Lawrence Berkeley National Laboratory
Recent Work

Title

FOURIER TRANSFORMS AND THE HUMAN VISUAL PROCESS: AN APPROACH VIA OPTICAL ILLUSIONS

Permalink

https://escholarship.org/uc/item/7mt7c6k6

Author

Pearlman, Andrew L.

Publication Date

1972-06-16

FOURIER TRANSFORMS AND THE HUMAN VISUAL PROCESS: AN APPROACH VIA OPTICAL ILLUSIONS

DONNER LABORATORY

Andrew L. Pearlman
(Senior Thesis)

June 16, 1972

DISCLAIMER

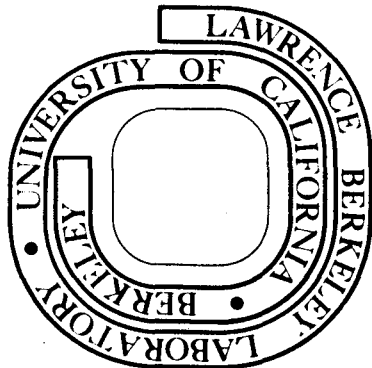# FOURIER TRANSFORMS AND THE HUMAN VISUAL PROCESS: AN APPROACH VIA OPTICAL ILLUSIONS

FOREWORD

I wish to express my deep gratitude to Dr. Robert Glaeser, friend and advisor to me during the many months of work that have gone into this thesis, whose patient and skillful assistance has been indispensable to me in this and other endeavors in the Medical Physics Department of U.C. Berkeley. I thank, also, all those who, each in his own way, have contributed to the completion of this thesis: secretaries, computer programmers, personal friends, and others. Their help has been deeply appreciated.

This thesis is dedicated to all those who make science a living, creative endeavor; and especially to the beloved memory of the late Professor Aharon Katzir-Katchalsky, a true giant in this right. May the boundless energy and depth of his remarkable personality continue to inspire those who were fortunate enough to have known him, to keep science a creative expression of life itself, and not a bland extraction thereof.

Submitted 16 June 1972

Andrew L. Pearlman, B. A.

ABSTRACT

Much interest has been expressed, of late, in the possible involvement of the Fourier transform in the primate visual process. Such involvement is investigated here, via Fourier transformation of well known optical illusions, which are used as inputs to a Fourier model of the visual process because they constitute known failures of the real visual system. Through study of such failures, knowledge of the working system may be gained, as is the case in the field of genetics, where mutations have been valuable tools of research. The essential features of the visual pathway are outlined, and an extensive introduction to the Fourier transform and its applications presented. A current model involving Fourier transforms is discussed, and arguments for the significance of optical illusions in testing that, and other models, are presented. A common geometrical illusion, the Müller-Lyer illusion, is mathematically described and its Fourier transform calculated analytically. Finally, the effects of spatial filtering of a Fourier representation of the Müller-Lyer illusion are investigated, using computer methods. Numerous different filter functions are tested, to accommodate different neurophysiological schemes, and it is shown that any centrosymmetric filter applied to the Müller-Lyer Fourier array will give rise to the known illusory effect. It is concluded that as far as geometrical illusions and ordinary geometric figures are concerned, the hypothesis that Fourier transformation is involved in the primate visual process is confirmed, provided

that some biologically realistic method of spatial filtering is
included in the hypothesis.

ILLUSTRATIONS BY SECTION

INTRODUCTION

For many years, a precise understanding of the nature of human visual information processing has evaded the efforts of researchers, largely due to the enormous complexity of the neural networks involved. The lack of detailed information on the morphology and electrophysiology of the visual pathway continues to prevent any conclusive model for the system as a whole, though much progress has been made towards elucidation of numerous elements thereof. The externally verifiable properties of the human visual system, such as invariance of object recognition with wide variance in object position, orientation, or visual context, have played a dual role in the efforts to form a model: they constitute a major challenge to any prospective model of human pattern recognition, but also provide useful guidelines for engineers and information scientists attempting to apply the general principles of pattern recognition to the visual pathway.

One of the most controversial and interesting models under current investigation is that in which a Fourier transformation takes place within the visual pathway. That is, according to the model, the light information striking the surface of the retina is transformed into a different kind of data by the time signals reach the cortex of the brain. As the visual pathway is so complicated, there is no direct method of verifying the model on the basis of the neurophysiology involved--no "poll" of cells in the visual pathway can be realistically taken which would indicate

whether the cells contain Fourier components or not--therefore, researchers are forced to resort to more indirect methods of verification. One such means is to give the model an input stimulus which is known in the actual human visual system to give rise to an illusory effect, and see if the model detects the illusion. This approach is the one taken here, for reasons given at a later point, and it is suggested that it has proven fruitful in this case.

The experimental approach chosen for this thesis is based on several important concepts, which in turn are based on a somewhat extensive mathematical and biological framework. In order that these concepts can be made clear, an outline of the essential features of the primate visual pathway is presented, as well as a fairly thorough introduction to the Fourier transform and its applications. In the description of the visual pathway, it is presumed that the reader has some familiarity with the vocabulary of neurophysiology, while the mathematical introduction to the Fourier transform is written with the assumption that the reader has some familiarity with calculus, complex numbers and linear algebra (vectors and matrices). It is intended that after reading the mathematical introduction, the reader will feel comfortable with the basic properties of Fourier transforms and their applications, and will be equipped to understand the various manipulations of transforms that are involved in the experimental approach.

The reader should note that all illustrations are arranged by section, and are located at the end of each section (unless otherwise indicated).

OUTLINE OF THE VISUAL PATHWAY

The principle features of the enormously complex network that constitutes the visual pathway should be understood before any models of the system are proposed or studied, if the models are to have any basis in physiological reality. A schematic and simplified summary of this network is outlined below for monocular vision. The author recognizes that the brevity of this discussion requires that it be incomplete, but believes that the most important elements are covered sufficiently for the reader to understand the material in this thesis.

Light waves passing through the pupil and the lens are focused on the retina through a liquid intermediary medium called the vitreous humor (Fig. 1). The photoreceptive layer consists of ~126 million photosensitive cells, of which ~120 million are rod-shaped and are thus called rods, and ~6 million are called cones, for their conical shape. These cells convert light energy into electrochemical energy, excite the neural cells attached to them, known as the bipolar cells, and stimulate surrounding rods and cones via the horizontal cells.

The bipolar cells, of which there are many types, link the rods and cones to the ganglion cells, usually directly, but often via amacrine cells, whose numerous extensions spread the signal to many ganglion cells. The complex network of interconnections between the photoreceptors and the optic nerve is illustrated schematically in Fig. 2. As can be seen in Fig. 2, the flow of information is by no means unidirectional from retina to optic nerve, and even involves transmission of information from the higher centers of the brain to the rods and cones, via so-called "centrifugal"-type bipolar cells.

In the mammalian retina, which is a multilayered structure, the photoreceptive layer of rods and cones lies farthest away from the object in space, and thus it is not surprising that all the other neural layers which precede it are highly transparent (1). The bipolar, horizontal and amacrine cells respond to excitation not by generating an action potential or impulse, but by continuously varying the degree of polarization of their cell membranes (1). The first cells in the visual pathway to translate the light intensity information at the level of rods and cones into neural impulse frequency, or firing rate, are the ganglion cells, whose long cell bodies (axons) merge to form the optic nerve.

The retinal surface is non-homogeneous in many respects--one of particular importance is that the cells in the portion lying on the visual axis (Fig. 1), which constitute the fovea, are much smaller and more densely packed than in the extremities of the retina. In this region, each cone is connected to a single, corresponding bipolar cell, whereas in other retinal areas, many photoreceptors feed into a bipolar cell. As a result, the foveal region is the area of highest visual acuity on the retina, and this accounts for the fact that the best visual detail is seen when the object of interest is centered in the field of view.

As shown in Fig. 2, <u>each</u> ganglion cell is linked to <u>many</u> photoreceptors via the network of bipolar, horizontal and amacrine cells; the portion of the retina to which a given ganglion cell responds is called its <u>receptive field.</u> The receptive field of a ganglion cell is usually composed of two parts: an excitatory portion, in which illumination will cause an increase in the rate of electrical impulse generation by the ganglial cells (increased firing rate), and an inhibitory portion, which has the opposite effect when illuminated. Illumination of a randomly selected point in the receptive field many either increase or decrease the firing rate of a given ganglion cell, so that a single ganglion may simultaneously receive several excitatory and several inhibitory impulses from a complicated image on the retina. The response of the ganglion to such stimulus is an averaged or net response to the multiplicity of inputs (10).

When the excitatory part of the receptive field of a ganglion cell is illuminated with light intensity I, the ganglion responds by firing impulses of equal amplitude at a frequency dependent upon I; in general, the higher is I, the greater the frequency (but not the amplitude) of the ganglion impulses (1, 10). The precise mathematical relationship between the intensity distribution (the focused image) on the retina and the response of the corresponding ganglion cells has yet to be elucidated, but it is clear that ganglion cell behavior is not a simple function of the intensity distribution. For example,

it has been shown that the information at the ganglion level corresponds to a modified version of the focused image, in which borders and boundaries of objects have been greatly enhanced, and the sensitivity to movement in the visual scene has been increased (10). The axons of the ganglion cells merge to form the optic nerve, and carry the retinal information to the lateral geniculate body via the chiasm, a junction region wherein half the optic nerve fibers from each eye cross over to the opposite side of the brain (Fig. 3).

In the lateral geniculate body, a multilaminar structure, the retinal ganglion cells synapse with the so-called geniculo-calcarine neurons, which connect the lateral geniculate to the visual cortex. Recent evidence indicates that neurons carrying information from other senses than vision also synapse with the ganglion and geniculo-calcarine neurons in the lateral geniculate body (14), supporting the theory that there is direct interaction between the senses prior to processing in the brain. In trans-ferring the information from ganglion to geniculo-calcarine cells, the lateral geniculate further modifies the infor-mation, although the mathematical nature of the modification is not precisely known. However, the information passed by the geniculo-calcarine neurons to the striate cortex is thought to be still, essentially, a one-to-one mapping of the retinal information (4, 1), and thus the cortex receives as its visual input a modified but more or less direct transfer of the focused image on the retina. One should not underestimate the importance

of the modifications, though, as they are surely essential to the visual process and are enormously complicated. For example, there is considerable evidence that many axons of the optic nerve carry signals from the brain to the retina, suggesting a feedback system (1, p. 109).

Before discussing the striate or visual cortex, some important additions should be made to the above comments on the subcortical part of the visual pathway. First, as has been pointed out by Westheimer (17), the input to the visual system is resolution-limited from the start by the diameter of the pupil, which acts as a variable aperture on the lens, and by the constant, random movement of the eye. These factors combine to produce an effect known as line spread, whereby a thin line placed in the field of view is degraded to appear spread out (thicker) in a gaussian or other distribution on the retina. For a pupil of diameter 3 mm, one experimentally determined best-fit line spread function is $f(x) = e^{-(.7|x|)}$ (17), which means that a vertical line is distorted as follows:

Intensity

Cross section of vertical line before entering eye

Line spread

Intensity

Cross section of vertical line on retina, after line spread

Minutes of arc on retina
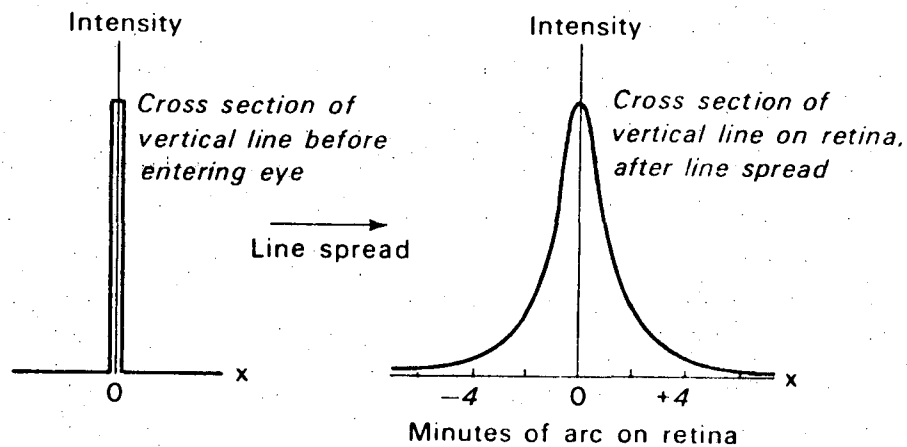
Fig. 4

Other experiments have shown that the minimum angular separation

between parallel lines (angle measured on the surface of the

curved retina) that the retina can detect is 0.6 minutes of arc,

under the most favorable conditions for high resolution viewing

(17). This is then an upper limit one could expect for the

resolution of the entire visual system.

Another important point is the nature of the receptive

fields of the ganglion and geniculate cells. The ganglion cell

responds to a more or less circular region of the retina, the

region having a smaller concentric subset (Fig. 5). If illumi-

nation of the central portion gives rise to increased ganglial

firing rate, illumination of the concentric surround inhibits

the rate, and this type of receptive field is called an "on"-

center field. Similarly, an "off"-center receptive field has

an inhibitory center and an excitatory surround. For a given

cell and intensity of illumination, the response to illumination

of the retina varies widely according to the portion of the

receptive field illuminated. Two beams of light striking

different portions of the "on" central region produce a greater

increase in firing rate than either beam alone, whereas the two

beams have little combined effect if one strikes the center and

the other strikes the surround (10) at the same time. Diffuse

illumination of the entire retina has far less effect on a

given ganglion cell than a single spot of light that strikes

only the excitatory center of its receptive field. Geniculate

cells have very similar receptive fields to those of ganglion

cells, the basic shape being circular with concentric excitatory and inhibitory regions. A significant difference between the two, however, is the greatly enhanced capacity of a geniculate cell's periphery to cancel the effects of its center (10). This implies that geniculate cells are even more sensitive to the precise distribution of intensity in the focused image than the ganglion cells, and thus the lateral geniculate body effectively increases the disparity, already present in ganglion cells, between response to small spots of light or sharp edges, and diffuse light.

The visual cortex is a vastly more complicated structure than the lateral geniculate body, consisting of a multilayered sheet of billions of cells, folded and convoluted on itself to form a compact structure. The geniculo-calcarine neurons attach to the cortex in the layer of cells that is fourth from the top of the cortex, the layer known as area 17 (4). From here the information is eventually disseminated to all the many layers of the cortex by rich interconnections between them. The cells of the visual cortex respond to illumination of definite areas of the retina, but their receptive fields are not concentric--a distinct difference from the ganglion and geniculate cells (10). Hubel and Wiesel (10) found two types of cortical cell, termed "simple" and "complex", for their respective responses to optical stimuli at the retinal level. A simple cell responds to line stimuli--such shapes as slits (bright lines on dark background), dark bars (dark lines on a light background) and

edges (straight boundaries between a light and a dark area), when located in a specific region of its receptive field, and respond maximally when the line is in a specific orientation. Such cells respond to movement of a stimulus on the retina only as the stimulus crosses a very narrow boundary between "on" and "off" regions, and it is thus said that simple cells have long, narrow receptive fields ("on" centers). Complex cells also respond to line stimuli, but are far more sensitive to angular orientation than simple cells, while being far less specific as to position of the stimulus on the retina. Complex cells respond to stimuli of a particular orientation, respond _maximally_ to those of a _particular_ width, and show virtually no variation in their excited firing rate as line stimuli of the proper orientation are moved across large segments (around 20% each) of the total retinal area. Thus, their receptive fields cannot be understood in terms of clearly demarked "on" and "off" regions. Rather, such behavior strongly implies superposition of similarly-oriented receptive fields, which would be the case if each complex cell received its input from many simple cells whose receptive fields all had the same orientation (10, p. 63). Such a scheme would call for inter-connections between similar simple cells and their complex counterparts, and Hubel (10) has found strong physiological evidence for such interconnections: "Functionally, the cortex is subdivided like a beehive into tiny columns or segments, each of which extends from the surface [of the cortex] to the white matter lower in the brain. A column is defined not by any

anatomically obvious wall--no columns are visible under the

microscope--but by the fact that the thousands of cells it con-

tains all have the same receptive-field orientation". (10, p. 62)

(Fig. 6). The electrophysiological data on which these state-

ments are based can be summarized as follows: as a microelectrode

is pushed into the cortex and the receptive fields of the cells

through which it passes recorded in sequence, all cells show the

same field orientation when the penetration is made in a direc-

tion perpendicular to the surface of the cortical segment. When

the microelectrode is inserted and samples taken at an oblique

angle, the data show first a few cells with the same receptive

field orientation, then several cells with a common, new

orientation, and so on, as though the electrode were passing

from column to column. Receptive field orientation appears to

be the only common denominator of the cells of a given column,

however, as they may be simple or complex cells, respond to

slits only, or respond preferentially to dark bars or edges.

Anatomically, the columns are irregular in cross sectional shape

but average about 0.5 mm^2 in cross-sectional area. Also, it is

well known that the preponderance of interconnections among

cortical cells is in a direction perpendicular to the cortical

surface, and this fits well with the long, narrow, more or less

cylindrical shape of the columns. The relative lack of

connections between cells of different columns implies that a

column may be considered an independent functional unit of the

cortex, in which simple cells receive connections from lateral

geniculate cells and send projections to complex cells. The
result of this is that a line stimulus with a given orientation
will most strongly stimulate the cells of the column whose
receptive fields correspond to that orientation. As the areas
of the retina represented in each column overlap with one another
to a great extent, each small region of visual field is represented
"over and over again, in column after column [of the cortex], first
for one receptive-field orientation and then for another." (10,
p. 63). This extensive redundancy will be mentioned again, as a key
ingredient to the Fourier model for visual information processing
and storage.

This completes the brief sketch of the visual pathway
intended for this thesis. At least one major point remains to
be emphasized, though, and it should be borne carefully in mind.
The foregoing description is of the machinery of vision and
pattern recognition, and some of the phenomenology thereof. The
precise mechanisms by which the phenomenology is linked to the
machinery have, in most cases, yet to be elucidated, and many
models are presently attempting to make the connection. However,
there is a clear concensus of scientific opinion on the visual
process as a whole, that there is a distinction between the
process of visual information processing and that of interpreting
the processed information. Anatomically, the former is believed
to take place between the retina and striate cortex (the sub-
cortical part of the pathway) and the latter in centers deeper
within the brain, though feedback networks between the two are
a virtual certainty. The mutual interdependence of the two
systems is clearly indicated by the fact that a change in the

information being processed visually causes a change in the interpretation thereof, and conversely, that prior knowledge about an input to the system may affect the way it is processed. An example of the latter is the phenomenon of figure-ground reversal, wherein an ambiguous picture is viewed, in which it is difficult to distinguish between object and surroundings. After one has been informed which is which, the same input elicits a different interpretation from that when uninformed.
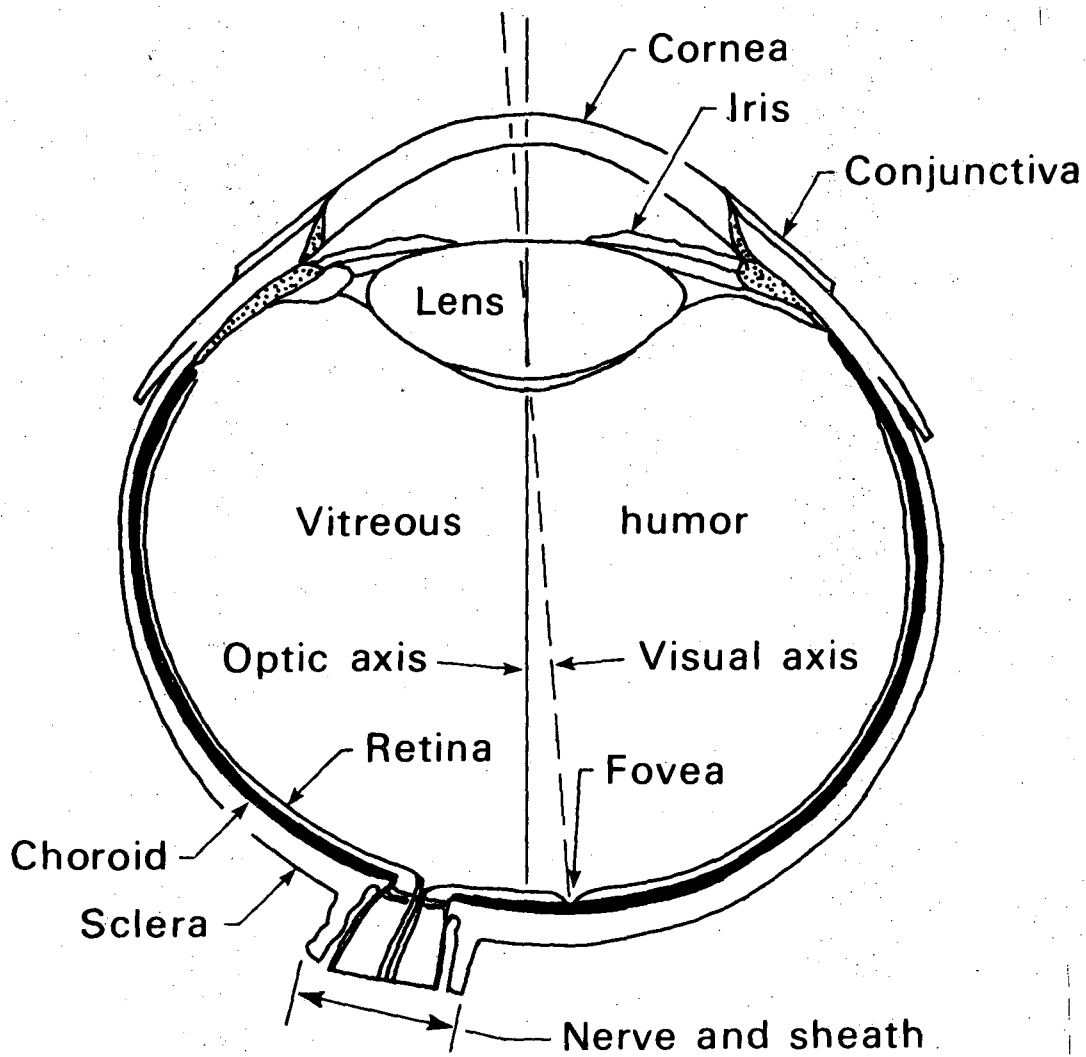
Fig. 1

DBL 728 5423

FIGURE 2 (EXPLANATION)

The structure of the primate retina reduced to its essentials, including the synopsis of the propagation of the retinal impulses from the photoreceptors to other parts of the retina, to the brain, and from the brain back to the retina (direction indicated by arrows).

Labling of the cells: a, b, rods and cones, or the photoreceptors; c, horizontal cells by means of which the impulses are transmitted to the surrounding rods and cones; d, e, f, h, centripetal bipolar cells of the mop, brush, flat, and midget varieties, which "transmit" the impulses from the photoreceptors to the ganglion cells, the bipolars serving as "analyzers"; i, centrifugal bipolar cell, a variety of the "amacrine cells," which probably receives the impulses from the centripetal bipolars, from the ganglion cells, and also from the brain by way of the centrifugal or efferent fibers (t) and transmits them back upon the photoreceptors (a, b); 1, an "amacrine cell" which possibly intercepts a part of the bipolar impulses and spreads them over the surrounding territory; m, n, o, p, s, ganglion cells which receive impulses from the centripetal bipolars and transmit them to the brain along their axons called "optic nerve fibers." (Polyak, The Retina, University of Chicago Press, 1941.)
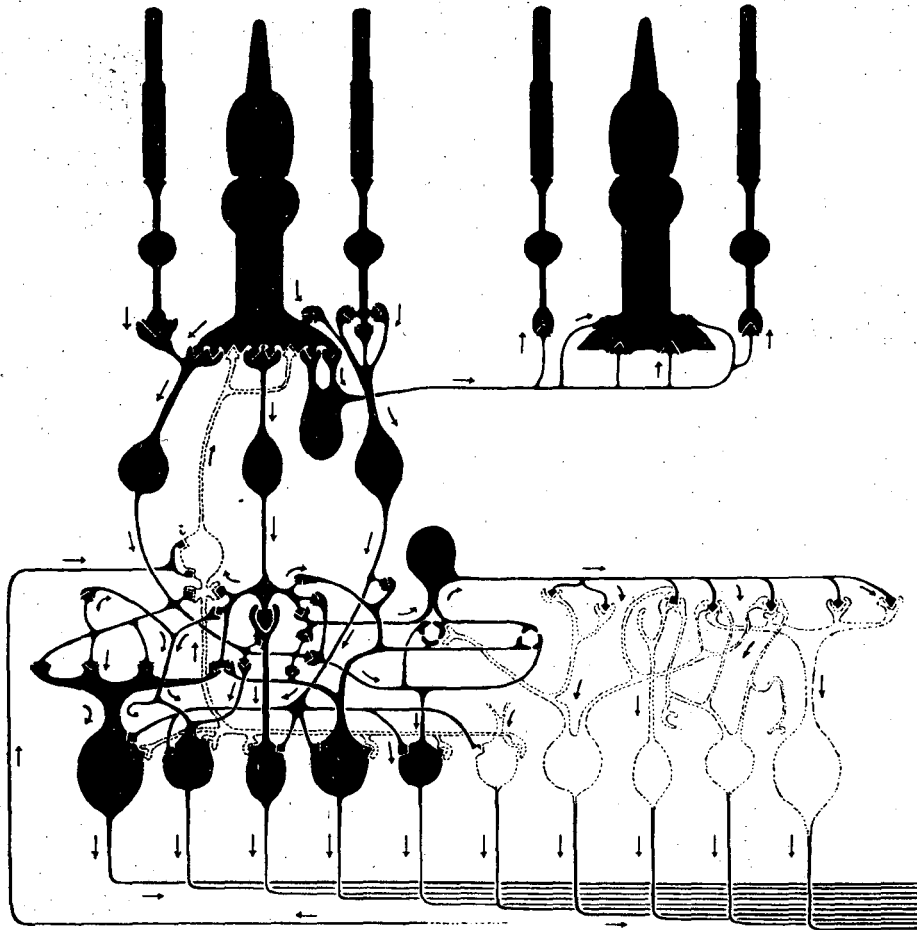
Fig. 2
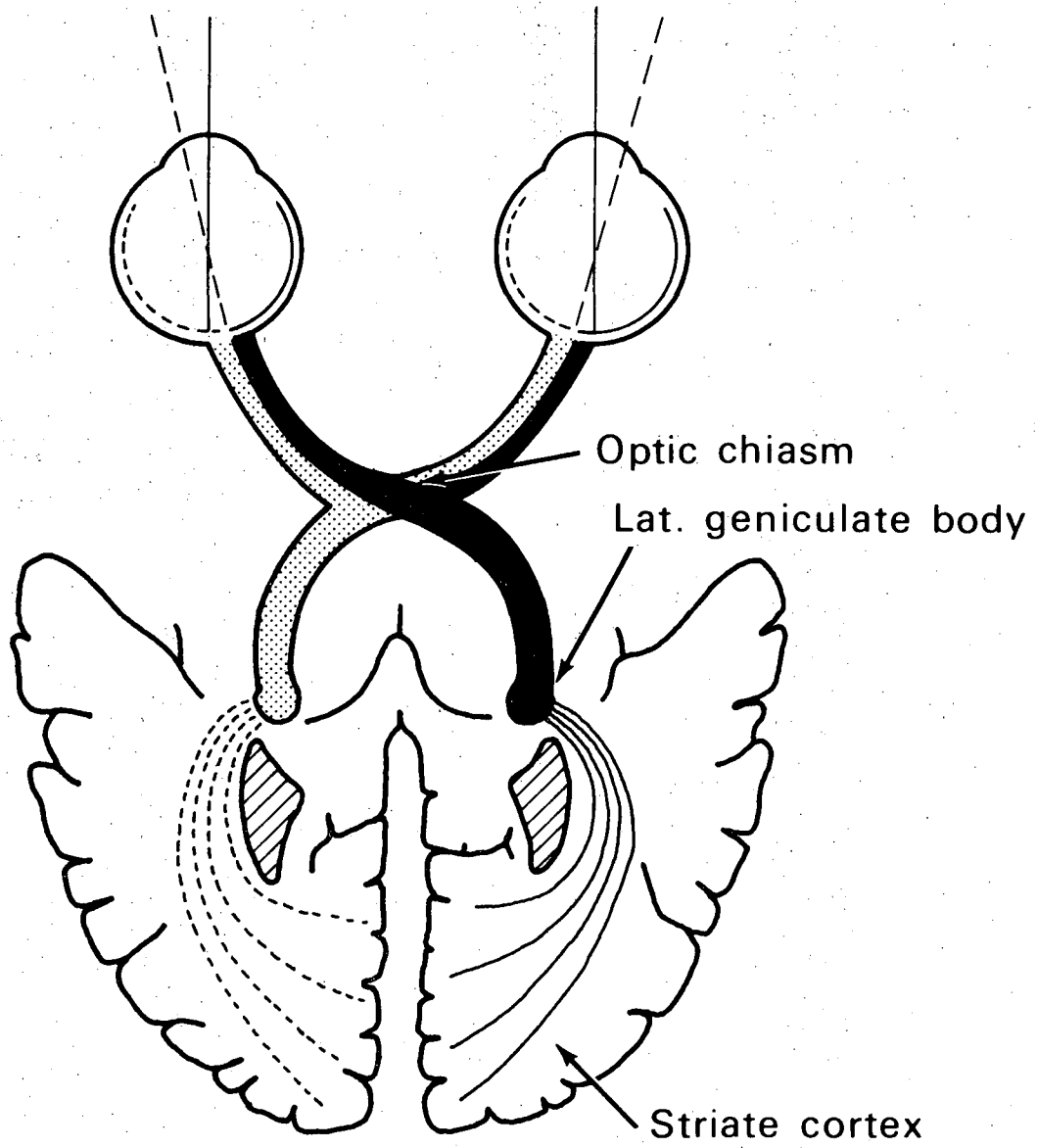
Fig. 3

DBL 728 5422
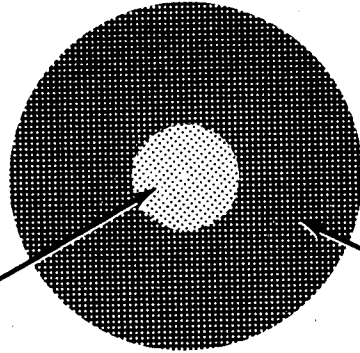
"ON" CENTER FIELD

Stimulation here excites nerve

Stimulation here suppresses nerve

"OFF" CENTER FIELD

DBL 728-5424

Fig. 5

Afferent fibers from lateral geniculate body

Axons of cortical cells

WHITE MATTER

GRAY MATTER

Electrolytic lesion
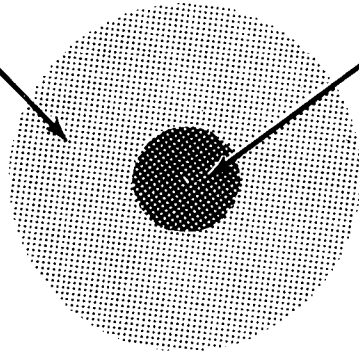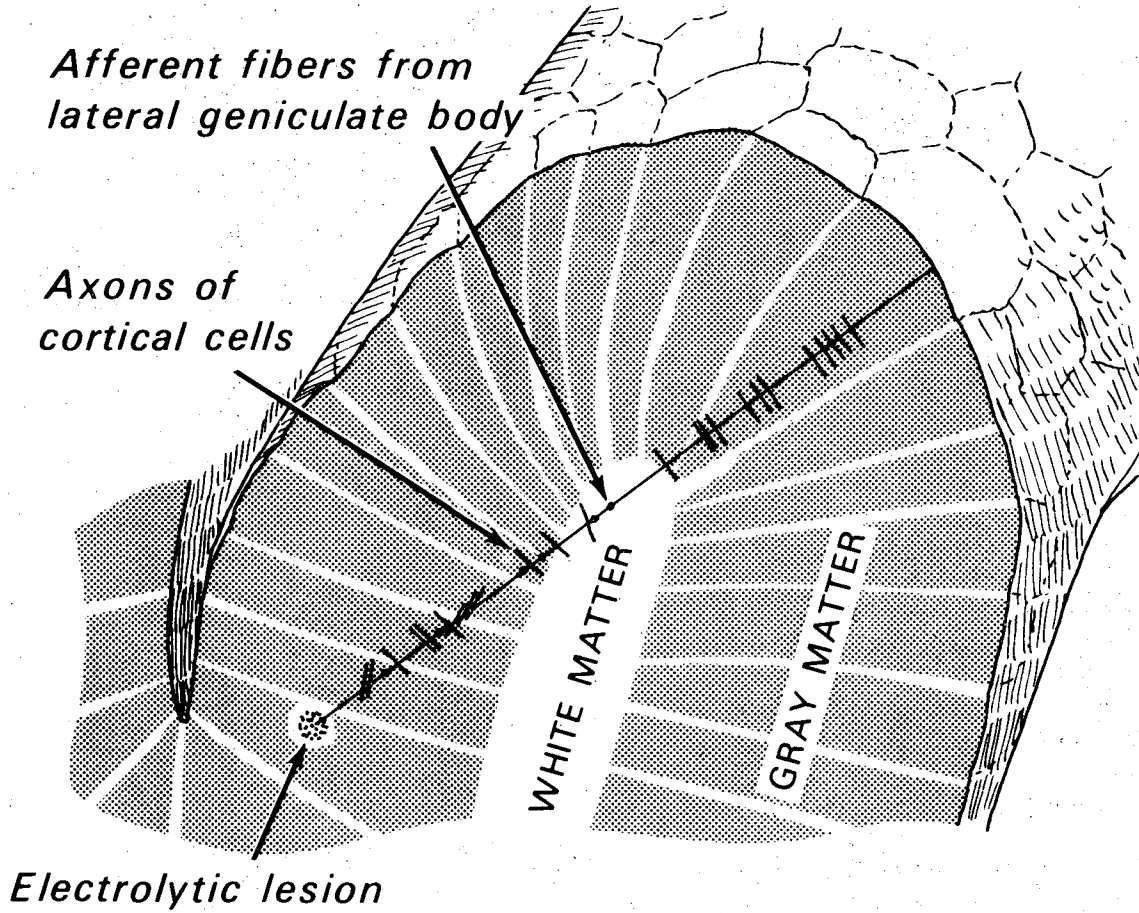
Redrawn from: SCIENTIFIC AMERICAN

DBL 728 5421

Fig. 6

FIGURE 6 (EXPLANATION)

Functional arrangement of cells in the visual cortex

resembled columns, although columnar structure is not apparent

under a microscope. Lines A and B show paths of two micro-

electrode penetrations: colored lines show receptive-field

orientations encountered. Cells in a single column had same

orientation; change of orientation showed new column.

(From Hubel, Reference 10, p. 62)

MATHEMATICAL BACKGROUND

One of the basic concepts of linear algebra is that of a vector space. A two-dimensional vector space, for example, consists of the plane in which two nonparallel vectors $\vec{A}$ and $\vec{B}$ lie. That is, the set of all points (vectors) which can be located by linear combinations $C_1\vec{A} + C_2\vec{B}$ constitutes a two dimensional vector space, and that space is a plane. The two vectors, $\vec{A}$ and $\vec{B}$ are said to constitute a <u>basis set</u> for the space in which they lie, and if they are perpendicular to one another, form an <u>orthogonal</u> basis set for the plane. In either case, $\vec{A}$ and $\vec{B}$ are said to <u>span</u> the plane, meaning that they linearly combine to completely determine all points on the plane. It should be noted that there are infinitely many pairs of orthonormal basis vectors in a given plane. In general, a space which is spanned by n independent vectors (i. e., no one vector can be generated by linear combinations of the others), is said to have dimension n, and the set of n such vectors constitutes a basis set for vector n-space.

The extent to which one vector "overlaps" with another is expressed by the <u>projection</u> of the one onto the other, as seen in the illustration below:

$\|\vec{A}\| \cos \theta$
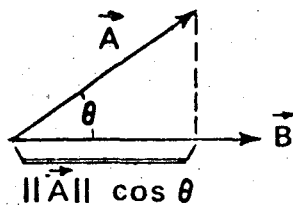
Fig. 1

A measure of the amount of projection is the <u>inner product</u> between $\vec{A}$ and $\vec{B}$, also known as the dot product because of the symbolic form in which it is usually written, $\vec{A} \cdot \vec{B}$. If $\vec{A}$ is (a_1, a_2) and $\vec{B}$ is (b_1, b_2), $\vec{A} \cdot \vec{B} \equiv a_1 b_1 + a_2 b_2 = \|A\| \|B\| \cos \theta$, where $\|A\|$ means the <u>length</u> of $\vec{A}$ and is given by $\|A\| = \sqrt{a_1^2 + a_2^2}$, and θ is the angle between $\vec{A}$ and $\vec{B}$ (see figure above). If $\vec{A}$ and $\vec{B}$ are orthogonal, then $\theta = \frac{\pi}{2}$

and $\|A\| \cdot \|B\| \cos \theta = 0$. Thus, two orthogonal vectors have an inner product of 0. The inner product can be applied to vectors in n-space by extending the definition to be:

If $\vec{A} = (a_1, a_2, \cdots, a_n)$ and $\vec{B} = (b_1, b_2, \cdots, b_n)$, then

$$\vec{A} \cdot \vec{B} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i. \text{ If } \sum_{i=1}^{n} a_i b_i = 0,$$

the two vectors are orthogonal, as in the two dimensional case. If $\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n$ are an orthogonal basis set for vector n-space and <u>each</u> <u>have length 1</u>, they constitute an <u>orthonormal basis set</u> for vector n-space. Any vector $\vec{A} = (a_1, a_2, \cdots, a_n) = a_1 \vec{i}_1 + a_2 \vec{i}_2 + \cdots + \vec{a}_n \vec{i}_n$, where the vectors $\vec{i}_1, \vec{i}_2, \cdots, \vec{i}_n$ are one basis set for the space, must also be expressable as a linear combination of the basis vectors $\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n$. That is, $\vec{A} = C_1 \vec{e}_1 + C_2 \vec{e}_2 + \cdots + C_n \vec{e}_n$, where the $C_1, C_2, \cdots, C_n$ are constants. What are the constants? The sum of n vectors is another vector, so it is appropriate to speak of a dot product (inner product) of the right side of the above equation, with another vector. Let's take the inner product of both sides of the equation with the vector $\vec{e}_j$, one of the basis vectors. Since the inner product is distributive over addition, we have:

$$\vec{A} \cdot \vec{e}_j = (C_1 \vec{e}_1 + C_2 \vec{e}_2 + \cdots + C_j \vec{e}_j + \cdots + C_n \vec{e}_n) \cdot \vec{e}_j$$

$$= C_1 \vec{e}_1 \cdot \vec{e}_j + C_2 \vec{e}_2 \cdot \vec{e}_j + \cdots + C_j \vec{e}_j \cdot \vec{e}_j + C_n \vec{e}_n \cdot \vec{e}_j.$$

By their mutual orthogonality, the dot product between one basis vector and another must be 0, ans since each vector has length 1, $\vec{e}_j \cdot \vec{e}_j = 1$, so the sum of the right side becomes $C_j \vec{e}_j \cdot \vec{e}_j = C_j$. Thus, we have $\vec{A} \cdot e_j \equiv <\vec{A}, \vec{e}_j> = C_j$, where $<,>$ is another common notation for the inner product of two vectors, in this case $\vec{A}$ and $\vec{e}_j$. So, each constant, C_j, is determined by the inner product of

$\vec{A}$ with the basis vector $\vec{e_j}$, and therefore,

$$\vec{A} = \sum_{j=1}^{n} <\vec{A}, \vec{e_j}> \vec{e_j} = (C_1, C_2, \cdots, C_n)_e \cdots. \qquad (1)$$

Such a change of basis sets within a space, with its corresponding change in the n-tuple representation of a vector, is called a linear transformation, and will be seen to have direct analogies in the discussion to follow.

The concepts of a vector space can be extended to other domains than the set of real numbers, to that of a function space. To understand the connection, one must first know what constitutes a space. A space is a set or collection of elements that have certain properties in common. A vector n-space is a collection of elements, each of which is an ordered set of n numbers; the number of numbers, n, in each element is the dimension of the space. For example, each point in a plane is defined by an ordered pair of numbers where each "slot" can have any real number $x \epsilon R^1$ (the real line). To say that the dimension of this vector space is "two" is to say that the combination (x_1, x_2), where both x_1 and x_2 are free to take on any real value, yields a collection of all possible points that are elements of R^2, the real plane. The method of combination is called the Cartesian product and is written $R^1 \times R^1 = R^2$. Now R^1, being the set of all real numbers, can be generated and regenerated in many ways; e.g., a real-valued, monotone function $f(x)$ takes a number in R^1 and manipulates it to give another number in R^1, so that R^1 is both the set of all values of x and the set of all values of $f(x)$. For any particular value of x, say $x = 1$, one can conceive of a set of functions $\{f(1)\}$ such that the values of these functions "fill up" the whole real line. Thus, an equivalent way of expressing R^1 is the set $\{f(1)\}$ of all real-valued functions of the number 1. So, upon looking at the concept of a vector space from this standpoint, one can write

an equivalent form of a pair of numbers (x_1, x_2) as a pair of

sets $(\{f(1)\}, \{f(2)\})$, where f(1) is the set of all real

valued functions of 1, and $\{f(2)\}$ is the set of all real-valued

functions of 2. The advantage of this more general formulation

is that vector n-space can be understood as the set of all

functions on the ordered integers $1,2,3,\ldots,n = (\{f(1)\}, \{f(2)\},$

$\{f(3)\}, \ldots \{f(n)\})$, and an important conceptual extension can

schematically be grasped. A vector n-space could then be represented

as follows:

Fig. 2 | ORDERED INTEGERS | ⟶ | ALL FUNCTIONS | ⟶ | VECTOR n-SPACE |
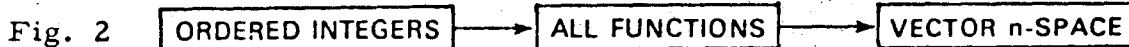
This is extendable to indefinitely large dimension by letting

n increase without bound. A question arises, though: what happens

if instead of using the integers as a start, we use the set of all

real numbers, R^1? The result is called a FUNCTION SPACE over the

real line and has analogies to all the properties of vector n-space.

However, the set of all functions of a real variable is certainly

an enormous set, and we need not deal with so vast and varied a

collection. The set that will be dealt with is the set of all

functions that obey the following condition:

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx < \infty .$$

This set constitutes a space, called an L_2 space, which is a

particular version of a more general notion called a Hilbert space.

We will allow f(x) to be complex-valued, meaning that the function

f takes a real number x and converts it to a number a+ib, where

a and b are real numbers and $i = \sqrt{-1}$. The set of all complex

numbers, $\{a+ib\}$ constitutes the complex plane C^1, whose horizontal

axis is the real line, R^1, and whose vertical axis is the set of all real numbers multiplied by i. The latter set is referred to as the set of imaginary numbers; the figure below depicts the complex plane and some important features thereof:

Fig. 3
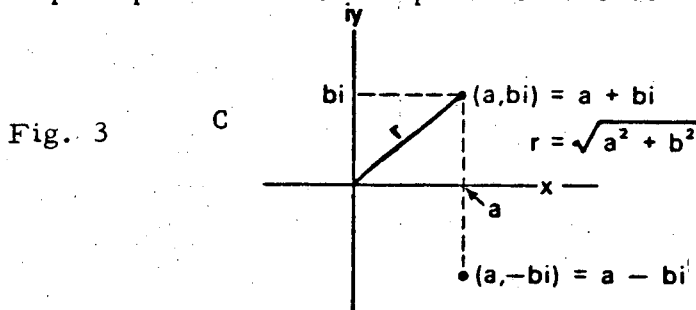
The point a-bi, the reflection of a+bi through the real axis, is called the <u>complex conjugate</u> of a+bi, and is written $a-bi = (a+bi)^*$. Since each member of the range of a complex-valued function is composed of one real and one imaginary term, a small a complex-valued function can be considered as the sum of a purely real-valued function and a purely imaginary-valued function, $f(x) = y(x) + iZ(x)$. The inner product of two functions is analogous to that of two vectors, i.e., if

$$f_1(x) = Y_1(x) + iZ_1(x) \text{ and } f_2 = Y_2(x) + iZ_2(x)$$

where $Y_1(x)$, $Z_1(x)$, $Y_2(x)$ and $Z_2(x)$ are real, then the inner product of f_1 & f_2, is

$$< f_1(x) , f_2(x) > = \int_{-\infty}^{+\infty} f_1^*(x) \cdot f_2(x) \, dx,$$

where $\quad f_1^*(x) = Y_1(x) - iZ_1(x)$.

Also, $\quad < f_2(x) , f_1(x) > = \int_{-\infty}^{+\infty} f_2^*(x) \, f_1(x) \, dx$

$$= (< f_1(x), f_2(x) >)^* .$$

If $f_1(x)$ and $f_2(x)$ are orthogonal, $\int_{-\infty}^{+\infty} f_1^*(x) \cdot f_2(x) \, dx = 0$, and conversely, if $\int_{-\infty}^{+\infty} f_1^*(x) \cdot f_2(x) \, dx = 0$, $f_1(x)$ and $f_2(x)$ are orthogonal. A set of functions $h_n(x)$ forms a basis set for the <u>space of functions</u> <u>defined</u> <u>on a finite interval</u> of the real line, $a \leq x \leq b$, if <u>any</u>

function f(x) on that interval can be expressed as $f(x) = \sum_{n=-\infty}^{+\infty} C_n h_n(x)$. If the set $h_n(x)$ is such that

$< h_i(x), h_j(x) > = 0$ for $i \neq j$, then $h_n(x)$ is an

orthogonal basis set, and if $< h_i(x), h_i(x) > = 1$, then the set

$h_n(x)$ is an <u>orthonormal</u> basis set, and the values of C_n in the

above summation can be determined in precisely analgous manner

to the situation in vector n-space. Thus, $C_j = < f(x), h_j(x) >$

and $f(x) = \sum_{n=-\infty}^{\infty} <f(x), h_n(x) > h_n(x)$ . . . (2)

There are infinitely many sets of orthogonal basis sets

$\left\{ h_n(x) \right\}$, among them being the Taylor, Laguerre, Hermite and other

polynomials, of great importance in quantum mechanics. Each set

may be used to expand f(x) into an infinite series, with

appropriate coefficients of the basis functions being summed.

The large number of functional basis sets is analogous to the situation in a

two-dimensional vector space, where <u>any</u> two perpendicular vectors form

an orthogonal basis set for the space.

It can be shown that the set of complex-valued functions

$\left\{ e^{2\pi i n x} \right\}$, $n = \pm 1, \pm 2, \pm 3, \cdots$, forms an orthogonal basis set for all L_2 functions of

the interval $0 < x < 1$. The proof that $\left\{ e^{2\pi i n x} \right\}$ forms a basis set

for the space is beyond the scope of this paper but can be found

in (13). The proof that $\left\{ e^{2\pi i n x} \right\}$ are orthogonal is straight-

forward, as will be shown: First, note that $e^{2\pi i n x} = \cos(2\pi n x)$

$+ i \sin(2\pi n x)$ by Euler's identity.

Then $(e^{2\pi i n x})^* = (\cos(2\pi n x) + i \sin(2\pi n x))^*$

$= \cos(2\pi n x) - i \sin(2\pi n x) = \cos(-2\pi n x) + i \sin(-2\pi n x) = e^{-2\pi i n x}$.

Then:

The inner product of two functions $f_1(x)$ and $f_2(x)$ over the interval $a \leq x \leq b$, is defined as $< f_1(x), f_2(x) >_{[a,b]}$ $= \int_a^b f_1^*(x) \cdot f_2(x)\, dx$, and will be 0 if and only if f_1 and f_2 are <u>orthogonal</u> on the interval $a \leq x \leq b$. So, putting these notes together, we have that $\left\{ e^{2\pi i n x} \right\}$ will be orthogonal on the interval [0,1] if and only if for every pair $e^{2\pi i m x}$, $e^{2\pi i \ell x}$ (ℓ & m integers $< n$), the inner product is 0 .

Thus:

$$< e^{2\pi i m x},\ e^{2\pi i \ell x} >_{[0,1]} = \int_0^1 e^{-2\pi i m x} e^{2\pi i \ell x} dx$$

$$= \int_0^1 e^{2\pi i (\ell - m) x} dx$$

and if $\ell - m \neq 0$ ($\ell \neq m$), we have

$$\int_0^1 e^{2\pi i (\ell - m) x} dx = \left[\left(\frac{1}{2\pi i (\ell - m)} \right) \left(e^{2\pi i (\ell - m) x} \right) \right]_0^1$$

$$= \frac{1}{2\pi i (\ell - m)} \left[\cos(2\pi(\ell - m)x) + i \sin(2\pi(\ell - m)x) \right]_0^1 . \quad \text{Now}$$

$(\ell - m)$ is an integer, so $\sin(2\pi(\ell - m)) = \sin(0 \cdot (\ell - m)) = 0$, and $\cos(2\pi(\ell - m)) - \cos(0 \cdot (\ell - m)) = 0$, so $< e^{2\pi i m x},\ e^{2\pi(\ell - m) x} >_{[0,1]} = 0$ if $\ell \neq m$. If $\ell = m$, we have $\int_0^1 e^{-2\pi i m x} e^{2\pi i m x} dx = \int_0^1 e^0 dx = 1$.

Another way of writing this is

$$< e^{2\pi i m x},\ e^{2\pi i \ell x} >_{[0,1]} = \delta_{\ell m} , \text{ where } \delta_{\ell m} = \begin{cases} 1 & \text{if } \ell = m \\ 0 & \text{if } \ell \neq m \end{cases}$$

With this background, the Fourier Transform becomes a simple application of the principles stated above.

THE FOURIER TRANSFORM

Suppose $f(x)$ is a function which satisfies the following restriction:

$$\int_0^1 |f(x)|^2 dx < \infty .$$

Then the set of all such functions comprises a function space, and that space is spanned by many different sets of eigenfunctions, among them being the set $\{e^{2\pi inx}\}$. By the previously mentioned property of function spaces, any function defined on a finite interval of the space can be expressed as a linear combination of the functions $\{e^{2\pi inx}\}$,

so
$$f(x) = \sum_{n=-\infty}^{+\infty} C_n e^{2\pi inx} \quad \cdots \cdot (2),$$
where the C_n are complex constants. The C_0 term, also known as the "DC term", is of particular interest, as shall be shown: let the length of the interval on which the function is defined be 1, for some appropriate unit of length. Integrating over this period, we have:

$$\int_0^1 f(x)\, dx = \int_0^1 \sum_{n=-\infty}^{\infty} C_n e^{2\pi inx} dx = \sum_{n=-\infty}^{\infty} C_n \int_0^1 e^{2\pi inx} dx .$$

For any $k \neq 0$, the integral is

$$\int_0^1 e^{2\pi ikx} dx = \left[\frac{1}{ik} e^{2\pi ikx}\right]_0^1 = \frac{1}{ik}\left[\cos 2\pi kx + i\sin 2\pi kx\right]_0^1$$

$= 0$. When $k=0$, the integral is $\int_0^1 e^0 dx = 1$ (length of interval), so,

$$\int_0^1 f(x) dx = C_0 \cdot (\text{interval}), \text{ and } C_0 = \frac{1}{(\text{interval})} \int_0^{(\text{interval})} f(x)\, dx = F(x), \text{ the}$$

average value of the function.

Now, to determine the general constant C_m, we multiply both sides of equ. (2) by $e^{-2\pi imx}$ and integrate from $x = 0$ to $x = 1$. Thus,

$$\int_0^1 e^{-2\pi imx} f(x) \, dx = \int_0^1 \left(\sum_{n=-\infty}^{+\infty} C_n e^{2\pi inx} \right) e^{-2\pi imx} \, dx$$

$$= \sum_{n=-\infty}^{+\infty} C_n \int_0^1 e^{2\pi inx} e^{-2\pi imx} \, dx . \text{ But } \int_0^1 e^{2\pi inx} e^{-2\pi imx} \, dx \text{ is}$$

just $< e^{2\pi imx}, e^{2\pi inx} >_{[0,1]}$ and from previous arguments,

$$< e^{2\pi imx}, e^{2\pi inx} > = \delta mn.$$ So, all terms in the series will be zero except where $n = m$, so

$$\int_0^1 f(x) \, e^{-2\pi imx} dx = 1 \cdot C_m \implies C_m = \int_0^1 f(x) \, e^{-2\pi imx} dx \ldots (3)$$

The set of constants $\left\{ C_m \right\}$ constitutes the DISCRETE FOURIER TRANSFORM of $f(x)$.

Rewriting equ. (2) using the new values for C_m, we have

$$f(x) = \sum_{n=-\infty}^{+\infty} < f, e^{-2\pi inx} >_{[0,1]} e^{2\pi inx} \quad \ldots \tag{4}$$

This is precisely analgous to expansion of a vector, seen earlier. [equ. (1)]. The discrete Fourier transform was defined for the set of all L_2 functions which map the interval $[0,1]$ of the real line into the complex plane. What if functions are to be considered that map (i.e., are defined on) the whole real line into the complex plane? The discrete summation then becomes an integral and the discrete constants C_m become functions $g(u)$ of a continuous variable u instead of the integer-valued K, as follows:

$$\sum_{k=-\infty}^{\infty} C_k e^{2\pi ikx} \longrightarrow \int_{-\infty}^{\infty} g(u) e^{2\pi iux} du \quad , \quad \text{where } g(u) = \int_{-\infty}^{\infty} f(x) e^{-2\pi iux} dx$$

so,

$$f(x) = \int_{-\infty}^{\infty} g(u) e^{+2\pi iux} du, \quad \text{and } g(u) = \int_{-\infty}^{\infty} f(x) e^{-2\pi iux} dx$$

$$g(u) = F\left\{f(x)\right\} = \text{"the Integral Fourier transform of } f(x)\text{"}$$

$$f(x) = F^{-1}\left\{g(u)\right\} = \text{"the inverse Fourier transform of } g(u)\text{"}$$

All of the foregoing arguments can be extended to functions
of more than one variable, x. A two-dimensional L_2 space
is the set of functions defined on the real plane, that
obey the condition:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x,y)|^2 \, dxdy < \infty \, .$$ All the arguments for

inner products, basis sets and expansions follow in an
analgous manner to those for the one-dimensional case, and
the set $\left\{e^{2\pi i \, (kx+ny)}\right\}$ forms a basis set for the L_2 space
on the region $0 < x < 1, 0 < y < 1$ of the real plane. Thus,
the expansion of a function f (x,y) on that region is given
by $f(x,y) = \sum_{k=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} C_{kn} e^{2\pi i \, (kx+ny)}$, where $C_{kn} = \int_{0}^{1}\int_{0}^{1} f(x,y) e^{-2\pi i(kx+ny)} dxdy$

When the interval of the real plane becomes the whole plane, this
becomes an integral transform, with Fourier variables u, v
instead of integers k,n, and we have

$$f(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(u,v) \, e^{2\pi i \, (ux+vy)} dudv \quad , \quad = F^{-1} \left\{g(u,v)\right\} . \quad (5)$$

where

$$g(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) \, e^{-2\pi i \, (ux + vy)} \, dxdy = F\{f(x,y)\} \ldots (6)$$

$F\{f(x,y)\}$ means, "the Fourier transform of $f(x,y)$" and

$F^{-1}\{g(u,v)\}$ means "the inverse Fourier transform of $g(u,v)$".

Several important properties of Fourier transforms emerge from

the definitions:

1. INVERTIBILITY OF THE FOURIER TRANSFORM:

$$F^{-1}\left\{F\left\{f(x,y)\right\}\right\} = f(x,y)$$

and

$$F\left\{F^{-1}\left\{g(u,v)\right\}\right\} = g(u,v)$$

The detailed proofs of these two identities are too lengthy

to be of use in this paper, but it is hoped that these properties

seem at least to be reasonable results of the definitions.

2. LINEARITY OF THE FOURIER TRANSFORM:

$$F\left\{f(x,y) + g(x,y)\right\} = F\left\{f(x,y)\right\} + F\left\{g(x,y)\right\} \text{ and}$$

$$F\left\{cf(x,y)\right\} = cF\left\{f(x,y)\right\}, \text{ c being a constant.}$$

This property follows directly from the linearity of the

integrals involved in the definitions.

3. THE SHIFT THEOREM:

The Fourier transform of a function that has been shifted

with respect to the origin is the same as that of the non-

shifted function, multiplied by a phase factor:

$$F\left\{f(x-a,y-b)\right\} = F\left\{f(x,y)\right\} e^{-2\pi i \, (ua + vb)} \ldots (7)$$

This is shown as follows for the one dimensional case, which is

simpler but precisely analgous to the two dimensional case:

$$\int_{-\infty}^{\infty} f(x-a)\ e^{-2\pi iux}\ dx = F\left\{f(x-a)\right\}$$

let $y = x-a$; then $dx = dy$ and $x = y+a$

$$F\left\{f(y)\right\} = \int_{-\infty}^{\infty} f(y)\ e^{-2\pi iu\,(y+a)}\ dy = \int_{-\infty}^{\infty} f(y)e^{-2\pi i\,(uy)}e^{-2\pi i\,(ua)}\ dy$$

$$= e^{-2\pi iua}\int_{-\infty}^{\infty} f(y)\ e^{-2\pi iuy}\ dy = F\left\{f(y)\right\}e^{-2\pi iua} = F\left\{f(x)\right\}e^{-2\pi iua}$$

(y and x are arbitrary names for all points on the real

line and are thus "dummy variables" that are interchangable)

An interesting and important consequence of the shift theorem

is the positional invariance of the Fourier transform, except

for a phase term $e^{-2\pi i\,(au+bv)}$, where a and b represent

the x and y displacements of the function $f(x,y)$ from the

origin of the x-y plane. This is a crucial property for any

proposed mathematical model of the visual system, as it is

well-known that recognition of objects is invariant to their

linear translation in the field of view. However, the phase

term <u>does</u> indicate a function's <u>position</u> relative to the origin.

A good physical insight into the action of the Fourier

transform on a function can be had from an examination of the

complex exponential factor, known as the PHASE FACTOR in the

integral. Note that when $ux + vy = n$, any integer, the phase

term is
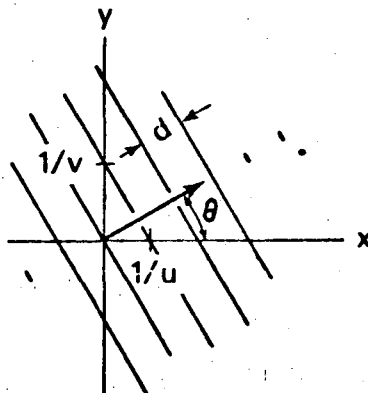
Fig. 4

$$e^{-i2\pi n} = \cos(-2\pi n) + i\sin(-2\pi n) = \cos 2\pi n = 1 = e^0 \; .$$

Such a situation is referred to as zero phase, and for any u, v we have

$$ux + vy = n \quad \text{and} \quad y = \frac{n}{v} - (\frac{u}{v})x, \quad n = 0, \pm 1, \pm 2, \cdots$$

The angle θ, the "direction" of these zero phase lines, is given by

$$\theta = \tan^{-1}(\frac{v}{u})$$

and d, the "SPATIAL PERIOD", is given by

$$d = \frac{1}{\sqrt{u^2 + v^2}} \; . \quad \text{(See Figure 4, previous page)}$$

Also, between these lines of zero phase, the phase factor varies sinusoidally, yielding a sinusoidal grating in two dimensions for each (u, v). As u and v vary, θ and d vary, so that sets of parallel lines at all possible angles and at all possible spacings cover the entire plane. It is thus often said that a Fourier Transform DECOMPOSES a function of two variables into such sine wave gratings, each grating appropriately weighted by f(x, y), the function being transformed.

Another important insight is gained from further examination of the phase term: the argument of the exponential must be dimensionless, as sines and cosines are only defined to operate on real numbers and not on seconds, meters, etc. This means that ux + vy is dimensionless, and since x and y are usually measured in units of distance, u and v must be in units of INVERSE DISTANCE or SPATIAL FREQUENCY. Similarly, a Fourier transform of a function of time would yield a function

of frequency, measured in Hertz. It is because of these relationships that Fourier space is often called Reciprocal Space, as the units of measure are the inverse of the corresponding units in Euclidean space. A particularly significant implication of this is that minute features of a function on the x-y plane, that occur within a small interval of the plane, will be spread into the extremities (i. e. , the high spatial frequencies) of the Fourier u-v plane. On the other hand, gross or overall features of the same function will be represented in the lower spatial frequencies in Fourier Space. As an example, consider a simple step function in one dimension, $f(x) = \begin{cases} 1/2\ell & \text{if } -\ell \leq x \leq \ell \\ 0 & \text{otherwise} \end{cases}$

$$F\left\{f(x)\right\} = \int_{-\infty}^{\infty} f(x)e^{-2\pi iux}\,dx = \frac{1}{2\ell} \int_{-\ell}^{\ell} e^{-2\pi iux}\,dx = \frac{-1}{4\pi iu\ell} e^{-2\pi iux}\Big]_{-\ell}^{\ell}$$

$$= \frac{-1}{4\pi iu\ell} \left[e^{-2\pi iu\ell} - e^{2\pi iu\ell}\right] = \frac{1}{2\pi u\ell} \left[\frac{e^{2\pi iu\ell} - e^{-2\pi iu\ell}}{2i}\right] ,$$

but $\dfrac{e^{ix} - e^{-ix}}{2i} = \sin x$, so we have:

$$F\left\{f(x)\right\} = \frac{\sin(2\pi u\ell)}{2\pi u\ell} .$$

Both the $\frac{1}{2\pi u\ell}$ envelope and the frequency of the sine function vary with ℓ, so that large ℓ will imply a steeply decreasing envelope and a large number of oscillations within a small interval of u centered around u = 0. Thus, most of the area under the curve will be in the low spatial frequencies of the Fourier variable u. A small ℓ will bring about the opposite effect, with a

slowly decreasing envelope and oscillations of significant amplitude

spread out to high spatial frequencies. The above arguments are

illustrated in Figure 5. Thus, a particular region of the Fourier u-v

plane does not correspond to a particular region in the x-y plane, but

rather to a particular set of frequency characteristics of the function

$f(x, y)$. This can be understood by the fact that at each point (u, v) on

the Fourier plane, the precise height of the Fourier representation

$g(u, v)$ is a result of adding contributions from all points in the x-y

plane. This situation involves extensive multiplicity, as $f(x, y)$ is

represented at each point (u, v). Likewise, each point on the x-y plane

exerts its influence over the entire u-v plane. This kind of multiplicity,

then, is a key ingredient to the Fourier transform.

Fig. 5

DBL-728-5428
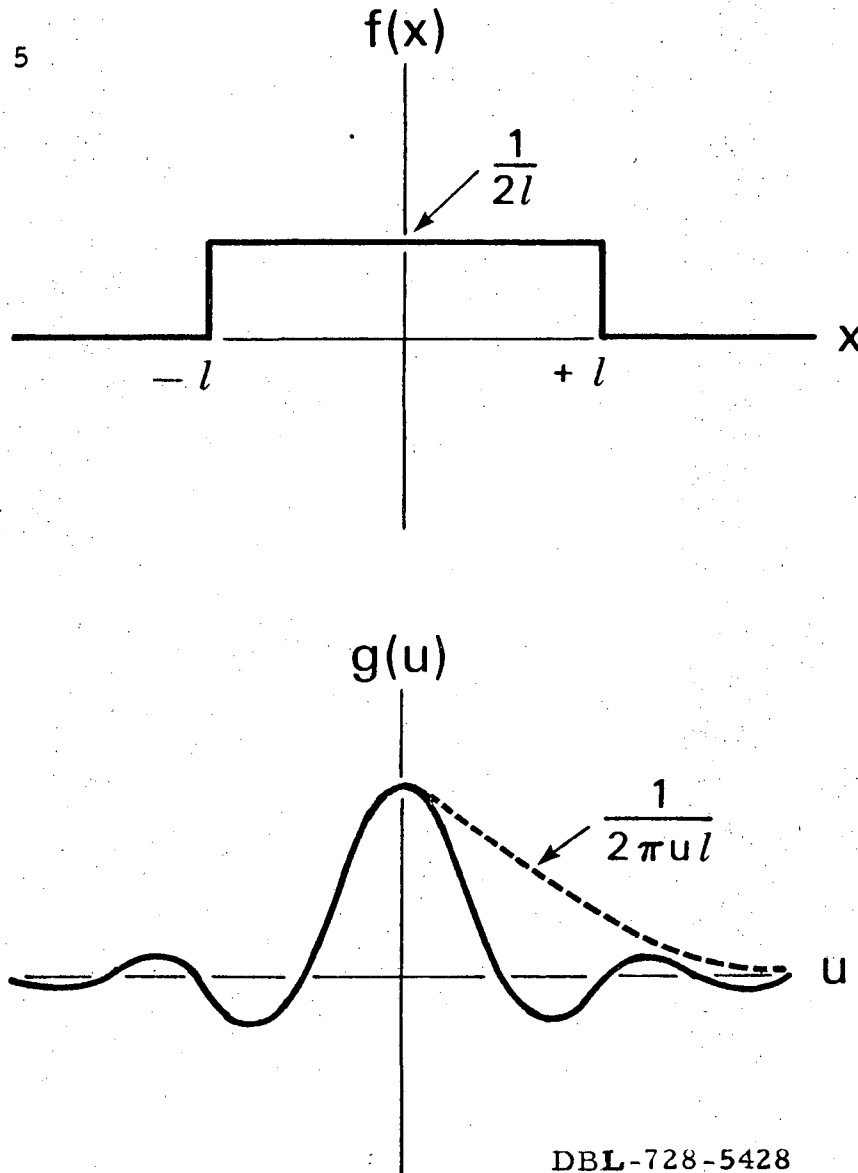
FURTHER INTERPRETATION AND APPLICATIONS
OF FOURIER TRANSFORMS

The process of Fourier Transforming a function is analogous

to translating a story from English to Spanish, assuming that a

translation could be done so that the story in each language is

exactly identical. This assumption is further based on the

premise that there exists a universal set of concepts that are

precisely equivalent, regardless of the language in which they

are expressed. A few important notions can be grasped via this

analogy: The equality of the length or total content of the story

in each language is analogous to the integrated area under

$|f(x)|^2$ being equal to $\int_{-\infty}^{\infty} |F(u)|^2 du$, which can be thought of as

the conservation of "total energy" of the function (story);

the fact that the stories in each language are equivalent raises

the question of which is the "real" story, corresponding to the

fact that Euclidean and Fourier space representations of a given

function are completely equivalent; the fact that there are many

other languages into which the story can be translated corresponds

to the existence of many different orthogonal basis sets for the

space of all square integrable (L_2) functions, of which $e^{2\pi inx}$

is only one; two stories which are very similar in one language

are very similar in another language, corresponding to the fact

that if $f(x)$ and $g(x)$ are very "similar" (i.e., their shapes and

scale are very similar), their Fourier transforms will also be

very similar; the fact that the words of the story in English

may bear little resemblance to the corresponding words in

Spanish parallels the fact that functions and their Fourier representations rarely show any resemblance to one another when viewed graphically.

An important and useful mathematical concept that fits in well with these others is the <u>convolution</u> of one function with another, sometimes referred to as the <u>cross-correlation</u> between two functions. A precise definition of "convolution" will be given shortly, but first some brief background. Let $f(x)$ and $g(x)$ be L_2 functions such that

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = c \int_{-\infty}^{\infty} |g(x)|^2 dx,$$ where c is a constant. That is, both

f and g cover a finite amount of area, but not necessarily the same. It is stipulated that $\int_{-\infty}^{\infty} |f(x)|^2 dx = c \int_{-\infty}^{\infty} |g(x)|^2 dx$. We shall show that this relation implies a precise analogy in Fourier space, namely

$$\int_{-\infty}^{\infty} |F(u)|^2 du = c \int_{-\infty}^{\infty} |G(u)|^2 du,$$

where the same c applies, due to Parseval's theorem. That theorem states that

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = \int_{-\infty}^{\infty} |F(u)|^2 du, \cdots \qquad (8)$$

and expresses the conservation of energy (area bounded by f on the x axis or F on the u axis) spoken of earlier. Now

$$<f(x),\ f(x)> = ||f(x)||^2 = \int_{-\infty}^{\infty} |f(x)|^2 dx,$$ and similarly

for $g(x)$, $F(u)$ and $G(u)$, so Parseval's theorem says

$||f(x)|| = ||F(u)||$ and $||g(x)|| = ||G(u)||$; therefore, in

substituting for the original stipulation,

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = c \int_{-\infty}^{\infty} |g(x)|^2 dx, \quad \text{we now have}$$

$$||f(x)|| = c||g(x)||,$$

and thus, $||f(x)|| = ||F(u)|| = c||g(x)|| = c||G(u)||.$

Therefore, $||F(u)|| = c||G(u)||.$

The convolution of $f(x)$ with $g(x)$, denoted $h(x) = f(x) * g(x)$ is defined as

$$h(x) = \int_{-\infty}^{\infty} f(u) g(x-u) du \quad \text{(u here is a} \qquad (9)$$

dummy variable, <u>not</u> a Fourier variable).

The significance of $h(x)$ is that it is a quantitative measure of how similar $f(x)$ and $g(x)$ are. This will be shown as follows: It was argued earlier that due to the equivalence of Euclidean and Fourier representations, similarity between functions in one space implies similarity between their representations in the other space. If it can be shown that the <u>Fourier representation</u> of $h(x)$ is a measure of the similarity between $F\left\{f(x)\right\}$ and $F\left\{g(x)\right\}$, then it would follow that $h(x)$ is a measure of similarity between $f(x)$ and $g(x)$. First, we need some important relations: The magnitude of a function $f(x)$ is given by $||f(x)|| = (<f(x), f(x)>)^{1/2}$, and expresses the amount of area bounded by $f(x)$. If $f(x)$ and $g(x)$ are two functions, we have the Schwartz Inequality,

$$<f(x), g(x)> \leq ||f(x)|| \cdot ||g(x)||. \quad \text{(13, Chapter 1)}$$

The equality holds when $f(x) = g(x)$, in which case

$$<f(x), f(x)> = ||f(x)|| \cdot ||f(x)|| = <f(x), f(x)>^{1/2} \cdot <f(x), f(x)>^{1/2},$$

from the definition of $||f(x)||$ above. Now, let's look at the

Fourier transform of h(x):

$$F\{h(x)\} = \int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} f(u)\,g(x-u)du\right\} e^{-2\pi isx}dx = F(s)$$

let $v = x-u$, so $dv = dx$ and $x = u + v$; Here, s is the Fourier variable, and u and v are merely dummy variables. Then, interchanging the order of integration yields:

$$F\{h(x)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(u)g(v)\,e^{-2\pi is(u+v)}dudv$$

$$= \int_{-\infty}^{\infty} f(u)\,e^{-2\pi isu}du \int_{-\infty}^{\infty} g(v)e^{-2\pi isv}dv$$

$$= F[f(x)] \cdot F[g(x)] = H(s) = F(s) \cdot G(s) \cdots \qquad (10)$$

So the Fourier transform of the convolution of f with g is the

product of the separate transforms of f and g. Thus, the

Fourier representation of h(x)   will have a large magnitude,

i.e.,

$$||H(s)|| = <H(s), H(s)>^{1/2}$$

will be **large** when $<F(s)\cdot G(s), F(s)\cdot G(s)>^{1/2} = ||F(s)G(s)||$

is large.

But by the Schwartz inequality (13),

$$||F(s)G(s)|| \le ||F(s)||\ ||G(s)||,$$

we see that $||F(s)G(s)||$ is maximized when $F(s) = G(s)$

(since we only stipulated that $||F(s)|| = c||G(s)||$, when $F(s) = c G(s)$, maximum will be reached). Short of maximum, we see that the greater the similarity between $F(s)$ and $G(s)$, the greater will be $||H(s)||$, so it can be said that $||H(s)||$ is a measure of the similarity of $F(s)$ and $G(s)$. Since $H(s)$, $F(s)$, and $G(s)$ are equivalent representations of $h(x)$, $f(x)$ and $g(x)$ respectively, relationships among each set of three must be equivalent. Therefore, the fact that $||H(s)||$ is large, when $F(s)$ and $G(s)$ are very similar, corresponds to the fact that $||h(x)||$ is large when $f(x)$ and $g(x)$ are similar. Thus, the convolution is a convenient device for comparing one function with another. The mechanics of the convolution can be viewed as follows: Graphically the convolution can be seen as holding $f(u)$ in a fixed position relative to the origin on the u axis, and sliding $g(-u)$ along the u axis; at each $g(-u)$ position a new function is formed by the product of $g(-u)$ and $f(u)$, and the integrated area bounded by the new function is calculated. That is, for each x,

$$\int_{-\infty}^{\infty} f(u) g(x-u) du$$

is calculated, and the magnitude of the convolution, $||h(x)||$, is the total area accumulated as $g(x-u)$ is shifted over the entire u axis by varying x from $-\infty$ to $+\infty$.
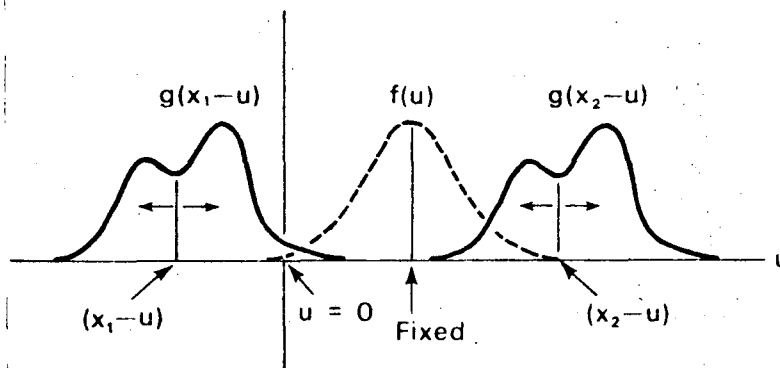
Fig. 6

Convolution is particularly useful in the pattern recognition process. By pattern recognition, it is meant that a system trying to identify a given "input function" or "message" does so by comparing it with a library of previously identified functions, to find that function to which the input is most similar. One way of making a systematic comparison of this kind is by convoluting the input function with library functions, and identifying the input as most like the library function with which $||h(x)||$ is maximized. However, such a process in Euclidean space requires computation of a difficult integral, $\int_{-\infty}^{\infty} f(u)g(x-u)du$, while in Fourier space the convolution process consists of a simple multiplication of two functions. An intrinsic advantage of Fourier representation over Euclidean representation of functions is thus seen, and this has constituted one of the prime reasons why students of the human visual system have sought evidence of a Fourier transform of optical information as it passes from the retina to the brain—were the retinal information coded in the brain as a Fourier representation of the retinal image, a method of pattern recognition involving convolutions would be vastly simplified.

Let us return, for a moment, to the discrete Fourier Transform. It was shown earlier that high spatial frequencies correspond to minute details of a function, while low spatial frequencies represent general features of the function as a whole. In a Fourier expansion of a function, it is necessary to add up the contributions of all terms to represent the given function in all

its detail. However, the elimination of the highest frequency
terms may be advantageous in certain instances where precise
details are not needed, as an infinite sum is simply not calculable.
It would be expected in such a curtailment that the overall fea-
tures and a lot of detail of the given function would still be repre-
sented, and one could design the extent of curtailment around the
desired amount of detail, using the appropriate equations. The
point is that a pattern recognition system may not need all of the
details of a given input function to distinguish it from all other
functions, and thus in Fourier space it would need only a certain
number of terms (called "harmonics"). A question which arises
is, what is the method of curtailment? Multiplication of the array
of Fourier terms by different kinds of step functions is one possible
method where the terms are multiplied by 1 if they are within a certain
region of the array, and by zero if they lie outside the region.
In general, curtailing or modifying functions are known as filters, and
they need not be only step functions. To see what happens to
the Fourier representation of a function as a filter is applied, it
is most convenient to use the continuous transform. If
$F[g(x)] = G(u)$ is the Fourier transform of $g(x)$, and $P(u)$ is a
filter function by which $G(u)$ is multiplied, we see the following
interesting relationships: $P(u)$ must be the Fourier transform of
some function of x, say $p(x)$, (because there is a one-to-one cor-
respondence between functions in Fourier space and functions in
Euclidean space), so the product $G(u) P(u)$ is a product of the
Fourier transforms of $g(x)$ and $p(x)$ respectively. From previous

arguments, this is the same as the Fourier transform of the con-
volution of g(x) and p(x), $F\left\{g(x) * p(x)\right\}$, where p(x) is the inverse
transform of the filter function P(u). If one were now to inverse
transform the filtered Fourier representation (i. e., to attempt to
reconstruct g(x)), the result would be

$$F^{-1}[\ F\left\{g(x) * p(x)\right\}] \ = g(x) * p(x) \cdots \ , \qquad (11)$$

a convolution of g(x) with the Euclidean equivalent of P(u). Thus,
in two dimensions, the imposition of a square filter on a Fourier
transform of g(x, y) will result in a Fourier representation of
g(x, y) * (convoluted with) $F^{-1}\{\square\}$

$$\text{which is} \quad g(x, y) * \left[\frac{\sin(2\pi \ell x)\sin(2\pi\ell y)}{\pi^2 \ell^2_{xy}}\right] .$$

Such a convolution in real space will give rise to distorting
oscillations in g(x, y), whose amplitude will depend on ℓ, the
length of a side of the square filter.

The Projection Theorem and Strip Integration

Let $f(x, y)$ be a function of two variables, x and y. The portion of $f(x, y)$ which lies above a line $x = k_1$ (constant) is a function of y and the constant, k_1, $f(k_1, y)$. The integral of $f(k_1, y)$ over all y is called a "line integral" on the line $x = k_1$, and results in the area under $f(k_1, y)$ (a number) being stored in the location $(k_1, 0)$.

If such integrals are computed at all values of x, and a corresponding number recorded at each x, the result would be a function of x,

$$f_L(x) = \int_{-\infty}^{\infty} f(x, y)dy \cdots \tag{12}$$

$f_L(x)$ is the projection of $f(x, y)$ onto the x-axis, which is the line $y = 0$. If we take the Fourier transform of $f_L(x)$ we have

$$F\left\{f_L(x)\right\} = \int_{-\infty}^{\infty} f_L(x)\, e^{-2\pi i u x}dx = \int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} f(x, y)dy\right\} e^{-2\pi i u x}dx$$

and assuming order of integration can be reversed,

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, e^{-2\pi i u x}\, dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, e^{-2\pi i(ux+oy)}\, dxdy = G(u, 0),$$

the portion of the <u>two</u> dimensional Fourier transform of $f(x, y)$ which lies above the line $v=0$ in Fourier space. Were we to choose another line [See illustration in fig. 7, next page].

f(x,y)

x = k₂

x = k₁

y

x
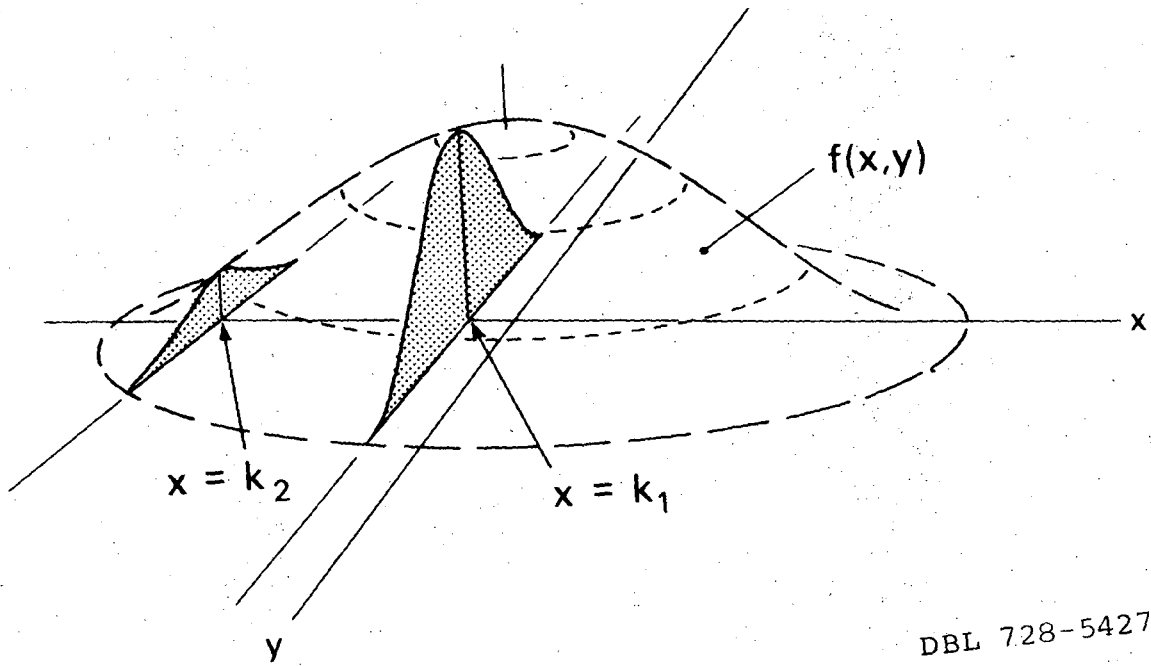
DBL 728-5427

Fig. 7

through the x-y origin, say y = ax, on which to project $f(x, y)$, the one-dimensional transform of that projection would yield that portion of the Fourier representation of $f(x, y)$ which lies above the line v = au in Fourier space, where "a" is the same constant in each case. One can see that the full two-dimensional Fourier representation of $f(x, y)$ could be assembled by superimposing the one-dimensional transforms of projections in the x-y plane, provided the projections are onto all possible lines through the x-y origin. This is a special case of the <u>Fourier Projection Theorem</u>, which asserts that the full Fourier transform of a function of n dimensions can be completely determined by appropriate combination of (n-1) dimensional transforms of projections of the original function. Such a technique is of great importance in the fields of radio astronomy and crystallography, and in information processing in general. However, in all physically realizable applications of the theorem, an important approximation must be used: the line samples, $f(k, y)$ from which the projections are constructed, are not physically realizable, due to their infinitesimal width, and must be approximated by thin "strip" samples of $f(x, y)$. Such projections are called "strip integrals." The use of finite samples introduces error into the mathematically precise technique outlined above, but offers computational compensation (aside from the fact that it is physically unavoidable): the two-dimensional transform of $f(x, y)$ can be approximated to any desired degree of resolution,

provided the function $f(x, y)$ is identically zero outside a finite cirlce of radius X. If this condition is met, then $F\{f(x,y)\}$ can be determined to within resolution β by one-dimensional transforms of strip integrals taken at equal angles, $\frac{2\pi}{n}$, through the x-y origin, where $n \geq X/2\beta$. (Reference 3)

COMPUTER FOURIER TRANSFORMS

The analytical or integral Fourier transform is directly

calculable for only a very restricted set of functions—those for

which the difficult integrals involved are solvable in closed form.

However, a complicated function can be approximated by an array

of sampled points from the function, and there are computer

algorithms to calculate the discrete Fourier transform of such

an array. One such algorithm, known as the Fast Fourier

Transform Algorithm, makes use of the shift theorem and the

linearity of the transform to significantly reduce the computational

time necessary to perform a given transform. That is, using

the definition of the discrete Fourier transform, an array of N

data points requires N^2 computational operations, employing

the formula:

$$c_n = \sum_{k=1}^{N} f(k) e^{-2\pi i k n} \quad \ldots \ldots \ldots \ldots \quad (12)$$

However, the sampled points can be divided into two parts, as

follows: let $f_1, f_2, \cdots, f_{64}$ be 64 sampled points from the desired

function f(x), at equal intervals of x, and divide the set of points

into two subsets, $\{f_1, f_3, \cdots, f_{63}\}$ and $\{f_2, f_4, \cdots, f_{64}\}$. Now by

the linearity of the Fourier transform,

$$F\{f_1, f_2, \cdots, f_{64}\} = F\{f_1, f_3, \cdots, f_{63}\} + F\{f_2, f_4, \cdots, f_{64}\},$$

and by the shift theorem, each $f_2, f_4, f_6, \cdots$ etc. can be shifted

to the left on the x axis so that it corresponds to the same x value

as f_1, f_3, f_5, etc., respectively. The transform of such a shifted

array would be the same as the unshifted array, multiplied by the phase factor $e^{+2\pi i n}$, as the shift for each element is one unit in the negative x direction, so

$$F\{f_1, f_2, \cdots, f_{64}\} = F\{f_1, f_3, \cdots, f_{63}\} + F\{f_2^1, f_4^3, \cdots, f_{64}^{63}\} e^{2\pi i n},$$

where f_k^{k-1} means, "the sampled <u>value</u> at the $k\underline{\text{th}}$ location on the x axis, shifted to the $(k-1)\underline{\text{th}}$ position on the x axis." Now, were the algorithm simply to use the formula (12) at this time, the time necessary to compute the transform would be about half the original required time, as each of the transforms on the right side would take 32^2 operations to perform. Thus, $64^2 = 2(32^2 + 32^2)$, and not counting the relatively few steps needed to calculate the exponential shift term, speed of calculation has been roughly doubled. Further binary divisions can be made, each time reducing the previously needed computation time by a factor of two, until the transforms of single points are left. The final computation of these single point transforms and the corresponding phase terms requires a total of $N\log_2 N$ computational operations, as opposed to N^2 for the straightforward method of formula (12), which for N = 64 is 384 as opposed to 4,096, operations. It should be noted that since binary divisions are the basis for the reduction in computation time, it is necessary that the input array of sampled points have a total of 2^M points, where M = 0, 1, 2, 3, $\cdots$. Other fast Fourier algorithms could be based on ternary, quaternary or other divisions of the original data set, requiring the input to have 3^M, 4^M or n^M, points, respectively.

MODEL OF PROPOSED FOURIER TRANSFORM

One mathematical appeal of the Fourier transform as an element in the visual process was shown to be in the shift theorem, which allows for invariance of functional shape with linear translation in Euclidean space. The phase factor, $e^{-2\pi i(au+bv)}$, represents the relative position of the input function, and thus between the structural representation $(g(u,v))$ and the phase term, the Fourier transform of $f(x+a, y+b)$ completely describes both what and where $f(x,y)$ is. In addition to the attractive property of positional invariance, the Fourier transform was shown to greatly simplify the process of convolution or correlation between functions. Impressive evidence has been accumulated to suggest that cross-correlations do occur in the visual process.[8]

From an anatomical-physiological approach, certain elements of the visual pathway are highly suggestive of an information processing method that incorporates a vast amount of informational redundancy. For example, Hubel and Wiesel have shown that the striate cortex of monkeys primarily contains columns of cells whose receptive fields overlap to the extent that, within any column, a large section of the total visual field is represented many times over.[10] As was stated earlier, this kind of multiplicity is essential to the Fourier Transform. Psychophysical experiments on visual detection of various types of gratings (square wave, sine wave, etc.) show that over a wide range of spacial frequencies, detection is determined only by the amplitude of the fundamental Fourier component of their wave forms.[6]

These findings and others have contributed to the promotion of the hypothesis that somewhere in the visual pathway, a Fourier transform of the information takes place. In lieu of a direct anatomical method for determining the nature of visual information processing, which is presently beyond reach due to the phenomenal complexity of neural networks in the visual pathway, indirect clues such as these are often combined with what may be termed, "seductive reasoning," to reach conclusions that <u>may</u> be premature. Given the seemingly hopeless complexity of the overall problem of visual information processing, and the rather attractive nature of the above-cited data, it is understandable that a great deal of enthusiasm has been generated by a fairly comprehensive model, involving the Fourier transform, that has recently been proposed by Pollen and Lee.[11] As one of the major goals of this thesis has been to test the validity of their model, it is appropriate that the essential elements of their proposal be outlined here.

Pollen and Lee assert that the output from the lateral geniculate into the striate cortex is essentially a topographical (i. e. , one-to-one) representation of the retinal image, (which is a two-dimensional intensity distribution, $f(x, y)$), and that transformation of visual information takes place primarily in the cortex and higher centers of the brain. They propose that the information coded in the "complex cells," noted by Hubel and Wiesel, is a partial, if not complete, Fourier transform of the information coded on the retina by the focused image. (They assert that the transform is

completed by the time the information reaches the higher centers
of the brain). This, they claim, is achieved via Fourier trans-
formation of strip integrals of the focused image information,
which are gotten, in turn, by combinations of signals from simi-
larly-oriented "simple" cells. It will be recalled that "simple"
cells have long, narrow receptive fields—Pollen and Lee cite
data to show that the firing rate of such cells is a monotonically
increasing function of the total intensity illuminating their receptive
fields. They assert that as such, each cell's output is equivalent
to an element of the projection of $f(x, y)$ onto a line-perpendicular
to the receptive field. Combinations of outputs from many
similarly-oriented cells would thus constitute the projection or
strip integral of $f(x, y)$ onto a line perpendicular to the receptive
fields. According to Pollen and Lee, such strip integration
takes place, along with at least the start of a Fourier transform
of the strip integrals, between the simple and complex cell stages
of the cortex.

The precise method by which the Fourier transform of a
given strip integral is performed is not elaborated by Pollen and
Lee, but they do offer some suggestions regarding the important
"ingredients" of the transform. Phase information, which is a
method of coding position of a function relative to an origin, could
be coded by delay of the response of a given complex cell to
stimuli, depending upon the position of the stimulus in the
receptive field of the cell. Indeed, Pollen and Lee cite data
to support such a coding of phase information, provided that the

visual process SCANS the incoming information at discrete intervals, as opposed to continuously processing the information. They proceed to show that a given complex cell demonstrates superposition of responses and interference, when presented with multiple stimuli within its receptive field, supporting their claim that response latency could map phase (positional) information. Spatial frequency information, they claim, is coded in the individual complex cells, as each complex cell has been found to respond maximally to a slit of a specific width (and orientation).

It should be noted that their model seeks only to account for stationary visual phenomena, not those associated with motion (i. e. , time-varying stimuli), and they argue convincingly that a firm understanding of such time-independent processes is a necessary prerequisite to dealing with moving-or otherwise time-dependent-stimuli.

OPTICAL ILLUSIONS AND MODELS

In addition to their intellectually fascinating and aesthetically enjoyable qualities, optical illusions can be valuable scientific tools, for they can be approached as true shortcomings of the visual system. That is, their varied origins notwithstanding, all illusions constitute contradictions between what is perceived and what is known through experience to be the case, and thus represent improper recognition by the optical system. In order to learn how a complex system operates, it is sometimes advantageous to examine how it fails — such was the case in the study of genetics, which has been done essentially entirely via mutations, which constitute failures of the reproductive system. It is hoped by many students of the visual process that a similar approach will prove fruitful in elucidating the methods of optical information processing, using illusions.

One important class of illusions pertains to the geometrical illusions of size and shape. They can be generally characterized as a distortion in shape or size of an otherwise non-distorted figure, brought about by placing it in a peculiar background. Apparently, there is some interaction between figure and background during the processing of optical information, which gives rise to misinterpretation of the visual scene. Several examples of well known geometrical illusions are illustrated on the following 2 pages. [11] They are divided into two basic groups: illusions of shape and illusions of size. The two are certainly related, as each involves apparent distortion of the relative positions of

DISTORTIONS OF SIZE

Some examples that might be considered are:

The Müller-Lyer illusion
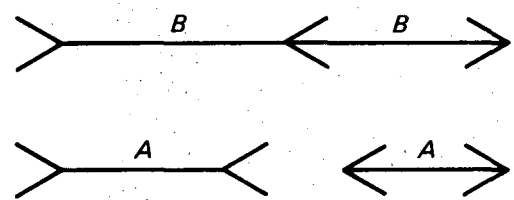
The B's are of equal length, as are the A's

Angles affect the apparent length of lines
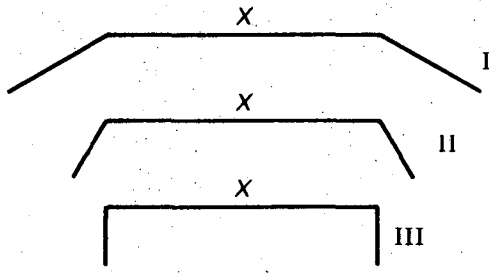
X has the same length in all three cases

Two equal diagonals which appear unequal
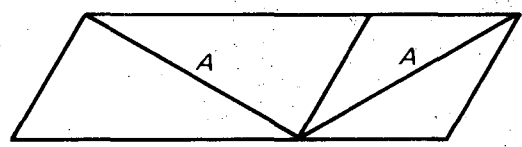
An illusion of contrast
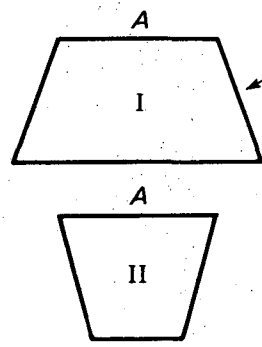
The A's are of equal length

Contrasting angles
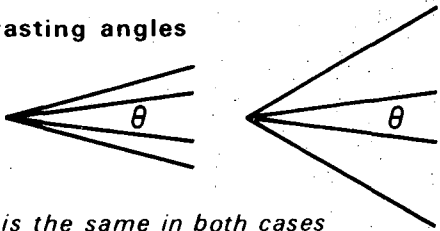
θ is the same in both cases

Fig. 1

DISTORTION OF SHAPE
Some examples that might be considered are:

Parallel lines which do not appear so
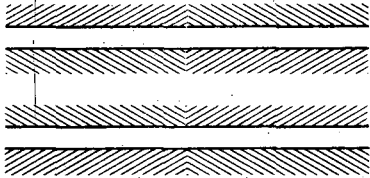
Wundt's illusion of direction
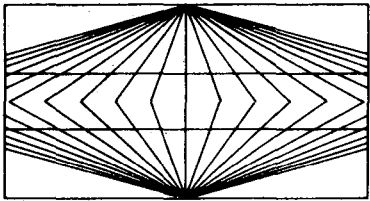
Parallel lines are distorted

Hering's illusion of direction
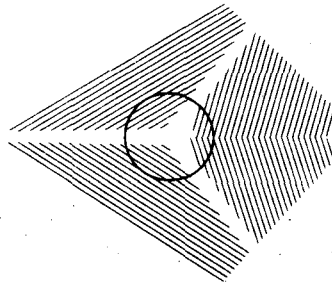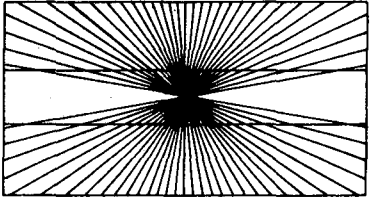
Distortion of a circle due to superposed lines
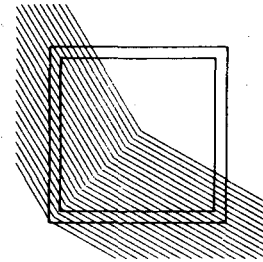
Distortion of a square due to superposed lines

"Twisted cord" illusions

See following two sheets.

These are concentric circles   *These are straight cords*

DBL 728-5419
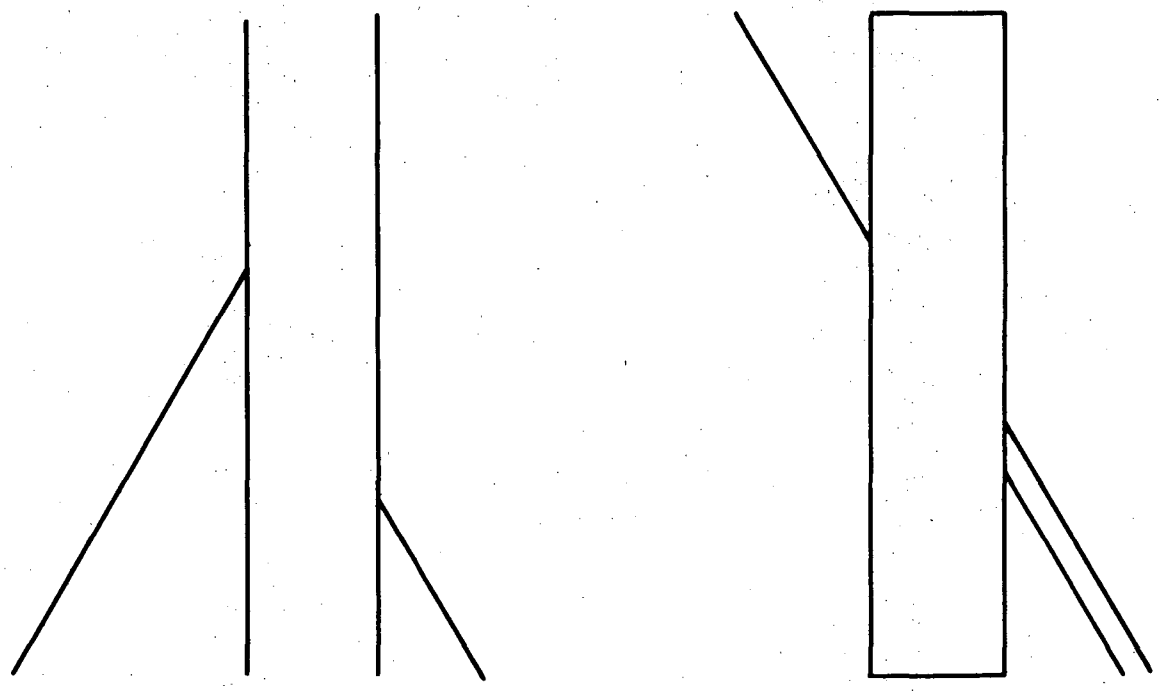
Fig. 2

DBL 728 5429

Fig. 2a

constituent parts of the illusory figure. The basic distinction
between them, however, is that a size illusion preserves the
overall shape of the illusory figure, whereas a shape illusion
preserves the overall size thereof. Both types, however, are
sufficiently common among humans that they cannot be dismissed
as mere statistical deviation from the normal visual process.
This is not to say, however, that such illusions must be ascrib-
able to only the visual pathway, as it is certainly conceivable
that memory, other sensory information, learned responses or
inhibitions, or other factors could influence the illusory effect
in any given individual.

Indeed, wide variation in response to illusory figures has been
noted among people of widely differing cultural backgrounds. Such
findings, however, do not preclude the possibility of a fundamentally
physiological origin of the geometrical illusions, if illusions are
understood as "causes" or "stimuli", rather than as specific effects.
That is, illusory figures constitute unusual stimuli which may be
interpreted and responded to by the viewer in more ways than one.
Their unusual quality is that they have a high probability of eliciting
responses that do not correspond to commonly accepted reality; but
such responses are not absolutely certain consequences of viewing
an illusory figure.

One way of testing the validity of a proposed model of the
visual system is to use common illusory figures as inputs, and

see if the output of the model shows a response to the illusory

effect. In the case of the Müller-Lyer illusion, for example,

the model should interpret the open-ended figure as having a

longer center line than that of the corresponding closed-ended

figure. An affirmative result of such a test constitutes one im-

portant piece of evidence in a long list of anatomical, physio-

logical and behavioral qualifications for a fitting model, the con-

ditions under which the logical property of SUFFICIENCY is

established. That is, a model which manifests all the known

characteristics of the visual system may be regarded as satisfying

the requirements for "candidacy", though its uniqueness has yet

to be established through the logical exclusion of all other models.

However, it is very doubtful whether, in a complex and empirical

study such as that of the visual process, all conceivable alternatives

to a proposed model could even be listed, let alone logically dis-

proved, so the question of logical uniqueness is somewhat restricted.

In most empirical science, "unique" refers to that single

model which has withstood, and continues to withstand, the

passage of time and further experimentation. Exclusion of rival

models happens by virtue of new evidence that such rivals fail

to predict, and it is in this sense that the passage of time selects

out the "unique" model. In the case of models of the visual

system, however, the field is so relatively young and the number

of proposed models so vast that it is doubtful that even the re-

stricted sense of uniqueness just described will be established in

the near future (as of 1972). Most of the present efforts are

toward the goal of establishing sufficiency, and even this is formidable in view of the lack of precise knowledge on much of the physiology and anatomy of the visual pathway.

The hypothesis that Fourier transforms are involved in the visual process, discussed in the previous section, is an example of a proposed model (or set of models) for which sufficiency has yet to be fully demonstrated. As a test of this hypothesis, one could use geometrical illusions as inputs and check to see if their Fourier representations, or some physiologically reasonable modification thereof, demonstrate the illusory effects. By "physiologically reasonable" is meant a modification that could be effected with the accepted machinery of the visual pathway, and that would be a functionally useful part of a model of the visual pattern recognition process. For example, in the interpretation of visual information a process of selective attention may be employed, in which only certain portions of the information are examined at any one time. If the information were coded as a Fourier transform of the retinal information, this would mean examining only part of the transform at a particular instant, and this would constitute a modification of the transform data reaching the brain.
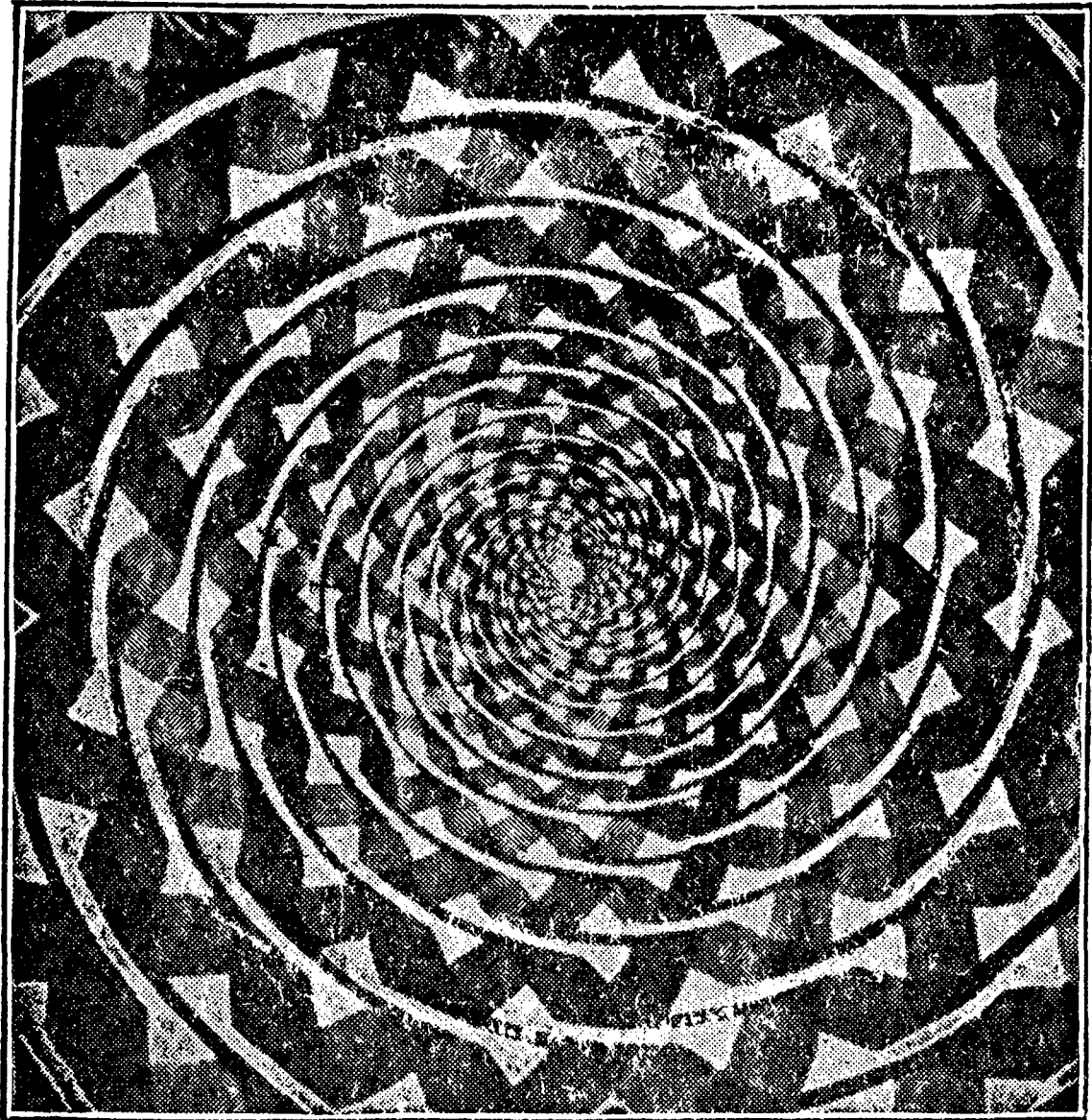
XBB 738-4851

Fig. 3 - "Twisted Cord" Illusion #1

These are concentric circles:
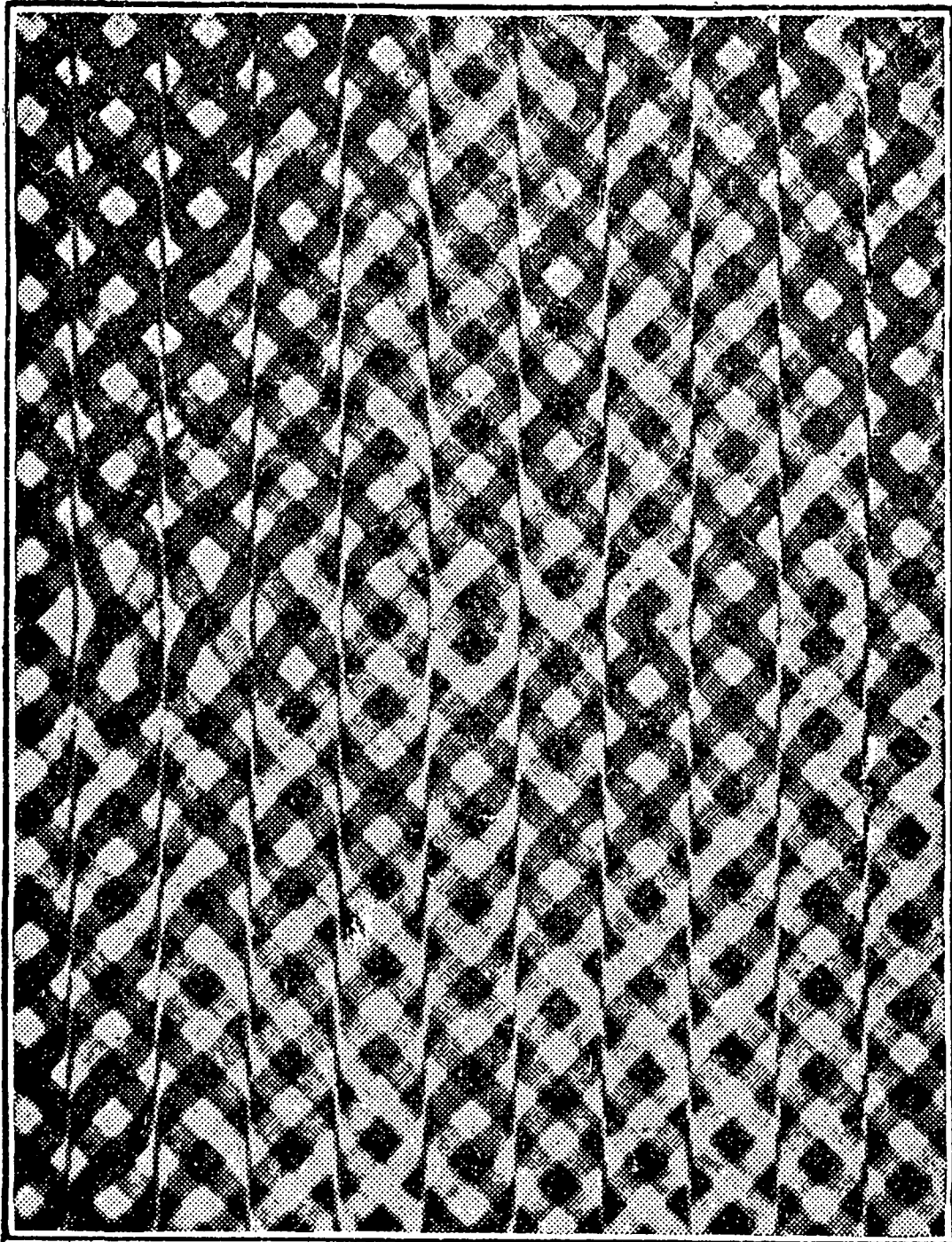
XBB 738-4852

Fig. 3a - "Twisted Cord" Illusion #2

These are straight, parallel lines:

THE MÜLLER-LYER ILLUSION: ANALYTICAL FOURIER TRANSFORM

In accordance with the reasoning outlined in the previous section, the MÜLLER-LYER illusion was selected as a convenient illusory input by which to test the sensitivity of the Fourier Transform to geometrical illusory effects. Two approaches were taken: (1) a purely analytical approach in which the illusory figure is expressed in terms of a mathematically canonical function of two variables, Fourier transformed, and comparisons made between the transforms of the two Müller-Lyer figures; (2) digitized versions of the figures are computer transformed, the transforms modified by various filters, and the resulting arrays inverse transformed to produce modified versions of the original figure. The first approach is discussed here.

As seen in Fig. 1, $\boxed{\Omega}$ represents a "primitive" form of the Müller-Lyer figures, from which they can be derived, as in $\boxed{A}$ and $\boxed{B}$. Now, the notation $\textcircled{2}$ (a, 0) refers to line $\textcircled{2}$ shifted in the positive x direction by an amount "a", and not shifted in the y direction. Using the shift theorem and other properties of the Fourier transform, we now proceed to calculate $||F(\boxed{A})||^2$, $||F(\boxed{B})||^2$, and their difference. As we are interested in the effect on line $\textcircled{3}$, particular attention will be paid to terms involving $\textcircled{3}$.

$$F\{\boxed{A}\} = F\{\textcircled{1}\}\, e^{-2\pi i(-au)} + F\{\textcircled{2}\}\, e^{-2\pi i(au)}$$

$$F + \{ ③ \} + F\{ ④ \}e^{-2\pi i(-au)} + F\{ ⑤ \}e^{-2\pi i(au)} .$$

$$F\{ \boxed{B} \} = F\{ ① \}e^{-2\pi i(au)} + F\{ ② \}e^{-2\pi i(-au)} + F\{ ③ \}$$

$$+ F\{ ④ \}e^{-2\pi i(au)} + F\{ ⑤ \}e^{-2\pi i(-au)} .$$

Let $F_1(s)$ denote $F\{ ① \}$, where s represents <u>u and v.</u>
Note that $F_1^*(s) = F_1(-s)$. Do similarly for ②, ③, ④ and
⑤ . Then,

$$|| F\{ \boxed{A} \} ||^2 = F\{ \boxed{A} \} F^*\{ \boxed{A} \}$$

$$= \left[F_1(s)F_1(-s) + F_4(s)F_4(-s) + F_1(s)F_4(-s) + F_4(s)F_1(-s) \right]$$

$$+ e^{-4\pi iau} \left[F_2(s)F_1(-s) + F_2(s)F_4(-s) + F_5(s)F_1(-s) + F_5(s)F_4(-s) \right]$$

$$+ e^{4\pi iau} \left[F_1(s)F_2(-s) + F_1(s)F_5(-s) + F_4(s)F_2(-s) + F_4(s)F_5(-s) \right]$$

$$+ F_3(s) \{ \left[F_1(-s) + F_4(-s) \right] e^{-2\pi iau}$$

$$+ \left[F_2(-s) + F_5(-s) \right] e^{2\pi iau} \} + F_3(-s) \{ \left[F_1(s) + F_4(s) \right] e^{2\pi iau}$$

$$+ \left[F_2(s) + F_5(s) \right] e^{-2\pi iau} \} + F_3(s)F_3(-s)$$

Only the terms on this page involve line ③ , so let us concern ourselves only with them.

$$|| F\{ \boxed{B} \} ||^2 = F\{ \boxed{B} \} F^*\{ \boxed{B} \}$$

$$= F_3(s) \{ \left[F_2(-s) + F_5(-s) \right] e^{-2\pi iau}$$

$$+ \left[F_1(-s) + F_4(-s) \right] e^{2\pi iau} \} + F_3(s)F_3(-s)$$

$$+F_3(-s) \left\{ \left[F_1(s) + F_4(s) \right] e^{-2\pi i a u} + \left[F_2(s) + F_5(s) \right] e^{2\pi a i u} \right\}$$

+ other terms not involving ③ .

$$\left|\left| F\{ \boxed{B} \} \right|\right|^2 - \left|\left| F\{ \boxed{A} \} \right|\right|^2$$

$$= 2i \sin(2\pi a u) \left[F_3(s) \left\{ F_1(-s) - F_2(-s) \underline{\pm} F_4(-s) - F_5(-s) \right\} \right.$$

$$\left. + F_3(-s) \left\{ F_2(s) - F_1(s) + F_5(s) - F_4(s) \right\} \right]$$

$$= 2i \sin(2\pi a u) \left[F_3(s) \left\{ F_1(-s) - F_2(-s) + F_4(-s) + F_5(-s) + F_2(s) - F_1(s) \right.\right.$$

$$\left.\left. + F_5(s) - F_4(s) \right\} \right]$$

because $F_3(s) = F_3(-s)$ (③ is centrosymmetric).

Now $F(s)$, being a complex valued function of u and v, can be represented as the sum of a real-valued function, $x(u, v)$, and an imaginary-valued function $iy(u, v)$, where $i = \sqrt{-1}$ and $y(u, v)$ is real-valued.

$$F(-s) = F^*(s) = x(u, v) - iy(u, v)$$

So,

LET

$$F_1(s) = x_1 + y_1 i$$
$$F_2(s) = x_2 + y_2 i$$
$$F_3(s) = x_3 \qquad (y_3 = 0)$$

$$F_4(s) = x_4 + y_4 i$$

$$F_5(s) = x_5 + y_5 i$$

where $x_i(u, v)$ and $y_i(u, v)$, and s, once again, represents the variables u, v.

Therefore:

$$|\,|\,F\,\{\boxed{B}\}|\,|^2 - |\,|F\{\,\boxed{A}\,\}|\,|^2$$

$$= i2\pi\sin(2\pi au)\left[x_3(x_1-y_1i-x_2+y_2i+x_4-y_4i-x_5+y_5i+x_2+y_2i\right.$$

$$\left. -x_1-y_1i+x_5+y_5i-x_4-y_4i)\right]$$

$$= -4\pi\sin(2\pi au)\left[x_3(y_2+y_5-y_1-y_4)\right]$$

$$= 4\pi\sin(2\pi au)\left[x_3(\,y_2+y_5-y_1-y_4)\right] \cdot \quad \cdots \cdots \cdots \cdots \text{Eq. 1}$$

Now, if ① and ⑤ are considered as <u>one</u> function of (x, y) (see illustration below, showing finite "thickness") that function is centrosymmetric, which implies that its transform is real, or $y_1 = - y_5$. By analogous argumentation, $y_2 = - y_4$. Thus, Eq. 1 can be written as $4\pi\sin(2\pi au)\left[x_3(2y_1 + 2y_4)\right]$

$$= 8\pi\sin(2\pi au)\left[x_3(y_1+y_4)\right].$$

NOTE: <u>All</u> of the above analysis was made without specifying the angle between ① and ③ , etc., <u>or</u> the actual functions ① and ⑤ , etc. The only restriction is that ① and ⑤ , and ② and ④ by symmetric about the origin, and that ③ also be centro- symmetric.

Fig. 2
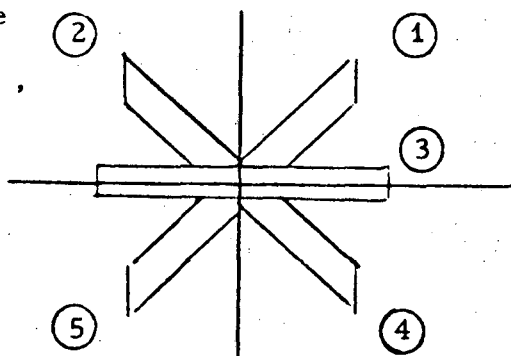
As can be seen from a comparison of figures III and VII (c) , the difference in the intensities of the two figures that is predicted by the above analysis is essentially what is seen in the computer plot of the terms of Eq. 1, shown in Figure VII. The dominant feature is the pair of spikes at the first order Fourier terms (exact location of the spikes was gotten from print-out), indicating a very significant difference in the intensities of the two figures at the very low frequencies.

It should be noted that subtraction of $F\{\longleftrightarrow\}$ from $F\{\succ\!\!\!-\!\!\!\prec\}$ without squaring the transforms, eliminates the representation of the center line, indicating that in the complete Fourier transform of each figure, the center line is identically represented. This, however, does not imply that there are no differences in the center line representation between the transforms of the two figures, as corresponding areas of the two Fourier representation may not contain identical information about the center line. This is investigated in the section on computer transforms of the Müller-Lyer illusory figures and the effects of spatial filtering them. What has been demonstrated here, however, is that from a purely mathematical standpoint, there is a difference in the Fourier intensities of the two figures, despite there seeming similarity in composition.
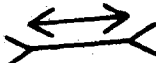
The results of the preceding calculations, and their comparison to those gained by computer techniques, are shown in the following 7 figures. The first is a graph of 4 sin(2πau), the first

factor (divided by π for illustrative convenience) in Eq. 1. Second is

a graph of x_3, the real part of the transform of the center line of the

illustory figure (i. e. , the complete transform of the line, since its

transform contains no imaginary terms); this term is the second

factor in Eq. 1. The third figure is a graph of Eq. 1 (divided by π);

note the prominent spikes at small values of u symmetric about the

origin, and their signs. Fourth is a computer-generated perspective

plot of the four arms of the canonical formulation of the illusory

figure (Figure 2, this section), prior to transformation. Fifth is a

perspective plot of the real terms in the transform of the central

line (now a function of <u>two</u> dimensions, whereas it was only single-

dimensioned in the previous graphs). Sixth is a perspective plot of

the terms $y_1 + y_4 - y_2 - y_5$, from Eq. 1, calculated via straightforward

computer transformation of the arms of the canonical figure. Last

is a perspective plot of the difference in Fourier intensities between

the two Müller-Lyer figures, calculated by computer transform.

Note the prominent spikes at low frequencies, and their signs, and

note their fundamental similarity to those of the second graph.
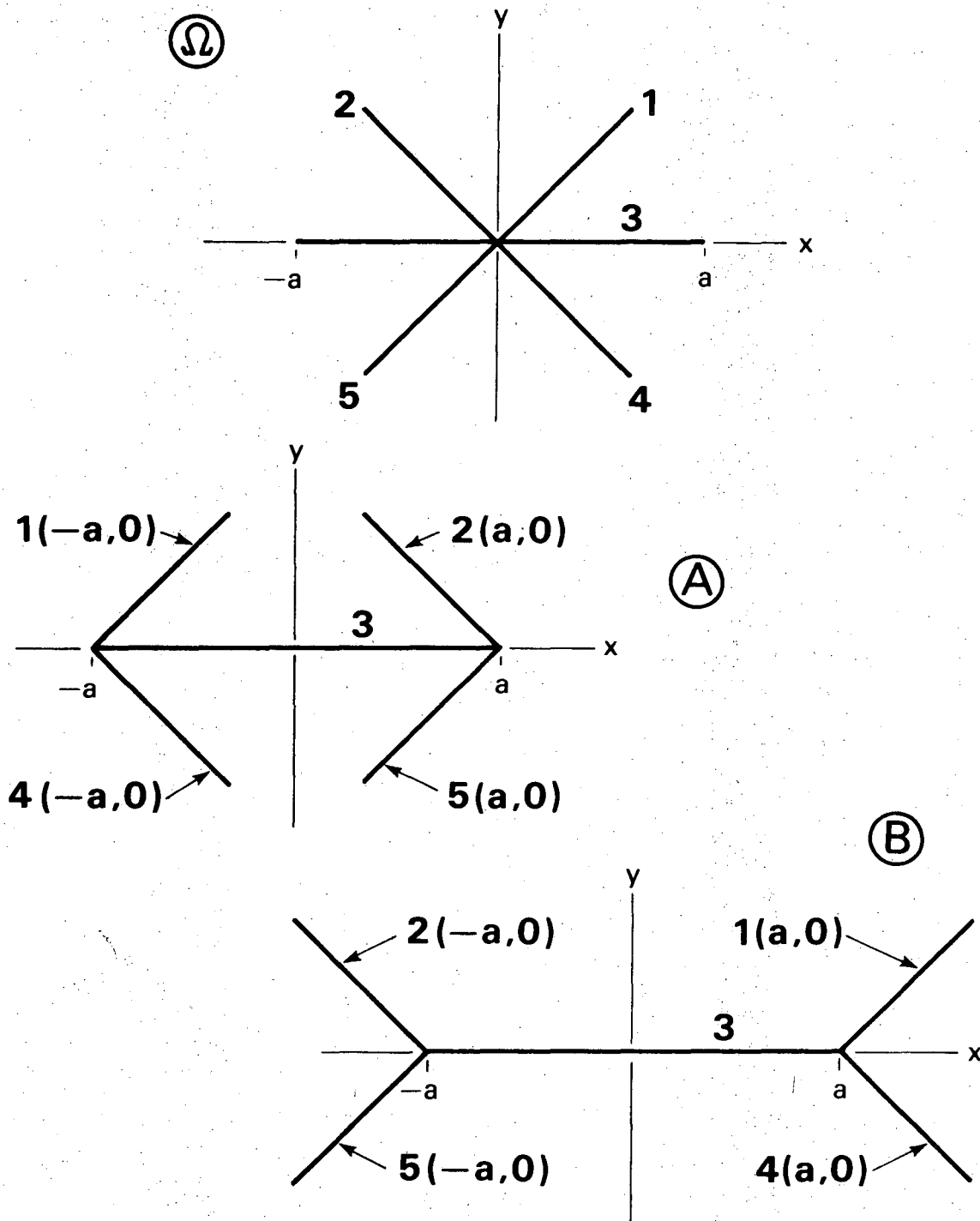
Fig. 1

DBL 728 5430

From 1-dimensional version of ⤡

$4 \sin (2 \pi a u)$
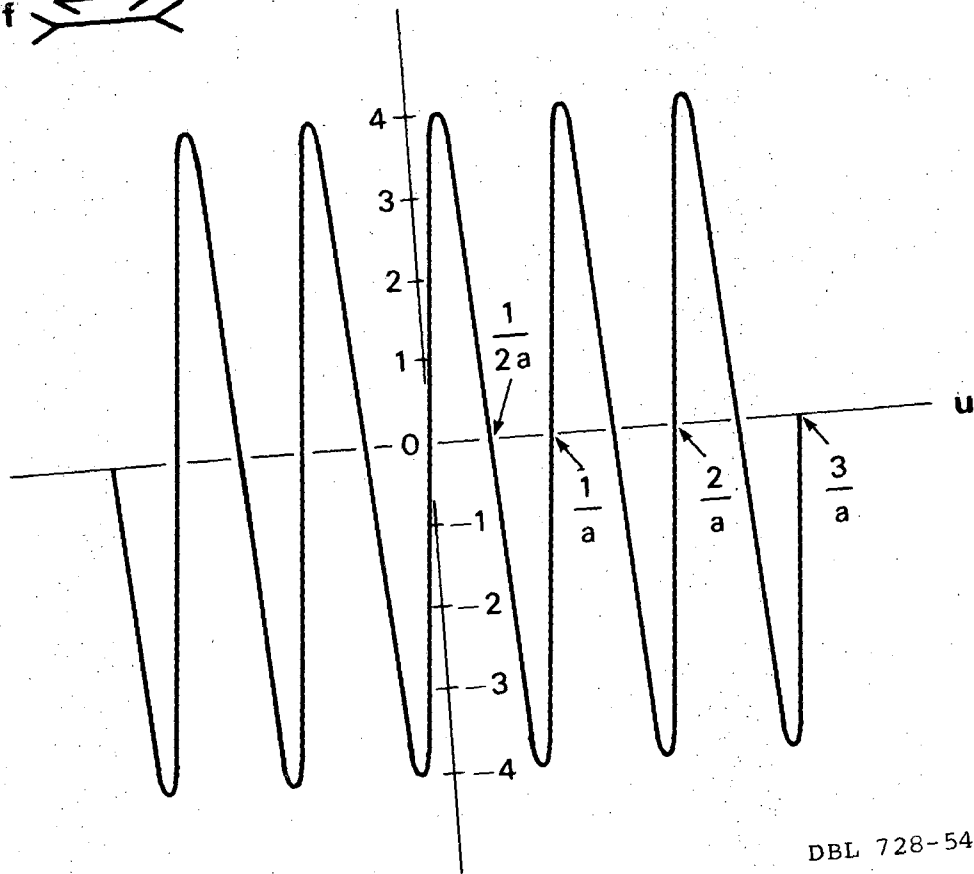
Fig. 3
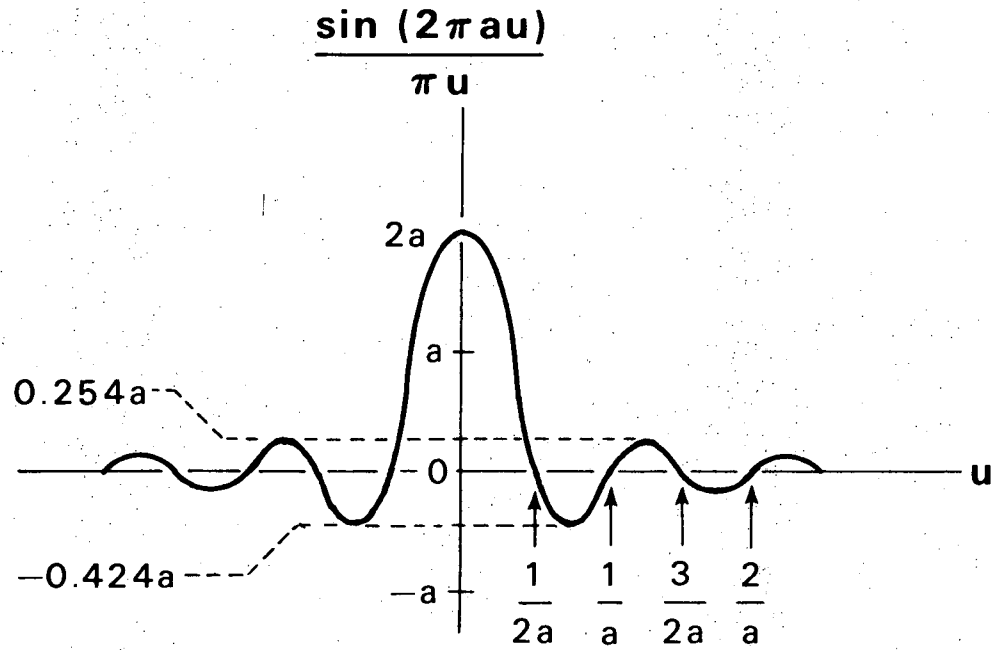
DBL 728-5432

DBL 728 5433

Fig. 4

$$4 \sin(2\pi au) \cdot \frac{\sin(2\pi au)}{\pi u} = 4 \frac{\sin^2(2\pi au)}{\pi u}$$
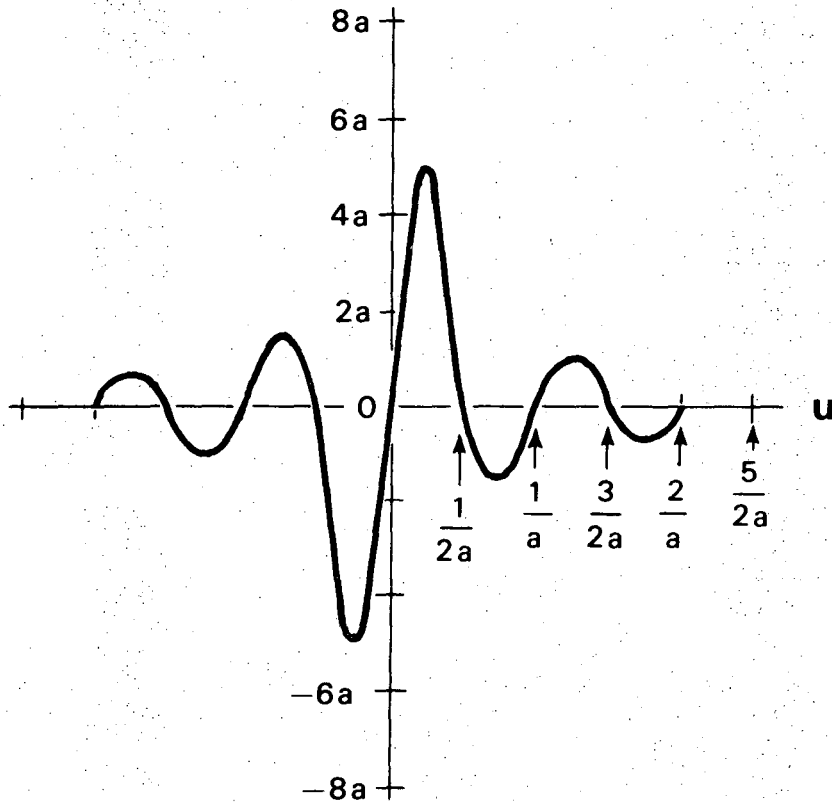
DBL 728 5431

Fig. 5

The Four Arms of the Müller-Lyer Illusion, in the x-y plane
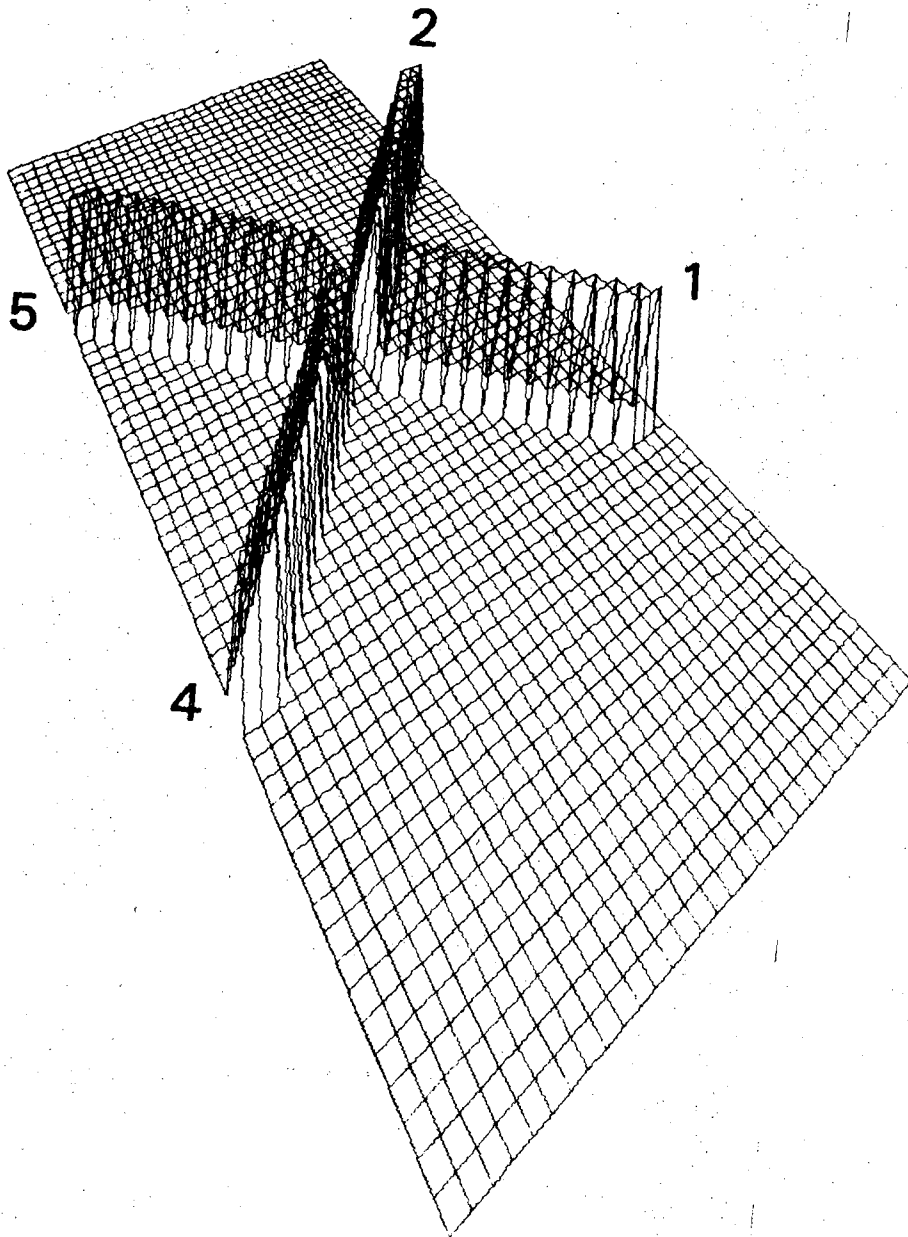
XBL 735-587

Fig. 6

$$\mathrm{Re}\left\{\mathscr{F}\xi - \xi\right\}$$
$$= x_3$$
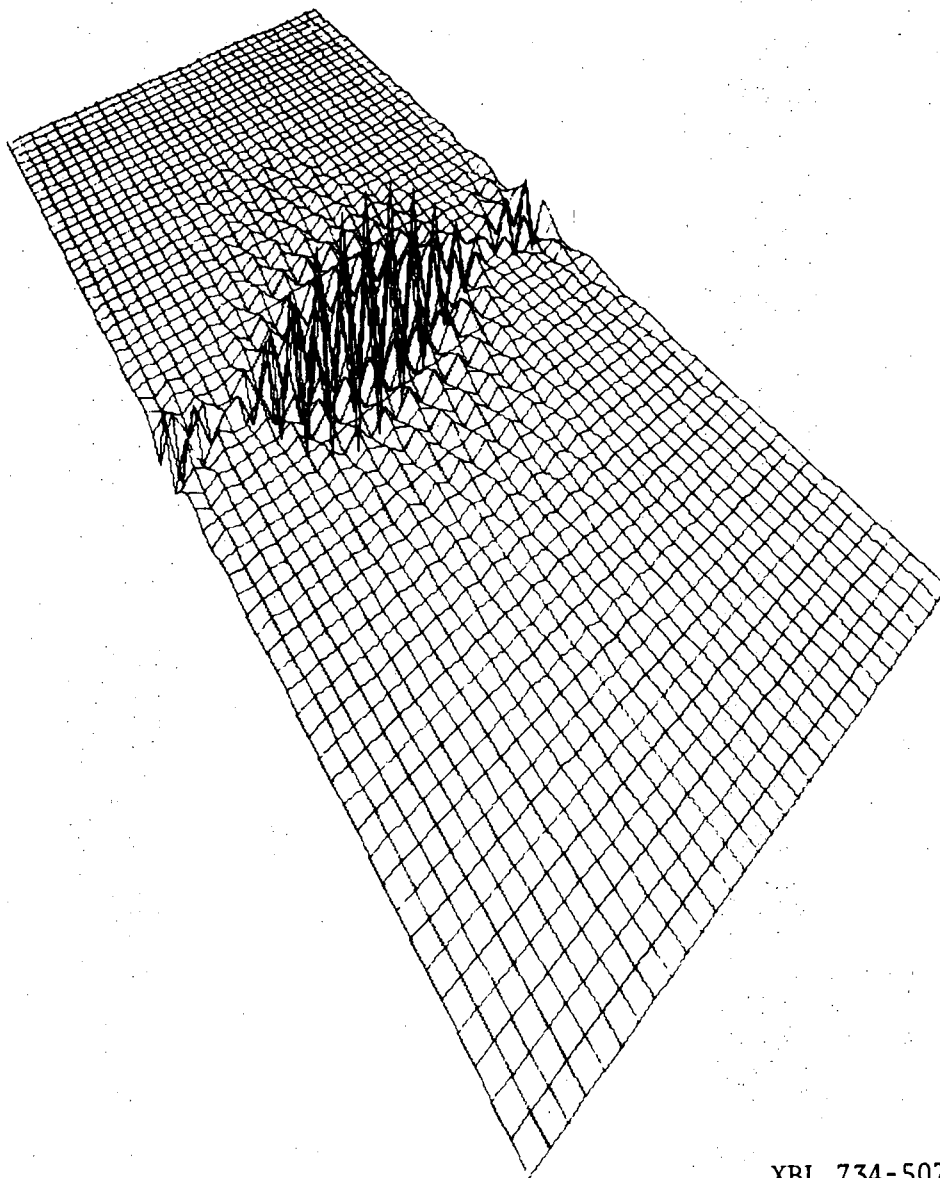
XBL 734-507

Fig. 7
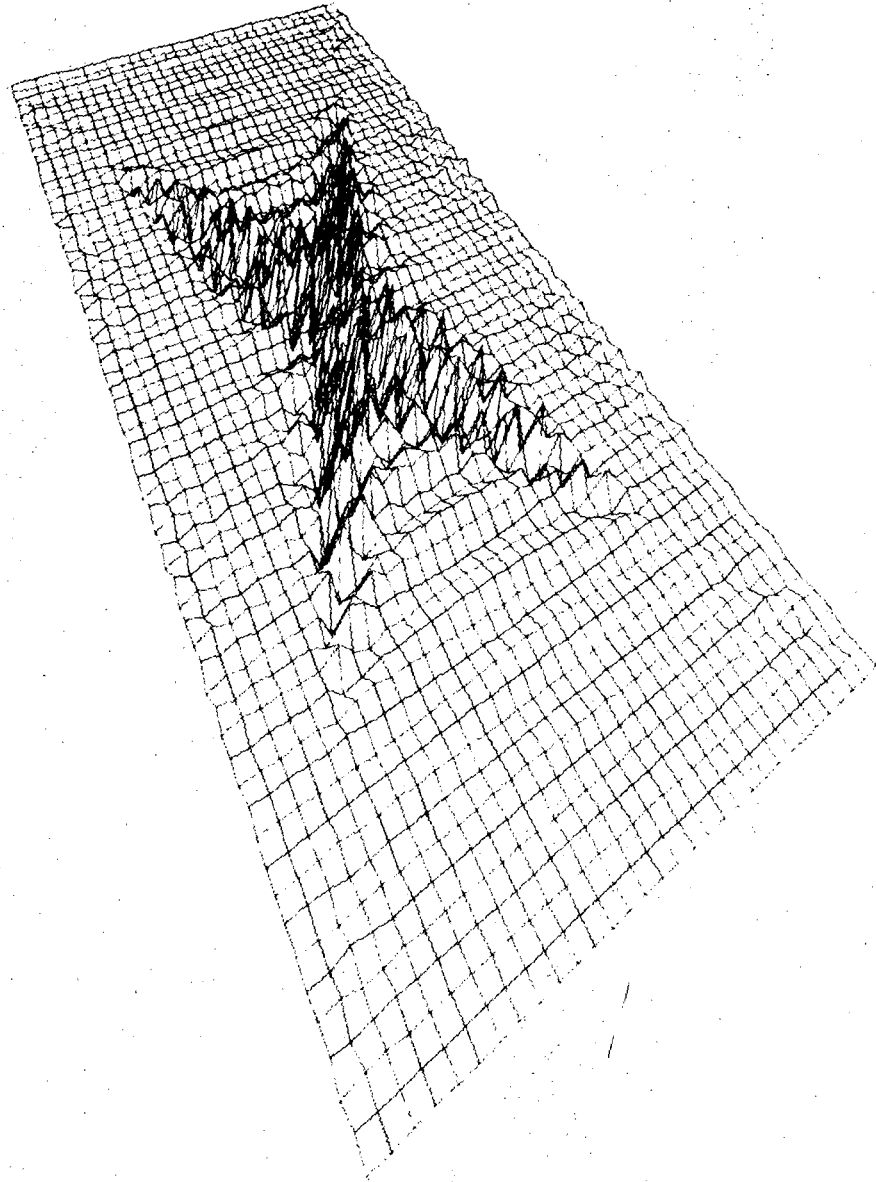
$$y_1 + y_4 - y_2 - y_5$$

$u - v$ PLANE

Fig. 8

XBL 734-506

$$X_3(y_1 + y_4 - y_2 - y_5)$$

$\mathcal{U} - \mathcal{V}$ PLANE
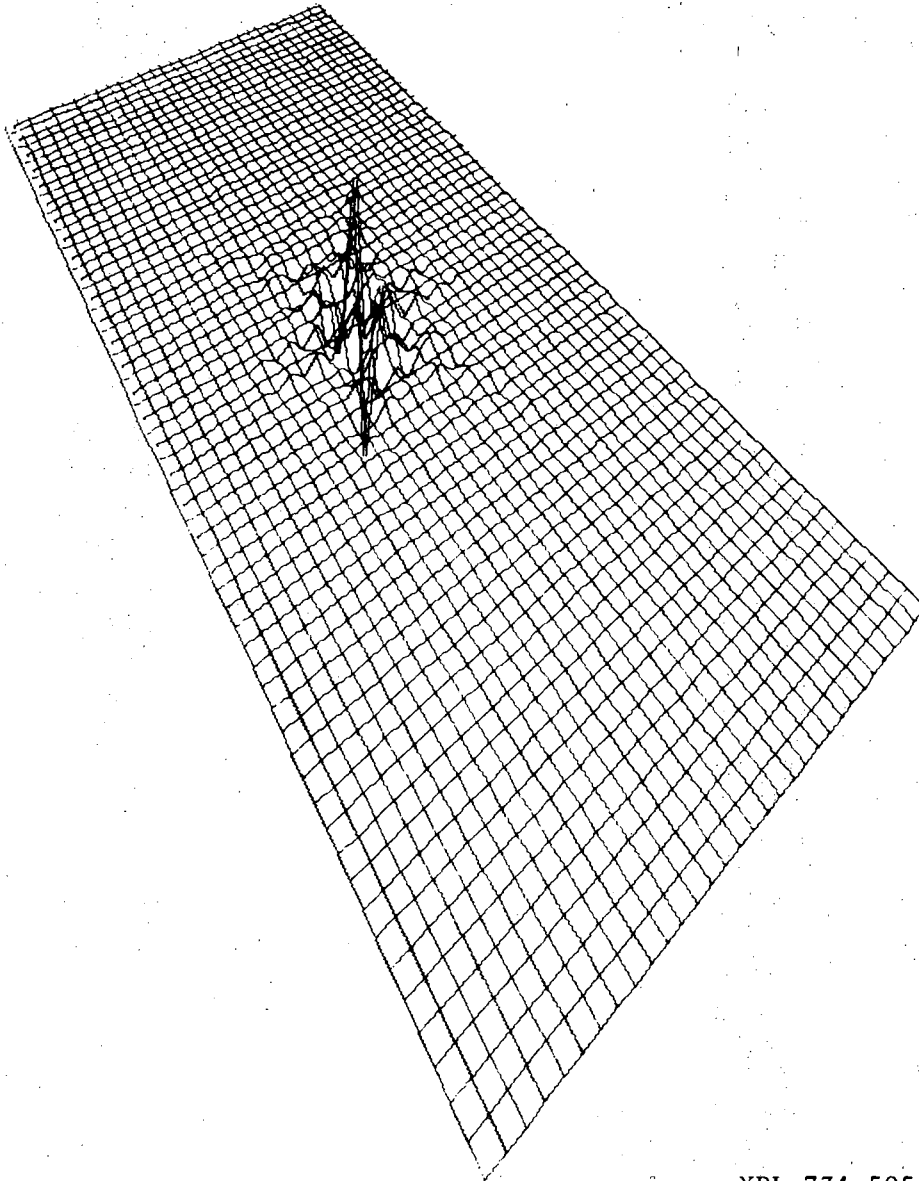
XBL 734-505

Fig. 9

COMPUTER TRANSFORM OF THE
MÜLLER-LYER ILLUSION

In this section, the second approach is discussed, wherein computer transforms of the illusory figures are performed.

In order to use the library subroutine for the fast Fourier Transform available, the illusory figures were digitized in 64 X 64 square arrays (see figure 1). The justification for this approximation to the straight line drawing shown earlier is the fact that the illusory effect is still evident in the digital version (examine figure 1a, b). Digitation was achieved by drawing the illusory figures, using 45° angles between branches and center line, on a 64 X 64 grid, and assigning to each square a value of 1 if a line traversed it and 0 otherwise. The digitized figure was then "thickened" to a line-width of three squares, to avoid the noise problems associated with Fourier Transforms of lines that are one element thick (such lines have the properties of δ functions,

which yield a constant amplitude, i.e., background noise, in Fourier space). An algorithm for generating a duplicate of the array was devised, and this was used as the input array to the transform subroutine. An output display package was assembled using a three-dimensional perspective plot subroutine called PLOT3D and numerous print subroutines.

The display system and transform routine were tested using a simple two-dimensional step function input. As the input function was centrosymmetric (i.e., symmetric about the origin so that $f(x, y) = f(-x, -y)$), no imaginary term was anticipated,[*] and to the limit of the resolution of the 64×64 array, none was obtained (see Figure 2).

$$\text{[*]} \quad F\{f(x, y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-2\pi i(ux+vy)} dx\, dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \cos 2\pi(ux+vy) dx\, dy$$

$$- i \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \sin 2\pi(ux+vy)\, dx\, dy.$$

$\sin(x, y)$ is an odd function, $f(x, y)$ is an even function, so their product is odd.

So, we have

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\text{odd function}) dx\, dy = \lim_{a \to \infty} \int_{-a}^{a} \int_{-a}^{a} (\text{odd function}) dx\, dy = 0.$$

The real part of the transform of the step function was expected to look like

$$\frac{\sin(2\pi\ell_x u)\sin(2\pi\ell_y v)}{\pi\ell_x u \cdot \pi\ell_y v}$$, where ℓ_x and ℓ_y are the lengths of the step function in the x and y directions, respectively,

and the fact that the plot showed this (see Fig. 2), established the dependability of the whole package. The digital versions of the illusory figures were transformed (see Figures 3-6), and the transform arrays compared. As can be seen in illustrations 3 through 6, there is little obviously discernable difference between the transforms of the two figures. It should be noted that, relative to the real part, the imaginary part of the Fourier array has very small amplitude, a fact that is obscured by the normalization of amplitudes within each array, but is evident in the intensity plots (Fig. 6), where real and imaginary parts are added, and which are virtually identical to the real plots. To further investigate the transforms, it was decided to truncate the high frequencies of each transform array and examine the effect on the original figure. This was done by inverse transforming each truncated array, calculating the values of the intensities from the resulting array, and plotting the array of intensities.

The calculated <u>intensities</u> were plotted instead of the complex <u>amplitudes</u>, because the amplitudes of the modified output array, unlike the input, contained imaginary terms. As the purpose of the reconstruction procedure was to <u>view</u> the modified, but still real-valued, input figure now represented by the truncated Fourier array, no imaginary terms could appear in the final output array.

Calculating the intensity was a good compromise, as contributions

from both real and imaginary parts were taken into account, as

follows: the amplitude array was arranged so that the odd rows

of elements were real-valued, the even rows imaginary. A com-

plex point in the amplitude array consisted of a pair of elements:

a real-valued term in an odd row and the corresponding imaginary-

valued term directly below it in the next row. The complex ampli-

tude array consisted of 32 complex rows and 64 vertical columns,

and was converted to a 32 by 64 real-valued intensity array by

converting each complex point to a real point. This was done by

adding the squares of the real and the imaginary elements of each

complex point (i, j), to obtain a new real value, and storing it in

the corresponding location (i, j) in the intensity array. The plots

shown in figs. 2-46 are perspective plots of the intensity arrays

calculated from the various truncations. It is in no way claimed

that such a reconstruction actually occurs in the visual process,

as there would be no need for translation of the Fourier informa-

tion back into Euclidean terms; the brain need operate only in the

Fourier domain, and learn to associate Fourier representations

with particular stimuli in Euclidean space. These plots are

strictly a visual aid to the researcher in interpreting modifications

in Fourier space. Truncation was achieved by imposing a square-shaped

filter (step-function of height 1 and variable side length) on the transform

array, which set all array elements equal to zero that lay outside

a square centered on the zero-order Fourier element (the point

(33,33)). As was discussed in the section on Mathematical Background, the inverse transform of such a filtered transform array yields a convolution of the original figure with the Fourier transform of the square step-function used as the filter, i.e., $\{ \longleftrightarrow \} \div F\{ \square \}$. The length of the square side was varied from 12 elements (Fourier terms) to 32, which is about half maximum resolution, and some of the results are shown in illustrations 7 thru 10. What is clearly demonstrated is that the illusory effect has been brought about by clipping high frequency terms in the transform array with a square filter—the center line in figure 10A is only 29 elements long, while that of figure 10B is either 33 or 35 long, depending on where the ends of the center line are defined to be. The center lines in the original figures were identical in every way, so the process of truncation, using a square filter, has been shown to give rise to the illusory effect. The question of distortion in the reconstruction due to the <u>specific filter shape</u>, rather than the general process of filtering the trans- form array, remains, because of the oscillatory properties of the particular convolution involved. However, the inverse Fourier transform of a square step-function in Fourier Space is a sinusoi- dally oscillating function in Euclidean space,

$$\frac{\sin(2\pi\ell x)\sin(2\pi\ell y)}{\pi^2\ell^2 xy}$$

where ℓ is the length of the square's side. Now, if ℓ is small (such as when $\ell = 12$, (Fig. 7A)), the frequency of oscillation is low, and the $1/\ell^2 xy$ envelope drops off slowly, so sinusoidal

distortion extends over the whole transform array; large ℓ yields a much higher frequency oscillation, but concentrated within a small region of the x-y plane. Thus, when ℓ is large, the oscillatory effects on the convolution are vastly reduced compared to the case when ℓ is small. It was thus postulated that the effects noted at a square size of 32 elements were due virtually only to loss of high frequency terms.

To test the above postulate an algorithm was written to generate a variable Gaussian filter (Fig. 34A) to impose on the transform array, centered arround the zero order Fourier term. The Gaussian function, $e^{-(x^2+y^2)/\sigma^2}$, is particularly desirable, because its Fourier transform is another Gaussian function in the Fourier variables u and v. As no distorting oscillatory phenomena are involved in the convolution of reconstruction, the theoretical difficulties of the square filter are avoided by using the Gaussian filter. The algorithm included a variable " standard deviation", σ, for the Gaussian function, which was varied from 2 terms to 64 terms. On the circle whose radius is σ, that is, $\sqrt{x^2+y^2} = \sigma$, the value of the Gaussian function is $1/e$ times the value at x, y = 0. By varying σ, the width of the filter is varied, and thus the amount of high frequency information used in the reconstruction is varied. To see how quickly the function $e^{-(x^2+y^2)/\sigma^2}$ approaches zero as the radius, $\sqrt{x^2+y^2}$, is increased, consider the following table:

$\sqrt{x^2+y^2}$	$e^{-(x^2+y^2)/\sigma^2}$
0	1
σ	$1/e \quad \sim 1/3$
2σ	$1/e^4 \quad \sim 1/55$
3σ	$1/e^9 \quad \sim 1/8000$

Multiplication of the elements of a transform array by this func-
tion, centered about the zero order term, smoothly eliminates
Fourier information beyond a radius of about 2σ (measured in
number of elements, or terms from center). Thus, reconstruc-
tion from a Gaussian filtered transform array with $\sigma = 2$ uses
essentially only the Fourier information within a radius of 2
terms around the zero order term and strongly diminished terms
out to a radius of 4 elements. With $\sigma = 64$, the Gaussian func-
tion imposed on the transform array leaves the Fourier informa-
tion essentially unmodified. Some results of this experiment are
shown in Figs. 11 thru 17: at any given σ up to $\sigma = 28$, the
center line in the second Müller-Lyer figure is longer than that
of the first, as measured in number of array elements. It is of
particular note that at any given σ up to $\sigma = 28$, it is far easier
to distinguish a bracket vertex from its adjacent center-line
endpoint in the open-ended figure than in the closed-ended figure.
This can be explained by the fact that the closed figure contains
elements of higher spatial frequency than the open figure, at the
bracket vertices. The intersection of three lines within a 45°
angle, as in the closed figure, involves very high spatial fre-
quencies at the vertex, and this situation is lacking in the open-

ended figure, where three lines intersect within a 360° angle.

When high spatial frequency terms are truncated in the Fourier

domain, the closed figure is thus more affected by a given amount

of truncation than the open figure. This would seem, in part, to

account for the illusory effect.

One might ask what biologically realizable form such

truncations could take, and how they would be brought about. It

was mentioned in the section on optical illusions that the processes

of information manipulation and of pattern interpretation interact

with another, and it was suggested that one such interaction

would be a selective attention mechanism. With such a mechanism,

the brain would evaluate only certain portions of the input at any

given time, and this could take the form of a filter. Were the

interpretation process of the brain to proceed from low resolution

to high, processing more input information as it is needed, this

could be accomplished in the Fourier domain by imposing a centro-

symmetric filter of variable diameter on the array of Fourier com-

ponents. In the Pollen and Lee model, this would correspond to

selective "read out" of the information in the complex cells of the

cortex, which could be effected by selective inhibition of the firing

rates of those cells connected to adjacent cells leading further into

the brain. The precise nature of the inhibition would determine

the mathematical shape of the imposed filter, and several filter

shapes have been used in this experiment to accommodate various

inhibitory schemes. The square filter corresponds to a precise

inhibitory mechanism in which individual neurons are separately and binarily controlled, a model of somewhat dubious biological relevance because of the sharpness of the cutoff between "on" and "off" cells. The Gaussian filter corresponds to a situation in which readout of cortical information is inhibited most of the time and occurs only when a specific "excitatory signal" is sent to the desired region of the cortical array. The Gaussian shape would arise if the excitatory signal had non-local effects which stimulate readout of surrounding neurons in a manner which decreases as a Gaussian function of distance from the signal center. Other inhibitory schemes involving Gaussian-like filters are also possible, but the above example should suffice to illustrate the biological relevance of a Gaussian filter.

Another filter to be tested was a truncated Gaussian filter, in which the "flat top" had a height of 1 and a variable radius equal to the "σ" of the Gaussian function which was truncated (Fig. 34B). Beyond a radius of σ, the function steeply falls off as $e^{-(x^2+y^2/\sigma^2)}$. This filter corresponds to a biological model similar to that of the Gaussian filter, in which the "readout" neurons have a maximum output threshold, which is reached when an excitatory signal equals $1/e$ times the strength of the signal at its center. This would occur within a radius of σ of the signal center, and sub-threshold neural firing rates would continue beyond σ, falling off as $e^{-(x^2+y^2)/\sigma^2}$. The sharp transition between the circular step function for distances $\leq \sigma$, and the steep

Gaussian drop-off thereafter, predictably gave rise to oscillations in the reconstructed figure, somewhat similar to the square filter oscillations. However, as can be seen in figures 18—22, this like the other filters, produced the illusory effect out to high resolutions, indicating that virtually any sort of symmetric selective inhibitory mechanism in a Fourier transform neural array would give rise to the Müller—Lyer illusory effects. Indeed, Ginsberg has shown that the use of either square, rectangular, or circular filters in the Fourier domain gives rise to the known illusory effects of all major geometrical illusions, for filter sizes of about four to eight Fourier terms.[7]

The last filter to be tested was a combination of a Gaussian function with a cubic function of $\sqrt{x^2+y^2}$, $f(\sqrt{x^2+y^2})$ = $f(r) = ar^3 + br^2 + cr + d$, so that the two-dimensional filter looked like a " smooth volcano" (see Fig. 34C and D). The odd shape was suggested, qualitatively, in a recent article on contrast illusions (Scientific American, June, 1972), in which the author proposed a model for visual information processing involving an extensive amount of lateral inhibition in the retina and other neural networks. The result of such inhibition, the author claimed, would be supression of low spatial frequency amplitudes, relative to intermediate frequency amplitudes, and other factors would cause high spatial frequency filtering, the the net combination leading to emphasis of intermediate spatial frequencies over all. As the proposed filter was only qualitatively described, a certain amount of liberty was taken in mathematically designing a function which effectively emphasizes the intermediate

frequency range. Two approaches were taken: The low frequency terms were actually reduced in amplitude, while intermediate terms were essentially unchanged and high frequency terms filtered by a Gaussian function. The actual function used was:

$$f(x,y)=\begin{cases} \dfrac{2(d-1)}{\sigma^3}\left(\sqrt{x^2+y^2}\right)^3 + \dfrac{3(1-d)}{\sigma^2}\left(\sqrt{x^2+y^2}\right)^2 + d, & 0 \le \sqrt{x^2+y^2} < \sigma \\ \exp-\left[\left(x^2+y^2+\sigma^2-2\sigma\sqrt{x^2+y^2}\right)/\sigma^2\right], & \sigma \le \sqrt{x^2+y^2} \end{cases}$$

where "d" is the adjustable height of the z-intercept, $f(0,0)$, and "σ" is the adjustable width parameter for the Gaussian function. Thus, as can be seen in Fig. 31C, for the interval $0 \le r < \sigma$ along any radius from the origin in the x-y plane, the filter is a cubic function of distance along the radius, with z-intercept at the adjustable value "d", and height 1 at a distance of σ units from the origin. At both ends of its domain, $0 \le r < \sigma$, the function has derivatives equal to zero, and at $r = \dfrac{\sigma}{2}$, the second derivative is zero. For $r \ge \sigma$, the function is "Gaussian-like", meaning that its height is 1 at $r = \sigma$ and decreases as $e^{-(r-\sigma)^2/\sigma^2}$ thereafter, for increasing r. Thus, "d" varies the number of both low and high frequency terms that are suppressed.

The second approach was to leave the low frequency terms essentially unchanged, while boosting intermediate frequency amplitudes and, again, filtering high frequency terms with a Gaussian function. This took a mathematical form nearly identical to the first approach:

$$f(x,y) = \begin{cases} \frac{2(1-d)}{\sigma^3}(\sqrt{x^2+y^2})^3 + \frac{3(d-1)}{\sigma^2}(\sqrt{x^2+y^2})^2 +1, & 0 \le \sqrt{x^2+y^2} < \sigma \\ d \cdot \exp- \left[(x^2+y^2+\sigma^2 -2\sigma\sqrt{x^2+y^2})/\sigma^2 \right], & \sigma \le \sqrt{x^2+y^2} \end{cases}$$

Here, however, the z-intercept is fixed at 1, while the peak of the Gaussian has height "d", which is adjustable. Once again, σ varies both the peak and the width of the Gaussian function (see Fig. 34D).

For both approaches, it should be noted that terms out to a radius of about 2σ around the zero-order term are taken into account in reconstruction, using this filter, as opposed to a radius of σ in the previous case using a pure Gaussian filter. Some illustrative results of the first approach are shown in Figs. 23—30. These figures are representative of the results obtained when "d" assumed the values 0.25, 0.50 and 0.75, while "σ" was varied from 2 to 22. As can be seen in all those marked "A", the center line is clearly separated from the vertices of the closed brackets, giving the line a shortened length of 29 or 30 elements, as compared with the "B" figures, whose lengths are all 33 to 34, depending on where the endpoints are defined to be. It is of particular interest that when $\sigma > 10$ and d = 0.25, (Figs. 26 and 28) the center line and vertices of both Müller-Lyer figures are significantly reduced in height, an effect which is most striking at the intersections of center and bracket lines in the "closed" figure (those labeled "A"). It will be recalled that when the pure Gaussian filter was used, at any but the highest reconstruction resolution, these points of

intersection were greatly magnified in heigh with respect to the

rest of the figure, and for the closed figure, it was difficult to

distinguish center line from brackets at such points. This is

more or less the case with the new filter for $\sigma \leq 10$, even when

d is a small as 0.25 (Figs. 23, 24 and 27). However, when

$\sigma > 10$ and d = 0.25, the effect is reversed dramatically. Some

insight into this effect can be seen in Fig. 35A. In the figure,

plots of the cross-section of the filter function along the positive

x-axis are seen for d = 0.25 and $\sigma = 6, 10, 14$ and 18. The number

of Fourier terms which each function diminishes by 50% or more

is indicated by the vertical arrow, and the number of terms

diminished by 25% is indicated for $\sigma = 10$ and $\sigma = 14$. When

$\sigma = 10$, the Fourier terms inside a radius of 4 from the zero-

order term are reduced by 50% or more, while when $\sigma = 14$,

terms within a radius of 6 terms are 50% diminished in amplitude.

Terms are diminished by 25% or more within a radius of 6 for

$\sigma = 10$, and within a radius of 8 when $\sigma = 14$. So, if "significant

reduction in amplitude" is defined as between 25% and 50% re-

duction, then a critical region of difference in the low frequencies

appears to be the "ringed area" bounded by radii of 4 Fourier

terms and 8 Fourier terms around the zero-order term (see

Fig. 35B). Somewhere in this region, it would seem, is crucial

information about the center line and intersection, which is not

eliminated when $\sigma \leq 10$, but is missing when $\sigma = 14$. Upon

examining the Fourier representations of the two Müller-Lyer

figures (Figs. 4 and 5, in particular), one notes that in the plots

of the imaginary terms (Fig. 5), the terms within a radius of about three elements around the center (zero-order term) of the array have zero amplitude. After the third term, significant amplitudes are seen, and at a radius of five terms, the peaks start to diverge from one another. Perhaps the "critical region" spoken of is this area where the low frequency imaginary terms begin to branch apart. Branching of a much more subtle sort is visible in the real-valued plots (Fig. 4), though divergence begins at $\sigma = 2$ in 4B (the open-ended figure) and at $\sigma = 4$ in 4A. Since branching occurs at terms that appear to be outside (prior to) the "critical region" of the real array, and no clearly significant changes in peak features for terms within the region are evident, it appears that if there is a critical region for the center line and intersections of the Müller-Lyer figures, it has much to do with the imaginary terms and little to do with the real terms. These conclusions, however, are only tentative, and call for further investigation. In any event, the Müller-Lyer illusion is clearly demonstrated by use of this filter, giving further support to the general argument that high frequency spatial filtering in the Fourier domain will give rise to the Müller-Lyer illusory effect.

With the second approach, the height of the Gaussian peak, equal to "d", was set to the values 1.25, 1.50, 1.75 and 2.0, while σ was varied from 2 to 22 terms. Figures 31, 32, and 33 are representative of the general results, which lacked the striking effects derived in the first approach when low frequency

terms were diminished. As the main purpose of the filter in this case was to boost intermediate frequency without reducing low frequency terms, it would not be predicted that the center line and intersections would undergo the same diminution as in the previous case, and the fact that such diminution is virtually absent lends support to the arguments given previously. However, the Müller-Lyer illusory effect is clearly evident in this case as well, as the center line in the "A" figures is about 29 terms in length, while that of the "B" figures is 33 to 34 terms long.

Thus, to the knowledge of this author, all tests done to date on the hypothesis that Fourier transforms are involved in the visual process, insofar as optical illusions are concerned, are positive, provided that some method of spatial filtering is postulated as part of the model. It should be pointed out that this scheme also allows for the phenomenon of "learning not to re-spond" to an illusory figure. In the case of the Müller-Lyer illusion, for example, one can greatly reduce or eliminate com-pletely the illusory effect by sketching and staring repeatedly, and for long intervals, at the figures involved. Such concentration would be represented in the model by increase in filter diameter, allowing terms of higher and higher spatial frequency to be "read-out", eventually eliminating the illusory effect (see Figs. 17, 22).

The question arises, can spatial filtering in the Fourier domain cause illusory effects in figures that are normally non-illusory? Clearly, the answer must be affirmative, as any figure

reconstructed from its Fourier representation will be distored

if the transform array has been suitably filtered. What, then, is

to distinguish illusory figures from non-illusory? There must,

apparently, be a " spectrum of illusory quality," with figures as

non-illusory as simple geometric figures at one end, and well-

known illusions such as the Muller-Lyer at the other. The

criterion for measuring the illusory extent of a figure, by the

above model, would be the number of Fourier terms needed to

eliminate the discrepancies noted at low resolution. Thus, figures

whose partial Fourier representations (Fourier arrays truncated

with symmetric filters) non-isotropically describe the figure,

so that some portions of the figure are exaggerated over others,

would require high frequency terms to give an accurate representa-

tion of even the gross features of the figure. Such a figure would

be highly sensitive to Fourier truncation and would tend to be

illusory. It was stated in the section on optical illusions that

size and shape illusions can be regarded as distortions brought

about in otherwise normal figures by placing those figures in a

peculiar background. In light of the foregoing arguments, there

would appear to be an important distinction in types of geometrical

illusions: those where the figure and the background interact

directly by intersection of lines of one with the other, etc. , and

those where there is no contact between figure and background.

Clearly, it is not the latter type that is of major concern to the

Fourier hypothesis, as such illusions most certainly depend on

other visual clues, such as perspective,(Ref. 5) which are not considered

here. Illusions of the former type may be understood, in light of the previous arguments, as resulting when figure and background interact so that new low frequency terms are generated which render the low-resolution Fourier representation nonisotropic, and therefore require higher frequency terms to accurately present the overall features of the figure-background combination.

As a test of this hypothesis, two non-illusory figures were treated in exactly the same manner as the Müller-Lyer illusion: a square, and a "plus sign." Figures 36 and 37 show the results of a Gaussian filter of increasing radius on the representation of a square, while Fig. 38 shows the effects of the truncated Gaussian filter on the same square. The pure Gaussian is seen to effect the square isotropically, with no portion of the figure exaggerated over any other. The truncated Gaussian gives rise to large oscillations in the figure, but does not emphasize one part of it over any other. This is clearly seen in 38B, which is at $\sigma = 10$, in which the oscillations are of smaller amplitude and average out evenly over the surface of the square. Figures 39 and 40 show the effect of a Gaussian filter on a "plus sign". The effect is seen to be completely uniform, with no part of the figure misrepresented relative to other parts. Figures 41 and 42 illustrate the uniform effect of a truncated Gaussian on the same "plus sign," with the associated induced oscillations averaging out at higher resolution, as in the case of the square (Fig. 38).

One further test of the hypothesis that spatial filtering accounts for the Müller-Lyer illusory effect was carried out, in which the branches in each figure were initially separated from the center line by three units (elements), before the arrays were transformed. The modified input (Fig. 43) no longer gives rise to any illusory effect when viewed, and if the previous data are to be useful, the spatial filtering should not give rise to an illusory effect. As seen in Figs. 44 and 45, which have been filtered by a square–and Gaussian–function, respectively, the center lines all have identical lengths, so there is no illusory effect. Only when the Fourier arrays are filtered with very small radii (two or four terms) is there any merging of the center line and branches, and it is highly doubtful that such an effect can be regarded as illusory.

Finally, the importance of a filter being symmetric was tested and the results shown in Fig. 46, where the open-figure of the Muller-Lyer illusion has been reconstructed from a modified version of its transform array. The array was filtered by a Gaussian filter centered at the rear left corner of the array (the point (0, 0)), which eliminated terms that were not in the region of that corner of the array. The original figure was the same as Fig. 3B, with four symmetric branches and a center line, and there remain now only remnants of two opposing branches. Thus, imposition of a non-symmetric filter has completely altered the figure; this cannot, however, be fairly termed an "illusion", as the filter used was highly biased from the start.

CONCLUSION AND CLOSING REMARKS

Through the use of many different spatial filters, it has been shown that spatial filtering of Fourier representations of the Müller-Lyer figures gives rise to the known illusory effect. More significantly, the effect has been generated by <u>symmetric</u> spatial filtering, which has as little biase as possible regarding <u>which</u> frequencies are to be eliminated. Asymmetric filtering has been shown to cause effects that have no parallel in nature, and while such effects may be dramatic, they are artifically produced and do not represent fundamental illusory properties of the figure thus modified. What has emerged is an understanding of <u>what constitutes</u> an illusory figure, in terms of a model of the visual system which involves Fourier transforms. A figure whose Fourier representation is so asymmetric that filtering with a centro-symmetric function leads to a distorted reconstruction, where some portions of the figure are exaggerated over others, in understood to be illusory. By this criterion, many figures that are not com-monly regarded as illusory would be deemed illusions—examples of this are Moiré patterns and similar figures involving intersecting sets of closely spaced, parallel lines. Several biologically reasonable models for spatial filtering have been outlined, to give some credibility to the hypothesis that such filtering could take place in the visual pathway, but there clearly remains a great deal of work to be done in this area. Finally, analysis of the Fourier transform arrays for the Müller-Lyer figures has led to the implication of the

imaginary parts of 4th through 8th order Fourier terms as crucial in the representation of the center line and points of intersection of the figures. This analysis, however, is far from complete, and there is good reason to believe that further work in this direction would lead to a clearer understanding of how geometrical illusory effects are generated in a Fourier model of the visual system.

It has been the goal of this thesis to discuss the mathematical and biological background of, the motivation for, and some implications of an important contemporary model of the human visual pathway - a model involving the Fourier transformation of visual information. The author hopes that, at the very least, this work will help the reader to dissipate any existing "clouds of mystery" surrounding the Fouier transform and its numerous uses, especially insofar as its possible involvement in neural information processing is concerned. It is further hoped that this work will aid future researchers in their efforts to elucidate the phenomenal complexities of the human mind, and finally, that such research will lead to a significant contribution to the improvement of the quality of life itself.

PLOT LABELING AND NOTATION

1. $\mathcal{F}\left\{\longleftrightarrow\right\}$ = two-dimensional Fourier transform of the figure in brackets.

2. $Re\left\{\mathcal{F}\left[\longleftrightarrow\right]\right\}$ = the real part of the transform of the figure in brackets.

3. $I_{m^2}\left\{\mathcal{F}\left[\longleftrightarrow\right]\right\}$ = the square of the imaginary part of the transform of the bracketed figure.

4. $\mathcal{F}^2\left\{\longleftrightarrow\right\}$ = the square of the transform (Fourier intensity) of the bracketed figure.

5. $\mathcal{F}^{-1}\left\{\mathcal{F}\left[\longleftrightarrow\right]\right\}$ = inverse transform of the transform of the bracketed figure (reconstruction)

6. $\mathcal{F}^{-1}_{r=10}\left\{\mathcal{F}\left[\longleftrightarrow\right]\right\}$ = same as 5, but the transform array of the figure in brackets has first been truncated by a square (step-function) filter, centered at the origin of the u - v plane, with a side length of 20 Fourier terms (r = "radius" = 10 terms). This is equivalent to the convolution between the figure in brackets and the transform of the square filter function (see section on the Convolution of Two Functions, in Mathematical Background).

7. $\left\{ \longleftrightarrow \right\} * \mathcal{F} \left\{ \triangle \right\}$ = the <u>convolution</u> of the figure in the brackets to the left with the transform of the bracketed figure on the right.

8. $\left\{ \longleftrightarrow \right\} * \mathcal{F} \left\{ \curvearrowright \quad \sigma = 4 \right\}$ = <u>convolution</u> of the bracketed figure on the left with transform of a two-dimensional <u>Gaussian function</u> (Fig. 34 A), whose width parameter, σ, has a value of 4 Fourier terms.

9. $\left\{ \longleftrightarrow \right\} * \mathcal{F} \left\{ \overbrace{} \quad \sigma = 4 \right\}$ = convolution of the bracketed figure on the left with the transform of a "flat-top" <u>Gaussian function</u>, illustrated in <u>figure 34 B</u>, where the "top" has a radius of $\sigma = 4$ Fourier terms.

10. $\left\{ \longleftrightarrow \right\} * \mathcal{F} \left\{ \begin{array}{c} 1- \\ \text{—d—} \end{array} , \quad d = 0.05, \quad \sigma = 18 \right\}$ = <u>convolution</u> of the bracketed figure on the left with the transform of the <u>"volcano function"</u> (Fig. 34 C, illustration; explanation of origin on p. 10 of the section on the <u>COMPUTER TRANSFORM OF THE MÜLLER-LYER ILLUSION</u>), in which the central minimum has height $d = 0.05$, and the "ridge maximum" occurs at a radius of $\sigma = 18$ Fourier terms from the center of the transform array. Beyond

the radius of σ, the function drops
off as a Gaussian exponential.

11. $\left\{ \longleftrightarrow \right\} * \mathcal{F} \left\{ \begin{array}{l} d- \\ \sim\sim\sim -1 \end{array} \right.$, $d = 2.0$, $\sigma = 18 \left. \right\} =$

the same as #10, but the "volcano
function" has a central minimum of
height 1, and a "ridge maximum of
height d = 2.0 (Fig. 34 D).

a.

b.
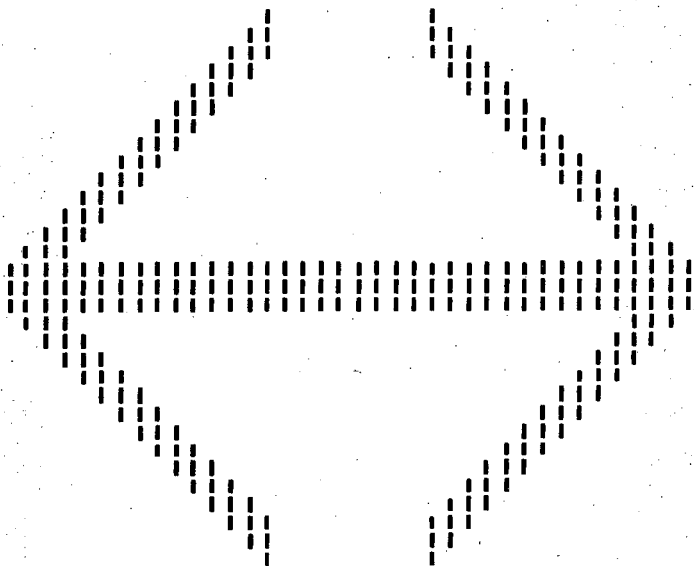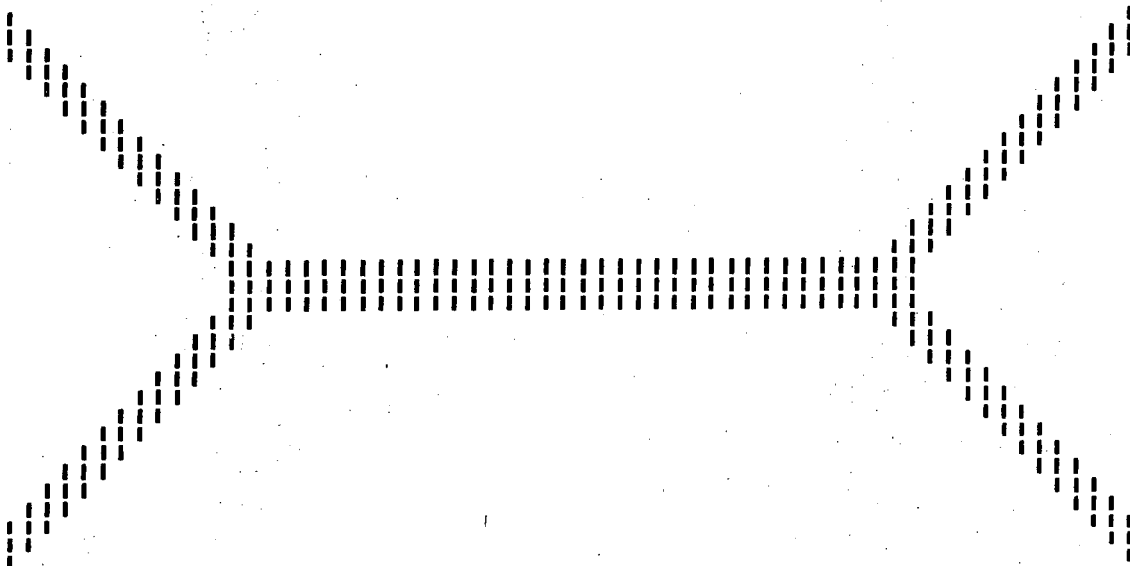
Fig. 1

$$Re^2\left\{\mathcal{F}\left\{-\right\}\right\} = \mathcal{F}^2\left\{-\right\}$$

(NO IMAGINARY TERMS)

XBL 734-508

Fig. 2

A $\{\leftrightarrow\}$ x-y PLANE

B $\{Y\}$ x-y PLANE

Fig. 3

XBL 735-589

Fig. 4

$$\text{A} \quad I_m^2 \left\{ \mathcal{F} \left\{ \longleftrightarrow \right\} \right\}$$

$$\text{B} \quad I_m^2 \left\{ \mathcal{F} \left\{ \curlyvee \right\} \right\}$$

Fig. 5

XBL 735-591

A

$$(Re^2 + Im^2)\left\{\mathcal{F}\{\updownarrow\}\right\} = \mathcal{F}^2\{\updownarrow\}$$

B

$$(Re^2 + Im^2)\left\{\mathcal{F}\{\}\right\} = \mathcal{F}^2\{\}$$

Fig. 6

XBL 735-592

Fig. 7

XBL 735-593

A

$\mathcal{F}^{-1}_{r=10}\left\{\left\{\mathcal{F}_\xi\longleftrightarrow\right\}\right\}$

B

$\mathcal{F}^{-1}_{r=10}\left\{\left\{\mathcal{F}_\xi\curlyvee\right\}\right\}$

Fig. 8

XBL 735-594

Fig. 9

XBL 735-595

Fig. 10

Fig. 11

XBL 735-597

Fig. 12

Fig. 13

Fig. 14

XBL 735-600

Fig. 15

XBL 735-601

Fig. 16

XBL 735-602

Fig. 17

XBL 735-606

XBL 735-605

Fig. 18

Fig. 19

XBL 735-604

Fig. 20

-120-

Fig. 21

XBL 735-607

Fig. 22

XBL 735-608

Fig. 23

XBL 735-609

A $\{\xi \leftrightarrow \xi * \mp \xi^{-1} \sqrt{\sqrt{\xi}}\, d=0.25, \sigma=10\}$

B $\{\xi \curlyvee \xi * \mp \xi^{-1} \sqrt{\sqrt{\xi}}\, d=0.25, \sigma=10\}$

XBL 735-610

Fig. 24

Fig. 25

XBL 735-611

Fig. 26

XBL 735-612

Fig. 27

XBL 735-613

A $\{\xi \longleftrightarrow \xi\} * \mathcal{F}\xi^{*}_{\xi}\int \mathcal{M}_{\mathcal{A}} \ d=0.25, \sigma=18\}$

B $\{\xi \curlyvee \xi\} * \mathcal{F}\xi^{*}_{\xi}\int \mathcal{M}_{\mathcal{A}} \ d=0.25, \sigma=18\}$

Fig. 28

XBL 735-614

-128-

Fig. 29

XBL 735-615

A $\{\xi \leftrightarrow \xi\} * \mathcal{T}\xi \mathcal{T}\Lambda d=0.75, \sigma=18\}$

B $\{\xi \leftrightarrow \xi\} * \mathcal{T}\xi \mathcal{T}\Lambda d=0.75, \sigma=18\}$

Fig. 30

XBL 735-616

Fig. 31

XBL 735-617

A $\{\xi \longleftrightarrow \{\xi * \underset{\sigma}{\mathscr{F}} \xi \underset{\sigma}{\$} \eta\}_{-1}$ $d=2.0, \sigma=10\}$

B $\{\xi \longrightarrow \{\xi * \underset{\sigma}{\mathscr{F}} \xi \underset{\sigma}{\$} \eta\}_{-1}$ $d=2.0, \sigma=10\}$

Fig. 32

XBL 735-618

Fig. 33

XBL 735-619

a.

Gaussian

1

1/e

x

y

b.

e

1

x

y

c.

1

d

-3σ

-2σ

-σ

σ

2σ

3σ

x

y

d.

d

1

-3σ

-2σ

-σ

σ

2σ

3σ

x

y

Fig. 34

A

Fig. 35 A

B

DBL 728·5435

Fig. 35 B

A $\{\square\} * \mathcal{F}\{\sigma_{SMALL}\}$

B $\{\square\} * \mathcal{F}\{\sigma_{INTERMEDIATE}\}$

XBL 735-620

Fig. 36 A

Fig. 36 B

-136-

A $\{\square\}*\mathcal{F}\} \wedge \sigma\text{ LARGE}\}$

Fig. 37 A

B $\{\square\}*\mathcal{F}\} \wedge \sigma=64\}$

XBL 735-621

Fig. 37 B

A $\{\square\}*\{\mathcal{I}\}\sim \sigma=5\}$

B $\{\square\}*\{\mathcal{I}\}\sim \sigma=15\}$

Fig. 38 A

Fig. 38 B

XBL 735-622

Fig. 39 A

Fig. 39 B

XBL 735-623

Fig. 40 A

Fig. 40 B

XBL 735-624

A

$\{+\} * \mathcal{F}\{ \sim \sigma=7\}$

B

$\{+\} * \mathcal{F}\{ \sim \sigma=15\}$

XBL 735-625

Fig. 41 A

Fig. 41 B

A $\left\{\left\{\right\} + \left\{\right\} * \left\{\right\} \underbrace{}_{\sigma = 63} \right\}$

Fig. 42 A

B $\left\{\left\{\right\} + \left\{\right\} * \left\{\right\} \underbrace{}_{\sigma = 25} \right\}$

Fig. 42 B

XBL 735-588

$\left\{\dfrac{}{}\right\}$ x, y PLANE

A

Fig. 43·A

$\left\{\underset{}{}\right\}$ x, y PLANE

B

XBL 735-626

Fig. 43 B

A $\mathcal{F}^{-1}_{r=16}\{\mathcal{F}\{\mathcal{W}\}\}$

B $\mathcal{F}^{-1}_{r=16}\{\mathcal{F}\{\mathcal{W}\}\}$

XBL 735-627

Fig. 44

A $\left\{\leftrightarrow\right\} * \mathcal{F}\left\{\bigwedge_{\sigma=16}\right\}$

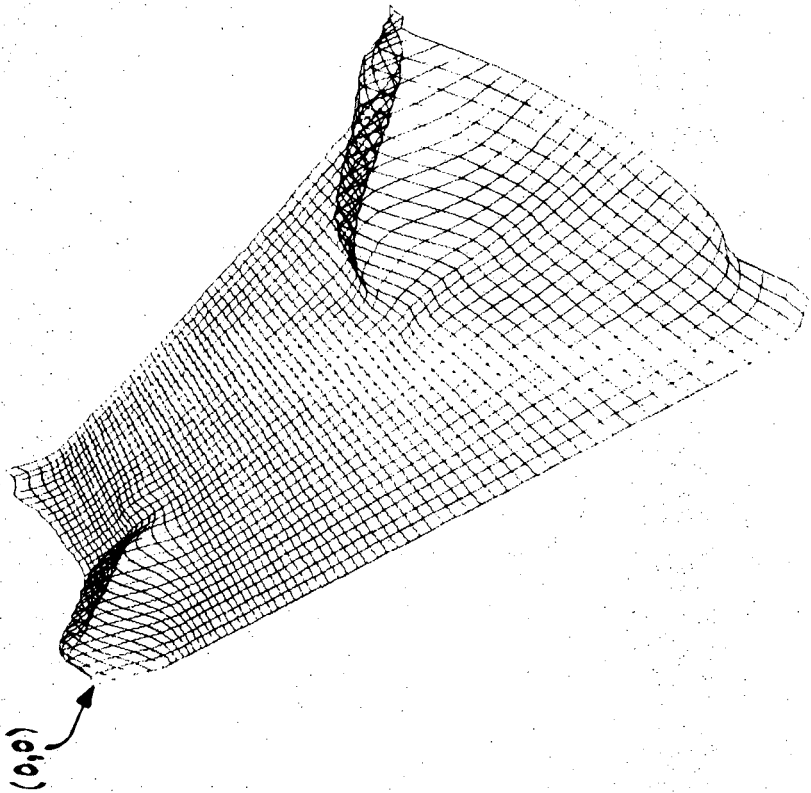B $\left\{\rangle-\langle\right\} * \mathcal{F}\left\{\bigwedge_{\sigma=16}\right\}$

Fig. 45 A

Fig. 45 B

XBL 735-628

Fig. 40

XBL 735-629

REFERENCES

1. Adler, Physiology of the Eye, 1959. Pp. 574-616.

2. Bracewell, R. The Fourier Transform and Its Applications. McGraw-Hill, 1965.

3. Bracewell, R. Australian Journal of Physics, 9, 198 (1956).

4. Brindley, Physiology of the Eye and Visual Pathway, 1970.

5. Day, R. H. Visual Spatial Illusions: A General Explanation, Science, Vol. 175, pp. 1335-1340, March, 1972.

6. Campbell, F. W. and Robson, J. G., J. Physiology, 197, 551 (1968).

7. Ginsberg, A. Physiological Correlates of a Model of the Human Visual System. Masters Thesis, USAF, Faculty of Engineering,

8. Glass, Leon, Nature, Vol. 223, Aug. 9, 1969. P. 578.

9. Goodman, J. W. Introduction to Fourier Optics. New York, McGraw-Hill, 1968.

10. Hubel, D. The Visual Cortex of the Brain, Scientific American, Nov. 1963, pp. 54-63.

11. Luckiesh, Visual Illusions New York, 1922.

12. Pollen, D., and Lee, How Does the Striate Cortex Begin the Reconstruction of the Visual World? Science, Vol. 173, July 2, 1971. Pp. 74-77.

13. Prugovevki, Quantum Mechanics in Hilbert Space, Academic Press, 1972.

14. Spinelli, Prof. Nico D., Dept. of Psychology, Stanford Univ. School of Medicine private communication.

15. Taylor, J. H., The Astrophysical Journal, Vol. 150, Nov., 1957.

16. Werblin, Dept. of Physiology, UC Berkeley. Ongoing Re-
    search, conveyed to author by Jeff Goldberg, one of Dr.
    Werblin's graduate students.

17. Westheimer, Gerald. "Optical and Motor Factors in the Formation
    of the Retinal Image." Journal of the Optical Society of America,
    Vol. 53, No. 1, Pp. 86-93. January, 1963.