NASA CR 3151 c.1

NASA Contractor Report 3151

An Investigation for the Development of an Integrated Optical Data Preprocessor

C. M. Verber, D. W. Vahey, R. P. Kenan,
V. E. Wood, N. F. Hartman, and C. M. Chapman

NASA

NASA Contractor Report 3151

An Investigation for the Development of an Integrated Optical Data Preprocessor

C. M. Verber, D. W. Vahey, R. P. Kenan,
V. E. Wood, N. F. Hartman, and C. M. Chapman
Battelle Memorial Institute
Columbus, Ohio

NASA

National Aeronautics
and Space Administration

**Scientific and Technical
Information Office**

1979

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

LIST OF FIGURES

SUMMARY

This report discusses the results of a two-year effort which culminated in the successful fabrication and testing of a laboratory model of a multichannel integrated-optical data preprocessor. The preprocessor is designed to operate on the set of analog voltages which are the outputs of a set of remote sensors. By comparing the data sets to one or more reference sets, a signal can be generated which will result in useless data being discarded before transmission. In the final device, the criteria for flagging data to be discarded will be capable of being preset or remotely manipulated.

The preprocessor which utilizes holographic subtraction to simultaneously compare an optical phase analog of N signal voltages with N predetermined reference voltages, is wholly integrated on a $LiNbO_3$ single crystal with the present exceptions of the optical source, lens, and detector. It has the potential for three distinct operating modes, a "screening mode" in which incoming data are compared to a single, holographically recorded reference set, an "identification mode" in which incoming data are compared to a large number of reference sets and a best fit is identified, and a "self-subtraction mode" in which the preprocessor automatically adjusts to slowly changing input signals and responds only to rapidly varying voltages.

The present program dealt first with the fabrication of the individual components required for the preprocessor and then with the problem of integrating them all on a common substrate to form a complete integrated optical circuit. Such a circuit, complete except for source detector and detector optics was fabricated. It consists of a beam splitter, a TIR mirror, a phase shifter, a set of three data-input electrodes and a hologram region all integrated on a 14 x 25 mm $LiNbO_3$ outdiffused waveguide. Tests on this laboratory model showed that a voltage which produced a one degree phase shift in a single signal channel produced an output from the preprocessor which had a 2:1 signal-to-noise ratio. This result indicates that the preprocessor has sufficient dynamic range to be of use in a variety of data screening and sorting applications.

I. INTRODUCTION

This report summarizes a program which resulted in the successful fabrication and demonstration of an integrated optical circuit designed to perform a parallel processing operation on sets of analog voltages. The device was conceived while carrying out investigations associated with the NASA-sponsored program entitled "Feasibility Investigation of Integrated Optics Fourier Transform Devices" (NASA CR-2869, 1977) in response to the NASA requirement to alleviate transmission, storage and processing loads by performing, at the sensor site, some preprocessing of data taken by remote sensors.

The preprocessor, which utilizes holographic subtraction to simultaneously compare N analog signal voltages with N predetermined reference voltages, is wholly integrated on a $LiNbO_3$ single crystal with the present exceptions of the optical source, lens, and detector. During the first phase of the program the principal goal was to ascertain that the materials and the individual components of the preprocessor would function as required. In achieving this goal, it was necessary to consider the compatibility of the various components with one another and the mutual compatibility of the individual fabrication techniques as well as the performance of the components considered singly. In the course of carrying out the first phase of this program, we made a number of major accomplishments in the fabrication of integrated optics components. Among these were:

- Fabrication of the first $LiNbO_3$ waveguide geodesic lens,
- Development of techniques for polishing TIR mirrors on $LiNbO_3$ waveguides,
- Fabrication of high efficiency metal-over-photoresist gratings for waveguide beam splitters,
- Demonstration of high S/N holographic subtraction using waveguide holograms.

As the result of this work it was concluded that the individual component fabrication should present no severe problems and that integration techniques compatible with all components could be devised. Consequently, the second phase of this program was undertaken with the goals of fabricating and

demonstrating the operation of a three-channel laboratory model of the pre-processor. In addition to satisfying these goals, the following have been accomplished.

- Discovery and first suggested use of holographic self-subtraction in $LiNbO_3$
- Development of alignment techniques for fabrication of integrated optic circuits
- Development of mathematical description of operating modes of preprocessor
- Development of theory for diffraction efficiency and beam quality of two dimensional beam-defined gratings.

This report deals primarily with Design and Fabrication of Pre-processor Components, System Integration, and System Test and Operation (Sections IV, V and VI). These main sections of the report are preceded by discussions of the Function and Operation of the Preprocessor and of its several Modes of Operation. We conclude that the results of this program are sufficiently encouraging that an effort should be made to construct a 16-channel preprocessor which can be tested by NASA with data inputs simulating the outputs of a variety of sensors.

Use of trade names or names of manufacturers in this report does not constitute an official endorsement of such products or manufacturers, either expressed or implied, by the National Aeronautics and Space Administration.

3

II. FUNCTION AND OPERATION OF THE PREPROCESSOR

ROLE OF THE PREPROCESSOR

The various NASA missions which are dedicated to the remote sensing
of terrestrial features are characterized by the acquisition, transmission,
storage and processing of enormous amounts of data. as is suggested in
Figure II-1, a large fraction of this data is, for one reason or another con-
sidered useless and is therefore discarded. A major failing of present data
handling systems is that processing this useless data accounts for a substan-
tial part of the total time lag between data acquisition and delivery of the
data to the customer in a usable form. The preprocessor is designed to alle-
viate this situation in cases where the data are in the form of sets of analog
voltages delivered simultaneously from a number of sensors. In its simplest
form, the function of the preprocessor is to provide a signal indicating
whether an incoming data set is useful or useless according to some predeter-
mined criteria (Fig. II-2). A more advanced version of the preprocessor should
be capable of performing a classification or identification function (Fig. II-3),
thereby affecting even greater economies in the transmission, storage and
processing of the data acquired by remote sensors.

The basic operation which the preprocessor performs is to generate
an output signal which is proportional to the sum of the absolute values of a
function of the channel by channel difference of the signal and reference
voltages. As will be discussed in detail below, this operation is accomplished
by holographic subtraction after first utilizing the electrooptic effect to
generate an optical phase analog of the N simultaneous voltages which comprise
the input data set. The simplest application of this operation is in the
screening mode which is indicated schematically in Figure II-4. As an example
of its utility, consider the case of a multispectral scanner whose mission is
to look for perturbations in the spectral signature of the sunlight reflected
from the ocean due to oil spills, plankton, etc. If the preprocessor is pro-
grammed with the reference set characteristic of clear sea water, then it will
generate a signal greater than some predetermined threshold value only when
the incoming data indicate a significant departure from the clear water
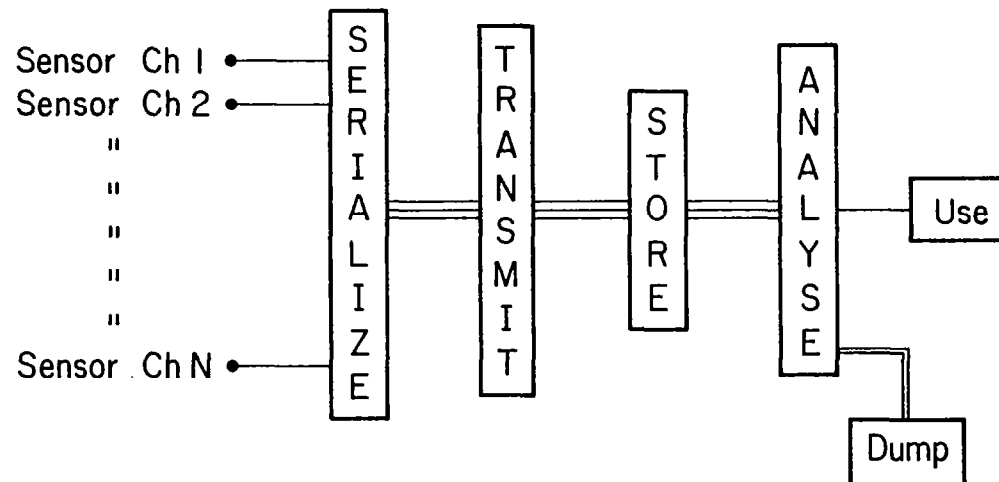
FIGURE II-1. CURRENT PROCEDURE FOR HANDLING REMOTELY ACQUIRED, MULTI-
SENSOR DATA. UNDESIRED DATA REPRESENT A SIGNIFICANT
FRACTION OF THE SYSTEM LOAD.
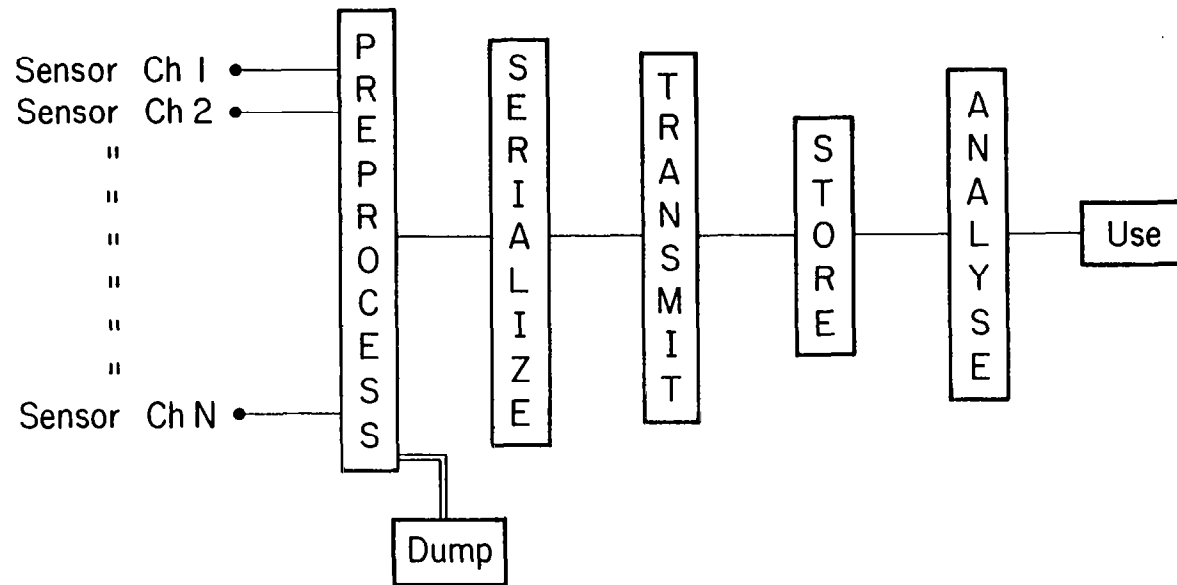
FIGURE II-2.   USE OF A PREPROCESSOR TO ELIMINATE USELESS DATA PRIOR
TO TRANSMISSION.
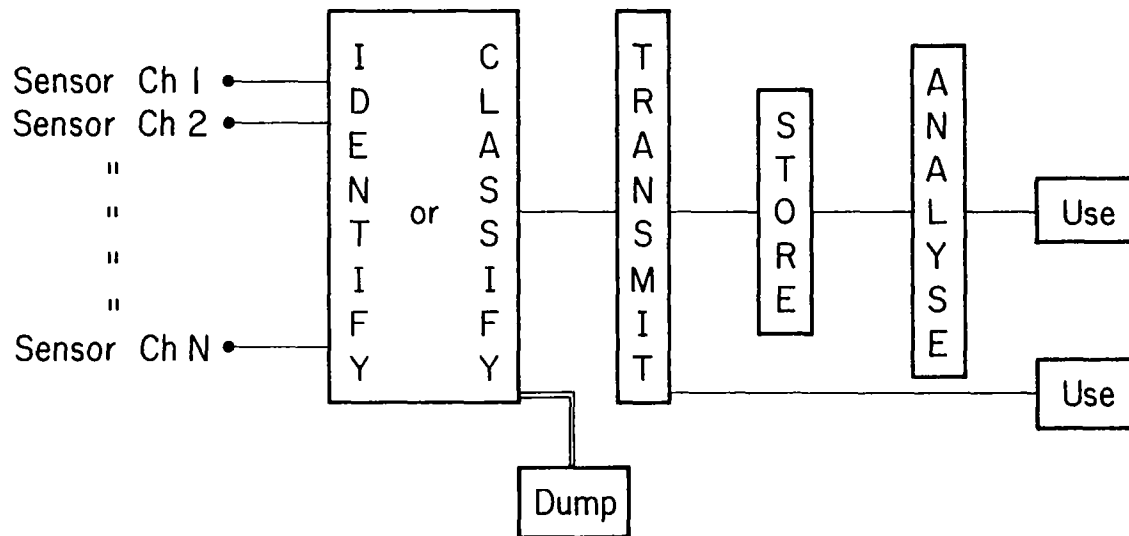
FIGURE II-3. USE OF A MORE ADVANCED PREPROCESSOR TO CLASSIFY DATA
BEFORE TRANSMISSION.

Signal in
(n-channels)

S_1 S_2 · · · · · · · · · S_n

R_1 R_2 R_n

Comparator with
pre-set reference
levels

$$\sum_{i=1}^{n} |S_i - R_i|^2$$

Computes sum
of differences

>L   ≤L

Indicates magnitude
with respect to
predetermined level
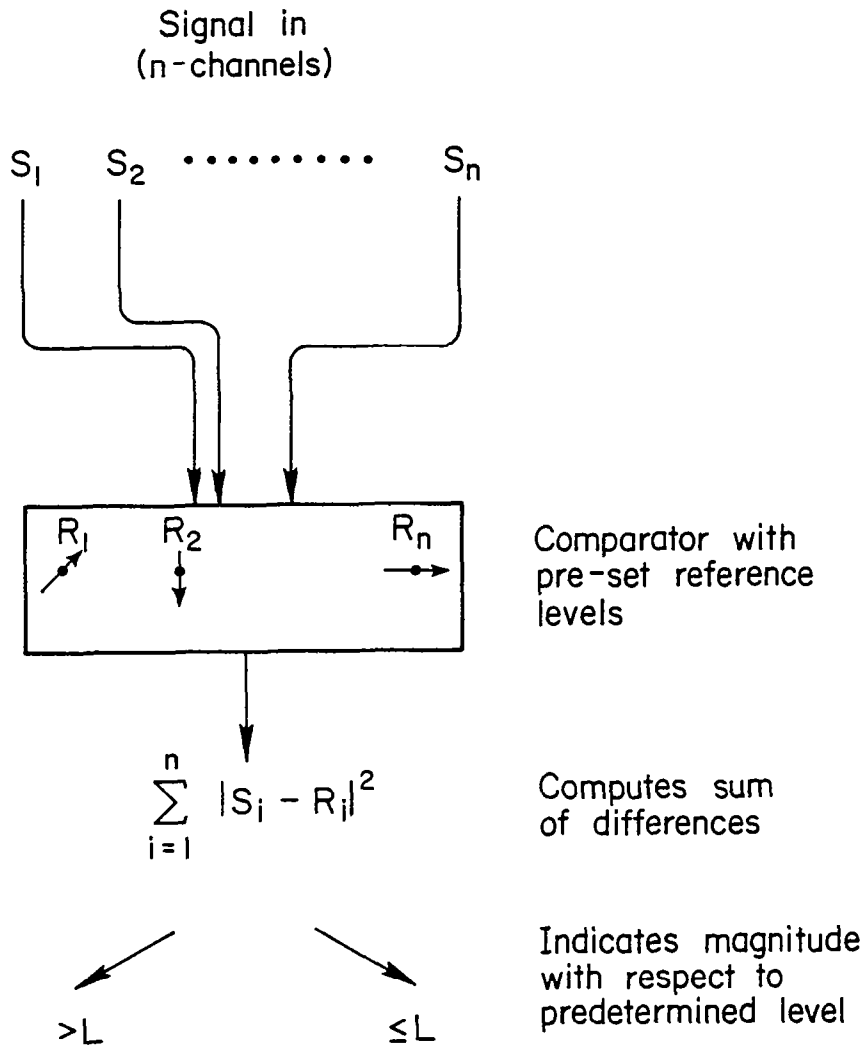
FIGURE II-4.   SCHEMATIC REPRESENTATION OF A PREPROCESSOR
OPERATING IN A SCREENING MODE.

spectrum. All unflagged data will be dumped before transmission, effecting an estimated 95% reduction in data load.

A more sophisticated mode of operation of the preprocessor is the classification or identification mode (Fig. II-5). Here, the data set to be identified is compared to a large number of reference sets which can be read into the preprocessor from some parallel storage medium, e.g., an N-track mangetic tape. The identification occurs when the comparison of the data to a particular reference set results in a null output from the preprocessor. Obviously this mode of operation will be useful when it is anticipated that the incoming data will represent one of a finite number of previously characterized situations.

A final mode of operation of the preprocessor is based upon the phenomenon of "holographic self-subtraction" which was discovered during the course of this program. This phenomenon, which is the result of the specific mechanism by which a hologram is written in $LiNbO_3$, provides a mechanism by which the preprocessor can adjust itself to a slowly varying background and, in effect, teach itself the reference set. It then provides an output signal which indicates a sudden departure from this learned reference set.

As an example of the application of the self-subtraction mode, consider again the case of a multispectral sensor system looking at the ocean. In the self-subtraction mode, the preprocessor can adjust to the slow spectral change which results from the change in solar elevation. Furthermore, as illustrated in Fig. II-6, after a short time the system adjusts to changes in the background signals due, for example, to the transition from the main body of the ocean to the Gulf Stream. The response time of this effect is controlled by material and optical parameters and can be less than 1 second. Its operation and potential for application have not yet been explored in detail. However, it is expected that all three modes of preprocessor operation will find application in monitoring of the output of remote sensor systems.

OPERATION OF THE PREPROCESSOR

Holographic subtraction is a natural choice of phenomenon to implement the parallel comparison of N signals to N standards. The

$$\sum_{i=1}^{n} |S_i - R_i^a|^2 = L_a$$

$$\sum_{i=1}^{n} |S_i - R_i^b|^2 = L_b$$

$$\sum_{i=1}^{n} |S_i - R_i^c|^2 = L_c$$
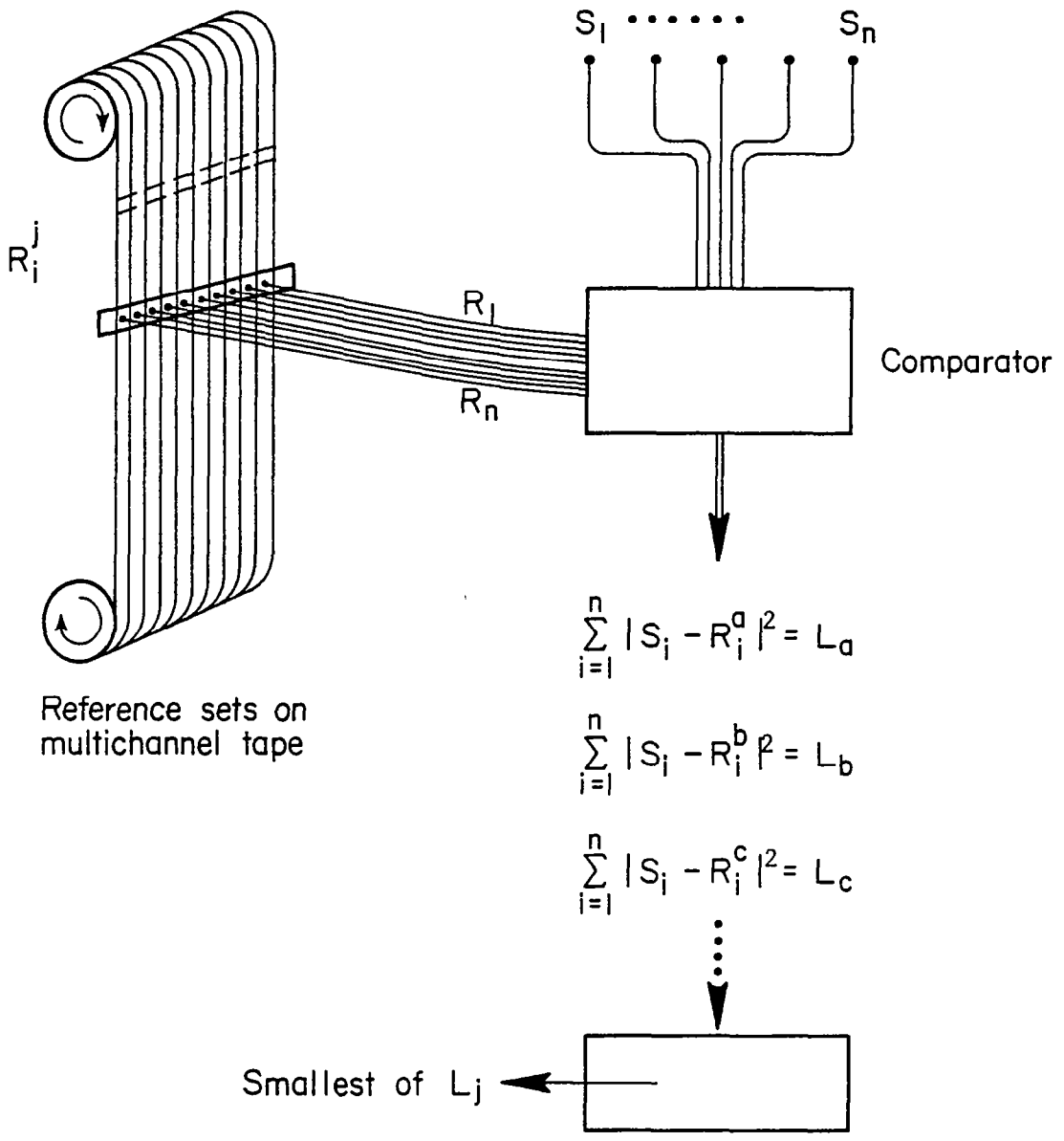
FIGURE II-5.   SCHEMATIC REPRESENTATION OF A PREPROCESSOR
OPERATING IN AN IDENTIFICATION MODE.
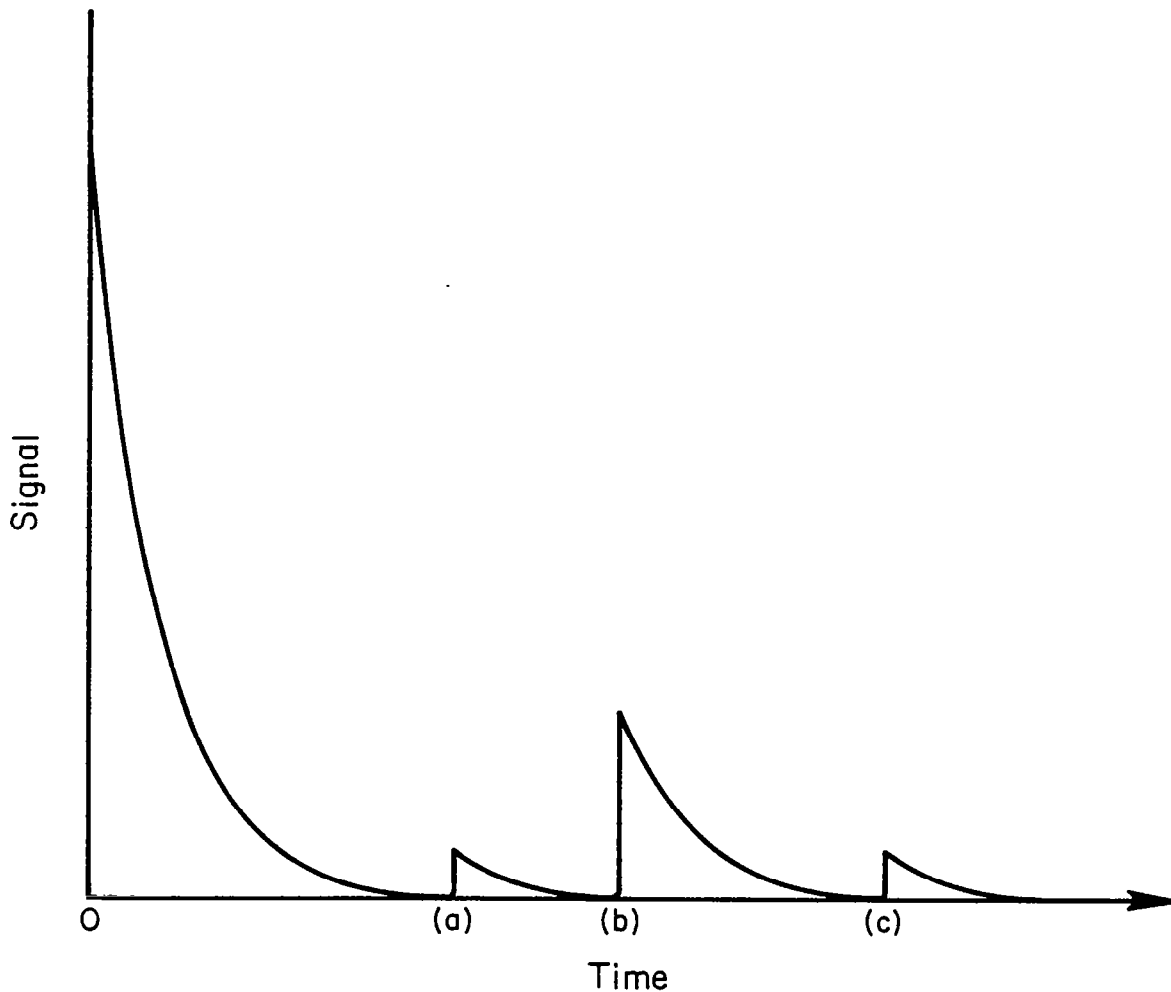
FIGURE II-6.   PREPROCESSOR OUTPUT IN SELF-SUBTRACTION MODE
SHOWING INITIAL HOLOGRAM WRITING FOLLOWED BY
(a) RESPONSE TO AN ANOMALY (b) ADJUSTMENT TO
ALTERED BACKGROUND AND (c) RESPONSE TO ANOMALY
IN NEW BACKGROUND.

11

incorporation of the entire holographic subtraction system in a monolithic integrated optic structure imbues the system with the qualities of rigidity, small size and manufacturability which are required for space deployment. In this section we will outline the operation of the integrated optical holographic subtraction system. Details of component design and fabrication and of the hologram formation process in $LiNbO_3$ will be deferred to later sections of this report. However, before preceding to the description of the device it will be necessary to digress briefly on the subjects of the electrooptic information input and holographic subtraction.

Electrooptic Information Input

The electrooptic effect is the term used to describe the change in index of refraction which is caused in certain materials by the application of an electric field. In particular, the index change due to the linear electro-optic effect or Pockel's effect is described by

$$\Delta n = \frac{1}{2} n^3 rE \qquad (II-1)$$

where n is the index of refraction, r is the appropriate term of the electro-optic tensor and E is the applied field.

As an optical plane wave of (vacuum) wavelength λ_o passes through a transparent medium of length ℓ and index of refraction n the phase upon leaving the medium $\phi(\ell)$ is related to the phase upon entering the medium $\phi(0)$, by

$$\phi(\ell) = \phi(0) + 2\pi n\ell/\lambda \quad . \qquad (II-2)$$

Obviously, if the material is electrooptic $\phi(\ell)$ can be altered by the application of an electric field. The phase difference due to the application of a field E is

$$\Delta\phi = \frac{2\pi\ell}{\lambda_o} \Delta n(E) = \frac{\pi\ell n^3 rE}{\lambda_o} \qquad (II-3)$$

To provide a feeling for the magnitude of this effect we note that for $\lambda_o = .633$ μm, to achieve a π phase shift in the optimum direction in

LiNbO$_3$ (r = r$_{33}$ = 32.2 x 10^{-12} V/m, n = 2.2) it is necessary to apply 9.2 x 10^5 V/m or 0.92 V/μm along a 2 mm path.

A sketch of a stylized three channel electrooptic data input device in shown in Fig. II-7. Here we make the assumption that we can generate a channelized plane wave (i.e., diffraction is ignored), and we show explicitly the fact that positive as well as negative phase shifts can result from the Pockel's effect. An integrated-optical version of such a phase shifter is used as the data input device for the preprocessor.

Holographic Subtraction

Optical subtraction is possible because of the fact that the amplitudes of two superimposed, mutually polarized coherent beams add algebraically. Thus, if the amplitudes of two beams are equal in magnitude but differ by $(2n + 1)\pi$ in relative phase, there will be a local cancellation. This is, for example, the cause of the central dark spot in a well-aligned Michelson interferometer when the path length variations between the arms differ by $(2n + 1)\lambda/2$.

Consider now a hologram produced in the conventional manner depicted in Fig. II-8(a). As shown in Fig. II-8(b), reconstruction of this hologram enables a viewer to see a virtual image of the object in its original position.

In the third part of the figure, the shutter has been removed from the signal beam so that the viewer sees both the virtual object and, by looking through the hologram, the real object illuminated by the signal beam. If a conventional recording material is used and if there has been no change in the object or in the optical system between the hologram formation and this viewing, the observer will see an object which is brighter than that which is seen with either R or S beam blocked. However, if a phase shift is introduced into one of the beams, the brightness will change, passing through one minimum and one maximum as the introduced phase shift goes from 0 to 2π. In the simplest case, where both the direct and the holographically reconstructed beams are of equal intensity, the brightness can be made to pass through zero. This is referred to as holographic subtraction. Any region of the object which has changed between the hologram formation and the subtraction process will distort the wavefront and will produce a bright spot on the dark background.

13

$\phi = \text{const}$

Plane
wave

ϕ

ϕ_3

ϕ_2

ϕ_1

V_1 V_2 V_3
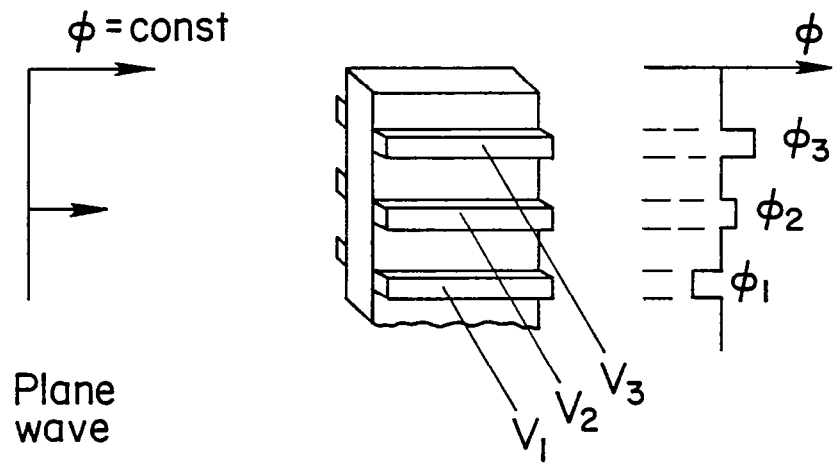
FIGURE II-7.   SCHEMATIC REPRESENTATION OF A THREE-CHANNEL
ELECTROOPTIC PHASE SHIFTER USED TO CREATE AN
OPTICAL PHASE ANALOG OF A SET OF VOLTAGES.
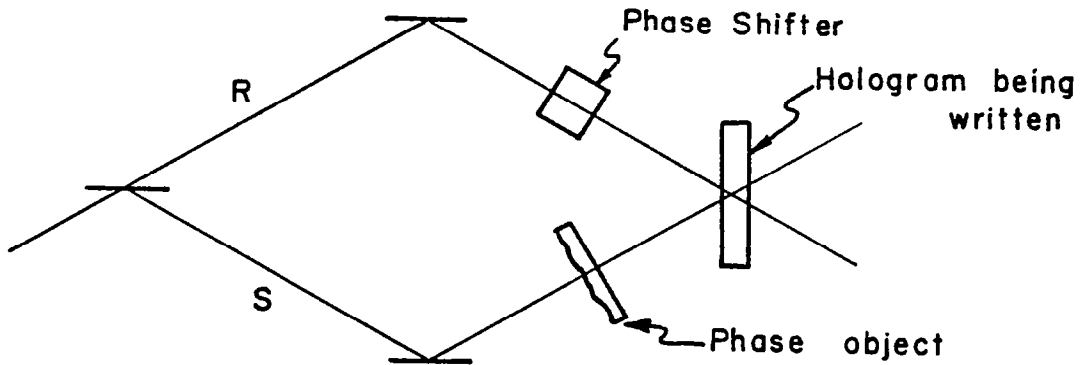
(a) THE FORMATION OF THE HOLOGRAM OF A PHASE OBJECT
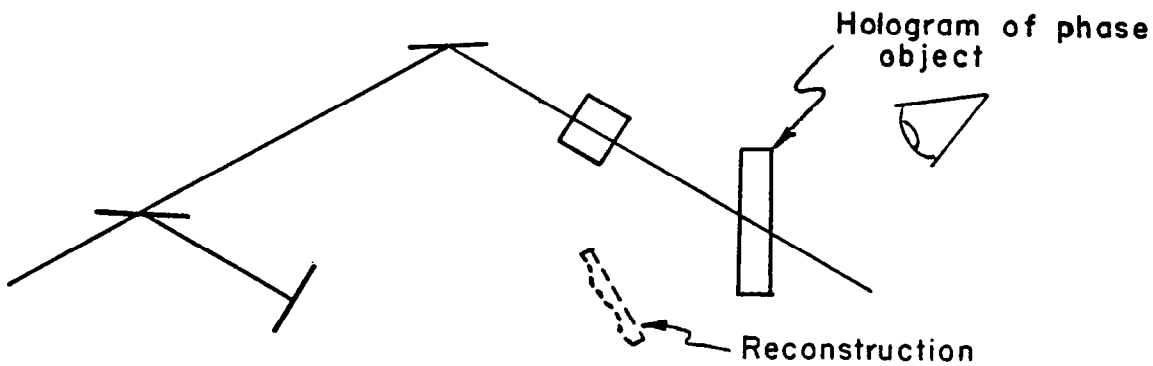
(b) RECONSTRUCTION OF OBJECT WITH SIGNAL BEAM BLOCKED
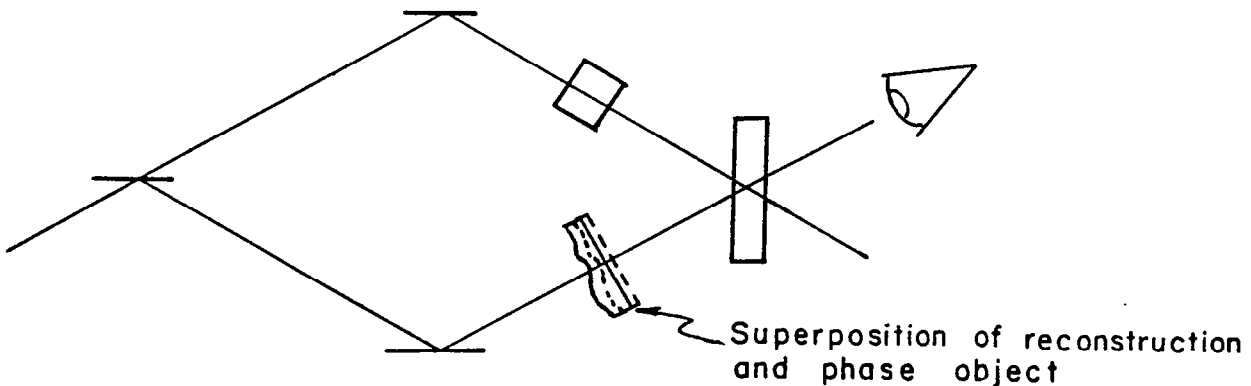
(c) SUPERPOSITION OF OBJECT AND IMAGE. AN APPROPRIATE
PHASE SHIFT WILL RESULT IN THE OBSERVER SEEING
A DARK FIELD

FIGURE II-8. THE HOLOGRAPHIC SUBTRACTION PROCESS.

The Integrated Optics Data Preprocessor Concept

The techniques of integrated optics will allow the concepts discussed above to be utilized in a compact, rigid structure capable of comparing a multiplicity of parallel analog inputs to a standard set at exceedingly high data rates. We first present a straightforward version of this device. In the following section we will discuss some of the system considerations which will influence the final design choice.

A schematic of one configuration of the comparator is shown in Fig. II-9. The waveguide is an outdiffused layer on the surface of a Y-cut $LiNbO_3$ slab. Light is coupled in via a grating coupler, and split into two beams by a grating beam splitter. The signal beam passes under a set of data input electrodes which, via the electrooptic effect, impress a signal pattern upon the guided wave. Assume, for the present, that the reference electrodes do not exist. The reference beam passes between two electrodes which, when energized, impart a suitable phase shift to the reference beam. The two beams then intersect in the hologram region.

In order to use the device, a set of reference voltages is applied to the data-input electrodes. In the ocean monitoring example, these voltages would be those corresponding to the clear water spectrum. A hologram is then made using the index-of-refraction pattern induced by this set of voltages as the object. This hologram is made in such a way as to be stable over some predetermined time period. Now, an appropriate phase shift is introduced into the reference beam and voltages representative of the unknown data set are impressed upon the electrodes. If they are identical to the reference set, the holographic subtraction process ensures that no light reaches the detector. If the data do not match the reference set, light arrives at the detector in proportion to the mismatch. Thus, a simple discriminator at the detector output can be set to flag data which exceed a predetermined degree of mismatch so that this data can be transmitted or stored for further analysis.

To implement the identification mode requires only a simple modification of the system which consists of the addition of reference electrodes, identical to the first set and located over the signal beam between the first set of electrodes and the hologram region. The reference hologram is made with

16
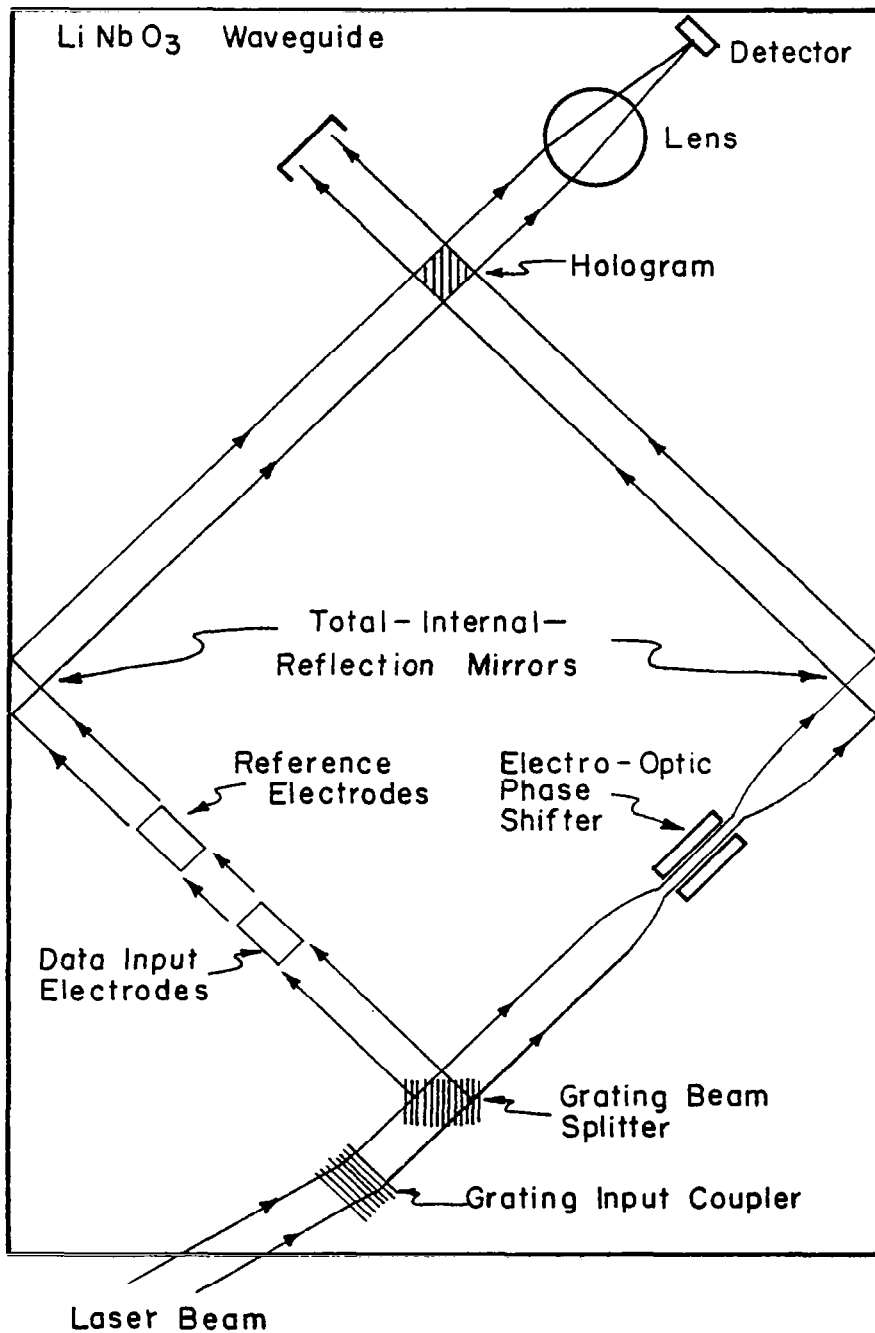
FIGURE II-9.   SCHEMATIC OF ONE CONFIGURATION OF THE
               INTEGRATED OPTICS DATA PREPROCESSOR.

no voltages on any of the electrodes. It therefore contains only information about the various optical imperfections in the system. During the subtraction steps these imperfections will be compensated for since they have been written into the hologram.

To use this modification of the device, the data voltages are impressed upon the first electrode set, and a negative set of reference voltages are impressed upon the second electrode set. If the two sets are equal in magnitude but opposite in sign, there will, at least to zero order, be a complete cancellation of their respective effects upon the wave front. If a library of reference voltages exists, it should be possible to cycle through the entire library of say, 10^3 to 10^4 reference sets during 10^{-3} sec to find a best fit to the unknown set. Thus, the device has the potential of being able to identify as well as flag interesting data.

III. THE PHOTOREFRACTIVE PROCESS AND HOLOGRAPHIC SUBTRACTION IN $LiNbO_3$

The operation of the preprocessor, its dynamic range, signal-to-noise characteristics, sensitivity, and speed of response all depend, to some extent upon the details of the photorefractive process by which a hologram is formed in $LiNbO_3$. In this section the phenomenological theory of photorefraction, the dependence of the effect upon material parameters and the details of the hologram formation process in $LiNbO_3$ are described and related to device performance and modes of operation.

THE PHYSICS OF THE PHOTOREFRACTIVE PROCESS

In order to record a hologram in either a three dimensional or a waveguide configuration, a photosensitive material is required. Ideally, this material should undergo an index-of-refraction change rather than an optical density change as the result of optical irradiation, so that holograms with high diffraction efficiencies can be written. A material which exhibits very low scattering is also desirable. In addition, for an integrated optics device, the material must be capable of existing in or on the waveguide, or ideally, actually be the waveguiding material. Crystalline $LiNbO_3$, in which holograms are recorded by means of a photorefractive mechanism unique to pyroelectric materials, satisfies each of the above criteria.

The term photorefraction designates photoinduced refractive-index changes. In $LiNbO_3$, the refractive-index changes arise from the generation of a nonuniform distribution of electrons trapped at impurity sites within the crystal.[1] When the crystal is illuminated, electrons are excited into the conduction band where they drift and diffuse until they are retrapped[2] as illustrated in Fig. III-1. For a spatially structured light beam, such as an interference pattern, the general tendency is for the population of trapped electrons to increase in the dark regions of the crystal and to decrease in the illuminated regions. The process is shown schematically in Fig. III-2.

The associated space-charge electric fields produce a refractive-index change by means of the linear electrooptic effect. The process
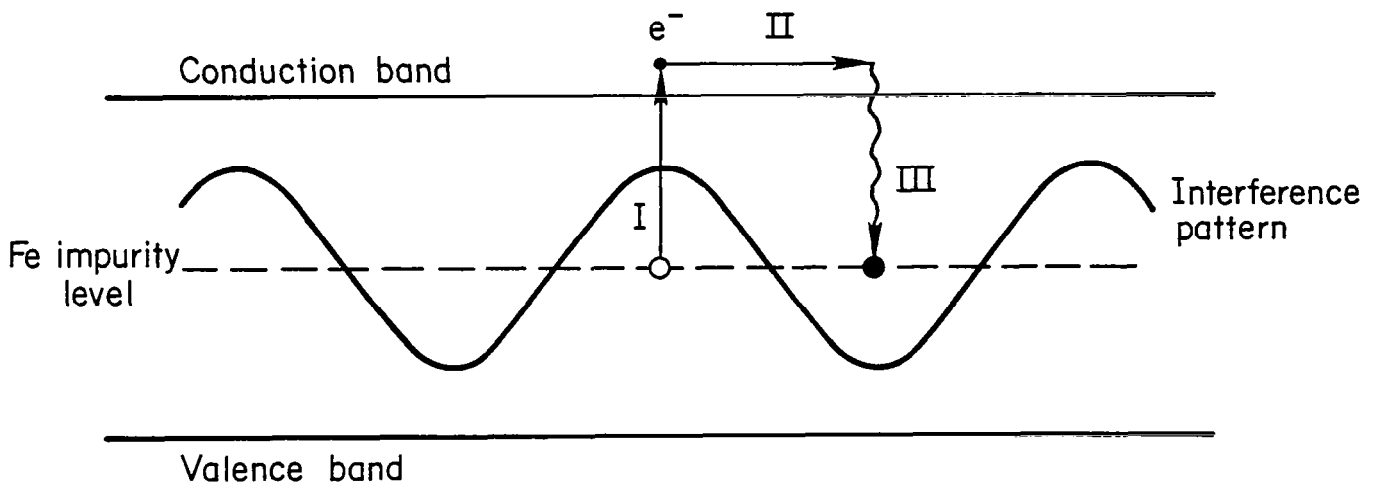
FIGURE III-1.   PHOTOREFRACTIVE MECHANISM FOR THE FORMATION
OF HOLOGRAMS IN LiNbO$_3$, CONSISTING OF (I)
ELECTRON EXCITATION, (II) ELECTRON MIGRA-
TION, (III) ELECTRON RETRAPPING.
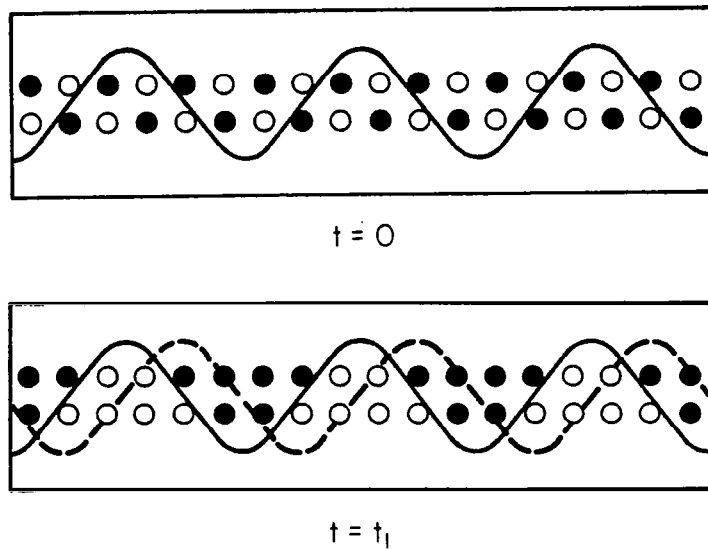
FIGURE III-2. CONCEPTUAL DRAWING SHOWING REDISTRIBUTION OF
FILLED TRAPS DURING ILLUMINATION BY AN INTER-
FERENCE PATTERN. OPEN CIRCLES ARE UNFILLED
TRAPS. CLOSED CIRCLES ARE FILLED TRAPS. SOLID
CURVE DESIGNATES INTENSITY VARIATION. DASHED
CURVE DESIGNATES REFRACTIVE INDEX VARIATION.

continues uniformly until the space-charge fields become so large that they oppose further migration of electrons.[3]   Often the fields are as large as 10^4 V/cm, corresponding to refractive-index changes greater than 10^{-4}.   These index changes are large enough to produce holograms having diffraction efficiencies approaching unity.

The Origin of Donors and Traps

Photorefractive sensitivity is observed to be largest in $LiNbO_3$ crystals that have a high concentration of iron.[1]   Since iron is a naturally occurring impurity in $LiNbO_3$, even samples that are described as nominally pure have sufficient iron for the formation of efficient holograms. Fe^{2+} impurities are known electron donors, and Fe^{3+} impurities are known traps.[4]   However, photorefractive effects have also been observed in crystals doped with other ions, such as Cu and Mn,[5]   and holograms formed by two-photon processes appear to rely on transient photoinduced color centers for sensitization.[6] Trapping sites necessary for the formation of long-duration holograms may be associated with crystal imperfections other than impurities, such as oxygen-site vacancies.

The Migration of Photo-Excited Electrons

After an electron is excited into the conduction band, it may be retrapped at the same location or it may migrate to other regions of the crystal.   This migration is partially diffusion controlled, but a more significant mechanism is electron drift associated with a bulk photovoltaic effect.[1,7] This effect originates in the pyroelectric nature of the $LiNbO_3$ host.   Fe impurities are separated from neighboring Nb ions by distances that vary in the positive and negative c-directions.   Electrons excited from Fe impurities consequently exhibit a preference to motion in one of the two directions, along which they drift until they are retrapped.   For uniform illumination the migration establishes a photovoltage, the sign of which shows that electron migration occurs along the + c-axis.   The effect is similar to that which would be observed if the electrons drifted in response to a negatively directed

electric field. Indeed, the migration can be speeded or slowed by applying a suitable voltage to the crystal. The field required to counteract the photovoltaic effect is found to be about 5×10^4 V/cm.[7]

For periodic illumination, as from an interference pattern, electron migration establishes a periodic distribution of trapped space charge. By means of the electrooptic effect, the space charge induces a proportional modulation of the refractive index of the material. This constitutes the hologram. The amplitude of the hologram replicates the local visibility of the optical interference pattern, as required, but the phase of the hologram index pattern may be shifted laterally with respect to the phase of the interference pattern. As we shall see, this gives rise to some interesting effects which have important consequences for the development of an integrated optical data preprocessor based on holographic subtraction.

As an example of the spatial phase shift between the hologram index pattern and the optical interference pattern used to record the hologram, consider Fig. III-2. After exposure, positive (negative) space charge regions are seen to be in phase with intensity maxima (minima). However the space-charge electric fields which give rise to the hologram are, by Maxwell's equations, proportional to the gradient of the space charge. Consequently, they are in quadrature with the interference pattern, as is the hologram grating. The grating phase shift, designated ϕ_g, has the value $\pi/2$.

The situation illustrated in Fig. III-2 arises when the mean free path length of conduction electrons, L, is large compared to the period of the interference pattern, Λ. When $L \ll \Lambda$, retrapping occurs very close to the initial site. In this case, the space charge distribution pattern has maxima and minima where the interference pattern slope has maxima and minima. The space charge distribution is accordingly in quadrature with the interference pattern, but the hologram, being in quadrature with the space charge, has $\phi_g = 0$. In general $\phi_g = \tan^{-1} \kappa L$, where $\kappa = 2\pi/\Lambda$. The derivation of this result is found in Appendix A, where the theory of photorefractivity is presented in some detail.

The value of L may be externally controlled by varying the concentration of trapping sites. For a given density of Fe impurities, the concentration of Fe^{3+} traps can be regulated by subjecting the crystal to oxidizing or reducing

heat treatments, as described by Philips and Staebler.[4] Figure III-3 shows the expected relationship between L and Fe^{3+} concentration.

Photorefractive Sensitivity

The time required to form a hologram of specified diffraction efficiency depends upon the rate at which the space charge electric fields evolve. This depends, in turn, on the rate at which trapped electrons are excited into the conduction band and on the average length L that conduction electrons travel travel before they are retrapped. The dynamics are described by a continuity equation that has been presented by Kim et al.[3] and discussed in Appendix A.

If η is the hologram diffraction efficiency and

$$I = I_o(1 + m \cos ky) \qquad (III-1)$$

is the intensity pattern used to generate the hologram, we find, for η<<1,

$$\eta = (\Gamma I_o mt)^2 \qquad (III-2)$$

where t is the writing time and Γ is a measure of photorefractive sensitivity. The expression for Γ derived in the appendix in terms of parameters defined there is

$$\Gamma = (2\pi e/hc)(d/\cos\theta)(n_e^3 r_{33}/4\varepsilon)[\sigma qNL(1+\kappa^2 L^2)^{-1/2}]. \qquad (III-3)$$

The last bracket on the right hand side contains the relationship between Γ and photorefractive parameters, some of which can be controlled by the experimenter. σ is the absorption cross section of filled traps, q is the fraction of absorptions that result in electron migrations, and N is the concentration of filled traps. Since L is inversely proportional to the concentration of unfilled traps, it is apparent that Γ can be increased by thermal reducing treatments of the crystal which lead to a high concentration of Fe^{2+} at the expense of Fe^{3+}.[4] The maximum practical value of Γ suggested by Eq. III-3 is
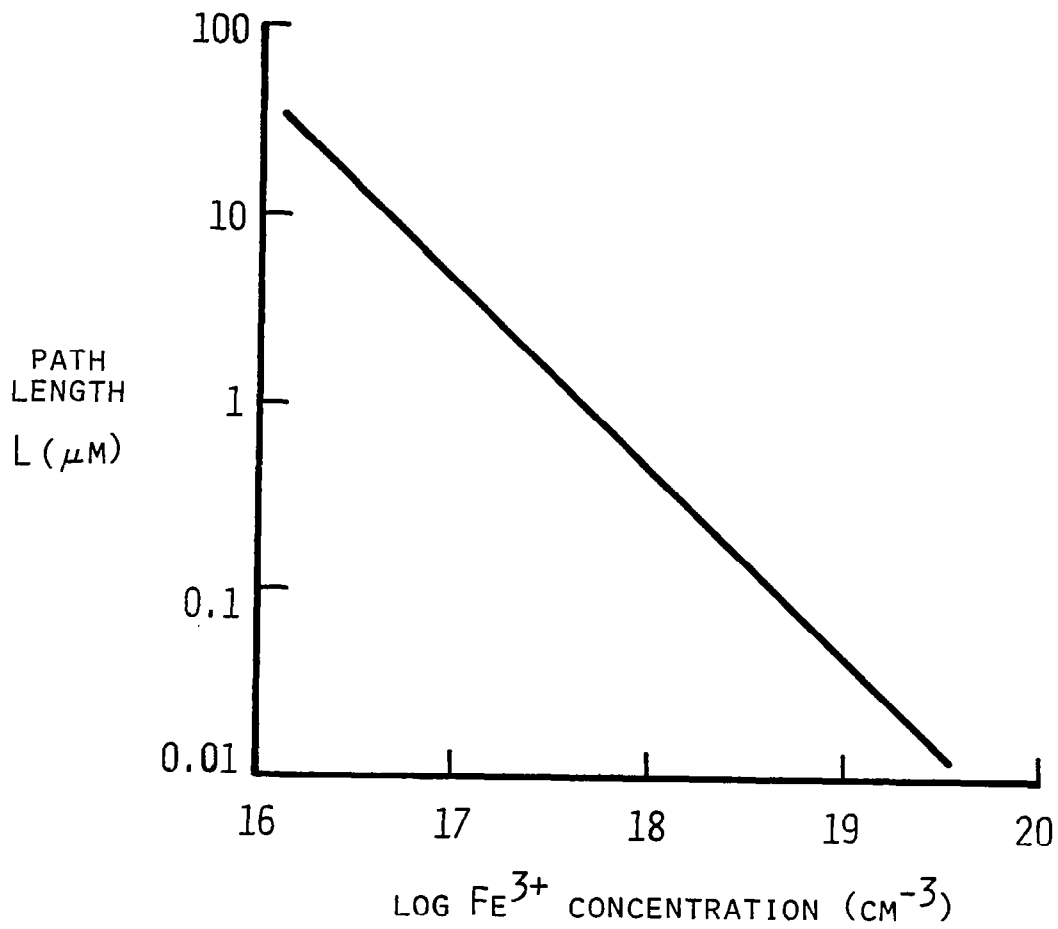
FIGURE III-3. DEPENDENCE OF PATH LENGTH ON Fe^{3+} TRAP CONCENTRATION.

$$\Gamma_{max} \simeq (e/hc)(n_e^3 r_{33}/4\epsilon)[\Lambda q] \qquad\qquad \text{(III-4)}$$

obtained by setting $N\sigma d/\cos\theta \simeq 1$ and $L \gg \Lambda$. Taking $\Lambda = 10^{-6}$ m, $q = 0.005$[8] and other parameters relevant to $LiNbO_3$, we find $\Gamma_{max} \simeq 0.04$ m^2/J. The implication of Eq. (2) is that the potential exists for wirting holograms having $\eta = 10\%$ in a waveguide format ($I_o = 10^5$ W/m^2) in times on the order of 0.1 msec.

Hologram Lifetime

Because of the small but finite conductivity of $LiNbO_3$, holograms formed by the photorefractive process will deday as the associated space charge distribution relaxes. The dielectric relaxation time ϵ/σ, where σ is now the conductivity, is reduced by heating the crystal, by illuminating it, and by increasing the impurity concentration, since all these processes serve to increase the conductivity. In crystals that are sufficiently impure as to show coloration, hologram lifetimes are typically measured in hours. In high purity crystals, the lifetimes can approach months. This is evidenced by the results of Fig. III-4, showing the decay of diffraction efficiency as a function of time for a hologram formed in "high purity" $LiNbO_3$, obtained from Crystal Technology.

If stable holograms are required for certain modes of operation of an integrated optical data preprocessor, several appraoches are feasible.

Thermal fixing of holograms by heating to about 110°C has been demonstrated by Staebler and Amodei.[9] At this temperature positive ionic impurities become sufficiently mobile to neutralize the electronic space charge distribution. As the hologram is subsequently read out, trapped electrons migrate out of the hologram region or, at least, are redistributed by the read beam in a quasi-uniform manner. The positive ions are not affected by light. They remain fixed in the periodic distribution of the original hologram and replicate its diffractive properties. There is some evidence[10] that the ions responsible for thermal fixing in $LiNbO_3$ are Si impurities.

An alternative method for obtaining long-lived holograms in $LiNbO_3$ is to exploit the large dielectric relaxation times of very pure crystals.
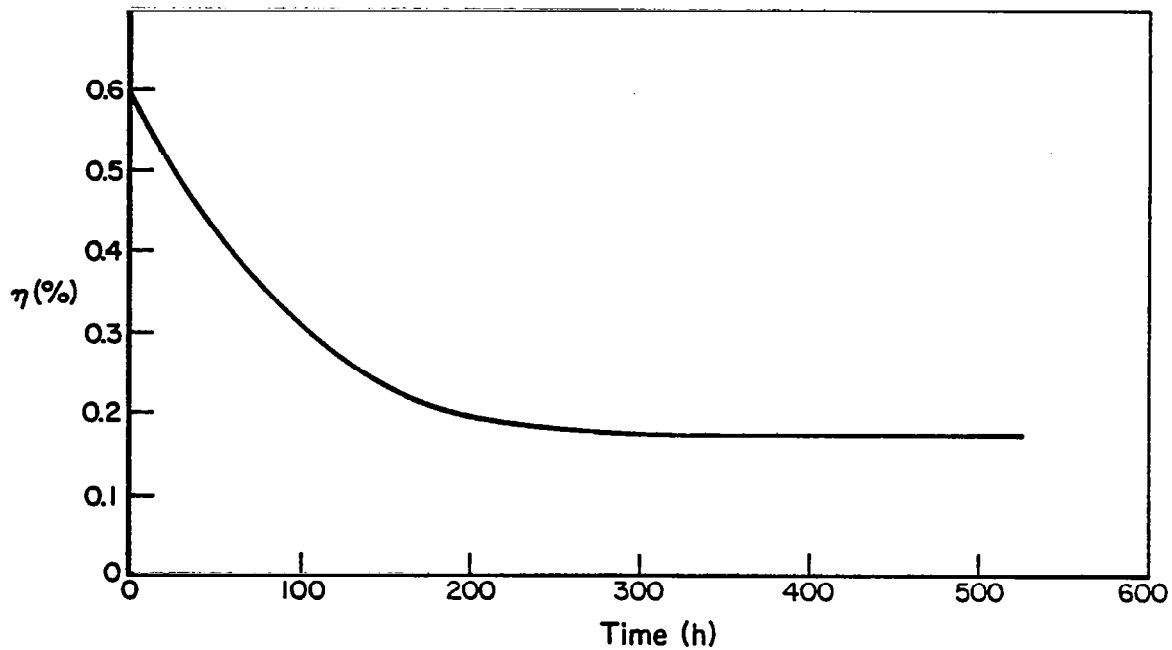
FIGURE III-4. DECAY OF DIFFRACTION EFFICIENCY AS A FUNCTION
OF TIME IN HIGH PURITY LiNbO$_3$.

An apparent problem is that such crystals have limited photorefractive sensitivity. However, a two photon photorefractive mechanism may be used to generate a hologram in pure crystals. The technique, first described by von der Linde[11] and later demonstrated in a waveguide during this program[12], relies on the fact that the simultaneous absorption of two photons can initiate an interband transition in $LiNbO_3$. Thus, an electron can be promoted directly from the valence band into the conduction band even in the absence of the impurity ions which are involved in single photon absorption.

The strength of the two-photon process is a quadratic function of peak laser power, approximately 10^6 W/cm^2 being required to produce detectable results. Both bulk and waveguide two-photon experiments have been performed during this program. The results are discussed in Appendix D.

The principle advantage of the two photon formation mechanism is that holograms so formed are impervious to light used to read the hologram. Moreover, lifetimes on the order of months appear feasible, owing to the poor conductivity of the pure crystals. A disadvantage at present is that the large optical powers required for the two-photon effect make space fabrication difficult. The technique may therefore find its best use with month-long missions in which the required hologram is formed on the ground in advance.

PHYSICS OF THE HOLOGRAM FORMATION PROCESS

During the formation of holograms in photorefractive $LiNbO_3$, the optical interference pattern used to write a hologram will in general be altered by diffraction in the hologram. This feedback causes a continual exchange of energy between the two writing beams which can significantly effect the operation of a processor based on holographic subtraction. The nature of the feedback is found to depend on the phase shift ϕ_g between the hologram index pattern and the optical interference pattern. In the case $\phi_g = 0$, slight exchange of energy between writing beams takes place as the hologram is recorded, and we can identify a normal mode of holographic subtraction in which a null is obtained by applying an external phase shift to one of the writing beams. In the case $\phi_g = \pi/2$, the evolving hologram continually diffracts energy out of one writing beam into the other. Eventually the null

signal associated with holographic subtraction is obtained without the need for an externally applied phase shift. This leads to what we refer to as the self-subtraction mode of operation.

To understand the two operating modes, consider the simple one-channel processor shown in Fig. III-5. The input data set contains only the single time-varying voltage V. We suppose that $V = V_o$ describes reference information, and we plan to use the one-channel processor to generate a signal whenever $V \neq V_o$.

In order to store reference information as a hologram, the voltage $V = V_o$ is applied to the signal-beam electrode in Fig. III-5. This imparts information to the beam by advancing it in phase by an amount GV_o, where G depends on the geometry of the electrodes and on the electrooptic properties of $LiNbO_3$. No voltage is applied to the electrode associated with the reference beam. This electrode will be used after the hologram is formed to achieve the null associated with holographic subtraction.

At the entrance to the hologram region, signal and reference beams intersect and form an interference pattern having the intensity

$$I = R_o^2 + S_o^2 + 2R_o S_o \cos(2ky\sin\theta - GV_o), \qquad \text{(III-5)}$$

where R_o and S_o are the reference and signal beam amplitudes, respectively, θ is one-half the angle of intersection, $k = 2\pi/\lambda$, where λ is the wavelength in the material, and y is the direction normal to the bisector of the beams. Henceforth we will use $K = 2k\sin\theta$.

The material response to the interference pattern is a hologram having a diffraction efficiency η and a phase shift ϕ_g. For short writing times η is small and we can ignore any change of the optical interference pattern caused by diffraction. As shown in Appendix B the signal and reference beam amplitudes at the output of the hologram region are then described by

$$S = (S_o + i\, \eta^{1/2} R_o\, e^{i\phi_g}) e^{iGV_o} \qquad \text{(III-6)}$$

$$R = R_o + i\, \eta^{1/2} S_o e^{-i\phi_g}$$
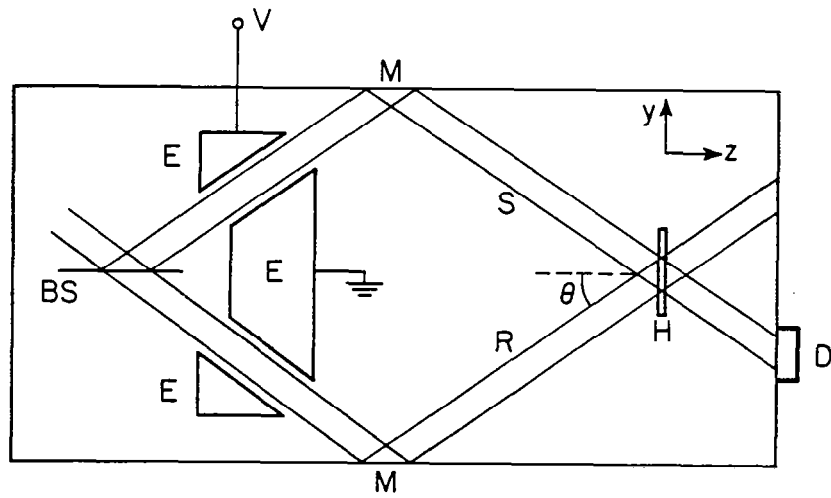
FIGURE III-5.   ONE-CHANNEL PREPROCESSOR.   S = SIGNAL
                BEAM; R = REFERENCE BEAM: H = HOLOGRAM;
                D = DETECTOR; M = MIRROR; E = ELECTRODE
                PAD; BS = BEAM SPLITTER; V = VOLTAGE
                DATA.

As a first iteration of the effect of hologram formation on the writing interference pattern, we can calculate $I = RR^* + SS^*$ using the results of Eq. III-6. For the interesting case $R_o >> S_o$, we find

$$I = I_o[1 + m \cos(ky - GV_o - \phi)] \tag{III-7}$$

$$I_o = R_o^2 + S_o^2$$

$$I_o m \simeq 2R_o[S_o^2 + \eta R_o^2 - 2\eta^{1/2}S_o R_o \sin\phi_g]^{1/2}$$

$$\phi \simeq \tan^{-1}[\eta^{1/2}R_o \cos\phi_g / (S_o - \eta^{1/2}R_o \sin\phi_g)]$$

OPERATING MODES OF HOLOGRAPHIC SUBTRACTION PROCESSORS

Self-Subtraction Mode of Hologram Formation

The variation of $I_o m$ and ϕ with η shows the extent to which diffraction in the hologram begins to effect the optical interference pattern of the writing beams, at least for small values of η. An exact analysis requires a numerical solution of Maxwell's Equations coupled to material equations describing photorefractivity in LiNbO$_3$.[13] A phenomenological approach can be appealed to in the special case $\phi_g = \pi/2$, since in that case $\phi = 0$ for all values of η. The indication is that the location of interference fringes remains constant, and only the modulation ratio m changes in time.

On the basis of work by Young et al.[14] we write

$$d(\eta^{1/2})/dt = \Gamma I_o m \tag{III-8}$$

during the initial stages of hologram formation, where Γ is the photorefractive sensitivity defined by Eq. 3, and

$$I_o m = 2R_o(S_o - \eta^{1/2}R_o) \tag{III-9}$$

for $\phi_g = \pi/2$. Upon integrating, we obtain

$$\eta^{1/2} = (S_o/R_o)(1 - e^{-2\Gamma R_o^2 t}) \quad . \tag{III-10}$$

By inserting this result into the first of Eqs. III-5 , it is seen that the condition of holographic subtraction, $SS^* = 0$, is achieved in the steady state. There is no need to terminate the writing process and fix the hologram at any finite time. As writing continues, the detected wavefront decreases exponentially, and the signal generated at the detector is proportional to

$$SS^* = S_o^2 \, e^{-4R_o^2 \Gamma t} \quad . \tag{III-11}$$

The hologram may be regarded as completed at a time beyond which this intensity is less than the system noise intensity. Since there will be no further change in the detected signal after this time, the hologram may be regarded as fixed, even though the material itself remains photosensitive. Since no external interruption of the writing process is needed to establish a condition of holographic subtraction, we term the process self-subtraction.

Normal Mode of Hologram Formation

For values of ϕ_g unequal to $\pi/2$, a solution for $\eta(t)$ based on Eq. (8) is possible but not readily justified in view of the fact that the phase of the optical interference pattern changes as η evolves. Ignoring this complication we find, for $\phi_g = 0$,

$$\eta^{1/2} = (S_o/R_o) \sinh 2\Gamma R_o^2 t. \tag{III-12}$$

From Eq. 6, the amplitude of the signal beam is

$$S = S_o(1 + ie^{i\phi} R \sinh 2\Gamma R_o^2 t)e^{iGV_o} \quad . \tag{III-13}$$

In this expression, ϕ_R is a phase shift applied to the reference beam using the lower electrode in Fig. III-5. To obtain a condition of holographic subtraction, the writing process must be terminated when $\sin 2\Gamma R_o^2 t = 1$. Also,

since $\phi_g = 0$ in this example, ϕ_R should be switched from 0 to $\pi/2$ to insure that the signal beam passing through the hologram and the reference beam generated by diffraction in the hologram will destructively interfere.

We refer to this situation as the normal mode of holographic subtraction because it applies to some extent in all cases except that for which $\phi_g = \pi/2$. It remains to be determined how close one must operate to the $\pi/2$ phase angle to effectively obtain self-subtraction performance. Since some amount of operator attention is required to implement holographic subtraction in the normal mode, it is seen to be the less "Smart" of the two operating modes.

Potential for Smart Operation of a Processor Based on Self-Subtraction

Figure III-6 shows the signal generated by a processor based on self subtraction as a function of time. The reference hologram is established in the interval $0 < t < t_1$, following which a voltage $V_1 = V_o$ is delivered to the processor. The initial signal generated is proportional to

$$SS^* = 2S_o^2[1 - \cos(GV_1 - GV_o)] \tag{III-14}$$

However, a new hologram will begin to form immediately, eventually, the generated signal will fall below the threshold level of the detection system (dashed line in Fig. III-6), and V_1 will be regarded as a new reference voltage.

The evolution of the new hologram is described phenomenologically by an equation similar to Eq. (8),

$$d(\eta_1^{1/2})/dt = \Gamma I_o m, \qquad t > t_1 \quad , \tag{III-15}$$

where η_1 is the diffraction efficiency of the new hologram, and $I_o m$ is taken to be the modulated component of intensity at the output of the hologram region,
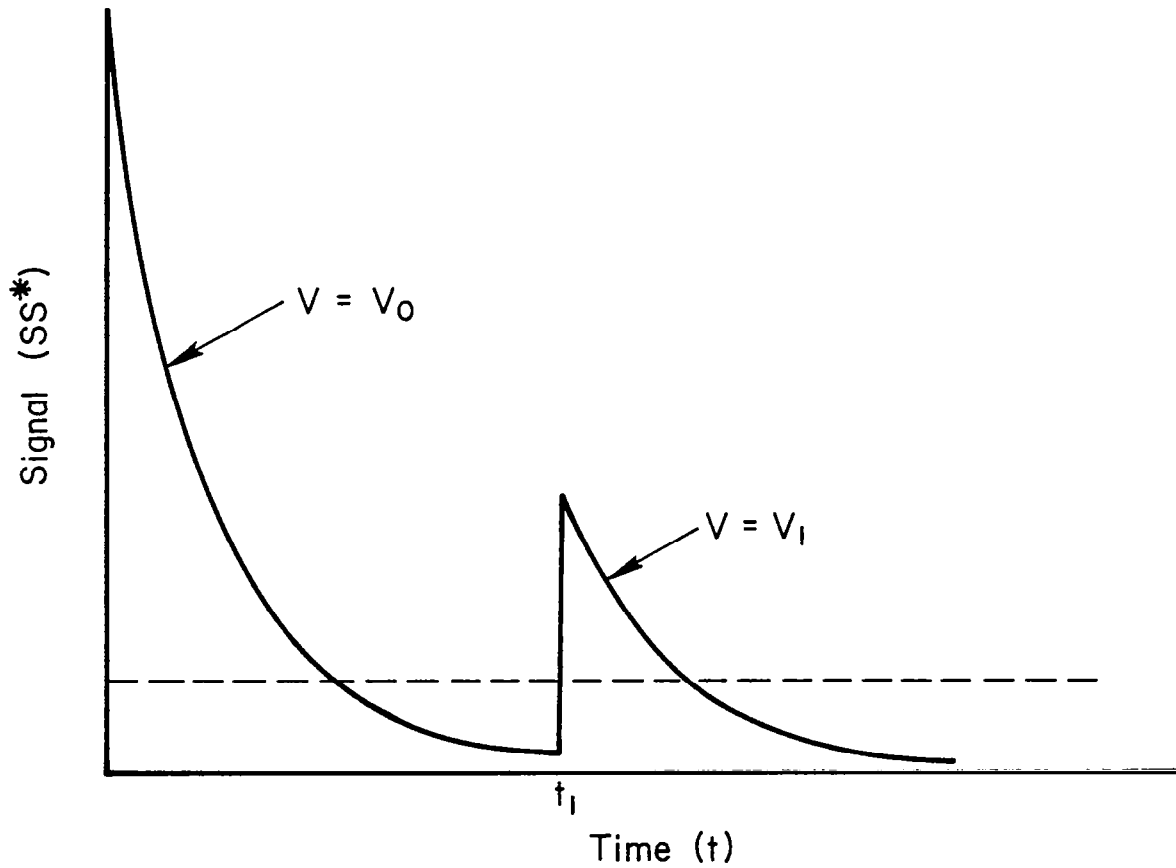
FIGURE III-6.   DETECTED SIGNAL FOR A HOLOGRAPHIC
SELF-SUBTRACTION PROCESSOR.

$$I_o m = 2R_o S_{eff} \qquad \text{(III-16)}$$

$$S_{eff} \simeq S_o [2-2\cos G(V_1-V_o)]^{1/2} - \eta_1^{1/2} R_o \quad .$$

Substituting these expressions into Eq. III-15 and integrating, we obtain

$$\eta_1^{1/2} = (S_o/R_o)[2-2\cos G(V_1-V_o)]^{1/2}(1-e^{-2\Gamma R_o^2(t-t_1)}) \qquad \text{(III-17)}$$

The signal generated at the detector is proportional to

$$SS^* = S_o^2(2-2\cos G(V_1-V_o)]e^{-4\Gamma R_o^2(t-t_k)} \quad . \qquad \text{(III-18)}$$

For large values of $4\Gamma R_o^2 (t-t_1)$, the processor is educated to the new input voltage V_1 and is ready to signal the arrival of voltages different from this value.

The processor is seen to be "Smart" in that it adapts to changes in reference information whenever this can be defined as information that persists longer than $(4\Gamma R_o^2)^{-1}$. All changes which last a time short compared to $(4\Gamma R_o^2)^{-1}$ do not erase existing reference information. To optimize the processor's utility, then, one must vary the photorefractive parameter Γ and the beam intensity R_o^2 to produce a time constant long in relation to the interesting events that one expect to observe, but short in relation to persistent events that may be classified as background.

An important capability of a "Smart" processor will be to discriminate against a slowly varying background signal. In the example under consideration, we suppose that V_1 is a fluctuating voltage that remains close to V_o, such that

$$(V_1-V_o) = v(t) \qquad \text{(III-19)}$$

For convenience, we restrict consideration to the case $Gv(t) \ll 1$. Solving Eqs. III-17 and III-18, we find the signal generated at the detector to be

$$SS^* \simeq S_o^2 G^2 [v(t) - \int_{-\infty}^{t} v(t')e^{-2\Gamma R_o^2(t-t')}2\Gamma R_o^2 \, dt']^2 \qquad \text{(III-20)}$$

35

when v(t) is slowly varying in comparison with the hologram time constant $T_h = (2\Gamma R_o^2)^{-1}$, the approximation

$$v(t') \simeq v(t) - \dot{v}(t)(t-t') \qquad \text{(III-21)}$$

may be used to evaluate the integral in Eq. 20, with the result

$$SS^* \simeq S_o^2 G^2 [\dot{v}(t) T_h]^2 \qquad \text{(III-22)}$$

In the normal mode of operation, the comparable signal is $SS^* \simeq S_o^2 G^2 v^2(t)$. If we identify $T_b \sim v/\dot{v}$ as the characteristic period of the background fluctuations, use of the self subtraction mode of operation is seen to reduce the background signal by a factor $(T_h/T_b)^2$. This makes the self-subtraction processor attractive from the point of view of detecting weak signals with sharp derivatives that might otherwise be lost.

IV. COMPONENT DESIGN AND FABRICATION

During the first phase of this program the major goal was to ascertain that the individual components required for the preprocessor could be fabricated and would function as required. In addition, attention was given to the ultimate necessity for component compatibility both in operation and in fabrication. This section summarizes the individual component studies, and establishes guidelines for source and detector selection.

Optical waveguides can be fabricated in a simple planar geometry in which a uniform guide exists on the entire substrate, or in a channelized geometry in which the light is confined to channels which are several micrometers wide and which provide path definition as well as lateral confinement. The planar and channel geometries can also be combined on the same substrate to form a hybrid structure.

A major choice to be made relates to the type of waveguide geometry to be employed with the preprocessor. Channel waveguides are a natural choice for the systems which make use of electrooptic phase perturbation of light, since these perturbations can be accomplished at low voltages and power levels when the light is transversely confined to narrow channels. On the other hand, the process of holographic subtraction works best when the reference beam is transversely unconfined, as in a guided plane wave.   These considerations appear to dictate a hybrid waveguide for an operational preprocessor. However, as the goal of the present program was to demonstrate initial feasibility for a 3-channel device, a simple planar structure was chosen. Plans for the hybrid device will be discussed in Section VII.

SUBSTRATE AND WAVEGUIDE

The choice of substrate and waveguide for the preprocessor is largely dictated by the device concept. The most restrictive requirement is that it must be possible to write high-quality waveguide holograms which have the ability to be changed when necessary. In addition, the requirements for a planar shifter and real time input of data in the form of voltages strongly suggest an electrooptic material. These two requirements are uniquely

satisfied by $LiNbO_3$, although $LiTaO_3$, either suitably doped or with a photo-sensitive overlay in the hologram region, might also be considered. Other advantages of $LiNbO_3$ are its ability to support high quality waveguides, its commercial availability, and the fact that $LiNbO_3$ waveguides have been well characterized in our laboratory and elsewhere. For these reasons, this material has been emphasized throughout this program.

Both Ti-infused and Li-effused $LiNbO_3$ waveguides have been studied. Each has advantages with regard to the operation of certain components of the preprocessor. For example, Ti-infused waveguides are useful for channelized components like the phase-shifter and for contoured components like the geodesic lens; Li-effused waveguides are deeper and hence more appropriate for use with edge-polished total-internal-reflection mirrors. A clearer choice as to the best waveguide to use is afforded by systems considerations. The preprocessor waveguide should exhibit low scattering and low photorefractivity except in the hologram region, where the photorefractivity should be controllable.

Waveguide Scattering

Two sources of scattering in optical waveguides are surface roughness and volume inhomogeneity. An approximate treatment of the former has been on the Rayleigh criteria for scattering from rough surfaces.[15] The attenuation is approximately proportional to $(\Delta n/n)^{3/2}/W$, where Δn is the waveguide index discontinuity, n is the substrate index of refraction, and W is the waveguide thickness. For Li-effused waveguides, W is an order of magnitude larger than for Ti-infused waveguides. Consequently, the scattering from surface roughness should be a factor of about 300 less in the effused case.

In addition to their lower surface scattering, it also appears likely that effusion is preferable to infusion from the point of view of limiting volume-scattering inhomogeneities. In connection with an Air Force sponsored program "Optical Waveguide Scattering Reduction",[16] in progress at Battelle, we have observed that $LiNbO_3$:Ti waveguides have a mottled surface appearance that may be readily observed using Nomarski microscopy. Burns, et al, have also found that part of the diffused metal goes toward the formation of

Li-Ti-O compounds, such as Li-Ti-O, within the top 0.3 µm of the waveguide surface. Neither of these phenomena are present in Li-effused waveguides. Both can contribute to scattering and reduce the signal-to-noise performance of an integrated optical data preprocessor.

In spite of the problems inherent in the Ti-infused waveguides, we have found that by polishing their surface scattering can be reduced to levels which suggest potentially very high signal-to-noise perfromance. For example, we find that when light coupled out from a $LiNbO_3$:Ti waveguide is focused on a narrow slit and detected, the signal for light scattered at 1° off axis is reduced by 40 dB from the on-axis component. As the slit width increases to a value greater than the width of the focal spot, the signal generated by unscattered light is limited by the total power in the beam. However, the noise signal from scattered light increases proportionately with the slit width. Our results suggest that for an integrated optical data preprocessor having a 20-µm-wide detector at the focal plane of a 10-mm-focal-length waveguide lens, $P_{noise} \simeq 10^{-3} P_{total}$. If this optical noise power is not added to or diminished by the process of holographic subtraction, we find that the total power incident on the detector may be doubled from its noise value by phase changes in the signal beam as small as 1°.

There are several reasons why performance levels both better and worse than indicated by this calculation could be achieved. First, methods for reducing waveguide scattering are expected as a result of the Air Force program. Second, the angular selectivity of the hologram could reduce the amount of scattered light diffracted toward the detector. Both of these possibilities argue in favor of improved sensitivity. On the other hand, scattering associated with optical damage could degrade performance. The mechanism for this is as follows: Light scattered by surface roughness and material defects interferes with unscattered light and forms a hologram. By the phenomenon of self subtraction, the hologram takes energy from the unscattered beam and diffracts it in the direction of the scattered beam. The net result is that the effects of weak scattering centers are amplified by photorefractivity. The harmful effects of this could be minimized if necessary by using a waveguide that is sensitized only in that region reserved for the information hologram.

Waveguide Photorefractivity

Even in a scatter-free waveguide, photorefractivity in regions not devoted to the hologram can decrease the signal-to-noise performance of the processor. The photorefractive mechanism causes space charge to accumulate along the edges of the signal and reference beams. The electric fields associated with the space charge produce a distributed negative lens by means of the electrooptic effect. Although the focal length of the lens may be on the order of meters, the angular selectivity of the information hologram is so great that the slightly diverging rays will fall outside its acceptance angle and frustrate holographic subtraction, thus constituting a source of optical noise.

It is anticipated that the beam profiles associated with the lens effect will eventually reach a steady state. If this occurs well before the evolution of the hologram is completed, the fringes of the hologram should adjust to the altered beam profiles, and the resulting optical noise will be minimal. If the steady-state beam profiles are not achieved until well after the hologram has formed, the optical noise associated with photorefractive lensing will have its largest value.

In order to estimate the magnitude of the effect in this worst-case limit, we calculate the photorefractively induced curvature radius of the reference beam wavefront as it impinges on a uniform hologram. At each point along the hologram we calculate the deviation from the Bragg angle associated with the curvature of the wavefront. Using Kogelnik's theory for diffraction by thick uniform gratings,[17] we relate the deviation from the Bragg angle to variations in the diffracted wavefront relative to the wavefront required for perfect holographic subtraction. The result of the calculation is the formula

$$P_{noise} \simeq P_{signal} \, (4/45)(kL\Delta n)^4 \qquad\qquad (IV-1)$$

where P_{signal} is the power in the signal beam passing through the hologram, P_{noise} is the noise power associated with the photorefractive lens effect, k is the magnitude of the wave vector, L is the physical path length of the reference beam, and Δn is the photorefractively induced index change.

If the waveguide is sufficiently photorefractive that holograms of 10% diffraction efficiency can be formed in it, $\Delta n \simeq 10^{-5}$. If in addition, $k \simeq 22 \ \mu m^{-1}$ (for $\lambda = 0.63 \ \mu m$), then L need be only 0.5 cm for intolerably high noise levels to result. In practice, a reference-beam path-length L of several centimeters is required.

The large noise levels calculated from Eq. IV-1 do not appear consistent with the performance we have achieved from the experimental processors fabricated to date. A likely reason is that the photorefractive lens forms at the same time that the hologram forms. Consequently, the hologram is able to compensate for wavefront variations in the writing beams as they occur.

In spite of this fortunate situation, optimum performance is sure to result when photorefractive snesitivity is limited to only that region of the waveguide required for the information hologram. Since this can be readily accomplished by local doping with impurities or by local illumination with a sensitizing beam (two-photon photorefractive effect), we investigated the minimum amount of background photorefractivity that could be expected in outdiffused $LiNbO_3$ waveguides in commercially available nominally pure and highly pure crystals. Our experiments made use of a sensitive interferometric system, shown in Fig. IV-1, to measure any photorefractively induced index changes and to monitor these changes over a long period of time.

The operation of the interferometer is described as follows: A silver-coated optical flat is used to generate two parallel beams coupled into the waveguide. The more intense beam optically damages the waveguide, thereby inducing a refractive index change which is to be measured. If this change is Δn and L is the path length through the crystal, the resulting phase change is $2\pi \Delta n L / \lambda$ which shows up as a shift of the interference pattern at the detector. Owing to intensity variations of the laser, it was found best to measure the photorefractive phase change by electrooptically inducing an equal phase change on the low intensity reference beam so that the same fringe was centered on the detector whenever a reading was taken. The sensitivity of this method was about 0.1 fringe, which translates to $\Delta n = 10^{-6}$. It is estimated that the sensitivity could be improved by a two orders-of-magnitude if lock-in techniques were employed.
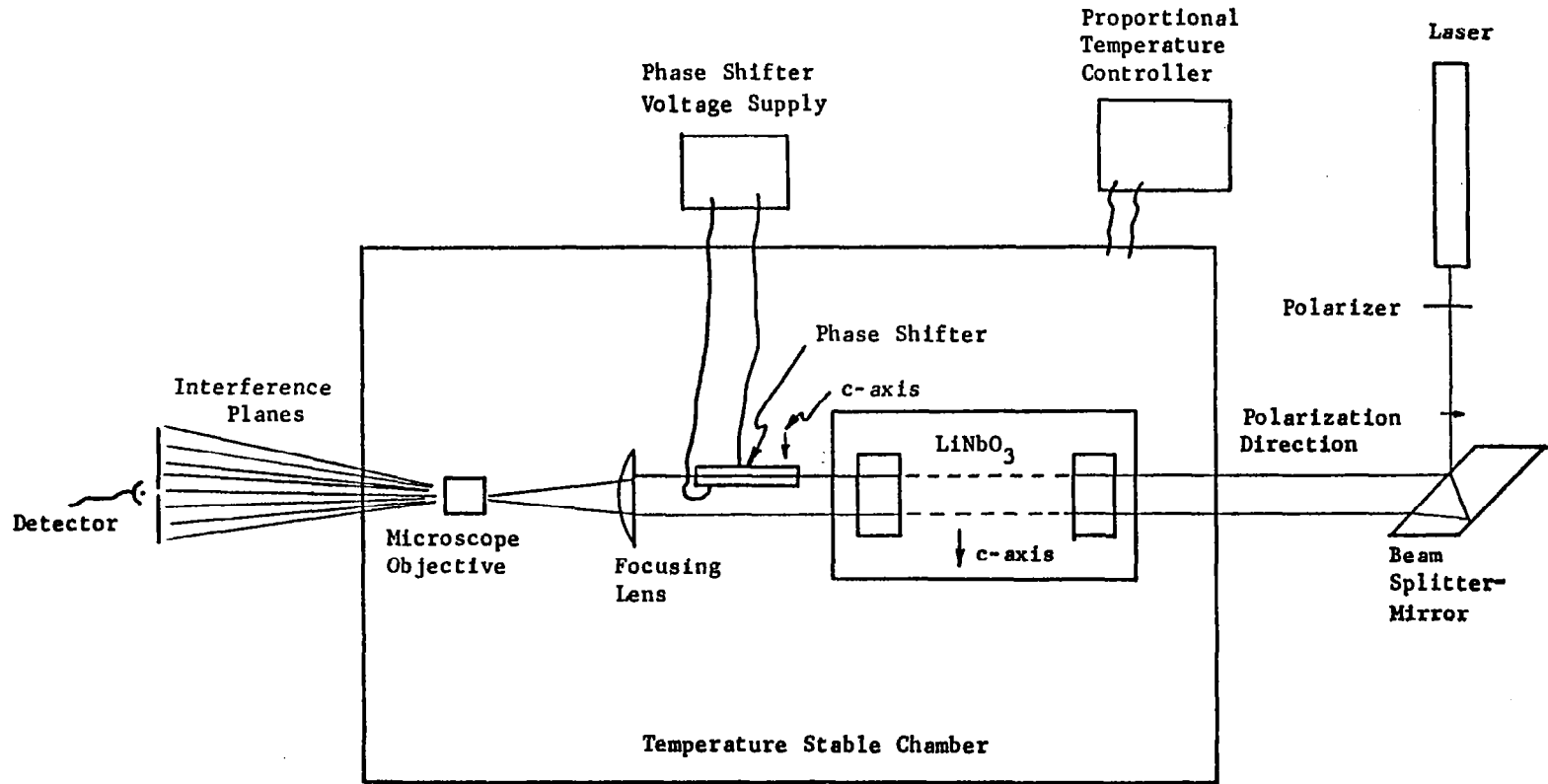
FIGURE IV-1. MODIFIED INTERFEROMETER FOR WAVEGUIDE STABILITY MEASUREMENTS.

Despite temperature stabilization of the waveguide region, temperature fluctuations experienced by the laser and beam-forming optics resulted in spurious data in all experiments that were run overnight. The best reasonable interpretation of the data is that the upper limit of Δn generated by optical damage was about 3×10^{-5}. In some experiments it appeared that this was a saturation value; in others, it was not clear. One possible source of complication was the effect of waveguide scattering enhanced, as described earlier, by the optical damage mechanism. This scattering can conceivably generate fluctuations in the beam that stabilize only slowly, or not at all, and make data interpretation difficult.

In view of the problems of experimentation and interpretation encountered during this investigation, it was decided to postpone further work on waveguide photorefractivity until its importance could be fully ascertained. Recent success in fabricating a number of three-channel processors suggests that optical damage that occurs outside the hologram region need not be of concern, at least in first-generation systems.

Substrate Orientation

In writing a hologram in $LiNbO_3$, it is necessary to have a significant component of the grating vector parallel to the cyrstalline c-axis. This implies that only x- or y- cut crystals can be used, and that the long axis of the device must be perpendicular to the c-axis. Most of our experimental work has been performed on y-cut slabs because of the ready availability of this material. However, because of the nature of the $LiNbO_3$ growth process, long x-cut slabs are easier to acquire than long y-cut slabs. We have therefore verified experimentally that the waveguide formation process and the hologram writing process are similar in both x- and y- cut materials. The application of results of this program will therefore not be limited by the availability of properly oriented crystals.

INPUT COUPLERS

The major disadvantage of using $LiNbO_3$ for any integrated optical data processing system is that it is impossible to construct a system with a truly integrated source or detector. It is therefore necessary to incorporate

some means of coupling light into and out of the waveguide. The most common means of coupling light into a waveguide is by means of the prism coupler (Fig. IV-2).

The prism coupler consists of a high-index-of-refraction material which is separated from the surface of the waveguide by a small airgap, typically less than the wavelength of the light being coupled. A laser beam incident upon the coupling face at an angle $\theta > \theta_c$, where θ_c is the critical angle, will undergo total internal reflection. There is an evanescent wave associated with this total internal reflection which has a transverse velocity $\frac{c}{n_p} \sin\theta$, where n_p is the prism index. When this velocity equals a given mode velocity $\frac{c}{n_m}$, there is a strong resonant transfer of energy into that specific mode. The strength of the transfer is a very sensitive function of g, the gap dimension and it is strongly modulated by prism or waveguide surface roughness or small dust particles in the gap. The prism coupler also requires a mechanical clamping device to hold it in contact with the waveguide.

In spite of the disadvantages of the prism coupler, it is the most convenient method of coupling into $LiNbO_3$ waveguides and has been used throughout the experimental portion of this program. In later versions of the preprocessor, it is anticipated that simple butt coupling will be utilized. This will certainly be the choice of output coupler since butt coupling of PIN diode detectors to $LiNbO_3$ waveguides is a well-accepted, efficient technique.

SURFACE-GRATING BEAM SPLITTERS

In order to conveniently provide two coherent guided beams for the processor, a simple beam splitter is needed. These can be fabricated in a number of ways, e.g., a holographic grating formed in the waveguide, a surface grating, or an etched groove. One of the more promising such devices is a grating formed in a photoresist layer on the surface of the waveguide. The theory of operation of these gratings and experimental work on realizing them are described in this section.

The basic mechanism operating in photoresist gratings is perturbation of the effective refractive index of the guided wave by modulation of the index of the waveguide superstrate. When properly exposed and developed, the
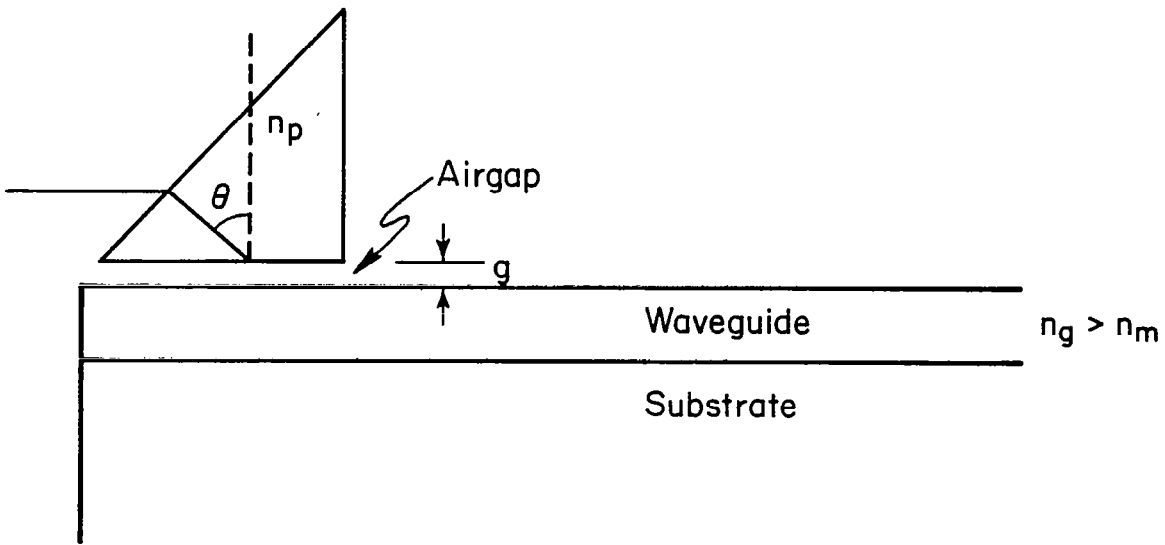
FIGURE IV-2. PRISM COUPLER GEOMETRY.

photoresist grating has the form of a periodic array of stripes of residual photoresist material lying on the waveguide surface, with bare waveguide between the stripes. Thus, the superstrate is alternately photoresist or air. Since the effective waveguide index depends on the superstrate index, the waveguide index is thereby modulated. This modulation comprises a phase grating as seen by the guided wave. The magnitude of this index modulation is, of necessity, less than the difference of the substrate and surface indices which, in LiNbO$_3$ outdiffused guides is only a few parts per thousand. The modulation amplitude may be enhanced by overcoating the grating with another material, e.g., a metal. Then the superstrate is alternately photoresist and other material. An overcoating may also serve to protect the grating.

The analysis of surface gratings proceeds from the WKB waveguide equation[18] using normalized parameters:

$$V \int_0^{x_t} \sqrt{f(x) - b} \; dx = \Phi + (m + 1/4) \, \pi \; ; \; f(x_t) = b \qquad \text{(IV-2)}$$

with

$$V = k_o D \sqrt{n_g^2 - n_s^2} \qquad \text{(IV-3)}$$

$$b = (N^2 - n_s^2)/(n_g^2 - n_s^2) \qquad \text{(IV-4)}$$

where n_g is the index of the waveguide material at the crystal surface, n_s is its value far from the surface, $k_o = 2\pi/\lambda_o$, λ_o is the free-space wavelength of the light, and D is the diffusion depth. $f(x)$ is a smooth function satisfying $f(0) = 1$, $f(\infty) = 0$; it determines the index profile of the waveguide:

$$n^2(y) = n_s^2 + (n_g^2 - n_s^2) \, f(y/D); \; x = y/D \; . \qquad \text{(IV-5)}$$

N is the effective index of the guided mode, m is the mode number (usually m = 0, since we are primarily concerned with single-mode guides). The phase shift on reflection, Φ, is determined from

$$\Phi = \tan^{-1} \sqrt{\frac{a + b}{1 - b}} \qquad\qquad \text{(IV-6)}$$

where the asymmetry parameter, a, is defined by

$$a = \frac{n_s^2 - n_a^2}{n_g^2 - n_s^2} \qquad\qquad \text{(IV-7)}$$

and n_a is the superstrate index.

Equation IV-2 determines the allowed values of b, given a, V and a function f(x). Since n_a occurs explicitly only in a, the modulation of b and, through b, of N, can be regarded as occurring through perturbation of the parameter a. For outdiffused guides in $LiNbO_3$, a will be around 600 or more, whereas b always lies between 0 and 1. Thus, Φ can be well-approximated by

$$\Phi \simeq \tan^{-1} \sqrt{\frac{a}{1 - b}} \approx \frac{\pi}{2} - \mu \sqrt{1 - b} \qquad\qquad \text{(IV-8)}$$

where we have defined a reciprocal assymmetry measure

$$\mu \equiv 1/\sqrt{a} \qquad\qquad \text{(IV-9)}$$

Then, equation (1) becomes

$$V\int_0^{X_t} \sqrt{f(x) - b} \; dx = (m + 3/4) \pi - \mu \sqrt{1 - b} \quad . \qquad\qquad \text{(IV-10)}$$

Define

$$I_j = \int_0^{X_t} (f(x) - b)^{j/2} \quad \text{with } j = \pm 1 \quad . \qquad\qquad \text{(IV-11)}$$

We want to find the change in b occasioned by a small change in μ. If we set

$$\mu = \mu_o + \mu_1$$

$$b = b_o + \mu_1 b_1 \tag{IV-12}$$

we can expand Eq. IV-10 in powers of μ_1 to obtain

$$V \cdot I_1(b_o) = (n + 3/4)\pi - \mu_o \sqrt{1 - b_o} \tag{IV-13}$$

and

$$b_1 = 2\sqrt{1 - b_o} / [V \cdot I_{-1}(b_o) + (1-b_o)^{-1/2}] \quad . \tag{IV-14}$$

Eq. IV-13 can be used to eliminate V in favor of I_1 and the other parameters. Now, for our LiNbO$_3$ waveguides, Eq. IV-4 gives approximately

$$b \simeq \frac{N - n_s}{n_g - n_s} \equiv \frac{N - n_s}{\Delta n} \tag{IV-15}$$

because $N \approx n_s \approx n_g$. Hence, the change in N corresponding to b_1 is

$$\delta N = \mu_1 b_1 \Delta n \qquad (\mu_1 = \mu - \mu_o) \quad . \tag{IV-16}$$

While these equations may be used to estimate the efficiency of surface gratings, using numerical methods to evaluate $I_{\pm 1}$, the uncertainty in the waveguide parameters may make the estimates relatively inaccurate. However, they can be used to estimate the effect of an overcoating on a grating. The efficiency of a Bragg grating may be written as

$$\eta = \sin^2 \nu \qquad \nu = c \, \delta N \tag{IV-17}$$

where δN is the depth of modulation of the grating and c is a constant. If we construct a grating of efficiency η_o using photoresist and use the photoresist-covered waveguide as the reference condition in our expansion (μ_o, b_o), we have for the overcoated grating

$$\frac{\delta N'}{\delta N} = (\mu' - \mu_o)/(\mu_{air} - \mu_o) \tag{IV-18}$$

48

where primed quantities go with the coated grating. Thus,

$$\eta' = \sin^2 [c\delta N \cdot \frac{\mu' - \mu_o}{\mu_{air} - \mu_o}] \qquad (IV-19)$$

It should be noted that, while the individual μ's each depend on $n_g^2 - n_s^2$, a small number, the ratio in Eq. (IV-19) does not. Measurement of η_o determines $c\delta N$, and the effect of various overcoatings can then be studied using (18).

A series of experiments on fabrication of photoresist gratings and the effect of Al films evaporated onto them have been carried out. The characteristics of these gratings and their beam-splitting efficiency, expressed as

$$\eta = \frac{\text{diffracted intensity}}{\text{(diffracted + undiffracted) intensity}} \qquad (IV-20)$$

are tabulated in Tables IV-1 and IV-2. The initial gratings were fabricated by holographic exposure of the photoresist through a circular mask. After the initial success of these gratings, subsequent beam splitters (#3-10) were fabricated with a rectangular mask. As the period of the gratings was decreased, the efficiency tended to decrease, and it was found that higher exposures were necessary to obtain good development of the photoresist. This is seen in gratings pairs no. 7-8 and 9-10 where reduced exposure drastically reduced the efficiency.

On gratings 5-10, two efficiencies are indicated. It was observed that the efficiency of the gratings was asymmetric, i.e., the efficiencies for the two first-order Bragg incidence angles ($\pm \theta_B$) were not the same. Where both efficiencies are given. It is not understood at this time why there should be this asymmetry. Microscopic examination of the gratings reveals no discernable blazing.

These results are very encouraging. Metallizing the gratings definitely improves their efficiency, and grating no. 9 could be used in a device since both its efficiency and its Bragg angle are sufficiently large. Calculations using Eq. (IV-19), carried out on some of the data, indicate that the real part of the dielectric constant of the evaporated Al films is in the

TABLE IV-1. GRATING PARAMETERS

No.	Exposure (J/cm^2)	Beam Splitter	Beam Intensity Ratio	Period (μm)	Length (mm)
1	8.0	Cube	1:1	1.7	Round, 7 mm
2	8.0	Cube	1:1	0.83	Round, 7 mm
3	8.5	Cube	1:1	1.25	3.73
4	8.5	Cube	1:1	1.25	3.73
5	10.5	Wedge	2:1	1.17	3.73
6	10.5	Wedge	2:1	1.17	3.73
7	11.2	Wedge	2:1	1.17	3.73
8	8.4	Wedge	2:1	1.17	3.73
9	11.2	Wedge	1:1	0.60	3.73
10	8.4	Wedge	1:1	0.60	3.73

TABLE IV-2. GRATING MEASUREMENTS

No.	Period (μm)	Calc. Bragg Angle (°)	Bare Photoresist η(%)	Aluminized η(%)	Remarks
1	1.7	5.1	24.0	47.0	Initial test-round gratings
2	0.83	10.4	3-4	44	Initial test-round gratings
3	1.25	6.9	1.85	11.9	
4	1.25	6.9	1.81	7.9	Degraded by multiple reflections
5	1.17	7.4	1.2	12.8/17.1	
6	1.17	7.4	1.3	5.5/15.4	
7	1.17	7.4	2.8/5.2	12/11.4	
8	1.17	7.4	.53/1.12	7.8/11.9	Reduced exposure
9	0.60	14.5	1.7	10.6/27	
10	0.60	14.5	.94	1.44/1.23	Reduced exposure

is in the range -5 to -15, smaller in magnitude than for bulk Al (-40). Improvements in fabrication technique to achieve more reproducible results are needed, but it appears that aluminized photoresist grating beam splitters will be acceptable for the preprocessor.

MIRRORS

The integrated optical data preprocessor inherently requires a folded optical path. Other integrated optical systems do not require folded paths but may use them to conserve space. For this reason, we perceive the realization of mirrors based on the total internal reflection (TIR) of waveguided light as one of the principle component developments of the program. The major alternative to a TIR mirror is a surface grating, of the type which has been described for use as a beam splitter. The less-than-unity reflectivity, angular selectivity, and size of these gratings make them acceptable for use as beam splitters, but less than desirable for use as mirrors. In comparison, TIR mirrors have unit reflectivity, no angular selectivity beyond the critical angle, and occupy no surface area. They do, of course, suffer the disadvantage of having to be located at the edge of the integrated optic chip.

TIR Mirror Fabrication

In order to fabricate a TIR mirror on the edge of a $LiNbO_3$ substrate, two slabs are clamped together between steel gauge blocks, as shown in Fig. IV-3. The slabs should be flat and parallel, and the gauge blocks should have at least two faces that intersect at an angle of precisely 90°. We employed $LiNbO_3$ crystals obtained from Crystal Technology that had flatness specified to within λ/4, and parallelism specified to within 30 seconds of arc. Gauge blocks, fabricated in the Battelle Machine Shop, had right-angle faces to within a precision of 1 minute of arc. These tolerances were found to be satisfactory for the production of TIR mirrors in Li-effused and titanium-infused waveguides.

Polishing jig

Reference blocks
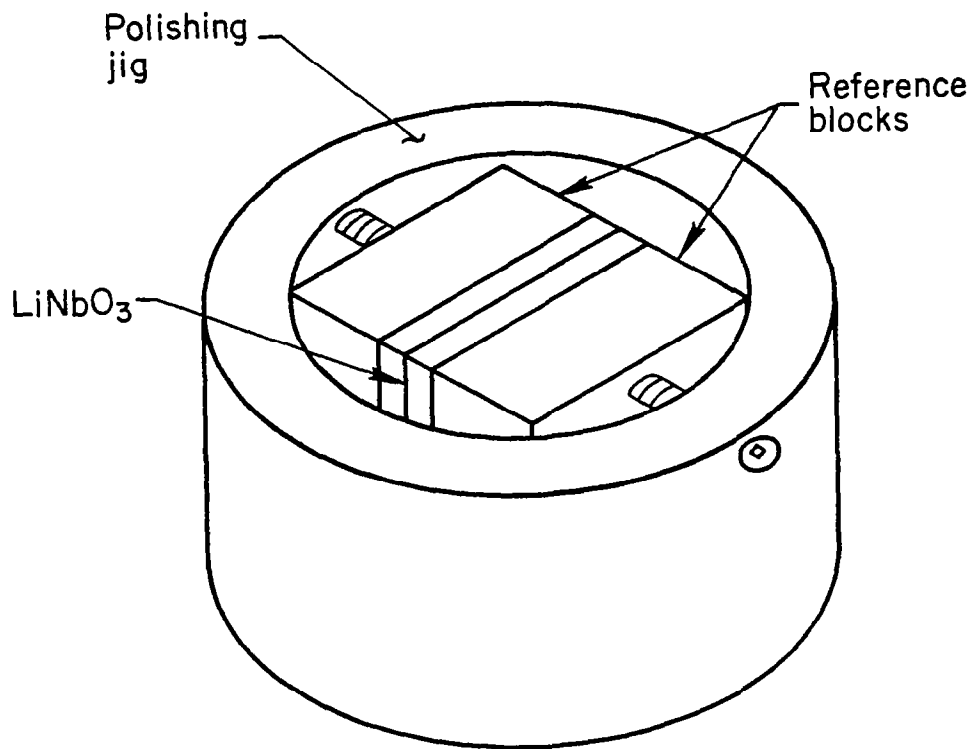
$LiNbO_3$

FIGURE IV-3.   JIG USED IN THE FABRICATION OF TIR
              WAVEGUIDE MIRRORS IN $LiNbO_3$.

The mirrors were fabricated by bringing the exposed slab edges to a state of high polish. This was done by the successive use of increasingly smaller abrasives. The final polish is best accomplished by the use of Syton HT-30, which is a colloidal suspension of silicon. The typical particle size is 0.03 μm. We have found that a soft lap like Buehler Microcloth on a brass base enhances the polishing characteristics of Syton. However glass, wax or metal laps should probably be used with larger-size abrasives to preserve flatness. Table IV-3 shows a combination of laps and abrasives that we have successfully employed in fabricating TIR mirrors. During the intermediate polishing, the 1-1/2-μm diamond and the Linde C on wax are used interchangeably. The diamond paste provides a fast polish while the wax wheel preserves flatness.

TABLE IV-3. TIR MIRROR FABRICATION

Abrasive	Lap	Finish
#240 paper	Grinding wheel	Rough grind
#320 Alundum	Glass plate	Rough grind
#600 Alundum	Glass plate	Fine grind
Dia. 1-1/2 μm	Forstman's Wool on brass wheel	Intermediate polish
Linde C (1 μm)	Wax wheel	Intermediate polish
Syton (0.030 μm)	Microcloth on vibrating polisher	Final polish

TIR Mirror Characterization

The TIR mirrors have been characterized by examining the surface finish and by measuring the angle of intersection between the waveguide surface and the mirror surface. The two means that we have employed to

characterize TIR mirror finishes are interference microscopy and electron microscopy. Figure IV-4 shows an interference micrograph of the polished area of two clamped samples. In the photograph, the fringes make an oblique angle to the interface. In this orientation, it is possible to see a slight perturbation of the fringes at the interface. When the fringes are normal to the interface no fringe shift is discernable, indicating that any rounding of the polished edge is small compared to the wavelength of visible light. Figure IV-5 is an electron micrograph at 5000X of the mirror edge of one of the two samples. Once again the edge rounding is observed to be minimal (< 0.1 μm radius of curvature) and there is no evidence of chipping. The debris evident on the surface of the waveguide is left over from the photolithographic fabrication of the electrode structures on the waveguide surface. The angular measurements made using a prism spectrometer table indicated the waveguide surface and the mirror surface intersected at an angle of 90° 1' $\pm$ 30".

TIR Mirror Performance

The two LiNbO$_3$ samples associated with the interferogram of Fig. IV-4 were heat-treated to form waveguides. In one of the samples an out-diffused guide was fabricated while a Titanium-infused guide was formed in the second sample. The reflectance was determined by measuring the intensity of the guided beam coupled out by a rutile prism before and after reflection from the TIR mirror. The outdiffused guide supported two modes at λ = 632.8 nm with the polarization parallel to the sample plane. The reflectance was found to be approximately 90% for both modes. The Ti-infused guide supported a guided mode with polarization either perpendicualr or parallel to the plane of the sample. The measured reflectance was approximately 90% for both modes. The angle of incidence of the guided beam with the TIR mirror was 70° for these measurements. Based on these experiments, we believe that TIR mirrors suitable for use with a high-performance data preprocessor can be fabricated simply and reproducibly.
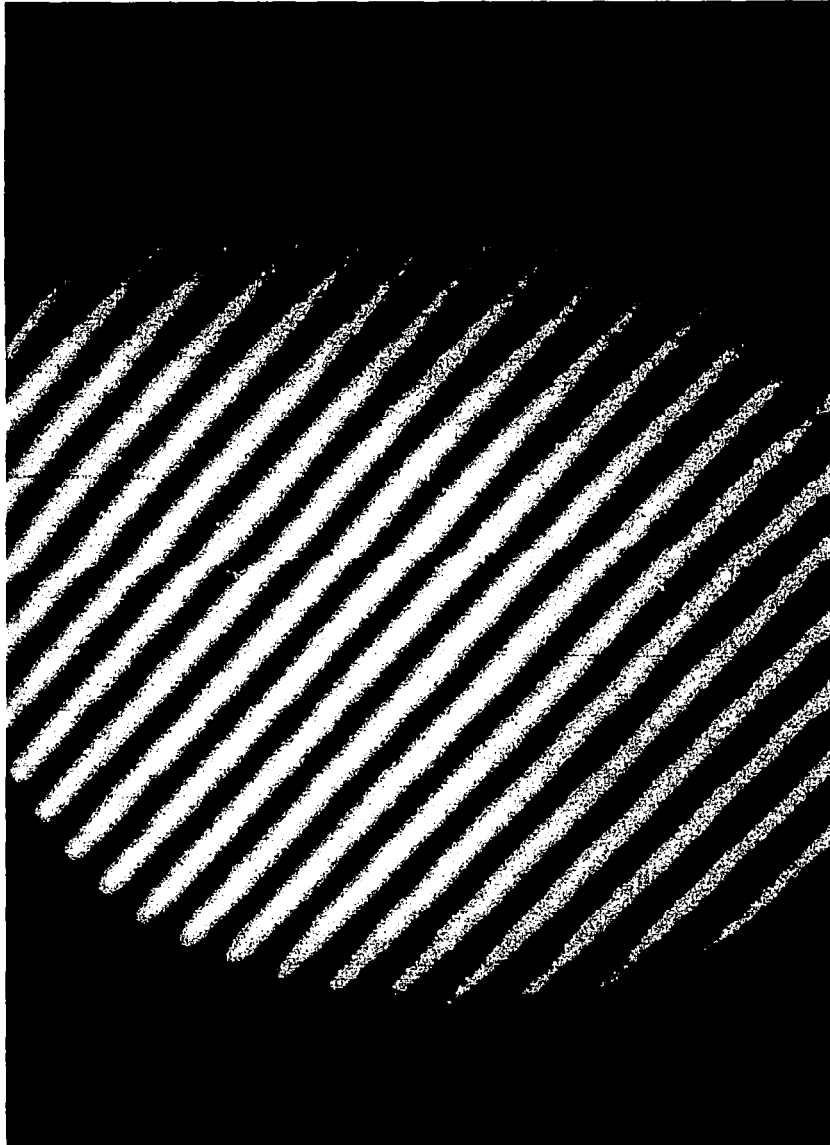
FIGURE IV-4. INTERFERENCE MICROGRAPH OF THE POLISHED SURFACES
OF TWO FACING LiNbO$_3$ SLABS. THE VERY SMALL FRINGE
SHIFT AT THE INTERFACE INDICATES EDGE ROUGHNESS
$<< \lambda$.
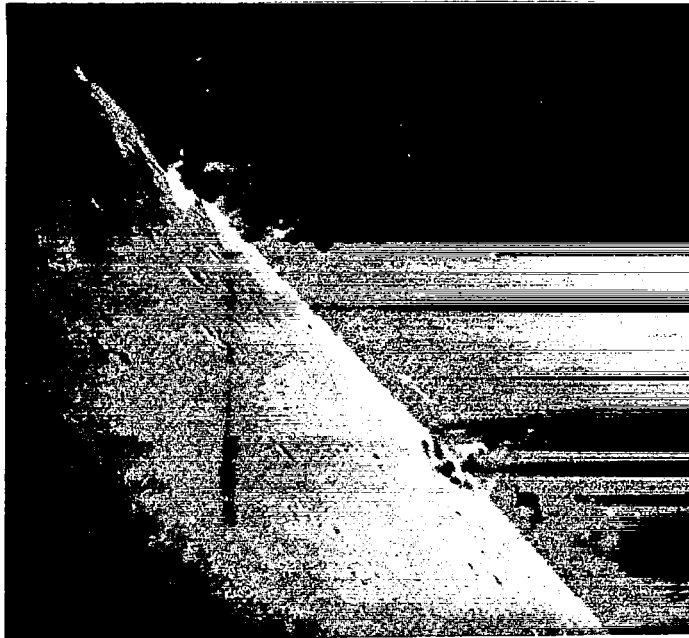
FIGURE IV-5.   ELECTRON MICROGRAPH (5000X) OF POLISHED EDGE
SHOWING RADIUS OF CURVATURE OF LESS THAN 1000 A.

PHASE SHIFTER

The role of the phase shifter is to introduce a known relative phase shift between the reference and signal beam. Proper choice of this phase shift is essential to obtaining good subtraction in the hologram region, except in the self-subtraction mode which is discussed elsewhere. Several means for producing phase shifts are known, and a choice must be made based on definite criteria. The criteria relevant to the needs of the preprocessor are

- Ease of selection of the phase shift
- Ease of fabrication
- Uniformity of the output wavefront
- Range of phase shifts available
- Operation with moderate voltages
- Physical size.

Because of the excellent electrooptical properties of $LiNbO_3$, considerations have been limited to electrooptic phase shifters. The precise form of phase shifter used will, of course, depend on the choice of planar-vs-channel waveguides. Two configurations for diffused waveguides and one for etched or ridge channel waveguides have been considered.

The simplest form for phase-shifter electrodes is the surface electrode pair shown in Fig. IV-6(a). This configuration can be used for planar or channel guides. The phase shift it produces can be written in the form

$$\frac{\Delta\phi}{V} = c(\theta,x,y)\left(\frac{L}{w}\right) \qquad \text{(IV-21)}$$

where L is the electrode length, w the gap width, $\Delta\phi$ the phase shift, V the applied voltage, and θ is the angle between the direction of propagation of the beam and the crystalline c-axis; x and y are referred to the axes in Fig. IV-6(a). The parameter c is given by

$$c(\theta,x,y) \equiv (2N^3 r_{eff}/\lambda_o)g(x,y) \qquad \text{(IV-22)}$$

(a) SURFACE ELECTRODES   (b) PARALLEL PLATE ELECTRODES

(c) RIDGE GUIDE
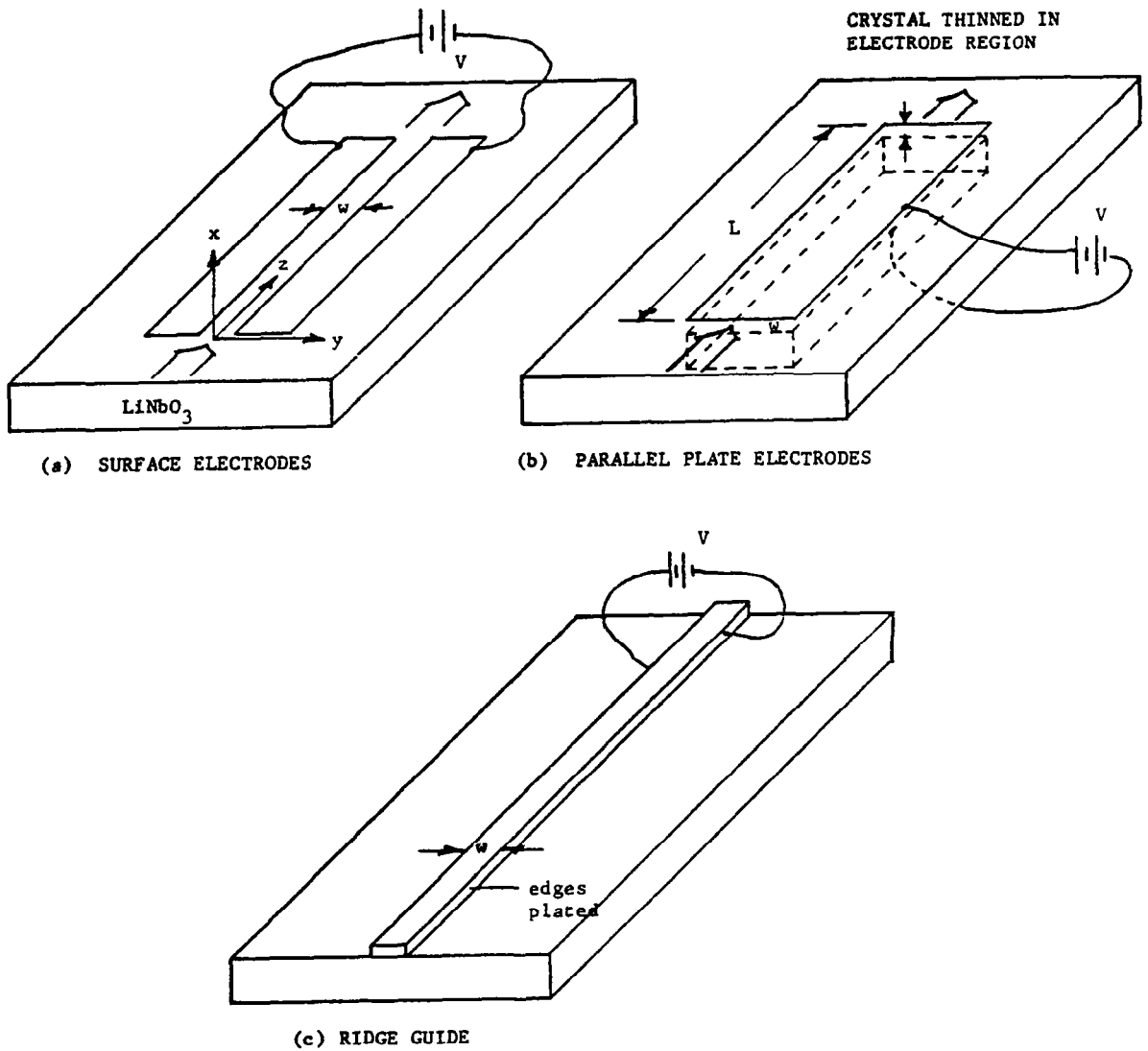
FIGURE IV-6. THREE PHASE-SHIFTER STRUCTURES.

For TE modes,

$$r_{eff}(\theta) = [r_{33}\sin^3\theta + r_{13}\sin\theta\cos^2\theta + r_{42}\sin\theta\sin2\theta] \qquad \text{(IV-23)}$$

The r_{ii} are electrooptic coefficients which for LiNbO$_3$ are

$$r_{33} = 3.0 \times 10^{-5}\frac{\mu m}{V}, r_{13} = 1.0 \times 10^{-5}\frac{\mu m}{V}, r_{42} = 3.2 \times 10^{-5}\frac{\mu m}{V}. \quad \text{(IV-24)}$$

N is the modal effective refractive index, and λ_o is the vacuum wavelength of the light. The function $g(x,y)$ describes the variation of the in-plane part of the electric field in the gap relative to that at $(x,y) = (0,0)$. There is also an x- component of the field that has been neglected in Eq. (IV-21). This neglect is justified if the light is confined to the central portion of the gap. The uniformity of the output beam will depend on the variation in g. For phase-shifters, the gap width of surface electrodes will be much larger than the confinement depth of the waveguide in planar guides, so one can take x = 0 and use a result from electrostatics to obtain

$$g(0,y) = w(w^2 - 4y^2)^{-1/2} = (1 - \frac{2y^2}{w})^{-1/2} \qquad \text{(IV-25)}$$

One finds a 10% variation in g for that part of the beam inside of the central 42% of the gap. A curve of percent-variation-in-g vs the normalized parameter 2y/w is shown in Fig. IV-7.

In Fig. IV-6(b) is shown an alternative electrode structure for planar waveguides. Here, the crystal is thinned in the electrode region and electrodes are deposited above and below the waveguide. In this case, $g = \pi/2$, and $r_{eff} = r_{22} = 6.6 \times 10^{-5}\frac{\mu m}{V}$. Beam uniformity is excellent for this case, but is obtained at the sacrifice of more complex fabrication. The effective electrooptic coefficient is lower in this case, but this is mitigated by having a higher field because of the lower w, as compared to Fig. IV-6(a).

In Fig. IV-6(c) is pictured a ridge guide with its vertical edges plated, so that w is the width of the guide. The effective electrooptic coefficient is again that of Eq. IV-24 and $g = \pi/2$. This electrode type has been investigated by Kaminow[8] and shown to be very effective with good

59
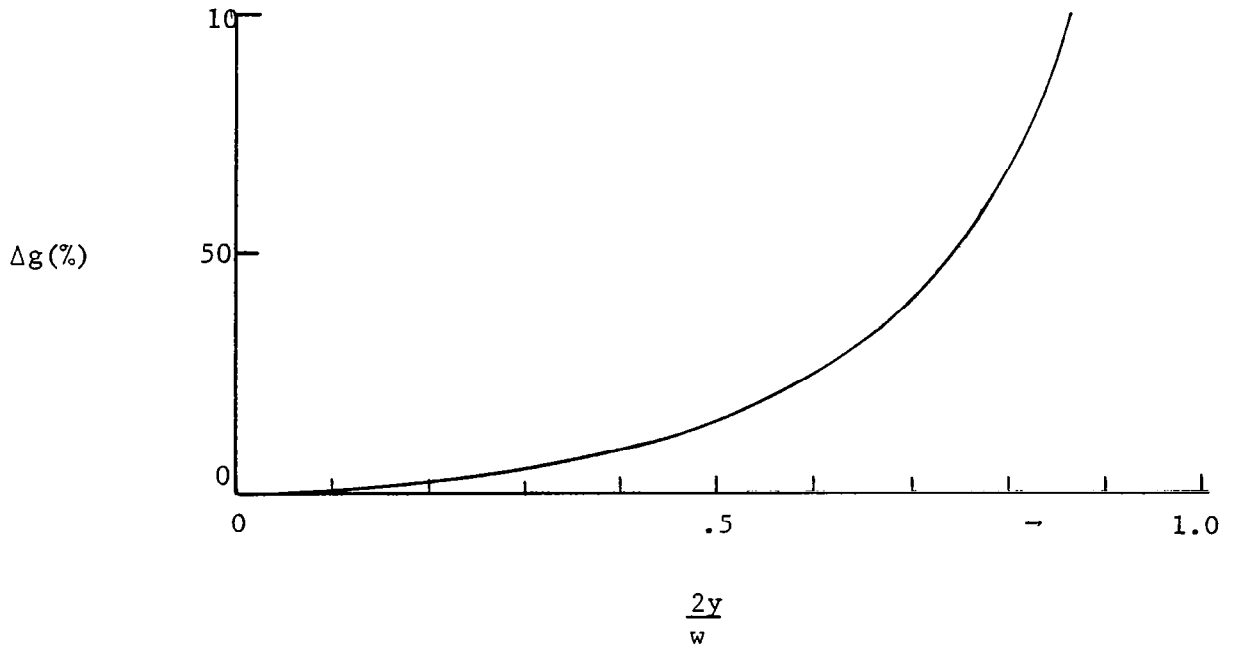
FIGURE IV-7.   PERCENTAGE VARIATION IN g-vs-$^{2y}/w$.

uniformity and low operating voltages. These advantages are, again, obtained at the expense of ease of fabrication, since ridge guides are more complicated to make than simple planar ones. However, if overall system considerations make ridge guides desirable, then the actual placement of electrodes is only somewhat more complex than placement of surface electrodes.

Experimental investigation of the actual phase shift obtained with simple surface electrode show good agreement with calculated values. No investigation of the parallel-plate electrodes of Fig. IV-5(b) has been carried out because crystal thinning techniques are not fully developed and the thining process may seriously weaken the crystal. However, $LiNbO_3$ grinding techniques developed under another program [19] show great promise and should be easily adaptable to phase shifter fabrication.

<center>DATA INPUT ELECTRODES</center>

  The data input electrodes impress the externally generated data onto
the optical beam by inserting a phase shift, proportional to the data voltage,
into selected portions of the input beam. They are, therefore, phase shifters
and the remarks made in above apply equally well to them. However,
there are several effects and considerations that strongly affect the design
of input electrodes which were not discussed in connection with phase shifters
because they are either irrelevant or only weakly affect their design. The
subjects to be discussed are geometry, diffraction, signal-to-noise ratio, and
field effects. In channelized versions of the preprocessor, these effects
are absent or are less severe than for planar versions, so we restrict our
attention to the latter. Further, for definiteness, we limit the discussion
to surface electrodes. Similar analyses will apply to the parallel-plate
electrodes.

  In the following paragraphs, we discuss each of the subjects listed
above and give relevant formulas useful in design.

<center>Geometry</center>

  The relevant geometrical parameters for surface electrodes are shown
in Fig. IV-8. In the drawing, the total channel width is W and is made up
of one full data channel of width w and one full guard band of width d. The
ratio f = w/W will be called the packing fraction. If a planewave is incident
from the left on a set of such electrodes, a fraction f enters signal channels,
the rest passing under the guard band. The part of the beam passing under the
guard bands continues to the hologram, unmodulated, and is rejected by the
holographic subtraction process. To the extent that this process is imperfect,
this part of the input beam contributes to the noise. While it is desirable
to have f as large as possible, increasing f must decrease d (because the
electrodes cannot be arbitrarily long) thus allowing more cross talk.

  It was shown above that the phase shift can be written

$$\Delta\phi = cV(\frac{L}{w}) \equiv cAV \qquad\qquad \text{(IV-26)}$$
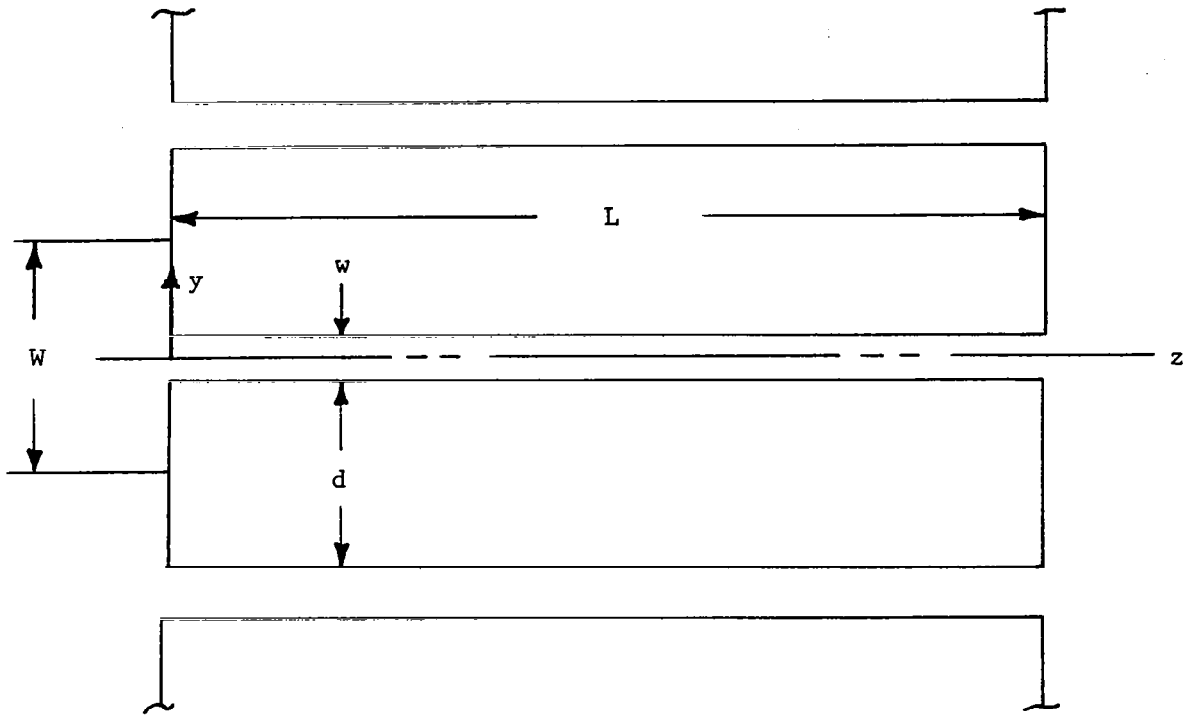
FIGURE IV-8.   SURFACE ELECTRODE GEOMETRY.

where A is the aspect ratio and c is a function of θ, x, y. Fairly large values of A may be desirable in order to increase $\Delta\phi$ and thus the sensitivity. However, larger A means smaller w resulting in higher diffraction which has the consequence of more cross talk or smaller packing fraction.

It will be seen below, as it is suggested above, that the maximum allowable electrode length, L_{max}, is one of the dominant design factors. For most designs, one expects to have an L_{max} of a few millimeters, although some designs may allow larger values. The overall electrode width for N channels is also an important parameter. If L_{max} is given, the total width will be

$$W_{Total} = NW = NL_{max}/Af \qquad (IV-27)$$

If W_{Total} is restricted to be less than some W_{max} (which is determined by the available crystal area), then Eq. (IV-27) restricts Af to

$$Af \geqslant N \frac{L_{max}}{W_{max}} \qquad . \qquad (IV-28)$$

Diffraction

The refractive index of the electrode gap region differs from that of the guard band; the channel therefore is, in effect, an aperture of width w. If the field in the gap was uniform and the length, L, was small, then the effect of the aperture would be to produce a diffracted light distribution whose amplitude just beyond the structure has the form

$$(e^{i\varphi}-1)W \, \text{sinc} \left(\frac{w}{\lambda}\, \alpha \right) \qquad (IV-29)$$

where φ is the phase shift due to the gap, α is the sine of the diffraction angle, λ is the wavelength, and

$$\text{sinc}(X) = \sin\pi X/\pi X \qquad (IV-30)$$

For larger values of L, the distribution becomes more complex in detail, but the main features still follow Eq. (IV-29). Nonuniform fields further complicate the pattern. However, since the field-induced refractive index changes are rather small, we will assume the effects of diffraction and of

64

field nonuniformity can be separately calculated and superimposed. This assumption greatly simplifies the discussion.

The sinc function is graphed in Fig. IV-9, from which we see that it has a central peak of height 1, zeros at $x = \pm m \frac{\lambda}{w}$, $m = 1,2,3,\ldots$, and subsidiary peaks near $\alpha = (m + \frac{1}{2}) \lambda/w$ of height approx. $1/(m + \frac{1}{2})\pi$. The energy in each lobe decreases as m increases and is given by the integral of the absolute square of Eq. (IV-29). Thus, the central lobe contains 90% of the diffracted energy, the next lobes contain 2.4% each, the next .8% each, etc.

Diffraction effects impact directly on the operation of the hologram because a divergent beam tends to reduce the overall efficiency of holographic subtraction. They also impact on cross talk, because diffractive spreading can send light from one channel into another if the guard bands are too narrow.

The maximum length and width of the electrode structure again comes into play here through Eq. (IV-26). For $LiNbO_3$, the phase shift in the gap center for propagation perpendicular to the $\hat{c}$-axis of the crystal can be obtained by taking $c \approx 10^{-3} \frac{rad}{volt}$. Then, for specified $\Delta\phi_{max}$/channel, we have

$$A \geqslant \frac{\Delta\phi_{max}}{c V_{max}} = \frac{1000}{V_{max}} \Delta\phi_{max} \qquad \text{(IV-31)}$$

where V_{max} is the maximum voltage available per channel. For TTL compatible systems, $V_{max} = 5$ Volts, so we have

$$A \geqslant 200 \, \Delta\phi_{max} \qquad \text{(IV-32)}$$

Since $\Delta\phi_{max}$ will be an appreciable part of 1 radian, A is likely to be fairly large. If cross talk requirements demand that the first $(2\ell-1)$ lobes of the sinc function are contained in the channel, then we have

$$\frac{\ell\lambda}{w} \sim \frac{w}{2L} = \frac{d}{L} + \frac{1}{2A} \qquad \text{(IV-33)}$$

Eq. (IV-33) specifies d for given A and degree of cross talk protection. It would also be desirable to keep, say, the first (2k-1) lobes inside the gap angle, 1/2A. This would lead to

1.0

0.13

-3   -2   -1   1   2   3   X

-0.22

-0.09
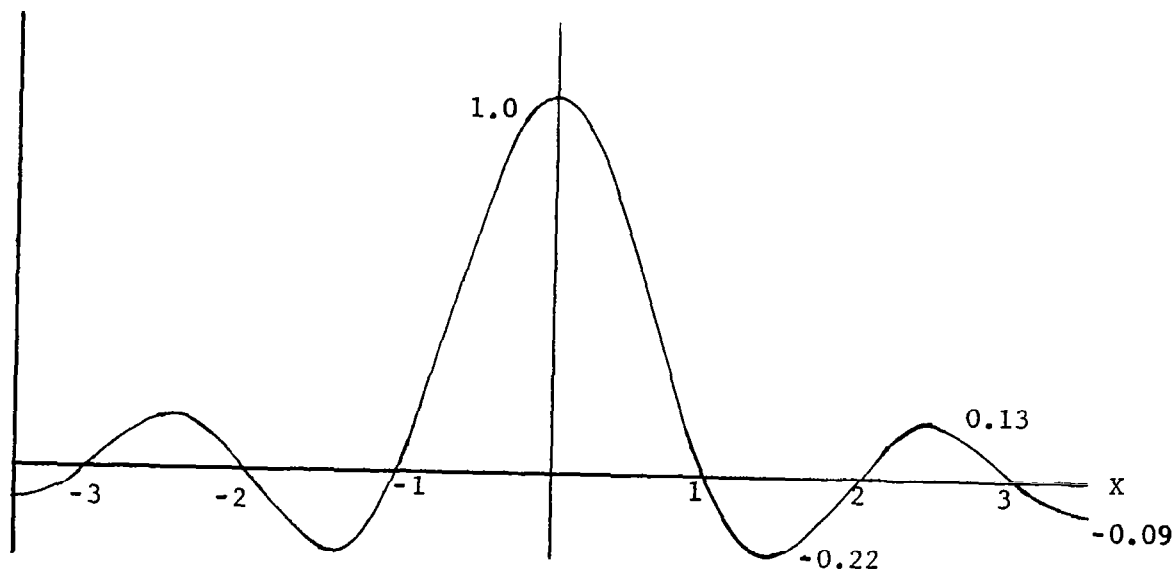
FIGURE IV-9. Sinc(X)-VS-X.

$$\frac{k\lambda}{w} \sim \frac{w}{2L} = \frac{1}{2A} \qquad\qquad (IV-34)$$

Then, Eq. (IV-33) would become

$$(\ell-k)\ (\lambda/w) \sim d/L \qquad\qquad (IV-35)$$

Given k and ℓ, Eqs. (IV-34) and (IV-35) determine w and d in terms of A:

$$w \sim 2k\ A\lambda;\ d \sim (\ell-k)\ L/(2kA) = w(\ell-k)/2k \qquad (IV-36)$$

It will not prove generally possible to accommodate Eq. (IV-34) with integer k and Eq. (IV-32) for moderate $\Delta\phi_{max}$ simultaneously. Since Eq. (IV-32) must take priority, we may be required to relax Eq. (IV-34) to allow fractional values of k. This will cause unavoidable variations in the phase shift for that part of the diffracted energy that leaves the gap, and will therefore decrease the signal-to-noise ratio.

Signal-to-Noise Ratio

The signal-to-noise ratio is affected by the electrodes because they introduce diffraction, which affects the efficiency of holographic subtraction, and because they scatter both the diffracted (signal) and the undiffracted beam. If necessary, the scattering can be reduced by use of a buffer layer between the metallic electrodes and the waveguide surface, and by careful preparation of the other parts of the surface. Beam divergence through diffraction can be reduced by using the largest w consistent with the other constraints.

The magnitude of the signal-to-noise ratio, S/N, determines the minimum detectable signal, which, in turn, determines the phase shift required. The direct impact on electrode design is, therefore, on the aspect ratio.

The detected signal is generally proportional to $1 - \cos(\Delta\phi) \approx \frac{1}{2}\Delta\phi^2$. If we assume these contributions can be added, channel by channel, then we must specify two signal levels: 1) the phase change, $\Delta\phi^{(1)}$, in a single channel that is just detectable, all other channels having no change; 2) the phase change, $\Delta\phi^{(N)}$ that is just detectable if it appears on all N channels. Writing the response of the system in the form

$$s = \alpha \sum_{i=1}^{N} (\Delta\phi_i)^2 \qquad\qquad\text{(IV-37)}$$

we have

$$s_{min} = \alpha \, (\Delta\phi^{(1)})^2 \qquad\qquad\text{(IV-38)}$$

with s_{min} determined by S/N. Also,

$$s_{min} = N\alpha \, (\Delta\phi^{(N)})^2 \qquad\qquad\text{(IV-39)}$$

so we find

$$\Delta\phi^{(N)} = \Delta\phi^{(1)}/\sqrt{N} \qquad\qquad\text{(IV-40)}$$

under the assumptions we have made.

Field Effects

There are two related field effects of importance: (1) ray bending because of the nonuniform fields, and (2) lateral waveguiding. Ray bending is unavoidable when surface electrodes are used, because the fields are fringing fields whose variation cannot be controlled. The bending may, depending on the sizes of the applied voltage, cause the beam to spread or to converge. If the beam is converged with sufficient strength, the gap region becomes a waveguide, and all of the light incident upon the gap becomes trapped in it. This is a desirable effect because it guarantees that the light entering the gap receives the full phase shift intended for it. Upon exit, the beam will spread with the pattern characteristic of the effective guide width. In most cases, this should not be much different from w.

The field at the waveguide surface can be written in the form

$$E(y) = E(0)/\sqrt{1-\eta^2} \qquad\qquad\text{(IV-41)}$$

where $\eta = 2y/w$ and E (0) is the field in the center of the gap. The resulting index is pictured schematically in Fig. IV-10(a) and (b) for the two possible polarities. The field of Eq. (IV-41) is infinite at $\eta = \pm 1$, but

68
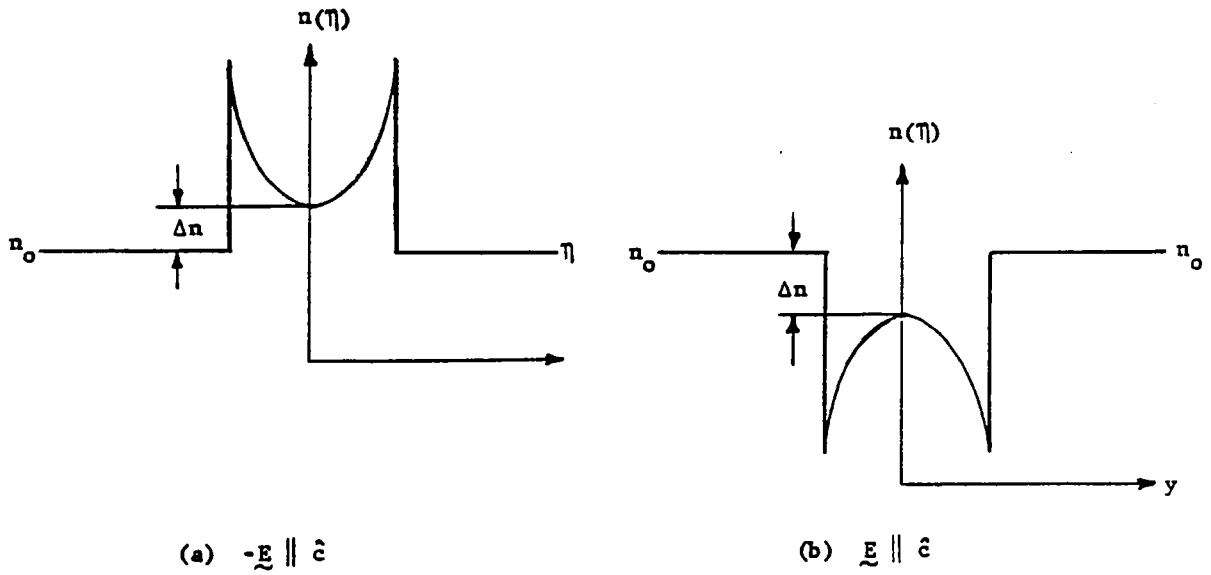
FIGURE IV-10. SCHEMATIC OF VARIATION IN INDEX ACROSS ELECTRODE GAP.

in a real system, these values would be finite. In Fig. IV-10(a), the energy would be trapped in the gap but would tend to reside in the wings. In Fig. IV-10(b), the energy would not be trapped, but would tend to reside in the gap center and leak through the wings. The wings shown for the surface of the waveguide rapidly diminish as one looks deeper into the guide, and finally disappear. Since the fields in an outdiffused $LiNbO_3$ waveguide penetrate to about 10-12 μm, it is probably reasonable to ignore the wings. Finally, it should be noted that the lateral waveguide shown in Fig. IV-10(a) is highly symmetrical, so its lowest mode should persist even for very weak fields. The wave is not well-confined for such weak guiding, but the entire guided wave partakes of the phase shift. Thus, even if fractional values of k in Eq. (IV-34) are used, good electrode operation should result if care is taken to avoid cross talk by using sufficiently wide guard bands.

Electrode configurations other than that of Fig. IV-8 are possible and may have important advantages. One such configuration is discussed in detail in Section VII of this report.

GEODESIC LENS DEVELOPMENT

The design of the integrated optical data preprocessor shown in Fig. II-9 employs a short focal-length geodesic lens to image the hologram output onto a single detector. Although the fabrication of geodesic lenses in $LiNbO_3$ waveguides was addressed in a previous NASA program,[20] further developmental work was necessitated by the fact that the radiative loss of light passing through a $LiNbO_3$ geodesic lens can be quite severe. During initial studies, this fact was not apparent because a considerable portion of the light radiated from the lens depression was coupled out of the slab in the form of a diverging spherical wavefront, just as expected for a properly focused guided wave. This anomaly was detected and corrected early in the program. Subsequently, and chiefly as a result of this work for NASA, Battelle was selected by the Air Force to perform a program "Optical Geodesic Waveguide Lenses".[19] During the course of this program, techniques for fabricating geodesic lenses in $LiNbO_3$ substrates were perfected, and methods of reducing aberrations in these lenses were developed. The lenses are judged to be better than required for the

operation of a holographic subtraction processor. Because of the relevance of the Air Force lens program to the present work, we have included a summary of it in Appendix C. In the body of this report, we restict discussion to work performed for NASA under the present contract, which was restricted to the important problem of forming geodesic waveguide lenses with low radiation losses.

Radiation Losses in Geodesic Lenses

The mechanism by which light is radiated from a geodesic lens depression is fundamentally the same as the mechanism by which light is radiated from a bent optical fiber, a phenomenon which has received considerable attention in the literature.[21,22] For a spherical waveguide surface so rounded that the center of curvature is in the low-index superstrate, guided light will radiate into the substrate at a rate that increases as the substrate index n_b approaches the mode index of guided light n_g. For waveguides in $LiNbO_3$, $(n_g-n_s)/n_b$ is generally much smaller than for thin-film or ion-exchanged glass waveguides. Consequently, radiation losses are more severe in $LiNbO_3$, as we have found.

However, radiation losses can be kept to a tolerable level by utilizing waveguide fabrication techniques that result in values of $(n_g-n_b)/n_b$ on the order of 1%. Titanium indiffusion is the most convenient means for accomplishing this. Lithium-outdiffused waveguides generally have $(n_g-n_b)/n_b \simeq 0.1\%$ and are not useful for geodesic lenses of the required short focal lengths.

Geodesic Lens Experiments

We have fabricated approximately spherical depressions in two slabs of $LiNbO_3$, and have tested the focusing capabilities of these depressions when several different types of waveguides were formed on the surfaces. Average radii of curvature of the depressions were 6.7 mm and 14.4 mm. When outdiffused guides of depth 380 μm and 14 μm, respectively, were formed in the samples, virtually all guided light was radiated from the depression regions and no focusing could be observed. When a 30 nm layer of Ti was indiffused for 3 h at 950°C into the slab containing the smaller radius depression, an

identical result was obtained. Subsequently, an 87 nm layer of Ti was e-beam evaporated onto the slab containing the larger radius depression and indiffused for 18 h at 960°C. This operation resulted in a 50-mm focal-length geodesic lens of high optical quality. Figure IV-11 shows a photometric scan of the focal region obtained when a collimated 2 mm-wide guided beam was incident on the lens. The beam was incident to one side of lens center to avoid an imperfection introduced there by manual grinding. The measured FWHH focal spot size was 13.5 μm, whereas the predicted theoretical value is 6.1 μm. About 2.1 μm of the difference is accounted for by diffraction spreading introduced by the imaging system. Consequently, the performance of the lens is within a factor of 2 of diffraction-limited. While this is adequate for the needs of a holographic subtraction processor, superior lenses are now routinely formed by a two-step machining and polishing process developed for the Air Force, in which fabrication time and the risk of breakage are considerably reduced, while adherence to dimensional tolerances is greatly improved.

Waveguide Compatibility

The use of a Ti-infused waveguide in the lens region of the preprocessor is inconvenient, in view of the overall advantages that pertain to the use of Li-effused waveguides elsewhere in the system. A potential problem to be overcome is that the thermal treatment necessary to diffuse Ti into the lens region will simultaneously result in an undesirable degree of Li-effusion elsewhere. The best way to circumvent this potential difficulty is to optimize the Ti metal thickness so that a single heat treatment suffices for the simultaneous fabrication of both Li-effused and Ti-infused waveguides on the same substrate. The use of faster diffusants in place of Ti (Ni is one possibility) could facilitate this process: alternatively, one could employ a variety of procedures that have been devised for inhibiting the Li-effusion process during any metal diffusion step.[23,24] The simplest of these involves coating the sample with MgO prior to the metal infusion. After the metal-infused-waveguide is established, the MgO coating is removed and a Li-effused waveguide having the required properties is fabricated.
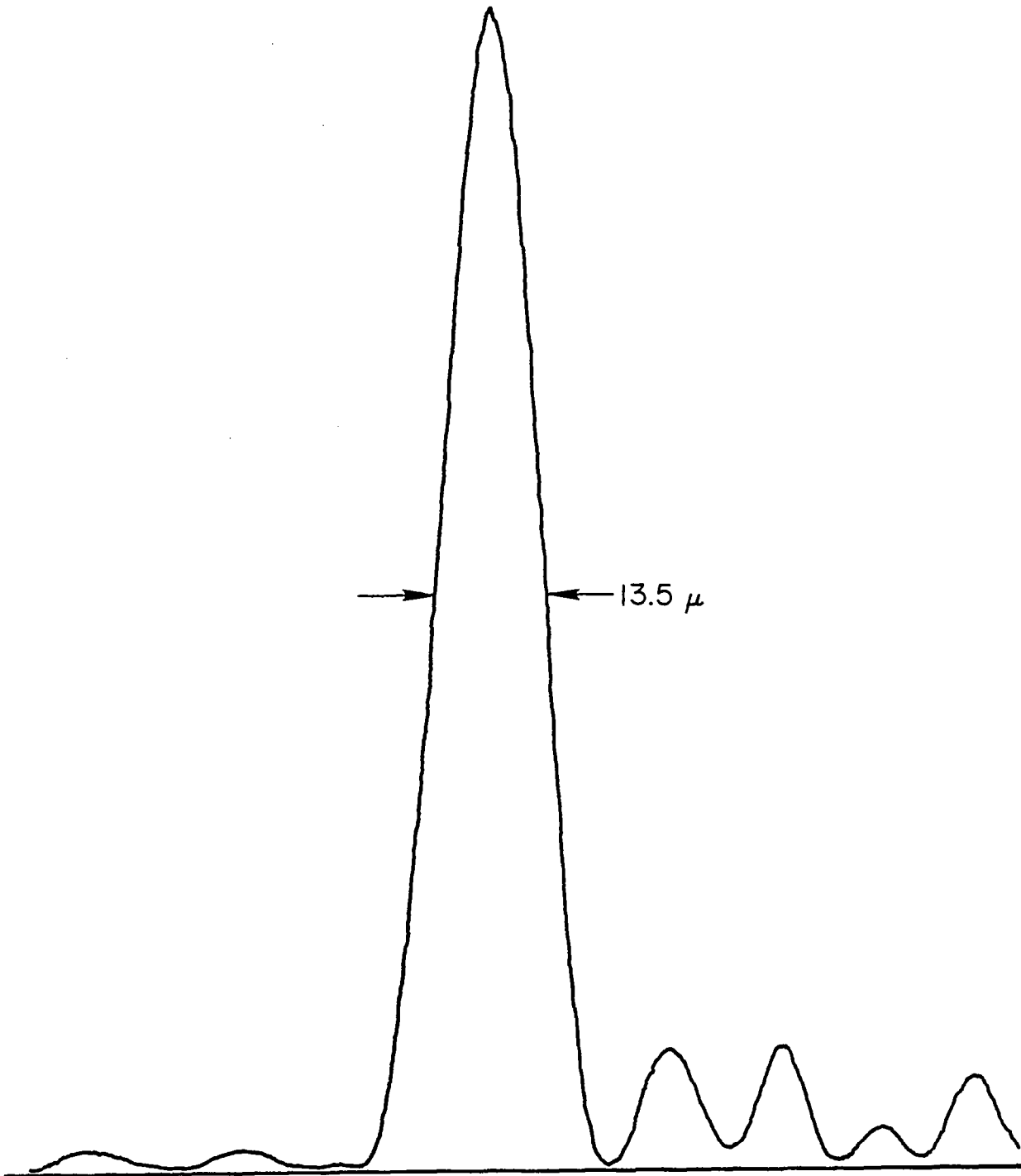
FIGURE IV-11.   FOCAL SCAN OF A GEODESIC LENS IN LiNbO$_3$
(50 mm FOCAL LENGTH, 2 mm INPUT APERTURE).

It may prove feasible to eliminate the need for a lens entirely by using tapered channels to project light onto a detector. This possibility will be investigated in connection with ongoing development of channel waveguide fabrication techniques.

Source and Detector

During the experimental phases of this program, a prism-coupled He-Ne laser has been used to investigate the behavior of the various preprocessor components and to characterize the three-channel laboratory demonstration model of the preprocessor. In operational models of the preprocessor it may be desirable for reasons of space, weight and power economies to replace the He-Ne laser with semiconducting laser diodes. It is therefore desirable at this time to consider the source characteristics necessary for the various modifications of the preprocessor. These are summarized below.

Preprocessor Operating Modes

Screening Mode. The incoming data are compared to a holographically recorded data set. The source must be capable of writing as well as reading the hologram. For an outdiffused waveguide in nominally pure $LiNbO_3$, the maximum wavelength for which low power (< 100 μW) writing has been demonstrated is .6328 μm. The use of longer wavelengths would surely be possible if the hologram region were doped with Fe^{2+}, but data are not available. In the screening mode it may be advantageous to write at a high power and read at a low power to prolong the usable life of the hologram.

Identification Mode. The incoming data are compared to a series of reference sets which are entered on a separate electrode set while the data voltages are present. In this case the primary function of the hologram is to act as a beam splitter. A secondary function is to compensate for undesired constant-wavefront perturbation due to the presence of the electrodes or to waveguide imperfections. The hologram may be written by a two-photon process or it may be replaced by a grating beam splitter. In this latter case, there

are no wavelength limitations on the light source. However, coherence length, which will be discussed below, is still a consideration.

Self-Subtraction Mode. In this mode the hologram is written and read continuously. The comments applicable to the screening mode are valid here also, except that only a single laser power will be required. Experiments are currently underway at the NASA Langley Research Center and at the Battelle Columbus Laboratories to determine the optimum laser and material parameters for operation in the self-subtraction mode.

Laser Characteristics

Power. The experiments on the three-channel preprocessor were carried out with approximately 100 μW of .6328 μm radiation coupled into the waveguide. Since the experiments were carried out in undoped $LiNbO_3$, it is reasonable to expect that the power requirement could be reduced by a factor of 100 by doping the hologram region with iron.

Spectral Width. The operation of the preprocessor, in any of its functional modes, depends upon the coherent interaction of the signal and reference beams in the hologram region. The coherence length requirements are therefore as severe when reading as when writing the hologram. These coherence requirements set a fundamental limitation on the maximum allowable spectral width of the source to be used in the preprocessor as can be shown by the following argument.

We first define the coherence length as the path difference in the two arms of the interferometer which will reduce the fringe contrast ratios to 2:1. The fringe intensity is given by

$$I = A_S^2 + A_R^2 + 2A_S A_R \cos(\phi_S - \phi_R) \qquad \text{(IV-42)}$$

where A_S and A_R and ϕ_S and ϕ_R are the amplitudes and phases of the signal and reference beams, respectively, and $\Delta\phi$ is their phase difference at the hologram region. As we move across the interference pattern, $(\phi_S - \phi_R)$ varies linearly. If $A_S = A_R \equiv A$ then $I_{max} = 4A^2$ and $I_{min} = 0$ and the phase contrast is $I_{max}/I_{min} = \infty$. The effect of a finite value of the laser coherence length

is to introduce a random phase variation in Equation (IV-42). If we denote this random phase variation by $\delta\phi$, then

$$I_{max} = 2A^2(1 + \cos\delta\phi)$$

and

$$I_{min} = 2A^2(1 + \cos[\pi + \Delta\delta])$$

The contrast is now

$$\frac{I_{max}}{I_{min}} = \frac{1 + \cos\delta\phi}{1 + \cos(\pi + \delta\phi)} \qquad (IV-43)$$

which for an allowable minimum contrast ratio of 2:1 gives

$$\delta\phi \overset{\sim}{<} 1.23 \text{ radians}$$

The phase uncertainty is related to the path length difference ℓ and laser bandwidth $\Delta\lambda$ by

$$\delta\phi = \frac{2\pi\ell}{\lambda^2} \Delta\lambda \qquad (IV-44)$$

To minimize fabrication problems the path length difference should be allowed to be as great as 0.5 mm. In addition, a path length difference of $w\tan\theta_B$ is encountered because of the finite beam width w. For a beam width of 2 mm and a 30° Bragg angle this adds an additional 1.15 mm to the coherence length requirement. Using the sum of these two numbers for the minimum coherence length, a wavelength of .82 μm and the previously derived value of $\delta\phi$, we see that

$$\Delta\lambda \overset{\sim}{<} 0.8 \text{ Å}$$

is required for satisfactory preprocessor operation.

The above derived constraint is easily satisfied by commercially available gas lasers. Conventional semiconductor diode lasers typically have line width much greater than the maximum allowable value and are currently

unsuitable for the preprocessor. However, recent advances in semiconductor laser technology have resulted in several successful approaches for limiting the laser linewidth and enhancing optical quality and stability. In a recent paper[25] on "Transverse Mode Stabilized $Al_x Ga_{x-1}$ As Injection Lasers with Channeled-Substrate-Planar Structure", K. Aiki, et al have demonstrated single mode stable operation with $\sim$ 40 MHz ($\sim$.001A) line widths. Another approach is the distributed feedback configuration. Here linewidths of about 0.2 Å have been demonstrated[26] at 320°K. In either case, powers should be sufficient for identification-mode operation.

V. SYSTEM INTEGRATION

The goal of the system integration effort was to demonstrate the fabrication of an integrated three-channel preprocessor utilizing the phenomena and the components discussed in the preceding sections. The construction of a laboratory model having all of the features illustrated in the schematic of Fig. II-9 would have been ideal. However, programmatic constraints necessitated the construction of a modified version of the preprocessor which was designed to demonstrate all of the operating characteristics of the device in a configuration which minimized the expenditure of time, money and materials. A photograph of the three-channel laboratory model which was fabricated and successfully tested is shown in Fig. V-1a; a line drawing on which the model was based is shown in Fig. V-1b. The design will be described and the details of its fabrication and testing will be presented in this Section.

CONFIGURATION OF THE THREE-CHANNEL LABORATORY MODEL

Function and Mode of Operation

The three-channel laboratory model of the preprocessor was designed to demonstrate that holographic subtraction could be performed in real time in integrated optics configuration. The model therefore incorporates the abilities to form a hologram of a given phase front and to detect departures from this phase front. It functions in the screening mode as described in Section II.

Geometry

The major distinction between the schematic of Fig. II-9 and the device actually constructed is that the symmetric geometry of the former was replaced by an unsymmetrical triangular geometry. This was done to minimize the amount of waveguide material required and to eliminate one of the mirrors. The asymmetric interferometer requires a laser with a larger coherence length than does the symmetric arrangement. It also lacks the insensitivity to

Output Beam — to Detector

Hologram Region

Data Input Electrodes

I mm

Outdiffused LiNbO$_3$ Waveguide

Polished TIR Mirror

Electrooptic Phase Shifter

Metallized Photoresist Grating Beamsplitter

Input Coupled He-Ne Laser
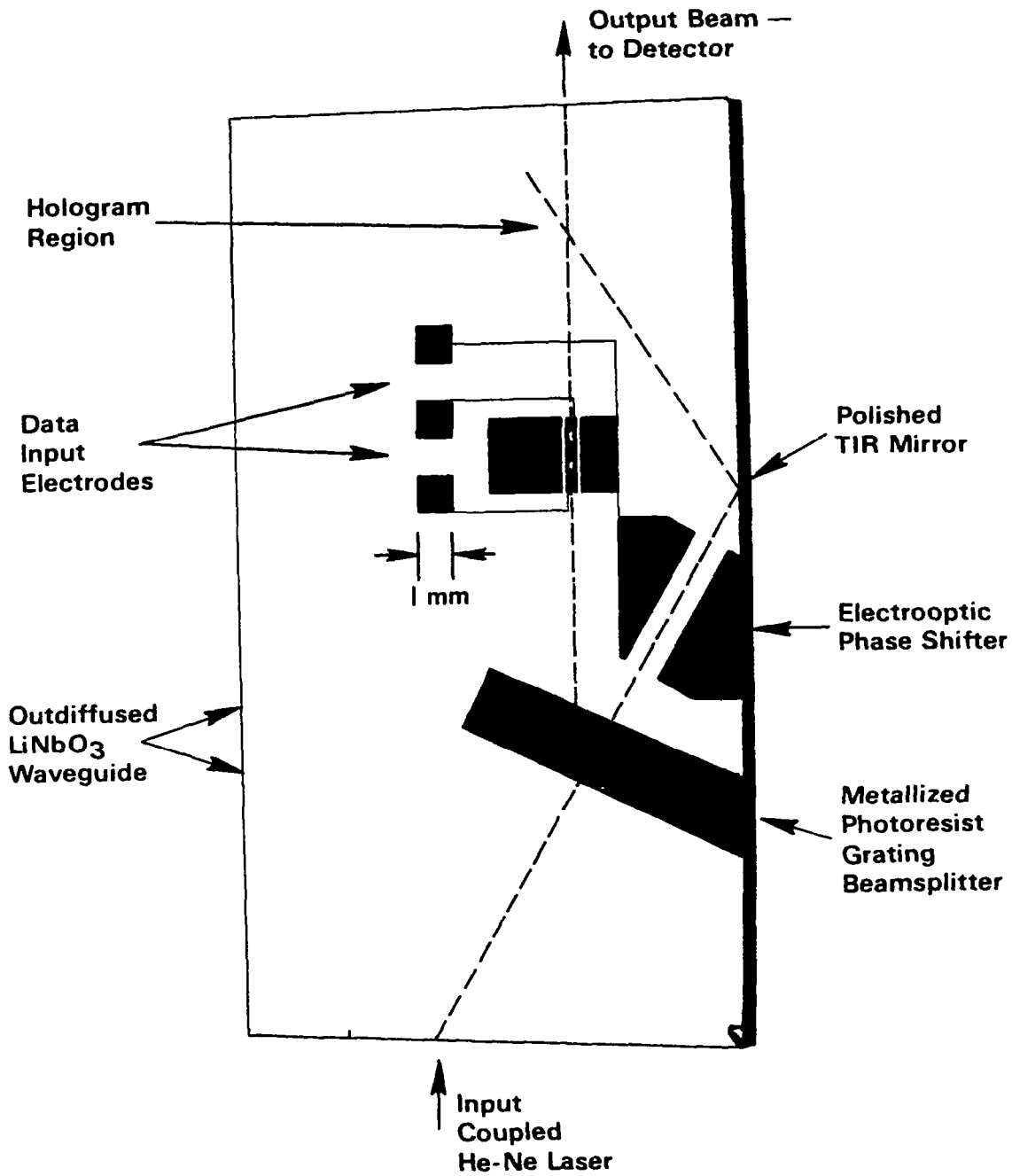
FIGURE V-1(a). PHOTOGRAPH OF THREE-CHANNEL LABORATORY MODEL SHOWING ELECTRODE STRUCTURES.
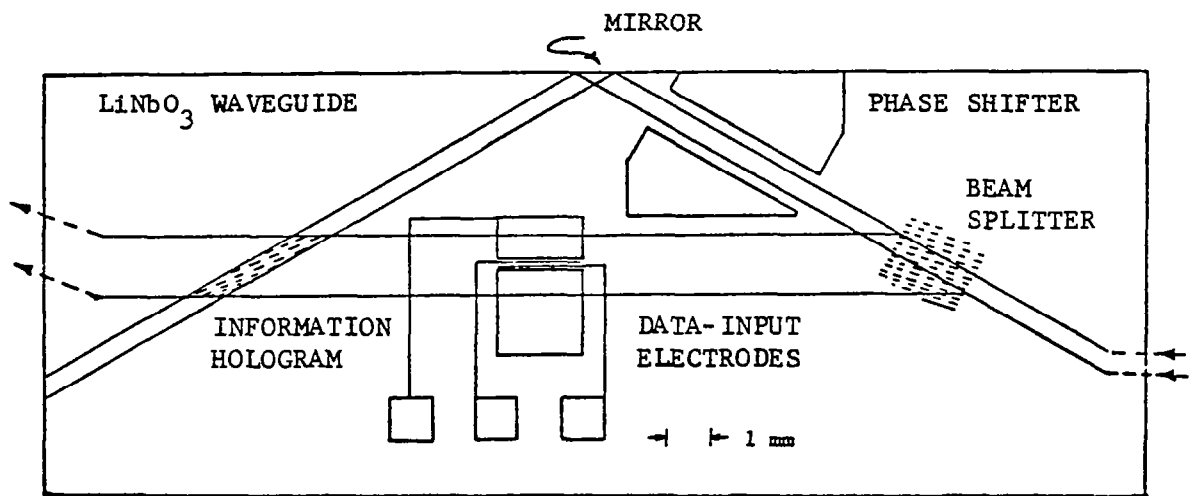
FIGURE V-1(b). LINE DRAWING OF THE 3-CHANNEL MODEL

thermal change inherent in the symmetric design. Neither of these features is relevant in a device designed for laboratory use with a long-coherence-length He-Ne laser.

Components not Integrated

The following components were omitted from the monolithic structure.

Laser. A remote He-Ne laser was used in place of a butt-coupled diode laser for several reasons. Most important is the fact that we had established the fact that waveguide holograms could be written and read with the He-Ne laser but had not demonstrated operation at the diode-laser wavelength of 0.8 μm. In addition, diode lasers with coherence length and stability required for the preprocessor are not yet readily available.

Lens. In spite of the fact that we have developed the capability for making geodesic waveguide lenses in $LiNbO_3$, the lens was not incorporated in the waveguide, primarily in an effort to conserve material and fabrication time. The preprocessor lens requirements are sufficiently lenient that lens quality and positioning should present no difficulty. The only problem associated with the incorporation of a lens in the waveguide may be associated with the necessity for infusing titanium in the lens region, while maintaining an outdiffused waveguide over the rest of the $LiNbO_3$ surface. This problem has been discussed in Section IV. Appropriate procedures are available and the problem may not exist at all in the next version of the device which may contain Ti-infused channel waveguides.

Detector. Due to the remote placement of the lens, it was obviously impossible to butt-couple the detector to the waveguide. A remote PIN diode was used. It had characteristics similar to detectors which might be coupled to the waveguide.

Phase Shifter

The simple phase-shifter structure shown in Fig. IV-6(a) was used, rather than the channelized structure indicated in Fig. II-9, or either of the

more complex structures shown in Fig. IV-6. The tradeoff involved here is that higher voltages are required but fabrication and alignment are much simpler. This is an appropriate tradeoff for a laboratory model. As the preprocessor technology evolves, a phase shifter requiring lower voltages will be used.

Data Input Electrodes

The dimensions of data input electrodes used in the three-channel laboratory model are shown in Figure V-2. Initially, in an effort to minimize interchannel cross talk, 10-μm signal channels and 100-μm guard bands were used. However, this approach resulted in a very poor signal-to-noise ratio which we attribute to the fact that only 10 percent of the light was in the signal channels. In addition, the narrower channel width increased diffraction effects and made the alignment of the electrodes with respect to the other components more difficult.

FABRICATION OF THE PREPROCESSOR

Fabrication of the preprocessor in $LiNbO_3$, exclusive of the source and detector, comprises 4 major operations:

(1) Formation of mirror on slab edge,

(2) Formation of a planar waveguide in the $LiNbO_3$,

(3) Formation of metallized phase shifter and data input electrodes on slab surface,

(4) Formation of grating beam splitter on slab surface.

Techniques employed in these fabrication steps included lapping and polishing, heat treatment, thin-film metallization, photoreduction, photo-lithography, metal etching, and precision alignment of preprocessor elements. Some of these techniques are conventional, others were developed for this fabrication task, as will be evident in the following sections in which fabrication steps are described in detail, from initial purchase of $LiNbO_3$ slabs to attaching coupling prisms to the preprocessor element.
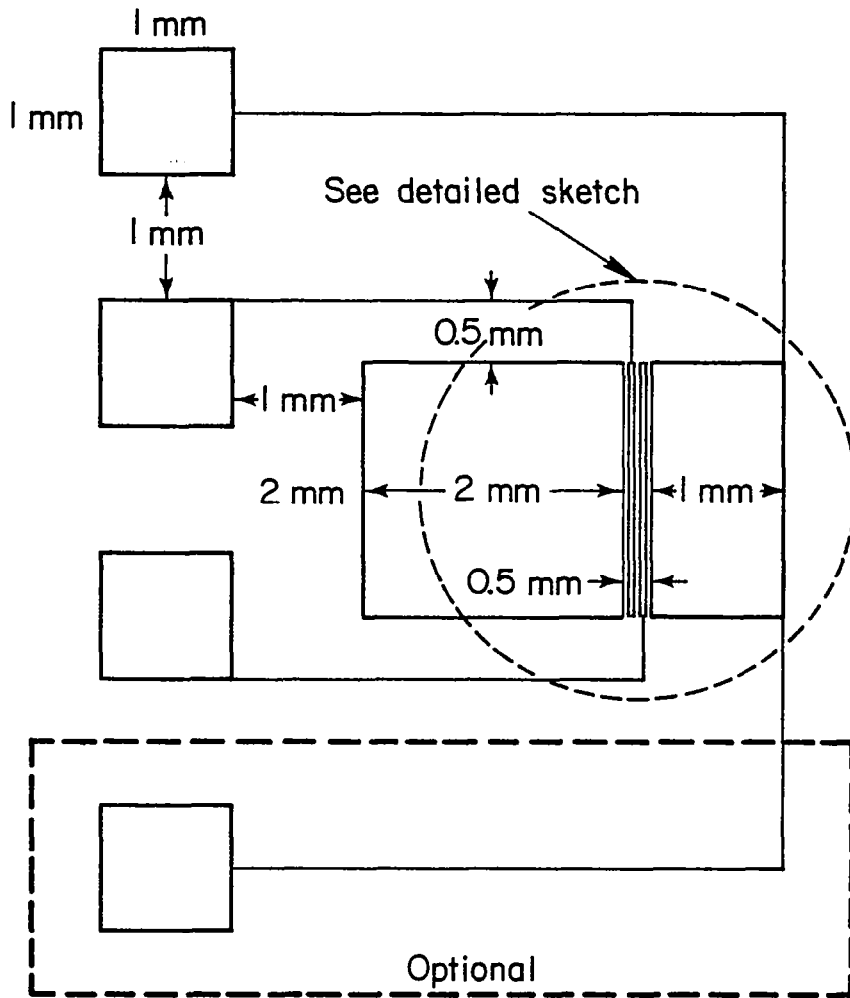
FIGURE V-2(a). OVERVIEW OF DATA-INPUT ELECTRODES SHOWING
BONDING PADS. AREA ENCIRCLED WITH DASHED
LINE IS SHOWN IN DETAIL IN FIG. V-2(b).
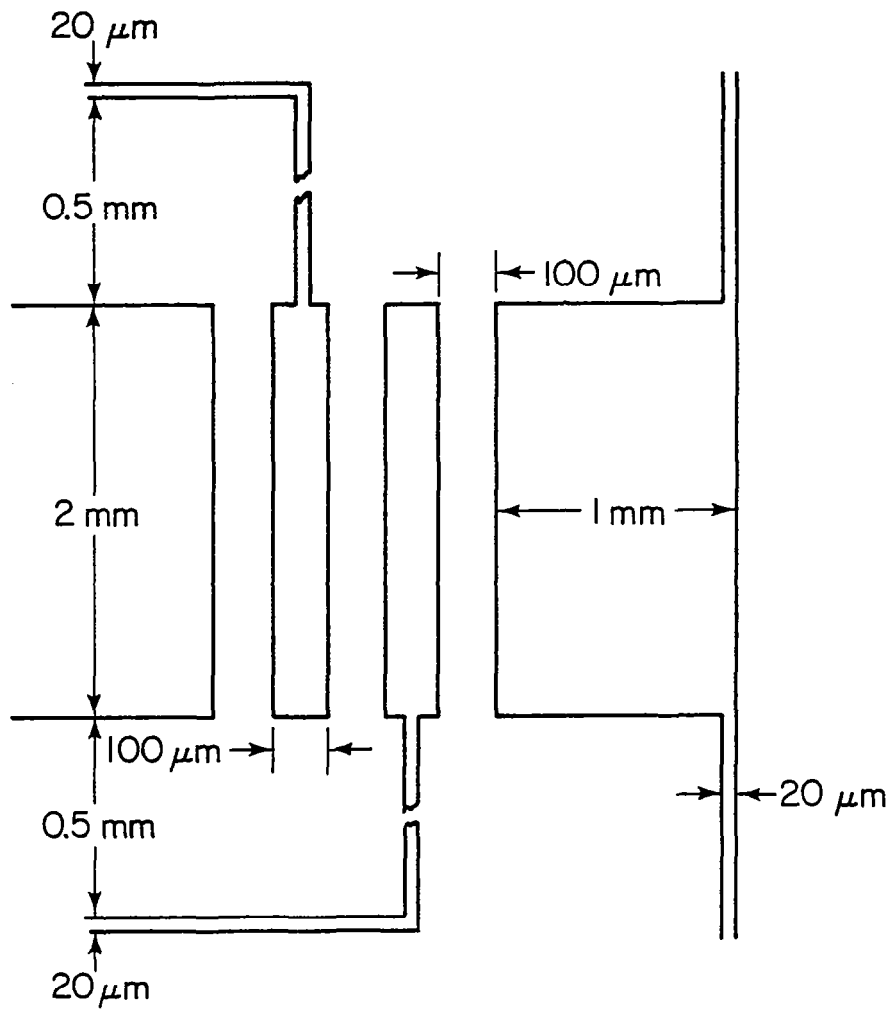
FIGURE V-2(b). DETAIL OF ELECTRODES.

Selection and Purchase of $LiNbO_3$

$LiNbO_3$ slabs are available from manufacturers with either the x,y, or z crystallographic axis perpendicular to the broad face. As discussed previously, both the x- and y- cut crystals are suitable for the preprocessor. Since y-cut slabs of the desired size are widely used for surface-acoustic-wave devices, and therefore readily available, they were the ones obtained for this work. A slab size 1-inch square by 3 mm thick was large enough to accomodate the preprocessor elements, was sufficiently rugged for handling, and was compatible with the various pieces of laboratory apparatus employed in the fabrication process.

$LiNbO_3$ slabs were obtained from Crystal Technology, Inc., a major and usually reliable supplier of various crystalline materials. Optical grade was specified, which, along with an optical polish and flatness to better than $\lambda/4$, included freedom from visible surface and internal flaws. Nevertheless, orders had to be returned on several occasions because of striations in the material which extended from surface to surface. A number of similarly flawed slabs, cut from a specially grown "high-purity" boule, have been used as expendable experimental samples and have not as yet proven to be unsatisfactory because of the striation-type of flaw. Until further proof is in hand that such flaws are, indeed, of no consequence, samples containing them will not be considered acceptable.

Waveguide Formation

Waveguides have been formed in $LiNbO_3$ both before and after lapping and polishing the edge. Thus far, no damage to the mirror surface has resulted from lithium effusion during heat treatment after polishing. However, there is no evidence that the out-diffused waveguide thus formed beneath the polished surface has a beneficial effect on reflecting qualities, and it might, indeed, prove later to be detrimental. Therefore, waveguide-before-mirror formation is the preferred order in device fabrication.

Waveguides were formed either by outdiffusion of lithium or infusion of titanium. The comparative characteristics of the two types of waveguides,

and the bases for choice, are discussed in Section IV. $LiNbO_3$ slabs with out-diffused waveguides have been used in preprocessors fabricated thus far. Nevertheless, both types of waveguide formation will be described.

Preparatory to waveguide formation, the $LiNbO_3$ slabs were cleaned. For virgin slabs, a rinse in isopropanol or methanol, or a utility mixture of equal parts of isopropanol, acetone, and trichlorethylene, usually has sufficed. A 5-minute soak in a solution of 64 parts distilled water to 1 part MICRO (a proprietary liquid detergent similar to LIQUINOX) followed by a warm tap-water rinse in turn followed by a final rinse in distilled or demineralized water was used in more stubborn cases. Agitation in an ultrasonic cleaner accelerates the MICRO-soak step. The slab was dried in a hot air stream. Drying by pressurized nitrogen or Freon gas frequently was substituted for the hot air method until occasional pinholes in metal films were traced to contamination of substrates dried in this manner. Pressurized gas drying has been abandoned until the nature and true source of the contamination has been determined.

Out-Diffused Waveguide in $LiNbO_3$

Formation of an out-diffused $LiNbO_3$ planar waveguide was accomplished simply by maintaining the slab at 950 C in a flowing oxygen atmosphere for several hours. A single-mode waveguide is produced with a 1-hour heat treatment, while 3 hours at 950 C produced a 2-mode waveguide. The slab, resting on narrow platinum or paladium shims on an alumina boat, was placed in the hot zone of the cold furnace and brought to temperature in the 30-minute furnace-heating time. At the end of the desired period of time, the slab was quickly withdrawn to a 600-650 C region of the quartz-tube furnace and retained there in an argon atmosphere as the furnace cooled to room temperature. No harm was done if the slab, after 20 minutes or thereabouts at 600-650 C, was withdrawn in 100 C increments (10 minutes at each temperature) to effect more rapid removal of the sample from the tube.

It should be emphasized that considerable care was taken to preclude subjecting the $LiNbO_3$ slab to thermal shock when changing environments either in cleaning or heat-treating. The critical temperature interval possibly is

as low as 20-30 C within the range from room temperature to 100-150 C. Not all slabs exhibit this tendency to fracture under moderately rapid temperature changes; a method for determining this sensitivity before the fact has not evolved.

Titanium Indiffused Waveguide in LiNbO$_3$

Formation of a planar waveguide in LiNbO$_3$ by indiffusion of titanium was accomplished in a two-step process: (1) deposition of a thin film of titanium on one surface of the slab, and (2) heat-treating the metallized slab. An electron-beam vacuum evaporation system was used to deposit the Ti film. As the result of a number of trials, with film thicknesses ranging from 150 to 500 A, it was determined that 400 A is the proper film thickness for formation of a 2-mode waveguide. The evaporator is equipped with a crystal film-thickness monitor to permit real-time determination of the evaporated film thickness. The as-prepared films exhibited a higher transparency than expected for a pure Ti layer. Undobutedly, the film comprised transparent TiO$_2$ as well as Ti as would be expected for an evaporation carried out at chamber pressures in the 10^{-6} Torr range, as these were. The presence of TiO$_2$ is not considered detrimental to waveguide formation, since the Ti becomes completely oxidized in the subsequent heat-treatment and is believed to diffuse into the LiNbO$_3$ as TiO$_2$ rather than Ti.

Heat-treatment for infusing Ti into LiNbO$_3$ is the same as for the out-diffusion process described above, namely, 3 hours at 950 C in flowing oxygen, and rapid quench to 600-650 C with subsequent cooling in an argon atmosphere. The Ti-infused surface thus formed appears free, to the unaided eye, of residual titanium or TiO$_2$. A haze is visible on Ti-coated slabs heat-treated for less than 3 hours; this haze usually can be removed by additional heat treatment.

Mirror Formation

This preparation step is described in Section IV, Component Development.

Photolithographic Masks

The 3-channel data input electrodes and the phase shifter (see Figure V-1(b)) were formed using photolithographic techniques. Initially, a separate mask was used for each of the two patterns and the metallized copies formed in a 2-step procedure. Later, the two patterns were combined on one mask, thereby eliminating an alignment problem and the second photolithographic operation. Two black chrome-on-glass masks for the data input electrodes with 10-μm channel width, one positive and one negative, were purchased from Qualitron Corporation. Copies of these masks then were made by contact printing on Kodak 649F 35-mm high resolution film. Such copies were used in the mask aligner, thereby preserving the glass masks as masters. Also, the film masks could be cropped easily for placement outside the mask aligner positioning range.

The masks for the phase-shifter pattern, and later for the data input processor having 100-μm channel widths, were made at Battelle. The master was laid out on 1/4-inch-thick clear acrylic stock, using black craft paper and 1/16-inch wide gummed black draftsman's tape. The master then was photoreduced onto Kodak 649F 35-mm high-resolution (more than 2000 lines per mm) film in a one-step operation using an Olympus OM-1 single-lens-reflex camera with f/1.4 lens. Photoreduction, in the neighborhood of 100, was carried out with the camera at an appropriate distance from the master panel placed against a sheet of translucent Plexiglas back-illuminated by a bank of six 100-W photofloodlamps in reflectors. A camera lens opening of f/5.6 gave the sharpest image; exposure times of 8-10 seconds were adequate.

Since positive photoresist was employed for all of the more recent photolighography, the Battelle-made masters were negatives, thus producing a positive photoreduced mask on film. For future patterns, it might prove simpler to make a positive rather than a negative master. In this case, a positive film mask can be made by contact printing the initial negative image on film. A two-step photoreduction procedure is being investigated for preparing masks if structures finer than 10 μm are called for. A graphic arts process camera is available for photoreductions up to 5X; the 35 mm film-type camera then would be used for the second and final photoreduction.

The film mask combining both the data-input-electrode and phase-shifter patterns was prepared by first trimming then butting or partially lapping the two individual film masks in their appropriate relationship. The two were attached to one another with transparent tape. This technique can be employed for making masks larger than can be photoreduced easily from one master. However, the relatively small-area patterns of immediate interest have been formed on film from one master, complete with alignment marks for photoresist grating formation subsequent to the photolithographic operation. Masks of this design will be used in future work.

Metallization of LiNbO$_3$

A photolithographic process in which the metal pattern is formed by etching the metallized slab through a developed pattern in a photoresist layer has given a higher yield of acceptable samples than has the lift-off technique, in which metallization follows pattern/formation in photoresist. Two problems have been encountered in the lift-off process which are not compensated for by any positive aspects: (1) assurance of a clean LiNbO$_3$ surface after resist pattern development, necessary for good metal adherence, and (2) removal of the unwanted resist after metal deposition without damage to adjacent metallized areas of KiNbO$_3$. These problems were significantly greater with Kodak KTFR negative resist but persist with the positive resists now employed.

Both aluminum and chromium have been used for slab metallization. Chromium adheres better to LiNbO$_3$ than does aluminum, and is more stable in an unprotected environment. However, it is anticipated that ultrasonic bonding of aluminum leads to aluminum pads can be employed (rather than conducting silver paint) as electrode density increases in patterns of the near future and as development extends beyond demonstration devices. An aluminum-over-chromium metallization is envisaged, for good film adherence with ultrasonic bonding capability, which may require a two-step etching technique for pattern delineation.

Aluminum was vapor-deposited onto LiNbO$_3$ using hot-boat evaporation at pressures in the 10^{-5} Torr range. Both tungsten boats and multi-loop tungsten helixes were satisfactory hot surfaces. Film thickness, in the range 500 A

to 1000 A, determined roughly by evaporating a given charge to completion, has not been considered critical but may become more so when ultrasonic bonding is attempted. A crystal thickness monitor can be introduced into the vacuum chamber when need arises.

Chromium was vapor-deposited from a current-carrying tungsten rod electroplated with chromium. A film thickness estimate was made by observing a back-light monitor slide to determine when opacity was achieved. As with aluminization, metal thickness beyond a few hundred Angstroms was not considered important.

Sample cleaning prior to metallization was the same as that described above in the section on Waveguide Formation. Surface cleanliness is a critical factor in promoting good film adherence and has been improved through care in surface visual examination under appropriate illumination and also microscopic examination before pronouncing the surface clean.

Photolithography

The photolithographic process comprises: (1) application of photo-resist to the metallized $LiNbO_3$ surface, (2) baking the photoresist, (3) exposing the photoresist to ultraviolet light through the pattern mask, (4) developing the pattern in the exposed photoresist, (5) removing the exposed metal by chemical etching, and (6) removing the residual resist. Sample cleaning, if any, was limited to a rinse in isopropanol or methanol.

Positive photoresist, in which the UV-exposed regions are soluble in developer, has been used exclusively in preparing the preprocessor patterns. Earlier experience with Kodak KTFR negative resist presented problems in removing the residual, exposed (cross-linked polymer) resist by unheated solvents; the safest procedure was to manually strip the resist softened by immersion in acetone or developer, a tedious and sometimes damaging process. Reuse of samples once put through the photolithographic process seldom was possible because of essentially permanent surface contamination. Furthermore, the Kodak resist patterns were dimensionally unstable in process solutions.

Shipley 1350J positive photoresist was satisfactory for forming data-input electrodes and the phase shifter. Best results were obtained with

dilution by an equal amount of Shipley AZ Thinner. One drawback, or possibly an asset in certain cases, of the Shipley resist was that the highly alkaline Shipley AZ-300 developer etched underlying aluminum. Thus, development and etching could be carried out in one step, if desired, but attack on the aluminum, if undesired, was difficult to prevent. Also, Shipley resist, when thinned to the degree necessary for making resist gratings (described later), exhibited poor adherence to the $LiNbO_3$ surface. This problem led to the use of Polychrome PC-129 resist for grating formation. Since it gave good results in the less demanding application to metal-pattern formation, this resist, diluted by an equal amount of Polychrome Thinner T-10, has been used consistently, largely to keep photoresist operations within the same family of materials.

Photoresist was applied to the slab by a spinner with vacuum holddown chuck (PLAT Engineering Company). The large mass of the slab, relative to a thin silicon wafer for which the spinner was designed required the use of double-stick masking tape on the chuck to secure the slab. Resist was applied by eyedropper or by a hypodermic needle with Millipore filter to the entire surface at rest. Under spinner acceleration and subsequent uniform rotation rate for 30 seconds, the resist was distributed fairly uniformly over most of the surface. The problem of resist buildup at the edges of the slab, and thickness distortion in the corner regions attributed to non-circular slab geometry, has not been solved. Buildup along the edges is a nuisance at best, since the phase-shifter is located, in part, at or close to one edge. The increased resist thickness precludes proper development and etching, necessitating a "cleanup" operation described later.

After photoresist application, the coated slab was placed in an oven at 90 C, with forced air ventilation, for about 20 minutes to bake out resist solvents.

Exposure of the resist-coated sample to UV-light through the mask was made on a Kulicke and Soffa Model 675 Mask Aligner. This equipment is particularly well-adapted to laboratory experimentation where flexibility rather than through-put by assembly line personnel is important. The aligner is designed to accommodate a 1-inch square, or 1.5-inch round, substrate. A retrofit mask holder has been constructed at Battelle to accept slabs up to 3 inches diameter. Both a microscope and UV-light source are mounted on arms which

alternately can be positioned over the mask. An Olympus Model 900 metallo-
graphic microscope shortly will replace the original to provide a choice of
magnifying powers and greater sample illumination. The UV source is a 100-
watt high-pressure mercury-arc-lamp with only a barrel-type light guide. If
future work requires higher lamp intensity and a high degree of light col-
limation or control of angle of incidence, an off-the-shelf illuminator hav-
ing these qualities can be fitted to this aligner. At present, the aligner
makes contact exposures. Although damage to the mask is risked when forced
into contact with the resist-coated substrate, contact exposure usually is
superior to non-contacting projection and proximity printing in terms of
pattern fidelity.

Positioning of the composite film mask (data-input-electrodes and
phase-shifter patterns) on the $LiNbO_3$ slab presents no problem since one edge
of the phase shifter is parallel to, and can be at, the mirror edge. Lateral
positioning of the mask can be approximate within a 2-mm interval about a
reference mark on the sample holder; this can be accomplished by the unaided
eye.

When two masks are used in successive exposures, one for the phase
shifter and the second for the data-input electrodes, mask registration be-
comes more complex. Registration marks can be included when the masks are
made; marks on the first mask to be processed will be transferred to the sub-
strates as copies in the metal over which the second mask can be registered.

A second, somewhat more flexible, method for registration proved
satisfactory and may be reemployed in future work. After the slab was
initially aluminized, a layer of copper (for good contrast to aluminum or
chromium) was vapor-deposited through a mask having an edge parallel to the
mirror edge and positioned at the juncture of the two patterns. Thus the
surface was copper-coated between the mirror edge and a line parallel to it
through this juncture. The phase-shifter pattern was formed first in the
copper-coated region, with the region above totally masked off; alignment was
at the mirror edge. The phase-shifter pattern region then was masked off,
leaving the balance of the copper-coated, as well as the non-copper-coated,
area exposed. The data-input-electrode mask was positioned such that the

92

bottom edge of the pattern was registered along the copper edge at a measurable distance from the phase shifter. Both patterns in the resist were developed simultaneously. On the other hand, the first could have been developed before the second was formed. After development, a copper etchant was used, followed by the aluminum etchant.

Resist exposure to the 100-watt lamp is a function of resist thickness; a typical exposure time is 75 seconds. If too short an exposure time results in areas of incomplete resist removal in the development step, an overall (no mask) exposure of 10 to 15 seconds usually completes exposure without harm to the initially unexposed resist.

Development of the photoresist after exposure usually was complete following a 30-second immersion in the developer appropriate to the resist used followed by a 60-second immersion in distilled or demineralized water. The cycle was repeated in shorter intervals if development was not complete. The sample finally was dried in a warm airstream. No baking was required following development of either Shipley or Polychrome resist.

Microscopic examination of the developed pattern was made before etching to detect any flaws in the pattern. If flaws were found, the resist was removed by an acetone rinse and the procedure repeated, starting with the spun-on coating step.

Etching was by immersion in an etchant appropriate to the metal film. An acidic etchant was used for aluminum, since alkaline etchants dissolved the photoresist in some cases. Following are the etchants found to be satisfactory when used at room temperature:

<u>Aluminum Etchant</u>

Phosphoric acid (concentrated)	800 ml
Acetic acid (concentrated)	50
Nitric acid (concentrated)	10

<u>Chromium Etchant</u>

Ceric ammonium nitrate	530 g
Perchloric acid	100 ml
Deionized or demineralized water	2000 ml

<u>Copper Etchant</u>

Ammonium persulfate or	20 percent solution
Ferric chloride	42° Baume solution

The time required to remove metal to the LiNbO$_3$ surface depends in part, of course, on metal thickness. The process was monitored visually. Typical etch periods were from 1 to 2 minutes. Etching was quenched by a water rinse. None of the etchants above attacked the photoresist.

After etching, the remaining unexposed resist was removed by an acetone rinse. A final surface cleanup usually was required to remove traces of photoresist, etch products, and metal protected by resist not fully removed in the development step because of excessive thickness (along the slab edges, for example). Immersion of the slab in the various processing liquids contributes to surface contamination because of pickup from the surface as the slab is withdrawn. Processing under flowing liquids to minimize this pickup has, at times, caused damage to fine metal structures and therefore has not been continued.

The simplest method for removing metal remnants was to "paint", with a pointed wooden Q-Tip shaft, the desired metal pattern with photoresist diluted with 2 parts of thinner to one of resist. This necessarily was done under low-power magnification. After a 20-minute bake at 90 C, the slab was immersed in the appropriate etchant, then rinsed in water and warm-air dried. An acetone rinse removed the protective resist coating. An alternate method for removal of unwanted metal was to selectively apply small droplets of etchant to such regions then rinse quickly when etching was observed to be complete. Etchant vapors hovering over the slab can damage the pattern; hence, a minimum of etching time is called for. A laminar-airflow environment for this work helps to disperse these vapors.

With the slab surface free of unwanted metal, final cleanup was accomplished by immersion in fresh developer or thinner or the appropriate stripping agents, Polychrome S-50 or Shipley 1112A.

Fabrication of the Metallized
Grating Beam Splitters

The integration of the surface-grating beam splitter with other elements of the preprocessor necessitated the development of specialized fabrication and alignment procedures to ensure satisfactory operation of

94

the overall device. The integrated beam splitter not only is required to diffract with a reasonable efficiency, but it must also direct the diffracted beam along a prescribed path. Thus the integration step places strict requirements on the spatial frequency of the grating, its orientation and its location. To meet these requirements the following fabrication techniques and procedures were developed.

The beam splitters were fabricated by recording holographic gratings in positive photoresist films deposited on the waveguide surface. The recording arrangement is shown schematically in Figure V-3. For good diffraction efficiency it is essential that the grating be developed over the entire grating region down to the $LiNbO_3$ substrate; therefore, it is important that the spatial intensity of the recording beam be very uniform. The spatial filters shown in Figure V-3 improve the uniformity by removing intensity variations due to diffraction from dust particles or imperfections in the optical elements. The collimating lens is then used to pick out only the central and most uniform part of the beam. Another source of nonuniformity of the exposed region is the reflected beam from the back surface of the $LiNbO_3$ slab. The reflected beam interferes with the incoming beams and generates coarse interference fringes causing some areas to be either under-exposed or overexposed. This problem was overcome by coating the back side of the $LiNbO_3$ slab with a high index matching fluid which significantly reduced the reflectivity at that interface. Furthermore, the index matching fluid utilized absorbs at the writing wavelength of 488.0 nm and little or none of the write beam transmitted by the resist film gets scattered or reflected back to the resist layer.

The recording medium is Polychrome PC-129 positive photoresist diluted by three parts thinner to one part resist. It was spun on the wave-guide surface at 3000 rpm resulting in a film thickness of .090 to .120 μm thickness. As noted previously, the writing wavelength was 0.488 μm and the exposure approximately 3500 mj/cm^2. The development procedure consisted of immersion for 20 seconds in Polychrome D-900 developer, a 30-second rinse in distilled water, and air drying.
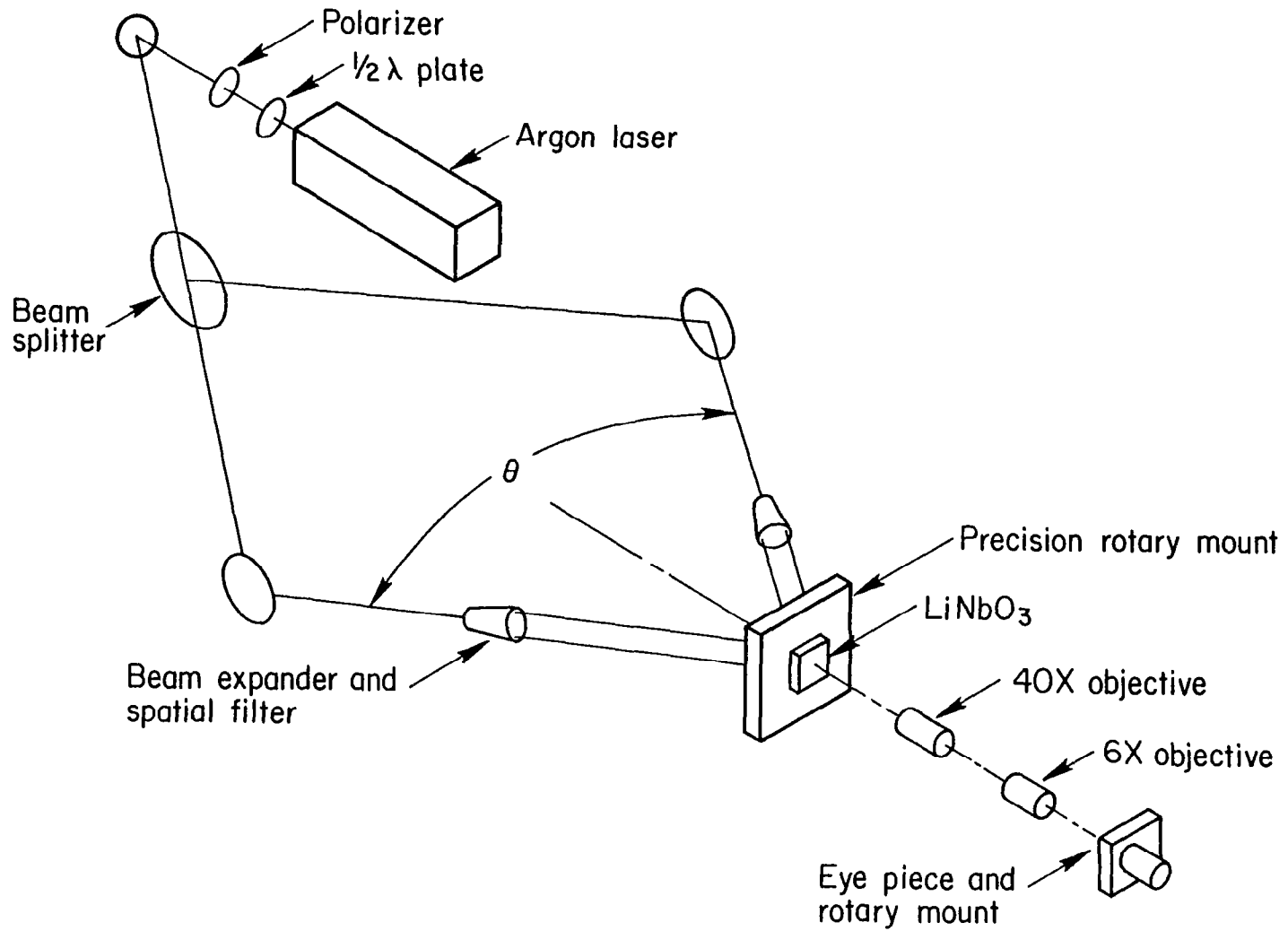
FIGURE V-3. ARRANGEMENT FOR ORIENTING AND EXPOSING
PHOTORESIST BEAM SPLITTER.

The grating frequency and its orientation are especially critical to the success of the integration step. In the present three-channel device, the guided beam is coupled into the waveguide so that it will pass through the channel formed by the phase-shifter electrode structure while the grating beam splitter deflects part of the input beam through the data electrodes. To function properly, the grating beam splitter must be precisely oriented at its Bragg angle, θ_B, with respect to the guided beam and θ_B must be carefully chosen to ensure that the diffracted beam will be directed along the prescribed path through the data electrodes. The angular separation between the diffracted guided beam and the zero-order guided beam ($2\theta_B$) is chosen to be exactly equal to the angular separation between the optical paths through the phase shifter and the data electrodes. To fabricate a resist grating of the appropriate spatial frequency, the angular separation between the two write-beams (θ), as shown in Figure V-3, is set according to the following equation:

$$Sin(\theta/2) = \frac{\lambda_w}{\lambda_g} \, n_g \, sin \, (\theta_B)$$

where

$\theta/2$ = one half the required angular separation of the write beams

λ_w = wavelength of the write beam

λ_g = wavelength in air of the guided beam

and θ_B = Bragg angle for light of λ_g inside the waveguide.

To ensure that the fringe lines of the grating are oriented properly with respect to the optical path through the electrode structure requires the use of a precision rotatable sample holder and the viewing optics shown in Figure V-3. The viewing optics consisted of two microscope objectives, one 40X and the other 6X, and a traveling micrometer eyepiece also mounted in a precision rotary mount. The two objectives, in tandem, are used to observe the interference fringes from the overlapping write beams. The eyepiece is rotated until its reticle is aligned with the interference fringes. With the precision rotary mount, the reticle can be aligned to the fringes with an

accuracy of approximately $\pm$ 0.8 milliradians. The resist-coated LiNbO$_3$ waveguide then is mounted on its precision rotary mount. After removing the 40X objective, the data electrodes are viewed with the 6X objective. The resist-coated sample now is rotated until the edges of the data electrodes are aligned with eyepiece reticle. After that alignment is completed, the LiNbO$_3$ waveguide is rotated in the appropriate direction by an amount equal to the required Bragg angle. The sample now is ready for exposure after coating the backside with the high-index matching fluid. The actual exposed grating region is much larger than necessary. To restrict the physical size of the grating and confine it to the appropriate region on the waveguide surface, a mask is placed over the waveguide and the sample is exposed to a UV lamp prior to development. This step not only positions the grating and restricts its size but also results in removal of the resist over the remainder of the surface area. The removal operation is required for prism coupling and attachment of the electrical leads.

The resist grating is then aluminized through an evaporation mask which exposed only the grating area. Metal-film thickness beyond opacity has not been shown to be critical, therefore an arbitrary thickness of 200-300 A has been used.

The remaining step before attaching coupling prisms to the slab is attachment of leads to the data-input and phase-shifter pads. Air-dry conducting silver paint, applied by an artist's brush, has proven satisfactory for bonding fine copper wires to the aluminum pads. As mentioned earlier in this section, ultrasonic bonding of aluminum wires to the pads is expected to be feasible and can be done at Battelle.

VI. SYSTEM TEST AND OPERATION

IDEAL BEHAVIOR

The goals of the testing program were to
    a) demonstrate the validity of the device principle
    b) get initial assessments of the sensitivity of the
       preprocessor
    c) investigate the sources of noise and other
       operational problems.

In order to assess the test results, a simple mathematical model of an ideal three channel device was used. This model was based upon the assumption of a "segmented plane wave", as shown in Fig. VI-1. The model assumes three channels in each of which there is a uniform phase shift of an idealized plane wave front. There are four guard bands which are of the same width as the data channels and it is assumed that there exists no diffraction or cross talk.

If we first consider the case of two plane waves being recombined by the hologram, we see that the intensity at the detector is

$$I = A_S^2 + A_R^2 + 2A_S A_R \cos(\phi_S - \phi_R) \qquad (VI-1)$$

where A_S, ϕ_S and A_R, ϕ_R are the amplitude and phase of the transmitted signal beam and diffracted reference beam, respectively. Both phases are measured with respect to the same arbitrary reference plane. If we maintain the simplifying assumption of no cross talk then for the three-channel device

$$I = \sum_{j=1}^{2N+1} I_j \qquad (VI-2)$$

where $j = 1,3,5,7$ refer to the guard bands and $j = 2,4,6$ refer to the three signal channels. If all channels have the same illumination and $I_{S_j} = I_{R_j}$, then
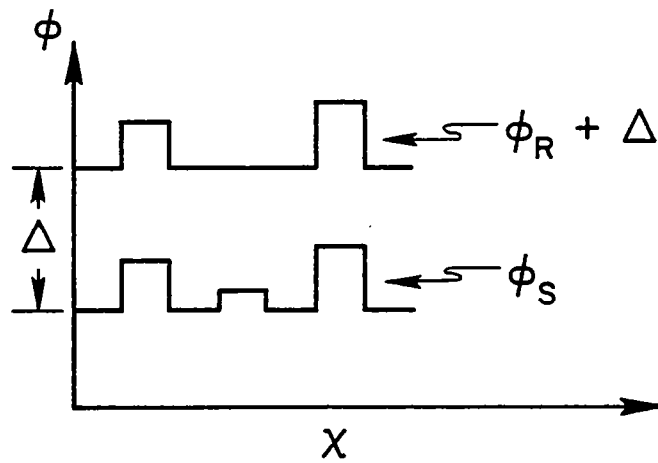
FIGURE VI-1. SEGMENTED PLANE WAVE REPRESENTATION OF
THE PHASES OF THE TRANSMITTED SIGNAL BEAM
ϕ_S AND THE DIFFRACTED REFERENCE BEAM ϕ_R.
THE PHASE Δ IS INTRODUCED BY THE PHASE
SHIFTER.

$$I_j = \frac{I_o}{2(2N+1)} \ [1 + \cos(\delta_j - \Delta)] \tag{VI-3}$$

where $\delta_j = \phi_{S_j} - \phi_{R_j}$, is the phase shift in the jth channel which represents the difference between the holographically recorded reference and the signal. Δ is the uniform phase shift introduced in the reference beam by the phase shifter (see Fig. VI-1).

The response of this idealized system to variations in ϕ_j and Δ can be easily derived from Eq. VI-2. For example, in the case $\delta_j = 0$ the output is simply

$$I = \sum_{j=1}^{2N+1} I_j = \frac{I_o}{2} \ [1 + \cos\Delta] \quad . \tag{VI-4}$$

Therefore as Δ is varied between zero and π, I varies between I_o and zero. The ideal device therefore has an infinite contrast ratio.

The sensitivity S, of the device is the change in output in response to a vanishingly small change in the phase shift in a single channel.

$$S = \frac{\partial I}{\partial \delta_k}\Bigg|_{\delta_k = o} = \frac{I_o}{2(2N+1)} \ \sin\Delta \tag{VI-5}$$

The sensitivity is seen to be a sinusoidal function of Δ which vanishes when the condition for complete subtraction ($\cos\Delta = -1$) is satisfied. The implication of this statement is that, to obtain a finite output it is necessary to operate off the null or to have a large enough value of δ_R that Eq. (VI-5) is replaced by an expression using finite differences.

TEST RESULTS

The system was tested using the arrangement shown in Fig. VI-2. Approximately 100 μW of 0.6328 μm radiation from a He-Ne laser was prism coupled into the waveguide. The output was detected by a PIN diode and was directly observed on an oscilloscope. The device response as a function of Δ
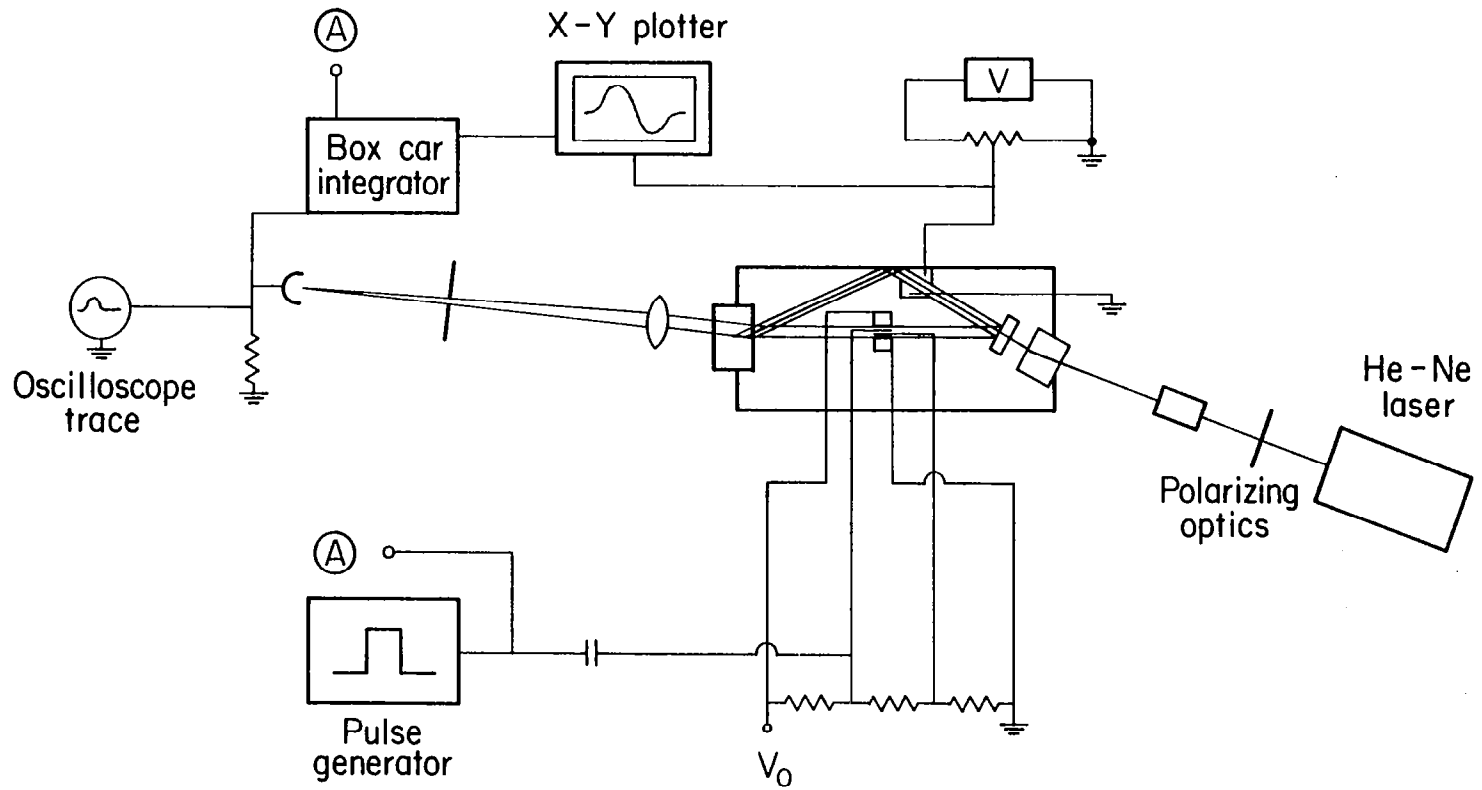
FIGURE VI-2.   SCHEMATIC OF ARRANGEMENT FOR TESTING THE
3-CHANNEL LABORATORY MODEL.

was recorded using an x-y plotter. The PIN diode was connected directly to the plotter for d.c. measurements. The boxcar integrator was used in recording the response to a pulse on a single channel.

In Fig. VI-3, the d.c. output is shown as a function of phase shifter voltage V_ϕ, where the phase shift Δ is proportional to V_ϕ. Although the curve approximates a cosinusoidal shape, it departs from the ideal cosine form predicted in Eq. IV-4 in two obvious ways. The contrast ratio is only 8, as opposed to the predicted infinite value, and the curve is not symmetric. The poor contrast ratio is believed to have resulted from the fact that these data were taken on a device which had a rather inefficient beam splitter ($\sim$ 3%). It has been our experience that unless the efficiency exceeds 10%, the hologram which is subsequently formed does not have a high enough diffraction efficiency to produce a good contrast ratio. Using an external beam splitter we have written waveguide holograms which produced contrast ratios as high as 160:1. In an integrated device containing all the components but the data input electrodes, a contrast of 80:1 was observed.

The assymetry of the curve is believed to result from the fact that the application of a voltage across the phase shifter induces enough of an index change to perturb the waveguide. Thus as the voltage is increased, an electrooptic enhancement (or destruction) of the planar waveguide occurs which results in more (or less) light in the channel and therefore in the reference beam at the hologram. This problem should not occur if channel waveguides are used.

In Figure IV-4 the response to a pulse on a single channel is displayed as a function of V_ϕ. As expected the response is proportional to the derivative of the curve in Fig. IV-3. At maximum sensitivity ($\Delta = \pi/2$) a pulse which produced a 1° phase shift on a single signal channel resulted in a S/N ~2 as observed directly on the oscilloscope.

The actual signal-to-noise ratio of a specific device will be a function of both the detector noise and the quality of the null which can be obtained. Obviously if the system is run at $\Delta = \pi$ the signal will be small, but if a good null is achieved, the noise will be minimal. For the system tested the S/N maximum occurred at $\Delta = \pi/2$ since a good null was not obtained.
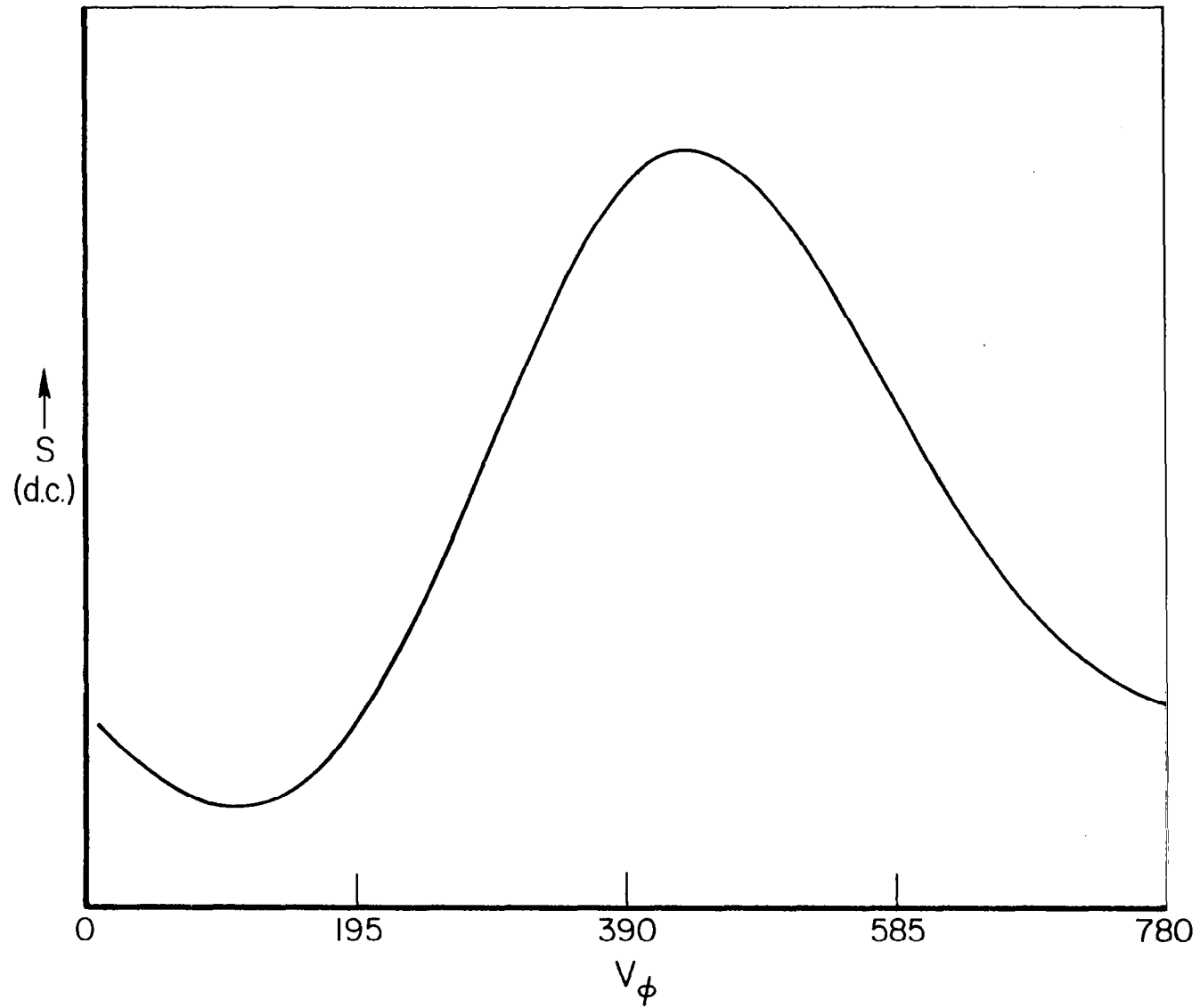
FIGURE VI-3. D.C. OUTPUT OF THE PREPROCESSOR ($\delta_j = 0$)
AS A FUNCTION OF PHASE SHIFTER VOLTAGE.
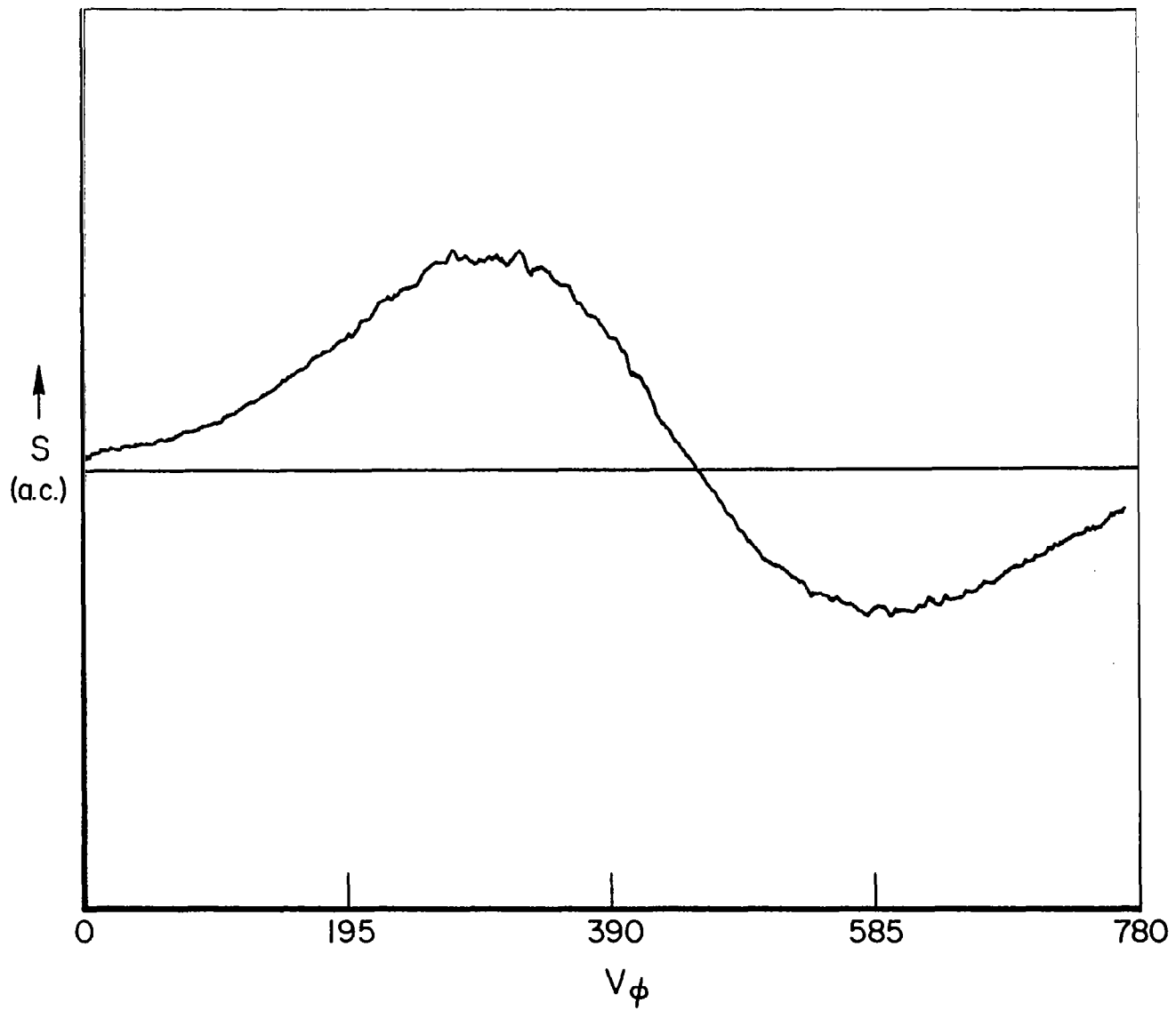
FIGURE VI-4. A.C. OUTPUT OF PROCESSOR IN RESPONSE TO PULSE ON
A SINGLE SIGNAL CHANNEL AS A FUNCTION OF PHASE
SHIFTER VOLTAGE.

An additional problem which had to be overcome before the device could be tested was the lack of stability of the hologram. After the reference hologram is written, the introduction of a phase shift produces conditions under which a new hologram is being written which erases the initial reference hologram. In making the measurements this effect was overcome by using a much lower laser power to take the data than was used to write the hologram. The useful life of the hologram could be extended to several minutes. The effect is aggravated at the null ($\Delta = \pi$) since the new grating is exactly out of phase with the original reference hologram.

Several approaches can be taken to overcome the stability problem. In a single purpose device a fixed hologram (thermally fixing or two-photon writing) can be used. Alternatively, reduced laser read powers and periodic refreshing of the reference hologram can be employed. In the identification mode, a permanent grating beam combiner can be used, while the self subtraction mode actually depends upon continual hologram updating and therefore requires continued response of the material to the laser beams.

Estimates of the preprocessor response time band upon the electrode capacitance and a 50 Ω driver impedance indicate that the rise time in response to a step input pulse should be on the order of several nanoseconds. The device was tested and was found to follow pulses with 1 μsec rise time with no apparent pulse broadening. The preprocessor has not yet been tested with sharper pulses.

VII. FUTURE PREPROCESSOR DEVELOPMENT

The successful testing of the three-channel laboratory model of the preprocessor in the recognition mode served to establish the validity of the operating principles and to demonstrate the existence of suitable fabrication techniques. Logical next steps would be to increase the number of channels and to demonstrate the identification mode. Some preliminary work has been done which suggest the feasibility of combining these two steps in the fabrication of a 16-channel identification-mode device.

In the identification mode, the signal voltages may be though of as impressing a distortion upon a plane wave while the reference voltages try to restore the original plane wave. Since there is no need to store a reference wavefront it is possible to replace the hologram with a simple beam splitter for recombining the signal and reference beams. Of course, the hologram's ability to account for the perturbations introduced by the electrode structure and the ability to efficiently diffract divergent beams will be sacrificed. The advantage here is that a passive grating beam splitter is more stable than the hologram. The tradeoffs will have to be evaluated as part of future studies.

A second design decision, one which has been considered as part of the present program, concerns the choice between building the preprocessor in a strictly planar-waveguide configuration or in a hybrid planar-channel waveguide configuration. The former makes use of electrode structures optimized to perform holographic subtraction with reduced cross talk and minimal deleterious effects resulting from diffraction. The latter makes use of simpler electrode structures possible with two-dimensional optical confinement, and has the potential advantages of lower voltage, better light utilization and less cross talk at the expense of increased waveguide fabrication complexity.

In this section we discuss some preliminary electrode designs that appear to be appropriate for a planar-waveguide configuration. As a first step in generating similar designs for a planar-channel waveguide configuration, we have addressed the impact of beam channelization on signal-to-noise performance of the preprocessor. Our calculations presented here, show the expected result of better performance for the channelized geometry. However, the

differences are not large, and theoretical performance limits for either geometry are well in excess of what we have observed to date. We conclude that both planar waveguide and hybrid planar-channel waveguide configurations are viable candidates for use with a 16-channel preprocessor employed in the identification mode.

POTENTIAL ELECTRODE CONFIGURATIONS FOR USE WITH A PLANAR WAVEGUIDE PROCESSOR CONFIGURATION

The accompanying Figures VII-1 through VII-4 list electrode configurations and some advantages and disadvantages of each configuration. These four types seem to encompass just about the full range of possibilities for the processor using a planar waveguide. Electrodes that use slanted connections in the region where the signal beam will be have been omitted, since such slanted connections may give rise to signal-dependent beam deflections.

It is currently envisioned that one electrode unit (one gap plus one guard region) will likely have dimensions in the ranges indicated below (refer to Fig. VII-1 for symbol definitions):

$$10 \ \mu m < a < 30 \ \mu m$$
$$50 \ \mu m < w < 300 \ \mu m$$
$$1 \ mm < L < 3 \ mm$$

Figure VII-1 shows a simple electrode-gap structure like that used for the experiments performed during this program. For use with a processor operating in the identification mode, twin electrode structures are required, one placed close behind the other in the direction of beam propagation. Proximity is necessary to minimize the effects of diffraction spreading.

We summarize the advantages and disadvantages of the design of Fig. VII-1 as follows:

Advantages:

1) Simple fabrication
2) Clean gap definition
3) Uninterrupted guard regions
4) Simple, nonperturbing connections.
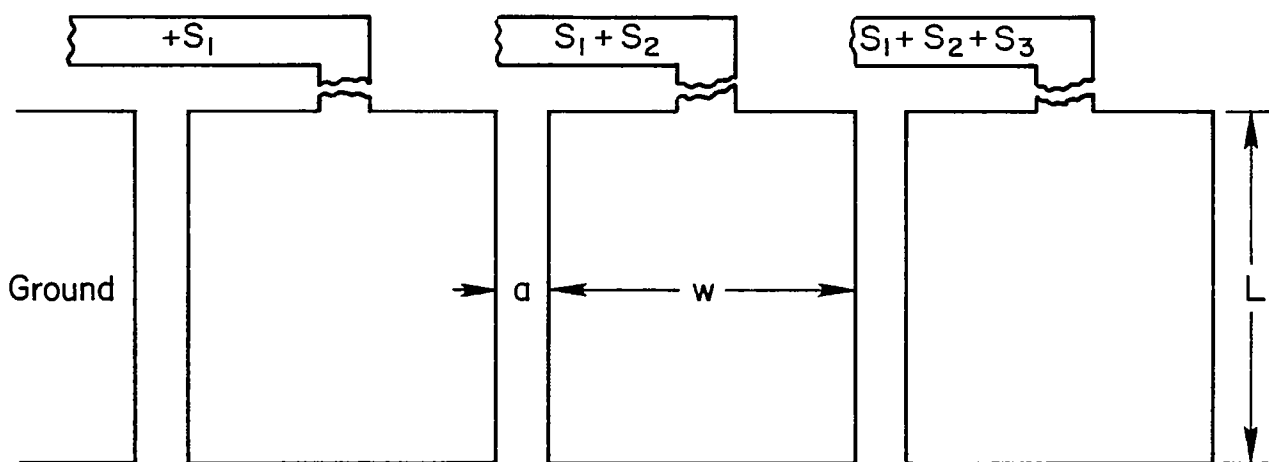   (if $w \gg a$)
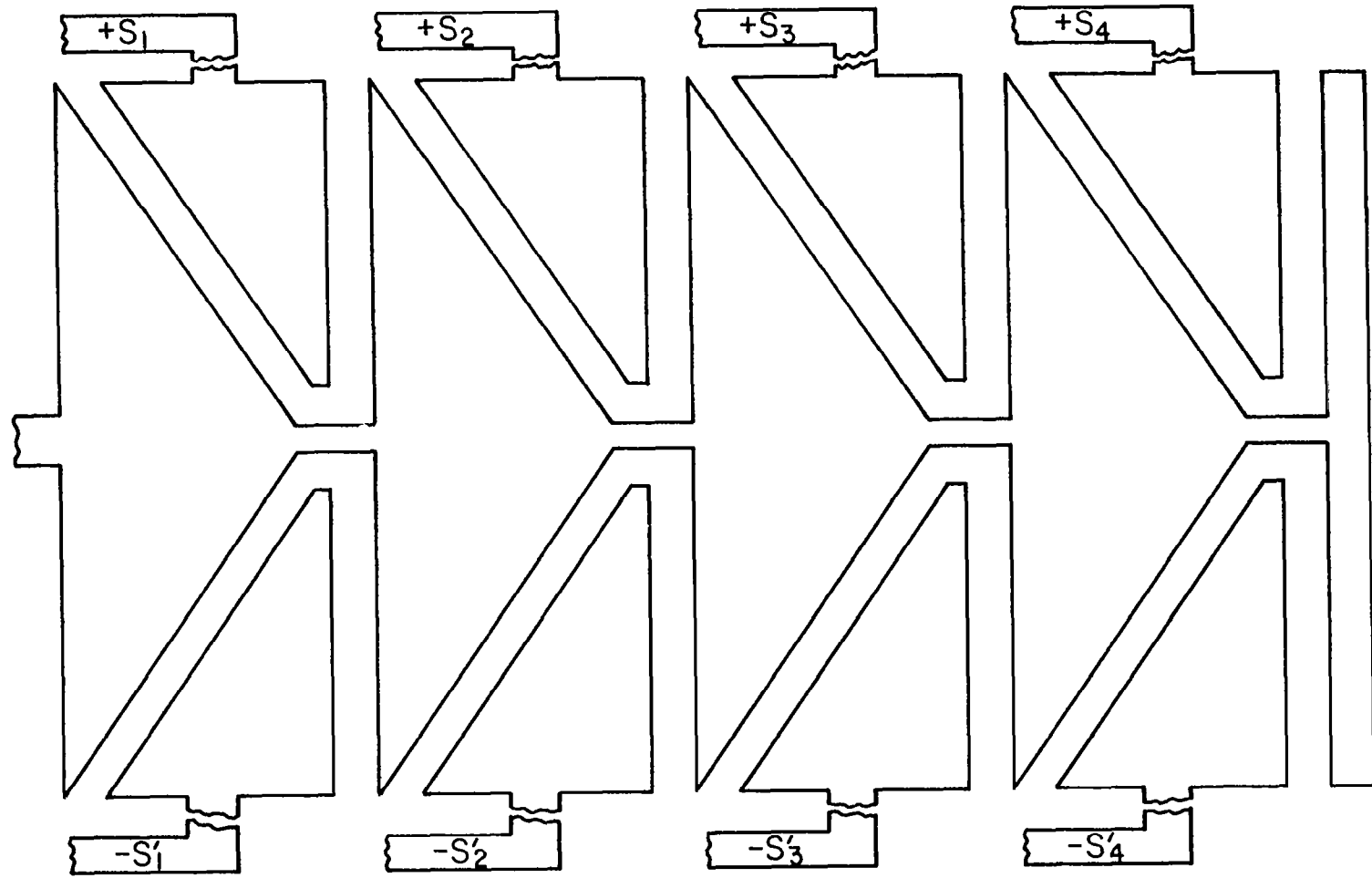
FIGURE VII-1.   BASIC SIMPLE GAP STRUCTURE.

FIGURE VII-2. EXTENDED GROUND ELECTRODE-REVERSED-S' TYPE.
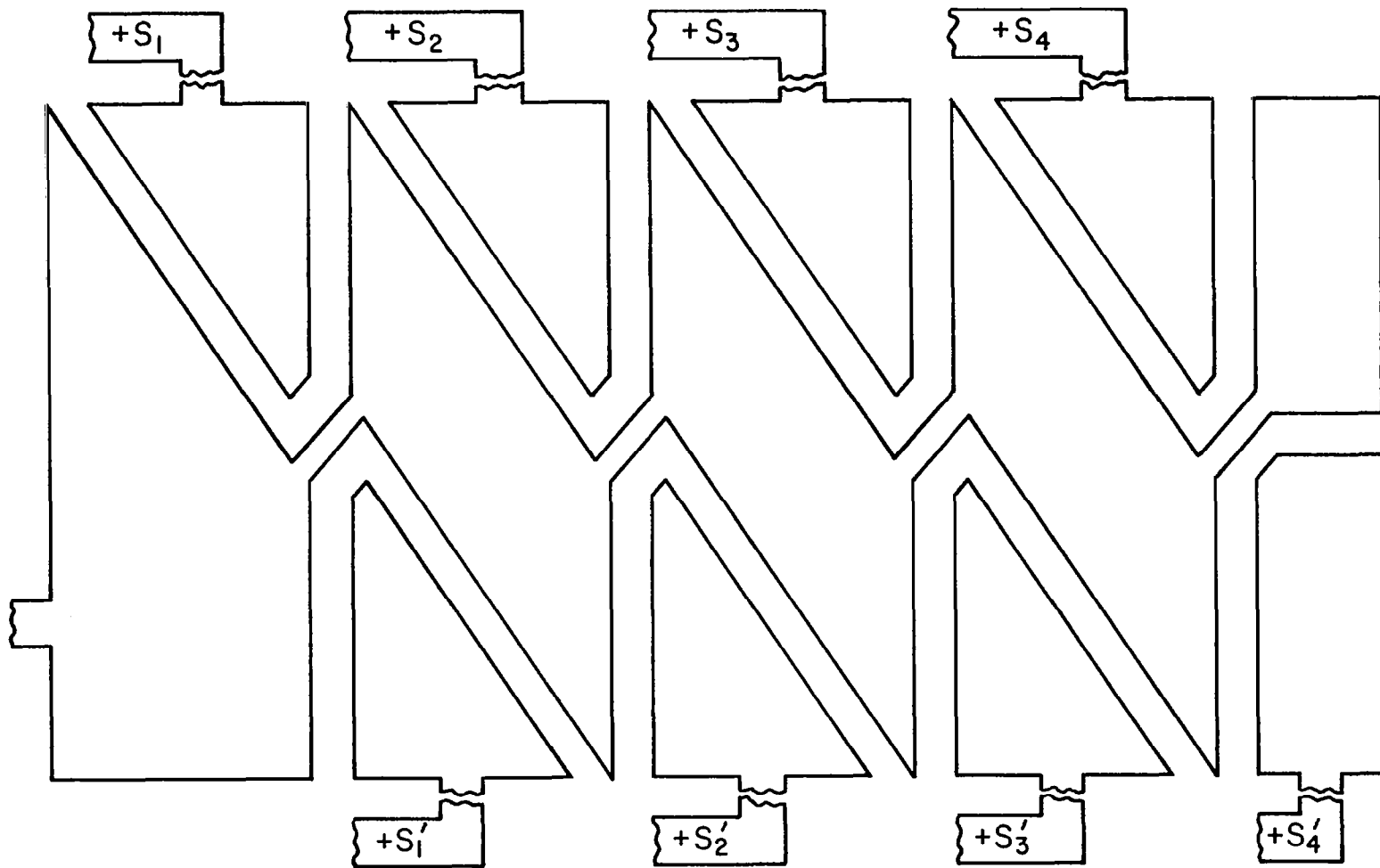
FIGURE VII-3.   EXTENDED GROUND ELECTRODE-NONREVERSED-S' TYPE.
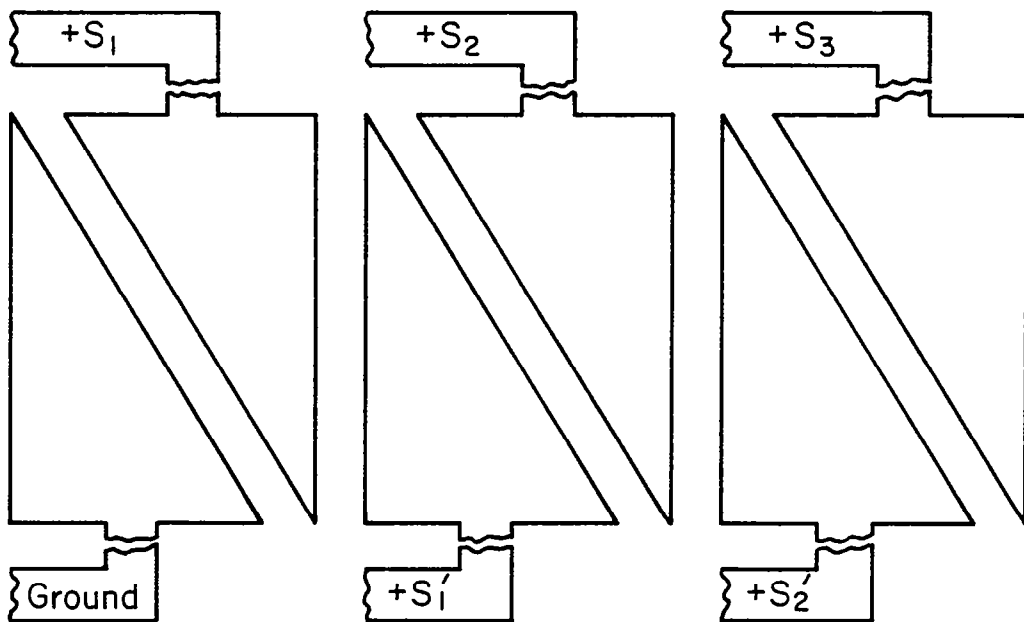
FIGURE VII-4.   BASIC SPLIT ELECTRODE DESIGN.

Disadvantages:

       1)   Requires use of summed signal voltages.

Figures VII-2 and VII-3 show electrode designs specifically intended for use with a preprocessor operating in the identification mode. As in the simple gap structure, information is imparted to the beam as it propagates through the narrow vertical channels. However, the introduction of an extended ground electrode in Figs. VII-2 and VII-3 makes the channel-by-channel summation of signal voltages unnecessary.

Advantages and disadvantages of the Extended Ground Electrode-Reversed-S' type structure of Fig. VII-2 are summarized as follows:

      Advantages:

         1)   Fairly clean gap definition
         2)   Simple, nonperturbative signal connections
         3)   Close-spacing of two sets possible
         4)   Distributed interruption of guard region minimizes unwanted phase shifts.

      Disadvantages:

         1)   Interrupted guard region increases noise possibilities
         2)   Reduced gap definition
         3)   Fringing fields at points of signal electrodes increases noise possibilities
         4)   Requires inverter for S' voltages
         5)   Complex structure.

For the "Extended Ground Electrode-Nonreversed-S' Type structure of Fig. VII-3, we have the following tradeoffs:

      Advantages:

         1)   Simple, nonperturbative signal connections
         2)   Close spacing of two sets possible
         3)   Distributed interruption of guard region minimizes unwanted phase shifts
         4)   No inverter needed for S' voltages.

      Disadvantages:

         1)   Interrupted guard region increases noise possibilities
         2)   Poor gap definition
         3)   Fringing fields at points of signal electrodes increases noise possibilities
         4)   Very complex structure.

The structure shown in Fig. VII-4 is currently the preferred configuration, since its eliminates the need for a distributed ground, for inverted S' voltages, and for channel-by-channel summation of voltages, all problems associated with the designs of Figs. VII-1, 2, and 3. A complete listing of the advantages and disadvantages of the Basic Split Electrode Design of Fig. VII-4 is presented here:

Advantages:

1) Fairly simple to fabricate
2) Clean gap definition
3) Can be operated in "groundless" mode so twin electrode structures are not needed
4) Simple nonperturbing connections (if $w \gg a$)
5) Distributed interruption of guard region minimizes unwanted phase shifts.

Disadvantages:

1) Interrupted guard regions - produce phase shift that is signal-dependent. (so: more noise; variable noise)

All electrode designs presented here have a potential difficulty in that information is imposed on only a fraction $a/(w+a)$ of the incident beam, where $a/(w+a)$ is the ratio of the gap width to the total channel width. Since $a \simeq 10$ μm and $w \simeq 90$ μm for a 16-channel processor using a 1.6 mm wide optical beam, about 90% of the beam will have the capability for contributing to the system noise without having the capability for contributing to the system signal. The noise can be reduced by making the electrodes lossy to waveguided light, however this causes potentially useful light to be discarded. Channel waveguides offer the attractive possibility of removing light from the guard regions and distributing it in the gap regions, thus providing more efficient electrooptic transfer of information. However, it is not certain that this sophistication will be required for a useful system.

As part of a continuing effort to evaluate the desirability of employing a channel-waveguide geometry, we have calculated the optical phase shifts associated with light passing through the guard regions of the split-electrode design of Fig. VII-4, and found that it could, under certain circumstance, approximate those associated with the gap region. The implication is that one

114

would accomplish electrooptic information transfer in the guard region, so that the fraction of the beam that carries information becomes $w/(w+a)$. Smaller values of w, say in the range 30-50 μm, will be required. Still the fraction of the beam that carries information is increased to about 80% without the need for channelization. A further advantage is that light which leaves the narrow gap region diffracts rapidly. A grating beam combiner, with its narrow acceptance angle, could possibly prevent this light from reaching the detector and thus contributing to system noise.

CHANNELIZED VS PLANAR GEOMETRY: SIGNAL-TO-NOISE CALCULATIONS

The above mentioned tradeoffs between channelized-versus-planar-waveguide geometries can be put into a more concerete form by actually calculating the signal-to-noise performance associated with each. For simplicity, we calculate the noise associated with a one-channel processor and multiply it by N, the total number of channels. Our results are presented in terms of ϕ_{min}, the minimum phase shift in any one of N channels which results in S/N = 1.

The relevant part of a possible hybrid system using channel waveguides is shown schematically in Fig. VII-5. The light in an information channel is initially confined to a channel waveguide of width a. The channel terminates in a planar waveguide region, and the signal beam spreads by diffraction on its way to the hologram. The reference beam is initially confined to a narrower channel of width b, so that it fills the entire hologram region and is seen by the hologram as approximately a plane wave.

Owing to the fact that the emerging reference beam is not a true plane wave, the holograpically reconstructed wavefront of the signal beam is distorted. There is no phase shift of the reference beam that can generate a holographically subtracted null in the presence of this distortion. The residual noise output is equivalent to the signal that would be generated by shifting the phase of any one information channel by an amount

$$\phi_{min} \simeq (0.79) \sqrt{N} \left[\frac{r_o b}{r_o' a} \right]^2 \quad , \tag{VII-1}$$

Electrodes

Signal
Channel

a

Hologram

r_o

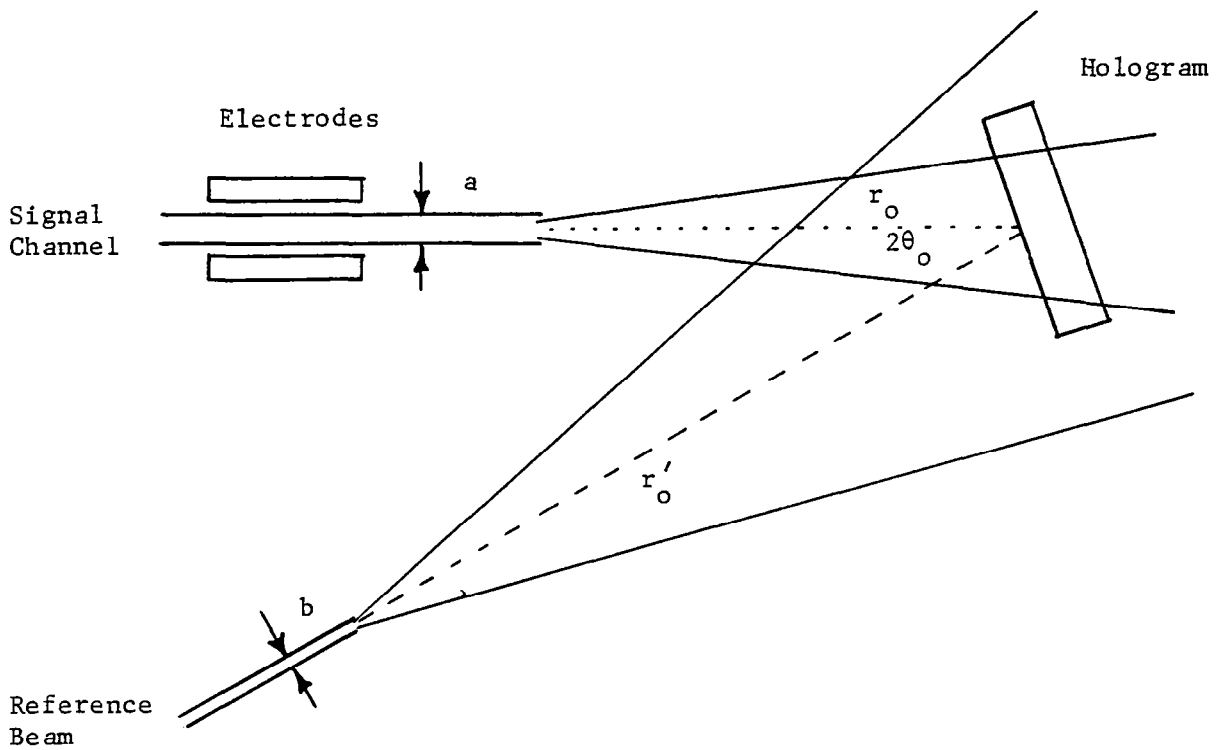$2\theta_o$

r_o'

b

Reference
Beam

FIGURE VII-5. HYBRID PLANAR-CHANNEL WAVEGUIDE CONFIGURATION.
ONLY ONE SIGNAL CHANNEL IS ILLUSTRATED FOR
SIMPLICITY.

where N is the number of information channels and $r_o(r_o')$ is the channel-hologram separation for the signal (reference) beam. Taking N = 16, r_o = 1 cm r_o' = 2 cm, a = 20 µm, and b = 4 µm, $\phi_{min} \simeq 2.0°$. The voltage change required to generate this phase shift can be as low as 0.1 V. Since this small voltage change is considered to be applied to only one of sixteen channels, the performance level indicated by this calculation is quite high. However, implementation of the system requires the fabrication of seventeen channel waveguides, with all but the reference beam channel in close, parallel registration.

This complexity in fabrication points to the desirability of a planar waveguide geometry in which information is carried by a relatively broad (several millimeters wide) unconfined signal beam. Figure VII-6 shows a device in which a broad signal beam passes beneath two electrodes separated by a narrow gap. Phase information is electrooptically introduced on light in the gap region, which diffracts on its way to the hologram. The problem addressed is that all of the signal beam can contribute to noise, while only that part of the signal beam in or near the gap region can carry information.

The analytical situation is complicated by the presence of three distinct beams in the hologram region. The background and information carrying components of the signal beam strongly interfere with each other to produce a hologram having a spatially modulated diffraction efficiency. Noise is found to be reduced as the overall signal beam intensity I_S is decreased relative to the reference beam intensity I_R. We have found

$$\phi_{min} \simeq 2 \sqrt{N} (I_S/I_R) \sin\psi/2 \quad , \qquad \text{(VII-2)}$$

where ψ is the phase shift associated with propagation beneath the metal electrode overlay; that is,

$$\psi = (2\pi/\lambda_o)\Delta n \, L, \qquad \text{(VII-3)}$$

where Δn is the index change associated with the electrode material, and L is the electrode length. In connection with an Air Force sponsored program to develop metal corrector-plate overlays for reducing aberrations in geodesic
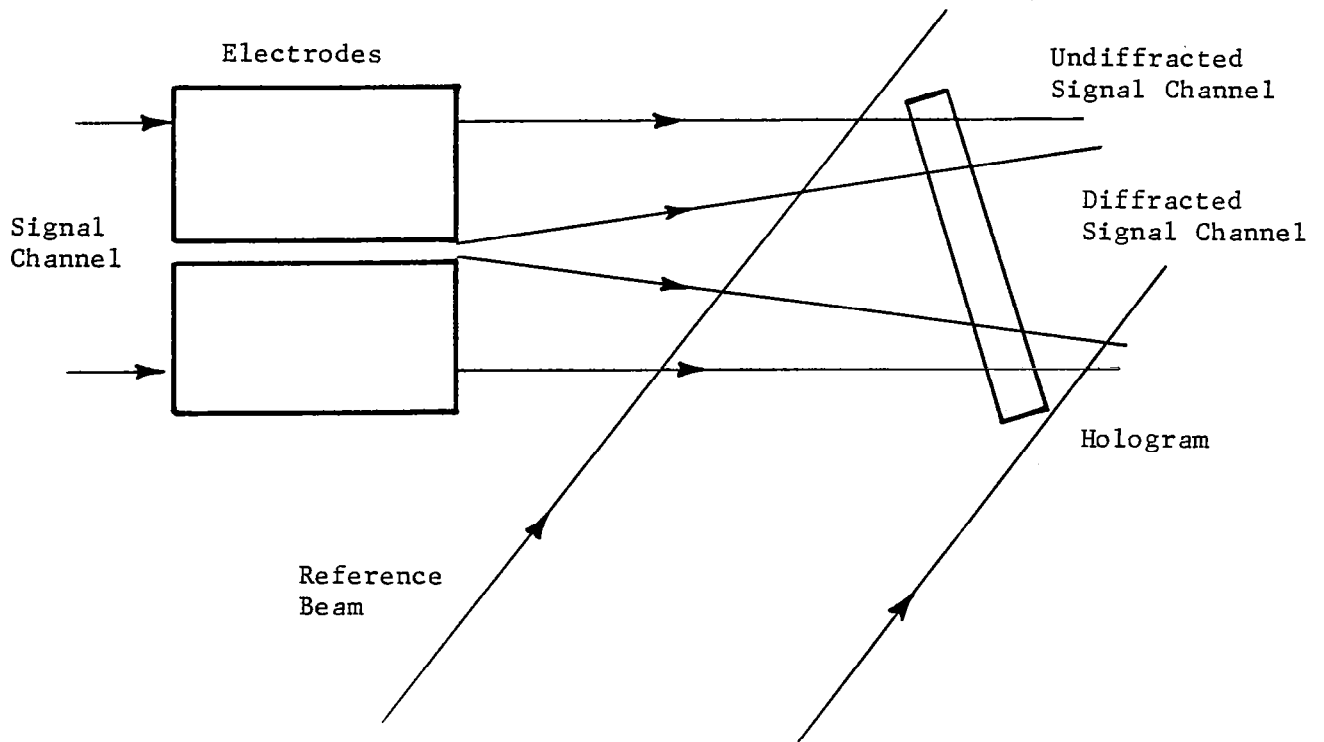
Electrodes

Undiffracted
Signal Channel

Signal
Channel

Diffracted
Signal Channel

Hologram

Reference
Beam

FIGURE VII-6.   WAVEGUIDE CONFIGURATION EMPLOYING AN UNCONFINED
SIGNAL BEAM.   ONLY ONE SIGNAL CHANNEL IS ILLUS-
TRATED FOR SIMPLICITY.

waveguide lenses[19] we have measured $\Delta n \simeq 3 \times 10^{-4}$ in $LiNbO_3:Ti$ waveguides. The condition for optimum processor sensitivity, $\psi = 2\pi m$, where m is in integer, is therefore achieved for electrode lengths that are integral multiples of L = 2.11 mm. Assuming a measurement error of 3%, $\psi = 2\pi(1 \pm 0.03)$ for L = 2.11 mm the minimum detectable phase change is

$$\phi_{min} = 0.18\sqrt{N} \ (I_S/I_R). \qquad\qquad (VII-4)$$

For $I_S = I_R/10$ and N = 16, we find $\phi_{min} \simeq 4.1°$, again for a unity signal-to-noise ratio. This value is only twice that achieved for the channel-waveguide configuration. It shows that, with care in system design, there need be no inherent disadvantage in employing a planar waveguide processor configuration. For this reason, it appears that the optimum configuration for a 16-channel processor will be determined by considerations other than predicted signal-to-noise characteristics.

CHANNEL-WAVEGUIDE FABRICATION METHODS

In the event that a hybrid planar-channel waveguide geometry is chosen for the preprocessor, a number of options are available for fabricating the requisite channel waveguides in $LiNbO_3$ substrates. Among these are the following:

1) In-diffusion of metal, usually titanium,[27,28] or oxide[29] from a narrow stripe source deposited on the crystal surface. The diffusion[30,31] of Nb metal into $LiTaO_3$ may also be included in this category.

2) Formation of a physical ridge[32,33] by ion-milling on sputter-etching a previously diffused planar guide.

3) Out-diffusion of Li from a thin stripe,[34] diffusion from the remainder of the surface being inhibited by an overlay of inert oxide.

4) Stripline-confinement via placing a thin stripe of high-index dielectric material atop the waveguide.[35]

5) Stripline-confinement via metallizing the waveguide surface away from the channel region.[36]

6) Ion-implantation to reduce the index throughout the crystal except in the waveguide region.[37]

Of these techniques, the indiffusion of titanium has been by far the most widely applied. This is because it was discovered quite early that low-loss waveguides supporting only a few transverse modes (only one, if desired) could be made rather straightforwardly in this way. Furthermore, with proper experimental technique the cyrstal surface is left much as it was before the waveguide was formed; so it is easy to couple to the guide using prism or grating couplers and to apply electrooptic modulator electrodes whereever desired. The procedure does suffer, though, from a vexing amount of variability from sample to sample in the lateral (surface) diffusion coefficient of the titanium. This problem has been studied intensively for some time now,[38,39,40] but still cannot be said to have been solved.

The ion-milling method is capable of producing waveguides with all the desirable properties described above, and has been the second most popular method of channel guide formation. It might have received more attention than it has if the required equipment were more widely available and if early work had not indicated possible difficulties in producing steep-side-wall ridges. We have made some preliminary experiments on ion-milling ridge waveguides in out-diffused layers; these experiments are described in Appendix E.

The other methods listed above for production of channel waveguides have been much less used, for the following reasons:

1) they are somewhat more complicated to carry out;

2) they frequently lead to more complicated device designs; and

3) they have been less well known than the in-diffusion or ion milling methods since they are relatively recent or have not been so widely publicized.

REFERENCES

1. D. von der Linde and A. M. Glass, "Photorefractive Effects for Reversible Holographic Storage of Information", Appl. Phys., $\underline{8}$ 85-100 (1975).

2. D. L. Staebler and J. J. Amodei, "Coupled-Wave Analysis of Holographic Storage in $LiNbO_3$", J. Appl. Phys., $\underline{43}$ (3), 1042-1049 (1972).

3. D. M. Kim, R. R. Shah, T. A. Rabson, and F. K. Tittel, "Nonlinear Dynamic Theory for Photorefractive Phase Hologram Formation", Appl. Phys. Lett., $\underline{28}$ (6), 338-340 (1976).

4. W. Phillips and D. L. Staebler, "Control of the Fe^{2+} Concentration in Iron-Doped Lithium Niobate", J. Electron. Mater., $\underline{3}$ (2), 601-617 (1974).

5. H. Kurz, "Wavelength Dependence of the Photorefractive Process in $LiNbO_3$", Ferroelectrics, $\underline{8}$, 437-439 (1974).

6. D. von der Linde, O. F. Schirmer, and H. Kurz, "Intrinsic Photorefractive Effect of $LiNbO_3$", Appl. Phys., $\underline{15}$, 153-156 (1978).

7. A. M. Glass, D. von der Linde, and T. J. Negran, "High Voltage Bulk Photovoltaic Effect and the Photorefractive Process in $LiNbO_3$", Appl. Phys. Lett., $\underline{25}$ (4), (1974).

8. D. W. Vahey, "A Nonlinear Coupled-Wave Theory of Holographic Storage in Ferroelectric Materials", J. Appl. Phys., $\underline{46}$ (8), 3510-3515 (1975).

9. D. L. Staebler and J. J. Amodei, "Thermally Fixed Holograms in $LiNbO_3$", Ferroelectrics, $\underline{3}$, 107-113 (1972).

10. W. J. Burke, W. Phillips, D. L. Staebler, and B. F. Williams, "Materials for Phase Holographic Storage", Final Contractor Report on Contract N00019-74-C-0312, Naval Air Systems Command, April 1975.

11. D. von der Linde, A. M. Glass, K. F. Rodgers, "Multiphoton Photorefractive Processes of Optical Storage in $LiNbO_3$", Appl. Phys. Lett., $\underline{25}$, 155-157 (1974).

12. C. M. Verber, N. F. Hartman, and A. M. Glass, "Formation of Integrated Optical Components by Multiphoton Photorefractive Processes", Appl. Phys. Lett., $\underline{30}$ (6), 272-273 (1977).

13. R. Magnusson and T. K. Gaylord, "Use of Dynamic Theory to Describe Experimental Results from Volume Holography", J. Appl. Phys., $\underline{47}$ (1), 190-199 (1976).

14. L. Young, W.K.Y. Wong, M.L.W. Thewalt, and W. D. Cornish, "Theory of Formation of Phase Holograms in Lithium Niobate", Appl. Phys. Lett., $\underline{24}$ (6), 264-265 (1974).

15. P. K. Tien, "Light Waves in Thin Films and Integrated Optics", Appl. Opt., $\underline{10}$ (11), 2395-2413 (1971).

16. "Optical Waveguide Scattering Reduction", Air Force Avionics Lab. Contract No. F33615-78-C-1426.

17. H. Kogelnick, "Coupled-Wave Theory for Thick Hologram Gratings", Bell Syst. Tech. J., $\underline{48}$ (9), 2909-2947 (1969).

18. G. R. Hocker and W. K. Burns, "Modes in Diffused Waveguides of Arbitrary Index Profile", IEEE J. Quant. Elec., $\underline{QE-11}$, 270-276 (1975).

19. David W. Vahey, "Optical Waveguide Geodesic Lenses", Air Force Avionics Lab Contract No. F33615-77-C-1153, Report No. AFAL-TR-78-85, (1978). See also Appendix C.

20. C. M. Verber, D. W. Vahey, V. E. Wood, R. P. Kenan, and N. F. Hartman, "Feasibility Investigation of Integrated Optics Fourier Transform Devices", NASA CR-2869, 1977.

21. E. F. Kuester and D. C. Chang, "Surface-Wave Radiation Loss from Curved Dielectric Slabs and Fibers", IEEE J. Quantum Electronics, $\underline{QE-11}$ (11), 903-907 (1975).

22. D. Marcuse, "Curvature Loss Formula for Optical Fibers", J. Opt. Soc. Am., $\underline{66}$ (3), 216-220 (1976).

23. T. R. Ranganath and S. Wang, "Suppression of Li_2O Out-Diffusion from Ti-Diffused $LiNbO_3$ Optical Waveguides", Appl. Phys. Lett., $\underline{30}$ (8), 376-379 (1977).

24. B. U. Chen and A. C. Pastor, "Elimination of Li_2O Out-Diffusion Waveguide in $LiNbO_3$ abd $LiTaO_3$", Appl. Phys. Lett., $\underline{30}$ (11), 570-571 (1977).

25. K. Aiki, et al, "Transverse Mode Stabilized Al_xGa_{1-x} As Injection Lasers with Channeled-Substrate-Planar Structure", IEEE J. Quant. Elec., $\underline{QE-14}$, 89 (1978).

26. M. Nakamura, K. Aiki, and J. Umeda, "cw Operation of Distributed-Feedback GaAs-GaAlAs Diode Lasers at Temperatures up to 300 K", Appl. Phys. Lett., $\underline{27}$ (7), 403-405 (1975).

27. M. Papuchon, et al, "Electrically Switched Optical Directional Coupler, COBRA:, Appl. Phys. Lett., $\underline{27}$ (5), 289-291 (1975).

28. I. P. Kaminow, L. W. Stulz and E. H. Turner, Efficient Strip-Waveguide Modulator", 27 (10), 555-557 (1975).

29. J. Noda, N. Uchida, S. Saito, T. Saku, and M. Minakata, "Electro-optic Amplitude Modulation Using Three-Dimensional $LiNbO_3$ Waveguide Fabricated by TiO_2 Diffusion", Appl. Phys. Lett., 27 (1), 19-21 (1975).

30. R. D. Standley and V. Ramaswamy, "Nb-Diffused $LiTaO_3$ Optical Waveguides: Planar and Embedded Strip Guides", Appl. Phys. Lett., 25 (12), 711-713 (1974).

31. J. M. Hammer, W. Phillips, and C. C. Neil, "Stripe Guides of $LiNb_xTa_{1-x}O_3$: Physical and Optical Characterization for Accurate Mode Control", Digest of Technical Papers, OSA/IEEE Topical Meeting on Integrated and Guided-Wave Optics, Salt Lake City, Jan. 16, 1978, pp WA21-WA2-4.

32. J. Noda, N. Uchida, M. Minakata, et al, "Electro-optic Intensity Modulation in $LiTaO_3$ Ridge Waveguide", Appl. Phys. Lett., 26 (6), 298-300 (1975).

33. L. P. Kaminow, V. Ramaswamy, R. V. Schmidt and E. H. Turner, "Lithium Niobate Ridge Waveguide Modulator", Appl. Phys. Lett., 24 622-624 (1974).

34. L. Goldberg and S. H. Lee, "Optically Activated Switch/Modulator Using a Photoconductor and Two Channel Waveguides", Radio Science, 12 (4), 537-542 (1977).

35. J. Noda, S. Zembutsu, S. Fukunishi, and N. Uchida, "Stri-Loaded Waveguide Formed in a Graded-Index $LiNbO_3$ Planar Waveguide", Appl Opt., 17 (12), 1953-1958 (1978).

36. H. Hayashi and Y. Fujii, "An Efficient Acousto-optic TE $\overset{\rightarrow}{\leftarrow}$ TM Mode Converter Utilizing a Doubly Confined Optical and Acoustic Waveguide Structure", J. Appl. Phys., 49 (8), 4534-4539 (1978).

37. G. L. Destefanis, P. D. Townsend, and J. P. Gailliard, "Optical Waveguides in $LiNbO_3$ Formed by Ion Implantation of Helium", Appl. Phys. Lett., 32 (5), 293-294 (1978).

38. M. Fukuma, J. Noda, and H. Iwasaki, "Optical Properties in Titanium-Diffused $LiNbO_3$ strip Waveguides", J. Appl. Phys., 49 (7) 3693-3698 (1978).

39. J. Noda, M. Fukuma, and S. Saito, "Effect of Mg Diffusion on Ti-Diffused $LiNbO_3$ Waveguides", J. Appl. Phys., 49 (6), 3150-3154 (1978).

40. W. K. Burns, C. H. Bulmer, and E. J. West, "Application of Li_2O Compensation Techniques to Ti-Diffused $LiNbO_3$ Planar and Channel Waveguides", Appl. Phys. Lett., 33 (1), 70-72 (1978).

APPENDIX A

THE PHOTOREFRACTIVE EQUATION

In this appendix we present an approximate solution of the continuity equation describing the photorefractive process in $LiNbO_3$. Our principle interest is in relating parameters of the material to the parameters describing the hologram formed when $LiNbO_3$ is exposed to an optical interference pattern. These include ϕ_g, the spatial phase shift between the hologram index pattern and the interference pattern, and Γ, a measure of photorefractive sensitivity.

At any point within the crystal, the concentration of conduction electrons n can be described by a continuity equation. Optical excitation generates conduction electrons, while retrapping, photovoltaic drift, and diffusion remove them. Kim et al[1] have employed an equation which appears to account for all relevant processes:

$$[\partial/\partial t + 1/\tau - D\partial^2/\partial y^2 + \mu\partial/\partial y(E + \mathcal{E}(y))]n = I(y)N\sigma q/h\nu. \tag{A-1}$$

The terms in this equation have the following meaning:

$\partial n/\partial t$ = rate of change of electron concentration;

n/τ = rate of electron retrapping;

$D\partial^2 n/\partial y^2$ = spatial derivative of the diffusion current, $D\partial n/\partial y$, where D is the diffusion coefficient;

$\mu\partial/\partial y(En + \mathcal{E}n)$ = spatial derivative of the drift current, where E is the effective photovoltaic field, $\mathcal{E}$ is the space charge field, and μ is the electron mobility;

$I(y)N\sigma q/h\nu$ = rate of generation of conduction electrons per unit volume, where N is the trapped electron concentration, σ is the absorption cross section, q is the quantum efficiency for processes resulting in migration, and $h\nu$ = photon energy.

For simplicity we eliminate diffusion from further consideration because holographic sensitivity in $LiNbO_3$ is much greater for geometries in

124

which both drift and diffusion operate than for geometries in which only diffusion operates. We also neglect the effects of $\mathcal{E}$ on electron drift, since it is small during the initial stages of hologram formation in which we have interest. Finally, we neglect $\partial n/\partial t$ in comparison to $1/\tau$, since the retrapping time is expected to be on the order of picoseconds while changes in n from the holographic process occur over a span of seconds. This simplifies Eq. (A-1) to

$$n + L(\partial n/\partial y) = I(y)N\sigma q\tau/h\nu \equiv g(y)\tau \tag{A-2}$$

where $L = \mu E\tau$ is the characteristic length that a conduction electron drifts before it is retrapped, and g(y) is a simplified notation for the generation rate. In the case of illumination by the interference pattern formed when two coherent beams intersect at an included angle 2θ,

$$g(y) = g_o(1 + m \cos 2ky\sin\theta) \quad , \tag{A-3}$$

the solution to Eq. (A-2) is

$$n(y) = n_o + n_s \sin(2ky\sin\theta) + n_c \cos(2ky\sin\theta) \quad , \tag{A-4}$$

$$n_o = g_o\tau \quad ,$$

$$n_s = g_o\tau\, m\, X/(1 + X^2) \quad ,$$

$$n_c = g_o\tau\, m/(1 + X^2) \quad ,$$

$$X = 2kL\sin\theta \equiv kL$$

Our greatest interest is in the concentration of trapped space charge ρ described by the continuity equation $\rho = e\mu E\partial n/\partial y$. After determining ρ with the help of Eq. (A-4), Maxwell's equation $\partial\mathcal{E}/\partial y = \rho/\epsilon$ is integrated to obtain the space-charge electric field. The associated refractive index perturbation is $n_1 = (1/2)\, n_e^3 r_{33}\mathcal{E}$, where $n_e = 2.2$ is the extraordinary refractive index of $LiNbO_3$, and $r_{33} = 30 \times 10^{-12}$ m/V is the appropriate electro-optic coefficient. We find

$$n_1 = n_{10} \cos(2kys\sin\theta - \phi_g) \qquad\qquad (A-5)$$

$$n_{10} = t(n_e^3 r_{33}/2)(eLg_0/\epsilon)m \cos\phi_g$$

$$\phi_g = \tan^{-1} 2kLs\sin\theta \quad .$$

The Photorefractive Parameters

The Phase Shift ϕ_g

The phase shift between the hologram index pattern and the optical interference pattern is seen to arise as a natural consequence of electron drift. The value $\phi_g = \pi/2$ required for self-subtra̲̲̲ ̲̲̲ ̲̲ ̲̲̲̲̲ ̲̲ ̲̲̲̲̲ ̲̲̲ ̲ crystal for which $2kLs\sin\theta \gg 1$; that is, the drift le̲̲̲̲̲̲ ̲ ̲ ̲̲̲̲̲ ̲̲ the hologram fringe spacing.

The value of L can be experimentally increased by reducing the number of empty traps. This can be done by increasing the relative number of filled traps, using reducing treatments of the type described by Phillips and Staebler.[2] It is also possible to increase L by using crystals with a lower density of impurities. However, this simultaneously reduces photorefractive sensitivity and may be undesirable.

It is worth noting that in crystals for which electron diffusion is the dominant transport mechanism a phase shift $\phi_g = \pi/2$ automatically occurs. This may be seen by keeping D and eliminating μ in Eq. (A-1), and following procedures similar to those we have already used in analysing that equation.

The Sensitivity Γ

In Section III we introduced the parameter Γ to describe the temporal evolution of hologram diffraction efficiency. If we now make the identification $\Gamma = \dot{n}_{10}\, kd/2I_0 m\, n_e\cos\theta$ and use the expression for n_{10} given in Eq. (A-5), we find

$$\Gamma = (e/hc)(d/\cos\theta)(n_e^3 r_{33}/4\epsilon)(LN\sigma q\cos\phi_g) \quad , \qquad\qquad (A-6)$$

valid during initial stages of hologram formation. The factors on the right hand side of this equation include geometric, material, and photorefractive parameters. In the limit $\phi_g = \pi/2$, $L\cos\phi_g = (2ks\sin\theta)^{-1} = \Lambda/2\pi$, where Λ is the hologram fringe spacing. The only photorefractive parameter that influences Γ is then the product $N\sigma q$.

References

1. D. M. Kim, R. R. Shah, T. A. Rabson, and F. K. Tittel, "Nonlinear Dynamic Theory for Photorefractive Phase Hologram Formation", Appl. Phys. Lett., 28 (6), 338-340 (1976).

2. W. Phillips and D. L. Staebler, "Control of the Fe^{2+} Concentration in Iron-Doped Lithium Niobate", J. Electron. Mater., 3 (2), 601-617 (1974).

APPENDIX B

DIFFRACTION BY THIN HOLOGRAMS IN PHOTOREFRACTIVE MEDIA

In this appendix we present a simplified theory of diffraction by thin holograms in photorefractive media. Our objective is to relate the diffraction efficiency η to the modulation of the refractive index, and to show the dependence of the diffracted output on the hologram phase shift ϕ_g.

At the entrance to the hologram region, signal and reference beams intersect and form an interference pattern having the intensity

$$I = R_o^2 + S_o^2 + 2R_o S_o \cos ky. \tag{B-1}$$

Owing to the phenomenon of photorefractivity, the material response to the interference pattern is a refractive-index modulation[1]

$$n_1(y) = n_{10} \cos(ky - \phi_g) \quad , \tag{B-2}$$

where n_{10} and ϕ_g depend on photorefractive parameters of $LiNbO_3$. As a result of this modulation, the signal and reference beams have the following approximate form at the exit to the hologram region:

$$S = S_o e^{iKV_o} e^{i\underline{k}_- \cdot \underline{r}} e^{i(kd/\cos\theta)n_1(y)/n} \tag{B-3}$$

$$R = R_o e^{i\underline{k}_+ \cdot \underline{r}} e^{i(kd/\cos\theta)n_1(y)/n}$$

$$\underline{k}_\pm = k(0, \pm \sin\theta, \cos\theta) \quad ,$$

where n is the average refractive index, d is the hologram thickness, k is the magnitude of the wave vector in the material, and θ is the Bragg angle.

The final exponential in the expressions for S and R simply represents the added phase associated with the modulated refractive index $n_1(y)$. If we expand this exponential in a Taylor series, then terms other than the zerp-order term in the resulting summations for S and R may be identified

as representing diffracted components of the incident plane waves S_o and R_o. A more precise analysis shows that the thick holograms we deal with in practice, the only diffracted components of significance are those traveling in the directions defined by $\underline{k}_\pm$.[2] If we keep only these components and work to first order in $kdn_{10}/2n\cos\theta = \eta^{1/2}$, we find

$$S \simeq S_o e^{iKV_o} e^{i\underline{k}_- \cdot \underline{r}} [1 + i\eta^{1/2} e^{i(ky - \phi_g)}] \tag{B-4}$$

$$R \simeq R_o e^{i\underline{k}_+ \cdot \underline{r}} [1 + i\eta^{1/2} e^{-i(ky - \phi_g)}] \quad .$$

Since $\underline{k}_- + yk = \underline{k}_+$, both S and R have components proportional to $e^{i\underline{k}_+ \cdot \underline{r}}$ and $e^{i\underline{k}_- \cdot \underline{r}}$ it is appropriate to redefine the signal amplitude S to represent all waves traveling in the direction $\underline{k}_-$ and R to represent all waves traveling in the direction $\underline{k}_+$, so that

$$S = (S_o + i\eta^{1/2} R_o e^{i\phi_g}) e^{iKV_o} e^{i\underline{k}_- \cdot \underline{r}} \tag{B-5}$$

$$R = (R_o + i\eta^{1/2} S_o e^{-i\phi_g}) e^{i\underline{k}_+ \cdot \underline{r}} \quad .$$

If S_o is now set equal to zero; that is, if the input signal beam is blocked, the output signal beam amplitude is

$$S = i\eta^{1/2} R_o e^{i\phi_g} \tag{B-6}$$

Note that the phase of the reconstructed signal beam is different from that of the input signal beam by an amount $\phi_g + \pi/2$. $\pi/2$ of this phase shift is associated with the phenomenon of diffraction, and ϕ_g is associated with the spatial offset between the hologram index pattern and the optical interference pattern used to form the hologram.

Since $\eta = SS^*/R_o R_o^*$ from Eq. B-6, we may identify

$$\eta = (kdn_{10}/2n\cos\theta)^2 \tag{B-7}$$

with the hologram diffraction efficiency. The relation between the photo-refractive sensitivity Γ and the index modulation amplitude is found by comparing Eq. B-7 and Eq. III-2, with the result

$$\Gamma = \dot{n}_{10} kd / 2 I_o m \; n\cos\theta \quad . \tag{B-8}$$

This expression was used in Appendix A to obtain Γ in terms of photorefrac-tive parameters.

References

1. D. W. Vahey, "A Nonlinear Coupled-Wave Theory of Holographic Storage in Ferroelectric Materials", J. Appl. Phys., 46 (8), 3510-3515 (1975).

2. H. Kogelnick, "Coupled-Wave Theory for Thick Hologram Gratings", Bell Syst. Tech. J., 48 (9), 2909-2947 (1969).

APPENDIX C

OPTICAL WAVEGUIDE GEODESIC LENSES

The goal of this program has been the development of optical geodesic waveguide lenses having performance characteristics sufficient for the needs of wide-band, integrated optical, RF spectrum analyzers in $LiNbO_3$ and Si substrates.[1] The means selected for achieving this goal was the formation of "corrector-plate" waveguide overlays adjacent to spherical depression lenses. The material and shape of the overlays is chosen in such a way that the corrector-plate-lens combination produces near diffraction-limited focal characteristics when employed within its field of view.

In operation, a guided beam of light passing beneath the corrector plate has its phase wavefront distorted in a manner equal and opposite to the phase distortion associated with lens aberrations. Determination of the appropriate geometric shape for the corrector plate and an analysis of fabrication tolerances was based on thin-lens theory and included only the effects of third-order spherical aberrations. A more exact analysis was found to be unnecessary for the needs of the program.

Parameters of the corrector-plate design equation include the coefficient of third-order lens aberrations and the change in waveguide mode index associated with the corrector-plate overlay. Imprecise knowledge of the shape of the lens and the refractive-index profile of the waveguide make it impractical to calculate these quantities, so we have developed techniques that permit their accurate measurement. Third-order spherical aberrations are measured by imaging the focal spot obtained using a narrow pencil of parallel input rays. Variations in the image distance as the pencil is translated across the geodesic lens aperture are directly correlated with the desired aberration coefficient. A similar technique has been used to measure residual aberrations in a corrected lens.

The change in waveguide mode index associated with a given corrector-plate material can be measured by fabricating a diamond-shaped overlay of the material on the waveguide, and using this overlay to refract a single incident guided wave propagating parallel to the long axis of the pattern into two symmetrically disposed guided waves propagating at small angles to the axis.

Measurement of the included angle between beams permits the determination of the mode-index change for the overlay, Δn_g. This change should be made as large as possible without, however, resorting to overlay materials having a refractive index greater than that of the waveguide. We found $\Delta n_g \simeq 3 \times 10^{-4}$ for Al overlays on our $LiNbO_3$:Ti waveguides. This value was large enough to permit the use of corrector plates conveniently limited in area to 5-10 mm^2. Conventional photolithographic techniques were employed in their fabrication.

Spherical depressions in $LiNbO_3$ substrates were generated using a wheel embedded with 20 μm-sized diamond chips to achieve a fine-ground finish. This was followed by final polishing with a microcloth lap and a suspension of 30-nm sized SiO_2 particles. The conventional intermediate polishing stage was eliminated. Use of a cloth lap was seen to be undesirable in that it resulted in a loss of precise figure of the lens depression by 5-10%. However, conventional polishing with a pitch lap was tested and found to be unsuited to the refractory materials' characteristics of $LiNbO_3$. Cloth produced a good surface finish and at the same time resulted in a rounded rim, facilitating coupling in and out of the depression region. The loss of figure that was tolerated to attain high optical quality resulted in lens focal lengths 10% longer than desired, and spherical aberrations that were 14-20% less than expected.

In the case of a 55-mm focal-length lens, focusing occurred well outside the output prism coupler; yet prism-associated aberrations were calculated to be only 1.5% of those of the uncorrected lens. They were, however, comparable to and opposite in sign from small residual aberrations in the 4%-overcorrected lens. Consequently, they may have served to enhance the quality of the focal profiles obtained, which showed near diffraction-limited f/19 performance over a full field of view exceeding 1.22°, a figure consistent with theory developed in the program. The full width of the central peak was 13.25 μm, about 17% above the diffraction-limited value for a 2.8-mm input aperture. Sidelobes, except for the first, were also larger than expected for a diffraction-limited lens; however, this observation could be interpreted as indicating the existence of beam nonuniformities introduced during prism input and output coupling, as well as residual aberrations.

The results of this program lead to the conclusion that integrated optical corrector plates are a practical way to improve the focal characteristics of any waveguide lens, geodesic or otherwise. More generally, the introduction of corrector plates into planar integrated optical systems appears to provide a useful mechanism for relaxing the severe tolerance requirements that have plagued the development of these systems to the present time. We believe that lenses of the type fabricated in this program are emminently suitable for one of the more ambitious planar optical systems now under consideration, that of an integrated optical RF Spectrum Analyzer.[1]

(1) M. C. Hamilton, D. A. Wille, and W. J. Miceli, "An Integrated Optical RF Spectrum Analyzer", Optical Engineering, 16 (5), 475-478 (1977).

APPENDIX D

THE TWO-PHOTON PHOTOREFRACTIVE EFFECT

One of the more difficult problems to overcome in the development of the preprocessor is that of utilizing the photorefractive effect to write a hologram in a $LiNbO_3$ waveguide without causing "optical damage", i.e., undesirable photorefractive effects in the remainder of the waveguide. One approach to overcoming this problem has been to locally enhance the photorefractive sensitivity in the hologram region by iron doping. However, the enhanced photorefractive sensitivity brought about by iron doping is accompanied by a short hologram lifetime and unavoidable sensitivity to the signal and reference beams during the operation of the preprocessor. In some modes of operation, for example, the self subtraction mode, this effect is desirable. However, for the more conventional system, it is highly desirable to be able to write a stable hologram in an otherwise damage-resistant waveguide. Toward this end, nonlinear photorefractive processes have been investigated.

An ideal situation is depicted in Figure D-1, which shows a high-purity and therefore highly damage resistant $LiNbO_3$ waveguide. During the period of hologram formation there is a sensitizing beam, incident from above, which allows the hologram to be written. After writing the sensitizing beam is turned off and a stable, nondamageable hologram remains for subsequent use.

It is believed that the two-photon photorefractive effect involves an interband transition with a subsequent stable electron distribution. In the ideal case the information-carrying guided waves would be well beyond the $LiNbO_3$ intrinsic band edge (.6328 or even 0.8 μm). The writing would then be sensitized by a shorter wavelength beam. Although we have not yet been able to accomplish this, we have written two-photon holograms under a number of conditions. The waveguide experiments are described in the reprint which is at the end of this Appendix. Some bulk experiments which were performed to further illustrate the nature of the phenomenon are described below.

The experimental arrangement is shown in Fig. D-2. For a variety of reasons these experiments were done at a (frequency-doubled) wavelength of 0.659 μm, rather than the wavelength 0.532 μm used in the experiment reported in the reprint and in the work of von der Linde et al. referred to

Sensitizing
radiation

(a)

Stable hologram
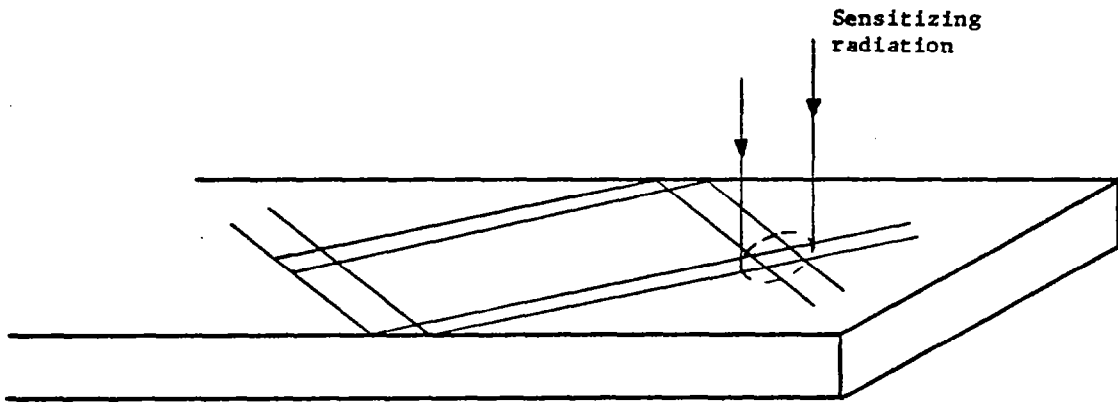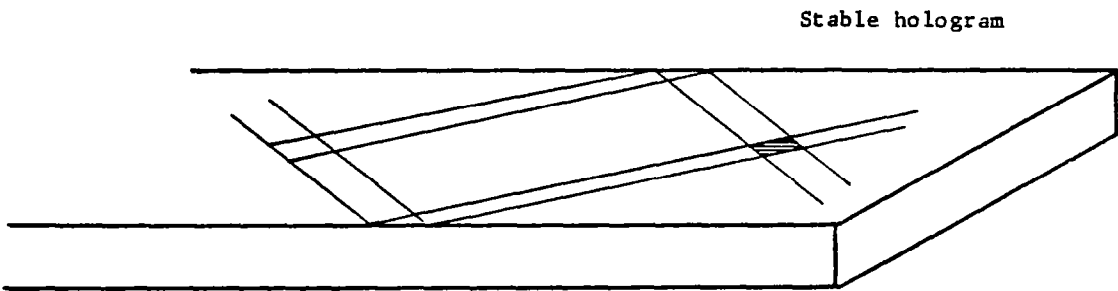
(b)

FIGURE D-1.   IDEAL HOLOGRAM FORMATION PROCESS IN A
"DAMAGE-PROOF" LiNbO$_3$ WAVEGUIDE.
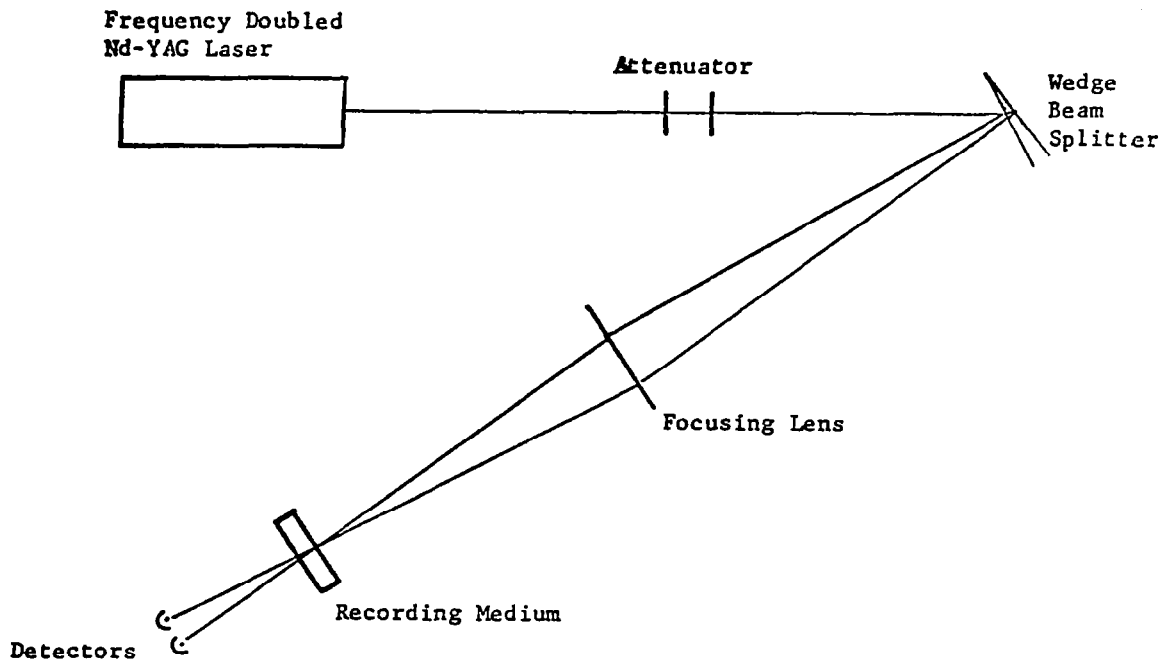
FIGURE D-2. EXPERIMENTAL ARRANGEMENT FOR BULK TWO-PHOTON
HOLOGRAPHIC RECORDING.

therein. Experiments on one sample at 0.532 μm gave similar results, in slope and sensitivity, to those reported here. The wedge beam-splitter produces a rather coarse grating, of around 10 μm wavelength. In all the experiments, the polarization, the grating vector, and the crystal c-axis were parallel. The laser emitted 160 ns pulses at a rate of 25 per second.

Four samples were used. The first (A) was from the same boule of nominally undoped $LiNbO_3$ as that used for the waveguide experiment, and it was used as received—that is, it was not heat treated to form a waveguide. The second sample (B) was from a specially grown boule of high-purity $LiNbO_3$. The third sample (C) was a piece adjoining (B) from the same boule; it was heat treated in oxygen at 950°C for one hour. The fourth sample (D) was a crystal doped with 0.015% iron.

Results for all 4 samples are shown in Fig. D-3. The straight lines are just "eyeballed" fits to the data. Two-photon photorefractive changes are readily seen in all four samples despite the relatively low peak power, the bulk configuration, and the relatively long wavelength used for excitation. It is also apparent that the iron-doped sample is for some reason more sensitive than the other samples. In Fig. D-4, the data for samples (A) and (D) are plotted on an expanded scale along with a straight line representing a fit to the results obtained by workers at Bell Labs[†] at higher peak powers (and at 0.533 μm) on a sample of Cu-doped $LiNbO_3$. Our iron-doped sample appears to be of roughly comparable sensitivity to the Bell Labs crystal; the lower senstivity observed for sample (A) tends to reinforce the conclusions concerning the greater sensitivity of the waveguide configuration, although more work is required before any final conclusions can be drawn.

---

† Ref. 5 of the appended reprint; the line has been adjusted to correspond to the pulse length of the present experiments.
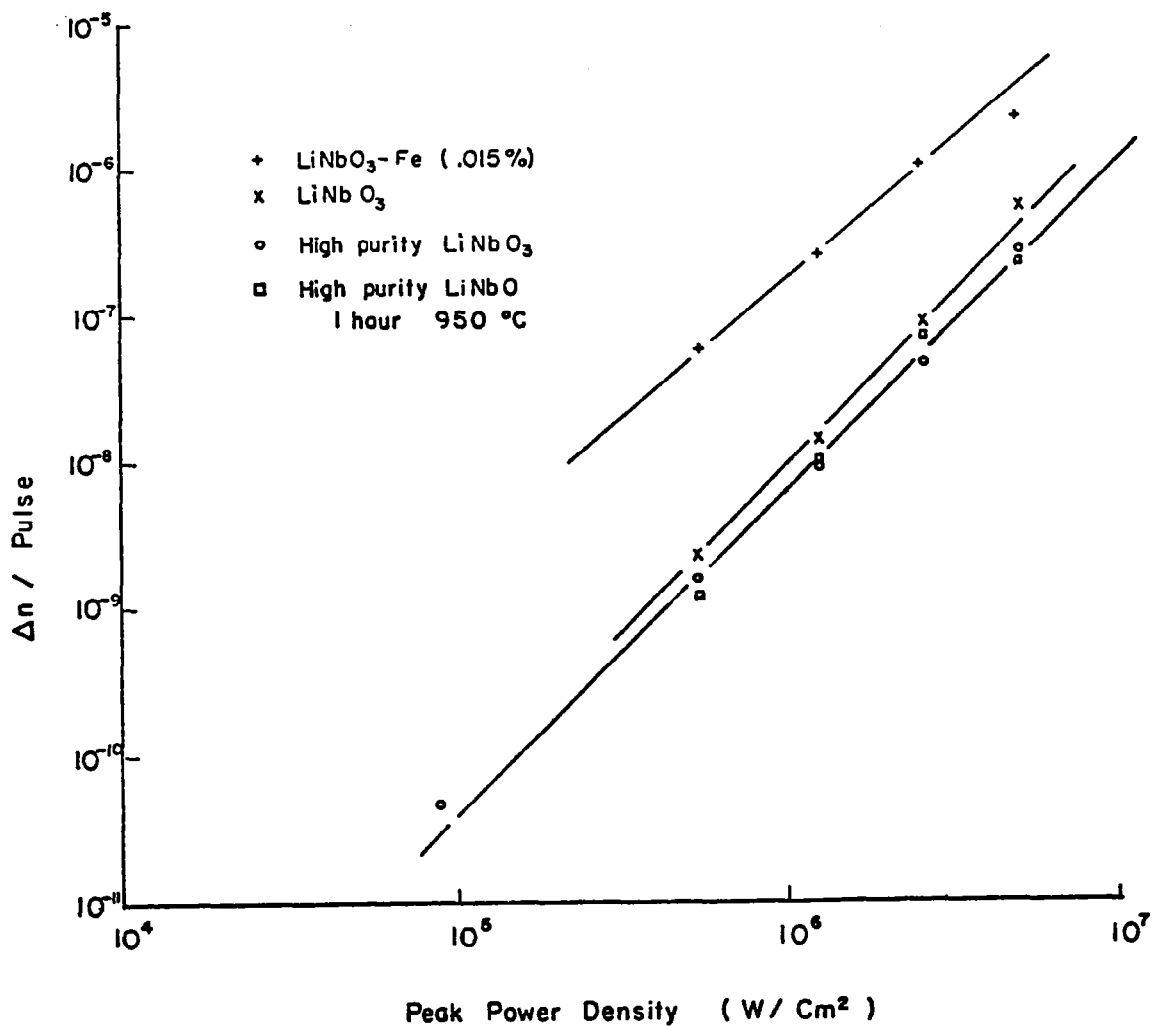
FIGURE D-3.   REFRACTIVE INDEX CHANGE, AS DETERMINED FROM HOLOGRAPHIC
             DIFFRACTION EFFICIENCY, PER 160 ns PULSE OF INCIDENT
             0.659 μm RADIATION AT SEVERAL POWER DENSITIES, IN BULK
             TWO-PHOTON HOLOGRAM RECORDING EXPERIMENTS ON 4 LiNbO$_3$
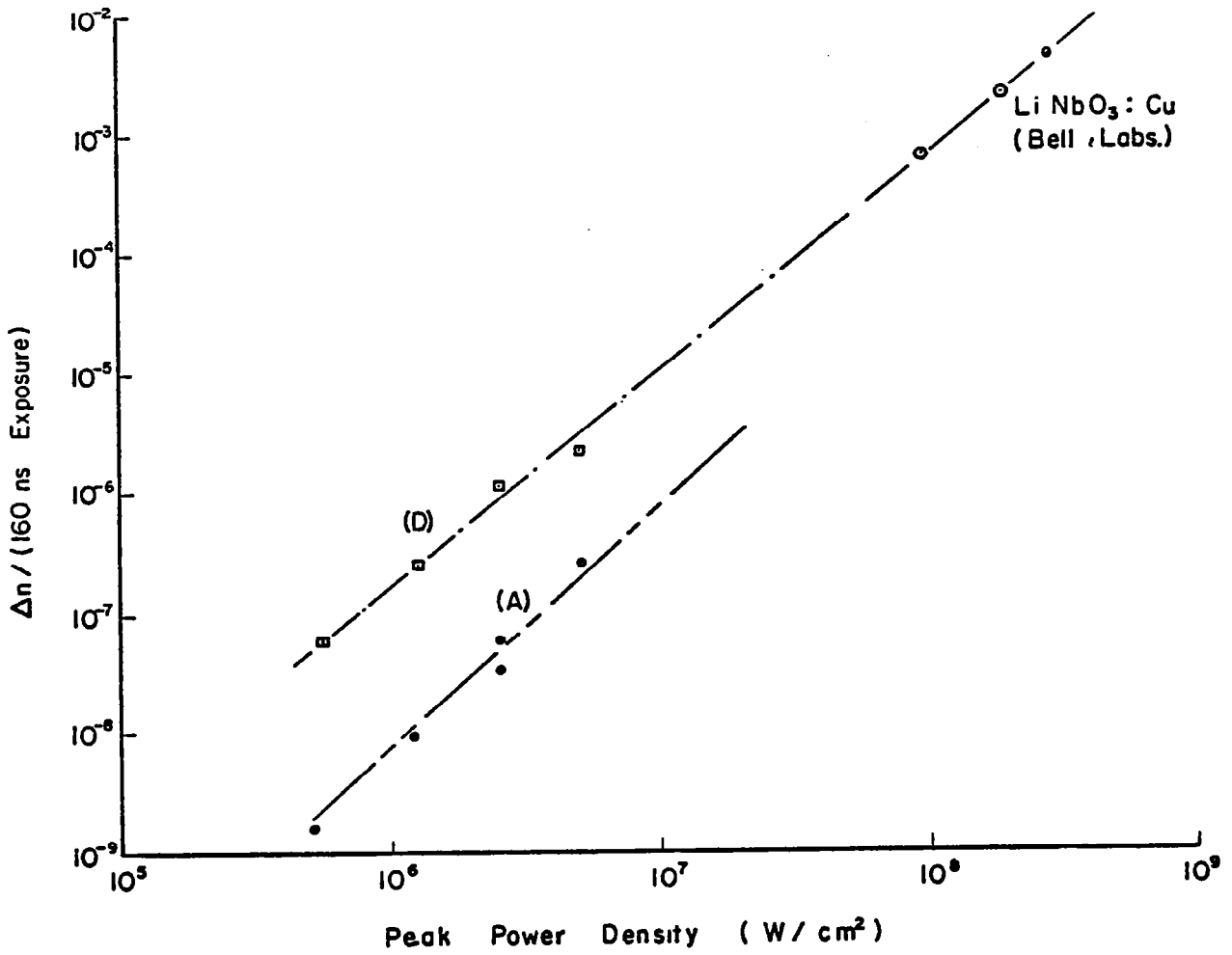   CRYSTALS.

FIGURE D-4. REFRACTIVE INDEX CHANGE PER PULSE, AS IN FIG. D-3, FOR TWO LiNbO$_3$ CRYSTALS COMPARED WITH SIMILAR DATA OBTAINED AT HIGHER POWER DENSITY BY VON DER LINDE ET AL (BELL LABORATORIES).

Formation of integrated optics components by multiphoton photorefractive processes*

C. M. Verber and N. F. Hartman

Battelle Columbus Laboratories, Columbus, Ohio 43201

A. M. Glass

Bell Laboratories, Murray Hill, New Jersey 07974
(Received 15 November 1976)

Phase holograms have been formed in pure outdiffused LiNbO₃ waveguides using the photorefractive effect initiated by two-photon absorption. Efficient holograms were recorded with two guided waves, of less than 100 W peak power, and less than 10 μJ optical energy. This technique automatically compensates for wave front perturbations in the guide and allows the formation of various integrated components such as wavelength multiplexers and optical switches.

It has previously been demonstrated[1] that the photorefractive effect may be used to record thick phase holograms in optical waveguides for applications such as mirrors, beam splitters, and large-angle switches.[2] The unique advantage of this mechanism is that thick phase holograms can be recorded *in situ* using guided waves having the same mode structure and often derived from the same source as the light used during the operation of the device. Such holograms automatically compensate for wave-front perturbations introduced by the mode structure or by other waveguide structures or imperfections.

In the previous demonstrations the photorefractive process has been initiated by linear absorption of the writing light either by the transition-metal ions used to make infused waveguides[3] in LiNbO₃ or, presumably, by the residual impurities in outdiffused LiNbO₃ waveguides.[1] Although high diffraction efficiencies have been achieved, components formed in this way have the disadvantage of possible degradation during use since the guided waves present during the operation of the device can initiate the same mechanism used to create the hologram.

Improved stability can be achieved by thermal fixing[4] although this has certain practical complications for waveguide gratings. LiNbO₃ cannot be locally heated without fracturing, and heating the entire assembly is awkward and makes it impossible to fabricate more than one holographic element on a single substrate. An additional disadvantage of thermal fixing is the high optical energy required to recover the hologram after fixing.

An alternative linear photorefractive technique is the use of short-wavelength light to form the holograms and long-wavelength light, for which the photorefractive sensitivity is negligible, as the operating wavelength. This technique is satisfactory only for simple plane-grating holograms since complex thick holograms suffer large losses in fidelity and efficiency if the read and write wavelengths differ.

A holographic writing technique which has the useful advantages of avoiding destructive readout, producing stable holograms, and retaining the low optical loss of outdiffused waveguides in undoped crystals is based

upon the use of multiphoton absorption[5] for initiating the photorefractive process. We have now demonstrated that holograms may be recorded by a two-photon absorption process in outdiffused LiNbO₃ waveguides by intersecting two guided waves, and that the required energy and intensity are readily achieved in the waveguide using commercially available lasers.

The absorption of a beam of light of intensity I (W/cm²) by a two-photon process is described by

$$dI = -\alpha_2 I^2(x)\,dx, \tag{1}$$

where α_2 is the second-order absorption coefficient and x is depth in the crystal measured from the surface upon which the beam is incident. The index change associated with the photorefractive effect is proportional to the number of electrons excited into the conduction band. For the two-photon process, this number will be proportional to $\tfrac{1}{2}N$, where N is the number of photons absorbed per cm³. The index change is then

$$\Delta n = S(\tfrac{1}{2}N), \tag{2}$$

where S is a proportionality constant characteristic of the material and the geometry. For an optically thin sample, Δn is not a function of x and

$$N(t) = -\frac{\Delta I}{\Delta x}\frac{1}{h\nu}t = \frac{\alpha_2}{h\nu}tI^2, \tag{3}$$

where t is the time during which the sample is irradiated and $h\nu$ is the photon energy. If the irradiation occurs in the form of M equal rectangular pulses of duration Δt, we have

$$\Delta n = kI^2 M\,\Delta t \tag{4}$$

where all constants have been absorbed into k. The two-photon process is then indicated by a quadratic dependence of $\Delta n/M$ upon I.

This quadratic dependence was observed using a Nd:YAG laser with an intercavity doubler which produced 140-nsec pulses of 0.53-μ radiation. After reflection from a wedge beam splitter the laser output was prism coupled into an outdiffused waveguide in the surface of an undoped LiNbO₃ slab. The value of the induced Δn was monitored by measuring the diffraction efficiency of the holographic grating formed in the
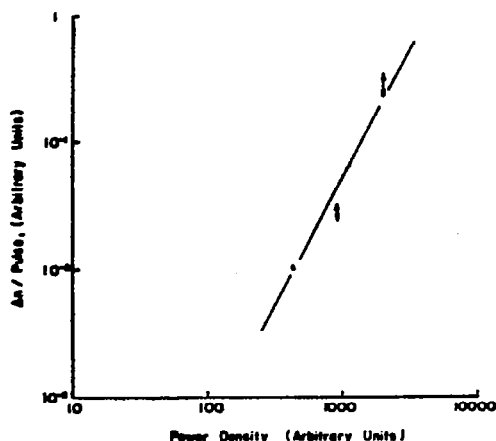
140

FIG. 1. Log-log plot of Δn/pulse versus power density for a grating written in an outdiffused $LiNbO_3$ waveguide with pulsed 0.53-μ radiation. The solid line has a slope of 2. The grating spacing is 0.55 μ.

beam overlap region. The data, shown in the log-log plot in Fig. 1, clearly display the quadratic behavior indicative of the two-photon effect. Only relative values are given due to the difficulty in accurately determining power densities and the diffraction efficiency in the waveguide. However, from the maximum power incident upon the coupling prism of 2 kW and estimates of the coupling efficiency and effective waveguide thickness, we can estimate a maximum power density of 10^6 W/cm^2 in the waveguide. Diffraction efficiencies of several percent were observed with no sign of saturation. Comparison of these results with the previous data[5] obtained using the same wavelength but in a bulk, rather than a waveguide, configuration shows a far greater sensitivity for the waveguide case. The discrepancy is larger than can be accounted for by experimental errors or errors in estimating the power density in the waveguide. We feel that the discrepancy may be due to changes induced by the outdiffusion process or to other compositional differences in the samples used for the two sets of measurements.

In an additional experiment performed with a flash-lamp-pumped dye laser, waveguide gratings were formed via the nonlinear process using a writing wavelength of 5974 Å. Diffraction efficiencies exceeding 10% were achieved with a single pulse. Measurement of the wavelength tuning curve of these gratings confirmed that they were thick phase gratings. The wavelength selectivity suggests that this process could be used to fabricate an integrated wavelength multiplex/demultiplex system. A typical configuration would consist of a linear array of tuned holographic gratings each of which would operate upon a specific wavelength for insertion into or removal from the multiplexed bundle.

It is evident that, since the sensitivity of the two-photon writing process is quadratic in optical power density, the holograms written by this process will not be optically altered at normal read powers. Moreover, two-photon absorption in undoped $LiNbO_3$ using dissimilar photons, which was reported in Ref. 5, can also be used to advantage in the waveguide configuration. By using a long-wavelength guided (information carrying) wave and a normally incident shorter-wavelength beam for sensitizing the photorefractive effect, one can achieve the dual goals of writing a permanent optically inert hologram with no risk of "optically damaging" the remainder of the guided wave structure and matching the output wavelengths of the He-Ne, GaAs, and solid-state Nd lasers.

The authors would like to acknowledge fruitful discussions with W.J. Tomlinson, R.P. Kenan, and V.E. Wood.

[1]V.E. Wood, N.F. Hartman, C.M. Verber, and R.P. Kenan, J. Appl. Phys. 46, 1214 (1975).
[2]R.P. Kenan, D.W. Vahey, N.F. Hartman, V.E. Wood, and C.M. Verber, Opt. Eng. 15, 12 (1976).
[3]C.M. Verber, V.E. Wood, R.P. Kenan, and N.F. Hartman, Ferroelectrics 10, 253 (1976).
[4]D.L. Staebler and J.J. Amodei, Ferroelectrics 3, 107 (1972).
[5]D. Von der Linde, A.M. Glass, and K.F. Rodgers, Appl. Phys. Lett. 25, 155 (1974).

141

APPENDIX E

ION-BEAM MILLING OF LiNbO$_3$

Experiments in ion-beam milling of LiNbO$_3$ were carried out on the IMMI-V system manufactured by Commonwealth Scientific Corporation. A Type 2-50, ion gun was employed, which can provide a total beam current of 1 mA, along with an eccentric rotating stage providing a 2-dimensional quasi-linear motion of the sample in the ion beam. Circular areas up to 1-inch diameter can be milled in this fashion. The gun is equipped with a hot-filament-type beam neutralizer to preclude charge buildup on dielectric materials.

Preliminary milling experiments pointed up the necessity of excluding all organic materials from the chamber to minimize sample surface contamination from deposition of foreign species from the chamber environment. Organics proved to be virtually impossible to remove from the surface. This condition for milling operations precluded the use of a photoresist milling mask.

Narrow ridges in outdiffused LiNbO$_3$ have successfully been formed by milling away the surrounding area. The slab first was coated with a 2-μm-thick layer of aluminum by vacuum deposition. A positive photolithographic mask of a set of horn-coupled waveguides of different widths, plus a 100-μm wide uncoupled line, was used to form a pattern in a positive resist coating. The resist was masked further to prevent removal of aluminum, in the subsequent chemical etching process, from the areas extending from the horn couplers to the sample edge in-line with the optical path. Thus, planar waveguide surfaces were retained for mounting the coupling prisms. Milling time in an 0.8 mA, neutralized argon-ion beam, with sample normal to the beam, was 27 hours.

Three ridges thus milled were 8 mm long and 100 μm, 12 μm, and 10 μm wide. A transmission photomicrograph of the horn-coupler ends is shown in Figure E-1 while a reflection photomicrograph of the 10-μm ridge, at higher magnification, is shown in Figure E-2. The photomicrographs were made before attempting to remove residual aluminum from the LiNbO$_3$ surface. A granular structure on top of the ridges, as well as over the Al-masked portions of the milled region, is believed to be aluminum. However, this structure did not yield to conventional aluminum etchants which readily dissolved the aluminum
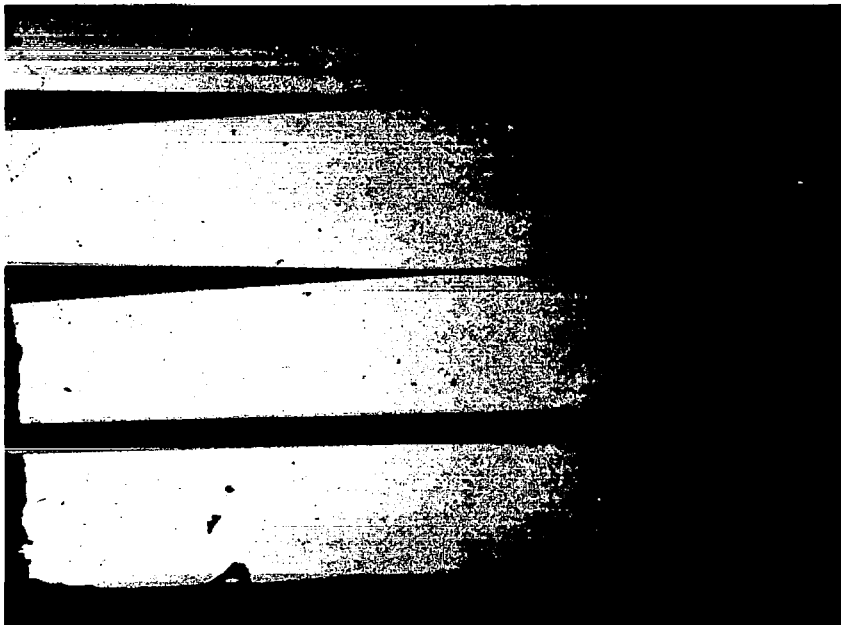
FIGURE E-1. TRANSMISSION PHOTOMICROGRAPH OF ION—MILLED RIDGES IN LiNbO$_3$ (24X), SHOWING A 100 μm WIDE STRAIGHT RIDGE AND 12 μm AND 10 μm WIDE RIDGES WITH HORN COUPLERS. THE RIDGES ARE DARK WHEN VIEWED WITH TRANSMITTED LIGHT BECAUSE THE REMANTS OF THE ALUMINUM MASK (ORIGINAL THICKNESS 2 μm) HAVE NOT YET BEEN REMOVED.
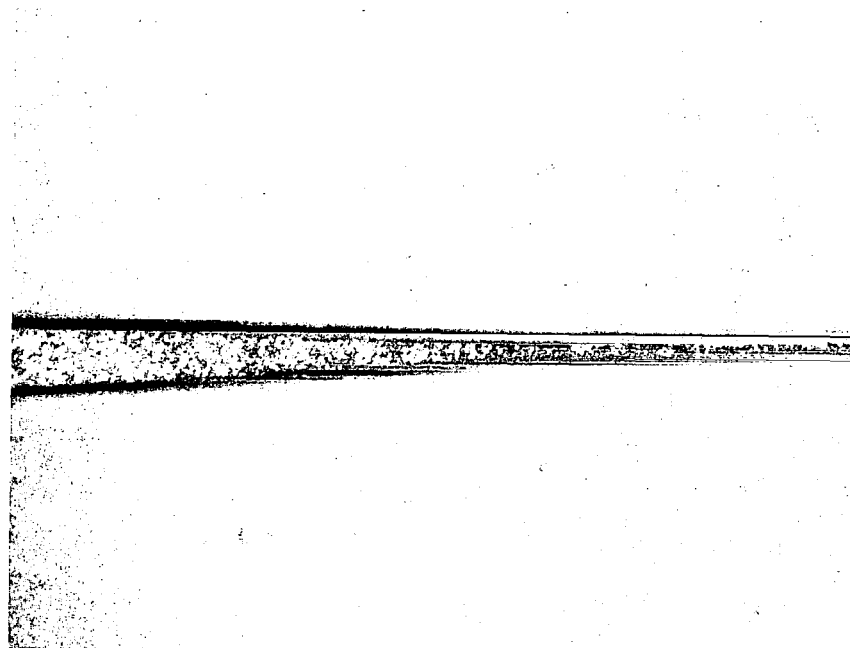
FIGURE E-2. REFLECTION PHOTOMICROGRAPH OF A PORTION OF
THE 10 μm RIDGE (240X). THE BRIGHT REGION
IS THE REMNANT OF THE ALUMINUM MASK. THE
DARK BORDERS ARE THE SIDEWALLS.

layer not exposed to the ion beam. Other etchants, short of formulations and immersion periods known to etch $LiNbO_3$, have proven unsuccessful.

A Talysurf surface profile normal to, and approximately midpoint of, the ridges indicated the following heights:

Ridge Height, μm	Ridge Width, μm	
	0.8 height	base
4.7	90	105
2.0	8	20
0	--	--

The peak-to-peak roughness of the granular structure was about 0.375 μm. This roughness obtained to within 400 μm of the perimeter of the milled area, whereupon it decreased to below the instrument noise level over the bare $LiNbO_3$. The granular structure was present on top the ridge of "zero" height, as well as on the other two. The milled $LiNbO_3$ between ridges was as smooth (by Talysurf examination) as the unmilled surface, although some pits and "crow's foot" defects were evident by Nomarski microscope inspection.

Waveguiding was observed in the milled ridges, but the light guided by the ridges was highly attenuated, presumably because of the presence of the granular material. Metal masks appear to be feasible for future ion-beam-milling experiments on $LiNbO_3$. Whether or not aluminum is a suitable metal will be determined from further analysis of the sample described, including scanning-electron-microscope assessment of the refractory coating over the ridges and other aluminum-masked areas exposed to the ion beam; aluminum may react with lithium, for example, under ion bombardment after aluminum thickness has been reduced to a few tens of atomic layers. Careful lapping and polishing of the milled slab, followed by re-formation of an outdiffused waveguide, may be possible should no way be found to prevent the surface damage.