Chapter 1

The Nonlinear Optical Susceptibility

1.1. Introduction to Nonlinear Optics

Nonlinear optics is the study of phenomena that occur as a consequence of the modification of the optical properties of a material system by the presence of light. Typically, only laser light is sufficiently intense to modify the optical properties of a material system. In fact, the beginning of the field of nonlinear optics is often taken to be the discovery of second-harmonic generation by Franken *et al.* in 1961, shortly after the demonstration of the first working laser by Maiman in 1960. Nonlinear optical phenomena are "nonlinear" in the sense that they occur when the response of a material system to an applied optical field depends in a nonlinear manner upon the strength of the optical field. For example, second-harmonic generation occurs as a result of the part of the atomic response that depends quadratically on the strength of the applied optical field. Consequently, the intensity of the light generated at the second-harmonic frequency tends to increase as the square of the intensity of the applied laser light.

In order to describe more precisely what we mean by an optical nonlinearity, let us consider how the dipole moment per unit volume, or polarization $\tilde{P}(t)$, of a material system depends upon the strength $\tilde{E}(t)$ of the applied optical field.* In the case of conventional (i.e., linear) optics, the induced polarization depends linearly upon the electric field strength in a manner that can often be

* Throughout the text, we use the tilde to denote a quantity that varies rapidly in time. Constant quantities, slowly varying quantities, and Fourier amplitudes are written without the tilde. See, for example, Eq. (1.2.1).

described by the relationship

$$\tilde{P}(t) = \chi^{(1)}\tilde{E}(t), \tag{1.1.1}$$

where the constant of proportionality $\chi^{(1)}$ is known as the linear susceptibility. In nonlinear optics, the optical response can often be described by generalizing Eq. (1.1.1) by expressing the polarization $\tilde{P}(t)$ as a power series in the field strength $\tilde{E}(t)$ as

$$\begin{aligned}\tilde{P}(t) &= \chi^{(1)}\tilde{E}(t) + \chi^{(2)}\tilde{E}^2(t) + \chi^{(3)}\tilde{E}^3(t) + \cdots \\ &\equiv \tilde{P}^{(1)}(t) + \tilde{P}^{(2)}(t) + \tilde{P}^{(3)}(t) + \cdots .\end{aligned} \tag{1.1.2}$$

The quantities $\chi^{(2)}$ and $\chi^{(3)}$ are known as the second- and third-order nonlinear optical susceptibilities, respectively. For simplicity, we have taken the fields $\tilde{P}(t)$ and $\tilde{E}(t)$ to be scalar quantities in writing Eqs. (1.1.1) and (1.1.2). In Section 1.3 we show how to treat the vector nature of the fields; in such a case $\chi^{(1)}$ becomes a second-rank tensor, $\chi^{(2)}$ becomes a third-rank tensor, etc. In writing Eqs. (1.1.1) and (1.1.2) in the form shown, we have also assumed that the polarization at time t depends only on the instantaneous value of the electric field strength. The assumption that the medium responds instantaneously also implies (through the Kramers–Kronig relations)* that the medium must be lossless and dispersionless. We shall also see in Section 1.3 how to generalize these equations for the case of a medium with dispersion and loss. In general, the nonlinear susceptibilities depend on the frequencies of the applied fields, but under our present assumption of instantaneous response we take them to be constants.

We shall refer to $\tilde{P}^{(2)}(t) = \chi^{(2)}\tilde{E}(t)^2$ as the second-order nonlinear polarization and to $\tilde{P}^{(3)}(t) = \chi^{(3)}\tilde{E}(t)^3$ as the third-order nonlinear polarization. We shall see later in this section that the physical processes that occur as a result of the second-order polarization $\tilde{P}^{(2)}$ are distinct from those that occur as a result of the third-order polarization $\tilde{P}^{(3)}$. In addition, we shall show in Section 1.5 that second-order nonlinear optical interactions can occur only in noncentrosymmetric crystals, that is, in crystals that do not display inversion symmetry. Since liquids, gases, amorphous solids (such as glass), and even many crystals do display inversion symmetry, $\chi^{(2)}$ vanishes identically for such media, and consequently they cannot produce second-order nonlinear optical interactions. On the other hand, third-order nonlinear optical interactions

* See, for example, Loudon (1973) Chapter 4 or the discussion in Section 1.7 of the present book for a discussion of the Kramers–Kronig relations.

(i.e., those described by a $\chi^{(3)}$ susceptibility) can occur both for centrosymmetric and noncentrosymmetric media.

We shall see in later sections of this book how to calculate the values of the nonlinear susceptibilities for various physical mechanisms that lead to optical nonlinearities. For the present, we shall make a simple order-of-magnitude estimate of the size of these quantities for the common case in which the nonlinearity is electronic in origin (see, for instance, Armstrong *et al.*, 1962). One might expect that the lowest-order correction term $\tilde{P}^{(2)}$ would be comparable to the linear response $\tilde{P}^{(1)}$ when the amplitude of the applied field $\tilde{E}$ is of the order of the characteristic atomic electric field strength $E_{\text{at}} = e/a_0^2$, where $-e$ is the charge of the electron and $a_0 = \hbar^2/me^2$ is the Bohr radius of the hydrogen atom (here $\hbar$ is Planck's constant divided by 2π, and m is the mass of the electron). Numerically, we find that $E_{\text{at}} = 2 \times 10^7$ statvolt/cm.* We thus expect that under conditions of nonresonant excitation the second-order susceptibility $\chi^{(2)}$ will be of the order of $\chi^{(1)}/E_{\text{at}}$. For condensed matter $\chi^{(1)}$ is of the order of unity, and we hence expect that $\chi^{(2)}$ will be of the order of $1/E_{\text{at}}$, or that

$$\chi^{(2)} \simeq 5 \times 10^{-8} \frac{\text{cm}}{\text{statvolt}} = 5 \times 10^{-8} (\text{cm}^3/\text{erg})^{1/2} = 5 \times 10^{-8} \text{ esu}. \quad (1.1.3)$$

Similarly, we expect $\chi^{(3)}$ to be of the order of $\chi^{(1)}/E_{\text{at}}^2$, which for condensed matter is of the order of

$$\chi^{(3)} \simeq 3 \times 10^{-15} \frac{\text{cm}^2}{\text{statvolt}^2} = 3 \times 10^{-15} \text{cm}^3/\text{erg} = 3 \times 10^{-15} \text{ esu}. \quad (1.1.4)$$

These predictions are in fact quite accurate, as one can see by comparing these values with actual measured values of $\chi^{(2)}$ (see for instance Table 1.5.3) and $\chi^{(3)}$ (see for instance Table 4.3.1). For certain purposes, it is useful to express the second- and third-order susceptibilities in terms of fundamental physical constants. Noting that the number density N of condensed matter is of the order of $(a_0)^{-3}$, we find that $\chi^{(2)} \simeq \hbar^4/m^2e^5$ and $\chi^{(3)} \simeq \hbar^8/m^4e^{10}$. See Boyd (1999) for further details.

The most common procedure for describing nonlinear optical phenomena is based on expressing the polarization $\tilde{P}(t)$ in terms of the applied electric field strength $\tilde{E}(t)$, as we have done in Eq. (1.1.2). The reason why the polarization plays a key role in the description of nonlinear optical phenomena is that a

* Except where otherwise noted, we use the gaussian system of units in this book. Note that in the scientific literature the units of an electrical quantity expressed in the gaussian system are often not given explicitly, but rather are simply said to be stated in electrostatic units (esu). As an example, in the present instance one would say that $E_{\text{at}} = 2 \times 10^7$ esu. See also the discussion in the appendix to this book on the conversion between the systems of units.

time-varying polarization can act as the source of new components of the electromagnetic field. For example, we shall see in Section 2.1 that the wave equation in nonlinear optical media often has the form

$$\nabla^2 \tilde{E} - \frac{n^2}{c^2}\frac{\partial^2 \tilde{E}}{\partial t^2} = \frac{4\pi}{c^2}\frac{\partial^2 \tilde{P}^{\mathrm{NL}}}{\partial t^2}, \tag{1.1.5}$$

where n is the usual linear refractive index and c is the speed of light in vacuum. We can interpret this expression as an inhomogeneous wave equation in which the polarization $\tilde{P}^{\mathrm{NL}}$ associated with the nonlinear response drives the electric field $\tilde{E}$. This equation expresses the fact that, whenever $\partial^2 \tilde{P}^{\mathrm{NL}}/\partial t^2$ is nonzero, charges are being accelerated, and according to Larmor's theorem from electromagnetism accelerated charges generate electromagnetic radiation.

It should be noted that the power series expansion expressed by Eq. (1.1.2) need not necessarily converge. In such circumstances the relationship between the material response and the applied electric field amplitude must be expressed using different procedures. One such example is that under resonant excitation of an atomic system, an appreciable fraction of the atoms can be removed from the ground state. Saturation effects of this sort can be described by procedures developed in Chapter 6. Even under nonresonant conditions, Eq. (1.1.2) loses its validity if the applied laser field strength becomes comparable to the characteristic atomic field strength E_{at}, because of strong photoionization that can occur under these conditions. For future reference, we note that the laser intensity associated with a peak field strength of E_{at} is given by

$$I_{\mathrm{at}} = \frac{c}{8\pi}E_{\mathrm{at}}^2 = 5 \times 10^{23}\ \mathrm{erg/cm^2 s} = 5 \times 10^{16}\ \mathrm{W/cm^2}. \tag{1.1.6}$$

We shall see in later sections of this book how nonlinear optical processes display qualitatively distinct features when excited by such super-intense fields.

1.2. Descriptions of Nonlinear Optical Interactions

In the present section, we present brief qualitative descriptions of a number of nonlinear optical interactions. In addition, for those processes that can occur in a lossless medium, we indicate how they can be described in terms of the nonlinear contributions to the polarization described by Eq. (1.1.2).* Our motivation is to provide the reader with an indication of the variety of nonlinear optical phenomena that can occur. These interactions are described in greater detail in later sections of this book. In this section we also introduce some notational conventions and some of the basic concepts of nonlinear optics.

* Recall that Eq. (1.1.2) is valid only for a medium that is lossless and dispersionless.

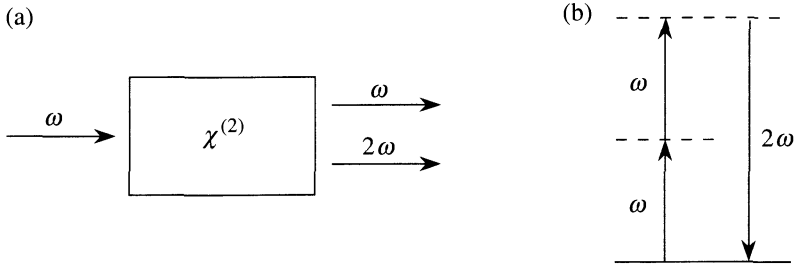

FIGURE 1.2.1 (a) Geometry of second-harmonic generation. (b) Energy-level diagram describing second-harmonic generation.

Second-Harmonic Generation

As an example of a nonlinear optical interaction, let us consider the process of second-harmonic generation, which is illustrated schematically in Fig. 1.2.1. Here a laser beam whose electric field strength is represented as

$$\tilde{E}(t) = Ee^{-i\omega t} + \text{c.c.} \tag{1.2.1}$$

is incident upon a crystal for which the second-order susceptibility $\chi^{(2)}$ is nonzero. The nonlinear polarization that is created in such a crystal is given according to Eq. (1.1.2) as $\tilde{P}^{(2)}(t) = \chi^{(2)}\tilde{E}^2(t)$ or as

$$\tilde{P}^{(2)}(t) = 2\chi^{(2)}EE^* + \left(\chi^{(2)}E^2e^{-2\omega t} + \text{c.c.}\right). \tag{1.2.2}$$

We see that the second-order polarization consists of a contribution at zero frequency (the first term) and a contribution at frequency 2ω (the second term). According to the driven wave equation (1.1.5), this latter contribution can lead to the generation of radiation at the second-harmonic frequency. Note that the first contribution in Eq. (1.2.2) does not lead to the generation of electromagnetic radiation (because its second time derivative vanishes); it leads to a process known as optical rectification in which a static electric field is created within the nonlinear crystal.

Under proper experimental conditions, the process of second-harmonic generation can be so efficient that nearly all of the power in the incident radiation at frequency ω is converted to radiation at the second-harmonic frequency 2ω. One common use of second-harmonic generation is to convert the output of a fixed-frequency laser to a different spectral region. For example, the Nd:YAG laser operates in the near infrared at a wavelength of 1.06 μm. Second-harmonic generation is routinely used to convert the wavelength of the radiation to 0.53 μm, in the middle of the visible spectrum.

Second-harmonic generation can be visualized by considering the interaction in terms of the exchange of photons between the various frequency components of the field. According to this picture, which is illustrated in part (b) of Fig. 1.2.1, two photons of frequency ω are destroyed and a photon of frequency 2ω is simultaneously created in a single quantum-mechanical process. The solid line in the figure represents the atomic ground state, and the dashed lines represent what are known as virtual levels. These levels are not energy eigenlevels of the free atom, but rather represent the combined energy of one of the energy eigenstates of the atom and of one or more photons of the radiation field.

The theory of second-harmonic generation is developed more fully in Section 2.6.

Sum- and Difference-Frequency Generation

Let us next consider the circumstance in which the optical field incident upon a nonlinear optical medium characterized by a nonlinear susceptibility $\chi^{(2)}$ consists of two distinct frequency components, which we represent in the form

$$\tilde{E}(t) = E_1 e^{-i\omega_1 t} + E_2 e^{-i\omega_2 t} + \text{c.c.} \tag{1.2.3}$$

Then, assuming as in Eq. (1.1.2) that the second-order contribution to the nonlinear polarization is of the form

$$\tilde{P}^{(2)}(t) = \chi^{(2)} \tilde{E}(t)^2, \tag{1.2.4}$$

we find that the nonlinear polarization is given by

$$\begin{aligned} \tilde{P}^{(2)}(t) = \chi^{(2)}\big[&E_1^2 e^{-2i\omega_1 t} + E_2^2 e^{-2i\omega_2 t} + 2E_1 E_2 e^{-i(\omega_1+\omega_2)t} \\ &+ 2E_1 E_2^* e^{-i(\omega_1-\omega_2)t} + \text{c.c.}\big] + 2\chi^{(2)}[E_1 E_1^* + E_2 E_2^*]. \end{aligned} \tag{1.2.5}$$

It is convenient to express this result using the notation

$$\tilde{P}^{(2)}(t) = \sum_n P(\omega_n) e^{-i\omega_n t}, \tag{1.2.6}$$

where the summation extends over positive and negative frequencies ω_n. The complex amplitudes of the various frequency components of the nonlinear polarization are hence given by

$$\begin{aligned} P(2\omega_1) &= \chi^{(2)} E_1^2 \quad \text{(SHG)}, \\ P(2\omega_2) &= \chi^{(2)} E_2^2 \quad \text{(SHG)}, \\ P(\omega_1 + \omega_2) &= 2\chi^{(2)} E_1 E_2 \quad \text{(SFG)}, \\ P(\omega_1 - \omega_2) &= 2\chi^{(2)} E_1 E_2^* \quad \text{(DFG)}, \\ P(0) &= 2\chi^{(2)}(E_1 E_1^* + E_2 E_2^*) \quad \text{(OR)}. \end{aligned} \tag{1.2.7}$$

Here we have labeled each expression by the name of the physical process that it describes, such as second-harmonic generation (SHG), sum-frequency generation (SFG), difference-frequency generation (DFG), and optical rectification (OR). Note that, in accordance with our complex notation, there is also a response at the negative of each of the nonzero frequencies given above:

$$\begin{aligned} P(-2\omega_1) &= \chi^{(2)} E_1^{*2}, & P(-2\omega_2) &= \chi^{(2)} E_2^{*2}, \\ P(-\omega_1 - \omega_2) &= 2\chi^{(2)} E_1^* E_2^*, & P(\omega_2 - \omega_1) &= 2\chi^{(2)} E_2 E_1^*. \end{aligned} \tag{1.2.8}$$

However, since each of these quantities is simply the complex conjugate of one of the quantities given in Eq. (1.2.7), it is not necessary to take explicit account of both the positive and negative frequency components.*

We see from Eq. (1.2.7) that four different nonzero frequency components are present in the nonlinear polarization. However, typically no more than one of these frequency components will be present with any appreciable intensity in the radiation generated by the nonlinear optical interaction. The reason for this behavior is that the nonlinear polarization can efficiently produce an output signal only if a certain phase-matching condition (which is discussed in detail in Section 2.7) is satisfied, and usually this condition cannot be satisfied for more than one frequency component of the nonlinear polarization. Operationally, one often chooses which frequency component will be radiated by properly selecting the polarization of the input radiation and orientation of the nonlinear crystal.

Sum-Frequency Generation

Let us now consider the process of sum-frequency generation, which is illustrated in Fig. 1.2.2. According to Eq. (1.2.7), the complex amplitude of the

* Not all workers in nonlinear optics use our convention that the fields and polarizations are given by Eqs. (1.2.3) and (1.2.6). Another common convention is to define the field amplitudes according to

$$\tilde{E}(t) = \tfrac{1}{2}(E_1' e^{-i\omega_1 t} + E_2' e^{-i\omega_2 t} + \text{c.c.}),$$

$$\tilde{P}^2(t) = \tfrac{1}{2} \sum_n P'(\omega_n) e^{i\omega_n t},$$

where in the second expression the summation extends over all positive and negative frequencies. Using this convention, one finds that

$$\begin{aligned} P'(2\omega_1) &= \tfrac{1}{2}\chi^{(2)} E_1'^2, & P'(2\omega_2) &= \tfrac{1}{2}\chi^{(2)} E_2'^2, \\ P'(\omega_1 + \omega_2) &= \chi^{(2)} E_1' E_2', & P'(\omega_1 - \omega_2) &= \chi^{(2)} E_1' E_2'^*, \\ P'(0) &= \chi^{(2)}(E_1' E_1'^* + E_2' E_2'^*). \end{aligned}$$

Note that these expressions differ from Eqs. (1.2.7) by factors of 1/2.

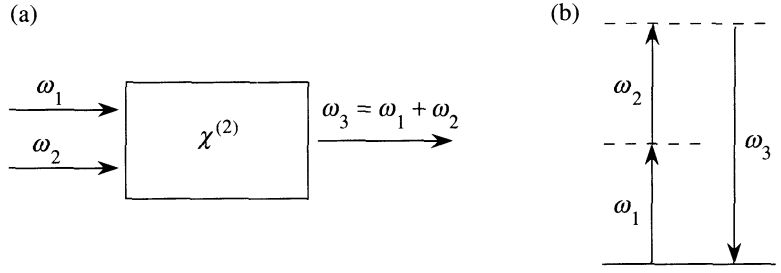

FIGURE 1.2.2 Sum-frequency generation. (a) Geometry of the interaction. (b) Energy-level description.

nonlinear polarization describing this process is given by the expression

$$P(\omega_1 + \omega_2) = 2\chi^{(2)} E_1 E_2. \tag{1.2.9}$$

In many ways the process of sum-frequency generation is analogous to that of second-harmonic generation, except that in sum-frequency generation the two input waves are at different frequencies. One application of sum-frequency generation is to produce tunable radiation in the ultraviolet spectral region by choosing one of the input waves to be the output of a fixed-frequency visible laser and the other to be the output of a frequency-tunable visible laser. The theory of sum-frequency generation is developed more fully in Sections 2.2 and 2.4.

Difference-Frequency Generation

The process of difference-frequency generation is described by a nonlinear polarization of the form

$$P(\omega_1 - \omega_2) = 2\chi^{(2)} E_1 E_2^* \tag{1.2.10}$$

and is illustrated in Fig. 1.2.3. Here the frequency of the generated wave is the difference of those of the applied fields. Difference-frequency generation can be used to produce tunable infrared radiation by mixing the output of a frequency-tunable visible laser with that of a fixed-frequency visible laser.

Superficially, difference-frequency generation and sum-frequency generation appear to be very similar processes. However, an important difference between the two processes can be deduced from the description of difference-frequency generation in terms of a photon energy-level diagram (part (b) of Fig. 1.2.3). We see that conservation of energy requires that for every photon that is created at the difference frequency $\omega_3 = \omega_1 - \omega_2$, a photon at the higher input frequency (ω_1) must be destroyed and a photon at the lower input

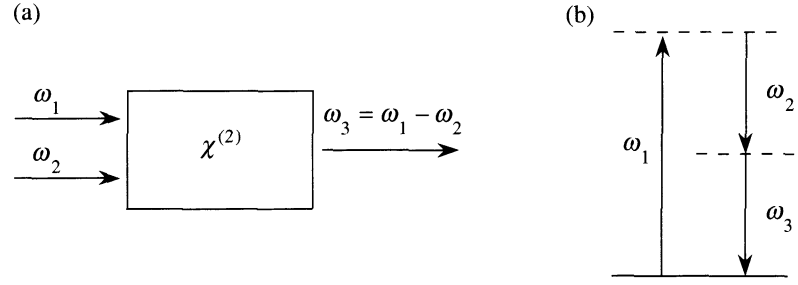

FIGURE 1.2.3 Difference-frequency generation. (a) Geometry of the interaction. (b) Energy-level description.

frequency (ω_2) must be created. Thus, the lower-frequency input field is amplified by the process of difference-frequency generation. For this reason, the process of difference-frequency generation is also known as optical parametric amplification. According to the photon energy-level description of difference-frequency generation, the atom first absorbs a photon of frequency ω_1 and jumps to the highest virtual level. This level decays by a two-photon emission process that is stimulated by the presence of the ω_2 field, which is already present. Two-photon emission can occur even if the ω_2 field is not applied. The generated fields in such a case are very much weaker, since they are created by *spontaneous* two-photon emission from a virtual level. This process is known as parametric fluorescence and has been observed experimentally (Harris *et al.*, 1967; Byer and Harris, 1968).

The theory of difference-frequency generation is developed more fully in Section 2.5.

Optical Parametric Oscillation

We have just seen that in the process of difference-frequency generation the presence of radiation at frequency ω_2 or ω_3 can stimulate the emission of additional photons at these frequencies. If the nonlinear crystal used in this process is placed inside an optical resonator, as shown in Fig. 1.2.4, the ω_2

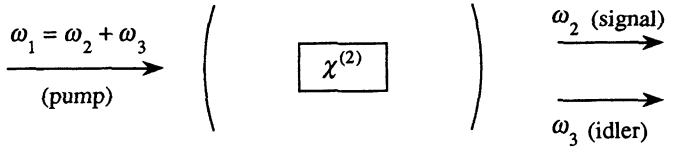

FIGURE 1.2.4 The optical parametric oscillator. The cavity end mirrors have high reflectivities at frequencies ω_2 and/or ω_3.

and/or ω_3 fields can build up to large values. Such a device is known as an optical parametric oscillator. Optical parametric oscillators are frequently used at infrared wavelengths, where other sources of tunable radiation are not readily available. Such a device is tunable because any frequency ω_2 that is smaller than ω_1 can satisfy the condition $\omega_2 + \omega_3 = \omega_1$ for some frequency ω_3. In practice, one controls the output frequency of an optical parametric oscillator by adjusting the phase-matching condition, as discussed in Section 2.7. The applied field frequency ω_1 is often called the pump frequency, the desired output frequency is called the signal frequency, and the other, unwanted, output frequency is called the idler frequency.

Third-Order Polarization

We next consider the third-order contribution to the nonlinear polarization

$$\tilde{P}^{(3)}(t) = \chi^{(3)}\tilde{E}(t)^3. \tag{1.2.11}$$

For the general case in which the field $\tilde{E}(t)$ is made up of several different frequency components, the expression for $\tilde{P}^{(3)}(t)$ is very complicated. For this reason, we first consider the simple case in which the applied field is monochromatic and is given by

$$\tilde{E}(t) = \mathscr{E}\cos\omega t. \tag{1.2.12}$$

Then, through use of the identity $\cos^3\omega t = \frac{1}{4}\cos 3\omega t + \frac{3}{4}\cos\omega t$, the nonlinear polarization can be expressed as

$$\tilde{P}^{(3)}(t) = \tfrac{1}{4}\chi^{(3)}\mathscr{E}^3\cos 3\omega t + \tfrac{3}{4}\chi^{(3)}\mathscr{E}^3\cos\omega t. \tag{1.2.13}$$

The significance of each of the two terms in this expression is described briefly below.

Third-Harmonic Generation

The first term in Eq. (1.2.13) describes a response at frequency 3ω that is due to an applied field at frequency ω. This term leads to the process of third-harmonic generation, which is illustrated in Fig. 1.2.5. According to the photon description of this process, shown in part (b) of the figure, three photons of frequency ω are destroyed and one photon of frequency 3ω is created in each elementary event.

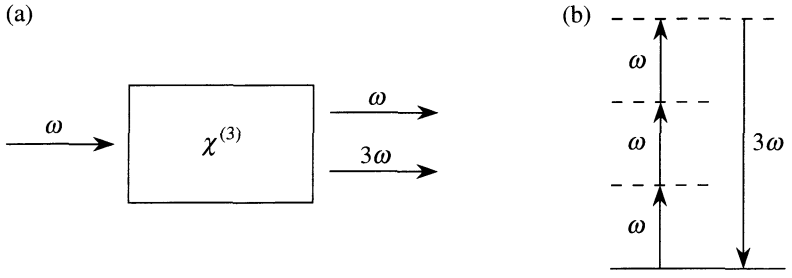

FIGURE 1.2.5 Third-harmonic generation. (a) Geometry of the interaction. (b) Energy-level description.

Intensity-Dependent Refractive Index

The second term in Eq. (1.2.13) describes a nonlinear contribution to the polarization at the frequency of the incident field; this term hence leads to a nonlinear contribution to the refractive index experienced by a wave at frequency ω. We shall see in Section 4.1 that the refractive index in the presence of this type of nonlinearity can be represented as

$$n = n_0 + n_2 I \tag{1.2.14a}$$

where n_0 is the usual (i.e., linear or low-intensity) refractive index, where

$$n_2 = \frac{12\pi^2}{n_0^2 c}\chi^{(3)} \tag{1.2.14b}$$

is an optical constant that characterizes the strength of the optical nonlinearity, and where $I = (n_0 c/8\pi)\mathscr{E}^2$ is the intensity of the incident wave.

Self-Focusing. One of the processes that can occur as a result of the intensity-dependent refractive index is self-focusing, which is illustrated in Fig. 1.2.6. This process can occur when a beam of light having a nonuniform transverse intensity distribution propagates through a material in which n_2 is positive. Under these conditions, the material effectively acts as a positive lens, which causes the rays to curve toward each other. This process is of great practical importance because the intensity at the focal spot of the self-focused beam is usually sufficiently large to lead to optical damage of the material. The process of self-focusing is described in greater detail in Section 7.1.

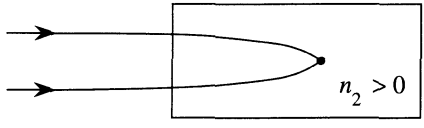

FIGURE 1.2.6 Self-focusing of light.

Third-Order Polarization (General Case)

Let us next examine the form of the nonlinear polarization

$$\tilde{P}^{(3)}(t) = \chi^{(3)}\tilde{E}(t)^3 \tag{1.2.15a}$$

induced by an applied field that consists of three frequency components:

$$\tilde{E}(t) = E_1 e^{-i\omega_1 t} + E_2 e^{-i\omega_2 t} + E_3 e^{-i\omega_3 t} + \text{c.c.} \tag{1.2.15b}$$

When we calculate $\tilde{E}(t)^3$, we find that the resulting expression contains 44 different frequency components, if we consider positive and negative frequencies to be distinct. Explicitly, these frequencies are

$$\omega_1, \omega_2, \omega_3, 3\omega_1, 3\omega_2, 3\omega_3, (\omega_1 + \omega_2 + \omega_3), (\omega_1 + \omega_2 - \omega_3),$$
$$(\omega_1 + \omega_3 - \omega_2), (\omega_2 + \omega_3 - \omega_1), (2\omega_1 \pm \omega_2), (2\omega_1 \pm \omega_3), (2\omega_2 \pm \omega_1),$$
$$(2\omega_2 \pm \omega_3), (2\omega_3 \pm \omega_1), (2\omega_3 \pm \omega_2),$$

and the negative of each. Again representing the nonlinear polarization as

$$\tilde{P}^{(3)}(t) = \sum_n P(\omega_n) e^{-i\omega_n t}, \tag{1.2.16}$$

we can write the complex amplitudes of the nonlinear polarization for the positive frequencies as

$$P(\omega_1) = \chi^{(3)}(3E_1E_1^* + 6E_2E_2^* + 6E_3E_3^*)E_1,$$
$$P(\omega_2) = \chi^{(3)}(6E_1E_1^* + 3E_2E_2^* + 6E_3E_3^*)E_2,$$
$$P(\omega_3) = \chi^{(3)}(6E_1E_1^* + 6E_2E_2^* + 3E_3E_3^*)E_3,$$

$$P(3\omega_1) = \chi^{(3)}E_1^3, \qquad P(3\omega_2) = \chi^{(3)}E_2^3, \qquad P(3\omega_3) = \chi^{(3)}E_3^3,$$

$$P(\omega_1 + \omega_2 + \omega_3) = 6\chi^{(3)}E_1E_2E_3, \qquad P(\omega_1 + \omega_2 - \omega_3) = 6\chi^{(3)}E_1E_2E_3^*,$$
$$P(\omega_1 + \omega_3 - \omega_2) = 6\chi^{(3)}E_1E_3E_2^*, \qquad P(\omega_2 + \omega_3 - \omega_1) = 6\chi^{(3)}E_2E_3E_1^*,$$
$$P(2\omega_1 + \omega_2) = 3\chi^{(3)}E_1^2E_2, \qquad P(2\omega_1 + \omega_3) = 3\chi^{(3)}E_1^2E_3,$$
$$P(2\omega_2 + \omega_1) = 3\chi^{(3)}E_2^2E_1, \qquad P(2\omega_2 + \omega_3) = 3\chi^{(3)}E_2^2E_3,$$
$$P(2\omega_3 + \omega_1) = 3\chi^{(3)}E_3^2E_1, \qquad P(2\omega_3 + \omega_2) = 3\chi^{(3)}E_3^2E_2,$$
$$P(2\omega_1 - \omega_2) = 3\chi^{(3)}E_1^2E_2^*, \qquad P(2\omega_1 - \omega_3) = 3\chi^{(3)}E_1^2E_3^*,$$
$$P(2\omega_2 - \omega_1) = 3\chi^{(3)}E_2^2E_1^*, \qquad P(2\omega_2 - \omega_3) = 3\chi^{(3)}E_2^2E_3^*,$$
$$P(2\omega_3 - \omega_1) = 3\chi^{(3)}E_3^2E_1^*, \qquad P(2\omega_3 - \omega_2) = 3\chi^{(3)}E_3^2E_2^* \tag{1.2.17}$$

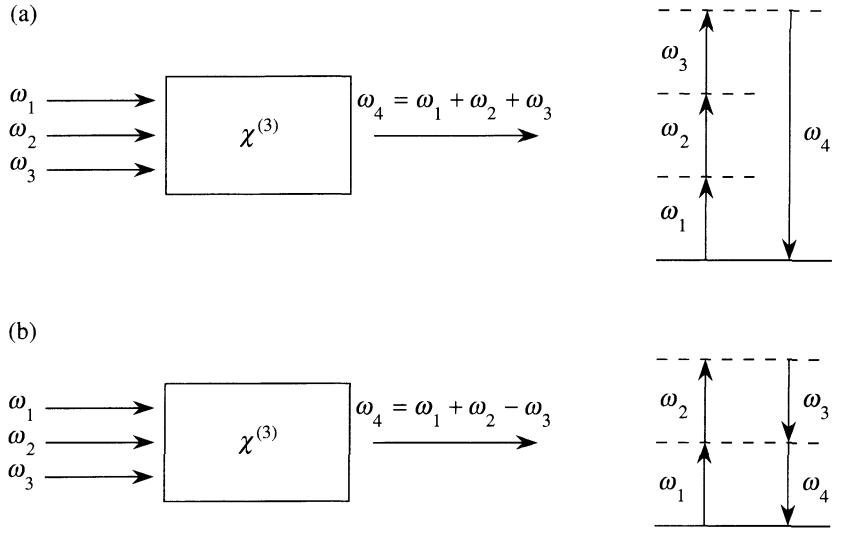

FIGURE 1.2.7 Two of the possible mixing processes described by Eq. (1.2.17) that can occur when three input waves interact in a medium characterized by a $\chi^{(3)}$ susceptibility.

We have displayed these expressions in complete detail because it is very instructive to study their form. In each case the frequency argument of P is equal to the sum of the frequencies associated with the field amplitudes appearing on the right-hand side of the equation, if we adopt the convention that a negative frequency is to be associated with a field amplitude that appears as a complex conjugate. Also, the numerical factor (1, 3, or 6) that appears in each term on the right-hand side of each equation is equal to the number of distinct permutations of the field frequencies that contribute to that term.

Some of the nonlinear optical mixing processes described by Eq. (1.2.17) are illustrated in Fig. 1.2.7.

Parametric versus Nonparametric Process

All of the processes described thus far in this chapter are examples of what are known as parametric processes. The origin of this terminology is obscure, but the word parametric has come to denote a process in which the initial and final quantum-mechanical states of the system are identical. Consequently, in a parametric process population can be removed from the ground state only for those brief intervals of time when it resides in a virtual level. According to the uncertainty principle, population can reside in a virtual level for a time interval of the order of $\hbar/\delta E$, where δE is the energy difference between the

virtual level and the nearest real level. Conversely, processes that do involve the transfer of population from one real level to another are known as nonparametric processes. The processes that we describe in the remainder of the present section are all examples of nonparametric processes.

One difference between parametric and nonparametric processes is that parametric processes can always be described by a real susceptibility; conversely, nonparametric processes are described by a complex susceptibility by means of a procedure described in the following section, Section 1.3. Another difference is that photon energy is always conserved in a parametric process; photon energy need not be conserved in a nonparametric process, because energy can be transferred to or from the material medium. For this reason, photon energy level diagrams of the sort shown in Figs. 1.2.1, 1.2.2, 1.2.3, 1.2.5, and 1.2.7 to describe parametric processes play a less definitive role in describing non-parametric processes.

As a simple example of the distinction between parametric and nonparametric processes, we consider the case of the usual (linear) index of refraction. The real part of the refractive index is a consequence of parametric processes, whereas its imaginary part is a consequence of nonparametric processes, since the imaginary part of the refractive index describes the absorption of radiation, which results from the transfer of population from the atomic ground state to an excited state.

Saturable Absorption

One example of a nonparametric nonlinear optical process is saturable absorption. Many material systems have the property that their absorption coefficient decreases when measured using high laser intensity. Often the dependence of the measured absorption coefficient α on the intensity I of the incident laser radiation is given by the expression*

$$\alpha = \frac{\alpha_0}{1 + I/I_s}, \tag{1.2.18}$$

where α_0 is the low-intensity absorption coefficient, and I_s is a parameter known as the saturation intensity.

Optical Bistability. One consequence of saturable absorption is optical bistability. One way of forming a bistable optical device is to place a saturable absorber inside a Fabry–Perot resonator, as illustrated in Fig. 1.2.8. As the

* This form is valid, for instance, for the case of homogeneous broadening of a simple atomic transition.

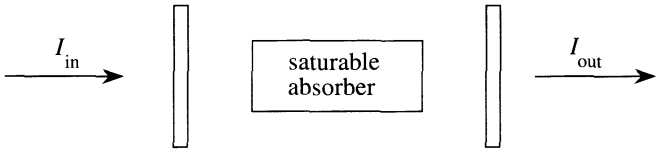

FIGURE 1.2.8 Bistable optical device.

input intensity is increased, the field inside the cavity also increases, lowering the absorption that the field experiences and thus increasing the field intensity still further. If the intensity of the incident field is subsequently lowered, the field inside the cavity tends to remain large because the absorption of the material system has already been reduced. A plot of the input-versus-output characteristics thus looks qualitatively like that shown in Fig. 1.2.9. Note that over an appreciable range of input intensities more than one output intensity is possible. The process of optical bistability is described in greater detail in Section 7.3.

Two-Photon Absorption

In the process of two-photon absorption, which is illustrated in Fig. 1.2.10, an atom makes a transition from its ground state to an excited state by the simultaneous absorption of two laser photons. The absorption cross section σ describing this process increases linearly with laser intensity according to the relation

$$\sigma = \sigma^{(2)} I, \tag{1.2.19}$$

where $\sigma^{(2)}$ is a coefficient that describes two-photon absorption. (Recall that in conventional, linear optics the absorption cross section σ is a constant.) Consequently, the atomic transition rate R due to two-photon absorption scales

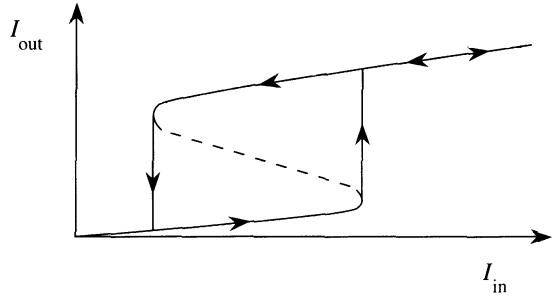

FIGURE 1.2.9 Typical input-versus-output characteristics of a bistable optical device.

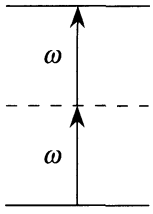

FIGURE 1.2.10 Two-photon absorption.

as the square of the laser intensity, since $R = \sigma I/\hbar\omega$, or as

$$R = \frac{\sigma^{(2)} I^2}{\hbar\omega}. \tag{1.2.20}$$

Two-photon absorption is a useful spectroscopic tool for determining the positions of energy levels that are not connected to the atomic ground state by a one-photon transition. Two-photon absorption was first observed experimentally by Kaiser and Garrett (1961).

Stimulated Raman Scattering

In stimulated Raman scattering, which is illustrated in Fig. 1.2.11, a photon of frequency ω is annihilated and a photon at the Stokes shifted frequency $\omega_s = \omega - \omega_v$ is created, leaving the molecule (or atom) in an excited state with energy $\hbar\omega_v$. The excitation energy is referred to as ω_v because stimulated Raman scattering was first studied in molecular systems, where $\hbar\omega_v$ corresponds to a vibrational energy. The efficiency of this process can be quite large, with often 10% or more of the power of the incident light being converted to the Stokes frequency. In contrast, the efficiency of normal or spontaneous Raman scattering is typically many orders of magnitude smaller. Stimulated Raman scattering is described more fully in Chapter 9.

Other stimulated scattering processes such as stimulated Brillouin scattering and stimulated Rayleigh scattering also occur and are described more fully in Chapter 8.

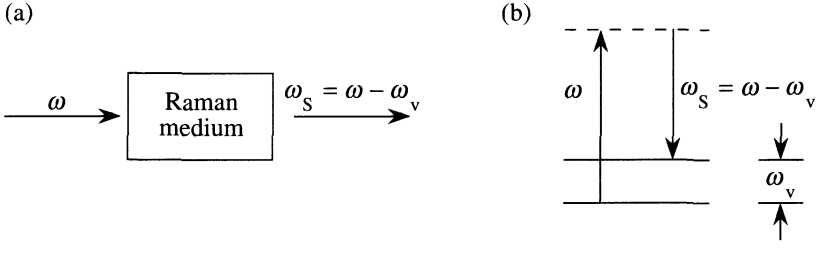

FIGURE 1.2.11 Stimulated Raman scattering.

1.3. Formal Definition of the Nonlinear Susceptibility

Nonlinear optical interactions can be described in terms of the nonlinear polarization given by Eq. (1.1.2) only for a material system that is lossless and dispersionless. In the present section, we consider the more general case of a material with dispersion and/or loss. In this general case the nonlinear susceptibility becomes a complex quantity relating the complex amplitudes of the electric field and polarization.

We assume that we can represent the electric field vector of the optical wave as the discrete sum of a number of frequency components as

$$\tilde{\mathbf{E}}(\mathbf{r}, t) = \sum_n{}' \tilde{\mathbf{E}}_n(\mathbf{r}, t). \tag{1.3.1}$$

The prime on the summation sign of Eq. (1.3.1) indicates that the summation is to be taken over positive frequencies only. It is often convenient to represent $\tilde{\mathbf{E}}_n(\mathbf{r}, t)$ as the sum of its positive- and negative-frequency parts as

$$\tilde{\mathbf{E}}_n = \tilde{\mathbf{E}}_n^{(+)} + \tilde{\mathbf{E}}_n^{(-)}, \tag{1.3.2}$$

where

$$\tilde{\mathbf{E}}_n^{(+)} = \mathbf{E}_n e^{-i\omega_n t} \tag{1.3.3a}$$

and

$$\tilde{\mathbf{E}}_n^{(-)} = \tilde{\mathbf{E}}_n^{(+)*} = \mathbf{E}_n^* e^{i\omega_n t}. \tag{1.3.3b}$$

By requiring $\tilde{\mathbf{E}}_n^{(-)}$ to be the complex conjugate of $\tilde{\mathbf{E}}_n^{(+)}$ we are assured that the quantity of $\tilde{\mathbf{E}}(\mathbf{r}, t)$ of Eq. (1.3.1) will be real, as it must be in order to represent a physical field. It is also convenient to define the spatially slowly varying field amplitude $\mathbf{A}_n$ by means of the relation

$$\mathbf{E}_n = \mathbf{A}_n e^{i\mathbf{k}_n \cdot \mathbf{r}}. \tag{1.3.4}$$

The total electric field of Eq. (1.3.1) can thus be represented in terms of these field amplitudes by either of the expressions

$$\begin{aligned} \tilde{\mathbf{E}}(\mathbf{r}, t) &= \sum_n{}' \mathbf{E}_n e^{-i\omega_n t} + \text{c.c.} \\ &= \sum_n{}' \mathbf{A}_n e^{(i\mathbf{k}_n \cdot \mathbf{r} - \omega_n t)} + \text{c.c.} \end{aligned} \tag{1.3.5}$$

On occasion, we shall express these field amplitudes using the alternative notation

$$\mathbf{E}_n = \mathbf{E}(\omega_n) \quad \text{and} \quad \mathbf{A}_n = \mathbf{A}(\omega_n). \tag{1.3.6}$$

In terms of this new notation, the reality condition of Eq. (1.3.3b) becomes

$$\mathbf{E}(-\omega_n) = \mathbf{E}(\omega_n)^* \quad \text{or} \quad \mathbf{A}(-\omega_n) = \mathbf{A}(\omega_n)^*. \tag{1.3.7}$$

Using this new notation, we can write the total field in the more compact form

$$\begin{aligned} \tilde{\mathbf{E}}(\mathbf{r}, t) &= \sum_n \mathbf{E}(\omega_n) e^{-i\omega_n t} \\ &= \sum_n \mathbf{A}(\omega_n) e^{i(\mathbf{k}_n \cdot \mathbf{r} - \omega_n t)}, \end{aligned} \tag{1.3.8}$$

where the unprimed summation symbol denotes a summation over all frequencies, both positive and negative.

Note that according to our definition of field amplitude, the field given by

$$\tilde{\mathbf{E}}(\mathbf{r}, t) = \mathscr{E} \cos(\mathbf{k} \cdot \mathbf{r} - \omega t) \tag{1.3.9}$$

is represented by the complex field amplitudes

$$\mathbf{E}(\omega) = \tfrac{1}{2}\mathscr{E} e^{i\mathbf{k}\cdot\mathbf{r}}, \qquad \mathbf{E}(-\omega) = \tfrac{1}{2}\mathscr{E} e^{-i\mathbf{k}\cdot\mathbf{r}}, \tag{1.3.10}$$

or alternatively by the slowly varying amplitudes

$$\mathbf{A}(\omega) = \tfrac{1}{2}\mathscr{E}, \qquad \mathbf{A}(-\omega) = \tfrac{1}{2}\mathscr{E}. \tag{1.3.11}$$

In either representation, factors of $1/2$ appear because the physical field amplitude $\mathscr{E}$ has been divided equally between the positive- and negative-frequency components.

Using a notation similar to that of Eq. (1.3.8), we can express the nonlinear polarization as

$$\tilde{\mathbf{P}}(\mathbf{r}, t) = \sum_n \mathbf{P}(\omega_n) e^{-i\omega_n t}, \tag{1.3.12}$$

where, as before, the summation extends over all positive- and negative-frequency components.

We now define the components of the second-order susceptibility tensor $\chi^{(2)}_{ijk}(\omega_n + \omega_m, \omega_n, \omega_m)$ as the constants of proportionality relating the amplitude of the nonlinear polarization to the product of field amplitudes according to

$$P_i(\omega_n + \omega_m) = \sum_{jk} \sum_{(nm)} \chi^{(2)}_{ijk}(\omega_n + \omega_m, \omega_n, \omega_m) E_j(\omega_n) E_k(\omega_m). \tag{1.3.13}$$

Here the indices ijk refer to the cartesian components of the fields. The notation (nm) indicates that, in performing the summation over n and m, the sum $\omega_n + \omega_m$ is to be held fixed, although ω_n and ω_m are each allowed to vary. Since the amplitude $E(\omega_n)$ is associated with the time dependence $\exp(-i\omega_n t)$, and the amplitude $E(\omega_m)$ is associated with the time dependence $\exp(-i\omega_m t)$, their product $E(\omega_n)E(\omega_m)$ is associated with the time dependence $\exp[-i(\omega_n + \omega_m)t]$. Hence the product $E(\omega_n)E(\omega_m)$ does in fact lead

to a contribution to the nonlinear polarization oscillating at frequency $\omega_n + \omega_m$, as the notation of Eq. (1.3.13) suggests. Following convention, we have written $\chi^{(2)}$ as a function of three frequency arguments. This is technically unnecessary in that the first argument is always the sum of the other two. To emphasize this fact, the susceptibility $\chi^{(2)}(\omega_3, \omega_2, \omega_1)$ is sometimes written as $\chi^{(2)}(\omega_3; \omega_2, \omega_1)$ as a reminder that the first argument is different from the other two; or it may be written symbolically as $\chi^{(2)}(\omega_3 = \omega_2 + \omega_1)$.

Let us examine some of the consequences of the definition of the nonlinear susceptibility as given by Eq. (1.3.13) by considering two simple examples.

1. *Sum-frequency generation.* We let the input field frequencies be ω_1 and ω_2 and the sum frequency be ω_3, so that $\omega_3 = \omega_1 + \omega_2$. Then, by carrying out the summation over ω_n and ω_m in Eq. (1.3.13), we find that

$$P_i(\omega_3) = \sum_{jk} \big[\chi^{(2)}_{ijk}(\omega_3, \omega_1, \omega_2) E_j(\omega_1) E_k(\omega_2) + \chi^{(2)}_{ijk}(\omega_3, \omega_2, \omega_1) E_j(\omega_2) E_k(\omega_1)\big]. \tag{1.3.14}$$

This expression can be simplified by making use of the intrinsic permutation symmetry of the nonlinear susceptibility (this symmetry is discussed in more detail in Eq. (1.5.6) below), which requires that

$$\chi^{(2)}_{ijk}(\omega_m + \omega_n, \omega_m, \omega_n) = \chi^{(2)}_{ikj}(\omega_m + \omega_n, \omega_n, \omega_m). \tag{1.3.15}$$

Through use of this relation, the expression for the nonlinear polarization becomes

$$P_i(\omega_3) = 2 \sum_{jk} \chi^{(2)}_{ijk}(\omega_3, \omega_1, \omega_2) E_j(\omega_1) E_k(\omega_2), \tag{1.3.16}$$

and for the special case in which both input fields are polarized in the x direction the polarization becomes

$$P_i(\omega_3) = 2\chi^{(2)}_{ixx}(\omega_3, \omega_1, \omega_2) E_x(\omega_1) E_x(\omega_2). \tag{1.3.17}$$

2. *Second-harmonic generation.* We take the input frequency as ω_1 and the generated frequency as $\omega_3 = 2\omega_1$. If we again perform the summation over field frequencies in Eq. (1.3.13), we obtain

$$P_i(\omega_3) = \sum_{jk} \chi^{(2)}_{ijk}(\omega_3, \omega_1, \omega_1) E_j(\omega_1) E_k(\omega_1). \tag{1.3.18}$$

Again assuming the special case of an input field polarization along the x direction, this result becomes

$$P_i(\omega_3) = \chi^{(2)}_{ixx}(\omega_3, \omega_1, \omega_1) E_x(\omega_1)^2. \tag{1.3.19}$$

Note that a factor of two appears in Eqs. (1.3.16) and (1.3.17), which describe sum-frequency generation, but not in Eqs. (1.3.18) and (1.3.19), which describe second-harmonic generation. The fact that these expressions remain different even as ω_2 approaches ω_1 is at first sight surprising, but is a consequence of our convention that $\chi^{(2)}_{ijk}(\omega_3, \omega_1, \omega_2)$ must approach $\chi^{(2)}_{ijk}(\omega_3, \omega_1, \omega_1)$ as ω_1 approaches ω_2. Note that the expressions for $P(2\omega_2)$ and $P(\omega_1+\omega_2)$ that apply for the case of a dispersionless nonlinear susceptibility (Eq. (1.2.7)) also differ by a factor of two. Moreover, one should expect the nonlinear polarization produced by two distinct fields to be larger than that produced by a single field (both of the same amplitude, say), because the total light intensity is larger in the former case.

In general, the summation over field frequencies ($\sum_{(nm)}$) in Eq. (1.3.13) can be performed formally to obtain the result

$$P_i(\omega_n + \omega_m) = D \sum_{jk} \chi^{(2)}_{ijk}(\omega_n + \omega_m, \omega_n, \omega_m) E_j(\omega_n) E_k(\omega_m), \quad (1.3.20)$$

where D is known as the degeneracy factor and is equal to the number of distinct permutations of the applied field frequencies ω_n and ω_m.

The expression (1.3.13) defining the second-order susceptibility can readily be generalized to higher-order interactions. In particular, the components of the third-order susceptibility are defined as the coefficients relating the amplitudes according to the expression

$$\begin{aligned} P_i(\omega_o + \omega_n + \omega_m) = \sum_{jkl} \sum_{(mno)} \chi^{(3)}_{ijkl}(\omega_0 + \omega_n + \omega_m, \omega_o, \omega_n, \omega_m) \\ \times E_j(\omega_o) E_k(\omega_n) E_l(\omega_m). \end{aligned} \quad (1.3.21)$$

We can again perform the summation over m, n, and o to obtain the result

$$\begin{aligned} P_i(\omega_o + \omega_n + \omega_m) = D \sum_{jkl} \chi^{(3)}_{ijkl}(\omega_0 + \omega_n + \omega_m, \omega_o, \omega_n, \omega_m) \\ \times E_j(\omega_0) E_k(\omega_n) E_l(\omega_m), \end{aligned} \quad (1.3.22)$$

where the degeneracy factor D represents the number of distinct permutations of the frequencies ω_m, ω_n, and ω_o.

[redacted]

$$\tilde{P}(t) = \chi^{(2)}\tilde{E}(t)^2,$$

Contracted Notation

We now introduce a notational device that is often used when the Kleinman symmetry condition is valid. We introduce the tensor

$$d_{ijk} = \tfrac{1}{2}\chi^{(2)}_{ijk} \tag{1.5.20}$$

and for simplicity suppress the frequency arguments. The nonlinear polarization can then be written as

$$P_i(\omega_n + \omega_m) = \sum_{jk}\sum_{(nm)} 2d_{ijk}E_j(\omega_n)E_k(\omega_m). \tag{1.5.21}$$

We now assume that d_{ijk} is symmetric in its last two indices. This assumption is valid whenever Kleinman's symmetry condition is valid and in addition is valid in general for second-harmonic generation, since in this case ω_n and ω_m are equal. We then simplify the notation by introducing a contracted matrix d_{il} according to the prescription

$$\begin{array}{cccccccc} jk: & 11 & 22 & 33 & 23,32 & 31,13 & 12,21 \\ l: & 1 & 2 & 3 & 4 & 5 & 6 \end{array} \tag{1.5.22}$$

The nonlinear susceptibility tensor can then be represented as the 3×6 matrix

$$d_{il} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \end{bmatrix}. \tag{1.5.23}$$

If we now *explicitly* introduce the Kleinman symmetry condition, i.e., we assert that the indices d_{ijk} can be freely permuted, we find that not all of the 18 elements of d_{il} are independent. For instance, we see that

$$d_{12} \equiv d_{122} = d_{212} \equiv d_{26} \tag{1.5.24a}$$

and that

$$d_{14} \equiv d_{123} = d_{213} \equiv d_{25}. \tag{1.5.24b}$$

By applying this type of argument systematically, we find that d_{il} has only 10 independent elements when the Kleinman symmetry condition is valid; the form of d_{il} under these conditions is

$$d_{il} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{16} & d_{22} & d_{23} & d_{24} & d_{14} & d_{12} \\ d_{15} & d_{24} & d_{33} & d_{23} & d_{13} & d_{14} \end{bmatrix}. \tag{1.5.25}$$

We can describe the nonlinear polarization leading to second-harmonic generation in terms of d_{il} by the matrix equation

$$\begin{bmatrix} P_x(2\omega) \\ P_y(2\omega) \\ P_z(2\omega) \end{bmatrix} = 2 \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \end{bmatrix} \begin{bmatrix} E_x(\omega)^2 \\ E_y(\omega)^2 \\ E_z(\omega)^2 \\ 2E_y(\omega)E_z(\omega) \\ 2E_x(\omega)E_z(\omega) \\ 2E_x(\omega)E_y(\omega) \end{bmatrix}. \tag{1.5.26}$$

When the Kleinman symmetry condition is valid, we can describe the nonlinear polarization leading to sum-frequency generation (with $\omega_3 = \omega_1 + \omega_2$) by the equation

$$\begin{bmatrix} P_x(\omega_3) \\ P_y(\omega_3) \\ P_z(\omega_3) \end{bmatrix} = 4 \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \end{bmatrix} \times \begin{bmatrix} E_x(\omega_1)E_x(\omega_2) \\ E_y(\omega_1)E_y(\omega_2) \\ E_z(\omega_1)E_z(\omega_2) \\ E_y(\omega_1)E_z(\omega_2) + E_z(\omega_1)E_y(\omega_2) \\ E_x(\omega_1)E_z(\omega_2) + E_z(\omega_1)E_x(\omega_2) \\ E_x(\omega_1)E_y(\omega_2) + E_y(\omega_1)E_x(\omega_2) \end{bmatrix}. \quad (1.5.27)$$

As described above in relation to Eq. (1.3.17), the extra factor of 2 comes from the summation over n and m in Eq. (1.3.13).

Effective Value of d (d_{eff})

For a fixed geometry (i.e., for fixed propagation and polarization directions) it is possible to express the nonlinear polarization giving rise to sum-frequency generation by means of the scalar relationship

$$P(\omega_3) = 4d_{\text{eff}}E(\omega_1)E(\omega_2), \quad (1.5.28)$$

and analogously for second-harmonic generation by

$$P(2\omega) = 2d_{\text{eff}}E(\omega)^2, \quad (1.5.29)$$

where

$$E(\omega) = |\mathbf{E}(\omega)| = \left[\sum_j E_j(\omega)^2\right]^{1/2}.$$

In each case, d_{eff} is obtained by evaluation of the summation $\sum_{jk}$ in the general equation (1.3.13).

A general prescription for calculating d_{eff} for each of the crystal classes has been presented by Midwinter and Warner (1965); see also Table 3.1 of Zernike and Midwinter (1973). They show, for example, that for a negative uniaxial

crystal of crystal class $3m$ the effective value of d is given by the expression

$$d_{\text{eff}} = d_{31} \sin\theta - d_{22} \cos\theta \sin 3\phi \tag{1.5.30a}$$

under conditions (known as type I conditions) such that the two lower-frequency waves have the same polarization, and by

$$d_{\text{eff}} = d_{22} \cos^2\theta \cos 3\phi \tag{1.5.30b}$$

under conditions (known as type II conditions) such that the polarizations are orthogonal. In these equations, θ is the angle between the propagation vector and the crystalline z axis (the optic axis), and ϕ is the azimuthal angle between the propagation vector and the xz crystalline plane.

Spatial Symmetry of the Nonlinear Medium

The form of the linear and nonlinear susceptibility tensors is constrained by the symmetry properties of the optical medium. To see why this should be so, let us consider a crystal for which the x and y directions are equivalent but for which the z direction is different. By saying that the x and y directions are equivalent, we mean that if the crystal were rotated by 90 degrees about the z axis, the crystal structure would look identical after the rotation. The z axis is then said to be a fourfold axis of symmetry. For such a crystal, we would expect that the optical response would be the same for an applied optical field polarized in either the x or the y direction, and thus, for example, that the second-order susceptibility components $\chi^{(2)}_{zxx}$ and $\chi^{(2)}_{zyy}$ would be equal.

For any particular crystal, the form of the linear and nonlinear optical susceptibilities is determined by considering the consequences of all of the symmetry properties for that particular crystal. For this reason, it is necessary to determine what types of symmetry properties can occur in a crystalline medium. By means of the mathematical method known as group theory, crystallographers have found that all crystals can be classified as belonging to one of 32 possible crystal classes depending on what is called the point group symmetry of the crystal. The details of this classification scheme lie outside of the subject matter of the present text.* However, by way of example, a crystal is said to belong to point group 4 if it possesses only a fourfold axis of symmetry, to point group 3 if it possesses only a threefold axis of symmetry, and to belong to point group $3m$ if it possesses a threefold axis of symmetry and in addition a plane of mirror symmetry perpendicular to this axis.

* The reader who is interested in the details should consult Buerger (1963) or any of the other books on group theory and crystal symmetry listed in the bibliography at the end of the present chapter.

Influence of Spatial Symmetry on the Linear Optical Properties of a Material Medium

As an illustration of the consequences of spatial symmetry on the optical properties of a material system, let us first consider the restrictions that this symmetry imposes on the form of the linear susceptibility tensor $\chi^{(1)}$. The results of a group theroretical analysis of this problem shows that five different cases are possible depending on the symmetry properties of the material system. These possibilities are summarized in Table 1.5.1. Each entry is labeled by the crystal system to which the material belongs. By convention, crystals are categorized in terms of seven possible crystal systems on the basis of the form of the crystal lattice. (Table 1.5.2 below gives the correspondence between crystal system and each of the 32 point groups.) For completeness, isotropic materials (such as liquids and gases) are also included in Table 1.5.1. We see from this table that cubic and isotropic materials are isotropic in their linear optical properties, because $\chi^{(1)}$ is diagonal with equal diagonal components. All of the

TABLE 1.5.1 Form of the linear susceptibility tensor $\chi^{(1)}$ as determined by the symmetry properties of the optical medium, for each of the seven crystral classes and for isotropic materials. Each nonvanishing element is denoted by its cartesian indices

Triclinic	$\begin{bmatrix} xx & xy & xz \\ yx & yy & yz \\ zx & zy & zz \end{bmatrix}$
Monoclinic	$\begin{bmatrix} xx & 0 & xz \\ 0 & yy & 0 \\ zx & 0 & zz \end{bmatrix}$
Orthorhombic	$\begin{bmatrix} xx & 0 & 0 \\ 0 & yy & 0 \\ 0 & 0 & zz \end{bmatrix}$
Tetragonal Trigonal Hexagonal	$\begin{bmatrix} xx & 0 & 0 \\ 0 & xx & 0 \\ 0 & 0 & zz \end{bmatrix}$
Cubic Isotropic	$\begin{bmatrix} xx & 0 & 0 \\ 0 & xx & 0 \\ 0 & 0 & xx \end{bmatrix}$

other crystal systems are anisotropic in their linear optical properties (in the sense that the polarization **P** need not be parallel to the applied electric field **E**) and consequently display the property of birefringence. Tetragonal, trigonal, and hexagonal crystals are said to be uniaxial crystals because there is one particular direction (the z-axis) for which the linear optical properties display rotational symmetry. Crystals of the triclinic, monoclinic, and orthorhombic systems are said to be biaxial.

Influence of Inversion Symmetry on the Second-Order Nonlinear Response

One of the symmetry properties that some but not all crystals possess is inversion symmetry. For a material system that is centrosymmetric (i.e., possesses a center of inversion) the $\chi^{(2)}$ nonlinear susceptibility must vanish identically. Since 11 of the 32 crystal classes possess inversion symmetry, this rule is very powerful, as it immediately eliminates all crystals belonging to these classes from consideration for second-order nonlinear optical interactions.

Although the result that $\chi^{(2)}$ vanishes for a centrosymmetric medium is general in nature, we shall demonstrate this fact only for the special case of second-harmonic generation in a medium that responds instantaneously to the applied optical field. We assume that the nonlinear polarization is given by

$$\tilde{P}(t) = \chi^{(2)}\tilde{E}^2(t), \tag{1.5.31}$$

where the applied field is given by

$$\tilde{E}(t) = \mathscr{E}\cos\omega t. \tag{1.5.32}$$

If we now change the sign of the applied electric field $\tilde{E}(t)$, the sign of the induced polarization $\tilde{P}(t)$ must also change, because we have assumed that the medium possesses inversion symmetry. Hence the relation (1.5.31) must be replaced by

$$-\tilde{P}(t) = \chi^{(2)}[-\tilde{E}(t)]^2, \tag{1.5.33}$$

which shows that

$$-\tilde{P}(t) = \chi^{(2)}\tilde{E}^2(t). \tag{1.5.34}$$

By comparison of this result with Eq. (1.5.31), we see that $\tilde{P}(t)$ must equal $-\tilde{P}(t)$, which can occur only if $\tilde{P}(t)$ vanishes identically. This result shows that

$$\chi^{(2)} = 0. \tag{1.5.35}$$

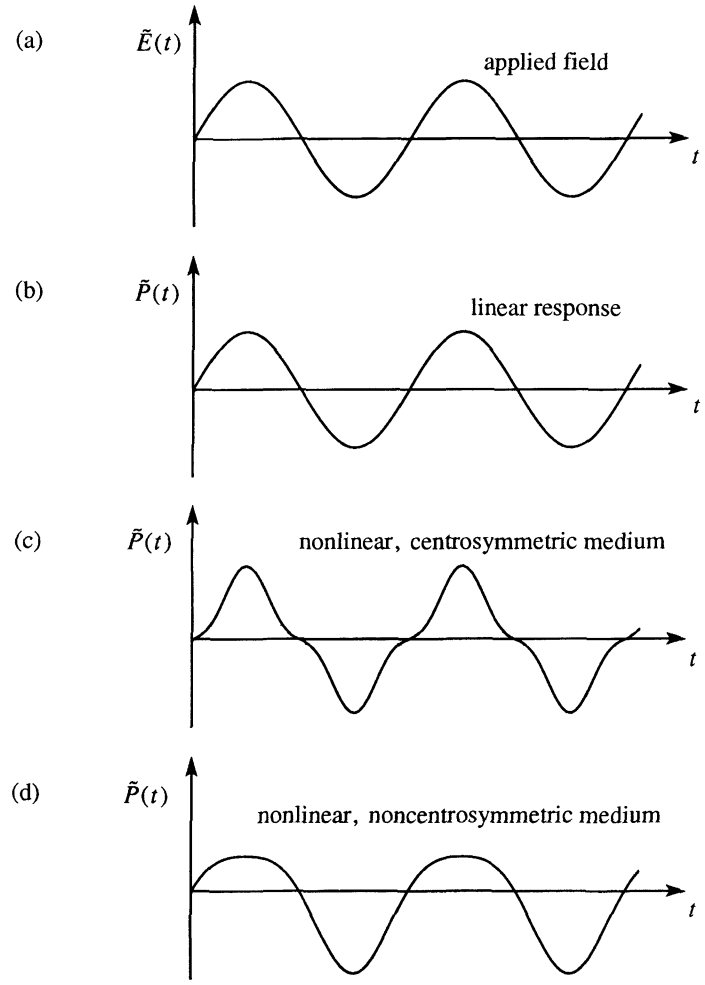

FIGURE 1.5.1 Waveforms associated with the atomic response.

This result can be understood intuitively by considering the motion of an electron in a nonparabolic potential well. Because of the nonlinearity of the associated restoring force, the atomic response will show significant harmonic distortion. Part (a) of Fig. 1.5.1 shows the waveform of the incident monochromatic electromagnetic wave of frequency ω. For the case of a medium with linear response (part b), there is no distortion of the waveform associated with the polarization of the medium. Part (c) shows the induced polarization for the case of a nonlinear medium that possesses a center of symmetry and whose potential energy function has the form shown in Fig. 1.4.2. Although significant

waveform distortion is evident, only odd harmonics of the fundamental frequency are present. For the case (part d) of a nonlinear, noncentrosymmetric medium having a potential energy function of the form shown in Fig. 1.4.1, both even and odd harmonics are present in the waveform associated with the atomic response. Note also the qualitative difference between the waveforms shown in parts (c) and (d). For the centrosymmetric medium (part c), the time-averaged response is zero, whereas for the noncentrosymmetric medium (part d) the time-average response is nonzero, because the medium responds differently to an electric field pointing, say, in the upward direction than to one pointing downward.*

Influence of Spatial Symmetry on the Second-Order Susceptibility

We have just seen how inversion symmetry when present requires that the second-order vanish identically. Any additional symmetry property of a nonlinear optical medium can impose additional restrictions on the form of the nonlinear susceptibility tensor. By explicit consideration of the symmetries of each of the 32 crystal classes, one can determine the allowed form of the susceptibility tensor for crystals of that class. The results of such a calculation for the second-order nonlinear optical response, which was performed originally by Butcher (1965), are presented in Table 1.5.2. Under those conditions (described following Eq. (1.5.21)) where the second-order susceptibility can be described using contracted notation, the results presented in Table 1.5.2 can usefully be displayed graphically. These results, as adapted from Zernike and Midwinter (1973), are presented in Fig. 1.5.2. Note that the influence of Kleinman symmetry is also described in the figure. As an example of how to use the table, the diagram for a crystal of class $3m$ is meant to imply that the form of the d_{il} matrix is

$$d_{il} = \begin{bmatrix} 0 & 0 & 0 & 0 & d_{31} & -d_{22} \\ -d_{22} & d_{22} & 0 & d_{31} & 0 & 0 \\ d_{31} & d_{31} & d_{33} & 0 & 0 & 0 \end{bmatrix}$$

The second-order nonlinear optical susceptibilities of a number of crystals are summarized in Table 1.5.3.

* Parts (a) and (b) of Fig. 1.5.1 are plots of the function $\sin \omega t$, part (c) is a plot of the function $\sin \omega t - 0.25 \sin 3\omega t$, and part (d) is a plot of $-0.2 + \sin \omega t + 0.2 \cos 2\omega t$.

TABLE 1.5.2 Form of the second-order susceptibility tensor for each of the 32 crystal classes. Each element is denoted by its cartesian indices

Crystal system	Crystal class	Nonvanishing tensor elements
Triclinic	$1 = C_1$	All elements are independent and nonzero
	$\bar{1} = S_2$	Each element vanishes
Monoclinic	$2 = C_2$	$xyz, xzy, xxy, xyx, yxx, yyy, yzz, yzx, yxz, zyz,$ zzy, zxy, zyx (twofold axis parallel to $\hat{y}$)
	$m = C_{1h}$	$xxx, xyy, xzz, xzx, xxz, yyz, yzy, yxy, yyx, zxx,$ zyy, zzz, zzx, zxz (mirror plane perpendicular to $\hat{y}$)
	$2/m = C_{2h}$	Each element vanishes
Orthorhombic	$222 = D_2$	$xyz, xzy, yzx, yxz, zxy, zyx$
	$mm2 = C_{2v}$	$xzx, xxz, yyz, yzy, zxx, zyy, zzz$
	$mmm = D_{2h}$	Each element vanishes
Tetragonal	$4 = C_4$	$xyz = -yxz, xzy = -yzx, xzx = yzy, xxz = yyz,$ $zxx = zyy, zzz, zxy = -zyx$
	$\bar{4} = S_4$	$xyz = yxz, xzy = yzx, xzx = -yzy, xxz = -yyz,$ $zxx = -zyy, zxy = zyx$
	$422 = D_4$	$xyz = -yxz, xzy = -yzx, zxy = -zyx$
	$4mm = C_{4v}$	$xzx = yzy, xxz = yyz, zxx = zyy, zzz$
	$\bar{4}2m = D_{2d}$	$xyz = yxz, xzy = yzx, zxy = zyx$
	$4/m = C_{4h}$	Each element vanishes
	$4/mmm = D_{4h}$	Each element vanishes
Cubic	$432 = O$	$xyz = -xzy = yzx = -yxz = zxy = -zyx$
	$\bar{4}3m = T_d$	$xyz = xzy = yzx = yxz = zxy = zyx$
	$23 = T$	$xyz = yzx = zxy, xzy = yxz = zyx$
	$m3 = T_h, m3m = O_h$	Each element vanishes
Trigonal	$3 = C_3$	$xxx = -xyy = -yyz = -yxy, xyz = -yxz, xzy = -yzx,$ $xzx = yzy, xxz = yyz, yyy = -yxx = -xxy = -xyx,$ $zxx = zyy, zzz, zxy = -zyx$
	$32 = D_3$	$xxx = -xyy = -yyx = -yxy, xyz = -yxz,$ $xzy = -yzx, zxy = -zyx$
	$3m = C_{3v}$	$xzx = yzy, xxz = yyz, zxx = zyy, zzz, yyy = -yxx =$ $-xxy = -xyx$ (mirror plane perpendicular to $\hat{x}$)
	$\bar{3} = S_6, \bar{3}m = D_{3d}$	Each element vanishes
Hexagonal	$6 = C_6$	$xyz = -yxz, xzy = -yzx, xzx = yzy, xxz = yyz,$ $zxx = zyy, zzz, zxy = -zyx$
	$\bar{6} = C_{3h}$	$xxx = -xyy = -yxy = -yyx,$ $yyy = -yxx = -xyx = -xxy$
	$622 = D_6$	$xyz = -yxz, xzy = -yxz, zxy = -zyx$
	$6mm = C_{6v}$	$xzx = yzy, xxz = yyz, zxx = zyy, zzz$
	$\bar{6}m2 = D_{3h}$	$yyy = -yxx = -xxy = -xyx$
	$6/m = C_{6h},$	Each element vanishes
	$6/mmm = D_{6h}$	Each element vanishes

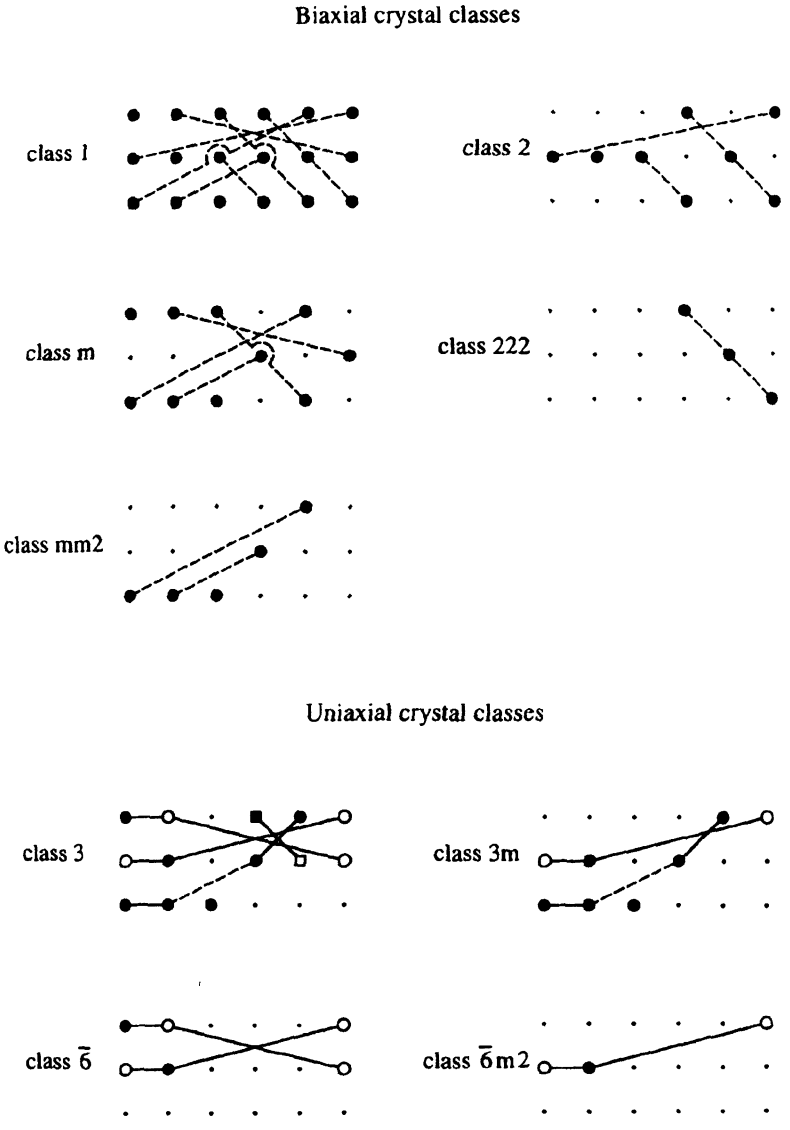

FIGURE 1.5.2 Form of the d_{il} matrix for the 21 crystal classes that lack inversion symmetry. Small dot: zero coefficient; large dot: nonzero coefficient; square: coefficient that is zero when Kleinman's symmetry condition is valid; connected symbols: numerically equal coefficients, but the open-symbol coefficient is opposite in sign to the closed symbol to which it is joined. Dashed connections are valid only under Kleinman's symmetry conditions. (After Zernike and Midwinter, 1973.)

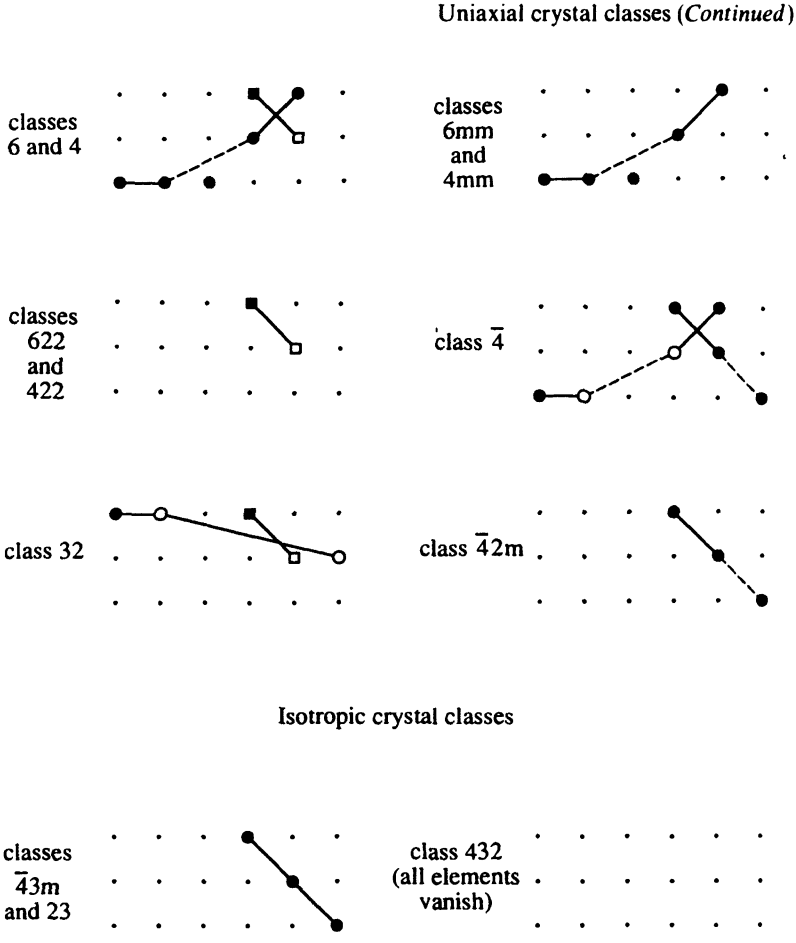

FIGURE 1.5.2 (*Continued*)

Number of Independent Elements of $\chi^{(2)}_{ijk}(\omega_3, \omega_2, \omega_1)$

We remarked above in relation to Eq. (1.5.1) that as many as 324 complex numbers must be specified in order to describe the general interaction of three optical waves. In practice, this number is often greatly reduced.

Because of the reality of the physical fields, only half of these numbers are independent (see Eq. (1.5.5)). Furthermore, the intrinsic permutation symmetry of $\chi^{(2)}$ (Eq. (1.5.6)) shows that there are only 81 independent parameters. For a lossless medium, all elements of $\chi^{(2)}$ are real and the condition of full permutation symmetry is valid, implying that only 27 of these numbers are

TABLE 1.5.3 Second-order nonlinear optical susceptibilities for several crystals

Material	Point group	d_{il} (10^{-9}cm/statvolt)
Quartz	$32 = D_3$	$d_{11} = 0.96$
		$d_{14} = 0.02$
$Ba_2NaNb_5O_{15}$	$mm2 = C_{2v}$	$d_{31} = -35$
		$d_{32} = -35$
		$d_{33} = -48$
$LiNbO_3$	$3m = C_{3v}$	$d_{22} = 7.4$
		$d_{31} = 14$
		$d_{33} = -98$
$BaTiO_3$	$4mm = C_{4v}$	$d_{15} = -41$
		$d_{31} = -43$
		$d_{33} = -16$
KH_2PO_4 (KDP)	$\bar{4}2m = D_{2d}$	$d_{14} = 1.2$
		$d_{36} = 1.1$
$LiIO_3$	$6 = C_6$	$d_{35} = -13$
		$d_{36} = -10$
GaAs	$\bar{4}3m$	$d_{36} = 406$
KD_2PO_4 (KD*P)	$\bar{4}2m = D_{2d}$	$d_{36} = 1.26$
		$d_{14} = 1.26$
CdS	$6mm = C_{6v}$	$d_{33} = 86$
		$d_{31} = 90$
		$d_{36} = 100$
Ag_3AsS_3 (proustite)	$3m = C_{3v}$	$d_{22} = 68$
		$d_{31} = 36$
$CdGeAs_2$	$\bar{4}2m = D_{2d}$	$d_{36} = 1090$
$AgGaSe_2$	$\bar{4}2m = D_{2d}$	$d_{36} = -81$
$AgSbS_3$ (pyrargyrite)	$3m = C_{3v}$	$d_{31} = 30$
		$d_{22} = 32$
beta-BaB_2O_4 (beta barium borate)		$d_{11} = 4.6$

Notes: Values are obtained from a variety of sources. Some of the more complete tabulations are those of S. Singh in *Handbook of Lasers*, Chemical Rubber Company, Cleveland, Ohio 1971, that of A. V. Smith, available at http://www.sandia.gov/imrl/XWEB1128/xxtal.htm, and the data sheets of Cleveland Crystals, Inc, available at http://www.clevelandcrystals.com.

To convert to the MKS system using the convention that $P = dE^2$, multiply each entry by $4\pi\epsilon_0/(3 \times 10^4) = 3.71 \times 10^{-15}$ to obtain d in units of C/V^2.

To convert to the MKS system using the convention that $P = \epsilon_0 dE^2$, multiply each entry by $4\pi(3 \times 10^4) = 4.189 \times 10^{-4}$ to obtain d in units of m/V.

In any system of units, $\chi^{(2)} = 2d$ by convention.

Chapter 2

Wave-Equation Description of Nonlinear Optical Interactions

2.1. The Wave Equation for Nonlinear Optical Media

We have seen in the last chapter how nonlinearity in the response of a material system to an intense laser field can cause the polarization of the medium to develop new frequency components not present in the incident radiation field. These new frequency components of the polarization act as sources of new frequency components of the electromagnetic field. In the present chapter, we examine how Maxwell's equations describe the generation of these new components of the field, and more generally we see how the various frequency components of the field become coupled by the nonlinear interaction.

Before developing the mathematical theory of these effects, we shall give a simple physical picture of how these frequency components are generated. For definiteness, we consider the case of sum-frequency generation as shown in part (a) of Fig. 2.1.1, where the input fields are at frequency ω_1 and ω_2. Because of nonlinearities in the atomic response, each atom develops an oscillating dipole moment which contains a component at frequency $\omega_1 + \omega_2$. An isolated atom would radiate at this frequency in the form of a dipole radiation pattern, as shown symbolically in part (b) of the figure. However, any material sample contains an enormous number N of atomic dipoles, each oscillating with a phase that is determined by the phases of the incident fields. If the relative phasing of these dipoles is correct, the field radiated by each dipole will add constructively in the forward direction, leading to radiation in the form of a well-defined beam, as illustrated in part (c) of the figure. The system

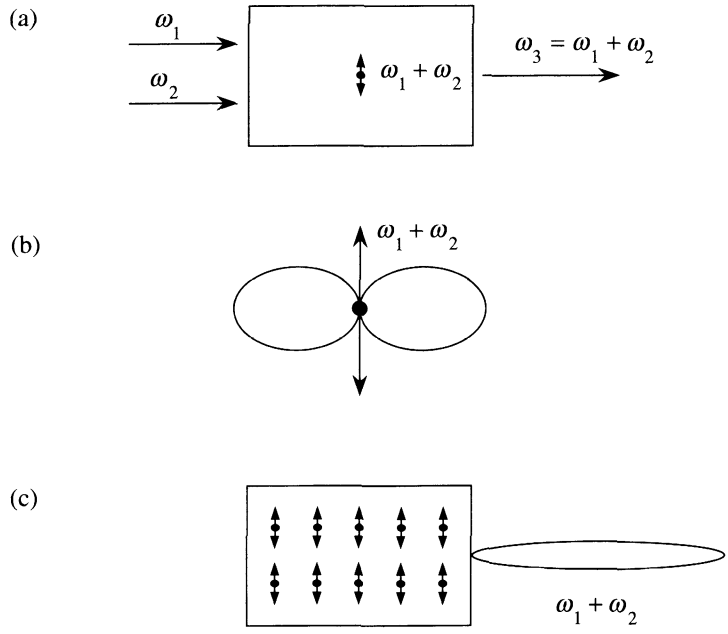

FIGURE 2.1.1 Sum-frequency generation.

will act as a phased array of dipoles when a certain condition, known as the phase-matching condition (see Eq. (2.2.15) in the next section), is satisfied. Under these conditions, the electric field strength of the radiation emitted in the forward direction will be N times larger than that due to any one atom, and consequently the intensity will be N^2 times as large.

Let us now consider the form of the wave equation for the propagation of light through a nonlinear optical medium. We begin with Maxwell's equations, which we write in gaussian units in the form*

$$\nabla \cdot \tilde{\mathbf{D}} = 4\pi\tilde{\rho}, \tag{2.1.1}$$

$$\nabla \cdot \tilde{\mathbf{B}} = 0, \tag{2.1.2}$$

$$\nabla \times \tilde{\mathbf{E}} = -\frac{1}{c}\frac{\partial \tilde{\mathbf{B}}}{\partial t}, \tag{2.1.3}$$

$$\nabla \times \tilde{\mathbf{H}} = \frac{1}{c}\frac{\partial \tilde{\mathbf{D}}}{\partial t} + \frac{4\pi}{c}\tilde{\mathbf{J}}. \tag{2.1.4}$$

* Throughout the text we use a tilde to denote a quantity that varies rapidly in time.

We are primarily interested in the solution of these equations in regions of space that contain no free charges, so that

$$\tilde{\rho} = 0, \tag{2.1.5}$$

and that contain no free currents, so that

$$\tilde{\mathbf{J}} = 0. \tag{2.1.6}$$

We assume that the material is nonmagnetic, so that

$$\tilde{\mathbf{B}} = \tilde{\mathbf{H}}. \tag{2.1.7}$$

However, we allow the material to be nonlinear in the sense that the fields $\tilde{\mathbf{D}}$ and $\tilde{\mathbf{E}}$ are related by

$$\tilde{\mathbf{D}} = \tilde{\mathbf{E}} + 4\pi\tilde{\mathbf{P}}, \tag{2.1.8}$$

where in general the polarization vector $\tilde{\mathbf{P}}$ depends nonlinearly upon the local value of the electric field strength $\tilde{\mathbf{E}}$.

We now proceed to derive the optical wave equation in the usual manner. We take the curl of the curl-$\tilde{\mathbf{E}}$ Maxwell equation (2.1.3), interchange the order of space and time derivatives on the right-hand side of the resulting equation, and use Eqs. (2.1.4), (2.1.6), and (2.1.7) to replace $\nabla \times \tilde{\mathbf{B}}$ by $(1/c)(\partial\tilde{\mathbf{D}}/\partial t)$, to obtain the equation

$$\nabla \times \nabla \times \tilde{\mathbf{E}} + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\tilde{\mathbf{D}} = 0. \tag{2.1.9a}$$

We now use Eq. (2.1.8) to eliminate $\tilde{\mathbf{D}}$ from this equation, and we thereby obtain the expression

$$\nabla \times \nabla \times \tilde{\mathbf{E}} + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\tilde{\mathbf{E}} = -\frac{4\pi}{c^2}\frac{\partial^2\tilde{\mathbf{P}}}{\partial t^2}. \tag{2.1.9b}$$

This is the most general form of the wave equation in nonlinear optics. Under certain conditions it can be simplified. For example, by using an identity from vector calculus, we can write the first term on the left-hand side of Eq. (2.1.9b) as

$$\nabla \times \nabla \times \tilde{\mathbf{E}} = \nabla(\nabla \cdot \tilde{\mathbf{E}}) - \nabla^2\tilde{\mathbf{E}}. \tag{2.1.10}$$

In the linear optics of isotropic source-free media, the first term on the right-hand side of this equation vanishes because the Maxwell equation $\nabla \cdot \tilde{\mathbf{D}} = 0$ implies that $\nabla \cdot \tilde{\mathbf{E}} = 0$. However, in nonlinear optics this term is generally nonvanishing even for isotropic materials, as a consequence of the more general relation (2.1.8) between $\tilde{\mathbf{D}}$ and $\tilde{\mathbf{E}}$. Fortunately, in nonlinear optics the first term on the right-hand side of Eq. (2.1.10) can usually be dropped for cases of interest.

For example, if $\tilde{\mathbf{E}}$ is of the form of a transverse, infinite plane wave, $\nabla \cdot \tilde{\mathbf{E}}$ vanishes identically. More generally, the first term can often be shown to be small, even when it does not vanish identically, especially when the slowly-varying amplitude approximation (see Section 2.2) is valid. For the remainder of this book, we shall usually assume that the contribution of $\nabla(\nabla \cdot \tilde{\mathbf{E}})$ in Eq. (2.1.10) is negligible so that the wave equation can be taken to have the form

$$-\nabla^2 \tilde{\mathbf{E}} + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\tilde{\mathbf{E}} = -\frac{4\pi}{c^2}\frac{\partial^2 \tilde{\mathbf{P}}}{\partial t^2}. \tag{2.1.11}$$

It is often convenient to split $\tilde{\mathbf{P}}$ into its linear and nonlinear parts as

$$\tilde{\mathbf{P}} = \tilde{\mathbf{P}}^{(1)} + \tilde{\mathbf{P}}^{\mathrm{NL}}. \tag{2.1.12}$$

Here $\tilde{\mathbf{P}}^{(1)}$ is the part of $\tilde{\mathbf{P}}$ that depends linearly on the electric field strength $\tilde{\mathbf{E}}$. We can similarly decompose the displacement field $\tilde{\mathbf{D}}$ into its linear and nonlinear parts as

$$\tilde{\mathbf{D}} = \tilde{\mathbf{D}}^{(1)} + 4\pi \tilde{\mathbf{P}}^{\mathrm{NL}}, \tag{2.1.13a}$$

where the linear part is given by

$$\tilde{\mathbf{D}}^{(1)} = \tilde{\mathbf{E}} + 4\pi \tilde{\mathbf{P}}^{(1)}. \tag{2.1.13b}$$

In terms of this quantity, the wave equation (2.1.11) becomes

$$-\nabla^2 \tilde{\mathbf{E}} + \frac{1}{c^2}\frac{\partial^2 \tilde{\mathbf{D}}^{(1)}}{\partial t^2} = -\frac{4\pi}{c^2}\frac{\partial^2 \tilde{\mathbf{P}}^{\mathrm{NL}}}{\partial t^2}. \tag{2.1.14}$$

To see why this form of the wave equation is useful, let us first consider the case of a lossless, dispersionless medium. We can then express the relation between $\tilde{\mathbf{D}}^{(1)}$ and $\tilde{\mathbf{E}}$ in terms of a real, frequency-independent dielectric tensor $\boldsymbol{\epsilon}^{(1)}$ as

$$\tilde{\mathbf{D}}^{(1)} = \boldsymbol{\epsilon}^{(1)} \cdot \tilde{\mathbf{E}}. \tag{2.1.15a}$$

For the case of an isotropic material, this relation reduces to simply

$$\tilde{\mathbf{D}}^{(1)} = \epsilon^{(1)} \tilde{\mathbf{E}}, \tag{2.1.15b}$$

where $\epsilon^{(1)}$ is a scalar quantity. For this (simpler) case of an isotropic, dispersionless material, the wave equation (2.1.14) becomes

$$-\nabla^2 \tilde{\mathbf{E}} + \frac{\epsilon^{(1)}}{c^2}\frac{\partial^2 \tilde{\mathbf{E}}}{\partial t^2} = -\frac{4\pi}{c^2}\frac{\partial^2 \tilde{\mathbf{P}}^{\mathrm{NL}}}{\partial t^2}. \tag{2.1.16}$$

This equation has the form of a driven (i.e., inhomogeneous) wave equation; the nonlinear response of the medium acts as a source term which appears on the right-hand side of this equation. In the absence of this source term,

Eq. (2.1.16) admits solutions of the form of free waves propagating with velocity c/n, where $n = [\epsilon^{(1)}]^{1/2}$ is the (linear) index of refraction.

For the case of a dispersive medium, we must consider each frequency component of the field separately. We represent the electric, linear displacement, and polarization fields as the sums of their various frequency components:

$$\tilde{\mathbf{E}}(\mathbf{r}, t) = \sum_n{}' \tilde{\mathbf{E}}_n(\mathbf{r}, t), \tag{2.1.17a}$$

$$\tilde{\mathbf{D}}^{(1)}(\mathbf{r}, t) = \sum_n{}' \tilde{\mathbf{D}}_n^{(1)}(\mathbf{r}, t), \tag{2.1.17b}$$

$$\tilde{\mathbf{P}}^{\mathrm{NL}}(\mathbf{r}, t) = \sum_n{}' \tilde{\mathbf{P}}_n^{\mathrm{NL}}(\mathbf{r}, t), \tag{2.1.17c}$$

where the summation is to be performed over positive field frequencies only, and we represent each frequency component in terms of its complex amplitude as

$$\tilde{\mathbf{E}}_n(\mathbf{r}, t) = \mathbf{E}_n(\mathbf{r}) e^{-i\omega_n t} + \text{c.c.}, \tag{2.1.18a}$$

$$\tilde{\mathbf{D}}_n^{(1)}(\mathbf{r}, t) = \mathbf{D}_n^{(1)}(\mathbf{r}) e^{-i\omega_n t} + \text{c.c.}, \tag{2.1.18b}$$

$$\tilde{\mathbf{P}}_n^{\mathrm{NL}}(\mathbf{r}, t) = \mathbf{P}_n^{\mathrm{NL}}(\mathbf{r}) e^{-i\omega_n t} + \text{c.c.} \tag{2.1.18c}$$

If dissipation can be neglected, the relationship between $\tilde{\mathbf{D}}_n^{(1)}$ and $\tilde{\mathbf{E}}_n$ can be expressed in terms of a real, frequency-dependent dielectric tensor according to

$$\tilde{\mathbf{D}}_n^{(1)}(\mathbf{r}, t) = \epsilon^{(1)}(\omega_n) \cdot \tilde{\mathbf{E}}_n(\mathbf{r}, t). \tag{2.1.19}$$

When Eqs. (2.1.17a) through (2.1.19) are introduced into Eq. (2.1.14), we obtain a wave equation analogous to (2.1.16) that is valid for each frequency component of the field:

$$-\nabla^2 \tilde{\mathbf{E}}_n + \frac{\epsilon^{(1)}(\omega_n)}{c^2} \frac{\partial^2 \tilde{\mathbf{E}}_n}{\partial t^2} = -\frac{4\pi}{c^2} \frac{\partial^2 \tilde{\mathbf{P}}_n^{\mathrm{NL}}}{\partial t^2}. \tag{2.1.20}$$

The general case of a dissipative medium is treated by allowing the dielectric tensor to be a complex quantity that relates the complex field amplitudes according to

$$\mathbf{D}_n^{(1)}(\mathbf{r}) = \epsilon^{(1)}(\omega_n) \cdot \mathbf{E}_n(\mathbf{r}). \tag{2.1.21}$$

This expression, along with Eqs. (2.1.17) and (2.1.18), can be introduced into the wave equation (2.1.14), to obtain

$$-\nabla^2 \mathbf{E}_n(\mathbf{r}) - \frac{\omega_n^2}{c^2} \epsilon^{(1)}(\omega_n) \cdot \mathbf{E}_n(\mathbf{r}) = \frac{4\pi\omega_n^2}{c^2} \mathbf{P}_n^{\mathrm{NL}}(\mathbf{r}). \tag{2.1.22}$$

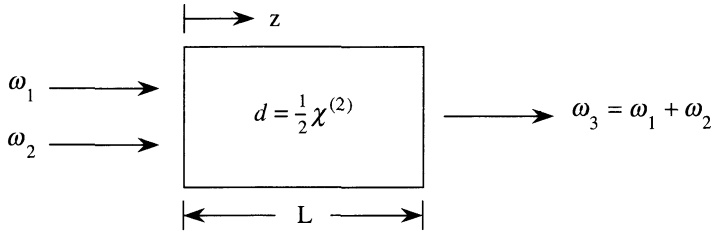

FIGURE 2.2.1 Sum-frequency generation.

2.2. The Coupled-Wave Equations for Sum-Frequency Generation

We next study how the nonlinear optical wave equation that we derived in the previous section can be used to describe specific nonlinear optical interactions. In particular, we consider sum-frequency generation in a lossless nonlinear optical medium involving collimated, monochromatic, continuous-wave input beams. We assume the configuration shown in Fig. 2.2.1, where the applied waves fall onto the nonlinear medium at normal incidence. For simplicity, we ignore double refraction effects. The treatment given here can be generalized straightforwardly to include nonnormal incidence and double refraction.*

The wave equation in the form (2.1.20) must hold for each frequency component of the field and in particular for the sum-frequency component at frequency ω_3. In the absence of a nonlinear source term, the solution to this equation for a plane wave at frequency ω_3 propagating in the $+z$ direction is

$$\tilde{E}_3(z, t) = A_3 e^{i(k_3 z - \omega_3 t)} + \text{c.c.}, \tag{2.2.1}$$

where†

$$k_3 = \frac{n_3 \omega_3}{c}, \qquad n_3 = \left[\epsilon^{(1)}(\omega_3)\right]^{1/2}, \tag{2.2.2}$$

and where the amplitude of the wave A_3 is a *constant*. We expect on physical grounds that, when the nonlinear source term is not too large, the solution to Eq. (2.1.20) will still be of the form of Eq. (2.2.1), except that A_3 will become a slowly varying function of z. We hence adopt Eq. (2.2.1) with A_3 taken to be a function of z as the form of the trial solution to the wave equation (2.1.20) in the presence of the nonlinear source term.

We represent the nonlinear source term appearing in Eq. (2.1.20) as

$$\tilde{P}_3(z, t) = P_3 e^{-i\omega_3 t} + \text{c.c.}, \tag{2.2.3}$$

* See, for example, Shen (1984), Chapter 6.

† For convenience, we are working in the scalar field approximation; n_3 represents the refractive index appropriate to the state of polarization of the ω_3 wave.

where according to Eq. (1.5.28)

$$P_3 = 4d_{\text{eff}}E_1E_2. \tag{2.2.4}$$

If we represent the applied fields as

$$\tilde{E}_i(z,t) = E_i e^{-i\omega_i t} + \text{c.c.}, \qquad i = 1, 2, \tag{2.2.5}$$

with

$$E_i = A_i e^{ik_i z}, \qquad i = 1, 2, \tag{2.2.6}$$

the amplitude of the nonlinear polarization can be written as

$$P_3 = 4d_{\text{eff}}A_1A_2e^{i(k_1+k_2)z} \equiv p_3 e^{i(k_1+k_2)z}. \tag{2.2.7}$$

We now substitute Eqs. (2.2.1), (2.2.3), and (2.2.7) into the wave equation (2.1.20). Since the fields depend only on the longitudinal coordinate z, we can replace ∇^2 by d^2/dz^2. We then obtain

$$\begin{aligned}\left[\frac{d^2A_3}{dz^2} + 2ik_3\frac{dA_3}{dz} - k_3^2A_3 + \frac{\epsilon^{(1)}(\omega_3)\omega_3^2A_3}{c^2}\right]e^{i(k_3z-\omega_3t)} + \text{c.c.}\\ = \frac{-16\pi d_{\text{eff}}\omega_3^2}{c^2}A_1A_2e^{i[(k_1+k_2)z-\omega_3t]} + \text{c.c.}\end{aligned} \tag{2.2.8}$$

Since $k_3^2 = \epsilon^{(1)}(\omega_3)\omega_3^2/c^2$, the third and fourth terms on the left-hand side of this expression cancel. Note that we can drop the complex conjugate terms from each side and still maintain the equality. We can then cancel the factor $\exp(-i\omega_3 t)$ on each side and write the resulting equation as

$$\frac{d^2A_3}{dz^2} + 2ik_3\frac{dA_3}{dz} = \frac{-16\pi d_{\text{eff}}\omega_3^2}{c^2}A_1A_2e^{i(k_1+k_2-k_3)z}. \tag{2.2.9}$$

It is usually permissible to neglect the first term on the left-hand side of this equation on the grounds that it is very much smaller than the second. This approximation is known as the slowly-varying amplitude approximation and is valid whenever

$$\left|\frac{d^2A_3}{dz^2}\right| \ll \left|k_3\frac{dA_3}{dz}\right|. \tag{2.2.10}$$

This condition requires that the fractional change in A_3 in a distance of the order of an optical wavelength must be much smaller than unity. When this approximation is made, Eq. (2.2.9) becomes

$$\begin{aligned}\frac{dA_3}{dz} &= \frac{8\pi i d_{\text{eff}}\omega_3^2}{k_3c^2}A_1A_2e^{i\Delta kz}\\ &= \frac{2\pi i\omega_3}{n_3c}p_3e^{i\Delta kz},\end{aligned} \tag{2.2.11}$$

where we have introduced the quantity

$$\Delta k = k_1 + k_2 - k_3, \tag{2.2.12}$$

which is called the wavevector (or momentum) mismatch. Equation (2.2.11) is known as a coupled-amplitude equation, because it shows how the amplitude of the ω_3 wave varies as a consequence of its coupling to the ω_1 and ω_2 waves. In general, the spatial variation of the ω_1 and ω_2 waves must also be taken into consideration, and we can derive analogous equations for the ω_1 and ω_2 fields by repeating the derivation given above for each of these frequencies. We hence find two additional coupled-amplitude equations given by

$$\frac{dA_1}{dz} = \frac{8\pi i d_{\text{eff}} \omega_1^2}{k_1 c^2} A_3 A_2^* e^{-i\Delta kz} \tag{2.2.13}$$

and

$$\frac{dA_2}{dz} = \frac{8\pi i d_{\text{eff}} \omega_2^2}{k_2 c^2} A_3 A_1^* e^{-i\Delta kz}. \tag{2.2.14}$$

Note that, in writing these equations in the forms shown, we have assumed that the medium is lossless. For a lossless medium, no explicit loss terms need be included in these equations, and furthermore we can make use of the condition of full permutation symmetry (Eq. (1.5.8)) to conclude that the coupling coefficient has the same value d_{eff} in each equation.

Phase-Matching Considerations

For simplicity, let us first assume that the amplitudes A_1 and A_2 of the input fields can be taken as constants on the right-hand side of Eq. (2.2.11). This assumption is valid whenever the conversion of the input fields into the sum-frequency field is not too large. We note that, for the special case

$$\Delta k = 0, \tag{2.2.15}$$

the amplitude A_3 of the sum-frequency wave increases linearly with z, and consequently that its intensity increases quadratically with z. The condition (2.2.15) is known as the condition of perfect phase matching. When this condition is fulfilled, the generated wave maintains a fixed phase relation with respect to the nonlinear polarization and is able to extract energy most efficiently from the incident waves. From a microscopic point of view, when the condition (2.2.15) is fulfilled the individual atomic dipoles that constitute the material system are properly phased so that the field emitted by each dipole adds coherently in the forward direction. The total power radiated by the ensemble of atomic dipoles thus scales as the square of the number of atoms that participate.

When the condition (2.2.15) is not satisfied, the intensity of the emitted radiation is smaller than for the case of $\Delta k = 0$. The amplitude of the sum-frequency (ω_3) field at the exit plane of the nonlinear medium is given in this case by integrating Eq. (2.2.11) from $z = 0$ to $z = L$, yielding

$$A_3(L) = \frac{8\pi i d_{\text{eff}} \omega_3^2 A_1 A_2}{k_3 c^2} \int_0^L e^{i\Delta k z} dz = \frac{8\pi i d \omega_3^2 A_1 A_2}{k_3 c^2} \left(\frac{e^{i\Delta k L} - 1}{i\Delta k} \right). \tag{2.2.16}$$

The intensity of the ω_3 wave is given by the magnitude of the time-averaged Poynting vector, which for our definition of field amplitude is given by

$$I_i = \frac{n_i c}{2\pi} |A_i|^2, \quad i = 1, 2, 3. \tag{2.2.17}$$

We thus obtain

$$I_3 = \frac{32\pi d_{\text{eff}}^2 \omega_3^4 |A_1|^2 |A_2|^2 n_3}{k_3^2 c^3} \left| \frac{e^{i\Delta k L} - 1}{\Delta k} \right|^2. \tag{2.2.18}$$

The squared modulus that appears in this equation can be expressed as

$$\left| \frac{e^{i\Delta k L} - 1}{\Delta k} \right|^2 = L^2 \left(\frac{e^{i\Delta k L} - 1}{\Delta k L} \right) \left(\frac{e^{-i\Delta k L} - 1}{\Delta k L} \right) = 2L^2 \frac{(1 - \cos \Delta k L)}{(\Delta k L)^2} \tag{2.2.19}$$

$$= L^2 \frac{\sin^2(\Delta k L/2)}{(\Delta k L/2)^2} \equiv L^2 \text{sinc}^2(\Delta k L/2).$$

Finally, our expression for I_3 can be written in terms of the intensities of the incident fields by using Eq. (2.2.17) to express $|A_i|^2$ in terms of the intensities, yielding the result

$$I_3 = \frac{512\pi^5 d_{\text{eff}}^2 I_1 I_2}{n_1 n_2 n_3 \lambda_3^2 c} L^2 \text{sinc}^2(\Delta k L/2), \tag{2.2.20}$$

where $\lambda_3 = 2\pi c/\omega_3$ is the vacuum wavelength of the ω_3 wave. Note that the effect of wavevector mismatch is included entirely in the factor $\text{sinc}^2(\Delta k L/2)$. This factor, which is known as the phase mismatch factor, is plotted in Fig. 2.2.2.

It should be noted that the efficiency of the three-wave mixing process decreases as $|\Delta k| L$ increases, with some oscillations occurring. The reason for this behavior is that if L is greater than approximately $1/\Delta k$, the output wave can get out of phase with its driving polarization, and power can flow from the ω_3 wave back into the ω_1 and ω_2 waves (see Eq. (2.2.11)). For this reason, one sometimes defines

$$L_c = 2/\Delta k \tag{2.2.21}$$

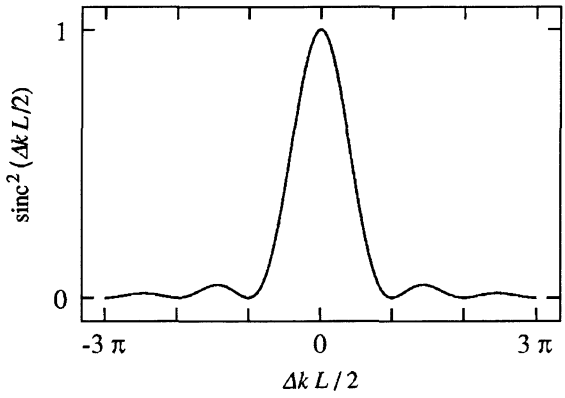

FIGURE 2.2.2 Effects of wavevector mismatch on the efficiency of sum-frequency generation.

to be the coherent buildup length of the interaction, so that the phase mismatch factor in Eq. (2.2.20) can be written as

$$\mathrm{sinc}^2(L/L_c). \tag{2.2.22}$$

2.3. The Manley–Rowe Relations

Let us now consider, from a general point of view, the mutual interaction of three optical waves propagating through a lossless nonlinear optical medium, as illustrated in Fig. 2.3.1.

We have just derived the coupled-amplitude equations (Eqs. (2.2.11) through (2.2.14)) that describes the spatial variation of the amplitude of each wave. Let us now consider the spatial variation of the *intensity* associated with each of these waves. Since

$$I_i = \frac{n_i c}{2\pi} A_i A_i^*, \tag{2.3.1}$$

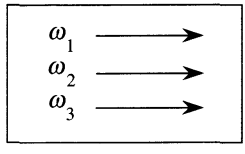

FIGURE 2.3.1 Optical waves of frequencies ω_1, ω_2, and $\omega_3 = \omega_1 + \omega_2$ interact in a lossless nonlinear optical medium.

the variation of the intensity is described by

$$\frac{dI_i}{dz} = \frac{n_i c}{2\pi}\left(A_i^* \frac{dA_i}{dz} + A_i \frac{dA_i^*}{dz}\right). \tag{2.3.2}$$

Through use of this result and Eq. (2.2.13), we find that the spatial variation of the intensity of the wave at frequency ω_1 is given by

$$\begin{aligned}\frac{dI_1}{dz} &= \frac{n_1 c}{2\pi}\frac{8\pi d_{\text{eff}}\omega_1^2}{k_1 c^2}(iA_1^* A_3 A_2^* e^{-i\Delta kz} + \text{c.c.}) \\ &= 4d_{\text{eff}}\omega_1(iA_3 A_1^* A_2^* e^{-i\Delta kz} + \text{c.c.})\end{aligned}$$

or by

$$\frac{dI_1}{dz} = -8d_{\text{eff}}\omega_1 \,\text{Im}(A_3 A_1^* A_2^* e^{-i\Delta kz}). \tag{2.3.3a}$$

We similarly find that the spatial variation of the intensities of the waves at frequencies ω_2 and ω_3 is given by

$$\frac{dI_2}{dz} = -8d_{\text{eff}}\omega_2 \,\text{Im}(A_3 A_1^* A_2^* e^{-i\Delta kz}), \tag{2.3.3b}$$

$$\begin{aligned}\frac{dI_3}{dz} &= -8d_{\text{eff}}\omega_3 \,\text{Im}(A_3^* A_1 A_2 e^{i\Delta kz}) \\ &= 8d_{\text{eff}}\omega_3 \,\text{Im}(A_3 A_1^* A_2^* e^{-i\Delta kz}).\end{aligned} \tag{2.3.3c}$$

We see that the sign of dI_1/dz is the same as that of dI_2/dz but is opposite to that of dI_3/dz. We also see that the direction of energy flow depends on the relative phases of the three interacting fields.

The set of equations (2.3.3a), (2.3.3b), and (2.3.3c) shows that the total power flow is conserved, as expected for propagation through a lossless medium. To demonstrate this fact, we define the total intensity as

$$I = I_1 + I_2 + I_3. \tag{2.3.4}$$

We then find that the spatial variation of the total intensity is given by

$$\begin{aligned}\frac{dI}{dz} &= \frac{dI_1}{dz} + \frac{dI_2}{dz} + \frac{dI_3}{dz} \\ &= -8d_{\text{eff}}(\omega_1 + \omega_2 - \omega_3)\text{Im}(A_3 A_1^* A_2^* e^{i\Delta kz}) = 0,\end{aligned} \tag{2.3.5}$$

where we have made use of Eqs. (2.3.3a), (2.3.3b), and (2.3.3c) and where the last equality follows from the fact that $\omega_3 = \omega_1 + \omega_2$.

The set of equations (2.3.3a), (2.3.3b), and (2.3.3c) also implies that

$$\frac{d}{dz}\left(\frac{I_1}{\omega_1}\right) = \frac{d}{dz}\left(\frac{I_2}{\omega_2}\right) = -\frac{d}{dz}\left(\frac{I_3}{\omega_3}\right), \tag{2.3.6}$$

as can be verified by inspection. These equalities are known as the Manley–Rowe relations (Manley and Rowe, 1959). Since the energy of a photon of frequency ω_i is $\hbar\omega_i$, the quantity I_i/ω_i that appears in these relations is proportional to the intensity of the wave measured in photons per unit area per unit time. The Manley–Rowe relations can alternatively be expressed as

$$\frac{d}{dz}\left(\frac{I_2}{\omega_2}+\frac{I_3}{\omega_3}\right)=0, \quad \frac{d}{dz}\left(\frac{I_1}{\omega_1}+\frac{I_3}{\omega_3}\right)=0, \quad \frac{d}{dz}\left(\frac{I_1}{\omega_1}-\frac{I_2}{\omega_2}\right)=0. \tag{2.3.7}$$

These equations can be formally integrated to obtain the three conserved quantities (conserved in the sense that they are spatially invariant) M_1, M_2, and M_3, which are given by

$$M_1=\frac{I_2}{\omega_2}+\frac{I_3}{\omega_3}, \quad M_2=\frac{I_1}{\omega_1}+\frac{I_3}{\omega_3}, \quad M_3=\frac{I_1}{\omega_1}-\frac{I_2}{\omega_2}. \tag{2.3.8}$$

These relations tell us that the rate at which photons at frequency ω_1 are created is equal to the rate at which photons at frequency ω_2 are created and is equal to the rate at which photons at frequency ω_3 are destroyed. This result can be understood intuitively by means of the energy level description of a three-wave mixing process, which is shown in Figure 2.3.2. This diagram shows that, for a lossless medium, the creation of an ω_1 photon must be accompanied by the creation of an ω_2 photon and the annihilation of an ω_3 photon. It seems at first sight surprising that the Manley–Rowe relations should be consistent with this quantum-mechanical interpretation, when our derivation of these relations appears to be entirely classical. Note, however, that our derivation implicitly assumes that the nonlinear susceptibility possesses full permutation symmetry in that we have taken the coupling constant d_{eff} to have the same value in each of the coupled-amplitude equations (2.2.11), (2.2.13), and (2.2.14). We remarked earlier (following Eq. (1.5.9)) that in a sense the condition of full permutation symmetry is a consequence of the laws of quantum mechanics.

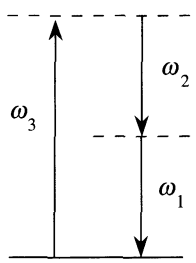

FIGURE 2.3.2 Photon description of the interaction of three optical waves.

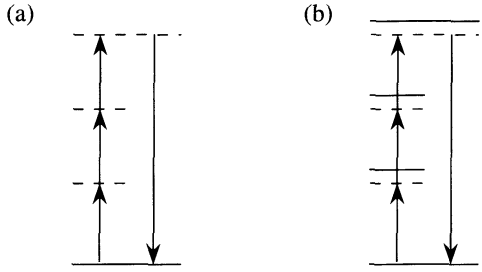

FIGURE 3.1.1 Third-harmonic generation described in terms of virtual levels (a) and with real atomic levels indicated (b).

coupling between the radiation and the atom is particularly strong and the nonlinear optical susceptibility becomes large.

Three possible strategies for enhancing the efficiency of third-harmonic generation through the technique of resonance enhancement are illustrated in Fig. 3.1.2. In part (a), the one-photon transition is nearly resonant, in part (b) the two-photon transition is nearly resonant, and in part (c) the three-photon transition is nearly resonant. The formulas derived later in this chapter demonstrate that all three procedures are equally effective at increasing the value of the third-order nonlinear susceptibility. However, the method shown in part (b) is usually the preferred way in which to generate the third-harmonic field with high efficiency, for the following reason: For the case of a one-photon resonance (part a), the incident field experiences linear absorption and is rapidly attenuated as it propagates through the medium. Similarly, for the case of the three-photon resonance (part c), the generated field experiences linear absorption. However, for the case of a two-photon resonance (part b), there is no linear absorption to limit the efficiency of the process.

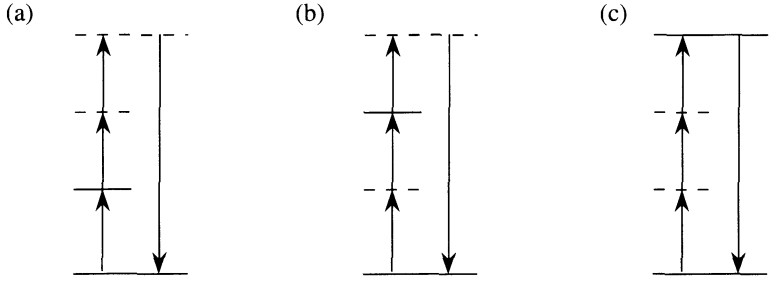

FIGURE 3.1.2 Three strategies for enhancing the process of third-harmonic generation.

3.2. Schrödinger Equation Calculation of the Nonlinear Optical Susceptibility

In this section, we present a derivation of the nonlinear optical susceptibility based on quantum-mechanical perturbation theory of the atomic wave function. The expressions that we derive using this formalism can be used to make accurate predictions of the *nonresonant* response of atomic and molecular systems. Relaxation processes, which are important for the case of near-resonant excitation, cannot be adequately described by this formalism. Relaxation processes are discussed later in this chapter in connection with the density matrix formulation of the theory of the nonlinear optical susceptibility. Even though the density matrix formalism provides results that are more generally valid, the calculation of the nonlinear susceptibility is much more complicated when performed using this method. For this reason, we first present a calculation of the nonlinear susceptibility based on the properties of the atomic wavefunction, since this method is somewhat simpler and for this reason gives a clearer picture of the underlying physics of the nonlinear interaction.

One of the fundamental assumption of quantum mechanics is that all of the properties of the atomic system can be described in terms of the atomic wavefunction $\psi(\mathbf{r}, t)$, which is the solution to the time-dependent Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = \hat{H}\psi. \tag{3.2.1}$$

Here $\hat{H}$ is the Hamiltonian operator

$$\hat{H} = \hat{H}_0 + \hat{V}(t), \tag{3.2.2}$$

which is written as the sum of the Hamiltonian $\hat{H}_0$ for a free atom and an interaction Hamiltonian, $\hat{V}(t)$, which describes the interaction of the atom with the electromagnetic field. We usually take the interaction Hamiltonian to be of the form

$$\hat{V}(t) = -\hat{\boldsymbol{\mu}} \cdot \tilde{\mathbf{E}}(t), \tag{3.2.3}$$

where $\hat{\boldsymbol{\mu}} = -e\hat{\mathbf{r}}$ is the electric dipole moment operator and $-e$ is the charge of the electron, and where we assume that $\tilde{\mathbf{E}}(t)$ can be represented as a discrete sum of (positive and negative) frequency components as

$$\tilde{\mathbf{E}}(t) = \sum_p \mathbf{E}(\omega_p) e^{-i\omega_p t}. \tag{3.2.4}$$

Energy Eigenstates

For the case in which no external field is applied to the atom, the Hamiltonian $\hat{H}$ is simply equal to $\hat{H}_0$, and Schrödinger's equation (3.2.1) possesses solutions in the form of energy eigenstates. These states are also known as stationary states, because the time of evolution of these states is given by a simple exponential phase factor. These states have the form

$$\psi_n(\mathbf{r}, t) = u_n(\mathbf{r})e^{-i\omega_n t}. \tag{3.2.5a}$$

By substituting this form into the Schrödinger equation (3.2.1), we find that the spatially varying part of the wavefunction $u_n(\mathbf{r})$ must satisfy the eigenvalue equation (known as the time-independent Schrödinger equation)

$$\hat{H}_0 u_n(\mathbf{r}) = E_n u_n(\mathbf{r}), \tag{3.2.5b}$$

where $E_n = \hbar\omega_n$. For future convenience, we assume that these solutions are chosen in such a manner that they constitute a complete, orthonormal set satisfying the condition

$$\int u_m^* u_n d^3r = \delta_{mn}. \tag{3.2.6}$$

Perturbation Solution to Schrödinger's Equation

For the general case in which the atom is exposed to an electromagnetic field, Schrödinger's equation (3.2.1) usually cannot be solved exactly. In such cases, it is often adequate to solve Schrödinger's equation through the use of perturbation theory. In order to solve Eq. (3.2.1) systematically in terms of a perturbation expansion, we replace the Hamiltonian (3.2.2) by

$$\hat{H} = \hat{H}_0 + \lambda\hat{V}(t), \tag{3.2.7}$$

where λ is a continuously varying parameter ranging from zero to unity that characterizes the strength of the interaction; the value $\lambda = 1$ corresponds to the actual physical situation. We now seek a solution to Schrödinger's equation in the form of a power series in λ:

$$\psi(\mathbf{r}, t) = \psi^{(0)}(\mathbf{r}, t) + \lambda\psi^{(1)}(\mathbf{r}, t) + \lambda^2\psi^{(2)}(\mathbf{r}, t) + \cdots. \tag{3.2.8}$$

By requiring that the solution be of this form for any value of λ, we assure that $\psi^{(N)}$ will be that part of the solution which is of order N in the interaction energy V. We now introduce Eq. (3.2.8) into Eq. (3.2.1) and require that all terms proportional to $\lambda^{(N)}$ satisfy the equality separately. We thereby obtain

the set of equations

$$i\hbar \frac{\partial \psi^{(0)}}{\partial t} = \hat{H}_0 \psi^{(0)}, \tag{3.2.9a}$$

$$i\hbar \frac{\partial \psi^{(N)}}{\partial t} = \hat{H}_0 \psi^{(N)} + \hat{V} \psi^{(N-1)}, \quad N = 1, 2, 3 \ldots. \tag{3.2.9b}$$

Equation (3.2.9a) is simply Schrödinger's equation for the atom in the absence of its interaction with the applied field; we assume for definiteness that initially the atom is in state g (typically the ground state) so that the solution to this equation is

$$\psi^{(0)}(\mathbf{r}, t) = u_g(\mathbf{r}) e^{-iE_g t/\hbar}. \tag{3.2.10}$$

The remaining equations in the perturbation expansion (Eq. (3.2.9b)) are solved by making use of the fact that the energy eigenfunctions for the free atom constitute a complete set of basis functions, in terms of which any function can be expanded. In particular, we represent the Nth-order contribution to the wavefunction $\psi^{(N)}(\mathbf{r}, t)$ as the sum

$$\psi^{(N)}(\mathbf{r}, t) = \sum_l a_l^{(N)}(t) u_l(\mathbf{r}) e^{-i\omega_l t}. \tag{3.2.11}$$

Here $a_l^{(N)}(t)$ gives the probability amplitude that, to Nth order in the perturbation, the atom is in energy eigenstate l at time t. If Eq. (3.2.11) is substituted into Eq. (3.2.9b), we find that the probability amplitudes obey the system of equations

$$i\hbar \sum_l \dot{a}_l^{(N)} u_l(\mathbf{r}) e^{-i\omega_l t} = \sum_l a_l^{(N-1)} \hat{V} u_l(\mathbf{r}) e^{-i\omega_l t}, \tag{3.2.12}$$

where the dot denotes a total time derivative. This equation relates all of the probability amplitudes of order N to all of the amplitudes of order $N-1$. To simplify this equation, we multiply each side from the left by u_m^* and we integrate the resulting equation over all space. Then through use of the orthonormality condition (3.2.6), we obtain the equation

$$\dot{a}_m^{(N)} = (i\hbar)^{-1} \sum_l a_l^{(N-1)} V_{ml} e^{i\omega_{ml} t}, \tag{3.2.13}$$

where $\omega_{ml} \equiv \omega_m - \omega_l$ and where we have introduced the matrix elements of the perturbing Hamiltonian, which are defined by

$$V_{ml} \equiv \langle u_m | \hat{V} | u_l \rangle = \int u_m^* \hat{V} u_l \, d^3 r. \tag{3.2.14}$$

The form of Eq. (3.2.13) demonstrates the usefulness of the perturbation technique; once the probability amplitudes of order $N-1$ are determined, the

amplitudes of the next higher order (N) can be obtained by straightforward time integration. In particular, we find that

$$a_m^{(N)}(t) = (i\hbar)^{-1} \sum_l \int_{-\infty}^{t} dt' V_{ml}(t') a_l^{(N-1)}(t') e^{i\omega_{ml}t'} \tag{3.2.15}$$

We shall eventually be interested in determining the linear, second-order, and third-order optical susceptibilities. To do so, we shall require explicit expressions for the probability amplitudes up to third order in the perturbation expansion. We now determine the form of these amplitudes.

To determine the first-order amplitudes $a_m^{(1)}(t)$, we set $a_l^{(0)}$ in Eq. (3.2.15) equal to δ_{lg} (corresponding to an atom known to be in state g in zeroth order) and, through use of Eqs. (3.2.3) and (3.2.4), replace $V_{ml}(t')$ by $-\sum_p \boldsymbol{\mu}_{ml} \cdot \mathbf{E}(\omega_p)\exp(-i\omega_p t')$, where $\boldsymbol{\mu}_{ml} = \int u_m^* \hat{\boldsymbol{\mu}} u_l d^3r$ is known as the electric dipole transition moment. We next evaluate the integral appearing in Eq. (3.2.15) and assume that the contribution from the lower limit of integration vanishes; we thereby find that

$$a_m^{(1)}(t) = \frac{1}{\hbar} \sum_p \frac{\boldsymbol{\mu}_{mg} \cdot \mathbf{E}(\omega_p)}{\omega_{mg} - \omega_p} e^{i(\omega_{mg} - \omega_p)t}. \tag{3.2.16}$$

We next determine the second-order correction to the probability amplitude by using Eq. (3.2.15) once again, but with N set equal to 2. We introduce Eq. (3.2.16) for $a_m^{(1)}$ into the right-hand side of this equation and perform the integration to find that

$$a_n^{(2)}(t) = \frac{1}{\hbar^2} \sum_{pq} \sum_m \frac{[\boldsymbol{\mu}_{nm} \cdot \mathbf{E}(\omega_q)][\boldsymbol{\mu}_{mg} \cdot \mathbf{E}(\omega_p)]}{(\omega_{ng} - \omega_p - \omega_q)(\omega_{mg} - \omega_p)} e^{i(\omega_{ng} - \omega_p - \omega_q)t}. \tag{3.2.17}$$

Analogously, through an additional use of Eq. (3.2.15), we find that the third-order correction to the probability amplitude is given by

$$a_v^{(3)}(t) = \frac{1}{\hbar^3} \sum_{pqr} \sum_{mn} \frac{[\boldsymbol{\mu}_{vn} \cdot \mathbf{E}(\omega_r)][\boldsymbol{\mu}_{nm} \cdot \mathbf{E}(\omega_q)][\boldsymbol{\mu}_{mg} \cdot \mathbf{E}(\omega_p)]}{(\omega_{vg} - \omega_p - \omega_q - \omega_r)(\omega_{ng} - \omega_p - \omega_q)(\omega_{mg} - \omega_p)}$$
$$\times e^{i(\omega_{vg} - \omega_p - \omega_q - \omega_r)t}. \tag{3.2.18}$$

Linear Susceptibility

Let us now use the results just obtained to determine the linear optical properties of a material system. The expectation value of the electric dipole moment is given by

$$\langle \tilde{\mathbf{p}} \rangle = \langle \psi | \hat{\boldsymbol{\mu}} | \psi \rangle, \tag{3.2.19}$$

We find that the lowest-order contribution to $\langle\tilde{\mathbf{p}}\rangle$ (i.e., the contribution linear in the applied field amplitude) is given by

$$\langle\tilde{\mathbf{p}}^{(1)}\rangle = \langle\psi^{(0)}|\hat{\boldsymbol{\mu}}|\psi^{(1)}\rangle + \langle\psi^{(1)}|\hat{\boldsymbol{\mu}}|\psi^{(0)}\rangle, \tag{3.2.20}$$

where $\psi^{(0)}$ is given by Eq. (3.2.10) and $\psi^{(1)}$ is given by Eqs. (3.2.11) and (3.2.16). By substituting these forms into Eq. (3.2.20) we find that

$$\langle\tilde{\mathbf{p}}^{(1)}\rangle = \frac{1}{\hbar}\sum_p\sum_m\left(\frac{\boldsymbol{\mu}_{gm}[\boldsymbol{\mu}_{mg}\cdot\mathbf{E}(\omega_p)]}{\omega_{mg}-\omega_p}e^{-i\omega_p t} + \frac{[\boldsymbol{\mu}_{mg}\cdot\mathbf{E}(\omega_p)]^*\boldsymbol{\mu}_{mg}}{\omega_{mg}^*-\omega_p}e^{i\omega_p t}\right). \tag{3.2.21}$$

In writing Eq. (3.2.21) in the form shown, we have formally allowed the possibility that the transition frequency ω_{mg} is a complex quantity. We have done this because a crude way of incorporating damping phenomena into the theory is to take ω_{mg} to be the complex quantity $\omega_{mg} = \omega_{mg}^0 - i\Gamma_m/2$, where ω_{mg}^0 is the (real) transition frequency and Γ_m is the population decay rate of the upper level m. This procedure is not totally acceptable, because it cannot describe the cascade of population among the excited states nor can it describe dephasing processes that are not accompanied by the transfer of population. Nonetheless, for the remainder of the present section, we shall allow the transition frequency to be a complex quantity in order to provide an indication of how damping effects could be incorporated into the present theory.

Equation (3.2.21) is written as a summation over all positive and negative field frequencies ω_p. This result is easier to interpret if we formally replace ω_p by $-\omega_p$ in the second term, in which case the expression becomes

$$\langle\tilde{\mathbf{p}}^{(1)}\rangle = \frac{1}{\hbar}\sum_p\sum_m\left(\frac{\boldsymbol{\mu}_{gm}[\boldsymbol{\mu}_{mg}\cdot\mathbf{E}(\omega_p)]}{\omega_{mg}-\omega_p} + \frac{[\boldsymbol{\mu}_{gm}\cdot\mathbf{E}(\omega_p)]\boldsymbol{\mu}_{mg}}{\omega_{mg}^*+\omega_p}\right)e^{-i\omega_p t}. \tag{3.2.22}$$

We next use this result to calculate the form of the linear susceptibilty. We take the linear polarization to be $\tilde{\mathbf{P}}^{(1)} = N\langle\tilde{\mathbf{p}}^{(1)}\rangle$, where N is the number density of atoms. We next express the polarization in terms of its complex amplitude as $\tilde{\mathbf{P}}^{(1)} = \sum_p \mathbf{P}^{(1)}(\omega_p)\exp(-i\omega_p t)$. Finally, we introduce the linear susceptibility defined through the relation $P_i^{(1)}(\omega_p) = \sum_j \chi_{ij}^{(1)} E_j(\omega_p)$. We thereby find that

$$\chi_{ij}^{(1)}(\omega_p) = \frac{N}{\hbar}\sum_m\left(\frac{\mu_{gm}^i\mu_{mg}^j}{\omega_{mg}-\omega_p} + \frac{\mu_{gm}^j\mu_{mg}^i}{\omega_{mg}^*+\omega_p}\right). \tag{3.2.23}$$

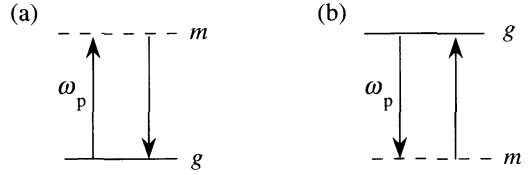

FIGURE 3.2.1 The resonant (a) and antiresonant (b) contributions to the linear susceptibility of Eq. (3.2.23).

The first and second terms in Eq. (3.2.23) can be interpreted as the resonant and antiresonant contributions to the susceptibility, as illustrated in Fig. 3.2.1. In this figure we have indicated where level m would have to be located in order for the corresponding term to become resonant. Note that if g denotes the ground state, it is impossible for the second term to become resonant, which is why it is called the antiresonant contribution.

Second-Order Susceptibility

The expression for the second-order susceptibility is derived in a manner analogous to that used for the linear susceptibility. The second-order contribution (i.e., the contribution second order in $\hat{V}$) to the induced dipole moment per atom is given by

$$\langle \tilde{\mathbf{p}}^{(2)} \rangle = \langle \psi^{(0)} | \hat{\boldsymbol{\mu}} | \psi^{(2)} \rangle + \langle \psi^{(1)} | \hat{\boldsymbol{\mu}} | \psi^{(1)} \rangle + \langle \psi^{(2)} | \hat{\boldsymbol{\mu}} | \psi^{(0)} \rangle, \tag{3.2.24}$$

where $\psi^{(0)}$ is given by Eq. (3.2.10), and $\psi^{(1)}$ and $\psi^{(2)}$ are given by Eqs. (3.2.11), (3.2.16), and (3.2.17). We find that $\langle \tilde{\mathbf{p}}^{(2)} \rangle$ is given explicitly by

$$\begin{aligned} \langle \tilde{\mathbf{p}}^{(2)} \rangle = \frac{1}{\hbar^2} \sum_{pq} \sum_{mn} \Bigg(& \frac{\boldsymbol{\mu}_{gn} [\boldsymbol{\mu}_{nm} \cdot \mathbf{E}(\omega_q)][\boldsymbol{\mu}_{mg} \cdot \mathbf{E}(\omega_p)]}{(\omega_{ng} - \omega_p - \omega_q)(\omega_{mg} - \omega_p)} e^{-i(\omega_p + \omega_q)t} \\ & + \frac{[\boldsymbol{\mu}_{ng} \cdot \mathbf{E}(\omega_q)]^* \boldsymbol{\mu}_{nm} [\boldsymbol{\mu}_{mg} \cdot \mathbf{E}(\omega_q)]}{(\omega_{ng}^* - \omega_q)(\omega_{mg} - \omega_p)} e^{-i(\omega_p - \omega_q)t} \\ & + \frac{[\boldsymbol{\mu}_{ng} \cdot \mathbf{E}(\omega_q)]^* [\boldsymbol{\mu}_{nm} \cdot \mathbf{E}(\omega_p)]^* \boldsymbol{\mu}_{mg}}{(\omega_{ng}^* - \omega_q)(\omega_{mg}^* - \omega_q - \omega_q)} e^{i(\omega_p + \omega_q)t} \Bigg). \end{aligned} \tag{3.2.25}$$

As in the case of the linear susceptibility, this equation can be rendered more transparent by replacing ω_q by $-\omega_q$ in the second term and by replacing ω_q by $-\omega_q$ and ω_p by $-\omega_p$ in the third term; these substitutions are permissible because the expression is to be summed over frequencies ω_p and ω_q.

We thereby obtain the result

$$\langle \tilde{\mathbf{p}}^{(2)} \rangle = \frac{1}{\hbar^2} \sum_{pq} \sum_{mn} \Bigg(\frac{\boldsymbol{\mu}_{gn}[\boldsymbol{\mu}_{nm} \cdot \mathbf{E}(\omega_q)][\boldsymbol{\mu}_{mg} \cdot \mathbf{E}(\omega_p)]}{(\omega_{ng} - \omega_p - \omega_q)(\omega_{mg} - \omega_p)} + \frac{[\boldsymbol{\mu}_{gn} \cdot \mathbf{E}(\omega_q)]\boldsymbol{\mu}_{nm}[\boldsymbol{\mu}_{mg} \cdot \mathbf{E}(\omega_p)]}{(\omega_{ng}^* + \omega_q)(\omega_{mg} - \omega_p)} + \frac{[\boldsymbol{\mu}_{gn} \cdot \mathbf{E}(\omega_q)][\boldsymbol{\mu}_{nm} \cdot \mathbf{E}(\omega_p)]\boldsymbol{\mu}_{mg}}{(\omega_{ng}^* + \omega_q)(\omega_{mg}^* + \omega_q + \omega_q)} \Bigg) e^{-i(\omega_p + \omega_q)t}. \tag{3.2.26}$$

We next take the second-order polarization to be $\tilde{\mathbf{P}}^{(2)} = N\langle \tilde{\mathbf{p}}^{(2)} \rangle$ and represent it in terms of its frequency components as $\tilde{\mathbf{P}}^{(2)} = \sum_r \mathbf{P}^{(2)}(\omega_r)\exp(-i\omega_r t)$. We also introduce the standard definition of the second-order susceptibility (see also Eq. (1.3.13)):

$$P_i^{(2)} = \sum_{jk} \sum_{(pq)} \chi_{ijk}^{(2)}(\omega_p + \omega_q, \omega_q, \omega_p) E_j(\omega_q) E_k(\omega_p)$$

and find that the second-order susceptibility is given by

$$\chi_{ijk}^{(2)}(\omega_p + \omega_q, \omega_q, \omega_p) = \frac{N}{\hbar^2} \mathscr{P}_I \sum_{mn} \Bigg(\frac{\mu_{gn}^i \mu_{nm}^j \mu_{mg}^k}{(\omega_{ng} - \omega_p - \omega_q)(\omega_{mg} - \omega_p)} + \frac{\mu_{gn}^j \mu_{nm}^i \mu_{mg}^k}{(\omega_{ng}^* + \omega_q)(\omega_{mg} - \omega_p)} + \frac{\mu_{gn}^j \mu_{nm}^k \mu_{mg}^i}{(\omega_{ng}^* + \omega_q)(\omega_{mg}^* + \omega_p + \omega_q)} \Bigg). \tag{3.2.27}$$

In this expression, the symbol $\mathscr{P}_I$ denotes the intrinsic permutation operator. This operator tells us to average the expression that follows it over both permutations of the frequencies ω_p and ω_q of the applied fields. The cartesian indices j and k are to be permuted simultaneously. We introduce the intrinsic permutation operator into Eq. (3.2.27) to ensure that the resulting expression obeys the condition of intrinsic permutation symmetry, as described in the discussion of Eqs. (1.4.52) and (1.5.6). The nature of the expression (3.2.27) for the second-order susceptibility can be understood in terms of the energy level diagrams shown in Fig. 3.2.2, which show where the levels m and n would have to be located in order for each term in the expression to become resonant.

The quantum-mechanical expression for the second-order susceptibility actually comprises six terms; through use of the intrinsic permutation operator $\mathscr{P}_I$, we have been able to express the susceptibility in the form (3.2.27), in which only three terms are displayed explicitly. For the case of highly

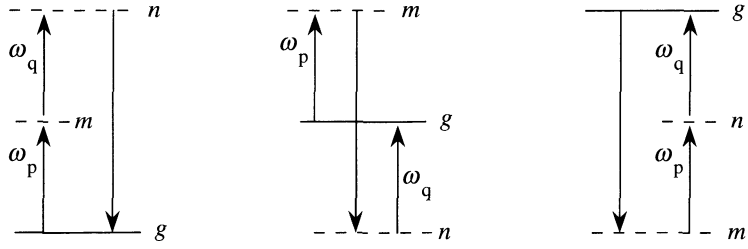

FIGURE 3.2.2 Resonant structure of the three terms of the second-order susceptibility of Eq. (3.2.27).

nonresonant excitation, such that the resonance frequencies ω_{mg} and ω_{ng} can be taken to be real quantities, the expression for $\chi^{(2)}$ can be simplified still further. In particular, under such circumstances Eq. (3.2.27) can be expressed as

$$\chi^{(2)}_{ijk}(\omega_\sigma, \omega_q, \omega_p) = \frac{N}{\hbar^2}\mathscr{P}_F \sum_{mn} \frac{\mu^i_{gn}\mu^j_{nm}\mu^k_{mg}}{(\omega_{ng} - \omega_\sigma)(\omega_{mg} - \omega_p)} \tag{3.2.28}$$

where $\omega_\sigma = \omega_p + \omega_q$. Here we have introduced the full permutation operator, $\mathscr{P}_F$ defined such that the expression that follows it is to be summed over all permutations of the frequencies ω_p, ω_q, and $-\omega_\sigma$, that is, over all input and output frequencies. The cartesian indices are to be permuted along with the frequencies. The final result is then to be divided by the number of permutations of the input frequencies. The equivalence of Eqs. (3.2.27) and (3.2.28) can be verified by explicitly expanding the right-hand side of each equation into all six terms. The six permutations denoted by the operator $\mathscr{P}_F$ are

$$(-\omega_\sigma, \omega_q, \omega_p) \rightarrow (-\omega_\sigma, \omega_p, \omega_q), (\omega_q, -\omega_\sigma, \omega_p), (\omega_q, \omega_p, -\omega_\sigma),$$
$$(\omega_p, -\omega_\sigma, \omega_q), (\omega_p, \omega_q, -\omega_\sigma).$$

Since we can express the nonlinear susceptibility in the form of Eq. (3.2.28), we have proven the statement made in Section 1.5 that the nonlinear susceptibility of a lossless medium possesses full permutation symmetry.

Third-Order Susceptibility

We now calculate the third-order susceptibility. The dipole moment per atom, correct to third order in perturbation theory, is given by

$$\langle \tilde{\mathbf{p}}^{(3)} \rangle = \langle \psi^{(0)} | \hat{\boldsymbol{\mu}} | \psi^{(3)} \rangle + \langle \psi^{(1)} | \hat{\boldsymbol{\mu}} | \psi^{(2)} \rangle + \langle \psi^{(2)} | \hat{\boldsymbol{\mu}} | \psi^{(1)} \rangle + \langle \psi^{(3)} | \hat{\boldsymbol{\mu}} | \psi^{(0)} \rangle. \tag{3.2.29}$$