Chapter 12: Plane waves (1): Propagation, polarization, and superposition. (19 Feb 2022)

A. Perspective.

Up to now we have considered phenomena that either did not depend on time or involved only a first derivative with respect to time. We now investigate wave phenomena, which involve a second derivative with respect to time. In this case the solutions of Maxwell's Equations are plane waves $e^{i\vec{k}\cdot\vec{r}-i\omega t}$, so this can also be described as applied Fourier analysis. Although c appears as c^{-2}, wave phenomena are more accurately described as c^{-1} retardation physics than c^{-2} relativistic physics, because the laboratory time frame applies throughout. Special relativity does not enter until time dilation needs to be considered.

While this level of approximation does not allow synchrotron radiation to be treated, the topic is nevertheless broad enough to keep us busy for the next five chapters. As usual, the applied-physics rules remain: get the math right, determine the relevant parameters, then extract all the physics from the math. In this chapter we cover propagation, polarization, and superposition, both coherent and incoherent. Topics evolve with increasing degrees of complexity, starting with the propagation of a single wave through a material, which covers crystal optics, then polarization and the elementary interferometry physics underlying LIGO, where the two field directions share a common phase factor. The next degree of complexity is coherent *multiwavelength* superposition, which underlies short-pulse optics and therefore the physics underlying information transfer on the Internet. This is followed with incoherent multiwavelength superposition, which describes the intensity-based world that we perceive. We end by including a short digression into collective effects such as electromagnetically induced transparency, which is usually covered only in treatments of atomic and molecular optics (AMO), although the mathematics is classical.

Chapter 13 deals with phenomena associated with planar interfaces between dissimilar materials, that is, reflection, transmission, thin-film interference, and interface plasmons. Metamaterials, including negative-index materials, are covered here as well. The topic of Ch. 14 is lateral confinement, seen for example in waveguides and optical fibers. The

category also includes some aspects of subwavelength imaging, which was recently recognized by a Nobel Prize in Chemistry. It also includes Gaussian and vortex beams, which can be considered self-confined, and are becoming increasingly significant in technology. Chapter 15 covers radiation and scattering. Diffraction gets its own chapter, 16.

Before getting into details, it is worth providing an overview of the approaches that we take, and why we take them. Up to now we have focused nearly exclusively on potentials rather than fields, but Chs. 12, 13, and 14, are based on fields, described specifically by the homogeneous wave equation

$$\nabla(\nabla\cdot\vec{E})-\nabla^2\vec{E}+\frac{1}{c^2}\frac{\partial^2(\varepsilon\vec{E})}{\partial t^2}=0\,. \tag{12.1}$$

In Chs. 15 and 16 we return to potentials, working with the inhomogeneous wave equation

$$\left(\nabla^2-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\left(\phi,\vec{A}\right)=-\frac{4\pi}{c}\left(c\rho,\vec{J}\right). \tag{12.2}$$

The reasons follow from both mathematics and applications. In Eq. (12.1) the coordinate system is Cartesian and fields are planar, having the forms $\vec{E}(\vec{r},t)=\vec{E}_o e^{i\vec{k}\cdot\vec{r}-i\omega t}$ and $\vec{H}(\vec{r},t)=\vec{H}_o e^{i\vec{k}\cdot\vec{r}-i\omega t}$, whereas in Eq. (12.2) both coordinate system and waves are spherical. With no source terms Eq. (12.1) is ideally suited to describe how waves propagate, including reflection and transmission at planar interfaces, where boundary conditions are written in terms of fields not potentials. In contrast, with source terms Eq. (12.2) is ideally suited to describe how waves originate, including not only radiation but also diffraction, which by Huygens' Principle is just another form of radiation.

Elaborating, Eq. (12.1) contains a material property, the dielectric function ε. This can be anisotropic, allowing applications to crystal optics, an important consideration in a universe where the vast majority of materials are anisotropic. It also provides partial entry in the much broader topic of metamaterials, where the permeability μ is generally anisotropic as well. You recall from Ch. 1 that Eq. (12.2) cannot be generalized in this way, so its applications are restricted to empty space. Consistent with these applications, Eq. (12.1) has no simple Green function, whereas the Green function for Eq. (12.2),

$$G(\vec{r},\vec{r}\,',t,t')=\frac{\delta(t-t'-|\vec{r}-\vec{r}\,'|/c)}{|\vec{r}-\vec{r}\,'|}, \tag{12.3}$$

can be written down by inspection (although its mathematical proof involves two pages of fairly dense math, see Ch. 15.) Consequently, Eqs. (12.2) and (12.3) form the natural basis for the description of radiation, which also gets us entry into the physics of retardation. In earlier works (including Jackson) heroic efforts were made to describe radiation with fields, but it makes as much sense to do radiation with fields as to do classical mechanics with algebra.

Despite these fundamental differences, the two approaches converge at the locally flat approximation,

$$e^{ikr-i\omega t} \cong e^{i\vec{k}\cdot\vec{r}-i\omega t}, \tag{12.4}$$

which is sufficiently accurate if the distance from the source is large compared to the lateral dimension of the region being considered.

We begin by deriving the dispersion equations for $\vec{E}$ for general values of μ and ε. These are the equations that must be satisfied if a wave is to propagate in a material. In conventional treatments μ is usually assumed to be equal to 1, but the world has since moved on: $\mu \neq 1$ for metamaterials. Hence the more general case must be considered. Limitations imposed by tensorial μ and ε are noted. Attenuation and energy flow are shown to be consistent with causality in that both occur only in the propagation direction. Discussion includes propagation in dielectrically anisotropic materials (crystal optics), where polarization states are determined by crystal symmetry.

This chapter also overlaps parts of Jackson's Chs. 6 and 7. Much of Jackson's Ch. 7 deals with dielectric response, which we covered in the first semester. However, many developments have occurred since the 3rd Edition appeared. Our recent capabilities to engineer materials on a nanometer scale allow us to create metamaterials for which $\mu \neq 1$, even though much of the work is currently being done in the microwave region where wavelengths are compatible with machine-shop capabilities for manufacturing structures. While many consider metamaterials a relatively recent development, integrated-circuits technology has been working with "metamaterials" ever since feature sizes dropped below visible-near uv optical wavelengths, specifically the 180 nm process node in 1999 (the production mode is currently 5 nm, with 3 nm under investigation in laboratories.). Optical metrology, which is the determination of the properties of materials, thin, films, and nanostructures by measuring and analyzing light scattered from lithographically patterned structures, is now indispensable in this area.

B. Basics.

As usual, the foundation is Maxwell's Equations, in this case the macroscopic versions. Chapters 12-14 are based on the homogeneous equations:

$$\nabla\cdot\vec{D} = 0; \qquad \nabla\cdot\vec{B} = 0; \tag{12.5a,b}$$

$$\nabla\times\vec{E} + \frac{1}{c}\frac{\partial\vec{B}}{\partial t} = 0; \qquad \nabla\times\vec{H} - \frac{1}{c}\frac{\partial\vec{D}}{\partial t} = 0; \tag{12.5c,d}$$

supplemented by the constitutive relations

$$\vec{D} = \varepsilon\vec{E}; \quad \vec{B} = \mu\vec{H}, \tag{12.6a,b}$$

where in general ε and μ are tensorial. As discussed in Ch. 1, these equations give the complete framework for propagation of electromagnetic fields in materials. We will consider only plane-wave solutions with fields $\sim e^{i\vec{k}\cdot\vec{r}-i\omega t}$, so finite conductivities σ are incorporated in ε as

$$\varepsilon \to \varepsilon + \frac{4\pi i\sigma}{\omega}. \tag{12.7}$$

Being a homogeneous set, a perfectly good solution to Eqs. (12.5) is

$$\vec{D} = \vec{E} = \vec{B} = \vec{H} = 0 . \tag{12.8}$$

However, this is not particularly useful. Accordingly, we seek more substantive results. These fall automatically into the eigenvalue/eigenvector category, where only waves with specific eigenvalues $|\vec{k}|$ and eigenvectors $\vec{E}_{\vec{k}}$ and $\vec{H}_{\vec{k}}$ can propagate in a given material. Our objectives here are to develop methods by which these eigenvalues and eigenvectors can be obtained, then examine some of their properties.

Although solutions could be based on $\vec{D}$ and $\vec{B}$ instead of $\vec{E}$ and $\vec{H}$, the latter two are more convenient. When we discuss non-normal-incidence reflection we find that field components tangential to the surface lead to the most efficient solutions, and the fields that satisfy the tangential conditions are $\vec{E}$ and $\vec{H}$. Moreover, the Poynting vector $\vec{S}$ is written in terms of $\vec{E}$ and $\vec{H}$, so the boundary conditions ensuring continuity of tangential $\vec{E}$ and $\vec{H}$ also ensure that energy cannot accumulate at interfaces. For propagation we generally work with $\vec{E}$ rather than $\vec{H}$ because, as noted below, the equations for $\vec{E}$ allow ε to be tensorial, whereas those for $\vec{H}$ do not. This is an important consideration in the optical-frequency range in a world where μ is usually equal to 1, but 95% of all crystals are optically anisotropic. However, in treating reflectance in the next chapter, we base calculations on both $\vec{E}$ and $\vec{H}$, the former for describing transverse-electric (TE) modes and the latter for the transverse-magnetic (TM) case.

C. Dispersion equation.

All analyses start with dispersion equations. Assuming that a plane wave is propagating in a material of permeability μ and dielectric function ε, we take the curl of the Faraday-Maxwell Equation, obtaining

$$\nabla \times (\nabla \times \vec{E}) + \frac{1}{c}\frac{\partial}{\partial t}(\nabla \times \vec{B})$$

$$= \nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E} + \frac{1}{c}\frac{\partial}{\partial t}(\nabla \times \mu \vec{H}) = 0 . \tag{12.9}$$

The next step is to invoke the Ampère-Maxwell Equation, but we face a problem. If μ is tensorial this cannot be done directly because $\nabla \times (\underset{\sim}{\mu} \cdot \vec{H}) \neq \underset{\sim}{\mu} \cdot \nabla \times \vec{H}$. But if we assume that μ is a scalar, it can be removed from the curl operation leaving

$$\nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E} + \frac{\mu}{c}\frac{\partial}{\partial t}(\nabla \times \vec{H}) = 0 . \tag{12.10}$$

We are now free to substitute Ampère's Equation, and the result is

$$\nabla(\nabla \cdot \vec{E}) - \nabla^2 \vec{E} + \frac{\omega^2 \mu}{c^2}\underset{\sim}{\varepsilon} \cdot \vec{E} = 0 . \tag{12.11}$$

Note that this development allows $\varepsilon = \underset{\sim}{\varepsilon}$ to be tensorial. For this reason, most of what follows is based on Eq. (12.11). Had we started by taking the curl of the Ampère-Maxwell Equation, we would have obtained

$$\nabla(\nabla \cdot \vec{H}) - \nabla^2 \vec{H} + \frac{\omega^2 \varepsilon}{c^2} \underset{\sim}{\mu} \cdot \vec{H} = 0 . \tag{12.12}$$

This would allow $\underset{\sim}{\mu}$ to be tensorial, but would restrict ε to be a scalar. If $\underset{\sim}{\mu}$ and $\underset{\sim}{\varepsilon}$ are both tensors, as is the common situation in metamaterials, then neither of the above simplifications is possible and separate equations for $\vec{E}$ and $\vec{H}$ cannot be obtained. Maxwell's Equations must then be used in their entirety. While these remain eigenvalue/eigenvector calculations, they are much more difficult.

We now consider plane-wave space and time dependences explicitly. Then $\nabla \to i\vec{k}$ and $\frac{\partial}{\partial t} \to -i\omega$, casting the original differential equations into the Fourier domain and turning the derivative operations into vector algebra. Working with Eq. (12.11) from now on, the plane-wave substitution reduces it to the *dispersion equation*

$$\vec{k}(\vec{k} \cdot \vec{E}) - \vec{k}^2 \vec{E} + \frac{\omega^2 \mu}{c^2} \underset{\sim}{\varepsilon} \cdot \vec{E} = 0 . \tag{12.13}$$

While $\vec{E} = 0$ is a perfectly good solution, it is not very informative so we seek an alternative. Writing Eq. (12.13) as

$$\vec{k}(\vec{k} \cdot \vec{E}) + \frac{\omega^2 \mu}{c^2} \underset{\sim}{\varepsilon} \cdot \vec{E} = \vec{k}^2 \vec{E} , \tag{12.14}$$

we realize that this is a standard eigenfunction/eigenvalue equation, where the operator on the left must yield the original vector (the eigenvector) times a constant (the eigenvalue). Equation (12.13) is the foundation of most of what follows.

Regarding the parameters in Eqs. (12.13), in practice ω is defined by the incident radiation, μ and $\underset{\sim}{\varepsilon}$ by the material in which the wave propagates, and the direction $\hat{k}$ of $\vec{k} = k\hat{k}$ either specified or established by boundary conditions. This leaves *k* as the parameter to be determined, after which the eigenvectors $\vec{E}$ follow. Once a coordinate system is selected, it is important to realize that Eq. (12.14) produces an eigenvalue equation for *k* no matter what direction is specified for $\vec{k}$ or whether the coordinate system conforms to the principal axes of the dielectric tensor. We give several examples below.

The imaginary part k_i of $k = k_r + ik_i$ describes how the wave attenuates, and the real part k_r the wavelength in the material. We show below that attenuation and net energy flow occur only in the propagation direction, as required by energy conservation and causality The refractive index associated with the eigenmode is $n = ck/\omega$. This is also generally separated into real and imaginary parts $n = n_r + in_i$, where n_r is termed the

"ordinary" index of refraction and $n_i = \kappa$ is the extinction coefficient. The term "dispersion equation" is used because *n* is necessarily a function of ω. This terminology will become more meaningful in Sec. G, where we find that dispersion is responsible for the broadening of short optical pulses as they propagate, thereby limiting for example the distance between repeater stations in optical-fiber communications systems.

The equivalent equation for $\vec{H}$, assuming ε is a scalar, is

$$\vec{k}(\vec{k}\cdot\vec{H}) - \vec{k}^2\vec{H} + \frac{\omega^2\varepsilon}{c^2}\underset{\sim}{\mu}\cdot\vec{H} = 0. \tag{12.15}$$

In either case the equations needed to calculate $\vec{H}$ from $\vec{E}$ and *vice versa* are the Faraday- and Ampère-Maxwell Equations:

$$\vec{H} = \frac{-ic}{\mu\omega}\nabla\times\vec{E}\,; \quad \vec{E} = \frac{ic}{\omega\varepsilon}\nabla\times\vec{H}\,; \tag{12.16a,b}$$

or in plane-wave form

$$\vec{H} = \frac{c}{\mu\omega}\vec{k}\times\vec{E}\,; \quad \vec{E} = -\frac{c}{\omega\varepsilon}\vec{k}\times\vec{H}\,. \tag{12.17a,b}$$

In the latter case it is important to take note of the propagation direction. The symmetry of the above equations with respect to $\vec{E}$ and $\vec{H}$, and ε and μ, is obvious.

In standard textbook derivations it is customary to discard the first terms in Eqs. (12.11) and (12.12) on the grounds that waves are transverse, that is, $\vec{k}\cdot\vec{E} = 0$ and $\vec{k}\cdot\vec{H} = 0$. While this is correct for isotropic materials, where ε is a scalar, this assumption is not general. The divergence equations give the general expressions, and show that

$$\vec{k}\cdot\vec{D} = 0\,; \quad \vec{k}\cdot\vec{B} = 0\,. \tag{12.18a,b}$$

Thus $\vec{D}$ and $\vec{B}$, not $\vec{E}$ and $\vec{H}$, are perpendicular to $\vec{k}$. If we discard the $\vec{k}(\vec{k}\cdot\vec{E})$ and $\vec{k}(\vec{k}\cdot\vec{H})$ terms we discard crystal optics, which is unacceptable in a graduate-level course. Examples involving crystal optics are provided below, and a double-refraction example is given as a homework assignment.

D. Dispersion equation: general solution and properties.

From now on we restrict attention to Eq. (12.13), and assume that μ is a scalar unless noted otherwise. We also restrict our discussion to Cartesian coordinates. Solutions proceed in either of two ways. First, unless a material is optically active or magnetic fields are present, dielectric tensors are diagonal. We take advantage of this by defining a local coordinate system where the *x*, *y*, and *z* axes are aligned with these principal axes. The direction $\hat{k}$ of $\vec{k}$ is then defined relative to these axes. Alternatively, we can define the local coordinate system to place $\hat{k}$ along one axis, usually $\hat{z}$, and rotate $\underset{\sim}{\varepsilon}$ accordingly. For uniaxial or biaxial materials, this generates off-diagonal elements in the

dielectric tensor. Although examples of both are provided below, we start with the simple case of an isotropic material and $\vec{k} = k\hat{z}$ to establish general properties, then elaborate from there.

With the material parameters identified and the direction of $\vec{k}$ specified, the solution of Eq. (12.13) proceeds as follows. Because we have no *a priori* knowledge about the eigenvectors, we make the most general assumption

$$\vec{E} = \hat{x}E_x + \hat{y}E_y + \hat{z}E_z \,. \tag{12.19}$$

Because in this example we assume that ε and μ are scalars, then

$$\underset{\sim}{\varepsilon} \cdot \vec{E} = \hat{x}\varepsilon E_x + \hat{y}\varepsilon E_y + \hat{z}\varepsilon E_z \,. \tag{12.20}$$

Equation (12.5) then becomes

$$k\hat{z}(kE_z) - k^2(\hat{x}E_x + \hat{y}E_y + \hat{z}E_z) + \frac{\omega^2}{c^2}(\hat{x}\varepsilon E_x + \hat{y}\varepsilon E_y + \hat{z}\varepsilon E_z) = 0 \,. \tag{12.21}$$

This is three equations in one, one for each coordinate direction. These are most conveniently represented in matrix form, starting with the $\hat{x}$ direction and ending with $\hat{z}$:

$$\begin{pmatrix} \frac{\omega^2}{c^2}\varepsilon - k^2 & 0 & 0 \\ 0 & \frac{\omega^2}{c^2}\varepsilon - k^2 & 0 \\ 0 & 0 & \frac{\omega^2}{c^2}\varepsilon \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} = 0 \,. \tag{12.22}$$

One solution is obviously

$$k^2 = \frac{\omega^2}{c^2}\varepsilon; \;\; E_x \text{ and } E_y \text{ arbitrary}; \;\; E_z = 0. \tag{12.23}$$

This is a transverse wave. Because E_x and E_y can independently assume arbitrary values but share the plane-wave factor $e^{\pm ikz - i\omega t}$, any combination $(\hat{x}E_x + \hat{y}E_y)$ is preserved indefinitely during propagation. We call this primitive superposition a polarization state. How it is manipulated and applied is discussed below.

Bringing μ back into the picture, the complete solution Eq. (12.23) is

$$k = \pm\omega\sqrt{\mu\varepsilon}\,/c \,. \tag{12.24}$$

The $\pm$ sign is the direction of propagation. Thus a solution that propagates in one direction has a companion solution that propagates in the opposite direction. Additional physics follows from the phase of the remaining part $\omega\sqrt{\mu\varepsilon}\,/c$ of Eq. (12.24). This is

$$\arg(k) = \arg(\sqrt{\mu\varepsilon}) = \frac{1}{2}\arg(\mu) + \frac{1}{2}\arg(\varepsilon)\,. \tag{12.25}$$

Now by causality, in any passive material $0 \le \arg(\mu), \arg(\varepsilon) \le \pi$. Hence $0 \le \arg(k) \le \pi$. Thus $k_i \ge 0$, and taking the positive sign (propagation in the positive *z* direction), it follows that attenuation, if it occurs, is in the propagation direction, as expected for passive material.

Using either Eq. (12.16a) or (12.16d) and the definition of the Poynting vector, it can be shown that the net energy flow must also be in the propagation direction. The calculation is straightforward, and is left as a homework assignment. As a word of caution, as in any nonlinear expression, the real projections of $\vec{E}$ and $\vec{H}$ must be taken before the Poynting vector or energy densities are evaluated.

The critical observer will also have noticed that Eq. (12.22) has a third solution: if $\varepsilon = 0$ then $E_x = E_y = 0$ but E_z is allowed to be nonzero. This longitudinal solution is a plasmon, and can only occur at specific frequencies. If the term $\nabla(\nabla \cdot \vec{E})$ had been discarded, this solution would have been missed.

The math associated with tensorial dielectric functions and arbitrary directions of $\vec{k}$ is not difficult but requires some bookkeeping. The approach least likely to cause errors begins by writing the general expression for $\vec{D}$:

$$\vec{D} = \begin{pmatrix} D_x \\ D_y \\ D_z \end{pmatrix} = \begin{pmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{xy}^* & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{xz}^* & \varepsilon_{yz}^* & \varepsilon_{zz} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}, \tag{12.26}$$

accepting for the moment that $\underset{\sim}{\varepsilon}$ is Hermitian and that off-diagonal elements may be present. Next, define the propagation vector $\vec{k}$ in the same coordinate system. Third, introduce the new terms brought in by the dispersion equation. The result is the dispersion equation in matrix form. Finally, set the determinant of the matrix equal to zero and solve for the values *k* of $\vec{k} = k\hat{k}$. The results are the eigenvalues. Substituting these back into the matrix equation then gives the associated eigenvectors.

For example, if $\vec{k} = k\hat{z}$ the eigenvalues $n^2 = c^2k^2/\omega^2$ of Eq. (12.26) are given by solving

$$\det \begin{pmatrix} \varepsilon_{xx} - n^2 & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{xy}^* & \varepsilon_{yy} - n^2 & \varepsilon_{yz} \\ \varepsilon_{xz}^* & \varepsilon_{yz}^* & \varepsilon_{zz} \end{pmatrix} = 0\,. \tag{12.27}$$

The eigenvalues n_1^2 and n_2^2 of the "transverse" solutions are found in the usual manner by setting the determinant of the matrix equal to zero, noting in the general case the fields that are transverse are $\vec{D}$ and $\vec{B}$, not $\vec{E}$ and $\vec{H}$. The third ("longitudinal") solution

occurs only for specific values of the elements of the $\underset{\sim}{\varepsilon}$ tensor, usually at a single frequency.

We next consider general characteristics that will be important in applications later in this chapter. First, the equations are linear, so solutions can be added to form new solutions having different properties. Short-pulse optics depends on this, where plane waves of different frequencies are superposed coherently at a given place and time to approximate delta-function behavior.

Second, the real and imaginary projections of the solutions are also solutions. While complex functions such as $e^{i\vec{k}\cdot\vec{r}-i\omega t}$ are far more efficient in evaluating potentials and fields than their real and imaginary counterparts, it is well known that real projections must be taken when reducing results to observables. Thus for an electric field propagating in an isotropic material, it and its associated magnetic field

$$\vec{E}(z,t) = \hat{x}E_o e^{ikz-i\omega t}\ ,\ \vec{H}(z,t) = \hat{y}nE_o e^{ikz-i\omega t}\ , \tag{12.28a,b}$$

reduce to the observables

$$\vec{E}_{obs}(z,t) = \hat{x}E_o\cos(kz-\omega t)\ ,\ \vec{H}_{obs}(z,t) = \hat{y}nE_o\cos(kz-\omega t)\ . \tag{12.29a,b}$$

Third, the simple example give above illustrates some points that are actually general: the solution of the dispersion equation always yields a pair of values n^2 and the normal modes that go with them. Reflectance configurations are somewhat more complicated in that the propagation direction in the material may need to be determined as part of the overall solution. We will consider this more complicated situation later.

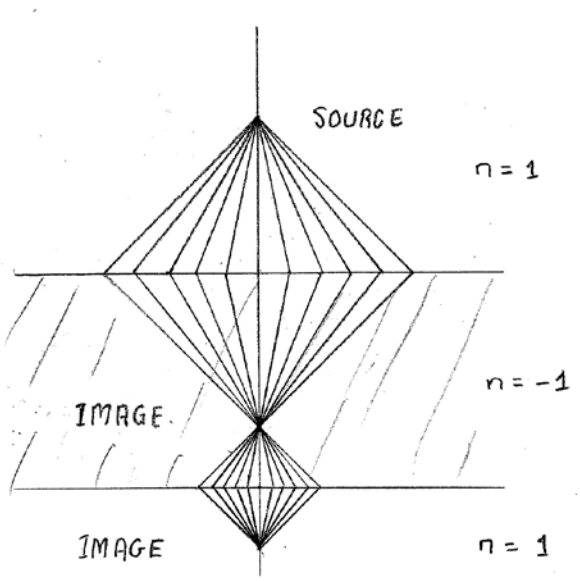
A comment on negative-index materials, which is a topic of much current interest in physics. A negative-index material is that for which $k_r < 0$. In such materials propagation results in the accumulation of negative phase in the propagation direction, which can be used to cancel positive accumulations in other media. Among the resulting bizarre effects is the Veselago lens, which is shown in the diagram at the top of the next page. Consider two materials, one with $n_1 = 1$ (air) and the other with $n_2 = -1$ (a metamaterial). From Snell's Law

$$n_1\sin\theta_1 = n_2\sin\theta_2\ . \tag{12.30}$$

where θ_1 is the angle of incidence and θ_2 is the angle of refraction. Rays leaving a point source in the upper part of the figure exhibit *negative* refraction at the air-material interface, giving rise to a focus *inside* the material even though the interface is a plane, as shown in the figure. If the material is sufficiently thick, then a second focus appears on the other side.

These negative-index materials do not exist in nature, but metamaterials that exhibit this property can be made in the laboratory. The Russian theorist Veselago predicted this effect in 1964, but his prediction was so far out of step with the topics of interest at the time that his paper was not even translated into English until 1968. Negative-index materials became a hot topic in the 2000, when Pendry published a paper that nominally proved (incorrectly) that the "Veselago lens" focused light on a subwavelength scale. Pendry had simply rediscovered the coupled interface plasmons that Raether had

described 20 years earlier. (You will examine their properties in a homework assignment later this semester.) As of 02 Feb 2019 the Veselago and Pendry papers have been cited 7143 and 7875 times , respectively. Veselago (now deceased) was for a time a member of the physics faculty of Princeton, which appeared to be betting that a Nobel Prize was on the horizon.



The Veselago lens has a focus only in the sense that propagation in the negative-index material reverses the phase that accumulates in propagation through the positive-index material, yielding a point in the material where all rays exhibit the same phase, namely zero. However, this focus does not satisfy the more rigorous Abbé criterion, which defines a focus as a point where all rays arrive at the same time. A Veselago lens does not have this property and hence cannot be used to focus short-pulse radiation.

E. Examples; anisotropic materials, normal modes, and crystal optics.

This section elaborates on the introduction to the topic given in Secs. C and D, and makes a short reference to quantum mechanics at the end. As a review, at the microscopic level ε is defined as

$$\vec{D} = \varepsilon \vec{E} = \vec{E} + 4\pi \vec{P} = \vec{E} + 4\pi \chi_E \vec{E} = \vec{E} + 4\pi n q \Delta\vec{r} \,, \tag{12.31}$$

where χ_E is the electric susceptibility, and in the classical approach used in Ch. 7, $\Delta\vec{r}$ is the solution of the force equation for a particular class of charge. A tensorial response results if the material is anisotropic on a macroscopic scale, such that the response $\Delta\vec{r}_j$ summed over the different bonds *j* in a crystal is not necessarily parallel to the applied field $\vec{E}$. Although macroscopic anisotropy originates on the atomic scale, a material may be isotropic on the macroscopic scale even though it is anisotropic on an atomic scale. A good example is Si, where each atom is tetrahedrally bonded to four nearest neighbors. However, when the individual bond responses are summed over a unit cell, the resulting macroscopic symmetry is cubic.

As the first example, consider a material that has different dielectric-tensor components along each of its principal axes:

$$\vec{\vec{\varepsilon}} = \begin{pmatrix} \varepsilon_a & 0 & 0 \\ 0 & \varepsilon_b & 0 \\ 0 & 0 & \varepsilon_c \end{pmatrix}. \tag{12.32}$$

For propagation in the *z* direction, the eigenvalues and normal modes are determined from

$$\begin{pmatrix} \varepsilon_a - n^2 & 0 & 0 \\ 0 & \varepsilon_b - n^2 & 0 \\ 0 & 0 & \varepsilon_c \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} = 0 . \tag{12.33}$$

We now have two values $n^2 = n_a^2$ and $n^2 = n_b^2$, which are the eigenvalues of the electric fields linearly polarized in the *x* and *y* directions, respectively. Since $k_a^2 \neq k_b^2$ the normal modes $\hat{x}E_x$ and $\hat{y}E_y$ propagate at different speeds, so in this case a given state of polarization is not preserved except in the trivial case where only one normal mode is involved. This example shows that the important quantity regarding n^2 is not the propagation direction, as we might have expected, but the *field* direction. Optically uniaxial materials, where $\varepsilon_a = \varepsilon_b \neq \varepsilon_c$, are a subclass of anisotropic materials that includes calcite, MgF_2, and crystal quartz, all of which are commonly used in polarizers.

The second example involves a tensorial dielectric function. Consider a material for which $\underset{\sim}{\varepsilon}$ in the lab frame contains off-diagonal elements:

$$\underset{\sim}{\varepsilon} = \begin{pmatrix} \varepsilon_a & \Delta\varepsilon & 0 \\ \Delta\varepsilon & \varepsilon_a & 0 \\ 0 & 0 & \varepsilon_c \end{pmatrix} . \tag{12.34}$$

We continue to assume $\vec{k} = k\hat{z}$. The matrix equation is now

$$\begin{pmatrix} \varepsilon_a - n^2 & \Delta\varepsilon & 0 \\ \Delta\varepsilon & \varepsilon_a - n^2 & 0 \\ 0 & 0 & \varepsilon_c \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} = 0 . \tag{12.35}$$

This equation cannot be solved by inspection. However, setting $\det\{\vec{M}\} = 0$ yields

$$\left((\varepsilon_a - n^2)^2 - \Delta\varepsilon^2\right)\varepsilon_c = 0 . \tag{12.36}$$

The solution of Eq. (12.31) is

$$n_{\pm}^2 = \varepsilon_a \pm \Delta\varepsilon . \tag{12.37}$$

The normal mode for $n_+^2 = \varepsilon + \Delta\varepsilon$ is therefore:

$$\begin{pmatrix} \varepsilon - \varepsilon - \Delta\varepsilon & \Delta\varepsilon \\ \Delta\varepsilon & \varepsilon - \varepsilon - \Delta\varepsilon \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} -\Delta\varepsilon & \Delta\varepsilon \\ \Delta\varepsilon & -\Delta\varepsilon \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix} = 0 , \tag{12.38}$$

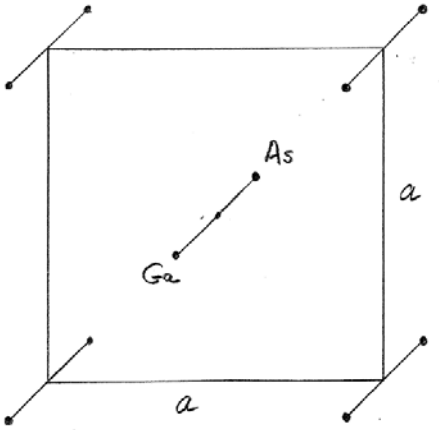
so $E_x = E_y$. In unit-vector form, this normal mode is

$$\vec{E}_+ = \left(\frac{\hat{x} + \hat{y}}{\sqrt{2}}\right) E_+ . \tag{12.39a}$$

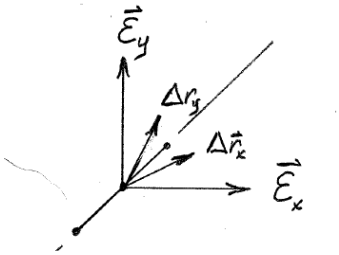
For $n_-^2 = \varepsilon - \Delta\varepsilon$, the result is

$$\vec{E}_- = \left(\frac{\hat{x} - \hat{y}}{\sqrt{2}}\right) E_- . \tag{12.39b}$$

For each mode, the $\hat{x}$ and $\hat{y}$ vectors of the laboratory frame are locked together in the polarization state determined by the value of n^2 for that mode. Consequently, the coefficients $E_\pm$ specifying the amplitudes of the above modes multiply the appropriate linear combinations of $\hat{x}$ and $\hat{y}$, as shown in Eqs. (12.35). The plasmonic mode is defined by $\varepsilon_c = 0$ with propagation in the *z* direction, i.e., a longitudinal excitation propagating parallel to $\vec{k}$.



In this example the principal axes of the material are rotated ±45° with respect to the laboratory coordinate system, as can be shown by standard matrix-diagonalization methods. Had we rotated our local system by 45° to conform to the principal axes of the dielectric tensor, the dispersion matrix would have been diagonal with the eigenvalues determined above. The dielectric tensor of a given material depends not only on the material, but also on its orientation relative to the local coordinate system, even though the eigenvalues themselves are unaffected by the rotation.



A physical realization of this configuration is found on the (001) surface of atomically clean GaAs in ultrahigh vacuum. Depending on conditions the surface atoms form Ga-Ga or As-As dimer bonds that are oriented at 45° relative to the cubic axes in the surface plane. Polarization is achieved more easily along the bond than perpendicular to it, so if the crystallographic axes are aligned to the laboratory frame of reference, the dielectric tensor has the above form. The lower figure indicates schematically the resulting displacements for fields applied in the *x* and *y* directions. While the displacements favor the directions of the applied field, components also appear in the orthogonal direction.

As a third example, consider propagation in the *x* direction in the above material. The matrix equation is now

$$\begin{pmatrix} \varepsilon_a & \Delta\varepsilon & 0 \\ \Delta\varepsilon & \varepsilon_a - n^2 & 0 \\ 0 & 0 & \varepsilon_c - n^2 \end{pmatrix} \begin{pmatrix} E_{ox} \\ E_{oy} \\ E_{oz} \end{pmatrix} = 0 . \tag{12.40}$$

The solutions of the equation $\det\left(\vec{\vec{M}}\right)=0$ are now $n_1^2=\varepsilon_c$ and $n_2^2=\varepsilon_a-(\Delta\varepsilon)^2/\varepsilon_a$. Substituting $n_1^2=\varepsilon_c$ in Eq. (12.40) yields

$$\begin{pmatrix}\varepsilon_a & \Delta\varepsilon & 0\\ \Delta\varepsilon & \varepsilon_a-\varepsilon_c & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}E_{ox}\\ E_{oy}\\ E_{oz}\end{pmatrix}=0. \tag{12.41}$$

This set of equations is obviously satisfied for any value of E_{oz}. Although the equation

$$\varepsilon_a E_{ox}+\Delta\varepsilon\, E_{oy}=0 \tag{12.42}$$

of line 1 appears to give a second mode that connects E_{ox} and E_{oy}, these components must also satisfy the equation on line 2:

$$\Delta\varepsilon\, E_{ox}+(\varepsilon_a-\varepsilon_c)E_{oy}=0. \tag{12.43}$$

When we take the second equation into account, the result is

$$\left(\varepsilon_a(\varepsilon_a-\varepsilon_c)-\Delta\varepsilon^2\right)E_{oy}=0. \tag{12.44}$$

Thus in general $E_{ox}=E_{oy}=0$. This confirms that the mode $n^2=\varepsilon_c$ is fully polarized along *z*. This is what we would expect: the electric field of this mode is orthogonal to the directions containing the off-diagonal elements, so it is unaffected by them.

A more interesting situation occurs with the second mode. By following the above procedure, we verify that $E_{0z}=0$ for this mode and the normalized polarization vector is

$$\hat{e}=\frac{-\hat{x}\,\Delta\varepsilon+\hat{y}\,\varepsilon_a}{\sqrt{\varepsilon_a^2+\Delta\varepsilon^2}}E_o. \tag{12.45}$$

Thus $\vec{E}_o$ is seen to have a component *in* the propagation direction. Although this appears to violate the rule that propagating waves are transverse waves, we recall that it is $\vec{D}$, not $\vec{E}$, that is perpendicular to $\vec{k}$. We can easily confirm this by evaluating

$$\vec{D}=\vec{\vec{\varepsilon}}\cdot\vec{E}_o=\begin{pmatrix}\varepsilon_a & \Delta\varepsilon & 0\\ \Delta\varepsilon & \varepsilon_a & 0\\ 0 & 0 & \varepsilon_c\end{pmatrix}\begin{pmatrix}-\Delta\varepsilon\, E_o\\ \varepsilon_a E_o\\ 0\end{pmatrix} \tag{12.46a}$$

$$=\begin{pmatrix}(-\varepsilon_a\Delta\varepsilon+\varepsilon_a\Delta\varepsilon)E_o\\ (-\Delta\varepsilon^2+\varepsilon_a^2)E_o\\ 0\end{pmatrix}=\begin{pmatrix}0\\ (-\Delta\varepsilon^2+\varepsilon_a^2)E_o\\ 0\end{pmatrix}. \tag{12.46b}$$

The orthogonality of $\vec{D}$ and $\vec{k}$ is therefore proven. The physics behind the math follows from our work in the first semester, where we learned that $\nabla\cdot\vec{D}$ picks up only the *external* charge, which is zero in this case, whereas $\nabla\cdot\vec{E}$ picks up *all* charge, external

and induced. Because the induced charge lies along a line that is not perpendicular to $\vec{k}$, the result follows.

The fact that $\vec{E}$ is not perpendicular to $\vec{k}$ leads to double refraction in optically uniaxial and biaxial materials. In the above example the eigenvector $\vec{E} = \hat{z}E_o$ for $n^2 = \varepsilon_c$ is orthogonal to $\vec{k}$, so the Poynting vector is parallel to $\vec{k}$, In contrast, for the second mode $\vec{S}$ is not parallel to $\vec{k}$, so the intensity "walks off." In visual terms this causes the appearance of a double image for light propagating through the material, with the "ordinary" ray traversing as if the material were isotropic and the "extraordinary" ray following a different direction.

What else can we learn from this? Returning to our definition of ε in Ch. 8,

$$\vec{D} = \vec{\vec{\varepsilon}} \cdot \vec{E} = \vec{E} + 4\pi n q \Delta\vec{r}\,, \tag{12.47}$$

we see that for a given mode the average displacement vector is

$$\Delta\vec{r}_o = \frac{(\vec{D} - \vec{E})}{4\pi n q} = \frac{1}{4\pi n q}(\vec{\vec{\varepsilon}} - \vec{\vec{I}}) \cdot \vec{E}_o\,, \tag{12.48}$$

where $\vec{\vec{I}}$ is the unit tensor. Thus we can reverse-engineer the macroscopic result to obtain the effective microscopic displacement that gives the result.

At this stage we have basically ignored possibilities when both μ and ε are tensorial. If we have total control over the tensor components of both μ and ε, then new phenomena canoccur. Arguably, the most spectacular of these is cloaking, where with proper choices of μ and ε, light will pass around an object as if it does not exist. This was proposed in a Science article by Pendry, Schurig, and Smith in 2006 (J. B. Pendry, D. Schurig, and D. R. Smith, Science **312**, 1780 (2006)), and is illustrated on the next page with a figure from their paper. The theory is based on three-dimensional conformal mapping, noting that a configuration that satisfies Maxwell's Equations in the laboratory frame $\vec{r} = (x, y, z)$ also satisfies it in a mapped frame in $\vec{r}' = (u, v, w)$, where $u = u(x, y, z)$; $v = v(x, y, z)$; $w = w(x, y, z)$. However, the transformation requires the dielectric function, permeability, and fields to be scaled according to

$$\mu'_u = \mu_u \frac{Q_u Q_v Q_w}{Q_u^2}; \quad \mu'_v = \mu_v \frac{Q_u Q_v Q_w}{Q_v^2}; \quad \mu'_w = \mu_w \frac{Q_u Q_v Q_w}{Q_w^2}; \tag{12.49a,b,c}$$

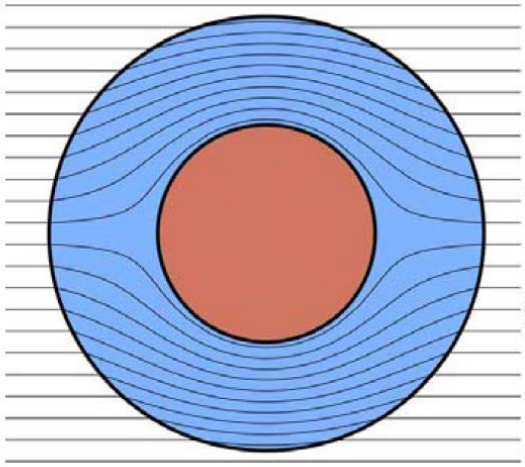
$$\varepsilon'_u = \varepsilon_u \frac{Q_u Q_v Q_w}{Q_u^2}; \quad \varepsilon'_v = \varepsilon_v \frac{Q_u Q_v Q_w}{Q_v^2}; \quad \varepsilon'_w = \varepsilon_w \frac{Q_u Q_v Q_w}{Q_w^2}; \tag{12.49d,e,f}$$

$$E'_u = Q_u E_u; \quad H'_u = Q_u H_u, \text{ etc.;} \tag{12.49g,h}$$

where

$$Q_u^2 = \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2 + \left(\frac{\partial z}{\partial u}\right)^2 \text{ etc.} \tag{12.49i}$$

The authors provide an analytic example where light bends around a sphere and continues unimpeded thereafter. The situation is realized because the refractive index is less than 1 over the longer ray paths, so the wave effectively speeds up there. The important quantity here is the phase: when a given ray emerges from the transformed region, it must have exactly the same phase as if the object were not there and the path were unimpeded. The necessary constraints on phase cannot be achieved with ε alone, but require μ to be modified as well.



It is worth making some comments about equivalent normal-mode calculations in quantum mechanics. The basic equation is the Hamiltonian

$$H_0\varphi_n = \left(\frac{p^2}{2m} + V(\vec{r})\right)\varphi_n = E_n\varphi_n, \tag{12.50}$$

where H_0 is the unperturbed Hamiltonian and φ_n is one of a set of orthonormal eigenfunctions of H_0 with eigenvalue E_n. The E_n and φ_n are assumed to be known.

Either perturbation theory or matrix diagonalization is required when a perturbation term H' is added to the Hamiltonian. We consider diagonalization here. In this case

$$H\Psi = (H_0 + H')\Psi = E\Psi, \tag{12.51}$$

and the goal is to find Ψ and E. We start by writing

$$\Psi = \sum_n a_n\varphi_n . \tag{12.52}$$

Bringing the right-hand term to the left side and applying the orthogonality conditions gives

$$< \varphi_{n'} | \left(H_0 + H' - E\right)\sum_n a_n | \varphi_n > = \left((E_{n'} - E)a_{n'} + \sum_{n'} H_{n'n} = 0\right). \tag{12.53}$$

These equations can be written in matrix form. Considering only the first 4 coefficients, Eqs. (12.45) this become

$$\begin{pmatrix} E_1 - E & H'_{12} & H'_{13} & H'_{14} \\ H'_{21} & E_2 - E & H'_{23} & H'_{24} \\ H'_{31} & H'_{32} & E_3 - E & H'_{34} \\ H'_{41} & H'_{42} & H'_{43} & E_4 - E \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = 0 . \tag{12.54}$$

Higher accuracy obviously means more coefficients and a bigger matrix.

The next steps are the same as in E&M: set the determinant of the matrix equal to zero to determine the eigenvalues E, then substitute these back into the original matrix one at a time to get the coefficients (eigenvectors). Although the procedure is the same, the QM solution can involve an infinite number of equations leading to matrices of infinite dimension. Also, the eigenvalue appears in each diagonal term, not just two. Obviously, in E&M we have it easy.

One loose end remains. For the off-diagonal elements in ε, we have assumed that $\varepsilon_{xy} = \varepsilon_{yx} = \Delta\varepsilon$. Can this assumption be justified? Suppose that the material is transparent, so these terms are real. Now consider the work done in applying a force $\vec{F}$. On the microscopic scale, the result is a displacement $\Delta\vec{r}$. The work done is given by

$$\Delta W = \int_0^{\Delta\vec{r}} \vec{F} \cdot d\vec{l} \ . \tag{12.55}$$

If the work is independent of the path, then it follows that $\vec{F} = -\nabla U$, where U is a potential energy. If the system is linear, we can find a set of axes where the gradient is diagonal:

$$\vec{F} = -\hat{x}\frac{\partial U}{\partial x} - \hat{y}\frac{\partial U}{\partial y} \, . \tag{12.56}$$

Returning to the lab frame, we calculate the effect of a field on the material by first rotating into the coordinate system of the material, apply Eq. (12.52), then rotate back. The result is

$$\begin{pmatrix} D_x \\ D_y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \varepsilon_{xx} & 0 \\ 0 & \varepsilon_{yy} \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix} \tag{12.57a}$$

$$= \begin{pmatrix} \varepsilon_{xx}\cos^2\theta + \varepsilon_{yy}\sin^2\theta & (\varepsilon_{xx} - \varepsilon_{yy)}\sin\theta\cos\theta \\ (\varepsilon_{xx} - \varepsilon_{yy)}\sin\theta\cos\theta & \varepsilon_{xx}\sin^2\theta + \varepsilon_{yy}\cos^2\theta \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix}, \tag{12.57b}$$

and the result is demonstrated. More generally, for complex values the off-diagonal elements in the laboratory frame satisfy $\varepsilon_{xy} = -\varepsilon_{yx}^*$. You found this to be the case when obtaining the dielectric response in a magnetic field.

F. Superposition 1: Polarization and applications; Jones vectors and matrices.

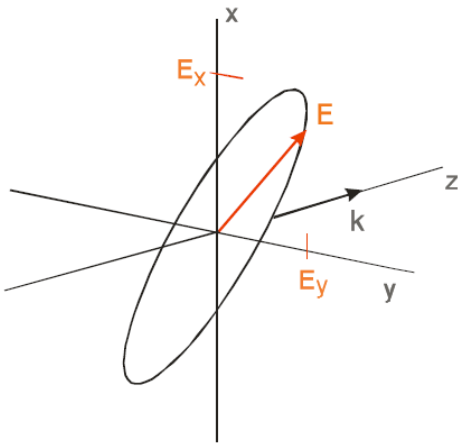
The most elemental form of superposition is that where the *x* and *y* components of a field share a common phase factor. Write $\tilde{E}_x = E_x e^{i\theta_x}$ and $\tilde{E}_y = E_y e^{i\theta_y}$, where in temporary notation $\tilde{E}_x$ and $\tilde{E}_y$ are complex and E_x and E_y are amplitudes. Taking the real projection

$$\begin{aligned} \mathrm{Re}\left(\vec{E}(z,t)\right) &= \mathrm{Re}\left(\vec{x}E_x e^{i\theta_x} + \hat{y}E_y e^{i\theta_y}) \, e^{ikz - i\omega t}\right) \\ &= \hat{x}E_x \cos(kz - \omega t + \theta_x) + \hat{y}E_y \cos(kz - \omega t + \theta_y) \, , \end{aligned} \tag{12.58}$$

it follows that the path $\vec{E}(0,t)$ traced as a function of time is elliptical, as shown in the figure. The shape of the path is given by the complex ratio

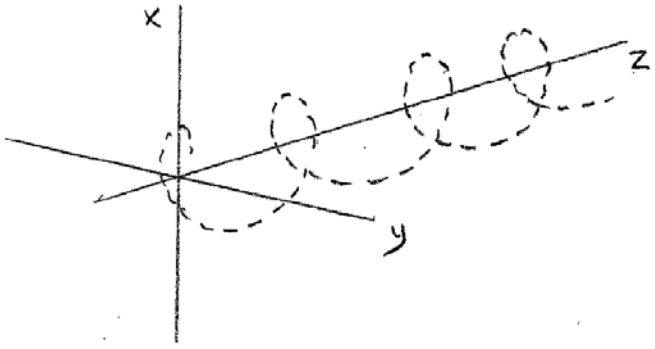
$$\tilde{\chi} = \frac{\tilde{E}_y}{\tilde{E}_x}. \tag{12.59}$$

This is termed the polarization state, and is independent of the magnitude of the wave. Because the two constituent plane waves have the same frequency and in isotropic media are traveling at the same speed and in the same direction, $\tilde{\chi}$ is a fundamental property of the combination. It remains invariant as the wave propagates.



Limiting cases are linear polarization, where $\theta_x = \theta_y$, and circular polarization, where $E_x = E_y$ and $\theta_y = \theta_x \pm \pi/2$. In the former case the trajectory is a straight line of azimuth angle $\varphi = \tan^{-1}(E_y/E_x)$. In the latter case it is a circle of radius $E_x = E_y$. Regarding terminology, elliptical and circular polarizations have positive or negative helicity, or are left- or right-handed, respectively. Helicity is determined by examining $\vec{E}(0,t)$. Positive helicity corresponds to the field rotating in the direction of increasing φ in cylindrical coordinates. This is clockwise when looking in the direction that the wave is propagating. It is also consistent with the quantum-mechanical sign convention for the *z*-axis projection of the angular momentum of a photon. Negative helicity is the opposite. Our viewing convention is different from that used by Jackson, who takes the perspective of a viewer looking into the beam (*never* a good idea!)

Handedness is assessed by considering $\vec{E}(z,0)$. The behavior for positive helicity is shown in the figure. It is that of a screw with left-handed thread. Accordingly, polarization of positive helicity is termed left-handed.



Analysis and control of polarization is important in a wide range of applications. In the optical-frequency range, polarization-conditioning components include polarizers and retarders, also called compensators or wave plates. Changes of the polarization state upon non-normal incidence reflection from a specular surface are a fundamental source of information about processes for control in the fabrication of integrated circuits.

The formal description of polarization is based on Jones calculus, which provides a systematic means of keeping track of the effect of different optical components on a plane wave as it propagates through an optical system. To establish the convention, the Jones vector of the field is

$$\vec{E}(z,t) = \begin{pmatrix} E_x \\ E_y \end{pmatrix} e^{ikz - i\omega t}, \tag{12.60}$$

where E_x and E_y are components of $\vec{E}$ in the local coordinate system. E_x and E_y are once again considered complex.

The action of a component on $\vec{E}(z,t)$ is described by the Jones matrix

$$J = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}. \tag{12.61}$$

The primary actions are polarization, phase shifting, reflection, and rotation. The Jones matrix of a linear polarizer that passes E_x and blocks E_y is

$$J_P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{12.62}$$

as can be verified by direct inspection. The Jones matrix of a compensator is

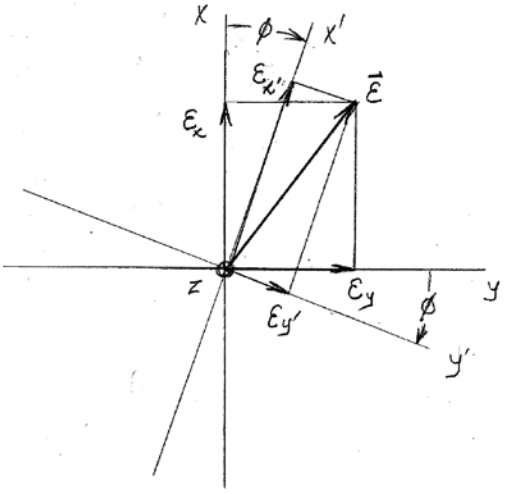
$$J_C = \begin{pmatrix} e^{i\delta_x} & 0 \\ 0 & e^{i\delta_y} \end{pmatrix}, \tag{12.63}$$

thereby describing phase delays δ_x and δ_y as the wave propagates through it. The "fast" axis, defined as $\delta_x < \delta_y$, is generally assigned to *x*. The Jones matrix describing reflection is

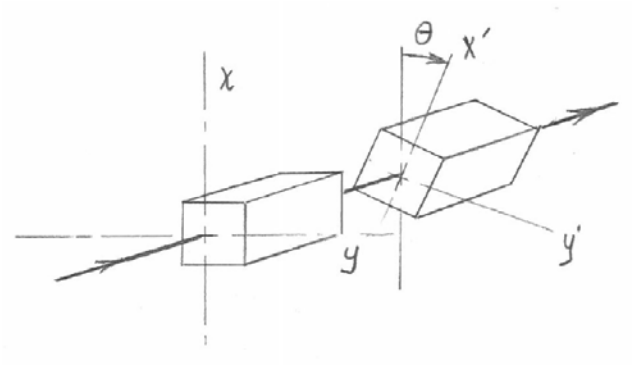
$$J_R = \begin{pmatrix} r_p & 0 \\ 0 & r_s \end{pmatrix}, \tag{12.64}$$

where r_p and r_s are the (field) reflectances for light polarized in parallel (German *parallel*) and perpendicular (German *senkrecht*), respectively, to the plane of incidence. The Jones matrix describing rotation is

$$J_R(\varphi) = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}, \tag{12.65}$$



where the direction of propagation is into the plane of the paper, and the rotation angle φ satisfies the right-hand rule according to the diagram on the right.

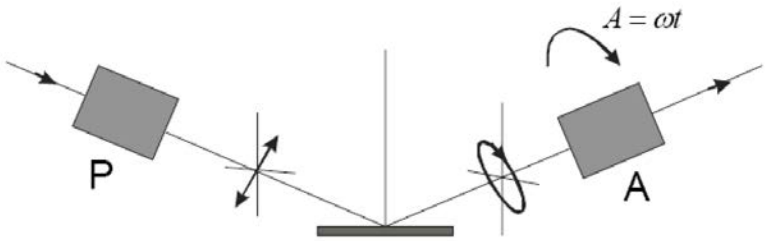


As an example of the use of Jones matrices, we consider Malus' Law, which states that the intensity of initially linearly polarized light transmitted through a polarizer is proportional to the square of the cosine of the angle between the polarization direction and the principal axis of the polarizer. The figure shows two polarizer prisms with the principal axes of the second rotated by θ with respect to those first. The first polarizer ensures that the beam is linearly polarized. The rotation matrix then projects this component E_x into the local coordinate system of the second polarizer, which passes the component E_x and blocks $E_{y'}$. Written as Jones matrices, this operation is

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix}_{out} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix}_{in} \tag{12.66a}$$

$$= \begin{pmatrix} E_x \cos\theta \\ 0 \end{pmatrix} \tag{12.66b}$$

Because $\vec{S} \sim |E^2|$, the proposition is proved.



As a second example of the use of Jones matrices to describe an optical configuration, we consider the rotating-analyzer ellipsometer (RAE) shown at the right. An unpolarized beam enters the configuration from the left, then passes through the polarizer *P* to place it in a definite state of polarization. The principal axis of the polarizer is oriented at an azimuth angle φ_P with respect to the plane of incidence. (All azimuth angles in the following are measured with respect to the plane of incidence.) The emerging field is then projected by a rotation matrix into the reference axes of the reflecting surface. After reflection, the beam passes through a second polarizer oriented at an azimuth angle φ_A relative to the plane of incidence before proceeding to the detector. The description requires a second rotation matrix. The Jones-matrix product for this configuration is therefore

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix}_{out} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \cos\varphi_A & \sin\varphi_A \\ -\sin\varphi_A & \cos\varphi_A \end{pmatrix} \begin{pmatrix} r_p & 0 \\ 0 & r_s \end{pmatrix} \begin{pmatrix} \cos\varphi_P & -\sin\varphi_P \\ \sin\varphi_P & \cos\varphi_P \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix}_{in} . \tag{12.67}$$

Carrying out the multiplications explicitly shows that $E_{y,out} = 0$ and

$$E_{x,out} = \left(r_p \cos\varphi_A \cos\varphi_P + r_s \sin\varphi_A \sin\varphi_P\right) E_{x,in} . \tag{12.68}$$

The expression is symmetric in φ_A and φ_P, as expected, because reversing the direction of the beam simply yields the time-reversed version of the original configuration. Writing $r_p' = r_p \cos\varphi_P$ and $r_s' = r_s \sin\varphi_P$, and taking the absolute square of $E_{x,out}$ and the rest of Eq. (12.68), we obtain the system transfer function, which describes the intensity $I_{x,out}$ that reaches the detector. This is

$$I_{x,out} = \left(\frac{1}{2}(|r_p'|^2 + |r_s'|^2) + \frac{1}{2}(|r_p'|^2 - |r_s'|^2)\cos 2\varphi_A + \mathrm{Re}(r_p' r_s'^{*}) \sin 2\varphi_A \right) I_{x,in} \tag{12.69a}$$

$$= \frac{1}{2}(|r_p'|^2 + |r_s'|^2)\left(1 + \alpha_2 \cos 2\varphi_A + \beta_2 \sin 2\varphi_A\right) I_{x,in} , \tag{12.69b}$$

where $I_{x,in}$ is the source intensity. In Eqs. (12.69) we have assumed that r_p and r_s are complex, which is the usual situation.

In the absence of a calibrated measurement of the transmitted intensity, which is the usual case, information about the reflecting surface is contained in the normalized Fourier coefficients α_2 and β_2. Solving Eqs. (12.69) we find

$$\left|\frac{r_p}{r_s}\right| = \tan\psi = \left(\frac{1+\alpha_2}{1-\alpha_2}\right)^{\frac{1}{2}} \cot P ; \tag{12.70a}$$

$$\mathrm{Re}\left(\frac{r_p}{r_s}\right) = \cos\Delta = \frac{\beta_2}{1-\alpha_2} . \tag{12.70b}$$

The fact that the results are expressed in terms of the angles ψ and Δ is historical: at one time these were angles read off divided circles. Only the cosine projection of Δ is obtained. We could have predicted this, because polarizers do not couple to the circularly polarized component and thus cannot measure it directly. The sine projection and therefore Δ must be inferred from

$$\sin\Delta = \pm\sqrt{1 - \cos^2\Delta} . \tag{12.71}$$

For anisotropic samples such as organics or optically uniaxial or biaxial materials, the fields of the normal modes are not necessarily in or perpendicular to the plane of incidence, so the Jones matrix for reflection has off-diagonal elements as well:

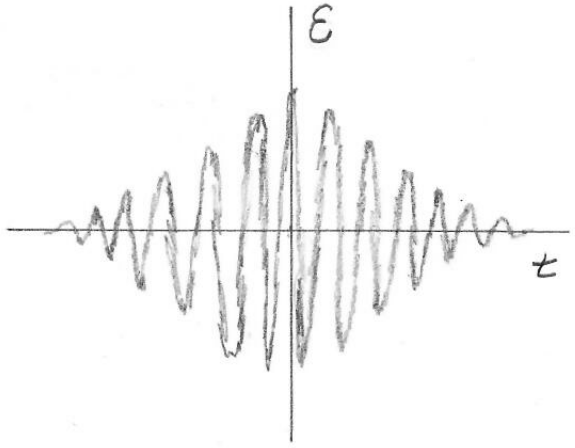
$$R = \begin{pmatrix} r_{pp} & r_{ps} \\ r_{sp} & r_{ss} \end{pmatrix} . \tag{12.72}$$

More complicated configurations involving compensators as well as polarizers are needed to analyze materials of this complexity. Study of these configurations will be left as problems.

The above relates only peripherally to Jackson's treatment of polarization in Secs. 7.1 and 7.2. We defer the treatment of partial polarization and Stokes parameters to Sec. I. Also, the treatment of phase velocity is best worked into a discussion of short-pulse optics, which deals with the coherent superposition of plane waves of different frequencies and is covered in the next section.

G. Superposition 2: Short-pulse electrodynamics and dispersion.

Short electromagnetic pulses are vitally important in a wide range of areas. Radar obviously depends entirely on the generation and reception of short pulses. In telecommunications, information is sent down the optical-fiber backbone of the Internet by short pulses. How these pulses propagate is critical in determining the distance between repeaters and the maximum rate by which information can be transferred. Short electromagnetic pulses are also used to achieve the extremely high peak powers that are needed in nonlinear optics and (hopefully) nuclear fusion. The study of attosecond phenomena demands exceedingly short pulses. At the other extreme, high-resolution measurements require the duration of the wave train to be as great as possible, making frequency spreads as narrow as possible. Thus it is worth examining the mathematics and physics of short pulses, and how they propagate. Generation is covered in Ch. 15.



From our understanding of delta functions we know that short pulses are formed by the coherent superposition of plane waves covering as broad a spectral range as possible under the condition that all waves are in phase at a particular location, say $z = 0$, at a particular time, say $t = 0$. We further expect that these pulses will be built around a central frequency ω_o and be described as an amplitude-modulated carrier wave with envelope $E_{env}(z)$, as shown in the figure.

Because the plane waves are solutions of the wave equation, the most general superposition is

$$\vec{E}(\vec{r},t) = \int_{-\infty}^{\infty} d^3k\, \vec{E}(\vec{k}) e^{i\vec{k}\cdot\vec{r} - i\omega t}\,, \tag{12.74}$$

where all components satisfy the dispersion relation.

$$\frac{c^2\vec{k}^2}{\omega^2} = \varepsilon(\omega)\,. \tag{12.75}$$

The maximum-coherency requirement can be satisfied by making all $\vec{E}(\vec{k})$ real at $z = 0$ and $t = 0$, in which case the in-phase requirement is satisfied automatically.

Although we are currently interested in one-dimensional phenomena, the general expansion must be written in terms of the three-dimensional vector $\vec{k}$ instead of the one-dimensional frequency ω, because the existence of paraxial rays demonstrates that waves

can exhibit lateral localization. This requires the packet to contain wave vectors with components orthogonal to the propagation direction. However, in this chapter we make the simplifying assumption that all $\vec{k} = k\hat{z}$ point in the same direction. We also assume that all ε are real and scalar (no absorption or crystal optics.) We also drop the vector notation, writing $\vec{E}(\vec{r},t) \to E(z,t)$. With these assumptions Eq. (12.74) becomes

$$E(z,t) = \int_{-\infty}^{\infty} dk\, E(k) e^{ikz - i\omega t}\,. \tag{12.76}$$

$E(z,t)$ is clearly a function of time and space. Its intrinsic behavior can be viewed in two ways. An observer located at $z = 0$ can watch the wave $E(0,t)$ pass by. Alternatively, time can be frozen at, for example, $t = 0$ and the spatial dependence $E(z,0)$ can be examined. In practice the packet is often specified by this spatial dependence, in which case $E(k)$ follows from the inverse transformation

$$E(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dz E(z,0) e^{-ikz}\,. \tag{12.77}$$

At some point the factor $1/2\pi$ must be accommodated between Eqs. (12.76) and (12.77). We follow the usual convention, describing $E(z,t)$, the quantity of primary interest, as a straight Fourier transform and assign the $1/2\pi$ to Eq. (12.77).

If $E(k)$ were viewed as $E(\omega)$, the inverse transformation would be

$$E(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dt E(0,t) e^{i\omega t}\,. \tag{12.78}$$

Since k and ω are connected by $ck = \omega n$, it is possible in principle to convert Eq. (12.76) to an integral over ω. However, since *n* is usually a function of ω, the differentials relate according to

$$cdk = d(\omega n(\omega)) \tag{12.79a}$$

$$= \left(n + \omega \frac{dn(\omega)}{d\omega} \right) d\omega\,. \tag{12.79b}$$

This is a simple direct proportion only if *n* is a constant. Hence for any given *z*, $E(\omega)$ should be calculated from $E(z,t)$ at that *z*, noting that in general that $E(\omega) = E(z,\omega)$ is a function of both ω and *z*.

We now consider short pulses in detail. Let $\omega = \omega(k)$, and let the superposition center about a wavenumber k_o to be defined below. Then Eq. (12.76) can be written

$$E(z,t) = \int_{-\infty}^{\infty} dk'\, E(k_o + k') e^{i(k_o + k')z} e^{i\omega(k_o + k')t}\,, \tag{12.80}$$

where $k = k_o + k'$, and $|k'| << k_o$. To proceed further, expand $\omega(k)$ as

$$\omega(k) = \omega(k_o) + k' \frac{d\omega}{dk} + \frac{1}{2} k'^2 \frac{d^2\omega}{dk^2} + \ldots, \tag{12.81}$$

where $k' = k - k_o$ is small compared to k_o. In practice the expansion can stop at second order, as shown below. Substituting Eq. (12.81) into (12.80) yields

$$\vec{E}(z,t) = \int_{-\infty}^{\infty} dk\, \vec{E}(k_o + k')\, e^{i(k_o+k')z - i\omega_o t - ik'(d\omega/dk)t - ik'^2(d^2\omega/dk^2)t/2 - \ldots}, \tag{12.82}$$

where $\omega_o = \omega(k_o)$. The phase factors involving k_o and ω_o are independent of k' and can be taken out of the integrand. Next, group the phase factors that are proportional to k' into a single term. Third, noting that k_o in $\vec{E}(k_o + k')$ is simply a placeholder, define a new coefficient $E_{env}(k')$ according to $\vec{E}(k_o + k') = \vec{E}_{env}(k')$. Ignoring for the moment the second-derivative term in the phase factor, the intermediate result is

$$\vec{E}(z,t) = e^{ik_o z - i\omega_o t} \int_{-\infty}^{\infty} dk'\, \vec{E}_{env}(k')\, e^{ik'(z - (d\omega/dk)t) - ik'^2(d^2\omega/dk^2)t/2 - \ldots}. \tag{12.83}$$

Next, noting that (ω_o / k_o) and $(d\omega/dk)$ both have the dimensions of velocity, define the *phase velocity* v_p and the *group velocity* v_g as

$$v_p = \omega_o / k_o; \tag{12.84a}$$

$$v_g = d\omega/dk; \tag{12.84b}$$

where the derivative is evaluated at $k = k_o$. Making these substitutions and ignoring for the moment the second derivative, Eq. (12.83) is now reduced to

$$E(z,t) = e^{ik_o(z - v_p t)} \int_{-\infty}^{\infty} dk'\, E_{env}(k')\, e^{ik'(z - v_g t)}. \tag{12.85}$$

We can now interpret Eq. (12.85) with respect to the figure above. The prefactor is the carrier wave moving with the phase velocity v_p. The integral has the classic form of a spatial Fourier transform, but one moving with a group velocity v_g. In spatial form this is

$$\vec{E}(z,t) = e^{ik_o z - i\omega_o t} \vec{E}_e(z - v_g t). \tag{12.86}$$

Hence $\vec{E}_e$ is the envelope, and $\vec{E}_{env}(k')$ is its Fourier transform (our nomenclature is correct). Given that the envelope is usually specified at $t = 0$,

$$E_{env}(k') = \frac{1}{2\pi} \int_{-\infty}^{\infty} dz E_{env}(z) e^{-ik'z}. \tag{12.87}$$

The group velocity is obtained by differentiating Eq. (12.79a) explicitly with respect to *k* and using the chain rule. We find

$$c = n(\omega)\frac{d\omega}{dk} + \omega\frac{dn}{d\omega}\frac{d\omega}{dk} \tag{12.88}$$

or

$$v_g = \frac{d\omega}{dk} = \frac{c}{n + \omega(dn/d\omega)}. \tag{12.89}$$

If ε is independent of ω, then the denominator reduces to *n*, and $v_g = v_p$. However, if ε varies rapidly with ω, as for example near the absorption edge of a material or one exhibiting electromagnetically induced transparency, then the derivative term can be equivalent to, or even orders of magnitude larger than, *n*. This is the origin of "slow light" as discussed in Sec. K.

The group velocity is also the scaling factor relevant for casting Eq. (12.85) into the frequency domain. Write $\omega' = k' v_g$ in which case Eq. (12.85) becomes

$$\vec{E}(z,t) = e^{ik_o z - i\omega_o t}\frac{1}{v_g}\int_{-\infty}^{\infty} d\omega' \vec{E}_e(\omega'/v_g)\, e^{i\omega'(z/v_g - t)}. \tag{12.90}$$

For the description of the envelope, the scaling factor between k' and ω' is v_g, not the usual c/n. An observer at $z = 0$ would therefore observe the time dependence

$$\vec{E}(0,t) = e^{-i\omega_o t}\frac{1}{v_g}\int_{-\infty}^{\infty} d\omega' \vec{E}_e(\omega'/v_g)\, e^{-i\omega' t}. \tag{12.91a}$$

$$= e^{-i\omega_o t}\vec{E}_e(-v_g t). \tag{12.91b}$$

As an example suppose the envelope has the Lorentzian shape

$$\vec{E}_e(z,0) = \frac{\vec{E}_o\,\Delta z^2}{z^2 + \Delta z^2}. \tag{12.92}$$

Then by Eq. (12.82)

$$\vec{E}_{env}(k') = \frac{1}{2}\vec{E}_o \Delta z\, e^{-|k'|\Delta z}. \tag{12.93}$$

This expression shows that the extent of the envelope Fourier coefficient is not infinite, but limited by a cutoff k'_{ce} that we shall define by $k'_{ce}\,\Delta z = 1$, or $k'_{ce} = 1/\Delta z$. Typical of Fourier analysis, the narrower the extent in one space the broader the extent in the conjugate space. Again ignoring the second derivative, Eq. (12.80) shows that the wave packet can also be written as

$$\vec{E}(z,t) = e^{ik_o(z - v_p t)}\frac{\vec{E}_o \Delta z^2}{(z - v_g t)^2 + \Delta z^2}. \tag{12.94}$$

We now consider the second-derivative term in the phase. The effect of this term requires a consideration of cutoffs, one type of which (k'_{ce}) has just been introduced by example. Being proportional to k'^2 this term in principle is highly oscillatory for large k', but since it is also linear in *t* we can safely ignore it at small *t*, which is what we have done.

The consequences of its oscillatory nature can be made more specific by performing a contour integration in the complex plane to turn the oscillatory function into a Gaussian ("method of steepest descents"). Specifically, replace the actual term according to

$$e^{-ik'^2(d^2\omega/dk^2)t/2} \to e^{-k'^2\omega''t/2}. \tag{12.95}$$

where $\omega'' = d^2\omega/dk^2$. This shows that k' cannot increase without measure, but is subjected to a second cutoff k'_{c2} that is given by setting the exponent equal to 1:

$$k'_{c2} = \sqrt{\frac{2}{\omega''t}}. \tag{12.96}$$

Because *t* appears in the denominator, at some point this phase term *must* become more important than k'_{ce}. At this point the k' integration range decreases, and the pulse broadens. The cutoff time t_c at which this occurs is obtained by replacing k'_{c2} with k'_{ce}. Then

$$t_c = \frac{2}{\omega''(k'_{ce})^2}. \tag{12.97}$$

Because the distance the pulse travels is $z = v_g t_c$, we see that both time and distance decrease as the square of the width of the pulse.

Thus the shape and extent of the packets become critical design parameters in information transfer. Shorter packets allow more packets to be sent per unit time, but shorter packets also broaden more rapidly, thereby more quickly reaching the point where separate packets can no longer be distinguished. This means shorter distances between regenerators, thereby raising costs, although dispersion compensators, short pieces of fiber of opposite compensation, can be used to regenerate the envelopes. Also, the broader frequency ranges needed to generate packets pulses means that carrier frequencies of multiplexed channels must be further apart, reducing the number of channels.

A second critical design parameter is $d^2\omega/dk^2$, and much effort has been invested in fiber-optic technology to reduce this as far as possible. The present state-of-the-art (2016) is perhaps best exemplified by the optical fiber manufactured by NICT, with c-band ($1.53\ \mu m \le \lambda \le 1.565\ \mu m$) specifications of 171 Gbits/s, 432 channels, and a length limit of 240 km. This gives an overall bit transfer rate of $(171\times10^9 s^{-1})\times 432 = 73.9\,\text{Tbits}/s$, although NT&T specifications claim "only" $69.1\,\text{Tbits}/s$. Either way the technology is truly astounding: given that a full-length

movie consists of approximately 20 Gbits of data, this corresponds to transmitting 3,500 full-length movies over 240 km *per second*.

To get a better idea of effectiveness of the engineering involved, consider using pure silica as the core. The refractive index of SiO_2 is 1.500 at 1.545 μm. The pulse transmission speed in the fiber is therefore approximately $v_g = 2\times10^8$ m/s. The time Δt required to propagate 240 km at $v_g = 2\times10^8$ m/s is $\Delta t = 1.20$ ms. At the stated transmission rate the separation between pulses in the fiber is about $4\Delta L = (2\times10^8 \text{m/s})/(1.71\times10^{11}\, s^{-1}) = 1.17$ mm. This allows a few hundred cycles per pulse, which is enough to define it accurately. If we assume that each packet is basically Lorentzian and the above limits apply, then the dispersion of the highly engineered version of SiO_2 at this wavelength is

$$\frac{d^2\omega}{dk^2} \approx \frac{2}{\Delta t\,\Delta k^2} = \frac{2}{\Delta t}\left(\frac{\Delta L}{\ln 2}\right)^2 = 0.0012\,(\text{m}^2/\text{s}). \tag{12.98}$$

As a reference, the dielectric function of pure SiO_2 is given approximately by

$$\varepsilon(\omega) \cong 2.016 + \frac{E_p^2}{E_g^2 - E^2}, \tag{12.99}$$

where E_p = 5.26 eV and E_g = 10.93 eV. At the operating wavelength of $\hbar\omega$ = 0.800 eV, dropping the term $\omega(dn/d\omega)$ relative to $n(\omega)$ results in an error of only 0.06%, so we can safely neglect it. Then

$$\frac{d^2\omega}{dk^2} \approx 0.34\,(\text{m}^2/\text{s}). \tag{12.100}$$

Thus the dispersion in the NT&T fiber is more than 300 times better than that of SiO_2. This improvement is accomplished by inserting impurities that do not affect the transmission properties but influence the dispersion. Not having these improvements, our fiber allows propagation for only about 0.9 km before the transmitted packets must be regenerated, so our SiO_2 fiber is clearly noncompetitive. It is interesting to see how close communications technology is coming to pushing fundamental limits.

Why Fourier analysis is essential in describing short pulses can be appreciated by considering the wave equation where μ and ε are constants, independent of frequency. For simplicity we consider them to be scalars as well. Assuming that the wave is propagating in the *z* direction, the field $\vec{E}(z,t)$ must satisfy

$$\left(\frac{\partial^2}{\partial z^2} - \frac{\mu\varepsilon}{c^2}\frac{\partial^2}{\partial t^2}\right)\vec{E}(z,t) = 0. \tag{12.101}$$

Consider the trial solution

$$\vec{E}(z,t) = \vec{E}(z - vt) = \vec{E}(u), \tag{12.102}$$

where $u = z - vt$. Then invoking the chain-rule of differentiation

$$\frac{\partial}{\partial z} = \frac{\partial}{\partial u}\frac{\partial u}{\partial z} = \frac{\partial}{\partial u}; \quad \frac{\partial}{\partial t} = \frac{\partial}{\partial u}\frac{\partial u}{\partial t} = -v\frac{\partial}{\partial u}; \tag{12.103a,b}$$

and completing the algebra we find

$$\left(\frac{\partial^2}{\partial z^2} - \frac{\mu\varepsilon}{c^2}\frac{\partial^2}{\partial t^2}\right)\vec{E}(z,t) = \left(\frac{\partial^2}{\partial u^2} - \frac{\mu\varepsilon}{c^2}v^2\frac{\partial^2}{\partial u^2}\right)\vec{E}(u) = 0. \tag{12.104}$$

This is an identity if

$$v = c/\sqrt{\mu\varepsilon}. \tag{12.105}$$

Hence we conclude that if ε and μ are both constants, any wave form will propagate unchanged through the material.

So far, we have said nothing about k_o. We can guess that it is somewhere near the midpoint of the distribution of the wave vectors forming the packet. More formally, recalling $E_{env}(k') = E(k_o + k') = E(k)$, let k_o be defined as that value that minimizes the mean-square deviation of $E(k)$:

$$\frac{d}{dk_o}\int_{-\infty}^{\infty} dk(k-k_o)^2 E_k(k) = -2\int_{-\infty}^{\infty} dk(k-k_o)E_k(k) = 0. \tag{12.106}$$

Therefore

$$k_o = \frac{\int_{-\infty}^{\infty} kdkE_k(k)}{\int_{-\infty}^{\infty} dkE_k(k)}. \tag{12.107}$$

As another example, let $E_k = E_{ko}$ be constant for $k_1 \le k \le k_2$ and zero elsewhere (square Fourier envelope). Then $k_o = (k_1 + k_2)/2$, $\Delta k_o = (k_2 - k_1)/2$, and

$$E(z,t) = e^{ik_o z - i\omega_o t}\left(\int_{-\Delta k_o}^{+\Delta k_o} d\Delta k E_{ko} e^{i\Delta k(z - v_g t)}\right) \tag{12.108a}$$

$$= E_{ko}e^{ik_o z - i\omega_o t}\frac{e^{i\Delta k_o(z - v_g t)} - e^{-i\Delta k_o(z - v_g t)}}{i(z - v_g t)} \tag{12.108b}$$

$$= 2E_{ko}\Delta k_o e^{ik_o z - i\omega_o t}\,\mathrm{sinc}\left(\Delta k_o (z - v_g t)\right). \tag{12.108c}$$

Here, the full-width-half-maximum of $\Delta z = 3.79/\Delta k_c$ follows from $\mathrm{sinc}(1.895) = 0.5$. The solution exhibits the Gibbs phenomenon, the appearance of oscillations when the function being transformed has hard cutoffs. The oscillations attenuate slowly as

$\sim 1/(z - v_g t)$ for large $(z - v_g t)$. Given that $E_k(k) = E_{ko} u(\Delta k - |k|)$, the frequency distribution is

$$E_\omega(\omega) = \frac{E_{ko}}{v_g} u\left((\omega - \omega_o)/v_g\right). \tag{12.109}$$

As another example where $E_k(k)$ is given, let

$$E_k(k) = E_{ko} \cos\left(\frac{\pi k}{2\Delta k_o}\right), \ \frac{|k'|}{\Delta k_o} \le 1; \tag{12.110a}$$

$$= 0, \qquad \text{elsewhere.} \tag{12.110b}$$

This expression is artificial, but it is also symmetric about $k = 0$ and has hard cutoffs at $k = \pm\Delta k_o$. However, the *k*-space distribution is narrower than that of the square distribution, so we expect the envelope function to be wider in direct space. Doing the math we find

$$E(z,t) = e^{ik_o z - i\omega_o t} \int_{-\Delta k_o}^{\Delta k_o} dk E_{ko} \cos\left(\frac{\pi k}{2\Delta k_o}\right) e^{ik(z - v_g t)} = E_{ko} \Delta k_o \left(\frac{\pi \cos\theta}{(\pi/2)^2 - \theta^2}\right), \tag{12.111}$$

where $\theta = \Delta k_o (z - v_g t)$. This looks impossibly divergent, but the denominator factors into simple poles at $\pm\pi/2$, where the numerator also vanishes linearly in θ. Thus the result is regular everywhere. The term in braces has the value 0.5 at $\theta = \pm 2.93$, so the FWHM is $\Delta z = 5.86/\Delta k$. This is substantially broader than the square distribution, which is not surprising because its reciprocal-space width is effectively narrower. The oscillations in the transform now die off as $1/z^2$, because the discontinuities in the source function occur in its derivative, not in its value. In this case the frequency distribution is

$$E(\omega) = E_{ko} \cos\left(\pi(\omega - \omega_o)/2\Delta k v_g\right). \tag{12.112}$$

H. Superposition 3: Incoherent superposition and detector averaging.

Short-pulse electrodynamics is a consequence of superposing fields phase-coherently. Although Nature works only with fields, our experience shows us that in our uncorrelated world the relevant superposed quantity is intensity. For example, even though any given wave that emerges from a light bulb is in a definite state of polarization, if the phases, amplitudes, frequencies, and/or polarization states of different waves are uncorrelated, superpositions of fields must reduce in some way to superpositions of intensities. The objective of this section is to identify these conditions. These will be found to be dependent on statistical averaging and in some cases on the finite response times of detectors.

We start with the general superposition

$$\vec{E}(\vec{r},t) = \vec{E}(z,t) = \sum_i \left(\hat{x} E_{ix} e^{i\varphi_{ix}} + \hat{y} E_{iy} e^{i\varphi_y}\right) e^{ik_i z - i\omega_i t}, \tag{12.113}$$

where E_{ix} and E_{iy} are magnitudes and φ_{ix} and φ_{iy} the respective corresponding phases. Propagation in air and in the *z* direction is assumed. Also assumed is that the individual waves are completely uncorrelated, i.e., have no definite phase relationship, and hence add incoherently. The output of a laser is an exception, but we do not consider the laser case here. Note that $E_{ix}e^{i\varphi_{ix}}$ and $E_{iy}e^{i\varphi_{iy}}$ of a given wave *i* are also exceptions, because they share a common propagation factor $e^{ik_i z - i\omega_i t}$ and hence are automatically correlated. The use of a discrete summation rather than a continuum integration is not due to our desire as physicists to count photons, but as mathematicians to avoid a later need to introduce densities of states.

We next consider the detector. The operation of a linear detector can be written

$$< \vec{E} >_{out} = \int_{-\infty}^{\infty} dt' G(t-t')\vec{E}(t'), \tag{12.114}$$

where the Green function $G(t-t') = 0$ for $t < t'$. The detector acts as a low-pass filter, delivering an output voltage or current proportional to $< \vec{E} >_{out}$. For our purposes its key property is its finite response time τ. An illustrative possibility is a decreasing-exponential response

$$G(t-t') = \frac{1}{\tau} e^{(t-t')/\tau} u(t-t'). \tag{12.115}$$

Assuming an incoming wave with a time dependence $E(0,t) = E_o \cos(\omega t)$, the output is

$$< \vec{E}(t) > = \frac{E_o}{1+(\omega\tau)^2}. \tag{12.116}$$

Considering that the constituents of visible-light beams have angular frequencies $\omega \sim 10^{15} s^{-1}$, it is clear that a detector with the ~5 ms response time of an eye is hopelessly inadequate for field detection. Some form of rectification is necessary.

Accordingly, we investigate intensity. To avoid notational complexity we consider without loss of either math or physics only the *x* component, equivalent to assuming that the beam has been passed through a linear polarizer. From the Faraday-Maxwell Equation

$$ik_{ix}\hat{z} \times \hat{x} E_{ix} e^{ik_{ix}x - i\omega_{xi}t + i\varphi_{ix}} - \frac{i\omega_i}{c}\vec{H}_{ix} = 0, \tag{12.117}$$

whence

$$\vec{H}_{ix} = \hat{y} E_{ix} e^{ik_i z - i\omega_i t + \varphi_{ix}}. \tag{12.118}$$

Therefore

$$\vec{S}_x = \frac{c\hat{z}}{4\pi}\left(\sum_i E_{ix}\cos(k_{ix}x - \omega_{ix}t + \varphi_{ix})\right)\left(\sum_j E_{jx}\cos(k_{jx}x - \omega_{jx}t + \varphi_{jx})\right). \tag{12.119}$$

Using the trigonometric cosine-product identity, Eq. (12.119) can be written

$$\vec{S}_x = \frac{c\hat{z}}{4\pi}\sum_{ij} E_{ix}E_{jx}\frac{1}{2}\Big(\cos\big((k_{ix}-k_{jx})x-(\omega_{ix}-\omega_{jx})t+(\varphi_{ix}-\varphi_{jx})\big) + \cos\big((k_{ix}+k_{jx})x-(\omega_{ix}+\omega_{jx})t+(\varphi_{ix}+\varphi_{jx})\big)\Big). \qquad (12.120)$$

Given that the waves are uncorrelated, $(\varphi_{ix}-\varphi_{jx})$ and $(\varphi_{ix}+\varphi_{jx})$ are effectively random variables except for $i = j$. Assuming the beam contains enough waves, Eq. (12.120) therefore reduces to

$$\vec{S}_x = \frac{c\hat{z}}{4\pi}\sum_i E_{ix}^2\frac{1}{2} = \frac{c\hat{z}}{8\pi}\sum_i E_{ix}^2, \qquad (12.121a,b)$$

$$= \hat{z}\sum_i I_{ix} = \hat{z}I_x. \qquad (12.121c)$$

Thus in the uncorrelated limit the total intensity I_x of the beam is just the sum of the intensities I_{ix} of the individual constituents of the beam. Note also that rectification is perfect: no time averaging is necessary. Further, in a large enough ensemble the random nature also eliminates "speckle" patterns that are characteristic of coherent sources.

Repeating the same derivation for polarization along *y* leads to the analogous result

$$\vec{S}_y = \hat{z}\sum_i I_{iy} = \hat{z}I_y. \qquad (12.122)$$

Working with intensities rather than fields is a huge advantage in many applications. The processing power that would be required to extract relevant information from incident light at the field level in our everyday view of the world would be enormous. On the other hand, cell-phone technology among others relies on interference among fields to reduce power consumption and improve range by beam steering, as discussed in Ch. 14. This is the reason for multiple antennas on cell-phone towers.

I. Superposition 4: Stokes Parameters.

Stokes (of Stokes' Theorem fame) found that the polarization state of any optical beam, coherent or not, can be described quantitatively by 4 parameters calculated from 6 intensities measured by passing the beam through a polarizer and a 90° retarder. Two of the intensities, $I_0 = I_x$ and $I_{90} = I_y$, are described above. The remaining four are developed below.

To do this, recognize first that we have not completed the task outlined in the previous section. We must determine the intensity when the polarization is oriented at a general azimuth angle *P* relative to the laboratory system. In this case the polarizer acts as a nonlinear element mixing the *x*- and *y*- polarized components of the beam. We have already seen that all cross terms in $\vec{S}$ vanish, so the result is obtained by considering plane waves one at a time. We find that this projection provides information on the

relative phase $(\varphi_{ix} - \varphi_{iy})$. This difference is part of the definition of the polarization state of the wave, hence establishes additional critical information about the beam.

With polarization along *P*, the unit vector in the field direction is not $\hat{x}$ or $\hat{y}$ but rather $\hat{u} = \hat{x}\cos P + \hat{y}\sin P$. The polarizer projects the *x* and *y* components of the beam along $\hat{u}$, so Eq. (12.108) becomes

$$\vec{E}(z,t) = \sum_i \hat{u}\left(E_{ix}e^{i\varphi_{ix}}\cos P + E_{iy}e^{i\varphi_y}\sin P\right)e^{ik_i z - i\omega_i t}\,, \tag{12.123}$$

Working through the math Eq. (12.119) becomes

$$\vec{S}_P = \frac{c\hat{z}}{4\pi}\left\{\left(\sum_i\left(E_{ix}\cos P\cos(k_i z - \omega_i t + \varphi_{ix}) + E_{iy}\sin P\cos(k_i z - \omega_i t + \varphi_{iy})\right)\right)\right.$$

$$\left.\times\left(\sum_j\left(E_{jx}\cos P\cos(k_j z - \omega_j t + \varphi_{jx}) + E_{jy}\sin P\cos(k_j z - \omega_j t + \varphi_{jy})\right)\right)\right\}. \tag{12.124}$$

In employing the trigonometric identity new cross terms arise. After eliminating all terms $i \neq j$ the result is

$$I_P = \frac{c}{8\pi}\sum_i\left(E_{ix}^2\cos^2 P + E_{iy}^2\sin^2 P + 2E_{ix}E_{iy}\sin P\cos P\cos(\varphi_{ix} - \varphi_{iy})\right)$$

$$= I_x\cos^2 P + I_y\sin^2 P + \left(\sum_i\sqrt{I_{ix}I_{iy}}\cos(\varphi_{ix} - \varphi_{iy})\right)\sin 2P \tag{12.125}$$

where $I_p = |S_P|$. Note that Eq. (12.125) is completely general, containing Eqs. (12.121c) and (12.122) when $P = 0$ and $P = \pi/2$, respectively. While the sums over I_{ix} and I_{iy} can be expressed in closed form, as indicated, in general the cross term cannot.

The new term is the interesting one. It is maximized by setting $P = \pi/4$, which is not surprising since that optimizes the mixing. Relevant to the next paragraph, if the y component is passed through a retarder that adds a relative phase $\pi/2$, the cosine term is converted to a sine term according to

$$\cos(\varphi_{ix} - \varphi_{iy} - \pi/2) = \sin(\varphi_x - \varphi_y)\,. \tag{12.126}$$

We are now ready to define the Stokes Parameters. Define the intensities I_0, I_{90}, I_{45}, and I_{-45} obtained from Eq. (12.125) with $P = 0^\circ$, 90°, 45°, and -45°, respectively, and $I_{q,45}$ and $I_{q,-45}$ obtained from Eq. (12.125) with the 90°-retarder in place and $P = 45^\circ$ and -45°, respectively. The list is

$$I_0 = I_x\,; \tag{12.127a}$$

$$I_{90} = I_y\,; \tag{12.127b}$$

$$I_{45} = \frac{1}{2}(I_x + I_y) + \sum_i \sqrt{I_{ix} I_{iy}} \cos(\varphi_{ix} - \varphi_{iy}) ; \tag{12.127c}$$

$$I_{-45} = \frac{1}{2}(I_x + I_y) - \sum_i \sqrt{I_{ix} I_{iy}} \cos(\varphi_{ix} - \varphi_{iy}) ; \tag{12.127d}$$

$$I_{q,45} = \frac{1}{2}(I_x + I_y) + \sum_i \sqrt{I_{ix} I_{iy}} \sin(\varphi_{ix} - \varphi_{iy}) ; \tag{12.127e}$$

$$I_{q,-45} = \frac{1}{2}(I_x + I_y) - \sum_i \sqrt{I_{ix} I_{iy}} \sin(\varphi_{ix} - \varphi_{iy}) ; \tag{12.127f}$$

Given Eqs. (12.120), the Stokes parameters are the obvious linear combinations:

$$S_0 = I_0 + I_{90} = I_{45} + I_{-45} = I_{q,45} + I_{q,-45} ; \tag{12.128a}$$

$$S_1 = I_0 - I_{90} ; \tag{12.128b}$$

$$S_2 = I_{45} - I_{-45} ; \tag{12.128c}$$

$$S_3 = I_{q,45} - I_{q,-45} ; \tag{12.128d}$$

The significance of the Stokes parameters is as follows. First, S_0 is the total intensity, which can be obtained three different ways. Second, S_1 and S_2 describe linearly polarized light, and S_3 circularly polarized light. The identification of S_1 and S_2 with linear-polarization follows because in this case $\sin(\varphi_x - \varphi_y) = 0$, hence $S_3 = 0$. Third, for circularly polarized light $I_0 = I_{90}$ and $\cos(\varphi_x - \varphi_y) = 0$, hence $S_1 = S_2 = 0$. Fourth, for unpolarized light $S_1 = S_2 = S_3 = 0$. This follows because in this case $I_x = I_y$, which eliminates S_1, and summing the (random) phase differences eliminates S_2 and S_3. Fifth, for totally polarized light

$$S_0^2 = S_1^2 + S_2^2 + S_3^2 . \tag{12.129}$$

This follows because in this case the sums reduce to single terms $\sqrt{I_x I_y} \cos(\varphi_x - \varphi_y)$ for Eqs. (12.127c) and (12.127d), and $\sqrt{I_x I_y} \sin(\varphi_x - \varphi_y)$ for Eqs. (12.127e) and (12.127f). The remaining math is elementary.

The limits of zero and S_0^2 on $S_1^2 + S_2^2 + S_3^2$ leads to the definition of the degree of polarization *D*:

$$D = \sqrt{S_1^2 + S_2^2 + S_3^2} \Big/ S_0 , \tag{12.130}$$

where $0 \le D \le 1$, and the fractions f_l and f_c of linear and circular polarized light,

$$f_l = (S_1^2 + S_2^2) \Big/ S_0^2 , \tag{12.131a}$$

$$f_c = S_3^2 / S_0^2 \,. \tag{12.131b}$$

The determination of the expression for the azimuth P of the major axis of the linearly polarized component is more difficult, but also begins with Eq. (12.125). Assuming that the beam is linearly polarized, Eq. (12.125) reduces to

$$I(P) = \frac{1}{2}(I_x + I_y) + \frac{1}{2}(I_x - I_y)\cos 2P + \sqrt{I_x I_y}\,\sin 2P \tag{12.132a}$$

$$= \frac{1}{2}\left(S_0 + S_1 \cos 2P + S_2 \sin 2P\right). \tag{12.132b}$$

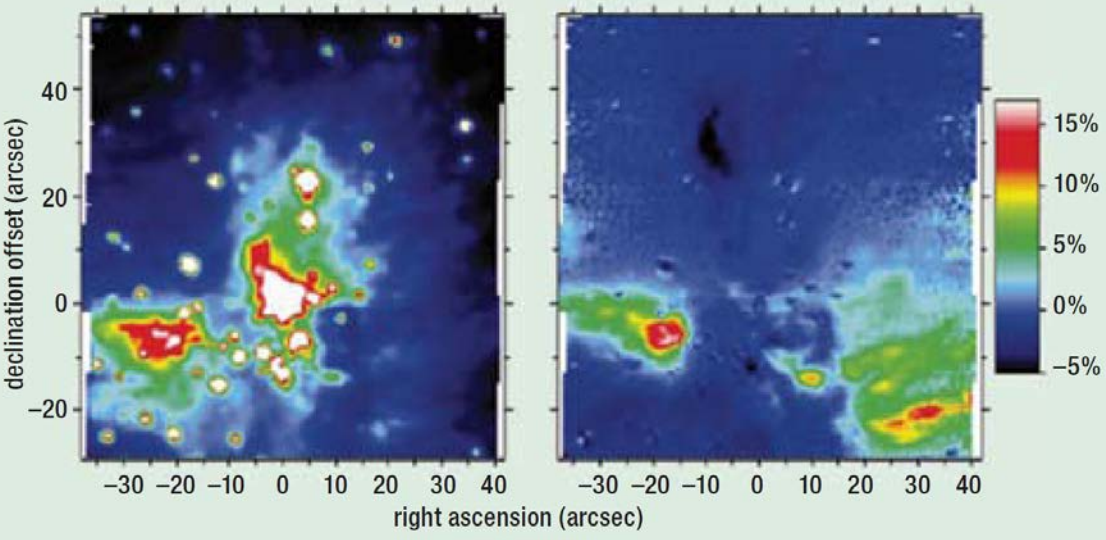
Defining $\cos\theta = S_1 / \sqrt{S_1^2 + S_2^2}$ and $\sin\theta = S_2 / \sqrt{S_1^2 + S_2^2}$, it follows that

$$\theta = \tan^{-1}(S_2 / S_1) = 2P \tag{12.133}$$

is the angle needed to rotate the major axis back to the x axis in the laboratory frame. At the same time the derivation reinforces the conclusion that $S_1^2 + S_2^2$ is the fraction of linearly polarized light.

The advantage of the Stokes parameters is that they provide a complete description of a light beam even if the beam is partially polarized. This allows partially polarized light to be analyzed, a capability that is becoming increasingly important as more and more complicated materials and structures, particularly organics, are of interest. "Partial polarization" is a general term that describes the result of depolarization scattering as well as the incoherent superposition of two plane waves in different states of polarization.

The following is an example from astronomy. The figures below are images obtained at $\lambda = 2.2$ μm of the OMC-1 star-forming region in Orion (Bailey et al., Science 280, 672 (1998)) (from J. Hough, Astronomy and Geophysics 47, 3.31-3.35 (2006)). The image on



7: Circular polarization image of the OMC-1 star-forming region in Orion at 2.2μm. Total intensity is shown on the left with the bright Becklin-Neugebauer (BN) object at co-ordinate (0,0). Percentage circular polarization is shown on the right, from –5% (black) to +17% (white). (Bailey *et al.* 1998)

the left shows the intensity, while that on the right shows S_3, the circular polarization. The bright Becklin-Neugebauer (BM) object at coordinate (0,0) exhibits essentially no circular polarization, while the object to its left is substantially circularly polarized. Likewise, extensive circular polarization is emerging at the lower right from a region that is no strong source of radiation. As seen in an earlier homework assignment, circular polarization is a consequence of charged particles interacting with magnetic fields. Thus this information can be interpreted in terms of magnetic fields in the region.

The capability of analyzing partially polarized light underlies much of current optical metrology of complex materials and structures. Procedures work as follows. Form a 4-vector of the Stokes parameters, i.e., $\vec{S} = (S_1, S_2, S_3, S_4)$, measure the Stokes parameters of a beam, then let the beam interact with a system. The system turns the Stokes parameters of the incident beam into those of the exit beam. The process can be described in matrix form as

$$\begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}_{out} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{pmatrix} \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}_{in} . \qquad (12.134)$$

The 4×4 matrix M of transfer-function elements m_{ij} is termed the Mueller matrix. The m_{ij} are analyzed as a function of wavelength to obtain information about the system. The major advantage of Mueller-matrix spectroscopy is that it can deal not only with specular reflection and homogeneous transmission, but also scattering.

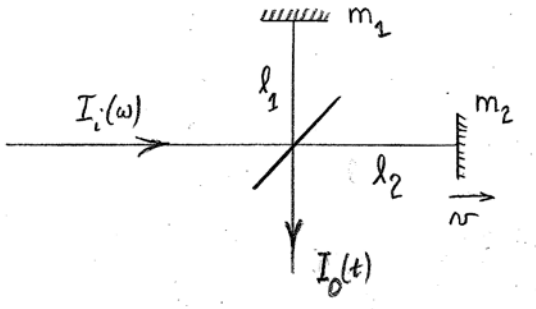
In practice, configurations typically consist of a polarizer to establish a specific polarization, a rotating retarder plate with 90° retardation generally near midrange, the sample under study, a second retarder plate rotating at a different speed, a polarizer, then finally a detector. The last polarizer may be a beamsplitting element that sends the individual polarization intensities to separate detectors. The detected intensities are Fourier analyzed to obtain coefficients that are converted to the elements m_{ij} of the Mueller matrix. The result is an optically complete picture of the sample.

We have assumed throughout that a beam is uniform over its cross section. However, situations arise where the polarization state of a beam may differ over this section. In that case spatial variation also needs to be considered. For example, if half the beam is linearly polarized in the vertical direction and the other half horizontal, a detector would sense no change of intensity as a polarizer inserted in the beam is rotated, whether or not a quarter-wave plate is inserted. In this case the three polarization-sensitive Stokes parameters of the beam are exactly zero.

J. Superposition 5: Interferometry.

Interferometry deals with a different type of superposition, one where a plane wave is split into two parts that travel along different paths through an optical system. These are then recombined so the plane wave interferes with itself (Pauli's famous comment: a

photon only interferes with itself). The classic configuration is the Michelson interferometer shown in the diagram. It consists of a beamsplitter that divided the beam into two parts, a fixed mirror m_1, a moving mirror m_2, and a detector to determine the intensity of the reconstructed beam. The beamsplitter divides the beam into approximately equal fractions, and the mirrors m_1 and m_2 are considered ideal.



Because Nature combines fields, the interferometer is sensitive to phase, which is essential information in many physical applications.

In this section we develop the basic equation, then use it to analyze several applications. One is the LIGO gravitational-wave detector, which uses a stabilized Nd laser source operating at 1064 nm. Its purpose is to detect gravity waves by measuring path-length differences to precisions of 10^{-18} m. A second application is the Fourier-transform infrared spectrometer (FTIR), where the output is analyzed for the frequency distribution $\sum_i (I(\omega_i))$ of the incident beam.

Our interest follows in part because this is a different type of superposition than those discussed so far. To analyze the interferometer, we track the central ray of a plane wave that starts with an amplitude E_o a distance l_0 away from the beamsplitter. Considering path 1 in the above diagram, the ray is reflected at a right angle with a complex reflectance $\tilde{r} = re^{i\varphi_r}$, travels a distance l_1 where it is retroreflected by mirror m_1 with a reflectance $\tilde{r}_m = r_m e^{i\varphi_m}$, travels a return distance l_1, is then transmitted through the beamsplitter with a transmission coefficient $\tilde{\tau} = \tau e^{i\varphi_\tau}$, and finally reaches the detector after traveling a distance l_2. The action of the path can be summarized as

$$\tilde{E}_{1d} = E_o e^{ikl_o} re^{i\varphi_r} e^{ikl_1} r_m e^{i\varphi_m} e^{ikl_1} \tau e^{i\varphi_\tau} e^{ikl_3} e^{-i\omega t}. \tag{12.135}$$

The action of the second path is identical except for the distance to the second mirror:

$$\tilde{E}_{2d} = E_o e^{ikl_o} \tau e^{i\varphi_\tau} e^{ikl_2} r_m e^{i\varphi_m} e^{ikl_2} re^{i\varphi_r} e^{ikl_3} e^{-i\omega t}. \tag{12.136}$$

Factoring out common terms, the field reaching the detector is

$$\tilde{E}_{1d} + \tilde{E}_{2d} = E_{tot} = E_o e^{ikl_o} re^{i\varphi_r} r_m e^{i\varphi_m} \tau e^{i\varphi_\tau} e^{ikl_3} \left(e^{2ikl_1} + e^{2ikl_2}\right) e^{-i\omega t}. \tag{12.137}$$

Lumping all the common phase factors together as φ, the real projection is

$$\mathrm{Re}(\tilde{E}) = |E_o r r_m \tau| \left(\cos(\varphi + 2kl_1 - \omega t) + \cos(\varphi + 2kl_2 - \omega t)\right) \tag{12.138}$$

Recognizing that this must be squared in calculating $\vec{S}$, we use a trigonometric identity to make the calculation more efficient. Write

$$\varphi + 2k_{l1} - \omega t = (\varphi + kl_1 + kl_2 - \omega t) + (kl_1 - kl_2)\,; \tag{12.139a}$$

$$\varphi + 2kl_2 - \omega t = (\varphi + kl_1 + kl_2 - \omega t) - (kl_1 - kl_2)\,; \tag{12.139b}$$

Then Eq. (12.130) can be transformed into

$$\mathrm{Re}(\tilde{E}) = 2\,|\,E_o r r_m \tau\,|\cos(\varphi + kl_1 + kl_2 - \omega t)\cos(kl_1 - l_2)\,. \tag{12.140}$$

In the time average of its square, the first cosine is replaced by $(1/2)$, leaving

$$< I >_{out} = \frac{c}{8\pi}\,|\,E_o r r_m \tau\,|^2\; 4\cos^2 k(l_1 - l_2) \tag{12.141a}$$

$$= \frac{c}{4\pi}\,|\,E_o r r_m \tau\,|^2\,\left(1 + \cos 2k(l_2 - l_1)\right). \tag{12.141b}$$

For an ideal beamsplitter $r = \tau = 1/\sqrt{2}$, and for an ideal mirror $r_m = -1$. With these substitutions

$$< I >_{out} = \frac{c}{16\pi}\,|\,E_o\,|^2\,\left(1 + \cos 2k(l_2 - l_1)\right) \tag{12.142a}$$

$$= \frac{1}{2} I_{in}\left(1 + \cos 2k(l_2 - l_1)\right). \tag{12.142b}$$

Equation (12.142b) defines the operation of the Michelson interferometer for an input wave of wavevector $k = \omega/c$. It is positive definite, as expected of an intensity, and consists of a cosine term that ranges from zero to 2 about a mean value of 1 and averages to zero.

Equation (12.142b) is easily generalized to the situation where the input consists of collection of incoherent plane waves of different wave vectors $\boldsymbol{k}_i$

$$< I_{tot} >_{out} = \frac{1}{2}\sum_i I_{in,i}\left(1 + \cos 2k_i(l_2 - l_1)\right). \tag{12.143}$$

Again, Eq. (12.143) is consistent with expectations: all terms are positive definite, as required for intensities. All waves, regardless of frequency, add coherently at $l_1 = l_2$. However, if the source has a localized spread of wavelengths about a central value, then interference becomes increasingly destructive as $|\,l_1 - l_2\,|$ increases, and the oscillatory term collapses to zero. The separation $|\,l_1 - l_2\,|$ where the oscillatory term reaches half of its peak value is termed the *coherence length*.

The functional dependence of the decrease with $|\,l_1 - l_2\,|$ provides information about the physics causing the distribution of wavelengths. If the source is a stabilized gas laser with well-defined atomic transitions, the coherence length may be of the order of km and the functional form of the decrease is gaussian. This is the signature of Doppler broadening due to the motion of the atoms or molecules in the gas. If the source is a solid-state laser, the coherence length is much less and the decrease is exponential, which is characteristic of lifetime broadening.

We consider lifetime broadening as an example. Suppose that the incoming radiation has a Lorentzian distribution of width Δk about a central value k_o, such that

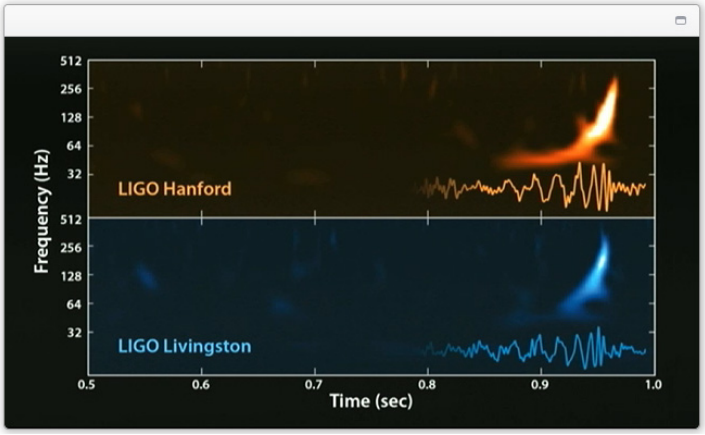
$$< dI_{in}(k)/dk > = \frac{I_o \Delta k}{(k-k_o)^2 + \Delta k^2}. \qquad (12.144)$$

Replacing the sum with an integral and performing the integration yields

$$< I_{out}(l_1 - l_2) > = I_o(1 + e^{-|l_1 - l_2|\Delta k} \cos(k_o z)). \qquad (12.145)$$

Gaussian broadening leads to a similar expression, and is left as a homework assignment.

The most famous applications of interferometry, the Michelson and LIGO experiments, use a single-wavelength source. The quantity of interest to Michelson and Morley was $k = \omega/c$, because if an "ether" were present, a daily or seasonal dependence of *c* might be expected as the earth rotated and orbited around the sun, both of which would change the velocity of the interferometer relative to the ether. Of course, no such variation was found. This null result helped lead Einstein to special relativity. As a result, the Michelson-Morley experiment goes down as one of the most important "failed" experiments in history.



In the LIGO interferometer, the quantity of interest is $(l_1 - l_2)$ itself. This can presently be measured to a precision of approximately 10^{-18} m, or about 1000 times less than the diameter of the proton. The LIGO configuration uses 1064 nm radiation from a highly stabilized Nd:YAG laser. After passing through the beamsplitter, each wave makes approximately 400 round trips down its 4 km arm before returning to the beamsplitter, for a total path length of about 3200 km. The path-length difference is adjusted to maximize sensitivity to it, a calculation that you do as a homework assignment. As everyone knows, the experiment is a success. The facility recently detected the merger of two black holes, as shown in the picture (from physicsworld.com). (As an aside, this is a phase-space plot similar to a musical score: time and frequency are plotted on the horizonal and vertical axes, respectively.)

Although this is never mentioned, LIGO also provides the most critical test to date of Einstein's postulate that *c* is independent of the speed of the source or of the observer. This test is many orders of magnitude more stringent than Michelson and Morley could achieve. As presently constituted, the LIGO interferometer can detect variations δc of *c* of the order of

$$\frac{\delta c}{c} \sim \frac{\delta L}{L} = \frac{10^{-18}\ \text{m}}{3.2\times 10^{9}\ \text{m}} \sim 3\times 10^{-28}, \tag{12.146}$$

or about $10^{-19}\ \text{m/s}$. The fact that this is never mentioned shows how thoroughly we have accepted Einstein's postulate. Nevertheless, it is even safer now to say that no evidence of an ether has ever been found.

A major application today is the Fourier transform infrared spectrometer, or FTIR. Here, the source is a spectrum typically consisting of a wide range of wavelengths, and the mirror m_2 is moved over a distance range Δl_2 at a velocity *v*. Although the output is usually expressed as a function of time, it is less confusing to do the analysis as a function of $z = (l_1 - l_2)$, writing

$$< I_{out}(z) > = \frac{1}{2}\sum_i < I_{in,i} > \big(1 + \cos(k_i z)\big). \tag{12.147}$$

The cosine transform of Eq. (12.137) with respect to *z* over the length $2L$ is

$$\begin{aligned}
&\frac{1}{2\pi}\int_{-L}^{L} dz\, \cos(kz)\left(\frac{1}{2}\sum_i < I_{in,i} > (1+\cos(k_i z))\right) \\
&\quad = \frac{1}{4\pi}\sum_i < I_{in,i} > \int_{-L}^{L} dz\, \cos(kz)(1+\cos(k_i z)) \\
&\quad = \frac{L}{4\pi}\sum_i < I_{in,i} > \big(\text{sinc}(kL) + \text{sinc}((k+k_i)L) + \text{sinc}((k-k_i)L)\big) \\
&\quad \cong \frac{L}{4\pi}\sum_i < I_{in,i} > \text{sinc}((k-k_i)L)\,.
\end{aligned} \tag{12.148}$$

Therefore, the Fourier cosine transform $\cos kz$ isolates a band of intensities $< I_{in,i} >$ centered about *k* with a resolution (FWHM width) of $|\, k_i - k\,| = 3.79/L$, where $2L$ is the total excursion of the mirror. As we might expect, resolution increases with the length of mirror travel. The transform is repeated with different values of $k = 2\pi/\lambda$ to cover the wavelength range λ of interest. Conversion of the results to energy follows from $ck/\omega = 1$.

As a historical sidelight, the principle of operation of the FTIR was known long before it came into widespread use. Practical implementations had to await the invention of the HeNe laser. Mechanical positioning was simply not accurate enough to create interferograms that could be decoded reliably. Sufficient accuracy could not be realized until the mirror position could be determined from a second interferogram obtained by reflecting a HeNe laser beam from the same mirror.

Regarding the Michelson interferometer, the reader has probably noticed that we have not accounted for all the input energy. The reason is that a second output beam exists. This beam returns to the source. By considering the second beam the characteristics of

the ideal beam splitter can be obtained. Following the same development that leads to Eq. (12.149), this beam also consists of a superposition of two beams, with one reflected twice and the other transmitted twice. Thus the field returning to the source is given by

$$E_{\text{det}} = -E_o e^{2ikl_0}\left(r^2 e^{2ikl_1} + \tau^2 e^{2ikl_2}\right)e^{-i\omega t} \tag{12.149a}$$

$$= -E_o e^{2ikl_0}\,\frac{1}{2}e^{i2\phi_r}\,(e^{2ikl_1} + e^{2ikl_2+2i(\phi_\tau-\phi_r)})e^{-i\omega t} \tag{12.149b}$$

where $r = |r|e^{i\phi_r} = e^{i\phi_r}/\sqrt{2}$ and $\tau = |\tau|e^{i\phi_\tau} = e^{i\phi_\tau}/\sqrt{2}$. The math shows that the "missing" energy is recovered if

$$\phi_r - \phi_\tau = \pm\pi/2 + 2\nu\pi \tag{12.150}$$

where ν is an integer. Carrying forward the calculation gives the result

$$< S_{source} > = \sum_i \frac{cE_{oi}^2}{16\pi}\left(1-\cos(2k_i(l_1-l_2))\right). \tag{12.151}$$

The sum of the two intensities of Eqs. (12.147) and (12.151) is equal to the input intensity, as expected. The $\pi/2$ phase shift between reflection and transmission is characteristic of the ideal beamsplitter, and will be discussed further in the next chapter.

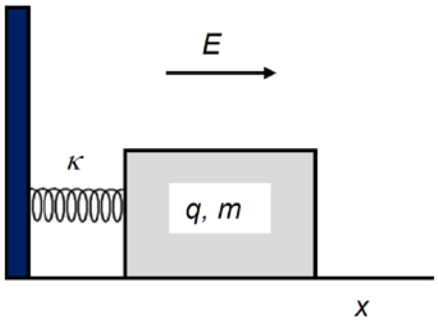
K. Superposition 6: Collective effects.

Condensed-matter and atomic, molecular, and optical (AMO) physicists are currently actively investigating the physics of collective effects, which are those associated with interacting oscillators. The Clausius-Mossotti effect, which we discussed last semester, is one. Superconductivity is another. A third is electromagnetically induced transparency (EIT), discussed below. Superposition here assumes a different form: it is a superposition of sources represented as different types of oscillators that can store and deliver energy in different ways. These are collective effects that generate characteristic features in dielectric response spectra, and can be approached with mechanical models through $\vec{F} = m\vec{a}$

In E&M collective effects are usually the result of the long range of the Coulomb interaction. The distinctive physical feature is feedback from one oscillator to another, and the corresponding mathematics features matrix inversion. Solutions of coupled oscillators are linear combinations of the "bare" states that exist in the absence of coupling. This not only opens up the possibility of interference, but also generates new types of resonances depending on how denominators go to zero.

Collective effects are particularly evident in plasmonics, where structures can be engineered to possess multiple modes, some of which are dipole-active ("bright") and others not ("dark"). The large number of options results in a wide range of physical phenomena, including not only EIT but also lineshape distortions, enhancement of oscillator strengths, and so forth. Probably the most famous work is the 1961 paper of Fano, which is now one of the most-cited (and probably one of the most misunderstood) papers in all of physics.

Although many collective systems are quantum-mechanical in nature, as noted above, most have classical analogs consisting of masses connected by springs. These are typically minor extensions of what we used in Ch. 7 to develop the atomic-scale basis of ε. Although mechanical models miss some essential features (zero-point energies, for example), the solutions are easier to understand. Here, we develop several classical analogs of quantum phenomena and use them to extract the physics. The classical analog to Fano's treatment of the absorption of radiation of a He atom in a weakly absorbing continuum background also exists, but the math is somewhat intimidating. I reserve that discussion until later.



The main points follow by comparing the classical and quantum-mechanical solutions of the displacement of a mass *m*, carrying a charge *q*, connected to a wall by a spring of spring constant κ, and acted on by a time-independent electric field $\vec{E}$, as shown in the figure. The displacement Δx is calculated classically by solving

$$F = m\frac{\partial^2 \Delta x}{\partial t^2} = qE - \kappa \Delta x = 0 . \tag{12.152}$$

The solution is

$$\Delta x = \frac{qE}{\kappa} . \tag{12.153}$$

The quantum-mechanical calculation is based on energy, not force, and starts with the Hamiltonian

$$\left(\frac{p^2}{2m} + \frac{1}{2}\kappa x^2 - qEx\right)\Psi = \left(-\frac{h^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2}\kappa x^2 - qEx\right)\Psi = E_\Psi \Psi . \tag{12.154}$$

The wavefunction

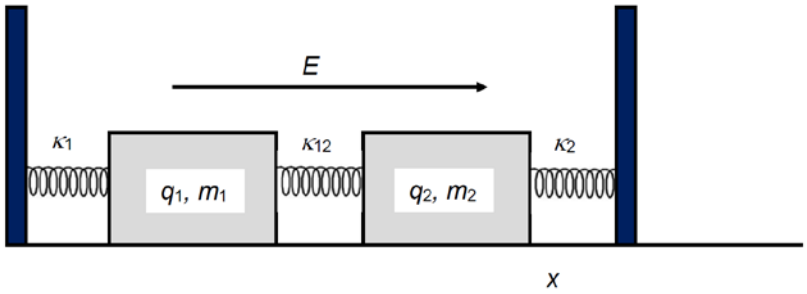
$$\Psi = \varphi_o + \sum_{n\neq 0} a_n \varphi_n \tag{12.155}$$

is expanded as a series of eigenfunctions φ_n about the ground state φ_o. For the harmonic oscillator, the eigenfunctions are related to the Hermite polynomials. After the coefficients a_n are determined, *x* is given to first order by its expectation value $< x >$:

$$< x > = qE \sum_{n\neq 0} \frac{2 |< n | x | o >|^2}{E_n - E_o} . \tag{12.156}$$

To interpret the result, we return to the classical model, where we recognize the sum as κ^{-1}. Thus while mechanical models may not provide as much detail, they provide a quicker path to the solution, and make the quantum-mechanical solution more easily understood.

We now consider collective motion. The model is shown in the figure. It consists of a second spring-mass combination coupled to the first with a third spring of spring constant κ_{12}.



Assuming a general time dependence $\varepsilon^{-i\omega t}$, the coupled one-dimensional equations of motion are

$$-m_1\omega^2\Delta x_1 = q_1 E_o + i\omega b_1 \Delta x_1 - \kappa_1 \Delta x_1 - \kappa_{12}(\Delta x_1 - \Delta x_2); \tag{12.157a}$$

$$-m_2\omega^2\Delta x_2 = q_2 E_o + i\omega b_2 \Delta x_2 - \kappa_2 \Delta x_2 - \kappa_{12}(\Delta x_2 - \Delta x_1). \tag{12.157b}$$

where Δx_1 and Δx_2 are the displacements relative to the equilibrium positions. As usual, ε is given in terms of Δx_1 and Δx_2 by

$$D_o = \varepsilon E_o = E_o + 4\pi P = E_o + 4\pi n_1 q_1 \Delta x_1 + 4\pi n_2 q_2 \Delta x_2 . \tag{12.158}$$

Defining

$$a_{11} = \kappa_1 - m_1\omega^2 - i\omega b_1 + \kappa_{12}, \tag{12.159a}$$

$$a_{22} = \kappa_2 - m_2\omega^2 - i\omega b_2 - \kappa_{12}, \tag{12.159b}$$

Eqs. (12.160) can be written as

$$\begin{pmatrix} q_1 E_o \\ q_2 E_o \end{pmatrix} = \begin{pmatrix} a_{11} & -\kappa_{12} \\ -\kappa_{12} & a_{22} \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix}. \tag{12.160}$$

with the solution

$$\Delta x_1 = \frac{q_1 a_{22} + q_2 \kappa_{12}}{a_{11}a_{22} - \kappa_{12}^2} E_o ; \tag{12.161a}$$

$$\Delta x_2 = \frac{q_2 a_{11} + q_1 \kappa_{12}}{a_{11}a_{22} - \kappa_{12}^2} E_o . \tag{12.161b}$$

The eigenstates are linear combinations of the original "bare" states, and can be reduced to the original equations by setting $\kappa_{12} = 0$.

The new physics arises because two original states can now interfere with each other, as described by the numerator, while the final states share a common set of poles, as described by the denominator. Both aspects have significant consequences. And despite their apparent simplicity, Eqs. (12.162) exhibit a wide range of behavior depending on the values of their nine parameters. This is one reason why the electrodynamics of collective systems can be confusing.

The scope of Eqs. (12.162) is best appreciated via special cases. To simplify notation we define the original "bare" eigenfrequencies as $\omega_{01}^2 = \kappa_1/m_1$ and $\omega_{02}^2 = \kappa_2/m_2$, and the new "dressed" eigenfrequencies as $\omega_1^2 = (\kappa_1 + \kappa_{12})/m_1$ and $\omega_2^2 = (\kappa_2 + \kappa_{12})/m_2$. To reduce notational complexity we will also define a mean frequency $\omega_{12}^2 = \kappa_{12}^2/(4m_1m_2\omega_1\omega_2) \sim \kappa_{12}^2$. The "dressed" frequencies are higher than their "bare" equivalents, an artifact of the mechanical model. In the numerical results that follow, we ignore this shift because we are interested in lineshapes (this is equivalent to redefining ω_1^2 and ω_2^2 for each κ_{12}. In addition, we assume that q_1 and q_2 have the same sign.

We divide the following into effects where both states are "bright" (cases 1-3) and where oscillator #2 is "dark", i.e., where $q_2 = 0$ (cases 4-6). In the former situation we work with both Δx_1 and Δx_2, whereas in the latter case we only need Δx_1.

Case 1: no interaction, $\kappa_{12} = 0$:

$$\Delta x_1 = \frac{q_1}{a_{11}} E_o = \frac{q_1}{m_1(\omega_1^2 - \omega^2 - i2\omega\Gamma_1)} E_o \cong \frac{q_1}{2m_1\omega_1(\omega_1 - \omega - i\Gamma_1)} E_o, \qquad (12.162)$$

with $b_1 = 2\Gamma_1$. A similar expression is obtained for Δx_2. This gives the standard dielectric response of two independent oscillators. The approximation on the right is valid for $\Gamma_1 << \omega_{01}$, which is the usual situation.

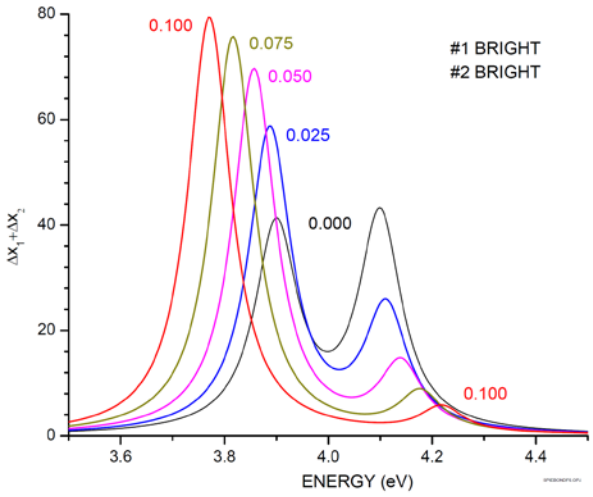
Case 2: κ_{12} "small", and a_{11} and a_{22} mostly real ($\Gamma_1, \Gamma_2 \cong 0$). Equation (12.163) reduces to

$$\Delta x_1 \cong \frac{q_1}{2m_1\omega_1}\left(\frac{(\omega_2 - \omega) \; + \; q_2\kappa_{12}/(2q_1m_2\omega_2)}{(\omega_1 - \omega)(\omega_2 - \omega) \; - \; \kappa_{12}^2/(4m_1m_2\omega_1\omega_2)}\right) E_o, \qquad (12.163)$$

with a corresponding expression Δx_2. What does this mean?

Denominator: the product term is an upward-facing parabola centered on the average oscillator frequency $\overline{\omega} = (\omega_1 + \omega_2)/2$. It thus generates two poles. For $\kappa_{12} = 0$, these poles occur at the "bare" frequencies ω_{01} and ω_{02}. Because the term proportional to κ_{12}^2 is positive definite, as κ_{12} increases the parabola moves downward. As a result, the intersections that generate the two structures move further apart. This is the geometric version of the well-known result that interactions between two quantum levels increases their energy separation.

Numerator: for $\kappa_{12} = 0$ the oscillators are independent with approximately equal lineshapes, as shown in the figure. The figure shows $\mathrm{Im}(\Delta x_1 + \Delta x_2) \sim \varepsilon_2$ for increasing κ_{12}, starting with $\kappa_{12} = 0$. As κ_{12} increases, Interference becomes more and more of a factor. It is constructive for oscillator #1 and destructive for oscillator #2. The result is a shift of oscillator strength from oscillator #2 to oscillator #1. At the same time the features move farther apart. While $\mathrm{Im}(\Delta x_1 + \Delta x_2) > 0$ for all κ_{12}, as required by causality, for larger values Δx_1 actually goes negative in the frequency range of the upper structure. This shows that the energy stored kinetically in mass #1 feeds into mass #2 in this region of the spectrum.



Those familiar with condensed-matter physics recognize the figure on the previous page as the quantum-mechanical two-level model. If a coupling interaction occurs between two levels, the levels move apart and oscillator strength shifts from the upper transition to the lower transition. For this reason the energy gap of Ge is less than that of GaAs.

Case 3: Mixing of real and imaginary parts of the "bare" response of a feature in a resonance lineshape. This occurs if one oscillator is strongly absorbing, with a resonance frequency well removed from the transition of interest. Let Γ_2 be so large that $a_{22} \cong -2i m_2 \omega_{02} \Gamma_2$ and the term $\sim \kappa_{12}^2$ in the denominator can be ignored. Then

$$\Delta x_1 \cong \frac{q_1}{2m_1\omega_1} \frac{1 + i\kappa_{12}/(2\omega_{01} m_1 \Gamma_2)}{\omega_{01} - \omega - i\Gamma_1} E_o . \qquad (12.164)$$

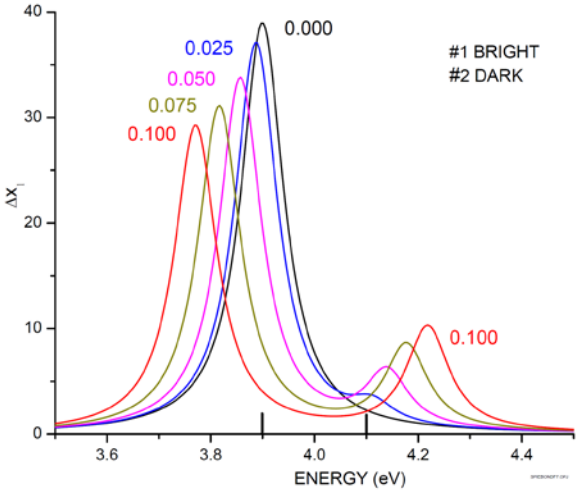
Δx_1 now has a significant imaginary part.

In the remaining examples, oscillator #2 is "dark", which is realized by setting $q_2 = 0$. Thus oscillator #2 is entirely parasitic, being neither driven by *E* nor contributing to ε. The results now depend strongly on Γ_1 and Γ_2, so they are retained throughout. The relevant expression is

$$\Delta x_1 \cong \frac{q_1}{2m_1\omega_1} \left(\frac{\omega_2 - \omega - i\Gamma_2}{(\omega_1 - \omega - i\Gamma_1)(\omega_2 - \omega - i\Gamma_2) - \kappa_{12}^2/(4m_1 m_2 \omega_1 \omega_2)} \right) E_o . \qquad (12.165)$$

Case 4: Appearance of a feature at the resonance frequency of the parasitic oscillator.

The figure is calculated with the same parameters used previously. Note that the "dark" feature does not exist for $\kappa_{12} = 0$, but increases as κ_{12} increases. This is not surprising, because as κ_{12} increases an increasing amount of the response of oscillator #2 is mixed into that of oscillator #1.

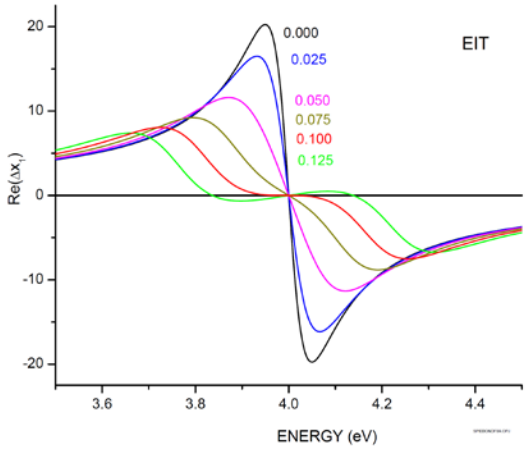


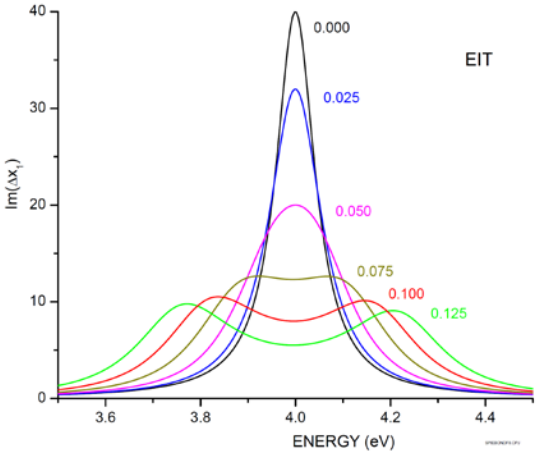
Case 5: $\mathrm{Re}(\varepsilon) = 1$ with zero dispersion: the case of the disappearing oscillator.

The broadening parameters play an essential role in this and the next (EIT) case, so we cast Eq. (12.166) in a more useful form for interpretation. Define $\delta\omega = \omega - \omega_{01}$, and rationalize the denominator. The algebra is tedious but straightforward, and after the math is over, Eq. (12.166) is

$$\Delta x_1 \cong \frac{q}{2m_1\omega_{01}}\left(\frac{\delta\omega(\omega_{12}^2 - \Gamma_2^2 - \delta\omega^2) + i(\delta\omega^2\Gamma_1 + \omega_{12}^2\Gamma_2 + \Gamma_1\Gamma_2^2)}{(\delta\omega^2 - \omega_{12}^2 - \Gamma_1\Gamma_2)^2 + \delta\omega^2(\Gamma_1 + \Gamma_2)^2}\right)E_o \qquad (12.166)$$

If $\omega_{12}^2 = \Gamma_2^2 = \kappa_{12}^2/(4m_1m_2\omega_1\omega)$, the real part of Δx_1 vanishes identically to third order in $\delta\omega$ at $\delta\omega = 0$. This is a remarkable result. If these are the only two oscillators in the configuration, then $\varepsilon = 1$ in the vicinity of ω_1, the group velocity is *c*, dispersion vanishes, and except for the absorptive component, the material is perfectly matched to air and simply disappears. The value of the imaginary part at the minimum is





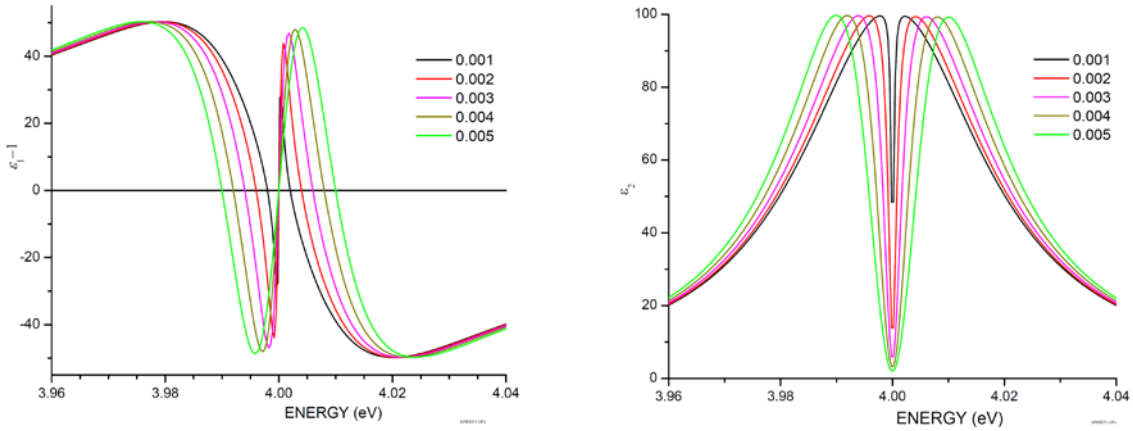
$$\mathrm{Im}(\Delta x_1) = \frac{q}{2m_1\omega_{01}}\frac{\Gamma_2}{\Gamma_1\Gamma_2 + \omega_{12}^2}E_o = \frac{q}{2m_1\omega_{01}}\frac{1}{\Gamma_1 + \Gamma_2}E_o, \qquad (12.167)$$

using the value of Γ_2 under these conditions. Thus as the broadening is increased, the minimum value is reduced. This effect is due in part to the increased separation between the energies of the oscillators as κ_{12} increases, noting that Γ_2 is tied to κ_{12}.

This is illustrated in the two figures, which show this behavior for the parameters used in the previous two figures except $\omega_{01} = \omega_{02} = 4.0$ and $\Gamma_2 = 0.20$. The maximally flat condition occurs at $\omega_{12} = 0.10$. At this value absorption is significantly reduced, approximately by a factor of 5. It is seen that reduction occurs because of the separation of the resonance energies of the two oscillators.

Case 6: electromagnetic-induced transparency (EIT). This is probably the strangest situation of all. The optimal effect is realized if the configuration has a *lossless* parasitic oscillator resonating at exactly the same frequency as the "bright" oscillator. With $\omega_{01} = \omega_{02}$ and $\delta\omega$ defined as in case 5, the absorptive part of Eq. (12.166) reduces to

$$\mathrm{Im}(\Delta x_1) = \frac{q}{2m_1\omega_1}\left(\frac{\Gamma_2}{\omega_{12}^2 + \Gamma_1\Gamma_2}\right)E_o . \tag{12.168}$$



This is also a remarkable result, because the absorption vanishes completely for $\Gamma_2 = 0$, even though the active oscillator is strongly absorbing. Supposing that $\Gamma_2 = 0$, we can evaluate the maximum value of $\mathrm{Im}(\Delta x_1)$ about $\delta\omega = 0$, which occurs at two locations: $\omega_{\pm} = \pm\omega_{12}$. The value itself is

$$\mathrm{Im}(\Delta x_1)_{\max} = \frac{q}{2m_1\omega_1\Gamma_1}E_o . \tag{12.169}$$

Also relevant is $\mathrm{Re}(\Delta x_1)$ at $\delta\omega = 0$:

$$\mathrm{Re}(\Delta x_1) = \frac{q}{2m_1\omega_1}\frac{\delta\omega}{\omega_{12}^2}E_o . \tag{12.170}$$

The group velocity here is therefore

$$v_g = \frac{d\omega}{dk} = \frac{c}{n + \omega(dn/d\omega)}\bigg|_{\omega_{10}} . \tag{12.171}$$

Now for small $\delta\omega$, we have

$$\varepsilon = 1 + \frac{\omega_p^2 \delta\omega}{2\omega_{10}\omega_p^2}. \tag{12.172}$$

The group velocity for $\delta\omega \sim 0$ is therefore

$$v_g = \frac{c}{n + \omega(dn/d\omega)} = \frac{4c\omega_{12}^2}{\omega_p^2 + 4\omega_{12}^2} \cong 4c\frac{\omega_{12}^2}{\omega_p^2}. \tag{12.173}$$

Thus for weak interactions $v_g << c$. "Slow light" experiments capitalize on this effect.