

An important feature of coupled-mode equations is *phase matching*, which represents the degree to which the induced mode coupling terms in the polarization have the same phase as the field modes they affect. To the extent that the phases differ, the mode coupling is reduced. Phase matching involves differences in wave vectors and amounts to conservation of momentum. This is distinct from the frequency differences of the last section, which amount to conservation of energy.

Indeed, the resonance denominators appearing in the nonlinear susceptibilities can be interpreted as a consequence of energy conservation. In contrast, the spatial phase factors are a result of momentum conservation. This is particularly apparent when the electromagnetic field is quantized, as will be the case in the second part of the book, because in that case, it is easy to show that the energy and momentum of a photon of frequency ν and momentum \mathbf{K} are $k\nu$ and $k\mathbf{K}$, respectively. In a vacuum, one has that $\nu = Kc$ for all frequencies, so that energy conservation automatically guarantees momentum conservation. But this ceases to be true in dispersive media, where the factor of proportionality between ν and K is frequency-dependent.

Consider for instance the case of difference-frequency generation, where two incident waves at frequencies ν_1 and ν_2 combine in the nonlinear medium to generate a wave at the difference frequency ν_d . In the slowly-varying amplitude approximation and in steady state (we neglect the $\partial/\partial t$ terms), (1.43) becomes

$$\frac{dE_n}{dz} = \frac{iK_n}{2\epsilon_n} P_n, \quad (2.12)$$

where ϵ_n is the permittivity at the frequency ν_n of the host medium in which our oscillators are found. From (1.64, 1.65), we use the linear solutions for modes 1 and 2

$$\frac{dE_n}{dz} = \frac{N(z)e^2}{c i K_n 2\epsilon_n m} \frac{E_n}{\omega^2 - \nu_n^2 - 2i\nu_n\gamma} = -\alpha_n E_n. \quad (2.13)$$

Note that in this example we assume that the host medium is purely dispersive; otherwise another absorption term would have to be included. Using (2.10, 2.12), we find for the difference-frequency term at the frequency $\nu_d = \nu_1 - \nu_2$

$$\frac{dE_d(z)}{dz} = \frac{iK_d N}{2\epsilon_d} \frac{(e/m)E_d - a x_1^{(1)} [x_2^{(1)}]^* e^{i(K_1 - K_2 - K_d)z}}{\omega^2 - \nu_d^2 - 2i\nu_d\gamma}. \quad (2.14)$$

The coupled-mode equations (2.13, 2.14) take a rather simple form, since we have neglected the back action of E_d on E_1 and E_2 . Equation (2.13) simply describes the linear absorption and dispersion of E_1 and E_2 due to the nonlinear oscillators, as does the first term on the right-hand side of (2.14) for E_d . In many cases these are small effects compared to those of the host medium accounted for by ϵ_n , and we neglect them in the following. E_1 and E_2

then remain constant and (2.14) (without the leading term) can be readily integrated over the length L of the nonlinear medium to give

$$E_d(L) = GE_1E_2^* \frac{e^{i\Delta KL} - 1}{i\Delta K}, \quad (2.15)$$

where

$$G = - \frac{iK_d a N e^{-x^{(1)}[x^{(1)*}/E_1E^*]}}{2\varepsilon_d} \frac{1}{\omega^2 - v_d^2 - 2iv_d\gamma},$$

and the K -vector mismatch $\Delta K = K_1 - K_2 - K_d$. The resulting intensity $I_d = |E_d|^2$ is

$$I_d(L) = |G|^2 I_1 I_2 \frac{\sin^2(\Delta KL/2)}{(\Delta K/2)^2}. \quad (2.16)$$

If $\Delta K = 0$, $I_d(L)$ reduces to $|G|^2 I_1 I_2 L^2$, but for $\Delta K \neq 0$, it oscillates periodically. To achieve efficient frequency conversion, it is thus crucial that

$(K_1 - K_2)L$ be close to $K_d L$. For $\Delta K \neq 0$ the maximum intensity I_d is reached

for a medium of length $L = \pi/\Delta K$. For larger values of ΔKL , the induced polarization at the frequency ν_d and the wave propagation at that frequency start to interfere destructively, attenuating the wave. For still larger values L , the interference once again becomes positive, and continues to oscillate in this fashion. Since nonlinear crystals are expensive, it is worth trying to achieve the best conversion with the smallest crystal, namely for $\Delta KL = \pi$. In the plane-wave, collinear propagation model described here, perfect phase matching requires that the wave speeds $u_n = v_n/K_n$ all be equal, as would be the case in a dispersionless medium. More generally we have the difference $\Delta K = K_1 - K_2 - K_d = (n_1\nu_1 - n_2\nu_2 - n_d\nu_d)/c \neq 0$ since the n 's differ. For

noncollinear operation the vectorial phase matching condition

$$\Delta KL = |\mathbf{K}_1 - \mathbf{K}_2 - \mathbf{K}_d| L \leq \pi \quad (2.17)$$

must be fulfilled for maximum I_d . There are a number of ways to achieve this, including appropriate geometry, the use of birefringent media, and temperature index tuning.

2.3 Cubic Nonlinearity

We already mentioned that quadratic nonlinearities such as described in the preceding sections do not occur in isolated atoms, for which the lowest order nonlinear effects are cubic in the fields. These can be described in our classical model by keeping the $b\mathbf{x}^3$ term instead of $a\mathbf{x}^2$ in the nonlinear oscillator equation (2.1). In the presence of two strong pump fields at frequencies ν_1 and ν_2 , the third-order polarization given by $b\mathbf{x}^3$ includes contributions at the frequencies ν_1 and ν_2 and at the *sideband* frequencies $\nu_0 = \nu_1 - \Delta$ and $\nu_3 = \nu_2 + \Delta$ as well, where

$$\Delta = \nu_2 - \nu_1. \quad (2.18)$$

The generation of these sidebands is an example of four-wave mixing. To describe the initial growth of the sidebands, we write the anharmonic term $b\mathbf{x}^3$ to third-order in \mathbf{x}_1 and \mathbf{x}_2 , and first-order in the small displacements \mathbf{x}_0 and \mathbf{x}_3 , that is,

$$\begin{aligned} [\mathbf{x}^{(1)}]^3 = & \frac{1}{8} [\mathbf{x}^{(1)} e^{i(K_1 z - \nu_1 t)} + \mathbf{x}^{(1)} e^{i(K_2 z - \nu_2 t)} \\ & + 3\mathbf{x}^{(1)} e^{i(K_0 z - \nu_0 t)} + 3\mathbf{x}_3 e^{i(K_3 z - \nu_3 t)} + \text{c.c.}] \\ & \times \frac{1}{2} \mathbf{x}^{(1)} e^{i[(K_1 + K_2)z - (\nu_1 + \nu_2)t]} \\ & + 2\mathbf{x}^{(1)} [\mathbf{x}_2]^* e^{i[(K_1 - K_2)z - (\nu_1 + \nu_2)t]} \\ & + \left\{ [\mathbf{x}^{(1)}]^2 e^{2i(K_n z - \nu_n t)} + |\mathbf{x}_n^{(1)}|^2 \right\} + \text{c.c.}, \end{aligned} \quad (2.19)$$

$n=1$

where the terms in $\{\}$ are similar to those in (2.9). The factor of 3 results from the three ways of choosing the \mathbf{x}_0 and \mathbf{x}_3 from the triple product.

The curly braces in (2.19) contain two dc terms, a contribution oscillating at the difference frequency Δ , and three rapidly oscillating contributions oscillating at the frequencies $\nu_n + \nu_m$. These time-dependent terms are sometimes called (complex) index gratings, and the nonlinear polarization may be interpreted as the scattering of a light field E_n from the grating produced by two fields E_m and E_k . In this picture, the dc terms are “degenerate” gratings produced by the fields E_m and E_m^* . Equation (2.19) readily gives the third-order contributions to the components of the polarization P_n at the frequencies of interest.

One can interpret (2.19) as the scattering of components in the $[\]$ of the first lines off the slowly varying terms in the $\{\}$. Specifically the $|\mathbf{x}_n^{(1)}|^2$ terms in (2.19) contribute nonlinear changes at the respective frequencies of the components in the $[\]$. In contrast, the scattering off the “Raman-like” term $\exp[i(K_2 - K_1)z - i\Delta t]$ and its complex conjugate contribute corrections at frequencies shifted by $\pm\Delta$. Taking $\nu_2 > \nu_1$, we see that the ν_2 term in the $[\]$ scatters producing components at both the lower frequency ν_1 (called a *Stokes* shift) and the higher frequency $\nu_3 = \nu_2 + \Delta$ (called an *anti-Stokes* shift). Similarly the ν_1 term in the $[\]$ leads to contributions at the frequencies $\nu_0 = \nu_1 - \Delta$ and at ν_2 . The induced polarization components at the frequencies ν_0 and ν_3 are called *combination tones*. They are generated in the nonlinear medium from other frequencies. If the two pump beams at ν_1 and ν_2 are copropagating, the index grating represented by the $K_2 - K_1$ term propagates at approximately the velocity of light in the host medium, but if the beams are counterpropagating, the grating propagates at the relatively slow speed $v = -\Delta/(K_1 + K_2)$. In particular, it becomes stationary for the degenerate case $\nu_1 = \nu_2$. (Compare with the ponderomotive force acting on the electrons

in the free electron laser, (1.126)! –Can you draw an analogy between the two situations?)

We are often only interested in induced polarizations near or at the fundamental frequencies $\nu_n, n = 0, 1, 2, 3$. Keeping only these in (2.19) and neglecting combination tones involving x_0 and x_1 in the pump-mode polarizations (Prob. 2.7), we find

$$\begin{aligned}
 |[\mathbf{x}^{(1)}]^3|_{\text{fund}} = & \frac{3}{8} \mathbf{x}^{(1)} (|\mathbf{x}_1^{(1)}|^2 + 2|\mathbf{x}_2^{(1)}|^2) e^{i(K_1 z - \nu_1 t)} \\
 & + \frac{3}{8} \mathbf{x}^{(1)} (|\mathbf{x}_2^{(1)}|^2 + 2|\mathbf{x}_1^{(1)}|^2) e^{i(K_2 z - \nu_2 t)} \\
 & + \frac{1}{8} [|\mathbf{x}_1^{(1)}|^2 + |\mathbf{x}_2^{(1)}|^2] [\mathbf{x}_0^{(1)} e^{i(K_0 z - \nu_0 t)} + \mathbf{x}_3^{(1)} e^{i(K_3 z - \nu_3 t)}] \\
 & + \frac{1}{8} \mathbf{x}^{(1)} (1) (1) e^{i[(K_1 + K_2 - K_3)z - \nu_0 t]} \\
 & + \frac{1}{8} \mathbf{x}_2^{(1)} [\mathbf{x}_3^{(1)}]^* e^{i[(K_1 + K_2 - K_0)z - \nu_3 t]} \\
 & + \frac{1}{8} \mathbf{x}_2^{(1)} [\mathbf{x}_0^{(1)}]^* e^{i[(2K_1 - K_2)z - \nu_0 t]} \\
 & + \frac{1}{8} [\mathbf{x}_1^{(1)}]^2 [\mathbf{x}_2^{(1)}]^* e^{i[(2K_2 - K_1)z - \nu_3 t]} \\
 & + \frac{1}{8} [\mathbf{x}_1^{(1)}]^2 [\mathbf{x}_1^{(1)}]^* e^{i[(2K_2 - K_1)z - \nu_3 t]} + \text{c.c.} \quad (2.20)
 \end{aligned}$$

Combining the various terms, we find that the third-order polarization components are given by

$$\begin{aligned}
 P_0 = & \frac{3}{8} \text{Neb} [|\mathbf{x}_1^{(1)}|^2 + |\mathbf{x}_2^{(1)}|^2] \mathbf{x}_0^{(1)} \\
 & + \frac{1}{8} \text{Neb} \mathbf{x}_1^{(1)} \mathbf{x}_2^{(1)} [\mathbf{x}_3^{(1)}]^* e^{i(K_1 + K_2 - K_3 - K_0)z} \\
 & + \frac{1}{8} \mathbf{x}_2^{(1)} [\mathbf{x}_3^{(1)}]^* e^{i(K_1 + K_2 - K_0)z} \\
 P_1 = & \frac{3}{8} \text{Neb} \mathbf{x}_1^{(1)} [|\mathbf{x}_1^{(1)}|^2 + 2|\mathbf{x}_2^{(1)}|^2] \quad (2.21b)
 \end{aligned}$$

$$P_2 = \frac{3}{8} \text{Neb} \mathbf{x}_2^{(1)} [2|\mathbf{x}_1^{(1)}|^2 + |\mathbf{x}_2^{(1)}|^2] \quad (2.21c)$$

$$\begin{aligned}
 P_3 = & \frac{3}{8} \text{Neb} [|\mathbf{x}_1^{(1)}|^2 + |\mathbf{x}_2^{(1)}|^2] \mathbf{x}_3^{(1)} \\
 & + \frac{1}{8} \text{Neb} \mathbf{x}_1^{(1)} \mathbf{x}_2^{(1)} [\mathbf{x}_0^{(1)}]^* e^{i(K_1 + K_2 - K_0 - K_3)z} \\
 & + \frac{1}{8} \text{Neb} [\mathbf{x}_1^{(1)}]^2 [\mathbf{x}_2^{(1)}]^* e^{i(2K_2 - K_1 - K_3)z} . \quad (2.21d)
 \end{aligned}$$

The polarization components $P^{(3)}$ and $P^{(3)}$ are solely due to the existence of index gratings, which are also responsible for the factors of 2 in the cross coupling terms for $P^{(3)}$ and $P^{(3)}$. This asymmetry is sometimes called nonlinear nonreciprocity and was discovered in quantum optics by Chiao

et al. (1966). It also appears in the work by van der Pol (1934) on coupled vacuum-tube tank circuits. In the absence of index gratings, the factors of 2 in (2.21b, c) are replaced by 1, and $|x_1^{(r)}|^2$ and $|x_2^{(r)}|^2$ play symmetrical roles in $P_1^{(3)}$ and $P_2^{(3)}$.

The polarizations P_n lead to coupled-mode equations for the field envelopes. The procedure follows exactly the method of Sect. 2.2 and we obtain (Prob. 2.2)

$$\frac{dE_0}{dz} = -E_0 \left[\alpha_0 - \theta_{01}|E_1|^2 - \theta_{02}|E_2|^2 \right] + \vartheta_{0121} E_1^2 E_2^* e^{i(2K_1 - K_2 - K_0)z} \\ - \vartheta_{0122} E_1^2 E_2^* e^{i(K_1 + K_2 - K_3 - K_0)z}, \quad (2.22a)$$

$$\frac{dE_1}{dz} = -E_1 [\alpha_1 - \beta_1 |E_1|^2 - \theta_{12}|E_2|^2], \quad (2.22b)$$

$$\frac{dE_2}{dz} = -E_2 [\alpha_2 - \beta_2 |E_2|^2 - \theta_{21}|E_1|^2], \quad (2.22c)$$

$$\frac{dE_3}{dz} = -E_3 \left[\alpha_3 - \theta_{31}|E_1|^2 - \theta_{32}|E_2|^2 \right] + \vartheta_{3212} E_2^2 E_1^* e^{i(2K_2 - K_1 - K_3)z} \\ - \vartheta_{3211} E_1^2 E_2^* e^{i(K_1 + K_2 - K_0 - K_3)z}. \quad (2.22d)$$

Here E_n is the complex amplitude of the field at frequency ν_n , and the $-\alpha_n E_n$ terms allow for linear dispersion and absorption.

Equations (2.22b, c) for the pump modes amplitudes are coupled by the cross-coupling (or cross-saturation) coefficients θ_{nj} . To this order of perturbation, they are independent of the sidemode amplitudes E_0 and E_3 . Because E_1 and E_2 always conspire to create an index grating of the correct phase, the evolution of these modes is not subject to a phase matching condition. Equations of this type are rather common in nonlinear optics and laser theory. In Sect. 7.4, we obtain an evolution of precisely this type for the counterpropagating modes in a ring laser. We show that the cross-coupling between modes can lead either to the suppression of one of the modes or to their coexistence, depending on the magnitude of the coupling parameter $C = \theta_{12}\theta_{21}/\beta_1\beta_2$ and relative sizes of the α_n .

In contrast, the sidemodes E_0 and E_3 are coupled to the strong pump fields E_1 and E_2 only, and not directly to each other. They have no back-action on the pump modes dynamics, and their growth is subject to a phase-matching condition.

2.4 Four-Wave Mixing with Degenerate Pump Frequencies

In many experimental situations, it is convenient to drive the nonlinear medium with two pump fields of the same frequency ν_2 , but with opposite propagation directions given by the wave vectors $\mathbf{K}_{2\downarrow}$ and $\mathbf{K}_{2\uparrow}$. The pump

waves cannot by themselves generate polarization components at sideband frequencies. However one can still take advantage of the index gratings produced by the pump beams with weak waves at frequencies symmetrically detuned from ν_2 by a small amount $\pm\Delta$ (see Fig. 2.1). This procedure has gained considerable popularity in connection with optical phase conjugation. In optical phase conjugation, one of the sidebands is called the probe (at $\nu_1 = \nu_2 - \Delta$) and the other the signal (at $\nu_3 = \nu_2 + \Delta$), and we adopt this notation here in anticipation of Chaps. 9, 10.

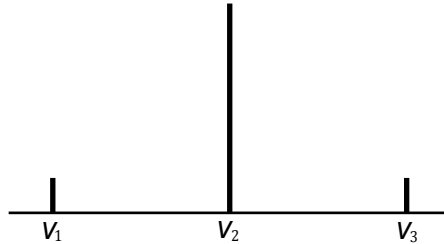


Fig. 2.1. Mode spectrum in four-wave mixing for optical phase conjugation

We consider the wave configuration in Fig. 2.2 with two counterpropagating pump beams along one direction, and counterpropagating signal and “conjugate” waves along another direction, which we call the z axis. The electric field for these four waves has the form

$$E(\mathbf{r}, t) = \frac{1}{2} [E_1 e^{i(K_1 z - \nu_1 t)} + E_{2\downarrow} e^{i(\mathbf{K}_{2\downarrow} \cdot \mathbf{r} - \nu_2 t)} + E_{2\uparrow} e^{i(\mathbf{K}_{2\uparrow} \cdot \mathbf{r} - \nu_2 t)} + E_3 e^{i(-K_3 z - \nu_3 t)}] + \text{c.c.}, \quad (2.23)$$

where we take $\mathbf{K}_{2\uparrow} = -\mathbf{K}_{2\downarrow}$. The field fringe patterns resulting from interference between the various waves can induce index gratings. The corresponding linear displacement $\mathbf{x}^{(1)}(t)$ contains components proportional to each of the field amplitudes, and the third-order nonlinear displacement $\mathbf{x}^{(3)}$ consists of the sum of all terms proportional to the products of three fields, each of which can be anyone of the four waves or their complex conjugates. This gives a grand total of $8 \cdot 8 \cdot 8 = 512$ terms. Fortunately, we’re only interested in a relatively small subset of these terms, namely those with the positive frequency ν_1 linear in E_1 . This gives a third-order signal polarization $\mathcal{P}_1^{(3)}$ proportional to

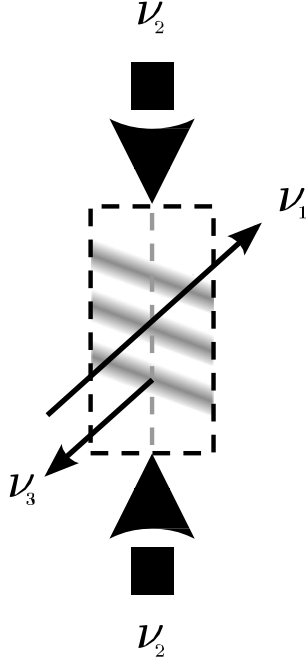


Fig. 2.2. Diagram of interaction between standing-wave pump beam (ν_2) with probe (ν_1) and conjugate (ν_3) beams used in phase conjugation

$$\begin{aligned}
 & E_1 E_2 \downarrow E^* + E_1 E^* E_2 \downarrow + E_2 \downarrow E_1 E^* + E^* E_1 E_2 \downarrow \\
 & + E_2 \downarrow E^* E_1 \downarrow + E^* E_2 \downarrow E_1 + E_1 E_2 \uparrow E^* + E_1 E^* E_2 \uparrow \\
 & + E_2 \uparrow E_1 E^* + E^* E_1 E_2 \uparrow + E_2 \uparrow E^* E_1 + E^* E_2 \uparrow E_1 \\
 & + [E_3^* E_2 \uparrow E_2 \downarrow + E_3^* E_2 \downarrow E_2 \uparrow + E_2 \uparrow E_3^* E_2 \downarrow + E_2 \downarrow E_3^* E_2 \uparrow \\
 & + E_2 \uparrow E_2 \downarrow E_3^* + E_2 \downarrow E_2 \uparrow E_3^*] e^{i(K_3 - K_1)z} \quad (2.24a)
 \end{aligned}$$

$$= 6E_1(|E_2 \downarrow|^2 + |E_2 \uparrow|^2) + 6E_3 E_2 \uparrow E_2 \downarrow e^{i(K_3 - K_1)z} \quad (2.24b)$$

The various terms in (2.24a) have simple physical interpretations. For instance, the first term results from the product in $[\chi^{(1)}]^3$

$$E_1 e^{i(K_1 z - \nu_1 t)} E_2 \downarrow e^{i(\mathbf{K}_2 \downarrow \cdot \mathbf{r} - \nu_2 t)} E_2^* \downarrow e^{-i(\mathbf{K}_2 \downarrow \cdot \mathbf{r} - \nu_2 t)} .$$

Note that the pump phase dependencies $\mathbf{K}_2 \downarrow \cdot \mathbf{r}$ cancel identically, as they do for all terms in (2.24). The first and second terms represent contributions to the nonlinear refraction of the field E_1 due the nonlinear index (Kerr effect) induced by the pump field intensity $I_2 \downarrow$. The third term can be understood as originating from the scattering of the field $E_2 \downarrow$ off the grating produced by E_1 and $E_2 \downarrow$, etc. Its effect on the polarization is precisely the same as that of the

first two terms, but we have intentionally written it separately in anticipation of the quantum mechanical discussion of Chap. 10, where the order in which the fields are applied to the medium matters. Indeed in our classical model, the terms in the first three lines in (2.24a) are all proportional to the product of a pump field intensity and the probe field E_1 , and can be globally described as nonlinear absorption and refraction terms.

The last two lines in (2.24a) couple the sidemode E_3 to E_1 via the following scattering mechanism: the field E_3^* interferes with the pump fields $E_{2\uparrow}$ and $E_{2\downarrow}$ to induce two complex index gratings that scatter $E_{2\downarrow}$ and $E_{2\uparrow}$, respectively, into E_1 . This process, which is essentially the real-time realization of holographic writing and reading, is called phase conjugation and is discussed in detail in Chap. 10. The process retroreflects a wavefront, sending it back along the path through which it came (see Fig. 10.1). It can be used to compensate for poor optics. Note that although the pump phase dependencies cancel one another as they do for the terms in the first four lines of (2.24a), the induced polarization has the phase $\exp[i(K_3 z - v_1 t)]$, while Maxwell's equations require $\exp[i(K_1 z - v_1 t)]$. This gives the phase mismatch factor $\exp[i(K_3 - K_1)z]$, which is important except in the degenerate frequency case $v_1 = v_2 = v_3$, for which $K_3 = K_1$.

Neglecting the depletion of the pump beams $E_{2\uparrow}$ and $E_{2\downarrow}$, we find the coupled-mode equations for E_1 and E_3^*

$$\frac{dE_1}{dz} = -\alpha_1 E_1 + \chi_1 E_3^* e^{2i\Delta K z}, \quad (2.25a)$$

$$-\frac{dE_3^*}{dz} = -\alpha_3^* E_3^* + \chi_3^* E_1 e^{-2i\Delta K z}, \quad (2.25b)$$

where

$$\Delta K = \frac{K_3 - K_1}{2}. \quad (2.26)$$

Here $-dE_3^*/dz$ appears since E_3^* propagates along $-z$, and we use χ_n for the coupling coefficient to agree with later usage, although it is only a part of a susceptibility.

To solve these equations, we proceed by first transforming away the phase mismatch by the substitution

$$E_1 = A_1 e^{2i\Delta K z} \quad (2.27)$$

into (2.25). In particular, (2.25a) becomes

$$\frac{dA_1}{dz} = -(\alpha_1 + 2i\Delta K)A_1 + \chi_1 E_3^*. \quad (2.28)$$

We seek solutions of (2.25b, 2.28) of the form $e^{\mu z}$. Substituting $A_1 = e^{\mu z}$ into (2.28), solving for E_3^* , and substituting the result into (2.25b), we find the eigenvalues

$$\begin{aligned}\mu_{\pm} &= -\frac{1}{2}(\alpha_1 - \alpha_3^* + 2i\Delta K) \pm [(\alpha_1 + \alpha_3^* + 2i\Delta K)^2/4 - \chi_1\chi_3^*]^{1/2} \\ &= -a \pm [\alpha^2 - \chi_1\chi_3^*]^{1/2} = -a \pm w.\end{aligned}\quad (2.29)$$

Hence the general solutions are

$$A_1(z) = e^{-az}[Ae^{wz} + Be^{-wz}] \quad (2.30a)$$

and

$$E_3^*(z) = e^{-az}[Ce^{wz} + De^{-wz}], \quad (2.30b)$$

where the coefficients A, B, C and D are determined by the boundary conditions of the problem.

We suppose here that a weak signal weak field $E_1(0)$ is injected inside the nonlinear medium at $z = 0$, and we study the growth of the counterpropagating conjugate wave E_3^* , which is taken to be zero at $z = L$. This means that $A_1(0) = E_1(0) = \text{constant}$, and $E_3^*(L) = 0$, in which case one has immediately $B = E_1(0) - A$ and $D = -Ce^{wL}$. Matching the boundary conditions of (2.25b, 2.28) at $z = L$ yields

$$A = \frac{1}{2}A_1(0)e^{-\frac{wL}{2}}(w - \alpha)/(w \cosh wL + \alpha \sinh wL), \quad (2.31)$$

$$2wCe^{wL} = \chi_3^*(A \sinh wL + A_1(0)e^{-wL}). \quad (2.32)$$

Further manipulation yields finally

$$E_1(z) = E_1(0)e^{-(a+w+2i\Delta K)z} \left[1 + \frac{(w - \alpha)e^{w(z-L)} \sinh wz}{w \cosh wL + \alpha \sinh wL} \right], \quad (2.33)$$

$$E_3^*(z) = \chi_3^* E_1(0) \frac{e^{-az} \sinh w(z-L)}{w \cosh wL + \alpha \sinh wL}. \quad (2.34)$$

In particular the *amplitude reflection coefficient* $r = E_3^*(0)/E_1(0)$ is given by

$$r = \frac{E_3^*(0)}{E_1(0)} = -\chi_3^* \frac{\sinh wL}{w \cosh wL + \alpha \sinh wL}. \quad (2.35)$$

See Chap. 10 on phase conjugation for further discussion of these equations.

Coupled Modes and Squeezing

A popular topic is the “squeezing”, i.e., deamplifying, of noise in one quadrature of an electromagnetic wave at the expense of amplifying the noise in the orthogonal quadrature. One way to achieve such squeezing is through the use of mode coupling mechanisms such as described by (2.25a, b). To see under which conditions the χ_n coupling factors can lead to this quadrature-dependent amplification, let's drop the α_n term in (2.25) and put the time dependencies back in. We find for example the schematic equation

$$\{\chi^{(3)} E_2^2 e^{-2i\nu t}\} [E_3 e^{i\nu t}]^* \rightarrow E_1 e^{-i\nu t}, \quad (2.36)$$

where $\chi^{(3)}$ is a third-order susceptibility. Suppose that at a time t , $e^{i\nu t} = 1$ and that $\{\chi\} = 1$. According to (2.36), this tends to amplify E_1 . Now wait until the orthogonal quadrature phasor $\exp(i\nu t - i\pi/2) = 1$. At this time, the second-harmonic (two-photon) phasor $\exp(-2i\nu t)$ has precessed through *two* times $\pi/2$, that is, $\{\chi\} = -1$. Hence a two-photon coupling $\{\chi\}$ flips the sign of the coupling between orthogonal quadratures. This is the signature of a coupling process that can lead to squeezing. It is equally possible that a $\chi^{(2)}$ process with a single pump photon having the value 2ν can cause squeezing. This $\chi^{(2)}$ process is known as parametric amplification. Chapter 17 discusses the squeezing of quantum noise by four-wave mixing.

2.5 Nonlinear Susceptibilities

So far we have used an anharmonic oscillator to introduce some aspects of nonlinear optics that are useful in the remainder of this book. Such a simple model is surprisingly powerful and permits us to understand numerous nonlinear optics effects intuitively. In general, however, first principle quantum mechanical calculations are needed to determine the response of a medium to a strong electromagnetic field. A substantial fraction of this book addresses this problem under resonant or near-resonant conditions, i.e., under conditions such that the frequency (ies) of the field(s) are near an atomic transition. Perturbative analyses such as sketched in this chapter are usually not sufficient to describe these situations.

In many cases, however, the incident radiation is far from resonance with any transition of interest, and/or the material relaxation rate is exceedingly fast. In such cases, perturbation theory based on the concept of nonlinear susceptibility may be of great advantage. This is the realm of conventional nonlinear optics, and the reader should consult the recent treatises by Shen (1984), by Hopf and Stegeman (1986), and by Boyd (1992), as well as the classic book by Bloembergen (1965), for detailed descriptions of these fields. Here we limit ourselves to a brief introduction to the formalism of nonlinear susceptibility.

In linear problems, the polarization of the medium is (by definition) a linear function of the applied electric fields. The most general form that it can take is given by the space-time convolution of a linear susceptibility tensor $\chi^{(1)}$ with the electric field:

$$\mathbf{P}(\mathbf{r}, t) = \epsilon_0 \int_{-\infty}^t d^3 r' \int_{-\infty}^t dt' \chi^{(1)}(\mathbf{r} - \mathbf{r}', t - t') : \mathbf{E}(\mathbf{r}', t'). \quad (2.37)$$

Taking the four-dimensional Fourier transform of this expression for a monochromatic wave $\mathbf{E}(\mathbf{r}, t) = \mathbf{E}(\mathbf{K}, \nu) e^{i(\mathbf{K} \cdot \mathbf{r} - \nu t)}$, we find

$$\mathbf{P}(\mathbf{K}, \nu) = \varepsilon_0 \chi^{(1)}(\mathbf{K}, \nu) : \mathbf{E}(\mathbf{K}, \nu) , \quad (2.38)$$

with

$$\chi^{(1)}(\mathbf{K}, \nu) = \int_{-\infty}^{\infty} d^3 r \int_{-\infty}^t dt' \chi^{(1)}(\mathbf{r}, t') e^{i(\mathbf{K} \cdot \mathbf{r} - \nu t')} . \quad (2.39)$$

The linear dielectric constant is related to $\chi^{(1)}(\mathbf{K}, \nu)$ via

$$\varepsilon(\mathbf{K}, \nu) = \varepsilon_0 [1 + \chi^{(1)}(\mathbf{K}, \nu)] . \quad (2.40)$$

In the nonlinear case, and for electric fields sufficiently weak that perturbation theory is valid, one gets instead

$$\begin{aligned} \mathbf{P}(\mathbf{r}, t) = & \varepsilon_0 \int_{-\infty}^{\infty} d^3 r' \int_{-\infty}^t dt' \chi^{(1)}(\mathbf{r} - \mathbf{r}', t - t') \cdot \mathbf{E}(\mathbf{r}', t') \\ & + \varepsilon_0 \int_{-\infty}^{\infty} d\mathbf{r}_1 d\mathbf{r}_2 dt_1 dt_2 \chi^{(2)}(\mathbf{r} - \mathbf{r}_1, t - t_1; \mathbf{r} - \mathbf{r}_2, t - t_2) : \mathbf{E}(\mathbf{r}_1, t_1) \mathbf{E}(\mathbf{r}_2, t_2) \\ & + \varepsilon_0 \int_{-\infty}^{\infty} d\mathbf{r}_1 d\mathbf{r}_2 dt_1 dt_2 d\mathbf{r}_3 dt_3 \chi^{(3)}(\mathbf{r} - \mathbf{r}_1, t - t_1; \mathbf{r} - \mathbf{r}_2, t - t_2; \mathbf{r} - \mathbf{r}_3, t - t_3) \\ & \cdot \mathbf{E}(\mathbf{r}_1, t_1) \mathbf{E}(\mathbf{r}_2, t_2) \mathbf{E}(\mathbf{r}_3, t_3) + \dots , \end{aligned} \quad (2.41)$$

where $\chi^{(n)}$ the n th-order susceptibility. If $\mathbf{E}(\mathbf{r}, t)$ can be expressed as a sum of plane waves,

$$\mathbf{E}(\mathbf{r}, t) = \sum_n \mathbf{E}(\mathbf{K}_n, \nu_n) e^{i(\mathbf{K}_n \cdot \mathbf{r} - \nu_n t)} , \quad (2.42)$$

then as in the linear case, the Fourier transform of (2.41) gives

$$\mathbf{P}(\mathbf{K}, \nu) = \mathbf{P}^{(1)}(\mathbf{K}, \nu) + \mathbf{P}^{(2)}(\mathbf{K}, \nu) + \mathbf{P}^{(3)}(\mathbf{K}, \nu) + \dots \quad (2.43)$$

with $\mathbf{P}^{(1)}(\mathbf{K}, \nu)$ given by (2.38) and

$$\begin{aligned} \mathbf{P}^{(2)}(\mathbf{K}, \nu) &= \chi^{(2)}(\mathbf{K} = \mathbf{K}_n + \mathbf{K}_m, \nu = \nu_n + \nu_m) : \mathbf{E}(\mathbf{K}_n, \nu_n) \mathbf{E}(\mathbf{K}_m, \nu_m) \\ \mathbf{P}^{(3)}(\mathbf{K}, \nu) &= \chi^{(3)}(\mathbf{K} = \mathbf{K}_n + \mathbf{K}_m + \mathbf{K}_A, \nu = \nu_n + \nu_m + \nu_A) \\ &\quad \cdot \mathbf{E}(\mathbf{K}_n, \nu_n) \mathbf{E}(\mathbf{K}_m, \nu_m) \mathbf{E}(\mathbf{K}_A, \nu_A) \end{aligned}$$

and

$$\begin{aligned} \chi^{(n)}(\mathbf{K} = \mathbf{K}_1 + \mathbf{K}_2 + \dots + \mathbf{K}_n, \nu = \nu_1 + \nu_2 + \dots + \nu_n) \\ = \int_{-\infty}^{\infty} d^3 r_1 d\mathbf{r}_2 dt_1 dt_2 \dots d^3 r_n dt_n \chi^{(n)} \\ \times (\mathbf{r} - \mathbf{r}_1, t - t_1; \mathbf{r} - \mathbf{r}_2, t - t_2; \dots \mathbf{r} - \mathbf{r}_n, t - t_n) \\ \times \exp[i\mathbf{K}_1 \cdot (\mathbf{r} - \mathbf{r}_1) - i\nu_1(t - t_1) + \dots + i\mathbf{K}_n \cdot (\mathbf{r} - \mathbf{r}_n) - i\nu_n(t - t_n)] . \end{aligned}$$

Problems

2.1 Solve the coupled-mode equations

$$\frac{dE_1}{dz} = -\alpha_1 E_1 + \chi_1 E_3^*, \quad (2.44)$$

$$\frac{dE_3^*}{dz} = -\alpha_3^* E_3^* + \chi_3^* E_1, \quad (2.45)$$

valid for phase-matched forward three-wave mixing. Ans:

$$E_1(z) = e^{-az} [E_1(0) \cosh wz + (-\alpha E_1(0) + \chi_1 E_3^*(0)) \sinh wz/w] \quad (2.46)$$

$$E_3^*(z) = e^{-az} [E_3^*(0) \cosh wz + (a E_3^*(0) + \chi_3^* E_1(0)) \sinh wz/w], \quad (2.47)$$

where $a = (\alpha_1 + \alpha_3^*)/2$, $\alpha = (\alpha_1 - \alpha_3^*)/2$, and $w = \alpha^2 + \chi_1 \chi_3^*$.

2.2 Derive the coefficients in the coupled-mode equations (2.22).

2.3 Calculate all wavelengths generated in a $\chi^{(3)}$ nonlinear medium by a combination of 632.8 nm and 388 nm laser light.

2.4 Calculate the coupling coefficient χ_n for four-wave mixing based on an anharmonic oscillator.

2.5 Write the propagation equations for second-harmonic generation. Comment on phase matching.

2.6 Calculate the phase mismatch for a conjugate wave of frequency $\nu_3 = \nu_2 + (\nu_2 - \nu_1)$ generated by signal and pump waves with frequencies ν_1 and ν_2 , respectively, and propagating in the same direction. Include the fact that the indices of refraction for the three waves are in general different, that is, $\eta(\nu_1) = \eta(\nu_2) + \delta\eta_1$ and $\eta(\nu_3) = \eta(\nu_2) + \delta\eta_3$.

2.7 Show that (2.20) contains all the fundamental contributions from the third-order expression (2.19).

2.8 Evaluate the reflection coefficient r of (2.35) in the limit of large L . Answer: $r = \mp \chi_3^*/(w \pm \alpha)$ for $\text{Re}(w) \gtrless 0$.

3 Quantum Mechanical Background

Chapters 1, 2 describe the interaction of radiation with matter in terms of a phenomenological classical polarization \mathbf{P} . The question remains as to when this approach is justified and what to do when it isn't. Unexcited systems interacting with radiation far from the system resonances can often be treated purely classically. The response of the system near and at resonance often deviates substantially from the classical descriptions. Since the laser itself and many applications involve systems near atomic (or molecular) resonances, we need to study them with the aid of quantum mechanics.

In preparation for this study, this chapter reviews some of the highlights of quantum mechanics paying particular attention to topics relevant to the interaction of radiation with matter. Section 3.1 introduces the wave function for an abstract quantum system, discusses the wave function's probabilistic interpretation, its role in the calculation of expectation values, and its equation of motion (the Schrödinger equation). Expansions of the wave function in various bases, most notably in terms of energy eigenstates, are presented and used to convert the Schrödinger partial differential equation into a set of ordinary differential equations. The Dirac notation is reviewed and used to discuss the state vector and how the state vector is related to the wave function. System time evolution is revisited with a short review of the Schrödinger, Heisenberg and interaction pictures.

In Chaps. 4–12, we are concerned with the interaction of classical electromagnetic fields with simple atomic systems. Section 3.2 lays the foundations for these chapters by discussing wave functions for atomic systems and studying their evolution under the influence of applied perturbations. Time dependent perturbation theory and the rotating wave approximation are used to predict this evolution in limits for which transitions are unlikely. The Fermi Golden Rule is derived. Section 3.3 deals with a particularly simple atomic model, the two-level atom subject to a resonant or nearly resonant classical field. We first discuss the nature of the electric-dipole interaction and then use the Fermi Golden Rule to derive Einstein's A and B coefficients for spontaneous and stimulated emission. We then relax the assumption that the interaction is weak and derive the famous Rabi solution.

In Chaps. 13–19, we discuss interactions for which the electromagnetic field as well as the atoms must be quantized. In particular, Chap. 13 shows

that electromagnetic field modes are described mathematically by simple harmonic oscillators. In addition, these oscillators can model the polarization of certain kinds of media, such as simple molecular systems. In preparation for such problems, Sect. 3.4 quantizes the simple harmonic oscillator. The section writes the appropriate Hamiltonian in terms of the annihilation and creation operators, and derives the corresponding energy eigenstates.

This chapter is concerned with the quantum mechanics of single systems in pure states. Discussions of mixtures of systems including the decay phenomena and excitation mechanisms that occur in lasers and their applications are postponed to Chap. 4 on the density matrix.

3.1 Review of Quantum Mechanics

According to the postulates of quantum mechanics, the best possible knowledge about a quantum mechanical system is given by its wave function $\psi(\mathbf{r}, t)$. Although $\psi(\mathbf{r}, t)$ itself has no direct physical meaning, it allows us to calculate the *expectation values* of all *observables* of interest. This is due to the fact that the quantity

$$\psi(\mathbf{r}, t)^* \psi(\mathbf{r}, t) d^3r$$

is the probability of finding the system in the volume element d^3r . Since the system described by $\psi(\mathbf{r}, t)$ is assumed to exist, its probability of being somewhere has to equal 1. This gives the normalization condition

$$\int \psi(\mathbf{r}, t)^* \psi(\mathbf{r}, t) d^3r = 1, \quad (3.1)$$

where the integration is taken over all space.

An observable is represented by a *Hermitian* operator $O(\mathbf{r})$ and its expectation value is given in terms of $\psi(\mathbf{r}, t)$ by

$$(O)(t) = \int d^3r \psi^*(\mathbf{r}, t) O(\mathbf{r}) \psi(\mathbf{r}, t). \quad (3.2)$$

Experimentally this expectation value is given by the average value of the results of many measurements of the observable O acting on identically prepared systems. The accuracy of the experimental value for (O) typically depends on the number of measurements performed. Hence enough measurements should be made so that the value obtained for (O) doesn't change significantly when still more measurements are performed. It is crucial to note that the expectation value (3.2) predicts the average from many measurements; in general it is unable to predict the outcome of a single event with absolute certainty. This does not mean that quantum mechanics in other ways is unable to make some predictions about single events.

The reason observables, such as position, momentum, energy, and dipole moment, are represented by Hermitian operators is that the expectation values (3.2) must be real. Denoting by (ϕ, ψ) the inner or scalar product of two vectors ϕ and ψ , we say that a linear operator O is Hermitian if the equality

$$(\phi, O\psi) = (O\phi, \psi) . \quad (3.3)$$

holds for all ϕ and ψ . In this notation, (3.2) reads $(O) = (\psi, O\psi)$.

An important observable in the interaction of radiation with bound electrons is the electric dipole \mathbf{er} . This operator provides the bridge between the quantum mechanical description of a system and the polarization of the medium \mathbf{P} used as a source in Maxwell's equations for the electromagnetic field. According to (3.2) the expectation value of \mathbf{er} is

$$(\mathbf{er})(t) = \int d^3r \mathbf{r} e |\psi(\mathbf{r}, t)|^2 , \quad (3.4)$$

where we can move \mathbf{er} to the left of $\psi(\mathbf{r}, t)^*$ since the two commute (an operator like ∇ cannot be moved). Here we see that the dipole-moment expectation value has the same form as the classical value if we identify $\rho = e|\psi(\mathbf{r}, t)|^2$ as the charge density.

In nonrelativistic quantum mechanics, the evolution of $\psi(\mathbf{r}, t)$ is governed by the Schrödinger equation.

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = H\psi(\mathbf{r}, t) , \quad (3.5)$$

where H is the Hamiltonian for the system and $\hbar = 1.054 \times 10^{-34}$ joule-seconds is Planck's constant divided by 2π . The Hamiltonian of an unperturbed system, for instance an atom not interacting with light, is the sum of its kinetic and potential energies

$$H_0 = \frac{\hbar^2 \nabla^2}{2m} + V(\mathbf{r}) , \quad (3.6)$$

where m is its mass and $V(\mathbf{r})$ the potential energy. As we shall see shortly, in the coordinate representation we are considering here the momentum operator $\hat{\mathbf{p}}$ is expressed in terms of the system's position operator \mathbf{r} as

$$\hat{\mathbf{p}} = -i\hbar \nabla , \quad (3.7)$$

so that we recognize that the first term of the Hamiltonian (3.6) is nothing but the kinetic energy of the system. Note also the important relationship

$$[\hat{x}_i, \hat{p}_j] = i\hbar \delta_{ij} \quad (3.8)$$

where \hat{x}_i and \hat{p}_j are cartesian coordinates of the operators $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$ and $[\hat{a}, \hat{b}] \equiv \hat{a}\hat{b} - \hat{b}\hat{a}$ is the *commutator* between the operators \hat{a} and \hat{b} . Observables which satisfy this commutation relation are generally called *conjugate variables*.

In view of (3.7), we see that the Schrödinger equation (3.5) is a partial differential equation. The time and space dependencies in (3.5) separate for functions having the form

$$\psi_n(\mathbf{r}, t) = u_n(\mathbf{r})e^{-i\omega_n t} \quad (3.9)$$

for which the $u_n(\mathbf{r})$ satisfy the energy eigenvalue equation

$$Hu_n(\mathbf{r}) = E_n u_n(\mathbf{r}). \quad (3.10)$$

The eigenfunctions $u_n(\mathbf{r})$ can be shown to be orthogonal, and we take them to be normalized according to (3.1), so that they are then orthonormal,

$$\int u_n^*(\mathbf{r}) u_m(\mathbf{r}) d\mathbf{r} = \delta_{n,m} = \begin{cases} 1 & n = m \\ 0 & n \neq m \end{cases} \quad (3.11)$$

and complete

$$\sum_n u_n^*(\mathbf{r}) u_n(\mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}'), \quad (3.12)$$

where $\delta_{n,m}$ and $\delta(\mathbf{r} - \mathbf{r}')$ are the Kronecker and Dirac delta functions, respectively. The completeness relation (3.12) means that any function can be written as a superposition of the $u_n(\mathbf{r})$. Problem 3.1 shows that this definition is equivalent to saying that any wave function can be expanded in a complete set of states.

In particular the wave function $\psi(\mathbf{r}, t)$ itself can be written as the superposition of the $\psi_n(\mathbf{r}, t)$:

$$\psi(\mathbf{r}, t) = \sum_n C_n(t) u_n(\mathbf{r}) e^{-i\omega_n t}. \quad (3.13)$$

The expansion coefficients $C_n(t)$ are actually independent of time for problems described by a Hamiltonian satisfying the eigenvalue equation (3.10). We have nevertheless included a time dependence in anticipation of adding an interaction energy to the Hamiltonian. Such a modified Hamiltonian wouldn't quite satisfy (3.10), thereby causing the $C_n(t)$ to change in time.

Substituting (3.13) into the normalization condition (3.1) and using the orthonormality condition (10), we find

$$\sum_n |C_n|^2 = 1. \quad (3.14)$$

The $|C_n|^2$ can be interpreted as the probability that the system is in the n th energy state. The C_n are complex probability amplitudes and completely determine the wave function. To find the expectation value (3.2) in terms of the C_n , we substitute (3.13) into (3.2). This gives

$$\langle O \rangle = \sum_{n,m} C_n C_m^* O_{mn} e^{-i\omega_{nm} t}, \quad (3.15)$$

where the operator matrix elements O_{mn} are given by

$$O_{mn} = \int d^3r u_m^*(\mathbf{r}) O u_n(\mathbf{r}) , \quad (3.16)$$

and the frequency differences

$$\omega_{nm} = \omega_n - \omega_m . \quad (3.17)$$

Typically we consider the interaction of atoms with electromagnetic fields. To treat such interactions, we add the appropriate interaction energy to the Hamiltonian, that is

$$H = H_0 + V . \quad (3.18)$$

If we expand the wave function in terms of the eigenfunctions of the “unperturbed Hamiltonian” H_0 , rather than those of the total Hamiltonian H , the probability amplitudes $C_n(t)$ change in time. To find out just how, we substitute the wave function (3.13) and Hamiltonian (3.18) into Schrödinger’s equation (3.5) to find

$$\sum_n (\mathbf{k}\omega_n + V) C_n u_n(\mathbf{r}) e^{-i\omega_n t} = \sum_n (\mathbf{k}\omega_n C_n + i\mathbf{k}\dot{C}_n) u_n(\mathbf{r}) e^{-i\omega_n t} . \quad (3.19)$$

Cancelling the $\mathbf{k}\omega_n$ terms, changing the summation index n to m , multiplying through by $u_n^*(\mathbf{r}) \exp(i\omega_n t)$, and using the orthonormality property (3.11), we find the equation of motion for the probability amplitude $C_n(t)$

$$\dot{C}_n(t) = -\frac{i}{\hbar} \sum_m (n|V|m) e^{i\omega_{nm}t} C_m(t) , \quad (3.20)$$

where the matrix element

$$(n|V|m) = \int d^3r u_n^*(\mathbf{r}) V u_m(\mathbf{r}) . \quad (3.21)$$

Note that instead of the form (3.13), we can also expand the wave function $\psi(\mathbf{r}, t)$ as

$$\psi(\mathbf{r}, t) = \sum_n c_n(t) u_n(\mathbf{r}) , \quad (3.22)$$

for which the $\mathbf{k}\omega_n$ time dependence in (3.19) doesn’t cancel out. The $c_n(t)$ then obey the equation of motion

$$\dot{c}_n(t) = -i\omega_n c_n(t) - \frac{i}{\hbar} \sum_m (n|V|m) c_m(t) . \quad (3.23)$$

In terms of the c_n , the expectation value (3.2) becomes

$$\langle O \rangle = \sum_{n,m} c_n c_m^* O_{mn} . \quad (3.24)$$

Equation (3.20) and equivalently (3.23) shows how the probability amplitudes for the wave function written as a superposition of energy eigenfunctions changes in time. They are equivalent to the original Schrödinger equation (3.5), but are no longer concerned with the precise position dependence, which is already accounted for by the \mathbf{r} -dependence of the eigenfunctions $u_n(\mathbf{r})$. In particular if we're only concerned about how a system such as an atom absorbs energy from a light field, this development is completely described by the changes in the C_n or c_n .

The choice of using the relatively slowly varying C_n versus using the rapidly varying c_n is a matter of taste and convenience. The time dependence of the C_n is due to the interaction energy V alone, while that of the c_n is due to the total Hamiltonian H . To distinguish between the two, we say that the C_n are in the *interaction picture*, while the c_n are in the *Schrödinger picture*. We discuss this more formally at the end of this section.

Armed with (3.20) or (3.23), you can skip directly to Sect. 3.2, which shows how systems evolve in time due to various interactions. Before going ahead, we review the Dirac notation and some other aspects of the wave function and of its more abstract form, the state vector $|\psi\rangle$. This material is needed for our discussions involving quantized fields in Chaps. 13–19, and is useful in proving various properties of the density operator in Chap. 4.

Up to now we have used the so-called coordinate representation, where all operators, as well as the wave function, are expressed as functions of the system's position \mathbf{r} . Alternatively, one can work in a number of other representations, a rather common one being the momentum representation. Here, operators and wave functions are expressed as functions of the system's momentum \mathbf{p} . As we shall see shortly, one can transform the system's wave function from the coordinate to the momentum representation by a simple Fourier transformation of $\psi(\mathbf{r}, t)$,

$$\varphi(\mathbf{p}, t) = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3r \psi(\mathbf{r}, t) e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar}. \quad (3.25)$$

Here $\varphi(\mathbf{p}, t)$ describes the same dynamical state as $\psi(\mathbf{r}, t)$. It doesn't make any difference in principle which representation we choose to use, and as we see with the \hat{C}_n , we sometimes don't have to worry about the coordinate representation at all.

We now turn to a discussion of a general formalism which does away with the explicit use of representations, and allows us to switch from one to another representation, when it is desirable.

Dirac Notation

The formalism that permits one to achieve this goal is the Dirac notation. Roughly speaking Dirac's formulation is analogous to using vectors instead of coordinates. The notation has an additional advantage in that one can label

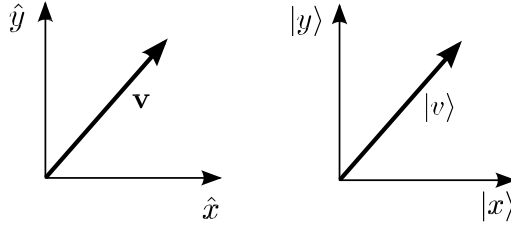


Fig. 3.1. A two dimensional vector written in ordinary vector notation and in Dirac notation

the basis vectors much more conveniently than with ordinary vector notation. We start our discussion with a comparison between ordinary notation for a vector in a two-dimensional space and Dirac's version. As shown in Fig. 3.1, a vector \mathbf{v} can be expanded as

$$\mathbf{v} = v_x \hat{\mathbf{x}} + v_y \hat{\mathbf{y}}, \quad (3.26)$$

where $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are unit vectors along the x and y axes, respectively. In Dirac notation, this reads

$$|v\rangle = v_x |x\rangle + v_y |y\rangle, \quad (3.27)$$

the component v_x given in ordinary vector notation by the dot product $\hat{\mathbf{x}} \cdot \mathbf{v}$ is given in Dirac notation by

$$v_x = \langle x | v \rangle. \quad (3.28)$$

The Dirac vector $|v\rangle$ is called a “ket” and the vector $\langle v|$ a “bra”, which come from calling the inner product (3.28) a “bra c ket”. With the notation (3.28), (3.27) reads

$$|v\rangle = |x\rangle \langle x | v \rangle + |y\rangle \langle y | v \rangle. \quad (3.29)$$

This immediately gives the identity diadic (outer product of two vectors)

$$|x\rangle \langle x| + |y\rangle \langle y| = I. \quad (3.30)$$

Equations (3.27, 3.30) can be immediately generalized to many dimensions as in

$$|v\rangle = \sum_n |n\rangle \langle n | v \rangle, \quad (3.31)$$

$$I = \sum_n |n\rangle \langle n|, \quad (3.32)$$

where the $\{|n\rangle\}$ are a complete orthonormal set of vectors, i.e., a basis. The inner products $\langle n | v \rangle$ are the expansion coefficients of the vector $|v\rangle$ in this basis. The bra $\langle n|$ is the adjoint of the ket $|n\rangle$ and the expansion coefficients have the property

$$(k|v) = (v|k)^* . \quad (3.33)$$

Unlike the real spaces of usual geometry, quantum mechanics works in a complex vector space called a Hilbert space, where the expansion coefficients are in general complex.

The basis $\{|n\rangle\}$ is discrete. Alternatively, we can expand vectors in terms of the coordinate basis $\{|\mathbf{r}\rangle\}$ which like the $\{|n\rangle\}$ forms a complete basis, albeit a continuous one. For such a situation we need to use continuous summations in (3.31, 3.32), that is, integrals. For example, the identity operator of (3.32) can be expanded as

$$I = \int d^3r |\mathbf{r}\rangle \langle \mathbf{r}| . \quad (3.34)$$

One major advantage of the bra and ket notation is that you can label the vectors with as many letters as desired. For example, you could write $|r\theta\phi\rangle$ in place of $|\mathbf{r}\rangle$.

The vector of primary interest in quantum mechanics is the state vector $|\psi(t)\rangle$. The wave function is actually the expansion coefficient of $|\psi\rangle$ in the coordinate basis

$$|\psi\rangle = \int d^3r |\mathbf{r}\rangle \langle \mathbf{r}|\psi\rangle , \quad (3.35)$$

where the wave function

$$\psi(\mathbf{r}, t) = \langle \mathbf{r}|\psi\rangle . \quad (3.36)$$

Hence the state vector $|\psi\rangle$ is equivalent to the wave function $\psi(\mathbf{r}, t)$, but doesn't explicitly display the coordinate dependence.

Instead of using the position expansion of (3.35), we can expand the state vector in the discrete basis $\{|n\rangle\}$ as

$$|\psi\rangle = \sum_n c_n |n\rangle . \quad (3.37)$$

The most common basis to use consists of the eigenstates of the unperturbed Hamiltonian operator H_0 . For this basis, the expansion coefficients c_n are just those in (3.22), and the energy eigenfunctions are related to the eigenvectors by

$$u_n(\mathbf{r}) = \langle \mathbf{r}|n\rangle . \quad (3.38)$$

A useful trick in transforming from one basis to another is to think of the vertical bar as an identity operator expanded either as in (3.32) or as in (3.34). Using the form of (3.34) in (3.37) on both sides of the equation along with (3.38), we recover (3.22).

The expectation value of the operator O is given in terms of the state vector by

$$\langle O \rangle(t) = \langle \psi(t) | O | \psi(t) \rangle . \quad (3.39)$$