{E}})}{k^2} \right] - \left(\tfrac{\omega^2}{k^2 c^2} \right) \epsilon_\parallel(k, \omega) \mathbf{k}(\mathbf{k} \cdot \boldsymbol{\mathcal{E}}) = 0. \tag{2.49}$$

Eqs. (2.48) and (2.49) then imply the existence of a family of purely transverse plane waves,

$$\mathbf{k} \cdot \boldsymbol{\mathcal{E}} = 0, \tag{2.50}$$

with the dispersion relation,

$$k = \pm\tfrac{\omega}{c}\sqrt{\epsilon_\perp(k,\omega)}, \tag{2.51}$$

and a family of the waves which have longitudinal component(s) of the electric field, $\mathbf{k}\cdot\boldsymbol{\mathcal{E}} \neq 0$, with the dispersion relation determined by a *common* solution of Eq. (2.51) and of the following equation

$$\epsilon_\parallel(k,\omega) = 0. \tag{2.52}$$

In Eq. (2.51) the two signs on the right-hand side correspond to two plane waves at a given frequency ω propagating the the opposite directions.

Exercise. 2.3. *As we will see in Sec. 2.5., dielectric response of metals at high frequencies can be modeled by the permittivity*

$$\epsilon_{ij}(\omega) = \delta_{ij}\left(1 - \frac{\omega_p^2}{\omega^2}\right), \tag{2.53}$$

where ω_p is the so-called plasma frequency. Determine the frequency(s) and dispersion relation of longitudinal electromagnetic waves propagating in metals at such ultraviolet frequencies and interpret your results in physical terms. Show that transverse electromagnetic waves can only propagate if $\omega > \omega_p$. What is their dispersion relation?

Note that the dispersion relation (2.51) is, in general, in the implicit form due to spatial dispersion of the medium. It is then instructive to examine the limiting case of local media which lack spatial dispersion. In reality the vast majority of inorganic media are made of atoms or molecules with the size significantly smaller than the optical wavelength. Hence, spatial nonlocality of their dielectric response to the applied field is negligible, resulting in the absence of spatial dispersion in such media. Under the circumstances, the permittivity tensor can be simplified as

$$\epsilon_{ij}(\mathbf{r}-\mathbf{r}', t-t') = \delta(\mathbf{r}-\mathbf{r}')\epsilon_{ij}(t-t'). \tag{2.54}$$

It then follows at once from Eqs. (2.34) and (2.54) that the permittivity tensor in Fourier space is independent of k, implying that

$$\epsilon(\mathbf{k},\omega) = \epsilon(\mathbf{k}=0,\omega) \equiv \epsilon(\omega). \tag{2.55}$$

The dispersion relation for transverse electromagnetic waves can be expressed in the explicit form as

$$k = \pm\tfrac{\omega}{c}\sqrt{\epsilon_\perp(\omega)}, \tag{2.56}$$

and the generalized transversality condition states

$$\epsilon_\parallel(\omega) = 0. \tag{2.57}$$

In the following sections, we will explore several commonly occurring types of linear optical media.

2.2.1 Plane waves in homogeneous isotropic media with no spatial dispersion

Medium isotropy and locality imply a greatly simplified form of the permittivity tensor,

$$\epsilon_{ij}(\mathbf{k},\omega) = \epsilon(\omega)\delta_{ij}. \tag{2.58}$$

It then follows from Eqs. (2.47) and (2.58) that $\epsilon_{\|}(\omega) = \epsilon_{\perp}(\omega) = \epsilon(\omega)$. Assuming further that in the spectral range of interest, $\epsilon(\omega) \neq 0$, we conclude that in this case, the only allowed plane waves in such media must be transverse, governed by the dispersion relation

$$k = \pm\tfrac{\omega}{c}\sqrt{\epsilon(\omega)}, \tag{2.59}$$

Representing the dielectric function in terms of its real and imaginary parts,

$$\epsilon(\omega) = \epsilon'(\omega) + i\epsilon''(\omega), \tag{2.60}$$

we can express the wave number of the propagating wave as

$$k = \beta_{\pm} + i\alpha_{\pm}/2. \tag{2.61}$$

Here

$$\beta_{\pm} = \pm\tfrac{\omega}{2c}\sqrt{\sqrt{\epsilon'^2 + \epsilon''^2} + \epsilon'}, \tag{2.62}$$

and

$$\alpha_{\pm} = \pm\tfrac{\omega}{c}\sqrt{\sqrt{\epsilon'^2 + \epsilon''^2} - \epsilon'}. \tag{2.63}$$

Exercise 2.4. *Derive the equations (2.62) and (2.63).*

Let us choose the z-axis of our coordinate system along propagation direction of the wave, $\mathbf{k} = k\mathbf{e}_z$. It then follows from the Maxwell equations (2.39) through (2.42) that the electric and magnetic field amplitudes are related as

$$\boldsymbol{\mathcal{E}} = -\eta(\mathbf{e}_z \times \boldsymbol{\mathcal{H}}), \tag{2.64}$$

or, alternatively,

$$\boldsymbol{\mathcal{H}} = \frac{(\mathbf{e}_z \times \boldsymbol{\mathcal{E}})}{\eta}, \tag{2.65}$$

where η is a complex impedance of the lossy medium, defined as

$$\eta(\omega) = \sqrt{\frac{\mu_0}{\epsilon_0\epsilon(\omega)}} = \frac{\eta_0}{\sqrt{\epsilon(\omega)}}. \tag{2.66}$$

To illustrate the plane wave propagation in such a medium, let us focus now on a particular case of a linearly polarized in the x-direction plane wave which propagates in the positive z-direction. The electric and magnetic fields of the wave can then be represented as

$$\mathbf{E}(z,t) = \tfrac{1}{2}\mathbf{e}_x[\mathcal{E}e^{-\alpha_+ z/2}e^{i(\beta_+ z - \omega t)} + c.\,c.,] \tag{2.67}$$

and

$$\mathbf{H}(z,t) = \tfrac{1}{2}\mathbf{e}_y\left[\tfrac{\mathcal{E}}{|\eta|}e^{-\alpha_+ z/2}e^{i(\beta_+ z - \omega t - \theta_\eta)} + c.\,c\right], \tag{2.68}$$

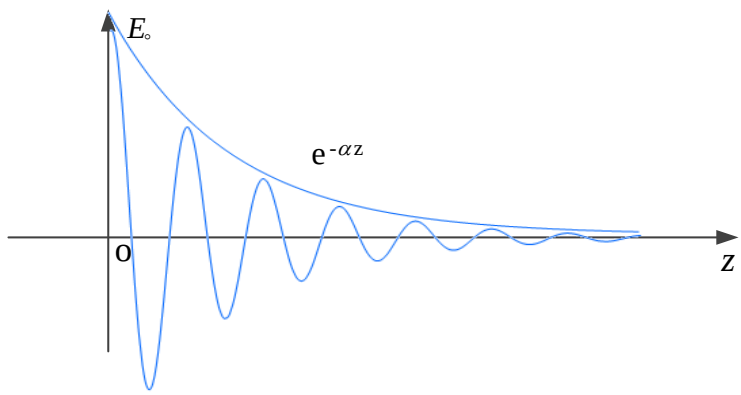

Figure 2.5: Inhomogeneous plane wave propagating in a lossy medium.

which describe inhomogeneous plane waves thanks to losses. Here we chose β_+ and α_+ which describe a plane wave propagating to the right and exponentially decaying on propagation into the medium; the magnitude and phase of the complex impedance can be expressed as

$$|\eta| = \frac{\eta_0}{(\epsilon'^2 + \epsilon''^2)^{1/4}}, \qquad \tan\theta_\eta = -\epsilon''/\epsilon'. \tag{2.69}$$

We can then infer from Eqs. (2.67) and (2.68) that the presence of losses introduces a phase lag between the magnetic and electric fields in such media as well.

Finally, we can work out the time-averaged energy flux density (Poynting vector), and hence the optical intensity, associated with the inhomogeneous plane wave. On substituting from Eqs. (2.67) and (2.68) into Eq. (1.26), we obtain for the optical intensity

$$I = |\langle S \rangle| = \frac{|\mathcal{E}|^2}{2|\eta|} e^{-z/\delta} \cos\theta_\eta. \tag{2.70}$$

Eq. (2.70) is known as Beer's absorption law, and by measuring the intensity extinction, one can infer the Beer absorption length, or skin depth

$$\delta = \frac{1}{\alpha_+}. \tag{2.71}$$

We note that Beer's absorption length is then a directly measurable quantity. We can also define a complex refractive index by the expression

$$\mathcal{N}(\omega) = \sqrt{\epsilon(\omega)} = n(\omega) + i\kappa(\omega), \tag{2.72}$$

where n is a real refractive index which can be determined from reflectivity measurements and κ is a so-called extinction coefficient, closely related to Beers' absorption length. In fact, it readily follows from Eqs. (2.59), (2.61) and (2.72) that

$$\delta^{-1}(\omega) = \frac{2\kappa(\omega)\omega}{c}. \tag{2.73}$$

The magnitudes of real and imaginary parts of ϵ can then be inferred from the knowledge of n and κ, i.e.,

$$\epsilon' = n^2 - \kappa^2, \qquad \epsilon'' = 2n\kappa. \tag{2.74}$$

In particular, in the transparent regions of the spectrum, where $\epsilon'' \ll \epsilon'$, $\epsilon' \simeq n^2$ and the optical intensity of a plane wave can be expressed as

$$I = \frac{\epsilon_0 n c}{2}|\mathcal{E}|^2. \tag{2.75}$$

2.2.2 Plane waves in uniaxial crystals

We will now explore the families of plane waves that can propagate in transparent dispersionless anisotropic media. Most crystals fall into this category in the optical frequency range. We will limit ourselves to the case of uniaxial crystals. Dielectric properties of uniaxial crystals along a special axis, usually defined by a unit vector $\mathbf{n}$, are different from those in any direction orthogonal to the axis. The special direction is called an optical axis of the crystal. In the absence of spatial dispersion, the dielectric permittivity tensor can only depend on δ_{ij} and $n_i n_j$ and can be conveniently expressed in terms of transverse $\epsilon_\perp$ and longitudinal $\epsilon_\parallel$ components as

$$\epsilon_{ij} = \epsilon_\perp(\delta_{ij} - n_i n_j) + \epsilon_\parallel n_i n_j. \tag{2.76}$$

One can always choose a coordinate systems such that the optical axis of the crystal coincides with one of the axes, the z-axis, say. In these coordinates, the dielectric tensor transforms to its canonical (diagonal) form represented by the matrix

$$\epsilon_{ij} = \begin{pmatrix} \epsilon_\perp & 0 & 0 \\ 0 & \epsilon_\perp & 0 \\ 0 & 0 & \epsilon_\parallel \end{pmatrix} \tag{2.77}$$

If $\epsilon_\parallel > \epsilon_\perp$, the crystal is said to be a positive uniaxial crystal, and if $\epsilon_\parallel < \epsilon_\perp$ the crystal is referred to as a negative uniaxial one.

Let us assume, for simplicity that the wave vector lies in the xz-plane, $\mathbf{k} = k_x\mathbf{e}_x + k_z\mathbf{e}_z$. It then follows from Eqs. (2.44) and Eq. (2.77) that

$$\left(k_z^2 - \tfrac{\omega^2}{c^2}\epsilon_\perp\right)\mathcal{E}_x - k_x k_z \mathcal{E}_z = 0, \tag{2.78}$$

$$-k_x k_z \mathcal{E}_x + \left(k_x^2 - \tfrac{\omega^2}{c^2}\epsilon_\parallel\right)\mathcal{E}_z = 0, \tag{2.79}$$

and

$$\left(k^2 - \tfrac{\omega^2}{c^2}\epsilon_\perp\right)\mathcal{E}_y = 0. \tag{2.80}$$

The generalized transversality condition (2.43) can then be cast into the form

$$k_x\epsilon_\perp\mathcal{E}_x + k_z\epsilon_\parallel\mathcal{E}_z = 0. \tag{2.81}$$

The analysis of Eqs. (2.78) through (2.81) reveals that there are two possible polarizations: ordinary and extraordinary one. For the ordinary polarization, it follows at once from Eq. (2.80) that the ordinarily polarized wave is transverse,

$$\boldsymbol{\mathcal{E}} = \mathcal{E}_y\mathbf{e}_y, \tag{2.82}$$

and ita dispersion relation is given by the expression

$$k_{\rm o} = \tfrac{\omega}{c}\sqrt{\epsilon_\perp}. \tag{2.83}$$

We observe that ordinary waves in uniaxial crystals have all the same properties as plane waves supported by transparent isotropic media.

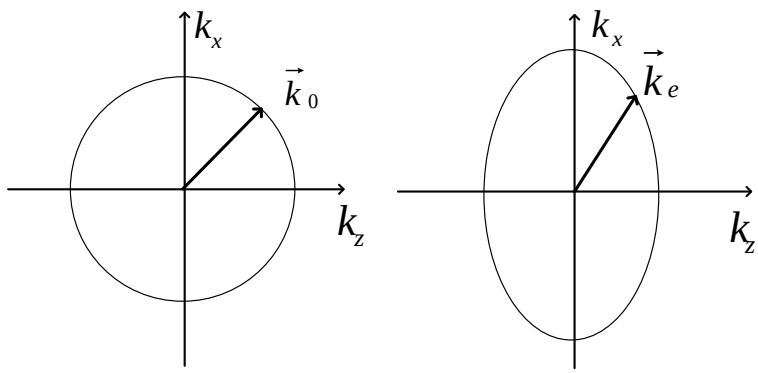

Figure 2.6: Graphical representation of the wave vectors of ordinary (left) and extraordinary (right) waves in a uniaxial crystal.

The polarization of the extraordinary waves can be inferred from

$$\boldsymbol{\mathcal{E}} = \mathcal{E}_x \mathbf{e}_x + \mathcal{E}_z \mathbf{e}_z, \tag{2.84}$$

where $\mathcal{E}_x$ and $\mathcal{E}_z$ are related by Eq. (2.81). We can also derive their dispersion relation from the determinant condition for Eqs. (2.78) and (2.79). The resulting dispersion relation reads

$$\frac{k_x^2 c^2}{\omega^2 \epsilon_\parallel} + \frac{k_z^2 c^2}{\omega^2 \epsilon_\perp} = 1. \tag{2.85}$$

Using $k_x = k_e \sin\theta$, and $k_z = k_{\rm e}\cos\theta$, we can cast Eq. (2.85) into the form

$$\frac{\omega^2}{k_{\rm e}^2 c^2} = \frac{\sin^2\theta}{\epsilon_\parallel} + \frac{\cos^2\theta}{\epsilon_\perp}. \tag{2.86}$$

Thus the wave vector magnitude of an extraordinary wave depends on its propagation direction which is a novel propagation feature arising in anisotropic media. The difference between ordinary and extraordinary waves can be best visualized by comparing their dispersion relations. It is seen from Eqs. (2.83) and (2.86) that in the k-plane the dispersion relations of ordinary and extraordinary waves can be represented by a sphere of radius $(\omega/c)\sqrt{\epsilon_\perp}$ and ellipse with the semi-axes $(\omega/c)\sqrt{\epsilon_\perp}$ and $(\omega/c)\sqrt{\epsilon_\parallel}$, respectively. The situation is schematically depicted in the figure above.

Exercise 2.5. *Using Maxwell's equations show that the wave vector of the extraordinary wave is not parallel to the Pointing vector,* $\mathbf{S} = \mathbf{E} \times \mathbf{H}$*. In other words, demonstrate that the direction of propagation of such a wave does not, in general, coincide with the direction of the energy flow.*

2.2.3 Faraday effect and polarization rotation

We will now consider light propagation in an isotropic, weakly dispersive–and hence lossless–dielectric medium with a weak homogeneous static magnetic field, $\mathbf{B}$, applied along the $z-$axis such that $\mathbf{B} = B\mathbf{e}_z$. We assume that the influence of magnetic field can be treated as a perturbation and we seek a phenomenological expression for a dielectric permittivity tensor of an isotropic medium with a small correction due to the magnetic field. The lowest-order correction is assumed to linear in the magnetic field. Therefore, the second-order permittivity tensor can only be comprised of δ_{ij} and a component linear in B_i. Recall that both $\mathbf{D}$ and $\mathbf{E}$ are physical vectors that change their sign upon reflections with respect to the origin of a coordinate system. It then follows from Eq. (2.31) that ϵ_{ij} should be invariant upon reflections. To respect the reflectional invariance of the permittivity tensor, the correction term can only be of the form $\sum_k e_{ijk}B_k$, where

$$e_{ijk} = \begin{cases} 1, & \text{clockwise permutation} \\ -1, & \text{counterclockwise permutation} \end{cases} \tag{2.87}$$

is an antisymmetric Levi-Chivita symbol; $e_{xyz} = 1$, $e_{yxz} = -1$ and so on up to a cyclic permutation. Thus, on phenomenological grounds, the dielectric permittivity tensor describing an isotropic dispersionless medium perturbed by a weak magnetic field can be written as

$$\epsilon_{ij}(\omega) = \epsilon(\omega)\delta_{ij} + ig(\omega)\sum_{k=x,y,z} e_{ijk}B_k, \qquad |gB| \ll \epsilon. \tag{2.88}$$

where $g(\omega)$ is a phenomenological constant. In the end of this chapter, we will derive Eq. (2.88) using a simple classical microscopic model of a medium. The permittivity tensor (2.88) can be written in a matrix form as

$$\epsilon_{ij}(\omega) = \begin{pmatrix} \epsilon(\omega) & ig(\omega)B & 0 \\ -ig(\omega)B & \epsilon(\omega) & 0 \\ 0 & 0 & \epsilon \end{pmatrix}. \tag{2.89}$$

Let us now assume, for simplicity that the wave propagates along the magnetic field, $\mathbf{k} = k\mathbf{e}_z$. It then follows from Eqs. (2.44) and (2.43) that

$$\left[k^2 - \tfrac{\omega^2}{c^2}\epsilon(\omega)\right]\mathcal{E}_x - ig(\omega)B\left(\tfrac{\omega^2}{c^2}\right)\mathcal{E}_y = 0, \tag{2.90}$$

$$ig(\omega)B\left(\tfrac{\omega^2}{c^2}\right)\mathcal{E}_x + \left[k^2 - \tfrac{\omega^2}{c^2}\epsilon(\omega)\right]\mathcal{E}_y = 0, \tag{2.91}$$

and

$$-\tfrac{\omega^2}{c^2}\mathcal{E}_z = 0. \tag{2.92}$$

Exercise 2.6. *Derive Eqs. (2.90) – (2.92) from Eqs. (2.44), (2.43), and (2.88).*
We can then show that up to the first order in $fB/\epsilon \ll 1$, wave vector magnitude of any wave existing in such media is given by

$$k_\pm = k \pm \Delta k, \tag{2.93}$$

where we introduced the notations

$$k = \frac{\omega}{c}\sqrt{\epsilon(\omega)}; \qquad \Delta k = \frac{\omega g(\omega) B}{2c\sqrt{\epsilon(\omega)}}. \tag{2.94}$$

The plane waves supported by the media must be circularly polarized, i. e.,

$$\mathcal{E}_z = 0, \qquad \mathcal{E}_y = \pm i\mathcal{E}_x, \tag{2.95}$$

where the upper (lower) sign on the right-hand side of Eq. (2.95) corresponds to the upper (lower) subscript on the left-hand side of Eq. (2.93). In other words, the medium supports left- and right-circularly polarized waves with slightly different wave numbers.

We will now explore how a linearly polarized wave evolves in the medium. Assuming the wave is polarized along the x-axis, say, at the entrance to the medium,

$$\mathbf{E}_0 = \tfrac{1}{2}\mathbf{e}_x\mathcal{E}_0 e^{-i\omega t} + c.c, \tag{2.96}$$

we can represent the incident electric field as

$$\mathbf{E}_0 = \tfrac{1}{2}\tfrac{\mathcal{E}_0}{\sqrt{2}}(\mathbf{e}_+ + \mathbf{e}_-)e^{-i\omega t} + c.c, \tag{2.97}$$

where

$$\mathbf{e}_\pm = \frac{\mathbf{e}_x \pm i\mathbf{e}_y}{\sqrt{2}}, \tag{2.98}$$

are the unit vectors associated with the two circular polarizations. We can now examine wave propagation in the medium. The electric field in any transverse plane $z = const$ can be written as

$$\mathbf{E} = (A\mathbf{e}_+ e^{ik_+ z} + B\mathbf{e}_- e^{ik_- z})e^{-i\omega t} + c.c. \tag{2.99}$$

It follows from the initial conditions that $A = B = \mathcal{E}_0/2\sqrt{2}$. Thus, we obtain subsequently the propagated wave expression in the form

$$\mathbf{E} = \tfrac{1}{2}\tfrac{\mathcal{E}_0}{\sqrt{2}}(\mathbf{e}_+ e^{i\Delta kz} + \mathbf{e}_- e^{-i\Delta kz})e^{i(kz-\omega t)} + c.c. \tag{2.100}$$

We can transform Eq. (2.100) to

$$\mathbf{E} = \tfrac{1}{2}\mathcal{E}_0\mathbf{e}_p(z)e^{i(kz-\omega t)} + c.c, \tag{2.101}$$

where

$$\mathbf{e}_p(z) = \mathbf{e}_x \cos\Delta kz - \mathbf{e}_y \sin\Delta kz. \tag{2.102}$$

Exercise 2.7. *Derive Eq. (2.101) from Eq. (2.100)*

It can be inferred from Eqs. (2.101) and (2.102) that the wave remains linearly polarized, but the plane of polarization rotates. Alternatively, one can conclude that the polarization vector rotates in the transverse plane as the wave propagates along the z-axis. This phenomenon is called Faraday rotation. The rate of rotation is customary characterized by the Verdet constant V defined by the expression

$$\Delta k = VB, \tag{2.103}$$

It then follows at once from Eqs. (2.94) and (2.103) that for a plane wave propagating along the magnetic field, the Verdet constant is given by

$$V = \frac{\omega g(\omega)}{2c\sqrt{\epsilon(\omega)}}. \tag{2.104}$$

Exercise 2.8. *Generalize the discussion of this section to the case when a plane wave propagates at an angle θ to the magnetic field. Derive a generalized dispersion relation and determine the Verdet constant in this case.*

2.3 Refraction and reflection of plane waves at the interface of homogeneous media

2.3.1 Reflection of plane waves at oblique incidence: Generalized Snell's law

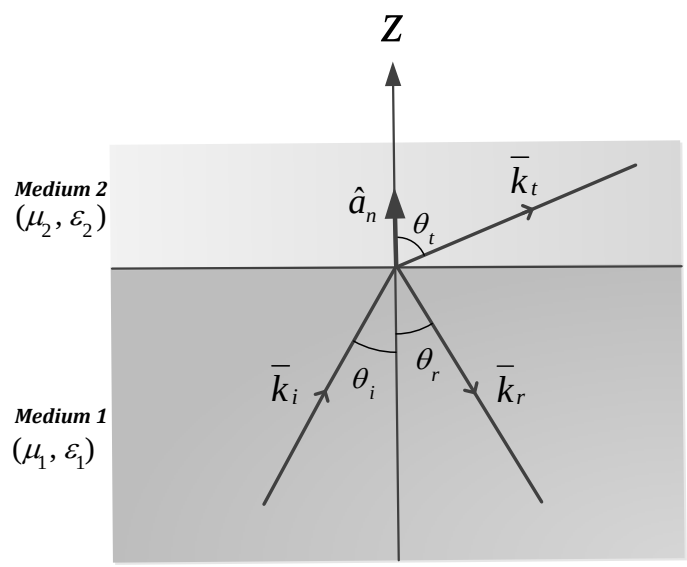

Figure 2.7: Illustrating Snell's law for oblique incidence of a plane wave.

We now explore refraction and reflection of plane electromagnetic waves at an interface of two homogeneous media. To reflect a typical physical situation, we will assume that a plane wave is incident from a transparent medium with the permittivity ϵ_1 onto a flat interface separating the medium from a lossy medium 2, characterized by the complex permittivity $\epsilon_2(\omega) = \epsilon_2'(\omega) + i\epsilon_2''(\omega)$. We choose a coordinate system with the

unit normal to the interface pointed along the z-axis. The geometry of the problem is sketched in Fig. 2.3.1. Note that the incidence, refraction, and transmission angles, θ_i, θ_r, and θ_t, respectively, are real angles only in the transparency window of the second medium, $\epsilon_2''(\omega) = 0$. Otherwise, all sines and cosines of θ_t are complex. For this reason, we will not use the angles hereafter. Rather, we will derive the Fresnel formulas for transmission and reflection amplitudes in terms of the corresponding projections of k-vectors which can, in general, be complex. Next, it will prove convenient hereafter to introduce the notations

$$k_1 = k_0 n_1, \qquad k_2 = k_0 \mathcal{N}_2, \tag{2.105}$$

where $k_0 = \omega/c$ and $\mathcal{N}_2$ is a complex refractive index of medium 2, c. f. Eq. (2.72). The boundary conditions at the flat interface $z = 0$ should hold at any point in the xz-plane and at any instant of time t, implying that

$$e^{i(\mathbf{k}_i \cdot \mathbf{r} - \omega_i t)}|_{z=0} = e^{i(\mathbf{k}_r \cdot \mathbf{r} - \omega_r t)}|_{z=0} = e^{i(\mathbf{k}_t \cdot \mathbf{r} - \omega_t t)}|_{z=0}. \tag{2.106}$$

Here the subscripts i, r, and t stand for incident, reflected and transmitted waves, respectively. It follows at once from Eq. (2.106) that

$$\omega_i = \omega_r = \omega_t = \omega, \tag{2.107}$$

that is the frequencies of the incident, reflected and transmitted waves must match. Further, it can be inferred from the boundary conditions (2.106) that

$$k_{ix} = k_{rx} = k_{tx} = k_x, \tag{2.108}$$

In other words, the in-plane components of the wave vectors must match as well. Notice that since medium 1 is assumed to be transparent, Eq. (2.108) stipulates that in-plane components of the wave vectors of all the involved waves be real. We stress that Eq. (2.108) is a generalized Snell's law. We shall also introduce the notations

$$k_{iz} = -k_{ir} = k_{1z}, \qquad k_{tz} = k_{2z}. \tag{2.109}$$

It then follows from Eqs. (2.105) (2.108), and (3.30) that

$$\mathbf{k}_i = k_x \mathbf{e}_x + k_{1z} \mathbf{e}_z, \tag{2.110}$$

$$\mathbf{k}_r = k_x \mathbf{e}_x - k_{1z} \mathbf{e}_z, \tag{2.111}$$

and

$$\mathbf{k}_t = k_x \mathbf{e}_x + k_{2z} \mathbf{e}_z, \tag{2.112}$$

where

$$k_{1z} = \sqrt{k_1^2 - k_x^2}, \quad \text{and} \quad k_{2z} = \sqrt{k_2^2 - k_x^2}. \tag{2.113}$$

It is easy to see from the geometry of Fig. 2.3.1 that in the transparency window of medium 2, all angles are real and Eq. (2.108 reduces to

$$\theta_i = \theta_r \equiv \theta_1, \qquad \theta_t \equiv \theta_2 \tag{2.114}$$

and

$$n_1 \sin\theta_1 = n_2 \sin\theta_2. \tag{2.115}$$

Put another way, the incidence and reflection angles should be equal and the Snell law should simplify to its familiar form for refraction at the interface of two transparent media.

2.3.2 Reflection of plane waves at oblique incidence: Fresnel Formulae

There are two important special cases of the incident polarization that should be distinguished: transverse magnetic (TM), or p-polarization, and transverse electric (TE), or s-polarization. In the first instance, the magnetic field of an incident wave is directed perpendicular to the plane of incidence, whereas in the second case it is the incident electric field that is orthogonal to this plane. We will examine the two cases separately. Note that an arbitrarily polarized incident field can be decomposed into a TM and TE polarized components which are mutually orthogonal.

Transverse magnetic (TM) or p-polarization. – Consider first the TM case. Magnetic fields of the incident, reflected and transmitted TM waves are assumed to be polarized along the y-axis, such that we can express their complex amplitudes as

$$\boldsymbol{\mathcal{H}}_s = H_s \mathbf{e}_y, \qquad s = i, r, t. \tag{2.116}$$

Since the magnetic field of a TM-polarized wave has only one component, it is convenient to express the electric field in terms of the magnetic one. It follows at once from the Maxwell equations (2.41) and (2.42) that

$$\boldsymbol{\mathcal{E}}_s = -\eta_s(\mathbf{e}_{ks} \times \boldsymbol{\mathcal{H}}_s), \tag{2.117}$$

where $\eta_{i,r} = \sqrt{\mu_0/\epsilon_1}$ and $\eta_t = \sqrt{\mu_0/\epsilon_2}$ are relevant media impedances. We can then infer from Eqs. (2.116) and (2.117) as well as Eqs. (2.109) through (2.113) that the complex amplitudes of the incident, reflected, and transmitted fields can be represented as

$$\begin{aligned} \boldsymbol{\mathcal{H}}_i &= H_i \mathbf{e}_y, \\ \boldsymbol{\mathcal{E}}_i &= \frac{\eta_0 H_i}{k_0 \epsilon_1}(k_{1z}\mathbf{e}_x - k_x \mathbf{e}_z), \end{aligned} \tag{2.118}$$

$$\begin{aligned} \boldsymbol{\mathcal{H}}_r &= H_r \mathbf{e}_y, \\ \boldsymbol{\mathcal{E}}_r &= \frac{\eta_0 H_r}{k_0 \epsilon_1}(-k_{1z}\mathbf{e}_x - k_x \mathbf{e}_z), \end{aligned} \tag{2.119}$$

and

$$\begin{aligned} \boldsymbol{\mathcal{H}}_t &= H_t \mathbf{e}_y, \\ \boldsymbol{\mathcal{E}}_t &= \frac{\eta_0 H_t}{k_0 \epsilon_2}(k_{2z}\mathbf{e}_x - k_x \mathbf{e}_z), \end{aligned} \tag{2.120}$$

respectively.

The boundary conditions for the tangential components of the fields across the interface state

$$H_i + H_r = H_t \tag{2.121}$$

and

$$\frac{H_i}{\epsilon_1} k_{1z} - \frac{H_r}{\epsilon_1} k_{1z} = \frac{H_t}{\epsilon_2} k_{2z}. \tag{2.122}$$

It then follows from Eqs. (2.121) and (2.122) that

$$H_r = \frac{\epsilon_2 k_{1z} - \epsilon_1 k_{2z}}{\epsilon_2 k_{1z} + \epsilon_1 k_{2z}} H_i, \tag{2.123}$$

and

$$H_t = \frac{2\epsilon_2 k_{1z}}{\epsilon_2 k_{1z} + \epsilon_1 k_{2z}} H_i. \tag{2.124}$$

Using (2.117) we arrive at the expressions for the electric fields in the form

$$E_i = \eta_1 H_i, \qquad E_r = \eta_1 H_r, \qquad E_t = \eta_2 H_t. \tag{2.125}$$

Finally, the complex reflectivity and transmittance can be represented as

$$r_{\mathrm{p}} \equiv \frac{E_r}{E_i} = \frac{\epsilon_2 k_{1z} - \epsilon_1 k_{2z}}{\epsilon_2 k_{1z} + \epsilon_1 k_{2z}}, \tag{2.126}$$

and

$$t_{\mathrm{p}} \equiv \frac{E_t}{E_i} = \frac{2\epsilon_2 k_{1z}}{\epsilon_2 k_{1z} + \epsilon_1 k_{2z}} \sqrt{\frac{\epsilon_1}{\epsilon_2}}. \tag{2.127}$$

Equations (2.126) and (2.127) are the celebrated Fresnel formulas for the TM case.

Let us now focus on the situation when the wave is incident normally to the interface, such that $k_x = 0$, $k_{sz} = k_s$, $s = 1, 2$. It then follows from Eqs. (2.105), (2.126) and (2.127) that

$$r_\perp = \frac{\mathcal{N}_2 - n_1}{\mathcal{N}_2 + n_1}, \qquad \text{and} \qquad t_\perp = \frac{2n_1}{\mathcal{N}_2 + n_1}. \tag{2.128}$$

There are two instructive limiting cases here. First, the second medium is transparent, $\mathcal{N}_2 = n_2$, such that the reflectivity and transmittance are purely real,

$$r_\perp = \frac{n_2 - n_1}{n_2 + n_1}, \qquad \text{and} \qquad t_\perp = \frac{2n_1}{n_2 + n_1}, \tag{2.129}$$

and the latter relations simply quantify the relative amplitudes of the reflected and transmitted waves. Note that no energy will be lost in transmission in this case.

Exercise 2.9. *A plane wave is normally incident at an interface separating two transparent media. Show that the electromagnetic energy fluxes on both sides of the interface are the same.*

Another interesting situation arises when medium 2 behaves as a good conductor in a certain spectral range. As is seen from Eqs. (2.38), 2.60), (2.72), and (2.74), $\kappa_2 \gg$

$\max(n_1, n_2)$ in this case. Thus, the reflectivity and transmittance may be approximated as

$$r_\perp \simeq 1 - \frac{2in_2}{\kappa_2} \quad \text{and} \quad t_\perp \simeq -\frac{2in_1}{\kappa_2}. \tag{2.130}$$

It follows that most of the incident wave power is reflected from the interface of a good conductor; only is its tiny fraction transmitted into the conductor.

Exercise 2.10. *Consider a plane wave incident normally at the interface separating air from a good conductor. Determine the portion of the incident wave power absorbed by the conductor.*

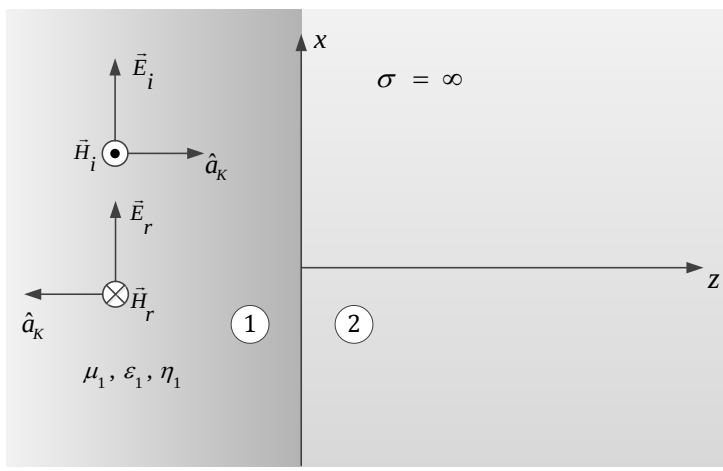

Figure 2.8: Normal incidence of a plane wave onto an interface separating a dielectric and a perfect conductor.

In the extreme case of a perfect conductor, $\kappa_2 \to \infty$, such that $r_\perp \to 1$–the wave is perfectly reflected from the interface. The situation is sketched in the Fig. 2.3.2. The electric and magnetic fields of the incident and reflected waves can then be represented as

$$\mathbf{E}_i(z,t) = \mathbf{e}_x \eta_1 H_i e^{i(k_1 z - \omega t)}, \tag{2.131}$$

$$\mathbf{H}_i(z,t) = \mathbf{e}_y H_i e^{i(k_1 z - \omega t)}. \tag{2.132}$$

and

$$\mathbf{E}_r(z,t) = -\mathbf{e}_x \eta_1 H_i e^{-i(k_1 z + \omega t)}, \tag{2.133}$$

$$\mathbf{H}_r(z,t) = \mathbf{e}_y H_i e^{-i(k_1 z + \omega t)}, \tag{2.134}$$

respectively. The total electric and magnetic fields in medium 1 can then be transformed to

$$\mathbf{E}_1 = \mathrm{Re}(\mathbf{E}_i + \mathbf{E}_r) = 2\mathbf{e}_x \eta_1 |H_i| \sin k_1 z \sin \omega t, \tag{2.135}$$

and

$$\mathbf{H}_1 = \mathrm{Re}(\mathbf{H}_i + \mathbf{H}_r) = 2\mathbf{e}_y |H_i| \cos k_1 z \cos \omega t. \tag{2.136}$$

These equations describe standing waves carrying no energy which conforms to our intuitive picture for reflection from a perfect conductor: The counterpropagating incident and reflected waves of equal amplitudes interfere to form a standing wave pattern in medium 1.

Exercise 2.11. *A right-hand circularly polarized wave, propagating in the positive z-direction is normally incident on a perfect conductor wall $z = 0$. Determine (a) the polarization of the reflected wave and (b) the induced current on the conducting wall.*

Transverse electric (TE) or s-polarization. – In the TE case, the electric field is normal to the incidence plane,

$$\boldsymbol{\mathcal{E}}_i = E_i \mathbf{e}_y, \tag{2.137}$$

and it is convenient to work with complex amplitudes of electric fields, expressing the the magnetic field amplitudes as

$$\boldsymbol{\mathcal{H}}_s = \frac{(\mathbf{e}_s \times \boldsymbol{\mathcal{E}}_s)}{\eta_s}; \qquad s = i, r, t. \tag{2.138}$$

Similarly to the p-polarization case, we can obtain the expressions

$$\boldsymbol{\mathcal{E}}_i = E_i \mathbf{e}_y$$
$$\boldsymbol{\mathcal{H}}_i = \frac{E_i}{\eta_0 k_0}(-k_{1z}\mathbf{e}_x + k_x \mathbf{e}_z), \tag{2.139}$$

$$\boldsymbol{\mathcal{E}}_r = E_r \mathbf{e}_y,$$
$$\boldsymbol{\mathcal{H}}_r = \frac{E_r}{\eta_0 k_0}(k_{1z}\mathbf{e}_x + k_x \mathbf{e}_z), \tag{2.140}$$

and

$$\boldsymbol{\mathcal{E}}_t = E_t \mathbf{e}_y,$$
$$\boldsymbol{\mathcal{H}}_t = \frac{E_t}{\eta_0 k_0}(-k_{2z}\mathbf{e}_x + k_x \mathbf{e}_z), \tag{2.141}$$

for the complex amplitudes of incident, reflected, and transmitted fields, respectively. The continuity of tangential components of electric and magnetic fields across the interface leads to

$$E_i + E_r = E_t, \tag{2.142}$$

and

$$(-E_i + E_r)k_{1z} = -E_t k_{2z} \tag{2.143}$$

Solving the last pair of equations, we arrive at the complex reflectivity and transmittance of an s-polarized incident wave in the form

$$r_{\mathrm{s}} \equiv \frac{E_r}{E_i} = \frac{k_{1z} - k_{2z}}{k_{1z} + k_{2z}}, \tag{2.144}$$

and

$$t_{\mathrm{s}} \equiv \frac{E_t}{E_i} = \frac{2k_{1z}}{k_{1z} + k_{2z}}. \tag{2.145}$$

2.3.3 Brewster angle and surface plasmon polaritons

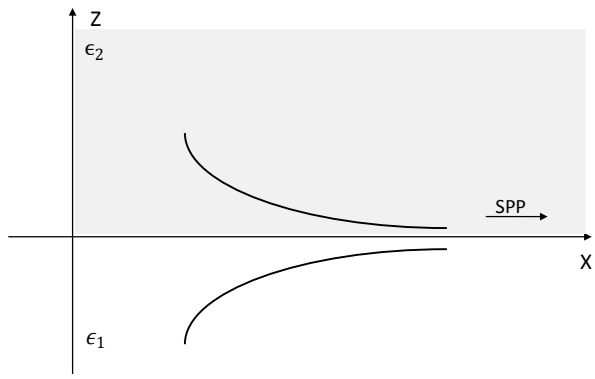

Figure 2.9: Surface electromagnetic wave (surface plasmon polariton) at a metal-dielectric interface. The electric and magnetic field decay fast away from the interface

Let us return to the general case of p-polarized wave reflection form the interface and study the behavior of reflectivity in more detail. We will assume both media to be transparent, for simplicity. It can be inferred from Eq. (2.126) that the reflectivity attains zero under the condition

$$\epsilon_2 k_{1z} = \epsilon_1 k_{2z}. \tag{2.146}$$

Solving Eq. (2.146), together with (2.113), we obtain expressions for the in-plane and normal components of the wave vectors as

$$k_x = \frac{\omega}{c}\sqrt{\frac{\epsilon_1\epsilon_2}{\epsilon_1+\epsilon_2}}, \tag{2.147}$$

and

$$k_{jz} = \frac{\omega}{c}\sqrt{\frac{\epsilon_j^2}{\epsilon_1+\epsilon_2}}, \qquad j = 1, 2. \tag{2.148}$$

The analysis of Eqs. (2.147) and (2.148) reveals two options. First, if both media permittivities are positive, $\epsilon_j > 0$, we may introduce real refractive indices, $n_j = \sqrt{\epsilon_j}$. It then follows at once from Eqs. (2.147) and (2.148) that there exists a special incidence angle θ_{B}, given by the expression

$$\tan\theta_{\mathrm{B}} = k_x/k_{1z} = n_2/n_1, \tag{2.149}$$

such that there is no p-polarized reflected wave. This special incidence angle is known as the Brewster angle. Alternatively, Eqs (2.147) and (2.148) describe a surface wave propagating along the interface, $k_x^2 > 0$ and exponentially decaying in the direction normal to the interface such that k_{jz} is purely imaginary (for transparent media), $k_{jz}^2 < 0$, Eqs. (2.147) and (2.148) imply that this is possible under the conditions,

$$\epsilon_1 + \epsilon_2 < 0, \tag{2.150}$$

and

$$\epsilon_1\epsilon_2 < 0. \tag{2.151}$$

In other words, at least one of the permittivities must be negative. Usually, the wave is incident form a dielectric medium, $\epsilon_1 > 0$, implying that $\epsilon_2 < 0$. The latter condition can be realized for metals as we will see in Sec. 5.

These surface electromagnetic waves are known as surface plasmon polaritons (SPP). Using Eqs. (2.118), (2.120), and (2.148), the electromagnetic fields of SPPs on each side of the interface can be expressed as

$$\mathbf{H}(\mathbf{r},t) = \begin{cases} \mathbf{e}_y H_i e^{-|k_{1z}|z} e^{i(k_x x - \omega t)}, & z > 0; \\ \mathbf{e}_y H_i e^{|k_{2z}|z} e^{i(k_x x - \omega t)}, & z < 0, \end{cases} \tag{2.152}$$

and

$$\mathbf{E}(\mathbf{r},t) = \begin{cases} \frac{\eta_0 H_i}{k_0 \epsilon_1}(i|k_{1z}|\mathbf{e}_x - k_x \mathbf{e}_z) e^{-|k_{1z}|z} e^{i(k_x x - \omega t)}, & z > 0, \\ \frac{\eta_0 H_i}{k_0 \epsilon_2}(-i|k_{2z}|\mathbf{e}_x - k_x \mathbf{e}_z) e^{|k_{2z}|z} e^{i(k_x x - \omega t)}, & z < 0. \end{cases} \tag{2.153}$$

Thus, SPP fields propagate along the interface and exponentially decay away from the interface which is a characteristic signature of surface electromagnetic waves. In case the second medium is an ideal metal, its permittivity can be successfully modeled by the expression

$$\epsilon_2(\omega) = 1 - \frac{\omega_p^2}{\omega^2}, \tag{2.154}$$

where ω_p is the so-called plasma frequency. It can be seen from Eqs. (2.147) that in the short wavelength approximation, $k_x \to \infty$, the SPP frequency tends to a constant value, ω_∞ given by the expression

$$\omega_\infty = \frac{\omega_p}{\sqrt{1+\epsilon_1}}. \tag{2.155}$$

In this case, the SPP approaches its quasi-static limit termed a surface plasmon (SP).

Exercise 2.12. *Show that Eq. (2.155) can be derived in the quasi-static limit by solving Laplace's equation for the electrostatic potential and matching the appropriate boundary conditions.*

So far, we have assumed that the SPPs propagate on the interface of two transparent media. In reality, of course, all metals are lossy, albeit losses are usually small at optical frequencies. Realistic metals can then be properly described by complex dielectric permittivities to account for Joules' losses. Introducing a complex permittivity of medium 2 viz.,

$$\epsilon_2 = \epsilon_2' + i\epsilon_2'', \tag{2.156}$$

and assuming that under at optical frequencies of interest $|\epsilon_2''| \ll |\epsilon_2'|$, we can express the in-plane component of the SPP wave vector as

$$k_x = k_x' + ik_x'', \tag{2.157}$$

where

$$k_x' \simeq k_0\sqrt{\frac{\epsilon_1\epsilon_2'}{\epsilon_1+\epsilon_2'}}, \tag{2.158}$$

and

$$k_x'' \simeq k_0 \sqrt{\frac{\epsilon_1 \epsilon_2'}{\epsilon_1 + \epsilon_2'}} \left[\frac{\epsilon_2'' \epsilon_1}{2\epsilon_2'(\epsilon_1 + \epsilon_2')} \right]. \tag{2.159}$$

Here the imaginary part specifies a characteristic inverse damping distance of the SPP,

$$L_{SPP} = 1/k_x''. \tag{2.160}$$

Exercise 2.13. *Derive Eqs. (2.158) and (2.159).*

It follows from Eq. (2.147) that a plane wave in the air with $\epsilon_1 = 1$ can never excite a plasmon because of the wave vector mismatch: the plasmon wave vector component along the interface is always greater than that of a plane wave in the air. One way to generate an SPP then will be to nano-engineer the surface by creating periodic imperfections such as grooves. The modified surface can serve as a diffraction grating by shifting the in-plane wave vector component of the incident wave to achieve phase matching. Introducing the lattice constant of the grooves a and assuming that the light wave is incident from air, we can write down the matching condition

$$k_{xSPP} = k_0 \sin\theta_i + 2\pi/a, \tag{2.161}$$

where θ_i is the incidence angle. This excitation scheme is sketched in the figure.

2.3.4 Total internal reflection

We saw in the previous section that a TM-polarized surface electromagnetic wave can be excited at an interface of a metal and transparent dielectric. In this section, we show that surface wave generation is also possible at an interface of two transparent media with refractive indices n_1 and n_2, when light is incident form a more optically dense medium, $n_1 > n_2$. This phenomenon is referred to as total internal reflection. It follows from Eq. (2.113) and the geometry of Fig. 2.3.1 that

$$k_{2z}^2 = k_0^2 (n_2^2 - n_1^2 \sin^2\theta_1). \tag{2.162}$$

It can be readily inferred from Eq. (2.162) that the in-plane component of the wave vector in medium 2 becomes purely imaginary,

$$k_{2z} = i|k_{2z}| = ik_2 \sqrt{\tfrac{n_1^2}{n_2^2} \sin^2\theta_1 - 1}, \tag{2.163}$$

whenever the incidence angle exceeds the threshold,

$$\theta_{\mathrm{c}} = \sin^{-1}(n_2/n_1), \tag{2.164}$$

It then follows at once from Eqs. (2.126) and (2.163) that the for any wave incident at an angle grater than the critical angle given by Eq. (2.164), the reflectivity is unimodular, i. e.,

$$\overline{r}_{\mathrm{p}*} = \frac{\epsilon_2 k_{1z} - i\epsilon_1 |k_{2z}|}{\epsilon_2 k_{1z} + i\epsilon_1 |k_{2z}|}. \tag{2.165}$$

Alternatively, the reflectivity of a totally internally reflected wave can be expressed as

$$\boxed{\overline{r}_{\mathrm{p}*} = e^{-2i\phi_{\mathrm{p}*}}}, \tag{2.166}$$

where the phase can be expressed in terms of the incidence angle and refractive indices of the media as

$$\phi_{\mathrm{p}*} = \tan^{-1}\left(\frac{\epsilon_1|k_{2z}|}{\epsilon_2 k_{1z}}\right). \tag{2.167}$$

To better understand the behavior of the transmitted wave, we derive explicit expressions for its electric and magnetic fields. Using Eq. (2.163) in Eqs. (2.120), we can cast complex amplitudes of the transmitted magnetic and electric fields into

$$\mathbf{H}_t(\mathbf{r},t) = H_t \mathbf{e}_y e^{-|k_{2z}|z} e^{i(k_x x - \omega t)}, \tag{2.168}$$

and

$$\mathbf{E}_t(\mathbf{r},t) = \frac{\eta_0 H_t}{\epsilon_2 k_0}(i|k_{2z}|\mathbf{e}_x - k_x \mathbf{e}_z) e^{-|k_{2z}|z} e^{i(k_x x - \omega t)}. \tag{2.169}$$

We can conclude from Eqs. (2.168) and (2.169) that the transmitted wave fields exponentially decay into medium 2. Next, let us determine the magnitude and direction of the energy flow specified by the time-averaged Poynting vector. It follows from Eqs. (1.26) (2.168), and (2.169) after some algebra that

$$\langle \mathbf{S}_t(z) \rangle = \mathbf{e}_x \frac{4\epsilon_2 k_{1z}^2 k_x}{k_0(\epsilon_2^2 k_{1z}^2 + \epsilon_1^2 |k_{2z}|^2)} I_i e^{-2|k_{2z}|z}, \tag{2.170}$$

where I_i is an optical intensity of the incident wave. It can be concluded from Eq. (2.170) that the power of the wave incident at an angle greater than the total internal reflection angle does not flow into the less optically dense medium. Rather, it propagates along the interface separating the two media, exponentially decaying in the direction normal to the interface. This is a signature of a surface wave. Such surface waves generated by total internal reflection are known as evanescent waves. The evanescent waves play a prominent role in generating surface plasmon polaritons in the laboratory. Indeed, one of the approaches to SPP generation employs evanescent waves. In practice, one uses a device referred to as a Kretschmann prism shown in the figure below. The refractive index of the prism makes it possible to match the in-plane wave vector components for a plane wave launched through the Kretschmann prism under the conditions of total internal reflection to that of the SPP. The launch angle is then determined by the matching condition,

$$k_{x\mathrm{pr}} = k_{x\mathrm{SPP}}, \tag{2.171}$$

implying that

$$n_{\mathrm{pr}} \sin\theta_{\mathrm{SPP}} = \sqrt{\frac{\epsilon_{\mathrm{d}}\epsilon_{\mathrm{m}}}{\epsilon_{\mathrm{d}} + \epsilon_{\mathrm{m}}}}, \tag{2.172}$$

where n_{pr} is the refractive index of the prism, typically it is equal to 1.5 for a glass prism, and ϵ_{d} and $\epsilon\mathrm{m}$ are the permittivities of the dielectric and metal on the two sides of the interface supporting the SPP. In the figure, the SPP is produced at the metal-air

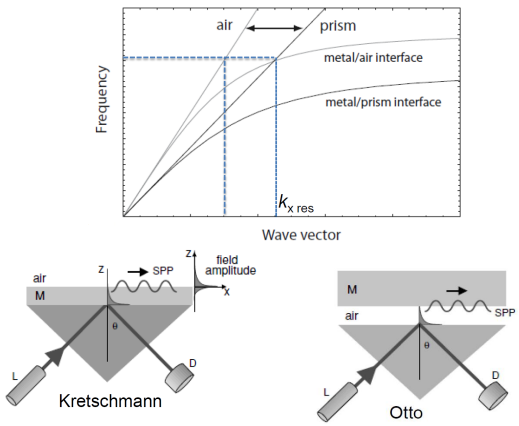

Figure 2.10: Illustrating SPP excitation with Kretschmann method. Reproduced from Novotny& Hecht, *Principles of Nanooptics*.

interface by an evanescent wave tunneling across the metal film from the glass prism.

Exercise 2.14. *Show that the reflectivity of a totally internally reflected TE-wave is given by the expression*

$$r_{s*} = e^{-2i\phi_{s*}}, \tag{2.173}$$

where

$$\phi_{s*} = \tan^{-1}(|k_{2z}|/k_{1z}). \tag{2.174}$$

Derive an expression for the transmitted energy flux.

2.4 Refraction and reflection from dielectric slab: Multi-wave interference

We will now examine a situation when two unbounded, homogeneous isotropic media–media 1 and 3–are separated by a slab of finite thickness d filled with a third medium, medium 3; for simplicity, we assume that the plane coincides with the xz-plane. The situation is illustrated in the figure below. Suppose further that a plane wave is incident from medium 1 onto the interface separating media 1 and 2 and limit ourselves to the instructive case of a p-polarized incident wave throughout this section. We will seek to determine the complex reflectivity and transmittance of the system. Next, we introduce the reflectivity and transmittance of each individual interface, r_{ij} and t_{ij}, $i, j = 1, 2, 3$, respectively, which are determined by Eqs. (2.126) and (2.127).

The incident, reflected, and transmitted magnetic field amplitudes can be expressed as

$$\mathbf{H}_s = H_s \mathbf{e}_y, \qquad s = i, r, t. \tag{2.175}$$

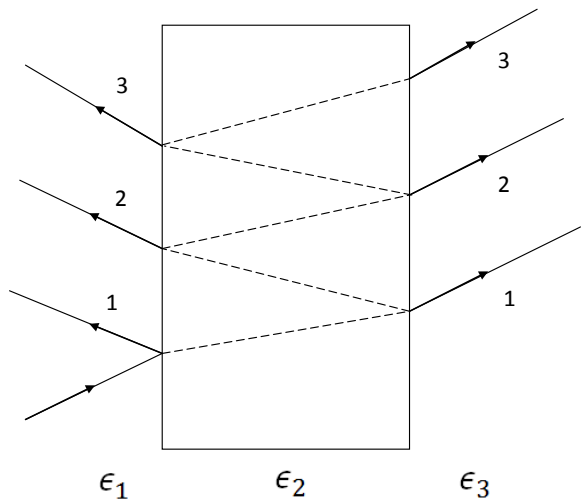

Figure 2.11: Illustrating the multi-wave reflection and transmission through a film.

We will then find the reflected magnetic field in terms of the incident field by adding up the contributions from reflected waves of all orders. Let us consider several lowest-order reflected waves, labeling the reflection order with the corresponding superscript assigned to H_r. The first-order reflected field is simply the field reflected from the first interface once. Thus,

$$\mathbf{H}_r^{(1)} = r_{12} H_i \mathbf{e}_y. \tag{2.176}$$

Next, the second-order reflected field is twice transmitted trough the first interface and once reflected from the second one, i. e.,

$$\mathbf{H}_r^{(2)} = t_{12} t_{21} r_{23} e^{i2k_{2z}d} H_i \mathbf{e}_y, \tag{2.177}$$

where we also included the accrued phase due to the optical path difference. By the same token, the third- and fourth-order reflected waves can be represented as

$$\mathbf{H}_r^{(3)} = t_{12} t_{21} r_{23}^2 r_{21} e^{i4k_{2z}d} H_i \mathbf{e}_y \tag{2.178}$$

and

$$\mathbf{H}_r^{(4)} = t_{12} t_{21} r_{23}^3 r_{21}^2 e^{i6k_{2z}d} H_i \mathbf{e}_y. \tag{2.179}$$

Summing up the contributions to all orders, we obtain

$$\mathbf{H}_r = \mathbf{e}_y H_i \left(r_{12} + r_{23} t_{12} t_{21} e^{i2k_{2z}d} \sum_{s=0}^{\infty} r_{21}^s r_{23}^s e^{i2sk_{2z}d} \right). \tag{2.180}$$

Observe that as follows from Eqs. (2.126) and (2.127),

$$r_{12} = -r_{21}, \quad \text{and} \quad t_{12} = t_{21}. \tag{2.181}$$

Hence, performing the summation on the right-hand side of Eq. (2.180) and employing Eq. (2.181), yields, after minor algebra, the expression

$$\mathbf{H}_r = \mathbf{e}_y H_i \frac{r_{12} + r_{23} e^{i2k_{2z}d}}{1 + r_{12} r_{23} e^{i2k_{2z}d}}. \tag{2.182}$$

Thus introducing the complex reflectivity,

$$\overline{r} \equiv E_r / E_i, \tag{2.183}$$

and using Eqs. (2.125), (2.182), we arrive at

$$\overline{r} = \frac{r_{12} + r_{23} e^{2ik_{2z}d}}{1 + r_{12} r_{23} e^{2ik_{2z}d}}. \tag{2.184}$$

Eq. (2.184) gives the reflectivity of the slab. The outlined method for reflectivity calculation using reflected wave summations of all orders is known as Airy technique.

The analysis of Eqs. (2.184) reveals two instructive particular cases which emerge whenever the reflectivity attains zero,

$$r_{12} + r_{23} e^{2ik_{2z}d} = 0. \tag{2.185}$$

First, we consider the reflectionless transmission of a homogeneous plane wave through a transparent film. This is a multi-wave analog of the Brewster regime except it can occur even for normal incidence. Indeed, as follows from Eq. (2.185) the reflectionless transmission is possible for normal incidence, $k_{2z} = k_2$, provided that

$$2k_2 d = \pi, \tag{2.186}$$

implying a constraint on the slab thickness,

$$d = \frac{\lambda}{4n_2}. \tag{2.187}$$

Eqs. (2.185) and (2.187) are compatible if the refractive index of the slab satisfies the condition

$$n_2 = \sqrt{n_1 n_3}. \tag{2.188}$$

The constraints (2.187) and (2.188) establish requirements for reflectionless transmission of a normally incident plane wave through a dielectric film. In practice, these conditions are taken advantage of in fabricating antireflection coatings of dielectric surfaces such as antireflection glass coating to protect against glare or improve night vision.

The second instance of no reflectivity corresponds to the generation of SPPs on both surfaces of the film. Under the circumstances, the waves multiply reflected from the film interfere constructively to transfer their energy into the SPPs. Thus all power of the incident wave is channeled into the surface waves, resulting in no reflection. In this case, the normal components of all wave vectors must be purely imaginary, a

signature of surface waves. In particular, the normal components of $\mathbf{k}$ in media 1 and 3 can be represented as

$$k_{1z} = -iq_1 = -i\sqrt{k_x^2 - k_1^2}, \tag{2.189}$$

and

$$k_{3z} = iq_3 = i\sqrt{k_x^2 - k_3^2}, \tag{2.190}$$

to ensure the exponential decay of the waves away from the interfaces. Note that these definitions imply that $q_{1,3} > 0$ since the positive root is taken on the right-hand sides of Eqs. (2.189) and (2.190). On the other hand, there exist both exponentially growing and decaying waves inside the slab, implying that

$$k_{2z} = iq_2 = \pm i\sqrt{k_x^2 - k_2^2}. \tag{2.191}$$

The SPP dispersion relation follows at once from Eqs. (2.126), (2.185) and Eqs. (2.189) through (2.191):

$$e^{-2q_2 d} = \left(\frac{\epsilon_1 q_2 + \epsilon_2 q_1}{\epsilon_1 q_2 - \epsilon_2 q_1}\right)\left(\frac{\epsilon_3 q_2 + \epsilon_2 q_3}{\epsilon_3 q_2 - \epsilon_2 q_3}\right). \tag{2.192}$$

In general, Eq. (2.192) describes a rather complicated dispersion relation. To gain a better insight into the SPPs in the film, let us consider a particular case when media 1 and 3 are the same such that $\epsilon_1 = \epsilon_3$ and $q_1 = q_3$. It can then be inferred from Eq. (2.192) after a minor algebra that two families of SPPs exist in this case with the dispersion relations governed by the equations

$$\tanh\left(\frac{q_2 d}{2}\right) = -\frac{\epsilon_1 q_2}{\epsilon_2 q_1}, \tag{2.193}$$

and

$$\tanh\left(\frac{q_2 d}{2}\right) = -\frac{\epsilon_2 q_1}{\epsilon_1 q_2}. \tag{2.194}$$

Exercise 2.15. *Derive Eqs. (2.193)and (2.194).*

It follows at once from Eqs. (2.193) and (2.194) that as the film thickness increases without limit, $d \to \infty$, both dispersion relations reduce to

$$\epsilon_1 q_2 = -\epsilon_2 q_1. \tag{2.195}$$

Since in this case, $q_2 < 0$, one of the permittivities ought to be negative, $\epsilon_2 < 0$, say. Comparison of Eq. (2.195) with (2.146) leads to the conclusion that the SPPs on both sides of a very thick film are uncoupled and have the same dispersion relation as the SPP at the interface of two unbounded media.

In the other extreme of very thin films, $d \to 0$, particularly simple results can be obtained under the condition

$$\tfrac{1}{2} q_2 d \ll 1. \tag{2.196}$$

In other words, the characteristic penetration depth in medium 2, $\delta \simeq |q_2|^{-1}$ is much smaller than half the film thickness. In physical terms, this condition implies strong

coupling between SPPs propagating on both sides of the film. Eqs. (2.193) and (2.196) then yield an approximate expression

$$q_1 \simeq -\frac{2\epsilon_1}{\epsilon_2 d}, \tag{2.197}$$

for the normal component of the wave vector in medium 1. Further, the in-plane component of the wave vector is given by

$$k_x \simeq \sqrt{k_0^2\epsilon_1 + \frac{4\epsilon_1^2}{\epsilon_2^2 d^2}}, \tag{2.198}$$

and the other normal component of the wave vector is

$$q_2 \simeq \pm\sqrt{k_0^2(\epsilon_1 - \epsilon_2) + \frac{4\epsilon_1^2}{\epsilon_2^2 d^2}}. \tag{2.199}$$

In particular, we apply our results to a thin metal film sandwiched between insulator media (IMI). Such a thin-film IMI geometry implies the following conditions

$$0 < \epsilon_1 \ll |\epsilon_2|, \qquad \epsilon_2 < 0. \tag{2.200}$$

Eqs. (2.196) through (2.200) will be consistent for genuinely thin films $d \ll \lambda_0$, yielding

$$|q_2| \simeq \sqrt{k_0^2|\epsilon_2| + \frac{4\epsilon_1^2}{\epsilon_2^2 d^2}}. \tag{2.201}$$

such that the light penetration depth into the metal and dielectric are approximately given by

$$\delta_{\rm m} \simeq \frac{1}{|q_2|}, \qquad \delta_{\rm d} \simeq \frac{|\epsilon_2| d}{2\epsilon_1}. \tag{2.202}$$

This case would correspond to a 20 nm thin metal film, say, with $\epsilon_2 \sim -20$ illuminated from glass $\epsilon_1 \simeq 1.25$ by a light beam with $\lambda_0 \sim 500$ nm, for example. The SPP confinement is still rather tight $\delta_{\rm d} \sim 200$ nm, and $\delta_{\rm m} \sim 50$ nm.

Exercise 2.16. *Plot an explicit dispersion relation curve $\omega = \omega(k_x)$ given by Eq. (2.198). What happens in the static limit, $k_x \to \infty$? You may assume an ideal metal with $\epsilon(\omega) = 1 - \omega_p^2/\omega^2$.*

Exercise 2.17. *Use the Airy technique to show that the transmittance of the slab examined in this section is given by the expression*

$$\bar{t} = \frac{t_{12}t_{23}e^{ik_{2z}d}}{1 + r_{12}r_{23}e^{2ik_{2z}d}}. \tag{2.203}$$

Suppose a dielectric film made of a transparent material is placed in the air. Define the transmission coefficient of the film by the expression

$$\mathcal{T} \equiv \frac{|E_t|^2}{|E_i|^2}, \tag{2.204}$$

and show that $\mathcal{T}$ can be expressed as

$$\mathcal{T} = \frac{T^2}{(1-R)^2}\,\frac{1}{1+F\sin^2\delta}. \tag{2.205}$$

Here

$$\delta = \delta_r + \frac{2\pi n d}{\lambda}\cos\theta_t, \tag{2.206}$$

and we introduced the transmission and reflection coefficients for each interface of the slab, T and R, respectively, and the interferometer finesse F by the expression

$$F = \frac{4R}{(1-R)^2}. \tag{2.207}$$

The considered system serves as a basis for a Fabry-Perot interferometer used to precisely measure the wavelength of light. It can be inferred from Eq. (2.205) –(2.207) that for large enough reflectance, $R \simeq 1$, $\mathcal{T}$ has very sharp maxima at

$$\delta = \pi m, \qquad m = 0, 1, 2, \ldots. \tag{2.208}$$

In the ideal case, $\mathcal{T} = 1$ at the maxima and $\mathcal{T} = 1/F$ at the minima. Thus boosting the finesse, one can increase the contrast of the interferometer. The distance between the adjacent maxima can be determined from Eqs. (2.206) and (2.208) to be

$$d_{m+1} - d_m = \frac{\lambda}{2n\cos\theta}. \tag{2.209}$$

For sufficiently small angles, $\theta \simeq 0$, the latter reduces to

$$d_{m+1} - d_m \simeq \frac{\lambda}{2n}. \tag{2.210}$$

Eq. (2.210) can be used to infer the value of λ from the measurements of the maxima positions.

2.5 Classical theory of optical dispersion and absorption

2.5.1 Lorentz-Kramers expression for dielectric permittivity

As we saw in Sec. 2.2, atoms or molecules of realistic media do not respond instantaneously to an applied external electric field. The time lag between the applied electric field and induced polarization manifests itself as frequency dispersion when one examines the frequency behavior of medium response to a harmonic applied electric field,

$$\mathbf{E}(t) = \mathbf{E}_\omega e^{-i\omega t}. \tag{2.211}$$

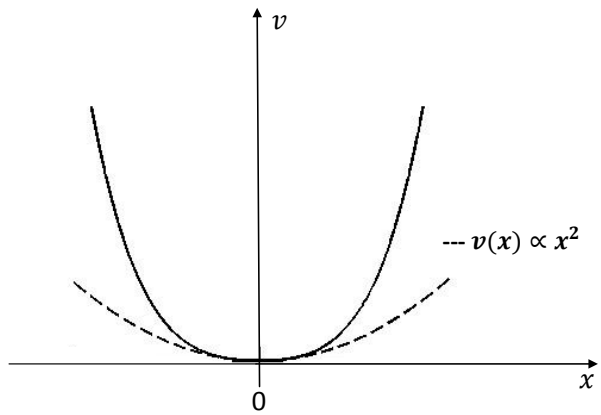

Figure 2.12: Schematic of a trapping Coulomb potential (solid) for an electron in an atom and its harmonic approximation (dashed) near the electron equilibrium position x is a scalar displacement away from equilibrium.

To drive this point home, we develop a simple classical model of matter response to an external time-harmonic field. In this model atoms are treated as simple harmonic oscillators. A linear restoring force proportional to an electron displacement from its equilibrium position–in the classical sense, of course–is due to a harmonic interaction potential between an electron and the other electrons in an atom as well as the atomic nucleus. In reality each atomic electron is trapped by a complicated electrostatic potential which is strongly anharmonic. However, so long as the applied electric field is sufficiently weak such that the electron displacement from its equilibrium position is small compared to the characteristic atomic size, the electrostatic Coulomb potential in the vicinity of the electron equilibrium position can be well approximated by a harmonic one. The situation is schematically depicted in the figure.

Further, we assume that each atom has Z bound electrons. Assume also that there are $f_{\rm s}$ electrons per atom having the binding frequency $\omega_{\rm s}$ which corresponds to a particular type of the trapping harmonic potential. The quantities $\{f_{\rm s}\}$ are referred to as the oscillator strengths.

Whenever an electron having the binding frequency ω_s is displaced by the displacement vector $\mathbf{r}_{\rm s}$ in response to the external electric field, it experiences three forces: the restoring force, $\mathbf{F}_{\rm r} = -m\omega_{\rm s}^2\mathbf{r}_{\rm s}$, the damping force, $\mathbf{F}_{\rm d} = -2m\gamma_s\dot{\mathbf{r}}_{\rm s}$–where $\gamma_{\rm s}$ is a phenomenological damping constant–and the force due to the external electric field, $\mathbf{F}_{\rm e} = -e\mathbf{E}_\omega e^{-i\omega t}$.

The equation of electron motion (second law of Newton) is then

$$m\ddot{\mathbf{r}}_{\rm s} = -m\omega_{\rm s}^2\mathbf{r}_{\rm s} - 2m\gamma_{\rm s}\dot{\mathbf{r}}_{\rm s} - e\mathbf{E}_\omega e^{-i\omega t}. \tag{2.212}$$

Here each "dot" stands for a time derivative. We seek a driven solution to Eq. (2.212) in the form,

$$\mathbf{r}_{\rm s}(t) = \mathbf{r}_{{\rm s}\omega} e^{-i\omega t}. \tag{2.213}$$

It follows from Eqs. (2.212) and (2.213) that the electron displacement amplitude is

$$\mathbf{r}_{\mathrm{s}\omega} = -\frac{e\mathbf{E}_\omega}{m(\omega_{\mathrm{s}}^2 - \omega^2 - 2i\omega\gamma_{\mathrm{s}})}, \tag{2.214}$$

implying that

$$\mathbf{r}_{\mathrm{s}}(t) = -\frac{e\mathbf{E}(t)}{m(\omega_{\mathrm{s}}^2 - \omega^2 - 2i\omega\gamma_{\mathrm{s}})}. \tag{2.215}$$

The induced individual dipole moment of the electron of this type will be $\mathbf{p}_{\mathrm{s}} = -e\mathbf{r}_{\mathrm{s}}$. Next, if there are N atoms per unit volume, the induced polarization is

$$\mathbf{P}(t) = N\sum_{\mathrm{s}} f_s\mathbf{p}_{\mathrm{s}}(t) = -Ne\sum_{\mathrm{s}} f_{\mathrm{s}}\mathbf{r}_{\mathrm{s}}(t) = \frac{Ne^2}{m}\sum_{\mathrm{s}}\frac{f_{\mathrm{s}}\mathbf{E}(t)}{(\omega_{\mathrm{s}}^2 - \omega^2 - 2i\omega\gamma_{\mathrm{s}})}. \tag{2.216}$$

Note that the oscillator strengths satisfy the so-called sum rule

$$\sum_{\mathrm{s}} f_{\mathrm{s}} = Z. \tag{2.217}$$

On comparing Eqs. (1.6), (2.36) and (2.216), we infer that

$$\epsilon(\omega) = 1 + \frac{Ne^2}{\epsilon_0 m}\sum_{\mathrm{s}} f_{\mathrm{s}}\mathcal{L}_s(\omega), \tag{2.218}$$

where we introduced a complex Lorentzian line-shape factor by the expression

$$\mathcal{L}_s(\omega) = \frac{1}{(\omega_{\mathrm{s}}^2 - \omega^2 - 2i\omega\gamma_{\mathrm{s}})}. \tag{2.219}$$

Eqs. (2.218) and (2.219) give a classical expression for the dielectric permittivity of materials as a function of frequency of the applied electric field. The real part describes dispersion while the imaginary part accounts for light absorption by medium atoms. The latter simply because we identified the imaginary part of ϵ with losses as the light propagates through the medium (c.f. Sec. 2.3). Clearly, the light wave loses its energy to the medium atoms which is a classical picture of light absorption.

Let us now explore what happens if the frequency of the applied electric field is close to a particular resonant frequency of the material. For the sake of clarity, let that be the lowest bound frequency of the dielectric, $\omega_0 \neq 0$, i.e, $\omega \approx \omega_0$. In this case, we can single out the resonant term in Eq. (2.218) implying that

$$\epsilon(\omega) = \epsilon_{\mathrm{NR}}(\omega) + \frac{Ne^2 f_0}{\epsilon_0 m}\frac{1}{(\omega_0^2 - \omega^2 - 2i\omega\gamma_0)}. \tag{2.220}$$

As typically $\gamma_s \ll \omega_s$, the contribution to the permittivity due to non-resonant terms, ϵ_{NR} is a purely real and only weakly frequency dependent. It can be expressed as

$$\epsilon_{\mathrm{NR}}(\omega) \simeq \sum_{s\neq 0}\frac{Ne^2 f_{\mathrm{s}}}{\epsilon_0 m(\omega_{\mathrm{s}}^2 - \omega^2)}. \tag{2.221}$$

Notice that close to resonance, we can approximate

$$-\omega^2 + \omega_0^2 - 2i\gamma_0\omega \simeq 2\omega(\omega_0 - \omega - i\gamma_0) \simeq 2\omega_0(\omega_0 - \omega - i\gamma_0). \tag{2.222}$$

It can be inferred from Eqs. (2.221) and (2.222) that the permittivity near optical resonance can be represented as

$$\epsilon(\omega) = \epsilon'(\omega) + i\epsilon''(\omega), \tag{2.223}$$

where

$$\epsilon'(\omega) = \epsilon_{\mathrm{NR}}(\omega) + \frac{Ne^2 f_0}{2\epsilon_0 m\omega_0}\left[\frac{(\omega - \omega_0)}{(\omega - \omega_0)^2 + \gamma_0^2}\right], \tag{2.224}$$

and

$$\epsilon''(\omega) = \frac{Ne^2 f_0}{2\epsilon_0 m\omega_0}\left[\frac{\gamma_0}{(\omega - \omega_0)^2 + \gamma_0^2}\right]. \tag{2.225}$$

The real and imaginary parts of the permittivity are sketched as functions of the frequency in Fig. 2.5.

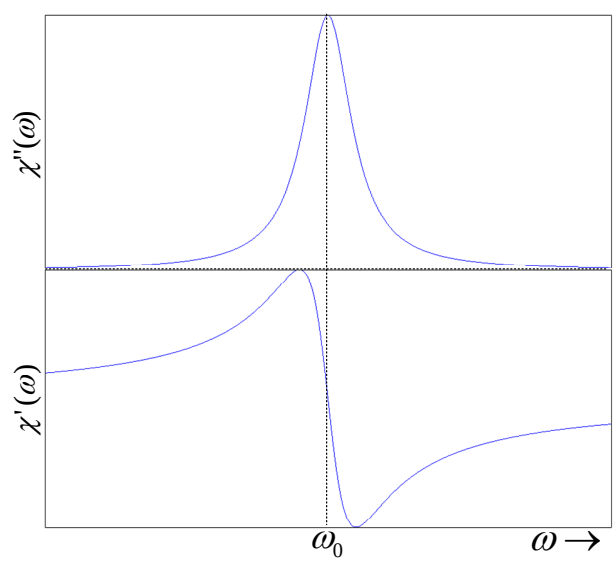

Figure 2.13: Imaginary (top) and real (bottom) parts of the electric permittivity as functions of frequency near resonance.

As is seen in Fig. 2.5., the real part of the permittivity sufficiently far below and above the resonance frequency increases with the frequency. Such a behavior is known as normal dispersion. In the vicinity of resonance, however, ϵ' decreases with the frequency which is referred to as anomalous dispersion. Optical absorption is generally small far from resonance, but is seen to sharply increase as we approach the resonance frequency. Notice also that in regions of weak dispersion are nearly transparent, whereas strong dispersion is accompanied with pronounced absorption as well. This

connection is not accidental. In fact, we show in the following chapters that there are fundamental quantitative relations, the Kramers-Kronig relations that link dispersive and absorptive properties of optical media.

The difference between realistic conductors and dielectrics can be attributed to the presence of free electrons in the former. Indeed, by looking into the low-frequency limit, we notice that for pure dielectrics the lowest bound frequency must be nonzero, while conductors can have a fraction of electrons, f_0, say, that have $\omega_0 = 0$; those are essentially free electrons. Consequently, the dielectric permittivity of conductors is given by the expression

$$\epsilon_{\rm c}(\omega) = \epsilon_{\rm b}(\omega) + i\frac{Nf_0e^2}{\epsilon_0 m\omega(2\gamma_0 - i\omega)}, \tag{2.226}$$

where $\epsilon_{\rm b}$ is the overall contribution of the bound electrons with $\omega_s \neq 0$. Since free electrons can conduct currents, we can use Eq. (2.215) to determine the current density to be

$$\mathbf{J} = -Nef_0\dot{\mathbf{r}}_0 = \frac{Nf_0e^2}{m(2\gamma_0 - i\omega)}\mathbf{E}. \tag{2.227}$$

On comparing Eqs (2.37) and (2.227), we infer the expression for the conductivity,

$$\sigma(\omega) = \frac{Nf_0e^2}{m(2\gamma_0 - i\omega)}. \tag{2.228}$$

It is seen from Eq. (3.86) that in the dc limit $\omega \to 0$, we arrive at

$$\sigma \to \frac{Nf_0e^2}{2m\gamma_0} = \sigma_0, \tag{2.229}$$

the conductivity is real, describing dc currents. In view of Eq. (2.229), the expression for σ can be cast into the form

$$\sigma(\omega) = \frac{\sigma_0}{1 - i\omega\tau}, \tag{2.230}$$

where $\tau = 1/2\gamma_0$ is a characteristic time for current relaxation in conductors.

Next, comparing Eqs. (2.226) and (2.228), we can express the former as

$$\epsilon_{\rm c}(\omega) = \epsilon_{\rm b}(\omega) + i\frac{\sigma}{\epsilon_0\omega}. \tag{2.231}$$

Eq. (2.231) implies that losses in real conductors/metals come in two guises: absorption of electromagnetic waves by bound electrons–which is described by the imaginary part of ϵ_b–and ohmic losses due to generating electric currents as described by the second term on the right-hand side of Eq. (2.231).

Exercise 2.18. *Use the limiting case of Eq. (2.38) for isotropic media with no spatial dispersion and Eq. (2.231) to relate real and imaginary parts of permittivity and conductivity. Thus, you may argue that the distinction between conductors and dielectrics is rather artificial at optical frequencies.*

Next, we note that at the frequencies far exceeding the highest bound frequency, $\omega \gg \max(\omega_s)$, dielectrics and conductors respond to the applied electric field the same

wave. In this limit, we can neglect all $\{\omega_s\}$ and $\{\gamma_s\}$ in the denominator of Eq. (2.218), leading to

$$\epsilon_{\rm c}(\omega) = 1 - \frac{\omega_p^2}{\omega^2}, \tag{2.232}$$

where we used Eq. (2.217) and introduced the plasma frequency

$$\omega_{\rm p} = \sqrt{\frac{NZe^2}{m}}. \tag{2.233}$$

Incidentally, Eq. (2.232) is a simplified form of the so-called Drude expression for a dielectric constant of a metal. The Drude model describes well noble metals; it follows from Eq. (2.232) that ϵ becomes negative for the frequencies above the plasma frequency.

Finally, we note that the polarization caused by a monochromatic applied electric field in an isotropic linear medium can be represented as

$$\mathbf{P}(\mathbf{r},\omega) = \epsilon_0 \chi(\omega) \mathbf{E}(\mathbf{r},\omega), \tag{2.234}$$

where $\chi(\omega)$ is a linear susceptibility of the medium. In case of an optical pulse, consisting of many monochromatic components, the electric field of the pulse can be represented as a Fourier integral viz.,

$$\mathbf{E}(\mathbf{r},t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, \boldsymbol{\mathcal{E}}(\mathbf{r},\omega) e^{-i\omega t}, \tag{2.235}$$

where $\boldsymbol{\mathcal{E}}(\mathbf{r},\omega)$ is the spectral amplitude of the pulse. The polarization field induced by each spectral component of the pulse is given by

$$\boldsymbol{\mathcal{P}}(\mathbf{r},\omega) = \epsilon_0 \chi(\omega) \boldsymbol{\mathcal{E}}(\mathbf{r},\omega). \tag{2.236}$$

It follows at once from Eqs. (2.235) and (2.236) that the overall polarization field induced by the pulse is given by a time convolution,

$$\mathbf{P}(\mathbf{r},t) = \epsilon_0 \int_{-\infty}^{\infty} dt' \chi(t-t') \mathbf{E}(\mathbf{r},t'). \tag{2.237}$$

We will return to Eq. (2.237) in Chap. 4 where we will present a general theory of optical response of nonlocal noninstantaneous nonlinear media to electromagnetic pulses.

2.5.2 Classical theory of Faraday effect

Let us now consider the optical response of an isotropic dielectric to an applied static magnetic field $\mathbf{B}_0$. We will use the Lorentz-Kramers harmonic oscillator model of the medium elaborated in the previous subsection. We will assume that the external magnetic field is weak enough such that it can be treated as a small perturbation. The driven harmonic oscillator equation of motion of each electron can then be rewritten as

$$\ddot{\mathbf{r}}_{\rm s} + 2\gamma_{\rm s}\dot{\mathbf{r}}_{\rm s} + \omega_{\rm s}^2 \mathbf{r}_{\rm s} = -\frac{e}{m}\mathbf{E}_\omega e^{-i\omega t} - \nu \frac{e}{m}[\dot{\mathbf{r}}_{\rm s} \times \mathbf{B}_0]. \tag{2.238}$$

Here we assume that the Lorentz force experienced by an electron due to the external magnetic field $\mathbf{B}_0$ is a small perturbation to the force exerted by the driving harmonic electric field. Instead of using an explicit small dimensionless parameter, we introduced a book-keeping parameter ν to aid keeping track of the same order terms in $\mathbf{B}_0$; we will let $\nu = 1$ at the end of our calculation. We can then represent the electron displacement as a perturbation series in the formal parameter ν as

$$\mathbf{r}_{\rm s} = \mathbf{r}_{\rm s}^{(0)} + \nu \mathbf{r}_{\rm s}^{(1)} + \nu^2 \mathbf{r}_{\rm s}^{(2)} + \dots \tag{2.239}$$

We will seek a driven solution to Eq.(2.238) in the form (2.213). On substituting Eq. (2.239) into (2.238), we can recover, to the first order in ν, the result of the previous subsection, i. e.,

$$\mathbf{r}_{{\rm s}\omega}^{(0)} = -\frac{e}{m}\mathcal{L}_s(\omega)\mathbf{E}_\omega. \tag{2.240}$$

To the first order in ν, we obtain from Eq. (2.238)

$$\ddot{\mathbf{r}}_{\rm s}^{(1)} + 2\gamma_{\rm s}\dot{\mathbf{r}}_{\rm s}^{(1)} + \omega_{\rm s}^2 \mathbf{r}_{\rm s}^{(1)} = \frac{e}{m}[\dot{\mathbf{r}}_{\rm s}^{(0)} \times \mathbf{B}_0]. \tag{2.241}$$

Solving Eq. (2.241) in the steady-state regime, we arrive at the correction term as

$$\mathbf{r}_{{\rm s}\omega}^{(1)} = -\frac{ie^2\omega}{m^2}\mathcal{L}_s^2(\omega)[\mathbf{E}_\omega \times \mathbf{B}_0]. \tag{2.242}$$

Combining Eqs. (1.6) and (2.36, which furnish a macroscopic description of permittivity, with the classical microscopic picture of Eqs. (2.216) as well as with Eqs. (2.239) through (2.242), we finally obtain the following expression for the permittivity tensor,

$$\epsilon_{ij}(\omega) = \epsilon(\omega)\delta_{ij} + ig(\omega)\sum_p e_{ijp}B_{0p}. \tag{2.243}$$

Here

$$\epsilon(\omega) = 1 + \frac{Ne^2}{\epsilon_0 m}\sum_s f_s \mathcal{L}_s(\omega), \tag{2.244}$$

is a dielectric permittivity of an isotropic medium and

$$g(\omega) = \frac{Ne^3\omega}{\epsilon_0 m^2}\sum_s f_s \mathcal{L}_s^2(\omega), \tag{2.245}$$

is a Faraday coefficient which determines the rate of Faraday polarization rotation; it is related to the previously introduced Verdet constant, c. f., Sec. 2.2.3. Notice that Eq. (2.243) is identical to the expression (2.88) which we have introduced before on purely phenomenological grounds. Thus, the presented classical theory of Faraday's effect justifies the phenomenological approach of Sec. 2.3.3. Note also that the microscopic theory furnishes a classical expression for the rotation coefficient g as well.

Exercise 2.19. *Fill in missing steps in the derivation of Eq. (2.243).*

Exercise 2.20. *Extend the above discussion to determine the permittivity tensor correct to the second-order of perturbation theory. Show that the quadratic correction solely determines the rate of polarization rotation of a wave propagating orthogonally to the external magnetic field. This is known as Cotton-Mouton effect.*

Chapter 3

Beams and pulses in linear optics

3.1 Paraxial wave equation and Gaussian beam optics

We consider evolution of a monochromatic electromagnetic field in free space. The electric and magnetic fields can be represented as

$$\mathbf{E}(\mathbf{r},t) = \boldsymbol{\mathcal{E}}(\mathbf{r},\omega)e^{-i\omega t}, \qquad \mathbf{H}(\mathbf{r},t) = \boldsymbol{\mathcal{H}}(\mathbf{r},\omega)e^{-i\omega t}. \tag{3.1}$$

Thus Maxwell's equations for the field envelopes read

$$\nabla \times \boldsymbol{\mathcal{E}} = i\mu_0\omega\boldsymbol{\mathcal{H}}, \tag{3.2}$$

$$\nabla \times \boldsymbol{\mathcal{H}} = -i\epsilon_0\omega\boldsymbol{\mathcal{E}}, \tag{3.3}$$

and

$$\nabla \cdot \boldsymbol{\mathcal{E}} = 0, \qquad \nabla \cdot \boldsymbol{\mathcal{H}} = 0. \tag{3.4}$$

Eliminating the magnetic field in favor of the electric in Eqs. (3.2) – (3.4), we arrive at the equation for the electric field envelope in the form

$$\nabla^2\boldsymbol{\mathcal{E}} + k^2\,\boldsymbol{\mathcal{E}} = 0, \tag{3.5}$$

where $k = \omega/c$.

We seek a plane polarized beam-like solution to (3.41):

$$\boldsymbol{\mathcal{E}} = \mathbf{e}_y\mathcal{E}(x,z)e^{ikz}. \tag{3.6}$$

Physically, the solution (3.34) represents a beam of light propagating in the z-direction with an homogeneous electric field in the y-direction and an inhomogeneous intensity distribution in the x-direction. It automatically satisfies the transversality conditions (3.4). Note that in the limiting case when $\mathcal{E} = const$, we have a plane wave.

The beam is different in that its field amplitude should in some sense be a slowly varying function of coordinates. To make this requirement more quantitative we stipulate that for the intensity distribution to represent a beam, the complex envelope $\mathcal{E}$ change slowly at the wavelength scale, i. e.,

$$\partial_z \mathcal{E} \ll k\mathcal{E}, \tag{3.7}$$

The latter condition is referred to as a slowly-varying amplitude approximation (SVEA). On substituting from Eq. (3.34) and taking the SVEA into account, we arrive at the *paraxial wave equation* for the beam envelope in the form

$$2ik\partial_z \mathcal{E} + \partial^2_{xx}\mathcal{E} = 0. \tag{3.8}$$

Let us now study the evolution of the beam with a Gaussian field profile in the source plane $z = 0$,

$$\mathcal{E}(x, 0) = \mathcal{E}_0 e^{-x^2/2w_0^2}, \tag{3.9}$$

where w_0 characterizes the width of the source intensity profile. We use a Fourier transform method to address the problem. Consider a Fourier decomposition of the beam amplitude in the transverse direction,

$$\mathcal{E}(x, z) = \int_{-\infty}^{+\infty} dq e^{iqx}\, \tilde{\mathcal{E}}(q, z), \tag{3.10}$$

where the Fourier (spectral) amplitude can be determined by the inverse transformation,

$$\tilde{\mathcal{E}}(q, z) = \int_{-\infty}^{+\infty} \frac{dx}{2\pi} e^{-iqx}\, \mathcal{E}(x, z). \tag{3.11}$$

In particular, for the Gaussian beam of (3.9), we can obtain

$$\tilde{\mathcal{E}}(q, 0) = \mathcal{E}_0 \sqrt{\frac{w}{2\pi}} e^{-q^2 w_0^2/2}. \tag{3.12}$$

Here we used the following standard integral

$$\int_{-\infty}^{+\infty} dx e^{-ax^2+bx} = \sqrt{\frac{\pi}{a}} e^{b^2/4a}, \tag{3.13}$$

where a and b are arbitrary complex numbers.

Next, we use the properties of Fourier transforms to convert Eq. (3.44) to the k-space,

$$2ik\partial_z \tilde{\mathcal{E}} - q^2 \tilde{\mathcal{E}} = 0. \tag{3.14}$$

Solving the latter, we obtain

$$\tilde{\mathcal{E}}(q, z) = \tilde{\mathcal{E}}(q, 0) \exp\left(-\frac{iq^2 z}{2k}\right). \tag{3.15}$$

Combining Eqs. (3.12) and (3.15) and using the inverse Fourier transform (3.11), we obtain after some algebra the expression for the Gaussian beam envelope at any z,

$$\mathcal{E}(x,z) = \frac{\mathcal{E}_0}{\sqrt{1+i\zeta}} \exp\left[-\frac{x^2}{2w_0^2(1+i\zeta)}\right], \tag{3.16}$$

where

$$\zeta = z/L_d, \qquad L_d = kw_0^2. \tag{3.17}$$

Exercise 21. *Derive Eq. (3.16).*

To discuss the solution (3.16) it is convenient to represent it in the form where the complex phase and real amplitude are expressed explicitly as

$$\mathcal{E}(x,z) = \mathcal{E}_0 \sqrt{\frac{w_0}{w(z)}}\, e^{i\Phi(z)} \exp\left[\frac{ikx^2}{2R(z)}\right] \exp\left[-\frac{x^2}{2w^2(z)}\right]. \tag{3.18}$$

Exercise 3.1. *Derive Eq. 3.18).*
Here we introduced the beam width $w(z)$ as

$$w(z) = w_0\sqrt{1+z^2/L_d^2}, \tag{3.19}$$

the radius of the wavefront curvature $R(z)$,

$$R(z) = z(1+L_d^2/z^2), \tag{3.20}$$

and the accrued phase $\Phi(z)$,

$$\Phi(z) = -\tfrac{1}{2}\arctan(z/L_d). \tag{3.21}$$

Notice first that although the intensity of a Gaussian beam steadily decreases upon diffraction in free space, the beam profile remains Gaussian in any transverse plane $z = const$. Further, the diffraction length L_d sets the characteristic spatial scale for the problem. It is equal to the distance over which the beam width doubles from its minimal value w_0 at the source. The plane where the beam width is the smallest is called the beam waist and the diffraction length is often referred to as the Rayleigh range.

Consider now the wavefront $\Psi(x,z)$ of the beam which is defined as a surface of constant phase. It follows from Eq . (3.18) that

$$\Psi(x,z) = \Phi(z) + \frac{kx^2}{2R(z)} = const \tag{3.22}$$

We observe that near the waist of the beam, $z \ll L_d$, the radius of the curvature is very large, $R \simeq L_d^2/z$, implying that in the limit $z \to 0$, $R \to \infty$, and the wavefront is flat. In the opposite limit, $z \to +\infty$, the accrued phase is $\Phi = -\pi/4$. This is the so-called Gouy phase shift of a Gaussian beam. Finally for large but finite propagation distances, $z \ll L_d$ such that $R(z) \simeq z$, the wavefront is parabolic

$$z \propto x^2/\lambda, \tag{3.23}$$

with the curvature decreasing in the inverse proportion to the propagation distance. The curvature attains its maximum at the Rayleigh distance, $z_R = L_d$.

Finally, we mention that a natural generalization of the paraxial equation to two transverse dimensions is

$$2ik\partial_z \mathcal{E} + \nabla_\perp^2 \mathcal{E} = 0, \tag{3.24}$$

where $\nabla_\perp^2$ is a Laplacian operator in the transverse plane defined as

$$\nabla_\perp^2 \equiv \partial_{xx}^2 + \partial_{yy}^2. \tag{3.25}$$

3.2 Plane wave decomposition of beams: Angular spectrum

Let us now approach beam propagation in free space from a different perspective. To this end, we consider any linearly polarized electromagnetic field–which, for simplicity, is assumed to be uniform in the polarization direction–as a linear superposition of plane waves in the form

$$\boldsymbol{\mathcal{E}}(x,z) = \mathbf{e}_y \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} dk_x dk_z \, \tilde{\mathcal{A}}(k_x,k_z) e^{i(k_x x + k_z z)}. \tag{3.26}$$

The electromagnetic field is supposed to propagate in free space into the half space $z > 0$. The representation of the field by Eq. (3.26) is known as the angular spectrum: The field is composed of plane waves propagating at different angles to the z-axis.

Substituting from Eq. (3.26) into the wave equation, we obtain the equation for the spectral amplitude $\mathcal{A}$ as

$$\tilde{\mathcal{A}}(k_x,k_z)(-k_x^2 - k_z^2 + k^2) = 0. \tag{3.27}$$

It follows at once from Eq. (4.211) that $\mathcal{A}$ is constrained to lie on the circle in the k-space, i.e,

$$\tilde{\mathcal{A}}(k_x,k_z) = \mathcal{A}(k_x)\delta(k_x^2 + k_z^2 - k^2). \tag{3.28}$$

The circle in the k-space determines the dispersion relation for the wave vector components,

$$k_x^2 + k_z^2 = k^2 \implies k_z = \sqrt{k^2 - k_x^2}. \tag{3.29}$$

It can then be inferred from Eq. (3.29) that

$$k_z = \begin{cases} \sqrt{k^2 - k_x^2}, & k_x < k \\ \pm i\sqrt{k_x^2 - k^2}, & k_x > k \end{cases} \tag{3.30}$$

Combining Eqs. (3.26) and (3.30), we arrive at the angular spectrum representation of **any** linearly polarized $(1+1)$D electromagnetic field in the half-space $z > 0$

$$\boldsymbol{\mathcal{E}}(x,z) = \mathbf{e}_y \underbrace{\int_{k_x<k} dk_x \, \mathcal{A}(k_x) e^{i(k_x x + \sqrt{k^2-k_x^2}\, z)}}_{homogeneous\ waves} + \mathbf{e}_y \underbrace{\int_{k_x>k} dk_x \, \mathcal{A}(k_x) e^{ikx} e^{-\sqrt{k_x^2-k^2}\, z}}_{evanescent\ waves}. \tag{3.31}$$

The first and second terms provide contributions of homogeneous and evanescent plane waves; the latter exponentially decay away from the source plane $z = 0$. Notice incidentally that we chose "$+$" sign to have the evanescent waves decay into $z > 0$ as the exponentially growing solution does not obviously make any sense.

Next, the evanescent waves quickly damp out as the field propagates sufficiently far from the source and their contribution is negligible outside of the source vicinity. Thus, we have

$$\boldsymbol{\mathcal{E}}(x,z) = \mathbf{e}_y \int_{k_x<k} dk_x\, \mathcal{A}(k_x) e^{i(k_x x + \sqrt{k^2-k_x^2}\, z)}. \tag{3.32}$$

Let us now specialize to the beam case whereupon all the plane waves making up the field propagate close to the z-axis such that $k_x \ll k$. It then follows upon a Taylor series expansion in Eq. (3.30) that

$$\sqrt{k^2 - k_x^2} \simeq k - \frac{k_x^2}{2k},$$

Therefore we can rewrite our plane wave decomposition as

$$\boldsymbol{\mathcal{E}}(x,z) \simeq \mathbf{e}_y e^{ikz} \int_{-\infty}^{+\infty} dk_x\, \mathcal{A}(k_x)\, e^{ik_x x}\, \exp\left(-\frac{ik_x^2 z}{2k}\right). \tag{3.33}$$

On comparing Eqs. (3.33) and

$$\boldsymbol{\mathcal{E}}(x,z) = \mathbf{e}_y \mathcal{E}(x,z) e^{ikz}, \tag{3.34}$$

we conclude that we can represent electric fields of optical beams as

$$\mathcal{E}(x,z) = \int_{-\infty}^{+\infty} dk_x\, \mathcal{A}(k_x)\, e^{ik_x x}\, \exp\left(-\frac{ik_x^2 z}{2k}\right). \tag{3.35}$$

It then follows from the Fourier transform definition that

$$\mathcal{E}(x,z) = \int_{-\infty}^{+\infty} dk_x\, \tilde{\mathcal{E}}(k_x,0)\, \exp\left(-\frac{ik_x^2 z}{2k}\right)\, e^{ik_x x}. \tag{3.36}$$

Hence,

$$\tilde{\mathcal{E}}(k_x,z) = \tilde{\mathcal{E}}(k_x,0)\, \exp\left(-\frac{ik_x^2 z}{2k}\right), \tag{3.37}$$

which coincides with Eq. (15) of Lecture 6. Thus our angular spectrum representation treatment is equivalent to the paraxial equation approach. While the latter is usually more convenient to solve practical problems and is straightforwardly generalized to nonlinear situations, the former brings up more insight into the physics of beam propagation in free space.

Finally, applying the convolution theorem of Fourier transforms to Eq. (3.37) and using Eq. (3.13) we can derive the Fresnel representation for any $(1+1)$D beam evolution in free space:

$$\mathcal{E}(x,z) = \sqrt{\frac{k}{2\pi i z}} \int_{-\infty}^{+\infty} dx'\, \mathcal{E}(x',0) \exp\left[\frac{ik(x-x')^2}{2z}\right]. \tag{3.38}$$

Exercise 3.2. *Derive Eq. (3.38).*
A natural generalization of the latter to two transverse dimensions is

$$\mathcal{E}(\boldsymbol{\rho}, z) = \left(\frac{k}{2\pi i z}\right) \int d\boldsymbol{\rho}\, \mathcal{E}(\boldsymbol{\rho}', 0) \exp\left[\frac{ik(\boldsymbol{\rho} - \boldsymbol{\rho}')^2}{2z}\right], \tag{3.39}$$

where $\boldsymbol{\rho} = x\mathbf{e}_x + y\mathbf{e}_y$ is a radius vector in the transverse plane of the beam.

3.3 Pulse propagation in dispersive media: non-resonant case

Let us consider propagation of electromagnetic waves in nonmagnetic media with frequency dispersion. The constitutive relation for the electric flux density in the space-frequency representation reads

$$\tilde{\boldsymbol{\mathcal{D}}}(\mathbf{r}, \omega) = \epsilon_0 \epsilon(\omega) \tilde{\boldsymbol{\mathcal{E}}}(\mathbf{r}, \omega), \tag{3.40}$$

where frequency dispersion enters through the dependence of the dielectric permittivity on the wave frequency. The corresponding wave equation takes the form

$$\nabla^2 \tilde{\boldsymbol{\mathcal{E}}} + \epsilon(\omega) \tfrac{\omega^2}{c^2} \tilde{\boldsymbol{\mathcal{E}}} = 0. \tag{3.41}$$

We seek a linearly polarized spatially homogeneous frequency-dependent wave propagating in the positive z-direction, i.e,

$$\tilde{\boldsymbol{\mathcal{E}}}(\mathbf{r}, \omega) = \mathbf{e}_x \tilde{\mathcal{E}}(\omega, z) e^{ik_0 z}. \tag{3.42}$$

Here k_0 is a wave number associated with the carrier frequency ω_0, and a slowly-varying envelope is assumed such that

$$\partial_z \tilde{\mathcal{E}} \ll k_0 \tilde{\mathcal{E}}, \tag{3.43}$$

Eqs. (3.42) and (3.43) represent a spectral envelope amplitude of a slowly varying optical pulse. On substituting from Eqs. (3.42) and (3.98) into Eq. (3.41), we arrive at the paraxial wave equation in the space-frequency representation,

$$2ik_0 \partial_z \tilde{\mathcal{E}} + [k^2(\omega) - k_0^2] \tilde{\mathcal{E}} = 0, \tag{3.44}$$

where we introduced the frequency-dependent wave number viz.,

$$k^2(\omega) = \epsilon(\omega) \frac{\omega^2}{c^2}. \tag{3.45}$$

Suppose now the bandwidth of the pulse is small compared to the carrier frequency, i.e,

$$\Delta\omega = 2|\omega_{\max} - \omega_0| \ll \omega_0, \tag{3.46}$$

where ω_{max} is the frequency of the highest harmonic within the pulse associated with a finite amplitude. The combined approximations (3.98) and (3.99) constitute the slowly varying envelope approximation for optical pulses. The SVEA implies that

$$\tilde{\mathcal{E}}(\omega, z) \simeq \tilde{\mathcal{E}}(\omega - \omega_0, z) = \tilde{\mathcal{E}}(\omega', z), \tag{3.47}$$

that is the pulse envelope changes slowly over an optical cycle. To this level of accuracy, we can then expand the wave number in a Taylor series as

$$k(\omega) \simeq k_0 + \underbrace{k'(\omega_0)}_{k_1}(\omega - \omega_0) + \frac{1}{2!}\underbrace{k''(\omega_0)}_{k_2}(\omega - \omega_0)^2. \tag{3.48}$$

Assuming further that

$$k(\omega) + k_0 \simeq 2k_0,$$

we can cast Eq. (3.44) into the form

$$i\partial_z\tilde{\mathcal{E}} + k_1\omega'\tilde{\mathcal{E}} + \tfrac{1}{2}k_2\omega'^2\tilde{\mathcal{E}} = 0. \tag{3.49}$$

The overall electric field can then be factorized into a fast carrier wave and slowly varying pulse envelope as

$$\mathbf{E}(t, z) = \mathbf{e}_x \underbrace{e^{i(k_0 z - \omega_0 t)}}_{carrier\ wave} \underbrace{\int_{-\infty}^{+\infty} d\omega' e^{-i\omega' t}\tilde{\mathcal{E}}(\omega', z)}_{slow\ envelope}. \tag{3.50}$$

Introducing a Fourier transform of the pulse envelope spectrum by

$$\mathcal{E}(t, z) = \int_{-\infty}^{+\infty} d\omega' e^{-i\omega' t}\tilde{\mathcal{E}}(\omega'), \tag{3.51}$$

we can derive, using Fourier transform properties, a paraxial wave equation for the temporal envelope

$$2i(\partial_z\mathcal{E} + k_1\partial_t\mathcal{E}) - k_2\partial_{tt}^2\mathcal{E} = 0. \tag{3.52}$$

It is now convenient to transfer to a moving reference frame by introducing the coordinate transformation

$$\zeta = z; \qquad \tau = t - k_1 z, \tag{3.53}$$

One can then re-calculate the derivatives using the chain rules

$$\partial_t\mathcal{E} = \partial_\tau\mathcal{E}; \qquad \partial_{tt}^2\mathcal{E} = \partial_{\tau\tau}^2\mathcal{E}, \tag{3.54}$$

and

$$\partial_z\mathcal{E} = \partial_\zeta\mathcal{E} - k_1\partial_\tau\mathcal{E}, \tag{3.55}$$

to arrive at the final form of the governing pulse propagation equation in linear dispersive media

$$2i\partial_\zeta\mathcal{E} - k_2\partial_{\tau\tau}^2\mathcal{E} = 0. \tag{3.56}$$

To elucidate physical meaning of each term in Eq. (3.56), we observe that if one assumes that at the carrier frequency, $k_2(\omega_0) = 0$, we arrive at the greatly simplified equation

$$\partial_\varsigma \mathcal{E} = 0, \tag{3.57}$$

with the solution

$$\mathcal{E}(t, z) = \mathcal{E}_0(t - z/v_g), \tag{3.58}$$

where $\mathcal{E}_0(t)$ is a pulse envelope in the source plane, and we introduced

$$k_1 \equiv v_g^{-1}. \tag{3.59}$$

It can be concluded from Eq. (3.58) that the pulse maintains its shape and its peak moves inside the medium with the speed v_g. This velocity is referred to as the **group velocity** of the pulse. To understand the role of k_2, it is sufficient to observe that Eq. (3.56) is a temporal analog of the paraxial wave equation governing beam diffraction in free space we have studied before. Hence the second derivative term in Eq. (3.56) describes pulse spreading in dispersive media. The **group velocity dispersion** coefficient k_2 then sets a spatial scale of the problem, the so-called dispersion length, $L_{dis} = t_p^2/k_2$, where t_p is a characteristic duration of the pulse in the source plane $z = 0$.

In the preceding development, we ignored spatial distribution of the pulse, which is justified in a plane wave geometry. Alternatively, pulse propagation in single-mode dispersive fibers can be of interest. In this case, the spatial distribution of the pulse is dictated by the fiber mode such that a more appropriate Ansatz for the field,

$$\tilde{\boldsymbol{\mathcal{E}}}(\mathbf{r}, \omega) = \mathbf{e}_x \tilde{\mathcal{E}}(\omega, z) \phi(\mathbf{r}_\perp, \omega)\, e^{i\beta_0 z}, \tag{3.60}$$

should be considered instead. Here β_0 is a carrier propagation constant in the fiber and $\phi(\mathbf{r}_\perp, \omega)$ is a fiber mode field distribution. Substituting from Eq. (3.60) into (3.41), separating spatial and temporal degrees of freedom and assuming the SVEA (3.98), we obtain the set of equations for the field amplitude

$$2i\beta_0 \partial_z \tilde{\mathcal{E}} + [\beta^2(\omega) - \beta_0^2]\tilde{\mathcal{E}} = 0, \tag{3.61}$$

and the fiber mode

$$\nabla_\perp^2 \phi + [k^2(\omega) - \beta^2(\omega)]\phi = 0. \tag{3.62}$$

Next, assuming (3.99) and that the only allowed fiber mode is excited at the carrier frequency, we can approximate

$$\phi(\mathbf{r}_\perp, \omega) \simeq \phi(\mathbf{r}_\perp, \omega_0), \tag{3.63}$$

and replace $k(\omega)$ and $\beta(\omega)$ in the equation for the fiber mode by their values at the carrier frequency, i.e,

$$\nabla_\perp^2 \phi + [k^2(\omega_0) - \beta_0^2]\phi = 0. \tag{3.64}$$

The resulting eigenvalue equation, subject to the appropriate boundary conditions at the fiber boundaries, determines the spatial distribution of the fiber mode and the mode

propagation constant. Further, expanding the frequency dependent propagation constant $\beta(\omega)$ in a Taylor series up to the second order

$$\beta(\omega) \simeq \beta_0 + \underbrace{\beta'(\omega_0)}_{\beta_1}(\omega - \omega_0) + \frac{1}{2!}\underbrace{\beta''(\omega_0)}_{\beta_2}(\omega - \omega_0)^2, \tag{3.65}$$

and following exactly the same procedure as before, we can arrive at the paraxial wave equation for pulse propagation in linear fibers as

$$2i\partial_\zeta \mathcal{E} - \beta_2 \partial^2_{\tau\tau}\mathcal{E} = 0. \tag{3.66}$$

3.4 Resonant pulse propagation in linear absorbers

3.4.1 Resonant interaction of short pulses with linear media: Homogeneous line broadening

Let us now discuss a more general case of a near-resonant optical pulse, propagating in the medium in the positive z-direction. The displacement x of each Lorentz oscillator induced by the pulse is governed by the equation

$$\partial_t^2 x + 2\gamma\partial_t x + \omega_0^2 x = -eE/m, \tag{3.67}$$

where E is the electric field of the pulse in the scalar approximation. In the slowly-varying envelope approximation (SVEA), the pulse field and atomic dipole moments can be represented as

$$E(z,t) = \tfrac{1}{2}[\mathcal{E}(z,t)e^{i(kz-\omega t)} + c.c]; \qquad ex(z,t) = \tfrac{1}{2}[d_0\sigma(z,t)e^{i(kz-\omega t)} + c.c], \tag{3.68}$$

where ω is a carrier frequency of the pulse, and $d_0 = ex_0$ is a characteristic dipole moment amplitude. Further, $\mathcal{E}$ and σ are slowly varying envelope fields in the sense that

$$\partial_z \mathcal{E} \ll k\mathcal{E}, \qquad \partial_t \mathcal{E} \ll \omega\mathcal{E} \tag{3.69}$$

and

$$\partial_t \sigma \ll \omega\sigma. \tag{3.70}$$

On substituting from (3.85) into (3.67) and using (3.99), we obtain the equation

$$-\omega^2\sigma - 2i\omega\partial_t\sigma - 2i\gamma\omega\sigma + \omega_0^2\sigma = -e\mathcal{E}/mx_0. \tag{3.71}$$

Next, we have near resonance,

$$\omega_0^2 - \omega^2 \simeq 2\omega(\omega_0 - \omega) = 2\omega\Delta, \tag{3.72}$$

where Δ is a detuning of the carrier wave frequency ω from the atomic resonance frequency ω_0. On substituting from Eq. (3.87) into (3.71), we obtain, after some algebra, the SVEA equation for atomic dipole envelope evolution as

$$\partial_t\sigma = -(\gamma + i\Delta)\sigma + i\Omega, \tag{3.73}$$

where we introduced the field envelope in frequency units, $\Omega = -e\mathcal{E}/2m\omega x_0$.

Alternatively, Eq. (3.86) can be written in a real form by introducing the *in-phase* U and *quadrature* V components of the dipole moment viz.,

$$\sigma(t,z) = U(t,z) - iV(t,z), \tag{3.74}$$

such that provided $\Omega^* = \Omega$,

$$\partial_t U = -\gamma U + \Delta V, \tag{3.75}$$

and

$$\partial_t V = -\gamma V - \Delta U + \Omega. \tag{3.76}$$

Thus in the absence of pulse modulation, only the imaginary part of the dipole moment is directly coupled to the electric field amplitude, and it determines the pulse intensity evolution. For this reason, V is termed the absorptive part of σ. The real part U is referred to as dispersive part because it is coupled to the field only via the absorptive part. It will however govern pulse modulation dynamics, if any initial pulse modulation is present.

To better understand physical implications of Eq. (3.86), let us study a particular case of a cw electric field–which has induced the atomic dipole moments in the past–being suddenly switched off. In this case, $\Omega(t) = \theta(-t)\Omega_0(z)$, where $\theta(t)$ is a unit step function. It then follows that for $t > 0$, $\Omega = 0$ and, as follows from Eq. (3.86), each dipole moment exponentially decays with time according to

$$\sigma(t,z) = \sigma(0,z)\theta(t)e^{-\gamma t}e^{i\omega_0 t}. \tag{3.77}$$

This is called free-induction decay of an individual dipole moment. One can introduce a characteristic time $T_0 = 1/\gamma$ which is known as a dipole relaxation time.

A Fourier transform of σ can be defined as

$$\tilde{\sigma}(\omega,z) \equiv \int_{-\infty}^{\infty} dt\, \sigma(t,z)e^{-i\omega t}. \tag{3.78}$$

The spectral response, $S_0(\omega,z) \propto |\tilde{\sigma}(\omega,z)|^2$, obtained in a typical set of absorption measurements, is then given by

$$S_0(\omega,z) \propto \frac{|\sigma(0,z)|^2}{(\omega-\omega_0)^2+\gamma^2} \tag{3.79}$$

The characteristic absorption spectral width is thus $\gamma = 1/T_0$ and is referred to as the width of *homogeneous broadening* as it is the same for each individual atom.

3.4.2 Inhomogeneous broadening

Consider the polarization of a macroscopic sample of atoms. Generally, in solid state samples, the resonant frequency ω_0 of atoms will vary from atom to atom due to local defects which perturb the atomic transition frequencies. As a result, the polarization is determined as an average over the resonant frequency fluctuations such that

$$P(t,z) = \tfrac{1}{2}[\mathcal{P}(t,z)e^{i(kz-\omega t)} + c.c], \tag{3.80}$$

where

$$\mathcal{P}(t,z) = Nd_0\langle\sigma(t,z,\omega_0)\rangle, \tag{3.81}$$

and the averaging is defined as

$$\langle\sigma(t,z,\omega_0)\rangle = \int_0^\infty d\omega_0\, g(\omega_0)\sigma(t,z,\omega_0). \tag{3.82}$$

Here the distribution function $g(\omega_0)$ is normalized to unity as

$$\int_0^\infty d\omega_0\, g(\omega_0) = 1.$$

In reality, the distribution function is often sharply peaked around some value of ω_0 which we denote by $\overline{\omega}_0$, say, i. e.,

$$g(\omega_0) \simeq g(\omega_0 - \overline{\omega}_0) = g(\Delta).$$

It then follows by changing the integration variable to Δ that for any average,

$$\int_0^\infty d\omega_0 g(\omega_0)(\ldots) = \int_{-\overline{\omega}_0}^\infty d\Delta\, g(\Delta)(\ldots) \simeq \int_{-\infty}^\infty d\Delta\, g(\Delta)(\ldots).$$

Thus,

$$\mathcal{P}(t,z) = Nd_0\int_{-\infty}^\infty d\Delta\, g(\Delta)\sigma(t,z,\Delta). \tag{3.83}$$

In gases or atomic vapors, Doppler's effect is at the origin of the frequency detuning distribution. To make this point clear, suppose a plane wave propagating in a laboratory frame has the form $e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$. In the reference frame moving with the atom at the velocity $\mathbf{v}$, the plane wave has the form $e^{i(\mathbf{k}\cdot\mathbf{r}'-\omega' t)}$, where $\mathbf{r}' = \mathbf{r} - \mathbf{v}t$ is a position of the atom at time t. It then follows that the wave form will be the same in the two frames–which it should as it is the same wave–if the frequencies ω' and ω in the moving and laboratory frames, respectively, are related as $\omega' = \omega - \mathbf{k}\cdot\mathbf{v}$. The frequency shift of the wave in a moving reference frame is known as the Doppler effect. For a plane wave propagating in the positive z-direction, the Doppler shifted frequency is

$$\omega' = \omega - kv_z. \tag{3.84}$$

Next, the pulse field and atomic dipole moment distributions in the moving reference frame are

$$E(z,t) = \tfrac{1}{2}[\mathcal{E}(z,t)e^{i[kz-(\omega-kv_z)t]}+c.c]; \qquad ex(z,t) = \tfrac{1}{2}[d_0\sigma(z,t)e^{i[kz-(\omega-kv_z)t]}+c.c], \tag{3.85}$$

and we dropped the prime over z to simplify the notation. The derivation along the lines outlined in the previous Lecture would yield the dipole evolution equation in the form

$$\partial_t\sigma = -(\gamma + i\Delta)\sigma + i\Omega, \tag{3.86}$$

where

$$\Delta = \omega_0 - \omega + kv_z. \tag{3.87}$$

Assuming that $\omega = \omega_0$–the light is tuned to the atomic transition at rest–we obtain the dependence of the detuning on the atom velocity,

$$\Delta = kv_z. \tag{3.88}$$

The atom velocities are distributed according to Maxwell's distribution such that for the z-component of velocity, we have

$$g(v_z) \propto \exp\left(-\frac{mv_z^2}{2k_BT}\right), \tag{3.89}$$

where k_B is the Boltzmann constant and T is the temperature. It then follows from Eqs. (3.88) and (3.89) that the detuning distribution is Maxwellian in this case,

$$g(\Delta) \propto \exp\left(-\frac{m\Delta^2}{2k^2k_BT}\right), \tag{3.90}$$

Let us now revisit the free-induction decay experiment and examine the polarization evolution,

$$\mathcal{P}(t,z) = Nd_0 \int_{-\infty}^{\infty} d\Delta\, g(\Delta)\sigma(t,z), \tag{3.91}$$

which can be rewritten in the free-induction decay as

$$P(t,z) \propto Nd_0 e^{-t/T_0} e^{i\omega_0 t} \int_{-\infty}^{\infty} d\Delta\, g(\Delta) e^{i\Delta t} + c.c. \tag{3.92}$$

Suppose, for definiteness, the detuning distribution is Lorentzian,

$$g(\Delta) \propto \frac{1}{\Delta^2 + 1/T_\Delta^2}, \tag{3.93}$$

where $1/T_\Delta$ characterizes the width of $g(\Delta)$. Using a Fourier transform table integral,

$$\mathcal{F}\left\{\frac{1}{\Delta^2 + 1/T_\Delta^2}\right\} \propto e^{-|t|/T_\Delta},$$

we obtain for $t > 0$,

$$P(t,z) \propto Nd_0 e^{-t/T_{eff}} e^{i\omega_0 t} + c.c. \tag{3.94}$$

Here

$$\frac{1}{T_{eff}} = \underbrace{\frac{1}{T_0}}_{homogeneous} + \underbrace{\frac{1}{T_\Delta}}_{inhomogeneous}. \tag{3.95}$$

The second term on the rhs describes *inhomogeneous broadening* which would occur in the spectral domain due to fluctuations of atomic detunings; its nature is atom specific (distribution of resonant frequencies, velocity distributions, etc.) The functional form of $g(\Delta)$ and the magnitude of a characteristic damping time T_Δ associated with inhomogeneous broadening depend on a specific broadening mechanism.

3.4.3 Maxwell-Lorentz pulse evolution equations and classical area theorem

We start by considering propagation of an optical pulse in a resonant medium. Assuming linear polarization, the electromagnetic field E of the pulse obeys the wave equation in the form

$$\partial_{zz}^2 E - c^{-2}\partial_{tt}^2 E = \mu_0 \partial_{tt}^2 P, \tag{3.96}$$

where the medium polarization P can be expressed as

$$P = -Ne\langle x\rangle. \tag{3.97}$$

In Eq. (3.97), the angle brackets denote averaging over detunings of the pulse from the resonance frequency ω_0.

In the slowly varying envelope approximation, we can use the representation (3.85) and assume that

$$\partial_z \mathcal{E} \ll k\mathcal{E}, \qquad \partial_t \mathcal{E} \ll \omega\mathcal{E} \tag{3.98}$$

and

$$\partial_t \sigma \ll \omega\sigma. \tag{3.99}$$

On substituting from Eq. (3.85) into (3.96) and using the SVEA (3.98), we can obtain the reduced wave equation for the slowly-varying field envelope as

$$\partial_z \Omega + c^{-1}\partial_t \Omega = i\kappa\langle\sigma\rangle. \tag{3.100}$$

which should be coupled with the derived dipole moment evolution equation (3.86). In Eq. (3.100), we introduced a coupling constant, $\kappa = \omega_{pe}^2/4c$, where $\omega_{pe} = (Ne^2/\epsilon_0 m)^{1/2}$ is the electron plasma frequency.

Exercise 3.3. *Derive Eq. (3.100).*

Transforming to the moving reference frame via $\tau = t - z/c$ and $\zeta = z$ just as we did in the derivation of nonresonant pulse propagation equation, we finally arrive at the coupled Maxwell-Lorentz propagation equations

$$\partial_\zeta \Omega = i\kappa\langle\sigma\rangle, \tag{3.101}$$

and

$$\partial_\tau \sigma = -(\gamma + i\Delta)\sigma + i\Omega. \tag{3.102}$$

To solve Eqs. (3.101) and (3.102) we use the familiar now Fourier transform technique. First, we introduce temporal Fourier transforms of the field and dipole moment as

$$\Omega(\tau,\zeta) = \int_{-\infty}^{\infty} d\omega\, \tilde{\Omega}(\omega,\zeta) e^{-i\omega\tau}, \tag{3.103}$$

and

$$\sigma(\tau,\zeta) = \int_{-\infty}^{\infty} d\omega\, \tilde{\sigma}(\omega,\zeta) e^{-i\omega\tau}, \tag{3.104}$$

Substituting those back into our evolution equations, we obtain the algebraic expression for $\tilde{\sigma}$ in the form

$$\tilde{\sigma}(\omega,\zeta) = \frac{i\tilde{\Omega}(\omega,\zeta)}{\gamma + i(\Delta - \omega)}. \tag{3.105}$$

It then follows from Eq. (3.105) and a Fourier transformed Eq. (3.101) that

$$\partial_\zeta \tilde{\Omega} = -\kappa \mathcal{R}\, \tilde{\Omega}, \tag{3.106}$$

where the spectral material response function is defined as

$$\mathcal{R}(\omega) = \left\langle \frac{1}{\gamma + i(\Delta - \omega)} \right\rangle. \tag{3.107}$$

Integrating Eq. (3.106) at once, we arrive at

$$\tilde{\Omega}(\omega, \zeta) = \tilde{\Omega}(\omega, 0) \exp[-\kappa \mathcal{R}(\omega)\zeta]. \tag{3.108}$$

Hence the field envelope at any propagation distance can be expressed as

$$\mathcal{E}(\tau, \zeta) = \int_{-\infty}^{\infty} d\omega\, \tilde{\mathcal{E}}(\omega) \exp[-i\omega\tau - \kappa \mathcal{R}(\omega)\zeta], \tag{3.109}$$

where

$$\tilde{\mathcal{E}}(\omega) = \int_{-\infty}^{\infty} \frac{dt'}{2\pi}\, e^{i\omega t'} \mathcal{E}(t', 0). \tag{3.110}$$

On combining Eqs. (3.109) and (3.110), we can express the answer in the original variables in the form

$$\mathcal{E}(t, z) = \int_{-\infty}^{\infty} \frac{dt'}{2\pi}\, \mathcal{E}(t', 0) \int_{-\infty}^{\infty} d\omega\, e^{i\omega(t'-t)} \exp[i\omega z/c - \kappa \mathcal{R}(\omega) z]. \tag{3.111}$$

Exercise 3.4. *Fill in missing steps in the derivation of Eq. (3.111).*
Note that in the absence of inhomogeneous broadening (the so-called sharp line limit), $g(\Delta) = \delta(\Delta)$ and

$$\mathcal{R}_{hom}(\omega) = \frac{1}{(\gamma - i\omega)}. \tag{3.112}$$

The so-called classical area theorem follows directly from Eq. (3.111). Indeed, let us introduce the classical area, $\mathcal{A}$ as

$$\mathcal{A}(z) = \int_{-\infty}^{\infty} dt\, \mathcal{E}(t, z). \tag{3.113}$$

Integrating Eq. (3.111) over time and using the integral representation of the delta function,

$$\delta(\omega) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega t}, \tag{3.114}$$

we arrive at the area theorem

$$\mathcal{A}(z) = \mathcal{A}_0 \exp[-\kappa \mathcal{R}(0) z], \tag{3.115}$$

where $\mathcal{A}_0 = \mathcal{A}(0)$ is the initial area under the pulse profile. In general, the area theorem can be cast into the form

$$\mathcal{A}(z) = \mathcal{A}_0 e^{-\alpha z/2} e^{i\beta z/2}, \tag{3.116}$$

where we introduced a characteristic attenuation decrement α and the phase accumulation factor β by the expressions

$$\alpha = \left\langle \frac{2\kappa\gamma}{\gamma^2 + \Delta^2} \right\rangle, \tag{3.117}$$

and

$$\beta = \left\langle \frac{2\kappa\Delta}{\gamma^2 + \Delta^2} \right\rangle. \tag{3.118}$$

Thus, regardless of the incident pulse shape, the area under the pulse will exponentially decay on pulse propagation in linear resonant absorbers as a consequence of medium absorption manifested, in general, through homogeneous and inhomogeneous broadening.

Exercise 3.5. *Derive Eqs. (4.211) and (3.116).*

Finally, we examine the case of very long pulses such that the characteristic pulse width T_p is much longer than the longer of homogeneous or inhomogeneous damping times,

$$T_p \gg \max(T_0, T_\Delta). \tag{3.119}$$

It then follows from Eq. (3.102) that the dipole moment can be adiabatically eliminated: It decays fast to its dynamic equilibrium value determined by the pulse amplitude. Mathematically, we can formally set $\partial_\tau \sigma \simeq 0$ in Eq. (3.102) and conclude that

$$\sigma \simeq \frac{i\Omega}{\gamma + i\Delta}, \tag{3.120}$$

On substituting back into Eq. (3.101) we arrive at the pulse evolution equation as

$$\partial_\zeta \mathcal{E} = -\kappa \left\langle \frac{1}{\gamma + i\Delta} \right\rangle \mathcal{E}. \tag{3.121}$$

The latter implies that

$$\mathcal{E}(t, z) = e^{-\alpha z/2} e^{i\beta z/2} \mathcal{E}_0(t - z/c), \tag{3.122}$$

where $\mathcal{E}_0(t)$ is a pulse profile in the source plane.

Exercise 3.6. *Derive Eq. (3.122).*

Equation (3.122) is Beer's absorption law, familiar from elementary optics treatment of absorbers. It states that sufficiently long pulses propagate in absorbers undistorted except that their amplitudes decay exponentially with the propagation distance; the typical damping distance is known as Beers' absorption length, $L_B = \alpha^{-1}$.

Chapter 4

Nonlinear optics

4.1 Introduction to nonlinear optics

Whenever an external electric field is applied to matter, it induces or reorients dipole moments of atoms or molecules of the matter, resulting in a nonzero average dipole moment per unit volume or polarization of the material. If the applied electric field is not too large, the polarization is proportional to the field strength, i.e,

$$P = \epsilon_0 \chi^{(1)} E, \tag{4.1}$$

where $\chi^{(1)}$ is the usual susceptibility of linear optics. In writing Eq. (4.1) we ignored, for simplicity, the vector nature of both the applied field and the resulting polarization.

As the magnitude of the field increases though, the simple linear relation (4.1) no longer holds. However, typical electric fields generated by all but most powerful modern lasers are in the range of 10^6 to 10^7 V/cm, whereas the electrons bound to atoms or molecules experience far greater fields of the order of 10^9 to 10^{10} V/cm. Consequently, one can assume the induced electron displacements in laser fields to be rather small; the latter circumstance justifies using a power series representation for the induced polarizations as

$$P = \epsilon_0 (\chi^{(1)} E + \chi^{(2)} E^2 + \chi^{(3)} E^3 + \ldots), \tag{4.2}$$

where $\chi^{(2)}$ and $\chi^{(3)}$ are referred to as second- and third-order susceptibilities, respectively.

To estimate orders of magnitude of the nonlinear susceptibilities, we consider nonlinearity of electronic origin. In this case, the nonlinear polarization depends on the displacements of the electrons from the nuclei. One could expect that the second-order contribution to the polarization would definitely be of the same order as the first one if the electrons are displaced a distance as large as the atomic size, which is roughly of the order of the Bohr radius, $a_0 = \hbar^2/me^2 \simeq 5 \times 10^{-9}$ cm. The corresponding electric field would be comparable with the field binding electrons to a nucleus, $E_{at} = e/4\pi\epsilon_0 a_0^2 \simeq 5 \times 10^{11}$ V/m. As the linear susceptibility is of the order of unity,

$\chi^{(1)} \sim 1$, it follows that the second-order susceptibility can be estimated as

$$\chi^{(2)} \sim E_{at}^{-1} \sim 10^{-12}, \text{ m/V}. \tag{4.3}$$

By the same token, a typical value of the third-order susceptibility for condensed-matter systems would be

$$\chi^{(3)} \sim 10^{-21} \text{ to } 10^{-22}, \text{ m}^2/\text{V}^2. \tag{4.4}$$

It can be readily inferred from Eqs. (4.3) and (4.4) that (a) one needs very large fields indeed to probe nonlinear response of dielectric materials and (b) for most laser field strengths encountered in practice, each higher-order contribution to the polarization field P is much smaller than the corresponding lower-order one, enabling us to take into account only the lowest order nonvanishing contribution to P in a given nonlinear medium.

In the following subsection, we are going to discuss nonlinear optical susceptibilities semi-quantitatively. A note of caution is due before we proceed any further: The just introduced expansion (4.2) fails in the vicinity of any internal atomic resonance of the medium, where nonlinear saturation effects start playing a role. Hence, a more subtle quantum theory has to be developed to describe such resonant light-matter interactions. Hereafter, we assume that frequencies of all electric fields involved are far away from any material resonance.

4.1.1 Qualitative description of nonlinear optical processes

We now qualitatively examine second-order processes, starting with the second harmonic generation (SHG). To this end, consider a monochromatic input field,

$$E(t) = \tfrac{1}{2}(\mathcal{E}e^{-i\omega t} + c.c),$$

The second-order polarization associated with the field is

$$P^{(2)}(t) = \epsilon_0\chi^{(2)}E^2(t) = \tfrac{1}{2}\epsilon_0\chi^{(2)}|\mathcal{E}|^2 + \tfrac{1}{4}(\epsilon_0\chi^{(2)}\mathcal{E}^2e^{-i2\omega t} + c.c).$$

The first process describes generation of a dc field, **optical rectification** while the second is **second harmonic generation**. It is schematically illustrated in the block-diagram below.

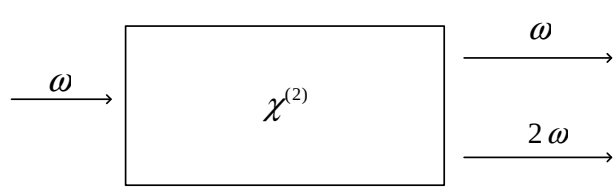

Figure 4.1: Illustrating the second harmonic generation.

In the SHG process an input wave of frequency ω generates an output at double frequency in a nonlinear medium. **Sum-** and **difference-frequency generation** are

more general processes taking place if two different input frequencies ω_1 and ω_2 are present. The input field is then

$$E(t) = \tfrac{1}{2}(\mathcal{E}_1 e^{-i\omega_1 t} + \mathcal{E}_2 e^{-i\omega_2 t} + c.c).$$

The generated output polarization takes the form

$$P^{(2)}(t) = \tfrac{1}{2}\sum_s \mathcal{P}(\omega_s) e^{-i\omega_s t} + c.c, \tag{4.5}$$

where the summation is over all possible combinations s of two frequency components and

$$\mathcal{P}_{SHG}(2\omega_j) = \tfrac{1}{2}\epsilon_0 \chi^{(2)} \mathcal{E}_j^2,$$
$$\mathcal{P}_{SFG}(\omega_1 + \omega_2) = \epsilon_0 \chi^{(2)} \mathcal{E}_1 \mathcal{E}_2,$$
$$\mathcal{P}_{DFG}(\omega_1 - \omega_2) = \epsilon_0 \chi^{(2)} \mathcal{E}_1 \mathcal{E}_2^*,$$
$$\mathcal{P}_{OR}(0) = \epsilon_0 \chi^{(2)} (|\mathcal{E}_1|^2 + |\mathcal{E}_2|^2).$$

While the first and last terms describe SHG and OR, the second and third correspond

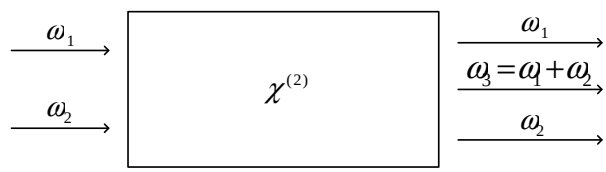

Figure 4.2: Schematic illustration of the sum-frequency generation process.

to new processes of sum- and difference frequency generation, to be abbreviated as (SFG) and (DFG), respectively. The block diagrams of the processes are displayed in Figs. 2 and 3.

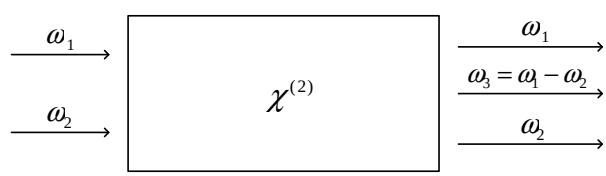

Figure 4.3: Schematic illustration of the difference-frequency generation process.

The fundamental difference between the two processes can be seen from the energy-level diagrams below.

In the SFG process two input photons at frequencies ω_1 and ω_2 annihilate giving rise to one photon at the sum frequency, $\omega_3 = \omega_1 + \omega_2$. In the DFG process, however, annihilation of a pump photon at frequency ω_1 and generation of a difference frequency photon $\omega_3 = \omega_1 - \omega_2$–sometimes referred to as signal–go hand in hand with generation

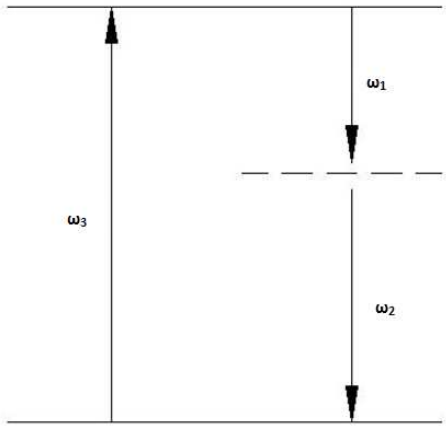

Figure 4.4: Energy-level description of sum-frequency generation.

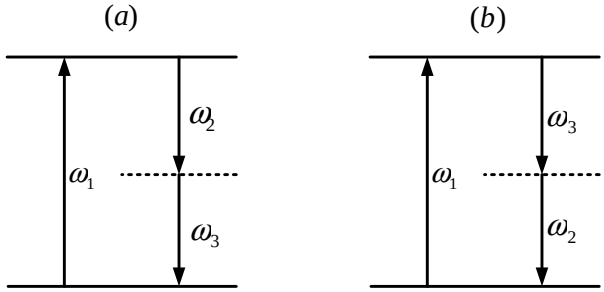

Figure 4.5: Energy-level diagram of difference-frequency generation.

of an idler photon at frequency ω_2, say. Thus the DFG production is accompanied by the amplification of one of input fields at the expense of the other. For this reason, DFG is often referred to as optical parametric amplification. SHG, SFG and DFG are collectively known as three-wave mixing processes.

Next, we briefly consider another three-wave mixing process, **stimulated Raman scattering** (SRS) which can be quantitatively described quantum-mechanically. In the SRS a pump photon of frequency ω gets blue-(Stokes mode) or red-shifted (anti-Stokes mode) such that $\omega_S = \omega - \omega_v$ and $\omega_A = \omega + \omega_v$ exciting some medium degrees of freedom on the way. As it was first studied in molecules where SRS causes medium vibrations, we used the subscript "v" to indicate the frequency ω_v of generated molecular vibrations. The process can be described by energy-level diagrams to be presented in class.

Further, we consider the third-order processes, associated with $\chi^{(3)}$. As there are plethora of those–all falling into a general category of four-wave mixing–we will limit ourselves in this course to only third harmonic generation (THG) and self-focusing

(SF), both excited by a monochromatic input field,

$$E(t) = \frac{1}{2}(\mathcal{E}e^{-i\omega t} + c.c),$$

The third-order polarization,

$$P^{(3)}(t) = \epsilon_0 \chi^{(3)} E^3(t)$$

The application of the trigonometric identity, $\cos^3 \omega t = \frac{1}{4}\cos 3\omega t + \frac{3}{4}\cos \omega t$ results in

$$P^{(3)}(t) = \frac{1}{2}[\mathcal{P}(3\omega)e^{-i3\omega t} + \mathcal{P}(\omega)e^{-i\omega t} + c.c.],$$

where the THG polarization field is

$$\mathcal{P}_{THG}(3\omega) = \frac{1}{2}\epsilon_0 \chi^{(3)} \mathcal{E}^3,$$

and the SF polarization field takes the form

$$\mathcal{P}_{SF}(\omega) = \frac{3\epsilon_0}{2}\chi^{(3)}|\mathcal{E}|^2\mathcal{E}.$$

The THG process is a third-order analog of the THG process; the THG block diagram is as follows The SF process is so called because the input field modifies the refractive

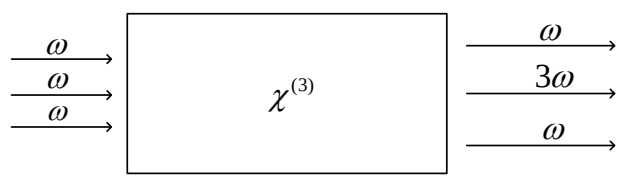

Figure 4.6: Illustrating the third harmonic generation.

index of the medium to

$$n = n_0 + n_2|\mathcal{E}|^2,$$

leading to self-lensing of a light beam. The self-induced "medium lens" is a positive one if $n_2 > 0$ and a negative one otherwise. Thus, either self-focusing or self-defocusing ensues. Another third-order process that, in general, accompanies SF is **two-photon absorption** (TPA). In the TPA process, two photons can be absorbed from a light wave by a medium atom, promoting the latter to an excited state which cannot be related to the ground state by a dipole transition.

4.2 Anharmonic oscillator model

Let us now generalize Lorentz's model to treat a nonlinear response of the medium. A natural generalization is to model atoms as anharmonic oscillators whose equation of motion takes the form

$$m\ddot{x} = -2\gamma m\dot{x} - eE + F_r, \tag{4.6}$$

where the first term on the r.h.s of Eq. (4.6) is a velocity-dependent radiation-reaction force, while the second term

$$F_r = -\frac{\partial V}{\partial x}, \tag{4.7}$$

represents a restoring force modeling the nonlinear coupling between the electron and the nucleus. $V(x)$ is the potential energy of the electron-nucleus interaction, which, in general, contains quadratic as well as higher-order terms:

$$V(x) = \frac{1}{2}m\omega_0^2 x^2 + \frac{1}{3}max^3 + \frac{1}{4}mbx^4 + \dots \tag{4.8}$$

For sufficiently small electron displacements, the restoring force is linear, and using the Lorentz model is justified. As the strength of the driving field E increases, larger electron displacements become possible and nonlinear contributions to the restoring force must be taken into account. If we restrict our consideration to the second-order nonlinearity, the equation of motion reduces to

$$\ddot{x} + \omega_0^2 x + 2\gamma\dot{x} + ax^2 = -\lambda\frac{eE}{m}. \tag{4.9}$$

Here we introduced a formal bookkeeping perturbation parameter λ which is needed to keep track of the same order terms in the driving field powers; we will let $\lambda = 1$ at the end.

In general, the driving field governing any second-order process is bi-chromatic, i.e,

$$E(t) = \tfrac{1}{2}(\mathcal{E}_1 e^{-i\omega_1 t} + \mathcal{E}_2 e^{-i\omega_2 t} + c.\ c.). \tag{4.10}$$

We seek a driven solution to Eq. (4.9) to a given order in λ. Such a solution can be sought in the form

$$x = \lambda x^{(1)} + \lambda^2 x^{(2)} + \lambda^3 x^{(3)} + \dots. \tag{4.11}$$

To the first order in the perturbation parameter, we obtain

$$\ddot{x}^{(1)} + \omega_0^2 x^{(1)} + 2\gamma\dot{x}^{(1)} = -\frac{e\mathcal{E}_1}{m}e^{-i\omega_1 t} - \frac{e\mathcal{E}_2}{m}e^{-i\omega_2 t} + c.\ c., \tag{4.12}$$

We look for a solution to Eq. (4.12) in the form

$$x^{(1)} = \tfrac{1}{2}(x_{\omega_1} e^{-i\omega_1 t} + x_{\omega_2} e^{-i\omega_2 t} + c.\ c.), \tag{4.13}$$

and we arrive, after minor algebra, at the expression

$$x_{\omega_j} = -\frac{e\mathcal{E}_j}{m\mathcal{D}(\omega_j)}. \tag{4.14}$$

Here we introduced the spectral denominator

$$\mathcal{D}(\omega_j) \equiv -\omega_j^2 + \omega_0^2 - 2i\gamma\omega_j. \tag{4.15}$$

To the second order in λ, the anharmonic oscillator equation of motion is

$$\ddot{x}^{(2)} + \omega_0^2 x^{(2)} + 2\gamma\dot{x}^{(2)} = -a[x^{(1)}]^2. \tag{4.16}$$

A driven solution to Eq. (4.16) should be sought in the form

$$\begin{aligned} x^{(2)} &= \tfrac{1}{2}\left[x_{2\omega_1}e^{-i2\omega_1 t} + x_{2\omega_2}e^{-i2\omega_2 t} + x_{\omega_1+\omega_2}e^{-i(\omega_1+\omega_2)t}\right. \\ &\quad \left. + x_{\omega_1-\omega_2}e^{-i(\omega_1-\omega_2)t} + x_{\omega_2-\omega_1}e^{-i(\omega_2-\omega_1)t} + c.\ c.\right] \end{aligned} \tag{4.17}$$

On substituting from Eqs. (4.17) into Eq. (4.16), we obtain the displacement terms corresponding to the second-harmonic generation (SHG), sum- and difference-frequency generation, (SFG) and (DFG), respectively, as

SHG :
$$x_{2\omega_j} = -\frac{ae^2\mathcal{E}_j^2}{2m^2\mathcal{D}^2(\omega_j)\mathcal{D}(2\omega_j)}. \tag{4.18}$$

SFG :
$$x_{\omega_1+\omega_2} = -\frac{ae^2\mathcal{E}_1\mathcal{E}_2}{m^2\mathcal{D}(\omega_1)\mathcal{D}(\omega_2)\mathcal{D}(\omega_1+\omega_2)}. \tag{4.19}$$

DFG :
$$x_{\omega_j-\omega_{3-j}} = -\frac{ae^2\mathcal{E}_j\mathcal{E}_{3-j}^*}{m^2\mathcal{D}(\omega_j)\mathcal{D}^*(\omega_{3-j})\mathcal{D}(\omega_j-\omega_{3-j})}. \tag{4.20}$$

Here $j = 1, 2$.

Recall the definition of the polarization fields:

$$\mathcal{P}(2\omega_j) = -Nex_{2\omega_j}, \tag{4.21}$$

$$\mathcal{P}(\omega_1+\omega_2) = -Nex_{\omega_1+\omega_2}, \tag{4.22}$$

and

$$\mathcal{P}(\omega_j-\omega_{3-j}) = -Nex_{\omega_j-\omega_{3-j}};\ j = 1, 2. \tag{4.23}$$

On the other hand, the appropriate susceptibilities are determined from

$$\mathcal{P}(2\omega_j) = \tfrac{1}{2}\epsilon_0\chi^{(2)}(-2\omega_j;\omega_j,\omega_j)\mathcal{E}_j^2, \tag{4.24}$$

$$\mathcal{P}(\omega_1+\omega_2) = \epsilon_0\chi^{(2)}(-\omega_1-\omega_2;\omega_1,\omega_2)\mathcal{E}_1\mathcal{E}_2, \tag{4.25}$$

$$\mathcal{P}(\omega_j-\omega_{3-j}) = \epsilon_0\chi^{(2)}(-\omega_j+\omega_{3-j};\omega_j,-\omega_{3-j})\mathcal{E}_j\mathcal{E}_{3-j}^*;\ j = 1, 2. \tag{4.26}$$

It can then be concluded by inspection that

$$\chi_{SHG}^{(2)}(-2\omega_j;\omega_j,\omega_j) = \frac{Nae^3}{\epsilon_0 m^2\mathcal{D}^2(\omega_j)\mathcal{D}(2\omega_j)}, \tag{4.27}$$

and

$$\chi_{SFG}^{(2)}(-\omega_1-\omega_2;\omega_1,\omega_2) = \frac{Nae^3}{\epsilon_0 m^2\mathcal{D}(\omega_1)\mathcal{D}(\omega_2)\mathcal{D}(\omega_1+\omega_2)}, \tag{4.28}$$

as well as

$$\chi_{DFG}^{(2)}(-\omega_j+\omega_{3-j};\omega_j,-\omega_{3-j}) = \frac{Nae^3}{\epsilon_0 m^2\mathcal{D}(\omega_j)\mathcal{D}(-\omega_{3-j})\mathcal{D}(\omega_j-\omega_{3-j})}, \tag{4.29}$$

where

$$\mathcal{D}(-\omega_{3-j}) = \mathcal{D}^*(\omega_{3-j}). \tag{4.30}$$

Equations (4.27) – (4.29) can be used to estimate the magnitude of $\chi^{(2)}$ for a typical non-centrosymmetric system far off resonance. We can easily infer from (4.27) – (4.29) – for a far off resonance driving frequencies we can safely assume $\omega_0 \gg \omega, \gamma$, resulting in the estimate $\mathcal{D} \sim \omega_0^2$ – that

$$\chi^{(2)} \simeq \frac{Ne^3 a}{m^2 \epsilon_0 \omega_0^6}. \tag{4.31}$$

To estimate a, we notice that the electron displacement is expected to be of the order of the atom size – which is also roughly equal to the interatomic distance d for solids – and a is determined from the requirement that the linear and nonlinear restoring forces be of the same order, that is

$$m\omega_0^2 d \sim mad^2, \tag{4.32}$$

implying that

$$a \sim \omega_0^2/d. \tag{4.33}$$

Noting finally that $d \sim N^{-1/3}$, we obtain

$$\chi^{(2)} \simeq \frac{N^{4/3} e^3}{m^2 \epsilon_0 \omega_0^4}. \tag{4.34}$$

Using the values $\omega_0 \simeq 10^{16}$ rad/s, $N \simeq 10^{28}$, m^{-3}, $e \simeq 10^{-19}$ C, and $m \simeq 10^{-30}$ kg, we arrive at the estimate

$$\chi^{(2)} \sim 10^{-12}\text{m/V}. \tag{4.35}$$

Exercise 4.1. *The third-order nonlinear response is determined from the equation of motion*

$$\ddot{x} + \omega_0^2 x + 2\gamma\dot{x} + ax^2 + bx^3 = \lambda \frac{eE}{m}. \tag{4.36}$$

Find an expression for $\chi^{(3)}(-\omega, \omega, -\omega, \omega)$ for the degenerate four-wave mixing process and estimate the magnitude of $\chi^{(3)}$ from it.

Exercise 4.2. *Use the above anharmonic oscillator model to determine a frequency dependence of the nonlinear susceptibility $\chi^{(3)}(-3\omega, \omega, \omega, \omega)$, corresponding to the the third harmonic generation.*

4.3 Nonlinear optical processes generated by cw driving fields: A general approach

We first examine generic nonlinear response of media to a quasi cw input wave: The input spectrum consists of a finite number of monochromatic components. In particular, we consider the nonlinear polarization field resulting from a general three-wave mixing process,

$$P_i^{(2)}(t) = \tfrac{1}{2} \textstyle\sum_s \mathcal{P}_i(\omega_s) e^{-i\omega_s t} + c.c, \tag{4.37}$$

where the summation is over all possible combinations s of two two monochromatic components with frequencies ω_1 and ω_2. The complex amplitude of the second-order

polarization is then

$$\mathcal{P}_i(\omega_3) = \epsilon_0 c^{(2)}(\omega_1, \omega_2) \sum_{jk} \chi^{(2)}_{ijk}(-\omega_3; \omega_1, \omega_2) \mathcal{E}_j(\omega_1) \mathcal{E}_k(\omega_2), \tag{4.38}$$

Here the the subscripts stand for Cartesian components of the fields, $c^{(2)}(\omega_1, \omega_2)$ is a degeneracy factor corresponding to a number of permutations of the two frequency components giving the same ω_3, and the following convention is adopted

$$\mathcal{E}_l(-\omega_s) = \mathcal{E}^*_l(\omega_s).$$

Summarizing the results of Sec. 4.1.1, we readily conclude that

$$c^{(2)}(\omega_1, \omega_2) = \begin{cases} 1, & \omega_1 \neq \omega_2; \\ 1/2, & \omega_1 = \omega_2. \end{cases} \tag{4.39}$$

Similarly, the complex amplitude of the third-order polarization can be represented as

$$\mathcal{P}_i(\omega_4) = \epsilon_0 c^{(3)}(\omega_1, \omega_2, \omega_3) \sum_{jkl} \chi^{(3)}_{ijkl}(-\omega_4; \omega_1, \omega_2, \omega_3) \mathcal{E}_j(\omega_1) \mathcal{E}_k(\omega_2) \mathcal{E}_l(\omega_3), \tag{4.40}$$

The tensors $\chi^{(2)}_{ijk}$ and $\chi^{(3)}_{ijkl}$ are known as the second- and third-order nonlinear susceptibility tensors, respectively, and are the formal generalizations of the corresponding nonlinear susceptibilities studied in Sec. 4.1.1.

4.4 Nonlinear processes generated by arbitrary fields: Spatial and temporal dispersion

In general, the input field can have an arbitrary space-time dependence. Suppose, however, that the medium is stationary and homogeneous. This is a fairly general assumption which holds for most situations of practical interest. Under these conditions, the most general form of linear response is as follows

$$\mathbf{P}^{(1)}(\mathbf{r}, t) = \epsilon_0 \int d\mathbf{r}' \int_{-\infty}^{\infty} dt' \boldsymbol{\chi}^{(1)}(\mathbf{r} - \mathbf{r}', t - t') \vdots \mathbf{E}(\mathbf{r}', t'), \tag{4.41}$$

where we have assumed that the polarization is invariant with respect to translations in space and shifts in time, thanks to stationarity and homogeneity of the medium. By the same token, the second-order nonlinear polarization can be represented as

$$\mathbf{P}^{(2)}(\mathbf{r}, t) = \epsilon_0 \int d\mathbf{r}_1 \int d\mathbf{r}_2 \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2$$
$$\times \boldsymbol{\chi}^{(2)}(\mathbf{r} - \mathbf{r}_1, \mathbf{r} - \mathbf{r}_2; t - t_1, t - t_2) \vdots \mathbf{E}(\mathbf{r}_1, t_1) \mathbf{E}(\mathbf{r}_2, t_2), \tag{4.42}$$

The expressions for higher-order nonlinear polarization fields can be expressed in a similar fashion.

To proceed further, we will assume the medium response to be spatially local; this is a reasonably good approximation for a vast majority of optical media which we will rely on hereafter. In these conditions, the susceptibility tensors can be simplified to

$$\boldsymbol{\chi}^{(1)}(\mathbf{r}-\mathbf{r}', t-t') = \delta(\mathbf{r}-\mathbf{r}')\boldsymbol{\chi}_t^{(1)}(t-t'), \tag{4.43}$$

and

$$\boldsymbol{\chi}^{(2)}(\mathbf{r}-\mathbf{r}_1, \mathbf{r}-\mathbf{r}_2; t-t_1, t-t_2) = \delta(\mathbf{r}-\mathbf{r}_1)\delta(\mathbf{r}-\mathbf{r}_2)\boldsymbol{\chi}_t^{(2)}(t-t_1, t-t_2). \tag{4.44}$$

The corresponding contributions to the polarization field are greatly simplified as well:

$$\mathbf{P}^{(1)}(\mathbf{r}, t) = \epsilon_0 \int_{-\infty}^{\infty} dt' \boldsymbol{\chi}^{(1)}(t-t') \vdots \mathbf{E}(\mathbf{r}, t'), \tag{4.45}$$

$$\mathbf{P}^{(2)}(\mathbf{r}, t) = \epsilon_0 \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 \boldsymbol{\chi}^{(2)}(t-t_1, t-t_2) \vdots \mathbf{E}(\mathbf{r}, t_1)\mathbf{E}(\mathbf{r}, t_2). \tag{4.46}$$

In equations (4.45) and (4.46) we dropped, for brevity, the subscript "t" for the temporal parts of the linear and nonlinear susceptibilities.

The manifest translational invariance of susceptibilities prompts the introduction of Fourier transforms

$$\tilde{\boldsymbol{\chi}}^{(1)}(\omega) = \int_{-\infty}^{\infty} dt \boldsymbol{\chi}^{(1)}(t) e^{i\omega t}, \tag{4.47}$$

and

$$\tilde{\boldsymbol{\chi}}^{(2)}(\omega_1, \omega_2) = \prod_{s=1}^{2} \int_{-\infty}^{\infty} dt_s \boldsymbol{\chi}^{(2)}(t_1, t_2) e^{i \sum_{s=1}^{2} \omega_s t_s}. \tag{4.48}$$

An obvious generalization to the nth order is

$$\tilde{\boldsymbol{\chi}}^{(n)}(\omega_1, \ldots \omega_n) = \prod_{s=1}^{n} \int_{-\infty}^{\infty} dt_s \boldsymbol{\chi}^{(n)}(t_1, \ldots t_n) e^{i \sum_{s=1}^{n} \omega_s t_s}. \tag{4.49}$$

Using Eqs. (4.45) – (4.48), we can obtain in the component form

$$\tilde{P}_i^{(1)}(\mathbf{r}, \omega) = \epsilon_0 \sum_j \tilde{\chi}_{ij}^{(1)}(\omega) \tilde{E}_j(\mathbf{r}, \omega), \tag{4.50}$$

and

$$\tilde{P}_i^{(2)}(\mathbf{r}, \omega_3) = \epsilon_0 \sum_{jk} \int_{-\infty}^{\infty} \frac{d\omega_1}{2\pi} \tilde{\chi}_{ijk}^{(2)}(-\omega_3, \omega_1, \omega_2) \tilde{E}_j(\mathbf{r}, \omega_1) \tilde{E}_k(\mathbf{r}, \omega_2), \tag{4.51}$$

where $\omega_3 = \omega_1 + \omega_2$.

Exercise. 13.1 *Derive Eq. (4.51).*

Generalizing Eq. (4.51) to any order n, we can write down

$$\begin{aligned} \tilde{P}_{j_n}^{(n)}(\mathbf{r}, \omega_n) &= \epsilon_0 \sum_{j_1 \ldots j_{n-1}} \prod_{s=1}^{n-1} \int_{-\infty}^{\infty} \frac{d\omega_s}{2\pi} \tilde{\chi}_{j_n j_1 j_2 \ldots j_{n-1}}^{(n)}(-\omega_n, \omega_1, \omega_2, \ldots \omega_{n-1}) \\ &\quad \times \tilde{E}_{j_1}(\mathbf{r}, \omega_1) \ldots \tilde{E}_{j_{n-1}}(\mathbf{r}, \omega_{n-1}), \end{aligned} \tag{4.52}$$

with $\omega_n = \sum_{s=1}^{n-1} \omega_s$. In particular, the third-order polarization contribution is

$$\tilde{P}_i^{(3)}(\mathbf{r},\omega_4) = \epsilon_0 \sum_{jkl} \prod_{s=1}^{2} \int_{-\infty}^{\infty} \frac{d\omega_s}{2\pi} \tilde{\chi}_{ijkl}^{(3)}(-\omega_4,\omega_1,\omega_2,\omega_3)\tilde{E}_j(\mathbf{r},\omega_1)\tilde{E}_k(\mathbf{r},\omega_2)\tilde{E}_l(\mathbf{r},\omega_3), \tag{4.53}$$

where $\omega_4 = \omega_1 + \omega_2 + \omega_3$.

4.5 Formal properties of nonlinear optical susceptibilities

We now list generic properties of optical susceptibilities which follow from their definitions.

- *Intrinsic permutational symmetry*:

$$\tilde{\chi}_{jj_1\ldots j_n}^{(n)}(-\omega,\omega_1\ldots\omega_n) = P_t \cdot \tilde{\chi}_{jj_1\ldots j_n}^{(n)}(-\omega,\omega_1\ldots\omega_n). \tag{4.54}$$

where $\omega = \sum_{s=1}^{n} \omega_s$, and $P_t \cdot (\ldots)$ stands for a permutation of the n index pairs $(j_1,\omega_1)\ldots(j_n,\omega_n)$ with the exclusion of the pair $(j, -\sum_s \omega_s)$. This property follows at once from the definition of nonlinear optical susceptibilities (4.52): Indeed the indices $(j_1\ldots j_n)$ are dummy ones, and hence the polarization field does not change upon interchanging any pair of them as long as we simultaneously exchange the corresponding frequencies.

$$\text{Example: } \tilde{\chi}_{ijk}^{(2)}(-\omega,\omega_1,\omega_2) = \tilde{\chi}_{ikj}^{(2)}(-\omega,\omega_2,\omega_1).$$

- *Reality of χ in the time-domain*:

The reality of $\boldsymbol{\chi}^{(n)}$ in time domain implies the following relation in the Fourier domain

$$\tilde{\chi}_{jj_1\ldots j_n}^{(n)*}(-\omega,\omega_1,\ldots,\omega_n) = \tilde{\chi}_{jj_1\ldots j_n}^{(n)}(\omega,-\omega_1,\ldots,-\omega_n), \tag{4.55}$$

where $*$ denotes, as usual, complex conjugation.

$$\text{Example: } \tilde{\chi}_{ijk}^{(2)*}(-\omega,\omega_1,\omega_2) = \tilde{\chi}_{ijk}^{(2)}(\omega,-\omega_1,-\omega_2).$$

Exercise 4.3. *Derive Eq. (4.55).*

- *Causality*:

For the response of a physical medium to be causal, the polarization field must be equal to zero at any instant before the electric field is applied, which implies, in accord with Eq.(4.52) that

$$\chi_{jj_1\ldots j_n}^{(n)}(t-\tau_1,\ldots t-\tau_n) = 0, \qquad \text{for any} \quad \tau_s > t. \tag{4.56}$$

Let us now exhibit very tangible constraints on the functional form of the real and imaginary parts of the susceptibility functions in the Fourier domain, stemming from causality.

We begin by considering the linear susceptibility. It follows from Eq. (4.56) that a causal linear response function must obey

$$\boldsymbol{\chi}^{(1)}(\tau) = \boldsymbol{\chi}^{(1)}(\tau)\theta(\tau), \tag{4.57}$$

where $\theta(\tau)$ is a Heaviside step function defined as

$$\theta(\tau) = \begin{cases} 1 & \tau \geq 0, \\ 0 & \tau < 0. \end{cases} \tag{4.58}$$

On introducing Fourier transforms of $\boldsymbol{\chi}$ and θ by the expressions

$$\tilde{\boldsymbol{\chi}}^{(1)}(\omega) = \int_{-\infty}^{\infty} d\tau\, \boldsymbol{\chi}^{(1)}(\tau) e^{i\omega\tau}, \tag{4.59}$$

and

$$\tilde{\theta}(\omega) = \int_{-\infty}^{\infty} d\tau\, \theta(\tau) e^{i\omega\tau}, \tag{4.60}$$

we conclude from Eq. (4.57) that

$$\tilde{\boldsymbol{\chi}}^{(1)}(\omega) = \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi} \tilde{\boldsymbol{\chi}}^{(1)}(\omega')\tilde{\theta}(\omega - \omega'). \tag{4.61}$$

Recall further that

$$\tilde{\theta}(\omega - \omega') = \mathcal{P}\left[\frac{1}{i(\omega - \omega')}\right] + \pi\delta(\omega - \omega'), \tag{4.62}$$

where $\mathcal{P}$ stands for a principal value, excluding the singularity in the denominator. It follows from Eqs. (4.61) and (4.62), after simple algebra, that

$$\tilde{\boldsymbol{\chi}}^{(1)}(\omega) = \frac{1}{\pi i}\mathcal{P}\int_{-\infty}^{\infty} d\omega'\, \frac{\tilde{\boldsymbol{\chi}}^{(1)}(\omega')}{\omega - \omega'}. \tag{4.63}$$

Eq. (4.63) implies that real and imaginary parts of the linear susceptibility tensor are related vie the following *Kramers-Kronig* relations

$$\mathrm{Re}\,\tilde{\boldsymbol{\chi}}^{(1)}(\omega) = \frac{1}{\pi}\mathcal{P}\int_{-\infty}^{\infty} d\omega'\, \frac{\mathrm{Im}\,\tilde{\boldsymbol{\chi}}^{(1)}(\omega')}{\omega - \omega'}, \tag{4.64}$$

and

$$\mathrm{Im}\,\tilde{\boldsymbol{\chi}}^{(1)}(\omega) = -\frac{1}{\pi}\mathcal{P}\int_{-\infty}^{\infty} d\omega'\, \frac{\mathrm{Re}\,\tilde{\boldsymbol{\chi}}^{(1)}(\omega')}{\omega - \omega'}. \tag{4.65}$$

Relations (4.64) and (4.65) not only impose a constraint on the functional form of the real and imaginary parts of the linear susceptibility tensor, but they also enable one to

reconstruct the real part – describing dispersion – from the imaginary one, which is much easier to measure as it relates to absorption in the medium.

Kramers-Kronig relations can also be derived for some second-order susceptibilities. In particular, starting from the causality condition

$$\boldsymbol{\chi}^{(2)}(\tau_1, \tau_2) = \boldsymbol{\chi}^{(2)}(\tau_1, \tau_2)\theta(\tau_1)\theta(\tau_2), \tag{4.66}$$

and following the same line of argument as above, we obtain

$$\tilde{\boldsymbol{\chi}}^{(2)}(-\omega_3, \omega_1, \omega_2) = \frac{1}{\pi i}\mathcal{P}\int_{-\infty}^{\infty} d\omega_1' \, \frac{\tilde{\boldsymbol{\chi}}^{(2)}(-\omega_3', \omega_1', \omega_2)}{\omega_1 - \omega_1'}. \tag{4.67}$$

Here $\omega_3 = \omega_1 + \omega_2$ and $\omega_3' = \omega_1' + \omega_2$. This process is referred to as a sum-frequency generation. By the same token, the Kramers-Kronig relations for a difference-frequency generation are

$$\tilde{\boldsymbol{\chi}}^{(2)}(-\omega_3, \omega_1, -\omega_2) = \frac{1}{\pi i}\mathcal{P}\int_{-\infty}^{\infty} d\omega_2' \, \frac{\tilde{\boldsymbol{\chi}}^{(2)}(-\omega_3', \omega_1, -\omega_2')}{\omega_2 - \omega_2'}, \tag{4.68}$$

where in this case, $\omega_3 = \omega_1 - \omega_2$ and $\omega_3' = \omega_1 - \omega_2'$.

Exercise 4.4. *Derive Eqs. (4.67) and (4.68) .*

Exercise 4.5.* *Consider a degenerate case of the sum-frequency generation, $\omega_1 = \omega_2 = \omega$, and derive the following Kramers-Kronig relations*

$$\tilde{\boldsymbol{\chi}}^{(2)}(-2\omega, \omega, \omega) = \frac{1}{\pi i}\mathcal{P}\int_{-\infty}^{\infty} d\omega' \, \frac{\tilde{\boldsymbol{\chi}}^{(2)}(-2\omega', \omega', \omega')}{\omega - \omega'}. \tag{4.69}$$

This case corresponds to an important second-order nonlinear process we will study in detail later on – it is referred to as the second-harmonic generation.

Unfortunately, no general Kramers-Kronig relations can be derived for higher-order nonlinear susceptibilities. Moreover, there are nonlinear processes for which no Kramers-Kronig relations exist, one of the most prominent cases being the self-focusing/self-defocusing process – specified by $\chi^{(3)}(-\omega, \omega, -\omega, \omega)$ – which is the most common nonlinear process in isotropic media with inversion symmetry.

The symmetry properties of nonlinear susceptibilities we have studied so far hold quite generally. In addition, there are other symmetry properties of $\boldsymbol{\chi}$ which depend on the symmetries of underlying physical systems. First, consider the multitude of orthogonal transformations – such as rotations, translations and inversions – that leave the medium unchanged. It follows that the corresponding susceptibility tensor of any rank must be invariant with respect to such transformations, implying for any n

$$\chi^{(n)}_{ii_1\ldots i_n} = \sum_{jj_1\ldots j_n} \mathcal{T}_{ij}\mathcal{T}_{i_1j_1}\ldots\mathcal{T}_{i_nj_n}\,\chi^{(n)}_{jj_1\ldots j_n}, \tag{4.70}$$

where the summation over the dummy indices is implied as usual. For instance,

$$\chi^{(1)}_{ij} = \sum_{kl} \mathcal{T}_{ik}\mathcal{T}_{jl}\,\chi^{(1)}_{kl}, \tag{4.71}$$

or

$$\chi^{(2)}_{ijk} = \sum_{lsm} \mathcal{T}_{is}\mathcal{T}_{jl}\mathcal{T}_{km}\,\chi^{(2)}_{slm}, \tag{4.72}$$

and so on.

Exercise 4.6. *A rotation with respect to the z-axis can be described by the matrix*

$$T_{ij} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Assume the medium is invariant with respect to rotations by $\theta = \pi/2$. Determine the constraints on the components of $\boldsymbol{\chi}^{(1)}$ imposed in this case.

One of the most important orthogonal transformations is *inversion* such that for an every point in the medium $\mathbf{r} \to -\mathbf{r}$ implying $\mathcal{T}_{ij} = -\delta_{ij}$. It follows at once from Eq. (4.70) that if the medium is symmetric with respect to inversions – i.e., if it has an inversion center – then for any susceptibility tensor of *odd* rank , or for an *even* $n = 2k$, we obtain

$$\chi^{(2k)}_{ii_1\ldots i_{2k}} = -\chi^{(2k)}_{ii_1\ldots i_{2k}} = 0. \tag{4.73}$$

In particular, in media with the inversion centers the lowest-order nonlinear response is cubic, described by $\chi^{(3)}_{ijkl}$. Such inversion symmetric media are referred to as *centrosymmetric*. Most gases and liquids as well as many solids possess such properties.

Another important constraint is imposed by requiring that media be lossless. In lossless media, equations of motions are symmetric with respect to time reversal – there are no losses and the microscopic evolution can in principle be reversed. Under such conditions,

$$\boldsymbol{\chi}^{(n)}(\tau_1 \ldots \tau_n) = \boldsymbol{\chi}^{(n)}(-\tau_1 \ldots -\tau_n). \tag{4.74}$$

It can then be readily inferred from Eq. (4.49) that

$$\tilde{\chi}^{(n)}_{jj_1\ldots j_n}(-\omega, \omega_1, \ldots \omega_n) = \chi^{(n)*}_{jj_1\ldots j_n}(-\omega, \omega_1, \ldots \omega_n), \tag{4.75}$$

that is a Fourier image of $\boldsymbol{\chi}$ is real.

Exercise 4.7. *Derive Eq. (4.75).*

Exercise 4.8. *Show that in lossless media ϵ_{ij} must be symmetric.*

Moreover, in lossless nonlinear media, there is an *overall permutation symmetry* of the susceptibility tensor, similar to that expressed in Eq. (4.54), except the pair $(j, -\sum_s \omega_s)$ is included.

$$\text{Example: } \tilde{\chi}^{(3)}_{ijkl}(-\omega_4, \omega_1, \omega_2, \omega_3) = \tilde{\chi}^{(3)}_{jlik}(\omega_1, \omega_3, -\omega_4, \omega_2).$$

Finally, if all frequencies involved in the interaction are *well below* the lowest resonant frequency of the medium, there exists a permutation symmetry of the Cartesian indices alone, known as the *Kleinman symmetry*.

$$\text{Example: } \tilde{\chi}^{(2)}_{ijk}(-\omega_3, \omega_1, \omega_2) = \tilde{\chi}^{(2)}_{jki}(-\omega_3, \omega_1, \omega_2) = \tilde{\chi}^{(2)}_{kij}(-\omega_3, \omega_1, \omega_2).$$

We stress though that Kleinman's symmetry is only an approximation valid far from any internal resonances where dispersive properties of nonlinear media are negligible such that one can virtually neglect frequency dependence of the nonlinear susceptibilities. The Kleinman symmetry breaks down, for instance, if there is an absorption band sandwiched between a pair of frequencies involved with a nonlinear interaction. In the latter case, dispersive properties of the medium would be important at those frequencies near the absorption band.

4.6 Nonlinear wave equation approach: Second-order processes

4.6.1 Classical coupled-wave equations

We now proceed to deriving general nonlinear wave equations governing second-order nonlinear processes. To this end, we represent the polarization field $\mathbf{P}$ as

$$\mathbf{P} = \mathbf{P}_L + \mathbf{P}_{NL}, \tag{4.76}$$

where the first and second terms correspond to linear and nonlinear contributions, respectively. Eliminating $\mathbf{H}$ from Eqs. (**??**) – (**??**), we obtain a nonlinear wave equation for the electric field alone in the form

$$\nabla \times (\nabla \times \mathbf{E}) = -\epsilon_0 \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} - \mu_0 \frac{\partial^2 \mathbf{P}}{\partial t^2}. \tag{4.77}$$

Using a well-known vector identity, $\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$, as well as Eq. (4.76), we arrive at the equation

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\mu_0 \frac{\partial^2 (\epsilon_0 \mathbf{E} + \mathbf{P}_L)}{\partial t^2} - \mu_0 \frac{\partial^2 \mathbf{P}_{NL}}{\partial t^2}. \tag{4.78}$$

Further, we can introduce a linear flux density by the expression

$$\mathbf{D}_L = \epsilon_0 \mathbf{E} + \mathbf{P}_L = \epsilon \otimes \mathbf{E}, \tag{4.79}$$

where the symbol $\otimes$ denotes the temporal convolution. Using Eq. (4.79) as well as $\epsilon_0 \mu_0 = c^{-2}$, we can transform Eq. (4.78) into

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\frac{1}{\epsilon_0 c^2} \frac{\partial^2 \mathbf{D}_L}{\partial t^2} - \mu_0 \frac{\partial^2 \mathbf{P}_{NL}}{\partial t^2}. \tag{4.80}$$

A further simplification of Eq. (4.80) is possible. To this end, we employ Eq. (4.79) to rewrite the third Maxwell equation (**??**) as $\nabla \cdot \mathbf{E} \approx -(\nabla \epsilon / \epsilon)\mathbf{E}$, and, provided $|\nabla \epsilon / \epsilon| \ll 1$, we can conclude that $\nabla \cdot \mathbf{E} \approx 0$. The preceding line of argument tacitly implies that linear optical properties of the medium are weakly inhomogeneous and isotropic – or weakly anisotropic at worst – which constitutes the so-called weak guidance approximation. Hereafter, we will assume the weak guidance approximation

holds true. Under the circumstances, the l.h.s. of Eq. (4.80) is considerably simplified, yielding

$$\nabla^2 \mathbf{E} - \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \mathbf{D}_L}{\partial t^2} = \mu_0 \frac{\partial^2 \mathbf{P}_{NL}}{\partial t^2}. \tag{4.81}$$

Equation (4.81) is a classical nonlinear wave equation, governing electromagnetic wave propagation in nonlinear optical media.

The polarization field $\mathbf{P}_{NL}$ plays the role of a driving source which generates optical fields oscillating at new frequencies. To illustrate this point and to derive a set of coupled wave equations governing second-order nonlinear interactions, we will consider a cw optical field which is a linear superposition of monochromatic components:

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{2} \sum_s \tilde{\mathbf{E}}(\mathbf{r}, \omega_s) e^{-i\omega_s t} + c.\ c. \tag{4.82}$$

Such an electric field gives rise to the polarization field in the form

$$\mathbf{P}_{NL}(\mathbf{r}, t) = \frac{1}{2} \sum_s \tilde{\mathbf{P}}_{NL}(\mathbf{r}, \omega_s) e^{-i\omega_s t} + c.\ c., \tag{4.83}$$

On substituting from Eqs. (4.82) and (4.83) into (4.81, we obtain the wave equation in the space-frequency domain as

$$\nabla^2 \tilde{\mathbf{E}} + k^2(\omega_s) \tilde{\mathbf{E}} = -\mu_0 \omega_s^2 \tilde{\mathbf{P}}_{NL}. \tag{4.84}$$

Let us now consider a linearly polarized optical beam whose spatial profile changes very slowly – at the wavelength scale – in the plane, transverse to the propagation direction of the carrier wave which we choose to coincide with the z-axis. Under such a *paraxial* approximation, we seek an electric field in the form

$$\tilde{\mathbf{E}}(\mathbf{r}, \omega_s) = \mathbf{e}(\omega_s) \mathcal{E}(\mathbf{r}_\perp, z, \omega_s) e^{ik_s z}, \tag{4.85}$$

which induces the polarization field such that

$$\tilde{\mathbf{P}}_{NL}(\mathbf{r}, \omega_s) = \mathbf{e}(\omega_s) \mathcal{P}_{NL}(\mathbf{r}_\perp, z, \omega_s) e^{ik_s z}, \tag{4.86}$$

where

$$k^2(\omega_s) = \epsilon(\omega_s) \frac{\omega_s^2}{c^2}. \tag{4.87}$$

Substituting from Eqs. (4.85) – (4.87), and using the *slowly-varying envelope approximation* (SVEA),

$$\frac{\partial \mathcal{E}}{\partial z} \ll k_s \mathcal{E}; \qquad \frac{\partial^2 \mathcal{E}}{\partial z^2} \ll k_s^2 \mathcal{E}, \tag{4.88}$$

we arrive at the paraxial wave equation

$$2ik_s \frac{\partial \mathcal{E}}{\partial z} + \nabla_\perp^2 \mathcal{E} = -\mu_0 \omega_s^2 \mathcal{P}_{NL}. \tag{4.89}$$

Our treatment has been general so far. We will now specialize to the second-order processes. Recall that

$$\tilde{P}_i^{(2)}(\mathbf{r},\omega_s) = \epsilon_0 c^{(2)}(\omega_1,\omega_2)\sum_{jk}\tilde{\chi}_{ijk}^{(2)}(-\omega_s;\omega_1,\omega_2)\tilde{E}_j(\mathbf{r},\omega_1)\tilde{E}_k(\mathbf{r},\omega_2), \tag{4.90}$$

with $\omega_s = \omega_1 + \omega_2$ and

$$c^{(2)}(\omega_1,\omega_2) = \begin{cases} 1, & \omega_s \neq 2\omega_1; \\ 1/2, & \omega_s = 2\omega_1. \end{cases} \tag{4.91}$$

Using Eqs. (4.90) and (4.86), we obtain for the slowly-varying second-order polarization field the expression

$$\begin{aligned} \mathcal{P}^{(2)}(\mathbf{r},\omega_s) &= \epsilon_0 c^{(2)}(\omega_1,\omega_2)\sum_{ijk}\tilde{\chi}_{ijk}^{(2)}(-\omega_s;\omega_1,\omega_2)e_i(\omega_s) \\ &\quad \times e_j(\omega_1)e_k(\omega_2)\mathcal{E}(\mathbf{r}_\perp,\omega_1)\mathcal{E}(\mathbf{r}_\perp,\omega_2)e^{i\Delta kz}, \end{aligned} \tag{4.92}$$

where

$$\Delta k \equiv k(\omega_1) + k(\omega_2) - k(\omega_s). \tag{4.93}$$

Utilizing Eq. (4.92) and introducing

$$\chi_{eff}^{(2)}(-\omega_s;\omega_1,\omega_2) \equiv c^{(2)}(\omega_1,\omega_2)\sum_{ijk}\tilde{\chi}_{ijk}^{(2)}(-\omega_s;\omega_1,\omega_2)e_i(\omega_s)e_j(\omega_1)e_k(\omega_2), \tag{4.94}$$

we finally arrive at the paraxial wave equation governing the second-order nonlinear processes:

$$\frac{\partial\mathcal{E}_s}{\partial z} - \frac{i}{2k(\omega_s)}\nabla_\perp^2\mathcal{E}_s = \frac{i\omega_s^2}{2k(\omega_s)c^2}\chi_{eff}^{(2)}(-\omega_s;\omega_1,\omega_2)\mathcal{E}_1\mathcal{E}_2 e^{i\Delta kz}. \tag{4.95}$$

Here we adopted the convention

$$\mathcal{E}_j(\mathbf{r}_\perp,z,-\omega_j) = \mathcal{E}_j^*(\mathbf{r}_\perp,z,\omega_j).$$

introduced short-hand notations

$$\mathcal{E}_j \equiv \mathcal{E}(\mathbf{r}_\perp,z,\omega_j); \qquad j = s,1,2.$$

4.7 Second-harmonic generation

4.7.1 Coupled wave equations and phase matching considerations

The process of second harmonic generation involves the interaction of two waves at frequency ω to produce a wave with the frequency 2ω. It is schematically illustrated in Fig. 1 below.

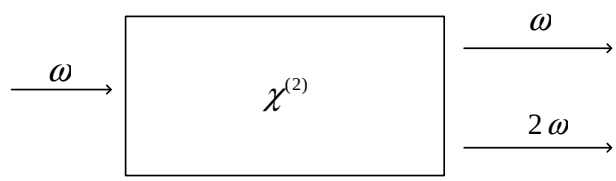

Figure 4.7: Illustrating the second harmonic generation.

The coupled wave equations governing the second harmonic generation (SHG) in lossless media can be obtained directly from the general coupled-mode equations derived in the previous Lecture by specializing to the case of two identical mixing frequencies. The resulting wave equations for the fundamental $\mathcal{E}_\omega$ and the second harmonic $\mathcal{E}_{2\omega}$ fields are

$$\frac{\partial \mathcal{E}_\omega}{\partial z} - \frac{i}{2k_\omega}\nabla_\perp^2 \mathcal{E}_\omega = \frac{i\omega^2}{2k_\omega c^2}\chi_{eff}^{(2)}(-\omega, 2\omega, -\omega)\mathcal{E}_{2\omega}\mathcal{E}_\omega^* e^{-i\Delta kz}. \tag{4.96}$$

and

$$\frac{\partial \mathcal{E}_{2\omega}}{\partial z} - \frac{i}{2k_{2\omega}}\nabla_\perp^2 \mathcal{E}_{2\omega} = \frac{i4\omega^2}{2k_{2\omega} c^2}\chi_{eff}^{(2)}(-2\omega, \omega, \omega)\, \mathcal{E}_\omega^2\, e^{i\Delta kz}, \tag{4.97}$$

where the wave number mismatch is now defined as

$$\Delta k = 2k_\omega - k_{2\omega}. \tag{4.98}$$

In Eqs. (4.96) – (4.98), we have introduced the notations

$$k_\omega = \frac{\omega n(\omega)}{c}, \qquad k_{2\omega} = \frac{2\omega n(2\omega)}{c}. \tag{4.99}$$

It follows from general properties of susceptibilities in the absence of losses that

$$\chi_{eff}^{(2)}(-\omega, 2\omega, -\omega) = 2\chi_{eff}^{(2)}(-2\omega, \omega, \omega) \equiv \chi_{eff}^{(2)}. \tag{4.100}$$

Using (4.100), we can transform the SHG coupled wave equations in the plane wave geometry to

$$\frac{d\mathcal{E}_\omega}{dz} = \frac{i\omega^2}{2k_\omega c^2}\chi_{eff}^{(2)}\, \mathcal{E}_{2\omega}\mathcal{E}_\omega^* e^{-i\Delta kz}. \tag{4.101}$$

and

$$\frac{d\mathcal{E}_{2\omega}}{dz} = \frac{i\omega^2}{k_{2\omega} c^2}\chi_{eff}^{(2)}\, \mathcal{E}_\omega^2\, e^{i\Delta kz}. \tag{4.102}$$

Let us now study the second harmonic generation in the undepleted pump approximation, which implies that the power of the fundamental wave is high enough and the efficiency of the second harmonic generation is low enough that we can neglect the power depletion of the fundamental wave. As the efficiency η_{SHG} of the second

harmonic generation can be defined as the ratio of the second harmonic intensity at the output to the input intensity of the fundamental,

$$\eta_{SHG} \equiv \frac{I_{2\omega}(L)}{I_{\omega}(0)}, \tag{4.103}$$

we can define a quantitative criterion for the undepleted pump approximation to hold:

$$\eta_{SHG} \ll 1. \tag{4.104}$$

In the undepleted pump approximation, Eq. (4.102) can be integrated at once with the result

$$\mathcal{E}_{2\omega}(L) = \frac{i\omega^2}{k_{2\omega}c^2}\chi^{(2)}_{eff}\,\mathcal{E}^2_{\omega}\frac{e^{i\Delta kL}-1}{i\Delta k} = \frac{\omega^2 L\chi^{(2)}_{eff}}{k_{2\omega}c^2}\mathcal{E}^2_{\omega}e^{i\Delta kL/2}\frac{e^{i\Delta kL/2}-e^{-i\Delta kL/2}}{2i(\Delta kL/2)}, \tag{4.105}$$

where L is the length of the interaction region and $\mathcal{E}_{\omega} = const$. Further, equation (4.105) can be simplified as

$$\mathcal{E}_{2\omega}(L) = \frac{\omega^2 L\chi^{(2)}_{eff}\mathcal{E}^2_{\omega}}{k_{2\omega}c^2}e^{i\Delta kL/2}\frac{\sin(\Delta kL/2)}{\Delta kL/2}. \tag{4.106}$$

It can be readily inferred from Eq. (4.106) that the intensity of the second harmonic is given by

$$I_{2\omega}(L) = \frac{\omega^2 L^2\chi^{(2)2}_{eff}I^2_{\omega}}{2\epsilon_0 n_{2\omega}n^2_{\omega}c^3}\mathrm{sinc}^2\left(\frac{\Delta kL}{2}\right), \tag{4.107}$$

where we defined

$$\mathrm{sinc}(x) \equiv \frac{\sin x}{x}. \tag{4.108}$$

The analysis of Eq. (4.107) reveals that if the phases of the fundamental and second harmonic waves are matched, the intensity of the second harmonic is proportional to the square of the interaction length, $I_{2\omega}(L) \propto L^2$. Physically, it can be interpreted by observing that if all N polarized atomic dipoles in the interaction volume – whose total number is proportional to L – radiate in phase, their resulting fields interfere constructively; consequently the total intensity of the second harmonic is such that $I_{2\omega}(L) \propto N^2 \propto L^2$. On the other hand, if the phase matching condition (4.109) is not met, the efficiency of the second harmonic generation decreases dramatically, as is shown in Fig. 2.

Let us now discuss the efficiency of the SHG process. It follows from Eqs. (4.104) and (4.107) that under the best possible condition of the perfect *phase matching*

$$\Delta k = 0, \tag{4.109}$$

the undepleted pump approximation is valid provided

$$\eta_{SHG} = \frac{\omega^2 L^2\chi^{(2)2}_{eff}I_{\omega}}{2\epsilon_0 n_{2\omega}n^2_{\omega}c^3} \ll 1, \tag{4.110}$$

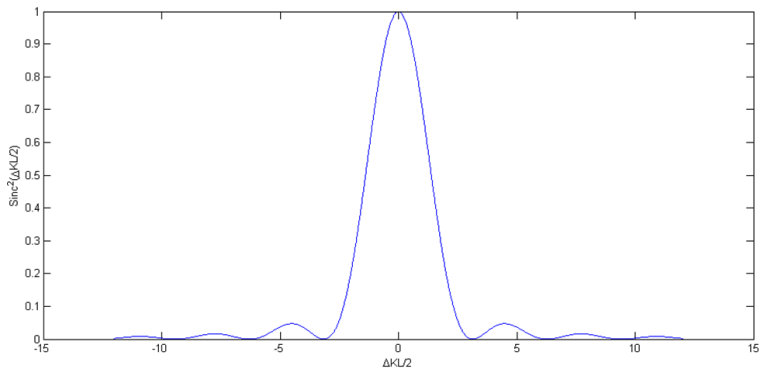

Figure 4.8: Second harmonic output as a function of the interaction length in the undepleted pump approximation

which can be physically interpreted as a limitation on the allowed interaction length for a (large) given power of the fundamental wave: the power depletion of the fundamental can no longer be neglected for sufficiently large interaction lengths. To estimate the efficiency of the SHG under typical experimental conditions, we can estimate the intensity of the fundamental as

$$I_\omega = \frac{P}{\pi w_0^2}, \tag{4.111}$$

where P is the laser power and w_0 is the spot size of the laser output beam, which we choose by stipulating that the corresponding diffraction length, $L_d \simeq k w_0^2$, be much greater that the interaction length,

$$L_d \gg L, \tag{4.112}$$

for the plane wave approximation to hold. Using typical values, for moderate-to-high power lasers $P \sim 1$ W, and $\chi^2_{eff} \sim 5 \times 10^{-23}$ m^2/V^2, for $LiNbO_3$, say; with the other parameters being chosen as follows: $L \sim 1$ cm, $n_\omega \sim n_{2\omega} \sim 2$, $\lambda \sim 5 \times 10^{-5}$ cm, and the spot size $w_0 \sim 100$ μm, such that $L_d \sim 10$ cm, we obtain the order-of-magnitude estimate as $\eta_{SHG} \sim 10^{-3} \ll 1$. Clearly, the undepleted pump approximation is a good one even for relatively high power laser sources in the plane wave geometry. To increase the SHG conversion efficiency, it is advised that (a) pulsed lasers be employed to augment the input power and (b) source light beam be tightly focused into the interaction volume to significantly increase the intensity of the fundamental input wave. In general, the analysis of the SHG with such tightly focused laser beams requires a more careful consideration of diffraction effects. With this in mind, however, we could still make a rough order-of-magnitude estimate of the efficiency using Eq. (4.110) by taking the spot size of a focused beam to be $w_0 \sim 10$ μm, even though $L_d \ll L$. The resulting efficiency is of the order of 10%, which is already quite an improvement.

Due to the importance of phase matching, we briefly discuss the ways of realizing the condition (4.109), which, when translated in terms of the refractive indices, implies

$$n(2\omega) = n(\omega). \tag{4.113}$$

First, we note that the requirement (4.113) cannot be satisfied in an isotropic medium because of frequency dispersion: Typically, the refractive index of a nonlinear medium far below absorption resonances is a monotonically increasing function of frequency, a phenomenon referred to as normal dispersion. Thus isotropic media are in general not phase matchable.

Phase matching can be realized in anisotropic media, which is referred to as **birefringence phase matching**. As we previously mentioned, the distribution of the ordinary wave vectors is spherically symmetric–which is graphically illustrated in Fig. 3–where we assumed, for simplicity, the wave vector lies in the xz-plane–and one can introduce the corresponding frequency-dependent refractive index $n_o(\omega)$ by the expression

$$n_o(\omega) \equiv \frac{k_o c}{\omega} = \sqrt{\epsilon_\perp(\omega)}. \tag{4.114}$$

The extraordinary wave vector, on the other hand, does depend on the propagation direction, and the associated extraordinary refractive index is given by

$$n_e(\theta, \omega) \equiv \frac{k_e c}{\omega} = \left(\frac{\sin^2\theta}{\epsilon_\perp(\omega)} + \frac{\cos^2\theta}{\epsilon_\parallel(\omega)} \right)^{-1/2}. \tag{4.115}$$

The surface $n_e(\theta, \omega) = const$ is, in general, an ellipsoid, but it reduces to an ellipse if we restrict the extraordinary wave vector to lie in the $xz-$ plane, see Fig. 3.

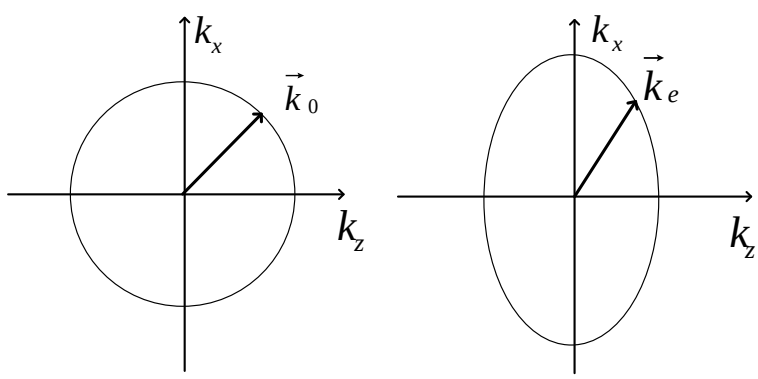

Figure 4.9: Graphical representation of the wave vectors of ordinary (left) and extraordinary (right) waves in a uniaxial crystal.

Assume now that the fundamental is an ordinary wave and the second harmonic is an extraordinary one. It can then be inferred from Fig. 4 that provided the extraordinary refractive index for the SH along the crystal axis is smaller than the ordinary refractive index of the fundamental, which can be mathematically expressed by the inequality

$$\epsilon_\parallel(2\omega) < \epsilon_\perp(\omega), \tag{4.116}$$

the phase matching is possible at the angle θ_* which can be determined from Eqs. (4.113),

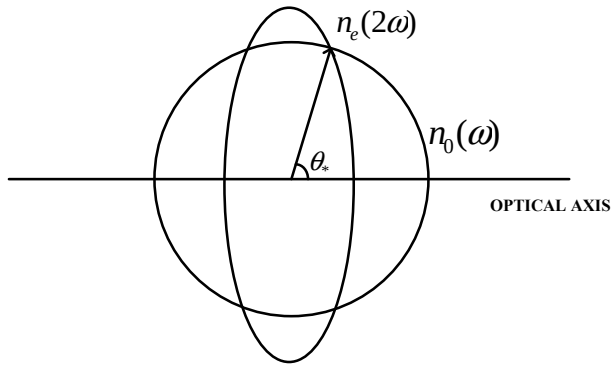

Figure 4.10: Illustrating phase matching for the SHG in uniaxial crystals.

(4.114) and (4.115) to be

$$\tan\theta_* = \sqrt{\frac{\frac{1}{\epsilon_{\parallel}(2\omega)} - \frac{1}{\epsilon_{\perp}(\omega)}}{\frac{1}{\epsilon_{\perp}(\omega)} - \frac{1}{\epsilon_{\perp}(2\omega)}}}. \tag{4.117}$$

Unfortunately, whenever the angle between the ordinary and extraordinary wave vectors is other than 90^{deg}, a spatial walkoff accrues on propagation of the two waves as a consequence of directional mismatch between the Poynting vector and propagation direction of an extraordinary wave. The walkoff reduces spatial overlap between the polarization modes, thereby drastically reducing the SHG efficiency. Fortunately, some nonlinear crystals, such as lithium niobate, have a pronounced dependence of their birefringence on the temperature. Thus, one can achieve phase matching by keeping the angle between the modes fixed at 90^{deg} and varying the temperature of the crystal. This is called **temperature phase matching**.

In the situations when neither birefringence nor temperature phase matching is possible, the most powerful phase matching technique is used, the so-called **quasi-phase-matching**. The technique involves periodically polling $\chi^{(2)}$ samples to modulate the second-order susceptibility. The latter can then be expanded in a Fourier series

$$\chi^{(2)}(z) = \sum_{m=-\infty}^{\infty} \chi_m^{(2)} e^{i2\pi mz/\Lambda},$$

where Λ is a spatial period of the structure. The phase mismatch is then modified to $\Delta k_{eff} = \Delta k - 2\pi m/\Lambda$. As $\chi_m^{(2)}$ decreases with m, reducing the SH intensity, it is preferable to work with $m = 1$ harmonic and choose the period Λ to phase match the interaction, i.e.,

$$\Lambda = 2\pi/\Delta k.$$

If Δk is so large, $\Delta k \sim k$ that it is impossible to attain perfect phase matching, quasi-phase-matching allows to extend, at least, the effective interaction length to

$$L_{eff} = L(1 + 2\pi/\Lambda\Delta k),$$

where the smallest available Λ should be used.

4.7.2 Second-harmonic generation: Beyond the undepleted pump approximation

In this section, we describe the second harmonic generation process under general conditions. To this end, we rewrite the governing coupled wave equations in the form

$$\frac{d\mathcal{E}_\omega}{dz} = \frac{i\omega^2}{2k_\omega c^2}\chi^{(2)}_{eff}\,\mathcal{E}_{2\omega}\mathcal{E}^*_\omega e^{-i\Delta kz}, \tag{4.118}$$

$$\frac{d\mathcal{E}_{2\omega}}{dz} = \frac{i\omega^2}{k_{2\omega}c^2}\chi^{(2)}_{eff}\,\mathcal{E}^2_\omega\, e^{i\Delta kz}. \tag{4.119}$$

Let us now introduce the total optical intensity of the fundamental and second harmonic waves as

$$I = I_1 + I_2. \tag{4.120}$$

It is convenient to transform to dimensionless real amplitudes $\mathcal{A}$ and phases ϕ, related to the complex amplitudes of the fundamental and second harmonic waves by the expressions

$$\mathcal{E}_\omega = \sqrt{\frac{I}{n_\omega\epsilon_0 c}}\mathcal{A}_\omega e^{i\phi_\omega}, \tag{4.121}$$

and

$$\mathcal{E}_{2\omega} = \sqrt{\frac{I}{n_{2\omega}\epsilon_0 c}}\mathcal{A}_{2\omega} e^{i\phi_{2\omega}}. \tag{4.122}$$

Using the definitions (4.121) and (4.122), one can derive from Eqs. (4.118) and (4.119) the equations for the real amplitudes as

$$\frac{d\mathcal{A}_\omega}{dz} = \frac{\mathcal{A}_\omega\mathcal{A}_{2\omega}}{l}\sin\theta, \tag{4.123}$$

$$\frac{d\mathcal{A}_{2\omega}}{dz} = -\frac{\mathcal{A}^2_\omega}{l}\sin\theta, \tag{4.124}$$

where

$$\theta = 2\phi_\omega - \phi_{2\omega} + \Delta kz, \tag{4.125}$$

and we have introduced the characteristic spatial period l of the power exchange between the fundamental and second harmonic by the expression

$$\frac{1}{l} = \frac{\omega\chi^{(2)}_{eff}}{2c}\sqrt{\frac{I}{n^2_\omega n_{2\omega}\epsilon_0 c}}. \tag{4.126}$$

Similarly, the equations for the phases take the form

$$\frac{d\phi_\omega}{dz} = \frac{\mathcal{A}_{2\omega}}{l}\cos\theta, \tag{4.127}$$

$$\frac{d\phi_{2\omega}}{dz} = \frac{\mathcal{A}_\omega^2}{l\mathcal{A}_{2\omega}} \cos\theta. \tag{4.128}$$

Introducing $\zeta = z/l$, we can cast our equations into the following dimensionless form

$$\frac{d\mathcal{A}_\omega}{d\zeta} = \mathcal{A}_\omega \mathcal{A}_{2\omega} \sin\theta, \tag{4.129}$$

$$\frac{d\mathcal{A}_{2\omega}}{d\zeta} = -\mathcal{A}_\omega^2 \sin\theta, \tag{4.130}$$

$$\frac{d\phi_\omega}{d\zeta} = \mathcal{A}_{2\omega} \cos\theta, \tag{4.131}$$

$$\frac{d\phi_{2\omega}}{d\zeta} = \frac{\mathcal{A}_\omega^2}{\mathcal{A}_{2\omega}} \cos\theta. \tag{4.132}$$

It can be inferred at once from Eqs. (4.131) and (4.132) as well as from Eq. (4.125) that θ obeys the equation

$$\frac{d\theta}{d\zeta} = \Delta s + \left(2\mathcal{A}_{2\omega} - \frac{\mathcal{A}_\omega^2}{\mathcal{A}_{2\omega}}\right) \cos\theta, \tag{4.133}$$

where we have introduced the quantity

$$\Delta s = \Delta k l. \tag{4.134}$$

We can easily see from Eqs. (4.129) and (4.130) that the set possesses the integral of motion

$$\mathcal{A}_\omega^2 + \mathcal{A}_{2\omega}^2 = 1, \tag{4.135}$$

which implies the power conservation in the SHG process in a lossless medium. It then follows from Eqs. (4.129) and (4.130) that

$$\mathcal{A}_{2\omega} = \frac{1}{\sin\theta} \frac{d}{d\zeta} \ln \mathcal{A}_\omega, \tag{4.136}$$

and

$$\frac{\mathcal{A}_\omega^2}{\mathcal{A}_{2\omega}} = -\frac{1}{\sin\theta} \frac{d}{d\zeta} \ln \mathcal{A}_{2\omega}. \tag{4.137}$$

Substituting from the last two equations into Eq. (4.125), we obtain the equation for the phase difference in the form

$$\frac{d\theta}{d\zeta} = \Delta s + \cot\theta \frac{d}{d\zeta} \ln(\mathcal{A}_\omega^2 \mathcal{A}_{2\omega}). \tag{4.138}$$

Hereafter we focus on the perfect phase matching situation, $\Delta s = 0$. In this case, we can transform Eq. (4.138), with the aid of Eq. (4.130) to

$$\frac{d \ln\cos\theta}{d\zeta} = -\frac{d}{d\zeta} \ln(\mathcal{A}_\omega^2 \mathcal{A}_{2\omega}), \tag{4.139}$$

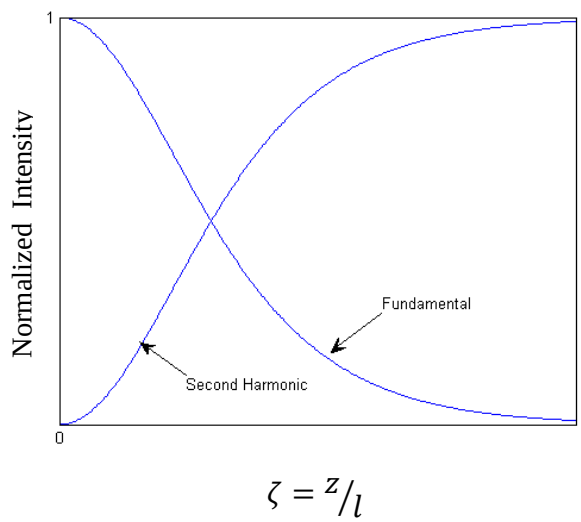

Figure 4.11: Intensity of the fundamental and second harmonic as functions of the interaction distance in the case of perfect phase matching.

which can be integrated at once yielding the second integral of motion as

$$\mathcal{A}_\omega^2 \mathcal{A}_{2\omega} \cos\theta = \Gamma. \tag{4.140}$$

Suppose now that $\Gamma = 0$ implying a fixed phase difference between the FW and SH, $\theta = -\pi/2$. It then follows that the equations of motion for the mode amplitudes simplify to

$$\frac{d\mathcal{A}_\omega}{d\zeta} = \mathcal{A}_\omega \mathcal{A}_{2\omega}, \tag{4.141}$$

$$\frac{d\mathcal{A}_{2\omega}}{d\zeta} = -\mathcal{A}_\omega^2, \tag{4.142}$$

Using Eq. (4.135), we can eliminate the fundamental from Eq. (4.142) resulting in

$$\frac{d\mathcal{A}_{2\omega}}{d\zeta} = -(1 - \mathcal{A}_{2\omega}^2), \tag{4.143}$$

which can be integrated at once yielding

$$\mathcal{A}_{2\omega} = \tanh\zeta; \qquad \mathcal{A}_\omega = \text{sech}\zeta. \tag{4.144}$$

The intensities of the fundamental and second harmonic are displayed in the Fig. 5.

4.8 Sum-frequency generation

4.8.1 Coupled wave equations and their solution in the undepleted pump approximation

In this Lecture, we examine the sum-frequency generation (SFG), which involves mixing a signal wave at frequency ω_1 with a pump wave at frequency ω_2 to yield a harmonic oscillating at $\omega_3 = \omega_1 + \omega_2$, to be referred to as the sum-frequency (SF) wave. The SFG process is schematically illustrated in Fig. 1.

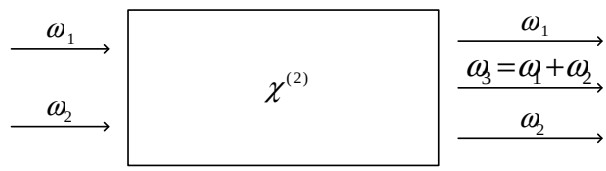

Figure 4.12: Schematic illustration of the sum-frequency generation process.

The wave equations governing the SFG can be readily obtained from the general coupled wave equations, yielding the following set

$$\frac{\partial \mathcal{E}_1}{\partial z} - \frac{i}{2k_1}\nabla_\perp^2 \mathcal{E}_1 = \frac{i\omega_1^2}{2k_1c^2}\chi_{eff}^{(2)}(-\omega_1;\omega_3,-\omega_2)\mathcal{E}_3\mathcal{E}_2^* e^{-i\Delta kz}. \tag{4.145}$$

$$\frac{\partial \mathcal{E}_2}{\partial z} - \frac{i}{2k_2}\nabla_\perp^2 \mathcal{E}_2 = \frac{i\omega_2^2}{2k_2c^2}\chi_{eff}^{(2)}(-\omega_2;\omega_3,-\omega_1)\mathcal{E}_3\mathcal{E}_1^* e^{-i\Delta kz}. \tag{4.146}$$

and

$$\frac{\partial \mathcal{E}_3}{\partial z} - \frac{i}{2k_3}\nabla_\perp^2 \mathcal{E}_3 = \frac{i\omega_3^2}{2k_3c^2}\chi_{eff}^{(2)}(-\omega_3;\omega_1,\omega_2)\,\mathcal{E}_1\mathcal{E}_2\, e^{i\Delta kz}. \tag{4.147}$$

Here $\mathcal{E}_j = \mathcal{E}(\boldsymbol{\rho}, z, \omega_j)$, $k_j = k(\omega_j)$; we also introduced the wave number mismatch Δk

$$\Delta k = k_1 + k_2 - k_3. \tag{4.148}$$

Exercise 4.9. *Using general symmetry properties of the second-order susceptibilities, show that*

$$\chi_{eff}^{(2)}(-\omega_2;\omega_3,-\omega_1) = \chi_{eff}^{(2)*}(-\omega_3;\omega_1,\omega_2), \tag{4.149}$$

and

$$\chi_{eff}^{(2)}(-\omega_1;\omega_3,-\omega_2) = \chi_{eff}^{(2)*}(-\omega_3;\omega_1,\omega_2). \tag{4.150}$$

The situation is further simplified if we neglect diffraction by focusing on a plane wave geometry. In these circumstances and taking account of the properties (4.149) and (4.150) to drop arguments of $\chi_{eff}^{(2)}$, we can reduce Eqs. (4.145) – (4.147) to the set of ODEs in the form

$$\frac{d\mathcal{E}_1}{dz} = \frac{i\omega_1^2}{2k_1c^2}\chi_{eff}^{(2)*}\mathcal{E}_3\mathcal{E}_2^* e^{-i\Delta kz}, \tag{4.151}$$

$$\frac{d\mathcal{E}_2}{dz} = \frac{i\omega_2^2}{2k_2c^2}\chi_{eff}^{(2)*}\mathcal{E}_3\mathcal{E}_1^* e^{-i\Delta kz}, \tag{4.152}$$

$$\frac{d\mathcal{E}_3}{dz} = \frac{i\omega_3^2}{2k_3c^2}\chi_{eff}^{(2)}\mathcal{E}_1\mathcal{E}_2 e^{i\Delta kz}. \tag{4.153}$$

Although Eqs. (4.151) – (4.153) can be solved in general, the solution is very complicated and not too instructive. Instead, we will study the SFG process in the so-called *undepleted pump approximation*, i.e, when the amplitude of the pump wave $\mathcal{E}_2$ is so much larger than those of the other waves that we can neglect the pump depletion – that is we will assume $\mathcal{E}_2 = const$ – which enables us to rewrite Eqs. (4.151) – (4.153) as

$$\frac{d\mathcal{E}_1}{dz} = \kappa_1\mathcal{E}_3 e^{-i\Delta kz}, \tag{4.154}$$

and

$$\frac{d\mathcal{E}_3}{dz} = \kappa_3\mathcal{E}_1 e^{i\Delta kz}. \tag{4.155}$$

Here we introduced the notations

$$\kappa_1 = \frac{i\omega_1^2\chi_{eff}^{(2)*}}{2k_1c^2}\mathcal{E}_2^*, \qquad \kappa_3 = \frac{i\omega_3^2\chi_{eff}^{(2)}}{2k_3c^2}\mathcal{E}_2. \tag{4.156}$$

Let us then assume perfect phase matching, $\Delta k = 0$. In this case, we can eliminate one of the fields from Eqs. (4.154) and (4.155) in favor of the other, reducing the set to a second-order ODE; for instance,

$$\frac{d^2\mathcal{E}_1}{dz^2} + \kappa_{eff}^2\mathcal{E}_1 = 0, \tag{4.157}$$

with

$$\kappa_{eff}^2 = -\kappa_1\kappa_3 = \frac{\omega_1^2\omega_3^2|\chi_{eff}^{(2)}|^2|\mathcal{E}_2|^2}{4k_1k_3c^4}. \tag{4.158}$$

A general solution to (4.157) is

$$\mathcal{E}_1 = C_1\cos\kappa_{eff}z + C_2\sin\kappa_{eff}z, \tag{4.159}$$

where C_1 and C_2 are arbitrary constants. It then follows from Eqs. (4.154), (4.155) and (4.159) that

$$\mathcal{E}_3 = -\frac{\kappa_{eff}C_1}{\kappa_1}\sin\kappa_{eff}z + \frac{\kappa_{eff}C_2}{\kappa_1}\cos\kappa_{eff}z. \tag{4.160}$$

Specifying the initial conditions, $\mathcal{E}_1(z=0) = \mathcal{E}_1(0)$ and $\mathcal{E}_3(z=0) = 0$ – there is no SF at the entrance to the medium – we obtain the expressions for the signal and the SF waves as

$$\mathcal{E}_1 = \mathcal{E}_1(0)\cos\kappa_{eff}z, \tag{4.161}$$

and

$$\mathcal{E}_3 = -\mathcal{E}_1(0)\frac{\kappa_{eff}}{\kappa_1}\sin\kappa_{eff}z, \tag{4.162}$$

In physical terms, the SFG in the undepleted pump approximation describes periodic power exchange between the signal and the SF waves. The periodic character of the power exchange between the signal and the SF can be explained by observing that to create an SF photon, a signal photon has to be annihilated, $\omega_3 = \omega_1 + \omega_2$, such that the more the power residing with the SF, the less the power of the signal and vice versa.

Exercise 4.10. *Solve Eqs. (4.154) and (4.155) for $\Delta k \neq 0$ in the case when initially all power resides with ω_1 harmonic. Determine the SF intensity and show that its maximum reduces precipitously as Δk increases. Comment on the importance of phase matching for efficient SFG. Hint: look for solutions in the form*

$$\mathcal{E}_1 = \mathcal{A}_1 e^{-i\Delta k z/2}, \qquad \mathcal{E}_3 = \mathcal{A}_3 e^{i\Delta k z/2}, \tag{4.163}$$

and show that (4.154) and (4.155) reduce to homogeneous equations

$$\frac{d\mathcal{A}_1}{dz} = \frac{i\Delta k}{2}\mathcal{A}_1 + \kappa_1 \mathcal{A}_3, \tag{4.164}$$

and

$$\frac{d\mathcal{A}_3}{dz} = -\frac{i\Delta k}{2}\mathcal{A}_3 + \kappa_3 \mathcal{A}_1, \tag{4.165}$$

which can be solved by usual methods.

4.8.2 Manley-Rowe relations

Consider now the SFG in a lossless medium such that

$$\chi^{(2)}_{eff} = \chi^{(2)*}_{eff}. \tag{4.166}$$

The wave equations in the plane wave geometry, (4.151) – (4.153), can then be cast into the form

$$\frac{d\mathcal{E}_1}{dz} = \frac{i\omega_1^2}{2k_1c^2}\chi^{(2)}_{eff}\mathcal{E}_3\mathcal{E}_2^* e^{-i\Delta k z}, \tag{4.167}$$

$$\frac{d\mathcal{E}_2}{dz} = \frac{i\omega_2^2}{2k_2c^2}\chi^{(2)}_{eff}\mathcal{E}_3\mathcal{E}_1^* e^{-i\Delta k z}, \tag{4.168}$$

and

$$\frac{d\mathcal{E}_3}{dz} = \frac{i\omega_3^2}{2k_3c^2}\chi^{(2)}_{eff}\mathcal{E}_1\mathcal{E}_2 e^{i\Delta k z}. \tag{4.169}$$

Let us now study relations among the energy fluxes associated with the mixing waves. To this end, we derive the following equations for the wave intensities

$$\frac{d|\mathcal{E}_1|^2}{dz} = \frac{\omega_1}{n_1 c}\chi^{(2)}_{eff}\mathrm{Im}(\mathcal{E}_1^*\mathcal{E}_2^*\mathcal{E}_3 e^{i\Delta k z}), \tag{4.170}$$

$$\frac{d|\mathcal{E}_2|^2}{dz} = \frac{\omega_2}{n_2 c}\chi^{(2)}_{eff}\mathrm{Im}(\mathcal{E}_1^*\mathcal{E}_2^*\mathcal{E}_3 e^{i\Delta k z}), \tag{4.171}$$

$$\frac{d|\mathcal{E}_3|^2}{dz} = -\frac{\omega_3}{n_3 c^2}\chi^{(2)}_{eff}\mathrm{Im}(\mathcal{E}_1^*\mathcal{E}_2^*\mathcal{E}_3 e^{i\Delta k z}), \tag{4.172}$$

where we introduced $k_j = n_j\omega_j/c$.

Further, we introduce the optical intensities of the signal, pump and the SF waves as

$$I_j = \frac{\epsilon_0 n_j c}{2}|\mathcal{E}_j|^2, \tag{4.173}$$

with $j = 1, 2, 3$. It can then be inferred from Eqs. (4.170) – (4.173) that

$$\frac{dI_1}{dz} = \frac{\epsilon_0\omega_1}{2}\chi^{(2)}_{eff}\text{Im}(\mathcal{E}_1\mathcal{E}_2\mathcal{E}_3^* e^{i\Delta kz}), \tag{4.174}$$

and

$$\frac{dI_2}{dz} = \frac{\epsilon_0\omega_2}{2}\chi^{(2)}_{eff}\text{Im}(\mathcal{E}_1\mathcal{E}_2\mathcal{E}_3^* e^{i\Delta kz}), \tag{4.175}$$

as well as

$$\frac{dI_3}{dz} = -\frac{\epsilon_0\omega_3}{2}\chi^{(2)}_{eff}\text{Im}(\mathcal{E}_1\mathcal{E}_2\mathcal{E}_3^* e^{i\Delta kz}). \tag{4.176}$$

It follows at once by adding Eqs. (4.174), (4.175) and (4.176) that

$$\sum_{j=1}^{3} I_j = const, \tag{4.177}$$

which is tantamount to energy conservation for the SFG in lossless media. We can also infer from Eqs. (4.174) – (4.176) that

$$\frac{d}{dz}\left(\frac{I_1}{\omega_1} - \frac{I_2}{\omega_2}\right) = 0, \tag{4.178}$$

$$\frac{d}{dz}\left(\frac{I_1}{\omega_1} + \frac{I_3}{\omega_3}\right) = 0, \tag{4.179}$$

and

$$\frac{d}{dz}\left(\frac{I_2}{\omega_2} + \frac{I_3}{\omega_3}\right) = 0. \tag{4.180}$$

The preceding differential laws are equivalent to the three new invariants for the SFG process, which are known as the Manley-Rowe relations; the latter take the form

$$\frac{I_1}{\omega_1} - \frac{I_2}{\omega_2} = \mathcal{M}_1 = const, \tag{4.181}$$

$$\frac{I_1}{\omega_1} + \frac{I_3}{\omega_3} = \mathcal{M}_2 = const, \tag{4.182}$$

$$\frac{I_2}{\omega_2} + \frac{I_3}{\omega_3} = \mathcal{M}_3 = const. \tag{4.183}$$

The physical interpretation of Eqs. (4.181) – (4.183) can be best furnished using the photon picture. To this end, one can introduce the photon number fluxes – the number of photons at frequency ω_j created or annihilated per second – by the expression, $N_j = I_j/\hbar\omega_j$. It then follows from Eq. (4.181) – (4.183) that the numbers of signal and idler photons generated per unit time in any SFG process must be separately equal to the number of pump photons destroyed per unit time. Summarizing, we can say that to generate one SF photon, a signal and a pump photon must be destroyed. The qualitative photon picture of the SFG is exhibited in the form of a simple three-photon diagram in Fig. 2.

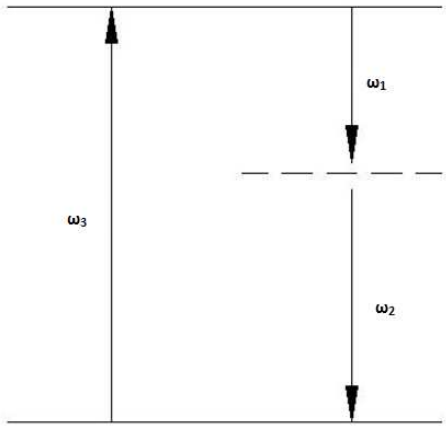

Figure 4.13: Illustrating Manley-Rowe relations with a photon diagram.

4.9 Difference-frequency generation

Let us now look into the difference-frequency generation (DFG), a second-order process of generating a difference frequency (DF) wave at frequency $\omega_3 = \omega_1 - \omega_2$ from the pump wave at frequency ω_1 and the idler wave at frequency ω_2 so named as its mere presence is required for realization of the process. The DF wave is often referred to as the signal. The DF generation is schematically illustrated in Fig. 1.

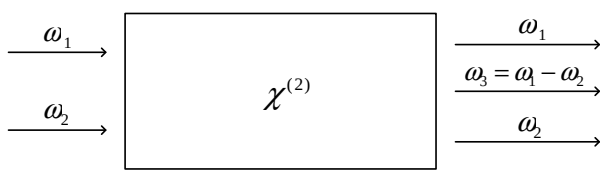

Figure 4.14: Schematic illustration of the difference-frequency generation process.

The paraxial wave equations governing DFG can be shown to take the form

$$\frac{\partial \mathcal{E}_1}{\partial z} - \frac{i}{2k_1}\nabla_\perp^2 \mathcal{E}_1 = \frac{i\omega_1^2}{2k_1 c^2}\chi_{eff}^{(2)}(-\omega_1;\omega_2,\omega_3)\mathcal{E}_2\mathcal{E}_3 e^{i\Delta kz}, \tag{4.184}$$

and

$$\frac{\partial \mathcal{E}_2}{\partial z} - \frac{i}{2k_2}\nabla_\perp^2 \mathcal{E}_2 = \frac{i\omega_2^2}{2k_2 c^2}\chi_{eff}^{(2)}(-\omega_2;\omega_1,-\omega_3)\mathcal{E}_1\mathcal{E}_3^* e^{-i\Delta kz}, \tag{4.185}$$

as well as

$$\frac{\partial \mathcal{E}_3}{\partial z} - \frac{i}{2k_3}\nabla_\perp^2 \mathcal{E}_3 = \frac{i\omega_3^2}{2k_3 c^2}\chi_{eff}^{(2)}(-\omega_3;\omega_1,-\omega_2)\,\mathcal{E}_1\mathcal{E}_2^*\, e^{-i\Delta kz}, \tag{4.186}$$

where the wave number mismatch is now defined as

$$\Delta k = k_1 - k_2 - k_3, \tag{4.187}$$

and the signal frequency is given by

$$\omega_3 = \omega_1 - \omega_2. \tag{4.188}$$

Using general symmetries of nonlinear susceptibilities it can be demonstrated that

$$\chi^{(2)}_{eff}(-\omega_3;\omega_1,-\omega_2) = \chi^{(2)*}_{eff}(-\omega_1;\omega_2,\omega_3) = \chi^{(2)}_{eff}(-\omega_2;\omega_1,-\omega_3). \tag{4.189}$$

It follows from Eqs. (4.184) – (4.186) and (4.189) that in the plane wave geometry, one can obtain the following set of DFG wave equations

$$\frac{d\mathcal{E}_1}{dz} = \frac{i\omega_1^2}{2k_1c^2}\chi^{(2)}_{eff}\mathcal{E}_2\mathcal{E}_3e^{i\Delta kz}, \tag{4.190}$$

$$\frac{d\mathcal{E}_2}{dz} = \frac{i\omega_2^2}{2k_2c^2}\chi^{(2)*}_{eff}\mathcal{E}_1\mathcal{E}_3^*e^{-i\Delta kz}, \tag{4.191}$$

and

$$\frac{d\mathcal{E}_3}{dz} = \frac{i\omega_3^2}{2k_3c^2}\chi^{(2)*}_{eff}\mathcal{E}_1\mathcal{E}_2^*e^{-i\Delta kz}. \tag{4.192}$$

We will restrict ourselves to studying DFG in the undepleted pump approximation, $\mathcal{E}_1 = const$, implying that

$$\frac{d\mathcal{E}_2}{dz} = \zeta_2\mathcal{E}_3^*e^{-i\Delta kz}, \tag{4.193}$$

and

$$\frac{d\mathcal{E}_3}{dz} = \zeta_3\mathcal{E}_2^*e^{-i\Delta kz}. \tag{4.194}$$

Here we introduced the quantities

$$\zeta_j = \frac{i\omega_j^2}{2k_jc^2}\chi^{(2)*}_{eff}\mathcal{E}_1, \qquad j = 2,3. \tag{4.195}$$

Assuming, for simplicity, there is perfect phase matching, $\Delta k = 0$, we can reduce Eqs. (4.193) and (4.194) to

$$\frac{d^2\mathcal{E}_3}{dz^2} - \zeta^2_{eff}\mathcal{E}_3 = 0, \tag{4.196}$$

where

$$\zeta^2_{eff} = \frac{\omega_2^2\omega_3^2|\chi^{(2)}_{eff}|^2|\mathcal{E}_1|^2}{4k_2k_3c^4}. \tag{4.197}$$

A general solution to (4.196) is

$$\mathcal{E}_3(z) = D_1\cosh\zeta_{eff}z + D_2\sinh\zeta_{eff}z. \tag{4.198}$$

Stipulating that initially all power reside with the idler, $\mathcal{E}_3(z=0)=\mathcal{E}_3(0)$ and $\mathcal{E}_2(z=0)=0$, yields the solution

$$\mathcal{E}_3(z)=\mathcal{E}_3(0)\cosh\zeta_{eff}z, \tag{4.199}$$

and

$$\mathcal{E}_2(z)=\frac{\zeta_{eff}\mathcal{E}_3^*(0)}{\zeta_3^*}\sinh\zeta_{eff}z. \tag{4.200}$$

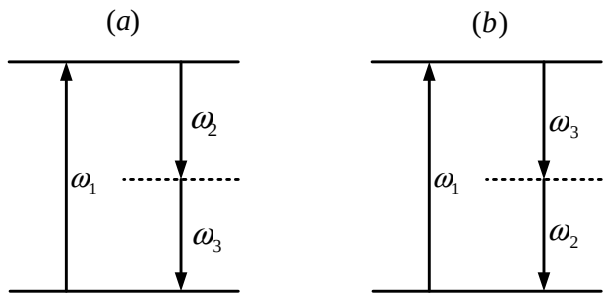

Figure 4.15: Schematic illustration of the difference-frequency generation process.

It can be easily inferred from Eq. (4.199) and (4.200) that both the signal and the idler monotonically grow with the distance z. Such a behavior – which is in sharp contrast with the SFG – is graphically presented in the diagram in Fig. 2. To explain the diagram, it is sufficient to notice that in the DFG process, the signal and idler photons are created and annihilated in pairs, $\omega_1=\omega_3+\omega_2$. In other words, the greater the power of one wave – be it the signal or the idler – the greater the power of the other. The two possibilities are illustrated in Figs. 2(a) and 2(b), respectively.

We can show that the monotonic character of the signal and idler wave growth depends on the pump power level in case of finite mismatch $\Delta k\neq 0$. To this end, we transform Eqs. (4.193) and (4.194) to

$$\frac{d\mathcal{E}_2^*}{dz}=\zeta_2^*\mathcal{E}_3e^{-i\Delta kz}. \tag{4.201}$$

$$\frac{d\mathcal{E}_3}{dz}=\zeta_3\mathcal{E}_2^*e^{-i\Delta kz}. \tag{4.202}$$

Introducing the new variables viz.,

$$\mathcal{E}_2^*=\overline{\mathcal{E}}_2e^{i\Delta kz/2},\qquad \mathcal{E}_3=\overline{\mathcal{E}}_3e^{-i\Delta kz/2}, \tag{4.203}$$

we arrive at the equations

$$\overline{\mathcal{E}}_2'+\tfrac{1}{2}i\Delta k\overline{\mathcal{E}}_2=\zeta_2^*\overline{\mathcal{E}}_3, \tag{4.204}$$

and

$$\overline{\mathcal{E}}_3'-\tfrac{1}{2}i\Delta k\overline{\mathcal{E}}_2=\zeta_3\overline{\mathcal{E}}_2. \tag{4.205}$$

Seeking solutions to Eqs. (4.204) and (4.205) in the form

$$\overline{\mathcal{E}}_2(z) = \overline{\mathcal{E}}_2^{(0)} e^{\Omega_{eff} z}; \qquad \overline{\mathcal{E}}_3(z) = \overline{\mathcal{E}}_3^{(0)} e^{\Omega_{eff} z}, \tag{4.206}$$

we obtain from the determinant condition, the expression for Ω_{eff}:

$$\Omega_{eff} = \sqrt{\zeta_{eff}^2 - (\Delta k/2)^2}. \tag{4.207}$$

The latter implies that in the presence of phase mismatch, there exists a power threshold for simultaneous amplification of the signal and idler modes,

$$I_{th} = \frac{2\epsilon_0 n_1 n_2 n_3 c^3}{\omega_2 \omega_3 |\chi_{eff}^{(2)}|^2} \left(\frac{\Delta k}{2}\right)^2. \tag{4.208}$$

Thus, for a given phase mismatch, the pump intensity must be greater than a certain critical value, $I_1 \geq I_{th}$, for parametric amplification to take place.

Next, general solutions for the idler and signal modes can be expressed above threshold as

$$\overline{\mathcal{E}}_2(z) = \mathcal{E}_2^*(0) \cosh \Omega_{eff} z + A \sinh \Omega_{eff} z, \tag{4.209}$$

and

$$\overline{\mathcal{E}}_3(z) = \mathcal{E}_3(0) \cosh \Omega_{eff} z + B \sinh \Omega_{eff} z. \tag{4.210}$$

Substituting from Eqs. (4.209) and (4.210) into Eqs. (4.204) and (4.205), we determine the coefficients A and B:

$$A = \frac{\zeta_2^* \mathcal{E}_3(0) - \frac{1}{2}\Delta k \mathcal{E}_2^*(0)}{\Omega_{eff}}, \tag{4.211}$$

and

$$B = \frac{\zeta_3 \mathcal{E}_2^*(0) + \frac{1}{2}\Delta k \mathcal{E}_3(0)}{\Omega_{eff}}, \tag{4.212}$$

Finally, on substituting from Eqs. (4.211) and (4.212) into (4.209) and (4.210) we obtain, upon a slight rearrangement, the signal and idler fields in the form

$$\mathcal{E}_3(z) = \left[\mathcal{E}_3(0)\left(\cosh \Omega_{eff} z + \frac{i\Delta k}{2\Omega_{eff}} \sinh \Omega_{eff} z\right) + \frac{\zeta_3 \mathcal{E}_2^*(0)}{\Omega_{eff}} \sinh \Omega_{eff} z\right] e^{-i\Delta k z/2}, \tag{4.213}$$

and

$$\mathcal{E}_2(z) = \left[\mathcal{E}_2(0)\left(\cosh \Omega_{eff} z + \frac{i\Delta k}{2\Omega_{eff}} \sinh \Omega_{eff} z\right) + \frac{\zeta_2 \mathcal{E}_3^*(0)}{\Omega_{eff}} \sinh \Omega_{eff} z\right] e^{-i\Delta k z/2}. \tag{4.214}$$

Exercise 4.11. *Show that below threshold, the solutions can be obtained with the substitutions,*

$$\Omega_{eff} \to i\Omega_{eff}; \qquad \cosh i\Omega_{eff} z \to \cos \Omega_{eff} z, \qquad \sinh i\Omega_{eff} z \to i \sin \Omega_{eff} z,$$

yielding

$$\mathcal{E}_3(z) = \left[\mathcal{E}_3(0)\left(\cos\Omega_{eff}z + \frac{i\Delta k}{2\Omega_{eff}}\sin\Omega_{eff}z\right) + \frac{\zeta_3\mathcal{E}_2^*(0)}{\Omega_{eff}}\sin\Omega_{eff}z\right]e^{-i\Delta kz/2},$$

and

$$\mathcal{E}_2(z) = \left[\mathcal{E}_2(0)\left(\cos\Omega_{eff}z + \frac{i\Delta k}{2\Omega_{eff}}\sin\Omega_{eff}z\right) + \frac{\zeta_2\mathcal{E}_3^*(0)}{\Omega_{eff}}\sin\Omega_{eff}z\right]e^{-i\Delta kz/2}.$$

How can you reconcile the periodic power exchange between the signal and idler modes with the photon diagram of Fig. 2 demanding that signal and idler photons be created or annihilated in pairs?

The DFG process is also known as parametric down-conversion: A high-frequency pump photon generates a signal-idler photon pair at lower frequencies. It is the key process to generate a pair of entangled photons from a single pump photon in $\chi^{(2)}$ nonlinear media; the latter finds numerous applications in quantum optics.

Exercise 4.12. *Show that the Manley-Rowe relations for the DFG without the undepleted pump approximation take the form*

$$\frac{I_1}{\omega_1} + \frac{I_2}{\omega_2} = \mathcal{M}_1 = const,$$
$$\frac{I_1}{\omega_1} + \frac{I_3}{\omega_3} = \mathcal{M}_2 = const,$$
$$\frac{I_2}{\omega_2} - \frac{I_3}{\omega_3} = \mathcal{M}_3 = const.$$

and interpret your results using the photon diagram of Fig. 2.

4.10 Four-wave mixing: General considerations

In general, third-order nonlinear processes are much weaker than their second-order counterparts. For example, in solids the ratio of the third- to the second-order susceptibility is of the order of 10^{-9} m/V, implying that the fields as large as 10^3 MV/m are required to make the influence of the third-order nonlinearities felt in presence of the second-order ones. On the other hand, if the medium atoms do have the center of inversion, the third-order nonlinearity makes the dominant contribution to the nonlinear polarization. As most isotropic nonlinear media fall into this category, the third-order nonlinear interactions, involving mixing of four waves with, in general, different frequencies, are of the utmost importance both for our fundamental understanding of nonlinear optical processes and in potential applications. Quite generally, the third-order nonlinear processes are commonly referred to as *four-wave mixing*. We will study four-wave mixing in isotropic lossless media by deriving coupled wave equations describing the interaction of four quasi-monochromatic paraxial waves with different carrier frequencies. We will assume the beams to be linearly or circularly polarized. The electric field of a beam with the carrier frequency ω_s can be represented as

$$\tilde{\mathbf{E}}(\mathbf{r},\omega_s) = \mathbf{e}(\omega_s)\mathcal{E}(\mathbf{r}_\perp, z, \omega_s)e^{ik_s z}. \tag{4.215}$$

Since a linearly –or circularly – polarized field maintains its state of polarization in an isotropic medium, the induced polarization field is then given by the expression

$$\tilde{\mathbf{P}}_{NL}(\mathbf{r},\omega_s) = \mathbf{e}(\omega_s)\mathcal{P}_{NL}(\mathbf{r}_\perp, z, \omega_s)e^{ik_s z}, \tag{4.216}$$

where

$$k_s^2 = \epsilon(\omega_s)\frac{\omega_s^2}{c^2}. \tag{4.217}$$

The paraxial wave equation governing the field evolution is

$$2ik_s\frac{\partial \mathcal{E}_s}{\partial z} + \nabla_\perp^2 \mathcal{E}_s = -\mu_0\omega_s^2\mathcal{P}_{NL}. \tag{4.218}$$

Here we introduced the notation

$$\mathcal{E}_s \equiv \mathcal{E}(\mathbf{r},\omega_s). \tag{4.219}$$

Recall that the third-order polarization field can be expressed as

$$\begin{aligned}\tilde{P}_i^{(3)}(\mathbf{r},\omega_s) &= \epsilon_0 c^{(3)}(\omega_1,\omega_2,\omega_3)\sum_{jkl}\tilde{\chi}_{ijkl}^{(3)}(-\omega_s;\omega_1,\omega_2,\omega_3) \\ &\quad \times\tilde{E}_j(\mathbf{r},\omega_1)\tilde{E}_k(\mathbf{r},\omega_2)\tilde{E}_l(\mathbf{r},\omega_3),\end{aligned} \tag{4.220}$$

with $\omega_s = \omega_1 + \omega_2 + \omega_3$. Using Eqs. (4.220) and (4.216), we obtain for the slowly-varying third-order polarization field the expression

$$\begin{aligned}\mathcal{P}_i^{(3)}(\mathbf{r},\omega_s) &= \epsilon_0 c^{(3)}(\omega_1,\omega_2,\omega_3)\sum_{jkl}\tilde{\chi}_{ijkl}^{(3)}(-\omega_s;\omega_1,\omega_2,\omega_3)e_i(\omega_s) \\ &\quad \times e_j(\omega_1)e_k(\omega_2)e_l(\omega_3)\mathcal{E}(\mathbf{r},\omega_1)\mathcal{E}(\mathbf{r},\omega_2)\mathcal{E}(\mathbf{r},\omega_3)e^{i\Delta kz},\end{aligned} \tag{4.221}$$

where the phase mismatch is defined as

$$\Delta k \equiv k(\omega_1) + k(\omega_2) + k(\omega_3) - k(\omega_s). \tag{4.222}$$

Introducing the notation

$$\chi_{eff}^{(3)}(-\omega_s;\omega_1,\omega_2,\omega_3) \equiv c^{(3)}(\omega_1,\omega_2,\omega_3)\sum_{ijkl}\tilde{\chi}_{ijkl}^{(3)}(-\omega_s;\omega_1,\omega_2,\omega_3)\mathbf{e}_i(\omega_s)\mathbf{e}_j(\omega_1)\mathbf{e}_k(\omega_2)\mathbf{e}_l(\omega_3), \tag{4.223}$$

we finally arrive at the paraxial wave equation governing the four-wave mixing processes

$$\frac{\partial \mathcal{E}_s}{\partial z} - \frac{i}{2k(\omega_s)}\nabla_\perp^2\mathcal{E}_s = \frac{i\omega_s^2}{2k(\omega_s)c^2}\chi_{eff}^{(3)}(-\omega_s;\omega_1,\omega_2,\omega_3)\mathcal{E}_1\mathcal{E}_2\mathcal{E}_3 e^{i\Delta kz}. \tag{4.224}$$

The family of third-order processes is very large; each particular process is specified by a choice of four mixing frequencies. In the following, we will only consider two commonly encountered processes: third-harmonic generation and self-focusing.

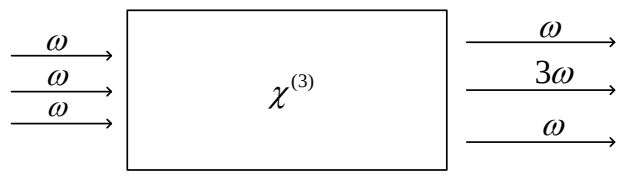

Figure 4.16: Illustrating the third harmonic generation.

4.11 Third harmonic generation

Third harmonic generation (THG) is a process of producing a wave that oscillates at the frequency 3ω by mixing three waves, each having the same carrier frequency ω as is indicated in the diagram below.

The set of THG governing equations can be easily obtained from (4.224) to be

$$\frac{\partial \mathcal{E}_\omega}{\partial z} - \frac{i}{2k_\omega}\nabla_\perp^2 \mathcal{E}_\omega = \frac{i\omega^2}{2k_\omega c^2}\chi_{eff}^{(3)}(-\omega; 3\omega, -\omega, -\omega)\mathcal{E}_{3\omega}\mathcal{E}_\omega^{*2}e^{-i\Delta kz}. \tag{4.225}$$

and

$$\frac{\partial \mathcal{E}_{3\omega}}{\partial z} - \frac{i}{2k_{3\omega}}\nabla_\perp^2 \mathcal{E}_{3\omega} = \frac{9i\omega^2}{2k_{3\omega}c^2}\chi_{eff}^{(3)}(-3\omega; \omega, \omega, \omega)\,\mathcal{E}_\omega^3\, e^{i\Delta kz}, \tag{4.226}$$

where the phase mismatch is given by

$$\Delta k = 3k(\omega) - k(3\omega). \tag{4.227}$$

The analysis reveals that the degeneracy factors associated with the corresponding mixing processes, ($3\omega = \omega + \omega + \omega$) and ($\omega = 3\omega - \omega - \omega$) are related as

$$c^{(3)}(3\omega, -\omega, -\omega) = 3c^{(3)}(\omega, \omega, \omega), \tag{4.228}$$

implying the relation between the effective susceptibilities as

$$\chi_{eff}^{(3)}(-\omega; 3\omega, -\omega, -\omega) = 3\chi_{eff}^{(3)}(-3\omega; \omega, \omega, \omega) \equiv 3\chi_{eff}^{(3)} \tag{4.229}$$

Using Eq. (4.229) and assuming a plane wave geometry, we can transform the governing coupled wave equations, Eqs. (4.225) and (4.226), into the form

$$\frac{d\mathcal{E}_\omega}{dz} = \frac{3i\omega}{2n_\omega c}\chi_{eff}^{(3)}\,\mathcal{E}_{3\omega}\mathcal{E}_\omega^{*2}e^{-i\Delta kz}. \tag{4.230}$$

and

$$\frac{d\mathcal{E}_{3\omega}}{dz} = \frac{3i\omega}{2n_{3\omega}c}\chi_{eff}^{(3)}\,\mathcal{E}_\omega^3\, e^{i\Delta kz}. \tag{4.231}$$

The last equations are very similar to those describing second harmonic generation. Unfortunately, though, third harmonic generation is a rather weak process. Therefore relatively high optical intensities are required to generate THG in a crystal with a

reasonable efficiency. To estimate the THG efficiency, we consider the THG process in the undepleted pump approximation, $\mathcal{E}_\omega = const$. Under these conditions, equations (4.230) and (4.231) can be easily integrated to give an expression for the third harmonic field in the form

$$\mathcal{E}_{3\omega}(L) = \frac{i3\omega}{2n_{3\omega}c}\chi^{(3)}_{eff}\mathcal{E}^3_\omega e^{i\Delta kL/2}\text{sinc}(\Delta kL/2). \tag{4.232}$$

In complete analogy with the SHG theory, we introduce the THG efficiency by the expression

$$\eta_{THG} = \frac{I_{3\omega}(L)}{I_\omega(0)}. \tag{4.233}$$

It follows from Eqs. (4.232) and (4.233), assuming perfect phase matching that in the undepleted pump approximation,

$$\eta_{THG} = 36\pi^2\left(\frac{L}{\lambda}\right)^2\frac{\chi^{(3)2}_{eff}I^2_\omega}{n_{3\omega}n^3_\omega\epsilon^2_0c^2}. \tag{4.234}$$

Even if we assume the fundamental field intensity is as large as a typical breakdown intensity in solids, $I \sim 100$ MW/cm^2 and take realistic values of the other parameters: $L \sim 1$ cm, $n_\omega \sim n_{3\omega} \sim 1.5$, $\lambda \sim 5\times10^{-5}$ cm, and $\chi^{(3)}_{eff} \sim 10^{-21}$ m^2/W^2, we arrive at an estimate

$$\eta_{THG} \sim 5\times10^{-7} \ll 1, \tag{4.235}$$

which is tiny for all practical purposes. Moreover, it is hard to achieve phase matching in crystals; all of which effectively precludes the laboratory THG realization in most solid media.

However, THG can be generated in gases, such as sodium or rubidium vapors, in the vicinity of an optical resonance where the magnitude of $\chi^{(3)}$ is significantly enhanced. Unfortunately, such an enhancement is, in general, accompanied by the increase in linear as well as nonlinear absorption that must also be reckoned with whenever third harmonic generation in gases is attempted. As linear absorption dominates at resonance, the best way to boost the THG efficiency is to tune the laser to a two-photon resonance as is indicated in Fig. 2(a).

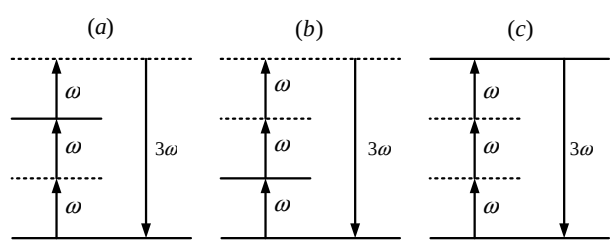

Figure 4.17: Illustrating the third-harmonic generation in gases under resonant excitation conditions. The laser is tuned to either two- or one- or else three-photon resonance in parts (a), (b) and (c), respectively.

4.12 Self-focusing and spatial solitons

Whenever a light beam propagates inside a nonlinear medium whose refractive index depends on the beam intensity, the light rays near the beam center, where the intensity is the highest, experience stronger refraction – assuming the nonlinear refractive index of the medium increases with the intensity – causing the rays to bend toward the center. As a result, the intensity increases toward the beam center on propagation in the medium. The light evolution looks as if the rays were focused by a positive lens toward the beam center. Such a behavior is termed *self-focusing* of light in a nonlinear medium, and the medium with a positive nonlinear refractive index forming a focusing lens, self-focusing. As a consequence of self-focusing, the beam narrows and its peak intensity is enhanced with the propagation distance. On the other hand, every beam tends to spread due to diffraction which tends to decrease light intensity at the center. The two opposing trends are characterized by different spatial scales. We can easily estimate such scales – referred to as nonlinear and diffraction lengths, respectively – from elementary considerations. The characteristic diffraction length was defined before in the studies of Gaussian beam diffraction in free space:

$$L_D \simeq k w_0^2, \tag{4.236}$$

where $k = n_0\omega/c$, ω being the carrier frequency of the beam.

On the other hand, the intensity-dependent nonlinear refractive index modulates the optical phase of the beam electric field. This phenomenon is known as the *self-phase modulation*. Further, due to coupling of the intensity and phase dynamics of the field in nonlinear media, the change in the phase of the optical field induces modifications of the beam intensity profile. The influence of nonlinearity becomes important over distances such that the phase accretion is of the order of, at least, one radian, i. e.,

$$k\Delta n_{NL} L_{NL} \sim 1, \tag{4.237}$$

where the nonlinear change in the refractive index Δn_{NL} can be estimated using the peak intensity of the beam as

$$\Delta n_{NL} \sim \overline{n}_2 I_0. \tag{4.238}$$

Here I_0 is the peak intensity and $\overline{n}_2 > 0$ is a nonlinear refractive index coefficient to be discussed in greater detail below. It follows from Eqs. (4.237) and (4.238) that

$$L_{NL} \sim \frac{1}{k\overline{n}_2 I_0}. \tag{4.239}$$

The beam evolution scenario entirely depends on the relative sizes of the two characteristic lengths. In particular, if $L_D < L_{NL}$, diffraction dominates, and the beam spreads. However, if the diffraction and nonlinearity operate at the characteristic scales of the same order, exact balance of the two opposing trends is possible, leading to the formation of *spatial solitons*, i. e., the beams whose spatial profiles and widths do not change upon propagation in self-focusing nonlinear media. A soliton can be formed if the optical power of the beam is exactly equal to a certain critical power such that

the nonlinearity can arrest diffraction-induced spreading. We can estimate the critical power necessary for soliton formation by imposing the balance condition

$$L_D \simeq L_{NL}. \tag{4.240}$$

It follows at once from Eqs. (4.236), (4.239) and (4.240) that the critical power, $P_{cr} = I_{cr}\pi w_0^2$ is given by

$$P_{cr} \simeq \frac{\lambda_0^2}{4\pi n_0 \overline{n}_2}, \tag{4.241}$$

where $\lambda_0 = 2\pi/k_0 = 2\pi c/\omega$.

Exercise 4.13. *The magnitude of the nonlinear refractive index for carbon disulfide (CS_2) is $\overline{n}_2 \simeq 3 \times 10^{-14}$ cm^2/W, the linear refractive index is equal to 1.63. Estimate the critical power for spatial soliton formation at $\lambda_0 \simeq 1$ μm. Compare your results with P_{cr} for silica glass for which $\overline{n}_2 \simeq 5 \times 10^{-16}$ cm^2/W, and $n_0 \simeq 1.4$.*

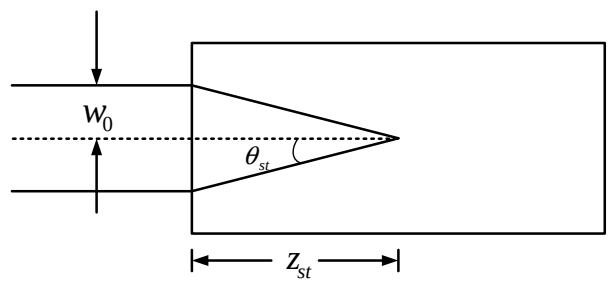

Figure 4.18: Illustrating the focal length in the self-focusing regime.

If the characteristic nonlinear length is smaller than the diffraction length, the nonlinearity prevails, causing self-focusing of the beam. One can estimate a characteristic self-focusing distance in the limit $L_{NL} \ll L_D$. In this case, diffraction is negligible, and geometrical optics approach would suffice for a rough estimate. According to Fermat's principle, any ray traveling from the wavefront up to the focusing point must traverse the same optical path, $\int dsn(s) = const$. As a result, we obtain for the paths exhibited in Fig. 3,

$$(n_0 + \delta n)z_f = \left(n_0 + \frac{\delta n}{2}\right)\sqrt{z_f^2 + w_0^2} \simeq n_0 z_f \left(1 + \frac{\delta n}{2n_0}\right)\left(1 + \frac{w_0^2}{2z_f^2}\right), \tag{4.242}$$

where we have assumed that the refractive index along the central ray is $n_0 + \delta n$, whereas the peripheral ray experiences the refractive index strength of roughly $n_0 + \delta n/2$. It then follows from (4.242) after simple algebra that the self-focusing distance is

$$z_f \simeq w_0\sqrt{\frac{n_0}{\delta n}} \simeq w_0\sqrt{\frac{n_0}{\overline{n}_2 I_0}}. \tag{4.243}$$

Finally, using the expressions for the beam power and the critical power as

$$P = I\pi w_0^2, \qquad P_{cr} = I_0 \pi w_0^2, \tag{4.244}$$

we obtain the estimate

$$z_f \simeq \frac{L_D}{2}\sqrt{\frac{P_{cr}n_0}{P}}, \qquad P \gg P_{cr}. \tag{4.245}$$

Note that our approximate result (4.245) is consistent with our premise that diffraction is negligible, $z_f \ll L_D$ in the given power range. In reality, high-power optical beams, $P \gg P_{cr}$, will disintegrate into multiple filaments, each carrying approximately the power of P_{cr}, long before the self-focusing distance is reached. The filamentation is caused by a transverse instability resulting from the growth of tiny imperfections of the beam wave front.

To describe self-focusing and soliton formation, we can derive the nonlinear wave equation corresponding to the self-action process by a fundamental wave of frequency ω. Mathematically, the corresponding susceptibility tensor is $\boldsymbol{\chi}^{(3)}(-\omega;\omega,-\omega,\omega)$. The resulting equation takes the form

$$\frac{\partial \mathcal{E}_\omega}{\partial z} - \frac{i}{2k_\omega}\nabla_\perp^2 \mathcal{E}_\omega = \frac{i\omega^2}{2k_\omega c^2}\chi^{(3)}_{eff}(-\omega;\omega,-\omega,\omega)|\mathcal{E}_\omega|^2\mathcal{E}_\omega, \tag{4.246}$$

where we have introduced the quantities

$$\chi^{(3)}_{eff} = \frac{3}{4}\chi^{(3)}, \tag{4.247}$$

and

$$\chi^{(3)} \equiv \sum_{ijkl}\tilde{\chi}^{(3)}_{ijkl}(-\omega;\omega,-\omega,\omega)\mathbf{e}_i(\omega)\mathbf{e}_j(\omega)\mathbf{e}_k(\omega)\mathbf{e}_l(\omega). \tag{4.248}$$

Notice that there is no phase mismatch involved in the process of self-focusing, $\Delta k = k(\omega) + k(-\omega) + k(\omega) - k(\omega) = 0$, because $k(-\omega) = -k(\omega)$ in lossless media. Physically, this is the consequence of the fact that there is only one fundamental wave involved in the process which implies automatic conservation of the energy and momenta at the photon level.

We assume, for simplicity, the linear polarization of the beam – such that any polarization effects can be ignored – and introduce the scalar nonlinear polarization $\mathcal{P}_{tot} \equiv \mathcal{P}_i e_i$, by the expression

$$\mathcal{P}_{tot} = \epsilon_0\left(\chi^{(1)}\mathcal{E} + \frac{3}{4}\chi^{(3)}|\mathcal{E}|^2\mathcal{E}\right) = \epsilon_0\chi_{tot}\mathcal{E}, \tag{4.249}$$

where the total susceptibility is given by

$$\chi_{tot} = \chi^{(1)} + \frac{3}{4}\chi^{(3)}|\mathcal{E}|^2. \tag{4.250}$$

We can then defines the total refractive index as

$$n^2 = 1 + \chi_{tot}, \tag{4.251}$$

and the Kerr nonlinear refractive index by the expression

$$n = n_0 + n_2|\mathcal{E}|^2. \tag{4.252}$$

It follows from Eq. (4.250) and (4.251) and the fact that the nonlinear refraction is always a small effect as compared with the linear one, we obtain

$$(n_0 + n_2|\mathcal{E}|^2)^2 \simeq n_0^2 + 2n_0 n_2|\mathcal{E}|^2. \tag{4.253}$$

On comparing Eqs. (4.250) and (4.253), we infer that

$$n_2 = \frac{3\chi^{(3)}}{8n_0}, \tag{4.254}$$

which provides a relation between the third-order susceptibility and the nonlinear refractive index. In terms of the latter, the nonlinear wave equation for self-focusing can be rewritten as

$$i\frac{\partial \mathcal{E}}{\partial z} + \frac{1}{2k}\nabla_\perp^2 \mathcal{E} + \frac{kn_2}{n_0}|\mathcal{E}|^2\mathcal{E} = 0. \tag{4.255}$$

Equation (4.255) is referred to as the *nonlinear Schrödinger equation* (NLSE) because of its formal similarity with the Schrödinger equation in quantum mechanics.

Before we proceed further, we note that often the nonlinear refractive index associated with the optical intensity is introduced viz.,

$$n = n_0 + \overline{n}_2 I, \tag{4.256}$$

where I is the optical intensity – the energy density flux – defined as

$$I = \frac{\epsilon_0 c n_0}{2}|\mathcal{E}|^2. \tag{4.257}$$

The nonlinear nonlinear refractive index n_2 has the units of m^2/V^2 whereas the other one, $\overline{n}_2$, is measured in m^2/W^2. The two indices are related as

$$n_2 = \frac{\epsilon_0 c n_0}{2}\overline{n}_2. \tag{4.258}$$

We can now introduce dimensionless variables, $Z = z/L_D$, $U = \mathcal{E}/\mathcal{E}_0$, $\mathbf{R}_\perp = \mathbf{r}_\perp/w_0$, $\mathcal{E}_0 = (2I_0/\epsilon_0 c n_0)^{1/2}$, and transform the NLSE to the dimensionless form

$$i\frac{\partial U}{\partial Z} + \frac{1}{2}\nabla_\perp^2 U + \mathcal{N}^2|U|^2 U = 0. \tag{4.259}$$

Here we have introduced the only dimensionless parameter – the soliton parameter $\mathcal{N}$, governing the dynamics of the system. It is defined as follows

$$\mathcal{N}^2 \equiv \frac{L_D}{L_{NL}}, \tag{4.260}$$

where the diffraction and nonlinear lengths, L_D and L_{NL} are given by the expressions

$$L_D = kw_0^2, \qquad L_{NL} = \frac{1}{k\overline{n}_2 I_0}. \tag{4.261}$$

A numerical analysis of Eq. (4.259) confirms formation of a spatial soliton for the beam power such that $\mathcal{N} = 1$. However, the soliton turns out to be unstable with

respect to small perturbations. Stable solitons can be formed in two-spatial dimensions, provided the saturation of nonlinear refractive index is allowed. Stable spatial solitons can be generated in Kerr-like nonlinear media in a planar waveguide geometry where trapping in one spatial dimension is realized by the nonlinear medium whereas the other spatial dimension is trapped by the waveguide. The dimensionless NLSE in the planar waveguide geometry takes the form

$$i\frac{\partial U}{\partial Z} + \frac{1}{2}\frac{\partial^2 U}{\partial X^2} + \mathcal{N}^2|U|^2 U = 0. \tag{4.262}$$

The lowest order soliton corresponds to the exact balance between the nonlinearity and diffraction, $\mathcal{N} = 1$, and its spatial profile is given by

$$U(Z, X) = \text{sech}\, X\, e^{-iZ/2}. \tag{4.263}$$

Higher-order solitons also exist. They correspond to more intense beams, $\mathcal{N} > 1$. In such cases, the nonlinearity dominates at first, causing self-focusing of the beam. However, in $(1+1)$D geometry – indicating one transverse dimension plus one dimension along the waveguide unaffected by the waveguide trapping – the initial self-focusing can be slowed down and eventually reversed by increased diffraction of a more tightly focused beam. As a result, the periodic pattern of contraction and expansion of the soliton manifests itself, with the soliton returning to its initial shape and transverse size every half-period. Such solitons are called optical breathers. An example of a breather is displayed in Fig. 4(b) for $\mathcal{N} = 3$.

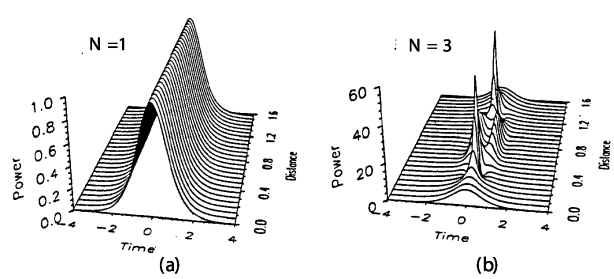

Figure 4.19: Intensity of the fundamental (a) and the third-order (b) soliton as function of the propagation distance.

Exercise 4.14. *Show that the 1D NLSE is invariant with respect to the Galilean transformation,*

$$X' = X - vZ; \qquad Z' = Z,$$

for an arbitrary speed v. In other words, demonstrate that Eq. (4.262) has the same form in the "primed" variables, provided the fields in the two coordinate systems are related by a gauge transformation,

$$U(Z, X) = V(Z', X')e^{if(Z', X')}.$$

Determine the phase f. Draw conclusions about the functional form of a moving soliton field.

4.13 Polarization dynamics of third-order processes

So far we have ignored tensor properties of nonlinear optical susceptibilities by considering linearly or circularly polarized light whose polarization properties do not change on propagation in isotropic media. Whenever elliptically polarized light is launched into such media, its state of polarization does in general change despite the isotropy of the medium. Thus, we shall be interested in polarization dynamics of light propagating in isotropic nonlinear media. If the isotropic medium possesses reflectional symmetry, the lowest order of the optical susceptibility tensor is the third. Remarkably, the mere isotropy and reflectional symmetry of the medium are sufficient to determine a general form of the third-order susceptibility tensor which we will do following a seminal work of Maker and Tehrune. In the next subsection, we examine tensor properties of the third-order susceptibility in the media with isotropic linear and nonlinear responses, while we will then explore the influence of linear anisotropy of the nonlinear medium – whose nonlinear properties can still be assumed isotropic – on light polarization dynamics in such media.

4.13.1 Isotropic nonlinear media with inversion symmetry

We begin by observing that since there is no privileged direction in such a medium, the third-order susceptibility tensor cannot have an index – corresponding to a given Cartesian coordinate – repeat an odd number of times: In other words, $\chi^{(3)}_{ijjj} = 0$ for any $j = x, y, z$. To demonstrate this property, consider a polarization component, P_x, say. If $\chi^{(3)}_{xyyy} \neq 0$, it follows that $P_x = \chi^{(3)}_{xyyy} E_y E_y E_y \neq 0$. On the other hand, polarization along the $x-$axis in an isotropic medium should not be affected by reflections with respect to the $xz-$plane. The latter affect the y-component of the field, though, $E_y \to -E_y$. Consequently, $P_x = \chi^{(3)}_{xyyy} E_y E_y E_y = (-1)^3 \chi^{(3)}_{xyyy} E_y E_y E_y$, implying that $\chi^{(3)}_{xyyy} = 0$. By the same token, all the other tensor components containing three repeated indices can be shown to be zero.

Further, we conclude by inspection that there are four kinds of nonzero tensor elements which are mutually related by the symmetry relations as

$$\chi^{(3)}_{xxxx} = \chi^{(3)}_{yyyy} = \chi^{(3)}_{zzzz}, \tag{4.264}$$

$$\chi^{(3)}_{xxyy} = \chi^{(3)}_{xxzz} = \chi^{(3)}_{yyxx} = \chi^{(3)}_{yyzz} = \chi^{(3)}_{zzyy} = \chi^{(3)}_{zzxx}, \tag{4.265}$$

$$\chi^{(3)}_{xyxy} = \chi^{(3)}_{xzxz} = \chi^{(3)}_{yzyz} = \chi^{(3)}_{zxzx} = \chi^{(3)}_{zyzy} = \chi^{(3)}_{yxyx}, \tag{4.266}$$

$$\chi^{(3)}_{xyyx} = \chi^{(3)}_{yxxy} = \chi^{(3)}_{xzzx} = \chi^{(3)}_{zxxz} = \chi^{(3)}_{yzzy} = \chi^{(3)}_{zyyz}. \tag{4.267}$$

Moreover, as $\chi^{(3)}_{ijkl}$ must be invariant with respect to rotations, the diagonal and off-diagonal elements of the susceptibility tensor can be shown to satisfy the relations

$$\chi^{(3)}_{xxxx} = \chi^{(3)}_{xxyy} + \chi^{(3)}_{xyyx} + \chi^{(3)}_{xyxy}, \tag{4.268}$$

with similar ones for $\chi^{(3)}_{yyyy}$ and $\chi^{(3)}_{zzzz}$. We can then infer from Eqs. (4.264) – (4.267) as well as Eq. (4.268) that the third-order susceptibility tensor in isotropic media with

inversion symmetry takes a general form

$$\chi^{(3)}_{ijkl} = \chi^{(3)}_{xxyy}\delta_{ij}\delta_{kl} + \chi^{(3)}_{xyxy}\delta_{ik}\delta_{jl} + \chi^{(3)}_{xyyx}\delta_{il}\delta_{jk}. \tag{4.269}$$

The expression (4.269) can be simplified even further for particular nonlinear processes if one recalls intrinsic symmetries of $\chi^{(3)}$ with respect to frequency permutations. We will focus here on the self-focusing (SF) process, $\omega_1 = \omega_2 = -\omega_3 = -\omega_4 = \omega$. The intrinsic permutation symmetry then implies

$$\chi^{(3)}_{ijkl}(-\omega, \omega, -\omega, \omega) = \chi^{(3)}_{ilkj}(-\omega, \omega, -\omega, \omega). \tag{4.270}$$

It follows at once that

$$\chi^{(3)}_{xxyy}(-\omega, \omega, -\omega, \omega) = \chi^{(3)}_{xyyx}(-\omega, \omega, -\omega, \omega). \tag{4.271}$$

We can then arrive at the final form for the third-order susceptibility for SF in isotropic media,

$$\begin{aligned} \chi^{(3)}_{ijkl}(-\omega, \omega, -\omega, \omega) \quad = \quad & \chi^{(3)}_{xxyy}(-\omega, \omega, -\omega, \omega)(\delta_{ij}\delta_{kl} + \delta_{il}\delta_{jk}) \\ & + \chi^{(3)}_{xyxy}(-\omega, \omega, -\omega, \omega)\delta_{ik}\delta_{jl}. \end{aligned} \tag{4.272}$$

Exercise 4.15. *Use intrinsic permutation symmetries of $\chi^{(3)}_{ijkl}$ to determine a general form of the susceptibility tensor $\chi^{(3)}_{ijkl}(-3\omega, \omega, \omega, \omega)$ for the third harmonic generation in isotropic media.*

The third-order polarization field for self-focusing takes the form

$$\mathcal{P}_i(\omega) = \frac{3\epsilon_0}{4}\sum_{jkl}\chi^{(3)}_{ijkl}(-\omega, \omega, -\omega, \omega)\mathcal{E}_j(\omega)\mathcal{E}_k(-\omega)\mathcal{E}_l(\omega), \tag{4.273}$$

where

$$\mathcal{E}_k(-\omega) = \mathcal{E}^*_k(\omega). \tag{4.274}$$

Substituting from Eqs. (4.272) into (4.273) we obtain, after some algebra, the expression

$$\mathcal{P}_i = \frac{3\epsilon_0}{2}\,\chi^{(3)}_{xxyy}\mathcal{E}_i\sum_k \mathcal{E}_k\mathcal{E}^*_k + \frac{3\epsilon_0}{4}\,\chi^{(3)}_{xyxy}\mathcal{E}^*_i\sum_l \mathcal{E}_l\mathcal{E}_l. \tag{4.275}$$

The latter can be written in the vector form as

$$\boldsymbol{\mathcal{P}}_{NL} = A(\boldsymbol{\mathcal{E}}\cdot\boldsymbol{\mathcal{E}}^*)\boldsymbol{\mathcal{E}} + \tfrac{1}{2}B(\boldsymbol{\mathcal{E}}\cdot\boldsymbol{\mathcal{E}})\boldsymbol{\mathcal{E}}^*, \tag{4.276}$$

where we have introduced the notations of *Maker and Terhune (1965)*

$$A \equiv \frac{3\epsilon_0}{2}\chi^{(3)}_{xxyy}, \qquad\qquad B \equiv \frac{3\epsilon_0}{2}\chi^{(3)}_{xyxy}. \tag{4.277}$$

Equation (4.276) gives the most general form of the third-order polarization response to an applied field of any polarization for a self-focusing process in an isotropic non-chiral nonlinear medium.

To better understand the role of the two terms entering the expression (4.276), we consider an elliptically polarized wave, propagating in the positive z-direction, which can be conveniently represented as a linear superposition of the right- and left-handed circular polarizations as

$$\boldsymbol{\mathcal{E}} = \mathcal{E}_+ \mathbf{e}_+ + \mathcal{E}_- \mathbf{e}_-, \tag{4.278}$$

where the unit vectors associated with the circular polarizations are defined as

$$\mathbf{e}_\pm = \frac{\mathbf{e}_x \pm i\mathbf{e}_y}{\sqrt{2}}; \qquad \mathbf{e}_- = \mathbf{e}_+^*. \tag{4.279}$$

It follows from (4.279) that

$$\mathbf{e}_\pm \cdot \mathbf{e}_\pm = 0, \qquad \mathbf{e}_\pm \cdot \mathbf{e}_\mp = 1. \tag{4.280}$$

The dot product of the two electric field vectors can then be expressed as

$$\boldsymbol{\mathcal{E}} \cdot \boldsymbol{\mathcal{E}} = 2\mathcal{E}_+\mathcal{E}_-, \qquad \boldsymbol{\mathcal{E}} \cdot \boldsymbol{\mathcal{E}}^* = |\mathcal{E}_+|^2 + |\mathcal{E}_-|^2. \tag{4.281}$$

Let us define the nonlinear polarization field in the circular polarization basis by the expression

$$\boldsymbol{\mathcal{P}}_{NL} = \mathcal{P}_{NL}^{(+)} \mathbf{e}_+ + \mathcal{P}_{NL}^{(-)} \mathbf{e}_-, \tag{4.282}$$

It can be inferred from (4.276), (4.281) and (4.282) that

$$\mathcal{P}_{NL}^{(\pm)} = [A|\mathcal{E}_\pm|^2 + (A+B)|\mathcal{E}_\mp|^2]\mathcal{E}_\pm. \tag{4.283}$$

Further, we can represent the total polarization field as a linear superposition of the circular polarization components as

$$\boldsymbol{\mathcal{P}} = \mathcal{P}_+ \mathbf{e}_+ + \mathcal{P}_- \mathbf{e}_-, \tag{4.284}$$

where $\mathcal{P}_+$ and $\mathcal{P}_-$ are effectively decoupled – there is an indirect coupling, though, via the nonlinear susceptibility – such that each polarization component is proportional to the corresponding electric field viz.,

$$\mathcal{P}_\pm = \left[\chi_L + \chi_{NL}^{(\pm)}\right] \mathcal{E}_\pm. \tag{4.285}$$

Here the nonlinear susceptibility of each component is given by

$$\chi_{NL}^{(\pm)} = A|\mathcal{E}_\pm|^2 + (A+B)|\mathcal{E}_\mp|^2. \tag{4.286}$$

The corresponding effective refractive index, including linear as well as nonlinear parts, can be defined as

$$n_\pm^2 = 1 + \chi_L + \chi_{NL}^{(\pm)}. \tag{4.287}$$

Since in practice, $\chi_{NL} \ll \chi_L$, we can make the approximation

$$n_\pm \simeq n_L + \frac{\chi_{NL}^{(\pm)}}{2n_L}. \tag{4.288}$$

The analysis of Eqs. (4.284) – (4.288) reveals that in the circular polarization basis, the nonlinear wave equation in isotropic media can be effectively decoupled into the two as

$$\frac{\partial^2 E_\pm}{\partial t^2} - \frac{n_\pm^2}{c^2}\frac{\partial^2 E_\pm}{\partial z^2} = 0, \tag{4.289}$$

where we have neglected any spatial dependence in the transverse directions. Equations (4.289) are satisfied by the plane wave solutions

$$E_\pm(z,t) = \mathcal{E}_\pm e^{i(k_\pm z - \omega t)}, \tag{4.290}$$

where

$$k_\pm = \frac{n_\pm \omega}{c}. \tag{4.291}$$

On substituting from Eq. (4.288) into (4.290), and using the identities

$$n_\pm = \overline{n} \pm \Delta n/2, \tag{4.292}$$

where

$$\overline{n} \equiv \frac{n_+ + n_-}{2}, \qquad \Delta n = n_+ - n_-; \tag{4.293}$$

we obtain the expression for the total field as

$$\mathbf{E}(z,t) = [\mathcal{E}_+ e^{i\Delta n\omega z/2c}\mathbf{e}_+ + \mathcal{E}_- e^{-i\Delta n\omega z/2c}\mathbf{e}_-]e^{i\omega(\overline{n}z/c - t)}. \tag{4.294}$$

Here we have introduced the average effective refractive index

$$\overline{n} = n_L + \frac{(2A+B)}{4n_L}(|\mathcal{E}_+|^2 + |\mathcal{E}_-|^2), \tag{4.295}$$

and the refractive index difference,

$$\Delta n = n_+ - n_- = \frac{B}{2n_L}(|\mathcal{E}_-|^2 - |\mathcal{E}_+|^2), \tag{4.296}$$

respectively. The electric field can be represented as

$$\mathbf{E}(z,t) = [\mathcal{E}_+\mathbf{e}_+(z) + \mathcal{E}_-\mathbf{e}_-(z)]e^{i\omega(\overline{n}z/c - t)}, \tag{4.297}$$

where the rotating circular polarization basis is

$$\mathbf{e}_\pm(z) = \frac{\mathbf{e}_x(z) \pm i\mathbf{e}_y(z)}{\sqrt{2}}, \tag{4.298}$$

with

$$\mathbf{e}_x(z) = \cos(\Delta n\omega z/2c)\mathbf{e}_x + \sin(\Delta n\omega z/2c)\mathbf{e}_y, \tag{4.299}$$

$$\mathbf{e}_y(z) = \cos(\Delta n\omega z/2c)\mathbf{e}_y - \sin(\Delta n\omega z/2c)\mathbf{e}_x. \tag{4.300}$$

Exercise 4.16. *Verify that the representation of the field in terms of rotating polarization vectors, given by Eqs. (4.297) – (4.300), does indeed correspond to our field of*

Eq. (4.294).

Analyzing Eqs. (4.298) – (4.300), we can conclude that the electric field is elliptically polarized at any position z, according to Eq. (4.297); yet the polarization ellipse rotates in the $xy-$ plane at the rate proportional to the differences of refractive indices along the two principal axes. The latter is referred to as *birefringence*; it is the nonlinear birefringence of the medium that gives rise to polarization rotation even in isotropic optical media. Notice also that the rate of polarization rotation depends only on the coefficient B as is evidenced by Eqs. (4.296) and (4.299), (4.300). Hence the second term on the r.h.s. of Eq. (4.276) is wholly responsible for nonlinear birefringence effects. The first term on the r.h.s of (4.276) contributes to the overall phase accretion factor which is proportional to $\overline{n}$, but it does not affect polarization rotation.

Exercise 4.17. *In the fiber optical case, the nonlinear response of the medium is of electronic type such that $A = B$. Silica-glass optical fibers can serve as an important particular example. Linear birefringence of the fiber is typically introduced – either intentionally or inadvertently – at the fabrication stage. On account of linear birefringence, the most general field propagating in such a fiber can be represented as*

$$\mathbf{E} = \frac{1}{2}\left(\mathbf{e}_x\mathcal{E}_x e^{i\beta_x z} + \mathbf{e}_y\mathcal{E}_y e^{i\beta_y z}\right)e^{-i\omega t} + c.\,c,$$

where $\beta_{x,y}$ is the propagation constant of the corresponding linear polarization component; the field components are assumed to be polarized along the principal axes of the fiber. Show that the polarization field at the frequency ω is then given by

$$\mathbf{P}_{NL} = \frac{1}{2}\left(\mathbf{e}_x\mathcal{P}_x e^{i\beta_x z} + \mathbf{e}_y\mathcal{P}_y e^{i\beta_y z}\right)e^{-i\omega t} + c.\,c., \tag{4.301}$$

where

$$\mathcal{P}_x = \frac{3\epsilon_0}{4}\chi^{(3)}_{xxxx}\left[(|\mathcal{E}_x|^2 + \frac{2}{3}|\mathcal{E}_y|^2)\mathcal{E}_x + \frac{1}{3}\mathcal{E}_x^*\mathcal{E}_y^2 e^{-2i\Delta\beta z}\right], \tag{4.302}$$

$$\mathcal{P}_y = \frac{3\epsilon_0}{4}\chi^{(3)}_{xxxx}\left[(|\mathcal{E}_y|^2 + \frac{2}{3}|\mathcal{E}_x|^2)\mathcal{E}_x + \frac{1}{3}\mathcal{E}_y^*\mathcal{E}_x^2 e^{2i\Delta\beta z}\right]. \tag{4.303}$$

Here $\Delta\beta = \beta_x - \beta_y$.

4.14 Electro-optical Kerr effect

In this section, we study the electro-optical Kerr effect which manifests itself in the modification of a linear refractive index of an isotropic non-chiral media in presence of an electrostatic field. The effect becomes possible due to the second-order (Kerr) nonlinearity – which is the leading nonlinearity in such media – and hence the name, electro-optical Kerr effect. Classically, the corresponding polarization reads

$$\mathcal{P}_{NLi} = 3\epsilon_0\sum_{jkl}\chi^{(3)}_{ijkl}(-\omega,\omega,0,0)\mathcal{E}_j(\omega)E_k(0)E_l(0). \tag{4.304}$$

Substituting from Eq. (4.269) into Eq. (4.304), we obtain

$$\begin{aligned}\mathcal{P}_{NLi} &= 3\epsilon_0\chi^{(3)}_{xxyy}\mathcal{E}_i(\omega)\sum_k E_k^2(0) + 3\epsilon_0\chi^{(3)}_{xyxy}E_i(0)\sum_j E_j(0)\mathcal{E}_j(\omega) \\ &\quad + 3\epsilon_0\chi^{(3)}_{xyyx}E_i(0)\sum_j \mathcal{E}_j(\omega)E_j(0). \end{aligned} \tag{4.305}$$

The intrinsic permutation symmetry, $\chi^{(3)}_{ijkl}(-\omega,\omega,0,0) = \chi^{(3)}_{ijlk}(-\omega,\omega,0,0)$, implies that

$$\chi^{(3)}_{xyxy} = \chi^{(3)}_{xyyx}, \tag{4.306}$$

Using Eq. (4.306), Eq. (4.305) can be written in the vector form as

$$\boldsymbol{\mathcal{P}}_{NL} = 3\epsilon_0[\chi^{(3)}_{xxyy}\,\boldsymbol{\mathcal{E}}(\mathbf{E}_0\cdot\mathbf{E}_0) + 2\chi^{(3)}_{xyxy}\,\mathbf{E}_0(\boldsymbol{\mathcal{E}}\cdot\mathbf{E}_0)], \tag{4.307}$$

where we have introduced the notations,

$$\boldsymbol{\mathcal{E}} \equiv \boldsymbol{\mathcal{E}}(\omega) \quad \text{and} \quad \mathbf{E}_0 \equiv \mathbf{E}(0). \tag{4.308}$$

Equation (4.307) represents the general form of the polarization associated with the electro-optical Kerr effect for the electrostatic and optical fields of any polarizations. Let us focus on the case of linear polarization of the dc field such that

$$\mathbf{E}_0 = E_0\mathbf{e}_x, \qquad \boldsymbol{\mathcal{E}} = \mathcal{E}_x\mathbf{e}_x + \mathcal{E}_y\mathbf{e}_y. \tag{4.309}$$

Under the circumstances, the polarization components take the form

$$\begin{aligned}\mathcal{P}_{NLx} &= 3\epsilon_0[\chi^{(3)}_{xxyy}\mathcal{E}_xE_0^2 + 2\chi^{(3)}_{xyxy}E_0^2\mathcal{E}_x] \\ &= 3\epsilon_0[(\chi^{(3)}_{xxyy} + 2\chi^{(3)}_{xyxy})]E_0^2\mathcal{E}_x = 3\epsilon_0\chi^{(3)}_{xxxx}E_0^2\mathcal{E}_x, \end{aligned} \tag{4.310}$$

and

$$\mathcal{P}_{NLy} = 3\epsilon_0\chi^{(3)}_{xxyy}E_0^2\mathcal{E}_y. \tag{4.311}$$

It can be inferred from Eqs. (4.310) and (4.311) that the components of the total polarization field can be represented as

$$\mathcal{P}_{x,y} = \epsilon_0\chi_{x,y}\mathcal{E}_{x,y}, \tag{4.312}$$

where the components of the effective susceptibility tensor are

$$\chi_x = \chi^{(1)} + 3\chi^{(3)}_{xxxx}E_0^2, \tag{4.313}$$

and

$$\chi_y = \chi^{(1)} + 3\chi^{(3)}_{xxyy}E_0^2. \tag{4.314}$$

The corresponding components of the total refractive index are given by

$$n_x \simeq n + \frac{3\chi^{(3)}_{xxxx}}{2n}E_0^2, \tag{4.315}$$

and

$$n_y \simeq n + \frac{3\chi^{(3)}_{xxyy}}{2n} E_0^2. \tag{4.316}$$

It follows from Eqs. (4.312) – (4.314) that in electro-optical Kerr effect with the dc field breaking the azimuthal symmetry, it is the Cartesian components of the optical field that are decoupled; each satisfies the wave equation – neglecting spatial dependence in the transverse plane – of the form

$$\frac{\partial^2 E_{x,y}}{\partial t^2} - \frac{n^2_{x,y}}{c^2}\frac{\partial^2 E_{x,y}}{\partial z^2} = 0, \tag{4.317}$$

The plane-wave solutions to Eq. (4.317) are

$$E_{x,y}(z,t) = \mathcal{E}_{x,y} e^{i(k_{x,y}z - \omega t)}, \tag{4.318}$$

where

$$k_{x,y} = \frac{n_{x,y}\omega}{c}. \tag{4.319}$$

The evolution of the optical field can then be represented as

$$\mathbf{E}(z,t) = \mathcal{E}_x[\mathbf{e}_x + \mathbf{e}_y \tan\theta e^{-i\Delta n\omega z/c}]e^{i\omega(n_x z/c - t)}, \tag{4.320}$$

where $\tan\theta = \mathcal{E}_y/\mathcal{E}_x$, and

$$\Delta n = \frac{3\chi^{(3)}_{xyxy}E_0^2}{n}. \tag{4.321}$$

The analysis of Eq. (4.320) reveals that the Kerr effect is present in two guises: the dc field breaks the symmetry of the isotropic medium turning the medium into a uniaxial one and it generates effective linear birefringence which manifests itself in the polarization rotation. The latter can be seen by observing, for instance, that if the wave is initially linearly polarized at $45°$ to the dc field, such that $\tan\theta = 1$, it can acquire a circular polarization provided, $e^{-i\Delta n\omega L/c} = \pm i$ at the exit to the medium, $z = L$. In general, the polarization rotation angle for the beam having traversed a distance L inside the medium is given by

$$\Delta\phi_L = \frac{\Delta n\omega L}{c} = \frac{3\omega}{nc}\chi^{(3)}_{xyxy}E_0^2 L. \tag{4.322}$$

In experimental work, the so-called *Kerr constant* K is often introduced via the relation

$$\Delta n = K\lambda E_0^2. \tag{4.323}$$

The Kerr constant is related to relevant components of the susceptibility tensor by the expression

$$K = \frac{3\chi^{(3)}_{xyxy}}{n\lambda}. \tag{4.324}$$

Exercise 4.18. *Determine the dc field strength needed to produce a circular polarization by a 10 cm long Kerr cell filled with carbon disulfide, CS_2. For carbon disulfide, $K = 3.6 \times 10^{-14}$ m/V^2.*