Ultra-fast all-optical switching

C. N. Ironside

A review is presented of ultrafast all-optical switching in guided wave formats for broad bandwidth all-optical communication networks. The physics underlying ultrafast optical non-linearities is discussed and the application of these non-linearities to data transmission and all-optical switching is covered.

1. Introduction

Broadly speaking an information system has three fundamental functions that it has to perform; it has to transmit, store and process information. Optical technology already has made substantial inroads into the first two of these functions. It is the purpose of this article to explain how by employing nonlinear optics it is also possible that optics will replace electronics in processing. Probably this will happen first of all in some specialized functions in ultra-high-speed communication systems because this area of optical technology has the most advanced high-speed optoelectronic technology and has the most urgent demand for high-speed processing. At present there appears to be an almost limitless demand for information-handling capacity and in optical communication fibres there is a large available information-handling capacity, or bandwidth. In the communication wavelength window between 1.3 and 1.5 µm there is a bandwidth of 40 THz (40 × 10^{12} Hz) in a single fibre or enough information-handling capacity to allow everyone on the planet to have a telephone conversation simultaneously. However, because of the limitations imposed by both electronics and optoelectronics, only a small fraction of this bandwidth is at present available (the next generation of optical fibre systems are planned to have bit rates of around a maximum of 2.5 Gbits s^{-1} (2.5 × 10^9 bits s^{-1})). The current challenge is to make use of the available bandwidth in the optical fibre. One approach to this is simply to increase the speed at which the fundamental element of processing, the switch, operates. If this is to be achieved, we need a clear view of the physics underlying information technology and why it has evolved into its current state. Information technology makes extensive use of two fundamental particles: the electron and the photon. The electron is a fermion and has charge. The photon is a boson and has no charge. Electrons interact strongly through photons. Photons are non-interacting. These features make the electron suitable for processing information and the photon suitable for the transmission of information. Currently therefore, in information technology, electrons are used in switches, usually in the form of semiconductor devices such as bipolar transistors and field-effect transistors, and photons are used to transmit information. Photons of a wide range of energies are employed spanning radio to optical frequencies which is a range of 10^{11} and each band of frequency has its own associated technology. In this article we are interested in optical frequencies and the associated technology of optoelectronics which converts information in electronic form to optical form and vice versa. However, it would save time and energy if the information could remain in either the electronic or optical form and the functions of processing and transmission were carried out without the need for endlessly converting information from electronic form to photonic form. It is interesting to note that recently there has been an upsurge of interest in high-temperature superconductors because electrons in superconductors adopt some photon-like (or, strictly speaking, boson-like) features and become non-interacting. In nonlinear optics the opposite process is apparent and photons adopt some electron-like features and become interacting.

At this point it is a good idea to ask what is meant by ultra-fast and what we mean by switching. Ultra-fast is a relative term and in this context the comparison is with current state-of-the-art electronics which can demonstrate transistors (the metal semiconductor field-effect transistors (MESFETs) and high-electron-mobility transistors (HEMTs)) designed for microwave applications which have switching characteristics of around 100 GHz or in the time domain 10 ps (10^{-11} s) switching...
times. However, the real difficulty with electronics is that, because electrons are highly interactive, before an electronic signal can reach the gate of a transistor to affect the electrons in the gate, it has to interact with all the electrons close to the gate in the surrounding insulator and to some extent the surrounding universe. This problem is known as capacitance and the so-called stray capacitance is the curse of high-speed electronics. Stray capacitance limits speed and consequently bandwidth, and in effect switching times of around < 500 ps can be considered high speed. Although bearing in mind the limits demonstrated so far in this article we shall consider < 10 ps as ultra-fast. A switching operation is where a signal controls another signal either to turn it on or off or to divert it. There are various categories of logical switching operation such as AND, OR and NAND. An information system, for example a computer or communication system, will generally be composed of a collection of connected switches. The type of connection and the software controlling the system make up the architecture of the system. Gradually ultra-fast all-optical switching at this stage of its development is concerned with simple switches for simple (so-called pipeline) architectures where the data are processed serially.

The development of optical communications has been led by advances in the guiding of optical signals by optical fibres and much of the technology associated with optical communications is determined by the requirements of the optical fibres. The technology of optical guiding in fibre, planar and channel waveguides has seen outstanding advances in the past 20 years. The key feature of the optical guiding is that high irreverances can be maintained over long distances and are not limited by diffraction. This is crucial for ultra-fast all-optical switching. It means that relatively weak nonlinear optical effects can be utilized and devices are not limited by diffraction but are limited by other factors such as absorption and dispersion. All the devices discussed in this review employ fibre or channel waveguide configurations to confine light.

There are other technologies associated with ultra-fast all-optical switching and most closely related are the ultra-fast technologies of ultra-fast optical pulse generation and ultra-fast optical pulse propagation. In the last few years these technologies have formed a 'virtuous triangle' illustrated in figure 1 where advances in one have led to advances in the other, with the underlying theme of nonlinear optics. In this review we shall touch on the other aspects of ultra-fast optical technology as they affect ultra-fast all-optical switching.

From the above introduction the themes of this article become clear. We are concerned with nonlinear optical effects in guided wave formats at wavelengths where optical fibres are most transparent, that is 1.3–1.5 μm and employing high-powered ultra-short (< 10 ps) pulses of light.

The outline of this article is as follows: nonlinear optics is developed from linear optics; the Kramers–Krönig transformation is introduced and applied to nonlinear optics. The physical origin of optical nonlinearities for both second- and third-order effects are treated. The resonant and non-resonant third-order effects are covered and the switching figure of merit is discussed. Ultra-fast all-optical switches employing the non-resonant optical nonlinearity in both fibre and waveguide formats are treated. A promising new method of employing the second-order optical nonlinearity to obtain a nonlinear phase shift, the so-called cascaded second-order effect, is discussed.

2. Nonlinear optics

Nonlinear optics is the study of irradiance-dependent (intensity-dependent) optical properties of materials and it only really became possible with the invention of the laser in 1960 which could produce the high optical irradiances required for what are generally weak effects (for a recent review of nonlinear optics see Butcher and Cotter 1990). There are a whole variety of phenomena which come under this heading and range in time scale from very slow phenomena such as changes in optical properties induced by heating of the material by light (the so-called thermally-induced optical nonlinearities which recover in ~ 1 ms) to optical nonlinearities due to the displacement of valence electrons in the material which respond in times comparable with a cycle of light (10⁻¹⁴ s). Clearly in ultra-fast switching applications we are interested in the fast optical nonlinearities.

3. Linear optics and the Kramers–Krönig transformation

We start with linear optics. The usual way of categorizing nonlinear optical phenomena depends upon extending the concepts introduced by Maxwell in his set of equations which describe the propagation of electromagnetic radiation in a polarizable medium. Maxwell
postulated that the electric field $E$ of the light field induces a polarization $P$ in the medium and that

$$P(\omega) = \chi(\omega)E(\omega).$$  \hspace{1cm} (1)

$\chi(\omega)$ is a frequency-dependent variable material parameter known as the electric susceptibility. In general, $\chi(\omega)$ is a complex number with the real part giving rise to refractive index and the imaginary part giving rise to absorption.

Equation (1) can be treated like a linear systems equation. The material response can be regarded as shown in figure 2 where the electric field $E(\omega)$ is regarded as an input to the material which produces an output polarization $P(\omega)$. Equation (1) can be transformed into the time domain by applying the convolution theorem which states that inverse Fourier transforming the product of two functions in the frequency domain is equal to the convolution of the inverse Fourier transforms of the functions. This can be more neatly expressed as

$$\mathcal{F}\{x(t)\cdot y(t)\} = \mathcal{F}\{x(t)\} \cdot \mathcal{F}\{y(t)\},$$  \hspace{1cm} (2)

where $X(\omega)$ is the Fourier transform of $x(t)$, $Y(\omega)$ is the Fourier transform of $y(t)$, and $\otimes$ is the symbol for convolution.

Applying the convolution theorem to equation (1) and transforming it to the time domain give the following:

$$p(t) = \chi(t) \otimes e(t).$$  \hspace{1cm} (3)

where we have adopted the convention of using capital letters for the frequency domain and lower case letters for the time domain. Rewriting the above equation (3) in full integral form gives

$$p(t) = \int_{-\infty}^{\infty} \chi(\tau)e(t - \tau) \, d\tau,$$  \hspace{1cm} (4)

where $\tau$ is a dummy variable. $\chi(\tau)$ is the impulse response function of the material as it gives the resulting polarization from a delta function input electric field, that is if the $e(t) = \delta(t)$ (the delta function or unit impulse function) then $p(t) = \chi(t)$. Therefore $\chi(t)$ describes the material response to the ultimate ultra-short pulse of light.

The principle of causality states that the effect cannot precede the cause and this can be restated mathematically as

$$\chi(t) = \chi(t)\delta(t),$$  \hspace{1cm} (5)

where $\theta(t)$ is the step function defined as

$$\theta(t) = \begin{cases} 0, & t < 0, \\ 1, & t > 0. \end{cases}$$  \hspace{1cm} (6)

Equation (5) restated in the frequency domain through the convolution theorem becomes

$$\chi(\omega) = \chi(\omega) \otimes \left( \frac{\delta(\omega)}{2} + \frac{i}{2\pi\omega} \right)$$

$$= \frac{\chi(\omega)}{2} + \frac{i}{2\pi} \mathcal{P}\left( \int_{-\infty}^{\infty} \frac{\chi(\Omega)}{\Omega - \omega} \, d\Omega \right)$$

$$= \frac{1}{2\pi} \mathcal{P}\left( \int_{-\infty}^{\infty} \frac{\chi(\Omega)}{\Omega^2 - \omega^2} \, d\Omega \right),$$

where $\mathcal{P}$ refers to the integral over the principal part. This is the Kramers–Krönig relation for the linear susceptibility and thus the Kramers–Krönig relation can simply be derived by a restatement of causality in the frequency domain. The real part of $\chi(\omega)$ is associated with the refractive index and the imaginary part of $\chi(\omega)$ is associated with absorption and these are related through a Kramers–Krönig relationship as follows:

$$\text{Im}(\omega) - 1 = \frac{c}{\pi} \mathcal{P}\left( \int_{0}^{\infty} \frac{\omega(\Omega)}{\Omega^2 - \omega^2} \, d\Omega \right),$$  \hspace{1cm} (8)

which means that, if a complete absorption spectrum of a material is known, then the complete dependence of the refractive index on frequency, that is the dispersion of the refractive index can be calculated. All this is standard linear response theory and can be summarized as shown in figure 3.

4. Nonlinear optics and Kramers–Krönig transformation

When considering nonlinear optical effects, higher-order terms are introduced into equation (1) as follows:

$$p = \chi^{(1)}E + \chi^{(2)}EE + \chi^{(3)}EEE + \ldots,$$  \hspace{1cm} (9)

where the equation has been made more general by renaming $\chi$ as $\chi^{(1)}$ and introducing the higher-order terms $\chi^{(2)}$ and $\chi^{(3)}$, which take account of the field dependence of the electric susceptibility. $\chi^{(2)}$ is known as
the second-order nonlinear susceptibility and \( \chi^{(3)} \) is the third-order nonlinear susceptibility. The nonlinear optical phenomena are usually categorized in terms of second- or third-order nonlinearities (higher than third are not usually considered since the intensities have to be very large, approaching the damage thresholds of the medium, before terms higher than \( \chi^{(3)} \) become significant).

The analysis that was given for linear optics, where we derived the Kramers–Krönig relationship between absorption and refractive index, has recently been extended to nonlinear optics. Hutchings et al. (1992) show how the analysis that we have given above for linear optics can be extended to nonlinear optics. They demonstrated that by the simple conceptual trick of combining the external stimulus, which causes a change in the absorption of the medium, with the medium itself into a linear system which is stimulus plus medium then it is possible to employ the Kramers–Krönig transformation in nonlinear optics. Their main result can be stated as follows:

\[
\Delta n(\omega; \xi) = \frac{c}{\pi} \mathcal{P} \left( \int_0^\infty \frac{\Delta \alpha(\omega'; \xi)}{\omega' \omega' - \omega^2} d\omega' \right),
\]

where \( \Delta n \) is the change in refractive index and \( \Delta \alpha \) is the change in absorption coefficient brought about by a perturbation represented by \( \xi \). For this transformation to be valid and in order to be consistent with causality it is essential that the perturbation be held constant and independent of \( \omega' \). To clarify this, consider the transformation to be operationally equivalent to an experiment where some excitation is applied to the medium at a frequency \( \xi \) and this produces a change in the absorption spectrum at all other frequencies \( \omega \) given by \( \Delta \alpha(\omega; \xi) \); this in principle could be measured by a tuneable light source which is tuned over the full spectrum of the induced absorption change. The consequent change in the refractive index spectrum can be calculated by using the Kramers–Krönig transformation as indicated in equation (10). The point of all the above is that, if we know \( \Delta \alpha(\omega; \xi) \), we can calculate \( \Delta n(\omega; \xi) \). This turns out to be an important result which we can employ to gain considerable insight into nonlinear optics.

5. Physical origins of optical nonlinearities

The equation for the polarization \( P \) links the material properties with Maxwell's equations for the propagation of electromagnetic waves. It does not give any indication as to the physical mechanisms at work which give rise to the optical nonlinearities. In this section we cover the types of physical mechanisms which give rise to the second-order nonlinearity \( \chi^{(2)} \) and the third-order nonlinearity \( \chi^{(3)} \), particularly with regard to semiconductors. Semiconductors have a mature technology and are extensively employed in opto-electronics for both detection and generation of light. They are therefore natural candidate materials for many all-optical ultra-fast switching applications and this review concentrates on the nonlinear optical properties of in particular III–V semiconductors such AlGaaS, GaAs although the physical principles discussed apply to many other types of materials.

The most obvious way of approaching the physical origin of optical nonlinearities and to gain some appreciation of the scales involved is simply to consider that the electric field obtainable from focusing a laser beam of 20 W average power down to a spot size of 20 \( \mu \)m results in an electric field of which is of the order of \( 10^7 \) V m\(^{-1} \) which can significantly perturb the interatomic field in a typical solid. Therefore it is not surprising that there is some dependence of the optical properties of the solid on the optical field applied because the optical field from a modestly powered laser is sufficient to affect the bonds (i.e. move the valence electrons) which bind condensed matter.

6. Second-order optical nonlinearities

The second-order optical nonlinearity has been extensively studied in III–V semiconductors and a theory has been developed which is quite successful in predicting the magnitude of \( \chi^{(2)} \) for semiconductors, particularly those with a crystal structure known as zincblende (an example of this is GaAs). The zincblende structure is closely related to the tetrahedrally bonded diamond structure. The difference is that the zincblende structure consists of two atoms and there is only one atom present in the diamond structure. As we shall see, this difference is crucial to the second-order optical nonlinearity.

The zincblende structure is illustrated in figure 4. The uneven charge distribution along the bond between the
metal and non-metal atoms give rise to the second-order optical nonlinearity in this material. The large optical field applied interacts with the built-in dipole and polarizes the bond, that is the field can redistribute the charge along the bond and thereby alter the optical properties of the material. Another important consideration is that this effect is not balanced by another asymmetric charge distribution in the unit cell of the crystal, or in other words the crystal must lack a centre of symmetry if it is to have a second-order optical nonlinearity. Therefore a zinblende structure such as GaAs exhibits a second-order optical nonlinearity but diamond structures such as Si do not have a second-order optical nonlinearity.

The detailed theory of the second-order nonlinearity in the zinblende materials and how it relates to the linear properties has been worked out with considerable success by several workers (for example Flytzanis and Ducuing (1969) and Harrison (1980). We shall just quote a typical result, for example the expression obtained by Tang (1973) for the second-order susceptibility $\chi^{(2)}_{14}$ (the subscript refers to the fact that actually the second-order susceptibility is a tensor similar to an electro-optic tensor; the convention used is explained in the book by Nye (1979)) from simple molecular-orbital considerations is

$$\chi^{(2)}_{14} = \frac{8R^2}{6e} \left( \frac{E_g}{E_h} \right)^2 \frac{(E_g - E_h)^{1/2}}{E_g} |\chi^{(1)}|^2,$$

where $R$ is the bond length, $E_g$ is the energy gap (note this is not the usual energy gap (Tang 1973)) and $E_h$ is the energy gap of the equivalent homopolar (group IV) semiconductor. The magnitudes of the energy gap parameters $E_g$ and $E_h$ can be determined on the basis of linear optical results.

It turns out that, for most materials, reasonable agreement with experimentally determined values can be obtained with the above theory. The theory works best with sp$^3$ hybrid bonds. If d electrons play a significant role in the bond, as with copper-containing materials, then there is a more significant divergence between theory and experiment.

7. Third-order optical nonlinearities

In this section we cover the types of physical mechanism which give rise to third-order optical nonlinearities in particular again with regard to semiconductors.

As with the linear susceptibility, the imaginary part of the third-order nonlinear susceptibility $\chi^{(3)}$ shows an absorptive nonlinearity which can be described by the equation

$$\alpha = \alpha_0 + \beta_2 I,$$

where $\alpha$ is the total absorption, $\alpha_0$ is the linear absorption coefficient, $\beta_2$ is the coefficient characterizing the nonlinear absorption and $I$ is the irradiance of light. The real part of $\chi^{(3)}$ shows a refractive nonlinearity which can be described by the equation

$$n = n_0 + n_2 I,$$

where $n$ is the total refractive index, $n_0$ is the linear absorption coefficient and $n_2$ is the coefficient characterizing the nonlinear refractive index. The nonlinear absorption and the nonlinear refractive index are related through the Kramers-Krönig transformation given in equation (10). This form of the refractive optical nonlinearity is referred to as the optical Kerr effect after John Kerr who did the original work on the field dependence of the refractive index in the last century.

8. Resonant optical nonlinearities

The term resonant optical nonlinearities refers to optical nonlinearities at wavelengths close to linear absorption. In semiconductors, resonant optical nonlinearities are at wavelengths close to the bandgap absorption edge. Associated with resonant optical nonlinearities are transitions of electrons from the valence band to the conduction band. The electrons have a finite lifetime in the conduction band and then return to the valence band. This leads to a finite recovery time for the nonlinear optical effects. In good-quality semiconductors this time can be of the order of nanoseconds and this leads to switch speeds which are too slow to be considered ultra-fast. However, it is possible to speed up switching
recovery times by introducing other effects such as recombination traps that allow electrons to return to the valence band and electric fields that sweep out carriers from the active part of the switch. Another difficulty with resonant optical nonlinearities is the absorption of light which leads to a low figure of merit for switching and also introduces generally unwanted thermal effects.

8.1. Passive resonant optical nonlinearities

The passive resonant nonlinear optical effects received a great deal of attention at the start of the 1980s because of the large values of the intensity-dependent refractive index that were measured in semiconductors at wavelengths close to the bandgap resonance. A typical mechanism which leads to passive resonant optical nonlinearities is the so-called dynamic Moss–Burstein shift illustrated in figure 5. In this mechanism the absorption close to the band edge is reduced by the electrons at the bottom of the conduction band which are created by the light and block further transitions into the states that they occupy.

In particular, optical bistability employs these resonant optical nonlinearities (see the book by Gibbs (1985) for a review). However, in general, optical bistability uses low powers and recovery times of around a nanosecond which makes optically bistable devices suitable for parallel switching but unsuitable for ultra-fast switching. Some attempts were made to increase the recovery time by incorporating defect sites where electrons can non-radiatively recombine with holes and return to the valence band. These generally encounter other problems associated with increased heating of the material; this is neatly illustrated in figure 6 which summarizes the results of an extensive series of pump-probe measurements of nonlinear refractive index on a material known as semiconductor-doped glass. Semiconductor-doped glass consists of microcrystallites of CdS,Se_{1-x} semiconductor embedded in a glass host. The microcrystallites contain many surface defects which act as recombination centres and speed up the recovery of the optical nonlinearity. The curves in figure 6 show the total change in refractive index induced by a large pump pulse as a function of pump fluence (the fluence is the energy in a pulse divided by area). Each curve refers to a different delay time of probe with respect to pump. The 260-ns curve remains long after the refractive index change induced by electronic effects have faded and the observed refractive index change is entirely due to thermal effects, that is the refractive index changes owing to the rise in temperature associated with the absorption of the pump pulse. The 4-ps curve shows the refractive index changes due to both electronic and thermal effects. If the thermal effects are subtracted by subtracting the 260-ns curve from the 4-ps curve, then we are left with effects due to entirely electronic effects and that is the third curve labelled electronic. The electronic effects clearly saturate. However, the thermal effects do not saturate but are opposite in sign. So the 4-ps curve shows that for some input fluences it is possible for the thermal effect to balance the electronic effect and for there to be no net change in refractive index. The electronic

![Figure 5](image)

**Figure 5.** Moss–Burstein effect, showing the energy against the wave-vector \( k \) for conduction and valence electrons. Electron and hole states of energy \( E_{Fc} \) and \( E_{Fv} \) are separated by \( hf \). Electrons at the bottom of the conduction band are created by the light and block further transitions into these states.

![Figure 6](image)

**Figure 6.** Induced refractive index against pump fluence in semiconductor-doped glass waveguides for two different values of probe delay time. The difference between them is labelled electronic. (After Finlayson et al. 1989)
nonlinearity saturates at values too small to be useful in all-optical switching. Also the absorption in passive resonant optical nonlinearities tends to be large, which means that the throughput of devices is generally small.

8.2. Active resonant optical nonlinearities

Some interesting work has recently been reported (Hultgren and Ippe 1991) on using laser amplifiers as the nonlinear medium and again operating at a wavelength close to the bandgap resonance where the gain in the amplifier can be utilized. It has been shown that there are genuinely ultra-fast optical nonlinearities with a relatively large value of the nonlinear refractive index, although the mechanisms at work are not clear. This work shows a good deal of promise since it seems that it may be possible to have large fast optical nonlinearities with gain rather than loss.

9. Nonresonant optical nonlinearities

Nonresonant nonlinearity refers to the intensity-dependent refractive index and intensity-dependent absorption measured at photon energies well away from the bandgap absorption. In semiconductors this usually means at photon energies well below the bandgap energy. The non-resonant nonlinearity is smaller than the resonant nonlinearity but has a very fast response compared with the resonant nonlinearity. Recent work (Sheik-Bahae et al. 1991) has given considerable insight into the origins of the non-resonant refractive nonlinearity and has shown that it is linked to two-photon absorption. Below we outline the theory of both non-resonant nonlinearity and two-photon absorption.

The nonlinear refractive index can be regarded as related to the real part of \( \chi^{(3)} \) and the two-photon absorption coefficient can be regarded as related to the imaginary part of \( \chi^{(3)} \). This is analogous to linear optics where the refractive index is related to the real part of the linear susceptibility \( \chi \), and the absorption coefficient is related to the imaginary part of \( \chi \). In a similar way to the linear theory it has been shown that the real and imaginary part of \( \chi^{(3)} \) are related by a Kramers–Krönig transformation; that is the nonlinear refractive index can be calculated from a knowledge of the frequency dependence of the two-photon absorption coefficient. It has been shown that this applies to a wide range of semiconductors both II–V and III–V semiconductors and indeed to materials that are insulators such as silica.

In this section this important result is explained in more detail.

It is possible for light with photon energies in the range given by

\[
\frac{1}{2} E_g < h\omega < E_g
\]

(14)

to be absorbed. If the light is sufficiently intense, this can be a significant effect and is known as two-photon absorption. A two-photon process is a process where two photons are either emitted or absorbed per real transition in the material (figure 7). It can be regarded as a process whereby one photon makes a transition to a virtual intermediate state and another photon makes a transition from the intermediate state to a final real state. A virtual state is a state which is very short lived compared with any dephasing lifetime in the material and in accordance with the uncertainty principle has an ill-defined energy.

Two-photon absorption is observed as an intensity-dependent absorption according to the following equation:

\[
\frac{dI}{dz} = -(\alpha_0 I + \beta_2 I^2),
\]

(15)

where \( z \) is the propagation direction, \( \alpha \) the linear absorption coefficient, \( \beta_2 \) the two-photon absorption coefficient and \( I \) the intensity of light. The two-photon absorption coefficient is given by

\[
\beta_2 = \frac{2\hbar\omega_p}{I^2} M_{12},
\]

(16)

where \( \omega_p \) is the operating frequency and \( M_{12} \) is the transition rate of electrons from the valence band to the conduction band via the two-photon absorption process. The transition rate can be worked out from second-order

![Figure 7. Two-photon concepts in a semiconductor. There are two photon- involved per transition of an electron. Both (a) two-photon absorption and (b) two-photon absorption gain are shown. If the photons have the same energy, this is referred to as a degenerate two-photon process.](image)
perturbation theory (Pidgeon et al. 1979) and the following expression for \( \beta_2 \) derived:

\[
\beta_2 = 3.1 \times 10^3 \frac{E_p^{1/2}}{n^2E_g^3} F \left( \frac{2\hbar\omega}{E_g} \right),
\]

(17)

where \( E_p \) is the Kane momentum energy and is nearly constant for a wide range of semiconductors, \( E_g \) is the bandgap energy and \( n \) is the refractive index. The constant \( 3.1 \times 10^3 \) gives the \( \beta_2 \) coefficient in square centimetres per gigawatt if \( E_p \) and \( E_g \) are given in electron–volts. The function \( F(x) \) is determined by the band structure of the semiconductor and has been calculated for a number of generic band structures. For example, \( F(x) \) for semiconductors with parabolic energy bands is given by

\[
F(x) = \frac{(x - 1)^{3/2}}{x^5}
\]

(18)

Equation (17) has been applied to a wide range of semiconductors and has been shown to be accurate within 26\% (van Stryland et al. 1985).

The theory of two-photon absorption is reasonably accurate and indeed, considering the approximations involved, surprisingly accurate. Furthermore it is surprisingly wide ranging; although it was developed for III–V semiconductors, it has also been shown to give good results for II–V semiconductors and for wide-band gap materials normally regarded as insulators.

Equation (17) can be used to predict the two-photon absorption as a function of \( \omega_0 \) and, in a similar manner to linear optical properties, a Kramers–Krönig relationship can be employed to predict the induced refractive index change from the induced absorption coefficient change (Sheik-Bahae et al. 1991):

\[
\Delta n(\omega, \zeta) = \frac{c}{\pi} \int_0^\infty \frac{\Delta \alpha(\omega', \zeta)}{\omega'^2 - \omega^2} \, d\omega',
\]

(19)

where \( \zeta \) is the parameter which causes the change in absorption; for nonlinear effects this is the light intensity. From this relationship the following equation is derived which gives the magnitude and dispersion of \( n_2 \):

\[
n_2(\text{esu}) = \frac{K'G_2(\hbar\omega/E_g)}{n_0E_g^4},
\]

(20)

where \( K' = 6.9 \times 10^{-8} \) and \( E_g \) is in electronvolts. \( G_2 \) is a function approximately given by

\[
G_2(x) = \frac{-2 + 6x - 3x^2 - x^3 - \frac{3}{4}x^4 - \frac{3}{8}x^5 + 2(1 - 2x)^{3/2}\theta(1 - 2x)}{64x^6}
\]

(21)

and \( \theta \) is the step function.

Figure 8. A log–log plot showing the expected \( E_g^{-4} \) dependence of \( n_2 \); (---), indicates the fit to the wide gap solids. (After Sheik-Bahae et al. 1991.)

These equations (20) and (21) have been used to fit the nonlinear refraction of a number of materials. The results are given in figure 8. As can be seen from figure 8, the fit is remarkably good for a wide range of semiconductors and also for other solids such as SiO₂ and LiF₂ generally regarded as insulators. Therefore it would appear that equations (20) and (21) can be used to predict the magnitude and dispersion of \( n_2 \) for a wide range of solids. In particular it should be noted from equation (20) that \( n_2 \) scales as \( E_g^{-4} \) and that this has been verified over a wide range of materials.

In summary, the above shows that the non-resonant refractive nonlinearity and two-photon absorption are closely linked through the Kramers–Krönig relationship for a wide range of solids. The theory of third-order optical nonlinearity in semiconductors has been shown to be reasonably accurate for the dispersion of both the two-photon absorption coefficient \( \beta_2 \) and the refractive part of the nonlinearity \( n_2 \). \( \beta_2 \) relates to the imaginary part of \( \chi^{(3)} \), and \( n_2 \) relates to the real part of \( \chi^{(3)} \). The effect of the electric field on the semiconductor is essentially instantaneous, or at least the semiconductor appears to respond on a time scale shorter than a cycle of light, that is faster than around \( 10^{-14} \) s as far as can be determined by experiment. Another useful feature of the non-resonant nonlinearity is that the linear absorption can be kept low and devices employing the non-resonant nonlinearity can have a large transmission of light.
10. Figures of merit for ultra-fast all-optical switching

If we are comparing materials for ultra-fast all-optical switching, it is convenient to have some figure of merit. The all-optical switches usually rely for their operation on a nonlinear refractive index effect and the general requirement is that the induced refractive index effect should occur before the light is absorbed by either linear or nonlinear absorption. Another consideration is the speed of relaxation of the material which is required to be as short as possible. Taking account of these requirements needs a generic figure $M$ of merit, given as

$$M = \frac{\Delta n}{\alpha \tau}, \quad (22)$$

where $\Delta n$ is the induced refractive index, $\alpha$ is the absorption of either a linear or nonlinear nature and $\tau$ is the recovery time. This figure of merit is generally adopted for a particular switch and type of nonlinearity; for example Delong and Stegeman (1991) consider semiconductors and the effect of two-photon absorption on the non-resonant optical nonlinearity as do Sheik-Bahae et al. (1991). The response time of the non-resonant nonlinearity, which is considered to be comparable with a cycle of light, is no longer an issue and the absorption is considered as dominated by nonlinear two-photon absorption. The figure of merit is a function of the ratio of the photon energy to the bandgap energy and from the theory presented above can be given as

$$M = \frac{\hbar \omega}{E_g} \left| \frac{G_2(\hbar \omega/E_g)}{F(\hbar \omega/E_g)} \right|, \quad (23)$$

which is a dimensionless figure of merit that must be $>1$ for all-optical switching. This figure of merit is plotted in figure 9 which is taken from the work of Sheik-Bahae et al. (1991).

From the point of view of ultra-fast all-optical switching devices (for example the nonlinear directional coupler and the nonlinear interferometers in the following section) the important parts of figure 9 are around $\hbar \omega/E_g = 1$ and $\hbar \omega/E_g = 0.5$ where the figure of merit is large enough that all-optical switching is feasible. However, the theory does not take account of the thermal effects which are apparent in figure 6 and will particularly affect the region where the resonant nonlinearity is in operation, that is around $\hbar \omega/E_g = 1$; these thermal effects will in general lower the value of $M$. Therefore from figure 9 it is predicted that, for ultra-fast all-optical switching, the regime in which to operate is where the photon energy is less than half the bandgap energy. This regime avoids two-photon absorption and so far all ultra-fast all-optical switching devices have operated in this regime.

11. All-optical switching devices

In this section we consider how the optical nonlinearities discussed above can be utilized in optical fibre and optical waveguide format to produce ultra-fast all-optical switching devices.

11.1. Fibre devices

There has been an enormous amount of investment into optical fibres and the losses in optical fibres have been brought down to 0.2 dB km$^{-1}$. Furthermore the losses can now be compensated for with optical fibre amplifiers which consist of optical fibre doped with erbium and pumped by semiconductor lasers. Almost by accident a very good nonlinear medium has been produced because, although the optical nonlinearity in fibres is small (see figure 8), with $n_2$ for fused silica (SiO$_2$) equal to $3.2 \times 10^{-16}$ cm$^2$ W$^{-1}$, there are essentially no losses, and therefore high irradiances can be maintained over a long distance and sufficient nonlinear phase change achieved for all-optical switching. There have been some extremely exciting developments (Islam 1992) in fibre nonlinear optics and ultra-fast all-optical switching with optical fibres. We outline some of them here.

11.1.1. Propagation of ultra-short pulses

An important consideration in all-optical switching is ensuring that the controlling and data ultra-short pulses remain ultra-short even after propagating in what will generally be dispersive media. A difficulty with linear pulse propagation in optical fibres or waveguides is the dispersion of pulses, which is due to both waveguide and material dispersion, that is the effective refractive index is a function of the frequency of the light. The different
Fourier components which make up a pulse of light therefore travel at different speeds and the output of an optical fibre has the Fourier components arriving in a different order compared with the input and this results in pulse distortion.

The pulses used for digital transmission in optical fibres can therefore in general be expected to be distorted. For short distances and low data rates this may not present any difficulty. However, the drive in telecommunication is towards higher data rates and longer distances, and therefore distortion-free transmission at high data rates is required. There are several approaches to achieving this; one of these is to utilize the third-order optical nonlinearity of optical fibres to compensate for the dispersion and to produce soliton pulses which propagate without distortion.

Solitons are nonlinear waves which propagate without changing form (or at least which change form in a manner that is predictably periodic). A physically intuitive interpretation of a soliton can be understood through the three-trucks-on-a-mattress model illustrated in figure 10 where the analogy between the optical nonlinearity and a gravity potential well is made. The intense pulse of light creates its own ‘potential well’ for the various Fourier components of the pulse via the intensity-dependent refractive index. These components are held in the correct order by the optical nonlinearity. In the three-trucks-on-a-mattress model the potential well is equivalent to the nonlinearity \( n_2 \) multiplied by \( U(t) \); the time-dependent shape of the pulse amplitude and the different powered trucks represent the dispersion of the light in the optical fibre. One of the features immediately noticeable from this model is the importance of the relationship between the sign of the dispersion and the sign of nonlinearity. If, for example, \( n_2 \) was negative in silica, then the ‘potential well’ would be a ‘potential hill’ and the nonlinearity would tend to increase the dispersion of the pulse rather than to compensate for it.

It turns out that, in the correct wavelength regime, optical solitons can be created in optical fibres with modest powers easily obtainable from laser sources. The solitons in optical fibres exhibit several properties which make them attractive not only for the transmission of digital data in optical fibres but also for the processing (all-optical switching) of data. They do not strongly interact and they also tend to stay together as a complete entity and are not easily disturbed by switching operations. These properties have led some workers to state that solitons are ‘natural bits’.

11.2. Nonlinear loop mirror

Although optical fibre has a small optical nonlinearity, it does have a small absorption and therefore long interaction lengths are possible. An all-optical switch called the nonlinear loop mirror switch employing optical fibre has recently been demonstrated (Blow and Doran 1990). It utilizes the long interaction lengths available in optical fibres to build up the required nonlinear phase shift from the small optical nonlinearity present in optical communication fibres.

11.2.1. Loop mirrors

Figure 11 illustrates a Mach–Zehnder interferometer concept for an optical fibre. The nonlinear loop mirror is based on a nonlinear Mach–Zehnder interferometer. The device operates as an all-optical switch because in each arm of the interferometer a nonlinear phase change is induced by the intensity-dependent refractive index of the fibre. To have sufficient nonlinear phase shift for all-optical switching (e.g. 0.5\( \pi \)) requires long lengths of

![Figure 10. The three-trucks-on-a-mattress model of a soliton. The three trucks are low, medium and high powered and are travelling on a mattress-like material. As they go along, they create their own potential well, but the low-powered truck is always travelling downhill, the medium-powered truck is always on the flat and the high-powered truck is always travelling uphill. So the shape of the potential well is maintained as the soliton travels along. In the negative dispersion regime the low-powered truck is analogous to the red light (larger refractive index) in the soliton pulse, the high-powered truck corresponds to the blue light (lower refractive index) and the medium-powered truck to the average-wavelength light. In the real soliton the potential well is created by the intensity of the light changing the refractive index via the intensity-dependent refractive index \( n_2 \).](image)

![Figure 11. A fibre-optic version of a Mach–Zehnder interferometer. The fibre couplers consist of two fibres brought into close proximity and act like the fibre equivalent of beam splitters. The light is split between the two fibres which form the two arms of the interferometer in a ratio that is determined by the design of the couplers.](image)
fibres because $n_2$ is small in optical fibre. The difficulty with an interferometer made from long lengths of optical fibre is that it is subject to environmental fluctuations which make its operation unpredictable and essentially impossible to control. To overcome this difficulty the device is designed as an all-fibre Sagnac interferometer as illustrated in figure 12. In effect it is the same as a Mach–Zehnder interferometer but the two paths are the clockwise and counter-clockwise routes around the loop and the two couplers are replaced by one which is traversed twice. A pulse of light which is split at the input and travels in a clockwise direction around the loop experiences the same environmental fluctuations as the other part of the pulse which travels in a counter-clockwise direction. Thereby the environmental fluctuations are cancelled out, and the device has a predictable repeatable behaviour, making it suitable for experimental study.

The equation which describes operation of the loop mirror is derived from a consideration of an all-fibre interferometer where the power ratio in the arms of the interferometer is $\delta:1-\delta$ (Blow and Doran 1990). The output from the device is given by

$$|E_{out}|^2 = |E_{in}|^2(1 - 2\delta(1 - \delta))$$

$$\{1 + \cos^2[(1 - 2\delta)n_2k_0L|E_{in}|^2]\}. \quad (24)$$

As can be seen from the above equation, the transmission of the device is intensity dependent and an ultra-short control pulse of high intensity laser light can be used to switch the loop mirror on and off. The loop mirror can act as an all-optical AND gate which responds at very high speeds. Furthermore it is a very ‘cold’ switch, that is the switch does not consume a significant amount of power. These loop mirror switches could be employed in solitonic communication systems where the data rates are such that conventional electronic logic is too slow.

11.3. Solitons

The Kerr optical nonlinearity in fibres is essentially instantaneous and therefore even ultra-short pulses will tend to reshape or break up on switching because the various intensities present in the pulse will experience different nonlinear phase shifts in the interferometer and thereby different switching conditions. Pulse reshaping and break-up is undesirable in all-optical signal processing. A proposed solution to the above problem is to use solitons in these all-optical switching systems. The soliton acquires a uniform nonlinear phase shift and does not significantly reshape or break-up in the loop mirror (Blow and Doran 1990, Islam 1992). To some extent the soliton exhibits particle-like behaviour and one soliton can switch another soliton in a ‘billiard ball’ interaction that makes truly conservative logic possible where information is switched between channels in an ultra-fast fashion with essentially no or very little energy expenditure. As Blow and Doran (1990) have commented, the soliton becomes the ‘natural bit’ not only for information transmission but also for information processing. Islam (1992) has shown that billiard ball interactions can take place in so-called soliton interaction gates where solitons are switched intact without any pulse break-up. The energy required for the pulses for this type of interaction is around 15 pJ. Interesting questions arise as to whether in the future solitronics can take over from electronics in some applications where ultra-fast switching is required.

11.4. Waveguide devices

Optical waveguides fabricated on a wide range of materials provide more flexibility in both device design and material choice compared with optical fibres. If materials with large optical nonlinearities are selected, then it is possible to scale down some of the ultra-fast all-optical switching devices which were first realized in fibre format. It is also possible to introduce new device designs and to utilize both third- and second-order optical nonlinearities.

11.5. Third-order devices

Largely because of the commercial success of the semiconductor laser and associated optoelectronics there is a readily available technology for both the growth of high-quality semiconductor layers and the fabrication of channel optical waveguides with these semiconductor layers. Recent developments in the theory of optical nonlinearities outlined in the section on non-resonant optical nonlinearities indicate that operating with photon energies just below half the bandgap is optimum. This is illustrated in figure 13 for an alloy III–V...
The semiconductor waveguide structure is designed as an optimum structure for all-optical switching at 1.55 μm. Imperfections on the side walls of the semiconductor waveguide produce scattering losses and the linear losses in semiconductor waveguides are much larger than optical fibre. Losses of 1 or 2 dB cm\(^{-1}\) are typical compared with 0.2 dB km\(^{-1}\) for fibre. However, because the bandgap energy of semiconductors is much smaller than the glass used in optical fibre and also because the ratio \(h \omega / E_g\) of the photon energy to the bandgap energy is larger in the case of the semiconductor, then \(n_2\) is much larger by a factor of two or three orders of magnitude. Therefore the devices can be scaled down in size by approximately two to three orders of magnitude.

11.5.1. Asymmetric Mach–Zehnder interferometer

The first device to consider is simply a scaled-down version of the nonlinear loop mirror. It is the integrated asymmetric Mach–Zehnder interferometer (AMZI). The top view of the device is illustrated in figure 15. The AMZI device consists of an input waveguide, an asymmetric Y junction which splits the input irradiance into two uneven parts in the single-mode waveguides 1 and 2 as a function of the Y-junction angle \(\gamma\), and a second Y junction, also asymmetric but in the opposite sense, to recombine the two optical fields into the output waveguide. The output power density of the AMZI depends on the phase difference between the two beams incident on the output Y junction. This phase difference \(\Delta \phi = \phi_1 - \phi_2\) is achieved by means of an intensity-dependent refractive index in the waveguide material. The output transmission of a lossless AMZI is given by

\[
T = 4\delta(1 - \delta) \cos^2\left(\frac{1}{2} \Delta \phi + \theta\right).
\]  

(25)

\(\theta\) takes account of any built-in phase differences in the device due to slightly different optical paths in the arms caused by fabrication inaccuracies. The differential
phase change $\Delta \phi$ between the two arms due to the optical nonlinearity is given by

$$\Delta \phi = \frac{2 \pi n_2 I_{in} (1 - 2 \delta)}{J_0},$$

(26)

where $n_2$ is the nonlinear coefficient, $I_{in}$ is the pump beam intensity, $\delta$ and $1 - \delta$ are the optical power splitting between the two arms and $l$ is the length of the interferometer arms. The output response of the AMZI with the pump beam intensity is a $\cos^2$ function with amplitude and oscillating strength dependent on the Y-junction split ratio between the two arms.

The device was fabricated using the waveguides shown in figure 14. Its operation was characterized by employing a tuneable mode-locked F-centre laser. The laser was operated in two mode-locking regimes which provide pulse of either 30 ps or 300 fs ($300 \times 10^{-15}$ s). Both regimes produced pulse trains at the same repetition rate around 100 MHz and around the same average power. The 300 fs pulses were more intense by the ratio of the pulse widths, that is a factor of 100 more intense for the same average power. Figure 16 shows the transmission of the device for both types of pulse.

Two types of pulse are used in order to eliminate the possibility that thermal nonlinearities are responsible for the switching characteristic observed. Since the pulse trains have the same average power the thermal effects will be equivalent. Therefore the switch-off characteristic observed is due to the electronic nonlinearity. The value of $n_2$ deduced from the measurement is $5.4 \pm 0.5 \times 10^{-14}$ $\text{cm}^2 \text{W}^{-1}$ which is approximately half that predicted from the theory described in the section on the non-resonant nonlinearity and shown in figure 13 but within the expected accuracy of the theory. The device shows a modified type of $\cos^2$ switch-off characterized. The $\cos^2$ behaviour is modified by the fact that the pulse employed in testing the device contains a range of irradiances. Because fundamentally the material response is ultra-fast, even for 300-fs pulses the high-irradiance centre part of the pulse may be switched off but the low-irradiance wings are still transmitted. This effect causes the pulse to break up. Pulse break-up is treated in more detail in the section on the nonlinear directional coupler. In this device, pulse break-up accounts for the fact that the device does not completely switch off. The irradiance required to switch off the device is a peak intensity of about 3.92 $\text{GW cm}^{-2}$ in the waveguide which, if we take a waveguide cross sectional area of $4 \times 10^{-8}$ $\text{cm}^2$, translates into a peak power of 100 W in the pulse.

11.5.2. Nonlinear directional coupler

A nonlinear directional coupler is illustrated in figure 17. The nonlinear directional coupler is another waveguide all-optical switching device. The device has been realized in several types of material; glass, semiconductor-doped glass and semiconductors. The nonlinear optical mechanism responsible for the switching includes resonant and non-resonant optical nonlinearities. The

Figure 16. Transmission of the integrated AMZI at a wavelength of 1.53 $\mu$m with both 300-fs and 30-ps pulses for a device with interferometer arms 6 mm long. The 300-fs pulses are a factor of 100 larger in irradiance for the same average power. For the 300-fs pulses, several experimental runs are shown to give an indication of the spread in experimental data. The solid line is a theoretical fit to the data that takes account of pulse break up. (From Al-Hemyari 1992, and Al-Hemyari et al. 1992.)

Figure 17. Top view of the nonlinear directional coupler which consists of two optical waveguides in close proximity fabricated in a nonlinear optical medium. At low intensities the device is designed so that light coupled into the bar waveguide couples across to the cross waveguide. At high intensities the coupling is destroyed because the waveguides no longer have the same effective refractive index.
device was first proposed by Jensen (1982) who analysed its operation using coupled-mode theory described below.

At low powers, in the linear regime, light will couple across and back again into the original guide in a length known as the transfer length. If the device is designed so that the length is half a transfer length (a so-called interaction length), then the light coupled into the bar guide couples across completely to the cross guide at the output. This happens because the waveguides have the same propagation constant and there is a small spatial overlap field between them which provides the coupling. The nonlinear effect results in an intensity-dependent propagation constant. If the intensity input to the bar guide is sufficiently high, the two guides no longer have the same propagation constant and coupling between them is detuned.

The behaviour is analogous to two coupled pendulums of the same length; the pendulums will exchange energy in a regular fashion. If one of the pendulums is set swinging, it will stop after a time determined by the strength of coupling between the pendulums and the other pendulum starts to swing. In turn it will stop and transfer energy back to the original pendulum. This behaviour only happens if the pendulums are in tune, that is have the same resonance frequency determined by the length of the pendulum. If the pendulums are of different lengths, no exchange of energy occurs. In the nonlinear directional coupler the waveguides start off ‘in tune’, that is with the same effective refractive index. When the irradiance inside the input waveguide is increased, the effective refractive index is altered by the nonlinear optical effect and the waveguides are no longer ‘in tune’. Consequently coupling between the waveguides can no longer take place. Returning to our analogy of the pendulums, the nonlinear effect is equivalent to altering the length of one of the pendulums.

The operation of the nonlinear coupler is outlined below using coupled-mode theory following Jensen (1982). The amplitudes of the modes in each guide are defined as \(a\) and \(a'\); then the spatial evolution including the coupling between these modes is as follows:

\[
\frac{\partial a}{\partial z} = Q_1 a + Q_2 a' + (Q_3 |a|^2 + 2Q_4 |a'|^2)a, \quad (27)
\]

\[
\frac{\partial a'}{\partial z} = Q_1 a' + Q_2 a + (Q_3 |a'|^2 + 2Q_4 |a|^2)a', \quad (28)
\]

where \(Q_1\) to \(Q_4\) are coupling coefficients.

If all the power is launched into the bar guide, then the power in the bar as a function of propagation distance \(z\) is given by

\[
I_b(z) = \frac{1}{2} I_b(0) \left[ 1 + \text{cn} \left( \frac{\pi z}{L_c} m \right) \right], \quad (29)
\]

where \(m = \left[ I(0)/I_c \right]^2\), \(L_c\) is the linear coupling length (i.e. the distance required for the intensity to couple over completely from the bar guide to the cross guide in the linear regime) and \(\text{cn}(\mu|m)\) is a Jacobi elliptic function. \(I_c\) is the critical intensity which is defined as the intensity required to switch a coupler with an interaction length equal to \(L_c\) to the condition where equal intensities emerge from the bar and cross guides and is given by

\[
I_c = \frac{\lambda}{L_c n_2} \quad \text{.} \quad (30)
\]

Recently, a nonlinear directional coupler has been shown to operate with semiconductor waveguides at photon energies close to half the bandgap energy where there is a resonant enhancement of \(n_2\) but more importantly there is greatly reduced nonlinear absorption (figure 13) (Aitchison et al. 1991). The semiconductor chosen was Al_{0.18}Ga_{0.82}As which has half of its bandgap energy at a wavelength of around 1.55 \(\mu\)m which is in the optical communication band. The nonlinear directional coupler was 6.25 mm long and consisted of waveguides of 4 \(\mu\)m width separated by 5 \(\mu\)m. The device was tested by high-intensity pulses from a mode-locked F-centre laser producing a 10-ps pulse of light at a wavelength of around 1.55 \(\mu\)m. Figure 18 shows the transmission characteristics of each guide as a function of input intensity. As can be seen from figure 18, for a given waveguide the device has an intensity dependent transmission and acts as a high speed all-optical switch.

Just how high-speed switching is possible from this device was revealed by an experiment designed to investigate the pulse break-up effect alluded to in the section on the nonlinear interferometer. The pulse

---

**Figure 18.** Transmission characteristic of an Al_{0.18}Ga_{0.82}As nonlinear directional coupler. The coupler is 0.7 \(L_c\) long. The wavelength is 1.56 \(\mu\)m. The relative transmission of the bar and cross waveguides are shown as a function of input intensity (irradiance) to the bar guide.
break-up effect in the nonlinear directional coupler is illustrated in figure 19. There is sufficient light throughput from the directional coupler for use in other nonlinear measurements; about 30% of the input light is transmitted (most of the loss comes from coupling light into the waveguides). The pulses used in the experiment are too short to be measured electronically. The only technique which gives an indication of the pulse width is known as a second-harmonic autocorrelation technique which employs the ultra-fast second-order nonlinearity to measure the pulses. The technique does not directly measure the pulse shape but allows the pulse to measure itself. The pulses emitted from the device are split into two by a beam splitter. One is put through a variable-delay line which is simply a moveable reflector which changes the path length. The other is directed into a second-harmonic generation crystal. By varying the delay between the pulses they can be overlapped in time in the second-harmonic crystal or moved apart in time. The second-harmonic signal is proportional to the amount of overlap in time between the pulses. The product of overlap in time between the pulses is the autocorrelation of a pulse and thus the second-harmonic signal is proportional to the autocorrelation of the pulse. By scanning the delay the autocorrelation can be measured as a function of time and some information about the pulse shape can be deduced. Although this is not ideal, it is the only way to obtain information about the pulse shape of ultra-fast pulses.

Figure 20 (Villeneuve et al. 1992) shows the autocorrelation of the pulses from both the bar and the cross waveguides. The cross-waveguide autocorrelation is measured for a series of input irradiances. The input pulse has a full width at half-height of <500 fs. The autocorrelation of the cross-guide pulses show that the pulses have been broken up into two pulses as shown in figure 19. It is a feature of the autocorrelation that the maximum second harmonic signal is obtained at zero delay, that is maximum overlap between the pulses and therefore the autocorrelation of a double peaked pulse shows three peaks, as can be seen in figure 20 d. The autocorrelation shows that the input pulse can be split up into shorter pulses of <100 fs. The switch acts on

![Figure 19. The pulse break-up effect in a nonlinear directional coupler. The input pulse is assumed to have a Gaussian distribution of intensity with time with the peak 1.5 critical intensity. The central part of the pulse stays in the bar waveguide but the low-intensity wings cross over to the cross waveguide.](image)

![Figure 20. Autocorrelation of the pulses from a nonlinear directional coupler and the fit of a numerical simulation of the device. (a) the input pulse and output pulse for the bar guide; (b)-(d) the output pulses in the bar state when the input pulses have peak intensities of 5.5, 6.9 and 9 GW cm⁻² respectively.](image)
time scales <100 fs which makes it the fastest semiconductor switch yet demonstrated. However, in real systems, pulse break-up is undesirable and there have been various proposals to get round this difficulty including employing square pulses and using solitons (Islam 1992, Blow and Doran 1990).

11.6. Second-order devices

Second-order optical nonlinearity has been routinely employed in the laboratory for second-harmonic generation where light at a fundamental frequency is frequency doubled in a crystal with a $\chi^{(2)}$ optical nonlinearity to produce second-harmonic light. Recently it has been observed that second-order optical nonlinearity can also give rise to a nonlinear phase shift and effective $n_2$ via the so-called cascaded second-order optical nonlinearity (DeSalvo et al. 1992). To understand how this comes about requires some further consideration of the second-harmonic process. The process is parametric in nature which means that the relative phase relationship between the fundamental and the second harmonic is crucial to the efficiency of the process. Normally, owing to dispersion, the fundamental wave and the second-harmonic wave will have different refractive indices, and second-harmonic light generated at one end of the crystal will be out of phase with light generated at the other end and will destructively interfere with it. In order for the second-harmonic generation to be reasonably efficient, the mismatch between the second harmonic and the fundamental must be kept at zero. This condition is known as phase matching. There are various techniques for obtaining it. A recently introduced method has been to alter periodically the sign of the $\chi^{(2)}$ optical nonlinearity in the crystal which is known as quasi-phase matching. The phase mismatch is given by

$$\Delta k = k_{2\omega} - 2k_\omega - K,$$

where $k_{2\omega}$ is the phase propagation coefficient of the second-harmonic light, $k_\omega$ is that of the fundamental light and $K = \frac{2\pi}{\Lambda}$ where $\Lambda$ is the periodicity of the $\chi^{(2)}$ grating.

The nonlinear phase shift can be understood in terms of the second-harmonic generation process illustrated in figure 21. At the fundamental wavelength there is an effective nonlinear (i.e. intensity-dependent) loss and this is a maximum at $\Delta k = 0$ (the phase match condition) and varies as a sinc (i.e. $\sin x/x$) function. The nonlinear loss at the fundamental is equivalent to the two-photon loss discussed in the section on the third-order optical nonlinearity where we saw that a nonlinear loss gives rise to a nonlinear refractive index. The wavelength dependence of the nonlinear refractive index can therefore be calculated from the nonlinear loss at the fundamental by taking a Kramers–Krönig transformation as illustrated. It should also be noted that this nonlinear phase shift takes place at the fundamental wavelength. It is called the cascaded second-order effect because it comes about in the phase mismatch condition $\Delta k \neq 0$ where the fundamental is converted to the second harmonic and then the second harmonic is down-converted to the fundamental. The second-order optical nonlinearity acts twice or is 'cascaded'.

11.6.1. Push–pull switch

How can we employ the cascaded optical nonlinearity in ultra-fast all-optical switching? It has been proposed by Ironside et al. (1993) that the effect can be employed in an integrated Mach–Zehnder device which they term the push–pull switch. The device is illustrated in figure 22. The device is closely related to the AMZI discussed previously in the section on third-order nonlinear waveguide devices. With this second-order cascaded nonlinearity there is an extra design variable; the sign of the nonlinear phase shift can be controlled via the sign of the phase mismatch. Therefore by having a phase mismatch of equal magnitude but opposite sign in the arms of the interferometer the nonlinear phase shift can be positive in one arm of the interferometer and negative in the other arm, or one arm pushes and the other pulls.
The transmission of the device is given by

\[ T = (1 - D) \cos^2 \left( \frac{\theta_{n1} - \theta_{n2} + \zeta}{2} \right), \]  

(32)

where \( D \) is the depletion of the fundamental, \( \theta_{n1} \) and \( \theta_{n2} \) are the nonlinear shifts in arm 1 and arm 2 respectively of the interferometer and \( \zeta \) is the phase offset. The nonlinear phase shift is found by solving the coupled differential equations which describe second-harmonic generation. The details of the solution will not concern us here but the results of the calculations for Al\(_x\)Ga\(_{1-x}\)As is shown in figure 23. The calculated results shown that it is possible for the device to undergo an on-to-off transition with powers of around 2 W which is approximately \( 10^{-2} \) of the switching powers required using the third-order nonlinear effect. Furthermore the cascaded second-order optical nonlinearity is intrinsically low loss and ultra-fast.

12. Conclusions

Ultra-fast all-optical switching has been transformed in the last few years from a point where no such switches had been demonstrated to the present position where there are now several all-optical switches demonstrated in both fibre and waveguide format. The switches are all operated at photon energies below half the bandgap energy of the material which is crucial if strong nonlinear absorption is to be avoided. The figure of merit of the third-order optical nonlinearity plotted in figure 9 has been shown to be very important in choosing the optimum material for a given wavelength. It turns out that the optimum ratio of photon energy to bandgap energy is around 0.5. For optical communications wavelengths, Al\(_{0.5}\)Ga\(_{0.8}\)As material has the optimum nonlinear figure of merit but, because of its excellent linear optical properties, optical fibre is also a very good material system for ultra-fast all-optical switches.

A recent development has been the realization that large nonlinear phase shifts can be achieved via cascaded second-order nonlinearity. This type of nonlinearity can also have the sign of the nonlinear phase shift controlled which is very convenient for the 'push–pull' switch based on the integrated Mach–Zehnder interferometer. This switch may dramatically bring down the switching power in compact waveguide devices.

As switches, apart from their all-optical nature, the switches described here are unique in many respects. They operate at unprecedented speed, and high-speed data can be delivered to them in convenient form. Unlike electronic switches, they consume very little energy. Although they require energy to be present, the energy in the ultra-short pulses is not converted to heat. They are the so-called non-demolition switches (Yamamoto and Haus 1986) where the controlling pulse is not substantially absorbed and this type of switch is predicted to have interesting quantum optical properties. If their potential is to be fully utilized, it may well be that new concepts in both physics and system architecture will have to be developed.

Acknowledgements

The author thanks his many collaborators for their help in preparing this article and thanks the Science and
Engineering Research Council for their financial support of his research on waveguide switches through the Nonlinear Optics Initiative and the Nonlinear Waveguide Consortium.

References


