INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in “sectioning” the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University Microfilms International
300 N. ZEEB ROAD, ANN ARBOR, MI 48106
18 BEDFORD ROW, LONDON WC1R 4EJ, ENGLAND
BIENVENU, MICHAEL PATRICK

CONDUCTIVITY STUDIES IN IRON-DOPED LITHIUM-NIOBATE BY PHASE HOLOGRAPHY

Rice University

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106

Ph.D. 1981
RICE UNIVERSITY

CONDUCTIVITY STUDIES IN IRON DOPED LITHIUM NIOBATE

BY PHASE HOLOGRAPHY

by

MICHAEL PATRICK BIENVENU

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

DOCTOR OF PHILOSOPHY

APPROVED, THESIS COMMITTEE

Dr. Thomas A. Rabson, Chairman
Dept. of Electrical Engineering

Dr. Dae Han Kim, Professor of
Electrical Engineering

Dr. William L. Wilson, Professor
of Electrical Engineering

Dr. Barry F. Dunning, Professor
of Space Physics

HOUSTON, TEXAS

MAY, 1981
Conductivity Studies of Iron Doped Lithium Niobate by Phase Holography

by

Michael Patrick Bienvenu

ABSTRACT

By the use of holographic storage techniques, the conductivity of iron doped lithium niobate has been investigated. The results indicate that the conductivity of the material immediately after hologram formation is much greater than previously suspected. This increased conductivity is shown to be a direct result of the multivalent properties of the iron dopant molecules, and the underlying theory necessary to account for the effect is presented, this theory being an elaboration of the bulk photovoltaic theory of von der Linde and Glass.
Acknowledgements

It is difficult to know from where one receives the inspiration and desire to pursue such a project as this. Most certainly, my parents were instrumental, and for this they deserve the utmost thanks. My advisor, and mentor, Dr. Rabson, must be credited, not only with helping with this work in particular, but also with relieving me of many of my delusions about scientific research (and graduate work). My many friends have helped me through these past years, and they unfortunately number too many to list here. Dr. P. Pazzaglia deserves special mention, for she provided a very valuable moral support, and acted as a sounding board for my various ideas.

This thesis is then dedicated to the folks named above. But it is also dedicated to the entire scientific community, which I am honored to join.
## TABLE OF CONTENTS

### Chapter 1 - Introduction

I. Background  \hspace{1cm} 1

II. Structure of Lithium Niobate

   1. Pure Crystals  \hspace{1cm} 2

   2. Metal Doped Crystals  \hspace{1cm} 4

III. Crystal Properties of LiNbO₃  \hspace{1cm} 6

### Chapter 2 - Current Status of Understanding of Holographic Storage of LiNbO₃

I. The Holographic Storage Process  \hspace{1cm} 13

II. Theory of Electron Transport  \hspace{1cm} 14

III. Hologram Formation Process  \hspace{1cm} 15

IV. Diffraction by a Thick Phase Transmission Grating  \hspace{1cm} 17

V. Summary  \hspace{1cm} 20

### Chapter 3 - Experimental Work  \hspace{1cm} 24

I. Experimental Setup  \hspace{1cm} 26

II. Experimental Procedure  \hspace{1cm} 29

III. Experimental Data

   1. Dopant Level vs. Absorption  \hspace{1cm} 30

   2. Time Varying Decay Rate  \hspace{1cm} 31

   3. Electric Field Dependence  \hspace{1cm} 33

### Chapter 4 - Theory  \hspace{1cm} 49

I. Modified Build Photorefractive Effect  \hspace{1cm} 56

II. Conclusion  \hspace{1cm} 59
II. Rate Equations for Donor Sites and Conduction Electrons 61

III. Asymmetric Potential Well 71
   1. The Fe$^{2+}$ Site 72
   2. The Fe$^{3+}$ Site 75

IV. Limitations of the Model 77

Chapter 5 - Discussion 79

I. Data 79
   1. Field Dependence 81

II. Comparison of the Results to the Works of Other Researchers 84
   1. Phonon Generation 84
   2. Induced Absorption 86
   3. Luminescence 86
   4. "Thermal Fixing" 87
   5. Electric Field Effects 89

III. Conclusions 89

Appendix A 90

Appendix B 95

References 96
List of Figures

<table>
<thead>
<tr>
<th>Fig.</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Crystal structure of LiNbO$_3$</td>
<td>3</td>
</tr>
<tr>
<td>1-2</td>
<td>Absorption spectra of Fe-doped LiNbO$_3$</td>
<td>5</td>
</tr>
<tr>
<td>1-3</td>
<td>Energy band gap diagram</td>
<td>7</td>
</tr>
<tr>
<td>2-1</td>
<td>Spatial relationships between important quantities</td>
<td>19</td>
</tr>
<tr>
<td>2-2</td>
<td>Thick phase transmission grating</td>
<td>21</td>
</tr>
<tr>
<td>2-3</td>
<td>Time behavior of $\Delta n(t)$ and $\eta(t)$</td>
<td>24</td>
</tr>
<tr>
<td>3-1</td>
<td>Layout of experimental setup</td>
<td>27</td>
</tr>
<tr>
<td>3-2</td>
<td>Optical attenuation vs Fe$^{2+}$ concentration</td>
<td>32</td>
</tr>
<tr>
<td>3-3</td>
<td>Two decay curves in X9, showing reproducibility</td>
<td>34</td>
</tr>
<tr>
<td>3-4</td>
<td>Decay curve of hologram written to saturation</td>
<td>35</td>
</tr>
<tr>
<td>3-5</td>
<td>Decay curve for a low-doped crystal</td>
<td>36</td>
</tr>
<tr>
<td>3-6</td>
<td>Decay curve for a moderately-doped crystal</td>
<td>37</td>
</tr>
<tr>
<td>3-7</td>
<td>Decay curve for a moderately-doped crystal</td>
<td>38</td>
</tr>
<tr>
<td>3-8</td>
<td>Decay curve for a strongly-doped crystal</td>
<td>39</td>
</tr>
<tr>
<td>3-9</td>
<td>Schematic of parametric model</td>
<td>41</td>
</tr>
<tr>
<td>3-10a</td>
<td>Initial section of plot of decay rate vs decay time (Crystal X6)</td>
<td>42</td>
</tr>
<tr>
<td>3-10b</td>
<td>Intermediate section of plot of decay rate vs decay time (Crystal X6)</td>
<td>43</td>
</tr>
<tr>
<td>3-10c</td>
<td>Final section of plot of decay rate vs decay time (Crystal X6)</td>
<td>44</td>
</tr>
<tr>
<td>3-11</td>
<td>Decay rate vs decay time (Crystal X6)</td>
<td>45</td>
</tr>
<tr>
<td>3-12</td>
<td>Decay rate vs decay time (Crystal X16)</td>
<td>46</td>
</tr>
<tr>
<td>3-13</td>
<td>Anomalous decay curve for Crystal X2</td>
<td>50</td>
</tr>
<tr>
<td>Fig. 3-14</td>
<td>Decay parameters vs $\text{Fe}^{2+}$ concentration</td>
<td>51</td>
</tr>
<tr>
<td>Fig. 3-15</td>
<td>Decay parameters vs electric field for Crystal X16</td>
<td>52</td>
</tr>
<tr>
<td>Fig. 3-16</td>
<td>Decay parameters vs electric field for Crystal X9</td>
<td>53</td>
</tr>
<tr>
<td>Fig. 3-17</td>
<td>Decay parameters vs electric field for Crystal X6</td>
<td>54</td>
</tr>
<tr>
<td>Fig. 4-1</td>
<td>Asymmetric potential well model</td>
<td>73</td>
</tr>
<tr>
<td>Fig. 4-2</td>
<td>Configuration coordinate diagram for the $\text{Fe}^{2+}_0$ site</td>
<td>74</td>
</tr>
<tr>
<td>Fig. 4-3</td>
<td>Configuration coordinate diagram for the $\text{Fe}^{3+}_0$ site</td>
<td>76</td>
</tr>
<tr>
<td>Fig. 5-1</td>
<td>Configuration coordinate diagram for the $\text{Fe}^{2+}_0$ site with applied electric field</td>
<td>83</td>
</tr>
<tr>
<td>Fig. 5-2</td>
<td>Configuration coordinate diagram for the $\text{Fe}^{3+}_0$ site with applied electric field</td>
<td>85</td>
</tr>
</tbody>
</table>
Chapter 1
Introduction

I. Background

Lithium niobate has recently found great popularity in numerous applications. The crystal exhibits several large non-linear effects, and these make it useful as a material for acoustic waveguides [1], electro-optic devices [2], and optical data storage devices [3, 4]. The crystal may be doped in order to modify certain effects [5, 6, 7, 8]. While employing the doped crystal in the optical data storage role, unusual conductivity effects have been discovered, and this work examines and presents theories for these effects.

Elementary diffraction gratings were written in LiNbO$_3$ crystals, and the decay of the non-uniform electron density thus created was observed in order to gain information on the conductivity of the crystals. Experimental data indicates the existence of a transient photoconductivity which dominates the hologram decay process immediately after the cessation of illumination. The conductivity after illumination can be described by

$$\sigma(t) = \sigma_0 + \sigma_1 e^{-t/\tau_1} + \sigma_2 e^{-t/\tau_2}$$

This form is shown to be a result of the multivalent properties of the iron dopant ions.

The conductivity is also affected by external electric fields, and these data give additional information about the nature of the iron ion sites, and the electron transport process.
The structure and basic properties of the crystal are presented in Sections 2 and 3 of Chapter 1. Chapter 2 follows with the history and background work in the fields of the photorefractive effect, hologram formation, and grating diffraction. The experimental work, including the setup, the techniques, and the actual data, make up Chapter 3. The underlying theory is presented in Chapter 4, along with several predictions arising from the theory. Discussion of the results, and comparison to the previous works of other researchers, makes up Chapter 5.

II. Structure of Lithium Niobate

1. Pure Crystals

Lithium niobate is a uniaxial, non-centric crystal which is ferroelectric, that is, it possesses a non-zero spontaneous polarization. The material is ferroelectric at all temperatures below its Curie point ($T_c$) of $\approx 1200^\circ C$, where it undergoes a phase transition to its paraelectric (zero spontaneous polarization) phase [8]. In the ferroelectric phase, the crystal has a rhombohedral structure[8].

The exact ferroelectric structure consists of distorted oxygen octahedra stacked one on another along the $c$-axis (also called the optic axis), and is illustrated in Fig. 1-1. As one progresses along the $c$-axis, the interiors of the octahedra are occupied by, in turn, a lithium ion, a niobium ion, or a vacancy. The Li and Nb ions are not centered within their octahedras, but are significantly offset. This feature defines the crystal as polar and non-centric. The structure has, roughly, a hexagonal basis, so that the crystal is uniaxial. The symmetry is $C_{3v}$ [9, 10, 11, 12].
Fig. 1-1 Crystal Structure of LiNbO$_3$. a) Stacked oxygen octahedra along c-axis, with octahedra occupied by niobiums, lithiiums, and vacancies. b) Ferroelectric phase ($T < T_c$). Niobiums (large hatched) and lithiiums shown along c-axis relative to oxygen planes (solid lines). c) Paraelectric phase ($T > T_c$). Li Nb ions shift to form non-polar crystal.
The structure in the paraelectric phase is not completely determined, since the phase transition occurs only 50°C below the melting point. There is evidence that the Li ions move into the nearest oxygen planes, and the Nb ions move to centralized positions within their octahedra, thus forming a non-polar crystal [8].

Extensive studies of the ferroelectric phase have completely determined the exact cation positions. At room temperatures, the Li ion is closer to the oxygen plane in the -c direction, while the Nb is closer to its +c direction plane [9, 10, 11, 12]. The exact interior distances are listed in Appendix A, along with other miscellaneous properties, such as the dielectric constant and mobility.

The energy band gap is 3.72 eV, which corresponds to an absorption edge at approximately 3500 Å. The valence band is based on oxygen 2p\textsubscript{x} orbitals, and the conduction band on niobium 3d\textsubscript{c} orbitals [5]. The intrinsic, or dark, conductivity of the pure material is extremely low, making exact measurements difficult, but estimates range from 10⁻¹⁴ to 10⁻²⁰ (ohm-cm)⁻¹ [5, 13, 14, 15].

2. Metal Doped Crystals

Lithium niobate is readily doped with almost any of the transition metals. In fact, it is difficult to avoid accidental metal contamination of the melt during the growth process. Doping usually creates new absorption centers which alter the absorption spectrum. Figure 1-2 illustrates the changes caused by iron doping. The doping produces a large, broad peak in the spectrum at approximately 4900 Å. Other metals produce absorption peaks at other points in the spectrum [6, 7, 13, 16,
Fig. 1-2 Absorption spectra of iron-doped LiNbO$_3$. The small numbers labelling each curve indicate the iron doping concentration in %wt/mole. All samples are unannealed (reduced).
The iron-induced absorption peak is quite wide, and covers the 4880 Å laser line from an argon ion laser, which is why this wavelength was selected for the hologram writing.

Iron ions may exist in LiNbO₃ as a substitutional impurity in either of two valence states: the 2+ state or the 3+ state. From stoichiometry, the Fe²⁺ ion is allowed to occupy the Li site, while the Fe³⁺ is allowed to enter either the Li or the Nb sites [5]. However, the data on energy levels from the electronic spectra and magnetic susceptibility experiments imply that the iron ions occupy the lithium sites [5].

The ground energy levels of the iron impurities are located within the energy gap of the host crystal. The Fe²⁺ level is 1.02 eV above the valence band [18], and the Fe³⁺ level is 0.6 eV below the conduction band [5]. The resulting energy level diagram appears in Fig. 1-3.

III. Crystal Properties of LiNbO₃

The piezoelectric effect exists in twenty of the twenty-one crystal classes which lack a center of symmetry [8]. This effect is defined as a linear electric polarization which appears upon the application of stress. The converse is also true, that is, application of an electric field linearly produces a strain in the crystal.

Ten of these twenty piezoelectric classes possess a spontaneous electric polarization, and are termed the polar crystal classes [8]. These crystal classes are listed in Table 1-1. The polarization in these crystals is temperature dependent, which gives rise to the
Fig. 1-3 Energy band gap diagram for Fe:LiNbO$_3$.
(from Zylbersztejn, Ref. 18)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cubic</td>
<td>23</td>
<td>T</td>
<td>T_h</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>m3</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>432</td>
<td>O</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>43m</td>
<td>T_d</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>m3m</td>
<td>O_h</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Tetragonal</td>
<td>4</td>
<td></td>
<td>C_4</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>4'</td>
<td>S_4</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>4/m</td>
<td>C_4h</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>422</td>
<td>D_4</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>4mm</td>
<td></td>
<td>C_4v</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>42m</td>
<td>D_2d</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4/mmm</td>
<td></td>
<td>D_4h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orthorhombic</td>
<td>222</td>
<td></td>
<td>D_2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mm2</td>
<td></td>
<td>C_2v</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>mmm</td>
<td></td>
<td>D_2h</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Hexagonal</td>
<td>6</td>
<td></td>
<td>C_6</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>6'</td>
<td></td>
<td>C_3h</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>6/m</td>
<td></td>
<td>C_6h</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>622</td>
<td></td>
<td>D_6</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>6mm</td>
<td></td>
<td>C_6v</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>6m2</td>
<td></td>
<td>D_3h</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>6/mmm</td>
<td></td>
<td>D_6h</td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>

(continued on next page)
<table>
<thead>
<tr>
<th>Crystal System</th>
<th>Internat’l</th>
<th>Schoenflies</th>
<th>Symbol</th>
<th>Centro.*</th>
<th>Piezo.+</th>
<th>Pyro.#</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monoclinic</td>
<td>2</td>
<td>$C_2$</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>$C_s$</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2/m</td>
<td>$C_{2h}$</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Triclinic</td>
<td>1</td>
<td>$C_1$</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\overline{1}$</td>
<td>$C_1$</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trigonal</td>
<td>3</td>
<td>$C_3$</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\overline{3}$</td>
<td>$S_6$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>$D_3$</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3m</td>
<td>$C_{3v}$</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>$\overline{3}$m</td>
<td>$D_{3d}$</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>

**TABLE 1-1**

The 32 crystal point groups.

* Centrosymmetric, + Piezoelectric, # Pyroelectric
pyroelectric effect, so that these crystal classes are also called the pyroelectric classes. If the polarization vector of a pyroelectric crystal is bistable, that is, if the vector can be switched between at least two stable states, then the crystal is called ferroelectric [8]. Lithium niobate is such a crystal.

While in the ferroelectric phase \((T < T_c)\), the spontaneous polarization vector may exist in one of two states, either parallel or antiparallel to the optic axis. The switching of polarization states is accomplished by applying large electric fields along the c-axis while cooling the crystal from above the Curie point to a temperature below \(T_c\) [8]. This technique is called "poling", and is also used to align all polarization domains in the crystal to the same orientation. All crystals used in this work were poled by the manufacturers.

The crystal is optically birefringent, which is a direct consequence of its pyroelectricity. The crystal is uniaxial, so that the refractive indices for light polarized perpendicular to the polar axis \((n_1 \text{ and } n_2)\) are different from the refractive index for parallel polarized light \(n_3\) as

\[
 n_1 = n_2 \neq n_3
\]

The polar axis is often referred to as the c-axis, or the z-axis, and the subscript 3 refers to this same axis.

The crystal also possesses a non-linear optical susceptibility as defined by

\[
P_i(t) = d_{ijk} E_j(t) E_k(t)
\]
Here, $P_i$ is the polarization vector in the $i$ direction, $E_j$ and $E_k$ are electric fields in the $j$ and $k$ directions, and $d_{ijk}$ is the second-rank susceptibility tensor. This can be simplified by noting that

$$d_{ijk} = d_{ikj}$$

since no physical change is realized when $E_j$ and $E_k$ are interchanged.

This realization is expressed most easily by the contracted Voigt notation, which is an abbreviated notation scheme for the last two subscripts of $d$. The abbreviations are

$$\begin{align*}
xx &= 1 \\
yy &= 2 \\
zz &= 3 \\
yz &= zy = 4 \\
xz &= zx = 5 \\
xy &= yx = 6
\end{align*}$$

Use of this convention in Eq. 1.3 results in a tensor equation for the polarization with $d$ represented as a $6 \times 3$ matrix

$$P_i = d_{ik} E_k$$

This susceptibility tensor $d$ is related to the electro-optic tensor by

$$d_{jkl} = -\frac{1}{2} \frac{\varepsilon_l \varepsilon_1}{\varepsilon_0} r_{jlk}$$

The electro-optic tensor relates the electric fields to the changes in the refractive indices through the equation

$$\Delta \left( \frac{1}{n^2} \right)_{i} = \sum_{j=1}^{3} r_{ij} E_j$$

Usually only a few of the terms in a materials electro-optic tensor are non-zero, and in LiNbO$_3$ the tensor has the form [19]
The largest of the electro-optic coefficients in LiNbO₃ is r₃₃, which is more than three times the size of any of the others. Hence, the most efficient holographic gratings are made by having the grating planes perpendicular to the c-axis (z-axis). This produces an electric field in the c direction, and the tensor equation reduces to

\[
\frac{\Delta \left( \frac{1}{n^2} \right)_{x}}{\Delta \left( \frac{1}{n^2} \right)_{y}} = r_{13} E_{z} \\
\Delta \left( \frac{1}{n^2} \right)_{z} = r_{33} E_{z}
\]

The second equation can be rewritten, since the quantity of interest is not nₓ, but the change in nₓ, which is usually termed \( \Delta n_{x} \). Since \( \Delta n_{z} \) is usually on the order of \( 10^{-3} \) \( (\Delta n_{z} \ll n_{z}) \), it can be approximated as

\[
\Delta n_{z} = \frac{1}{2} n_{3}^{3} r_{33} E_{z}
\]

Maximum advantage of these equations is obtained by having the probe beam polarized parallel to the c-axis, where the beam will "see" the larger r₃₃ coefficient. Also, by polarizing the writing beams perpendicularly to the optic axis, the beam coupling effects [20] can be minimized.
Chapter 2

Current Status of Understanding of Holographic Storage in LiNbO$_3$

Holographic storage was first thought to be possible in LiNbO$_3$ when it was discovered that the optical damage to the refractive index produced by an incident laser beam was reversible, and represented a phenomenon of interesting and useful possibilities. It was soon found that illumination of the material with the interference pattern of two laser beams created a non-uniform charge distribution which copied the spatial pattern of the illuminating light. Through the electro-optic effect, this charge pattern gave rise to a modulated refractive index, which is essentially a holographic phase grating. Thus, information in holographic form could be stored in the crystal and read out at a later date with an appropriate reference beam. This process was, at one time, thought to be of great potential for high density computer memories, but recent interest has turned to other methods instead.

Much work was done in the field of holographic storage in crystals, and several other crystals besides LiNbO$_3$ were found to be usable. LiTaO$_3$, a related compound, was found to behave similarly in this respect. In both of these substances, it was found that the charge transport occurred automatically upon illumination, without the necessity of applying any external electric fields. This implied the existence of some process in the crystal which was providing a transport mechanism for the photo-excited charges. Other crystals were found that could record holograms, but lacked this mechanism, and these crystals required the application of external fields in order to record the holograms with
any efficiency. BaSiO$_{12}$ is such a crystal, and much work was done with it because of its greater sensitivity to the recording light than LiNbO$_3$ [21].

I. The Holographic Storage Process

The underlying physics of the holographic storage phenomenon involves several specific processes. First, there must exist in the crystal, donors which can be photo-excited by an optical interference pattern. Next, some transport process must be operational in order to move the photo-excited electrons to form the space-charge field. At this point, the electrons require some form of trap site in order to "freeze" the charge distribution and allow the space-charge field to persist for some useful period of time.

The specific form of the grating depends on several parameters, and there has been considerable work done on grating formation dynamics. A suitable formation theory must include all mechanisms of electron transport, and must also take into account the fact that a buildup of the space-charge fields tends to inhibit the further buildup of change; therefore, the writing process will saturate for sufficiently long exposure times.

Finally, the specific form of the grating, as found from the formation theory, must be utilized in conjunction with a coupled-wave theory of diffraction in order to arrive at the diffraction efficiency of the grating, which is the physically measurable quantity of interest.

This chapter presents the background work in these three areas
(electron transport, grating formation dynamics, and diffraction theory), and briefly summarizes the most crucial findings as related to this work.

II. Theory of Electron Transport

The two most common causes of electron movement are diffusion and drift. Diffusion is caused by a non-uniform spatial distribution and the diffusion current has the form

\[ J_{\text{diff}} = -D \frac{\partial n}{\partial x} \]  \hspace{1cm} (2.1)

where \( D \) is the diffusion constant, and \( n \) is the conduction electron density. Drift is the movement of charges in an electric field, and the drift current has the form

\[ J_{\text{drift}} = ne \mu E \]  \hspace{1cm} (2.2)

where \( e \) is the charge of one electron, \( \mu \) is the electron mobility, and \( E \) is the electric field.

Initially, it was thought that the predominant electron transport mechanism in \( \text{LiNbO}_3 \) was either drift or diffusion, and that it was possible to explain most of the experimental data by means of drift and diffusion \([4, 22]\). An early work by F.S. Chen \([23]\) showed that the effect then called "reversible optical damage" was actually an example of charge migration, causing variations in the refractive index through the electro-optic effect. At the time, Chen hypothesized an "internal drift field" \((6.7 \times 10^4 \text{ V/cm})\) of unspecified origin that was responsible for causing the migration of the photo-excited charges.
Chen showed that the change in refractive index, $\Delta n$, increased upon illumination until a saturation value was reached. His data also showed that this saturation value of $\Delta n$ increased with increasing intensity [23]. However, there was no obvious mechanism available to account for such a large static internal field as required by his hypothesis. There were also no hypotheses presented for the donor or trap mechanisms.

In 1973, Clark, et al [5], presented their work on the electronic structure of iron-doped LiNbO$_3$. This work first identified the electron donor sites as transition metal impurity atoms, which exist in almost all samples. They also identified the orbitals comprising the conduction and valence bands. By a variety of experiments, the explicit electronic transitions involved in the photo-excitation process were identified and linked to the features in the absorption spectra. For the iron doped samples in particular, the prime transition involved in the photorefractive process was determined to be an intervalence transfer from the Fe$^{2+}$ d$_{x^2}$ orbital to the Nb$^{5+}$ d$_{x^2}$ orbital. The Fe$^{3+}$ ions were also shown to act as the acceptor sites necessary to the theory.

A wide variety of transition metals were found to exhibit this effect when in the crystal, including iron, cobalt, manganese, rhodium, nickel, and copper [6, 7, 16, 24, 25]. These dopants were found to have different absorption spectra, and slight differences in behavior upon excitation.

Although the work of Staebler and Amodei was very important to the overall picture, one key phenomenon not yet explained was the generation of a steady-state, uniaxial current in the crystal when illuminated by light of the proper wavelength [23]. If the crystal was in an open-
circuit condition, a photo-voltage was set up between the \( z_c \) faces of the crystal that could be as much as several KV/cm. Various theories sought to attribute this to an internal electric field in the crystal, but estimates of the size of the field went up to \( 10^5 \) V/cm, and the origins of such fields could never be completely justified.

It was soon realized that the existence of the photovoltage in LiNbO\(_3\) was not the result of ordinary material characteristics, but was an example of a previously unknown phenomenon. This was described by von der Linde and Glass [13, 26], and termed the "bulk photovoltaic effect". In this model, the photovoltages and photocurrents were attributed to asymmetries in the local crystal potential surrounding the electron donor sites. Electrons photoexcited in such situations will have a preferential velocity relative to the crystal c-axis, hence creating a photocurrent, in the short-circuit condition, or a photovoltage in the open-circuit condition. This model has been able to satisfactorily explain almost all existing data, and has met with wide acceptance. The theory of the bulk photovoltaic effect also forms the basis of the theory of this thesis, and modifications to the BPVE theory will be presented in the discussion section (Chapter 4).

III. Hologram Formation Process

The BPVE is utilized to create a holographic grating in the following manner. An interference pattern is created by intersecting two laser beams in the crystal. These beams must be at a wavelength which falls within the absorption spectra of the metal dopant present in the crystal. Also, for optimum effectiveness, the fringes of the interference pattern should be perpendicular to the optic axis of the crystal.
In materials not possessing a spontaneous polarization, an electric field can be applied, with the field aligned perpendicularly to the interference fringes [21].

When the crystal is illuminated by such an intensity pattern, the donor sites are photo-excited in a corresponding spatial pattern (see Fig. 2-1). Upon excitation, the electrons may also be transported some distance along the c-axis due to drift or the BPVE. The net result is the creation of an electron density pattern that matches the initial illumination pattern, except for a possible shift along the optic axis. These electrons will tend to remain in this nonuniform density pattern for quite a while because the electrical conductivity of these crystals is exceedingly low. Estimates and measurements of the dark conductivity range from $10^{-14}$ to $10^{-20}$ mho/cm [13, 14, 27].

The nonuniform electron density, $n(x)$, gives rise to an electric field (by Gauss' Law), which is termed a "space charge" field, $E(x)$.

$$\tilde{E}(x) = \frac{\partial n(x)}{\partial x} \quad 2.3$$

Since these crystals exhibit the linear electro-optic effect, the space charge field creates a modulated term, $\Delta n(x)$, in the refractive indices of the crystals.

$$\Delta n(x) = \frac{1}{2} r_{ij} E(x) \quad 2.4$$

The modulation of the refractive index is a thick holographic phase grating, and will diffract light of the proper wavelength when incident at the proper Bragg angle.
Fig. 2-1 Spatial relationships between illumination, electron density, space-charge field, and refractive index variation as a result of hologram formation. (Phase shift of $\pi$ assumed between illumination and electron density.)
IV. Diffraction by a Thick Phase Transmission Grating

The diffraction theory underlying the diffraction of beams by thick holographic gratings has gone through many improvements in the past few years. The general starting point is a coupled wave theory, as in the paper by Kogelnik [28]. Basically, the method consists of writing the wave equations for the two waves, the incident beam and the diffracted one, and then incorporating a coupling constant to couple the two equations. Thus the diffraction efficiency of the grating may be found. Referring to Fig. 2-2, the two coupled equations appear as

\[ \cos \theta R' + \alpha R = -j\kappa S \]  \hspace{1cm} 2.5a

and

\[ (\cos \theta - \frac{K}{\beta} \cos \phi) S' + (\alpha + j\nu) S = -j\kappa R \]  \hspace{1cm} 2.5b

with the coupling coefficient \( \kappa \) defined as

\[ \kappa = \frac{\pi \Delta n}{\lambda} - j \frac{\alpha_1}{2} \]  \hspace{1cm} 2.6

where \( \alpha_1 \) is the modulation of the absorption coefficient, and \( \lambda \) is the wavelength of the light. The solution to this general case is not exceedingly illuminating, but the particular solution for the case of the transmission phase grating is very useful. The diffraction efficiency, \( \eta \), is defined as the intensity of the diffracted beam divided by the intensity of the incident beam (lossless case), and can be shown to be equal to [28]

\[ \eta = e^{-\alpha z} \sin^2\left(\frac{\pi d \Delta n}{\lambda \cos \theta}\right) \]  \hspace{1cm} 2.7

where \( \alpha \) is the attenuation coefficient, \( d \) is the thickness of the crystal, \( \lambda \) is the wavelength of the light, and \( \theta \) is the Bragg angle for \( \lambda \).
Fig. 2-2 Model of thick phase transmission grating used by Kogelnik (Ref. 28).
The assumptions made in this derivation are that only the first order of diffracted light is considered, all other orders being assumed to not meet the Bragg condition, and that the incident wave enters the grating at the Bragg angle.

It is important to note here that the process is insensitive to the phase of the grating, i.e., shifting the grating in the x-direction does not affect the diffraction efficiency at all. In fact, unless some sort of interferometric arrangement is used, it is not possible to sense the phase of the grating.

The dynamics of the electron transport process lead to an eventual saturation of $\Delta n$. This saturation level is proportional to the intensity of the writing beams. The behaviour of $\Delta n$ versus exposure (Intensity * time) is illustrated in Fig. 2-3a. $\Delta n$ is initially linear with exposure, but as $\Delta n$ increases, it begins to asymptotically approach its saturation level. Using this dependence, and Eq. 2.8 for the relation between $\Delta n$ and $\eta$, the dependence of $\eta$ on exposure may be found, and this relationship is illustrated in Fig. 2-3b. During the initial linear part of $\Delta n$ vs $I^*t$, $\eta$ follows a $\sin^2$ form very closely. As $\Delta n$ begins to approach its saturation value, the periods of $\eta$ become more widely spaced, until, when $\eta$ reaches saturation and becomes constant, $\eta$ stops oscillating and also becomes constant.

The initial region of these curves is of prime interest. By writing holograms with only small exposures, one remains in the early linear portion of the $\Delta n$ vs $I^*t$ curve, and so the more complicated dependences of high exposures are avoided. Also, if the exposure is further limited such that $\Delta n$ is small enough to assure the validity of the assumption
Fig. 2-3 Time behavior of $\Delta n(t)$ and $\eta(t)$ during writing and decay.
\[ \sin^2 \left( \frac{\pi d \Delta n}{\lambda \cos \theta} \right) \approx \left( \frac{\pi d \Delta n}{\lambda \cos \theta} \right)^2 \]

then the process is said to be in the quadratic region. Operation in this region is of interest because it allows a simple mathematical description and treatment.

V. Summary

Taken together, these individual results form the process of hologram recording. When initially illuminated by the optical interference pattern, the electrons are photo-excited proportionally to the spatial dependence of the light. Then, the electrons migrate out of the high intensity areas to the dark areas by some combination of drift, diffusion, and the BPVE. Once reaching the dark regions, the retrapping of the electrons gives rise to a non-uniform charge distribution. The inclusion of the BPVE means that the average electron migration length is affected by the asymmetries of the potential of the Fe$^{2+}$ site. This leads to a phase shift between the initial illumination pattern and the resulting space charge field that is partially dependent on the magnitude of the asymmetry of the potential of the donor.

Once the space charge field exists, the crystal's electro-optic coefficients determine the refractive index variation $\Delta n$. This index variation can then be viewed as a volume holographic grating of the phase transmission type. The diffraction efficiency, $\eta$, of this type of holographic grating has been analysed and for small values of $\Delta n$, one finds that $\eta$ can be easily related to $(\Delta n)^2$. Thus, by studying the behaviour of the diffraction efficiency, one can learn much about the behaviour of the space charge field, and therefore deduce information
about the electron transport mechanisms. This method of measurement has been used before, but only extremely infrequently, and it has never been utilized to its full potential.
Chapter 3
Experimental Work

I. Experimental Setup

The experimental set-up for the writing and decay monitoring of the holograms is shown in Fig. 3-1. The argon writing beams are at a wavelength of 4880 Å, and intersect in the crystal at a writing angle of 12°, so that each beam was incident at an angle of 6°. The bisector of the writing angle was normal to the crystal face. The writing plane, formed by the two writing beams, also contains the c-axis of the crystal. The writing beams were polarized perpendicular to the crystal axis. The helium-neon probe beam ($\lambda = 6328$ Å) was also in the writing plane, but was incident at its own Bragg angle of 7.23°. The He-Ne beam was polarized parallel to the crystal c-axis. The writing beam diameter was 3 mm, while probe beam diameter was 1 mm. The probe beam was centered on the writing area [29].

The entire experiment was constructed on an optical isolation bench, and covered by a plexiglass box. This was done to prevent mechanical vibrations and air currents from affecting the hologram formation process. The process has been found to be severely disturbed if preventative measures were not taken [20, 30, 31].

The temperature of the crystal can be considered constant since the writing process takes very little time. During the decay, the temperature was constant to within 2° C. All work was done at room temperature.
Fig. 3-1 Layout of Experimental Setup
Since the depth of the hologram affects the diffraction efficiency [32], all crystals used had the same thickness of 1 mm, although their other dimensions varied.

Electric fields, when applied, were parallel to the c-axis. The field was applied through silver paint electrodes on the ± c-axis faces of the crystal.

During the entire experiment, all extraneous light was kept away from the crystal to prevent interference with the processes under investigation.

The intensity of the He-Ne probe beam was kept at least two orders of magnitude lower than the writing beam intensities to avoid interference. Since the absorption at 6328 Å is also much lower than at 4880 Å, the total absorbed energy at 6328 Å is less than at 4880 Å by at least three orders of magnitude. The writing wavelength of 4880 Å was chosen because it is near the peak in the photocurrent spectrum (see Fig. 1-2).

Control of all three beams (two argon and the He-Ne) was by means of three remote controlled electronic shutters. The shutters could be operated independently or linked together for simultaneous action. There was also a timing circuit which could open a shutter for two seconds once in a preselected period. The period was controllable from 0.5 minutes to 60.25 minutes.

All equipment was mounted on the optical bench, except for the shutter and the electronics. This kept the experiment mechanically isolated. The lasers used were a Coherent Radiation Model 528 Argon Ion, operated single-line with an intracavity etalon, for the $\lambda = 4880 \, \text{Å}$
writing beam, and a Spectra Physics Model 212 Helium-Neon for the probe beam. The detectors used were two Coherent Radiation Model 212 Power Meters, and their outputs were recorded on an Omniscribe Dual-Channel Strip Chart Recorder. The shutter system was homemade, and hung from overhead supports extending through small ports in the ceiling of the plexiglass enclosure. The shutters were designed so as to minimize air disturbance when activated.

II. Experimental Procedure

Before using a crystal in an experiment, all previous perturbations of the refractive index were erased by exposing the crystal to broad-spectrum high intensity light for several hours. The crystal was then allowed to "relax" for several hours (approx. 24) in complete darkness. The refreshed crystal was then installed in its mount.

The experiment was started by turning on the He-Ne beam. This probe was on continuously throughout the entire writing phase. The writing phase (formation of the grating) began when the two argon beams were simultaneously turned on. The diffraction efficiency, $\eta$, of the grating was continuously monitored by the probe beam until $\eta$ reached a predetermined point at which writing was halted by closing the writing beam shutters. At this point, any electric fields were turned on, and the He-Ne probe beam was switched from continuous mode to pulsed mode. The probe pulses had a duration of two seconds, which was long enough to allow recording on the strip chart recorder, but were short enough to keep interference with the hologram to a minimum. The time at which the writing beams were terminated is defined as the start of decay. The decay was allowed to continue, with periodic checks of $\eta$ by the probe
beam, until \( \eta \) became too small to measure and was lost in the noise. The noise referred to here is electronic noise in the detector and recorder circuits. The limits of detectability for the equipment used was approximately \( \eta = 10^{-2}\% \), or \( 10^{-4} \). The intensity of the probe beam was 50 \( \mu \text{W}/\text{cm}^2 \), giving a minimum detectable signal of \( I = 5\text{nW}/\text{cm}^2 \).

One of the prime differences between experiments was the value of \( \eta \) that the hologram was allowed to reach before the decay was started. In some runs, the writing process was continued until \( \eta \) had reached its saturation value. Since a long decay time was then needed for \( \eta \) to reach its quadratic region, these runs were used to measure the long-time asymptotic decay rate. For measurements of the short-time decay rate, only holograms with \( \eta \) still in the initial quadratic region were used.

Once recorded on the strip charts, the data was analyzed manually, with the aid of programmable calculators. The recorded data was in the form of a measurement of the transmitted beam intensity and a measurement of the diffracted beam intensity. The diffraction efficiency was determined by dividing the diffracted intensity by the sum of the transmitted and diffracted intensities. This compensated for any fluctuations in the incident intensity of the probe beam.

III. Experimental Data

The experimental data is in three sections. The first section simply shows a relationship between the Fe\(^{2+}\) concentration and \( a \), the absorption coefficient at 4880 Å. The second section is concerned with the basic time dependence of the decay rate as a function of the dopant
(Fe) concentration in the crystal. Much of this data has been previously presented in past works by this author [15, 33]. In conjunction with this, data are also presented concerning the relationship between the dark conductivity and the absorption coefficient at \( \lambda = 4880 \) Å. The third section presents the results of experiments investigating the effects on the decay rate of an electric field applied either parallel or anti-parallel to the crystal c-axis.

1. Dopant Level vs. Absorption

The dopant concentrations for the crystals used in this work, and their optical absorption coefficients (at 4880 Å) are listed in Table 3-I. The dopant concentrations for the crystals have been determined by previous researchers [34]. Absorption is not noticeably influenced by the Fe\(^{3+}\) concentration, and a plot of \( \alpha \) vs. Fe\(^{3+}\) level shows no organized behaviour. However, when \( \alpha \) is graphed as a function of the Fe\(^{2+}\) concentration, as in Fig. 3-2, a significant relationship is revealed. The absorption is linear with \( \ln(\text{Fe}^{2+}) \). Although no crystals with doping higher than 0.25% were used in this work, Clark, et al [5], have noted that crystals with doping levels of 3% have essentially a black color, and extremely high absorption. The doping level of such a crystal was estimated at \( 10^{21} \text{ atoms/cm}^3 \), and the conductivity was approximately \( 10^{-7} \text{ (ohm-cm)}^{-1} \).

The crystals represented in Table 3-I and Fig. 3-2 comprise the entire stock of doped crystals used in these experiments, and more experimental data on very high doping effects ( \( > 0.25\% \) ) is, at present, unavailable. Although pure \( \text{LiNbO}_3 \) is readily obtainable, there are no manufacturers of the doped crystals at this time.
<table>
<thead>
<tr>
<th>Crystal</th>
<th>$Fe_{T}$ (cm$^{-3}$)</th>
<th>$Fe^{2+}$ (cm$^{-3}$)</th>
<th>$Fe^{3+}$ (cm$^{-3}$)</th>
<th>$Fe^{2+}/Fe_{T}$ (%)</th>
<th>$\alpha$ (cm$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>$6 \times 10^{18}$</td>
<td>$4.5 \times 10^{17}$</td>
<td>$5.55 \times 10^{18}$</td>
<td>6.5</td>
<td>3.83</td>
</tr>
<tr>
<td>7</td>
<td>$2 \times 10^{19}$</td>
<td>$4.1 \times 10^{17}$</td>
<td>$1.96 \times 10^{19}$</td>
<td>2.0</td>
<td>4.03</td>
</tr>
<tr>
<td>4</td>
<td>$8 \times 10^{18}$</td>
<td>$4.9 \times 10^{17}$</td>
<td>$7.5 \times 10^{18}$</td>
<td>6.0</td>
<td>5.41</td>
</tr>
<tr>
<td>14</td>
<td>$5.6 \times 10^{19}$</td>
<td>$5.2 \times 10^{17}$</td>
<td>$5.55 \times 10^{19}$</td>
<td>9.09</td>
<td>7.54</td>
</tr>
<tr>
<td>5</td>
<td>$4 \times 10^{19}$</td>
<td>$8.2 \times 10^{17}$</td>
<td>$3.92 \times 10^{19}$</td>
<td>2.0</td>
<td>5.53</td>
</tr>
<tr>
<td>16</td>
<td>$1 \times 10^{20}$</td>
<td>$5 \times 10^{18}$</td>
<td>$9.5 \times 10^{19}$</td>
<td>2.05</td>
<td>16.96</td>
</tr>
<tr>
<td>13</td>
<td>$5.6 \times 10^{19}$</td>
<td>$3 \times 10^{18}$</td>
<td>$5.3 \times 10^{19}$</td>
<td>52.4</td>
<td>20.16</td>
</tr>
<tr>
<td>8</td>
<td>$2 \times 10^{19}$</td>
<td>$4.51 \times 10^{18}$</td>
<td>$1.55 \times 10^{19}$</td>
<td>22.</td>
<td>13.9</td>
</tr>
<tr>
<td>15</td>
<td>$1.4 \times 10^{20}$</td>
<td>$4.6 \times 10^{18}$</td>
<td>$1.35 \times 10^{20}$</td>
<td>33.3</td>
<td>24.05</td>
</tr>
<tr>
<td>9</td>
<td>$2 \times 10^{19}$</td>
<td>$5.7 \times 10^{18}$</td>
<td>$1.43 \times 10^{19}$</td>
<td>28.</td>
<td>25.1</td>
</tr>
<tr>
<td>12</td>
<td>$4 \times 10^{19}$</td>
<td>$8 \times 10^{18}$</td>
<td>$3.2 \times 10^{19}$</td>
<td>19.5</td>
<td>27.7</td>
</tr>
<tr>
<td>6</td>
<td>$4 \times 10^{19}$</td>
<td>$9.84 \times 10^{18}$</td>
<td>$3.02 \times 10^{19}$</td>
<td>24.</td>
<td>29.4</td>
</tr>
</tbody>
</table>

**TABLE 3-I**

Dopant Concentrations and Absorption Coefficients
Fig. 3-2 Measured optical attenuation at $\lambda=4880$ Å vs. $\text{Fe}^{2+}$ concentration.
This relation between $a$ and the $\text{Fe}^{2+}$ level is not surprising, since it has already been implied by the absorption spectra for doped and undoped crystals (Fig. 1-2). The spectra do show a slight increase in absorption for doping with an annealing treatment (unreduced Fe ions, i.e., primarily doped with $\text{Fe}^{3+}$), but the large characteristic peak near 4880 Å only appears after the crystal has been reduced (i.e., is unannealed), so that the $\text{Fe}^{2+}$ density is significantly increased.

2. Time Varying Decay Rate

The basic form of the experimental data is a plot of the hologram decay, presented as $\ln \eta$ vs. the decay time $t$. A few decay curves appear in Figs. 3-3 to 3-8. These six experiments are only a few examples of the total number of experiments performed. Immediately it can be noted that all experiments done on moderately to highly doped crystals show strong deviations from simple exponential behaviour. Since this type of information forms the basis of this work, extreme care was taken to ascertain that this was a true effect occurring within the crystal and not an experimental artifact.

All electronics associated with the detection of the probe beam were checked to assure that their responses to the input beams were linear with intensity. Also, the fact that the low-doped crystals produced essentially exponential decay curves is more evidence that the electronics were in fact operating properly and that the observed phenomenon truly does occur because of some process internal to the moderately-to-highly doped crystals.

The time dependence of $\eta$ cannot be easily described by a simple
Fig. 3-3 Two decay curves taken using X9 (0.05%Fe) showing varying decay rate and also showing the reproducibility of the curves.
Fig. 3-4 Decay curve of hologram written to saturation, showing the "retracing" of the writing curve, and constant slope of final decay period. (Crystal X12, 0.1%Fe)
Fig. 3-5 Decay curve for hologram in a low-doped crystal. Slope is essentially constant. (Crystal X4, 0.02%Fe)
Fig. 3-6 Decay curve for a moderately-doped crystal.
(Crystal X9, 0.05\%Fe)
Fig. 3-7 Decay curve for a moderately-doped crystal.
(Crystal X9, 0.05%Fe)
Fig. 3-8 Decay curve for a strongly-doped crystal.
(Crystal X6, 0.1%Fe)
exponential, or even a sum of exponentials. The \( \ln \eta \) vs. \( t \) plot shows a continuous smooth curvature, and gives no indication of exponential behavior. However, if the slope of \( \ln \eta \) vs. \( t \), \( \left( \frac{2}{\tau} \ln \eta \right) \) is plotted as a function of the decay time, the time dependence becomes clear. An example of the slope, \( r(t) \), is plotted versus decay time in Fig. 3-9. The decay rate plots of short experiments can be fit by two single exponential functions. Long time experiments show the same two exponential functions, but also show the final constant slope indicative of the final dark conductivity.

The decay rate \( r(t) \) initially exhibits one exponential decay rate for the first thirty to sixty seconds, at which time the second slower decay rate becomes dominant. This slower exponential continues until reaching a constant value, indicating a final constant decay rate. This change from the second exponential to the constant rate usually occurs at approximately \( 10^4 \) seconds. The decay rate stays at this constant level for all further decay times.

The decay rate \( r(t) \) can be fit well to the form

\[
    r(t) = a_1 e^{-b_1 t} + a_2 e^{-b_2 t} + a_3. \tag{3.1}
\]

The fit of Eq. 3.1 to the experimental data is illustrated in Figs. 3-10 through 3-12. Figure 3-10 shows the best match obtained between Eq. 3.1 and the data, Figure 3-11 shows the average fit quality, and Figure 3-12 shows a low quality fit.

The parameters \( a_1, a_2, a_3, b_1 \) and \( b_2 \) are extracted from the data curves in the following manner. First, exponential lines are fit to the
Fig. 3-9 Schematic of parametric model used to characterize the data, showing the definitions of the five parameters $a_{1-3}$, and $b_{1-2}$. 
Fig. 3-10a Decay rate of hologram in Crystal X6 (0.1%Fe) vs time. The initial section of an experiment, showing the transition from initial slope to intermediate slope. High quality fit between theory and experiment. This decay curve is continued in Figs. 3-10b, c.
Fig. 3-10b Intermediate section of the decay curve of Fig. 3-10a. Longer time scale shows intermediate slope.
Fig. 3-10c Final section of the decay curve of Figs. 3-10a,b. Transition from intermediate slope to final slope is shown.
Fig. 3-11 Decay rate of hologram in Crystal X6 (0.1%Fe). Initial and intermediate slopes are shown. Average quality fit between theory and experiment.
Fig. 3-12 Decay rate of hologram in X16 (0.25%Fe). Initial intermediate slopes are shown. Poor quality fit between theory and experiment.
initial and middle sections of the decay curve. Also, for decay curves that are long enough, and have reached the constant slope region, a constant line may be added. These three lines intercept the vertical axis at the intercept points, \( y_1 \), \( y_2 \), and \( y_3 \). The slopes of the two exponential lines are termed \( \lambda_1 \) and \( \lambda_2 \). The three \( a \) and two \( b \) parameters are now

\[
a_3 = y_3 \\
a_2 = y_2 - y_3 \\
a_1 = y_1 - y_2 \\
b_2 = \lambda_2 \\
b_1 = \lambda_1 - \lambda_2
\]

These parameters are different for each crystal, and are also affected by any electric fields applied during the decay. The values for the crystals used with zero applied field (E=0) are summarized in Table 3-II.

Table 3-II shows that, in almost all cases, all decay parameters increase in magnitude for increasing \( \text{Fe}^{2+} \) concentrations. The \( a_3 \) parameters for \( \text{Fe}^{2+} \) levels greater than \( \approx 3.0 \times 10^{18} \) seem to not support this conclusion, although \( a_3 \) parameters for lower doping levels, as well as almost all other parameters, do support this conclusion. For crystals of low \( \text{Fe}^{2+} \) concentration, the decay rate was essentially constant, that is, \( a_1 = a_2 = 0 \).

Several experiments were performed using different writing beam
<table>
<thead>
<tr>
<th>Crystal</th>
<th>Fe^{2+} concen.</th>
<th>$a_1$</th>
<th>$b_1$</th>
<th>$a_2$</th>
<th>$b_2$</th>
<th>$a_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>7*</td>
<td>$4.1 \times 10^{17}$</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>$8.0 \times 10^{-7}$</td>
</tr>
<tr>
<td>11*</td>
<td>$4.5 \times 10^{17}$</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>$2.05 \times 10^{-6}$</td>
</tr>
<tr>
<td>4*</td>
<td>$4.9 \times 10^{17}$</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>$1.35 \times 10^{-5}$</td>
</tr>
<tr>
<td>5*</td>
<td>$8.2 \times 10^{17}$</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>$5.87 \times 10^{-6}$</td>
</tr>
<tr>
<td>14*</td>
<td>$5.2 \times 10^{17}$</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>$1.68 \times 10^{-5}$</td>
</tr>
<tr>
<td>13*</td>
<td>$3.0 \times 10^{18}$</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>$9.55 \times 10^{-5}$</td>
</tr>
<tr>
<td>12</td>
<td>$8.0 \times 10^{18}$</td>
<td>$3.12 \times 10^{-4}$</td>
<td>$7.55 \times 10^{-4}$</td>
<td>$5.6 \times 10^{-5}$</td>
<td>$5.26 \times 10^{-5}$</td>
<td>$2.15 \times 10^{-5}$</td>
</tr>
<tr>
<td>8</td>
<td>$4.5 \times 10^{18}$</td>
<td>$4.48 \times 10^{-4}$</td>
<td>$4.66 \times 10^{-3}$</td>
<td>$8.2 \times 10^{-5}$</td>
<td>$2.7 \times 10^{-4}$</td>
<td>$7.2 \times 10^{-6}$</td>
</tr>
<tr>
<td>9</td>
<td>$5.7 \times 10^{18}$</td>
<td>$4.66 \times 10^{-3}$</td>
<td>$2.33 \times 10^{-2}$</td>
<td>$4.73 \times 10^{-4}$</td>
<td>$3.85 \times 10^{-4}$</td>
<td>$2.75 \times 10^{-5}$</td>
</tr>
<tr>
<td>6</td>
<td>$9.84 \times 10^{18}$</td>
<td>$6.63 \times 10^{-3}$</td>
<td>$1.03 \times 10^{-1}$</td>
<td>$1.29 \times 10^{-3}$</td>
<td>$6.85 \times 10^{-3}$</td>
<td>$1.94 \times 10^{-5}$</td>
</tr>
<tr>
<td>16</td>
<td>$5 \times 10^{18}$</td>
<td>$1.64 \times 10^{-2}$</td>
<td>$4.93 \times 10^{-2}$</td>
<td>$7.9 \times 10^{-3}$</td>
<td>$5.34 \times 10^{-2}$</td>
<td>not reached</td>
</tr>
</tbody>
</table>

**TABLE 3-II**

Decay rate parameters for zero applied electric field (in units of seconds^{-1})

* These crystals exhibited a simple exponential decay, slope given as $a_3$ parameter.
intensities, with the total exposure for each experiment kept constant. No significant variations were found in any of the decay parameters. The intensity was varied over 3 orders of magnitude, from 100 µW/cm² to 100 mW/cm².

Two lightly doped crystals showed very anomalous decay curves. A sample of this anomalous behaviour appears in Fig. 3-13. This type of behaviour is exhibited by crystal X13, which has 0.014% Fe (partially reduced), and by crystal X7, which has 0.05% Fe (annealed). As can be seen, the curves look almost random. The reasons behind this behaviour are unknown.

The relations between the decay parameters and the Fe²⁺ doping density are illustrated in Fig. 3-14. These figures show a somewhat linear relation between each of the parameters and the doping density.

3. Electric Field Dependence

Crystals showing a variable decay rate also exhibit a variation in the decay parameters with the applied electric field. In crystals showing an essentially time-invariant decay rate, no variation in the decay rate with field was found. The relation between the field and the parameters is not a simple one, for in almost every case, the parameters have a maximum for E = 0, two minima for E = ±(1.5 - 3.0) KV/cm, and then high values again for high field magnitudes. This behaviour gives the curves a characteristic "W" shape. The values of field used were not greater than ±6 KV/cm because of equipment limitations.

The actual data is shown is Figs. 3-15 through 3-17. Each figure shows the variation in a single parameter as a function of field, for
Fig. 3-13 Diffraction efficiency vs decay time for Crystal (0.015%Fe). This type of anomalous behavior was also seen in X2 (0.02%Fe). This was not exactly repeatable, and is unexplained.
Fig. 3-14 Decay parameters vs Fe$^{2+}$ doping concentration.

$a_1$, $a_2$, $a_3$, $b_1$, $b_2$. 
Fig. 3-15 Plots of the decay parameters as a function of the applied electric field for Crystal X16 (0.25%Fe).
Fig. 3-16 Plots of the decay parameters as a function of the applied electric field for Crystal X9 (0.05%Fe).
Fig. 3-17 Plots of the decay parameters as a function of the applied electric field for Crystal X6 (0.1%Fe).
each crystal examined. The "W" shape is very visible for most curves. There is also a slight asymmetry to the curves, that is, the parameter values for negative fields are generally a bit larger in magnitude than the values for positive fields.
Chapter 4

Theory

The data presented in Chapter 3 clearly established that the decay rate of the diffraction efficiency of holograms in moderate-to-strongly doped crystals exhibits a strong time dependence. The time dependence is not due to exceeding the quadratic region of the $\eta$ vs. exposure curve. By keeping the total exposure low, complicated decay behaviour, such as that illustrated in Fig. 2-3d, does not occur, and the validity of the relationships

$$\Delta n \propto it \quad \text{and} \quad \eta \propto (\Delta n)^2$$

are assured. Hence, the time dependence of $\frac{3}{3t}(\ln \eta)$ is due to a time dependence in $\frac{3}{3t}(2 \ln \tilde{E})$ where $\tilde{E}$ is the space charge field. Additional confirmation of this conclusion is established by the fact that holograms in lightly doped crystals show a constant decay rate when written to the same initial $\eta$. Thus, it is valid to conclude that the time dependence is an effect due to the higher doping present in some crystals.

Since $\Delta n$ is given by

$$\Delta n = \frac{1}{2} n_3^3 r_{33} \tilde{E}_z$$

then a time variation in $\Delta n$ should also exist in $r_{33}$ or $\tilde{E}_z$. Assuming that the crystal properties, including the electro-optic coefficients, do not change significantly, if at all, due to the formation of a hologram, then the time dependence must be present in the space charge field, $\tilde{E}_z$. 
The space charge field is dependent on the charge density, \( \rho \), and the material permittivity, \( \varepsilon \), as

\[
\nabla \mathbf{E} = -\frac{\rho}{\varepsilon}
\]

Now, \( \varepsilon \) does undergo changes due to the hologram grating, since

\[
\varepsilon = \left( \frac{n_o}{\Delta n} \right)^{1/2}
\]

\[
\approx \sqrt{\frac{n_o}{\Delta n}} \approx \frac{\Delta n}{2 n_o}
\]

assuming that \( \Delta n \ll n_o \). This assumption is valid, since \( \Delta n \approx 10^{-3} \) and \( n_o \approx 2.3 \). It can now be seen that the changes in \( \varepsilon \) are extremely small, compared to the magnitude of the changes in the charge density, \( \rho \), and therefore \( \varepsilon \) can be assumed to be essentially time-invariant in Eq. 4.3.

The only remaining source for the time dependence of the decay rate on \( \eta \) is the decay rate of the charge density \( \rho \). The decay rate of \( \rho \) is governed by the electrical conductivity, which has both spatial and temporal dependences. The conductivity, \( \sigma \), is defined as

\[
\sigma = ne\mu
\]

where \( n \) is the electron density, \( e \) is the electronic charge, and \( \mu \) is the electron mobility. It must be noted that \( n \) is specifically the conduction electron density\(^5\).

To discover the reasons for the temporal and spatial dependences of the conductivity, it is necessary to solve for the electron density, \( n \). It is assumed that the electron mobility, \( \mu \), is constant.

The electron density is determined by using a model incorporating two types of donor sites and two types of trap sites. All four types of
sites are assumed to be at different energy levels (i.e., all levels are non-degenerate). Electrons are excited from the donors (by either thermal or optical processes) into the conduction band, and then are retrapped at one of the trapping sites. Simultaneously, when a donor releases an electron, a trap is formed, and when a trap accepts an electron, a donor is formed. Additionally, any site, donor or trap, may be situated in a crystal potential field which is asymmetric with respect to the crystal c-axis. At this point the BPVE theory of Glass and von der Linde[26] comes into play. According to this theory, the electron photo-excited from the donor will have a preferential velocity with respect to the c-axis. Likewise, the acceptors will preferentially trap electrons of certain velocities (either positive or negative velocities with respect to the c-axis). The electron velocities are assumed to be completely randomized after the first collision after excitation. Thus there are components of the electron current due to the electron's velocity when excited, and also when trapped. Additionally, there is an ionic current term due to the ionic relaxation after excitation or trapping. These current components (ionic, at excitation, and at trapping) may be treated as one single drift current produced by an "internal field", $E_{\text{INT}}$. This field does not actually exist, but the net effect of the individual current components is the same as if it did exist.

The electron density is found by first solving the rate equations for the donor sites in the crystal and then implementing these solutions as the sources in the continuity equation for the electron density. The

---

1 The estimate of the conduction electron lifetime is very crude, and it is unknown as to whether the excited electron can actually last long enough to experience multiple collisions.
two donor sites are necessary because of the initial multivalent characteristic of the iron dopant. This factor has not been recognized in any previous work, and gives rise to an additional exponential factor in the final answer. This fact, when incorporated into the bulk photovoltaic theory of von der Linde and Glass [26] and the dynamic rate equations of Kim, et al [35,36] yield solutions for the conduction electron density, and hence the hologram decay rate, which agree with the experimental data.

The introduction of this elaboration to the basic theories has not been required before now because all previous research has concentrated almost exclusively on the writing process. During writing, the photoconductivity created by the writing beams is so much greater than the dark and transient conductivities that the multivalent effects of the iron ions are completely masked. Since previous works have all focussed on the writing process, and the decay process has received only cursory attention, these effects have gone unnoticed. The new hypotheses presented here also provide a coherent framework which is consistent, not only with the experimental data of this work and previous holographic works, but is also consistent with research on the luminescence properties of iron doped lithium niobate.

I. Modified Bulk Photorefractive Effect

In the original theory of von der Linde and Glass [13,26] the Fe$^{2+}$ sites are treated as completely interchangable with Fe$^{3+}$ sites. However, a rigourous approach demands that Fe$^{2+}$ sites, in the initial state, must be considered differently from initial Fe$^{3+}$ sites. The reason for this difference is based on the principle of charge neutrality.
Consider a doped crystal that contains both types of sites, and has not been illuminated for a long period. This crystal can be considered to be in its initial, or unperturbed, state. The principle of charge neutrality is valid, not only for the entire crystal, but also for local regions surrounding particular Fe sites. The compensating charges surrounding the Fe\textsuperscript{3+} ion must have one electron's worth of charge more than the charges surrounding the Fe\textsuperscript{2+} ion. The combination of an ion and its surrounding compensation charges is defined here as a site. In the zero-input steady-state, there exist two types of compensated sites, the Fe\textsubscript{0}\textsuperscript{2+} site and the Fe\textsubscript{0}\textsuperscript{3+} site. This occurs simply because one cannot dope the crystal with one type of site only, that is, after doping, one finds that the iron naturally goes into both valence states. Although the ratio of Fe\textsubscript{0}\textsuperscript{2+} to Fe\textsubscript{0}\textsuperscript{3+} can be altered by reduction or oxidation, there is always some of each state present.

The following notation system is adopted:

\[
\text{Fe}\textsuperscript{(iron ion charge)}\textsubscript{(net site charge)}
\]

It is assumed that all compensation charges are stationary at room temperature, and are not affected by any excitation or trapping processes. Thus, photoexcitation of the iron ion produces a change in the total charge of, not only the iron ion itself, but the associated site as well.

Kurz, et al\cite{37}, have found that the electric field gradient at the iron sites is primarily parallel to the crystalline c-axis. Hence, the extra electron must be along the c-axis from its associated iron site. This rules out charge compensation by the nearest neighbor oxygen ions.
[37]. It is possible that the charge compensation is by lithium deficiencies [5], which could produce a field gradient of this type. These processes can be summarized as

\[
\begin{align*}
\text{Fe}^{2+} & + \nu_p l + \text{Fe}^{3+} + e^- \\
\text{Fe}^{3+} + e^- & + \text{Fe}^{2+} \\
\text{Fe}_0^{3+} + e^- & + \text{Fe}_-^{2+} \\
\text{Fe}_-^{2+} + \nu_p l_2 & + \text{Fe}_0^{3+}
\end{align*}
\]

The charged situations are not the normal steady-state conditions of the ions; hence, the charged sites may be expected to be at higher energies than their uncharged counterparts. Conversely, this means that the uncharged sites will be at lower energy levels, and therefore will be more stable. It should also be expected that the thermal generation rates of the two forms of donors should differ, as well as the trapping rates for the two acceptors.

II. Rate Equations for Donor Sites and Conduction Electrons

The iron concentration is specified by a variable for each of the four types of sites:

\[
\begin{align*}
N_0^2 & \rightarrow \text{density of Fe}_0^{2+} \text{ sites} \\
N_0^3 & \rightarrow \text{density of Fe}_0^{3+} \text{ sites} \\
N_-^2 & \rightarrow \text{density of Fe}_-^{2+} \text{ sites} \\
N_+^3 & \rightarrow \text{density of Fe}_+^{3+} \text{ sites}
\end{align*}
\]

In all concentration variables, the subscript indicates the net
site charge, while the superscript is two for donors and three for acceptors. Additionally, a subscript \( s \) may be added to indicate background (steady-state) values with no previous input.

The rate equations for these four quantities can immediately be simplified by noting that Fe\(^{2+}\) sites convert one-for-one with Fe\(^{3+}\) sites, and likewise for Fe\(^{3+}\) and Fe\(^{2+}\) sites. This is expressed as

\[
\frac{\partial N_0^2}{\partial t} = -\frac{\partial N_+^3}{\partial t} \quad \text{and} \quad \frac{\partial N_0^3}{\partial t} = -\frac{\partial N_-^2}{\partial t}.
\] (4.7, 4.8)

The behaviour of the crystal may be completely described by solving the continuity equations for the two donor concentrations and the conduction electron density. The first equation is written for the \( N_0^2 \) donor site density. This density is increased when an electron is trapped by a \( N_+^3 \) site, thus forming a new \( N_0^2 \) site. The \( N_0^2 \) density is decreased when an electron is excited from an \( N_0^2 \) site, thus forming an \( N_+^3 \) site and decreasing the \( N_0^2 \) density by one. Since only the decay period of the hologram is being considered, the only excitation mechanism is thermal generation, and the thermal generation rates are represented by \( g_1 \). The change is the \( N_0^2 \) density caused by thermal excitation is the generation rate times the density of excitable sites, namely, the \( N_0^2 \) density. The change in the \( N_0^2 \) density due to trapping of electrons at \( N_+^3 \) sites is equal to a trapping rate \( \kappa \) times the \( N_+^3 \) density. Thus, the \( N_0^2 \) rate equation is written as

\[
\frac{\partial N_0^2(t)}{\partial t} = -g_1 N_0^2(t) + \kappa N_+^3(t)
\] (4.9)
A similar line of reasoning involving the $N_2$ and $N_0^2$ sites results in the $N_2$ rate equation, which is

$$\frac{3N_2(t)}{3t} = -g_2N_2(t) + \kappa_2N_0^3(t) \quad 4.10$$

The equation for the time rate of change of the electron density contains the two thermal generation terms (of the form $gN$), the drift and diffusion terms ($\mu \frac{3E}{3x}$ and $D \frac{3^2}{3x^2}$), and the recombination term ($\frac{3}{\tau}$). This equation appears as

$$\left(\frac{3}{3t} + \frac{1}{\tau} + \mu \frac{3(E + \bar{E})}{3x} - D \frac{3^2}{3x^2}\right)n(x,t) = g_1N_0^2(x,t) + g_2N_2(x,t) \quad 4.11$$

where $n$ is the electron density, $g_1, g_2$ represent thermal generation rates from $Fe_0^{2+}$ and $Fe_-^{2+}$ sites respectively, and $\kappa_1, \kappa_2$ are the trapping rates. In Eq. 4.11, $\tau$ is the mean electron lifetime, $\mu$ is the electron mobility, and $E$ represents the sum of $E_{ext}$, the external applied field, and $E_{INT}$, the internal field. $\bar{E}$ is the space charge field, and $D$ is the diffusion constant of the electrons. $N_0^2(x,t)$ and $N_2^2(x,t)$ are the solutions to Eqs. 4.9 and 4.10. Since the $Fe_0^{2+}$ sites are more stable than the $Fe_-^{2+}$ sites,

$$g_2 > g_1 \quad 4.12$$

and since a positively charged acceptor is more likely to attract an electron,

$$\kappa_2 > \kappa_1 \quad 4.13$$
The steady-state densities can be found by setting the time derivatives in Eqs. 4.9 and 4.10 equal to 0. This gives the steady-state values as

\[ N_{0s}^2 = \frac{\kappa_1}{g_1} N_{t+s}^3 \] 4.14

and

\[ N_{-s}^2 = \frac{\kappa_2}{g_2} N_{0s}^3 \] 4.15

where the subscript \( s \) indicates the steady-state value. The background electron density due to these background donor densities, and is given by

\[ n_o = \tau (g_1 N_{0s}^2 + g_2 N_{-s}^2) \] 4.16

Thus the dark conductivity is

\[ \sigma_o = e \mu \tau (g_1 N_{0s}^2 + g_2 N_{-s}^2) \] 4.17

This conductivity term is spatially and temporally invariant.

The total number of sites must always remain constant, i.e.,

\[ N_0^2 + N_t^3 = N_{0s}^2 + N_{t+s}^3 \] 4.18

and

\[ N_-^2 + N_0^3 = N_{-s}^2 + N_{0s}^3 \] 4.19

Eqs. 4.18 and 4.19 can be rearranged as

\[ N_0^2 - N_{0s}^2 = -(N_t^3 - N_{t+s}^3) \] 4.18a

and
\[ N^2_- - N^2_{-s} = -(N^3_0 - N^3_{0s}) \]  

The rate equations for \( N^2_0 \) and \( N^2_- \) are identical with the rate equations for the quantities \( (N^2_0 - N^2_{0s}) \) and \( (N^2_- - N^2_{-s}) \), respectively, so that Eqs. 4.9 and 4.10 are now

\[ \frac{3}{3t}(N^2_0 - N^2_{0s}) = -g_1(N^2_0 - N^2_{0s}) + \kappa_1(N^3_+ - N^3_{+s}) \]  

and

\[ \frac{3}{3t}(N^2_- - N^2_{-s}) = -g_2(N^2_- - N^2_{-s}) + \kappa_2(N^2_- - N^3_{-s}) \]  

Substituting Eqs. 4.18a and 4.19a into 4.20 and 4.21, the rate equations become

\[ \frac{3}{3t}(N^2_0 - N^2_{0s}) = -g_1(N^2_0 - N^2_{0s}) - \kappa_1(N^2_0 - N^2_{0s}) \]  

and

\[ \frac{3}{3t}(N^2_- - N^2_{-s}) = -g_2(N^2_- - N^2_{-s}) - \kappa_2(N^2_- - N^2_{-s}) \]  

The solutions to Eqs. 4.22 and 4.23 are, respectively,

\[ N^2_0(t) - N^2_{0s} = (N^2_0(t=0) - N^2_{0s})e^{-(g_1 + \kappa_1)t} \]  

and

\[ N^2_-(t) - N^2_{-s} = (N^2_-(t=0) - N^2_{-s})e^{-(g_2 + \kappa_2)t} \]  

At this point, it becomes convenient to define two new density terms, each representing the deviation of a donor density from its steady-state level,

\[ \Delta N_0(x,t) = N^2_0(x,t) - N^2_{0s} \]  

for the Fe\(^2\)\(_0\) sites, and
\[ \Delta N_-(x,t) = N^2_-(x,t) - N^2_- \]  
for the Fe\(^{2+}\) sites.

The spatial dependences of these solutions may be found from inspection. The excitation pattern of the Fe\(^{2+}\) sites is exactly identical with the illumination pattern. The illumination pattern is assumed, for the sake of simplicity, to be

\[ I(x) = \frac{1}{2} I_0 (1 + \cos Kx) \]  

In order to determine the donor densities resulting from this illumination pattern, the following two assumptions are made:

1) electrons are uniformly shifted some distance along the c-axis upon excitation, and do not recombine, with any significant probability, at the same sites that they were excited from, and

2) the density of Fe\(^{3+}_0\) sites is much greater than the Fe\(^{3+}_+\) sites, so that most retrapping occurs at Fe\(^{3+}_0\) sites. This assumption is the same as assuming an infinite supply of Fe\(^{3+}_0\) sites.

The resulting Fe\(^{2+}_0\) density is simply equal to the initial uniform distribution minus those sites that have been excited. Therefore, the Fe\(^{2+}_0\) distribution has exactly the same spatial variation as the excitation pattern. One may write

\[ \Delta N_0(x,t) = \Delta N_0(t) \frac{1}{2} (1 + \cos Kx) \]  

The density of the Fe\(^{2+}\) sites that are created by the retrapping of
the shifted electrons is spatially periodic with the same period as the illumination, but can deviate slightly from an exact sinusoid, depending on writing intensity, exposure time, and the ratio of the transport length to the grating spacing. However, for the restricted class of holograms used in this work, the deviation from pure sinusoidal form is not too severe, and the density is approximated by a sinusoidal distribution which is phase shifted with respect to the illumination, resulting in

$$\Delta N_-(x,t) = N_{2e}^- (t) \frac{1}{2}(1 + \cos(Kx + \phi))$$

For convenience, the following quantities are defined

$$g_1 + \kappa_1 = G_1$$ and $$g_2 + \kappa_2 = G_2$$

Eqs. 4.24 through 4.31 now allow Eq. 4.13 to be written as

$$\left( \frac{\partial}{\partial t} + \frac{1}{\tau} + \mu \frac{\partial^2}{\partial x^2}(E + \tilde{E}) - D \frac{\partial^2}{\partial x^2} \right) n(x,t) =$$

$$\frac{1}{2\bar{g}_1} \Delta N_0 (1 + \cos Kx) e^{-G_1 t} + \frac{1}{2\bar{g}_2} \Delta N_-(1 + \cos(Kx + \phi)) e^{-G_2 t}$$

At this point, it is convenient to partition $$n(x,t)$$ as

$$n(x,t) = n_0(t) + n_1(t)e^{iKx} + n_2(t)e^{-iKx}$$

Also, since the writing process does not approach saturation for the range of $$\Delta n$$ of interest, the space-charge field, $$\tilde{E}$$, will be small enough to ignore, because $$\tilde{E}$$ only affects the writing process (and the decay process) when $$\Delta n$$ is approaching its saturation level. Hence,

$$\mu \frac{\partial^2}{\partial x^2}(E + \tilde{E}) n = \mu E \frac{\partial^2}{\partial x^2} n$$
since \( \frac{\partial E}{\partial x} = 0 \). Eq. 4.32 is now

\[
\left( \frac{3}{3t} + \frac{1}{\tau} + \mu E \frac{\partial^2}{\partial x^2} - \frac{D \partial^2}{\partial x^2} \right) n(x,t) = \frac{1}{2} \left\{ g_1 \Delta N_0 e^{-G_1 t} + g_2 \Delta N_0 e^{-G_2 t} \right\} \\
+ \frac{1}{4} \left\{ g_1 \Delta N_0 e^{-G_1 t} \left[ e^{iKx} + e^{-iKx} \right] \\
+ g_2 \Delta N_0 e^{-G_2 t} \left[ e^{i(Kx + \phi)} + e^{-i(Kx + \phi)} \right] \right\} 4.35
\]

By grouping together the terms with identical \( x \)-dependences, the solutions for \( n_0(t) \), \( n_1(t) \), and \( n_2(t) \) can be found. The equation for \( n_0(t) \)
is

\[
\left( \frac{3}{3t} + \frac{1}{\tau} \right) n_0(t) = g_1 \Delta N_0 e^{-G_1 t} + g_2 \Delta N_0 e^{-G_2 t} 4.36
\]

Multiplying both sides of this equation by \( e^{t/\tau} \) and integrating over time from 0 to \( t \) yields

\[
n_0(t) = \frac{g_1 \Delta N_0}{1/\tau - G_1} \left[ e^{-G_1 t} - e^{-t/\tau} \right] + \frac{g_2 \Delta N_0}{1/\tau - G_2} \left[ e^{-G_2 t} - e^{-t/\tau} \right] + C e^{-t/\tau} 4.37
\]

which is the time solution for \( n_0(t) \). The constant of integration, \( C \), may be determined by considering the boundary conditions occurring at the cessation of illumination. The boundary conditions are completely specified by three quantities: the conduction electron density, \( n(x,t) \), and the two donor densities, \( \Delta N_0(x,t) \) and \( \Delta N_-(x,t) \). During the writing period, there exists a dynamic balance between the excitation processes and the relaxation processes. The primary excitation process during the writing period is the illumination. At a time \( t=0^- \) there exists a very large number of electrons in the conduction band, which gives rise to a (relatively) large conductivity. With the cessation of illumination,
this dynamic balance is disturbed, and the conduction electron density be must drastically reduced to restore equilibrium. Hence, a large number of electrons are trapped immediately after the light is removed. Note that the physical situation requires that

\[ C > \frac{g_1 \Delta N_0}{1/\tau - G_1} + \frac{g_2 \Delta N_1}{1/\tau - G_2} \] 4.37a

The exact values of the boundary conditions at \( t=0^- \) ( \( n(x,t), \Delta N_0(x,t), \) and \( \Delta N_1(x,t) \)) can only be determined by the use of dynamic theory for the writing period (i.e., Kim, et al.)[35,36]. However, since these experiments are restricted to writing only in the quadratic region, the approximation of sinusoidal forms is quite valid. The equation for \( n_1(t) \) derived from Eq. 4.35 is

\[
\left[ \frac{3}{3t} + \frac{1}{\tau} + i\mu E K + DK^2 \right] n_1(t) = \frac{1}{2} \left[ g_1 \Delta N e^{-G_1 t} + g_2 \Delta N e^{-G_2 t} e^{i\phi} \right] 4.38
\]

Defining \( \alpha = DK^2 + \frac{1}{\tau} \) and \( \beta = \mu E K \), the equation is multiplied by \( \exp ((\alpha + i\beta)t) \) and integrated as before, giving

\[
n_1(t) = \frac{1}{2} \left[ \frac{g_1 \Delta N_0}{\alpha - G_1 + i\beta} \left[ e^{-G_1 t} + e^{-t(\alpha + i\beta)} \right] \right] 4.39
\]

as the solution for \( n_1(t) \). Following the identical procedure, the solution for \( n_2(t) \) is found to be

\[
n_2(t) = \frac{1}{2} \left[ \frac{g_2 \Delta N_1}{\alpha - G_2 - i\beta} \left[ e^{-G_2 t} + e^{-t(\alpha - i\beta)} \right] \right] 4.40
\]

The boundary conditions involved in the solutions for \( n_1(t) \) and \( n_2(t) \) are the same as for the solution to \( n_0(t) \), namely, the conduction elec-
tron density, and the donor site densities.

The recombination time, $\tau$, has been estimated to be on the order of $10^{-11}$ seconds, so that it is assumed that the terms including $e^{-at}$ go to zero for all times of physical interest ($t \gg \tau$). Eqs. 4.39 and 4.40 can now be written

$$n_1(t) = \frac{1}{2} \left(\frac{g_1 N_0 e^{-G_1 t}}{\alpha - G_1 + i\beta} + \frac{g_2 N_0 e^{i\phi} e^{-G_2 t}}{\alpha - G_2 + i\beta}\right)$$

and

$$n_2(t) = \frac{1}{2} \left(\frac{g_1 N_0 e^{-G_1 t}}{\alpha - G_1 - i\beta} + \frac{g_2 N_0 e^{i\phi} e^{-G_2 t}}{\alpha - G_2 - i\beta}\right)$$

The conduction electron density, $n(t)$, is now seen to consist of a spatially constant part ($n_0(t)$), and a modulated part consisting of $n_1(t) e^{iKx} + n_2(t) e^{-iKx}$. The spatially modulated electron density can be compactly expressed as

$$n_1(t)e^{iKx} + n_2(t)e^{-iKx} =$$

$$\frac{g_1 N_0 e^{-G_1 t}}{(\alpha - G_1)^2 + \beta^2 \left[(\alpha - G_1)^2 + \beta^2\right]^{1/2}} \cos \left(Kx + \tan^{-1}\frac{\beta}{\alpha - G_1}\right)$$

$$- \frac{g_2 N_0 e^{-G_2 t}}{(\alpha - G_2)^2 + \beta^2 \left[(\alpha - G_2)^2 + \beta^2\right]^{1/2}} \cos \left(Kx + \tan^{-1}\frac{\beta}{\alpha - G_2}\right)$$

The complete electron density is

$$n(x,t) = \tau(g_1 N_0^2 + g_2 N_0^2)$$

$$+ \frac{g_1 N_0 e^{-G_1 t}}{\frac{1}{\tau} - G_1} + \frac{1}{\left[(\alpha - G_1)^2 + \beta^2\right]^{1/2}} \cos \left(Kx + \tan^{-1}\frac{\beta}{\alpha - G_1}\right)$$
\[ + g_2^N n e^{-G_2 t} \left( \frac{1}{1 - G_2} + \frac{1}{[a - G_2]^2 + \beta^2]^{1/2}} \cos \left( Kx + \tan^{-1} \frac{\beta}{a - G_1} \right) \right) \]

The transient density is seen to consist of four terms, two being spatially varying and two being spatially constant. All terms exhibit an exponential decay depending on the generation/trapping coefficients \( G_1 \) and \( G_2 \). The spatially varying terms also have a dependence on the applied electric field.

Thus the conductivity

\[ \sigma(x,t) = e\mu n(x,t) \]

is seen to consist of:

1) a time-invariant, spatially-invariant "dark" conductivity,

2) two time-varying, spatially-invariant terms, and

3) two time-varying, spatially-varying terms.

III. ASYMETRIC POTENTIAL WELL

The bulk photovoltaic effect (BPVE), as theorized by von der Linde and Glass [26] is based on the idea that the crystal field potential surrounding an iron ion is assymetric with respect to the \( \pm c \)-axis directions. They illustrated their theory by way of a one-dimensional model, as shown in Fig. 4-1. Although this model only contains the axial dimension, it may be more accurate than would initially be expected, because of the findings of Kurz, et.al. [37], i.e., that the electric field gradient at the iron site is parallel to the \( c \)-axis. This implies that there is a general cylindrical symmetry to the field surrounding the ion, which allows the one-dimensional model to be used
with more certainty.

As can be seen from Fig. 4-1, when an electron is excited in this potential to an energy $E_1 < E < E_2$, it then may propagate freely in the $+c$ direction, but sees a potential barrier in the $-c$ direction. This fact can be expressed by assigning different probabilities, $p_\pm$, for travel of the excited electron in the $+$ and $-$ directions. If the mean free paths for the $+$ and $-$ directions are specified as $l_\pm$, then after a large number of excitations we can write the average electron movement as

$$p_+ l_+ - p_- l_-$$

This is an important quantity and will be referred to as the electron excitation length. This is theorized to be the primary transport process during hologram formation.

There is another component of the photocurrent which is due to relaxation of the charged ion after photoexcitation. This process is best illustrated by means of a configuration coordinate diagram, in which the energy of an ion is plotted versus some interatomic, or intramolecular distance. The c.c. diagram for the $\text{Fe}^{2+}_0$ site is shown in Fig. 4-2.

1. The $\text{Fe}^{2+}_0$ Site

The site is initially in the relaxed ground state at point A on the diagram. Absorption of a pump photon ($h\nu_p$) excites an electron and leaves the site in a charged $3^+$ state with high vibrational energy.
**Fig. 4-1** Basic asymmetric potential well model surrounding the iron donor ion, as proposed by Glass and von der Linde (Ref. 13).
Fig. 4-2 Configuration coordinate diagram for the Fe\(^{2+}\) donor site, showing the transitions to and from the Fe\(^{3+}\) acceptor site.
(point B). Nonradiative relaxation brings the site to its new equilibrium point at C. When an electron is trapped, a luminescence photon ($h\nu_2$) is emitted and the site is at point D, where non-radiative relaxation returns the site to point A again.

Note that the ion travels from B to C in a +3 charge state, and from D to A in a +2 state. Thus it transports one electron worth of charge a distance of $d_1$. Analogously to Eq. 4.39, this distance $d_1$ is termed the ionic excitation distance.

1. The Fe$^{3+}$ Site

The configuration coordinate diagram for the Fe$^{3+}_0$ sites (Fig. 4-3) differs from the Fe$^{2+}_0$ diagram in two aspects. The first is that, on the Fe$^{3+}_0$ diagram, the 3+ curve has a lower minimum than the 2+. The second difference is that the minimum of the upper curve (the $2^+_c$ in this case) is outside of the lower curve.

The trapping of an electron from the conduction band causes the state of the site to change from point A to point B on the diagram. Nonradiative transitions then relax the site to point C. Note that at this time in the process, a quasi-stable donor, Fe$^{2+}_-$, has been created. As with the Fe$^{2+}_0$ site, the absorption of a pump photon at $h\nu_p2$ can excite an electron, and the site is excited to point D, where non-radiative relaxation returns the site to the starting point A. The photon frequency $\nu_p2$ is not constrained to be the same as $\nu_p1$. 
Fig. 4-3 Configuration coordinate diagram for the $\text{Fe}^{3+}$ acceptor site, showing transitions to and from the $\text{Fe}^{2+}$ donor site.
It is interesting to note the effects of high temperature as predicted by these diagrams. If either excited site receives sufficient thermal energy to raise its energy from point C to point E, then a transition can occur without emission or absorption of a photon. In the case of the Fe$_{2}^{2+}$ site (Fig. 4-2), this thermally activated transition will prevent the luminescence transition. This is called "thermal quenching" of luminescence. In the case of the Fe$_{0}^{3+}$ site (Fig. 4-3), sufficient thermal excitation will, in essence, prevent the formation of any Fe$_{2}^{2+}$ donors.

IV. Limitations of the Model

There are several key assumptions made in this theoretical derivation that must be mentioned. The first concerns the rate equations for the donor densities. The trapping terms in both equations do not show an explicit dependence on the free electron density, which would be required for a completely rigorous treatment. Instead, it is assumed the trapping terms show only a dependence on the trap density.

Second, $\tau$, the electron recombination time, has been assumed to be essentially time- and space-invariant. In the long-time limit of Eq. 4.44, this approximation will make little difference, and variations in $\tau$ from a constant value will be visible only in the period immediately after the cessation of illumination. It is also assumed that the recombination time will not be drastically affected by variations in the trap density.

A third assumption used in this derivation is that the shape of the refractive index grating remains essentially sinusoidal throughout the
hologram decay process. If, however, the decay is such that higher frequency components of the grating appear during decay, then the final answer arrived at will not give the proper description for the observed diffraction efficiency. However, there is some experimental evidence supporting the conclusion that there is no development of higher orders of the grating. Gratings of higher order will produce higher orders of diffraction of the incident light. During the experiments of this work, these higher orders were not seen, thus showing that these components were much smaller than the fundamental, if they existed at all. If, however, the gratings do not maintain a sinusoidal form throughout the decay process, then the magnitude of the first order diffracted spot will be decreased. These higher frequency gratings would decay faster than the fundamental, since the grating spacing would be much smaller. So it can be seen that although the development of such higher order components would lessen the observed diffraction efficiency, these components would decay faster, restoring the energy to the first order diffracted beam.

This model of field affected sites is also dependent upon the accuracy of the assumption that the space-charge field does not affect the decay process. This assumption is valid if the space-charge field is less than the applied field used in this work. This assumption is valid, and a short proof appears in Appendix B.
Chapter 5
Discussion

I. Data

The experimental data of Ch. 3 is well described by the theory developed in Ch. 4. The basic form of the data, as described by Eq. 3.1, matches the form of the theoretical solution as given by Eq. 4.42. The theoretical form also shows a dependence on the doping density which appears in the data. However, the electric field dependence shown in the data is slightly more complicated than what the theory would indicate.

Equating Eq. 3.1 with Eq. 4.42 leads to the following correlations between the physical quantities of Ch. 4 and the data parameters of Ch. 3:

\[ a_1 = g_1 \Delta N_0 \left( (\alpha - G_1)^2 + \beta^2 \right)^{-1/2} \]

\[ a_2 = g_2 \Delta N_0 \left( (\alpha - G_2)^2 + \beta^2 \right)^{-1/2} \]

\[ b_1 = G_1 \]

\[ b_2 = G_2 \]

\[ a_3 = \tau (g_1 N_{0s}^2 + g_2 N_{s-}^2) \]

Noting that

\[ \alpha = \frac{1}{\tau} + DK^2 \]

\[ \beta = \mu E K \]

\[ G_1 = g_1 + \kappa_1 \]

\[ G_2 = g_2 + \kappa_2 \]
the data parameters can be written as

\[
a_1 = g_1 \Delta N_0 \left( \frac{1}{\tau} + DK^2 - g_1 - \kappa_1 \right)^2 + \left( \mu EK \right)^2 \right)^{-1/2}
\]

\[
a_2 = g_2 \Delta N_0 \left( \frac{1}{\tau} + DK^2 - g_2 - \kappa_2 \right)^2 + \left( \mu EK \right)^2 \right)^{-1/2}
\]

\[
b_1 = g_1 + \kappa_1^m
\]

\[
b_2 = g_2 + \kappa_2^m
\]

where \( m \) represents a dimensionless variable which is proportional to the doping concentration (\( 0 < m < 1 \)). The doping dependence predicted by the theory is linear in \( Fe^{2+} \) density, while the data show generally increasing "a" parameters for increasing \( Fe^{2+} \) density. This deviation from linear behaviour could be due to some sort of saturation effects. One possibility is that the crystal may be able to contain only a limited amount of iron without suffering alterations in the stoichiometry. Another possibility is that the conductivity increases rapidly at the point when increased doping allows the iron orbitals to begin to overlap.

The strong variations visible in the "b" parameters due to doping density (Fig. 3-17) lead to the conclusion that the trapping rates \( \kappa_1 \) and \( \kappa_2 \) are both larger than their respective generation rates, \( g_1 \) and \( g_2 \). Fig. 3-17 also shows that the values for \( b_2 \) cover a wider range than the values for \( b_1 \). Since the doping dependence of \( b_1 \) and \( b_2 \) is due to the \( \kappa^m \) terms (the generation rates are constant), the data imply that

\[
\kappa_2 > \kappa_1
\]
which is consistent with the expectations of Ch. 4 (Eq. 4.15).

It is difficult to draw conclusions with any certainty about the generation rates since they are so much smaller than their trapping rates.

The estimated magnitudes of the various parameters are

\[ \kappa_1 = 7 \times 10^{-2} \text{sec}^{-1} \]
\[ \kappa_2 = 5 \times 10^{-2} \text{sec}^{-1} \]
\[ g_1 = 10^{-6} \text{sec}^{-1} \]
\[ g_2 = 3 \times 10^{-6} \text{sec}^{-1} \]

The values for the generation rates are very approximate.

1. Field Dependence

The field dependence predicted by the theory for the "a" parameters is of the general form

\[ a_1 = C \frac{1}{(d + cE^2)^{1/2}} \quad 5.5 \]

where \( C, c \) and \( d \) are constants and \( E \) is the field. This function predicts that the parameters would monotonically decrease in magnitude with increasing field magnitude. However, this has not been observed. Therefore, one is forced to conclude that the effects of the electric field are not limited to merely affecting the electron continuity equation (Eq. 4.11).

It should be noted that the parameters undergo greater variations due to the field when the doping level is high, in other words, a high
doping concentration produces a stronger field dependence.

One possible hypothesis to account for this involves a shift of the donor and acceptor energy levels by the applied field. Assuming that the donors and acceptors are in essentially parabolic energy wells (the standard harmonic oscillator model), then an applied field would affect both the thermal generation and the trapping rates. These effects are most easily visualized by use of the configuration coordinate (c.c.) diagrams introduced in Fig. 4-2 and 4-3. In these diagrams, energy of an ion is plotted versus some spatial coordinate (usually not rigorously defined). The spatial coordinate used in these diagrams is the c-axis.

Fig. 5-1 illustrates the effect of an applied field on an Fe\textsuperscript{2+}\textsubscript{0} site. The field actually affects both states (Fe\textsuperscript{2+}\textsubscript{0} and Fe\textsuperscript{3+}\textsubscript{+}), but the 3+ state has the greater charge, and so is affected to a greater extent. For simplicity then, this is illustrated by having only the 3+ curve shifted by the field. As can be seen, the energy minimum of the 3+ curve shifts both in energy, by an amount ΔE, and in position, by an amount Δx. Also, the energy gap determining the optical absorption (pt.A to pt.B) is slightly altered. This produces four results. First, the optical absorption peak (as induced by the iron doping) and the thermal generation rate should both change slightly. Second, the wavelength of the observed luminescence should change, since, with no applied field

\[ hν_{LUM} = \text{Energy (pt.C)} - \text{Energy (pt.E)} \]

where, with the field

\[ hν_{LUM} = \text{Energy (pt.C')} - \text{Energy (pt.E')} \]

5.6

5.7
Fig. 5-1 Configuration coordinate diagram for the Fe\textsubscript{2+} donor site showing the energy shifts of the Fe\textsubscript{3+} acceptor site due to an applied electric field.
Third, the "thermal quenching" of luminescence would be affected. This phenomenon is caused by thermal excitation of the sites at pt. C up to pt. D, where vibration relaxation returns the ion to the ground state without emitting a photon. This phenomenon is dependent upon the energy separation between the C and D points. A change in this energy difference would be seen as a change in the temperature at which thermal quenching occurred. Fourth, the ionic component of the current will be changed. However, it is difficult to separate the individual current components, so that only the first three effects would be easily observable.

The applied field would also affect the Fe$_{0}^{3+}$ sites, and this is illustrated in Fig. 5-2.

II. Comparison of the Results to the Works of Other Researchers

The results presented here agree quite well with the results of many other researchers when the two donor-two acceptor version of the BPVE is adopted.

1. Phonon Generation

The optical absorption processes described by the c.c. diagrams of Ch. 4 imply that the process would generate a large amount of phonons, since the iron ions move back and forth with each event. This process has been described in more detail by Chanussot and Glass [38]. It has been calculated, from a completely different approach, that 77% of the optical transition energy goes into the excitation of phonons [5]. The large number of phonons may be involved in the scattering of large
Fig. 5-2 Configuration coordinate diagram for the Fe$_3^+$ acceptor site showing the energy shifts of the Fe$_2^+$ donor site due to an applied electric field.
phenomenon of thermally activated trapping. The high amounts of thermal energy cause \( \text{Fe}^{2+} \) sites to decay as soon as they are formed, and also provides the energy needed to raise the \( \text{Fe}^{3+} \) sites from pt.C to pt.D (see Fig. 4-3). By raising the level of the \( \text{Fe}^{3+} \) above its minima, necessary thermal energy required to reach pt.D is lessened, so that the capture of an electron is more probable. This can be seen more clearly when one thinks of the sites as being thermally-activated trapping sites [18]. Hordvik, et al [42], have found that 175°C is sufficient to almost completely quench all luminescence, and this result is also in agreement with the present theory, thus giving confirmation that at these temperatures, thermal excitation of all sites is nearly complete. The \( \text{Fe}^{3+} \) sites are prevented from trapping any electrons for any length of time, so that no space charge field can form.

4. "Thermal Fixing"

The phenomenon of thermal fixing has been established for quite a while\[14\] and is seen when recording a hologram in a crystal while the crystal is at approximately 100°C. Note that this temperature is below that required for thermal quenching of luminescence. After cooling the crystal to room temperature, the diffraction efficiency goes to 0, i.e., the heat treatment seems to erase the hologram. But when re-illuminated with a readout beam, \( n \) was found to increase from 0 to a fairly high level as illumination continued. In contrast, holograms written at room temperature, and then illuminated with the same readout beam were optically erased, as would be expected. The hologram written at 100°C was not optically erased by the readout beam, even after long exposures.
This process has been interpreted by Staebler and Amodei[14] as being indicative of ionic movement at 100°C. The ionic movement occurs so as to cancel out the non-uniform charge distribution produced by the initial interference pattern. Cooling the crystal to room temperature then freezes the ions into their new positions, so that the space charge fields are zero (or nearly so). Illumination with a readout beam initially sees no diffraction grating, but then re-excites the redistributed electrons, and a new space-charge field comes into being, but this time produced by the displaced ions. As such, it shows very little decay at temperatures below 100°C.

This theory is plausible, except for the fact that ionic movement in LiNbO₃ at only 100°C seems difficult. No diffusion effects, such as out-diffusion of Li, or in-diffusion of dopants (Ti, V, or Ni) takes place at any temperature much below 600°C[43] so it is hard to envision ionic movement occurring at such a low temperature. However, the two donor–two acceptor model can readily provide a model for this effect.

A basic characteristic of this model is that the Fe²⁺ sites are distinguished from the Fe³⁺ sites because the Fe²⁺ sites have an extra electrons worth of charge compensation in their local vacinity. Although this extra charge is constrained to be somewhere along the c-axis from the Fe site, no other assumptions can be made about it. If this compensating charge were to leave the Fe²⁺ site, the site would be transformed into a Fe³⁺ site. It seems much more plausible that such a compensating charge would become mobile at 100°C than that an ion would become mobile at the same temperature. Thus it is proposed that thermal
fixing is due to the movement of the charge compensation charges, and not of ions.

5. Electric Field Effects

The only previous work done on the effects of electric fields was done for fields applied during the writing period, and examining the effects produced on the writing process itself. Since illumination causes a photo-current of relatively large magnitude [3, 44], this current completely masks any non-linear effects. The main result reported so far is that the value of electric field which is required to reduce the photo-current to 0 varies with the intensity of the illumination [26, 34, 45]. The exact form of this dependence has been debated [36, 46], and it was unclear from the data whether the field required has an intensity dependence or not. If, however, the field affects the generation and trapping rates, as concluded here, then the results of those experiments must be completely reconsidered.

III. Conclusions

The decay of holograms stored in iron-doped LiNbO₃ has been found to deviate from the exponential decay expected of a non-uniform charge distribution in dielectrics. By adopting the bulk photovoltaic effect theory of Glass and von der Linde, and introducing modifications due to the divalent iron ions in the crystal, a form for the decay of the diffraction efficiency has been developed which matches the experimental data quite well. This theoretical form predicts the time dependence of the results found, and shows general agreement with the results found for the dependence on the doping density. However, disagreement was
found in the predicted and measured effects of an electric field applied during decay. This result is interpreted as implying that the field can affect the basic properties of the iron sites, which function as the donors and acceptors.

The general results found here are in agreement with many other works, and no direct disagreements between these findings and the results of other researchers have been found.

In summary, a modification to the bulk photovoltaic effect theory of Glass and von der Linde has been found to be required for the theory to account for effects seen during the decay of holograms. This modification has the added benefit of allowing this one theory to account for several phenomenon, such as thermal fixing, and luminescence, which were not previously accounted for by the theory.
APPENDIX A
MATERIAL PROPERTIES OF LITHIUM NIOBATE

Formula Weight: 147.83
Mass Density [49]: 4.64 gm/cm³ (at 25°C)
Melting Point [9]: 1260 ± 10°C
Curie Temperature [49]: 1210 ± 10°C
Crystal Structure [9]: Rhombohedral, C₃ᵥ (R3c) (at 25°C)
Lattice Constants [9]:
  \[ a_H = 5.14829 \pm 2 \text{ Å} \]
  \[ c_H = 13.861 \pm 4 \text{ Å} \]

  corresponding to
  \[ a_R = 5.4944 \text{ Å} \]
  \[ \alpha = 55.867° \]

Atomic Arrangement [48,9-12]: Constant up to 1200°C

(i) Oxygen atoms arranged on planar sheets forming a network of
distorted octahedra (based on hexagonal close-packing of oxygens)

(ii) The centers of the octahedra are occupied, in moving along
the +c axis, by Nb, vacancy, Li, Nb, vacancy, Li, etc.

(iii) Interatomic distances:

<table>
<thead>
<tr>
<th></th>
<th>Formal Charge</th>
<th>Ionic Radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nb - O</td>
<td>+5</td>
<td>0.69 Å</td>
</tr>
<tr>
<td>Li - O</td>
<td>+1</td>
<td>0.68 Å</td>
</tr>
</tbody>
</table>

(iv) Isotropic thermal vibrations for all atoms; Debye temperature
of 503°C.

Defects
(i) [50]: Assuming substitutional impurities:
+1, +2 ions enter lithium sites
+4, +5, +6 ions enter niobium sites
+3 ions enter either site
(ii) [25]: Evidence exists for an oxygen vacancy - metal dopant defect center with chemochromic properties.

Band Structure [52]: Oxygen 2p\nu valence band and a niobium 3d\nu conduction band. Band gap of 3.72 eV (~3500 Å). Energy difference between band centers is ~5 eV.

Refractive Index [53]: For pure LiNbO\textsubscript{3}, the refractive indices are temperature and wavelength dependent:
(for T < 600°C)

\[
(n_o)^2 = 4.913 + \frac{1.173 \times 10^5 + 1.65 \times 10^{-2}T^2}{\lambda^2 - (2.12 \times 10^2 + 2.7 \times 10^{-5}T^2)^2}
- 2.78 \times 10^{-8} \lambda^2
\]

\[
(n_e)^2 = 4.5567 + 2.605 \times 10^{-7}T^2 + \frac{0.97 \times 10^5 + 2.7 \times 10^{-2}T^2}{\lambda^2 - (2.01 \times 10^2 + 5.4 \times 10^{-5}T^2)^2}
- 2.24 \times 10^{-8} \lambda^2
\]

T = temperature in K \quad \lambda = \text{wavelength in nm}
at $T = 25^\circ$C $\lambda = 6328$ Å $4880$ Å

$n_0$ 
2.2884 
2.3504

$n_e$ 
2.2019 
2.2544

birefringence 
-0.0865 
-0.0960

Dielectric Constant [49]: ($T = 25^\circ$C)

(i) $78 \varepsilon_0 = 6.91 \times 10^{-12}$ F/cm perpendicular to the optic axis

(ii) $32 \varepsilon_0 = 2.83 \times 10^{-12}$ F/cm parallel to the optic axis

Electro-Optic Tensor [19,54]:

$$
\begin{bmatrix}
0 & -r_{22} & r_{13} \\
0 & r_{22} & r_{13} \\
0 & 0 & r_{33} \\
0 & r_{42} & 0 \\
r_{42} & 0 & 0 \\
-r_{22} & 0 & 0
\end{bmatrix}
$$
in units of $10^{-10}$ cm/V: $r_{13} = 8.6$, $r_{22} = 3.4$,

$r_{33} = 30.8$, $r_{42} = 2.8$

Piezo-Electric Tensor [55]:

$$
\begin{bmatrix}
0 & -e_{22} & e_{31} \\
0 & e_{22} & e_{31} \\
0 & 0 & e_{33} \\
0 & e_{15} & 0 \\
e_{15} & 0 & 0 \\
-e_{22} & 0 & 0
\end{bmatrix}
$$
in units of coul/m²: \( e_{15} = 3.7, e_{22} = 2.5 \)

\[ e_{31} = 0.2, e_{33} = 1.3 \]

Elastic Tensor [55]:

\[
\begin{bmatrix}
C_{11} & C_{12} & C_{13} & C_{14} & 0 & 0 \\
C_{12} & C_{11} & C_{13} & -C_{14} & 0 & 0 \\
C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\
C_{14} & -C_{14} & 0 & C_{44} & 0 & 0 \\
0 & 0 & 0 & 0 & C_{44} & C_{14} \\
0 & 0 & 0 & 90 & C_{14} & C_{66}
\end{bmatrix}
\]

in units of \(10^{11}\) N/m²: \( C_{11} = 2.03, C_{12} = 0.53 \)

\( C_{13} = 0.75, C_{14} = 0.09 \)

\( C_{33} = 2.45, C_{44} = 0.60 \)

\( C_{66} = \frac{1}{2}(C_{11} - C_{12}) = 0.75 \)

For the equations concerning the above tensors, see Ch. 7, Ref. 55.

Pyro-Electric Coefficient [4]:

\[ \approx 10^{-8} \text{C/(m}^2\text{-deg)} \] (at \( T = 100^\circ\text{C} \))

Mobility [51]:

15. \( \text{cm}^2/\text{V-sec} \) (extrapolated from \( 1000^\circ\text{C} \))

Trapping Time [47]:

\( 4 \times 10^{-11} \text{sec} \) (using the above mobility)
Inverse Susceptibility [49]:

\(\frac{1}{\chi}\) linear from 300 to 950\(^\circ\) C

Curie-Weiss intercept \(\approx 1080\(^\circ\) C\)
APPENDIX B

The maximum magnitude of the space-charge field achieved in the holograms used in this work may be easily calculated. It is found to be on the order of 400 V/cm, so that its affect on the hologram decay process will be much less than the applied fields.

Beginning with the formula relating $\eta$ to $\Delta n$,

$$\eta = \sin^2 \left( \frac{\pi d \Delta n}{\lambda \cos \theta} \right)$$

substitute the values appropriate to these experiments, i.e.,

$$\eta = 0.1 \quad d = 1 \text{ mm} \quad \lambda = 4880 \text{ Å} \quad \theta = 6^\circ$$

gives a value for $\Delta n$ of

$$\Delta n = 4.943 \times 10^{-5}$$

Then, with the equation relating $\Delta n$ to the space-charge field $E$, (Eq. 1.11),

$$\Delta n = \frac{1}{2} n_0^3 r_{33} E_3$$

with the following values,

$$n_0 = 2.3 \quad r_{33} = 30.8 \times 10^{-10} \text{ cm/V}$$

the value of $E$ is found to be

$$E = 263 \text{ V/cm}$$

This value is much less than the fields applied to the crystal during the decay experiments, therefore the assumption that the space-charge field affects the decay process less than the applied field is a valid one.
REFERENCES


