Contents

Chapter 10

Nonlinear Optics

Version 1210.1.K, 10 January 2013 *Please send comments, suggestions, and errata via email to kip@caltech.edu or on paper to Kip Thorne, 350-17 Caltech, Pasadena CA 91125*

Box 10.1
Reader's Guide

- This chapter depends substantially on Secs. 7.2, 7.3 and 7.7.1 of Chap. 7, Geometric Optics.
- Sec. 10.6, on wave-wave mixing, is an important foundation for Chap. 23 on the nonlinear dynamics of plasmas, and (to a lesser extent) for the discussions of solitary waves (solitons) in Secs. 16.3 and 23.6. Nothing else in this book relies substantially on this chapter.

10.1 Overview

Communication technology is undergoing a revolution, and computer technology may do so soon — a revolution in which the key devices used (e.g., switches and communication lines) are changing from radio and microwave frequencies to optical frequencies. This revolution has been made possible by the invention and development of lasers (most especially semiconductor diode lasers) and other technology developments such as nonlinear media whose polarization P_i is a nonlinear function of the applied electric field, $P_i = \epsilon_0(\chi_{ij}E^j + d_{ijk}E^jE^k + \chi_{ijkl}E^jE^kE^l + \cdots)$. In this chapter we shall study lasers, nonlinear media, and various *nonlinear optics* applications that are based on them.

Most courses in elementary physics idealize the world as linear. From the simple harmonic oscillator to Maxwell's equations to the Schrödinger equation, most all the elementary physical laws one studies are linear, and most all the applications one studies make use of this linearity. In the real world, however, nonlinearities abound, creating such phenomena as avalanches, breaking ocean waves, holograms, optical switches, and neural networks; and

in the past three decades nonlinearities and their applications have become major themes in physics research, both basic and applied. This chapter, with its exploration of nonlinear effects in optics, serves as a first introduction to some fundamental nonlinear phenomena and their present and future applications. In later chapters, we shall revisit some of these phenomena, and shall meet others, in the context of fluids (Chaps. 16 and 17), plasmas (Chap. 23), and spacetime curvature (Chaps. 25–28).

Since highly coherent and monochromatic laser light is one of the key foundations on which modern nonlinear optics has been built, we shall begin in Sec. 10.2 with a review of the basic physics principles that underlie the laser: the pumping of an active medium to produce a molecular population inversion, and the stimulated emission of radiation from the inverted population. Then we shall briefly describe the wide variety of lasers now available, how a few of them are pumped, and the characteristics of their light. As an important example (crucial for the optical frequency combs of Sec. 9.4.3 above), we shall give details about mode-locked lasers.

In Sec. 10.3, we shall meet our first example of an application of nonlinear optics: holography. In the simplest variant of holography, a three-dimensional, monochromatic image of an object is produced by a two step process: recording a hologram of the image, and then passing coherent light through the hologram to reconstruct the image. We shall analyze this recording and reconstruction and then shall describe a few of the many variants of holography now available and some of their practical applications.

Holography differs from more modern nonlinear optics applications in not being a real-time process. Real-time processes have been made possible by nonlinear media and other new technologies. In Sec. 10.4, we shall study an example of a real-time, nonlinear-optics process: phase conjugation of light by a phase-conjugating mirror (though we will delay a detailed discussion of how such mirrors work until Sec. 10.8.2). In Sec. 10.4, we shall also see how phase conjugation can be used to counteract the distortion of images and signals by media though which they travel.

In Sec. 10.5, we shall introduce nonlinear media and formulate Maxwell's equations for waves propagating through such media. As an example, we shall very briefly discuss electro-optic effects, where a slowly changing electric field modulates the optical properties of a nonlinear crystal, thereby modulating light waves that propagate through it. Then in Sec. 10.6, we shall develop a detailed description of how such a nonlinear crystal couples two optical waves to produce a new, third wave — so-called three-wave mixing. Three-wave mixing has many important applications in modern technology. In Sec. 10.7, we describe and analyze several: frequency doubling (e.g., in a green laser pointer), optical parametric amplification of signals, and driving light into a squeezed state—e.g., the squeezed vacuum of quantum electrodynamics.

In an isotropic medium, three-wave mixing is suppressed, but a new, fourth wave can be produced by three incoming waves. In Sec. 10.8, we describe and analyze this four-wave mixing, and how it is used in phase conjugate mirrors and produces unwanted problems in the optical fibers widely used to transmit internet, television and telephone signals.

These topics just scratch the surface of the exciting field of nonlinear optics, but they will give the reader an overview and some major insights into this field, and into nonlinear phenomena in the physical world.

10.2 Lasers

10.2.1 Basic Principles of the Laser

In quantum mechanics one identifies three different types of interaction of light with material systems (atoms, molecules, atomic nuclei, electrons, ...): (*i*) *Spontaneous emission*, in which a material system in an excited state spontaneously drops into a state of lesser excitation and emits a photon in the process. (*ii*) *Absorption*, in which an incoming photon is absorbed by a material system, exciting it. (*iii*) *Stimulated emission*, in which a material system, initially in some excited state, is "tickled" by passing photons and this tickling stimulates it to emit a photon of the same sort (in the same state) as the photons that tickled it.

As peculiar as stimulated emission may seem at first sight, it in fact is easily understood and analyzed classically. It is nothing but "negative absorption": In classical physics, when a light beam with electric field $E = \Re[Ae^{i(kz-\omega t+\varphi)}]$ travels through an absorbing medium, its real amplitude A decays exponentially with the distance propagated, $A \propto e^{-\mu z/2}$ (corresponding to an energy-flux decay $F \propto e^{-\mu z}$), while its frequency ω, wave number k, and phase φ remain very nearly constant. For normal materials, the absorption rate $\mu = F^{-1}dF/dz$ is positive and the energy lost goes ultimately into heat. However, one can imagine a material with an internally stored energy that amplifies a passing light beam. Such a material would have a negative absorption rate, $\mu < 0$, and correspondingly the amplitude of the passing light would grow with the distance traveled, $A \propto e^{+|\mu|z/2}$, while its frequency, wave number, and phase would remain very nearly constant. Such materials do exist; they are called "active media" and their amplification of passing waves is called "stimulated emission."

This elementary, classical description of stimulated emission is equivalent to the quantum mechanical description in the domain where the stimulated emission is strong: the domain of large photon occupation numbers $\eta \gg 1$ (which, as we learned in Sec. 3.2.5, is the domain of classical waves).

The classical description of stimulated emission takes for granted the existence of an active medium. To understand the nature of such a medium, we must turn to quantum mechanics:

As a first step toward such understanding, consider a beam of monochromatic light with frequency ω that impinges on a collection of molecules (or atoms or charged particles) that

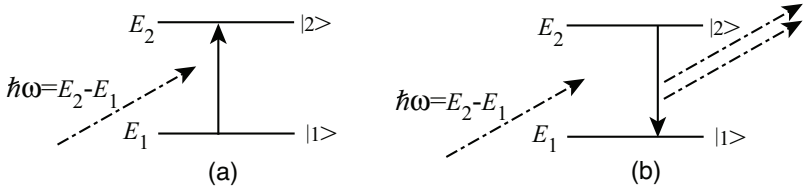

Fig. 10.1: (a) *Photon Absorption:* A photon with energy $\hbar\omega = E_2 - E_1$ excites a molecule from its ground state, with energy E_1 to an excited state with energy E_2 (as depicted by an energy-level diagram). (b) *Stimulated Emission:* The molecule is initially in its excited state, and the incoming photon stimulates it to deexcite into its ground state, emitting a photon identical to the incoming one.

are all in the same quantum mechanical state $|1\rangle$. Suppose the molecules have a second state $|2\rangle$ with energy $E_2 = E_1 + \hbar\omega$. Then the light will resonantly excite the molecules from their initial state $|1\rangle$ to the higher state $|2\rangle$, and in the process photons will be absorbed (Fig. 10.1a). The strength of the interaction is proportional to the beam's energy flux F. Stated more precisely, the rate of absorption of photons is proportional to the number flux of photons in the beam $dn/dAdt = F/\hbar\omega$ and thence proportional to F, in accord with the classical description of absorption.

Suppose, next, that when the light beam first arrives, the atoms are all in the higher state $|2\rangle$ rather than the lower state $|1\rangle$. There will still be a resonant interaction, but this time the interaction will deexcite the atoms, with an accompanying emission of photons (Fig. 10.1b). As in the absorption case, the strength of the interaction is proportional to the flux of the incoming beam, i.e., the rate of emission of new photons is proportional to the number flux of photons that the beam already has, and thence also proportional to its energy flux F. A quantum mechanical analysis shows that the photons from this *stimulated emission* come out in the same quantum state as is occupied by the photons of the incoming beam (Bose-Einstein statistics: photons, being bosons, like to congregate in the same state). Correspondingly, when viewed classically, the beam's flux will be amplified at a rate proportional to its initial flux, with no change of its frequency, wave number, or phase.

In nature molecules usually have their energy levels populated in accord with the laws of statistical (thermodynamic) equilibrium. Such thermalized populations, as we saw at the end of Sec. 4.4.1, entail a ratio $N_2/N_1 = \exp[-(E_2 - E_1)/k_B T] < 1$ for the number N_2 of molecules in state $|2\rangle$ to the number N_1 in state $|1\rangle$. Here T is the molecular temperature, and for simplicity it is assumed that the states are nondegenerate. Since there are more molecules in the lower state $|1\rangle$ than the higher one $|2\rangle$, an incoming light beam will experience more absorption than stimulated emission.

On the other hand, occasionally in nature and often in the laboratory a collection of molecules develops a "population inversion" in which $N_2 > N_1$. The two states can then be thought of as having a negative temperature with respect to each other. Light propagating through population-inverted molecules will experience more stimulated emission than absorption; i.e., it will be amplified. The result is "light amplification by stimulated emission," or "laser" action.

This basic principle underlying the laser has been known since the early years of quantum mechanics, but only in the 1950s did physicists succeed in designing, constructing, and operating real lasers. The first proposals for practical devices were made, independently, in the U.S. by Weber (1953) and Gordon, Zeiger, and Townes (1954), and in Russia by Basov and Prokhorov (1954, 1955). The first successful construction and operation of a laser was by Gordon, Zeiger, and Townes (1954, 1955), and soon thereafter by Basov and Prokhorov —though these first lasers actually used radiation not at optical frequencies but rather at microwave frequencies (based on a population inversion of ammonia molecules[1]) and thus was called a *maser*. The first optical frequency laser, one based on a population inversion of chromium ions in a ruby crystal, was constructed and operated by Maiman (1960).

[1]For the basic principles of the ammonia maser, see, e.g., Chap. 9 of *The Feynman Lectures on Physics* (Feynman, Leighton, and Sands 1965).

The key to laser action is the population inversion. Population inversions are incompatible with thermodynamic equilibrium; thus, to achieve them, one must manipulate the molecules in a nonequilibrium way. This is usually done by some concrete variant of the process shown in the energy level diagram of Fig. 10.2. Some sort of pump mechanism (to be discussed in the next section) rapidly excites molecules from the ground state into some group of "absorption" states. The molecules then decay rapidly from the absorption states into the state $|2\rangle$, which is metastable (i.e., has a long lifetime against spontaneous decay), so the molecules linger there. The laser transition is from state $|2\rangle$ into state $|1\rangle$. Once a molecule has decayed into state $|1\rangle$, it quickly decays on down to the ground state and then may be quickly pumped back up into the absorption states. This is called "four-level pumping". It is called, instead, "three-level pumping" if state $|1\rangle$ is the ground state.

If the pump acts suddenly and briefly, this process will produce a temporary population inversion of states $|2\rangle$ and $|1\rangle$, with which an incoming, weak burst of "seed" light can interact to produce a burst of amplification. The result is a pulsed laser. If the pump acts continually, the result may be a permanently maintained population inversion with which continuous seed light can interact to produce continuous-wave laser light.

As the laser beam travels through the active medium (the population-inverted molecules), its flux F builds up with distance z as $dF/dz = F/\ell_o$, so $F(z) = F_o e^{z/\ell_o}$. Here F_o is the initial flux, and $\ell_o \equiv 1/|\mu|$, the e-folding length, depends on the strength of the population inversion and the strength of the coupling between the light and the active medium. Typically ℓ_o is so long that strong lasing action cannot be achieved by a single pass through the active medium. In this case, the lasing action is enhanced by placing the active medium inside a Fabry-Perot cavity (Fig. 10.3 and Sec. 9.4.3). The length L of the cavity is adjusted so the lasing transition frequency $\omega = (E_2 - E_1)/\hbar$ is an eigenfrequency of the cavity. The lasing action then excites a standing wave mode of the cavity, from which the light leaks out through one or both cavity mirrors. If $\mathcal{F}$ is the cavity's finesse [approximately the average number of times a photon bounces back and forth inside the cavity before escaping through a mirror; cf. Eq. (9.39)], then the cavity increases the distance that typical photons travel through the active medium by a factor $\sim \mathcal{F}$, thereby increasing the energy flux of the light output by a factor $\sim e^{\mathcal{F}L/\ell_o}$.

Typically many modes of the Fabry Perot cavity are excited, so the laser's output is

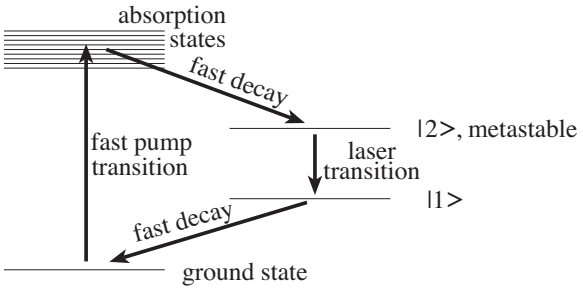

Fig. 10.2: The mechanism for creating the population inversion that underlies laser action. The horizontal lines and band represent energy levels of a molecule, and the arrows represent transitions in which the molecules are excited by pumping or decay by emission of photons.

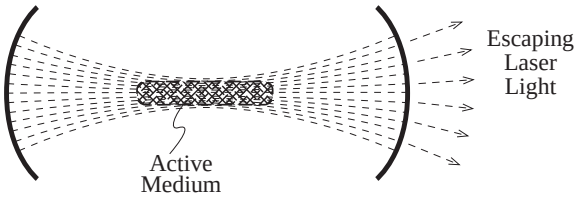

Fig. 10.3: The use of a Fabry-Perot cavity to enhance the interaction of the light in a laser with its active medium.

multimode and contains a mixture of polarizations. When a single mode and polarization are desired, the polarization is made pure by oblique optical elements at the ends of the laser that transmit only one, and all the modes except one are removed from the output light by a variety of techniques, for example filtering with a second Fabry Perot cavity (Sec. 9.4.3).

For an ideal laser (one, e.g., with a perfectly steady pump maintaining a perfectly steady population inversion that in turn maintains perfectly steady lasing), the light comes out in the most perfectly classical state that quantum mechanics allows. This state, called a *quantum mechanical coherent state*, has a perfectly sinusoidally oscillating electric field on which is superimposed the smallest amount of noise (the smallest wandering of phase and amplitude) allowed by quantum mechanics: the noise of quantum electrodynamical vacuum fluctuations. The value of the oscillations' well defined phase is determined by the phase of the seed field from which the coherent state was built up by lasing. Real lasers have additional noise due to a variety of practical factors, but nevertheless, their outputs are usually highly coherent, with long coherence times.

10.2.2 Types of Lasers, Their Performances and Applications

Lasers can have continuous, near monochromatic output, or can be pulsed. Their active media can be liquids, gases (ionized or neutral), or solids (semiconductors, glasses, or crystals; usually carefully doped with impurities). Lasers can be pumped by radiation (e.g. from a flash tube), by atomic collisions that drive the lasing atoms into their excited states, by nonequilibrium chemical reactions , or by electric fields associated with electric currents (e.g. in semiconductor diode lasers that can be powered by ordinary batteries and are easily modulated for optical communication).

Lasers can be made to pulse by turning the pump on and off, by mode-locked operation (next subsection), or by Q-switching (turn off the lasing, e.g. by inserting into the Fabry Perot cavity an electro-optic material that absorbs light until the pump has produced a huge population inversion, and then suddenly applying an electric field to the absorber, which makes it transparent and restores the lasing).

Laser pulses can be as short as a few fs (thus enabling experimental investigations of fast chemical reactions) and they can carry as much as 20,000 J with duration a few ps and pulse power $\sim 10^{16}$ W (at the U.S. National Ignition Facility for controlled fusion).

The most powerful continuous laser in the U.S. was the Mid-Infrared Advanced Chemical Laser (MIRACL), developed by the Navy to shoot down missiles and satellites, with ~ 1 MW power in a 14×14 cm beam lasting ~ 70 s. Continuous CO_2 lasers with powers ~ 3 kW

are used industrially for cutting and welding metal.

The beam from a Q-switched CO_2 laser with ~ 1 GW power can be concentrated into a region with transverse dimensions as small as one wavelength ($\sim 1\mu$m) yielding a local energy flux of 10^{21} W m^{-2}, an rms magnetic field strength of ~ 3 kT, an electric field ~ 1 TV m^{-1}, and an electrical potential difference across a wavelength ~ 1 MeV. It then should not be surprising that high power lasers can create electron-positron pair plasmas!

For most applications, large power is irrelevant or undesireable, but high frequency stability (a long coherence time) is often crucial. By locking the laser frequency to an optical-frequency atomic transition (e.g. in the Al^+ atomic clock; footnote 8 in Fig. 6.11), one can achieve a frequency stability $\Delta f/f \sim 10^{-17}$, i.e. $\Delta f \sim 3$ mHz, for hours or longer, corresponding to coherence times of ~ 100 sec and coherence lengths of $\sim 3 \times 10^8$ km. By locking the frequency to a mode of a physically highly stable Fabry Perot cavity (e.g. PDH locking, Sec. 9.4.3), stabilities have been achieved as high as $\Delta f/f \sim 10^{-16}$ for times ~ 1 hr in a physically solid cavity (the superconducting cavity stabilized oscillator), and $\Delta f/f \sim 10^{-21}$ for a few ms in LIGO's 4km long cavity with freely hanging mirrors and sophisticated seismic isolation (Sec. 9.5).

When first invented, lasers were called "a solution looking for a problem." Now they permeate everyday life and high technology. Examples are supermarket bar-code readers, laser pointers, DVD players, eye surgery, laser printers, laser cutting and welding, laser gyroscopes (which are standard on commercial aircraft), laser-based surveying, Raman spectroscopy, laser fusion, optical communication, optically based computers, holography, maser amplifiers, and atomic clocks.

10.2.3 Ti:Sapp Mode-Locked Laser

As a concrete example of a modern, specialized laser, we shall discuss the Titanium Sapphire (Ti:Sapp) mode-locked laser that is used to generate the optical frequency comb that we described in Sec. 9.4.3. Recall that this laser's light must be concentrated in a very short pulse that travels back and forth between its Fabry Perot mirrors, unchanged. The pulse is made from phase-locked (Gaussian) modes of the optical cavity that extend over a huge frequency band, $\Delta\omega \sim \omega$. Among other things, this mode-locked laser will illustrate the use of an optical nonlinearity called the Kerr effect, whose underlying physics we shall describe later in this chapter (Sec. 10.8.3.

As we discussed in Sec. 9.4.3, this mode-locked laser must (i) more strongly amplify modes with high energy flux than with low (this pushes the light into the short, high-flux pulse), and (ii) its (Gaussian) modes must have a group velocity V_g that is independent of frequency over the frequency band $\Delta\omega \sim \omega$ (this enables the pulse to stay short rather than disperse).

Figure 10.4 illustrates the Ti:sapph laser that achieves this. The active medium is a sapphire crystal doped with titanium ions. This medium exhibits the optical Kerr effect, which means that its index of refraction $\mathfrak{n}$ is a sum of two terms, one independent of the light's energy flux; the other proportional to the flux [Eq. (**??**. The flux-dependent term slows the light's speed near the beam's center and thereby focuses the beam, making its cross section smaller. A circular aperture attenuates large light beams but not small. As a

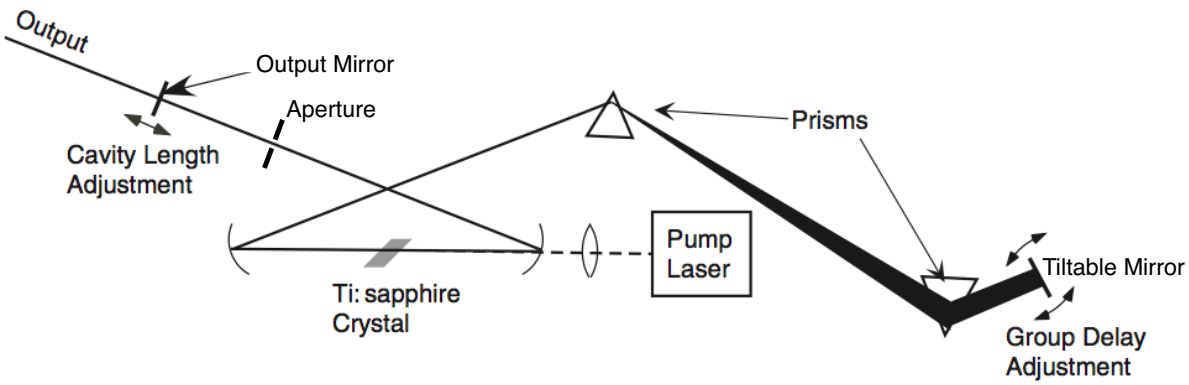

Fig. 10.4: The Ti:Sapph mode locked laser. From Cundiff (2002).

result, the lasing is stronger the smaller the beam, which means the higher its flux. This drives the lasing light into the desired short, high-flux pulse.

The Ti:sapph crystal has a group velocity that increases with frequency. The two prisms and tiltable mirror compensate this. The first prism bends low-frequency light more strongly than high, assuring that the high-frequency light traverses more glass in the second prism and is thereby slowed. By adjusting the mirror tilt, one adjusts the amount of slowing so as to keep the round-trip-averaged phase velocity the same at high frequencies as at low. The laser then happily generates the multimode, high-intensity, short-pulsed light whose output is the optical frequency comb of Sec. 9.4.3. For additional detail see, e.g., Cundiff (2002).

****

EXERCISES

Exercise 10.1 *Challenge: Nuclear Powered X-Ray Laser*

A device much ballyhooed in America during the reign of Ronald Reagan, but never built, was a futuristic, super-powerful X-ray laser pumped by a nuclear explosion. As part of Reagan's Strategic Defense Initiative ("Star Wars"), this laser was supposed to shoot down Soviet missiles.

How would *you* design a nuclear powered X-ray laser? The energy for the pump comes from a nuclear explosion that you set off in space above the earth. You want to use that energy to create a population inversion in an active medium that will lase at X-ray wavelengths; and you want to focus the resulting X-ray beam onto an intercontinental ballistic missile that is rising out of the earth's atmosphere. What would you use for the active medium? How would you guarantee that a population inversion is created in the active medium? How would you focus the resulting X-ray beam? (*Note*: This is a highly nontrivial exercise, intended more as a stimulus for thought than as a test of one's understanding of things taught in this book.)

****

10.3 Holography

Holography is an old and well-explored example of nonlinear optics—an example in which the nonlinear interaction of light with itself is produced not in real time, but rather by means of a recording followed by a later readout.

By contrast with ordinary photography (Fig. 10.5), which produces a colored, 2-dimensional image of 3-dimensional objects, holography (Figs. 10.6 and 10.8 below) normally produces a monochromatic 3-dimensional image of 3-dimensional objects. Roughly speaking, the two processes contain the same amount of information, two items at each location in the image. For ordinary photography, they are the energy flux and color; for holography, the energy flux and phase of monochromatic light.

It is the phase, lost from an ordinary photograph but preserved in holography, that carries the information about the third dimension. Our brain deduces the distance to a point on an object from the difference in the directions of propagation of the point's light as it arrives at our two eyes. Those propagation directions are encoded in the light as variations of the phase with transverse location [see, e.g., the point-spread function for a thin lens, Eq. (8.29)].

In an ordinary photograph (Fig. 10.5), white light scatters off an object, with different colors scattering at different strengths. The resulting colored light is focused through a lens to form a colored image on a photographic plate or a CCD. The plate or CCD records the color and energy flux at each point or pixel in the focal plane, thereby producing the ordinary photograph.

In holography, one records a hologram with flux and phase information (Fig. 10.6 below), and one then uses the hologram to reconstruct the 3-dimensional, monochromatic, holographic image (Fig. 10.8 below).

10.3.1 Recording a Hologram

Consider, first, the recording of the hologram. Monochromatic, linearly polarized plane-wave light with electric field

$$E = \Re[\psi(x, y, z)e^{-i\omega t}] \,, \tag{10.1}$$

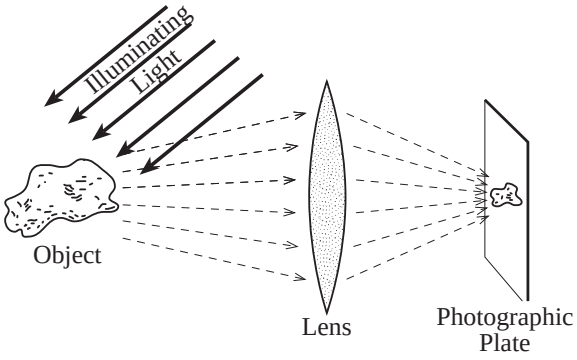

Fig. 10.5: Ordinary photography.

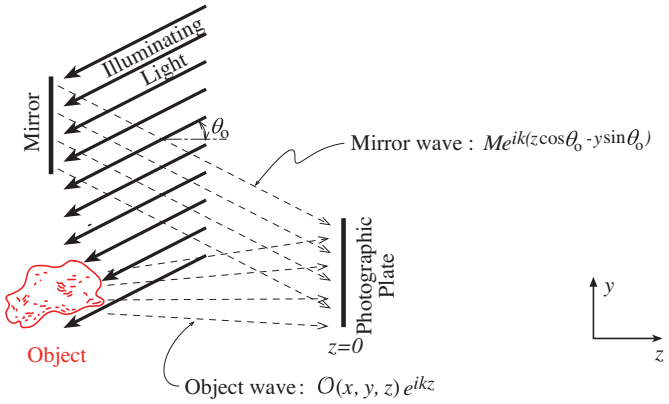

Fig. 10.6: Recording a hologram.

angular frequency ω and wave number $k = \omega/c$, illuminates the object and also a mirror as shown in Fig. 10.6. The light must be spatially coherent over the entire region of mirror plus object. The propagation vector $\mathbf{k}$ of the illuminating light lies in the y–z plane, at some angle θ_o to the z axis, and the mirror lies in the x–y plane. The mirror reflects the illuminating light, producing a so-called *reference beam*, which we shall call the *mirror wave*:

$$\psi_{\rm mirror} = Me^{ik(z\cos\theta_o - y\sin\theta_o)} \ , \tag{10.2}$$

where M is a real constant. The object (shown red) scatters the illuminating light, producing a wave that propagates in the z direction toward the recording medium (a photographic plate for concreteness). We shall call this the *object wave* and shall denote it

$$\psi_{\rm object} = \mathcal{O}(x, y, z)e^{ikz} \ . \tag{10.3}$$

It is the slowly varying complex amplitude $\mathcal{O}(x, y, z)$ of this object wave that carries the three-dimensional, but monochromatic, information about the object's appearance, and it thus is this $\mathcal{O}(x, y, z)$ that will be reconstructed in the second step of holography.

In the first step (Fig. 10.6), the object wave propagates along the z-direction to the photographic plate at $z = 0$, where it interferes with the mirror wave to produce the transverse pattern of energy flux

$$\begin{aligned} F(x, y) &\propto |\mathcal{O} + Me^{-iky\sin\theta_o}|^2 \\ &= M^2 + |\mathcal{O}(x, y, z = 0)|^2 + \mathcal{O}(x, y, z = 0)Me^{iky\sin\theta_o} + \mathcal{O}^*(x, y, z = 0)Me^{-iky\sin\theta_o}. \end{aligned} \tag{10.4}$$

(Here and throughout this chapter a * denotes complex conjugation.) The plate is blackened at each point in proportion to this flux. The plate is then developed and a positive or negative print (it doesn't matter which because of Babinet's principle) is made on a transparent sheet of plastic or glass. This print, the *hologram*, has a transmissivity as a function of x and y that is proportional to the flux distribution (10.4):

$$\mathfrak{t}(x, y) \propto M^2 + |\mathcal{O}(x, y, z = 0)|^2 + \mathcal{O}(x, y, z = 0)Me^{iky\sin\theta_o} + \mathcal{O}^*(x, y, z = 0)Me^{-iky\sin\theta_o} \ . \tag{10.5}$$

In this transmissivity we meet our first example of nonlinearity: $\mathfrak{t}(x, y)$ is a nonlinear superposition of the mirror wave and the object wave. Stated more precisely, the superposition is not a linear sum of wave fields, but instead is a sum of products of one wave field with the complex conjugate of another wave field. A further nonlinearity will arise in the reconstruction of the holographic image, Eq. (10.7) below.

Figure 10.7 shows an example. Figure 10.7a is an ordinary photograph of an object, 10.7b is a hologram of the same object, and 10.7c is a blow-up of a portion of that hologram. The object is not at all recognizable in the hologram because the object wave $\mathcal{O}$ was not focused to form an image at the plane of the photographic plate. Rather, light from each region of the object was scattered to and recorded by all regions of the photographic plate. Nevertheless, the plate contains the full details of the scattered light $\mathcal{O}(x, y, z = 0)$, including its phase. That information is recorded in the piece $M(\mathcal{O}e^{iky\sin\theta_o} + \mathcal{O}^* e^{-iky\sin\theta_o}) = 2M\,\Re(\mathcal{O}e^{iky\sin\theta_o})$ of the hologram's transmissivity. This piece oscillates sinusoidally in the y direction with wavelength $2\pi/k\sin\theta_o$; and the amplitude and phase of its oscillations are modulated by the object wave $\mathcal{O}(x, y, z = 0)$. Those modulated oscillations show up clearly when one magnifies the hologram (Fig. 10.7c); they make the hologram into a sort of diffraction grating, with the object wave $\mathcal{O}(x, y, z = 0)$ encoded as variations of the darkness and spacings of the grating lines.

What about the other pieces of the transmissivity (10.5), which superpose linearly on the diffraction grating? One piece, $\mathfrak{t} \propto M^2$, is spatially uniform and thus has no effect except to make the lightest parts of the hologram slightly grey rather than leaving it absolutely transparent (since this hologram is a negative rather than a positive). The other piece, $\mathfrak{t} \propto |\mathcal{O}|^2$, is the flux of the object's unfocussed, scattered light. It produces a greying and whitening of the hologram (Fig. 10.7b) that varies on lengthscales long compared to the grating's wavelength $2\pi/k\sin\theta_o$, and that thus blots out the diffraction grating a bit here and there, but does not change the amplitude or phase of the grating's modulation.

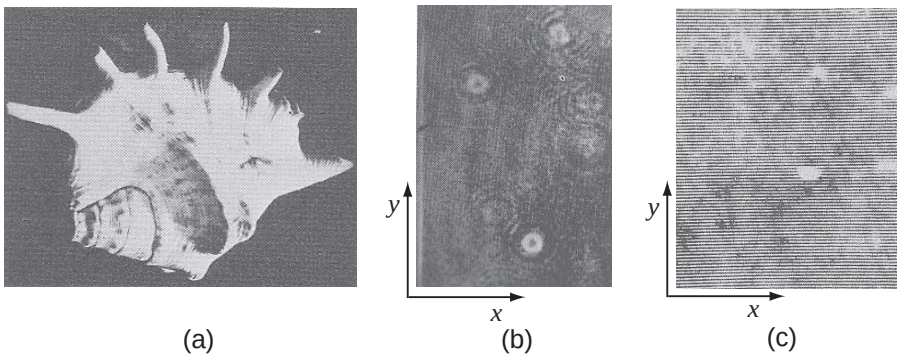

Fig. 10.7: (a) Ordinary photograph of an object. (b) Hologram of the same object. (c) Magnification of the hologram. [Photographs courtesy Professor R.S. Sirohi, adapted from Fig. 8.3 of Ghatak and Thyagarajan (1978), and from Fig. 21.3 of Ghatak (2010)]

10.3.2 Reconstructing the 3-Dimensional Image from a Hologram

To reconstruct the object's 3-dimensional wave, $\mathcal{O}(x,y,z)e^{ikz}$, one sends through the hologram monochromatic, plane-wave light identical to the mirror light used in making the hologram; Fig. 10.8. If, for pedagogical simplicity, we place the hologram at the same location $z=0$ as was previously occupied by the photographic plate, then the incoming light has the same form (10.2) as the original mirror wave, but with an amplitude that we shall denote as R corresponding to the phrase *reference beam* that is used to describe this incoming light:

$$\psi_{\rm reference} = Re^{ik(z\cos\theta_o - y\sin\theta_o)} \ . \tag{10.6}$$

In passing through the hologram at $z=0$, this reference beam is partially absorbed and partially transmitted. The result, immediately upon exiting from the hologram, is a "reconstructed" light-wave field whose value and normal derivative are given by [cf. Eq. (10.5)]

$$\begin{aligned}
\psi_{\rm reconstructed}\Big|_{z=0} \equiv \mathcal{R}(x,y,z=0) &= \mathfrak{t}(x,y)Re^{-iky\sin\theta_o} \\
&= \left[M^2 + |\mathcal{O}(x,y,z=0)|^2\right] Re^{-iky\sin\theta_o} \\
&\quad + MR\mathcal{O}(x,y,z=0) \\
&\quad + MR\mathcal{O}^*(x,y,z=0)e^{-i2ky\sin\theta_o} \ ; \\
\psi_{\rm reconstructed\ ,z}\Big|_{z=0} \equiv \mathcal{Z}(x,y,z=0) &= ik\cos\theta_o\mathcal{R}(x,y,z=0).
\end{aligned} \tag{10.7}$$

This field and normal derivative act as initial data for the subsequent evolution of the reconstructed wave. Note that the field and derivative, and thus also the reconstructed wave,

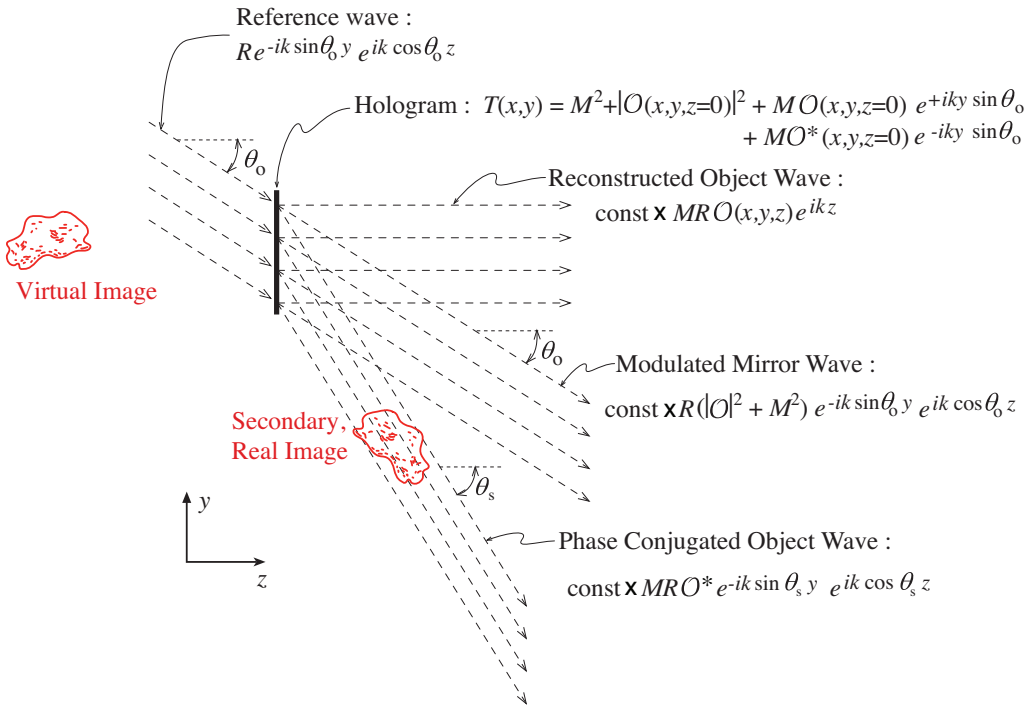

Fig. 10.8: Reconstructing the holographic image from the hologram. Note that $\sin\theta_s = 2\sin\theta_o$.

are triply nonlinear: each term in Eq. (10.7) is a product of (i) the original mirror wave M used to construct the hologram or the original object wave $\mathcal{O}$, times (ii) $\mathcal{O}^*$ or $M^* = M$, times (iii) the reference wave R that is being used in the holographic reconstruction.

The evolution of the reconstructed wave beyond the hologram (at $z > 0$) can be computed by combining the initial data (10.7) for $\psi_{\rm reconstructed}$ and $\psi_{\rm reconstructed\ ,z}$ at $z = 0$ with the Helmholtz-Kirchhoff formula (8.4); see Exs. 10.2 and 10.6. From the four terms in the initial data, Eq. (10.7) [which arise from the four terms in the hologram's transmissivity $\mathfrak{t}(x, y)$, Eq. (10.5)], the reconstruction produces four wave fields; see Fig. 10.8. The direction of propagation of each of these waves can easily be inferred from the vertical spacing of its phase fronts along the outgoing face of the hologram, or equivalently from the relation $\partial\psi_{\rm reconstructed}/\partial y = ik_y\psi_{\rm reconstructed} = -ik\sin\theta\psi$, where θ is the angle of propagation relative to the horizontal z direction. Since, immediately in front of the hologram, $\psi_{\rm reconstructed} = \mathcal{R}$, the propagation angle is

$$\sin\theta = \frac{\partial\mathcal{R}/\partial y}{-ik\mathcal{R}} \ . \tag{10.8}$$

Comparing with Eqs. (10.5) and (10.7), we see that the first two, slowly spatially varying terms in the transmissivity, $\mathfrak{t} \propto M^2$ and $T \propto |\mathcal{O}|^2$, both produce waves that propagate in the same direction as the reference wave, $\theta = \theta_o$. This combined wave has an uninteresting, smoothly and slowly varying energy-flux pattern.

The two diffraction-grating terms in the hologram's transmissivity produce two interesting waves. One, arising from $\mathfrak{t} \propto \mathcal{O}(x, y, z = 0)Me^{iky\sin\theta_o}$ [and produced by the $MR\mathcal{O}$ term of the initial conditions (10.7)], is *precisely the same object wave* $\psi_{\rm object} = \mathcal{O}(x, y, z)e^{ikz}$ *(aside from overall amplitude) as one would have seen while making the hologram if one had replaced the photographic plate by a window and looked through it.* This object wave, carrying [encoded in $\mathcal{O}(x, y, z)$] the famous holographic image with full 3-dimensionality, propagates in the z-direction, $\theta = 0$.

The transmissivity's second diffraction-grating term, $\mathfrak{t} \propto \mathcal{O}^*(x, y, z = 0)Me^{-iky\sin\theta_o}$, acting via the $MR\mathcal{O}^*$ term of the initial conditions (10.7), gives rise to a *secondary* wave which [according to Eq. (10.8)] propagates at an angle θ_s to the z-axis, where

$$\sin\theta_s = 2\sin\theta_o \ . \tag{10.9}$$

(If $\theta_o > 30^{\rm o}$, then $2\sin\theta_o > 1$ which means θ_s cannot be a real angle, and there will be no secondary wave.) *This secondary wave, if it exists, carries an image that is encoded in the complex conjugate* $\mathcal{O}^*(x, y, z = 0)$ *of the transverse (i.e., x, y) part of the original object wave.* Since complex conjugation of an oscillatory wave just reverses the sign of the wave's phase, this wave in some sense is a *"phase conjugate" of the original object wave.*

When one recalls that the electric and magnetic fields that make up an electromagnetic wave are actually real rather than complex, and that we are using complex wave fields to describe electromagnetic waves only for mathematical convenience, one then realizes that this phase conjugation of the object wave is actually a highly nonlinear process. There is no way, by linear manipulations of the real electric and magnetic fields, to produce the phase-conjugated wave from the original object wave.

In Sec. 10.4 we shall develop in detail the theory of *phase-conjugated waves*, and in Ex. 10.6, we shall relate our holographically constructed secondary wave to that theory.

As we shall see, our secondary wave is not quite the same as the "phase-conjugated object wave," but it *is* the same aside from some distortion along the y direction and a change in propagation direction. More specifically: *If one looks into the object wave with one's eyes (i.e. if one focuses it onto one's retinas), one sees the original object in all its three-dimensional glory, though single colored, sitting behind the hologram at the object's original position* (shown red in Fig. 10.8). Because the image one sees is behind the hologram, it is called a *virtual image. If, instead, one looks into the secondary wave with one's eyes* (i.e. if one focuses it onto one's retinas), *one sees the original three-dimensional object, sitting in front of the hologram but turned inside out and distorted* (also shown red in the figure). For example, if the object is a human face, the secondary image looks like the interior of a mask made from that human face, with distortion along the y direction. Because this secondary image appears to be in front of the hologram, it is called a *real image*—even though one can pass one's hands through it and feel nothing but thin air.

Other Types of Holography, and Applications

There are many variants on the basic holographic technique that we have described in Figs. 10.6–10.8. These include, among others:

Phase holography. Here, instead of darkening the high-flux regions of the hologram as in photography, one produces a phase-shifting screen, whose phase shift (due to thickening of the hologram's material) is proportional to the incoming flux. Such a phase hologram transmits more of the reference-wave light than a standard, darkened hologram, thus making a brighter reconstructed image.

Volume holography. Here the hologram is a number of wavelengths deep rather than being just two-dimensional. For example, it could be made from a thick photographic emulsion, in which the absorption length for light is longer than the thickness. Such a hologram has a three-dimensional grating structure (grating "surfaces" rather than grating "lines"), with two consequences: When one reconstructs the holographic image from it in the manner of Fig. 10.8, (i) *the third dimension of the grating suppresses the secondary wave while enhancing the (desired) object wave so more power goes into it;* and (ii) the reference wave's incoming angle θ_o must be controlled much more precisely, as modest errors suppress the reconstructed object wave. This second consequence enables one to record multiple images in a volume hologram, each using its own angle θ_o for the illuminating light and reference wave.

Reflection holography. Here one reads out the hologram by reflecting light off of it rather than transmitting light through it, and the hologram's diffraction grating produces a three-dimensional holographic image by the same process as in transmission; see Ex. 10.3

White-light holography. Here the hologram is recorded with monochromatic light as usual, but it is optimized for reading out with white light. Even for the simple two-dimensional hologram of Fig. 10.8, if one sends in white light at the angle θ_o, one will get a three-dimensional object wave: The hologram's grating will diffract various wavelengths in various directions. In the direction of the original object wave (the horizontal direction in Fig. 10.8), one gets a 3-dimensional reconstructed image of the same color as was used when constructing the hologram. When one moves away from that direction (vertically in Fig. 10.8), one sees the color of the 3-dimensional image continuously change; Ex. 10.2c. White-light reflection holograms are used on credit cards, and money as impediments to counterfeiting, and have even been used on postage stamps.

Computational holograms. Just as one can draw two-dimensional pictures numerically, pixel-by-pixel, so one can also create and modify holograms numerically, then read them out optically.

Full-color holography. A full-color holographic image of an object can be constructed by superposing three monochromatic holographic images with the three primary colors, red, green and blue. One way to achieve this is to construct a single volume hologram using illuminating light from red, green, and blue laser beams, each arriving from a different 2-dimensional direction $\boldsymbol{\theta}_o$. Each beam produces a diffraction grating in the hologram with a different orientation and with spatial wave number corresponding to the beam's color. The thee-dimensional image can then be reconstructed using three white-light reference waves, one from each of the original three directions $\boldsymbol{\theta}_o$. The hologram will pick out of each beam the appropriate primary color, and produce the desired three overlapping images, which the eye will interpret as having approximately the true color of the original object.

Holographic interferometry. One can observe changes in the shape of a surface at the $\sim \mu$m level by constructing two holograms, one of the original surface and the other of the changed surface, and then interfering the reconstructed light from the two holograms. This is called holographic interferometry, and it is used, e.g., to observe small strains and vibrations of solid bodies — for example, sonic vibrations of a guitar in Fig. 10.9.

Holographic lenses. Instead of designing a hologram to reconstruct a 3-dimensional image, one can design it to manipulate light beams in most any way one wishes. Such a hologram is called a holographic lense. As a simple example (Ex. 10.4c), one can construct a holographic lens that splits one beam into two and focuses each of the two beams onto a different spot. Holographic lenses are widely used in everyday technology, e.g., to read bar codes in supermarket checkouts, and to read the information off CDs, DVDs, and BDs (Ex. 10.5).

Future applications. Major applications of holography that are under development include (i) dynamically changing volume holograms for three-dimensional movies (which, of course,

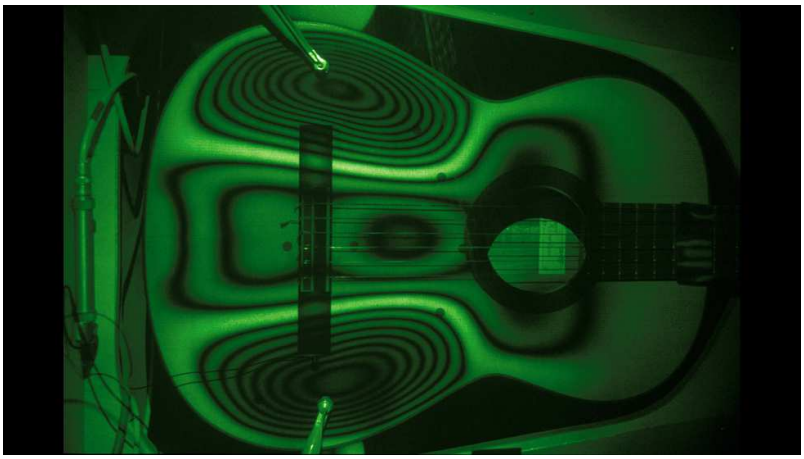

Fig. 10.9: In and out vibrations of a guitar body visualized via holographic interferometry with green light. The dark and bright curves are a contour map, in units of the green light's wavelength, of the amplitude of vibration. Courtesy Bernard Richardson, University of Cardiff. [DON'T YET HAVE PERMISSION]

will require no eye glasses), and (ii) voume holograms for storage of large amounts of data — up to terabytes cm^{-3}.

****

EXERCISES

Exercise 10.2 *Derivation and Problem: The Holographically Reconstructed Wave*

(a) Use the Helmholtz-Kirchhoff integral (8.4) or (8.6) to compute all four pieces of the holographically reconstructed wave field. Show that the piece generated by

$$\mathfrak{t} \propto \mathcal{O}(x, y, z \ = \ 0) M e^{iky \sin\theta_o}$$

is the same (aside from overall amplitude) as the field $\psi_{\text{object}} = \mathcal{O}(x, y, z)e^{-i\omega t}$ that would have resulted, when making the hologram (Fig. 10.6), had the mirror wave been absent and the photographic plate been replaced by a window. Show that the other pieces have the forms and propagation directions indicated heuristically in Fig. 10.8.

(b) We shall examine the secondary wave, generated by $\mathfrak{t} \propto M\mathcal{O}^* e^{-iky \sin\theta_o}$, in Ex. 10.6.

(c) Suppose that plane-parallel white light is used in the holographic reconstruction of Fig. 10.8. Derive an expression for the direction in which one sees the object's three-dimensional image have a given color (or equivalently wave number). Assume that the original hologram was made with green light and $\theta_o = 45$ degrees. What are the angles at which one sees the image as violet, green, yellow and red?

Exercise 10.3 *Problem: Recording a Reflection Hologram*

How would you record a hologram if you want to read it out via reflection? Draw diagrams illustrating this, similar to Figs. 10.6 and 10.8. [Hint: The mirror wave and object wave can impinge on the photographic plate from either side; it's your choice.]

Exercise 10.4 *Example: Holographic Lens to Split and Focus a Light Beam*

A holographic lens, like any other hologram, can be described by its transmissivity $\mathfrak{t}(x, y)$.

(a) What $\mathfrak{t}(x, y)$ will take a reference wave, impinging from the θ_o direction (as in Fig. 10.8) and produce from it a primary object wave that converges on the spot $(x, y, z) = (0, 0, d)$? [Hint: consider, at the hologram's plane, a superposition of the incoming mirror wave and the point spread function (8.28), which represents a beam that diverges from a point source, and phase conjugate the point spread function so it converges to a point instead of diverging.]

(b) Draw a contour plot of this lens's transmissivity $\mathfrak{t}(x, y)$. Notice the resemblance to the Fresnel zone plate of Sec. 8.4.4. Explain the connection of the two, paying attention to how the holographic lens changes when one alters the chosen angle θ_o of the reference wave.

(c) What $\mathfrak{t}(x, y)$ will take a reference wave, impinging from the θ_o direction, and produce from it a primary wave that splits in two, with equal light powers converging onto the spots $(x, y, z) = (-a, 0, d)$ and $(x, y, z) = (+a, 0, d)$?

Exercise 10.5 *** Problem: Compact Disks, DVDs and Blue Ray Disks*

Information on compact disks (CDs), on DVDs and on Blue Ray disks (BDs) is recorded and read out using holographic lenses, but it is not stored holographically. Rather, it is stored in a linear binary code consisting of pits and no-pits (for 0 and 1) along a narrow spiraling track. In each successive generation, the laser light has been pushed to a shorter wavelength (λ = 760 nm for CDs, 650 nm for DVDs, 405nm for BDs), and in each generation, the efficiency of the information storage has been improved. In CDs, the information is stored in a single holographic layer on the surface of the disk; in DVDs and BDs, it is usually stored in a single layer, but can also be stored in two layers, one above the other, though with a modest price in access time and efficiency of storage.

(a) Explain why one can expect to record in a disk's recording layer, at the very most, (close to) one bit of information per square wavelength of the recording light.

(b) The actual storage capacities are up to 900 MB for CDs, 4.7GB for DVDs, and 25 GB for BDs. How efficient are each of these technologies relative to the maximum of part (a)?

(c) Estimate the number of volumes of the Encyclopedia Britannica that can be recorded on a CD, on a DVD and on a BD.

****

10.4 Phase-Conjugate Optics

Nonlinear optical techniques make it possible to phase conjugate an optical wave in real time, by contrast with holography where the phase conjugation requires recording a hologram and then reconstructing the wave later. In this section, we shall explore the properties of phase conjugated waves of any sort (light, sound, plasma waves, ...), and in the next section we shall discuss technology by which real-time phase conjugation is achieved for light.

The basic ideas and foundations for phase conjugation of waves were laid in Moscow, Russia by Boris Yakovovich Zel'dovich[2] and his colleagues (1972) and at Caltech by Amnon Yariv (1977).

Phase conjugation is the process of taking a *monchromatic* wave

$$\boxed{\Psi_{\rm O} = \Re[\psi(x, y, z)e^{-i\omega t}] = \frac{1}{2}(\psi e^{-i\omega t} + \psi^* e^{+i\omega t})}\,, \tag{10.10a}$$

[2]Zel'dovich is the famous son of a famous Russian/Jewish physicist, Yakov Borisovich Zel'dovich, who with Andrei Dmitrievich Sakharov fathered the Soviet hydrogen bomb and then went on to become a dominant figure internationally in astrophysics and cosmology.

and from it constructing the wave

$$\boxed{\Psi_{\rm PC} = \Re[\psi^*(x,y,z)e^{-i\omega t}] = \frac{1}{2}(\psi^* e^{-i\omega t} + \psi e^{+i\omega t})} \,. \tag{10.10b}$$

Notice that the phase conjugated wave $\Psi_{\rm PC}$ is obtainable from the original wave $\Psi_{\rm O}$ by *time reversal*, $t \to -t$. This has a number of important consequences. One is that $\Psi_{\rm PC}$ propagates in the opposite direction to $\Psi_{\rm O}$. Others are explained most clearly with the help of a *phase-conjugating mirror*:

Consider a wave $\Psi_{\rm O}$ with spatial modulation (i.e., a wave that carries a picture or a signal of some sort). Let the wave propagate in the z-direction (rightward in Fig. 10.10), so

$$\psi = \mathcal{A}(x,y,z)e^{i(kz-\omega t)} \,, \quad \text{where} \quad \mathcal{A} = Ae^{i\varphi} \tag{10.11}$$

is a complex amplitude whose modulus A and phase φ change slowly in x, y, z (slowly compared to the wave's wavelength $\lambda = 2\pi/k$). Suppose that this wave propagates through a time-independent medium with slowly varying physical properties [e.g. a dielectric medium with slowly varying index of refraction $\mathfrak{n}(x,y,z)$]. These slow variations will distort the wave's complex amplitude as it propagates. The wave equation for the real, classical field $\Psi = \Re[\psi e^{-i\omega t}]$ will have the form $\mathcal{L}\Psi - \partial^2\Psi/\partial t^2 = 0$, where $\mathcal{L}$ is a real spatial differential operator that depends on the medium's slowly varying physical properties. This wave equation implies that the complex field ψ satisfies

$$\mathcal{L}\psi + \omega^2\psi = 0 \,. \tag{10.12}$$

This is the evolution equation for the wave's complex amplitude.

Let the distorted, rightward propagating wave $\Psi_{\rm O}$ reflect off a mirror located at $z = 0$. If the mirror is a phase-conjugating one, then very near it (at z near zero) the reflected wave will have the form

$$\Psi_{\rm PC} = \Re[\mathcal{A}^*(x,y,z=0)e^{i(-kz-\omega t)}] \,, \tag{10.13}$$

while if it is an ordinary mirror, then the reflected wave will be

$$\Psi_{\rm R} = \Re[\pm\mathcal{A}(x,y,z=0)e^{i(-kz-\omega t)}] \,. \tag{10.14}$$

(Here the sign, plus or minus, depends on the physics of the wave. For example, if Ψ is the transverse electric field of an electromagnetic wave and the mirror is a perfect conductor, the sign will be minus to guarantee that the total electric field, original plus reflected, vanishes at the mirror's surface.)

These two waves, the phase-conjugated one $\Psi_{\rm PC}$ and the ordinary reflected one $\Psi_{\rm R}$, have very different surfaces of constant phase (*phase fronts*): The phase of the incoming wave $\Psi_{\rm O}$ [Eq. (10.11)] as it nears the mirror ($z = 0$) is $\varphi + kz$, so (taking account of the fact that φ is slowly varying), the surfaces of constant phase are $z = -\varphi(x,y,z=0)/k$. Similarly, the phase of the wave $\Psi_{\rm R}$ [Eq. (10.14)] reflected from the ordinary mirror is $\varphi - kz$, so its surfaces of constant phase near the mirror are $z = +\varphi(x,y,z=0)/k$, which are reversed from those of the incoming wave as shown in the upper right of Fig. 10.10. Finally, the phase of the wave $\Psi_{\rm PC}$ [Eq. (10.13)] reflected from the phase-conjugating mirror is $-\varphi - kz$, so its

surfaces of constant phase near the mirror are $z = -\varphi(x, y, z = 0)/k$, which are the same as those of the incoming wave (lower right of Fig. 10.10), even though the two waves are propagating in opposite directions.

The phase fronts of the original incoming wave and the phase conjugated wave are the same not only near the phase conjugating mirror; they are the same everywhere. More specifically, as the phase-conjugated wave $\Psi_{\rm PC}$ propagates away from the mirror [near which it is described by Eq. (10.13)], the propagation equation (10.12) forces it to evolve in such a way as to remain always the phase conjugate of the incoming wave:

$$\Psi_{\rm PC} = \Re[\mathcal{A}^*(x, y, z)e^{-ikz}e^{-i\omega t}] \ . \tag{10.15}$$

This should be obvious from the fact that, because the differential operator $\mathcal{L}$ in the propagation equation (10.12) for $\psi(x, y, z) = \mathcal{A}e^{ikz}$ is real, $\psi^*(x, y, z) = \mathcal{A}^*e^{-ikz}$ will satisfy this propagation equation whenever $\psi(x, y, z)$ does.

This fact that the reflected wave $\Psi_{\rm PC}$ remains always the phase conjugate of the incoming wave $\Psi_{\rm O}$ means that *the distortions put onto the incoming wave, as it propagates rightward through the inhomogeneous medium, get removed from the phase conjugated wave as it propagates back leftward;* see Fig. 10.10.

This removal of distortions has a number of important applications. One is for image transmission in optical fibers. Normally when an optical fiber is used to transmit an optical image, the transverse spatial variations $\mathfrak{n}(x, y)$ of the fiber's index of refraction (which are required to hold the light in the fiber; Ex. 7.8) distort the image somewhat. The distortions can be eliminated by using a sequence of identical segments of optical fibers separated

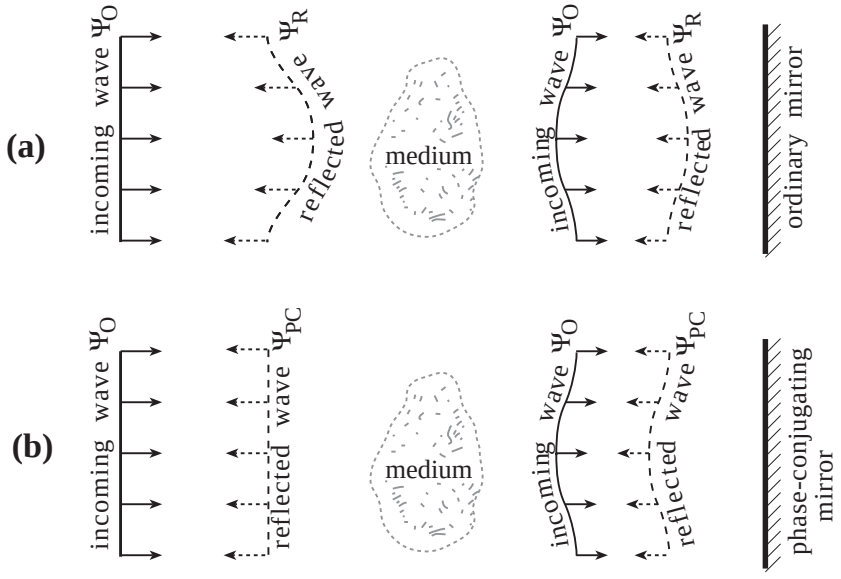

Fig. 10.10: A rightward propagating wave and the reflected wave produced by (a) an ordinary mirror and (b) a phase-conjugating mirror. In both cases the waves propagate through a medium with spatially variable properties, which distorts their phase fronts. In case (a) the distortion is reinforced by the second passage through the variable medium; in case (b) the distortion is removed by the second passage.

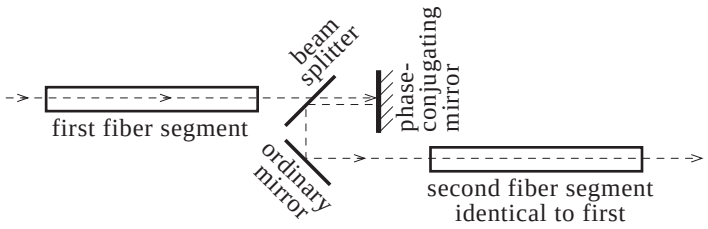

Fig. 10.11: The use of a phase-conjugating mirror in an optical transmission line to prevent the fiber from distorting an optical image. The distortions put onto the image as it propagates through the first segment of fiber are removed during propagation through the second segment.

by phase-conjugating mirrors (Fig. 10.11). A few other applications include (*i*) real time holography, (*ii*) removal of phase distortions in Fabry-Perot cavities by making one of the mirrors a phase conjugating one, with a resulting improvement in the shape of the beam that emerges from the cavity, (*iii*) devices that can memorize an optical image and compare it to other images, (*iv*) the production of squeezed light (Ex. 10.16), and (*v*) improved focusing of laser light for laser fusion.

As we shall see in the next section, phase conjugating mirrors rely crucially on the sinusoidal time evolution of the wave field; they integrate up that sinusoidal evolution coherently over some timescale $\hat{\tau}$ (typically microseconds to nanoseconds) in order to produce the phase conjugated wave. Correspondingly, if an incoming wave varies on timescales τ long compared to this $\hat{\tau}$ (e.g., if it carries a temporal modulation with bandwidth $\Delta\omega \sim 1/\tau$ small compared to $1/\hat{\tau}$), then the wave's temporal modulations will *not* get time reversed by the phase conjugating mirror. For example, if the wave impinging on a phase conjugating mirror has a frequency that is ω_a initially, and then gradually, over a time τ, increases to $\omega_b = \omega_a + 2\pi/\tau$, then the phase conjugated wave will *not* emerge from the mirror with frequency ω_b first and ω_a later. Rather, it will emerge with ω_a first and ω_b later (same order as for the original wave). When the incoming wave's temporal variations are fast compared to the mirror's integration time, $\tau \ll \hat{\tau}$, the mirror encounters a variety of frequencies during its integration time, and ceases to function properly. Thus, *even though phase conjugation is equivalent to time reversal in a formal sense, a phase conjugating mirror cannot time reverse a temporal signal. It only time reverses monochromatic waves (which might carry a spatial signal).*

****

EXERCISES

Exercise 10.6 *Derivation and Example: The Secondary Wave in Holography*
Consider the secondary wave generated by $\mathfrak{t} \propto M\mathcal{O}^* e^{-iky\sin\theta_o}$ in the holographic reconstruction process of Fig. 10.8, Eqs. 10.7 and Ex. 10.2.

(a) Assume, for simplicity, that the mirror and reference waves propagate nearly perpendicular to the hologram, so $\theta_o \ll 90^\circ$ and $\theta_s \simeq 2\theta_o \ll 90^\circ$; but assume that θ_s is still large enough that fairly far from the hologram the object wave and secondary waves

separate cleanly from each other. Then, taking account of the fact that the object wave field has the form $\mathcal{O}(x, y, z)e^{ikz}$, show that the secondary wave is the phase conjugated object wave defined in Sec. 10.4, except that it is propagating in the $+z$ direction rather than $-z$, i.e. it has been reflected through the $z = 0$ plane. Then use this, and the discussion of phase conjugation in Sec. 10.4, to show that the secondary wave carries an image that resides in front of the hologram and is turned inside out, as discussed near the end of Sec. 10.3. Show, further, that if θ_o is not $\ll 90^o$ degrees (but is $< 30^o$, so θ_s is a real angle and the secondary image actually exists), then the secondary image is changed by a distortion along the y direction. What is the nature of the distortion, a squashing or a stretch?

(b) Suppose that a hologram has been made, with $\theta_o < 30^o$. Show that it is possible to perform image reconstruction with a modified reference wave (different from Fig, 10.8) in such a manner that the secondary, phase-conjugated wave emerges precisely perpendicular to the hologram and undistorted.

****

10.5 Maxwell's Equations in a Nonlinear Medium; Nonlinear Dielectric Susceptibilities; Electro-Optic Effects

In nonlinear optics, one is often concerned with media that that are electrically polarized with *polarization* (electric dipole moment per unit volume) $\mathbf{P}$, but that have no free charges or currents and are unmagnetized. In such a medium, the charge and current densities associated with the polarization are

$$\rho_P = -\boldsymbol{\nabla} \cdot \mathbf{P} \ , \quad \mathbf{j}_P = \frac{\partial \mathbf{P}}{\partial t} \ , \tag{10.16a}$$

and Maxwell's equations in SI units take the form

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{\rho_{\mathbf{P}}}{\epsilon_{\mathbf{0}}}, \quad \boldsymbol{\nabla} \cdot \mathbf{B} = \mathbf{0} \ , \quad \boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial \mathbf{t}} \ , \quad \boldsymbol{\nabla} \times \mathbf{B} = \mu_{\mathbf{0}} \left(\mathbf{j}_{\mathbf{P}} + \epsilon_{\mathbf{0}} \frac{\partial \mathbf{E}}{\partial \mathbf{t}} \right) , \tag{10.16b}$$

which should be familiar. When rewritten in terms of the electric displacement vector

$$\mathbf{D} \equiv \epsilon_0 \mathbf{E} + \mathbf{P} \ , \tag{10.17}$$

these Maxwell equations take following the alternative form

$$\boldsymbol{\nabla} \cdot \mathbf{D} = 0 \ , \quad \boldsymbol{\nabla} \cdot \mathbf{B} = 0 \ , \quad \boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial \mathbf{t}} \ , \quad \boldsymbol{\nabla} \times \mathbf{B} = \mu_{\mathbf{0}} \frac{\partial \mathbf{D}}{\partial \mathbf{t}} \ , \tag{10.18}$$

which should also be familiar. By taking the curl of the third Maxwell equation (10.16b), using the relation $\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\mathbf{E} = -\nabla^2\mathbf{E} + \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{E})$, and combining with the time derivative of the fourth Maxwell equation (10.16b) and with $\epsilon_0\mu_0 = 1/c^2$ and $\mathbf{j}_P = \partial\mathbf{P}/\partial t$, we obtain the following wave equation for the electric field, sourced by the medium's polarization:

$$\boxed{\nabla^2\mathbf{E} - \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{E}) = \frac{1}{c^2}\frac{\partial^2(\mathbf{E}+\mathbf{P}/\epsilon_0)}{\partial t^2}} \ . \tag{10.19}$$

If the electric field is sufficiently weak and the medium is homogeneous and isotropic (the case treated in most textbooks on electromagnetic theory), the polarization $\mathbf{P}$ is proportional to the electric field: $\mathbf{P} = \epsilon_0\chi_0 E_j$, where χ_0 is the medium's electrical susceptibility. In this case the medium does not introduce any nonlinearities into Maxwell's equations, the right side of Eq. (10.19) becomes $[(1+\chi_0)/c^2]\partial\mathbf{E}/\partial t^2$, the divergence of (10.19) implies that the divergence of $\mathbf{E}$ vanishes, and therefore (10.19) becomes the standard dispersionless wave equation

$$\nabla^2\mathbf{E} - \frac{\mathfrak{n}^2}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2} = 0 \ , \quad \text{with} \quad \mathfrak{n}^2 = 1+\chi_0 \ . \tag{10.20}$$

In many dielectric media, however, a strong electric field can produce a polarization that is nonlinear in the field. In such "nonlinear media," the general expression for the (real) polarization in terms of the (real) electric field is

$$\boxed{P_i = \epsilon_0(\chi_{ij}E_j + 2d_{ijk}E_jE_k + 4\chi_{ijkl}E_jE_kE_l + \ldots)} \ . \tag{10.21}$$

Here χ_{ij}, the linear susceptibility, is proportional to the 3-dimensional metric, $\chi_{ij} = \chi_0 g_{ij} = \chi_0\delta_{ij}$, if the medium is isotropic (i.e., if all directions in it are equivalent), but otherwise is tensorial; and the d_{ijk} and χ_{ijkl} are nonlinear susceptibilities, referred to as *second-order* and *third-order* because of the two and three $\mathbf{E}$'s that multiply them in Eq. (10.21). The normalizations used for these second- and third-order susceptibilities differ from one researcher to another: sometimes the factor ϵ_0 is omitted in Eq. (10.21); occasionally the factors of 2 and 4 are omitted. A compressed 2-index notation is sometimes used for the components of d_{ijk}; see Box 10.2 below.

With $\mathbf{P}$ given by Eq. (10.21), the wave equation (10.19) becomes

$$\boxed{\nabla^2\mathbf{E} - \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{E}) - \frac{1}{c^2}\boldsymbol{\epsilon}\cdot\frac{\partial^2\mathbf{E}}{\partial t^2} = \frac{1}{c^2\epsilon_0}\frac{\partial^2\,\mathbf{P}^{\rm NL}}{\partial t^2} \ , \quad \text{where} \quad \epsilon_{ij} = \delta_{ij} + \chi_{ij}} \tag{10.22a}$$

and $\mathbf{P}^{\rm NL}$ is the nonlinear part of the polarization:

$$\boxed{P_i^{\rm NL} = \epsilon_0(2d_{ijk}E_jE_k + 4\chi_{ijkl}E_jE_kE_l + \ldots)} \ . \tag{10.22b}$$

When $\mathbf{P}^{\rm NL}$ is strong enough to be important and a monochromatic wave at frequency ω enters the medium, the nonlinearities lead to harmonic generation—i.e., the production of secondary waves with frequencies 2ω, 3ω, ...; see below. As a result, an electric field in the medium cannot oscillate at just one frequency, and each of the electric fields in

expression (10.22b) for the nonlinear polarization must be a sum of pieces with different frequencies. Because *the susceptibilities can depend on frequency*, this means that, when using expression (10.21), one sometimes must break P_i and each E_i up into its frequency components and use different values of the susceptibility to couple the different frequencies together. For example, one of the terms in Eq. (10.22b) will become

$$P_i^{(4)} = 4\epsilon_0 \chi_{ijkl}\, E_j^{(1)}\, E_k^{(2)}\, E_l^{(3)} \ , \tag{10.23}$$

where $E_j^{(n)}$ oscillates at frequency ω_n, $P_i^{(4)}$ oscillates at frequency ω_4, and χ_{ijkl} depends on the four frequencies $\omega_1, \ldots, \omega_4$. Although this is complicated in the general case, in most practical applications, resonant coupling (or equivalently energy and momentum conservation for photons) guarantees that only a single set of frequencies is important, and the resulting analysis simplifies substantially. See, e.g., Sec. 10.6.1 below.

Because all the tensor indices on the susceptibilities except the first index get contracted into the electric field in expression (10.21), we are free to (and it is conventional to) define the susceptibilities as symmetric under interchange of any pair of indices that does not include the first. When [as has been tacitly assumed in Eq. (10.21)] there is no hysteresis in the medium's response to the electric field, the energy density of interaction between the polarization and the electric field is

$$\boxed{U = \epsilon_0 \left(\frac{\chi_{ij} E_i E_j}{2} + \frac{2 d_{ijk} E_i E_j E_k}{3} + \frac{4 \chi_{ijkl} E_i E_j E_k E_l}{4} + \cdots \right)} \ , \tag{10.24a}$$

and the polarization is related to this energy of interaction, in Cartesian coordinates, by

$$\boxed{P_i = \frac{\partial U}{\partial E_i}} \ , \tag{10.24b}$$

which agrees with Eq. (10.21) providing *the susceptibilities are symmetric under interchange of all pairs of indices, including the first.* We shall assume such symmetry.[3] *If the crystal is isotropic (as will be the case if it has cubic symmetry and reflection symmetry), then each of its tensorial susceptibilities is constructable from the metric $g_{ij} = \delta_{ij}$ and a single scalar susceptibility;* see Ex. 10.7:

$$\boxed{\chi_{ij} = \chi_0 g_{ij} \ , \quad d_{ijk} = 0 \ , \quad \chi_{ijkl} = \frac{1}{3}\chi_4 (g_{ij}g_{kl} + g_{ik}g_{jl} + g_{il}g_{jk}) \ , \quad \chi_{ijklm} = 0 \ , \quad \cdots} \ . \tag{10.25}$$

[There is a caveat to these symmetry arguments: When the nonlinear susceptibilities depend significantly on the frequencies of the three or four waves, then these simple symmetries can be broken. For example, the third-order susceptibility, χ_{ijkl}, for an isotropic medium depends on which of the three input waves is paired with the output wave in Eq. (10.25); so when one orders the input waves with wave 1 on the j index, 2 on the k index and 3 on the l index

[3] For further details see, e.g., Secs. 16.2 and 16.3 of Yariv (1989) or Sec. 14.3 of Sharma (2006). In a lossy medium, symmetry on the first index is lost; see Sec. 8.1 of Yariv and Pochi (2007).

(and output 4 on the i index), the three terms in χ_{ijkl} [Eq. (10.25)] have different scalar coefficients. We shall ignore this subtlety in the remainder of this chapter. For details see, e.g., Sec. 14.3 of Sharma (2006).]

A simple model of a crystal which explains how nonlinear susceptibilities can arise is the following. Imagine each ion in the crystal as having a valence electron that can oscillate in response to a sinusoidal electric field. The electron can be regarded as residing in a potential well which, for low-amplitude oscillations, is very nearly harmonic (potential energy quadratic in displacement; restoring force proportional to displacement; "spring constant" independent of displacement). However, if the electron's displacement from equilibrium becomes a significant fraction of the interionic distance, it will begin to feel the electrostatic attraction of the neighboring ions, and its spring constant will weaken. This means that the potential the electron sees is really not that of a harmonic oscillator, but rather that of an *anharmonic oscillator*, $V(x) = \alpha x^2 - \beta x^3 + \cdots$, where x is the electron's displacement from equilibrium. The nonlinearities in this potential cause the electron's amplitude of oscillation, when driven by a sinusoidal electric field, to be nonlinear in the field strength, and that nonlinear displacement causes the crystal's polarization to be nonlinear.[4] For most crystals, the spatial arrangement of the ions causes the electron's potential energy V to be different for displacements in different directions, and this causes the nonlinear susceptibilities to be anisotropic.

Because the total energy required to liberate the electron from its lattice site is roughly one eV and the separation between lattice sites is $\sim 10^{-10}$ m, the characteristic electric field for strong instantaneous nonlinearities is $\sim 1\mathrm{V}/10^{-10}\mathrm{m} = 10^{10}\,\mathrm{V\ m^{-1}} = 1\mathrm{V}\,(100\mathrm{pm})^{-1}$. Correspondingly, since d_{ijk} has dimensions 1/(electric field) and χ_{ijkl} has dimensions 1/(electric field)^2, *rough upper limits on their Cartesian components are*

$$\boxed{d_{ijk} \lesssim 100\,\mathrm{pm\,V^{-1}}\ , \quad \chi_4 \sim \chi_{ijkl} \lesssim \left(100\,\mathrm{pm\,V^{-1}}\right)^2}\ . \tag{10.26}$$

For comparison, because stronger fields will pull electrons out of solids, the strongest continuous-wave electric fields that occur in practical applications are $E \sim 10^6$ V m^{-1} corresponding to maximum intensities $F \sim 1$ kW / mm^2 $= 1$ GW/m^2. These numbers dictate that, *unless the second-order d_{ijk} are suppressed by isotropy, they will produce much larger effects than the third-order χ_{ijkl}, which in turn will dominate over all higher orders.*

In the next few sections, we shall explore how the nonlinear susceptibilities produce nonlinear couplings of optical waves. There is, however, another application that we must mention in passing. When a slowly changing, non-wave electric field E_k is applied to a nonlinear medium, it can be thought of as producing a change in the linear dielectric tensor for waves $\Delta\chi_{ij} = 2(d_{ijk} + d_{ikj})E_k +$ quadratic terms [cf. Eq. (10.20)]. This is an example (e.g. Boyd 2008) of an ***electro-optic effect***: the modification of optical properties of a medium by an applied electric field. Electro-optic effects are very important in modern optical technology. For example, Pockels cells (used to modulate Gaussian light beams), optical switches (used in Q-switched lasers) and liquid crystal displays (used for computer screens and television screens) are based on electro-optic effects. For some details of several

[4] Quantitative details are worked out, e.g., in Sec. 16.3 of Yariv (1989).

important electro-optic effects and their applications, see, e.g. Chap. 9 of Yariv and Yeh (2006).

****

EXERCISES

Exercise 10.7 *Derivation and Example: Nonlinear Susceptibilities for an Isotropic Medium*

Explain why the nonlinear susceptibilities for an isotropic medium have the forms given in Eq. (10.25). [Hint: Use the facts that the χ's must be symmetric in all their indices, and that, because the medium is isotropic, the χ's must be constructable from the only isotropic tensors available to us, the (symmetric) metric tensor g_{ij} and the (antisymmetric) Levi-Civita tensor ϵ_{ijk}.] What are the corresponding forms, in an isotropic medium, of χ_{ijklmn} and $\chi_{ijklmnp}$?

****

10.6 Three-Wave Mixing in Anisotropic, Nonlinear Crystals

10.6.1 Resonance Conditions for 3-Wave Mixing

When a beam of light is sent through a nonlinear crystal, the nonlinear susceptibilities produce wave-wave mixing. The mixing due to the second-order susceptibility d_{ijk} is called three-wave mixing because three electric fields appear in the polarization-induced interaction energy, Eq. (10.24a). The mixing produced by the third-order χ_{ijkl} is similarly called four-wave mixing. Three-wave mixing dominates in an anisotropic medium, but is suppressed when the medium is isotropic, leaving four-wave mixing as the leading-order nonlinearity.

For use in our analyses of three-wave mixing, in Box 10.2 we list the second-order susceptibilities and some other properties of several specific nonlinear crystals.

Let us examine three-wave mixing in a general anisotropic crystal. Because the nonlinear susceptibilities are so small [i.e., because the input wave will generally be far weaker than 10^{10} V m^{-1} = 1 V (100pm)$^{-1}$], the nonlinearities can be regarded as small perturbations. Suppose that two waves, labeled $n = 1$ and $n = 2$, are injected into the anisotropic crystal, and let their wave vectors be $\mathbf{k}_n$ when one ignores the (perturbative) nonlinear susceptibilities but keeps the large linear χ_{ij}. Because χ_{ij} is an anisotropic function of frequency, the dispersion relation for these waves (ignoring the nonlinearities), $\Omega(\mathbf{k})$, will typically be anisotropic. The frequencies of the two input waves satisfy this dispersion relation, $\omega_n = \Omega(\mathbf{k}_n)$, and the waves' forms are

$$E_j^{(n)} = \Re\left(\mathcal{A}_j^{(n)} e^{i(\mathbf{k}_n\cdot\mathbf{x}-\omega_n t)}\right) = \frac{1}{2}\left(\mathcal{A}_j^{(n)} e^{i(\mathbf{k}_n\cdot\mathbf{x}-\omega_n t)} + \mathcal{A}_j^{(n)*} e^{i(-\mathbf{k}_n\cdot\mathbf{x}+\omega_n t)}\right) , \qquad (10.27)$$

Box 10.2
Properties of Some Anisotropic, Nonlinear Crystals

Notation for Second-Order Susceptibilities

In tabulations of the second-order nonlinear susceptibilities d_{ijk}, a compressed two-index notation d_{ab} is often used, with the indices running over

$$\begin{aligned} a: &\quad 1 = x,\ 2 = y,\ 3 = z, \\ b: &\quad 1 = xx,\ 2 = yy,\ 3 = zz,\ 4 = yz = zy,\ 5 = xz = zx,\ 6 = xy = yx\ . \end{aligned} \tag{1}$$

Crystals with Large Second-Order Susceptibilities

The following crystals have especially large second-order susceptibilities:

$$\begin{aligned} \text{Te: Tellurium} &\quad d_{33} = d_{zzz} = 650\,\text{pm V}^{-1} \\ \text{CdGeAs}_2 &\quad d_{36} = d_{zyx} = 450\,\text{pm V}^{-1} \\ \text{Se: Selenium} &\quad d_{33} = d_{zzz} = 160\,\text{pm V}^{-1}\ . \end{aligned} \tag{2}$$

However, they are not widely used in nonlinear optics because some of their other properties are unfavorable. On the other hand, glasses containing Tellurium or Sellenium have moderately large nonlinearities and are useful.

KH_2PO_4, Potassium Dihidrogen Phosphate; also called KDP

This is among the most widely used nonlinear crystals in 2013, not because of its nonlinear susceptibilities (which are quite modest) but because (i) it can sustain large electric fields without suffering damage, (ii) it is highly birefringent (different light speeds in different directions and for different polarizations), which as we shall see in Sec. 10.6.3 is useful for phase matching), and (iii) it has large electro-optic coefficients (end of Sec. 10.5). At linear order, it is axisymmetric around the z-axis, and its indices of refraction and susceptibilities have the following dependence on wavelength λ (measured in microns), which we shall use in Sec. 10.6.3:

$$\begin{aligned} (\mathfrak{n}_o)^2 &= 1 + \chi_{xx} = 1 + \chi_{yy} = 2.259276 + \frac{0.01008956}{\lambda^2 - 0.012942625} + \frac{13.005522\lambda^2}{\lambda^2 - 400}\ , \\ (\mathfrak{n}_e)^2 &= 1 + \chi_{zz} = 2.132668 + \frac{0.008637494}{\lambda^2 - 0.012281043} + \frac{3.2279924\lambda^2}{\lambda^2 - 400}\ . \end{aligned} \tag{3}$$

The second-order nonlinearities break the axisymmetry of KDP, giving rise to

$$d_{36} = d_{zyx} = 0.44\,\text{pm V}^{-1}\ . \tag{4}$$

Although this is three orders of magnitude smaller than the largest nonlinearities available, its smallness is compensated by its ability to sustain large electric fields.

Box 10.2 (continued)

$KTiOPO_4$, Potassium Titanyl Phosphate; also known as KTP

This is quite widely used in 2013, e.g. in green laser pointers (Ex. 10.13). At linear order it is non-axisymmetric, but with only modest birefringence: its indices of refraction along its three principal axes, at the wavelengths $\lambda = 1064\mu$m and 532μm, are

$$\begin{aligned} 1064\mu\text{m}: &\quad \mathfrak{n}_x = \sqrt{1+\chi_{xx}} = 1.740\ ,\ \mathfrak{n}_y = \sqrt{1+\chi_{yy}} = 1.747\ ,\ \mathfrak{n}_z = \sqrt{1+\chi_{zz}} = 1.830\ ; \\ 532\mu\text{m}: &\quad \mathfrak{n}_x = \sqrt{1+\chi_{xx}} = 1.779\ ,\ \mathfrak{n}_y = \sqrt{1+\chi_{yy}} = 1.790\ ,\ \mathfrak{n}_z = \sqrt{1+\chi_{zz}} = 1.887\ ; \end{aligned} \tag{5}$$

Its third order nonlinearities are moderately large: in units pm V^{-1}

$$\begin{aligned} &d_{31} = d_{zxx} = 6.5\ , \quad d_{32} = d_{zyy} = 5.0\ , \quad d_{33} = d_{zzz} = 13.7\ , \\ &d_{24} = d_{xyz} = d_{xzy} = 7.6\ , \quad d_{15} = d_{xxz} = d_{xzx} = 7.6\ . \end{aligned} \tag{6}$$

Notice that symmetry on the first index is modestly broken: $d_{zxx} = 6.5 \neq d_{xxz} = 7.6$. This symmetry breaking is caused by the crystal's dissipating a small portion of the light power that drives it.

Evolution of Materials

Over the past three decades materials scientists have found and developed nonlinear crystals with ever improving properties. By the time you read this book, the most widely used crystals are likely to have changed.

where we have denoted their vectorial complex amplitudes by $\mathcal{A}_j^{(n)}$. We shall adopt the convention that wave 1 is the wave with the larger frequency, so $\omega_1 - \omega_2 \geq 0$.

These two input waves couple, via the second-order nonlinear susceptibility d_{ijk}, to produce the following contribution to the medium's nonlinear polarization vector:

$$\begin{aligned} P_i^{(3)} &= 2\epsilon_0 d_{ijk}\, 2E_j^{(1)} E_k^{(2)} \\ &= \epsilon_0 d_{ijk} \left(\mathcal{A}_j^{(1)} \mathcal{A}_k^{(2)} e^{i(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{x}} e^{i(\boldsymbol{\omega}_1+\omega_2)t} \ +\ \mathcal{A}_j^{(1)} \mathcal{A}_k^{(2)*} e^{i(\mathbf{k}_1-\mathbf{k}_2)\cdot\mathbf{x}} e^{i(\boldsymbol{\omega}_1-\omega_2)t} + \text{cc} \right) \end{aligned} \tag{10.28}$$

where "cc" means complex conjugate.[5] This sinusoidally oscillating polarization produces source terms in Maxwell's equations (10.16b) and the wave equation (10.19): an oscillating, polarization-induced charge density $\rho_P = -\boldsymbol{\nabla}\cdot\mathbf{P}^{(3)}$ and current density $\mathbf{j}_P = \partial\mathbf{P}^{(3)}/\partial t$. This polarization charge and current, like $\mathbf{P}^{(3)}$ itself [Eq. (10.28)], consist of two traveling waves, one with frequency and wave vector

$$\boxed{\omega_3 = \omega_1 + \omega_2\ , \quad \mathbf{k}_3 = \mathbf{k}_1 + \mathbf{k}_2}\ ; \tag{10.29a}$$

the other with frequency and wave vector

$$\boxed{\omega_3 = \omega_1 - \omega_2\ , \quad \mathbf{k}_3 = \mathbf{k}_1 - \mathbf{k}_2}\ . \tag{10.29b}$$

[5] The reason for the factor 2 in the defintion $P_i = 2\epsilon_0 d_{ijk} E_j E_k$ is to guarantee a factor unity in Eq. (10.28) and in the resulting coupling constant κ of Eq. (10.38) below.

If either of these $(\omega_3, \mathbf{k}_3)$ *satisfies the medium's dispersion relation* $\omega = \Omega(\mathbf{k})$*, then the polarization will generate an electromagnetic wave* $E_j^{(3)}$ *that propagates along in resonance with its polarization-vector source in the wave equation*

$$\nabla^2 \mathbf{E}^{(3)} - \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{E}^{(3)}) + \frac{\omega_3^2}{c^2} \boldsymbol{\epsilon} \cdot \mathbf{E}^{(3)} = \frac{1}{c^2 \epsilon_0} \frac{\partial^2 \mathbf{P}^{(3)}}{\partial t^2} \tag{10.30}$$

[the frequency-ω_3 part of Eq. (10.22a)]. Therefore, this new electromagnetic wave, with frequency ω_3 and wave vector $\mathbf{k}_3$, will grow as it propagates.

For most choices of the input waves, i.e. most choices of $\{\mathbf{k}_1, \omega_1 = \Omega(\mathbf{k}_1), \mathbf{k}_2, \omega_2 = \Omega(\mathbf{k}_2)\}$, neither of the polarizations $\mathbf{P}^{(3)}$ will have a frequency $\omega_3 = \omega_1 \pm \omega_2$ and wave vector $\mathbf{k}_3 = \mathbf{k}_1 \pm \mathbf{k}_2$ that satisfy the medium's dispersion relation, and thus neither will be able to create a third electromagnetic wave resonantly; the wave-wave mixing is ineffective. However, for certain special choices of the input waves, resonant coupling *will* be achieved, and a strong third wave will be produced. See Sec. 10.6.3 below for details.

In nonlinear optics, enforcing the resonance conditions (10.29), with all three waves satisfying their dispersion relations, is called *phase matching*, because it guarantees that the new wave propagates along in phase with the polarization produced by the the two old waves.

The resonance conditions (10.29) have simple quantum mechanical interpretations—a fact that is not at all accidental: quantum mechanics underlies the classical theory that we are developing. Each classical wave is carried by photons that have discrete energies $\mathcal{E}_n = \hbar\omega_n$ and discrete momenta $\mathbf{p}_n = \hbar\mathbf{k}_n$. The input waves are able to produce, resonantly, waves with $\omega_3 = \omega_1 \pm \omega_2$ and $\mathbf{k}_3 = \mathbf{k}_1 \pm \mathbf{k}_2$, if those waves satisfy the dispersion relation. Restated in quantum mechanical terms, the condition of resonance with the "+" sign rather than the "−" is

$$\mathcal{E}_3 = \mathcal{E}_1 + \mathcal{E}_2 \ , \quad \mathbf{p}_3 = \mathbf{p}_1 + \mathbf{p}_2 \ . \tag{10.31a}$$

This has the quantum mechanical meaning that one photon of energy $\mathcal{E}_1$ and momentum $\mathbf{p}_1$, and another of energy $\mathcal{E}_2$ and momentum $\mathbf{p}_2$ combine together, via the medium's nonlinearities, and are annihilated (in the language of quantum field theory), and by their annihilation they create a new photon with energy $\mathcal{E}_3 = \mathcal{E}_1 + \mathcal{E}_2$ and momentum $\mathbf{p}_3 = \mathbf{p}_1 + \mathbf{p}_2$. Thus, the classical condition of resonance is the quantum mechanical condition of energy-momentum conservation for the sets of photons involved in a quantum annihilation and creation process. For this process to proceed, not only must energy-momentum conservation be satisfied, but all three photons must have energies and momenta that obey the photons' semiclassical Hamiltonian relation $\mathcal{E} = H(\mathbf{p})$ (i.e., the dispersion relation $\omega = \Omega(\mathbf{k})$ with $H = \hbar\Omega$, $\mathcal{E} = \hbar\omega$, and $\mathbf{p} = \hbar\mathbf{k}$).

Similarly, the classical conditions of resonance with the "−" sign rather than the "+" can be written (after bringing photon 2 to the left-hand side) as

$$\mathcal{E}_3 + \mathcal{E}_2 = \mathcal{E}_1 \ , \quad \mathbf{p}_3 + \mathbf{p}_2 = \mathbf{p}_1 \ . \tag{10.31b}$$

This has the quantum mechanical meaning that one photon of energy $\mathcal{E}_1$ and momentum $\mathbf{p}_1$ gets annihilated, via the medium's nonlinearities, and from its energy and momentum there are created two photons, with energies $\mathcal{E}_2$, $\mathcal{E}_3$ and momenta $\mathbf{p}_2$, $\mathbf{p}_3$ that satisfy energy-momentum conservation.

10.6.2 Three-Wave-Mixing Evolution Equations in a Medium that is Dispersion-Free and Isotropic at Linear Order

Consider the simple, idealized case where the linear part of the susceptibility χ_{jk} is isotropic and frequency-independent, $\chi_{jk} = \chi_0 g_{jk}$, and correspondingly Maxwell's equations imply $\boldsymbol{\nabla}\cdot\mathbf{E} = 0$. The Track-One part of this chapter will be confined to this idealized case. In the next section (Track Two), we shall treat the more realistic case, which has dispersion and anisotropy at linear order.

In our idealized case, the dispersion relation, ignoring the nonlinearities, takes the simple, nondispersive form [which follows from Eq. (10.20)]

$$\omega = \frac{c}{\mathfrak{n}}k, \quad \text{where } k = |\mathbf{k}|, \quad \mathfrak{n} = \sqrt{1+\chi_0}\,. \tag{10.32}$$

Consider three-wave mixing for waves 1, 2, and 3 that all propagate in the same z direction with wave numbers that satisfy the resonance condition $k_3 = k_1 + k_2$. The dispersion-free dispersion relation (10.32) guarantees that the frequencies will also resonate, $\omega_3 = \omega_1 + \omega_2$. If we write the new wave as $E_i^{(3)} = \Re(\mathcal{A}_i^{(3)} e^{i(k_3 z - \omega_3 t)}) = \frac{1}{2}\mathcal{A}_i^{(3)} e^{i(k_3 z-\omega_3 t)} + \text{cc}$, then its evolution equation (10.30), when combined with Eqs. (10.27) and (10.28), will take the form

$$\nabla^2\left(\mathcal{A}_i^{(3)} e^{i(k_3 z-\omega_3 t)}\right) + \frac{\mathfrak{n}^2\omega_3^2}{c^2}\mathcal{A}_i^{(3)} e^{i(k_3 z-\omega_3 t)} = -2\frac{\omega_3^2}{c^2} d_{ijk}\mathcal{A}_j^{(1)}\mathcal{A}_k^{(2)} e^{i(k_3 z-\omega_3 t)}\,. \tag{10.33}$$

Using the dispersion relation (10.32) and the fact that the lengthscale on which wave 3 changes is long compared to its wavelength (which is always the case because the fields are always much weaker than 10^{10} V m^{-1}), the left-hand side becomes $2ik_3 d\mathcal{A}_i^{(3)}/dz$, and Eq. (10.33) then becomes (with the aid of the dispersion relation) $d\mathcal{A}_i^{(3)}/dz = i(k_3/\mathfrak{n}^2) d_{ijk}\mathcal{A}_j^{(1)}\mathcal{A}_k^{(2)}$. This and similar computations for evolution of the other two waves (Ex. 10.8) give the following equations for the rates of change of the three waves' complex amplitudes:

$$\frac{d\mathcal{A}_i^{(3)}}{dz} = i\frac{k_3}{\mathfrak{n}^2} d_{ijk}\mathcal{A}_j^{(1)}\mathcal{A}_k^{(2)} \quad \text{at } \omega_3 = \omega_1+\omega_2\,,\ \ k_3 = k_1 + k_2\,; \tag{10.34a}$$

$$\frac{d\mathcal{A}_i^{(1)}}{dz} = i\frac{k_1}{\mathfrak{n}^2} d_{ijk}\mathcal{A}_j^{(3)}\mathcal{A}_k^{(2)*} \quad \text{at } \omega_1 = \omega_3-\omega_2\,,\ \ k_1 = k_3 - k_2\,; \tag{10.34b}$$

$$\frac{d\mathcal{A}_i^{(2)}}{dz} = i\frac{k_2}{\mathfrak{n}^2} d_{ijk}\mathcal{A}_j^{(3)}\mathcal{A}_k^{(1)*} \quad \text{at } \omega_2 = \omega_3-\omega_1\,,\ \ k_2 = k_3 - k_1\,. \tag{10.34c}$$

Therefore, each wave's amplitude changes with distance z travelled at a rate proportional to the product of the field strengths of the other two waves.

It is instructive to rewrite the evolution equations (10.34) in terms of renormalized scalar amplitudes $\mathfrak{A}_n$ and unit-normed polarization vectors $f_j^{(n)}$ for the three waves $n = 1, 2, 3$:

$$\mathcal{A}_j^{(n)} = \sqrt{\frac{2k_n}{\epsilon_0\mathfrak{n}^2}}\,\mathfrak{A}_n\, f_j^{(n)} = \sqrt{\frac{2\omega_n}{\epsilon_0 c\,\mathfrak{n}}}\,\mathfrak{A}_n\, f_j^{(n)}\,. \tag{10.35}$$

This renormalization is motivated by the fact that $|\mathfrak{A}_n|^2$ is proportional to the flux of quanta $dN_n/dAdt$ associated with wave n. Specifically: the energy density in wave n is (neglecting

nonlinearities) $U = \epsilon_o(1+\chi_o)\overline{\mathbf{E}^2} = \frac{1}{2}\epsilon_o \mathfrak{n}^2 |\mathcal{A}^{(n)}|^2$ (where the bar means time average); the energy flux is this U times the wave speed $c/\mathfrak{n}$:

$$F_n = \frac{1}{2}\epsilon_o \mathfrak{n} c |\mathcal{A}^{(n)}|^2 = \omega_n |\mathfrak{A}_n|^2 \; ; \tag{10.36}$$

and the flux of quanta is this F_n divided by the energy $\mathcal{E}_n = \hbar\omega_n$ of each quantum: $dN_n/dAdt = |\mathfrak{A}_n|^2/\hbar$.

The 3-wave-mixing evolution equations (10.34), rewritten in terms of the renormalized amplitudes, take the simple form

$$\frac{d\mathfrak{A}_3}{dz} = i\kappa\, \mathfrak{A}_1 \mathfrak{A}_2 \; , \quad \frac{d\mathfrak{A}_1}{dz} = i\kappa\, \mathfrak{A}_3 \mathfrak{A}_2^* \; , \quad \frac{d\mathfrak{A}_2}{dz} = i\kappa\, \mathfrak{A}_3 \mathfrak{A}_1^* \; ; \quad \kappa = \sqrt{\frac{2\omega_1\omega_2\omega_3}{\epsilon_0 c^3 \mathfrak{n}^3}}\, d_{ijk}\, f_i^{(1)} f_j^{(2)} f_k^{(3)} \; . \tag{10.37}$$

It is straightforward to verify that these evolution equations guarantee energy conservation $d/dz(F_1+F_2+F_3) = 0$, with F_n given by Eq. (10.36). Therefore, at least one wave will grow and at least one wave will decay due to three-wave mixing.

When waves 1 and 2 are the same wave, the three-wave mixing leads to *frequency doubling*: $\omega_3 = 2\omega_1$. In this case, the nonlinear polarization that produces the third wave is $P_i = \epsilon_0 d_{ijk} E_j^{(1)} E_k^{(2)}$, by contrast with that when waves 1 and 2 are different, $P_i = 2\epsilon_0 d_{ijk} E_j^{(1)} E_j^{(2)}$ [Eq. (10.28)]. [In the latter case the factor 2 arises because we are dealing with cross terms in $(E_j^{(1)} + E_j^{(2)})(E_k^{(1)} + E_k^{(2)})$.] Losing the factor 2 and making wave 2 the same as wave 1 leads to an obvious modification of the evolution equations (10.37):

$$\frac{d\mathfrak{A}_3}{dz} = \frac{i\kappa}{2}(\mathfrak{A}_1)^2 \; , \quad \frac{d\mathfrak{A}_1}{dz} = i\kappa\, \mathfrak{A}_3 \mathfrak{A}_1^* \; ; \quad \kappa = \frac{1}{c}\sqrt{\frac{\omega_1^2\omega_3}{\mathfrak{n}^3}}\, d_{ijk}\, f_i^{(1)} f_j^{(1)} f_k^{(3)} \; . \tag{10.38}$$

Once again, it is easy to verify energy conservation, $d/dz(F_1+F_3) = 0$. We shall discuss frequency doubling in Sec. 10.7.1.

****

EXERCISES

Exercise 10.8 *Derivation: Evolution Equations in Idealized Three-Wave Mixing*

Derive Eqs. (10.34b) and (10.34c) for the amplitudes of waves 1 and 2 produced by three-wave mixing.

****

10.6.3 T2 Three-Wave Mixing in a Birefringent Crystal: Phase Matching and Evolution Equations

Ordinary Waves, Extraordinary Waves, and Dispersion Relations

In reality, all nonlinear media have frequency-dependent dispersion relations and most are anisotropic at linear order and therefore birefringent (different wave speeds in different directions). An example is the crystal KDP (Box 10.2), which has indices of refraction

$$\mathfrak{n}_o = \sqrt{1+\chi_{xx}} = \sqrt{1+\chi_{yy}}\,, \quad \mathfrak{n}_e = \sqrt{1+\chi_{zz}};, \tag{10.39}$$

that depend on the light's wave number $k = 2\pi/\lambda$ in the manner shown in Fig. 10.12a and in Eq. (3) of Box 10.2. The subscript "o" stands for *ordinary*; e, for *extraordinary*; see below.

Maxwell's equations imply that, in this crystal, for plane, monochromatic waves propagating in the $x-z$ plane at an angle θ to the symmetry axis [$\mathbf{k} = k(\sin\theta\mathbf{e}_x + \cos\theta\mathbf{e}_z)$], there are two dispersion relations corresponding to the two polarizations of the electric field: (i) If $\mathbf{E}$ is orthogonal to the symmetry axis, then (as is shown in Ex. 10.9), it must also be orthogonal to the propagation direction (i.e., must point in the $\mathbf{e}_y$ direction), and the dispersion relation is

$$\frac{\omega/k}{c} = (\text{phase speed in units of speed of light}) = \frac{1}{\mathfrak{n}_o}\,, \tag{10.40a}$$

independent of the angle θ. These waves are called *ordinary*, and their phase speed (10.40a) is the lower curve in Fig. 10.12a; at $k = 10\mu\text{m}^{-1}$ (point A), the phase speed is $0.663c$, while at $k = 20\mu\text{m}^{-1}$, it is $0.649c$. (ii) If $\mathbf{E}$ is not orthogonal to the symmetry axis, then (Ex. 10.9) it must lie in the plane formed by $\mathbf{k}$ and the symmetry axis (the $x-z$) plane, with $E_x/E_z = -(\mathfrak{n}_e/\mathfrak{n}_o)^2\cot\theta$ [which means that $\mathbf{E}$ is *not* orthogonal to the propagation direction unless the crystal is isotropic, $\mathfrak{n}_e = \mathfrak{n}_o$, which it is not]; and the dispersion relation is

$$\frac{\omega/k}{c} = \frac{1}{\mathfrak{n}} = \sqrt{\frac{\cos^2\theta}{\mathfrak{n}_o^2} + \frac{\sin^2\theta}{\mathfrak{n}_e^2}}\,. \tag{10.40b}$$

In this case the waves are called *extraordinary*. As the propagation direction varies from parallel to the symmetry axis ($\cos\theta = 1$) to perpendicular ($\sin\theta = 1$), this extraordinary phase speed varies from $c/\mathfrak{n}_o$ (the lower curve in Fig. 10.12; $0.663c$ at $k = 10\mu\text{m}^{-1}$), to $c/\mathfrak{n}_e$ (the upper curve; $0.681c$ at $k = 10\mu\text{m}^{-1}$).

Phase Matching for Frequency Doubling in a KDP Crystal

This birefringence enables one to achieve phase matching (satisfy the resonance conditions) in three-wave mixing. As an example, consider the resonance conditions for a *frequency-doubling device* (discussed in further detail in the next section): one in which the two input waves are identical, so $\mathbf{k}_1 = \mathbf{k}_2$ and $\mathbf{k}_3 = 2\mathbf{k}_1$ point in the same direction. Let this common propagation direction be at an angle θ to the symmetry axis. Then the resonance conditions reduce to the demands that the output wave number be twice the input wave number, $k_3 = 2k_1$, and the output phase speed be the same as the input phase speed, $\omega_3/k_3 = \omega_1/k_1$. Now, for waves of the same type (both ordinary or both extraordinary), the phase speed is a monotonic decreasing function of wave number [Fig. 10.12a and Eqs. (10.40)], so there is no choice of propagation angle θ that enables these resonance conditions to be satisfied. The only way to satisfy them is by using ordinary input waves and extraordinary output waves, and then only for a special, frequency-dependent propagation

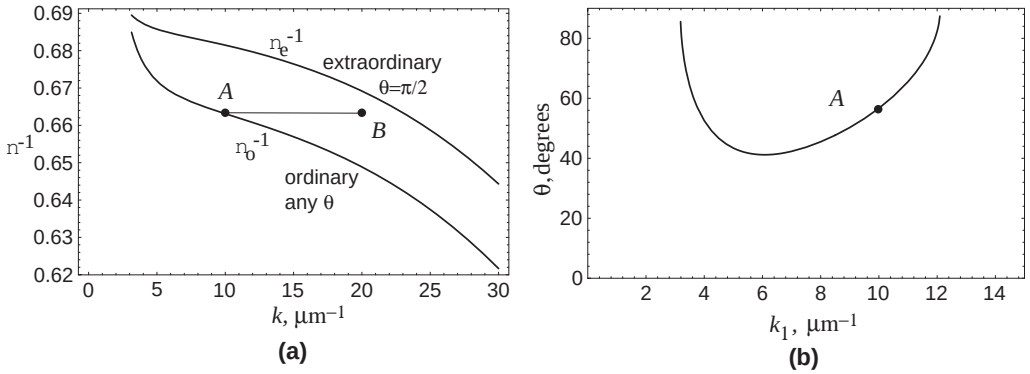

Fig. 10.12: (a) The inverse of the index of refraction $\mathfrak{n}^{-1}$ (equal to the phase speed in units of the speed of light) for electromagnetic waves propagating at an angle θ to the symmetry axis of a KDP crystal, as a function of wave number k in reciprocal microns. See Eq. (10.40a) for lower curve and Eq. (10.40b) with $\theta = \pi/2$ for upper curve. For extraordinary waves propagating at an arbitrary angle θ to the crystal's symmetry axis, $\mathfrak{n}^{-1}$ is a mean [Eq. (10.40b)] of the two plotted curves. The plotted curves are fit by the analytical formulae (3) of Box 10.2. (b) The angle θ to the symmetry axis at which ordinary waves with wave number k_1 (e.g. point A) must propagate in order that 3-wave mixing be able to produce frequency doubled or phase conjugated extraordinary waves (e.g. point B).

direction. This technique is called *type I phase matching*; "type I" because there are other techniques for phase matching (i.e. for arranging that the resonance conditions be satisfied); see, e.g., Ex. 10.13d.

As an example, if the input waves are ordinary, with $k_1 = 10\mu\text{m}^{-1}$ (approximately the value for light from a ruby laser; point A in Fig. 10.12a), then the output waves must be extraordinary and must have the same phase speed as the input waves (same height in Fig. 10.12a) and have $k_3 = 2k_1 = 20\mu\text{m}^{-1}$ (i.e., point B). This phase speed is between $c/\mathfrak{n}_e(2k_1)$ and $c/\mathfrak{n}_o(2k_1)$, and thus can be achieved for a special choice of propagation angle: $\theta = 56.7^\circ$ (point A in Fig. 10.12b). In general, Eqs. (10.40) imply that *the unique propagation direction θ at which the resonance conditions can be satisfied is the following function of the input wave number k_1:*

$$\sin^2\theta = \frac{1/\mathfrak{n}_o^2(k_1) \; - \; 1/\mathfrak{n}_o^2(2k_1)}{1/\mathfrak{n}_e^2(2k_1) - 1/\mathfrak{n}_o^2(2k_1)} \; . \tag{10.41}$$

This resonance angle is plotted as a function of wave number for KDP in Fig. 10.12b.

This special case of identical input waves illustrates the very general phenomenon, that *at fixed input frequencies, the resonance conditions can be satisfied only for special, discrete input and output directions.*

For our frequency doubling example, the extraordinary dispersion relation (10.40b) for the output wave can be rewritten as

$$\omega = \frac{ck}{\mathfrak{n}} = \Omega_e(\mathbf{k}) = c\sqrt{\frac{k_z^2}{\mathfrak{n}_o(k)^2} + \frac{k_x^2}{\mathfrak{n}_e(k)^2}} \; , \quad \text{where} \quad k = \sqrt{k_x^2 + k_z^2} \; . \tag{10.42}$$

Correspondingly, the group velocity[6] $V_g^j = \partial\Omega/\partial k_j$ for the output wave has components

$$\begin{aligned} V_g^x &= V_{\rm ph}\sin\theta\left(\frac{\mathfrak{n}^2}{\mathfrak{n}_e^2} - \frac{\mathfrak{n}^2\cos^2\theta}{\mathfrak{n}_o^2}\frac{d\ln\mathfrak{n}_o}{d\ln k} - \frac{\mathfrak{n}^2\sin^2\theta}{\mathfrak{n}_e^2}\frac{d\ln\mathfrak{n}_o}{d\ln k}\right) , \\ V_g^z &= V_{\rm ph}\cos\theta\left(\frac{\mathfrak{n}^2}{\mathfrak{n}_o^2} - \frac{\mathfrak{n}^2\cos^2\theta}{\mathfrak{n}_o^2}\frac{d\ln\mathfrak{n}_o}{d\ln k} - \frac{\mathfrak{n}^2\sin^2\theta}{\mathfrak{n}_e^2}\frac{d\ln\mathfrak{n}_o}{d\ln k}\right) , \end{aligned} \tag{10.43}$$

where $V_{\rm ph} = \omega/k = c/\mathfrak{n}$ is the phase velocity. For an ordinary input wave with $k_1 = 10\mu\text{m}^{-1}$ (point A in Fig. 10.12) and an extraordinary output wave with $k_3 = 20\mu\text{m}^{-1}$ (point B), these formulae give for the direction of the output group velocity (direction along which the output waves grow) $\theta_g = \arctan(V_g^x/V_g^z) = 58.4^\circ$, compared to the direction of the common input-output phase velocity $\theta = 56.7^\circ$; and they give for the magnitude of the group velocity $V_g = 0.628c$, compared to the common phase velocity $v_{\rm ph} = 0.663c$. Thus, *the differences between the group velocity and the phase velocity are small, but they do differ.*

Evolution Equations

Once one has found wave vectors and frequencies that satisfy the resonance conditions, the evolution equations for the two (or three) coupled waves have the same form as in the idealized dispersion-free, isotropic case [Eqs. (10.38) or (10.37)], but with minor modifications. Specifically:

Let planar input waves impinge on a homogeneous, nonlinear crystal at some plane $z = 0$ and therefore (by symmetry) have energy fluxes inside the crystal that evolve as functions of z only: $\mathbf{F}_n = \mathbf{F}_n(z)$ for waves $n = 1$ and 3 in the case of frequency doubling (or 1, 2, 3 in the case of three different waves). Then energy conservation dictates that

$$\frac{d}{dz}\sum_n F_{n\,z} = 0 , \tag{10.44}$$

where $F_{n\,z}(z)$ is the z component of the energy flux for wave n. It is convenient to define a complex amplitude $\mathfrak{A}_n$ for wave n that is related to the wave's complex electric field amplitude by an analog of Eq. (10.35):

$$\mathcal{A}_j^{(n)} = \zeta_n\sqrt{\frac{2\omega_n}{\epsilon_0 c\,\mathfrak{n}}}\,\mathfrak{A}_n f_j^{(n)} . \tag{10.45}$$

Here $f_j^{(n)}$ is the wave's polarization vector, $\mathfrak{n}_n$ is its index of refraction (defined by $\omega_n/k_n = c/\mathfrak{n}_n$), and ζ_n is some positive real constant that depends on the relative directions of $\mathbf{k}_n$, $\mathbf{f}^{(n)}$, $\mathbf{e}_z$ and has a value which ensures that

$$F_{n\,z} = \omega_n|\mathfrak{A}_n|^2 \tag{10.46}$$

[same as Eq. (10.36) but with F_n replaced by $F_{n\,z}$]. Since the energy flux is $\hbar\omega_n$ times the photon number flux, this equation tells us that $|\mathfrak{A}_n|^2/\hbar$ is the photon number flux (just like the idealized case).

[6]Since a wave's energy travels with the group velocity, it must be that $\mathbf{V}_g = \mathbf{E}\times\mathbf{H}/U$, where U is the wave's energy density, $\mathbf{E}\times\mathbf{H}$ is its Poynting vector (energy flux), and $\mathbf{H} = \mathbf{B}/\mu_0$ (in our dielectric medium). It can be shown explicitly that this, indeed, is the case.

Because the evolution equations involve the same photon creation and annihilation processes as in the idealized case, they must have the same mathematical form as in the idealized case [Eqs. (10.38) or (10.37)], except for the magnitude of the coupling constant. (For a proof see Ex. 10.10.) Specifically, for frequency doubling of a wave 1 to produce wave 3 ($\omega_3 = \omega_1 + \omega_2$), the resonant evolution equations and coupling constant are:

$$\boxed{\frac{d\mathfrak{A}_3}{dz} = \frac{i\kappa}{2}(\mathfrak{A}_1)^2 \ , \quad \frac{d\mathfrak{A}_1}{dz} = i\kappa\,\mathfrak{A}_3\mathfrak{A}_1^* \ ; \quad \kappa = \beta\sqrt{\frac{2\omega_1^2\omega_3}{\epsilon_0 c^3 \mathfrak{n}_1^2 \mathfrak{n}_3}}\, d_{ijk}\, f_i^{(1)} f_j^{(1)} f_k^{(3)}} \tag{10.47}$$

[cf. Eqs. (10.38) for the idealized case]. For resonant mixing of three different waves ($\omega_3 = \omega_1 + \omega_2$), they are:

$$\boxed{\frac{d\mathfrak{A}_3}{dz} = i\kappa\,\mathfrak{A}_1\mathfrak{A}_2 \ , \quad \frac{d\mathfrak{A}_1}{dz} = i\kappa\,\mathfrak{A}_3\mathfrak{A}_2^* \ , \quad \frac{d\mathfrak{A}_2}{dz} = i\kappa\,\mathfrak{A}_3\mathfrak{A}_1^* \ ; \quad \kappa = \beta'\sqrt{\frac{2\omega_1\omega_2\omega_3}{\epsilon_0 c^3 \mathfrak{n}_1\mathfrak{n}_2\mathfrak{n}_3}}\, d_{ijk}\, f_i^{(1)} f_j^{(2)} f_k^{(3)}} \tag{10.48}$$

[cf. Eqs. (10.37) for the idealized case]. Here β and β' are constants of order unity that depend on the relative directions of $\mathbf{e}_z$ and the wave vectors $\mathbf{k}_n$ and polarization vectors $\mathbf{f}^{(n)}$; see Ex. 10.10.

It is useful to keep in mind the following magnitudes of the quantitites that appear in these three-wave-coupling equations (Ex. 10.11)

$$\boxed{F_n \lesssim 1\mathrm{GW\,m^{-2}} \ , \quad |\mathfrak{A}_n| \lesssim 10^{-3}\sqrt{\mathrm{J\,m^{-1}}} \ , \quad \kappa \lesssim 10^5 \mathrm{J^{-1/2}} \ , \quad |\kappa\,\mathfrak{A}_n| \lesssim 1/\mathrm{cm} \ .} \tag{10.49}$$

We shall use the evolution equations (10.47) and (10.48) in the next section to explore several applications of three-wave mixing.

****

EXERCISES

Exercise 10.9 ***T2*** ***Example: Dispersion Relation for an Anisotropic Medium*

Consider a wave propagating through a dielectric medium that is anisotropic, but not necessarily—for the moment—axisymmetric. Let the wave be sufficiently weak that nonlinear effects are unimportant. Define the wave's displacement vector in the usual way, $D_i = \epsilon_0 E_i + P_i$ [Eq. (10.17)].

(a) Show that

$$D_i = \epsilon_o \epsilon_{ij} E_j \ , \quad \text{where } \epsilon_{ij} \equiv \delta_{ij} + \chi_{ij} \equiv \text{ "dielectric tensor"}; \tag{10.50}$$

ϵ_o is often absorbed into the dielectric tensor, but we find it more convenient to normalize ϵ_{ij} so that in vacuum $\epsilon_{ij} = \delta_{ij}$.

(b) Show that the wave equation (10.19) for the electric field takes the form

$$-\nabla^2\mathbf{E} + \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{E}) = -\frac{1}{c^2}\boldsymbol{\epsilon}\cdot\frac{\partial^2\mathbf{E}}{\partial t^2} \ . \tag{10.51}$$

(d) Now specialize to a monochromatic plane wave with angular frequency ω and wave vector $\mathbf{k}$. Show that the wave equation (10.51) reduces to

$$L_{ij}E_j = 0 \ , \quad \text{where} \quad L_{ij} = k_i k_j - k^2\delta_{ij} + \frac{\omega^2}{c^2}\epsilon_{ij} \ . \tag{10.52a}$$

This equation says that E_j is an eigenvector of L_{ij} with vanishing eigenvalue, which is possible if and only if

$$\det||L_{ij}|| = 0 \ . \tag{10.52b}$$

This vanishing determinant is the waves' dispersion relation. We shall use it in Chap. 21 to study waves in plasmas.

(e) Next specialize to an axisymmetric medium and orient the symmetry axis along the z direction so the only nonvanishing components of ϵ_{ij} are $\epsilon_{11} = \epsilon_{22}$ and ϵ_{33}, and let the wave propagate in a direction $\hat{\mathbf{k}}$ that makes an angle θ to the symmetry axis. Show that in this case L_{ij} has the form

$$||L_{ij}|| = k^2 \left|\left| \begin{matrix} (\mathfrak{n}_o/\mathfrak{n})^2 - \cos^2\theta & 0 & \sin\theta\cos\theta \\ 0 & (\mathfrak{n}_o/\mathfrak{n})^2 - 1 & 0 \\ \sin\theta\cos\theta & 0 & (\mathfrak{n}_e/\mathfrak{n})^2 - \sin^2\theta \end{matrix} \right|\right| \ , \tag{10.53a}$$

and the dispersion relation (10.52b) reduces to

$$\left(\frac{1}{\mathfrak{n}^2} - \frac{1}{\mathfrak{n}_o^2}\right)\left(\frac{1}{\mathfrak{n}^2} - \frac{\cos^2\theta}{\mathfrak{n}_o^2} - \frac{\sin^2\theta}{\mathfrak{n}_e^2}\right) = 0 \ , \tag{10.53b}$$

where $1/\mathfrak{n} = \omega/kc$, $\mathfrak{n}_o = \sqrt{\epsilon_{11}} = \sqrt{\epsilon_{22}}$, and $\mathfrak{n}_e = \sqrt{\epsilon_{33}}$, in accord with Eq. (10.39).

(f) Show that this dispersion relation has the two solutions (ordinary and extraordinary) discussed in the text, Eqs. (10.40a) and (10.40b), and show that the electric fields associated with these two solutions point in the directions described in the text.

Exercise 10.10 ***T2*** ***Derivation and Example: Evolution Equations for Realistic Wave-Wave Mixing*

Derive the evolution equations (10.48) for 3-wave mixing. [The derivation of those (10.47) for frequency doubling is similar.] You could proceed as follows:

(a) Insert expressions (10.27) and (10.28) into the general wave equation (10.30) and extract the portions with frequency $\omega_3 = \omega_1 + \omega_2$, thereby obtaining the generalization of Eq. (10.33):

$$\begin{aligned} &\nabla^2\left(\mathcal{A}_i^{(3)}e^{i(k_3 z - \omega_3 t)}\right) - \frac{\partial^2}{\partial x^i \partial x^j}\left(\mathcal{A}_j^{(3)}e^{i(k_3 z - \omega_3 t)}\right) + \frac{\omega_3^2}{c^2}\epsilon_{ij}\mathcal{A}_j^{(3)}e^{i(k_3 z - \omega_3 t)} \\ &\quad = -2\frac{\omega_3^2}{c^2}d_{ijk}\mathcal{A}_j^{(1)}\mathcal{A}_k^{(2)}e^{ik_3 z - \omega_3 t} \ . \end{aligned} \tag{10.54}$$

(b) Infer from the homogeneous wave equation (10.51) for $\mathbf{E}^{(3)}$ that $e^{i(\mathbf{k}_3\cdot\mathbf{x}-\omega_3 t)}$ satisfies that same wave equation. Then, splitting each wave into its scalar field and polarization vector, $\mathcal{A}_i^{(n)} \equiv \mathcal{A}^{(n)} f_i^{(n)}$, and letting each $\mathcal{A}^{(n)}$ be a function of z (because of the boundary condition that the 3-wave mixing begins at the crystal face $z = 0$), show that Eq. (10.54) reduces to $\alpha_3 d\mathcal{A}^{(3)}/dz = i(k_3/\mathfrak{n}_3^2) d_{ijk} f_j^{(1)} f_k^{(2)} \mathcal{A}^{(1)}\mathcal{A}^{(2)}$, where α_3 is a constant of order unity that depends on the relative orientations of the unit vectors $\mathbf{e}_z$, $\mathbf{f}^{(3)}$, and $\hat{\mathbf{k}}_3$. Note that, aside from α_3, this is the same evolution equation as for our idealized isotropic, dispersion-free medium, Eq. (10.34a). Show that, similarly, $\mathcal{A}^{(1)}(z)$ and $\mathcal{A}^{(2)}(z)$ satisfy the same equations (10.34b) and (10.34c) as in the idealized case, aside from multiplicative constants α_1 and α_2.

(c) Adopting the renormalizations $\mathcal{A}^{(n)} = \zeta_n \sqrt{\omega_n/\mathfrak{n}_n}\, \mathfrak{A}_n$ [Eq. (10.45)] with ζ_n so chosen that the photon number flux for wave n is proportional to $|\mathfrak{A}_n|^2$, show that your evolution equations for $\mathcal{A}_n$ become Eqs. (10.48), except that the factor β' and thence the value of κ might be different for each equation.

(d) Since the evolution entails one photon with frequency ω_1 and one with frequency ω_2 anihilating to produce a photon with frequency ω_3, it must be that $d|\mathfrak{A}_1|^3/dz = d|\mathfrak{A}_2|^3/dz = -d|\mathfrak{A}_3|^2/dz$. (These are called *Manley-Rowe relations*.) By imposing this on your evolution equations in part (c), deduce that all three coupling constants κ must be the same, and thence also all three β' must be the same; and therefore the evolution equations take precisely the claimed form, (10.48).

Exercise 10.11 ***T2*** ***Derivation: Magnitudes of 3-Wave-Mixing Quantities*

Derive Eqs. (10.49). Hints: The maximum energy flux in a wave arises from the limit $E \lesssim 10^6 \mathrm{V\,m^{-1}}$ on the wave's electric field to ensure that it not pull electrons out of the surface of the nonlinear medium; and the maximum coupling constant κ arises from the largest values $|d_{ijk}| \lesssim 10\mathrm{pm\,V^{-1}}$ for materials typically used in 3-wave mixing (Box 10.2).

****

10.7 Applications of Three-Wave Mixing: Frequency Doubling, Optical Parametric Amplification, and Squeezed Light

10.7.1 Frequency Doubling

Frequency doubling (also called *second harmonic generation*) is one of the most important applications of wave-wave mixing. As we have already discussed briefly in Secs. 10.6.2 (Track One) and 10.6.3 (Track Two), it can be achieved by passing a single wave (which plays the role of both wave $n = 1$ and wave $n = 2$) through a nonlinear crystal, with the propagation direction chosen to satisfy the resonance conditions. As we have also seen in the previous

section (Track Two), the crystal's birefringence and dispersion have little influence on the growth of the output wave, $n = 3$ with $\omega_3 = 2\omega_1$; it grows with distance inside the crystal at a rate given by Eqs. (10.47), which is the same as in the Track-One case of a medium that is isotropic at linear order, Eqs. (10.38). By doing a sufficiently good job of phase matching (satisfying the resonance conditions) and choosing the thickness of the crystal appropriately, one can achieve close to 100% conversion of the input-wave energy into frequency-doubled energy. More specifically, if wave 1 enters the crystal at $z = 0$ with $\mathfrak{A}_1(0) = \mathfrak{A}_{1o}$ which we choose (without loss of generality) to be real, and if there is no incoming wave 3 so $\mathfrak{A}_3(0) = 0$, then the solution to the evolution equations (10.47) or (10.38) is

$$\mathfrak{A}_3 = \frac{i}{\sqrt{2}}\,\mathfrak{A}_{1o}\tanh\left(\frac{\kappa}{\sqrt{2}}\,\mathfrak{A}_{1o}\,z\right)\ , \quad \mathfrak{A}_1 = \mathfrak{A}_{1o}\,\mathrm{sech}\left(\frac{\kappa}{\sqrt{2}}\,\mathfrak{A}_{1o}\,z\right)\ . \tag{10.55}$$

It is easy to see that this solution has the following properties: (i) It satisfies energy conservation, $2|\mathfrak{A}_3| + |\mathfrak{A}_1|^2 = |\mathfrak{A}_{1\,o}|^2$. (ii) At a depth $z = 1.246/(\kappa\mathfrak{A}_{1o})$ into the crystal, half the initial energy has been frequency doubled. (iii) As z increases beyond this half-doubling depth, the light asymptotes to fully frequency doubled.

One might expect the frequency doubling to proceed onward to $4\omega_1$ etc. However, it typically does not because these higher-frequency waves typically fail to satisfy the crystal's dispersion relation.

As an example, the Neodymium:YAG (Nd^{3+}:YAG) laser, which is based on an Yttrium Aluminum Garnet crystal with trivalent Neodymium impurities, is among the most attractive of all lasers for a combination of high frequency stability, moderately high power, and high efficiency. However, this laser operates in the infrared, at a wavelength of 1.064 microns. For many purposes, one wants optical light. This can be achieved by frequency doubling the laser's output, thereby obtaining 0.532 micron (green) light. This is how green laser pointers, used in lecturing, work (though in 2012 they are typically driven not by Nd:YAG but rather a relative; see Ex. 10.13.

Frequency doubling also plays a key role in laser fusion, where intense, pulsed laser beams, focused on a pellet of fusion fuel, compress and heat the pellet to high densities and temperatures. Because the beam's energy flux is inversely proportional to the area of its focused cross section, and because the larger the wavelength, the more seriously diffraction impedes making the cross section small, it is important to give the beam a very short wavelength. This is achieved by multiple frequency doublings, which *can and do* occur in the experimental setup of laser fusion, though not in typical nonlinear crystals.

****

EXERCISES

Exercise 10.12 *Derivation: Saturation in Frequency Doubling*

Derive the solution (10.55) to the evolution equations (10.47) for frequency doubling, and verify that it has the claimed properties.

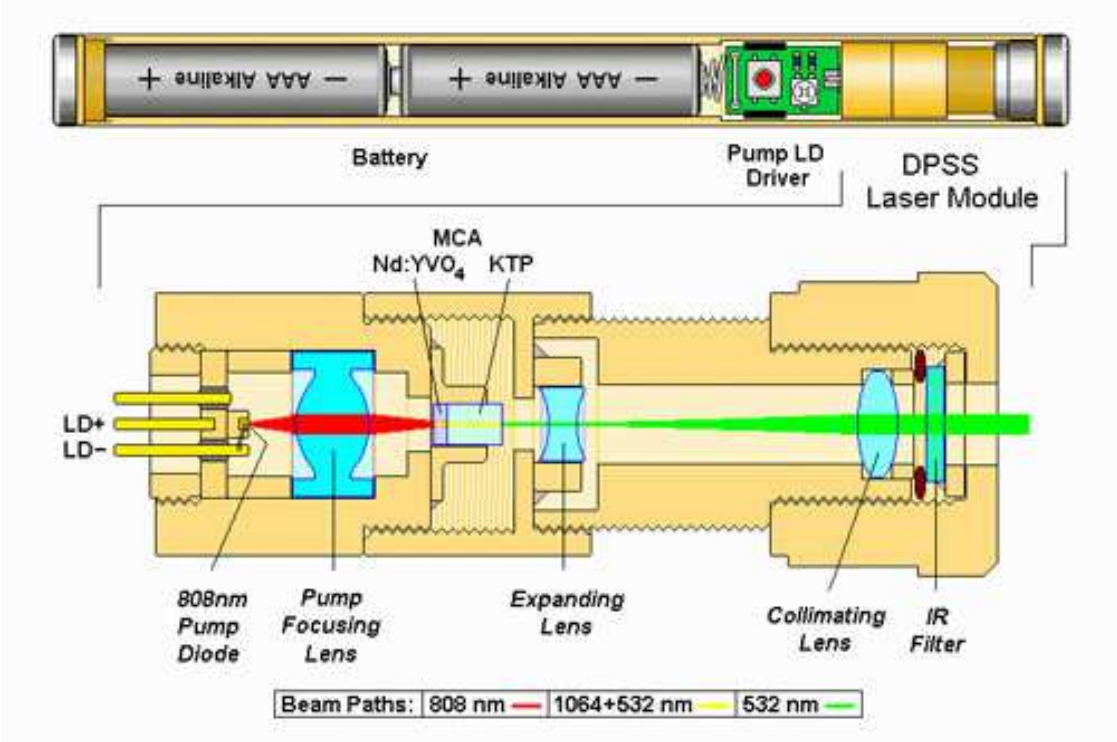

Fig. 10.13: Structure of a green laser pointer, circa 2012. [From Wikimedia Commons: File:Green-laser-pointer-dpss-diagrams.jpg .]

Exercise 10.13 ***Example: Frequency Doubling in a Green Laser Pointer*

Green laser pointers, popular in 2013, have the structure shown in Fig. 10.13. A battery-driven infrared diode laser puts out 808 nm light that pumps a Nd:YVO_4 laser crystal (Neodymium-doped ytrium vanadate; a relative of Nd:YAG). The 1064 nm light beam from this Nd:YVO_4 laser is frequency doubled by a KTP crystal, resulting in 532 nm green light. An infrared filter removes all the 880 nm and 1064 nm light from the output, leaving only the green.

(a) To make the frequency doubling as efficient as possible, the light is focused to as small a beam radius ϖ_o as diffraction allows, as it travels through the KTP crystal. Assuming that the crystal length is $L \simeq 1$ cm, show that $\varpi_o \simeq \sqrt{\lambda L/2\pi} \simeq 10\mu$m (about 10 times larger than the 1064 nm wavelength). [Hint: use the properties of Gaussian beams; Sec. 8.5.5]

(b) The 1064 nm beam has an input power $W_{1o} \simeq 100$ mW as it enters the KTP crystal. Show that its energy flux and its electric field strength are $F \simeq 30$ MW m^{-2} and $\mathcal{A}^{(1)} \simeq 160$ kV m^{-1}.

(c) Assuming that phase matching has been carried out successfully (i.e., photon energy and momentum conservation have been enforced), use Eq. (10.47)) or (10.55) for the growth of the green output light, and KTP's nonlinear susceptibility [Eq. (6) of Box 10.2], to show that the output green beam has $|\mathfrak{A}_3|^2 \sim 0.1\,\mathfrak{A}_{1o}^2$, corresponding to an

output power $W_3 \sim 0.02\, W_{1o} \simeq 2$ mW. This is close to the maximum legal power for such a laser pointer in the United States in 2013: 5 mW.

(d) Using the indices of refraction for KTP at 1064 nm and 532 nm listed in Eq. (5) of Box 10.2, work out the details of how the phase matching (enforcement of energy and momentum conservation for photons) can be achieved for this frequency doubling in KTP.

****

10.7.2 Optical Parametric Amplification

In optical parametric amplification, the energy of a *pump wave* is used to amplify an initially weak *signal wave* and also amplify an uninteresting *idler wave*. The waves satisfy the resonance conditions with $\omega_p = \omega_s + \omega_i$. The pump wave and signal wave are fed into an anisotropic nonlinear crystal, propagating in (nearly) the same direction, with nonzero renormalized amplitudes $\mathfrak{A}_{po}$ and $\mathfrak{A}_{so}$ at $z = 0$. The idler wave has $\mathfrak{A}_{io} = 0$ at the entry plane. Because the pump wave is so strong, it is negligibly influenced by the three-wave mixing; i.e., $\mathfrak{A}_p$ remains constant inside the crystal.

The evolution equations for the (renormalized) signal and idler amplitudes are

$$\frac{d\mathfrak{A}_s}{dz} = -i\kappa \mathfrak{A}_p \mathfrak{A}_i^* \ , \quad \frac{d\mathfrak{A}_i}{dz} = -i\kappa \mathfrak{A}_p \mathfrak{A}_s^* \tag{10.56}$$

[Eqs. (10.48) or (10.37)]. For the initial conditions of weak signal wave and no idler wave, the solution to these equations is

$$\mathfrak{A}_s = \mathfrak{A}_{so} \cosh(|\gamma| z) \ , \quad \mathfrak{A}_i = \frac{\gamma}{|\gamma|} \mathfrak{A}_{so}^* \sinh(|\gamma| z) \ ; \quad \gamma \equiv -i\kappa \mathfrak{A}_p \ . \tag{10.57}$$

Thus, the signal field grows exponentially, after an initial pause, with an e-folding length $1/|\gamma|$, which for strong 3-wave nonlinearities is of order a centimeter [Ex. (10.14)].

****

EXERCISES

Exercise 10.14 *Derivation: e-folding Length for an Optical Parametric Amplifier*

Estimate the magnitude of the e-folding length for an optical parametric amplifier that is based on a strong 3-wave nonlinearity.

****

10.7.3 Degenerate Optical Parameteric Amplification: Squeezed Light

Consider optical parametric amplification with the signal and idler frequencies identical, so the idler field is the same as the signal field and the pump frequency is twice the signal frequency: $\omega_p = 2\omega_s$. This condition is called *degenerate*. Adjust the phase of the pump field so that $\gamma = -i\kappa\mathfrak{A}_p$ is real and positive. Then the equation of evolution for the signal field is the same as appears in frequency doubling [Eq. (10.47) or (10.38)]:

$$d\mathfrak{A}_s/dz = \gamma\mathfrak{A}_s^* \ . \tag{10.58}$$

The resulting evolution is most clearly understood by decomposing $\mathfrak{A}_s$ into its real and imaginary parts (as we did in Ex. 6.23 when studying thermal noise in an oscillator): $\mathfrak{A}_s = X_1 + iX_2$. Then the time evolving electric field is

$$E \propto \Re(\mathfrak{A}_s e^{i(k_s z - \omega_s t)}) = X_1 \cos(k_s z - \omega_s t) + X_2 \sin(k_s z - \omega_s t) \ . \tag{10.59}$$

Therefore, X_1 is the amplitude of the field's cosine quadrature and X_2 is the amplitude of its sine quadrature. Equation (10.58) then says that $dX_1/dz = \gamma X_1$, $dX_2/dz = -\gamma X_2$, so

$$X_1 = X_{1o} e^{\gamma z} \ , \quad X_2 = X_{2o} e^{-\gamma z} \ . \tag{10.60}$$

Therefore, the wave's cosine quadrature gets amplified as the wave propagates, and its sine quadrature gets attenuated. This is called *squeezing* because X_2 is reduced (squeezed) while X_1 is increased. It is a phenomenon known to children who swing; see Ex. 10.15.

Squeezing is especially interesting when it is applied to noise. Typically, a wave has equal amounts of noise in its two quadratures, i.e. the standard deviations ΔX_1 and ΔX_2 of the two quadratures are equal (as was the case in Ex. 6.23). When such a wave is squeezed, its two standard deviations get altered in such a way that their product is unchanged:

$$\boxed{\Delta X_1 = \Delta X_{1o}\, e^{\gamma z} \ ; \quad \Delta X_2 = \Delta X_{2o}\, e^{-\gamma z} \ , \quad \Delta X_1 \Delta X_2 = \text{constant}} \ . \tag{10.61}$$

see Fig. 10.14 below. When, as here, the standard deviations of two quadratures differ, the light is said to be in a *squeezed state*.

In quantum theory, X_1 and X_2 are complementary observables; they are described by Hermitian operators that do not commute. The uncertainty principle associated with their noncommutation implies that their product $\Delta X_1 \Delta X_2$ has some minimum possible value. This minimum is achieved by the wave's vacuum state, which has $\Delta X_1 = \Delta X_2$ with values corresponding to one half quantum of energy (vacuum fluctuations) in the field mode that we are studying. When this "quantum electrodynamic vacuum" is fed into a degenerate optical parametric amplifier, the vacuum noise gets squeezed in the same manner (10.60) as any other noise.

Squeezed states of light, including this *squeezed vacuum*, have great promise for fundamental physics experiments and technology. For example, they can be used to reduce the photon shot noise of an interferometer or a laser below the "standard quantum limit"

of $\Delta N = \sqrt{N}$ (Poisson statistics), thereby improving the signal to noise ratio in certain communications devices, and in laser interferometer gravitational-wave detectors.[7]

We explore some properties of squeezed light in Ex. 10.16.

****

EXERCISES

Exercise 10.15 *** Example: Squeezing by Children Who Swing*

A child, standing in a swing, bends her knees then straightens them twice per swing period, making the distance ℓ from the swing's support to her center of mass oscillate as $\ell = \ell_0 + \ell_1 \sin 2\omega_0 t$. Here $\omega_0 = \sqrt{g\ell_0}$ is the swing's mean angular frequency.

(a) Show that the swing's angular displacement from vertical, θ, obeys the equation of motion

$$\frac{d^2\theta}{dt^2} + \omega_0^2\theta = -\omega_1^2 \sin(2\omega_0 t)\theta \ , \tag{10.62}$$

where $\omega_1 = \sqrt{g\ell_1}$ and θ is assumed small, $\theta \ll 1$.

(b) Write $\theta = X_1 \cos\omega_0 t + X_2 \sin\omega_0 t$. Assuming that $\ell_1 \ll \ell_0$ so $\omega_1 \ll \omega_0$, show that the child's knee bending (her "pumping" the swing) squeezes θ:

$$X_1(t) = X_1(0)e^{-(\omega_1^2/8\omega_o)t} \ , \quad X_2(t) = X_2(0)e^{+(\omega_1^2/8\omega_o)t} \tag{10.63}$$

(c) Explain how this squeezing is related to the child's conscious manipulation of the swing — i.e., to her strategy for increasing the swing's amplitude when she starts up, and her strategy for reducing the amplitude when she wants to quit swinging.

Exercise 10.16 ***Squeezed States of Light*

Consider a plane, monochromatic electromagnetic wave with angular frequency ω, whose electric is expressed in terms of its complex amplitude $X_1 + iX + 2$ by Eq. (10.60). Because the field (inevitably) is noisy, its quadrature amplitudes X_1 and X_2 are random processes with means $\bar{X}_1$, $\bar{X}_2$ and variances ΔX_1, ΔX_2.

(a) Normal, classical light has equal amounts of noise in its two quadratures. Explain why it can be represented by Fig. 10.14a.

(b) Explain why Fig. 10.14b represents *phase-squeezed light*, and show that its electric field as a function of time has the form shown in Fig. 10.15.

(c) Explain why Fig. 10.14c represents *amplitude-squeezed light*, and construct a diagram of its electric field as a function of time analogous to Fig. 10.15.

[7] For detailed discussions see, e.g., Walls (1983), Wu *et al.* (1986), and LaPorta *et al.* (1989).

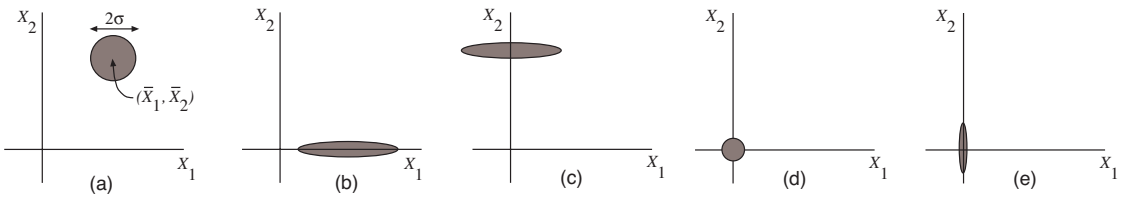

Fig. 10.14: Error boxes in the complex amplitude plane for several different electromagnetic waves: (a) Classical light. (b) Phase-squeezed light. (c) Amplitude-squeezed light. (d) The quantum electrodynamical vacuum. (e) The squeezed vacuum.

(d) Figure 10.14d represents the vacuum state of light's frequency-ω plane-wave mode. Give a formula for the diameter of the mode's circular error box. Construct a diagram of the electric field as a function of time analogous to Fig. 10.15.

(e) Figure 10.14e represents the squeezed vacuum. Construct a diagram of its electric field as a function of time analogous to Fig. 10.15.

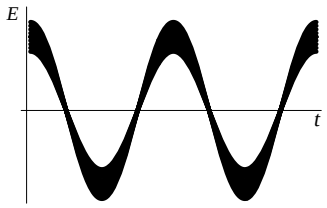

Fig. 10.15: The error band for the electric field $E(t)$, as measured at a fixed location in space, when phase-squeezed light passes by.

****

10.8 Four-Wave Mixing in Isotropic Media

10.8.1 Third-Order Susceptibilities and Field Strengths

The nonlinear polarization for four-wave mixing, $P_i^{(4)} = 4\epsilon_0 \chi_{ijkl} E_j E_k E_l$, is typically smaller than that $P_i^{(3)} = 2\epsilon_0 d_{ijk} E_j E_k$ for 3-wave mixing by $\sim E|\chi/d| \sim (10^6\,\mathrm{V\,m^{-1}})(100\mathrm{pm\,V^{-1}}) \sim 10^{-4}$. (Here we have used the largest electric field that solids typically can support.) Therefore (as we have already discussed), only when d_{ijk} is greatly suppressed by isotropy of the nonlinear material does χ_{ijkl} and 4-wave mixing become the dominant nonlinearity. And in that case, we expect the propagation lengthscale for strong, cumulative 4-wave mixing to be $\sim 10^4$ larger than that (~ 1 cm) for the strongest 3-wave mixing; i.e., $\ell_{4w} \gtrsim 100$ m.

In reality, as we shall see in the next subsection, this estimate is overly pessimistic. In special materials, ℓ_{4w} can be less than a meter (though still much bigger than the 3-wave mixing's centimeter). Two factors enable this: (i) If the nonlinear material is a fluid (e.g. CS_2) confined by solid walls, then it can support somewhat larger electric field strengths

than a nonlinear crystal's maximum, $10^6\,\mathrm{V\,m^{-1}}$. (ii) If the nonlinear material is made of molecules significantly larger than 10^{-10} m (e.g., organic molecules), then the molecular electric dipoles induced by an electric field can be significantly larger than our estimates (10.26); and correspondingly, $|\chi_{ijkl}|$ can significantly exceed $(100\ \mathrm{pm\ V^{-1}})^2$; see Table 10.1.

Material	Wavelength	$\mathfrak{n}$	χ_{1111} $(\mathrm{pm/V})^2$	n_2 $(10^{-20}\mathrm{m^2\ W^{-1}})$
Fused Silica	0.694 μm	1.455	56.4	3.
SF_6 glass	1.06	1.77	587.	21.
CS_2 liquid	1.06	1.594	6,400.	290.
2-Methyl-4-nitroaniline (MNA) organic crystal†		1.8	1.69×10^5	5,800.
PTS polydiacetylene polymeric crystal†		1.88	5.53×10^5	1.8×10^4

†Also has large d_{ijk}.

Table 10.1: Materials used in 4-wave mixing. At the indicated light wavelength, $\mathfrak{n}$ is the index of refraction, χ_{1111} is the third-order nonlinear susceptibility, and n_2 is the Kerr coefficient of Eq. (10.70). Adapted from Table 8.8 of Yariv and Yeh (2007).

In the next two subsections, we shall give (i) an example with strong 4-wave mixing: phase conjugation by a half-meter-long cell containing CS_2 liquid; and then (ii) an example with weak but important 4-wave mixing: light propagation in a multi-kilometer long fused-silica optical fiber.

10.8.2 Phase Conjugation Via Four-Wave Mixing in CS_2 Fluid

As an example of four-wave mixing, we discuss phase conjugation in a rectangular cell that contains CS_2 liquid (Fig. 10.16a).[8] The fluid is pumped by two strong waves, 1 and 2, propagating in opposite directions with the same frequency as the incoming wave 3 that is to be phase conjugated. The pump waves are planar without modulation, but wave 3 has a spatial modulation (slow compared to the wave number) that carries, for example, a picture; $\mathcal{A}_3 = \mathcal{A}_3(x, y; z)$. As we shall see, nonlinear interaction of the two pump waves 1 & 2 and the incoming wave 3 produces outgoing wave 4, which is the phase conjugate of wave 3. All four waves propagate in planes of constant x and have their electric fields in the x direction, so the relevant component of the third-order nonlinearity is $\chi_{xxxx} = \chi_{1111}$.

The resonance conditions (photon energy and momentum conservation) for this 4-wave mixing process are $\omega_4 = \omega_1 + \omega_2 - \omega_3$ and $\mathbf{k}_4 = \mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3$. Since the three input waves all have the same frequency, $\omega_1 = \omega_2 = \omega_3 = \omega$, the output wave 4 will also have $\omega_4 = \omega$, so this is *fully degenerate four-wave mixing*. Since the pump waves propagate in opposite directions, they satisfy $\mathbf{k}_1 = -\mathbf{k}_2$, whence the output wave has $\mathbf{k}_4 = -\mathbf{k}_3$; i.e., it propagates in the opposite direction to the input wave 3 and has the same frequency, as it must, if it is to be (as we claim) the phase conjugate of 3.

[8]This example is adapted from Sec. 8 of Yariv and Yeh (2007).

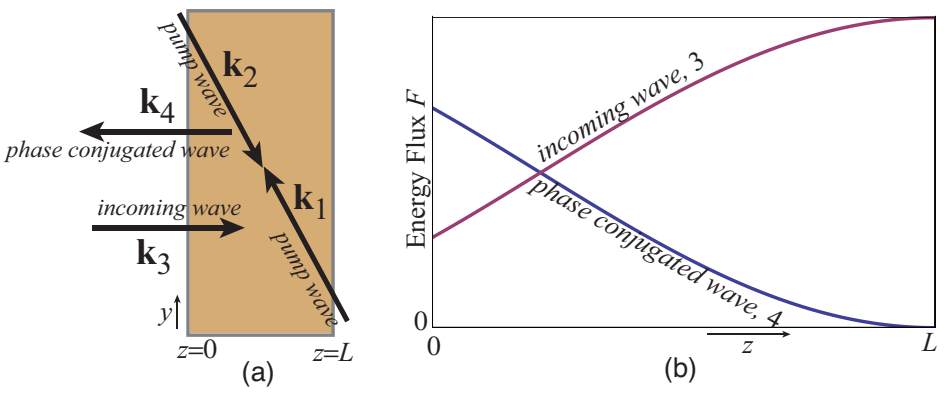

Fig. 10.16: (a) A phase conjugating mirror based on 4-wave mixing. (b) The evolution of the incoming wave's flux and the phase conjugated wave's flux inside the mirror (the nonlinear medium).

The nonlinear polarization that generates wave 4 is $P_x^{(4)} = 4\epsilon_0\chi_{1111}(E_x^{(1)}E_x^{(2)}E_x^{(3)} + E_x^{(2)}E_x^{(3)}E_x^{(1)} + \ldots)$. There are six terms in the sum (six ways to order the three waves), so $P_x^{(4)} = 24\epsilon_0\chi_{1111}E_x^{(1)}E_x^{(2)}E_x^{(3)}$. Inserting $E_x^{(n)} = \frac{1}{2}(\mathcal{A}^{(n)}e^{i(\mathbf{k}_n\cdot\mathbf{x}-\omega_n t)} + \mathcal{A}^{(n)*}e^{i(-\mathbf{k}_n+\omega_n t)})$ into this $P_x^{(4)}$ and plucking out the relevant term for our phase conjugation process (with the signal wave 3 phase conjugated and the pump waves not), we obtain

$$P_x^{(4)} = 3\epsilon_0\chi_{1111}\mathcal{A}^{(1)}\mathcal{A}^{(2)}\mathcal{A}^{(3)*}e^{-i(\mathbf{k}_4\cdot\mathbf{x}-\omega_4 t)} \ . \tag{10.64}$$

Inserting these into the wave equation for wave 4 in an isotropic medium, $(\nabla^2 + \mathfrak{n}^2\omega_4^2/c^2)$ $(\mathcal{A}^{(4)}e^{-i(\mathbf{k}_4\cdot\mathbf{x}-\omega_4 t)}) = -\omega_4^2 P_x^{(4)}$ [analog of the isotropic-medium wave equation (10.33) for 3-wave mixing] and making use of the fact that the lengthscale on which wave 4 changes is long compared to its wavelength, we obtain the following evolution equation for wave 4, which we augment with that for wave 3, obtained in the same way:

$$\frac{d\mathcal{A}^{(4)}}{dz} = -\frac{3ik}{\mathfrak{n}^2}\chi_{1111}\mathcal{A}^{(1)}\mathcal{A}^{(2)}\mathcal{A}^{(3)*} \ , \quad \frac{d\mathcal{A}^{(3)}}{dz} = -\frac{3ik}{\mathfrak{n}^2}\chi_{1111}\mathcal{A}^{(1)}\mathcal{A}^{(2)}\mathcal{A}^{(4)*} \ . \tag{10.65}$$

Here we have dropped subscripts from k since all waves have the same scalar wave number, and we have used the dispersion relation $\omega/k = c/\mathfrak{n}$ common to all four waves since they all have the same frequency. We have not written down the evolution equations for the pump waves because in practice they are very much stronger than the incoming and phase conjugated waves, so they change hardly at all during the evolution.

It is convenient to change normalizations of the wave fields, as in 3-wave mixing, from $\mathcal{A}^{(n)}$ (the electric field) to $\mathfrak{A}_n$ (the square root of energy flux divided by frequency, $\mathfrak{A}_n = \sqrt{F_n/\omega_n}$):

$$\mathcal{A}^{(n)} = \sqrt{\frac{2k}{\epsilon_0\mathfrak{n}^2}}\mathfrak{A}_n = \sqrt{\frac{2\omega_n}{\epsilon_0 c\mathfrak{n}}}\mathfrak{A}_n \tag{10.66}$$

[Eq. (10.35)]. Inserting these into the evolution equations (10.65) and absorbing the constant pump-wave amplitudes into a coupling constant, we obtain

$$\boxed{\frac{d\mathfrak{A}_4}{dz} = -i\kappa\mathfrak{A}_3^* \ , \quad \frac{d\mathfrak{A}_3}{dz} = -i\kappa^*\mathfrak{A}_4 \ ; \quad \kappa = \frac{6\omega^2}{c^2\mathfrak{n}^2\epsilon_0}\chi_{1111}\mathfrak{A}_1\mathfrak{A}_2} \ . \tag{10.67}$$

These are our final, very simple equations for the evolution of the input and phase conjugate waves in our isotropic, nonlinear medium (CS_2 fluid). Inserting the index of refraction $\mathfrak{n} = 1.594$ and nonlinear susceptibility $\chi_{1111} = 6400$ (pm/V)^2 for CS_2 (Table 10.1), the angular frequency corresponding to 1.06 μm wavelength light from, say, a Nd:YAG laser, and letting both pump waves $n = 1$ and 2 have energy fluxes $F_n = \omega|\mathfrak{A}_n|^2 = 5 \times 10^{10}$ W m^{-2} (corresponding to electric field amplitudes 6.1×10^6 V m^{-1}, six times larger than good nonlinear crystals can support), we obtain for the magnitude of the coupling constant $|\kappa| = 1/0.59$ m. Thus, the CS_2 cell of Fig. 10.16a need only be a half meter thick in order to produce strong phase conjugation.

For an input wave $\mathfrak{A}_{3o}(x, y)$ at the cell's front face $z = 0$ and no input wave 4, the solution to the evolution equations (10.67) is easily found to be, in the interior of the cell:

$$\mathfrak{A}_4(x, y, z) = \frac{\kappa}{|\kappa|}\left(\frac{\sin[|\kappa|(z - L)]}{\cos[|\kappa|L]}\right)\mathfrak{A}^*_{3o}(x, y) \ , \quad \mathfrak{A}_3(x, y, z) = \left(\frac{\cos[|\kappa|(z - L)]}{\cos[|\kappa|L]}\right)\mathfrak{A}_{3o}(x, y) \ . \tag{10.68}$$

The corresponding energy fluxes, $F_n = \omega|\mathfrak{A}_n|^2$ (ignoring the spatial modulation) are plotted in Fig. 10.16b above, for a crystal thickness $L = 1/|\kappa| = 0.59$ m. Notice that the pump waves amplify the rightward propagating input wave, so it grows from the crystal front to the crystal back; and at the same time, the interaction of the input wave with the pump waves generates the leftward propagating phase conjugated wave, which begins with zero strength at the back of the crystal and grows (when $L \sim 1/|\kappa|$) to be stronger than the input wave at the crystal front.

****

EXERCISES

Exercise 10.17 ***Problem: Photon Creation and Annihilation in a Phase-Conjugate Mirror*

Describe the creation and annihilation of photons that underlies in a phase-conjugating mirror's four-wave mixing. Specifically: how many photons of each wave are created or annihilated? [Hint: See the discussion of photon creation and annihilation for three-wave mixing at the end of Sec. 10.6.1.]

Exercise 10.18 ***Problem: Spontaneous Oscillation in 4-Wave Mixing*

Suppose the thickness of the nonlinear medium of the text's 4-wave mixing analysis is $L = \pi/2\kappa$, so the denominators in Eqs. (10.68) are zero. Explain the physical nature of the resulting evolution of waves 3 and 4.

Exercise 10.19 *Problem: Squeezed Light Produced by Phase Conjugation*

Suppose a light beam is split in two by a beam splitter. One beam is reflected off an ordinary mirror, and the other off a phase conjugating mirror, and the beams are then recombined at the beam splitter. Suppose that the powers returning to the beam splitter are nearly the same; they differ by a fractional amount $\Delta P/P = \epsilon \ll 1$. Show that the recombined light is in a strongly squeezed state, and discuss how one can guarantee it is phase squeezed, or (if one prefers) amplitude squeezed.

****

10.8.3 Optical Kerr Effect and Four-Wave Mixing in an Optical Fiber

Suppose that an isotropic, nonlinear medium is driven by a single input plane wave polarized in the x direction, $E_x = \Re[\mathcal{A}e^{-(kz-\omega t)}]$. This input wave produces the following polarization that propagates along with itself in resonance (Ex. 10.20):

$$P_x = \epsilon_0\chi_0 E_x + 6\epsilon_0\chi_{1111}\overline{E^2}E_x \ . \tag{10.69}$$

Here the second term is due to four-wave mixing, and $\overline{E^2}$ is the time average of the square of the electric field, which can be expressed in terms of the energy flux as $\overline{E^2} = F/\epsilon_0\mathfrak{n}c$. The four-wave-mixing term can be regarded as a nonlinear correction to χ_0: $\Delta\chi_0 = 6\chi_{1111}\overline{E^2} = (6\chi_{111}/\mathfrak{n}c\,\epsilon_0)F$; and since the index of refraction is $\mathfrak{n} = \sqrt{1+\chi_0}$, this corresponds to a fractional change of index of refraction given by

$$\boxed{\Delta\mathfrak{n} = n_2 F \ , \quad \text{where} \quad n_2 = \frac{3\chi_{1111}}{\mathfrak{n}^2 c\,\epsilon_0}} \ . \tag{10.70}$$

This nonlinear change of $\mathfrak{n}$ is called the *optical Kerr effect*, and the coefficient n_2 is called the *Kerr coefficient* and has dimensions of $1/$(energy flux), i.e. $\mathrm{m}^2\ \mathrm{W}^{-1}$. Values for n_2 for several materials are listed in Table 10.1 above.

We have already briefly discussed an important application of the optical Kerr effect: the self focusing of a laser beam, which plays a key role in mode locked lasers (Sec. 10.2.3) and also in laser fusion.

The optical Kerr effect is also important in the optical fibers used in modern communication (e.g., to carry telephone, television, and internet signals to your home). Such fibers are generally designed to support just one spatial mode of propagation: the fundamental Gaussian mode of Sec. 8.5.5 or some analog of it. Their light-carrying cores are typically made from fused silica, doped with particular impurities, so their Kerr coefficients are $n_2 \simeq 3\times10^{-20}\ \mathrm{m}^2\ \mathrm{W}^{-1}$ (Table 10.1). Although the fibers are not spatially homogeneous and the wave is not planar, one can show (and it should not be surprising) that the fibers nontheless exhibit the optical Kerr effect, with $\Delta\mathfrak{n} = n_2 F_{\rm eff}$. Here $F_{\rm eff}$, the *effective energy flux*, is the light beam's power P divided by an effective cross sectional area, $\pi\sigma_0^2$, with σ_0 the Gaussian beam's radius, defined in Eq. (8.39): $F_{\rm eff} = P/\pi\sigma_0^2$; see Sec. 14.1 of Yariv and Yeh (2007).

As a realistic indication of the importance of the optical Kerr effect in communication fibers, consider a signal beam with mean power $P = 10$ mW and a beam radius $\sigma_0 = 5\mu$m and thence effective energy flux $F_{\rm eff} = 127\ \mathrm{MW\ m}^{-2}$. If the wavelength is $2\pi/k = 0.693\mu$m, then Table 10.1 gives $\mathfrak{n} = 1.455$ and $n_2 = 3\times10^{-20}\ \mathrm{m}^2\ \mathrm{W}^{-1}$. When this beam travels a distance $L = 50$ km along the fiber, its light experiences a phase shift

$$\Delta\phi = \frac{\Delta n}{n}kL = \frac{n_2}{\mathfrak{n}}FkL \simeq 1.2\,\text{radians} \ . \tag{10.71}$$

A phase shift of this size or larger can cause significant problems for optical communication. As examples: (i) Variations of the flux, when pulsed signals are being transmitted, cause time-varying phase shifts that modify the signals' phase evolution (*self-phase modulation*). One consequence of this is broadening of each pulse; another is a nonlinearly induced chirping of each pulse (slightly lower frequency at beginning and higher at end). (ii) Fibers generally carry many channels of communication, with slightly different carrier frequencies, and the time-varying flux of one channel can modify the phase evolution of another (*cross-phase modulation*). Various techniques have been developed to deal with these issues. See, e.g., Chap. 14 of Yariv and Yeh (2007).

In long optical fibers, pulse broadening due to the nonlinear optical Kerr effect can be counterbalanced by a narrowing of a pulse due to linear dispersion (dependence of group velocity on frequency). The result is an *optical soliton*, i.e. a pulse of light with a very special shape that travels down the fiber without any broadening or narrowing [see, e.g., Sec. 1.45 of Yariv and Yeh (2007)]. In Sec. 16.3, we shall study, in full mathematical detail, this same soliton phenomenon for nonlinear waves on the surface of water; and in Sec. (23.6), we shall study it for nonlinear waves in plasmas.

****

EXERCISES

Exercise 10.20 *Derivation: Optical Kerr Effect*

(a) Derive Eq. (10.69) for the polarization induced in an isotropic medium by a linearly polarized electromagnetic wave.

(b) Fill in the remaining details of the derivation of Eq. (10.70) doe the optical Kerr effect.

****

Bibliographic Note

For a lucid and detailed discussion of lasers and their applications, see Saleh and Teich (2007), and at a more advanced level, Yariv and Yeh (2007). For less detailed but clear discussions see standard optics textbooks, such as Ghatak (2010), Hecht (2002), and Jenkins and White (2001).

For a lucid and very detailed discussion of holography and its applications, see Goodman (2005). Most optics textbooks contain less detailed but clear discussions. We like Brooker (2003), Jenkins and White (2001), Ghatak (2010), Hecht (2002), and Sharma (2006).

Wave-wave mixing in nonlinear media is discussed in great detail, with many applications, by Yariv and Yeh (2007). Some readers might find an earlier book by Yariv (1989) pedagogically easier; it was written when the subject was less rich, but the foundations were already in place, and it has a more quantum mechanical focus. The fundamental concepts of wave-wave mixing and its underlying physical processes are treated especially nicely by Boyd (2008). A more elementary treatment with focus on applications is given by Saleh and Teich (2007). Among treatments in standard optics texts, we like Sharma (2006).

Box 10.3
Important Concepts in Chapter 10

- Lasers, Sec. 10.2
  - Spontaneous emission, stimulated emission, and absorption, Sec. 10.2.1
  - Population inversion and the basic principles underlying a laser, Sec. 10.2.1
  - Types of lasers and their properties, Sec. 10.2.2
- Holography, Sec. 10.3
  - Recording a hologram Sec. 10.3.1; Fig 10.6
  - Using a hologram to reconstruct a 3D image, Sec. 10.3.2, Fig. 10.8
  - Every day applications of holography, Sec. 10.3.2, Ex. 10.5
- Phase conjugation, its relationship to time reversal, and practical applications, Sec. 10.4; Figs. 10.10, 10.11
  - How phase conjugation is achieved via four-wave mixing, Sec. 10.8.2
- Nonlinear susceptibilities in a dielectric crystal, Eqs. (10.21), (10.24a)
  - their magnitudes, Eqs. (10.26)
  - their influence on Maxwell's equations, Sec. 10.5
- Opto-electric effects and some applications, end of Sec. 10.5
- Resonance conditions for three-wave mixing and their relationship to creation and annihilation of quanta, Sec. 10.6.1
  - Why they can only be satisfied for a special and restricted set of waves, Sec. 10.6.3
- Three-wave mixing: general form of evolution equations, Secs. 10.6.2 and 10.6.3
- Applications of three-wave mixing
  - Frequency doubling, Secs. 10.6.2, 10.7.1
  - Optical parametric amplification, Sec. 10.7.2
  - Squeezed light, and its relationship to degenerate optical parametric amplification, Sec. 10.7.3
  - Squeezing by a swinging child, Ex. 10.15
- Four-wave mixing in an isotropic medium, Sec. 10.8
  - Phase conjugation via four-wave mixing, Sec. 10.8.2
  - Optical Kerr effect, and its consequences in an optical fiber, Sec. 10.8.3

Bibliography

Basov, N. G. & Prokhorov, A.M. 1954. JETP 27 431

Basov, N. G. & Prokhorov, A.M. 1955. JETP 28 249

Boyd, R. W. 2008. *Nonlinear Optics*, third edition, New York: Academic Press

Cathey, W. T. 1974. *Optical Information Processing and Holography* New York: Wiley

Cundiff, Steven T. 2002. "Phase stabilization of ultrashort optical pulses," *Journal of Physics D: Applied Physics* **35**, R43–R59.

Feynman, R. P., Leighton, R. B., & Sands, M. 1965. *The Feynman Lectures on Physics* New York: Addison Wesley

Ghatak, Ajoy 2010. *Optics*, New York: McGraw-Hill

Ghatak, A. K. & Thyagarajan, K. 1978. *Contemporary Optics* New YorK: Plenum

Goodman, Joseph W. 1985. *Statistical Optics*, New York: Wiley

Gordon, J. P., Zeiger, H. J., & Townes, C. H. 1954. Phys. Rev. 95 282

Gordon, J. P., Zeiger, H. J., & Townes, C. H. 1955. Phys. Rev. 99 1264

Hecht, E. 2002 *Optics*, fourth edition, New York: Addison Wesley

Iizuka, K. 1987. *Engineering Optics* Berlin: Springer-Verlag

Jackson, J. D. 1999. *Classical Electrodynamics*New York: Wiley

La Porta, A., Slusher, R. E., & Yurke, B. 1989. Phys. Rev. Lett. 62 28

Maiman, T.H. 1960 *Nature* 187 493

Saleh, B. E. A. and Teich, M. C. 2007. *Fundamentals of Photonics*, New York: Wiley

Sharma, K. K. 2006. *Optics: Principles and Applications*, Amsterdam: Elsevier.

Walls, D.F. 1983. Nature 306 141

Weber, J. 1953. *IRE Trans. Prof. Group on Electron Devices* 3 1

Wu, L.-A., Kimble, H. J., Hall, J .L., & Wu, H. 1986. Phys. Rev. Lett. 57 2520

Yariv, A. 1977. J. Opt. Soc. Amer. 67 1

Yariv, A. 1989. *Quantum Electronics*, New York: Wiley

Yariv, A. and Yeh, P. 2007 *Photonics: Optical Electronics in Modern communications*, Oxford: Oxford University Press

Zel'dovich, B.Y., Popovichev, V.I., Ragulskii, V.V., & Faisullov, F.S. 1972. JETP 15 109