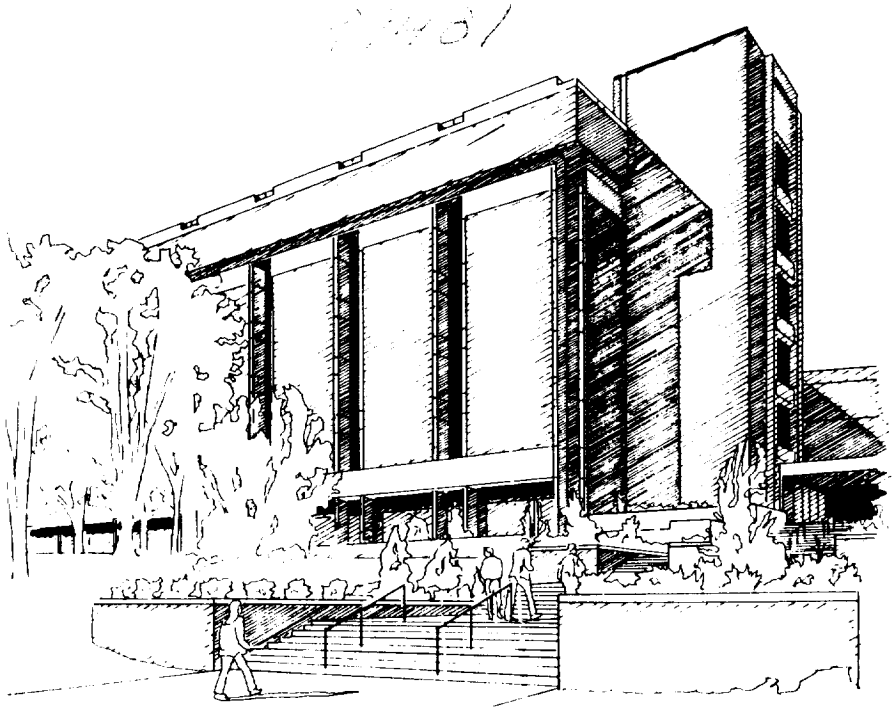
NASA-CR-202535

NAG3-852



UNIVERSITY OF CINCINNATI
COLLEGE OF ENGINEERING

Integrated Optical Interferometers with Micromachined Diaphragms for Pressure Sensing

Gregory N. De Brabander

Department of
Electrical & Computer Engineering
and Computer Science

ABSTRACT

Optical pressure sensors have been fabricated which use an integrated optical channel waveguide that is part of an interferometer to measure the pressure-induced strain in a micromachined silicon diaphragm. A silicon substrate is etched from the back of the wafer leaving a rectangular diaphragm. On the opposite side of the wafer, ring resonator and Mach-Zehnder interferometers are formed with optical channel waveguides made from a low pressure chemical vapor deposited film of silicon oxynitride. The interferometer's phase is altered by pressure-induced stress in a channel segment positioned over the long edge of the diaphragm. The phase change in the ring resonator is monitored using a link-insensitive swept frequency laser diode, while in the Mach-Zehnder it is determined using a broad band super luminescent diode with subsequent wavelength separation. The ring resonator was found to be highly temperature sensitive, while the Mach-Zehnder, which had a smaller optical path length difference, was proportionally less so. The quasi-TM mode was more sensitive to pressure, in accord with calculations. Waveguide and sensor theory, sensitivity calculations, a fabrication sequence, and experimental results are presented.

ACKNOWLEDGMENTS

No project such as this is performed without the involvement and stimulation of others. I wish to thank my advisor J. T. Boyd for the freedom to pursue tangential topics of interest and fellow lab mates Dave, Jie, Kevin, and Mukesh for numerous consultations and discussions. An appreciation is extended to Glenn Beheim of NASA Lewis for equipment loans and technical help, and to his organization as well as the NASA Space Engineering Research Center for funding the project. Professor Henderson's equipment was heavily used in this project. The nurturing and support of my parents were invaluable.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I Introduction and Literature Review of Optical Sensors

1 INTRODUCTION

Optic sensors have been used to measure pressure, temperature, flow, the presence of gases, acceleration, rotation, and electric and magnetic fields. They are used because (1) they can offer superior performance, such as for gyroscopes, (2) when measurements are best made by optics, such as for determining temperature by remote pyrometry, or (3) when sensors are needed in electrically hostile or noisy environments. In the first two cases, electrical wired connections may accompany or be packaged with the fiber. For the last case, however, the sensor head contains no electronics (passive) and is fed only by fiber optic cables from an electronic package located in an electrically benign environment (Fig. 1-1). Such sensors are immune to and do not cause electromagnetic interference, are nonsparking, and have electrical isolation between the sensor head and electronics.

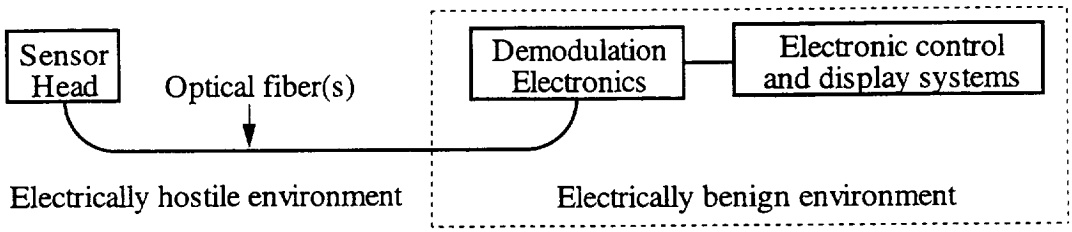


Fig. 1-1 All-optic sensor configuration.

This thesis is concerned with the design, fabrication, and testing of an integrated optic pressure sensor suited for category (3) above. The device is constructed using semiconductor and micromachining fabrication techniques. Pictured in Fig. 1-2, it consists of an integrated optic interferometer having one arm positioned over a micromachined diaphragm. Light will be provided through single mode optical fibers.

No electrical connections are made to the chip. Fluid pressure applied to the diaphragm alters the optical path length of a channel waveguide passing over it, changing the phase shift of light in the channel. The interferometer converts this into a change of the ratio of light leaving the chip to that which entered (transmissivity change). Various methods can be used to measure this change.

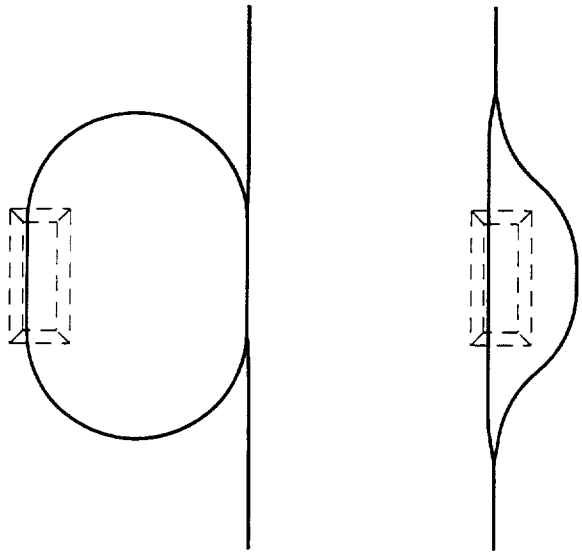


Fig. 1-2 Top view of pressure sensor configurations in which an integrated optic interferometer is used to detect pressure applied to a micromechanical diaphragm.

With conventional electronic micromachined sensors, pressure is measured by a resistance change in a strain-sensitive resistor located on the edge of a diaphragm or as a capacitance change between electrodes whose spacing is altered by diaphragm deflection. Our all optic sensor is similar to the first type of device, but a waveguide measures diaphragm strain instead of a piezoresistor. For electronic sensors, value changes (resistance, capacitance) in the sensing components are monotonic with pressure. In addition, on-board compensation, linearization, and signal processing may enhance ease of use. All-optic sensors use interferometry, which can be nonlinear and nonmonotonic, or intensity modulation, which may be affected by fiber transmission. In addition, the

limited functionality of optic devices and the passivity requirement precludes on-board signal processing.

Table 1-1 lists the advantages of all optic and electronic pressure sensors. For most applications, electronic sensors are preferred, as demonstrated by their greater usage. However, where electromagnetic immunity is mandatory, users can either shield an electronic sensor or employ an all-optical sensor. In contrast to the lightweight optical fiber, the shielded cable to an electronic sensor head is potentially heavier and more bulky. Weight considerations in applications such as aircraft make all optic sensors attractive.

Optic Pressure Sensors	Electronic Pressure Sensors
Electromagnetic interference immune (head and fiber link only)	Less expensive
Does not generate electromagnetic interference	More familiar, especially in connectorization
May be multiplexed	Standard interfaces available (4-20 mA, 0-10 VDC)
Non-sparking	Widely available
	Signal conditioning at the head possible

Table 1-1 Advantages of optic and electronic technologies. The converse of the advantages of one are the disadvantages of the other.

2 DESIGN REQUIREMENTS

Passive optic sensors suited for field use must meet a common set of requirements.

- Link independence: Insensitivity to variation in transmission through the fibers. A major problem faced when designing a fiber optic sensor is that, unlike electrical wires, fiber optic transmission can vary by 10% with large temperature changes and remates of optical fiber connectors [1]. Since constant fiber transmission cannot be relied upon, researchers have been driven to pursue more complicated sensor schemes which do not count on constant link transmissivity. Such schemes may also make the sensor independent of fluctuations in source brightness.

- Temperature insensitivity: The accuracy of the sensor will be no better than the uncorrectable drift in scale factor and zero experienced in the sensor environment.
- No ambiguity problem: When interrogated by a fixed wavelength source, an interferometric sensor's output is periodic with measurand (Fig. 1-3). Movement of the phase through a minima or maxima creates an ambiguity. One solution is to restrict the range. Other techniques, such as sweeping the source wavelength, monitoring transmittance at multiple wavelengths, or designing the device to produce simultaneous cosine-sine outputs can allow phase tracking over more than π radians phase change. Other measurement applications which use interferometry, such as acoustic sensing, do not require DC information as does pressure sensing. In these applications low frequency thermal drifts move the interferometer to an extremum where the slope of the transmittance function (see Fig. 1-3) and thus small signal sensitivity go to zero. Such an event is called signal fading. Active biasing, not permitted in our passive device at the head, although limited in range, can be used to maintain the interferometer at the point of maximum sensitivity.
- Existence of zero: Phase tracking techniques can permit following an interferometer's phase change over an extremum. However, when the source is first activated, and after a power interruption, the measurand may have to be set to a known state to permit determination of the correct fringe order. In such a case, the sensor is said to lack a zero. Once the sensor has been initialized, phase tracking can enable determination of the measurand wherever it may venture. It is desirable that the sensor possess a zero to eliminate initialization procedure.

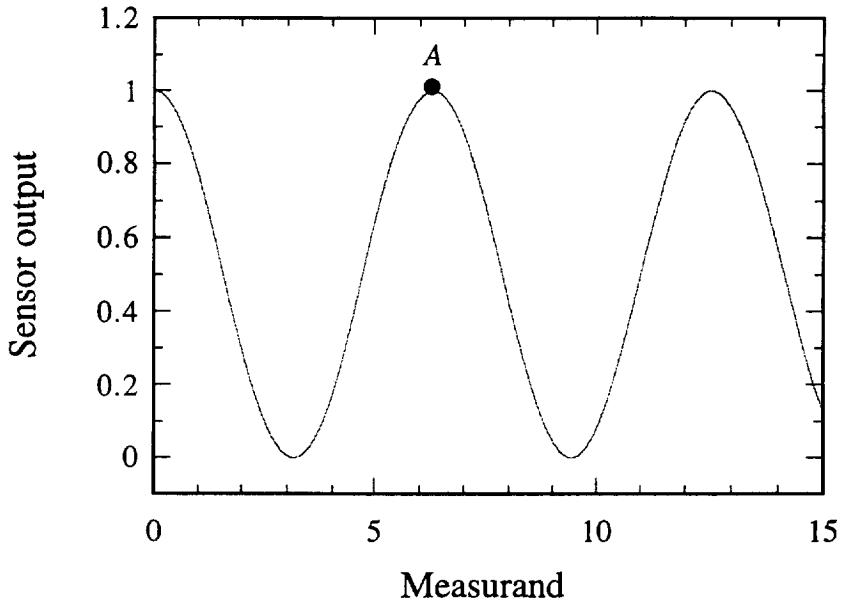


Fig. 1-3 The ambiguity problem: If sensor output is at *A* and falls, has the measurand increased or decreased? Existence of zero problem: If the measurand could lie between 0 and 15, and at turn-on the sensor output is 0.5, what is the measurand? Link independence problem: If the DC signal from the output fiber is 0.5 and the measurand lies between 4 and 6 but the fiber transmissivity could have changed since calibration, what is the measurand?

Several other items, although of lesser importance are also worthy of consideration:

- Does the source have to be temperature stabilized to keep its wavelength from drifting?
- What kind of fibers link the sensor head, and how many?
- How complex is extracting (demodulating) the measurement from the sensor output?
- What is the size of the sensor head?
- Can the head or demodulation electronics be replaced without recalibrating?
- What is the sensor hysteresis and creep?

Because optic sensors serve in applications where electronic sensors are unsatisfactory, optical sensors do not have to be less expensive that their electronic

counterparts to be practical. Configurations have been proposed which use thousands of dollars of equipment to support the sensor.

3 REVIEW OF PASSIVE OPTIC SENSORS

A distinction should be made between a transduction technique and a sensor. Some papers propose transduction technologies that, as demonstrated, do not meet the previous requirements but might if a different interrogation technique was used. For example, consider the sensor pictured in Fig. 1-4. Pressure applied to a diaphragm raises a knife edge into a light beam coupling from one fiber to another and reducing the transmission between them. This sensor would be interrogated by launching light from a laser diode or LED into one fiber and measuring the light power leaving the other fiber. This sensor is not link insensitive; changes in light leaving the output fiber due to connector or fiber transmission changes will be indistinguishable from coupling changes between the two fibers due to diaphragm deflection. Suppose the knife edge in Fig. 1-4 is replaced by a two-segment filter in Fig. 1-5. The filter would be positioned so that over the entire range of diaphragm deflection, some light from each filter would couple into the output fiber. Movement of the filter pair would alter the ratio of λ_1/λ_2 emerging from the output fiber, while link transmissivity changes which affect both wavelengths approximately equally would not influence λ_1/λ_2, making the sensor more link insensitive (Fig. 1-6). If a multimode feed-fiber is used, the sensor would still be sensitive to changes in the spatial distribution of light (modal noise), which can affect the ratio of light passing through the filters. Clearly, both the interrogation scheme and the sensor design need be considered together in assessing the merits of a sensor.

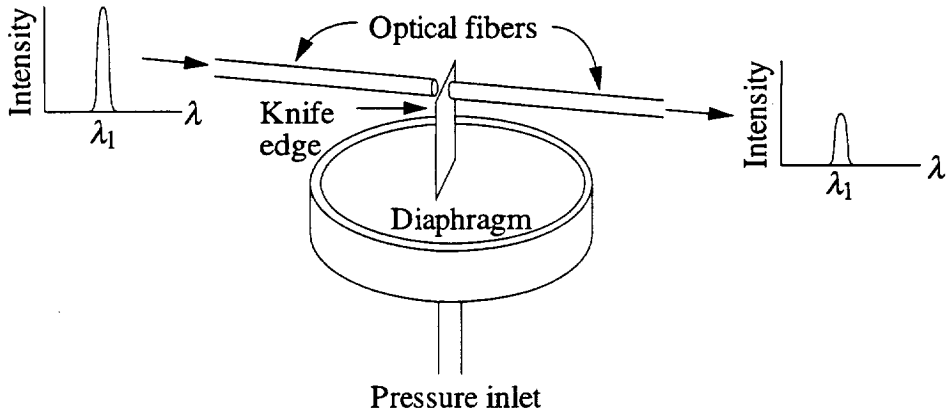


Fig. 1-4 Link sensitive example of a pressure sensor interrogated by measuring light transmitted between two fibers.

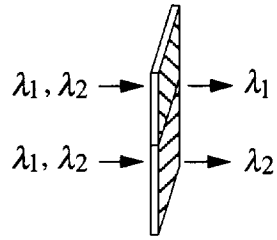


Fig. 1-5 Two-segment filter to replace the knife edge of Fig. 1-4.

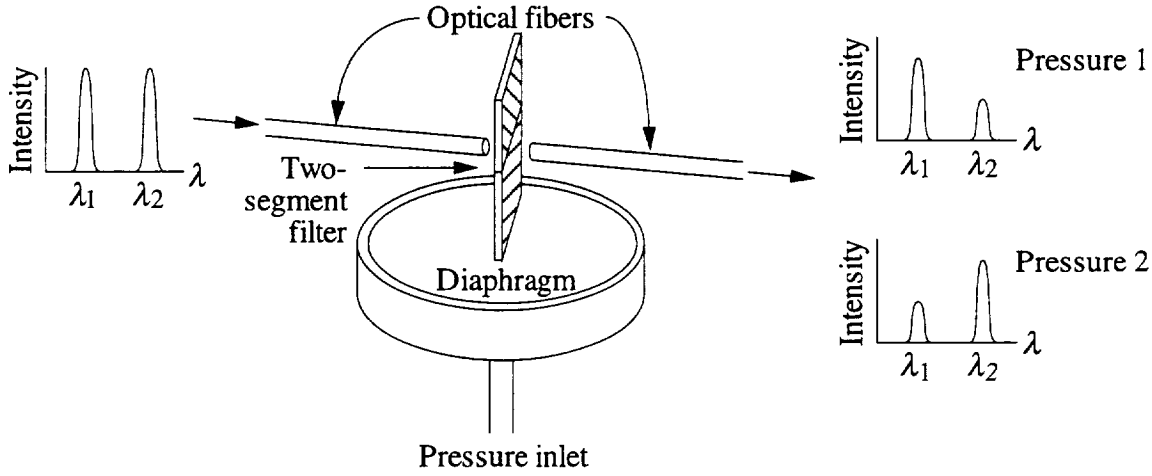


Fig. 1-6 The sensor of Fig. 1-4 has been made more link insensitive by changing to a more complex interrogation scheme. Light is launched by two sources (or one broad source) and the transmitted intensity ratio is a function of the measurand. Disadvantages with this sensor include the need to measure the launched λ_1/λ_2 ratio (or keep it constant), larger diaphragm mass, sensitivity to vibration, and sensitivity to modal noise.

The next sections review various passive optical techniques proposed in the literature for measuring pressure. Examples of various measurement schemes will be presented.

Vibrating element

These sensors consist of an element that is made to vibrate by periodic application of heat produced from optical fiber delivered light. Surprisingly, the vibration has been induced by sources as weak as LED's. The greater the Q of the resonator, the less power necessary to produce the same vibration [2], and the easier it is to identify the resonant frequency [3]. The measurand effects the tension on the element which changes the resonant frequency of vibration. The vibration is detected by measuring the change in reflected light from a second constant intensity optical beam delivered via the same cable. These sensors output a sinusoidally varying signal whose frequency is largely independent of fiber transmission. Micromachining is often used to make these sensors since it can produce the small resonators needed with ease.

Fig. 1-7 shows the layout of a micromachined vibrating resonator pressure sensor [4]. The bridge, 577 $\mu m \times 6\ \mu m \times 2.8\ \mu m$, consists of a layer of plasma silicon nitride coated with 30 nm of chromium. Pressure applied to the diaphragm moves the pillars up, and because of diaphragm curvature, the tips of the pillars move away from each other. This increases the tensile stress in the bridge and its resonant frequency.

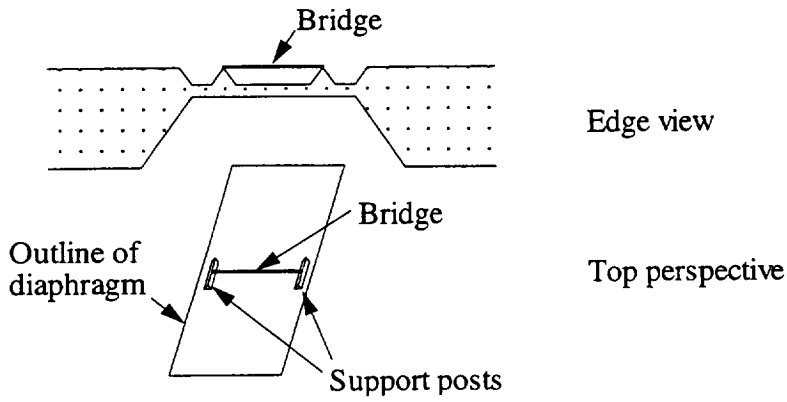


Fig. 1-7 Layout of the vibrating bridge pressure sensor [4].

To interrogate the structure, the authors used a bulk optic Twyman-Green interferometer (a modified Michelson interferometer), pictured in Fig. 1-8. The 830 nm laser diode is modulated while the 780 nm laser diode is maintained on continuously. Light from the 780 nm diode, after reflecting from the bridge, interferes with that reflected off the beam splitter and mirror, combining at the photodiode. A filter rejects the 830 nm excitation light. Temperature can shift the difference in distance between the bridge and beam splitter and mirror and beam splitter causing signal fading. A three axis stage can be adjusted to compensate. Because temperature was found to shift the bridge resonant frequency, a second bridge, sensitive only to temperature but too far away to interrogate with the same fiber, was included off the diaphragm.

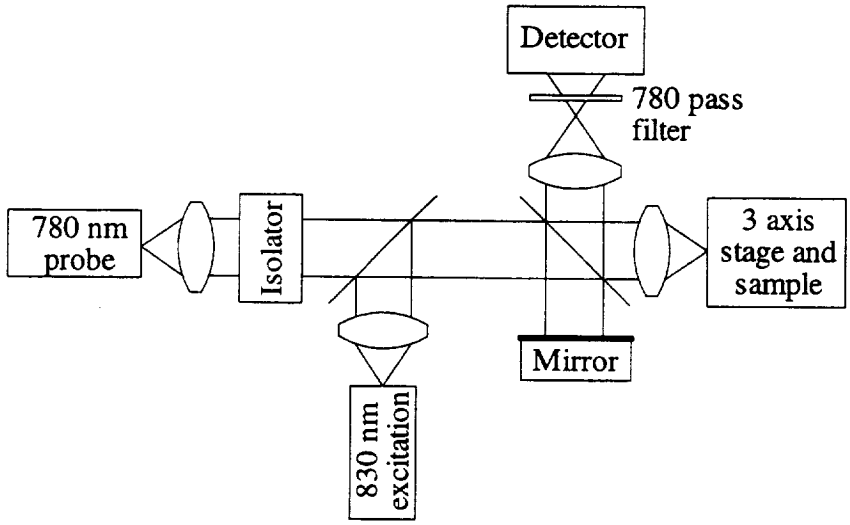


Fig. 1-8 Twyman-Green interferometer for excitation and readout [4].

Angelidis and Parsons have described a similar sensor [2] containing a bridge and a beam. The bridge was affected by pressure and temperature, while the beam, because its stress was unaffected by diaphragm distortion, measured only temperature and could be used to compensate the pressure sensor. The beam and bridge were fabricated in a silicon wafer fusion bonded to a diaphragm made from the same (Fig. 1-9). The authors intended that the sensor operate to 450°C. The beam/bridge wafer was then thinned to 5 μm and enclosed in an evacuated package (to raise the Q) that also held the optical fiber, allowing both sensors to be simultaneously interrogated with the same fiber. A laser diode excited the resonators, and their deflection was measured with an LED, whose broad bandwidth allowed a signal-fade-free interferometric measurement to be made. The sensors could be distinguished because their resonant frequencies were different. The resonant frequency of the bridge changed by 0.12 kHz/kPa for a 0-130 kPa range sensor and had a zero pressure frequency of 119 kHz, while the beam frequency changed by -0.0081 kHz/°C and had a room temperature value of 200.3 kHz. The temperature sensitivity of the bridge was not stated.

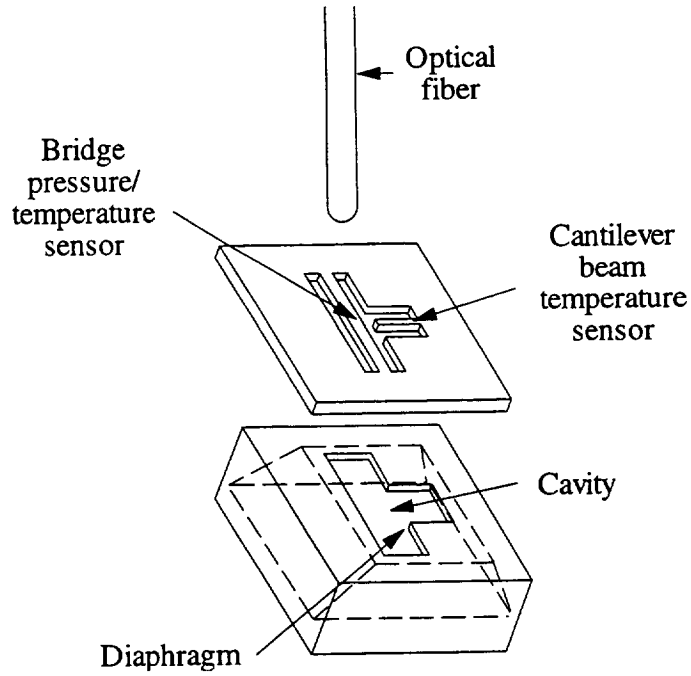


Fig. 1-9 Temperature compensated resonant beam pressure sensor [2].

Alternatively, the bridge may be dispensed with, and the diaphragm vibrated directly by a modulated semiconductor laser. A group that did this with a 750 μm × 750 μm × 5 μm diaphragm formed in a silicon wafer found that the resonant frequency changed from 121.5 to 162.9 kHz when the pressure varied from 0 to 75 kPa [5]. They measured a relatively low Q of 40, however. Another design consisted of a resonator made from a flattened Pyrex tube (10 mm × 0.24 mm × 0.06 mm) in which the pressure produced both a dimensional change and damping, with outside pressure noted to produce more damping than inside pressure [6]. The resonant frequency changed -0.0029 Hz/kPa of inside pressure and +0.0108 Hz/kPa of outside pressure. Since the sum of these is not zero, the sensor would produce a large common mode signal when used in a differential mode.

Bois *et al.* [7] described a hybrid force sensor incorporating electrical circuitry but fed only by a multimode optical fiber (Fig. 1-10). Their sensor consisted of a stainless steel double ended tuning fork whose tines, spread by impulses of current through a coil,

modulate the intensity of a probe beam reflected back into a fiber. Force applied by a pressure cell changes the tensile stress in the tines and their resonant frequency. A dichroic mirror designed to separate the pulsed 850 nm light, converted to electricity by the photodiode for the coil, from the continuous 660 nm beam, whose reflectance the tines modulate, allows one fiber operation. Since only the 660 nm probe is reflected back into the fiber, the photodiode at the source end needs no filter. Although the fiber optic link is immune to interference, the head is not.

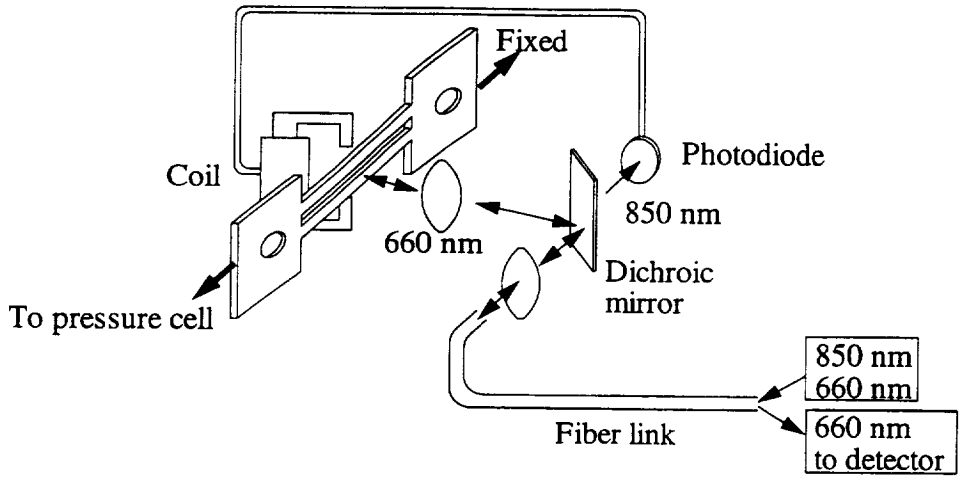


Fig. 1-10 Hybrid resonant tuning fork pressure sensor [7].

Although these sensors appear promising, designs where the measured fluid wets the resonant element are liable to suffer from changes in the resonant frequency due to condensate formation or fouling. In addition, high Q resonators must be operated in an ambient of less than 40 Pa to reduce damping to that of the resonator material [3] Nonlinearity and hysteresis have been found in resonant sensors and they were ascribed to a nonlinear restoring force [8]. Nonlinearity will make the sensor link-dependent.

Although the typically high thermal coefficient makes temperature compensation mandatory, it has been shown that temperature and pressure information can be multiplexed on the same fiber via the different resonant frequency of two resonators [2]. If the nonlinearity and hysteresis problems can be overcome, this design has the potential

of being a good sensor platform. Other measurands, such as acceleration [3], that can be made to stress a resonant element can also be measured.

Fabry-Perot

A pressure sensor head which is compact and often described in the literature consists of a thick slab, a spacer layer, and a thin diaphragm cap arranged to form a Fabry-Perot (F-P) cavity (Fig. 1-11). The cap and substrate surfaces inside the cavity may be coated to enhance their reflectivity. The spacer layer may be sputter deposited and lithographically defined (~$\frac{1}{2}$ μm thick), consist of a washer (~100 μm thick), or be dispensed with, and instead an etched recess in the substrate made. The head is attached to the tip of a fiber into which light is coupled, and light reflected back from the head is measured (Fig. 1-12). If the absolute intensity of light reflected is measured, a detector may monitor the source brightness. For ease of connectorization, multimode fibers are typically used.

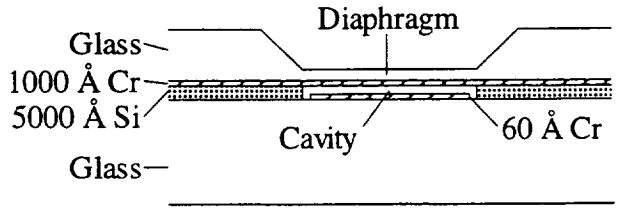


Fig. 1-11 F-P pressure sensor head [9]. In this example, cavity spacing is less than the coherence length of an LED source.

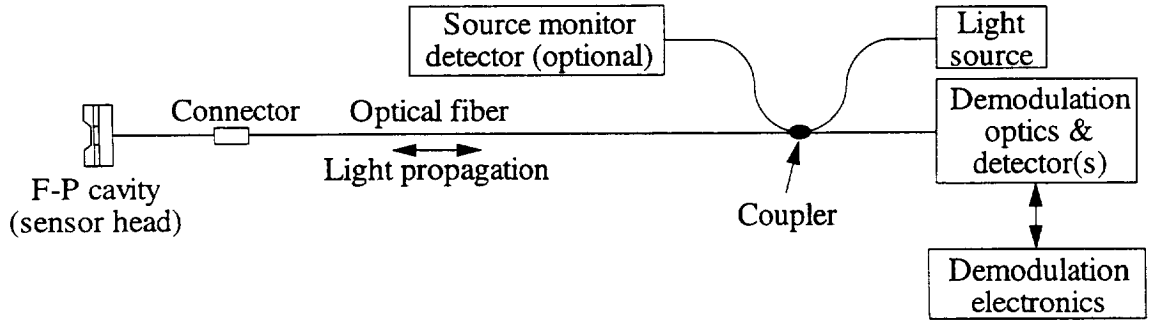


Fig. 1-12 Typical implementation of F-P type sensor.

An important design consideration is whether the interferometer's optical path length (OPL), which is twice the mirror spacing in the Fabry-Perot cavity, is greater than the source coherence length L_{coh}. The coherence length may be calculated from the spectral line width ($\Delta\lambda$)

$$L_{coh} = \frac{\lambda_0^2}{\Delta\lambda} \quad , \tag{1.3.1}$$

where λ_0 is the wavelength at the peak spectral power density. As the cavity mirror spacing is increased from 0 μm, the reflectance will periodically vary between a minima and a maxima. As the separation increases, approaching $L_{coh}/2$, the amount of variation will diminish and monotonically approach 0. For cavity lengths significantly greater than $L_{coh}/2$, the power reflected will be independent of cavity length. In this case, the reflectance minima are spaced closely enough in wavelength that the separation between them is less than the line width of the source. Wavelengths not reflected by the cavity are missing from the envelope emitted by the source (Fig. 1-13).

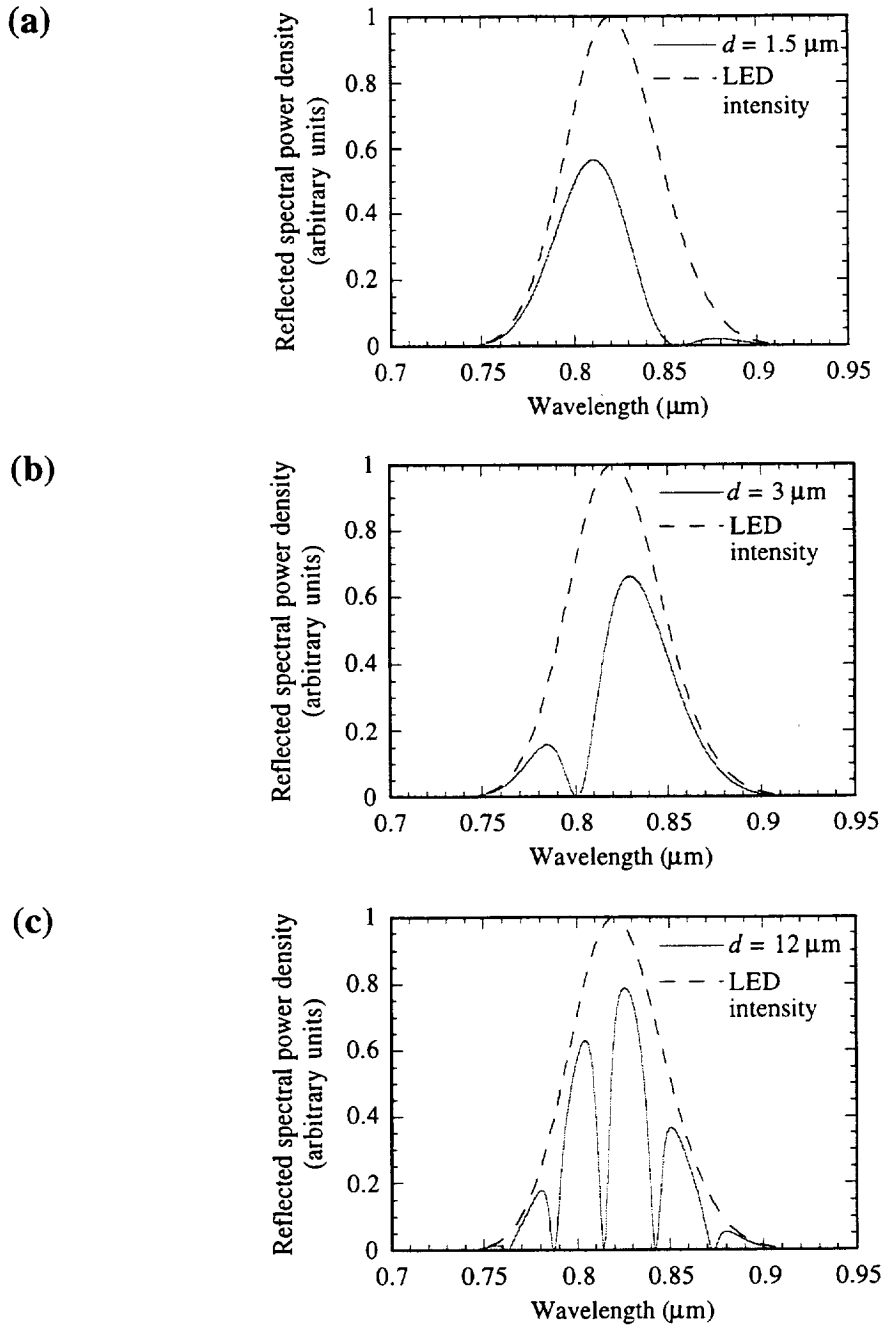


Fig. 1-13 Calculated reflected spectral power density (solid curve) for a F-P interferometer possessing 40% reflectance mirrors separated by d μm and interrogated by an LED having $L_{coh} = 6$ μm. As the distance between the mirrors increases, the reflectance spectra minima and maxima shift to longer wavelengths and the separation between minima decreases, causing more of them to appear under the LED envelope. Since the power, proportional to the area under the solid curve, is then modulated less by shifts of the curve, the fringe visibility decreases.

If the cavity separation is > $L_{coh}/2$, amplitude variations due to changes in cavity separation (fringes) may still be observed if light reflected from the sensor cavity is fed through another interferometer (the analyzing interferometer) such that the difference between its optical path length and that of the sensor head is < L_{coh}. Because the path lengths of the interferometers nearly match, all wavelengths reflected by the head can be canceled by the analyzing interferometer. As the difference of the path lengths increase, wavelengths reflected by the head will coincide less with those blocked by the analyzing interferometer and the fringe visibility will decrease. This technique is known as path matched differential interferometry [10].

Various techniques have been have been proposed to interrogate this sensor head. Lieberman and Blonder [9] propose using a LED to measure a 0.5 μm cavity and monitoring the coherently reflected intensity. This technique, while simple to demodulate, is not link independent and the cavity reflectance will depend on the wavelength emitted by the LED, which will vary 0.3 nm/°C [11]. In addition, reflection from connectors will be indistinguishable from that of the head.

Saaski *et al.* [12] and Dakin *et al.* [13] solve the link sensitivity problem by calculating the diaphragm deflection, and thus pressure, from the intensity ratio of two reflected wavelengths. Saaski *et al.* used a single LED source with a 1 to 2 μm cavity spacing and split the reflected light into two wavelengths which were separately monitored. Since effects which alter the fiber transmissivity equally at both ends of the LED spectrum will not effect this ratio, the technique is link insensitive.

Although ratioing the intensity of the two wavelengths returned from the sensor head can reduce link sensitivity, connector back reflection will add equally to the ratio's numerator and denominator, altering it. In addition, while the source monitor intensity and back reflected head signals emerge from different fibers (see Fig. 1-12) some cross coupling can occur. Dakin *et al.* [13] reduced this problem by connecting their head to the coupler of Fig. 1-12 via 1.1 km of optical fiber and pulsing the LED's at a frequency

(23 kHz) which would place the light pulses reflected from the head electrically 90° out of phase both from cross coupled signal used to monitor the LED and from reflections from connectors close to the coupler. Processing relying on this phase shift employing synchronous demodulation, the technique used in lock-in amplifiers, can remove the cross coupled signals. Rejection of light reflected from connectors placed near the electronics (far from the head) will also occur.

Diaphragm deflections in devices presented so far are restricted to $\lambda_0/4$ to prevent ambiguity and ensure a zero, regardless of what movement may be permitted from diaphragm breakage considerations. Temperature changes will cause variations in cavity separation, so it is desirable to permit as large a deflection as possible. Halg [14] has described a sensor configuration which consists of a F-P head with a 3 μm cavity spacing ($> L_{coh}/2$) which may be fed by a single multimode optical fiber (Fig. 1-14). An LED illuminates the cavity and the reflected light falls on three detectors simultaneously. Two of the detectors have F-P filters on them; one filter is $\lambda_0/8$ thicker than the other. The difference between the optical path length of the filters and the head F-P cavities is $< L_{coh}$ so the light falling on the filtered detectors varies as pressure alters the cavity gap. As the head cavity length changes one detector will see maximum intensity while the other will see an intensity in between the maximum and minimum. If one filtered detector's normalized output is plotted vs. the other, the curve will spiral out, and then back. If the head optical path length is confined to be either greater or less than the reference cavities, each point on the spiral will correspond to a unique head gap. Back reflection from connectors will add to the signal from the filtered and normalization detectors, which could cause an incorrect turn on the spiral to be chosen and a large error in the pressure measurement.

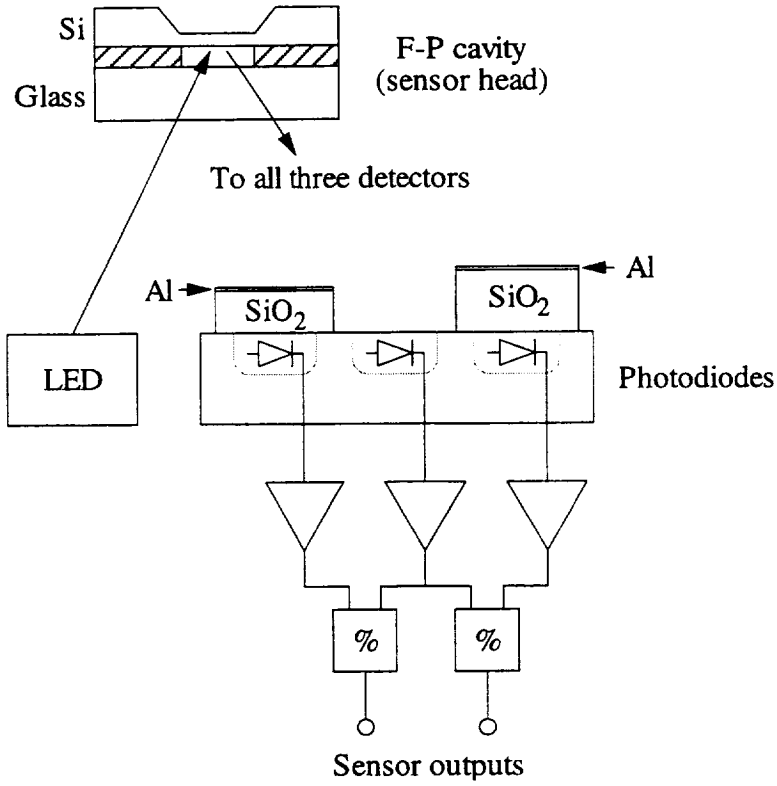


Fig. 1-14 Halg pressure sensor [14]. The reference cavities are made by thin film deposition onto silicon detectors. The detector having no cavity is used to compensate for link transmissivity changes.

Another demodulation technique which works with cavity movements in excess of $\lambda_0/2$ and uses a wide linewidth source is to send the light reflected from the sensor head, which has its OPL >> L_{coh}, into an analyzing interferometer and adjust it for maximum fringe visibility. Then the analyzing interferometer's OPL matches that of the sensor. Such a technique is insensitive to variations in source wavelength, link transmissivity, and light reflected from connectors but requires either moving parts or an electrooptic modulator. It is also necessary to accurately ascertain the OPL of the analyzing interferometer. A demonstration of the technique is given by Beheim *et al.* [15]. Their analyzing interferometer consists of two metal coated glass plates whose separation can be controlled with a piezoelectric translator. This separation is determined by incorporating the metal coated plates, which form a capacitor, in a circuit whose frequency is proportional to the plate's spacing. To set the plate's separation to match the

head, a sinusoidal dither, which varies the plate separation by ~2 μm, is applied to the piezoelectric translator, and the fringe movement (AC output of the interferometer) measured with a photodiode. The fringe amplitude will increase as the dither moves the mirrors to the place of maximum fringe visibility.

In the dither cycle, as the plate separation changes, a rapidly varying sinusoid, corresponding to interference between the two plates, will be noted. This variation will come under an envelope, and its peak will coincide with the glass plate separation providing maximum fringe visibility. At the peak, the head and analyzing interferometer's OPL will match. To make the average plate separation coincide with this peak, the analyzing interferometer's photodiode output is high-pass filtered (giving it zero DC content), rectified, and peak followed (low pass filtered). This signal, which will follow the envelope, is then multiplied by the sign of the dither sinusoid. If the envelop peak coincides with the zero crossing of the dither, the peak will be split, with one-half being multiplied by +1 and one half by -1 and the product will have zero DC value. If the peak falls on one side of the dither (say the -1 side) then the multiplication will produce a signal having a negative average value which when low pass filtered gives a negative DC voltage. The DC signal is fed back to the piezoelectric translator which adjusts the mirror spacing, making its average position identical to that of the sensor head.

An alternative approach by Belsley *et al.* is to replace the piezoelectrically adjusted cavity with a spectrometer-microprocessor [16]. In this scheme (Fig. 1-15) LED light with a coherence length less than twice the spacer thickness interrogates a F-P cavity at the end of an optical fiber. The reflected light, spectrally modulated by the cavity (Fig. 1-13c), is wavelength-separated. To decrease the time between pressure measurements, a linear photodetector array simultaneously measures the wavelength-segment's amplitude. This information is fed into a microprocessor, which determines the pressure (they are not specific how that's done). It is claimed that the instrument possesses a 0.002% resolution, a 0.02% hysteresis and a temperature zero point shift of

0.005%/°C. Since the microprocessor must process many intensity points and may need to use a computationally complex algorithm to extract the pressure, a concern is that the time between readings may be excessive. No data rates are provided, however.

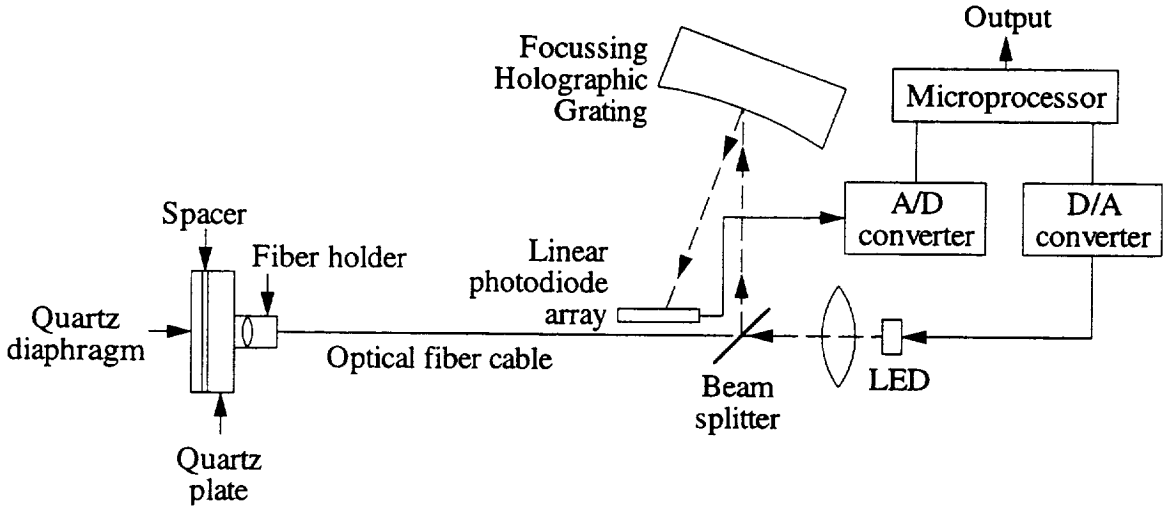


Fig. 1-15 Pressure sensor employing a F-P head and a spectrometer to determine cavity separation.[16]

In a separate paper, Velluet *et al.* [17] suggests that a spectrophotometer-determined cavity separation can give the absolute cavity size, with no restriction that it move less than $\lambda_0/4$. The "period" of the sinusoidal variation in wavenumber (reciprocal wavelength) space can provide a coarse estimate, with the "phase" providing the refinement.

One possible problem if the cavity is vented to make a relative pressure sensor is that air between the two plates can squeeze-film dampen rapid pressure variations, limiting the device's frequency response [18].

The Four-Fiber Compensation Technique and the Bulk Photoelastic Head

An intensity-type pressure sensor which satisfies the link and temperature insensitivity requirement, possesses a zero, and has no ambiguity problem if operated over a certain pressure range, has been demonstrated. This sensor uses a four-fiber

compensation technique which corrects for link, source-brightness, and detector-sensitivity variations. The four-fiber compensation technique employs two sources, a variable coupler whose coupling coefficient is affected by the measurand, and two detectors. Referring to Fig. 1-16, if the coupler has the transfer function,

$$\begin{aligned} P_{out1} &= (1-K)\,P_{in1} + K\,P_{in2} \\ P_{out2} &= K\,P_{in1} + (1-K)\,P_{in2} \end{aligned} \quad , \tag{1.3.2}$$

where P_{in} and P_{out} are the powers into and out of the coupler ports, the power entering the detectors (P_C and P_D) can be written in terms of the power emitted by the sources (P_A and P_B), the fiber link transmissivities (T_1, T_2, T_3, and T_4) and the power coupling coefficient (K) as

$$\begin{aligned} P_C &= P_A\,T_1\,(1-K)\,T_3 + P_B\,T_2\,K\,T_3 \\ P_D &= P_A\,T_1\,K\,L_4 + P_B\,T_2\,(1-K)\,T_4 \end{aligned} \quad . \tag{1.3.3}$$

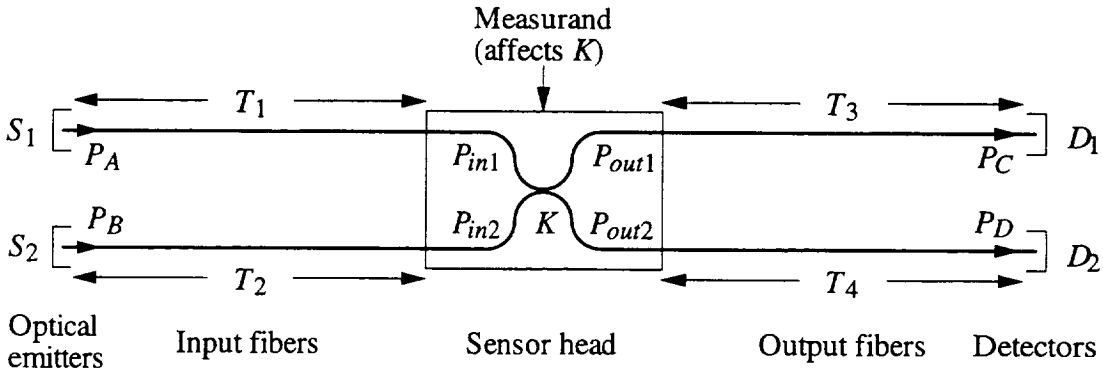


Fig. 1-16 Four-fiber link-insensitive sensor layout in which the measurand affects the power coupled between two fibers [19].

By time division multiplexing the sources, it is possible to separate P_A and P_B's contribution to P_C and P_D. To recover K, form the ratio

$$\frac{P_D|_{S_1\text{ on},\,S_2\text{ off}} \cdot P_C|_{S_1\text{ off},\,S_2\text{ on}}}{P_C|_{S_1\text{ on},\,S_2\text{ off}} \cdot P_D|_{S_1\text{ off},\,S_2\text{ on}}} \quad , \tag{1.3.4}$$

where P_C and P_D is the power incident on detector D_1 and D_2 [19]. Substituting (1.3.3) into (1.3.4), the following is obtained:

$$\frac{P_A T_1 K T_4 \cdot P_B T_2 K T_3}{P_A T_1 (1-K) T_3 \cdot P_B T_2 (1-K) T_4} = \left(\frac{K}{1-K}\right)^2 . \tag{1.3.5}$$

All the quantities cancel except for K, showing that K can be recovered from the ratio given by (1.3.4) independently of the loss of the fibers, the brightness of the source or the responsivity of the detectors.

A slightly different topology, which can link insensitize a transducer whose transmissivity T_A (not coupling coefficient) depends on the measurand is given in Fig. 1-17 [20]. Forming (1.3.4) using Fig. 1-17, we find

$$\frac{P_A T_1 T_B T_4 \cdot P_B T_2 T_C T_3}{P_A T_1 T_A T_3 \cdot P_B T_2 T_D T_4} = \frac{T_B T_C}{T_A T_D} , \tag{1.3.6}$$

which is dependent only on transmissivities internal to the sensor head. Other than T_A, these must remain constant under all conditions. A head employing this technique might be implemented with integrated optics using in addition to the transducer which controls T_A, three Y-branches, one cross and some interconnecting channels.

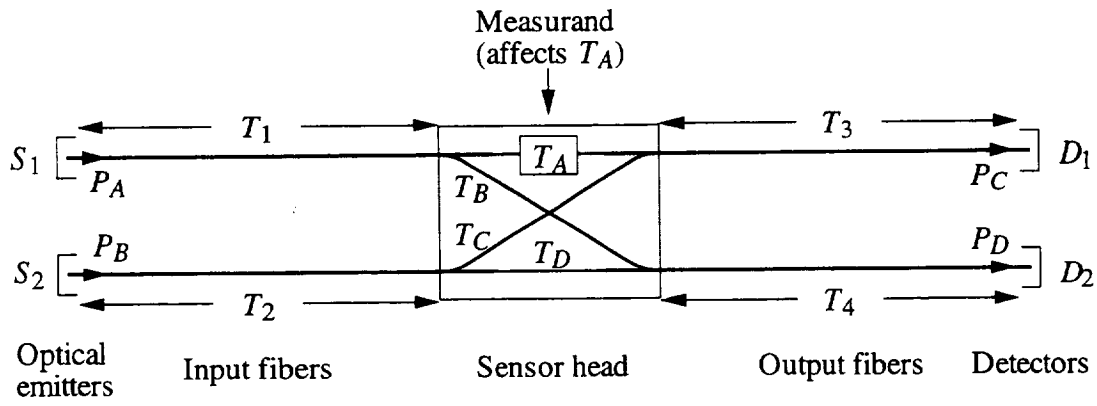


Fig. 1-17 A four-fiber layout which makes a simple intensity attenuation sensor link-insensitive [20].

Wang has demonstrated a pressure sensor using a bulk photoelastic slab, and fed by four multimode fibers tipped with graded index rod lenses, positioned around a polarizing beam splitting cube (Fig. 1-18). The light from fiber 1, unpolarized after leaving the grin lens, passes through a polarizer positioned so that light emerging will be polarized 45° with respect to the x-axis. Pressure applied to a diaphragm squeezes the photoelastic material (silica) along the y-axis, causing a negative strain (decrease in length) in the y-direction. Simultaneously, it causes a positive strain in x and z equal to -ν·(strain in y), where ν is Poisson's ratio. Through the photoelastic effect, the strain makes the material birefringent, with the slow axis along y. The photoelastic material becomes a waveplate, with the retardation proportional to the force in y, which in this sensor, is derived from pressure on a plate. Light from the polarizer, with its electric field vector oriented at a 45° angle with respect to the slow axis, launches half its power into the slow wave, polarized along y, and half into the fast wave, polarized along x. After propagating through the photoelastic material, a phase shift, caused by strain, will accumulate between the x and y polarizations--the light will become elliptically polarized. The quarter wave plate, oriented with its slow axis parallel to the electric field-transmissive axis of the polarizer, converts the phase shift into an intensity change in the x and y polarized electric fields. The polarizing beam splitting cube reflects the y polarization into fiber 4, while the x polarization is transmitted into fiber 3. Unpolarized time multiplexed light from fiber 2 allows the transmissivity ratio of fibers 3 and 4 to be determined and removed from the photoelastic material stress measurement.

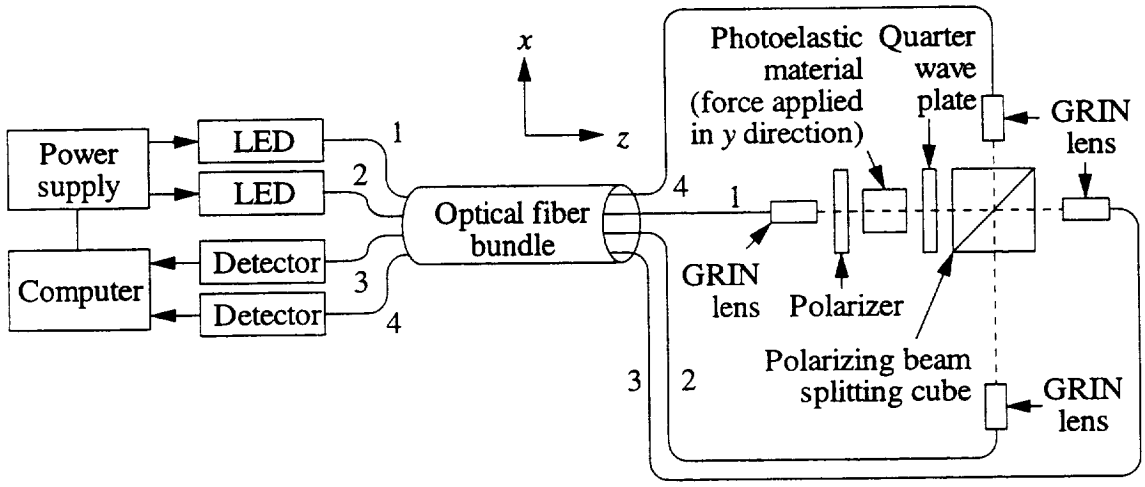


Fig. 1-18 Four-fiber link-compensated pressure sensor in which pressure causes a change in the state of polarization of light passing through photoelastic material [21].

The intensity received by the detectors through fibers 3 and 4 is [21]

$$I_3 = I_1\, T_1\, T_3 \sin^2\left(\frac{\phi}{2} + \frac{\pi}{4}\right) + I_2\, \frac{T_2\, T_3}{2}$$

$$I_4 = I_1\, T_1\, T_4 \cos^2\left(\frac{\phi}{2} + \frac{\pi}{4}\right) + I_2\, \frac{T_2\, T_4}{2} \qquad (1.3.7)$$

where T is the fiber transmissivity, I is the LED intensity, and ϕ is the photoelastically induced phase shift between the x and y polarized waves. The link insensitive expression then becomes

$$\frac{\dfrac{I_3|_{I_1 \text{ on}, I_2 \text{ off}}}{I_3|_{I_1 \text{ off}, I_2 \text{ on}}} - \dfrac{I_4|_{I_1 \text{ on}, I_2 \text{ off}}}{I_4|_{I_1 \text{ off}, I_2 \text{ on}}}}{\dfrac{I_3|_{I_1 \text{ on}, I_2 \text{ off}}}{I_3|_{I_1 \text{ off}, I_2 \text{ on}}} + \dfrac{I_4|_{I_1 \text{ on}, I_2 \text{ off}}}{I_4|_{I_1 \text{ off}, I_2 \text{ on}}}} = \sin\phi \quad . \qquad (1.3.8)$$

The authors examined the effect of varying link loss, by altering the distance between two fiber end faces at a splice, and varying modal power distribution, by laterally offsetting the multimode fibers. A 30 dB link loss increase altered the measurement by 0.1%, while laterally offsetting the fibers by 60 μm produced a

measurement perturbation of 10%. Despite this, in an outdoor setting, the sensor deviated from its calibration curve by no more then 0.25% over 1 year.

This sensor type is attractive because it is constructed from off-the-shelf parts and taking a reading requires only measuring four signals and doing simple math. However, it requires four cables and the head can not be made as small as in other approaches.

Intensity

In intensity sensors, the transmission between fibers is controlled by the measurand. Although interferometric sensors may be used in this manner, intensity sensors typically do not take advantage of techniques such as spectral encoding or two wavelength ratioing. Unless other techniques such as four-fiber compensation are employed, the sensor will not be link insensitive. Nevertheless, their simplicity, ease of demodulation, output monotonicity, and unusually high temperature immunity make them attractive.

Microbending a fiber is a way to induce loss without creating joints and potentially compromising its strength or temperature withstanding. A sensor employing this technique is illustrated in Fig. 1-19 [22]. Two Al coated multimode fibers run into a sensor head made of a steel super alloy. One fiber is squeezed by corrugations having a 2 mm period, causing some of its guided power to radiate. The greater the pressure applied to the diaphragm, the harder the corrugations squeeze the fiber and the greater the loss. A reference fiber, routed along side the first, passes through the head but avoids the corrugations. The pressure is derived from the difference of the logarithms of the intensities of the squeezed fiber and the reference fiber $\left(=\log \frac{\text{Signal}}{\text{Reference}}\right)$, removing link transmissivity variations common to both fibers. The diaphragm center (and the distance between the corrugations) deflects linearly with pressure, moving 10 μm from a full scale pressure of 3000 psi.

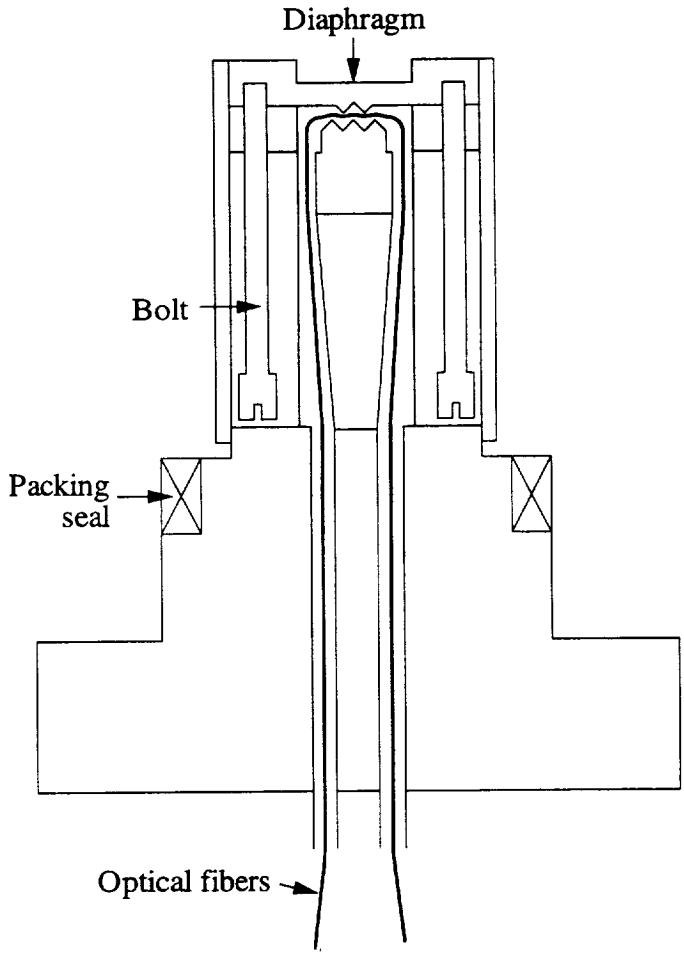


Fig. 1-19 Intensity-type pressure sensor intended for high temperature use [22].

The sensor was fitted to a pipe placed in an oven and tested at temperatures between 23°C and 428°C and pressures between 0 psi and 3000 psi. The light ratio varied by 15% of full scale when the temperature was changed from 23°C to 428°C and the pressure held constant. However, if temperature (T) was electrically measured, and it together with the light ratio (R) and fitted constants (c) were substituted into

$$P = c_1 + c_2R + c_3R^2 + (c_4 + c_5R + c_6R^2)T + (c_7 + c_8R + c_9R^2)T^2 \quad , \qquad (1.3.9)$$

the optically measured pressure (P) was found to agree with that from an accurate gauge placed outside the oven to within 1.2% of full scale.

Later on, this device was tested in a coal liquefaction line at temperatures up to 425°C [23]. As before, the temperature was monitored electronically. After an initial

zero point adjustment, pressure measured electronically in a cool location was found to agree within 2% of full scale with that measured optically. The device survived for 20 days, after which it failed because the squeezed fiber broke. It was speculated that a fiber coated with a different metal might survive longer.

Time division multiplexing techniques can be used to make intensity sensors link insensitive. In the sensor in Fig. 1-20 [24], an LED simultaneously modulated at 10 MHz and 20 MHz launches light into a coupler. Half of the launched power enters a branch and is reflected off of a mirror whose spacing from the fiber tip varies with pressure. As this distance increases, the amount of light back reflected into the fiber is reduced. The other branch's mirror causes all of its light to be back reflected. Because a 5 m fiber length (15 m OPL) difference exists between the two arms, light from the immobile mirror is delayed by 50 ns. Since the OPL difference greatly exceeds the coherence length of the LED, reflected light from these two arms sums incoherently (intensities simply add). Although the device was demonstrated with sinusoidal LED modulation, square wave modulation illustrates it best. Referring to Fig. 1-21, if the movable mirror reflects all of the light back into the fiber, the fiber delay causes cancellation of the 10 MHz signal, while reinforcing the 20 MHz one. If a pressure change moves the mirror to reduce the light back reflected into the fiber, the 10 MHz signal increases, while the 20 MHz one is reduced. The ratio of these two signals is independent of links external to the head.

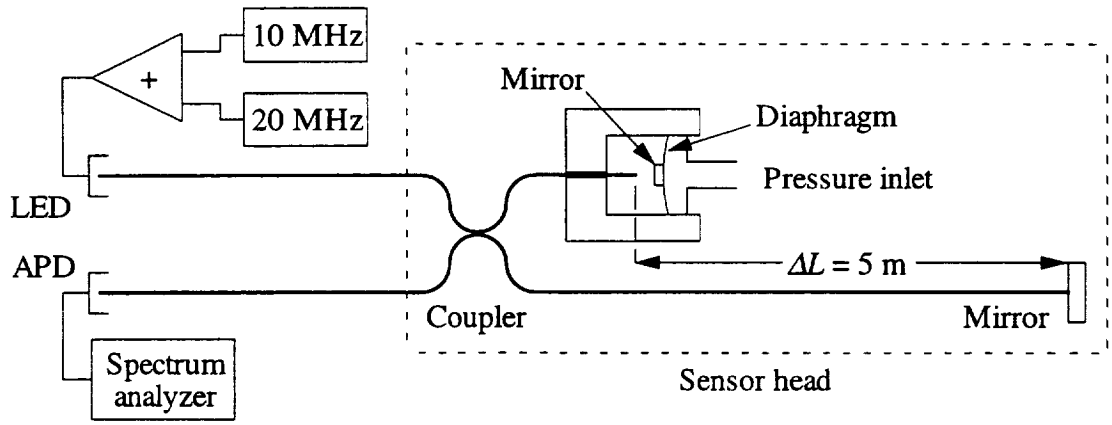


Fig. 1-20 Link insensitive intensity sensor using an LED source and an avalanche photo diode [24].

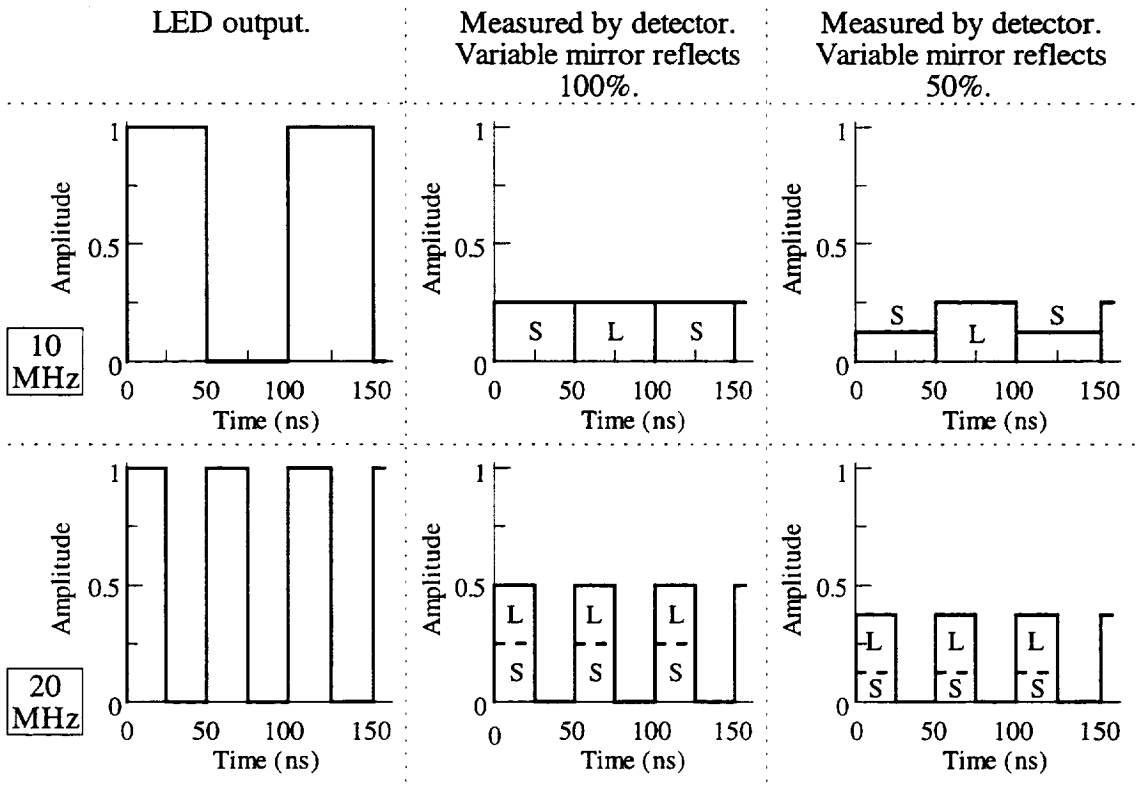


Fig. 1-21 Operation of the sensor illustrated with square wave modulation of the LED. Electrical signal amplitude is plotted vs time. Power in the "S" and "L" regions is back reflected from the short and long fiber.

Integrated Optic Interferometer on a Diaphragm

The sensing heads shown so far consist of multiple parts assembled together. A pressure sensor in which an interferometer and diaphragm are fabricated on the same wafer and thus permanently aligned has been presented by several groups. This sensor, shown in Fig. 1-22, typically uses a Mach-Zehnder interferometer and the diaphragm is formed by anistropic wet etching into a silicon substrate.

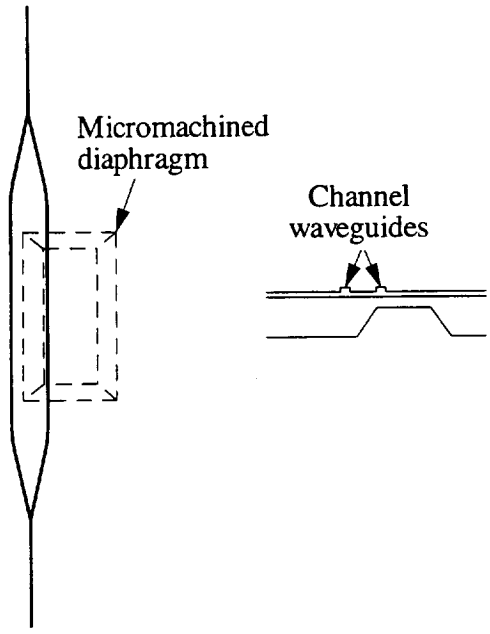


Fig. 1-22 Layout of integrated optic interferometer on a diaphragm.

In the conventional electronic piezoresistive pressure sensor, a diaphragm is created by anistropic etching into a silicon substrate. The etch may stop on layers grown on the front face, establishing diaphragm thickness, or etching may be discontinued when the proper silicon diaphragm thickness is reached. Piezoresistors diffused into the highest stress regions of the diaphragm alter their resistance, through the piezoresistive effect, when strained by pressure applied to the diaphragm. These piezoresistors and others having an opposite sensitivity to stress, all on the same diaphragm, are connected in a bridge configuration to measure pressure. For the optical sensor, an integrated optic

waveguide is placed over the highest stress regions of the diaphragm. Pressure applied to the diaphragm creates stress which alters, through the photoelastic effect, the refractive index of films forming the waveguide. The waveguide segment over the diaphragm is actually part of an integrated optic interferometer, and so an imbalance (or change in imbalance) alters the interferometer's transmittance to a particular wavelength of light.

The integrated optic interferometer on a diaphragm can suffer from link sensitivity, ambiguity, and zero problems. As with other interferometric sensors, these problems can be overcome by interrogating and demodulating with an appropriate technique.

In one of the first demonstrations of the device, Ohkawa *et al.* [25] built a balanced Mach-Zehnder interferometer with one arm placed over the long edges of a series of seven diaphragms placed one after another. Their waveguides were formed from sputtered 7059 glass and their 1 mm × 0.22 mm × 7.9 μm diaphragms used a p^+ etch stop. They provide design curves which show that the quasi-TM polarization experiences more phase shift than the quasi-TE, and that both are more affected when the channel is routed over the diaphragm's long edge rather than over its middle.

Their paper also suggests how to calculate the phase change. First the deflection is determined for an applied pressure and from this the strain at the location of the waveguide found. The incremental phase change is then determined by integrating the change in the dielectric constant over the cross section of the waveguide. Once this is done, the change in the phase of the light is found by integrating the incremental changes over the length of segment on the diaphragm. The interferometer transfer function is then applied and the device's optical response found.

In Ohkawa *et al.* [25] the strain's variation throughout the cross section of the guide is taken into account in an intergal. However, assuming that the strain in all the layers (except the air) is the same as that calculated in the core simplifies the calculation and introduces little error. This is because for the small deflections encountered in this

device, the strain varies linearly through the cross section of the diaphragm, so the greater magnitude of index change above the core is offset by the lesser change beneath.

Ohkawa *et al.* [25] measured a phase shift of 0.039 rad/kPa, while calculating 0.11 rad/kPa. They assumed that the diaphragm was made entirely of SiO_2, however, which would inflate their calculated result. Their measured transmission as a function of applied pressure was nonlinear, possibly indicating that the bond between the support plate and sensor chip was not rigid.

Wagner *et al.* [26] constructed a Mach-Zehnder interferometer with one of its legs over the middle of a micromachined diaphragm, and integrated onto their chip a photodiode for measuring the intensity of light passing through their interferometer. They too measured a greater sensitivity for the quasi-TM mode. They applied a sufficient pressure to make their interferometer operate over several fringes and nonlinearity could be noted in their response, as the minima spacing in the sinusoidal intensity variation decreased as the pressure increased.

A pressure sensor consisting of an antiresonant reflecting optical waveguide positioned across a diaphragm has been proposed by Vadekar *et al.* [27]. For computational simplicity they use a circular diaphragm model in their study and conclude that when the waveguide is routed across the diaphragm diameter, the change of phase experienced by light due to diaphragm bulging predominates over photoelastic induced refractive index change.

Fischer *et al.* [28] patterned an integrated optic Mach-Zehnder interferometer on the front of a diaphragm which had a square post protruding from the middle of its back. The interferometer was configured with one of its arms near the diaphragm edge, while another ran near the diaphragm's middle. A pin pressing on the protrusion created stress of opposite polarity in each of the waveguide's arms, and the interferometer's transmissivity to $\lambda_0 = 0.633\ \mu m$ light as a function of pin force was recorded. This device was used to measure the photoelastic constants of the silicon oxynitride

waveguide core material which was deposited, on different samples, using plasma and thermal chemical vapor deposition techniques. The photoelastic constants of the thermal films were smaller than the plasma films.

In all of these devices the Mach-Zehnder was nominally balanced, and monochromatic light was used for interrogation. Since both polarizations respond differently to pressure, their ratio potentially could be used to remove link transmissivity changes. Another approach is to interrogate an unbalanced Mach-Zehnder interferometer with a broad band source and spectrally separate the output, permitting the phase of both polarizations to be determined independently of link transmissivity. An unbalanced interferometer will be temperature sensitive, and so this approach can polarization multiplex both temperature and pressure information on the same fiber. Of course a polarization preserving fiber must transmit the output of the device to the demodulation unit.

Electromagnetic interference immunity is lost when a detector is mounted on the sensing head chip. In addition, both wire and fiber connections must be made, increasing package complexity.

4 CONCLUSIONS

Common problems with optical sensing systems have been presented. Examples of optical pressure sensors of various designs have been provided. Intensity sensors, while simple to interrogate, require more effort to link insensitize. The Fabry-Perot sensor is compact and can be interfaced with one multimode fiber. However, when measuring relative pressure, where gas is allowed into the cavity between the two parallel elements, damping may prevent a rapid response. Vibrating element sensors must be temperature compensated to be useful and require circuitry to determine their resonant frequency. Sensors formed by placing an integrated optic interferometer on a diaphragm require fiber interfacing with polarization preserving fibers (at least on the output end).

5 REFERENCES

[1] L. A. Johnson and S. C. Jensen, "Problems and Approaches for Remote Fiber Optic Absolute Sensors," presented at Fiber Optic and Laser Sensors III, San Diego, CA, and published in Proceedings of SPIE, vol. 566, pp. 45-53, 1985.

[2] D. Angelidis and P. Parsons, "Optical Micromachined Pressure Sensor for Aerospace Applications," *Opt. Eng.*, vol. 31, pp. 1638-1641, 1992.

[3] M. J. Tudor, M. V. Andres, K. W. H. Foulds, and J. M. Naden, "Silicon Resonator Sensors: Interrogation Techniques and Characteristics," *IEE Proceedings*, vol. 135, Pt. D, pp. 364-368, 1988.

[4] H. Unzeitig and H. Bartelt, "All-Optical Pressure Sensor with Temperature Compensation on Resonant PECVD Silicon Nitride Microstructures," *Electron. Lett.*, vol. 28, pp. 400-402, 1992.

[5] D. Uttamchandani, K. E. B. Thornton, J. Nixon, and B. Culshaw, "Optically Excited Resonant Diaphragm Pressure Sensor," *Electron. Lett.*, vol. 23, pp. 152-153, 1987.

[6] M. Johnson, "Optical-Actuator Frequency-Coded Pressure Sensor," *Opt. Lett.*, vol. 11, pp. 587-589, 1986.

[7] E. Bois, R. C. Spooncer, and B. E. Jones, "A Hybrid Resonant Differential Pressure Transmitter with Wavelength Multiplexed Power and Data Channels," presented at Optical Fiber Sensors, Proceedings of the 6th International Conference, OFS'89, Paris, France, and published in Springer Proceedings in Physics, vol. 44, pp. 478-483, Springer-Verlag publishers, 1989.

[8] M. V. Andres, K. W. H. Foulds, and M. J. Tudor, "Nonlinear Vibrations and Hysteresis of Micromachined Silicon Resonators Designed as Frequency-Out Sensors," *Electron. Lett.*, vol. 23, pp. 952-954, 1987.

[9] R. A. Lieberman and G. E. Blonder, "An Improved Interferometric Pressure Optrode," presented at Fiber Optic and Laser Sensors V, San Diego, CA, and published in proceedings of SPIE, vol. 838, pp. 49-59, 1987.

[10] A. D. Kersey, A. Dandridge, and A. B. Tveten, "Recent Advances in Demodulation/Multiplexing Techniques for Interferometric Fiber Sensors," presented at Fiber Optics '87, London, England, and published in proceedings of SPIE, vol. 734, pp. 261-269, 1987.

[11] B. Masterson and J. Webb, "Better Diode Lasers Will Expand Applications in Many Areas," *Photonics Spectra*, pp. 119-124, May 1992.

[12] E. W. Saaski, J. C. Hartl, G. L. Mitchell, R. A. Wolthuis, and M. A. Afromowitz, "Multimode Fiber Optic Pressure Sensor with Extended Range," presented at Fiber Optic and Laser Sensors V, San Diego, CA, and published in proceedings of SPIE, vol. 838, pp. 46-48, 1987.

[13] J. P. Dakin, C. A. Wade, and P. B. Withers, "An Optical Fiber Pressure Sensor," presented at Fiber Optics '87, London, England, and published in proceedings of SPIE, vol. 734, pp. 194-201, 1987.

[14] B. Halg, "A Silicon Pressure Sensor With an Interferometric Optical Readout," presented at International Conference on Solid State Sensors and Actuators, pp. 682-684, 1991.

[15] G. Beheim, K. Fritsch, and R. N. Poorman, "Fiber-Linked Interferometric Pressure Sensor," *Rev. Sci. Instr.*, vol. 58, pp. 1655-1659, 1987.

[16] K. L. Belsley, D. R. Huber, and J. Goodman, "All-Passive Fiber-Optic Pressure Sensor," *InTech*, vol. 33, pp. 39-42, 1986.

[17] M. T. Velluet, P. H. Graindorge, and H. J. Arditty, "Fiber Optic Pressure Sensor Using White Light Interferometry," presented at Fiber Optic and Laser Sensors V, San Diego, CA, and published in proceedings of SPIE, vol. 838, pp. 78-83, 1987.

[18] M. Andrews, I. Harris, and G. Turner, "A Comparison of Squeeze-Film Theory With Measurements on a Microstructure," *Sensors and Actuators A*, vol. 36, pp. 79-87, 1993.

[19] G. Martens, J. Kordts, and G. Weidinger, "A Photo-Elastic Pressure Sensor with Loss-Compensated Fiber Link," presented at Optical Fiber Sensors, Proceedings of the 6th International Conference, OFS'89, Paris, France, and published in Springer Proceedings in Physics, vol. 44, pp. 458-463, Springer-Verlag publishers, 1989.

[20] I. P. Giles, S. McNeill, and B. Culshaw, "A Stable Remote Intensity Based Optical Fibre Sensor," *J. Phys. E: Sci. Instr.*, vol. 18, pp. 502-504, 1985.

[21] A. Wang, S. He, X. Fang, X. Jin, and J. Lin, "Optical Fiber Pressure Sensor Based on Photoelasticity and its Application," *J. Lightwave Technol.*, vol. 10, pp. 1466-1472, 1992.

[22] J. W. Berthold, W. L. Ghering, and D. Varshneya, "Design and Characterization of a High Temperature Fiber-Optic Pressure Transducer," *J. Lightwave Technol.*, vol. 5, pp. 870-875, 1987.

[23] J. W. Berthold, "Field Test Results on Fiber-Optic Pressure Transmitter System," presented at Fiber Optic and Laser Sensors IX, Boston, MA, and published in proceedings of SPIE, vol. 1584, pp. 39-47, 1991.

[24] J. G. Eustace, G. A. Coghlan, C. M. Yorka, E. F. Carome, and G. Adamovsky, "Fiberoptic Sensing Technique Employing RF Modulated Interferometry," presented at Fiber Optic and Laser Sensors IX, Boston, MA, and published in proceedings of SPIE, vol. 1584, pp. 320-327, 1991.

[25] M. Ohkawa, M. Izutsu, and T. Sueta, "Integrated Optic Pressure Sensor on Silicon Substrate," *Appl. Opt.*, vol. 28, pp. 5153-5157, 1989.

[26] C. Wagner, J. Frankenberger, and P. P. Deimel, "Optical Pressure Sensor Based on a Mach-Zehnder Interferometer Integrated with a Lateral a-Si:H p-i-n Photodiode," *IEEE Photon. Technol. Lett.*, vol. 5, pp. 1257-1259, 1993.

[27] A. Vadekar, A. Nathan, and W. P. Huang, "Analysis and Design of an Integrated Silicon ARROW Mach-Zehnder Micromechanical Interferometer," *J. Lightwave Technol.*, vol. 12, pp. 157-162, 1994.

[28] K. Fischer, J. Muller, R. Hoffmann, F. Wasse, and D. Salle, "Elastooptical Properties of SiON Layers in an Integrated Optical Interferometer Used as a Pressure Sensor," *J. Lightwave Technol.*, vol. 12, pp. 163-169, 1994.

II Diaphragm Modeling

1 INTRODUCTION

The pressure sensors examined in this work are illustrated in Fig. 2-1. They consists of an integrated optic channel waveguide fabricated on a micromachined diaphragm. This channel waveguide is a segment in an integrated optic interferometer, such as a ring resonator or Mach-Zehnder. Gas or liquid pressure applied to the diaphragm induces strain, which through the photoelastic effect, alters the refractive index of the film making up the channel waveguide segment. Although this influences the channel transmissivity negligibly, it affects the phase of the light. Changing the film refractive index changes the effective index of the channel, which affects the path length (the product of the physical length and the effective index) of the interferometer. A change in path length may be detected as a shift in the wavelength maximally transmitted through the device.

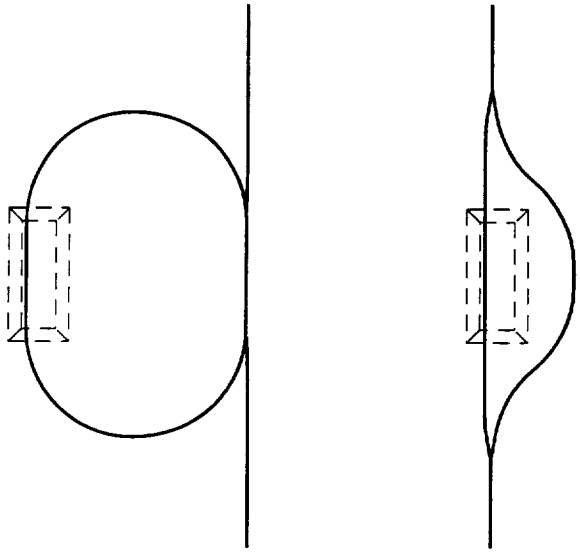


Fig. 2-1 Top view of integrated optic pressure sensors with a rectangular diaphragm.

To model the device it is necessary to: (1) calculate the strain in the waveguide, (2) find the resulting change this produces in refractive index, (3) determine the affect this has on the waveguide effective index, and, (4) determine how this perturbs the interferometer. Accordingly, mechanical models of the diaphragm are presented in this chapter. A 2-*D* beam, in which strain is easily calculated, is first presented. This structure accurately mimics the diaphragm's behavior over most of its area. Subsequently, equations to allow calculation of the deflection and strain in a 3-*D* plate, which will be compared with the beam, are presented. In the next chapter, waveguide models are derived for the three and four-layer planar waveguides which allow determination of bending loss, layer thickness, and channel width. The photoelastic effect, which couples mechanical strain into waveguide path length change will also be given there.

Symbol	Meaning	MKS unit
$\epsilon_x, \epsilon_y, \epsilon_z$	strain	unitless
$\sigma_x, \sigma_y, \sigma_z$	stress	Pa
E	isotropic Young's modulus	Pa
ν	Poisson's ratio	unitless
z	direction normal to diaphragm, $z = 0$ at diaphragm midplane	m
$w(x,y)$	deflection of diaphragm	m
$\frac{1}{\rho}$	radius of curvature	m
M_a	moment at built-in beam edge	N-m
V_a	clamping force at edge	N
I	moment of inertia	m^4
p	pressure applied to beam or diaphragm	Pa
b	width of beam	m
q	force per unit beam length	N/m
t	full beam thickness	m
h	full plate thickness	m
x	distance from beam built-in edge	m
L	length of beam	m
D	plate flexural rigidity	m
a, b	lengths of side of the plate	m
i, m	counting integers in plate solution	unitless

Table 2-1 Beam and plate modeling symbols.

2 BEAM MODEL

Before embarking on the plate theory used to analyze the diaphragm, the simpler beam theory will be discussed. The deflection and stress in a beam is much more easily calculated than in a plate. In addition, some of the elasticity relationships derived for the beam can be applied to a plate. Finally, straightforward theory exists to determine the deflection and stress in beams consisting of multiple layers and/or undergoing large deflections.

A film subject to a uniform gas pressure and held by its edges against an opening may be modeled in two ways. If film deflection is large compared to thickness, then stretching gives rise to a tensile stress uniform throughout the film's thickness (balloon stress) and the film is considered to act as a membrane. If, however, the film is sufficiently rigid that it deflects little, then the film acts as a plate and pressure is resisted by bending, with stress varying linearly with distance from the film's center plane which remains free from normal stress. Because deflection is so small in this last case, balloon stress, associated with elongation of the center plane of the plate, remains negligible. Even though our typical diaphragm is roughly 10 μm thick, it is this latter case that most closely approximates our diaphragm's behavior.

A diaphragm fabricated by anistropic etching from a silicon wafer will have its edges built-in or clamped, meaning that they are not free to displace or rotate. This may be contrasted to the simply-supported case, (for example a plate resting over hole) where no torques are applied by the support to the plate's edges, so they may rotate, but not displace.

In a plate having one dimension much larger than the other (Fig. 2-2), the deflection, stress, and strain away from the short edge will be the same as in a beam, provided that the beam is constructed of the same layers and has the same length as the short edge of the diaphragm. Although this ignores the effect on the plate of the short clamped edges, constraining them does not affect the diaphragm except in their

immediate vicinity. For near unity aspect ratio diaphragms, the single beam approximation can be modified to increase its accuracy [1].

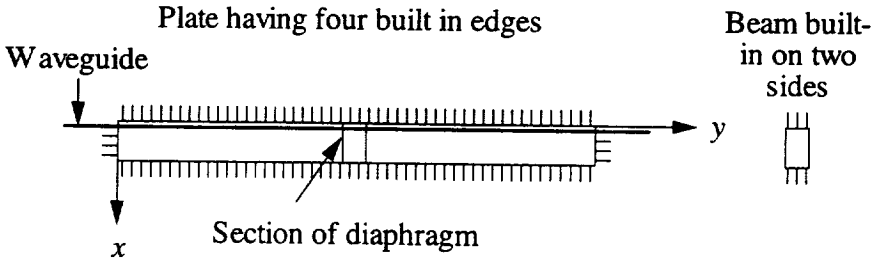


Fig. 2-2 Top view of beam and large aspect ratio plate. The deflection, stress, and strain in the shaded portion of the diaphragm is approximately the same as in the beam. $z = 0$ at the midplane.

Our stress analysis begins with the infinitesimal volume element, pictured in Fig. 2-3, which is used to describe stress at a point. Nine stresses may be applied to this element comprised of three normal and six shear stresses. A normal stress is assumed positive if it is in the same direction as the area vector associated with the face (if it is tensile). A shear stress is assumed positive if it is applied to a positive face (a face through which a positive axis passes) in a positive direction or a negative face in a negative direction. Since the element is in static equilibrium, an equal and oppositely directed stress must be applied to its reverse face.

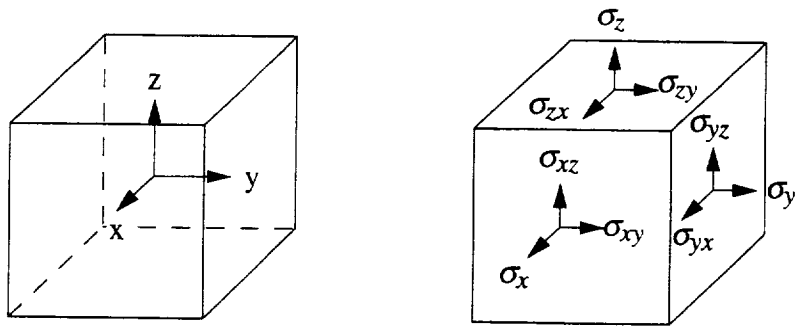


Fig. 2-3 Notation for stress applied to a volume element. The element coordinate origin is at the center of the cube.

Applied stresses will distort the element. For example, the normal stress σ_x will

change the element's length in the x direction by

$$\epsilon_x = \frac{\sigma_x}{E} \quad , \tag{2.2.1}$$

where E is Young's modulus and ϵ is strain (change in length per unit length). However, this stress will also cause constriction in the y and z directions. The coupling constant is Poisson's ratio defined by

$$\epsilon_y = \epsilon_z = -\nu\epsilon_x = -\frac{\nu}{E}\sigma_x \quad . \tag{2.2.2}$$

Poisson's ratio is between 0 and 0.5 with a typical value of 0.25. While strain and Poisson's ratio are unitless, stress and Young's modulus have the units of pressure (force/unit area). In a similar manner, a separate application of normal stress in the other directions will produce strain. When only one stress is applied at a time, these may be written [2]:

$$\begin{array}{lll} \epsilon_x = \frac{\sigma_x}{E} & \epsilon_y = -\frac{\nu\sigma_x}{E} & \epsilon_z = -\frac{\nu\sigma_x}{E} \\ \epsilon_x = -\frac{\nu\sigma_y}{E} & \epsilon_y = \frac{\sigma_y}{E} & \epsilon_z = -\frac{\nu\sigma_y}{E} \\ \epsilon_x = -\frac{\nu\sigma_z}{E} & \epsilon_y = -\frac{\nu\sigma_z}{E} & \epsilon_z = \frac{\sigma_z}{E} \end{array} \quad . \tag{2.2.3}$$

If more than one stress is applied the strains can be superposed

$$\epsilon_x = \frac{1}{E}[\sigma_x - \nu(\sigma_y + \sigma_z)] \quad , \tag{2.2.4a}$$

$$\epsilon_y = \frac{1}{E}[\sigma_y - \nu(\sigma_x + \sigma_z)] \quad , \tag{2.2.4b}$$

and

$$\epsilon_z = \frac{1}{E}[\sigma_z - \nu(\sigma_x + \sigma_y)] \quad . \tag{2.2.4c}$$

Consider the cross section of a beam in pure bending shown in Fig. 2-4 to which pressure has been applied, causing it to bow upward. Normal stress at the neutral axis is zero. Curvature and not slope evidences the applied pressure. In the figure, the volume elements near the top of the beam will be in tension in x, while those at the bottom will be under compression. To relate curvature to strain, note that an element at distance *bd* from the beam neutral axis having original length *ab* elongates by *cd*. The normal strain can be

written

$$\text{strain} = \epsilon = \frac{cd}{ab} = \frac{bd \cdot d\theta}{\rho \cdot d\theta} = \frac{z}{\rho} \quad . \tag{2.2.5}$$

A fiber's strain at a distance z from the neutral plane of a beam having radius of curvature ρ is z/ρ.

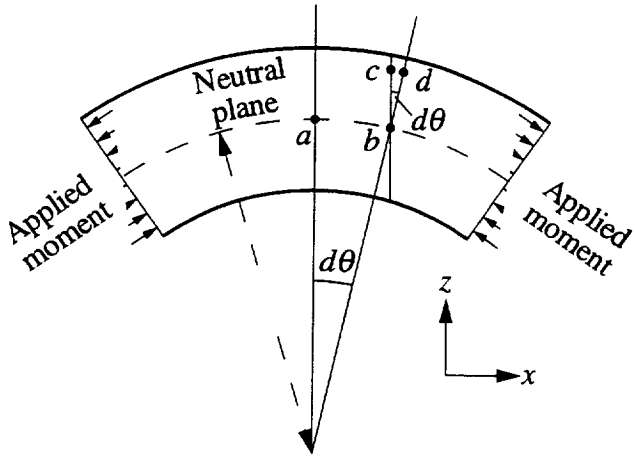


Fig. 2-4 Geometry of a bent beam.

From calculus, the radius of curvature may be written

$$\frac{1}{\rho} = \frac{-\frac{d^2w}{dx^2}}{\left[1 + \left(\frac{dw}{dx}\right)^2\right]^{\frac{3}{2}}} \quad , \tag{2.2.6}$$

where w is the deflection of the beam in the z direction and a positive radius of curvature has been drawn in Fig. 2-4. Since the slope of the beam is small, (2.2.6) can be approximated as

$$\frac{1}{\rho} \approx -\frac{d^2w}{dx^2} \quad . \tag{2.2.7}$$

Using this, equation (2.2.5) can be recast as

$$\epsilon_x = -z\frac{d^2w}{dx^2} \quad . \tag{2.2.8}$$

In these calculations we are concerned with strain in the diaphragm z distance from the

neutral plane (near the surface) and in the x-y region where light passes in the channel waveguide (see Fig. 2-2). Away from the short edge on our diaphragm, $\epsilon_y = 0$ because there is bending only about y and shear stress in the beam will prevent expansion or contraction of the top or bottom of the beam with respect to the neutral plane. In addition, $\sigma_z = 0$ because stress applied to the diaphragm by gas pressure is negligible compared to that experienced due to bending. Although $\epsilon_y = 0$, $\sigma_y \neq 0$ because of Poisson coupling of ϵ_x. Since more than one stress is acting at a time, equations (2.2.4a-c) must be used instead of (2.2.1)-(2.2.3). To find σ_x and σ_y, use (2.2.4b) giving,

$$\epsilon_y = 0 = \frac{1}{E}[\sigma_y - \nu\sigma_x] \tag{2.2.9}$$

or

$$\sigma_y = \nu\sigma_x \quad . \tag{2.2.10}$$

Equating (2.2.8) and (2.2.4a),

$$-z\frac{d^2w}{dx^2} = \epsilon_x = \frac{1}{E}[\sigma_x - \nu\sigma_y] \quad . \tag{2.2.11}$$

Substituting (2.2.10) into (2.2.11),

$$\sigma_x = \frac{-zE\frac{d^2w}{dx^2}}{1-\nu^2} \quad , \tag{2.2.12}$$

or

$$\sigma_x = \frac{z\frac{E}{\rho_x}}{1-\nu^2} \quad . \tag{2.2.13}$$

To find ϵ_z, use (2.2.10), (2.2.4c), and (2.2.12) to get

$$\epsilon_z = \frac{\nu}{1-\nu} z\frac{d^2w}{dx^2} \quad . \tag{2.2.14}$$

Although the change in thickness of the beam is zero because z strain on one side of the neutral plane is negated by reverse z strain on the other side, our concern is with strain where the light is confined; ϵ_z, given by (2.2.14), will not be zero there.

To calculate the deformation of the beam, note that a moment applied to one built-in edge can be written

$$M = \int z \cdot \sigma \cdot dA \quad , \tag{2.2.15}$$

where σ is the stress distribution across the edge. Inserting the results of (2.2.13) into (2.2.15), and using the coordinate system of Fig. 2-5, which places the z origin where $\int_{-t/2}^{t/2} \sigma_x \, dz = 0$, the moment about the x axis may be written

$$M_x = \frac{E}{\rho_x(1-\nu^2)} \int_{y=0}^{y=b} \int_{z=-t/2}^{z=t/2} z^2 \, dz \, dy \quad . \tag{2.2.16}$$

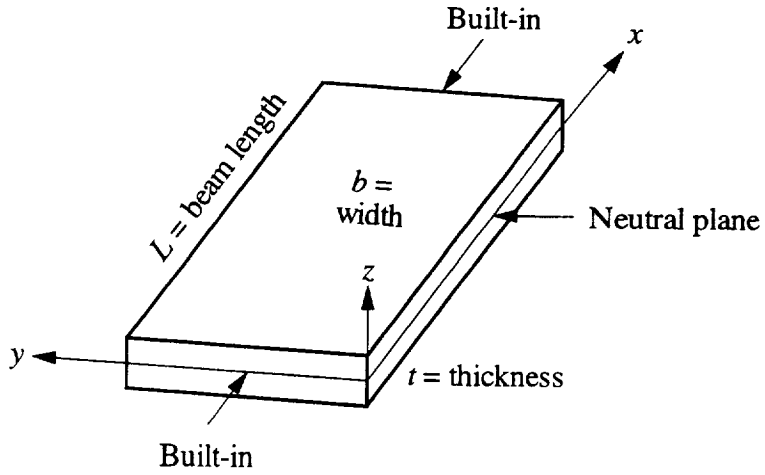


Fig. 2-5 Dimensions and coordinate system in beam model.

The moment of inertia is defined as

$$I = \int_{y=0}^{y=b} \int_{z=-t/2}^{z=t/2} z^2 \, dz \, dy \quad . \tag{2.2.17}$$

Carrying out this integration for the beam in Fig. 2-5,

$$I = \frac{bt^3}{12} \quad , \tag{2.2.18}$$

$$M_x = \frac{EI}{\rho_x\left(1-\nu^2\right)} \quad , \tag{2.2.19}$$

and using (2.2.7),

$$\frac{1}{\rho_x(x)} = -\frac{d^2w(x)}{dx^2} = \frac{M_x(x)}{E}\frac{\left(1-\nu^2\right)}{I} \quad . \tag{2.2.20}$$

Equation (2.2.20) is true at any cross section in the beam, not only at the ends, as the explicit x dependence indicates. The moment $M(x)$ must have the same direction as in Fig. 2-4 for the right hand side of (2.2.20) to have the correct sign. Strength of materials texts often ignore Poisson coupling and omit $1\text{-}\nu^2$ in deriving this equation. For a typical $\nu = 0.25$, only a small error is made. To find the deflection of a beam as a function of distance from its edge, a "cut" is made in the beam and the moment (torque) applied across the cut found. Using the moment and (2.2.20), the deflection of the beam may be found.

Consider the uniformly loaded beam built-in at both sides in Fig. 2-6. The pressure sensor diaphragm is similarly loaded and clamped. Since deflection is much less than t, the load is resisted by torques at the beam edge and elongation and stress in the center plane is negligible. Because the beam is clamped, the deflection and slope at both ends is zero. To find the moment as a function of x, a "cut" is made, resulting in the free-body diagram in the bottom of Fig. 2-6. Summing the torques at the right end of this beam segment gives,

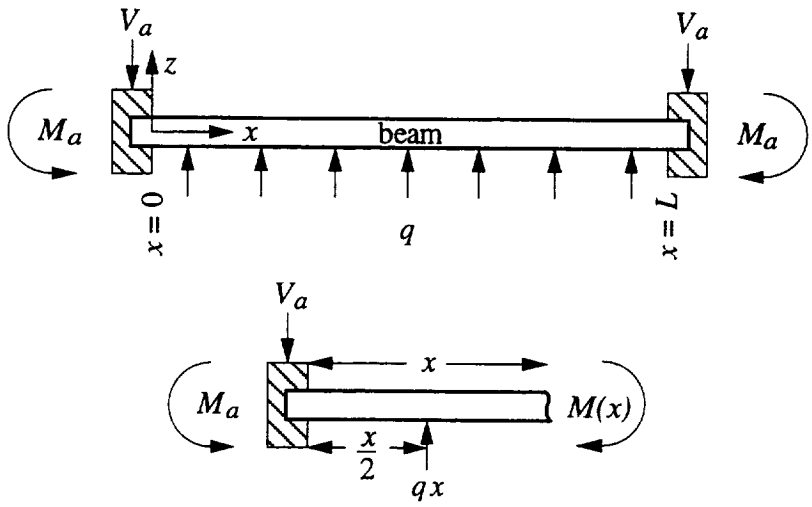
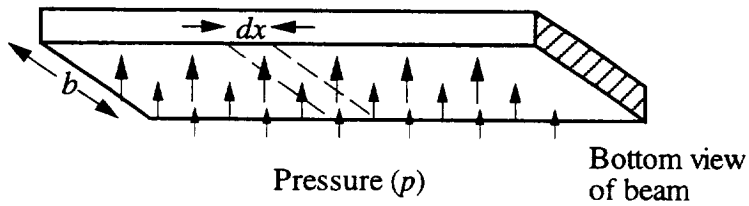


Fig. 2-6 Free-body diagram of beam. Bottom sketch is of beam segment used to find $M(x)$.

$$M(x) = M_a + V_a x - \frac{q}{2} x^2 \quad , \tag{2.2.21}$$

where q is the force per unit length per total beam width (N/m) (see Fig. 2-7).



$$q = \frac{force}{unit\ length} = \frac{p\,b\,dx}{dx} = pb$$

Fig 2-7 Definition of q.

Inserting (2.2.21) in (2.2.20),

$$\frac{EI}{1-\nu^2} \frac{d^2w(x)}{dx^2} = -M_a - V_a x + \frac{q}{2} x^2 \quad . \tag{2.2.22}$$

Integrating once gives

$$\frac{EI}{1-\nu^2} \frac{dw(x)}{dx} = -M_a x - \frac{V_a}{2} x^2 + \frac{q}{6} x^3 + C_1 \quad . \tag{2.2.23}$$

At $x = 0$ the slope is zero so $C_1 = 0$. Integrating again,

$$\frac{EI}{1-\nu^2}w(x) = \frac{-M_a}{2}x^2 - \frac{V_a}{6}x^3 + \frac{q}{24}x^4 + C_2 \quad . \tag{2.2.24}$$

Since $w(0) = 0$, $C_2 = 0$. Using $w(L) = 0$ and $\frac{dw(L)}{dx} = 0$, V_a and M_a are determined to be

$$V_a = \frac{1}{2}qL \tag{2.2.25}$$

and

$$M_a = -\frac{1}{12}qL^2 \quad . \tag{2.2.26}$$

Inserting these into (2.2.24) gives

$$w(x) = \frac{(1-\nu^2)qL^4}{24EI}\left(\frac{x^2}{L^2} - 2\frac{x^3}{L^3} + \frac{x^4}{L^4}\right) \quad . \tag{2.2.27}$$

Inserting (2.2.18) into (2.2.27), and using the definition of q from Fig. 2-7,

$$w(x) = \frac{(1-\nu^2)pL^4}{2Et^3}\left(\frac{x^2}{L^2} - 2\frac{x^3}{L^3} + \frac{x^4}{L^4}\right) \quad . \tag{2.2.28}$$

The maximum deflection occurs at $x = \frac{L}{2}$. It is

$$w_{max} = \frac{(1-\nu^2)pL^4}{32Et^3} \quad . \tag{2.2.29}$$

Normal stress and strain is found by taking the second derivative of (2.2.28) and substituting this into (2.2.12), (2.2.8), and (2.2.14) to find σ_x, ϵ_x, and ϵ_z:

$$\sigma_x(x,z) = \frac{-zpL^2}{t^3}\left(1 - \frac{6x}{L} + \frac{6x^2}{L^2}\right) \tag{2.2.30}$$

$$\epsilon_x(x,z) = -\frac{z(1-\nu^2)pL^2}{Et^3}\left(1 - \frac{6x}{L} + \frac{6x^2}{L^2}\right) \tag{2.2.31}$$

$$\epsilon_z(x,z) = \frac{z\nu(1+\nu)pL^2}{Et^3}\left(1 - \frac{6x}{L} + \frac{6x^2}{L^2}\right) \quad . \tag{2.2.32}$$

A diaphragm consists of a silicon core, and outer layers of silicon dioxide/silicon oxynitride; these have different Young's modulus. An equivalent beam, having identical resistance to bending will result if the different Young's modulus layers are replaced by layers with the same Young's modulus but having different widths (Fig. 2-8). This can be seen by substituting (2.2.18) into (2.2.19) giving

$$M_x = \frac{Ebt^3}{12\rho_x(1-\nu^2)} \quad . \tag{2.2.33}$$

The bending moment consists of the product of Young's modulus E and width b; so long as this product remains the same, resistance to bending will be unaltered. When calculating deflection and strain, the moment of inertia equation (2.2.17) takes into account the different widths. Although strain will be calculated correctly in the transformed material, the calculation of stress will require determining the transformed beam radius of curvature and inserting this in (2.2.13) with the original E.

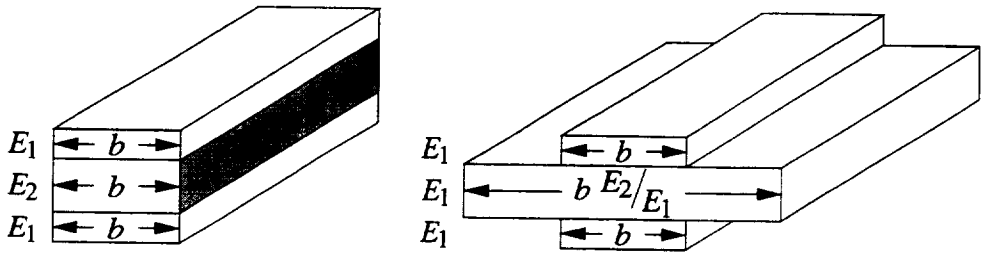


Fig. 2-8 Strain and resistance to bending is identical in these beams.

Equations (2.2.31)-(2.2.32), can also be written with the moment of inertia by substituting the second derivative of (2.2.27) into (2.2.8), and (2.2.14) giving

$$\epsilon_x(x,z) = -\frac{z(1-\nu^2)pbL^2}{12EI}\left(1 - \frac{6x}{L} + \frac{6x^2}{L^2}\right) \tag{2.2.34}$$

$$\epsilon_z(x,z) = \frac{z\nu(1+\nu)pbL^2}{12EI}\left(1 - \frac{6x}{L} + \frac{6x^2}{L^2}\right) \quad . \tag{2.2.35}$$

The quantity b in the numerator of (2.2.34)-(2.2.35) is the width to which pressure is applied in the original diaphragm. These equations, together with the definition of the moment of inertia (2.2.17) are used to solve the transformed beam of Fig. 2-8.

3 PLATE ANALYSIS

If the diaphragm does not have a large aspect ratio, or if there is an interest in including the variation in strain near the short built-in edges, the plate deflection must be

calculated. Once the plate deflection is determined, stress and strain may be found. As before, the assumption is made that deflection is small compared to plate thickness so that midplane stress may be neglected. To calculate plate deflection, the fourth order biharmonic differential equation

$$\frac{\partial^4 w}{\partial x^4} + 2\frac{\partial^4 w}{\partial x^2 \partial y^2} + \frac{\partial^4 w}{\partial y^4} = \frac{p}{D} \tag{2.3.1}$$

is solved subject to the boundary conditions resulting from the four built-in edges. Here p is a uniform pressure applied to the plate and D is given by

$$D = \frac{Et^3}{12(1-\nu^2)} \quad , \tag{2.3.2}$$

where t is the plate's full thickness and ν is Poisson's ratio. Timoshenko approached this problem by superimposing the deflections of a uniformly loaded simply-supported plate (which is more easily determined) with a plate which has a distributed edge torque [3]. The edge torque is selected to give the opposite edge slope of the simply-supported plate. Because the biharmonic differential equation is linear, a plate which has both a uniform load and the edge torques will have zero edge slope and deflection, thus satisfying the boundary conditions of the built-in plate. The deflection of this plate will therefore be the same as the built-in case.

In accord with the diaphragm dimensions and coordinate system given in Fig. 2-9, Timoshenko's moments are formulated as:

$$(M_y)_{y=\pm b/2} = \sum_{m=1,3,5,\cdots}^{\infty} (-1)^{(m-1)/2} E_m \cos\left(\frac{m\pi x}{a}\right) \tag{2.3.3}$$

$$(M_x)_{x=\pm a/2} = \sum_{m=1,3,5,\cdots}^{\infty} (-1)^{(m-1)/2} F_m \cos\left(\frac{m\pi y}{b}\right) . \tag{2.3.4}$$

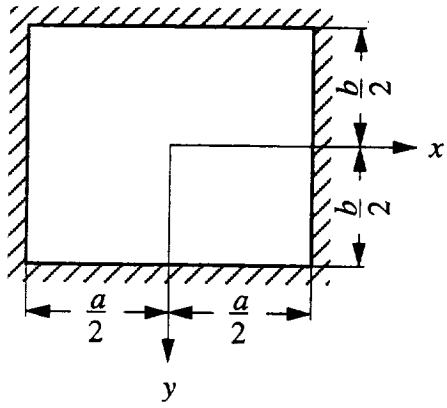


Fig 2-9 Dimensions and coordinate system used in Timoshenko's plate solution [3]. This coordinate system applies to equations (2.36)-(2.60). $z = 0$ is at the plate midplane.

To satisfy boundary conditions of the clamped plate, edge rotation along $y = \frac{b}{2}$ and $x = \frac{a}{2}$ must be zero (symmetry will make the edge rotation zero at the other two sides):

$$\left(\frac{\partial w_{s\text{-}s}}{\partial y} + \frac{\partial w_1}{\partial y} + \frac{\partial w_2}{\partial y}\right)_{y=b/2} = 0 \quad , \tag{2.3.5}$$

and

$$\left(\frac{\partial w_{s\text{-}s}}{\partial x} + \frac{\partial w_1}{\partial x} + \frac{\partial w_2}{\partial x}\right)_{y=a/2} = 0 \quad , \tag{2.3.6}$$

where $\frac{\partial w_{s\text{-}s}}{\partial y}$ and $\frac{\partial w_{s\text{-}s}}{\partial x}$ are the slopes of the simply supported plate at $y=b/2$ and $x=a/2$, $\frac{\partial w_1}{\partial y}$ and $\frac{\partial w_1}{\partial x}$ are the slopes at $y=b/2$ and $x=a/2$ resulting from a moment $My(x)$ applied to the $y = \pm b/2$ edges, and $\frac{\partial w_2}{\partial y}$ and $\frac{\partial w_2}{\partial x}$ are the slopes at $y = b/2$ and $x = a/2$ resulting from a moment $M_x(y)$ applied to the $x = \pm a/2$ edges. After substituting the proper expressions into (2.3.5) and (2.3.6), the following equations are obtained:

$$\frac{4pa^2}{\pi^3}\frac{1}{i^4}\left(\frac{\alpha_i}{\cosh^2(\alpha_i)} - \tanh(\alpha_i)\right) - \frac{E_i}{i}\left(\tanh(\alpha_i) + \frac{\alpha_i}{\cosh^2(\alpha_i)}\right) - \frac{8ia}{\pi b}\sum_{m=1,3,5,\ldots}^{\infty}\frac{F_m}{m^3}\frac{1}{\left(\frac{a^2}{b^2}+\frac{i^2}{m^2}\right)^2} = 0 \tag{2.3.7}$$

and

$$\frac{4pb^2}{\pi^3}\frac{1}{i^4}\left(\frac{\beta_i}{\cosh^2(\beta_i)}-\tanh(\beta_i)\right)-\frac{F_i}{i}\left(\tanh(\beta_i)+\frac{\beta_i}{\cosh^2(\beta_i)}\right)-\frac{8ib}{\pi a}\sum_{m=1,3,5,\cdots}^{\infty}\frac{E_m}{m^3}\frac{1}{\left(\frac{b^2}{a^2}+\frac{i^2}{m^2}\right)^2}=0 \tag{2.3.8}$$

where $\alpha_i=\frac{i\pi b}{2a}$, $\beta_i=\frac{i\pi a}{2b}$ and i ranges over the limits of m, *ie*, $i=1,3,5,\cdots$. (The accuracy increases with the number of terms.) From (2.3.7) and (2.3.8) a system of equations may be written from which E_i and F_i can be found. If m and i run to 5, for example, they will take the form

$$\begin{aligned}
&(\cdot)E_1+0E_3+0E_5+(\cdot)F_1+(\cdot)F_3+(\cdot)F_5=(\cdot)\\
&0E_1+(\cdot)E_3+0E_5+(\cdot)F_1+(\cdot)F_3+(\cdot)F_5=(\cdot)\\
&0E_1+0E_3+(\cdot)E_5+(\cdot)F_1+(\cdot)F_3+(\cdot)F_5=(\cdot)\\
&(\cdot)E_1+(\cdot)E_3+(\cdot)E_5+(\cdot)F_1+0F_3+0F_5=(\cdot)\\
&(\cdot)E_1+(\cdot)E_3+(\cdot)E_5+0F_1+(\cdot)F_3+0F_5=(\cdot)\\
&(\cdot)E_1+(\cdot)E_3+(\cdot)E_5+0F_1+0F_3+(\cdot)F_5=(\cdot)
\end{aligned} \tag{2.3.9}$$

where $(\cdot)$ are quantities evident from (2.3.7) and (2.3.8) and depend only on a, b, and p. Once E_m and F_m have been determined, the are used in

$$w_1(x,y)=-\frac{a^2}{2\pi^2 D}\sum_{m=1,3,5,\cdots}^{\infty}E_m\frac{(-1)^{(m-1)/2}}{m^2\cosh(\alpha_m)}\cos\left(\frac{m\pi x}{a}\right)\left[\frac{m\pi y}{a}\sinh\left(\frac{m\pi y}{a}\right)-\alpha_m\tanh(\alpha_m)\cosh\left(\frac{m\pi y}{a}\right)\right] \tag{2.3.10}$$

$$w_2(x,y)=-\frac{b^2}{2\pi^2 D}\sum_{m=1,3,5,\cdots}^{\infty}F_m\frac{(-1)^{(m-1)/2}}{m^2\cosh(\beta_m)}\cos\left(\frac{m\pi y}{b}\right)\left[\frac{m\pi x}{b}\sinh\left(\frac{m\pi x}{b}\right)-\beta_m\tanh(\beta_m)\cosh\left(\frac{m\pi x}{b}\right)\right] \tag{2.3.11}$$

to determine the deflection due to the applied moment. The deflection of the simply-supported plate is given by

$$w(x,y)_{s\text{-}s} = \frac{4pa^4}{\pi^5 D} \sum_{m=1,3,5,\ldots}^{\infty} \frac{(-1)^{(m-1)/2}}{m^5} \cos\left(\frac{m\pi x}{a}\right)\left[1 - \frac{\alpha_m \tanh(\alpha_m) + 2}{2\cosh(\alpha_m)} \cosh\left(\frac{m\pi y}{a}\right) + \frac{1}{2\cosh(\alpha_m)} \frac{m\pi y}{a} \sinh\left(\frac{m\pi y}{a}\right)\right] \quad (2.3.12)$$

The deflection of the plate built-in on all four sides is given by superposing the three,

$$w(x,y)_{built\text{-}in} = w(x,y)_{s\text{-}s} + w_1(x,y) + w_2(x,y) \quad . \quad (2.3.13)$$

To determine strain, needed later to calculate the effect of diaphragm pressure on the channel waveguide, the second partial derivative of (2.3.13) is required. This is assembled from the partials of (2.3.10)-(2.3.12),

$$\frac{\partial^2 w_1(x,y)}{\partial x^2} = \frac{1}{2D} \sum_{m=1,3,5,\ldots}^{\infty} E_m \frac{(-1)^{(m-1)/2}}{\cosh(\alpha_m)} \cos\left(\frac{m\pi x}{a}\right)\left[\frac{m\pi y}{a} \sinh\left(\frac{m\pi y}{a}\right) - \alpha_m \tanh(\alpha_m) \cosh\left(\frac{m\pi y}{a}\right)\right] \quad (2.3.14)$$

$$\frac{\partial^2 w_2(x,y)}{\partial x^2} = -\frac{1}{2D} \sum_{m=1,3,5,\ldots}^{\infty} F_m \frac{(-1)^{(m-1)/2}}{\cosh(\beta_m)} \cos\left(\frac{m\pi y}{b}\right)\left\{\frac{m\pi x}{b} \sinh\left(\frac{m\pi x}{b}\right) + \left[2 - \beta_m \tanh(\beta_m)\right] \cosh\left(\frac{m\pi x}{b}\right)\right\} \quad (2.3.15)$$

$$\frac{\partial^2 w(x,y)_{s\text{-}s}}{\partial x^2} = -\frac{4pa^2}{\pi^3 D} \sum_{m=1,3,5,\ldots}^{\infty} \frac{(-1)^{(m-1)/2}}{m^3} \cos\left(\frac{m\pi x}{a}\right)\left[1 - \frac{\alpha_m \tanh(\alpha_m) + 2}{2\cosh(\alpha_m)} \cosh\left(\frac{m\pi y}{a}\right) + \frac{1}{2\cosh(\alpha_m)} \frac{m\pi y}{a} \sinh\left(\frac{m\pi y}{a}\right)\right] \quad (2.3.16)$$

and

$$\frac{\partial^2 w_1(x,y)}{\partial y^2} = -\frac{1}{2D} \sum_{m=1,3,5,\ldots}^{\infty} E_m \frac{(-1)^{(m-1)/2}}{\cosh(\alpha_m)} \cos\left(\frac{m\pi x}{a}\right)\left\{\frac{m\pi y}{a} \sinh\left(\frac{m\pi y}{a}\right) + \left[2 - \alpha_m \tanh(\alpha_m)\right] \cosh\left(\frac{m\pi y}{a}\right)\right\} \quad (2.3.17)$$

$$\frac{\partial^2 w_2(x,y)}{\partial y^2} = \frac{1}{2D} \sum_{m=1,3,5,\ldots}^{\infty} F_m \frac{(-1)^{(m-1)/2}}{\cosh(\beta_m)} \cos\left(\frac{m\pi y}{b}\right)\left[\frac{m\pi x}{b}\sinh\left(\frac{m\pi x}{b}\right) - \beta_m \tanh(\beta_m)\cosh\left(\frac{m\pi x}{b}\right)\right] \tag{2.3.18}$$

$$\frac{\partial^2 w(x,y)_{s\text{-}s}}{\partial y^2} = \frac{2pa^2}{\pi^3 D} \sum_{m=1,3,5,\ldots}^{\infty} \frac{(-1)^{(m-1)/2}}{m^3 \cosh(\alpha_m)} \cos\left(\frac{m\pi x}{a}\right)\left[\alpha_m \tanh(\alpha_m)\cosh\left(\frac{m\pi y}{a}\right) + \frac{m\pi y}{a}\sinh\left(\frac{m\pi y}{a}\right)\right] \tag{2.3.19}$$

From the previous derivation for the beam,

$$\epsilon_x(x,y,z) = -\, z\frac{\partial^2 w}{\partial x^2} \quad . \tag{2.3.20}$$

The presence of a nonzero second partial with respect to y will similarly produce strain,

$$\epsilon_y(x,y,z) = -\, z\frac{\partial^2 w}{\partial y^2} \quad . \tag{2.3.21}$$

Furthermore, as with the beam,

$$\sigma_z = 0 \quad . \tag{2.3.22}$$

Substituting (2.3.20) and (2.3.22) into (2.2.4a), and (2.3.21) and (2.3.22) into (2.2.4b) gives

$$\epsilon_x(x,y,z) = -\, z\frac{\partial^2 w}{\partial x^2} = \frac{1}{E}(\sigma_x - \nu\sigma_y) \tag{2.3.23}$$

and

$$\epsilon_y(x,y,z) = -\, z\frac{\partial^2 w}{\partial y^2} = \frac{1}{E}(\sigma_y - \nu\sigma_x) \quad . \tag{2.3.24}$$

Solving these equations simultaneously gives

$$\sigma_x(x,y,z) = \frac{-zE}{1-\nu^2}\left(\frac{\partial^2 w}{\partial x^2} + \nu\frac{\partial^2 w}{\partial y^2}\right) \tag{2.3.25}$$

and
$$\sigma_y(x,y,z) = \frac{-zE}{1-\nu^2}\left(\nu\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}\right) . \tag{2.3.26}$$

Inserting (2.3.25) and (2.3.26) into (2.2.4c) gives

$$\epsilon_z(x,y,z) = \frac{\nu z}{1-\nu}\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}\right) . \tag{2.3.27}$$

To allow comparison of the plate and beam models, the deflection, stress, and strain in a plate composed of silicon dioxide and having the same dimensions as a diaphragm incorporated in the ring resonator pressure sensor has been computed and is given Table 2-2 - Table 2-7. Because of symmetry, only $\frac{1}{4}$ of the diaphragm data is shown. The core center of the waveguides fabricated for this particular sensor were located underneath 0.677 μm of film. To accurately determine the stress and strain in the waveguide core, once the second derivatives were calculated [equations (2.3.14)-(2.3.19)], z in (2.3.20), (2.3.21), and (2.3.25)-(2.3.27) was set to the distance from the diaphragm midplane to the core.

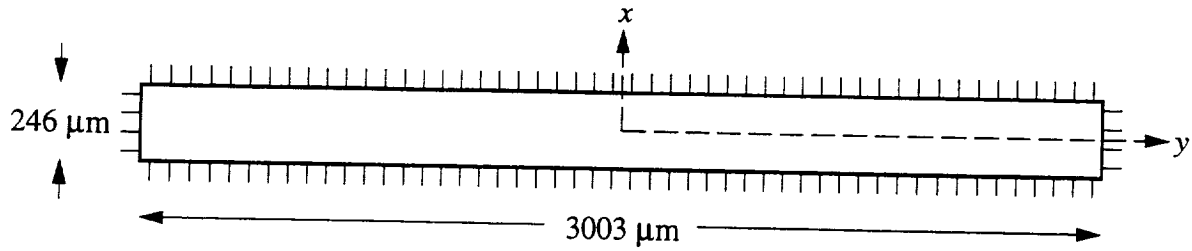


Fig. 2-10 Top view of plate built-in on four sides used in example calculation. Plate is constructed of SiO_2 ($E = 7.3\times10^{10}$ Pa [4], $\nu = 0.17$ [5]), has thickness $t = 13.87$ μm, and $p = 4.1\times10^5$ Pa (60 pisg) air pressure is applied to its back. The top left entry in tables (2-2)-(2-7) describes the behavior of the top left corner of the plate, while the bottom-right entry goes with the center. The spacing of the data is 75.075 μm in y (between columns) and 6.15 μm in x (between rows). Plate deflection, stress and strain are symmetric about the x and y axis of Fig. 2-8. Tables 2-3 and 2-7 were calculated at the waveguide core; to determine top surface values, multiply each entry by $\frac{t/2}{6.258\ \mu m}$. Although the indices i and m used in (2.3.10)-(2.3.12) and (2.3.14)-(2.3.19) ranged over 1, 3, ...27, some ripple is visible in Table 2-5.

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0	4	8	8	9	8	8	9	8	8	9	8	8	8	8	8	8	8	8	8	8
0	9	16	17	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18
0	15	27	30	31	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
0	22	39	44	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45
0	29	53	59	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61
0	37	67	76	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78
0	46	82	94	96	96	96	96	96	96	96	96	96	96	96	96	96	96	96	96	96
0	54	98	111	114	114	114	114	114	114	114	114	114	114	114	114	114	114	114	114	114
0	61	112	128	132	132	132	132	132	132	132	132	132	132	132	132	132	132	132	132	132
0	69	127	145	149	149	149	149	149	149	149	149	149	149	149	149	149	149	149	149	149
0	76	140	161	165	165	165	165	165	165	165	165	165	165	165	165	165	165	165	165	165
0	83	152	175	180	181	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180
0	88	164	189	194	194	194	194	194	194	194	194	194	194	194	194	194	194	194	194	194
0	93	173	200	206	206	206	206	206	206	206	206	206	206	206	206	206	206	206	206	206
0	98	181	210	216	216	216	216	216	216	216	216	216	216	216	216	216	216	216	216	216
0	101	188	217	224	224	224	224	224	223	224	223	224	224	223	224	224	223	224	224	223
0	103	193	223	230	230	229	229	229	229	229	229	229	229	229	229	229	229	229	229	229
0	105	195	226	233	233	233	233	233	233	233	233	233	233	233	233	233	233	233	233	233
0	105	196	228	234	235	234	234	234	234	234	234	234	234	234	234	234	234	234	234	234 ...
																				⋮

Table 2-2 Deflection (w) in nm of diaphragm in Fig. 2-10.

0	-30	-51	-58	-58	-58	-59	-58	-58	-58	-58	-58	-58	-58	-58	-58	-58	-58	-58	-58	-58
1	-25	-43	-49	-50	-50	-50	-49	-50	-50	-50	-50	-50	-50	-50	-50	-50	-50	-50	-50	-50
0	-20	-35	-41	-41	-42	-42	-41	-42	-42	-41	-42	-42	-42	-42	-42	-42	-42	-42	-42	-42
0	-15	-28	-33	-34	-34	-34	-34	-34	-34	-34	-34	-34	-34	-34	-34	-34	-34	-34	-34	-34
-1	-11	-22	-26	-27	-27	-27	-27	-27	-27	-27	-27	-27	-27	-27	-27	-27	-27	-27	-27	-27
-1	-8	-16	-19	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20
-2	-5	-10	-13	-14	-14	-14	-14	-14	-14	-14	-14	-14	-14	-14	-14	-14	-14	-14	-14	-14
-2	-2	-5	-7	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
-3	1	0	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
-4	3	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
-4	5	7	8	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
-5	6	11	12	12	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
-5	8	14	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
-5	9	16	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18
-6	10	19	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21
-6	11	20	23	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24
-6	12	22	25	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26
-7	12	23	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
-7	13	24	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28
-7	13	24	28	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29
-7	13	25	28	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29 ...
																				⋮

Table 2-3 σ_x in MPa of diaphragm in Fig. 2-10.

0	-5	-9	-10	-10	-10	-10	-10	-10	-10	-10	-10	-10	-10	-10	-10	-10	-10	-10	-10	-10
3	-4	-7	-8	-8	-8	-9	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
2	-3	-6	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7
-1	-2	-4	-6	-6	-6	-6	-6	-6	-6	-6	-6	-6	-6	-6	-6	-6	-6	-6	-6	-6
-4	-1	-3	-4	-4	-5	-5	-4	-5	-5	-5	-5	-5	-5	-5	-5	-5	-5	-5	-5	-5
-7	0	-2	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
-11	1	0	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
-14	2	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-17	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-21	3	3	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
-24	4	4	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
-27	4	5	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
-29	5	6	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
-32	5	7	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
-34	5	7	5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
-36	6	8	5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
-37	6	8	6	5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
-38	6	9	6	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
-39	6	9	6	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
-40	6	9	7	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
-40	6	9	7	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5 ...
																				⋮

Table 2-4 σ_y in MPa of diaphragm in Fig. 2-10.

0	-402	-682	-775	-771	-776	-778	-770	-775	-776	-771	-776	-775	-772	-776	-774	-772	-776	-773	-773	-776
0	-330	-573	-658	-659	-663	-664	-658	-662	-662	-659	-662	-661	-659	-662	-661	-660	-662	-660	-660	-662
0	-264	-471	-548	-551	-555	-556	-551	-554	-554	-552	-554	-554	-552	-554	-553	-552	-554	-553	-553	-555
0	-206	-377	-445	-450	-453	-454	-451	-452	-453	-451	-453	-452	-451	-453	-452	-451	-453	-452	-451	-453
0	-153	-291	-348	-355	-357	-358	-355	-356	-357	-355	-357	-356	-355	-357	-356	-356	-357	-356	-356	-357
0	-106	-211	-258	-265	-267	-267	-265	-266	-267	-265	-266	-266	-266	-266	-266	-266	-267	-266	-266	-267
0	-64	-139	-174	-181	-182	-183	-181	-182	-182	-182	-182	-182	-182	-182	-182	-182	-182	-182	-182	-182
0	-28	-72	-97	-102	-104	-104	-103	-104	-104	-103	-104	-104	-103	-104	-103	-103	-104	-103	-103	-104
0	5	-12	-26	-30	-31	-31	-31	-31	-31	-31	-31	-31	-31	-31	-31	-31	-31	-31	-31	-31
0	33	43	39	37	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36
0	58	92	98	98	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97
0	79	136	151	153	152	152	152	152	152	152	152	152	152	152	152	152	152	152	152	152
0	98	174	198	202	202	201	201	201	201	201	201	201	201	201	201	201	201	201	201	201
0	113	208	240	245	245	245	245	245	245	245	245	245	245	245	245	245	245	245	245	245
0	126	236	275	283	283	283	282	283	283	282	283	283	282	283	283	282	283	283	282	283
0	137	260	305	314	315	315	314	315	315	314	315	315	314	315	314	314	315	314	314	315
0	146	280	330	340	341	341	341	341	341	341	341	341	341	341	341	341	341	341	341	341
0	152	295	349	360	362	361	361	361	361	361	361	361	361	361	361	361	361	361	361	361
0	157	306	363	375	376	376	375	376	376	375	376	375	375	376	375	375	376	375	375	376
0	160	312	371	384	385	385	384	384	384	384	384	384	384	384	384	384	384	384	384	384
0	161	314	374	386	388	388	387	387	387	387	387	387	387	387	387	387	387	387	387	387 ...
																				⋮

Table 2-5 ϵ_x in microstrain ($1\ \mu\epsilon$ = one microstrain $\rightarrow \epsilon = 10^{-6}$) of diaphragm in Fig. 2-10.

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
39	2	1	-1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	5	3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-11	8	6	0	1	0	-1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
-47	12	10	1	2	0	-1	1	0	0	1	0	0	0	0	0	0	0	0	0	0
-93	16	14	2	2	0	-1	1	0	0	1	0	0	0	0	0	0	0	0	0	0
-141	21	19	4	2	0	-1	1	0	0	1	0	0	0	0	0	0	0	0	0	0
-185	25	24	6	3	0	-1	1	0	0	1	0	0	0	0	0	0	0	0	0	0
-231	29	30	8	3	0	-1	1	0	0	1	0	0	0	0	0	0	0	0	0	0
-276	34	35	10	3	0	-1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
-315	38	40	12	4	0	-1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
-355	41	45	14	4	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-390	45	50	16	4	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-421	48	54	18	5	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-450	51	58	19	5	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-474	53	62	21	5	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-493	55	65	22	6	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-510	56	67	23	6	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-520	57	69	24	6	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-526	58	70	24	6	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-530	58	70	25	6	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0 ...
																				⋮

Table 2-6 ϵ_y in microstrain of diaphragm in Fig. 2-10.

0	82	140	159	158	159	159	158	159	159	158	159	159	158	159	159	158	159	158	158	159
-8	67	117	135	135	136	136	135	136	136	135	136	135	135	136	135	135	136	135	135	136
-5	53	96	112	113	114	114	113	114	114	113	114	113	113	114	113	113	114	113	113	114
2	40	76	91	92	93	93	92	93	93	92	93	93	92	93	93	92	93	93	92	93
10	29	57	71	72	73	73	73	73	73	73	73	73	73	73	73	73	73	73	73	73
19	18	40	52	54	55	55	54	55	55	54	55	55	54	55	54	54	55	54	54	55
29	9	24	35	37	37	38	37	37	37	37	37	37	37	37	37	37	37	37	37	37
38	1	10	19	20	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21
47	-7	-4	4	6	6	7	6	6	6	6	6	6	6	6	6	6	6	6	6	6
56	-14	-16	-10	-8	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7
65	-20	-27	-23	-21	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20
73	-25	-37	-34	-32	-31	-31	-31	-31	-31	-31	-31	-31	-31	-31	-31	-31	-31	-31	-31	-31
80	-29	-46	-44	-42	-41	-41	-41	-41	-41	-41	-41	-41	-41	-41	-41	-41	-41	-41	-41	-41
86	-33	-54	-53	-51	-50	-50	-50	-50	-50	-50	-50	-50	-50	-50	-50	-50	-50	-50	-50	-50
92	-36	-60	-60	-59	-58	-58	-58	-58	-58	-58	-58	-58	-58	-58	-58	-58	-58	-58	-58	-58
97	-39	-66	-67	-65	-65	-64	-64	-64	-64	-64	-64	-64	-64	-64	-64	-64	-64	-64	-64	-64
101	-41	-71	-72	-71	-70	-70	-70	-70	-70	-70	-70	-70	-70	-70	-70	-70	-70	-70	-70	-70
104	-43	-74	-76	-75	-74	-74	-74	-74	-74	-74	-74	-74	-74	-74	-74	-74	-74	-74	-74	-74
106	-44	-77	-79	-78	-77	-77	-77	-77	-77	-77	-77	-77	-77	-77	-77	-77	-77	-77	-77	-77
108	-45	-78	-81	-80	-79	-79	-79	-79	-79	-79	-79	-79	-79	-79	-79	-79	-79	-79	-79	-79
108	-45	-79	-82	-80	-80	-79	-79	-79	-79	-79	-79	-79	-79	-79	-79	-79	-79	-79	-79	-79 ...
																				⋮

Table 2-7 ϵ_z in microstrain of diaphragm in Fig. 2-10.

Distance from top plate edge or built-in edge of beam (μm)	Deflection (nm)		σ_x (MPa)		ϵ_x (μϵ)		ϵ_z (μϵ)	
	beam	plate	beam	plate	beam	plate	beam	plate
0	0	0	-58	-58	-774	-776	159	159
6.15	2	2	-50	-50	-661	-662	135	136
12.3	8	8	-42	-42	-553	-555	113	114
18.45	18	18	-34	-34	-452	-453	93	93
24.6	30	30	-27	-27	-356	-357	73	73
30.75	45	45	-20	-20	-266	-267	55	55
36.9	61	61	-14	-14	-182	-182	37	37
43.05	78	78	-8	-8	-104	-104	21	21
49.2	96	96	-2	-2	-31	-31	6	6
55.35	114	114	3	3	36	36	-7	-7
61.5	132	132	7	7	97	97	-20	-20
67.65	149	149	11	11	152	152	-31	-31
73.8	165	165	15	15	201	201	-41	-41
79.95	180	180	18	18	245	245	-50	-50
86.1	194	194	21	21	283	283	-58	-58
92.25	206	206	24	24	314	315	-64	-64
98.4	216	216	26	26	341	341	-70	-70
104.55	224	223	27	27	361	361	-74	-74
110.7	229	229	28	28	375	376	-77	-77
116.85	233	233	29	29	384	384	-79	-79
123	234	234	29	29	387	387	-79	-79

Table 2-8 Deflection stress and strain at $z = 6.258$ μm and $y = 0$ in the plate of Fig. 2-10 compared to that in a beam having the same composition and thickness and at the same depth. The beam's length is 246 μm.

4 REFERENCES

[1] E. Volterra and J. H. Gaines, *Advanced Strength of Materials*. Englewood Cliffs, N. J.: Prentice-Hall, pp. 471-472, 1971.

[2] A. K. Mal and S. J. Singh, *Deformation of Elastic Solids*. Englewood Cliffs, NJ: Prentice-Hall, p. 12, 1991.

[3] S. Timoshenko and S. Woinowsky-Krieger, *Theory of Plates and Shells*, 2nd edition: McGraw-Hill, p. 197, 1959.

[4] K. E. Petersen, "Silicon as a Mechanical Material," *Proc. IEEE*, vol. 70, pp. 420-457, 1982.

[5] G. B. Hocker, "Fiber-Optic Sensing of Pressure and Temperature," *Appl. Opt.*, vol. 18, pp. 1445-1448, 1979.

III Modeling of Integrated Optic Waveguides and Devices

1 INTRODUCTION

Channel waveguide design was carried out using the effective index technique. Determined were bending loss, directional coupler transfer length, layer thickness, etch depth, and the path length change due to the pressure-induced photoelastic effect. The easily-implemented effective index technique has been shown to agree with other more involved methods for waveguide designs considered in this study [1, 2]. An additional advantage is that the technique can be reversed, allowing a determination of the waveguide structure which will have a desired property.

To apply this method, the effective index of both the planar waveguide which would exist if the channel were infinitely wide, and the planar region adjacent to the channel must be determined. The planar effective indices are calculated numerically with relations derived from Maxwell's equations. The notation used in these derivations follows that used by Lee [3], and is given in Table 3-1.

Symbol	Meaning	MKS unit
$\mathbf{E}$	Vector form of electric field	V/m
$\mathbf{H}$	Vector form of magnetic field	A/m
μ	permeability	H/m
ε	permittivity	F/m
ω	optical frequency	Hz
k_x, k_y, k_z	wavevector in x, y, z	m^{-1}
$k_0 = 2\pi/\lambda_0$	free space wavenumber	m^{-1}
λ_0	free space wavelength	m
n, n_1, n_2, n_3, n_4	refractive index	unitless
α_{1x}	field decay constant in the x direction in medium 1	1/m
k_{2x}	field constant in the x direction in medium 2	1/m
c_0	free space speed of light	m
z	direction of propagation in waveguide coordinates. direction normal to wafer in plate coordinates	
E_x, E_y, E_z	Electric field polarized in x, y, z	V/m
n_{e1}	effective refractive index of planar film outside the channel	unitless
n_{e2}	effective refractive index of the planar guide which would result if the channel were infinitely wide	unitless
n_{et}	total or channel effective index	unitless
k_{zt}	total or channel wavenumber (k_z for channel) $k_{zt} = k_0 n_{et}$	1/m
w	width of channel	m
p, m	mode order (0 for fundamental)	unitless
p_{11}, p_{12}	strain photoelastic constants	unitless
t	three layer guide thickness	m
t_1, t_2, t_2'	four layer (overlayer guide) cladding, guide layer under ridge and guide layer away from ridge thickness	m
t_g, t_i	four layer (substrate coupling planar guide problem) guide and isolation layer thickness	m
P	power density	W/m^2
a	offset between two bent waveguides	m
Δk	difference between the isolated channel propagation constants (for directional coupler)	1/m
C, C_{AB}, C_{BA}	coupling coefficients	1/m
κ	power coupling coefficient	unitless
a_A, a_B	field amplitude in channels A and B	V/m
L_{DC}	normalized coupling length. In synchronous couplers, if equal to $\pi/2$, 100% coupling occurs	radians
L_{PT}	power transfer length. The distance required for complete power transfer in directional couplers	m
L_{coh}	coherence length	m
ϵ	strain	unitless
ϕ	phase change (negative for increasing phase delay)	radians

Table 3-1 Symbols used in this chapter.

2 THE WAVE EQUATION

Maxwell's equations in phasor form in an isotropic medium are:

$$\nabla\times\mathbf{E} = -j\omega\mu\mathbf{H} \quad , \tag{3.2.1}$$

$$\nabla\times\mathbf{H} = \sigma + j\omega\varepsilon\mathbf{E} \quad , \tag{3.2.2}$$

$$\nabla\cdot\mathbf{D} = \nabla\cdot\varepsilon\mathbf{E} = \rho \quad , \tag{3.2.3}$$

and

$$\nabla\cdot\mathbf{B} = 0 \quad . \tag{3.2.4}$$

Substituting (3.2.1) into (3.2.2) gives

$$\nabla\times\left(\frac{\nabla\times\mathbf{E}}{-j\omega\mu}\right) = \sigma + j\omega\varepsilon\mathbf{E} \quad . \tag{3.2.5}$$

In a nonconducting, nonmagnetic, and charge-free medium, σ=0, μ=μ_0, and ρ=0. Rewriting (3.2.5), the equation is obtained

$$\nabla\times(\nabla\times\mathbf{E}) = \omega^2\mu_0\varepsilon\mathbf{E} \tag{3.2.6}$$

which, using the identity, $\nabla\times(\nabla\times\mathbf{E}) = \nabla(\nabla\cdot\mathbf{E}) - \nabla^2\mathbf{E}$, becomes

$$\nabla^2\mathbf{E} - \nabla(\nabla\cdot\mathbf{E}) + \omega^2\mu_0\varepsilon\mathbf{E} = 0 \quad . \tag{3.2.7}$$

Using another identity $\nabla\cdot(\varepsilon\mathbf{E}) = \mathbf{E}\cdot\nabla\varepsilon + \varepsilon\nabla\cdot\mathbf{E}$, where ε is a scalar field, (*i.e.* $\varepsilon = \varepsilon(x,y,z)$), and (3.2.3), the relationship can be established [4] that

$$-\mathbf{E}\cdot\frac{\nabla\varepsilon}{\varepsilon} = \nabla\cdot\mathbf{E} \quad . \tag{3.2.8}$$

Equation (3.2.8), which contains $\frac{\nabla\varepsilon}{\varepsilon}$ $\left(=\frac{1}{\varepsilon}\frac{\partial\varepsilon}{\partial x}\hat{x} + \frac{1}{\varepsilon}\frac{\partial\varepsilon}{\partial y}\hat{y} + \frac{1}{\varepsilon}\frac{\partial\varepsilon}{\partial z}\hat{z}\right)$, can be rewritten by noting that $\frac{\partial}{\partial x}\ln[f(x)] = \frac{1}{f(x)}\frac{\partial f(x)}{\partial x}$, as $-\mathbf{E}\cdot\nabla\ln\varepsilon = \nabla\cdot\mathbf{E}$. Equation (3.2.7) then becomes

$$\nabla^2\mathbf{E} + \nabla(\mathbf{E}\cdot\nabla\ln\varepsilon) + \omega^2\mu_0\varepsilon\mathbf{E} = 0 \quad . \tag{3.2.9}$$

This is the vector wave equation. In the uniform layer waveguide, the gradient of ε is zero everywhere except at the boundaries, where the discontinuity is taken into account by the boundary conditions [5]. Thus (3.2.9) can be reduced to the scalar wave equation

and solved in each homogeneous region separately:

$$\nabla^2\mathbf{E} + \omega^2\mu_0\varepsilon\mathbf{E} = \frac{\partial^2\mathbf{E}}{\partial x^2} + \frac{\partial^2\mathbf{E}}{\partial y^2} + \frac{\partial^2\mathbf{E}}{\partial z^2} + \omega^2\mu_0\varepsilon\mathbf{E} = 0 \quad . \tag{3.2.10}$$

A solution to (3.2.10) for a plane wave is [6]

$$E_r = A\exp[-j(k_x x + k_y y + k_z z)] \quad , \tag{3.2.11}$$

where A is the amplitude, r indicates its direction of polarization, and $\mathbf{k} = k_x\hat{x} + k_y\hat{y} + k_z\hat{z}$ is the wave vector parallel to the direction of propagation. Substituting (3.2.11) into (3.2.10) gives

$$k_x^2 + k_y^2 + k_z^2 = \mathbf{k}^2 = \omega^2\varepsilon\mu_0 \quad . \tag{3.2.12}$$

Using the relationships

$$\varepsilon = \varepsilon_0\varepsilon_r \quad , \tag{3.2.13a}$$

$$\frac{1}{c_0^2} = \varepsilon_0\mu_0 \quad , \tag{3.2.13b}$$

$$n^2 = \varepsilon_r \quad , \tag{3.2.13c}$$

and

$$c_0 = \frac{\omega}{2\pi}\lambda_0 \quad , \tag{3.2.13d}$$

where c_0 is the speed of light, ε_r and ε_0 are the relative and free space permittivities, and λ_0 is the vacuum wavelength, (3.2.12) can be written:

$$k_x^2 + k_y^2 + k_z^2 = \mathbf{k}^2 = \left(\frac{2\pi n}{\lambda_0}\right)^2 \quad . \tag{3.2.14}$$

This relationship will be applied separately to each of the regions in the planar waveguide.

3 THREE-LAYER PLANAR WAVEGUIDE

Consider the planar waveguide in Fig. 3-1 where $n_1, n_3 < n_2$. The coordinate system is arranged so that propagation is in z, $x = 0$ is in the middle of n_2, and there is no field variation in y. A light ray incident on the n_2-n_1 and n_2-n_3 interfaces at greater than the critical angle will be totally reflected. In order for a stable field pattern (a mode) to

exist, the bounce angle of the ray must be such that its phase at points A and B, which are located at the same x-position, is an integral multiple of 2π radians. Although the properties of a planar waveguide can be derived by analyzing this phase shift, it is more easily approached by solving the wave equation.

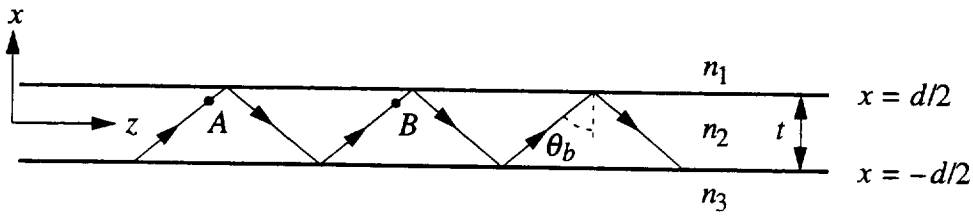


Fig. 3-1 Planar waveguide configuration. The phase shift from A to B must be an intergal multiple of 2π radians. θ_b is the bounce angle.

Although light propagating in polarizations with either **E** or **H** in y will be able to satisfy the criterion of Fig. 3-1, they will have different bounce angles and therefore different phase velocities because the phase shift upon reflection at the film interfaces will be different. If both polarizations are excited, the polarization state of light at a position z will depend on the length of the waveguide. Treating the case where **E** is polarized in y, so that $k_y = 0$, the phase will only vary in z, while the amplitude will have an x dependence $E(x)$

$$E_y(x,z) = E(x)\exp(-jk_z z) \quad . \tag{3.3.1}$$

Inserting (3.3.1) into (3.3.10), gives

$$\frac{d^2E(x)}{dx^2} + \left(\omega^2\mu_0\varepsilon - k_z^2\right)E(x) = 0 \quad . \tag{3.3.2}$$

As before, using $(2\pi n/\lambda_0)^2 = \omega^2\mu_0\varepsilon$,

$$\frac{d^2E(x)}{dx^2} + \left[\left(\frac{2\pi n}{\lambda_0}\right)^2 - k_z^2\right]E(x) = 0 \quad . \tag{3.3.3}$$

Equation (3.3.3) can be simplified by defining

$$k_0 = \frac{2\pi}{\lambda_0} \quad , \tag{3.3.4}$$

and
$$k_z = k_0 n_e \quad , \tag{3.3.5}$$

giving

$$\frac{d^2E(x)}{dx^2} + k_0^2\left(n^2 - n_e^2\right)E(x) = 0 \quad . \tag{3.3.6}$$

According to the solution to (3.3.6), $E(x)$ is sinusoidal when $n^2 - n_e^2 > 0$ and exponential otherwise. To find $E(x)$, equation (3.3.6) is written for each of the layers and the boundary condition that tangential **E** and **H** fields are continuous across the interfaces applied. E_y must be a decaying exponential in the semi-infinite claddings since the amount of power in the mode is finite, and sinusoidal in the core in order to match the electric field derivative (proportional to the **H** field). Thus, **E** is given by [7]

$$E_y(x,z) = \begin{Bmatrix} A_1 \exp(-\alpha_{1x} x) \\ A_2 \cos(k_{2x} x + \psi) \\ A_3 \exp(+\alpha_{3x} x) \end{Bmatrix} \exp(-jk_z z) \qquad \begin{matrix} x > d/2 \\ -d/2 < x < d/2 \\ x < -d/2 \end{matrix} \quad , \tag{3.3.7}$$

where

$$\alpha_{1x} = k_0 \sqrt{n_e^2 - n_1^2} \quad , \tag{3.3.8a}$$

$$k_{2x} = k_0 \sqrt{n_2^2 - n_e^2} \quad , \tag{3.3.8b}$$

and
$$\alpha_{3x} = k_0 \sqrt{n_e^2 - n_3^2} \quad . \tag{3.3.8c}$$

Using the definition $\nabla \times \mathbf{E} = \left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\right)\hat{x} + \left(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}\right)\hat{y} + \left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right)\hat{z}$ and (3.2.1), the **H** fields are found to be:

$$H_x(x,z) = \frac{-k_z}{\omega\mu} \begin{Bmatrix} A_1 \exp(-\alpha_{1x} x) \\ A_2 \cos(k_{2x} x + \psi) \\ A_3 \exp(+\alpha_{3x} x) \end{Bmatrix} \exp(-jk_z z) \qquad \begin{matrix} x > d/2 \\ -d/2 < x < d/2 \\ x < -d/2 \end{matrix} \tag{3.3.9}$$

$$H_z(x,z) = \frac{j}{\omega\mu} \begin{Bmatrix} -A_1 \alpha_{1x} \exp(-\alpha_{1x} x) \\ -A_2 k_{2x} \sin(k_{2x} x + \psi) \\ A_3 \alpha_{3x} \exp(+\alpha_{3x} x) \end{Bmatrix} \exp(-jk_z z) \qquad \begin{matrix} x > d/2 \\ -d/2 < x < d/2 \\ x < -d/2 \end{matrix} \quad . \tag{3.3.10}$$

At $x = d/2$, the tangential **H** (H_z) and **E** fields are continuous:

$$-A_1 \alpha_{1x} \exp(-\alpha_{1x} d/2) = -A_2 k_{2x} \sin(k_{2x} d/2 + \psi) \tag{3.3.11}$$

$$A_1 \exp(-\alpha_{1x} d/2) = A_2 \cos(k_{2x} d/2 + \psi) \quad . \tag{3.3.12}$$

Taking the ratio of (3.3.11) and (3.3.12) gives

$$\frac{\alpha_{1x}}{k_{2x}} = \tan(k_{2x} d/2 + \psi) \quad . \tag{3.3.13}$$

Taking the inverse tangent of both sides,

$$\tan^{-1}\left(\frac{\alpha_{1x}}{k_{2x}}\right) = k_{2x} d/2 + \psi \pm m\pi \tag{3.3.14}$$

is found, where m is an integer. Similarly, at $x = -d/2$:

$$A_3 \alpha_{3x} \exp(-\alpha_{3x} d/2) = -A_2 k_{2x} \sin(-k_{2x} d/2 + \psi) \tag{3.3.15}$$

$$A_3 \exp(-\alpha_{3x} d/2) = A_2 \cos(-k_{2x} d/2 + \psi) \tag{3.3.16}$$

$$\tan^{-1}\left(\frac{\alpha_{3x}}{k_{2x}}\right) = k_{2x} d/2 - \psi \pm q\pi \tag{3.3.17}$$

where q is an integer. Adding (3.3.13) and (3.3.17), and noting that $t = d$ gives

$$-k_{2x} t + \tan^{-1}\left(\frac{\alpha_{1x}}{k_{2x}}\right) + \tan^{-1}\left(\frac{\alpha_{3x}}{k_{2x}}\right) + p\pi = 0 \quad , \tag{3.3.18}$$

where p, an integer (0, 1, 2..), is interpreted as the mode order, with the fundamental mode having $p = 0$. The sign in front of $p\pi$ is chosen to be the opposite of $k_{2x}t$ in (3.3.18) because the waveguide will support a higher order mode having the same effective index as a lower order one for an appropriate, but larger, value of t..

For the TM mode, a parallel derivation will result in the scalar wave equation,

$$\frac{\partial^2 \mathbf{H}}{\partial x^2} + \frac{\partial^2 \mathbf{H}}{\partial y^2} + \frac{\partial^2 \mathbf{H}}{\partial z^2} + \omega^2 \mu_0 \varepsilon \mathbf{H} = 0 \tag{3.3.19}$$

with the assumed field given by

$$H_y(x,z) = H(x) \exp(-jk_z z) \quad . \tag{3.3.20}$$

Solving the similar differential equation

$$\frac{d^2 H(x)}{dx^2} + k_0^2\left(n^2 - n_e^2\right) H(x) = 0 \tag{3.3.21}$$

will result in the solution

$$H_y(x,z) = \begin{Bmatrix} A_1 \exp(-\alpha_{1x} x) \\ A_2 \cos(k_{2x} x + \psi) \\ A_3 \exp(+\alpha_{3x} x) \end{Bmatrix} \exp(-jk_z z) \qquad \begin{matrix} x>d/2 \\ -d/2<x<d/2 \\ x<-d/2 \end{matrix} \quad . \tag{3.3.22}$$

The tangential electric field calculated from (3.2.2) is

$$E_z(x,z) = \frac{-j}{\omega} \begin{Bmatrix} -A_1 \dfrac{\alpha_{1x}}{n_1^2} \exp(-\alpha_{1x} x) \\ -A_2 \dfrac{k_{2x}}{n_2^2} \sin(k_{2x} x + \psi) \\ A_3 \dfrac{\alpha_{3x}}{n_3^2} \exp(+\alpha_{3x} x) \end{Bmatrix} \exp(-jk_z z) \qquad \begin{matrix} x>d/2 \\ -d/2<x<d/2 \\ x<-d/2 \end{matrix} \quad . \tag{3.3.23}$$

The guidance condition used to determine n_e is

$$-k_{2x} t + \tan^{-1}\left(\frac{\alpha_{1x} n_2^2}{k_{2x} n_1^2}\right) + \tan^{-1}\left(\frac{\alpha_{3x} n_2^2}{k_{2x} n_3^2}\right) + p\pi = 0 \quad , \tag{3.3.24}$$

where α_{1x}, k_{2x}, and α_{3x} were defined in (3.3.8a-c). Once n_e is found from (3.3.18) or (3.3.24), the other constants are determined from

$$\text{TE: } \psi = \tan^{-1}\left(\frac{\alpha_{1x}}{k_{2x}}\right) - k_{2x} t/2 \qquad \text{TM: } \psi = \tan^{-1}\left(\frac{\alpha_{1x}}{n_1^2}\frac{n_2^2}{k_{2x}}\right) - k_{2x} t/2 \tag{3.3.25a}$$

$$\text{TE and TM: } A_1 = \frac{A_2 \cos(k_{2x} + \psi)}{\exp(-\alpha_{1x} t/2)} \tag{3.3.25b}$$

$$\text{TE and TM: } A_3 = \frac{A_2 \cos(k_{2x} - \psi)}{\exp(-\alpha_{3x} t/2)} \quad , \tag{3.3.25c}$$

and A_2 establishes the power in the mode.

Some observations can be made about the guidance conditions (3.3.18) and (3.3.24). If t is increased, the $\tan^{-1}$ expressions must compensate by reducing k_{2x} and increasing α_{1x} and α_{3x}; this will require an increase in n_e. The bounce angle θ_b, determined from $k_z/k_{2x} = \tan(\theta_b)$, will therefore increase more beyond the critical angle. Increasing t will eventually allow a second mode to propagate when its bounce angle exceeds the critical angle. This mode will have a smaller n_e and α than the fundamental mode and its evanescent tail will protrude into the cladding further. Examination of the

guidance conditions also reveals a scaling rule. Wavelength cancels in the $\tan^{-1}$ expressions and only appears in the $-k_{2x}t$ term; wavelength may therefore be changed with no affect on n_e so long as t/λ_0 remains constant (aside from material dispersion). Finally, a larger core-cladding index difference will permit a ray at a smaller θ_b to be totally reflected, with a resulting decrease in the maximum thickness for single mode operation.

Solution method

Equation (3.3.24) is best solved with a bisection technique. With this method, two estimates of n_e, one above, and one below the true value are selected. One will produce a positive right hand side in (3.3.24) and the other a negative. Bound modes will occur for values of n_e between n_2 and the larger of n_1 and n_3. A slightly smaller interval, to avoid divide by zero errors, serves as the initial bracket. The average of the n_e bracket values is taken and the bracket value that has the same right hand side sign in (3.2.24) is replaced by this average. The process is continued until the bracket is sufficiently small. Since the bracket size is halved each iteration, roughly 30 are required to achieve, using double precision arithmetic, a 10^{-10} uncertainty in n_e.

An advantage of the guidance condition presented here is that, even for multimode guides, there always exists at most one n_e which will satisfy (3.3.24). In contrast, the guidance condition presented by Hunsperger containing tan(∗) [8] possesses multiple zeros for multimode guides, requiring that the iteration be started in a carefully chosen subbracket to assure convergence on the correct mode order's n_e.

4 FOUR LAYER PLANAR WAVEGUIDE

Although channel confinement may be created by partially etching a three layer (substrate-film-air) waveguide, the etch depth for single mode operation are very small making the channel difficult to distinguish. If the ridge is etched into a film having a lower index which is deposited onto the guide layer, the field will be weaker at the ridge,

and the etch depth can be larger for single mode operation. To design one of these, the guidance equation of the four layer planar waveguide depicted in Fig. 3-2 was determined.

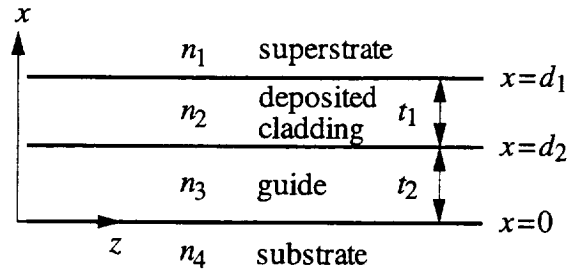


Fig. 3-2 Planar four layer waveguide. Propagation is in z.

Solution of the wave equation for TE modes yields

$$E_y(x,z) = \begin{cases} A_1 \exp(-\alpha_{1x} x) & d_1 < x \\ A_2 \exp(-\alpha_{2x} x) + B_2 \exp(+\alpha_{2x} x) & d_2 < x < d_1 \\ A_3 \cos(k_{3x} x + \psi) & 0 < x < d_2 \\ A_4 \exp(+\alpha_{4x} x) & x < 0 \end{cases} \exp(-jk_z z) \quad , \quad (3.4.1)$$

where it is assumed that $n_e > n_2$ so the fields are exponential in the deposited cladding and also that $n_3 > n_1, n_2, n_4$, making n_3 the guide layer. A sufficient condition for exponential fields in the deposited cladding is that $n_2 \leq n_4$; this would occur if silicon dioxide is selected for n_2 and n_4. It is also possible for $n_e < n_2$; this will result in a different guidance condition than shown below. The guidance condition for $n_e > n_2$ can be derived in a similar manner as before,

$$-k_{3x} t_2 + \tan^{-1}\left[\frac{\alpha_{2x}}{k_{3x}} \frac{\frac{\alpha_{2x}}{\alpha_{1x}} + 1 - \left(\frac{\alpha_{2x}}{\alpha_{1x}} - 1\right)\exp(-2\alpha_{2x} t_1)}{\frac{\alpha_{2x}}{\alpha_{1x}} + 1 + \left(\frac{\alpha_{2x}}{\alpha_{1x}} - 1\right)\exp(-2\alpha_{2x} t_1)}\right] + \tan^{-1}\left[\frac{\alpha_{4x}}{k_{3x}}\right] + p\pi = 0 \quad (3.4.2)$$

where

$$\alpha_{1x} = k_0 \sqrt{n_e^2 - n_1^2} \quad (3.4.3a)$$

$$\alpha_{2x} = k_0 \sqrt{n_e^2 - n_2^2} \quad (3.4.3b)$$

$$k_{3x} = k_0 \sqrt{n_3^2 - n_e^2} \tag{3.4.3c}$$

$$\alpha_{4x} = k_0 \sqrt{n_e^2 - n_4^2} \ . \tag{3.4.3d}$$

For TM modes, the **H** field equation will be identical to (3.4.1), except it will represent $H_y(x,z)$. The guidance condition is

$$-k_{3x}t_2 + \tan^{-1}\left[\frac{\dfrac{\alpha_{2x}n_1^2}{\alpha_{1x}n_2^2} + 1 - \left(\dfrac{\alpha_{2x}n_1^2}{\alpha_{1x}n_2^2} - 1\right)\exp(-2\alpha_{2x}t_1)}{\dfrac{\alpha_{2x}n_3^2}{k_{3x}n_2^2}\left[\dfrac{\alpha_{2x}n_1^2}{\alpha_{1x}n_2^2} + 1 + \left(\dfrac{\alpha_{2x}n_1^2}{\alpha_{1x}n_2^2} - 1\right)\exp(-2\alpha_{2x}t_1)\right]}\right] + \tan^{-1}\left[\frac{\alpha_{4x}n_3^2}{k_{3x}n_4^2}\right] + p\pi = 0 \ , \tag{3.4.4}$$

which is used with equations (3.4.3). The field profile is determined using (3.4.1) together with the following constants:

$$\text{TE: } \psi = \tan^{-1}\left(\frac{\alpha_{4x}}{k_{3x}}\right) \qquad \text{TM: } \psi = \tan^{-1}\left(\frac{\alpha_{4x}}{n_4^2}\frac{n_3^2}{k_{3x}}\right) \tag{3.4.5a}$$

$$\text{TE: } C = \frac{B_2}{A_2} = \frac{1 - \dfrac{\alpha_1}{\alpha_2}}{1 + \dfrac{\alpha_1}{\alpha_2}}\exp[-2\alpha_2(t_1 + t_2)] \qquad \text{TM: } C = \frac{B_2}{A_2} = \frac{1 - \dfrac{\alpha_1}{n_1^2}\dfrac{n_2^2}{\alpha_2}}{1 + \dfrac{\alpha_1}{n_1^2}\dfrac{n_2^2}{\alpha_2}}\exp[-2\alpha_2(t_1 + t_2)] \tag{3.4.5b}$$

$$\text{TE and TM: } A_1 = A_3\frac{\cos(k_3 t_2 + \psi)}{\exp[-\alpha_1(t_1 + t_2)]}\frac{\exp[-\alpha_2(t_1 + t_2)] + C\exp[+\alpha_2(t_1 + t_2)]}{\exp(-\alpha_2 t_2) + C\exp(+\alpha_2 t_2)} \tag{3.4.5c}$$

$$\text{TE and TM: } A_2 = A_3\frac{\cos(k_3 t_2 + \psi)}{\exp(-\alpha_2 t_2) + C\exp(+\alpha_2 t_2)} \tag{3.4.5d}$$

$$\text{TE and TM: } A_4 = A_3\cos(\psi) \tag{3.4.5e}$$

5 A NOTE ON POWER DENSITY IN THE PLANAR WAVEGUIDE

The power density (W/m^2) in the three layer asymmetric guide may be calculated from

$$P = \frac{1}{2}\mathrm{Re}(E \times H^*) \tag{3.5.1}$$

where E and H are the phasor representations of the fields. For example, for TE modes, in the region defined by n_1, inserting (3.3.7), (3.3.9) and (3.3.10) into (3.5.1) gives

$$P = \frac{1}{2}\exp(-2\alpha_{1x}x)\frac{|A_1|^2}{\omega\mu}\mathrm{Re}(k_z\hat{z} + j\alpha_{1x}\hat{x}) \quad . \tag{3.5.2}$$

The $\hat{x}$ component of P, which is derived from H_z, is removed by the operation Re $(\cdot)$. Since only H_x contributes to the power, and since, from (3.3.7) and (3.3.9) it is evident that $H_x = \frac{-k_z}{\omega\mu}E_y$, the power density for TE modes can be written $P = \frac{-k_z}{2\omega\mu}E_y^2$. Because E_y is continuous across the boundaries, so too is the power density. In addition, the ratio of power between x_1 and x_2 to that carried by the guide is just

$$\eta = \frac{\int_{x_1}^{x_2} E_y^2(x)\,dx}{\int_{-\infty}^{+\infty} E_y^2(x)\,dx} \quad , \tag{3.5.3}$$

where $E_y(x)$ is given by (3.3.7), and the phase factor $\exp(-jk_z z)$ is ignored.

For TM modes, it will be found that only E_x, which can be written $E_x = \frac{k_z}{\omega\varepsilon_0 n^2}H_y$, will contribute to the power density calculated to be $P = \frac{k_z}{2\omega\varepsilon_0 n^2}H_y^2$. Unlike TE modes, the power density will not be continuous across boundaries because n jumps abruptly. The ratio of power between x_1 and x_2 to that carried by the guide is given by

$$\eta = \frac{\int_{x_1}^{x_2} \frac{1}{n^2(x)}H_y^2(x)\,dx}{\int_{-\infty}^{+\infty} \frac{1}{n^2(x)}H_y^2(x)\,dx} \quad . \tag{3.5.4}$$

6 CHANNEL SOLUTION USING THE EFFECTIVE INDEX METHOD

Using the planar waveguide guidance relations, together with the effective index method, permits a determination of the 2-*D* field profile, which is needed for bending loss and directional coupler calculations, as well as channel width and etch depth determinations. Referring to Fig 3-3, the method is applied by first finding the effective index for the appropriate polarization of the planar guides adjacent to the channel (n_{e1}), and the planar guide which would be present if the channel width w was infinite (n_{e2}). The method assigns the channel effective index (n_{et}) the value of the effective index of the three layer planar guide which has its light polarization rotated 90° (*i.e.* TM if previous planar guides were TE), thickness w, and indices n_{e1}-n_{e2}-n_{e1}. The field decay in y is determined from this second planar guide calculation. If the region outside the channel is cutoff, its refractive index may be used in place of n_{e1} [4, 9], although the accuracy of the field decay constant in y may be poor, causing large errors in directional coupler transfer length calculations [2].

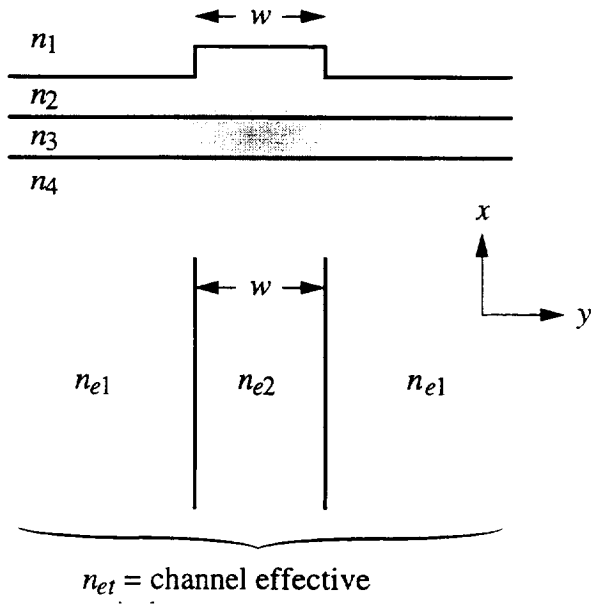


Fig. 3-3 Application of the effective index technique. The shaded region represents the approximate power distribution.

To understand this method, recall that the scalar wave equation for TE modes is

$$\frac{\partial^2 \mathbf{E}}{\partial x^2} + \frac{\partial^2 \mathbf{E}}{\partial y^2} + \frac{\partial^2 \mathbf{E}}{\partial z^2} + (2\pi n/\lambda_0)^2 \mathbf{E} = 0 \tag{3.6.1}$$

Assume that the **E** field distribution in the channel can be described by [1]

$$E_{2y} = A \cos(k_{2x} x) \cos(k_{2y} y) \exp(-j k_{zt} z) \tag{3.6.2}$$

where k_{zt} is the "total" or channel propagation constant. Substituting (3.6.2) into (3.6.1) gives

$$k_{zt} = \sqrt{k_0^2 n_2^2 - k_{2x}^2 - k_{2y}^2} \tag{3.6.3}$$

where n is the refractive index of the core, here assumed to be homogeneous. The same result can be found using the effective index technique. Here we observe that under the stripe,

$$k_{2x} = k_0 \sqrt{n_2^2 - n_{e2}^2} \tag{3.6.4}$$

According to the effective index method, considering the sideways waveguide depicted in Fig. 3-3 gives the other transverse field constant

$$k_{2y} = k_0 \sqrt{n_{e2}^2 - n_{et}^2} \tag{3.6.5}$$

Using (3.6.4) to eliminate n_{e2} in (3.6.5) gives

$$n_{et} k_0 \equiv k_{zt} = \sqrt{k_0^2 n_2^2 - k_{2x}^2 - k_{2y}^2} \tag{3.6.6}$$

which is the same result as when (3.6.2) was assumed for the solution to (3.6.1).

7 THE SUBSTRATE COUPLING PROBLEM

Waveguides formed for this thesis were fabricated on silicon substrates for reasons which include:

- Precision diaphragms can be micromachined into this substrate using wet anistropic etching.
- Silicon does not soften or deteriorate at the elevated temperatures used to deposit waveguide films (common glass would).

- Silicon substrates are both inexpensive and possess a good surface finish which reduces loss due to scattering.
- Silicon's opacity to light wavelengths used in this work (< 1 μm) suppressed substrate waveguiding and permitted easy identification of film waveguiding.
- Silicon's coefficient of thermal expansion (CTE=2.33 ppm/°C) [10] is slightly larger than that of Si_3N_4 (CTE=0.8 ppm/°C) [10] and SiO_2 (0.55 ppm/°C) [10]. LPCVD-deposited SiON having refractive index > 1.6 (for films deposited from $SiCl_2H_2$-NH_3-O_2) [11] or > 1.5 (for films deposited from SiH_4-CO_2-NH_3-H_2) is deposited by LPCVD under tensile stress, which is reduced during cooling because the silicon substrate contracts more than the film. This enables thicker crack-free films to be deposited than can be on silica wafers.
- Silicon wafers having a <100> orientation and a semiconductor standard major flat, which is along the {110} family of planes, preferentially cleave both perpendicular and parallel to this flat. These are also the optimum directions to align the sides of anisotropically etched rectangular diaphragms, making initial pattern alignment easier. Having a rectangular diaphragm side, a cleavage plane, and the flat all parallel means that a waveguide which travels along the edge of a diaphragm will also be perpendicular to a cleaved edge, simplifying end-fire coupling.

In order to make optical waveguides with films having a lower refractive index than silicon on silicon substrates, it is necessary to insert a transparent film having refractive index lower than the guide layer between it and the substrate. Such an isolation layer should have sufficient thickness that the evanescent field tail in the guide layer has a very small amplitude at the silicon interface. If this layer is too thin, the propagation loss due to substrate coupling dominates. The loss can be substantial, reaching hundreds of dB/cm. The equations used to calculate this loss and thereby determine the minimum isolation thickness will be presented in this section.

Consider the planar waveguide depicted in Fig. 3-4. Unlike the four layer

structure of the previous section, the guide layer is adjacent to the superstrate, while the bottom layer is the semi-infinite substrate. A typical structure could consist of Air ($n = 1$), SiON $n = 1.55$, SiO_2 ($n = 1.46$), and Si ($n = 3.85$). The problem is to find the propagation loss due to coupling of guided modes into the substrate. This loss can then be made negligible by making the isolation layer sufficiently thick.

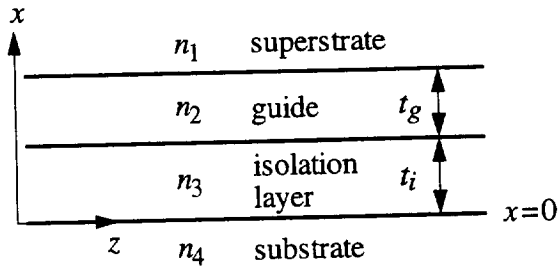


Fig. 3-4 Substrate coupling problem. The refractive index arrangement is $n_4, n_2 > n_e$, and $n_3, n_1 < n_e$.

The TE electric and TM magnetic field F_y has the form

$$F_y(x,z) = \begin{Bmatrix} A_1 \exp(-\alpha_1 x) \\ A_2 \cos(k_2 x + \psi) \\ A_3 \exp(-\alpha_3 x) + B_3 \exp(+\alpha_3 x) \\ A_4 \exp(+j k_4 x) \end{Bmatrix} \exp(-jk_z z) \qquad \begin{matrix} t_i+t_g<x \\ t_i<x<t_i+t_g \\ 0<x<t_i \\ x<0 \end{matrix} \tag{3.7.1}$$

where

$$\alpha_1 = k_0 \sqrt{n_e^2 - n_1^2} \tag{3.7.2a}$$

$$k_2 = k_0 \sqrt{n_2^2 - n_e^2} \tag{3.7.2b}$$

$$\alpha_3 = k_0 \sqrt{n_e^2 - n_3^2} \tag{3.7.2c}$$

and

$$k_4 = k_0 \sqrt{n_4^2 - n_e^2} \quad . \tag{3.7.2d}$$

Equations (3.7.1) and (3.7.2) are the solution of the scalar wave equation (3.3.6) in each of the layers of the structure in Fig. 3-4. The condition has been applied in both semi-infinite layers that the field must not exponentially increase with distance away from the waveguide. The effective index is complex,

$$n_e = n_r + jn_i \ , \tag{3.7.3}$$

and provides both the phase velocity and the propagation loss:

$$F_y(z) = F_y(z=0)\exp(-jn_r k_0 z)\exp(n_i k_0 z) \ , \tag{3.7.4}$$

where $k_z = k_0 n_e$ has been used. For the fundamental mode, the guidance equation can be derived in a similar manner as before,

$$0 = -\tan^{-1}\left[\frac{\alpha_3\sigma_3}{k_2\sigma_2}\frac{(\alpha_3\sigma_3 + jk_4\sigma_4) - (\alpha_3\sigma_3 - jk_4\sigma_4)\exp(-2\alpha_3 t_i)}{(\alpha_3\sigma_3 + jk_4\sigma_4) + (\alpha_3\sigma_3 - jk_4\sigma_4)\exp(-2\alpha_3 t_i)}\right] - \tan^{-1}\left(\frac{\alpha_1\sigma_1}{k_2\sigma_2}\right) + k_2 t_g \tag{3.7.5}$$

where $\sigma_x = 1$ for TE modes and $\sigma_x = 1/n_x^2$ for TM modes.

Iterating for the effective index using (3.7.5) is substantially more involved than in previous problems because the α's, k's and n_e are complex. Equation (3.7.5) can be written $E_r + jE_i = f(n_r + jn_i)$, where E_r and E_i, the error terms, represent the real and imaginary part of the right hand side of (3.7.5). A 2-dimensional Newton's method can be used to solve simultaneously for the values of n_r and n_i which will set a linearized version of the right hand side of (3.7.5) to $0 + j0$ [12]. This method operates by expressing (3.7.5) as

$$f_r(n_r,n_i) = f_r(n_{r0},n_{i0}) + (n_r - n_{r0})\frac{\partial f_r(n_{r0},n_{i0})}{\partial n_r} + (n_i - n_{i0})\frac{\partial f_r(n_{r0},n_{i0})}{\partial n_i} \tag{3.7.6a}$$

$$f_i(n_r,n_i) = f_i(n_{r0},n_{i0}) + (n_r - n_{r0})\frac{\partial f_i(n_{r0},n_{i0})}{\partial n_r} + (n_i - n_{i0})\frac{\partial f_i(n_{r0},n_{i0})}{\partial n_i} \ , \tag{3.7.6b}$$

where the functions f_r and f_i are the real and imaginary part of the right hand side of (3.7.5), and the partial derivatives are evaluated numerically. The point n_{r0},n_{i0} is an initial estimate of the complex effective index. Once it has been selected, the left hand side of (3.7.6) is set to zero and new values of n_r and n_i solved for:

$$n_r = n_{r0} + \frac{f_i(n_{r0},n_{i0})\dfrac{\partial f_r(n_{r0},n_{i0})}{\partial n_i} - f_r(n_{r0},n_{i0})\dfrac{\partial f_i(n_{r0},n_{i0})}{\partial n_i}}{\dfrac{\partial f_r(n_{r0},n_{i0})}{\partial n_r}\dfrac{\partial f_i(n_{r0},n_{i0})}{\partial n_i} - \dfrac{\partial f_i(n_{r0},n_{i0})}{\partial n_r}\dfrac{\partial f_r(n_{r0},n_{i0})}{\partial n_i}} \tag{3.7.7a}$$

$$n_i = n_{i0} + \frac{f_r(n_{r0},n_{i0})\dfrac{\partial f_i(n_{r0},n_{i0})}{\partial n_r} - f_i(n_{r0},n_{i0})\dfrac{\partial f_r(n_{r0},n_{i0})}{\partial n_r}}{\dfrac{\partial f_r(n_{r0},n_{i0})}{\partial n_r}\dfrac{\partial f_i(n_{r0},n_{i0})}{\partial n_i} - \dfrac{\partial f_i(n_{r0},n_{i0})}{\partial n_r}\dfrac{\partial f_r(n_{r0},n_{i0})}{\partial n_i}} \quad . \tag{3.7.7b}$$

n_r and n_i are then assigned to n_{r0} and n_{i0}. This process is continued until the values converge.

Because the iteration uses Newton's method, convergence is not guaranteed as it is with the three and four (upper cladding) layer cases which use non-complex n_e's that can be bracketed and found with the bisection method. If the initial guess is far from the correct answer, it may not converge.

Perturbation analysis

Using a perturbation method, the complex n_e for the four layer waveguide in Fig. 3-4 may be found without iterating (3.7.5). The method exploits the small real and imaginary effective index difference between the 4 layer problem and the three layer one in which n_3 in Fig. 3-4 is infinite.

Using the identity $\tan(\alpha+\beta) = \dfrac{\tan(\alpha)+\tan(\beta)}{1-\tan(\alpha)\tan(\beta)}$, (3.7.5) can be transformed to

$$\begin{aligned}&(\alpha_3\sigma_3 + jk_4\sigma_4)\,[(k_2^2\sigma_2^2 - \alpha_1\alpha_3\sigma_1\sigma_3)\tan(k_2 t_g) - k_2\sigma_2(\alpha_3\sigma_3 + \alpha_1\sigma_1)]\\ &+ \exp(-2\alpha_3 t_i)\,(\alpha_3\sigma_3 - jk_4\sigma_4)\,[(k_2^2\sigma_2^2 + \alpha_1\alpha_3\sigma_1\sigma_3)\tan(k_2 t_g) + k_2\sigma_2(\alpha_3\sigma_3 - \alpha_1\sigma_1)] = 0\end{aligned} \tag{3.7.8}$$

where σ has been defined previously in this section. The three layer asymmetric guidance conditions (3.3.18) and (3.3.24) may likewise be transformed into

$$(k_2^2\sigma_2^2 - \alpha_1\alpha_3\sigma_1\sigma_3)\tan(k_2 t_g) - k_2\sigma_2(\alpha_3\sigma_3 + \alpha_1\sigma_1) = 0 \quad . \tag{3.7.9}$$

This function is apparent in (3.7.15). A method of finding the complex propagation constant (which gives the substrate coupling) without iterating (3.7.15) is to rewrite it as

$$\begin{aligned}&(k_2^2\sigma_2^2 - \alpha_1\alpha_3\sigma_1\sigma_3)\tan(k_2 t_g) - k_2\sigma_2(\alpha_3\sigma_3 + \alpha_1\sigma_1) = F_3\\ &= -\exp(-2\alpha_3 t_i)\frac{(\alpha_3\sigma_3 - jk_4\sigma_4)}{(\alpha_3\sigma_3 + jk_4\sigma_4)}\,[(k_2^2\sigma_2^2 + \alpha_1\alpha_3\sigma_1\sigma_3)\tan(k_2 t_g) + k_2\sigma_2(\alpha_3\sigma_3 - \alpha_1\sigma_1)],\end{aligned} \tag{3.7.10}$$

linearize the left hand side of (3.7.10) about the three layer solution, and explicitly solve for the complex k_z. Since k_z is expected to depart only slightly from the three layer case, the left hand side is expected to be close to zero and can be written $F_3 = \kappa \Delta k_z$, where the four layer $k_z = \Delta k_z + \overline{k_z}$, $\overline{k_z}$ is the propagation constant of the three layer case ($t_i \rightarrow \infty$), and κ is a function of the waveguide parameters. Once κ is determined, (3.7.10) can be rewritten $\Delta k_z = \dfrac{\text{Right hand side of (3.7.10)}}{\kappa}$ and the four layer k_z explicitly found. The left hand side of (3.7.10) is linearized by first writing α_1, k_2, and α_3 in terms of the three layer $n_e, \overline{n}_e$:

$$\alpha_1 = k_0 \sqrt{(\overline{n}_e + \Delta n_e)^2 - n_1^2} \tag{3.7.11}$$

$$\alpha_1 = k_0 \sqrt{\overline{n}_e^2 + 2\overline{n}_e \Delta n_e + \Delta n_e^2 - n_1^2} \tag{3.7.12}$$

Throughout the derivation, Δ terms which are squared will be neglected.

$$\alpha_1 = k_0 \sqrt{(\overline{n}_e^2 - n_1^2)\left(1 + \frac{2\overline{n}_e \Delta n_e}{(\overline{n}_e^2 - n_1^2)}\right)} \tag{3.7.13}$$

$$\alpha_1 = \overline{\alpha}_1 \sqrt{1 + \frac{2\overline{n}_e \Delta n_e k_0^2}{\overline{\alpha}_1}} \tag{3.7.14}$$

$\overline{\alpha}_1$ corresponds to the three layer ($t_i \rightarrow \infty$) α_1. Using the approximation $\sqrt{1 + \Delta x} \approx 1 + \Delta x/2$,

$$\alpha_1 = \overline{\alpha}_1 + \frac{\overline{n}_e \Delta n_e k_0^2}{\overline{\alpha}_1} \tag{3.7.15}$$

$$\alpha_1 = \overline{\alpha}_1 + \frac{\overline{k}_z \Delta k_z}{\overline{\alpha}_1} \; . \tag{3.7.16}$$

Similarly,

$$k_2 = \overline{k}_2 - \frac{\overline{k}_z \Delta k_z}{\overline{k}_2} \tag{3.7.17}$$

$$\alpha_3 = \overline{\alpha}_3 + \frac{\overline{k}_z \Delta k_z}{\overline{\alpha}_3} \; , \tag{3.7.18}$$

where the bar indicates the easily obtained 3 layer quantity. Inserting (3.7.16)-(3.7.18)

into the left hand side of (3.7.10) gives

$$F_3 = \left[\left(\bar{k}_2 - \frac{\bar{k}_z}{\bar{k}_2}\Delta k_z\right)^2 \sigma_2^2 - \left(\bar{\alpha}_1 + \frac{\bar{k}_z}{\bar{\alpha}_1}\Delta k_z\right)\left(\bar{\alpha}_3 + \frac{\bar{k}_z}{\bar{\alpha}_3}\Delta k_z\right)\sigma_1\sigma_3\right]\tan\left[\left(\bar{k}_2 - \frac{\bar{k}_z}{\bar{k}_2}\Delta k_z\right)t_g\right]$$
$$-\left(\bar{k}_2 - \frac{\bar{k}_z}{\bar{k}_2}\Delta k_z\right)\sigma_2\left[\left(\bar{\alpha}_3 + \frac{\bar{k}_z}{\bar{\alpha}_3}\Delta k_z\right)\sigma_3 + \left(\bar{\alpha}_1 + \frac{\bar{k}_z}{\bar{\alpha}_1}\Delta k_z\right)\sigma_1\right] . \qquad (3.7.19)$$

Using the approximation

$$\tan\left[\left(\bar{k}_2 - \frac{\bar{k}_z}{\bar{k}_2}\Delta k_z\right)t_g\right] \approx \frac{\tan(\bar{k}_2 t_g) - \frac{\bar{k}_z}{\bar{k}_2}\Delta k_z t_g}{1 + \tan(\bar{k}_2 t_g)\frac{\bar{k}_z}{\bar{k}_2}\Delta k_z t_g} \approx \left[\tan(\bar{k}_2 t_g) - \frac{\bar{k}_z}{\bar{k}_2}\Delta k_z t_g\right]\left[1 - \frac{\bar{k}_z}{\bar{k}_2}\Delta k_z t_g \tan(\bar{k}_2 t_g)\right]$$
$$\approx \tan(\bar{k}_2 t_g) - \left[1 + \tan^2(\bar{k}_2 t_g)\right]\frac{\bar{k}_z}{\bar{k}_2}\Delta k_z t_g \ ,$$

(3.7.19) can be rewritten

$$F_3 = (\bar{k}_2^2\sigma_2^2 - \bar{\alpha}_1\bar{\alpha}_3\sigma_1\sigma_3)\tan(\bar{k}_2 t_g) - \bar{k}_2\sigma_2(\bar{\alpha}_3\sigma_3 + \bar{\alpha}_1\sigma_1)$$
$$-\bar{k}_z\Delta k_z\left[2\sigma_2^2 + \sigma_1\sigma_3\left(\frac{\bar{\alpha}_1}{\bar{\alpha}_3} + \frac{\bar{\alpha}_3}{\bar{\alpha}_1}\right)\right]\tan(\bar{k}_2 t_g) - \bar{k}_z\Delta k_z(\bar{k}_2^2\sigma_2^2 - \bar{\alpha}_1\bar{\alpha}_3\sigma_1\sigma_3)\left[1 + \tan^2(\bar{k}_2 t_g)\right]\frac{t_g}{\bar{k}_2}$$
$$-\bar{k}_z\Delta k_z\left[\sigma_1\sigma_2\left(\frac{\bar{k}_2}{\bar{\alpha}_1} - \frac{\bar{\alpha}_1}{\bar{k}_2}\right) + \sigma_2\sigma_3\left(\frac{\bar{k}_2}{\bar{\alpha}_3} - \frac{\bar{\alpha}_3}{\bar{k}_2}\right)\right] \qquad (3.7.20)$$

The first line of (3.7.20) is zero, according to the three layer guidance condition (3.7.20). Although it is now possible to rewrite (3.7.10) and solve for Δk_z, significant simplification of the resulting expression will occur in subsequent steps. Define:

$$\tilde{\alpha}_1 = \bar{\alpha}_1\sigma_1 \qquad (3.7.21a)$$

$$\tilde{k}_2 = \bar{k}_2\sigma_2 \qquad (3.7.21b)$$

$$\tilde{\alpha}_3 = \bar{\alpha}_3\sigma_3 \qquad (3.7.21c)$$

$$\tilde{k}_4 = \bar{k}_4\sigma_4 \qquad (3.7.21d)$$

(3.7.20) now becomes:

$$F_3 = -\bar{k}_z \Delta k_z \left\{ (\tilde{k}_2^2 - \tilde{\alpha}_1 \tilde{\alpha}_3) \left[1 + \left(\frac{\tilde{k}_2 (\tilde{\alpha}_1 + \tilde{\alpha}_3)}{\tilde{k}_2^2 - \tilde{\alpha}_1 \tilde{\alpha}_3} \right)^2 \right] \frac{t_g \sigma_2}{\tilde{k}_2} \right.$$
$$+ \left[2\sigma_2^2 + \sigma_1 \sigma_3 \left(\frac{\tilde{\alpha}_1 \sigma_3}{\tilde{\alpha}_3 \sigma_1} + \frac{\tilde{\alpha}_3 \sigma_1}{\tilde{\alpha}_1 \sigma_3} \right) \right] \frac{\tilde{k}_2 (\tilde{\alpha}_1 + \tilde{\alpha}_3)}{\tilde{k}_2^2 - \tilde{\alpha}_1 \tilde{\alpha}_3}$$
$$\left. + \sigma_1 \sigma_2 \left(\frac{\tilde{k}_2 \sigma_1}{\tilde{\alpha}_1 \sigma_2} - \frac{\tilde{\alpha}_1 \sigma_2}{\tilde{k}_2 \sigma_1} \right) + \sigma_2 \sigma_3 \left(\frac{\tilde{k}_2 \sigma_3}{\tilde{\alpha}_3 \sigma_2} - \frac{\tilde{\alpha}_3 \sigma_2}{\tilde{k}_2 \sigma_3} \right) \right\} \tag{3.7.22}$$

$$F_3 = -\bar{k}_z \Delta k_z \left\{ \frac{(\tilde{k}_2^2 - \tilde{\alpha}_1 \tilde{\alpha}_3)^2 + \tilde{k}_2^2 (\tilde{\alpha}_1 + \tilde{\alpha}_3)^2}{\tilde{k}_2^2 - \tilde{\alpha}_1 \tilde{\alpha}_3} \frac{t_g \sigma_2}{\tilde{k}_2} \right.$$
$$+ \left[\frac{2\sigma_2^2 \tilde{\alpha}_1 \tilde{\alpha}_3 + \tilde{\alpha}_1^2 \sigma_3^2 + \tilde{\alpha}_3^2 \sigma_1^2}{\tilde{\alpha}_1 \tilde{\alpha}_3} \right] \frac{\tilde{k}_2 (\tilde{\alpha}_1 + \tilde{\alpha}_3)}{\tilde{k}_2^2 - \tilde{\alpha}_1 \tilde{\alpha}_3}$$
$$\left. + \frac{\tilde{k}_2^2 \tilde{\alpha}_3 \sigma_1^2 + \tilde{k}_2^2 \tilde{\alpha}_1 \sigma_3^2 - \tilde{\alpha}_1^2 \tilde{\alpha}_3 \sigma_2^2 - \tilde{\alpha}_1 \tilde{\alpha}_3^2 \sigma_2^2}{\tilde{\alpha}_1 \tilde{k}_2 \tilde{\alpha}_3} \right\} \tag{3.7.23}$$

$$F_3 = \frac{-\bar{k}_z \Delta k_z}{\tilde{\alpha}_1 \tilde{k}_2 \tilde{\alpha}_3 (\tilde{k}_2^2 - \tilde{\alpha}_1 \tilde{\alpha}_3)} \left\{ (\tilde{k}_2^2 + \tilde{\alpha}_1^2)(\tilde{k}_2^2 + \tilde{\alpha}_3^2)\, t_g \sigma_2 \tilde{\alpha}_1 \tilde{\alpha}_3 \right.$$
$$+ (\sigma_1^2 \tilde{\alpha}_3^2 + 2\sigma_2^2 \tilde{\alpha}_1 \tilde{\alpha}_3 + \sigma_3^2 \tilde{\alpha}_1^2)\, \tilde{k}_2^2 (\tilde{\alpha}_1 + \tilde{\alpha}_3)$$
$$\left. + \left[\tilde{k}_2^2 (\tilde{\alpha}_3 \sigma_1^2 + \tilde{\alpha}_1 \sigma_3^2) - \sigma_2^2 \tilde{\alpha}_1 \tilde{\alpha}_3 (\tilde{\alpha}_1 + \tilde{\alpha}_3) \right] (\tilde{k}_2^2 - \tilde{\alpha}_1 \tilde{\alpha}_3) \right\} \tag{3.7.24}$$

$$F_3 = \frac{-\bar{k}_z \Delta k_z (\tilde{k}_2^2 + \tilde{\alpha}_1^2)(\tilde{k}_2^2 + \tilde{\alpha}_3^2) \tilde{\alpha}_1 \tilde{\alpha}_3}{\tilde{\alpha}_1 \tilde{k}_2 \tilde{\alpha}_3 (\tilde{k}_2^2 - \tilde{\alpha}_1 \tilde{\alpha}_3)} \left(t_g \sigma_2 + \frac{(\tilde{\alpha}_1 + \tilde{\alpha}_3)}{(\tilde{k}_2^2 + \tilde{\alpha}_1^2)(\tilde{k}_2^2 + \tilde{\alpha}_3^2)\, \tilde{\alpha}_1 \tilde{\alpha}_3} \right.$$
$$\cdot \left\{ \tilde{k}_2^2 (\sigma_1^2 \tilde{\alpha}_3^2 + 2\sigma_2^2 \tilde{\alpha}_1 \tilde{\alpha}_3 + \sigma_3^2 \tilde{\alpha}_1^2) \right.$$
$$\left. \left. + \left[\tilde{k}_2^2 \frac{(\tilde{\alpha}_3 \sigma_1^2 + \tilde{\alpha}_1 \sigma_3^2)}{\tilde{\alpha}_3 + \tilde{\alpha}_1} - \sigma_2^2 \tilde{\alpha}_1 \tilde{\alpha}_3 \right] (\tilde{k}_2^2 - \tilde{\alpha}_1 \tilde{\alpha}_3) \right\} \right) \tag{3.7.25}$$

$$F_3 = \frac{-\bar{k}_z \Delta k_z (\tilde{k}_2^2 + \tilde{\alpha}_1^2)(\tilde{k}_2^2 + \tilde{\alpha}_3^2)\tilde{\alpha}_1 \tilde{\alpha}_3}{\tilde{\alpha}_1 \tilde{k}_2 \tilde{\alpha}_3 (\tilde{k}_2^2 - \tilde{\alpha}_1 \tilde{\alpha}_3)} \left(t_g \sigma_2 + \left[\frac{1}{\tilde{\alpha}_3} + \frac{1}{\tilde{\alpha}_1} \right] \right.$$
$$\cdot \frac{1}{(\tilde{k}_2^2 + \tilde{\alpha}_1^2)(\tilde{k}_2^2 + \tilde{\alpha}_3^2)} \left\{ \tilde{k}_2^2 (\sigma_1^2 \tilde{\alpha}_3^2 + 2\sigma_2^2 \tilde{\alpha}_1 \tilde{\alpha}_3 + \sigma_3^2 \tilde{\alpha}_1^2) \right.$$
$$\left. \left. + \left[\tilde{k}_2^2 \frac{(\tilde{\alpha}_3 \sigma_1^2 + \tilde{\alpha}_1 \sigma_3^2)}{\tilde{\alpha}_3 + \tilde{\alpha}_1} - \sigma_2^2 \tilde{\alpha}_1 \tilde{\alpha}_3 \right] (\tilde{k}_2^2 - \tilde{\alpha}_1 \tilde{\alpha}_3) \right\} \right) \qquad (3.7.26)$$

By defining Γ as the last two lines of (3.7.26)

$$\Gamma = \frac{\sigma_2^2 \sigma_1 \sigma_3}{\gamma^2 + \mu^2} \left[\bar{k}_2^2 (\sigma_1 \sigma_3 \bar{\alpha}_3^2 + 2\sigma_2^2 \bar{\alpha}_1 \bar{\alpha}_3 + \sigma_1 \sigma_3 \bar{\alpha}_1^2) + \gamma \left(\bar{k}_2^2 \frac{\sigma_1 \bar{\alpha}_3 + \sigma_3 \bar{\alpha}_1}{\sigma_3 \bar{\alpha}_3 + \sigma_1 \bar{\alpha}_1} - \bar{\alpha}_1 \bar{\alpha}_3 \right) \right] \qquad (3.7.27)$$

where

$$\gamma = \bar{k}_2^2 \sigma_2^2 - \bar{\alpha}_1 \bar{\alpha}_3 \sigma_1 \sigma_3 \ , \qquad (3.7.28)$$

$$\mu = \sigma_2 \bar{k}_2 (\sigma_1 \bar{\alpha}_1 + \sigma_3 \bar{\alpha}_3) \ , \qquad (3.7.29)$$

and

$$\omega_e = \sigma_2 t_g + \Gamma \left(\frac{1}{\sigma_3 \bar{\alpha}_3} + \frac{1}{\sigma_1 \bar{\alpha}_1} \right) \ , \qquad (3.7.30)$$

F_3 can be rewritten as

$$F_3 = -\bar{k}_z \Delta k_z \frac{\gamma^2 + \mu^2}{\sigma_2 \bar{k}_2 \gamma} \omega_e \ . \qquad (3.7.31)$$

Substituting the three layer guidance (3.7.19) condition into the right hand side of (3.7.10) to remove the tan(*) gives

$$F_3 = -\exp(-2\bar{\alpha}_3 t_i) \frac{\sigma_3 \bar{\alpha}_3 - j\sigma_4 \bar{k}_4}{\sigma_3 \bar{\alpha}_3 + j\sigma_4 \bar{k}_4} \frac{2\sigma_2 \sigma_3 \bar{k}_2 \bar{\alpha}_3 (\sigma_2^2 \bar{k}_2^2 + \sigma_1^2 \bar{\alpha}_1^2)}{\sigma_2^2 \bar{k}_2^2 - \sigma_1 \sigma_3 \bar{\alpha}_1 \bar{\alpha}_3} \ . \qquad (3.7.32)$$

Equating (3.7.31) and (3.7.32) produces the final result

$$\Delta k_z = \frac{\sigma_3 \bar{\alpha}_3 - j\sigma_4 \bar{k}_4}{\sigma_3 \bar{\alpha}_3 + j\sigma_4 \bar{k}_4} \frac{2\sigma_2^2 \bar{k}_2^2 \sigma_3 \bar{\alpha}_3}{\bar{k}_z (\sigma_2^2 \bar{k}_2^2 + \sigma_3^2 \bar{\alpha}_3^2) \omega_e} \exp(-2\bar{\alpha}_3 t_i) \qquad (3.7.33)$$

where, again, the quantities with the bar are calculated using the three layer n_e, which is

real, and (3.7.2a-d). For the TE case, in addition to having all the σ's become 1, Γ also simplifies to 1.

The loss of a five layer waveguide consisting of a guide with upper cladding as in Fig. 3-2 which sits on a layer isolating it from the substrate can be estimated from the above analysis. Suppose the field varies exponentially in the upper cladding. Employing the layer designations in Fig. 3-2, for a thin upper cladding, the effective index will be close to that of the three layer guide consisting of n_1-n_3-n_4. As the cladding thickness is increased, n_e will approach that of the three layer guide consisting of n_2-n_3-n_4. If the effective index is calculated to lie much closer to one of these limits than the other, the evanescent field decay rate in the isolation layer, and the substrate coupling loss will likewise be similar to having this three layer guide on the high index substrate. At any rate, the substrate coupling loss will lie between the that of the two extremes.

Propagation loss

The propagation loss can be calculated from Δk_z. Writing

$$\Delta k_z = \Delta k_{zr} + j\,\Delta k_{zi} \quad , \tag{3.7.34}$$

the electric field amplitude varies as

$$E(z) = E_0 \exp(k_z z) = E_0 \exp[-j(\bar{k}_z + \Delta k_{zr})\, z] \exp(+\Delta k_{zi}\, z) \quad . \tag{3.7.35}$$

For a lossy waveguide, $\Delta k_{zi} < 0$. The waveguide loss in dB/cm is found from

$$\begin{aligned} \text{loss (dB/cm)} &= 10\log_{10}\left(\frac{\text{Power at 0 cm}}{\text{Power at 1 cm}}\right) = 20\log_{10}\left(\frac{E \text{ at 0 cm}}{E \text{ at 1 cm}}\right) \\ &= 20\log_{10}\left(\frac{1}{\exp(+\Delta k_{zi} \text{ in cm}^{-1})}\right) \end{aligned} \tag{3.7.36}$$

For a Δk_{zi} in μm^{-1}, the relationship is

$$\text{loss (dB/cm)} = -86{,}859\, \Delta k_{zi} \text{ (in } \mu m^{-1}) \quad . \tag{3.7.37}$$

Waveguides in this study were formed by depositing SiON film on oxidized silicon substrates. Since the refractive index of the silicon substrate exceeds that of the waveguide film, total internal reflection will not occur and the power will leak into the

substrate. By growing a thermal oxide of sufficient thickness which has a refractive index below that of the guide film, the loss due to substrate coupling can be made as low as desired. For this work, a negligible substrate coupling loss is 0.1 dB/cm.

Figures 3-5 through 3-7 plot the thermal oxide isolation thickness required for a loss of 0.1 dB/cm due to substrate coupling vs guide thickness for various guide refractive indices. The plots range over values of t_g from where the fundamental TM mode cuts off to where a second TM mode ($m = 1$) can propagate. As the guide approaches cutoff, the mode broadens and the required isolation thickness increases. The plots are for the TM mode, which spreads out more than the TE and so is lossier. Thus the TE loss due to substrate coupling will be less than 0.1 dB/cm.

As Fig. 3-8 might suggest, substrate coupling can be exploited to produce single polarization waveguides, where the TM mode will be rapidly attenuated through substrate coupling, while the TE mode propagation loss will be much lower, and remain dominated by scattering and film absorption [13]. This differential loss can be made greater in highly birefringent waveguides, as the decay rates of the field and thus the amplitude at the substrate-isolation layer interface will then be most different.

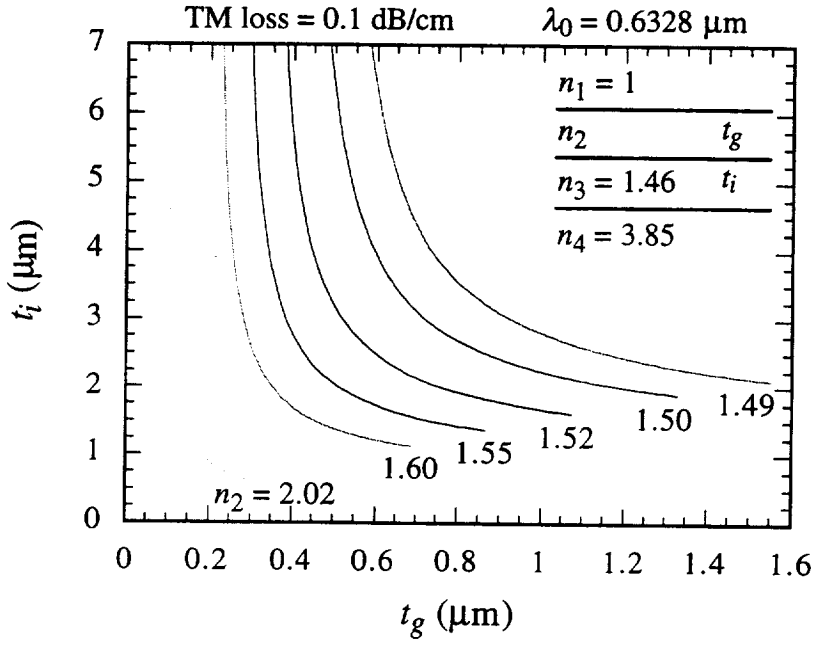


Fig. 3-5 Thickness of the isolation layer to have a fundamental ($m = 0$) mode TM substrate coupling loss of 0.1 dB/cm for vs. core layer t_g thickness. Curves cover t_g between TM fundamental cutoff and TM $m = 1$ mode cutoff.

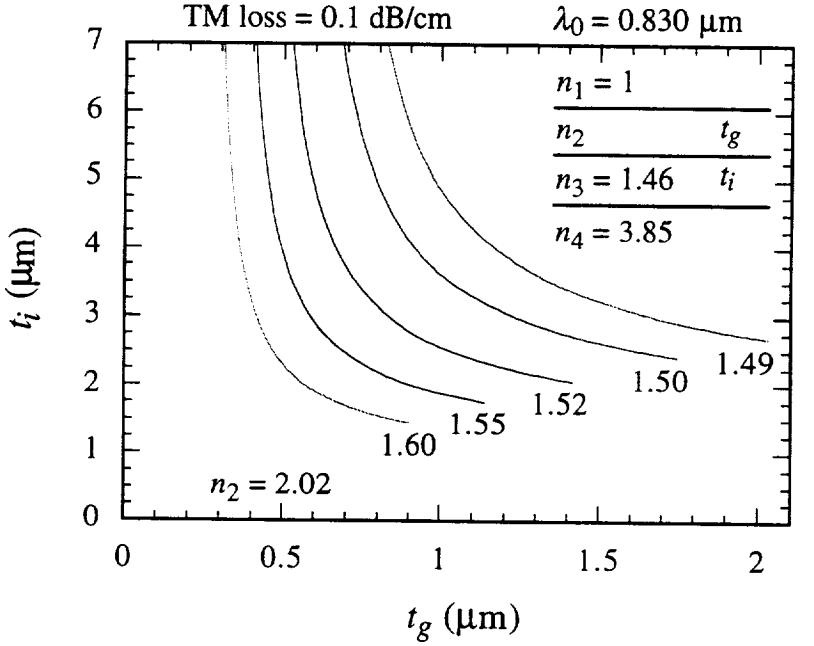


Fig. 3-6 Same as Fig. 3-5 except for $\lambda_0 = 0.830$ μm.

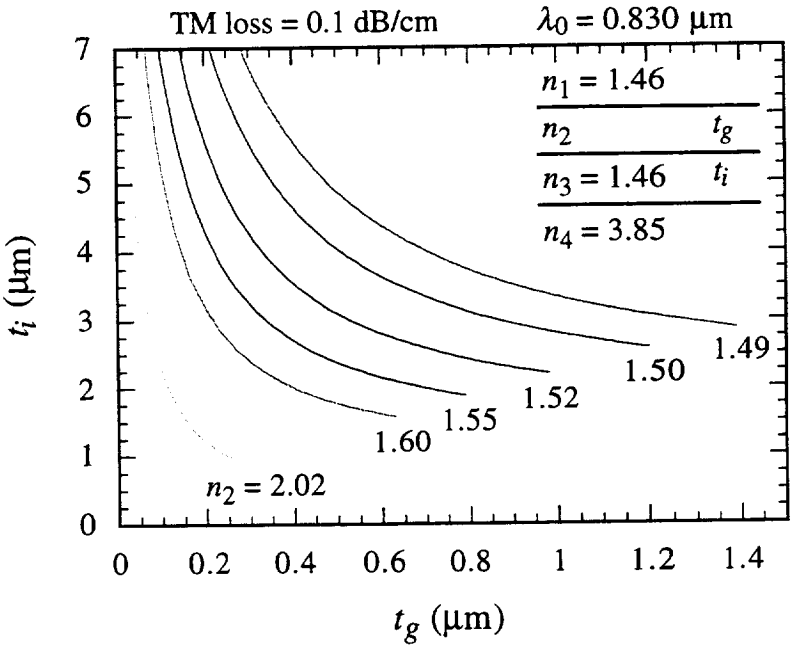


Fig. 3-7 Same as Fig. 3-6 except air has been replaced by an SiO_2 overlayer.

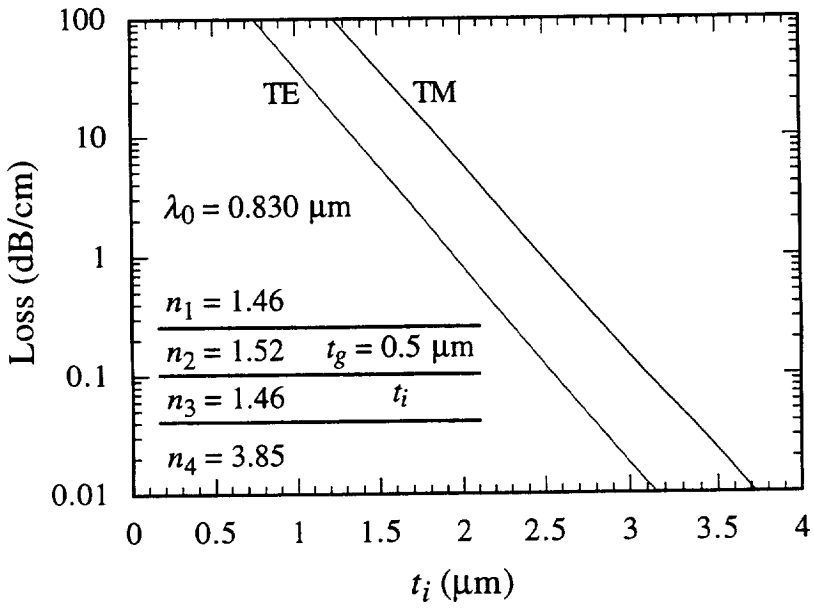


Fig. 3-8 Substrate coupling loss vs. isolation layer thickness for the planar guide shown in the inset.

8 DIRECTIONAL COUPLER

A directional coupler consists of two waveguides placed in close enough proximity that their evanescent fields partially overlap, permitting power to be transferred from one guide to the other. They are fabricated from two fibers by embedding each in a glass block, lapping a side of each block until the field is accessible and bonding the blocks together with the fibers adjacently positioned. Coupling can be controlled by adjusting the alignment between the two blocks. In integrated optics, where coupler trimming is more difficult, the *Y*-branch is often preferred for 50%/50% splitting. For other ratios, the directional coupler is favored. Because the directional coupler is wavelength sensitive, it can be used as a wavelength division multiplexer. An interferometer's ambiguity problem can be solved by using a three coupler which can simultaneously produce a sine and cosine signal [14, 15]. In this thesis, the directional coupler was configured with a ring resonator interferometer to form one of the pressure sensor designs.

Lee [16] has calculated the coupling coefficient between the fundamental modes of the two slab waveguides pictured in Fig. 3-9 using coupled mode theory; equations (3.8.1)-(3.8.4) and (3.8.6)-(3.8.15) were taken from this reference but are rewritten for use with channel waveguides approached with the effective index technique. Quasi-TE means that the electric field is parallel to the layers composing the channel structure (if the channel were infinitely wide this same polarization would be referred to as TE) and quasi-TM is with the magnetic field parallel to the layers. If the waveguides are not too close together, the fields over the directional coupler can be written in terms of the normalized isolated e and h fields of each channel:

$$\begin{Bmatrix} E_{DC} \\ H_{DC} \end{Bmatrix} = a_A(z) \begin{Bmatrix} e_A(x,y) \\ h_A(x,y) \end{Bmatrix} e^{-jzk_{z\,tA}} + a_B(z) \begin{Bmatrix} e_B(x,y) \\ h_B(x,y) \end{Bmatrix} e^{-jzk_{z\,tB}} \quad . \tag{3.8.1}$$

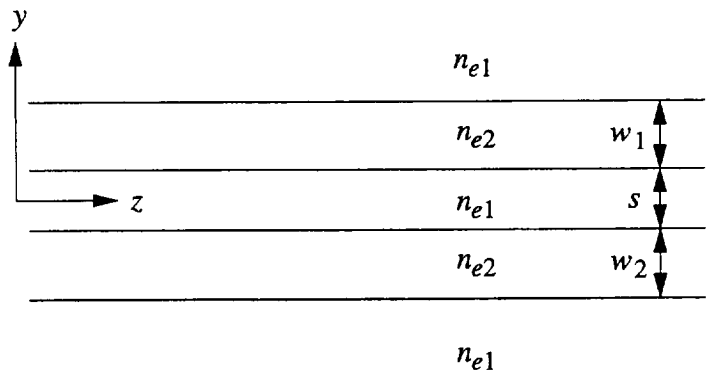


Fig. 3-9 Top view of directional coupler [16] made with channel waveguides.

The coupled mode equations are

$$\frac{da_A(z)}{dz} = -jC_{BA}\, e^{jz\Delta k} a_B(z) \tag{3.8.2}$$

$$\frac{da_B(z)}{dz} = -jC_{AB}\, e^{-jz\Delta k} a_A(z) \tag{3.8.3}$$

Differentiating (3.8.2) and substituting (3.8.3) and (3.8.2) into it gives

$$\frac{d^2a_A(z)}{dz^2} - j\Delta k \frac{da_A(z)}{dz} + C_{BA}\, C_{AB}\, a_A(z) = 0 \tag{3.8.4}$$

where $\Delta k = k_{z\,tA} - k_{z\,tB}$.

For the phase matched ($\Delta k = 0$) case, where power may be launched into both channels, the solution can be written

$$\begin{bmatrix} a_A(z) \\ a_B(z) \end{bmatrix} = \begin{bmatrix} \cos(Cz) & -j\sin(Cz) \\ -j\sin(Cz) & \cos(Cz) \end{bmatrix} \begin{bmatrix} a_{A0} \\ a_{B0} \end{bmatrix}, \tag{3.8.5}$$

where a_{A0} and a_{B0} are the field amplitudes at $z = 0$ and $C = C_{AB} = C_{BA}$.

For the non-phase matched case where all the optical power initially resides in slab B, $a_{A0} = 0$ and (3.8.4) yields

$$a_A(z) = 2jA_1\, e^{jz\Delta k/2} \sin(\mathcal{S}z) \tag{3.8.6}$$

and $$a_B(z) = \frac{2jA_1}{C_{BA}} e^{-jz\Delta k/2}\left[-\frac{\Delta k}{2}\sin(\mathcal{S}z) + j\mathcal{S}\cos(\mathcal{S}z)\right], \tag{3.8.7}$$

where $\mathcal{S}$ is defined as $\mathcal{S} = \sqrt{\left(\frac{\Delta k}{2}\right)^2 + C_{AB}C_{BA}}$. If the guides are synchronous or phase matched, $\Delta k = 0$ and the power in the guides is found to be

$$P_A(z) = P_0 \sin^2(Cz) \tag{3.8.8}$$

$$P_B(z) = P_0 \cos^2(Cz) \ . \tag{3.8.9}$$

For the asynchronous case,

$$P_A(z) = P_0 \frac{|C_{BA}|^2}{\mathcal{S}^2}\sin^2(\mathcal{S}z) \tag{3.8.10}$$

$$P_B(z) = P_0\left[\left(\frac{\Delta k}{2\mathcal{S}}\right)^2 \sin^2(\mathcal{S}z) + \cos^2(\mathcal{S}z)\right] \tag{3.8.11}$$

Complete power transfer is possible only between synchronous guides.

Continuing on, Lee [16] calculates the coupling coefficient between the two synchronous guides to be

$$C = C_{AB} = C_{BA} = \frac{2\omega^2\mu(\varepsilon_2 - \varepsilon_1)\alpha_y \cos^2(k_{2y}d/2)}{k_{zt}w_{eff}(\alpha_y^2 + k_{2y}^2)} e^{-\alpha_y s} \tag{3.8.12a}$$

where

$$w_{eff} = w + \frac{2}{\alpha_y}\left[\frac{k_{2y}^2 + \alpha_y^2}{k_{2y}^2 + \left(\frac{n_{e2}}{n_{e1}}\right)^4\alpha_y^2}\right]\left(\frac{n_{e2}}{n_{e1}}\right)^2 \quad \text{for quasi TE,} \tag{3.8.12b}$$

and $$w_{eff} = w + \frac{2}{\alpha_y} \quad \text{for quasi TM.} \tag{3.8.12c}$$

Using $$\omega^2\mu_0\varepsilon_0 = \left(\frac{2\pi}{\lambda_0}\right)^2 , \tag{3.8.13a}$$

$$k_{zt} = k_0 n_{et} , \tag{3.8.13b}$$

$$\alpha_x = k_0\sqrt{n_e^2 - n_1^2} , \tag{3.8.13c}$$

$$k_{2x} = k_0\sqrt{n_2^2 - n_e^2} , \tag{3.8.13d}$$

and $$\varepsilon_r = n^2 \ , \tag{3.8.13e}$$

(3.8.12a) can be simplified to

$$C = \frac{2\,\alpha_y \cos^2(k_{2y}\,w/2)\,e^{-\alpha_y s}}{k_0 n_{et} w_{eff}} \ . \tag{3.8.14}$$

From equations (3.8.8) and (3.8.9), the distance over which complete power transfer will occur (the *power transfer length* = L_{PT}) is evidently

$$L_{PT} = \frac{\pi}{2C} \ . \tag{3.8.15}$$

Practical directional coupler structures often take form shown in Fig. 3-10. Curved or angled waveguides are used to connect the parallel closely spaced channels. The coupling occurring between waveguides having a variable separation can be considered by referring to Fig. 3-11, a magnified view of the curve. The steps are normally present due to quantization of the electron beam position (typically to 0.1 μm) during mask writing. The steps are assumed to be so small that they induce negligible loss. To find the power in each guide at the end of the structure, suppose all the power is initially in guide B (this could be either one) and the coupling coefficients of the three sections are given by C_1, C_2, and C_3. At the end of ℓ_1,

$$P_A = P_0 \sin^2(C_1 \ell_1) \tag{3.8.16a}$$

$$P_B = P_0 \cos^2(C_1 \ell_1) \ . \tag{3.8.16b}$$

The same power could have been brought about, for example, if the first segment had length $\frac{C_1 \ell_1}{C_2}$ and coupling coefficient C_2. Using this as the second segment's input, its output is

$$P_A = P_0 \sin^2\left[\left(\frac{C_1 \ell_1}{C_2} + \ell_2\right) C_2\right] = P_0 \sin^2(C_1 \ell_1 + C_2 \ell_2) \tag{3.8.17a}$$

$$P_B = P_0 \cos^2\left[\left(\frac{C_1 \ell_1}{C_2} + \ell_2\right) C_2\right] = P_0 \cos^2(C_1 \ell_1 + C_2 \ell_2) \tag{3.8.17b}$$

In a similar manner, the power at the end of the third segment is

$$P_A = P_0 \sin^2\left[\left(\frac{C_1\ell_1 + C_2\ell_2}{C_3} + \ell_3\right)C_3\right] = P_0 \sin^2(C_1\ell_1 + C_2\ell_2 + C_3\ell_3) \quad (3.8.18a)$$

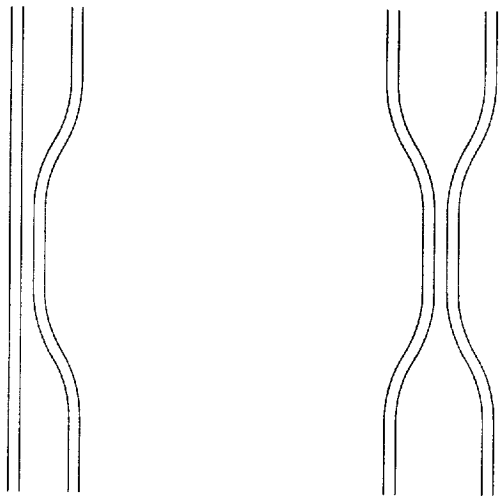
$$P_B = P_0 \cos^2\left[\left(\frac{C_1\ell_1 + C_2\ell_2}{C_3} + \ell_3\right)C_3\right] = P_0 \cos^2(C_1\ell_1 + C_2\ell_2 + C_3\ell_3) \quad (3.8.18b)$$

The rule is evidently

$$P_A = P_0 \sin^2\left(\sum_{i=1,2...}^{m} C_i\ell_i\right) \quad (3.8.19a)$$

$$P_B = P_0 \cos^2\left(\sum_{i=1,2...}^{m} C_i\ell_i\right) \quad (3.8.19b)$$

The quantity $L_{DC} = \sum_{i=1,2...}^{m} C_i\ell_i$ is here defined as the normalized coupling length. In phase matched directional couplers, when $L_{DC} = \pi/2$, complete power transfer occurs.



a) b)

Fig. 3-10 Designs where coupling can occur in connecting guides.

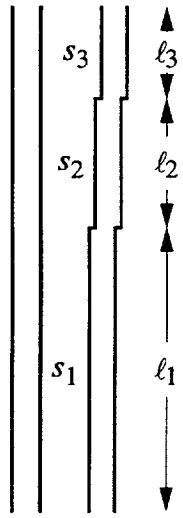


Fig. 3-11 Example for finding normalized coupling length for separating waveguides.

The additional coupling due to an arc connecting the coupled waveguides to the feed guides may be calculated with the aid of Fig. 3-8. The waveguide path, quantized to the mask generator grid, is assumed to snap to the nearest grid column, which are separated by distance s_c from each other. Assuming that the section ℓ_0 was drawn to exactly lie on a grid, after the arc passes through an angle θ_1, the nearest grid lies s_t distance from the coupler-channel grid line. From the figure, the relevant quantities are:

$$\begin{aligned} \theta_1 &= \cos^{-1}\left[1 - \frac{s_t/2}{R}\right] \\ \theta_2 &= \cos^{-1}\left[1 - \frac{3s_t/2}{R}\right] \\ &\vdots \\ \theta_m &= \cos^{-1}\left[1 - \frac{(2m-1)s_t/2}{R}\right] \end{aligned} \tag{3.8.20}$$

$$\begin{aligned} \ell_1 &= R\sin(\theta_1) \\ \ell_2 &= R\sin(\theta_2) - R\sin(\theta_1) \\ &\vdots \\ \ell_m &= R\sin(\theta_m) - R\sin(\theta_{m-1}) \end{aligned} \tag{3.8.21}$$

$$\begin{aligned} s_1 &= s_0 \\ s_2 &= s_0 + s_t \\ &\vdots \\ s_2 &= s_0 + (m-1)s_t \end{aligned} \qquad (3.8.22)$$

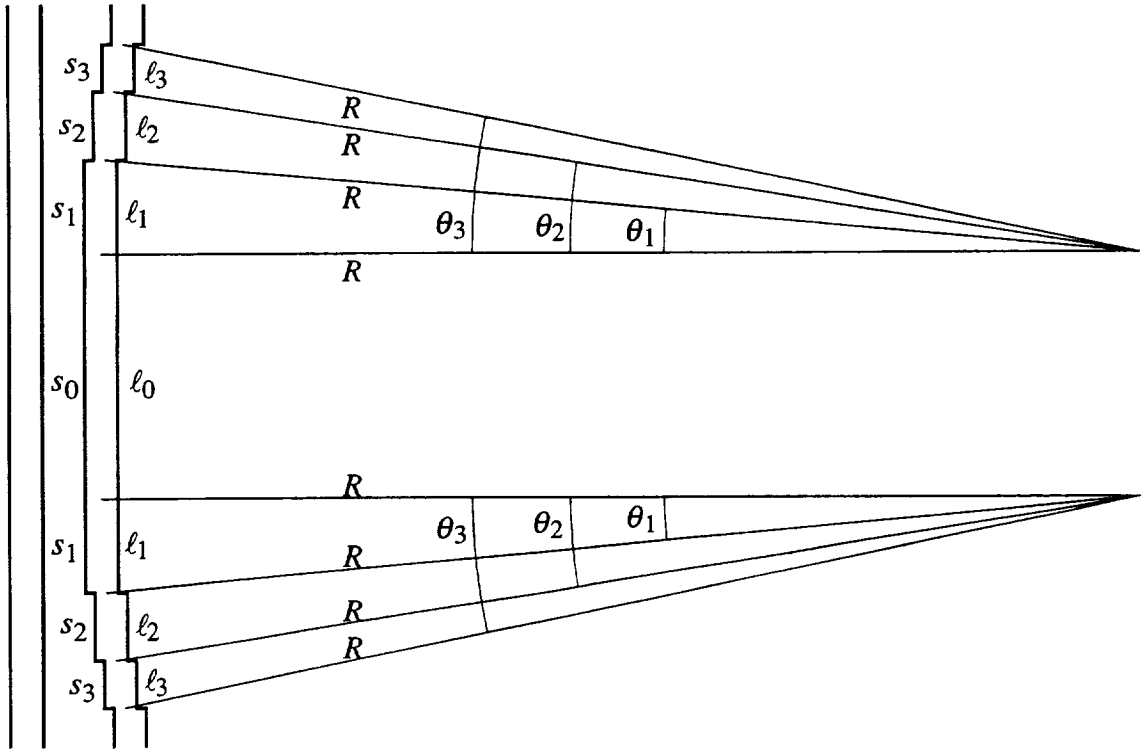


Fig. 3-12 Construction of quantization steps in directional coupler arc feeds.

If both channels of the coupler simultaneously arc away, (see Fig. 3-10) then (3.8.22) changes to

$$\begin{aligned} s_1 &= s_0 \\ s_2 &= s_0 + 2s_t \\ &\vdots \\ s_2 &= s_0 + 2(m-1)s_t \end{aligned} \qquad (3.8.23)$$

After the segment lengths and separations have been found, the coupling coefficient is calculated for each segment. The normalized coupling length of the directional coupler structure, including the curved feed channels is then given by

$$L_{DC} = \ell_0 C_0 + 2 \sum_{i=1,2...}^{m} C_i \ell_i \qquad (3.8.24)$$

where the summation is doubled because it is assumed that the lead-in and lead-out bends are identical. The transfer function of the phase matched directional coupler is given by (3.8.5) with L_{DC} is substituted for Cz. The summation in (3.8.24) is evaluated numerically on a computer to a sufficiently large m such that negligible coupling occurs in subsequent segments. If no quantization steps are present, they may still be assumed for computational purposes; the normalized coupling length is not a strong function of s_t.

Coupling in the curve has been calculated by assuming that the coupling coefficient's value at each position along the separating waveguides is the same as for two parallel guides having that separation. This is only accurate if the modes change adiabatically--if the change is so slow that the fields can adapt to the new structure [17]. Also, the field shift toward the edge of the waveguide away from the center of curvature when the light travels around a bend is ignored in this calculation.

Up till now, the field in the directional coupler has been described in terms of that in two isolated channels [see (3.8.1)]. Alternatively, we can calculate the fields in the five layer planar structure (Fig. 3-13) exactly [18]. These will be subsequently referred to as super modes since they encompass both guides. Typically each of the channels in isolation supports a single quasi-TE and quasi-TM mode. If $w_1 = w_2$, and the two guides are brought together so that $s = 0$, the structure will support two modes. The field for mode order $m = 0$ will be symmetric about $x = 0$ and possess the higher n_e while mode order $m = 1$ will be antisymmetric. As the waveguides are separated, two modes will continue to propagate but the difference between their effective indices will decrease, merging to the same value--their effective indices in isolation [19].

Fig. 3-14 illustrates these fields. For this example assume the power was coupled into the right channel. The initial amplitude and phase of the two fields is set by that combination which overlaps with least loss to the input field. Their phase relationship causes cancellation of the electric field in the right channel, producing the appearance of power only in the left one. When the super mode's phase difference changes by π, power

will appear to have transferred to the right channel, so that

$$L_{PT} = \frac{\pi}{k_0(n_{e\,m=0} - n_{e\,m=1})} \,. \tag{3.8.25}$$

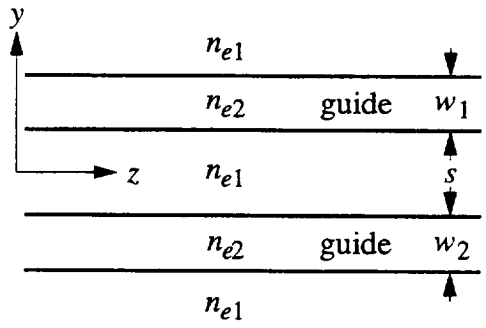


Fig 3-13 Five layer planar guide for finding coupling coefficient between two channels (using effective index method). The origin is taken as the midpoint between the inside edges of the two guides.

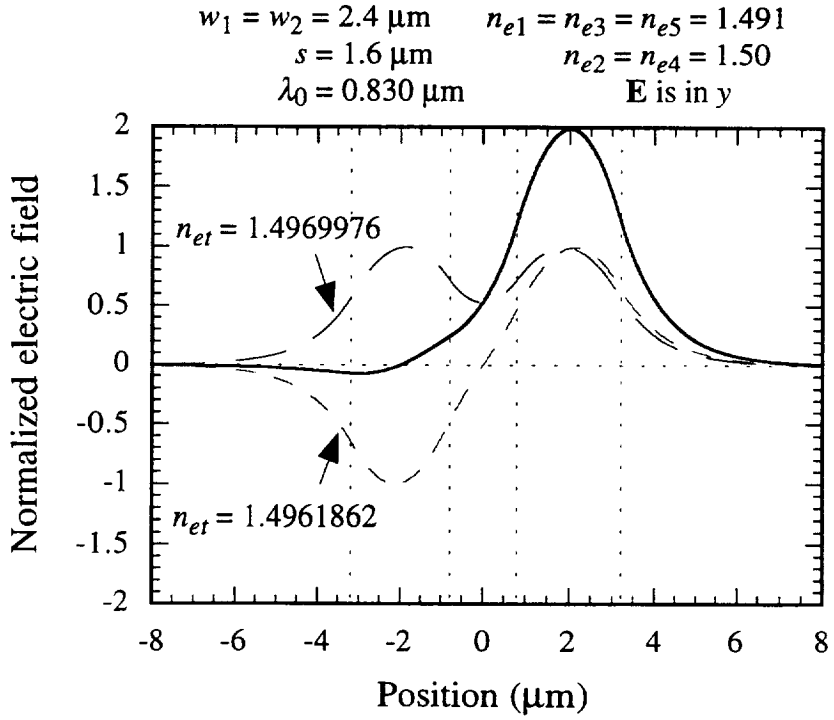


Fig. 3-14 Symmetric and antisymmetric quasi-TE super mode's electric field in the directional coupler plotted with their sum. The vertical dotted lines indicate the waveguide's extent. The coordinate system is given in Fig. 3-13.

The guidance equation for finding the super modes will now be presented. The quasi-TE (**E** in y) coupler's electric field is given by

$$E_y(y,z) = \begin{cases} A_1 \exp(-\alpha_{1y} y) & s/2 + w_1 < y < +\infty \\ A_2 \cos(k_{2y} y + \theta) & s/2 < y < s/2 + w_1 \\ A_3 \exp(-\alpha_{3y} y) + B_3 \exp(+\alpha_{3y} y) & -s/2 < y < s/2 \\ A_4 \cos(k_{4y} y + \psi) & -s/2 - w_2 < y < -s/2 \\ A_5 \exp(+\alpha_{5y} y) & -\infty < y < -s/2 - w_2 \end{cases} \exp(-jk_z z) \qquad (3.8.26a)$$

where

$$\begin{aligned} \alpha_{1y} &= k_0 \sqrt{n_{et}^2 - n_{e1}^2} \\ k_{2y} &= k_0 \sqrt{n_{e2}^2 - n_{et}^2} \\ \alpha_{3y} &= k_0 \sqrt{n_{et}^2 - n_{e1}^2} \\ k_{4y} &= k_0 \sqrt{n_{e2}^2 - n_{et}^2} \\ \alpha_{5y} &= k_0 \sqrt{n_{et}^2 - n_{e1}^2} \end{aligned} \quad . \qquad (3.8.26b)$$

As before, equation (3.8.26) can also represent the TM super mode's **H** field. The guidance condition can be derived in an analogous manner as the previous cases. The result is:

$$\tan^{-1}\left[\frac{\rho_2 \alpha_{3y}}{k_{2y}\rho_3} \frac{1 - C\exp(+\alpha_{3y} s)}{1 + C\exp(+\alpha_{3y} s)}\right] - \tan^{-1}\left(\frac{\rho_2 \alpha_{1y}}{k_{2y}\rho_1}\right) + k_{2y} w_1 = 0 \quad , \qquad (3.8.27a)$$

where
$$C \equiv \frac{B_3}{A_3} = \frac{1 - \dfrac{\rho_3 k_{4y}}{\alpha_{3y}\rho_4} \tan\left(-k_{4y}\dfrac{s}{2} + \psi\right)}{1 + \dfrac{\rho_3 k_{4y}}{\alpha_{3y}\rho_4} \tan\left(-k_{4y}\dfrac{s}{2} + \psi\right)} \exp(+\alpha_{3y} s) \qquad (3.8.27b)$$

and
$$\psi = -\tan^{-1}\left(\frac{\alpha_{5y}\rho_4}{\rho_5 k_{4y}}\right) + k_{4y}\left(\frac{s}{2} + w_2\right) , \qquad (3.8.27c)$$

with $\rho_{1\text{-}5} = n_{e1\text{-}e5}^2$ for the TM modes and $\rho_{1\text{-}5} = 1$ for the TE modes. Once n_{et} is iteratively determined, the field is found from (3.8.26a) and

$$\theta = \tan^{-1}\left(\frac{\alpha_{1y}\rho_2}{\rho_1 k_{2y}}\right) - k_{2y}\left(\frac{s}{2} + w_1\right) \qquad (3.8.28a)$$

$$A_1 = \frac{A_2 \cos\left[k_{2y}\left(\frac{s}{2}+w_1\right)+\theta\right]}{\exp\left[-\alpha_{1y}\left(\frac{s}{2}+w_1\right)\right]} \tag{3.8.28b}$$

$$A_3 = \frac{A_2 \cos\left[k_{2y}\frac{s}{2}+\theta\right]}{\exp\left(-\alpha_{3y}\frac{s}{2}\right)+C\exp\left(+\alpha_{3y}\frac{s}{2}\right)} \tag{3.8.28c}$$

$$A_4 = \frac{A_3\left[\exp\left(+\alpha_{3y}\frac{s}{2}\right)+C\exp\left(-\alpha_{3y}\frac{s}{2}\right)\right]}{\cos\left(-k_{4y}\frac{s}{2}+\psi\right)} \tag{3.8.28d}$$

$$A_5 = \frac{A_4 \cos\left[-k_{4y}\left(\frac{s}{2}+w_2\right)+\psi\right]}{\exp\left[-\alpha_{5y}\left(\frac{s}{2}+w_2\right)\right]} \tag{3.8.28e}$$

where A_2 is an arbitrary constant.

Unlike the guidance conditions for the three and four layer waveguides, no simple way was found to bracket each n_{et} to permit (3.8.27) to be solved using a bisection method. Therefore the left hand side of (3.8.27a) was plotted vs n_{et}, the zero crossings roughly identified, and a bisection method used to home-in on the precise zero crossing. Such a plot (Fig. 3-15) was used to find the n_{et}'s and obtain the electric field of Fig. 3-14.

The coupled mode solution [(3.8.15)] for the structure whose description is given in Fig. 3-15 predicts a power transfer length $L_{PT} = 515.8$ μm, in good agreement with $L_{PT} = 511.4$ μm found from (3.8.25). However, the five layer theory has the advantage of just as easily determining the coupling coefficient for asynchronous guides.

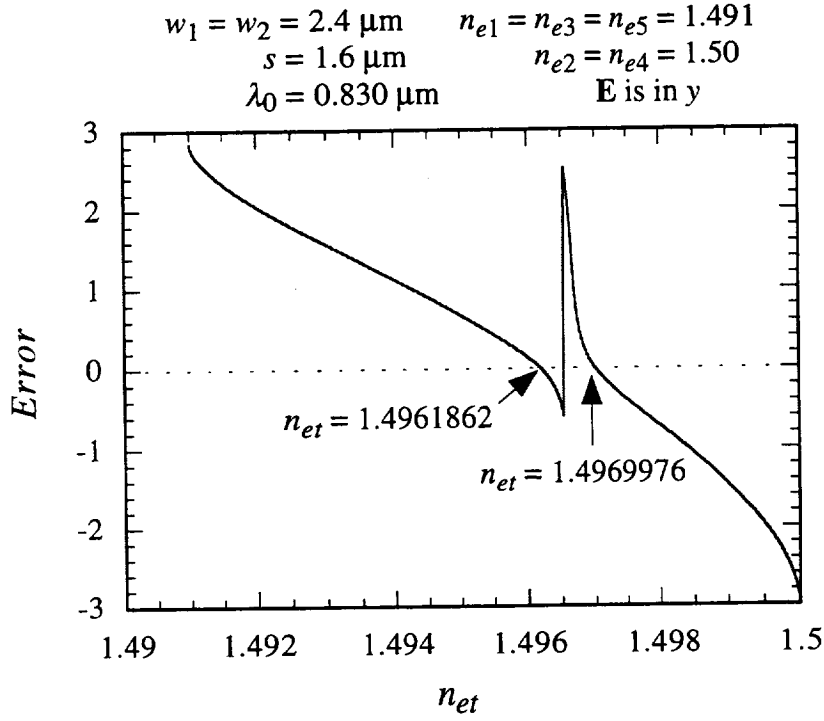


Fig. 3-15 Non-zeroness of the left hand side of (3.8.27a) vs n_{et} for the quasi-TE modes. The discontinuity occurs when $C\exp(+\alpha_{3y}s) = -1$ in (3.8.27a). The n_{et} of each of these guides in isolation is 1.4966118, approximately midway between the values above.

It is possible to accurately measure the coupling coefficient (or, where coupling occurs as the guides approach each other, the normalized coupling length) even if the transmissivity of the connecting waveguides is unknown. The four fiber link technique previously discussed in chapter 1 was shown to remove link transmissivity variations in transducers where the power split into two output fibers from two input fibers is a function of the measurand. Fig. 3-16 recalls the method. Treating the directional coupler in the box in Fig. 3-16 as lossless, or alternatively, lumping its loss with the channels, the ratio of coupled power to that which entered the directional coupler is determined from (3.8.5) to be

$$\kappa = \sin^2(C\ell_1) \quad , \tag{3.8.29}$$

where ℓ_1 is the length of the directional coupler, C is the coupling coefficient, and only one channel is excited at a time to avoid interference from both inputs. Since the coupler

is assumed lossless, the power which is transmitted through the incoupled channel is 1 κ. If coupling in the curved approached to the directional coupler contribute, a more general statement is

$$\kappa = \sin^2(L_{DC}) \quad , \tag{3.8.30}$$

where L_{DC} was defined in (3.8.24). The measured ratio R can be used to calculate the power coupling coefficient κ: independent of transmissivities T_1-T_4:

$$R = \frac{I_3|I_1 \text{ on}, I_2 \text{ off} \cdot \; I_4|I_1 \text{ off}, I_2 \text{ on}}{I_3|I_1 \text{ off}, I_2 \text{ on} \cdot \; I_4|I_1 \text{ on}, I_2 \text{ off}} \quad . \tag{3.8.31}$$

Inserting the transmissivities associated with the paths gives

$$R = \frac{I_1 T_1 (1-\kappa) T_3 \cdot I_2 \;\; T_2 (1-\kappa) T_4}{I_2 T_2 \kappa T_3 \cdot I_1 \;\; T_1 \kappa T_4} = \frac{(1-\kappa)^2}{\kappa^2} \quad . \tag{3.8.32}$$

κ is then

$$\kappa = \frac{1-\sqrt{R}}{1-R} \quad , \tag{3.8.33}$$

and the normalized coupling length is calculated to be

$$L_{DC} = \sin^{-1}(\sqrt{\kappa}) \quad . \tag{3.8.34}$$

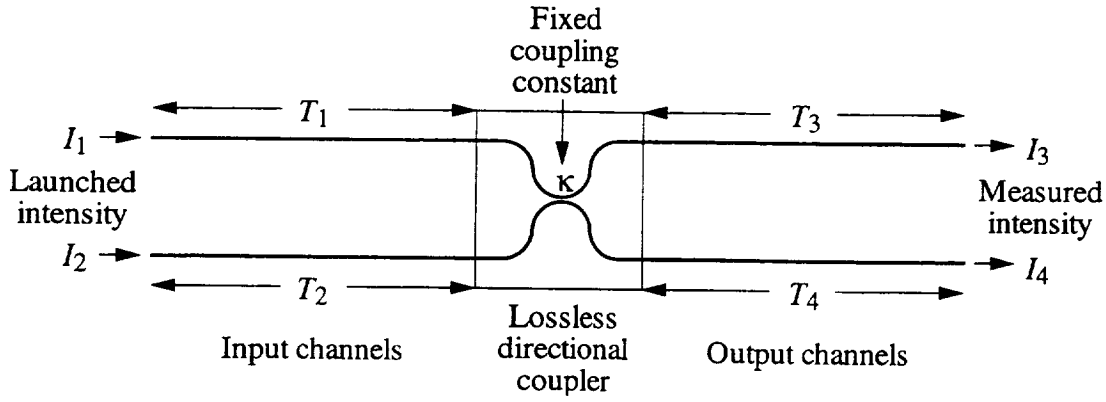


Fig. 3-16 Symbol definition.

It is necessary that I_1 and I_2 be constant while measurement of I_3 and I_4 are being made, so that I_1 and I_2 will truly cancel themselves in (3.8.32). In multimode channels,

the power each mode carries, and therefore the coupling coefficient, will be dependent on the coupling conditions. The value of κ calculated will be valid only for identical mode power ratios.

9 Y-JUNCTION

The Y-junction is one of the most commonly used structures in integrated optics. Compared to the directional coupler, it is less sensitive to fabrication and wavelength variations. Pictured in Fig. 3-17, it consists of a single mode guide of width w which tapers at a half angle θ_T to a guide of width $2w$. The waveguide of width $2w$ may or may not be single mode. The taper splits into two branches of width w which propagate away at half angle θ_B.

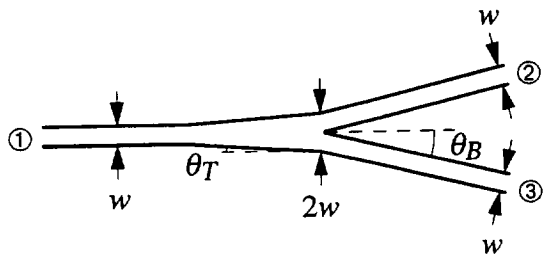


Fig. 3-17 The Y-junction.

Light may enter the device from either the taper or branch ends. Considering the former case, for low loss, the channel mode should expand adiabaticly in the taper, meaning the power should remain in the fundamental ($m = 0$) local normal mode, ensuring that a high percentage of the energy is delivered the end of the taper in this mode. Least loss will then occur, since, except for a slight angle, the splitter begins as two waveguides having zero separation; this has a similar field profile as one waveguide of width $2w$. Symmetry will prevent any of the energy from entering the channel mode $m = 1$, although if the taper is too abrupt, some of the energy may enter the channel mode $m = 2$, and radiate, since it is cutoff. Tapers for narrow guides can have larger θ_T; this is not surprising since the more highly confined and narrow the guide, the more rapidly the

light spreads out after leaving it. The same truism holds for laser beams--the smaller the spot they are focused to, the more rapidly they diverge afterward. The splitter loss is determined by how well the field at the taper termination overlaps with that guided by the branches, with a larger θ_B producing more loss. However, a shorter device with the same waveguide separation can be had when using a larger θ_B, an important consideration in Mach-Zehnder interferometers employed as electrooptic modulators [20]. In an ideal Y-branch the loss can be very low, but it is practically limited by the mechanisms discussed above and tip blunting [21], brought about in lithography [22] and etching.

Coupling can occur between the guides as they separate. However, in a symmetric device, when the initial electric field phase and amplitude is the same in each channel, only the symmetric super mode is excited (see the discussion on directional couplers). If the taper angle is small, it can adiabaticly evolve as the guides propagate away from each other. In this case, no periodic power transfer is expected. If dissimilar electric fields flow into ② and ③, both the antisymmetric and symmetric super modes will be excited. In the region where the guides couple, power exchange may appear to occur, but this can be viewed as beating between the two super modes--which does not affect the power each individually carries. The taper will radiate the second order super mode, and again coupling in the Y-branch will not affect the power which emerges from ①. However, in asymmetric Y-branches, mode splitting has been observed [23].

If the branches are independently excited, their electric fields may not possess the same phase and magnitude. Both the symmetric and anti-symmetric super modes will be created in the approaching branches. The anti-symmetric super mode will become the anti-symmetric ordinary mode ($m = 1$) which will be radiated as the channel tapers to single mode [24]. This energy will propagate away in the film adjacent to the channel if it planar guides [25] or be radiated into the substrate if it does not. If only one channel is excited, half the power will go into the symmetric and half into the antisymmetric super modes. Therefore half the incident power will be lost [26].

Based on this discussion, the power flowing into ① (see Fig. 3-17) will be split, with equal parts flowing out through ② and ③, and so

$$P_2 = P_3 = \frac{P_1}{2}(1-\gamma) \tag{3.9.1}$$

may be written. Here γ, which gives the power loss, lies in the range 0 (lossless) $< \gamma <$ 1 (all power lost). Since, for TE modes (and approximately so for TM modes) power is proportional to the electric field squared, and this proportionality constant is the same for all guides,

$$E_2^2 = E_3^2 = \frac{E_1^2}{2}(1-\gamma) \tag{3.9.2}$$

may be written. Taking the square root,

$$E_2 = E_3 = \frac{E_1}{\sqrt{2}}(1-\gamma)^{1/2} \tag{3.9.3}$$

is found for power traveling into ①. For power entering ② and ③,

$$E_1 = \left(\frac{E_2}{\sqrt{2}} + \frac{E_3}{\sqrt{2}}\right)(1-\gamma)^{1/2} \tag{3.9.4}$$

holds. It is instructive to analyze (3.9.4) for the lossless case ($\gamma = 0$) more closely. If $E_2 = E_3$, the power coming in ($2E_2^2$) equals the power leaving ($E_1^2 = \left(\frac{2}{\sqrt{2}}E_2\right)^2$). If, however, $E_2 = -E_3$, zero power leaves. When only one branch is energized, $E_1^2 = \frac{1}{2}E_2^2$ and half the power is radiated. The formulas will be employed in determining the transfer function of *Y*-branch equipped devices.

Both the taper angle and branch angle are needed for a *Y*-junction design. A waveguide having width w should be single mode but tighter low-loss bends can be achieved if the guide is nearly multimode. For this reason, tapers and branches at the single mode limit will be discussed. Beats and Lagasse used a beam propagation method to determine conversion efficiency for various taper angles from the fundamental mode at the input to the fundamental mode at the output of abrupt and tapered transitions as a function of the geometric mean of normalized channel width [25]. Among others, they treated the transition in which the input guide width was reduced by half. The

normalized channel width parameter V is defined as

$$V = k_0 w \sqrt{n_{e2}^2 - n_{e1}^2} \quad , \tag{3.9.5}$$

where w is the width of the channel and n_{e2} and n_{e1} are the effective indices of the planar guides if the channel were infinitely wide and zero width. Since $V = \pi$ in a guide at the single mode limit, and our taper posses this normalized width at its narrowest end, the geometric mean of the normalized thickness for the guide at the input and output of the taper is:

$$V_m = \sqrt{V_{in} V_{out}} = 4.44 \quad . \tag{3.9.6}$$

For this case, the following curve can be constructed from their data:

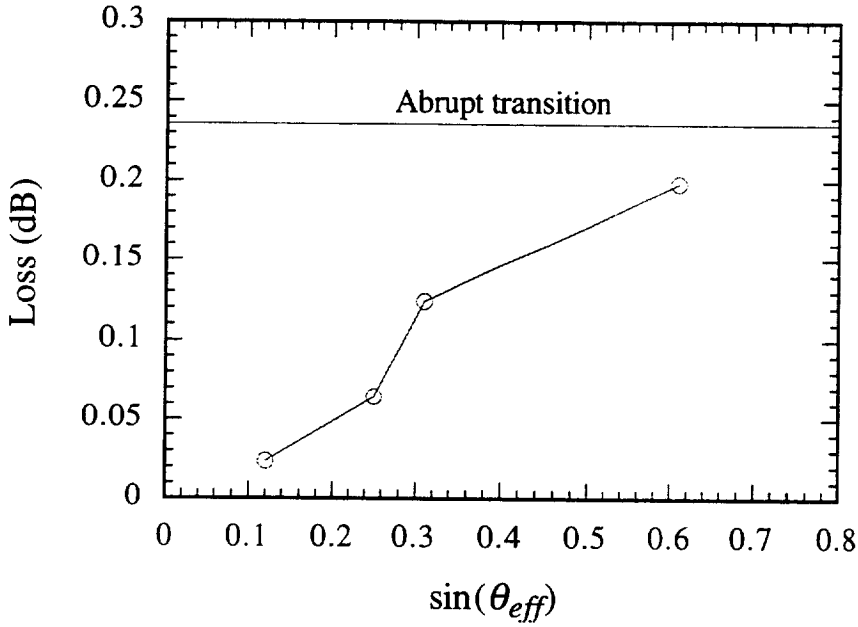


Fig. 3-18 Loss in a taper whose narrowest end is fed by a guide at the single mode limit. This data was read from Fig. 6A in reference [25].

where

$$\textit{effective taper angle} = \frac{\sin\theta_T}{NA} \tag{3.9.6}$$

and

$$NA = \sqrt{n_{e2}^2 - n_{e1}^2} \quad . \tag{3.9.7}$$

It can be seen that none of the transitions are very lossy.

Sasaki and Mikoshiba [21] have presented curves of the transmission in a Y-

branch for various normalized input guide thicknesses as a function of normalized branch angle. Such general curves could be calculated because the overlap intergal could be expressed in terms of normalized thickness and the product of the branch width, half angle, and propagation constant. They restricted their calculations to where the guide feeding the *Y*-branch (the channel emerging from the wider end of the taper) remained single mode. The guides feeding the taper would thus be restricted to being less than one-half of the width where they became multimode. However, tighter bends can be made if the wide end of the taper is allowed to become multimode, so a set of curves, shown in Fig. 3-19, were constructed which extended their analysis to this region.

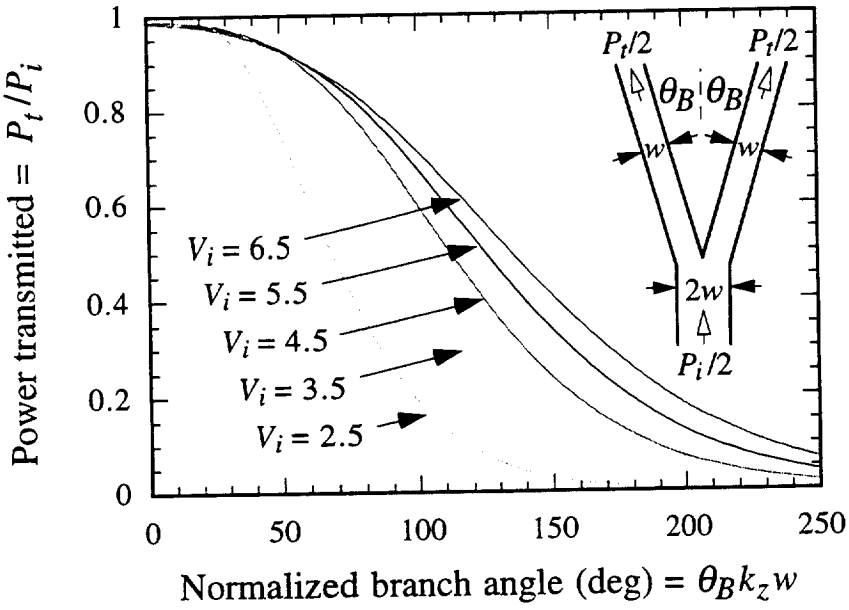


Fig. 3-19 Power transmitted vs normalized branch angle. The normalized input width is $V_i = 2k_0 w\sqrt{n_{e2}^2 - n_{e1}^2}$. When the guide of width w feeding the taper (not shown in the inset) is at the single mode limit, $V_i = 2\,\pi$. After [21].

Two approximations are made in this analysis [27]. First, the loss caused by the branches separating is ignored. Secondly, the field accepted by the branches in the overlap intergal is assumed to be the superposition of the fields of the waveguides in

isolation, with a phase factor to account for their tilt. However, at the input to the branch, as θ_B goes to 0, the field of these two waveguides does not converge to the field of a waveguide having twice each of their width. This is true because the field of a waveguide of width $2w$ is not equal to the fields of two isolated waveguides of width w which are separated by w and summed. This last assumption accounts for the failure of the power transmission in Fig. 3-14 to converge to 1 as the normalized branch angle approaches 0 degrees [27].

10 RING RESONATOR

Ring resonators have been employed as wavelength filters, in gyros, and to measure propagation loss of low loss light guides. While the transmittance of a Mach-Zehnder interferometer varies as $1 + \cos(\phi)$, where ϕ is the phase difference between the arms, the transmittance of a low loss ring resonator is high away from resonance and plunges abruptly at resonance. A figure of merit is finesse, the ratio of the separation between minima (in Hz) to the full width at half maximum (in Hz). High finesse ring resonators possess low loss rings. For wavelength division multiplexing, ring resonators permit notch filters to be constructed which have a flat transmittance and narrow blocking. Gyros benefit from the steep slope near resonance, which allows the precise phase difference between counter propagating waves to be discerned. Since in low loss ring resonators, finesse is sensitively affected by ring loss, finesse can be used to measure propagation loss.

A ring resonator consists of a light guide arranged as shown in Fig. 3-20. In device a, light launched into the straight channel couples into the ring via the directional coupler. Equation (3.8.5), which expresses the phase matched directional coupler input-output field relationship can be rewritten for this device:

$$E_3(z) = +E_1\cos(Cz) - jE_2\sin(Cz) \qquad (3.10.1a)$$

$$E_4(z) = -jE_1\sin(Cz) + E_2\cos(Cz)\ . \qquad (3.10.1b)$$

Evident in this equation is a property of the directional coupler: the phase of light coupled to the adjacent guide will lag 90° behind the guide from which it came [28]. Assume that after propagating around the ring and coming to ②, the light's phase lags by 90° that entering ①, a condition called resonance. This phase could be achieved by adjusting the wavelength of light, the refractive index of the ring or the physical length of the ring. After coupling over, the field from ① will be in phase with that which came over from ② and interfere constructively at port ④. However, in coupling, the field from ② will accumulate an additional 90° retardance and interfere destructively at ③. With each successive pass, the power will build in the ring until the loss of the ring matches the power added [29]. If out-scattering from the channels is significant enough to make the power in them visible, at resonance, the ring will appear bright and the output guide dim. Altering the phase at ② by 180° will reverse this situation.

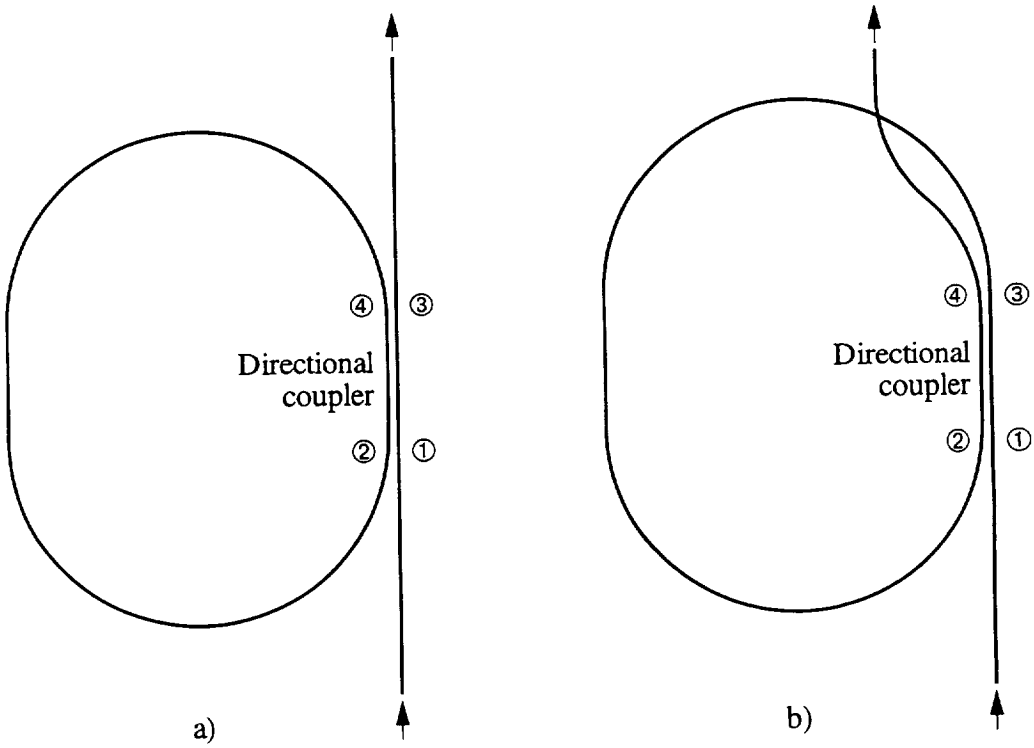


Fig. 3-20 Two styles of ring resonators with directional couplers. Style a is used with integrated optics to avoid the loss of a waveguide crossing, while b avoids a splice loss when implemented with optical fibers.

To simplify the analysis of the ring resonator in Fig. 3-20a, (3.10.1) can be rewritten by assigning

$$\kappa = \sin^2(C\ell_1) \tag{3.10.2}$$

where κ equals the ratio of power transferred into the adjacent channel to that launched into the source channel when only one channel is excited, and ℓ_1 is the length of the coupler. This substitution was discussed in conjunction with determining the coupling coefficient of the directional coupler in section 7. It follows that $\sqrt{\kappa} = \sin(C\ell_1)$ and $\sqrt{1-\kappa} = \cos(C\ell_1)$. (To include in κ coupling in the bends as the guides approach each other, the normalized coupling length L_{DC} may be calculated from (3.8.24) and L_{DC} substituted for $C\,\ell_1$ in these assignments.) After coupling for ℓ_1 distance, equations (3.10.1a,b) may be rewritten

$$E_3 = \sqrt{1-\kappa}\,E_1 - j\sqrt{\kappa}\,E_2 \tag{3.10.3a}$$

and

$$E_4 = -j\sqrt{\kappa}\,E_1 + \sqrt{1-\kappa}\,E_2 \ . \tag{3.10.3b}$$

The transfer function of the device in Fig. 3-20a will be determined with the use of the directional coupler transfer function (3.10.3) and the equation

$$E_2 = E_4 \exp(-\alpha\ell)\exp(-jk_{zt}\ell) \tag{3.10.4}$$

This method has been used by Stokes *et al.* [29] to find the transmissivity of the resonator shown in Fig. 3-16b. The waveguide in the directional coupler will be lumped with the ring and the coupler assumed to possess zero length, so that ℓ in (3.10.4) is the distance around the ring, including the segment in the coupler. The analysis of Stokes *et al.* [29] included a factor γ representing the loss of the coupler. However, their ring consisted of an unspliced piece of optical fiber, while their coupler fabrication required lapping the fiber nearly to the core; their coupler loss could be significant compared to the propagation loss of a short length of fiber. In the integrated optic ring resonator of Fig. 3-16a, the coupler loss is expected to be similar to the propagation loss of an equivalent length of waveguide segment. Using (3.10.4) to eliminate E_2 from (3.10.3) gives

$$E_3 = \sqrt{1-\kappa}\,E_1 - j\sqrt{\kappa}\,E_4 \exp(-\alpha\ell)\exp(-jk_{zt}\ell) \tag{3.10.5a}$$

$$E_4 = -j\sqrt{\kappa}\,E_1 + \sqrt{1-\kappa}\,E_4 \exp(-\alpha\ell)\exp(-jk_{zt}\ell) \quad . \tag{3.10.5b}$$

Equation (3.10.5b) can be simplified to be

$$E_4 = \frac{-j\sqrt{\kappa}\,E_1}{1-\sqrt{1-\kappa}\exp(-\alpha\ell)\exp(-jk_{zt}\ell)} \quad , \tag{3.10.6}$$

which is then used to eliminate E_4 from (3.10.5a), giving

$$\frac{E_3}{E_1} = \frac{\sqrt{1-\kappa} - \exp(-\alpha\ell)\cos(k_{zt}\ell) + j\exp(-\alpha\ell)\sin(k_{zt}\ell)}{1-\sqrt{1-\kappa}\exp(-\alpha\ell)\cos(k_{zt}\ell) + j\sqrt{1-\kappa}\exp(-\alpha\ell)\sin(k_{zt}\ell)} \tag{3.10.7}$$

At resonance, and for certain coupling κ, the transmissivity is zero. This occurs when the real and imaginary parts in the numerator of (3.10.7) are zero:

$$\sin(k_{zt}\ell) = 0 \tag{3.10.8}$$

so that $k_{zt}\ell = 0,\ \pi,\ 2\pi,\ \ldots$ The other condition is that

$$\sqrt{1-\kappa} = \exp(-\alpha\ell)\cos(k_{zt}\ell) \quad , \tag{3.10.9}$$

narrowing the choice of $k_{zt}\ell$ to $0,\ 2\pi,\ 4\pi\ldots$, and requiring that

$$\kappa = 1 - \exp(-2\alpha\ell) \quad . \tag{3.10.10}$$

The transmissivity is determined from (3.10.7) using the rule for complex numbers $\left|\frac{a+jb}{c+jd}\right|^2 = \frac{a^2+b^2}{c^2+d^2}$, which gives

$$\left|\frac{E_3}{E_1}\right|^2 = \frac{1-\kappa-2\sqrt{1-\kappa}\exp(-\alpha\ell)\cos(k_{zt}\ell)+\exp(-2\alpha\ell)}{1-2\sqrt{1-\kappa}\exp(-\alpha\ell)\cos(k_{zt}\ell)+(1-\kappa)\exp(-2\alpha\ell)} \quad . \tag{3.10.11}$$

11 RING RESONATOR WITH Y-JUNCTIONS

Ordinarily the ring resonator is constructed with a directional coupler and a low loss ring, since this produces the maximum finesse. However, the directional coupler power transfer length is sensitive to fabrication variations which change the lateral evanescent field tail (etch depth, width, etc). In addition, our waveguide technology will not produce a low loss ring (< 0.1 dB total loss), so even with a directional coupler we

will not obtain high finesse. Finally high finesse is not essential to the operation of our device. Therefore, the ring resonator configured with Y-branches depicted in Fig. 3-21 was constructed along with the more conventional directional coupler ring resonator.

The Y-junction ring resonator transfer function $\left|\frac{E_0}{E_i}\right|^2 = \frac{P_0}{P_i}$ can be determined with the aid of Fig. 3-21. The easiest approach is to assume some electric field E_0 at the output port ⑥ and calculate the input field E_i which must be present at ② to produce it. Using the formula which applies to the power splitting of symmetrical Y-junctions, (3.8.3), the E field which must be present at ④ is $\frac{E_0\sqrt{2}}{(1-\gamma)^{1/2}}$. As before, γ is the loss factor for each of the Y-junctions, assumed identical. To get this field, $\frac{E_0\sqrt{2}}{(1-\gamma)^{1/2}}\exp(+\alpha\ell_1)\exp(+jk_{zt}\ell_1)$ must have come from ③, where α is the attenuation of the channel waveguide. When the field at ⑤ traverses the ring to ①, it will become $E_0\exp(-\alpha\ell_2)\exp(-jk_{zt}\ell_2)$. This establishes all the fields shown in Fig. 3-21.

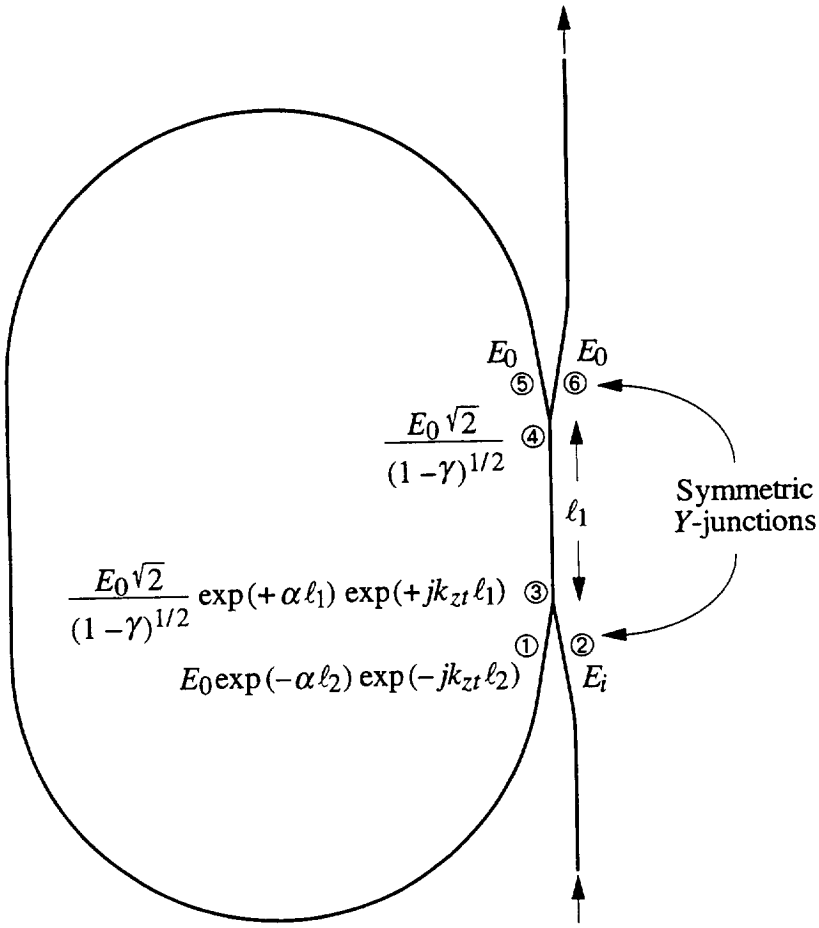


Fig. 3-21 Fields present in Y-branch ring resonator made up of single mode channels. ℓ_2 is the distance around the ring from ⑤ to ①.

Now the input field E_i will be determined. Writing (3.9.4) for the input Y-junction gives

$$\frac{E_0\sqrt{2}}{(1-\gamma)^{1/2}}\exp(+\alpha\ell_1)\exp(+jk_{zt}\ell_1)=\left[\frac{E_0\exp(-\alpha\ell_2)\exp(-jk_{zt}\ell_2)}{\sqrt{2}}+\frac{E_i}{\sqrt{2}}\right](1-\gamma)^{1/2} \quad . \quad (3.11.1)$$

or

$$\frac{E_0}{E_i}=\frac{\exp(-\alpha\ell_1)\exp(-jk_{zt}\ell_1)}{\frac{2}{1-\gamma}-\exp[-\alpha(\ell_1+\ell_2)]\cos[k_{zt}(\ell_1+\ell_2)]+j\exp[-\alpha(\ell_1+\ell_2)]\sin[k_{zt}(\ell_1+\ell_2)]} \quad . \quad (3.11.2)$$

Recalling that $\left|\frac{a+bj}{c+dj}\right|^2=\frac{a^2+b^2}{c^2+d^2}$, the power transmitted is

$$\left|\frac{E_0}{E_i}\right|^2 = \frac{\frac{1}{4}(1-\gamma)^2 \exp(-2\alpha\ell_1)}{1+\frac{1}{4}(1-\gamma)^2 \exp[-2\alpha(\ell_1+\ell_2)] - (1-\gamma)\exp[-\alpha(\ell_1+\ell_2)]\cos[k_{zt}(\ell_1+\ell_2)]} \quad . \tag{3.11.3}$$

This is the transmittance of the Y-junction ring resonator. The affect of loss on the transmittance function may be understood by defining $A = (1-\gamma)\exp[-\alpha(\ell_1+\ell_2)]$ in (3.11.3), where 0 (all power lost) $\leq A \leq 1$ (no loss). Equation (3.11.3) becomes

$$\left|\frac{E_0}{E_i}\right|^2 = \frac{\frac{1}{4}(1-\gamma)^2 \exp(-2\alpha\ell_1)}{1+\frac{1}{4}A^2 - A\cos[k_{zt}(\ell_1+\ell_2)]} \quad . \tag{3.11.4}$$

When $A = 0$, transmittance becomes constant, independent of path length. In contrast, the transmittance of a lossless device ($\gamma = \alpha = 0$, $A = 1$) is plotted in Fig. 3-22. The transmissivity varies between 1 and 1/9, and the maxima occur when $\cos[k_{zt}(\ell_1+\ell_2)] = 1$ or $k_{zt}(\ell_1+\ell_2) = 0, 2\pi, 4\pi, \cdots$. The minima occur when $k_{zt}(\ell_1+\ell_2) = \pi, 3\pi, 5\pi, \cdots$. Although waveguide dispersion gives n_e a λ_0 dependence, the effect is small and in the expression $k_{zt} = n_e k_0$, n_e may treated as a fixed constant even though λ_0 varies.

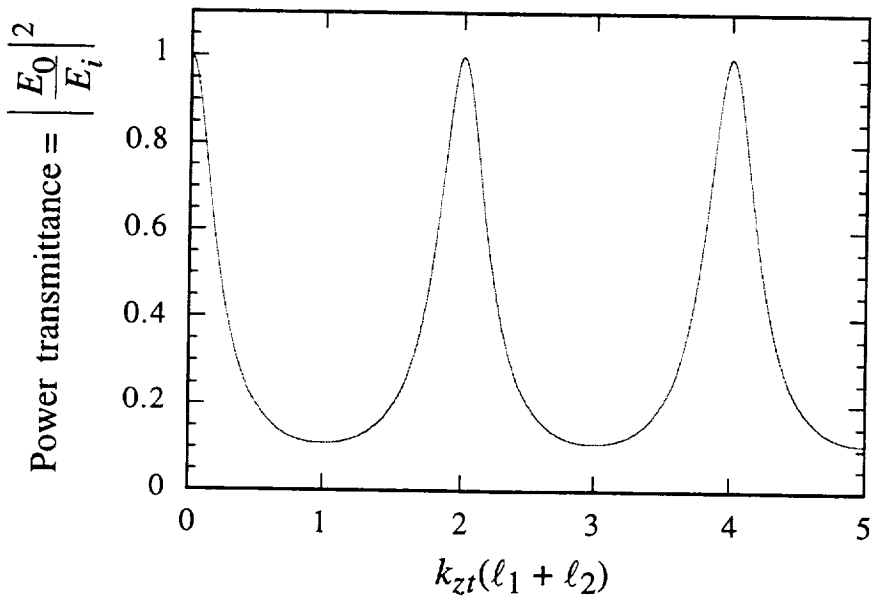


Fig. 3-22 Transmittance of a lossless Y-junction ring resonator.

An increase in *Y*-junction loss causes a drop in maximum transmissivity, makes the transmissivity curve (Fig. 3-22) appear more sinusoidal, and reduces the maxima-to-mimima transmittance ratio. This last quantity, which can be more accurately measured than absolute transmittance, can be used to calculate A in (3.11.4). The ratio is

$$Ratio = \frac{1 + A + \frac{1}{4}A^2}{1 - A + \frac{1}{4}A^2} \quad . \tag{3.11.5}$$

This can be inverted, becoming

$$A = \frac{2(\sqrt{Ratio} - 1)^2}{Ratio - 1} \quad . \tag{3.11.6}$$

12 MACH-ZEHNDER INTERFEROMETER

One of the most commonly implemented integrated optic structures is the Mach-Zehnder interferometer. Often used to construct electrooptic modulators, it can be constructed in zero and near-zero path length configurations without large angle bends. Fig. 23 shows the device. Light, split into two equal intensity beams by the first *Y*-junction, propagates through the arms and recombines at the second *Y*-junction. If the beams are in phase, only the symmetric mode is excited in the second *Y*-junction taper, and this is adiabatically compressed into the fundamental mode of the output waveguide. If the beams recombine out of phase, only the antisymmetric mode will be excited, and, for a single mode output channel, the taper will radiate the light into the planar guide adjacent to the channel or, if this does not exist, into the substrate. The input channel should also be single mode to insure that only the symmetric mode is excited in the input taper. Otherwise, the power split at the first *Y*-junction will be a function of coupling conditions and the length of the input channel, which will alter the phase relationship between the modes.

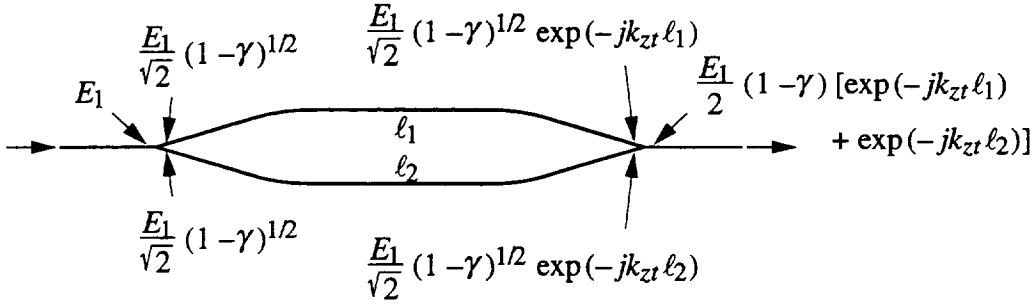


Fig. 3-23 Mach-Zehnder interferometer. The equations come from the discussion of *Y*-junctions. The factor γ supplies the power loss of the *Y*-junction.

Mach-Zehnder transfer function can be obtained from Fig. 3-23. The field ratio is

$$\frac{E_{in}}{E_{out}} = \frac{(1-\gamma)}{2}\{\cos(k_{zt}\ell_1) + \cos(k_{zt}\ell_2) - j[\sin(k_{zt}\ell_1) + \sin(k_{zt}\ell_2)]\} \quad . \quad (3.12.1)$$

The power transmissivity is then

$$\left|\frac{E_{in}}{E_{out}}\right|^2 = \frac{(1-\gamma)^2}{2}\{1 + \cos[k_{zt}(\ell_1 - \ell_2)]\} \quad , \quad (3.12.2)$$

where γ provides the *Y*-junction loss (0 for the lossless case), ℓ is the physical length of the arm, and k_{zt} is the channel wavenumber.

13 WAVEGUIDE BENDS

Waveguide bends are an important part of ring resonators, large path difference Mach-Zehnder interferometers, and integrated optic interconnects. There are two kinds of losses associated with them: 1) bending loss, which occurs continuously throughout the bend, and is the result of the evanescent field on the side of the channel away from the arc center radiating, and 2) transition loss, which occurs at discontinuities in the radius of curvature [30]. Both will be examined.

In a straight channel, the phase fronts are planar. When the mode propagates around a bend, the velocity of a point on the phase front will be proportional to its distance from the center of curvature. Beyond some radius, the phase front's velocity will

exceed c/n, where c is the speed of light and n is the film's refractive index. Energy in this portion of the tail will radiate, but is continuously resupplied by the power in the mode, producing bending loss. Since the exponential field tail, in theory, never attains zero amplitude, all bends are lossy. The problem is to determine the loss due to the bend, or conversely, to find the tightest bend at which this loss remains negligible.

Lee [31] analyzed the problem of the bent slab waveguide, shown in Fig. 3-24, by treating the bend as a radiating aperture. Marcuse [32], using much more involved analysis, determined the bending loss by solving the wave equation in cylindrical coordinates. His equation is similar to Lee's except a 2/3 coefficient is missing from Lee's equation. Cast in effective index notation, Marcuse's equation is:

$$\alpha = \frac{\alpha_{1y}^2}{k_{zt}\left(1+\alpha_{1y}\frac{w}{2}\right)\left(n_{e2}^2-n_{e1}^2\right)k_0^2}\frac{k_{2y}^2}{}\exp(\alpha_{1y}w)\exp\left(-\frac{2}{3}\frac{\alpha_{1y}^3}{k_{zt}^2}R\right) \qquad (3.13.1)$$

where α is the power decay coefficient, *e.g.*, $I = I_0\exp(\alpha\ell)$, where I and I_0 are the intensities at the beginning and end of the arc of circumferential length ℓ. Marcuse's electric field is tangential to the slab's interface (TE modes) so his equation applies to the bent channel's quasi-TM mode. However, since n_{e1} is very close to n_{e2}, little error should result if quasi-TE n_{e1} and n_{e2} are used in his expression. As is the case with other polarization dependent quantities such as the coupling coefficient of the directional coupler, the difference between the bending loss of the two channel polarizations is due to differences in the effective index contrast $(n_{e2} - n_{e1})$.

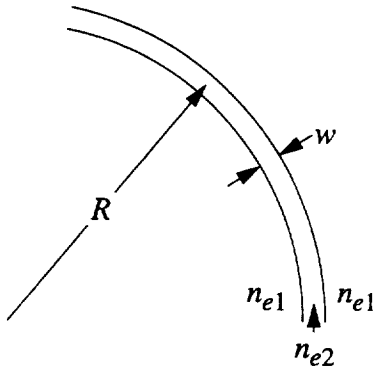


Fig. 3-24 Bent waveguide.

Calculating bending loss using (3.13.1) requires knowledge of the lateral field profile. However, it has been shown with directional coupler power transfer calculations that the effective index method can only accurately estimate this profile if the region outside the channel planar guides [2].

Equation (3.13.1) has been compared to bending loss calculated using a matrix method in which the bent channel was replaced by a straight one having an effective index profile consisting of the original effective index profile summed with a function which increases with R [33]. Such a waveguide exhibits the field distortion induced by the bend, including the peak shift away from the center of curvature. In addition, since at some radius of curvature the refractive index increases above that under the ridge, the channel will be leaky, similar to a planar guide constructed on an isolation layer deposited on a higher index substrate. These authors considered bending radii where the propagation loss was less than 2 dB/cm. Near the maximum bending loss the authors considered, 2 dB/cm, Marcuse's formula overestimated the loss by a factor of 2. At bending radii where the loss was around 10^{-4} dB/cm, Marcuse's formula overestimated the loss by a factor of 20. In this last case, however, the loss levels are well below that of most channels, and so the difference would not be discernible.

Gu *et al.* [34] compared bending loss calculated using the effective index method and a theory yielding similar results as (3.13.1) to a numerical solution of the wave

equation in cylindrical coordinates. They also compared these two theories to the bending loss data Deri *et al.* [35] measured on a GaAlAs rib waveguide which calculations indicate would planar guide outside the channel. Although they found the loss calculated from a numerical solution of the wave equation to agree better with the measured loss, the difference between the two theories was slight, especially around 1 dB/cm.

Bending loss can be reduced by increasing the decay rate of the exponential tail extending radially out from the channel. This can be achieved, for example, by increasing channel effective index contrast. In order to maintain single mode guiding, the ridge width must also be reduced. Fig. 3-25 shows the bending loss vs waveguide width for a silicon oxynitride channel waveguide. For each ridge width plotted, the etch depth was chosen so that the waveguide was single mode for both polarizations. For this particular waveguide the quasi-TE polarization was found to become multimode before the quasi-TM and the former thus constrained the etch depth. Somewhat greater ridge heights were found to reverse which first became multimode. The effective index method was used to calculate the etch depth and (3.13.1) to find the bending radius. Reducing the width permits significant decreases in bending radius for the same bending loss. In Fig. 3-26, the ridge width was also varied, but the etch depth was fixed to the value for single mode operation at 4 μm. Here the bending radius reaches a minimum which is slightly less than that when the width was at single mode cutoff for the quasi TE mode. The "U" shape occurs because for a constant etch depth (constant difference between n_{e2} and n_{e1}), the lateral field width is a minimum when the channel width is near single mode cutoff [25].

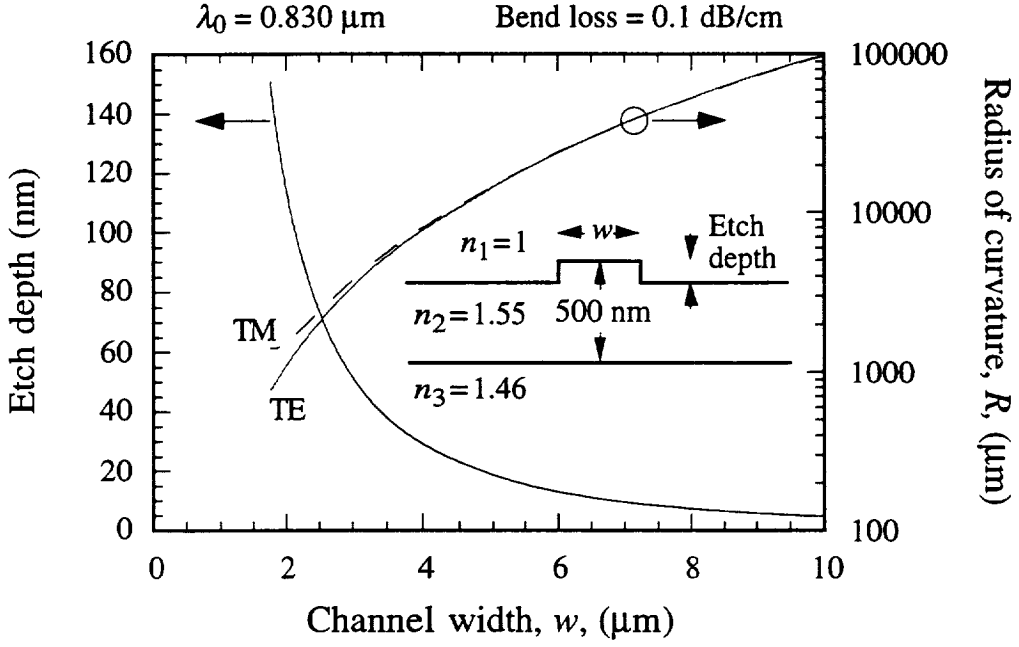


Fig. 3-25 Radius of curvature for the quasi-TE and quasi-TM modes which produces a bending loss of 0.1 dB/cm. The etch depth was adjusted as the width was varied to maintain the channel at the single mode limit.

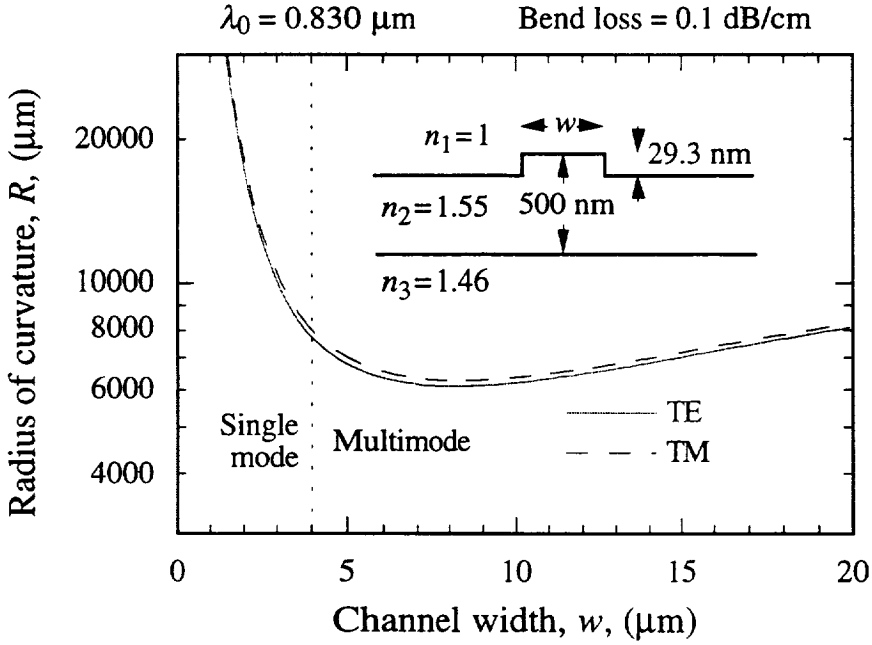


Fig. 3-26 Radius of curvature for the quasi-TE and quasi-TM modes which produces a bending loss of 0.1 dB/cm. The etch depth was fixed to that value which made the channel single mode when its width was 4 μm.

A second loss mechanism in curved optical waveguides, termed transition loss, is due to curvature discontinuity. When traveling around a bend, the portion of the phase front on the inside of the bend must move more slowly than that on the outside of the bend. The field accomplishes this by shifting its midpoint toward the outer channel wall. Thus more of the field on the inside of the bend sees a higher effective index, causing it to slow down [30]. At a discontinuity in curvature, the field distributions do not match (Fig. 3-27 left), producing loss. A solution easily implemented in integrated optics (Fig. 3-27 right) is to offset the channels [30, 36-38]. Another technique is to use a sinusoidal bend shape [39]. Although this eliminates discontinuities in curvature, it increases the length of the transition, since some of the guide in the bend possesses a larger of radius of curvature.

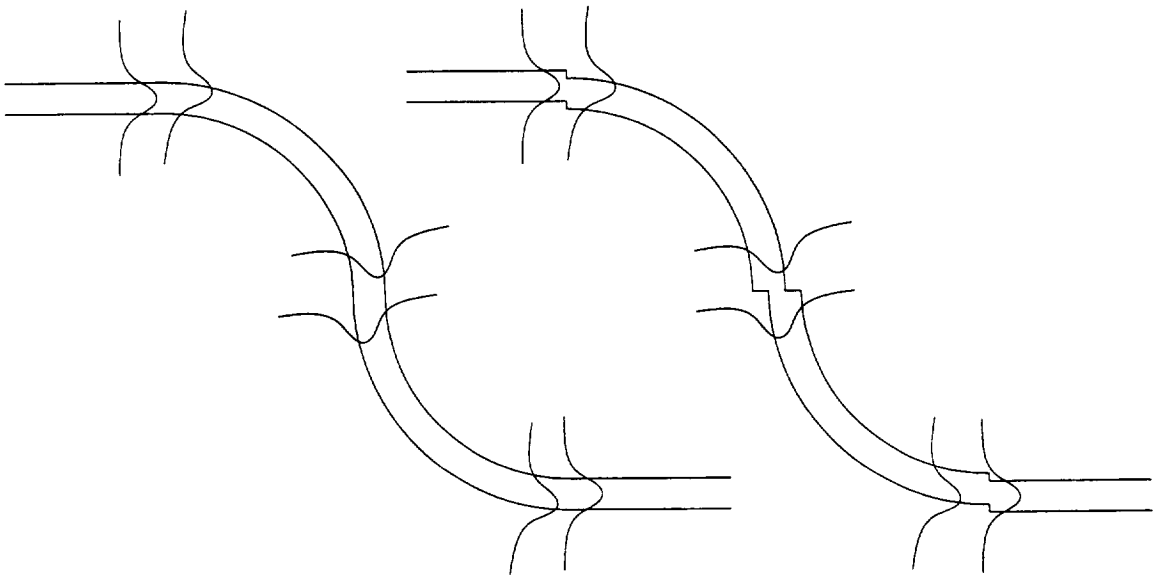


Fig. 3-27 (Left) Field mismatch at abrupt changes in curvature, and (right), offsetting channels to reduce curvature loss.

The loss due to curvature discontinuity does not vary with radius as rapidly as bending loss, and because calculations indicated that this loss would be negligible for most devices, the corrections shown in Fig. 3-27 were omitted in most channel bends in the mask designed for this work. The loss it contributes will now be considered. An

analog to transition loss in optical fibers is termed microbending loss. Petermann [40] developed an approximation for the electric field profile in bent optical fibers which can be applied to bent channel waveguides. Later Gambling *et al.* [41], using Petermann's results, determined the peak displacement and coupling loss. Their formulas will now be presented, recast for channel waveguides. The electric field profile in a straight guide is assumed Gaussian

$$E_0(x) = A \exp\left(\frac{-x^2}{2\omega^2}\right) \tag{3.13.2}$$

but in a bent guide takes the form [40, 41]

$$E(x) = E_0(x)\left[1 + \frac{1}{R}(k_0 n_{e2}\omega)^2 x\right] , \tag{3.13.3}$$

where E_0 is given in (3.13.2), A is an arbitrary constant, and ω represents the field width. To use these results for channel waveguides the transverse field must be fit to a Gaussian.

The field peak displacement d can be determined from (3.13.3) by setting $dE(x)/dx = 0$. This results in

$$d = \frac{\sqrt{1 + 4(k_0 n_{e2}\omega)^4 \omega^2/R^2} - 1}{2(k_0 n_{e2}\omega)^2/R} . \tag{3.13.4}$$

Since the quantity summed with 1 under the square root is small compared to 1, (3.13.4) can be approximated

$$d \approx \frac{\omega^4 (n_{e2} k_0)^2}{R} . \tag{3.13.5}$$

This is the field peak displacement in a bent channel waveguide. The transmittance between a straight and a bent guide can be determined from the overlap integral

$$\eta = \frac{\left|\int_{-\infty}^{\infty} E(x)\, E_0^*(x)\, dx\right|^2}{\int_{-\infty}^{\infty} E(x)\, E^*(x)\, dx \int_{-\infty}^{\infty} E_0(x)\, E_0^*(x)\, dx} . \tag{3.13.6}$$

Inserting (3.13.2) and (3.13.3) into (3.13.6) gives

$$\eta = \frac{1}{1 + \frac{\omega^2}{2}\frac{(k_0 n_{e2}\omega)^4}{R^2}} \quad . \tag{3.13.7}$$

If the quantity added to 1 is much smaller than 1, (3.13.7) can be written

$$\eta \approx 1 - \frac{1}{2}\omega^6 \frac{(k_0 n_{e2})^4}{R^2} \quad . \tag{3.13.8}$$

The transition loss is approximately four times greater when two waveguides having the same radius of curvature and bending in opposite directions join [38].

A method of reducing transition loss is to offset the channels, as depicted in Fig. 3-27. The transmittance η between curved waveguides bent in opposite directions and having a different radius of curvature can be calculated as before, yielding

$$\eta = \frac{\exp\left(-\frac{a^2}{2\omega^2}\right)\left\{\left[1 + \frac{(k_0 n_{e2}\omega)^2 a}{2R_1}\right]\left[1 + \frac{(k_0 n_{e2}\omega)^2 a}{2R_2}\right] - \frac{(k_0 n_{e2}\omega)^4 \omega^2}{2R_1 R_2}\right\}^2}{\left[1 + \frac{(k_0 n_{e2}\omega)^4 \omega^2}{2R_1^2}\right]\left[1 + \frac{(k_0 n_{e2}\omega)^4 \omega^2}{2R_2^2}\right]} \quad , \tag{3.13.9}$$

where the a and R are defined in Fig. 3-28. This is a more general result and the cases discussed previously can be derived from it. It can be applied to straight-to-arc connections by setting one of the R's to infinity.

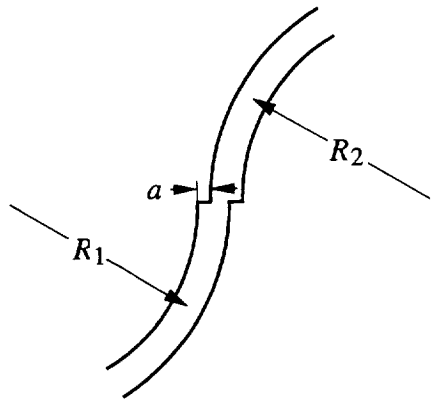


Fig. 3-28 Definitions of variables in (3.13.9). Drawing is for $a > 0$.

Waveguides typical of those used in this study consist of silicon oxynitride layers

and have an effective index around 1.5. Fig. 3-29 compares the calculated lateral electric field of the quasi TM mode in a waveguide at lateral single mode cutoff having the effective indices and width shown in the inset with a best fit Gaussian. The agreement obtained for the width parameter $\omega = 1.22$ μm suggests that the Gaussian approximation is reasonable. Transition loss for various offsets between two channel arcs having oppositely directed bends and between a straight guide and an arc is plotted in Fig. 3-25. The radius of curvature is 3500 μm and this design possesses negligible bending loss. Transition loss is minimized when the offset is equal to the peak displacement d, for connections between the straight channel and the arc, or $2d$ for connections between two arcs. Because of the small field width ω, Fig. 3-30 shows that transition loss is negligible even without the offsets.

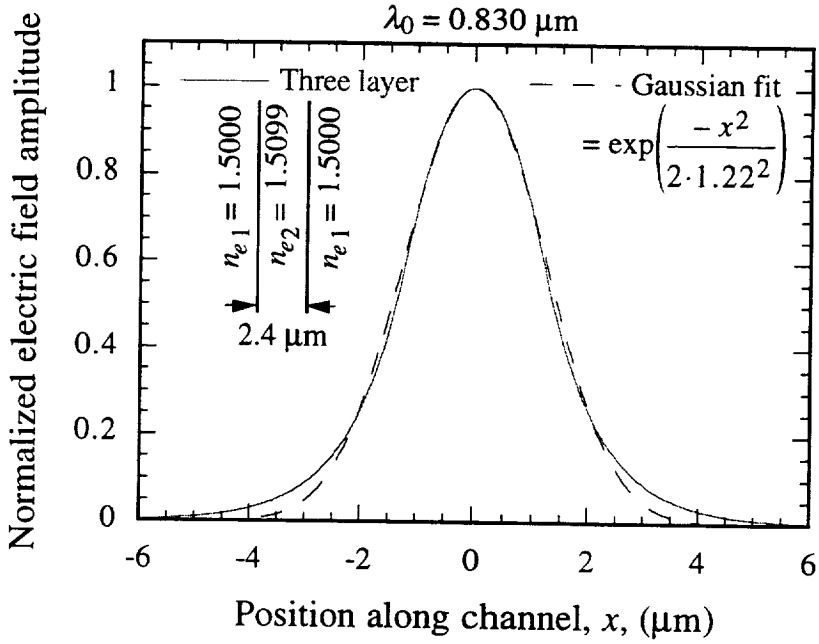


Fig. 3-29 Comparison of the three layer symmetric waveguide model with a best fit Gaussian.

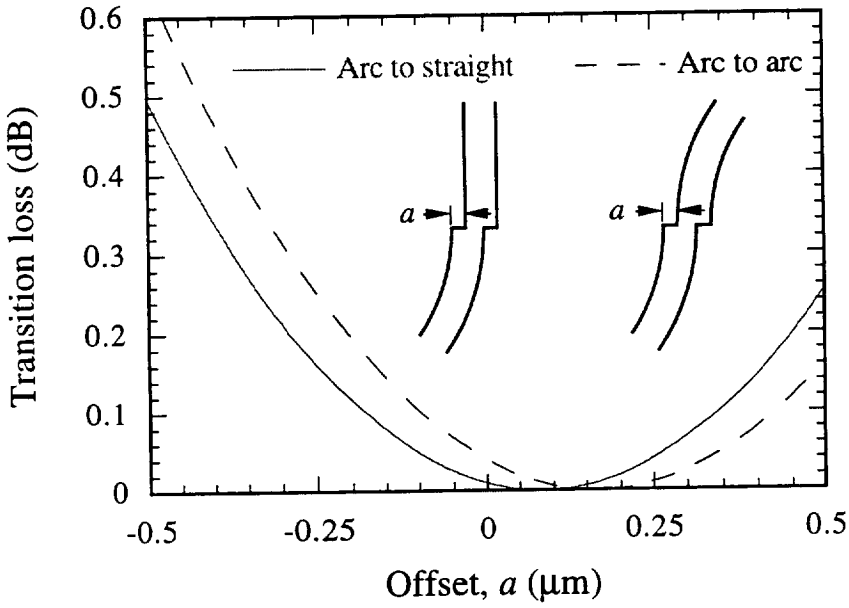


Fig. 3-30 Transition loss for various offsets between the channels shown in Fig. 3-29.

14 COHERENCE LENGTH

The sensor explored in this project is an interferometer whose path length is a function of the measurand. There are many ways such an interferometer can be interrogated (its path length determined). The scheme chosen depends on whether the interferometer's path length is longer than the coherence length of the source. Some of the schemes, along with their advantages and disadvantages will be discussed.

All optical sources may be classified according to their coherence length, or broadness of linewidth. A source's coherence length is equal to the maximum path length an interferometer can have and still have its fringes resolved by the source. Consider the transfer function of a lossless Mach-Zehnder interferometer, as derived in the previous section,

$$\frac{I_{out}}{I_{in}} = 1 + \cos(\phi) \tag{3.14.1}$$

where $\phi = k_0 OPLD$, the optical path length difference is given by $OPLD = n_{et}\Delta\ell$ and $\Delta\ell$ is the difference between the physical lengths of the arms. Using $c_0 = \lambda_0 f$, where c_0 is

the speed of light in a vacuum, and f is the optical frequency, k_0 can be written $k_0 = 2\pi \frac{f}{c_0}$. The wavenumber k_0 and thus ϕ are proportional to optical frequency. Mach-Zehnder minima, which occur at $\phi = \pi$, $\phi = 3\pi$, etc. are thus separated by wavenumber $k = 2\pi/OPLD$ (Fig. 3-31). If the source has significant energy over a wavenumber span greater than this, movement of the interferometer minima due to changes in its OPLD will not affect the integrated intensity reaching the detector. Rather, movement will simply alter the spectral distribution reaching the detector. Therefore, a source having coherence length L_{coh} will have spectral broadness $\Delta k = 2\pi/L_{coh}$.

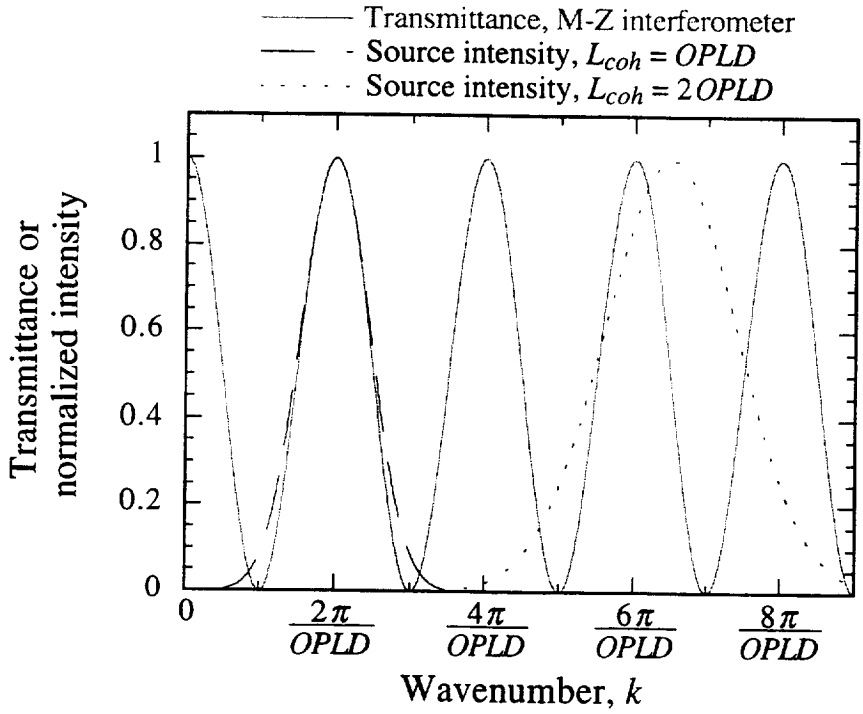


Fig. 3-31 Mach-Zehnder transmittance and the spectral power densities of sources having coherence length equal to OPLD and twice OPLD.

The relation $\Delta k = 2\pi/L_{coh}$ may be rewritten in a simpler form by noting

$$k_0 + \frac{\Delta k}{2} = \frac{2\pi}{\lambda_{min}} \tag{3.14.2}$$

and

$$k_0 - \frac{\Delta k}{2} = \frac{2\pi}{\lambda_{max}} \; . \tag{3.14.3}$$

Subtracting (2) from (1) and using $\Delta k = 2\pi / L_{coh}$ gives

$$\frac{1}{L_{coh}} = \frac{1}{\lambda_{min}} - \frac{1}{\lambda_{max}} \tag{3.14.4}$$

or

$$L_{coh} = \frac{\lambda_{max}\lambda_{min}}{\lambda_{max} - \lambda_{min}} \quad . \tag{3.14.5}$$

Defining $\lambda_{max} - \lambda_{min}$ to be $\Delta\lambda$ and $\lambda_{max}\lambda_{min}$ to be λ_0 gives a definition for coherence length

$$L_{coh} = \frac{\lambda_0^2}{\Delta\lambda} \quad . \tag{3.14.6}$$

15 PHOTOELASTIC EFFECT

This section provides the relationships to determine how mechanical strain, discussed in chapter 2, alters the refractive index of the waveguide layers. Once the change in the refractive index is found, the change in the effective index is calculated using the methods given in this chapter.

Strain-induced change in the refractive index of matter is called the photoelastic effect. In most materials decreasing their density (positive strain) makes them more vacuum like and reduces their refractive index. The phenomenon has been used to study, in transparent models, the flow of stress; it can be used to determine if residual stress exists in fabricated glass objects; and it permits the sound-induced creation of gratings which can diffract light. It is of interest here because it is the mechanism by which mechanical strain, induced by pressure, perturbs the refractive index of the channel waveguide, altering its path length.

The photoelastic effect can be understood with the aid of the index ellipsoid shown in Fig. 3-32 and described by

$$\frac{x^2}{n_x^2} + \frac{y^2}{n_y^2} + \frac{z^2}{n_z^2} = 1 \quad . \tag{3.15.1}$$

This construct may also be used to determine the behavior of light propagating in

optically anistropic crystals. In fact, isotropic materials become anistropic in the presence of a strain field. A plane wave propagating in an optically anistropic material does not simply see some effective index. Rather, the light couples into two eigenpolarizations, called the ordinary and extraordinary rays; these propagate through the material, and upon reaching the exit face, recombine into a plane wave. These rays have different propagation velocities since they see different refractive indices, and light traversing an anistropic material can emerge having a different polarization.

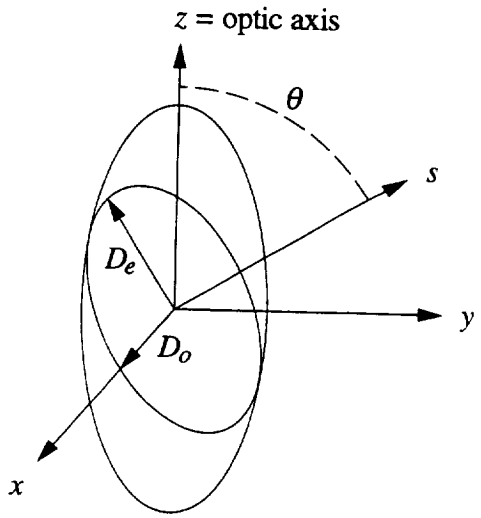


Fig 3-32 The index ellipsoid, shown here for a uniaxial material. The ordinary ray is polarized in x and sees refractive index D_o, while the extraordinary ray, polarized parallel to D_e, sees a refractive index equal to D_e. The light propagates in s and the optic axis is in z [44].

To find the polarization direction and refractive index associated with the two rays, a plane perpendicular to the direction of propagation s and passing through the origin is constructed (Fig. 3-32) [45]. The intersection of this plane with the index ellipsoid will be an ellipse. The half-length and direction of the semi-major and major axes of this ellipse gives the refractive index and electric flux density (D) vector orientation of each of the two permitted rays. In optically anistropic crystals, the coordinate frame of Fig. 3-32 should coincide with the crystal frame, with the z direction parallel to the crystal's optic axis. In optically isotropic materials such as an unstressed

waveguide films, the index ellipsoid reduces to a sphere with $n_x=n_y=n_z=n$. A single normal strain will make the material uniaxial, and its optic axis will be parallel to the direction of applied strain; the cross section of the ellipsoid in the x-y plane will be circular so $n_x=n_y\neq n_z$. Two or more normal strains will make the material biaxial, with $n_x\neq n_y\neq n_z$, and the x-y cross section will be elliptical. Shear strains will rotate the index ellipsoid. For the uniaxial case shown in Fig. 3-32, with propagation in the y-z plane, light in the ordinary ray polarized in x will see a refractive index independent of θ. The extraordinary ray, polarized perpendicular to x, will see a refractive index dependent on θ and but it will equal that of the ordinary ray when $\theta=0$.

When the optical material becomes strained, the index ellipsoid is written [46]

$$\begin{aligned}
&x^2\left(\frac{1}{n_x^2}+p_{11}\,\epsilon_x+p_{12}\,\epsilon_y+p_{13}\,\epsilon_z+p_{14}\,\epsilon_{yz}+p_{15}\,\epsilon_{xz}+p_{16}\,\epsilon_{xy}\right)\\
&+y^2\left(\frac{1}{n_y^2}+p_{21}\,\epsilon_x+p_{22}\,\epsilon_y+p_{23}\,\epsilon_z+p_{24}\,\epsilon_{yz}+p_{25}\,\epsilon_{xz}+p_{26}\,\epsilon_{xy}\right)\\
&+z^2\left(\frac{1}{n_z^2}+p_{31}\,\epsilon_x+p_{32}\,\epsilon_y+p_{33}\,\epsilon_z+p_{34}\,\epsilon_{yz}+p_{35}\,\epsilon_{xz}+p_{36}\,\epsilon_{xy}\right)\\
&+2yz(p_{41}\,\epsilon_x+p_{42}\,\epsilon_y+p_{43}\,\epsilon_z+p_{44}\,\epsilon_{yz}+p_{45}\,\epsilon_{xz}+p_{46}\,\epsilon_{xy})\\
&+2xz(p_{51}\,\epsilon_x+p_{52}\,\epsilon_y+p_{53}\,\epsilon_z+p_{54}\,\epsilon_{yz}+p_{55}\,\epsilon_{xz}+p_{56}\,\epsilon_{xy})\\
&+2xy(p_{61}\,\epsilon_x+p_{62}\,\epsilon_y+p_{63}\,\epsilon_z+p_{64}\,\epsilon_{yz}+p_{65}\,\epsilon_{xz}+p_{66}\,\epsilon_{xy})=1
\end{aligned}, \qquad (3.15.2)$$

where p_{11}-p_{66} are the photoelastic constants. In an optically isotropic material the matrix of photoelastic constants simplifies to [46]

$$\begin{array}{cccccc}
p_{11} & p_{12} & p_{12} & 0 & 0 & 0\\
p_{12} & p_{11} & p_{12} & 0 & 0 & 0\\
p_{12} & p_{12} & p_{11} & 0 & 0 & 0\\
0 & 0 & 0 & (p_{11}-p_{12})/2 & 0 & 0\\
0 & 0 & 0 & 0 & (p_{11}-p_{12})/2 & 0\\
0 & 0 & 0 & 0 & 0 & (p_{11}-p_{12})/2
\end{array}. \qquad (3.15.3)$$

Ignoring the small shear stress present in the vicinity of the channel waveguide on the diaphragm, the index ellipsoid becomes

$$\begin{aligned} & x^2\left(\frac{1}{n^2}+p_{11}\,\epsilon_x+p_{12}\,\epsilon_y+p_{12}\,\epsilon_z\right) \\ & +y^2\left(\frac{1}{n^2}+p_{12}\,\epsilon_x+p_{11}\,\epsilon_y+p_{12}\,\epsilon_z\right) \\ & +z^2\left(\frac{1}{n^2}+p_{12}\,\epsilon_x+p_{12}\,\epsilon_y+p_{11}\,\epsilon_z\right)=1 \end{aligned} \quad . \tag{3.15.4}$$

In the diaphragm coordinate system, the TE wave is polarized principally in x and will couple only into the x eigenpolarization. This polarization sees a refractive index equal to the location on the x axis of Fig. 3-32 intersected by the index ellipsoid,

$$x = n_{TE} = \left(\frac{1}{n^2}+p_{11}\,\epsilon_x+p_{12}\,\epsilon_y+p_{12}\,\epsilon_z\right)^{-1/2} \quad . \tag{3.15.5}$$

Since strain's contribution to the refractive index is small, (3.15.5) can be linearized using

$$n_{TE} = \left(\frac{1}{n^2}+\mathrm{f}(\epsilon)\right)^{-1/2} = n_{TE}\Big|_{\epsilon=0} + \frac{\partial n_{TE}}{\partial \epsilon}\,\epsilon \tag{3.15.6}$$

becoming

$$n_{TE} = n - \frac{1}{2}\left[\frac{1}{n^2}+\mathrm{f}(\epsilon)\right]^{-3/2}\frac{\partial \mathrm{f}(\epsilon)}{\partial \epsilon}\,\epsilon \quad . \tag{3.15.7}$$

Since $\frac{1}{n^2} >> \mathrm{f}(\epsilon)$, (3.15.7) reduces to

$$n_{TE} = n - \frac{n^3}{2}(p_{11}\,\epsilon_x+p_{12}\,\epsilon_y+p_{12}\,\epsilon_z) \quad . \tag{3.15.8}$$

In a similar manner the TM mode polarized in the diaphragm coordinate system principally perpendicular to the substrate (in z) will see a refractive index change given by

$$n_{TM} = n - \frac{n^3}{2}(p_{12}\,\epsilon_x+p_{12}\,\epsilon_y+p_{11}\,\epsilon_z) \quad . \tag{3.15.9}$$

At the x edge of the diaphragm ϵ_y is 0.

To calculate strain's effect on the waveguide path length or phase, the refractive index of each of the layers making up the waveguide is determined using (3.15.8) or (3.15.9). A good approximation is to assume the value of strain calculated in the middle of the core is constant throughout and the same in each of the layers except the air where

it is zero. This will underestimate the strain in the upper cladding and over estimate it in the lower cladding, errors which will mostly cancel if these two layers have similar refractive indices and photoelastic constants. (A more exact treatment integrates the change in dielectric permittivity with the light intensity in the channel [47]). The channel effective index (n_{et}) is calculated using both the strained and unstrained values of film refractive index. The phase change

$$\Delta\phi = -\int \frac{2\pi}{\lambda_0}\left[n_{et_{strained}}(\ell) - n_{et_{unstrained}}\right] d\ell \tag{3.15.10}$$

is then integrated along the channel path over the diaphragm. This definition of phase is in harmony with that commonly accepted in electrical engineering, namely that an increase in delay constitutes a negative phase. If the waveguide traverses the diaphragm over its long edge and the short end's influence is neglected because, for example, the beam model was used to determine strain, (3.15.10) reduces to

$$\Delta\phi = -\frac{2\pi}{\lambda_0}\left[n_{et_{strained}} - n_{et_{unstrained}}\right] L \quad , \tag{3.15.11}$$

where L is the length of the long diaphragm edge.

16 NUMERICAL CALCULATION OF PRESSURE AND TEMPERATURE SENSITIVITY

Numerical calculation of the phase change in the waveguide running across the long edge and center of the diaphragm (Fig. 3-33) due to pressure applied to the back of the diaphragm will be performed. The phase change due to a temperature change will also be determined. The structure used in these calculations, shown in Fig. 3-34, was built and tested, and those details and a comparison of the calculated response and that measured is given in chapter 4.

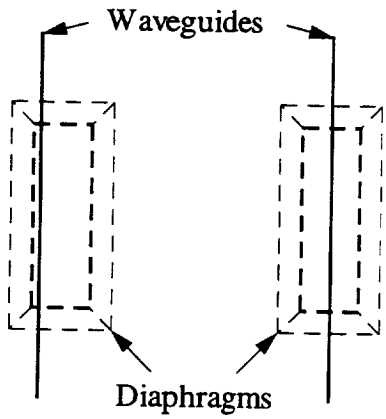


Fig. 3-33 Waveguide routed along the diaphragm's long edge (a), and over its center (b).

Photoelastic and elastic constants of the low pressure chemical vapor deposited SiON waveguide film were not available. Therefore in the following calculations, the elastic and photoelastic constants of SiON will be assumed to be the same as those of SiO_2.

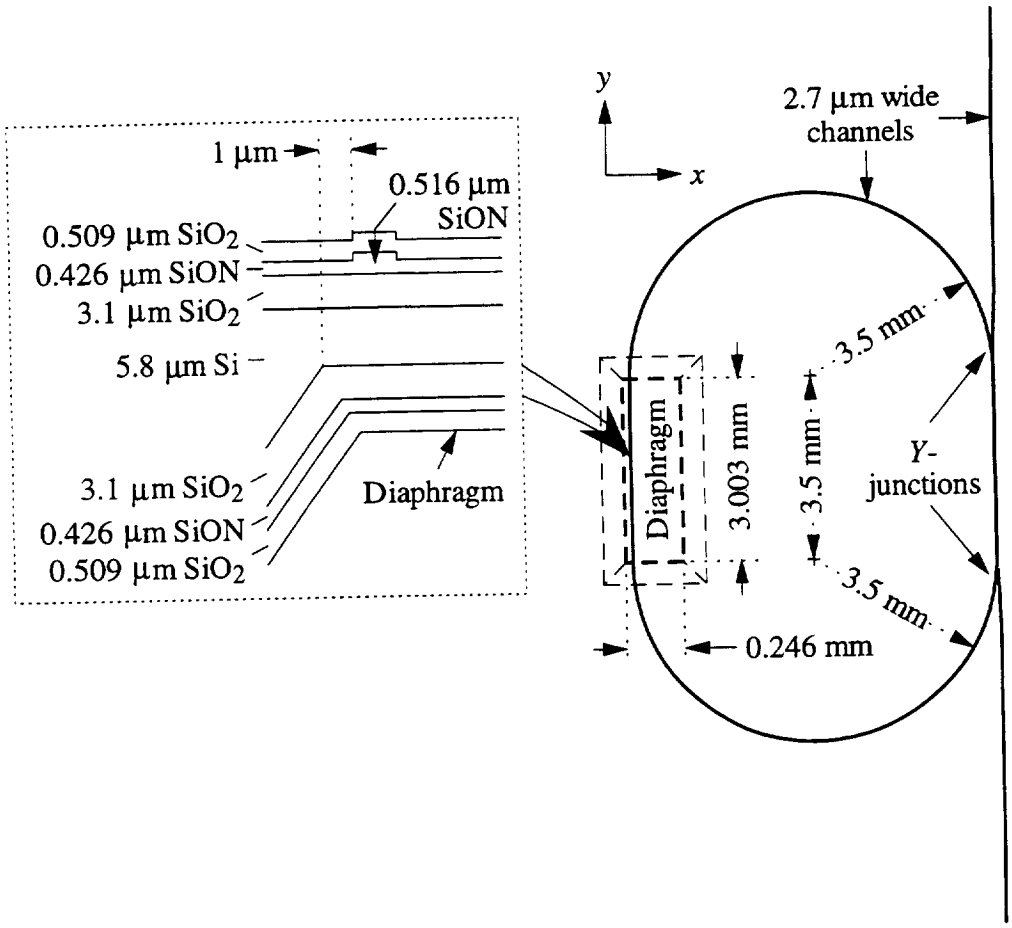


Fig. 3-34 Diaphragm and waveguide structure used for the calculation. The coordinate axis apply to the top view.

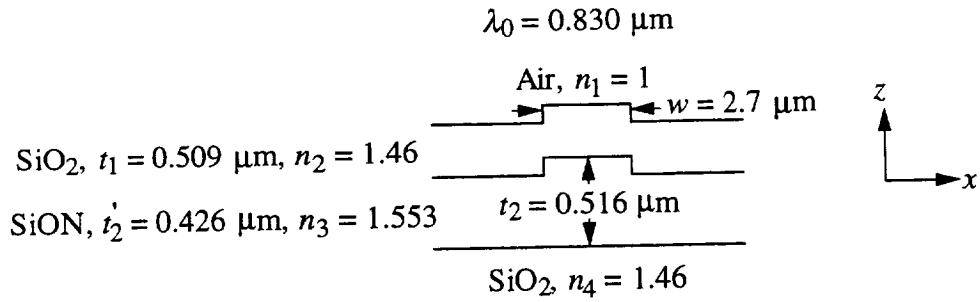


Fig. 3-35 Definition of waveguide variables. Light of 0.830 μm wavelength is available from laser diodes.

The customary coordinate systems for plate and waveguide analysis given in the literature and in this thesis will disagree when applied together to the structure of

Fig. 3-34. The z axis is customarily taken to be perpendicular to the plate, but in waveguide discussions, is parallel to the direction of propagation. This section will use the plate coordinate system in Fig. 3-34. Table 3-2 gives relevant constants to be used in the calculation.

Material	Refractive index ($\lambda = 0.63$ to 0.83 μm)	Thermal change in refractive index ($°C^{-1}$) [48]	Thermal Expansion (ppm/°C) [10]	Photo-elastic coefficients [49]		Poisson ratio [49]	Young's modulus (Pa) [10]	Yield strength (Pa) [10]
				p_{11}	p_{12}			
Si	3.85	1.5×10^{-4}	2.33	—	—	—	1.9×10^{11}	7×10^{9}
SiO_2	1.46	1.28×10^{-5}	0.55	0.121	0.270	0.17	7.3×10^{10}	8.4×10^{9}
Si_3N_4	2.02	?	0.8	?	?	?	3.85×10^{11}	1.4×10^{10}

Table 3-2 Material constants. — indicates that the constant is unimportant to this work and ? means that it could not be found. Reference applies to whole column.

Pressure sensitivity

Pressure applied to the diaphragm back creates stress in the waveguide films which alters, through the photoelastic effect, their refractive index. Stress also slightly changes the waveguide's dimensions. Both of these effects alter the channel effective index, n_{et}. If the waveguide is routed over the diaphragm center bulging stretches it. A quantitative determination of these effects will be performed using the effective index technique, described in section 6 of this chapter. Although the waveguide is actually a five layer structure, it will be assumed that the SiO_2 isolation layer is sufficiently thick that the silicon substrate can be neglected in the waveguide analysis, but not in the diaphragm analysis. The waveguide then reduces to a four layer structure, treated in section 4 of this chapter. The analysis will proceed by finding the difference between n_{et} in the strain-altered waveguide and that in the unaltered one and applying (3.15.11). The quantity n_{et} is found by determining n_{e2} and n_{e1} by solving iteratively (3.4.2) or (3.4.4)

and then iteratively solving (3.3.24) or (3.3.18) for n_{et}.

As shown in Chapter 2, the beam model accurately predicts deflection only away from the short edges, while the plate theory did not include the diaphragm's differing Young's modulus layers. Both can be included by modeling the plate as a beam composed of different layers and possessing width slightly less than the plate's long dimension.

It will be found that the strain ϵ_x dominates the pressure-induced phase change and that this phase change is largest when the waveguide is routed over the long edge of the diaphragm. In chapter 2, the deflection, stress, and strain in a 3.003 mm × 0.246 mm × 13.87 μm diaphragm made from SiO_2 was tabulated. This calculation was for the same diaphragm in the device depicted in Fig. 3-34, but the plate was assumed to be made from SiO_2, the layer deemed to most influence its stiffness. Referring back to Table 2-5, the strain at the long edge of the diaphragm moving in 75.075 μm increments in y from the corner takes on the values 0 $\mu\epsilon$, -402 $\mu\epsilon$, -682 $\mu\epsilon$ and then becomes, although possessing some slight oscillation, -774 $\mu\epsilon$, where 1 $\mu\epsilon$ means the strain is 10^{-6}. Assuming a linear change in strain in between these points, no difference would be felt by the waveguide, linearly influenced by strain, if the diaphragm possessed –774 $\mu\epsilon$ over a 2839 μm distance and 0 $\mu\epsilon$ for each 82 μm end piece (Fig. 3-36). Indeed the diaphragm aspect ration was chosen to provide a long sustained region of constant strain. Accordingly, the beam, whose strain is constant along its clamped edge, is made 2839 μm wide, thus taking into account the roll off in strain in the vicinity of the diaphragm's short clamped edge.

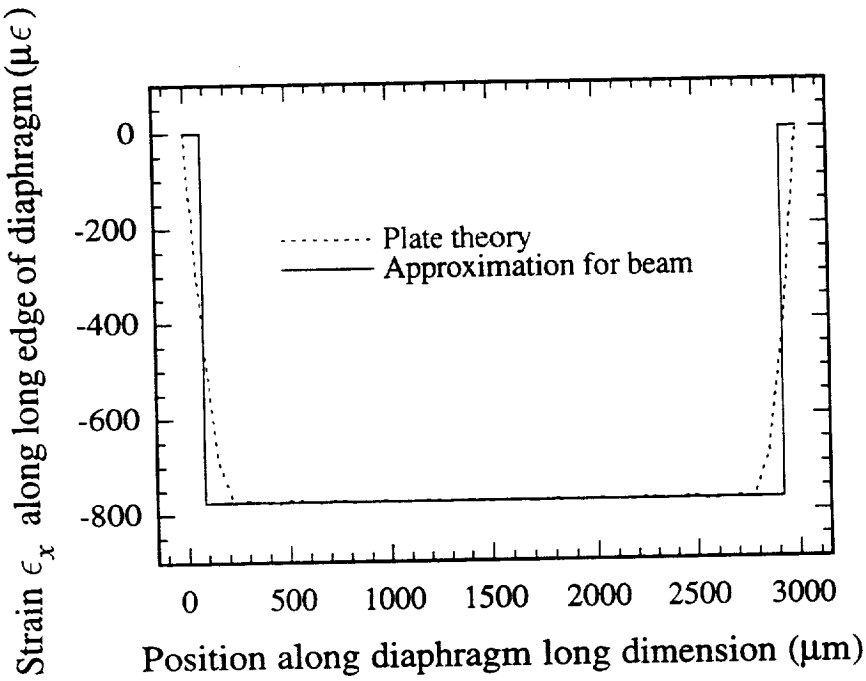


Fig. 3-36 Strain ϵ_x along the diaphragm edge as tabulated in Table 2-5 and the beam-modeled strain.

To take into account the silicon layer in the diaphragm an equivalent beam is constructed in which the stiffer silicon layer is replaced by an SiO_2 layer of the same thickness but having width scaled by E_{Si}/E_{SiO_2} (see Fig. 2-8). Such a beam, shown in Fig. 3-37b, possesses the same strain and deflection as the one shown in Fig. 3-37a. The wider part accounts for silicon's higher Young's modulus but does not increase the surface area exposed to pressure.

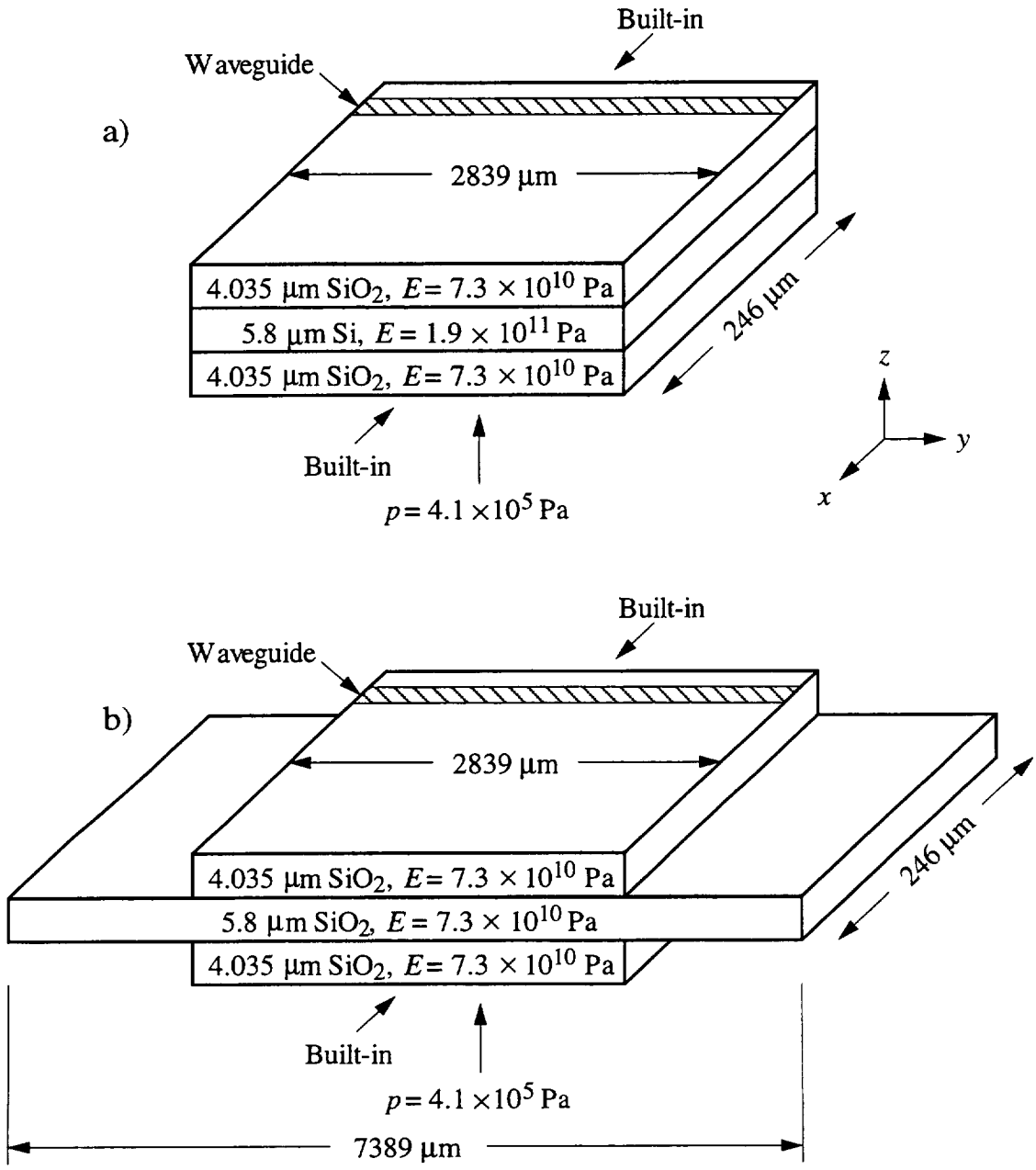


Fig. 3-37 a) Layered beam model for the plate shown in Fig. 3-34. b) Beam model having same deflection and strain but made from one material. This beam is more easily analyzed.

Continuing with the calculation, the moment of inertia I is next determined for the beam of Fig. 3-37b using (2.2.17). The intergal is written

$$I = \int_{y=0}^{y=2839\ \mu m} \int_{z=-6.935\ \mu m}^{z=-2.9\ \mu m} z^2\, dz\, dy + \int_{y=0}^{y=7389\ \mu m} \int_{z=-2.9\ \mu m}^{z=2.9\ \mu m} z^2\, dz\, dy + \int_{y=0}^{y=2839\ \mu m} \int_{z=2.9\ \mu m}^{z=6.935\ \mu m} z^2\, dz\, dy \quad , \qquad (3.16.1)$$

giving $I = 7.05 \times 10^5$ mm^4 or $I = 7.05 \times 10^{-19}$ m^4. It is necessary when applying this intergal that the z origin be in the beam's midplane so that $\int \sigma_x\, dz = 0$.

All strains are zero at the diaphragm midplane ($z = 0$) and linearly increase toward either surface. Fig. 3-34 shows that the center of the waveguide core is $\frac{5.8}{2} + 3.1 + \frac{0.516}{2} = 6.258$ μm from the midplane, and it is here the strains will be determined. The strains ϵ_x and ϵ_z are found using (2.2.34) and (2.2.35) with $b = 2839 \times 10^{-6}$ m, $p = 4.1 \times 10^5$ Pa, $L = 246 \times 10^{-6}$ m, $\nu = 0.17$, $E = 7.3 \times 10^{10}$ Pa, and $x = 0$ or L for the waveguide over the 3003 μm long edge or $x = L/2$ for the waveguide over the center of the diaphragm and running parallel to the 3003 μm long edge. They are, over the edge, $\epsilon_{x,\,edge} = -0.000693$, and $\epsilon_{z,\,edge} = 0.000142$ and over the center, $\epsilon_{x,\,center} = 0.000347$ and $\epsilon_{z,\,center} = -0.000071$. The strains over the center are –1/2 of those at the edge because the curvature at the center is –1/2 of that at the edge.

These strains alter the waveguide dimensions, and through the photoelastic effect, the film refractive indices. Equations (3.15.8) and (3.15.9), written in the diaphragm coordinate system, are used to find the change in film refractive index (not waveguide effective index) produced by this strain, and the definition strain = $\frac{\text{change in length}}{\text{length}}$ used to find the altered waveguide dimensions.

Table 3-3 gives the formula and parameter changes in the waveguide of Fig. 3-34 with 4.1×10^5 Pa of pressure applied to the diaphragm back. There are four components to the pressure induced change in waveguide path length:

- 1) Strain photoelastically alters the refractive index of the films.
- 2) The strain ϵ_x alters the width of the channel.
- 3) The strain ϵ_z change the layer thickness.

- 4) Away from the edge, the film bulges out and stretches the channel, lengthening it.

If the waveguide is routed along the edge of the diaphragm, the last effect will be zero. Since the changes in index and thickness are small, the effect of imposing all of them simultaneously will be the same as imposing them separately and summing the change each produces in n_{et}. The latter approach shows the relative importance of each of the contributions.

Parameter	Formula for change in parameter	Calculated change in parameter	Change this produces in quasi-TE n_{et}	Change this produces in quasi-TM n_{et}
n_2	$\frac{-n_2^3}{2}\left(\epsilon_{x,\,edge}\,p_{11} + \epsilon_{z,\,edge}\,p_{12}\right)$	7.08×10^{-5}	9.14×10^{-6}	NA
n_3	$\frac{-n_3^3}{2}\left(\epsilon_{x,\,edge}\,p_{11} + \epsilon_{z,\,edge}\,p_{12}\right)$	8.52×10^{-5}	6.12×10^{-5}	NA
n_4	$\frac{-n_4^3}{2}\left(\epsilon_{x,\,edge}\,p_{11} + \epsilon_{z,\,edge}\,p_{12}\right)$	7.08×10^{-5}	1.19×10^{-5}	NA
n_2	$\frac{-n_2^3}{2}\left(\epsilon_{x,\,edge}\,p_{12} + \epsilon_{z,\,edge}\,p_{11}\right)$	2.64×10^{-4}	NA	3.90×10^{-5}
n_3	$\frac{-n_3^3}{2}\left(\epsilon_{x,\,edge}\,p_{12} + \epsilon_{z,\,edge}\,p_{11}\right)$	3.18×10^{-4}	NA	2.10×10^{-4}
n_4	$\frac{-n_4^3}{2}\left(\epsilon_{x,\,edge}\,p_{12} + \epsilon_{z,\,edge}\,p_{11}\right)$	2.64×10^{-4}	NA	5.57×10^{-5}
w	$w\ \epsilon_{x,\,edge}$	-1.87×10^{-3} μm	-2.24×10^{-6}	-2.31×10^{-6}
t_1	$t_1\ \epsilon_{z,\,edge}$	7.23×10^{-5} μm	3.83×10^{-7}	6.15×10^{-7}
t_2	$t_2\ \epsilon_{z,\,edge}$	7.33×10^{-5} μm	5.19×10^{-6}	5.59×10^{-6}
t_2'	$t_2'\ \epsilon_{z,\,edge}$	6.05×10^{-5} μm	9.21×10^{-7}	8.96×10^{-7}
Total			8.65×10^{-5}	3.09×10^{-4}

Table 3-3 Change in film refractive index, layer thickness, and channel width produced by 4.1×10^5 Pa applied to the back of the diaphragm of Fig. 3-34 along with these changes affect on channel effective index, n_{et}. The waveguide is routed across the diaphragm's long edge.

Along the center line parallel to the long edge, the strains ϵ_x and ϵ_z are −1/2 of those over the long edge, so the change in the quasi-TE and quasi-TM modes of a channel routed over this path will be −1/2 of those calculated in Table 3-3. This neglects the strain ϵ_y, but it contributes little because it decays rapidly away from the short edge (Table 2-6). The change in the path length of the waveguide due to bulging is calculated using the data of Table 2-2. Although this neglects the silicon layer, little error will be introduced because the bulging effect contributes minimally to the total phase change of light in the waveguide. In the curved areas, the difference is found between the distance connecting each of the points and the undeflected diaphragm length:

$$\Delta\ell = 2\left[\sqrt{0.105^2 + 75.075^2} + \sqrt{(0.196-0.105)^2 + 75.075^2} + \sqrt{(0.228-0.196)^2 + 75.075^2} + \sqrt{(0.234-0.228)^2 + 75.075^2}\right] - 8(75.075) \tag{3.16.2}$$

$\Delta\ell = 0.00027\ \mu\text{m}$.

From the change in effective index the phase change is calculated using (3.15.11), while bulging, which increases the length of the channel changes the phase according to

$$\Delta\phi = -\frac{2\pi}{\lambda_0} n_{et} \Delta\ell \quad . \tag{3.16.3}$$

In (3.15.11), L = 2839 μm, the length of the material under constant beam modeled stress.

Table 3-4 summarizes the calculations presented in Table 3-3, interpreted using (3.15.11) in terms of phase change. The most sensitive device results from routing the channel over the long edge of the diaphragm. In Table 3-4, the bulk of the effect comes from the photoelastic effect. The calculated TM/TE phase change ratio is 3.5 owing to the way the photoelastic constants combine with the stress. For the quasi-TE mode, the product of a small (p_{11}) and a large negative (ϵ_x) number is added to the product of a large (p_{12}) and small positive (ϵ_z) one, yielding nearly zero. For the quasi-TM mode, the product of two large numbers $\epsilon_x p_{11}$, one of which is negative, is added to the product of two small numbers $\epsilon_z p_{12}$, yielding a larger number than for the TE case.

	Channel over long edge		Channel over middle	
	quasi-TE $\Delta\phi$ (rad)	quasi-TM $\Delta\phi$ (rad)	quasi-TE $\Delta\phi$ (rad)	quasi-TM $\Delta\phi$ (rad)
photoelastic	–1.77	–6.55	0.88	3.27
change in waveguide width	0.05	0.05	–0.02	–0.02
change in layer thickness	–0.14	–0.15	0.07	0.08
bulging	0	0	–0.003	–0.003
TOTAL	–1.86	–6.65	0.93	3.33

Table 3-4 Calculated change in ϕ due to a pressure of 4.1×10^5 Pa (60 psi) applied to the underside of the structure in Fig. 3-34 for two waveguide routings. A negative phase means that the application of pressure has increased the delay experienced by the light in propagating through the channel.

Temperature sensitivity

There are three different mechanisms by which temperature affects the effective refractive index of the channel:

- 1) The dimensions of the layers and channel width change with temperature.
- 2) Because the coefficient of thermal expansion (CTE) of the layers does not match that of the substrate, stress in the layers will vary with temperature. Through the photoelastic effect, this stress will perturb the film refractive index.
- 3) The layers's refractive indices will vary with temperature, apart from the photoelastic effect.

Because the films are thin compared to the substrate, their expansion is expected to be constrained to that of the silicon substrate, even for the diaphragm, which is built-in on four sides. Therefore the above three effects are expected to be the same in a waveguide on the diaphragm as off the diaphragm.

The temperature sensitivity of the channel structure depicted in Fig. 3-35 will be found in this analysis. Since the relevant thermal material constants for our formulation of SiON are not available, they will be taken to be the same as those of SiO_2. Because the films are much thinner than the substrate, changes in their dimensions parallel to the

substrate face will be the same as those in the substrate having no films: $\epsilon_{x,\,thermal} = \epsilon_{y,\,thermal} = 2.33 \times 10^{-6}/°C$, where the x and y directions are parallel to the wafer face. However, removed from the substrate, the oxide/oxynitride films would experience strain $\epsilon_{x,\,thermal} = \epsilon_{y,\,thermal} = \epsilon_{z,\,thermal} = 0.55 \times 10^{-6}/°C$. Therefore, the stress-inducing strain parallel to the substrate face, which photoelastically alters the refractive index is $\epsilon_{x,\,photoelastic} = \epsilon_{y,\,photoelastic} = (2.33 - 0.55) \times 10^{-6}/°C$. This substrate stretching or compressing of the film induces, through Poisson coupling, an additional strain in z :

$\epsilon_{z,\,photoelastic} = \frac{-\nu}{1-\nu}(\epsilon_{x,\,photoelastic} + \epsilon_{y,\,photoelastic}) = -0.73 \times 10^{-6}/°C$. The total dimensional change in the z direction is the sum of the thermal and Poisson coupled strains: $\epsilon_{z,\,total} = -0.18 \times 10^{-6}/°C$.

Using the effective index method, the change in n_{et} (channel effective index) with changes in layer thickness and refractive index were calculated for both quasi-TE and quasi-TM modes for a 1°C temperature rise. Since the changes in thickness and index are small, the effect of imposing all of them simultaneously will be the same as imposing them separately and summing the change each produces in n_{et}.

Parameter	Formula for change in parameter	Calculated change in parameter	Change this produces in quasi-TE n_{et}	Change this produces in quasi-TM n_{et}
w	$w\ \epsilon_{x,\,thermal}$	6.29×10^{-6} μm	7.5×10^{-9}	7.7×10^{-9}
t_1	$t_1\ \epsilon_{z,\,total}$	-9.2×10^{-8} μm	-4.9×10^{-10}	-7.8×10^{-10}
t_2	$t_2\ \epsilon_{z,\,total}$	-9.3×10^{-8} μm	-6.6×10^{-9}	-7.1×10^{-9}
t_2'	$t_2'\ \epsilon_{z,\,total}$	-7.7×10^{-8} μm	-1.2×10^{-9}	-1.1×10^{-9}
n_2 (photoelastic on quasi-TE)	$\frac{-n_2^3}{2}(\epsilon_{x,\,photoelastic}\,p_{11} + \epsilon_{y,\,photoelastic}\,p_{12} + \epsilon_{z,\,photoelastic}\,p_{12})$	-7.76×10^{-7}	-1.00×10^{-7}	NA
n_3 (photoelastic on quasi-TE)	$\frac{-n_3^3}{2}(\epsilon_{x,\,photoelastic}\,p_{11} + \epsilon_{y,\,photoelastic}\,p_{12} + \epsilon_{z,\,photoelastic}\,p_{12})$	-9.34×10^{-7}	-6.70×10^{-7}	NA
n_4 (photoelastic on quasi-TE)	$\frac{-n_4^3}{2}(\epsilon_{x,\,photoelastic}\,p_{11} + \epsilon_{y,\,photoelastic}\,p_{12} + \epsilon_{z,\,photoelastic}\,p_{12})$	-7.76×10^{-7}	-1.31×10^{-7}	NA
n_2 (photoelastic on quasi-TM)	$\frac{-n_2^3}{2}(\epsilon_{x,\,photoelastic}\,p_{12} + \epsilon_{y,\,photoelastic}\,p_{12} + \epsilon_{z,\,photoelastic}\,p_{11})$	-1.36×10^{-6}	NA	-2.01×10^{-7}
n_3 (photoelastic on quasi-TM)	$\frac{-n_3^3}{2}(\epsilon_{x,\,photoelastic}\,p_{12} + \epsilon_{y,\,photoelastic}\,p_{12} + \epsilon_{z,\,photoelastic}\,p_{11})$	-1.63×10^{-6}	NA	-1.08×10^{-6}
n_4 (photoelastic on quasi-TM)	$\frac{-n_4^3}{2}(\epsilon_{x,\,photoelastic}\,p_{12} + \epsilon_{y,\,photoelastic}\,p_{12} + \epsilon_{z,\,photoelastic}\,p_{11})$	-1.36×10^{-6}	NA	-2.87×10^{-7}
n_2 (thermal)	NA	1.28×10^{-5}	1.65×10^{-6}	1.89×10^{-6}
n_3 (thermal)	NA	1.28×10^{-5}	9.19×10^{-6}	8.45×10^{-6}
n_4 (thermal)	NA	1.28×10^{-5}	2.16×10^{-6}	2.70×10^{-6}
Total			1.21×10^{-5}	1.15×10^{-5}

Table 3-5 Change in layer thickness, refractive index and channel effective index for a 1°C temperature rise in the channel of Fig. 3-35.

Wafer expansion, apart from altering the film refractive index, also changes the optical path length by increasing the physical lengths of channels and their physical path length difference. The change in optical path length of a 1 mm long channel due to the total of the effects tabulated above and a change in the physical waveguide length is

$$\Delta\phi\,(\text{rad}) = -\frac{2\pi\Delta n_{et}\ell}{\lambda_0} - \frac{2\pi n_{et}\Delta\ell}{\lambda_0} = -\frac{2\pi\Delta n_{et}(1\text{ mm})}{0.83\times10^{-3}\text{ mm}} - \frac{2\pi n_{et}\dfrac{2.33\times10^{-6}\text{ mm}}{\text{mm}}\,1\text{ mm}}{0.83\times10^{-3}\text{mm}}, \tag{3.16.4}$$

where the value of Δn_{et} is totaled in Table 3-5.

	quasi-TE $\Delta\phi$ $\left(\frac{\text{rad}}{\text{°C-mm}}\right)$	quasi-TM $\Delta\phi$ $\left(\frac{\text{rad}}{\text{°C-mm}}\right)$
change in waveguide width	–0.0001	–0.0001
change in layer thickness	0.0001	0.0001
photoelastically induced through thermal mismatch in film-substrate expansions	0.0068	0.0119
thermal change in film index	–0.0984	–0.0987
elongation due to wafer expansion	–0.0265	–0.0264
TOTAL	–0.1181	–0.1132

Table 3-6 Calculated thermally induced change in ϕ over a 1 mm long waveguide due to a temperature change of 1°C. A negative phase means that the temperature increase has added to the delay experienced by the light in propagating through the channel.

Temperature sensitivity is attributed mostly to a change in material refractive index with temperature.

17 REFERENCES

[1] D. Marcuse, *Theory of Dielectric Optical Waveguides*, 2nd edition: Academic Press, Inc., pp. 330-334, 1991.

[2] M. J. Robertson, S. Ritchie, and P. Dayan, "Semiconductor Waveguides: Analysis of Optical Propagation in Single Rib Structures and Directional Couplers," *IEE*

Proceedings, vol. 132, Pt. J, pp. 336-342, 1985.

[3] D. L. Lee, *Electromagnetic Principles of Integrated Optics*: John Wiley & Sons, 1986.

[4] H. Kogelnik, "Theory of Optical Waveguides," in *Guided-Wave Optoelectronics*, vol. 26, *Springer Series in Electronics and Photonics*, T. Tamir, Ed.: Springer-Verlag, 1988, pp. 59-61.

[5] A. W. Snyder and J. D. Love, *Optical Waveguide Theory*: Chapman and Hall, p. 223, 1983.

[6] D. L. Lee, *Electromagnetic Principles of Integrated Optics*: John Wiley & Sons, p. 19, 1986.

[7] D. L. Lee, *Electromagnetic Principles of Integrated Optics*: John Wiley & Sons, p. 87, 1986.

[8] R. G. Hunsperger, *Integrated Optics: Theory and Technology*, vol. 33, 2nd edition: Springer-Verlag, pp. 32-33, 1985.

[9] A. Kumar, D. F. Clark, and B. Culshaw, "Explanation of Errors Inherent in the Effective-Index Method For Analyzing Rectangular-Core Waveguides," *Opt. Lett.*, vol. 13, pp. 1129-1131, 1988.

[10] K. E. Petersen, "Silicon as a Mechanical Material," *Proc. IEEE*, vol. 70, pp. 420-457, 1982.

[11] W. Gleine and J. Muller, "Laser Trimming of SiON Components for Integrated Optics," *J. Lightwave Technol.*, vol. 9, pp. 1626-1629, 1991.

[12] W. C. Borland, D. E. Zelmon, C. J. Radens, J. T. Boyd, and H. E. Jackson, "Properties of Four-Layer Planar Optical Waveguides Near Cutoff," *IEEE J. of Quantum Electron.*, vol. 23, pp. 1172-1179, 1987, and "Correction," *IEEE J. of Quantum Electron.*, vol. 26, pp. 960, 1990

[13] G. N. De Brabander, J. T. Boyd, and H. E. Jackson, "Single Polarization Optical Waveguide on Silicon," *IEEE J. of Quantum Electron.*, vol. 27, pp. 575-579, 1991.

[14] K. P. Koo, A. B. Tveten, and A. Dandridge, "Passive Stabilization Scheme for Fiber Interferometers Using (3 × 3) Fiber Directional Couplers," *Appl. Phys. Lett.*, vol. 41, pp. 616-618, 1982.

[15] S. K. Sheem, "Optical Fiber Interferometer with [3 × 3] Directional Couplers: Analysis," *J. Appl. Phys.*, vol. 52, pp. 3865-3872, 1981.

[16] D. L. Lee, *Electromagnetic Principles of Integrated Optics*: John Wiley & Sons, pp. 209-231, 1986.

[17] D. Marcuse, *Theory of Dielectric Optical Waveguides*, 2nd edition: Academic Press, Inc., pp. 263-264, 1991.

[18] D. Marcuse, *Theory of Dielectric Optical Waveguides*, 2nd edition: Academic Press, Inc., pp. 272-276, 1991.

[19] H. Nishihara, M. Haruna, and T. Suhara, *Optical Integrated Circuits*: McGraw-Hill, pp. 41-42, 1989.

[20] T. R. Ranganath and S. Wang, "Ti-Diffused $LiNbO_3$ Branched-Waveguide Modulators: Performance and Design," *IEEE J. of Quantum Electron.*, vol. 13, pp. 290-295, 1977.

[21] H. Sasaki and N. Mikoshiba, "Normalized Power Transmission in Single Mode Optical Branching Waveguides," *Electron. Lett.*, vol. 17, pp. 136-138, 1981.

[22] J. Y. Jeng and W. S. Wang, "Special Consideration on the Design of a *Y*-Junction Optical Waveguide," *Microwave Opt. Technol. Lett.*, vol. 5, pp. 592-597, 1992.

[23] W. K. Burns and A. F. Milton, "Mode Conversion in Planar-Dielectric Separating Waveguides," *IEEE J. of Quantum Electron.*, vol. 11, pp. 32-39, 1975.

[24] M. Izutsu, Y. Nakai, and T. Sueta, "Operation Mechanism of the Single-Mode Optical-Waveguide *Y* Junction," *Opt. Lett.*, vol. 7, pp. 136-138, 1982.

[25] R. Baets and P. E. Lagasse, "Calculation of Radiation Loss in Integrated-Optic Tapers and *Y*-Junctions," *Appl. Opt.*, vol. 21, pp. 1972-1978, 1982.

[26] H. Nishihara, M. Haruna, and T. Suhara, *Optical Integrated Circuits*: McGraw-

Hill, pp. 40-41, 1989.

[27] M. Kuznetsov, "Radiation Loss in Dielectric Waveguide *Y*-Branch Structures," *J. Lightwave Technol.*, vol. 3, pp. 674-677, 1985.

[28] R. G. Hunsperger, *Integrated Optics: Theory and Technology*, vol. 33, 2nd edition: Springer-Verlag, pp. 107-114, 1985.

[29] L. F. Stokes, M. Chodorow, and H. J. Shaw, "All-Single-Mode Fiber Resonator," *Opt. Lett.*, vol. 7, pp. 288-290, 1982.

[30] E.-G. Neumann, "Curved Dielectric Optical Waveguides With Reduced Transition Losses," *Proc. IEE*, vol. 129, Pt. H, pp. 278-280, 1982.

[31] D. L. Lee, *Electromagnetic Principles of Integrated Optics*: John Wiley & Sons, pp. 135-142, 1986.

[32] D. Marcuse, *Light Transmission Optics*. New York: Van Nostrand Reinhold Company, p. 404, 1972.

[33] K. Thyagarajan, M. R. Shenoy, and A. K. Ghatak, "Accurate Numerical Method for the Calculation of Bending Loss in Optical Waveguides Using a Matrix Approach," *Opt. Lett.*, vol. 12, pp. 296-298, 1987.

[34] J. S. Gu, P. A. Besse, and H. Melchior, "Novel Method for Analysis of Curved Optical Rib-Waveguides," *Electron. Lett.*, vol. 25, pp. 278-280, 1989.

[35] R. J. Deri, E. Kapon, and L. M. Schiavone, "Bend Losses in GaAs/AlGaAs Optical Waveguides," *Electron. Lett.*, vol. 23, pp. 845-847, 1987.

[36] E. C. M. Pennings, G. H. Manhoudt, and M. K. Smit, "Low-Loss Bends in Planar Optical Ridge Waveguides," *Electron. Lett.*, vol. 24, pp. 998-999, 1988.

[37] H. Van Brug, F. H. Groen, Y. S. Oei, J. W. Pedersen, and E. C. M. Pennings, "Low-Loss Straight and Curved Ridge Waveguides in LPE-Grown GaInAsP," *Electron. Lett.*, vol. 25, pp. 1330-1332, 1989.

[38] J. J. G. M. Van Der Tol, J. W. Verhoof, and M. B. J. Diemeer, "S-Bends Using Offsets in Fibre-Compatible K^+-Na^+ Ion-Exchanged Glass Waveguides," *Electron. Lett.*,

vol. 27, pp. 379-380, 1991.

[39] W. J. Minford, S. K. Korotky, and R. C. Alferness, "Low-Loss Ti:$LiNbO_3$ Waveguide Bends at λ = 1.3 μm," *IEEE J. Quantum Electron.*, vol. 18, pp. 1802-1806, 1982.

[40] K. Petermann, "Fundamental Mode Microbending Loss in Graded-Index and W Fibers," *Optical and Quantum Electronics*, vol. 9, pp. 167-175, 1977.

[41] W. A. Gambling, H. Matsumura, and C. M. Ragdale, "Field Deformation in a Curved Single-Mode Fiber," *Electron. Lett.*, vol. 14, pp. 130-132, 1978.

[42] G. Beheim, K. Fritsch, and R. N. Poorman, "Fiber-Linked Interferometric Pressure Sensor," *Rev. Sci. Instr.*, vol. 58, pp. 1655-1659, 1987.

[43] P. Sansonetti, M. Lequime, and H. Giovannini, "Interferometric Spectrally Encoded Sensor Using a Superradiant Diode," presented at Optical Fiber Sensors, Proceedings of the 6th International Conference, OFS'89, Paris, France, and published in Springer Proceedings in Physics, vol. 44, pp. 71-77, Springer-Verlag publishers, 1989.

[44] A. Yariv and P. Yeh, *Optical Waves in Crystals*: John Wiley & Sons, p. 86, 1983.

[45] A. Yariv and P. Yeh, *Optical Waves in Crystals*: John Wiley & Sons, p. 77, 1983.

[46] A. Yariv and P. Yeh, *Optical Waves in Crystals*: John Wiley & Sons, p. 318-328, 1983.

[47] M. Ohkawa, M. Izutsu, and T. Sueta, "Integrated Optic Pressure Sensor on Silicon Substrate," *Appl. Opt.*, vol. 28, pp. 5153-5157, 1989.

[48] *Optics Guide 5*: Melles Griot Corporation, p. 3-5, 1990.

[49] G. B. Hocker, "Fiber-Optic Sensing of Pressure and Temperature," *Appl. Opt.*, vol. 18, pp. 1445-1448, 1979.

IV Diaphragm and Waveguide Fabrication

1 INTRODUCTION

This chapter provides particulars of diaphragm and waveguide fabrication. The formation and characterization of the waveguide films, including a description of the low pressure chemical vapor deposition reactor, controlled etching for channel formation, propagation loss measurements, the method of preparing waveguide end faces by cleaving, a description of pitfalls in using ellipsometry, and the prism coupling setup, is given. The system for etching diaphragms and another for inferring their thickness from a measurement of their transmission to a Helium Neon laser is described. The chapter finishes up with a description of the two masks used to fabricate the devices and details the fabrication sequence.

2 DIAPHRAGM FABRICATION

Wafers having a <100> orientation and "SEMI standard flat" were used in this work. The SEMI standard states that the primary flat on a <100>, 2 inch diameter silicon wafer be along the {110} family of planes [1]. Fig. 4-1 illustrates such a wafer where the face and flat have been oriented parallel to the (100) and (011) planes. These are members of the <100> and {110} family of directions and planes. The {111} planes, on which anistropic etchants preferentially stop, make an angle with the surface of $\tan^{-1}(\sqrt{2}) = 54.74°$. Lines where this family of planes intersect the surface are either parallel or perpendicular to the {110} planes and therefore to the SEMI standard primary flat on <100> wafers. To prevent undercutting the masking oxide [2], an opening in it for anistropically etching a rectangular diaphragm should also have the opening's sides parallel or perpendicular to the SEMI standard flat, a situation which simplifies orienting the diaphragm mask on the wafer surface. A properly oriented opening in the masking

oxide of dimension $A \times B$ will yield a $\left(A - \frac{2d}{\sqrt{2}}\right) \times \left(B - \frac{2d}{\sqrt{2}}\right)$ diaphragm, where d is the well depth, not allowing for oxide undercut or <111> plane etching.

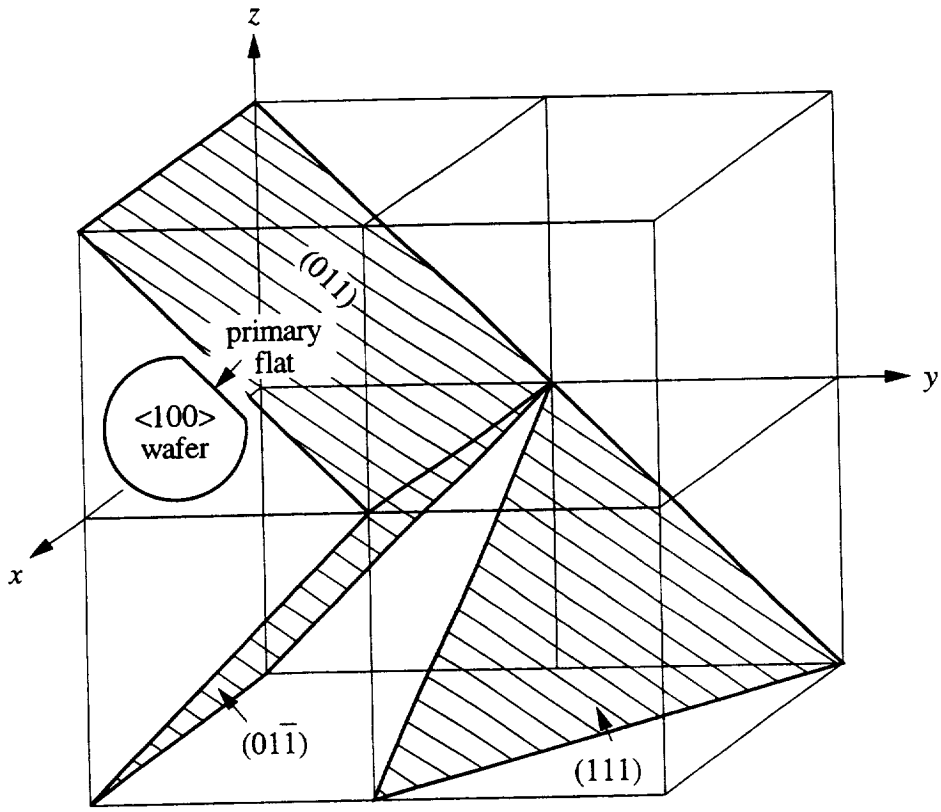


Fig. 4-1 Unit cells of silicon crystal and orientation of <100> face, (011) flat, and observed cleavage planes.

3 ETCHING SYSTEM

Although diaphragms can be fabricated using dry or wet etching, the control, minimal undercut, sidewall quality, equipment simplicity, and easy masking requirements of anistropic wet etchants makes them preferred, and one was chosen for this work. Diaphragms were fabricated by wet etching in straight, unstirred, PSE-200, a commercially prepared solution of isopropyl alcohol, KOH, and water purchased from Transene Co. [3]. The etching solution was maintained at 40°C by an etch system (Fig. 4-2) consisting of a hot plate (Cole-Parmer 04644) with the provision for closed loop control using an RTD temperature probe, here constructed from stainless steel and

enclosed by a Teflon sleeve closed at one end. The etchant in the system was contained in a 1000 ml Teflon beaker which rested directly on the hotplate surface. City water flowed through a Wollam condenser-beaker cap, also purchased from Transene, which condensed vapors evaporating from the solution, maintaining its concentration. The temperature probe passed through an opening in the center of the condenser-cap and its tip was immersed to a depth of 2.5 cm in the etchant. The wafers were held with their faces vertical by Teflon V-grooved plates which were tied to a polypropylene basket, and this was secured to an inverted beaker bottom. The V-grooved wafer holders could accommodate 5 wafers, holding them 0.5 cm apart.

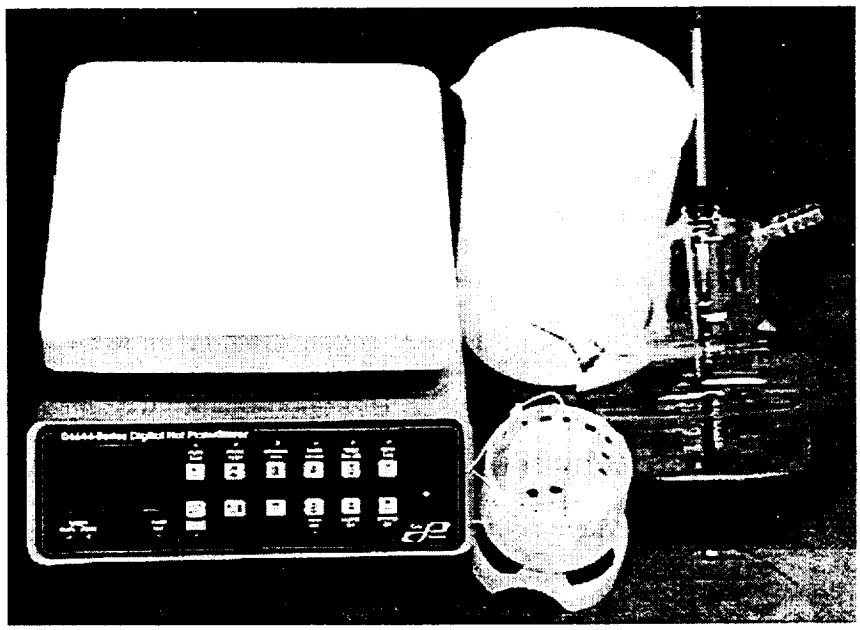


Fig. 4-2 Wet etching system for fabricating diaphragms. From left to right the parts are hotplate, Teflon beaker, basket with one wafer in it, and Wollam condenser/cap with temperature probe. A steam oxide mask and KOH etching, with no etch stop, was used to fabricate the diaphragms.

4 ETCHING CONSIDERATIONS

Two-inch, double side polished, <100> orientation, ~10 mil wafers having a <110> major flat were the starting substrates. Cleaves will be perpendicular or parallel to this flat, and rectangular openings in protective films on the wafer back for diaphragm

formation should be oriented with one of their sides parallel to the flat. The side of the wafer on which the waveguides will be fabricated should have an excellent surface finish. Using the conditions described before, the etch rate for <100> silicon was 4.39 μm/hr, while for 1050°C steam grown SiO_2, it was 9.50 nm/hr, giving a etch rate ratio $Si:SiO_2$ of 462:1. This large ratio, due in part to use of the etchant at such a low temperature, means that only 0.5 μm of masking oxide would be consumed for an etch from one side all the way through a 10 mil wafer. However a masking oxide of ~3 μm is recommended to provide scratch resistance. This requires a ~24 hr steam oxide growth (oxygen bubbled through boiling water) at 1050°C. A horizontal furnace tube at atmospheric pressure was used. The wafers were held in a quartz boat so that the tube axis was perpendicular to their face.

5 SIZING THE DIAPHRAGMS

Reproducible diaphragms are desired, but their length and width are affected by the opening in the oxide mask as well as etch depth, which varies with wafer thickness. In diaphragms where width is much less than length, strain at the surface along the long edge, due to an applied differential pressure, is proportional to the square of the width. Since sensor scale factor is proportional to the strain integrated along the waveguide path, accurate control of the diaphragm width (the shorter dimension) is more important than length (the longer dimension). To more accurately control the shorter dimension, the diaphragm mask was aligned to the major flat, positioned so the devices of interest would appear intact on the wafer (the masks contained more structures than could be printed on one 2 inch wafer), and contact-exposed. After development, the mask was realigned to the original pattern, and the wafer translated, using the aligner's calibrated differential micrometers, a distance parallel to the short edge which depended on wafer thickness and desired diaphragm short dimension. The wafer was then again contact exposed and developed. In the positive resist, the resulting rectangular openings will correspond to those exposed either time and will have the same long dimension as on the mask, but

their short dimension will increase beyond that on the mask by the amount the wafer was translated after realigning to the first print.

6 MEASURING DIAPHRAGM THICKNESS

The final diaphragm thickness for all diaphragms fabricated for this work was determined by the length of time the wafer remained in the etchant. Although anistropic etchants may be found which will stop on p^{++} boron layers, thus establishing diaphragm thickness, these layers can introduce stress and reduce the Young's modulus [4]. In addition, vendors of layers grown epitaxially have a high batch charge, making it difficult to try a variety of thicknesses. Of course for batch production having one thickness is highly desirable. For this work, however, a method was desired to accurately size diaphragm thickness to suit.

Diaphragm thickness was initially determined by subtracting wafer thickness, measured with a micrometer, from etched well depth, measured by noting the stage movement required to focus with an optical microscope alternatively on the well floor and substrate back. However, it was found that on the Nikon optical microscope, the stage movement produced by the fine focus adjust knob, from whose markings stage translation was inferred, only reproduced, over various segments of its travel, to within ± 1.5 μm over 200 μm of travel. The digital micrometer (Mitutoyo Digimatic) only resolved to one micron, giving an uncertainty of ± 2.5 μm in diaphragm thickness. From beam theory given in a previous chapter, strain along the long edge of the diaphragm varies inversely with the cube of diaphragm thickness, resulting in a potentially large difference in device sensitivity from that expected just due to uncertainties in diaphragm thickness.

To reduce these uncertainties, an optical method of measuring diaphragm thickness from its transmittance of a laser beam was developed, and the layout is given in Fig. 4-3. A plano-convex lens focused a 0.6328 μm laser beam to a waist measured to be $\omega_0 = 58$ μm, at which was located a pinhole aperture to reduce stray light scattered from

the lens. The definition for ω_0 comes from the assumed intensity profile $I = I_0 \exp\left(\frac{-2r^2}{\omega_0}\right)$, where r is radius. Using this definition, one can determine ω_0 from the distance a knife edge can be translated perpendicular to the beam to make the transmitted intensity vary between 16% and 84%: ω_0 was measured in this manner. The shortest dimension of the diaphragm was always at least four times greater than ω_0, permitting the full diameter of the beam to be transmitted. Afterward, a biased 5.05 mm diameter silicon photodiode (EG&G SGD-200) detected the transmitted beam. Transmittance through the diaphragm was taken to be

$$T_{meas} = \frac{I_{det}|\text{with sample in place} - I_{det}|\text{beam blocked}}{I_{det}|\text{with sample removed} - I_{det}|\text{beam blocked}} , \qquad (4.6.1)$$

where $I_{det}|\text{beam blocked}$ represents the detector dark current.

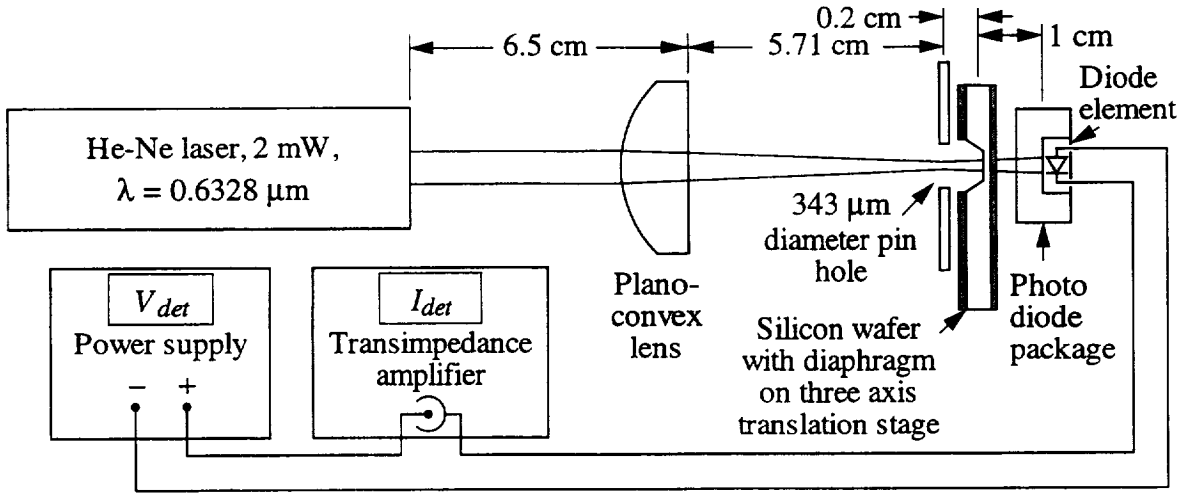


Fig. 4-3 Optical method for determining diaphragm thickness. The detector is placed close to the sample to minimize the scattered light not collected. The pinhole, although larger than the diaphragm, reduced the pickup of scattered light when the wafer was removed.

Fig. 4-4 shows the photodiode current as a function of beam radius for various reverse biases under full 2 mW HeNe fluence. No diaphragm was in the setup and the beam radius was altered by adjusting the distance between the detector and the beam

waist. For reverse biases < 40 V detector responsivity decreases as the spot size is reduced. Therefore, V_{det} = 40 V was selected for all diaphragm measurements.

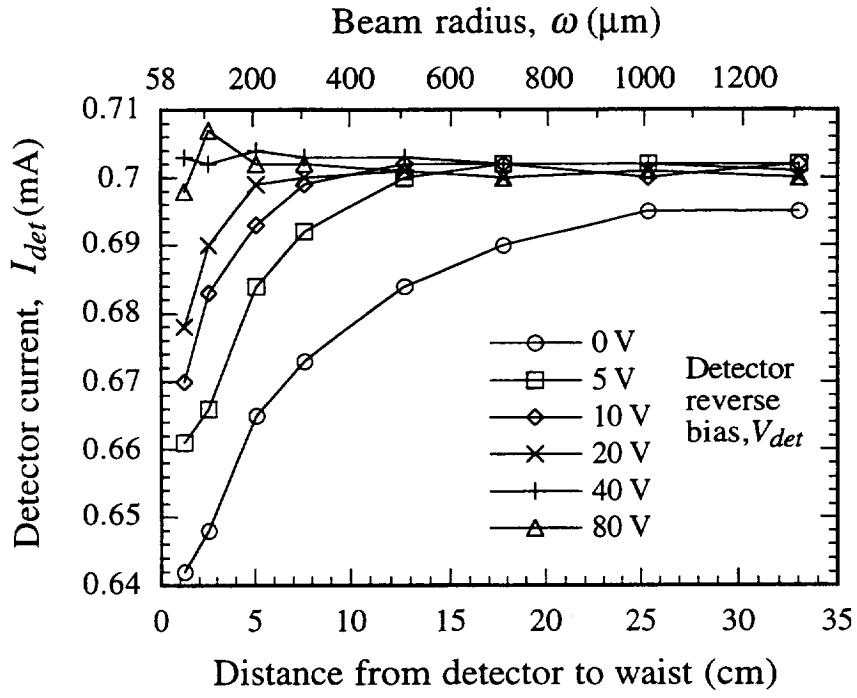


Fig. 4-4 Detector current for various spot sizes and reverse biases. In all cases, the beam underfilled the detector active area.

Because of the high attenuation of silicon for the wavelength used, multiple reflections between the interfaces (interference) of the silicon layer in the diaphragm can be ignored. In the masking oxide coating one side of the diaphragm, multiple beam interference between the air-SiO_2 and SiO_2-Si interfaces must be summed. This layer can act as an anti-reflection coating or absentee layer, depending on its thickness.

The transmittance T_{calc} of the diaphragm shown in Fig. 4-5, at normal incidence, is given by the product of the transmittances 1) into the silicon through the masking oxide, 2) through the silicon, and 3) through the silicon-air interface:

$$T_{calc} = T_{air\text{-}SiO_2\text{-}Si} \cdot T_{Si} \cdot T_{Si\text{-}air} \ , \qquad (4.6.2)$$

where

$$T_{air\text{-}SiO_2\text{-}Si} = 1 - \frac{r_{12}^2 + r_{23}^2 + 2r_{12}r_{23}\cos(2\phi)}{1 + r_{12}^2 r_{23}^2 + 2r_{12}r_{23}\cos(2\phi)} \quad . \tag{4.6.3}$$

Transmission is expected to be the same whether the laser is incident on the front or the back. The Fresnel coefficients for (4.3.3) are $r_{12} = \frac{n_{SiO_2} - n_{air}}{n_{SiO_2} + n_{air}}$, $r_{23} = \frac{n_{Si} - n_{SiO_2}}{n_{Si} + n_{SiO_2}}$, while the phase shift of light traveling through the oxide is $\phi = \frac{2\pi}{\lambda_0} n_{SiO_2} t_{SiO_2}$. The other quantities are

$$T_{Si} = \exp(-\alpha_{Si} t_{Si}) \tag{4.6.4}$$

and

$$\alpha_{Si} = \frac{4\pi k_{Si}}{\lambda_0} \quad . \tag{4.6.5}$$

To simplify calculations, the refractive index of silicon is assumed real for all but the calculation of T_{Si}. At our wavelength, $n_{air} = 1$, $n_{SiO_2} = 1.46$, and $n_{Si} = 3.85$. The quantity $T_{Si\text{-}air}$ can be calculated from (4.6.3) by setting $\phi = 0$ and is 0.6547. Transmittance measurements on a $10^{14}/cm^3$ boron doped, 13 μm thick silicon membrane formed using an epitaxial-etch stop have been used by workers at the National Institute of Standards and Technology to determine the value of extinction coefficient of silicon to be $k_{Si} = 0.01564 \pm 0.00031$ at $\lambda_0 = 0.633$ μm [5].

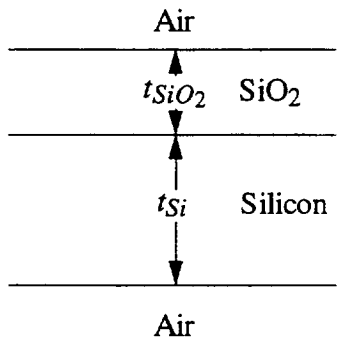


Fig. 4-5 Model for diaphragm.

For certain thicknesses, the SiO_2 film can function as an antireflection coating, permitting more light to pass through the diaphragm, or as an absentee layer, in which

case transmission through the diaphragm would be the same as if there were no oxide. Fig. 4-6 shows these two extremes. Although diaphragm thickness can be estimated to within about 1 μm without knowing the SiO_2 thickness, in this study the masking SiO_2 film was ellipsometrically measured.

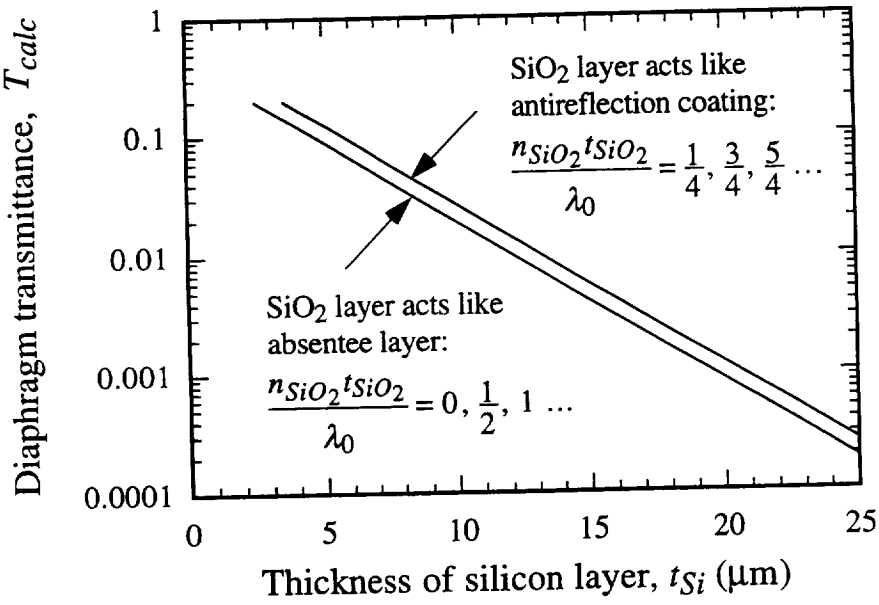


Fig. 4-6 Normal incidence transmission of HeNe laser through diaphragm coated on one face with SiO_2 and bare on the other. For a given silicon thickness, transmittance varies between two extremes depending on SiO_2 thickness.

Once oxide thickness has been determined, (4.6.3) is used to determine $T_{air\text{-}SiO_2\text{-}Si}$, and then the measured diaphragm transmission (4.6.1) is equated to T_{calc}. Then (4.6.2) is used to solve for T_{Si}, and (4.6.4) inverted to determine t_{Si}, the silicon layer thickness in the diaphragm.

Evaluation of diaphragm thickness measurement using laser

This method of measuring diaphragms was compared to the microscope focusing technique and to the definitive method consisting of cleaving through the diaphragm and measuring its thickness on edge with an optical microscope, as shown for one particular diaphragm in Fig. 4-7. A concern was whether roughness as shown on a typical

diaphragm back in Fig. 4-8, might make the method inaccurate by scattering or causing some of the beam to be total internally reflected. Indeed, it was noted that the photodiode needed to be located close to the diaphragm to intercept the beam, as its divergence was increased by diaphragm roughness.

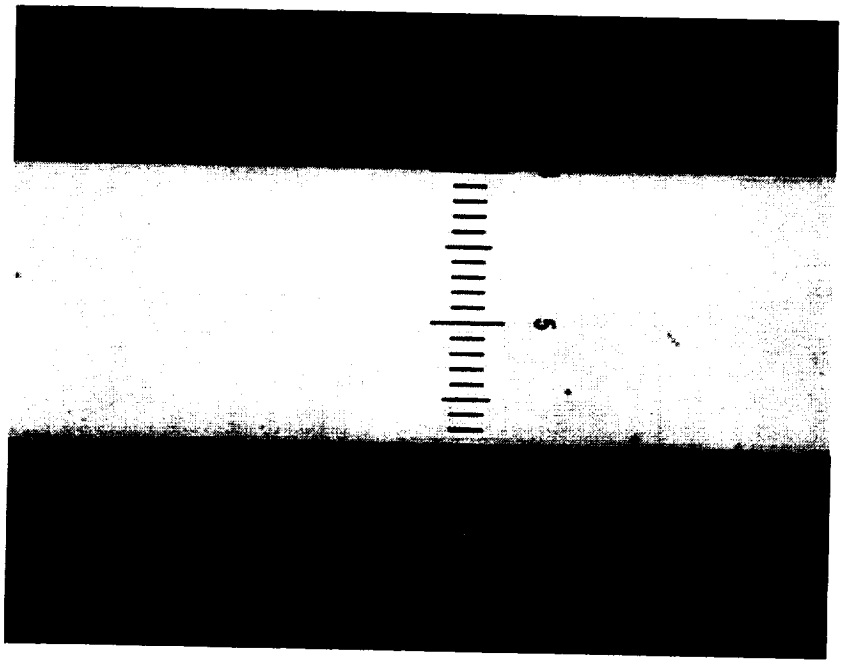


Fig. 4-7 A 17.4 μm thick silicon diaphragm (graticule is 0.96 μm/div) having a 2.1175 μm thick SiO_2 layer on one side. To avoid having the diaphragm appear thicker than actual, the film was exposed correctly for the silicon, and so the underexposed SiO_2 layer is not visible.

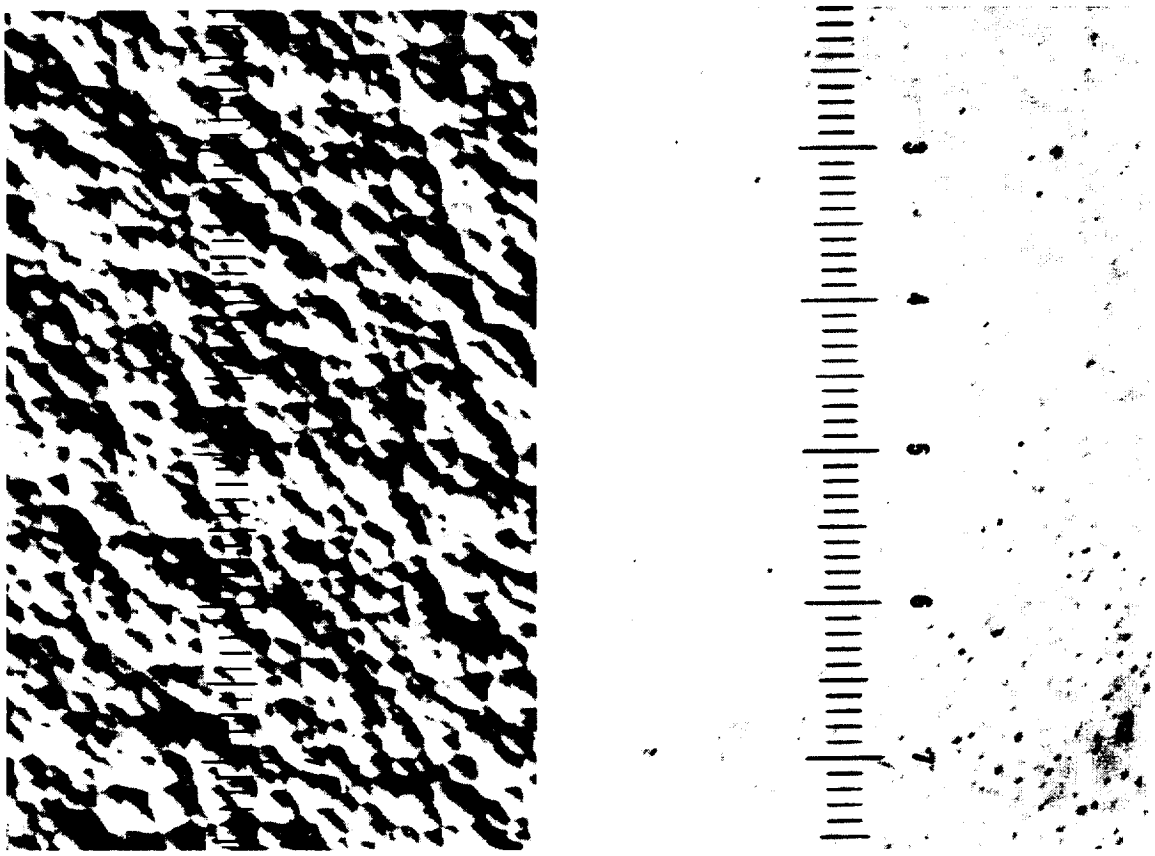


Fig. 4-8 Back of diaphragm produced by etching 246 μm into wafer. (a) Differential interference contrast image (Nomarski); (b) standard bright field image. Graticule is 4.75 μm/div.

A double-side polished, phosphorus doped, 50 Ω-cm, 262 μm average thickness, 2 inch diameter silicon wafer was oxidized, windows were opened in the masking oxide on one side, and the wafer was etched in PSE-200 until diaphragms 3000 μm × 275 μm × 18 μm remained. The diaphragms were separated from each other by cleaving the wafer into pieces and these were further etched as before for different times to create a variety of diaphragm thicknesses. The thickness of each piece was measured with a digital micrometer (Mitutoyo Digimatic), which resolved to 1 μm, and this reading was subtracted from the well depth measured by focusing alternately on the wafer back and diaphragm using an optical microscope (Nikon Optiphot). For the laser transmission measurement, the masking oxide on the unetched side was measured with an

ellipsometer. Finally the diaphragms were cleaved through and measured on edge with an optical microscope, as in Fig. 4-7.

Table 4-1 shows the results of this investigation. The wafer thickness measured using the micrometer includes the masking film of SiO_2 measured on both sides of the wafer. The silicon layer thickness component in the diaphragm measured after it was cleaved through agrees better with that measured by laser transmission than by microscope focusing. It matters little to the laser method whether the laser is incident on the silicon or silicon dioxide surfaces.

			I_{det} (μA) with beam incident on SiO_2 side			I_{det} (μA) with beam incident on Si side			Measured silicon layer thickness, t_{Si} (μm)			
Oxide thick (μm)	Wafer thick (μm)	Well depth by focusing (μm)	Through diaph-ragm	Beam blocked	Sample removed	Through diaph-ragm	Beam blocked	Sample removed	by focusing	by HeNe incident on SiO_2	by HeNe incident on Si	by cleaving
2.0769	265	256.6	137.3	4.6	705	136.3	4.6	703	4.2	3.63	3.64	3.7
2.0887	264	249.7	59.2	4.6	714	60.2	4.6	714	10.1	6.39	6.34	5.9
2.1017	259	247.2	25.5	4.4	714	26.1	4.5	712	7.6	9.27	9.18	8.6
2.0992	260	244.3	25.1	4.4	719	25.0	4.4	707	11.5	9.39	9.35	9.2
2.0858	265	247.7	12.46	4.55	713	12.35	4.39	719	13.1	12.6	12.7	12.4
2.1002	262	245.2	9.50	4.79	710	9.58	4.88	709	12.6	14.1	14.1	13.9
2.1145	263	239.4	8.11	4.16	724	8.24	4.21	718	19.4	14.5	14.4	14.2
2.1175	263	236.8	5.44	4.26	717	5.57	4.35	715	22.0	18.3	18.2	17.4

Table 4-1 Comparison of diaphragm thickness measured using three methods.

7 CLEAVING

To create edges suitable for waveguide incoupling, after all waveguide films were deposited, the wafer was cleaved using an inexpensive hand-held tungsten-carbide scriber (General Tool). The wafer was placed on a compliant surface, such as a magazine, and the scriber tip rested on the wafer face about 0.5 mm from the edge, with the handle making a 45° angle with respect to the table. Then, after applying about 100 g of force, the tip was drawn over the edge. Several attempts were usually necessary, but eventually either the wafer would suddenly cleave into two pieces, or a crack would initiate. In the latter case, the tip was brought down over the crack with the same handle angle as before

and a gradually increasing force applied until the cleave propagated across the wafer, separating the sample into two. It was found by a coworker that defects in the end face, present near the cleavage initiation site, tended to decay away more rapidly near the surface opposite the one that the scriber was pressed against. Therefore the point on the front edge where the cleave line was to begin was transferred to the wafer back edge with a pen. The cleaving procedure described above was then performed on the wafer back.

Fig. 4-9 shows a photograph of a good cleave on a <100> silicon wafer bearing a thermal oxide and waveguide films. In most cases, the cleaved edge was found to be approximately perpendicular to the face and parallel (along the (011 plane)) or perpendicular (along the $(01\bar{1})$ or $(0\bar{1}1)$ planes) to the primary flat. This is in contradiction to Sze who states that the {111} planes are the easiest planes of cleavage [6]. According to Fig. 4-1, a cleave along this family of planes would run either parallel or perpendicular to the flat, with the edge inclined by 35.3° from normal. Fig. 4-10 shows a poor cleave. Although the silicon appears smooth, the SiO_2 and waveguide film has chipped off in places. Edge fire coupling into the planar film at various locations along the edge by translating the sample perpendicular to the face would cause the angle the streak made with respect to the edge to rapidly vary. In addition, streak splitting, with the two beams propagating at different angles, and low coupling efficiency in some places would be noted.

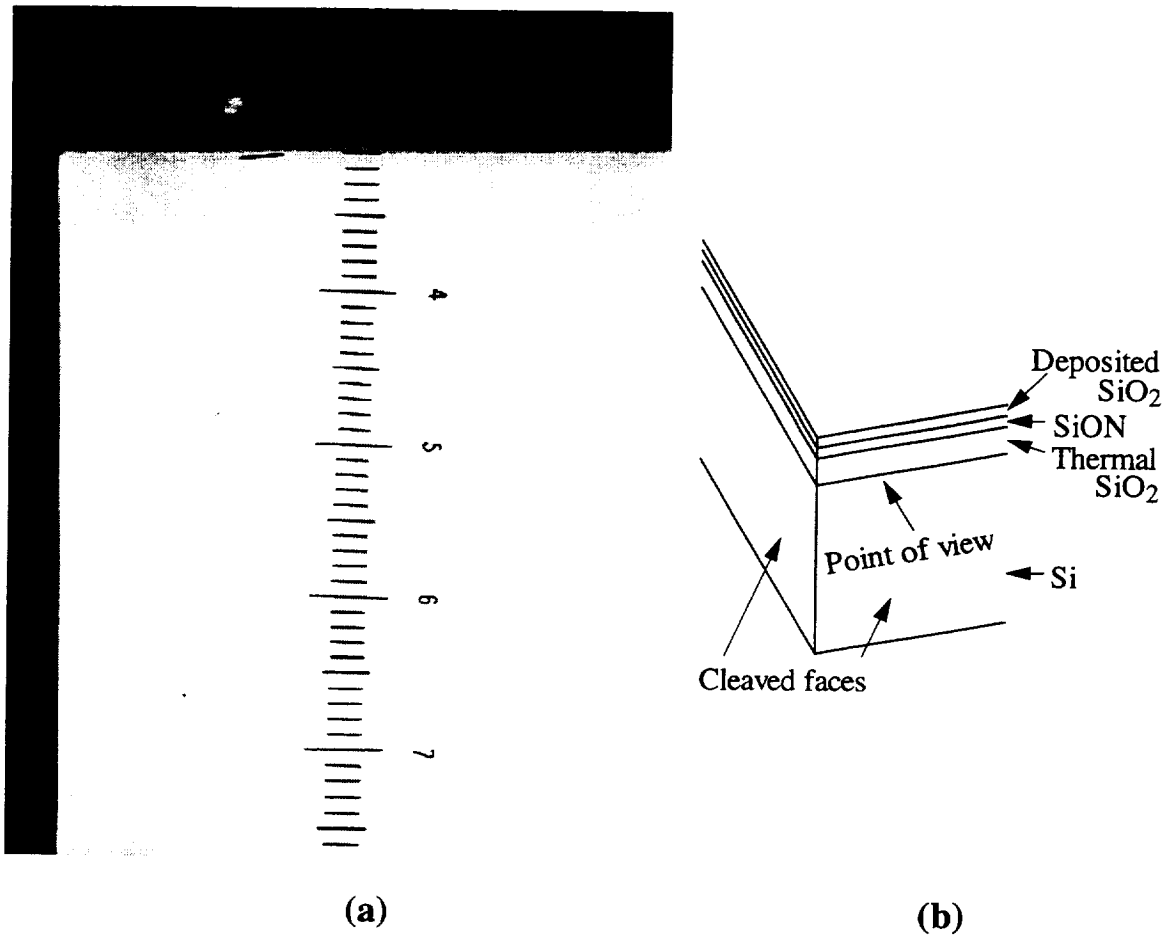


Fig. 4-9 Edge view of corner of wafer bearing waveguide films, (a), and perspective of photo, (b). Both the visible edge and the one meeting it at the corner are approximately perpendicular to the face. Both cleaves would be termed good.

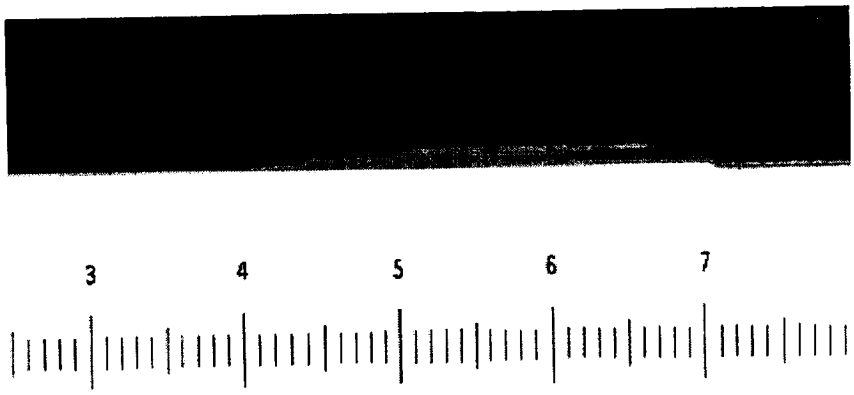


Fig. 4-10 Bad cleave. Although the edge of the silicon wafer appears satisfactory, the SiO_2 isolation and guide layer has chipped off in areas. The structure is the same as in Fig. 4-9

8 DEPOSITION AND CHARACTERIZATION OF WAVEGUIDE FILMS

Fabrication of the integrated optic sensor required forming channel waveguides on substrates having micromachined diaphragms. Passive waveguides have been fabricated by sputtering [7], spin casting polymers [8], plasma depositing organic films [9], flame hydrolysis deposition [10], and low pressure chemical vapor depositing glasses [11]. Polymer films, especially those deposited by spin coating, can be adversely affected by photoresist, and waveguides made from polymers usually experience a large change in their effective index with temperature [12]. Early on in this project RF reactive sputter deposited Si_3N_4 films were evaluated for waveguide use but found to be too lossy. Since a low pressure chemical vapor deposition system was available, a silicon oxynitride (SiON) film formed using this system was selected for the guide layer. Unlike SiO_2 films doped with phosphorus, this film can be deposited without phosphine.

Our LPCVD system for depositing SiON, SiO_2, and Si_3N_4 films is shown in Fig. 4-11, and photographs of the wafer loading and gas supply areas are given in Fig. 4-12 and Fig. 4-13. The system uses dichlorosilane ($SiCl_2H_2$), ammonia (NH_3), nitrous oxide (N_2O), and oxygen, all semiconductor or UHP purity. The waveguide isolation layer and KOH masking films were always grown on the wafer using steam oxidation, rather than deposited using LPCVD, because LPCVD films were found to be less smooth and required more effort to form. In our LPCVD system, N_2 purging is used to remove gases (especially the corrosive $SiCl_2H_2$) from the lines and mass flow controllers at the end of a day's use. Flow limiting valves will close if the flow becomes too high (if, for example, a pipe ruptures). Dichlorosilane is carried to the reactor in a separate line to prevent it from reacting with other gases in the lines. The reactor consists of a horizontal semiconductor diffusion furnace-heated quartz tube, 84 inches long, with 101.6 mm ID and 106.6 mm OD. The flange on the door end had an ID equal to the tube, an OD of 172 mm, and was 0.5 in thick. The other end had a 50/30 ball with O-ring joint which was clamped to a stainless steel socket. Quick clamp NW-40 O-ring sealed fittings carried the gases to the two-stage oil-sealed mechanical roughing pump. Initially, the pump was charged with an inert Fomblin oil. However, one by-product of the $NH_3/SiCl_2H_2$ reaction is ammonium chloride (NH_4Cl) [13], a whitish, water soluble substance which forms in the cooler exhaust end of the reactor and in the pump upon compression of the spent gases, necessitating frequent filter changes, each of which requires adding additional oil. In a study, it was determined that solvents present in a less expensive oil were capable of dissolving another jelly-like by-product of the LPCVD reaction which forms in the pump [13]. No disadvantages were noted to using the less expensive oil. Therefore, to reduce the cost of replenishing oil after a filter change, Fomblin was replaced by TKO-19, an inexpensive mineral oil from Kurt Lesker Co., after pump disassembly and clean out. The pump operated satisfactory subsequent to the oil changeover.

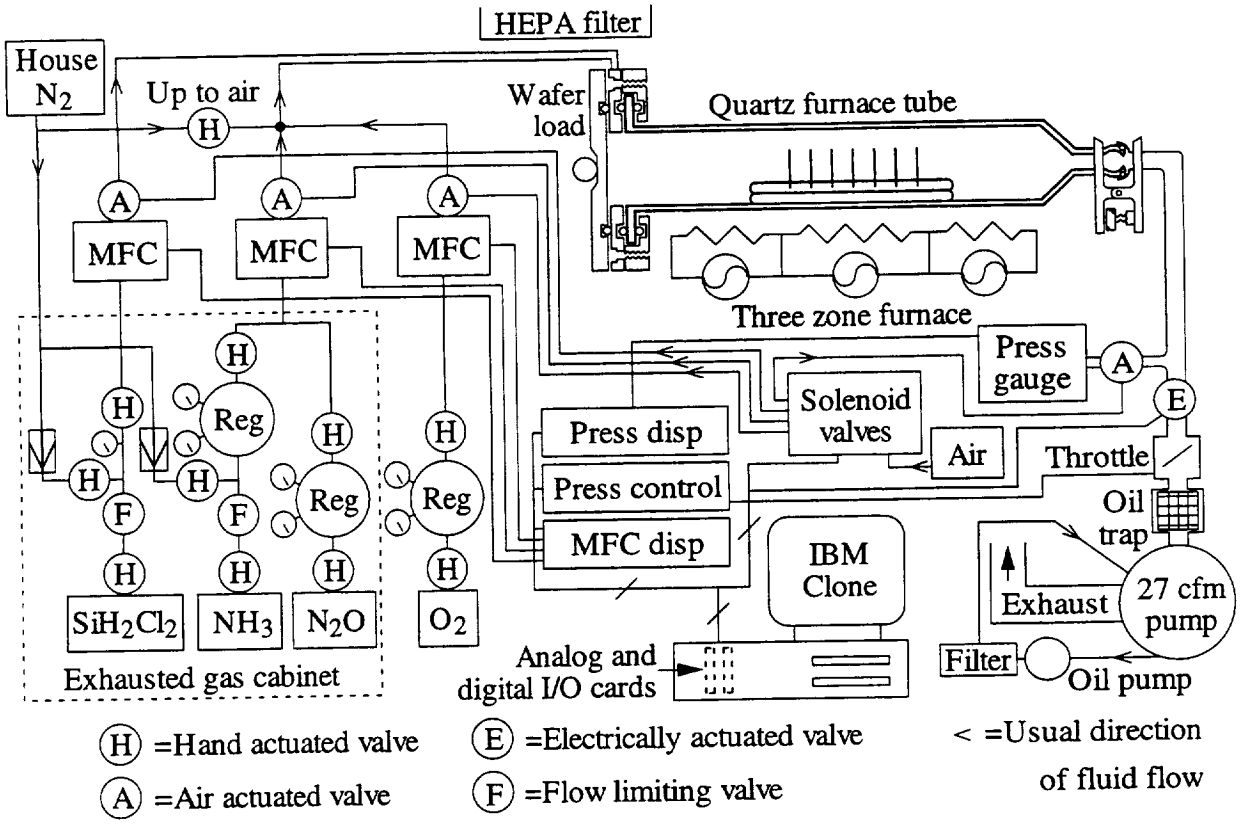


Fig. 4-11 Schematic drawing of the LPCVD system showing the gases used to deposit waveguide films.

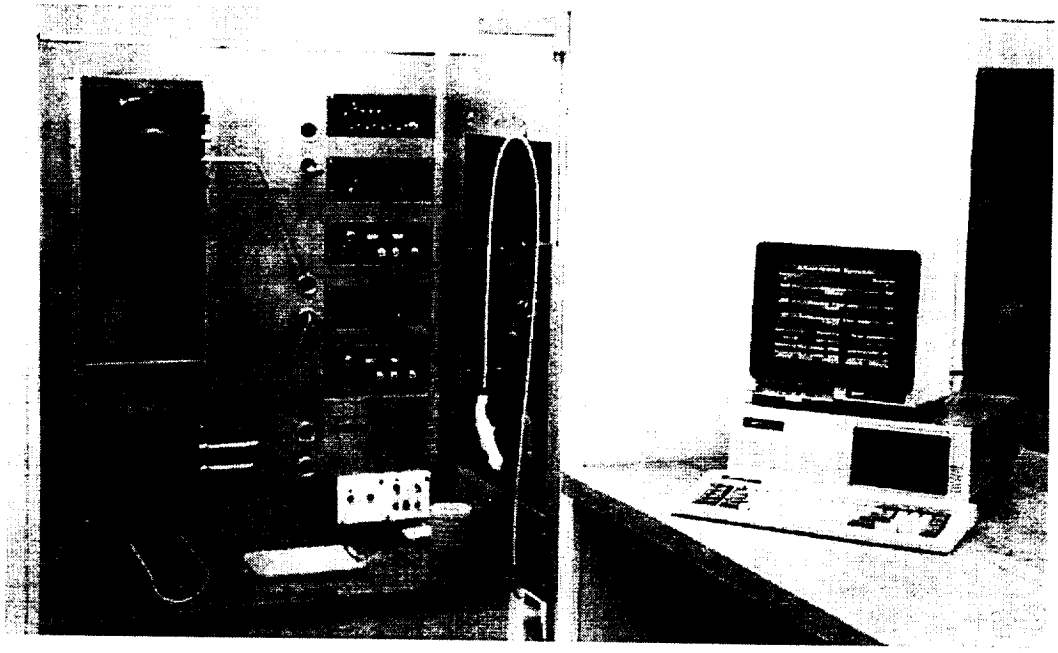


Fig. 4-12 Part of LPCVD system in clean room where wafers are loaded. The upper tube was used for polysilicon, the lower for SiON, Si_3N_4, and SiO_2 depositions. Another tube, not shown, was used for SiO_2 growth.

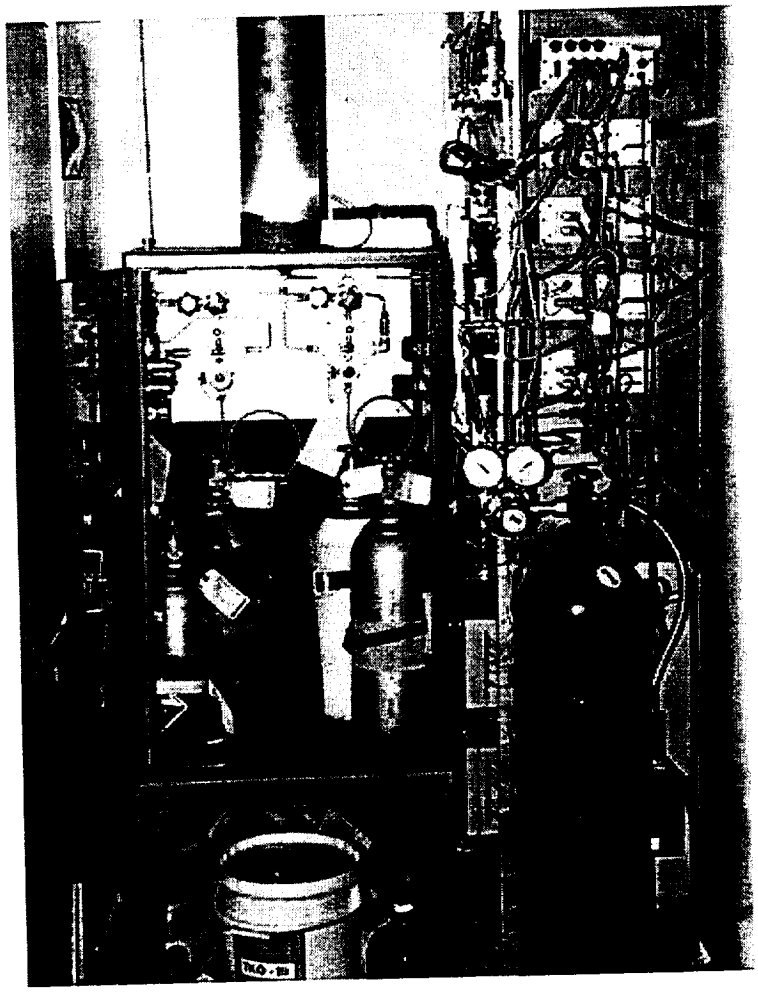


Fig. 4-13 Gas supply for LPCVD outside of clean room. From left to right the gas cylinders are dichlorosilane, ammonia, silane (for polysilicon) nitrous oxide and oxygen.

LPCVD deposition conditions are given in Table 4-2. As the deposition temperature is increased, the deposition rate increases and the uniformity decreases. The latter occurs because with increasing reaction rate, the reactant gases become more depleted in between the wafers, especially near where they are held in the boat. There the gases are restricted from entering by the boat and the smaller gap between the tube wall and the wafer edge.

Film	Temp-erature (°C)	Pressure (mTorr)	Gas flow rate (sccm)				Deposition rate (Å/min)	Refractive index
			$SiCl_2H_2$	NH_3	N_2O	O_2		
Si_3N_4	770	300	25	100	–	–	57	2.02
SiO_2	910	300	10	–	40	–	42	1.46
SiON	750	400	20	100	–	11	75	1.53

Table 4-2 Typical LPCVD deposition conditions for waveguide film formation. The refractive index was measured at $\lambda_0 = 0.6328$ μm.

At a deposition temperature of 770°C, silicon nitride deposited from ammonia and dichlorosilane was found to suffer from little of this depletion effect, even coating the areas of the wafer held in the boat slots. On oxidized silicon wafers low-loss, crack-free waveguides could be formed by keeping the Si_3N_4 thickness less than 2000 Å. An attempt was made to form LPCVD Si_3N_4 waveguides on polished fused silica discs but the film was found to crack and possess high loss. Silicon's coefficient of thermal expansion is higher than that of Si_3N_4 and the film's tensile stress is somewhat relieved on cool down after deposition, while silica, possessing a smaller coefficient of thermal expansion than Si_3N_4, effects the reverse. The ratio of ammonia to dichlorosilane, which was maintained at 4:1 to insure stoichiometry, cannot be manipulated to reduce the high tensile stress of the deposited film [14].

Silicon nitride was found to be easily deposited in uniform films, but silicon oxynitride was more problematic, as the prism coupling measurements made at $\lambda_0 = 0.6328$ μm in Fig. 4-14 show. The effective index of the film increases toward the part of the wafer nearest the top of the tube because the film is thicker there. A method of increasing uniformity and film thickness accuracy is to break the run into two and rotate and exchange the wafers in between. After the first half of the run, the wafers were removed from the reactor and usually base (5:1:1 by volume of H_2O:30% NH_4OH:30% H_2O_2) cleaned to remove particulates which might have come onto them in letting the

tube up to air. A bare monitor wafer also included in the run was ellipsometrically measured, and the duration of the second half of the run adjusted to hit the thickness target. Then the wafers were placed back into the boat rotated by 180° about an axis through their center, with the wafer in the first slot swapped with the wafer which used to be in the last slot, the wafer in the second slot swapped with the wafer in the second last slot, etc. To enable tracking the wafers, numbers were scribed on their faces and they were placed with their major flats up for the first half the run, and so had their flats down for the second half. The data in Fig. 4-14 shows that this procedure increases the film uniformity. Rotating and exchanging the wafers was also found to be useful to improve the thickness uniformity of the LPCVD deposited SiO_2 films.

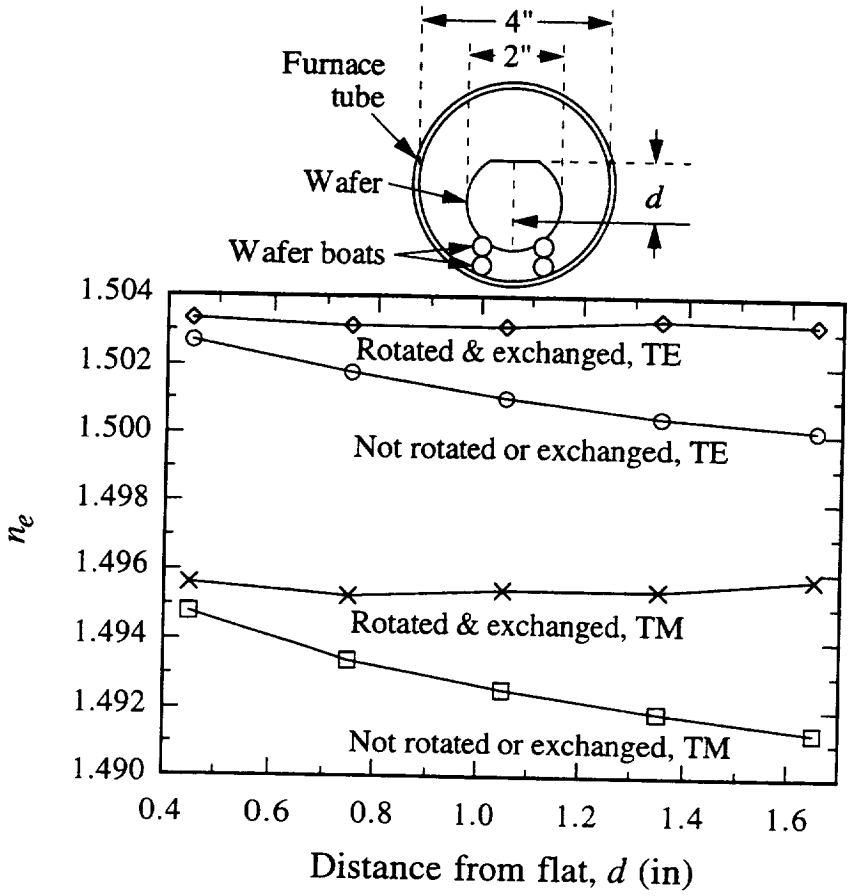


Fig. 4-14 Effective index measured by prism coupling light with $\lambda_0 = 0.6328\ \mu m$ into two LPCVD grown SiON films on thermally grown SiO_2. The films were approximately 0.5 μm thick and had a refractive index of 1.552. In the run in which the wafer was not rotated or exchanged, the flat was up, as shown in the drawing above the plot. In both runs, the wafer was positioned between two other 2" wafers, each of which was 1 cm away.

There are a number of gas systems which can be used to deposit various refractive index films of silicon oxynitride. One of these, suggested by Kuiper *et al.* [15] uses $SiCl_2H_2$, NH_3, and N_2O. However at a particular temperature, the deposition rate of Si_3N_4 using $SiCl_2H_2$ and NH_3 is higher than that of SiO_2 using $SiCl_2H_2$ and N_2O, so SiON films deposited using $SiCl_2H_2$, NH_3, and N_2O tend to have near-nitride indices for all but high flow ratios of N_2O/NH_3. It was found for example, that at a temperature of 920°C, a pressure of 333 mTorr, and with flow rates of 25 sccm $SiCl_2H_2$, 7 sccm NH_3, and 100 sccm N_2O, a film having refractive index of 1.54 deposited at a rate of 76 Å/min.

The SiON process used in this work was suggested by Gleine *et al.* [11]. It employs $SiCl_2H_2$, NH_3, and O_2, and it was found, within the range of flows which could be covered by mass flow controllers with the LPCVD, to produce films having a lower index of refraction than the process using N_2O. The refractive index of the Gleine SiON films could be adjusted by changing the O_2 gas flow, and near 11 sccm, a 1 sccm increase in flow rate produced a 0.01 decrease in the film refractive index. In order to avoid depositing Si_3N_4, oxygen, then ammonia, then dichlorosilane were introduced into the deposition tube and they were turned off in the reverse order. It was observed that without NH_3 present no film would deposit at 750°C from $SiCl_2H_2$ and O_2. The refractive index reported in Table 4-2 for the SiON film was that typically measured. The value ranged from 1.51 to 1.55 depending on aging (the films tended to creep, as discussed later) and the care taken to measure the flow and the zero offset in the mass flow controller. The ellipsometric measurement of refractive index was also sensitive to the film thickness, as discussed later. If hitting a specific index is critical, it is recommended that a test run be performed. If the film is to be characterized ellipsometrically, the target film thickness should be near an ellipsometric midperiod.

In all the runs, regardless of the film to be deposited, the 2" wafers were held in a fused silica boat placed atop another boat to move the wafer center closer to the center of the 4" tube (Fig. 4-14). The boat holding the wafers should be chosen so that the wafers do not flop over and occlude each other. A wafer-to-wafer separation of 1 cm was applied. The good side of the wafer faced toward the pump to keep the particulates which might be drawn out on pump down from hitting the good face. As Fig. 4-1 shows, the wafers were held vertically. A thicker film deposited on the side of the first and last wafer which did not face another wafer (Fig. 4-11) so a dummy wafer was usually placed at the beginning and end of the wafer pack.

After deposition, the ellipsometrically measured refractive index and thickness of the silicon oxynitride film was found to creep, as shown in Fig. 4-15. The film was

deposited before the conditions in Table 4-1 were settled on, as the Figure caption shows; the higher temperature of 770°C increased the deposition rate. It was found that annealing the film at 1100°C for 1/2 hour caused a reduction of the refractive index, as Fig. 4-15 shows. The lack of observed creepage in the film after the anneal might be attributed to the film having settled prior to the anneal. However, in another case in which the 1/2 hour anneal was begun within 6 minutes of the end of the LPCVD deposition, no change was measured between the end of the anneal and 1/2 hour later. Unlike in Fig. 4-15, this anneal reduced the refractive index by 0.02.

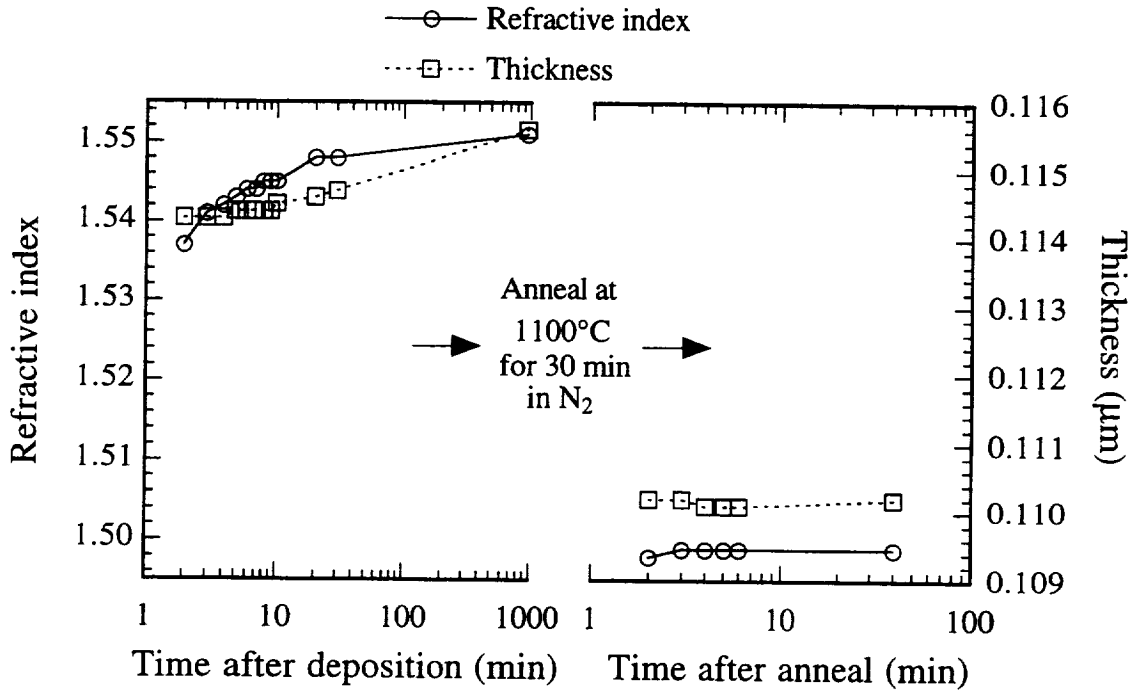


Fig. 4-15 Creep in the refractive index and thickness of LPCVD deposited SiON film. The deposition conditions were: temperature 770°C, flow rates: 25 sccm $SiCl_2H_2$, 100 sccm NH_3, 14 sccm O_2; 400 mTorr pressure; deposition time 10 min.

It is not known whether the creepage is due to a change in the exposed surface of the film or occurs throughout the bulk of the film. Since the ellipsometer measures the change in the state of polarization of reflected light, while the waveguide n_e "measures"

the overlap of the guided mode with the film refractive index profile, the two may not be equally sensitive to the creepage effect.

After the creepage problem was discovered, an attempt was made to stabilize the film by annealing. However, it was found that thicker, higher refractive index films contained crack lines after annealing. In one instance, a film having a refractive index of 1.569 and thickness of 0.6892 μm before anneal was observed to possess a few cracks after a 1/2 hour, 950°C anneal in N_2, which altered its index to 1.534 and thickness to 0.6921 μm. After stripping the film in HF, lines remained on the wafer which appeared similar to the cracks in the film. This could be due to etching of the substrate through the cracks or it could be a stress induced defect in the crystal. Film cracking appeared to be more of a problem on <111> than <100> substrates. To alleviate the cracking problem, only films having a low index of refraction were used and the annealing procedure was abandoned.

9 PATTERNING THE CHANNEL WAVEGUIDES

Channel waveguides were formed by wet chemical etching partially through a SiON film or partially through a capping oxide film while masking with a positive photoresist. The channel etch depth and width must be well controlled to obtain single mode guiding and low loss bends. Defects in the resist such as mouse bites, breaks, scumming, and line edge roughness will increase the channel guide loss. Although patterning using plasma etching with CHF_3 gas was attempted, the etched surface was found to be rougher than that left by wet chemical etching using buffered HF and so was not used.

Initially wet etching was performed using buffered HF acid consisting of room temperature 6:1 by volume of 40% NH_4F:49% HF. However, the etch rate for SiO_2 deposited by LPCVD using the conditions of Table 4-1 was about twice that of steam grown thermal oxide, or 2000 Å/min. For unannealed SiON grown according to

Table 4-1 and therefore having a refractive index of 1.53, the etch rate was measured to be 3800 Å/min.

To fabricate 2.4 μm wide single mode channels in SiON, the etch depth is typically chosen to be between 800 Å and 1500 Å, with the precise target value dependent on film thickness and whether a capping oxide is used. This implies an etch time of 13 sec - 24 sec. The time required to rinse the etchant off the wafers is a significant part of the etch time and this may cause a nonuniform etch depth. A solution purchased from Transene [3] called Timetch was found to etch more slowly, permitting better control of the etch depth. According to its material data safety sheet, this solution contains 20% ammonium fluoride and 5% ammonium phosphate in an aqueous solution. This solution was used unstirred and straight from the bottle. Fig. 4-16 and Fig. 4-17 show the etch rate of unannealed SiON deposited using the conditions in Table 4-2 as a function of solution temperature and reciprocal temperature (Arrhenius plot). Even though these fits require a different functional etch rate dependence to temperature, the fit to each is good because the temperature range is small. From the exponential fit, it can be observed that the etch rate doubles every 11°C and increases 6.4% per degree C, necessitating tight temperature control to achieve reproducible etch rates. The activation energy is typical in wet etching: another etch study measured a similar activation energy for Si_3N_4 etching in solutions of HF in glycerol at elevated temperatures (70-90°C) [16].

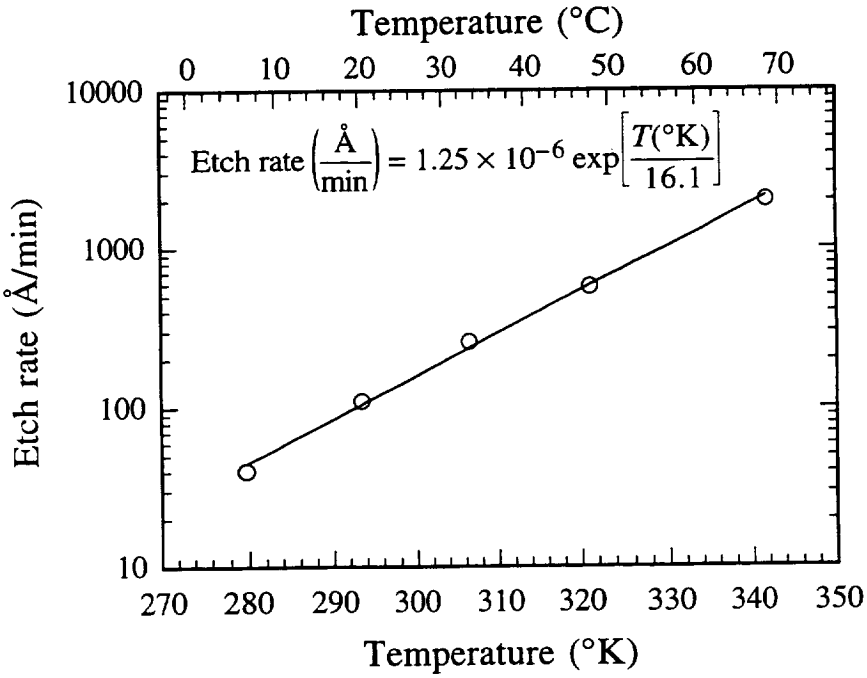


Fig. 4-16 Dependence of etch rate of unannealed SiON in Timetch as a function of solution temperature along with the best fitting exponential function.

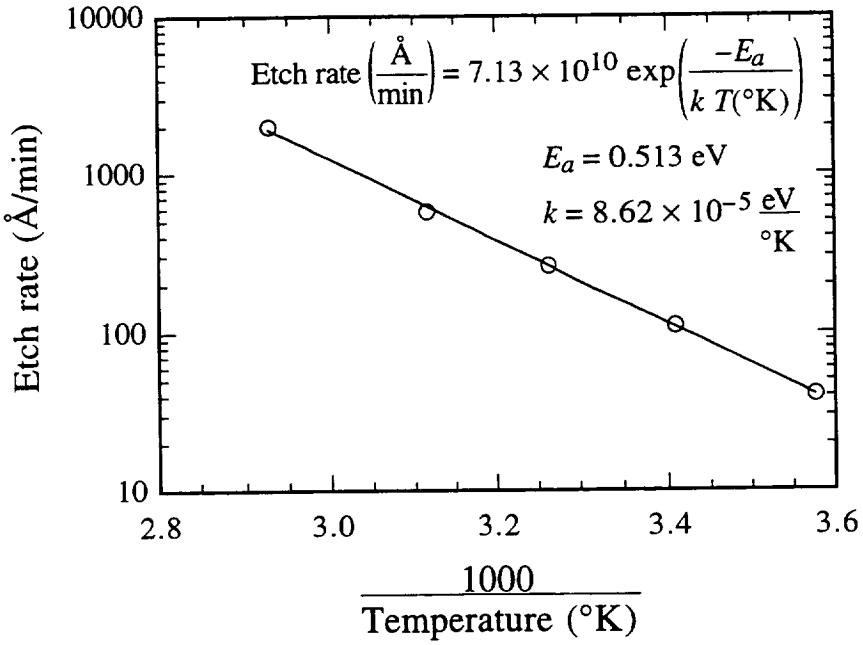


Fig. 4-17 Arrhenius plot of data in Fig. 4-16. The activation energy is 0.513 eV or 11.8 kcal/mol.

10 WAVEGUIDE LOSS MEASUREMENT

Propagation loss measurement of planar and channel SiON waveguides has been carried out by coupling light into the film and measuring the reduction in intensity of the outscattering streak. This method is nondestructive but generally requires channel power to go down by a dB or so over the measured distance. It also requires that if the substrate guides, that outscattering from it not obscure that from the film, and that bright outscattering points not perturb the measurement. The outscattering pickup must be shielded against light scattered from the incoupling edge or prism. The setup consists of a laser source, focusing lens or incoupling prism, and microscope having an eyepiece with pick up (Fig. 4-18). In the microscope, the objective forms a real image in a plane located just below the eyepiece where a bar containing a prism is located. Light is admitted to the prism through a 1.25 mm diameter aperture in the bar. The aperture position is referenced on the bar by a small protrusion on its three sides, and these, along with the image of the waveguide surface are simultaneously in focus in the eyepiece. Light from the waveguide is reflected off the prism's hypotenuse in the bar and coupled into a fiber bundle conveying it to a photomultiplier tube. If measurement is to be performed at the 830 nm diode laser wavelength, a camera having a lens to convert the virtual image into a real image is positioned above the eyepiece.

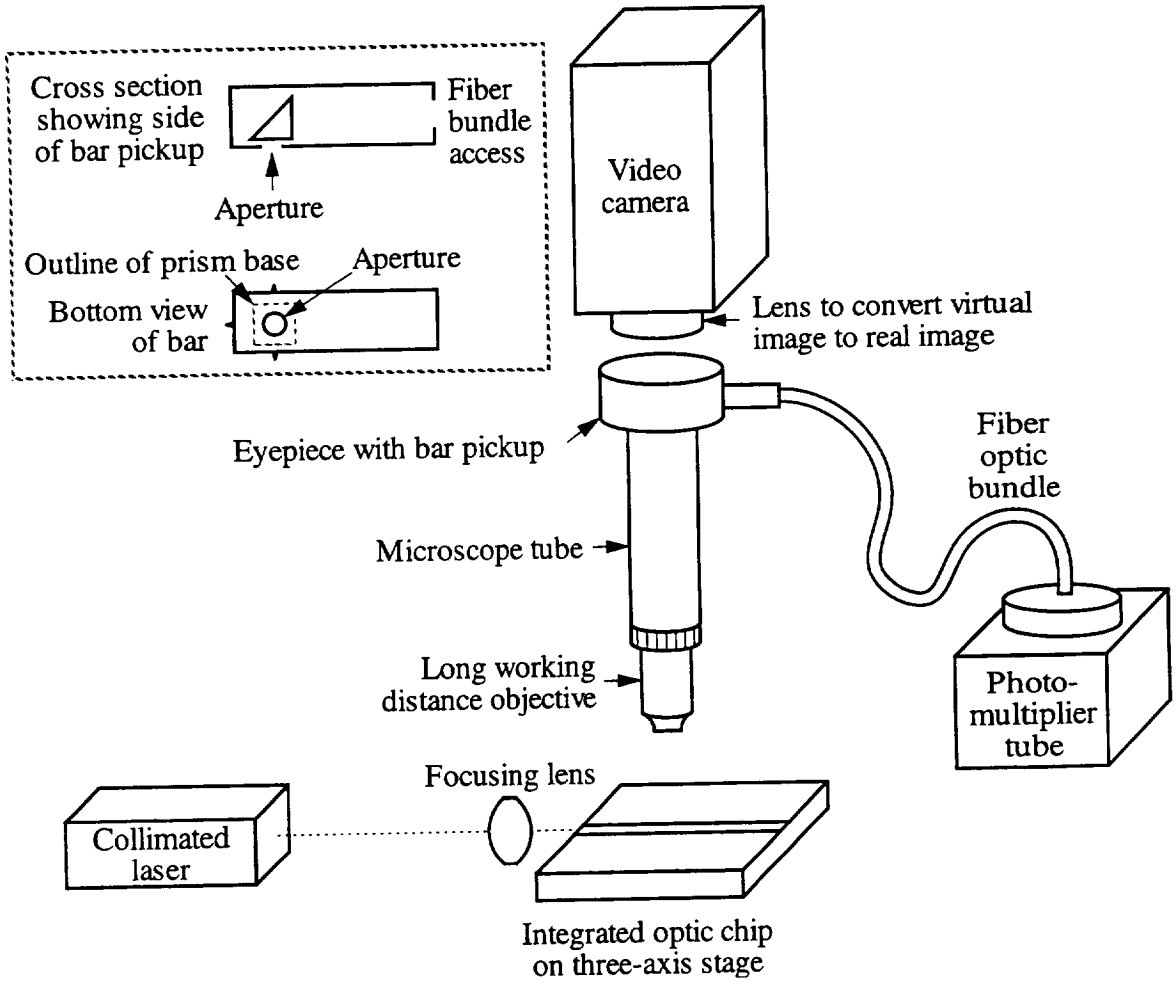


Fig. 4-18 Waveguide outscattering loss measurement setup.

To take a loss measurement, the microscope objective is positioned over the streak at least 2 mm from the incoupling point to assure that the mode has settled. The focus and lateral position of the microscope are adjusted to center the aperture over the streak. An intensity measurement is taken using the photomultiplier tube, and the stage holding the microscope and camera advanced using a micrometer. After several steps, the focus and lateral position are readjusted. The power lost per unit length is proportional to the power in the guide and so the streak intensity is expected to decay exponentially. The log of the intensity measurements are fitted to a line and its slope used to find the propagation loss, typically expressed in dB/cm.

The method assumes that the ratio of outscattered intensity to power in the guide remains constant along the measured length of the streak. Random defects or a gradation in surface roughness can cause a variation in this ratio and reduce the accuracy of the measurement (Fig. 4-19). To a degree, this problem can be overcome by averaging the loss measurements made with light coupled into each side separately.

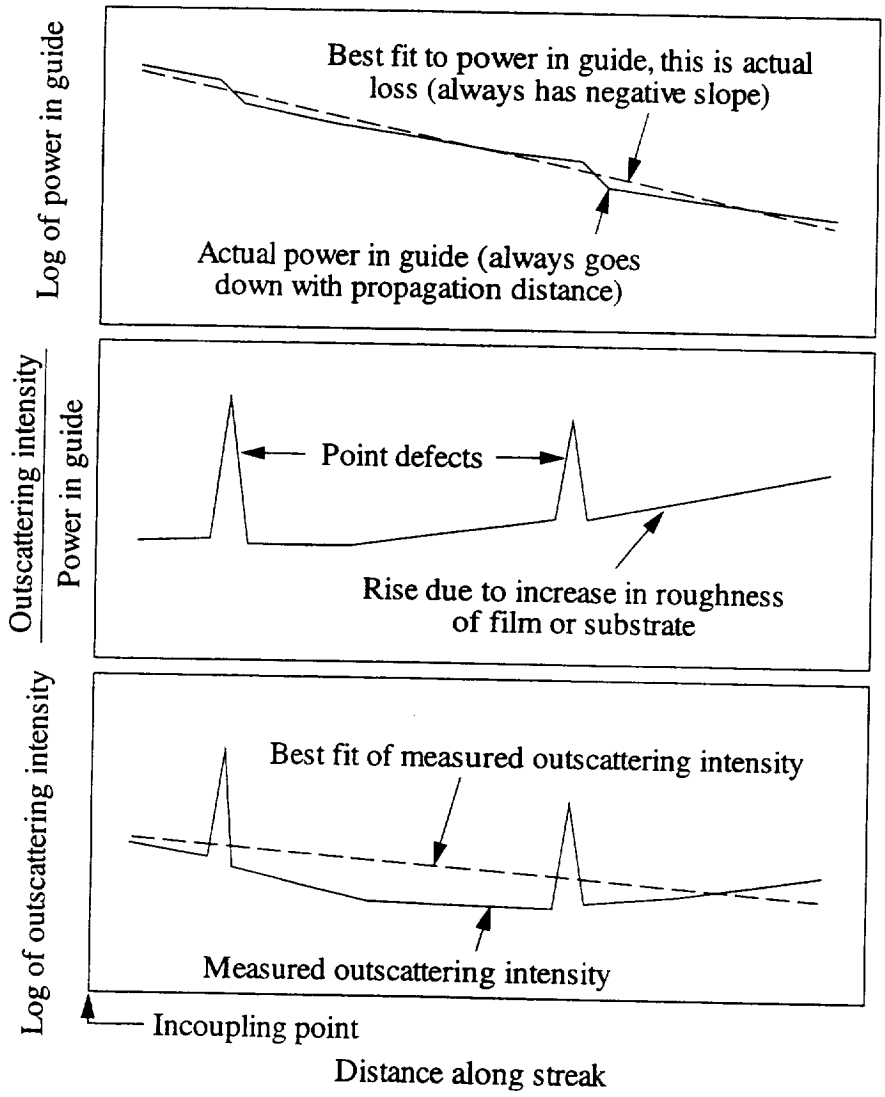


Fig. 4-19 Actual and measured propagation loss. In this model, outscattering intensity is the product of power in the guide and the ratio of outscattering intensity to power in the guide.

It will now be shown that the average of the dB/cm losses measured with light coupled into opposite sides of the waveguide one side at a time yields a more accurate measure of propagation loss than either of the two measurements. In this derivation, for simplicity, it will be assumed that the outscattering intensity is measured at only two places, location 1 and location 2, with light coupled into side 1 or side 2 (Fig. 4-20). These four outscattering intensity measurements, labeled A, B, C, and D are related to the power in the guide P_a, P_b, P_c, and P_d by

$$\begin{aligned} A &= P_a R_1 \\ B &= P_b R_2 \\ C &= P_c R_1 \\ D &= P_d R_2 \end{aligned} \quad , \tag{4.10.1}$$

where R_1 and R_2 are the ratio of outscattered light to power in the guide at locations 1 and 2. Here we will assume that the ratios R_1 and R_2 are independent of the direction of propagation of the light. With the outscattering technique, we measure, with light coupled into side 1,

$$\text{Loss (dB)} = 10 \log_{10}\left(\frac{A}{B}\right) = 10 \log_{10}\left(\frac{P_a R_1}{P_b R_2}\right) \quad , \tag{4.10.2}$$

while with light coupled into side 2, we obtain

$$\text{Loss (dB)} = 10 \log_{10}\left(\frac{D}{C}\right) = 10 \log_{10}\left(\frac{P_d R_2}{P_c R_1}\right) \quad . \tag{4.10.3}$$

In contrast, we desire

$$\text{True loss (dB)} = 10 \log_{10}\left(\frac{P_a}{P_b}\right) = 10 \log_{10}\left(\frac{P_d}{P_c}\right) \quad . \tag{4.10.4}$$

However, if $R_1 \neq R_2$ because of a defect or difference in surface roughness between locations 1 and 2, the loss measured when coupling into one side will be too high, and when into the other side, too low. Under the assumption that R is independent of direction of propagation, true loss may be recovered by averaging the two dB values measured while coupling into each side separately. To see this form

$$\text{Averaged loss (dB)} = \frac{10\log_{10}\left(\frac{A}{B}\right) + 10\log_{10}\left(\frac{D}{C}\right)}{2}$$

$$= \frac{10\log_{10}\left(\frac{P_a R_1}{P_b R_2}\right) + 10\log_{10}\left(\frac{P_d R_2}{P_c R_1}\right)}{2}$$

$$= \frac{10\log_{10}\left(\frac{P_a R_1}{P_b R_2}\,\frac{P_d R_2}{P_c R_1}\right)}{2}$$

$$= \frac{10\log_{10}\left(\frac{P_a P_d}{P_b P_c}\right)}{2} \qquad (4.10.5)$$

$$= \frac{10\log_{10}\left(\frac{P_a}{P_b}\right) + 10\log_{10}\left(\frac{P_d}{P_c}\right)}{2}$$

Since the ratio P_a/P_b equals P_d/P_c,

$$\text{Averaged loss (dB)} = 10\log_{10}\left(\frac{P_a}{P_b}\right) = 10\log_{10}\left(\frac{P_d}{P_c}\right) = \text{True loss} \qquad (4.10.6)$$

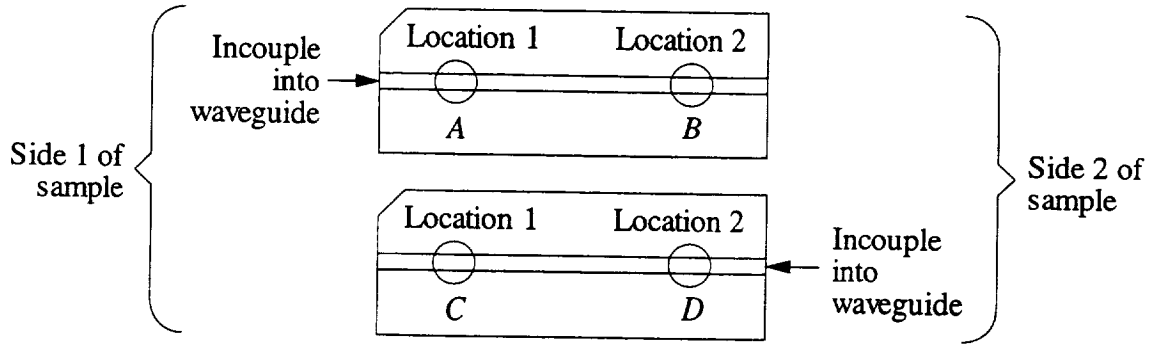


Fig. 4-20 Analysis of outscattering technique when averaging two-sided loss. *A-D* represent measured outscattering intensities.

This technique will work best if the ratio difference is due to a difference in surface roughness rather than an isolated point defect, since in the former case, one must take both sets of measurements at nearly identical locations. In practice, a propagation loss measurement is taken by coupling separately into opposite sides of the guide and averaging the two loss measurements obtained from a best fit found using many points.

This is done because of the difficulty of hitting the same spot in two separate measurements and the typically large random variation in R over the guide.

Fig. 4-21 shows loss measured by outscattering at $\lambda_0 = 0.6328$ μm on a planar, TE excited, SiON waveguide. The sample was fabricated before the conditions in Table 4-2 were settled on and so it was not rotated half way through deposition (which improves uniformity), the oxygen flow rate was slightly lower, giving a refractive index of 1.56, and the wafer was subsequently annealed in N_2 at 1100°C for 1/2 hour. To avoid having the beam spread out in the film, only weak focusing of the beam onto the prism was employed. The loss measured with the beam propagating away from the flat is apparently negative because the ratio of outscattering to power in the film increases as one moves away from the flat. By averaging these results, a planar loss of 0.29 dB/cm is obtained. Other planar waveguides grown in the same run showed a similar effect and their average was also similar.

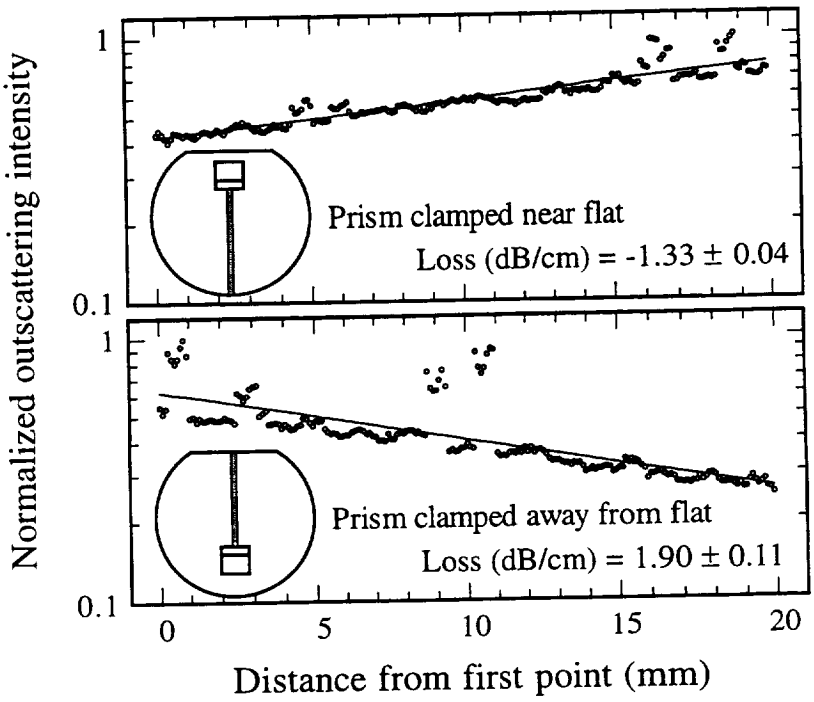


Fig. 4-21 Propagation loss measured by prism coupling into a planar waveguide so that the streak (the shaded line) would propagate away from the flat, and toward the flat.

Propagation loss of straight channels was also measured. Fig. 4-22 shows TE outscattering measured at $\lambda_0 = 0.830$ μm on channel ①-③ as a function of propagation distance. The channel, whose cross section is given above the plots, was made using the photomask described before. In contrast to the planar case, random fluctuations in outscattering intensity are seen over the entire streak. Apparently the loss in one direction is again negative because the ratio of outscattering to power in the guide increases toward the flat. In this deposition, the wafer was rotated and the flat was down for the last half of the run. The average of the losses is 0.57 dB/cm. A photo of a typical section of the channel used for loss measurements (Fig. 4-23) shows that the edges are free of almost all roughness and imperfections. When light was coupled into port ② a negative TE loss was also measured and the averaged loss for this channel was 0.51 dB/cm. The averaged TM losses for these two channels were 0.47 dB/cm and 0.27 dB/cm. These are the lowest channel losses that were obtained in this work.

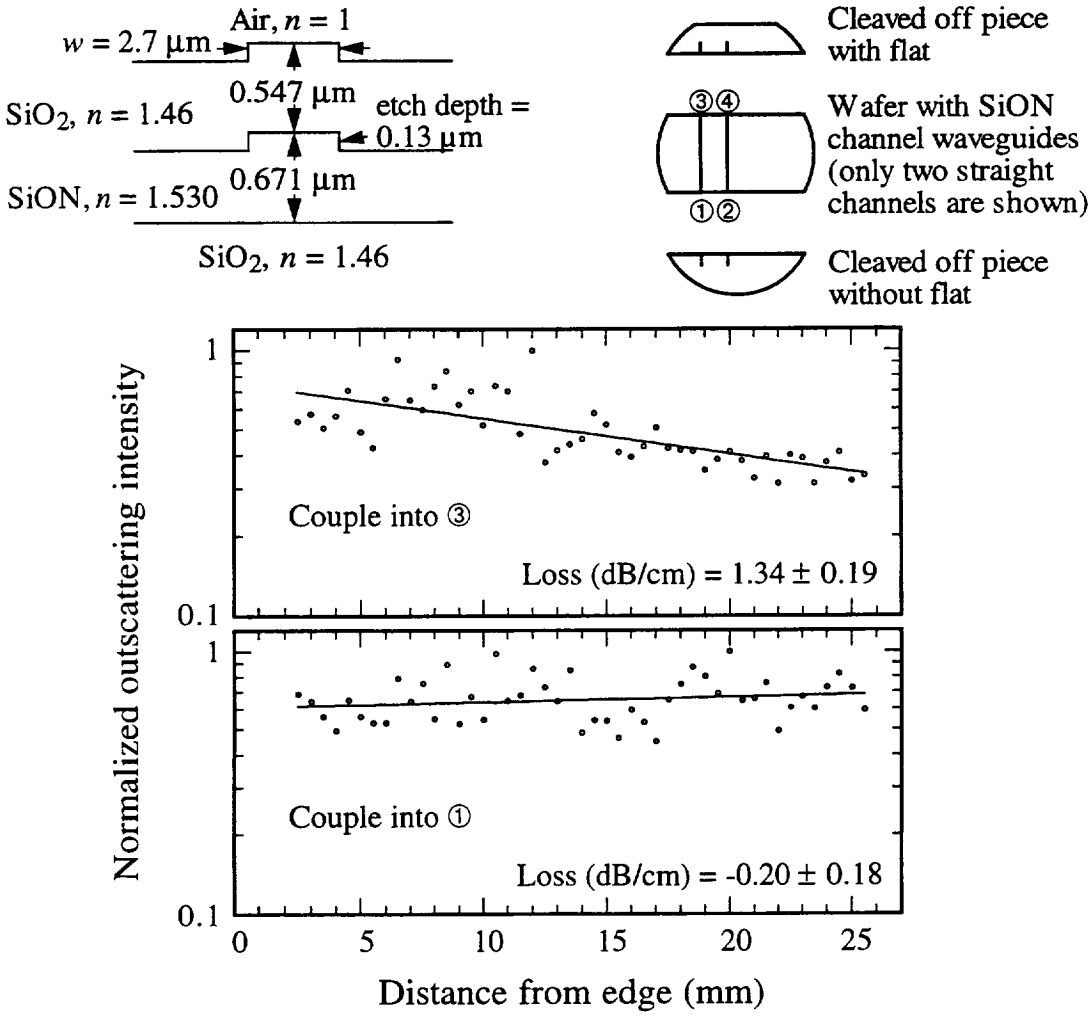


Fig. 4-22 Channel propagation loss.

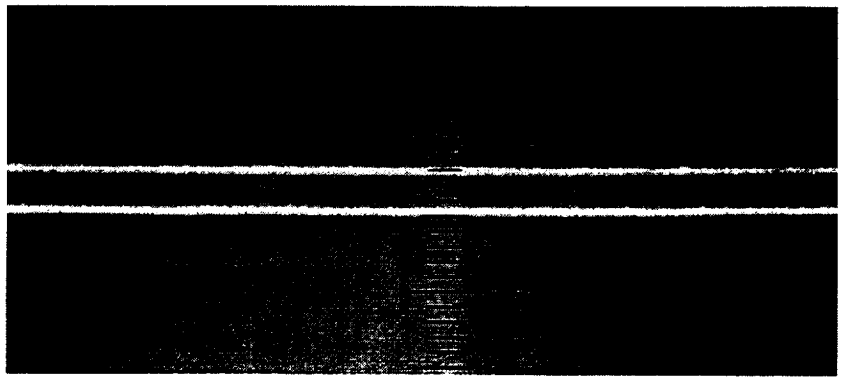


Fig. 4-23 Differential interference contrast photo of ①-③ channel whose loss is tabulated in Fig. 4-22. The channel edges are indicated by the white lines. The graticule is 0.96 μm/div.

Roughness in the substrate and channel sidewalls can increase the propagation loss. An experience was had with Si_3N_4 films deposited on thermally oxidized silicon substrates to form waveguides. The wafers had a polish which appeared to the unaided eye to be satisfactory but under the differential interference microscope the surface looked patchy. The planar waveguide loss was around 5 dB/cm, in contrast to the usual 0.3 dB/cm on substrates with a better surface. Afterward, all wafers used for waveguides were inspected under this microscope for surface finish and the better face on double side polished wafers was used for the waveguides.

Not all resist lines came out as well as that in Fig. 4-23. In order to keep the channel line width from shrinking, the least exposure required to clear the resist was used. Occasionally a scumming problem was encountered in which channel resist lines had curls of resist connected to them. Etching the film which had such line defects produced the result shown in Fig. 4-24. To avoid such lossy channels, the resist should be reworked, and the exposure time increased. In addition, the photomask should be cleaned after each print using acetone, and after several prints it should be lightly rubbed with an acetone soaked semiconductor wiper. Periodic examination under a microscope to check for residue near the chrome lines is also recommended.

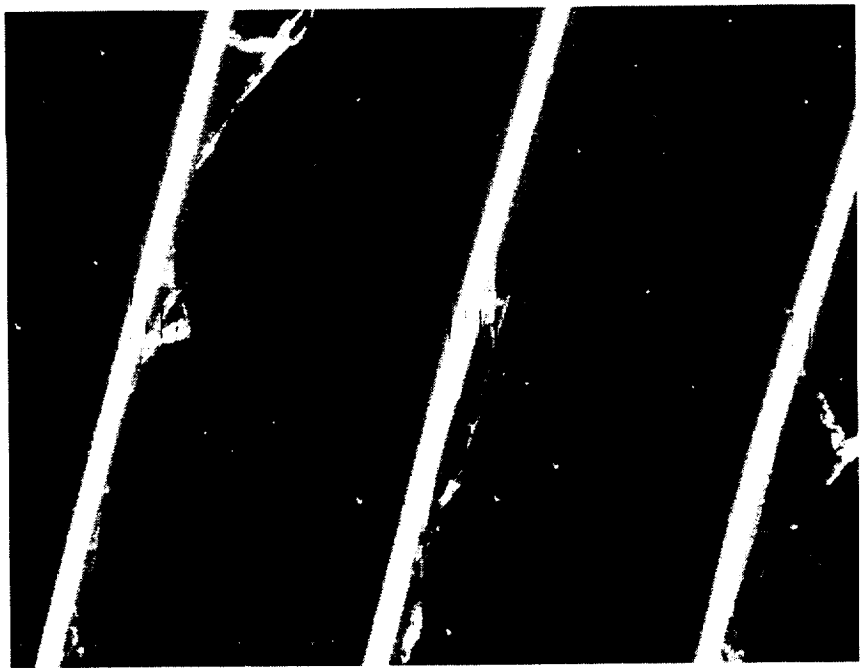
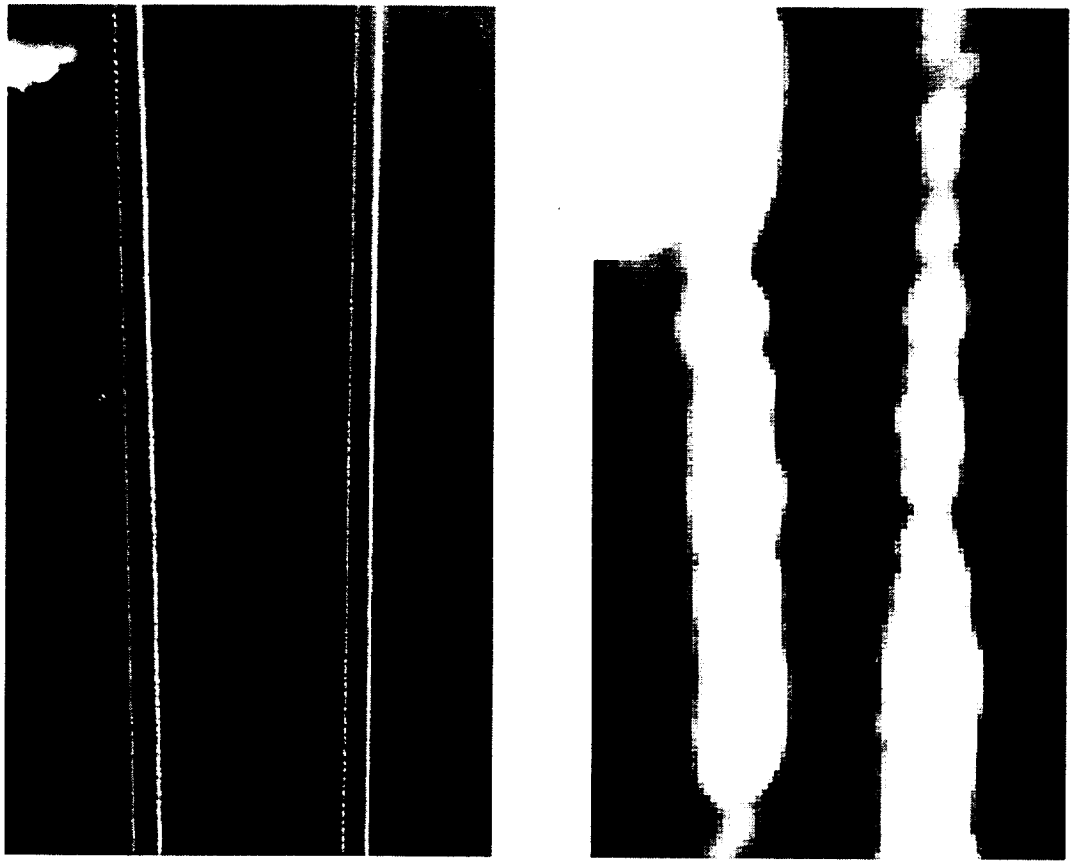


Fig. 4-24 Bad print and etch. The white lines are channels, made wider by overexposure to enable photographing the scumming. The graticule is 4.78 μm/div.

Fig. 4-25 shows roughness in the sidewall of a channel waveguide and how this scatters $\lambda_0 = 0.830$ μm guided light, as seen on a video monitor. The channels are 2.5 μm wide and the index contrast between unetched and etched planar guiding regions was 0.007, while to maintain single mode confinement, the index difference must be less than 0.00929. A slight amount of roughness increases the outscattering and therefore the propagation loss.



(a) (b)

Fig. 4-25 Photograph of channels in *Y*-junction showing roughness (a), and the scattering this produces to 830 nm guided light (b). A blemish in the substrate toward the upper left of the photos is visible in both.

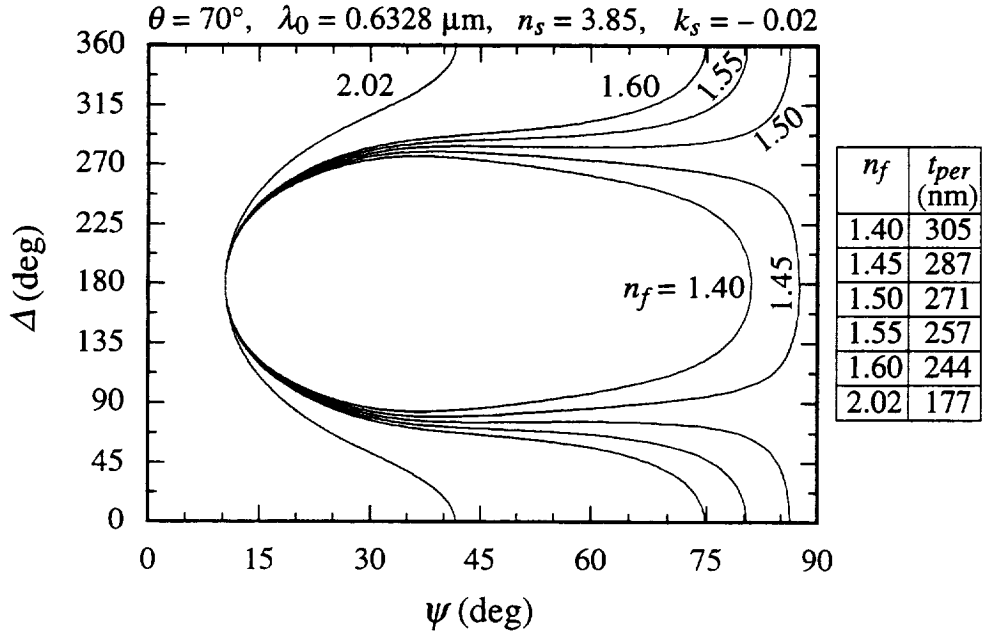
11 MEASURING FILMS WITH THE ELLIPSOMETER

A Gaertner L116A computer controlled ellipsometer was used to measure the thickness of thermal oxide grown to protect the wafer during KOH etching, as well as the thickness and index of LPCVD deposited films. This instrument, which uses a HeNe $\lambda_0 = 0.6328$ μm laser, can be manually set to 30°, 50°, 70°, or 90° angles of incidence. It uses a rotating analyzer (polarizer) and photodiode to measure the state of polarization of light reflected from the substrate. The substrate is illuminated with light linearly polarized 45° with respect to the plane of incidence, or by automatically switching a

compensator in the beam path, circularly polarized light. Software provided by Gaertner reads the intensity incident on the photodiode every 5° increment of the motorized analyzer drum and takes 72 consecutive readings without the compensator in the beam path, and 72 with it in, and then calculates ψ (psi) and Δ (del). From these two values and knowledge of the substrate refractive index (n_s) and extinction coefficient (k_s), the angle of incidence (θ), the source wavelength, that air ($n = 1$) is above the film, and that the film has extinction coefficient of 0, the film's refractive index n_f and thickness t_f can be determined. One n_f-t_f measurement takes about 5 sec. Software provided by the manufacturer allowed the instrument to measure transparent film thickness (within a known period) and index, and to determine the period order and thickness of SiO_2 and Si_3N_4 films on silicon by measuring at two angles of incidence.

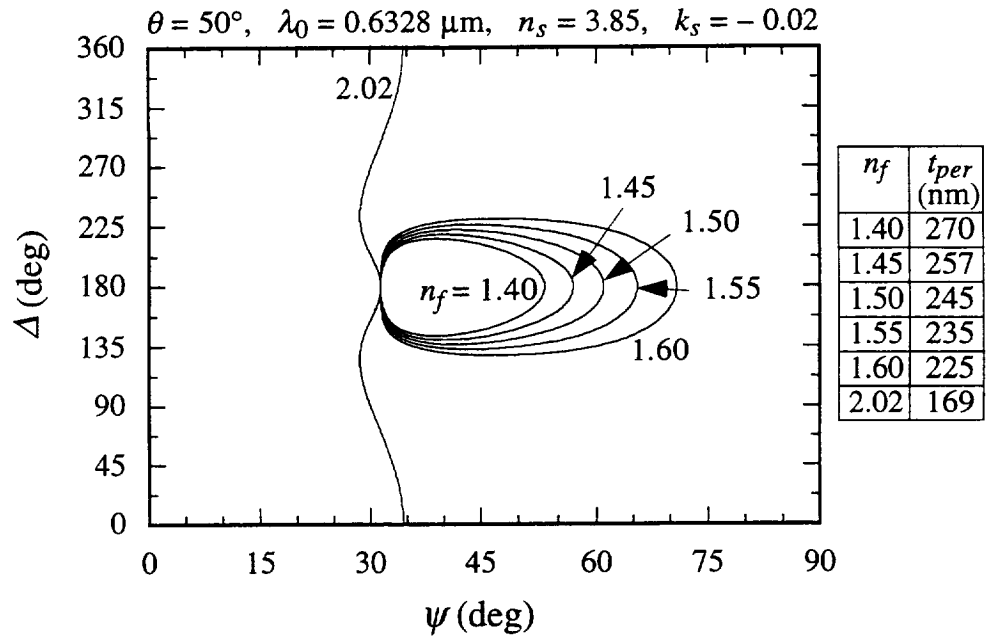
Figure 4-26 plots contours of constant index of transparent films of interest in this work in the ψ-Δ plane on silicon substrates at $\theta = 70°$ in a manner similar to Archer [17]. ψ and Δ are defined by $\tan(\psi)\, e^{j\Delta} = R_p/R_s$, where R_p and R_s are the complex amplitude p and s polarized reflection coefficients. As the film increases from zero thickness, a contour is traced counterclockwise beginning from $\psi = 10.34°$, $\Delta = 179.16°$. When the film hits a period thickness given by $t_{per} = \dfrac{\lambda_0}{2n_f \cos(\theta_f)}$, where θ_f is the angle of propagation of the beam in the film determined from Snell's law, a point on the contour in the ψ-Δ plane will have returned to the point it was located when $t_f = 0$. Increasing the film thickness beyond this results in retracing the same contour. There is thus a period ambiguity which can be resolved by either changing the angle of incidence and remeasuring ψ and Δ or from knowing the approximate film thickness. More serious is the difficulty of measuring the refractive index of the film in the vicinity of the period thickness. Here, slight errors in measuring ψ or in the assumed value of the substrate index or extinction coefficient will produce large errors in the estimated film index. Figure 4-27 shows the contour plot for $\theta = 50°$; the period thickness has become smaller and the curves more bunched together, making a determination of the correct contour

more difficult. However, because the period thickness at 70° and 50° is different, for a particular film thickness the contours may be better separated at 50°, enabling a more accurate measurement.



n_f	t_{per} (nm)
1.40	305
1.45	287
1.50	271
1.55	257
1.60	244
2.02	177

Fig. 4-26 Contours of ψ and Δ for films of practical interest in this work on a silicon wafer at an angle of incidence of 70°. The table on the right provides period thicknesses for various films at $\theta = 70°$ and $\lambda_0 = 0.6328$ μm. The value taken for k_s is the default of the ellipsometer and is slightly greater than that used in the section in which diaphragm thickness is measured in transmission with a laser. For the ellipsometer, this has little effect.



n_f	t_{per} (nm)
1.40	270
1.45	257
1.50	245
1.55	235
1.60	225
2.02	169

Fig. 4-27 Same as Fig. 4-26 except at $\theta = 50°$.

Ellipsometer inaccuracy in measuring a near-period thickness film is illustrated in Fig. 4-28, where a LPCVD deposited SiON film is repeatedly cycled through a 20 sec etch, water rinse, and ellipsometer measurement. A <100> substrate was coated with SiON film, and after being annealed to simulate the thermal experience of the LPCVD SiO_2 deposition, it was etched for 20 sec durations in 6:1 by volume of 49% HF:40% NH_4F. Period thickness locations, estimated by interpolating measured values of Δ to those of bare silicon, are indicated by arrows. The sensitivity to slight errors in ψ and Δ near the period thickness is apparent by the large error bar span there. A similar result was obtained when the experiment was repeated for thermally oxidized substrates. An enhanced sensitivity to the assumed values of substrate refractive index and extinction coefficient also exist at near period thicknesses. To compute the error bar span from the measured ψ and Δ, a computer program was written, because the ellipsometer software had no provision for finding n_f and t_f for values of ψ and Δ other than measured. When

run with the measured ψ and Δ, the program returned the same n_f and t_f as the ellipsometer software.

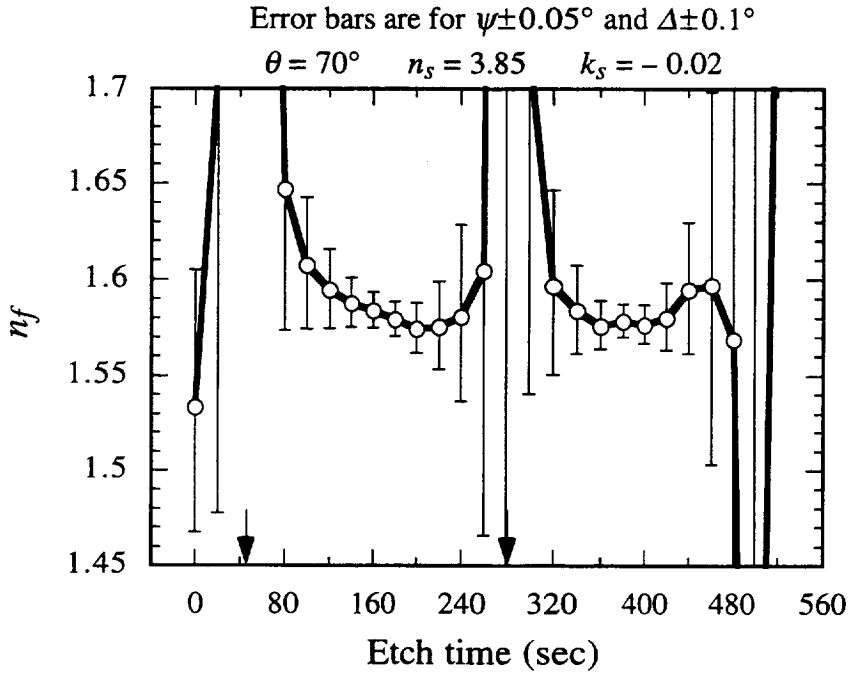


Fig. 4-28 Refractive index of annealed SiON film measured with a single angle and wavelength ellipsometer as a function of BHF etch time. The error bars bracket the refractive index limits found by assuming errors in measured ψ of –0.05°, 0°, and +0.05°, and in Δ of –0.1°, 0°, and +0.1°, in all nine possible combinations. Period thicknesses are indicated with arrows. The 540 sec of etching removed the film from the wafer.

In the program written to find n_f and t_f from ψ and Δ, in order to avoid having to numerically simultaneously solve for n_f and t_f, the ratio R_p/R_s, which is known from measured ψ and Δ values, was written [18]

$$\tan(\psi)\, e^{j\Delta} = \frac{R_p}{R_s} = \frac{r_{12p} + r_{23p} e^{-2j\beta}}{1 + r_{12p} r_{23p} e^{-2j\beta}} \frac{1 + r_{12s} r_{23s} e^{-2j\beta}}{r_{12s} + r_{23s} e^{-2j\beta}} \quad , \tag{4.11.1}$$

where $\beta = 2\pi \frac{t_f n_f}{\lambda_0} \cos(\theta_f)$, and r represents the Fresnel reflection coefficients at the air-film (r_{12}) and film-substrate (r_{23}) interface. The complex quantity $X = e^{-2j\beta}$ can also be further defined. For transparent films, X has unity magnitude. In (4.11.1), t_f appears only

in β and X and not in any r. To find n_f and t_f, substitute $X = e^{-2j\beta}$ into (4.11.1), multiply (4.11.1) by the denominator of its right hand side to form a quadratic in X, and solve for X using the quadratic formula. Then iterate n_f to make the magnitude of X become 1. Since the quadratic formula will find two values for X, a strategy is to iterate on one and if that root will not converge to magnitude 1, then on the other. Once n_f and X are determined, t_f follows without iteration from the definition of X. In the program, if the iteration which used Newton's method diverged, a set of $|X| - 1$ values using both solutions to the quadratic formula were generated by incrementing n_f over an interval and the zero crossings investigated with the bisection method.

12 PRISM COUPLING

Prism coupling can directly measure the waveguide effective refractive index. Its measurement is more accurate than that derived from an n_e calculation using ellipsometrically determined film index and thickness. If the waveguide supports two or more modes, then their n_e's can be used to determine the film index and thickness [19]. If the waveguide is single mode, it is possible to determine the film index and thickness from the TE and TM n_e, so long as there is no film birefringence. This latter use is less accurate than the former.

Our setup, shown in Fig. 4-29, allowed the substrate with clamped prism to be tilted using a machinist's rotary index table in order to achieve synchronism between the phase of light in the prism and that of the guided mode. The angle θ of the stage which maximized coupling can be used to determine the guided mode n_e. To make such a measurement the beam was made parallel to the optics table by adjusting laser tilt so that the distances h_1 and h_2 were the same. A dial indicator attached to a base which could be slid on the optics table was used to make the wafer table parallel to the optics table. With the dial indicator's plunger resting on the wafer table, the rotary stage was adjusted so that the dial indicator reading did not change as the sliding base was slid on the optics

table parallel to the beam. The position of the rotary stage where this occurred was taken as zero. Then the rotary stage was adjusted to maximize coupling into the waveguide. To enable coupling, the clamping force is initially high. Once coupled, the clamping force is reduced to narrow the range of angles over which coupling occurs [19], thus making angle determination more accurate.

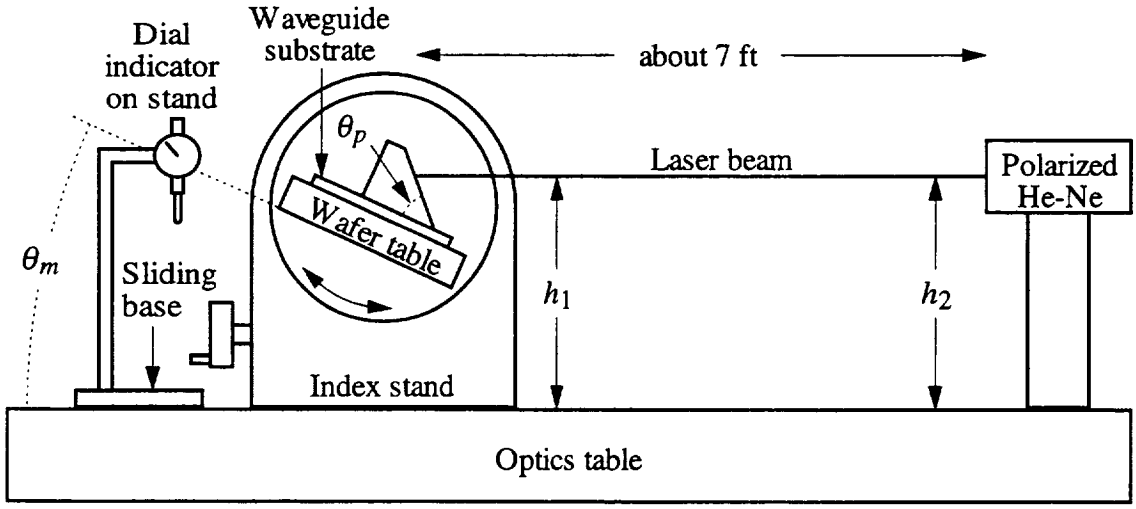


Fig. 4-29 Prism coupling setup.

To achieve coupling, phase matching between light in the prism and guide is required. Referring to Fig. 4-30, this will occur for beam propagation directions in the prism which cause its projected phase front separation on the base of the prism to be separated by the same amount as the phase fronts in the waveguide. Including the refraction of light at the prism air-interface, the guide's n_e is given by

$$n_e = n_p \sin\left\{\theta_p + \sin^{-1}\left[\frac{\cos(\theta_p + \theta_m)}{n_p}\right]\right\} , \qquad (4.12.1)$$

where n_p is the prism refractive index and θ_p and θ_m are defined in Fig. 4-30. Use of this formula requires an accurate knowledge of the prism refractive index n_p and apex angle θ_p. The latter can be obtained by attaching the prism to a rotary stage, and monitoring the angular rotation required to retroreflect a laser beam off the prism base and then off the

hypotenuse. The prism refractive index can be obtained from the method of minimum deviation [20].

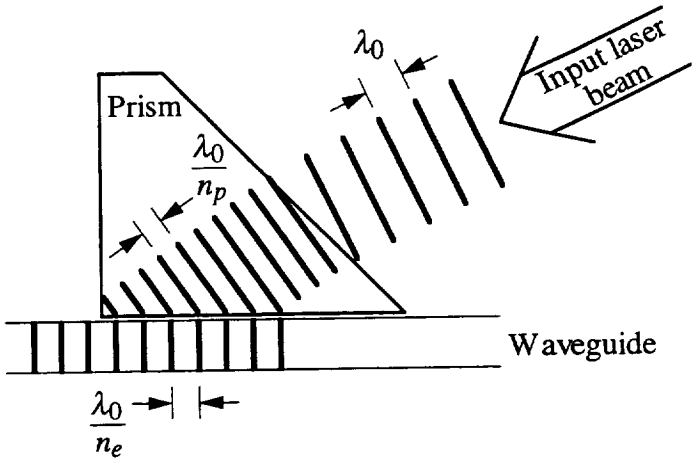


Fig. 4-30 Coupling from prism to waveguide. The thick lines represent phase fronts.

In this work, a Schott SF6 glass prism having a measured $\theta_p = 44.672°$ and $n_p = 1.80036$ (at $\lambda_0 = 0.6328$ μm) was used. For a typical SiON waveguide consisting of ~0.5 μm of SiON film having a refractive index of 1.53, resolving the coupling angle θ_m to within 0.01° means that the waveguide n_e is resolved to within ~0.00009. Prism coupling is the best method to determine the film effective index, and as can be seen from its sensitivity, it is more accurate than calculating n_e from ellipsometrically measured values. Channel confinement can be determined from the difference between the n_e in etched and unetched planar guiding areas.

13 MASKS

Fig. 4-31 shows the clear field channel waveguide mask designed for this project. The central 73,195 μm × 30,000 μm of a 4 in × 4 in plate was written with this pattern. It contains various types of test structures, in addition to "devices" such as ring resonators and Mach-Zehnder interferometers. Since this was our first integrated optic mask, many test structures were included as a diagnostic in the event that the devices didn't work as

expected. The radius of curvature of arcs incorporated into structures throughout the mask was 1000 μm, 3500 μm, 8000 μm, and 60,000 μm for the 1.4 μm, 2.4 μm, 3.4 μm and 5.4 μm channels. Referring to Fig. 31, the various sections are:

- 1) *Y*-junction test structures: 1.4 μm wide channels having 0.5, 1, 2, 5, 8, 12, 16, and 22° branch half angles; and 2.4 μm wide channels having 0.5, 1, 2, 5, 8, and 11.8° branch half angles. These branches begin after the channel tapers to double its width at a taper angle of 1°. The straight branch arms continue until they tangentially intersect an arc, which after bending through 1°, connects to two straight guides having 150 μm spacing. Also included are 1.4 and 2.4 μm 1° *Y*-junctions that abruptly join the channel without the 1° taper. Straight guides adjacent to the *Y*-junctions may enable subtracting straight guide loss from *Y*-junction loss.
- 2) Electro-optic switch: 2 and 3 μm channels formed into Mach-Zehnders with a separate channel for another laser feed-in. This is for another project involving nonlinear polymer films. A weak laser's emergence from the Mach-Zehnder should be modified by a bright laser's influence on the refractive index of one arm of the interferometer.
- 3) Electron beam quantization-induced propagation loss test structure: This test structure is formed of three lines, the middle one at a slight angle to the electron beam pattern generator stage axis. The pattern generator will write the middle line as a series of short lines parallel to the stage axis with sideways steps equal to the quantization size. The propagation loss due to these steps can be measured.
- 4) Curvature loss test structure: When light goes around a waveguide bend, the field peak shifts away from the center of curvature. The shift makes the field tail on the inside of the curve see a higher refractive index and slow down, maintaining a constant phase velocity across the phase front. When two curves meet, their field peaks are transversely displaced, and a loss occurs. At discontinuities in curvature, a

slight offset is interposed in these structures. The offset, not visible in the drawing, is different for each structure.

- 5) Bending loss test structures: A series of 180° arcs made with 1.4, 2.4, and 3.4 μm wide guides will allow comparison of calculated and measured bending loss. The bend radii goes from: 100 to 2000 μm for 1.4 μm guides; 100 to 4500 μm for 2.4 μm guides; and 200 to 8000 μm for 3.4 μm guides.
- 6) Corner mirror test structures: A series of 5.4 μm guide corner mirror test structures is provided. To fabricate these, the guide layer is printed and channels etched. Then, using reactive ion etching, another part of the this mask is employed to make the corner mirrors (a "dark field area"). Alignment marks allow this to be accurately accomplished.
- 7) Underneath the corner mirror test structures appear "box" resonators: Similar to ring resonators, they employ corner mirrors instead of waveguide bends. These devices use 5.4 μm guides and directional couplers. When the corner mirrors are aligned to the abrupt bends in structure 6, the corner mirrors of the box resonators will be aligned to the boxes. Three devices having different coupling strengths appear.
- 8) Waveguide stripper: If stray planar guiding light turns out to be a problem, then the Fresnel lens disperser may help. The total channel loss due to the prisms is conservatively estimated to be less than 1%.
- 9) First directional coupler set: Three of these 5.4 μm channel couplers are identical to those used in the box resonators. They are a check on the theory and a diagnostic if the box resonators don't work. Another edge alignment mark appears after them.
- 10) *Y*-junction ring resonator: A 1.4 μm channel and two symmetric *Y*-junctions are used to construct a ring resonator. The radius of curvature of the bends is 1000 μm. The physical path length of the ring is 13,283.4 μm.

- 11) *Y*-junction ring resonator: A 2.4 μm channel and two symmetric *Y*-junctions (half angle of 1°) are used to construct a ring resonator. The radius of curvature of the bends is 3500 μm. The physical path length of the ring is 28,991.3 μm. Two alignment marks are provided in the vicinity; one for a fiber groove mask and the other to allow alignment to back surface diaphragms via a square hole etched all the way through the wafer (an alternative back surface alignment method).
- 12) Directional coupler ring resonator: A 2.4 μm channel and a directional coupler are used to construct a ring resonator. The radius of curvature of the bends is 3500 μm. The physical path length of the ring is 28,991.1 μm. Unlike the other ring resonators, one of the feed channels (on the right) arcs away from the in-line position of the other so that stray planar guiding light propagating from the input coupling edge is avoided when measuring device transmissivity.
- 13) Second directional coupler set: Directional couplers are provided which have tighter and looser coupling than in the ring resonator. All employ 2.4 μm channels except a 1.4 μm one which mimics the *Y*-junction of the 1.4 μm ring resonator.
- 14) Nine Mach-Zehnders: Two balanced ones having different arm separations are provided. These Mach-Zehnder interferometers have physical path length differences of 0, 33.33, 100, 1040, 2080, and 4160 μm. A pair of 100 μm Mach-Zehnder interferometers in series with which one could do path matched differential interferometry is provided. All use 2.4 μm channels. Another pair of alignment marks similar to those near the *Y*-junction ring resonator are provided.
- 15) Edge alignment mark: Registration between the diaphragms and channels could be achieved by positioning this mark over a cleaved edge to which the back surface mask had been previously aligned. The other direction may be aligned with the aid of another cleaved edge and the opening edge in the chrome mask perpendicular to the dashed lines.

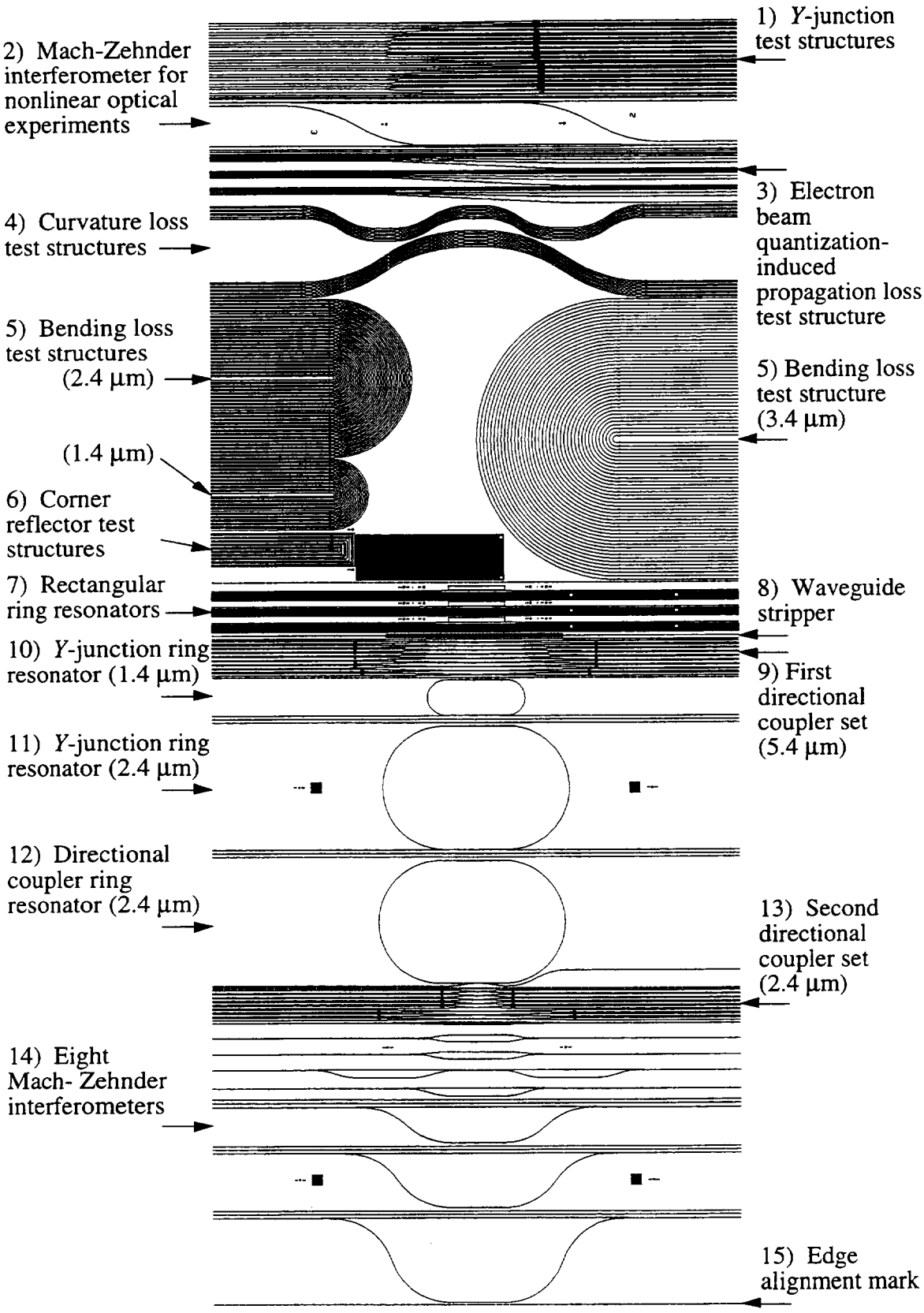
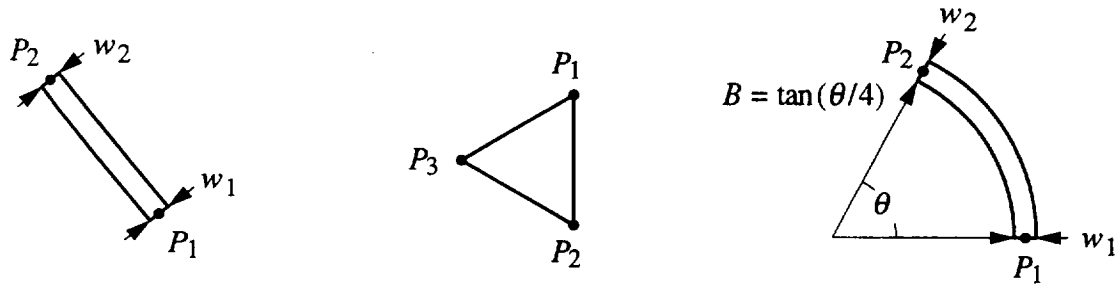


Fig. 4-31 Channel waveguide mask, facing glass side of plate.

The mask was drawn up in Autocad release 10, and the DXF (drawing interchange file) this program produced mailed to Microphase of Colorado Springs, Colorado, a commercial mask fabrication house, who wrote the mask using a 0.1 μm diameter electron beam. Although Autocad is a general purpose drawing program which can create many different types of structures including circles, donuts, ellipses, 3-*D* meshes, etc., only nonzero width open polylines, zero width closed polylines, and nonzero width open polyarcs were employed (Fig. 4-32).



Open nonzero width polyline specified by vertex points P_1 and P_2, and beginning and ending widths w_1 and w_2.

Closed zero width polyline specified by vertex points P_1, P_2, and P_3. The vertex points of each segment must precisely connect.

Open nonzero width polyarc specified by vertex points P_1, and P_2, beginning and ending widths w_1 and w_2, and bulge B.

Fig. 4-32 The three types of entities used in the original mask. In the polyarc, the bulge is positive if the arc is drawn counterclockwise.

Microphase translated the DXF data to GDSII, a popular format inputted by many programs which generate machine-readable mask generator instructions. Subsequently, GDSII was converted to ETEC, which contained the sorted raster data. Unfortunately, the GDSII language does not contain a descriptor for arcs, but for polygons, to which the polyarcs were converted. In the mask received from Microphase, each polyarc was represented by a polygon with segments having a departure angle of 1.9°. Under a microscope, these abrupt direction changes were most visible when the stage was moved at a constant speed while viewing a line. Fig. 4-33 shows a close up view of the 180°, 2.4 μm wide channel bending loss test structures (#5 in Fig. 4-31). The channels are in

100 μm radius of curvature increments. Holding the page at an oblique angle and sighting down the lines may make the abrupt discontinuities in slope more evident. Fig. 4-34 shows the scattered light from a channel made with this mask. Such bends are useless because of their high loss.

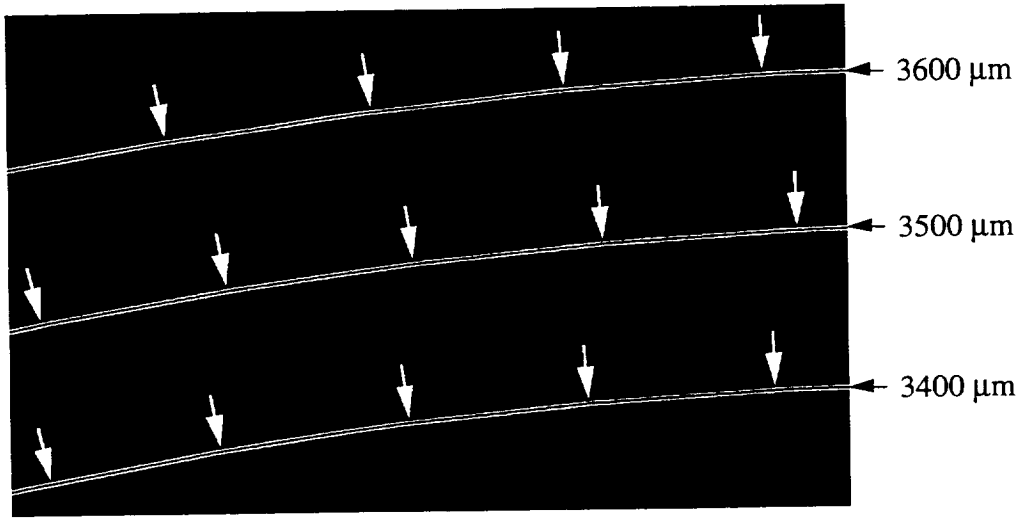


Fig. 4-33 Photo of section of 180°, 2.4 μm bends on mask. Abrupt changes in direction are indicated with arrows.

Fig. 4-34 Radiation from a 2.4 μm wide channel having a 4300 μm radius of curvature bend. The guide was constructed from 100 nm of Si_3N_4 capped with 185 nm of LPCVD SiO_2, all on an oxidized silicon wafer. To form the channel, the SiO_2 film surrounding it was etched to a thickness of 100 nm.

Since, in further attempts, Microphase was unable to reduce the departure angle, we wrote a program in Qbasic which would read in the DXF file and construct a new one, converting each polyarc into overlapping, nonjoined, straight, polylines, as well as disconnecting and overlapping the ends of straight, joined, polylines. The maximum theoretical error ε (Fig. 4-35) between a perfect arc and the end of our polyline was set to 0.04 μm, less than half the e-beam quantization size. Such an error should contribute negligibly to the loss, even though it results in a departure angle of 0.5° (Fig. 4-36) for the 3500 μm radius of curvature arc [21]. After the coordinates of polylines replacing a polyarc had been found and rounded to the nearest 0.1 μm, their endpoints were adjusted in 0.1 μm increments about their rounded value to minimize the offset between two

adjacent polylines. Approximately 75,000 polylines were needed to construct the mask, and the file's size was 16 MBytes.

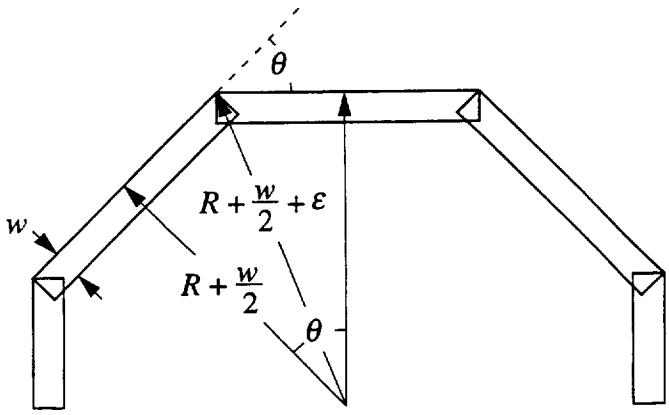


Fig. 4-35 Error criteria for selecting the number of polylines replacing a polyarc.

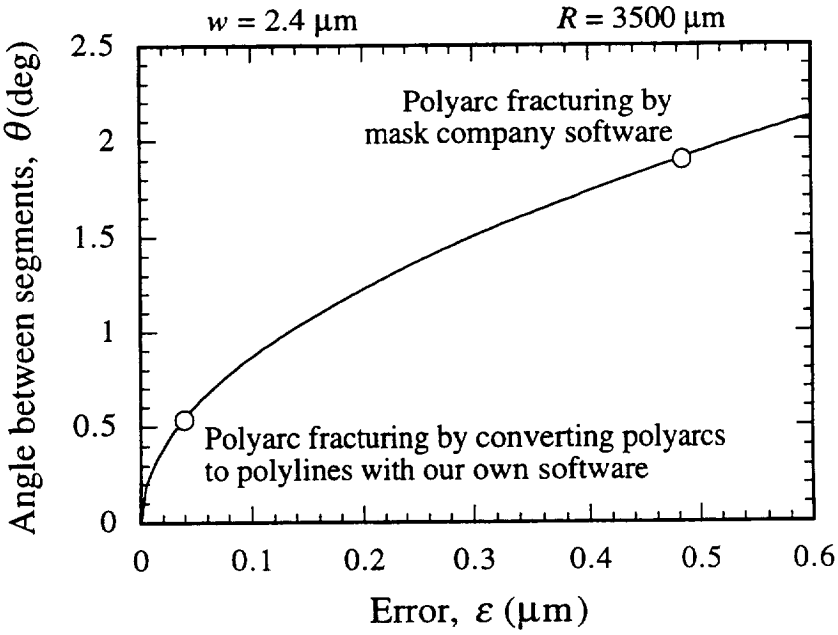


Fig. 4-36 "Error", as defined in Fig. 4-35, between a perfect polyarc and one composed of polyline segments.

Microphase generated another mask with the new DXF file. Abrupt changes in direction were not visible, nor were the radiated beams of Fig. 4-34. This mask was used to fabricate the ring resonator and Mach-Zehnder in the results section.

A dark field companion mask for opening windows in the masking oxide on the wafer's back surface to form diaphragms was also designed, and is shown in Fig. 4-37. This mask, which was also drawn up in Autocad, was fabricated here in the Department of Electrical Engineering at the University of Cincinnati using the TRE optical pattern generator made by Electromask. This unit could not have been used for the waveguide mask since the boxes it flashes can be no smaller than 5 μm. In addition, patterns drawn in Autocad translate to clear openings in the chrome mask, making it difficult to produce complex light field masks (one must draw the areas not occupied). In addition to diaphragm openings, this mask also contains alignment marks coinciding with those on the waveguide mask.

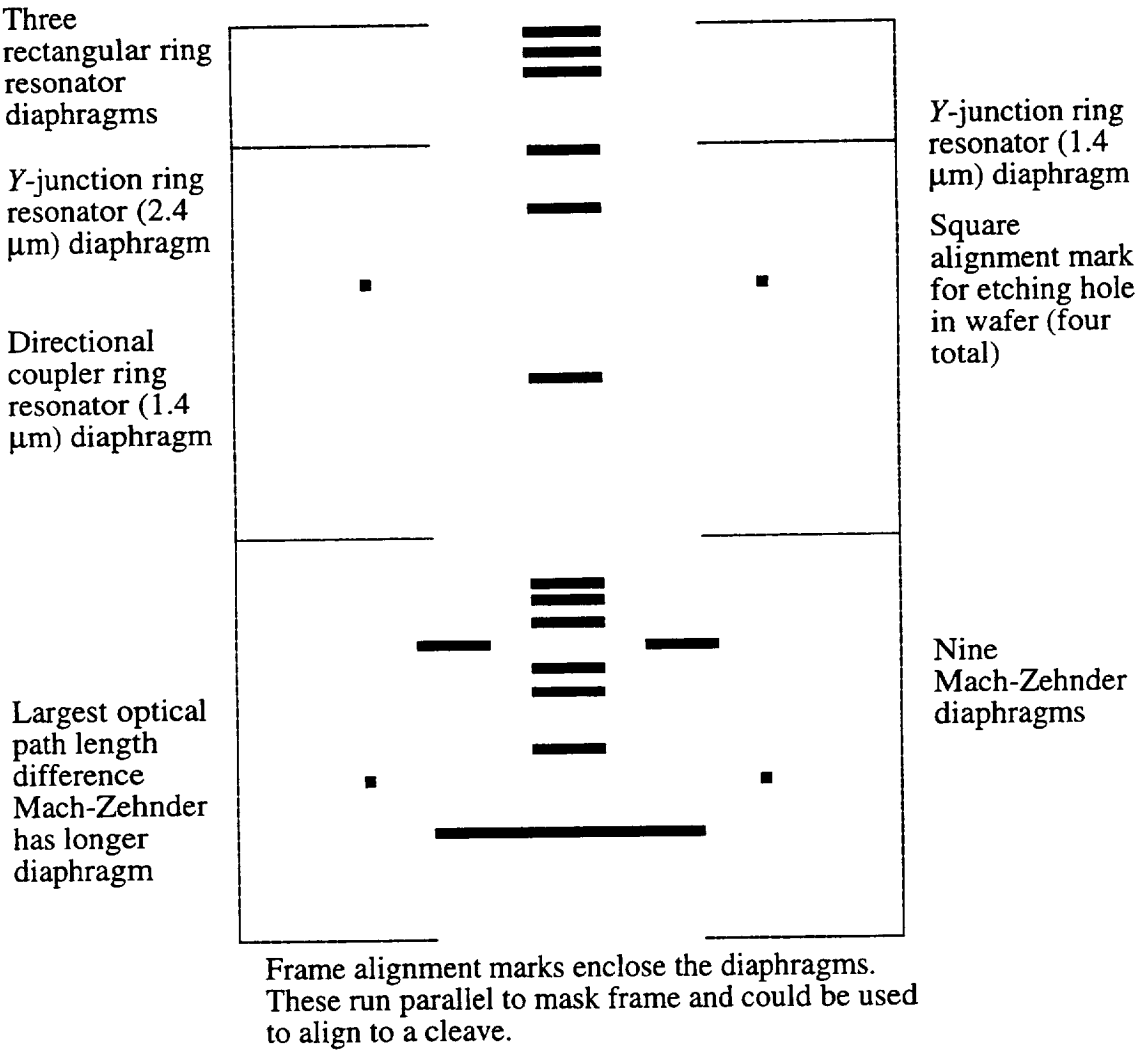


Fig. 4-37 Back surface mask facing either the glass or chrome side (it is symmetric). Black areas are openings in chrome. The large boxes were used for diaphragm formation and all except the long one at the bottom and three at the top were 3305.4 × 430 μm. The square boxes and the small connected boxes which frame the large rectangular boxes were to be used for alignment.

14 FABRICATION SEQUENCE

In the course of this work two full fabrication runs were performed. The first yielded the ring resonator and the second a Mach-Zehnder. Each device was attached to a chuck and tested. The following gives a process sequence which evolved after fabricating these devices.

The fabrication sequence for forming the pressure sensor consisted of forming the diaphragms, depositing the waveguide films, and then patterning the channels. Both lithographies were performed on a Karl Suss MJB 3 contact printing mask aligner. In this unit, the wafer is held to a stainless steel chuck with vacuum and its position adjusted with two differential micrometers. An additional micrometer produced a rotational adjustment about the chuck center (theta), and differing wafer and mask thicknesses could be compensated by a chuck height adjustment. The mask was held by vacuum to a plate above the wafer and could positioned over a range of locations on the plate. This capability was exploited in the design of the 4" waveguide mask, which had a wider written width (7.3195 cm) than the wafer's diameter (5.08 cm), to enable printing, in separate exposures, its entirety.

The aligner had three modes of exposure. The gentlest one consisted of simply bringing the lamp house over the mask and opening a shutter. In another mode, N_2 was blown through vacuum hold down holes beneath the wafer during exposure. The hard contact mode consisted of pulling vacuum between the wafer chuck and mask, which were sealed together with a gasket on the chuck. This pulls the mask against wafer with almost atmospheric pressure and results in the least separation and best resolution. However, it can damage the mask and resist on the wafer. Hard contact was always used for waveguide exposure, while one of the other two was used with the diaphragm mask exposure. Although an infrared aligner was available and could have been used to align the waveguide pattern to the diaphragms, its inability to do hard contact ruled it out.

To reduce the problem of resist sticking to the mask and ease mask cleaning, both masks were spin coated with Panax [22], which leaves a monolayer nonstick coating. The masks were thoroughly cleaned, given a brief dip in dilute HCl, baked, spin coated with Panax, baked again, and the excess Panax removed using a cleanroom wiper and acetone. The treatment, which lasted for ~20 prints, made the surface hydrophobic.

Diaphragm fabrication sequence

The fabrication sequence that follows describes two methods of etching into one side to fabricate diaphragms and then explains how channel waveguides are made on the opposite face aligned to them. An outline of the diaphragm fabrication sequence is given in Table 4-3; wafers thus processed are then sequenced through waveguide formation and chuck attachment.

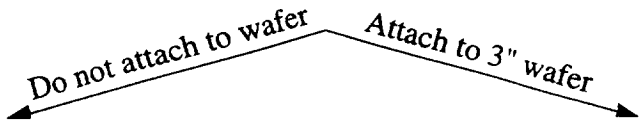
1) Start with <100> double side polished, 10 mil thick, <110> major flat, silicon wafer.	
2) Identify better surface using DIC microscope and, on it, scribe identification number near flat.	
3) Grow 3 μm thick steam oxide at 1050°C for 24 hr.	
4) Measure the oxide thickness with the ellipsometer.	
5) Measure wafer thickness with micrometer.	
6) Calculate the mask pattern displacement.	
7) Base clean.	
8) Prebake at 200°C for 5 hr.	
Alternative sequences: 	
9a) Spin resist on back, bake 95°C for 30 min.	9b) Attach front face to bare 3" wafer using photoresist.
10a) Align to edge and flat and expose.	10b) Bake 120°C for 1 hr.
11a) Partially develop resist to see imprint.	11b) Spin resist, bake 95°C for 30 min.
12a) Re-align to imprint, offset wafer and expose.	12b) Align to edge and flat and expose.
13a) Fully develop resist.	13b) Partially develop resist to see imprint.
14a) Bake 120°C for 1 hr.	14b) Re-align to imprint, offset wafer and expose.
15a) Lay wafer on flat surface and puddle BHF in openings in resist. Rinse in deionized water	15b) Fully develop resist.
16a) Rinse in acetone to strip resist. → Go to step 19.	16b) Bake 120°C for 1 hr.
	17b) Immerse in BHF until windows in oxide open. Rinse in deionized water.
	18b) Soak in acetone overnight to separate wafers.
19) Etch in 40°C KOH, (Transene PSE-200) until diaphragms remain. Silicon etch rate is 4.35 μm/hr. Rotate wafer 90° every ~12 hr. Measure diaphragm thickness using microscope focusing technique until diaphragms are within 4 μm of target thickness. Afterward, measure diaphragm thickness using laser.	
20) Strip oxide in BHF.	

Table 4-3 Diaphragm fabrication sequence summary.

1) Start with <100>, double side polished, nominal 10 mil thick silicon wafers having a <110> flat. The doping and resistivity are irrelevant except that doping should not be so high that etching is impeded. The wafer on which the Mach-Zehnder was

fabricated was 40-60 Ω-cm. When purchasing the wafers, specify the flat accuracy and total thickness variation as tightly as possible. None of the wafers used in this work employed etch stopping layers and it was found that thickness differences between diaphragms on the same wafer were due mostly to variation in wafer thickness rather than etch rate nonuniformity. The batch of wafers used for Mach-Zehnder fabrication had a specified total thickness variation of 0.0005 mils, or 13 μm, and was obtained from Virginia Semiconductor of Fredericksburg, Virginia. These wafers typically had a thickness uniformity of ~4 μm over the area used for diaphragms. Employing etch stop layers would of course decrease the diaphragm-to-diaphragm thickness variation. The wafer major flat was found to be aligned to the crystal planes to better than 0.5°.

2) Identify the better face of the wafer using a differential interference contrast microscope (DIC), and near the flat, scribe an identifying number on this face. This face will be called the waveguide or front surface and the other the back surface.

3) Steam oxidize the wafer at 1050°C for ~24 hours to grow 2.9 μm of oxide. No dry O_2 oxidations are needed Although this thickness is more than enough for masking, as only 0.5 μm of SiO_2 would be consumed to etch the diaphragms, the additional thickness affords scratch resistance.

4) Measure the oxide thickness with the ellipsometer, using either the double angle program, or with the single angle program at different angles with the index fixed to 1.46. Using a single angle and wavelength, the ellipsometrically measured thickness is ambiguous: it can only be known if the period order has been established. For SiO_2 with $n = 1.46$ at $\lambda_0 = 0.6328\,\mu m$ and at a 70° angle of incidence, the period is 0.283 μm; at 50° it is 0.255 μm. Thus at 70°, possible thicknesses could be $t = t_f + m\,(0.283\,\mu m)$, where t_f is the ellipsometrically measured fundamental thickness, which is angle of incidence dependent, and m is

an integer representing the period order. By measuring t_f at different angles of incidence and matching nearest t between the two angles, t can be determined to within a super period of a few microns. However a difference between the assumed or measured and actual refractive index, upon which the period thickness depends, can cause an incorrect m to be chosen which will result in an incorrect t. The double angle program furnished with the ellipsometer selected the film refractive index and period thickness either from one measured at one of the angles, or if both measurements were close to period thickness, it defaulted to 1.46 for SiO_2. If both angles were somewhat near period thickness, it might measure an index as high as 1.49 (see the section on ellipsometry in this chapter). In this case, a better approach is to use the single angle program at 50° and 70° and fix the index to 1.46. This tells the ellipsometer to disregard the measured ψ and use n_{fix} and Δ to determine t and the period thickness. This is effective near period thickness where curves of constant refractive index converge.

Alternative techniques to ellipsometry were also tried. One of these involved measuring absolute reflectance of s-polarized HeNe laser light as a function of angle of incidence, assuming a refractive index, and curve fitting to obtain the thickness. Another technique was to assume a film refractive index and measure reflectance as a function of wavelength; the thickness could be determined from the separation of reflectance minima. These techniques seemed more reliable, especially when the film thickness could range over 3~4 μm. However, they took significantly more time and required a dedicated setup.

5) Using a micrometer, measure wafer thickness at the places indicated in Fig. 4-38. The measurements should resolve to 0.0001" or better. These were typically made with a Mitutoyo Digimatic micrometer which resolved to 1 μm. The silicon surface will be more protected from scratches if this step is performed after oxidation.

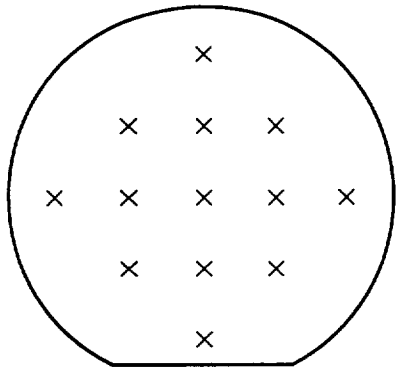


Fig. 4-38 Places where wafer thickness was measured.

6) Calculate the mask offset displacement to form diaphragms having a short dimension of 250 μm from an average of the wafer thicknesses measured in step 5. The back surface mask contains boxes which are, except for the rectangular ring resonators and the largest path difference Mach-Zehnder, 3305.4 μm × 430 μm. The mask opening is sized so that a wafer having a diaphragm well depth of 8.5 mils will have a finished diaphragm long dimension of 3000 μm. The diaphragm short dimension, upon which the pressure sensitivity is more dependent than the long, is established by two exposures which are displaced from each other. The first exposure and development creates openings in back surface resist, while the second, made in the same resist and overlapping the first, enlarges the short dimension of the rectangular opening. This insures that the diaphragm's short dimension will be approximately 250 μm, regardless of the wafer thickness. The displacement between the two exposed regions is calculated from (4.14.1) using Fig. 4-39.

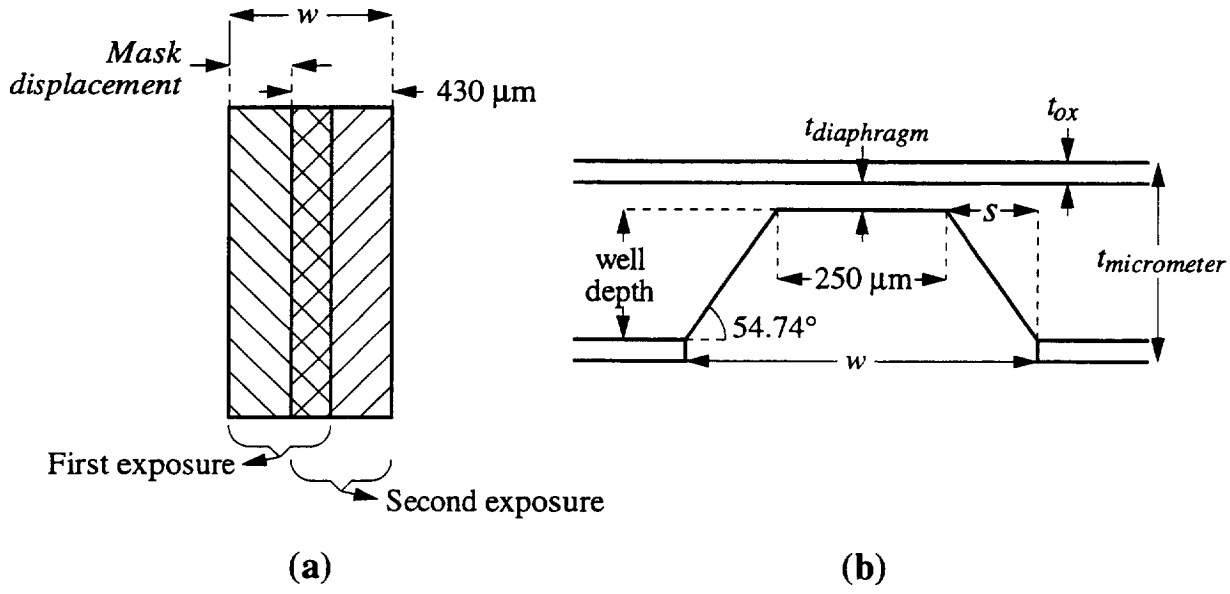


Fig. 4-39 (a) Arrangement of the exposures. (b) Dimensioned diaphragm.

$$Mask\ displacement = w - 430\ \mu m$$

$$w = 2s + 250\ \mu m$$

$$Mask\ displacement = \sqrt{2}\ (t_{micrometer} - 2t_{ox} - t_{diaphragm}) + 250\mu m - 430\ \mu m\ . \qquad (4.14.1)$$

7) Base clean the wafer to remove contamination left by the micrometer. Here "base clean" refers to the RCA base solution [23] prepared by heating deionized water to ~80°C, and then adding ammonium hydroxide and hydrogen peroxide to create the proportions of 5:1:1 by volume of deionized H_2O:30% NH_4OH:30% H_2O_2. The temperature of the solution should be maintained at approximately 70°C for a clean duration of 10 - 15 min. This clean is effective at removing trace organics.

8) Bake 200°C for ~5 hr or more. After a base clean, SiO_2 appears hydrophilic; in contrast, after growth, the oxide is weakly hydrophobic. It was found that buffered HF will severely undercut positive photoresist applied to wafers after a base clean without prior prolonged, high temperature baking, but will adhere well to an oxidized wafer freshly pulled from the furnace. The use of the adhesion promoter

HMDS prior to resist application on unbaked, base cleaned SiO_2 was not found to improve resist adhesion. While a bake of 120°C overnight was found insufficient, one at 200°C for ~5 hr. was satisfactory.

Several methods may be employed to prevent etching the front surface oxide when etching windows in the back. The front can be protected by coating it with resist or by attaching this face with resist to another wafer. Alternatively, the face can be left unprotected and windows opened in the back SiO_2 by puddling BHF in the resist openings using a dropper. All these methods were successfully used, but puddling the etchant was found to be the simplest and resulted in the fewest defects on the front face. This method cannot be used to pattern channel waveguides, where partial etching is required, as the etch rate could be nonuniform, and it could not be used if the area to be etched extended to the wafer edge. Attaching the 2" wafer to a bare 3" wafer using resist and etching by immersion was found to leave the front oxide defect free except for an occasional etched spot in the front oxide near the wafer edge. In addition, the wafer pair could be handled with tweezers without damaging resist anywhere on the 2" wafer. This method was used to strip films from the wafer back prior to electrostatic bonding, while protecting the waveguide films. Separating the two wafers required an overnight soak in acetone. Alternatively, spin coating both sides required making spinner chuck contact to a resist coated face, which was found to damage the resist. The front could be painted with resist, but the front oxide was found to have more defects after immersion etching than when adhered to a 3" wafer. Both of the better methods, outlined in Table 4-3 will now be discussed.

<u>Sequence in which the 2" wafer is not attached to a 3" one (puddle BHF etch)</u>

9a) Spin HMDS and positive resist on the back of the wafer. Use a resist which produces a thick film (2-3 μm), as this will reduce particle and void problems. Since the openings in the SiO_2 are large, fine feature imaging is not required.

When, a thick resist was not available Shipley 1813, which was used for waveguide patterning and spins to 1.2 μm at 5000 RPM, was substituted. To reduce defects such as comet tails, the resist was syringe filtered onto the substrate using a ground glass syringe with a 0.2 μm pore Gelman PTFE filter.

10a) Align the back mask to the wafer flat and edges. Before this can be done the orientation of the diaphragms on the wafer must be determined. The back surface mask places diaphragms in a line through the center of the wafer. In the mask drawing in Fig. 4-37, this direction is parallel to the left side of the page. To maximize diaphragm thickness uniformity, the wafer should be oriented so that this line runs along the more uniform wafer direction, as ascertained in step 5, either perpendicular or parallel to the flat.

The diaphragm mask had no alignment marks to allow it to be aligned to the flat directly. Instead the mask theta was adjusted to make the mask frame (not plate edge) parallel to the aligner microscope stage travel. The mask was placed on the vacuum plate mask holder and the aligner microscope was focused on a feature parallel to the mask frame. The alignment marks which framed the diaphragms were used for this purpose. The mask aligner microscope was scanned along this feature by releasing one of its brakes, and a determination was made of the direction of rotation required to make the mask parallel to the microscope travel. The mask holder plate was removed from the aligner, and the mask nudged to adjust its rotation. The holder was reinserted into the aligner and the mask's parallelism to the microscope stage reassessed. This was continued until the mask was satisfactorily oriented.

Next the wafer was rotated until its flat was parallel to the microscope stage translation, thus adjusting the wafer theta. To accomplish this, the mask holding plate was removed (the mask stayed attached and aligned) and the wafer and its

translation stage adjusted using the micrometers to move the flat toward an area accessible to the microscope.

With the mask holder back in the aligner, the wafer was positioned until its edge could be seen simultaneously through both corners of the alignment marks, as shown in Fig. 4-40(a). The wafer was then translated 2 mm, as shown in Fig. 4-40(b), by rotating the coarse wafer translation micrometer 4 turns. Positioning the diaphragms in this manner will result in all of the Mach-Zehnders and both 2.4 μm ring resonators being intact on the front face and testable after subsequent front mask alignment and processing. If other devices such as the rectangular ring resonators are to be tested, then a different offset adjustment will have to be used.

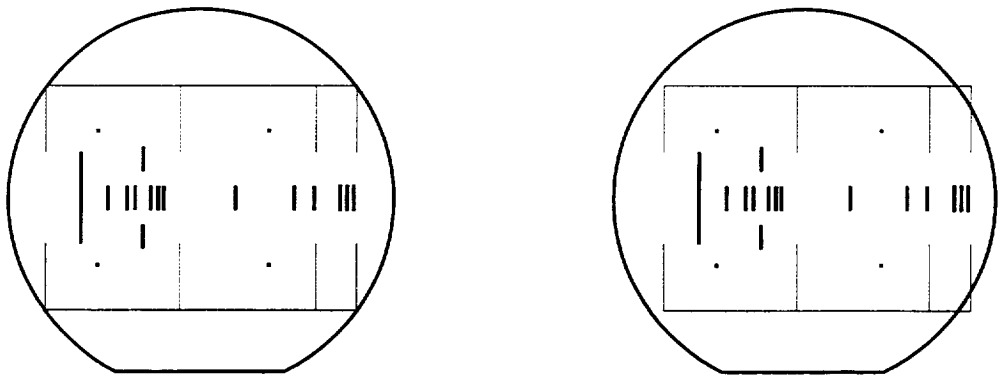


Fig. 4-40 (a) Initial alignment of mask to the wafer. The edge of the wafer should be visible simultaneously through the both corners of the alignment marks (Black areas are clear on the mask). (b) Final alignment. The wafer has been shifted to the left with respect to the mask by 2 mm. Shown is the case where the line of maximum uniformity is parallel to the flat.

A cardboard cutout was then laid on the glass side of the mask to cover the alignment marks and prevent their patterns from being subsequently etched into the substrate. The marks were far enough away from the diaphragm openings that the former were not exposed and the latter not affected.

11a) Perform the exposure and lightly develop the resist. The exposure dose is not critical, since the features are large. The developer was a solution of 5:1 by volume of deionized H_2O:Shipley 351 concentrate, which contained sodium hydroxide, and immersion development times of 5 sec were sufficient to see the exposed area. It was found that over developing made the resist more susceptible to undercutting in the buffered HF. After development the wafer was rinsed for ~1 min in deionized water, blown dry with N_2, and then placed back in the aligner.

12a) Realign to the partially developed pattern, and using the calibrated differential micrometers on the aligner, translate the wafer parallel to the short edge of the diaphragm by the amount calculated in step 6. The cardboard cutout was again in place during the exposure.

13a) Immersion develop fully, rinse, and inspect the resist for pinholes and the open areas for scumming or opaque defects. Reprocess if these are excessive.

14a) Hard bake the resist to improve its adhesion at 120°C for 1 hr.

15a) Place the wafer on a flat level surface and, using a plastic dropper, apply buffered HF to the openings in the resist. Buffered HF consists of 6:1 by volume of 40% NH_4F:49% HF and the steam oxide etch rate is 1000 Å/min. When the etch was presumed to be through the SiO_2, the wafer was rinsed in deionized water. If openings in the resist were not hydrophobic the etch was continued.

16a) Using acetone in a rinse bottle, strip resist. If the resist is difficult to remove, a more effective stripper such as hot Shipley 1165, which contains N-Methyl-2-Pyrrolidone, can be used.

<u>Sequence in which the 2" wafer is attached to a 3" one (immersion BHF etch)</u>

9b) Place one drop of photoresist in the center of the polished face of the 3" wafer and lay the 2" wafer centered on to the 3" wafer. If more than one wafer is being

processed, the 3" wafer should be marked to permit identification. Press lightly on the back face at the edge to seat the wafer.

10b) Bake the pair at 120°C for 1 hr. It was found that baking at 95°C for 30 min was insufficient to assure that the two would remain together during the spin coat.

11b) Spin resist onto the back of the 2" wafer. As mentioned in step 9a of the alternative sequence, the resist was typically syringe filtered onto the wafer. After the spin, additional resist was applied with a small paint brush to seal the interface between the two wafers.

12b)-16b) are the same as 10a)-14a).

17b) Immerse in BHF until the windows are hydrophobic. The solution is described in step 15a).

18b) To separate the wafers, soak the wafer pair in acetone overnight. Afterward, clean both wafers in fresh acetone or resist stripper.

19) Etch the wafers in 40°C KOH (PSE-200), until silicon diaphragms of the appropriate thickness remain. Every 12 hr, the wafers were rotated 90° to maintain uniformity. After being removed from the etchant, the wafers should be transferred to a deionized water rinse.

As the etch nears completion determine diaphragm thickness by taking the difference between the wafer thickness, measured using a micrometer, and the well depth, measured by focusing alternately on the diaphragm and wafer back with an optical microscope. When diaphragm thickness was ~4 μm greater than the target thickness, the laser transmission technique was invoked. The target thickness should take into account that the diaphragm will be reduced by 2×0.46 μm for every 1 μm of oxide grown on each side of the diaphragm when forming the waveguide isolation oxide, which is typically 3 μm thick.

20) Strip the oxide everywhere by immersing the wafer in a BHF solution until the surface is hydrophobic.

Aligning to the diaphragms

As the calculations in chapter 3 showed, maximum sensitivity to pressure results when a straight section of the channel in the interferometer is located over a long edge of the diaphragm. The sensitivity to pressure rapidly decreases as the channel is moved onto the portion of the wafer above the 54.74° sloping sidewall because substrate thickness increases rapidly there. The sensitivity to pressure does not fall off as rapidly as one moves away from the edge toward the diaphragm center. Therefore, when aligning the channels to the diaphragm edges, the channels were placed as close as possible to the edge while keeping the entire waveguide ridge on the diaphragm.

Waviness in the diaphragm edge was seen in varying degrees; a particularly bad case is shown in Fig. 4-41. This is believed due to misalignment between the masking oxide opening and the wafer's crystal planes. The misalignment is mostly due to the wafer flat departing from the <110> plane, rather than inaccuracy in mask to flat alignment. Such waviness results in a variable separation between the waveguide and the diaphragm edge. The waviness results in undulations in the sidewalls, which can also be seen in the Figure. A typical diaphragm would have ~3 μm peak-to-peak waviness although some had < 1 μm.

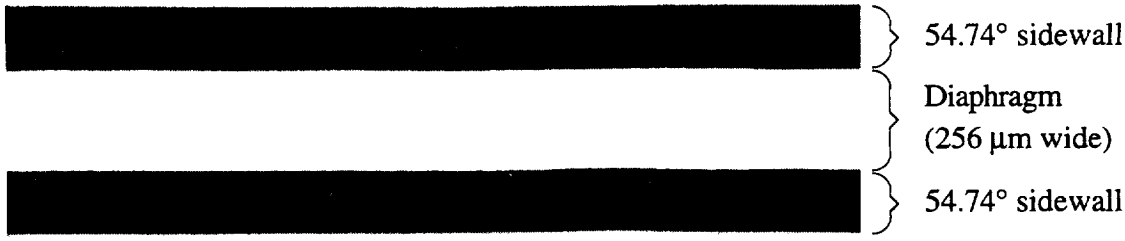
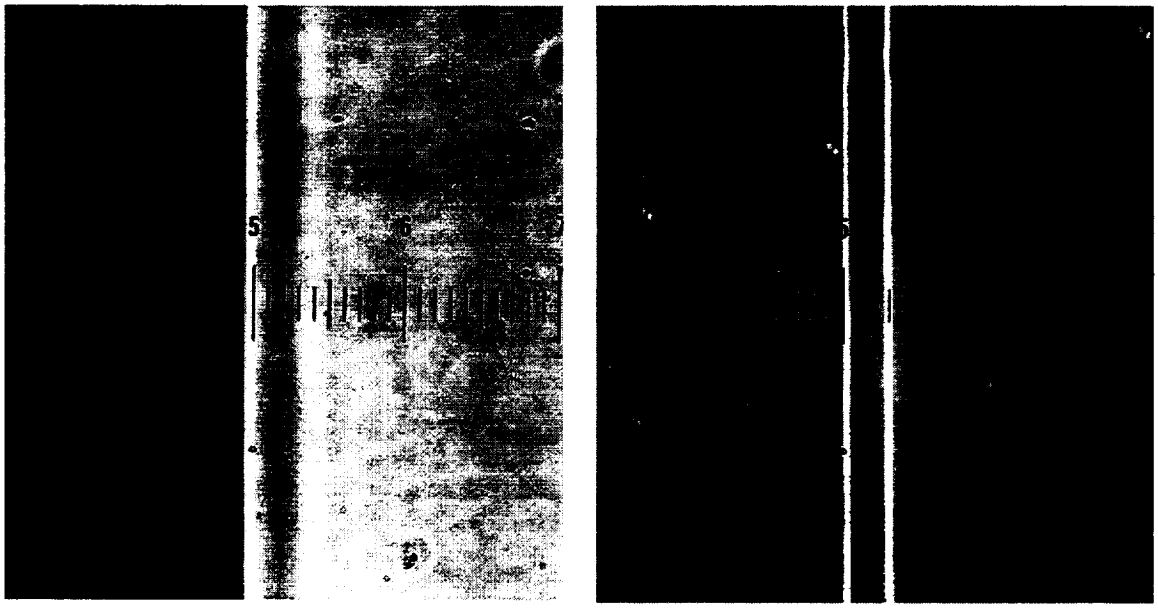


Fig. 4-41 Rectangular diaphragm showing approximately 9 μm peak-to-peak waviness.

Although both waveguide and diaphragm masks were designed with alignment marks which would allow them to be registered to two cleaved perpendicular edges or to two etched holes, the alignment marks were not used in this fashion. Instead, a chrome line on the mask, which could be used to form a channel waveguide over the diaphragm, was aligned to a diaphragm edge. For the first device made, the ring resonator, a diaphragm associated with an interferometer deemed to be less important was broken out with a pin and its waveguide segment aligned to the hole edge. Once done, all channels would be aligned to their respective diaphragms. Although with several attempts this was eventually accomplished, it was difficult because the broken area imprecisely defined the edge of the diaphragm. Additionally, the resist beaded on the edge of the hole closest to the center of the wafer. Resist could also end up on the wafer back and vacuum could be pulled through the hole; all these make alignment difficult.

Despite these problems, this technique can be used to achieve accurate front-back registration, as Fig. 4-42 shows. In the left half of the Figure, light was shone on the diaphragm back, and the microscope focused on the line made between the diaphragm back and the <111> sidewall. The two planes met at graticule 49, the edge of the diaphragm. Then the illumination was changed to come through the objective and the microscope was focused on the waveguide. The waveguide is seen to be on the diaphragm and the two edges are ~1 div (0.96 μm) apart.



(a) (b)

Fig. 4-42 Waveguide face view showing the alignment of the diaphragm edge and channel waveguide. The silicon component of the diaphragm was 5.8 μm thick. (a) The diaphragm back was illuminated, enabling the optical microscope to look through the diaphragm and focus on the its back at the edge. (b) Through the objective illumination was used and the microscope focused on the channel.

To permit use of this technique in the aligner, its stainless steel wafer chuck was copied in polycarbonate, with provision made for performing hard contact, and lamps placed inside. However, even when a lamp was directly under a diaphragm, there was insufficient light to determine its boundaries through the aligner microscope. Finally an adapter consisting of a tube with C-mount thread on one end and sleeve on the other was made to hold a TV camera collinear with the aligner's eyepiece socket. With the eyepiece removed, a real image was formed on the monochromatic Dage series 67 silicon target camera vidicon tube. Using the aligner's original stainless steel chuck and built-in through the objective illuminator, the edges of a diaphragm having a silicon layer thickness of 4.12 μm could be determined, as shown in Fig. 4-43. In one attempt, by alternately focusing on the waveguide edge and chrome mask line, the two could be

aligned to ±1.5 μm. The technique of aligning to a broken out diaphragm could do just as well but required multiple attempts. Alignments to other diaphragms having silicon layer thickness 1 μm greater were also successful and the upper usable diaphragm thickness was not identified. The camera was able to see the diaphragm edge while the unaided eye could not because of the better long wavelength sensitivity of the former.

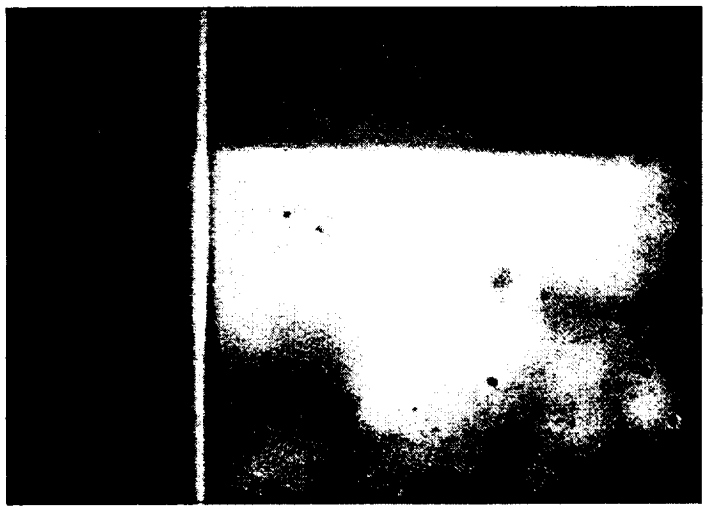


Fig. 4-43 TV monitor showing chrome waveguide line in vacuum contact with substrate having a diaphragm. The focus is on the diaphragm edge defined by the intersection of the <111> and <100> planes on the back of the diaphragm.

Predicting channel index contrast

A typical channel waveguide cross section is shown in Fig. 4-44. The lower oxide layer, grown in a furnace, serves to prevent the guided mode in the SiON film from leaking into the substrate. Its thickness is chosen according to the formulas given in Chapter 3. The effect of contamination on waveguide loss is reduced by the upper cladding oxide which was deposited using LPCVD. Although optional when testing the device by end fire coupling with lenses, it can reduce loss which would occur when a high index adhesive is used to attach optical fibers. Without this cladding, an index match between the SiON film and adhesive would eliminate channel index contrast where

the adhesive covered the ridge. The guided mode would also leak into the adhesive where it contacted the film surface in a manner analogous to substrate coupling, should the adhesive index exceed the SiON index. Such problems are of concern because UV curable adhesives are typically index matched to optical glasses which have an index (~1.55) near the SiON films.

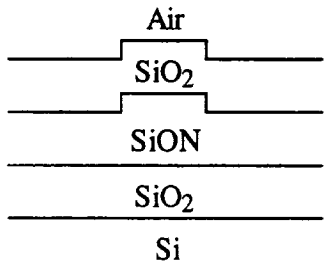


Fig. 4-44 Waveguide cross section.

The etch depth and SiON refractive index must be carefully controlled to obtain a particular channel index contrast $\Delta n_e = n_{e2} - n_{e1}$, where n_{e2} and n_{e1} are the effective planar guide indices which would result if the channel width was ∞ or 0. Bending loss and the number of guided lateral modes depends on the channel index contrast. The etch depth to attain this Δn_e depends in part on the SiON index. Accurately predicting channel index contrast after oxide overlayer deposition from either ellipsometric or prism measurement sometimes proved to be problematic. Although prism coupling provides an accurate measurement of n_e, predicting the reduction of Δn_e due to oxide overlayer deposition requires an accurate measure of the thickness and especially refractive index of the SiON film.

A comparison of the effective index contrast predicted on the finished wafer using ellipsometry and prism coupling will now be given. Referring to Table 4-4, the waveguide wafer #83 was processed with the aid of an SiON monitor, #85, and an oxide overlay monitor, #125. After growth of the isolation oxide on #83, both #83 and #85 were coated with LPCVD SiON film. For the SiON film index to be measured accurately

using ellipsometry on #85, the final SiON thickness should be near midperiod, while the planar waveguide resulting from the same SiON deposition should be single mode: these were simultaneously satisfied at $\lambda_0 = 0.6328$ μm for the ellipsometer and at $\lambda_0 = 0.83$ μm for the waveguide. The SiON refractive index and thickness was also extracted from prism coupling measurements of the TE and TM effective indices on #83 by assuming that the film was sandwiched between air ($n = 1$) and thermal SiO_2 ($n = 1.46$). Compared to the values determined ellipsometrically several days after deposition, the thickness was less and the index was slightly higher. However, effects such as stress could make the film anistropic, altering the planar guide effective indices, and reducing the accuracy of the extracted film index and thickness. To remedy this problem, the index and thickness can be extracted from measurement of several mode indices of the same polarization of a multimode guide [19]. Although anisotropy could reduce the prism measured accuracy of the refractive index and thickness, the accuracy of n_e and Δn_e is unaffected by this problem, and prism coupling was taken as the definitive measure of them. Afterward, #83 was masked with resist and contact printed with the U-bend patterns on the waveguide photomask for subsequent bending loss determination. To permit a prism determination of Δn_e on this sample, an area roughly half the size of the prism base was covered with resist and an adjacent area about the same size left open. The monitor was partially coated with resist using a paint brush. Both were etched in a timed, temperature controlled, slow, wet chemical solution of Timetch, an etchant described in a previous section of this chapter, and the resist was stripped. This etchant was slow enough (> 5 min to form the ridge) so that nonuniformity in rinsing off the etchant will not introduce significant variations in ridge height. Since the etch rate increases 6.4% per degree C, the solution should be controlled to within ±0.5°C. In order to ellipsometrically measure the film thickness after etching on the SiON monitor #85, it was necessary to instruct the program to "fix" the index of the ellipsometer to the value measured before on this sample. The ellipsometer will then ignore its measured ψ and

use the Δ and fixed index n_f to determine the film thickness t_f. This is an effective strategy for near period thickness films, where the curves of constant index, which were previously shown in Fig. 4-26 and 4-27, cluster together and small errors in measuring ψ will produce large changes in the measured film refractive index. The etch depth was also measured using prism coupling and it differed little from that measured ellipsometrically. The Δn_e calculated from the ellipsometrically measured values underestimated those measured on #83.

The channel width was measured using an optical microscope which had a graticule at an image plane in the eyepiece. It was found that the channel width measured on the etched SiON film was less than that similarly measured after the wafer was coated with LPCVD SiO_2. It is believed that the SiO_2 deposition coated the walls of the rib, increasing its width. If this deposited oxide is thick enough, the mode field will not see the air-oxide interface, and so the width measured before SiO_2 deposition is the proper one to use in waveguide calculations. In this study, the oxide will not be made that thick, but as shown for a similar channel in Chapter 3, n_e is near saturation, and so the additional thickness near the sidewalls will not perturb the n_e profile across the channel significantly.

Using LPCVD, SiO_2 was then coated onto #83 and another monitor, #125, which was uncoated prior to the deposition. The SiO_2 thickness was also chosen to be near midperiod of the ellipsometer. Coating a film which has an index of refraction greater than air and lower than SiON onto the SiON will reduce the channel index contrast; this change is sensitive to the index and thickness of the SiON layer. Measurements of the index contrast made using prism coupling after oxide deposition were found to agree well with that calculated from measurements of the thickness and index of the films made using ellipsometry.

#85, SiON monitor	#83, waveguide	#125, SiO_2 monitor
—	Grow 3.0 μm of SiO_2	—
Deposit SiON using LPCVD on #83 and #85 in same run. Conditions in Table 4-2 were used. The wafer pack was rotated and exchanged.		—
Ellipsometrically measure the SiON t and n at $\lambda_0 = 0.6328$ μm, $\theta = 70°$: $\psi = 65.54°$, $\Delta = 73.29°$ This gives $t =$ 0.6553 μm, $n =$ 1.511	Prism couple at $\lambda_0 = 0.6328$ μm: (See Fig. 4-29) $\theta_p = 44.672°$, $n_p = 1.80036$ $\theta_m\|_{TE} = 24.49° \rightarrow n_e\|_{TE} = 1.49375$ $\theta_m\|_{TM} = 24.99° \rightarrow n_e\|_{TM} = 1.48909$ Assume Air, $n = 1$ / SiON t, n / SiO_2, $n = 1.46$ and simultaneously solve for t and n of the SiON from $n_e\|_{TE}$ and $n_e\|_{TM}$ giving $n =$ 1.529 and $t =$ 0.6340 μm	—
Etch in 20°C Timetch for 15.25 min		—
Ellipsometrically measure the SiON t and n at $\lambda_0 = 0.6328$ μm, $\theta = 70°$: $\psi = 15.57°$, $\Delta = 244.33°$ This gives $t = 0.4468$ μm, $n = 1.637$, obviously in error since we are near period thickness. Fixing the index to $n = 1.511$ gives $t = 0.5066$ μm and an etch depth of 0.1487 μm. From this we calculate $\Delta n_e\|_{TE} =$ 0.00924 and $\Delta n_e\|_{TM} =$ 0.00983.	Prism couple at $\lambda_0 = 0.6328$ μm in etched area: $\theta_m\|_{TE} = 25.72° \rightarrow n_e\|_{TE} = 1.48220$ $\theta_m\|_{TM} = 26.39° \rightarrow n_e\|_{TM} = 1.47580$ Thus we get: $n = 1.530$ and $t = 0.4821$ μm and $\Delta n_e\|_{TE} =$ 0.01155 and $\Delta n_e\|_{TM} =$ 0.01329.	—
—	Deposit SiO_2 using LPCVD on #83 and #125 in same run. Conditions in Table 4-2 were used. The wafer pack was rotated and exchanged.	
—	Prism couple at $\lambda_0 = 0.6328$ μm in etched and unetched areas. Unetched: $n_e\|_{TE} = 1.49644$ $n_e\|_{TM} = 1.49413$ Etched: $n_e\|_{TE} = 1.49031$ $n_e\|_{TM} = 1.48797$ $\Delta n_e\|_{TE} =$ 0.00613 and $\Delta n_e\|_{TM} =$ 0.00616.	Ellipsometrically measure SiO_2 t and n at $\lambda_0 = 0.6328$ μm, $\theta = 70°$: $\psi = 53.06°$, $\Delta = 276.77°$, $t = 0.7386$ μm, $n = 1.453$.
—	Using the ellipsometer SiON index, thickness, and etch depth, and the SiO_2 index and thickness, the expected index contrast is $\Delta n_e\|_{TE} =$ 0.00616 and $\Delta n_e\|_{TM} =$ 0.00638. Using the prism coupling values, except for the SiO_2 upper cladding, the expected index contrast is $\Delta n_e\|_{TE} =$ 0.00848 and $\Delta n_e\|_{TM} =$ 0.00894.	—

Table 4-4 Agreement between the calculated and measured channel index contrast.

Fig. 4-45 shows quasi-TE and quasi-TM bending loss after final SiO_2 deposition as a function of bend radius measured on #83, the waveguide whose effective index was tracked in Table 4-4. The loss was determined by taking the ratio of the average of three outscattering intensity measurements made at each end of a U-bend, as shown in Fig. 4-45, using the same equipment described before for measuring propagation loss. Figure 4-46 shows the measured bending loss and compares it with calculations from a theoretical model by Marcuse [24], which was presented in this document in Chapter 3. Marcuse's model was also used to determine the bending loss sensitivity of the channel to small perturbations, such as could be due to fabrication uncertainties, each taken one at a time, and these are also plotted in Fig. 4-46. These show that small perturbations can lead to large changes in bending loss. Although the loss was measured at $\lambda_0 = 0.830\ \mu m$, calculations indicate that the channel index contrast, on which bending loss depends, is different by less than 1.5% from that measured in Table 4-4 at $\lambda_0 = 0.6328\ \mu m$. Loss due to scattering and abrupt transition in curvature is not included in the calculations. However, as shown in Chapter 3, the transition loss component is small. Propagation loss due to effects such as scattering and absorption may be estimated from the larger bending radius arcs to be negligible.

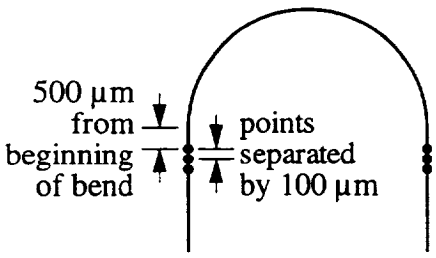


Fig. 4-45 Locations of outscattering points used to measure the loss of the bend.

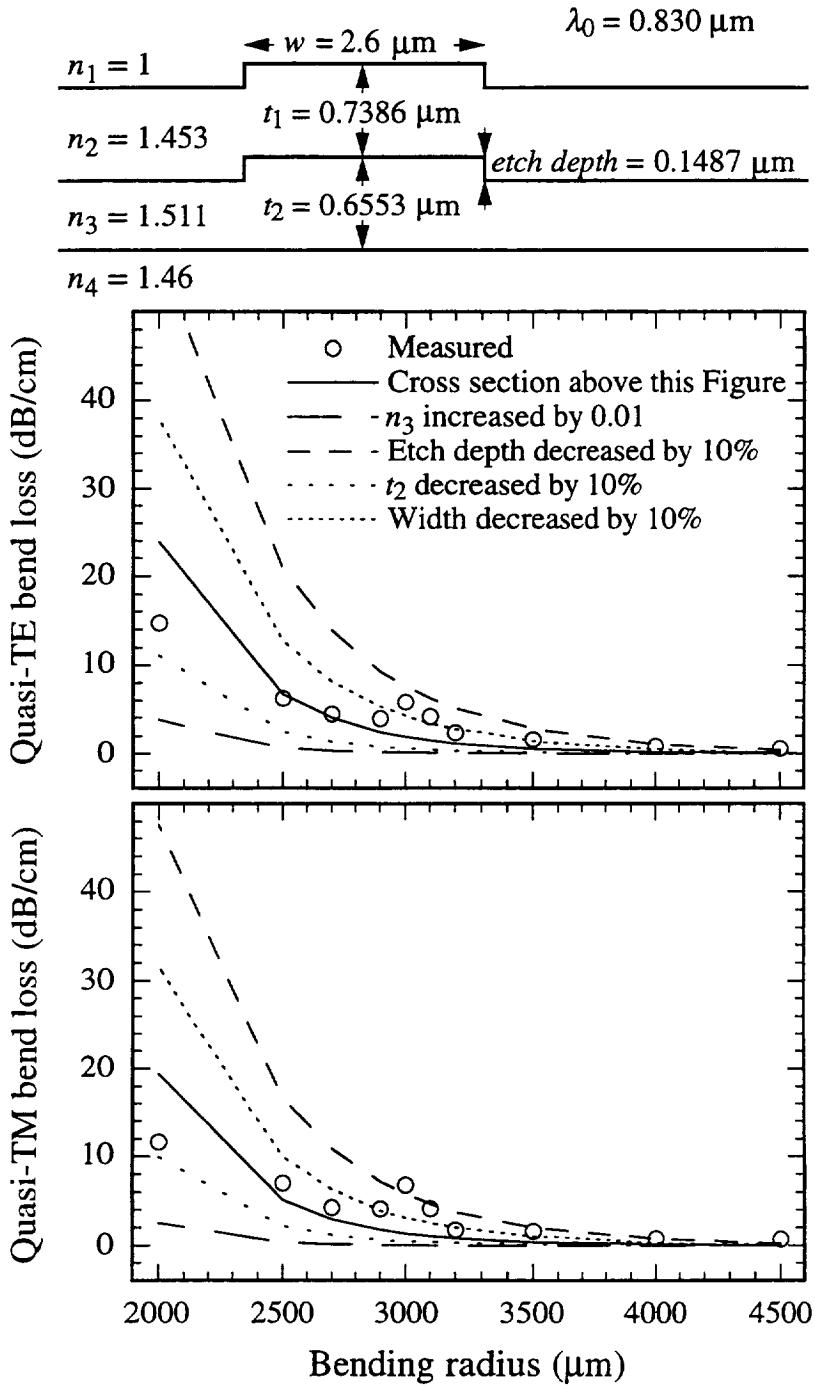


Fig. 4-46 Calculated and measured bending loss on sample #83. Bending loss is also calculated for perturbations made one at a time to the same sample. The width was that measured prior to oxide deposition.

Channel waveguide fabrication sequence

The following processing sequence was used to fabricate waveguides on wafers which have diaphragms or unprocessed single side polished substrates.

1) To make cleaving more reliable, start with <100> wafers which have a <110> major flat. Diaphragms may have already been etched into these wafers. A few monitor wafers should be inserted into the process here for subsequent prism measurement of Δn_e.

2) Scribe an identifying number on the front face if not done already.

3) Base clean to remove particles and any KOH residue.

4) Grow waveguide isolation oxide. The oxide thickness is chosen so that substrate coupling is negligible. Calculation of the propagation loss due to substrate coupling was discussed in Chapter 3.

5) Insert a few unoxidized wafers in the sequence here to aid in determining the LPCVD SiON thickness and index using ellipsometry.

6) Base clean all the wafers to receive the LPCVD SiON film if they are not immediately cycled from into the deposition after oxide growth.

7) Deposit 1/2 of the LPCVD oxynitride film thickness. Purge the gas lines, measure the temperature of the furnace-reactor, load the wafers into the boat with the flats up, the waveguide surface facing away from the door, and perform the deposition.

8) Using the ellipsometer, measure the thickness and refractive index on the monitor which went into the LPCVD unoxidized. Adjust the remaining time of the deposition to hit the target thickness.

9) Clean the substrates briefly in the base clean solution.

10) Put the wafers back into the boat flats down, waveguide surface facing away from the door, and swapped (first is made last, second is made second last), and perform

the second half of the deposition. Measure the index and thickness using the ellipsometer. For highest accuracy, the film should be near midperiod: films near period thickness will not be measured accurately by the ellipsometer. Etch depths based on the incorrect film index will produce an incorrect channel index contrast, which may result in the channel becoming multimode, or having a high bending loss.

11) If the wafer was not immediately cycled from the LPCVD to resist coating, bake at 200°C for 5 hr.

12) Spin coat resist. Use a resist that spins to ~1 μm and syringe filter it onto the substrate, as described in the diaphragm fabrication sequence.

13) If the wafer has no diaphragms, then to enable cleaving perpendicular to the channels, align the waveguides to be either perpendicular or parallel to the flat using the same technique described in step 10a for aligning the diaphragm mask to the flat. Wafers having diaphragms should be aligned using the infrared technique described previously in this chapter. Expose using vacuum contact.

14) Develop and inspect. Look for mouse bites, resist sidewall roughness, and measure line width. If the line is too wide, the channel could become multimode. On wafers which have diaphragms, inspect for alignment. Inspect for scumming under the DIC microscope. If problems are found, redo the resist. Clean the mask with acetone after each vacuum contact. Part of the waveguide substrate may be reserved for prism measurements. Here the resist can be removed using a swab moistened with acetone or covered by application of resist with a paint brush after channel exposure and development.

15) Bake the resist at 105°C for 1 hr.

16) Measure the refractive index and thickness of the SiON film on the monitor and calculate the etch depth required. This will depend on the thickness, index,

intended operating wavelength, and channel width. For interferometers, the channels should also be single mode. Make the etch as shallow as possible, while maintaining a low bending loss. The waveguide mask was designed so that the arc radius in everything but the U-bends with channels in interferometers is large enough that this etch depth will not make the channels multimode at $\lambda_0 = 0.830\ \mu m$.

17) Determine the etch rate of the monitor SiON film in the temperature stabilized, wet etchant selected. If only one monitor is available, it should be masked to leave an unetched area for monitoring when etching the device wafer.

18) Etch the channel patterns into the SiON film by leaving the monitor and device wafer in the etchant for about 80% of the time suggested by the etch rate measured on the monitor. Continue the etch for the amount of time estimated from measurement of the etch depth.

19) Strip the resist in acetone. Base clean the device wafers and SiO_2 monitors.

20) Adjust the LPCVD furnace for SiO_2 deposition as given in Table 4-2.

21) Deposit the SiO_2 film. Rotate and exchange the wafers halfway through the deposition.

22) Using prism coupling, measure the channel index contrast on the device wafer or on a monitor waveguide to determine whether the target Δn_e was hit.

In the next set of steps, a glass wafer will be drilled and electrostatically bonded to the device wafer having diaphragms. Films on the back of the device wafer must be removed prior to bonding. Two techniques can be used to protect the films on the diaphragm back, which were left unetched to keep the diaphragm thermally balanced. The first technique was to paint resist on the diaphragm back under a microscope using a small brush. However, it was difficult to coat the entire diaphragm back and not get any resist on the wafer back which would, subsequent to etching the films off the back, leave

a ~3 μm stack of layers to defeat bonding at the stack and outward from it for hundreds of microns. A technique for coating the diaphragm back which worked better was to flood the wafer back and diaphragm wells with resist and then spin the excess off. The resist remaining in the wells tended to be much thicker than on the wafer back. A blanket exposure and development removed the resist on the back, except for the edge bead, but left the resist in the diaphragm wells, protecting the diaphragm back. The edge bead was removed by a swab moistened with acetone. During the etch, to protect the waveguides, the wafer was attached with resist to a 3" wafer.

1) Bake the device wafers for 5 hr at 200°C

2) Place 1 drop of resist in the middle of a clean 3" wafer. Lay the 2" wafer waveguide side down onto the resist and lightly press on the face of the 2" wafer near its edge.

3) Bake at 120°C for 1 hr.

4) With the wafer pair on the spinner, flood the back of the wafer with thinned resist made by mixing 3:2 by volume of Shipley 1813 resist:type P thinner and spin the excess off at 750 RPM for 30 sec. Using a fine paint brush, flow the thinned resist onto the 3" wafer where it meets the 2" wafer to seal their interface.

5) Bake at 95°C for 1 hr.

6) Using a cotton swab moistened with acetone, remove the edge bead of resist from the 2" wafer.

7) Blanket expose on the aligner for 15 sec at 12 mW/cm^2. Develop for 60 sec in 5:1 by volume of H_2O:Shipley 351 concentrate.

8) Reapply resist to the 3" wafer at the edge of the 2" wafer.

9) Bake at 120°C for 30 min.

10) Etch in 6:1 by volume of 40% NH_4F:49% HF for 50 min or until the wafer back is bare.

11) Soak the wafer pair in acetone overnight to separate.

12) Inspect the wafer back for unetched film and repeat sequence to strip back of wafer if found.

The first device fabricated, the ring resonator, was epoxied directly to an aluminum chuck to apply pressure to the diaphragm back. If too much epoxy was used, it ended up on the diaphragm; too little and the wafer would not be rigidly attached to the chuck. For the second device, the Mach-Zehnder, a glass plate was electrostatically bonded to the wafer back and this epoxied to a sheet of rubber which was epoxied to the aluminum chuck. Such a design removed the requirement of a void free epoxy bond, reduced the chance of getting epoxy on the diaphragm back, and isolated the chip from the thermal expansion mismatch between the glass and chuck [25]. The fabrication sequence now continues with the implementation of this improved process.

The wafer back was then electrostatically bonded to a drilled, double side polished, Schott Borofloat (equivalent to Corning 7740) glass 1/4" thick × 1.75" diameter. Afterward, the glass was epoxied to a 0.032" thick rubber sheet which was epoxied to the machined aluminum chuck. The rubber sheet was used to reduce the shear strain which would be placed on the glass due to a temperature change, since the glass had a smaller (3.25×10^{-6}/°C) coefficient of thermal expansion than the aluminum (25×10^{-6}/°C) chuck. This glass disc was made thick enough (1/4") that even with poor bonding to the chuck, the wafer back would be held flat and diaphragm strain would be independent of the adhesive bond. This scheme was devised after experience with the ring resonator chip, which was attached using epoxy directly to the chuck, and was found to possess creep.

Drilling the glass

A hole must be bored in the glass disc to permit gas pressure to be applied to the diaphragm back. In order to electrostatically attach the glass to the wafer back up to the diaphragm opening, all damage made to the glass, including scratches, chips, and other defects must be kept within a frame defined by the opening in the back of the wafer, approximately 23 mil × 132 mil. If multiple holes are to be drilled into the glass, their positioning error and damaged regions must be such that the wafer back can be aligned to contact undamaged glass.

A 30 W focused CO_2 laser beam was found to be capable of drilling holes in 30 mil Corning Pyrex 7740 glass. In order to avoid cracking the glass, the laser had to be pulsed with a ~200 msec on, 2 sec off duty cycle. The hole had a clear aperture of about 70 μm in diameter and was surrounded by a 300 μm diameter damaged area. In addition, vaporization and melting of the glass sometimes produced a raised area at the hole periphery. However, attempts to bore through 1/4" thick glass proved to be unsuccessful, as once the hole was started the vaporized glass redeposited on the hole wall, impeding drilling further.

Small holes were successfully drilled using diamond plated drills purchased from Lunzer [26]. These drills consist of a 1/8" diameter rod which is ground down at its head to a smaller diameter and plated here with nickel containing diamonds. Since the ground down section of the small drills was less than 1/4" long, a recess was drilled into the glass from one side using a 35 mil diameter plated drill, which had 3/8" of head length, until 50 mils of glass remained. The hole was completed using a 12 mil diameter drill coming in from the other side. Alignment was performed visually. The 12 mil drills were found to be fragile and the diamond imbedded coating was easily broken off when they were used in a standard floor mount drill press. A sensitive drill press which had better feel and less vibration was used to drill the holes without breaking the coating. For drilling with either bit, the glass was waxed to a scrap piece of glass which was waxed to a metal

block clamped to the drill press. Drilling was performed through a few drops of deionized water which were occasionally replenished, and a pecking motion was used to flush the hole. Unlike those drilled using the CO_2 laser in thin glass, the hole's clear diameter was only slightly less than that of the damaged region and both, on the small hole face, could be kept smaller than the well opening in the wafer back. Prolonged soaking in solvent was required to remove wax in the hole, which was found to be ~4 mils larger than the bit diameter.

To prevent the glass from interfering with end fire coupling into the cleaved edge, two parallel cuts were made through the glass before bonding using a diamond saw. These cuts were closer together than the length of the cleaved sample. Both cleaved wafer edges overhung the sawn glass edges after electrostatic bonding. If fiber pigtailing was to be performed, the glass would not be cut, for these areas would be used to strain relieve the fibers.

Electrostatic bonding

Both electrostatic (anodic) bonding [27, 28] and direct wafer bonding [29] have been used to attach two substrates together without adhesive. Direct wafer bonding has several disadvantages when compared to electrostatic bonding: it requires higher temperatures (900°C); it requires purchasing thick silicon or quartz wafers instead of thick glass discs; the bond is more difficult to inspect if two silicon wafers are used; and voids can occur more readily than for electrostatic bonding [30].

In order to obtain void-free electrostatic bonds, both the diaphragm wafer and glass must be free of particles. After the glass was thoroughly cleaned of mounting wax, especially in the drilled hole, it and the diaphragm wafer were base cleaned, rinsed, and blown dry. They were mated under a clean hood and immediately carried to the electrostatic bonding setup, shown in Fig. 4-47. This system consisted of a hotplate constructed of a block of steel having cartridge heaters inserted into holes drilled parallel to the block's surface and powered by a Variac. In order to avoid scratching the

waveguide surface, the device wafer with the glass disc on top was set on a bare 3" wafer which rested on the heated steel block. A grounding electrode which contacted the back of the 2" wafer was connected to the steel block. A negative potential was applied via two probes to the glass disc.

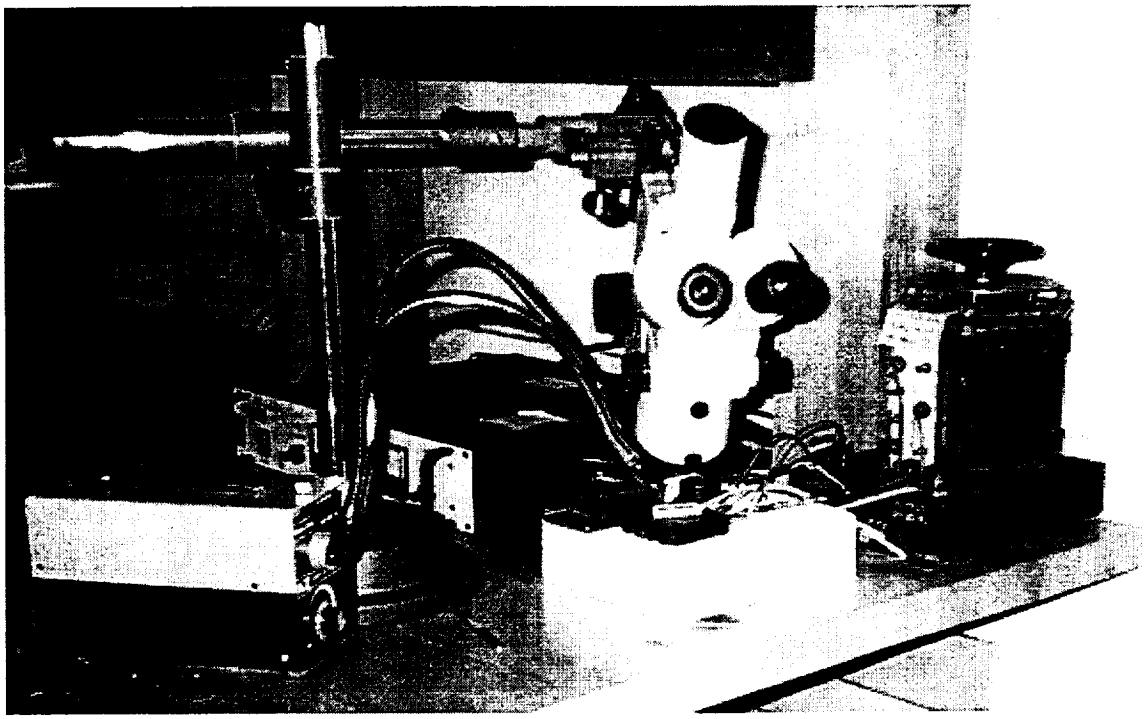


Fig. 4-47 Electrostatic bonding setup.

Once the stack was roughly positioned, a binocular zoom microscope was used to aid in the alignment, which was done by hand. Current to the hotplate was then applied, and after a 40 min warm up, the hotplate temperature reached 385°C. A voltage of −1200 VDC with respect to the grounded metal block and silicon wafer was then applied to both probes and bonding commenced immediately. At this temperature, the glass is slightly conductive, and the potential causes a drift of sodium to the probes, while driving oxygen to the substrate glass interface [31]. It has been observed that cavities sealed under vacuum end up containing oxygen after bonding [28]. The gray bonded area initiated near the probes, and within 5 min had swept over half the wafer. After 30 min,

bonding was complete, the hotplate was turned off, and when it had cooled 20 min later, the −1200 VDC was turned off. Except for one or two 1 mm diameter areas away from diaphragm openings, the entire face of the glass disc was bonded to the silicon wafer. Fig. 4-48 shows the drilled glass bonded to the wafer having diaphragms. The alignment established before the hotplate was turned on was not disturbed by bonding. The hole in the glass is entirely within the opening in the wafer back.

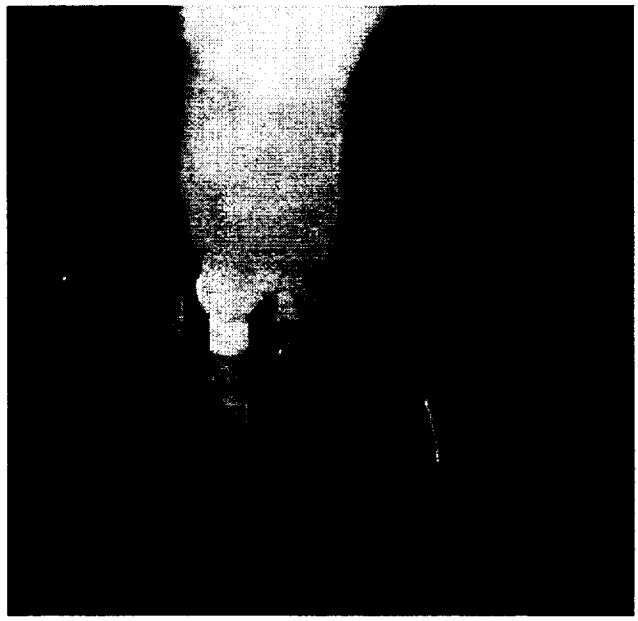


Fig. 4-48 Glass bonded to back of wafer. The larger and smaller holes appear white because of wall roughness. A nearby diaphragm back is also visible.

Once the wafer was bonded, it was epoxied to a rubber sheet, which was epoxied to the aluminum chuck, as shown in Fig. 4-49. The chuck, which was machined from one piece of aluminum, had a flat platform with drilled holes which allowed pressure to be applied to several diaphragms at once, if desired. It terminated in a 1/4" tubular stub which had a double compression fitting (Swagelok).

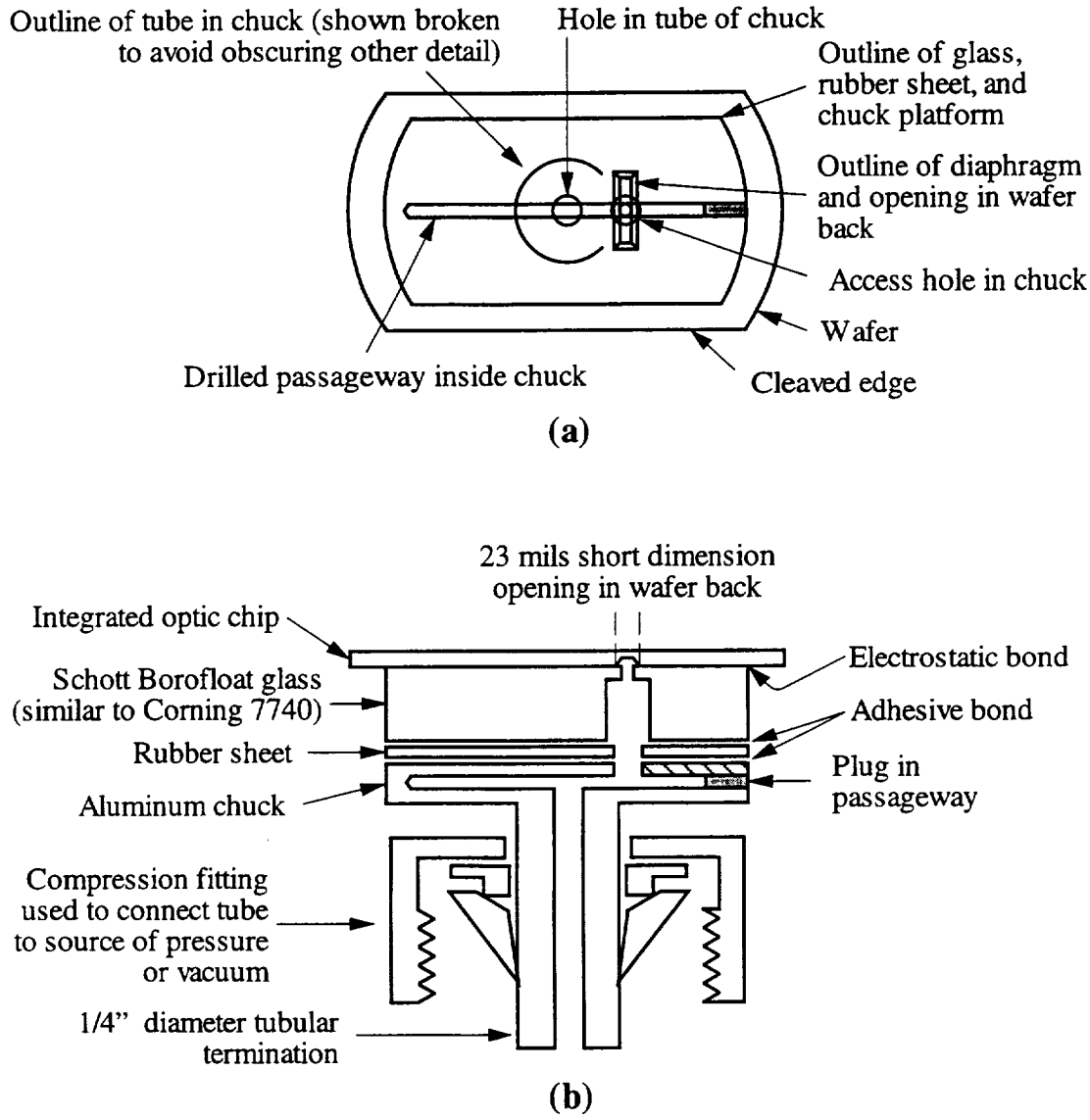


Fig. 4-49 (a) X-ray view looking through face of wafer onto the glass and chuck. (b) Cross section through middle of chuck.

This completed the fabrication and packaging of the device. The assemblage was placed in a setup which allowed the sensor to be optically interrogated while air pressure was applied to the back of the diaphragm. The setup, interrogation method, and test results are reported in the next chapter.

15 REFERENCES

[1] SEMI (Semiconductor Equipment Materials International), *Volume 3, Materials*, pp. Standard M1.1-89, "Standard for 2 inch polished monocrystalline silicon wafers," p. 21, 1991.

[2] K. E. Petersen, "Silicon as a Mechanical Material," *Proc. IEEE*, vol. 70, pp. 420-457, 1982.

[3] Transene, Co., Route One, Rowley, MA 01969.

[4] L. Tong, J. T. Hsu, W. H. Ko, and X. Ding, "The Analysis of Capacitive Pressure Sensors with Large Deflections," presented at International Conference on Solid-State Sensors and Actuators, pp. 185-187, 1991.

[5] J. Geist, A. R. Schaefer, J. F. Song, Y. H. Wang, and E. F. Zalewski, "An Accurate Value for the Absorption Coefficient of Silicon at 633 nm," *J. of Res. Natl. Inst. Stand. Technol.*, vol. 95, pp. 549-558, 1990.

[6] S. M. Sze, *Physics of Semiconductor Devices*, 2nd edition: John Wiley, p. 11, 1981.

[7] W. M. Paulson, F. S. Hickernell, and R. L. Davis, "Effects of Deposition Parameters on Optical Loss for RF-Sputtered Ta_2O_5 and Si_3N_4," *J. Vac. Sci. Technol.*, vol. 16, pp. 307-310, 1979.

[8] A. Mukherjee, B. J. Eapen, and S. W. Baral, "Very Low Loss Channel Waveguides in Polymethylmethacryalate," *Appl. Phys. Lett.*, vol. 65, pp. 3179-3181, 1994.

[9] P. K. Tien, G. Smolinsky, and R. J. Martin, "Thin Organosilicon Films for Integrated Optics," *Appl. Opt.*, vol. 11, pp. 637-642, 1972.

[10] T. Miyashita, S. Sumida, and S. Sakaguchi, "Integrated Optical Devices Based on Silica Waveguide Technologies," presented at Integrated Optical Circuit Engineering VI, and published in proceedings of SPIE, vol. 993, pp. 288-294, 1988.

[11] W. Gleine and J. Muller, "Low-Pressure Chemical Vapor Deposition Silicon-Oxynitride Films for Integrated Optics," *Appl. Opt.*, vol. 31, pp. 2036-2040, 1992.

[12] R. S. Moshrefzadeh, M. D. Radcliffe, T. C. Lee, and S. K. Mohapatra, "Temperature Dependence of Index of Refraction of Polymeric Waveguides," *J. Lightwave Technol.*, vol. 10, pp. 420-425, 1992.

[13] M. Hendriks, E. Granneman, M. Kuhn, and P. Bachmann, "Vacuum Pump Evaluation for an LPCVD Nitride Process," *Solid State Technol.*, pp. 123-126, November 1989.

[14] P. Pan and W. Berry, "The Composition and Physical Properties of LPCVD Silicon Nitride Deposited with Different NH_3/SiH_2Cl_2 Gas Ratios," *J. Electrochem. Soc.*, vol. 132, pp. 3001-3005, 1985.

[15] A. E. T. Kuiper, S. W. Koo, F. H. P. M. Habraken, and Y. Tamminga, "Deposition and Composition of Silicon Oxynitride Films," *J. Vac. Sci. Technol. B*, vol. 1, pp. 62-66, 1983.

[16] C. A. Deckert, "Pattern Etching of CVD Si_3N_4/SiO_2 Composites in HF/Glycerol Mixtures," *J. Electrochem. Soc.*, vol. 127, pp. 2433-2438, 1980.

[17] R. J. Archer, *Manual on Ellipsometry*. Chicago: Gaertner Scientific Corporation, pp. 14, 1968.

[18] R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light*: North-Holland, pp. 289-290 and p. 316, 1977.

[19] R. Ulrich and R. Torge, "Measurement of Thin Film Parameters with a Prism Coupler," *Appl. Opt.*, vol. 12, pp. 2901-2908, 1973.

[20] M. Born and E. Wolf, *Principles of Optics*, 6th edition: Pergamon Press, pp. 177-179, 1989.

[21] P. G. Flavin, "Fabrication of Curved Structures by Electron-Beam Lithography," *Electron. Lett.*, vol. 18, pp. 865-867, 1982.

[22] Bard Laboratories, Inc., Milford Rd., Amherst, NH 03031.

[23] W. Kern, "Purifying Si and SiO_2 Surfaces with Hydrogen Peroxide," *Semiconductor International*, pp. 94-99, April 1984.

[24] D. Marcuse, *Light Transmission Optics*. New York: Van Nostrand Reinhold Company, p. 404, 1972.

[25] Y.-C. Lin, P. J. Hesketh, and J. P. Schuster, "Finite-Element Analysis of Thermal Stresses in a Silicon Pressure Sensor for Various Die-Mount Materials," *Sensors and Actuators*, vol. A 44, pp. 145-149, 1994.

[26] Lunzer Inc., 280 N. Midland Ave., Building J1, Saddle Brook, NJ 07663.

[27] G. Wallis and D. I. Pomerantz, "Field Assisted Glass-Metal Sealing," *J. Appl. Phys.*, vol. 40, pp. 3946-3949, 1969.

[28] T. Rogers and J. Kowal, "Selection of Glass, Anodic Bonding Conditions and Material Compatibility for Silicon-Glass Capacitive Sensors," *Sensors and Actuators A*, vol. 46-47, pp. 113-120, 1995.

[29] K. Petersen, J. Brown, T. Vermeulen, P. Barth, J. Mallon, and J. Bryzek, "Ultra-Stable, High-Temperature Pressure Sensors Using Silicon Fusion Bonding," *Sensors and Actuators A*, vol. 21-23, pp. 96-101, 1990.

[30] R. Stengl, K. Y. Ahn, and U. Gosele, "Bubble-Free Silicon Wafer Bonding in a Non-Cleanroom Environment," *Japan J. Appl. Phys.*, vol. 27, pp. L2364-L2366, 1988.

[31] W. H. Ko, J. T. Suminto, and G. J. Yeh, "Bonding Techniques for Microsensors," in *Micromachining and Micropackaging of Transducers*, C. D. Fung, P. W. Cheung, W. H. Ko, and D. G. Fleming, Eds. Amsterdam: Elsevier Science Publishers, 1985, pp. 41-61.

V Temperature and Pressure Sensitivity Results

1 INTRODUCTION

This chapter shows the test setups, briefly describes the fabrication sequence, and provides the measured pressure and temperature sensitivities of the ring resonator and Mach-Zehnder pressure sensors. As recounted in Chapter 1, such all-optical sensors are immune to electromagnetic interference, and are not temperature limited by semiconductor junctions. They are intended to function in an electrically hostile environment, and be linked with optical fibers to an electronic controller in an electrically benign environment. In this work, for simplicity, light was launched into and collected from the waveguides using lenses.

The ring resonator was fabricated first, and its large temperature sensitivity prompted the fabrication of the smaller path length difference Mach-Zehnder interferometer. During Mach-Zehnder fabrication, improvements in channel etch depth control and chuck to chip bonding to solve a creep problem were implemented.

Both sensors consist of integrated optic interferometers in which part of their waveguide is positioned over the long edge of a rectangular diaphragm. In each of the devices, the interferometer optical path length imbalance is altered by pressure applied to the diaphragm back, and this change is measured optically. Because of the large optical path length of the ring resonator, it could be interrogated with a current-ramped, swept-frequency 830 nm laser diode. A shift in resonance peak with respect to the beginning of a current ramp was used to deduce the pressure. An isolator was necessary in the ring resonator setup to prevent feedback into the laser that otherwise produced satellite modes which washed out fringe visibility. The ring resonator was found to be much more temperature sensitive than the smaller path length Mach-Zehnder interferometer. The smaller path length difference between the two arms of the Mach-Zehnder interferometer

demanded a different interrogation technique. Here light from an 850 nm super luminescent diode was launched into the device, and that emerging was spectrally separated. The change in spectrum was used to deduce the pressure. In both devices, the quasi-TM mode was found to be much more pressure sensitive than the quasi-TE mode, in agreement with calculations made in Chapter 3 which assumed that the waveguide film had the same photoelastic constants as SiO_2. Creepage in the ring resonator was noted: after a pressure change, the ring-resonator optical path length was found to continue changing in the same direction. This effect was attributed to plasticity in the glue bond between the bottom of the chip and the aluminum chuck on which it was mounted. Such effects were avoided in the Mach-Zehnder by a thick, electrostatically bonded piece of glass interposed between the chuck and wafer.

*2 RING RESONATOR PRESSURE SENSOR**

Operation

The *Y*-junction ring resonator is pictured in Fig. 5-1 and the dimensioned pattern on the mask used to fabricate the device in Fig. 5-2. The location of the pattern in the mask layout was shown in the previous chapter, and its theoretical response function and expected pressure and temperature sensitivities were derived and calculated in Chapter 3. As shown in Chapter 3, a waveguide is most sensitive to pressure when it is placed over the long edge of a rectangular diaphragm, the location selected for a straight waveguide segment in the ring resonator. The device is bi-directional and so either channel may be used as input. To begin, assume that the *Y*-junction ring resonator is unexcited. For descriptive purposes, the *Y*-junction connected to the channel chosen for input will be termed "the first *Y*-junction". Light launched into the straight waveguide encounters the first *Y*-junction, and because no light yet exists in the ring, suffers a 3 dB loss. The light

* Some of the text and figures in this section have been published in: G.N. De Brabander, J.T. Boyd, and G. Beheim, "Integrated Optical Ring Resonator With Micromechanical Diaphragm for Pressure Sensing," IEEE Photon. Technol. Lett., vol. 6, pp. 671-673, 1994. Additional information has been added which was omitted from the paper in the interest of brevity.

not lost continues to the second *Y*-junction, where it is equally split, with half leaving through the output channel, and half going into the ring. Light in the ring propagates back to the first *Y*-junction, where it interferes with that launched into the channel selected for input. If the optical path length of the ring, which includes the straight channel linking the two *Y*-junctions and the loop connecting the branches of the two *Y*-junctions, is an integral multiple of wavelengths long, light in the ring will interfere constructively with that launched into the input channel. Much of the power launched into the straight channel linking the two *Y*-junctions will then be transmitted to the second *Y*-junction and to the output channel: the device transmissivity will be high. If the optical path length of the ring is an integral multiple of wavelengths plus one half wavelength long, light in the ring will destructively interfere with that launched into the input channel and the device transmissivity will be low. Note that unlike the ring resonator formed using a directional coupler, when the interferometer transmissivity is high, light lost by outscattering from the ring waveguide is expected to be most intense. Although a ring resonator fabricated using a directional coupler can have a higher finesse, this is not required for the pressure sensor, and the *Y*-junction power split is less fabrication dependent. Pressure applied to the diaphragm back creates stress in the waveguide films that increases, through the photoelastic effect, their refractive indices. This increases the effective index of the channel waveguide and leads to an increase in retardance of the light propagating in the channel. In terms of the optical path length, constructive interference will occur when

$$m\lambda = OPL = \int_0^L n_{et}(\ell)\, d\ell \quad , \tag{5.2.1}$$

where the channel effective index n_{et} is integrated around the ring of length L. In terms of the phase ϕ of light at the end of the ring where it joins the first *Y*-junction, the previous condition can be written

$$-2\pi m = \phi = -\frac{2\pi f}{c_0}\int_0^L n_{ef}(\ell)\, d\ell \quad , \tag{5.2.2}$$

where m is an integer, f is the optical frequency of the laser source, and c_0 is the speed of light in vacuum. This definition of phase is the same given in Chapter 3, Sections 15 and 16. According to that definition, when light is moving in the $+z$ direction, the electric field z distance from a reference point is given by $E_0 \exp(j\phi z)$. Also, if the optical path length of a channel increases, its change in phase is negative, and the light emerging from the guide is more delayed.

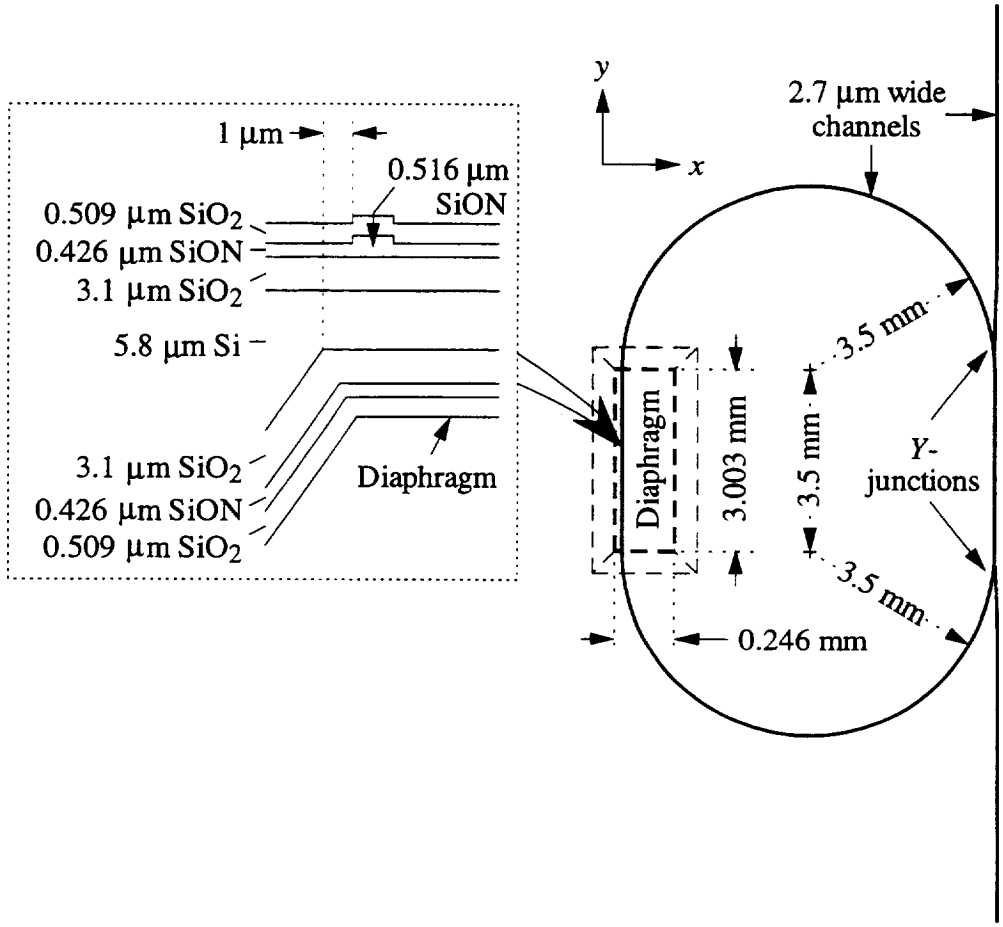


Fig. 5-1 The Y-junction ring resonator.

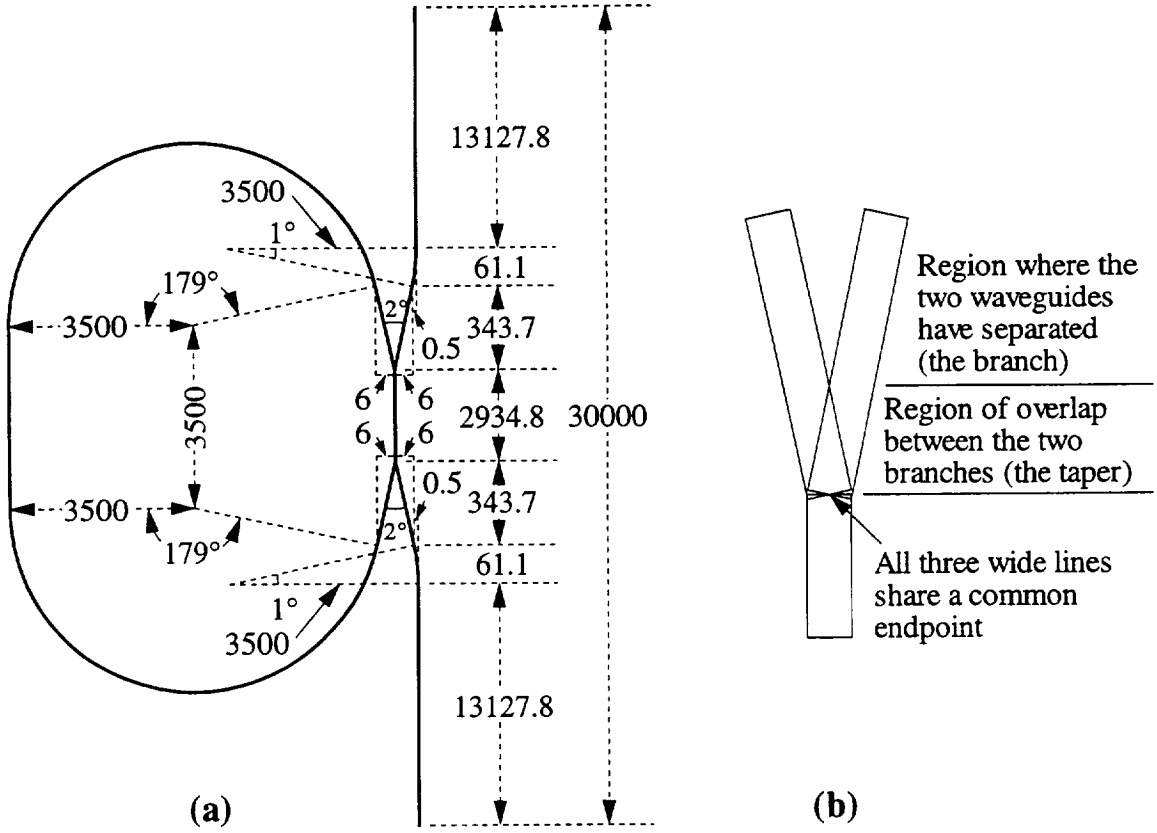


Fig. 5-2 (a) Dimensions in μm of the *Y*-junction ring resonator. In the Figure, the centerline of the waveguides, each of which was 2.4 μm wide, is drawn, and so the distances are measured from the middle of the ends of the channel segments (polylines). The distances are measured either parallel or perpendicular to the input channel direction. According to the detail in (b), which shows the three line segments used to make up the junction, the taper is the result of the overlap of two wide lines; the 343.7 μm dimension shown in (a) includes both the taper and branch. The *Y*-junctions were designed to be symmetric: the branches separate by 12 μm (center-to-center channel distance) before any dissimilarity is encountered. This is far enough so that the coupling between the branches is negligible, insuring an equal energy split. As discussed in Chapter 4, before the drawing file was submitted to the mask house, the arcs were broken into enough line segments (413 segments per 179° arc) that if a perfect arc were laid atop the line segments, no more than 0.04 μm of arc would protrude radially out from the edges of the lines. This avoided the course arc fragmentation produced by the mask house software.

Ring resonator interrogation scheme

The interrogation scheme, called pseudoheterodyne demodulation, consists of sweeping the frequency of light fed into the interferometer by ramping the drive current of the diode laser source, and measuring the time delay between the start of a ramp and a transmissivity minimum [1]. This is then normalized by the time delay between two adjacent minima and the phase ϕ at the beginning of a ramp determined. As the drive current to the laser increases during a ramp cycle, its wavelength becomes longer. Pressure applied to the diaphragm back will increase the ring optical path length or move its phase toward larger negative values. This will move the transmissivity minimum under observation to longer wavelengths and the minimum will be reached further along in the ramp cycle: the time delay between the start of a ramp and the minimum will increase. The pressure can be determined from this time delay using prior calibration.

In each cycle of the current ramp, the current rise produces a temperature rise in the GaAlAs laser diode which increases its refractive index, decreasing its frequency. The current ramp induced temperature rise tunes the laser frequency by -35 GHz/°C due to a refractive index change, and alters its gain curve by -140 GHz/°C [2]. Due to the difference in these values, the laser cannot be continuously tuned by more than 50 GHz, as this results in it jumping to a different cavity mode. Since a frequency sweep of at least the interferometer free spectral range $\left(FSR = \frac{c_0}{Ln_e}\right)$ is required to insure that a transmissivity minimum is observed, the smallest optical path length imbalance that can be interrogated with this technique is 6 mm; in the ring resonator it is 43.5 mm for a free spectral range of 6.9 GHz. Use of a ring resonator instead of Mach-Zehnder interferometer can provide the required optical path length from a more compact device.

This interrogation method provides a high degree of immunity to changes in the efficiency of coupling to and from the integrated optical circuit, which may be remotely located and linked by optical fibers whose losses, especially when connectorized, are in practice quite variable [3]. In addition, this technique eliminates the signal fading

problem and it allows the direction of the phase change to be determined at all times so that it may be tracked over a range greater than π.

Other interrogation schemes have been employed to detect the phase change of interferometers with arbitrary imbalance. For example, in the active homodyne, a fixed wavelength, long coherence source is used to excite the interferometer whose imbalance is actively biased [4]. For Mach-Zehnder or ring resonator interferometers made using optical fibers, a piezo electric fiber stretcher, constructed by wrapping a segment of the reference arm of the Mach-Zehnder or part of the ring around a tube which expands in response to voltage, can be used to actively ramp or track the phase. Thermooptic or electrooptic shifters could perform the same function in an integrated optic configuration. To accomplish this with the proposed sensors would require an electrical signal to be supplied to the head, causing the loss of electromagnetic immunity, however.

Ring resonator fabrication summary

The fabrication sequence is given in Fig. 5-3. Chapter 4 details the procedure for fabricating diaphragms and waveguide films, so only an abbreviated outline is given here. The starting substrates were <100>, double side polished, 2" diameter, 250 μm thick, silicon wafers having a <100> flat. The diaphragm was fabricated first to avoid exposing the waveguide films to the etchant, which would reduce their thickness. Steam oxide grown at 1050°C was used for masking, and the substrates were etched in 40°C KOH until the proper diaphragm thickness remained. No etch stop was employed. Diaphragm thickness was estimated by subtracting the well depth, measuring by focusing with an optical microscope alternatively on the wafer back and diaphragm, from the wafer thickness, measured with a micrometer. After diaphragm formation, the oxide was stripped everywhere and regrown using 1050°C steam oxidation to 3.1 μm, a thickness sufficient to make the TE and TM mode leakage into the substrate negligible. Using low pressure chemical vapor deposition (LPCVD), a silicon oxynitride waveguide film having a thickness of 0.516 μm and refractive index of 1.553 was deposited at a pressure of

400 mTorr in a horizontal tube diffusion furnace by flowing O_2, NH_3 and $SiCl_2H_2$ [5] at the flow rates 11, 100, and 20 sccm at a temperature of 750°C for 78 min. The index and film thickness was measured by prism coupling into a monitor and measuring the TE and TM mode angles. These were used to determine the TE and TM effective indices (see Chapter 4). Then the SiON index and thickness required to obtain these two n_e's was calculated. The ring resonator pattern was next defined in photoresist using contact print lithography. One of the interferometer mask patterns which was not to be tested was aligned to its respective diaphragm edge, which had been broken out with a pin. The unprotected SiON was then etched in BHF to a 90 nm depth, as determined by TE and TM prism coupling angle measurements on a monitor wafer. Then, with LPCVD, an SiO_2 cladding layer was deposited at 300 mTorr by flowing N_2O and $SiCl_2H_2$ at the flow rates of 40 and 10 sccm at 910°C for 106 min. Using $\lambda_0 = 0.633$ μm ellipsometry on a monitor wafer having only a deposited oxide, the cladding was determined to be 0.509 μm thick.

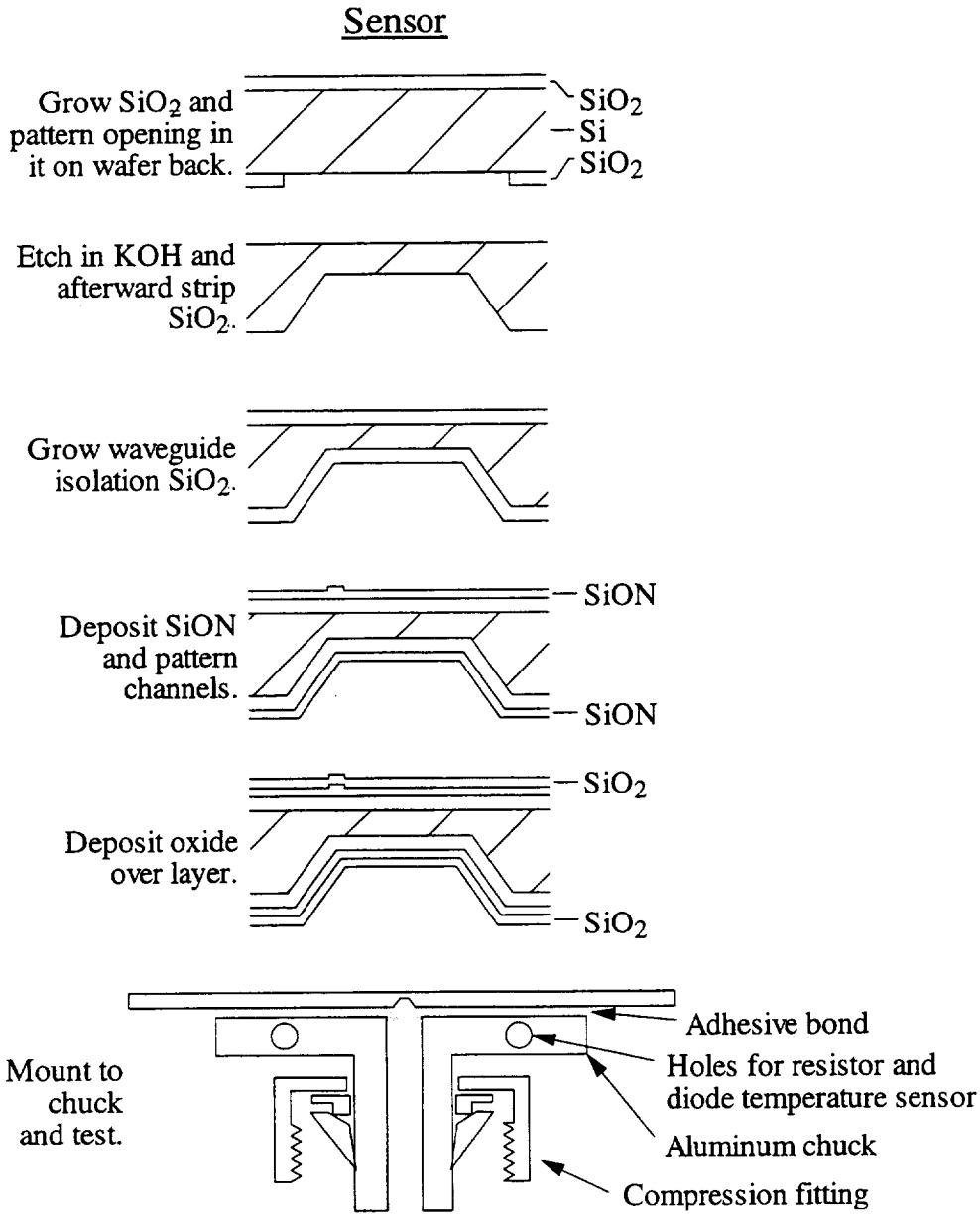


Fig. 5-3 Ring resonator fabrication sequence. Effective index and ellipsometer monitors were processed along with the device wafer.

The ring resonator was cleaved through the lead-in and lead-out channels and epoxied to an aluminum chuck. After all measurements were made, the wafer was removed from the chuck, the diaphragm cleaved through, and its silicon layer thickness measured to be 5.8 μm with an optical microscope.

Experimental setup

To measure the ring phase at the beginning of a ramp cycle (actually ϕ mod 2π), light from an 830 nm wavelength single longitudinal mode laser diode (Sharp LTO15MDO) was collimated and directed through an isolator consisting of a dichroic sheet polarizer, mica quarter wave plate and attenuator, and then end fire coupled into the input channel, as shown in Fig. 5-4. Since circularly polarized light was focused onto the cleaved end face, both quasi-TE and quasi-TM modes were excited; these propagated through the chip independently. A polarizer, located at the exit of the output channel, was used to select one of these modes, the intensity of which was measured by a photodiode.

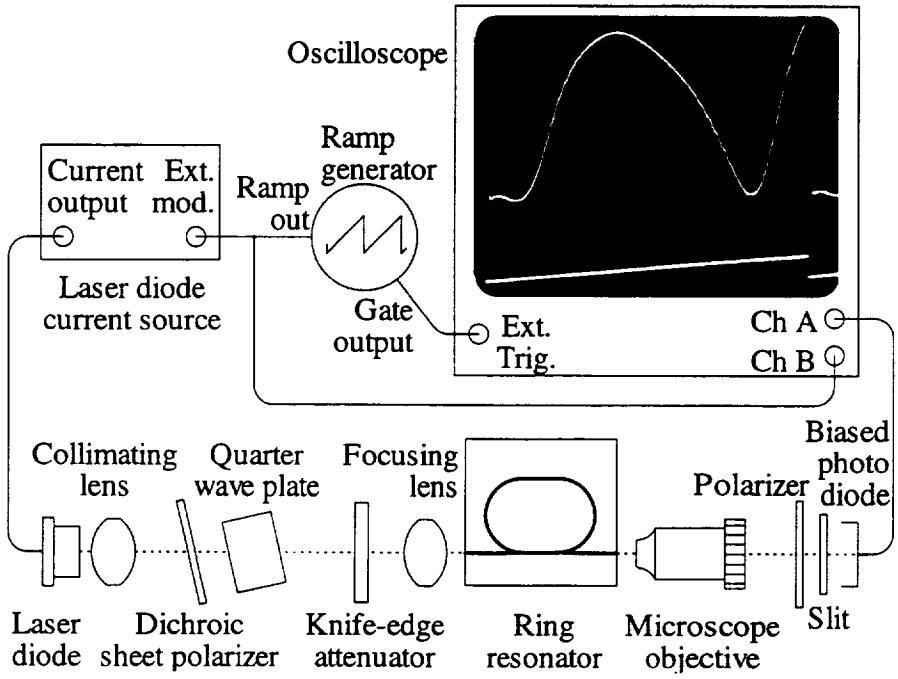


Fig. 5-4 Test setup for ring resonator sensor. Channel A is the top trace and B is the bottom one. The period if the ramp was 8 msec.

The laser was driven by a DC bias of approximately 60 mA, which maintained it above threshold, summed with a ramp function having a 3.3 mA peak-to-peak amplitude and a period of 8 ms. Over one cycle, as the drive current increased, the refractive index

of the laser cavity increased, decreasing the optical frequency. This ramp amplitude swept the laser frequency by 10 GHz, slightly more than the free spectral range of the interferometer [1], enabling at least one minima to be observed. An oscilloscope displayed the response of the ring resonator's photodiode, which for some measurements was normalized by the output of a photodiode that monitored laser intensity. The ring phase ϕ in radians at the start of a frequency sweep was determined by measuring the time delay between the start of a frequency sweep and a transmissivity minima, scaling this result by 2π, and then normalizing this product by the time delay between two adjacent minima. The sign of the phase change can be determined from the direction of the minima shift on the oscilloscope screen. In Chapter 3, the transfer function of the *Y*-junction ring resonator was derived and it was shown that a transmissivity minima will occur when $\frac{2\pi n_{et}}{\lambda_0}(\ell_1 + \ell_2) = \pi, 3\pi, 5\pi, \ldots = (2m+1)\pi$, where m is an integer, n_{et} is the channel effective index, and $\ell_1 + \ell_2$ is the path length around the ring including the straight section linking the two *Y*-junctions. For the positive slope current ramp depicted in Fig. 5-4, movement of a particular transmissivity minimum (fixed m) toward the right of the oscilloscope screen, and therefore toward higher laser drive currents and longer wavelengths, implies that the ring optical path length $n_{et}(\ell_1 + \ell_2)$ has increased and the phase has become more negative. The phase ϕ can therefore be defined

$$\phi = -2\pi\left(\frac{\text{time delay between the start of the ramp and a minimum}}{\text{time delay between two adjacent minima}} + \frac{1}{2} + m\right), \qquad (5.2.2)$$

where m is an integer, the 1/2 is used because the minima occur at odd integral π multiples, and a positive ramp is assumed to be applied to the laser. From this measurement, the phase can only be determined to mod 2π at turn on, but can be tracked over an arbitrary change.

Air pressure or vacuum applied to the back of the chip was controlled using a pressure regulator and monitored with a gauge, as shown in Fig. 5-5. The chuck screwed into an elbow which was soldered to a bar secured to a crossed-roller stainless steel stage.

This was placed into the setup of Fig. 5-4 to enable optically interrogating the chip while simultaneously applying pressure or vacuum to the diaphragm back.

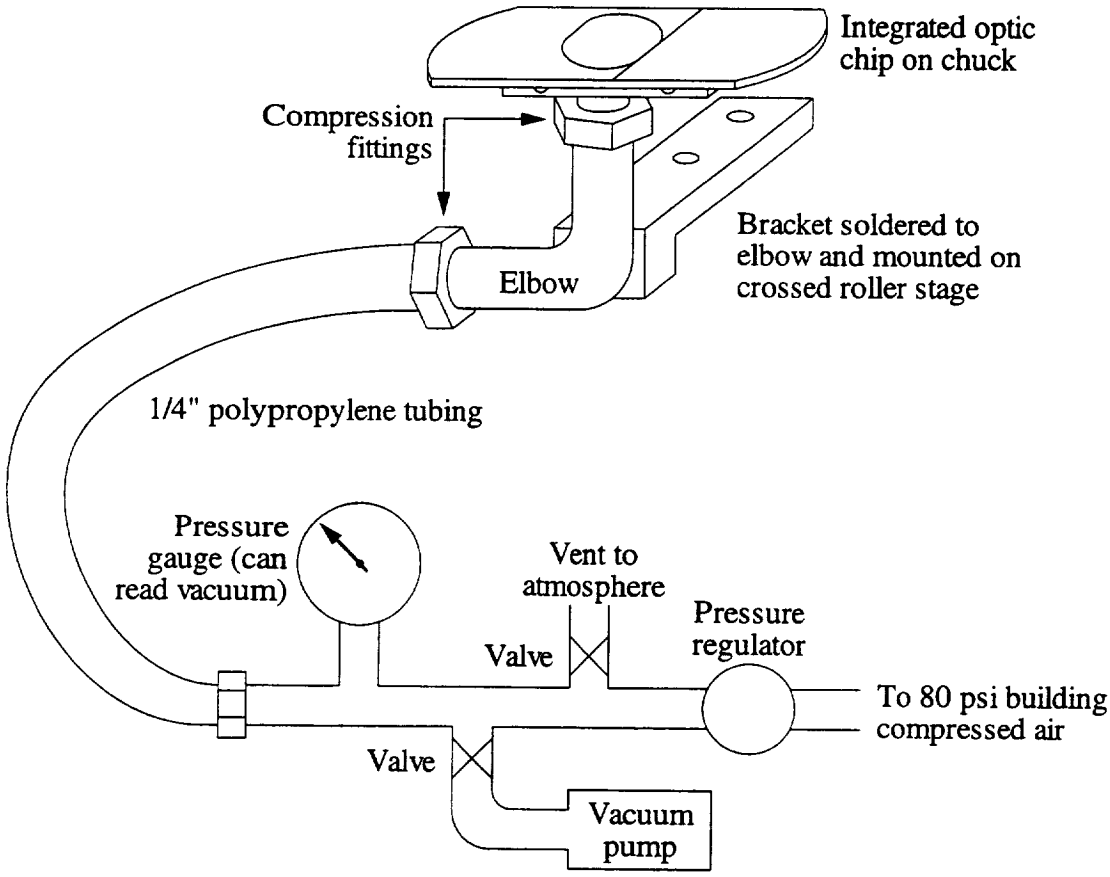


Fig. 5-5 Air pressure system and chuck mount. For ring resonator measurements, a ±3%, −15 psi to +60 psi mechanical gauge was used, while Mach-Zehnder characterization, which used this same configuration, was performed using a ±0.25%, -15 psi to 200 psi digital gauge from Cole-Parmer.

Laser isolation and mode hopping

End fire coupling by carefully aligning the chip to maximize channel power without the use of an isolator was found to wash out ring resonator fringe visibility: the ring resonator's photodiode measured only the increase in laser intensity with drive current. Monochromator measurement of the laser spectrum leaving the channel indicated that it had broadened to several nm or hundreds of GHz. To resolve ring

resonator transmissivity variations, light having a coherence length in excess of the ring optical path length of 4.3 cm or a frequency broadness much less than 6.9 GHz is needed. Fringes were observed when coupling was vastly reduced by defocusing the laser spot on the cleaved end face or lowering the waveguide almost out of the beam path. Even with the isolator it was necessary to reduce back reflected power to avoid Fabry-Perot effects between the chip and laser edge or unavoidable mode hopping over small current sweeps, problems depicted in Fig. 5-6, by use of a knife edge attenuator interposed in the beam path, moving the chip along the beam away from its waist, or lowering the wafer with respect to the beam. However, use of the isolator was found to increase the incoupled power before the onset of laser instability.

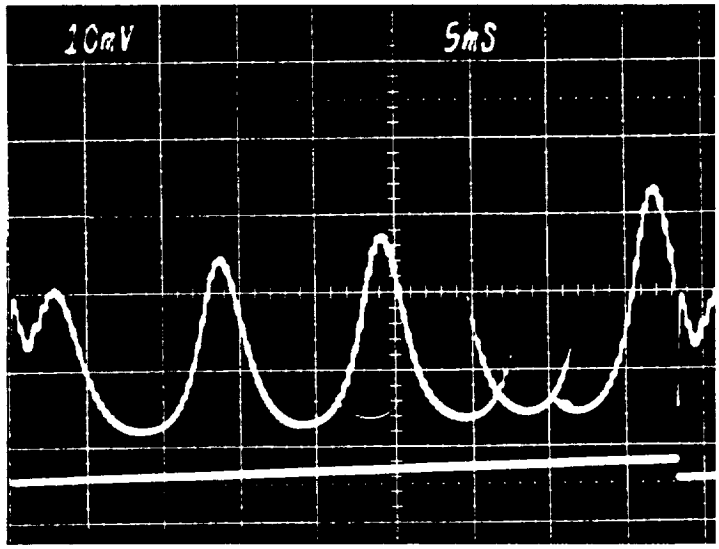


Fig. 5-6 Output of a *Y*-junction ring resonator using the setup of Fig. 5-4 with the isolator; coupling has been maximized. Ghost images in the photo are the result of an occasional mode hop during photograph exposure, which required more than one scope sweep to take. The small wiggles (roughly 1/20 of a division) could be translated along the ring resonator curve shape by slight movements of the laser or the chip parallel to the beam direction. The ring resonator in this Figure was single mode and so had a different transfer function than the one in Fig. 5-4.

The mode hopping problem was investigated further using the setup shown in Fig. 5-7. Part of the laser light was diverted using a metallic beam splitter to a

monochromator and its output monitored using a photodiode. For this experiment, the laser frequency sweep was turned off and the laser drive current was adjusted to be at a mode hop. In the single shot oscilloscope trace shown in Fig. 5-7, the laser frequency is observed to rapidly jump between at least two wavelengths, and one of them was transmitted by the monochromator to the detector. In contrast, the light intensity entering the monochromator as observed with the photodiode showed none of this rapid variation.

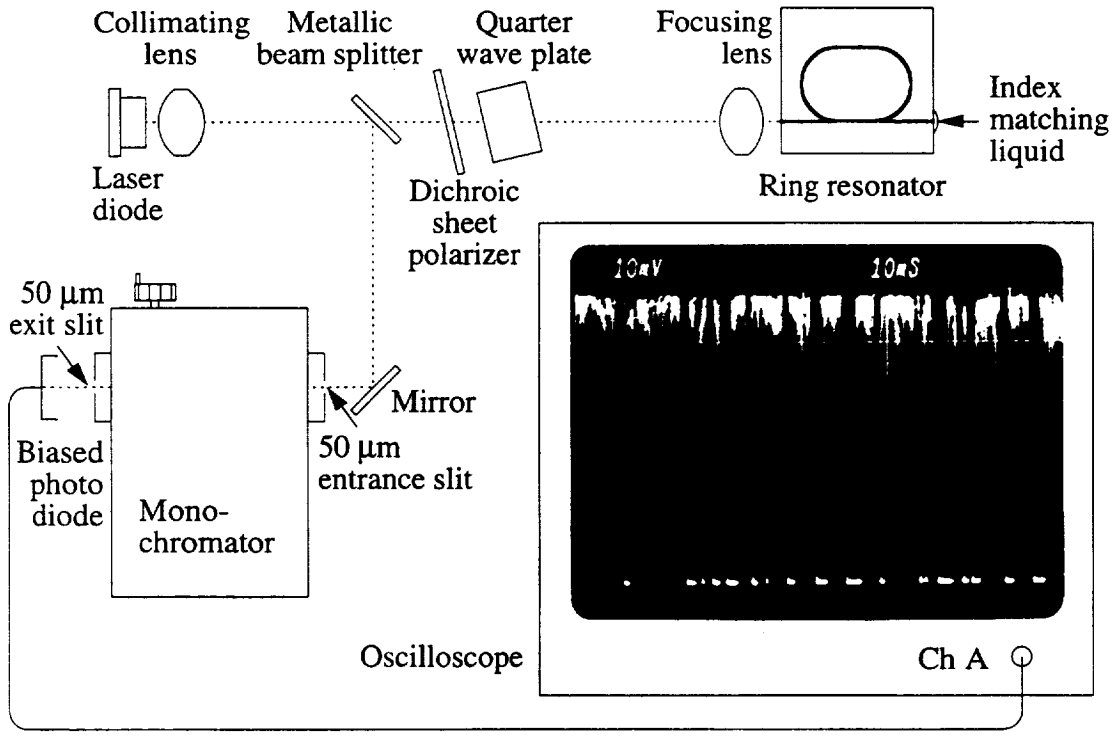


Fig. 5-7 Test setup and measurement of laser mode hopping.

The operation of the isolator is illustrated in Fig. 5-8. The linearly polarized laser diode could be oriented in any direction, but for best coupling, it was arranged so that the plane of widest divergence was parallel to the same plane for the chip. Without the isolator, this orientation would result in TE light launched into the chip. Light from the laser passes through the dichroic sheet polarizer oriented to maximally transmit the beam. At the quarter wave plate, the beam couples equally into the two eigenwaves parallel and perpendicular to the plate's fast axis, indicated by dots in the Figure. The wave polarized

along the slow axis is delayed by 90° on reaching the plate's exit face, and when it recombines with the fast wave, it creates a right hand polarized wave. If the plate were rotated by 90° about the beam axis, a left hand circularly polarized wave would be created instead. This wave couples into the channel, exciting approximately equally the quasi-TE and quasi-TM modes. Light reflected from the input edge of the chip propagates back to the waveplate, and upon passing through, the slow wave is delayed by an additional 90°. The double pass through the quarter wave plate alters the polarization similarly to a single pass through a half wave plate: the light ends up linearly polarized orthogonal to that emerging from the laser diode and is thus blocked by the polarizer.

Light which makes it through the waveguide to the end of the chip reflects off the output face, travels back to the input face, and emerges from the waveguide in a state of polarization dependent on the birefringence and length of the channel. This light will not have the same state of polarization as that directly reflected off the input face, and upon passing back through the quarter wave plate will not be as completely extinguished by the polarizer. Index matching liquid can be used to suppress reflection from the output face.

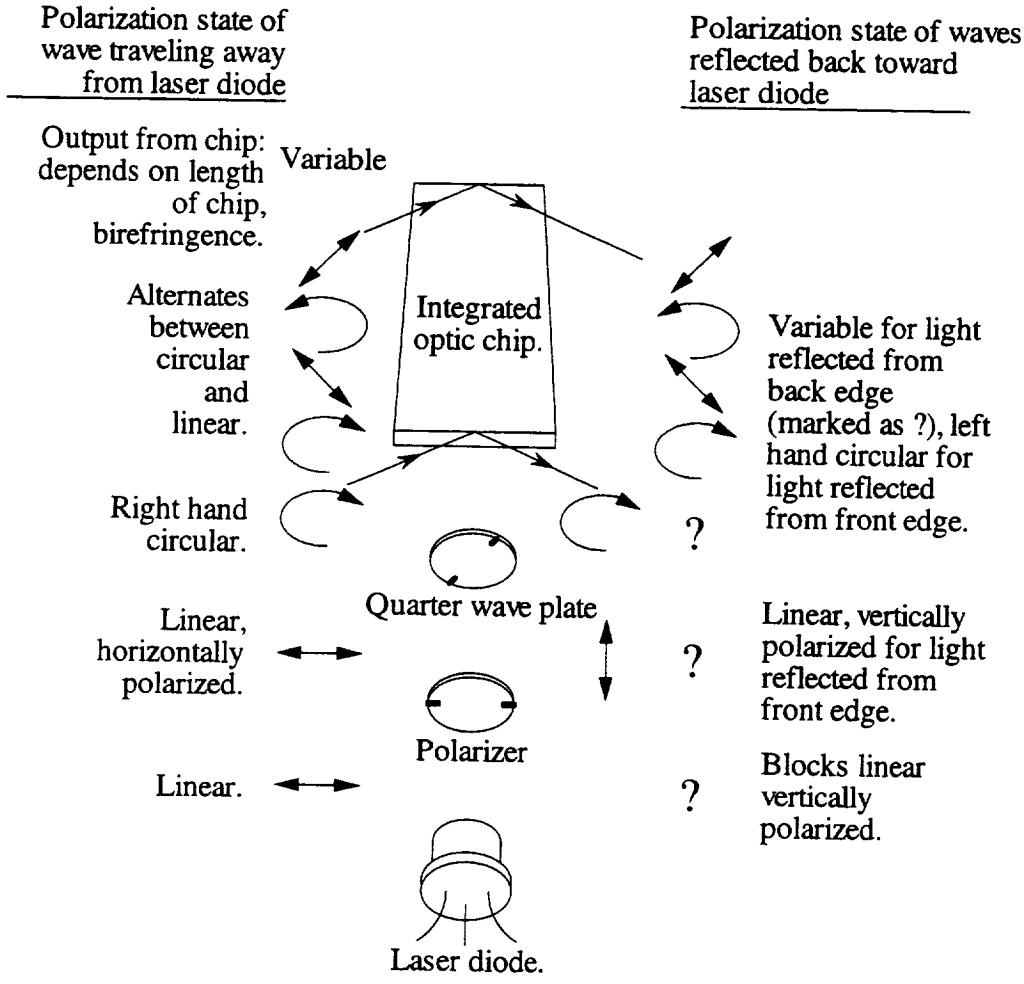


Fig. 5-8 Operation of isolator. The state of polarization is indicated by arrows when known and ? marks when it depends on the length of the sample.

The quarter wave plate was tuned in order to obtain maximum isolation by placing a mirror, cover glass beam splitter, and photodiode in the setup of Fig. 5-4 in the positions indicated in Fig. 5-9. The quarter wave plate was then rotated about its three axis to adjust the returned light's polarization angle and retardance [6] to minimize the DC signal on the DMM. To determine the degree of isolation achieved, the *s* and *p*-polarized reflectivities of the cover glass at 45° angle of incidence were measured before installation of the isolator. In addition, the maximum transmissivity of a polarizer, which would be used in front of the detector later, was also measured at the laser diode wavelength. Then with the isolator installed and the return loss maximized, the intensity

of the two reflected polarizations onto the detector positioned as shown in Fig. 5-8 was determined by placing the measured polarizer in front of the detector. By dividing these signals by the cover glass reflectivities for their respective polarizations and the polarizer transmissivity, the light incident on the cover glass for each polarization was determined. The sum of these, divided by the light incident on the mirror gave a return loss of 26 dB, while the insertion loss was measured to be 3 dB. These polarizers have been screened for extinction ratio: when crossed with polarized light at $\lambda_0 = 0.83$ μm, other similar looking types of dichroic sheet polarizers have been found to attenuate it by only 3 dB, while possessing high extinction at $\lambda_0 = 0.6328$ μm. Better polarizers have become available from Corning [7] whose Polarcor polarizer promises extinction ratios to 10,000:1 at $\lambda_0 = 0.83$ μm.

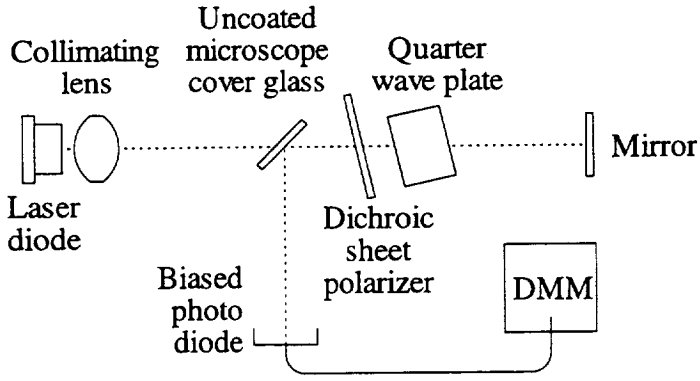


Fig. 5-9 Modification of the setup of Fig. 5-4 to adjust the quarter wave plate for maximum return loss.

Ring resonator pressure and temperature sensitivity

Although it was intended that the channel waveguide be single mode, it was found to support more than one lateral mode; the power split at the input *Y*-junction, as observed from outscattering intensities, was dependent on coupling conditions. This occurs because the modes, whose phase relationship and intensity ratio is coupling

dependent, interfere at the *Y*-junction and skew the intensity profile toward one side of the channel, thus favoring one of the *Y*-branches.

The calculated finished Δn_e from Fig. 5-1 and using 1.553 and 1.46 for the SiON and SiO_2 refractive indices was 0.0082 and 0.0087 for the quasi-TE and quasi-TM modes at $\lambda_0 = 0.83$ μm. In contrast, on a monitor wafer processed along with the device wafer and possessing the same stack of films, Δn_e of the quasi-TE and quasi-TM modes was measured to be 0.0117 and 0.0125 at $\lambda_0 = 0.6328$ μm. Calculations indicate that with these Δn_es, the channel will support two lateral modes for each polarization but not more. Calculations also indicate that the Δn_e will only change by a few percent at 0.830 μm. The maximum single mode Δn_e on a channel of this width is calculated to be 0.0079. Although there is a discrepancy between the expected Δn_e based on prism measured SiON index and thickness prior to the SiO_2 deposition and the Δn_e measured after deposition, the final prism coupling measurements do corroborate the multimode observation. In the Mach-Zehnder interferometer which was also investigated for sensing and will be reported on later in this chapter, the SiON film was made to be closer to the ellipsometric midperiod thickness to enable it to be measured using the ellipsometer and the etch depth was also measured ellipsometrically. In addition, the target etch depth was chosen to be shallower.

The sensor temperature response, shown in Fig. 5-10, was determined using a resistor heater and a forward biased diode to measure temperature, each mounted in its own hole drilled through the chuck, which was shown in Fig. 5-3. The temperature was increased to ~35°C and then allowed to slowly fall. As the device cooled, the phase slowly changed and the temperatures where the minima crossed a particular oscilloscope graticule were noted; thus temperature changes which produced a change in ϕ of $2\pi m$, where m is an integer, were recorded. Here the phase is taken with respect to that measured at the final temperature. The temperature was monitored using a Lake Shore Cryotronics DRC 82C temperature controller, which resolved to ~0.05°C. The negative

slope in Fig. 5-10 means that the path delay increases with temperature, and as was shown in the calculation at the end of Chapter 3, this was attributed to a thermally induced increase in film index with temperature. This calculation assumed the material and optical constants of SiO_2 for SiON. The magnitude of the measured temperature sensitivity normalized to the ring length is $0.13 \frac{\text{rad}}{\text{°C mm}}$ for the quasi-TE and $0.12 \frac{\text{rad}}{\text{°C mm}}$ for the quasi-TM modes. Fused silica optical fibers experience similar temperature-induced phase shifts; $0.083 \frac{\text{rad}}{\text{°C mm}}$ [8] and $0.1 \frac{\text{rad}}{\text{°C mm}}$ [9], both at a wavelength of 0.633 μm, have been reported.

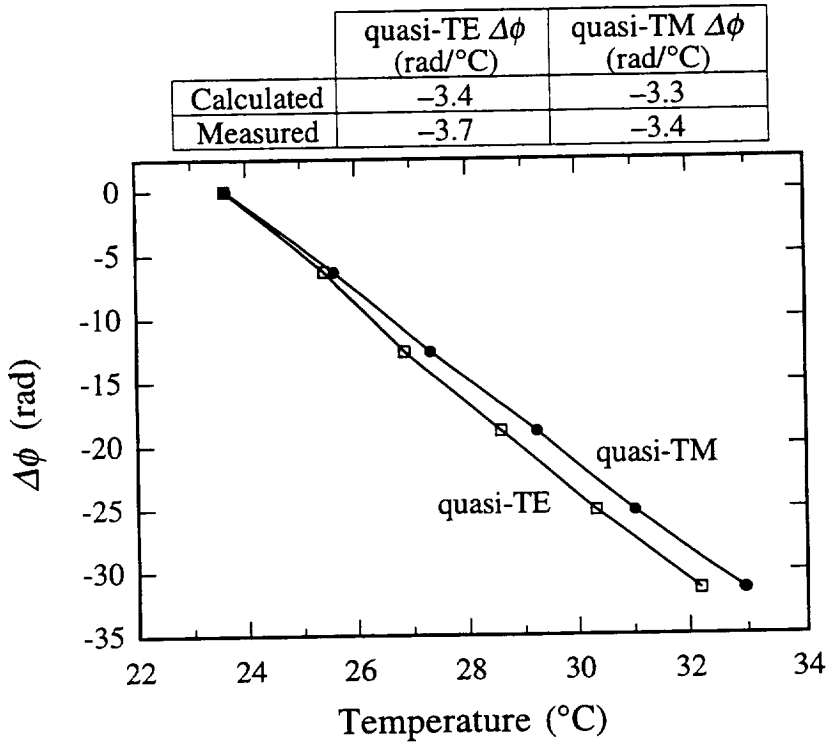


Fig. 5-10 Ring resonator temperature response. The measured values in the table above the plot were from a linear fit to the data, while the calculated were obtained by multiplying the temperature sensitivity values, calculated on a per mm basis at the end of Chapter 3, by the physical length of the ring, 28.9913 mm.

Figure 5-11 shows the pressure response of the sensor. The waveguide face was maintained at atmospheric pressure. To counter drifts caused by temperature variations

of the thermally unstabilized laser diode and device, each measurement consisted of the application of air pressure or vacuum to the back of the diaphragm, followed immediately by the measurement of ϕ; then the pressure was relieved and ϕ measured again. The difference between these two measurements, $\Delta\phi$, is plotted as a function of pressure in Fig. 5-11.

	quasi-TE $\Delta\phi$ (rad/kPa)	quasi-TM $\Delta\phi$ (rad/kPa)
Calculated	–0.0046	–0.016
Measured	–0.0004	–0.0094

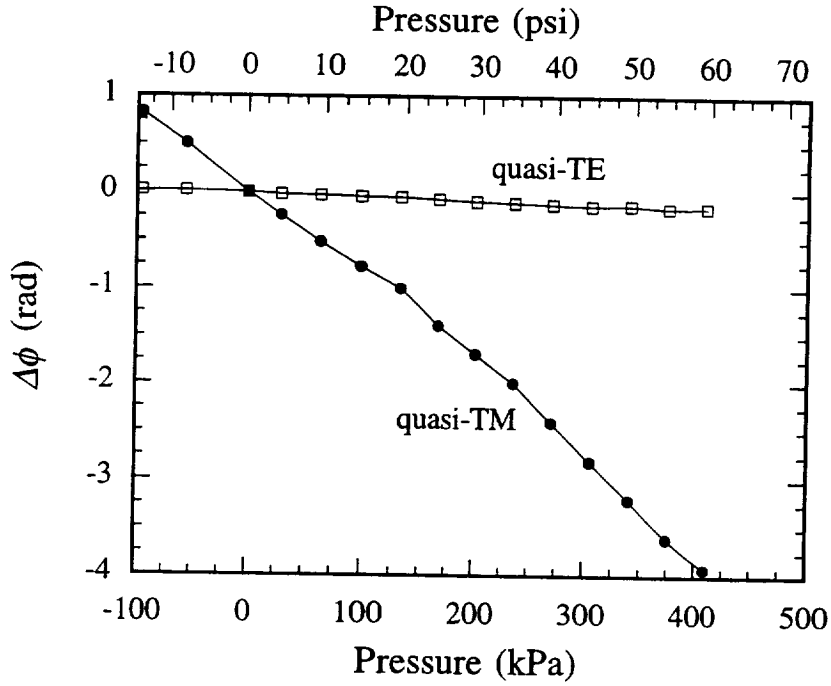


Fig. 5-11 Ring resonator pressure response. The negative slope of $\Delta\phi$ means that the optical path length of the ring increases with applied pressure.

The measured values in the table in Fig. 5-11 were obtained from a linear fit to the data, while the calculated were determined in the example at the end of Chapter 3 which assumed that the photoelastic and elastic constants of SiON are the same as those of SiO_2. Fischer *et al.* have measured the photoelastic constants of SiON deposited using LPCVD from the same gas system used in this study to form the ring resonator SiON

film ($SiCl_2H_2$-NH_3-O_2), but they deposited at a higher temperature [10]. To measure the constants, they fabricated a Mach-Zehnder interferometer that had both of its arms placed over a micromachined diaphragm in which a thick plate was left protruding from the diaphragm center. By measuring the interferometer phase change as a function of force applied to the plate using a pin, and with the aid of a finite element model, they calculated the photoelastic constants of the film. A large uncertainty was associated with these constants by the authors, however. Converting their constants to the form used in this document gave $p_{11} = 0.051$ and $p_{12} = 0.125$. Using these constants in place of those assumed for the SiON film and recalculating the pressure sensitivity gave –0.0026 rad/kPa and –0.010 rad/kPa for the quasi-TE and quasi-TM modes.

The ring resonator pressure response was observed to creep after a pressure change. In Fig. 5-12, the pressure was made to change between 0 and 60 psi with respect to atmosphere from time to time while the phase was monitored using the setup of Fig. 5-4. The phase can be seen to continue changing slowly in the same direction after the pressure transition. The creep, which had a time constant of ~5 min, was ~9% of the change induced immediately by pressure. It is believed that the creep is due plasticity in the adhesive bond or poorly bonded areas.

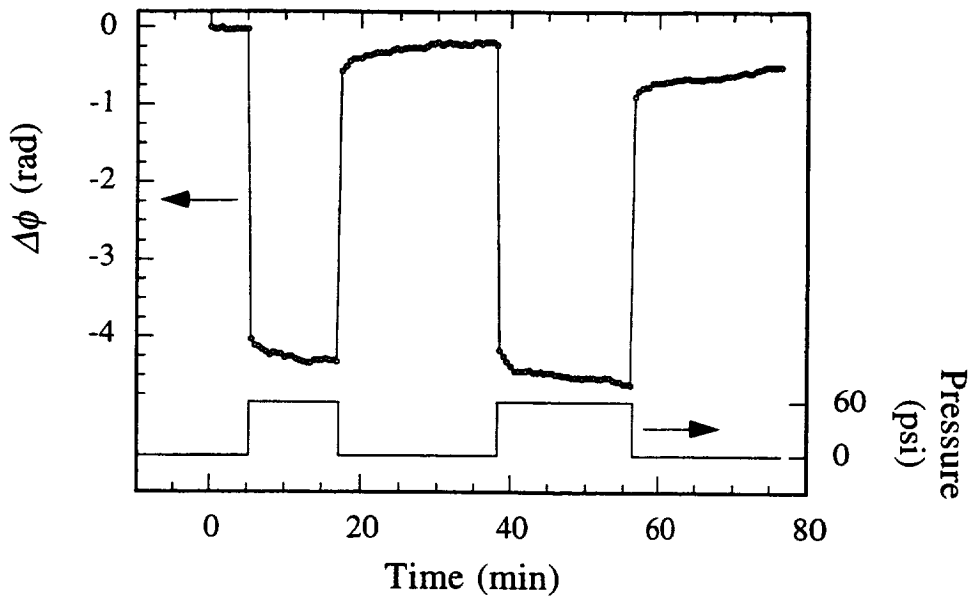


Fig. 5-12 Change in phase of the ring resonator with respect to that of the first point while the pressure alternated between 0 and 60 psi.

Several problems are apparent with the ring resonator pressure sensor. Its large optical path length (4.3 cm), although allowing easy interrogation using a swept frequency laser diode, makes the device highly temperature sensitive. However, since both modes are about equally sensitive to temperature, while the quasi-TM mode is much more pressure sensitive, both modes could be interrogated simultaneously, and pressure and temperature information recovered. Mathematically, the linear independence of pressure and temperature could allow their separation using

$$\begin{aligned} \text{Pressure (Pa)} &= A\,\phi_{TE} + B\,\phi_{TM} + C \\ \text{Temperature (°C)} &= D\,\phi_{TE} + E\,\phi_{TM} + F \end{aligned} \quad ,$$

where A-F are determined by the initial optical path length and the temperature and pressure sensitivities of the two polarizations. Although the phase can be tracked over an arbitrary change it can only be measured modulo 2π on power up, and in order to obtain absolute pressure information, the device will have to be initialized.

3 MACH-ZEHNDER PRESSURE SENSOR

The high temperature sensitivity of the ring resonator is due to its large path imbalance. An interferometer having a much smaller path imbalance is expected to have a similarly reduced temperature sensitivity. Because of minimum waveguide bend limitations, the ring resonator cannot be fabricated in the SiON waveguide technology with a much smaller path length than has been done. The Mach-Zehnder interferometer can be fabricated with a path imbalance down to zero. In the following sections the interrogation method, summarized fabrication sequence, and pressure and temperature response of a 150 μm optical path length difference Mach-Zehnder interferometer will be reported.

Interrogation method

Unfortunately, a small path imbalance Mach-Zehnder cannot be interrogated using the convenient current ramped laser diode, as its frequency excursion is too small to obtain a 2π phase sweep. Interferometer transmissivity at a fixed wavelength could be measured using a source with coherence length much longer than the path imbalance, but such a scheme would be link loss sensitive.

In the technique adopted, the interferometer is interrogated by a source whose coherence length is shorter than its optical path difference, and the resultant spectrum is wavelength separated. Such short coherence light is available from LEDs and super luminescent diodes (SLDs), with LED radiation typically having a shorter coherence length. However, due to the large emission aperture and highly divergent radiation pattern, only a small amount of LED power can be coupled into the channel. SLDs have a laser diode-like beam quality and short coherence length making them ideal for this experiment.

Figure 5-13 illustrates the interrogation method. In this example, the optical path length difference of the interferometer is greater than the coherence length of the source.

In other words, the source's spectral power density function is broader than the separation between two adjacent interferometer transmissivity minima. The signal leaving the interferometer, shown as the thick line, is the product of the interferometer transmissivity function and the source's spectral power density function. This spectrum will not be altered by wavelength independent losses in optical fibers which could be used to link the sensor.

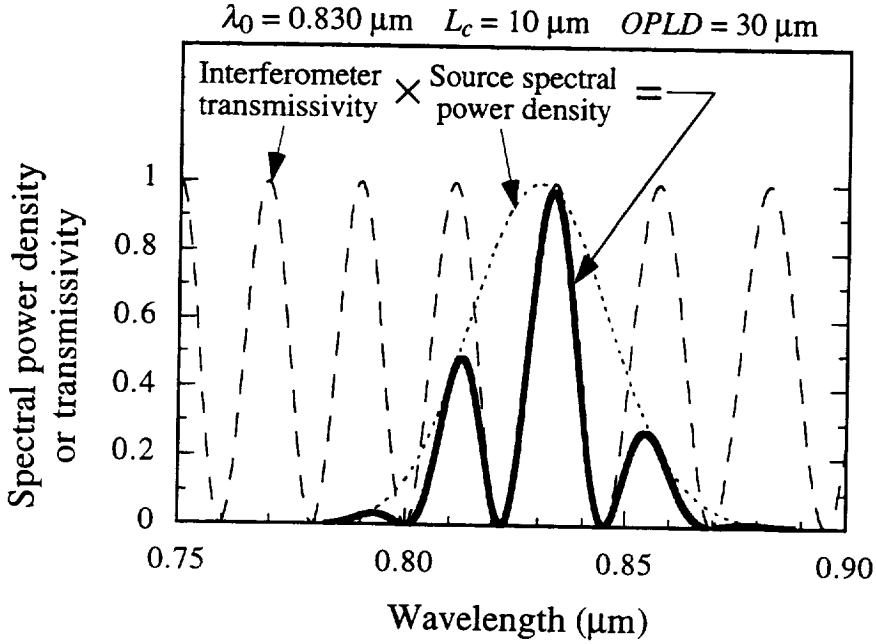


Fig. 5-13 Spectral power density from a Mach-Zehnder interferometer (thick line) with an optical path difference of 30 μm fed by a source having a coherence length of 10 μm. The curve is the product of the interferometer transmissivity function $\frac{1}{2}\left[1+\cos\left(\frac{2\pi\, OPLD}{\lambda}\right)\right]$ and the source's spectral power density function. An increase in the interferometer's OPLD causes a transmissivity maxima to shift to longer wavelengths and adjacent maxima or minima to become more closely spaced. A change in interferometer OPLD does not produce a significant change in the integrated output from the interferometer as detected by a photodiode. Rather the light must be spectrally separated and then detected.

Figure 5-14 shows the Mach-Zehnder interferometer and the dimensioned pattern on the mask used to create the device. This pattern was part of the integrated optic mask shown in the previous chapter and the Mach-Zehnder transmissivity function was derived in Chapter 3. The physical path length difference between the arms was 100 μm, and

using a nominal channel effective index of 1.5, the optical path length difference was 150 μm. The *Y*-junctions are similar to those of the ring resonator. As mentioned before, the arcs in this pattern, along with all the others on the mask, were fractured into straight segments so that the maximum error between a perfect arc and the segments was less than 0.04 μm. As with the ring resonator, the device is expected to be most pressure sensitive when one of its arms is positioned over the long edge of the diaphragm.

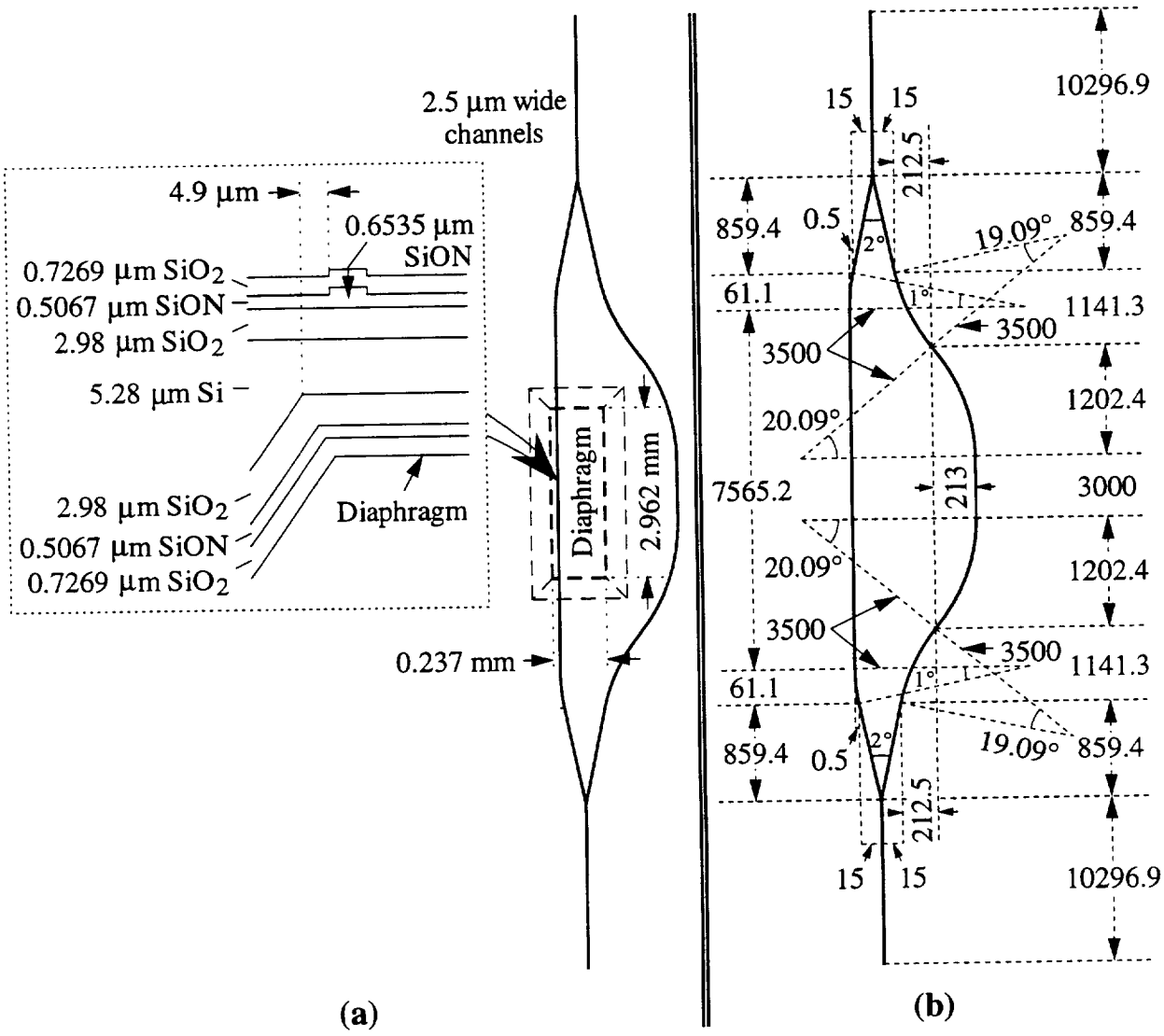


Fig. 5-14 (a) Cross section of Mach-Zehnder interferometer showing diaphragm detail, and (b) mask dimensions of the Mach-Zehnder in microns. The length of the input and output channels on the wafer is less than shown here because they were cleaved to permit end fire coupling.

Mach-Zehnder fabrication summary

The fabrication sequence was similar to that of the ring resonator and was detailed in Chapter 4. Improvements in the Mach-Zehnder processing over that of the ring resonator were that a slower etchant (Timetch) was used to form the channels, the deposited SiON and SiO_2 layers were made to be near an ellipsometric midperiod, allowing them to be measured more accurately, single mode channels were achieved, and an improved chuck was devised which used an electrostatically bonded glass interface. In addition, since the electrostatic bond cannot be reversed, the silicon component of the diaphragm was measured nondestructively with the laser transmission technique described in Chapter 4. Since this was performed prior to waveguide isolation oxide growth, the silicon thickness after oxide growth was inferred from the known consumption of 0.45 μm of silicon from each side of the diaphragm per 1 μm of oxide grown on each side [11]. The deposited oxide cladding and SiON layers had measured refractive indices of 1.451 and 1.513, as determined ellipsometrically on a monitor. The grown oxide refractive index was taken to be 1.46. Prism coupling was performed on the finished Mach-Zehnder wafer at a wavelength of 0.6328 μm and these results have been tabulated in Table 5-1 along with that calculated from Fig 5-14 and the measured film refractive indices. The Δn_e is expected to only slightly change at $\lambda_0 = 0.830$ μm where the device was used. The largest Δn_e can be for a channel of width 2.5 μm to remain single mode at $\lambda_0 = 0.830$ μm is 0.0093. The Δn_e measurement and the observation that the power splitting at *Y*-junctions was not coupling sensitive affirm that the device was single mode. The offset of 4.9 μm of the channel edge from the diaphragm edge was needed to assure that channels in other interferometers on this wafer were not off their diaphragms anywhere over their length, as determined using the infrared alignment technique described in Chapter 4. As it turned out, none of the other devices were investigated, so the alignment could have been optimized for this particular interferometer.

	Measured with prism coupling		Calculated from Fig. 5-14 and film indices in text	
	TE	TM	TE	TM
Unetched	1.4972	1.4952	1.4892	1.4881
Etched	1.4904	1.4875	1.4828	1.4815
Δn_e	0.0068	0.0077	0.0064	0.0066

Table 5-1 Measured and calculated effective indices for the Mach-Zehnder wafer at $\lambda_0 = 0.6328$ μm.

Characterization of the interferometer and source

The transfer function and fringe visibility of the Mach-Zehnder interferometer was evaluated using the setup of Fig. 5-15. Light from a GOLS 3000 super luminescent diode from General Optronics Corp. [12] was collimated and focused on the cleaved edge of the Mach-Zehnder input channel. The chip output was collimated, passed through a dichroic sheet polarizer, and focused on the input slit of a 1/4 meter Ebert design Jarrell-Ash monochromator with an 1180 grooves/mm grating and 50 μm slits. The resolution of this monochromator had previously been measured to be 0.25 nm full width at half maximum using an 0.830 μm wavelength laser diode driven above threshold. This laser diode was able to resolve fringes in the ring resonator, and so its linewidth was much less than 0.1 nm. The monochromator output was measured with a photodiode using lock-in detection. Since the SLD was weakly polarized, both quasi-TE and quasi-TM polarizations were excited in the interferometer. The SLD did not seem to be as sensitive to back reflection from the chip edge as the laser diode even under best coupling conditions and so no isolator was used. Figure 5-16 shows the measured spectrum. Fringe visibility is acceptable for the intended application. The optical path length difference was determined from the minima position of and fringe separation over 10 fringes to be 150.6 μm and 149.4 μm for the quasi-TE and quasi-TM modes, neglecting dispersion. The calculated values from Fig. 5-14, the measured film refractive indices,

and the designed physical path length difference of 100 μm is 147.9 μm and 147.8 μm for the quasi-TE and quasi-TM modes.

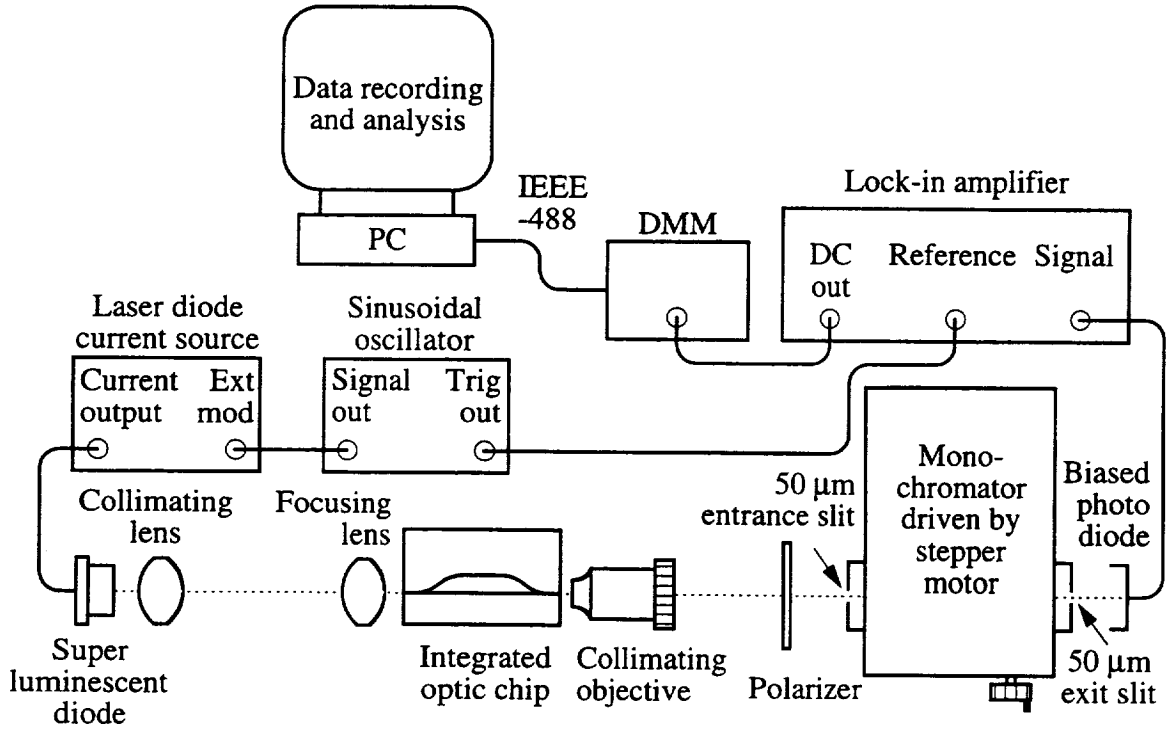


Fig. 5-15 Setup for acquiring Mach-Zehnder transmissivity spectrum.

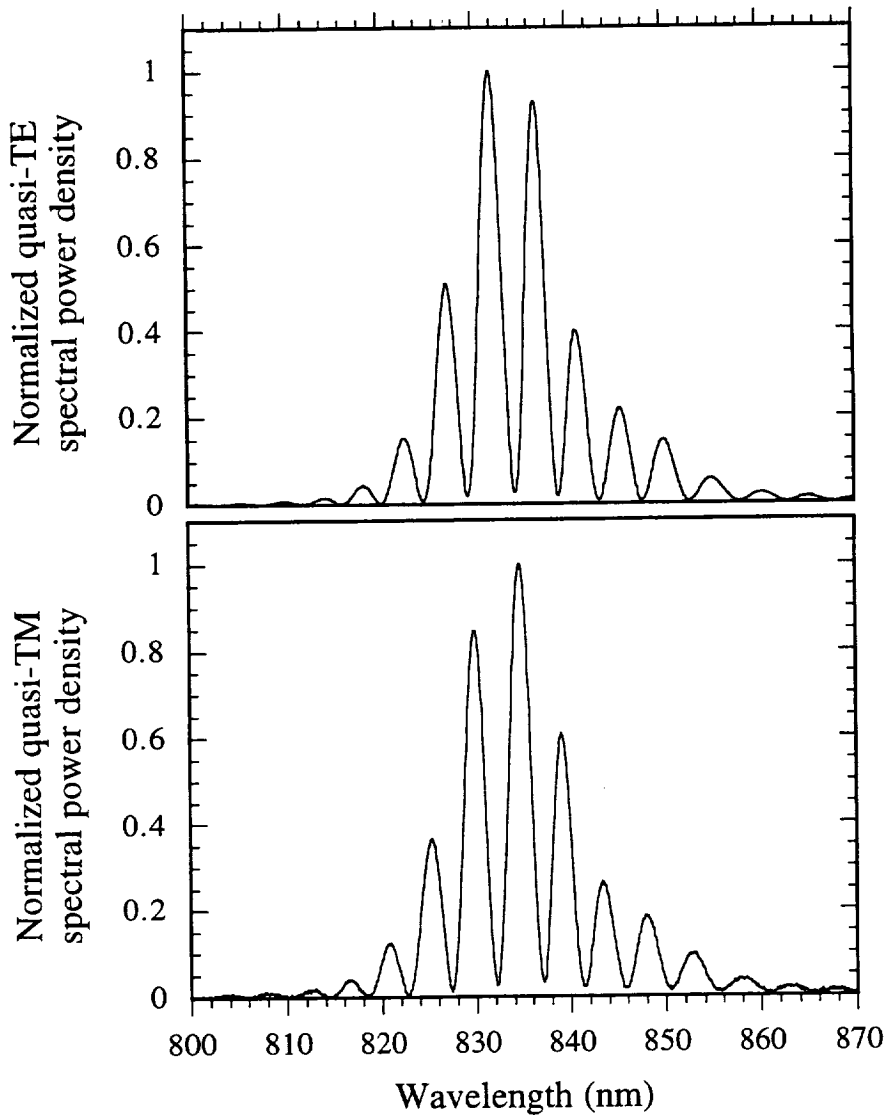


Fig. 5-16 Spectral power density as measured by the monochromator in Fig. 5-15. In the worst case, the minimum between the two larger peaks is 2.7% of the largest peak.

Coupling lenses designed for use with laser diodes might be employed to collimate and focus the SLD light into the channel waveguide. However, such optics were found to have excessive chromatic aberration. Figure 5-17 shows the spectrum obtained from a straight channel waveguide for two coupling conditions using the setup pictured in Fig. 5-15, except that a chopper was used rather than pulsing the source. In the upper plot of Fig. 5-17, a molded aspheric laser diode lens designed for nominal

780 nm wavelength (Thor Labs [13] C230TM-C) was used to collimate and focus the SLD beam onto the chip edge, while in the lower plot, collimation was performed using a 20× objective and focusing using a 40× objective. In the former case the spectral power density changes when the sample is lowered because the beam waist for wavelengths removed from the peak is displaced (chromatic aberration), and so its spot size is larger at the end face of the chip and better coupled than the smaller waist beam which misses the guide. Using objectives to incouple reduces this problem because they are achromatized. Because the intended demodulation technique is sensitive to source spectral power density changes, the guide was incoupled using objectives when characterizing the sensor's pressure and temperature sensitivity.

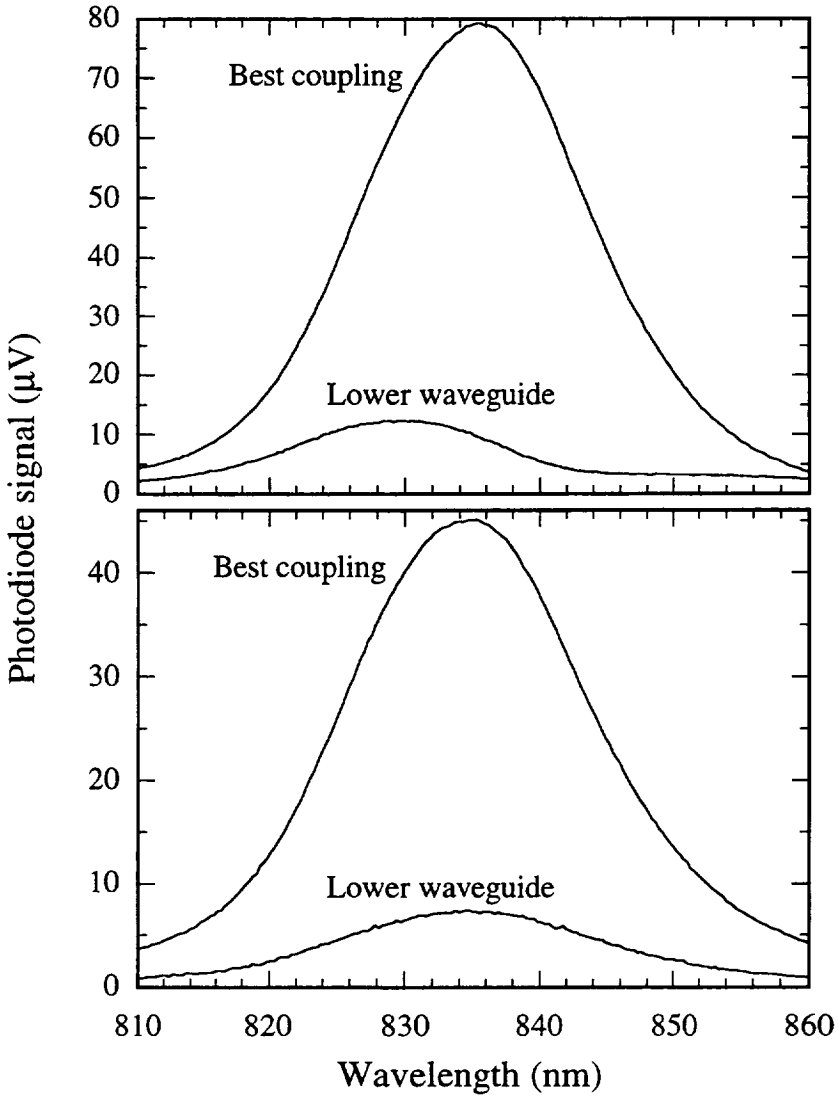


Fig. 5-17 Spectral power density of light emerging from a straight channel. Top plot: Aspheric lenses are used to collimate and focus. Bottom plot: Collimation was performed using a 20× objective and focusing with a 40× objective.

The normalized throughput of the Mach-Zehnder interferometer was measured by taking the ratio of the totality of the spectrum leaving the interferometer to that leaving a nearby straight channel when each separately was coupled into maximally with SLD light. The setup of Fig. 5-15 was used, except that the SLD beam was chopped rather than the source modulated, and the monochromator was omitted so that the entire SLD spectrum, modulated by the transmissivity function of the Mach-Zehnder, could be

measured directly with the photodiode. For our case where the coherence length of the source is shorter than the optical path length difference of the interferometer, one half of the SLD power coupled into the guide is theoretically transmitted through the interferometer. The output intensity from the Mach-Zehnder was measured to be 37.8% of that from a nearby straight channel. Thus, the transmissivity of the bends and two *Y*-junctions, neglecting the background propagation loss also present in straight channels, is 37.8% / 50% = 76%.

Mach-Zehnder pressure and temperature sensitivity

Both the sign and magnitude of the optical path length difference can be determined using the proposed interrogation technique. For the ring resonator, the phase of the light in the ring where it recombined at the *Y*-junction with the input channel was negative because of the ring's path delay. For the Mach-Zehnder, the phase will be defined as

$$\phi = \phi_{short\ arm} - \phi_{long\ arm} \quad , \tag{5.3.1}$$

where the phase in each arm is given by $\phi_{arm} = \frac{-2\pi}{\lambda_0} \int_{arm\ length} n_{et}\, d\ell$. The phase then becomes

$$\phi = \frac{2\pi}{\lambda_0} \left(n_{et} \ell_{long\ arm} - \int_{\ell_{short\ arm}} n_{et}\, d\ell \right) \quad , \tag{5.3.2}$$

where integration is taken over the length of the short arm, including the diaphragm, and the integral term in the phase of the long arm has reduced to $n_{et}\ell_{long\ arm}$ because the effective index of this channel segment does not contain a diaphragm and so is constant. The transmissivity of the Mach-Zehnder interferometer is then found from $\frac{1}{2}\left[1 + \cos(\phi)\right]$.

When no pressure is applied, (5.3.2) reduces to

$$\phi = \frac{2\pi n_{et}}{\lambda_0} \left(\ell_{long\ arm} - \ell_{short\ arm} \right) = \frac{2\pi}{\lambda_0} OPLD \quad , \tag{5.3.3}$$

where *OPLD* is the optical path length difference. A shift of a transmissivity minimum to longer wavelengths means that both the optical path length difference and phase have increased.

Although the sensor's response could have been investigated using the setup of Fig. 5-15, where the transmitted spectrum would be measured for each temperature or pressure, this was not done. Acquiring a spectrum in this manner took too long (30 min) to see potential transients such as creep, and slight shifts in the spectrum (0.1 nm) could be due to a difference in synchronization between the computer reading the photodiode and the monochromator drive. Instead, $\Delta\phi$ was measured in a link insensitive manner with the setup of Fig. 5-18 which used the same SLD and monochromator described before. Figure 5-19 shows a photograph of the setup and Fig. 5-20 shows a close up view of the chuck.

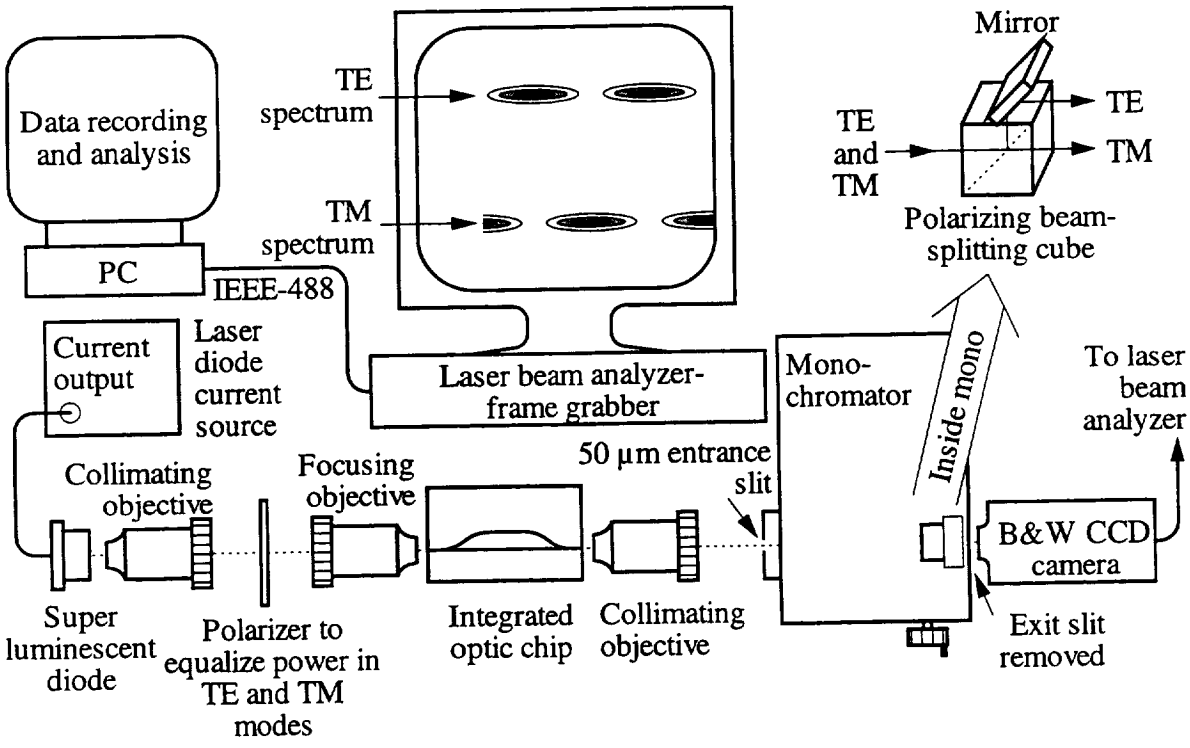


Fig. 5-18 Setup to simultaneously measure quasi-TE and quasi-TM $\Delta\phi$ as a function of temperature or pressure.

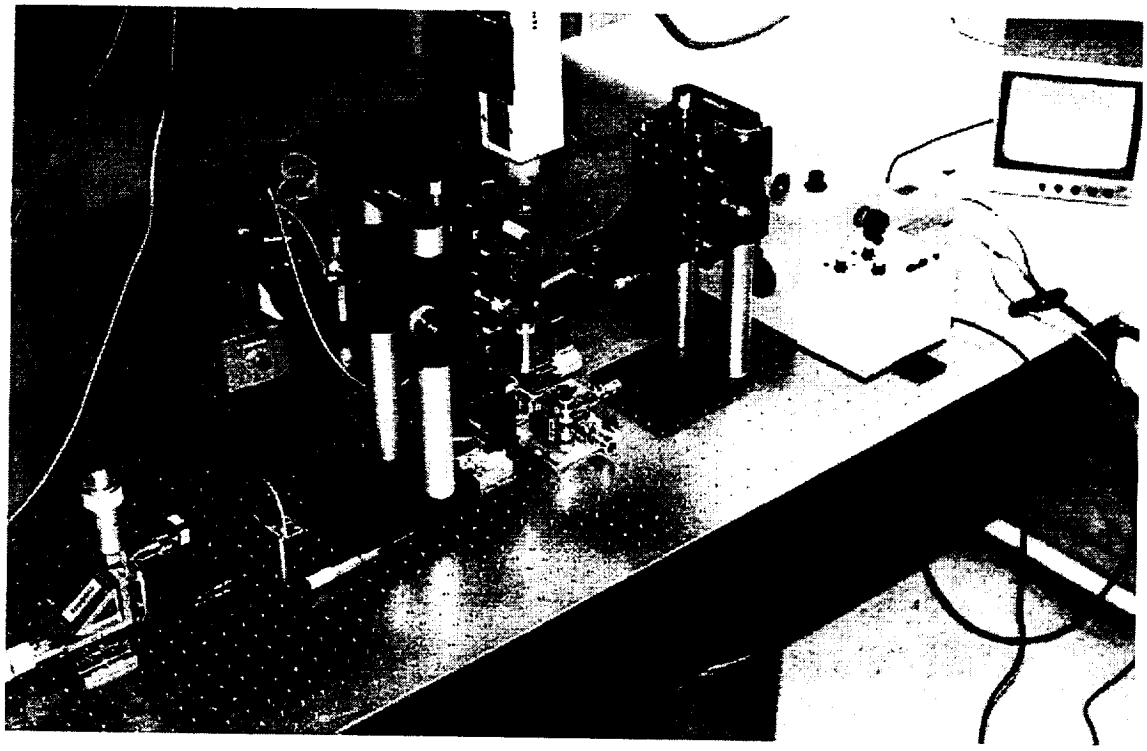


Fig. 5-19 Photograph of the setup in Fig. 5-18. The laser beam analyzer is not visible.

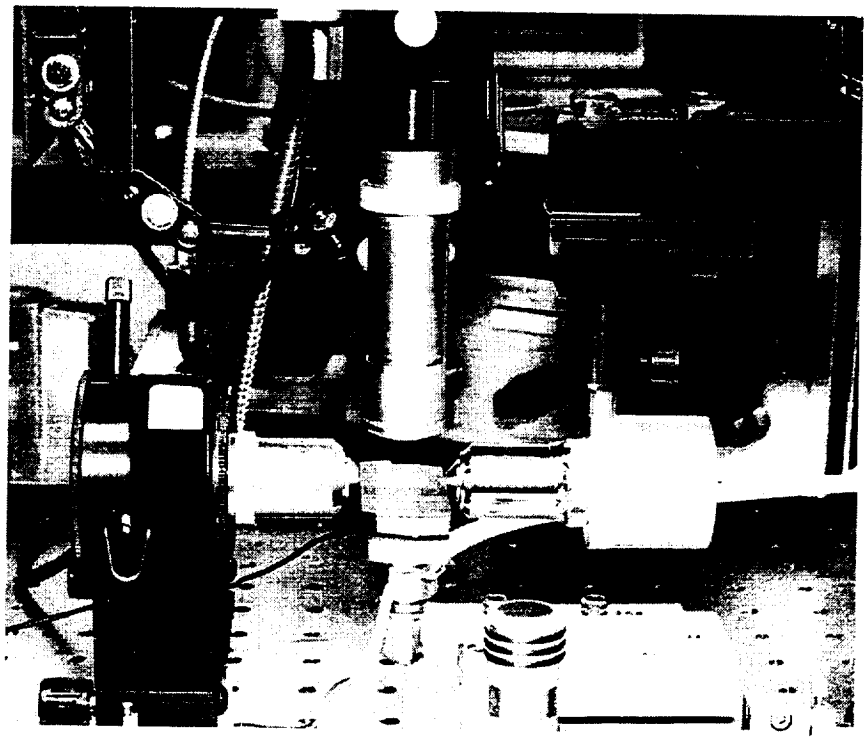


Fig. 5-20 Close up view of the wafer mounted on the chuck.

Light from the weakly polarized SLD was collimated by a 20× objective and passed through a dichroic sheet polarizer which was rotated to equalize the intensity of the quasi-TE and quasi-TM modes leaving the sensor. The chip was attached to the chuck, elbow, and support described in Chapter 4, and these were clamped to a stainless steel crossed-roller bearing stage (ULTRAlign [14]). This stage was equipped with three high precision adjusters (DS-4) which had negligible backlash. A flexible hose connected the elbow holding the chuck to the air manifold system, as shown in Fig. 5-5. A resistor and a thermocouple were glued to the underside of the chuck for temperature measurements. Incoupling and outcoupling from the chip was performed using 40× microscope objectives. The collimated beam from the chip overfilled the 50 μm entrance slit of the monochromator. Once inside the monochromator, the light was spectrally separated into a fan parallel to the horizontal plane of the TV camera. Located inside the monochromator and recessed from the plane where the slit would have been positioned was a 1/2", broad band, polarization-splitting cube which was designed for use over a 570 nm-890 nm wavelength range. A small first surface mirror was fixtured above the cube. The spectrally separated fan of light of TM polarization (polarized perpendicular to the top of the cube) passed through the cube, while the TE polarized component of the fan was reflected from the cube's coated diagonal onto the first surface mirror, which bent it parallel to the TM fan. This arrangement produced two fans of light, one above the other, which were separated by 1.6 mm. Both fans were simultaneously detected by a monochrome, windowless, CCD camera (Cohu 4915-2010/0000) which was read out using a laser beam analyzer frame grabber (Spiricon LBA-100). The camera and analyzer had previously been characterized for linearity by focusing the 830 nm wavelength collimated beam which emerged from a single mode channel waveguide alternately on the camera's CCD array and then on a photodiode, which was measured with lock-in detection. By comparing the camera's peak pixel reading to the photodiode current for various degrees of laser attenuation, the camera was determined to be linear.

The laser beam analyzer displayed the camera's images in false color contour plots. At the push of a button, the LBA-100 could be set to either continuously acquire images and display them at about 20 frames/sec or to acquire one or more frames. The instrument had 32 frame buffers, each of which could store a 120×120 array of pixels, with each pixel's intensity quantized to 8 bits (256 gray levels). The camera had at least four times as many pixels in a row and in a column as an LBA-100 frame. Three levels of zoom were available: 1) every pixel was taken, 2) every other pixel in a row and every other row was taken, or 3) every fourth pixel in a row and every fourth row was taken. Pixels were acquired until a 120×120 block was filled regardless of zoom. The area of the camera's CCD from which the block of pixels was taken could be shifted.

Characterization of the sensor was performed in the middle zoom where the sampled pixel pitch was 19.6 μm both vertically and horizontally. As shown in Fig. 5-18, the spectrum from both polarizations was monitored simultaneously, and the wavelength range covered by the 120 pixels was about 7.5 nm, with lower numbered pixels corresponding to longer wavelengths. With such a wavelength span, slightly over 1.5 fringes could be observed.

When pressure was applied to the sensor in the configuration of Fig. 5-18, with the analyzer continuously acquiring the camera's images, the lower set of spots on the display monitor, corresponding to wavelengths of maximum Mach-Zehnder transmissivity, were observed to move toward the right of the screen. Those above also moved in the same direction, but translated much less. Adjusting the monochromator to shorter wavelengths moved the spots in the opposite direction as the pressure increased. According to (5.3.2), since λ_0 had to decrease to stay on a particular fringe (maintain ϕ constant), the integral must have increased to reduce the difference in parenthesis: n_{et} of the channel segment over the diaphragm edge must have been increased by the pressure increase. The direction of change agrees with that calculated in Chapter 3.

Before measuring the sensor's pressure response, the setup of Fig. 5-18 was first used with a straight channel on the same sample. Here the camera, cube, and mirror were adjusted so that the fan of light in both polarizations was parallel to a horizontal row of camera pixels. The monochromator was adjusted so that the wavelength corresponding to the spectral peak from the SLD was approximately centered on the monitor screen and the SLD drive current and polarizer were also adjusted to equalize the TE and TM intensities and maintain them slightly below LBA-100 saturation. The straight channel spectrum was then acquired; a 120×120 screen of pixels was taken and this down loaded to computer. Without altering the SLD drive current or monochromator settings, the Mach-Zehnder was positioned and coupling optimized. This device's additional loss insured that the LBA-100 would not be saturated. Then spectra of the Mach-Zehnder were similarly acquired, each in its own frame buffer, as the pressure in psi was stepped through 0, 20.0, 40.0, 60.0, 80.0, 60.0, 40.0, 20.0, 0, and then another 0 was taken. Afterward, the frames were down loaded to computer. In order to remove any offset in the camera, a background frame taken with the SLD beam blocked was pixel-by-pixel subtracted from all frames.

As the sketch on the monitor screen in Fig. 5-18 suggests, most of the image contains pixels not illuminated by spectral fans of light from the monochromator. Two horizontal stripes, each twelve rows wide, were found to contain the TE and TM signals, with the rest containing noise. Fig. 5-21 plots the column averages of all twelve rows for both the TE and TM stripes, normalized by averages computed similarly on the straight channel for the pressures stated before.

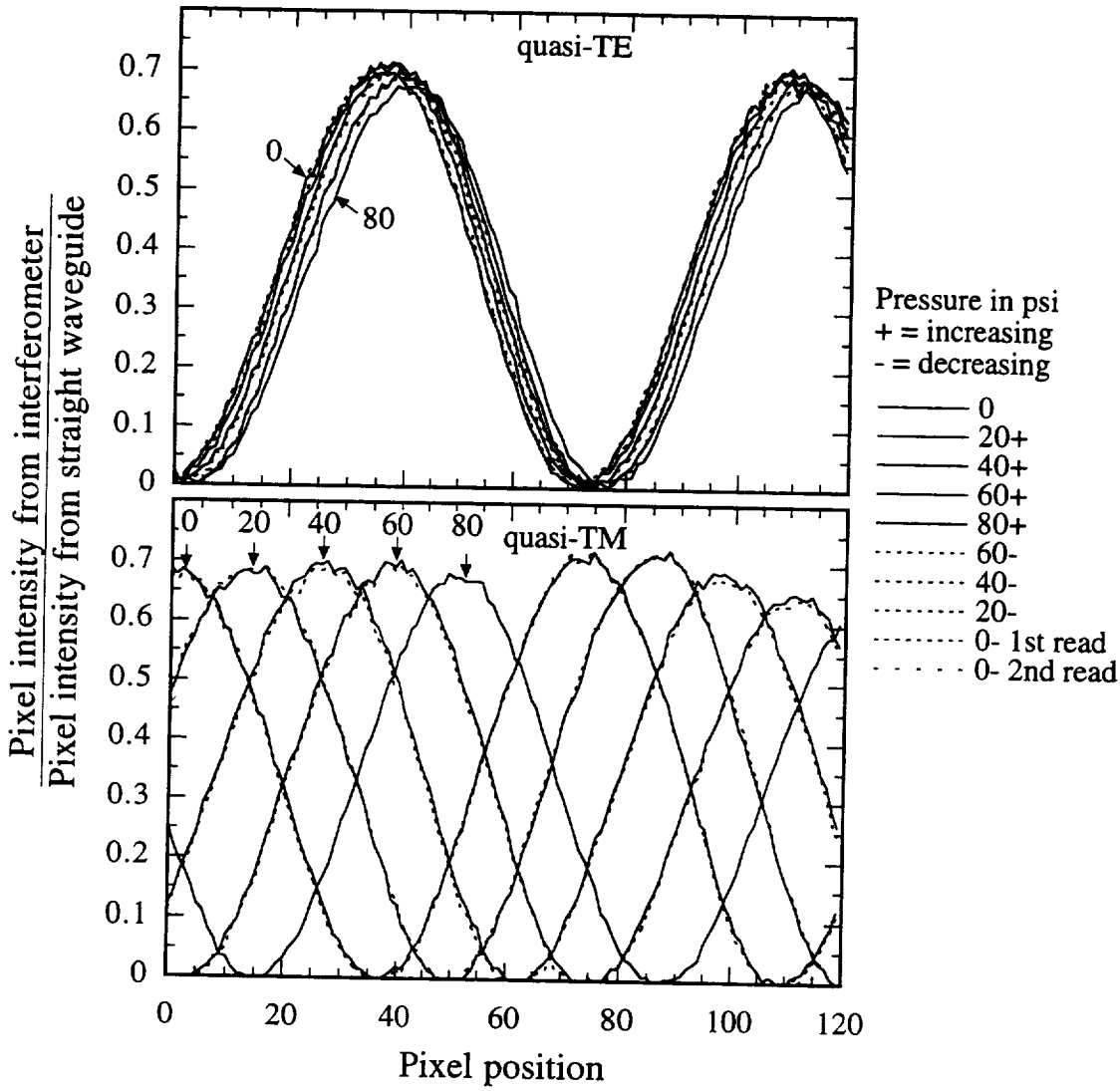


Fig. 5-21 Mach-Zehnder interferometer spectrum normalized to that from a straight channel as a function of pixel position (wavelength) for the indicated pressures.

The discrete Fourier transform was used to determine the phase shift. The algorithm finds the frequency spectrum $X(k)$ using

$$X(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n\,\Delta t) \exp(-j2\pi nk/N) \qquad k = 0, 1, .., N-1 \quad , \qquad (5.3.4)$$

where k is the frequency index, Δt the sampling interval, and $x(n\,\Delta t)$ the sampled data. The multiplier $1/N$, often omitted in DFT definitions, is required to obtain the proper

magnitude of $X(k)$. The real-valued time domain function can be recovered from the positive frequency spectra only using

$$x(t) = X(0) + 2\,|X(1)|\cos\left(\frac{2\pi t}{T} + \angle X(1)\right) + 2|X(2)|\cos\left(\frac{4\pi t}{T} + \angle X(2)\right)\ldots \quad , \quad (5.3.5)$$

where $\angle X(\cdot)$ is the phase. The algorithm will return a set of sinusoids which pass through all of the data points. This algorithm assumes that the function is periodic and that it has been sampled over precisely one period. If the algorithm were run on the full 120 point width of one TE spectrum in Fig. 5-21, it would assume that the function was periodic and that the next pixel following #119 was #0. It would therefore return a set of sinusoids which had a much high frequency content than actually present to bridge the discontinuity.

A nonlinear, iterating, data fitting algorithm was used to fit the function $y = a + b\cos\left(\frac{2\pi x}{c} + d\right)$ to several spectra to determine the period c, where x was pixel position, y was intensity, and a-d were coefficients to be determined. Although this algorithm could have been used directly to determine the phase, it was desired to use the DFT since it is more efficient, it could determine the phase in a fixed number of operations (no iteration), and it does not require starting values--advantages in an automatic phase finding algorithm. These fits indicated that the period of the TE and TM data sets was 73 pixels. To find the phase, the DFT algorithm was run over the middle of the measured set from pixel 23 to 95.

Fig. 5-22 shows the frequency spectrum taken by using pixels 23 through 95 from the TM curve "60+" in Fig. 5-21. Here the magnitudes $X(0)$, $2|X(1)|$, $2|X(2)|$... have been plotted. The fundamental, at the frequency $\frac{1}{73}$ cycles / pixel (or with a period of 73 pixels / cycle) is 48 times larger than the next harmonic.

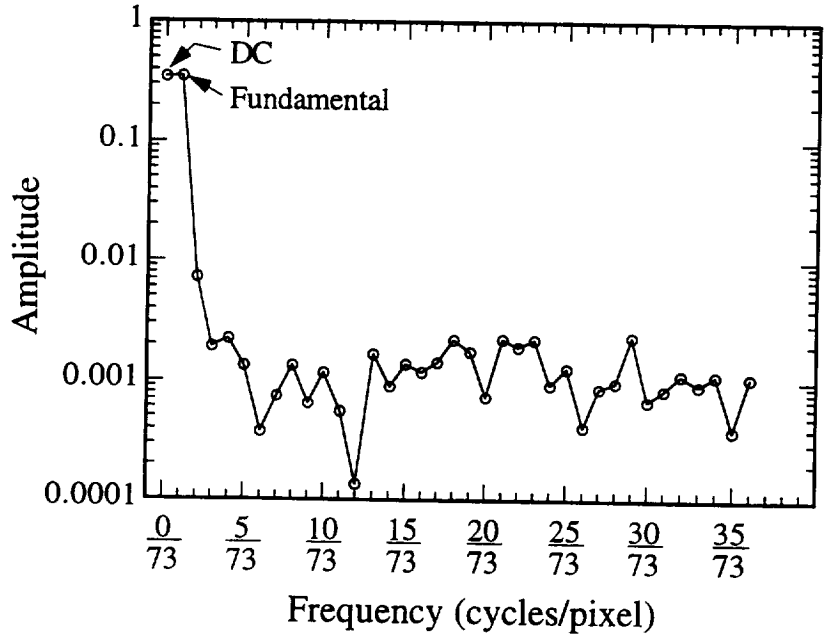


Fig. 5-22 Frequency spectrum of the TM curve "60+" in Fig. 5-21.

Using the DFT algorithm, the phase of the spectra in Fig. 5-21 was determined and they are plotted in Figure 5-23 with respect to the first spectrum taken at 0 psi. To allow comparison with theory, the sign of the phase change must be chosen to agree with the definition given in (5.3.2). As it turned out, however, the sign given by the DFT algorithm on the camera data agreed with that defined in (5.3.2). Hysteresis is seen to be negligible and linearity good.

The theoretical values were calculated at a wavelength of 0.830 μm from the cross section of Fig. 5-14 and the refractive indices provided in the Mach-Zehnder fabrication summary section in this Chapter. The calculation was similar to the example one at the end of Chapter 3 which used the ring resonator cross section and refractive indices. With the Mach-Zehnder, the larger separation of the channel center from the diaphragm edge was taken into account by calculating the beam strain 6.15 μm from the built-in edge, while the reduction of strain in the vicinity of the short edge was included by assuming that the waveguide saw the beam strain over only 2962 μm - (82 μm)(2) = 2798 μm length. The theoretical and calculated values compare as well as they did for the ring

resonator, again suggesting that the photoelastic constants for the LPCVD deposited SiON differ from those of SiO_2. The quasi-TE sensitivity is significantly less than the quasi-TM.

	quasi-TE $\Delta\phi$ (rad/kPa)	quasi-TM $\Delta\phi$ (rad/kPa)
Calculated	–0.0032	–0.012
Measured	–0.00076	–0.0079

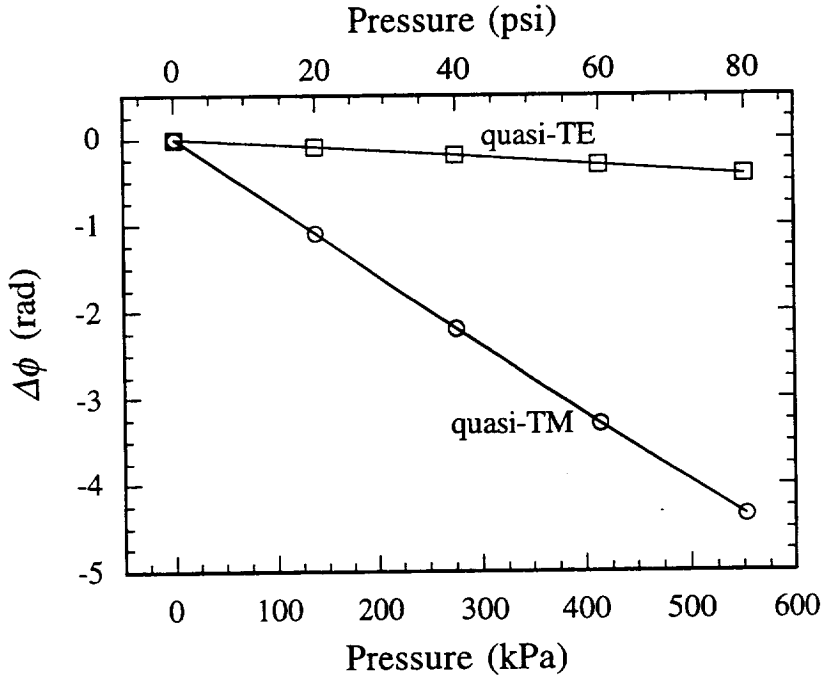


Fig. 5-23 Mach-Zehnder pressure response. Ten phase points $\angle X(1)$ were taken from the quasi-TE spectra and ten from the quasi-TM spectra of Fig. 5-21. These are plotted with respect to the first phase measured at 0 psi for their polarization. Multiple phase measurements taken at the same pressure fall on top of each other.

In order to determine the level of creep and noise in the sensor configuration, spectra were taken at 30 sec intervals, and in between two frames, the pressure was changed from 0 to 60.0 psi and then returned to 0 psi again in between two subsequent frames, as shown in Fig. 5-24. The same techniques discussed before were used to determine the phase change. The digital pressure gauge showed that the regulator permitted the pressure to relax by 0.6 psi from its initial setting of 60.0 psi. This relaxation was detected by the Mach-Zehnder. The initial pressure and phase measured

immediately after the pressure increase are indicated by arrows at the end of the pressure application; both the phase and pressure have declined similarly. Fig. 5-25 plots the phase measured after its return to 0 psi. No creep is seen and the level of noise is less than 0.01 rad.

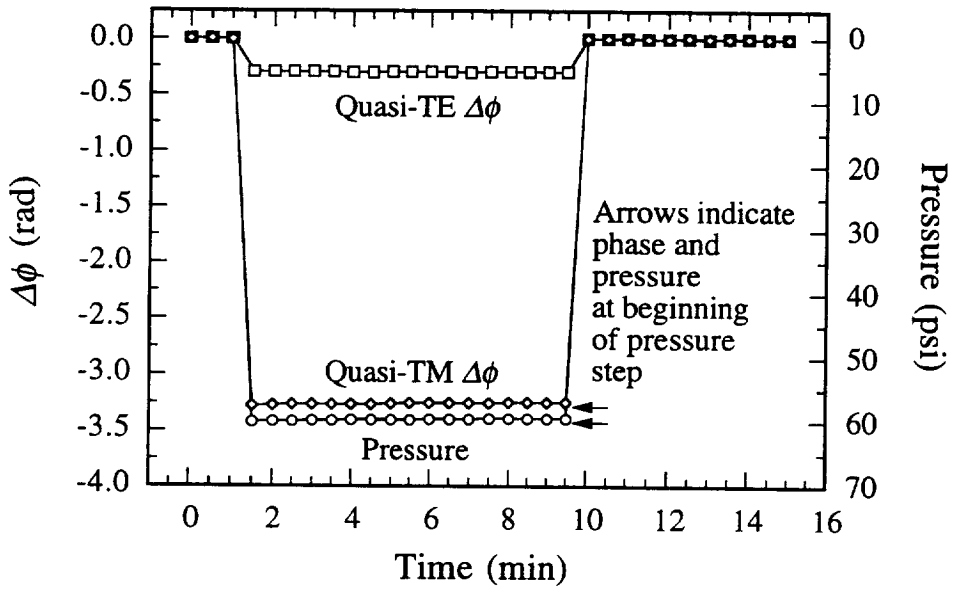


Fig. 5-24 Response of the Mach-Zehnder sensor to a pressure step.

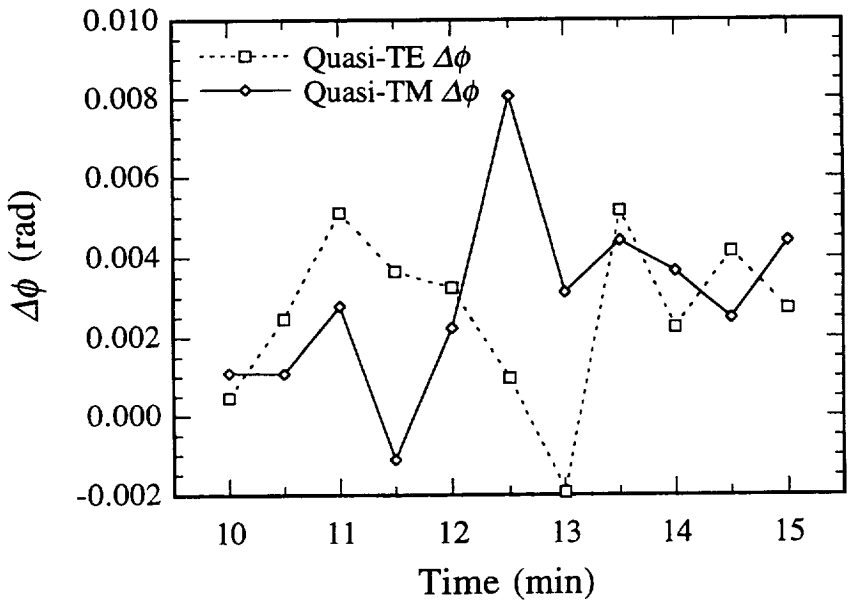


Fig. 5-25 Expanded view of Fig. 5-24 after the pressure has been set to 0 psi, showing the noise level in the sensor configuration.

Fig. 5-26 shows the temperature response of the sensor, which was evaluated using a resistor heater and thermocouple mounted on the back of the chuck. The quasi-TM polarization is seen to be about twice as sensitive as the quasi-TE.

	quasi-TE $\Delta\phi$ (rad/°C)	quasi-TM $\Delta\phi$ (rad/°C)
Calculated	0.118	0.114
Measured	0.137	0.243

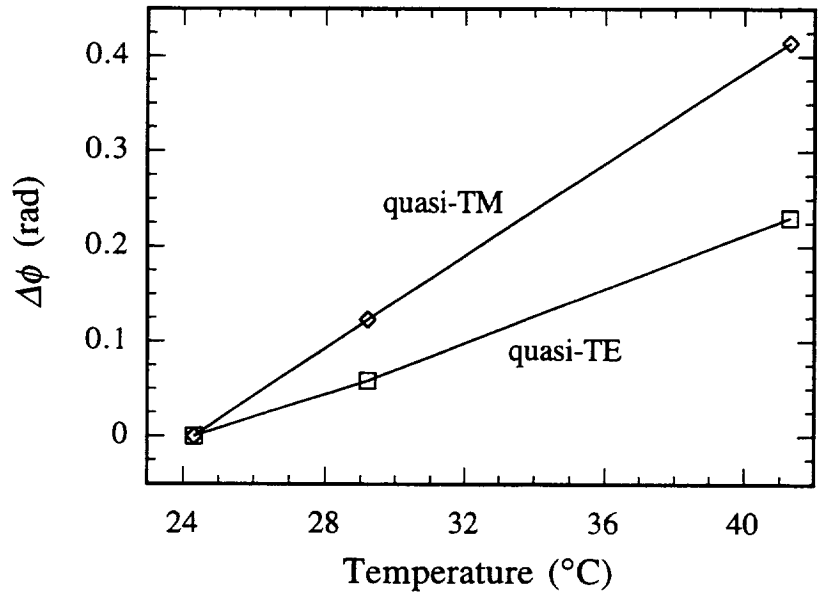


Fig. 5-26 Response of the Mach-Zehnder sensor to temperature change.

4 CONCLUSIONS

Two integrated optical pressure sensors have been fabricated. The ring resonator has a large enough path length difference to permit it to be interrogated using a swept frequency laser diode. This technique can determine the sign and magnitude of the change in path length even in the presence of variations in link transmissivity. In order to maintain a stable phase, the laser diode must be temperature controlled and isolation is required between the laser and chip. While the two polarizations have a large pressure sensitivity difference, but possess a similar sensitivity to temperature, thus allowing the two effects to be separated, the ring resonator will have to be initialized if only one laser diode wavelength is used, as its phase changes by 2π about every 2°C. Both the nearly balanced Mach-Zehnder interferometer and its source are much less temperature sensitive, and so the initialization requirement could be removed for suitable combinations of OPLD and limitations on temperature swing, such that the phase is

always confined to within 2π. This small phase swing is required since neither for the 150 μm OPLD Mach-Zehnder nor for the ring resonator can absolute fringe order be determined using the proposed interrogation techniques. Some temperature sensitivity may be desirable however, because temperature induced changes in parameters such as Young's modulus could then be accounted for.

Both interrogation techniques render the sensor insensitive to changes in link loss. For example, in Fig. 5-21, the 80 psi pressure spectra is slightly lower than the rest, perhaps because coupling into the chip decreased due to stiffening of the air hose. However, this did not affect the accuracy of the measurement. In a similar manner, coupling changes to the ring resonator will not affect its measurement.

Because the pressure response is polarization sensitive, single mode polarization preserving fibers will be required either on the output or input ports. When using a non polarization preserving fiber to incouple, polarization fading could be prevented by launching circularly polarized light into this fiber. Only on the output port would a polarization preserving fiber be needed. The electrostatically bonded glass disc could be sized to overhang the chip edge and function as a sturdy support for the fibers.

The electrostatic bond was found to remove the creep problems and difficulties with getting adhesive on the diaphragm back encountered when the wafer was directly glued to the chuck.

5 REFERENCES

[1] D. A. Jackson, A. D. Kersey, M. Corke, and J. D. C. Jones, "Pseudoheterodyne Detection Scheme for Optical Interferometers," *Electron. Lett.*, vol. 18, pp. 1081-1083, 1982.

[2] B. Masterson and J. Webb, "Better Diode Lasers Will Expand Applications in Many Areas," *Photonics Spectra*, pp. 119-124, May 1992.

[3] L. A. Johnson and S. C. Jensen, "Problems and Approaches for Remote Fiber Optic Absolute Sensors," presented at Fiber Optic and Laser Sensors III, San Diego, CA, and published in proceedings of SPIE, vol. 566, pp. 45-53, 1985.

[4] D. A. Jackson, "Monomode Fibre Optic Interferometers and their Application in Sensing Systems," in *NATO Advanced Study Institute on Optical Fiber Sensors*, vol. 132, A. N. Chester, Editor, Dordrecht: Martinus Nijhoff, 1987, pp. 13.

[5] W. Gleine and J. Muller, "Low-Pressure Chemical Vapor Deposition Silicon-Oxynitride Films for Integrated Optics," *Appl. Opt.*, vol. 31, pp. 2036-2040, 1992.

[6] *Optics Guide 5*: Melles Griot Corporation, pp. 14-28 to 14-30, 1990.

[7] Corning Incorporated, Corning, NY 14831.

[8] G. B. Hocker, "Fiber-Optic Sensing of Pressure and Temperature," *Appl. Opt.*, vol. 18, pp. 1445-1448, 1979.

[9] P. A. Leilabady, "Optical Fiber Point Temperature Sensor," presented at Fiber Optic and Laser Sensors V, San Diego, CA, and published in proceedings of SPIE, vol. 838, pp. 231-237, 1987.

[10] K. Fischer, J. Muller, R. Hoffmann, F. Wasse, and D. Salle, "Elastooptical Properties of SiON Layers in an Integrated Optical Interferometer Used as a Pressure Sensor," *J. Lightwave Technol.*, vol. 12, pp. 163-169, 1994.

[11] S. M. Sze, *Physics of Semiconductor Devices*, 2nd edition: John Wiley, pp. 66, 1981.

[12] General Optronics Corp., 2 Olsen Ave., Edison, NJ 08820.

[13] Thor Labs, Inc., P.O. Box 366, Newton, N.J. 07860.

[14] Newport Corporation, 1791 Deere Ave., Irvine, CA 92714.

Appendix A:
Waveguide Coupling and Near Field Measurement

Except for prism coupling to measure the effective refractive index, power was typically launched into the chip using end fire coupling from a semiconductor laser diode. For experiments carried out with only one polarization, such as propagation loss measurements, excitation of either the channel quasi-TE or quasi-TM mode was accomplished by adjusting the polarization of the light incident on the cleaved chip end face by rotation of the polarized laser source. The setup for coupling into the waveguide, and a beam trace through the lenses will be given. Measurements of the field profile were also carried out by imaging a cleaved end face of the excited waveguide onto a camera using a microscope objective (the near field technique). Because the objective broadens the image in object plane units to approximately the wavelength of light, and because channel dimensions are of this order, diffraction in the objective can distort the image. In this Appendix, the effect of diffraction on the image profile is studied. The conclusion is reached that for channels fabricated for this work, the field profile in the direction parallel to the wafer surface is not distorted, while the narrower one perpendicular to the wafer surface is only slightly broadened when imaging with a 100× $NA = 0.9$ objective.

1 INCOUPLING AND MEASUREMENT SETUP

Much of the waveguide work was performed in the near infrared at 830 nm using laser diode, SLD, or LED sources. It is more convenient to launch light into planar, and especially channel waveguides using end fire coupling than with prism coupling. A photo of the setup for launching and observing the outscattered radiation is shown in Fig. A-1. The fixture for holding the laser diode, shown in the photo, consists of a mounting plate/heat sink which is electrically insulated from the optical table. The laser

diode mounting plate, secured to an upright bracket with nylon screws, and insulated with a piece of printed circuit board, can be removed, rotated, and reattached at 90° increments, permitting the polarized laser light to be oriented parallel (TE) or perpendicular (TM) to the optical table. An 8 mm focal length collimating lens was mounted on a 5-axis positioner, which allowed adjustment of its distance from the laser diode, as well as permitting the beam to be steered so that it could be centered on the focusing lens. When measuring planar guide loss by the out scattering method, it is desirable to minimize spreading of the streak by keeping the beam waist large after the focusing lens, consistent with maintaining reasonable coupling efficiency. Here a 3× ($NA = 0.1$) objective was employed as the focusing lens. For channel coupling, either a high NA objective or another lens identical to the collimating lens was employed. The collimating and focusing lenses were set at approximately the same height above the optical table.

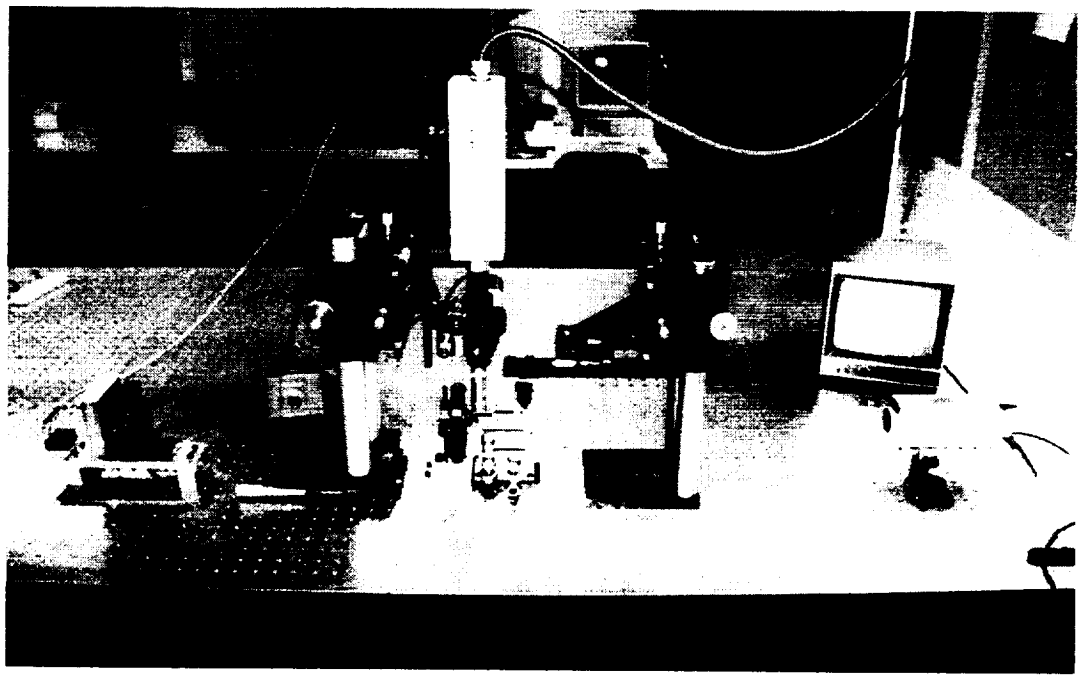


Fig. A-1 Photo of end fire coupling setup. The camera used to near field imaging is also shown. For comparison, the optical table has $\frac{1}{4}$ - 20 holes on 1" centers.

The wafer rested without vacuum hold down on a three-axis crossed roller stage equipped with fine adjustment micrometers. A 25× (NA = 0.31) long working distance objective was used to image outscattered light on to a vidicon monochrome camera, and the image was displayed on a 9 in black and white monitor. The objective and camera could be moved parallel to the incoming beam axis by a 2" travel crossed roller stage, over lesser distances normal to the substrate (for focusing) and normal to the streak. To measure the near-field intensity or output signal from the chip, another objective could be positioned at the output edge with a three axis stage.

To incouple light into a sample, the collimating lens was first adjusted using a phosphor card so that the beam entering the focusing lens was slightly smaller than its clear aperture. Then the collimating lens was adjusted in both directions parallel to the laser diode mounting plate so that the beam leaving the focusing lens expanded uniformly about the focusing lens axis. Typically the tilt in each direction for both lenses was set so that their axis coincided with the beam line. A waveguide chip was then placed on the sample stage waveguide face up and rotated so that its cleaved edge was perpendicular to the beam line. The waveguide face was maintained parallel to the optical table by the stage.

The chip position for optimal coupling was next found. First the chip was lowered until the beam missed its edge and passed over its surface. Then, the chip's height was increased until it occluded the beam. These two steps--lowering and raising the chip into the beam path--were repeated while the distance between the chip and incoupling lens was adjusted. As the chip edge horizontally approached the beam waist and optimum coupling, several effects could be seen on a phosphor card placed opposite the incoupling side. Diffraction lines, especially noticeable when the beam was just being cut, tended to spread out more. The amount of vertical stage movement necessary to take the beam from just being intercepted to completely occluded decreased. The power coupled into the waveguide when it was aligned with the beam, as revealed by

outscattered intensity, increased. After the chip was coarsely placed to intensify these effects, all three axes were iteratively adjusted to maximize coupled power, as evidenced by streak brightness.

2 GAUSSIAN BEAM PROPAGATION

One matter not dealt with in the previous discussion is that the waist size on the chip side of the focusing lens is affected by the diameter of the beam entering the focusing lens, which is controlled by the separation between the laser diode and collimating lens. Gaussian beam optics will now be used to quantify this interaction. This theory assumes thin lenses and a paraxial beam, and does not take into account lens aberration or diffraction effects. A good tutorial on Gaussian beam optics can be found in several optics catalogues [1, 2] and the full derivation is available in [3].

A Gaussian beam can be characterized by only two quantities: the waist position and diameter, where the waist is the place where the beam comes to a focus. Once these are known, the beam diameter and phase front curvature can easily be determined everywhere in the optical system (Fig. A-2). A Gaussian beam's field amplitude has the form

$$E(r,z) = E_0(z)\,\exp\left(-\frac{r^2}{\omega^2(z)}\right), \qquad \text{(A.2.1)}$$

where r is radial distance from the beam center, z is distance parallel to the direction of propagation, E_0 is the field amplitude at the center of the beam, and the state of polarization has been omitted from the field description. For elliptical beams, the field profile parallel to each axis of the ellipse is assumed be described by different Gaussians. The intensity is given by

$$I(r,z) = I_0(z)\,\exp\left(-\frac{2r^2}{\omega^2(z)}\right). \qquad \text{(A.2.2)}$$

The beam radius ω at any distance z from the waist ω_0 can be determined from

$$\omega^2(z) = \omega_0^2\left[1 + \left(\frac{\lambda z}{\pi\omega_0^2}\right)^2\right] , \tag{A.2.3}$$

while the radius R of curvature of the phase fronts is given by

$$R(z) = z\left[1 + \left(\frac{\pi\omega_0^2}{\lambda z}\right)^2\right] . \tag{A.2.4}$$

Gaussian beam optics use the thin lens approximation, where the lens' only attribute is its focal length. Here the lens is assumed to transform the beam's phase front curvature R while leaving its size ω unchanged. To find the new phase front curvature, the complex quantity

$$\frac{1}{q_{in}(z)} = \frac{1}{R(z)} - \frac{j\lambda}{\pi\omega^2(z)} \tag{A.2.5}$$

is evaluated. After passing from one side of a lens to the other (and moving through zero distance because of the thin lens assumption) or propagating over some distance, q can be recalculated using

$$q_{out} = \frac{q_{in}A + B}{q_{in}C + D} , \tag{A.2.6}$$

where, for a lens having focal length f_L,

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{-1}{f_L} & 1 \end{bmatrix} \tag{A.2.7}$$

and for propagation distance d,

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} . \tag{A.2.8}$$

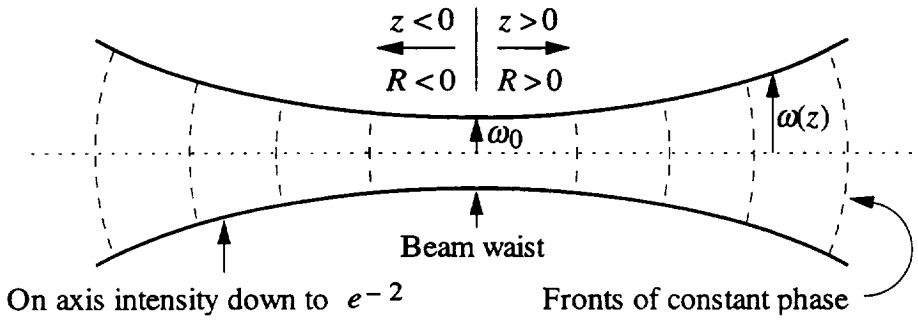


Fig. A-2 Sign convention for Gaussian beam analysis.

Matrices (A.2.7) and (A.2.8) can be used to find the beam (ω and R) leaving the focusing lens of the setup in Fig. A-1. To do this the matrices are multiplied together in reverse order. For example, suppose an incoming beam, described by q_{in}, propagates over distance d_1, through a lens having focal length f_L, and then over distance d_2. Its description q_{out} after this sequence is given by (A.2.6), where

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & d_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \dfrac{-1}{f_L} & 1 \end{bmatrix} \begin{bmatrix} 1 & d_1 \\ 0 & 1 \end{bmatrix} \tag{A.2.9}$$

provides the unknowns in (A.2.6). Subsequently, $1/q_{out}$ is separated into real and imaginary parts as in (A.2.5) and R and ω found.

Beam trace for incoupling setup

To determine the effect spot size at the entrance to the focusing lens has on the beam waist radius at the waveguide edge, Gaussian beam theory was applied to the setup of Fig. A-1. As (A.2.3) shows, beams having small waists rapidly approach a linear divergence asymptote. Our source is a Sharp LTO15MDO 830 nm laser diode, which has a far-field full angle to half intensity of 10° and 30° parallel and perpendicular to the laser junction plane. Using (A.2.2) and (A.2.3), the far-field full (radian) angle to half maximum is found to be $\theta = \dfrac{\lambda_0 \sqrt{-2\ln(0.5)}}{\pi \omega_0}$, so the waists ω_0 at the laser diode edge parallel and perpendicular to the junction are 1.8 μm and 0.6 μm. Both waists of the

elliptical laser diode beam are assumed to coincide with the laser diode edge, and therefore the astigmatic distance, defined as the difference in their position, is assumed zero. Since the laser diode has its electric field polarized parallel to the junction plane, the long axis of its far-field beam ellipse is parallel with the electric field polarization vector.

Two beam profiles were modeled: 1) when the laser-to-collimating lens separation is adjusted to bring the beam to a focus at the entrance of the focusing lens (Fig. A-3) and 2) when laser is placed at the collimating lens focal point (Fig. A-4). The separation between the two lenses, which have focal lengths of 8 mm, was set to 0.5 m, their approximate separation in the setup. The waist of both axis in Fig. A-3 is substantially larger than in Fig. A-4, where the beam returns to the size it had at the laser diode emission edge. Experimentally it was noted that maximum waveguide coupling efficiency and sensitivity to positioning occurred when the beam almost filled the focusing lens and when the TE mode was excited. The reason for the latter phenomenon is apparent from Fig. A-4. When the laser was mounted to excite quasi-TE in the channel, the intensity ellipse matched that guided by the waveguide, while for TM, the laser was mounted so that the long axis of its spot was parallel with the shortest dimension of the channel waveguide field profile.

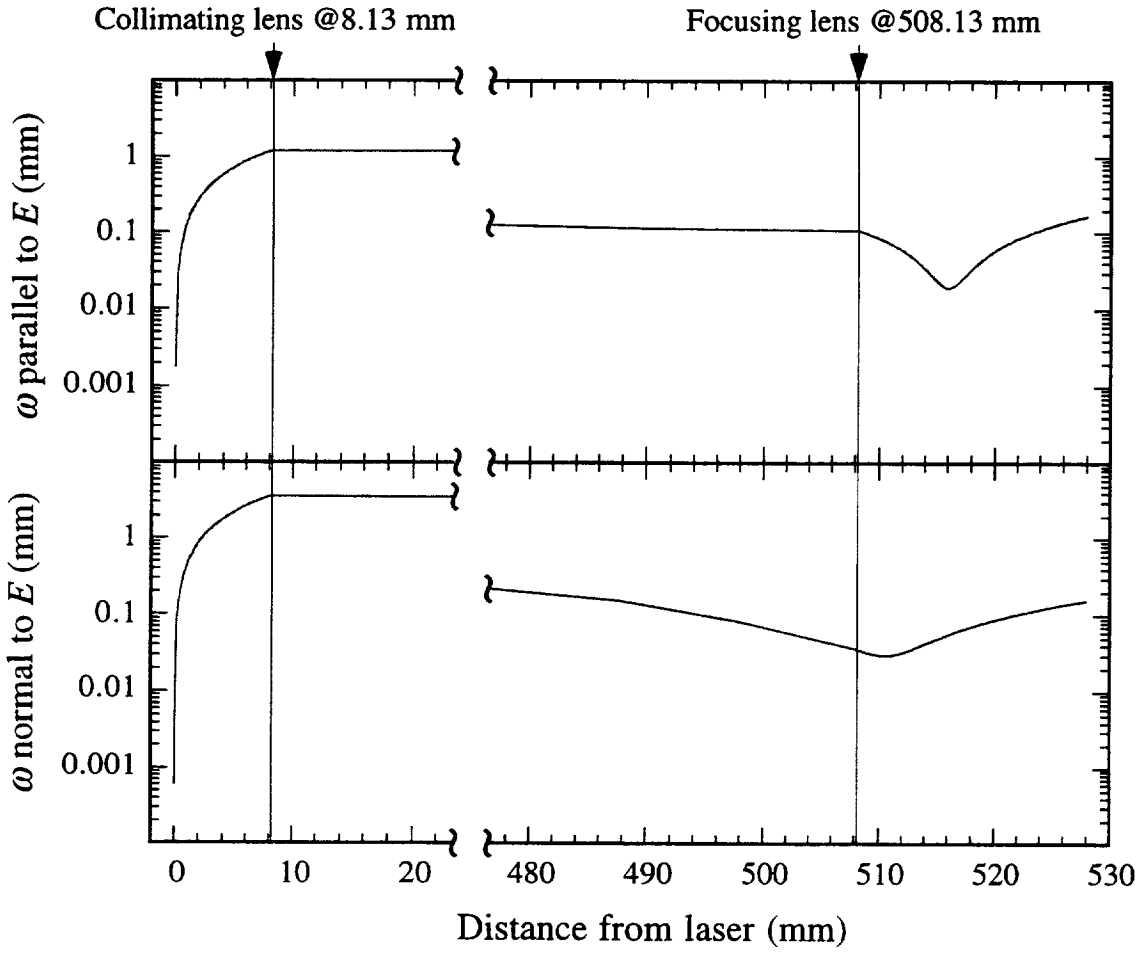


Fig. A-3 Calculated beam radius throughout system of Fig. A-1 with the laser-to-collimating lens separation adjusted to focus the beam (produce a waist) on the entrance of the focusing lens. Between the collimating lens and laser, the beam's increase in size is almost linear--the apparent curvature is the result of the log plot.

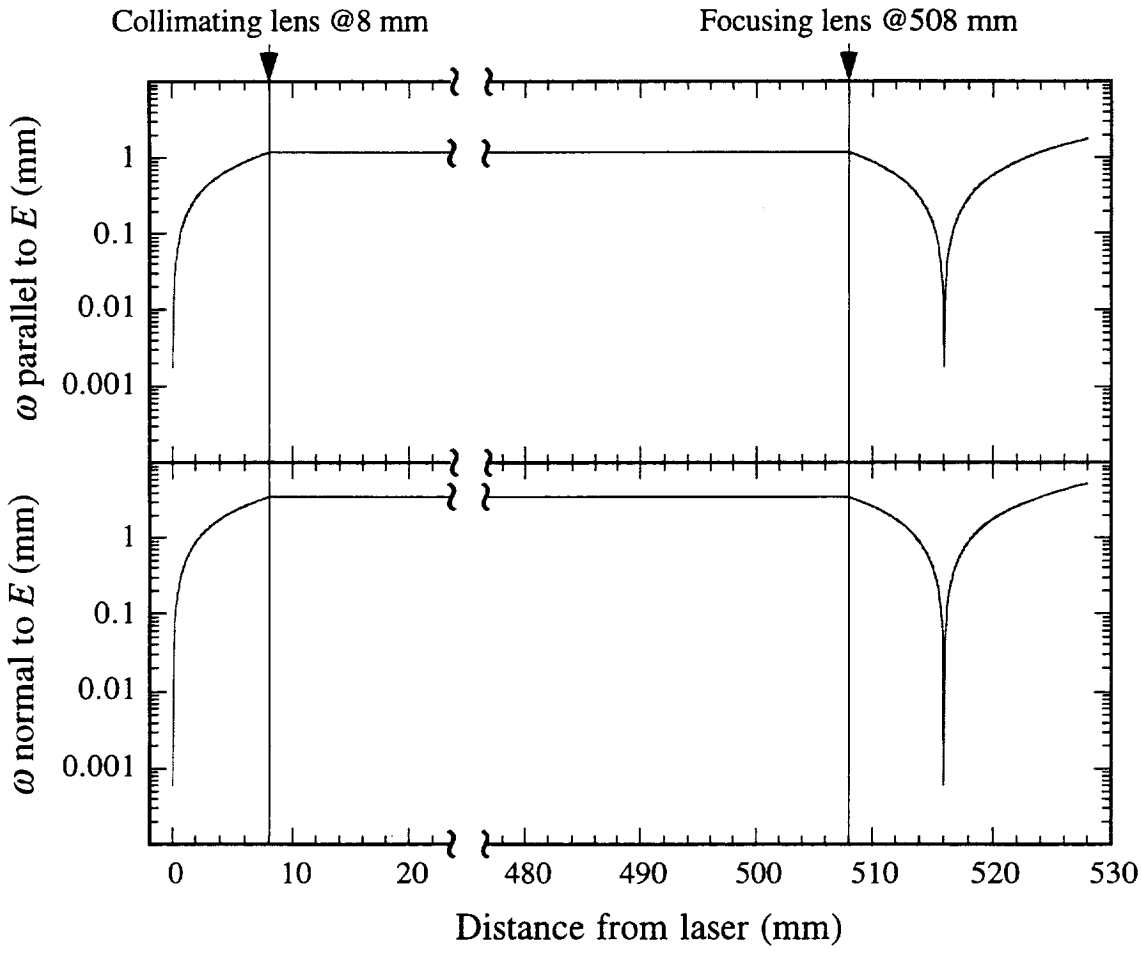


Fig. A-4 Same as Fig. A-3 but with the laser placed at the focal point of the collimating lens.

3 CHANNEL FIELD PROFILE

Until now, the field profile has been analytically investigated using the effective index method. For the rib waveguide, and using the channel coordinate system depicted in Fig. 3-3, the field profile has been approximated by the separable function $\phi = X(x)\ Y(y)$. $X(x)$ is identical to that of the three layer planar guide which would be formed if the channel width went to infinity, while $Y(y)$ is calculated for the waveguide shown in the lower part of Fig. 3-3. Marcuse [4] asserts that in the channel, the field spreads out more in the x direction than in the slab, and that the field in the y direction calculated by the effective index method is more spread out than in the actual channel.

Thus the field furnished by the effective index method is expected to be too confined in x and too spread out in y. Instead of evaluating the method's accuracy by comparing its field results with those from a more computationally intense (and presumable more accurate) technique, the effective index field was compared with the measured near-field intensity profile of a channel waveguide. Although the actual field is the "exact" answer, diffraction in the optics and differences between the estimated and actual channel dimensions will introduce uncertainty into the comparison.

The near-field method consists of measuring the intensity distribution in the image of a cleaved output edge of the wafer where the channel terminates. As will be seen, resolution can be maximized by selecting a high numerical aperture optic. Such a lens should be corrected for aberrations and easy to mount; a microscope objective provides this in a convenient package. The near-field method has the advantage of providing a direct although distorted view of the field. Other methods [5] which have been used for determining the spot size of optical fibers include the far-field method (no lens), measuring light coupled between two identical closely-spaced cleaved fibers as a function of offset, and measuring either the near or far-field intensity transmitted through a mask. The far-field method requires much computation to recover the field profile or an assumption of a Gaussian field profile. To apply the offset method, the channel would have to be positioned very close (<1 μm) to another channel and translated in submicron increments.

Since the field profile is of the order of the wavelength of light, its image will be smoothed and enlarged by diffraction effects in the objective lens. To account for these, this distortion will be quantified by comparing a known field profile from a single mode planar Si_3N_4 waveguide with its magnified image. Then the distortion will be mathematically applied to the calculated channel profile, and this compared with the near-field image.

Diffraction

Diffraction, which occurs even in a perfect lens due to its finite size, possesses the same functional form as that from an aperture having the same diameter as the lens [6]. The Fraunhofer diffraction pattern is typically assumed. The diffraction integral, which gives the field profile after the aperture, may be written as the inverse Fourier transform of the aperture function. To see this, the Fraunhofer diffraction pattern is found by summing the wavelets, which as Huygens principle states, are continuously produced at each point on the phase front of any wave [7], and in particular, on the one at the aperture plane (Fig. A-5):

$$E(X,Y) = \int\int_{aperture} \frac{E_a}{r} \exp[j(\omega t - k_0 r)]\, dx\ dy \quad . \tag{A.3.1}$$

The distance from the point in the aperture to a point in the image plane is

$$r = \sqrt{(x-X)^2 + (y-Y)^2 + s'^2} \quad . \tag{A.3.2}$$

To simplify the integral, expand (A.3.2), assign $R^2 = X^2 + Y^2 + s'^2$, assume the aperture is small ($x^2 \approx 0$, $y^2 \approx 0$), and approximate the square root function as the first two terms of its series expansion to get

$$r \approx R\left(1 - \frac{xX + y\,Y}{R^2}\right) \quad . \tag{A.3.3}$$

Inserting this into (A.3.1), using $k_0 = 2\pi/\lambda_0$, assuming that $R \approx s'$, and making the approximation $\frac{E_a}{r} \approx \frac{E_a}{s'}$ gives

$$E(X,Y) = \frac{E_a}{s'} \exp[j(\omega t - k_0 s')] \int\int_{aperture} \exp\left[2\pi j\left(X\frac{x}{\lambda_0 s'} + Y\frac{y}{\lambda_0 s'}\right)\right] dx\, dy \tag{A.3.4}$$

The fractions $\frac{x}{\lambda_0 s'}$ and $\frac{y}{\lambda_0 s'}$ have the units of spatial frequency (m^{-1}) so the quantities f_x and f_y are assigned to them. By defining a pupil function $P(x,y)$ which is 1 inside the aperture and 0 elsewhere,

$$P(x,y) = \begin{cases} 1 \text{ for } x^2 + y^2 \le \rho_m^2 \\ 0 \text{ for } x^2 + y^2 > \rho_m^2 \end{cases} , \tag{A.3.5}$$

(A.3.4) may be written

$$E(X,Y) = A \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(x,y) \exp[2\pi j(Xf_x + Yf_y)] \, dx \, dy , \tag{A.3.6}$$

where A is the multiplying constant apparent in (A.3.4). To put (A.3.6) in inverse Fourier transform form, the pupil function can be re-defined in the frequency domain

$$\widetilde{P}(f_x,f_y) = \begin{cases} 1 \text{ for } f_x^2 + f_y^2 \le \left(\dfrac{\rho_m}{\lambda_0 s'}\right)^2 \\ 0 \text{ for } f_x^2 + f_y^2 > \left(\dfrac{\rho_m}{\lambda_0 s'}\right)^2 \end{cases} , \tag{A.3.7}$$

and (A.3.6) becomes

$$E(X,Y) = A\,(\lambda_0\ s')^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \widetilde{P}(f_x,f_y) \exp[2\pi j(Xf_x + Yf_y)] \, df_x \, df_y . \tag{A.3.8}$$

With these manipulations (A.3.8) is evidently the 2-D inverse Fourier transform of the pupil function. Since the image plane field profile is of interest, while the absolute brightness is not, the constant $A\,(\lambda_0 s')^2$ is unimportant here.

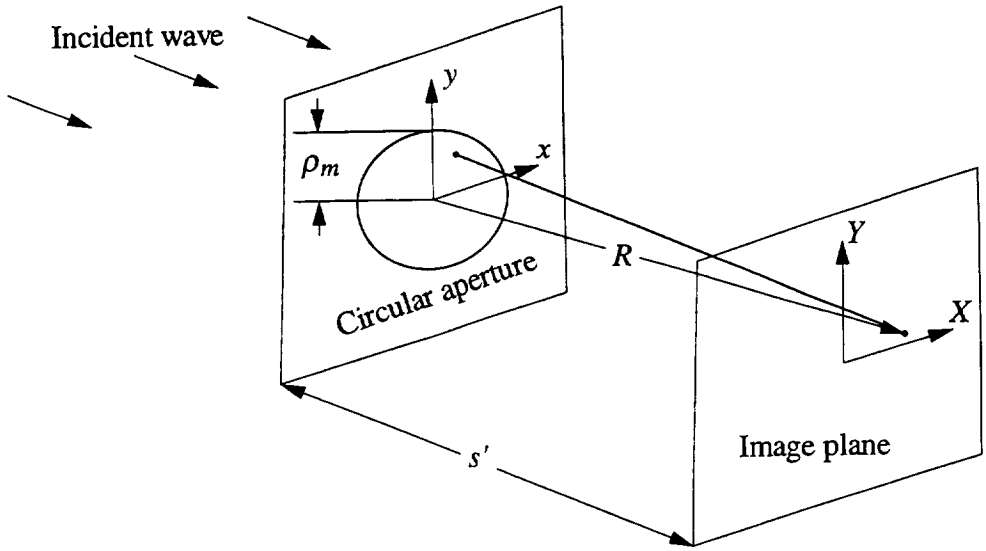


Fig. A-5 Arrangement for Fraunhofer diffraction [6].

Application to lenses

For a circular aperture, the integral (A.3.8), when performed in cylindrical coordinates with a circular pupil function, evaluates to [8]

$$E(r) = E_0 \frac{2 J_1\left(\frac{2\pi r \rho_m}{\lambda_0 s'}\right)}{\frac{2\pi r \rho_m}{\lambda_0 s'}}, \tag{A.3.9}$$

where E_0 is the maximum electric field amplitude (at the center of the distribution), r is radial distance from the distribution center, J_1 is the Bessel function of the first kind of order one, and ρ_m and s' are defined in Fig. A-5. An aberration-free spherical lens will act like an aperture and image a collimated beam not to a point of light at its focus but to the diffraction pattern given in (A.3.9). Here ρ_m and s' correspond to the lens radius and focal length. The intensity distribution

$$I(r) = 4 I_0 \left[\frac{J_1\left(\frac{2\pi r \rho_m}{\lambda_0 s'}\right)}{\frac{2\pi r \rho_m}{\lambda_0 s'}} \right]^2 \tag{A.3.10}$$

is the square of the field amplitude. Equations (A.3.9) and (A.3.10), which have been plotted in Fig. A-6, are called point spread functions since they represent the electric field amplitude and intensity distribution imaged from a point source.

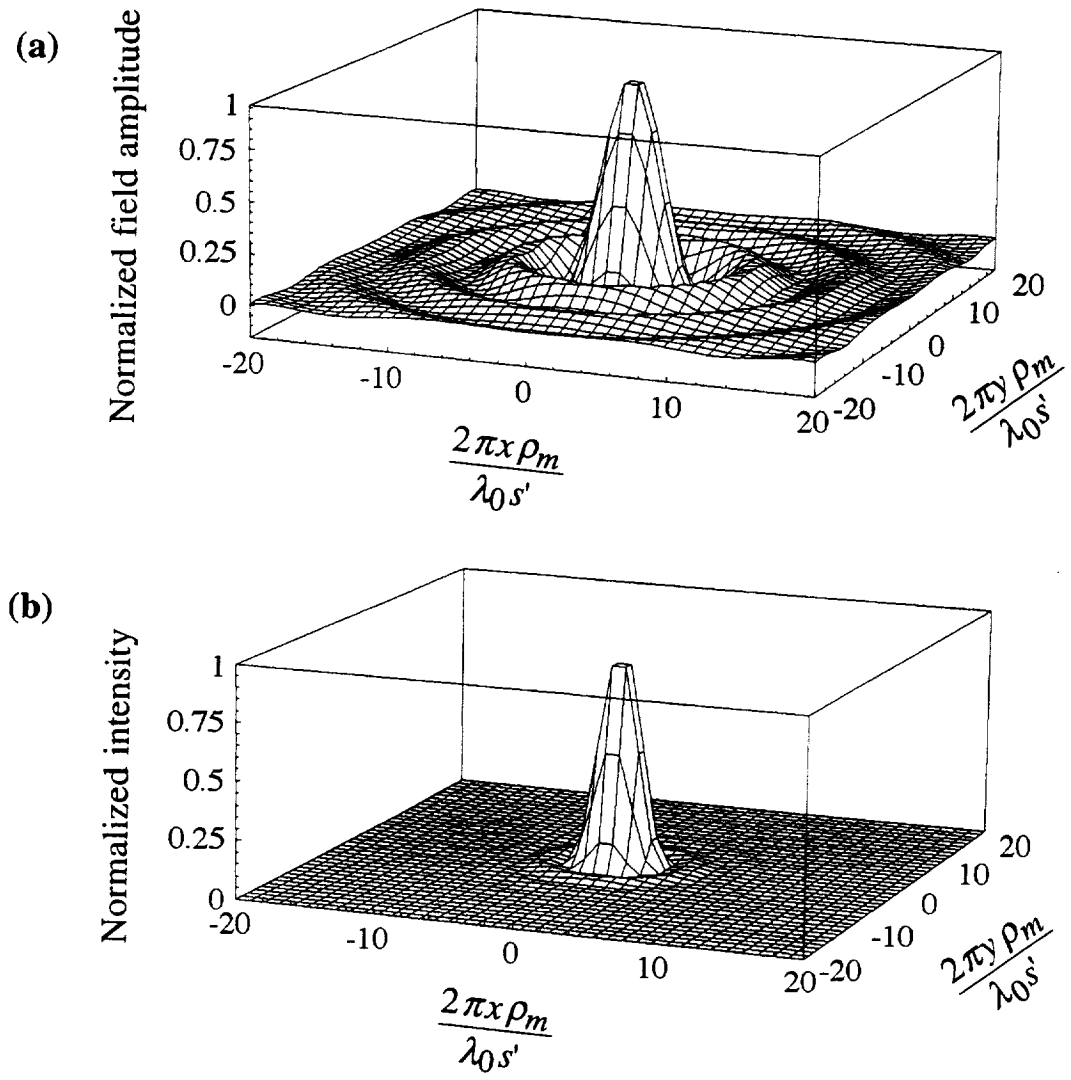


Fig. A-6 Plot of (A.3.9) and (A.3.10). They represent the field amplitude and intensity observed in the image plane from a point source.

A point object placed many times the focal length from a lens produces waves which are near-plane in the vicinity of the lens and approximates the situation just examined. The formulas (A.3.4)-(A.3.8) continue to apply as the point-object approaches the lens and even when it is placed near its focal point, a situation illustrated in Fig. A-7.

As the object nears the lens, the distance s' increase along with magnification $M = s'/s$. A measure of the diffracted spot size is the radial distance r to the first intensity minimum (Fig. A-6b) which occurs for r satisfying $\frac{2\pi r\rho_m}{\lambda_0 s'} = \frac{2\pi r\rho_m}{\lambda_0 Ms} = 3.83$ or $\frac{3.83\lambda_0 Ms}{2\pi\rho_m} = r$. As magnification increases, the diffraction pattern also grows. However the image size also does, and for convenience the image is usually scaled by the inverse of the magnification factor so one can work in object plane units. For large magnifications, $s \approx f_L$, where f_L is the focal length, while the objective's numerical aperture is given by $NA \equiv \sin(\theta_m) = \frac{\rho_m}{f_L}$. The field amplitude diffraction pattern in object plane units is then given by

$$E(r) = E_0 \frac{2J_1\left(\frac{2\pi rNA}{\lambda_0}\right)}{\frac{2\pi rNA}{\lambda_0}}, \tag{A.3.11}$$

and the intensity by

$$I(r) = 4I_0 \left[\frac{J_1\left(\frac{2\pi rNA}{\lambda_0}\right)}{\frac{2\pi rNA}{\lambda_0}}\right]^2. \tag{A.3.12}$$

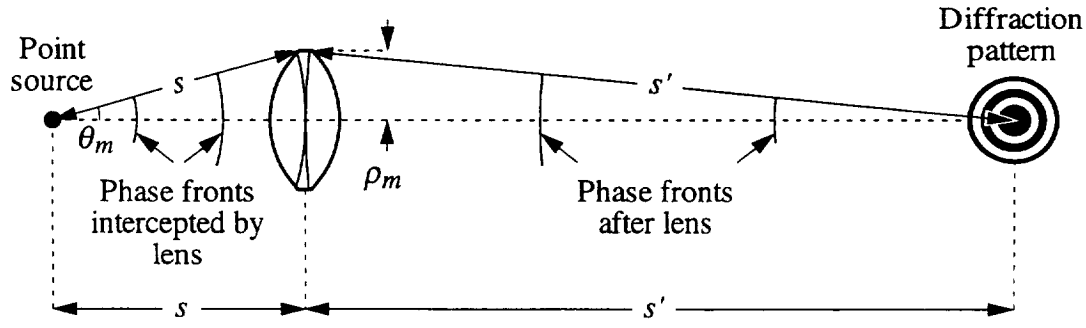


Fig. A-7 Diffraction by a lens.

Finding the affect of diffraction on the image

The waveguide field in the image plane, distorted by diffraction, is found by summing the overlapping diffraction patterns imaged from each point on the channel end face. Fig. A-8 shows two methods of doing this calculation, one for coherent light and the other for incoherent light. For either case or method, the computations are performed numerically. They are treated differently because with incoherent light, intensities add, while for coherent light, the complex field amplitudes must be summed, with proper account taken of phase [9]. Since our waveguide is excited with a diode laser, the coherent treatment is appropriate and was performed.

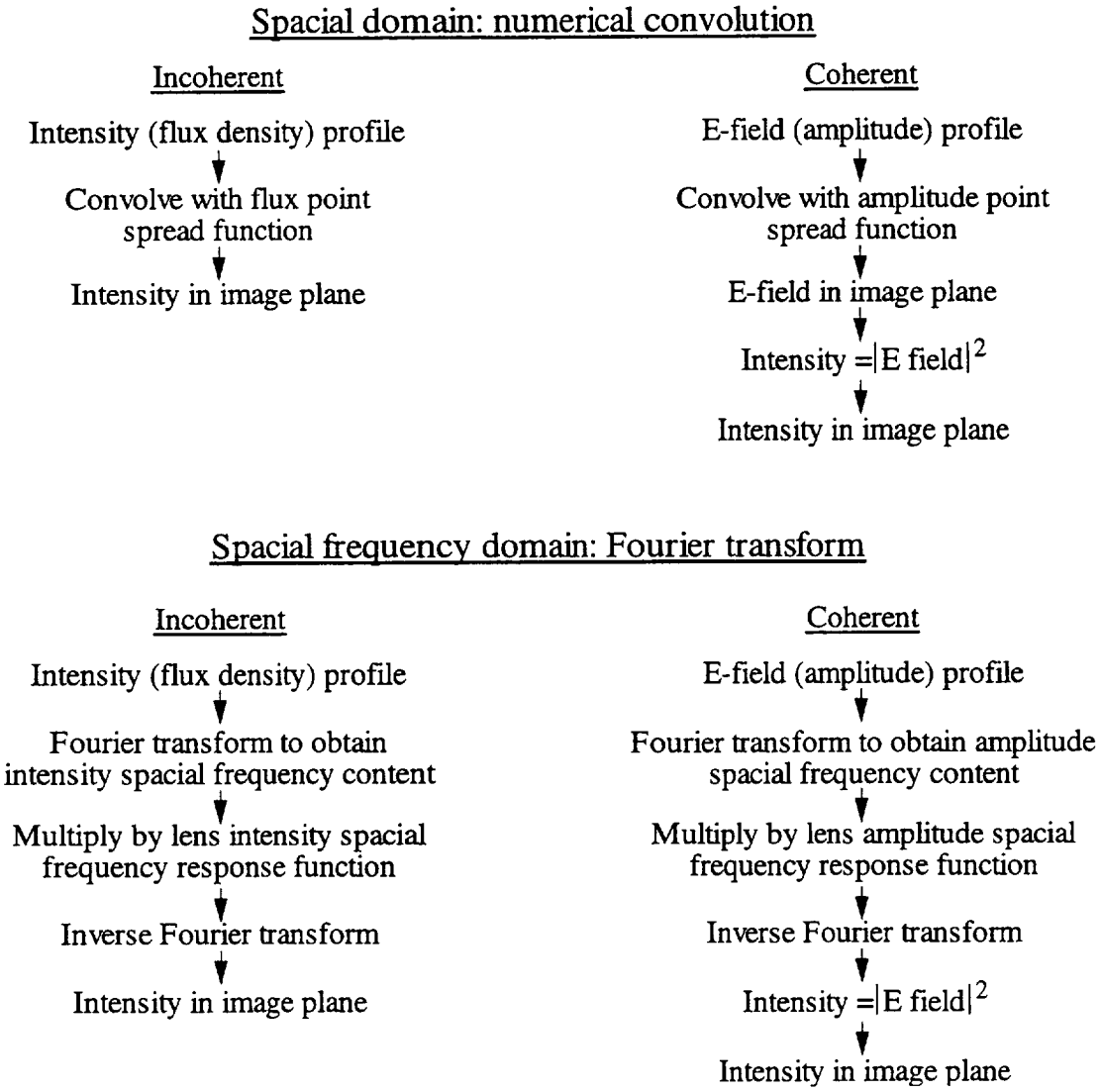


Fig. A-8 Sequence of calculations to include diffraction in calculating the field profile in the image plane from that at the exit face of the waveguide.

Table A-1 illustrates, using numbers which have no particular significance, a 2D numerical convolution $C = A \star B$. The center position in each field is bordered by a heavy line box. It can be seen that shifting either A or B fields with respect to the center box will also shift the result. For the lens calculation, the actual waveguide field would be in A, the point spread function (A.3.11) in B and the resultant blurred field in C. All should use the same grid spacing.

A		*B*			*C*			
			0.7			2.1	2.8	
3	4	0.9	1.2	0.8	2.7	3.6+3.6	4.8+2.4	3.2
			0.3	0.6		0.9	1.2+1.8	2.4

Table A-1 Example numerical convolution.

An alternate to numerical convolution, suggested in Fig. A-8, is to calculate the Fourier transform of the point spread function and multiply it by the Fourier transform of the field profile. The inverse Fourier transform of the product yields the diffraction-distorted image (in object units). The lens frequency response functions, found by taking the Fourier transform of the point spread functions, have been plotted in Fig. A-9. They can be analytically evaluated for both the coherent (A.3.11) and incoherent (A.3.12) point spread functions, obviating the need to find them numerically [9, 10]. For the coherent case, it takes the shape of a rotated step function, while the function $\frac{2}{\pi}\cos^{-1}\left(\frac{f\lambda_0 s'}{2\rho_m}\right)-\frac{f\lambda_0 s'}{\pi\rho_m}\left[1-\left(\frac{f\lambda_0 s'}{2\rho_m}\right)^2\right]^{1/2}$, where f represents spatial frequency, is rotated to generate the incoherent case plot. The cutoff frequencies for the coherent and incoherent cases are $\frac{\rho_m}{\lambda_0 s'}$ and $\frac{2\rho_m}{\lambda_0 s'}$.

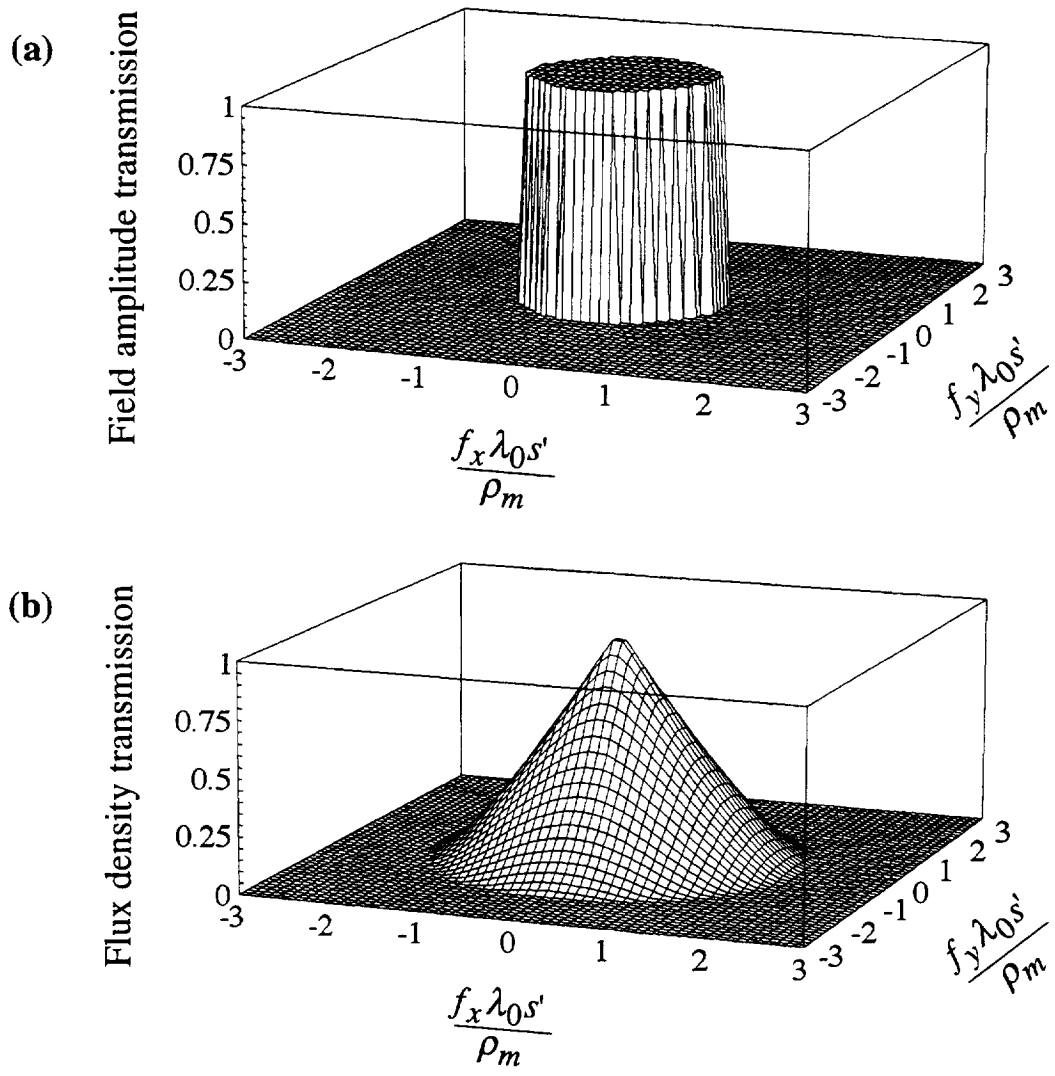


Fig. A-9 The Fourier transforms of the point spread functions in Fig. A-6. Figure (a) is a cylinder, while (b) is almost a cone. The functions represent transmission of either amplitude or flux density (intensity) as a function of spatial frequency. Spatial frequency is usually expressed in cycles/mm.

Defocusing and spherical aberration

There are seven aberrations in spherical lenses [11]: spherical aberration, coma, astigmatism, field curvature, distortion, longitudinal chromatic aberration, and lateral chromatic aberration. For on axis monochromatic illumination, only spherical aberration is significant. In addition, focus error can be treated similarly to an aberration. Spherical aberration was noted in the objective used to image the waveguide. In addition, since

defocusing can be used to partially offset spherical aberration, it too must be included in the analysis.

An aberration can be interrupted as a departure of the wavefront from its proper spherical shape. In Fig. A-10 movement of the object a distance ζ from its position of focus in a direction away from the lens produces a wavefront having too large a radius of curvature. From Fig. A-10 the relationship

$$\left(\sqrt{r_1^2-\rho^2}+\zeta\right)^2+\rho^2=(r_2-W_{defocus})^2 \tag{A.3.13}$$

can be determined, where $W_{defocus}$ is the wavefront error. Substituting $\zeta=r_2-r_1$ into (A.3.13) and simplifying results in

$$2\zeta\sqrt{r_1^2-\rho^2}=2\zeta r_1-2W_{defocus}r_1+W_{defocus}^2-2\zeta W_{defocus} \quad . \tag{A.3.14}$$

For a typical imaging situation, the phase front error $W_{defocus}$ is the order of a few wavelengths of light, defocusing (in the object plane) will be no more than a few micrometers, while r_1 and ρ will be on the order of hundreds of micrometers. To simplify (A.3.14), the small terms $W_{defocus}^2$ and $2\zeta W_{defocus}$ will be ignored and the approximation $\sqrt{1-\frac{\rho^2}{r_1^2}}\approx 1-\frac{\rho^2}{2r_1^2}$ used, producing

$$W_{defocus}=\frac{\zeta\rho^2}{2r_1^2} \quad . \tag{A.3.15}$$

Using $\phi=2\pi W_{defocus}/\lambda_0$ and $NA=\frac{\rho_m}{f}\approx\frac{\rho_m}{r_1}$, (A.3.15) can be written

$$\phi_{defocus}=\frac{\pi\zeta\rho^2 NA^2}{\lambda_0\rho_m^2} \quad . \tag{A.3.16}$$

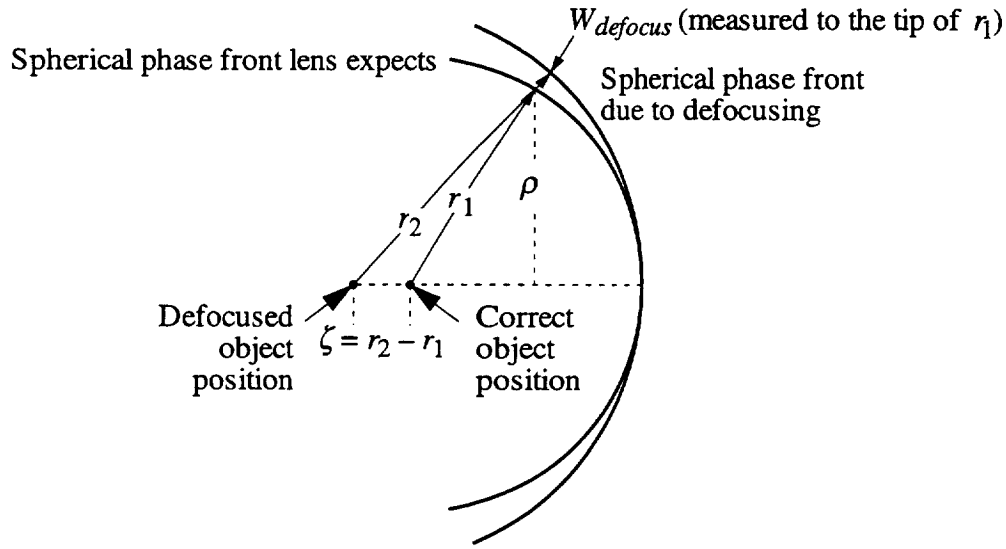


Fig. A-10 Phase front error from defocusing in the object plane.

Besides defocusing, spherical aberration can also be present. Its aberration function is

$$W_{spherical} = C_4\rho^4 + C_6\rho^6 + C_8\rho^8 + \ldots \quad , \tag{A.3.17}$$

but in this study, only the first term was assumed to be significant. The total phase error is the sum of these

$$\phi(\rho) = \frac{\pi\zeta\rho^2 NA^2}{\lambda_0\rho_m^2} + 2\pi C\frac{\rho^4}{\rho_m^4} \quad , \tag{A.3.18}$$

and the spherical aberration error term has been rewritten so that the constant C gives the number of waves of error at the lens margin. It can be seen from (A.3.18) that selecting a nonzero defocus ζ can reduce the average phase front error ϕ over the lens aperture. However this will only occur to the degree that a scaled ρ^2 term can offset a ρ^4 term over $0 \le \rho \le \rho_m$. Using this suggestion, (A.3.18) can be written

$$\phi(\rho) = \frac{\pi(\zeta_0 + \Delta\zeta)\rho^2 NA^2}{\lambda_0\rho_m^2} + 2\pi C\frac{\rho^4}{\rho_m^4} \quad , \tag{A.3.19}$$

where ζ_0 is the movement away from focus to maximally cancel the spherical aberration and $\Delta\zeta$ is the offset from this.

Transformation of wavefront error from ray angle to spatial frequency

To compute the image plane waveguide intensity profile from its calculated intensity profile (the object), the Fourier transform method was applied for coherent light to take into account diffraction, defocusing and spherical aberration in the objective (see the lower right part of Fig. A-8). Here the Fourier transform of the point spread function was multiplied by the Fourier transform of the object (the calculated field profile). Taking the inverse Fourier transform of the product gives the image field amplitude. The magnitude squared of the result gives the intensity, which could then be compared with the observed near field image. Since the magnitude of spherical aberration in the objective is unknown, it will be determined by comparing the calculated and observed intensity profiles in a planar waveguide having a thin, high refractive index core layer. The field profile from such a guide will contain spatial frequencies which extend well beyond those passed by the objective, and therefore serves as a convenient test of its frequency response (resolving ability).

To find the Fourier transform of the point spread function, including aberrations, note that the point spread function is simply the inverse Fourier transform of the pupil function $P(x,y)$. Therefore the Fourier transform of the point spread function recovers the pupil function. For a defocused lens with spherical aberration, the pupil function in cylindrical coordinates becomes

$$P(\rho) = \begin{cases} \phi(\rho) \text{ for } \rho \le \rho_m \\ 0 \text{ for } \rho > \rho_m \end{cases}, \qquad \text{(A.3.20)}$$

where $\phi(\rho)$ is given in (A.3.19). To put this function in the frequency domain, recall that the spatial frequencies in object plane units are given as a function of ρ by $f = \frac{\rho}{\lambda_0 s'} = \frac{\rho}{\lambda_0 M s}$ [see (A.3.4)] or by $\frac{\rho}{\lambda_0 s}$ in object plane units. This makes sense, since for magnification >1, the object becomes enlarged and spatial frequencies in image plane

units are scaled downward. Using $NA \approx \frac{\rho_m}{s}$, $f = \frac{\rho}{\rho_m}\frac{NA}{\lambda_0}$ or $\rho = \frac{f\lambda_0\rho_m}{NA}$ in object plane units. Substituting for ρ in (A.3.19) gives

$$\phi(f) = \pi(\zeta_0 + \Delta\zeta)f^2\lambda_0 + 2\pi C\left(\frac{f\lambda_0}{NA}\right)^4 . \qquad \text{(A.3.21)}$$

This equation provides the phase front error, due to spherical aberration and defocusing, for an image having spatial frequency f.

Calculation of near field image

The inverse Fourier transform of the product of the Fourier transforms of the point spread function and object field amplitude gives the field amplitude in the image plane. Since, as shown before, the pupil function is the Fourier transform of the point spread function, one can save a calculation step and incorporate aberrations by going directly to the pupil function (A.3.21), which is already written in the spatial frequency domain.

The Fast Fourier Transform algorithm was used to determine the Fourier Transform of the field amplitude [12]. When a 1-D FFT is taken, the amplitudes at spatial frequencies positioned as shown in Table A-2 are returned. The negative frequency components, shown in the latter half of the table, may be interpreted as belonging to one of the complex exponentials required to form, together with the positive frequency exponential, an arbitrary sinusoid:

$$\begin{aligned} A_m \cos(2\pi f_m x + \theta_m) = {} & \frac{A_m}{2}\exp(j\,\theta_m)\exp[j\,2\pi f_m x] \\ & + \frac{A_m}{2}\exp(-j\,\theta_m)\exp[j\,2\pi(-f_m)x] \end{aligned} \qquad \text{(A.3.22)}$$

If the sinusoid had real amplitude A_m, the positive and negative frequencies would be complex conjugates.

FFT term #	1	2	3	. . .	$\frac{N}{2}$	$\frac{N}{2}+1$	$\frac{N}{2}+2$	. . .	N-1	N
spatial frequency	0	$\frac{1}{N\Delta x}$	$\frac{2}{N\Delta x}$	. . .	$\frac{N-1}{N\Delta x}$	$\frac{1}{\Delta x}$ & $\frac{-1}{\Delta x}$	$\frac{-(N-1)}{N\Delta x}$	. . .	$\frac{-2}{N\Delta x}$	$\frac{-1}{N\Delta x}$

Table A-2 Correspondence between FFT terms and their spatial frequency for N even.

A 2-D FFT of an image may be taken by replacing the columns by their 1-D FFT and then taking the 1-D FFT of all of the rows, or the rows may be first operated on and then the columns [13]. The top row and left column contain the first FFT terms. For a planar guide having no horizontal field variation, an FFT first of the columns would yield a matrix having identical columns. Each row would contain the same value at all positions and its FFT would return only one nonzero term: the first DC term. For planar waveguides, then, it is apparently sufficient to multiply a 1-D FFT of the field profile normal to the wafer surface by the pupil function in one dimension to determine the near field image (which will also vary only in one dimension).

In (A.3.5) the spatial frequency response is independent of angle about the optic axis and depends only on ρ (note cylindrical symmetry in Fig. A-9). In a similar manner, phase front error depends only on $f = \sqrt{f_x^2 + f_y^2}$. Using this fact and Table A-2, the spatial frequencies would end up as shown in Fig. A-11. Here f is the spatial frequency increment given by $\frac{1}{N\Delta x}$,where N is the number of sampled intensities (256) and Δx is the separation between nodes in calculation grid, assumed to be the same horizontally and vertically, over which the field is numerically determined (0.05 μm). The complex frequency amplitudes A are calculated from (A.3.21), using $A(f) = \cos(-\phi(f)) + j\sin(-\phi(f))$. Note that this function is nonzero only for $f \le \frac{NA}{\lambda_0}$ as frequencies higher than this are not passed by the objective.

The convolution procedure deletes spatial frequencies from the image which the objective will not pass, and aberrations introduce a phase distortion into those passed. If aberrations were not to be considered for example, all one would have to do is take the

2-D FFT of the electric field profile, set those spatial frequencies $|f|$ higher than NA/λ_0 to zero, and inverse Fourier transform to get the image. The intensity profile would be computed by taking the square of the amplitudes. Note that since spatial frequencies higher than NA/λ_0 are missing from the image, and not just attenuated, one cannot recover the original waveguide field.

A 120×120 array of amplitudes on a 0.05 μm square grid was calculated using the effective index method. In order to prevent periodic convolution from occurring, the matrix was augmented with zeros such that each dimension became at least twice as large minus one (256×256 was selected) [13] (Fig. A-12). The 2-D FFT of this matrix was multiplied by the complex pupil function matrix (Fig. A-11) and the inverse FFT taken to determine the image plane field profile (in object plane units). The intensity profile was then taken from the magnitude squared of the complex field profile.

$$\begin{bmatrix}
0 & f & 2f & \cdots & \cdots & 2f & f \\
f & \sqrt{2}f & \sqrt{5}f & \cdots & \cdots & \sqrt{5}f & \sqrt{2}f \\
2f & \sqrt{5}f & \sqrt{8}f & \cdots & \cdots & \sqrt{8}f & \sqrt{5}f \\
 & \vdots & & & & \vdots & \\
 & \vdots & & & & \vdots & \\
 & \vdots & & & & \vdots & \\
 & \vdots & & & & \vdots & \\
 & \vdots & & & & \vdots & \\
 & \vdots & & & & \vdots & \\
2f & \sqrt{5}f & \sqrt{8}f & \cdots & \cdots & \sqrt{8}f & \sqrt{5}f \\
f & \sqrt{2}f & \sqrt{5}f & \cdots & \cdots & \sqrt{5}f & \sqrt{2}f
\end{bmatrix}$$

Fig. A-11 Arrangement of spatial frequencies in the matrix.

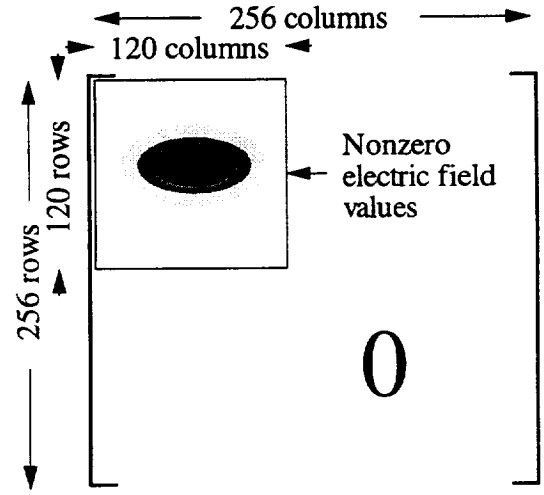


Fig. A-12 Construction of calculated electric field matrix.

Experimental setup

As shown in Fig. A-13, a CCD camera frame grabber and attenuator was added to the setup of Fig. A-1 to acquire the near field image. The integrated optic chip sat on a precision crossed-roller stage (Newport [14] ULTRAlign 561D) equipped with three Newport DS-4 adjusters. Fine adjustment of the objective-to-chip edge separation (focus) was performed by moving the sample stage. With the use of a graticule, it was found that the adjuster's fine movement was nonlinear; however a location existed where rotation of the fine-movement knob by one division gave 1 μm of movement. A Nikon 100×, 0.9 NA, differential interference contrast objective was used to image the output edge of the waveguide on to a lensless, windowless, monochrome, silicon, CCD camera (Cohu 4915-2010/0000). A 120×120 array of pixels was acquired from the camera and quantized to 8 bits (256 levels) by a Spiricon LBA-100 laser beam analyzer, an instrument which was described in Chapter 5. Since the camera had at least four times as many pixels as the analyzer, the analyzer captured one pixel in a 4×4 group, one in a 2×2 group, or every one over a particular area, depending on the degree of zoom selected, to fill in the 120×120 array. To reduce pixel-to-pixel intensity differences due to Fabry-Perot effects, the protective window over the CCD array had been removed. For

the same reason, an attenuator, needed to prevent the camera from saturating, was placed before the waveguide.

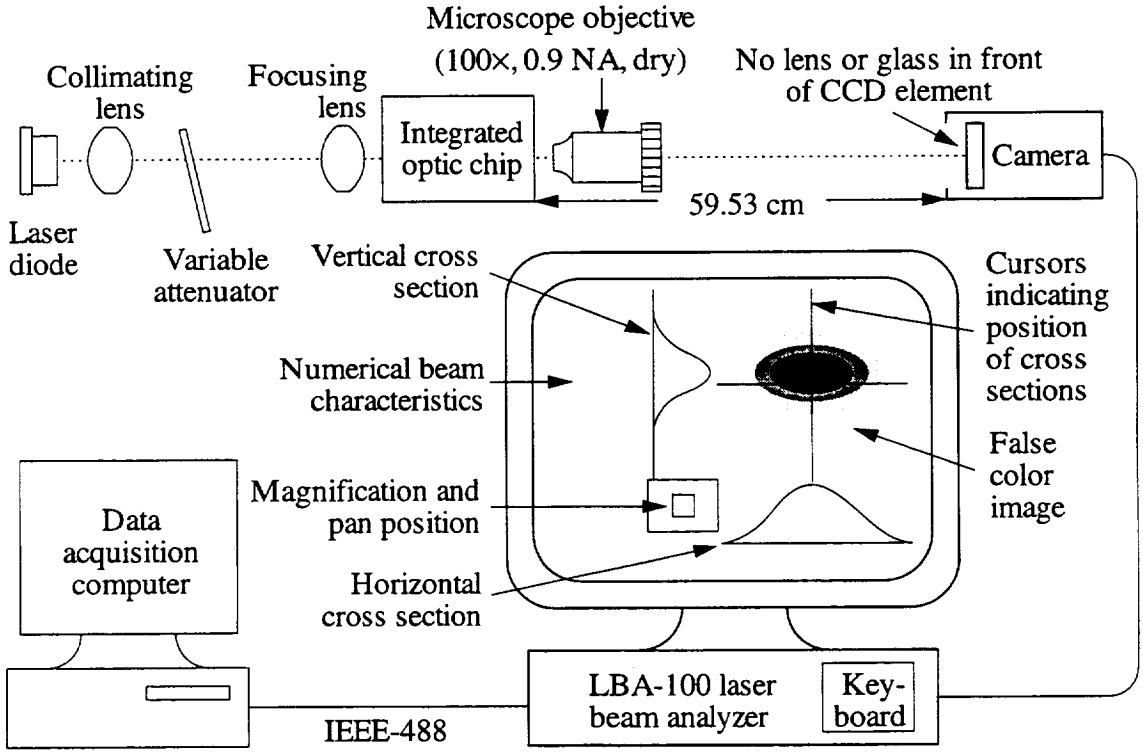


Fig. A-13 Setup to acquire near field image.

The LBA-100 presented the beam profile in the form of a false color image on a CRT monitor along with intensity plots through user selected cross sections and numerical beam characteristics. With an update rate of about 20 frames/sec, this display aided in fine focusing the 100× objective, for a change of 1 μm in its separation from the sample noticeably altered the image.

With a 32 frame memory, the LBA-100 could be set either to take one or a few frames at a time or continuously acquire frames. When making measurements, to reduce noise and provide an accurate zero reference, sixteen background frames were taken with the laser beam blocked and the same number of data frames with it not; then each frame was down loaded into a computer, and pixel by pixel, the sum of the background frames was subtracted from the sum of the data frames. All data was taken in a darkened room.

Near field measurements on planar waveguide

A single mode channel waveguide was placed in the setup shown in Fig. A-13 to determine the magnification factor. The camera and objective were mounted to micrometer driven translation stages. Magnification was taken to be the ratio of the camera translation to objective translation required to keep the imaged waveguide spot in the same location on the analyzer display: a magnification of 270 was measured. Next, a 3× objective, placed a short distance in front of the CCD, was used to focus the image of the waveguide to a spot about the size of one pixel. By translating the camera, a center-to-center pixel separation of 9.85 μm in the horizontal direction and 9.81 μm in the vertical direction was noted. The pixel separation was thus 0.0364 μm horizontally and 0.0363 μm vertically in object plane units.

The planar waveguide shown in Fig. A-14 was placed in the near field setup to assess the objective's spherical aberration. From the incoupling side, the finely focused spot of light had spread out to a fan when it reached the opposite cleaved edge of the sample. Fig. A-15 shows a contour plot of a near field image for an objective-to-chip separation which gave maximum peak intensity and the narrowest profile in the vertical direction ("best focus"). The waveguide was excited by 830 nm TE polarized light. Equation (A.3.21) predicts that if spherical aberration were zero ($C = 0$), the phase shift experienced by each spatial frequency f would be the same, except for a sign change, at the same distance $\zeta_0 + \Delta\zeta$ on either side of focus, and the intensity distribution would therefore also be the same. To determine C and the amount of defocusing ζ_0 for the image of Fig. A-15, intensity profiles were taken at other objective-to-chip separations. A slice from the middle of each of these scans, as well as from that of Fig. A-15 is given in Fig. A-16, together with the predicted intensity profile for the parameters $\zeta_0 = -1.9$ μm, and $C = 2.1$, which gave best agreement with all slices simultaneously. For $\Delta z < 0$, the objective-to-chip separation is larger than optimum. A Fast Fourier Transform method was used with $N = 256$, $\Delta x = 0.05$ μm, and to prevent circular convolution from

occurring, only the first 120 points entries of the matrix were nonzero. These points were sampled from $x = -3$ to 2.95 μm, where $x = 0$ lay at the center of the nitride layer. Little difference was noted when a denser grid was used. All of the experimentally measured curves have been scaled and shifted by the same amount to obtain maximum agreement with the calculated data. The agreement is reasonable given the uncertainty in setting the objective-to-chip displacement.

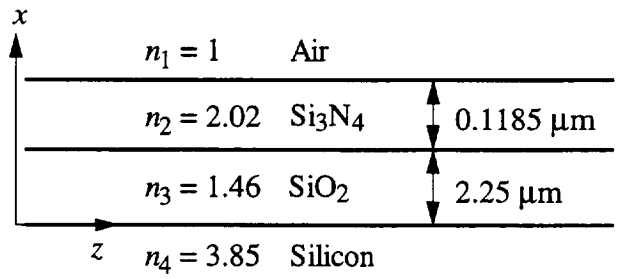


Fig. A-14 Silicon nitride planar waveguide on oxidized silicon substrate used for evaluating objective lens.

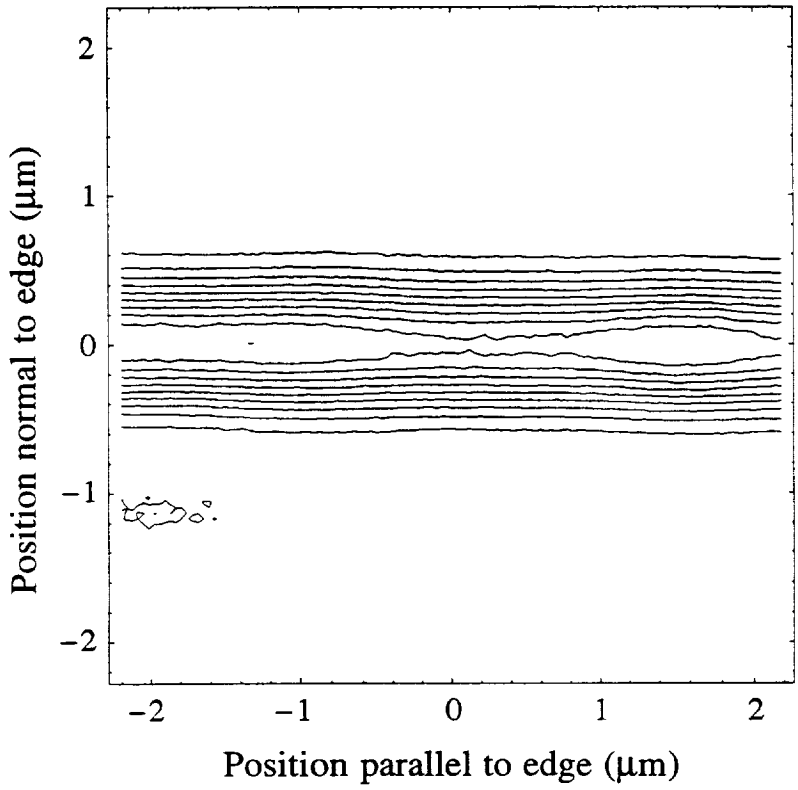


Fig. A-15 Intensity contours in 10% increments of maximum intensity measured on the chip of Fig. A-14 with the setup of Fig. A-13. The air side is toward the top of the plot, while SiO_2 is toward the bottom. The feature near −2, −1.1 is due to a side lobe just penetrating the 10% level (see Fig. A-16) while the 100% contour is near −1.3, 0. A slight horizontal variation in the intensity is also apparent.

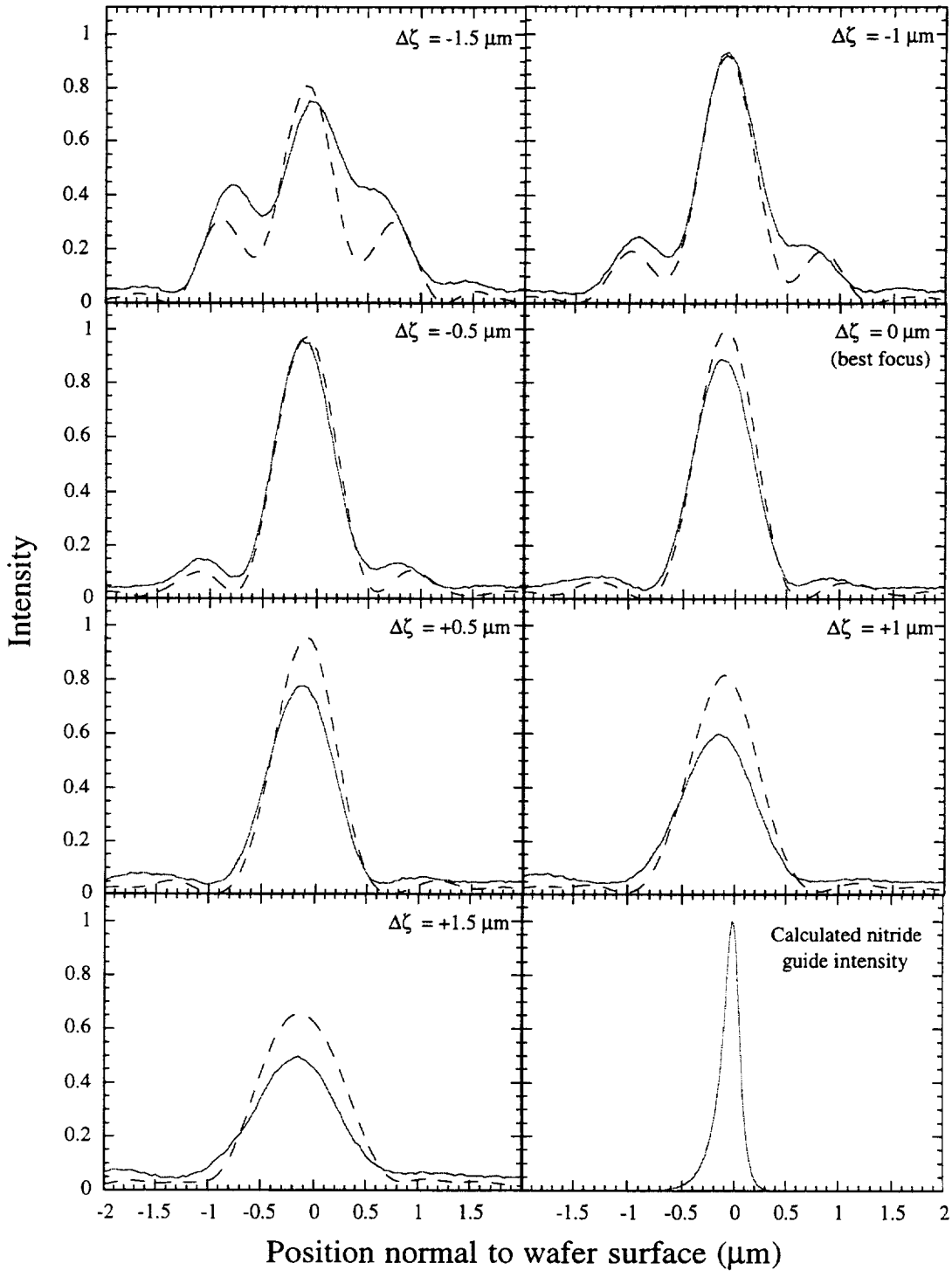


Fig. A-16 Measured and calculated near field intensity profiles for the planar guide of Fig. A-14. The air side is toward the right of the plots, while the SiO_2 is toward the left. The lower right plot is the theoretical intensity profile.

Near field measurements on channel waveguide

A silicon oxynitride channel waveguide typical of those used in this work, shown in Fig. A-17, was placed in the setup of Fig. A-13 and its near field image collected. Again the quasi-TE mode was excited with 830 nm light. Fig. A-18 shows the result, where the objective-to-chip separation has been set to provide the narrowest vertical field profile. Fig. A-19 shows the near field intensity calculated using the effective index method and Fig. A-20 includes in this calculation diffraction, spherical aberration $C = 2.1$, and defocusing $\zeta_0 = -1.9\ \mu m$ which partially offsets the aberration. Figure A-21, Fig. A-22, and Fig. A-23 repeat the previous set of comparisons using a channel waveguide on the same sample which has a channel width of 1.5 μm.

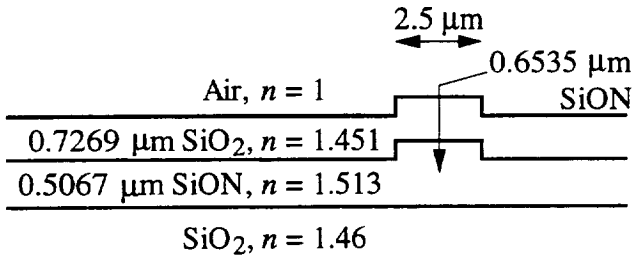


Fig. A-17 Near field imaged silicon oxynitride channel waveguide. This is one of the straight waveguides on the wafer bearing the Mach-Zehnder interferometer that was described in Chapter 5.

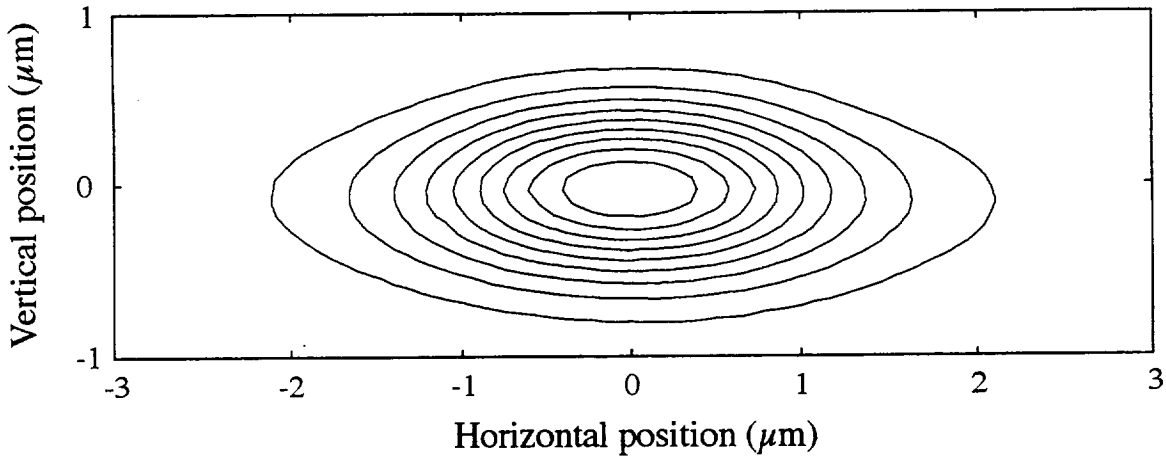


Fig. A-18 Measured near field intensity profile of the waveguide of Fig. A-17. Air is toward the top and the substrate toward the bottom. Intensity contours are plotted in 10% increments from 10% to 90% of the maximum.

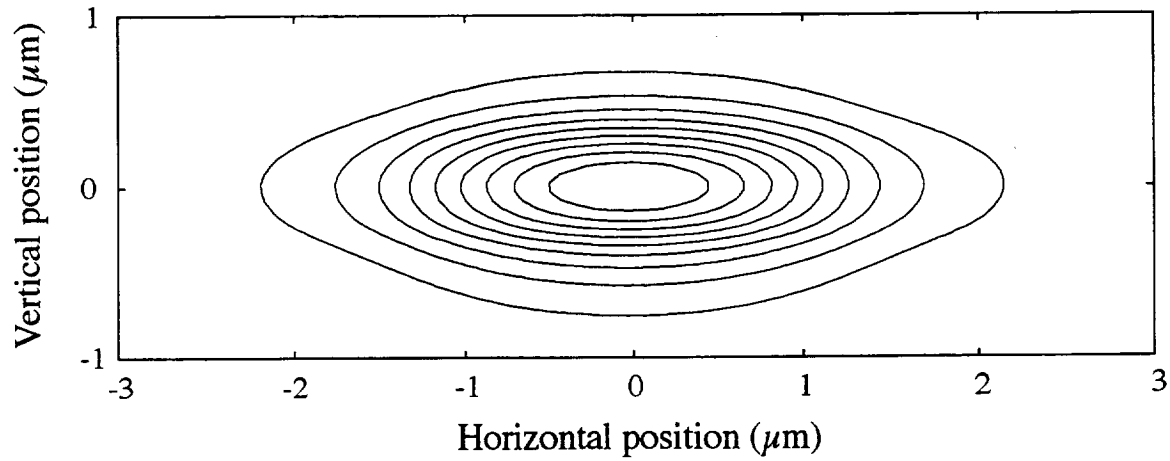


Fig. A-19 The intensity profile of the channel in Fig. A-17 calculated using the effective index method.

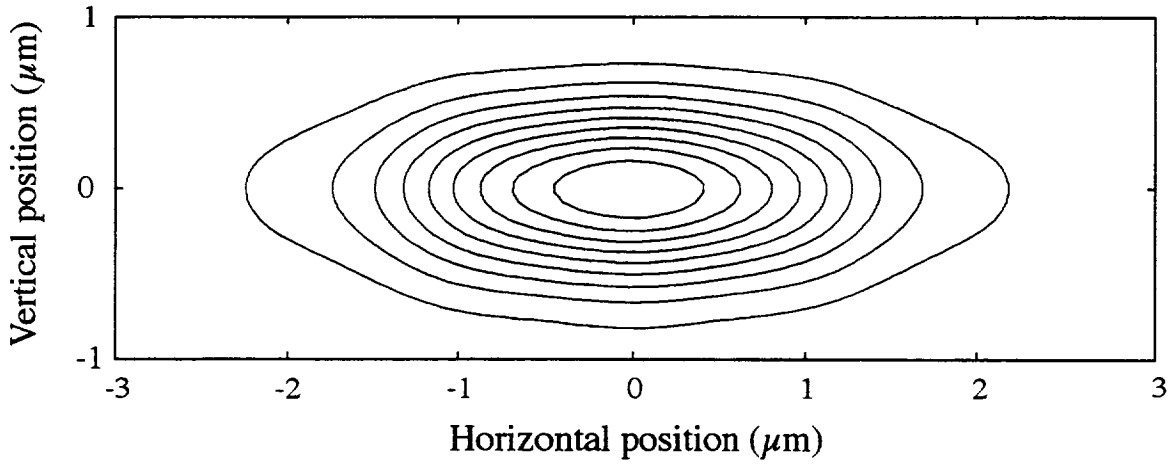


Fig. A-20 The same as Fig. A-19, except diffraction, spherical aberration, and defocusing (to partially offset the spherical aberration) has been included in the theoretical intensity profile.

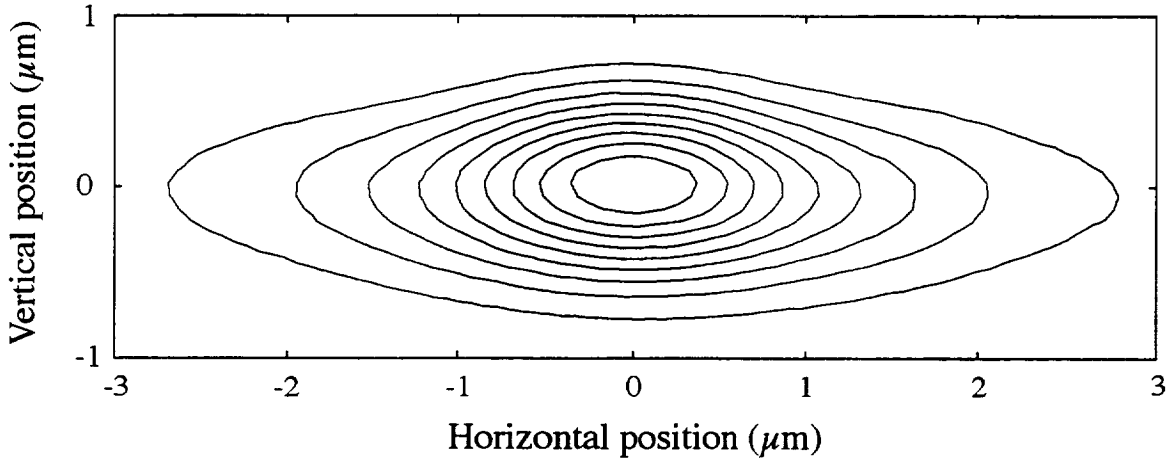


Fig. A-21 Measured near field intensity profile of the waveguide of Fig. A-17 except that the width was 1.5 μm.

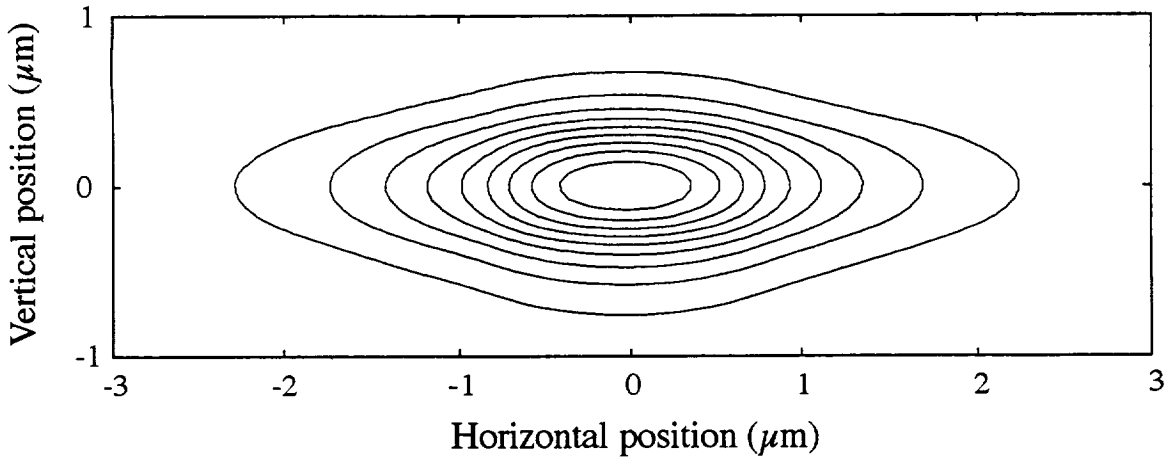


Fig. A-22 Same as Fig. A-21, except that this is the calculated effective index intensity contour.

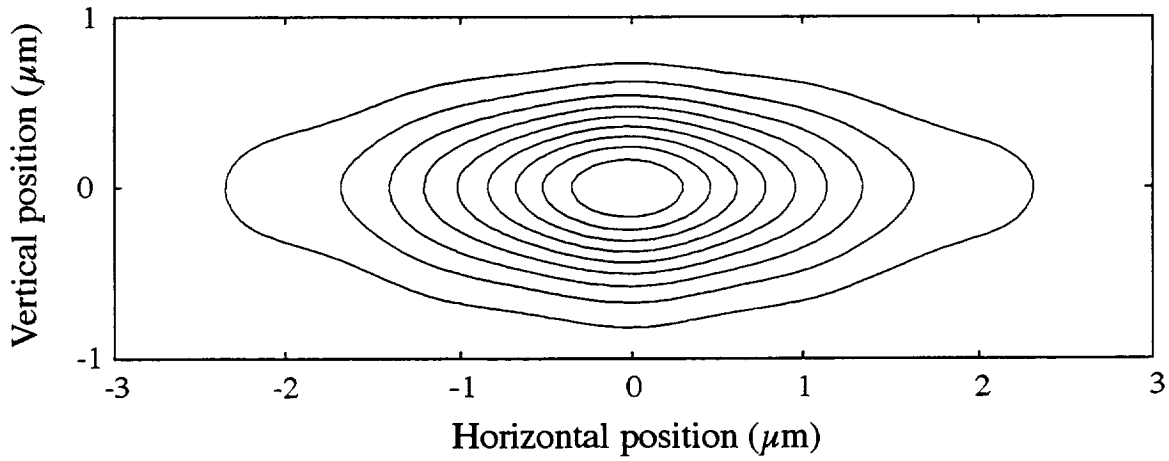


Fig. A-23 Same as Fig. A-22, except diffraction, spherical aberration, and defocusing has been included in the theoretical intensity profile.

4 CONCLUSIONS

Reasonable agreement was found between the imaged field profiles and those calculated using the effective index method, especially for the 2.5 μm wide channels. Diffraction, or the loss of higher spatial frequencies, together with spherical aberration in the objective, substantially distorts the field profile normal to the substrate in the thin

silicon nitride planar waveguides. It would be expected that when imaging a channel end face fabricated from such a planar waveguide that the intensity profile in a direction normal to the substrate would be significantly broadened, while that parallel to the wafer face could be faithfully reproduced, depending on the narrowness of the profile in that direction. The intensity profile in a direction normal to the wafer face of the SiON channels is only slightly broadened, while that parallel to the surface appears to be faithfully reproduced.

5 REFERENCES

[1] *Newport Catalog*: Newport Corporation, pp. 2.66 to 2.67, 1994.

[2] *Optics Guide 5*: Melles Griot Corporation, pp. 18-2 to 18-7, 1990.

[3] A. Yariv, *Optical Electronics*, 4th edition: Holt, Rinehart, and Winston, pp. 46-52, 1991.

[4] D. Marcuse, *Theory of Dielectric Optical Waveguides*, 2nd edition: Academic Press, Inc., pp. 330-334, 1991.

[5] E.-G. Neumann, *Single-Mode Fibers*, vol. 57: Springer-Verlag, pp. 374-392, 1988.

[6] E. Hecht, *Schaum's Outline of Theory and Problems of Optics*: McGraw-Hill, pp. 176-179, 1975.

[7] C. S. Williams and O. A. Becklund, *Optics: A Short Course for Engineers and Scientists*. Malabar, Florida: Robert E. Krieger publisher, pp. 100-104, 1984.

[8] J. W. Goodman, *Introduction to Fourier Optics*: McGraw-Hill, pp. 11-16, 1968.

[9] C. S. Williams and O. A. Becklund, *Optics: A Short Course for Engineers and Scientists*. Malabar, Florida: Robert E. Krieger publisher, pp. 215-221, 1984.

[10] J. W. Goodman, *Introduction to Fourier Optics*: McGraw-Hill, pp. 110-121, 1968.

[11] C. S. Williams and O. A. Becklund, *Optics: A Short Course for Engineers and Scientists*. Malabar, Florida: Robert E. Krieger publisher, pp. 173-195, 1984.

[12] C. D. McGillem and G. R. Cooper, *Continuous and Discrete Signal and System Analysis*: Holt, Rinehart, and Winston, pp. 459, 1991.

[13] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, pp. 118-120, 1975.

[14] Newport Corporation, 1791 Deere Ave., Irvine, CA 92714.