Introduction to Optical Electronics

Second Edition

AMNON YARIV

Introduction to Optical Electronics

Second Edition

AMNON YARIV

California Institute of Technology

HOLT, RINEHART AND WINSTON
New York Chicago San Francisco Atlanta
Dallas Montreal Toronto London Sydney

Library of Congress Cataloging in Publication Data

Yariv, Amnon.
Introduction to optical electronics.
Includes index.
1. Lasers. 2. Electro-optics. 3. Quantum electronics. I. Title.
TA1675.Y37 1976 537.5'6 76-11773
ISBN 0-03-089892-7

Printed in the United States of America
7 8 9 038 9 8 7 6 5 4 3 2

Preface

The rapidly growing field of quantum electronics can be divided roughly into two broad categories. The first is concerned chiefly with the atomic aspects of the problem. These include the study of energy levels, lifetimes, and transition rates in laser media, and also the mechanisms and the physical origins of phenomena such as Raman and Rayleigh scattering and second-harmonic generation. This branch of the field leans heavily on the formalism of quantum mechanics and has consequently become the domain of the physicist and, to a lesser extent, the physical chemist.

The second category deals with the *coherent interactions of optical radiation fields with various atomic media*. Here we tend to accept the existence of certain physical phenomena and concern ourselves with their implications and applications. The physical properties may now be represented by parameters characteristic of the material. Two typical examples are: (1) the analysis of power output and frequency pulling in laser oscillators in which the physical phenomena of spontaneous emission and atomic dispersion are important, and (2) the problem of optical second-harmonic generation and phase matching in which the complicated quantum mechanical considerations involved in understanding the optical nonlinearity are lumped into the nonlinear constant.

This second aspect of quantum electronics is more closely linked to applications and has consequently attracted the attention of the applied physicist and the electrical engineer. In this area, a good deal of the emphasis is on optics rather than on quantum physics and many of the concepts encountered here have their counterparts in radio and microwave electronics. For this reason I have decided to refer to the subject matter as optical electronics and to choose the same name for the book's title.

Since the appearance of the first edition, a number of topics have become important. These topics have been incorporated into this new edition and involve the following:

- Expanded treatment of Gaussian beam propagation in homogeneous and in focusing media
- Unstable optical resonators
- Heterojunction injection lasers
- High pressure CO_2 lasers
- Noise and detection error probability in binary communication channels
- Mode locking in homogeneously broadened laser systems
- Propagation in symmetric and asymmetric dielectric waveguides
- Mode coupling and directional coupling in dielectric waveguides
- Distributed feedback lasers

Although the first edition was aimed at students in the senior year or in the first year of graduate studies, it was used mostly by graduate students. To encourage this trend, I have augmented the level of mathematical sophistication used in some of the discussions. Nevertheless, I still believe that the ever-increasing role of coherent optics in science and technology will require an early exposure to this area on the part of most electrical engineering and applied physics students. With this in mind, I have undertaken to present the material without the use of quantum mechanics. Instead of inventing quasi-classical substitutes for quantum mechanical concepts, I decided to ask the student to accept on faith certain statements whose justification can only be provided by quantum mechanics. Somewhat to my own surprise, I found that this was necessary only when introducing the concepts of stimulated and spontaneous transitions. The rest of the material can then be treated using classical formalism. An introductory knowledge of atomic physics and of electromagnetic theory would be helpful, although the basic results are derived in the text.

I am grateful to Drs. U. Ganiel, R. MacAnally, and S. Kurtin for critical comments and suggestions and to Ruth Stratton and Dian Rapchak for their patient and competent typing of the manuscript.

Pasadena, California
June, 1976

AMNON YARIV

Contents

CHAPTER

1

Electromagnetic Theory

1.0 Introduction

In this chapter we derive some of the basic results concerning the propagation of plane, single-frequency, electromagnetic waves in homogeneous isotropic media, as well as in anisotropic crystal media. Starting with Maxwell's equations we obtain expressions for the dissipation, storage, and transport of energy resulting from the propagation of waves in material media. We consider in some detail the phenomenon of birefringence, in which the phase velocity of a plane wave in a crystal depends on its direction of polarization. The two allowed modes of propagation in uniaxial crystals—the "ordinary" and "extraordinary" rays—are discussed using the formalism of the index ellipsoid.

1.1 Complex-Function Formalism

In problems that involve sinusoidally varying time functions we can save a great deal of manipulation and space by using the complex function formalism. As an example consider the function

$$a(t) = |A| \cos(\omega t + \phi_a) \tag{1.1-1}$$

where ω is the circular (radian) frequency[1] and ϕ_a is the phase. Defining the complex amplitude of $a(t)$ by

$$A = |A|e^{i\phi_a} \tag{1.1-2}$$

we can rewrite (1.1-1) as

$$a(t) = \text{Re}\,[Ae^{i\omega t}] \tag{1.1-3}$$

We will often represent $a(t)$ by

$$a(t) = Ae^{i\omega t} \tag{1.1-4}$$

instead of by (1.1-1) or (1.1-3). This of course is not strictly correct so that when this happens *it is always understood* that what is meant by (1.1-4) is the *real part* of $A \exp(i\omega t)$. In most situations the replacement of (1.1-3) by the complex form (1.1-4) poses no problems. The exceptions are cases that involve the product (or powers) of sinusoidal functions. In these cases we must use the real form of the function (1.1-3). To illustrate the case where the distinction between the real and complex form is not necessary, consider the problem of taking the derivative of $a(t)$. Using (1.1-1) we obtain

$$\frac{da(t)}{dt} = \frac{d}{dt}[|A| \cos(\omega t + \phi_a)] = -\omega|A| \sin(\omega t + \phi_a) \tag{1.1-5}$$

If we use instead the complex form (1.1-4), we get

$$\frac{da(t)}{dt} = \frac{d}{dt}(Ae^{i\omega t}) = i\omega Ae^{i\omega t}$$

Taking, as agreed, the real part of the last expression and using (1.1-2), we obtain (1.1-5).

As an example of a case in which we have to use the real form of the function, consider the product of two sinusoidal functions $a(t)$ and $b(t)$, where

$$\begin{aligned} a(t) &= |A| \cos(\omega t + \phi_a) \\ &= \frac{|A|}{2}[e^{i(\omega t+\phi_a)} + e^{-i(\omega t+\phi_a)}] \\ &= \text{Re}\,[Ae^{i\omega t}] \end{aligned} \tag{1.1-6}$$

and

$$\begin{aligned} b(t) &= |B| \cos(\omega t + \phi_b) \\ &= \frac{|B|}{2}[e^{i(\omega t+\phi_b)} + e^{-i(\omega t+\phi_b)}] \\ &= \text{Re}\,[Be^{i\omega t}] \end{aligned} \tag{1.1-7}$$

---

[1] The radian frequency ω is to be distinguished from the real frequency $\nu = \omega/2\pi$.

with $A = |A| \exp(i\phi_a)$ and $B = |B| \exp(i\phi_b)$. Using the real functions, we get

$$a(t)b(t) = \frac{|A|\,|B|}{2}[\cos(2\omega t + \phi_a + \phi_b) + \cos(\phi_a - \phi_b)] \quad \textbf{(1.1-8)}$$

Were we to evaluate the product $a(t)b(t)$ using the complex form of the functions, we would get

$$a(t)b(t) = ABe^{i2\omega t} = |A|\,|B|e^{i(2\omega t+\phi_a+\phi_b)} \quad \textbf{(1.1-9)}$$

Comparing the last result to (1.1-8) shows that the time-independent (dc) term $\frac{1}{2}|A|\,|B| \cos(\phi_a - \phi_b)$ is missing, and thus the use of the complex form led to an error.

Time-averaging of sinusoidal products.[2] Another problem often encountered is that of finding the time average of the product of two sinusoidal functions of the same frequency

$$\overline{a(t)b(t)} = \frac{1}{T}\int_0^T |A| \cos(\omega t + \phi_a)|B| \cos(\omega t + \phi_b)\,dt \quad \textbf{(1.1-10)}$$

where $a(t)$ and $b(t)$ are given by (1.1-6) and (1.1-7) and the horizontal bar denotes time-averaging. $T = 2\pi/\omega$ is the period of the oscillation. Since the integrand in (1.1-10) is periodic in T, the averaging can be performed over a time T. Using (1.1-8) we obtain directly

$$\overline{a(t)b(t)} = \frac{|A|\,|B|}{2}\cos(\phi_a - \phi_b) \quad \textbf{(1.1-11)}$$

This last result can be written in terms of the complex amplitudes A and B, defined immediately following (1.1-7), as

$$\overline{a(t)b(t)} = \tfrac{1}{2}\,\mathrm{Re}\,(AB^*) \quad \textbf{(1.1-12)}$$

This important result will find frequent use throughout the book.

1.2 Considerations of Energy and Power in Electromagnetic Fields

In this section we derive the formal expressions for the power transport, power dissipation, and energy storage that accompany the propagation of electromagnetic radiation in material media. The starting point is

---

[2] The problem of the time average of two nearly sinusoidal functions is considered in Problems 1.1 and 1.2.

Maxwell's equations (in MKS units)

$$\nabla \times \mathbf{h} = \mathbf{i} + \frac{\partial \mathbf{d}}{\partial t} \tag{1.2-1}$$

$$\nabla \times \mathbf{e} = -\frac{\partial \mathbf{b}}{\partial t} \tag{1.2-2}$$

and the constitutive equations relating the polarization of the medium to the displacement vectors

$$\mathbf{d} = \epsilon_0 \mathbf{e} + \mathbf{p} \tag{1.2-3}$$

$$\mathbf{b} = \mu_0(\mathbf{h} + \mathbf{m}) \tag{1.2-4}$$

where $\mathbf{i}$ is the current density (amperes per square meter); $\mathbf{e}(\mathbf{r}, t)$ and $\mathbf{h}(\mathbf{r}, t)$ are the electric and magnetic field vectors, respectively; $\mathbf{d}(\mathbf{r}, t)$ and $\mathbf{b}(\mathbf{r}, t)$ are the electric and magnetic displacement vectors; $\mathbf{p}(\mathbf{r}, t)$ and $\mathbf{m}(\mathbf{r}, t)$ are the electric and magnetic polarizations (dipole moment per unit volume) of the medium; and ϵ_0 and μ_0 are the electric and magnetic permeabilities of vacuum, respectively. We adopt the convention of using lowercase letters to denote the time-varying functions, reserving capital letters for the amplitudes of the sinusoidal time functions. For a detailed discussion of Maxwell's equations, the reader is referred to any standard text on electromagnetic theory such as, for example, Reference [1].

Using (1.2-3) and (1.2-4) in (1.2-1) and (1.2-2) leads to

$$\nabla \times \mathbf{h} = \mathbf{i} + \frac{\partial}{\partial t}(\epsilon_0 \mathbf{e} + \mathbf{p}) \tag{1.2-5}$$

$$\nabla \times \mathbf{e} = -\frac{\partial}{\partial t}\mu_0(\mathbf{h} + \mathbf{m}) \tag{1.2-6}$$

Taking the scalar (dot) product of (1.2-5) and $\mathbf{e}$ gives

$$\mathbf{e} \cdot \nabla \times \mathbf{h} = \mathbf{e} \cdot \mathbf{i} + \frac{\epsilon_0}{2}\frac{\partial}{\partial t}(\mathbf{e} \cdot \mathbf{e}) + \mathbf{e} \cdot \frac{\partial \mathbf{p}}{\partial t} \tag{1.2-7}$$

where we used the relation

$$\frac{1}{2}\frac{\partial}{\partial t}(\mathbf{e} \cdot \mathbf{e}) = \mathbf{e} \cdot \frac{\partial \mathbf{e}}{\partial t}$$

Next we take the scalar product of (1.2-6) and $\mathbf{h}$:

$$\mathbf{h} \cdot \nabla \times \mathbf{e} = -\frac{\mu_0}{2}\frac{\partial}{\partial t}(\mathbf{h} \cdot \mathbf{h}) - \mu_0 \mathbf{h} \cdot \frac{\partial \mathbf{m}}{\partial t} \tag{1.2-8}$$

Subtracting (1.2-8) from (1.2-7) and using the vector identity

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot \nabla \times \mathbf{A} - \mathbf{A} \cdot \nabla \times \mathbf{B} \tag{1.2-9}$$

results in

$$-\nabla \cdot (\mathbf{e} \times \mathbf{h}) = \mathbf{e} \cdot \mathbf{i} + \frac{\partial}{\partial t}\left(\frac{\epsilon_0}{2}\mathbf{e} \cdot \mathbf{e} + \frac{\mu_0}{2}\mathbf{h} \cdot \mathbf{h}\right) + \mathbf{e} \cdot \frac{\partial \mathbf{p}}{\partial t} + \mu_0 \mathbf{h} \cdot \frac{\partial \mathbf{m}}{\partial t} \tag{1.2-10}$$

We integrate the last equation over an arbitrary volume V and use the Gauss theorem [1]

$$\int_V (\nabla \cdot \mathbf{A})\, dv = \int_S \mathbf{A} \cdot \mathbf{n}\, da$$

where $\mathbf{A}$ is any vector function, $\mathbf{n}$ is the unit vector normal to the surface S enclosing V, and dv and da are the differential volume and surface elements, respectively. The result is

$$\begin{aligned} -\int_V \nabla \cdot (\mathbf{e} \times \mathbf{h})\, dv &= -\int_S (\mathbf{e} \times \mathbf{h}) \cdot \mathbf{n}\, da \\ &= \int_V \left[\mathbf{e} \cdot \mathbf{i} + \frac{\partial}{\partial t}\left(\frac{\epsilon_0}{2}\mathbf{e} \cdot \mathbf{e}\right) + \frac{\partial}{\partial t}\left(\frac{\mu_0}{2}\mathbf{h} \cdot \mathbf{h}\right) + \mathbf{e} \cdot \frac{\partial \mathbf{p}}{\partial t} + \mu_0 \mathbf{h} \cdot \frac{\partial \mathbf{m}}{\partial t}\right] dv \end{aligned} \tag{1.2-11}$$

According to the conventional interpretation of electromagnetic theory, the left side of (1.2-11), that is,

$$-\int_S (\mathbf{e} \times \mathbf{h}) \cdot \mathbf{n}\, da$$

gives the total power flowing *into* the volume bounded by S. The first term on the right side is the power expended by the field on the moving charges, the sum of the second and third terms corresponds to the rate of increase of the vacuum electromagnetic stored energy $\mathcal{E}_{\text{vac}}$ where

$$\mathcal{E}_{\text{vac}} = \int_V \left[\frac{\epsilon_0}{2}\mathbf{e} \cdot \mathbf{e} + \frac{\mu_0}{2}\mathbf{h} \cdot \mathbf{h}\right] dv \tag{1.2-12}$$

Of special interest in this book is the next-to-last term

$$\mathbf{e} \cdot \frac{\partial \mathbf{p}}{\partial t}$$

which represents the power per unit volume expended by the field *on* the electric dipoles. This power goes into an increase in the potential energy stored by the dipoles as well as into supplying the dissipation that may accompany the change in $\mathbf{p}$. We will return to it again in Chapter 5, where we treat the interaction of radiation and atomic systems.

Dipolar dissipation in harmonic fields. According to the discussion in the preceding paragraph, the average power per unit volume expended by the field on the medium electric polarization is

$$\overline{\frac{\text{Power}}{\text{Volume}}} = \overline{\mathbf{e} \cdot \frac{\partial \mathbf{p}}{\partial t}} \tag{1.2-13}$$

where the horizontal bar denotes time-averaging. Let us assume for the sake of simplicity that $\mathbf{e}(t)$ and $\mathbf{p}(t)$ are parallel to each other and take their sinusoidally varying magnitudes as

$$e(t) = \text{Re}\,[Ee^{i\omega t}] \tag{1.2-14}$$

$$p(t) = \text{Re}\,[Pe^{i\omega t}] \tag{1.2-15}$$

where E and P are the complex amplitudes. The electric susceptibility χ_e is defined by

$$P = \epsilon_0 \chi_e E \tag{1.2-16}$$

and is thus a complex number. Substituting (1.2-14) and (1.2-15) in (1.2-13) and using (1.2-16) gives

$$\begin{aligned} \overline{\frac{\text{Power}}{\text{Volume}}} &= \overline{\text{Re}\,[Ee^{i\omega t}]\,\text{Re}\,[i\omega Pe^{i\omega t}]} \\ &= \tfrac{1}{2}\,\text{Re}\,[i\omega\epsilon_0\chi_e EE^*] \\ &= \frac{\omega}{2}\,\epsilon_0|E|^2\,\text{Re}\,(i\chi_e) \end{aligned} \tag{1.2-17}$$

where in going from the first to the second equality we used (1.1-12). Since χ_e is complex we can write it in terms of its real and imaginary parts as

$$\chi_e = \chi_e' - i\chi_e'' \tag{1.2-18}$$

which, when used in (1.2-17), gives

$$\overline{\frac{\text{Power}}{\text{Volume}}} = \frac{\omega\epsilon_0\chi_e''}{2}\,|E|^2 \tag{1.2-19}$$

which is the desired result.

We leave it as an exercise (Problem 1-3) to show that in anisotropic media in which the complex field components are related by

$$P_i = \epsilon_0 \sum_j \chi_{ij} E_j \tag{1.2-20}$$

the application of (1.2-13) yields

$$\overline{\frac{\text{Power}}{\text{Volume}}} = \frac{\omega}{2}\,\epsilon_0 \sum_{i,j} \text{Re}\,(i\chi_{ij}E_i^*E_j) \tag{1.2-21}$$

1.3 Wave Propagation in Isotropic Media

Here we consider the propagation of electromagnetic plane waves in homogeneous and isotropic media so that ϵ and μ are scalar constants. Vacuum is, of course, the best example of such a "medium." Liquids and glasses are material media that, to a first approximation, can be treated as homogeneous and isotropic.[3] We choose the direction of propagation as z and, taking the plane wave to be uniform in the x-y plane, put $\partial/\partial x = \partial/\partial y = 0$ in (1.2-1) and (1.2-2). Assuming a lossless ($\sigma = 0$) medium, (1.2-1) and (1.2-2) become

$$\nabla \times \mathbf{e} = -\mu \frac{\partial \mathbf{h}}{\partial t} \tag{1.3-1}$$

$$\nabla \times \mathbf{h} = \epsilon \frac{\partial \mathbf{e}}{\partial t} \tag{1.3-2}$$

$$\frac{\partial e_y}{\partial z} = \mu \frac{\partial h_x}{\partial t} \tag{1.3-3}$$

$$\frac{\partial h_y}{\partial z} = -\epsilon \frac{\partial e_x}{\partial t} \tag{1.3-4}$$

$$\frac{\partial e_x}{\partial z} = -\mu \frac{\partial h_y}{\partial t} \tag{1.3-5}$$

$$\frac{\partial h_x}{\partial z} = \epsilon \frac{\partial e_y}{\partial t} \tag{1.3-6}$$

$$0 = \mu \frac{\partial h_z}{\partial t} \tag{1.3-7}$$

$$0 = \epsilon \frac{\partial e_z}{\partial t} \tag{1.3-8}$$

From (1.3-7) and (1.3-8) it follows that h_z and e_z are both zero; therefore, a uniform plane wave in a homogeneous isotropic medium can have no longitudinal field components. We can obtain ,a self-consistent set of equations from (1.3-3) through (1.3-8) by taking e_y and h_x (or e_x and h_y) to be zero.[4] In this case the last set of equations reduces to Equations (1.3-4) and (1.3-5). Taking the derivative of (1.3-5) with respect to z and

[3] The individual molecules making up the liquid or glass are, of course, anisotropic. This anisotropy, however, is averaged out because of the very large number of molecules with random orientations present inside a volume $\sim\lambda^3$.

[4] More fundamentally it can be easily shown from (1.3-1) and (1.3-2) (see Problem 1.4) that, for uniform plane harmonic waves, **e** and **h** are normal to each other as well as to the direction of propagation. Thus, **x** and **y** can simply be chosen to coincide with the directions of **e** and **h**.