Micro-optoelectronics:
Materials,
Devices and Integration

a cura di Maurizio Ferrari
Lorenzo Pavesi
Giancarlo C. Righini

La Scuola è stata organizzata dal   The School was organized by:
Dipartimento di Fisica, Università di Trento

in collaborazione con   in cooperation with
Sez. di Trento, Istituto di Fotonica e Nanotecnologie – CNR (formerly CeFSA)
Sezione E dell'INFM
SIOF

con il patrocinio di   with the sponsorship of
Politecnico di Torino
TU Wien - Technische Universität Wien
INSA - Institut National des Sciences Appliquées de Lyon

In copertina (on the cover):
Silicon microbench for fibre-laser self-alignment: (clockwise from top left) schematic of the assembly; SEM photograph of the alignment ribs and V-groove; detail of the ribs and groove end; SEM photograph of the assembled laser chip and fibre. *(from the paper on pages 103-115)*

Basics on optical waveguides: materials, systems and technologies

Fred Hopper
Bookham Technology plc

Abstract

These lectures introduce the basic principles and technology of optical planar waveguides, concentrating on devices for the infrared region of the spectrum (1.3-1.5 μm) of interest in telecoms applications. The first part deals with the basics of optical slab and channel waveguides, the second introduces a number of waveguide components and the third illustrates some of the technologies that have developed based upon different materials systems.

Part I - Fundamentals

It is well known that light may propagate both in a vacuum and in a material medium. In the latter case the material modifies the characteristics of the light passing through it, in general reducing the speed of the light and absorbing energy from it. Contrast in properties between differing materials gives rise to the familiar effects of reflection and refraction as light passes

Fig 1 Snell's Law

from one medium to another. This contrast also leads to the phenomenon of waveguiding, in which it is possible to confine light to a plane or a channel under the influence of boundaries between regions of differing material. We are particularly interested here in wavelengths on the micron scale, where waveguides are conveniently fabricated in deposited layers on planar substrates. More diverse examples of the same underlying phenomenon include the guiding of microwaves through metal tubes and the reception of long wave radio transmissions, which follow the surface of the earth.
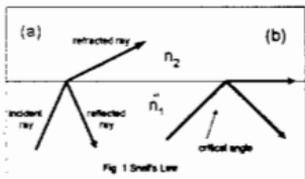
Ray description It is possible to visualise a guided wave using a simple ray description, familiar from geometrical optics. Fig. 1 illustrates the case of a ray propagating from a

Fig 2 Guidance by total internal reflection

medium of refractive index n_1 to a second medium with lower refractive index n_2. Reflection and refraction occur at the boundary, in such a way that the transmitted ray bends towards the boundary in the second material according to Snell's law (Fig. 1a). As the angle of incidence increases the ray reaches a critical angle (Fig.1b) beyond which there is no further propagation of energy into the second medium and total internal reflection occurs. If we add a second region of index n_2 (Fig. 2) the ray will be confined to the central high-index region by successive total internal reflections at the two boundaries.

Wave description In order to understand the guiding of light in more detail we need to appreciate the electromagnetic wave nature of light. Starting from a general description of electric and magnetic fields using Maxwell's equations (Fig. 3), a wave equation can be

deduced which describes the nature of light as a coupled oscillation of electric and magnetic fields. The simplest form of solution is a plane wave with electric and magnetic field vectors perpendicular to each other and to the direction of propagation (Fig. 4).

Considering firstly the case of propagation in a vacuum, the wave equation has the form shown in Fig. 5a. The plane wave solution can be conveniently represented using the complex notation of fig 5b. The wavevector $\mathbf{k}$ is defined such that it has a magnitude of $2\pi/\lambda$ (where λ is the wavelength) and lies in the direction of propagation of the wave. For light of a given

$$\text{curl } H = J + \frac{\partial D}{\partial t}$$

$$\text{curl } E = -\frac{\partial B}{\partial t}$$

$$\text{div } \mathbf{F} = 0$$

$$\text{div } D = \rho$$

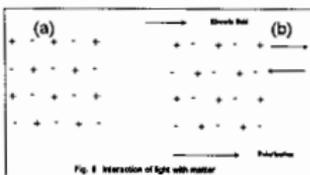Fig. 3 Maxwell's equations

Fig 4 Wave solution
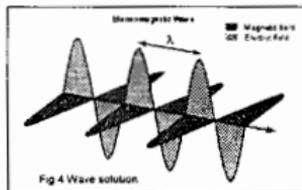
frequency the magnitude of the wavevector is given by $k = \omega/c$ (ω is the angular frequency, and c the velocity of light in vacuum). A physical interpretation of the vector nature of $\mathbf{k}$ is that the oscillatory time-dependence of the electric or magnetic field is balanced by a spatial oscillation which is shared among the x , y and z directions (fig 5c).

(a) $\quad \frac{\partial^2}{\partial x^2}E + \frac{\partial^2}{\partial y^2}E + \frac{\partial^2}{\partial z^2}E = \frac{1}{c^2}\frac{\partial^2}{\partial t^2}E$

(b) $\quad E = E_0 e^{i\omega t - \mathbf{k}\cdot\mathbf{D}}$

(c) $\quad k_x{}^2 + k_y{}^2 + k_z{}^2 = \omega^2/c^2$

Fig. 5 The wave equation and solution in a vacuum

Next we turn to the case in which light passes through a material medium. The effect can be understood by describing matter as a neutral assembly of electric charges (fig 6a). In the presence of an electric field positive and negative charges are displaced in opposite directions and the material becomes polarised (Fig. 6b), possessing a dipole moment, P, per unit volume. In this way the material develops its own electric field which opposes that originally applied, reducing the overall field in the region. This effect is conventionally described by defining the displacement vector, ($D = \varepsilon_0 E + P$) which appears in Maxwell's equations (Fig. 3). A dielectric constant, ε, is defined such that $D = \varepsilon\varepsilon_0 E$. The corresponding wave equation now has the form

(a) [diagram] Electric field (b) [diagram]

Fig. 6 Interaction of light with matter

shown in Fig. 7a, in which the refractive index, n, has been introduced and defined such that $n^2 = \varepsilon$. The consequence for wave propagation is that the phase velocity is reduced by the factor $1/n$ (Fig. 7b). Another aspect of this is that the wavelength of the light is reduced, or equivalently that the electric and magnetic fields oscillate more rapidly in space (possess higher k components, Fig. 7c) than was the case in a vacuum. It is this spatial frequency, and its dependence upon the material, which provides the basis for the guiding of waves between material boundaries.

(a) $\quad \frac{\partial^2}{\partial x^2}E + \frac{\partial^2}{\partial y^2}E + \frac{\partial^2}{\partial z^2}E = \frac{n^2}{c^2}\frac{\partial^2}{\partial t^2}E$

(b) $\quad E = E_0 e^{i\omega t - \mathbf{k}\cdot\mathbf{D}}$

(c) $\quad k_x{}^2 + k_y{}^2 + k_z{}^2 = n^2\omega^2/c^2$

Fig. 7 The wave equation and solution in matter

Boundary conditions In order to examine what happens when an electromagnetic wave reaches a boundary between dissimilar materials we recognise that the tangential components of E and H must be continuous along the boundary. Consider the arrival of a wave from a medium of refractive index n_1 at a boundary with medium of lower index n_2 (Fig. 8), and suppose that the electric field vector lies in a

direction parallel to the boundary, with the associated magnetic field perpendicular to the direction of propagation.
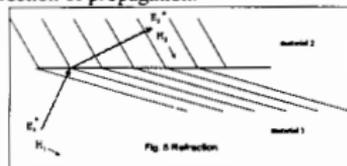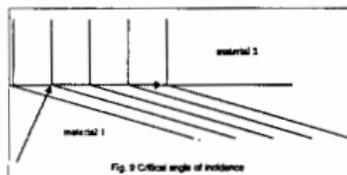

Fig. 8 Refraction

Fig. 9 Critical angle of incidence

Since the refractive index in region 1 exceeds that in region 2, the electric field in region 1 must vary more rapidly with distance than in region 2. At the same time field components must be continuous along the boundary, and must therefore share the same spatial frequency along the boundary. Equivalently, the propagation vectors in region 1 and region 2 must possess the same component along the boundary. In order to achieve this the transmitted wave in region 2 must be inclined towards the boundary as in Fig. 8. Snell's law of refraction follows from this requirement. The presence of a reflected wave, and the law of reflection also derive from the requirements of electric and magnetic field continuity.

As the angle of incidence increases the wave reaches a critical angle at which the entire spatial oscillation in region 2 takes place along the boundary (the wavevector k is directed along the boundary) with no component remaining to provide propagation in the perpendicular direction (Fig. 9)

As the angle of incidence increases further it would appear impossible to match fields at the boundary, since the spatial oscillation demanded in region 2 exceeds that of a freely propagating plane wave in that material. The situation is resolved, however, by allowing the field to decay exponentially in the perpendicular direction. This introduces a term of positive curvature – the evanescent field - to restore balance in the wave equation of Fig. 7c. Physically this means that energy no longer propagates away from the boundary into medium 2, and total internal reflection occurs. The solution remains oscillatory both parallel and perpendicular to the boundary in region 1.
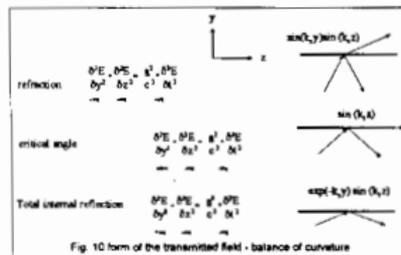
Fig. 10 form of the transmitted field - balance of curvature

The three situations are illustrated in Fig. 10, which gives expressions for the field in region 2, and shows the relative signs of the terms of the wave equation. It is apparent that the exponential decay in electric field perpendicular to the boundary in region 2 compensates for the fast sinusoidal variation along the boundary.

The slab waveguide. If we introduce a second region with refractive index lower than n_1 (as in Fig. 2) the additional constraint is imposed that the electric and magnetic fields must now match not only at the original boundary but also at the new one. The effect of this on totally internally reflected waves is to restrict the solutions to a finite number of allowed modes. The result is a slab waveguide, in which light is confined to the core layer of refractive index n_1, decaying exponentially in the surrounding regions of lower refractive index. Fig. 11 shows the first three modes for an example of the special case of a symmetric slab, in which the guiding core layer is bounded on either side by regions of identical material.

67

The fundamental mode has the highest wavevector component along the direction of propagation, and consequently the slowest transverse variation (in the x direction of Fig. 11).

Fig. 11 Equations and solutions of the slab

Successively higher modes possess larger transverse components of wavevector and correspondingly lower wavevector components in the propagation direction. Ultimately a cutoff condition is reached (corresponding to the critical angle) such that any waves with lower spatial frequency of oscillation along the boundary match to propagating waves in medium 2 and are no longer confined.
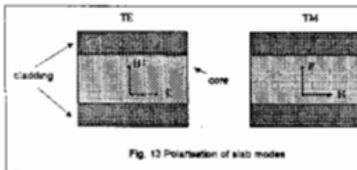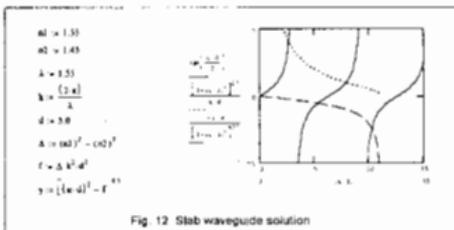
The equations of Fig. 11a and b express the wavevector relations (Fig. 7c) in the conventional language of texts on waveguides. β is the wavevector component in the direction of propagation, κ the transverse component in the guiding slab of index n_1, and the variable γ (which is equal to ik_y of Fig. 7) describes the exponential decay in the cladding region. A useful concept is that of the effective refractive index of a guided mode, defined by $n_{eff} = \beta/k$. The fundamental mode is the mode of highest effective refractive index (which must be less than n_1), and higher order modes possess successively lower index until the cutoff condition is reached at which the effective index falls to n_2.

Fig. 12 Slab waveguide solution

The field continuity conditions along the two boundaries lead to the equations 11c and d, which describe symmetric and antisymmetric solutions respectively. Fig. 12 evaluates 11c and d in terms of κd, after substituting from the relations 11a and b. It is apparent that the symmetric slab waveguide is completely defined by two parameters: (i) the ratio of core width to free space wavelength (via the term kd) and (ii) the contrast in refractive index between core and cladding, which is conveniently described by the term $(n_1^2 - n_2^2)$. Increasing kd and increasing the index contrast will both increase the number of guided modes.

Fig. 13 Polarisation of slab modes

An assumption at the beginning of this discussion was that the electric field is polarised parallel to the boundary. None of the boundary conditions has introduced any other vector components of electric field, hence the series of slab waveguide modes given by the equations of Fig. 11c and d retains this polarisation, and the modes are designated transverse electric (TE) as indicated in Fig 13 (where the direction of propagation is into the plane of the paper). The associated magnetic field is polarised perpendicular to the plane of the slab. Equally, if we start with the assumption of a wave with magnetic vector parallel to the boundary, we obtain the transverse magnetic (TM) solutions given by the equations of Fig. 14. In this case the electric field is perpendicular to

(a) $\tan(\kappa d/2) = (n_1/n_2)^2 \gamma/\kappa$

(b) $\tan(\kappa d/2) = -(n_1/n_2)^2 \kappa/\gamma$

Fig. 14 TM solutions of the slab

the slab (Fig. 13). The TM equations have similar form to the TE, but the factor (n_1/n_2)

appears which can be traced to the fact that while the tangential electric field must be continuous at the boundary, the normal component, which exists in the TM mode, changes by the factor n_1/n_2 at the boundary.
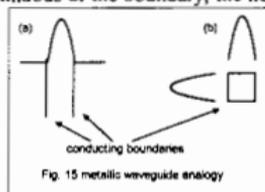

Fig. 15 metallic waveguide analogy

The channel waveguide Light propagating in a slab waveguide is confined in one dimension, but is free to propagate in two dimensions within the plane of the slab. Before turning to the case of confinement to a linear channel it is appropriate to mention briefly an analogous type of waveguide made using conducting metallic surfaces, as used traditionally to confine microwave radiation.

In the case of confinement by metallic walls the boundary conditions reduce to setting the electric field to zero along the conducting surfaces. Fig. 15a illustrates the equivalent of the slab waveguide, with confinement between two parallel conducting surfaces. A series of discrete modes is again obtained, with sinusoidal dependence normal to the surfaces. The solutions in this case are somewhat simpler than in the dielectric system, and can be constructed analytically by fitting an integral number of oscillations between the conducting planes. As in the dielectric case, the existence of transverse oscillation implies reduced index of propagation in comparison to a plane wave of the same frequency. Confinement to a one-dimensional channel by metallic surfaces works in the same way, (Fig. 15b) with the spatial oscillation now shared in all three dimensions, and solutions constructed again by fitting oscillation periods into the cross-sectional dimensions. Such waveguides have been used extensively at microwave frequencies, where the required dimensions are in the centimetre range.

Returning to the subject of dielectric optical waveguides, the analogue of fig 15b is the embedded dielectric channel shown in fig 16a. Such a structure does indeed act as a one-dimensional, or channel waveguide. In this case the situation is more complicated than in the microwave case, on account of the more complicated nature of the dielectric boundary conditions in comparison to their metallic analogues. In

Fig 16 Examples of channel waveguides

fact it is not necessary to surround the core region completely by region of lower refractive index. Thus the rib waveguide structure of Fig. 16b achieves the desired one-dimensional guiding by reducing the thickness of the high index layer at either side of the core region.
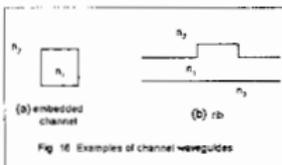
An analytical solution for the dielectric channel waveguide is not possible (even for the apparently simple embedded structure of fig 16a), though a wide range of approximations and numerical methods is available of varying accuracy. One of the most intuitive of these is the "effective index" method, which is illustrated in Fig. 17 for the case of the rib waveguide.

Using the effective index approach the two-dimensional confining geometry of the rib

slab 2   slab 1   slab 2   Fig 17 The effective index method

waveguide is analysed by two successive one-dimensional approximations. Firstly the cross-sectional structures of the central (B-B) and outer (A-A, C-C) regions are treated as independent slab waveguides (Fig. 17a), and are solved for effective index of propagation. Each slab waveguide is then rega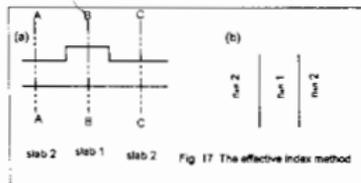rded as a region of uniform material possessing an effective refractive index equal to the index of propagation of that particular slab. Finally the three effective materials are thought of as the layers in a final slab waveguide which confines

69

light in the horizontal direction (Fig. 17b). Remarkably this approximation provides reasonably accurate results for effective index as a function of dimensions and index contrast, except at conditions close to cutoff.

Polarisation dependence We have seen that in the case of the slab waveguide it was possible to identify a series of modes with transverse electric polarisation and a series with transverse magnetic polarisation. This terminology is retained for channel waveguides, although the classification into TE and TM is now approximate. The presence of orthogonal boundaries causes mixing of polarisation via the term $(n_1/n_2)^2$ (Fig. 14), an effect which therefore becomes important as the index contrast increases. The various factors contributing to polarisation dependence in channel waveguides are (1) material birefringence, (2) cross-sectional geometry and (3) index contrast.

Loss The principal causes of loss in waveguides are those of intrinsic absorption and scattering within the material and scattering loss due to roughness of the edges. Techniques to measure loss include (1) detection of the scattered light by imaging the waveguide plane or scanning a detector alongside the waveguide, (2) the cut-back technique, in which the intensity of light tranmitted through a waveguide is repeatedly measured as the length of the waveguide is successively reduced and (3) a resonance technique in which loss is deduced from the finesse of a Fabry-Perot cavity fabricated by defining a pair of mirrors in the waveguide.

Coupling of light into waveguides Three methods will be described – endfire (butt) coupling, grating and prism coupling.

Endfire coupling requires a smooth facet to be prepared at the end of the waveguide, by polishing or cleaving. Light from a laser or optical fibre can then be coupled into the waveguide through the facet either by focusing a free space beam or by placing the source close to the prepared facet. Good coupling efficiency requires accurate positioning and orientation of the source and a close match of optical mode fields between the source of illumination and the waveguide. Considerable attention has been devoted to the problem of efficient coupling, and numerous solutions have been adopted, including the use of waveguide transitions or additional lenses to optimise the mode match.

A Bragg grating fabricated in the waveguide avoids the need for polished end faces by providing direct coupling of guided light into free space.

A third way in which to couple light, mainly used as a laboratory technique, is that of prism coupling (Fig. 18). It applies particularly to situations in which air provides the top cladding layer and the core is therefore exposed. A prism of refractive index higher than that

Fig. 18 Prism coupling
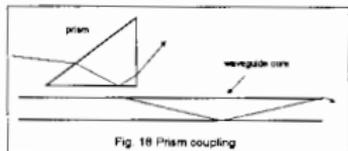
of the waveguide core is pressed close to the surface of the waveguide, in order to reduce the gap between prism and waveguide to the order of a wavelength. Light is introduced into the prism at angles chosen to allow total internal reflection to occur at the base of the prism. At certain discrete angles of incidence the wavevector component of the light resolved along the base of the prism will equal the index of propagation of a waveguide mode. Under these conditions light in the prism couples into the waveguide via the evanescent field in the narrow gap between prism and waveguide. Resonant dips in the reflectivity of the prism identify the waveguide modes.

Numerical solution A variety of methods has been developed to calculate the mode profiles and propagation indices of a channel waveguide, and also the evolution of a guided wave for cases such as tapers or branches in which the cross-section varies with distance. A simple way in which to solve for mode profile and index (Fig. 19) is to introduce an artificial boundary at some distance outside the region of confinement, where the field amplitudes are expected to be very low, and to set the field equal to zero on this new boundary. The field within can now be expanded in terms of two-dimensional sine and cosine functions, and the solution constructed algebraically. Care must be taken when choosing the location of the outer

Fig. 19 Numerical modelling

boundary: if too close to the waveguide core it will reduce the accuracy of results, and if too distant it will require too many terms to provide a good approximation.
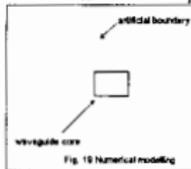
Part II - Components

The first lecture introduced slab and channel waveguides of constant cross-section. In this talk we consider some of the common waveguide components where in general the cross-sectional structure varies along the length of the waveguide.

Tapers It is frequently necessary to change the mode size of a guided wave. An example of this is the requirement to match the mode of a single-mode planar waveguide to that of a fibre in order to couple light efficiently between the two. One way in which to achieve this is to widen the planar waveguide gradually, as shown in Fig. 1a. As the waveguide expands, then

Fig. 1 Waveguide tapers

at some stage it will become multimoded, supporting an increasing number of modes as the width increases. If the propagating mode is to expand gradually, as depicted in Fig. 1a, then the light energy must remain in the fundamental mode as it propagates along the taper – in other words the transition must not excite any higher order modes.
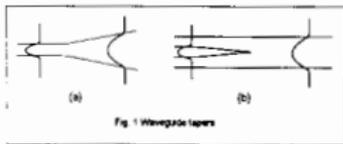
To study the problem in more detail one defines the "local normal modes" at each point along the taper – the set of modes that a waveguide of constant cross section would have at that particular width. If the taper is too abrupt it will cause mode coupling, exciting the undesirable higher order modes. To avoid this the taper must be made sufficiently gradual. In general a waveguide transition which is sufficiently gentle that it generates negligible coupling between local modes is said to adiabatic.

A second, and at first sight counter-intuitive example of a taper is shown in Fig. 1b. Again the object is provide mode expansion without local mode conversion. In this case the original waveguide core becomes thinner, causing the mode to expand into a wider, lower index contrast waveguide.

Bends and bend loss Bends are frequently required in order to accommodate more complex device structures. In order to visualise the physics of a waveguide bend, recall that the effect of the waveguide is to match the wavevector (or equivalently the spatial frequency) of light propagating in the core and cladding regions, and that this is achieved by introducing appropriate spatial
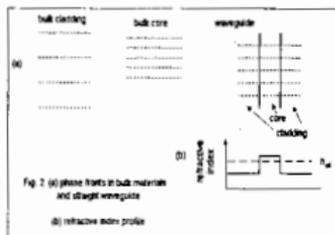

Fig. 2 (a) phase fronts in bulk materials and straight waveguide
(b) refractive index profile

71

dependence in the orthogonal directions. Thus an infinite plane wave of the same frequency would propagate with higher spatial frequency in bulk material of which the core is composed, and lower spatial frequency in the cladding material (Fig. 2a). Now consider a
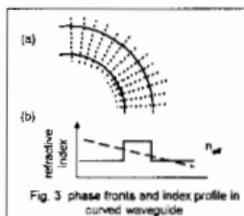

Fig. 3 phase fronts and index profile in curved waveguide

curved waveguide. While a rigorous solution of this case is complicated, a simple argument can be made as follows. In order to retain the character of a guided mode, light travelling along the curved waveguide must have a spatial frequency which varies inversely as the radial distance from the centre of the bend (Fig. 3a). Thus the mismatch in plane wave spatial frequencies between core and cladding is reduced at the outer boundary, in comparison to the straight waveguide case. The strength of the exponential decay of electric field in the cladding needed to match fields along the

outer boundary is thus reduced, hence the wave is less strongly confined in the cladding on the outside of the bend. At some particular bend radius the oscillation of field along the outer boundary will just match that of a plane wave in the cladding, and the mode will no longer be confined, but will be free to radiate into the cladding.

An alternative way to consider the bend problem, also illustrated in the diagrams, is to look at the refractive index profile through the waveguide cross-section, with the effective index superimposed. Fig 2b shows the effective index of the straight waveguide, while Fig. 3b shows the radially-dependent effective index needed to describe propagation around the curve. In the straight waveguide the mode is confined to the core by virtue of the exponential decay extending to infinity in each cladding region, where n_{eff} exceeds the local refractive index. In the case of the curved waveguide the effective index needed to provide guiding now falls with increasing distance from the centre of the bend. Consequently the region of exponential decay on the outside of the bend is now finite, extending to the point at which n_{eff} falls below the refractive index of the cladding. Beyond this point the field in the cladding is wave-like. Light is able to leak through the evanescent region and radiate into the cladding. As we would expect from this simple argument, the effect of curvature is more serious in waveguides with low index contrast.

A further effect of curvature is that the mode field within the core tends to be displaced towards the outside of the bend. In order to couple sections of different bend radius it is customary to introduce a stagger at the junction to compensate for differing displacements of the mode field, particularly in waveguides of low index contrast.

Y junctions A common building block in planar lightwave circuits, the Y junction allows the light energy in a guided mode to be divided between two output waveguides. It is another example of a gradual transition, in which the fundamental local mode evolves as shown in Fig. 4 for the case of a symmetric junction. Assuming that input and output waveguides are single-moded, then the usual design objective of such a transition is to minimise coupling to higher order modes or to radiation, and so maximise the efficiency of transfer . This is achieved by tapering the input waveguide and by choosing a low branching angle for the output waveguides. Under these conditions light propagating from left to right in the input waveguide will excite the two output waveguides with equal amplitude and phase, following the evolution of the local normal mode.
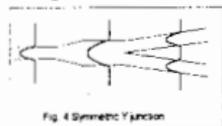

Fig 4 Symmetric Y junction

Fig. 5 What happens in reverse ?

Having considered the effect of the Y junction in the "forward" direction, it is instructive to ask what will happen if light propagates towards the junction in the "reverse" direction from one of the output waveguides (Fig. 5). In order to investigate this case we need to recognise that the excitation depicted in Fig. 5 is not a normal mode. We therefore need to identify the local normal modes in the output section. The symmetry of the problem dictates that the fundamental mode at the output involves a symmetric excitation of the individual output waveguides (which we have already encountered, Fig. 4) and that the second mode is antisymmetric, in which the two output waveguides are excited with equal amplitude but opposite phase. The state described by Fig. 5, in which a single output waveguide is excited, can be described by the superposition of symmetric and antisymmetric local normal modes (Fig. 6).
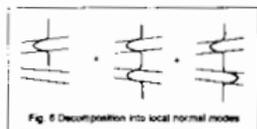

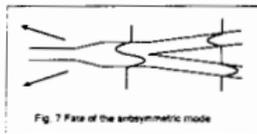Fig. 6 Decomposition into local normal modes

Fig. 7 Fate of the antisymmetric mode

We next consider the fate of each of these normal modes in turn. We know that the symmetric mode, being the fundamental mode, will evolve into the fundamental mode of the input side – this is seen to be the case simply by reversing the situation illustrated in Fig. 4. The antisymmetric mode, in contrast, evolves as shown in Fig. 7, coupling to radiation and so disappearing from the core at the input side of the junction. Since the arriving mode was made up of equal symmetric and antisymmetric components, it follows that half of the energy is transmitted to the input waveguide, and half is lost to radiation. The Y junction thus behaves as a 3dB attenuator for light propagating in the reverse direction.

Mach-Zehnder interferometer Using Y junctions and bends we can now construct the Mach-Zehnder interferometer illustrated in Fig. 8, which is the planar analogue of the well-known bulk optics construction. The operation of the interferometer can be understood from the forgoing explanation of the Y junction beam splitter. Light arriving from the left at the first Y junction is divided equally between the two branches and propagates independently in the two arms of the interferometer. If the two arms are perfectly symmetric, then the excitations arriving at the second Y junction will be in phase. Together they constitute the fundamental, symmetric mode of the second Y junction, and will evolve into the fundamental mode at the output end of the interferometer. The incident mode is therefore transmitted undiminished to the output.
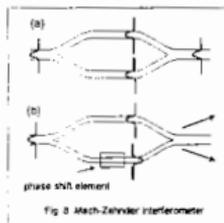

fig 8 Mach-Zehnder interferometer

Now suppose that some means is found to introduce a phase shift of π radians into one of the arms of the interferometer relative to the other. The excitation arriving at the output Y junction is now the antisymmetric local normal mode, which as we have seen couples into radiation and is lost from the waveguide. Thus the incident mode is fully attenuated, and nothing is transmitted to the output waveguide of the interferometer. The same reasoning can be extended to the general case in which an arbitrary phase difference, ϕ, is introduced between the two arms of the interferometer. Taking the amplitude of the mode propagating in the upper arm of the interferometer to be unity, and representing that in the phase-shifted arm as $e^{i\phi}$, the symmetric and antisymmetric modes can be expressed as $1+ e^{i\phi}$ and $1- e^{i\phi}$ respectively. To find the output of the interferometer we simply need the amplitude of the symmetric mode, which is proportional to $\cos(\phi/2)$. Thus the transmitted amplitude and therefore intensity vary sinusoidally with phase difference.

In practice the phase difference required to vary the output intensity can be introduced by locally changing the refractive index of the waveguide, exploiting the dependence of

Fig. 9 Path-imbalanced Mach-Zehnder interferometer

refractive index on temperature, electric field or carrier concentration, where the materials properties allow. Such a device performs the function of a modulator or variable attenuator.
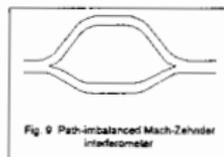
A second way in which to introduce a phase shift between the two arms of the interferometer is by arranging that the paths are of different length. Such a path-imbalanced interferometer will now have a phase difference which depends on wavelength, and hence, using the same argument as above, has wavelength-dependent transmission. Used in this way the Mach-Zehnder interferometer provides a way in which to implement filtering, for example to separate 1310 and 1550 nm signals.

Evanescent coupler Confinement of a guided wave has been shown to be analogous to the phenomenon of total internal reflection encountered in bulk optics at the boundary between semi-infinite regions of high and low refractive index (Fig. 10a). If the region of low refractive index is no longer infinite in extent, but instead a second region of high refractive index is allowed to approach the

Fig. 10 Frustrated total internal reflection

boundary, some light leaks across the gap into the second high index medium (Fig. 10b). Total internal reflection is "frustrated", and the reflectivity falls below 100%. The equivalent behaviour is found with guided waves. We have so far assumed the cladding layer of an ideal waveguide to be infinite in extent. If we relax this assumption and allow a second waveguide core to approach the first, the fields in the two will couple via the intervening evanescent field, allowing exchange of energy between the two waveguides. The strength of coupling depends exponentially on the separation of the two high-index regions, and only becomes appreciable for separations of the order of the wavelength. (The examples of prism coupling and of curved waveguides also display the phenomenon of evanescent coupling).
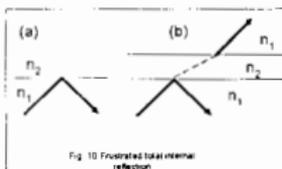
Fig. 11 shows the local normal modes of a system of two identical, parallel waveguides. Symmetry dictates that the modes are symmetric and antisymmetric, respectively. (We have already encountered this problem in discussing the Y junction.) At high separations the two modes are degenerate, in other words they have the same effective

Fig. 11 Evanescent coupler and local normal modes

index of propagation. As the separation between guides is reduced, evanescent coupling between the two waveguides lifts the degeneracy, giving the symmetric mode a slightly higher effective index of propagation than the antisymmetric mode.
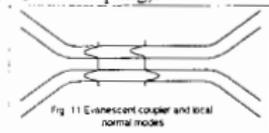
In view of the exponential dependence of coupling strength upon waveguide separation it is reasonable to define an ideal structure in which two uncoupled waveguides converge to a distance at which coupling becomes effective, running parallel at this distance, before diverging at the output, as illustrated in Fig. 11. Coupling can therefore be considered to begin abruptly at the start of the parallel section.

Consider a guided wave incident in the upper input arm of the coupler. In order to look at what happens in the coupling region we must first resolve the incident mode into the normal modes of the coupling region (as shown in Fig. 12 for both input arms). As the symmetric and antisymmetric modes propagate along the coupled section they travel with slightly different phase velocities. If the length of the coupler is chosen such that the

Fig. 12 Decomposition into local modes
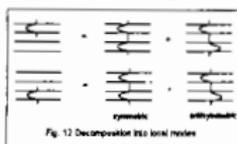
two local normal modes acquire a phase difference of π radians by the time they reach the output end, they will now cancel in the waveguide from which light originally entered the coupler and reinforce in the other waveguide. Thus the effect of the coupler in this case is to allow complete crossing from one waveguide into the other. In general the light energy oscillates back and forth from one waveguide to the other as it propagates, so that a given fraction of the incident light can be transferred by appropriate choice of coupler length. The result is a splitter which can be designed to achieve an arbitrary splitting ratio.

Mach-Zehnder switch Replacing the output Y junction of the Mach-Zehnder modulator described above by an evanescent coupler allows us to implement the function of an optical switch (Fig. 13). Its operation follows directly from the discussions of the interferometer and evanescent coupler. If the path lengths through the two arms of the device are identical, then the symmetric mode will be excited at the evanescent coupler. This mode will propagate

Fig 13 Mach-Zehnder switch

unchanged through the coupling region, and the emergent light will be equally divided between the two outputs (Fig. 14a). Similarly, a phase shift of π radians between the two arms will cause the antisymmetric mode to be excited at the output coupler. Again, this mode will pass unchanged through the coupling region, and the output will be equally divided between the two waveguides (Fig. 14b).

Now consider what happens when a phase shift of $\pi/2$ radians is applied between the two arms of the interferometer. In this case the situation is different, and is conveniently examined using a phasor approach (Fig. 15). The two individual waveguide modes, Φ_a and Φ_b, arriving at the start of the evanescent coupler are represented by phasors of equal
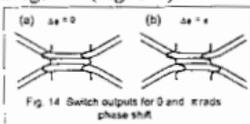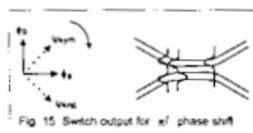

Fig. 14 Switch outputs for 0 and π rads phase shift

amplitude and $\pi/2$ phase difference. We can construct the symmetric and antisymmetric modes on the same diagram, and find that they must also have equal amplitude and $\pi/2$ radians phase difference. Supposing that the length of the coupling region is chosen to provide an additional phase difference of $\pi/2$ radians between symmetric and antisymmetric normal modes. The two modes will now have

Fig 15 Switch output for $\pi/2$ phase shift

accumulated a phase difference of either zero or π radians at the output end of the coupler, and the light which emerges will therefore be entirely in the upper or lower waveguide, according to the sign of the quadrature phase shift applied in the interferometer (Fig. 15). This result, and more general behaviour at intermediate coupling lengths and phase shifts, can be deduced directly by expressing the symmetric and antisymmetric normal modes as the sum and difference respectively of the modes of the individual waveguides.

Multimode interference coupler An alternative, and compact form of splitter (Fig. 16) is provided by the self-imaging property of a multimode waveguide. If a waveguide is made wide enough to support several modes, then the various modes travel with slightly different effective index. The sum of the individual mode amplitudes (in other words the modal interference pattern) therefore varies periodically along the waveguide. Fig. 17 shows how this effect may be exploited to design a coupler. A single-mode input waveguide joins the

Fig. 16 Multimode interference coupler

multimode section abruptly. In order to predict what happens subsequently we expand the input mode field in terms of the modes of the multimoded waveguide section. In this case the first and third modes are excited predominantly. As these two modes propagate they develop
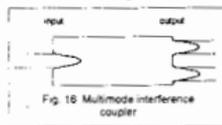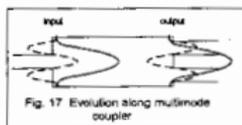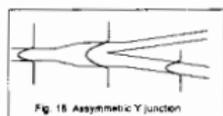
a phase difference, and the resulting interference pattern evolves as shown. Suitable positioning of two single mode output waveguides allows a 1 by 2 coupler to be obtained. Clearly higher order couplers can be designed in this way by taking advantage of higher order modes in the multimode section.


Fig. 17 Evolution along multimode coupler

Digital optical switch A disadvantage of the Mach-Zehnder switch described above is that the output state oscillates with phase shift and with coupling parameters, so that precise conditions must be maintained in order to hold the output stable. This property is characteristic of a device which relies for its effect on the interference between modes. If instead the output state can be defined by the evolution of a single mode then it is possible for the switch to display something approaching a threshold (or digital) response with respect to the controlling phase shift.

Fig. 18 Assymmetric Y junction

Consider the asymmetric Y branch of Fig. 18. We must first establish the nature of the local normal modes throughout the structure. In cases where two waveguides are equivalent, such as the symmetric Y junction and the evanescent coupler, we have seen that they contribute equally to the local normal modes, which have symmetric and antisymmetric character. However when strong asymmetry is present this no longer holds, and the individual waveguide modes themselves become the local normal modes of the composite structure. The wider of the two output branches has the higher effective index and contains the fundamental local mode, while the narrower branch conveys the second mode. Thus the fundamental mode of the input waveguide on the left in Fig. 18 evolves into the wider output branch. Guided light incident from the left will therefore emerge entirely in the wider output waveguide, provided that the transition is sufficiently gradual. Similarly, light arriving from the right in the wider branch will propagate through the structure in the fundamental mode and will emerge on the left in the fundamental mode of the input waveguide. In contrast light arriving via the narrow branch will couple to radiation and will not appear at the input waveguide.

Suppose that it is possible to increase the refractive index of the narrow waveguide to such an extent that the asymmetry is reversed – in other words so that the narrow guide now supports the fundamental mode by virtue of its higher effective index. In this case the light arriving from the left will emerge from the narrow waveguide, and the function

Fig. 19 DOS characteristic

of a switch is obtained. Further increase of refractive index in the narrow waveguide will not alter the state of the switch, which is more robust in this respect than the Mach-Zehnder interferometer. In general a more significant change in refractive index is needed to achieve this effect than to bring about switching using an interferometer. Fig 19 shows schematically the characteristics of the digital optical switch.
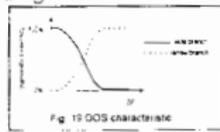
Star coupler All of the components discussed so far have been made from channel waveguides. The star coupler is an example in which both channel and slab waveguides are employed together. It is particularly useful for achieving 1 to n coupling for n larger than about 4. An input channel waveguide terminates in a slab region (Fig. 20) where light, while still confined to plane of the slab, is free to diffract within this plane and so illuminate an array of output waveguides, suitably arranged at some distance


Fig. 20 Star coupler
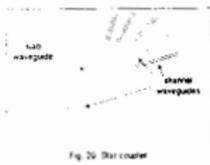
from the input. Large output port counts (up to 128) have been demonstrated in this way.

As might be expected it is also possible to make n by n and m by n couplers in the same way, by arranging suitable arrays of channel waveguides at both input and output.

Arrayed waveguide grating Of central importance for dense wavelength division multiplexing systems, this component consists of an array of waveguides in which the path length of each waveguide differs by a constant amount from that of its neighbour. A wavelength-dependent phase difference thus develops as light propagates along the array. Light is coupled into and out of the array using star couplers, as shown in Fig. 21. The diagram illustrates an example in which an array is connected to multiple input and output channel waveguides via the star couplers. In some cases a single channel waveguide provides the input.

In order to visualise the operation of the device, consider what happens to light at some wavelength λ_0, chosen such that the optical path difference between successive waveguides of the array is equal to $n\lambda_0$, where n is an integer. Assuming that light arriving at the input end of the array has the same phase at all waveguides, then light at the wavelength λ_0 will once again be

Fig. 21 Arrayed waveguide grating
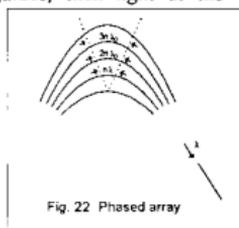
in phase across the array at the output end. This end of the array acts like a diffraction grating, and the uniform phase front ensures that light at the wavelength λ_0 interferes constructively on the central axis (Fig. 22) where the nth order of diffraction is directed. Light of a different wavelength will accumulate a phase difference which varies linearly across the array, and consequently its diffraction pattern will be shifted off the central axis.

Fig. 22 Phased array

In this way the angle at which light emerges from the array is a function of wavelength. Wavelength multiplexing and demultiplexing can therefore be accomplished by positioning channel waveguides to collect light appropriately from the final star coupler.
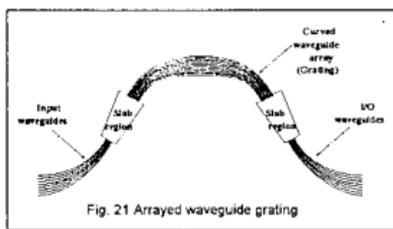
Part III - Technologies

Planar waveguide technology has developed around a number of materials systems. Passive components have been demonstrated in several materials systems at varying levels of integration. The same holds true for active devices in which the refractive index is varied by exploiting its dependence on temperature (for example in silica), electric field (as in lithium niobate) or carrier concentration (in semiconductors). Finally, near infrared sources and detectors have been integrated within III-V semiconductor planar waveguide circuits. In general planar waveguides are produced using the techniques of layer deposition, lithography, etching and diffusion adapted from the silicon integrated circuit industry. Features of several of the main technologies are summarised briefly.

Deposited silica
Layers of silica are deposited, most commonly by flame hydrolysis deposition or by chemical vapour deposition, typically on a silicon substrate. As in silica fibres the waveguide core is

defined by raising the refractive index of the silica using a suitable dopant such as germanium. In this way low refractive index contrast waveguides are typically achieved. The technology has the advantage of low loss and good fibre coupling, but the restriction of relatively high bend radius, and therefore large chip areas, imposed by the low index contrast.

Active functions such as switching have been demonstrated using thin film metal heaters to adjust the refractive index locally via the thermo-optic effect.

Lithium niobate

Lithium niobate is a mixed oxide, belonging to the Perovskite family of crystalline materials well known for their high and nonlinear dielectric response. Fig. 1 describes a general

$$P = \varepsilon_0(\chi^{(1)}E + \chi^{(2)}E^2 + \chi^{(3)}E^3 + \ldots)$$

Fig. 1 Nonlinear dependence of polarisation on electric field

nonlinear response, in which the polarisation induced in a medium is expressed as a power series in the applied electric field. Lithium niobate has the important property that its crystal structure lacks inversion symmetry and consequently the even terms in Fig. 1 are allowed. In particular the term $\chi^{(2)}$, which expresses a component of the dielectric polarisation proportional to the square of the electric field, is equivalent to a linear dependence of refractive index on the applied electric field. This is the linear electro-optic, or Pockels effect, which underlies the operation of electro-optic lithium niobate waveguides. A further consequence of the crystal symmetry is that lithium niobate is strongly birefringent.

Closely related to the electro-optic property of lithium niobate is its piezo-electric response. This is the variation of elastic strain with applied electric field, and is again allowed because of the lack of inversion symmetry. This property leads to an important class of accousto-optic devices such as the tunable Bragg grating. In this device a radio frequency signal establishes a periodic strain field in the crystal, which in turn modulates the refractive index, inducing a Bragg grating.

Waveguides are typically defined in lithium niobate by the in-diffusion of titanium, which increases the refractive index locally in a manner similar to that achieved in some of the earliest planar waveguides made in glass substrates by ion-exchange. An alternative technology exists in which protons are in-diffused, producing waveguides which guide only a single polarisation. In general these waveguides are of low index contrast, with the associated advantages and disadvantages.

Numerous types of electro-optic device have been demonstrated, including modulators operating at bit rates up to 40 Gb/s, switches and polarisation controllers.

Polymers

Polymers have attracted much interest for planar waveguide technology because of the possibility for the design of new materials by applying the appropriate polymer chemistry and because of the potential for low-cost manufacturing. A number of suitable transparent polymer families exist, although in many cases C-H vibration overtones lead to residual absorption in the telecoms window at 1.55 um.

It is possible to design electro-optic polymers by incorporating molecular species which possess a suitable nonlinear optical response such that the polarisability of the molecule, and therefore its contribution to refractive index, varies with electric field. These properties are found in many asymmetric conjugated molecules, in which the electronic orbitals responsible for the nonlinear polarisation are delocalised over the length of the molecule. Such molecules frequently possess permanent dipole moments which allow the material to be ordered under an electric field applied at high temperature and removed at lower temperature once the order has been frozen in place - a process known as poling (Fig. 2).

High speed modulation and switching have been demonstrated in polymers, which although less mature than lithium niobate possess the advantage of lower radio frequency dielectric constant, and hence better phase-matching of the rf electric field driving the modulator to the light propagating through it.
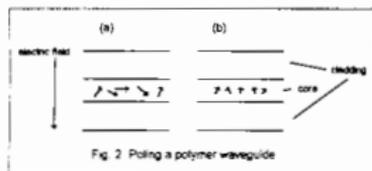
Among the challenges facing polymer technology are the demonstration of stability, both the inherent chemical stability of the organic molecules, and also the stability of the poled order of an electro-optic polymer, which in fact represents a metastable state susceptible to relaxing over time.

Fig 2 Poling a polymer waveguide

III-V Semiconductors

A rich variety of devices exits in the 1.3–1.5 μm telecom windows, based upon III-V semiconductor technology. Lasers, photodiodes, modulators and switches are possible, and more highly integrated devices have been demonstrated in the laboratory. The technology is covered by other series of lectures in the Winter School.

Silicon

Silicon provides a low-loss substrate for planar waveguides, and the maturity of silicon integrated circuit technology gives a route to low-cost, high-volume manufacturing. A particularly attractive realisation, developed by Bookham Technology, is that employing silicon-on-insulator substrates which are commercially available, and were developed in the first instance for high performance integrated circuits.
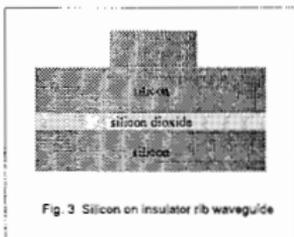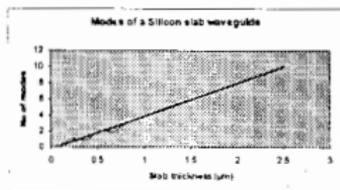

Fig. 3 Silicon on insulator rib waveguide

A rib structure allows single mode waveguides to be produced (Fig.3), and the high index contrast available provides a route to small radius bends and future chip size reduction.

It is instructive to consider the nature of the single mode rib waveguide in more detail. The silicon-on-insulator slab supports a large number of modes (Fig. 4), so that at first sight a single-mode rib waveguide looks implausible. Using the approach of the effective index method (Part 1, Fig. 17), the effecti

Fig. 4 SOI slab modes

ve indices of slab waveguides with the cross sections AA and BB are found to differ only slightly, despite significant height difference, and the rib can thus be shown to support a single mode for practical choices of width and step height.

Silicon-on-insulator technology offers three basic building blocks: passive waveguides, active waveguides and hybrid devices obtained by the addition of components

Fig. 5 Example AWG characteristics

such as lasers and photodiodes. The technology is modular, such that these three elements can be integrated as desired to perform complex functions.
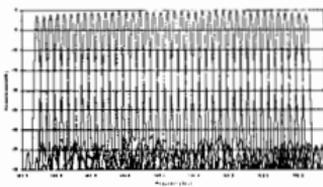
Passive waveguide components developed by Bookham Technology include wavelength multiplexers and demultiplexers for dense wavelength division multiplex systems, based on arrayed waveguide gratings. Fig. 5 shows an example of a 40 channel multiplexer transfer characteristic.
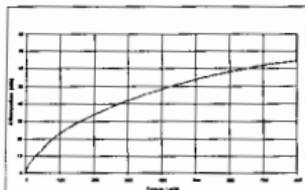

Fig. 6 Attenuator characteristic

Active components are possible by varying refractive index using either the thermo-optic effect or by injecting free carriers into the waveguide. An example of such a device is a variable optical attenuator in which attenuation of up to 30 dB is achieved as a result of the absorption of free carriers injected into the waveguide using a pin diode structure. Fig. 6 shows an example of an attenuator characteristic. Modulation and switching have also been demonstrated using this technology.

An example of a hybrid device is the transceiver of Fig. 7, in which a laser and two photodiodes are located in suitably etched cavities within the silicon chip, and interfaced to fibre via silicon rib waveguides. One photodiode acts as the receiver, while the second monitors the power emitted by the laser.

One of the attributes of silicon technology is that higher levels of integration are readily achievable. For example an optical channel monitor has been implemented by hybrid integration of an arrayed waveguide grating demultiplexer and a photodiode array.

An example of a highly-integrated monolithic device, incorporating both active and passive functions, is illustrated in Fig. 8. It consists of a 40 channel wavelength multiplexer with a variable optical attenuator at each input wavelength channel.
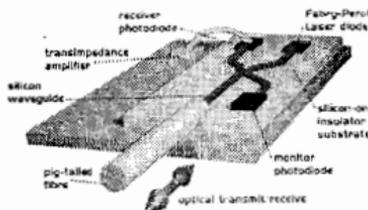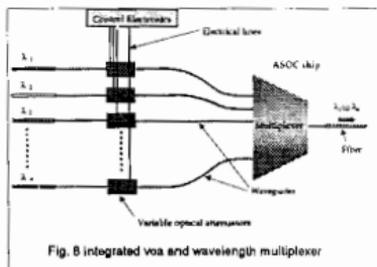

Fig. 7 Hybrid transceiver

Fig. 8 integrated voa and wavelength multiplexer

Acknowledgements