CHAPTER 12.5

OPTICAL MODULATION AND DEMODULATION

Joseph L. Chovan

MODULATION OF BEAMS OF RADIATION

This discussion of optical modulators is restricted to devices that operate on a directed beam of optical energy to control its intensity, phase, or frequency, according to some time-varying modulating signal. Devices that deflect a light beam or spatially modulate a light beam, such as light-valve projectors, are treated in Chap. 21.

Phase or frequency modulation requires a coherent light source, such as a laser. Optical heterodyning is then used to shift the received signal to lower frequencies, where conventional FM demodulation techniques can be applied.

Intensity modulation can be used on incoherent as well as coherent light sources. However, the properties of some types of intensity modulators are wavelength-dependent. Such modulators are restricted to monochromatic operation but not limited to the extremely narrow laser line widths required for frequency modulation.

Optical modulation depends on either perturbing the optical properties of some material with a modulating signal or mechanical motion to interact with the light beam. Modulation bandwidths of mechanical modulators are limited by the inertia of the moving masses. Optical-index modulators generally have a greater modulation bandwidth but typically require critical and expensive optical materials.

Optical-index modulation can be achieved with electric or magnetic fields or by mechanical stress. Typical modulator configurations are presented below, as in heterodyning, which is often useful in demodulation. Optical modulation can also be achieved using semiconductor junctions.

OPTICAL-INDEX MODULATION: ELECTRIC FIELD MODULATION

Pockels and Kerr Effects

In some materials, an electric field vector **E** can produce a displacement vector **D** whose direction and magnitude depend on the orientation of the material. Such a material can be completely characterized in terms of three independent dielectric constants associated with three mutually perpendicular natural directions of the material. If all three dielectric constants are equal, the material is *isotropic*. If two are equal and one is not, the material is *uniaxial*. If all three are unequal, the material is *biaxial*.

The optical properties of such a material can be described in terms of the *ellipsoid of wave normals* (Fig. 12.5.1). This is an ellipsoid whose semiaxes are the square roots of the associated dielectric constants. The behavior of any plane monochromatic wave through the medium can be determined from the ellipse formed by the intersection of the ellipsoid with a plane through the center of the ellipsoid and perpendicular to the direction of wave travel. The instantaneous electric field vector **E** associated with the optical wave has components along the two axes of this ellipse. Each of these components travels with a phase velocity that is inversely proportional to the length of the associated ellipse axis.

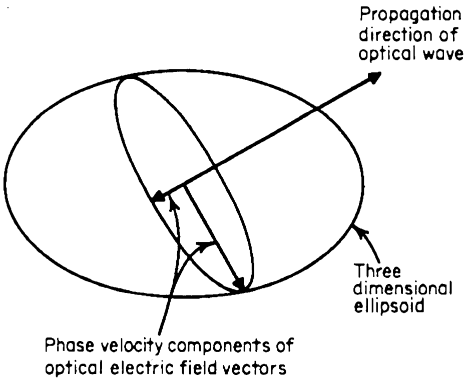

FIGURE 12.5.1 Ellipsoid of wave normals.

Thus there is a differential phase shift between the two orthogonal components of the electric field vector after it has traveled some distance through such a birefringent medium. The two orthogonal components of the vector vary sinusoidally with time but have a phase difference between them, which results in a vector whose magnitude and direction vary to trace out an ellipse once during each optical cycle. Thus linear polarization is converted into elliptical polarization in a birefringent medium.

In some materials it is possible to induce a perturbation in one or more of the ellipsoid axes by applying an external electric field. This is the electrooptical effect. The electrooptical effect is most commonly used in optical modulators presently available. More detailed configurations using these effects are discussed later. Kaminow and Turner (1966), present design considerations for various configurations and tabulates material properties.

Stark Effect

Materials absorb and emit optical energy at frequencies which depend on molecular or atomic resonances characteristic of the material. In some materials an externally applied electric field can perturb these natural resonances. This is known as the Stark effect.

Kaminow and Turner (1966) discusses a modulator for the CO_2 laser on the 3- to 22-μm region. The laser output is passed through an absorption cell whose natural absorption frequency is varied by the modulating signal, using the Stark effect. Since the laser frequency remains fixed, the amount of absorption depends on how closely the absorption cell is tuned to the laser frequency—intensity modulation results.

MAGNETIC FIELD MODULATION

Faraday Effect

Two equal-length vectors circularly rotating at equal rates in opposite directions in space combine to give a nonrotating resultant whose direction in space depends on the relative phase between the counterrotating components. Thus any linearly polarized light wave can be considered to consist of equal right and left circularly polarized waves.

In a material which exhibits the Faraday effect, an externally applied magnetic field causes a difference in the phase velocities of right and left circularly polarized waves traveling along the direction of the applied magnetic field. This results in a rotation of the electric field vector of the optical wave as it travels through the material. The amount of the rotation is controlled by the strength of a modulating current producing the magnetic field.

Zeeman Effect

In some materials the natural resonance frequencies at which the material emits or absorbs optical energy can be perturbed by an externally applied magnetic field. This is known as the Zeeman effect.

Intensity modulation can be achieved using an absorption cell modulated by a magnetizing current in much the same manner as the Stark effect absorption cell is used. The Zeeman effect has also been used to tune the frequency at which the active material in a laser emits.

MECHANICAL-STRESS MODULATION

In some materials the ellipsoid of optical-wave normals can be perturbed by mechanical stress. An acoustic wave traveling through such a medium is a propagating stress wave that produces a propagating wave of perturbation in the optical index.

When a sinusoidal acoustic wave produces a sinusoidal variation in the optical index of a thin isotropic medium, the medium can be considered, at any instant of time, as a simple phase grating. Such a grating diffracts a collimated beam of coherent light into discrete angles whose separation is inversely proportional to the spatial period of the grating.

This situation is analogous to an rf carrier phase-modulated by a sine wave. A series of sidebands results which correspond to the various orders of diffracted light. The amplitude of the *m*th order is given by an *m*th-order Bessel function whose argument depends on the peak phase deviation produced by the modulating signal. The phases of the sidebands are the appropriate integral multiples of the phase of the modulating signal.

The *m*th order of diffracted light has its optical frequency shifted by *m* times the acoustic frequency. The frequency is increased for positive orders and decreased for negative orders.

Similarly, a thick acoustic grating refracts light mainly at discrete input angles. This condition is known as *Bragg reflection* and is the basis for the *Bragg modulator* (Fig. 12.5.2). In the Bragg modulator, essentially all the incident light can be refracted into the desired order, and the optical frequency is shifted by the appropriate integral multiple of the acoustic frequency.

Figure 12.5.2 shows the geometry of a typical Bragg modulator. The input angles for which Bragg modulation occurs are given by

$$\sin\theta = m\lambda/2\Lambda$$

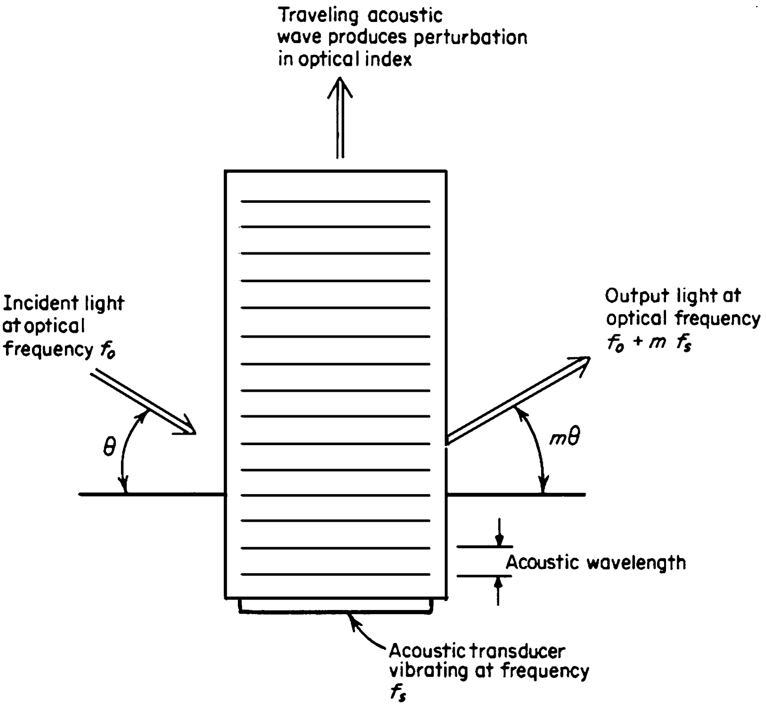

FIGURE 12.5.2 The Bragg modulator.

where θ = angle between propagation of input optical beam and planar acoustic wavefronts
λ = optical wavelength in medium
Λ = acoustic wavelength in medium
$m = \pm 1, \pm 2, \pm 3, \ldots$
$m\theta$ = angle between propagation direction of output optical beam and planar acoustic wavefronts

The ratio of optical to acoustic wavelength is typically quite small, and m is a low integer, so that the angle θ is very small. Critical alignment is thus required between the acoustic wavefronts and the input light beam.

If the modulation bandwidth of the acoustic signal is broad, the acoustic wavelength varies, so that there is a corresponding variation in the angle θ for which Bragg reflection occurs. To overcome this problem, a phased array of acoustic transducers is often used to steer the angle of the acoustic wave as a function of frequency in the desired manner.

A limitation on bandwidth is the acoustic transit time across the optical beam. Since the phase grating in the optical beam at any instant of time must be essentially constant frequency if all the light is to be diffracted at the same angle, the bandwidth is limited so that only small changes can occur in this time interval.

MODULATOR CONFIGURATIONS: INTENSITY MODULATION

Polarization Changes

Linearly polarized light can be passed through a medium exhibiting an electrooptical effect and the output beam passed through another polarizer. The modulating electric field controls the eccentricity and orientation of the elliptical polarization and hence the magnitude of the component in the direction of the output polarizer. Typically, the input linear polarization is oriented to have equal components along the fast and slow axes of the birefringent medium, and the output polarizer is orthogonal to the input polarizer. The modulating field causes a phase differential varying from 0 to π rad. This causes the polarization to change from linear (at 0) to circular (at $\pi/2$) to linear normal to the input polarization (at π). Thus the intensity passing through the output polarizer varies from 0 to 100 percent as the phase differential varies from 0 to π rad.

Figure 12.5.3 shows this typical configuration. The following equations relate the optical intensity transmission of this configuration to the modulation.

$$I_o/I_i = \tfrac{1}{2}(1 - \cos\phi)$$

where I_o = output optical intensity
I_i = input optical intensity
ϕ = differential phase shift between fast and slow axes.

In the Pockels effect the differential phase shift is linearly related to applied voltage; in the Kerr effect it is related to the voltage squared.

$$\phi = \begin{cases} \pi v/V & \text{Pockels effect} \\ \pi (v/V)^2 & \text{Kerr effect} \end{cases}$$

where v is the modulation voltage and V is the voltage to produce π rad differential phase shift.

Figure 12.5.4 shows the intensity transmission given by the above expression. The most linear part of the modulation curve is at $\phi = \pi/2$. Often a quarter-wave plate is added in series with the electrooptical material to provide this fixed bias at $\pi/2$. A fixed-bias voltage on the electrooptical material can also be used.

This arrangement is probably the most commonly used broadband intensity modulator. Early modulators of this type used a uniaxial Pockels cell with the electric field in the direction of optical propagation. In this arrangement, the induced phase differential is directly proportional to the opticalpath length, but the electric field is inversely proportional to this path length (at fixed voltage). Thus the phase differential is independent of the path length and depends only on applied voltage. Typical materials require several kilovolts for a differential phase shift of π in the visible-light region.

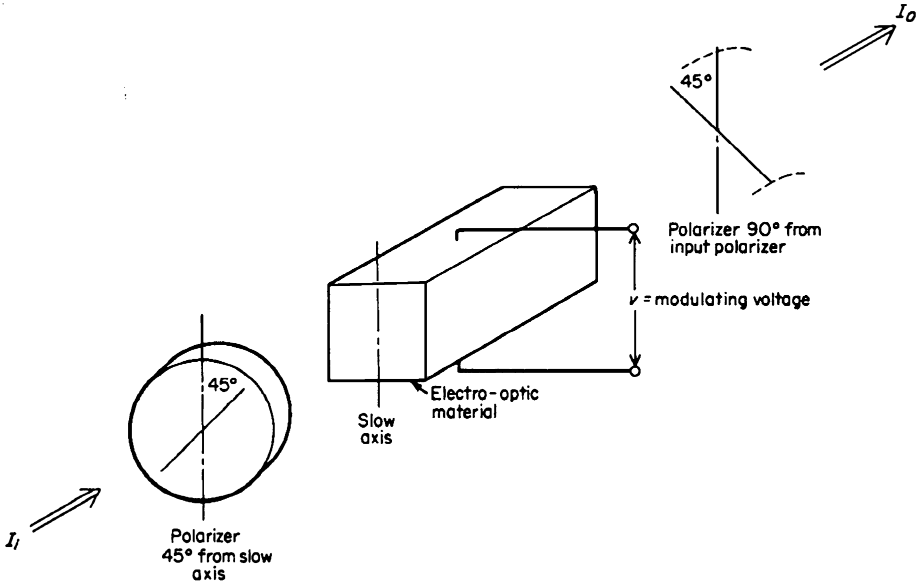

FIGURE 12.5.3 Electrooptical intensity modulator.

Since the Pockels cell is essentially a capacitor, the energy stored in it is $\frac{1}{2}CV^2$ where C is the capacitance and V is the voltage. This capacitor must be discharged and charged during each modulation cycle. Discharge is typically done through a load resistor, where this energy is dissipated. The high voltages involved mean that the dissipated power at high modulation rates is appreciable.

The high-voltage problem can be overcome by passing light through the medium in a direction normal to the applied electric field. This permits a short distance between the electrodes (so that a high-E field is obtained from a low voltage) and a long optical path in the orthogonal direction (so that the cumulative phase differential is experienced).

Unfortunately, materials available are typically uniaxial, having a high eccentricity in the absence of electric fields. When oriented in a direction that permits the modulating electric field to be orthogonal to the propagation direction, the material has an inherent phase differential which is orders of magnitude greater than that induced by the modulating field. Furthermore, minor temperature variations cause perturbations in this phase differential which are large compared with those caused by modulation.

This difficulty is overcome by cascading two crystals which are carefully oriented so that temperature effects in one are compensated for by temperature effects in the other. The modulation electrodes are then connected so that their effects add. Commercially available electrooptical modulators are of this type.

The Kerr effect is often used in a similar arrangement. Kerr cells containing nitrobenzene are commonly used as high-speed optical shutters.

Polarization rotation produced by the Faraday effect is also used in intensity modulation by passing through an output polarizer in a manner similar to that discussed above. The Faraday effect is more commonly used at wavelengths where materials exhibiting the electrooptical effect are not readily available.

Controlled Absorption

As noted above, the frequency at which a material absorbs energy because of molecular or atomic resonances can be tuned over some small range in materials exhibiting the Stark or Zeeman effect. Laser spectral widths are typically narrow compared with such an absorption line width. Thus the absorption of the narrow laser line can be modulated by tuning the absorption frequency over a range near the laser frequency. Although such modulators have been used, they are not as common as the electrooptical modulators discussed above.

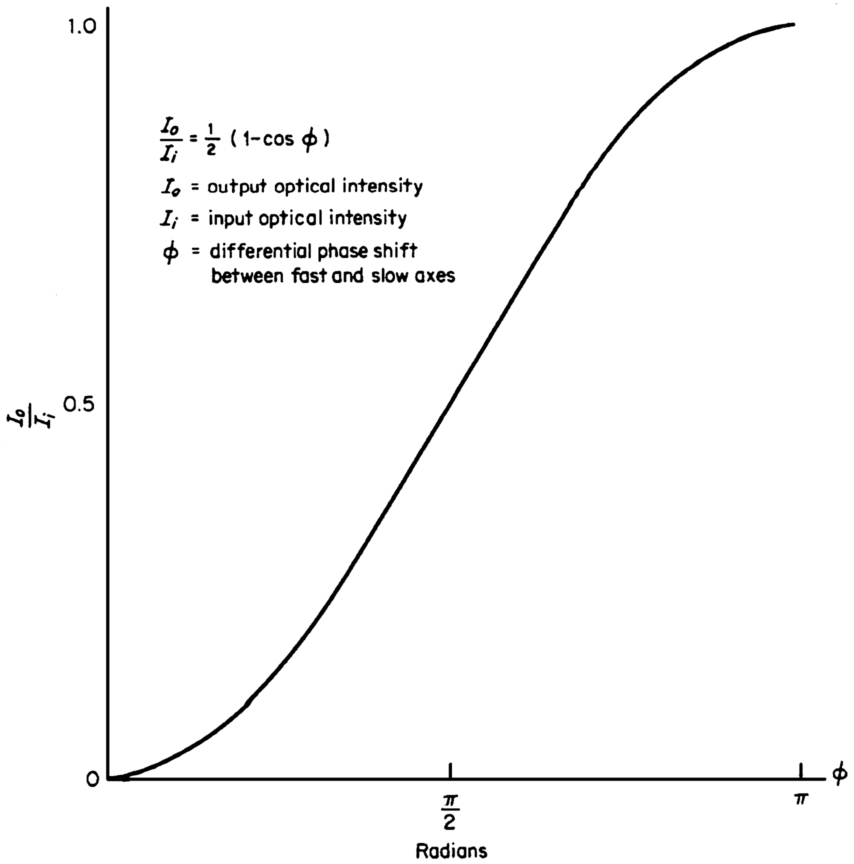

FIGURE 12.5.4 Transmission of electrooptical intensity modulator.

PHASE AND FREQUENCY MODULATION OF BEAMS

Laser-Cavity Modulation

The distance between mirrors in a laser cavity must be an integral number of wavelengths. If this distance is changed by a slight amount, the laser frequency changes to maintain an integral number. The following equation relates the change in cavity length to the change in frequency:

$$\Delta f = \frac{C}{L}\frac{\Delta L}{\lambda}$$

where Δf = change in optical frequency
ΔL = change in laser-cavity length
L = laser-cavity length
λ = optical wavelength of laser output
C = velocity of light in laser cavity.

In a cavity 1 m long, a change in mirror position of one optical wavelength produces about 300 MHz frequency shift. Thus a laser can be frequency-modulated by moving one of its mirrors with an acoustic transducer, but the mass of the transducer and mirror limit the modulation bandwidths that can be achieved.

An electrooptical cell can be used in a laser optical cavity to provide changes in the optical-path length. The polarization is oriented so that it lies entirely along the axis of the modulated electrooptical material. This produces the same effect as moving the mirror but without the inertial restrictions of the mirror's mass.

Under such conditions, the ultimate modulation bandwidth is limited by the Q of the laser cavity. A light beam undergoes several reflections across the cavity, depending on the Q, before an appreciable portion of it is coupled out. The laser frequency must remain essentially constant during the transit time required for these multiple reflections. This limits the upper modulation frequency.

Modulation of the laser-cavity length produces a conventional FM signal with modulating signal directly proportional to change in laser-cavity length. Demodulation is conveniently accomplished by optical heterodyning to lower rf frequencies where conventional FM demodulation techniques can be used.

EXTERNAL MODULATION

The Bragg modulator (Fig. 12.5.2) is commonly used to modulate the optical frequency. As such it produces a single-sideband suppressed-carrier type of modulation.

Demodulation can be achieved by optical heterodyning to lower rf frequencies, where conventional techniques can be employed for this type of modulation. It is also possible to reinsert the carrier at the transmitter for a frequency reference. This is done by using optical-beam splitters to combine a portion of the unmodulated laser beam with the Bragg modulator output.

Conventional double-sideband amplitude modulation has also been achieved by simultaneously modulating two laser beams (derived from the same source) with a common Bragg modulator to obtain signals shifted up and down. Optical-beam splitters are used to combine both signals with an unmodulated carrier. Conventional power detection demodulates such a signal.

Optical phase modulation is commonly accomplished by passing the laser output beam through an electrooptical material, with the polarization vector oriented along the modulated ellipsoid axis of the material. Demodulation is conveniently achieved by optical heterodyning to rf frequencies, FM demodulation, and integrating to recover the phase modulation in the usual manner.

For low modulation bandwidths, the electrooptical material can be replaced by a mechanically driven mirror. The light reflected from the mirror is phase modulated by the changes in the mirror position. This effect is often described in terms of the Doppler frequency shift, which is directly proportional to the mirror velocity and inversely proportional to the optical wavelength.

TRAVELING-WAVE MODULATION

In the electrooptical and magnetooptical modulators described thus far, it is assumed that the modulating signal is essentially constant during the optical transit time through the material. This sets a basic limit on the highest modulating frequency that can be used in a lumped modulator.

This problem is overcome in a traveling-wave modulator. The optical wave and the modulation signal propagate with equal phase velocities through the modulating medium, allowing the modulating fields to act on the optical wave over a long path, regardless of how rapidly the modulating fields are changing. The degree to which the two phase velocities can be matched determines the maximum interaction length possible.

OPTICAL HETERODYNING

Two collimated optical beams, derived from the same laser source and illuminating a common surface, produce straight-line interference fringes. The distance between fringes is inversely proportional to the angle between the beams. Shifting the phase of one of the beams results in a translation of the interference pattern, such that a 2π-rad phase shift translates the pattern by a complete cycle. An optical detector having a sensing area small compared with the fringe spacing has a sinusoidal output as the sinusoidal intensity of the interference pattern translates across the detector.

A frequency difference between the two optical beams produces a phase difference between the beams that changes at a constant rate with time. This causes the fringe pattern to translate across the detector at a constant rate, producing an output at the difference frequency. This technique is known as *optical heterodyning* in which one of the beams is the signal beam, the other the local oscillator.

The effect of the optical alignment between the beams is evident. As the angle between the two collimated beams is reduced, the spacing between the interference fringes increases, until the spacing becomes large compared with the overall beam size. This permits a large detector which uses all the light in the beam. If converging or diverging beams are used instead of collimated beams, the situation is similar, except that the interference fringes are curved instead of straight. Making the image of the local-oscillator point coincide with the image of the signal-beam point causes the desired infinite fringe spacing.

Optical heterodyning provides a convenient solution to several possible problems in optical demodulation. In systems where a technique other than simple amplitude modulation has been used (e.g., single-sideband, frequency, or phase modulation), optical heterodyning permits shifting to frequencies where established demodulation techniques are readily available.

In systems where background radiation, such as from the sun, is a problem, heterodyning permits shifting to lower frequencies, so that filtering to the modulation bandwidth removes most of the broadband background radiation. The required phase front alignment also eliminates background radiation from spatial positions other than that of the signal source.

Many systems are limited by thermal noise in the detector and/or front-end amplifier. Cooled detectors and elaborate amplifiers are often used to reduce this noise to the point that photon noise in the signal itself dominates. This limit also can be achieved in an optical heterodyne system with noncooled detector and normal amplifiers by increasing the local-oscillator power to the point where photon noise in the local oscillator is the dominant noise source. Under these conditions, the signal-to-noise power ratio is given by the following equation:

$$S/N = \eta \lambda P / 2hBC$$

where S/N = signal-power-noise-power ratio
η = quantum efficiency of photo detector
λ = optical wavelength
h = Planck's constant
C = velocity of light
B = bandwidth over which S/N is evaluated
P = optical signal power received by detector