Fundamentals of Photonics
Bahaa E. A. Saleh, Malvin Carl Teich

ISBNs: 0-471-83965-5 (Hardback); 0-471-2-1374-8 (Electronic)

CHAPTER

21

PHOTONIC SWITCHING AND COMPUTING

The ideas of **Johann (John) von Neumann (1903–1957)** had a major influence on the architecture of digital computers. He investigated the use of logic gates based on nonlinear dielectric constants. In 1953 he proposed that stimulated emission in a semiconductor material could be used to provide light amplification, which is the underlying principle for the operation of the semiconductor laser.

Switching is an essential operation in communication networks. It is also a basic operation in digital computers and signal processing systems. The current rapid development of high-data-rate fiber-optic communication systems has created a need for high-capacity repeaters and terminal systems for processing optical signals and, therefore, a need for high-speed photonic switches. Similarly, the potential for optical computing can only be realized if large arrays of fast photonic gates, switches, and memory elements are developed.

This chapter introduces the basic principles of the emerging technologies of photonic switching and optical signal processing. Many of the fundamental principles of photonics, which have been introduced in earlier chapters (Fourier optics and holography, guided-wave optics, electro-optics, acousto-optics, and nonlinear optics), find use here.

Section 21.1 provides a brief introduction to the general types and properties of switches and to photonic switching using opto-mechanical, acousto-optic, magneto-optic, and electro-optic devices. All-optical switches are described in Sec. 21.2. Section 21.3 is devoted to bistable optical devices. These are switches with memory—systems for which the output is one of only two states, depending on the current and previous values of the input. Section 21.4 covers optical interconnections and their applications in optical signal processing and in microelectronics. Finally, Sec. 21.5 outlines the basic features of optical processing and computing systems, both digital and analog.

21.1 PHOTONIC SWITCHES

A. Switches

A switch is a device that establishes and releases connections among transmission paths in a communication or signal-processing system. A control unit processes the commands for connections and sends a control signal to operate the switch in the desired manner. Examples of switches are shown in Fig. 21.1-1.

A 1×1 switch can be used as an elementary unit from which switches of larger sizes can be built. An $N \times N$ crosspoint-matrix (crossbar) switch, for example, may be constructed by using an array of N^2 1×1 switches organized at the points of an $N \times N$ matrix to connect or disconnect each of the N input lines to a free output line [see Fig. 21.1-1(*d*) and Fig. 21.1-2(*a*)]. The mth input reaches all elementary switches of the mth row, while the lth output is connected to outputs of all elementary switches of the lth column. A connection is made between the mth input and the lth output by activating the (m, l) 1×1 switch.

An $N \times N$ switch may also be built by use of 2×2 switches. An example is the 4×4 switch, made by the use of five 2×2 switches in the configuration shown in Fig. 21.1-2(*b*).

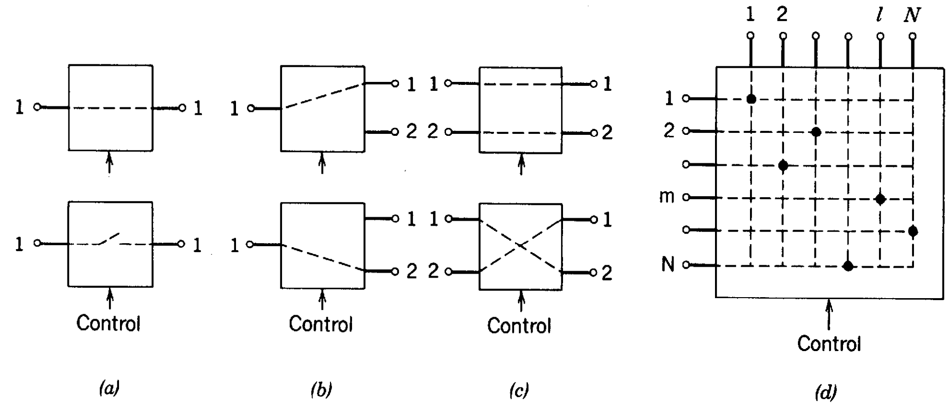


Figure 21.1-1 (*a*) 1×1 switch connects or disconnects two lines. It is an on–off switch. (*b*) 1×2 switch connects one line to either of two lines. (*c*) 2×2 crossbar switch connects two lines to two lines. It has two configurations: the bar state and the cross state. (*d*) $N \times N$ crossbar switch connects N lines to N lines. Any input line can always be connected to a free (unconnected) output line without blocking (i.e., without conflict).

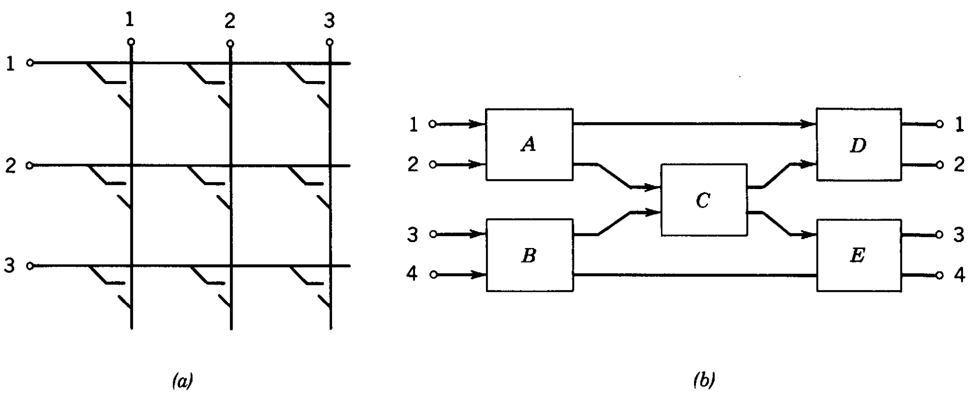


Figure 21.1-2 (*a*) A 3×3 switch made of nine 1×1 switches. (*b*) A 4×4 switch made of five 2×2 switches. Input line 1 is connected to output line 3, for example, if switches A and C are in the cross state and switch E is in the bar state.

A switch is characterized by the following parameters:

- *Size* (number of input and output lines) and *direction*(*s*), i.e., whether data can be transferred in one or two directions.
- *Switching time* (time necessary for the switch to be reconfigured from one state to another).
- *Propagation delay time* (time taken by the signal to cross the switch).
- *Throughput* (maximum data rate that can flow through the switch when it is connected).
- *Switching energy* (energy needed to activate and deactivate the switch).
- *Power dissipation* (energy dissipated per second in the process of switching).
- *Insertion loss* (drop in signal power introduced by the connection).
- *Crosstalk* (undesired power leakage to other lines).
- *Physical dimensions*. This is important when large arrays of switches are to be built.

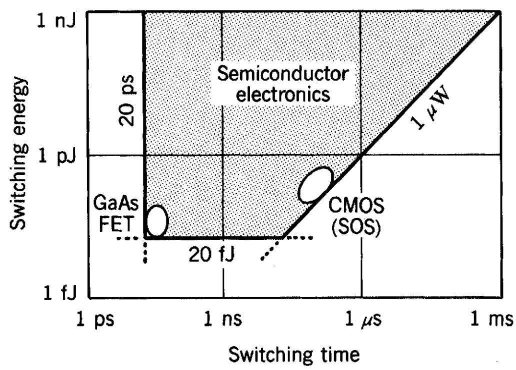


Figure 21.1-3 Limits of switching energy, switching time, and switching power for semiconductor devices. Both silicon-on-sapphire (SOS) complementary-symmetry metal-oxide-semiconductor (CMOS) and GaAs field-effect transistors (FET) are shown.

Electronic switches are used to switch electrical signals. The switch control is either electro-mechanical (using relays) or electronic (using semiconductor enabling logic circuits). Although it is difficult to provide precise limits on the minimum achievable switching time, switching energy, and switching power for semiconductor electronics technology, which continues to advance rapidly, the following bounds are representative of the orders of magnitude:

Minimum switching time	= 10–20 ps
Minimum energy per operation	= 10–20 fJ
Minimum switching power	≈ 1 μW.

Limits of Semiconductor Electronic Switches

These limits are shown schematically in Fig. 21.1-3.

Josephson devices can operate at lower energies (tens of aJ; 1 aJ = 10^{-18} J); a switching time of 1.5 ps has been demonstrated and subpicosecond operations are theoretically possible.

Optical signals may be switched by the use of electronic switches: the optical signals are converted into electrical signals using photodetectors, switched electronically, and then converted back into light using LEDs or lasers (Fig. 21.1-4). These optical/electrical/optical conversions introduce unnecessary time delays and power loss (in addition

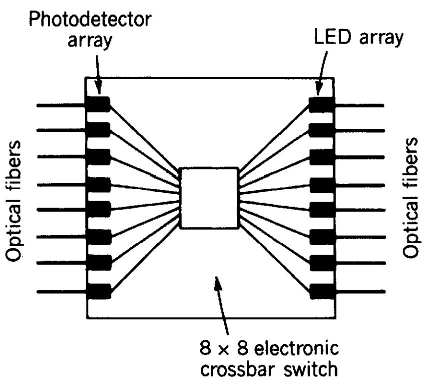


Figure 21.1-4 An optoelectronic 8 × 8 crossbar switch. Eight optical signals carried by eight optical fibers are detected by an array of photodetectors, switched using an 8 × 8 electronic crossbar switch, and regenerated using eight LEDs (or diode lasers) into eight outgoing optical fibers. The data rates that can be handled by silicon switches are currently a few hundred Mb/s, while GaAs switches can operate at rates exceeding 1 Gb/s.

to the loss of the optical phase caused by the process of detection). Direct optical switching is clearly preferable to electronic switching.

B. Opto-Mechanical, Electro-Optic, Acoustic-Optic, and Magneto-Optic Switches

Optical modulators and scanners can be used as switches. A modulator can be operated in the on–off mode as a 1×1 switch. A scanner that deflects an optical beam into N possible directions is a $1 \times N$ switch. These switches can be combined to make switches of higher dimensions.

Modulation and deflection of light can be achieved by the use of mechanical, electrical, acoustic, magnetic, or optical control; the switches are then called opto-mechanical (or mechano-optic), electro-optic, acousto-optic, magneto-optic, or opto-optic (all-optical), respectively. The remainder of this section provides a brief outline of opto-mechanical and magneto-optic switches, and a brief review of electro-optic and acousto-optic switches, which are discussed in Secs. 18.1B and 20.2, respectively. All-optical switches are covered in Sec. 21.2.

Opto-Mechanical Switches

Opto-mechanical switches use moving (rotating or alternating) mirrors, prisms, or holographic gratings to deflect light beams (Fig. 21.1-5). Piezoelectric elements may be

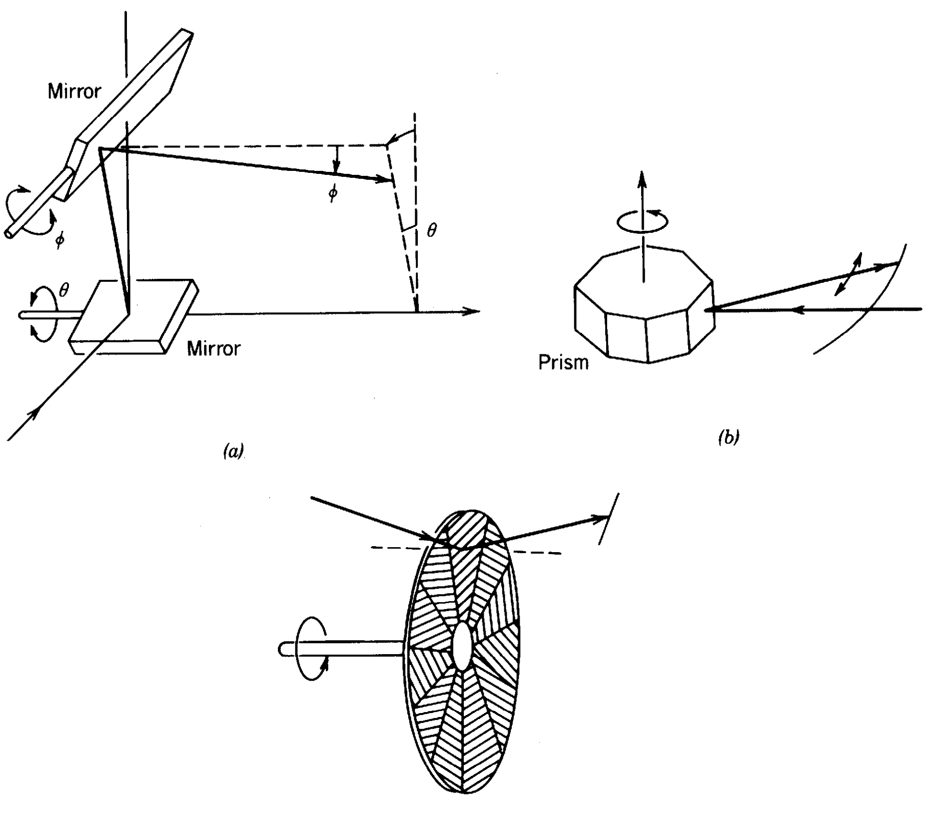


Figure 21.1-5 Deflecting light into different directions using (a) rotating mirrors; (b) a rotating prism; (c) a rotating holographic disk. Each sector of the holographic disk contains a grating whose orientation and period determine a scanning plane and scanning angle of the deflected light.

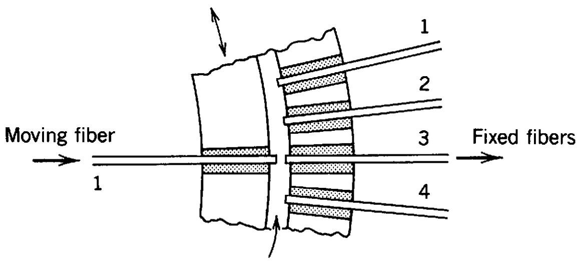


Figure 21.1-6 An optical fiber attached to a rotating wheel is aligned with one of a number of optical fibers attached to a fixed wheel. The fibers are placed in V-grooves. An index-matching liquid is used for better optical coupling.

used for fast mechanical action. A moving drop of mercury in a capillary cell can act as a moving mirror.

An optical fiber can be connected to any of a number of other optical fibers by mechanically moving the input fiber to align with the selected output fiber using a mechanism such as that illustrated in Fig. 21.1-6.

The major limitation of opto-mechanical switches is their low switching speeds (switching times are in the millisecond regime). Their major advantages are low insertion loss and low crosstalk.

Electro-Optic Switches

As discussed in Sec. 18.1, electro-optic materials alter their refractive indices in the presence of an electric field. They may be used as electrically controlled phase modulators or wave retarders. When placed in one arm of an interferometer, or between two crossed polarizers, the electro-optic cell serves as an electrically controlled light modulator or a 1×1 (on–off) switch (see Sec. 18.1B).

Since it is difficult to make large arrays of switches using bulk crystals, the most promising technology for electro-optic switching is integrated optics (see Chap. 7 and Sec. 18.1). Integrated-optic waveguides are fabricated using electro-optic dielectric substrates, such as $LiNbO_3$, with strips of slightly higher refractive index at the locations of the waveguides, created by diffusing titanium into the substrate.

An example of a 1×1 switch using an integrated-optic Mach–Zehnder interferometer is described in Sec. 18.1B and shown in Fig. 21.1-7(*a*). An example of a 2×2 switch is the directional coupler discussed in Sec. 18.1D and illustrated in Fig. 22.1-7(*b*). Two waveguides in close proximity are optically coupled; the refractive index is altered by applying an electric field adjusted so that the optical power either remains in the same waveguide or is transferred to the other waveguide. These switches operate at a few volts with speeds that can exceed 20 GHz.

An $N \times N$ integrated-optic switch can be built by use of a combination of 2×2 switches. A 4×4 switch is implemented by use of five 2×2 switches connected as in Fig. 21.1-2(*b*). This configuration can be built on a single substrate in the geometry shown in Fig. 21.1-8. An 8×8 switch is commercially available and larger switches are being developed.

The limit on the number of switches per unit area is governed by the relatively large physical dimensions of each directional coupler and the planar nature of the interconnections within the chip. To reduce the dimensions and increase the packing density of switches, intersecting (instead of parallel) waveguides are being investigated.

Because of the rectangular nature of integrated-optics technology, it is difficult to obtain efficient coupling to cylindrical waveguides (e.g., optical fibers). Relatively large insertion losses are encountered, especially when a single-mode fiber is connected to a directional coupler. Because the coupling coefficient is polarization dependent, the

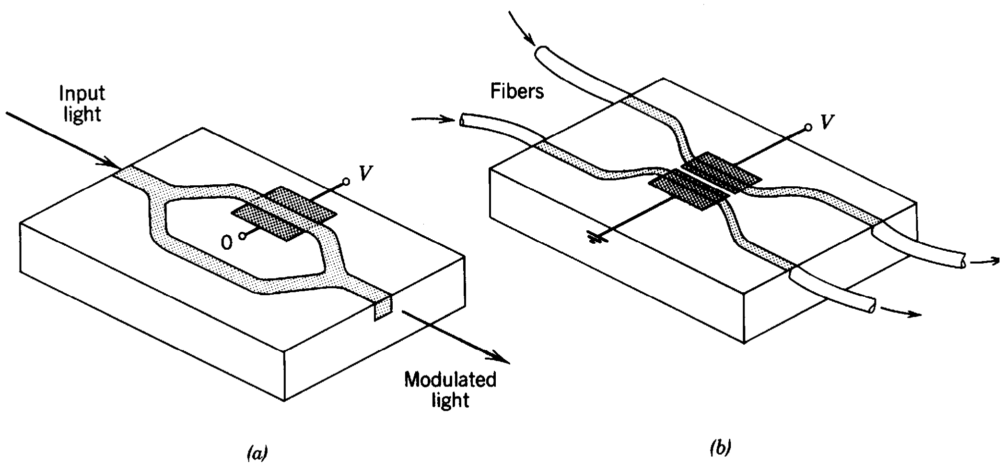


Figure 21.1-7 (*a*) A 1×1 switch using an integrated-optic Mach–Zehnder interferometer. (*b*) A 2×2 switch using an integrated electro-optic directional coupler.

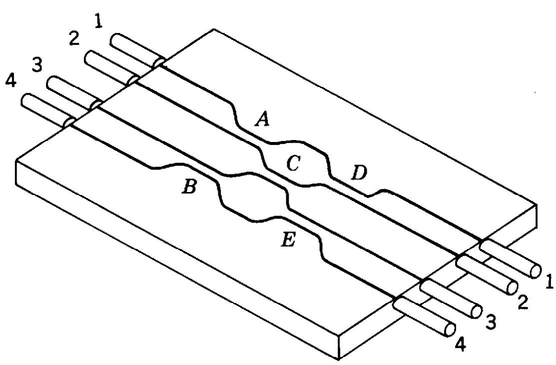


Figure 21.1-8 An integrated-optical 4×4 switch using five directional couplers A, B, C, D, and E on a single substrate.

polarization of the guided light must be properly selected. This imposes a restriction requiring that the input and output connecting fibers must be polarization maintaining (see Sec. 8.1C). Elaborate schemes are required to make polarization-independent switches.

Liquid crystals provide another technology that can be used to make electrically controlled optical switches (see Sec. 18.3). A large array of electrodes placed on a single liquid-crystal panel serves as a spatial light modulator or a set of 1×1 switches. The main limitation is the relatively low switching speed.

Acousto-Optic Switches

Acousto-optic switches use the property of Bragg deflection of light by sound (Chap. 20). The power of the deflected light is controlled by the intensity of the sound. The angle of deflection is controlled by the frequency of the sound. An acousto-optic modulator is a 1×1 switch. An acousto-optic scanner (Fig. 21.1-9) is a $1 \times N$ switch, where N is the number of resolvable spots of the scanner (see Sec. 20.2B). Acousto-optic cells with $N = 2000$ are available. If different parts of the acousto-optic cell carry sound waves of different frequencies, an $N \times M$ switch or interconnection device is

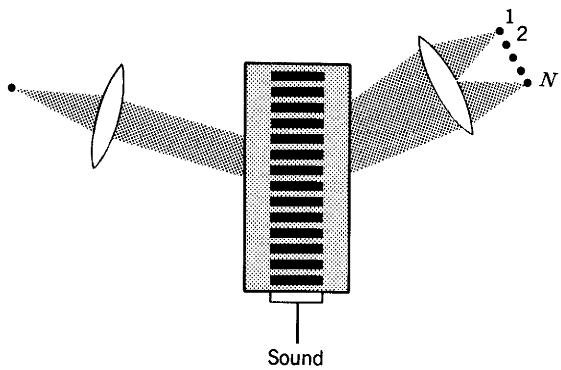


Figure 21.1-9 Acousto-optic switch.

obtained. Limitations on the maximum product NM achievable with acousto-optic cells have been discussed in Sec. 20.2C. Arrays of acousto-optic cells are also becoming available.

Magneto-Optic Switches

Magneto-optic materials alter their optical properties under the influence of a magnetic field. Materials exhibiting the Faraday effect, for example, act as polarization rotators in the presence of a magnetic flux density B (see Sec. 6.4B); the rotatory power ρ (angle per unit length) is proportional to the component of B in the direction of propagation. When the material is placed between two crossed polarizers, the optical power transmission $\mathscr{T} = \sin^2 \theta$ is dependent on the polarization rotation angle $\theta = \rho d$, where d is the thickness of the cell. The device is used as a 1×1 switch controlled by the magnetic field.

Magneto-optic materials have recently received more attention because of their use in optical-disk recording. In these systems, however, a thermomagnetic effect is used in which the magnetization is altered by heating with a strong focused laser. Weak linearly polarized light from a laser is used for readout.

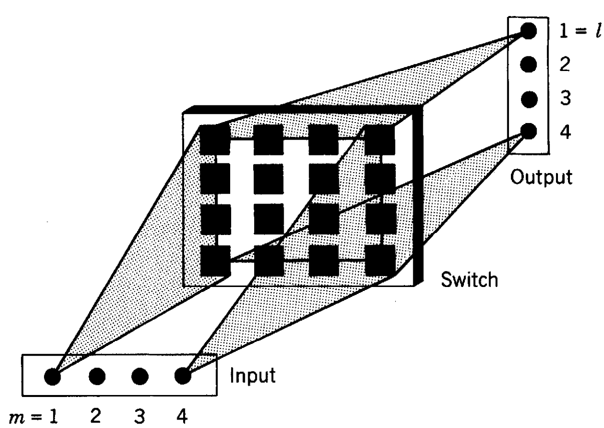


Figure 21.1-10 A 4×4 magneto-optic crossbar switch. Each of the 16 elements is a 1×1 switch transmitting or blocking light depending on the applied magnetic field. Light from the input mth point, $m = 1, 2, 3, 4$ is distributed to all switches in the mth column. Light from all switches of the lth row reaches the lth output point ($l = 1, 2, 3, 4$). The system is an implementation of the 4×4 switch depicted in Fig. 21.1-1(d).

The magneto-optic material is usually in the form of a film (e.g., bismuth-substituted iron garnet) grown on a nonmagnetic substrate. The magnetic field is applied by use of two intersecting conductors carrying electric current. The system operates in a binary mode by switching the direction of magnetization.

Arrays of magneto-optic switches can be constructed by etching isolated cells (each of size as small as 10×10 μm) on a single film. Conductors for the electric-current drive lines are subsequently deposited using usual photolithographic techniques. Large arrays of magneto-optic switches (1024×1024) have become available and the technology is advancing rapidly. Switching speeds of 100 ns are possible. Figure 21.1-10 illustrates the use of a 4×4 array of magneto-optic switches as a 4×4 switch.

21.2 ALL-OPTICAL SWITCHES

In an all-optical (or opto-optic) switch, light controls light with the help of a nonlinear optical material. Nonlinear optical effects may be direct or indirect. Direct effects occur at the atomic or molecular level when the presence of light alters the atomic susceptibility or the photon absorption rates of the medium. The optical Kerr effect (variation of the refractive index with the applied light intensity; see Sec. 19.3A) and saturable absorption (dependence of the absorption coefficient on the applied light intensity; see Sec. 13.3B) are examples of direct nonlinear optical effects.

Indirect nonlinear optical effects involve an intermediate process in which electric charges and/or electric fields play a role, as illustrated by the following two examples.

- In photorefractive materials (see Sec. 18.4), absorbed nonuniform light creates mobile charges that diffuse away from regions of high concentration and are trapped elsewhere, creating an internal space-charge electric field that modifies the optical properties of the medium by virtue of the electro-optic effect.
- In an optically-addressed liquid-crystal spatial light modulator (see Sec. 18.3B), the control light is absorbed by a photoconductive layer and the generated electric charges create an electric field that modifies the molecular orientation and therefore the indices of refraction of the material, thereby controlling the transmission of light.

In these two examples, optical nonlinear behavior is exhibited because of an intermediate effect: light creates an electric field that modifies the optical properties of the medium. Other indirect nonlinear optical effects will be discussed in Sec. 21.3 in connection with bistable optical devices.

Nonlinear optical effects (direct or indirect) may be used to make all-optical switches. The optical phase modulation in the Kerr medium (see Sec. 19.3A), for example, may be converted into intensity modulation by placing the medium in one leg of an interferometer, so that as the control light is turned on and off, the transmittance of the interferometer is switched between 1 and 0, as illustrated in Fig. 21.2-1.

The retardation between two polarizations in an anisotropic nonlinear medium may also be used for switching by placing the material between two crossed polarizers. Figure 21.2-2 illustrates an example of an all-optical switch using an anisotropic optical fiber exhibiting the optical Kerr effect.

An array of switches using an optically-addressed liquid-crystal spatial light modulator is illustrated in Fig. 21.2-3. The control light alters the electric field applied to the liquid-crystal layer and therefore alters its reflectance. Different points on the liquid-crystal surface have different reflectances and act as independent switches controlled

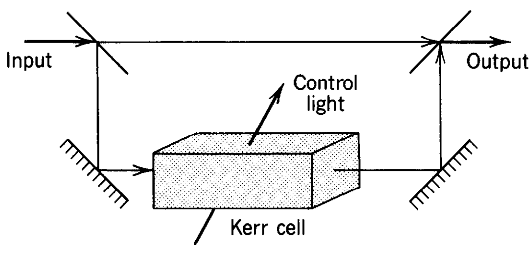


Figure 21.2.1 An all-optical on–off switch using a Mach–Zehnder interferometer and a material exhibiting the optical Kerr effect.

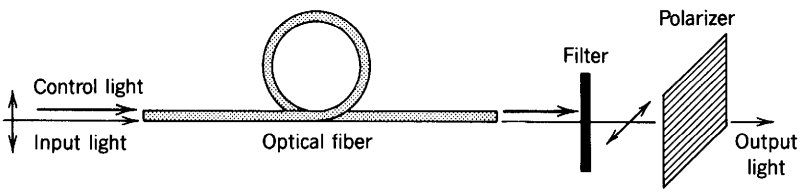


Figure 21.2-2 An anisotropic nonlinear optical fiber serving as an all-optical switch. In the presence of the control light, the fiber introduces a phase retardation π, so that the polarization of the linearly polarized input light rotates 90° and is transmitted by the output polarizer. In the absence of the control light, the fiber introduces no retardation and the light is blocked by the polarizer. The filter is used to transmit the signal light and block the control light, which has a different wavelength.

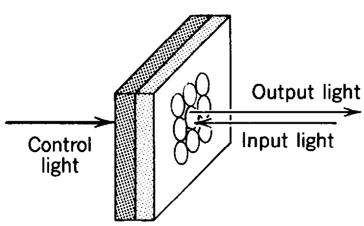


Figure 21.2-3 An all-optical array of switches using an optically addressed liquid-crystal spatial light modulator (light valve).

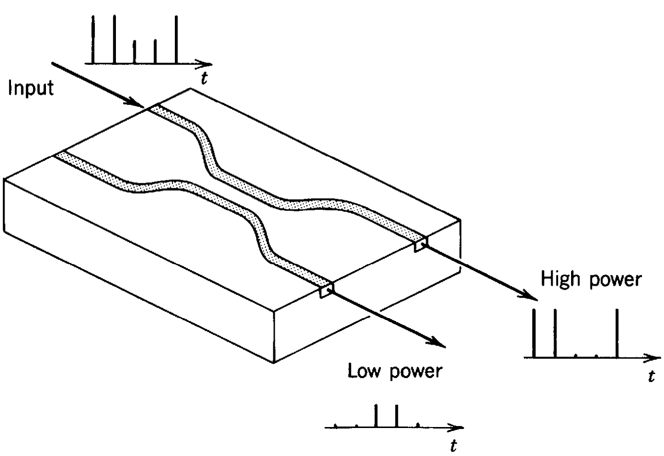


Figure 21.2-4 A directional coupler controlled by the optical Kerr effect. An input beam of low power entering one waveguide is channeled into the other waveguide; a beam of high power remains in the same waveguide.

by the input light beams. These devices can accommodate a large number of switches, but they are relatively slow.

It is not necessary that the control light and the controlled light be distinct. A single beam of light may control its own transmission. Consider, for example, the directional coupler illustrated in Fig. 21.2-4. The refractive indices and the dimensions may be selected so that when the input optical power is low, it is channeled into the other waveguide; when it is high the refractive indices are altered by virtue of the optical Kerr effect and the power remains in the same waveguide. The device serves as a self-controlled (self-addressed) switch. It can be used to sift a sequence of weak and strong pulses, separating them into the two output ports of the coupler. All-optical gates and optical-memory elements made of nonlinear optical materials will be discussed in Sec. 21.3.

Fundamental Limitations on All-Optical Switches

Minimum values of the switching energy E and the switching time T of all-optical switches are governed by the following fundamental physical limits.

Photon-Number Fluctuations. The minimum energy needed for switching is in principle one photon. However, since there is an inherent randomness in the number of photons emitted by a laser or light-emitting diode, a larger mean number of photons must be used to guarantee that the switching action almost always occurs whenever desired. For these light sources and under certain conditions (see Sec. 11.2C) the number of photons arriving within a fixed time interval is a Poisson-distributed random number n with probability distribution $p(n) = \bar{n}^n \exp(-\bar{n})/n!$, where $\bar{n}$ is the mean number of photons. If $\bar{n} = 21$ photons, the probability that no photons are delivered is $p(0) = e^{-21} \approx 10^{-9}$. An average of 21 photons is therefore the minimum number that guarantees delivery of at least one photon, with an average of 1 error every 10^9 trials. The corresponding energy is $E = 21h\nu$. For light of wavelength $\lambda_o = 1\ \mu\text{m}$, $E = 21 \times 1.24 \approx 26$ eV $= 4.2$ aJ. This is regarded as a lower bound on the switching energy; it should be noted, however, that this is a practical bound, rather than a fundamental limit, inasmuch as sub-Poisson light (see Sec. 11.3B) may in principle be used. To be on the less optimistic side, a minimum of 100 photons may be used as a reference. This corresponds to a minimum switching energy of 20 aJ at $\lambda_o = 1\ \mu\text{m}$. Note that, at optical frequencies, $h\nu$ is much greater than the thermal unit of energy $k_B T$ at room temperature (at $T = 300$ K, $k_B T = 0.026$ eV).

Energy–Time Uncertainty. Another fundamental quantum principle is the energy–time uncertainty relation $\sigma_E \sigma_T \geq h/4\pi$ [see (11.1-12)]. The product of the minimum switching energy E and the minimum switching time T must therefore be greater than $h/4\pi$ (i.e., $E \geq h/4\pi T = h\nu/4\pi\nu T$). This bound on energy is smaller than the energy of a photon $h\nu$ by a factor $4\pi\nu T$. Since the switching time T is not smaller than the duration of an optical cycle $1/\nu$, the term $4\pi\nu T$ is always greater than unity. Because E is chosen to be greater than the energy of one photon, $h\nu$, it follows that the energy–time uncertainty condition is always satisfied.

Switching Time. The only fundamental limit on the minimum switching time arises from energy–time uncertainty. In fact, optical pulses of a few femtoseconds (a few optical cycles) are readily generated. Such speeds cannot be attained by semiconductor electronic switches (and are also beyond the present capabilities of Josephson devices). Subpicosecond switching speeds have been demonstrated in a number of optical switching devices. Switching energies can also, in principle, be much smaller than in semiconductor electronics, as Fig. 21.2-5 illustrates.

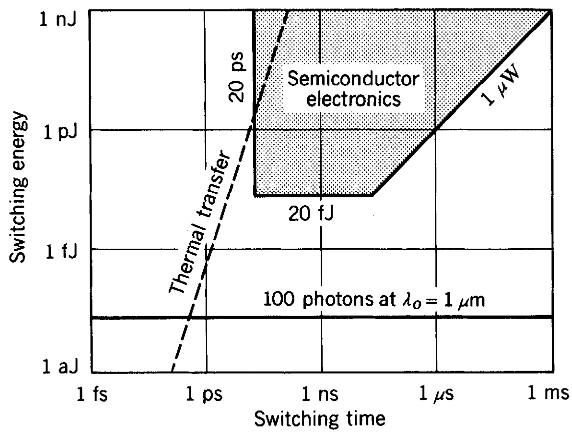


Figure 21.2-5 Limits on the switching energy and time for all-optical switches. Switching energy must be above the 100-photon line. If the switching is repetitive, points must lie to the right of the thermal-transfer line. Limits of semiconductor electronic devices are marked by the 1-μW, 20-fJ, and 20-ps lines.

Size. Limits on the size of photonic switches are governed by diffraction effects, which make it difficult to couple optical power to and from devices with dimensions smaller than a wavelength of light.

Practical Limitations

The primary limitation on all-optical switching is a result of the weakness of the nonlinear effects in currently available materials, which makes the required switching energy rather large. Another important practical limit is related to the difficulty of thermal transfer of the heat generated by the switching process. This limitation is particularly severe when the switching is performed repetitively. If a minimum switching energy E is used in each switching operation, a total energy E/T is used every second. For very short switching times this power can be quite large. The maximum rate at which the dissipated power must be removed sets a limit, making the combination of very short switching times and very high switching energies untenable. The thermal-transfer limit based on certain reasonable assumptions[†] is indicated on the diagram of Fig. 21.2-5. Note, however, that thermal effects are less restrictive if the device is operated at less than the maximum repetition rate; i.e., the energy of one switching operation has more than a bit time to be dissipated. The performance of a number of actual all-optical photonic switches is shown in Fig. 21.3-19 at the end of Sec. 21.3.

21.3 BISTABLE OPTICAL DEVICES

Highly sophisticated digital electronic systems (e.g., a digital computer) contain a large number of interconnected basic units: switches, gates, and memory elements (flip-flops). This section introduces bistable optical devices, which can be used as optical gates and

[†]See P. W. Smith and W. J. Tomlinson, Bistable Optical Devices Promise Subpicosecond Switching, *IEEE Spectrum*, vol. 18, no. 6, pp. 26–33, 1981.

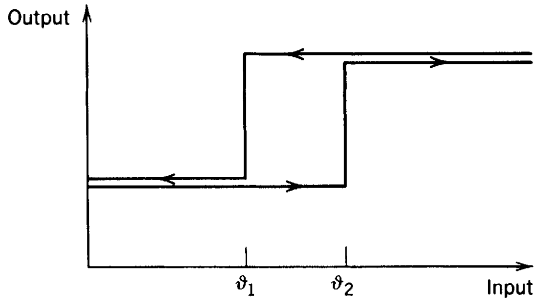


Figure 21.3-1 Input–output relation for a bistable system.

flip-flops. Potential applications in digital optical computing are discussed in Sec. 21.5A.

A. Bistable Systems

A bistable (or two-state) system has an output that can take only one of two distinct stable values, no matter what input is applied. Switching between these values may be achieved by a temporary change of the level of the input. In the system illustrated in Fig. 21.3-1, for example, the output takes its low value for small inputs and its high value for large inputs. When an increasing input exceeds a certain critical value (threshold) ϑ_2, the output jumps from the low to the high value. When the input is subsequently decreased, the output jumps back to the lower value when another critical value $\vartheta_1 < \vartheta_2$ is crossed, so that the input–output relation forms a hysteresis loop.

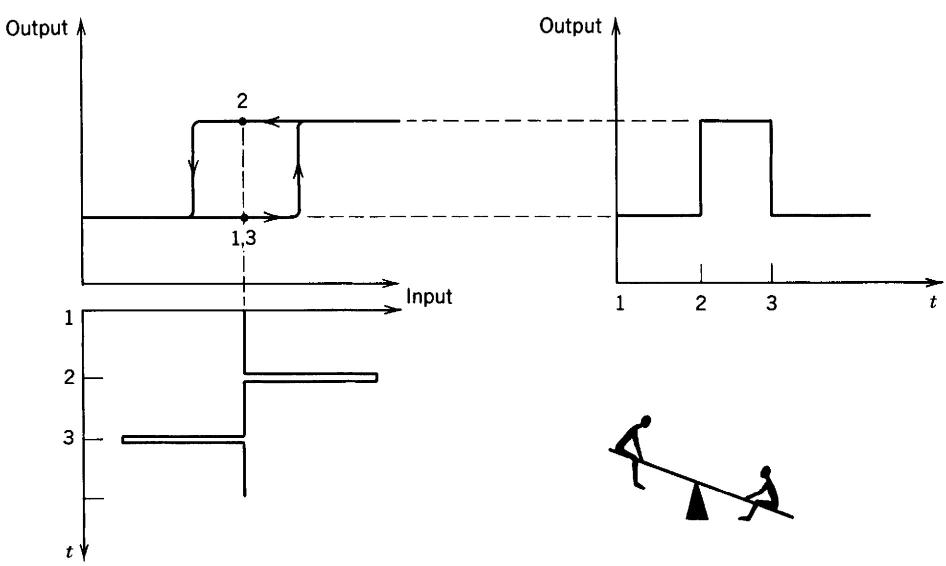


Figure 21.3-2 Flip-flopping of a bistable system. At time 1 the output is low. A positive input pulse at time 2 flips the system from low to high. The output remains in the high state until a negative pulse at time 3 flips it back to the low state. The system acts as a latching switch or a memory element.

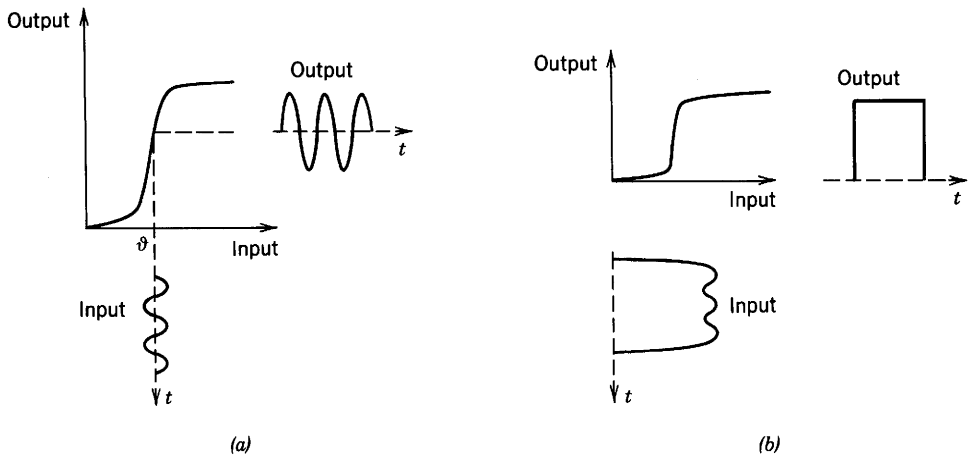


Figure 21.3-3 The bistable device as (*a*) an amplifier or (*b*) a thresholding device, pulse shaper, or limiter.

There is an intermediate range of input values (between ϑ_1 and ϑ_2) for which low or high outputs are possible, depending on the history of the input. Within this range, the system acts like a seesaw. If the output is low, a large positive input spike flips it to high. A large negative input spike flips it back to low. The system has a "flip-flop" behavior; its state depends on its history (whether the last spike was positive or negative; Fig. 21.3-2).

Bistable devices are important in the digital circuits used in communications, signal processing, and computing. They are used as switches, logic gates, and memory elements. The device parameters may be adjusted so that the two critical values (the thresholds ϑ_1 and ϑ_2) coalesce into a single value ϑ. The result is a single-threshold steep S-shaped nonlinear output–input relation. When biased appropriately the device can have large differential gain and can be used as an amplifier, like a transistor. It can

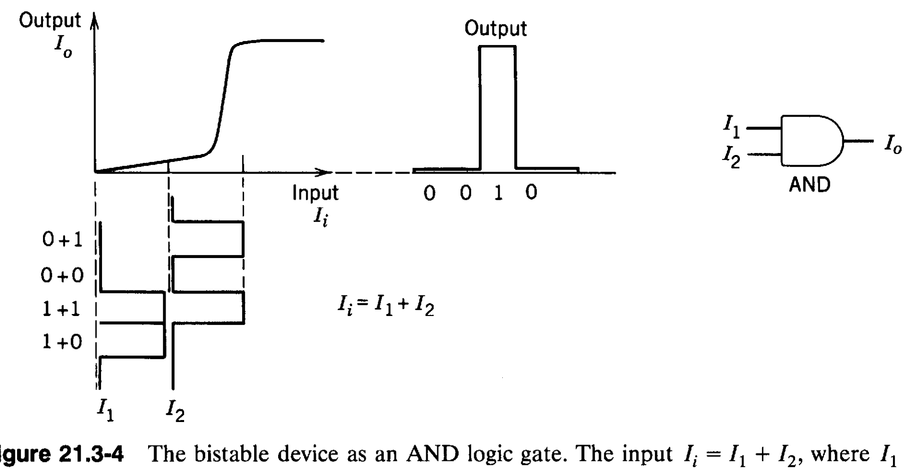


Figure 21.3-4 The bistable device as an AND logic gate. The input $I_i = I_1 + I_2$, where I_1 and I_2 are pulses representing the binary data. The output I_o is high if and only if both inputs are present.

also be used as a thresholding element in which the output switches between two values as the input exceeds a threshold, as a pulse shaper, or as a limiter (Fig. 21.3-3). A stable threshold and stable bias are necessary for these operations.

Bistable devices are also used as logic elements. The binary data are represented by pulses that are added and their sum used as input to the bistable device. With an appropriate choice of the pulse heights in relation to the threshold, the device can be made to switch to high only when both pulses are present, so that it acts as an AND gate, as illustrated in Fig. 21.3-4.

An *electronic* bistable (flip-flop) circuit is made by connecting the output of each of two transistors to the input of the other (see any textbook on digital electronic circuits). As will be explained subsequently, a photonic bistable system, on the other hand, uses a combination of a nonlinear optical material and optical feedback.

B. Principle of Optical Bistability

Two features are required for making a bistable device: *nonlinearity* and *feedback*. Both features are available in optics. If the output of a nonlinear optical element is fed back (by use of mirrors, for example) and used to control the transmission of light through the element itself, bistable behavior can be exhibited.

Consider the generic optical system illustrated in Fig. 21.3-5. By means of feedback the output intensity I_o is somehow made to control the transmittance $\mathcal{T}$ of the system, so that $\mathcal{T}$ is some nonlinear function $\mathcal{T} = \mathcal{T}(I_o)$. Since $I_o = \mathcal{T} I_i$,

$$\boxed{I_i = \frac{I_o}{\mathcal{T}(I_o)}.} \tag{21.3-1}$$

Input–Output Relation for a Bistable System

If $\mathcal{T}(I_o)$ is a nonmonotonic function, such as the bell-shaped function shown in Fig. 21.3-6(*a*), I_i will also be a nonmonotonic function of I_o, as illustrated in Fig. 21.3-6(*b*). Consequently, I_o must be a multivalued function of I_i; i.e., there are some values of I_i with more than one corresponding value of I_o, as illustrated in Fig. 21.3-6(*c*).

The system therefore exhibits bistable behavior. For small inputs ($I_i < \vartheta_1$) or large inputs ($I_i > \vartheta_2$), each input value has a single corresponding output value. In the intermediate range, $\vartheta_1 < I_i < \vartheta_2$, however, each input value corresponds to three possible output values. The upper and lower values are stable, but the intermediate value [the line joining points 1 and 2 in Fig. 21.3-6(*c*)] is unstable. Any slight perturbation added to the input forces the output to either the upper or the lower branch. Starting from small input values and increasing the input, when the threshold ϑ_2 is exceeded the output jumps to the upper state without passing through the

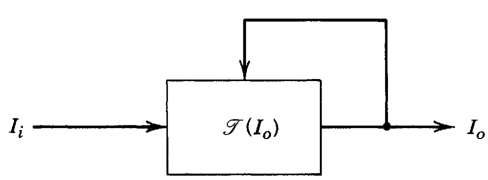


Figure 21.3-5 An optical system whose transmittance $\mathcal{T}$ is a function of its output I_o.

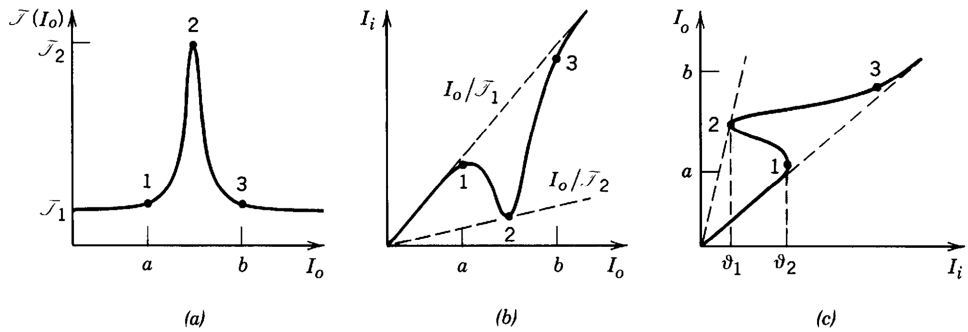


Figure 21.3-6 (a) Transmittance $\mathscr{T}(I_o)$ versus output I_o. (b) Input $I_i = I_o/\mathscr{T}(I_o)$ versus output I_o. For $I_o < a$ or $I_o > b$, $\mathscr{T}(I_o) = \mathscr{T}_1$ and $I_i = I_o/\mathscr{T}_1$ is a linear relation with slope $1/\mathscr{T}_1$. At the intermediate value of I_o for which $\mathscr{T}$ has its maximum value $\mathscr{T}_2$ (point 2), I_i dips below the line $I_i = I_o/\mathscr{T}_1$ and touches the lower line $I_i = I_o/\mathscr{T}_2$ at point 2. (c) The output I_o versus the input I_i is obtained simply by replotting the curve in (b) with the axes exchanged. (The diagram is rotated 90° in a counterclockwise direction and mirror imaged about the vertical axis.)

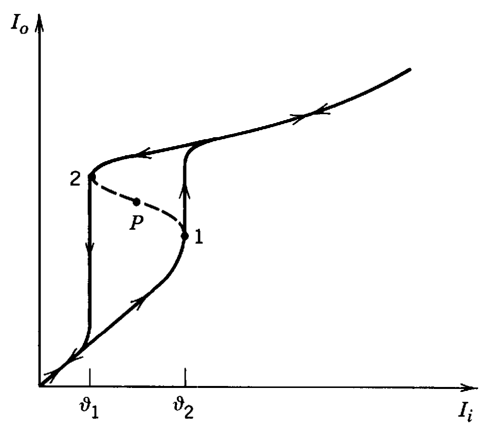


Figure 21.3-7 Output versus input of the bistable device shown in Fig. 21.3-5. The dashed line represents an unstable state.

unstable intermediate state. When the input is subsequently decreased, it follows the upper branch until it reaches ϑ_1 whereupon it jumps to the lower state, as illustrated in Fig. 21.3-7.

The instability of the intermediate state may be seen by considering point P in Fig. 21.3-7. A small increase of the output I_o causes a sharp increase of the transmittance $\mathscr{T}(I_o)$ since the slope of $\mathscr{T}(I_o)$ is positive and large [see Fig. 21.3-6(a) and note that P lies on the line joining points 1 and 2]. This, in turn, results in further increase of $\mathscr{T}(I_o)$, which increases I_o even more. The result is a transition to the upper stable state. Similarly, a small decrease in I_o causes a transition to the lower stable state.

The nonlinear bell-shaped function $\mathscr{T}(I_o)$ was used only for illustration. Many other *nonlinear* functions exhibit bistability (and possibly multistability, with more than two stable values of the output for a single value of the input).

EXERCISE 21.3-1

Examples of Nonlinear Functions Exhibiting Bistability. Use a computer to plot the relation between I_o and $I_i = I_o/\mathscr{T}(I_o)$, for each of the following functions:

(a) $\mathscr{T}(x) = 1/[(x-1)^2 + a^2]$
(b) $\mathscr{T}(x) = 1/[1 + a^2 \sin^2(x+\theta)]$
(c) $\mathscr{T}(x) = \frac{1}{2} + \frac{1}{2}\cos(x+\theta)$
(d) $\mathscr{T}(x) = \operatorname{sinc}^2[(a^2 + x^2)^{1/2}]$
(e) $\mathscr{T}(x) = (x+1)^2/(x+a)^2$.

Select appropriate values for the constants a and θ to generate a bistable relation. The functions in (b) to (e) apply to bistable systems that will be discussed subsequently.

C. Bistable Optical Devices

Numerous schemes can be used for the optical implementation of the foregoing basic principle. Two types of nonlinear optical elements can be used (Fig. 21.3-8):

- Dispersive nonlinear elements, for which the refractive index n is a function of the optical intensity.
- Dissipative nonlinear elements, for which the absorption coefficient α is a function of the optical intensity.

The optical element is placed within an optical system and the output light intensity I_o controls the system's transmittance in accordance with some nonlinear function $\mathscr{T}(I_o)$.

Dispersive Nonlinear Elements

A number of optical systems can be devised whose transmittance $\mathscr{T}$ is a nonmonotonic function of an intensity-dependent refractive index $n = n(I_o)$. Examples are interferometers, such as the Mach–Zehnder and the Fabry–Perot etalon, with a medium exhibiting the optical Kerr effect,

$$n = n_0 + n_2 I_o, \tag{21.3-2}$$

where n_0 and n_2 are constants.

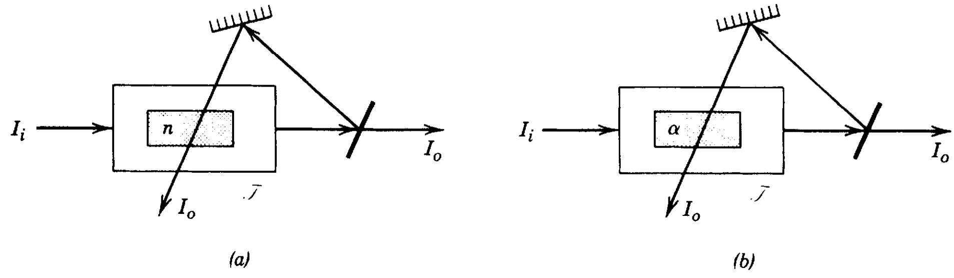


Figure 21.3-8 (*a*) Dispersive bistable optical system. The transmittance $\mathscr{T}$ is a function of the refractive index n, which is controlled by the output intensity I_o. (*b*) Dissipative bistable optical system. The transmittance $\mathscr{T}$ is a function of the absorption coefficient α, which is controlled by the output intensity I_o.

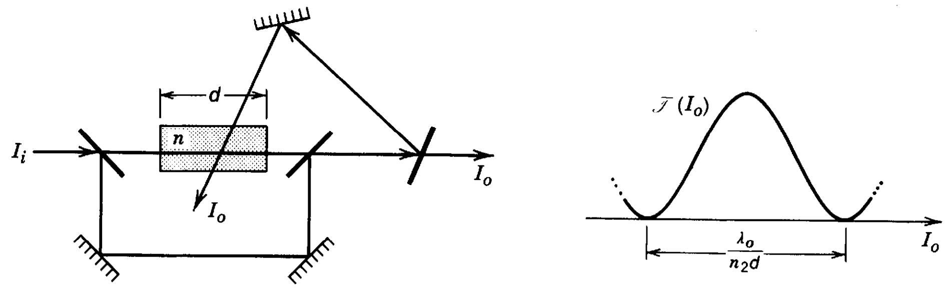


Figure 21.3-9 A Mach–Zehnder interferometer with a nonlinear medium of refractive index n controlled by the transmitted intensity I_o via the optical Kerr effect.

In the *Mach–Zehnder interferometer*, the nonlinear medium is placed in one branch, as illustrated in Fig. 21.3-9. The power transmittance of the system is (see Sec. 2.5A)

$$\mathcal{T} = \frac{1}{2} + \frac{1}{2}\cos\left(2\pi\frac{d}{\lambda_o}n + \varphi_0\right), \tag{21.3-3}$$

where d is the length of the active medium, λ_o the free-space wavelength, and φ_0 a constant. Substituting from (21.3-2), we obtain

$$\boxed{\mathcal{T}(I_o) = \frac{1}{2} + \frac{1}{2}\cos\left(2\pi\frac{d}{\lambda_o}n_2 I_o + \varphi\right),} \tag{21.3-4}$$

where $\varphi = \varphi_0 + (2\pi d/\lambda_o)n_0$ is another constant. As Fig. 21.3-9 shows, this is a nonlinear function comprising a periodic repetition of the generic bell-shaped function used earlier to demonstrate bistability [see Fig. 21.3-6(*a*)].

In a *Fabry–Perot etalon* with mirror separation d, the intensity transmittance is (see Sec. 2.5B)

$$\mathcal{T} = \frac{\mathcal{T}_{\max}}{1 + (2\mathcal{F}/\pi)^2 \sin^2[(2\pi d/\lambda_o)n + \varphi_0]}, \tag{21.3-5}$$

where $\mathcal{T}_{\max}$, $\mathcal{F}$, and φ_0 are constants and λ_o is the free-space wavelength. Substituting for n from (21.3-2) gives

$$\boxed{\mathcal{T}(I_o) = \frac{\mathcal{T}_{\max}}{1 + (2\mathcal{F}/\pi)^2 \sin^2[(2\pi d/\lambda_o)n_2 I_o + \varphi]},} \tag{21.3-6}$$

where φ is another constant. As illustrated in Fig. 21.3-10, this function is a periodic sequence of sharply peaked bell-shaped functions. The system is therefore bistable.

Intrinsic Bistable Optical Devices

The optical feedback required for bistability can be internal instead of external. The system shown in Fig. 21.3-11, for example, uses a resonator with an optically nonlinear medium whose refractive index n is controlled by the internal light intensity I within

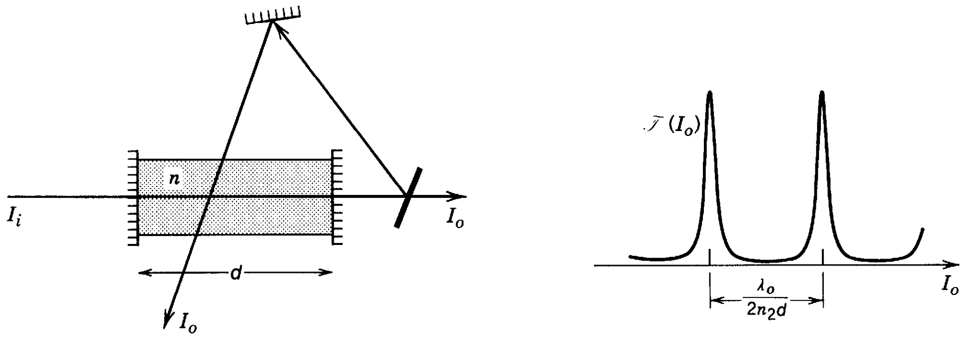


Figure 21.3-10 A Fabry–Perot interferometer containing a medium of refractive index n controlled by the transmitted light intensity I_o.

the resonator, instead of the output light intensity I_o. Since $I_o = \mathcal{T}_o I$, where $\mathcal{T}_o$ is the transmittance of the output mirror, the action of the internal intensity I has the same effect as that of the external intensity I_o, except for a constant factor. If the medium exhibits the optical Kerr effect, for example, the refractive index is a linear function of the optical intensity $n = n_0 + n_2 I$ and the transmittance of the Fabry–Perot etalon is

$$\mathcal{T}(I_o) = \frac{\mathcal{T}_{\max}}{1 + (2\mathcal{F}/\pi)^2 \sin^2[(2\pi d/\lambda_o) n_2 I_o/\mathcal{T}_o + \varphi]}. \tag{21.3-7}$$

Thus the device operates as a self-tuning system.

Dissipative Nonlinear Elements

A dissipative nonlinear material has an absorption coefficient that is dependent on the optical intensity I. The saturable absorber discussed in Sec. 13.3B is an example in which the absorption coefficient is a nonlinear function of I,

$$\alpha = \frac{\alpha_0}{1 + I/I_s}, \tag{21.3-8}$$

where α_0 is the small-signal absorption coefficient and I_s is the saturation intensity. If the absorber is placed inside a Fabry–Perot etalon of length d that is tuned for peak transmission (Fig. 21.3-12), then

$$\mathcal{T} = \frac{\mathcal{T}_1}{(1 - \mathcal{R} e^{-\alpha d})^2}, \tag{21.3-9}$$

where $\mathcal{R} = \sqrt{\mathcal{R}_1 \mathcal{R}_2}$ ($\mathcal{R}_1$ and $\mathcal{R}_2$ are the mirror reflectances) and $\mathcal{T}_1$ is a constant

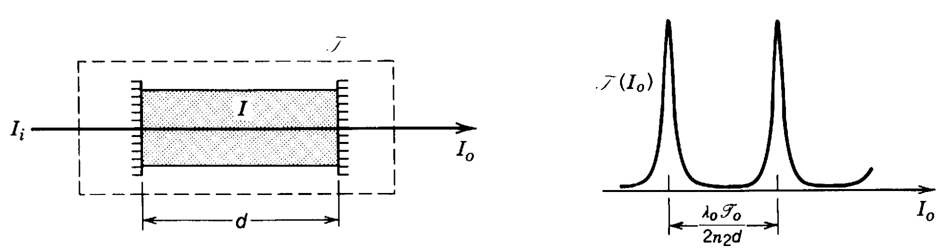


Figure 21.3-11 Intrinsic bistable device. The internal light intensity I controls the active medium and therefore the overall transmittance of the system $\mathcal{T}$.

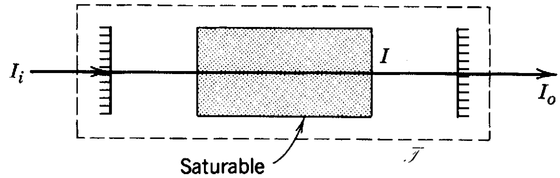


Figure 21.3-12 A bistable device consisting of a saturable absorber in a resonator.

(see Secs. 2.5B and 9.1A for details). If $\alpha d \ll 1$, i.e., the medium is optically thin, $e^{-\alpha d} \approx 1 - \alpha d$, and

$$\mathscr{T} \approx \frac{\mathscr{T}_1}{[1 - (1 - \alpha d)\mathscr{R}]^2}. \tag{21.3-10}$$

Because α is a nonlinear function of I, $\mathscr{T}$ is also a nonlinear function of I. Using the relation $I = I_o/\mathscr{T}_o$ and (21.3-8) and (21.3-10),

$$\mathscr{T}(I_o) = \mathscr{T}_2\left[\frac{I_o + I_{s1}}{I_o + (1 + a)I_{s1}}\right]^2, \tag{21.3-11}$$

where $\mathscr{T}_2 = \mathscr{T}_1/(1 - \mathscr{R})^2$, $a = \alpha_0 d\mathscr{R}/(1 - \mathscr{R})$, and $I_{s1} = I_s\mathscr{T}_o$. For certain values of a, the system is bistable [recall Exercise 21.3-1, example (e)].

Suppose now that the saturable absorber is replaced by an amplifying medium with saturable gain

$$\gamma = \frac{\gamma_0}{1 + I/I_s}. \tag{21.3-12}$$

The system is nothing but an optical amplifier with feedback, i.e., a laser. If $\mathscr{R}\exp(\gamma_0 d) < 1$, the laser is below threshold; but when $\mathscr{R}\exp(\gamma_0 d) > 1$, the system becomes unstable and we have laser oscillation. Lasers do exhibit bistable behavior. However, the theory of these phenomena is beyond the scope of this book.

In some sense, *the dispersive bistable optical system is the nonlinear-index-of-refraction (instead of nonlinear-gain) analog of the laser.*

Materials

Optical bistability has been observed in a number of materials exhibiting the optical Kerr effect (e.g., sodium vapor, carbon disulfide, and nitrobenzene). The coefficient of nonlinearity n_2 for these materials is very small. A long path length d is therefore required, and consequently the response time is large (nanosecond regime). The power requirement for switching is also high.

Semiconductors, such as GaAs, InSb, InAs, and CdS, exhibit a strong optical nonlinearity due to excitonic effects at wavelengths near the bandgap. A bistable device may simply be made of a layer of the semiconductor material with two parallel partially reflecting faces acting as the mirrors of a Fabry–Perot etalon (Fig. 21.3-13). Because of the large nonlinearity, the layer can be thin, allowing for a smaller response time.

GaAs switches based on this effect have been the most successful. Switch-on times of a few picoseconds have been measured, but the switch-off time, which is dominated by relatively slow carrier recombination, is much longer (a few nanoseconds). A switch-off time of 200 ps has been achieved by the use of specially prepared samples in

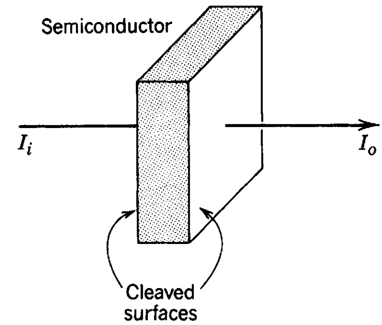


Figure 21.3-13 A thin layer of semiconductor with two parallel reflecting surfaces can serve as a bistable device.

which surface recombination is enhanced. The switching energy is 1 to 10 pJ. It is possible, in principle, to reduce the switching energy to the femtojoule regime. InAs and InSb have longer switch-off times (up to 200 ns). However, they can be speeded up at the expense of an increase of the switching energy. Semiconductor multiquantum-well structures (see Secs. 15.1G and 16.3G) are also being pursued as bistable devices, and so are organic materials.

The key condition for the usefulness of bistable optical devices, as opposed to semiconductor electronics technology, is the capability to make them in large arrays. Arrays of bistable elements can be placed on a single chip with the individual pixels defined by the light beams. Alternatively, reactive ion etching may be used to define the pixels. An array of 100×100 pixels on a 1-cm^2 GaAs chip is possible with existing technology. The main difficulty is heat dissipation. If the switching energy $E = 1$ pJ, and the switching time $T = 100$ ps, then for $N = 10^4$ pixels/cm^2 the heat load is $NE/T = 100$ W/cm^2. This is manageable with good thermal engineering. The device can perform 10^{14} bit operations per second, which is large in comparison with electronic supercomputers (which operate at a rate of about 10^{10} bit operations per second).

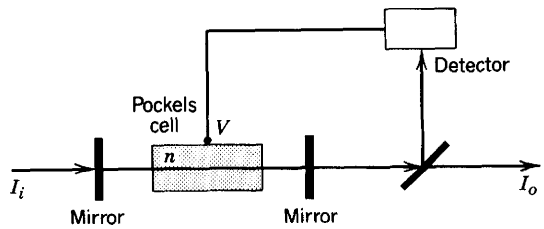
D. Hybrid Bistable Optical Devices

The bistable optical systems discussed so far are all-optical. Hybrid electrical/optical bistable systems in which electrical fields are involved have also been devised. An example is a system using a Pockels cell placed inside a Fabry–Perot etalon (Fig. 21.3-14). The output light is detected using a photodetector, and a voltage proportional to the detected optical intensity is applied to the cell, so that its refractive index variation is proportional to the output intensity. Using $LiNbO_3$ as the electro-optic material, 1-ns switching times have been achieved with $\approx$ 1-μW switching power and $\approx$ 1-fJ switching energy. An integrated optical version of this system [Fig. 21.3-14(*b*)] has also been implemented.

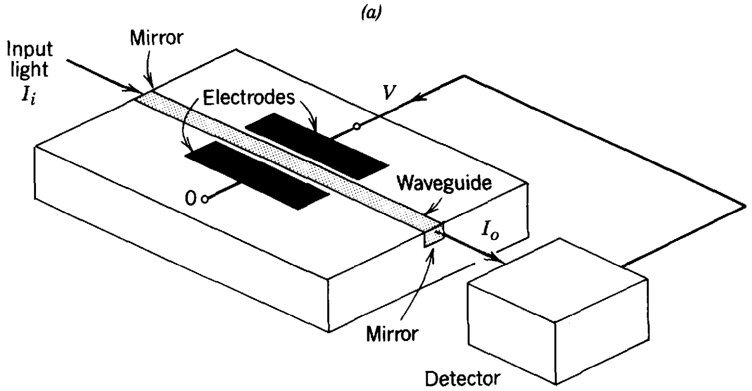
Another system uses an electro-optic modulator employing a Pockels cell wave retarder placed between two crossed polarizers (Fig. 21.3-15); see Sec. 18.1B. Again the output light intensity I_o is detected and a proportional voltage V is applied to the cell. The transmittance of the modulator is a nonlinear function of V, $\mathscr{T} = \sin^2(\Gamma_0/2 - \pi V/2V_\pi)$, where Γ_0 and V_π are constants. Because V is proportional to I_o, $\mathscr{T}(I_o)$ is a nonmonotonic function and the system exhibits bistability.

An integrated-optical directional coupler can also be used (Fig. 21.3-16). The input light I_i enters from one waveguide and the output I_o leaves from the other waveguide; the ratio $\mathscr{T} = I_o/I_i$ is the coupling efficiency (see Sec. 18.1D). Using (18.1-20) yields

$$\mathscr{T} = \left(\frac{\pi}{2}\right)^2 \operatorname{sinc}^2\left\{\frac{1}{2}\left[1 + 3\left(\frac{V}{V_0}\right)^2\right]^{1/2}\right\}, \qquad (21.3\text{-}13)$$



(a)



(b)

Figure 21.3-14 (*a*) A Fabry–Perot interferometer containing an electro-optic medium (Pockels cell). The output optical power is detected and a proportional electric field is applied to the medium to change its refractive index, thereby changing the transmittance of the interferometer. (*b*) An integrated-optical implementation.

where V is the applied voltage and V_0 is a constant. A bistable system is created by making V proportional to the output intensity I_o [see Exercise 21.3-1, example (d)].

Other nonlinear optical devices can also be used. An optically addressed liquid-crystal spatial light modulator (see Sec. 18.3B) can be used to create a large array of bistable elements (Fig. 21.3-17). The reflectance $\mathscr{R}$ of the modulator is proportional to the intensity of light illuminating its "write" side. The output reflected light is fed back

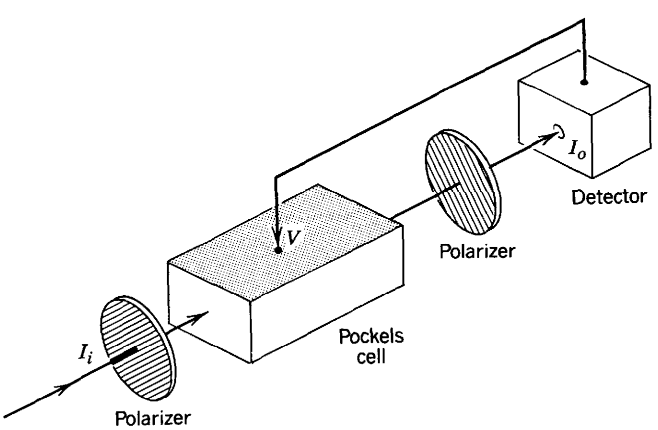


Figure 21.3-15 A hybrid bistable optical system uses an electro-optic modulator with electrical feedback.

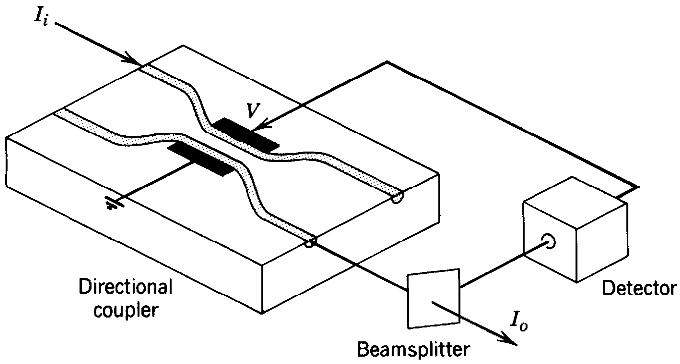


Figure 21.3-16 A bistable device uses a directional coupler with electrical feedback.

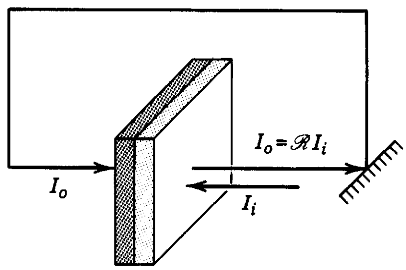


Figure 21.3-17 An optically addressed spatial light modulator operates as an array of bistable optical elements. The reflectance of the "read" side (right) of the valve at each position is a function $\mathscr{R} = \mathscr{R}(I_o)$ of the intensity I_o at the "write" side (left).

to "write" onto the device, so that $\mathscr{R} = \mathscr{R}(I_o)$. Since $\mathscr{R}(I_o)$ is a nonlinear function, bistable behavior is exhibited. Different points on the surface of the device can be addressed separately, so that the modulator serves as an array of bistable optical elements. Typical switching times are in the tens of milliseconds regime and switching powers are less than 1 μW.

The electro-optical properties of semiconductors offer many possibilities for making bistable optical devices. As mentioned earlier, the laser amplifier is an important example in which the nonlinearity is inherent in the saturation of the amplifier gain. InGaAsP laser-diode amplifiers have been operated as bistable switches with optical switching energy less than 1 fJ, and switching time less than 1 ns.

Self-Electro-Optic-Effect Device

Another electro-optic semiconductor device is the **self-electro-optic-effect device** (SEED). The SEED uses a heterostructure multiquantum-well semiconductor material made, for example, of alternating thin layers of GaAs and AlGaAs (Fig. 21.3-18). Because the bandgap of AlGaAs is greater than that of GaAs, quantum potential wells are formed (see Sec. 15.1G) which confine the electrons to the GaAs layers. An electric field is applied to the material using an external voltage source. The absorption coefficient is a nonlinear function $\alpha(V)$ of the voltage V at the wells. But V is dependent on the optical intensity I since the light absorbed by the material creates charge carriers which alter the conductance. Optical bistability is exhibited as a result of the dependence of the absorption $\alpha(V)$ on the internal optical intensity I.

This device operates without a resonator since the feedback is created internally by the optically generated electrons and holes. But it is not exactly an all-optical device since it involves electrical processes within the material and requires an external source of voltage. SEED devices can be fabricated in arrays operating at moderately high speeds and very low energies.

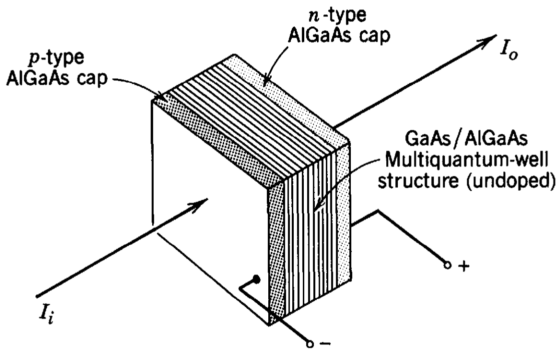


Figure 21.3-18 The self-electro-optic-effect device (SEED).

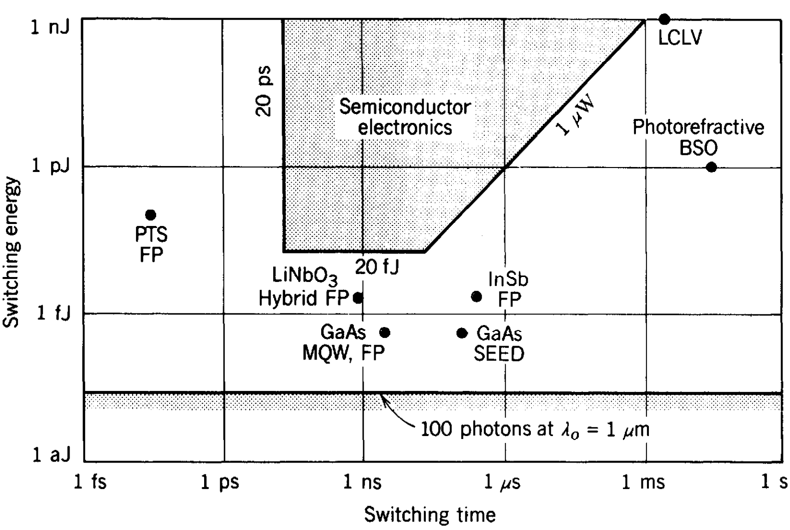


Figure 21.3-19 Switching energies and switching times of a number of optical bistable switches (LCLV = liquid-crystal light valve; FP = Fabry–Perot; SEED = self-electro-optic-effect device; MQW = multiquantum well; PTS = polymerized diacetylene, an organic material; BSO = bismuth silicon oxide). The photon-fluctuation limit on switching energy (100 photons of 1-μm wavelength) is marked. Limits of semiconductor electronic switches are also shown. (Data adapted from P. W. Smith and W. J. Tomlinson, Bistable Optical Devices Promise Subpicosecond Switching, *IEEE Spectrum*, vol. 18, no. 6, pp. 26–33, 1981 © IEEE.)

The performance of a number of bistable optical devices reported in the literature is summarized in Fig. 21.3-19.

21.4 OPTICAL INTERCONNECTIONS

Digital signal-processing and computing systems contain large numbers of interconnected gates, switches, and memory elements. In electronic systems the interconnections are made by use of conducting wires, coaxial cables, or conducting channels within semiconductor integrated circuits. Photonic interconnections may similarly be realized by use of optical waveguides with integrated-optic couplers (see Sec. 7.4B, and

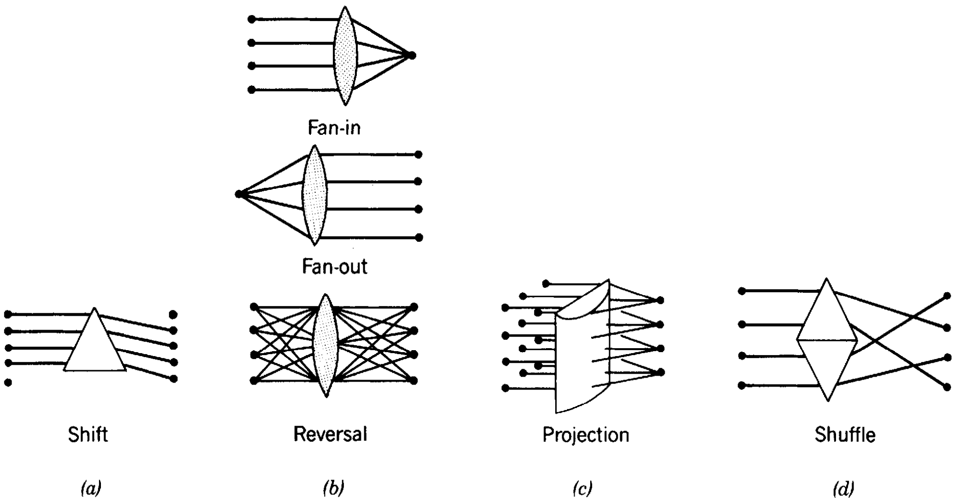


Figure 21.4-1 Examples of simple optical interconnection maps created by conventional optical components: (*a*) A prism bends parallel optical rays and establishes an ordered interconnection map with a shift. (*b*) A lens establishes a fan-in, a fan-out, or a reversal map. (*c*) An astigmatic optical system, such as a cylindrical lens, connects all points of each row in the input plane to a corresponding point in the output plane. (*d*) Two prisms are oriented to perform a perfect-shuffle interconnection map. The perfect shuffle is an operation used in sorting algorithms and in the fast Fourier transform (FFT).

Fig. 21.1-7, for example) or fiber-optic couplers and microlenses (see Sec. 22.2C and Fig. 22.2-12).

Free-space light beams may also be used for interconnections. This option is not available in electronic systems since electron beams must be in vacuum and cannot cross one another without mutual repulsion. This section is devoted to free-space optical interconnects.

Conventional optical components (mirrors, lenses, prisms, etc.) are used in numerous optical systems to establish optical interconnections, such as between points of the object and image planes of an imaging system. To appreciate the order of magnitude of the density of such interconnections, note that in a well-designed imaging system as many as 1000×1000 independent points per mm^2 in the object plane are connected optically by means of the lens to a corresponding 1000×1000 points per mm^2 in the image plane. For this to be implemented electrically, a million nonintersecting and properly insulated conducting channels per mm^2 would be required!

Conventional optical components may be used to create interconnection maps with simple patterns, such as shift, fan-in, fan-out, magnification, reduction, reversal, and shuffle, as Fig. 21.4-1 illustrates. Arbitrary optical interconnection maps, such as that illustrated in Fig. 21.4-2, require the design of custom optical components which may be quite complex and impractical. However, computer-generated holograms made of a

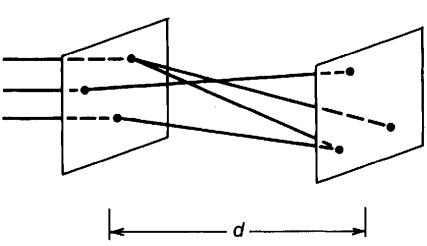


Figure 21.4-2 An arbitrary interconnection map.

large number of segments of phase gratings of different spatial frequencies and orientations have been used successfully to create high-density optical interconnections.

A. Holographic Interconnections

A phase grating is a thin optical element whose complex amplitude transmittance is a periodic function of unit amplitude, $t(x, y) = \exp[-j2\pi(\nu_x x + \nu_y y)]$, for example. The parameters ν_x and ν_y are the spatial frequencies in the x and y directions; they determine the period and orientation of the grating. It was shown in Secs. 2.4B and 4.1A that when a coherent optical beam of wavelength λ is transmitted through the grating, it undergoes a phase shift, causing it to tilt by angles $\sin^{-1}\lambda\nu_x \approx \lambda\nu_x$ and $\sin^{-1}\lambda\nu_y \approx \lambda\nu_y$, where $\lambda\nu_x \ll 1$ and $\lambda\nu_y \ll 1$, as illustrated in Fig. 21.4-3. By varying the spatial frequencies ν_x and ν_y (i.e., the periodicity and orientation of the grating) the tilt angles are altered.

As described in Sec. 4.1A, this principle may be used to make an arbitrary interconnection map by use of a phase grating made of a collection of segments of gratings of different spatial frequencies. Optical beams transmitted through the different segments undergo different tilts, in accordance with the desired interconnection map (Fig. 21.4-4). If the grating segment located at position (x, y) has frequencies

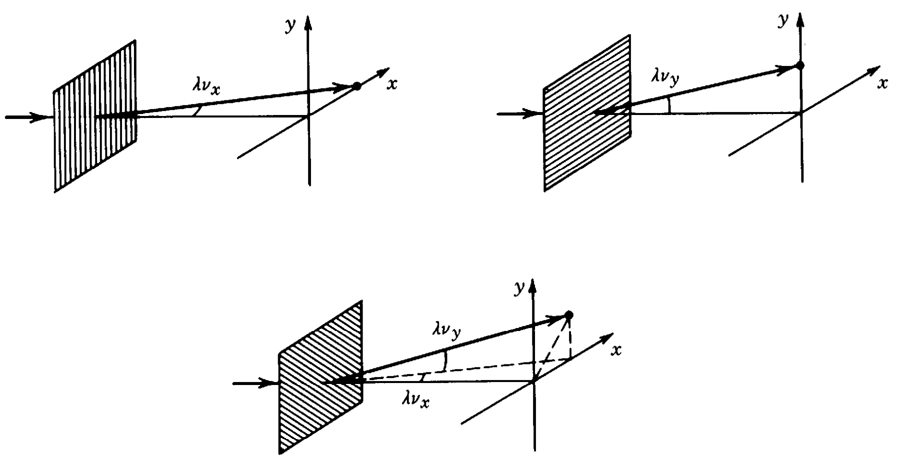


Figure 21.4-3 Bending of an optical wave as a result of transmission through a phase grating. The deflection angles, assumed to be small, depend on the spatial frequency and orientation of the grating.

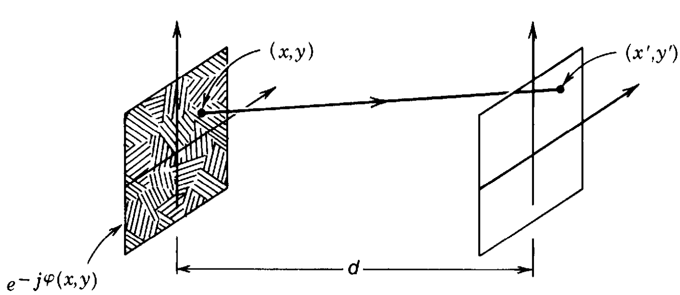


Figure 21.4-4 An interconnection map created by an array of phase gratings of different periodicities and orientations.

$\nu_x = \nu_x(x, y)$ and $\nu_y = \nu_y(x, y)$, the angles of tilt are approximately $\lambda\nu_x$ and $\lambda\nu_y$, and the beam hits the output plane at a point (x', y') satisfying

$$\frac{x' - x}{d} \approx \lambda\nu_x, \qquad \frac{y' - y}{d} \approx \lambda\nu_y, \tag{21.4-1}$$

where d is the distance between the hologram and the output plane and all angles are assumed to be small. Given the desired relation between (x', y') and (x, y), i.e., the interconnection map, the necessary spatial frequencies ν_x and ν_y may be determined at each position using (21.4-1).

In the limit in which the grating elements have infinitesimal areas, we have a continuous (instead of discrete) interconnection map: a geometric coordinate transformation rule that transforms each point (x, y) in the input plane into a corresponding point of the output plane (x', y'). If the desired transformation is defined by the two continuous functions

$$\boxed{x' = \psi_x(x, y), \qquad y' = \psi_y(x, y),} \tag{21.4-2}$$

the grating frequencies must vary continuously with x and y as in a frequency-modulated (FM) signal. Assuming that the grating has a transmittance $t(x, y) = \exp[-j\varphi(x, y)]$, the associated local (or instantaneous) frequencies are given by

$$2\pi\nu_x = \frac{\partial\varphi}{\partial x}, \qquad 2\pi\nu_y = \frac{\partial\varphi}{\partial y}. \tag{21.4-3}$$

(This is analogous to the instantaneous frequency of an FM signal.) Substituting into (21.4-1), we obtain

$$\boxed{\frac{\psi_x(x, y) - x}{d} = \frac{\lambda}{2\pi}\frac{\partial\varphi}{\partial x}, \qquad \frac{\psi_y(x, y) - y}{d} = \frac{\lambda}{2\pi}\frac{\partial\varphi}{\partial y}.} \tag{21.4-4}$$

These two partial differential equations may be solved to determine the grating phase function $\varphi(x, y)$.

EXAMPLE 21.4-1. ***Fan-In Map.*** Suppose that all points (x, y) in the input plane are to be steered to the point $(x', y') = (0, 0)$ in the output plane, so that a fan-in interconnection map is created. Substituting $\psi_x(x, y) = \psi_y(x, y) = 0$ in (21.4-4) and solving the two partial differential equations, we obtain $\varphi(x, y) = -\pi(x^2 + y^2)/\lambda d$. Not surprisingly, this is exactly the phase shift introduced by a lens of focal length d (see Sec. 2.4B).

EXERCISE 21.4-1

The Logarithmic Map. Show that the logarithmic coordinate transformation

$$x' = \psi_x(x, y) = \ln x$$

$$y' = \psi_y(x, y) = \ln y$$

is realized by a hologram with the phase function

$$\varphi(x, y) = \frac{2\pi}{\lambda d}\left(x \ln x - x - \tfrac{1}{2}x^2 + y \ln y - y - \tfrac{1}{2}y^2\right). \qquad (21.4\text{-}5)$$

Holographic interconnection devices are capable of establishing one-to-many or many-to-one interconnections (i.e., connecting one point to many points, or vice versa; Fig. 21.4-5). For example, if the grating centered at the location (x, y) is a superposition of two harmonic gratings so that its complex amplitude transmittance $t(x, y) = \exp[-j2\pi(\nu_{x1}x + \nu_{y1}y)] + \exp[-j2\pi(\nu_{x2}x + \nu_{y2}y)]$, the incident beam is split equally into two components, one tilted at angles $(\lambda\nu_{x1}, \lambda\nu_{y1})$ and the other at $(\lambda\nu_{x2}, \lambda\nu_{y2})$, where all angles are small. Weighted interconnections may be realized by assigning different weights to the different gratings. Arbitrary interconnection maps may therefore be created by appropriate selection of the grating spatial frequencies at each point of the hologram.

EXERCISE 21.4-2

Interconnection Capacity. The space–bandwidth product of a square hologram of size $d \times d$ is the product $N = (Bd)^2$, where B is the highest spatial frequencies that may be printed on the hologram. Show that if the hologram is used to direct each of L incoming beams to M directions, the product ML cannot exceed N,

$$\boxed{ML \le N.}$$

Hint: Use an analysis similar to that presented in Sec. 20.2C in connection with acousto-optic interconnection devices [see (20.2-7)].

What is the maximum number of interconnections per mm^2 if the highest spatial frequency is 1000 lines/mm and if every point in the input plane is connected to every point in the output plane?

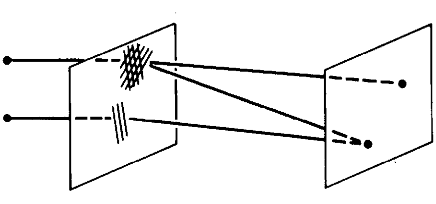


Figure 21.4-5 An arbitrary interconnection system containing one-to-many and many-to-one interconnections.

Once the appropriate phase $\varphi(x, y)$ is decided, the optical element is fabricated by using the techniques of **computer-generated holography**. This approach allows a complex function $\exp[-j\varphi(x, y)]$ to be encoded with the help of a binary function taking only two values, 1 and 0, or 1 and -1, for example. This is similar to encoding an image by use of black dots whose size or density vary in proportionality to the local gray value of the image (an example is the halftone process used for printing images in newspapers). With the help of a computer, the binary image is printed on a mask (a transparency) that plays the role of the hologram. The binary image may also be printed by etching grooves in a substrate, which modulate the phase of an incident coherent wave, a technology known as **surface-relief holography**. References discussing computer-generated holography are provided in the reading list.

Dynamic (reconfigurable) interconnections may be constructed using acousto-optic devices or magneto-optic devices. But the number of interconnection points is much smaller than is achievable by use of holographic gratings. Dynamic holographic interconnections may be achieved by use of nonlinear optical processes, such as four-wave mixing in photorefractive materials. Two waves interfere to create a grating from which a third wave is reflected. The angle between the two waves determines the spatial frequency of the grating, which determines the tilt of the reflected wave (Secs. 18.4 and 19.3C). These devices are the subject of current research.

B. Optical Interconnections in Microelectronics

The possibility of using optical interconnections to replace conventional electrical interconnections in microelectronics has led to a substantial research and development effort. With the successful use of fiber optics for computer-to-computer communications (in local area networks, for example) it is natural to consider the use of optical fibers for processor-to-processor, backplane-to-backplane, board-to-board, and chip-to-chip communications. However, the use of free-space optical communications at these different levels, and as well for intrachip interconnections, has also been explored.

Advances in high-speed high-density microelectronic circuitry and the emergence of parallel processing architectures have created communication bottlenecks so that interconnections have become a major problem. In very-large-scale integrated circuits (VLSI), interconnections occupy a large portion of the available chip area. To minimize the effect of interconnection time delays, which are becoming as long as, or even longer than, gate delays, considerable design effort is being devoted to the equalization of interconnect lengths. Optical interconnections have the potential for alleviating some of these problems.

Optical interconnections offer a number of basic advantages over electronic interconnections:

- *Density*. Electronic interconnections are planar or quasi-planar and cannot overlap or cross without proper insulation. Free-space optical interconnections can be three-dimensional. Optical beams can intersect (pass through one another) without mutual interference (provided that the medium is linear) and their size is limited only by optical diffraction. This allows for a much greater density of interference-free interconnections.
- *Delay*. Photons travel at the speed of light (0.3 mm/ps in free space). The propagation time delay is $\approx$ 3.3 ps/mm. By comparison, propagation delays of electrical signals in striplines fabricated on ceramics and polyimides are approximately 10.2 and 6.8 ps/mm, respectively. Whereas the velocity of light is independent of the number of interconnections branching from an interconnect, in electronic transmission lines the velocity is inversely proportional to the capacitance per unit length so that it depends on the total capacitive "load"; the

propagation delay time therefore increases with increase of the fan-outs. Optics offers a greater flexibility of fan-out and fan-in interconnections, limited only by the available optical power.

- *Bandwidth.* The density of optical interconnections is not affected by the bandwidth of the data carried by each connection. This is not the case in electronics for which the density of interconnections must be reduced sharply at high modulation frequencies to eliminate capacitive and inductive coupling effects between proximate interconnections. Optical interconnections have greater density–bandwidth products than those of electronic interconnections.
- *Power.* Electrical transmission lines must be terminated with their matched impedance to avoid reflections. This usually requires a larger expenditure of power. In optical interconnections, power requirements are limited by the sensitivity of photodetectors and the efficiencies of the electrical-to-optical and optical-to-electrical conversions as well as the power transmission efficiency of the routing elements (which also includes losses due to optical reflections).

Optical interconnections may be implemented within microelectronics by use of a number of electronic-optical transducers (light sources) acting as transmitters that beam the local electric signal to optical-electronic transducers (photodiodes) acting as receivers. A routing device (e.g., a reflection hologram) redirects the emitted light beams to the appropriate photodetector(s), as illustrated in Fig. 21.4-6. This idea can be applied to chip-to-chip or to intrachip interconnections.

There are a number of technical difficulties. Because light sources cannot be made using silicon (see Sec. 15.1D), another technology, GaAs for example, must be used. GaAs-on-Si technology (heteroepitaxy) must surmount the problems of lattice-parameter and thermal-expansion mismatch between the two materials. This is an area of ongoing research. Another difficulty is the design of light sources with sufficiently narrow beams. The design of efficient holograms and the problem of sensitivity to hologram misalignments are important, and must be addressed for this technology to become feasible.

A different approach is to replace the light sources with electro-optic modulators that modulate uniform light beams originating from an external source and reflect them onto a hologram, where they are routed back to the photodiodes on the chip (Fig. 21.4-7).

One-way optical interconnections may be achieved by use of an external light source that transmits information to a number of photodetectors on a silicon chip. One useful

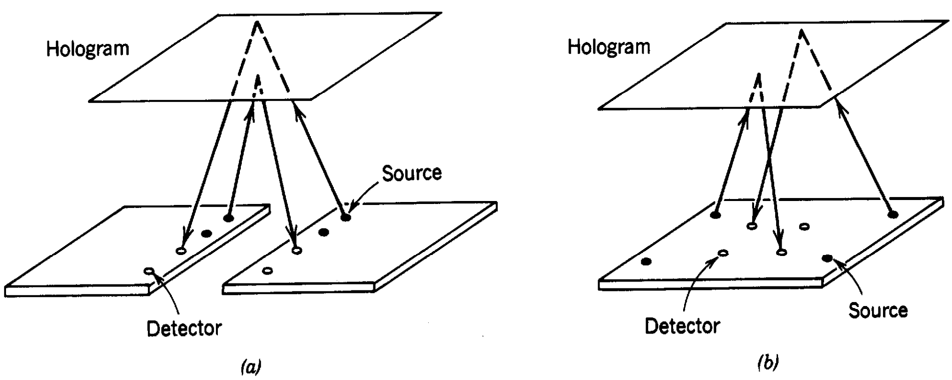


Figure 21.4-6 Optical interconnections using light sources (LEDs or diode lasers) connected optically to photodetectors by an external reflection hologram acting as a routing element: (*a*) chip-to-chip interconnections; (*b*) intrachip interconnections.

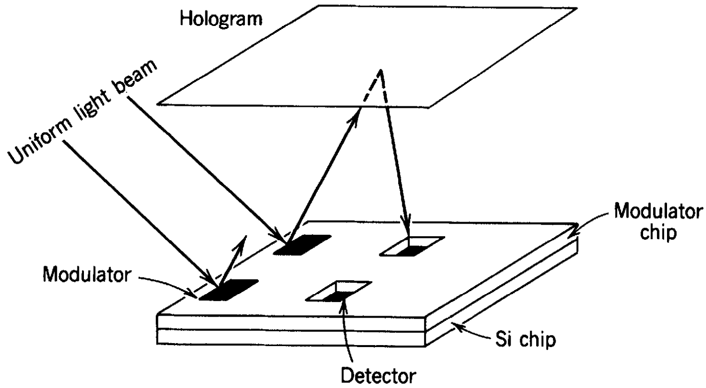


Figure 21.4-7 Interconnections using electro-optic modulators. Electrical signals are used to modulate light beams that are directed by a hologram onto photodiodes, where they are converted into electrical signals.

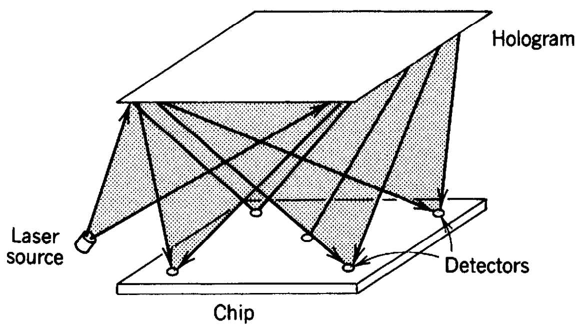


Figure 21.4-8 Clock pulses from an external light source are directed to photodetectors in a silicon chip. This reduces differential time delays and clock skew.

application is in optical clock distribution. This ensures accurate synchronization of high-speed synchronous circuits and alleviates the problem of clock skew that results from differential time delays (Fig. 21.4-8). The hologram may, of course, be eliminated and the light "broadcast" directly to all points on the chip. This creates a robust system that is insensitive to misalignment, but the power efficiency is low since a larger portion of the optical power is wasted.

Reprogrammable interconnections with dynamic holographic optical elements are also under investigation. Optical interconnections are likely to play an important future role in microelectronics.

21.5 OPTICAL COMPUTING

A. Digital Optical Computing

A digital electronic computer is made of a large number of interconnected logic gates, switches, and memory elements. Numbers are represented in a binary system and mathematical operations such as addition, subtraction, and multiplication are reduced to a set of logic operations. Instructions are encoded in the form of sequences of binary numbers. The binary numbers ("0" and "1") are represented physically by two values of the electric potential. The system operation is controlled by a clock that governs the flow of streams of "1" and "0" electrical pulses. Interconnections between the gates and switches are typically local or via a bus and the operation is sequential (i.e., time multiplexed).

It is natural to consider building an optical digital computer mimicking the electronic digital computer. The necessary optical hardware has already been introduced and discussed at length in this and earlier chapters. Electronic gates, switches, and memory elements are replaced by the corresponding optical devices; electrical interconnections within integrated circuits are replaced by waveguides in integrated optics; wires are replaced by optical fibers; the bits "1" and "0" are represented by two intensities of light, "bright" and "dark," for example; data enter the system in the form of light pulses at some clock rate; and the architecture is identical to that of the conventional electronic computer.

Although this straightforward duplication is possible (at least on a small scale), the size, speed, and switching energy and power of present state-of-the-art digital optical devices make the overall performance of the proposed optical computer significantly inferior to its electronic counterpart. As mentioned in Secs. 21.2 and 21.3, very fast optical switches are available, but not in large arrays, and the switching energy and power dissipation are prohibitively large. These limitations, however, are technological instead of fundamental. It is also important to note that the approach of mimicking the electronic computer does not exploit some basic differences between photonics and electronics, which, when properly utilized, could give photonics some important advantages.

Although it is necessary in electronic circuits to guide electrons within conduits (wires, microstrip lines, or planar conducting channels within planar integrated circuits), photons do not require such conduits and free-space three-dimensional global optical interconnections are possible, as described in Sec. 21.4. A large number of points in two parallel planes can be optically connected by a large three-dimensional network of free-space *global* interconnections established by use of a custom-made hologram. It is possible, for example, to have *each* of 10^4 points in the input plane interconnected to *all* 10^4 points of the output plane; or each point of 10^6 points in the input plane connected to an arbitrarily selected set of 100 points among 10^6 points in the output plane. This level of global interconnections is substantially greater than is possible in electronic circuits. A competitive optical computer can, and must, exploit this feature.

Consider, for example, the hypothetical computing system illustrated in Fig. 21.5-1, in which a two-dimensional array of N optical gates ($N = 10^6$, for example) are interconnected holographically. Each gate is connected locally or globally to a small or large number of other gates in accordance with a fixed wiring pattern representing, for example, arithmetic logic units, central processing units, or instruction decoders. The machine could be "programmed" or "reconfigured" by changing the interconnection hologram. This type of parallel architecture is significantly different from the usual bus-limited architecture typically used in very-large-scale integration (VLSI). The level of parallelism (i.e., size of global interconnections) envisioned in optical computers is also much higher than that possible in electronic array processors.

Such a system could, for example, be used to build an optical sequential logic machine. The gates are NOR gates. Each gate has two inputs and one output. The output optical beam from each gate is directed by the hologram to the appropriate input(s) of other gates. The electronic digital circuit is translated into a map of interconnections between output and input points in the gate plane. The interconnection map is coded on a fixed computer-generated hologram. Data arrive in the form of a number of optical beams connected directly to appropriate gate inputs and a number of gates deliver the final output of the processor.

If this parallelism of the processing elements and the interconnections were combined with high switching speeds, the result would yield staggering computational

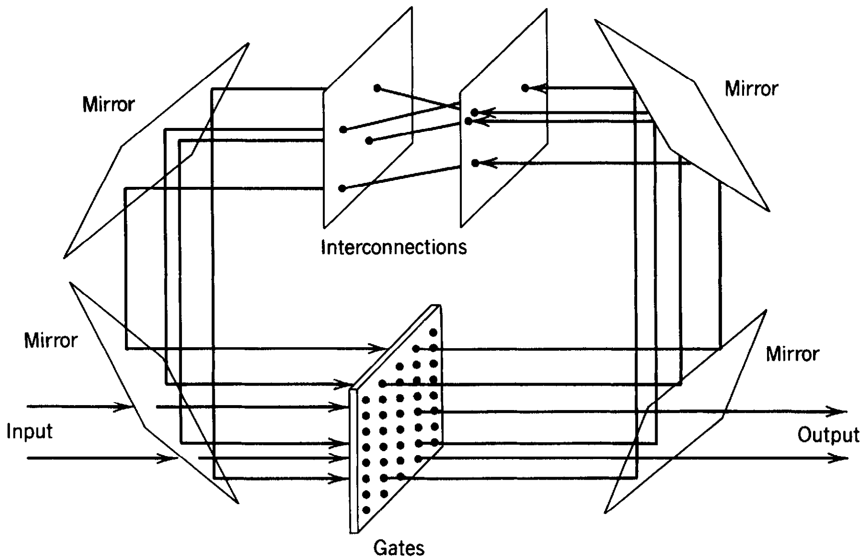


Figure 21.5-1 Possible architecture for all-optical digital computing. N gates are globally interconnected via a hologram.

power. Since the gates are operated in parallel, the data throughput is the product of the number of gates N and their switching speeds. If it were possible to have $N = 10^6$ optical gates operating with a switching time of 0.1 ns, the system could perform 10^{16} bit operations per second. This extremely high rate is approximately the same as that of the human brain and is orders of magnitude greater than the largest currently available electronic computer. As mentioned in Sec. 21.4, these numbers are, in principle, within the *fundamental* limits of photonics. The main technical difficulty remains: *the creation of large high-density arrays of fast optical gates that operate at sufficiently low switching energies and dissipate manageable powers.* Vigorous research toward achieving this goal is ongoing. If these optical machines become available, they are likely to be used in computational tasks suitable to this parallel architecture, e.g., digital image processing and artificial intelligence.

B. Analog Optical Processing

While effective application of the enormous capacity of optical global interconnections to *digital* computing awaits the development of large arrays of optical switches and gates, *analog* optical computing is presently a feasible technology with actual and potential applications in broadband signal processing, radar signal processing, image processing and machine vision, artificial intelligence, and associative memory operations in neural networks.

Most mathematical operations achievable with analog optical processors are combinations of the elementary operations of *addition* and *multiplication* performed many times in parallel by means of a large optical network of interconnections. Theoretically, all *linear* operations (weighted superpositions) can be implemented by use of these elementary operations. The routing elements used to establish the interconnections are usually conventional bulk optical component (lenses, for example), but holographic and acousto-optic devices are being used increasingly.

The variables on which the desired mathematical operation is to be performed are represented by physical (optical) quantities:

- In *incoherent optical processors*, the optical intensity, or the intensity reflectance or transmittance of a transparency or a spatial light modulator, may be used as the computation variable. These variables must be real values and cannot be negative.
- In *coherent optical processors*, the optical complex amplitude, or the complex amplitude transmittance or reflectance of a transparency or a modulator, may be used. These variables can be complex. Coherent optics permits the use of holograms as phase modulators and as interconnection elements.

Multiplication is achieved by transmitting the light through (or reflecting it from) a transparency or a modulator. In coherent processors the optical complex amplitude is multiplied by the amplitude transmittance of the transparency; in incoherent processors it is multiplied by the intensity transmittance. *Addition* is obtained when light beams are routed to the same point. In coherent processors, the complex amplitudes are added; in incoherent processors, the intensities are added.

Optical processors are inherently two-dimensional, so that data can be entered in the form of two-dimensional arrays, or images. This offers a great flexibility of interconnections and a variety of interesting signal-processing schemes. We shall illustrate these schemes by using examples from discrete processors operating on a finite number of variables. Examples of continuous processors operating on functions will then be presented.

Discrete Optical Processors

Summation. The operation $g = \Sigma_{lm} f_{lm}$ is performed by simple use of a fan-in interconnection map (implemented by a lens, for example; Fig. 21.5-2). The input variables f_{lm} $(l, m = 1, 2, \dots, N)$ are represented by the intensities of N^2 optical beams, which are added to produce a light intensity g at the output.

Projection. The operation $g_l = \Sigma_m f_{lm}$ is similarly performed by ordering the input variables f_{lm} in the form of rows and columns in the input plane and using an interconnection map (implemented by a cylindrical lens, for example) that routes the beams of each row into a single point at the output plane where they are added (Fig. 21.5-3).

Inner and Outer Products. The inner product $g = \Sigma_m f_m h_m$ is a transformation of two input vectors f_m and h_m into a scalar g. It is basically a sum of products. The outer

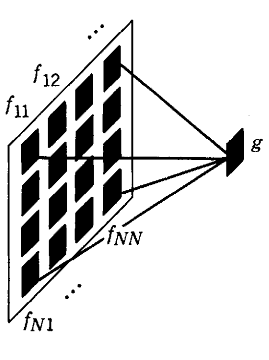


Figure 21.5-2 Optical summation.

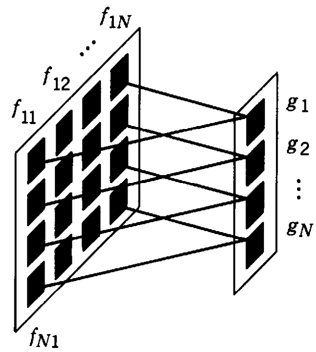


Figure 21.5-3 Optical projection.

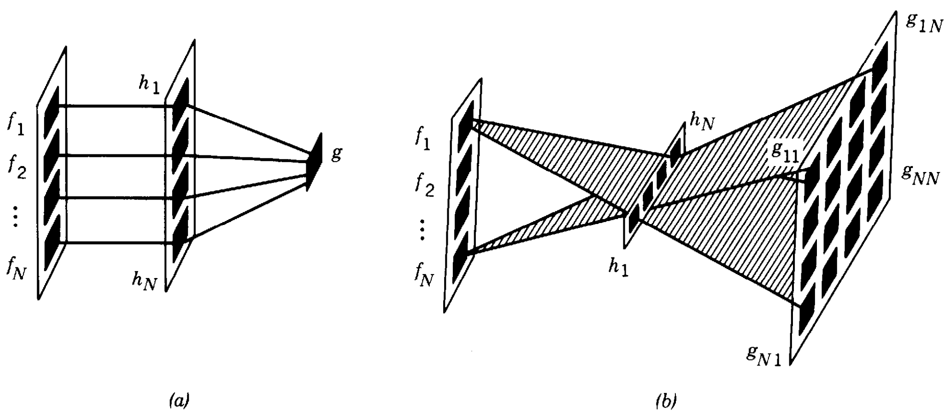


Figure 21.5-4 (a) Inner product. (b) Outer product.

product $g_{lm} = f_l h_m$ transforms two vectors into a matrix. These two operations are performed by use of a combination of a multiplication element and appropriate interconnections, fan-ins and fan-outs, for example (Fig. 21.5-4).

Matrix Multiplication. The operation $g_l = \sum_m A_{lm} f_m$ representing multiplication of a matrix of elements $\{A_{lm}\}$ by a vector of elements $\{f_m\}$ is a basic operation in linear algebra. It can be implemented by using a mask whose transmittances at an array of points are proportional to the elements $\{A_{lm}\}$ (Fig. 21.5-5). The elements are ordered

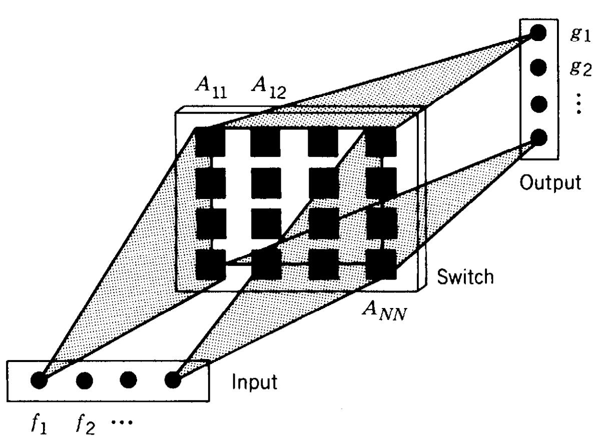


Figure 21.5-5 Optical matrix-vector multiplication.

in the form of a matrix. Two interconnection maps are used: One distributes (fans-out) the entry f_m to all elements of the mth column, where they are multiplied by $A_{1m}, A_{2m}, \ldots, A_{Nm}$; the second adds (fans-in) products in the lth row to obtain $g_l = \sum_m A_{lm} f_m$, for all $l = 1, 2, \ldots, N$. Fan-in and fan-out elements are implemented easily by use of conventional cylindrical lenses.

These five examples illustrate the flexibility of optical processors in performing the operations of linear algebra. Dynamic operations require the use of pulsed light sources. Dynamic transparencies are implemented by use of spatial light modulators.

Continuous Optical Processors

The generalization of these five operations to continuous functions is straightforward. The variables f_{lm}, g_l, and A_{lm} are replaced by the continuous functions $f(x, y)$, $g(x)$, and $A(x, y)$. The operations of integration, projection, inner product, outer product, and matrix-vector multiplication correspond to:

$$g = \int\int f(x, y)\, dx\, dy \qquad \text{(integration)}$$

$$g(x) = \int f(x, y)\, dy \qquad \text{(projection)}$$

$$g = \int f(x)h(x)\, dx \qquad \text{(inner product)}$$

$$g(x, y) = f(x)h(y) \qquad \text{(outer product)}$$

$$g(x) = \int A(x, y)f(y)\, dy \qquad \text{(linear filtering).}$$

The Fourier Transform as an Interconnection Map

The Fourier transform is an important mathematical tool used in the analysis of linear systems and employed in numerous signal processing applications (see Appendices A and B). In Chap. 4 a theory of wave optics based on the Fourier transform was presented. In Sec. 4.2 it was shown that if a transparency of complex amplitude transmittance $f(x, y)$ is illuminated by a plane wave of coherent light, the transmitted light takes the form of plane waves traveling in different directions; the amplitude of the wave that makes an angle (θ_x, θ_y) is $F(\theta_x/\lambda, \theta_y/\lambda)$, where

$$F(\nu_x, \nu_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left[j2\pi(\nu_x x + \nu_y y)\right] f(x, y)\, dx\, dy$$

is the Fourier transform of $f(x, y)$. If these plane waves are focused by a lens of focal length f, the Fourier transform forms an image $F(x/\lambda f, y/\lambda f)$ in the focal plane of the lens, as illustrated in Fig. 21.5-6.

We can also think of a transparency with amplitude transmittance $f(x, y)$ as a holographic interconnection element, like the ones considered in Sec. 21.4A, connecting each point in the output plane to the entire input plane. The function $f(x, y)$ is decomposed into a sum of harmonic functions of different spatial frequencies (ν_x, ν_y) with amplitudes $F(\nu_x, \nu_y)$. As an interconnection element, the transparency "routes" the amplitude $F(\nu_x, \nu_y)$ in a direction at angles $\theta_x \approx \lambda\nu_x$ and $\theta_y \approx \lambda\nu_y$. The natural rules of wave propagation correspond to a Fourier-transform interconnection map! The lens merely funnels all the rays coming from each direction to a single point, i.e., acts as a fan-in interconnection element. The recognition of this natural property of

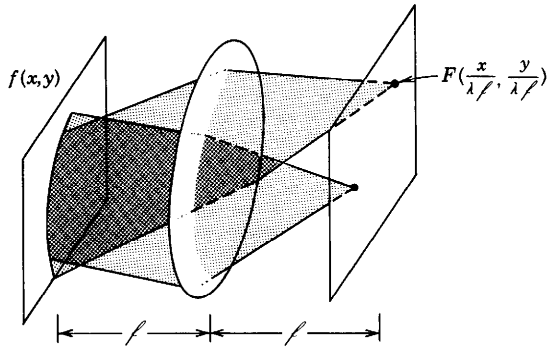


Figure 21.5-6 The optical Fourier transform as an interconnection map.

optical Fourier-transform generation has played an important historic role in motivating the use of optics for signal processing and computing.

Convolution and Correlation

The operation of convolution of two functions $f(x, y)$ and $h(x, y)$,

$$g(x, y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x', y')h(x - x', y - y')\, dx'\, dy'$$

(see Appendix B), represents the action of a spatial filter of impulse-response function $h(x, y)$ on an input function $f(x, y)$. This operation may be implemented by exploiting the property that the Fourier transform of $g(x, y)$ is the product of the Fourier transforms of $f(x, y)$ and $h(x, y)$, i.e., $G(\nu_x, \nu_y) = F(\nu_x, \nu_y)\mathcal{H}(\nu_x, \nu_y)$. Optical implementation involves three steps: Fourier transforming $f(x, y)$ using a lens, multiplication with $\mathcal{H}(\nu_x, \nu_y)$ using an appropriate holographic mask (see Sec. 4.5), and inverse Fourier transforming the product $F(\nu_x, \nu_y)\mathcal{H}(\nu_x, \nu_y)$ using another lens (see Sec. 4.4B for details). Arbitrary two-dimensional shift-invariant spatial filters may thus be implemented optically. Filters of this type have numerous applications in image processing (image enhancement and image deblurring, for example).

The operation of cross-correlation between two functions $h(x, y)$ and $f(x, y)$ is defined by

$$g(x, y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h^*(x', y')f(x' + x, y' + y)\, dx'\, dy'$$

[see (A.1-5)]. This operation may be implemented optically by exploiting the property that the Fourier transforms of $g(x, y)$, $f(x, y)$, and $h(x, y)$ are related by $G(\nu_x, \nu_y) = F(\nu_x, \nu_y)\mathcal{H}^*(\nu_x, \nu_y)$. The optical implementation is similar to that used in convolution. Cross-correlation is an important operation used in pattern recognition as a feature representing the degree of similarity between two images.

The multiplication operation $G = F\mathcal{H}^*$ may be implemented in real time by use of four-wave mixing in a nonlinear medium (see Sec. 19.3C). In accordance with (19.3-21), if the amplitude of waves 1 and 4 are proportional to $\mathcal{H}$ and F, respectively, and the amplitude of wave 3 is uniform, then the amplitude of wave 2 is proportional to the product $F\mathcal{H}^*$. As illustrated in Fig. 21.5-7, the Fourier transforms of $f(x, y)$ and

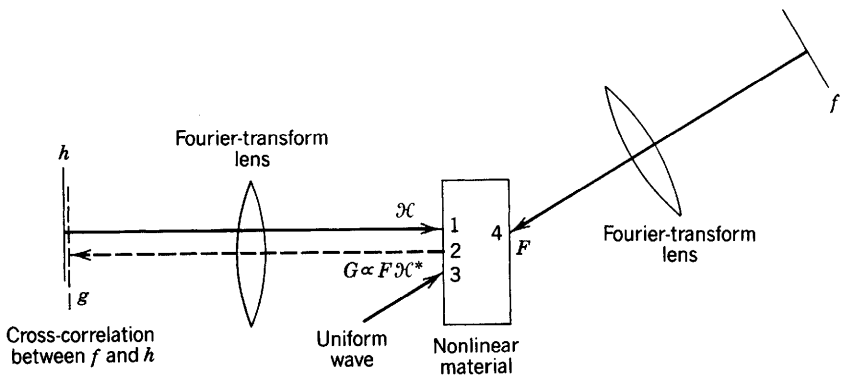


Figure 21.5-7 An optical system for performing the cross correlation between two spatial functions, $f(x, y)$ and $h(x, y)$, using two Fourier-transform lenses and four-wave mixing in a nonlinear optical material.

$h(x, y)$ are computed by use of Fourier transform lenses and the product $F(\nu_x, \nu_y)\mathcal{H}^*(\nu_x, \nu_y)$, which is generated by the mixing process, is inverse Fourier transformed by another Fourier transform lens, so that the cross correlation $g(x, y)$ is obtained in real time. The nonlinear material may be a Kerr medium or a photorefractive material (see Sec. 18.4).

Geometric Transformations

Another class of useful operations on two-dimensional signals (images) consists of geometric transformations. An image $f(x, y)$ is transformed into another image $g(x', y')$ by a change of the coordinate system $x' = \psi_x(x, y)$, $y' = \psi_y(x, y)$. These transformations include magnification, reduction, reversal, rotation, shift, perspective, and so on. The logarithmic transformation $\{x' = \ln x,\ y' = \ln y\}$ is useful in converting a change of scale in the original image into a displacement of the transformed image (because $\ln ax = \ln x + \ln a$). Similarly, the Cartesian-to-polar transformation maps a rotation of the original image into displacement in the transformed image. These operations are useful in scale-invariant and rotation-invariant pattern recognition. The optical implementation of geometric transformations using computer-generated holograms has been described in Sec. 21.4A.

Outlook

A large stock of discrete and continuous mathematical operations on arrays of variables and on two-dimensional functions may be implemented optically. Numerous other operations may be realized by serial and parallel combinations and cascades of these operations. The power of optical analog processors lies in the high degree of parallelism and the large size of the interconnection maps. However, analog computing has limited accuracy and dynamic range and is therefore suitable principally for computational tasks that are insensitive to error. A good example is the implementation of neural networks. These are networks with a high degree of global interconnection, involving simple operations of weighted superpositions and thresholding, that are cascaded and connected in a variety of forms. They implement algorithms which have an underlying redundancy, so that the limited accuracy of analog computing is tolerable.

The main challenge for optical processing lies in the development of high-resolution and fast-interface devices (spatial light modulators and array detectors) and in the design of robust and miniaturized optical systems.

READING LIST

PHOTONIC SWITCHING

Books

T. K. Gustafson and P. W. Smith, eds., *Photonic Switching*, Springer-Verlag, New York, 1988.

G. F. Marchall, ed., *Laser Beam Scanning: Opto-Mechanical Devices, Systems, and Data Storage Optics*, Marcel Dekker, New York, 1985.

H. A. Elion and V. N. Morozov, *Optoelectronic Switching Systems in Telecommunications and Computers*, Marcel Dekker, New York, 1984.

Articles and Special Issues of Journals and Proceedings

Nonlinear Optical Materials and Devices for Photonic Switching, *SPIE*, vol. 1216, 1990.

Y. Silberberg, Photonic Switching Devices, *Optics News*, vol. 15, no. 2, pp. 7–12, 1989.

Photonic Switching, *IEEE Journal of Selected Areas in Communications*, vol. 6, Aug. 1988.

S. F. Su, L. Jou, and J. Lenart, A Review on Classification of Optical Switching Systems, *IEEE Communications Magazine*, vol. 24, no. 5, pp. 50–55, 1986.

P. W. Smith, Applications of All-Optical Switching and Logic, *Philosophical Transactions of the Royal Society of London*, vol. A313, pp. 349–355, 1984.

OPTICAL BISTABILITY

Books

H. M. Gibbs, *Optical Bistability: Controlling Light with Light*, Academic Press, New York, 1985.

C. M. Bowden, M. Cifton, and H. R. Roble, eds., *Optical Bistability*, Plenum Press, New York, 1981.

Articles and Special Issues of Proceedings

High Speed Phenomena in Photonic Materials and Optical Bistability, *SPIE*, vol. 1280, 1990.

B. Chen, Integrated Optical Logic Devices, in *Integrated Optical Circuits and Components*, L. D. Hutcheson, ed., Marcel Dekker, New York, 1987, pp. 289–314.

H. M. Gibbs, Optical Bistability: Where Is It Headed? *Laser Focus*, vol. 21, Oct. 1985.

L. A. Lugiato, Theory of Optical Bistability, in *Progress in Optics*, vol. 21, E. Wolf, Ed., North-Holland, Amsterdam, 1984.

S. D. Smith and A. C. Walker, The Prospects for Optically Bistable Elements in Optical Computing, *SPIE Proceedings*, vol. 492, pp. 342–345, 1984.

P. W. Smith and W. J. Tomlinson, Bistable Optical Devices Promise Subpicosecond Switching, *IEEE Spectrum*, vol. 18, no. 6, pp. 26–33, 1981.

OPTICAL INTERCONNECTIONS

Articles and Special Issues of Journals and Proceedings

Optical Interconnects, *Applied Optics (Information Processing)*, vol. 29, no. 8, 1990.

Optical Interconnections and Networks, *SPIE*, vol. 1281, 1990.

Optical Interconnects in the Computer Environment, *SPIE*, vol. 1178, 1990.

J. E. Midwinter, Digital Optics, Smart Interconnect or Optical Logic?, *Physics in Technology*, vol. 119, part I, pp. 101–108; part II, pp. 153–165, May 1988.

Optical Interconnections, *Optical Engineering*, vol. 25, Oct. 1986.

P. R. Haugen, S. Rychnovsky, A. Husain, and L. D. Hutcheson, Optical Interconnects for High Speed Computing, *Optical Engineering*, vol. 25, pp. 1076–1085, 1986.

A. A. Sawchuk and B. K. Jenkins, Dynamic Optical Interconnections for Parallel Processors, *SPIE Proceedings*, vol. 625, pp. 143–153, 1986.

D. H. Hartman, Digital High Speed Interconnects: A Study of the Optical Alternative, *Optical Engineering*, vol. 25, pp. 1086–1102, 1986.

A. Husain, Optical Interconnect of Digital Integrated Circuits and Systems, *SPIE Proceedings*, vol. 466, pp. 10–20, 1984.

J. W. Goodman, F. I. Leonberger, S. Y. Kung, and R. A. Athale, Optical Interconnections for VLSI Systems, *Proceedings of the IEEE*, vol. 72, pp. 850–866, 1984.

J. W. Goodman, Optical Interconnections in Microelectronics, *SPIE Proceedings*, vol. 456, pp. 72–85, 1984.

COMPUTER-GENERATED HOLOGRAPHY

W. J. Dallas, Computer-Generated Holograms, in *The Computer in Optical Research*, B. R. Frieden, ed., Springer-Verlag, New York, 1980, pp. 291–366.

W.-H. Lee, Computer-Generated Holograms: Techniques and Applications, in *Progress in Optics*, vol. 16, E. Wolf, ed., North-Holland, Amsterdam, 1978, pp. 119–232.

OPTICAL COMPUTING AND PROCESSING

Books

A. D. McAulay, *Optical Computer Architectures*, Wiley, New York, 1991.

P. K. Das, *Optical Signal Processing*, Springer-Verlag, New York, 1990.

R. Arrathoon, ed., *Optical Computing: Digital and Symbolic*, Marcel Dekker, New York, 1989.

H. Arsenault, T. Szoplik, and B. Macukow, eds., *Optical Processing and Computing*, Academic Press, Orlando, FL 1989.

D. G. Feitelson, *Optical Computing*, MIT Press, Cambridge, MA, 1988.

J. L. Horner, ed., *Optical Signal Processing*, Academic Press, New York, 1987.

F. T. S. Yu, *White-Light Optical Signal Processing*, Wiley, New York, 1985.

F. T. S. Yu, *Optical Information Processing*, Wiley, New York, 1983.

H. Stark, ed., *Applications of Optical Fourier Transforms*, Academic Press, New York, 1982.

S. H. Lee, ed., *Optical Information Processing: Fundamentals*, Springer-Verlag, Berlin, 1981.

M. Françon, *Optical Image Formation and Processing*, Academic Press, New York, 1979.

D. Casasent, ed., *Optical Data Processing: Applications*, Springer-Verlag, New York, 1978.

W. E. Kock, G. W. Stroke, and Yu. E. Nesterikhin, *Optical Information Processing*, Plenum Press, New York, 1976.

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

A. R. Shulman, *Optical Data Processing*, Wiley, New York, 1970.

Special Issues of Journals and Proceedings

Digital Optical Computing, *SPIE Critical Reviews*, vol. CR35, 1990.

Advances in Optical Information Processing IV, *SPIE*, vol. 1296, 1990.

Digital Optical Computing II, *SPIE*, vol. 1215, 1990.

Selected Papers on Optical Computing, *SPIE*, vol. 1142, 1989.

Optical Computing '88, *SPIE*, vol. 963, 1989.

Optical Computing and Nonlinear Materials, *SPIE*, vol. 881, 1988.

Optical Computing, *Applied Optics*, vol. 27, May 1988.

Digital Optical Computing, *SPIE*, vol. 752, 1987.

Optical Information Processing II, *SPIE*, vol. 639, 1986.

Optical Computing, *SPIE*, vol. 625, 1986.

Digital Optical Computing, *Optical Engineering*, vol. 25, Jan. 1986.

Photonic Computing, *Applied Optics*, vol. 25, Sept. 15, 1986.
Optical and Hybrid Computing, *SPIE*, vol. 634, 1986.
Real Time Signal Processing VIII, *SPIE*, vol. 564, 1985.
Transformations in Optical Signal Processing, *SPIE*, vol. 373, 1984.
Optical Computing, *Proceedings of the IEEE*, vol. 72, July 1984.
Acoustooptic Signal Processing, *Proceedings of the IEEE*, vol. 69, Jan. 1981.
Optical Computing, *Proceedings of the IEEE*, vol. 65, Jan. 1977.

Articles

D. A. B. Miller, Optoelectronic Applications of Quantum Wells, *Optics and Photonics News*, vol. 1, no. 2, pp. 7–15, 1990.
B. S. Wherrett, The Many Facets of Optical Computing, *Computers in Physics*, vol. 2, pp. 24–27, Mar. 1988.
P. Batacan, Can Physics Make Optics Compute? *Computers in Physics*, vol. 2, pp. 9–15, Mar. 1988.
Y. S. Abumostafa and D. Psaltis, Optical Neural Computers, *Scientific American*, vol. 256, no. 3, pp. 88–95, 1987.
The Coming of the Age of Optical Computing, *Optics News*, Apr. 1986.
D. Casasent, Acoustooptic Linear Algebra Processors: Architectures, Algorithms, and Applications, *Proceedings of the IEEE*, vol. 72, pp. 831–849, 1984.
W. T. Rhodes and P. S. Guilfoyle, Acoustooptic Algebraic Processing Architectures, *Proceedings of the IEEE*, vol. 72, pp. 820–830, 1984.
E. Abraham, C. T. Seaton, and S. D. Smith, The Optical Digital Computer, *Scientific American*, vol. 248, no. 2, pp. 85–93, 1983.
J. Jahns, Concepts of Optical Digital Computing—A Survey, *Optik*, vol. 57, pp. 429–449, 1980.
J. W. Goodman, Operations Achievable with Coherent Optical Information Processing Systems, *Proceedings of the IEEE*, vol. 65, pp. 29–38, 1977.
L. J. Cutrona, E. N. Leith, C. J. Palermo, and L. J. Porcello, Optical Data Processing and Filtering Systems, *IRE Transactions on Information Theory*, vol. IT-6, pp. 386–400, 1960.

PROBLEMS

21.3-1 **Optical Logic.** Figure 21.3-4 illustrates how a nonlinear thresholding optical device may be used to make an AND gate. Show how a similar system may be used to make NAND, OR, and NOR gates. Is it possible to make an XOR (exclusive OR)? Can the same system be used to obtain the OR of N binary inputs?

21.3-2 **Bistable Interferometer.** A crystal exhibiting the optical Kerr effect is placed in one of the arms of a Mach–Zehnder interferometer. The transmitted intensity I_o is fed back and illuminates the crystal. Show that the intensity transmittance of the system is $I_o/I_i = \mathcal{T}(I_o) = \frac{1}{2} + \frac{1}{2}\cos(\pi I_o/I_\pi + \varphi)$, where I_π and φ are constants. Assuming that $\varphi = 0$, sketch I_o versus I_i and derive an expression for the maximum differential gain dI_o/dI_i.

21.4-1 **Interconnection Hologram for a Conformal Map.** Design a hologram to realize the geometric transformation defined by

$$x' = \psi_x(x, y) = \ln\sqrt{x^2 + y^2}$$

$$y' = \psi_y(x, y) = \tan^{-1}\frac{y}{x}.$$

This is a Cartesian-to-polar transformation followed by a logarithmic transformation of the polar coordinate $r = (x^2 + y^2)^{1/2}$. Determine an expression for the phase function $\varphi(x, y)$ of the hologram required.

21.5-1 **Optical Projection.** Design an optical system that implements the optical projection operation depicted in Fig. 21.5-3. Assume that the data $\{f_{lm}\}$ are entered by use of an array of LEDs. Use a spherical lens and a cylindrical lens, of appropriate focal lengths, to perform the necessary imaging in the vertical direction and focusing in the horizontal direction.