DIFFRACTION REALIZATION OF AN OPTICAL EXPANDER

Richard Barakat
Aiken Computation Laboratory
Division of Applied Sciences
Harvard University
Cambridge, Massachusetts 02138

and

Electro-Optics Research Center
Tufts University
Medford, Massachusetts 02155

John Reif
Department of Computer Science
Duke University
Durham, North Carolina 27706

ABSTRACT

An optical expander (i.e., an electro-optical device which takes d inputs and transmits c outputs (where $c \gg d$) is described. The one-dimensional version is worked out in detail using optical diffraction theory, thresholding (clipping), and singular value decomposition to generate the c linearly independent beams.

1. INTRODUCTION

We define an optical expander to be an electro-optical system that takes as input a series of spatial patterns (in our case, diffraction patterns of coherent sources) and transforms them into a number of linearly independent outputs capable of transmitting information with a minimum of "crosstalk". The actual number of coherent sources (call this number d) producing the diffractive images (call this number c) can be such that $c \gg d$, thereby justifying the word "expander". The present analysis is carried out in the context of a one-dimensional scenario as we are only interested in the basics of the concept at this point.

One of the major problems facing computer architecture today is that of interconnects; interconnection problems now dominate designs. Although many machines have successful interconnect schemes (eg., Connection Machine, BBN butterfly, etc.) all are limited in a variety of ways such as reconfiguration time, data word size, etc. Perhaps the most suitable interconnects are the $N \times N$ crossbars [1]. In their very important paper Goodman, Leonberger, Kung, and Athale [2] show that for a VLSI chip of area A the average maximum interconnect length $L_{\max}$ is given by

$$L_{\max} = \frac{1}{2} A^{1/2} \quad . \tag{1}$$

This equation implies that as VLSI geometries become smaller and chip sizes become larger, the interconnect distances and, consequently the maximum signal propagation delays will eventually become larger than the individual component switching times. Further Goodman, *et al.* call attention to Rent's rule which relates the number of logic devices on the VLSI chip to the number of interconnects. Leaving aside the actual details, Rent's rule has the consequence that interconnections eventually limit the useful growth of silicon-based

VLSI. There are a variety of motivations for employing optical interconnects; for details and references see the reviews of Goodman and associates [2-4] as well as Feitelson's book [5].

It can be shown [6] that based on the purely computational power of an IC a VLSI chip of area A and time T requires at least $AT^2 = \Omega(n)$ to solve problems such as FFT, convolution, and $m \times m$ matrix multiplication where $n^2 = m$. Here $f(n) = \Omega(g(n))$ means that there is a constant $c > 0$ such that for an infinite number of values of $n > 0$, we have $f(n) \geq cg(n)$; see Ullman [6]. This implies that as computational problems get larger, their solution in constant time will require larger ICs and some way of overcoming the interconnect problem. To overcome this restriction Barakat and Reif [7] show that (from a computational standpoint) by using optical technology to form interconnects across 3-space and that in such a system the lower bounds on computations becomes, $VT^{3/2} = \Omega(n^{3/2})$.

Compared to electrical interconnect optical interconnects provide greater electrical noise immunity, signal delay nearly independent of path length, and a higher bandwidth-space product. In addition, optical technology is compatible with GaAs technology and offers a variety of means of implementation (fiber, holographic, integrated optic, etc.). For a more detailed discussion of the engineering tradeoffs and a comparison between electrical and general optical interconnect technologies, see Feldman, *et al.* [8], Bergman, *et al.* [10], Anderson, *et al.* [11], as well as the Goodman references. The journals *Optical Engineering* and *Optics Communications* now contain many papers on optical interconnects.

We point out that holographic interconnect systems offer several advantages compared to fiber optic, integrated optic, or other forms of optical interconnects. HI can connect any location on an IC to any other location via free space whereas other approaches require a

physical connection (optical fiber or waveguide track) generally restricted to the edge of the IC. Holographic interconnects also offer the potential for an extremely large number of interconnects in a given volume.

Finally let us consider the message routing task: we assume there are N processors, where each processor has a distinct message with a given distinct message destination address. Simultaneously, each message is to be routed from its originating address to its destination address. The message routing problem is crucial to parallel machines, since this is how information is moved between distant processors. Furthermore, message routing allows a parallel machine with fixed connections to simulate arbitrary interconnects. In the following, we are concerned with the case when the length of each message is of some fixed constant length. This task can be accomplished (at least theoretically) by using holographic interconnection networks which require an optical expander. The purpose of the present paper is to discuss the "construction" of an optical expander; a second paper will discuss a class of highly efficient holographic message routing systems which require optical expanders.

2. THEORY OF THE EXPANDER

The basic idea of the optical expander under study in this communication is to propagate various configurations of 2 coherent sources through an optical system to form intensity configurations of these point sources on the receiving plane. Each configuration is then subject to a thresholding (clipping) operation in the manner to be described below. The effect of the thresholding is to produce (under certain conditions) linearly independent "beams" which can act as message carriers.

The geometry of the present optical expander is shown in Fig. 1. Consider a series of s ($|equiv$ even integer) coherent sources separated by an equal distance δ which form a diffraction image in the receiving plane with x denoting position. Each source is modulated by an on/off switch p_t such that

$$p_\ell = \begin{array}{l} 1 \quad \text{for } (d/2) \text{ of the } p_\ell \\ \\ 0 \quad \text{for } (d/2) \text{ of the } p_\ell \end{array} \quad . \tag{2}$$

The number of such source configurations is

$$\binom{d}{\frac{d}{2}} \quad . \tag{3}$$

However, only half these configurations are useful because each has a mirror configuration. Thus

$$c = \frac{1}{2}\binom{d}{\frac{d}{2}} = \frac{1}{2}\,\frac{d!}{(\frac{d}{2})!(\frac{d}{2})!} \tag{4}$$

is the number (always even since d is even) of allowable source configurations. We can employ Stirling's approximation of the factorial when d is large (say $d \geq 8$) to prove

$$c \sim (2\pi d)^{1/2} 2^d \quad . \tag{5}$$

Thus c is roughly proportional to $2^d!\Pi$ when the number of sources is large. Note that the number of configurations increases quite rapidly: $c_6 = 10,\ c_8 = 35,\ c_{10} = 126,\ c_{12} = 462.$

The diffracted amplitude of the coherent sources (all of equal strength) in the image receiving plane is

$$a(x) = \sum_{\ell=0}^{d-1} p_\ell a(x - \ell\delta) \quad . \tag{6}$$

where

$$a_\ell(x - \ell\delta) = \int_{-\alpha}^{d} e^{ikw(\zeta)} e^{i(k/f)(x-\ell\delta)\zeta}\, d\zeta \quad . \tag{7}$$

Here $w(\zeta)$ is the wavefront aberration function, k is the mean wavenumber of the light, f is the focal length of the system, and α is the half length of the exit pupil (taken to be a slit). In as much as we are only interested in the basic features, we specialize to an aberration-free, in-focus situation (i.e., $w(\zeta) \equiv 0$) so that

$$a_\ell(x - \ell\delta) = \frac{\sin(\frac{k}{f}(x - \ell\delta))}{(\frac{k}{f}(x - \ell\delta))} \quad . \tag{8}$$

At this point it is convenient to transform to normalized variables

$$v \equiv \left(\frac{k\alpha}{f}\right) x \quad , \quad L \equiv \left(\frac{k\alpha}{f}\right) \delta \tag{9}$$

and shift the origin of coordinates in the receiving plant to the paraxial axis of the optical system. The corresponding diffracted intensity is

$$h(v) \equiv |a(v)|^2 = \left| \sum_{\ell=0}^{d-1} p_\ell \frac{\sin\left(v + \left(\frac{d-1}{2}\right) L - L\ell\right)}{\left(v + \left(\frac{d-1}{2}\right) L - L\ell\right)} \right|^2 \tag{10}$$

On the image receiving plane, we set out an array of contiguous intensity detectors of normalized size D; an odd number $t = (1 + 2j_{\max})$ is employed. Let us consider a particular configuration of the sources, call it configuration #1 and form the resultant

diffraction image $h(v)$. Next evaluate the t integrated intensities as formed by the intensity detector array

$$H_0 \equiv \int_{-D/2}^{D/2} H(v)\, dv$$
$$H_{+j} = \int_{(j-1/2)D}^{(j+1/2)D} h(v)\, dv \quad , \quad H_{-j} = \int_{-(j+1/2)D}^{-(j-1/2)D} h(v)\, dv \tag{11}$$

for $j = 1, 2, \ldots, j_{\max}$. Now $j_{\max}$ is chosen to be large enough (at a fixed D) to guarantee that $H_{\pm j} \approx 0$ for all source configurations. Thus D is a variable parameter.

Subject these integrated intensities of configuration #1 to the nonlinear threshold operation:

$$B_{\pm j} = 1 \quad \text{if} \quad H_{\pm j} > g$$
$$= 0 \quad \text{if} \quad H_{\pm j} < g \tag{12}$$

where g is a specified intensity level. Store these t values as a column vector

$$\mathbf{B}^{(1)} \equiv \begin{vmatrix} B_{-j_{\max}}^{(1)} \\ \vdots \\ B_0^{(1)} \\ \vdots \\ B_{j_{\max}}^{(1)} \end{vmatrix} . \tag{13}$$

Next take source configuration #2 and repeat the above procedure leading to $\mathbf{B}_1$ and continue until all c configurations have been analyzed.

Since these vectors are going to be used to carry information, all c of them must be linearly independent in order to avoid "crosstalk". To determine the actual number of linearly independent vectors, we form the matrix $\mathbf{Q}$ of all $\mathbf{B}$ vectors of the ensemble of

source configurations:

$$\mathbf{Q} \equiv \begin{vmatrix} B_{-j_{max}}^{(1)} & B_{-j_{max}}^{(2)} & \dots & B_{-j_{max}}^{(c)} \\ \vdots & \vdots & & \vdots \\ B_0^{(1)} & B_0^{(2)} & \dots & B_0^{(c)} \\ \vdots & \vdots & & \vdots \\ B_{-j_{max}}^{(1)} & B_{-j_{max}}^{(2)} & \dots & B_{-j_{max}}^{(c)} \end{vmatrix} . \tag{14}$$

Note that $\mathbf{Q}$ has c rows and t columns where $t > c$. It can be shown [12] that all $!c$ of the $\mathbf{B}$ vectors are linearly independent if and only if the row rank of $\mathbf{Q}$ is c. Since $t > c$ and the row rank of $\mathbf{Q}$ must equal the column rank of $\mathbf{Q}$ by a well-known theorem in matrix theory [12], we will simply employ the work rank to denote the column rank.

Computation of the rank of a matrix is difficult and the method of singluar value decomposition (SVD) [12], the only known reliable method of computation. $\mathbf{Q}$ can be written as

$$\mathbf{Q} = \mathbf{U}\Sigma\mathbf{V}^{+} \tag{15}$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices

$$\mathbf{U}\mathbf{U}^{+} = \mathbf{U}^{+}\mathbf{U} = I$$

$$\mathbf{V}\mathbf{V}^{+} = \mathbf{V}^{+}\mathbf{V} = I \tag{16}$$

and

$$\Sigma = \begin{vmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ 0 & & & \sigma_k \\ & & 0 & \end{vmatrix} . \tag{17}$$

The σ's are the singular values of $\mathbf{Q}$ and can be ordered; they are real and nonnegative with ordering

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_c \geq 0 \quad > \tag{18}$$

The rank of $\mathbf{Q}$ is the number of nonzero singluar values. For a fixed source/system geometry the rank of $!\mathbf{Q}$ depends on: g, threshold level; L, normalized distance between sources; D, normalized size of intensity detectors; and $t = 1 + 2j_{\max}$, number of intensity detectors.

Let us now examine the role of g in determining the rank of $\mathbf{Q}$. When $g = 0$, all c of the $\mathbf{W}$'s are composed of 1's only, there is only one linearly independent $\mathbf{W}$ and thus the rank of $\mathbf{Q} = 1$. At the other extreme of g greater than the peak intensity of any of the diffracted intensities, all the $\mathbf{W}$'s are composed of 0's only and the rank of $\mathbf{Q} = 1$. Between these limiting cases of unit rank, we would expect the rank of $\mathbf{Q}$ to increase to c and stay at c for some range of g values. The actual behavior of the rank depends on the specifics of the problem; see next section for an example.

Our next problem (given that we have a $\mathbf{Q}$ of full rank) is to construct c orthogonal vectors each of length t. In addition to being orthogonal, these vectors must be linearly independent. One approach to constructing these vectors is the Gram-Schmidt procedure; however it seems easier to employ SVD since we have already used it to determine the rank of $\mathbf{B}$. To this end, we note that the $\mathbf{U}$-vectors of the SVD of $\mathbf{B}$ are c in number and of length t. If $\mathbf{B}$ is of full rank, then these $\mathbf{U}$-vectors are linearly independent as well as orthogonal. The only disadvantage of these vectors is that they do not possess the "natural ordering" of vectors constructed via the Gram-Schmidt procedure.

These vectors must be converted over to scalar functions in order to "code" them for use as information carrying beams in the hologram. We propose the following scheme based upon the Discrete Fourier Transform (DFT). Consider one of the above vectors, say $\mathbf{U}^{(j)}$, where we now relabel the elements so that they start with zero (i.e., $U_0^{(j)}, U_1^{(j)}, \ldots U_{c-1}^{(j)}$).

It has a DFT [13]

$$\mu_\ell^{(j)} = \sum_{k=0}^{c-1} U_k^{(j)} \, e^{-i2\pi\ell k/c} \tag{19}$$

where $\ell = 0, 1, \ldots, (c-1)$; and the $\mu_\ell^{(j)}$ are the spectral components of $\mathbf{U}^{(j)}$. Details will be given in a second paper.

3. NUMERICAL EXAMPLE

Let us consider the following example to illustrate the above arguments. The number of sources is taken to be $d = 6$ so that there are

$$c = \frac{1}{2}\binom{6}{3} = 10 \tag{20}$$

source configurations; they are listed in Table 1. In Table 2, we list singular values of $\mathbf{Q}$ for the case: $L = 3$, $j_{\max} = 14$, $g = 8$. Some of the intensity patterns for various source configurations are shown in Figs. 2-4. We caution the reader that no attempt has yet been made to undertake detailed parameter studies of trade-offs; the above parameters we chosen simply to illustrate the analysis. Note that $\mathbf{Q}$ is of full rank when $D = 1.25$, yet increasing the normalized detector size to $D = 1.50$ results in a rank deficient $\mathbf{Q}_0$. Only the $D = 1.25$ is admissible. The $\mathbf{U}_\ell^{(9)}$ vectors are listed in Table 3; they are orthogonal to each other as the reader can easily check.

Acknowledgement

R. Barakat was supported by internal research funds of RGB Associates.

REFERENCES

1) A.D. McAuly, *Optical Computer Architectures* (Wiley, New York, 1991).

2) J. Goodman, E. Leonberger, S. Kung, and B. Athale, "Optical interconnects for VLSI systems", Proc. IEE, 72, 850-866 (1984).

3) J. Goodman, R. Kostuk, and B. Clymer, "Optical interconnects, an overview", in Proc. Second Intl. VLSI Multilevel Interconnection Conf., 219-226 (1985).

4) J. Goodman, "Optics as an interconnect technology", in: *Optical Processing and Computing*, H. Arcenault, T. Szoplek, and B. Macukow (eds.) (Academic, New York, 1989).

5) D. Feitelson, *Optical Computing* (MIT Press, Cambridge, 1988) Chap. 3.

6) J. Ullman, *Computational Aspects of VLSI* (Computer Science Press, Rockville, MD, 1984) Chap. 3.

7) R. Barakat and J. Reif, "Lower bounds on the computational efficiency of optical computing systems", Appl. Opt., 26, 1015-1018 (1987).

8) M. Feldman, S. Essner, C. Guest, and S. Lee, "Comparison between optical and electrical interconnects based on power and speed considerations", Appl. Opt., 27, 1742-1751 (1988).

9) L. Bergman, W. Wu, A. Johnston, R. Nixon, S. Esener, C. Guest, P. Yu, T. Drabik, M. Feldman, and S. Lee, "Holographic optical interconnect for VLSI", Opt. Eng., 25, 1109-1118 (1986).

10) W. Wu, L. Bergman, A. Johnston, S. Esener, C. Guest, P. Yu, M. Feldman, and S. Lee, "Implementation of optical interconnects for VLSI", IEEE Trans. Elec. Dev.,

ED-34, 706-713 (1987).

item11) D. Anderson and D. Lininger, "Dynamic optical interconnects: volume holograms as optical two-port operators", Appl. Opt., 26, 5031-5038 (1987).

12) R.A. Horn and C.R. Johnson, *Topcis in Matrix Analysis* (Cambridge Univ. Press, New York, 1991).

13) E. Brigham, *The Fast Fourier Transform and its Applications* (Prentice-Hall, Englewood Cliffs, NJ, 1988).

Table 1

List of the $c = 10$ source configurations for $d = 6$ sources

p_ℓ	1	2	3	4	5	6
1	1	1	1	0	0	0
2	1	1	0	1	0	0
3	1	1	0	0	1	0
4	1	1	0	0	0	1
5	1	0	1	1	0	0
6	1	0	1	0	1	0
7	1	0	1	0	0	1
8	0	1	1	1	0	0
9	0	1	1	0	1	0
10	0	1	1	0	0	1

Table 2

Singular values of **Q** for $L = 3$, $j_{max} = 14$, and $g = .8$.

	D = 1.25	D = 1.50
0	5.102	7.970
1	2.772	2.504
2	2.628	2.332
3	2.132	2.044
4	1.848	1.957
5	1.189	1.335
6	1.000	0.7114
7	0.7638	0.5547
8	0.5358	0.2712
9	0.3702	$0.1709 \cdot 10^{-6}$
10–28	0	0
rank	10	9

Table 3

Singular vectors $U_j^{(9)}$, $j = 0, 1, \ldots, 9$ corresponding to the singular values in Table 2 for $D = 1.25$.

	0	1	2	3	4
0	-0.43353	-0.13083	0.08030	-0.07729	0.13226
1	-0.26146	0.12287	0.50844	-0.41977	-0.20018
2	-0.35416	-0.20251	0.43501	0.35260	-0.21131
3	-0.25076	0.12178	0.42612	-0.06453	0.47804
4	-0.22794	0.41661	-0.14163	-0.26358	-0.41112
5	-0.21620	0.32652	-0.04185	0.51803	-0.26102
6	-0.25520	0.64082	-0.22014	0.20221	0.31871
7	-0.37279	-0.10818	-0.39491	-0.44558	-0.20934
8	-0.37079	-0.40298	-0.22707	0.32712	-0.19514
9	0.33926	-0.21683	-0.29399	-0.07117	0.49863

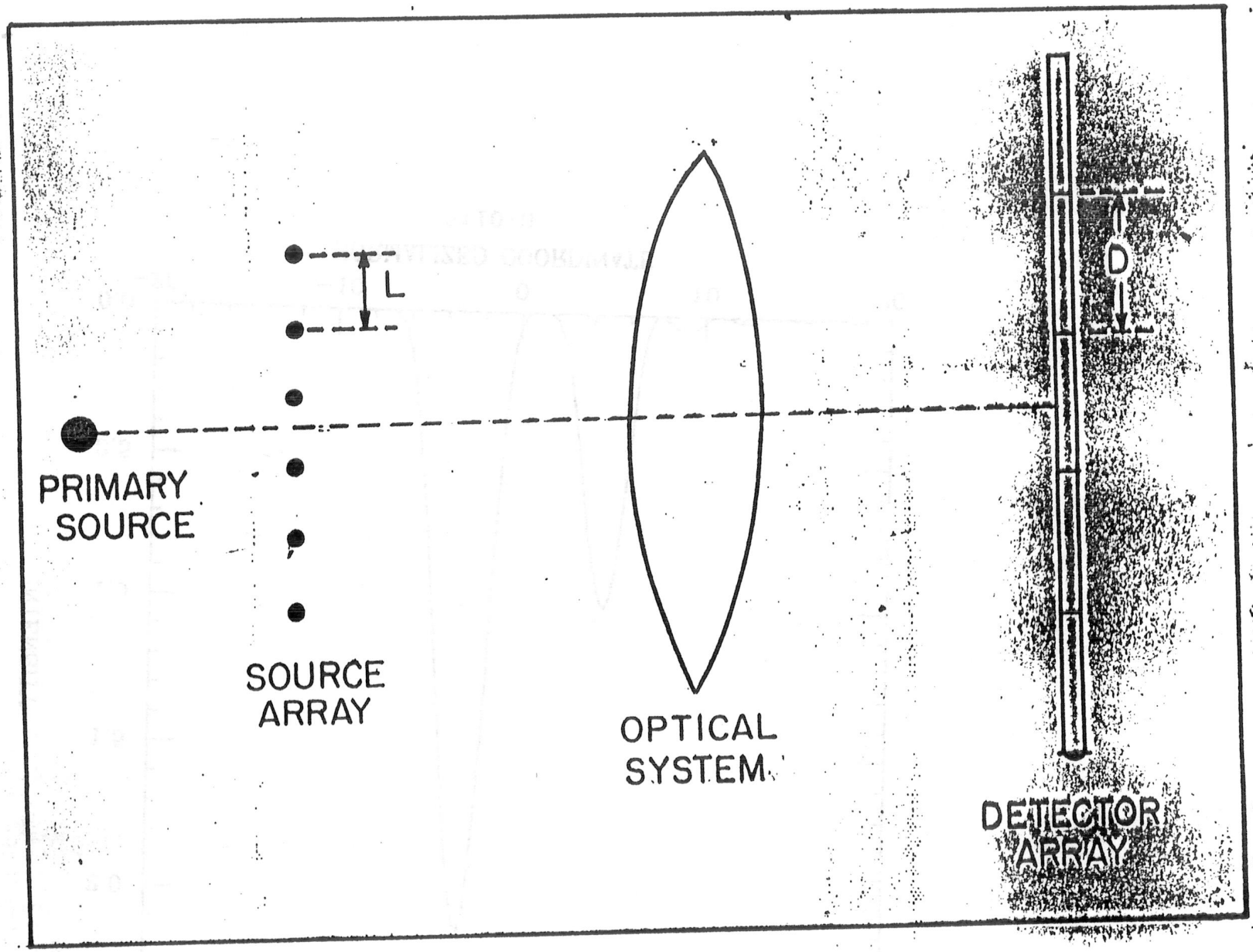
	5	6	7	8	9
0	-0.58050	0.55470	-0.30303	0.14795	0.08139
1	-0.10278	-0.55470	-0.29299	0.14795	0.08139
2	0.44858	0.27735	-0.03288	-0.21351	-0.38798
3	0.11651	0.00000	0.48920	-0.12898	0.48835
4	-0.07322	0.27735	0.51024	0.38578	-0.15644
5	-0.50546	-0.27735	0.18904	-0.37451	0.00035
6	-0.33212	0.00000	-0.45895	0.10832	0.02511
7	0.18495	0.00000	-0.09234	-0.60145	0.21570
8	0.15779	-0.27735	0.01020	0.47454	0.41994
9	-0.08598	-0.27735	-0.25804	0.04483	-0.58724

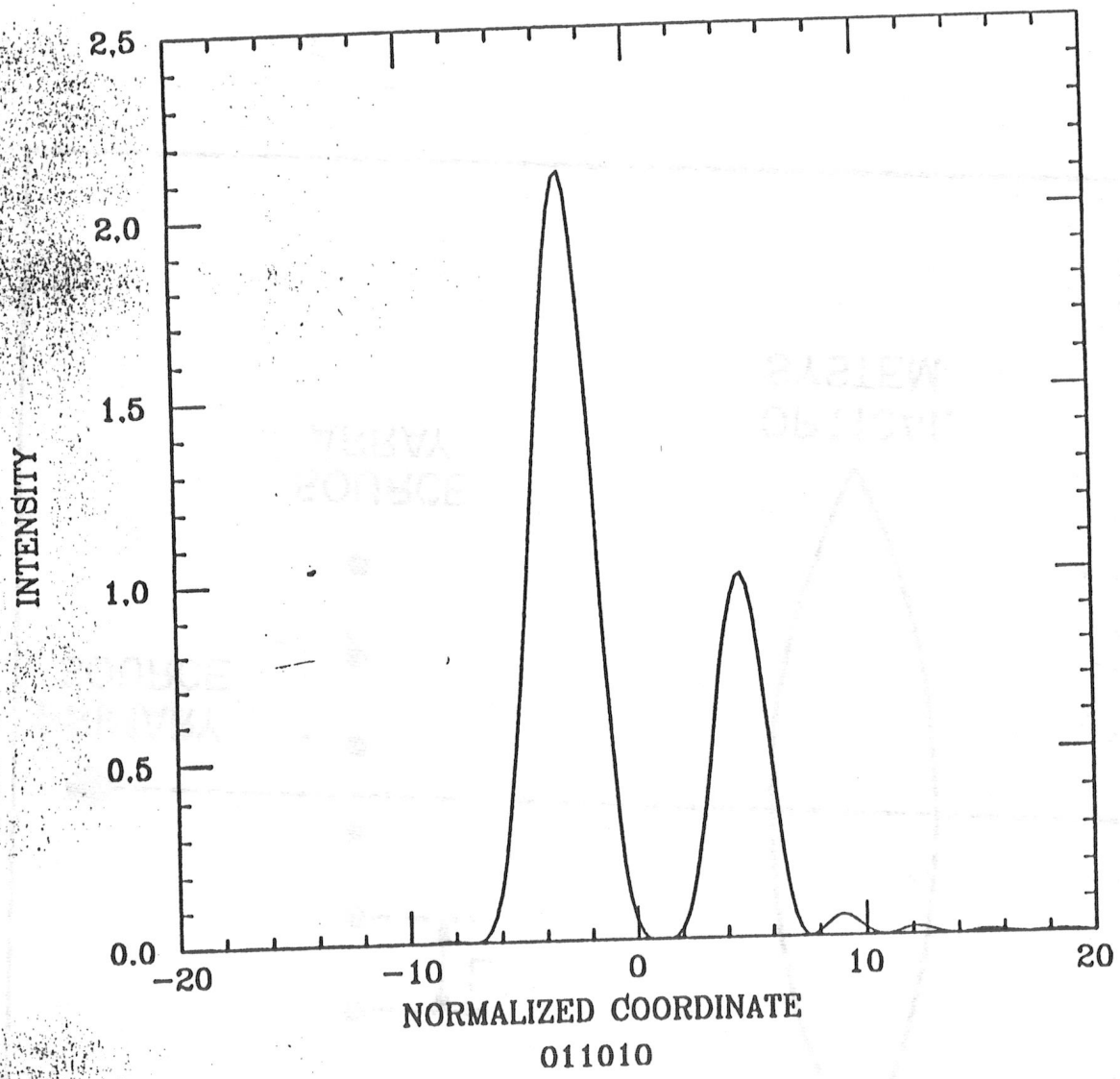
FIGURE LEGENDS

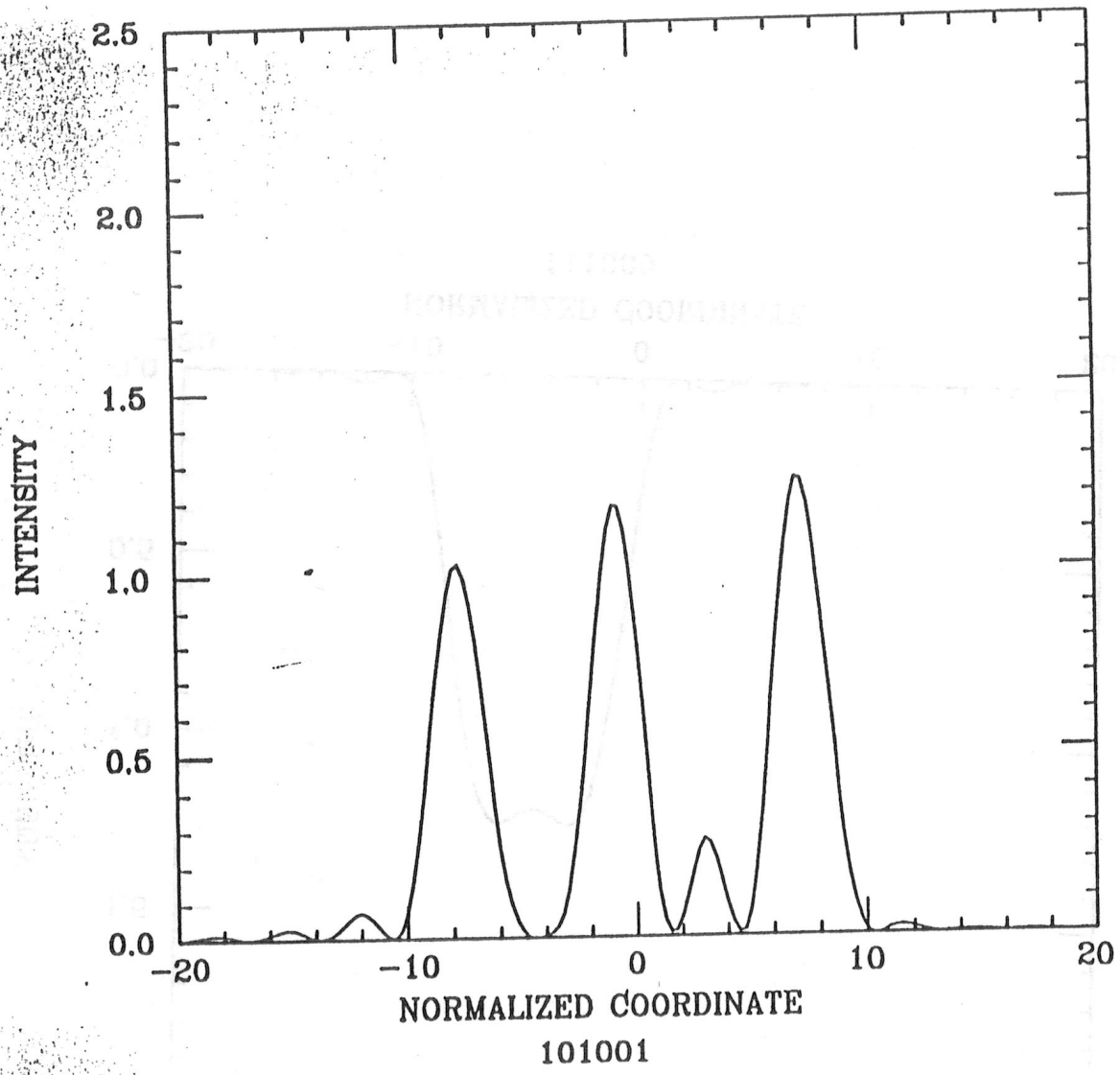
Fig. 1 Diffraction intensity corresponding to source configuration $p_\ell = 0, 1, 1, 0, 1, 0$.

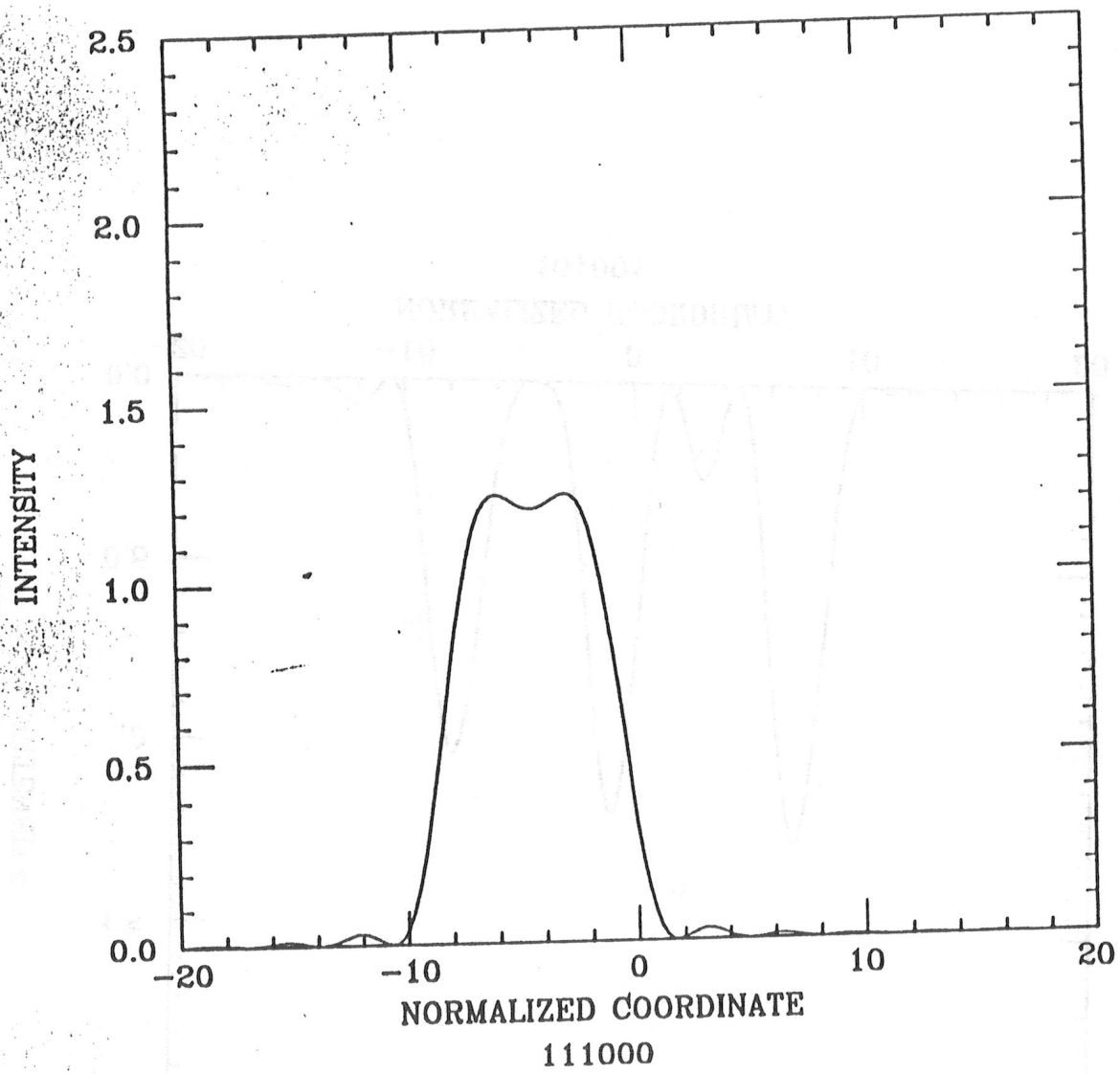
Fig. 2 Diffraction intensity corresponding to source configuration $p_\ell 1, 0, 1, 0, 0, 1$.

Fig. 3 Diffraction intensity corresponding to source configuration $p_\ell = 1, 1, 1, 0, 0, 0$.


L
D
PRIMARY SOURCE
SOURCE ARRAY
OPTICAL SYSTEM
DETECTOR ARRAY


2.5
2.0
1.5
1.0
0.5
0.0
INTENSITY
−20
−10
0
10
20
NORMALIZED COORDINATE
011010


2.5
2.0
1.5
1.0
0.5
0.0
INTENSITY
-20
-10
0
10
20
NORMALIZED COORDINATE
101001


2.5
2.0
1.5
1.0
0.5
0.0
INTENSITY
-20
-10
0
10
20
NORMALIZED COORDINATE
111000

H 25