"Optical Computing," eds. B. S. Wherrett and F. A. P. Tooley, Proceedings of the 34th Scottish Universities Summer School in Physics, Edinburgh, August 1988 (Adam Hilger, Bristol, 1989) pp 55 - 70

OPTICAL SWITCHING DEVICES: SOME BASIC CONCEPTS

D. A. B. Miller

AT&T Bell Laboratories, Holmdel, NJ 07733, USA

1.

INTRODUCTION

Optical switching devices, by which I mean devices with which light switches light, can take many forms. The many devices proposed can operate on many different principles and can perform many different functions. In this introductory summary, I will attempt to give a qualitative overview of this field. This will not be an exhaustive review, nor will it be a critical comparison of devices. Instead, I will try to explain briefly some of the concepts in the field, both in the physical principles and types of devices and in the requirements that systems impose on devices. I will also emphasize the importance of considering devices in the context of systems. It is important to realize that there is no "best" device; different systems require different device attributes.

Hopefully, in this article I will introduce much of the vocabulary and concepts in the field of optical switching devices. Space will not permit thorough discussions of any. I will start by discussing nonlinear optical processes in general; then I will review some physical attributes of devices, and finally I will summarize some systems requirements on devices.

2.

NONLINEAR OPTICAL PHENOMENA

For this discussion, I will take the broadest possible view on nonlinear optics; nonlinear optical effects are any phenomena that allow one light beam to affect another. Clearly such phenomena must be at the core of any optical switching device. I will treat such concepts as "real" and "virtual" transitions, local and nonlocal nonlinearities, nonlinearities that can be expanded as a power series in the optical field and those that cannot, all of which are relevant to optical switching devices.

What we could call "classical" nonlinear optics covers phenomena such as second harmonic generation and related wave mixing phenomena, and is based on the fact that, in all matter, the displacement of charge (i.e. the polarization) is not in general proportional to the applied field, especially when that field becomes comparable to the internal binding fields of the matter. Thus we can usefully think of this kind of effect in terms of nonlinear springs with charged masses attached; all harmonic generation and many other phenomena can be modelled this way. Important attributes of this kind of effect are as follows. (i) They are usually local, that is, fields at a given point only produce polarizations at that point; this is certainly not true on atomic distances, but is usually effectively true on length scales on the order of the light wavelength. (ii) They are usually fast, because the "resonant frequencies" of the springs are usually in the optical range (since they correspond to optical transitions in the material). (iii) They are weak, because fields of the order of the internal binding fields of matter are required for large effects. (iv) They rely on "virtual" transitions; once the springs are oscillating, the nonlinear process (such as second harmonic generation) can proceed without any dissipation in ideal lossless springs. Energy is however being exchanged, for example from the fundamental to the second harmonic, by being absorbed by the springs and coherently re-emitted. Quantum mechanically this proceeds through so-called "virtual" transitions. It is important to understand that such transitions do result in finite populations that are perfectly real in the general sense of the word, which correspond to the energy stored in the oscillating spring/mass system in our classical mechanical analog. The quantum-mechanical phase of these populations is however still coherent with the light field. Note that, although there is no steady state dissipation in such an ideal "virtual" system with monochromatic fields, there must be absorption of energy from the light field to start the springs oscillating in the first place. It is also in practice very difficult to recover all of that energy back into the light field when we are finished, and some fraction of it may be lost. As an illustrative analogy of this difficulty, if you drive your car, with no shock absorbers, over a bump, then some of your forward kinetic energy is converted into the vertical oscillations of your car. To convert this vertical energy back into forward kinetic energy requires a complementary bump of exactly the right size that is also positioned such that it catches your oscillation at exactly the right phase. (v) Such lossless systems are also intrinsically reversible in the thermodynamic sense (e.g. we can in principle run second harmonic generation backwards to generate the fundamental). This can be a problem when we desire irreversibility in our logic system.

As our optical frequency approaches a transition frequency, such virtual processes are resonantly enhanced, the time that the energy of an individual photon spends in the excited state of the matter becomes longer and longer (it is essentially given by the uncertainty time related to the detuning energy), and the probability that the quantum mechanical phase of the excited state will be interrupted by some random collision before re-emission becomes larger and larger. Eventually, this destroys the virtual transition physics, and we obtain "real" transitions, which can give absorption. It is the randomness of the collisions that makes the transition "real".

A typical real transition effect would be absorption saturation. This has received a lot of attention for optical switching, particularly in semiconductors, although it is more often the nonlinear refraction associated with this nonlinear absorption that is used for the switching. Here typically the absorption saturates because of the Pauli exclusion principle - the excited states become filled. The associated change in refractive index can be calculated from the Kramers-Krönig relations, which inescapably link absorption and refraction. The amount of intensity required for saturation can be low if the lifetime of the excited state is large, and there is generally a reciprocal relation between power and speed in such effects. We can sometimes approximately describe such effects in terms of nonlinear susceptibilities (e.g. the degenerate four wave $\chi^{(3)}(\omega,-\omega,\omega,-\omega)$), in which case we can obtain large numbers, but in doing so we must be aware of two potential problems: (i) because of diffusion of excitation (e.g. carrier diffusion), these effects can be non-local; (ii) the changes in properties are not actually induced directly by the field itself, but by the "degree of excitation" of the material (e.g. the carrier density), which in turn is dependent on the absorbed power. This latter effect has a non-trivial consequence, because it can lead to intrinsic bistability due to increasing absorption; if, for example, increasing carrier density gives increasing absorption (e.g. through band-gap renormalization) that in turn gives increasing carrier density and so on, we can have overall positive feedback and bistable switching. Such a bistable response of the material itself can never be described in terms of a power series in the internal optical field or intensity, because power series are always single-valued and bistability is triple-valued (there are two stable and one unstable values of the output for some range of values of the input). Another example of a real transition nonlinearity is a thermal nonlinearity; it also is non-local because of thermal conduction and in some cases can show bistability from increasing absorption. The "degree of excitation" in the thermal case is temperature.

Note that there is an important distinction between these intrinsic material bistabilities and bistabilities involving external feedback (e.g. Fabry-Perot bistability). In the latter case the whole system is bistable even although the response of the material (e.g. the polarization) is single-valued in the field inside the material, and hence can be described by a power series expansion.

Another example of non-local, non-power-series effects is provided by hybrid devices. Here we could electrically detect light in one place, communicate the resulting voltage to another, and use it through some electro-optic effect to change the optical transmission of another beam. It is also quite straightforward to make bistable systems, for example by incorporating electrical bistability between the detector and the electro-optic output device, or by some hybrid opto-electronic bistability as in simple self-electro-optic-effect devices (SEEDs) or bistable laser diodes, and hence the output of such systems cannot be described by a power series expansion in the field inside the photodetector.

In summary, it can be seen that there is a large variety of types of nonlinear effects that have been considered for use in optical switching devices. These effects have gone well beyond the "classical" nonlinear optics because of the necessity of providing devices operating with sufficiently low intensities, albeit at some compromise in speed of switching, and also partly because of the desire for sophisticated functionality and for behavior such as thermodynamic irreversibility in order to make system design easier.

3.

TYPES OF DEVICES

We can conveniently group many types of devices by means of opposite pairs. I will discuss several such pairs here: two-terminal v. three-terminal; bistable v. non-bistable; active v. passive; intrinsic v. hybrid; refractive v. absorptive. Between them, these categories include nearly all optical switching devices, and this discussion will also serve to introduce and define much of the associated terminology.

The terms "two-terminal" and "three-terminal" are simple to understand in their original electronic context: a two-terminal device, such as a tunnel diode, has only two connections, whereas a three-terminal device, such as a transistor, has three connections. Both of these types of devices can show gain and can switch. The real importance of the distinction lies in two qualitative attributes: (i) two-terminal devices themselves can make no distinction between input and output since both are connected to the same terminal (the other terminal is used for the ground or reference voltage for the signal in the electrical case), hence they have poor input/output isolation in that any fluctuations fed back into the output are liable to be amplified by the gain of the device; (ii) two-terminal devices usually suffer from critical biasing requirements because fluctuations in the biasing (e.g. power supply variations and load variations) also tend to be amplified by the device. Thus, for example, to exhibit signal gain a simple bistable device needs to be biased close to a switching threshold so that a small additional incident power can switch the device, making a large change in output power; this is obviously a critical biasing, and such a device can also be switched by small reflections into the output, showing poor input/output isolation. Three-terminal devices usually have neither of these attributes. In the optical context there is no direct analog of the "terminals" of electronic devices, but we can use a generalized definition in which "three-terminal" devices show input/output isolation and absence of critical biasing whereas "two-terminal" devices do not. In discussing input/output isolation, it is useful to consider the generalized "overlap" of input and output, with those devices with nearly "orthogonal" input and output (i.e. low overlap) having good input/output isolation. This kind of generalization is useful, for example, in discussing "time-sequential" gain as used in the recent symmetric-SEED devices; in this case, the input and output use the same physical connection, but when the device is sensitive to input it has little output, and when it subsequently has a large output, it has little input, so that the input and output have little overlap when the time integration is considered. Incidentally, it takes a three-dimensional graph to display the characteristics of a three-terminal device, whereas those of a two-terminal device can usually be displayed in two dimensions. The vast majority of three-terminal devices also utilize thermodynamic irreversibility to improve input/output isolation. Although three-terminal devices as defined here are obviously desirable for systems (as will be discussed below), it has proven difficult to make optical ones, and the majority of proposed devices are essentially "two-terminal" according to the present definition.

The most commonly investigated class of optical switching device is certainly bistable devices, that is devices that have two stable output states for a range of input powers. I discussed bistability from increasing absorption briefly in the previous section, which is a bistability that is intrinsic to the material. Most of the proposed bistability mechanisms, however, use a combination of a non-bistable optical nonlinearity with some external feedback. Of these, the most investigated system is Fabry-Perot nonlinear refractive bistability. In this system, the nonlinear Fabry-Perot resonator consists of a pair of plane, partially-reflecting mirrors surrounding a material whose index of refraction depends on the optical intensity within it. It is initially tuned slightly off resonance. With increasing input intensity, the intensity inside the resonator increases, thereby changing the refractive index of the cavity, and pulling it towards the resonant condition (in which an integral number of half wavelengths of light would fit inside the cavity). Because the cavity is now more nearly resonant, the intensity is magnified inside the cavity, giving yet further change in refractive index that pulls the system yet closer to resonance, thus further increasing the intensity inside the cavity and so on. This positive feedback mechanism, which results from the combination of the optical nonlinearity of the material with external feedback from the cavity, can become so strong as to switch the device into a highly resonant state. Once in this state, it can be maintained in a highly resonant condition with less incident light because the resonant cavity is magnifying the incident light intensity. Such a bistable system shows both hysteresis, in which the input/output characteristic is different for increasing and decreasing power, and also abrupt switching from one state to another at the switching powers. This bistability can be seen with many different materials and nonlinearities, and also with many different forms of cavity, thus giving a way of exploiting a large variety of materials and structures for optical switching. The resonance also reduces the incident intensity required for switching, therefore enhancing the usefulness of small optical nonlinearities. It has the disadvantages that it is basically still a two-terminal device as defined here and the cavity may have to be fabricated to very high precision, although this fabrication may be possible if it exploits highly-controllable layered growth technology. Many other bistable systems are possible by combining the general principle of nonlinearity combined with external feedback, although those not using resonators generally require higher intensities. Increasing absorption bistability does not normally take advantage of cavities, but can be interesting if used with very large nonlinearities, as in the case of SEEDs. In the case of SEEDs, three-terminal bistability is possible by making a device (the symmetric SEED) that is bistable in the ratio of two beam powers.

There is no fundamental system need for bistable devices, as will be discussed in the next section. Devices showing differential gain, in which small changes in one beam produce larger changes in another, could also be used as the basis for a switching system, and there are many ways of achieving such behavior. Many bistable devices can be adjusted to a regime where they are not quite bistable, but still show a strong "kink" in their input/output characteristics with a slope greater than one (i.e., differential gain), although again these are usually still two-terminal devices. Nonlinear Fabry-Perot resonators can be run with two different wavelengths, with light at one wavelength that is absorbed controlling the optical length of the cavity; hence much larger powers at a second wavelength where the material is transparent can be switched by moving a transmission resonance through this second wavelength. To make this system cascadable (see the next section) requires a complementary device that can be run with the wavelengths interchanged. Laser diode amplifiers can obviously give differential gain, although they can have input/output isolation problems since they amplify just as well in one direction as in the other. Opto-electronic systems including transistors can also obviously be set up to show differential gain without bistability.

A common conceptual difficulty is understanding how "passive" devices, that is devices which are not themselves a source of optical energy, can provide any kind of optical gain. "Active" devices, such as laser diode amplifiers or a device such as a phototransistor driving a light emitting diode, can clearly provide real optical gain since they convert some other source of power (e.g. electricity) into light. The key of course is that optical "passive" devices, just like the transistor in electronics, do only dissipate optical power, but can still induce large changes in power in one beam with small changes in power in another; such passive devices do require some external optical power source to generate the beam that will be modulated, but it may be possible to run many such passive devices off a single light source. (Note incidentally that the use of "active" and "passive" here is different from that in electronics, where "active" devices are components such as transistors, and "passive" refers to components such as resistors and capacitors.)

Another common distinction in switching devices is between "intrinsic" or "all-optical" devices and "hybrid" devices. There is some variability of meaning of these terms. "Intrinsic", when used in the sense of "not hybrid", usually means that there is no electronic transport (i.e. no movement of electrical currents), whereas "hybrid" means essentially opto-electronic in that some combination of optics and electronics is used. It should however be noted that the "classical" nonlinear optics actually involves coherent electrical transport in that the charge "clouds" are displaced by the field, and that this is the source of the nonlinearity itself. Perhaps "hybrid" therefore should be used only for those devices involving dissipative electronic transport. Hybrid devices have a reputation for being inefficient; it is a common opinion, for example, that it is inefficient to convert from optics to electronics and back to optics, and hence that the performance of such a device is fundamentally limited. In fact, the efficiency of hybrid devices depends on the degree to which they are integrated; there is nothing fundamentally very inefficient about photodetectors, laser diodes or some modulators (e.g. quantum well devices). The inefficiency comes about in practice from the energy required for electrical communication over macroscopic distances from one device to another; it must however be stated that the integration of all such components is a technologically very demanding task. We could however look upon "classical" nonlinear optical effects as being in some senses the ultimate integrated opto-electronic device since they are totally efficient in their optical-to-electronic-to-optical conversion, although unfortunately high overall intensities are usually required.

To further confuse the issue, sometimes the term "electronic nonlinearity" is used to emphasize that a nonlinearity works directly with the electronic transitions (e.g. as in absorption saturation of an atomic transition with real populations or as in second harmonic generation with virtual transitions), rather than through, say, some thermal effect. "All-optical" is sometimes used to mean "intrinsic" as defined here, but there is some semantic confusion. In one sense there is no such thing as an all-optical device; every optical interaction involves matter, and every optical interaction that interests us here involves electrons. Another reasonable use of the term is to describe systems in which there is essentially no electrical communication of information outside of the device, so that from the point of view of the system in which the device is embedded, the device is "all-optical".

Devices involving the mixing of waves of substantially different frequencies are possible, but unfortunately the nonlinearities are usually weak because they are usually virtual effects. Consequently, such devices have received less attention because of the severe optical power constraints. Also virtual transition devices can have severe problems with coherent back-coupling because of the inherent reversibility of the processes. The majority of devices can therefore be described as either absorptive or refractive (or both in some cases), since in the absence of wave mixing we can describe the effect of the medium on the beam in terms of the properties of the dielectric constant at the particular frequency of interest. Here we mean absorptive to include gain effects since gain is negative absorption. If we imagine a light beam making a single pass through an absorptive medium of transmission $\exp(-\alpha l)$, then it is clear that in order to make a substantial change (e.g. a factor of e or more) in the transmission of the light through the medium of thickness l (i.e. to "switch" it), we must change the number of absorption lengths (αl) by one or more, i.e., $\Delta\alpha l \geq 1$. The energy required to do this to the material will give us a measure of the minimum switching energy of the device. For a single pass refractive device such as a Mach-Zehnder interferometer, to turn the device from totally on (constructive interference) to totally off (destructive interference) requires a change in optical path of half a wavelength, i.e., $\Delta nl = \lambda/2$, and similarly the minimum switching energy is the energy required to effect this change in the material. Interestingly, the changes in the magnitude of the (complex) dielectric constant ε in each of these two cases are rather similar. (Strictly, the magnitude of the change that in the real part of $\sqrt{\varepsilon}$ would give $\lambda/2$ path change would instead alter the absorption of the material by 2π absorption lengths if made in the imaginary part of $\sqrt{\varepsilon}$.) Incidentally, although it is not immediately obvious, may other kinds of refractive device also require the same $\sim\lambda/2$ path change for switching, even although they are not obviously interferometric; this is generally true, for example, for self-focusing and self-defocusing devices, for devices in which a beam is deflected, and for devices in which a waveguide is changed from guiding to non-guiding.

By the use of multiple passes through the same material, we may reduce the required material volume (or the degree of change in the same volume) proportionately, and this is one of the great advantages of cavities, where the volume required reduces in proportion to the cavity finesse F. Cavities can be used both for absorptive and refractive devices. Note that when a cavity is used, the beam still acquires changes in path length or amplitude of the same magnitude as for the single pass device because it makes essentially F passes through the material. For all devices in which linear absorption of power results in the changes in refractive index required to switch the device (as is common in many proposed devices), we must obviously be able to acquire $\lambda/2$ path length change in less than, say, one absorption length in a single pass device so that the device has some reasonable transmission. From the above argument, we can see that this is still the criterion for the material even in the presence of a cavity. In general, for refractive devices, the linear absorption coefficient must be considered in evaluating the usefulness of nonlinear materials. This is not of course a problem in active devices such as laser diodes operated with gain rather than absorption.

4.

SYSTEMS REQUIREMENTS ON DEVICES

This article is not primarily about systems, but, as mentioned in the introduction, devices cannot usefully be considered without looking at them in the context of systems. Many apparently interesting devices are regretably not usable in large logic systems because they fail some basic requirement such as cascadability or are extremely inconvenient to utilize, and some device that is not apparently interesting because its physical performance is unexceptional compared, say, to electronic devices, might be very useful in a system because it would allow us to use the communications ability of optics to construct a better overall system. I will summarize very briefly below some of the considerations imposed by systems on devices.

Digital logic absolutely requires two attributes aside from the ability to perform a basic logic function: (i) cascadability, so that one gate can drive the next, which requires compatibility of input and output format (e.g. voltage in electronic systems, wavelength in optical systems) (ii) fanout, so that arbitrary logic systems can be constructed, and also to provide the gain necessary to overcome loss in the system. At least some of the devices must have a fanout of 2 or larger (i.e. the ability to drive at least two subsequent devices). Somewhere in the system we must also have the option of performing logical inversion. Any Boolean logic function can be constructed from combinations of two-input NOR (or NAND) gates with a fanout of two.

Also desirable for ease of system design are (i) absence of critical biasing, (ii) input-output isolation, (iii) logic level restoration, (iv) flexibility of functionality. The first two are rather obvious, and have been discussed in the previous section. The others require some more discussion. One of the important features of digital logic is that it is infinitely extensible, i.e. we can construct a logic system with an arbitrary number of levels of gates or logic. One reason for this is that the logic levels are "restored" by each gate, i.e. regardless of the precise input levels, as long as they fall within some defined acceptable range, the output logic levels are always essentially the same. In electronics, this primarily relates to voltage levels. In optics it would relate to power levels and also to beam quality, which is always only degraded by the optics of the system and must also be "restored". "Flexibility of functionality" is a more subtle concept. In designing any large and complex system, we must have complexity, i.e. if a system must perform a complex function, it must have involved a lot of choices to design it, because that complexity must be built into the design. The complexity may exist at a number of levels: in the software, in the physical interconnection and layout of the hardware; and in the devices themselves. It is consequently desirable to have choices at the device level as to the precise function the device will perform. Otherwise we must force all the complexity into the rest of the system, making it more difficult to design. In electronics for example, although we could build an arbitrary logic system using only NOR gates, in practice we will not do this, because the availability of devices with other functions (e.g. flip-flops, AND gates, etc.) allows us to make a simpler system overall. Very few classes of optical devices currently have much functional flexibility. Devices that under different biasing conditions can perform different functions do not necessarily solve this, because they simply transfer the complexity to the design of the biasing system; ideally, we would like to have a choice of devices that, under the *same* biasing conditions could perform different functions.

To make the system physically reasonable, we also require (i) sufficiently low optical operating energy, because of the finite optical power available, (ii) sufficiently low total operating energy to allow thermal dissipation, (iii) compatibility with relevant optical systems (e.g. planar waveguide devices are not readily suitable for two-dimensional optical processing optics since they are only conveniently available in one-dimensional arrays), (iii) sufficiently fast to allow the whole system to perform the desired task better than other methods. All of these properties relate very directly to the basic physics of the nonlinear processes in use. In discussing energies, it is very important that we consider the energy of the device in the system; we must consider not only the energy required to switch the device, but also the energy required to communicate the results to the next device. Here optics may have a considerable advantage, since it is not constrained by the same communications methods as electronics.

The various features described above have different degrees of importance in different kinds of system. Two extreme kinds of systems that illustrate these differences, and which are both of considerable interest, are (a) large, complex, two-dimensionally parallel processors, and (b) small, simple, fast serial processors. In the large complex processor, all the aspects related to convenience of design are particularly important, as is the possibility of two-dimensional operation, and low switching energies may be especially crucial. Very high speed may not be very important, because the speed of a large system may be constrained by the time taken to communicate from one side of it to the other, and also the limit on speed may be more our ability to provide and sink sufficient power to run the devices. As an extreme illustration, to a run a system with a total throughput of 1 Tb/s per chip with a 1 W laser requires less than 1 pJ system energy per bit (the device switching energy would probably have to be lower than this by a factor of 10 - 100 because of system losses and margins). Relatively simple optics could address 1000 - 10,000 such devices, thereby requiring operating speeds of 1 - 10 ns. Even with only such moderately fast absolute speeds, such a system, if it could be built, would have a level of total information on and off the chip that is very difficult to contemplate electronically. Such a system might be useful in processing tasks, such as telecommunications switching or image processing, that have a high total throughput with relatively simple processing. At the other extreme, we could imagine a very high speed serial task such as time-division multiplexing onto and off of an optical fiber. Here the system might be quite simple in that only a few devices might be required. We could therefore use waveguide devices. Because only a few devices are involved, dissipation of the energy is not problematic, and we can afford to work with slightly less convenient conditions because the system is not otherwise complex. In this particular operation, the total time required to perform the multiplexing, from the time the information enters the multiplexer to the time it leaves (i.e. the "latency") is also not important, and hence we may contemplate the use of very large devices, such as nonlinear optical fiber switches, to perform the task; because the latency is unimportant in this system, we may be able to use the weak but very fast virtual transition nonlinearities with a very long interaction length in the fiber. Overall operating power might also not be such a problem because we are only using a few devices, but clearly speed would be very important. It is debatable whether we currently have the optical devices to address either of these example cases, but the point is that different systems have very different device needs.

Finally, the system must be capable of performing the desired task sufficiently cheaply. One important aspect here is that device cost and device complexity are not closely related- except within a given technology. A complex device within one technology may be cheaper than a simple device in another (e.g. a silicon chip compared to a laser diode). In considering cost, again we must not consider the device in isolation from the system; we must also consider, for example, the cost of interconnecting the devices, and this again may turn out to be one of the major advantages of optics. A crucially important aspect of determining cost is "re-use" of existing technologies. Unless and until optics establishes a serious role in switching and processing, it will not be possible to justify developing a completely new technology for such applications, hence technology re-use is vital. It is no accident that there is considerable effort to use semiconductors as the optical materials because their technology is already well developed for other reasons. Similarly, we may try to make use of laser diodes developed for communications or other applications, or optical coating technology, or semiconductor lithographic or layered growth techniques for the fabrication of complex devices.

5.
CONCLUSIONS

In summary, we can see that there is a very large variety of types of nonlinear optical processes, and also a large selection of different devices that have been proposed. At the time of writing, we are just beginning to see the appearance of devices that may be seriously usable in systems. An important conclusion of this discussion is that no one parameter of a device can be used as a reliable guide to its usefulness. A device must be considered in the contexts of the system in which it will be used and the technologies that will be used to make it and the rest of the system.

Many of the advantages of optics come at the system level (e.g. in energy, speed and globality of interconnections), not at the device level, and consequently we should not always be looking simply for physical advantages (e.g. lower energy or higher speed) in the devices themselves. It is however important that we have devices that are good enough to make it worthwhile to contemplate practical systems. Some of the devices that are emerging now are seriously attractive candidates because they address many of the issues discussed above for "useful" digital logic devices. If they are successful, we can expect a strong pull for evolutionary improvement of these devices and we will see increasing technological resources devoted to optical systems in general that will be of much wider benefit both in the applications we can foresee and those we cannot.

BIBLIOGRAPHY

Because of the very large number of topics addressed here, I have not attempted to give a reference list. Instead, I will suggest some publications for further reading.

For a thorough discussion of "classical" nonlinear optics, see
Y. R. Shen, "The Principles of Nonlinear Optics" (Wiley, New York, 1984)

For recent discussions of semiconductor nonlinearities and devices, including quantum wells and self-electro-optic-effect devices, laser diode switching devices and other semiconductor systems, see the many chapters in
H. Haug (ed.), "Optical Nonlinearities and Instabilities in Semiconductors" (Academic, New York, 1988)

For a summary of work in optical bistability, see
H. M. Gibbs, "Optical Bistability: Controlling Light with Light" (Academic, New York, 1985)

For critical discussions of the requirements on logic devices from the perspective of electronics, see
R. W. Keyes, Science **230**, 138-44 (1985); Opt. Acta **32**, 525-35 (1985); Int. J. Theor. Phys. **21**, 263-73 (1982)