INTEGRAL OPTICS
LECTURE NOTES

*Recommended by the Methodical Council of Igor Sikorsky Kyiv Polytechnic
Instutute as a tutorial for master's degree students studying
for specialty 153 "Micro- and nanosystem technique"
educational program "Micro and nanoelectronics"*

*Stamp assigned Methodical Council of Igor Sikorsky Kyiv Polytechnic Instutute
(Protocol № 10 of June 18, 2020)
at the request of the Academic Council of the Faculty of Electronics
(Protocol № 05/2020 of May 25, 2020)*

Electronic online educational publication

INTEGRAL OPTICS
LECTURE NOTES

Compilers: *Svechnikov Georgii Serhiiovych,* Doctor of Philosophy, Associate Professor

Didenko Yurii Viktorovych, Doctor of Philosophy, Associate Professor

Responsible
Editor   *Tatarchuk, D. D.,* Doctor of Philosophy, Associate Professor

Reviewer: *Oleksenko, P. F.*, Doctor of Technical Sciences, Professor

An introduction is given to the principles of integrated optics and optical guided-wave devices. The characteristics of dielectric waveguides are summarized and methods for their fabrication are described. An illustration is given of recent work on devices including directional couplers, filters, modulators, light deflectors, and lasers.

The textbook reflects the latest achievements in the field of integrated optics, which have had a significant impact on the development of communication technology and methods for transmitting and processing information.

Contents

INTRODUCTION

The appearance of "Integrated Optics: An Introduction" by Stewart E. Miller in the September 1969 *Bell System Technical Journal* signalled the birth of an activity that now occupies thousands of researchers, Fig. I.1. Bell Labs, among others, had been concerned with optical communications for some time, encouraged in this activity by the advent of the laser. Miller pointed out that the then typical optical telephone repeater, involving a laser, modulator, detector, lenses, and so on, spread out on an optical bench, was a form of extremely short-range radio communication and as such suffered from a number of difficulties. The apparatus was sensitive to ambient temperature gradients, to temperature changes, to mechanical vibrations of the separately mounted parts. The elegant solution to these problems proposed by Miller was to combine the separate components on the same substrate or chip, connecting them by miniature transmission lines or waveguides. Because the size of the components need only be of the order of the wavelength of light in one, and possibly two, dimensions, the substrate could be quite small – centimeters or less. He proposed calling such an assembly of components an "integrated optical circuit," in view of its analogy to the assembly of electrical components on an integrated circuit chip.

Fig. I.1. 1969 BSTJ Cover and First Page of S. E. Miller article.

S. E. Miller, "Integrated optics: An introduction,"Bell Syst. Tech. J., vol. 48, p. 2059, 1969.

Integrated optics (IO) is one of the key enabling technologies of the 21st century, helping to overcome the bottlenecks faces by current electronics. Several decades ago, optical fibers revolutionized the communication field, providing affordable connectivity between people in different parts of the world. A prime example is the widespread of high-speed internet and mobile communications which was possible thanks to optical fiber technology. Integrated waveguides are on-chip versions of optical fibers. First proposed in the 1960's, integrated optical circuits are analogous to electronic integrated circuits. The main difference is that the information is processed in the form of "light" instead of by electrical signals, enabling therefore much higher transmission speed and processing capabilities. By selecting the correct set of materials, waveguiding structures, that confine and route light with dimensions in the micrometer and even nanometer scale, can be integrated on a chip.

The field of integrated optics is concerned with the theory, fabrication, and applications of guided wave optical devices. In these structures, light is guided along the surface region of a wafer by being confined in dielectric waveguides at or near the wafer surface. The light is confined to a cross-sectional region having a typical dimension of several wavelengths. Guided wave devices that perform passive operations analogous to classical optics (e. g., beam splitting) can be formed using microelectronic-based fabrication techniques. By fabricating devices in active materials such as ferroelectrics, modulators and switches based on the classical electro-optic effect can be formed. By fabricating the devices in compound semiconductors, passive and active guided wave devices can be monolithically combined with lasers, detectors, and optical amplifiers. The combination of both passive and active devices in a multicomponent circuit is referred to as an integrated optic circuit (IOC) or a photonic integrated circuit (PIC). In semiconductor materials, purely electronic devices can be integrated as well to form what is often referred to as an optoelectronic integrated circuit (OEIC).

Progress in the field of integrated optics has been rapid since its inception in 1970. Much of this progress is due to the availability of high-quality materials, microelectronic processing equipment and techniques, and the overall rapid advancement and deployment of fiber optic systems. The interest in integrated optics is due to its numerous advantages over other optical technologies. These include large electrical bandwidth, low power consumption, small size and weight, and improved reliability. Integrated optics devices also interface efficiently with optical fibers, and can reduce cost in complex circuits by eliminating the need for separate, individual packaging of each circuit element.

The original inspiration of integrated optics came from the technology of electronic integrated circuits, which has shown rapid development over several decades and has led to amazing achievements, such as complex and powerful microprocessors containing many millions of transistors, specialized signal processors and computer memory chips with huge data storage capacity. Unfortunately, integrated optics has not been able to match the progress of microelectronics in terms of the complexity of possible devices. This results from a number of technical limitations:

• While electronic circuits can contain extremely small wires, optical components need to be connected via waveguides, the dimensions of which usually cannot be much smaller than the wavelength, and which often cannot tolerate very sharp bends. (This limitation might be eliminated by using waveguides with very high index contrast, e.g. nanofibers or photonic bandgap waveguides.)

• Optical connections, e.g. between waveguides, and couplers are significantly more critical than electrical connections.

• Waveguides, device connections and passive optical components exhibit optical losses, which often need to be compensated with optical amplifiers. These are larger and more complex than electronic amplifiers based on transistors.

- Some types of optical components can hardly be miniaturized.

For these reasons, integrated optical circuits have not reached by far the complexity of electronic integrated circuits.

Nevertheless the applications for integrated optics are widespread. Generally these applications involve interfacing with single-mode fiber optic systems. Primary uses are in digital and analog communications, sensors (especially fiber optic gyroscopes), signal processing, and instrumentation. To a lesser extent, IO devices are being explored in nonfiber systems for laser beam control and optical signal processing and computing. IOCs are viewed in the marketplace as a key enabling technology for high-speed digital telecommunications, CATV (Cable television) signal distribution, and gyros. As such they will have a significant impact on commercially deployed fiber systems and devices in the years to come.

Welcome to the wonderful world of integrated optics – miniature planar elements and circuits for information technology.

1. RADIATION IN MATTER

The fundamental processes in integrated optics are due to the laws of radiation propagation in a solid and its interaction with matter, accompanied by changes in the quantitative and qualitative characteristics of radiation, phenomena of reflection, absorption, refraction, interference and diffraction. The mathematical description of this interaction is Maxwell's equations, particular solutions of which determine the ultimate characteristics and parameters of waveguide elements in integrated optics. Maxwell's equations relate the radiation characteristics to the material parameters of the medium, which allows one to obtain calculated relationships for all the main components of integrated optics: waveguides, lasers, led's, switches and photodetectors, etc.

The theoretical aspects of the processes of propagation and interaction of electromagnetic radiation with matter are covered in sufficient detail in the literature, therefore this chapter discusses only the basic provisions necessary for understanding the subsequent material.

1.1. Maxwell's equations and the wave equation

The range of electromagnetic oscillations in integrated optics, including the visible part of the spectrum, is limited by a quantum energy of $3 \leq h\nu \leq 0.6$ eV. Short-wave radiation is characterized by the transverse and two-dimensionality of electromagnetic waves, experimentally arising from polarization phenomena, as well as the reality of representing a stream of electromagnetic radiation in the form of a beam of rays directed along the wave propagation lines. In this case, the light beam is a series of plane waves limited from the sides and propagating along one direction.

$$U = U_1 \exp\left(i\mathbf{kr} - i\omega t\right) + U_2 \exp\left(-i\mathbf{kr} - i\omega t\right), \qquad (1.1)$$

where **r** is the radius vector of the considered point in space; ω is the angular frequency; **k** is the wave vector associated with the material parameters of the medium by the dispersion relation $\mathbf{k}^2 = \omega^2 / v^2 = \omega^2 \varepsilon \mu$. Here ε and μ are the permittivity and permeability, respectively.

In integrated optics, these features of the optical radiation are decisive for microwave technology, microminiaturization levels, and integration of integrated optical circuits. Function (1.1) describes plane waves with amplitudes U_1, U_2, propagating along the direction **r**, the propagation constant of which is generally a complex quantity. Such a plane wave is inhomogeneous. It corresponds to the propagation of radiation in media with a finite value of conductivity σ, as, for example, for semiconductors. Considering the propagation of an inhomogeneous plane wave along the positive coordinate **r**, for the electric field component from (1.1) it follows

$$E_x = E_0 \exp[i\omega(t - \beta z / t)], E_x = E_0 \exp(-\omega k^* z / c) \exp[i\omega(t - nz / c)],$$

where $\beta = n - ik^*$ is the propagation constant;

$$n = \sqrt{0.5\varepsilon \left(1 + \frac{\sigma^2}{\omega^2 \varepsilon_0 \varepsilon^2}\right)^{\frac{1}{2}} + \varepsilon} \quad \text{is the refractive index;}$$

$$k^* = \sqrt{0.5\varepsilon \left(1 + \frac{\sigma^2}{\omega^2 \varepsilon_0 \varepsilon^2}\right)^{\frac{1}{2}} - \varepsilon} \quad \text{is the excision coefficient, or absorption.}$$

The absorption coefficient is determined from the condition that the radiation energy in a substance, given by the square of the amplitude of the electric component $|\mathbf{E}|^2$, or its intensity at a distance of $1 / \alpha$ from the interface,

decreases by a factor of $E^2 \sim \exp(-\omega k^* z / c)$, whence, $\alpha = \dfrac{2\omega k^*}{c} = \dfrac{4\pi k^*}{\lambda_0}$ where

λ_0 is the radiation wavelength in vacuum. In this case, the specific losses in the material are determined by the value $(1/2) \, | \, \mathbf{E} \, |^2 \, \sigma$.

The frequency dependence of the refractive index determines the dispersion of the optical medium. If an optical signal passes through such a medium, then its various spectral components will have different propagation velocities, which leads to their displacement relative to each other over a certain path length. In this case, a signal is distorted, the spectrum of which is limited by a finite value $\Delta\lambda$. All solids and liquids have a greater or lesser dispersion, which is due to the finite amount of radiation absorption in them. In the medium transparency range, the group velocity of the electromagnetic wave packet is

$v_{th} = \dfrac{c}{n + \omega dn / d\omega}$, the delay of the energy packet with respect to the phase of the wave

$$\tau_{th} = \frac{1}{v_{th}} = \frac{n + \omega dn / d\omega}{c}$$

For solids and liquids in the frequency range of transparency

$$\frac{dn}{d\omega} = \omega_0 \omega_d \omega_{ae} (\omega_0^2 - \omega^2) \tag{1.2}$$

where ω_d is the dispersion frequency; ω_0 is the frequency of the oscillator; ω_{ae} is the frequency corresponding to the absorption edge. From relation (1.2) it follows that the less ω_d and more ω_0, the smaller the dispersion $dn / d\omega$. The order of $\omega_0 / 2\pi$ and $\omega_d / 2\pi$ are several thousand terahertz.

In dielectrics and semiconductors for which σ tends to zero at frequency ω, $k^* \to 0$, $\alpha \to 0$, $n^2 \to \varepsilon$. At this frequency, a homogeneous and isotropic material is transparent. This determines the application for dielectric waveguides

of the optical range, the losses in which are minimized. In the case of transparent media for which ε and μ are real quantities, the wave vector $\mathbf{k} = k\mathbf{n}$, where $k = 2\pi / \lambda$ is the wave number; $\mathbf{n}$ is the unit vector in the direction of propagation of the wave; λ is the wave length in the propagation medium. A plane homogeneous wave in a medium with σ = 0 propagates isotropically to infinity without changing the amplitude and phase relationships. In waveguide structures, radiation is always limited by their format.

In order for radiation to propagate with minimal losses along such a waveguide, for example along the r axis, the electromagnetic wave must undergo a total internal reflection from its walls. The phenomenon of total internal reflection underlies the operation of all integrated optical waveguide components.

1.2. Light propagation (reflection, refraction and transmission) through homogeneous media

Denote all quantities characterizing the incident wave by the index i, the reflected one by the index r, and the refracted by t. Let the plane of separation of media i and t coincide with the plane $z = 0$, and the direction of wave propagation lies in the xz plane. From the boundary conditions for the electromagnetic field vectors at the interface

$$(\mathbf{k}^i \mathbf{r})_{z=0} = (\mathbf{k}^r \mathbf{r})_{z=0} = (\mathbf{k}^t \mathbf{r})_{z=0}$$

or   (1.3)

$$k^i{}_x = k^r{}_x = k^t{}_x, \quad k^i{}_y = k^r{}_y = k^t{}_y.$$

From (1.3), the well-known laws of reflection and refraction follow $k^i \sin i_0 = k^r \sin r_0 = k^t \sin t_0$ where i_0, r_0, t_0 are the angles of incidence, reflection and refraction, respectively.

For homogeneous and anisotropic media, the angles of incidence and refraction obey Snell's law $\dfrac{\sin i_0}{\sin t_0} = \dfrac{n}{n'}$, n, n' are the refractive indices of the medium for $z > 0$ and for $z < 0$, respectively.

The polarization and energy characteristics of reflected and refracted waves are determined by the Fresnel formulas:

$$\boldsymbol{E}^t = \frac{2n\cos i_0}{n\cos i_0 + \Pi^*}\boldsymbol{E}^i, \quad \boldsymbol{H}^t = \frac{2n'^2\cos i_0}{2n'^2\cos i_0 + n\Pi^*}\boldsymbol{H}^i$$

$$(1.4)$$

$$\boldsymbol{E}^r = \frac{n\cos i_0 - \Pi^*}{n\cos i_0 + \Pi^*}\boldsymbol{E}^i, \quad \boldsymbol{H}^r = \frac{2n'^2\cos i_0 - n\Pi^*}{2n'^2\cos i_0 + \Pi^*}\boldsymbol{H}^i$$

where $\Pi^* = \left(2n'^2 - n^2\sin^2 i_0\right)$.

From Eqn. (1.4) it follows that when radiation passes from a denser medium to a less dense ($n' < n$), the reflection coefficient tends to unity when

$$\frac{n}{n'\sin i_0} \geq 0 \tag{1.5}$$

Inequality (1.5) corresponds to the total internal reflection of the electromagnetic wave from the interface for all angles $i_0 > i_{cr} = \arcsin\dfrac{n}{n'}$. In this case, the refraction angle t_0 becomes complex, and

$$\cos t_0 = \pm i\sqrt{\frac{n^2}{(n')^2}\sin^2 i_0 - 1}$$

14

is imaginary. This leads to the appearance in the expression for the refracted wave of a decaying factor and a phase jump. A refracted wave is a surface wave that attenuates along the z axis and propagates along the x axis.

$$\boldsymbol{E}^t = \boldsymbol{E}^t{}_0 \exp\left[\omega\sqrt{\varepsilon'\mu}\left(ix\frac{n}{n'}\sin i_0\right)\right] + z\sqrt{\frac{n^2}{(n')^2}\sin^2 i_0 - 1} - i\omega t.$$

In such an inhomogeneous wave, the planes of equal phases are represented by the planes $x = $ const, and the planes of equal amplitudes are parallel to the plane of the boundary. Full internal reflection, as follows from the Fresnel formulas, is realized for both polarizations.

The reflection coefficient $R = \dfrac{|E_r|^2}{|E_i|^2}$ in general is a complex value. Its real component

$$R = \left|\frac{n - ik^* - 1}{n - ik^* + 1}\right|^2 = \frac{(n-1)^2 + k^{*2}}{(n+1)^2 + k^{*2}}$$

Small k^* are inherent in dielectrics and wide bandgap semiconductors. For the visible and IR spectrum for dielectrics and semiconductors, $n^2 \gg k^{*2}$ and $n^2 \gg \sin^2 i_0$. With increasing k^*, the substance behaves like a metal and $R \to 0$. This allows the metallization of waveguides in integrated optics to improve their waveguide properties when using media with low refractive indices.

1.3. Radiation in anisotropic media

Anisotropic media are widely used in optical waveguides and integrated optical elements to create various types of optical signal converters. For monochromatic waves in anisotropic media, the plane wave equations and the

coupling equation $\mathbf{D} = \hat{\varepsilon}\,\mathbf{E}$, $\mathbf{B} = \hat{\mu}\,\mathbf{H}$ where $\hat{\varepsilon}$, $\hat{\mu}$, remain valid. $\hat{\varepsilon}$, $\hat{\mu}$ are tensor values. For non-magnetic optical media and non-gyrotropic transparent media

$$\varepsilon = \begin{pmatrix} \varepsilon_1 & 0 & 0 \\ 0 & \varepsilon_2 & 0 \\ 0 & 0 & \varepsilon_3 \end{pmatrix}, \quad \hat{\mu} \approx \mu = \mu_0 \tag{1.6}$$

Where $\varepsilon_{1,2,3}$ are the main values. In the presence of absorption or amplification of radiation in the medium, the dielectric constant tensor (1.6) becomes complex.

In the system of principal axes of the tensor ε_1, the relationship between the wave number k and the wave vector n is determined by the Fresnel normal equation

$$\frac{n_x^2}{\dfrac{1}{k^2} - \dfrac{1}{k_1^2}} - \frac{n_y^2}{\dfrac{1}{k^2} - \dfrac{1}{k_2^2}} - \frac{n_z^2}{\dfrac{1}{k^2} - \dfrac{1}{k_3^2}} = 0 \tag{1.7}$$

$k_j = \omega\sqrt{\varepsilon_j \mu}$, $j = 1,2,3$ are the main values of the propagation constant; $n_{x,y,z}$ are the direction cosines of the wave normal.

It follows from equation (1.7) that two waves with different phase velocities can propagate along any direction in the crystal.

All anisotropic crystals according to their optical properties are divided into uniaxial and biaxial. Uniaxial crystals have two identical principal values: $\varepsilon_1 = \varepsilon_2 = \varepsilon_3$, $\varepsilon_3 = \varepsilon_e$. In this case

$$n^2 = n_0^2, \tag{1.8}$$

$$\left(n_x^2 + n_y^2\right)\left(\frac{1}{n^2} - \frac{1}{n_e^2}\right) + n_z^2\left(\frac{1}{n^2} + \frac{1}{n_0^2}\right) = 0. \tag{1.9}$$

Equation (1.8) determines the refractive index of ordinary waves, and equation (1.9) determines extraordinary. For ordinary waves, the value of n does not depend on the direction of propagation. For extraordinary – it depends. By introducing the radius vector $r = nn$, Eqn. (1.9) can be reduced to the form

$$\frac{x^2}{n_e^2} + \frac{y^2}{n_e^2} + \frac{z^2}{n_e^2} = 1,$$ which is the equation of a spheroid, or ellipsoid of revolution.

Its surface is the surface of the ends of the vector r. If $n_e < n_0$, then the crystals are called negative, and if $n_e > n_0$ – positive.

For the optical axes of the crystal, the refractive indices of the ordinary and extraordinary waves coincide. Biaxial crystals have different principal values of permittivity and, accordingly, different principal values of refractive indices. In this case, the geometrical location of the points of the ends of the radius vector r forms a more complex than in uniaxial crystals, two-sheeted surface of refractive indices. Only the main sections of this surface have the form of circles and ellipses.

The above equations are fundamental for calculating the characteristics and parameters of integrated optical components, make it possible to evaluate the material classes of waveguide structures based on the specific conditions of the problems being solved.

2. DIELECTRIC MICROWAVEGUIDES

Optical devices are necessary to meet the anticipated future requirements for ultrafast and ultrahigh bandwidth communication and computing. All optical information processing can overcome optoelectronic conversions that limit both the speed and bandwidth and are also power consuming. The building block of an optical device/circuit is the optical waveguide, which enables low-loss light propagation and is thereby used to connect components and devices.

Central to integrated optics is the concept of guiding light in dielectric waveguide structures with dimensions comparable to the wavelength of the guided light.

A dielectric waveguide confines light to the core of the waveguide by somehow reflecting power back towards the waveguide core that would otherwise diffract or propagate away. While any means of reflection can accomplish this end (for example, glancing-incidence partial reflections from interfaces between different media can serve as the basis for leaky waveguides), the most common technique employs a 100 percent total internal reflection from the boundary of a high-index core and a lower-index cladding material. As light propagates down the axis of such a structure , the waveguide cross section can also be viewed as a lens-like phase plate that provides a larger retardation in the core region. Propagation down the guide then resembles a continuous refocusing of light that would otherwise diffract away.

The optical waveguide is the fundamental element that interconnects the various devices of an optical integrated circuit, just as a metallic strip does in an electrical integrated circuit. However, unlike electrical current that flows through a metal strip according to Ohm's law, optical waves travel in the waveguide in distinct optical modes. A mode, in this sense, is a spatial distribution of optical energy in one or more dimensions that remains constant in time.

In integrated optics, the basis of any device is a planar dielectric waveguide and various types of three-dimensional channel and strip waveguides. In slab waveguides, light is confined in only one dimension (Fig. 2.1). In most actual waveguides light is confined in two dimensions.
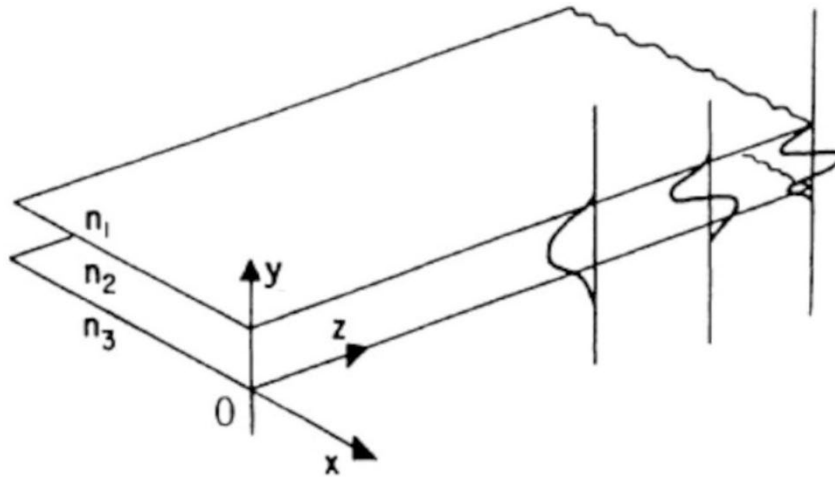
Fig. 2.1. Diagram of the basic three-layer planar waveguide structure. Three modes are shown, representing distributions of electric field in the *x* direction

Three-dimensional waveguides, unlike planar waveguides, provide an additional confiment of optical radiation in the transverse direction.

The limitation of the light flux in the transverse direction allows one to reduce the control voltages of the waveguide modulators and the operating powers of injection lasers, which are part of the structure of integrated optical circuits.

The basis of optical waveguides is thin dielectric layers with low losses at the working wavelengths. The format of such films is of the same order with the wavelength, which leads to high requirements for the technology of their manufacture, both in resolution and in the accuracy of their reproduction. Typical thicknesses of dielectric waveguides are in the range of $0.3…3$ µm with a width of $1…5$ µm. The effective waveguides requires that the standard deviation from linear dimensions does not exceed 0.5 nm. The technology should also provide low optical losses, which should not exceed units and tenths of decibels per centimeter.

Currently, the foundations of the theory and technology of thin-film waveguides are developed, which allow to calculate and fabricate waveguide structures with desired properties.

The list of materials suitable for fabrication waveguide structures includes various glasses, including chalcogenide, used for the medium and long-wave IR spectra, various polymers, ferrites, active dielectrics, electro-optical crystals, ceramics and semiconductors.

2.1. Dielectric waveguide modes

The mode of a dielectric waveguide at a frequency ω is the solution of the wave equation

$$\Delta^2 E(r) + k_0^2 n^2(r) E(r) = 0 \tag{2.1}$$

as

$$E(\mathbf{r},t) = E(x, y) \exp\left[i(\omega t - \beta z)\right] \tag{2.2}$$

when the continuity condition for the tangential components $\mathbf{E}$ and $\mathbf{H}$ is satisfied at the dielectric interface, where $k_0 = 2\pi / \lambda_0$ is the wave number; λ_0 is the wavelength in vacuum; n is the refractive index of the waveguide; $\beta = \omega / v_{ph}$ is the propagation constant or phase constant; ω, v_{ph} are the angular frequency and phase velocity of the wave, respectively.

The main properties of a dielectric waveguide can be determined using a model of a flat plate, infinite in one direction (y or x). Three-dimensional waveguides, the size of which is limited both in the x and y direction, approximately satisfy the planar waveguide model if one of the dimensions is much larger than the other. Let us consider the waveguide shown in Fig. 2.1.

Substituting (2.2) into equation (2.1) and assuming d / dy = 0, for regions I, II, and III we obtain

$$\frac{d}{dx^2}E(x,y) + \left(k_0^2 n_i^2 - \beta^2\right)E(x,y) = 0, \ i = 1,2,3, \qquad (2.3)$$

where $E(x,y)$ is the Cartesian component of $\mathbf{E}(x,y)$. For a given frequency ω, the form of the solution depends on the propagation constant β. Under the inequalities $k_0 n_3 < \beta < k_0 n_2$ and $\dfrac{1}{E}\dfrac{d^2 E}{dx^2} < 0$, solution (2.3) in region II is a harmonic function (Fig. 2.2, d,c). In regions I and III, the field decays exponentially. The obtained solutions $E(x)$ satisfy the given boundary conditions. The energy carried by these types of oscillations, called waveguide modes, is enclosed near the waveguide layer. A necessary condition for the existence of waveguide modes is the fulfillment of the inequality $k_0 n_1$ and $k_0 n_3 < \beta < k_0 n_2$ or $n_2 > n_1, n_3$, whence it follows that the inner layer of the waveguide structure should have a higher refractive index than the outer layers of regions I and III. For $k_0 n_1 < \beta < k_0 n_3$, solution (2.3) is exponential in region I and harmonic in regions II, III (Fig. 2.2, d), which corresponds to substrate radiative modes. If the inequalities $0 < \beta < k_0 n_1$ are satisfied, the solution for $E(x)$ becomes sinusoidal in all three regions (Fig. 2.2, e). This case corresponds to the radiative modes of the waveguide. When, the function $E(x)$ must be exponential in all three regions and only the mode shape shown as (a) in Fig. 2.2. could satisfy the boundary conditions at the interfaces. This mode is not physically realizable because the field increases unboundedly in regions I, III, implying infinite energy.

An analysis of the solutions of Eqs. (2.3) satisfying the boundary conditions at the interface shows that the spectrum β is discrete for the waveguide modes, and the continuum β belongs to radiative modes. The number of waveguide modes depends on the thickness of the waveguide d, the frequency ω, and the refractive indices of the media forming the waveguide. For a given ω, the number of waveguide modes increases with increasing d.
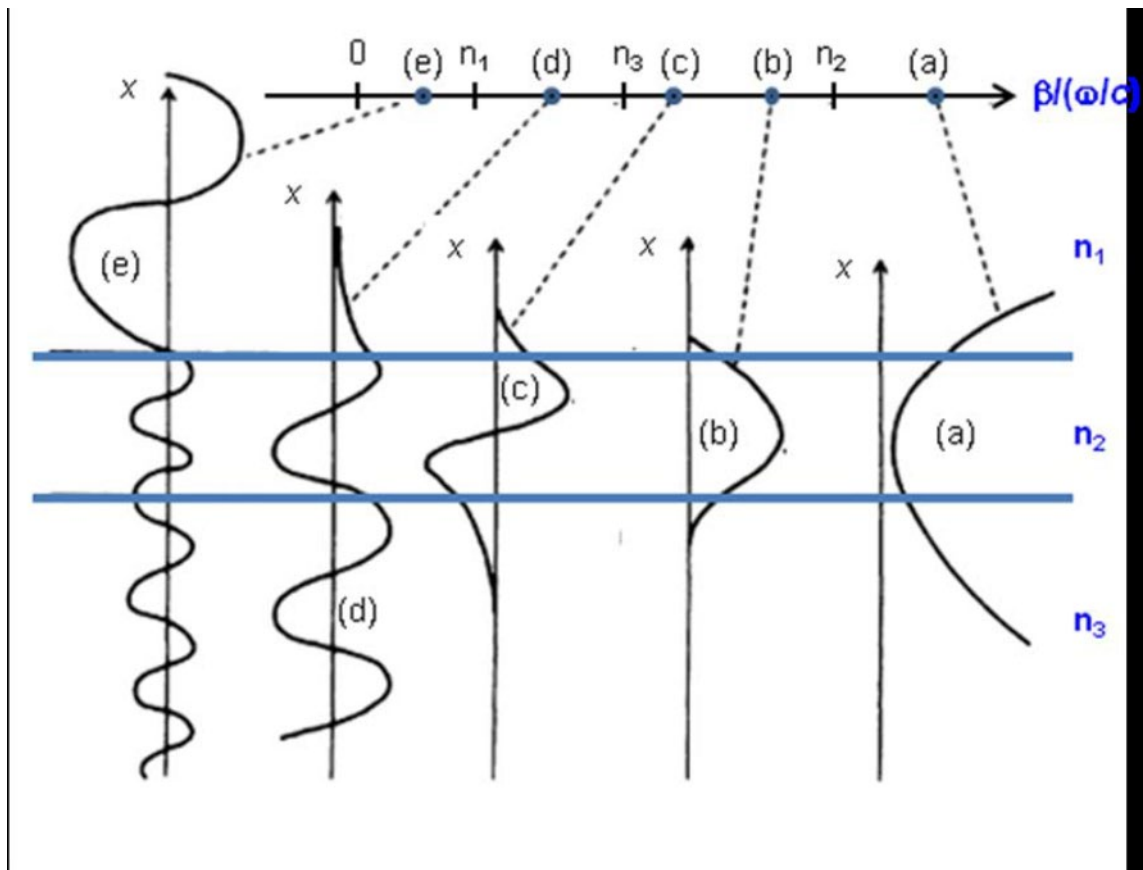
Fig. 2.2. Diagram of the possible modes in a planar waveguide

Depending on whether the component of the field, electric or magnetic, is perpendicular to the direction of wave propagation, transverse electrical TE and magnetic TM modes are distinguished. If the modes propagate along the z axis (Fig. 2.2) with the coordinate dependence of the amplitude proportional to $\exp{(-i\beta z)}$, then the TE mode will contain only E_y, H_x, H_z, and the TM mode H_x, E_y, E_z components of the electromagnetic field. In the first case, the electric field of the wave is bounded in the transverse plane of the waveguide, and the magnetic field in the second. For waveguide structures, this determines the various modes of their operation. Waveguide modes are the main working modes of integrated optical devices.

2.2. Planar optical waveguide

A planar optical waveguide consists of two layers with low values of the refractive index and a waveguide region located between them with a high value of the refractive index.

There are two main types of optical waveguide structures: the step index and the graded index. A waveguide in which the index profile changes abruptly between the core and the cladding is called a step-index waveguide, while one in which the index profile varies gradually is called a graded-index waveguide as shown in Fig. 2.3. Recently, hybrid index profile waveguide was shown combining both inverse-step index waveguide and graded index waveguides for high-power amplification of a Gaussian single-mode beam.
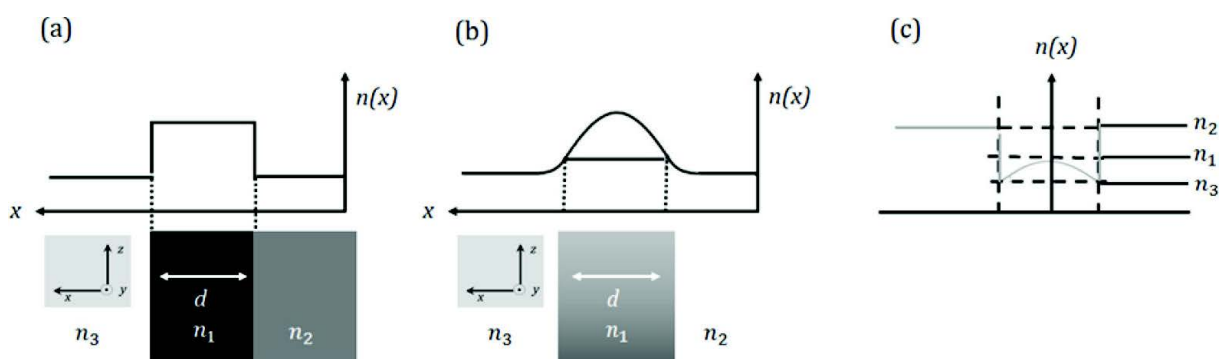
Fig. 2.3. (*a*) Step-index type waveguide, (*b*) Graded-index waveguide, and (*c*) Hybrid waveguide

In step index waveguides, the refractive index of the waveguide layer is constant. To implement the waveguide regime in it, it is necessary to fulfill condition $n_2 > n_1, n_3$.

When describing the characteristics and parameters of dielectric waveguides, it is convenient to use concept of the effective refractive index.

For plane waves in homogeneous transparent media, the refractive index n_{ef} can be used to quantify the increase in the wavenumber (phase change per unit length) caused by the medium: the wavenumber is n times higher than it

would be in vacuum. The effective refractive index n_{eff} has the analogous meaning for light propagation in a waveguide with restricted transverse extension: the phase constant β of the waveguide (for some wavelength) is the effective index times the vacuum wavenumber:

$$n_{eff} = \beta / k_0.$$

The mode-dependent and frequency-dependent β can be calculated with a mode solver software and depend on the refractive index profile of the waveguide.

Note that the effective refractive index depends not only on the wavelength but also (for multimode waveguides) on the mode in which the light propagates. For this reason, it is also called modal index. Obviously, the effective index is not just a material property, but depends on the whole waveguide design.

Figure 2.4 shows graphical solutions of the dispersion equation (2.3), which defines the propagation constant as a function of the frequency ω and the thickness of the waveguide d for the main modes of various orders of m. The normalized frequency $V^* = k_0 d \sqrt{n_2^2 - n_3^2}$ and the normalized refractive $b = \left\{ n_{eff}^2 - n_3^2 \right\} \left(n_2^2 - n_3^2 \right)^{-1}$. The parameter of the curves is the degree of asymmetry of the waveguide structure $\alpha^* = \left(n_3^2 - n_1^2 \right) \left(n_2^2 - n_3^2 \right)$. For a symmetric step index waveguide $n_3 = n_1$ and $\alpha^* = 0$. For waveguides with a pronounced asymmetry $n_3 \neq n_1$, $n_3 \to n_2$, $\alpha^* \to \infty$.
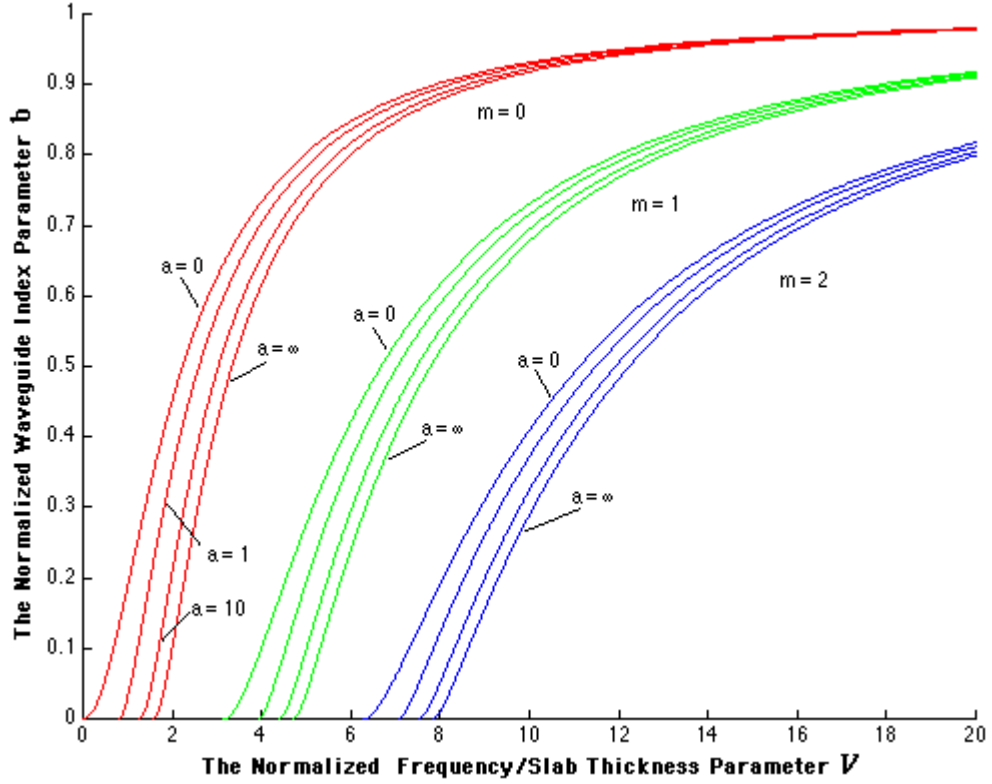
Fig. 2.4. Normalized diagram for a planar step index waveguide at various degree of its asymmetry

The maximum number of modes propagating along the waveguide is limited by the relation $M = \dfrac{1}{2} + \left(\dfrac{2d}{\lambda_0} \right) (2n_3 A_1)^{1/2}$, where $A_1 = n_2 - n_1$. In this case, the main mode of a symmetric step index waveguide has no cutoffs. For such a waveguide, the electromagnetic field within the waveguide layer changes according to a sinusoidal law, and outside it decays exponentially (see Fig. 2.5, *a*).

In a gradient waveguide, the interface between the substrate and the waveguide region is not pronounced and is determined by the profile of the refractive index specified by the relations:

$$n(x) = n_3 + \Delta n(x), \quad x > 0; \quad n(x) = n_1, \quad x < 0,$$

where x is the coordinate along which the gradient n is specified. The function of $\Delta n(x)$ depends on the method of manufacturing the waveguide. In most cases, with a sufficient degree of approximation, it can be approximated by a dependence of the form $\Delta n(x) = A_1 \exp\left(1 - \dfrac{x}{d_x}\right)$, where d_x is the characteristic thickness of the waveguide region. The number of modes propagating along such a waveguide is limited by the inequality $M \le \dfrac{1}{4} + \left(\dfrac{4d_1}{\lambda_0}\right)(2n_3 A_1)^{\frac{1}{2}}$. The propagation of the electric field in the range $0 < x < x_j$, where x_j is the mode turning point, corresponds to a periodic curve. Beyond these boundaries, it decays exponentially (Fig. 2.5, a).
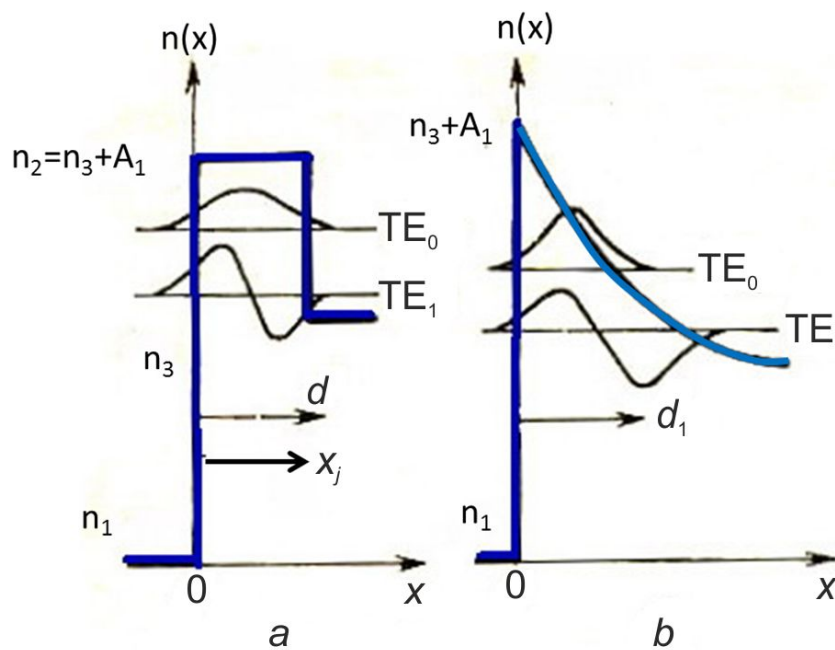
Fig. 2.5. Refractive index profile and electromagnetic field distribution of TE_0 and TE_1 modes in planar step-index (a) and gradient index (b) waveguides

In planar waveguides, the restriction of the light channel has it in only one plane. In integrated optics, this complicates the coupling/matching with other components and optical fibers, and reduces the efficiency of waveguide devices. Significantly large degrees of freedom are characterized by three-dimensional

waveguides, which are mainly used in integrated optics and integrated optoelectronic devices.

2.3. Three-dimensional waveguides

The main types of three-dimensional strip and channel waveguides are schematically shown in Fig. 2.6. The division of three-dimensional waveguides into strip and channel is somewhat arbitrary, nevertheless it reflects the structural features that affect the possible number of waveguide modes, the specifics of the field distribution and, therefore, their main characteristics. In channel waveguides of comb type (Fig. 2.6, *a*, *b*), a significant jump Δn in the transverse direction leads to their multimode mode. Submerged channel structures (Fig. 2.6, *c*, *d*) are characterized by intermediate properties compared with comb channel and strip waveguides. The number of propagating modes in them is determined by the channel size and the difference in the refractive indices of the materials forming the waveguide. A strip waveguide, or a strip-guided waveguide (Fig. 2.6, $f - i$), is a planar thin-film or diffusion waveguide with a narrow dielectric region deposited on its surface. In such waveguides, almost all light is localized under the strip of the dielectric and hardly penetrates into it. As a result, the roughness of the interface between the strip and the waveguide does not lead to significant radiation losses. Single-mode or low-mode operation is easily realized in strip waveguides.

The most promising for integrated optics are various types of strip and buried channel waveguides. Their main advantages are associated with wide technological and constructive capabilities for creating effective waveguide structures based on them. The advantages of strip structures include weak sensitivity to violations of the boundaries of the dielectric strip and an increase in the tolerance for its format in single-mode operation. In integrated optical circuits, its width can reach ten or more micrometers, which greatly simplifies the operation of coupling waveguides.
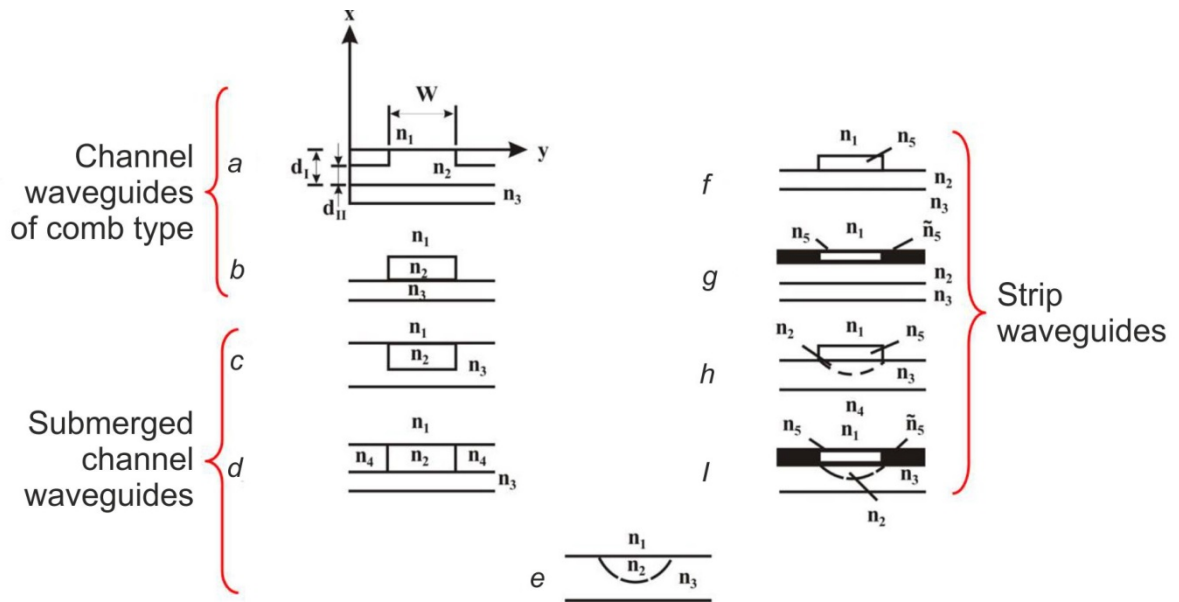
Fig. 2.6. The main types of three-dimensional strip (*f-i*) and channel (*a-d*)

waveguides $\tilde{n}_5 = n_5' + in_5''$ (rib embedded/buried)

For channel and strip structures, all waveguide characteristics, including dispersion ones, depend not only on the shape and relative dimensions of the waveguide, but also on the electrodynamic parameters of the materials that form the waveguide. Under these conditions, various approximate methods of calculation and analysis have become widespread. One of such approximations is the effective refractive index method, which has been successfully used to calculate the dispersion characteristics and mode spectrum of the main types of three-dimensional waveguides. This method reduce the calculation of the propagation constants of a three-dimensional waveguide to an analysis of the equivalent symmetric planar waveguide to which the original waveguide structure comes down. In most cases, they are limited to the numerical solution of comparing the one-dimensional wave equation when approximating the real profile along each transverse coordinate with piecewise continuous functions for which the wave equation has known solutions. The resulting solutions are then stitched at the corresponding points in the profile. The resulting distributions of the mode fields are represented as the product of the distributions along each transverse coordinate.

In integrated optical circuits, waveguides are communication lines. The input and output of radiation from them requires optical coupling between transmitting media, including waveguides, light sources, optical fibers, photodetectors, etc. The coupling efficiency is determined by the ratio of the power into the receiving structure to the total power emitted by the exciting structure. For multimode structures and uniform distribution of optical power over the modes of the exciting structure, the maximum matching efficiency is equal to the ratio of the number of modes of the receiving and exciting structures. When matching single-mode waveguides, the matching efficiency is limited by the degree of mismatch in the distribution of the fields of the waveguide modes.

To connect different planar and three-dimensional waveguide structures various devices have been developed that are based on the use of either direct or distributed coupling of surface waves. To connect planar waveguides, coupling devices based on smooth waveguide transitions, various types of diffraction-grating structures, and end connections of waveguides are most widely used. Three-dimensional waveguides of various types and with different cross-sectional sizes can be matched by smooth waveguide transitions from one section to another. As a rule, waveguide transitions are a narrowing of only one transverse dimension of the waveguide in the plane of the circuit. For the coupling of three-dimensional and planar waveguides, horn structures such as a smooth transition, which are used as intermediate matching elements when connecting channel waveguides with fibers, have found the greatest practical application.

2.4. Photonic crystal waveguide

Photonic crystals (PCs) are periodical optical micro- and nanostructures that receive increasing attention due to their ability to manipulate light propagation while maintaining high transmission efficiency. They are typically

designed based on their energy band gap to prohibit certain wavelengths. They can be arranged in one dimension (1D), two dimensions (2D), and three dimensions (3D) depending on guiding requirements. The 2D PC structure is the most popular choice for optical devices and has found various applications such as integrated circuits (ICs), wave filters or multiplexers, and waveguides.

The PC's excellent transmission properties and physical flexibility are attractive to minuscule waveguides fabrication. It can be bent to fit in a limited space and still achieve desirable performance (Fig. 2.7). A significant enhancement in operating bandwidth has also been reported.

a

b
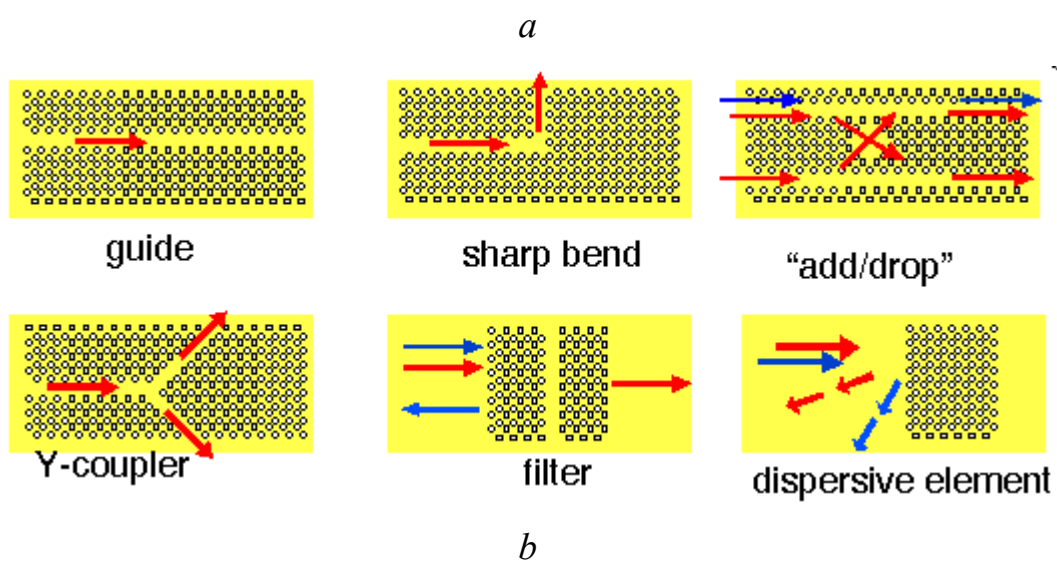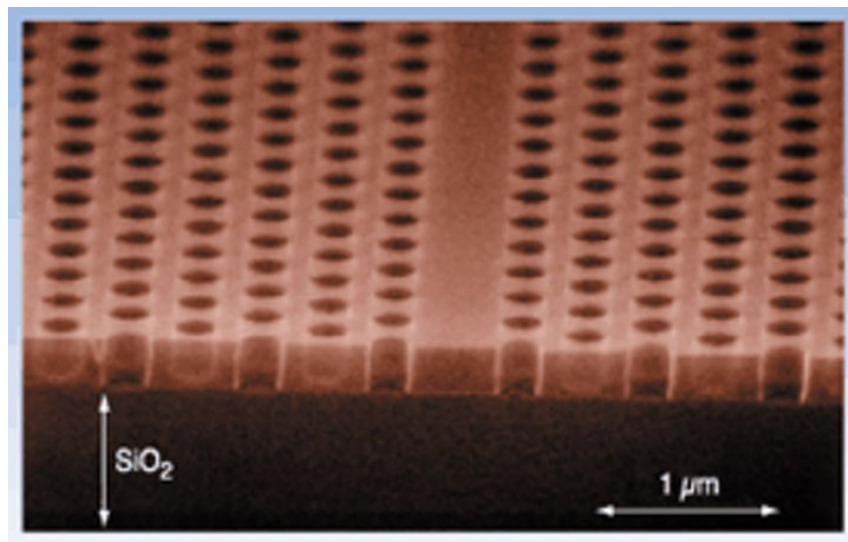
Fig. 2.7. Photonic crystal waveguides: *a* – SEM micrographs of photonic crystal waveguide pattern; *b* – possible realisations of photonic crystal waveguides in a various I/O elements

Photonic crystal waveguides guiding mechanism is different from that of a traditional waveguide, which is based on internal reflection. A photonic crystal is a periodic dielectric structure with a photonic band gap, that is, a frequency range over which there is no propagation of light.

The introduction of line defects into a photonic crystal structure creates an optical channel for propagation of light. If the line defect is properly designed, the resulting guiding mode falls within a photonic band gap, is highly confined, and can be used for guiding light. The guiding mode can also be designed to be broadband and thus gives rise to a compact, broadband photonic crystal waveguide. Application of these waveguides includes nanofluidic tuning, RI measurements, optical characterization of molecule orientation, and biosensing.

2.5. Waveguides on glass substrates

For the passive elements waveguides are usually formed on glass substrates. This is due, in particular, to the relative simplicity of manufacturing optical elements in such waveguides and their coupling with optical fibers. The advantages of glasses in this case are low cost, proven technology for the fabrication and processing, the ability to manufacture substrates of a given geometry and size.

One of the common methods for obtaining waveguide structures in glass is ion exchange. The ion exchange technology allows to obtain both planar and channel inhomogeneous waveguides with low losses. When a glass substrate is immersed in a molten salt containing heavy metal ions, glass ions are replaced with heavier ions from the molten salt. As a result, the refractive index of the surface layer of the glass increases compared to the refractive index of the glass in the bulk of the substrate (Fig. 2.8). By carrying out ion-exchange reactions in an electric field, it is possible to control the sequence of introduction of ions with different electronic polarizability (Fig. 2.9). Under the influence of an electric field, at the first stage, the introduction of ions with higher electron

polarizability, for example, thallium ions, takes place. In this case, a layer with a high refractive index is formed at the surface of the glass substrate. At the second stage, ions with low electron polarizability are introduced in the same way.
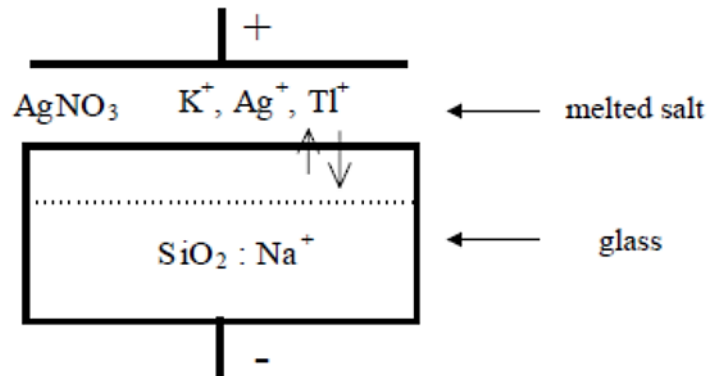
Fig. 2.8. Ion exchange - principle of the effect
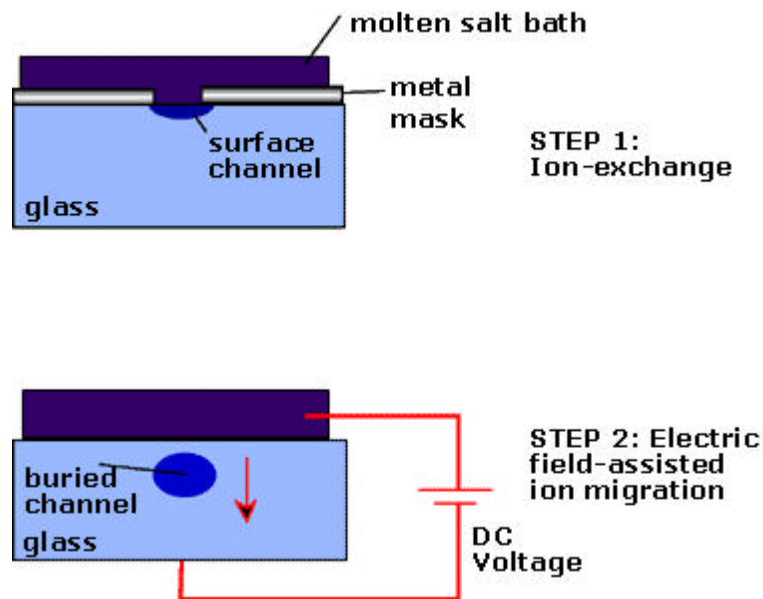
Fig. 2.9. A cross-sectional view showing how a buried round waveguide is created

These are usually sodium or potassium ions, which restore a relatively low refractive index of the glass substrate. As a result, the waveguide layer formed in the first stage appears to be overgrown in the substrate, which reduces the optical losses in the waveguide due to scattering at the substrate – air interface. By changing the conditions for the introduction of ions and their concentration in the salt, it is possible to control the Δn value at the interface

between the waveguide layer and the growing layer within the order of magnitude from 10^{-4} to 10^{-1}. Using this method, gradient waveguide structures have been fabricated, the coupling of which with optical fibers is greatly simplified due to the programmable shape of the waveguides.

The profile of the refractive index in ion-exchange waveguides is largely determined by the process conditions: temperature, melt composition, and the presence of stimulating electric fields. The most widely used processes of ion exchange from melts of nitrates Ag, K, Tl optimum temperatures are 250…365 °C). An increase in the temperature of the ion exchange reaction beyond the optimal for the salts of nitrates leads to blackening or etching of the surface layer, which affects the optical parameters of the waveguide layer.

The waveguides formed in glasses due to the diffusion of metal ions from melts of silver, thallium nitrates and mixtures of silver and sodium nitrates are not without drawbacks. In the first two cases, waveguides with relatively large differences in the refractive index $\Delta n >> 0.1$ are formed. As a consequence of this, lower waveguide modes have small effective penetration depths into the substrate (about 1 μm), which limits the possibility of using such waveguides. In addition, silver and thallium nitrates are an expensive and highly toxic raw material. To obtain large diffusion coefficients and waveguide structures of considerable depth, ion exchange from sulfate melts is used. In this case, the temperature of the exchange process rises to 530 °C. In this mode, it is possible to obtain waveguides with a depth of up to 150 μm and $\Delta n \approx 0.015$ in just 15…20 minutes. Ion-exchange waveguide structures can also be obtained from supersaturated aqueous solutions of nitrates.

When making waveguide structures in glasses, the composition of the substrate glass is essential. Depending on the composition during the heat treatment, in addition to direct, back diffusion can occur in it. In fluorinated glasses subjected to heat treatment, it is possible to obtain waveguide structures with relatively low losses due to back diffusion, it is possible to obtain waveguide structures with relatively small losses.

For the waveguide structures in glasses, along with ion exchange, solid-state diffusion of Ag and Pb can be used. However, it is difficult to obtain sufficiently deep structures ($d \geq 50$ μm) in this way.

For the waveguides in quartz glass, ion implantation methods are widely used. Irradiation with ions generally creates an inhomogeneous surface layer with a complex profile of the refractive index. Ions of H, He, Li, Ar, P, N are used as implants in quartz glass. At doses from 10^{13} to 10^{17} cm^{-2}. The ion energies range from 70 keV to mega electron volts. If saturation doses are achieved, the profile of the refractive index can already be approximated by a rectangular step. Changes in the refractive index of fused silica from the dose of implanted light ions with a good degree of approximation are determined by the empirical relation $\Delta n = n - n_0 \approx 2{,}1 \cdot 10^{-21} C_n$ where n_0 is the value of the refractive index in unirradiated material; C_n is the concentration of impurity ions. The relative change in the refractive index can reach 15%, and the optical loss in the waveguides subjected to post-implantation annealing is reduced to tenths of a decibel per centimeter (about 0.2).

Multimode waveguides with large depth and low losses are created in quartz substrates by the method of (Flame Hydrolsis) paraphase deposition during flame hydrolysis (SiO_2–TiO) waveguides and thermal oxidation (SiO_2–GeO_2) waveguide (Fig. 2.10). Such methods are well established for the production of fiber optical fibers and make it possible to ensure high adaptability of waveguide structures. Owing to the proximity to the composition and format of the optical fibers, the waveguides obtained by such methods can easily and effectively connect with them. Some parameters of technological processes for obtaining waveguide in glass substrates are given in table 2.1.

Along with waveguides in glass substrates in integrated optics, waveguides on various oxides and nitrides obtained by thin-film technology are used, among which are deposition from a molecular beam in a vacuum, sputtering in a gas-discharge plasma, chemical deposition from a gas phase, thermal oxidation, and also anodizing in solutions and discharge plasma. When

deposition of oxide films in vacuum, methods of RF ion-plasma and magnetron sputtering in an oxygen-argon atmosphere are used.
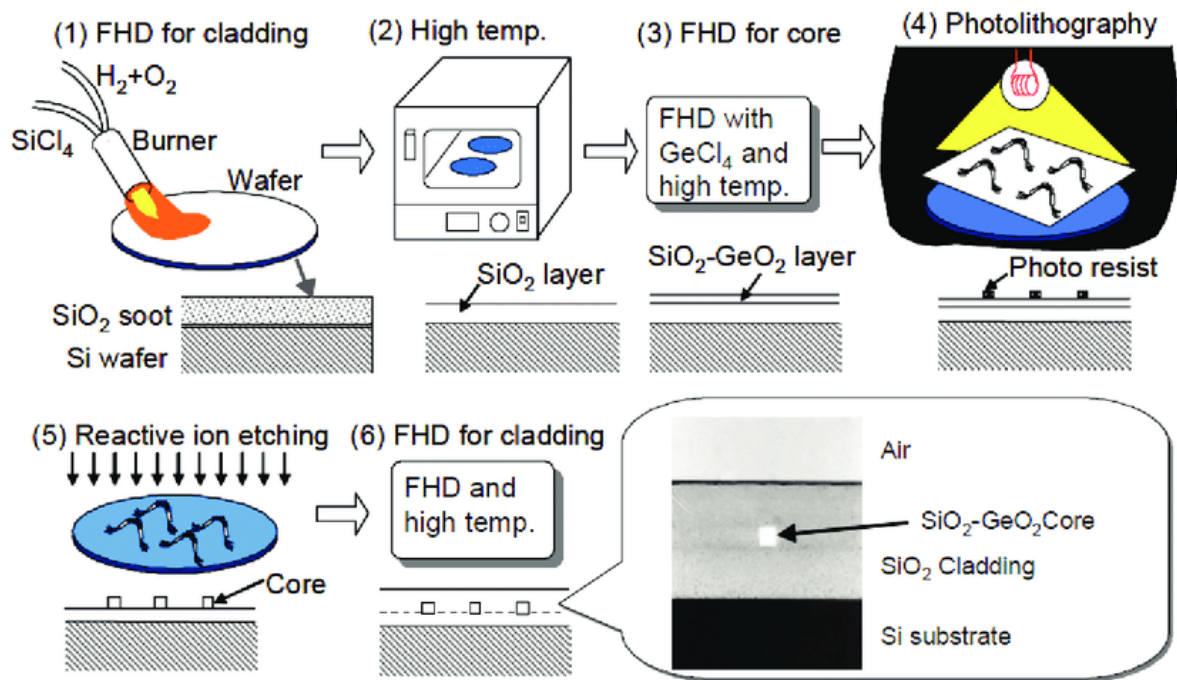
Fig. 2.10. Fabrication process and cross-section of the waveguide

Table 2.1

Parameters of technological processes for obtaining diffusion waveguide structures in glass substrates

Embedded ions	Refractive index increment	Diffusion temperature, °C	Diffusion time, h
Li	0.01	560	0.01
K	0.01	360	15
Rb-Cs	0.006	530	4
Cu	0.04	550	32
Pb	0.08	330	4*(150)
Ag	0.1	300	2
Ag	0.1	300	0.5*(50)
Tl	0.1	260	3.5

Compared to other plasma sputtering systems, magnetron systems provide one to two orders of magnitude higher atomization rates and lower (one to two orders of magnitude) pressures, which reduces the level of gas inclusions in the deposited film. The films thus obtained are about 10 μm homogeneous and are characterized by good adhesion to the substrate. Laser annealing is an effective method for reducing optical losses in oxide and nitride films. In this case, losses in oxide films decrease by an order of magnitude. Subsequent deposition of thin immersion films on the surface to reduce scattering by surface irregularities makes it possible to bring the level of losses in thin-film waveguides to 0.01 dB / cm.

As an example the fundamental planar lightwave circuits PLC fabrication process is shown in Fig. 2.10. First, SiO_2 and GeO_2 glass particles are deposited on a Si substrate by flame hydrolysis deposition (FHD). SiO_2 particles are deposited to form the under-cladding layer and GeO_2-doped SiO_2 particles are deposited for the core layer. Then, the substrate is heated to 1000 °C, which consolidates the glass particles into a transparent glass film. Next, a rectangular core is formed by reactive ion etching. Finally, over-cladding glass is deposited by FHD. The photograph in the bottom right of the figure shows that a well-defined rectangular core is embedded in the cladding glass.

For passive waveguides along with oxide and nitride films, polymer films are widely used. Polymer films, having good waveguide properties, are cheaper and more technologically advanced than oxide and nitride ones. Polymer films for waveguides are prepared by precipitation from solutions, gas phase polymerization, pyrolytic decomposition, photopolymerization. Of the organosilicon compounds, the most promising for waveguides are monomers: vinyltrimethylsilane (VTMS) $CH_2 = CH - Si(CH_3)_3$ and hexamethyldisiloxane (HDMS) $(CH_3)_3 Si - O - Si(CH_3)_3$, the refractive index of which at a film thickness of 1…4 μm at a wavelength of 0.63 μm is 1.527 and 1.488,

respectively. Within these limits, according to the linear law, the refractive index of the mixed monomers $(GDMS)_x$ $(VTMS)_{1-x}$ changes.

For polymer waveguides various photoresists are also used.

Table 2.2 shows the losses in polymer waveguides for various methods for their preparation.

Table 2.2

Characteristics of polymer waveguide layers

Waveguide material	Fabrication method	Refractive index (λ = 0.63 μm)	Average loss, dB/cm (λ = 0.63 μm)
Vinyl methylsilane	Plasma olymerization	1.52—1.53	0.04
Hexamethyldisiloxane	— » —	1.48	0.04
Octamethtrisiloxane	— » —	1.49—1.52	1.0
Octocylol	— » —	1.57	1.0
Toluene	— » —	1.59	1.0
Gelatin	Deposition from water solution	1.54	0.5
Tetraethoxy titanium	Pyrolysis	1.61	0.7
Polymonochloroparaxyls	Pyrolysis	1.66	1.0
Polyurethane	Solution polymerization	1.55	0.8
Epoxy resin	— » —	1.58	0.3
Photoresist	Photopolymerization	1.61	2.0

2.6. Waveguides on the active dielectrics substrate

The most widely used as substrates for the active waveguides are crystals of lithium niobate and tantalate. Based on the waveguides in $LiNbO_3$ and $LiTaO_3$, highly efficient electro-optical (EO) modulators, switches, bistable elements, etc. are created. The main methods for forming waveguide structures in these materials are diffusion of metal ions, reverse diffusion of lithium oxide, ion exchange, implantation and epitaxy.

In comparison with polymers or other ferroelectrics such as SBN (Strontium barium niobate), KTN ($KTa_xNb_{1-x}O_3$) , $BaTiO_3$, which show even higher EO sensitivity than $LiNbO_3$, the material benefits from technological maturity based on Czochralski's process, so that numerous suppliers can be found around the world. Moreover, its physical properties are compatible with cleanroom fabrication processes. For example, its high Curie temperature (~1200 °C) preserves the EO properties even during annealing steps, which is not the case for materials like SBN or KTN. Additionally, $LiNbO_3$ offers a wide transparency band (0.340…4.6 μm) which opens the range to applications from visible to mid-infrared. The low absorption losses (< 0.15 %/cm at 1.06 μm) and its weak optical dispersion in the transparency band also contribute to its success. As a result, lithium niobate has become indispensable for demanding applications such as broadband modulation for long-haul high-bit-rate optical telecommunication systems, electromagnetic sensors, precision gyroscopes, and astronomy. For each application, the challenge is to provide integrated configurations that are easy to implement, of low loss, low in energy consumption, and, if possible, compact. The basic element, namely the optical waveguide, is essential to meet these specifications.

The main methods for forming waveguides in these materials are diffusion of metal ions, reverse diffusion of lithium oxide, ion exchange, implantation and epitaxy.

Titanium diffusion and proton exchange (PE) constitute the two main commercial techniques for the manufacture of LiNbO$_3$-based optical waveguides.

Diffusion methods for the formation of waveguides in LiNbO$_3$ are most widely used, since they make it quite simple to create waveguides with a given mode composition and small (less than 1 dB/cm) losses. As diffusing impurities, Tl, Ni, V, W, Fe, Cr are used. The best results were obtained by diffusion of titanium ions. Today, almost all integrated optical devices on LiNbO$_3$ are performed on Ti-diffuse waveguides. In the LiNbO$_3$ crystal lattice, titanium replaces niobium. Since the ionic radius of titanium is smaller than the ionic radius of niobium, alloying with titanium leads to the appearance of mechanical stresses in the LiNbO$_3$ lattice. The most probable mechanism for the formation of a waveguide layer in it is also associated with this. As a result of the photoelastic effect caused by lattice deformation, an extraordinary and ordinary refractive index of the crystal increases. For the y-cut LiNbO$_3$

$$\Delta n = -\left(\frac{n_0^3}{2}\right)\left(P_{11} + P_{12}\right)\Delta y, \quad \Delta n_e = -n_e^3 P_{31}\Delta y,$$

where P_{11}, P_{12}, P_{31} are components of the photoelastic tensor; Δy is the relative deformation of the lattice proportional to the Ti concentration.

Diffusion of titanium is carried out at temperatures from 900 to 1100 °C (Fig. 2.11). The lower limit is established taking into account the diffusion time and the possibility of the formation of a stable phase from congruent crystals. The upper boundary of the temperature range of diffusion is determined by the Curie temperature (1210 °C) and the formation of lamellar structures at high temperatures. Thin films (several tens of nanometers) of titanium or its oxide sprayed through masks onto the surface of the dielectric are used as a diffusion source. Diffusion in argon atmosphere with subsequent cooling of samples in oxygen is considered optimal.
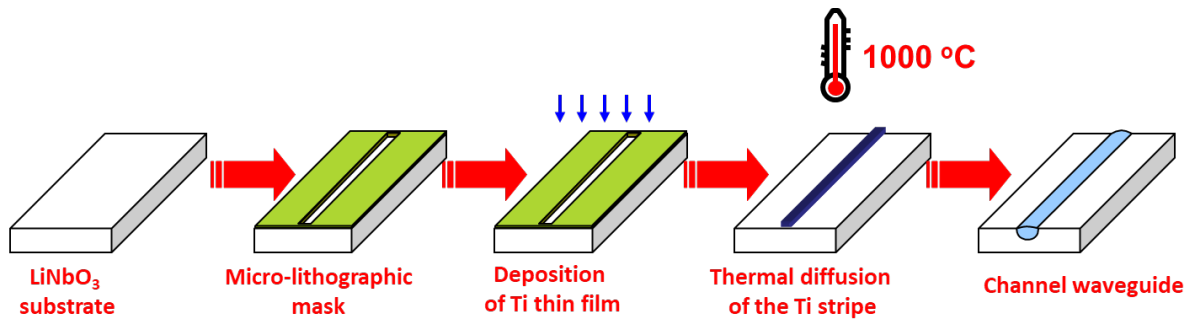
Fig. 2.11. Formation of diffusion waveguides in LiNbO$_3$

When forming active elements (modulators, switches, etc.) on the basis of diffusion waveguides, it is necessary to take measures to compensate for the out diffusion of Li$_2$O. In the case of lithium out diffusion in the process of forming a three-dimensional waveguide, a parasitic planar waveguide simultaneously emerges for an extraordinary wave. The necessary compensation can be obtained by introducing an additional source of lithium oxide vapor into the reactor zone.

For integrated optics, crystals of the LiNbO$_3$ *z*-cut are of great importance. This cut allows the use of a simpler electrodes configuration for modulators and switches and is characterized by Ti high diffusion rate. Advances and technologies for producing Ti: LiNbO$_3$ waveguides allow to obtain not only high-quality waveguides of a simple configuration, but also compact and complex circuits.

Lithium tantalate, which is also the base material of integrated optics, has greater hardness compared to niobate, is easier to polish, and is resistant to optical damage. Obtaining diffusion waveguides in it is carried out according to the same scheme as for lithium niobate. Its distinctive feature is the lower Curie temperature for lithium tantalate (620 °C), which limits the upper substrate heating temperature. At a diffusion temperature exceeding the Curie temperature, the electro-optical properties of the material deteriorates significantly. As a rule, the diffusion time is 5…10 hours. This creates certain difficulties, especially in mass (industrial) production. Recently new

technological solutions that can significantly reduce the process time without compromising the quality of the waveguides have been discovered.

So, for example, if a preheated (up to 300 °C) substrate is heated for 5…30 minutes by radiation from a carbon dioxide laser, the characteristics of the waveguides will be the same as if they were made using conventional thermal diffusion in a furnace for 5 hours.

In addition to thermal diffusion for passive and active devices on lithium niobate substrates proton exchange technology is widely used. In this case, similarly to the process of forming waveguides in glass, lithium niobate crystals are immersed in a benzoic acid melt. During the exchange of hydrogen and lithium ions, a waveguide layer is formed at the surface. The advantages of proton exchange technology include the possibility of obtaining large (~0.1) drops in the refractive index, a high speed of waveguide formation — a single-mode waveguide can be obtained within a few minutes — an increased threshold of optical damage.

In addition to lithium niobate and tantalate, promising materials for integrated optics include the so-called PLZT ceramics (chemical formula $(PbLa)(ZrTi)O_3$). By ceramics we mean a large group of dielectrics with various properties, which are compounded by a common technological cycle. The word ceramics in translation from Greek means "pottery clay". Now ceramics is understood not only as clay-containing, but also other inorganic materials with similar properties. Ceramics makes it possible to obtain predetermined characteristics by changing the composition and production technology. Common to ceramic products is high-temperature annealing. Waveguide structures in ceramics are obtained by high-frequency sputtering on sapphire (Al_2O_3) substrates. PLZT-ceramic has several times greater electro-optical coefficient than lithium niobate, which allows to create various active elements with a lower control voltage. The optical losses of a waveguide made of such ceramics are still about 4 dB/cm (which is not much more than in titanium-diffuse waveguides in lithium niobate). Further improvement of the technology

for producing waveguide layers will make it possible to reduce them to the level of the best diffuse waveguides in lithium niobate.

2.7. Semiconductor-based waveguides

Semiconductor optical waveguides are a very important part of modern integrated optoelectronic systems, especially for electrically active devices. Applications range from semiconductor lasers, optical filters, switches, modulators, isolators, and photodetectors. Semiconductor waveguides have many advantages for use in many I/O applications. They offer a significant enhancement of interaction length that, to first order, scales with the possible delay. With a tight confinement of the optical mode, the required optical power can be drastically reduced while the mode overlap with the active material is strongly enhanced. The use of semiconductor material is of particular interest since it offers compactness and enables for monolithic integration into optoelectronic devices using well established processing techniques. Furthermore, semiconductors are attractive since the operating wavelength, to a large extend, can be designed while performing with bandwidths in the GHz regime that is well suited for communication signals. Based on III–V, II–VI, or IV–VI group elements, two semiconductors with different refractive indices are generally synthesized for fabrication of optical waveguides. They must have different band gaps but same lattice constant. An attractive feature of the binary compounds is that they can be combined or alloyed to form ternary or quaternary compounds, or mixed crystals, for varying the band gap continuously and monotonically together with the variation of band structure, electronic, and optical properties. The formation of ternary and quaternary compounds of varying band gaps also enables the formation of heterojunctions, which have become essential for the design of high performance electronic and optoelectronic devices.

Semiconductor compounds $A^{III}B^V$ and their solid solutions are the most promising materials for the monolithic integrated optical circuits. To obtain waveguides in semiconductors, fabrication methods of epitaxial growth, diffusion, and ion implantation are used. The first of them was most widely distributed.

In semiconductors a change in the refractive index is achieved as a result of their doping with impurities, either due to the field-controlled conductivity, as is the case in space charge regions or in inversion layers. Another way to obtain a given n is to create multilayer structures with a smooth or stepwise change in composition from layer to layer.

The injection of a charge of free carriers into a semiconductor reduces the refractive index of the latter by an amount

$$\Delta n = -\frac{N\lambda_0^2 q^2}{\pi^2 \varepsilon_0 m^* c^2}, \qquad (2.4)$$

where n_0 is the refractive index of an undoped semiconductor; q is the electron charge; ε_0 is the dielectric constant of vacuum; N is the concentration of injected carriers, which determines the conductivity of the material, m^* is the effective mass of free carriers. In order for light to propagate in a waveguide of thickness d, the difference in the refractive index between the waveguide and the substrate must satisfy the inequality

$$\Delta n \geq \left[\frac{(2M-1)^2}{2n_0}\right]\left(\frac{\lambda_0}{4d}\right)^2, \qquad (2.5)$$

where M is the number of guided mods. Solving equation (2.4) and (2.5) determines the conditions for the waveguide cutoff

$$N_n - N_B > \frac{(2M-1)^2 \pi^2 \varepsilon_0 m^* c^2}{4q^2 d^2}, \qquad (2.6)$$

where N_n, N_B are the concentrations of free carriers in the substrate and in the waveguide layer, respectively. Condition (2.6) is independent of the wavelength. This means that, with a sufficiently large thickness of the semiconductor waveguide d, it retains waveguide properties for a relatively wide spectrum of wavelengths, in the limit of which it is transparnt.

The optical losses of the semiconductor waveguide are proportional to N. Accordingly, the optical losses in the substrate, the conductivity of which is lower than the conductivity of the waveguide layer, will be less. However, with a decrease in the conductivity of the substrate, Δn decreases. At the same time, the restriction of light worsens and, as a result, the overall losses in the structure increase. Thus, the choice of N_n and N_B is a compromise. The optimal values of N_n and N_B are determined in each case in accordance with the requirements for the characteristics of the integrated optical device.

To obtain semiconductor waveguide high-resistance layers, epitaxial technology is used, as a rule. Epitaxial waveguides of buried, rib type and with superimposed strip were obtained. Losses in such structures at a wavelength of $\lambda = 1.06 \, \mu m$ exceed 2 dB/cm. This is several times higher than the loss in waveguides on active dielectrics and glasses. This figure is not the limit. To reduce losses, it is necessary to improve the materials and technology of waveguide structures.

An example of a waveguide structure formed by the local epitaxy from the gas phase on a high-resistance substrate is shown in Fig. 2.12. A groove 1 μm wide and depth was etched in the substrate of the semi-insulating n-GaAs (110). The deposited Si_3N_4 film serves as a protective mask. The n-GaAs waveguide layer due to the strong anisotropy of the growth rate during chloride gas-phase epitaxy has the shape of a prism. This eliminates the need to cut the

crystal to input radiation and excite the waveguide directly through the substrate.
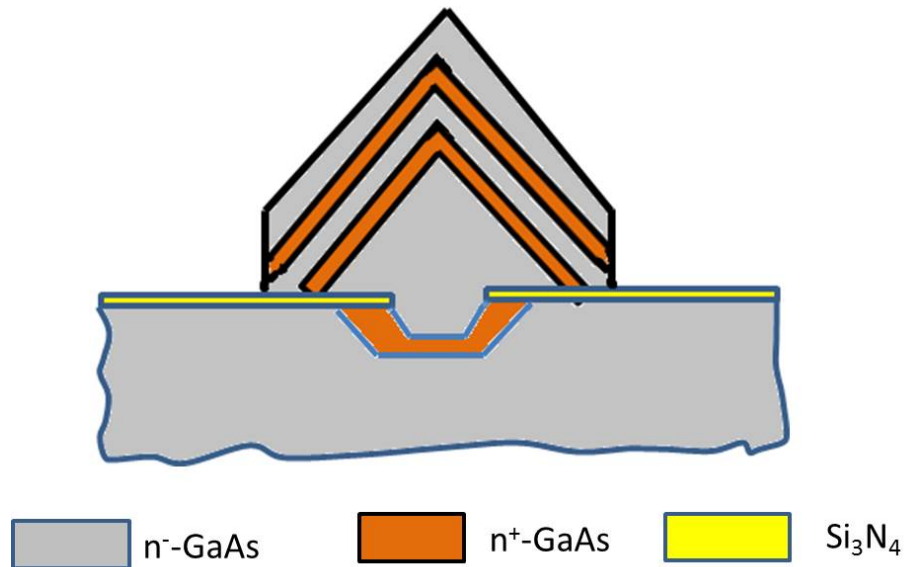
Fig. 2.12. GaAs waveguide with triangular cross section structure formed by local vapor phase epitaxy

To obtain thin high-resistance semiconductor layers, one can apply the method of bombarding their surface with protons (or other light ions) at energies of several hundred kiloelectron-volts. In this case, a partial amorphization of the semiconductor surface occurs, which requires additional annealing and its compensation, which leads to an increase in the layer resistance. The method is relatively simple, allows the processing of large surfaces with a high degree of uniformity. However, due to structural disturbances introduced during the ion bombardment, losses in waveguides also increase.

Another method of compensation in semiconductors is the diffusion of metals forming deep centers. Thermal diffusion is also the main method of doping (including local) semiconductors, specifying the sign and magnitude of their conductivity. To obtain diffusion waveguides on gallium arsenide and indium phosphide, diffusion into the n^+-substrate of zinc is used, which is characterized by a large diffusion coefficient (about 15 cm^2/s) and solid-state solubility $\geq 10^{20}$ cm^{-3} in these materials. Thus, high-quality planar waveguides

up to 5 mm long and 4 μm thick were obtained on n^+-GaAs substrates, and rib-type waveguides with losses not exceeding 10 dB/cm on n^+-InP substrates were obtained. The main disadvantages of the considered semiconductor waveguides are the inability to separate their waveguide and electrical characteristics and significant optical losses due to the finite conductivity of the materials.

The method that implements a given Δn profile due to programmable changes in the composition of the material turns out to be more flexible. It is based on the epitaxial technology (gas, liquid, and molecular) of layer buildup of various compositions selected so as to provide a waveguide effect. The method is the main one when creating semiconductor LEDs LD and photodetectors. Epitaxial technology allows to obtain films of binary, ternary and quaternary compounds $A^{III}B^V$ with thicknesses from a few nanometers to several tens of micrometers or more, uniformly doped, constant or variable in thickness. In contrast to thermal diffusion, epitaxy allows to obtain sharp boundaries between films with different values of the refractive index.

The $Al_xGa_{(1-x)}As$/GaAs system, in which the refractive index of the ternary compound steadily varies from 2.97 at $x = 1$ to 3.59 at $x = 0$, was the basis of the first waveguide structures used in lasers.

Compounds based on the InGaAs/InP and InGaAsP/InP structures form another group of $A^{III}B^V$ materials, which are promising for integrated optics. The level of processing of their technology for lasers, photodetectors and waveguides is close to the level of AlGaAs/GaAs technology.

Along with $A^{III}B^V$ semiconductors, $A^{II}B^{VI}$ semiconductor compounds that best correspond to visible spectra can also be used to create waveguide components. On wide-band $A^{II}B^{VI}$ materials — zinc and cadmium compounds with sulfur, selenium, and tellurium — various components of integrated optics were obtained. But in terms of manufacturability and a combination of electrophysical parameters, $A^{II}B^{VI}$ materials in integrated optics are still significantly inferior to $A^{III}B^V$ materials, and this limits their application to special tasks.

Comparative characteristics of the main waveguide structures used in integrated optical circuits are given in table 2.3.

Table 2.3

Comparative characteristics of the main waveguide structures

Waveguide material	Substrate	Refractive index ($\lambda=0.63$ μm)	Fabrication method	Losses average, dB/cm ($\lambda=0.63$ μm)
Glasses TK8—TK21	Glass	1.56–1.66	HF sputtering	0.1
KB_2O_5	— » —	2.27	— » —	0.01
ZnO	— » —	2.0	— » —	0.01
Vinyltrimethyl-silane	— » —	1.52–1.53	Glow discharge plasma polymerization	0.04
Gelatin	— » —	1.54	Deposition from solution	0.5
Ti : $LiNbO_3$	$LiNbO_3$	2.21	Diffusion	0.5
$LiNbO_3$	$LiTaO_3$	2.20	Epitaxy	1.0
$LiNbO_3$	$LiNbO_3$	2.35	Proton exchange	0.5
$LiNbO_3$	$LiTaO_3$	2.21	Ion exchange	2.0
GaAs	$Al_xGa_{1-x}As$	3.6*	Epitaxy	2**
$CdS_{1-x}Se_x$	CdS	2.46–2.5	Epitaxy	10
Cd_xSZn_{1-x}	ZnS	2.37	Diffusion	3.0
$Al_xGa_{1-x}As$	GaAs	3.2**	— » —	17*
$Ga_xIn_{1-x}As_yP_{1-y}$	InP	3.37**	— » —	6**

* – $\lambda = 1.15$ μm; ** $\lambda = 1.32$ μm

2.8. SOI waveguides

Silicon-on-Insulator (SOI) substrate offers the possibility to make waveguides with very high refractive index contrast of the order of 2 to 1 in all directions of confinement. With such strong confinement the typical dimensions of the Silicon core are 200 x 500 nm for singlemode waveguides. These waveguides – often called photonic wires – allow to reduce the bend radius

down to a few micrometer. For many types of photonic circuits this means that the die size can be miniaturized by 10-1000 times. The price paid for this dramatic reduction in size is the fact that the tolerances on the dimensions are also reduced. The geometrical precision of the topological layout needs to be of the order of 10 nm or better, so as to achieve a predictable and reproducible effective refractive index as well as to reduce the losses due to roughness. One of the major advantages however of SOI-based nanophotonic circuits is the fact that they can be fabricated by means of standard technologies normally used for the industrial manufacturing of microelectronic ICs, such as deep-UV lithography.

Two common types of SOI waveguides are shown in Fig. 2.13. In both cases a silicon waveguide core is resting on top of the SOI buried oxide (BOX) layer. The lower-index BOX optically separates the waveguide core from the silicon substrate: optical waveguide modes are never fully confined in the silicon core, and exponentially decaying "tails" extend into the silicon dioxide cladding. If the cladding is too thin, the mode will "feel" the underlying silicon substrate, which also has a high refractive index. The result will be a radiative leakage of the light into the substrate.
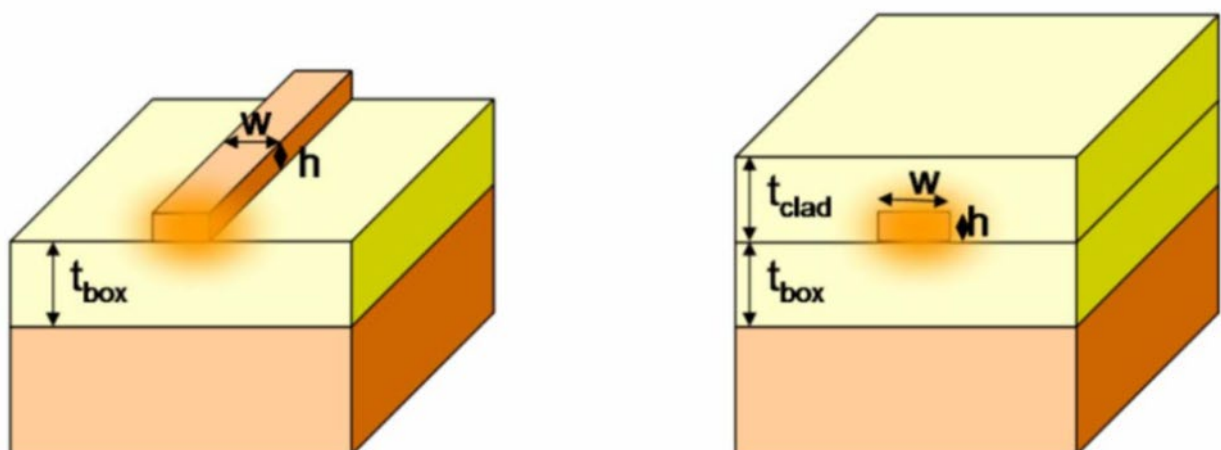

Fig. 2.13. Basic SOI photonic wire waveguide structure with air top cladding (left) and silica top cladding (right)

2.8.1. Large-core SOI waveguides

The large-core silicon waveguide is typically between 1 and 3 μm thick, and 2–5 μm wide. Such large cores with a high-index contrast of SOI are multi-mode, unless they are etched in such a way that they support only one guided mode: this is done using a rib geometry where the sides are only partially etched. The resulting waveguide has therefore a high-index contrast in the vertical direction, but in the plane of the chip the refractive index contrast is actually quite low (unetched vs partially etchcd silicon). Therefore, such waveguides are considered to be low-contrast waveguides, and they have bend radii in the order of millimeters.

The large index contrast and thickness in the vertical direction do give a good vertical confinement, and a relatively thin BOX layer is sufficient to insulate the optical mode from the substrate. BOX thicknesses are typically in the order of 300-600 nm.

Large-core SOI waveguides typically have a relative low loss, in the order of 0.1–1 dB/cm.The large core also makes them relatively insensitive to geometric variations. However, it has proven difficult to directly integrate many optical functions in this technology. The integration density is rather small, and active functions are difficult to achieve. Still, the technology has found an active use in simple photonic circuits and individual components and as an 'optical bench' technology, where individual optical components are positioned with pick-and-place techniques onto the SOI waveguide circuit.

We will not discuss further such large-core waveguides, instead focusing on the silicon nanophotonic waveguide technology.

2.8.2. Silicon photonic wires

Figure 2.14, b shows a nanophotonic silicon waveguide, also called a photonic wire. In contrast with the large-core waveguides, a photonic wire has a

high refractive index contrast in all directions. With this high contrast (3.48 to 1.44), the single-mode condition is met when the core width/height is of the order of 300–400 nm. A common choice of SOI layer thickness is 220 nm. because at this thickness the core layer itself only supports one mode (a so-called slab mode) for each polarization at a wavelength of 1550 nm. The corresponding maximum waveguide width for single-mode behavior is around 520 nm. Other thicknesses have been used as well, and in general a choice of thicker silicon will need a narrower core to maintain single-mode behavior.
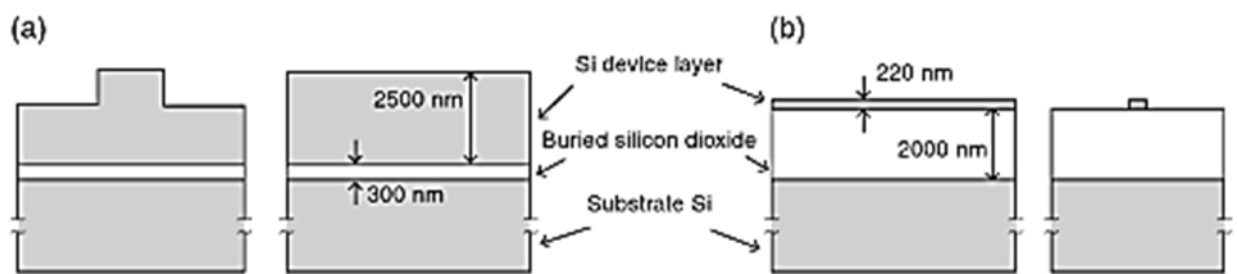
Fig. 2.14. Silicon in insulator waveguides (cross-sections): (a) large-core silicon waveguide and (b) small-core silicon waveguide ('photonic wire'). Schematic drawn to comparative scale

It needs to he noted that single-mode behavior is not an entirely accurate description. Even at these dimensions, a photonic wire will support two fundamental modes: one for the transverse electric (TE) polarization, and one for the transverse magnetic (TM) polarization. The intensity profile is plotted in Fig. 2.15. These modes behave very differently, because they will exhibit different field discontinuities at the top/bottom or left/right interfaces. In general the TE mode tends to be better confined, and it is considered the true fundamental mode. It has its dominant electric field vector in the plane of the SOI, and therefore field discontinuities at the waveguide sidewalls.
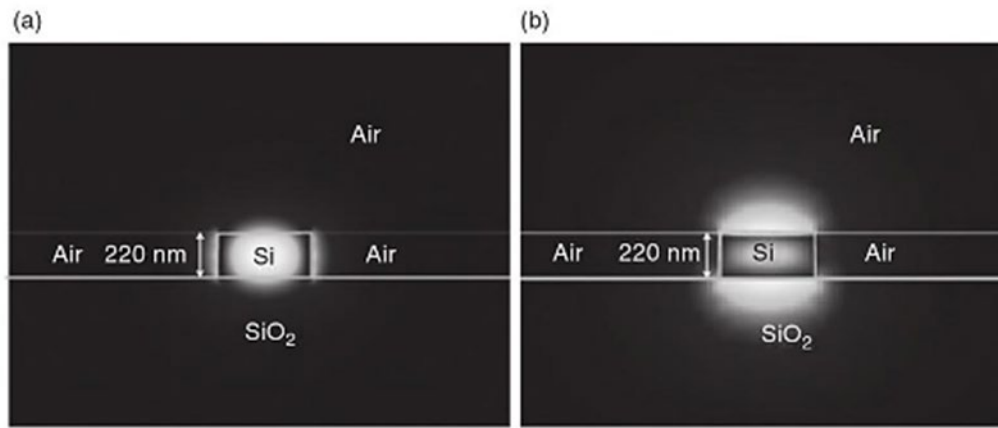
Fig. 2.15. Mode intensity plots for the fundamental TE (*a*) and TM (*b*) mode of a photonic wire waveguide

While both the TE and the TM optical mode are much smaller than with the large-core waveguides, they are less confined to the core and extend further into the oxide cladding. This imposes a requirement for a much thicker buried oxide. Figure 2.16 shows the propagation loss as a function of BOX thickness for TE and TM polarized light.
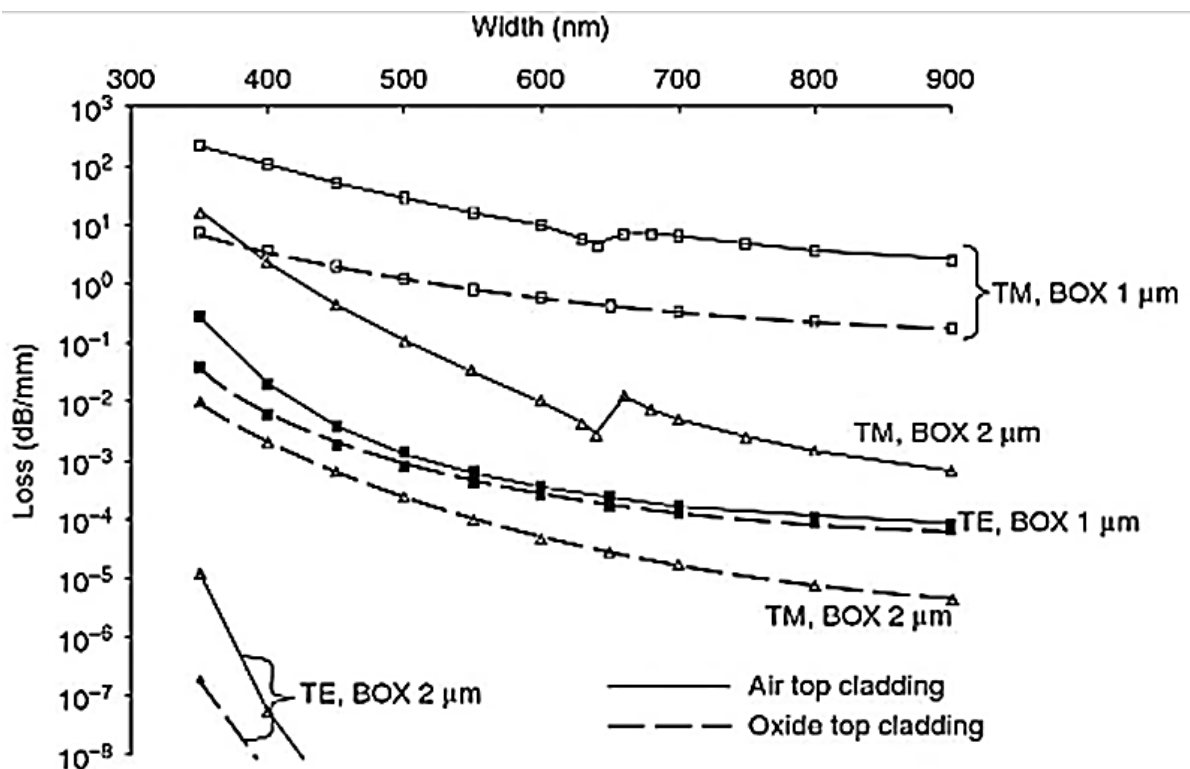
Fig. 2.16. Waveguide propagation losses for photonic wires and different BOX thicknesses

For TE polarized light at least 1000 nm of BOX is required to avoid leakage, while for the less confined TM mode a much thicker BOX is required.

Each waveguide mode is characterized by its propagation constant, or more commonly by an effective refractive index indicated by n_{eff}. Like an ordinary refractive index, the effective index describes the reduction in phase velocity of the light as it propagates through the medium. A higher refractive index will result in a slower propagation. The effective index is a weighted average between the refractive index of the core and the cladding: the more the optical mode is confined in the core, the higher its effective index. For a 450 x 220 nm waveguide, the effective index at 1550 nm is around 2.37.

Optical waveguides are dispersive, which means their effective index is wavelength dependent. One cause for this is the wavelength-dependence of the materials. But in a photonic wire, the main cause of dispersion is the geometry itself: because the dimensions of the core are on the same scale as the wavelength, a small change in wavelength will cause a change in confinement, and therefore effective index. The dispersion is described in the group index. n_g, linked to the group velocity, $v_g = c/n_g$. The group velocity is the propagation speed of a wave packet, and this is the relevant quantity for the transport of information (i.e., a modulated signal). Because of the dispersion, the group index in a silicon wire is, at $n_g = 4.3$. considerably higher than the effective index.

Figure 2.17 shows the effective refractive index, n_{eff}, of various modes in Si photonic wire waveguides as a function of waveguide width at 1550 nm for the TE and TM polarization. When n_{eff} is less than the index of the cladding n_{clad} the modes will start to leak into the cladding material. The width-dependence of n_{eff} is important for the performance of silicon photonic circuits.
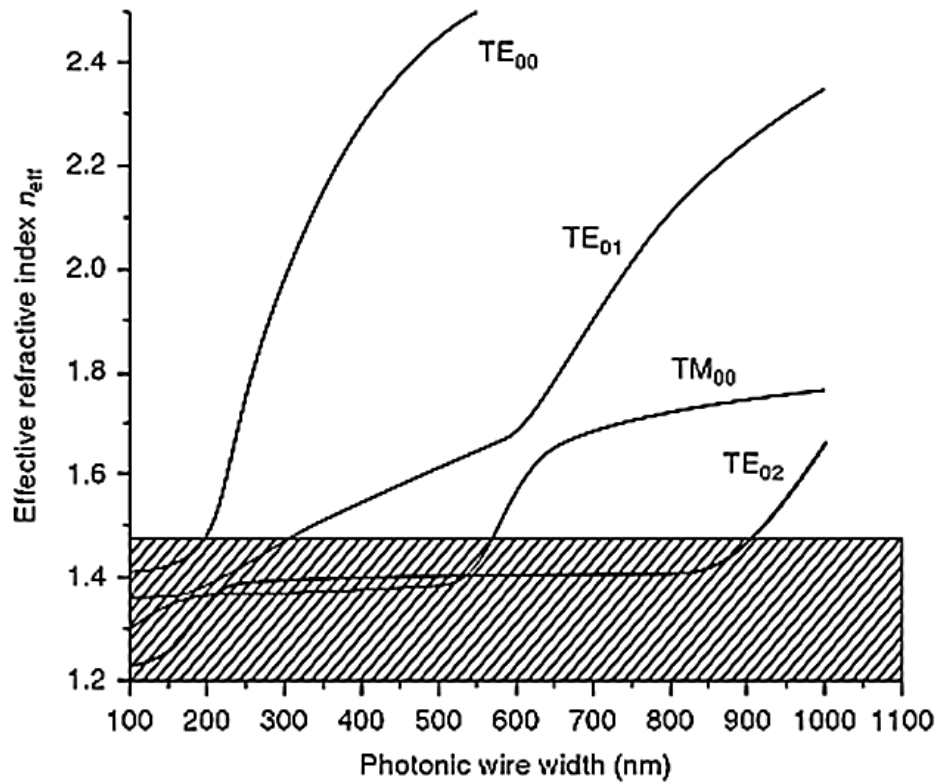
Fig. 2.17. Effective refractive index, n_{eff}, for the TE and TM mode of a 220 nm thick photonic wire as function of core width.

In addition to the thickness specification, other important requirements are thickness uniformity, doping concentration, surface roughness and defects, and sidewall roughness on the waveguides.

2.8.3. Fabrication of silicon waveguides

Even though silicon is the base material for both microelectronics and silicon photonics, fabrication process flow and specifications can be very different. Most nanophotonic waveguide circuits are made by high resolution lithography and dry etching. This flow is schematically represented in Fig. 2.18. Because of the submicron dimensions, the most commonly used lithography technologies are electron beam and 248 nm and 193 nm optical projection lithography.
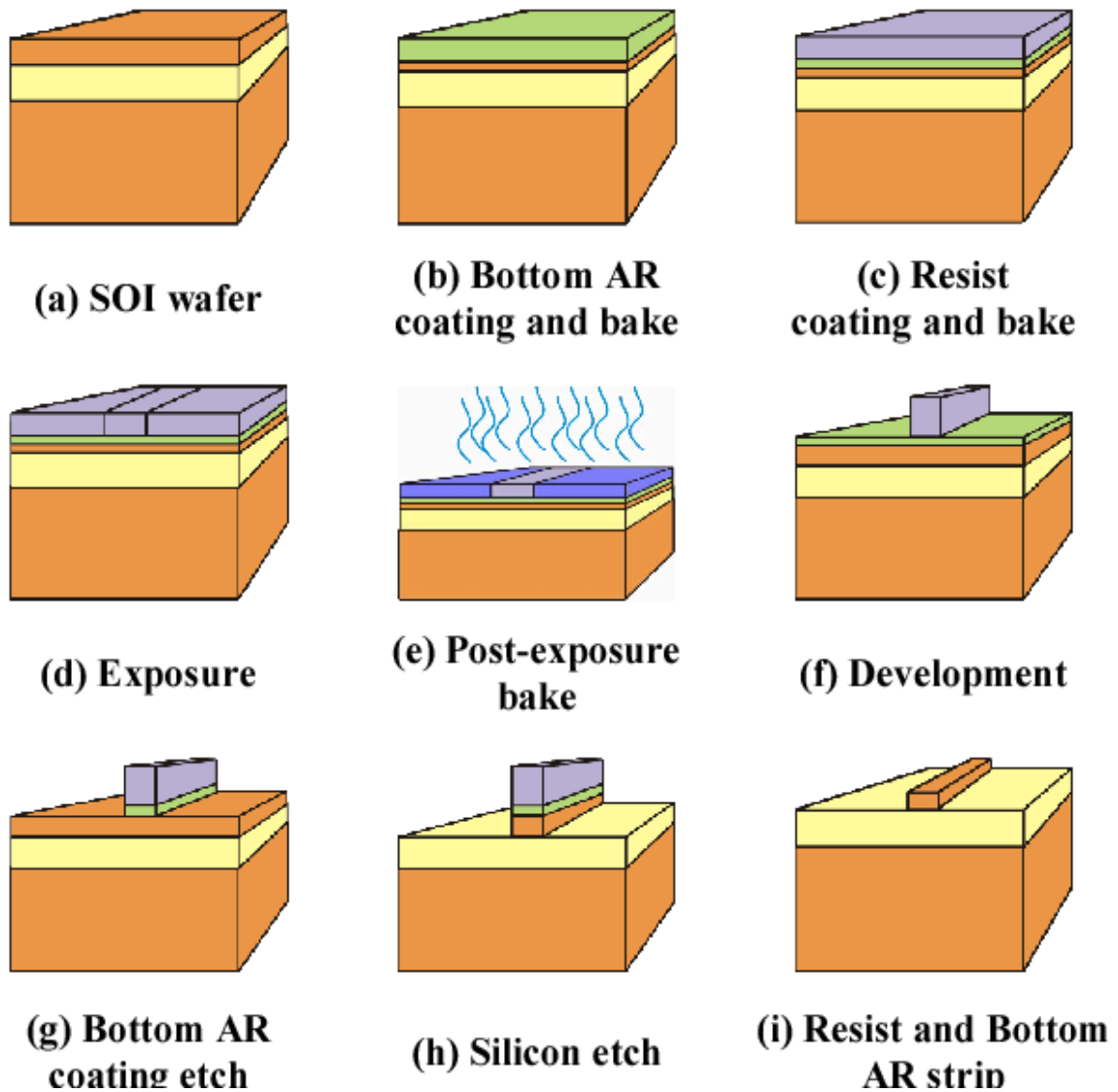
Fig. 2.18. Fabrication flow for SOI nanophotonic waveguides

The patterning requirements for passive photonic waveguides, that is, optical circuits that do not actively change the optical properties of the waveguides, are quite different from electronics. For instance, photonic circuits can contain anything from isolated lines of different widths, dense trenches and holes in one patterning layer. Lithography needs to be done in a single step as alignment accuracy between on-chip features is very important. This requires careful lithography optimization as well as biasing and proximity corrections. A variety of photonic waveguide structures is shown in Fig. 2.19.
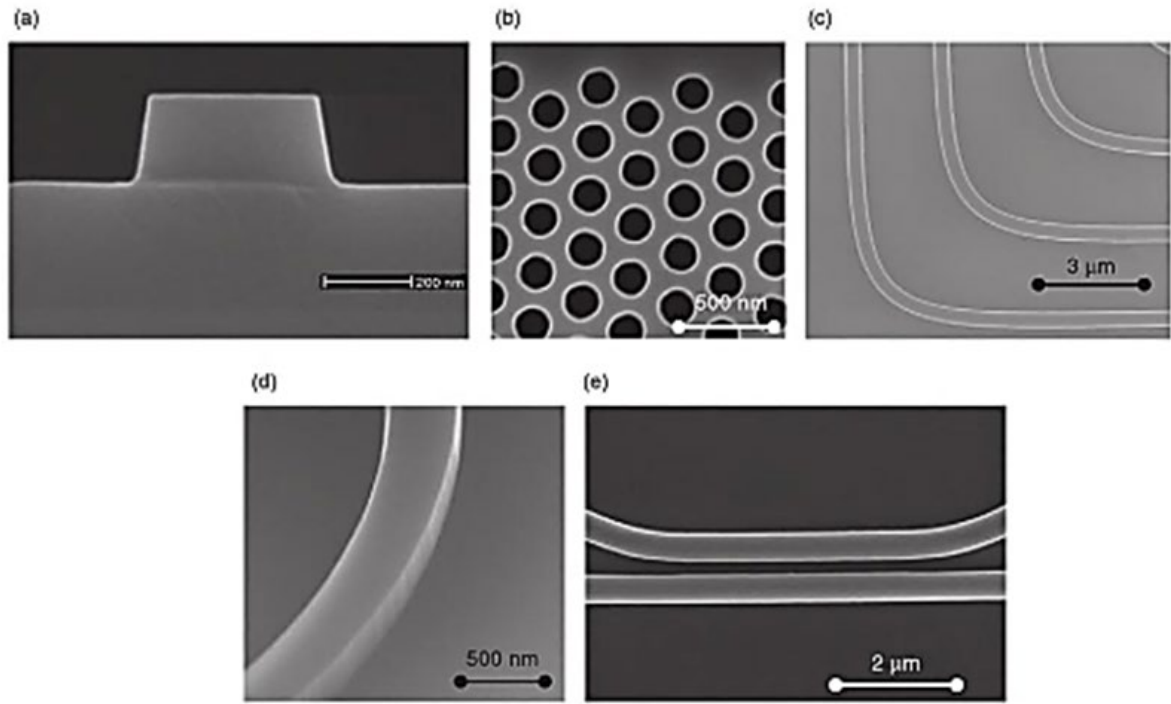
Fig. 2.19. Examples of photonic waveguide structures: *a* – cross-section of a photonic wire; *b* – photonic crystal consisting of a triangular lattice of holes; *c* – bend waveguide; *d* – waveguide with smooth sidewall; *e* – a directional coupler

As the effective index of the photonic wires is very dependent on the core width, good control of critical dimensions (CDs) is especially important. In CMOS patterning processes the process window is generally defined as a 5–10% variation on the nominal CDs. In photonics, CD control of the order of 1 nm is generally required.

Another important aspect of the photonic waveguide fabrication process is the quality at the core/cladding interface. Roughness at the interface will cause scattering and back reflection in the waveguide, which is considered to be the prime cause of propagation loss. Also, surface states at the interfaces could lead to absorption.

The top and bottom surfaces of the core layer in commercial SOI substrates is typically smooth to a sub-nanometer level. Etched sidewalls, however, can exhibit significant roughness, typically of a curtain type, as shown in Fig. 2.19. Improved surface quality results in lower propagation losses. State-of-the-art photonic wires perform between 1 and 2 dB/cm.

More complex waveguide geometries, such as the rib waveguide from Fig. 2.13, can be defined by introducing additional etch layers. Rib waveguides have the disadvantage of a lower confinement (like the large-core SOI waveguides), but have a smaller sidewall surface and therefore lower losses. Also, the shallow etch can be used to define diffractive gratings that can be used for coupling to optical fibers (see also Fig. 2.20).
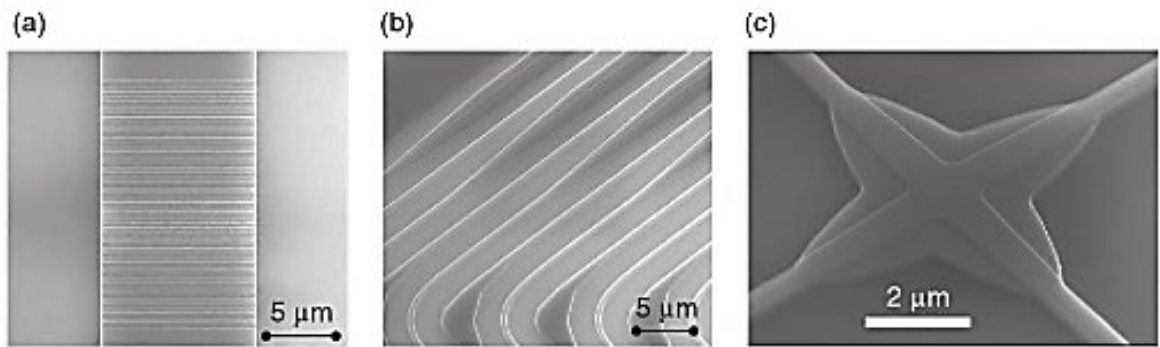
Fig. 2.20. Photonic structures defined by a deep and shallow etch process: *a* – a diffractive grating; *b* – hybrid rib/strip waveguides; *c* – a waveguide crossing

3. COUPLING ELEMENTS

From the successful solution of the coupling problem of elements in integrated optics, the efficiency of any integrated optical device or IOC strongly depends. In the simplest scheme, the number of various kinds of compounds reaches ten. As the structure of the device becomes more complicated, this number increases sharply. Therefore, even if the integrated optical circuit is composed of ideal elements whose internal losses are equal to zero, but the efficiency of their coupling with each other is low, the circuit as a whole will not work due to significant energy losses . The coupling problem arose already at the very early stage of the development of integrated optics, when, to evaluate the characteristics of the first waveguides, it was necessary to input and output radiation from it. It turned out that this task is not quite simple, and it is not possible to solve it with the desired efficiency by focusing the beam at the end or on the surface of the waveguide.

Various coupling elements have been developed, intended for both inputting and outputting radiation from the same waveguide-waveguide structures, and for matching optical fibers with thin-film integrated optical waveguides, a free light beam and a waveguide, etc. Also, some couplers selectively couple energy to a given waveguide mode, while others are multimode. Regretfully, it is typical for most components of integrated optics(as we saw with waveguide lenses) - none of the coupling elements satisfies the requirement of universality, each of them has both pros and cons. The choice of the most suitable coupling element is dictated by the characteristics of the particular integrated optical device or circuit in which it should operate

3.1. Fundamentals of optical coupling

The principal characteristics of any coupler are its efficiency and its mode selectivity. Coupling efficiency is usually given as the fraction of total power in

the optical beam, which is coupled into (or out of) the waveguide. Alternatively, it may be specified in terms of a coupling loss in dB. For a mode-selective coupler, efficiency can be determined independently for each mode, while multimode couplers are usually described by an overall efficiency. However, in some cases it is possible to determine the relative efficiencies for the various modes of a multimode coupler. Thus, the basic definition of coupling efficiency is given by

$$\eta_{cm} \equiv \frac{\text{power coupled into (out of) the } m\text{th order mode}}{\text{total power in optical beam prior to coupling}} \qquad (3.1)$$

and coupling loss (in dB) is defined as

$$L_{cm} \equiv 10\log\frac{\text{total power in optical beam prior to coupling}}{\text{power coupled into (out of) the } m\text{th order mode}}. \qquad (3.2)$$

If the power in each mode cannot be separately determined, overall values of η_{cm} and L_{cm} are used.

Coupling efficiency depends most strongly on the degree of matching between the field of the optical beam and that of the waveguided mode. This principle can be best illustrated by considering the case of the transeverse coupler.

3.2. Transverse Couplers

Transverse couplers are those in which the beam is focused directly onto an exposed cross-section of the waveguide. In the case of a free space (air) beam, this may be accomplished by means of a lens. Transverse coupling of two solid waveguides may be done by butting polished or cleaved cross-sectional faces together.

3.3. Direct Focusing

The simplest method of transverse coupling of a laser beam to a waveguide is the direct focusing or end-fire approach shown in Fig. 3.1. The waveguide may be of either the planar or channel type, but we assume a planar waveguide for the moment.
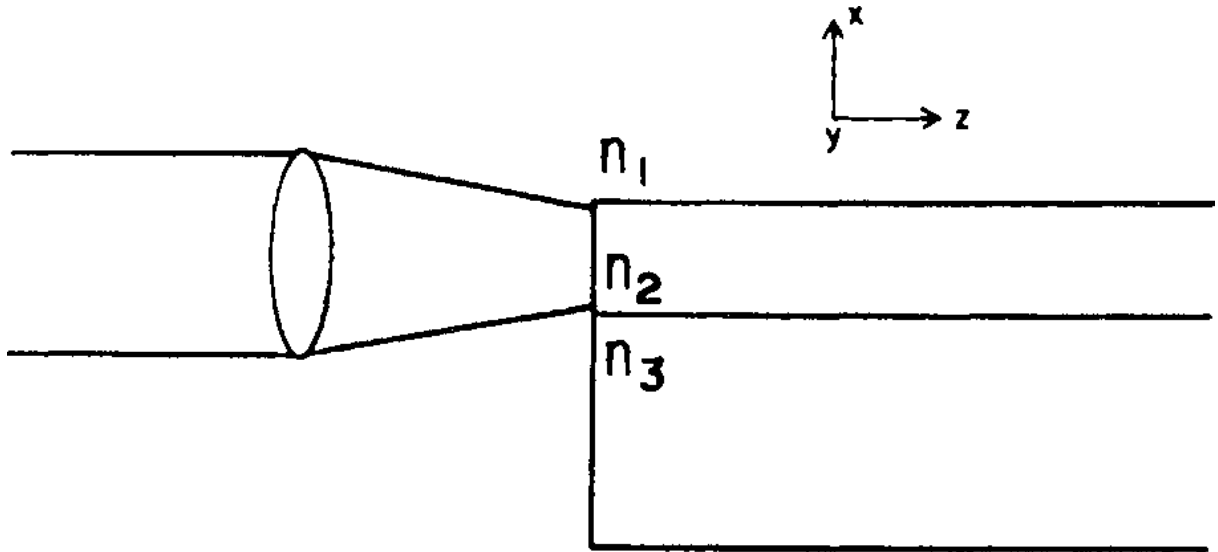
Fig. 3.1. The transverse coupling method, which is sometimes referred to as end-fire coupling

The transfer of beam energy to a given waveguide mode is accomplished by matching the beam-field to the waveguide mode field. The coupling efficiency can be calculated from the overlap integral of the field pattern of the incident beam and the waveguide mode, given by

$$\eta_{cm} = \frac{\left[\int A(x)B_m^*(x)dx\right]^2}{\int A(x)A^*(x)dx\int B_m^*(x)B_m(x)dx},\qquad (3.3)$$

where $A(x)$ is the amplitude distribution of the input laser beam, and $B_m(x)$ is the amplitude distribution of the mth mode.

The end-fire method is particularly useful for coupling gas-laser beams to the fundamental waveguide mode because of the relatively good match between the Gaussian beam profile and the TE0 waveguide mode shape. Of course, the beam diameter must be closely matched to the waveguide thickness for optimum coupling.

In principle, coupling efficiency could be nearly 100% if field contours were carefully matched. However, in practice, efficiencies of about 60% are usually achieved, because film thicknesses are on the order of 1 μm, and thus alignment is very critical.

End-fire coupling is often used in the laboratory because of its convenience. However, the difficulty of maintaining alignment without an optical bench limits its usefulness in practical applications

3.4. End-butt coupling

Transverse coupling does have a practical application in the case of coupling a waveguide to a semiconductor laser, or to another waveguide. A parallel end-butt approach can be used, as shown in Fig. 3.2. Very efficient coupling can be achieved, since the thickness of the waveguide can be made approximately equal to that of the light emitting layer in the laser, and since the field distribution of the fundamental lasing mode is well matched to the TE0 waveguide mode. The method is especially useful for coupling a laser diode to a planar waveguide, because efficient coupling of an injection laser to a thin film waveguide is difficult to achieve by using either a prism, grating, or tapered film coupler The reason for this is that the injection laser has a relatively uncollimated emitted beam which diverges at a half-angle of typically 10–20°. Prism, grating and tapered film couplers are all very sensitive to the angle of incidence of the light beam, requiring collimation to better than 1° for efficient coupling, as will be explained later in this chapter.
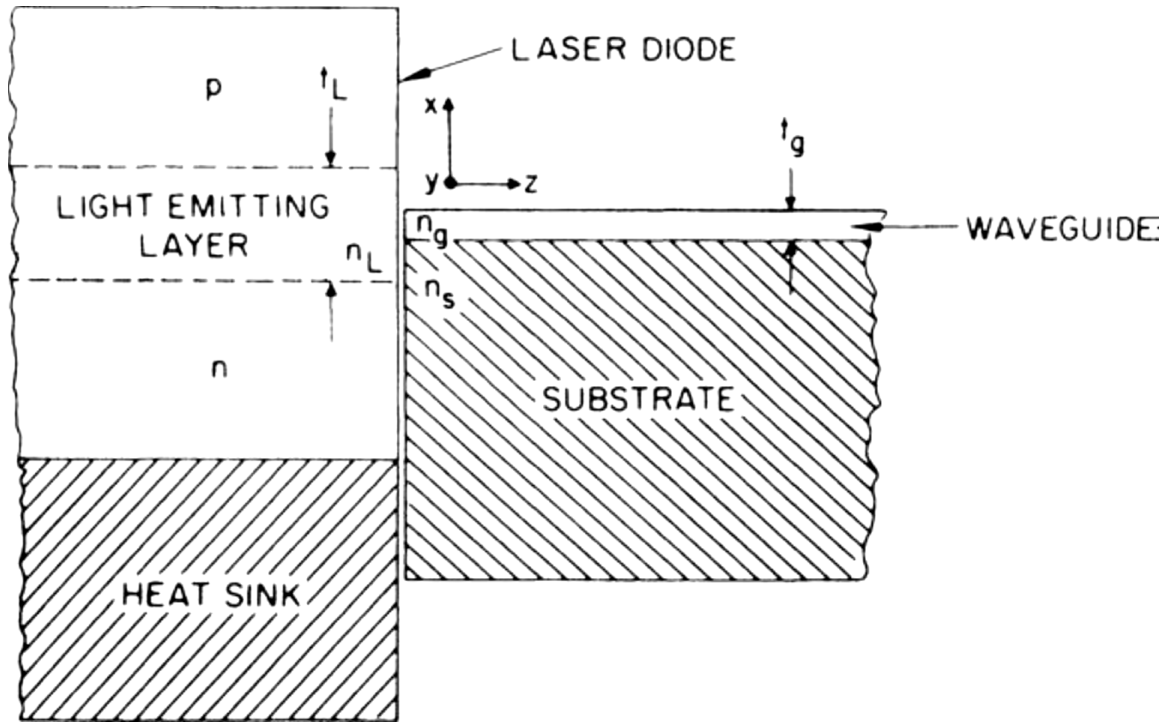
Fig. 3.2. Parallel end-butt coupling of a laser diode and thin-film waveguide

For the case of a laser diode, operating in the fundamental TE0 mode and coupled to a planar waveguide, as shown in Fig. 3.2, the coupling efficiency for the TE modes is given by

$$\eta_{cm} = \underbrace{\frac{64}{(m+1)^2 \pi^2}}_{\text{normalization}} \cdot \underbrace{\frac{n_L \cdot n_g}{(n_L + n_g)^2}}_{\text{reflection}} \cdot \cos^2\left(\frac{\pi t_g}{2 t_L}\right) \times$$

(3.4)

$$\times \underbrace{\frac{1}{\left[1 - \left(\frac{t_g}{(m+1) t_L}\right)^2\right]^2}}_{\text{overlap}} \cdot \underbrace{\frac{t_g}{t_L}}_{\substack{\text{area} \\ \text{mismatch}}} \cdot \cos^2\left(\frac{m\pi}{2}\right), \quad m = 0, 1, 2, 3,...$$

The above expression is based on the assumptions that all waveguide modes are well confined, and that $t_g \leq t_L$. It is interesting to note from the last factor of (3.4) that there is no coupling to odd-order waveguiding modes. This is

61

because the field distributions have cancelling lobes when their overlap integrals are taken with the even ($m = 0$) laser mode. The first factor of (3.4) is just a normalization term, while the second factor arises from reflections at the laser-waveguide interface. The other terms account for mismatch in the field distributions in the laser and waveguide.

Calculated curves of η_{cm} as a function of relative waveguide thickness are plotted in Fig. 3.3, along with experimental data for the case of a GaAs laser diode coupled to a Ta_2O_5 waveguide on a glass substrate. If $t_g \approx t_L$, coupling efficiency can theoretically approach 100% for the lowest-order waveguide mode. In that case, coupling into higher-order waveguide modes is nearly zero.
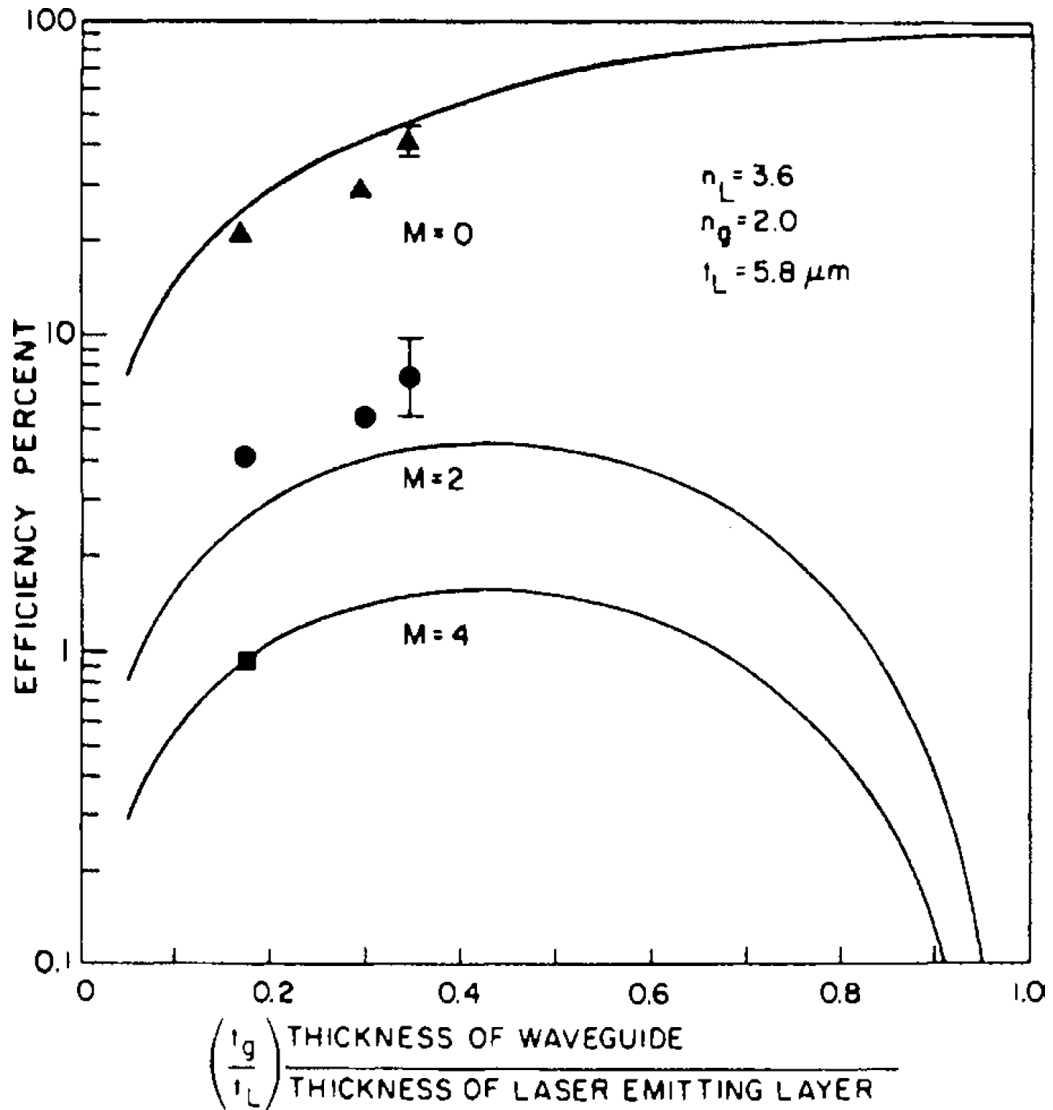
Fig. 3.3. Comparison of experimental coupling efficiency data with theoretical curves as a function of waveguide thickness

The coupling efficiencies shown in Fig. 3.3 are optimum values corresponding to perfect alignment of the laser and waveguide. Coupling efficiency is most sensitive to transverse lateral misalignment in the x direction. A displacement X of the waveguide relative to the laser, as shown in Fig. 3.4, reduces the coupling efficiency according to the relation

$$\frac{P}{P_0} = \cos^2\left(\frac{\pi X}{t_L}\right),$$
(3.5)

where P_0 is the coupled power for $X = 0$. The above expression assumes that $t_g < t_L$ and $X \leq (t_L - t_g)/2$. The dashed curve in Fig. 3.4 is the theoretically calculated P/P_0 for the case of $t_L = 5.8$ µm and $t_g = 2.0$ µm, while the solid curves represent experimentally measured data. A prism output coupler was used on the waveguide to determine the relative power that was butt coupled into each of the three modes that were observable.

The spacing between the laser and the waveguide in the z direction is also very critical, and must be controlled to a precision on the order of a wavelength for optimum coupling. Figure 3.5 shows the experimentally measured variation of coupled power as a function of z displacement. The oscillatory shape of the curve results from modulation of the effective reflectivity of the laser output face by resonance in the Fabry-Perot etalon formed by the plane parallel faces of the laser and waveguide.

In principle, this effect could be eliminated by using an index matching fluid between the laser and waveguide, so that the coupled power would vary smoothly with z displacement. as shown by the dashed curve in Fig. 3.5. The results presented so far demonstrate that end-butt coupling can be a very efficient means of coupling a diode laser to a waveguide. Similar results can be achieved for the case of butt coupling a laser to an optical fiber. Yet, it is also obvious that submicrometer alignment tolerances are required if optimum efficiency is to be obtained. Alignment to such tolerances can be achieved by

using piezoelectrically driven micrometer heads, which feature a small piezoelectric crystal stage bonded onto the end of a conventional screw micrometer. Coarse alignment is established with the micrometer screw, and then final alignment is produced by applying a voltage to the piezoelectric stage to move the laser (or waveguide), while coupled optical power is monitored by means of an appropriate output coupler and photodetector.

$$\frac{P}{P_o} = \cos^2\left(\frac{\pi x}{t_L}\right) \; ; \; t_L = 5.8 \, \mu m \quad t_g = 2.0 \, \mu m$$
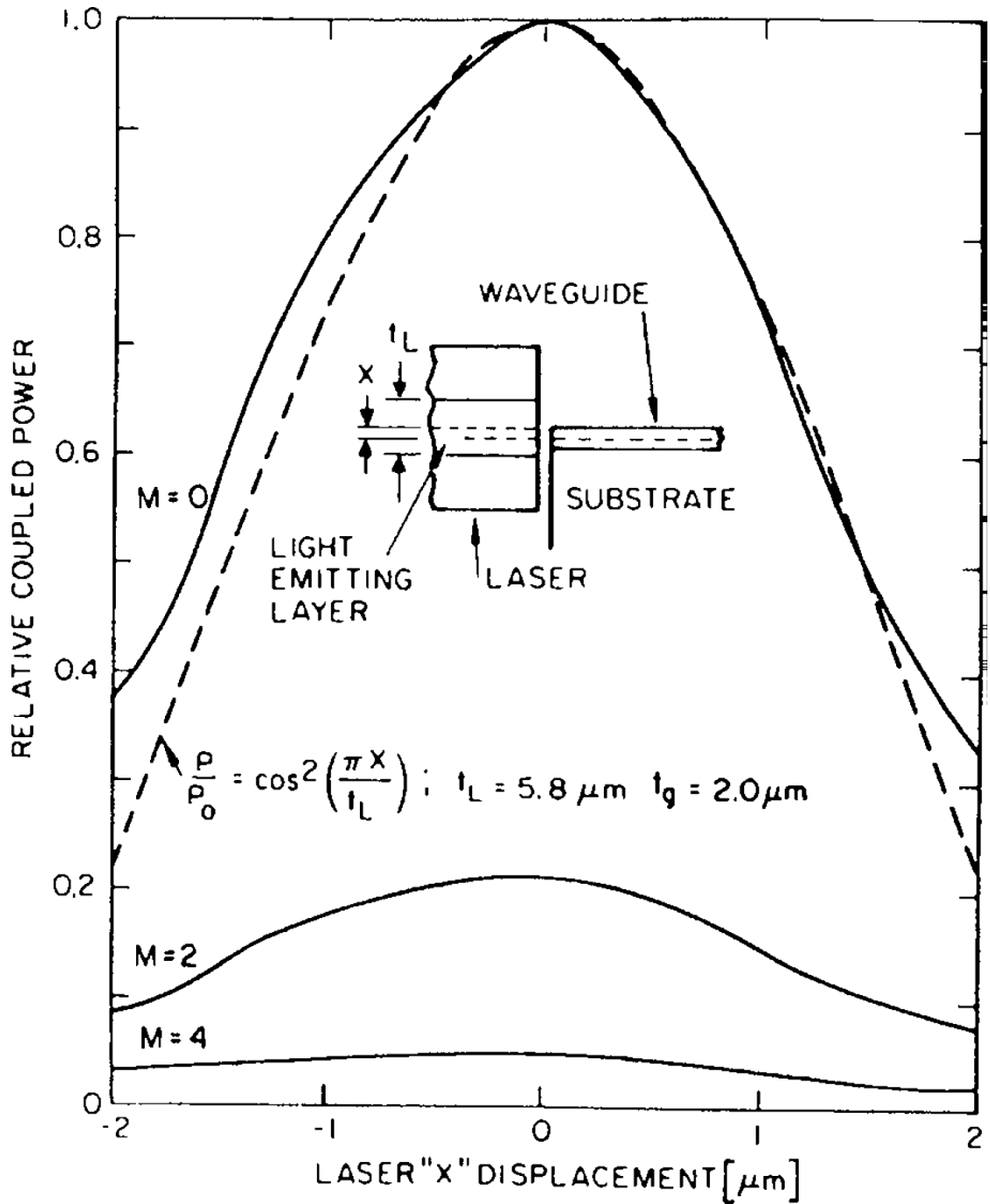
Fig. 3.4. Comparison of experimental coupling efficiency data (solid line) with theoretical curve (dashed) as a function of lateral misalignment of laser and waveguide
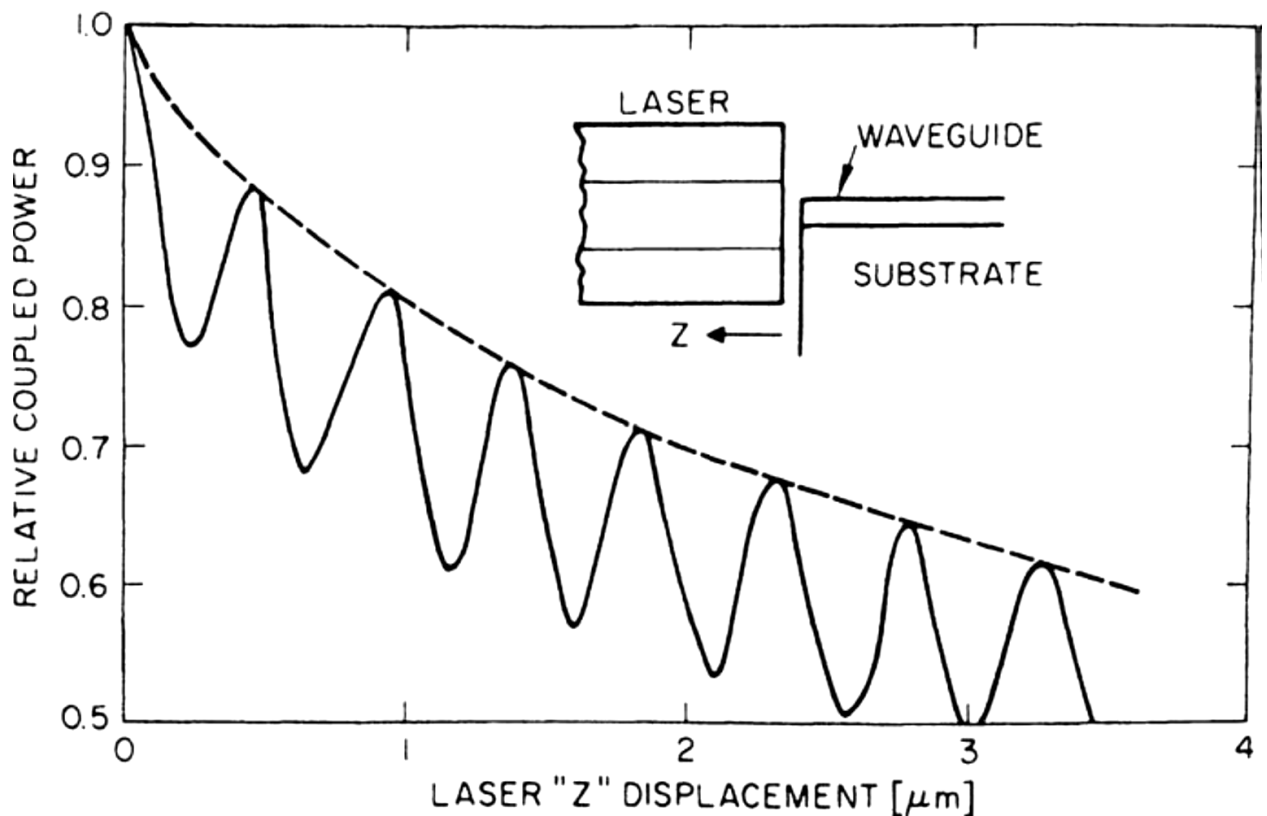
Fig. 3.5. Experimentally measured dependence of coupling efficiency on spacing between laser and waveguide

Piezoelectric micrometers with a sensitivity of better than 40Å per volt, over a 2000 V range, are commercially available. Thus, alignment to better than 0.1 μm can be permanently bonded to the waveguide support structure with epoxy, or with a metallic bond. Because the size and mass of both the laser and the OIC are relatively small, vibration sensitivity is not a significant problem, and reliable alignment can be maintained.

While the above-described alignment technique is an effective method of coupling a laser diode to a waveguide it is relatively time consuming, and hence expensive in a production line setting. To solve this problem, techniques of both hybrid and monolithic integrated fabrication have been developed which take advantage of automated, batch-fabrication procedures. An example of monolithic integration is the waveguide coupling of an InGaAsP distributed-feedback laser and electro-absorption modulator on an InP substrate. The method is based on carefully controlled selective-area Metal-Organic-Vapor-

Phase Epitaxy (MOVPE) of Multiple-Quantum-Well (MQW) structures. The integrated laser/modulator was employed to transmit data over an 80 km singlemode fiber at a data rate of 2.5 Gbit/s.

3.5. Prism couplers

Transverse coupling can be used only when a cross-sectional end face of the waveguide is exposed. In many cases, it is necessary to couple light into a waveguide that is buried within an OIC, with only the surface exposed. One could envision focusing the light onto the surface of the waveguide at an oblique angle, as shown in Fig. 3.6, but a fundamental problem is encountered in that case. For coupling to occur, it is necessary that the components of the phase velocities of the waves in the z direction be the same in both the waveguide and the beam. Thus, a phase-match condition must be satisfied, which requires

$$\beta_m = k n_1 \sin \theta_m = \frac{2\pi}{\lambda_0} n_1 \sin \theta_m. \tag{3.6}$$
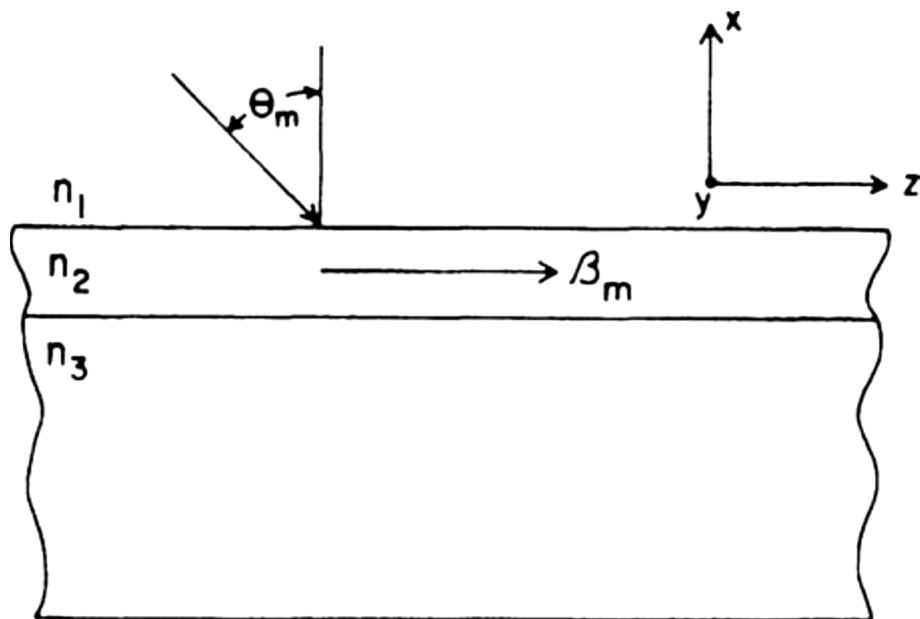
Fig. 3.6. Diagram of an attempt to obliquely couple light into a waveguide through its surface

It was shown that, for a waveguided mode

$$\beta_m > k n_1. \tag{3.7}$$

Combining (3.6) and (3.7) leads to the result that $\sin(\theta_m) > 1$, which is, of course, impossible.

One solution to the problem of phase matching is to use a prism, as shown in Fig. 3.7.
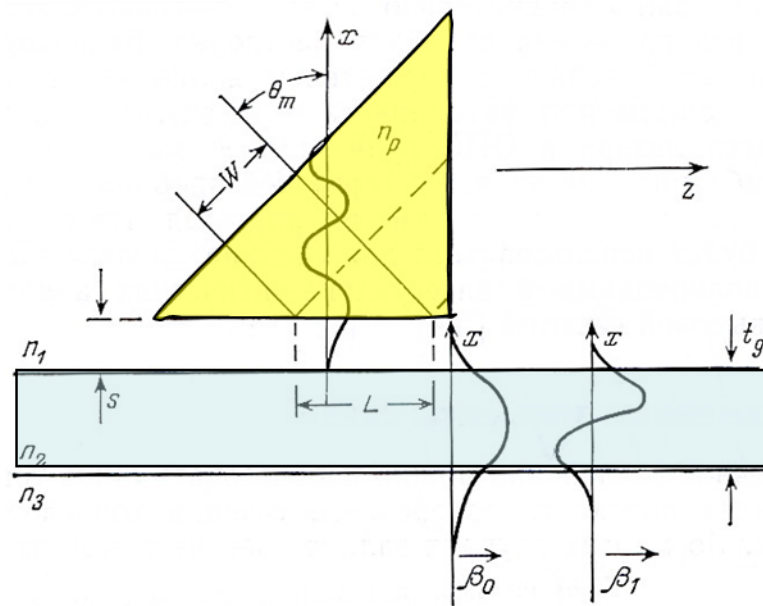
Fig. 3.7. Diagram of a prism coupler. The electric field distributions of the prism mode and the $m = 0$ and $m = 1$ waveguide modes in the x direction are shown

A beam of light of width W is directed into the face of the prism, which has $n_p > n_1$. The beam is totally internally reflected at the $n_p - n_1$ interface, setting up a standing wave mode in the prism, as shown in Fig. 3.7. This mode is stationary in the x direction, but moves in the z direction with a phase constant β_p. In the waveguide, various guided modes can exist, moving in the z direction with phase constants β_m. All of these guided modes have an evanescent tail extending slightly beyond the $n_1 - n_2$ interface. If the prism spacing s is small enough so that the tails of the waveguide modes overlap the tail of the prism

mode, there is coherent coupling of energy from the prism mode to the mth waveguide mode when θ_m is chosen so that $\beta_p = \beta_m$. The condition for matching of the β terms is given by

$$\frac{2\pi n_p}{\lambda_0}\sin\theta_m = \beta_m. \qquad (3.8)$$

Although θ_m must be carefully chosen in order to couple to a given mode, a single prism can be used to couple to many different modes by merely changing the angle of incidence of the optical beam. The beam need not be perpendicular to the prism surface, as shown in Fig. 3.7. However, if the beam is not perpendicular to the prism surface, refraction at that interface will require a modification of the expression for θ_m given by (3.8). The process of coupling energy via the overlapping mode tails, while the incident beam tends to be totally internally reflected in the prism, is sometimes called optical tunneling, because it is analogous to the quantum mechanical tunneling of a particle through an energy barrier. The modes in the waveguide are only weakly coupled to the mode in the prism. Hence, negligible perturbation of the basic mode shapes occurs. Of course, the condition

$$\theta_m > \theta_c = \sin^{-1}\left(\frac{n_1}{n_p}\right) \qquad (3.9)$$

must also be satisfied if total internal reflection is to occur in the prism, where θ_c is the critical angle.

Because of the size of the prism, the interaction between prism and waveguide modes can occur only over the length L. The theory of weakly coupled modes indicates that a complete interchange of energy between phase-matched modes occurs if the interaction length in the z direction satisfies the relation

$$\kappa L = \frac{\pi}{2} \qquad\qquad (3.10)$$

where κ is the coupling coefficient. The coefficient κ depends on n_p, n_1 and n_2, which determine the shape of the mode tails, and on the prism spacing s. From (3.10), the length required for complete coupling is given by

$$L = \frac{W}{\cos\theta_m} = \frac{\pi}{2\kappa} . \qquad\qquad (3.11)$$

For a given L, the coupling coefficient required for complete coupling is thus given by

$$\kappa = \frac{\pi\cos\theta_m}{2W} . \qquad\qquad (3.12)$$

This condition for complete coupling assumes that the amplitude of the electric field is uniform over the entire width W of the beam. In a practical case this is never true. For a Gaussian beam shape, it results that the maximum coupling efficiency is about 80%. For a more detailed discussion of the effect of beam width and shape on coupling efficiency. Also, it can be seen that, in order to get 100% coupling with a uniform beam, the trailing edge of the beam must exactly intersect the right-angle corner of the prism. If it intersects too far to the right, some of the incident power will be either reflected or transmitted directly into the waveguide and will not enter the prism mode. If the beam is incident too far to the left, some of the power coupled into the waveguide will be coupled back out into the prism.

The prism coupler is frequently used in integrated optics applications because of its versatility. It can be used as either an input or output coupler. When used as an output coupler, the prism is arranged exactly as in Fig. 3.7,

except that the direction of travel of the waveguided light would be in the negative z direction.

If more than one mode is propagating in the guide, light is coupled out at specific angles corresponding to each mode. Because of this characteristic, the prism coupler can be used as an analytical tool to determine the relative power in each waveguide mode. The prism can also be moved along the length of the waveguide to determine losses. However, care must be taken to apply the same mechanical pressure to the prism during each measurement of coupled power so that the spacing, and hence the coupling coefficient, will be constant.

One disadvantage of the prism coupler is that np must be not only greater than n_1 but also greater than n_2. This is true because the waveguide index n_2 is generally close to the substrate index n_3, which leads to the result that

$$\beta_m \cong kn_2 = \frac{2\pi}{\lambda_0} n_2.$$

(3.13)

Since $\sin \theta_m \leq 1$, (3.13) coupled with (3.8) implies that $n_p > n_2$. In the case of glass waveguides, with indices ≈ 1.5, it is easy to find a suitable prism material with $n_p > n_2$. However, semiconductor waveguides, which typically have indices ≈ 3 or 4, are more difficult to couple with prisms. Both the index and the transparency of the prism material must be considered at the wavelength of interest.

Another disadvantage of the prism coupler is that the incident beam must be highly collimated because of the critical angular dependence of coupling efficiency into a given mode. Due to this problem, prism couplers cannot be used effectively with semiconductor lasers, which have a beam divergence half angle of 10–20° unless a lens is used to collimate the beam.

Prism couplers can be used in some situations for effectively coupling optical beams to other than planar waveguide structures. They have been used to couple light into and out of the fundamental whispering-gallery mode (WGM)

of a dielectric microsphere. Light from a tunable narrow-linewidth diode laser emitting around 670 nm was focused by a lens to a small spot on the surface of a triangular prism that was in contact with a 59-μm-diameter glass microsphere. This produced an evanescent wave via frustrated total internal reflection within the prism, through which light was coupled to the WGMs of the microsphere resonator.

Prism couplers are very useful in laboratory applications where flexibility is desired regarding the position of the incident beam. However, the requirement of a stable mechanical pressure to hold the prism in place makes it less useful in practical applications, in which vibration and temperature variations are usually encountered.

The grating coupler can be used to avoid this problem, without giving up the advantage of mode selectivity.

3.6. Grating couplers

The grating coupler, like the prism coupler, functions to produce a phase matching between a particular waveguide mode and an unguided optical beam which is incident at an oblique angle to the surface of the waveguide, as shown in Fig. 3.8. It will be recalled that, without the grating, the phase matching condition is given by (3.6), and cannot be satisfied for any θ_m.

Because of its periodic nature, the grating perturbs the waveguide modes in the region underneath the grating, thus causing each one of them to have a set of spatial harmonics with z-direction propagation constants given by

$$\beta_\nu = \beta_0 + \frac{\nu 2\pi}{\Lambda} \tag{3.14}$$

where $\nu = 0, \pm 1, \pm 2, \ldots$, and where Λ is the periodicity of the grating. The fundamental factor β_0 is approximately equal to the β_m of the particular mode in

the waveguide region not covered by the grating. Because of the negative values of v, the phase matching condition (3.6) can now be satisfied so that

$$\beta_v = kn_1 \sin \theta_m, \qquad (3.15)$$

even though $\beta_m > kn_1$.
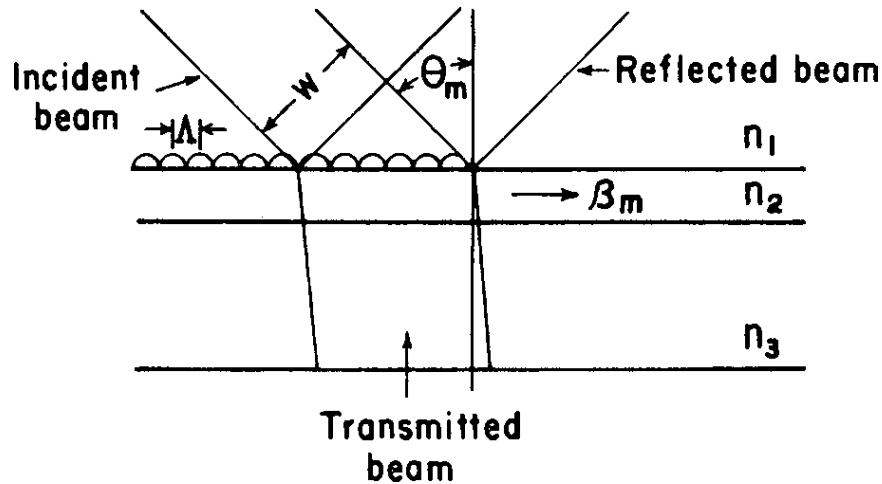
Fig. 3.8. Diagram of a grating coupler

Since all of the spatial harmonics of each mode are coupled to form the complete surface wave field in the grating region, energy introduced from the beam into any one of the spatial harmonics is eventually coupled into the fundamental ($v = 0$) harmonic as it travels to the right and past the grating region. This fundamental harmonic is very close to, and eventually becomes, the β_m mode outside of the grating region. Thus, the grating coupler can be used to selectively transfer energy from an optical beam to a particular waveguide mode by propertly choosing the angle of incidence. The grating can also be used as an output coupler, because, by reciprocity, energy from waveguide modes will be coupled out at specific angles, θ_m, corresponding to a particular mode.

The preceding paragraphs have described the operation of the grating coupler in relatively simple terms. However, the details of the coupling

phenomena are quite complex, and depend very strongly on the cross-sectional shape of the grating bars as well as on their spacing (Fig. 3.9).
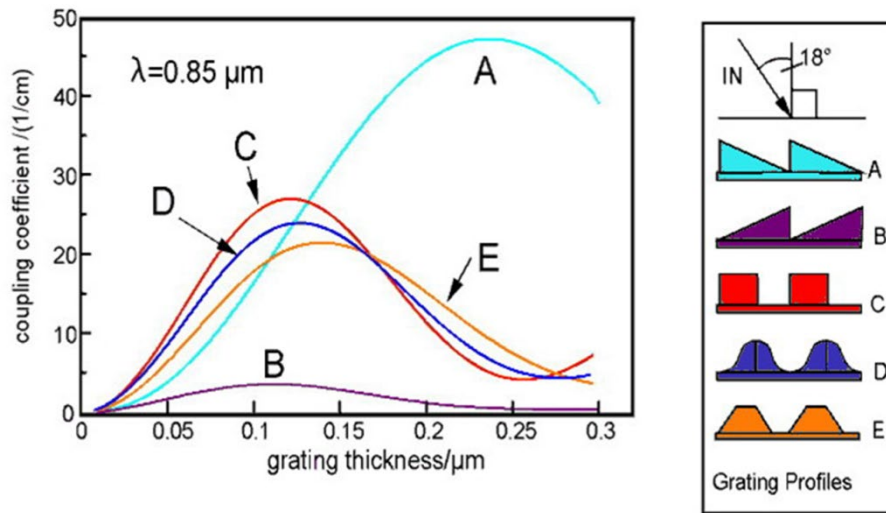
Fig. 3.9. Coupling efficiency as function of grate thickness and grating profiles

As in the case of the prism coupler, an optimum coupling efficiency of approximately 80% is theoretically possible when coupling a Gaussian beam with a grating. However, typical unblazed gratings (with symmetric profiles) generally have efficiencies of 10–30%. The principal reason for this is that much of the incident energy is usually transmitted through the guide and lost in the substrate, because, unlike a prism, the grating does not operate in a total internal reflection mode. Power can also be coupled into higher-order diffracted beams produced by the grating, unless the ratio of grating periodicity to guide wavelength is approximately 1.

The efficiency of a grating coupler can be greatly improved by shaping its profile asymmetrically to 'blaze' it for optimum performance at the coupling angle and wavelength of interest. For example, the theoretical maximum efficiency for coupling either the TE0 or TM0 mode to an air beam is roughly 50% for symmetric grating profiles, while an asymmetric saw-tooth profile can produce an efficiency greater than 95%. These theoretical predictions are supported by experimental results indicating very high efficiencies with blazed gratings. The principal advantage of the grating coupler is that, once fabricated,

it is an integral part of the waveguide structure. Hence, its coupling efficiency remains constant and is not altered appreciably by vibration or ambient conditions. Also, the grating coupler can be used on high-index semiconductor waveguides for which it is difficult to obtain a suitable prism material. However, since it is highly angle dependent, the grating coupler cannot be used effectively with the relatively divergent beam of a semiconductor laser. Perhaps the greatest disadvantage of the grating coupler is that it is difficult to fabricate, requiring the use of sophisticated masking and etching techniques.

3.7. Fiber to waveguide couplers

A key element of an integrated optic system is the coupler that transfers the optical wave between a fiber, used for long distance transmission, and a waveguide of the OIC used for signal processing. Research in recent years has resulted in the development of a number of different types of fiber-to-waveguide couplers.

The fiber may be directly butted in contact with the waveguide, without any interfacing device, in an end-on alignment. If the cross-sectional area of the fiber core and the waveguide are closely matched, high efficiency coupling can be achieved, as in the case of butt-coupled channel waveguides or a laser diode and a channel waveguide. An index matching fluid can be used to reduce reflection loss at the interface. The greatest problem with the butt coupling approach is that it is extremely difficult to establish and maintain correct alignment, since both the fiber core and waveguide typically have micron sized dimensions.

A mechanical arrangement for effectively coupling a single-mode fiber to a laser diode (or waveguide) is difficult because the fiber core diameter is only a few micrometers and the laser light emitting region is less than 1 micrometer thick.

Also, it is necessary to maintain the alignment while the fiber is permanently bonded in place. Piezoelectrically driven micromanipulators can be used to align the fiber, as mentioned previously in this chapter. To provide permanent bonding of the fiber and laser, the fiber is metallized and soldered to a fiber mount substrate. That substrate is, in turn, attached to a thermoelectric heat sink with solder. On the underside of the fiber mount substrate are metallization leads and a thin film resistor which are used to melt the solder by passage of an electrical current. A computerized feedback control system is used to move the laser and fiber into optimum alignment while current passing through the thin film resistor keeps the solder fluid. When the desired alignment has been achieved, the computer turns off the current, allowing the solder to cool and set. Thus the laser and fiber remain permanently bonded to the thermoelectric heat sink and alignment is maintained.

The silicon V-groove and flip-chip techniques, which have been used for years in the electrical integrated circuits industry, have been applied to the problem of aligning a channel waveguide with a single-mode fiber (Fig. 3.10).
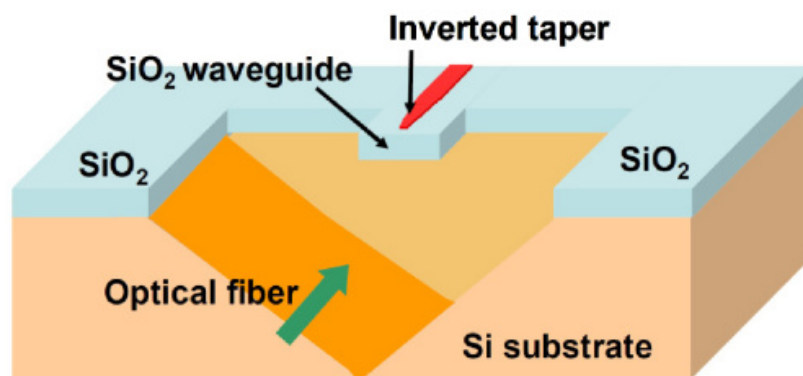
Fig. 3.10. Schematic of the SOI edge-coupling structure based on the use of a Si inverse taper and a SiO_2 waveguide implemented in the BOX. A V-groove is etched in the Si substrate to allow for fiber auto-alignment

Etched grooves can also be utilized to align optical fibers with either planar or channel waveguides in polymers. Precisely oriented and sized slots are generated by eximer-laser ablation to accept the core of an optical fiber from

which the cladding has been stripped. Angular precision of 0.3° and translational accuracy of 0.5 μm have been achieved, resulting in coupling losses less than 0.5 dB.

3.8. Coupling Between Waveguides

The phenomenon of optical tunneling can be used not only to couple energy from a fiber or a beam to a waveguide, but also to couple one waveguide to another. Couplers of this type are usually called directional couplers because the energy is transferred in a coherent fashion so that the direction of propagation is maintained. Directional couplers have been fabricated in two basic geometries: multilayer planar structures, and dual side-by-side channel waveguides.

3.8.1. Multilayer planar waveguide couplers

While butt coupling can be used to couple two planar waveguides the more common method is to bring the guides into close proximity and allow coupling to occur through phase coherent energy transfer (optical tunneling), as shown in Fig. 3.11. The indices n_0 and n_2 in the guiding layers must be larger than n_1 and n_3, and the thickness of the confining layer 1 must be small enough that the evanescent tails of the guided modes overlap. In order for energy transfer to occur between the two guides, they must have identical propagation constants. Thus, the indices and the thicknesses of the waveguiding layers must be very carefully controlled to provide matching propagation constants. This is difficult to do but has been accomplished with excellent results. As in the case of other devices embodying the synchronous coupling principle, such as the prism coupler, the interaction length must be carefully chosen for optimum coupling. The condition for total transfer of energy is again given by (3.10), but

the value of κ would obviously be different for a pair of overlapped planar waveguides than for a prism coupler.
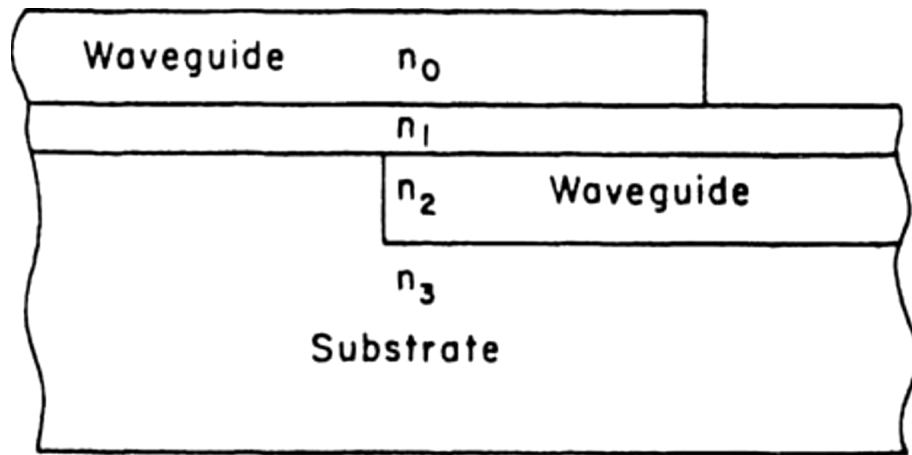
Fig. 3.11. Coupling between two planar waveguides by optical tunneling. Transfer of energy occurs by phase coherent synchronous coupling through the isolation layer with index n_1

This method of coupling is diffcult to use with deposited thin film waveguides because thicknesses and indices of refraction cannot be conveniently controlled.

However, in the case of epitaxially grown waveguides, the superior control of thicknesses and index (by controlling composition) make the problem of matching propagation constants much easier to solve. The method seems to be particularly well suited to coupling an integrated laser diode to a waveguide.

A very efficient laser can be obtained by using two separate, but coupled, waveguides to isolate the active region, where photons are generated, from the distributed feedback region.

3.8.2. Dual-channel directional couplers

The dual-channel directional coupler, which is analogous to the microwave dualguide multihole coupler, consists basically of parallel channel optical waveguides sufficiently closely spaced so that energy is transferred from one to the other by optical tunneling, as shown in Fig. 3.12. This energy is

transferred by a process of synchronous coherent coupling between the overlapping evanescent tails of the modes guided in each waveguide. Photons of the driving mode, say in guide 0, transfer into the driven mode in guide 1, maintaining phase coherence as they do. This process occurs cumulatively over a significant length; hence, the light must propagate with the same phase velocity in each channel in order for this synchronous coupling to occur. The fraction of the power coupled per unit length is determined by the overlap of the modes in the separate channels. Thus, it depends on the separation distance s, the interaction length L, and the mode penetration into the space between channels, which can be characterized by the extinction coefficients p and q.
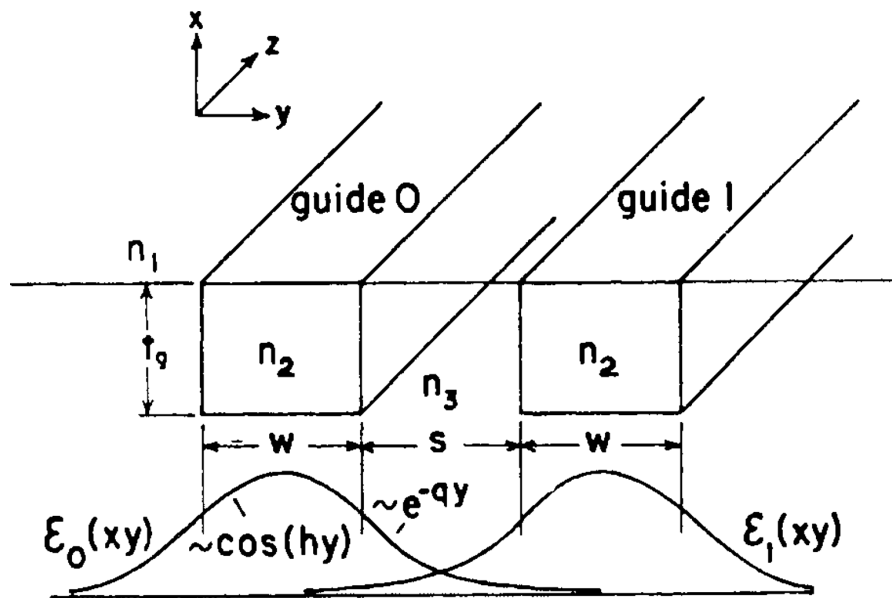
Fig. 3.12. Diagram of dual-channel directional coupler. The amplitudes of the electric field distributions in the guides are shown below them

In a dual-channel coupler, the energy transfers alternately from one waveguide to the other, and then back again if the interaction length is sufficient. If one were to measure the optical energy density while moving in the z direction along one channel of a direction coupler, a sinusoidal variation with distance would be observed.

For a coupler to transfer any given fraction of the energy, it is necessary only to bend away the secondary channel at the proper point. In this way, for

78

example, either a 10 dB coupler for measurement padding, a 3 dB coupler for beam splitting, or a 100% coupler for beam switching can be made. Dual-channel couplers can produce a very efficient transfer of optical energy from one guide to another over a relatively short interaction length.

A concise theory of operation of the dual-channel directional coupler can be developed by following the coupled mode theory approach according to which the power in guide number 1 is given by

$$P_1(z) = |A_1(z)|^2 = A_1(z) A_1^*(z) \tag{3.16}$$

where $A(z)$ is a complex amplitude.

The coupling between modes is given by the general coupled mode equations for the amplitudes of the two modes. Thus,

$$\frac{dA_0(z)}{dz} = -i\beta_0 A_0(z) + \kappa_{01} A_1(z), \tag{3.17}$$

and

$$\frac{dA_1(z)}{dz} = -i\beta_1 A_1(z) + \kappa_{10} A_0(z), \tag{3.18}$$

where β_0 and β_1 are the propagation constants of the modes in the two guides, and κ_{01} and κ_{10} are the coupling coefficients between modes.

Consider the guides shown in Fig. 3.12. Assume that the guides are identical and that they both have an exponential optical loss coefficient α. Thus,

$$\beta = \beta_r - i\frac{\alpha}{2}, \tag{3.19}$$

where $\beta = \beta_0 = \beta_1$, and β_r is the real part of β. For the case of identical guides, it is obvious from reciprocity that

$$\kappa_{01} = \kappa_{10} = -i\kappa, \tag{3.20}$$

where κ is real. Then, using (3.19) and (3.20), Eqs. (3.17) and (3.18) can be rewritten in the form

$$\frac{dA_0(z)}{dz} = -i\beta A_0(z) - i\kappa A_1(z), \tag{3.21}$$

and

$$\frac{dA_1(z)}{dz} = -i\beta A_1(z) - i\kappa A_0(z), \tag{3.22}$$

If it is assumed that light is coupled into guide 0 at the point $z = 0$, the power flow in the guides is given by

$$P_0(z) = A_0(z) A_0^*(z) = \cos^2(\kappa z) e^{-\alpha z} \tag{3.23}$$

and

$$P_1(z) = A_1(z) A_1^*(z) = \sin^2(\kappa z) e^{-\alpha z}. \tag{3.24}$$

It can be seen that the power does indeed transfer back and forth between the two guides as a function of length. From (3.23) and (3.24), it can be seen that the length L necessary for complete transfer of power from one guide to the other is given by

$$L = \frac{\pi}{2\kappa} + \frac{m\pi}{\kappa}, \tag{3.25}$$

where $m = 0, 1, 2, \ldots$. In a real guide, with absorption and scattering losses, β is complex. Hence, the total power contained in both guides decreases by a factor $\exp(-\alpha z)$.

The coupling coefficient κ is a strong function of the shape of the mode tails in the guides. For well confined modes, in which the overlapping of the

tails causes only a negligible perturbation of the basic mode shape, it can be shown that the coupling coefficient is given by

$$\kappa = \frac{2h^2 q e^{-qs}}{\beta W\left(q^2 + h^2\right)},$$

(3.26)

where W is the channel width, s is the separation, h and β are the propagation constants in the y and z directions, respectively, and q is the extinction coefficient in the y direction. It will be recalled that these parameters have been assumed to be identical for both waveguides. In a practical situation, it may be difficult to fabricate two identical waveguides to form a coupler. If, for example, the guides do not have exactly the same thickness and width, the phase velocities will not be the same in both. This will not necessarily destroy the coupling effect entirely. If the difference in phase constants $\Delta\beta$ is small, it can be shown that the power distributions in the two guides are given by

$$P_0(z)\cos^2(gz)e^{-\alpha z} + \left(\frac{\Delta\beta}{2}\right)^2 \frac{\sin^2(gz)}{g^2}e^{-\alpha z},$$

(3.27)

and

$$P_1(z) = \frac{\kappa^2}{g^2}\sin^2(gz)e^{-\alpha z},$$

(3.28)

where

$$g^2 \equiv \kappa^2 + \left(\frac{\Delta\beta}{2}\right)^2.$$

(3.29)

It can be seen from (3.27–3.29), that, in the presence of a phase constant difference $\Delta\beta$, transfer of power will still occur. However, the transfer will be incomplete, since (3.27) has no zeros for any z.

3.9. Tapered waveguides

Efficient low-loss optical coupling between waveguides made of dissimilar materials is important in providing seamless optical interfaces for heterogeneous photonic integration. Especially, optical coupling between optical fibers and integrated waveguides in the photonic integrated circuits (PICs) and their systems is of growing interest with the recent advances in integrated photonics. Optical fibers have many attractive properties as an optical waveguide, such as low transmission loss, low cost, flexibility, and electromagnetic immunity, and therefore can provide an ideal conduit between various optical components and systems, including PICs.

The integration density of optical waveguides in the PIC platform is closely related to the refractive index contrast between the core and the cladding materials. As waveguides based on high-index waveguide core materials with large index contrasts to the cladding have much smaller waveguide dimensions and bending radii, the integration density of such high-index PIC platforms can be significantly increased. As the demand for more compact integrated optical components increases, high refractive index materials, such as silicon and III-V compound semiconductors, are required for large-scale high-density PICs. For example, for single-mode operation at telecommunication wavelengths (e.g. 1550 nm), the cross-section of a silicon waveguide (a refractive index of 3.45) embedded in a silica cladding (a refractive index of 1.45) must be on a submicron scale (e.g. 400 nm×220 nm), and the corresponding guided optical mode profile is highly confined in the waveguide core with a typical mode area on the order of 0.1 μm^2. Because of the highly confined optical modes, a tight bending radius of less than 5 μm can be achieved, which results in ultracompact guided optical components while allowing more flexibilities in the photonic circuit layout.

In contrast to the small mode size of the high-index integrated waveguide, the typical mode-field diameters (MFDs) of standard single-mode fiber (SMF) at

telecommunication wavelengths are about 10 μm. The difference in the cross-sectional area between those two waveguides is about 1000 times, which poses a great challenge to efficient optical coupling. For comparison, fiber-to-PIC interfaces for silica-based planar lightwave circuits can achieve very high coupling efficiencies of up to −0.05 dB (98.9%) mainly because the integrated waveguides and optical fibers are made of the similar silica materials with low-index contrasts and they have similar waveguiding properties, such as mode profile and effective modal index.

Lateral, vertical and three-dimensional waveguide taper designs have been introduced to enlarge the effective MFD of the integrated waveguides with high-index core materials (e.g. III–V compound semiconductor and silicon). These types of intermediate coupling structures are also referred to as spot-size converters (SSCs or mode-size converters) and gradually increase the width and/or height of the integrated waveguide to provide a large terminating facet area up to 100 μm², which is close to that of the fiber core, as schematically shown in Fig. 3.13. The tapering angle of the SSC transition region needs to be carefully designed to prevent unwanted light coupling to the higher order modes for the tapered waveguide region. An extended waveguide facet dimension through the SSC can achieve an enlarged electromagnetic field profile, resulting in a higher mode overlap value (η).

Three-dimensionally tapered SSCs have been employed in the silicon-based PIC platform for efficient optical coupling between optical fibers and integrated silicon waveguides. Such SSC structures can be fabricated with conventional silicon waveguides on silicon-on-insulator (SOI) wafers using additional fabrication processing steps, such as polishing, etching and gray-tone shadow masking. As shown in a scanning electron microscope (SEM) in Fig. 3.13, *b*, a metal shadow mask fabrication process used to make a vertical waveguide taper. Starting from the silicon waveguide cross-section of 500 nm in height and 300 nm in width, a total taper length of 2 mm and a final waveguide width of 10 μm were achieved with a linear taper. The final waveguide width is

close to the standard single-mode fiber MFD. Amorphous hydrated silicon (α-Si:H) was deposited to extend the vertical mode dimensions in the laterally extended silicon waveguide. A measured net transmission loss of 0.5 dB was obtained at the input wavelength of 1.55 μm using a lensed fiber.
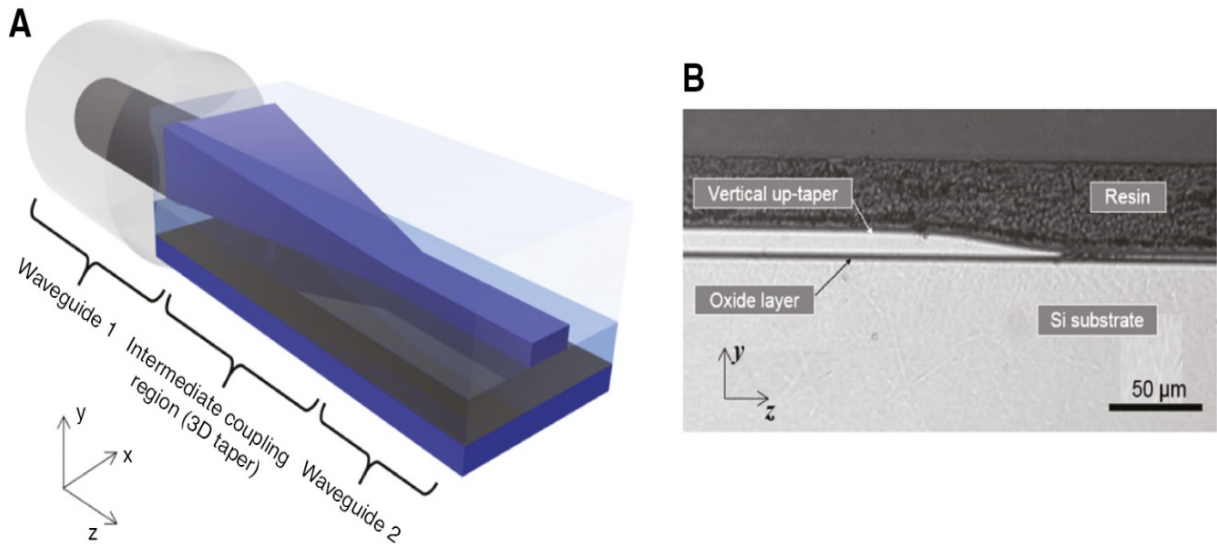
Fig. 3.13. Schematic diagram of a three-dimensionally tapered intermediate waveguide structure (*a*); an example of a vertically tapered waveguide fabricated by gray-tone shadow masking (*b*)

Although this type of guided mode expansion approaches offers good coupling efficiency of better than −1 dB and seems straightforward in concept, it requires additional dedicated fabrication steps, such as thick materials deposition and etching. It might also occupy more space when compared to other coupling schemes.

3.10. Enhanced optical coupling with intermediate waveguide regions

Optical mode profiles and their cross-sectional areas can be extended not only by expanding the waveguide dimensions with waveguide tapers as discussed in the previous section but also by shrinking the waveguide dimensions, as schematically shown in Fig. 3.14. When the waveguide cladding material has an intermediate refractive index (larger than air but smaller than the

waveguide core refractive index) and thus provide a smaller index contrast, the optical modes are less confined to the waveguide core as the core dimensions decrease, resulting in larger effective mode areas and smaller effective modal index. The low-index intermediate cladding layers, such as polymer, SiO_2 and Si_3N_4 or SiON with effective refractive indices comparable to optical fibers, can efficiently expand the optical mode dimensions and reduce the partial reflections at the interfaces . The input light gets adiabatically expanded through the inversely tapered high-index waveguide and can be finally coupled to the optical fiber with a high coupling efficiency.
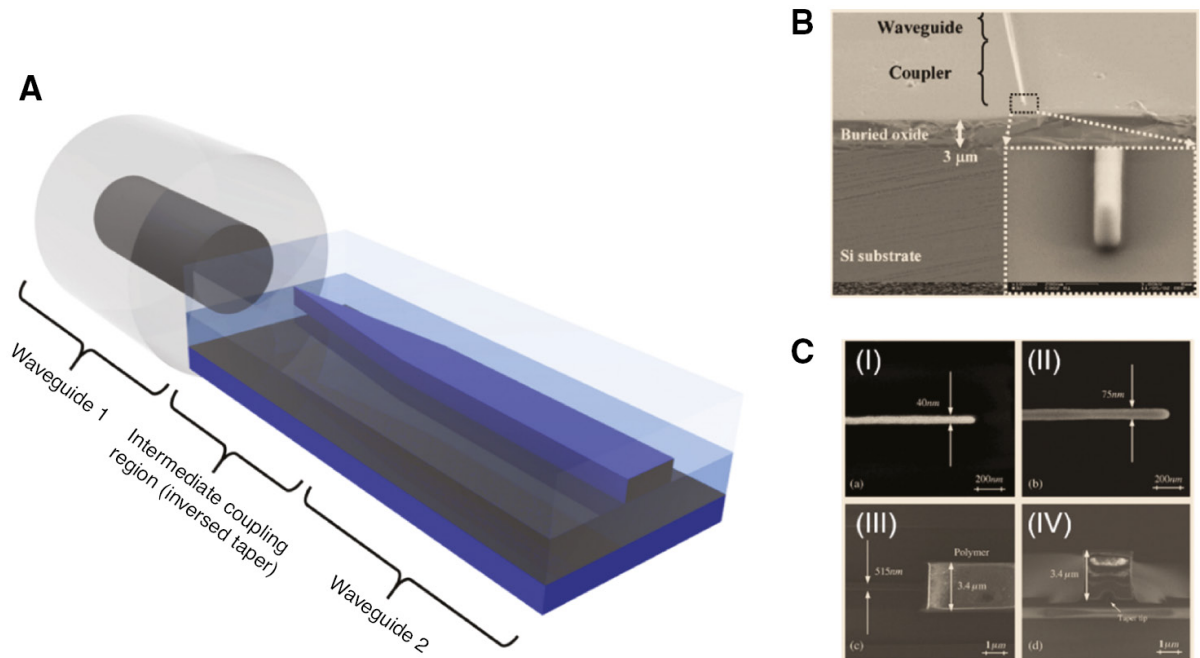
Fig. 3.14. Schematic diagram of an inversely tapered intermediate waveguide coupler with a low-index cladding. The integrated waveguide core region is represented with blue colors (*a*); a typical inversely tapered waveguide coupler with SiO_2 cladding layers (*b*); a coupler with a very narrow tip fabricated by silicon oxidation process and its SEM image (*c*). Silicon nano-taper tip end (I) before and (II) after oxidation process. (III) Top view of the junction point of a polymer core and a silicon waveguide. (IV) Cross-section view of the nano-taper coupler

Figure 3.14, *b* shows an example of an inversely tapered waveguide coupler, which transforms both the effective mode size and effective mode index between an optical fiber and an integrated silicon waveguide. The inversely

tapered silicon waveguide (smaller waveguide dimensions toward the terminating end) has a tapering length of ~40 μm, and the waveguide width reduces from ~470 nm (single-mode waveguiding condition) to ~100 nm with a waveguide height of 270 nm. Due to the expanded mode size at the subwavelength-scale waveguide tip, the maximum mode overlap with an input lensed fiber (MFD ≈ 5 μm) was calculated to be as high as 94%, and this corresponds to a mode mismatch loss of 0.26 dB. The total insertion losses with partial reflections and misalignment were measured to be ~3.3 dB and 6.0 dB for the TM (an effective refractive index of 1.51) and TE (an effective refractive index of 1.31) modes, respectively. Although there is a large index contrast between the silicon core material and the SiO_2 cladding layer, the propagating electromagnetic field for the very narrow inversely tapered waveguide structure resides mostly outside the core region and its effective modal index becomes close to the refractive index of the SiO_2 cladding layer.

Using a similar approach with a combination of an intermediate polymer region and a SiO_2 cladding layer, very low optical coupling losses of ~0.36 dB and ~0.66 dB for the TM and TE modes, respectively have been (Fig. 3.14, c). The ultra-low losses were obtained using a very narrow silicon taper tip whose width was reduced from ~40 nm to ~12 nm by silicon oxidation.

To further improve the fiber-to-chip optical coupling efficiency with intermediate waveguide structures, suspended waveguide structures with low-index cladding materials, such as SiO_2 and SiON, were investigated, as schematically shown in Figure 3.15, a. Figure 3.15, a shows a tapered silicon waveguide, surrounded by a SiO_2 cladding cantilever structure, is suspended. After formation of the tapered silicon waveguide and SiO_2 cladding cantilever structure (13 μm×10 μm), the underlying Si substrate was etched to form a trench, as shown in Figure 3.15, a. The optical power leakage to the substrate was minimized by providing a sufficiently large trench depth. The coupling efficiencies were measured to be −1.5 dB and −1.0 dB for the TM and TE modes, respectively. As the suspended waveguide structures have an alignment

tolerance of approximately ±2.0 μm for a 1 dB excess loss, which is better than other end-fire coupling schemes, it may have benefits in the fiber assembly and packaging processes. Multi-layered suspended tapering structures can further improve the coupling efficiencies by gradually transforming the effective refractive index between high-index silicon waveguides and low-index optical fibers.
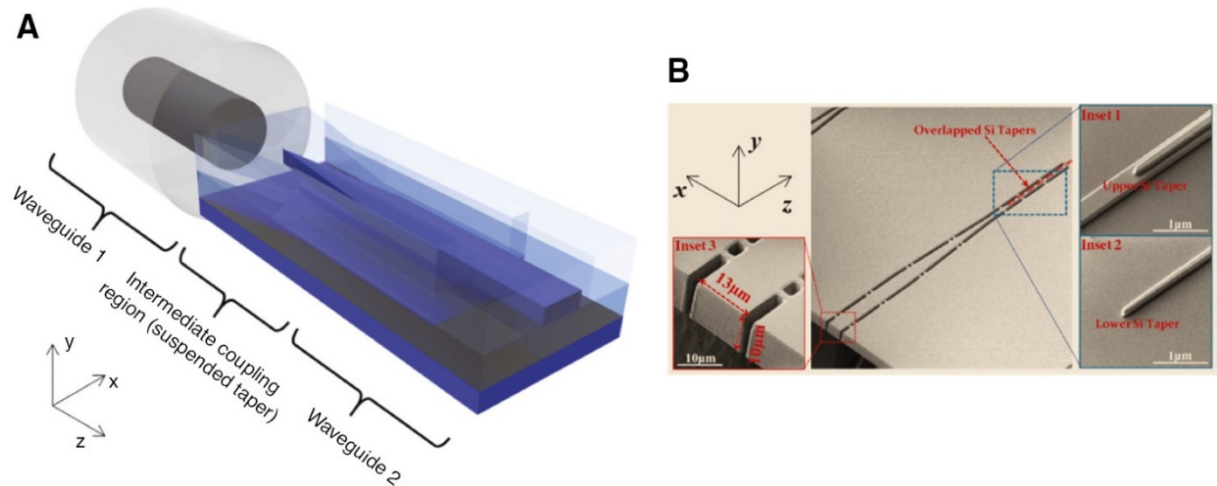
Fig. 3.15. (A) Schematic diagram of a suspended intermediate waveguide structure. (B) An example of an inversely tapered silicon waveguide surrounded by a SiO$_2$ cladding with a cantilever structure. The SEM images of the fabricated couplers are shown

This type of end-fire coupling with inversely tapered high-index waveguides can achieve efficient optical coupling with small footprints and low wavelength/polarization dependencies. Similar approaches in this category include waveguide ribbon layers photonic wire bonding and free-form lenses and mirrors. Figure 3.16, *a* and *b* show examples of single-channel SSC interconnects. Multi-port coupling examples using photonic wire bonds and free-form components are illustrated in Figure 3.16, *c* and *d*, respectively. Recently, several researchers have exploited subwavelength grating (SWG) converters based on high-index materials, such as silicon. By engineering the effective refractive index of the waveguides, SWG-based end-fire couplings can achieve coupling efficiencies as high as −0.32 dB (93%).
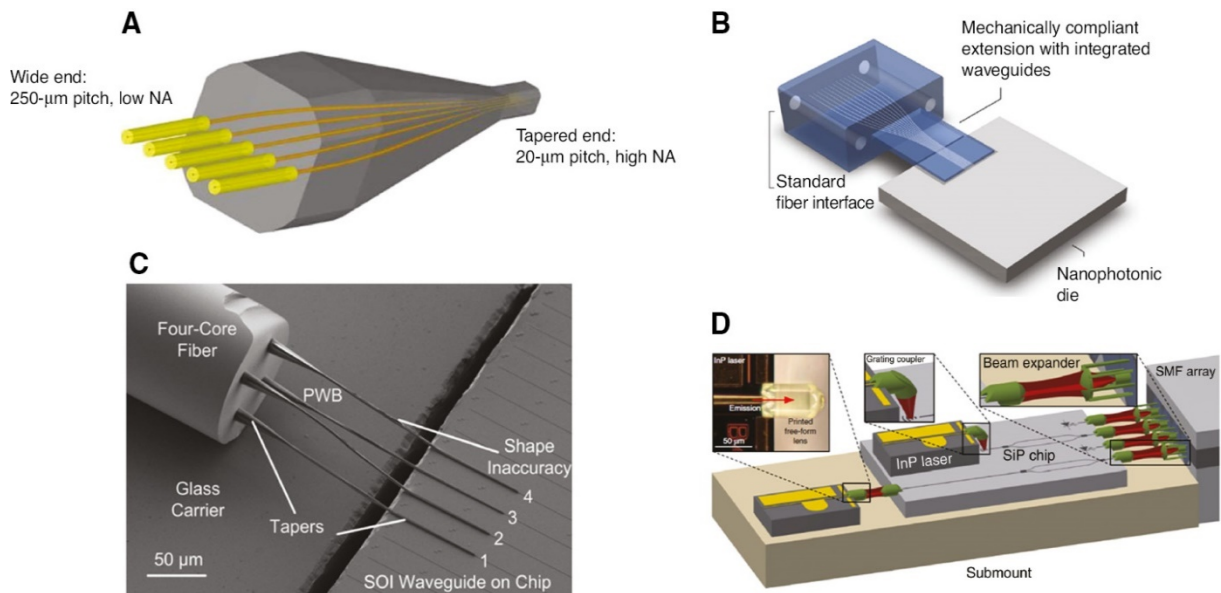
Fig. 3.16. Optical interfaces for multiple ports using intermediate mode transformation stages. (A) Schematic diagram of a multi-channel SSC interconnecting multiple fibers and waveguide arrays. (B) Schematic diagram of the polymer interface linking an array of standard SMFs to an array of nanophotonic waveguides. (C) SEM image of a multi-core fiber-to-chip interface using photonic wire bonds to connect the individual cores of the multi-core fiber to an array of silicon waveguides. (D) Schematic diagram of free-form lenses and expanders

3.11. Vertically curved and suspended waveguide structures

One of the main disadvantages of the conventional end-fire coupling methods is that the optical interfaces to the input/output fibers need to be located at the PIC chip edges and the dicing process is required to expose the end-fire coupler structures to the optical fibers. However, the out-of-plane light coupling can be achieved by physically bending the waveguides toward the upward direction from which the optical fibers are approached, as schematically shown in Figure 3.17, *a*. While providing the advantages of conventional end-fire coupling, such as low wavelength and polarization dependencies, this out-of-plane coupling scheme does not require the optical couplers to be located at the chip edges.
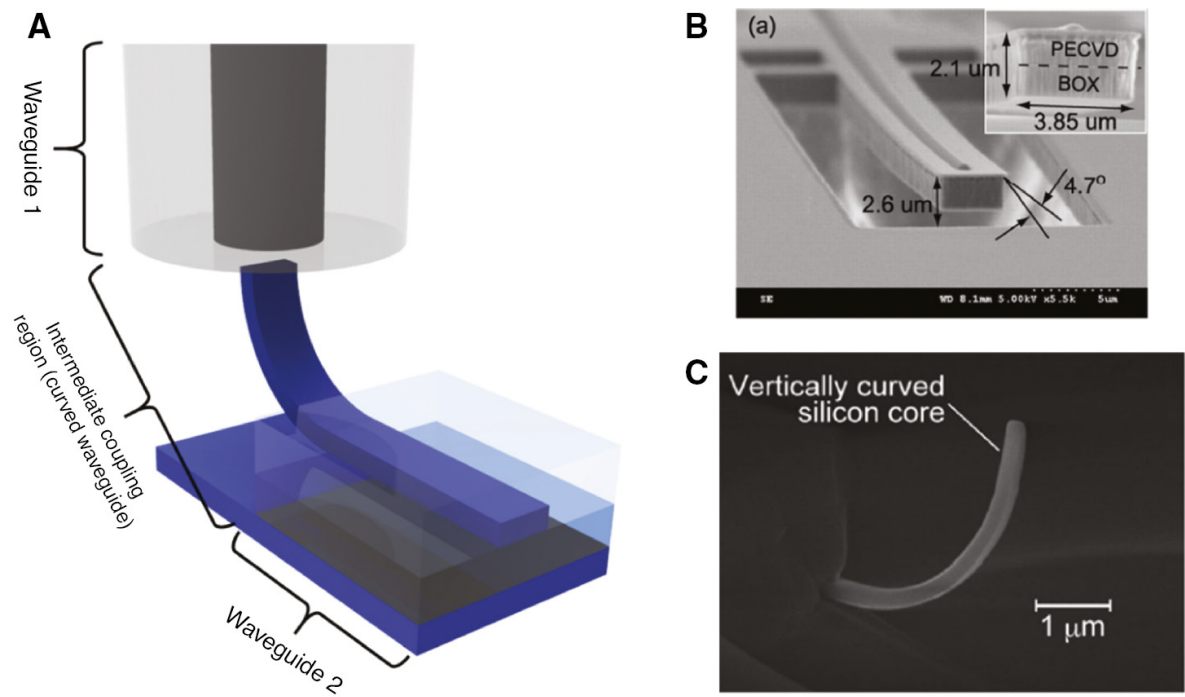
Fig. 3.17. (A) Schematic diagram of a vertically curved intermediate waveguide structure. (B) SEM image of a curved coupler using thermal annealing. (C) SEM image of a bent waveguide coupler with a radius of curvature of 3 μm using ion implantation

The bent waveguide couplers can be fabricated by introducing stress gradients to the waveguide structure so that the waveguide termination region can be bent upward when released from the substrate. The overall footprint for the bent waveguide coupler is closely related with its bending radius. As highly confined optical modes in the high-index waveguides allow small bending radii of less than 10 μm without too much radiation losses occurring, the vertically curved waveguide coupler design can be compact and scalable while providing the benefits of end-fire coupling.

For example, out-of-plane optical coupling to the optical fibers or free-space optical beams was obtained by stress-engineered Si/SiO_2 cantilevers deflected upward by thermal annealing stress, as shown in Fig. 3.17, *b*. Tapered silicon waveguides were embedded in the center of the SiO_2 cantilevers. After the deposition of PECVD SiO_2 layer, the SiO_2 cantilevers are released from the substrate by reactive ion etching and subsequently annealed for stress control. Maximum deflection and tilt angles at the end of a 40-μm-long SiO_2 cantilever

with silicon core in the center are 2.6 μm and 4.7°, respectively. Because of limitations in maximum stress level from the thermal annealing process, the bending radii obtained from this method were limited to approximately 100 μm. To introduce higher stress gradients and thereby reduce the bending radius of the curved waveguide structure for smaller coupler dimensions, an ion beam implantation method was applied to intentionally introduce lattice defects. As a result, a bending radius of approximately 3 μm was obtained, as shown in Fig. 3.17, *c*. The vertically curved waveguide coupler structures showed ≈1 dB (80%) optical coupling efficiencies with tapered fibers and lensed fibers.

4. WAVEGUIDE LENSES

Optics can not do without elements that perform operations such as integration, focusing, collimation, Fourier transforms, and others, as usually performed by conventional lens optics. In integrated optics, these devices should remain planar, and in their constructive basis should be a waveguide as the main component. Such lens devices, which are technologically and functionally integrated with a microwaveguide, belong to integrated optical focusing devices, analogues of conventional lens optics.

In integrated optics, focusing elements belong to its basic components, which are also made (like waveguides) using integrated technology. In integrated optics they provide formation of specified phase frontal optical beams, Fourier transforms, etc.

Focusing systems are one of the main functional units of integrated optical spectrum analyzers, correlators, and a number of other integrated optical systems for optical information processing.

It is well known that lens (focusing) properties are possessed by such systems in which different parts of the light beam pass through different optical paths. In integrated optical waveguides, this can be achieved by changing either the film thickness or the composition of the waveguide material, which causes a change in its refractive index. It is possible to obtain a lens effect even when the thickness and composition of the waveguide material remain unchanged. In this case, it is necessary to apply a waveguide film to a substrate of complex shape. If, in this case, a local change in the thickness or refractive index is programmed in a limited volume of the waveguide layer due to a change in the substrate profile, then an optical focusing element – a planar lens – is built into the waveguide.

The first experimental versions of planar lenses were created by changing the effective refractive index of the waveguide layer with a change in its

thickness in a section that mimics the shape of a cross section of a conventional bulk lens.

Such lenses possessed all kinds of optical distortions – aberrations characteristic of their bulk analogues. The resolution of such planar lenses is limited primarily by spherical aberration (its two-dimensional analogue) and coma aberration – this is due to their poor focusing properties and blurred image. These attempts once again confirmed that the similarity of integrated optics with a bulk lens is only formal. It cannot be constructively used to create high-quality planar focusing systems, which are based on all the same waveguide properties of integrated optical components.

Planar lenses currently used in integrated optics can be divided into three types (Fig. 4.1): Luneberg lenses (a), geodesic lenses (b), and diffraction grating type (c). Regardless of which of the listed types a lens belongs to, it must satisfy a number of general requirements for this type of integrated optics elements. First, the size of the focal spot of the lens should be as small as possible, limited only by the diffraction limit. Secondly, the focusing quality should not depend on the width of the input beam within the input aperture and, finally, the structure of the lens should be technologically advanced and provide the possibility of replication without compromising the basic characteristics.
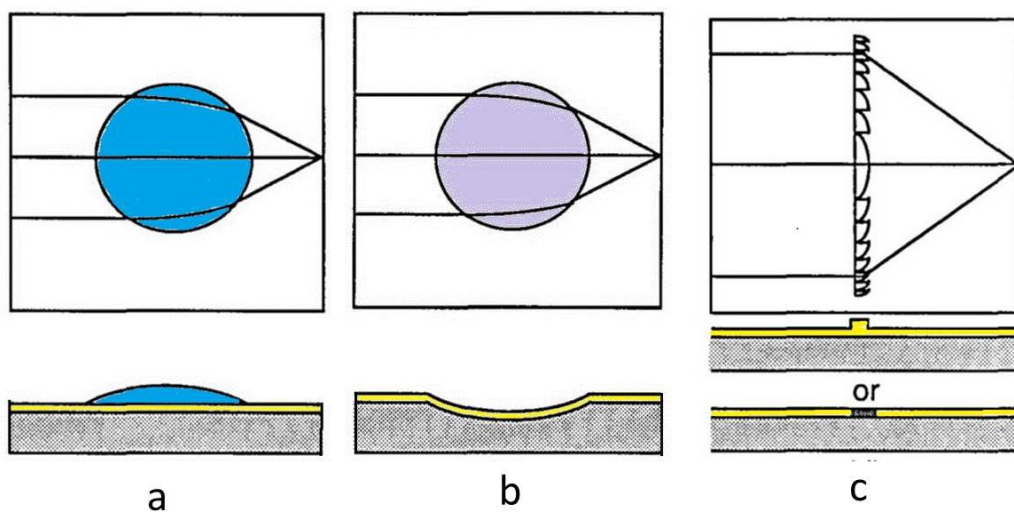
Fig. 4.1. Top and side views of standard types of waveguide lens: Luneburg (a), geodesic, (b) and diffractive (c)

92

One of the qualitative characteristics of lenses in integrated optical information processing devices is the magnitude of the relative level of the side lobes of the intensity diagram on the focus line. The better the lens, the more radiation is concentrated in the main lobe on its optical axis.

4.1. Luneberg lenses

This type of lens is named by analogy with the well-known Luneberg lens antennas, widely used in the microwave technology. An antenna is a sphere with such variation of the refractive index, which provides focusing of the incident parallel beam of rays on the opposite side of the lens.

There are several configurations of the waveguide layers of the integrated optical focusing element according to the Luneberg lens type. The most common structure is a substrate with a refractive index n_4, on which two waveguide layers with different refractive indices are successively applied: n_3 and n_2 (Fig. 4.2).
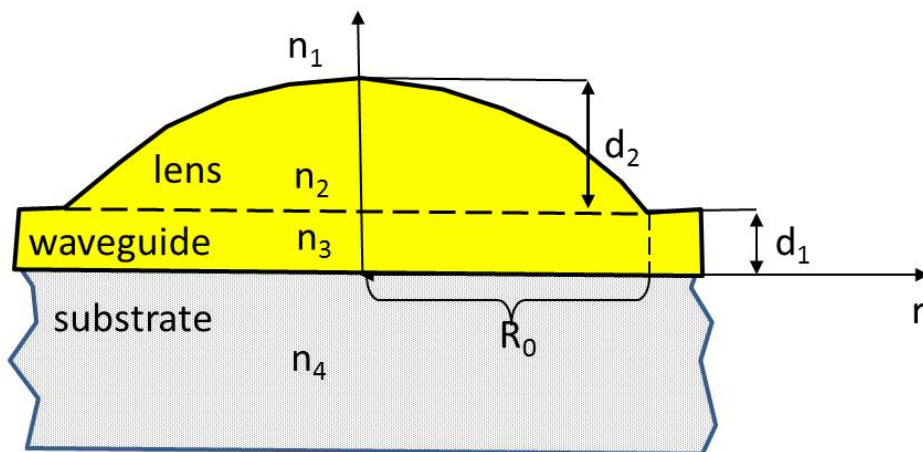
Fig. 4.2. Cross section of thin-film Luneberg lens

The upper waveguide layer n_1 has a thickness profile that creates a gradient of the effective waveguide refractive index. The focusing properties of such a multilayer system are determined by the specified gradient of its effective refractive index. the higher the focusing and resolution of the lens, the greater the change in the refractive index of the waveguide layer and the substrate. For

this, set the refractive index of the substrate to be minimal, and the waveguide layer as large as possible. Actually, a Luneberg lens is a structure in which the upper waveguide layer has circular symmetry. Examples of typical materials combinations that make it possible to obtain the necessary characteristics of film lenses – Nb_2O_5, ZnS, Ta_2O_5 – for lenses ($n = 2...3$) and waveguide layers in glass ($n \sim 1.5$).

In practice, there are often problems which there is no need to use the entire aperture of the lens – for example, when a device on a large-diameter lens works with narrow light beams. Then, when creating several lenses with different focal distances, you can use one mask, and the spraying time, which determines the thickness of the layer, can be set in accordance with the formed lens. So, to create several lens with a radius of 4 mm, limiting the aperture to 25% one can use a single mask to obtain a series of lenses with different focal lengths and a diffraction limited focal spot.

The circular symmetry of Luneberg lenses eliminates such type of aberration as coma (due to complete symmetry for any oblique beam). In addition, in the case of circuits for which the size is strictly limited, the Luneberg film lens together with a system of mirrors makes it possible to use several light beams simultaneously in different directions. Nevertheless, high-quality Luneberg film lenses require special measures to compensate for spherical aberration.

To compensate for spherical aberration Luneberg profile of the effective refractive index must satisfy the integral equation

$$n_{eff} = (r,m) = n_{eff}^0 (m) \exp\left[\omega(\rho, S)\right],$$

where $\omega(\rho, S) = \dfrac{1}{\pi} \displaystyle\int_{\rho}^{1} \dfrac{\arcsin x / S}{\sqrt{x^2 - \rho^2}} dx$, $\rho = \dfrac{r n_{eff}(r,m)}{R_0 n_{eff}^0(m)}$, R_0 is the lens radius; r is the radial coordinate from the center of the lens along the surface of the waveguide;

$n_{eff}(r,m)$, $n_{eff}^0(m)$ are the effective refractive indices of the mode with number m in the waveguide and in the lens, respectively; $S = F / R_0$ is the normalized focal length of the lens.

For integrated-optical focusing elements, as a rule, $S > 1$, which corresponds to the so-called inverted Luneberg lens. In this case, the focus of the lens is located outside, and

$$n_{eff}(r,m) = \frac{n_{eff}^0(m)\exp\left\{\left[n_{eff}^0(m)\right]^2 R_0^2 - r^2 n_{eff}^2(r)\right\}^{\frac{1}{2}}}{\pi n_{eff}^0(m) F}. \qquad (4.1)$$

Based on the results obtained for similar microwave lenses, an algorithm was developed for calculating the normalized n_{eff} profile for the Luneberg waveguide lens. The constructive parameter is the lens material profile over the thickness $d_2(r)$, which is determined by the numerical solution of the dispersion equation for a multilayer waveguide. The result of calculating the normalized n_{eff} and thickness profile for a Luneberg lens depending on the normalized radius of the lens is shown in Fig. 4.3.
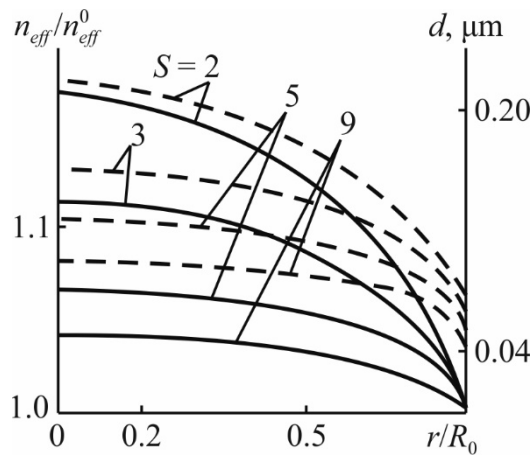

Fig. 4.3. Normalized effective refractive index and thickness profile of the Luneberg lens as a function of normalized focal length

Any deviation of the real profile of the effective refractive index from the calculated one leads to a deterioration in the characteristics of the lens. This requires the use of precision technology in its manufacture. Deviation from the calculated profile within a percent leads to a multiple (up to 5 times or more) excess of the focal spot size over the diffraction minimum. For example a lens with a radius of 5 mm from Ta_2O_5 on a waveguide made of Corning-7059 glass, changing its profile thickness by $1 \cdot 10^{-7}$ mm leads to a change in focal length by 0.25 mm. Permissible changes in the depth of field of waveguide film lenses in integrated optics do not exceed several tens of micrometers; otherwise, the very idea of integration and miniaturization of integrated optical circuits is lost.

Currently, the main technological method for fabrication a given lens profile is spraying through a mask. Attempts to use simple non-contact masks, composite masks with rotating parts, and masks with conical holes were unsuccessful. To ensure the necessary accuracy, a machine calculation of the cross-sectional shape of the mask hole as an integral part of the spraying process model as a whole was required. The lens thus obtained retains the diffraction limit of the focal spot when using up to 11% of its entire aperture.

Using a precision technology is due to the high sensitivity of the characteristics of Luneberg lenses to the technological variation of their thicknesses. This sensitivity can be weakened if the characteristics of the lens are calculated taking into account the characteristics of the modes of the waveguide components of the circuit.

Consider a lens with $n_2 = n_3$ (see Fig. 4.2). Such a lens is an asymmetric ridge waveguide, the behavior of TE modes in which is described by the equation

$$\tan\left(K_2 d + m\pi\right) = \frac{K_2\left(\gamma_1 + \gamma_3\right)}{K_2^2 - \gamma_1\gamma_3},\tag{4.2}$$

where $\gamma_{1,3} = k_0\left(n_{eff}^2 - n_{1,3}^0\right)^{\frac{1}{2}}$; $K_2 = k_0\left(n_2^2 - n_{1,3}^0\right)^{\frac{1}{2}}$; $m = 0,1,2\ldots$

The same relations are obtained for TM modes. Typical dispersion curves corresponding to equation (4.2) for the TE0 and TE1 modes are shown in Fig. 4.4. The section shown by the continuous line corresponds to the single-mode waveguide mode, which is typical for integrated optic focusing devices.
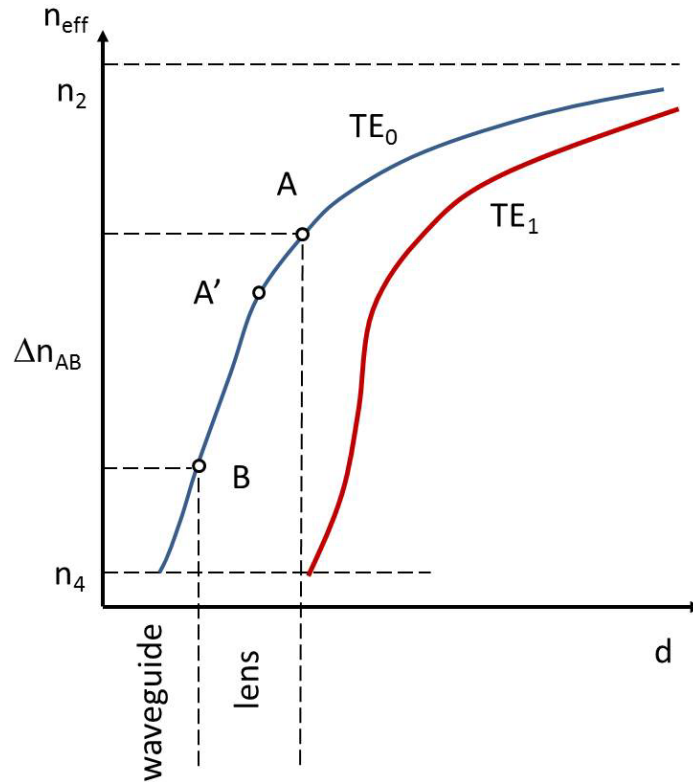
Fig. 4.4. Effective refractive index as a function of waveguide thickness (single mode)

The thickness of the waveguide layer is chosen such that the operating point B is simultaneously sufficiently far from the cutoff and from the upper boundary of the single-mode mode. The length of the BA section provides a change in the refractive index of the lens. A complete change in the refractive index Δn_{BA} depends on the selected focal length F. The smaller F, the greater the change in the refractive index Δn_{BA} and the thicker the lens in accordance with (4.1). For a fixed value of n_2 in the case of a telephoto lens, the working section on the dispersion curve BA' will be shorter than for a short focus BA. Since the slope of the dispersion curve at point A' is greater than at point A, the sensitivity of the focal length to a change in the thickness of the lens for longer-focus lenses is higher. The sensitivity can be reduced if select the minimum possible

n_2, which still allows to realize the full necessary change in the refractive index of the lens. The value of $k_0 d$ at the cutoff point of the TE mode is determined by the equation

$$k_0 d = \frac{1}{\sqrt{n_2^2 - n_3^2}} \left[\arctan\left(\sqrt{\frac{n_3^2 - n_1^2}{n_2^2 - n_3^2}} \right) + \pi \right] \tag{4.3}$$

which allows to calculate the optimal refractive index n_2. The latter does not depend on the radiation wavelength, but is a function of the parameter S. As the focal length increases, the value of n_2 decreases and approaches the refractive index of the substrate (for a silicon substrate, $n_4 = 1.47$). The joint solution of equations (4.2) and (4.3) allows us to present the calculated dispersion curves of the TE0 mode in the form of $f(n_2) = \tan K_2 d - \dfrac{K_2(\gamma_1 + \gamma_3)}{K_2^2 - \gamma_1 \gamma_3}$, the illustration of which is presented in Fig. 4.5.
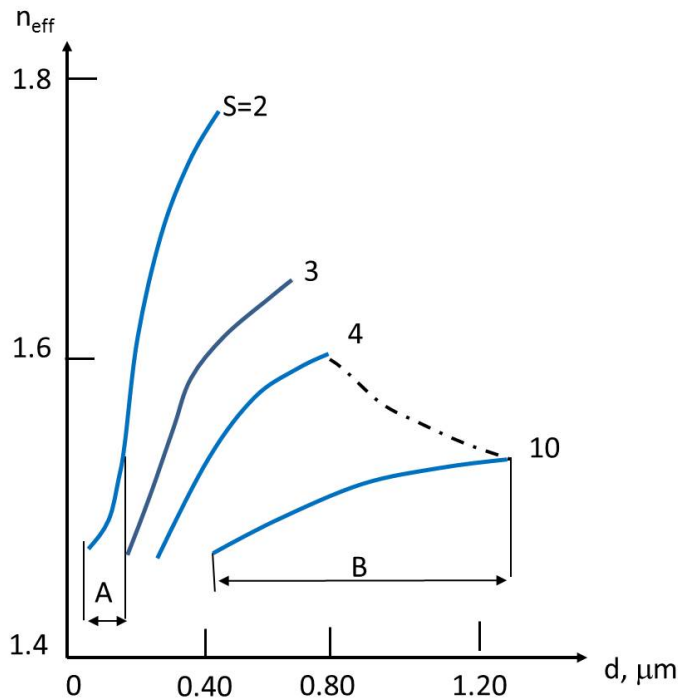

Fig. 4.5. Dispersion curves of the TE0 mode for different S values ($\lambda_0 = 0.9$ µm)

For optimal n_2 and λ_0 values of d in the graphs of Fig. 4.5 are also optimal. In this case, larger telephoto lenses correspond to larger waveguide thicknesses, in contrast to the case when n_2 is not optimized. In Fig. 4.5, the optimization effect of n_2 is clearly visible, which is especially pronounced for a lens with a large S. Thus, for $S = 10$ with an optimal n_2, the interval of permissible thicknesses B within which the lens profile can be realized is relatively large. Consequently, the technological deviation of thicknesses within these limits does not catastrophically degrade their characteristics. With increasing n_2 for smaller S (for example, $S = 2$), the interval of permissible thicknesses A decreases. In addition, the slope of the dispersion curves at the upper boundary points of interval A significantly increases compared to interval B. As a result, with increasing n_2 and decreasing S, the steepness of the dependence of F on d increases significantly.

Figure 4.6 shows the dependences of the normalized focal length S on the thickness of the lens in its central part, calculated for various optimal n_2 values. From the above curves it follows that the minimum sensitivity F to a change in thickness is observed in the region of relatively small values of S corresponding to the optimal values of n_2. At the points of the optimal regime corresponding to the minimum value of S, the slope for each curve remains approximately constant. Thus, for the optimized structure, the steepness $\Delta S/\Delta d$, which determines the dependence of F on d, does not depend, to a first approximation,

on S, $\dfrac{\Delta S}{\Delta\left(\dfrac{d}{\lambda_0}\right)} = \dfrac{\Delta S \lambda_0}{\Delta d}$, whence $\Delta S/\Delta d \sim 1/\lambda_0$. In order to minimize the

dependence of F on d, this leads to the choice of longer wavelength radiation as the working one for the device.

In order to the permissible limits of variation in the lens thickness does not lead to a loss of resolution, a change in the focal length $\Delta F_{\max}$ should not exceed its focus depth

$$\Delta F = \left(\frac{\alpha_i \lambda_0}{n_{eff}} \right) \left(\frac{F}{\alpha} \right)^2 \le \Delta F_{max} \qquad (4.4)$$

where α_i is the dimensionless coefficient associated with the intensity profile of the input beam; and α is the input aperture of the lens.
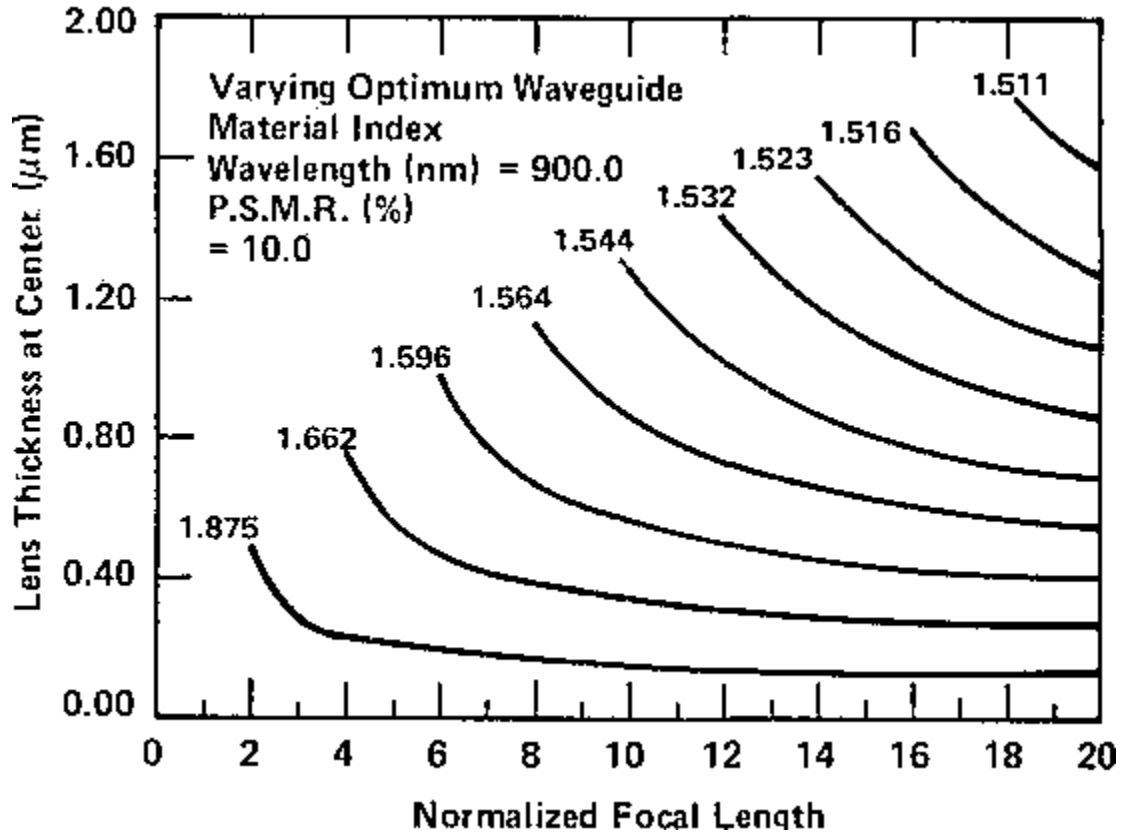
Fig. 4.6. Dependences of the normalized focal length S on the thickness of the lens in its central part, calculated for various optimal n_2

It follows from expression (4.4) that with an increase in the focal length of the lens, the restrictions imposed on the permissible changes in its thickness become less stringent. According to the definition of the normalized focal length

$\Delta F_{max} = \Delta S \dfrac{D}{2}$, where D is the diameter of the lens.

After substitution in eq. (4.4) we have

$$\frac{1}{2}\Delta SD \Big/ \left(\frac{\lambda_0 \alpha_i}{n_{eff}}\right)\left(\frac{F}{a}\right)^2 \le 1$$

which gives

$$\frac{\Delta S}{S} \le \frac{F\alpha_i \lambda_0}{\alpha^2 n_{eff}}. \tag{4.5}$$

Eq. (4.5) determines the permissible relative changes in the focal length of the lens $\Delta S/S$ from its main parameters, due to both geometric factors and the characteristics of the material and the device in which the lens operates.

A graphical solution of Eq. (4.5) for the transforming lens of an integrated optical spectrum analyzer is shown in Fig. 4.7. Its aperture is uniquely determined by the required frequency resolution $\Delta fa = v_a / \Delta f$ where v_a is the speed of the acoustic wave in the acousto-optic region. The lower the frequency resolution of the analyzer, the larger the aperture of the lens, which, in accordance with Eq. (4.5), also determines the smaller the tolerances for its thickness, the higher the precision of its manufacturing technology. The focal length is mainly limited by the size of the substrate. Dashed lines limit the allowable range of lens parameters, beyond which defocusing will reduce the frequency resolution of the analyzer. The choice of material for the required value n_2 is limited by the level of optical losses in it. For glasses (a very common lens material) $n_2 = 1.56$ (Corning 7059). To this value n_2 in accordance with Fig. 4.6 corresponds the optimal value $S = 8$. For a silicon substrate, $v_a = 3200$ m/s (for frequency resolutions 1.0 MHz $a = 3.2$ mm, for 1.5 MHz $a = 2.1$ mm), $F = 40$ mm and $n_{eff} = 1.41$. The value of F at a fixed value of S determines the diameter of the lens.

To minimize lateral aberrations at the edges of the lens caused by imperfection of its profile, the technological diameter of the lens is set approximately twice as large as the calculated aperture. For the $S = 8$ defined above, the minimum value is $D = 10$ mm. For such a lens at $\Delta f = 1$ MHz, the

accuracy of reproducing the calculated thickness in the center of the lens should not be lower than ±6.3 nm, which is 0.84% of the maximum thickness. In this case, the diffraction limited spot should be retained at $F = 40$ mm. Such accuracy is technologically realized at very low spraying rates of 1–5 nm/min, which makes it possible to obtain Luneberg lenses with characteristics satisfactory for integrated optics.
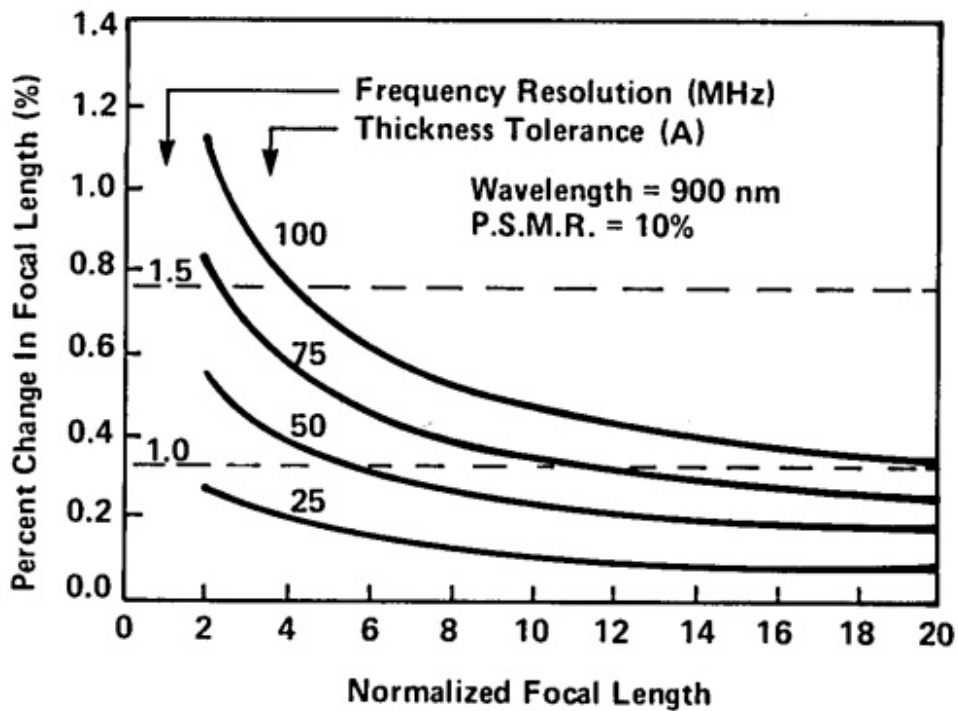
Fig. 4.7. Relative change in the focal length of the lens as a function of *S* for various tolerances of its thickness Δd of the frequency resolution of the spectrum analyzer ($\lambda = 0.9$ μm)

However, the process of their fabrication is still laborious, which greatly limits their application. One of the main disadvantages of Luneberg waveguide lenses is the difficulty in selecting material for them if a semiconductor or active dielectrics with relatively high refractive indices ($n = 2$–3) are used as a substrate. For such substrates $n_L > n_s$. At the same time, the lens material should have low losses in the operating wavelength range. The coupling of a lens with a waveguide, which has a high refractive index, is difficult, which complicates its technology. As such coupling devices, tapering multilayer films are used.

Luneberg lenses require correction of spherical aberration, which is associated with additional restrictions on the parameters of the mask, spraying speed and composition of the sprayed material. All this led to the search for other solutions and structures for creating focusing elements of integrated optics devices.

4.2. Focusing elements based on geodesic lenses

Geodesic lenses are based on the concept of configuration lenses that was developed for the application of microwave antennas. In integrated optics, they are one of the alternatives to Luneberg lenses. A geodetic lens is a depression or elevation of a spherical or aspherical shape in the substrate, formed by optical grinding and polishing. The waveguide in the substrate is created after making the desired relief in it. The waveguide layer is formed by diffusion and ion-exchange processes in the substrate or by depositing the corresponding waveguide films on the surface of the substrate. Optical polishing at the boundary of the plane of the waveguide and the lens reduces optical loss and is one of the stages of the lens manufacturing technology. The lens is based on the Fermat principle. The shortest optical path for the ray entering the lens will follow the geodesic line of the curved surface that forms the lens. The optical path length at the center of the lens will be longer than at its edges. Since the refractive index of the waveguide layer is constant over the entire length, the phase front behind the lens will be curved towards its axis. The constancy of the parameters of the waveguide layer determines the absence of chromatic aberration for geodetic lenses. This feature allows the use of geodetic lenses in multimode devices.

The axis of symmetry of the lens depression (Fig. 4.8) is perpendicular to the plane of the waveguide, and its edge forms a circular contour. The ABCEF curve corresponds to the axial path, and the WXYZF curve corresponds to the path of incident rays at the boundary of the lens at a distance $R_c \sin\phi$ from its axis. The focal length of the lens is determined from the equation

$$F = R_0 \left[-\frac{\sin \varphi}{\sin 2(\varphi + \theta)} \right] \tag{4.6}$$

where R_0 is the radius of the depression in the plane of the waveguide. The angles φ and θ are interconnected by the relation known from the methodology for calculating geodesic lines

$$\theta = \left(1 - \frac{e^2 h^2}{R_1^2} \right)^{-\frac{1}{2}} \left[\left(\frac{R_1}{h} \right) \Pi \left(\mu_1, 1 - \frac{R_1^2}{h^2}, r \right) - \left(\frac{e^2 h}{R_1} \right) F(\mu_1, r) \right] \tag{4.7}$$

where $\mu_1 = \arcsin \left[\dfrac{R_0^2 - h^2}{R_1^2 - h^2} \right]^{\frac{1}{2}}$; $r = e \left[\dfrac{R_1^2 - h^2}{R_1^2 - e^2 h^2} \right]^{\frac{1}{2}}$; $F(\mu_1, r)$; $\Pi \left(\mu_1, 1 - \dfrac{R_1^2}{h^2}, r \right)$ are the elliptic integrals, of the first and third kind respectively;

$e = \left[1 - \left(\dfrac{R_2^2}{R_1^2} \right)^2 \right]^{\frac{1}{2}}$ is the eccentricity of the generating ellipse; R_1, R_2 are the axis

of the ellipse; $h = \left(\dfrac{n_{eff}^0}{n_{eff}} \right) R_0 \sin \varphi$; n_{eff}^0; n_{eff} are the effective refractive index of

the waveguide outside and inside the lens depression, respectively.

In Figure 4.9 shows the dependences of the normalized focal length on the normalized aperture calculated in accordance with expressions (4.6) and (4.7) for several types of lenses. From the above patterns it follows that geodesic lenses of a spherical shape ($e = 0$) are characterized by significant spherical aberration, expressed in the strong dependence of the focal length on the aperture. For small values of a, the focal lengths tend to their paraxial focus

$F_0 = \dfrac{R_0}{2(1 - \cos \varphi)}$, which depend only on the lens parameters.
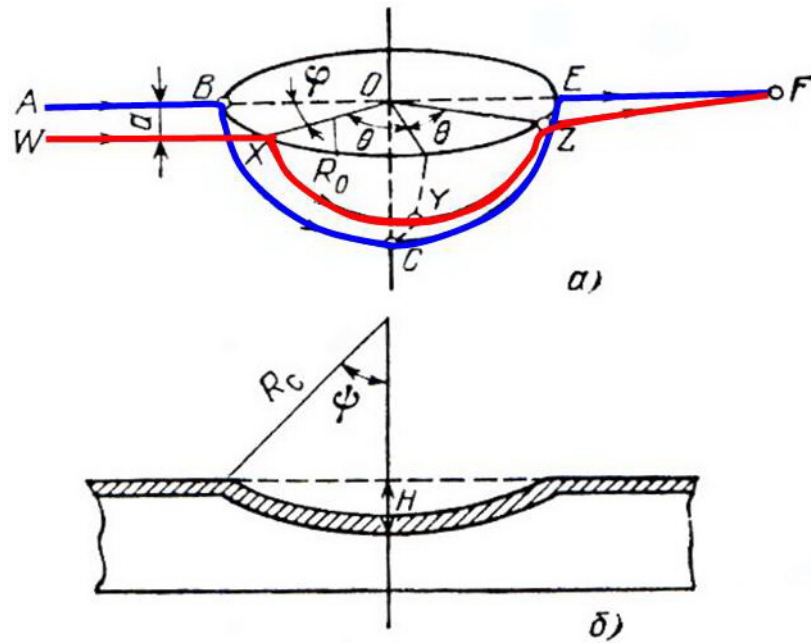
Fig. 4.8. Waveguide geodesic lens: *a* is the ray path, *b* is the cross-section
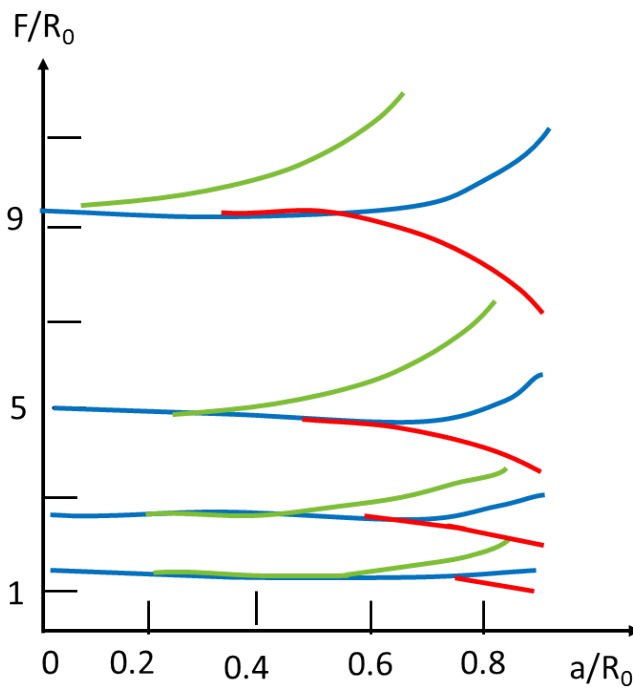
Fig. 4.9. Normalized focal length versus normalized aperture calculated in accordance with eqn. (4.6) and (4.7) for several types of lenses. ------ spheroidal and ------ spherical lens, respectively; ------ spherical lens with n_{eff} correction

    For an accurate assessment of the ultimate non-aberrational aperture sizes and the shape of the focal spot for a spherical lens, the calculation of aberrations is not necessary. In this case, the aberration function is determined

$$\Phi\left(\frac{a}{R_0}\right) = \Phi(x) = \Delta - \Delta_0 \qquad (4.8)$$

which is the difference in the optical paths of an infinitely thin ray entering the lens at a distance a from the axis in the ideal (all rays intersect at one point) Δ_0 and the real Δ .

As an example in Fig. 4.10 shows the calculation results for $\Phi(x)$ (curves 1, 2) for lenses made on glass and $LiNbO_3$ substrates. In both cases, large radius lenses are characterized by lesser aberration. Function (4.8) allows to calculate the distribution of light intensity near the focus of the lens. According to the Huygens-Fresnel principle, the distribution of light intensity in the transverse and longitudinal directions near F_0 of the geodesic lens is determined by the relations:

$$\frac{I(\xi)}{I_{max}} = \left\{ \int_{-a_1}^{+a_1} \exp\left[ik\left(\Phi(x) + \frac{\xi x}{F_0} \right) \right] dx \right\}^2 ,$$

$$\frac{I(\zeta)}{I_{max}} = \left\{ \int_{-a_1}^{+a_1} \exp\left[ik\left(\Phi(x) + \frac{\zeta x}{2F_0} \right) \right] dx \right\}^2 ,$$

where $2a_1$ is the aperture of the incident wave; k is the wave number of the wave in the waveguide; ξ, ζ — coordinates in the transverse and longitudinal directions with the center at the focal point of the lens; I_{max} is the maximum light intensity in the distribution.

Unless special measures are taken, the distribution of light intensity in the focal spot (at $\zeta = 0$) only for small apertures has the form close to the function of $\left(\sin\frac{\zeta}{\xi}\right)^2$. With an increase in the used aperture, an increase in the intensity side lobes and maximum width compared to the diffraction limit, which means a

deterioration in the optical quality of the lens take place. Fig. 4.10 shows a graph of the function $\Gamma_1 = \Gamma_1(x)$ (curve 3), which characterizes the excess of the central maximum of the light field distribution in the focal spot of the glass lens above the level of the first side lobe. The quality of such a lens is very low – a sharp deterioration in quality is already observed at a relative aperture of $x = 0.15$.
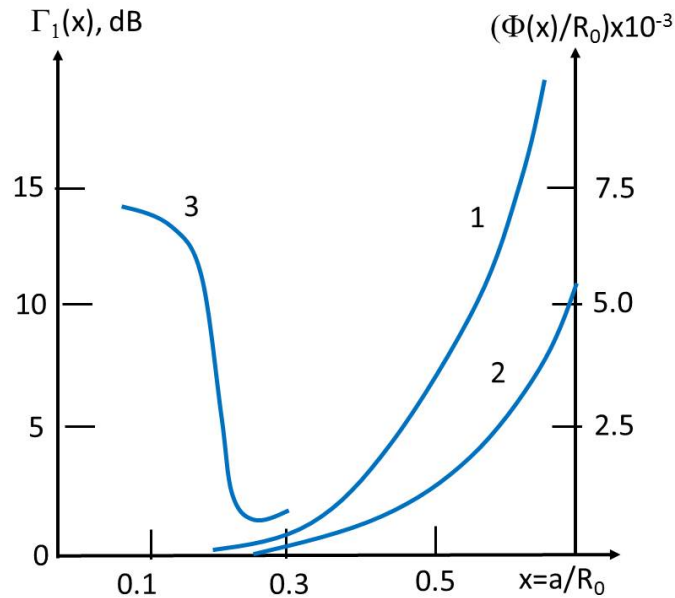
Fig. 4.10. Aberration characteristics (curve 1,2) and function $\Gamma_1(x)$ (curve 3) of a geodesic lens. 1,3 – glass K8, $R_0 = 8$ mm, $H = 0.8$ mm; 2 - LiNbO₃ – $H = 0.7$ mm

The presence of significant aberrations can lead not only to a deterioration in the quality of the diffraction spot at the paraxial focus, but also to a change in the light intensity along the axis. In this case, a shift of the focal point relative to F_0, the appearance of a diffraction focus, and even the formation of several focal spots are possible. To reduce aberrations in geodetic lenses, various compensation methods are used: either a spheroidal shape of the lenses, or a compensating thin film lens with an opposite aberration is formed in the region of the depression. The latter can be realized by thickening the waveguide layer inside the lens contour or by applying a material different from that for the waveguide refractive index. The smallest aberration is characterized by

spheroidal lenses: they allow the use of up to 50% of the aperture without compromising the focusing properties (see Fig. 4.9).

Another approach to calculating the shape of the non-aberrational lens is based on the fact that, knowing the distribution of n_{eff} for the Luberberg non-aberrational lens, one can calculate, based on the principle of asymmetric waveguides, the shape of the geodesic analogue:

$$H = -H_0 + \int\limits_0^\rho \left[\frac{\rho}{n_{eff}} \frac{dn_{eff}}{dr} \left(\frac{\rho}{n_{eff}} \frac{dn_{eff}}{dr} - 2 \right) \right]^{1/2} d\rho, \qquad (4.9)$$

where H_0 is the maximum depression of the lens; $\rho = r n_{eff}(r)$ is the radial coordinate; n_{eff} is the effective refractive index of the Luneberg lens. Equation (4.9) is solved by numerical methods. As a result, we obtain the profile of the geodesic lens $H(\rho)$ with parameters corresponding to its film counterpart.

Current technology allows to obtain lenses with a focal spot size close to the diffraction limit with a stable focal length F and using more than 90% of the entire aperture. In most cases, the quality of the lens is assessed by the ability to produce a calculated profile of the depression in the substrate and to control the uniformity of the waveguide thickness with an accuracy of the order of the wavelength. An analysis of modern machining methods shows that only diamond processing allows for super-micrometric accuracy of the profile and depth with the simultaneous high optical quality of the machined surfaces.

Geodetic lenses are compatible with substrates of active dielectrics ($LiNbO_3$, $LiTaO_3$), which are widely used in integrated optics devices.

4.3. Diffractive lenses

Diffraction lenses are one of the promising directions in creating high-quality focusing devices for integrated optics. They are technologically simpler

than geodesics and Luneberg lenses. Their technology can be attributed to the basic technology of integrated optics elements. Its basis is short-wave UV photolithography and electronic lithography. Electronic lithography allows to process large areas (≤ 10 mm^2), with a resolution of 0.3 microns and form almost any lens pattern. The most promising planar diffraction lenses include Fresnel lenses, Bragg lenses and lenses with a chirp-grating.

4.3.1. Fresnel lenses

In the classic Fresnel lens, focusing is due to two effects: Fresnel diffraction; introducing into the diffraction plane an appropriate phase shift or absorption to enhance interference at the focal point. In the waveguide lens, an essential phase shift is obtained by changing the effective refractive index of the waveguide, changing the parameters of the waveguide layer (d or n), or by applying metallization or a dielectric coating to the waveguide. Focus enhancement of diffraction can also be obtained by enhancing the absorption of that part of the radiation that destroys interference at the focal point. With increasing absorption, the efficiency of the lens decreases simultaneously, which makes the phase shift method more preferable.

The necessary phase shift between the focal point and any point in the lens plane is given by the equation

$$\Phi = \left(x^2 + F^2 \right)^{\frac{1}{2}} - F , \qquad (4.10)$$

where x is the distance from the axis of the lens to a point on the plane of the lens. If the lens plane is divided into such rectangular zones for which the phase factor at a distance x_m from the axis is

$$\Phi = \frac{m\lambda}{2}, \quad m = 1, 2, 3\ldots \tag{4.11}$$

Then the Fresnel amplitude lens is obtained by absorbing power in all odd zones. Accordingly, a phase lens is obtained by a phase shift equal to π in odd zones. Absorption or phase shift in odd zones is due to the deposition or etching of the corresponding areas of the waveguide (Fig. 4.11).
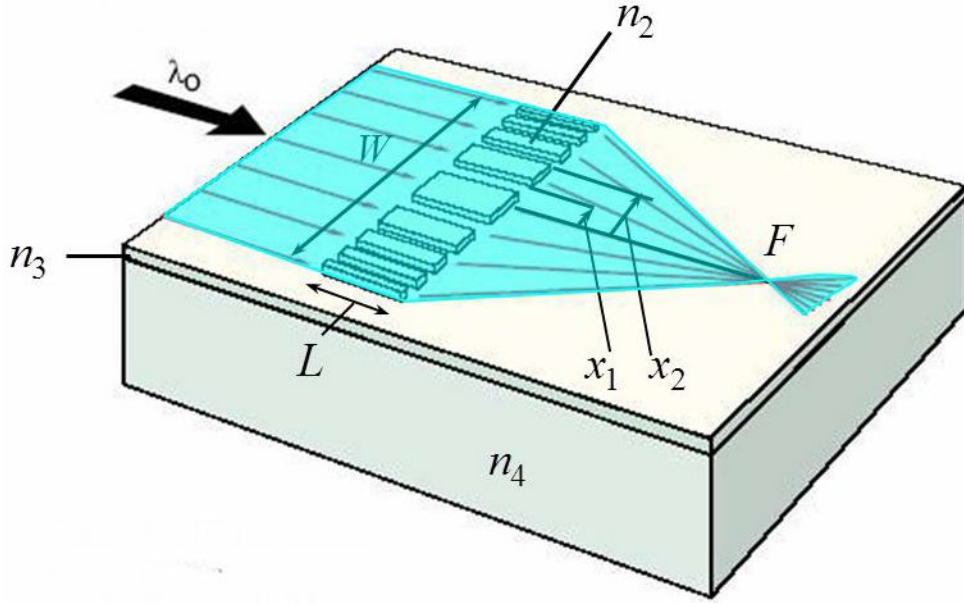
Fig. 4.11. Waveguide Fresnel lens

The boundaries of these zones are determined by the joint solution of equations (4.10) and (4.11) $x_m = (m\lambda F)^{1/2}$; $\lambda \ll F$; $m = 1, 2, 3\ldots$ and the minimum distance between the zones $d_{\min} = x_m - x_{m-1} \approx F\lambda / W$, where $W = 2(MF\lambda)^{1/2}$ is the width of the lens; M is the number of zones. The lens length L shown in Fig. 4.11 or the length of the zone plates can also be determined from the phase equation $\dfrac{2\pi L \Delta n_{eff}}{\lambda} = \pi$, where Δn_{eff} is the difference between the effective refractive indices in the waveguide outside and in the lens area, respectively.

If the lens parameters are known, then the distribution of the intensity in its focal plane is determined by summing the diffraction fields from all zones of the lens.

The main disadvantage of Fresnel waveguide phase lenses is their relatively low diffraction efficiency (≤0.35%). Low diffraction efficiency is a consequence of the stepwise nature of the change in the refractive index of the lenses along its width W. This disadvantage is partially eliminated in the so-called gradient lenses, in which the parameters of the waveguide layer in the Fresnel zone change according to a certain law.

For lenses whose refractive index is a gradient parameter, its change in the Fresnel zones is given by the equation $n(x) = n_{\min} + \Delta n \left[\dfrac{\Phi_F(x)}{2\pi} + 1 \right]$, where $n_{\min}$ and Δn are minimum of the refractive index and the amplitude of its change in the zones respectively; $\Phi_F(x) \approx -\left(\dfrac{k}{2F} \right) x^2 + 2m\pi$ is the modulation of the phase when x varies within $x_m \leq |x| \leq x_{m+1}$;

$$x_m = \left(\frac{2m\lambda F}{n_{\min}} \right)^{\frac{1}{2}}; \quad k = n_{\min} k_0 \tag{4.12}$$

is the coordinate of the boundary of the Fresnel zones. From eqn. (4.12) it follows that in gradient lenses the focal length depends only on the geometric dimensions, therefore, it can be determined with a high degree of accuracy. Calculated intensity distribution curves on the focal line as a function of the normalized distance from the optical axis of the lens are shown in Fig. 4.12. The parameter is the quantity $l = L/L_F$, which characterizes the deviation of the lens thickness L, caused by technological spread, from its optimal value $L_F = \lambda/\Delta n$.
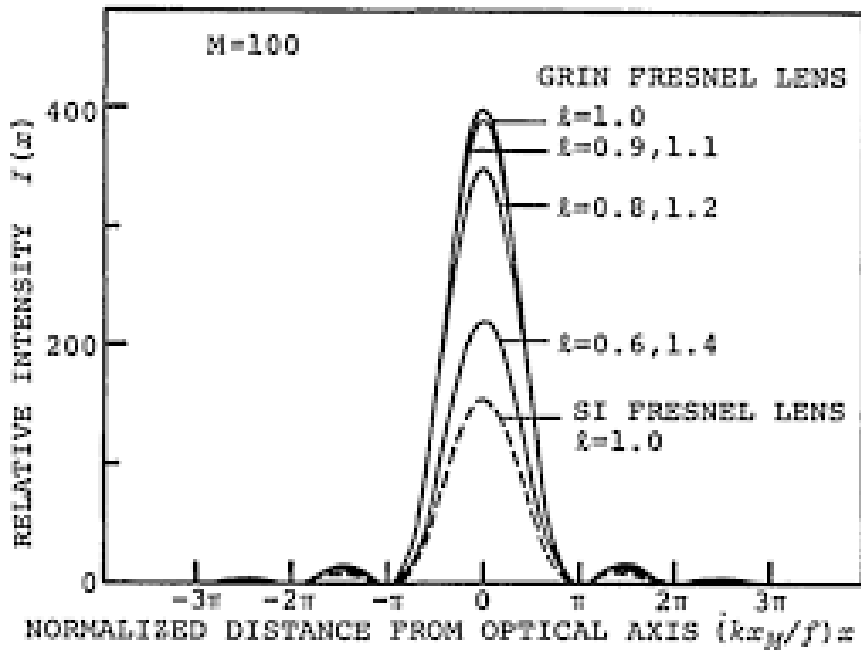
Fig. 4.12. Light intensity profiles on the focal line of Fresnel lenses

From comparison, the curves of Fig. 4.12, it follows that for Fresnel gradient lenses, the magnitude of the focal spot remains within the diffraction limit even with significant deviations of L from the optimal value ($l = 0.6; 1.4$). The technological spread ($l \neq 1$) leads only to a change in the maximum intensity in focus, reducing the diffraction efficiency of the lens, determined by the expression

$$\eta = \int\limits_{-\sigma}^{+\sigma} I(x)\,dx\,/\,2x_m.$$

Fig. 4.13, *a* shows the comparative dependences of $\eta(l)$ for gradient and step index Fresnel lenses. The maximum efficiency for gradient lenses reaches 90%, which is about 2-2.5 times higher than the η_{max} for step index lenses. For Fresnel gradient lenses, the dependence $\eta(l)$ is very weakly expressed; a change in l within $|1-l| \leq 0.25$ causes a decrease in η_{max} by about 30%. This is another advantage of gradient focusing structures, important for their use in integrated optical circuits.
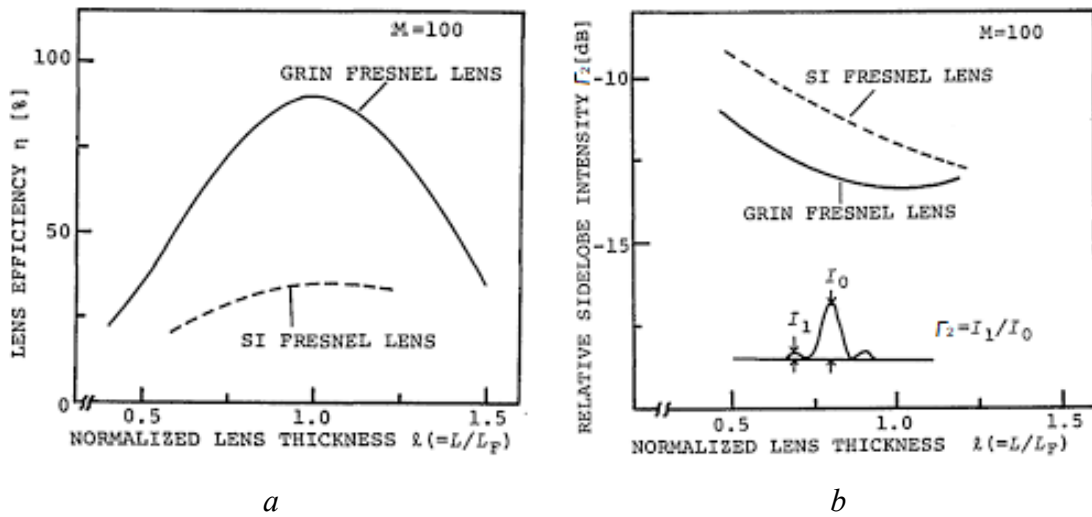
Fig. 4.13. Dependence of diffraction efficiency (*a*) and level of the side lobes of the intensity (*b*) upon the phase modulation for gradient and step index Fresnel lenses

One of the qualitative characteristics of lenses in integrated optical information processing devices is the level of the side lobes of the intensity diagram on the focus line. The better the lens, the more radiation is concentrated in the main lobe. With exact observance of the lens thickness ($l = 1$), the value $\Gamma_2 = \dfrac{I}{I_0}$ is −13.3 and 11.9 dB for gradient and step-index lenses, respectively (Fig. 4.13, *b*).

Thus, Fresnel gradient lenses are superior in all characteristics to similar lenses with a step- index profile, which determines their predominant use in integrated optical focusing devices. Realization of high-performance waveguide lenses in active dielectrics (LiNbO$_3$, LiTaO$_3$) has been strongly desired since they are considered to be one of the most suitable waveguide materials for optical IC's.

One of the disadvantages of diffraction lenses, which in some cases limits their use in integrated optics , is a significant deterioration in performance when working with inclined beams.

113

4.3.2. Bragg chirp-grating lenses

The Bragg lens is obtained as a result of the interference of a plane and two-dimensional cylindrical wave converging at a point. The phase difference of these two waves at $z = 0$ can be written as

$$\Phi(x,z) = k\left[\sqrt{x^2 + (F-z)^2} - (F-z)\right], \qquad (4.13)$$

where k is the wave vector of the waveguide. If $\Phi(x, z)$ is given by eqn. (4.13) then the necessary change in the effective refractive index is given by

$$\Delta n_{eff}(x,z) \propto \cos\left[\Phi(x,z)\right]. \qquad (4.14)$$

Bragg diffraction occurs if the structure is sufficiently long in z direction. Therefore, Bragg lenses are also called thick-lattice lenses, in contrast to Fresnel lenses, which are thin-lattice lenses.

For a Bragg lens, it is necessary to ensure the high accuracy of Δn_{eff} change in accordance with eqn. (4.14). Such a change can be obtained on a rectangular lattice. The slope and position of the center of each of its strips is determined by eqn. (4.13) with $\Phi = 2m\pi$, where m is the strip number. It has been experimentally established that such a grating diffracts efficiently at a certain ratio between the grating period Λ and the wavelength of incident light. If the lattice period changes according to the required law, then the plane wave diffracts into a converging cone, creating a focusing effect (Fig. 4.14). In this case, the angle between the direction of the incident and diffracted beams is given by:

$$\theta_m = 2\arcsin\left\{\frac{m\lambda}{2n_{eff}\Lambda(x)}\cos\left[0.5(\theta_i - \theta_d)\right]\right\}$$

where θ_i, θ_d are the incidence and diffraction angles, respectively;

$$\Lambda(x) = \frac{m\lambda}{2n_{eff}}\left\{\sin\left(\frac{\theta_m}{2}\right)\cos\left[\frac{1}{2}(\theta_i - \theta_d)\right]\right\}^{-1} \qquad (4.15)$$

is the variable period of the lattice. With perfect focusing for a collimated beam (θ_i = const), the angles θ_m and θ_d are proportional to arctan(x/F).
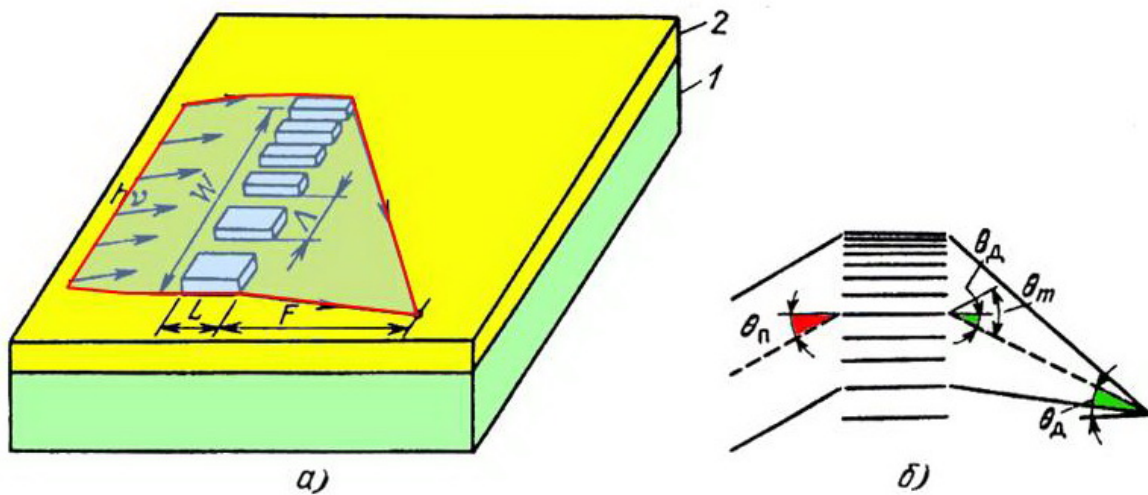
Fig. 4.14. Bragg lens with variable-period grating, *a* – structure; *b* – scheme for determining angles for calculating a variable period; *1* is the substrate; *2* is the waveguide

The focal length of the lens based on the Bragg grating with a variable step depends on its geometrical dimensions, in particular, on $\Lambda(x)$. The parameters that determine the diffraction efficiency (but do not affect F) are the width L and the depth d_1 of the grating. Maximum lens performance is realized when

$$\Delta n_{eff}(d_1)L = \pi. \qquad (4.16)$$

In this case, the interaction length (lattice width) must satisfy the condition

$$L \leq \frac{5.4 n_{eff} \Lambda_m^2}{\lambda} \qquad (4.17)$$

where Λ_m is the grating period at the high-frequency edge of the lens. Equations (4.15) – (4.17) allow us to calculate all the necessary parameters of the lens according to their given characteristics. Bragg lenses on variable gratings are based on passive (glass) and LiNbO$_3$ active substrates and are characterized by an almost diffraction limited focal spot at a diffraction efficiency of $\geq 75\%$. The high characteristics of the Bragg lenses and their manufacturability make them competitive with gradient lenses for a wide class of integrated optics devices.

Despite the fact that all types of discussed above lenses require a sufficiently high level of technology for their fabrication, the achieved characteristics allow the use of such lenses in practical devices of integrated optics. Today there is still no such universal design that would satisfy all the requirements, and each of the focusing elements considered has its own advantages and disadvantages. Their predominant use in a particular case is determined by the peculiarity of the task to be solved, and by the technology that can be implemented.

5. WAVEGUIDE MODULATORS

High-performance optical information systems are directly related to the problem of increasing the maximum signal frequency. The development of high-speed injection lasers with the highest speed in the technological sense is a very difficult task.

Industrial injection lasers have a modulation band that, as a rule, does not exceed several tens of gigahertz. Modulation of radiation with a frequency exceeding the cutoff frequency of direct laser modulation in integrated optics is carried out by an external modulator. Original solutions in the development of integrated optical modulators and radiation switches are due to the peculiarities of the propagation of radiation in microwaves. The creation of waveguide modulators made it possible to reduce the power consumed by them per unit width of the modulation band by several orders of magnitude to expand the frequency band of electro-optical modulators by more than an order of magnitude compared to bulk counterparts. The aforesaid applies equally to waveguide switches, the family of which has been replenished with representatives who do not have volume analogues. The principle of operation of waveguide modulators and radiation switches is essentially the same. The difference is determined by the operating modes and design features due to their specialization.

The operation of integrated optical modulators is based on the effect of refractive index change in waveguide under the influence of various external fields: electric, magnetic, and acoustic. Depending upon the external means used, we have electro-optic, acousto-optic, or magneto-optic modulators. All integrated optical modulators can be divided into two large groups. The first includes devices based on planar waveguides, the second – based on three-dimensional waveguides. The use of strip or channel waveguides instead of planar can reduce the consumed electrical power by more than an order of

magnitude and provide greater design flexibility in the development of integrated optical circuits and devices.

In terms of the physical and technical characteristics, the most promising are electro-optical modulators (EOM). These include modulators based on the effect of tunneling light pumping (modulators on coupled waveguides) and interference type modulators. Modulators based on coupled waveguides have, as a rule, a rather complicated configuration of electrodes, to which voltages of different amplitudes are applied to optimize the coupling length and adjust propagation constants in coupled waveguides. All this determines the criticality of the tunnel type modulators to the manufacturing conditions and operation mode. Interference-type modulators are preferable: their technology is less critical, and temperature characteristics are more stable. At interference-type modulators, a frequency band up to 200 GHz is reached.

5.1. General principles of a waveguide light modulators

Basic Operating Characteristics of Switches and Modulators

One important characteristic of modulators and switches is the modulation depth, or modulation index, η. In the case of an intensity modulator in which the applied electrical signal acts to decrease the intensity of the transmitted light, η is given by

$$\eta = \frac{I_0 - I}{I_0},$$ (5.1)

where I is the transmitted intensity and I_0 is the value of I with no electrical signal applied. If the applied electrical signal acts to increase the transmitted light intensity, η is given by

$$\eta = \frac{I - I_0}{I_m},$$ (5.2)

118

where I_m is the transmitted intensity when maximum signal is applied. The maximum modulation depth, or extinction ratio, is given by

$$\eta_{\max} = \frac{I_0 - I_m}{I_0} \quad \text{for} \quad I_m \leq I_0 \tag{5.3}$$

or by

$$\eta_{\max} = \frac{I_m - I_0}{I_m} \quad \text{for} \quad I_m \geq I_0. \tag{5.4}$$

It is also possible to define the modulation depth for phase modulators, as long as the phase change can be functionally related to an equivalent intensity change.

For the case of interference modulators, it can be shown that the modulation depth is given by

$$\eta = \sin^2\left(\frac{\Delta\varphi}{2}\right) \tag{5.5}$$

where $\Delta\varphi$ is the phase change.

Modulation depth has been defined for intensity modulators (and indirectly for phase modulators); however, an analogous figure of merit, the maximum deviation of a frequency modulator, is given by

$$D_{\max} = \frac{|f_m - f_0|}{f_0} \tag{5.6}$$

where f_0 is the optical carrier frequency, and f_m is the shifted optical frequency when the maximum electrical signal is applied.

Another important characteristic of modulators and switches is the bandwidth, or range of modulation frequencies over which the device can be

operated. By convention, that bandwidth of a modulator is usually taken as the difference between the upper and lower frequencies at which the modulation depth falls to 50% of its maximum value. In the case of a switch, frequency response is usually given in terms of the switching speed, or switching time. The switching time T is related to the bandwidth Δf by the expression

$$T = \frac{2\pi}{\Delta f}.$$ (5.7)

Minimizing switching time is most important when large-scale arrays of switches are used to route optical waves over desired paths. Similarly, modulation bandwidth is a critical factor when many information channels are to be multiplexed onto the same optical beam. Thus, the unusually fast switching speed and wide bandwidth of waveguide switches and modulators, which will be discussed later in this chapter, make them particularly useful in large telecommunications systems.

Insertion loss is another important characteristic of optical switches and modulators that must be known for system design. Insertion loss is generally stated in decibels, and for the case in which the modulating signal acts to decrease the intensity, it is given by

$$\alpha_i = 10 \log \left(\frac{I_t}{I_0} \right)$$ (5.8)

where I_t is the optical intensity that would be transmitted by the waveguide if the modulator were absent, and I_0 is the intensity transmitted with the modulator in place, but with no applied signal.

For a modulator in which the applied signal acts to increase the transmitted intensity, the insertion loss is given by

$$\alpha_i = 10 \log \left(\frac{I_t}{I_m} \right) \tag{5.9}$$

where I_m is the transmitted intensity when maximum signal is applied. Insertion loss is, of course, an optical power loss. However, it ultimately increases the amount of electrical power that must be supplied to the system, since higher power optical sources must be used.

Electrical power must also be supplied to drive the modulator or switch. In the case of modulators, the required drive power increases with modulation frequency. Hence, a useful figure of merit is the drive power per unit bandwidth, $P/\Delta f$, usually expressed in milliwatts per megahertz. A key advantage of channel-waveguide modulators is that they have a significantly lower $P/\Delta f$ than that required for bulk modulators.

The power requirements of optical switches operating at high clock rates, for example to time-division multiplex a number of different signals, can be evaluated in much the same way as is used for modulators. Hence, $P/\Delta f$ would still be a useful figure of merit in that case. However, if switching is done at relatively slow rates, a more important quantity is the amount of power required to hold the switch in a given state. An ideal switch would consume significant power only during the change of state; holding power would be negligible. Since electro-optic switches require the presence of an electric field to maintain at least one state, they could not be called ideal in that respect. However, except for leakage current, little power is needed to maintain a field in the small volume of a waveguide switch.

The degree of isolation between various inputs and outputs of a switch or modulator is a major design consideration. In a modulator, the isolation between input and output is merely the maximum modulation index, as defined previously. However, it is usually expressed in decibels when used to specify isolation. In the case of a switch, the isolation between two ports (either input or output) is given by

121

$$\text{isolation [dB]} = 10\log\frac{I_2}{I_1} \qquad (5.10)$$

where I_1 is the optical intensity in the driving port, and I_2 is the intensity at the driven port when the switch is in the off state with respect to port 1 and 2. Thus, a switch with a signal leakage, or crosstalk, of 1% with respect to two ports would have -20 dB isolation.

5.2. The electro-optic effect

The fundamental phenomenon that accounts for the operation of most electrooptic modulators and switches is the change in index of refraction produced by the application of an electric field. In the most general case, this effect is nonisotropic, and contains both linear (Pockels effect) and nonlinear (Kerr effect) components.

The manner in which the indicated change in the refractive index is converted to intensity modulation depends on the geometry of the particular device. The linear electro-optical effect Δn is related to the amplitude of the applied field by the equation $\Delta n = -\left(\dfrac{n^3}{2}\right)r_{ij}E$, where n is the refractive index in the absence of a field; E is the strength of the applied electric field; Δn is the electro-optical coefficient which is a tensor quantity. As a result of the action of the electric field, light passing through the distance l in the electro-optical material will accumulate a phase change $\Delta\varphi = \dfrac{2\pi\Delta nl}{\lambda}$, which leads to phase modulation of the radiation.

Materials for waveguide EOM must have the largest electro-optical coefficient and allow the fabrication of waveguide structures based on them with low losses. Active dielectrics, include lithium niobate and lithium tantalate, which have pronounced nonlinearity and high efficiency of electro-optical

conversion, satisfy these requirements in the best way. In addition to lithium niobate and tantalum, which are ferroelectrics, semiconductor materials such as $A^{III}B^V$, which are promising materials for integrated optics and high-speed integrated optoelectronic devices, have been widely used recently for waveguide modulators.

The electro-optical coefficients, which determine, with all other things being equal, the value of Δn, depend on the crystallographic orientation of the material used to create the modulator. For anisotropic LiNbO$_3$ crystals with 3m point symmetry, the optimal orientations for the waveguide structure are y- and z-cut (Fig. 5.1). In order to obtain maximum electro-optical coefficients, the control electric field should be directed along the z axis. When using the y-cut (Fig. 5.1, a), the electrodes are located symmetrically on both sides of the waveguide, which determines the longitudinal component of the electric field E modulating. Modulators on the z-cut (Fig. 5.1, b) to maintain the direction of the modulating field along the z axis, apply one of the electrodes on top of the waveguide, and the vertical component of the field is working. Since the waveguide modes decay in the region along the metal electrodes, the direct deposition of electrodes on the optical waveguide leads to additional losses and and change in the characteristics of the modulator. To reduce losses between the waveguide and the metal electrode, a buffer layer with a low refractive index is introduced. As a buffer layer, Al$_2$O$_3$ ($n = 1.6$) or SiO$_2$ ($n = 1.5$) films with a thickness of 0.15–0.2 µm are used. A buffer layer is especially necessary for the working TM mode, which is polarized perpendicular to the plane of the substrate.

Modulator with lumped electrode configuration (Fig. 5.1, a), the electrodes form a concentrated modulator capacitance. Its bandwidth is limited by the product of this geometric capacity and load resistance and light propagation time through the modulator. For modulators on LiNbO$_3$, the limiting is the modulator capacitance, which should be minimized.
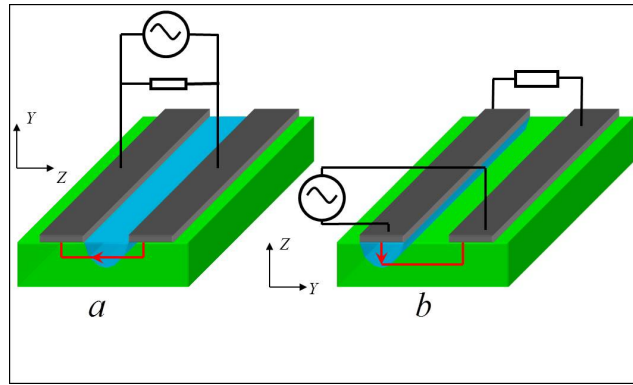
Fig. 5.1. Waveguide modulator structure: *a* – *y*-cut with lumped electrode configuration; *b* – *z*-cut and traveling wave electrode configuration type

For modulator with traveling wave electrode configuration (Fig. 5.1, *b*), the light and the modulating microwave signal propagate in the same direction. If the phase velocities of the modulation field and optical radiation are equal, then the capacitance of the electrodes and the propagation time of light through the modulator do not affect the bandwidth. In this case, the bandwidth is limited by the degree of mismatch of the velocities of the optical and microwave waves. A slight mismatch in speeds does not noticeably degrade the modulator characteristics. At all other parameters being equal, traveling wave modulators are characterized by greater efficiency than modulators with lumped parameters. For modulators on LiNbO$_3$ with coplanar parallel electrodes, the ratio of quality factors

$$\frac{\left(\dfrac{\Delta P}{\Delta f}\right)_{ww}}{\left(\dfrac{\Delta P}{\Delta f}\right)_{lp}} = \frac{1}{3}.$$

The control circuit of the modulator consists of two broadband lines: a strip type formed by a transducer with symmetrical electrodes and a coaxial line that connects the modulator with the generator. The electrical modulating circuit should not impair the properties of the modulator and its quality factor, which requires a high degree of matching of both lines. For a symmetric electrodes

type, it is not possible, as a rule, to achieve good match between coaxial lines and strip lines. Therefore, in high-quality integrated optical modulators, the electrodes are of the asymmetric type: one of them, most often grounded, is made much wider. With asymmetric electrodes, good line matching is achieved, which determines the high quality factor of the control circuit of the modulator.

5.3. Directional-coupler modulator

Two closely spaced channel waveguides could function as a directional coupler, in which optical, energy was synchronously transferred from one guide to the other (section 4). Such a coupler can be made into an electro-optic modulator by merely adding two electrodes, as shown in Fig. 5.2.
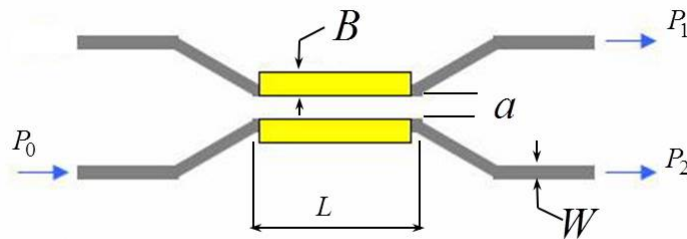

Fig. 5.2. Basic dual-channel modulator structure

When one of the waveguide is excited, energy is redistributed between the waveguides, which depends on the voltage applied to the guide electrodes. The redistribution of energy is due to the overlap of the field profiles of the waveguide modes, which thus provides a distributed connection between the waveguides. In coupled waveguides, the complete transfer of energy from one waveguide to another in the absence of a control field occurs at the coupling length l_c, which corresponds to the crossover state of the waveguides. If the control circuit mode is such that at a distance l_c energy transfer between the waveguides does not occur, then this state is called parallel or straight-through

state. A change in the state of coupled control waveguides underlies the modulation of the radiation propagating through them.

A necessary condition for 100% energy switching is the equality of mode propagation constants.

Therefore, the refractive indices and the thickness of the waveguide layers should be carefully monitored. To minimize l_c, it is necessary to minimize the distance between the waveguides (coupling depth) α and achieve maximum radiation confinement by increasing the jump in the refractive index at the boundary of the waveguide with the substrate.

The use of the TM mode in the modulator leads to a decrease in the value of the control voltage V_0 due to the larger value of the electro-optical coefficient r_{33} (for lithium niobate $r_{33} - 3.5r_{13}$). In this case, to implement the single-mode regime, it is necessary to reduce the width of the waveguides: the increment of the extraordinary refractive index (Δn_e) is greater than the ordinary (Δn_0), as a result the coupling coefficient decreases. To increase the coupling coefficient, it is necessary to reduce the distance between the waveguides.

The main disadvantage of the modulator-switches on coupled waveguides is that if during the manufacturing process it is not possible to get an ideally resonant waveguide with a length exactly corresponding to the coupling length, the possibility of achieving 100% switch will automatically disappear. And since there is almost always a certain mismatch between the waveguide and communication lengths, the switching-modulation efficiency does not reach a maximum.

To eliminate this drawback, it is proposed to use electrodes divided into two sections, to which voltages of opposite polarity are applied. In accordance with it, a change in the refractive indices in the sections will have the opposite sign. This configuration is called a modulator with sectional Δ-electrodes or Δβ-modulator.

The operation mode of the modulator with Δβ electrodes is determined by the state diagram shown in Fig. 5.3. In this diagram, the points lying on curves

1, 2, 3 correspond to the crossover state, and on curves 4–6 to the parallel state. For comparison, the same figure shows a crossover-state diagram for a modulator with non-sectional electrodes, marked with circles with crosses. From the above diagrams it follows that in the case of a modulator with $\Delta\beta$ electrodes with a length, for example, from one to three bond lengths, it is always possible, by applying different magnitudes and inverses in sign voltage, to obtain both crossed and parallel states. Thus, the electrical adjustment of the modulator is implemented in a relatively wide range of lengths.
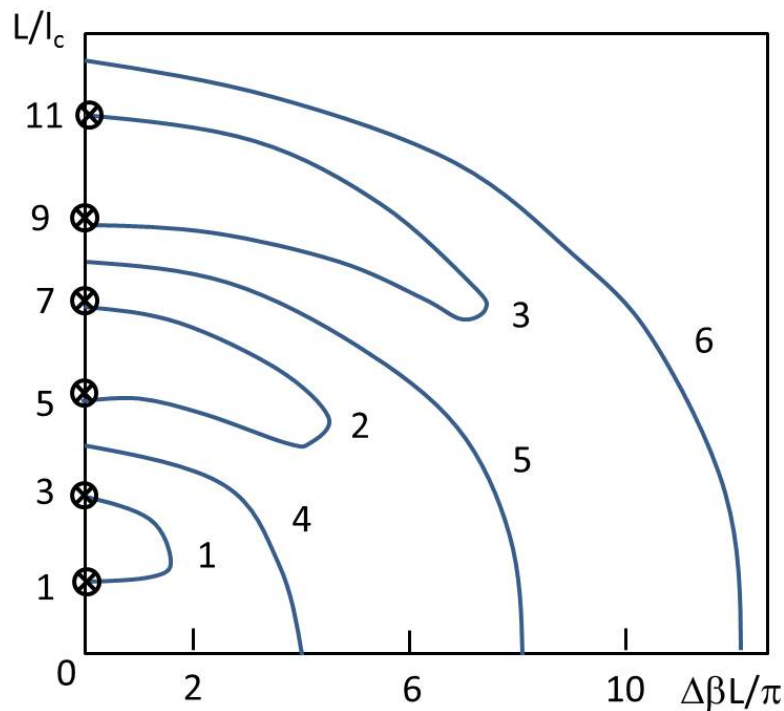

Fig. 5.3. Crossover-state diagram for a modulator with non-sectional electrodes:
1–3 – crossover state; *4–6* – parallel state

Figure 5.4 shows the structure and typical dimensions of a traveling wave modulator-switch with an asymmetric geometry of electrodes made on a *z*-cut of a lithium niobate crystal, whose characteristics are $\Delta P/\Delta f = 7.6$ mW / GHz, $\Delta f = 7.2$ GHz and control voltage $V_0 = 4.5$ V correspond to the best results achieved today.
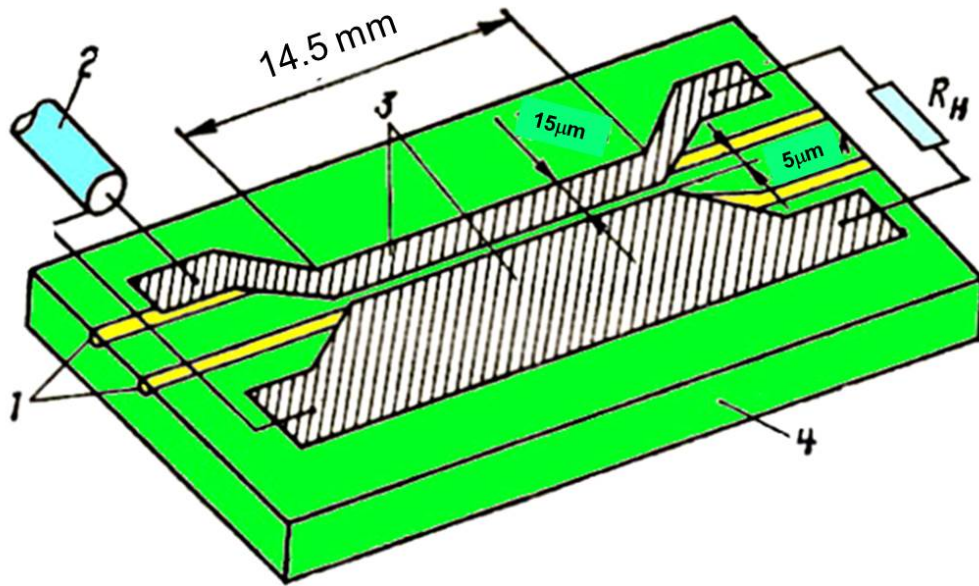
Fig. 5.4. Typical dimensions of a traveling wave modulator-switch with an asymmetric geometry of electrodes made on a *z*-cut of a lithium niobate crystal. *1* – Ti diffused waveguide, *2* – input signal, *3* – electrodes, *4* – substrate

Despite the fact that the majority of electro-optical active elements today are fabricated performed on crystals of active dielectrics, from the integration point of view the latter have a significant drawback. These materials do not allow the creation of radiation sources and receivers – integral components of any integrated optical circuit. The consequence of this is the degree of integration of the circuit is low, and therefore, its speed, reliability do not meet the highest requirements. From this point of view, the possibilities of semiconductor materials are much greater. That is why in recent years there has been a sharp increase in interest in structures based on semiconductor materials, in particular, compounds of the $A^{III}B^V$ type. The use of modern methods of epitaxial growth it is possible to obtain almost all the components of the integrated optical circuit. This opens up a real possibility of creating multifunctional integrated-optical circuits made on a common substrate and in a single technological cycle.

At the same time, diffusion waveguides are typical for active dielectric modulators, semiconductor modulators use comb-type and burred type waveguides (this is due to epitaxial manufacturing technology).

Depending on the circuit specification, waveguides can be fabricated both in homo- and heterostructures. Heteroepitaxial structures provide a more efficient guiding, and hence the best modulators parameters, all other things being equal. This determined heteroepitaxial technology in the development and fabrication of semiconductor waveguide modulators.

The first waveguide modulators, as well as injection emitters, were based on gallium arsenide. Depending on the working wavelength, the modulators are made on the basis of GaAs, GaAlAs, GaAsP with various types of waveguide structures: with metal gaps, comb type, MOS (metal – oxide – semiconductor) - structures, structures with a Schottky barrier.

The transfer of high-speed information systems to the wavelength range of 1.3–1.6 μm led to the development of waveguide modulators based on indium phosphide and InGaAsP/InP solid solutions.

As well as for gallium arsenide, in structures based on indium phosphide, only n-type conductivity materials are used as waveguide layers: the electro-optical change in the refractive index in p-type material is much smaller than that of n-type. Due to the small size of the barrier arising at the Schottky contact, including MIS (metal – dielectric – semiconductor) - contact, with indium phosphide, such structures were not widely used in waveguide modulators on this material.

Since all highly efficient semiconductor radiation sources used in fiber-optic communication and integrated optics are built on double heterostructures, the double heterostructure becomes basic for the development of waveguide modulators integrated with emitters. The use of double heterostructures for waveguide modulators, which are largely similar to structures of injection lasers, makes it possible to realize a high degree of controllability of refractive index

along the waveguide and in its cross section and effective overlap of electric and optical fields in the interaction region.

Despite the fact that the mismatch of the speeds of the optical and microwave waves in gallium arsenide and indium phosphide is very small, there are still no reports of the fabrication of a modulator on these materials with the traveling wave configuration. This is impeded by significant capacitances of the *p-n* junction, depending on the magnitude of the bias voltage, which complicates the calculation of the propagation of the modulating wave in the strip line; noticeable microwave losses of semiconductor waveguides, the nature of which is not fully understood and which limit the working length of the modulator.

However, these limitations are unprincipled, which does not preclude the development and creation in the near future of semiconductor modulator-switches according to the traveling wave scheme.

Semiconductor integrated optical modulators in terms of basic parameters are somewhat inferior to similar elements on active dielectrics. However, their limiting characteristics obtained today – low losses (~1 dB/cm), a wide frequency band exceeding gigahertz, and a relatively low level of control voltages (~10 V) – and in addition to other advantages of semiconductor switch modulators make them quite competitive active elements on crystals of active dielectrics.

5.4. Mach-Zehnder interferometric modulator

The basis of the integrated optical waveguide modulators of the interference type is the Mach-Zehnder interferometer. It is formed by two parallel waveguides connected at the output and input by *Y*-couplers (Fig. 5.5). At the input splitter, the input signal is divided in half. Both parts of the signal, having passed the different shoulders of the interferometer, are again summed at its output, taking into account their phases. Control electrodes are applied to the waveguides in the arms of the interferometer. An electric field through the

control circuit induces a phase shift in the arms of the interferometer of such magnitude that the waves arrive at the common point of the output splitter either in phase or in ant phase. In the first case, they excite the main mode at the junction, which propagates along the common output waveguide, in the second, the second-order odd mode, which is emitted in the transition region of the *Y* splitter and there is no signal at the output.
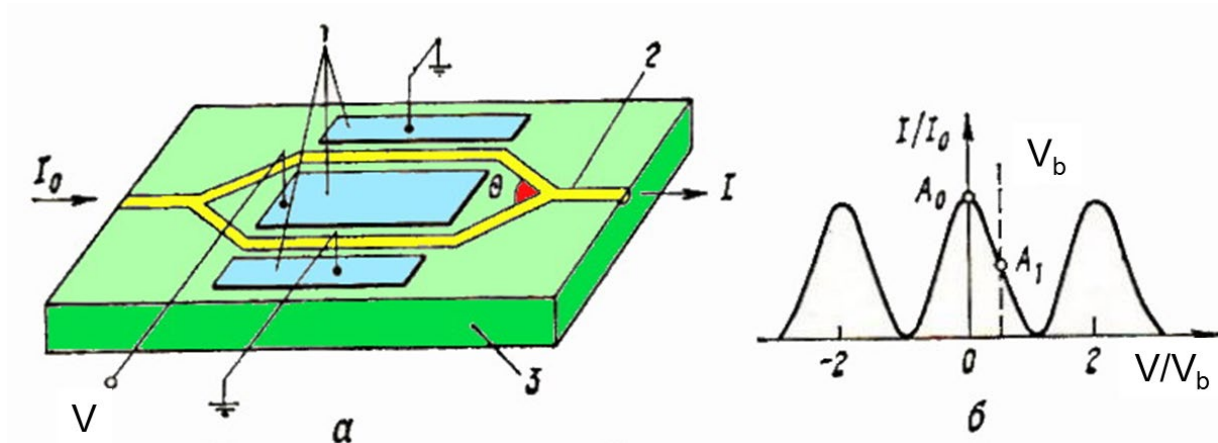

Fig. 5.5. Arrangement of *X*-cut Mach-Zehnder interferometric modulator:

a – schematic representation; *b* – the modulation characteristic;

1 – driving electrods, *2* – waveguide, *3* – substrate

The maximum modulation depth at the device output is determined by the ratio of the amplitudes of the waves in different arms of the interferometer. With equal amplitudes, the modulation coefficient is 100%. To obtain 100% modulation, an accurate signal separation on the *Y* splitter is required. There are various ways to actively control radiation power in *Y*-splitters. But they all require an additional power source and a relatively high control voltage (30 V).

It is possible to achieve almost 100% modulation depth without further complicating the circuit, however, this requires the manufacture of splitters with a high degree of symmetry and accuracy.

The radiation intensity at the output of the interferometer is determined by the following relation,

$$I = I_0 \cos^2\left(\frac{\Delta\varphi_0 + \Delta\varphi}{2}\right) \qquad (5.11)$$

where I_0 is the maximum value of the input intensity, $\Delta\varphi_0$, $\Delta\varphi$ is the phase difference in the arms of the interferometer in the absence and in the presence of a control voltage at the electrodes respectively.

The modulation characteristic of an interferometer with a symmetric structure, when its shoulders are completely identical, is determined by equ. (5.11), whence $\dfrac{I}{I_0} \cong \cos^2\left(\dfrac{\pi V}{2V_0}\right)$, where V_0 is the voltage corresponding to the $\Delta\varphi = \pi/2$ phase change of the signal. This situation is illustrated in Fig. 5.5, b The operating point on the transfer characteristic corresponds to the point A_0 located on the relatively shallow apex of the sine wave. Therefore, the sensitivity of the modulator in this mode is minimal, since small deviations of the control voltage do not give a noticeable change in intensity at the output of the interferometer. To obtain maximum sensitivity, bias voltage must be as

$$V_b = \frac{V_0}{2}.$$

In this case, the operating point on the transfer characteristic of the modulator is shifted from position A_0 to position A_1. This part of the characteristic is much steeper – small changes in the control voltage will give a significant change in the output intensity.

However, the electric displacement for lithium niobate-based waveguide modulators is undesirable due to the effect of a drift in the output signal intensity under the influence of a constant voltage at the electrodes. Therefore, in integrated optical modulators of this type, shift of the operating point to a steeper section of the transfer characteristic is carried out, as a rule due to its structural asymmetry. The necessary symmetry is achieved by shortening by Δ

one of the arms of the *Y*-splitter (Fig. 5.6, *a*). For small angles $\theta \leq 2^\circ$ the value of Δ, providing the required shift of the working point (Fig. 5.6, *b*), is in the range of 500 to 600 microns, which does not introduce additional difficulties – neither technological nor constructive.
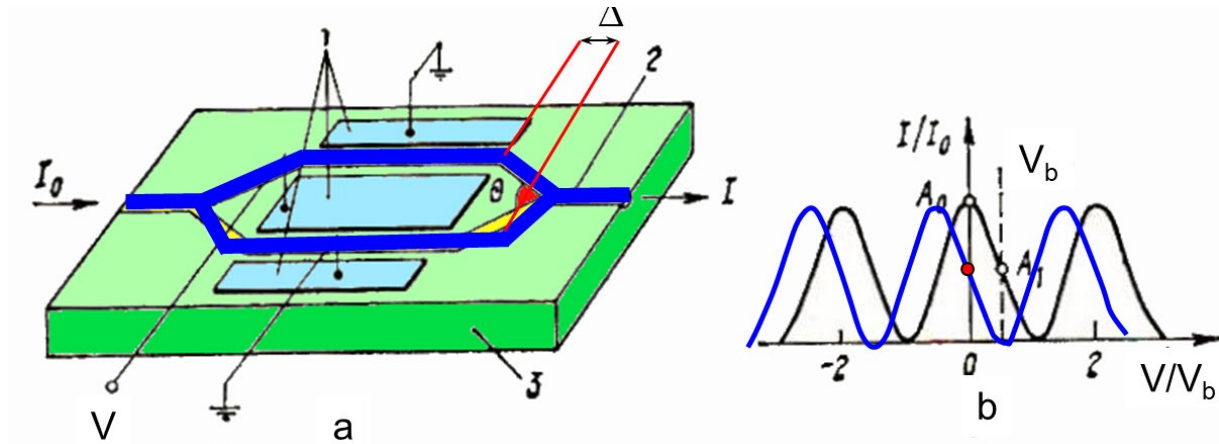

Fig. 5.6. An integrated Mach-Zehnder optical waveguide modulator with asymmetric interferometer arms: *a* – schematic representation; *b* – the modulation characteristic; *1* – driving electrods, *2* – waveguide, *3* – substrate

A constant phase shift can be obtained without changing the length of the arms of the interferometer. Phase shift is a function of the waveguide parameters, therefore, leaving the length unchanged, you can change any other waveguide indication (physical or geometric). In practice, the easiest way is to create a thickened section of the waveguide in one of the arms of the interferometer.

The most wideband interference modulators according to the traveling wave scheme were obtained on the *z*-cut of lithium niobate crystal, for which, at a control voltage of only 5 V, the bandwidth of the modulating frequencies reaches 50 GHz (Fig. 5.7). The values given are close to theoretical limits.

Fig. 5.7. Lithium Niobate Ultra High Bandwidth Electro-Optic Modulators, Fiber-Coupled

The technological difficulties of creating precision *Y*-couplers have hindered until recently the appearance of an interference type modulator based on semiconductor materials. Nevertheless, a number of structures of such modulators with sufficiently high characteristics have already been obtained. For example, in a modulator based on gallium arsenide with comb-type waveguides, optimization of the parameters of the control electric circuit and precise control of the geometry and composition of the waveguide layer made it possible to obtain a wide band of modulating frequencies (60 GHz) with a compact folded-optics configuration, which segregates RF and optical interfacing to opposing ends and helps preserve integrity of the RF feed (Fig. 5.8).
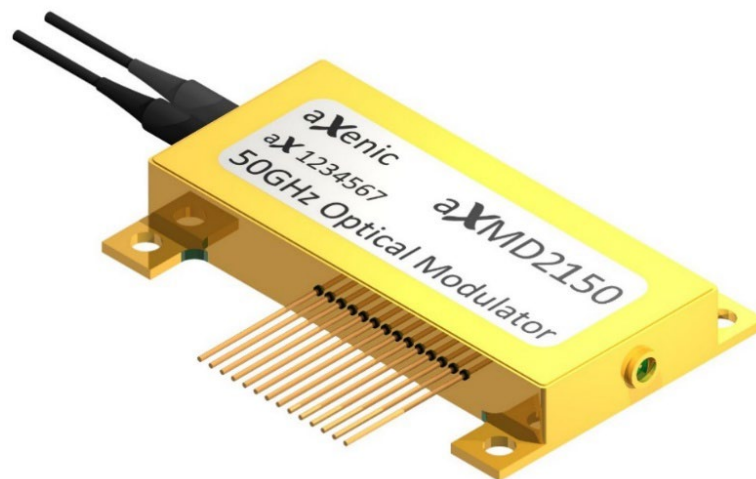

Fig. 5.8. GaAs Mach-Zehnder Broadband Optical Modulator

6. INTEGRATED OPTICAL DETECTORS

Detectors for integrated-optic applications must have high sensitivity, short response time, large quantum efficiency and low power consumption. The most common type of semiconductor optical detector, used in both integrated optic and discrete device applications, is the depletion-layer photo-diode. The depletion-layer photodiode is essentially a reverse-biased semiconductor diode in which reverse current is modulated by the electron-hole pairs produced in or near the depletion layer by the absorption of photons of light. The diode is generally operated in the photodiode mode, with relatively large bias voltage, rather than in photovoltaic mode, in which the diode itself is the electrical generator and no bias voltage is applied.

Aside from the reduction of quantum efficiency that results from poor matching of α (is the optical interband absorption coefficient), W (is the width of the depletion layer) and L_p (is the diffusion length for holes), there are some other limitations to depletion layer photodiode performance that are also important. Since W is usually relatively small (in the range from 0.1 to 1.0 µm), junction capacitance can limit high-frequency response through the familiar R-C time constant. Also, the time required for carriers to diffuse from depths between W and $(W + L_p)$ can limit the high frequency response of a conventional photodiode. The waveguide depletion layer photodiode, significantly mitigates many of these problems of the conventional photodiode.

6.1. Waveguide photodiodes

If the basic depletion layer photodiode is incorporated into a waveguide structure, as shown in Fig. 6.1, a number of improvements in performance are realized. In this case, the light is incident transversely on the active volume of the detector, rather than being normal to the junction plane. The diode photocurrent density is then given by

$$J = q\varphi_0\left(1 - e^{-\alpha L}\right), \qquad\qquad (6.1)$$

where L is the length of the detector in the direction of light propagation. Since W and L are two independent parameters, the carrier concentration within the detector volume and the bias voltage V_a can be chosen so that the depletion layer thickness W is equal to the thickness of the waveguide, while L can be made as long as necessary to make $\alpha L \gg 1$. Thus 100% quantum efficiency can be obtained for any value of α, by merely adjusting the length L. For example, for a material with the relatively small value of $\alpha = 30$ cm^{-1}, a length of $L = 3$ mm would give the quantum efficiency of the detector $\eta_q = 0.99988$. (Again, it has been tacitly assumed in (6.1) that scattering loss and free-carrier absorption are negligible.)
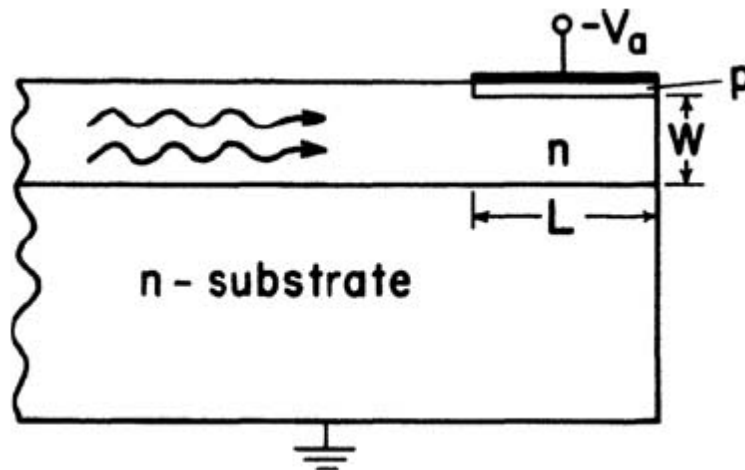

Fig. 6.1. Diagram of a waveguide detector

Because a waveguide detector can be formed in a narrow channel waveguide, the capacitance can be very small. Even if L is relatively large. For example, for a material with a relative dielectric constant $\varepsilon = 12$, such as GaAs, a 3 mm long detector formed in a 3 μm wide channel waveguide has a capacitance of only 0.32 pF. This capacitance is about a factor of ten less than that of a typical conventional mesa photodiode. Hence, the high frequency response can be expected to be correspondingly improved. Experimentally demonstrated bandwidth of 5 GHz and quantum efficiency of 83% have been

obtained with waveguide detectors on GaAs substrate material, and InGaAs waveguide photodetectors on InP substrates also have exhibited a 5 GHz bandwidth for light in the wavelength range of 1.3–1.6μm.

Computer simulation of waveguide photodetectors in AlGaInAs–GaInAs, designed for ultrawide-band operation at 60 and 100 GHz, predict internal quantum efficiencies as high as 94% and 75%, respectively, at 1.55 μm wavelength. Because all of the incident photons are absorbed directly within the depletion layer of a waveguide photodetector, not only is η_q improved, but also the time delay associated with the diffusion of carriers is eliminated. This result is a further improvement in high frequency response.

Due to the many improvements in performance inherent in the transverse structure of the waveguide detector, as compared to the axial geometry of the conventional mesa photodiode, waveguide detectors should be considered for use in discrete-device applications, as well as in optical integrated circuits. At the present time, waveguide detectors are not commercially available as discrete devices. However, they can be fabricated with relative ease in many laboratories. Hence availability should not long be a problem.

6.2. Specialized photodiode structures

There are two very useful photodiode structures that can be fabricated in either a waveguiding or conventional, nonwaveguiding form. These are the *p-i-n* photodiode and the avalanche photodiode.

6.2.1. Avalanche photodiodes

The gain of a depletion layer photodiode (i.e. the quantum efficiency), of either the *p-n* junction or Schottky-barrier type, can be at most equal to unity, under normal conditions of reverse bias. However, if the device is biased precisely at the point of avalanche breakdown, carrier multiplication due to

impact ionization can result in substantial gain in terms of increase in the carrier to photon ratio. In fact, avalanche gains as high as 10^4 are not uncommon. For relatively low reverse bias voltages, the diode exhibits a saturated dark current I_{d0} and a saturated photocurrent I_{ph0}. However, when biased at the point of avalanche breakdown, carrier multiplication results in increased dark current I_d, as well as increased photocurrent I_{ph}. It is possible to define a photomultiplication factor M_{ph}, given by

$$M_{ph} \equiv \frac{I_{ph}}{I_{ph0}} \tag{6.2}$$

and a multiplication factor M, given by

$$M \equiv \frac{I_{ph} + I_d}{I_{ph0} + I_{d0}}. \tag{6.3}$$

For the case of large photocurrent $I_{ph0} \gg I_{d0}$ it was shown that the multiplication factor is given by

$$M = \frac{1}{1 - \left(\dfrac{V_a - IR}{V_b}\right)^n} \tag{6.4}$$

where I is the total current, given by

$$I = I_d + I_{ph} \tag{6.5}$$

R being the series resistance of the diode (including space-charge resistance if significant).

The derivation of (6.4) assumes that $IR << V_b$. For the case of I_{d0} and I_d being negligibly small compared to I_{ph0} and I_{ph}, it can be shown that the maximum attainable multiplication factor is given by

$$M \cong M_{ph} \cong \sqrt{\frac{V_b}{nI_{ph0}R}}\ .\qquad\qquad (6.6)$$

Avalanche photodiodes are very useful detectors, not only because they are capable of high gain, but also because they can be operated at frequencies as high as 60 GHz. However, not every *p-n* junction or Schottky-barrier diode can be operated in the availanche multiplication mode, biased near avalanche breakdown. For example, the field required to produce avalanche breakdown in GaAs is approximately $4 \cdot 10^5$ V/cm. Hence, for a typical deplection width of 5 μm. V_b equals 120 V. Most GaAs diodes will breakdown at much voltages lower due to other mechanisms, such as edge breakdown or microplasma generation at localized defects, thus never reaching the avalanche breakdown condition. In order to fabricate an avalanche photo-diode, extreme care must be taken, beginning with a dislocation free substrate wafer of semiconductor material. Generally a guard ring structure must be employed to prevent edge breakdown.

Avalanche photodiodes are highly-stressed devices. Hence, reliability is a question of prime concern. Increasing leakage current due to poor surface passivation or the generation of internal defects during high current pulse operation can lead to degradation of performance as the devices age. Nevertheless, when diodes are carefully fabricated and are hermetically sealed into adequate packages, mean time to failure as high as 10^5 h at 170 °C has been observed, which projects to about 10^9 h at room temperature.

6.2.2. P-i-n photodiodes

Conventional photodiodes must be designed so as to have a large αW product in order to maximize η_q; but one doesn't have complete control over either the depletion width W, which depends on dopant concentrations, or the absorption coefficient, which depends mostly on the bandgap. In the *p-i-n* photodiode, a very lightly doped "intrinsic" layer is formed between the *p* and *n* sides of the diode. This layer generally has a carrier concentration of less than $10^{14}/cm^3$, but it is compensated by a balance of *p*- and *n*-type dopants rather than being truly intrinsic. Because of the low carrier concentration, the depletion layer in a *p-i-n* diode extends completely through the *i* layer so that the total thickness of the active layer is the sum of the *i*-layer thickness W_i and the deplection width on the lightly doped (*n*) side of the junction. Thus the device designer can adjust the total depletion width to produce a large αW product by varying the thickness of the *i*-layer. The presence of the relatively thick *i*-layer also reduces the junction capacitance and increases the R-C cutoff frequency of the diode, *p-i-n* photo-diodes are widely used as detectors in optical systems because of their high quantum efficiency (responsivity) and wide bandwidth. For example, a waveguide *p-i-n* photodiode operating at 1.55 µm wavelength with a quantum efficiency of 50% and a 3 dB bandwidth of 100 GHZ are commercially available (Fig. 6.2).

Fig. 6.2. Ultra-Fast 100 GHz Photodetector XPDV412xR

The fundamental problem of wavelength incompatibility is very significant with respect to waveguide detectors.An ideal waveguide should have minimal absorption at the wavelength being used.

However a detector depends on interband absorption for carrier generation. Hence, if a detector is monolithically coupled to a waveguide, some means must be provided for increasing the absorption of the photons transmitted by the waveguide within the detector volume. A number of different techniques have proven effective in this regard.

6.3. Hybrid structures

One of the most direct approaches to obtaining wavelength compatibility is to use a hybrid structure, in which a detector diode, formed in a relatively narrow bandgap material, is coupled to a waveguide fabricated in wider bandgap material. The two materials are chosen so that photons of the desired wavelength are transmitted freely by the waveguide, but are strongly absorbed within the detector material. An example of this type of hybrid waveguide/detector is the glass on silicon as shown in Fig. 6.3. The diode was formed by boron diffusion to a depth of about 1 μm into an n-type, 5 Ω·cm silicon substrate. A 1 μm thick layer of thermally grown SiO_2 was used as a diffusion mask.
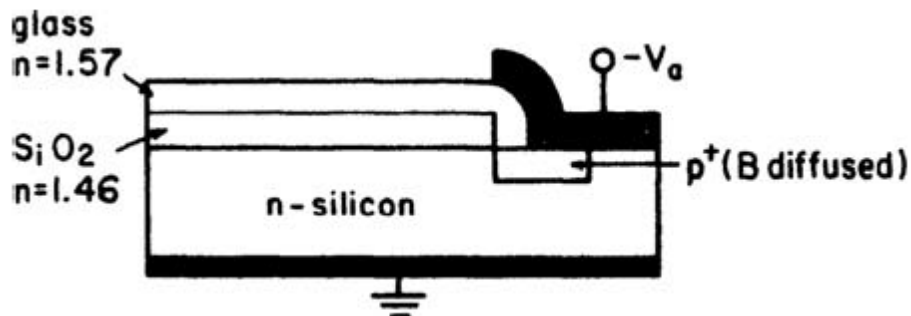

Fig. 6.3. Hybrid waveguide detector featuring a glass waveguide coupled to a silicon photodiode

The glass waveguide was then sputter-deposited and silver paint electrodes were added as shown. Total guide loss was measured to be 0.8 dB/cm ± 10% for light of 6328 Å wavelength. The efficiency of coupling between the waveguide and the detector was 80%. However, because the light enters the diode in the direction normal to the junction plane rather than parallel to it, this particular waveguide detector geometry does not have many of the advantages. Nevertheless, good high frequency response can be expected. These diffused diodes had a capacitance of only $3 \cdot 10^{-9}$ F/cm^2 when reverse biased with V_a equal to 10 V.

Thus a detector diode of approximately 10 μm radius, used in conjunction with a 50 Ω load resistance, would have an RC time constant of about 15 ps, implying that modulation of frequencies in excess of 10 GHz could be detected.

While hybrid detectors offer the possibility of choosing the waveguide and detector materials for optimum absorption characteristics, better coupling efficiency can be obtained with monolithic fabrication techniques. Monolithically fabricated waveguide detectors also have the advantage that light enters the device in the plane of the junction rather than normal to it.

The most popular method of monolithically integrating a waveguide and detector is to use heteroepitaxial growth of a relatively narrow bandgap semiconductor at the location where a detector is desired. An example of this approach is shown in Fig. 6.4 the InGaAs detector that has been integrated with a GaAs waveguide. In In$_x$Ga$_{(1-x)}$As, the bandgap can be adjusted to produce strong absorption of light at wavelengths in the range from 0.9 to 3.5 μm by changing the atomic fraction x of indium. The monolithic waveguide detector structure shown in Fig. 6.4 combines an epitaxially grown carrier concentration-reduction type waveguide with a platinum Schottky barrier detector.

A 6000 Å thick layer of pyrolytially deposited SiO$_2$ was used as a mask to etch a 125 μm diameter well into the 5–20 μm thick waveguide, and then grow the In$_x$Ga$_{(1-x)}$As detector material. A quantum efficiency of 60% was measured for this detector at a wavelength of 1.06 μm, for low bias voltages. The loss in

the waveguide was less than 1 dB/cm. At bias voltage greater than about 40 V, avalanche multiplication was observed, with multiplication factors as high as 50. Optimum performance was obtained at a wavelength of 1.06 μm for an In concentration of $x = 0.2$. The fact that quantum efficiency in this device did not approach 100% more closely was most likely caused by less than optimum depletion width in the Schottky-barrier diode. The carrier concentration in the waveguide must be very carefully controlled in order to make W equal to the waveguide thickness. Defect centers associated with stress at the GaAs-GaInAs interface may have also played a role in reducing η_q.
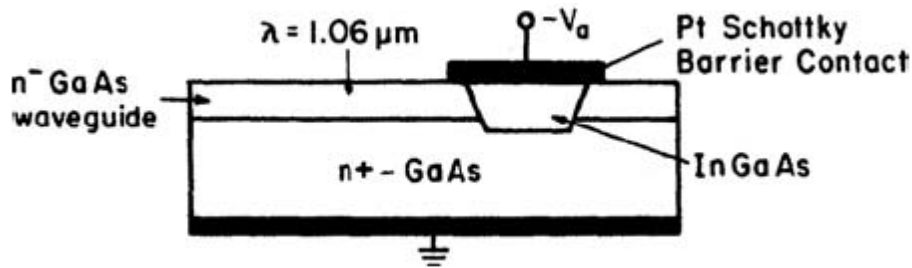

Fig. 6.4. Monolithically integrated InGaAs detector in a GaAs waveguide

In general, the III–V compound semiconductors and their associated ternary (and quaternary) alloys offer the device designer a wide range of bandgaps, and corresponding absorption edge wavelengths. Direct bandgap materials generally have interband absorption coefficients greater than 10^4 cm^{-1} for wavelengths shorter than the absorption edge, while α may be several orders of magnitude less in indirect gap materials. Nevertheless effective detectors can be made in indirect bandgap materials, especially when the waveguide detector geometry is used, so that the length L can be adjusted to compensate for small α.

The most popular material for the fabrication of detectors in the 1.0 to 1.6 μm wavelength range is GaInAsP. Figure 6.5 shows a planar embedded GaInAs *p-i-n* photodiode on an InP substrate. The device was fabricated by means of preferential ion beam etching and vapor phase epitaxy. Since the bonding pads lie mostly over semi-insulating substrate material, stray

capacitances have been reduced to the point at which total device capacitance is only 0.08 pF for a diameter of 20 μm of photosensitive area.
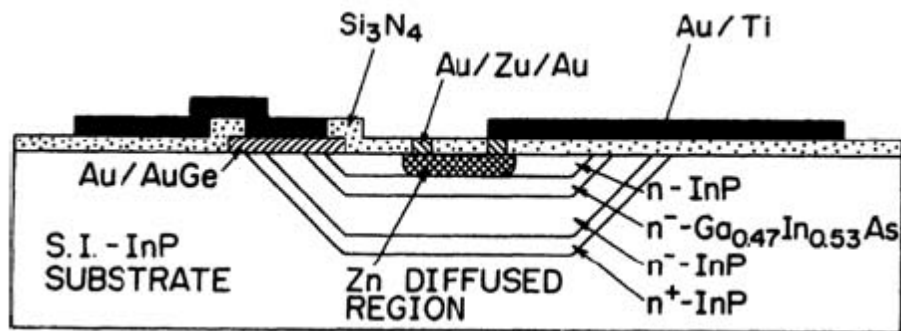

Fig. 6.5. Cross-section of the planar embedded photodiode

As a result, this device has a cutoff frequency of 14 GHz, limited by the carrier transit time. An integrated waveguide *p-i-n* photodetector coupled to a ridge waveguide on InP is shown in Fig. 6.6. The device was fabricated by means of a metal organic vapor phase epitaxial (MOVPE) regrowth technique.
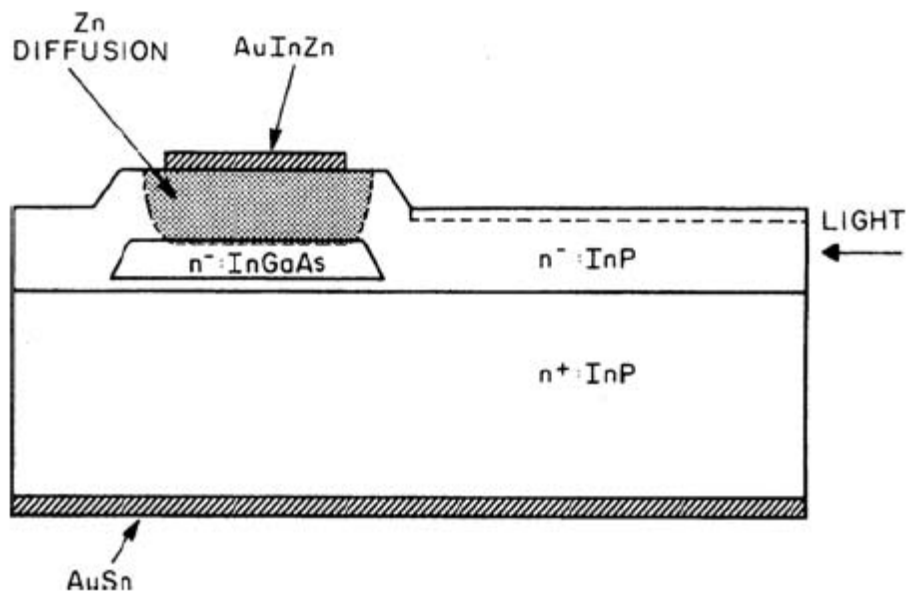

Fig. 6.6. A p-i-n photodiode integrated with a ridge waveguide

The detector can be used in the 1.0–1.6 μm wavelength range; and has a 3 dB bandwidth of 1.5 GHz with a pulse response (full-width-at-half-maximum)

of 80 ps at 1.3 μm wavelength. The waveguides had an average propagation loss of 3 dB/cm and 95% of the guided light was coupled into the photodetector.

Another approach to integrating a GaInAs detector and an InP waveguide is shown in Fig. 6.7. In this case an InGaAs detector was grown on top of the waveguide. Since InGaAs has a larger index of refraction than InP, light is coupled out of the waveguide up into the detector.
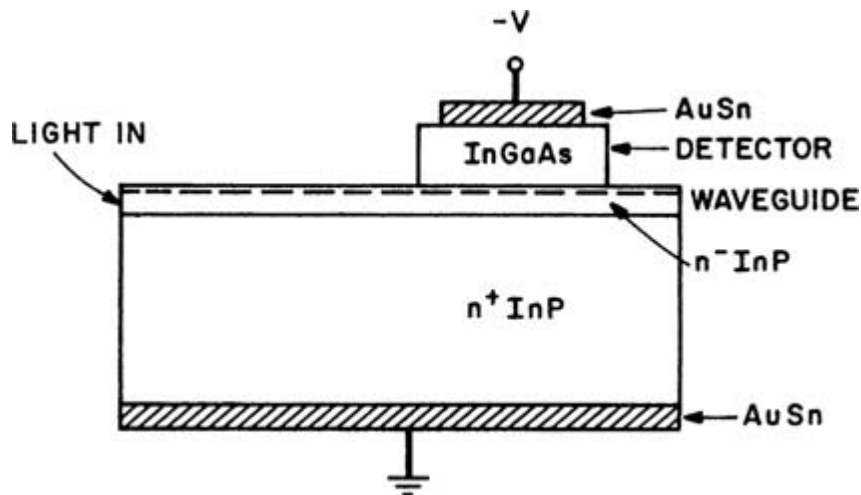

Fig. 6.7. An InGaAs detector grown on top of a ridge waveguide

6.4. Factors limiting performance of integrated detectors

In the design of an integrated optical detector, there are a number of mechanisms that can limit performance in various ways. Not all of these are important in every application. However, the designer (or user) should be aware of the limitations associated with different device types and geometries.

6.4.1. High Frequency Cutoff

The speed of detectors is essentially controlled by the RC time constant and the transit time for the carriers. The device speed design issues are thus similar to those for electronic devices. The devices are made as small as possible

and for high speed materials such as InGaAs, it is possible to have bandwidths approaching 150 GHz with current technology.

A number of the factors that can limit high frequency are summarized in Table 6.1, along with some additional frequency-limiting effects. Because of the small area of waveguide photodetectors, the RC time constant, which most often limits the response of conventional diodes, can be made small enough to allow frequencies of operation well above 10 GHz. In this case, other potentially limiting effects must be considered.

Table 6.1

Factors limiting high frequency response of a depletion layer photodiode

-	RC time constant due to bulk series resistance and junction capacitance
-	Carrier diffusion time from regions outside of the depletion layer
-	Carrier lifetime and diffusion length
-	Capacitance and inductance of the package
-	Carrier drift time across the depletion layer
-	Carrier trapping in deep levels

The drift time of carriers across the depletion layer can be minimized by designing the device so that the field in the depletion layer is high enough that carriers travel at the scattering limited velocity. For example, in GaAs the scattering limited velocity of $1 \cdot 10^7$ cm/s is reached in electric field strengths greater than about $2 \cdot 10^4$ V/cm. Hence, the transit time across a typical 3 μm wide depletion layer can be made as small as $3 \cdot 10^{-11}$ s. However, it is important that the detector be designed so that the depletion layer extends entirely through the waveguide to the substrate, so that all carriers are generated within the depletion layer itself. If depletion is incomplete, carriers generated in undepleted material must diffuse relatively slowly into the depletion region before being collected, a process that occurs over roughly a minority carrier lifetime (about 10^{-8} s in lightly doped GaAs). A short optical pulse will therefore appear to have

a long diffusion tail when detected. Carrier trapping in deep levels can also cause a tail on the detected waveforms of a short pulse, since detrapping times can be relatively long. Deep level traps are usually associated with defects present in the semiconductor crystal lattice. Hence, special care should be taken in materials selection and device fabrication to minimize the number of defects.

When waveguide detectors are properly designed and are appropriately fed with a microwave transmission line, bandwidths are greater than 80 GHz.

An important class of detectors for high speed and integration is the Schottky metal-semiconductor detector (see Tabl. 6.2). This detector is easy to fabricate and has an extremely high speed. Schottky detectors have been shown to have 3 dB bandwidths approaching 160 GHz.

6.4.2. Linearity

A depletion layer photodiode, when reverse biased by more than a couple of volts, operates in the photodiode (or photoconductive) mode. It then functions as a current source, with its current being proportional to the input optical power,up to power levels of typically 1 mW. Hence, in most applications it is a highly linear device. At higher power levels, saturation occurs when the concentration of photo-generated carriers is so large that the field in the depletion layer is reduced by space-charge effects. This field reduction is particularly important in high frequency applications, because it may then result in carriers traveling at less than the saturation-limited velocity.

6.4.3. Noise

The effects of noise in waveguide detectors are essentially the same as they are in conventional photodiodes. The major noise components are thermal noise, arising in bulk resistances of the device, shot noise, associated with nonuniformities of current flow, such as carrier generation and recombination,

and background noise, due to photons that are not part of the optical signal entering the detector.

From a relatively simple model, it can be shown that the signal-to-noise ratio in a depletion layer photodiode due to the effects of thermal and shot noise is given by

$$\left(\frac{S}{N}\right)_{\text{power}} = \frac{\eta_q}{4B} M^2 \varphi_0 A \left(1 + \frac{2kT}{q} \frac{(\omega RC)^2}{RI_s}\right)^{-1} \tag{6.7}$$

where η_q is the quantum efficiency, B is the bandwidth, M is the modulation index (in the case of an intensity modulated light beam), φ_0 is the incident photon flux density, A is the area of the input face, R is the diode bulk resistance, ω is the modulation frequency of the optical signal, I_s is the reverse saturation (dark) current, k is Boltzmann's constant and C is the capacitance. In the case of an avalanche photodiode, there are additional noise sources associated with the statistical nature of the avalanche.
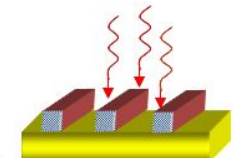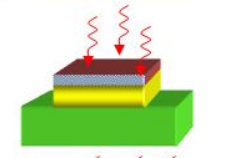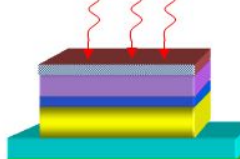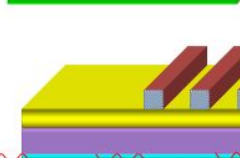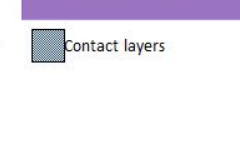
The waveguide detector has an inherent advantage over conventional detectors in regard to background noise, because the waveguide acts as a filter to eliminate much of the background light. Close matching to the signal wavelength also reduces background noise. For example, if light of 850 nm wavelength from a GaAlAs emitter is detected by a Si detector, background photons with wavelengths shorter than about 1.2 μm will also be detected. However if, instead, a GaAs detector is used, those photons with wavelengths in the range from 1.2 to 0.9 μm will not contribute to background noise.

Main types of photodetectors used in IOEC and their features are summarized in Table 6.2.

Table 6.2

Main types of photodetectors used in IOEC

Detector Technologies

	Layer Structure	Features
MSM (Metal Semiconductor Metal)	Semiinsulating GaAs	**Simple, Planar,** *Low Capacitance Low Quantum Efficiency*
PIN	Contact InGaAsP p 5×10^{18} Absorption InGaAs n 5×10^{14} Contact InP n 1×10^{19}	*Trade-off Between Quantum efficiency and Speed*
APD	Contact InP p 1×10^{18} Multiplication InP n 5×10^{16} Transition InGaAsP n 1×10^{16} Absorption InGaAs n 5×10^{14} Contact InP n 1×10^{18} Substrate InP Semi insulating	**Gain-Bandwidth: 420GHz** *Low Noise Difficult to make Complex*
Waveguide	Absorption Layer Guide Layers	**High efficiency High speed** *Difficult to couple into*

Key: Absorption Layer   Contact layers

7. SEMICONDUCTOR LASERS

The prospectivity of certain radiation sources for integrated optics is determined primarily by the possibility of their technological integration with other components, microminiaturization, the implementation of single-mode generation in a wide range of operating capacities, and high speed. Semiconductor injection lasers based on heterostructures, which are the main type of emitters in integrated optics, best satisfy these requirements.

The modern approach to solving the problem of creating laser heterostructures is based on the use of the isoperiodic principle of substitution in multicomponent solid solutions of semiconductor compounds. It is associated with the widespread use of ternary and quaternary systems of complex composition, covering an important spectral range from 0.8 to 1.7 μm. For monolithic integrated optical circuits, lasers with distributed feedback (DFB lasers) and lasers with distributed Bragg reflectors – (DBR lasers) are most suitable. Apparently, such structures will become the basis for the emitters of multifunctional monolithic integrated optical circuits of the future.

The physical structure of a typical DH laser in GaAlAs is shown in Fig. 7.1, along with a diagram of the index of refraction profile in the direction normal to the *p-n* junction plane.
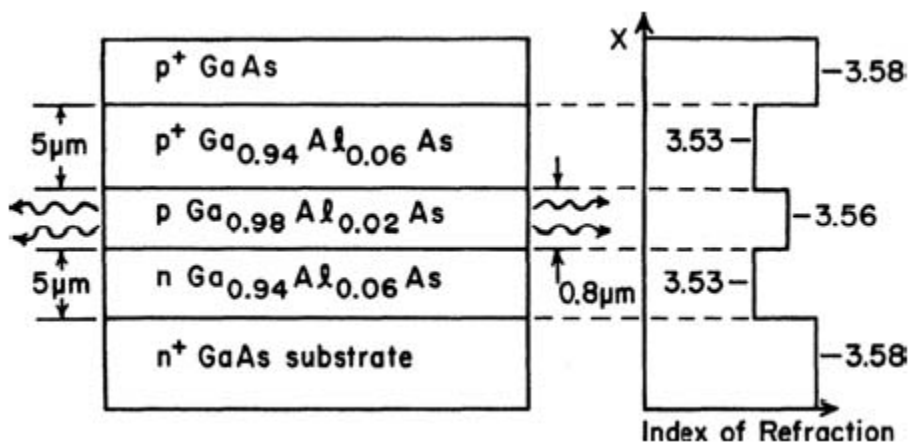

Fig. 7.1. Double heterostructure laser diode

A double heterostructure (DH) injection laser is a complex multilayer structure in which the geometric dimensions and composition of the layers must be maintained with great accuracy. The current level achieved in the field of semiconductor injection lasers is largely determined by the technology of their manufacture. All laser multilayer structures are grown by epitaxial methods.

In injection laser, an active layer of thickness d located between layers with a lower refractive index forms a planar waveguide, the reduced thickness of which is $D = \left(\dfrac{2\pi d}{\lambda_0}\right)\sqrt{n_1^2 - n_2^2}$ (a light emitting layer of thickness D, which is greater than the thickness d), where λ_0 is the radiation wavelength in vacuum; n_1, n_2 are the refractive indices of the active and emitter layers, respectively. The thickness of the active layer is a very important parameter of the structure, since laser efficiency depends on its optimization. The optimal value of D_{opt} is determined by the equation

$$\frac{1}{\Gamma_3}\left(1 - \frac{D_{opt}}{m_i \Gamma_3}\frac{d\Gamma_3}{dD_{opt}}\right) = C_\alpha \qquad (7.1)$$

where Γ_3 is the optical confinement parameter representing the ratio of the optical flux in the waveguide layer to the entire optical flux of a given mode; m_i is the indicator of the optical gain dependence on the pump current density; $C_\alpha = \dfrac{\alpha_{pn} - \alpha_0 - b}{\alpha_{pn} + \alpha_{exp}}$ is the coefficient of uniformity of optical losses across heterostructure layers; α_0 is the coefficient of nonresonant absorption in the active layer, α_{pn} is the average value of the absorption coefficient in passive p and n regions; α_{exp} is the coefficient characterizing the external losses in the cavity; b is a constant that is part of the formula for the volume index of refraction. For GaInAsP system

$$D_{opt} = \left[\frac{2(2 - m_i)}{m_i(1 - C_\alpha)} \right]; \quad D_{opt} = 6.594\, d_{opt} \qquad (6.2)$$

From eq. 6.2 the optimum value of the active layer thickness is 0.21 μm.

For injection lasers the parameters to be optimized are, first of all, the threshold current, the value of which should be as low as possible, the modulation frequency or speed, the mode composition of the radiation, which determines its monochromaticity and coherence. The minimum duration of the generation pulses τ_i is also significant. Since the modulation frequency band is inversely proportional to the pulse duration, the shorter the laser pulse, the more information it can transmit per unit time.

The characteristics of a laser as a transducer are determined by its differential quantum efficiency

$$\eta_d = q\left(\frac{dP}{dI} \right) \qquad (6.3)$$

where P, W is the output power; I, A is the pump current. For modern injection lasers, η_d lies in the range from 40 to 60% ($\eta_{d\max}$ = 70–75%) and is determined by the quality of the material and structure of the laser and its design.

Injection lasers are high-speed devices. The maximum modulation frequency of their radiation can reach 60 GHz The operating frequency of industrial lasers is already close to these limits and reaches 40 GHz.

Almost all elements of integrated optics operate in single mode – on the same wavelength or narrow frequency band of electromagnetic radiation. Single-mode radiation allows one to obtain its narrow radiation pattern. In addition, single-mode lasers are characterized by the highest rates with respect to noise and stability of their characteristics.
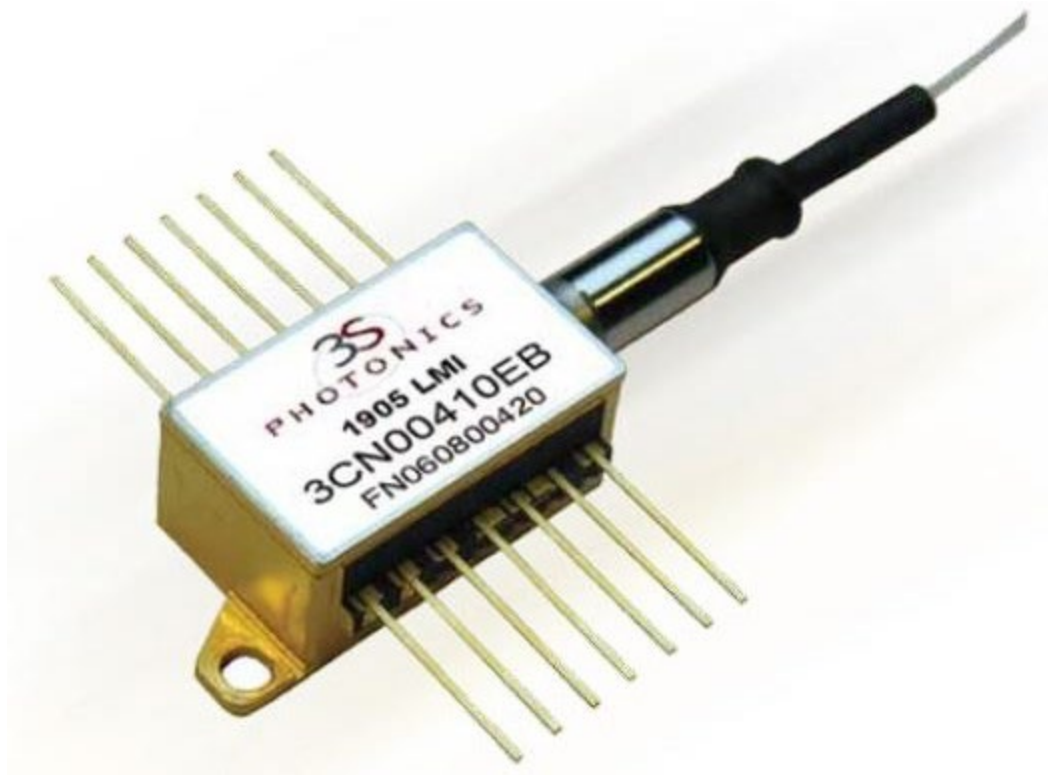
Fig. 7.2. 40Gb/s InGaAsP Distributed FeedBack laser

7.1. Peculiarity of integrated-optical lasers

Integrated optical lasers must be suitable for monolithic integration with optical waveguides and other elements of an optical integrated circuit with regard to their dimensions and material composition. However, four significant problems must be dealt with in the design of any specific monolithically integrated laser/waveguide structure. There must be efficient coupling of the light from the laser to the guide, and there must be some means of providing the optical feedback required by the laser. Also, the absorption loss of the waveguide must be low at the emitted wavelength. Finally, some provision must be made for the electrical contacts.

Integrated-optical lasers are characterized by operation at relatively low threshold current compared to discrete ones. The desire to reduce the latter is due to the need to solve the problem of heat removal from the laser. With small geometric dimensions of injection lasers, they emit heat with a small volume, it is removed directly through the substrate. Since the thermal conductivity of a

semi-insulated substrate is less than that of a conducting one, this shifts the emphasis in choosing a laser structures with minimized threshold current density. In addition, small threshold currents, all other things being equal, determine large laser modulation frequencies. This is very important for broadband integrated optics.

Most of the existing injection lasers use a Fabry-Perot resonator as a feedback element, formed by the cleaved edges of the laser crystal – semitransparent mirrors, from which the radiation generated in the crystal is reflected, and the positive feedback necessary for laser operation is created. This design of the resonator complicates the integration of the radiation source with other elements of the integrated optical circuit on a single substrate: the resonator size imposes strict restrictions on the geometry of the circuit in advance. Structures were developed in which the resonator was created without cleavage by selective (localized in place) epitaxy and etching methods, by micro-cleavage.

Masked chemical etching of a laser facet has also been used to produce a GaInAsP laser automatically integrated with a monitor photodetector as shown in Fig. 7.3. The laser in that case was a buried-heterostructure, stripegeometry device emitting at a wavelength of 1.3 μm. The isolation groove between the laser and photodiode was produced by masking with SiO_2, configured by standard photolithography, and etching with a methyl alcohol and bromine solution. The threshold current and external quantum efficiency of the laser were measured to be in the range $I_{th} = 25-35$ mA and $\eta_q = 0.06-0.09$ mW/mA. Comparable values for similar lasers with cleaved facets were $I_{th} = 20$ mA and $\eta_q = 0.19$ mW/mA.

The decreased performance of the lasers with etched facets was attributed to a slight tilt in the laser facet. Nevertheless, the I_{th} and η_q of the etched-facet lasers were reasonable for use in optical integrated circuits. A photocurrent of 15 nA per mW of laser optical output power was obtained from the photodiode, and

the response was linear. Since light emission was into the air absorption loss was not a problem.
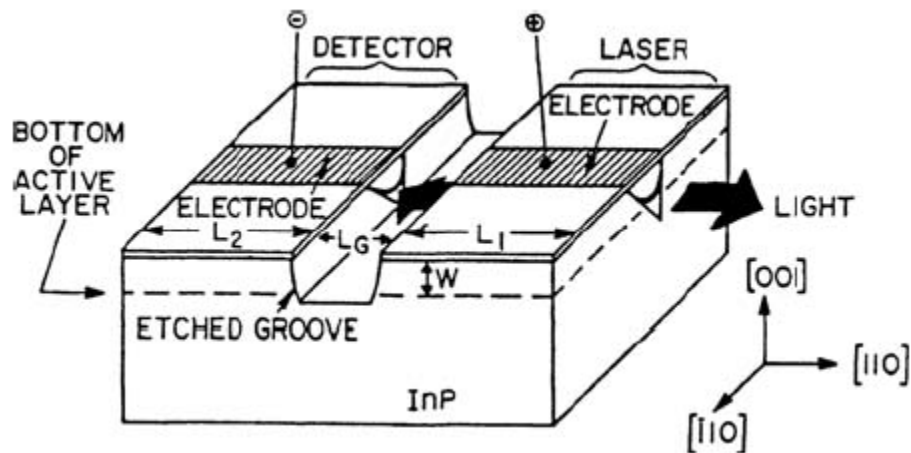

Fig. 7.3. A monolithically integrated GaInAsP laser and monitor photodiode

Another technique for producing monolithically integrated laser diodes with cleaved called "stop cleaving", is illustrated in Fig. 7.4, where the thick dark line marks the location of the laser cavity. Holes etched into the substrate are employed to stop the propagation along a cleavage plane, producing cleaved facets which are of comparable quality to conventionally cleaved laser facets.

To produce stop-cleaved InGaAsP lasers on an InP substrate, etchants of $4H_2SO_4:1H_2O_2:1H_2O$ and concentrated HCl were used to selectivity etch layers of InGaAsP and InP, respectively. Stripe-geometry, buriedheterostructure lasers fabricated by this method to emit at 1.3 μm had threshold currents as low as 20 mA and differential quantum efficiencies as high as 60%.

There is usually no difficulty in making electrical contact to the *p* and *n* sides of the junction of a discrete laser diode, but when the diode is monolithically integrated into an optical integrated circuit it may not be possible to obtain a return path for current through the substrate (for example, when a semi insulating substrate is used).

In such cases a transverse-junction-stripe (TJS) laser is called for. The TJS laser shown in Fig. 7.5, features a transverse junction which has both the *p*

and the *n* electrical contacts on the top surface. A double heterostructure, stripegeometry laser cavity is produced in n-type layers grown on a semi-insulating substrate.

PRIOR TO
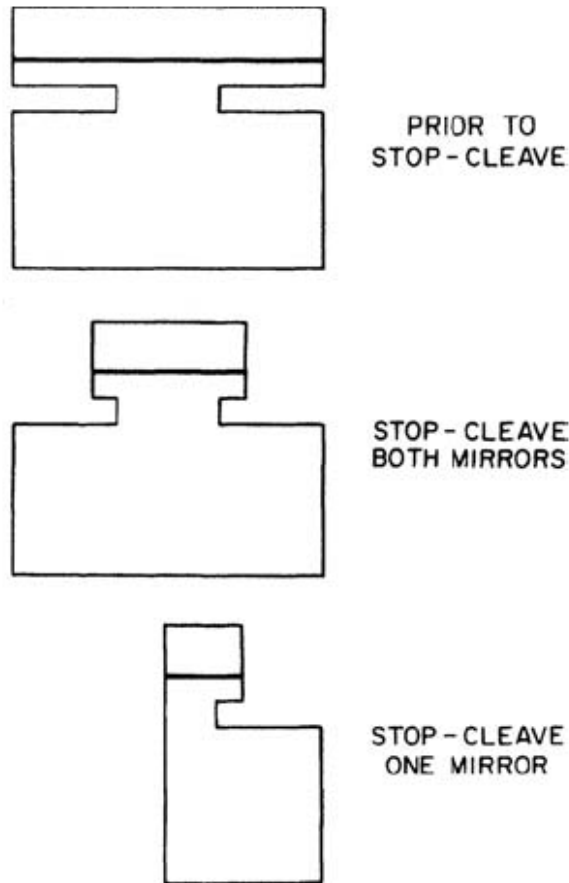STOP-CLEAVE

STOP-CLEAVE
BOTH MIRRORS

STOP-CLEAVE
ONE MIRROR

Fig. 7.4. Laser diodes with stopcleaved facets

METAL

SI-GaAs SUB

ACTIVE
REGION

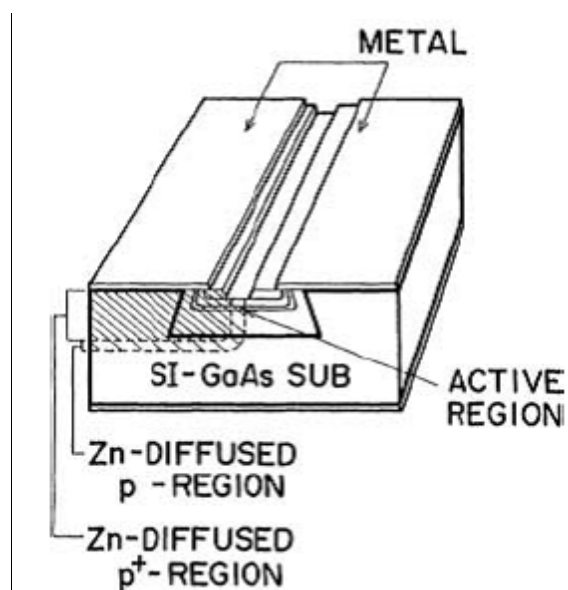Zn-DIFFUSED
p -REGION

Zn-DIFFUSED
p$^+$-REGION

Fig. 7.5. A transverse junction stripe (TJS) laser

Then masked Zn diffusion is used to produce a lateral *p-n* junction.

Unfortunately, the options considered do not completely solve the problem of technological integration of lasers – they relate, in fact, only to particular design and technological improvements of specific injection lasers that are not specifically designed for the tasks of integrated optics. Lasers with periodic modulation of optical characteristics are more suitable for this task.

7.2. Distributed-feedback lasers

Most of the lasers that have been described so are depend on optical feedback from a pair of reflecting surfaces, which form a Fabry-Perot etalon. In an optical integrated circuit, in which the laser diodes are monolithically integrated within the semiconductor wafer, it is usually very difficult to form such reflecting surfaces.

They can be formed by etching or cleaving, as described above. However, the planar surface of the wafer is then disrupted, which leads to difficulties in fabricating electrical connections and heat sinks. An alternative approach, which utilizes distributed-feedback (DFB) from a Bragg-type diffraction grating, provides a number of advantages while still utilizing a planar surface geometry.

To create positive feedback in lasers with periodic modulation of the optical characteristics, it is necessary that the wave, scattered by periodic inhomogeneities, change the direction of propagation in the opposite direction. Two coupled waves of the same frequency arise, propagating in opposite directions. This effect is realized if some optical parameter of the medium is modulated in the direction of wave propagation with a period Λ satisfying the Bragg condition

$$\Lambda = \frac{m\lambda_0}{2n_{eff}}, \qquad (7.4)$$

where *m* is the Bragg reflection order.

Such a Bragg resonator, in contrast to the Fabry-Perot resonator, is characterized by only one resonance and gain band. This determines its high spectral selectivity. In addition to this, in the Bragg resonator it is quite simple to carry out the emission of radiation in a plane perpendicular to its propagation in the waveguide. In some cases, this greatly simplifies the design of the emitter.

Radiators with periodic modulation of optical characteristics include distributed feedback lasers (DFB lasers) and distributed Bragg reflectors (DBR lasers). In DFB lasers, the medium is modulated directly in the active region or in a layer located close to it (Fig. 6, *a*). In DBR lasers, modulated regions are located outside the active region (Fig. 6, *b*). Laser structures with periodic modulation of the optical characteristics differ in diffraction order equal to an integer number of half-waves of laser radiation that fall within the period of inhomogeneity. Distributed reflectors are in the form of dielectric waveguides with a corrugated surface. A periodic inhomogeneity is a change in the thickness of the waveguide. Since the propagation constant depends on the thickness of the waveguide, a change in the latter is equivalent to a change in its refractive index.
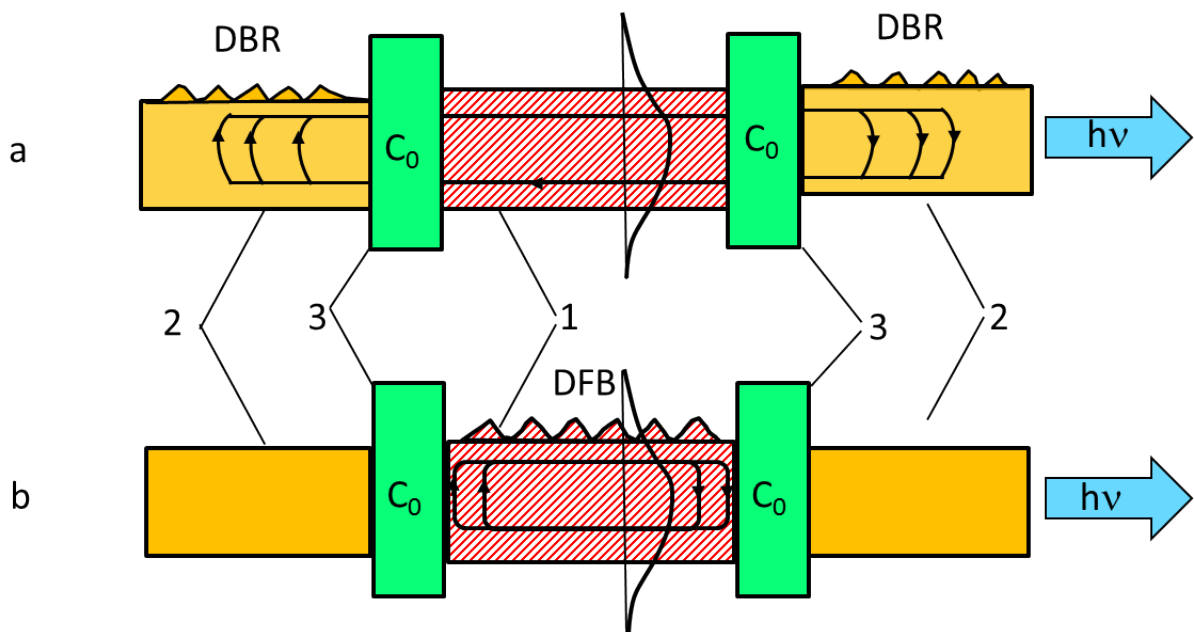

Fig. 7.6. DFB (*a*) and DBR (*b*) lasers constructing principle:
1 – active waveguide, *2* – coupling region, *3* – passive waveguide

The active (generator) and passive (waveguide) components that form the laser cavity are coupled through the transition region of coupling, the parameter of which is the coupling coefficient C_0. The radiation is output through an external waveguide. A necessary condition for high differential quantum efficiency in DBR lasers is low losses in the DBR regions and at the same time a highly efficient coupling between the external and internal waveguides. In DFB lasers, the quality of coupling with an external waveguide is not so significant for the generation process, but it determines the efficiency of radiation input into elements of an integrated optical circuit.

As a rule, five types of coupling are used in integrated lasers: tunneling, through a tapering edge to the joint, due to vanishing fields and the type of built-in end waveguide.

The first is implemented in a structure with a double integrated waveguide, in which the active and passive are fully connected by the principle of a directional coupler. This determines the equality of the phase velocities of the modes of both waveguides. A part of the light power generated in the active region is tunneled from the passive in to active waveguide through the layer with a lower refractive index separating them. The coupling efficiency is limited by a deviation from the conditions of exact phase matching and absorption losses in the waveguide. The requirement of matching the mode propagation constants in the waveguides imposes restrictions on the format and composition of the epitaxial layers of the structure. This complicates the process of its fabrication. The second type of structure is a bond due to the narrowing edge of the waveguide. The effectiveness of this type of coupling while minimizing losses in the narrowing region can approach 100%.

The disadvantage of this coupling is the need to carefully monitor the growth conditions of the structure or to grow it in two stages with intermediate etching to minimize losses in the narrowing area. The butt-coupling (end-to-end, third type), if you do not take special measures, is ineffective, however, due to the ease of implementation it is widely used in integrated optics devices. In

lasers with a large optical resonator, the active region is directly adjacent to the passive waveguide. The waveguides are coupled here due to modes evanescent tails. If the difference in the refractive index between the active region in the waveguide layer is small and the active region itself is quite thin, the field is not held in it and propagates in the waveguide formed by the active and passive regions. Without the use of special measures for the exact fitting of the external waveguide, the effectiveness of this connection is small.

A significant increase in the efficiency of coupling is provided by the build-in external waveguide. This structure is an optimized end-to-end coupling option. Improved docking characteristics are achieved by matching effective refractive indices and field profiles in both waveguides. The effect of the mismatch on coupling efficiency in such a structure is minimized. Maximum efficiency can reach 98%. In this case, when the thickness of the external waveguide changes from 0.41 to 0.61 μm, the coupling efficiency decreases only by 5%.

In single-mode lasers with a Fabry-Perot resonator, when the temperature changes by only a few degrees, the laser wavelength changes stepwise. The value of the jump is equal to the intermode interval (1–1.5 nm). Periodically modulated lasers are characterized by a weaker temperature dependence of the radiation wavelength: laser radiation with a flat resonator "monitors" the temperature dependence of the band gap, and in structures with periodic modulation, a weaker temperature dependence of the refractive index.

As the temperature changes in DBR lasers, the generated mode jumps to another one abruptly, while in the DFB lasers the emitting mode remains the same.

From the solution of the dispersion equation for waves satisfying the diffraction condition, it follows that the laser modes of the DFB lasers appear on the frequency scale symmetrically with respect to the Bragg frequency with the interval $\dfrac{\lambda_B^2}{2\pi L}$, i.e. $\lambda \neq \lambda_B$. In contrast to DFB lasers in DBR structures, the

reflection resonance falls precisely on the Bragg frequency λ_B and is not subject to splitting. In the general case, it is assumed that DBR lasers, in comparison with DFB, allow to obtain a lower lasing threshold and a narrower spectral radiation width for the same active medium properties.

The grating structure of a DFB/DBR laser is usually formed by masking and then etching the waveguide surface, either chemically or by ion-beam sputtering. Sometimes the two methods are combined in a single process called chemically assisted ion beam etching (CAIBE).

Periodic perturbation directly in the active region of the DBR laser is accompanied, as a rule, by structural defects that worsen its luminescent properties. Nonradiative recombination of injected carriers at the corrugated boundary of the active region significantly increases threshold currents and virtually eliminates their minimization.

The approach that is most often used to avoid the effects of lattice damage is to physically separate the grating from the active layer of the laser.

In practice, this is achieved by introducing an additional "separation" layer. To maintain effective interaction of the light wave with the grating, the thickness of such a layer should not exceed a micrometer.

For example the SCH (graded refractive index, separate confinement heterostructure) buried-heterostructure laser diode shown in Fig.6, which emits at 1.55 µm. This device features a first-order grating of 0.234 µm spacing fabricated by holographic photolithography. A buried-heterojunction stripe geometry was used for current confinement and for transverse mode control. A transparent InP window structure served to isolate one of the end facets to suppress Fabry-Perot oscillation. It should be noted that this type of window structure also suppresses a threshold degeneracy in which two longitudinal modes closely spaced in frequency about the ω_m.

Lasers fabricated with the structure shown in Fig. 7.7 have been shown to have threshold current as low as 11 mA and to exhibit stable single longitudinal mode CW operation at ambient temperatures up to 106 °C. Life tests on these

devices have indicated only minimal degradation over thousands of hours of operation. The average output degradation rate was 0.56% per thousand hours, which is comparable with that of conventional Fabry-Perot buried-heterostructure lasers.
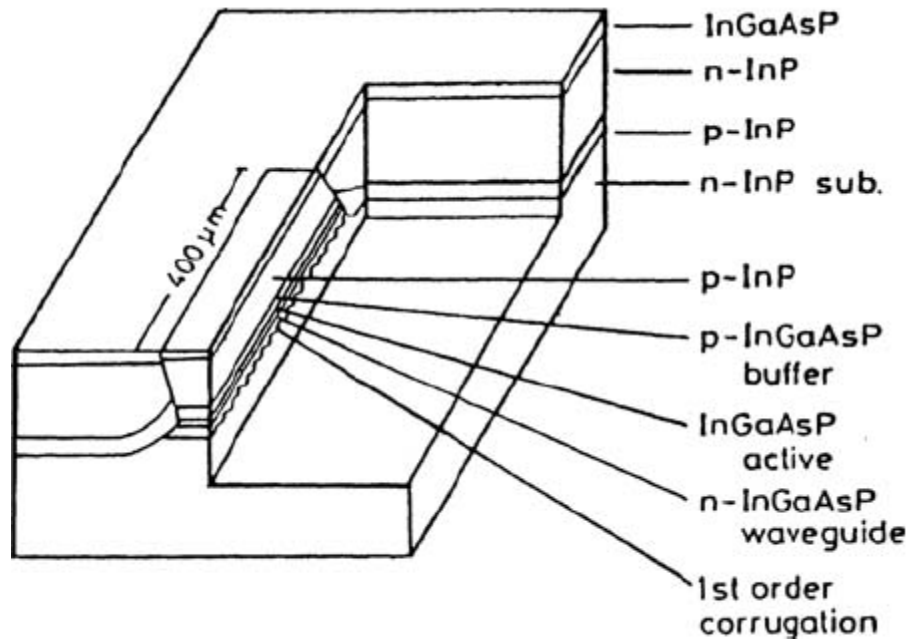

Fig. 7.7. An SCH-DFB-BH laser

The active region can also be isolated from the grating region by using the distributed Bragg reflection (DBR) structure shown in Fig. 7.8. In such a device, two Bragg gratings are employed, which are located at both ends of the laser and outside of the electrically-pumped active region. In addition to avoiding nonradiative recombination due to lattice damage, placement of two grating mirrors outside of the active region permits them to be individually tailored to produce single-ended output from the laser. In order to achieve efficient, single-longitudinal-mode operation, one distributed reflector must have narrow bandwidth, high reflectivity at the lasing wavelength, while the other must have relatively low reflectivity for optimal output coupling.

The need to create gratings with very precisely maintained corrugation parameters in such a multilayer and already complex structure as a modern injection laser, of course, complicates fabrication of DBR and DFB lasers.

In cleaved-end-face lasers, the light emission occurs at a wavelength determined jointly by the gain curve and the modal characteristics of the laser. Lasing occurs for the mode (or modes) that have the highest gain. When the laser is pumped well above the threshold level, usually a number of longitudinal modes lase simultaneously. It is very difficult, if not impossible, to obtain single mode oscillation.
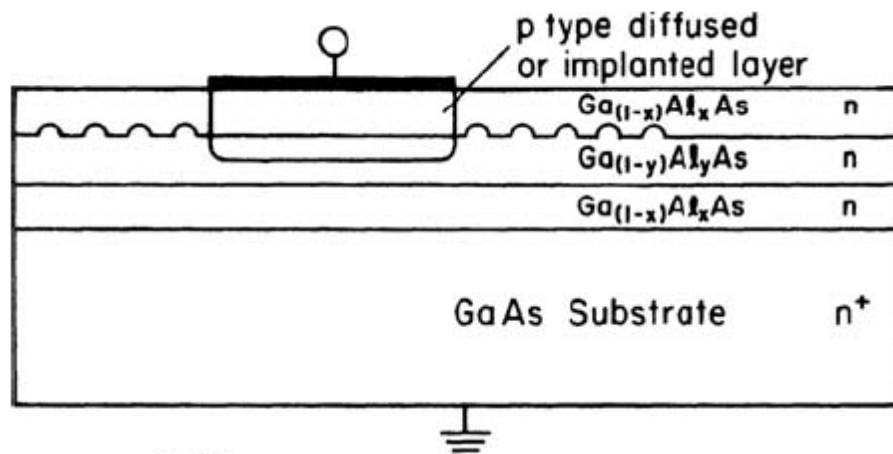

Fig. 7.8. Integrated DBR laser

In the case of DFB or DBR lasers, the emission wavelength is, of course, affected by the gain curve of the laser, but it is primarily determined by the grating spacing λ, as given by (7.4). The spacing between the m_{th} and the $m \pm 1$ modes is generally so large compared to the linewidth of the laser gain curve that only one mode has sufficient gain to lase. Thus single-longitudinal-mode operation is obtained relatively easily in distributed feedback lasers. This gives them a distinct advantage over reflectiveend-face lasers in many applications. The controllability of emission wavelength makes DFB lasers particularly useful in wavelength multiplexing applications.

In addition to providing a means of accurately selecting the peak emission wavelength, grating feedback also results in a narrower linewidth of the optical

emission. The spectral width of the emission line is established by a convolution of the laser gain curve with the mode-selective characteristics of the laser cavity. Since the grating is much more wavelength selective than a cleave or polished end-face, the resulting emission linewidth of a DFB or DBR laser is significantly less than that of reflective-end-face laser. The single-mode linewidth of a conventional cleaved-endface laser is typically about 1 or 2 Å, (about 50 GHz). The narrower linewidth obtainable with distributed feedback lasers is particularly important in optical communications applications, because the modulation bandwidth is ultimately limited by the linewidth of the laser source.

The improved temperature stability of distributed feedback lasers results from the fact that the shift of emitted wavelength in the cleaved laser follows the temperature dependence of the energy bandgap, while the shift in wavelength of the DFB laser follows only the temperature dependence of the index of refraction. This improved stability of the DFB laser makes it very useful in applications where wavelength filtering or wavelength divisionmultiplexing are employed.

The results of research and development of DFB/DBR lasers have been transferred to the commercial marketplace, where there are many suppliers. These lasers are now commercially available with emission wavelengths throughout the near-infrared range of 730–2800 nm.

7.2.1. Quantum dot DFB lasers

Semiconductor DFB or DBR lasers with narrow linewidths are of outmost importance for a variety of applications, the most important ones being communication and LIDAR. Conventional single mode lasers based on quantum wells have linewidths of the order of one to a few MHz; reducing the linewidth requires the addition of an external feedback, and a stabilization scheme by means of some control circuitry which enables to reach linewidths of about 100

kHz. A much better solution is a diode laser chip that can provide very narrow linewidths without the need for complex external additions. Such lasers are possible provided that the gain medium comprises quantum dots (QDs).

A quantum dot (QD) is a semiconductor structure with all three spatial dimensions less than 10 nm. Electrons, holes, and excitons are confined by the dot structure, which can significantly change the behavior of the material in which the quantum dots are formed.

By introducing quantum dots into the grating region of a DFB laser it is possible to improve the operating characteristics of the device. For example, self-assembled InAs/InAlGaAs QDs on InP/InGaAs grating structures have grown by molecular beam epitaxy and ridge-waveguide QD DFB lasers with a stripe width of 3 μm were fabricated. Single-mode cw lasing at a wavelength of 1.56 μm was achieved at room temperature, and pulsed-mode lasing was observed up to 70 °C.

The linewidth of the QD DFBs was more than one order of magnitude narrower than that of conventional quantum well (QW) DFBs at comparable output powers.

Threshold current was = 14 mA, differential efficiency = 0.33 W/A, and side-mode suppression ratios greater than 50 dB.

7.3. Vertical cavity lasers

Vertical cavity lasers are a relatively new type of semiconductor laser that emit light through their surface rather than through their end faces. In fact, they are often called vertical-cavity surface-emitting lasers (VCSELs). The typical device structure of a VCSEL is shown in Fig. 7.9. The key elements are an active layer, two DFB mirrors and a contact window to allow light to be emitted from the surface. In the particular device shown, a strain-compensated, quantum-well, 7-layer active region is sandwiched between two AlGaAs/GaAs mirrors by wafer fusion.
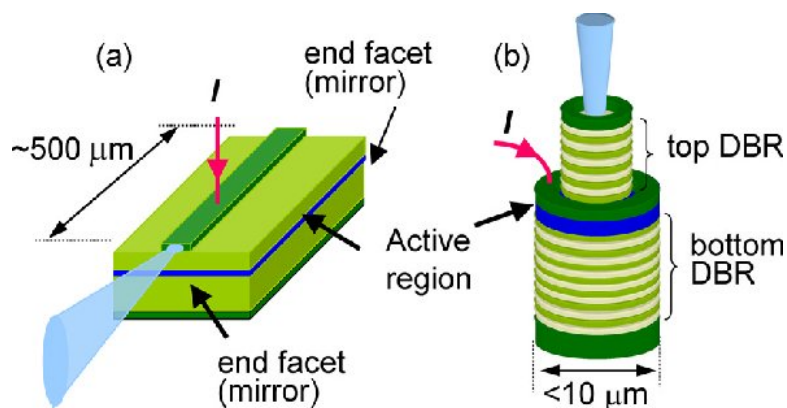
Fig. 7.9. Schematics of (*a*) edge-emitting laser and (*b*) vertical cavity surface-emitting laser

Typically, the active region is a multilayer, double heterostructure diode, as in a conventional Fabry-Perot end face laser. However, in this case, lightwaves travel in the direction perpendicular to the junction plane. They are reflected by top and bottom "mirrors" which consists of multilayer structures, alternating layers of materials with differing indices of refraction that are approximately one half wavelength thick (at the lasing wavelength in the material). For this spacing, the lightwaves reflected from each interface between layers positively reinforce waves reflected from all of the others in the reverse direction, so an effective "mirror" is formed. The overall reflectivity of the multilayer structure depends on the reflectivity at each interface and on the number of layers. The reflecting semiconductor layers are usually grown by either MOCVD or MBE, since sub-micron thickness is required. In some cases the reflecting layers can be thin films of dielectrics.

VCSELs have a number of advantages over Fabry-Perot endface diode lasers in certain applications. They can be made in a circular shape, as shown in Fig. 7.9, *b*, with a diameter matched to the core diameter of an optical fiber to optimize coupling. Because of their relatively largely emitting area, the beam remains relatively well collimated and does not diverge due to diffraction as does that of an endfaceemitting laser diode. Also, VCSELs can be conveniently arranged into a surface array to couple to a fiber ribbon or bundle, or to provide direct chip-to-chip coupling through the air. The surface orientation of emission also facilitates integration of a VCSEL with planar electronic circuitry.

The integration of VCSELs and CMOS circuitry provides an effective way to make IOC transceivers in which low-area, low-power, high-speed CMOS circuits are used to modulate VCSELs in a surface-emitting array.

VCSELs are part of the general class of lasers known as distributed Bragg reflector (DBR) lasers because their required positive feedback is provided by distribute Bragg-type reflections.

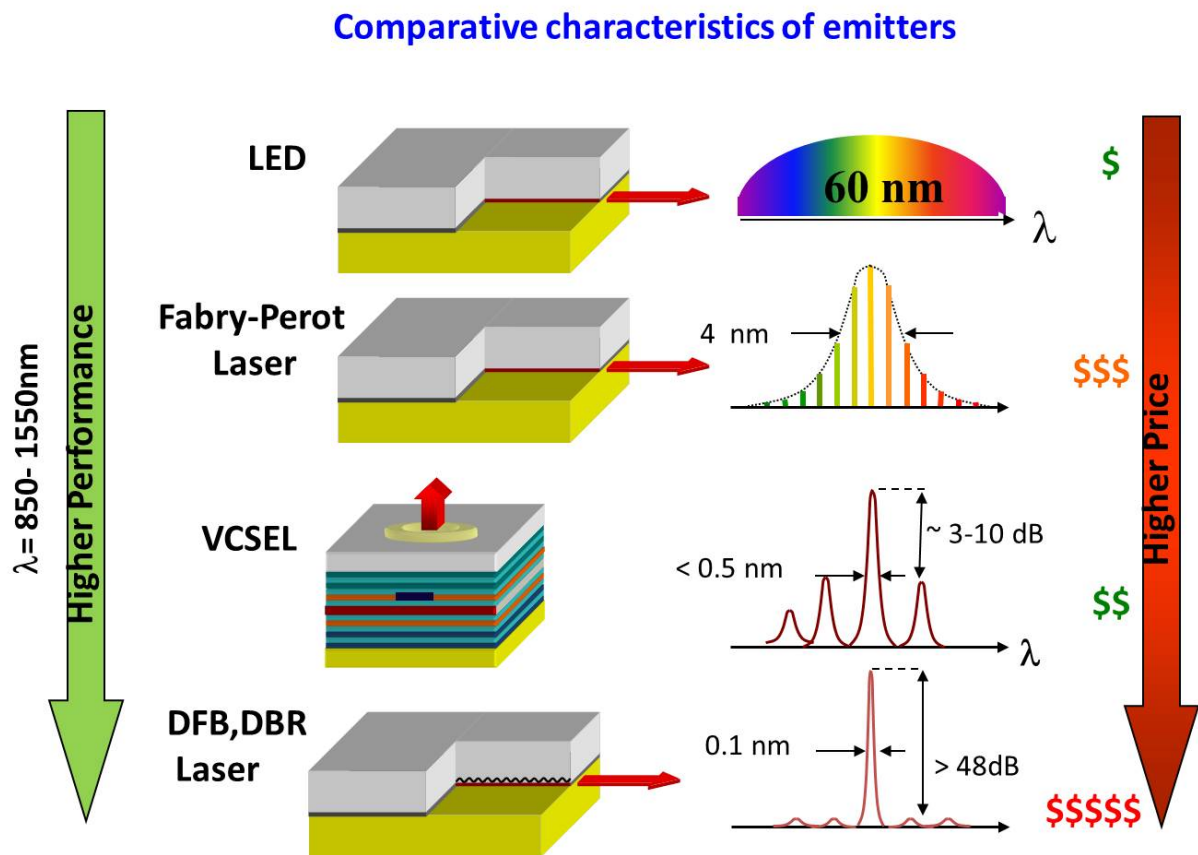Comparative characteristics of emitters for IOEC are shown in Fig. 7.10.

Fig. 7.10. Comparative characteristics of emitters for IOEC

One of the main requirements for integrated optical lasers is ability to integrate on a single substrate with other components of the integrated optical circuit. Semiconductor injection lasers are made on a substrate of a binary compound $A^{III}B^{V}$, on top of which multicomponent layers of various compositions epitaxially grow. As a rule, injection lasers are formed on a well-conducting, heavily doped substrate.

8. APPLICATIONS OF INTEGRATED OPTICS AND CURRENT TRENDS

Although integrated circuits is a relatively new field of endeavor, numerous applications of optical integrated circuits (OIC's) to the solution of current engineering problems have already been implemented and some OIC's are now available as "off-the-shelf" commercial products. Of course, optical fiber waveguides, the companion element of OIC's in an integrated-optic system, are already well recognized as being very useful consumer products.

8.1. RF spectrum analyzer

Probably the earliest demonstration of a multi-element OIC that was performed was the hybrid implementation of the real-time RF spectrum analyzer. The purpose of this spectrum analyzer, is to enable the pilot of a military aircraft to obtain an instantaneous spectral analysis of an incoming radar beam, in order to determine if his plane is being tracked by a ground station, air-to-air missile, etc. Obviously, such information is required if he is to be able to quickly take effective evasive action. Of course, the frequency content, or *signature*, of all enemy radar signals that are likely to be encountered would have to be available for comparison, probably stored in the memory of the plane's onboard computer.

A diagram of the integrated-optic spectrum analyzer is shown in Fig. 8.1. Light from a laser source is coupled into a planar waveguide, in which it passes first through a collimating lens, then through a Bragg-type acousto-optic modulator. The RF signal to be spectrally analyzed is applied to the acoustic transducer that generates the sound waves, causing them to have a time varying period. Thus the deflection angle of the optical beam at the output of the modulator is a function of the RF signal. A second lens is used to focus the optical beam onto an array of photodetectors. If more than one frequency

component is present in the rf signal, the light beam is divided into corresponding components that are focused onto different detector elements.
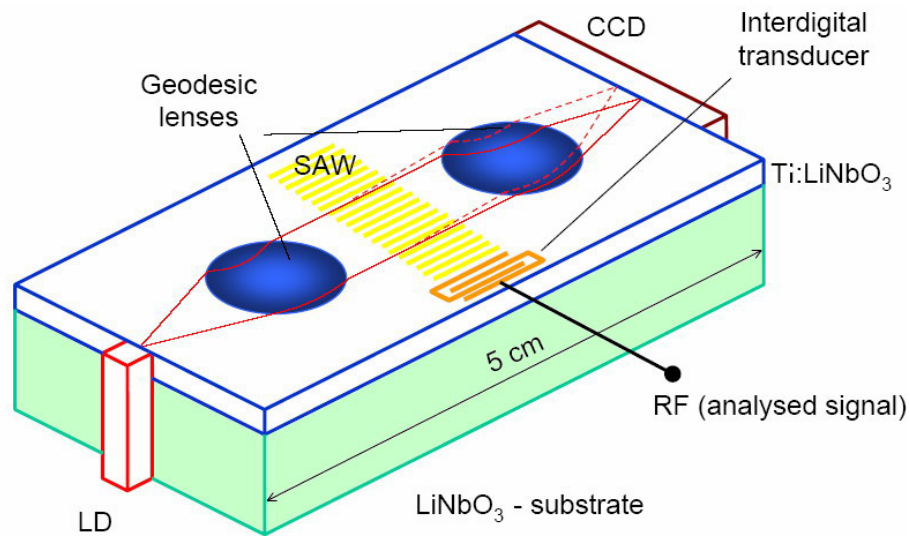

Fig. 8.1. Diagram of an integrated-optic rf spectrum analyzer

Each detector element represents a particular frequency channel, and, since photo-diodes generally have square law response, the output signal from any channel is proportional to the RF power at that frequency. The advantage of an integratedoptic spectrum analyzer, as compared to an electronic one, is that only a few optical elements are needed to perform a function that would otherwise require thousands of electronic elements.

The development of working models of the integrated optic RF spectrum analyzer took place at several different laboratories and extended over a number of years. The RF spectrum analyzer described above are excellent example of hybrid optical integrated circuit technology. By fabricating the laser diode in GaAlAs, the detector array in silicon. and the Bragg modulator in $LiNbO_3$. One can use the best features of all three materials to advantage. The major disadvantage of the hybrid approach is that all of these substrate materials must be carefully aligned and permanently bonded with micrometer-tolerance precision. Thermal expansion and vibration must somehow be prevented from destroying the alignment. Despite these difficulties, hybrid OIC's have been

demonstrated to be viable structures, and will continue to be used in many applications even after monolithic technology has been fully developed.

Technological difficulties in fabricating geodetic lenses and Luneberg lenses, which are part of the integrated optical spectrum analyzers, dramatically reduce the manufacturability of the latter. To reduce the requirements for the manufacturing technology of spectrum analyzers and further improve their characteristics, two new directions have been outlined. The use of focusing elements of the diffraction type — Fresnel lenses — and the type of gratings with a variable period, as well as the development of "lensless" designs of spectrum analyzers will greatly simplify their technology while maintaining the high efficiency of the device

Fig. 8.2 illustrates a diagram of an integrated optical spectrum analyzer based on amorphous arsenic sulfide chalcogenide As_2S_3 on a Si/SiO_2 substrate with flat Fresnel gradient lenses. Lenses are created in a waveguide from arsenic sulfide by electron beam lithography. For electro-acoustic transducer use a layer of zinc oxide with a thickness of 1.7 μm. With a focal length of 5 mm, diffraction lenses have a very high diffraction efficiency – up to 93% with a control signal power of 230 mW.
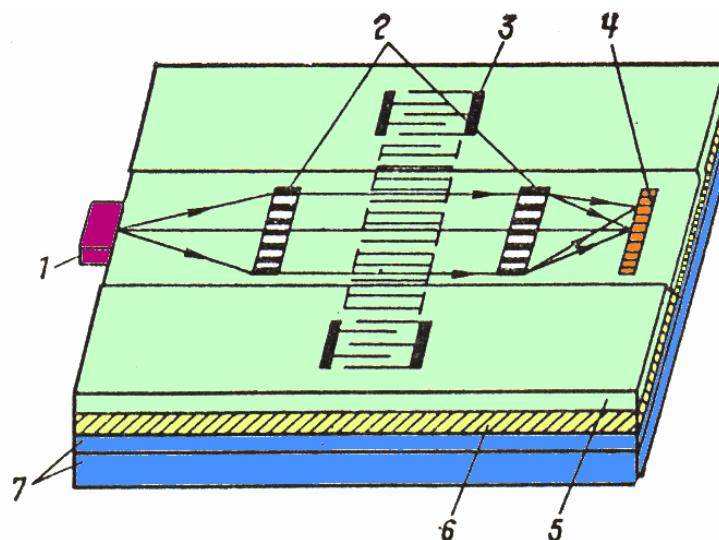

Fig. 8.2. Diagram of an integrated optical spectrum analyzer based on amorphous arsenic sulfide chalcogenide As_2S_3 on a Si/SiO_2 substrate with flat Fresnel gradient lenses:

1 – LD, *2* – Fresnel gradient lenses, *3* – electro-acoustic transducer,

4 – PD, *5* – waveguide layer, *6* – ZnO_2, *7* – substrate

8.2. Analog-to-digital converter (ADC)

An analog-to-digital conversion method has been implemented in an optical integrated circuit that is capable of one-bit electro-optical AD conversion at a 100 MHz rate. The OIC incorporates two 3-dB couplers and a phase shifter, formed in a pair of straight waveguides, as shown in Fig. 8.3. The waveguides were fabricated by Ti diffusion of a LiNbO₃ substrate. The phase shifter was formed by a Ti double-diffusion, as shown in Fig. 8.3. An Al₂O₃ 1100 Å thick buffer layer was used, separated by a 4 μm, gap between the waveguides to suppress dc drift. Waveguide spacing was 5.4 μm and device length was about 2 cm.
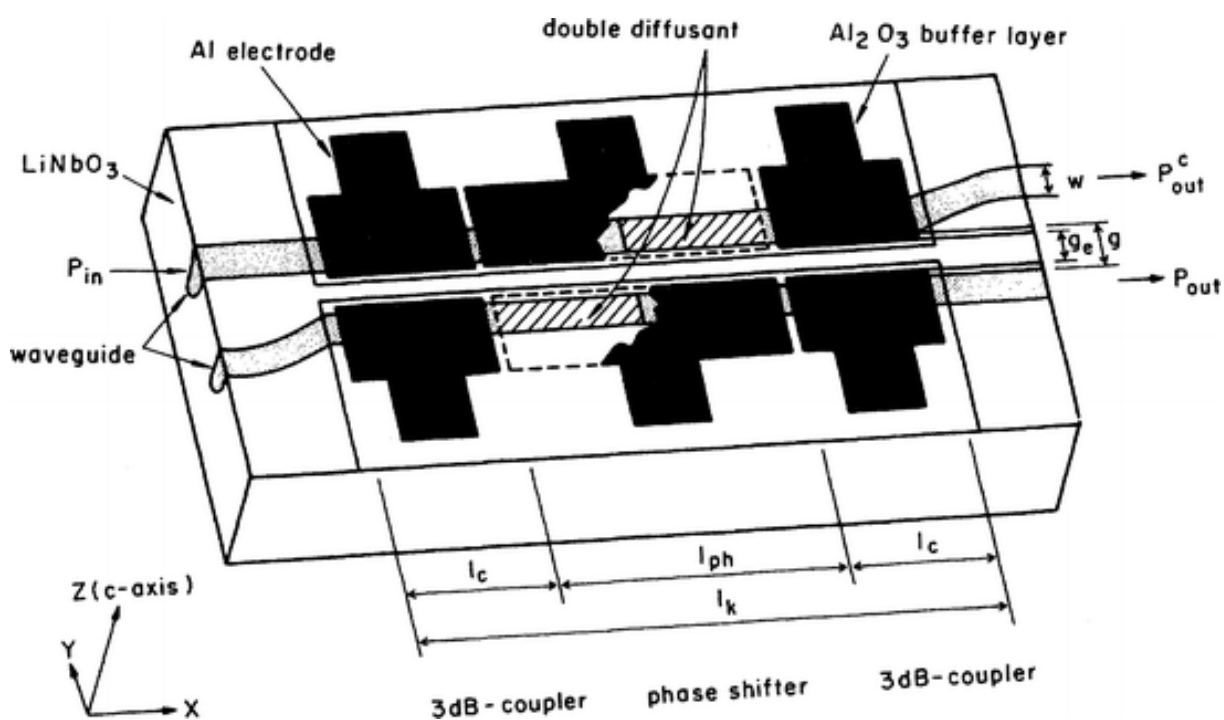

Fig. 8.3. Diagram of an integrated-optic analog to digital converter

The configuration of two electro-optic couplers and a phase shifter forms a balanced-bridge modulator, with two complementary outputs which are equally affected by fluctuation of the light source. Hence a serious source of conversion error is inherently eliminated in this OIC. The integrated ADC was operated with a 1.15 μm-wavelength He-Ne laser source at bit rates up to 100

MHz. This initial success of high speed analog-to-digital conversion points the way towards more sophisticated, multi-bit, and monolithic OIC's. However, much remains to be done, especially in regard to development of a monolithic, high-speed, electronic or optical comparator to be incorporated into a fully monolithic ADC system.

8.3. An IO optical disk readhead

Optical disk information storage has found widespread use for computer data, as well as for video and audio reproductions. High data density and low background noise are key advantages of this method. However, relatively sophisticated optics must be used to insure good resolution and tracking of the light beam that is used to read information off the disk. For example, the optical readheads used in commercially available audio compact disk (CD) players often have eight or nine discrete optical elements, all of which have to be held in exact alignment in the face of much shock and vibration.

As an alternative, an integrated-optic optical disk pickup device capable of detecting readout and focus/tracking error signals has been designed and fabricated , as shown in Fig. 8.4. The OIC was formed by depositing a planar #7059 glass waveguide on a SiO_2 buffer layer on a silicon substrate. The light source was a butt-coupled GaAlAs laser diode. A chirped and curved focusing grating pattern coupler fabricated by electron beam direct writing lithography was used to focus the beam onto the disk, as well as to refocus the reflected beam back into the waveguide. A twin-grating focusing beam splitter served to divide the reflected beam into two beams which were focused onto two pairs of photodiodes formed in the Si substrate.

In operation, the pickup head provides not only a readout signal but also focus and tracking error signals. When the light beam is focused, the return beams hit both diodes of each pair equally. If the readhead is too close to the disk, the beams fall more on the outer diodes, while if it is too far away they fall

more on the inner diodes. Tracking error is detected when the total intensity of the return beam reaching the left pair of diodes is not equal to that reaching the right pair. Thus conventional electronic comparators, sensing the photo-currents from the diodes, can be used to develop error signals to drive position correcting actuators. The OIC readhead of Fig. 8.4, which has dimensions of only 5×12 mm, obviously has the advantage of being relatively insensitive to shock and vibration, as compared to a readhead fabricated from discrete optical components. While this OIC was first proposed as an optical disk pickup device, the same basic arrangement can be used more generally as a fully integrated interferometer position/displacement sensor with direction discrimination. Such an interferometric sensor would be useful in a variety of high-precision positioning applications in which submicrometer accuracy is required.
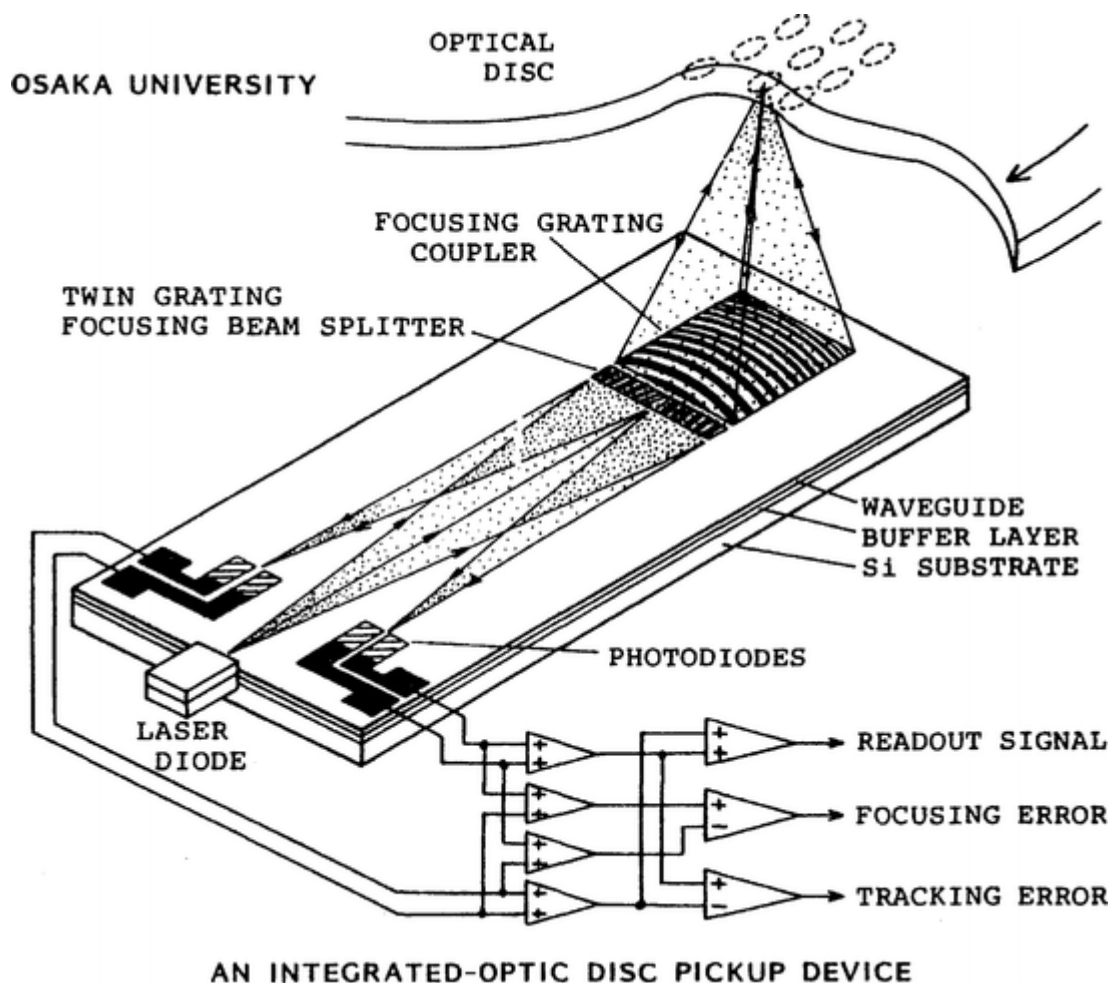

AN INTEGRATED-OPTIC DISC PICKUP DEVICE

Fig. 8.4. An integrated-optic disc pickup device

8.4. Integrated optical processors

Based on the principles of integrated optics, circuits that perform various mathematical operations can be fabricated. This makes it possible to solve many computational problems by the "optical" method (for example, simple arithmetic operations, matrix and vector multiplication).

The optical addition of two beams is performed using a beam splitting element – a waveguide *Y*-splitter. On *Y*-splitters, an optical signal power loss of approximately 3 dB (1.4 times) occurs. But this pays off by the fact that the result of the addition is represented by the optical signal, and, therefore, any optical operations can be performed on it further.

There are no direct methods for directly subtracting two optical amplitudes, as in addition. Several methods have been developed for generating an optical signal proportional to the difference of two voltages. One of them is based on the use of a strip waveguide with electrodes deposited on both sides of it. The phase shift of the transmitted light wave in this case is proportional to the voltage difference applied to each of the electrodes. To translate phase modulation into intensity modulation, which is most easily detectable, such a phase shifter must be placed in one of the arms of the interferometer.

Arithmetic operations also include multiplication. The multiplication of optical signals based on strip waveguides is reduced to a cascade connection of a pair of couplers equipped with electrodes to control the degree of branching. Multiplication can also be done by using two cascade, series-connected electro-optical deflectors. Then the intensity of the double-diffracted light beam will be proportional to the product of the voltages applied to each of the deflectors. Similarly, instead of an electro-optical deflector, appropriately oriented surface acoustic waves can be used.

The planarity of the elements of integrated optics somewhat limits the possibilities of modeling based on certain mathematical processes. Nevertheless, integrated optical circuits have been proposed for performing relatively complex

mathematical operations, such as subtracting and scalar product of vectors, multiplying the matrix by a vector and multiplying the matrices.

Matrix calculations are widespread mathematical operations that are often encountered in solving applied problems. The involvement of optical methods is due to the need for a significant increase in the speed of computation.

The architecture of integrated optical processes for performing matrix and vector operations (Fig. 8.5) (parallel processor) and counter-stream (counter-stream processor) differ in the methods of input, output, and execution of the calculation itself. To build an integrated optical processor, the choice between on-line and parallel approaches should also be based on taking into account the necessary speed of calculations and the permissible complexity of electronic equipment. An on-line processor for performing an operation with an N matrix (vector) requires an N-element deflector modulator. A parallel processor that implements a direct multiplication algorithm requires a modulator with the number of elements N^2. Therefore, in terms of total number of modulators, preference should be given to a processor with a counter-stream architecture. However, the integrated optical processor unit is framed by a large number of electronic circuits, which are much simpler in a parallel processor. Thus, the choice of processor architecture is always a compromise.
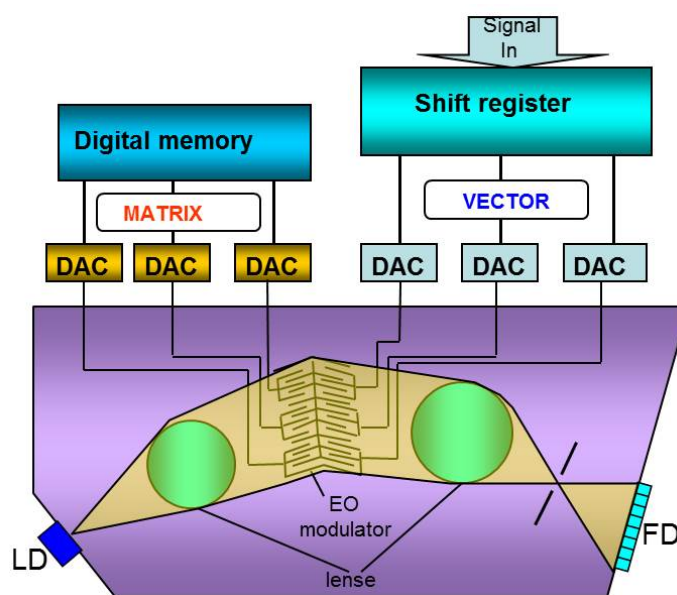

Fig. 8.5. An integrated-optic processor performing matrix and vector multiplication; DAC – digital to analog convertor

8.5. Heterodyne interferometer

The efficient interaction of surface acoustic waves with guided optical waves in combined optical and acoustical Ti:LiNbO$_3$ waveguide structures is the basis for wavelength-selective, tuneable polarization converters. By the combination of such converters with polarization splitters a whole family of wavelength-selective devices has been developed for applications in optical communications, e.g. tuneable filters, switches and (de-) multiplexers.

Moreover, acousto-optical mode conversion offers further attractive applications in optical metrology due to the frequency shift associated with the acousto-optical interaction. An example is the acousto-optical heterodyne interferometer (Fig. 8.6).
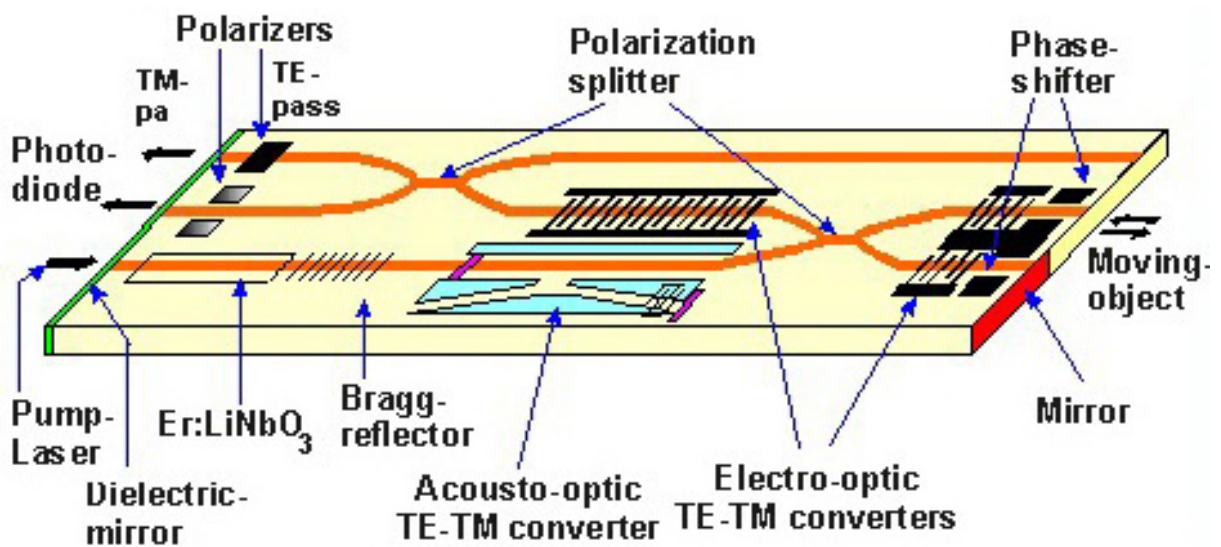

Fig. 8.6. Schematical diagram of an integrated optical heterodyne interferometer in LiNbO$_3$

Integrated optical heterodyne interferometers for applications such as Doppler velocimetry, frequency analysis of vibrating surfaces and contact-free distance measurements promise high sensitivity, rugged construction, small overall size and possibly low fabrication costs. Different versions of interferometers have recently been developed in glass, on silicon, with polymer waveguides and in LiNbO$_3$; the first optical systems with integrated optical

interferometric sensor on silicon and in glass are now commercially available. Contrary to glass, silicon and polymers, LiNbO$_3$ allows to take advantage of its excellent electro-optical and acousto-optical properties; both can be exploited to develop heterodyne interferometers of ultimate sensitivity using integrated frequency shifters, beam splitters and polarizing optics.

TM-polarized light of frequency f_0 is fed into the integrated optical heterodyne interferometer chip (see Fig. 8.6). It first passes the acousto-optical TE-TM mode converter; half of the optical power is converted into TE polarization. The frequency of the generated TE-mode is shifted by the acoustic frequency f_a (~170 MHz).

The Er-diffusion-doped DBR-waveguide laser was monolithically integrated on the interferometer chip. An Er-diffusion-doped DBR waveguide laser, pumped by a laser diode ($\lambda = 1480$ nm), was used as integrated coherent light source with $\lambda = 1530$ nm emission wavelength.

Both, TE- and TM-polarized waves are separated by the subsequent passive polarization splitter and fed into the reference and measuring arms of the (Michelson) interferometer. In both arms the electro-optic TE-TM mode converters and phase shifters are used to rotate the polarizations of the back-reflected waves by 90° without an additional frequency shift. The result is that both waves are recombined by the polarization splitter and fed into the output arm without any principle loss and without feedback to the optical source. The reference arm is terminated by a metallic end face mirror. In the output waveguide, the polarization of the reference and measuring waves are orthogonal, so the waves do not interfere with each other. Therefore, an additional electro-optic TE-TM converter serves to generate polarization components which can interfere. The further polarization splitter is used to separate the TE- and TM-polarized waves. A vibrating mirror at about 20 mm distance to the chip acts as object to be measured. Two PIN/FET detector/preamplifier modules are used as a single detector heterodyne receiver.

8.6. Monolithic wavelength-multiplexed optical source

One of the applications for which optical integrated circuits were proposed early in the history of the field is an optical-frequency-multiplexed transmitter, in which a number of DFB lasers, operating at different wavelengths, are coupled into a single fiber transmission line. An OIC of this type has, in fact, been fabricated using GaAlAs monolithic technology.

Six DFB lasers, operating at wavelengths separated by 20 Å, were fabricated on a 5 mm square GaAs substrate by a two-step LPE growth process. The lasers had a separate confinement heterostructure (SCH). Third-order gratings were made on the surface by chemical etching, by using a mask made by holographic lithography. The lasers were coupled to undoped $Ga_{0.9}Al_{0.1}As$ waveguides by direct transmission, as shown in Fig. 8.7. The lateral dimensions of the lasers and waveguides were denned by mesa etching down to the GaAs substrate to produce stripes that were 20 μm wide and 3 μm thick. The separation of the lasers was 300 μm, and the waveguides were curved through bends of minimum radius equal to 4 mm, in order to bring them together in a confluent coupler, as shown in in Fig. 8.8. The output of the coupler was obtained via a single waveguide that was butt coupled to an optical fiber.

The lasers were operated by applying 100 ns current pluses at a repetition rate of 1 kHz. The differential quantum efficiency of the lasers was measured to be 7%, and the waveguide loss coefficient was about 5 cm^{-1}. The threshold current densities of the lasers were in the range from 3 to 6 kA/cm^2 at room temperature. The wavelength separation between lasers was measured to be 20±5 Å. No difficulty was encountered in separately modulating the six lasers, and the overall differnetial quantum efficiency, measured at the launching output terminal, was about 30%. Thus, this early version of a monolithic chip represented a usable OIC, even though further refinements were to yield better efficiency.
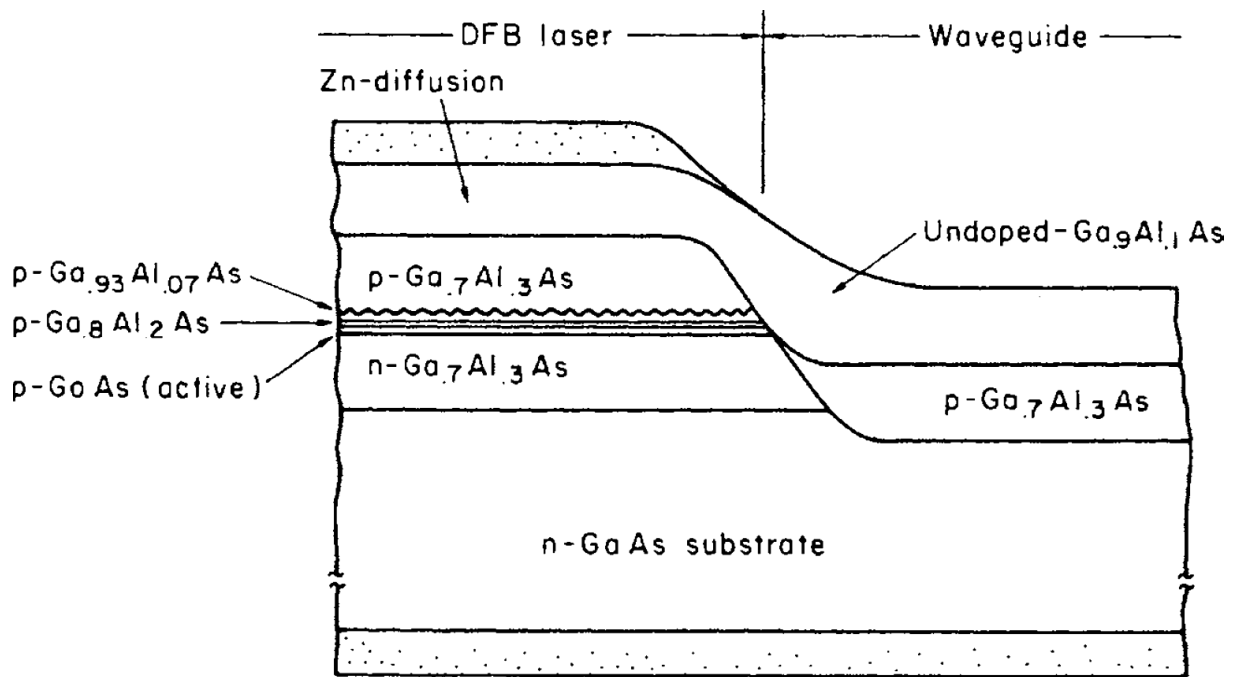
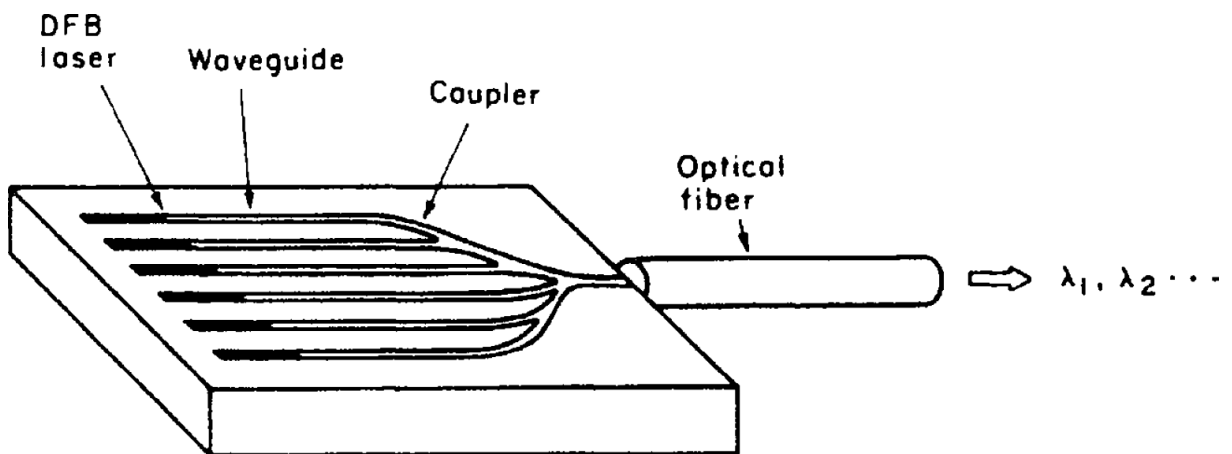*Fig. 8.7.* DFB laser coupled to a GaAlAs waveguide by direct transmission

Fig. 8.8. Schematic drawing of a wavelength-multiplexed light source

Work on monolithic wavelength-multiplexed optical sources employing DFB lasers has continued over the years, with more recent work being directed toward 1.3 μm and 1.55 μm wavelength lasers for optical-fiber telecommunication systems.

A study of wavelength-division multiplexed lighwave systems which has led to the conclusion that, in order to be cost effective, it is necessary to fabricate multi-wavelength laser transmitters by monolithic integration on one chip to reduce the cost of packaging and control circuitry by sharing them

among all of the wavelengths. An example of such monolithic integration is provided by the frequency-division multiplexed ten-channel tunable DFB laser array. The lasers are tunable, multi-section, quarter-wave-shifted, strained InGaAsP MQW devices. The lasing frequencies of channels are spaced within a 10 GHz range. The linewidth of each channel is less than 2.3 MHz. A monolithically integrated chip containing 21 DBR lasers has been produced.

The use sampled grating distributed Bragg Reflectors, in which the grating is not continuous, but rather is formed in bursts so that it has two inherent periodicities. This type of DBR grating makes the selection of laser emission wavelength more accurate. The 21 InP/InGaAsP MQW lasers on the chip had emission wavelengths spanning over 40 nm with 0.8 nm spacing, centered at approximately 1.56 μm.

8.7. Integrated-optic Doppler velocimeter

An integrated-optic Doppler velocimeter which employs both an optical fiber link and an OIC to measure velocity has been demonstrated, as shown in Fig. 8.9. The optical integrated circuit was fabricated in a z-propagation LiNbO$_3$ substrate with Ti diffused waveguides. Laser beam lithography, with 0.2 μm accuracy, was used for waveguide patterning. The light source was a linearly polarized He-Ne laser. TE polarized light was focused into the input waveguide by a 20% lens and then split by a Y-branch coupler into a signal beam and a reference beam.

In order to keep the outgoing beams separate from the reflected return beams. TE polarization was maintained on outgoing light while TM polarization was used for reflected light. Of course, polarization-maintaining optical fiber was used. In the case of the reference beam the TE/TM conversion was accomplished by means of an electro-optic mode converter. For the signal beam a quarterwave plate was used to produce TE/TM mode conversion. An absorptive TE/TM mode splitter was used to route the return signal beam to the

avalanche photodiode (APD) which served as a mixer and detector. An electro-optic modulator was used to impress the reference modulation frequency f_R onto the reference beam. The Doppler effect produced a shift in the signal beam frequency from f_0 to $f_0 + f_s$, where f_s is given by

$$f_s = \frac{2v}{\lambda_0},$$

where v is the velocity and λ_0 is the vacuum wavelength. After being recombined by a Y-branch coupler the reflected signal beam and reference beam were mixed in the avalanche diode. Because of the nonlinear response characteristic of the APD (it is a square-law device, i.e., response is proportional to the square of the electric field strength) the output photocurrent contained a beat frequency component $f_R - f_s$. Thus f_s, and hence the velocity, was determined. For a velocity of 8 mm/s the measured f_s was 25 kHz, and the signal to noise ratio was 25 dB.
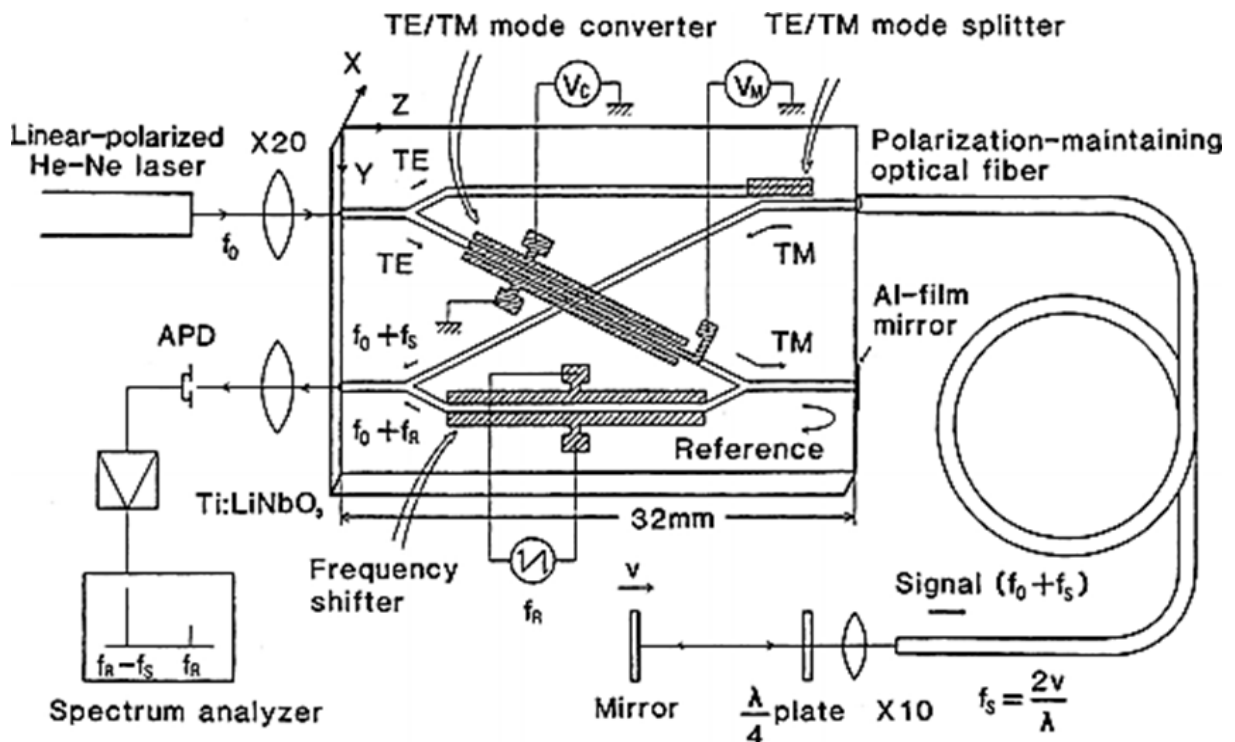

Fig. 8.9. An integrated-optic Doppler velocimeter

This Doppler velocimeter demonstrates that integrated-optic technology can provide the compact and rugged heterodyne optics need for highly accurate measurement of velocity and displacement. By adding a balanced-bridge waveguide optical switch to an interferometric circuit of this type it is possible to fabricate a time-division-multiplexed Doppler velocimeter which can measure two dimensional velocity components v_x and v_y.

8.8. OIC temperature sensor

The integrated-optic temperature sensor shown in Fig. 8.10, *a*. Requires no electrical connection, making it particularly useful in explosive or flammable environments in which an electrical sensor might be dangerous. The OIC is fabricated in a LiNbO$_3$ substrate with Ti diffused waveguides. It features a parallel array of three unequal arm-length Mach-Zehnder interferometers. The optical transmission of each interferometer varies sinusoidally with temperature, as shown in Fig. 8.10, *b*, with a period which is inversely proportional to the optical path length difference between the two arms.

The optical transmission P_{out}/P_{in} at the wavelength λ depends on both the effective index n_{eff} and the path length difference ΔL and is given by

$$\frac{P_{\text{out}}}{P_{\text{in}}} = \frac{\gamma}{z}\left[1 + m\,\cos\left(\frac{2\pi}{\lambda}b\Delta LT + \Delta\varphi_0\right)\right],$$

where the constant of proportionality b is given by

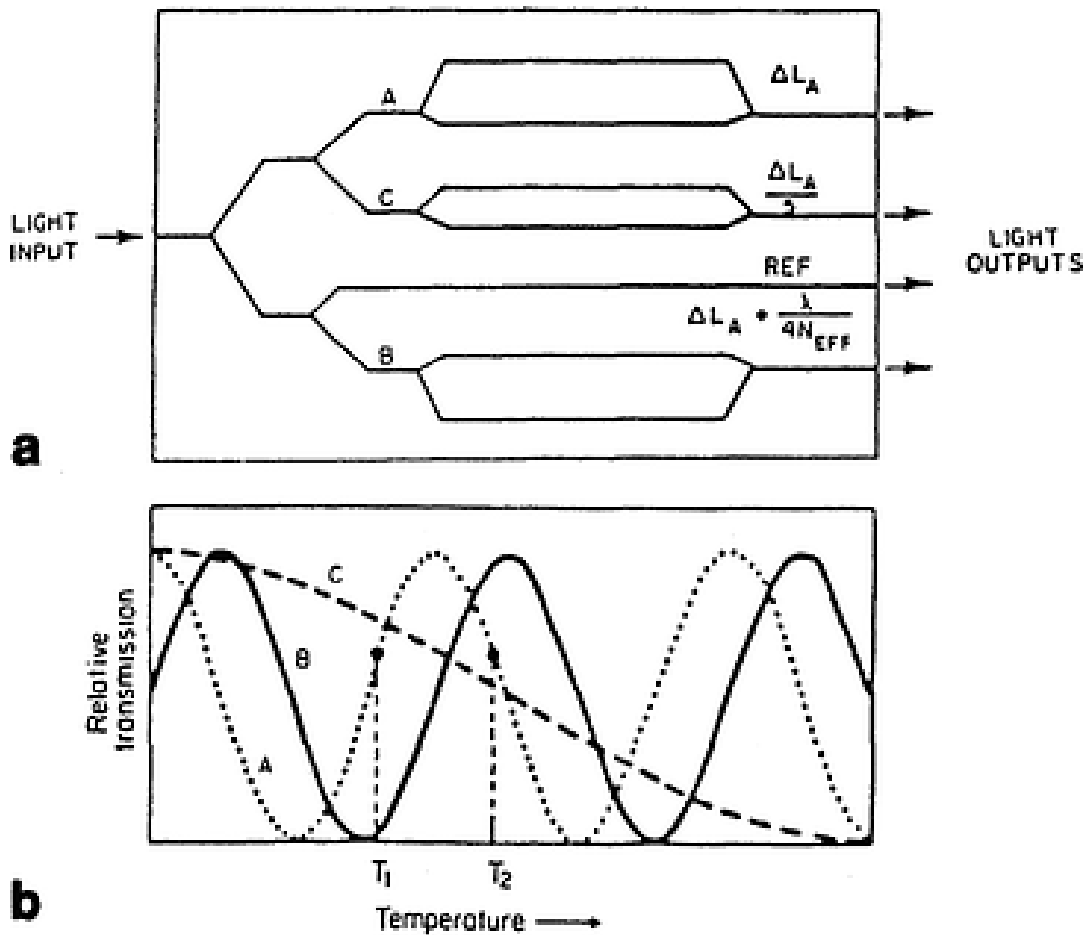$$b = \frac{dn_{eff}}{dT} + \frac{n_{eff}}{\Delta L}\frac{d(\Delta L)}{dT}.$$

Fig. 8.10. OIC temperature sensor:

a – device structure; *b* – optical transmission characteristics

Both n_{eff} and ΔL are functions of temperature T. The quantities γ and m are related to the insertion loss and depth of modulation of the interferometer, respectively. (For an ideal device $\gamma = m = 1.0$.) $\Delta\varphi_0$ is a constant for a given device.

By measuring the transmission of all three inteferometers one can determine the temperature. Two of the interferometers (A and B) have arm-length differences that are almost the same. Thus their transmission curves track close to one another, providing a high resolution in the temperature measurement. The third interferometer (C) has an arm-length difference which is only approximately one-fifth those of A and B. Thus one can determine which peak of the A and B transmission curves is being measured and temperature measurements can be made over a wide range. It has been reported that this

183

temperature sensing OIC can measure with an accuracy of $2 \cdot 10^{-3}$ °C over a 700 °C range, when used with a 6328 Å He-Ne laser as the light source. The sensor, of course, would be mounted at the point at which temperature measurement was desired and the optical input and output would be via optical fiber. Since the OIC chip is about 1 cm on a side, temperature measurements can be made on relatively small objects. Since the measurement signal is entirely optical, this device is relatively immune to electrical noise.

8.9. IO high voltage sensor

Integrated-optic Mach-Zehnder interferometers can be used to sense high voltage as well as temperature. A diagram of such a device is shown in Fig. 8.11. The waveguides are formed by Ti diffusion into a LiNbO$_3$ substrate. In this circuit the two branches of the interferometer are covered by metal electrodes which form a capacitive voltage divider. The electric field generated by the high voltage source induces a voltage on these electrodes which causes a relative phase shift between the optical waves in each arm, resulting in an intensity modulation of the output beam. The voltage-in/optical power-out transfer function is given by

$$P_{\text{out}} = \frac{\alpha P_{\text{in}}}{2} \left[1 + \gamma \cos \left(\frac{\pi V}{V_\pi} + \varphi_i \right) \right],$$

where P_{in} is the input power, φ_i is the intrinsic or zero voltage phase difference. V is the applied voltage and V_π is the half-wave phase shift voltage. The constants α and γ must be determined for a particular device. (For a perfect device $\alpha = \gamma = 1.0$.) Once the calibration curve has been determined for a particular sensor, the voltage can be accurately measured.
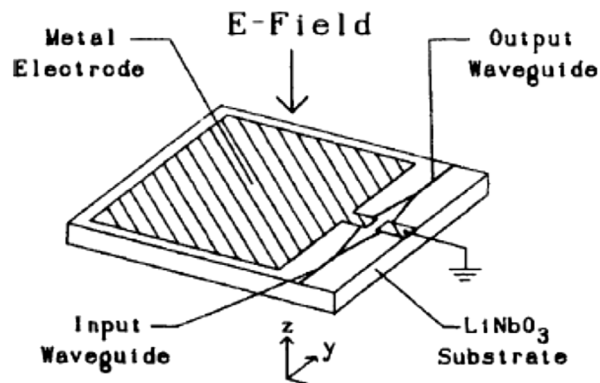
Fig. 8.11. IO high voltage sensor

Since this sensor operates on the principle of an induced voltage it is not necessary to make electrical contact to the high voltage source, and input to, and output from the sensor can be via optical fiber. Thus good high voltage isolation can be maintained. The immunity to electrical noise provided by an optical fiber link is also a particularly important advantage when operating in a high voltage environment. This integrated optical high voltage sensor could be used, for example, for monitoring line voltages in SF_6-gas insulated bus ducts such as are used in power plants and switching stations.

8.10. Wavelength meters and spectrum analyzers

By carefully arranging wavelength-selective optical elements in an OIC it is possible to make instruments that can measure an emission wavelength or spectrum.

For example a spectrophotometer that consists of two pn junction InGaAs/GaAs QW photodiodes stacked vertically with a DBR reflector between them, is shown in Fig. 8.12, *a*. The DBR reflector is composed of 101/2 pairs of AlAs/GaAs. Lightwaves entering the top surface pass through both photodiodes and are detected by them. However, because of the wavelength selective reflectance of the DBR, some of the incident light doesn't reach the bottom photodiode. The ratio of the top detector response to that of the bottom diode is given by:

185

$$\frac{I_{top}}{I_{bot}} = R_0 \frac{1+R(\lambda)}{1-R(\lambda)},$$

where R_0 is a constant of proportionality. The above ratio is a single-valued function of wavelength that can be used to measure it, as shown in Fig. 8.12, *b*, where the dotted curves are I_{top} and the solid curves are I_{bot}. The curve in (*c*) is a superposition of measurements at the three optical power levels shown in (*b*), indicating that measured wavelength is not sensitive to optical power
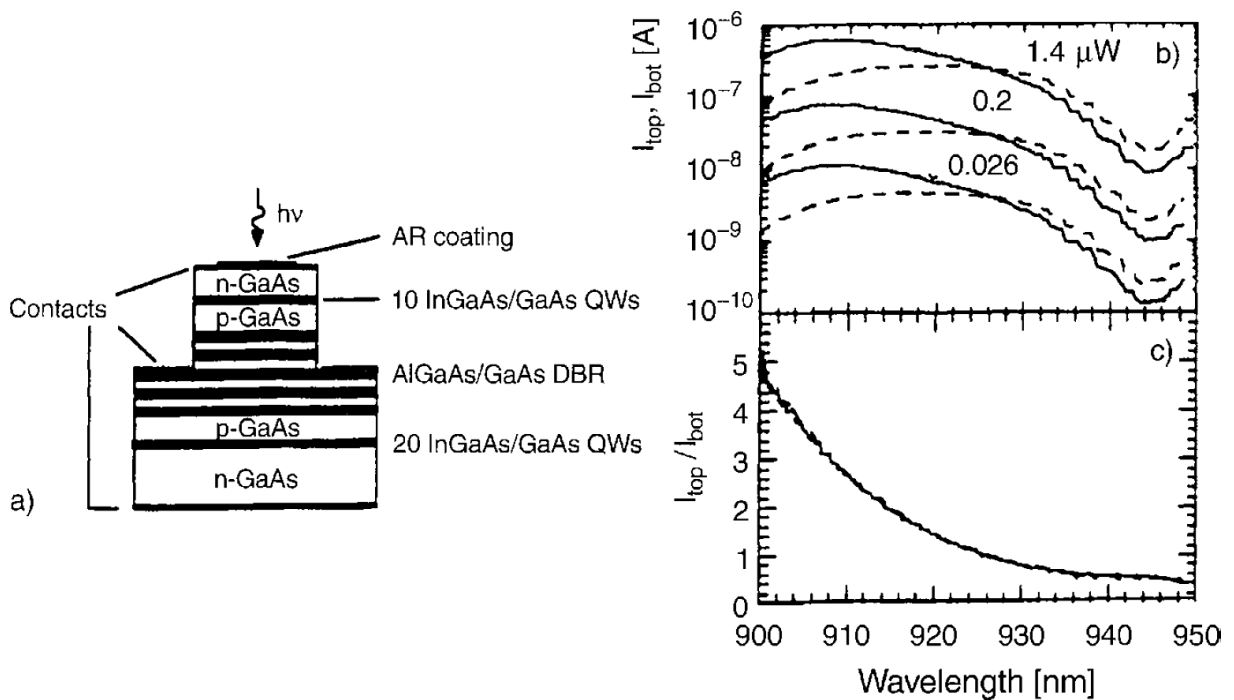

Fig. 8.12. Wavelength meter: *a* – device structure; *b* – response curves; *c* – response ratio

An integrated optic device also can be used to measure an optical spectrum as well as just a single wavelength. For example, an OIC optical spectrum analyzer device uses a tilted, chirped grating in a planar single-mode waveguide to diffract lightwaves of different wavelengths through different angles as they are coupled out of the waveguide. The light radiated out of the waveguide is captured by a vertical slab waveguide that directs it to a linear array of photodetectors. Each detector intercepts light of only a narrow range of wavelengths, so the optical spectrum of the light can be determined. The

fullwidth half-maximum resolution was 0.15 nm over a bandwidth range of 7.8 nm. In this device the grating, which has a chirp of −1.75 nm/cm and a focal length of 12 cm, performs both the spatial dispersion and imaging (focusing) functions.

8.11. Chemical and biological sensing applications

Integrated optic devices can be used to sense the presence and the concentration of various chemical elements. These sensors generally function by measuring the change in some optical property of the material in a waveguide that is produced by the presence of the chemical to be sensed. Both absorption spectro-photometry and attenuated total reflection spectrometry can be used, as well as fluorescence spectrometry. The waveguide may be either an optical fiber or a multilayer guide an OIC.

Many branches of human activity has broad needs to deploy better BOIChemical sensors for applications in medical diagnostics, environmental monitoring, rapid identification of threats, hazardous materials control, basic scientific research, and more. For example, point-of-care detection of biomarkers of injury or disease can provide rapid health status assessments and quicker medical responses in the field. Standoff detection of chemical and biological threats, such as nerve agents and bioweapons, using long-range sensors is an integral component of force protection strategies as weapons proliferate. Urban environment operations also require monitoring for toxic industrial chemicals and materials to ensure the health and safety of service members. BOIChemical sensors capable of tackling these challenges must detect specific molecules within complex mixtures at high confidence; they must also be relatively inexpensive, easy to operate, and sufficiently robust for operational field use. To this end, optical-based sensing using optical integrated circuits (OICs) has emerged as a promising technology.

OICs are devices analogous to electronic integrated circuits but use light rather than electrons for information signaling and transfer. While electronic integrated circuits are usually constructed as arrays of transistors, OICs employ a range of components (e.g., waveguides) to focus, split, isolate, polarize, couple, modulate, and, ultimately, detect light. Technical advancements in materials fabrication—realized especially over the past two decades—have enabled embedding these numerous functions in a single small-footprint OIC device. OICs are increasingly applied in telecommunications and sensing platforms.

While various materials are under exploration for use in OICs sensors silicon has received considerable attention. The high-refractive index contrast of silicon-on-insulator (SOI) waveguides provides a platform for the design and integration of high-density photonic circuitry. Use of silicon also leverages the existing commercial complementary metal-oxide-semiconductor (CMOS) infrastructure supporting the computing industry. Milestone advancements in this domain have occurred—Intel© announced the first fully integrated wavelength multiplexed silicon-based photonics chip in 2015, a disruptive advancement to present day information technology architectures. Silicon-based OIC development is also driven by the telecommunications industry due to its suitability for transmitting light in the near-infrared (NIR) region of the electromagnetic spectrum (a range relevant to the industry). Nevertheless, silicon is not appropriate for all applications and OICs that operate in the mid-infrared (IR) and other ranges may require different waveguide materials. Another key challenge with silicon is the realization of truly monolithic and portable lab-on-a-chip (LOC) systems, which include integrated optical sources and detectors on-chip.

Innovations of biochemical OIC-based sensors are taking place in the ealthcare/biomedical industry. OIC-based biochemical sensors essentially perform three general steps to: (1) prepare light, (2) direct light through a sample of interest, and (3) measure changes of the transmitted light due to the presence

of analytes in the sample. This simplified workflow provides several advantages over other biochemical sensing modalities, including immunity to electromagnetic interference, smaller footprints, label-free sensing, opportunities for multiplexing, and easier integration into LOC-type systems for rapid sample-to-answer capabilities.

While significant progress has been demonstrated, the transition of healthcare OIC sensors to field applications is nontrivial. Devices must operate with a low limit of detection (LOD) and high sensor selectivity in a potentially complex environment without access to laboratory-based sample preparations. OICs offer the potential for a direct sample-to-answer sensing workflow for a nontechnical user, eliminating the demand for training and the potential for analysis errors.

8.12. Fundamentals of OIC-based optical sensing

All IO-based biochemical sensors are optical transducers that encode the presence of an analyte onto the properties of light (i.e., intensity, phase, or frequency). As opposed to traditional free-space optics with long atmospheric path-lengths, IO-based sensors use light that is confined within, and propagates through, a semiconductor or dielectric waveguide.

As the light is reflected from the boundary back into the waveguide, part of the electromagnetic field leaks into and penetrates the sample medium where it can interact with analytes—this interaction is the fundamental mechanism of IO-based biochemical l sensing. The penetrating electromagnetic field, termed the evanescent field, exponentially decays with distance, dependent on material properties and geometry normal to the surface. This effectively confines the detection of analytes to a distance less than a wavelength from the surface (the field becomes too weak to sample the bulk solution farther from the surface). The basic detection principle based on evanescent fields is reviewed in Fig. 8.13.
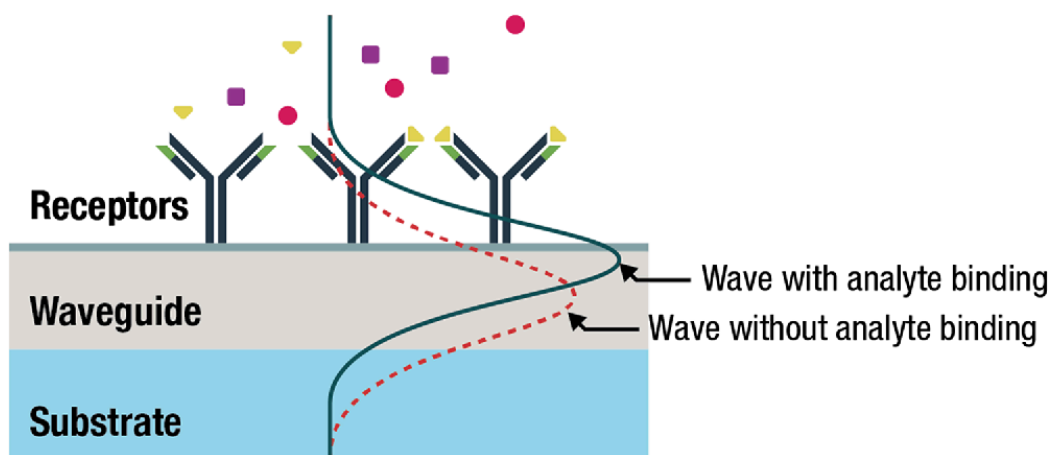
Fig. 8.13. Evanescent field detection principle

Light traveling through the waveguide produces an evanescent electromagnetic field in the lower refractive index environment above the waveguide. The presence of a target analyte (yellow triangles) bound to receptors (black) in the sampling area modify the properties of the evanescent wave and can be measured. The substrate has structural properties to physically support the waveguide and optical properties to both support the guided mode within the waveguide and the evanescent field above the waveguide.

To better elucidate the physics of IO-based sensing, consider a simple example: the detection of an optically absorbing analyte using an array of chemically specific receptors affixed to the waveguide surface. The presence of the bound analyte results in an attenuation of the total transmitted optical field, shown in Fig. 8.14, *a*, and, because a chemically specific receptor is used, the optical absorption of the molecular analyte can be known *a priori*. The total absorption for an analyte maximally proved by the evanescent field, described by the Beer–Lambert Law, can then be used to calculate the concentration (C) of the analyte,

$$\ln \frac{I(\lambda)}{I_0(\lambda)} = a(\lambda)CL,$$

190

where I_0 is the input intensity, I is the transmitted intensity, $a(\lambda)$ is the wavelength-dependent attenuation coefficient, and L is the length of the sensing region. Evident here is that sensor performance is determined not only by the analyte and its interaction strength with its receptor, but also by the physical design of the sensor, which determines the total analyte-light interaction (i.e., the length of the devices, the strength of the evanescent field, and its overlap with the optical cross section of the analyte).
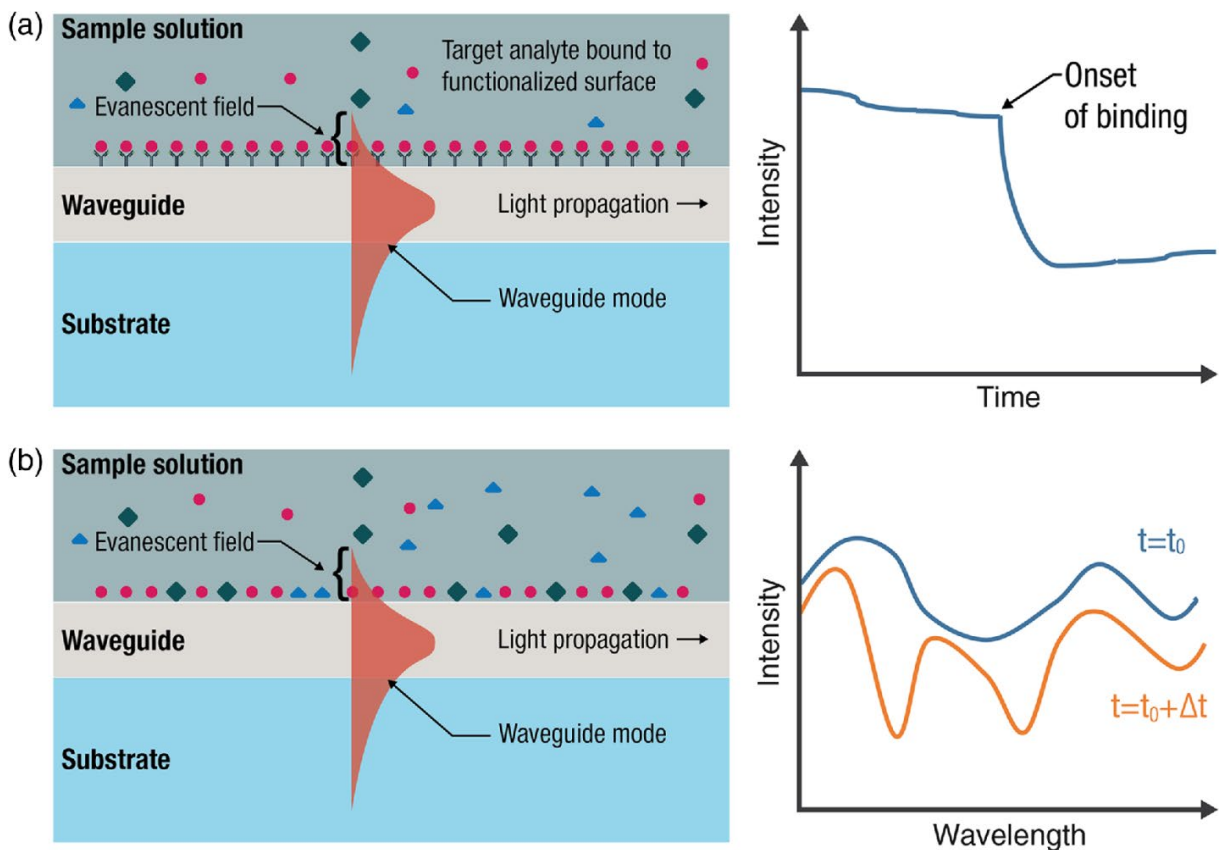

Fig. 8.14. Analyte detection using single-identification and spectroscopic sensors and the output spectrum. *a* – single-ID sensor waveguides are functionalized with a bOIChemical receptor spectroscopic sensors; *b* – no receptor is used

Single-ID sensor waveguides are functionalized with a bOIChemical receptor, which selectively binds to a specific analyte. These sensors operate at a single wavelength that is chosen based on the absorption of the analyte. When the analyte successfully binds to the receptor, the evanescent field interacts with

the analyte, and accordingly a reduction in the signal transmitted through the waveguide is detected.

For spectroscopic sensors, no receptor is used and a broadband optical source is used to probe the analytes. The resulting evanescent field probes all analytes near the surface of the waveguide. The output spectrum may have multiple features, which need to be deconvolved with known analytes in the environment in order to ascertain their concentrations. These spectra may change rapidly based on the complexity of the environment.

Sensor performance can generally be quantified in terms of specificity and sensitivity. In the example above, the use of a chemically specific antibody leads to a fundamentally high specificity and, correspondingly, allows for highly sensitive detectors to be engineered. This type of sensor is termed single-identification (single-ID) as shown in Fig. 8.14, *a*. Single-ID sensors have the additional advantage of operating at a single wavelength, thereby relaxing the bandwidth requirements on optical sources. When no receptor is available, or the analyte is not known *a priori*, broadband optical spectroscopy must be used; then through postprocessing, it is possible to mathematically deconvolve the measured spectrum with the spectrum of known analytes to determine the identity and concentration of those present in the sample [shown in Fig. 8.14, *b*].

It is challenging for spectroscopic methods, as just discussed, to achieve the same high specificity and sensitivity of single-ID methods, especially when detecting analytes in a complex mixture with interferents. Furthermore, the stability and bandwidth of the optical source become important factors to collect high-sensitivity and high-resolution spectra. It is also important to note that comparing sensitives across reports can be challenging: sometimes units such as mass per volume are given, which is difficult to extend to other analytes, and even if a normalized unit such as refractive index units (RIU) is given, it is not normalized to the physical size of the sensor.

8.13. Single-ID PIC-based BOIChemical sensors

As discussed above, single-ID sensors use functionalized waveguide surfaces to capture and detect target analytes. This type of detection is performed at a single wavelength because the optical properties of the analyte are known *a priori* and changes in the transmitted optical signal can be directly attributed to that analyte. Three IO-based sensor designs using evanescent field detection have been the most prevalent over the past two decades: (i) integrated interferometers, (ii) ring resonators, and (iii) photonic crystal (PhC)-based sensors.

8.13.1. Integrated interferometers

Interferometers are historically established as highly sensitive sensors based on the phase-sensitive detection of minute optical path differences; a striking recent example is the detection of gravitational waves. Interferometric IO-based sensors, like their traditional free-space counterparts, split light from one waveguide into two separate arms, a reference arm and a sensing arm, functionalized for analyte capture. If a target analyte is present in a sample, it will bind to receptors on the sensing arm, and, as a result, light passing through the sensing arm will be phase shifted relative to the reference light. The interference pattern formed when light from both arms is recombined can then be measured to determine the relative optical phase delay and calculate the amount of analyte present. Two standard configurations exist for interferometers integrated on a chip: the Mach–Zehnder interferometer (MZI) Fig. 8.15, *a* and the Young interferometer (YI) Fig. 8.15, *b*, which is similar but out-couples the light in both arms allowing the light to interfere in free space, instead of using a junction on-chip.
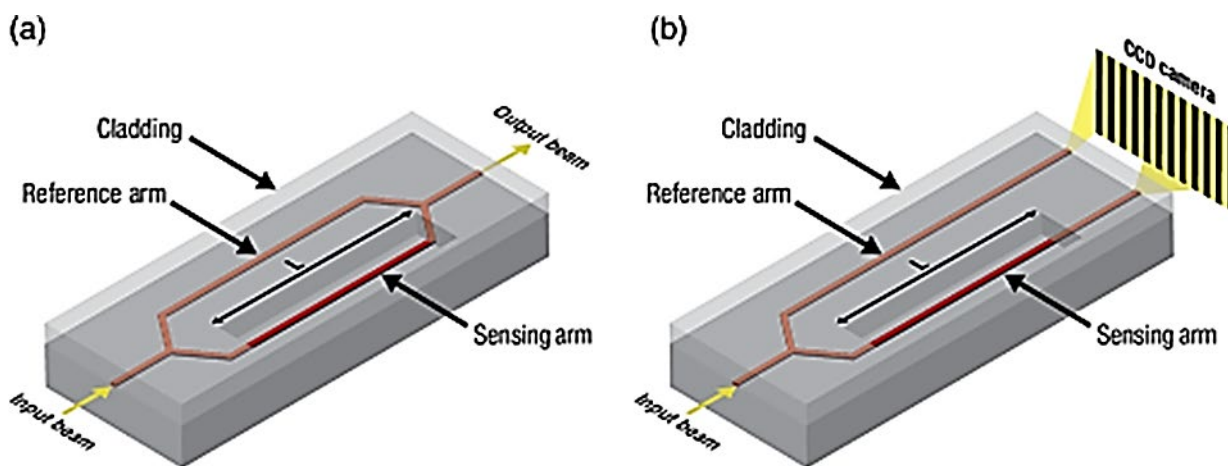
Fig. 8.15. Interferometric PIC sensors. A standard (*a*) MZI and (*b*) YI configurations

Light enters the interferometers on the left and is split into the sensing and reference arms. In the MZI case, the arms are recombined in a Y-junction and the interference measured in the output is used to quantify the presence of an analyte in the sensing arm. In the YI configuration, the light is not coupled, but is out coupled from sensing and reference arms and interfered in free space to create interference pattern. The presence of an analyte and its concentration is identified based on the interference pattern, reproduced.

Various waveguide designs, such as slot and bimodal waveguides (BiMW), have also been explored to improve sensor accuracy. In the slot design, the sensing arm is split into two waveguides separated by a nanometer-scale, low-refractive index "slot" region for light confinement. This design greatly enhances the light–matter interaction yielding higher sensitivity to the refractive index of the analyte.

In the BiMW configuration, a step junction in a single-mode waveguide excites two transverse modes, in a single waveguide, which interact with the sample before being measured by a photodetector. The sensitivity level of the BiMW sensor is comparable to other integrated interferometers.

To the best of our knowledge, the most sensitive bOIChemical interferometric sensor is a YI device with two Ta_2O_5 slab waveguides. However, the YI configuration requires off-chip detection and analysis, which may impact operationally relevant considerations for some applications.

8.13.2. Ring-resonator sensors

Ring-resonator transducers are becoming more common for bOIChemical sensing applications due to their high sensitivity and potential for multiplexing; the detection of multiple analytes is vital for testing complex fluids (e.g., blood, saliva, and urine) in diagnostics, monitoring, and toxicity screening. IO ring-resonator sensors use at least one linear waveguide to couple light to a closed-loop waveguide (the ring resonator) and, through evanescent-wave coupling, excite resonant modes in the loop waveguide. On resonance, constructive interference is generated in the multiple roundtrips over the ring circumference; however, off resonance the transmission rapidly decreases. This resonance effect considerably enhances sensor sensitivity. Analytes captured on the functionalized surface of the ring resonator shift the resonant wavelength only slightly, but this change is detected as a relatively large decrease in output intensity at a specific wavelength (Fig. 8.16, *a*).
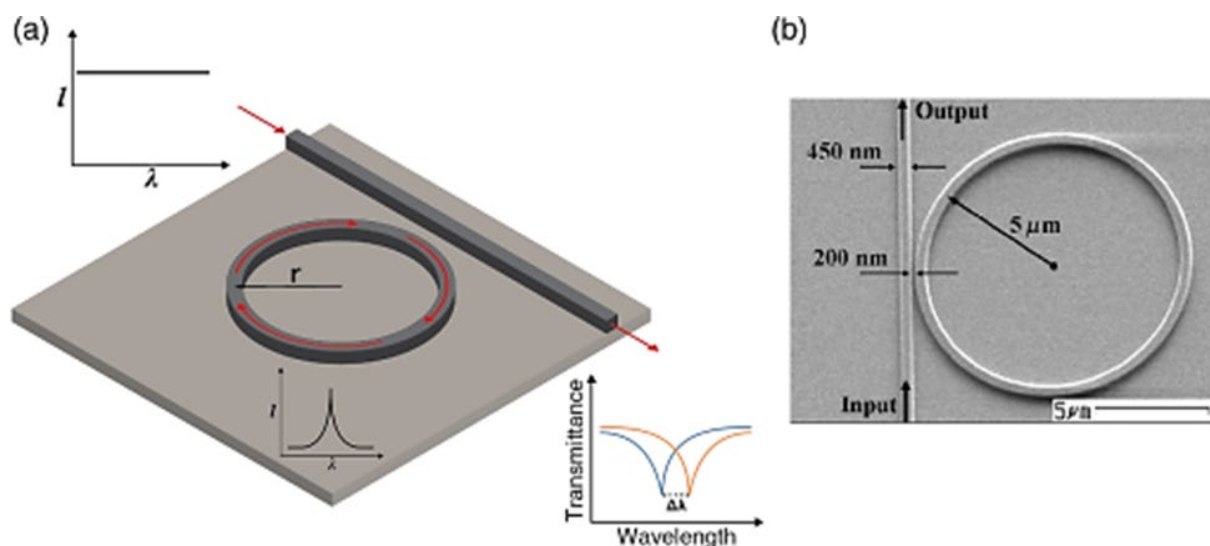

Fig. 8.16. Ring-resonator sensor:

a – schematic of a ring-resonator sensor; *b* – micrograph of a fabricated ring resonator

The linear waveguide provides input light to the loop waveguide at the coupling zone and captures output light from the loop for measurement. The

presence of analyte will shift the resonance wavelength within the loop, which is detected in the output spectra.

The number of revolutions light takes around the ring resonator yields an optical path length (OPL) orders of magnitude larger than the sensor's physical footprint, which is generally described by the resonator's quality (QQ) factor. High QQ-factors—considered to be at least 10^6—indicate low optical loss and correspondingly long photon lifetimes, which is translated into narrow linewidths and often high peak resolution (i.e., high sensitivity). QQ-factors in the 10^6 range have been reported for resonators around 50 to 200 μm in circumference, which has an OPL equivalent to a linear waveguide measuring several centimeters. In practice, a QQ-factor of 10^8 effectively means a molecule will be sampled more than 100,000 times, given by $L_{eff} = Q\lambda/2\pi n$, where L_{eff} is the effective free-space path length, QQ is the QQ-factor, λ; is the free-space wavelength, and n is the refractive index of the resonator.

Other important characteristics of resonators are the free spectral range (FSR) and the finesse. The FSR is the inverse of the round-trip time of an optical pulse around the resonator and defines the optical frequency range over which the resonator can be utilized. The finesse, which is determined by the resonator losses and in independent of the resonator circumference, is defined as the FSR divided by the full-width at half-maximum bandwidth of the resonance. Hence, L_{eff} can also be defined as the finesse times the ring circumference, which is a useful metric as the sensitivity generally scales with finesse for most sensing applications.

Therefore, despite the small physical size of the ring resonator, it can achieve higher sensitivities than straight waveguides while using orders of magnitude less surface area. Smaller waveguide sizes also facilitate the development of more complex, integrated devices. For example, an arrayed device using 32 ring-resonator sensors to detect streptavidin has been fabricated. Furthermore, pushing the technology toward detection of analytes in complex environments simultaneous detection of two DNA oligonucleotides was

demonstrated. Isolating 8 of the resonators to use as control sensors, and separated the remaining 24 sensors into two groups and functionalized each for a specific oligonucleotide and successfully multiplexed sensing was achieved. More recently, there have been demonstrations of functionalizing ring resonators with sorbent polymers to efficiently bind vapor-phase compounds, such as acetone and DMMP.

8.13.3. Photonic crystal-based sensors

PhC are structured matrices of materials that possess different dielectric constants, resulting in photonic bandgaps that determine the frequencies of light, which are reflected or transmitted by or through the substrate. The simplest type of PhC structure is a perfect array of holes, which was in fact one of the earliest geometries for PhC bOIChemical sensing; the PhC was designed to reflect a single wavelength of light and, upon analyte binding to immobilized receptors on the surface a change in the reflected light wavelength was measured with a reported detection limit.

The introduction of defects into the PhC structure (i.e., shifting or removing holes, or changing hole diameters) results in the confinement of light to the defect with high QQ-factors, enabling highly sensitive detection of extremely small sample volumes (e.g., a few μm^3). The prospect of high QQ-factors and small sample volumes has led to significant effort to explore the potential of PhCs for bOIChemical sensing, beyond the mirror geometry. Increasingly complex structures, such as that shown in Fig. 8.17, where the use defects (i.e., the absence of holes) introduces a waveguide, have been demonstrated to confine sensing volumes and increase sensitivity.

The top and bottom milled areas form a PhC with a photonic bandgap. Light is confined to defects which, here, is the region absent of holes, which forms the central waveguide. Introducing analytes near the PhC changes the RI and the cutoff frequencies of the bandgap.
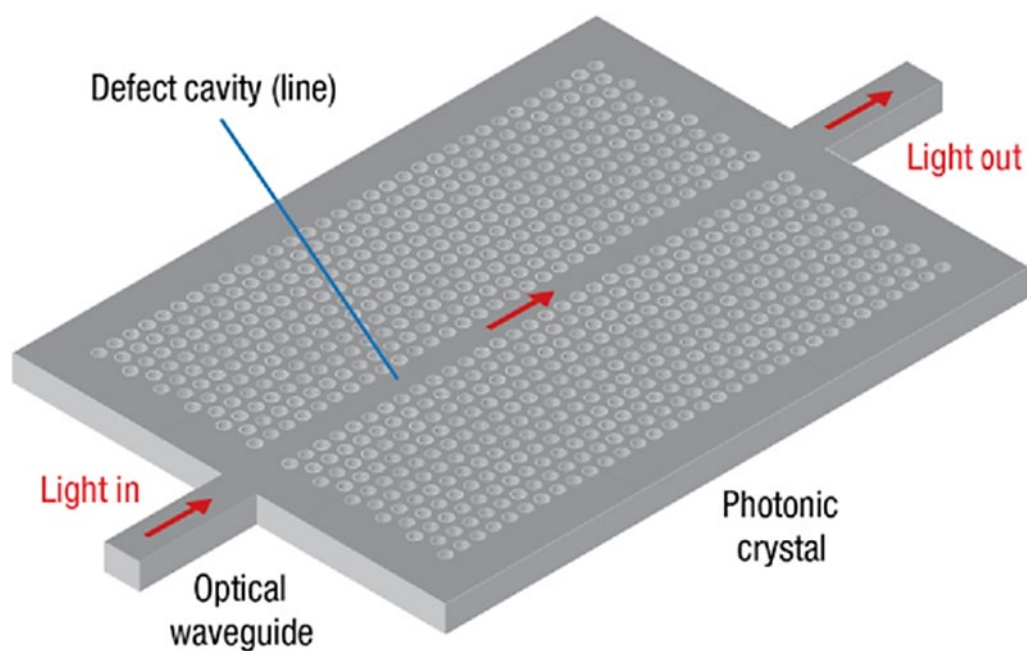
Fig. 8.17. A 2-D PhC waveguide

For example, slot waveguides have been integrated in PhCs to boost sensor performance. Similar PhC devices have been used to detect lung cancer cells down to 2 cells/μL, to measure biotin–streptavidin binding (a common analyte/receptor pairing to demonstrate biochemical assay viability), and to probe the binding kinetics of immunoglobulin G proteins. More recently, PhC biosensor devices have been developed for detection of HIV, biomarkers for iron deficiency anemia, and glucose monitoring.

8.14. Spectroscopic sensors

Spectroscopic IO-based sensors, in contrast to single-ID sensors, do not use a chemically specific receptor to capture a known analyte; instead, all molecules near the waveguide surface are measured and their respective properties are encoded into a broadband optical field. The resulting optical spectrum can then be postprocessed to mathematically deconvolve the spectral signatures of known molecules. While it is challenging for spectroscopic methods to achieve the same high specificity and sensitivity of single-ID methods, this type of sensing is well-suited for certain applications such as

environmental and air-quality monitoring, greenhouse-gas sensing, medical diagnostics, and standoff detection for chemical, biological, radiological, nuclear, and explosive (CBRNE) agents. Based on free-space spectroscopic detection geometries and leveraging silicon photonics technology, there have been demonstrations of IR spectroscopy [such as tunable diode laser absorption spectroscopy (TDLAS) or Raman spectroscopy] on integrated CMOS-compatible platforms for gas sensing. For example, using a dense spiral network of silicon waveguides to increase OPL on a SiO_2 platform sub-100 ppm methane gas sensing was achieved. The top of the spiral waveguide was exposed to the ambient atmosphere allowing the evanescent field of the optical mode to probe the rovibrational transitions of the surrounding environment directly. Optical absorption of the rovibrational transitions, in accordance with the Beer–Lambert law, provides an absorption spectrum that can be used for spectroscopic identification.

One challenge with spectroscopic detection is concentration analytes on the waveguide surface. Sorbent polymers where used designed to reversibly sorb organophosphates (OPs) and concentrate them by as much as 10^8, to functionalize IO sensor for trace gas Raman spectroscopy. By designing the waveguide mode to propagate within the sorbent polymer, in contrast to the previously described evanescent interactions, the light–analyte interaction and the resulting Raman scattering is enhanced along the length of a high-index Si_3N_4 waveguide; this design achieved single ppb detection limits for identifying trace compounds, such as dimethyl sulfoxide. A hyperbranched carbosilane sorbent polymer on silicon nitride waveguides was used for detection of four vapor-phased chemical warfare agent simulants as low as 5 ppb.

8.15. Senor using surface plasmons

Surface plasmons are surface electromagnetic waves that propagate in a direction parallel to the metal/dielectric (or metal/vacuum) interface. Surface

plasmons offer useful properties such as field localization and enhancement, high surface sensitivity, and sub-wavelength confinement, which have led surface plasmons into applications in, e.g., nanophotonics, biosensing, and integrated optical circuits.

Figure 8.18 shows the basic principle of operation of the chemical sensing using surface plasmons. Light propagating inside the waveguide interacts with the charge density waves of a thin metallic (usually silver) layer. Provided that the guided wave's wavenumber matches that of the surface plasmon, the optical energy is absorbed by the plasmons and is consequently dissipated. The plasmon wavenumber, in turn, depends on the conductivity and the surface condition of the metallic layer. A thin sensitive polymer film is usually deposited over the metallic film tha upon absorption of chemical and gases, changes the plasmon wavenumber of the metallic film.
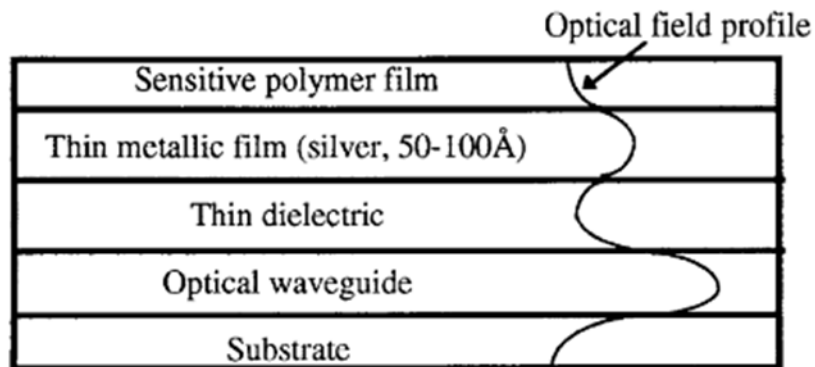
Optical field profile

| Sensitive polymer film |
| Thin metallic film (silver, 50-100Å) |
| Thin dielectric |
| Optical waveguide |
| Substrate |

Fig. 8.18. Schematic of a surface plasmon IO chemical sensor

Due to the large attenuation of the optical field that matches the plasmon wavenumber, this scheme cannot be used in interferometry. Using this method different gases including CH_2I_2 are detected.

The surface plasmons are so sensitive to the conditions of the polymer film that very large modulation-index light modulators with electro-optically active polymer films have been proposed and constructed.

9. PHOTONIC CRYSTALS

The explosive growth of the internet has led to unprecedented bandwidth requirements in the telecommunications network that can only be addressed with advanced optical technologies. One of the main limitations of these technologies is the size of devices. Photonic crystals can be used to overcome this limitation and enable a major breakthrough, because novel components based on photonic crystal waveguides will be orders of magnitude smaller, allowing a degree of integration that is hitherto unheard of. Making some the key demonstrations and developing miniaturised photonic components in conjunction with industry is the major driving force behind our work in this area.

When photon confining structures are multiply coupled on a single chip so as to produce a periodically varying index of refraction (dielectric constant) with a periodicity on the order of the wavelength of the photons, a photonic crystal is formed. An example of photonic crystal lattices is shown in Fig. 9.1. For light of 1 μm vacuum wavelength, both the lattice constant a and the radius of the perturbation r would typically be several hundreds of nanometers.
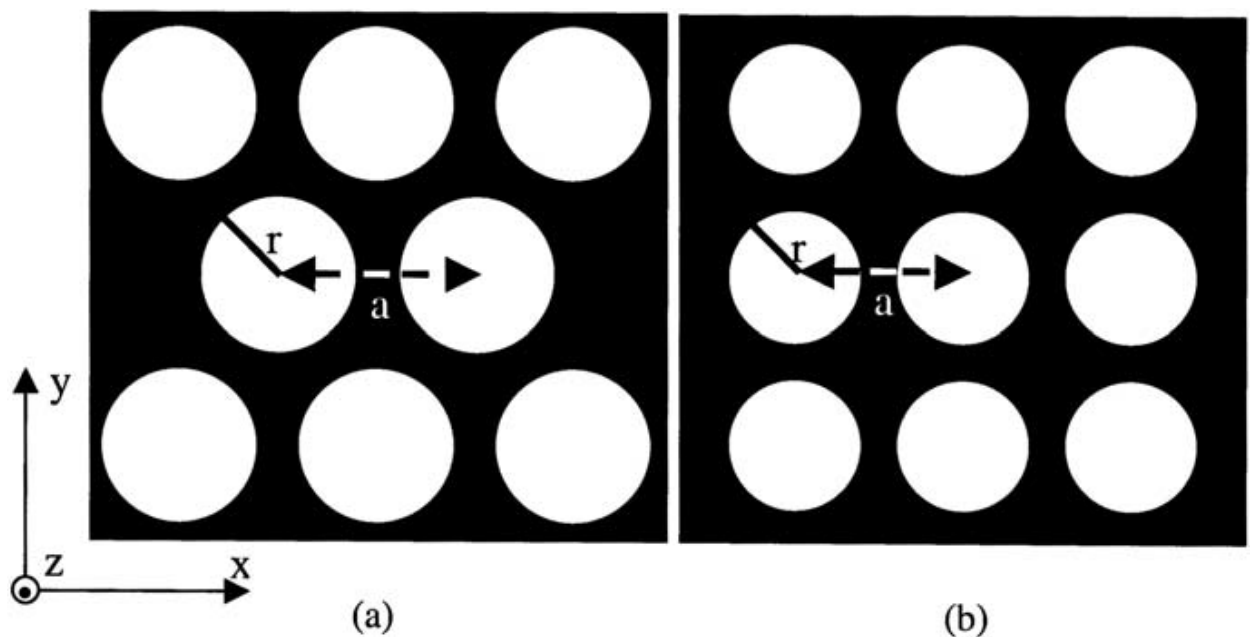

Fig. 9.1. Diagram of common two-dimensional photonic crystal structures:

a – a triangular lattice; *b* – a square lattice

Photonic crystals can be either one-, two- or three-dimensional (Fig. 9.2)
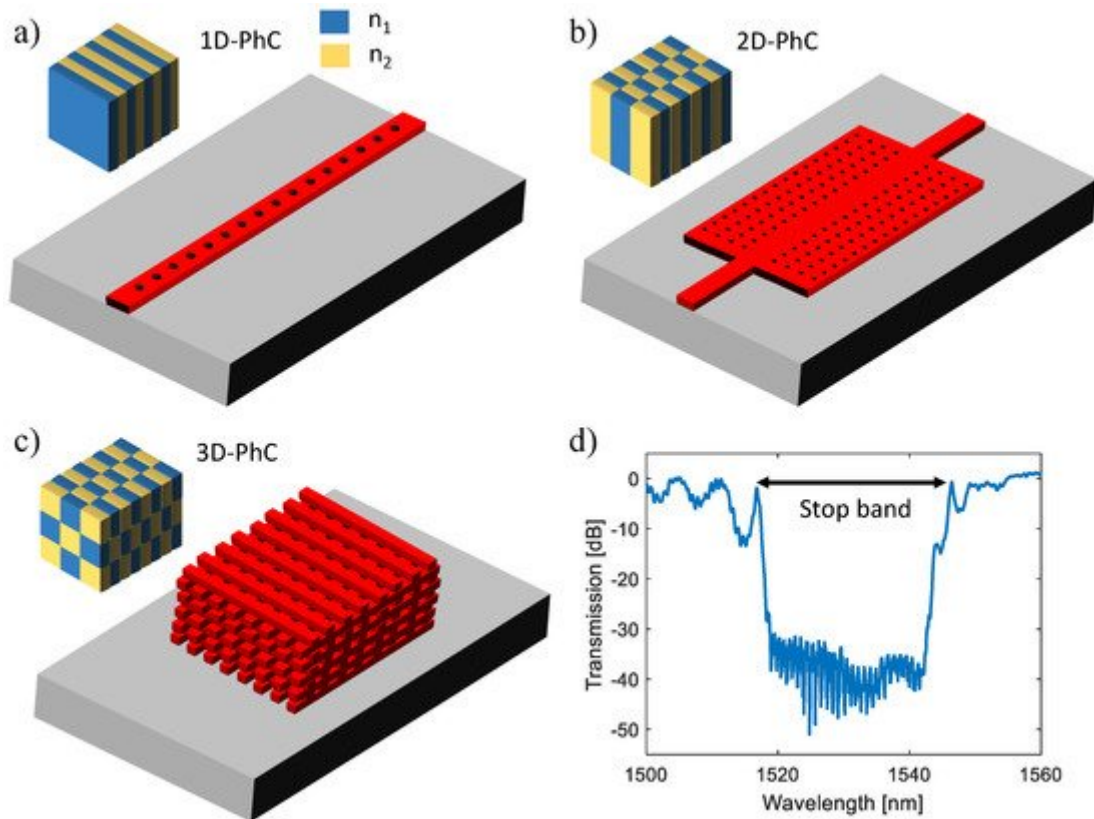

Fig. 9.2. Illustration of photonic crystals in: *a* – 1D conformation; *b* – 2D conformation; *c* – 3D conformation. Insert: Schematic representation of each format showing the periodic arrangements, different colors represent materials with different indices. *d* – measured transmission spectrum of a uniform photonic crystal (PhC) device after normalization

Technically speaking, examples of one-dimensional photonic crystals are a diffraction grating like that shown in Fig. 9.3.

However, since grating structures like those have been used for many years before the advent of modern photonic crystals, it is not common to call them photonic crystals.

Two-dimensional photonic crystals are usually fabricated by producing a lattice of air holes (lower index of refraction) in a layer of higher-index material or by producing a lattice of higher-index rods in a layer of lower-index material (such as air). Lattices of these two types are shown in Fig. 9.4.
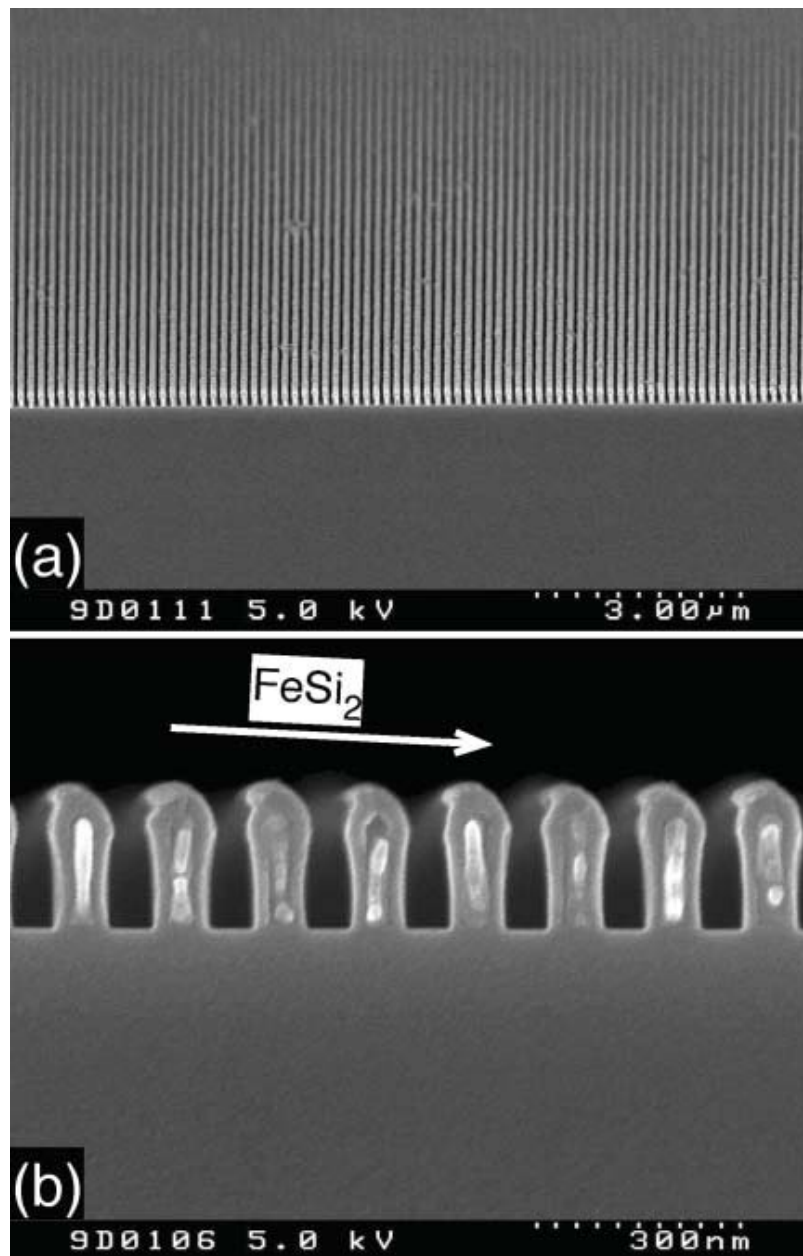
Fig. 9.3. One-dimensional photonic crystal:

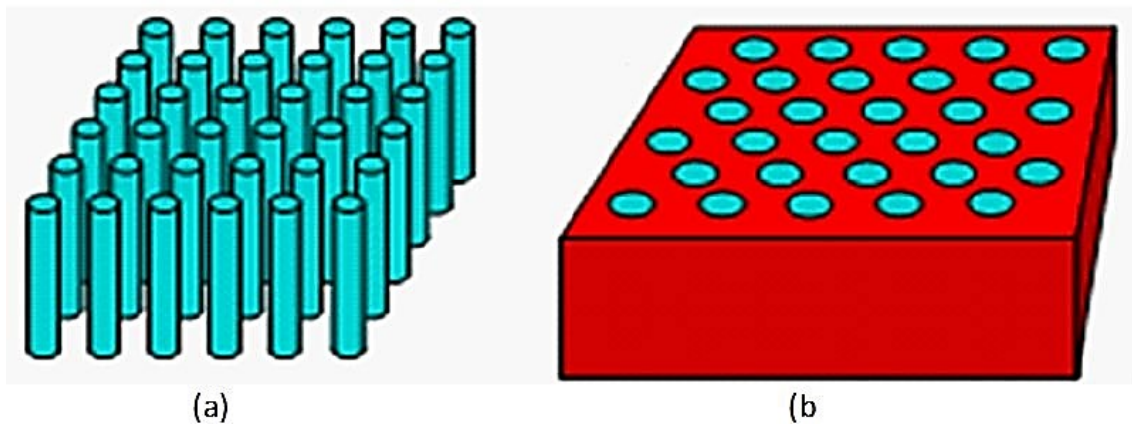*a* – diffraction grating; *b* – multilayer dielectric mirror

Fig. 9.4. Two-dimensional photonic crystals:

a – dielectric rods in air; *b* – holes in higher-index material

These lattice structures require a very large aspect ratio. (The aspect ratio of a feature on an integrated circuit chip is defined as its height divided by its width.) However, such high-aspectratio structures can be fabricated fairly routinely by using the recent advances in nanolithography and e-beam and ion-beam sputter etching.

The name, photonic crystal, derives from the fact that the behavior of photons in a photonic crystal is very similar to that of electrons in a crystalline semiconductor. In moving through a semiconductor crystal, an electron is affected by the periodic potential produced by its Coulomb interaction with the atoms of the semiconductor.

Solution of the Schrodinger equation for the wavefunction of the electron and its allowed energy states yields the familiar dispersion diagrams of allowed electron energy states vs. wave vector.

Key observable features are the valance and conduction bands and the bandgap E_g, within which there are no allowed electron energy states. When Maxwell's equations are solved for the case of photons in a photonic crystal, very similar dispersion diagrams of energy (frequency) vs. wave vector result, including the feature of a "bandgap" of frequencies within which no allowed states exist, i.e., no propagation can occur. The solution of Maxwell's equations for photons in a periodic dielectric lattice is a fairly complicated process. There are no simple closed-form equations that can be easily derived. However, accurate solutions can be obtained using computer-aided methodologies, such as the plane wave method (PWM) and finite-difference time-domain (FDTD) method. These mathematical methods can be adapted to include the vector nature of electromagnetic fields as they propagate within periodic lattice structures. The PWM works well for structures that have uniform periodicity and can be used for some structures that contain defects, as long as the defects can be incorporated into a periodic "super cell". However, structures containing non-periodic sets of defects, such as the waveguides and splitters are best handled by the FDTD method, which does not require periodicity. A typical

dispersion diagram calculated for a photons propagating in a 2D honeycomb GaAs rod photonic crystal is shown in Figure 9.5.
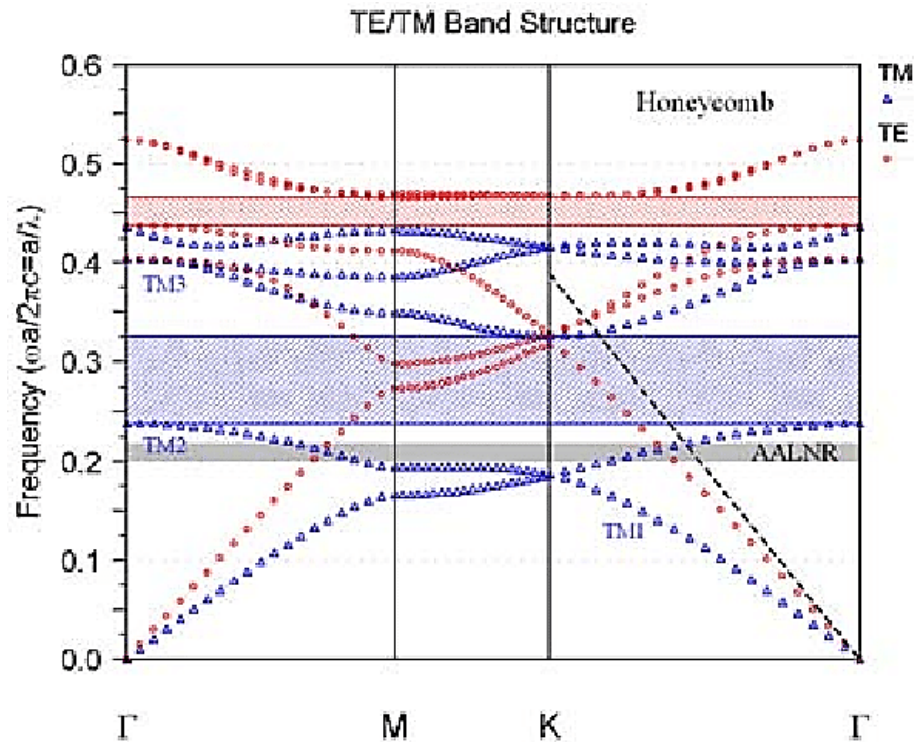

Fig. 9.5. Band structure of a 2D honeycomb GaAs rod photonic crystal for *r/a* > 0.24. The dashed line corresponds to the light line and the gray strip represents the region with AALNR in the TM2 band

The corner points of the irreducible Brillouin zone in k-space **Γ** (=0), **X** (=π/a**x**) and **M** (=π/**x**+**y**) are marked in the diagram. The blue band mark the bandgap for photons, within which there are no allowed states, i.e., propagation of photons of those frequencies is blocked. Figure 9.5 is a plot of allowed propagation energies frequencies) of photons in a PhC vs. their wave vectors, just as of the allowed energies of electrons in a semiconductor crystal vs. their wave vectors.

The blocking of the propagation of photons of certain frequencies in a photonic crystal can be easily understood without going through the full calculation of the dispersion diagram by considering it to be a phenomenon of Bragg scattering (reflection) similar to the reflection of photons in a DBR laser, as discussed in Chapter 7.

Consider a plane wave traveling in two-dimensional rectangular lattice as shown in Fig. 9.6.
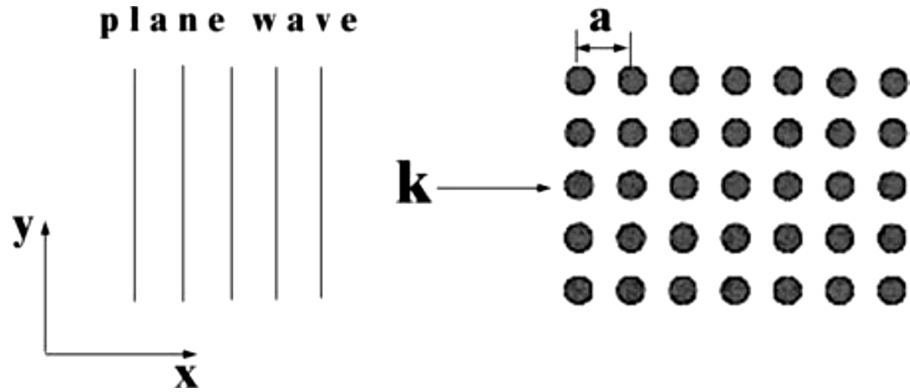

Fig. 9.6. Diagram of plane wave traveling in a rectangular dielectric lattice

The **E** and **H** fields of the wave are given by

$$\mathbf{E}, \mathbf{H} \sim \exp\left[i\left(\mathbf{kx} - \omega t\right)\right],$$

where

$$\left|\mathbf{k}\right| = \frac{\omega}{v} = \frac{2\pi}{\lambda}.$$

The basic condition for Bragg scattering through an angle $2\theta_B$ is given by $\sin\theta_B = \lambda/2\Lambda$, where θ_B is the Bragg angle and Λ is the periodicity. In this case $\Lambda = a$; thus waves of wavelength equal to $2a$ will produce $\sin\theta_B = 1$ and will be scattered through an angle $2\theta_B = 180°$. They will be reflected and will not propagate.

In a real PhC, small perturbations will spread the condition $\lambda = 2a$ over a narrow range of wavelengths approximating $2a$ to yield a bandgap. The transmission spectrum of a silicon PhC is shown in Fig. 9.7. It consists of a triangular lattice of air holes in a silicon layer with $a = 437$ nm and $r = 175$ nm. The resulting bandgap, with edges defined by transmission reduced by 20 dB, ranges from 1118 to 1881 nm. There is only 1% transmission (or less) permitted within this range of wavelengths.
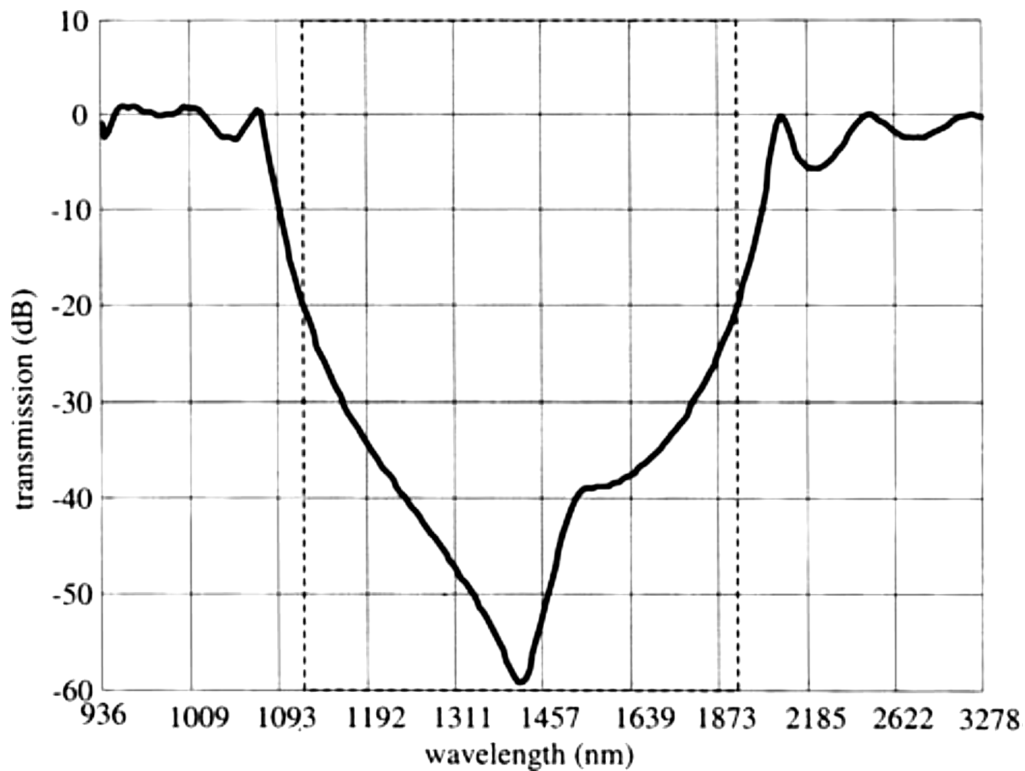
Fig. 9.7. Transmission spectrum of a silicon photonic crystal

For a two-dimensional PhC like this one, a bandgap may or may not exist. Sometimes the ω vs. *k* curves overlap so that there is no bandgap; it depends on the particular values of spacings and dielectric properties. A large *r/a* ratio and large difference in permittivities (dielectric constants) are required. (A one-dimensional PhC like a grating or a multilayer dielectric mirror always has a bandgap.) When the air holes of the PhC of Fig. 9.7 were filled with a material with dielectric constant >1, the width of the bandgap was reduced, as shown in Fig. 9.8.

Since the index of refraction of silicon is about 3.5 for the wavelengths used in PICs, the data of Fig. 9.8 demonstrate that a large index difference is required in order to produce a bandgap. When defects are controllably introduced into PhCs, disrupting the regularly ordered lattice, new propagation states are created that can be used to create devices such as waveguides, couplers, switches and filters.
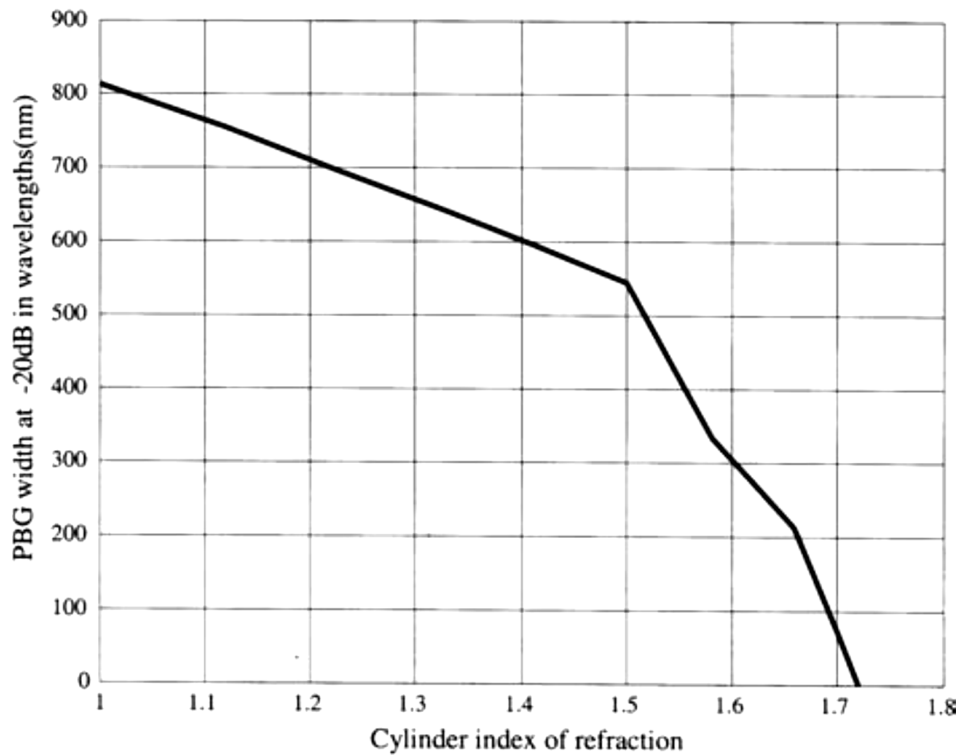
Fig. 9.8. Reduction of bandgap as a function of cylinder index of refraction

9.1. Fabrication technologies

There is a great variety of methods that have been demonstrated for producing either structures with features having dimensions on the order of nm or particles of nm size. For example, plasma arcing, sol-gels, electro-deposition, chemical vapor deposition, ball milling, self-assembly and natural nanoparticles have all been used. However, these have not found much application in the field of photonics. Instead, the traditional methods of semiconductor device and PIC fabrication, such as molecular beam epitaxy (MBE), metalorganic vapor phase epitaxy (MOVPE), e-beam lithography, ion-beam etching and reactive-ion etching(RIE), have been upgraded and adapted to the fabrication of nanophotonics structures and PhCs.

Certain additions and modifications must be made to the basic MBE system to qualify it for nanolayer fabrication. The remotely operated shutters covering the heated effusioncell sources must be capable of switching open and closed in 0.1 s, or less, in order to produce layers of thickness <100 nm. The

substrate must be rotated during growth to insure uniformity of the grown layer. The temperature of the substrate must be controlled to 0.1° accuracy to stabilize the growth rate. Also, the system should be fitted with analysis tools such as a mass-spectrometric flux monitor to measure layer composition and reflection high-energy electron diffraction (RHEED) to measure layer crystallinity. These measurements must be made while growth is going on in order to control the quality of the layer.

The layers grown by MBE can be patterned laterally to produce two-dimensional devices such as quantum wires, quantum dots or planar PhCs. This can be done by either growing on pre-patterned substrates or patterning the layer after growth by lithographic masking, exposure and etching. MBE growth can be used for III–V and II–VI semiconductors as well as for group IV semiconductors. Thin layers of both ternary and quarternary compounds, such as $Ga_xAl_{1-x}As$ and $In_{1-x}Ga_xAs_{1-y}P_y$, can be grown by MBE.

In MOVPE the constituent atoms are delivered to the substrate as a gaseous flow within a growth reactor furnace. A basic horizontal-flow MOVPE reactor is shown in Fig. 9.9.
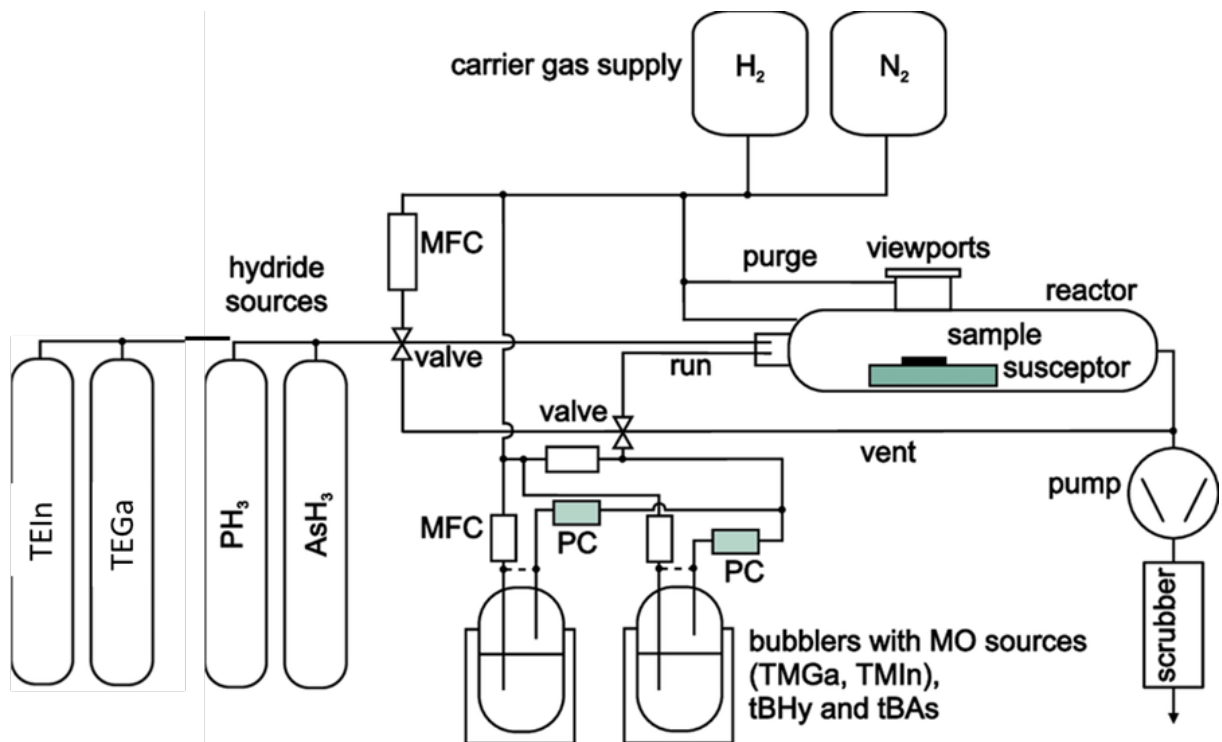

Fig. 9.9. A schematic diagram of a basic horizontal-flow MOVPE reactor

The MOVPE growth technique allows a precise control of layer thickness (up to fractions of a monolayer), chemical composition and doping by means of a sophisticated arrangement of mass flow controllers (MFC) and pressure controllers (PC).

MOVPE is a form of metalorganic chemical vapor deposition (MOCVD) in which the growth is epitaxial. The metalorganic gasses such as triethyl indium (TEIn) and triethyl gallium (TEGa) are usually transported along with H_2 as a carrier gas. A crystalline substrate must be provided to act as the seed for epitaxial growth. MOVPE is used for the growth of thin layers of III–V and II–VI semiconductors in order to maintain a high enough vapor pressure of the constituent elements to prevent dissociation of the layer during growth. It is not used for silicon vapor phase epitaxy because it is not necessary. Silicon is an elemental semiconductor which can be grown by conventional VPE using silicon tetrachloride, or silane, dichlorosilane and trichlorosilane as source gases.

For the growth of nanolayers by MOVPE special efforts must be made to provide the uniformity and control of growth rate that is required. Temperature within the reactor must be stabilized to within $0.1°$ and the remotely controlled valves must be switchable within 0.1 s. A premixing chamber must be added at the reactor input to insure uniformity of the layer.

Despite the apparent simplicity of the basic system practical MOVPE systems are quite complex and are usually purchased commercially rather than being assembled by the researcher in the laboratory. The major reason for this complexity is that the source gases are extremely dangerous and must be handled with strict adherence to established standards. In addition to a number of horizontal-flow MOVPE reactors that are marketed commercially, a closed coupled showerhead reactor is available. The showerhead consists of an array of uniformly distributed injection holes, spaced 1−2 cm away from the substrate surface. This type of reactor has been shown to be less sensitive to growth conditions like reactor temperature, pressure and rotation In regard to

semiconductor device and PIC fabrication, lithography is the technology that is used to transfer the lateral dimensions of structures to the active layer.

The basic approach used by the semiconductor industry is to coat the layer with a "resist" that is sensitive to either photons of light or electrons such that it becomes either removable in a developer solution when exposed (positive resist) or nonremovable (negative resist). The resist is exposed through a mask that defines lateral structures by blocking some areas and passing light through to others. The resist is "developed" in a solution that dissolves removable areas. The developed resist is then used as a mask against either liquid or sputter etching. The size of the features that can be created depends strongly on the wavelength of the exposure source. As one might expect, shorter wavelengths yield smaller feature sizes. The diffraction limited resolution of a circular lens is given by the empirically determined Rayleigh criterion, which states that

$$\sin\theta = 1.220\frac{\lambda}{D}, \tag{9.1}$$

where θ is the angular resolution, λ is the vacuum wavelength and D is the diameter of the lens. This leads to an expression for the spatial resolution Δl of an ideal lens, given by

$$\Delta l = 1.220 f\frac{\lambda}{D}, \tag{9.2}$$

where Δl is the size of smallest object that the lens can resolve and f is the focal length of the lens. Equations (9.1) and (9.2) assume an ideal lens. In a practical lithographic projection exposure system, imperfections in the optical train and the resist further limit the resolution. This results in the following adjusted form of the Rayleigh criterion:

$$W_{\min} \approx \mathbf{k} f\frac{\lambda}{D} \approx \mathbf{k}\frac{\lambda}{NA}, \tag{9.3}$$

where W_{min} is the minimum feature size that can be produced, *NA* is the numerical aperture of the optical system and **k** is a constant that characterizes the ability of the resist to distinguish between small changes in intensity. Typically **k** is approximately 0.75. Ultraviolet (UV) and deep ultraviolet (DUV) exposure systems are not suitable for fabricating nanoscale features. Extreme ultraviolet (EUV) systems are a possibility. However, because of high absorption associated with the very short wavelength no lenses are available. All focusing and directing must be done with mirrors and must be done in a vacuum to avoid absorption and scattering of the photons.

X-ray lithography can be used for nanostructure fabrication. In that case, special materials must be used in the exposure mask. Materials that are transparent to xrays, and hence can be used as mask substrates, are silicon, silicon carbide, silicon nitride, boron nitride and diamond. Materials that are strong absorbers of x-rays, and hence can be used to make blocking regions of the mask, are gold and tungsten. X-ray exposure systems are complicated by the fact that there are no known effective lenses or mirrors for x-rays. Thus the beam cannot be collimated and is radiating outward from a small point source to the larger mask. This introduces errors in the shape and position of the features that complicates mask design. A collimated x-ray beam that avoids these problems can be obtained from a synchrotron source, but they are very large and expensive.

The most widely used method for patterning the lateral dimensions of nanoscale photonic devices has been electron-beam (e-beam) lithography. Electron beams consist of charged particles for which the methods of collimating, focusing and directing have been well developed. There are also well-known electron-resist materials such as polymethyl methacrylate (PMMA) and a number of proprietary commercial resists. The use of an e-beam for exposure of nanoscale features is consistent with the use of short-wavelength photons. E-beam exposure is most often performed by direct writing using a focused beam with its lateral position relative to the target being controlled by a

preprogrammed computer. This approach avoids the need for a mask and eliminates problems associated with masks, such as distortion and wear.

Once the lateral dimensions of the nanostructure have been defined by lithography, it is necessary to controllably remove material in certain regions to shape the nanostructure. This might involve making a lattice of holes in a layer or shaping columns as in Fig. 9.4. Because of the high aspect ratios and close spacing desired in most cases, liquid etching is inadequate and special etching techniques must be used.

The basic method used is sputter etching, sometimes called dry etching. In sputter etching the layer to be etched is bombarded with energetic ions. Usually the ions of inert gases such as argon or krypton are used. Ion energy is typically 1–2 keV. The impinging ions transfer energy to the atoms near the surface of the layer causing them to be ejected. The rate of removal depends on the ion flux (number of ions/area·s) and the sputtering yield, which is defined as the number of atoms ejected from a target surface per incident ion. Sputtering yield generally increases with increasing ion energy up to about 2 keV, then decreases due to ion implantation. Heavier "bullet" ions generally increase sputtering yield for the same energy. Measured values of sputtering yield for a variety of the commonly used ions and target materials have been published. There are two major types of sputtering that are used, plasma and ion beam.

In plasma sputtering the ions are generated by establishing an electrical "glow" discharge between an anode and a cathode within a vacuum system containing one of the inert gases. A diagram of a plasma sputter etch system is shown in Fig. 9.10.

The pressure is typically $(2–20) \cdot 10^{-3}$ torr. The high voltage applied between the anode and the cathode can be either dc or rf with a dc component. A dc negative bias must be maintained on the cathode so that ions will be attracted to it. The impinging ions sputter away atoms from the layer in regions where it is not masked by the developed resist. The ions do cause some surface damage and also some ions may become implanted, particularly at higher sputtering

energies. However, if these effects prove detrimental to device performance they can be mitigated by annealing at about 200 °C for 10 or 20 min.
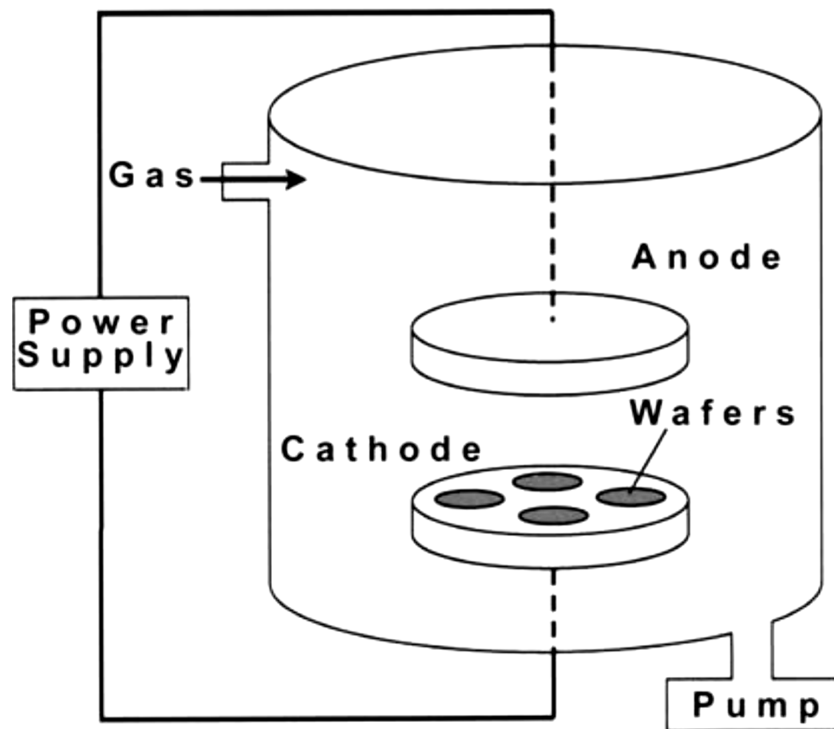

Fig. 9.10. Diagram of a plasma sputter etching system

An improvement can be made to the basic plasma sputter etching system by adding magnets so as to produce a magnetic field with its field lines aligned perpendicularly to the anode and cathode. This field causes the electrons in the plasma to travel a spiraling path, thus increasing the probability that they will encounter a gas atom and ionize it. (The path of the ions is not strongly affected by the magnetic field because of their much greater mass.) This type of sputtering is called magnetron sputtering. It results in the generation of a high-density plasma and increases the sputtering rate. It also allows reduction of the chamber pressure to the 10^{-5}–10^{-3} torr range, which reduces stray deposition of sputtered atoms within the system.

The second major type of sputter etching is ion-beam sputtering. In ion-beam sputtering the ions are generated in a localized ion source, extracted and then accelerated and focused to form a collimated beam. The inert gas ions are generally used just as in plasma sputtering.

One advantage of ion-beam etching as compared to plasma etching is that the beam is well collimated. Therefore there is less undercutting and higher aspect ratios can be achieved. Also, the ion-beam system can (and should) be operated in a hard vacuum of 10^{-7} or better. This reduces contamination. There is no incorporation of stray-contaminant atoms that might be present in the gas of a plasma sputtering system. The ions extracted from the source can be mass filtered for purity as in an ion implantation system. Another advantage of ion-beam sputtering is that the ion energy and beam current can be independently controlled by varying the accelerating voltage and extractor voltage, respectively. For these reasons ion-beam sputtering is generally preferred for fabrication nanostructures.

A modified plasma sputtering technique that is also capable of producing a high aspect ratio and vertical side walls is inductively coupled plasma (ICP) etching. ICP etching is a technique which uses rf energy coupled into a low-pressure gas such as SF_6 by an inductive coil mounted on the outside of a non-conducting, insulating window. There are two types of ICP geometries: planar and cylindrical. In planar geometry, the electrode is a coil of flat metal wound like a spiral. In cylindrical geometry, it is like a helical spring. When an rf current is passed through the coil, the resulting time-varying magnetic field induces breakdown of the gas within the reactor chamber and creates a high-density plasma. Two separate power supplies can be used in an ICP system for plasma sputter etching. This gives independent control of the ion energy and the ion-beam current. A low-frequency supply is used to generate a high-density plasma and a high-frequency supply is used to provide a wafer bias. This dual power supply arrangement allows the generation of a high density plasma for increased etch rate, along with a low wafer bias for low damage to the wafer surface. The wafer bias can be adjusted to control the decree of anisotropy of the etch process. In a conventional capacitively coupled sputter etch system with a single power supply, both the ion energy and the beam current are determined by the single applied voltage.

There is a wide selection of analysis tools available for the characterization of nanomaterials, nanostructures and nanodevices. These are the standard instruments that are found in many laboratories, so a detailed description of all of them is not necessary.

9.2. Photonic crystals devices

A fundamental element of any PIC is the waveguide. Thus, it is not surprising that the fabrication of waveguides in photonic crystals is a topic that was given early investigation. It has been demonstrated that introduction of a line defect into a two-dimensional, slab photonic crystal can be used to create a waveguide. A line defect is created by adding or removing dielectric material along a line in the lattice of the PhC. This is sometimes called "defect doping" by analogy with impurity doping of a semiconductor. In "acceptor" doping the effective index at a lattice site is reduced by replacing some higher-index material with a lower-index material (e.g., increasing the size of air holes in a slab). In "donor" doping the effective index at a lattice site is increased by replacing some lower-index material with a higher-index material (e.g., reducing the size of air holes in a slab). A defect of this type, formed by eliminating a line of holes in a slab PhC, is shown in Fig. 9.11. In this case, the slab is a silicon-on-insulator (SOI) layer.

Light is confined to the waveguide in the x–y plane for states that exist at certain frequencies within the photonic band gap. Light is confined in the z-direction, vertical to the plane, by total internal reflection (TIR) due to the changes in refractive index at the surfaces of the slab. (Acceptor doping is not used to confine light in a line-defect waveguide of this type because it would be confined to the regions of lower index, i.e., the air, and confinement to the slab in the vertical direction would be lost.)
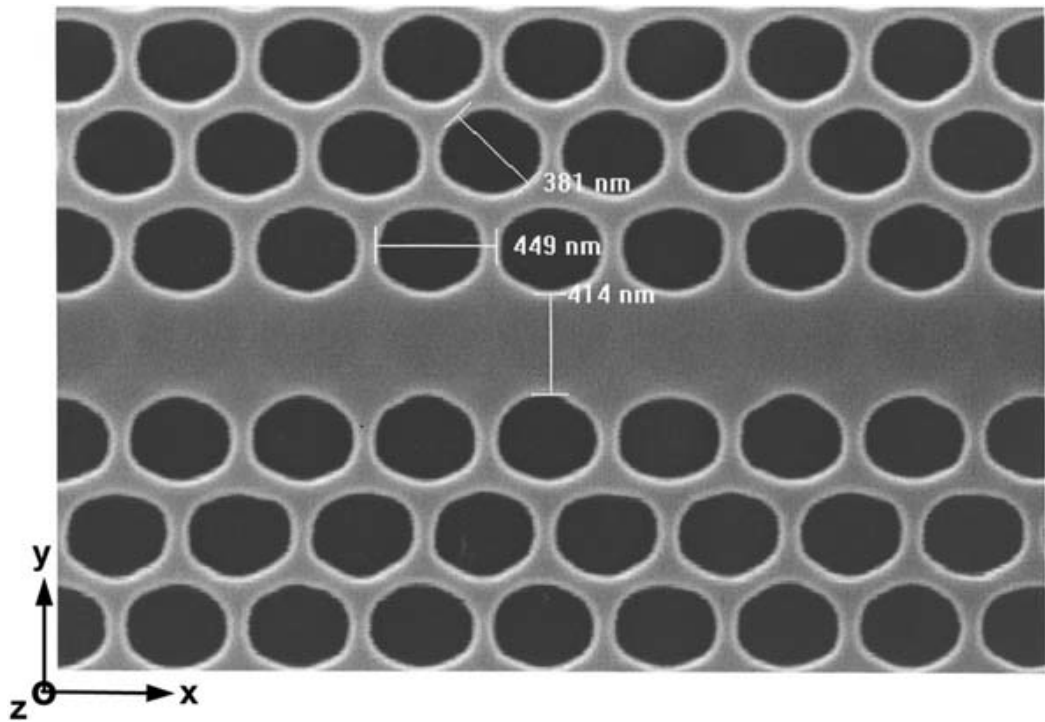
Fig. 9.11. An SEM photo of a line-defect waveguide in a slab PhC

The dispersion diagram for this waveguide is shown in Fig. 9.12, *a*. The light gray region marks the photonic band gap. The diagram reveals that this waveguide is multimode, with four propagating modes lying within the photonic band gap, and below the light line. The light line is established by the refractive indices of the cladding materials above and below the PhC slab. The modes below the light line ideally are not subject to propagation losses. Above the light line, in the dark gray region of the diagram, the spectrum becomes a continuum of states with resonances called quasi-guided modes. These modes correspond to leaky radiation modes of the waveguide that experience propagation losses. Fig. 9.12, *b* shows the steady state field distributions in the *x–y* plane of the four propagating modes calculated by the PWM. Two of the modes are even with respect to the horizontal mirror plane of the slab and two are odd.
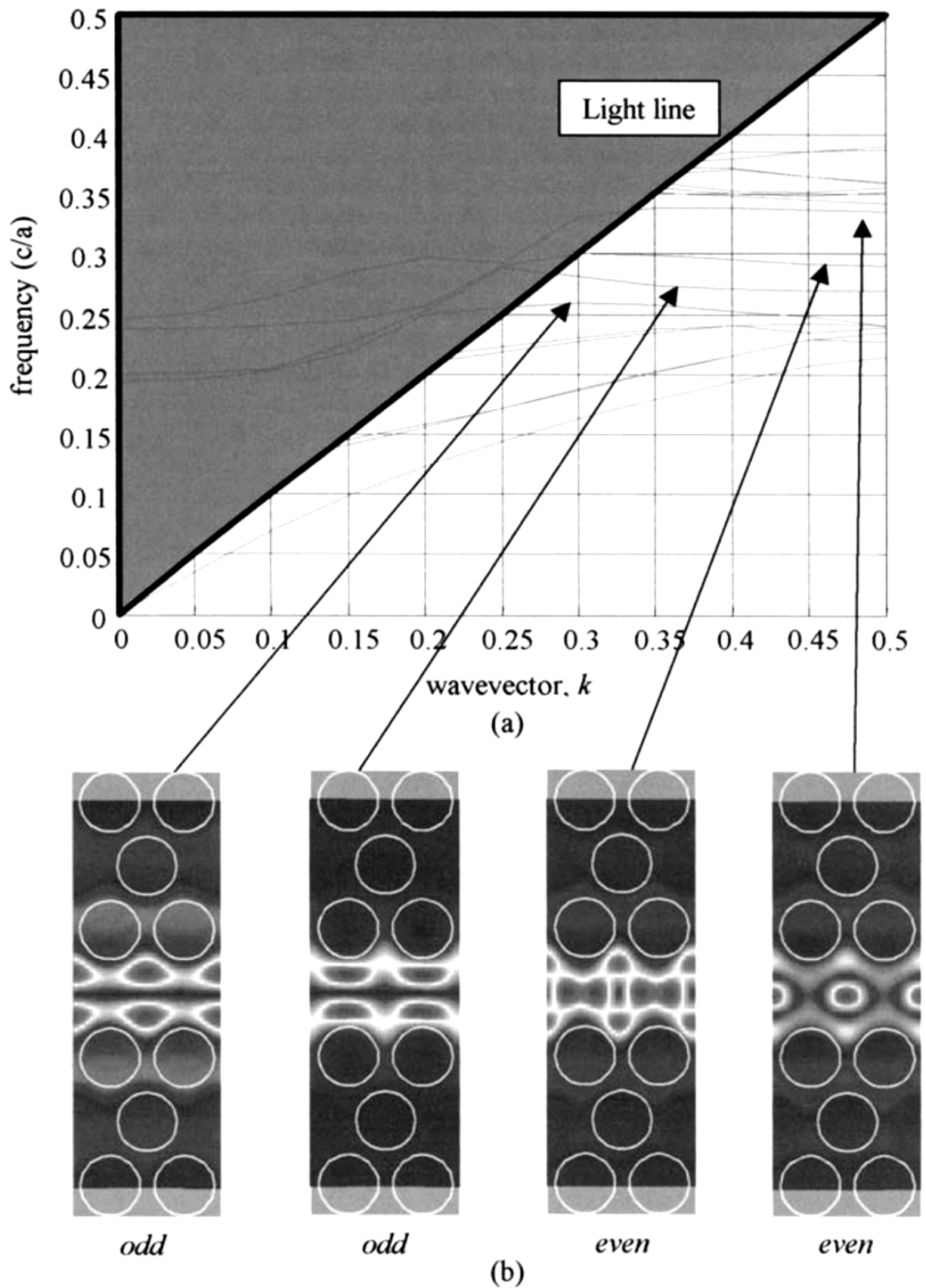
Fig. 9.12. Guided modes of a PhC single-line-defect waveguide:
a – dispersion diagram; *b* – field profiles of the guided modes

In a conventional index-guiding waveguide, in which guiding is by TIR, singlemode propagation can be obtained by reducing the cross-sectional dimensions of the guide until all but one mode is cut off. However, that technique does not work well in PhC waveguides, in which the guiding of different modes is determined by interference effects and resonances rather than index variations. One must use a resonance effect to block the propagation of unwanted modes. It was shown that adding a point defect to a line defect like that of Fig. 9.10, by removing a couple of additional holes on either side of the line, can produce a single-mode waveguide. If an additional defect is doped into the line defect, a mode with a resonant frequency at the cutoff range is strongly localized at the defect, while other modes overlapping with the waveguide band are radiated out through the waveguide. Thus a single propagating mode remains.

Line defect waveguides can be made to cross each other within the plane of the PhC slab without incurring excessive cross talk. For example, the structure shown in Fig. 9.13 has been shown to exhibit nearly 100% throughput and 0% cross talk.

The calculated field pattern of a standard 90° crossing exhibits a significant amount of cross talk. The modified crossing of Fig. 9.13 is seen to be essentially free of cross talk.

There are other types of optical waveguides in PhCs that are not based on line defects. A coupled-resonator optical waveguide (CROW) based on weak coupling between point-defect resonant cavities within a PhC have been proposed. The dispersion relation of the CROW band can be described by a small coupling parameter κ, and the spatial characteristics of the CROW modes remain the same as those of the single-resonator high Q modes. CROWs can be utilized in constructing waveguides without cross talk. They also exhibit enhanced efficiency of second-harmonic generation. It has been demonstrated that a section of CROW inserted into a line-defect waveguide alters its group velocity and positive/negative group velocity dispersion. This approach can be

used to produce delay lines or dispersion compensators in PICs connected by linear-defect PhC waveguides.
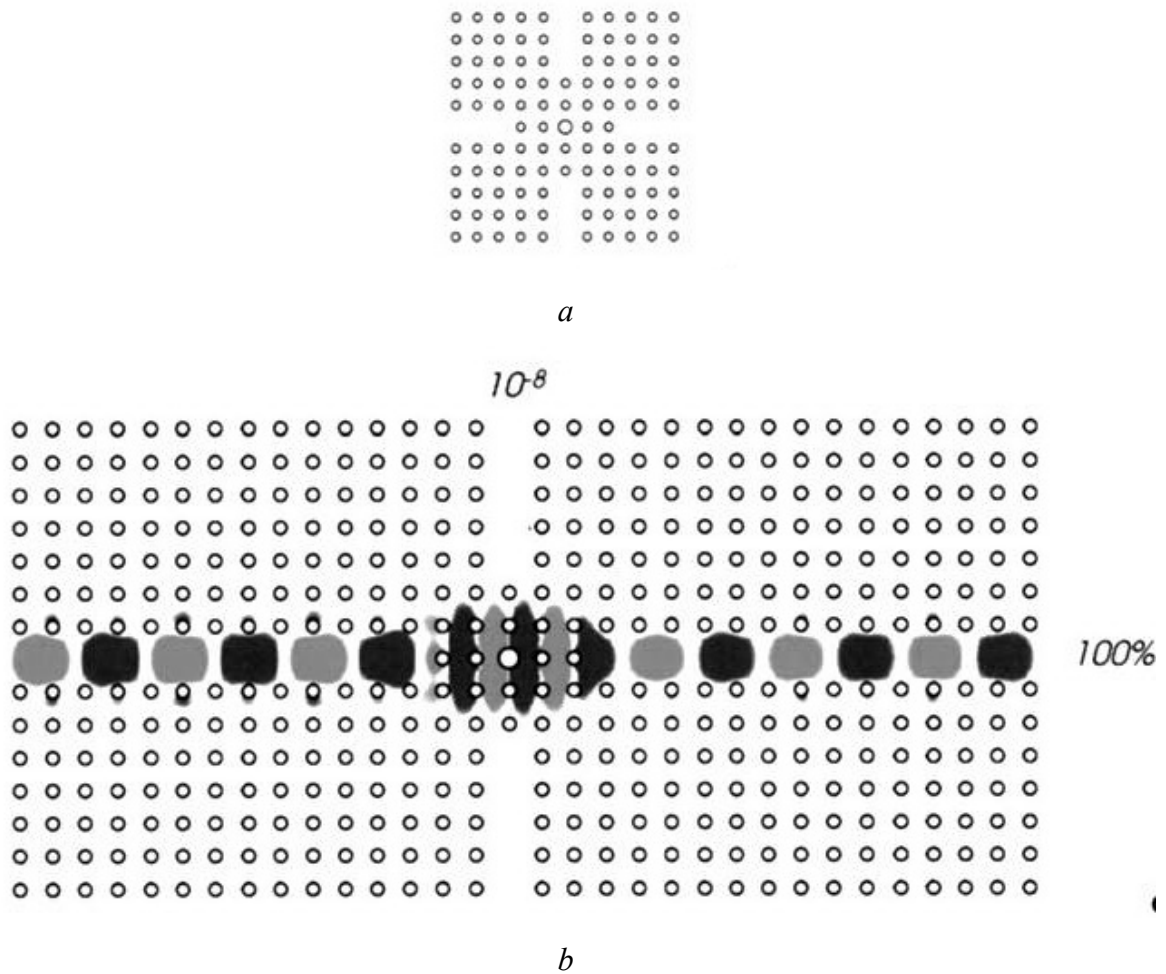

a

b

Fig. 9.13. Diagram of a line-defect-waveguide crossing:
a – structure perpendicular; *b* – field pattern in modified crossing

Because CROWs can significantly decrease the group velocity of light they are sometimes called slow light structures. CROWs may contain bends and branches, suggesting that efficient tunable CROW-based routers and switches may be realized. In fact, a CROW Mach–Zehnder interferometer has already been demonstrated Fig. 9.14 .
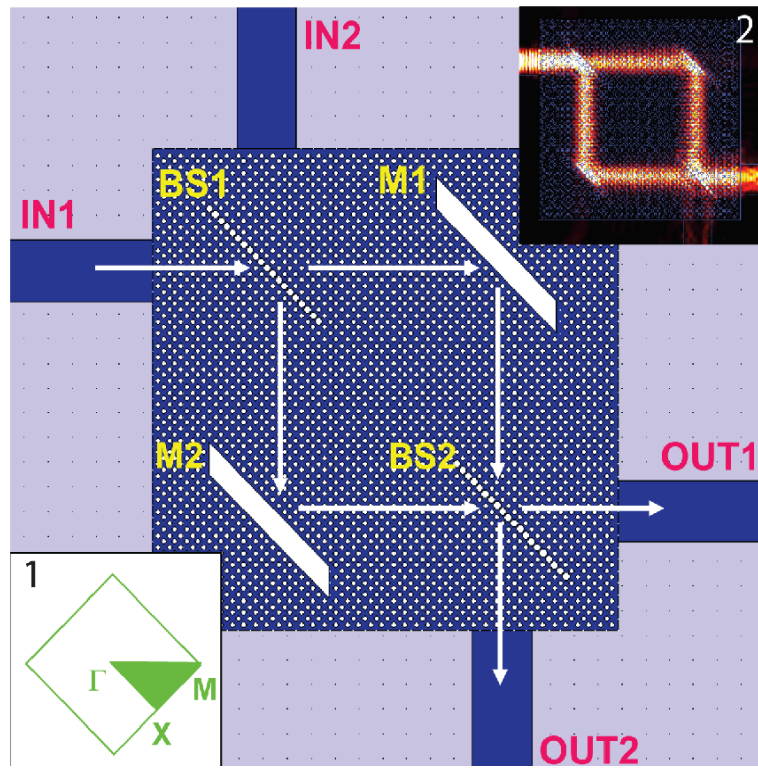
a
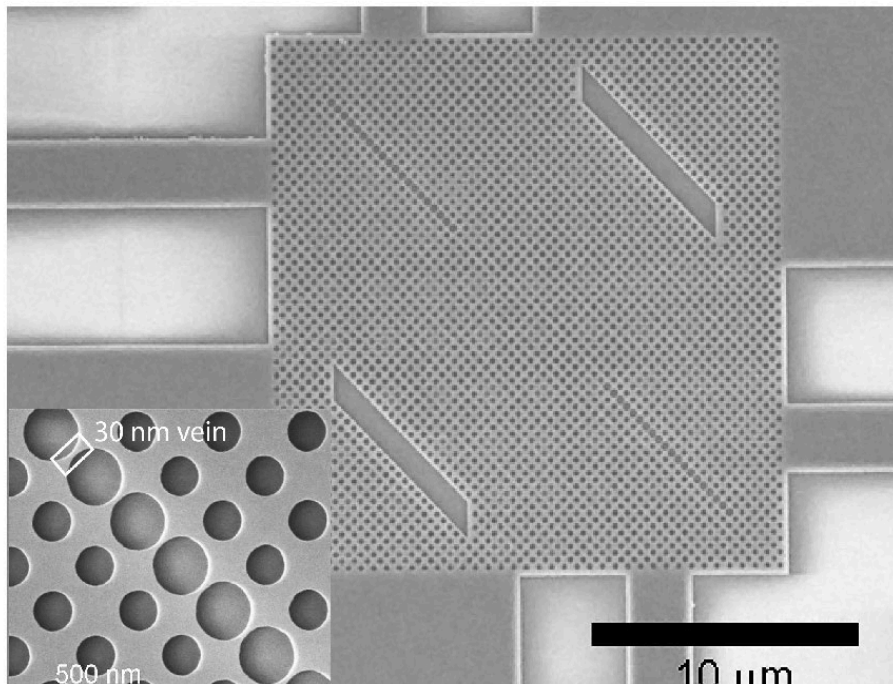
b

Fig. 9.14. PhC Mach–Zehnder interferometer: *a* – Design; *b* – SEM image

BS1 and BS2 are beam splitters, defined by a line defect. M1 and M2 are mirrors, defined by air regions. IN1, IN2 and OUT1, OUT2 are input and output

waveguides, respectively, for light coupling. White arrows indicate the direction of light propagation. Inset 1 (Fig. 9.14, *a*) shows the first Brillouin zone (BZ), with in green the irreducible BZ, which defines the symmetry directions ΓM and ΓX in the PhC. Inset 2 (Fig. 9.14, *a*) is a 2D FEFD simulation result at the collimation wavelength, showing the unidirectional output behaviour of the PhC MZI. SEM image of the PhC MZI, demonstrating accurate realization of the design. All structures (PhC holes, beam splitters, mirrors and aveguides) are etched in a single etching step. The inset (Fig. 9.14, *b*) is an SEM image of a beam splitter, featuring smooth circular holes.

Another class of nanophotonics waveguides is nanowires. Nanowires can be defined as structures that have cross-sectional dimensions of tens of nanometers or less and any length. Because of the quantum mechanical effects that occur at these dimensions they are sometimes also called quantum wires. Many different types of nanowires exist, including metallic (e.g., Ni, Pt, Au), semiconducting (e.g., Si, InP, GaN, etc.) and insulating (e.g., SiO_2, TiO_2).

Nanowire waveguides can be conveniently fabricated on silicon-on-insulator (SOI) substrates, which consist of a thin layer of silicon ($n = 3.45$) on top of an oxide buffer layer ($n = 1.45$) (Chapter 2). Either DUV, EUV or e-beam lithography is used to define waveguides with lateral core dimensions ~10 nm, and dry etching is used to produce waveguides surrounded by air on three sides and SiO_2 on the bottom. Thus $\Delta n > 2$ on all sides of the core and the required tight confinement is provided.

Using this type of nanowire waveguide an array waveguide gratings (AWGs), Mach–Zehnder lattice filters (MZLFs) and ring resonators have been elaborated. SOI wafers consisting of a 220-nm thin silicon layer on top of a 1-μm oxide cladding layer was used. Propagation losses in straight guides of this type have been measured as low as 0.24 dB/mm. For a 90° bend with a radius of 3 μm, the excess loss per bend was determined to be 0.016 dB. Compact couplers and array waveguides have both been made based on silicon nanowire waveguides.

A number of different types of couplers and splitters have been reported that efficiently couple lightwaves between linear-defect waveguides in PhCs. Numerical simulations of electromagnetic-wave propagation in waveguide branches in photonic crystals to identify structures with near-complete transmission, is shown in Fig. 9.15.
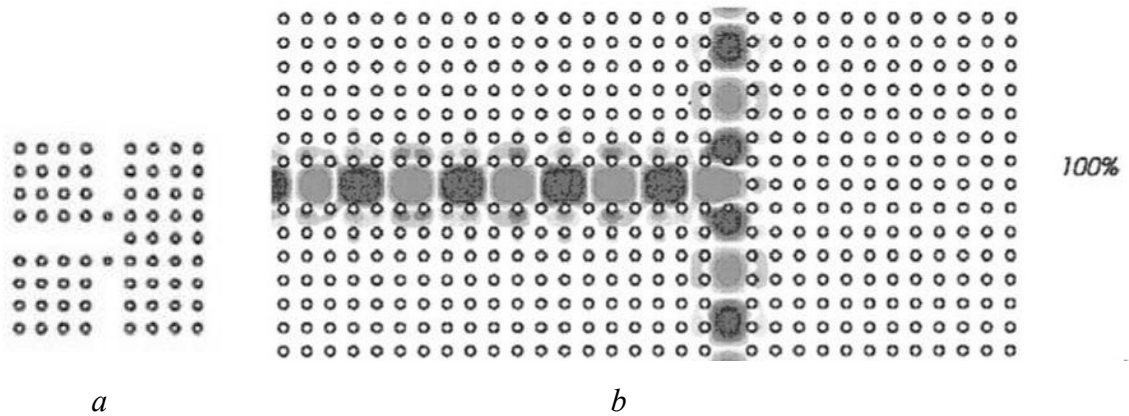

Fig. 9.15. Diagram of a line-defect-waveguide splitter:

a – 3-dB splitter structure; *b* – field pattern

The operation of this device depends on a tunneling process through localized resonant states between one-dimensional continuums. This same mechanism has been employed in the design of an efficient channel-drop coupler Fig. 9.16.
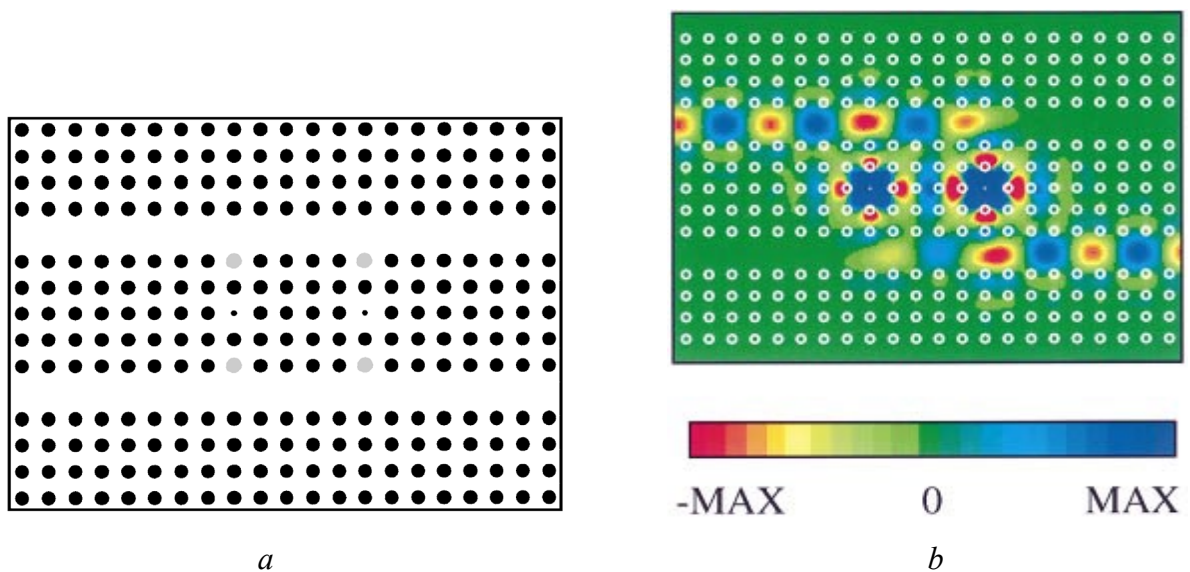

Fig. 9.16. Photonic crystal efficient channel-drop coupler:

a – photonic crystal structure; *b* – electric field pattern of the structure

Dual-channel directional couplers comprised of two closely spaced line-defect waveguides. The devices were fabricated in the 260-nm-thick silicon device layer of a SOI wafer. A triangular PhC lattice of air holes was produced by e-beam lithography with a PMMA resist (200 nm thick) followed by RIE. The lattice constant, a, was 472 nm and the cylinder diameter, d, was 380 nm. The parallel-waveguide structure was made by masking a line of air holes from the etch to make a first waveguide, and masking a second line of air holes parallel to the first guide and separated from it by a single line of air holes. Couplers were made with three different between-guide spacings by changing the diameter of the air holes in the row separating the two coupled waveguides. This changed the amount of overlap of the evanescent tails and hence the coupling coefficient.

The line-defect-waveguide couplers discussed so far have all coupled from one waveguide to another of similar dimensions. However, often it is necessary to couple a lightwave beam to a waveguide from the air, or from a waveguide of a different size. One way to accomplish this is to use a tapered coupler. The taper should be adiabatic in order to prevent mode conversion as lightwaves progress through it. An ideal structure would be one that adiabatically tapers in both the vertical and lateral directions and can be monolithically integrated with the PIC.

Coupling to a nanoscale waveguide in a PhC has also been done by using modified versions of some of the couplers described in earlier chapters, such as prisms, gratings and micro-lenses , as well as some new device geometries, such as the J Coupler .

9.3. Resonators

The waveguide splitter of Fig. 9.15 used a resonant structure to provide the necessary coupling from the main channel to the two secondary channels. And the addition of a single point defect to the waveguide of Fig. 9.11 was all that was necessary to form a resonant structure that blocked the higher-order

modes and made the waveguide single mode. Such resonant structures are found frequently in PhC devices and PICs. By tradition, they are sometimes called "cavities" even though they lack conventional cavity form.

The similarities between a point-defect structure in a PhC and a traditional cavity resonator can be seen in Fig. 9.17. This PhC cavity was created by reducing the radius of the lattice site within the cavity by a factor of 1/2 . Like a conventional optical cavity it consists of a localized mechanical structure with recognizable "walls". They both establish a time-invariant distribution of electromagnetic energy, or mode, and they both select a characteristic resonant frequency that is related to their dimensions and dielectric constant.
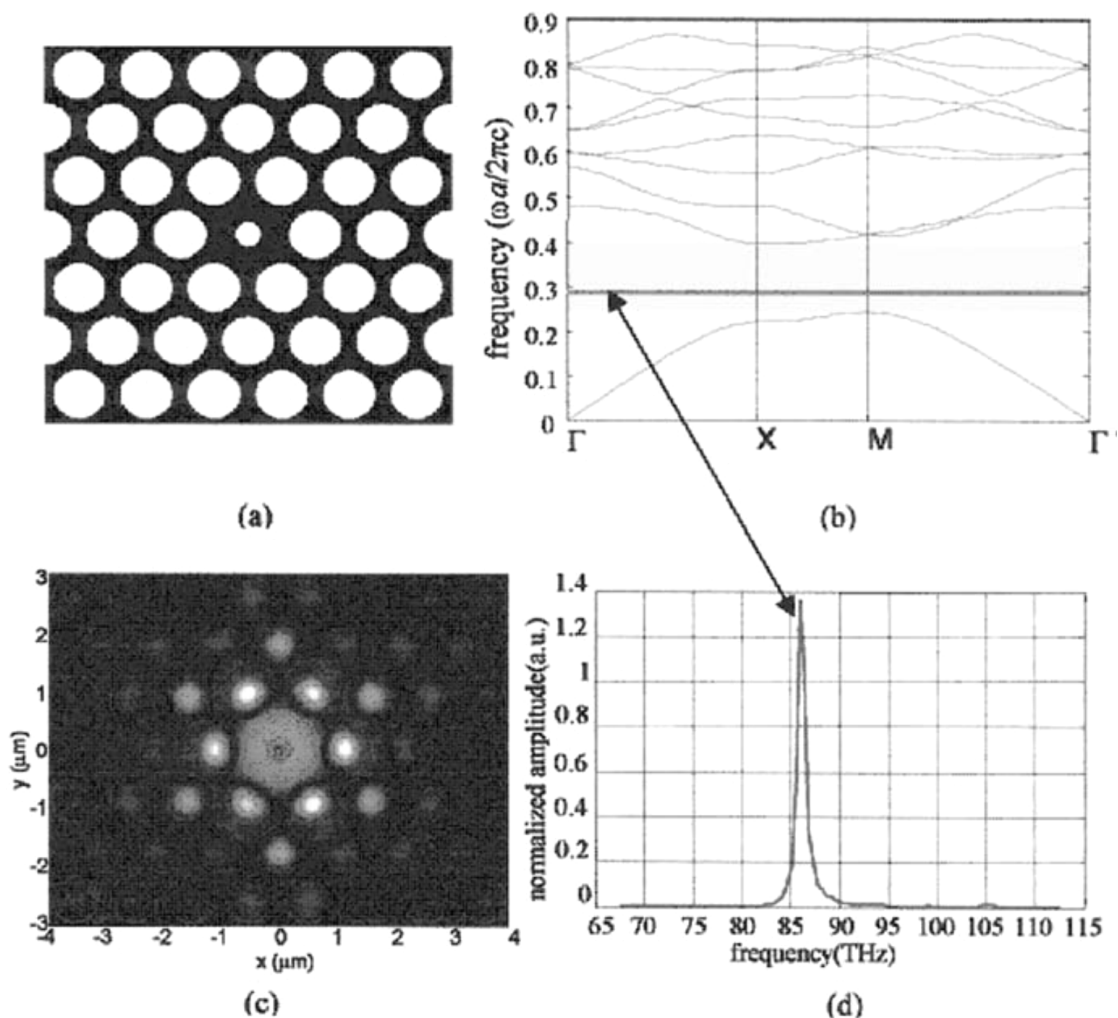

Fig. 9.17. A photonic crystal point-defect cavity:

a – a donor-type point defect in a PhC lattice; *b* – the dispersion where the state produced by the defect is marked by the *dark line*; *c* – steady-state field profile of the cavity mode for TE polarization; *d* – the frequency response

A PhC point-defect cavity can be tuned to a different frequency by changing either its dimensions or its dielectric constant. The air-bridge resonator shown in Fig. 9.18. it PhC consists of a linear array of regularly spaced holes in a GaAs slab that is supported on Al_xO_y piers to form an air-bridge resonator. A microcavity is formed by varying the spacing, d, of two central holes. The resonant wavelength depends on the size of d. Quality factors as high as 360 were measured at wavelengths near 1.55 μm, with modal volumes as small as 0.026 μm³.
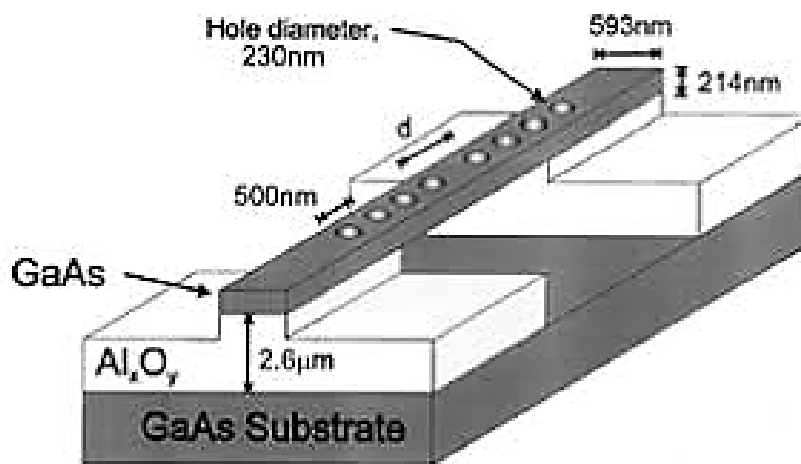

Fig. 9.18. Diagram of structure of Air-bridge resonator

Other types of PhC cavities with tunable resonant frequencies can also be made.

For example, by made a tunable cavity based on modulation of the near-field coupling between two sub-wavelength periodic nanostructures embedded in a Fabry–Perot cavity. The filter is composed of two planar dielectric mirrors forming a Fabry–P´erot cavity, with a field localization nanostructure affixed to the inner surface of each mirror. Each nanostructure consists of a transverse infinite sub-wavelength periodic grating with period $\Lambda = 0.6\ \lambda_0$, where λ_0 is the intended resonant optical wavelength. A small air gap of $0.01\ \lambda_0$ between the nanostructures allows near-field coupling between them. Either lateral or

longitudinal mechanical displacement of the two nanostructures causes a change in the resonant frequency of the cavity.

As follows from the foregoing, photonic crystals represent a very attractive object for a integrated optics. The main practical value is represented by photonic crystals with defects, with the help of which it is possible to very efficiently control light fluxes.

Such structures may possess the properties of waveguides, resonators, filters, splitters and other devices (Fig. 9.19) necessary to create a fully optical computer.
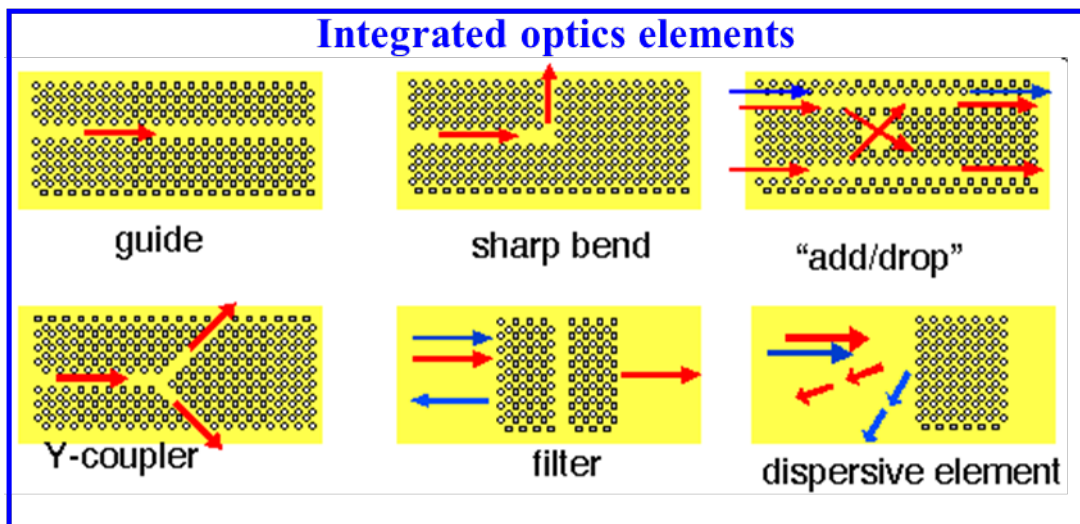

Fig. 9.19. Integrated optics elements based on photonic crystal structures

10. POLYMER-BASED INTEGRATED OPTICS DEVICES

Polymer waveguide technology has a great potential for economic mass production of complex planar photonic circuits that comply with the severe requirements imposed by applications in communication systems. The low-cost prospect arises from the availability of a wide range of cheap optical polymers and the simplicity of fabricating waveguides from them. A significant subset of optical polymer materials has shown excellent optical, chemical, and mechanical characteristics, that are very attractive for applications in long and short-distance data communication. For instance, most of these polymers are transparent in the wavelength range from 400 to 2000 nm, and losses of polymer waveguides can be as low as 0.1 dB/cm in the three telecommunications windows around 850, 1300, and 1550 nm. The available optical polymer materials have a refractive index in the range from 1.3 to 1.7, with the possibility of tailoring and fine-tuning the index of some of these materials. Polymers also can be used for active devices for routing, switching, and even high speed modulation of optical data, because all polymers have large thermo-optic coefficients and, in addition, a variety of polymers have shown good electro-optic properties. A large number of optical polymers have proven to be thermally, chemically, and mechanically stable, thus fulfilling the heavy requirements needed for operation under harsh environmental conditions.

The most appealing characteristic of polymer waveguide technology is the simplicity and flexibility of waveguide fabrication methods. Polymer thin films can be deposited in a wide thickness range by spin or dip coating using relatively simple equipment. A variety of channel waveguide fabrication methods exists, ranging from existing microtechnology techniques such as etching, to mass production methods developed especially for polymers, including molding and laser delineation.

Polymer waveguide technology for communications applications must satisfy severe requirements with regard to both the polymer materials as well as the waveguide fabrication methods.

For a polymer to be selected as a material for waveguide technology, it has to satisfy the following set of requirements:

• Low optical losses (not more than 0.1 dB/cm) in the communication spectral windows around 800, 1300, or 1500 nm.

• Low wavelength dispersion.

• Low birefringence.

• Low polarization-dependent losses.

• Thermally stable optical and mechanical properties.

• Resistant to humidity.

• Good mechanical properties such as flexibility and toughness.

• Low cost.

In the literature a large number of polymers have been used to fabricate waveguides.

10.1. Thermo-optic properties

Materials that have a large thermo-optic coefficient (dn/dT) have a great potential for use in active device applications like switches and tunable filters. Polymers have a relatively high thermo-optic coefficient (1 order of magnitude larger than that of silica, but with opposite sign). The thermo-optic coefficient of polymers is negative, which means that the refractive index of the polymer decreases as the temperature increases. The thermo-optical coefficients of some polymers and, for comparison, SiO_2 are shown in Table 10.1. The use of waveguides that employ polymers and silica layers simultaneously can lead to the design of a thermal waveguides, which allow critical functions, for example, wavelength multiplexing, without the need for active thermal stabilization of the device.

Table 10.1

Thermo-optic coefficient of selected polymers in comparison with SiO_2

Material	$(dn/dT)x10^{-4}$ C^{-1}
SiO_2	0.101
PMMA	−1.05
PS (polystyrene)	−1.4
PC (polycarbonate)	−1.07
SAN (styrene acrylonitrile)	−1.1

10.2. Electro-optic properties

We concentrate here on the well-known electro-optic Pockels effect, which describes a linear relationship between the refractive index and an externally applied electric field. This effect has been used extensively for modulators and switches in crystalline inorganic materials, especially lithium niobate (Chapter 5). In polymers, the linear electro-optic effect is based on electro-optic molecules or groups that, for a macroscopic effect, need to be aligned in a process called poling. Although this adds some complexity compared to the crystalline materials, which usually do not need poling, it also adds design flexibility by allowing selective poling. An important advantage of suitable polymer materials is that their permittivity does not vary strongly over a very broad frequency range from zero up to visible light frequencies.

This property facilitates the design of high-frequency devices where the propagation speed of the driving radio frequency or microwave signals should be matched to that of the optical wave.

An important issue to be dealt with in electro-optic devices is the birefringence of the electro-optic material that arises from the alignment of its molecules. In applications where the polarization of the input signal can be controlled, this is not a problem, but in other cases, for example, when input signals are randomly polarized with large fluctuations in time that are generated

by passing through a long length of standard telecom fiber, it is quite a challenge to obtain the desired polarization-independent behavior.

A special type of material that also must be mentioned here is electro-optic organic crystals. These materials do not need an additional poling treatment and have very high electro-optic coefficients (more than 100 pm/V). Some polymer formulations have been designed to have a large electro-optic coefficient (as large as 300 pm/V, the largest value achieved in any material system) These formulations are typically composed of standard polymers (e.g., polycarbonate) impregnated with specialty chromophores (e.g., CLD-1). They exhibit a large electro-optic effect once subjected to poling, a process where large electric fields (~200 V/μm) are applied to the material in order to orient the molecules. However, the result of the poling process is not stable with time or with environmental conditions, limiting the applications where polymer electro-optic modulators can be used.

10.3. Fabrication of polymer optical waveguides

In the foregoing text, the materials properties of optical polymers were discussed. The next step is the realization of optical waveguides based on these polymers, which includes the deposition of polymer films and their structuring. So as to end up with the desired waveguide pattern. Fabrication methods used to produce polymer waveguides have to comply with the following set of general demands:

• Optical losses introduced by fabrication, such as that caused by scattering on rough side walls, have to be kept to a minimum.

• The method has to be simple and reproducible.

• The fabrication process should be suitable for low-cost mass production.

• High precision, sometimes well below 1 μm, is required.

There are at least three fundamental categories of polymer channel waveguide fabrication techniques, namely, (1) optical lithography combined

with etching, (2) molding, and (3) photolithographic delineation. Important aspects of these three approaches are presented next.

10.3.1. Optical lithography combined with etching

The most common approach to realizing optical waveguides is deposition of a layer stack that has the desired optical properties, and structuring it by optical lithography and subsequent etching. The advantage of this approach is versatility, because all kinds of deposition methods and etching processes can be applied. The number of processing steps, in most cases more than 20 and some of them very critical, makes this method relatively expensive and less suitable for low-cost mass production.

The deposition of a polymer layer stack is far from trivial. The first problem is adhesion of the undercladding to the substrate and adhesion between subsequent polymer layers. Then the deposition process, in many cases spin coating, should result in homogeneous, crack-free layers that fully cover the entire chip area. Between subsequent deposition stages, the layers should be cured to avoid partial dissolution of existing layers by the freshly deposited layer.

Dry or wet etching can be used to structure polymer waveguides. Among the dry etching methods, reactive ion etching (RIE) is the most efficient and widely used, because it has good selectivity, little undercut, and high productivity. During the RIE process, a plasma discharge produces ions that directly or indirectly enhance the etch rates of selected material. Generally lower pressures and higher ion energies are required compared to plasma etching.

RIE has an inherent problem, namely, it creates side-wall roughness and changes the cross-sectional waveguide shape. The polymer debris that is formed by the etching process may be redeposited on the side walls where the ion bombardment is much weaker, thus causing roughness. The roughness can be

minimized by blowing with an inert gas, by heat treatment, or by dipping in a solvent.

Figure 10.1 a scanning electron microscope (SEM) picture is given of a highly multimode polyurethane waveguide after etching of the core layer (64×120 μm).
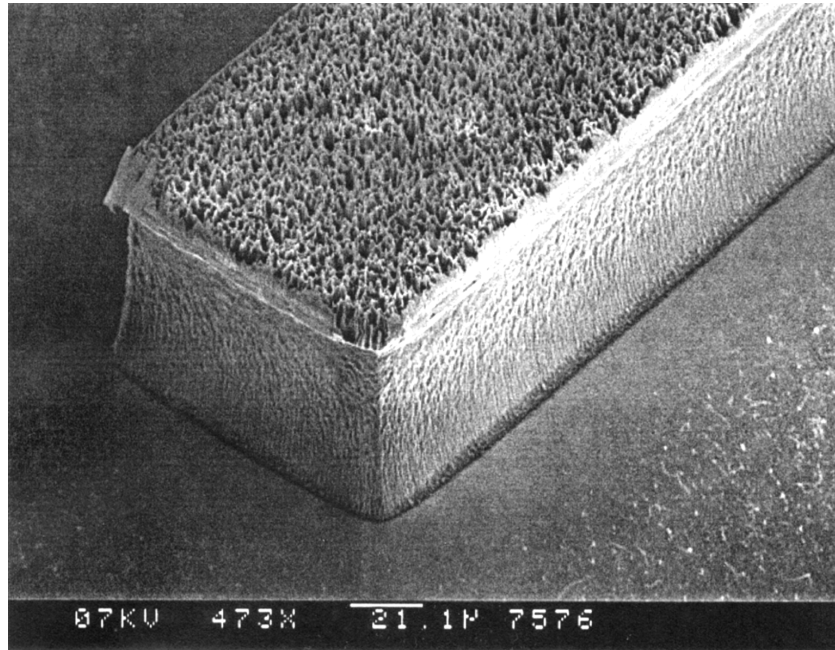

Fig. 10.1. SEM picture of a multimode polyurethane waveguide structured by RIE etching using a nonerodable Al mask. The waveguide dimensions are 64 × 120 μm

10.3.2. Molding techniques

The term "molding" categorizes a number of low-cost replication techniques, including casting, compression molding, injection molding, soft embossing, hot embossing. Molding is the most suitable technique for mass production. A structure can be fabricated in large numbers using only one mold.

Later we will consider the imprinting techniques in more detail as a variety of molding techniques.

10.3.3. Photolithographic delineation

Polymer waveguides can be realized by externally induced dopant diffusion. There are two types, namely, photolocking and selective polymerization. In both, a monomer is doped with a certain photosensitive dopant. The doped monomer is spun onto a substrate and the layer is exposed under a mask for UV light. The dopants are locked into the host polymer during the photo-polymerization reaction. When the polymer material is heated after removal of the mask, the exposed regions are polymerized, whereas out-diffusion of the dopants can occur in the unexposed regions. In photolocking, the core regions are exposed and the dopants in the cladding are out-diffused, which lowers the refractive index, because the dopants have higher refractive indices than the host polymer. PMMA has been doped with benzophenone and coumarin to form waveguides by this technique. In selective polymerization, a lower refractive index material is out-diffused from the core region, leaving a high-index waveguide. This method offers flexibility in creation of stable core layers by tailoring the molecular properties of the host or the dopant. It provides smooth top surfaces and side walls of the core layer. PMMA, polystyrene, and polyimide have been used for waveguide applications by utilizing this method.

A related method is called photolithography delineation or contact printing lithography because it is similar to negative-tone photoresist. A photosensitive polymerization initiator is added to a monomer to provide a means of photochemically initiated polymerization. The monomer is spun onto a substrate and exposure is performed using a mask. The exposed region is polymerized and the unexposed region remains monomeric. The pattern is developed by conventional wet etch of the unreacted material with a standard organic solvent such as methanol.

Internal polymerization-driven diffusion reactions in the absence of dopants is a method of polymer waveguide fabrication that utilizes diffusion of monomers during the polymerization process. The monomer layer is exposed by

using a mask for a certain dose of light with a wavelength of 350–400 nm. During exposure, the light initializes polymerization in the waveguide region and monomers begin to diffuse to that region. Due to this mass transfer, this region has a higher refractive index. The refractive index of the guiding layer can be controlled by the total dose of light, diffusion time, diffusion temperature, and presence of oxygen. Without using a mask, waveguides can be fabricated by laser writing. This technique utilizes a controllable focused.

UV laser beam and UV-curable polymers or crosslinked polymers. The UV-curable polymer is a photo polymer that is deposited by spin coating and cured with UV light exposure. The cross-linked PMMA becomes highly soluble in the developer solution after exposure. After exposure, the pattern is developed with a standard organic solvent such as acetone. In the case of UV-curable polymers, the unexposed region will be removed, leaving a channel waveguide pattern; for cross-linked polymers, the exposed region is soluble and can be dissolved by wet etching.

10.3.4. Imprinting techniques

Besides the conventional photolithography and etching technique, the imprinting technique can be utilized to fabricate polymer photonic waveguide devices, which can fully exert the plastic property of polymer materials. Several commercial polymer materials are available for imprinting craft.

Certainly, there are still some critical issues to be solved to obtain good performance imprinted structures, elements, and devices. First, the imprinting technique is a contact process. The pattern distortions can easily happen during the mold separation process. The flexible/soft molds can be applied, or the mold surface can be coated with a thin anti-adhesive film to facilitate the demolding process. Second, there is usually a residual layer left on the substrate after imprinting, which needs to be removed before subsequent processing. It can be

removed by using the etching process, or it can be decreased by using improved techniques, such as the selective imprint technique.

Certainly, not only the nanometer scale structures, but also various dimensions, ranging from the nanometer to the micrometer or even millimeter scale, can be formed using the imprinting technique. Therefore, besides the integrated circuits, the photonic and optoelectronic devices can be fabricated using the highly efficient and inexpensive imprinting technique. By using the master mold (hard mold or soft mold) fabricated by photolithography (for large-scale features) or electron-beam lithography (for small-scale features), a reverse replica of the mold pattern can be formed on the polymer layer through heating and apply pressure for solid plastic polymers or pressing and following UV curation for liquid polymers, as shown in Figure 10.2, which is known as the thermal imprinting technique and the UV imprinting technique, respectively.
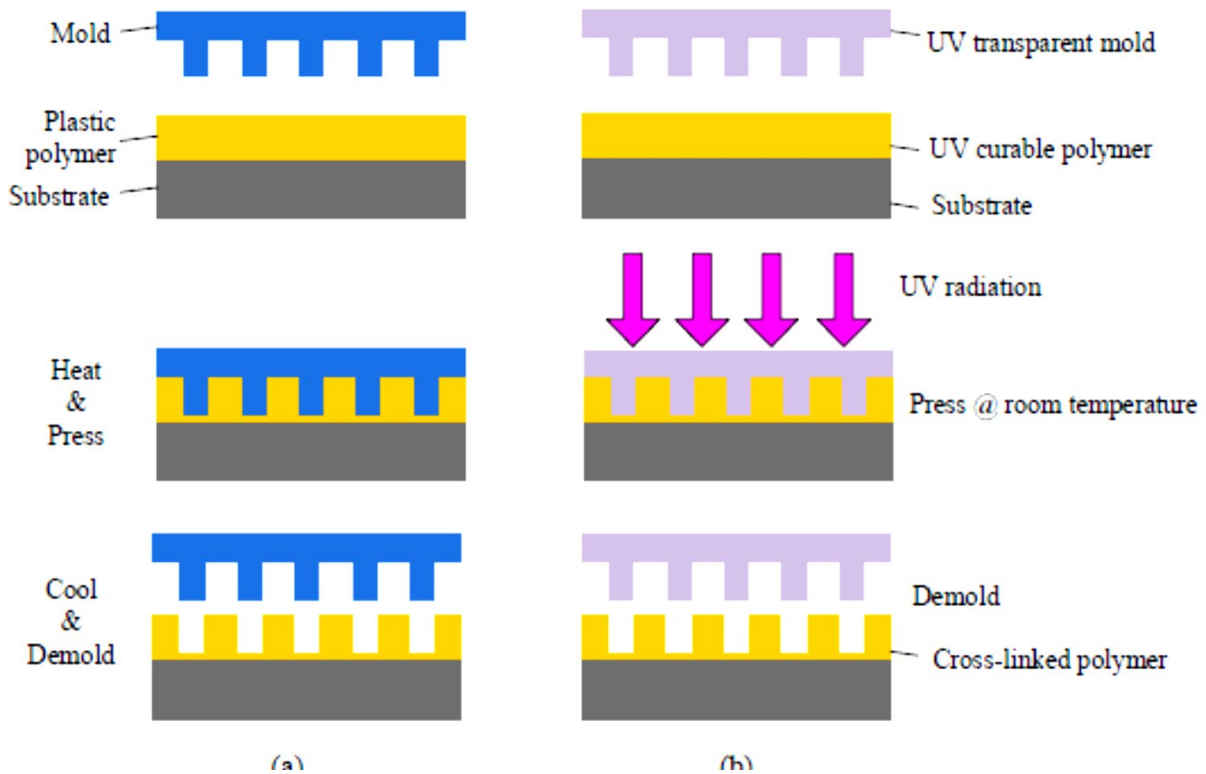

Fig. 10.2. Schematic of fabrication process: *a* – thermal imprinting; *b* – UV imprinting

10.4. Passive photonic integrated waveguide devices

The passive photonic integrated waveguide devices, such as microring resonators, waveguide splitters, arrayed waveguide gratings, long-period gratings, and microlenses, are key components consisting of large scale integration photonic circuits.

The basic microring resonator usually consists of a microring waveguide coupled with a single straight waveguide or double straight waveguides, as shown in Figure 10.3. They can be utilized as the key structures of filters, modulators, switches, and sensors in the photonic integrated circuit
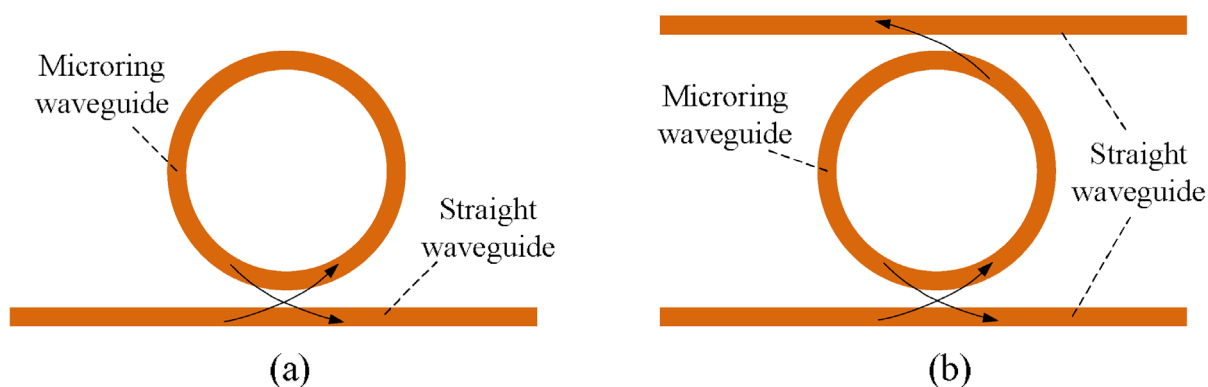

Fig. 10.3. Schematic structures of microring resonators:
a – "all-pass" type; *b* – "add-drop" type

For the imprinted polymer waveguide, the residual layer of waveguide core should be thin enough to confine the optical mode field in the core section. Otherwise, it results in a large radiation loss due to the expanding of the optical field into the residual layer, especially for the bent waveguide. The etching process is generally needed to reduce the thickness of the residual layer. However, it increases the complexity of waveguide fabrication. Two new kinds of imprinted polymer waveguide structures have been developed. One is the inverted ridge waveguide generated by imprinting the cladding, and the other is the ridge waveguide with a thin residual layer generated using the selective imprinting technique.

Figure 10.4 shows the fabrication process of the inverted ridge waveguides. The cladding layer rather than the core layer is imprinted to form the groove, and then the core waveguide is defined using spin-coating step to fill the groove. This process smartly avoids the difficulty of controlling the thickness of residual layer in the conventional imprinting process with ridge structure. The thickness of the slab layer can be made thin enough by increasing the spin-coating speed or using diluted core polymer material.
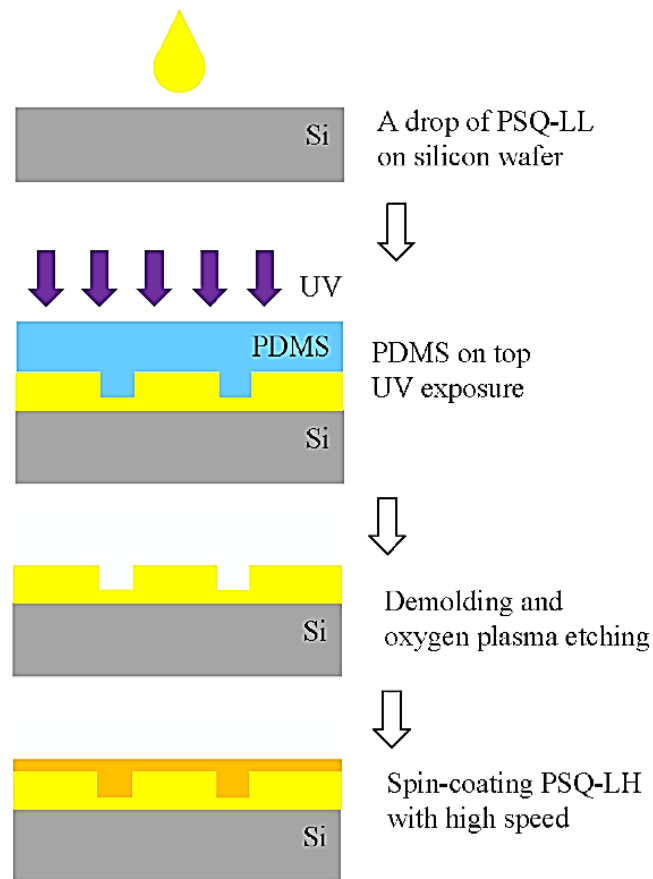

Fig. 10.4. Fabrication process of the inverted using UV soft imprinting technique

Photonic waveguide devices with ridge structure are more desirable for some applications,especially for biosensors. However, the residual slab layer of the imprinted ridge waveguide is usually very thick, which limits the minimum radius of microring, as well as the free spectrum range.

The selective imprinting technique was developed to reduce the thickness of the residual slab layer. The schematic processes of the conventional

imprinting technique and the selective imprinting technique are illustrated in Figure 10.5. Instead of using the single trench pattern mold to directly imprint the waveguide into the core layer as shown in Figure 10.5, *a*, the ridge shape waveguide is formed by imprinting two trenches on both sides of the waveguide (see Figure 10.5, *b*).
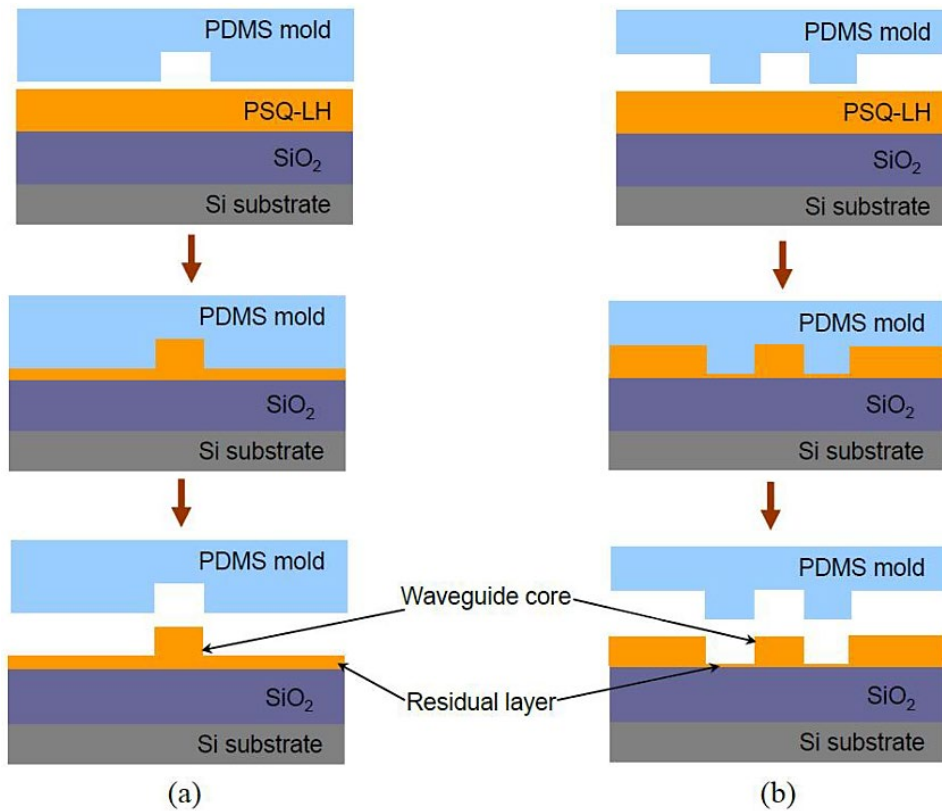

Fig. 10.5. The schematic process of (*a*) the conventional imprinting technique and (*b*) the selective imprinting technique

A polymer waveguide with good rectangular shape, as well as negligible residual slab layer thickness (below 100 nm), was fabricated with this technique.

10.4.1. Optical waveguide splitters

Optical splitters, also known as splitters, are one of the most important passive devices in optical fiber communication system. With the rapid development of telecommunication system, especially the fiber-to-the-home

(FTTH), the need for high-performance, low-cost splitters has become increasingly urgent. The fiber-based optical splitters with one input and multiple outputs are usually utilized. However, the output uniform of the divided power and the insertion loss will deteriorate when the number of outputs increases. The photonic integration technique provides a potential way with which to fabricate optical waveguide splitters with small size and good performance. With the help of imprinting technique, polymer-based optical waveguide splitters can be fabricated with low cost and low insertion loss.

The multimode interference (MMI)-based optical waveguide splitter in Figure 10.6, *a*, which avoids the narrow structure as *Y*-junction based splitter, is tolerant of the structure deviations by the fabrication process. A taper is introduced between the multimode waveguide and single mode waveguide for lower coupling loss. According to the self-imaging theory of multimode waveguide, the MMI-based waveguide splitter is optimized with the structure parameters $L = 850$ μm, $D = 22$ μm , $W = 40$ μm, and $L_{taper} = 350$ μm.
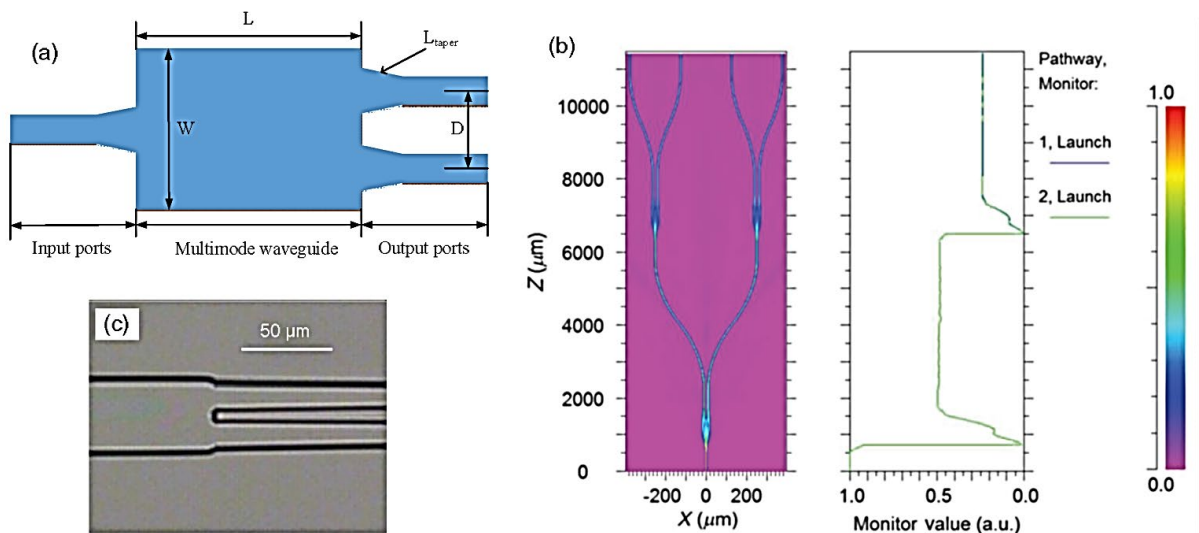

Fig 10.6. Schematic of the MMI-based waveguide splitter (*a*); the simulated beam propagation in 1x 4 waveguide splitter (*b*); the microscopic image of the connection section between the multimode waveguide and the two single mode waveguides (*c*)

Polarization-dependent loss (PDL) is measured as small as 0.05 dB due to the low birefringence of the polymer material and the square.

The simulated beam propagation in 1 x 4 waveguide splitter is illustrated in Figure 10.6, *b*. As the splitter is expected to connect with fiber array, S-bend waveguide is applied to enlarge the distance between output ports (to be 250 mm).

Figure 10.6, *c* shows the microscopic image of the connection section between the multimode waveguide and the two single mode waveguides. The average insertion loss of the splitter is 12.98 dB with uniformity of 1.08 dB.

10.4.2. Arrayed waveguide gratings

The arrayed waveguide grating (AWG) with the similar function of diffraction grating can be used as a wavelength multiplexer or demultiplexer. AWGs play an important role for the dense wavelength division multiplexing (DWDM) optical fiber communication system.

Figure 10.7, *a* shows the top view of the fabricated AWG, in which polymer material ZPU12-450 with refractive index of 1.45 at wavelength of 1.55 μm and ZPU12-430 with refractive index of 1.43 at wavelength of 1.55 μm are utilized as core and cladding, respectively. It can be seen from Figure 10.7, *a* that there are mainly five parts in the AWG, including a single input waveguide, two slab waveguide, a group of arrayed waveguides, and multiple output waveguides. Except the slab waveguide, the waveguide is basically a single mode one. In order to improve the lightwave coupling efficiency between the slab waveguide and the single mode waveguide, a tapered structure has been incorporated in between. A reverse imprint procedure is applied, as shown in Figure 10.7, *b*, incorporating a PDMS stamp, in which the waveguide patterns are formed in the core instead of the lower cladding without resorting to spin-coating process.
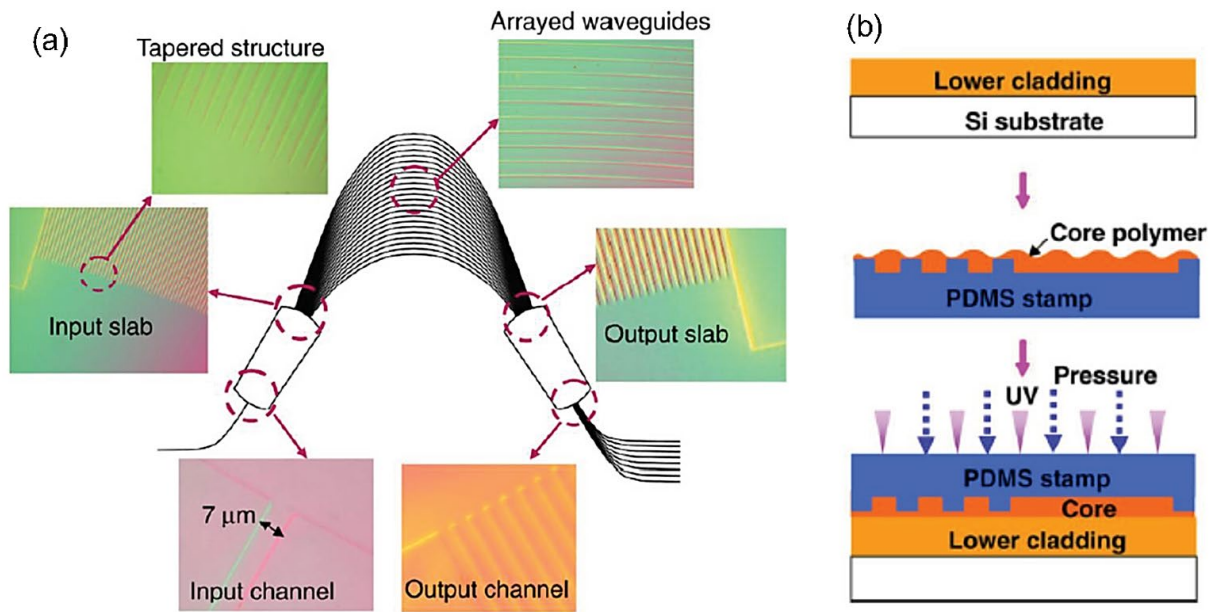
Fig. 10.7. Schematic diagram showing the plan view of the imprinted AWG device (*a*); the reverse imprint procedure for core layer patterning (*b*)

10.4.3. Microlenses

Microlenses and microlens arrays (MLAs) have shown promising ability for light collection and collimation, with a large range of applications including imaging sensors, interconnects, and photodetectors. Various fabrication methods for MLAs have been investigated. However, the fabrication cost is still too high for large-scale production due to the complex steps involved. Figure 10.8 shows polymer 40 × 40 microlens array with a pitch of 250 μm, a diameter of 240 μm, and a sag height of 24.5 μm. Microlenses are replicated into the UV curing polymer material PAK-01 by step and stamp UV imprinting technique on silicon substrates with a diameter of 150 mm. The resulting substrates are used as masters to cast PDMS templates. With the optimized dispense strategy, including the adjustment of the dispense height, pressure, and amplitude, air bubbles can be avoided in the dispensed resist layer. Figure 10.8, *a* shows the SEM image of the replicated microlenses into the resist PAK-01.
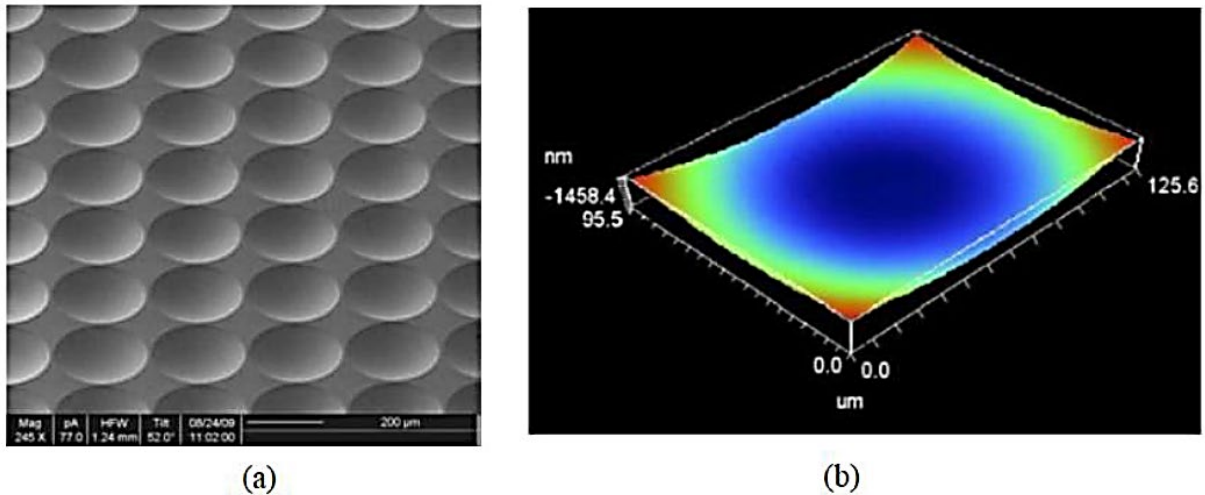
Fig. 10.8. SEM image of the replicated microlenses into the resist PAK-01 (*a*); surface profile of an imprinted microlens (*b*)

10.5. Active photonic integrated waveguide devices

Compared with the inorganic optical materials, such as silica, glass, and SiN_4, organic polymers have large TO coefficient, which has the potential to decrease the power consumption of TO integrated waveguide devices.

As mentioned in Section above the microring resonators are the key unit in the photonic integrated circuit. It is desired that the coupling coefficient or the resonant wavelength of the microring resonator can be tuned according to the requirement of applications. Figure 10.9, *b* shows the schematic structure of the tunable microring resonator, in which the Mach-Zehnder interferometer (MZI) with a micro-heater on one of its arms is introduced to replace the conventional directional coupler, as shown in Figure 10.9, *a*, while another micro-heater is placed on the microring waveguide. Consequentially, the coupling coefficient and the resonant waveguide can be tuned with micro-heaters.

The resonant wavelength of the fabricated polymer microring resonator can also be tuned. This is realized by applying the power to the micro-heater on the microring waveguide. The resonant wavelength may be tuned freely within the FSR around 0.13 nm.
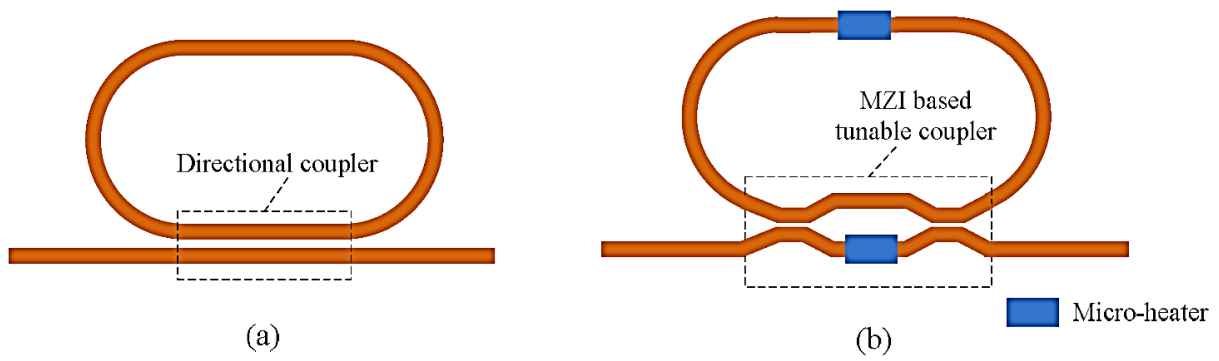
Fig. 10.9. Schematic structures of the microring resonators consisting of (*a*) a directional coupler and (*b*) a MZI based tunable coupler

10.5.1. Electro-optic modulators

Compared with inorganic materials such as the most widely used $LiNbO_3$, the polymeric EO materials have many advantages including lower cost, ease of processing, and larger EO coefficients. Especially, the polymeric materials have the potential for achiving high EO coefficients by virtue of synthetization at molecular level. For example, the long flexible chain-modified, julolidinyl-based chromophores linked by thiophene bridge and TCF acceptor (chromophore WJ5) afford the EO coefficient (r_{33}) value of 266 pm/V. The chromophore with thiophene-modified π-conjugation enhances the EO coefficient to be 337 pm/V.

Incorporating the highly nonlinear and stable chromophore doped in amorphous polycarbonate, a MZI based EO intensity modulator have been designed and fabricated using the UV imprinting technique. Figure 10.10, *a* shows the process steps for the fabrication of PDMS stamp (i–iii) and the molding of EO polymer layer (iv–vi). The cross-sectional view of the electro-optic waveguide structure is illustrated in Figure 10.10, *b*. The fabricated EO modulater is measured with the halfwave voltage (V_π) of 8.4 V at 1600 nm, and the on/off extinction ratio is better than 19 dB.
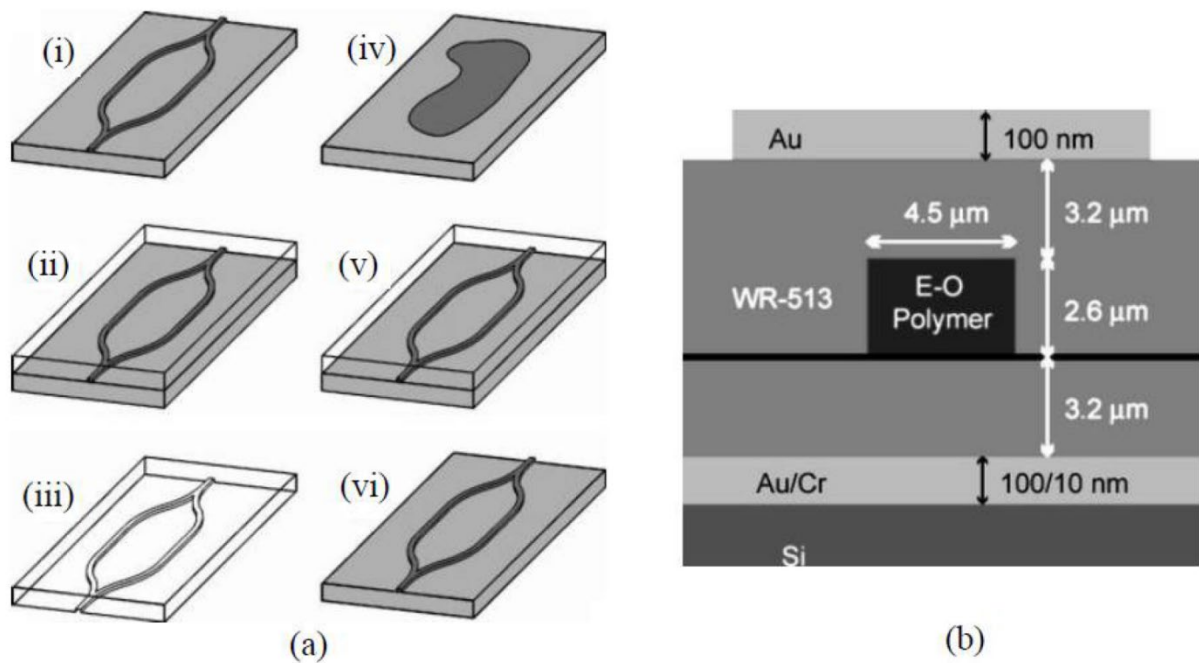
Fig. 10.10. a – fabrication of the PDMS mold (i–iii) and replication of the MZI structure (iv–vi). A master MZI device (i) is patterned using photolithography of SU-8. PDMS is poured on the master device (ii), cured, peeled, and diced (iii). A drop of electro-optic core polymer solution is placed on the silicon substrate (iv). The PDMS stamp is depressed (v) until the core polymer is cured. The stamp is peeled to reveal the replicated device (vi); *b* – cross-sectional view of the electro-optic waveguide structure with electrode

Although electro-optic polymers have been around for some decades, relatively few devices have been developed from proof-of-principle devices to real world applications.

Switches are very similar to modulators. They require at least two output ports and are optimized to have a high extinction ratio. Although extinction ratios of more than 20 dB have been measured without cascading devices, the realized values are still too low for most telecommunication system applications.

In addition to the standard application of polymeric electro-optic materials in modulators, some new devices: a photonically controlled RF phase shifter, a time stretcher for 10^2 GHz millimeter waves, and an integrated-optical spectrum analyzer also are considered.

11. GRAPHENE IN INTEGRATED OPTICS DEVICES

Graphene, a single layer of carbon atoms arranged in the form of hexagonal lattice, has many intriguing optical and electrical properties. However, due to the atomic layer thickness, light-matter interactions in the monolayer graphene are naturally weak when the light is normally incident to the material. To overcome this challenge, waveguide-integrated graphene optoelectronic devices have been proposed and demonstrated. In such coplanar configurations, the propagating light in the waveguide can significantly interact with the graphene layer integrated on the surface of the waveguide. The combination of photonic integrated circuits and graphene also enables the development of graphene devices by using silicon photonic technology,which greatly extends the scope of graphene's application. Moreover, the waveguide-integrated graphene devices are fully CMOS-compatible, which makes it possible to achieve low-cost and high-density integration in the future. As a result, the area has been attracting more and more attention in recent years.

Taking silicon modulators, for example, due to the limitation of carrier plasma dispersion effect in silicon, the theoretical upper limitation of the modulation bandwidth using *p-n* or *p-i-n* diodes may be $\sim$50 GHz, which cannot satisfy the growing demand in the modern information society. Graphene provides a promising solution to this issue. The carrier mobility of pristine grapheme can be as high as 200,000 cm^2/Vs which results in the maximum modulation bandwidth of 500 GHz. Besides, in waveguide configurations, light-graphene interactions are enhanced through evanescent field coupling, resulting in larger saturation energy and ultrafast optical pulses with higher energy.

11.1. Graphene-covered optical waveguides

Optical properties of graphene can be described by optical conductivity σ, calculated from Kubo formalism. The optical conductivity σ includes

contributions frominterband transition and intraband transition in graphene which can be expressed by

$$\sigma = -\frac{2ie^2(\Omega + 2i\Gamma)}{h}\left[\frac{1}{(\Omega + 2i\Gamma)^2}\right.$$

$$\cdot\int_\Delta^\infty \frac{\omega^2 + \Delta^2}{\omega}\left(\frac{\delta n_F(\omega)}{\delta\omega} - \frac{\delta n_F(-\omega)}{\delta\omega}\right)d\omega \qquad (11.1)$$

$$\left.-\int_\Delta^\infty \frac{\omega^2 + \Delta^2}{\omega}\left(\frac{n_F(-\omega) - n_F(\omega)}{(\Omega + 2i\Gamma)^2 - 4\omega^2}\right)d\omega\right],$$

where Ω is optical frequency, $n_F(\omega)$ is the Fermi distribution function, $\Gamma(\omega)$ is the scattering rate, ω is the energy of the relativistic Landau levels, and Δ is the exciton gap of Landau level energies.

By assuming a finite thickness to the graphene layer, the effective relative permittivity εeff of graphene can be calculated as follows

$$\varepsilon_{eff} = 1 + i\frac{\sigma}{\omega\varepsilon_0 d},$$

where d is the thickness of graphene layer. By using the calculated relative permittivity ε_{eff}, the effective complex refractive index (RI) of a graphene-covered waveguide can be numerically simulated such that we can design waveguide integrated graphene devices before fabrication.

Due to the different intensity distributions, graphene has distinct optical absorptions to TE and transverse magnetic (TM) modes propagating light. In this study, graphene introduces 7.7 dB higher losses for TM mode in a 150-μm-long waveguide.

The structure of graphene-on-silicon nitride microring resonator by tailoring the length of graphene integrated on the microring resonator to obtain

the maximum optical absorption, is shown in the schematic image of Figure 11.1, *a*.The variation in quality (*Q*) factor of the microring resonator as a function of the graphene length was measured experimentally, which agreed well with the theoretical calculation, as shown in Figure 11.1, *d*.
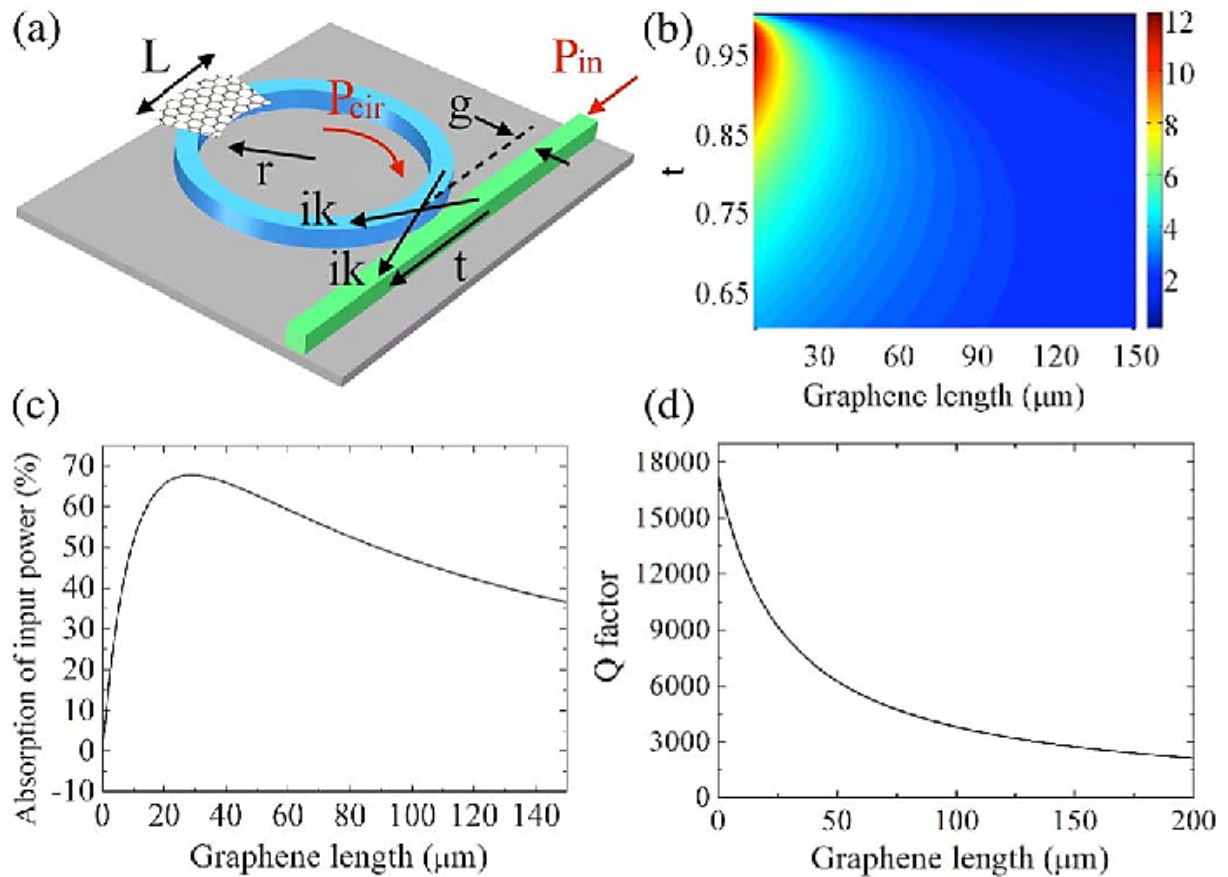

Fig. 11.1. Schematic illustration of the microring resonator for enhancement of graphene-on-Si$_3$N$_4$ waveguide absorption (*a*); modeling of *M* factor as a function of transmission coefficient (*t*) and graphene length (*L*) (*b*); calculated dependence of optical absorption on the graphene length (*c*); calculated influence of graphene length on the *Q* factor with *t* = 0.95 (*d*)

The original absorption of 25 dB/mm in a graphene-on-Si$_3$N$_4$ straight waveguide was significantly enhanced through the resonator structure, offering a quality factor from ~28200 to ~3800 depending on the graphene length in the cavity. Such resonators are useful in developing resonator structure for high-performance optical modulators and photodetectors.

11.2. Waveguide-integrated graphene modulators

Waveguide-integrated modulators with high speed, small footprint, and large spectral bandwidth are key components for applications of optical interconnects and optical communications, which can be classified as electro-refractive modulators or electro-absorptive modulators. In the electro-refractive modulators, an applied electrical field changes the real part of the effective RI of optical mode in the waveguide such that the phase of the propagating light is modulated. In the electro-absorptive modulators, an applied electrical field changes the imaginary part of the effective RI of optical mode in the waveguide, resulting in variations in optical loss. For traditional waveguide-integrated modulators based on silicon photonic technology, the most common method is based on the plasmon dispersion effect, in which some carriers are injected to silicon waveguide to alter the real and imaginary parts of the RI of waveguide mode. The intrinsic modulation bandwidth limit may be up to 50 GHz. Besides, due to the relatively weak light-matter interaction, usually mm-long interaction length is needed to achieve full transition between the maximum and minimum optical transmission. On the other hand, graphene has relatively low carrier density, which makes it possible to tune its permittivity by applying an external electrical field. When graphene is integrated on a waveguide, the changes of the complex effective RI of the waveguide mode can be achieved through electrically tuning the Fermi level of graphene, enabling electro-refractive modulators or electro-absorptive modulators. What is more, by optimizing the waveguide structures, an extremely strong light-graphene interaction may be obtained, leading to a significant reduction in device footprint. In this section, we summarize the recent progress on waveguide-integrated graphene modulators.

The first waveguide-integrated graphene modulator was demonstrated in 2011, in which electroabsorptive modulation was achieved by tuning the Fermi level of graphene through electro-gating from the lightly doped SOI waveguide.

The device had a bandwidth of 1.2 GHz. The modulator subsequently demonstrated the structure of dual-layer graphene sandwiching by a thin layer for the waveguide-integrated modulator, in which electrons and holes were injected into the graphene layers to form a p-oxide-n junction. The schematic of the cross section of dual-layer graphene modulator is shown in Figure 11.2. At zero drive voltage, the Fermi levels of both graphene layers were close to Dirac point such that the interband transition led to large optical absorption. At high drive voltage, the two layers of graphene were symmetrically doped with electrons and holes, respectively, shifting the Fermi levels beyond the half of the photon energy $|\hbar\omega/2|$, resulting in optical transparence of graphene. A 3-dB bandwidth of 1 GHz was experimentally achieved in this device.
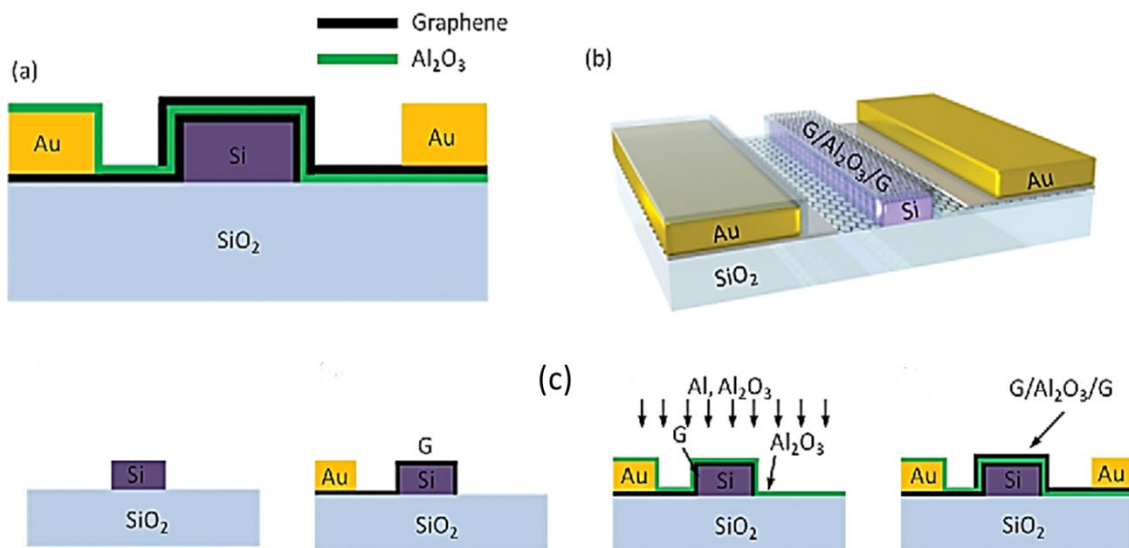

Fig. 11.2. Schematic of the dual-layer graphene modulator:
a, b – the final structure of the device; *c* – schematic illustration of the fabrication process

The configuration of dual-layer graphene has also been further explored with interferometric and resonant waveguide structures. Applied the duallayer graphene structure in an unbalanced Mach-Zehnder interferometer (MZI) allowed to achieve simultaneous optical modulation and detection, as shown in Figure 11.3. The graphene device is integrated on a silicon waveguide and is tunable with a gate made from another layer of graphene to achieve near-

infrared photodetection responsivity of 57 mA/W and modulation depth of 64%. The functionality of the device can be controlled with an integrated electrostatic gate also made of graphene. Fig. 11.3, *a* illustrates the configuration of the device which consists of two layers of graphene, separated by a dielectric layer, and integrated on a planarized silicon photonic waveguide. The device is in the configuration of a simple field effect transistor (FET): the bottom layer (the channel) acts as an optical absorber and can collect photo-generated carriers while the top layer acts as a transparent gate electrode which can tune the electrical and optical properties of the bottom graphene layer.
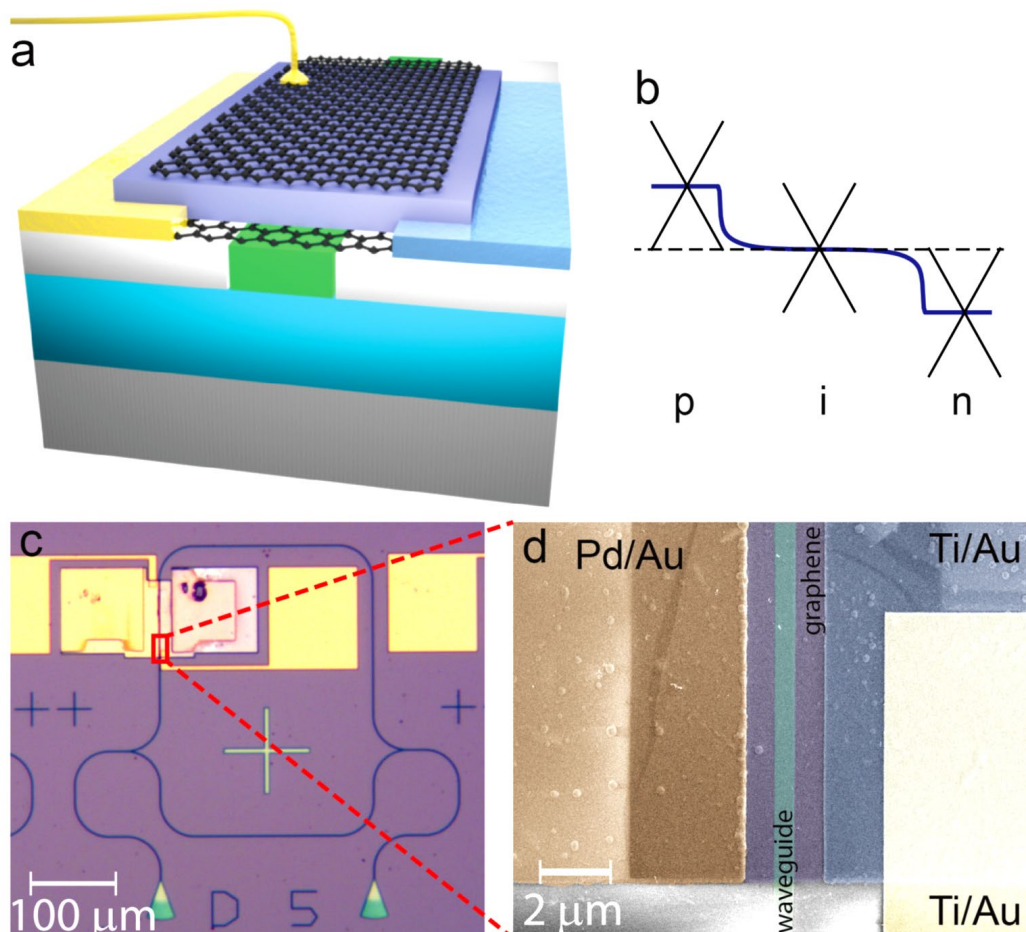

Fig. 11.3. Waveguide integrated graphene modulator and photodetector: *a* – schematic illustration of the dual layer graphene modulator/detector integrated on a planarized waveguide; *b* – illustration of the profile across the graphene channel highlighting the *p-i-n* junction and the build-in electric field; *c* – optical microscope image of the device; *d* – scanning electron microscope image of the device

Figs. 11.3, *c* and *d* show the optical and scanning electron microscope images of a typical device. To most accurately measure the net optical absorption of graphene in order to determine the performance, the device is embedded in one arm of an unbalanced Mach-Zehnder interferometer. Two asymmetric source and drain contacts made of titanium/gold (light blue) and palladium/gold (dark yellow) dope the graphene (light purple) to be n- and p-type, respectively. Contact (light yellow) to the top graphene gate is made of titanium/gold. The waveguide is in light green (Fig. 11.3, *d*).

A ring modulator based on few-layer graphene with large tunability and modulation efficiency is shown in Fig. 11.4. Compared with monolayer graphene, a few-layer graphene structure can enhance the tunability linearly. In the proposed structure, the shift of resonance wavelength can be enlarged to 20.3 nm, implying the proposed modulator can achieve the bandwidth of 100 GHz. In addition, a large extinction ratio of 15.5 dB and a good temperature tolerance of 45 K can be obtained in this modulator.
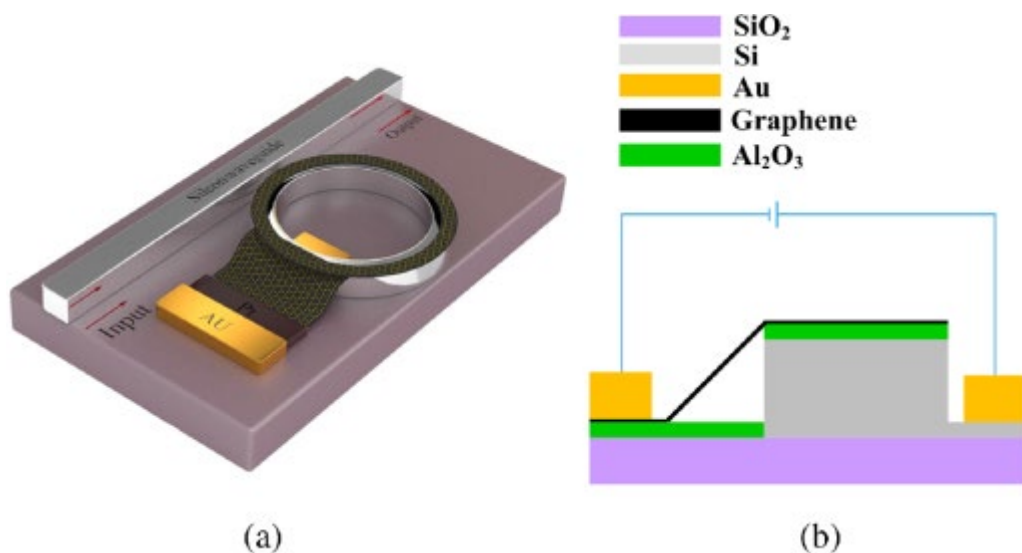

Fig. 11.4. A Ring modulator based on few-layer grapheme:

a – schematic of the structure; *b* – cros-section of the modulator

11.3. Mach-Zehnder (M-Z) graphene modulator

The M-Z modulator that based on a hybrid graphene silicon waveguide (GSW) that was ultra-fast and compatible with CMOS is shown in Figure 11.5, *a*. As shown in Fig. 11.5, *a*, GSW is made on a rectangular silica substrate, and its two arms are separated from each other at a fixed angle. Here, a lateral slot waveguide is formed by interposing Si_3N_4 inside the silicon. Three layers of graphene sheets were placed in Si_3N_4, and was chosen over single-layer graphene, because the effective they were separated from each other by 10 nm, like a sandwich. Three-layer graphene was chosen over single-layer graphene, because the effective index variation of the GSW can be larger. The metal electrode contacts the graphene sheet to form a capacitive structure between the second graphene layer and the first and third graphene layers. The graphene sheet placed in Si_3N_4 has a greater effect on the GSW effective index (n_{eff}) than the graphene sheet placed on top of the silicon waveguide, owing to most light energy is confined within the gap.
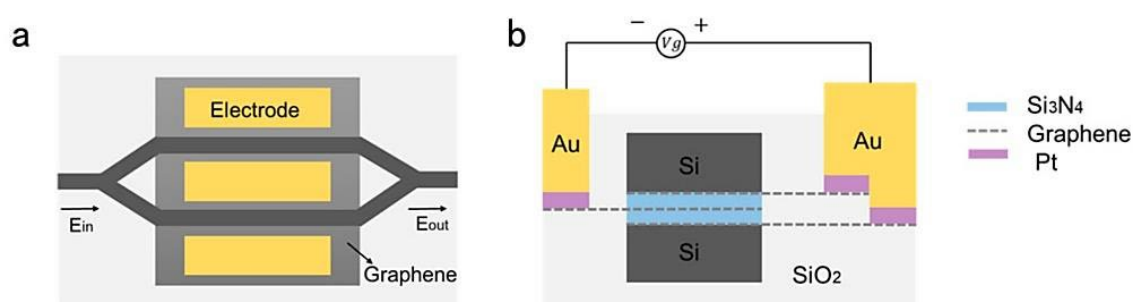

Fig. 11.5. Two-dimensional (2D) vertical view of the modulator (*a*) and 2D cross section of the modulator's arm (*b*). One metal electrode is brought into contact with the first and third graphene layers Another metal electrode is brought into contact with the second graphene layer. $n_{Si} = 3.47$, $n_{SiO2} = 1.447$, $n_{Si3N4} = 1.98$

When compared with the n_{eff} provided by the traditional material Kerr effect, the TM mode in the GSW has a larger n_{eff} change, that is, it has a strong phase modulation capability. In addition, the GSW loss and the real and

imaginary parts of n_{eff} change significantly around the epsilon near zero point For M-Z modulators, the phase of light on both arms can be uniformly moved in different directions simultaneously. At the same time, with the change of light phase, the loss difference between the two arms is still small. The phase change between the output signal and the input signal can be eliminated by using this characteristic. The modulator has excellent thermal stability at room temperature, but better performance at lower temperatures. When the environment becomes cold, GSW can obtain a large phase modulation capability, while the loss is still low. This can significantly reduce the insertion loss of the modulator. In addition, it is possible for this TM modulator to operate at ultra-high speed for ultra-high carrier mobility in graphene.

A polarization modulator that is based on graphene-coated interferometers that is independent of the polarization of incident light is shown in Fig. 11.6. By applying different bias voltages on the graphene layer, it is possible to achieve different changes in the effective index (n_{eff}) of the TE polarization and TM polarization modes. The M-Z interferometer can be used to adjust the light transmission. The device can be used as either a TE polarizer or a TM polarizer, depending on the external voltage applied to the graphene. Figure 11.6 shows the schematic of the modulator. The entire device is constructed on a silicon dioxide substrate, and the M-Z interferometer is made of a silicon waveguide coated with a graphene layer. The cross section of a silicon waveguide is a square with a side length of 300 nm (Fig. 11.6, *b*). The process of directional growth or bonding techniques constructed the silicon waveguide, and the top alumina transition layer and graphene can be placed on the top surface of the doped silicon waveguide.

Finally, in the same way, transitional alumina and graphene layers can be deposited on the sides of the waveguide. In addition, the graphene layers are separated from each other and extended to connect with the electrodes to provide an external oxide of graphene. The device provides a compact size of 70 μm × 2 μm, a high extinction ratio of 19 dB for the TE-pass state and of

21 dB for the TM-pass state. Furthermore, this device exhibits a wide working bandwidth from 1500 to 1800 nm.
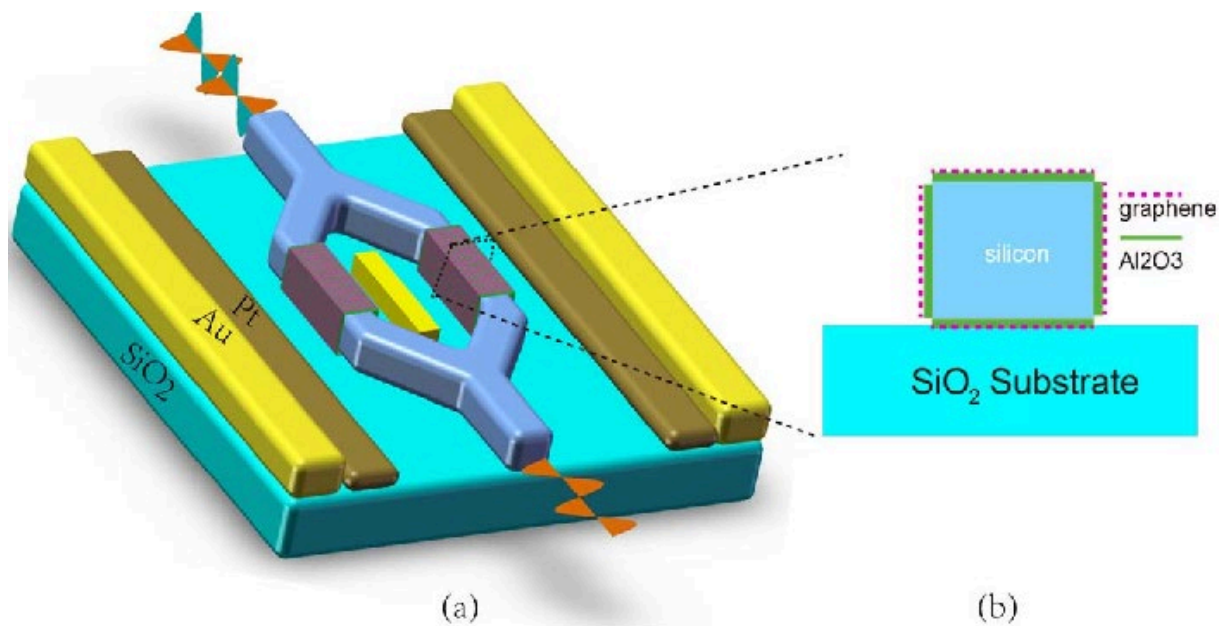

Fig. 11.6. The 3D view of the proposed polarizer (*a*); the schematic view of the proposed device (*b*)

11.4. Waveguide-integrated graphene photodetectors

With the features of zero bandgap and high electronmobility, graphene photodetectors are promising for broadband photodetection with ultrahigh bandwidth, which is only limited by the resistor-capacitor (RC) time constant of electrodes. Integration of graphene on an optical waveguide can dramatically increase the optical absorption through the in-plane light coupling, the detection efficiency of the grapheme photodetector therefore can be increased.

By integrating graphene on SOI waveguides, grapheme photodetectors with high responsivity and high bandwidth were demonstrated in telecommunication bands and midinfrared spectral region for the first time in 2013. The structure fabricated the lateral graphene-metal junction overlapped with the waveguide mode to form the built-in electrical field to separate

photogenerated electron-hole pairs. The device had a responsivity of 0.1 A/W with bandwidth of 20 GHz.

The silicon suspended membrane waveguide (SMW) can eliminate the absorption of buried oxide in the mid infrared spectral range and take full advantage of the silicon transparent window of 1–8 μm. So, such a graphene-on-silicon photodetectors can be operated in the mid-infrared spectral region. Moreover, the graphene/silicon heterostructure was employed in the work which reduced the dark current and achieving a high ON/OFF current ratio in a broadband spectral range.

Researchers have tried to use CVD growth graphene to develop high-performance waveguide integrated graphene photodetectors. The device had a 3-dB bandwidth of 41 GHz. The graphene/metal contacting with the silicon waveguide form a Schottky diode and explored the internal photon emission process, as shown in Figure 11.7. The maximum responsivity of 0.37 A/W and avalanche photo-gain of ~2 were achieved.
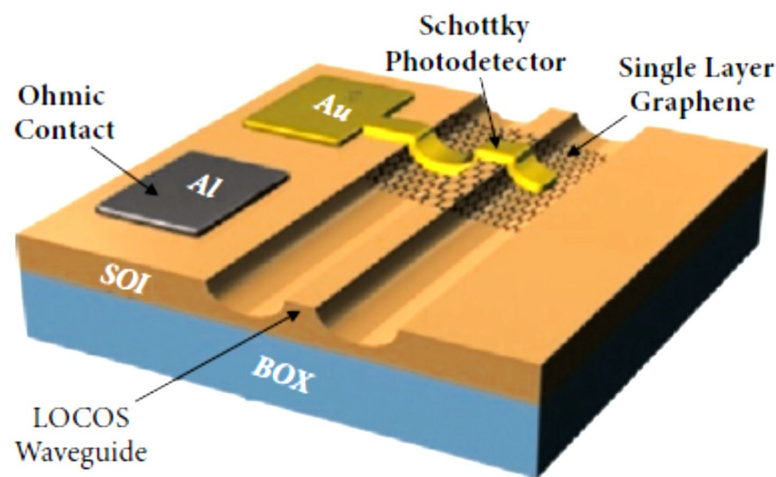

Fig. 11.7. Schematic of the graphene Schottky diode integrated on the SOI waveguide

Besides channel and rib waveguides numerous novel waveguide structures have exploited to develop high-performance graphene photodetectors.

Figure 11.8 demonstrated the high-responsivity graphene photodetector integrated on the silicon slot waveguide.
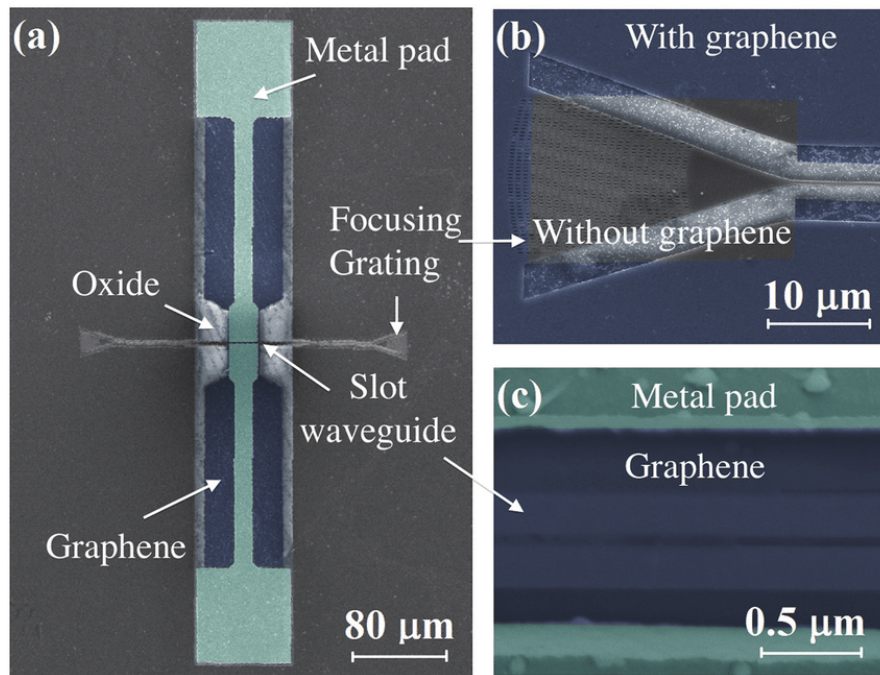
Fig. 11.8. SEM images of fabricated devices: *a* – SEM image of the graphene-on-silicon slot waveguide photodetector; *b* – SEM image of the subwavelength grating; *c* – SEM image of the metal pads set along the silicon slot waveguide

Compared to the channel/slab waveguide, the slot waveguide is able to confine and guide a higher intensity of light in the slot region due to the discontinuity of the electric field at the high- refractive index contrast interface. Thus, graphene will have a much stronger interaction with the propagating light after being integrated into the slot waveguide configuration. Moreover, there are some advantages for the graphene-on-silicon slot waveguide structure. First, the enhanced absorption in the graphene-on silicon slot waveguide can greatly reduce the device footprint,which will eliminate the requirement for the synthesis of large-area high-quality graphene films. Second, such a configuration is very suitable for investigating the nonlinear properties of graphene in the telecommunication band. It provides reinforced optical interactions as well as avoids the limitation imposed by two-photon absorption generated free carrier absorption loss in the silicon waveguide.Finally, small size devices also increase the integration density and reduce the power consumption for photonics integrated circuits.

Due to the increase in light intensity and decrease of optical mode confinement in the silicon slot, the graphene experienced an increased interaction to the in-plane light. A waveguide absorption of 0.935 dB μm^{-1} was measured at 1.55 μm wavelengths. Benefited from the lack of efficient electron cooling in the suspended graphene and the intensity enhancement effect in the nano slot, a maximum responsivity of 0.273 A W^{-1} was achieved in the telecommunication band.

Apart from graphene photodetectors integrated on the SOI platform, researchers have developed graphene photodetectors based on other integration platforms. For example, silicon nitride is a large bandgap (~5 eV) deposited material that is compatible with CMOS fabrication technology.

By using silicon nitride with the small relative permittivity the resistance–capacitance-limited bandwidth can be improved. A chemical vapor deposited graphene photodetector integrated on a silicon nitride waveguide is shown in Fig. 11.9. Interdigital metal contacts are used to reduce the channel spacing down to ~200 nm. Benefiting from the narrow channel spacing, a calculated transit-time-limited bandwidth of 111 GHz is derived.
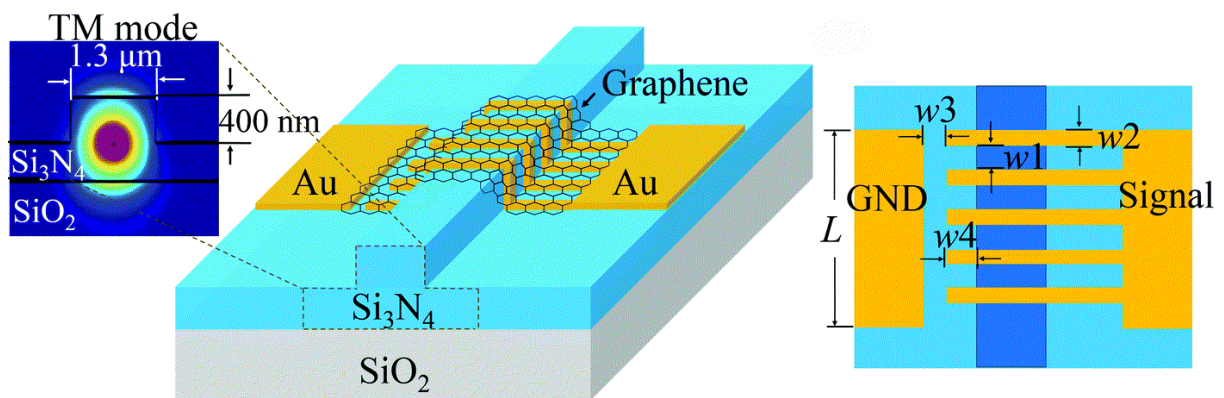

Fig. 11.9. Schematic of the chemical vapor deposited graphene photodetector integrated on a silicon nitride waveguide

At zero bias, a metal-graphene junction is used for photodetection, which is beneficial for an electro-optic bandwidth of ~33 GHz in the 1550 nm wavelength band. At a bias of 1 V, a photoconductive responsivity of

~2.36 A/W at 1550 nm was observed. The high speed and high responsivity make the device promising for photodetection in the telecommunication C-band. A diffusion model is applied to study the carrier transition process in the graphene channel. By adopting this model, the high performance of the device can be explained.

With the rapid development in graphene synthesis technologies, direct deposition of graphene on various types of substrates, especially silicon, will further boost advances in high-speed graphene optoelectronics and silicon photonics.

In conclusion, with the unique optical and electrical properties, graphene is expected to improve the performance of conventional PICs and bring new applications.

FURTHER READING

1. Haus H. A. Waves and Fields in Optoelectronics / H. A. Haus. – Englewood Cliffs, NJ : Prentice-Hall, 1984. – P. 456.

2. Introduction to Integrated Optics / Ed. M. Barnoski. – Springer, 2012. – P. 515.

3. Pollock C. Integrated Photonics / C. Pollock, M. Lipson. – Springer US, 2003. – P. 376.

4. Silicon Photonics II. Components and Integration / Eds: D. J. Lockwood, L. Pavesi. – Springer-Verlag Berlin Heidelberg, 2011. – P. 256.

5. Wasley N. A. Nano-photonics in III-V Semiconductors for Integrated Quantum Optical Circuits / N. A. Wasley. – Springer International Publishing, 2014. – P. 129.

6. Mitschke F. Fiber Optics. Physics and Technology / F. Mitschke. – Springer-Verlag Berlin Heidelberg, 2016. – P. 349.

7. Hunsperger R. Integrated Optics. Theory and Technology / R. Hunsperger. – Springer-Verlag Berlin Heidelberg, 2002. – P. 446.

8. Chang W. S. C. Fundamentals of Guided-Wave Optoelectronic Devices / W. S. C. Chang. – San Diego : Cambridge University Press, 2009. – P. 420.

9. Murphy D. P. Integrated Optical Circuits and Components: Design and Applications / D. P. Murphy. – CRC Press, 2019. – P. 468.

10. Alferness R. C. Guided-Wave Optoelectronics / R. C. Alferness, W. K. Burns, J. F. Donelly. – Berlin : Springer, 1988. – P. 401.

11. Fiber and Integrated Optics / Ed. D. B. Ostrowsky. – Springer US, 1979. – P. 417.

12. Buckman A. B. Guided-Wave Photonics / A. B. Buckman. – Saunders College Publishing, 1995. – P. 384.

13. Calvo M. L. Optical Waveguides : From Theory to Applied Technologies / Eds. : M. L. Calvo, V. Lakshminarayanan. – CRC Press, 2019. – P. 424.

14. Nishihara H. Optical Integrated Circuits / H. Nishihara, T. Suhara, M. Haruna. – McGraw-Hill Professional, 1989. – P. 454.

15. Encyclopedic Handbook of Integrated Optics / Eds. : K. Iga, Y. Kokubun. – CRC Press, 2005. – P. 528.

16. Gao J. Optoelectronic Integrated Circuit Design and Device Modeling / J. Gao. – John Wiley & Sons Inc., 2011. – P. 320.

17. Свечников Г. С. Элементы интегральной оптики / Г. С. Свечников. – М. : Радио и связь, 1987. – 105 с.

18. Свечников Г. С. Интегральная оптика / Г. С. Свечников. – К. : Наукова думка, 1988. – 168 с.

19. Тамир Т. Интегральная оптика / Т. Тамир. – М. : Мир, 1978. – 344 с.

20. Хансперджер Р. Интегральная оптика: теория и технология / Р. Хансперджер. – М. : Мир, 1985. – 384 с.

21. Гончаренко А. М. Основы теории оптических волноводов / А. М. Гончаренко, В. А. Карпенко. – М. : Едиториал УРСС, 2004. – 240 с.

22. Снайдер А. Теория оптических волноводов / А. Снайдер, Дж. Лав. – М. : Радио и связь, 1987. – 656 с.

23. Семенов А. С. Интегральная оптика для систем передачи и обработки информации / А. С. Семенов, В. Л. Смирнов, А. В. Шмалько. – М. : Радио и связь, 1990. – 224 с.