FREQUENCY VARIANT OPTICAL SIGNAL PROCESSING

By

James M. Florence

This work was supported by the
National Science Foundation
Grant No. ENG 75–17724

January 1979

GEORGIA INSTITUTE OF TECHNOLOGY
SCHOOL OF ELECTRICAL ENGINEERING
ATLANTA, GEORGIA 30332

1979

FREQUENCY-VARIANT OPTICAL SIGNAL PROCESSING

by

James M. Florence

January 1979

This work was supported by the
National Science Foundation
Under Grant: ENG75-17724
Principal Investigator: W. T. Rhodes

Optical Information Processing Laboratory
School of Electrical Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332

ACKNOWLEDGMENTS

I am greatly indebted to my thesis advisor, Dr. W. T. Rhodes, whose enthusiasm and guidance made this work possible.

I wish to thank the members of my reading committee, Dr. T. K. Gaylord and Dr. R. W. Schafer, for their careful reading and helpful suggestions.

I also wish to express my gratitude to Dr. A. M. Bush for his comments regarding general linear systems theory and to Mr. L. J. Laub for his numerous comments and suggestions regarding optical systems and optical processing.

Finally, my deepest appreciation goes to my wife, Maureen, for her patience and encouragement through the course of my Graduate studies.

This work was supported by the National Science Foundation.

DEDICATION

To my wife, Maureen

TABLE OF CONTENTS

TABLE OF CONTENTS (Concluded)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Continued)

LIST OF ILLUSTRATIONS (Continued)

LIST OF ILLUSTRATIONS (Concluded)

SUMMARY

Two-dimensional optical systems provide a second degree of freedom when processing one-dimensional information. This second degree of freedom can be exploited to increase either the information throughput of the processor or the kinds of operations performed by the optical system. In this thesis, the second degree of freedom is used to perform signal analysis and signal waveform processing operations that are frequency dependent.

Conventional coherent optical spectrum analyzers are characterized by a linear display of signal spectral components and uniform resolution over the analysis bandwidth. A general frequency-variant spectrum analyzer is developed with the capability of mapping signal spectral components along an arbitrary frequency axis and displaying these components with frequency dependent resolution characteristics. Log-frequency and log-frequency constant proportional bandwidth spectrum analyses are implemented with this system. A general description of the space-bandwidth product of the optical system is also developed applicable to the frequency-variant system.

The extension of the basic methods developed for implementing a frequency-variant spectrum analysis to the implementation of general one-dimensional linear operations is considered. A clarification of the two fundamental optical implementations of one-dimensional linear operations is presented.

A general optical system for processing temporal signal waveforms is developed. In this system, the frequency-variant spectrum analyzer

system is used in conjunction with an interferometric optical heterodyne processor. The capability of the frequency-variant analyzer to control the position and resolution of signal components allows for the very general mixing of a signal and a local oscillator. This mixing results in a frequency dependent redistribution of the signal components useful for bandwidth compression and expansion or for more general spread-spectrum signaling. The basic procedure for interpreting the operation of this system in terms of time-frequency signal components is described.

The practical problems associated with the experimental implementation of the signal waveform processor are considered. A channelized version of this system is developed greatly simplifying the input-output relationships of the system while retaining most of the operational generality of the non-channelized processor. The problems of channel crosstalk and channel independence in the channelized system are considered and the maximum number of channels possible in a system for a given degree of channel independence is determined. Common component optical processing configurations are also developed to eliminate the need for interferometric quality optical components and system isolation form vibrations. Finally, the basic operational capabilities of the processor are demonstrated.

CHAPTER I

INTRODUCTION

The application of Fourier methods and linear systems theory to optical problems was introduced in the late 1940's by Duffieux [1] and by Schade [2]. However, it was from the work of Elias, Grey, and Robinson [3], Rhodes [4], Marechal and Croce [5], and O'Neill [6], all in the early 1950's, that an extensive theory for optical systems based on communications theory has developed. The impact of this new theory - now generally referred to as Fourier optics theory - was extensive, changing not only the way in which optical systems were analyzed but the very ideas about what kinds of operations an optical system is capable of performing. Many of the predictions concerning the consequences of a blending of optical theory and communications theory were extravagant (for example, Elias foresaw the presence of independent base and treble control knobs on a camera lens [7]). Nevertheless, optical systems were correctly recognized to be information processing systems characterized by high speed of operation and the ability to handle large amounts of information in parallel.

Much of the original enthusiasm for the capabilities of optical processing systems (and, indeed, for all analog computing systems) died as the power and processing capabilities of digital computers increased. Yet, even using the present generation of high speed digital computers, there are processing operations, readily performable using optical systems, for which no fast digital algorithms have been developed. Recent interest in hybrid optical/digital system represents a willingness on the part of

the information processing community to sacrifice some of the extreme accuracy provided by the digital computer in order to gain the advantages of the very high speed and parallel processing capabilities of optical systems.

One class of operations for which optical systems can provide an alternative to digital processing is linear shift-variant operations. Examples of special cases in this class for which fast digital algorithms have been developed include the Fourier transform operation [8] and certain frequency-warping spectrum analysis operations [9]. However, a general-purpose high-speed algorithm for linear shift-variant operations is not presently available. A general-purpose optical technique for performing operations in this class on two-dimensional (2-D) information using generalized optical components has been described by Bryngdahl [10]. Unfortunately, this technique suffers severly from limitations imposed on the total amount of information that can be processed (there are severe bandwidth limitations using this method) [11,12]. If, however, we restrict our attention to strictly one-dimensional (1-D) information, a distinct advantage inherent in the optical systems makes itself evident. Specifically, the optical system is two-dimensional, and the second dimension can be exploited to provide an additional degree of freedom when processing 1-D information. In the work presented here, we investigate the use of this second degree of freedom to perform linear 1-D shift-variant operations. The range of problems that could be considered in this class is quite broad. We therefore restrict our attention, primarily, to frequency-variant systems (a sub-class of linear shift-variant systems that we define in this thesis). We develop the basic optical capacity for frequency-

variant processing of 1-D spatial signals, and extend these results to the frequency-variant processing of temporal signal waveforms.

The important results of this research lie in six main areas:

(1) the development of a frequency-variant spectrum analyzer;

(2) a clarification of the two fundamental methods used in optical systems to perform general 1-D linear operations;

(3) the development of a general frequency-variant optical processing system for temporal signal waveforms;

(4) a general interpretation of the operation of the signal waveform processor in both time and frequency;

(5) specification of a practical channelized version of the signal waveform processor; and

(6) the experimental demonstration of the basic processing capabilities of the channelized system.

CHAPTER II

FREQUENCY-VARIANT OPTICAL SPECTRUM ANALYSIS

Papoulis defines a spectrum analyzer as a physical system measuring the Fourier spectrum of an arbitrary signal [13]. Thus ideally, a coherent optical spectrum analyzer is an optical computer used to evaluate the Fourier transform integral

$$F(\nu) = \int_{-\infty}^{\infty} f(x) e^{-i2\pi\nu x} dx \quad , \tag{2-1}$$

where f(x) is the arbitrary signal input. Taking into account the finite extent of any actual optical system, it is clear that such a system can evaluate an integral only over finite limits. The output of an optical spectrum analyzer would therefore have the form

$$F(\nu) = \int_{-\infty}^{\infty} f(x) w(x) e^{-2\pi\nu x} dx \quad , \tag{2-2}$$

where w(x) is some windowing function that limits the input to a finite extent. Thus the output of an optical spectrum analyzer is the equivalent of the short-time analysis performed by electronic or digital spectrum analyzers.

We begin this chapter with a brief survey of the operation and essential characteristics of conventional optical spectrum analyzers for one-dimensional (1-D) signals. We discuss some of the important advantages and limitations of these systems. We then describe the extension

of optical spectrum analyzer capabilities through utilization of the two-dimensionality of the systems. In particular, we describe a class of spectrum analyzers for 1-D signals that exploit the additional degree of freedom available in the optical system to perform operations that are frequency-variant.

In the discussions below, we assume that the input signal, initially a function of time, is introduced into the optical system as 1-D spatial variations of a light amplitude distribution. One method of implementing this input operation is first to record the input signal as a density modulated soundtrack and process the film for light amplitude transmittance variations proportional to the recorded signal. Illuminating this soundtrack with a plane wave then produces the input light amplitude distribution. We will often discuss the corresponding time and temporal frequency characteristics of the recorded signal rather than the spatial or spatial frequency characteristics. Real-time operation of the analyzer systems is possible if the input recording is replaced with one of several real-time input transducers available. Radio-frequency signals, for example, can be input directly using a Bragg cell [14].

Conventional Optical Spectrum Analyzers

In a coherent optical system, there exists a two-dimensional (2-D) Fourier transform relationship between the light amplitude distributions in the front and back focal planes of a simple thin lens. If we consider the single lens system shown in Fig. 2-1, the light amplitude distributions in planes P_1 and P_2, the front and back focal planes of the lens L, respectively, are related by the expression [15]

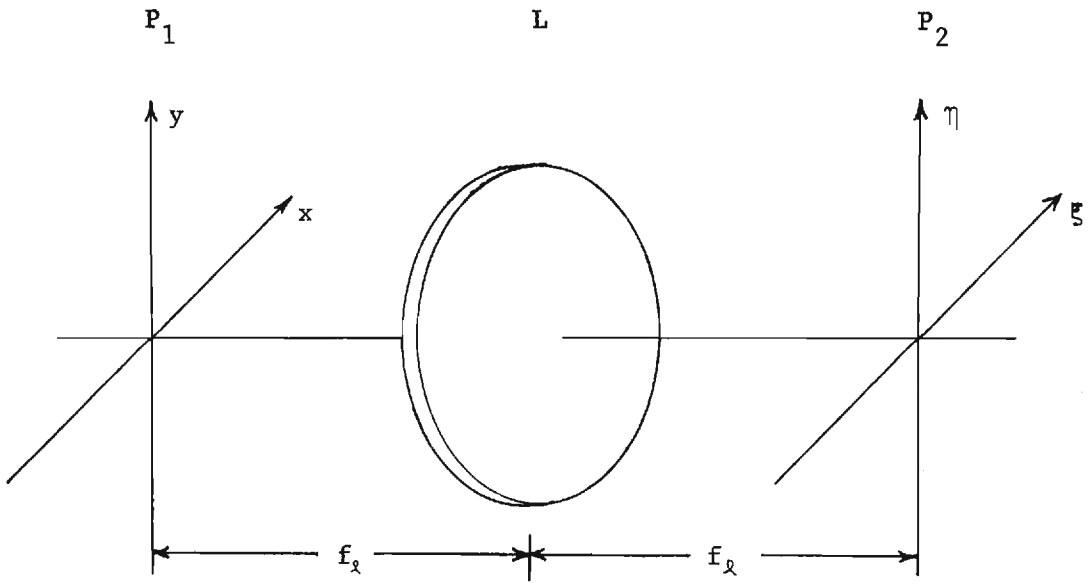


Figure 2-1. 2-D Fourier Transform System.

$$U_2(\xi,\eta) = \bar{A} \iint_{\sum} U_1(x,y) e^{-i2\pi\left(\frac{\xi}{\lambda f_\ell} x + \frac{\eta}{\lambda f_\ell} y\right)} dxdy \quad , \tag{2-3}$$

where λ is the light wavelength and f_ℓ is the focal length of the lens, and $\bar{A}$ is a complex constant. The integral in Eq. (2-3) is evaluated over the finite limits $\sum$ representing the area of the limiting aperture placed in P_1. The extent and limiting nature of this input aperture can be incorporated in a windowing function, $w(x,y)$, and the output rewritten as

$$U_2(\xi,\eta) = \bar{A} \int_{-\infty}^{\infty}\int U_1(x,y)w(x,y) e^{-i2\pi\left(\frac{\xi}{\lambda f_\ell} x + \frac{\eta}{\lambda f_\ell} y\right)} dxdy \quad . \tag{2-4}$$

The light amplitude distribution in P_2 is therefore expressible as the 2-D Fourier integral transform of the product function $U_1(x,y)w(x,y)$ evaluated at the frequencies $\nu = \xi/\lambda f_\ell$ and $\mu = \eta/\lambda f_\ell$.

Our primary concern is the spectral analysis of 1-D signals. We therefore consider the input light amplitude distribution in P_1 to be of the form

$$U_1(x,y) = f(x)1(y) \quad . \tag{2-5}$$

The function, $1(\cdot)$, is a constant function with value unity for all values of its argument. This function will be used to indicate explicitly the two-dimensional nature of distributions that have variations in only one dimension. With this input distribution, the output amplitude distribution is

$$U_2(\xi,\eta) = \bar{A} \iint_{-\infty}^{\infty} f(x)l(y)w(x,y)e^{-i2\pi\left(\frac{\xi}{\lambda f_\ell}x + \frac{\eta}{\lambda f_\ell}y\right)}\,dxdy \quad . \tag{2-6}$$

If we assume $w(x,y)$ to be separable in x and y, i.e., $w(x,y) = w_x(x)w_y(y)$, then the integral is separable:

$$U_2(\xi,\eta) = \bar{A} \int_{-\infty}^{\infty} l(y)w_y(y)e^{-i2\pi \frac{\eta}{\lambda f_\ell}y} \cdot \int_{-\infty}^{\infty} f(x)w_x(x)e^{-i2\pi \frac{\xi}{\lambda f_\ell}x}\,dxdy \quad , \tag{2-7}$$

or

$$U_2(\xi,\eta) = \bar{A}W_y(\eta/\lambda f_\ell)\int_{-\infty}^{\infty} f(x)w_x(x)e^{-i2\pi \frac{\xi}{\lambda f_\ell}x}\,dx \quad , \tag{2-8}$$

where $W_y(\cdot)$ is the Fourier transform of $w_y(\cdot)$. In most cases of interest, $w_y(y)$ will be a wide rectangle function, and most of the light in the output plane will be concentrated along the ξ-axis. Therefore, viewing the output distribution only along the ξ-axis, we have

$$U_2(\xi,0) = \bar{A}'\int_{-\infty}^{\infty} f(x)w_x(x)e^{-i2\pi \frac{\xi}{\lambda f_\ell}x}\,dx \quad , \tag{2-9}$$

where

$$\bar{A}' = \bar{A}W_y(0) \quad . \tag{2-10}$$

For convenience, we assume that the output axis is rescaled in terms of the frequency variable $\nu = \xi/\lambda f_\ell$ (this is equivalent to assuming unit

values for the light wavelength and lens focal length). Equation (2-9) can then be simplified to

$$U_2(\nu,0) = \bar{A}'\int_{-\infty}^{\infty} f(x)w_x(x)e^{-i2\pi\nu x}dx \ , \tag{2-11}$$

or

$$U_2(\nu,0) = \bar{A}'\mathcal{F}_x\{f(x)w_x(x)\} \ , \tag{2-12}$$

where $\mathcal{F}_x\{\cdot\}$ is the 1-D Fourier transform operator representing the integral of Eq. (2-11).

The single lens system of Figure 2-1 performs the short-time spectrum analysis with the result displayed along the ν-axis of the output plane. With this system, as with the other system described below, continuously updated real-time or delayed real-time spectral analysis is possible using the real-time input transducer or a moving soundtrack input. In the analytical expressions, f(x) then becomes f(x+vt), where v is the velocity associated with the input device (e.g., the acoustic propagation velocity in a Bragg cell). One distinct characteristic of the optical spectrum analyzer is the simultaneous nature of the output--the system acts as a bank of parallel filters displaying the output for all frequencies simultaneously. This is to be compared with sequential or scanning electronic systems where the frequency components are read out sequentially as the system scans through the analysis bandwidth.

A measure of the quality of the optical analyzer, for comparison with the electronic systems, is the time-bandwidth product (TW) of the

system. For a spectrum analyzer, the TW can be defined as the maximum number of resolvable frequency components present in the output. In a Fourier analysis over a signal duration Δt, frequency components spaced closer than $\delta\nu = 1/\Delta t$ cannot be resolved. If the analyzer performs the analysis over the bandwidth $\Delta\nu$, then the maximum number of resolvable frequency components in the analysis is

$$TW = \frac{\Delta\nu}{1/\Delta t} = \Delta t \Delta\nu \quad . \tag{2-13}$$

The analogous quantity in an optical system is the space-bandwidth product or SW. For the optical analyzer, the length of signal under analysis is determined by the width of the input aperture, Δx. The maximum workable bandwidth of an optical system is limited by vignetting; the situation for a single lens analyzer is shown in Fig. 2-2. The maximum angle through which light passing through the input aperture can be diffracted and be totally collected by lens is approximated as [16]

$$\theta = \tan^{-1}\left(\frac{d-\Delta x}{2f_\ell}\right) \simeq \frac{d-\Delta x}{2f_\ell} \quad , \tag{2-14}$$

where d is the lens diameter. Light diffracted through this angle represents components of the input with spatial frequency

$$\nu_m = \frac{\sin\theta}{\lambda} = \frac{\sin\left[\tan^{-1}\left(\frac{d-\Delta x}{2f_\ell}\right)\right]}{\lambda} \simeq \frac{d-\Delta x}{2\lambda f_\ell} \quad . \tag{2-15}$$

The quantity ν_m is therefore the maximum frequency that the single lens system can analyze in a signal of length Δx. This is equal to the analysis

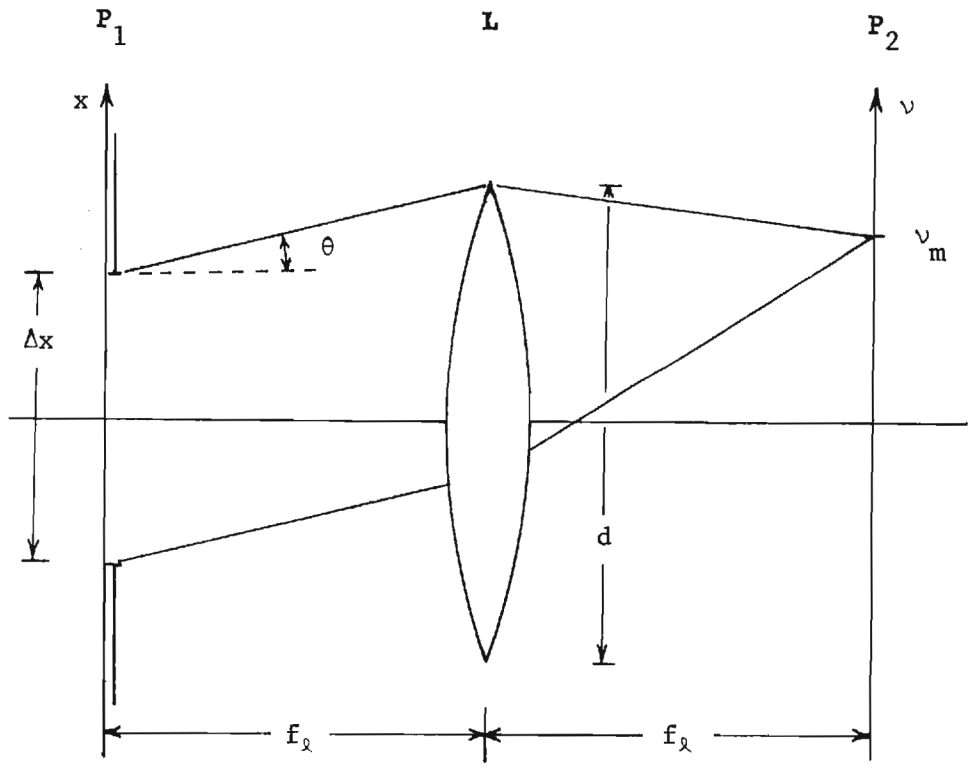


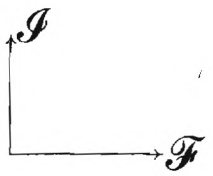
Figure 2-2. Vignetting in a Single Lens Fourier Transform System.

bandwidth of the system $\Delta\nu = \nu_m$. The system SW is then given by

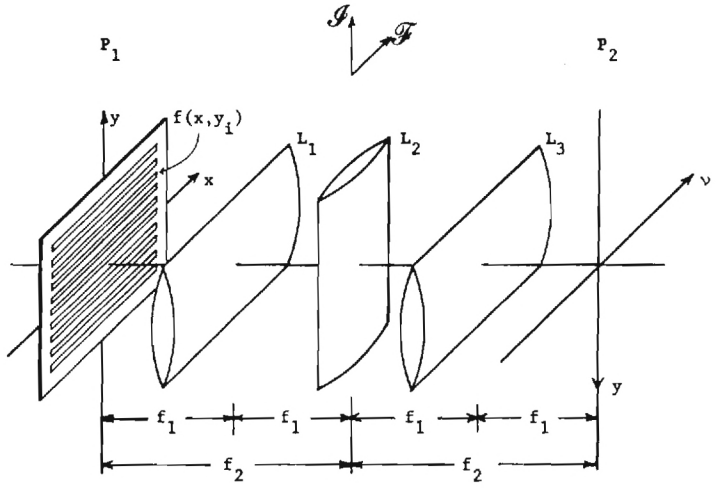
$$SW = \Delta\nu\Delta x = \frac{\Delta x}{\lambda} \sin[\tan^{-1}(\frac{d-\Delta x}{\lambda f_\ell})] \simeq \frac{\Delta x(d-\Delta x)}{\lambda^2 f_\ell} \quad . \tag{2-16}$$

Typical values of the SW for optical analyzers made with good quality optical components range from 10^3 to 10^4. These high values of SW provide the optical analyzer with the ability to achieve high frequency resolution for a given input signal bandwidth.

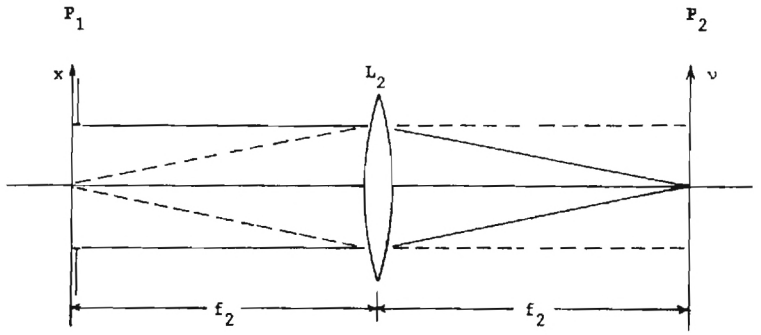
Perhaps the most important aspect of the optical spectrum analyzer besides the high resolution capabilities is the availability of a second dimension. Cutrona et al. demonstrated how the additional degree of freedom provided by the second dimension could be used to perform multichannel spectrum analysis [17]. The optical configuration of a multichannel analyzer is shown in Fig. 2-3(a). This system is seen to be fundamentally one that Fourier transforms in the horizontal direction while imaging in the vertical direction. This operation is indicated in the figure by the symbol



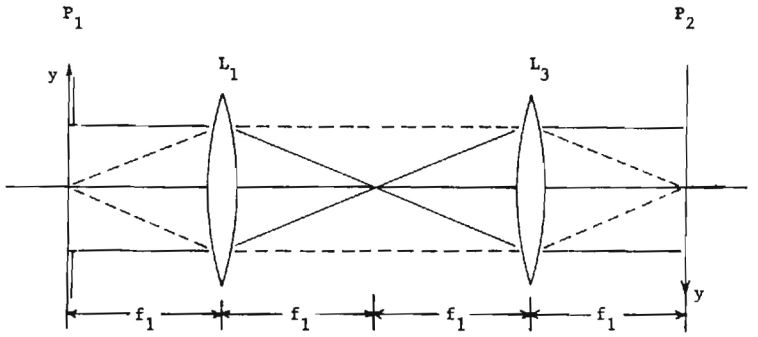
which represents the directions of operation for the imaging operator, $\mathcal{I}\{\cdot\}$, and the Fourier transform operator. A particularly easy method to determine the operation of the system is a ray-tracing method suggested by Mr. Leonard J. Laub. Using this method, we view the operation in the horizontal and vertical directions separately as shown in Figs. 2-3(b)



(a) Multichannel System



(b) Horizontal



(c) Vertical

Figure 2-3. Multichannel Spectrum Analyzer; (a) Optical Configuration, (b) System Operation in the Horizontal Direction, (c) System Operation in the Veritcal Direction.

and (c). In the horizontal direction, Fig. 2-3(b), the lenses L_1 and L_3 have no effect (they are cylindrical lenses with curvature oriented in the vertical direction). The lens L_2 then maps points (δ-functions) in plane P_1 into plane waves (linear phase factors) in plane P_2 and vice versa. This is the mapping that occurs when a Fourier transform operation is performed by the lens system. In the vertical direction, Fig. 2-3(c), the lens L_2 has no effect; lenses L_1 and L_3 map points in P_1 into points in P_2 and plane waves in P_1 into plane waves in P_2. This mapping is the imaging operation. Returning to Fig. 2-3(a), the input transparency introduces N one-dimensional signals, $f(x,y_i)$, i=1,N, where the signals are separated and equally spaced at the vertical positions y_i. The input is again limited in the horizontal direction by the windowing aperture so that the input amplitude distribution at y_i is

$$U_1(x,y_i) = f(x,y_i)w_x(x) \quad . \tag{2-17}$$

With the vertical imaging/horizontal Fourier transforming operation of the system, the output at the position y_i is

$$U_2(\upsilon,y_i) = \mathcal{F}_x\{f(x,y_i)w_x(x)\} \quad . \tag{2-18}$$

The maximum number of signals that can be analyzed is limited by the resolving power of the vertical imaging section of the system. This number is again equal to the SW of the imaging system with values of 10^3 to 10^4 possible. In the multichannel system, it is therefore possible to analyze 1000 signals simultaneously with each analysis having a SW of 10^3.

In another application of the second degree of freedom when processing 1-D signals, Thomas has shown that it is possible to use the entire 2-D space bandwidth product of the optical system in analyzing a single 1-D signal [18]. By using a raster recording of the 1-D input signal, as shown in Fig. 2-4, a much longer portion of the input signal can be analyzed than with the density modulated soundtrack. Exactly how much longer again depends on the SW for the lens--the individual raster lines must be resolvable. Using the value of 10^3 as a typical 1-D SW, approximately 1000 times more signal is present to be analyzed with the raster technique. Thomas described the technique for properly interpreting the output distribution, which may contain up to 10^6-10^7 individually resolvable spectral components. This method has made possible the application of optical spectrum analyzers in high resolution analysis of broadband signals [18-21].

As can be seen from the above discussions, the historical application of the second degree of freedom in spectrum analyzers has been one of increasing the number of signals or the amount of signal being processed. In other words, the second degree of freedom has been used to increase the amount of information throughput of the optical analyzer. However, the inflexibility of these conventional systems is still apparent: the analysis performed produces a spectral display along a linear frequency axis and with fixed resolution characteristics. It is often convenient to perform alternative forms of spectral analysis where a non-linear relationship exists between the position within the spectral display and the frequency of the component displayed at that position. An example of this type of analysis is a log-frequency spectrum analysis for signals that have

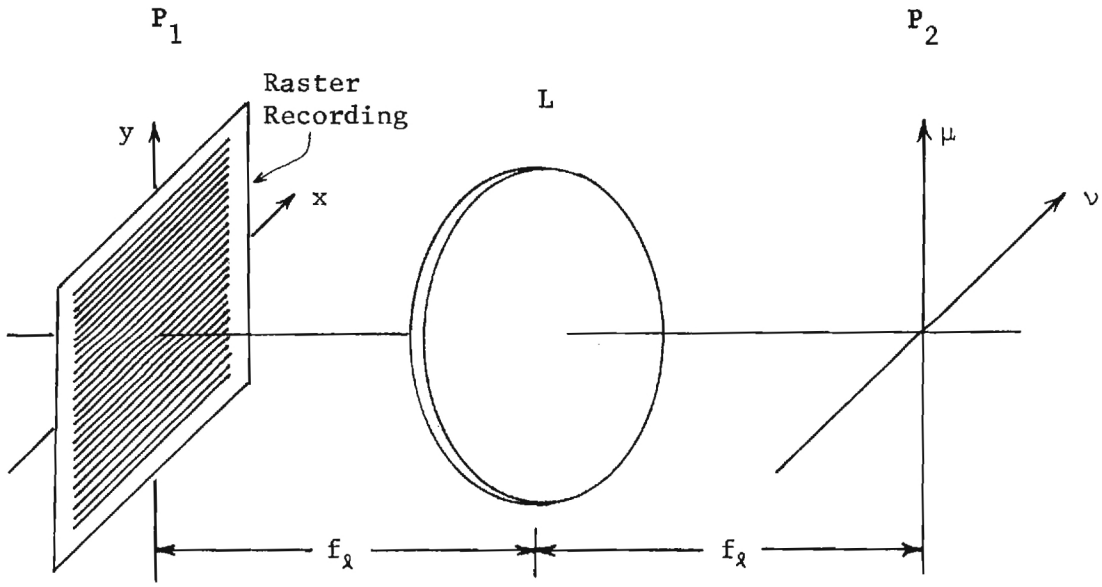


Figure 2-4. System for Spectrum Analysis of Broadband Signals.

undergone a significant Doppler shift. Another useful form of spectral analysis is one in which different frequency components are displayed with different resolution characteristics. Such a system would be useful in the analysis of speech, where good frequency resolution may be required in the low frequency region of voiced phonemes while good time resolution is desirable in the higher frequency region of stops and plosives [22].

In the remainder of this chapter, we describe how the second degree of freedom available in optical systems can be used to increase the kinds of operations that can be performed optically in the analysis of 1-D signals.

The Variable Resolution Spectrum Analyzer

In this section, we describe an optical spectrum analyzer characterized by continuously varying time and frequency resolution. The basic optical configuration, shown in Fig. 2-5, is that of the vertical imaging-horizontal Fourier transforming system used in the multichannel analysis. The distinguishing features in this case are the absence of multiple input signals and the use of a special input plane mask. Here one signal serves as the input and the input aperture is a non-separable function of x and y:

$$U_1(x,y) = f(x)w_t(x,y) \quad . \tag{2-19}$$

The resulting output distribution in the output plane is

$$U_2(\nu,y) = \mathcal{F}_x\{U_1(x,y)\} \quad , \tag{2-20}$$

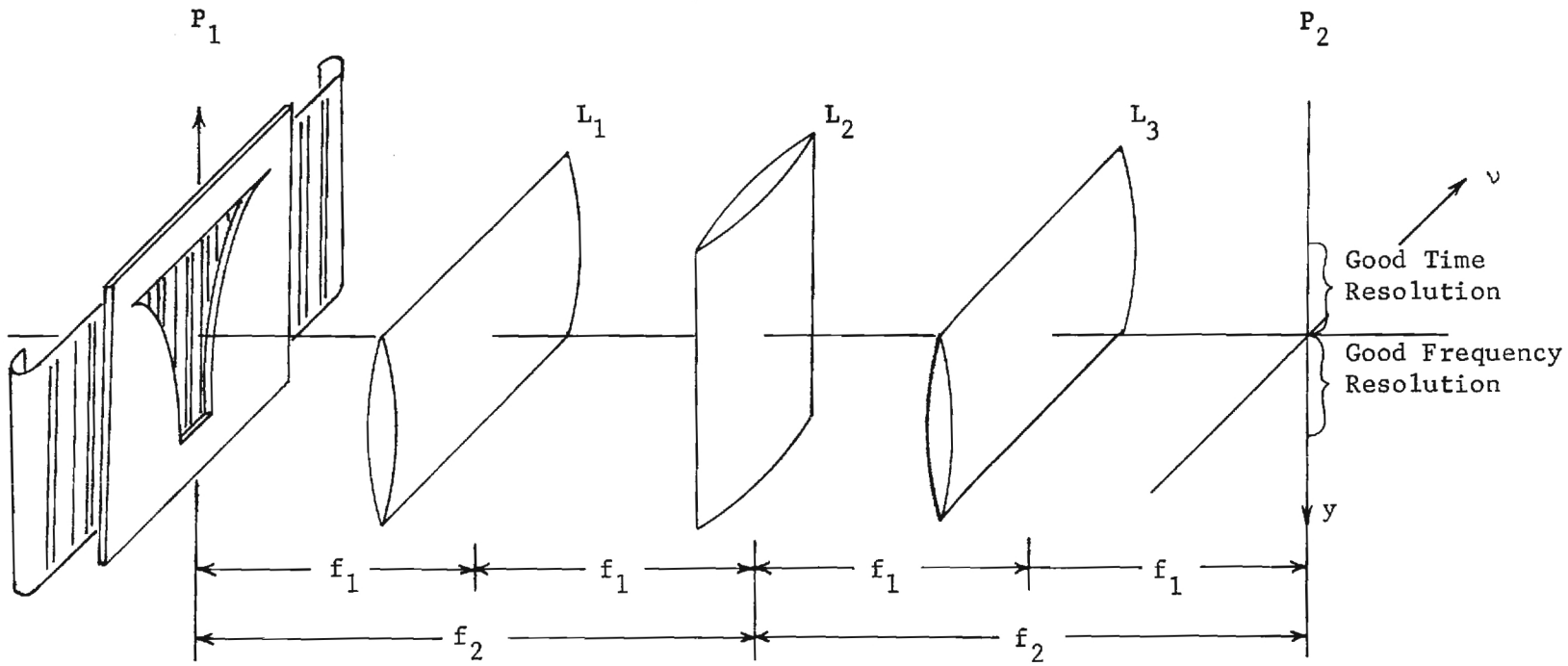


Figure 2-5. Variable Resolution Spectrum Analyzer.

where, for simplicity, we have ignored the proportionality constants and scale factors. Substituting Eq. (2-19) for $U_1(x,y)$, we obtain the output plane distribution

$$U_2(\nu,y) = \int_{-\infty}^{\infty} F(u)W_t(\nu-u,y)du \quad , \tag{2-21}$$

where

$$F(\nu) = \mathcal{F}_x\{f(x)\} \quad , \tag{2-22}$$

and

$$W_t(\nu,y) = \mathcal{F}_x\{w_t(x,y)\} \quad . \tag{2-23}$$

The convolution integral of Eq. (2-21) represents a smoothing of the signal transform $F(\nu)$, where the nature of the smoothing function depends on the value of y. The result is an output distribution that represents the spectral content of the input signal for continuously varying degrees of time and frequency resolution. Referring again to Fig. 2-5, toward the top appears the spectrum of a long segment of the signal (top because of the imaging inversion): frequency resolution is relatively good; temporal resolution relatively poor. At the bottom, the converse is true: temporal resolution is good; frequency resolution is poor. Such a system can be useful in the analysis of speech where a tradeoff is usually made between time and frequency resolution. The variable window spectrum analyzer can provide simultaneously a convenient look at the

spectral content of input speech segments for a wide range of signal durations.

As an example, let the input signal be a single cissoid

$$f(x) = e^{i2\pi\nu_o x} \quad , \tag{2-24}$$

(representing, for example, the positive frequency component of a sinusoid recorded on film). The resulting output distribution is then

$$U_2(\nu,y) = \int_{-\infty}^{\infty} \delta(u-\nu_o) W_t(\nu-u,y)\,du$$

$$= W_t(\nu-\nu_o,y) \quad , \tag{2-25}$$

which, for any particular value y, is simply a displaced version of the transform of the corresponding cross section window profile. Detailed characteristics of the input plane mask's cross sectional profile and its variation with y are a matter of choice. Of special interest is the case where the window is a function of the product xy, i.e.,

$$w_t(x,y) = w_t(xy) \quad . \tag{2-26}$$

This form would be chosen for a display of spectral content with constant proportional bandwidth. The output in this case assumes the form

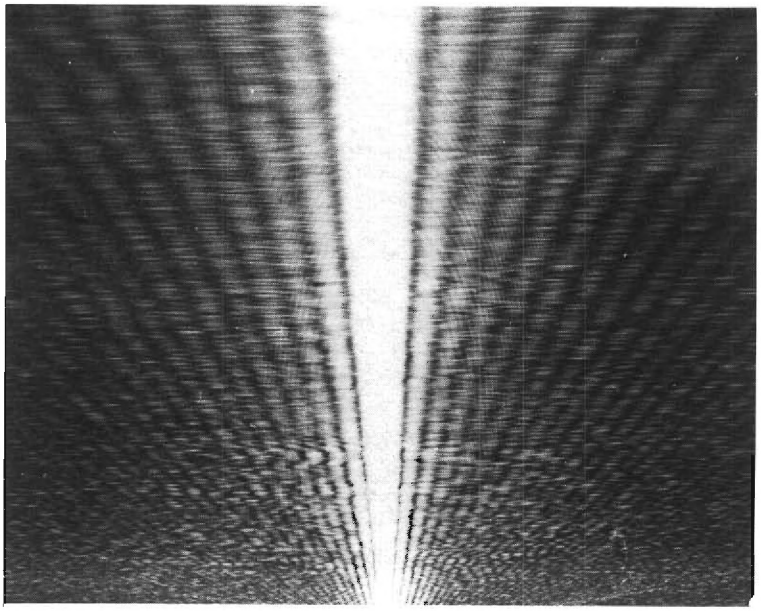
$$U_2(\nu,y) = \frac{1}{y} W_t\left(\frac{\nu-\nu_o}{y}\right) \quad . \tag{2-27}$$

The width of the input window, and therefore the time resolution, is proportional to 1/y; the width of the resulting output display, and therefore the frequency resolution, is proportional to y. In some applications a more general resolution vs. y relationship may be desirable. $w_t(x,y)$ can then be specified by $w_t[x/\Delta(y)]$, where $\Delta(y)$ specifies the window width for a given y. Binary windows of this type are easily fabricated. A tapered window, e.g., one with a Gaussian transmittance profile, can reduce or eliminate spectral component side lobes.

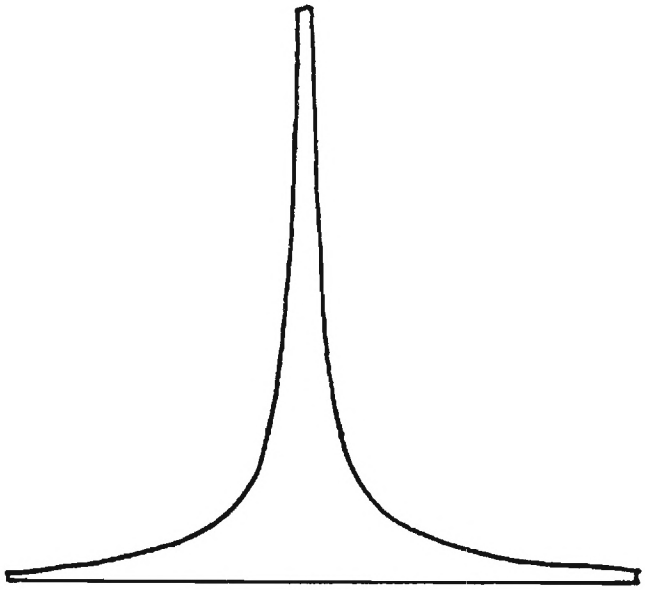
In Fig. 2-6, output spectra obtained experimentally using a variable resolution spectrum analyzer are shown. The appearance of a single spectral line at the output is shown in Fig. 2-6(a) when the input plane window is that shown in Fig. 2-6(b). The window is binary with a hyperbolic (i.e., 1/y) shape; it produces an output intensity distribution that has a $\text{sinc}^2(\cdot)$ cross section and a spectral width proportional to y. With a square wave (Ronchi ruling) input signal placed behind the 1/y window as shown in Fig. 2-6(d), the multiple line spectrum shown in Fig. 2-6(c) was obtained. Emulsion thickness variations were not compensated, with some resultant irregularities in the output spectral patterns.

The Frequency Mapping Spectrum Analyzer

The optical system illustrated in Fig. 2-7(a) is capable of displaying the spectral content of the input signal vs. an arbitrary function of frequency. As indicated in the ray tracing diagrams for the horizontal and vertical directions, Fig. 2-7(b) and (c), respectively, the lens combinations between planes P_1 and P_2 and between planes P_2 and

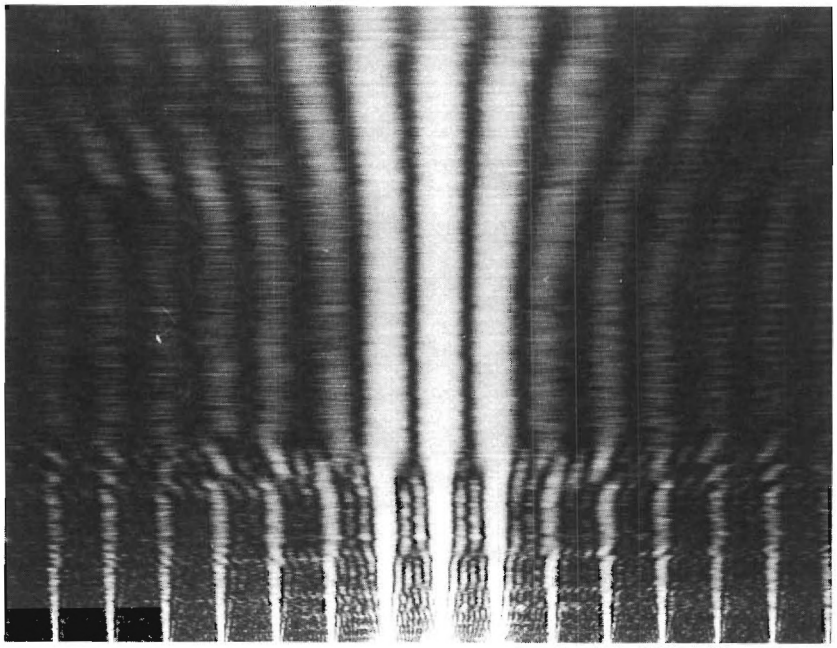


(a)

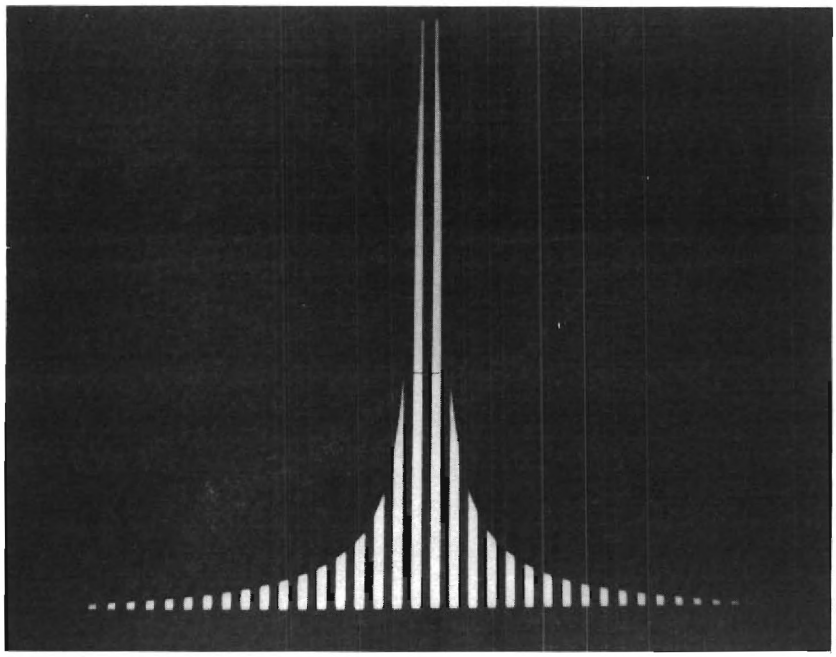


(b)

Figure 2-6. Output of Variable Resolution Spectrum Analyzer; (a) Single Spectral Component, (b) Input Window.

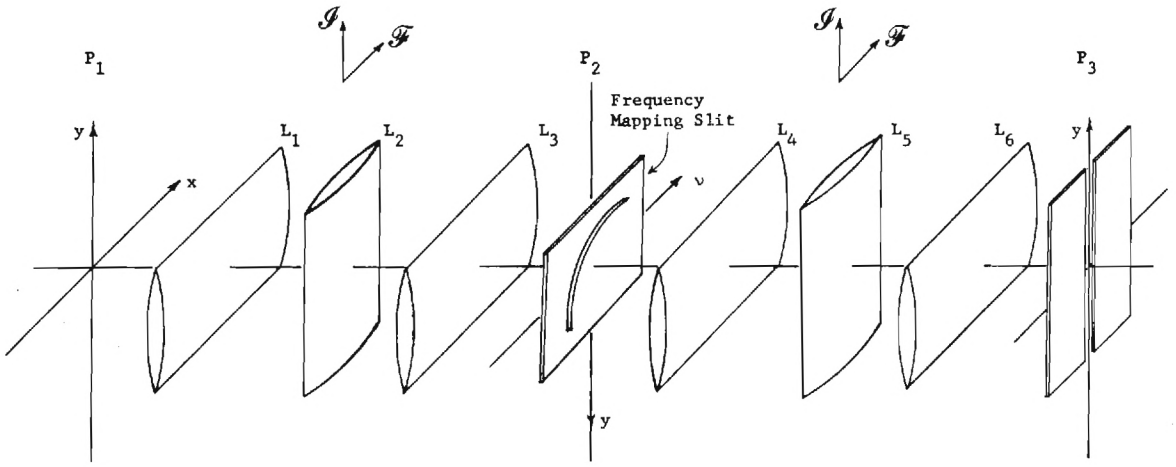


(c)

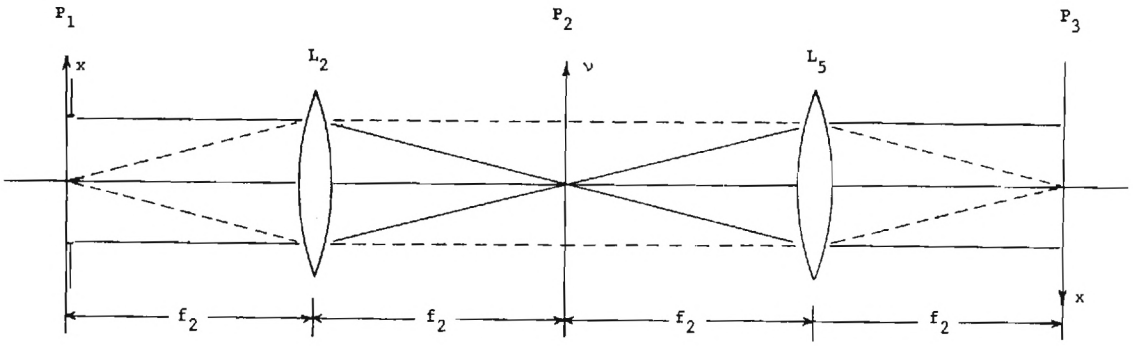


(d)

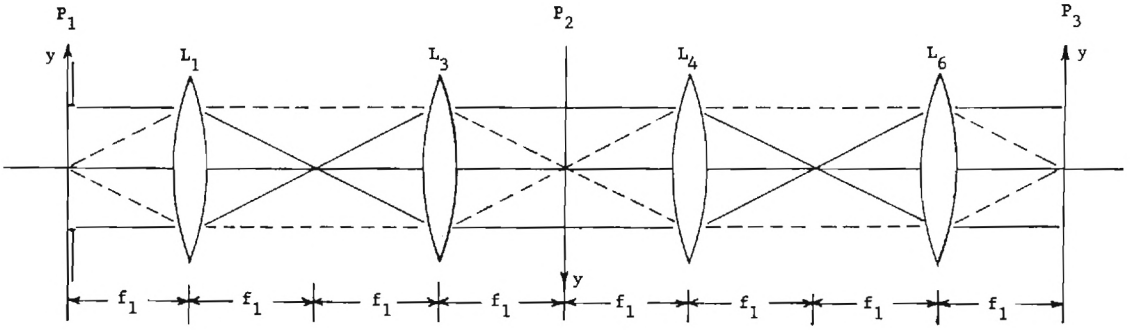
Figure 2-6(cont.) Output of Variable Resolution Spectrum Analyzer (c) Spectrum of Square Wave Input, (d) Windowed Input.



(a) Frequency Mapping Spectrum Analyzer



(b) Horizontal



(c) Vertical

Figure 2-7. Frequency Mapping Spectrum Analyzer; (a) Optical Configuration, (b) System Operation in the Horizontal Direction, (c) System Operation in the Vertical Direction.

P_3 both image in the vertical direction while Fourier transforming in the horizontal direction. The overall system operation with no transform plane mask is to image the input plane onto the output plane. This frequency-mapping spectrum analyzer is actually a special case of the more general frequency-variant analyzer described in the next section. We will therefore defer an analytical description of this system to the next section to avoid duplication. However, to facilitate a better understanding of the general analyzer operation, we include here a qualitative description of the frequency mapping operation.

For convenience of illustration, we assume that the desired output of the analyzer is a display of spectral content vs. log frequency and that the input signal transparency is a recording of sinusoids with frequencies in the proportions 1:2:4:8. To aid in the description of the overall operation of the system, we show several intermediate steps in Fig. 2-8. In the first step, the input signal amplitude distribution shown in Fig. 2-8(a) is imaged in the vertical direction and Fourier transformed in the horizontal direction. The resulting distribution, shown in Fig. 2-8(b), appears in plane P_2 of the system. Placed in this plane is a mask consisting of a narrow logarithmically curved slit. The four positive frequency spectral components associated with the sinusoids are masked as shown in Fig. 2-8(c). In the final stage of the system, a second vertical imaging/horizontal Fourier transform is performed. Each spot of light passing the mask, at its original vertical displacement from the origin (ignoring inversion through the optic axis), is converted into a horizontal smear about the y-axis. Since the light amplitude along that axis is proportional, for each value of y, to

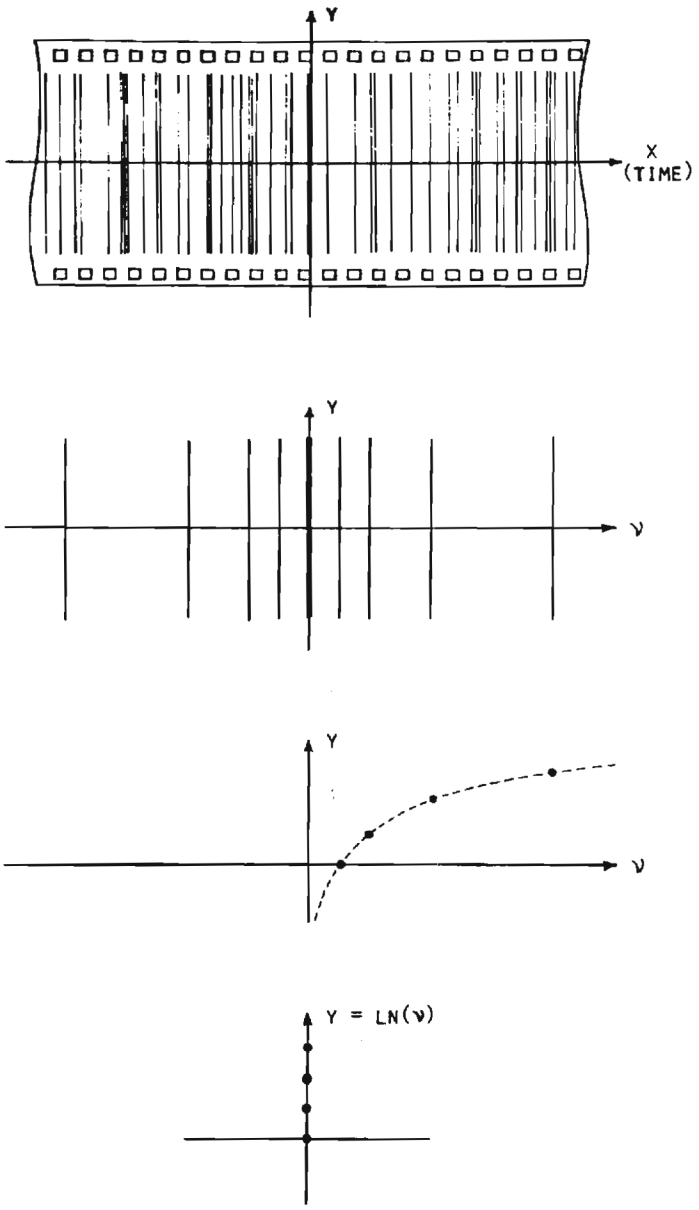


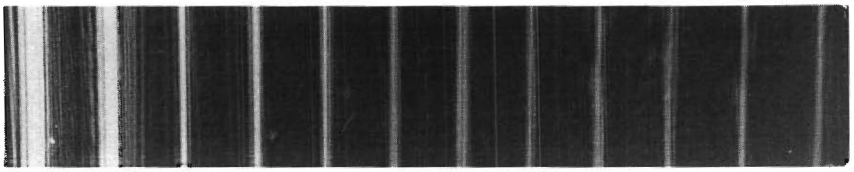
Figure 2-8. Operation of the Frequency Mapping Spectrum Analyzer.

the average value of the light amplitude in P_2, masking off all but a narrow region of the vertical axis in the output plane results in the amplitude distribution shown in Fig. 2-8(d), which displays the spectral content of the input signal vs. log frequency. Variable attenuation in the vertical direction can be introduced by one of the masks to compensate for the greater spot packing with larger values of frequency. It is possible to display signal spectral content vs. many different functions of frequency simply by changing the functional form of the mapping slit in plane P_2. The logarithmic function is useful in many applications, such as the aforementioned analysis of Doppler shifted signals, but other nonlinear mappings may be desired.

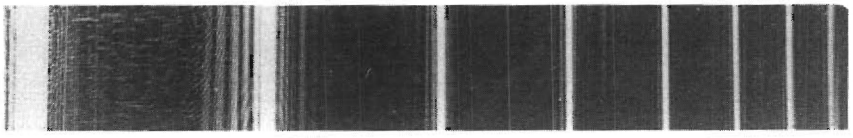
In Fig. 2-9 the output of the frequency mapping spectrum analyzer with a square-wave input is shown. A straight line mapping slit was used for Fig. 2-9(a) resulting in the display along a linear frequency axis. In Fig. 2-9(b), a curved slit was used to get the illustrated results approximately a log frequency mapping.

The Frequency-Variant Spectrum Analyzer

An optical system for general frequency-variant spectrum analysis, combining the essential aspects of the variable resolution analyzer and the frequency-mapping analyzer, is shown in Fig. 2-10. The operation of this system is characteristic of the full power of the two-dimensional optical approach to the spectral analysis of one-dimensional signals. The optical configuration of this system is essentially the same as the frequency-mapping analyzer of Fig. 2-7, but now the input signal transparency is masked as in the variable resolution analyzer of Fig. 2-5.



(a)



(b)

Figure 2-9. Output of Frequency Mapping Spectrum Analyzer;
(a)Output with Straight Line Mapping Slit,
(b)Output with Logarithmic Mapping Slit.

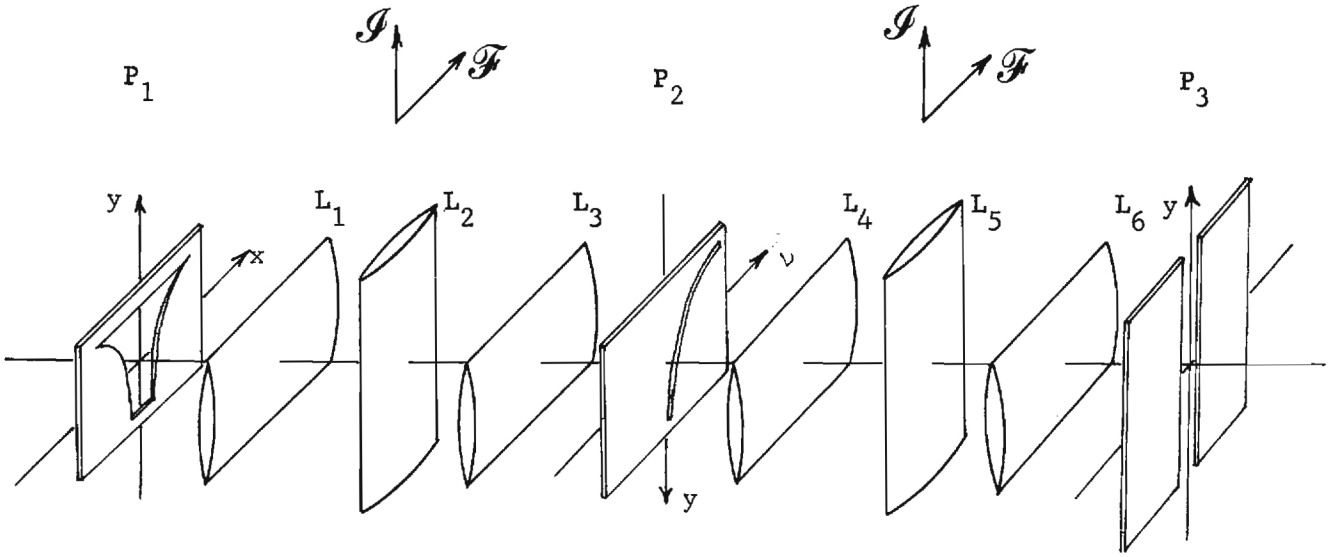


Figure 2-10. Frequency-Variant Spectrum Analyzer.

The characteristics of the slit mask in P_2 thus determines how spectral components are mapped, while the mask in P_1 determines the system response to each frequency component of the input signal.

The input to the frequency-variant analyzer is the same as the variable resolution analyzer, and $U_1(x,y)$ is given by Eq. (2-19). The resulting distribution incident on the frequency-mapping mask in the intermediate plane (plane P_2 of Fig. 2-10) is given by Eq. (2-21), which we rewrite here in the form

$$U_2(\nu,y) = F(\nu)*W_t(\nu,y) \quad , \tag{2-28}$$

where * is understood to denote the one-dimensional convolution operation with respect to the variable ν. Initially, we assume that the mask in the intermediate plane consists of an opaque background with a narrow transparent slit that follows the curve $y=g(\nu)$. In most cases of interest, $g(\nu)$ will be a strictly increasing monotonic function of ν with a unique inverse, $g^{-1}(\cdot)$. For convenience of notation, we define the mapping curve by the pair of equivalent equations

$$y = g(\nu) \tag{2-29}$$

$$\nu = h(y) = g^{-1}(\nu) \quad . \tag{2-30}$$

We assume the mapping slit to have a constant width in the direction of the curve normal. The light amplitude transmittance of the mask can then be represented by the function

$$t_m(\nu,y) = m(y)\delta[\nu-h(y)] \quad , \tag{2-31}$$

where the factor m(y) is given by [23]

$$m(y) = [1 + (\frac{dh}{dy})^2]^{\frac{1}{2}} \quad . \tag{2-32}$$

For a given y, m(y) is proportional to the width of the slit in the horizontal (ν) direction.

The amplitude distribution immediately behind the intermediate plane mask is

$$U_2'(\nu,y) = [F(\nu)*W_t(\nu,y)]m(y)\delta[\nu-h(y)] \quad . \tag{2-33}$$

The final stage of the optical system performs a second vertical imaging/horizontal Fourier transforming operation. Note that the horizontal axis of the output plane has been reversed in direction. We can therefore describe the second Fourier transforming operation as an inverse Fourier transform with the resulting output plane distribution given by

$$\begin{aligned} U_3(x,y) &= \mathcal{F}_\nu^{-1}\{[F(\nu)*W_t(\nu,y)]m(y)\delta[\nu-h(y)]\} \\ &= [f(x)w_t(x,y)]*[m(y)\exp(+i2\pi h(y)x)] \quad , \end{aligned} \tag{2-34}$$

where the convolution operation is with respect to the variable x. The output plane is masked with a narrow vertical slit to observe the distribution only for x=0. Writing out the convolution integral and setting

x=0, we obtain

$$U_3(0,y) = \int_{-\infty}^{\infty} f(\xi) w_t(\xi,y) m(y) e^{-i2\pi h(y)\xi} d\xi \quad . \tag{2-35}$$

Equation (2-35) is seen to have the basic form of the Fourier transform integral. Specifically, for any position y along the vertical axis, $U_3(0,y)$ represents the spectral content of the input signal f(x) at the frequency $\nu=h(y)$, f(x) having been windowed by the function $m(y)w_t(x,y)$. The factor m(y) reflects the greater packing of spectral components along the y-axis in regions where the slope of the mapping slit is small. (If the slit is constructed to have uniform horizontal width as a function of y, m(y) will then be a constant.)

This system can be used in a wide variety of applications requiring a frequency-variant analysis of signal spectral content. If, for example, the input or time window, $w_t(x,y)$, in mask M_1 has a constant cross sectional profile and a width that increases monotonically with y, each resolution cell in plane P_2 represents a different set of values $(\nu,\Delta\nu)$, where ν is the center frequency of the spectral component observed at that point and $\Delta\nu$ is the spectral resolution. If the mapping slit in mask M_2 is a diagonal straight line, spectral components are presented in a linear frequency display, but with controllable resolution. As an example we consider a constant proportional bandwidth (constant-Q) analysis. We assume the mapping slit to be along the straight line $y=\nu$ and to be of constant horizontal width. The inverse function, h(y), for this slit is then simply y. For the constant-Q analysis it is necessary that the frequency resolution be proportional to the frequency, i.e., $\Delta\nu \propto \nu$, and

therefore proportional to y. From our earlier analysis, we know that an input window function of the form $w_t(xy)$ provides a spectral display where frequency resolution is proportional to y. Thus, by using this type of input window and a linear slit, we perform a linear frequency constant Q spectrum analysis.

Alternatively, spectral resolution can be made to change in discrete jumps, say octave by octave. More generally, the functional form of the input window cross sectional profile can change with y, leading to a general frequency dependent response of the system to a cissoidal input. Indeed, the window can be real or complex, symmetric or asymmetric, anticipating or delayed, all (through y) as a function of the frequency of the spectral component observed.

Choice of the mapping slit mask M_2 is also quite flexible. It can be continuous or discontinuous, monotonic or multiple valued. The combination of more general mapping slit curvature and variations in the input window can further extend the system capabilities. As an example, assume we wish to perform a log-frequency spectral analysis and to retain the constant-Q feature described in the previous example. For the log-frequency display, we require that the slit lie along the line $y = \ell n(\nu)$. We therefore have

$$y = g(\nu) = \ell n(\nu) \tag{2-36}$$

and

$$\nu = h(y) = e^{y} \tag{2-37}$$

The question now arises: How do we determine the width of the input window as a function of y to produce a constant-Q analysis? First, we let $w_t(x,y)$ be given by

$$w_t(x,y) = w_t\left[\frac{x}{\Delta(y)}\right] \quad . \tag{2-38}$$

With this input window, the frequency resolution as a function of y is proportional to $1/\Delta(y)$. Then, for the constant-Q analysis, we again must have frequency resolution proportional to frequency, dictating

$$\frac{1}{\Delta(y)} \propto e^{y} \quad , \tag{2-39}$$

or

$$\Delta(y) \propto e^{-y} \quad . \tag{2-40}$$

A similar approach can be taken for other resolution and mapping requirements.

The flexibility in choice of the two system masks can lead to quite general forms of spectral analysis. There are, however, constraints, imposed by the interaction of the two masks, if mask M_2 is to serve principally to map spectral components and not to determine, at least in part, system time and frequency resolution. These constraints can be viewed as a direct consequence of the imaging nature of the overall optical system. It is apparent that as the horizontal width of the mapping slit increases, the display at the corresponding vertical displacement in the output plane

will look more like an image of the input and less like a measure of its spectral content. We present here a qualitative analysis of the necessary constraints, a general analytical treatment follows.

Assume the mapping plane to be divided up into a large number of cells as shown in Fig. 2-11. In the vertical direction, cell size is determined by the resolution of the one-dimensional imaging operation, and is constant throughout the plane. In the horizontal direction, each cell has a width equal to the spectral resolution at that location, being inversely proportional to the width of the corresponding horizontal slice of the input plane mask. We take the simplifying viewpoint that, for a given vertical location, each spectral resolution cell represents an independent sample of the input signal spectral content. The basic restriction on the mapping slit in mask M_2, then, is that for each vertical location the mapping slit must have a horizontal width narrower than the spectral resolution cell it transmits. For example, the mapping slit in Fig. 2-11, which has a constant cross-sectional width, is satisfactory in region (1) but is unsatisfactory in region (2), where, because of its small slope, it transmits several independent spectral components. Light from these cells interferes in the output plane in a manner that depends upon the horizontal position of the input. The result is a false representation of spectral content.

It would appear that the mapping slit should be made as narrow as possible, for the chances of interference between adjacent spectral resolution cells would then be minimized. For very narrow slits, only the dimension of the image resolution cell and the slope of the slit are important in determining whether more than one spectral resolution are

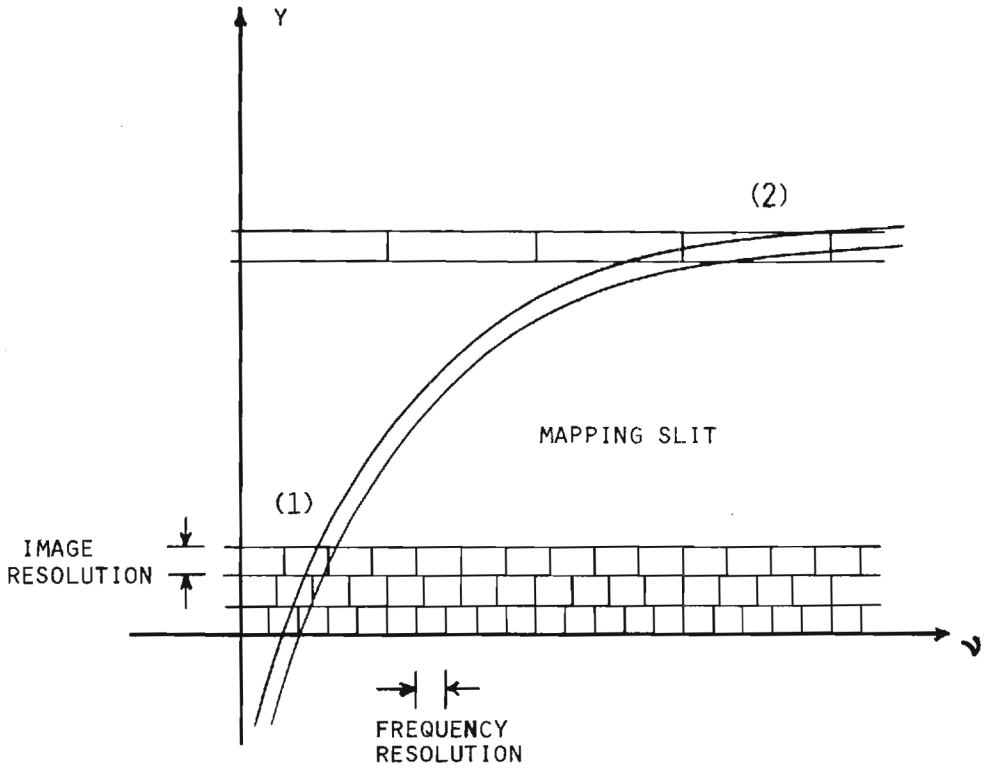


Figure 2-11. Fourier Transform Plane Resolution Cell Diagram.

important in determining whether more than one spectral resolution cell (per image cell) will be transmitted. As the slit becomes narrower, however, output display brightness diminishes. A more economical approach is to choose a mapping slit whose horizontal width at any vertical location equals some constant fraction of the spectral resolution cell at that location. Such a choice still assures insignificant interference from adjacent spectral component cells and at the same time works to equalize the output display brightness for both high resolution and low resolution spectral components.

We must now reconsider the analytical description of the system operation to determine the effect of a finite-width mapping slit on the output distribution. We return to an expression for the light amplitude distribution immediately in front of the intermediate plane mask,

$$U_2(\nu,y) = \mathcal{F}_x\{g(x)w_t(x,y)\} \quad , \tag{2-41}$$

and assume a mapping transmittance function of the form

$$t_m(\nu,y) = W_f(\nu-h(y),y) \quad . \tag{2-42}$$

This is the function $W_f(\nu,y)$ shifted by an amount ν_o, where $\nu_o = h(y)$. The amplitude distribution immediately behind the mask is then

$$U_2'(\nu,y) = \mathcal{F}_x\{f(x)w_t(x,y)\}W_f(\nu-h(y),y) \quad , \tag{2-43}$$

and the distribution in the output plane is

$$U_3(x,y) = \mathcal{F}_\nu^{-1}\{\mathcal{F}_x\{f(x)w_t(x,y)\}W_f(\nu-h(y),y)\}$$

$$= [f(x)w_t(x,y)]*[w_f(x,y)e^{i2\pi h(y)x}] \quad , \tag{2-44}$$

where

$$w_f(x,y) = \mathcal{F}_\nu^{-1}\{W_f(\nu,y)\} \quad , \tag{2-45}$$

and where the convolution is with respect to the variable x. Writing out the integral expression and setting x=0 (again, we observe the output along the y-axis), we obtain

$$U_3(0,y) = \int_{-\infty}^{\infty} f(\xi)w_t(\xi,y)w_f(-\xi,y)e^{-i2\pi h(y)\xi}d\xi \quad . \tag{2-46}$$

The output distribution for a given value y is seen to be the spectral content at frequency $\nu=h(y)$ with time and frequency resolution determined by an <u>equivalent</u> input window function, w(x,y), given by

$$w(x,y) = w_t(x,y)w_f(-x,y) \quad . \tag{2-47}$$

So long as the mapping slit is sufficiently narrow, $w_f(-x,y)$ will be quite broad in the x-direction, and resolution will be determined primarily by the input plane window. If, on the other hand, $W_f(\nu,y)$ becomes too broad in the ν-direction, $w_f(x,y)$ can be approximated by a δ-function in Eq. (2-43) with the result

$$U_3(x,y) \propto \int_{-\infty}^{\infty} f(\xi) w_t(\xi,y)\,\delta(x-\xi)\, e^{i2\pi h(y)(x-\xi)}\, d\xi$$

$$\propto f(x) w_t(x,y) \quad ; \tag{2-48}$$

i.e., the output plane distribution is simply an image of the masked input recording.

Spectra obtained with the frequency variant spectrum analyzer are shown in Figs. 2-12 and 2-13. In both cases the input was a Ronchi ruling giving a square-wave input signal. The output of the system configured for a log frequency, constant Q analysis is shown in Fig. 2-12. Each frequency component appearing in this display is viewed through a window whose length is inversely proportional to frequency. The example of a general analysis shown in Fig. 2-13 further illustrates the flexibility of the operations that can be performed. In this analysis, low frequency components are viewed through a twin aperture window (giving a sinusoidal variation to the spectral cross-section) and with a magnified frequency axis (the slope of the line segment mapping low frequencies is greater than unity). For the high frequency components a single aperture window is used with the frequency mapping resulting in a demagnification (slope less than unity) or compression of the components.

Space-Bandwidth Product for Frequency-Variant Analysis

As mentioned earlier, the measure of quality for a spectrum analyzer is the time- or space-bandwidth product, which measures the maximum number of resolvable frequency components in the analysis bandwidth. For the conventional optical analyzer, it was found that the SW was equal to the product of the analysis bandwidth and the length of the singal under

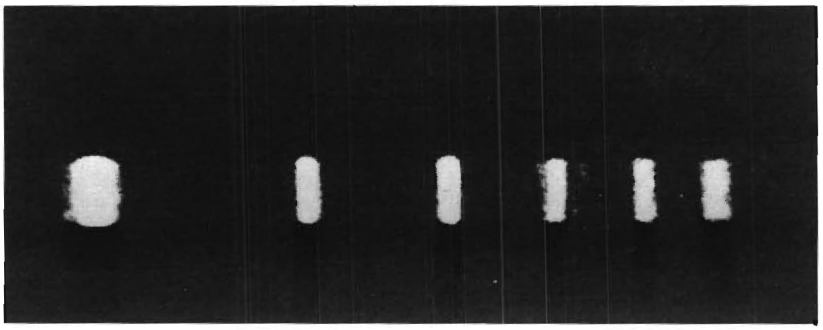


Figure 2-12. Output of Log-Frequency Constant-Q Spectrum Analyzer.

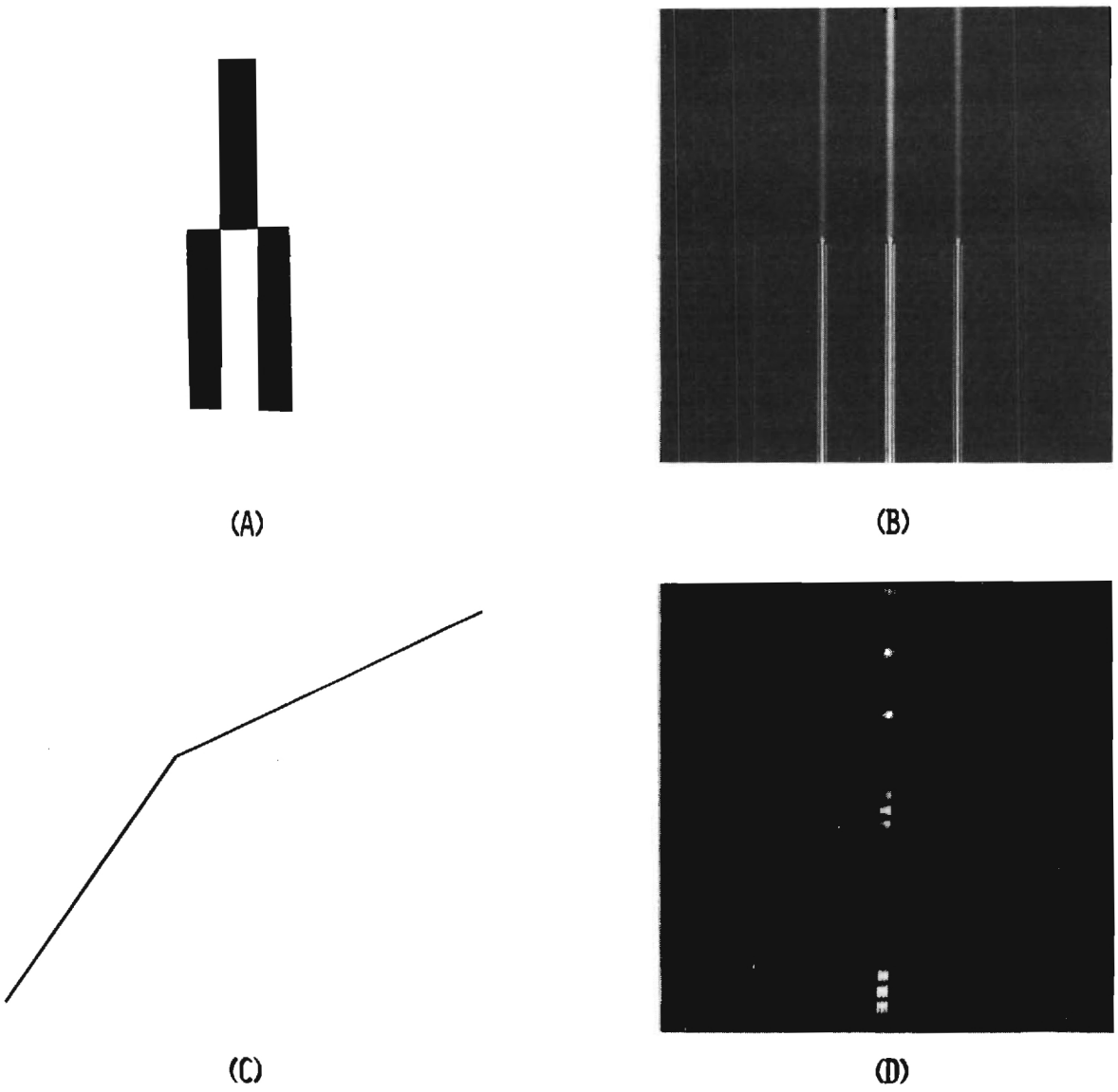


Figure 2-13. General Output of Frequency-Variant Spectrum Analyzer; (a)Input Window, (b) Transform Distribution with Square Wave Input, (c) Mapping Slit, (d) Analyzer Output with Square Wave Input.

analysis. However, in the frequency-variant analyzer, the length of the input signal being analyzed varies as a function of frequency resulting in continuously varying resolution characteristics as a function of frequency. The SW of the system is not given by the simple product, $\Delta x \Delta \nu$, but must be found by actually adding up the number of resolution cells over the total analysis bandwidth. The procedure for calculating the SW (valid for both frequency-variant and frequency-invariant analyzers) is as follows: At the frequency ν_o, the length of signal being analyzed is $\Delta x(\nu_o)$. Therefore, in the narrow band of frequencies $d\nu_o$ about ν_o the frequency components cannot be spaced closer than $1/\Delta x(\nu_o)$ apart. The number of resolution cells in the band $d\nu_o$ is then

$$\frac{d\nu_o}{1/\Delta x(\nu_o)} = \Delta x(\nu_o) d\nu_o \quad . \tag{2-49}$$

Adding up the number of resolution cells over the analysis bandwidth, we obtain

$$SW = \int_{\nu_{min}}^{\nu_{max}} \Delta x(\nu_o) d\nu_o \quad , \tag{2-50}$$

where ν_{min} and ν_{max} are the minimum and maximum frequencies of the analysis.

In the optical frequency-variant spectrum analyzer, the vertical level displaying the component of frequency ν_o is $y_o = g(\nu_o)$. If we assume that the input window has the form of Eq. (2-38), then the length of signal being analyzed at that level is

$$\Delta(y_o) = \Delta(g(\nu_o)) \tag{2-51}$$

giving

$$\Delta x(\nu_o) = \Delta[g(\nu_o)] \tag{2-52}$$

For this analyzer the SW is obtained from Eqs. (2-50) and (2-52) as

$$SW = \int_{\nu_{min}}^{\nu_{max}} \Delta(g(\nu_o))d\nu \quad . \tag{2-53}$$

As an example, consider the SW of an analyzer performing a constant-Q analysis. For the linear-frequency, constant-Q analysis, $\Delta(y_o) = \Delta x/y_o$ (where Δx is the constant of proportionality), and $y_o = \nu_o$. For the log-frequency, constant-Q analysis, $\Delta y_o = \Delta x e^{-y}$ and $y_o = \ell n \nu_o$. In either case, we have

$$\Delta x(\nu_o) = \Delta[g(\nu_o)] = \frac{\Delta x}{\nu_o} \tag{2-54}$$

Substituting Eq. (2-54) into Eq. (2-50), we obtain the SW for a constant Q analysis:

$$SW_{constant\text{-}Q} = \int_{\nu_{min}}^{\nu_{max}} \frac{\Delta x}{\nu_o} d\nu_o = \Delta x[\ell n \nu_{max} - \ell n \nu_{min}] \quad . \tag{2-55}$$

We note that, assuming $\nu_{min} \geq 1$, Eq. (2-55) indicates that a constant Q analysis has substantially fewer resolution cells than a conventional analysis of signal length Δx. However, this is certainly to be expected

in an analysis where resolution cell size increases proportionally with frequency from some (system determined) minimum.

The Holographic Frequency-Variant Spectrum Analyzer

The use of a slit to map the spectral components in the frequency variant analyzer is convenient because it is easy to fabricate and simple to design. However, the slit mapping technique is very inefficient in the utilization of light; only a small fraction of the available light in the system actually passes through the slit. Rhodes has described an alternative technique employing a holographic mapping element, which, although more difficult to fabricate, uses more of the available light [24]. The system configuration for the holographic frequency-variant analyzer is shown in Fig. 2-14. The optical configuration is similar to the slit mapping analyzer, but in this case the lens system between the intermediate plane P_2 and the output plane P_3 consists of a single spherical lens. The operation of this system is characteristic of the more general 1-D signal processing systems described in the next chapter. We note here, however, that with an intermediate plane mask with transmittance

$$t(\nu,y) = m(\nu,y)e^{i2\pi h(\nu)y} \quad , \tag{2-56}$$

the system output along the vertical axis is

$$u_3(0,\eta) = \int_{-\infty}^{\infty} F(\nu)M_y(\nu,\eta-g(\nu))d\nu \quad , \tag{2-57}$$

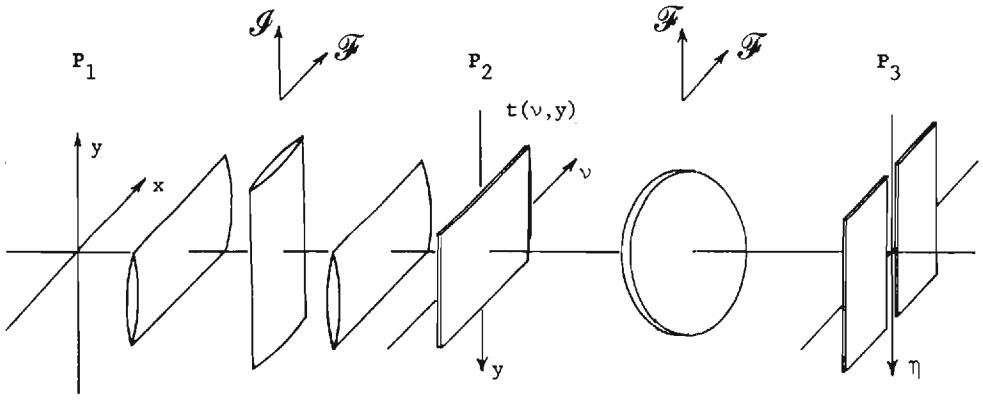


Figure 2-14. Holographic Frequency-Variant Spectrum Analyzer.

where

$$M_y(\nu,\eta) = \mathcal{F}_y\{m(\nu,y)\} \quad . \tag{2-58}$$

The response to an individual frequency component $\exp[i2\pi\nu_o x]$ ($F(\nu) = \delta(\nu-\nu_o)$) is then

$$U_3(0,\eta) = M_y(\nu_o, \eta-g(\nu_o)) \quad . \tag{2-59}$$

In this system, the resolution vs. frequency characteristics as well as the frequency mapping characteristics are controlled by the intermediate plane mask. The transmittance described in Eq. (2-56) is difficult to fabricate even for simple frequency mapping requirements. However, the transmittance can be incorporated into a holographic recording. Details concerning the construction of the holographic mask can be found in the paper by Rhodes [24].

Discussion

Conventional coherent optical spectrum analyzers are generally characterized by frequency-invariant operation: a cissoidal component, $\exp[i2\pi\nu x]$, of an input signal produces a response in the output plane whose position varies linearly with frequency and whose shape remains fixed. There are circumstances, however, when a frequency dependent analysis of signal spectral content is desirable.

We have shown how the second degree of freedom provided by an optical systems allows the frequency content of one-dimensional signals to be analyzed in operations that are frequency-variant. We have discussed in

general the requirements and constraints on the masks used in the system to control resolution vs. frequency characteristics and frequency mapping characteristics. Also, we have given an equivalent mask description simplifying the description of mask interaction in the system operation. Specific design examples considered include both linear frequency and log frequency constant proportional bandwidth analyses. Additionally, we have provided a general procedure for the calculation of space (time)-bandwidth products appropriate for both frequency invariant and frequency variant systems.

The descriptions presented in this chapter are intended to convey a general understanding of the operation and limitations of this class of signal analysis system. It has not been our purpose to specify design parameters; anyone skilled in the area of Fourier optics is capable of determining pertinent focal lengths and scale factors, and the choice of input window or mapping slit profiles--whether they should be binary or tapered, gaussian or raised cosine--relates to signal analysis topics adequately discussed elswhere [25-29]. The systems described enjoy relative ease of construction. Interferometric accuracy in the positioning of components is not necessary. The masks that control the frequency varying response and frequency mapping operations can be specified with great flexibility.

CHAPTER III

GENERAL LINEAR PROCESSING WITH COHERENT OPTICAL SYSTEMS

In this chapter, we discuss the problem of optical implementation of general 1-D linear superposition operations. Cutrona et al. in 1959 were first to recognize the capability of astigmatic optical systems to perform such operations [17]. The work reported by Rhodes and Florence on frequency-variant spectrum analysis, the major results of which were reported in the previous chapter, is the first significant additional investigation and application of 1-D shift-variant operations in optical systems [23,30-33]. This recent work has stimulated renewed interest in the topic, as is evident by the number of papers and reports describing the optical implementation of such operations. The work of Goodman, Kellman, et al. has closely paralleled that of Marks, Walkup et al., with both groups describing a number of systems for geometrical transformations [34,35] and for performing integral transform operations such as the inverse Abel transform [34,36]. Additionally, Kellman and Goodman have described a purely optical system for 1-D Mellin transforms [37], and Marks et al. have described a system to display the ambiguity function for 1-D signals [38].

The spectrum analysis operations discussed in the previous chapter were described as frequency-variant primarily to indicate that the output was a spectral display that had been distorted or altered in a manner that was somehow frequency dependent. Similarly, a geometrical distortion can be described as space-variant, indicating a distortion that is

spatially dependent. In many cases, however, optical systems have been described as space-variant simply to indicate that general linear operations are performed by the system on a signal without regard for what the operations are or how they are performed. In the discussion below we take the viewpoint that there is a fundamental difference between a space-variant implementation of a general linear operation and a frequency-variant implementation of a general linear operation. We demonstrate this difference later by considering both implementations of a single operation, the Mellin transform. First, however, we begin with a basic discussion of linear operations and the various optical methods for implementing these operations.

Linear Systems and Operations

The theory of linear systems has been well established for many years. The concept of linear shift-invariant systems and the relationships describing their operation are fundamental in all areas of science and engineering. In many areas (such as communications or electronics), shift-variant systems (e.g., time-variant communications channels or time-varying filters) have been studied and their basic operational relationships determined (see, for example Ref. [39]). In this section, we consider the distinctions between shift-invariant and shift-variant systems, and we introduce the concepts of frequency-invariance and frequency-variance in a linear system. In addition, we establish the relationships between these various properties in a general linear system.

Our primary interest here is optical systems and we will therefore make reference to the space domain and the spatial frequency domain. (The latter we often denote simply as the frequency domain.) Except where

optical systems are specifically mentioned, the discussion is equally valid for systems operating on functions of time, requiring simply the substitution of the words time for space and temporal frequency for spatial frequency. We note, however, that the major implications of the discussion are important only in the context of optical systems.

In order to establish notation, we begin this discussion with a very brief review of the basic theory. Referring to Fig. 3-1, a system can be considered as a device or process that operates on an input, $f(x)$, producing an output, $g(y)$. Analytically, a system can be characterized by an operator, $O\{\cdot\}$, that specifies the relationship between the input and the output according to

$$g(y) = O\{f(x)\} \quad . \tag{3-1}$$

A linear system is a system whose operator has the property of linearity:

$$O\{c_1 f_1(x) + c_2 f_2(x)\} = c_1 O\{f_1(x)\} + c_2 O\{f_2(x)\} \quad , \tag{3-2}$$

where c_1 and c_2 are arbitrary constants. In general, the response of a linear system to an arbitrary input is given by the superposition integral

$$g(y) = \int_{-\infty}^{\infty} f(x) h(y,x)\, dx \quad , \tag{3-3}$$

where $h(y,x)$ is the system response at y to a unit impulse input at x. The validity of Eq. (3-3) is easily verified by making the substitution

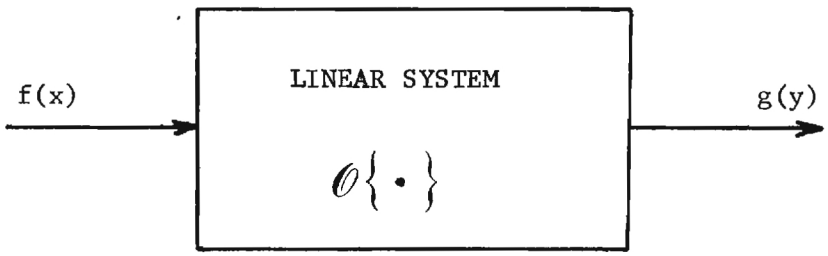


Figure 3-1. Linear System Model.

$$f(x) = \int_{-\infty}^{\infty} f(x')\delta(x'-x)dx' \quad , \tag{3-4}$$

in Eq. (3-1) and applying the linearity property. The function h(y,x) therefore completely characterizes the system for all inputs that can be expressed as in Eq. (3-4). If a shift in the system input results in an equivalent shift in the system output, with the output remaining otherwise unchanged, then the system is called space-invariant. In this case, the impulse response reduces to a function of y-x, i.e.,

$$h(y,x) \Rightarrow h(y-x) \quad , \tag{3-5}$$

and Eq. (3-3) becomes the convolution operation

$$g(y) = \int_{-\infty}^{\infty} f(x)h(y-x)dx \quad . \tag{3-6}$$

The expression for the input function given in Eq. (3-4) describes the input in terms of a continuum of delta-functions. This description decomposes f(x) into the simpler functions for which the system response is known. Due to the linearity property, the response to the general function is then the sum (integral) of the responses to the simpler inputs. An alternative decomposition of interest here is found by expressing f(x) in terms of its Fourier transform, i.e.,

$$f(x) = \int_{-\infty}^{\infty} F(\nu)e^{i2\pi\nu x}d\nu \quad , \tag{3-7}$$

where

$$F(\nu) = \int_{-\infty}^{\infty} f(x)e^{-i2\pi\nu x}dx \quad . \tag{3-8}$$

In this decomposition, f(x) is given as a weighted sum of single frequency cissoidal functions $e^{i2\pi\nu x}$. Substituting Eq. (3-7) for f(x) in Eq. (3-1) and applying the linearity property, we find the system output

$$g(y) = \int_{-\infty}^{\infty} F(\nu)O\{e^{i2\pi\nu x}\}d\nu \quad . \tag{3-9}$$

The system cissoid response can then be defined as

$$C(y,\nu) = O\{e^{i2\pi\nu x}\} \quad , \tag{3-10}$$

which is the system response at y to a cissoidal input at frequency ν. The output can then be rewritten as

$$g(y) = \int_{-\infty}^{\infty} F(\nu)C(y,\nu)d\nu \quad . \tag{3-11}$$

The function $C(y,\nu)$ completely characterizes the system for all inputs for which the Fourier transform exists. This description is particularly appropriate when dealing with optical systems because such systems can provide direct access to the frequency domain. Single frequency cissoidal components of an input can be isolated and manipulated in certain planes within an optical system just as a single spatial component (delta function) can be isolated and manipulated in other planes. Thus, unlike the situation in electronic systems, this characterization corresponds to a physically realizable configuration in the optical system.

If the cissoid response shifts linearly with a shift in frequency of an input cissoid but otherwise remains unchanged, then we refer to this system as frequency-invariant and

$$C(y,\nu) \Rightarrow C(y-\nu) \tag{3-12}$$

The general response of a frequency-invariant system can therefore be expressed as the convolution integral

$$g(y) = \int_{-\infty}^{\infty} F(\nu)C(y-\nu)d \quad . \tag{3-13}$$

An example of such a system would be a conventional spectrum analyzer.

We have described two different methods for determining the response of a linear system to an arbitrary input. One method specifies the output in terms of the system impulse response, while the dual method (the cissoid is an impulse in the frequency domain) describes the output in terms of the cissoid response. The equivalence of expressing the output of any linear system in terms of the impulse response or in terms of the cissoid response is found in a general statement of Parseval's theorem [40]:

$$\int_{-\infty}^{\infty} f(x)h^{*}(y,x)dx = \int_{-\infty}^{\infty} F(\nu)H^{*}(y,\nu)d\nu \quad , \tag{3-14}$$

where

$$F(\nu) = \mathscr{F}_{x}\{f(x)\} \quad , \tag{3-15}$$

and

$$H(y,\nu) = \mathcal{F}_x\{h(y,x)\} \quad . \tag{3-16}$$

We have also defined two special cases of linear systems, one being the space-invariant system, the other being the frequency-invariant system. It is easy to show that space-invariance and frequency-invariance are mutually exclusive properties of linear systems. Consider the response of a space-invariant system with impulse response h(y-x) to a single frequency cissoidal input $e^{i2\pi\nu x}$. The output of the system is, from Eq. (3-5)

$$O\{e^{i2\pi\nu x}\} = \int_{-\infty}^{\infty} e^{i2\pi\nu x} h(y-x)dx = H(\nu)e^{i2\pi\nu y} \quad , \tag{3-17}$$

where

$$H(\nu) = \mathcal{F}_x\{h(x)\} \quad . \tag{3-18}$$

The cissoid response is clearly not a function of y-ν alone. It can, similarly, be shown that a frequency-invariant system can not also be space-invariant.

Finally, we consider more closely the non-invariant cases. Simply stated, a system is <u>space-variant</u> if the system impulse response changes shape or size or shifts non-linearly (here we mean any non-equivalent shift) with a shift in the input impulse. Similarly, if the system cissoid response changes shape or size or shifts non-linearly with a change

in the frequency of the cissoidal input, then the system is frequency-variant. If we consider the cissoid response of a space-variant system, we find

$$O\{e^{i2\pi\nu x}\} = \int_{-\infty}^{\infty} e^{i2\pi\nu x} h(y,x)dx = H(y,-\nu) \quad , \tag{3-19}$$

where

$$H(y,\nu) = \mathcal{F}_x\{h(y,x)\} \quad . \tag{3-20}$$

It is clear that the space-variant system will be frequency-invariant only if the impulse response is of the form

$$h(y,x) = h(x)e^{i2\pi yx} \tag{3-21}$$

In the more general case, a space-variant system is also frequency-variant. Similarly, a frequency-variant system will in general be space-variant with the exception occurring when the cissoid response has the form of Eq. (3-17).

Optical Implementation of General 1-D Linear Operations

The equivalence of space-variant and frequency-variant expressions for a general linear operation is especially important when considering the optical implementation of that operation. Because a coherent optical system can provide direct access to the frequency domain, Eq. (3-14) suggests that there are two fundamentally different schemes for optically

implementing a given linear operation. One scheme operates directly on the input signal, altering or distorting the spatial components of the input according to the superposition integral on the left side of Eq. (3-14). In the other scheme, the input undergoes a Fourier transform, after which the spatial frequency components of the input are altered or distorted according to the superposition integral on the right side of Eq. (3-14).

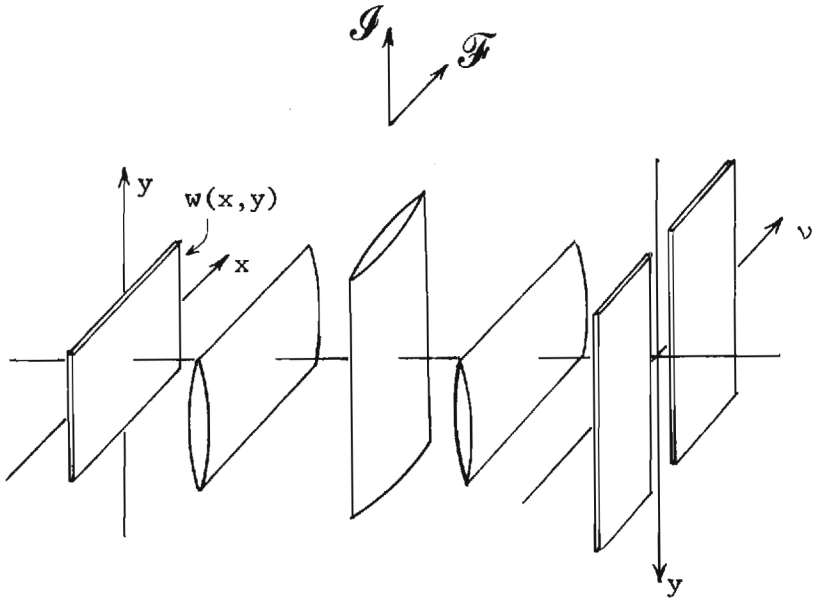
Two optical systems utilizing the first scheme are shown in Fig. 3-2. In both systems, the input transparency, f(x), is placed in contact with a mask, w(x,y), in plane P_1. In system A, this combination is imaged in the vertical direction and Fourier transformed in the horizontal direction. The transform distribution is masked with a vertical slit at $\nu=0$. The kernel of the transform operation becomes unity at $\nu=0$ resulting in the output given by the integral expression

$$U(0,y) = \int_{-\infty}^{\infty} f(x)w(x,y)dx \quad . \tag{3-22}$$

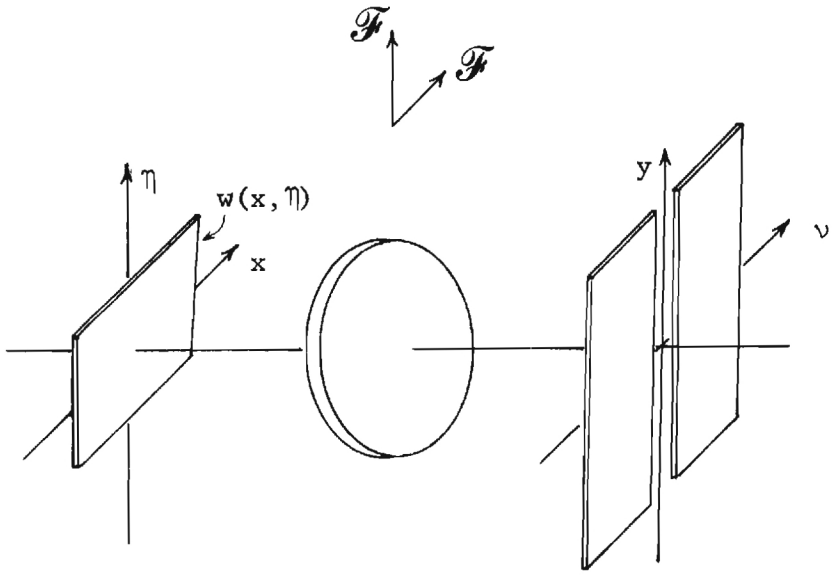
In system B, the single spherical lens performs a 2-D Fourier transform, which again becomes an integration with respect to x if we observe only at $\nu=0$. The vertical Fourier transform is not affected by the slit mask. The resulting output distribution is then

$$U_2(0,y)=\int_{-\infty}^{\infty}\int f(x)w(x,\eta)e^{-i2\pi y\eta}dxd\eta \quad , \tag{3-23}$$

which we can rewrite as



(A)



(B)

Figure 3-2. Two Systems for the Space-Variant Implementation of General 1-D Linear Operations.

$$U_2(0,y) = \int_{-\infty}^{\infty} f(x)W(x,y)dx \quad , \tag{3-24}$$

where

$$W(x,y) = \mathscr{F}_{\eta}\{w(x,\eta)\} \tag{3-25}$$

In system A, we see that the system impulse response is given directly by the mask transmittance

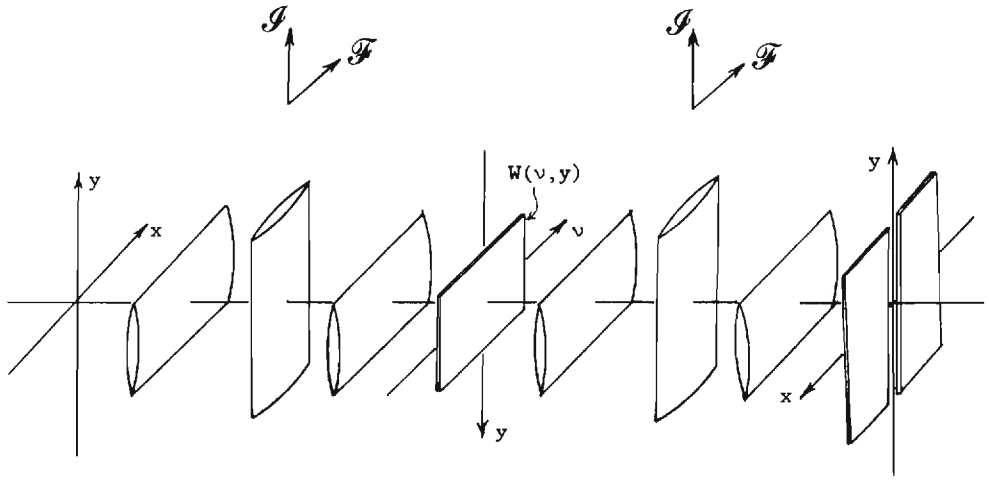
$$h(y,x) = w(x,y) \quad , \tag{3-26}$$

whereas in system B, the impulse response is

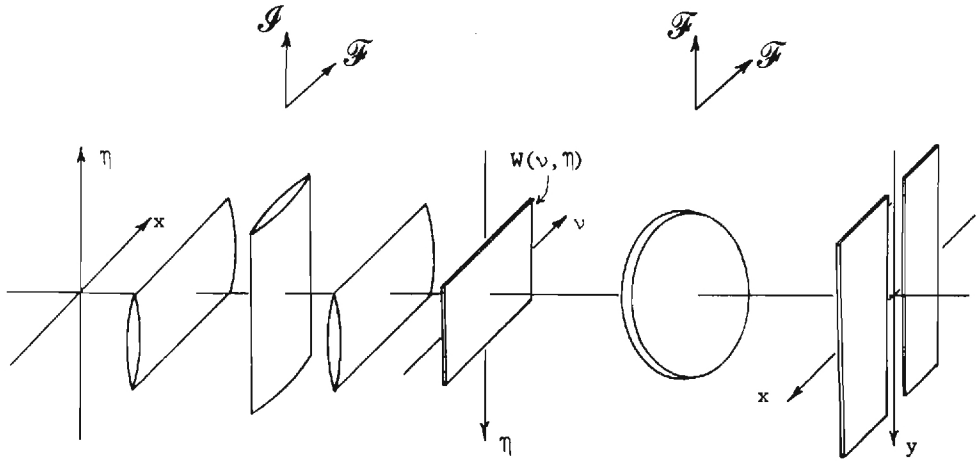
$$h(y,x) = W(x,y) \quad , \tag{3-27}$$

which can be determined from the mask transmittance according to Eq. (3-25). In both systems, the mask operates directly on the input signal, consistent with the space-variant expression for general linear operations. This scheme is therefore appropriately described as the space-variant implementation of a general linear operation.

The second approach for performing general linear operations on 1-D signals is utilized by the two systems shown in Fig. 3-3. In each system, the input, f(x), undergoes a vertical imaging/horizontal Fourier transforming operation. The resulting spectral distribution appears incident on a filter mask W. The general linear superposition operations



(C)



(D)

Figure 3-3. Two Systems for the Frequency-Variant Implementation of General 1-D Linear Operations.

are then performed in the same manner, respectively, as the system in Fig. 3-2. In system C, a second vertical imaging/horizontal Fourier transform operation is performed, with the output along the y-axis given by

$$U_3(0,y) = \int_{-\infty}^{\infty} F(\nu)W(\nu,y)d\nu \quad . \tag{3-28}$$

The cissoid response of system C is given directly by the mask transmittance

$$C(y,\nu) = W(\nu,y) \quad . \tag{3-29}$$

System D employs a final 2-D Fourier transform to perform the superposition, with the resulting output given by

$$U_3(0,y) = \int_{-\infty}^{\infty} F(\nu)W_\eta(\nu,y)d\nu \quad . \tag{3-30}$$

The system cissoid response is determined from the mask transmittance by

$$C(y,\nu) = W_\eta(\nu,y) = \mathcal{F}_\eta\{W(\nu,\eta)\} \quad . \tag{3-31}$$

Both of these systems perform the general linear superposition operation on the spectrum of the input signal as indicated in the frequency-variant representation of the operation. We therefore describe this scheme as the frequency-variant implementation of a general linear operation.

The general linear operation can also be implemented with the

systems of Fig. 3-3 using two masks, one in the input plane and one in the transform plane. These alternate configurations are shown in Fig. 3-4. The input signal is masked in the input plane by $w_1(x,y)$. The frequency-plane light distribution is then given by the convolution

$$U_2(\nu,y) = \int_{-\infty}^{\infty} F(u)W_1(\nu-u,y)du \quad , \tag{3-32}$$

in either system, where

$$W_1(\nu,y) = \mathcal{F}_x\{w_1(x,y)\} \quad . \tag{3-33}$$

These systems then perform like the frequency-variant implementations, multiplying the frequency-plane distributions by the filter $W_2(\nu,y)$, and performing an integration with respect to ν. In system E, the output is given by

$$U_3(0,y) = \int\int_{-\infty}^{\infty} F(u)W_1(\nu-u,y)W_2(\nu,y)dud\nu \quad . \tag{3-34}$$

This expression can be rewritten as

$$U_3(0,y) = \int_{-\infty}^{\infty} F(u)W_e(u,y)du \quad , \tag{3-35}$$

where

$$W_e(u,y) = \int_{-\infty}^{\infty} W_1(\nu-u,y)W_2(\nu,y)d\nu \quad . \tag{3-36}$$

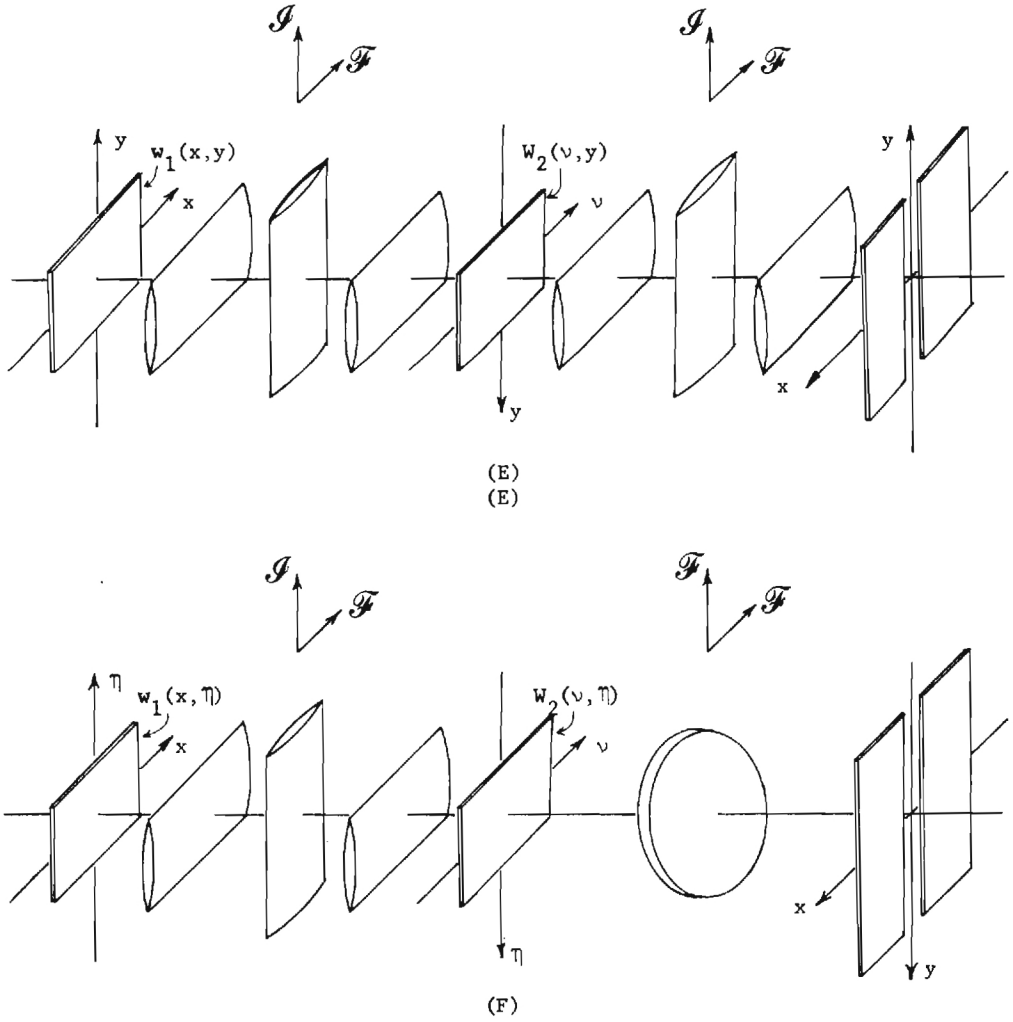


Figure 3-4. Two Systems for the Two-Mask Frequency-Variant Implementation of General 1-D Linear Operations.

Equation (3-36) is the expression for the equivalent frequency domain window for the system. This expression is also the system cissoid response. In system F, the filtered frequency-domain distribution undergoes a 2-D Fourier transform, with the output given by

$$U_3(0,y) = \iiint_{-\infty}^{\infty} F(u)W_1(\nu-u,\eta)W_2(\nu,\eta)e^{-i2\pi y\eta}dud\nu d\eta \quad . \tag{3-37}$$

This expression can also be simplified to the frequency-variant expression

$$U_3(0,y) = \int_{-\infty}^{\infty} F(u)W_e(u,y)du \quad , \tag{3-38}$$

where here

$$W_e(u,y) = \iint_{-\infty}^{\infty} W_1(\nu-u,\eta)W_2(\nu,\eta)e^{-i2\pi y\eta}d\nu d\eta \quad . \tag{3-39}$$

$W_e(u,y)$ is, again, the system cissoid response.

The two-mask implementations can have certain advantages in flexibility over an equivalent one mask system by dividing the control of the overall system operation between two planes. The spectrum analyzer system described in the previous chapter employed two masks to perform the operation. In that system, the transform plane mapping slit controlled the position vs. frequency characteristics of the output display, while at the same time the resolution vs. frequency characteristics were controlled by a variable width mask in the input plane. We note, however, that the operation of the two-mask implementations, like the one-mask frequency-variant implementation, relies inherently on access to the frequency

domain provided by the optical systems. We will therefore refer to these implementations as the two-mask frequency-variant implementations of the general linear operation.

The Mellin Transform: Two Implementations

To understand better the differences between frequency-variant and space-variant implementations, we consider both methods of implementing a single operation. The operation we choose to implement is the Mellin transform (with purely imaginary argument) given by

$$g(y) = \int_0^\infty f(x) x^{i2\pi y - 1} dx \quad . \tag{3-40}$$

The Mellin transform has been studied by several researchers for its possible application to optical pattern recognition [41,42], and Doppler signal processing [43]. The primary feature of this transform that makes it applicable to these problems is its scale invariance.

Casasent and his colleagues have demonstrated both 1-D and 2-D Mellin transforms using an electrooptical input device to perform part of the operation [41-44]. In this system, the input is introduced to the system using a spatial light modulator. The amplifiers controlling the position coordinates of the input distribution are logarithmic, so that the input function arguments become exponential functions of the input plane cartesian coordinates. The operation is then completed by performing a Fourier transform operation. That this operation produces the Mellin transform can be seen by making the substitution $x=e^{\eta}$ in Equation (3-40). Noting that $d\eta = dx/x$, we see that, after the variable substitution, Equation (3-40) becomes

$$g(y) = \int_{-\infty}^{\infty} f(e^{\eta}) e^{i2\pi\eta y} d\eta \quad , \tag{3-41}$$

as desired.

A purely optical space-variant implementation of the 1-D Mellin transform can be effected directly along these lines. To determine the required mask (and best space-variant system) to use, we rewrite Equation (3-41) as

$$g(y) = \int\int_{-\infty}^{\infty} f(\xi)\delta(\xi - e^{\eta}) e^{i2\pi\eta y} d\xi d\eta \quad . \tag{3-42}$$

Comparing this expression with Equations (3-23) and (3-25), it is evident that system B in Fig. 3-2 can be used to perform this operation. The mask required is simply a narrow slit along the line $x = e^{\eta}$. This system is shown in Fig. 3-4.

For a frequency-variant implementation, we need access to the spectral distribution of the input signal. The analytic form of the Mellin transform in terms of $F(\nu)$ can be determined from Equation (3-41) by noting that $f(e^{\eta})$ is the inverse Fourier transform of $F(\nu)$ evaluated at $x = e^{\eta}$; i.e.,

$$f(e^{\eta}) = \int_{-\infty}^{\infty} F(\nu) e^{i2\pi\nu e^{\eta}} d\nu \quad . \tag{3-43}$$

Substituting Equation (3-43) into Equation (3-41), we obtain an equivalent expression for the Mellin transform

$$g(y) = \int\int_{-\infty}^{\infty} F(\nu) e^{i2\pi\nu e^{\eta}} e^{i2\pi y\eta} d\nu d\eta \quad . \tag{3-44}$$

A system that performs this operation is shown in Fig. 3-5. This system is the second example (system D) of the frequency-variant implementations of Fig. 3-3. Comparing Equations (3-30), (3-31) and (3-44), the required frequency-plane mask is seen to be the phase-only transmittance function

$$W(\nu,\eta) = e^{i2\pi e^{\eta}\nu} \quad . \tag{3-45}$$

Phase-only transmittance filters, although efficient in the utilization of light, are somewhat difficult to construct directly. As a consequence, the phase functions are usually incorporated in the transmittance of holographic filters. Recently, Kellman and Goodman have demonstrated 1-D Mellin transforms using essentially the system in Fig. 3-5 and a holographic filter [37]. The filter is recorded from a double slit mask as shown in Fig. 3-6.* The transmittance of the carrier frequency holographic filter is

$$W(\nu,\eta) = B + \alpha\cos[2\pi\nu(e^{\eta} + X_o/2)] \quad , \tag{3-46}$$

where B is the bias transmittance, α is the fringe amplitude, and $X_o/2$ is the offset carrier for the recording determined by the minimum slit separation of the double slit mask. When this mask is used in the system of Fig. 3-5, the distribution in the output plane is

---

*This holographic mapping technique had been previously described by Rhodes for implementing a log-frequency spectrum analysis [24].

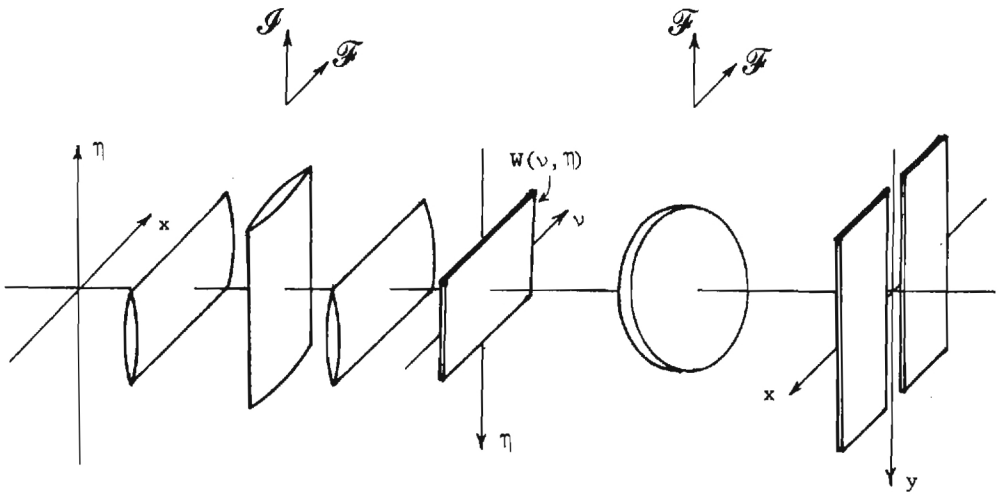
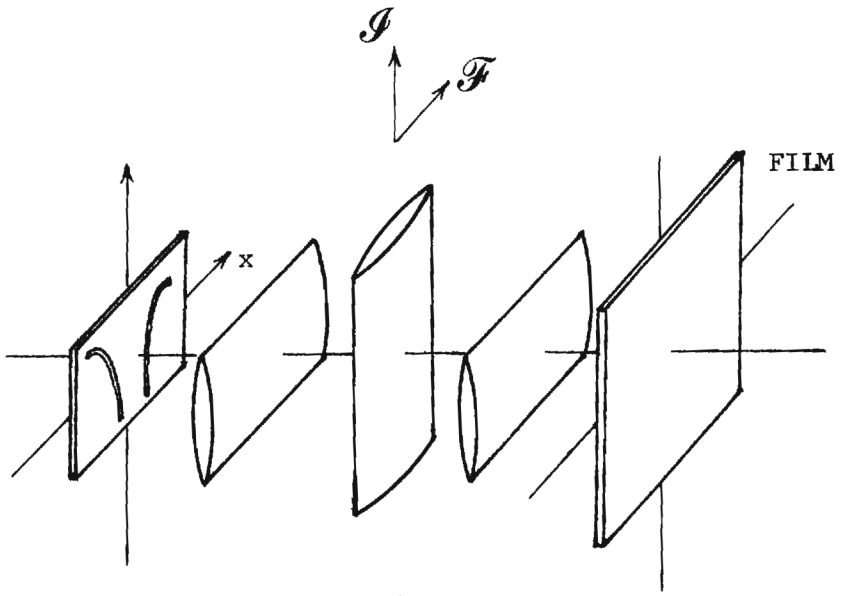
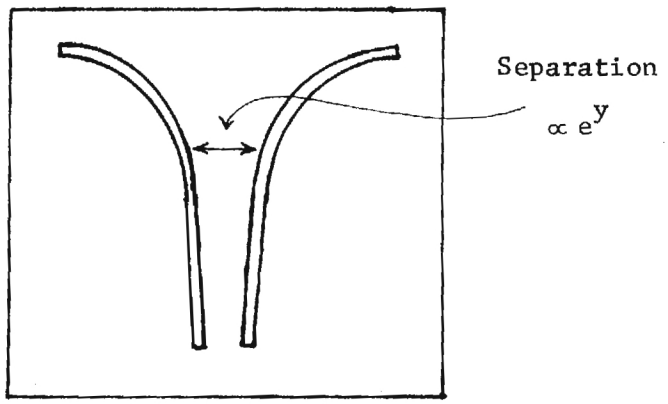


Figure 3-5. System for the Frequency-Variant Implementation of the 1-D Mellin Transform.



(a)



(b)

Figure 3-6. Recording Technique for Holographic Filter in 1-D Mellin Transforming System: (a) System for Exposing Film Plate; (b) Double Slit Mask Used in the Recording System.

$$U_3(x,y) = \iint_{-\infty}^{\infty} F(\nu)\{B + \alpha\cos[2\pi\nu(e^{\eta} + X_o/2)]\}e^{i2\pi(\nu x+\eta y)}d\nu d\eta \quad . \qquad (3\text{-}47)$$

In the vicinity of $x = -X_o/2$, the bias term and the negative frequency (-i) term of the cosine are negligible. (The distributions resulting from the bias transmittance and the negative frequency diffraction order of $W(\nu,\eta)$ are spatially separated from the positive-frequency diffraction order). From Equation (3-47), we then obtain

$$U_3(-X_o/2,y) \tilde{=} \iint_{-\infty}^{\infty} F(\nu)\,\frac{1}{2}e^{i2\pi\nu e^{\eta}}e^{i2\pi\nu\frac{X_o}{2}}e^{-i2\pi\nu\frac{X_o}{2}}e^{i2\pi y\eta}d\nu d\eta \ , \qquad (3\text{-}48)$$

or, noting that $e^{i2\pi\nu\frac{X_o}{2}}e^{-i2\pi\nu\frac{X_o}{2}} = 1$,

$$U_3(-\frac{X_o}{2},y) = \frac{1}{2}\iint_{-\infty}^{\infty} F(\nu)e^{i2\pi\nu e^{\eta}}e^{i2\pi y\eta}d\nu d\eta \quad . \qquad (3\text{-}49)$$

Comparing Equation (3-49) and Equation (3-44), we see that the output along $x = -X_o/2$ is, indeed, the Mellin transform of $f(x)$.*

Discussion

We have presented four different systems capable of performing general 1-D linear superposition operations (there are actually six systems if we include the two two-mask frequency-variant implementations).

---

*In the analysis of Ref. [37], Kellman and Goodman incorrectly conclude that this system evaluates the Mellin transform of $xf(x)$ rather than of $f(x)$ alone.

The systems described were classified--space-variant or frequency-variant--depending on how the operation was implemented. For each system there are different requirements of the system masks for a given operation, providing a good deal of flexibility in implementing that operation. General statements concerning which system is to be preferred are not possible. But in specific cases, the different implementations can be compared with attention to such matters as ease of implementation and efficiency of light utilization. The discussion of the two optical implementations of the Mellin transform made clear the essential differences between the two classes. The operation in a space-variant implementation relies intrinsically on a distortion or alteration of the spatial components of the input. A frequency-variant implementation is based on a distortion or an alteration of the Fourier spectral components of the input. Thus, for optical systems, the two fundamental expressions for a general linear operation translate directly into two fundamental methods of implementing the operation.

CHAPTER IV

FREQUENCY-VARIANT OPTICAL PROCESSING OF TEMPORAL SIGNALS

The processing operations discussed in the preceding chapters have all produced an optical display as the output, such as the spectral display. In this chapter, we investigate the extension of the frequency-variant mapping operations to the mapping of temporal frequencies from an input temporal signal waveform to an output temporal signal waveform. Optical techniques for performing frequency-invariant short-time spectral mappings have been studied previously [45, 46, 47, 48]. The new techniques described here, however, provide greater flexibility and generality in the mapping operations. The special nature of these techniques, which rely on the two-dimensional processing capability of optical systems, makes it possible to perform spectral mappings in a frequency-dependent manner. Applications of this type of signal processing include bandwidth compression or expansion operations where the amount of compression or expansion can vary for different portions of the signal spectrum. Additionally, it has applicability in more general frequency domain redistributions for spread-spectrum signaling or the correction of frequency domain distortions.

In this chapter, we speak of both temporal frequencies associated with signal waveforms and spatial frequencies associated with light distributions. To avoid confusion in the analytic relationships dealing with these quantities we introduce a notation slightly different from the preceding chapters. Specifically, the variable ν is used to denote temporal

frequency while u and v denote spatial frequencies. We make this change in order to retain ν as the variable associated with frequencies being mapped in the processing operation.

The Coherent Optical Signal Processor

The general format for a coherent optical signal processing system is shown schematically in Fig. 4-1. In such a system, we begin with an input signal, f(t), to be processed and proceed through three different mapping operations to an output signal g(t). As indicated in the figure, the input signal first undergoes an input mapping designed to produce an optical distribution, i(x,y;t) suitable for processing by a coherent optical system. Specifically, the input mapping maps f(t) into a light amplitude distribution of the form

$$i(x,y;t) = [C_b + C_a f(\frac{x+\beta t}{\alpha})]w(x,y), \tag{4-1}$$

where α and β are real constants, C_a is a constant of proportionality, C_b is a bias amplitude, and w(x,y) is an aperture function limiting the extent of i(x,y;t). A distribution of this form can be produced in several ways. If real time processing is not required, the input mapping can be accomplished by recording the signal as a density modulated film sound track. Recording and film processing are such that the resulting sound track has a light amplitude transmittance

$$t(x,y) = t_b + t_a f(\frac{x}{\alpha}) 1(y) \tag{4-2}$$

where here α represents the recording velocity of the sound track. When

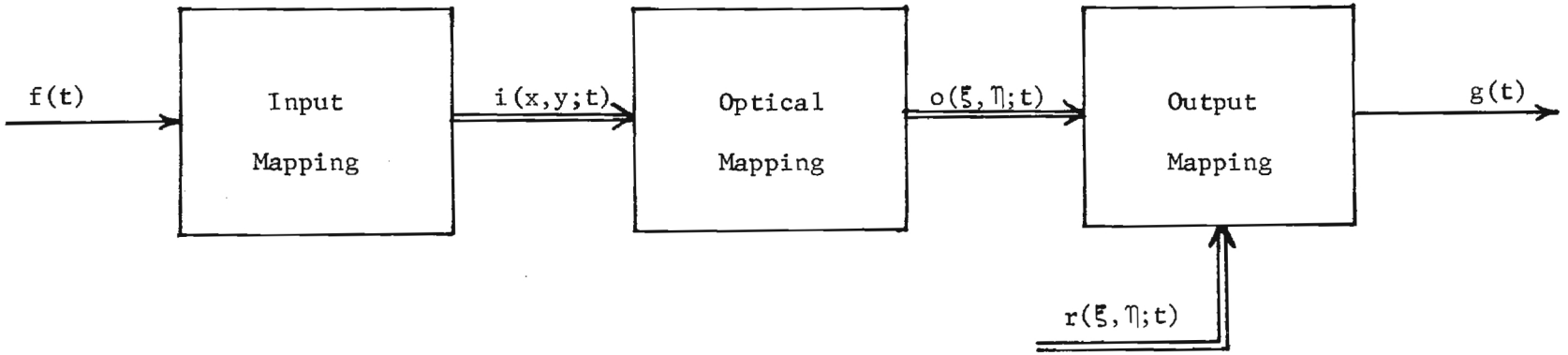


Figure 4-1. Coherent Optical Signal Processing System.

the sound track is illuminated and translated in the x-direction past the aperture $w(x,y)$ with a playback velocity β, an amplitude distribution of the form of Eq. (4-1) results. Note that the time-varying portion of the input amplitude observed at $x=0$ is proportional to $f(\beta t/\alpha)$. With a sound track, the input signal can be introduced with a transport velocity greater or less than the recording velocity for variable rate playback. For real-time processing applications, an acoustooptic cell can be used in the input mapping. In this case, α and β are equal and represent the acoustic transport velocity within the cell.

Associated with the input mapping is a time delay t_d. Taking this delay into account in the analytic expressions, the processor output has the form $g(t) = g_1(t-t_d)$. In the case of real-time processing using the acoustooptic cell, t_d is associated with the acoustic wave transit time through the cell, which is very small (typically microseconds). For the sound track, t_d might represent the time between the recording and the playback of the sound track - a time that could be minutes, hours, or even days. In either case, explicit retention of the time delay term is found to be analytically clumsy. We will therefore assume that the time delay associated with the input mapping to be understood and represent both the input signal and the input optical distribution as functions of the time variable t. In a further effort to simplify the analysis, we assume that $\beta=\alpha$, reducing the final expression for the input mapping to

$$f(t) \rightarrow i(x,y;t) = [C_b + C_a f(\frac{x+\alpha t}{\alpha})]w(x,y). \qquad (4-3)$$

All analytical results derived in this chapter are therefore applicable in

the cases of real-time input mapping or film sound track input mapping where recording and playback velocities are equal. For the case of variable-rate playback of sound track recordings, the time variable t in all expressions must be replaced with $(\frac{\beta}{\alpha})t$.

The second mapping occuring in the coherent processor is an instantaneous optical mapping that produces an output optical distribution $O(\xi,\eta;t)$ from the input optical distribution, i.e.

$$i(x,y;t) \rightarrow O(\xi,\eta;t). \tag{4-4}$$

The specific details of the optical mapping depend on the optical system used to perform the mapping. The function of this mapping in conjunction with the input mapping is to transform the input signal into an optical distribution of appropriate form for processing in the final system mapping.

The final mapping of the processing system is performed in the coherent detection of the output distribution. For this mapping, a second optical distribution, which for reasons that will become clear shortly we call the local oscillator distribution, is added to (mixed with) the output optical distribution. Denoting the local oscillator distribution as $r(\xi,\eta;t)$, the total output amplitude distribution is

$$O(\xi,\eta;t) + r(\xi,\eta;t), \tag{4-5}$$

giving an output irradiance distribution

$$I(\xi,\eta;t) = \left|O(\xi,\eta;t) + r(\xi,\eta;t)\right|^2. \tag{4-6}$$

This irradiance distribution is incident on the photosensitive surface of a large area photodetector such as a photomultiplier tube (PMT). The PMT responds to the incident irradiance producing a current signal

$$s(t) \propto \iint_{\Sigma} I(\xi,\eta;t)\,d\xi d\eta, \tag{4-7}$$

where Σ represents the extent of the photosensitive surface of the PMT. If a large area detector is used that collects all of the incident light, the limits of integration may be assumed infinite. Substituting Eq. (4-6) into the expression for s(t) and expanding the squared modulus term, we obtain

$$s(t) \propto \iint_{\Sigma} [\left|O(\xi,\eta;t)\right|^2 + \left|r(\xi,\eta;t)\right|^2 + O(\xi,\eta;t)r^*(\xi,\eta;t) + O^*(\xi,\eta;t)r(\xi,\eta;t)]\,d\xi d\eta, \tag{4-8}$$

which we rewrite as

$$s(t) \propto s_0(t) + s_r(t) + g(t), \tag{4-9}$$

with

$$s_0(t) = \iint_{\Sigma} \left|O(\xi,\eta;t)\right|^2 d\xi d\eta, \tag{4-10}$$

$$s_r(t) = \iint_{\Sigma} |r(\xi,\eta;t)|^2 d\xi d\eta, \tag{4-11}$$

and

$$g(t) = \iint_{\Sigma} [O(\xi,\eta;t)r^*(\xi,\eta;t) + O^*(\xi,\eta;t)r(\xi,\eta;t)]d\xi d\eta. \tag{4-12}$$

The signal terms $s_0(t)$ and $s_r(t)$, which we call the self-product terms, represent the signals that would be detected by the PMT from the amplitude distributions $O(\xi,\eta;t)$ and $r(\xi,\eta;t)$ individually (without mixing). The cross product term $g(t)$ is the desired output signal and represents the final mapping of the system. The detector output thus contains the desired output signal plus two additional signals which may degrade the processor output. We show in the next section that there are several ways to eliminate the self-product terms from the detector output. Assuming this is done, the final processor output is simply $g(t)$, given by Eq. (4-12). We can simplify this expression by noting that the integrand is the sum of a function and its complex conjugate; then

$$g(t) = 2\mathrm{Re}\{\iint_{\Sigma} O(\xi,\eta;t)r^*(\xi,\eta;t)d\xi d\eta\}, \tag{4-13}$$

where $\mathrm{Re}\{\cdot\}$ denotes the real part. It will be convenient in the analysis below to express the system operation in terms of the analytic signal

$$\hat{g}(t) = \iint_{\Sigma} O(\xi,\eta;t)r^*(\xi,\eta;t)d\xi d\eta, \tag{4-14}$$

with the final real output signal given by

$$g(t) = 2\mathrm{Re}\{\hat{g}(t)\}. \qquad (4\text{-}15)$$

The overall operation of the processor of Fig. 4-1 depends strongly on the optical mapping operation performed and on the nature of the local oscillator distribution. In the sections below, we investigate more specifically the operation of the coherent signal processor. We begin with a discussion of conventional optical heterodyne processing as exemplified by a system developed by Whitman, Korpel, and Lotsoff at the Zenith Radio Corporation [49]. This system, which we will refer to as the Zenith system, was initially designed to simulate the operation of complex electrical networks. We will discuss its operation and describe how the basic design can be extended to linear bandwidth compression and expansion applications. These extensions lead to the description of a considerably more versatile signal processor that utilizes the second degree of freedom available in the optical mapping system to perform general linear transformations (mappings) of signal spectral components.

Conventional Optical Heterodyne Signal Processing

The fundamental characteristics of conventional optical heterodyne signal processing can be demonstrated using the Zenith system, shown in Fig. 4-2. In this system, a laser provides coherent plane wave illumination at an optical frequency $\nu_0 \simeq 5 \times 10^{14}$ Hz. Through a series of beamsplitters and mirrors, in an arrangement related to a Mach-Zender interferometer, the illumination is divided (we assume equally), passes through two dissimilar optical subsystems and is recombined coherently on the photosensitive surface of a PMT. In one arm of the interferometer, the input mapping is performed in the front focal plane of a lens L, by placing

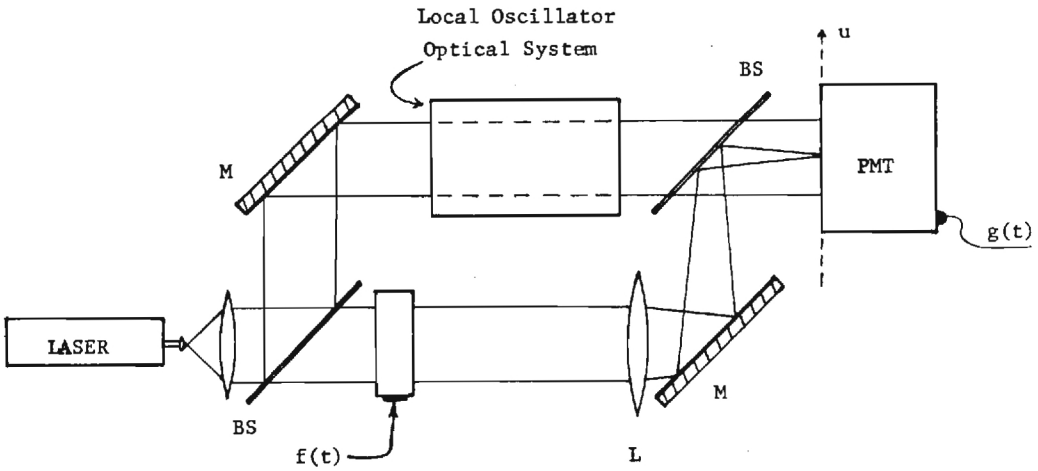


Figure 4-2. The Zenith System.

the input device (acoustooptic cell or the moving sound track) in that plane. The photosensitive surface of the PMT is coincident with the back focal plane of the lens*; the optical mapping operation of the system is therefore the 2-D Fourier transform operation

$$O(u,v;t) = \mathcal{F}\{i(x,y;t)\}. \tag{4-16}$$

In the other arm of the interferometer, an optical system modifies the illuminating plane wave to produce the local oscillator distribution that mixes with the output optical distribution at the PMT. In the output plane a narrow slit is placed along the u-axis, with a stop blocking the range $u \leq 0$. The PMT therefore collects only the light along the positive u-axis, and from Eq. (4-14) the analytic output signal is

$$\hat{g}(t) = \int_0^\infty O(u,0;t)\, r^*(u,0;t)\, du. \tag{4-17}$$

Before analyzing the system with a general input signal, we illustrate its operation by introducing a single sinusoid of temporal frequency ν_1 as the input. This input causes a sinusoidal grating-like variation (amplitude or phase) in the input device with a spatial frequency ν_1/α. The grating-like structure, moving with velocity α in the negative x-direction, introduces the temporal variations at frequency ν_1. As shown in Fig. 4-3, each of the two cissoidal components of the sinusoidal input

---

*It can be shown using conservation of energy arguments that the exact placement of the PMT in the back focal plane of the lens is not necessary - see, for example, Reference [50].

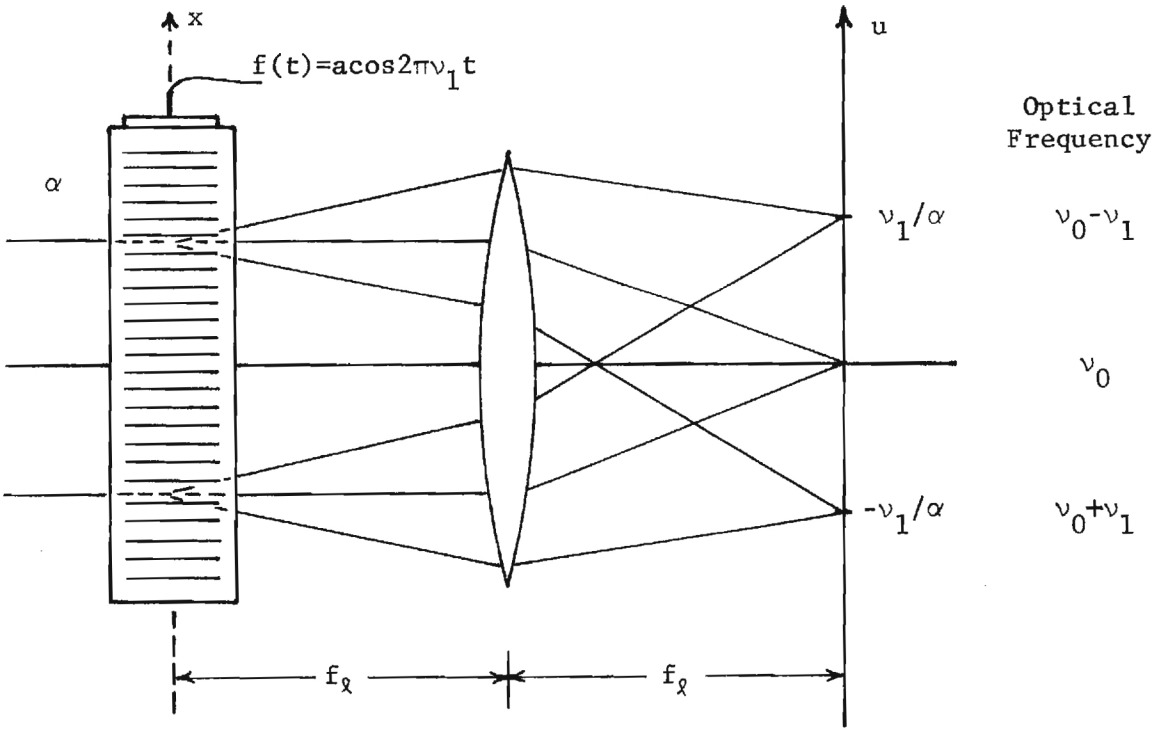


Figure 4-3. Diffraction of Planewave Illuminating Input Device with Single Sinusoid Input.

diffracts a portion of the illuminating plane wave. The diffracted waves are brought to a focus by the lens to produce two spots of light at $u = \pm\nu_1/\alpha$ on the u-axis. The undiffracted or bias portion of the input distribution focuses to a spot of light at u=0. These spots represent the spatial frequency components of the input distribution through the Fourier transform optical mapping. The spot at u=0 represents the average amplitude of the input distribution, and will be referred to as the d.c. spot. The spots at $u=\nu_1/\alpha$ and $u = -\nu_1/\alpha$ will be referred to as the positive and negative frequency components, respectively, of the input signal. The motion of the grating-like variations in the input plane results in a shift in the optical frequency of these two spots of light. The frequency shift can be interpreted as a Doppler shift: light in the positive frequency spot, being deflected away from the direction of motion of the grating, is downshifted in optical frequency from ν_0 to $\nu_0-\nu_1$. Similarly, the light in the negative frequency spot, being deflected in the opposite direction, is upshifted in optical frequency from ν_0 to $\nu_0+\nu_1$. The optical frequency of the undeflected light in the d.c. spot is unchanged. In the heterodyne mapping of the Zenith system, the unblocked positive frequency component is mixed with a local oscillator distribution at the original optical frequency ν_0. This mixing of the positive frequency spot at optical frequency $\nu_0-\nu_1$ with a distribution at ν_0 produces an irradiance distribution that varies sinusoidally at the difference frequency ν_1. The PMT responds to this distribution yielding as the output signal waveform a sinusoid at temporal frequency ν_1.

Adding additional sinusoidal components at frequencies ν_n to the input signal produces additional light spots in the output light distribu-

tion at points $u=\nu_n/\alpha$. The optical frequency of each spot is shifted from ν_0 by the temporal frequency of the corresponding input signal component. When these spots mix with the local oscillator distribution at ν_0, the output mapping reproduces the original input sinusoids in the output signal waveform. If the local oscillator distribution is uniform in amplitude and phase along the entire u-axis (e.g. a plane wave local oscillator), then the relative amplitude and phase of the input signal components are preserved in the output signal waveform. However, variations in the local oscillator distribution can alter the amplitude and/or phase of the components in the output signal. Thus, we can control the amplitude and phase response of the processor by controlling the amplitude and phase of the local oscillator distribution.

For an arbitrary input signal f(t), the input mapping produces the input optical distribution given in Eq. (4-3). The 2-D Fourier transform optical mapping then produces the output distribution

$$O(u,v;t) = C_b W(u,v) + C_a\{\alpha F(\alpha u) e^{+i2\pi\alpha ut} *_u W(u,v)\}. \qquad (4\text{-}18)$$

We block the d.c. spot and the negative frequency components of the signal with the stop and, through the slit, observe the output distribution only along the u-axis. Thus, for u>0,

$$O(u,0;t) = C_a\{\alpha F(\alpha u) e^{i2\pi\alpha ut} *_u W(u,0)\}. \qquad (4\text{-}19)$$

(In blocking the d.c. spot, we also block any average component of the input signal. We therefore assume that the input signal has zero average

value or that the signal of interest is placed on a carrier to insure zero average value.) The general local oscillator distribution in the Zenith system has the form

$$r(u,0;t) = B(u)e^{i\phi(u)}, \tag{4-20}$$

for $u>0$. Substituting Eqs. (4-19) and (4-20) into Eq. (4-17), we obtain the general output for the Zenith system,

$$\hat{g}(t) = C_a \int_0^{\infty}\int_{-\infty}^{\infty} \alpha F(\alpha\xi)e^{i2\pi\alpha\xi t}W(u-\xi,0)B(u)e^{-i\phi(u)}d\xi du. \tag{4-21}$$

We can assume that the local oscillator distribution varies slowly along the u-axis compared with the spot size $W(\cdot,0)$. Then, integrated Eq. (4-21) with respect to u, we obtain

$$\hat{g}(t) = C_a \int_0^{\infty} F(\nu)B(\frac{\nu}{\alpha})e^{-i\phi(\nu/\alpha)}e^{i2\pi\nu t}d\nu, \tag{4-22}$$

where we have made the variable substitution $\nu=\alpha\xi$. Assuming the input signal is real, Eq. (4-22) is the analytic expression for the output of a system whose input is f(t) and whose transfer function is

$$H(\nu) = \begin{cases} B(\nu/\alpha)e^{-i\phi(\nu/\alpha)}, & \nu \geq 0 \\ \\ B(-\nu/\alpha)e^{i\phi(-\nu/\alpha)}, & \nu < 0 \end{cases}. \tag{4-23}$$

(This form is required for the transfer function to be Hermetian.) The

real system output signal can then be expressed as

$$g(t) = C_a \int_{-\infty}^{\infty} f(\tau)h(t-\tau)d\tau, \tag{4-24}$$

where

$$h(t) = \mathcal{F}_{\nu}^{-1}\{H(\nu)\}, \tag{4-25}$$

is the system impulse response. The system response is that of a linear time-invariant system whose transfer function is determined by the local oscillator optical distribution along the u-axis.* If the input signal is not strictly real, the system output signal will have the form of Eq. (4-24) with f(t) now the real signal whose positive frequency spectrum is equal to that of the complex input signal. (The input signal can be complex-valued either by intentionally introducing a non-real input or as a result of phase problems associated with the input transducer.)

Control over the amplitude and phase of the local oscillator is quite flexible, indicating that general system responses can be simulated with this system by the proper choice of local oscillator distribution. A simple example of such a simulation is a local oscillator consisting of a plane wave that is uniform at the detector except at $\nu=\bar{\nu}$, where it is blocked by a thin wire placed across the surface of the PMT perpendicular

---

*The general nature of the relationship between the local oscillator optical distribution and the system transfer function suggests that non-causal filter response characteristics are possible. However, recalling that a time delay is associated with the input mapping, it can be shown that the system response is indeed causal.

to the ν-axis. (This wire will block both the local oscillator distribution and the signal distribution, but the result is the same as if only the local oscillator were blocked.) This system will pass all frequencies with unaltered amplitude and phase except the very small band of frequencies centered at temporal frequency $\bar{\nu}$. The system response is therefore that of a notch filter.

As a second example, consider the introduction of a thin lens into the local oscillator arm of the interferometer. This lens will change the illuminating plane wave into a spherical wave. If the focal length of the lens is chosen to produce a diverging spherical wave at the detector surface, the amplitude of the local oscillator will be uniform across the detector but the phase will vary quadratically with position u. The phase response of the system is therefore a quadratic function of frequency, corresponding to a signal delay proportional to the frequency of the signal. The system has the transfer characteristics of a dispersive delay line and can be used to simulate such a system.

Extensions of Conventional Optical Heterodyne Signal Processing: Frequency Shifting and Linear Bandwidth Compression

In the system developed at Zenith, the amplitude and phase of the various input signal components could be modified, but the frequencies of the components in the output signal remained unchanged. However, a shift in the optical frequency of the local oscillator can easily be used to introduce a uniform shift in the frequencies of all sinusoidal components of the system output. Such a frequency shift can be used, for example, to place the output components in an intermediate band for subsequent processing or to return a signal that has been upshifted, prior to input to

the optical system, to its original baseband range. The shift in optical frequency of the local oscillator can be accomplished with a variety of devices including moving mirrors, rotating quarter- and half-wave plates, and acoustooptic devices [51,52,53]. With an appropriate device introducing a downshift of ν_c, the local oscillator distribution along the u-axis will be

$$\hat{r}(u,0;t) = B(u)e^{i\phi(u)}e^{i2\pi\nu_c t}. \tag{4-26}$$

The system output then becomes

$$\hat{g}(t) = C_a\{\int_{-\infty}^{\infty} \hat{f}(t)\hat{h}(t-\tau)d\tau\}e^{-i2\pi\nu_c t}, \tag{4-27}$$

where we have returned to the analytic signal representations. The operation of this system is shown diagramatically in Fig. 4-4. The input signal is processed by a linear time-invariant filter whose impulse response is again found from Eqs. (4-23) and (4-25). The filter output is then downshifted in frequency by ν_c. In this diagram the analytic representation is used for all signals. The multiplication of the analytic filter output by the cissoid $e^{-i2\pi\nu_c t}$ is therefore equivalent to multiplication of the real filter output by a cosine at ν_c followed by low pass filtering to give only the difference frequencies. This corresponds closely with the actual optical heterodyne operation: sum frequencies of the mixing of two optical distributions are themselves at optical frequencies and are not detected by the PMT. We note that ν_c is a <u>down</u>shift applied to the local oscillator optical distribution. If we should instead apply an

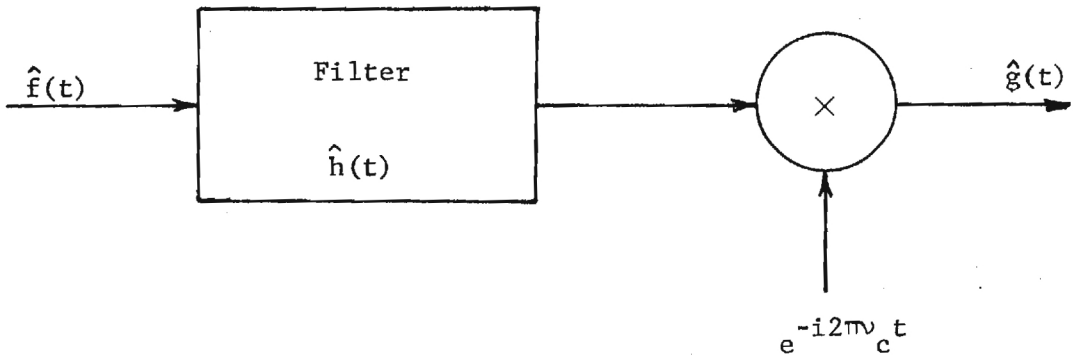


Figure 4-4. Diagram Model for Processor Operation with Frequency Shifted Local Oscillator.

upshift to the local oscillator, then ν_c is negative and the filter output is upshifted by an equivalent amount.

A uniform shift in the optical frequency of the local oscillator distribution produces a uniform shift in the frequency of output signal frequency components. If, however, the optical frequency of the local oscillator distribution changes as a function of position in the output plane, an entirely different type of signal processing occurs. The frequency components of the input signal are then processed with local oscillators of different frequencies. The amount of frequency shift is different for each component resulting in a rearrangement of the frequency components in the output signal. One method of introducing a local oscillator distribution of this type is to introduce a time-varying signal in the local oscillator arm, as shown in Fig. 4-5. An input mapping similar to that used for the input signal is employed and a lens is again used to form the spectral distribution of the local oscillator signal at the detector surface. The optical frequency of the local oscillator distribution at the PMT, like that of the signal distribution, varies linearly with position. Consider, for example, a local oscillator signal introduced by a moving slit as shown in Fig. 4-6. The spectral distribution for this signal (a moving impulse) is uniform at all frequencies. (The spectral distribution is a plane wave whose angle of incidence - assumed small - changes at a constant rate as the slit moves through the input plane.) Ignoring for the moment the finite extent of the input apertures of the lenses L_1 and L_2, the optical frequency vs. position for the local oscillator distribution and the input signal spectral distribution are shown in Fig. 4-7. A single Fourier component of the input signal at temporal frequency ν_1

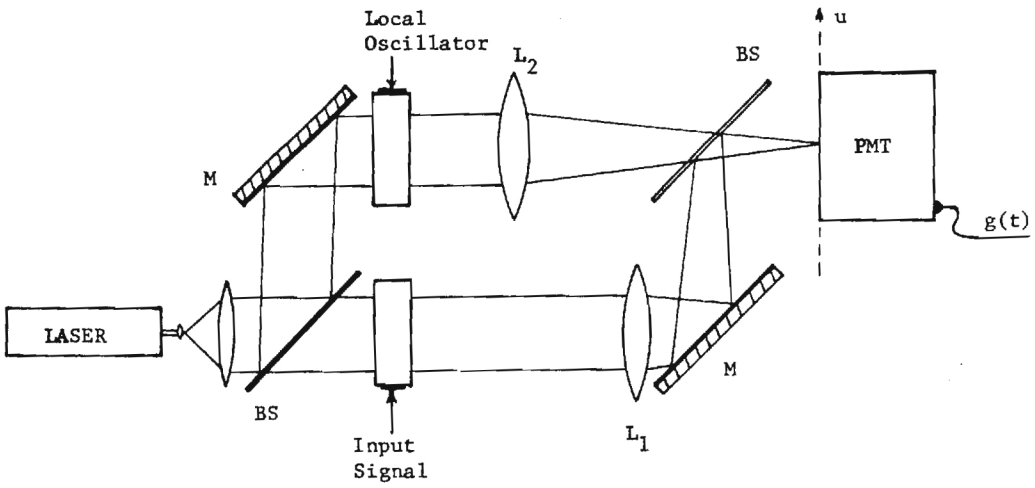


Figure 4-5. Optical Heterodyne Processing System with Introduction of a Local Oscillator Signal.

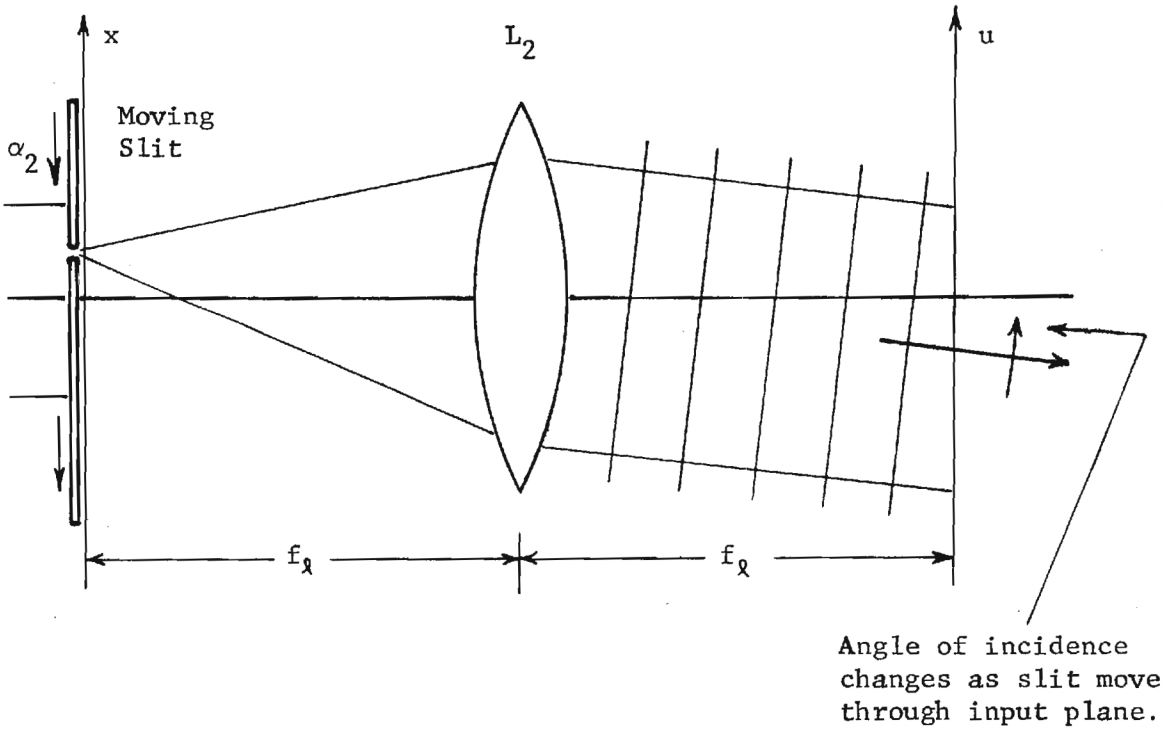


Figure 4-6. Impulse Local Oscillator Signal Introduced by a Moving Slit.

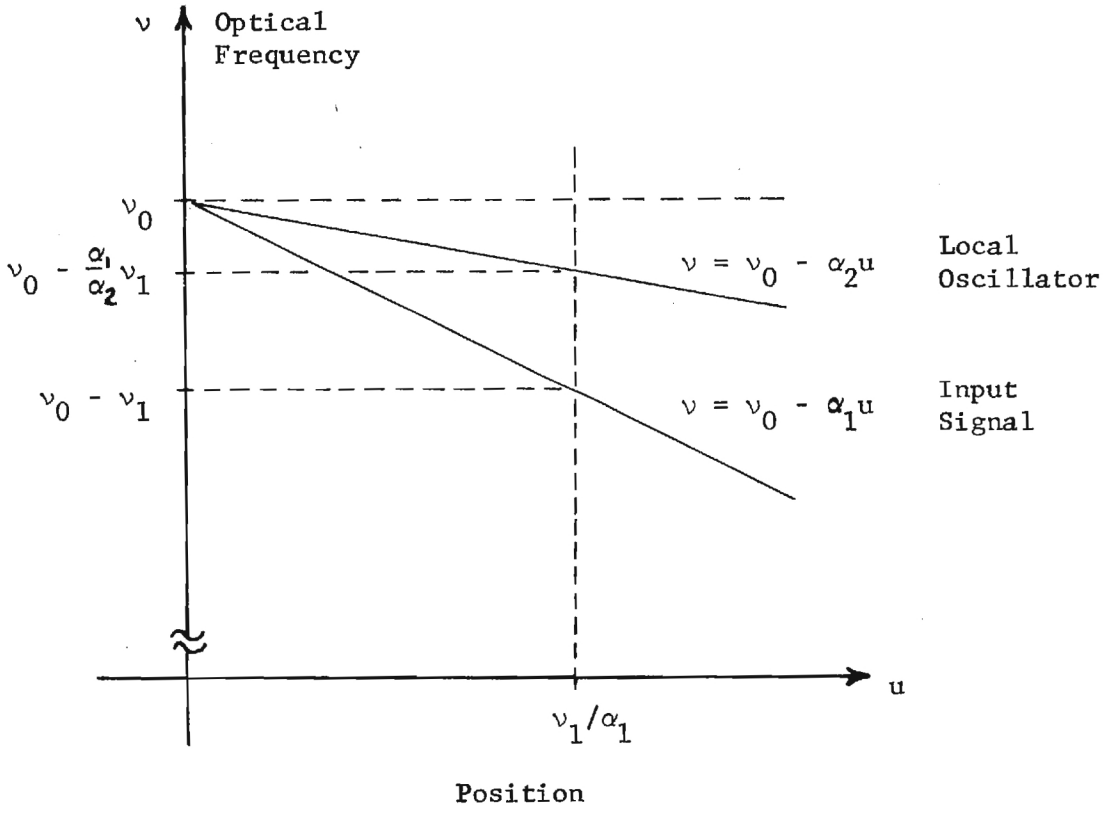


Figure 4-7. Optical Frequency vs. Position of Input Signal and Local Oscillator Spectral Distributions.

produces a spot of light in the spectral distribution at $u=\nu_1/\alpha_1$, where α_1 is the velocity associated with the input mapping of the input signal. The optical frequency of the light in this spot, downshifted by the signal frequency, is $\nu_0-\nu_1$. At the position $u=\frac{\nu_1}{\alpha_1}$ in the output plane, the light of the local oscillator distribution is at the optical frequency $\nu_0-\alpha_2\left(\frac{\nu_1}{\alpha_1}\right)$, where α_2 is the velocity of the moving slit. The output mapping then mixes these distributions producing an output sinusoidal signal at the difference frequency

$$\left(\nu_0 - \frac{\alpha_2}{\alpha_1}\right)\nu_1 \; - (\nu_0 - \nu_1) = \nu_1 - \frac{\alpha_2}{\alpha_1}\,\nu_1 = \left(\frac{\alpha_1-\alpha_2}{\alpha_1}\right)\nu_1. \qquad (4\text{-}28)$$

The various Fourier components of the input signal therefore appear in the output signal scaled in frequency by a factor $\left(\frac{\alpha_1-\alpha_2}{\alpha_1}\right)$ suggesting that a bandwidth compression or expansion can be performed.

The output of this processor is identical to what would be obtained with a noncoherent playback of a film strip recording with a moving slit, as shown in Fig. 4-8. In the noncoherent playback, we assume the intensity transmittance (not the amplitude transmittance) to be equal to the recorded signal. The film strip moves with velocity α_1 past a slit that is itself moving with velocity α_2. The light intensity transmitted through the slit and signal recording is collected by a detector, producing the output signal. Relative to the slit, the playback velocity for the recorded signal is $\alpha_1-\alpha_2$, and the Fourier components of the input are therefore all scaled by the factor $\left(\frac{\alpha_1-\alpha_2}{\alpha_1}\right)$, however it takes the moving slit $\left(\frac{\alpha_1}{\alpha_1-\alpha_2}\right)$ times as long to scan a given length of recorded signal as would a stationary slit. The frequency scaling is thus seen to result from a change in the signal

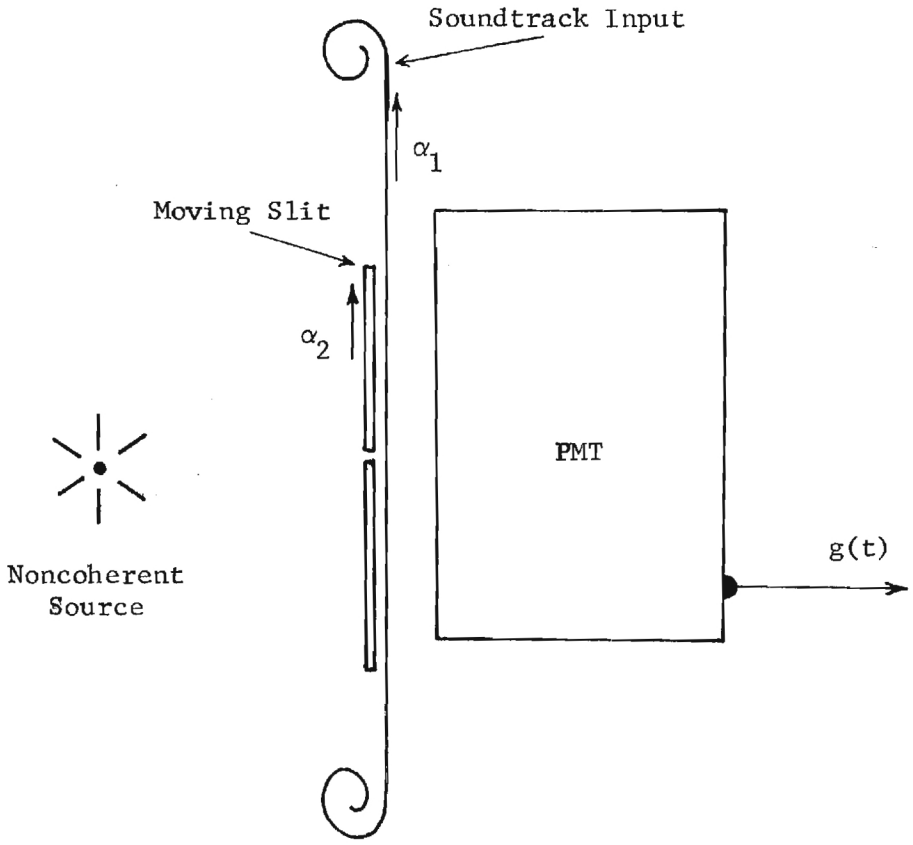


Figure 4-8. Equivalent Noncoherent System for Heterodyne Processor with Impulse Local Oscillator.

time base. If we include the effect of finite input apertures, the moving slit scans the recorded signal only during the time it moves through the fixed window. To obtain a continuous output, additional moving slits must be introduced, one entering the fixed window as the previous slit leaves. The frequency scaling of the recorded signal then proceeds on a section by section basis. Dennis Gabor first described such a processing technique for signal bandwidth compression or expansion in his classic 1946 paper, "Theory of Communication." [45]* There are several fundamental problems with this processing technique all discussed by Gabor. In the heterodyne system, these problems can be attributed to the impulse train nature of the signal introduced to produce the local oscillator distribution. We therefore turn our attention to an alternative processing technique that employs a more general local oscillator signal.

If we introduce an arbitrary signal, r(t), in the local oscillator arm of the system, the spectral distribution that appears in the output plane as the local oscillator distribution will retain (approximately) the linear optical frequency vs. position characteristic shown in Fig. 4-7. The distribution itself, however, will no longer be uniform in amplitude and phase along the u-axis as in the case of an impulse local oscillator signal. The system will still perform a frequency scaling of the input signal but there will also be a certain transfer function associated with the operation specified by the spectral distribution of the local oscillator signal. We must be careful in defining this transfer function because

---

*Gabor describes what has come to be known as the "sample-and-discard," or "sampling" method of time/bandwidth compression. For additional discussions, see, e.g. Reference [54].

at any time t, the transfer characteristics will be specified only by the spectrum of the local oscillator signal present in the window of the local oscillator spectrum analyzer. Furthermore, these transfer characteristics will apply only to the portion of the input signal under the window of the input signal spectrum analyzer at that time. Any description of system operation must therefore take into consideration the temporal and frequency characteristics of both input signal and local oscillator.

We need first to reformulate the analytical description for the spectrum analyzer output distribution to state explicitly the time-frequency characteristics of the distribution. From Eq. (4-3) we rewrite the input amplitude distribution for the input signal spectrum analyzer as

$$i(x,y;t) = C_a \hat{f}\left(\frac{x+\alpha_1 t}{\alpha_1}\right) w_1(x)w(y), \tag{4-29}$$

where we have ignored the bias term that results in the d.c. spot and we have for convenience assumed a separable input window $w(x,y) = w_1(x)w(y)$. For the present, we are interested in the light distribution only along the u-axis in the output plane of the spectrum analyzer:

$$O(u;t) \triangleq O(u,0;t) = C_a \int_{-\infty}^{\infty} f\left(\frac{x+\alpha_1 t}{\alpha_1}\right) w_1(x) e^{-i2\pi ux} dx. \tag{4-30}$$

Making the variable substitution $\tau=(x+\alpha_1 t)/\alpha_1$; this distribution reduces to

$$O(u;t) = C_a F_w(u;t) e^{+i2\pi\alpha_1 ut} \tag{4-31}$$

where

$$F_w(u;t) = \int_{-\infty}^{\infty} f(\tau) w_1[\alpha_1(\tau-t)] e^{-i2\pi\alpha_1 u\tau} \alpha_1 d\tau \tag{4-32}$$

is the spectrum of the signal under the analyzer window at time t, i.e. the short-time spectrum of the input signal evaluated for $\nu=\alpha_1 u$. To introduce the local oscillator signal we perform an input mapping and an optical mapping (spectrum analysis) similar to those applied to the input signal. The local oscillator spectral distribution can therefore be expressed as in Eq. (4-31),

$$\hat{r}(u;t) = C_a R_w(u;t) e^{+i2\pi\alpha_2 ut} \tag{4-33}$$

where

$$R_w(u;t) = \int_{-\infty}^{\infty} \hat{r}(\tau) w_2[\alpha_2(\tau-t)] e^{-i2\pi\alpha_2 u\tau} \alpha_2 d\tau. \tag{4-34}$$

Here, $w_2(x)$ is the input window for the local oscillator spectrum analyzer and α_2 is the velocity associated with the input mapping of the local oscillator signal. The quantity $R_w(u;t)$ is the short-time spectrum for the local oscillator signal evaluated for $\nu=\alpha_2 u$.

When the input signal and local oscillator spectral distributions are mixed and detected at the PMT, the system output signal is

$$\hat{g}(t) = C_a^2 \int_0^{\infty} F_w(u;t) R_w^*(u;t) e^{i2\pi(\alpha_1-\alpha_2)ut} du. \tag{4-35}$$

Expressing the system output in terms of the frequency $\nu=\alpha_1 u$ of the input signal short-time spectrum we obtain

$$\hat{g}(t) = \frac{1}{\alpha_1} C_a^2 \int_0^\infty F_w\left(\frac{\nu}{\alpha_1};t\right) R_w^*\left(\frac{\nu}{\alpha_1};t\right) e^{i2\pi\left[\frac{\alpha_1-\alpha_2}{\alpha_1}\right]\nu t} d\nu. \qquad (4\text{-}36)$$

From this expression it is clear how the short-time spectrum of the local oscillator establishes the transfer characteristics for the input signal short-time spectrum. The expression also shows explicitly that the frequencies of the input signal short-time spectrum will be scaled in the system output by the factor $\left(\frac{\alpha_1-\alpha_2}{\alpha_1}\right)$. From this frequency scaling factor, we can define a compression factor equal to the reciprocal of the scaling factor, i.e.

$$\gamma = \frac{\alpha_1}{\alpha_1-\alpha_2} . \qquad (4\text{-}37)$$

In terms of this compression factor, the system performs a bandwidth compression if $\gamma>1$ and a bandwidth expansion if $\gamma<1$. Clearly, if $\alpha_1=\alpha_2$, the compression factor becomes infinite. The meaning of an infinite γ can be seen directly from Eq. (4-36) for $\alpha_1=\alpha_2$:

$$\hat{g}(t) = \frac{1}{\alpha_1} C_a^2 \int_0^\infty F_w\left(\frac{\nu}{\alpha_1};t\right) R_w\left(\frac{\nu}{\alpha_1};t\right) d\nu. \qquad (4\text{-}38)$$

After substitution of the expressions for the short-time spectra and some analytical manipulation, the output expression will become

$$\hat{g}(t) = \alpha_1 C_a^2 \int_{-\infty}^\infty f(\tau) r^*(\tau) W_e(t-\tau) d\tau, \qquad (4\text{-}39)$$

where

$$w_e(t) = w_1[-\alpha_1 t]\, w_2^*[-\alpha_1 t]. \tag{4-40}$$

Equation (4-40) indicates that the system output with $\alpha_1=\alpha_2$ (γ infinite) is equal to the output of a filter whose impulse response is $W_e(t)$ when the input is the product function $f(t)r^*(t)$. The input windows $W_1(\cdot)$ and $W_2(\cdot)$ used in the two system spectrum analyzers are usually rectangular or slightly smoothed, such as a Gaussian or a raised cosine window. In such cases, the structure of the impulse response described in Eq. (4-40), which is equal to the product of the two time-windows, will be that of a low pass filter. The output is then a low pass filtered version of the product of input signal and local oscillator.

The Self-Product Terms and System Linearity

In the previous sections we have ignored the self-product terms in the system output and described system operation in terms of the linear operation Eq. (4-14). We now consider these terms in more detail. We recall from Eqs. (4-9) through (4-12) that the total system output is

$$s(t) \propto s_0(t) + s_r(t) + g(t), \tag{4-41}$$

where $g(t)$ is assumed to be the desired output signal and $s_0(t)$ and $s_r(t)$ are the self-product terms produced due to the square-law nature of the photodetector response. In the processor of the previous section where we observe the output only along the u-axis, these self-product terms have the form

$$\hat{s}_0(t) = \int_0^\infty \hat{O}(u,t)O^*(u,t)\,du, \tag{4-42}$$

and

$$\hat{s}_r(t) = \int_0^\infty \hat{r}(u,t)r^*(u,t)\,du, \tag{4-43}$$

where the " ^ " again indicates the analytic signal. Substituting Eq. (4-31) for $O(u,t)$ in the expression for $s_0(t)$ we obtain

$$\hat{s}_0(t) = C_a^2 \int_0^\infty F_w(u;t)F_w^*(u,t)\,du. \tag{4-44}$$

This signal has the same form as the cross-product signal Eq. (4-35) in the case where $R_w(u;t) = F_w(u;t)$ and with $\alpha_1=\alpha_2$; i.e. with γ infinite. The signal can then be expressed in a form similar to Eq. (4-39),

$$\hat{s}_0(t) = \alpha_1 C_a^2 \int_{-\infty}^\infty |\hat{f}(\tau)|^2 w_{e_1}(t-\tau)\,d\tau, \tag{4-45}$$

with

$$w_{e_1}(t) = |w_1[-\alpha_1 t]|^2. \tag{4-46}$$

Similarly, in the case of a local oscillator distribution,

$$\hat{s}_r(t) = \alpha_2 C_a^2 \int_{-\infty}^\infty |\hat{r}(\tau)|^2 w_{e_2}(t-\tau)\,d\tau, \tag{4-47}$$

with

$$w_{e_2}(t) = |w_w(-\alpha_2 t)|^2. \tag{4-48}$$

Both self-product terms are therefore low-frequency signals that can be removed by high-pass filtering the system output. If such a filtering would remove portions of the desired output signal, this latter component can be upshifted to a higher frequency band by uniformly shifting the optical frequency of the light illuminating the local oscillator signal input device. The self-products can then be removed by filtering and the frequency-scaled output recovered in the desired frequency band by a post-detection electronic heterodyne operation.

For a more physical interpretation of the intermodulation terms, we reintroduce the variable x in Eq. (4-45) by making the substitution $x=\alpha_1(\tau_1-t)$. Then, $\hat{s}_0(t)$ is given by

$$\hat{s}_0(t) = C_a^2 \int_{-\infty}^{\infty} \left|\hat{f}\left(\frac{x+\alpha_1 t}{\alpha_1}\right) w_1(x)\right|^2 dx. \tag{4-49}$$

Similary, we replace x with $x=\alpha_2(\tau-t)$ in Eq. (4-47), obtaining

$$\hat{s}_r(t) = C_a^2 \int_{-\infty}^{\infty} \left|\hat{r}\left(\frac{x+\alpha_2 t}{\alpha_2}\right) w_2(x)\right|^2 dx. \tag{4-50}$$

The integrands of Eqs. (4-49) and (4-50) are the non-bias portions of the light irradiance distributions passing through the signal aperture and local oscillator aperture respectively. These expressions then indicate that the intermodulation terms are measures of the variations in the light flux passing through the input signal aperture and the local oscillator aperture induced by the input signal and local oscillator signal. If we

assume a rectangular or Gaussian aperture transmittance, then several cycles of the higher frequency components of $\hat{f}(\cdot)$ and $\hat{r}(\cdot)$ will be present in the window at any time, and the integrals of Eqs. (4-49) and (4-50) will represent the average of these components over several cycles. This average will be essentially constant for the higher frequency terms. The only significant temporal variations of the signal-induced light flux comes from low frequency components where the period of variation is a large fraction of the effective width of the input aperture.

Frequency-Variant Optical Heterodyne Signal Processing

The optical heterodyne processing system described in the previous section can certainly be viewed as a frequency-variant processing system, since input signals of different frequencies are processed with different local oscillators. However, because conventional spectrum analyzers are used, the frequency vs. position characteristics are linear for both spectral distributions mixing in the system output plane. The operation performed by this system is therefore a uniform (i.e. frequency-independent) frequency scaling operation. In Chapter II we showed that, by exploiting the second degree of freedom available in an astigmatic system, a spectrum analyzer could be devised in which the frequency vs. position characteristics and the time-frequency resolution characteristics within the spectral display are frequency-dependent. We show in this section that by using frequency-variant spectrum analyzers in a heterodyne signal processing system the processing capabilities of the system for 1-D signals can be extended to include general frequency-domain redistributions of signal information.

The optical configuration for a general frequency-variant heterodyne

processing system is shown in Fig. 4-9. The input signal and the local oscillator signal are introduced in separate arms of the system by the conventional input mapping. The optical distributions produced by the input mappings are therefore

$$i_s(x,y;t) = C_a \hat{f}\left(\frac{x+\alpha_1 t}{\alpha_1}\right) w_1(x,y) \tag{4-51}$$

in the signal arm, and

$$i_r(x,y;t) = C_a \hat{r}\left(\frac{x+\alpha_2 t}{\alpha_2}\right) w_2(x,y) \tag{4-52}$$

in the local oscillator arm where we have ignored the bias amplitude in each distribution. The optical mapping performed on the input signal distribution is a frequency-variant spectrum analysis using a system like that described in Chapter II, with the output passing through a vertical slit onto the PMT. We assume a spatial frequency mapping characteristic $u=m_1(y)$ in this system with mapping slit narrow enough that $w_1(x,y)$ is the equivalent window for the operation. Then, from Eq. (2-35), the input signal spectral distribution along the y-axis is

$$O(y;t) = C_a \int_{-\infty}^{\infty} \hat{f}\left(\frac{x+\alpha_1 t}{\alpha_1}\right) w_1(x,y) e^{-i2\pi m_1(y)x} dx, \tag{4-53}$$

which, with an appropriate variable substitution, we can simplify to

$$O(y;t) = C_a F_w(m_1(y),y;t) e^{i2\pi\alpha_1 m_1(y)t}, \tag{4-54}$$

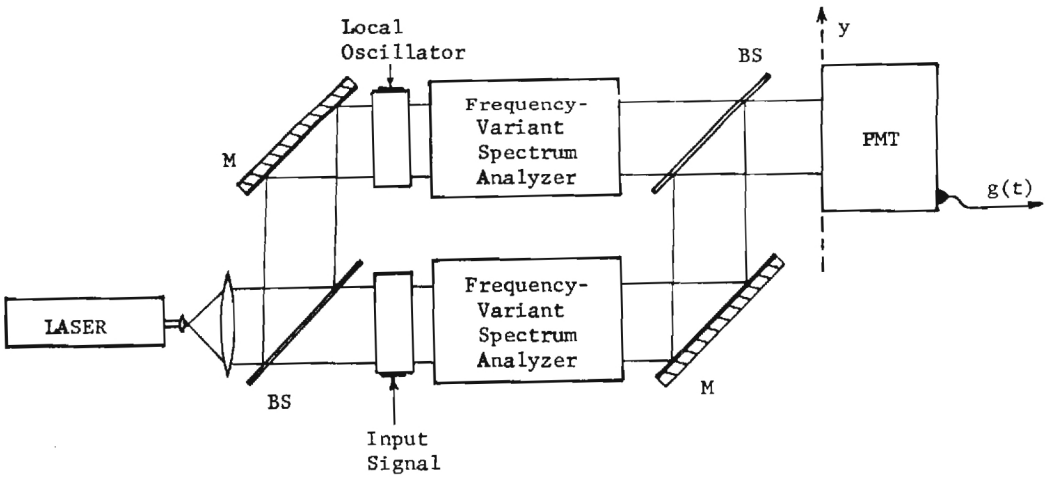


Figure 4-9. The Frequency-Variant Optical Heterodyne Signal Processing System.

where

$$F_w(m_1(y),y;t) = \int_{-\infty}^{\infty} \hat{f}(\tau)\alpha_1 w_1[\alpha_1(\tau-t),y]e^{-i2\pi\alpha_1 m_1(y)\tau}d\tau \qquad (4\text{-}55)$$

is the frequency-variant short-time spectrum of the input signal. For complete generality, we assume that the local oscillator input distribution is also mapped by a frequency-variant spectrum analyzer with spatial frequency mapping characteristic $u=m_2(y)$.* With $w_2(x,y)$ as the equivalent window for this mapping, the spectral distribution of the local oscillator system is

$$r(y;t) = C_a R_w(m_2(y),y;t)e^{i2\pi\alpha_2 m_2(y)t}, \qquad (4\text{-}56)$$

where

$$R_w(m_s(y),y;t) = \int_{-\infty}^{\infty} \hat{r}(\tau)\,\alpha_2 w_2\,[\alpha_2(\tau-t),y]e^{-i2\pi\alpha_2 m_2(y)\tau}d\tau \qquad (4\text{-}57)$$

is the frequency-variant short-time spectrum of the local oscillator signal.

The input signal and local oscillator distributions are superimposed at the PMT, which responds with the system output signal

$$\hat{g}(t) = C_a^2 \int_0^{\infty} F_w(m_1(y),y;t)R_w^*(m_2(y),y;t)e^{i2\pi[\alpha_1 m_1(y)-\alpha_2 m_2(y)]t}dy. \qquad (4\text{-}58)$$

---

*General short-time spectral mappings can be performed with only one nonlinear spatial frequency mapping.

In this expression, we note again that the local oscillator short-time spectrum establishes the transfer characteristics for the short-time spectrum of the input signal. In this case, however, the short-time spectra are frequency-variant: the system transfer characteristics are more general, and the frequency scaling becomes a nonuniform frequency mapping. In terms of the frequency ν_s of the input signal short-time spectrum, the system output takes on the rather formidable analytic form

$$\hat{g}(t) = C_a^2 \int_{-\infty}^{\infty} F_w[\nu_s/\alpha_1, y_s; t] R_w^*[m_2(y_s), y_s; t] e^{i2\pi[\nu_s - \alpha_2 m_2(y_s)]t} \frac{d\nu_s}{\alpha_1 m_1'(y_s)}, \tag{4-59}$$

where

$$y_s = m_1^{-1}(\nu_s/\alpha_1), \tag{4-60}$$

and

$$m_1'(y) = \frac{d}{dy} m_1(y). \tag{4-61}$$

The factor $1/\alpha_1 m_1'(y_s)$ is present in the integral expression to compensate for the non-uniform packing of the spectral distribution along the y-axis. From this expression the frequency mapping relationship for the system is readily determined:

$$\nu_s \rightarrow \nu_s - \alpha_2 m_2(y_s). \tag{4-62}$$

A uniform system compression factor cannot be defined for this system

because the frequency mapping relationship is no longer a simple frequency scaling. We can, however, define a compression function equal to the ratio of input frequency to the output frequency,

$$\gamma(\nu_s) = \frac{\nu_s}{\nu_s - \alpha_2 m_2(y_s)} \quad . \tag{4-63}$$

Using this compression function it is then possible to define the overall system compression factor as the ratio of the input signal bandwidth to the output signal bandwidth. If ν_{min} and ν_{max} are the minimum and maximum frequencies of the input signal, then the system compression factor is

$$\gamma_{sys} = \frac{\nu_{max} - \nu_{min}}{\int_{\nu_{min}}^{\nu_{max}} \gamma(\nu_s)\, d\nu_s} \quad . \tag{4-64}$$

Interpretation of System Operation With Time-Frequency Signal Components

The expressions derived for the operation of the general frequency-variant processing system, although mathematically correct, are largely uninformative. The relationship for the mapping of short-time spectral frequencies was easily determined from Eq. (4-59), but further interpretation of system operation is hidden by the complexity of the expression. If we are to properly interpret system operation, a mapping relationship should be established that describes the effect of the processing operation, in both time and frequency, on basic components of the input signal. This task is complicated by the presence of the local oscillator signal - the system operation can be viewed as a very general mixing of two signals

distinguished only by the labels we have placed on them. The mapping relationship must therefore describe the operation in terms of a mixing of basic components from two arbitrary signals.

The question now arises as to what basic components to use in the mapping relationship. Neither frequency components (sinusoids) nor temporal components (delta-functions) are wholely adequate for this representation. This is suggested by the operation of the frequency-variant optical mappings used in the system, in which different frequencies are processed with different temporal characteristics, and by the output expression Eq. (4-59), in which the system frequency transfer characteristics are seen to change with time. We turn therefore to basic components that are localized in both time and frequency. Gabor first used a superposition of elementary time-frequency components to represent an arbitrary signal. [45] His "elementary signals" were cissoidal pulses with Gaussian time envelopes. Lerner extended Gabor's technique of signal representation to include cissoidal pulses with non-Gaussian temporal envelopes or, in general, elementary time-frequency components of the form [53]

$$\varepsilon_{mn}(t) = w(t-t_m)e^{i2\pi\nu_n t}. \tag{4-65}$$

In the frequency domain, this signal has the form

$$E_{mn}(\nu) = W(\nu-\nu_n)e^{-i2\pi(\nu-\nu_n)tm} \tag{4-66}$$

where

$$W(\nu) = \mathcal{F}_t\{w(t)\}. \tag{4-67}$$

The duration of the elementary signal, Δt, can be taken as the effective width of $w(t)$, while the frequency spread, $\Delta\nu$, can be described as the effective width of $W(\nu)$. The elementary signal can therefore be viewed as occupying an area $\Delta t\Delta\nu$ in a time-frequency diagram centered at time t_m and frequency ν_n. Using a root-mean-square definition for effective width, Gabor was able to show that the area of a time-frequency component was restricted by an uncertainty relationship,

$$\Delta t\Delta\nu \geq \frac{1}{2} , \tag{4-68}$$

where the minimum area of $\frac{1}{2}$ is achieved only for elementary signals with Gaussian envelopes. The uncertainty product, he noted, relates to the inability of any receiver to define simultaneously time and frequency in an exact way.

An arbitrary signal $f(t)$ can be approximated by a superposition of elementary time-frequency components:

$$f(t) \simeq \sum_m \sum_n a_{mn}\, w(t-t_m) e^{i2\pi\nu_n t} . \tag{4-69}$$

Component spacing in time and frequency is usually taken to be uniform, i.e. $t_m = m\Delta t$ and $\nu_n = n\Delta\nu$. The accuracy of the approximation depends on both the choice of envelope $w(t)$ and on the choice of Δt and $\Delta\nu$. (There is no clear cut "best" choice for any of these quantities.) The expansion of Eq. (4-69) covers the entire time-frequency area for the original signal, as shown in Fig. 4-10. In this figure, each rectangle represents a single time-frequency component; the amount of shading in the rectangle indicates

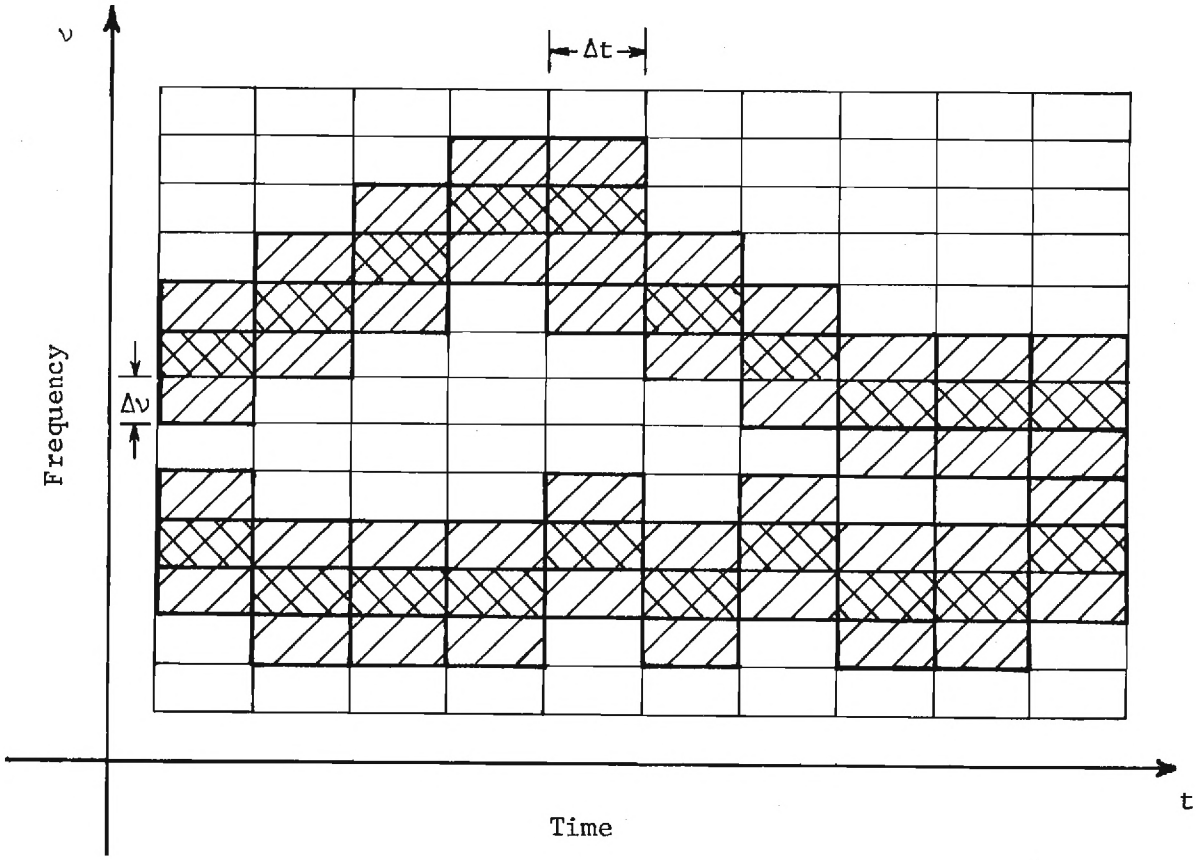


Figure 4-10. Time-Frequency Signal Diagram.

the component amplitude. In general, the elementary time-frequency components used in such an expansion are neither orthogonal nor do they form a complete set. Nevertheless, such expansions can be highly accurate. For example, Markel and Carey have demonstrated that the expansion of speech waveforms in terms of Gaussian elementary signals, although not exact, is virtually indistinguishable from the original speech, both qualitatively and quantitatively [56].

In Gabor's original expansion with elementary signals of minimum area, each elementary signal was viewed as carrying an elementary quantum (or logon) of signal information observable in time t. Each elementary signal of the general expansion can be associated with an elementary piece of signal information observable by a receiver or processor with time window w(t). In a system operating with this time window, signal expansion in terms of elementary signals with time envelopes equal to w(t) is therefore a natural choice.

We now consider an interpretation of the general system processor in terms of time-frequency components for the special case of frequency-invariant operation. In the frequency-invariant case, the input windows for both the input signal and the local oscillator signal are frequency-independent and the frequency mapping characteristics are linear: $m_1(y)=y$ and $m_2(y)=y$. We now assume that an adequate expansion for the input signal and local oscillator can be found in terms of the elementary time-frequency components associated with their respective input windows, i.e. we assume the expansion

$$\hat{f}(t) = \sum_m \sum_n a_{mn} w_1[\alpha_1(t-t_m)] e^{i2\pi\nu_n t} \qquad (4-70)$$

for the input signal and the expansion

$$\hat{r}(t) = \sum_j \sum_k b_{jk} w_2[\alpha_2(t-t_j)]e^{i2\pi\nu_k t} \tag{4-71}$$

for the local oscillator to be correct. In these expressions we recall that $w_1(x)$ and $w_2(x)$ are the input aperture transmittances for the input signal and local oscillator respectively and that α_1 and α_2 are the velocities associated with the input mappings of those two signals.

The optical mapping performed on the input signal is a conventional Fourier transform where signal components are mapped in frequency along the y-axis according to the relation $\nu=\alpha_1 y$. The time-frequency component at $t=t_m$ and $\nu=\nu_n$ appears in the transform distribution in the small spatial interval $\Delta\nu/\alpha_1$ centered about $y=\nu_n/\alpha_1$ and during the time interval Δt centered about $t=t_m$. Observing the transform distribution of the input signal along the y-axis is therefore equivalent to observing the time-frequency diagram of the input signal (with frequency axis scaled by $1/\alpha_1$) through a narrow vertical slit that slides along the time axis, as shown in Fig. 4-11. In order to describe analytically the distribution observed through the slit, as a function of space and time, we define the function

$$\varepsilon_{mn}(y,t) \triangleq w_1[\alpha_1(t-t_m)]W_1(y-\nu_n/\alpha_1)e^{i2\pi\nu_n t}, \tag{4-72}$$

which describes the distribution of a single time-frequency component at $t=t_m$ and $\nu=\nu_n$. The quantity $\varepsilon_{mn}(y,t)$ expresses explicitly both the time envelope, $w_1[\alpha_1(t-t_m)]$, and the frequency envelope (along the y-axis), $W_1(y-\nu_n/\alpha_1)$, of the component as well as the shift in optical frequency

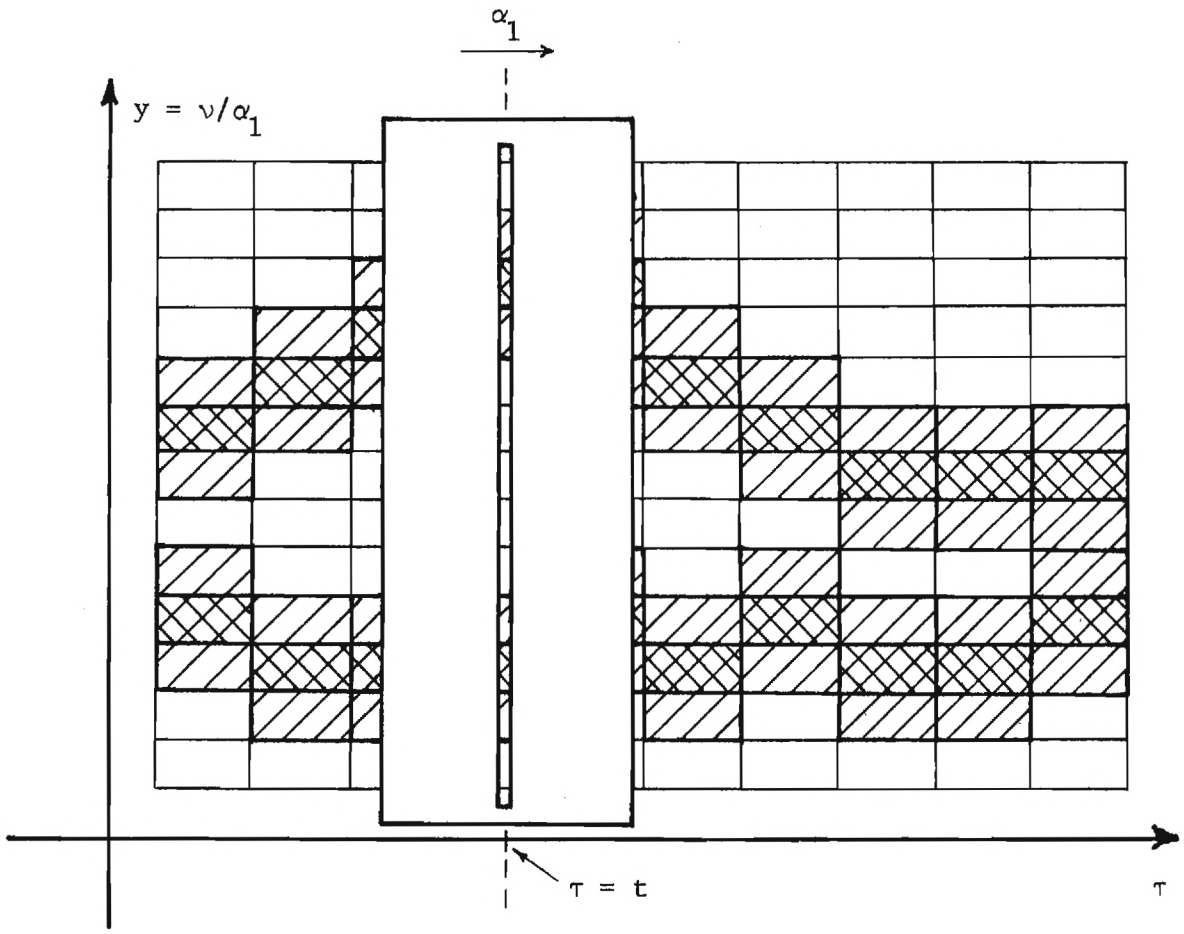


Figure 4-11. Scaled Version of Time-Frequency Diagram Viewed through Moving Slit — Equivalent to Transform Distribution of the Input Signal Along the y-axis at the PMT.

of the light in the distribution. In terms of this elementary distribution, the transform distribution of the input signal is

$$O(y;t) \simeq \sum_m \sum_n a_{mn} w_1[\alpha_1(t-t_m)] W_1(y-\nu_n/\alpha_1) e^{i2\pi\nu_n t}. \tag{4-73}$$

We should recover the input signal directly from this distribution by introducing a plane wave local oscillator. Assuming the plane wave to be normally incident and of unit amplitude, the system output would be

$$\hat{g}(t) = \int_0^\infty O(y;t)\,dy = \sum_m \sum_n a_{mn} w_1[\alpha_1(t-t_m)] e^{i2\pi\nu_n t} \cdot \{ \int_0^\infty W_1(y-\nu_n/\alpha_1)\,dy \}, \tag{4-74}$$

or

$$\hat{g}(t) = w_1(0) \cdot \sum_m \sum_n a_{mn} w_1[\alpha_1(t-t_m)] e^{i2\pi\nu_n t}. \tag{4-75}$$

Comparing Eqs. (4-70) and (4-75), we see that the input signal is indeed recovered.

The components of the local oscillator are also mapped into a spectral distribution along the y-axis, with frequencies mapped according to $\nu=\alpha_2 y$. The spectral distribution can be expressed in terms of elementary components similar to those for the input signal, i.e.

$$r(y;t) = \sum_j \sum_k b_{jk} w_2[\alpha_2(t-t_j)] W_2\left(y - \frac{\nu_k}{\alpha_2}\right) e^{i2\pi\nu_k t}. \tag{4-76}$$

The mixing of the two spatial distributions then produces the system output

$$\hat{g}(t) = \int_0^\infty O(y;t)\, r^*(y;t)\, dy$$

$$= \sum_j \sum_k \sum_m \sum_n a_{mn} b^*_{jk} c_{kn} w_1[\alpha_1(t-t_m)] w_2^*[\alpha_2(t-t_j)] e^{i2\pi(\nu_n-\nu_k)t}, \quad (4\text{-}77)$$

where

$$c_{kn} = \int_0^\infty W_1(y - \frac{\nu_n}{\alpha_1}) W_2^*(y - \frac{\nu_k}{\alpha_2})\, dy. \quad (4\text{-}78)$$

Equation (4-77) indicates how the various time-frequency components of the input signal and the local oscillator mix in the optical processing system. Again, interpretation is impeded by the complexity of the expression. Nevertheless, we can obtain some useful information from this expression by considering the mapping of a single input signal component at time t_m and frequency ν_n. From Eq. (4-77) we obtain

$$a_{mn} w_1[\alpha_1(t-t_m)] e^{i2\pi\nu_n t} \rightarrow \sum_j \sum_k a_{mn} b^*_{jk} c_{kn} w_1[\alpha_1(t-t_m)]$$

$$\cdot\, w_2[\alpha_2(t-t_j)] e^{i2\pi(\nu_n-\nu_k)t} \quad (4\text{-}79)$$

It is clear from this expression that an input component is mapped into several components of the output signal due to the mixing with different components of the local oscillator. However, the major contribution to the output from this input signal component will result from the mixing with the one local oscillator component that occurs at approximately the

same time and at approximately the same position along the y-axis. Thus the major contribution results from the local oscillator component at time $t_j \simeq t_m$ and at frequency ν_k for which $\frac{\nu_k}{\alpha_2} \simeq \frac{\nu_n}{\alpha_1}$. The nominal component mapping relationship can then be expressed as

$$a_{mn} w_1[\alpha_1(t-t_m)] e^{i2\pi\nu_n t} \rightarrow \bar{C}(t_m,\nu_n) a_{mn} w_1[\alpha_1(t-t_m)]$$

$$\cdot w_2[\alpha_2(t-t_m)] e^{i2\pi\left(\frac{\alpha_1-\alpha_2}{\alpha_1}\right)\nu_n t}, \qquad (4\text{-}80)$$

where the factor $\bar{C}(t_m,\nu_n)$ is determined by the amplitude of the local oscillator component at $t_j \simeq t_m$ and $\nu_k \simeq \frac{\alpha_2}{\alpha_1}\nu_n$. This relationship indicates that the input signal components are mapped with time-frequency resolution characteristics determined by the product window $w_1[\alpha_1 t] w_2[\alpha_2 t]$ and that the component frequencies are scaled by a factor $\left(\frac{\alpha_1-\alpha_2}{\alpha_1}\right)$.

Expressing the system output in terms of the nominal mapping relationship we obtain

$$\hat{g}(t) = \sum_m \sum_n \bar{C}(t_m,\nu_n) a_{mn} w_1[\alpha_1(t-t_m)] w_2[\alpha_2(t-t_n)] e^{i2\pi\left(\frac{\alpha_1-\alpha_2}{\alpha_1}\right)\nu_n t}. \qquad (4\text{-}81)$$

The quantity $\bar{C}(t_m,\nu_n)$ can be viewed as the system transfer function in both time and frequency and is established directly from the time-frequency characteristics of the local oscillator signal. As was mentioned, the time-frequency resolution characteristics in the system output are established by the product window.

A complete analytic description for the system output in the general

frequency-variant processor, analogous to Eq. (4-77), becomes extremely complex due to the use of separate input windows for components of different frequencies and the nonlinear frequency mapping relationships. The nominal mapping relationship can, however, be determined in the same manner as in the frequency-invariant case. In a frequency-variant processor, the input signal components are mapped in frequency onto the y-axis according to the relationship $\nu=\alpha_1 m_1(y)$. The input signal is processed with an input window whose characteristics are, through y, frequency dependent. The input signal component at the time t_m and frequency ν_n has the form

$$a_{mn} w_1[\alpha_1(t-t_m), y_n] e^{i2\pi\nu_n t}, \tag{4-82}$$

where y_n is found from

$$\nu_n = \alpha_1 m_1(y_n). \tag{4-83}$$

The local oscillator component mixing with this signal component has frequency $\nu_k=\alpha_2 m_2(y_n)$ and is processed by the time window $w_2[\alpha_2 t, y_n]$. The nominal mapping relationship is then given by

$$a_{mn} w_1[\alpha_1(t-t_m), y_n] e^{i2\pi\nu_n t} \rightarrow \bar{C}(t_m, \nu_n) a_{mn} w_1[\alpha_1(t-t_m)]$$

$$\cdot\, w_2[\alpha_2(t-t_m), y_n] e^{i2\pi[\nu_n-\alpha_2 m_2(y_n)]t}, \tag{4-84}$$

where again $\bar{C}(t_m, \nu_n)$ is determined by the amplitude of the local oscillator component at $t_j \simeq t_m$ and $\nu_k \simeq \alpha_2 m_2(t_n)$. From the nominal mapping relationship,

we see that components of the input signal are processed in the general system with frequency-dependent time-frequency resolution characteristics. Assuming the mapping function $m_1(\cdot)$ has a unique inverse, $m_1^{-1}(\cdot)$, we also find that the components are redistributed in the frequency domain according to

$$\nu_n \rightarrow \nu_n - \alpha_2 m_2[m_1^{-1}(\nu_n/\alpha_1)]. \tag{4-85}$$

Discussion

In this chapter, we have presented the essential aspects of optical heterodyne signal processing systems for 1-D signals. We discussed the basic properties of the conventional heterodyne processors exemplified by the system investigated at Zenith Corporation by Whitman, Korpel, and Lotsoff. We then described extensions of the fundamental systems to more general signal processing applications characterized by the mapping of short-time signal components. The general processing system, whose operation is based on the general mapping capabilities of the frequency-variant spectrum analyzers, displays a unique flexibility among systems of this class. In this system, general nonlinear frequency mapping characteristics are easily implemented and controlled by the design of the mapping slits of the spectrum analyzers. Additionally, the use of varying width input apertures allows the control of system time-frequency resolution characteristics as a function of frequency.

In a general processing operation (especially a bandwidth compression), certain components of the input signal could be mapped so as to be

indistinguishable in the system output. (A receiver of the output signal could not uniquely recover the original components.) Recalling that we can view each time-frequency component as carrying an elementary piece of signal information, we can then associate a loss of signal information with the loss of signal components in the processing operation. The frequency-dependence of the mapping operation in the frequency-variant processor indicates that the loss of signal components and therefore the loss of signal information can be nonuniform over the signal bandwidth.

These general capabilities of the frequency-variant heterodyne processor suggest its applicability to the processing of audio signals where the human ear is the final receiver. Gabor showed that the ear's ability to distinguish elementary signals decreases nonlinearly with increasing component frequency [45]. The frequency-variant processor could then be used to minimize preceived information loss in a bandwidth compression of audio signals by compressing high frequency components more than low frequency components. Also, the control of resolution characteristics as a function of frequency could possibly be used to match the resolution capabilities of the ear. Finally, the general nature of the frequency mapping characteristics suggests the potential applicability of this system to the correction of frequency-domain signal distortions such as those observed in speech in a helium atmosphere [57].

In addition to the applicability in the area of audio signal processing, the frequency-variant processor should be a useful tool in the processing of general communication signals. In television, radar, and sonar signals, the quasi-periodic nature of the signals results in spectral information concentrated in frequency bands separated by regions relatively

free of spectral content. The optical processor can be designed to process only those regions where spectral information exists, ignoring the separating regions. Bandwidth compression or expansion operations could therefore be performed for efficient signal transmission or storage with very little loss of signal information.

The general mixing of input signal and local oscillator can also be viewed as a time-variant modulation of the input signal. Careful choice of the local oscillator time and frequency characteristics can lead to a controlled time-variant modulation useful in secure communications problems as a spreading spectrum technique.

The operation of the general optical heterodyne signal processing system has been described essentially as a frequency-variant mixing of the short-time spectral components of an input signal and a local oscillator. As indicated in Eqs. (4-36), (4-59) and (4-77), the nature of this mixing requires that the analytic description of system operation be in terms of complicated analytic expressions. The complexities of these analytic relationships are not conducive to easy interpretation of general system operation. We have shown that knowledge of general frequency mapping and resolution characteristics can be applied for an intuitive description of system operation, but the complete analytic descriptions are readily interpretable only for elementary specific cases. In the next chapter, we describe optical heterodyne signal processors in which the mixing of input signal and local oscillator components is performed in a channelized manner. The interpretation of the channelized system operation is greatly simplified by the assumption of frequency-invariance within the individual processing channels. However, the flexibility of the processor provided by the 2-D

nature of the optical systems allows us to maintain a substantial degree of generality in the spectral mapping operations even with this channelized arrangement.

CHAPTER V

CHANNELIZED OPTICAL HETERODYNE PROCESSING SYSTEMS

In this chapter, we are concerned with the practical implementation of a frequency-variant heterodyne signal processing system. We have analyzed the general processor and found the relationships between system input and output to be unwieldy in their complexity. The primary cause of this complexity is the continuous nature of the variations in mapping parameters over the input signal bandwidth. Here we make the simplifying assumption that the mapping parameters change over the input signal bandwidth in a finite number of discrete steps, i.e. the mapping operation is channelized. We describe below two versions of the channelized system: first, a system whose channelized structure arises directly from the general processor described in Chapter IV when periodic local oscillator signals are used; second, a more general channelized configuration resulting when segmented or channelized masks are employed in the optical mapping portion of the system. In both cases, we show that the analytical description of individual channel operation is greatly simplified, and we develop appropriate channel models. The overall operation of the channelized processors can then be described in terms of the individual channel models.

In either a channelized or a non-channelized system, the interferometric configuration of the processor requires that optical components, particularly mirrors and beamsplitters, be of very high quality, implying high cost. The system must also be isolated from any mechanical vibrations.

We describe in this chapter two alternative configurations that eliminate the need for mirrors and conventional beamsplitters. These configurations also greatly reduce the requirements for mechanical stability by processing all optical distributions with a common optical system.

In a channelized processor, the finite imaging resolution capabilities of practical optical mapping systems introduce channel crosstalk terms into the system output. The magnitude of these crosstalk terms increases as the number of channels increases, limiting the number of independent channels obtainable in the optical system. We analyze the crosstalk terms and describe their effect in terms of an overall crosstalk noise signal. It is then possible to determine the maximum number of channels obtainable in the optical processor for a given signal-to-crosstalk noise power ratio. We also describe a processing technique with which we can substantially increase the number of independent channels for a given signal-to-crosstalk ratio.

Finally, we demonstrate some of the basic operational capabilities of the optical processing system with a series of experiments demonstrating spectral component mapping.

Channelized Processor with Periodic Local Oscillator Signals

In Chapter IV, we derived the analytical expressions for the output signal using an arbitrary input signal and an arbitrary local oscillator signal. We now consider the system output for the specific local oscillator

$$r(t) = a_1 \cos 2\pi\nu_1 t. \tag{5-1}$$

For simplicity, in this chapter, we assume that the velocities associated with the input mappings of both the input signal and the local oscillator are the same and equal to unity: $\alpha_1=\alpha_2=1$. The local oscillator frequency mapping characteristic is then $\nu=m_2(y)$. The single sinusoid local oscillator signal produces a single spot of light in the local oscillator spectral distribution at the position y_1 for which $\nu_1=m_2(y_1)$. This distribution mixes with the input signal spectral distribution at the PMT. We assume that the extent of the local oscillator spot is small in the y-direction compared with variations in the input signal spectral distribution as mapped along the y-axis. The effect of this local oscillator spot is then to sample the input signal spectral distribution at $y=y_1$. From Eq. (4-58), this sampling gives the output signal

$$\hat{g}(t) = C_a^2\, a_1 F_w(m_1(y_1), y_1; t) e^{i2\pi[m_1(y_1)-\nu_1]t}. \tag{5-2}$$

Substituting Eq. (4-55) for the frequency-variant short-time spectrum we obtain

$$\hat{g}(t) = C_a\{\hat{f}(t) * w_1(-t, y_1) e^{i2\pi m_1(y_1)t}\} a_1 e^{-i2\pi\nu_1 t}. \tag{5-3}$$

A block diagram system model for the optical processor with a single sinusoid local oscillator is shown in Fig. 5-1. The input signal is filtered by a narrowband filter whose center frequency is $m_1(y_1)$. The filter output is then scaled in amplitude by a_1 and downshifted in frequency by ν_1.

In general, if we introduce a local oscillator consisting of N separate sinusoids of different frequencies,

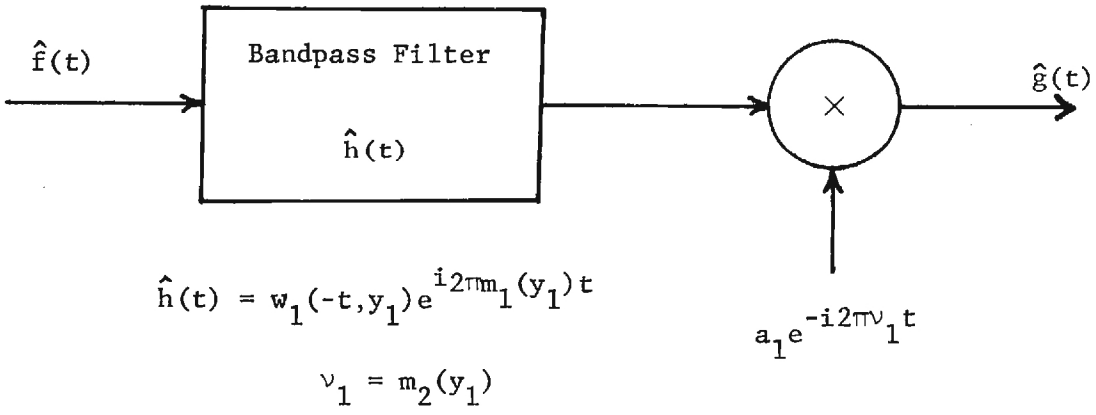


Figure 5-1. System Diagram for Heterodyne Processor with Single Sinusoid Local Oscillator.

$$r(t) = \sum_{n=1}^{N} a_n \cos 2\pi\nu_n t, \tag{5-4}$$

the processor output is

$$\hat{g}(t) = \sum_{n=1}^{N} \{\hat{f}(t) * w_1(-t,y_n) e^{i2\pi m_1(y_n)t}\} a_n e^{-i2\pi\nu_n t}, \tag{5-5}$$

where y_n is found from

$$\nu_n = m_2(y_n). \tag{5-6}$$

The operation of the optical processor with a local oscillator in the form of Eq. (5-4) can now be visualized in a relatively simple manner, as illustrated in Fig. 5-2. The input signal is first divided into N narrowband components;* the narrowband components are then rearranged in frequency, each component being shifted in frequency by a different amount. The center frequency of the n^{th} narrowband component is $\nu_{c_n} = m_1(y_n)$; the time and frequency resolution of the component are determined by the input window $w_1(-t,y_n)$. These quantities are, in turn, dependent on the local oscillator frequency ν_n through Eq. (5-6). This dependence establishes a set of relationships between component characteristics and the subsequent frequency shift applied to that component. In the general frequency-variant processor, these relationships are quite flexible: the rearrangement of components can represent a bandwidth compression or expansion operation, and the time-frequency resolution can vary from component to component.

---

*It is not necessary that the components be narrowband, this would, however, be a typical case.

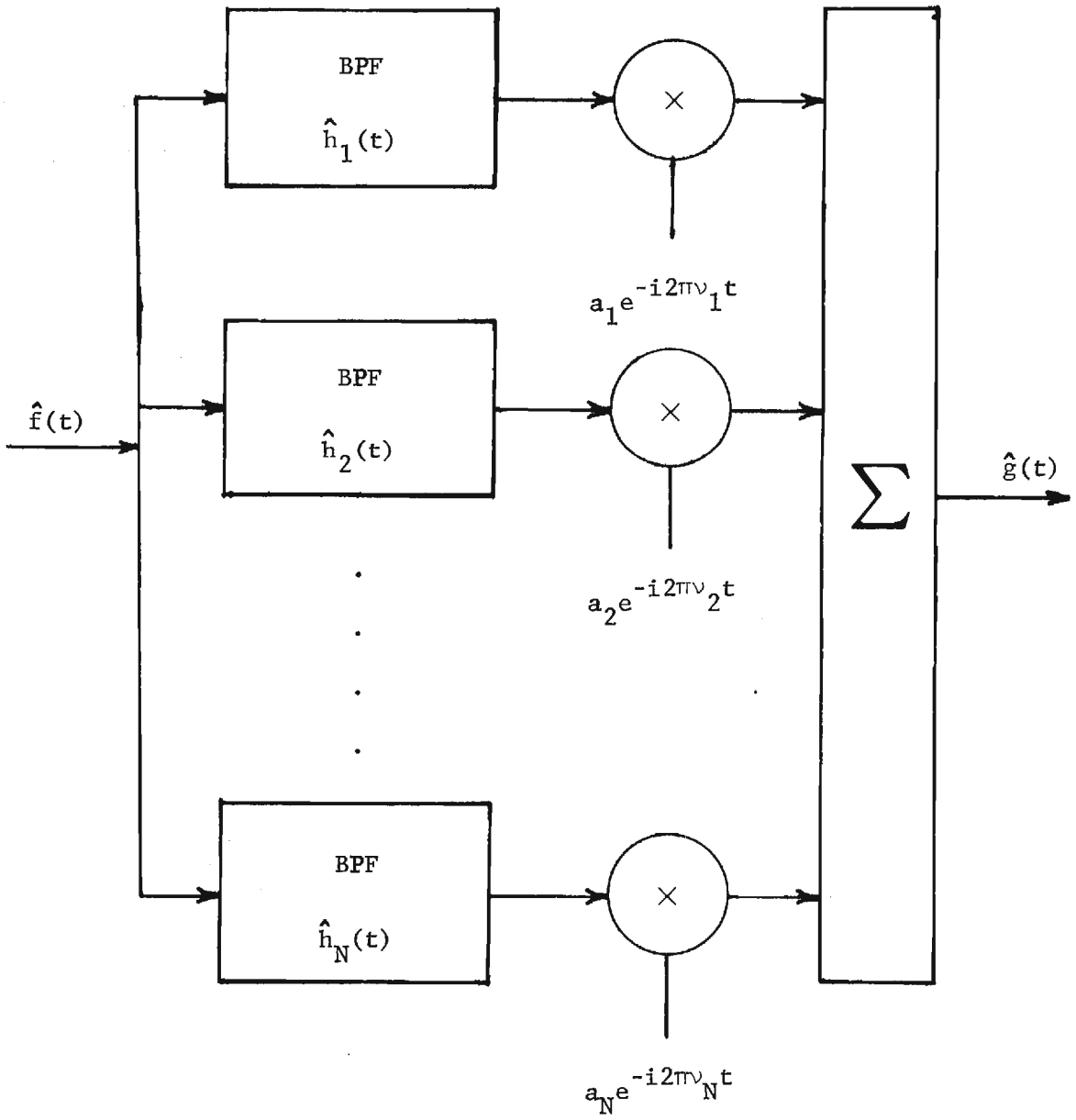


Figure 5-2. Channelized System Diagram for Heterodyne Processor with Multiple Sinusoid Local Oscillators.

In a practical implementation of this system, a local oscillator signal of the form given in Eq. (5-4) may cause problems due to dynamic range limitations of the input devices. If all or even a substantial portion of the frequency components are at integer multiples of some fundamental frequency, the local oscillator will exhibit a periodic, pulse-like structure with magnitude excursions that can be large compared with normal signal levels, especially if the number of components is large. Such variations in signal level can easily exceed the dynamic range of the input devices. The pulse-like structure of the local oscillator can be eliminated by introducing the local oscillator frequencies with different phases:

$$r(t) = \sum_{n=1}^{N} a_n \cos [2\pi\nu_n t + \phi_n], \qquad (5\text{-}7)$$

where the ϕ_n are random or otherwise chosen to minimize the maximum excursion of r(t). The introduction of the phase differences in the local oscillator frequencies introduces phase changes in the processed signal components. These changes must be incorporated as part of the overall mapping operation.

The General Channelized Processing System

A more general configuration for a channelized optical processing system is produced by channelizing the masks controlling the optical mapping operation within the processor. In this system, the vertical axis of the mapping system is divided into N regions centered at y_n with each region of nonoverlapping width Δy. Within each region the transmission characteristics of the optical mapping masks are invariant, i.e. they do not change in the y-direction. Thus, within the n^{th} region, the input optical distri-

bution is imaged by a shift-invariant narrowband imaging system as shown in Fig. 5-3. The local oscillator distribution in this region is imaged by a similar narrowband imaging system with the two images superposed at the detector. The two images combine coherently in the output plane, passing through a narrow slit at x=0 onto the photosensitive surface of the PMT.

From Eq. (4-58), the two narrowband images in the n^{th} channel at x=0 can be expressed as

$$[\hat{f}(t) * w_1(-t,y_n)e^{i2\pi m_1(y_n)t}]\,\mathrm{rect}\left(\frac{y-y_n}{\Delta y}\right) = \hat{f}_n(t)\,\mathrm{rect}\left(\frac{y-y_n}{\Delta y}\right), \qquad (5\text{-}8)$$

and

$$[\hat{r}(t) * w_2(-t,y_n)e^{i2\pi m_2(y_n)t}]\,\mathrm{rect}\left(\frac{y-y_n}{\Delta y}\right) = r_n(t)\,\mathrm{rect}\left(\frac{y-y_n}{\Delta y}\right), \qquad (5\text{-}9)$$

where rect(•) is the unit rectangle function. The signal detected by the PMT in the n^{th} channel is the

$$\hat{g}_n(t) = C_a^2 \int_{y_n-\Delta y/2}^{y_n+\Delta y/2} \hat{f}_n(t)\hat{r}^*(t)\,dy = C_a^2 \Delta y \hat{f}_n(t)\hat{r}_n^*(t). \qquad (5\text{-}10)$$

A diagram model for this channel is shown in Fig. 5-4. This channel model has the form of a Type I separable shift-variant system described by Kailath [39]. In such a system, a signal first passes through a linear shift-invariant filter and is then multiplied by a time-varying function. In this optical channel, the time-varying signal is produced by passing the local oscillator signal through a narrowband linear shift-invariant

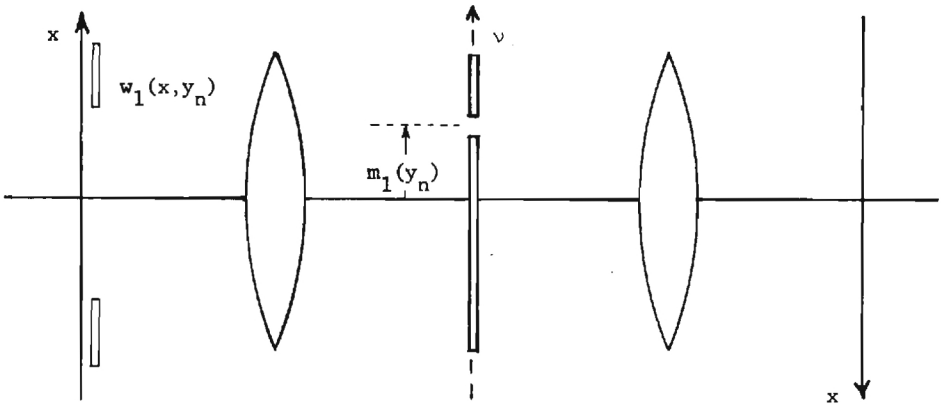


Figure 5-3. Narrowband Imaging System Configuration of the n^{th} Channel in a General Channelized Processor.

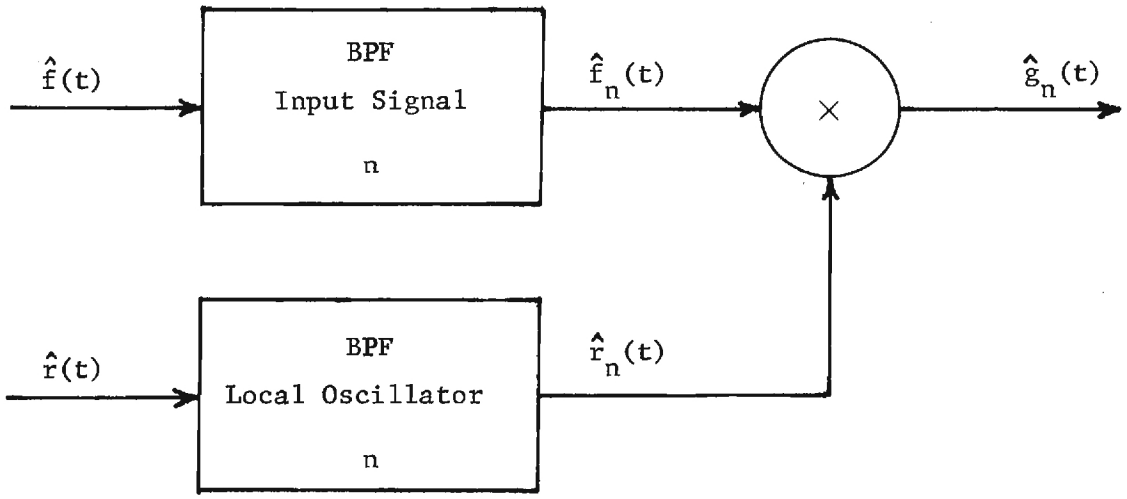


Figure 5-4. Channel Model for n^{th} Channel.

filter.

The total system output in the channelized processor is the sum of all the individual channel outputs

$$\hat{g}(t) = \Delta y C_a^2 \sum_{n=1}^{N} f_n(t)\hat{r}_n^*(t). \qquad (5\text{-}11)$$

The channelized system can then be modeled as shown in Fig. 5-5. The input signal and the local oscillator signal pass through separate banks of narrowband filters dividing each signal into N narrowband components. The components of the input signal are then mixed channel by channel with the local oscillator components to produce the total system output.

The overall operation of this channelized processor is a rearrangement of the input signal spectral components in a manner controllable by the choice of optical mapping within each channel and by the choice of local oscillator signal. The channelized processor of the previous section, which used periodic local oscillator signals, can be viewed as a special case of this general processor. However, channelization of the masks allows an extra degree of flexibility not possible in a system employing a continuously varying mapping slit. With certain restrictions we describe in the next section, the optical mapping operation performed in an individual channel can be viewed as completely independent of that performed in any other channel. The frequency mapping relations $m_1(\cdot)$ and $m_2(\cdot)$ do not have to describe continuous curves, and the input windows $w_s(\cdot,\cdot)$ and $w_r(\cdot,\cdot)$ can change drastically from one channel to the next. Thus, any narrowband component of the input signal can be mixed with any component of the local oscillator with any desired resolution characteristics by

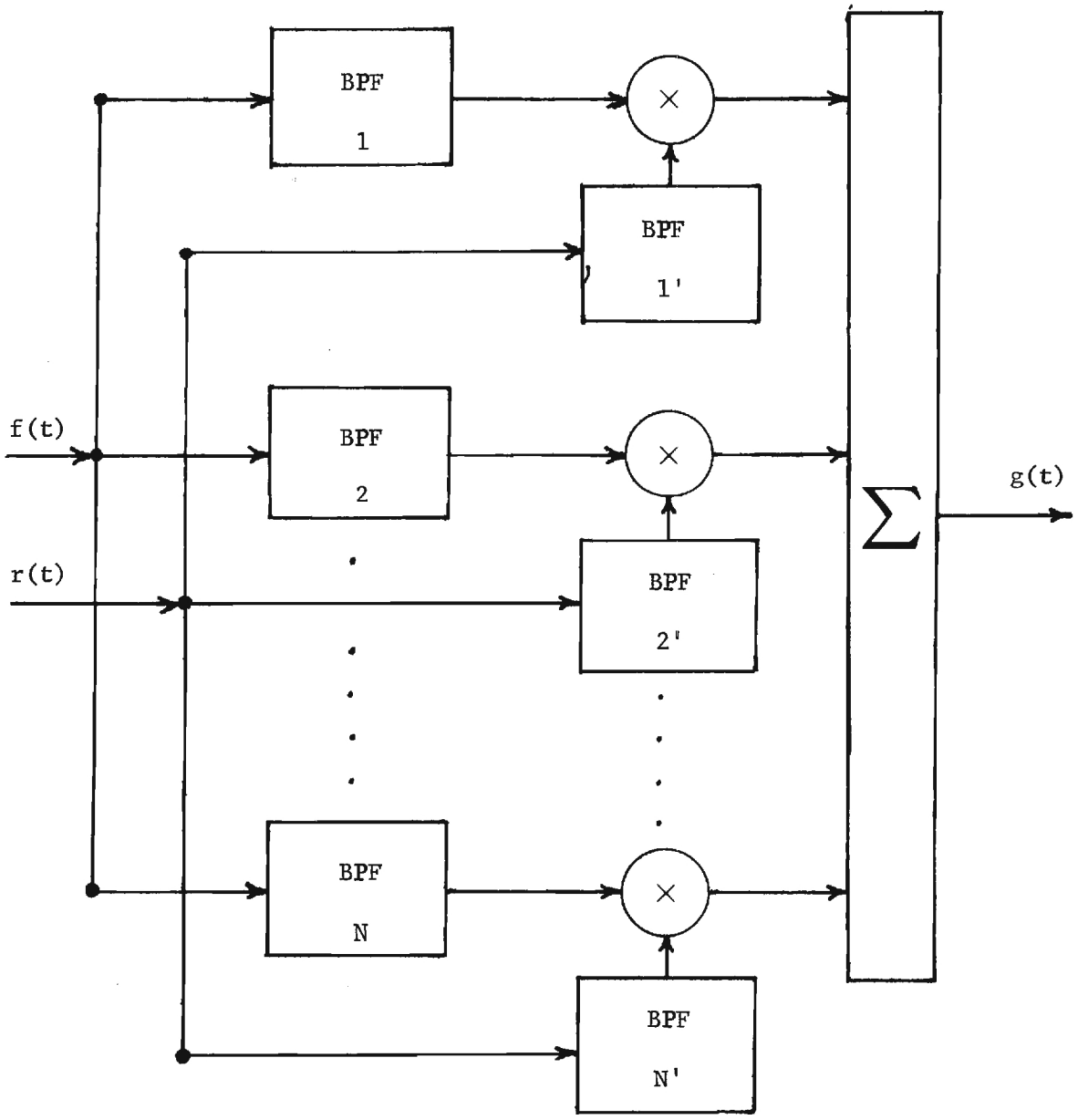


Figure 5-5. System Model for General Channelized System.

proper design of the channel mapping operations. Applications of such a system range from simple constant-compression-factor bandwidth compression to frequency domain scrambling for secure communications.

In addition to the mapping characteristics of the individual channels, the overall operation of the channelized processor is dependent on the total number of channels used in the processing system. More channels means a finer division of the total signal spectrum and the possibility of more general mapping characteristics. The total number of independent channels in the system depends on the imaging resolution capabilities of the optical mapping system and on requirements for channel independence. We discuss this aspect of the channelized processor in a later section; first, however, we consider alternative optical configurations for the heterodyne processing system.

Optical Heterodyne Processing with a Common Optical Mapping System

A major practical problem in the implementation of the optical heterodyne processor is the interferometric configuration used to combine the signal and local oscillator light distributions. Any slight irregularity in an optical component of the system (especially a mirror or a beamsplitter) will introduce a phase change in one or both optical distributions. The phase change can be observed in the output as a change in the output irradiance distribution at the PMT. Constant phase irregularities can be minimized by using optical components of very high (interferometric) quality, but the most serious problems in the heterodyne processor result from phase changes caused by vibration of individual optical components. The vibrations cause time-varying changes in the output irradiance distribution that are detected by the PMT as a noise-

like signal in the system output. Elimination of this vibration noise requires mechanical isolation of the system components from all vibration sources. This is usually done by mounting system components on an air suspended stable table. However, if moving film sound tracks are used to input the signal and local oscillator, it is virtually impossible to isolate the system from vibrations induced by the film drives or even by the moving film strips themselves.

We describe in this section two alternative configurations for the heterodyne processor that significantly reduce the effect of vibrations. In these configurations, mirrors and conventional beamsplitters are eliminated; the optical mappings of both the signal and local oscillator distributions and the subsequent combination of the two distributions are performed with a single optical system. All components of the optical system are common to both the signal and local oscillator distributions so that minor vibrations of individual components impart equal phase changes to both distributions. Equal phase changes, even though time varying, cause no change in the output irradiance distribution and therefore no noise signal in the system output.

The first common optical configuration we discuss is shown in Fig. 5-6. In this system, the two optical distributions resulting from the input mapping of the signal and the local oscillator are both positioned in the input plane of a single astigmatic processor. These distributions are separated in the vertical direction allowing the use of different input plane masks for each distribution. The first vertical imaging/ horizontal Fourier transforming operation of the astigmatic processor produces the transform distributions for the signal and local oscillator

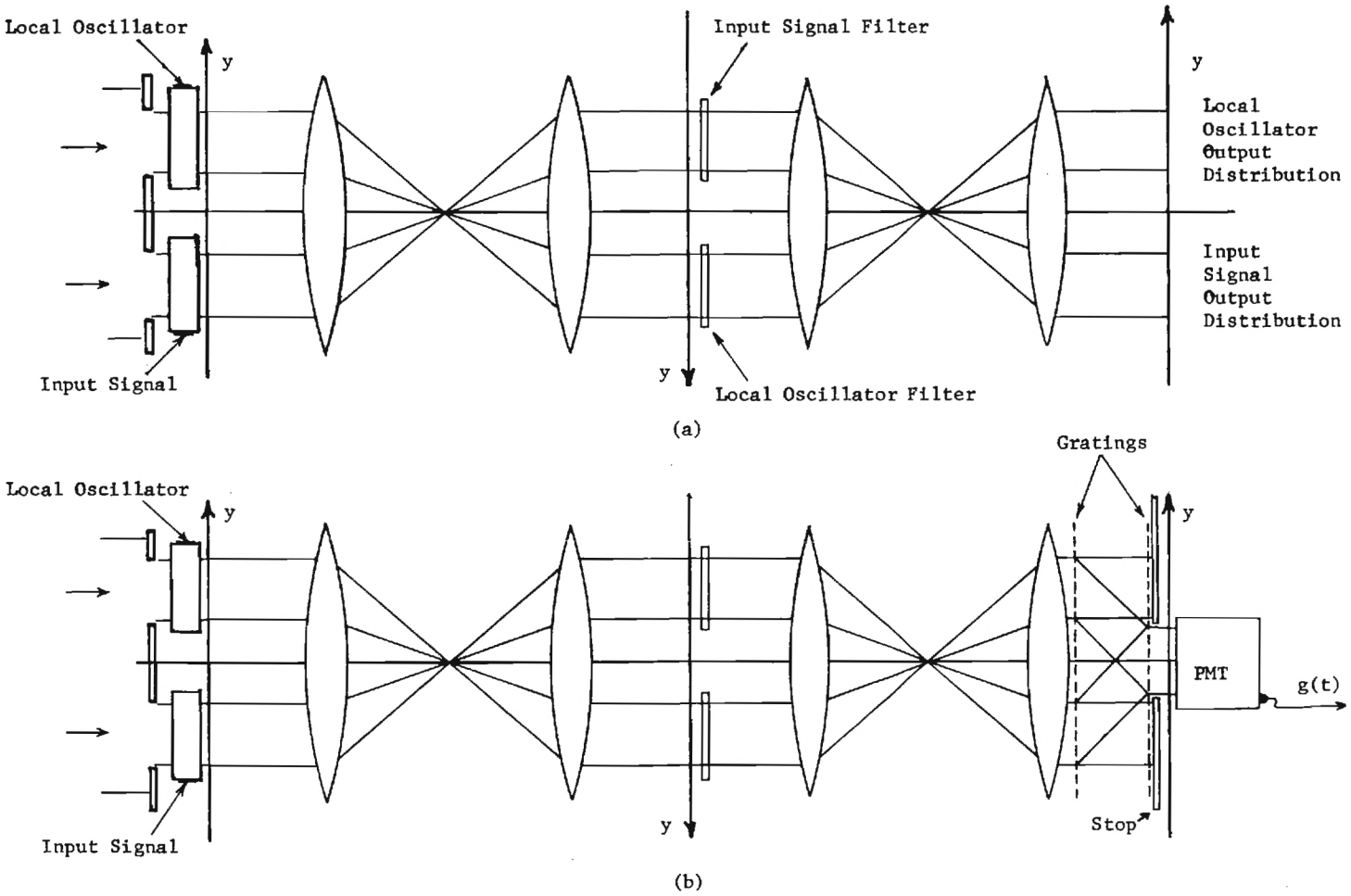


Figure 5-6. A Common Optical Component Heterodyne Processing System.

in the intermediate plane. The vertical imaging operation maintains the spatial separation of the two distributions. Separate frequency-domain masks can then be applied to each distribution to affect the desired mappings. The second vertical imaging/horizontal Fourier transforming operation produces the two output plane distributions still spatially separated in the vertical direction as indicated in Fig. 5-6(a). Portions of these distributions can be superimposed and aligned normally incident to the PMT by a pair of diffraction gratings as shown in Fig. 5-6(b).

In this system, the only components not common to both distributions are the input devices, which in the case of acoustooptic input devices introduce virtually no vibrations. If film sound tracks are used to input the signals, separate film strips and film drives can be eliminated by recording both the signal and the local oscillator on a single film strip as shown in Fig. 5-7. Certain limitations of this approach are, however, readily apparent. The gratings used to combine the signal and local oscillator distributions can diffract only a small portion of the total light distribution incident upon them. Since the output signal can only be detected from the light distributions that are diffracted by both gratings, the light levels of the output distribution will be low. These low light levels result in low level output signals that may be lost in the inherent low level noise signals inherent to the photodetector (e.g. dark current noise). Another problem with this configuration is a limitation in the total number of channels that can be processed by the system. In order to combine the two distributions properly with the gratings, they must originally be spatially separated by a distance at least equal to their vertical extent. This separation allows the diffracted portions to overlap

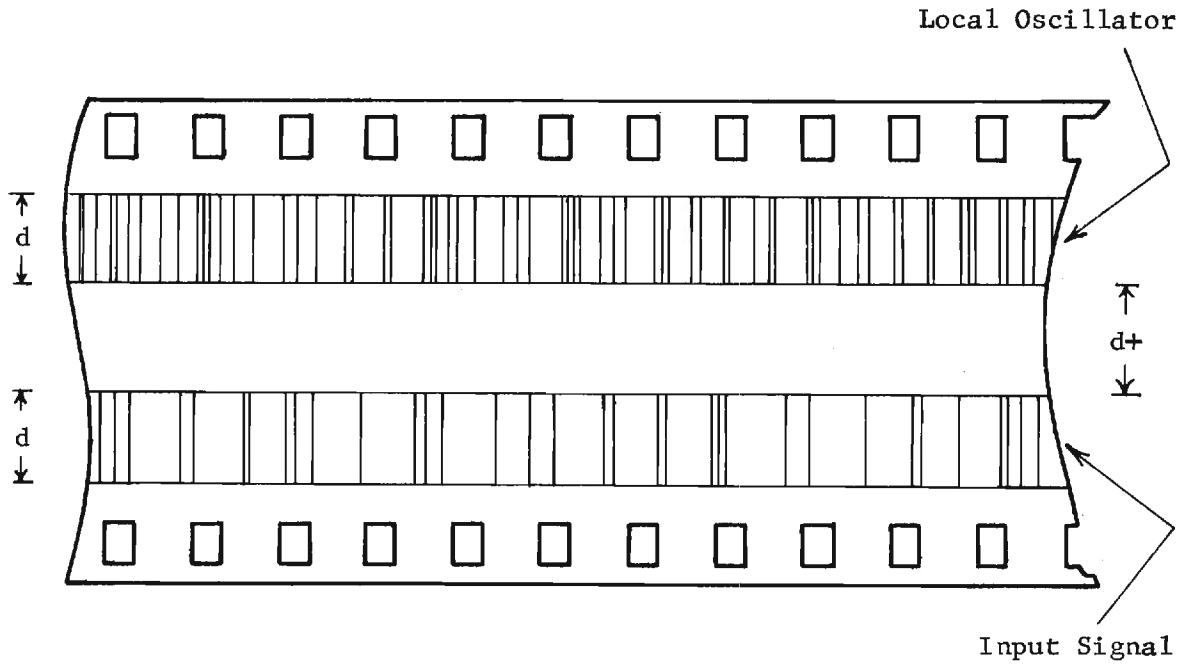


Figure 5-7. Dual Soundtrack for Input in Common Optical Processor.

without interference from the undiffracted portions of the two distributions. At most only one third of the total vertical extent of the optical mapping system can then be used to form individual channels. The total number of channels possible in this configuration is therefore reduced from that of the interferometric configuration by a factor of at least three.

The second common optical configuration we consider is shown in Fig. 5-8. In this system the signal and local oscillator are combined in a frequency multiplexed fashion as shown in Fig. 5-9(a). The signal components occupy the low frequency portion of the combined spectrum while local oscillator components are restricted to a separate high frequency region. The combined signal undergoes the input mapping with a single input device in the input plane of an astigmatic processor. Only one input plane window can be used, the same for both signal and local oscillator. However, the single-mask frequency-variant processor can easily incorporate the separate mapping operations in a single frequency-domain mask. An example of a mask for separate, continuous slit-mapping operations is shown in Fig. 5-9(b). The second vertical imaging/horizontal Fourier transforming operation performed in the optical system then combines the two output distributions in the proper manner at the PMT. No grating or beamsplitters are required to combine the two distributions.

With the frequency multiplexed configuration, the total vertical extent of the optical mapping system can be used in forming channels. The maximum number of channels possible in this system is the same as in the interferometric configuration. Output distribution light levels are higher in the frequency-multiplexed configuration than in either of the other configurations because no light is lost at beamsplitters or gratings. We do,

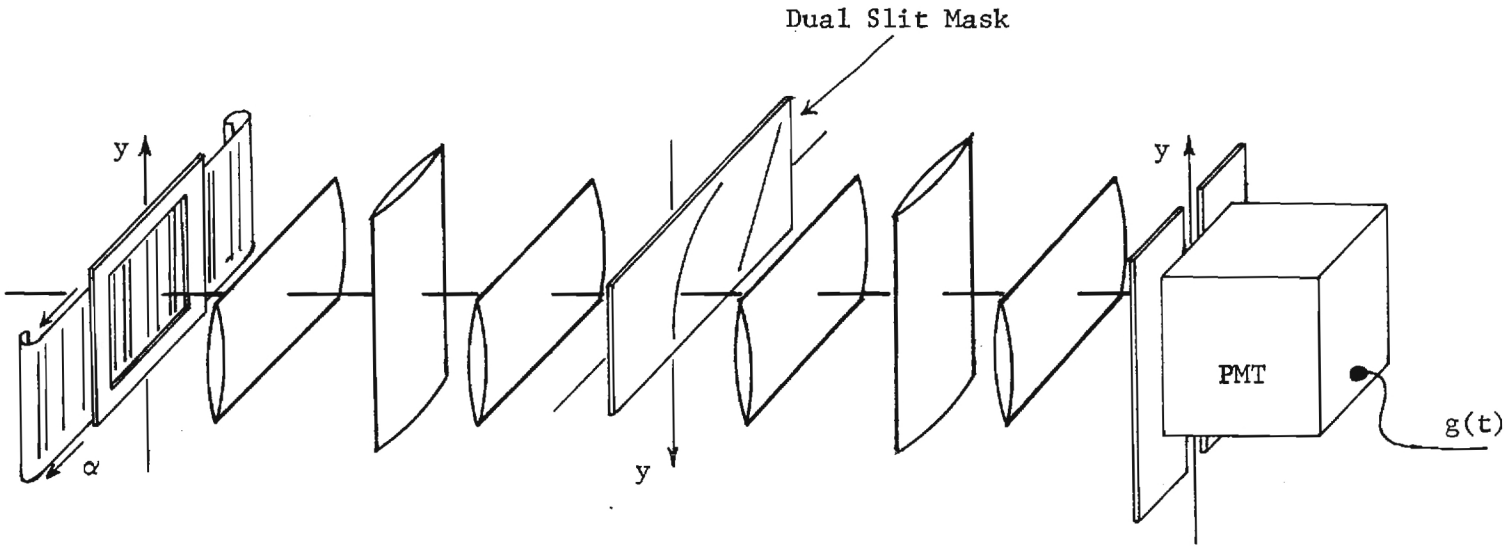


Figure 5-8. Frequency Multiplexed Common Optical Processor.

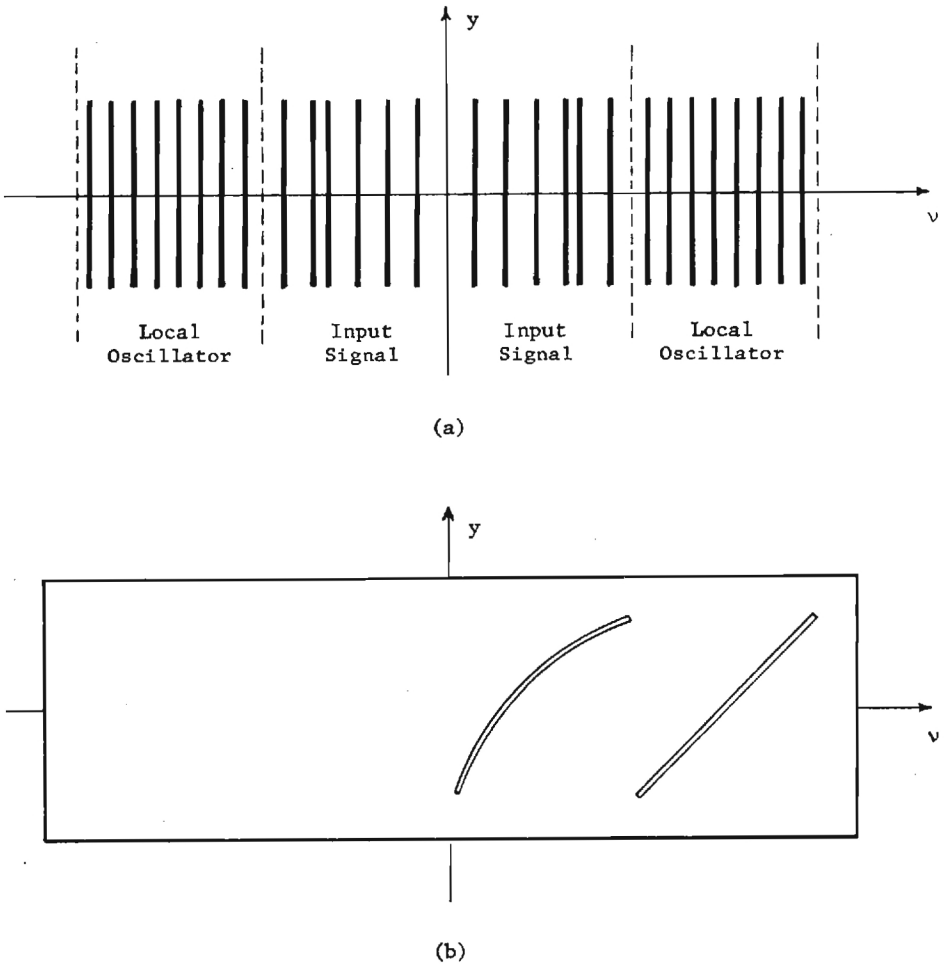


Figure 5-9. The Frequency-Domain Distribution of the Input Signal and Local Oscillator in the Frequency Multiplexed Processor; (a) Spectral Distribution for the Combined Signal, (b) Transform Plane Mask for the Frequency Multiplexed Processor.

however, sacrifice available bandwidth for the signal and local oscillator. Most input devices (especially film) have a finite bandwidth over which they will operate. If we multiplex the signal and local oscillator in the frequency domain, we must divide the available bandwidth between the two signals. Additionally, input signal and local oscillator levels must be restricted because of dynamic range limitations in the single input device.

Channel Crosstalk and Channel Independence

The system models presented in previous sections described channelized systems in which the operation of an individual channel is completely independent from that in any other channel. Inherent in this model is the assumption that the vertical imaging operation of the astigmatic processor is ideal, imaging point into point. The ideal light distributions (filtered input signal and filtered local oscillator) in the n^{th} channel were given in Eqs. (5-8) and (5-9). However, a practical imaging system has finite resolution - the image produced by such a system is the ideal image smoothed by some pointspread function p(y). The smoothed images then mix in the heterodyne operation to produce the desired output term for the n^{th} channel:

$$C_a^2\{\int_{-\infty}^{\infty}[\mathrm{rect}\left(\frac{y-y_n}{\Delta y}\right) *_y p(y)]^2 dy\}\ \hat{f}_n(t)\hat{r}_n^*(t) = C_a^2 A_0 \hat{f}_n(t) r_n^*(t). \qquad (5\text{-}14)$$

The signal and local oscillator light distributions in the other system channels are also smoothed by the imaging operation causing an overlap of light distributions from neighboring channels with light in the n^{th} channel. The overlap results in additional mixing in the heterodyne operation producing extraneous signal terms in the system output. We

consider the additional signal terms produced in the n^{th} channel from the overlap of light from the m^{th} channel. One additional term resulting from this overlap is produced by the mixing of the signal distribution in the n^{th} channel with the local oscillator distribution from the m^{th} channel. This mixing produces the output signal term

$$C_a^2\{\int_{-\infty}^{\infty}[\mathrm{rect}\left(\frac{y-y_n}{\Delta y}\right)* p(y)][\mathrm{rect}\left(\frac{y-y_m}{\Delta y}\right)*p(y)]^* dy\} \cdot \hat{f}_n(t)\hat{r}_n^*(t)$$

$$= C_a^2 A_{nm}\hat{f}_n(t)\hat{r}_m^*(t). \qquad (5\text{-}15)$$

A similar mixing occurs between the local oscillator component in the n^{th} channel and the signal distribution of the m^{th} channel, but we associate this term with the output of the m^{th} channel. Two additional terms are produced in the heterodyne operation due to this overlap, $C_a^2 A_{nm}\hat{f}_n(t)\hat{f}_m^*(t)$ and $C_a^2 A_{nm}\hat{r}_n(t)\hat{r}_m^*(t)$. The first of these terms results from the cross-channel mixing of input signal components, whereas the second results from cross-channel mixing of local oscillator components. We can assume that a portion of these last signal terms is present in output of both the n^{th} and the m^{th} channel. We therefore assign one half of these terms to each channel. The overlap of light from the m^{th} channel thus produces an additional signal in the output of the n^{th} channel given by

$$\rho_{nm}(t) = C_a^2 A_{nm}[\hat{f}_n\hat{r}_m^*(t) + \frac{1}{2}\hat{f}_n(t)\hat{f}_m^*(t) + \frac{1}{2}\hat{r}_n(t)\hat{r}_m(t)]. \qquad (5\text{-}16)$$

We refer to this extraneous signal as the crosstalk in the n^{th} channel from the m^{th} channel and can treat it as signal dependent noise in system output.

The total crosstalk signal in the n^{th} channel is the sum of the crosstalk from all other channels

$$\rho_n(t) = \sum_{\substack{m=1 \\ m\neq n}}^{N} \rho_{nm}(t), \tag{5-17}$$

and the total output associated with the n^{th} channel is the sum of the desired output and the crosstalk noise:

$$\hat{g}_n(t) = C_a^2 A_0 \hat{f}_n(t)\hat{r}_n^*(t) + \rho_n(t). \tag{5-18}$$

Each term of the total channel output is formed by the mixing of two narrowband signal components. The first term of the crosstalk signal Eq. (5-16) is similar in nature to the desired channel mapping, i.e. the term results from the mixing of an input signal component with a local oscillator component. The second and third term of the crosstalk signal have the form of self-product signals; however, since there are no restrictions on the change in filter passband from channel to channel, these terms are potentially higher in frequency than normal self-product terms. The possible frequency range of any term in the system output is the frequency difference possible between the two components that mix to form the term. With the aid of Fig. 5-10 we can determine the potential frequency range for each term of the channel output.

We assume that the frequency multiplexed processing configuration is used in the processing, with input signal and local oscillator components distributed in frequency as shown in Fig. 5-10(a). Initially, we assume that the desired mapping in all channels is performed by mixing a positive frequency component of the input signal (between 0 and ν_s) with

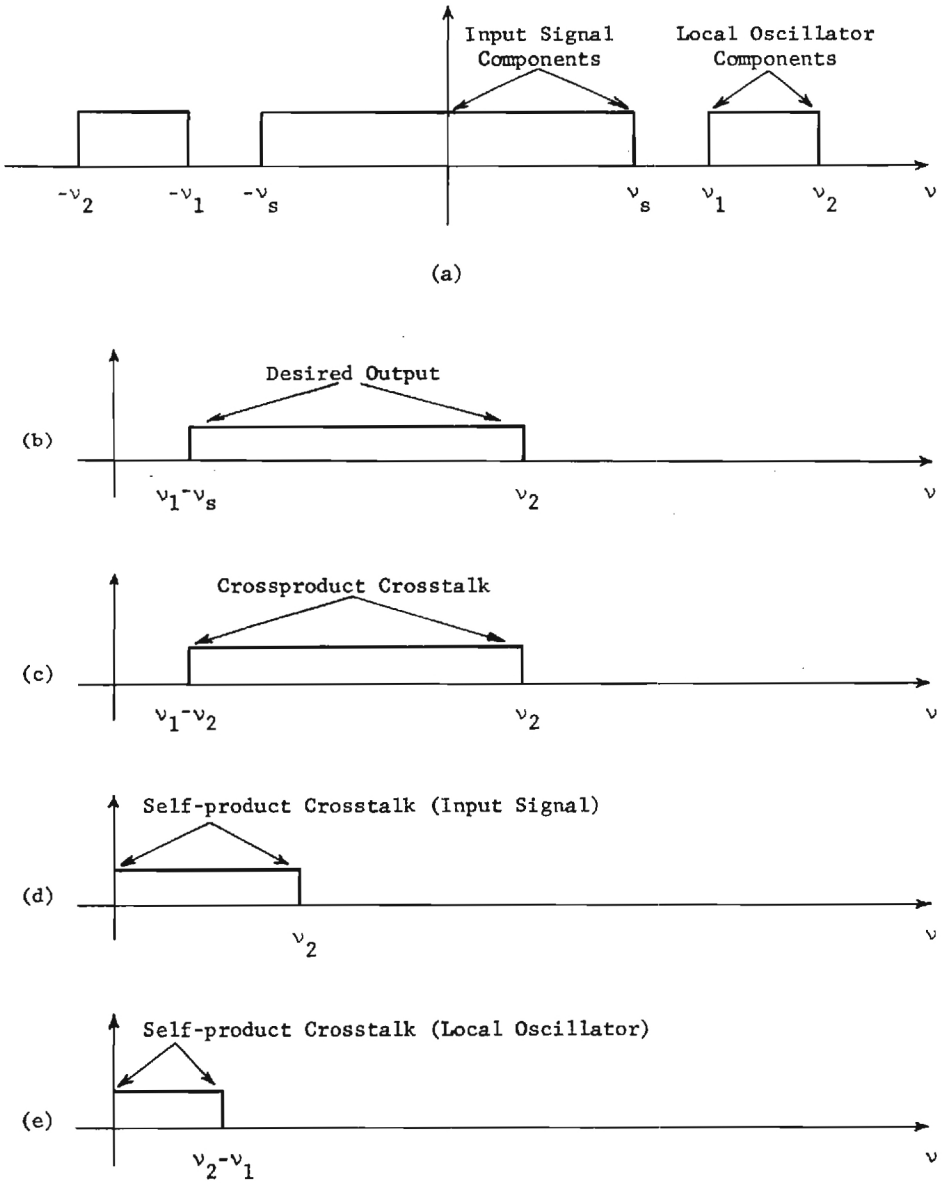


Figure 5-10. Potential Frequency Ranges of the System Output Terms when Positive Input Signal Components Mix with Positive Local Oscillator Components.

a positive frequency component of the local oscillator (between ν_1 and ν_2). The minimum frequency difference in the mixing of components in these ranges is $\nu_1-\nu_s$ while the maximum frequency difference is ν_2. The desired channel output signal therefore lies in the frequency range from $\nu_1-\nu_s$ to ν_2 as shown in Fig. 5-10(b). The first term of the crosstalk signal $\rho_{nm}(t)$ is also the result of mixing a signal component (0 to ν_s) and a local oscillator component (ν_1 to ν_2). The potential range of this crosstalk term, Fig. 5-10(c), is the same as the desired channel output. The first self product term of $\rho_{nm}(t)$ mixes two components of the input signal with possible frequency difference from 0 to ν_s, as shown in Fig. 5-10(d). Similarly the self product term mixing two local oscillator components has the possible frequency range from 0 to $\nu_2-\nu_1$, Fig. 5-10(e).

From Fig. 5-10, we see that any self-product terms below $\nu_1-\nu_s$ can be removed by highpass filtering the system output signal without affecting the desired channel output. Certainly, if $\nu_1 > 2\nu_s$, then all of the self-product crosstalk terms can be removed from the system output by highpass filtering. However, this requirement would put further restrictions on input signal and local oscillator bandwidths already limited in a frequency multiplexed processor due to the finite operational bandwidth of the input device.

An alternative to the use of such high frequency local oscillator components is to perform the channel processing by mixing positive frequency input signal components (between 0 and ν_s) with negative frequency local oscillator components (between $-\nu_2$ and $-\nu_1$) as indicated in Fig. 5-11(a). The desired channel mixing of an input signal component and a local oscillator component then produces an output signal in the range

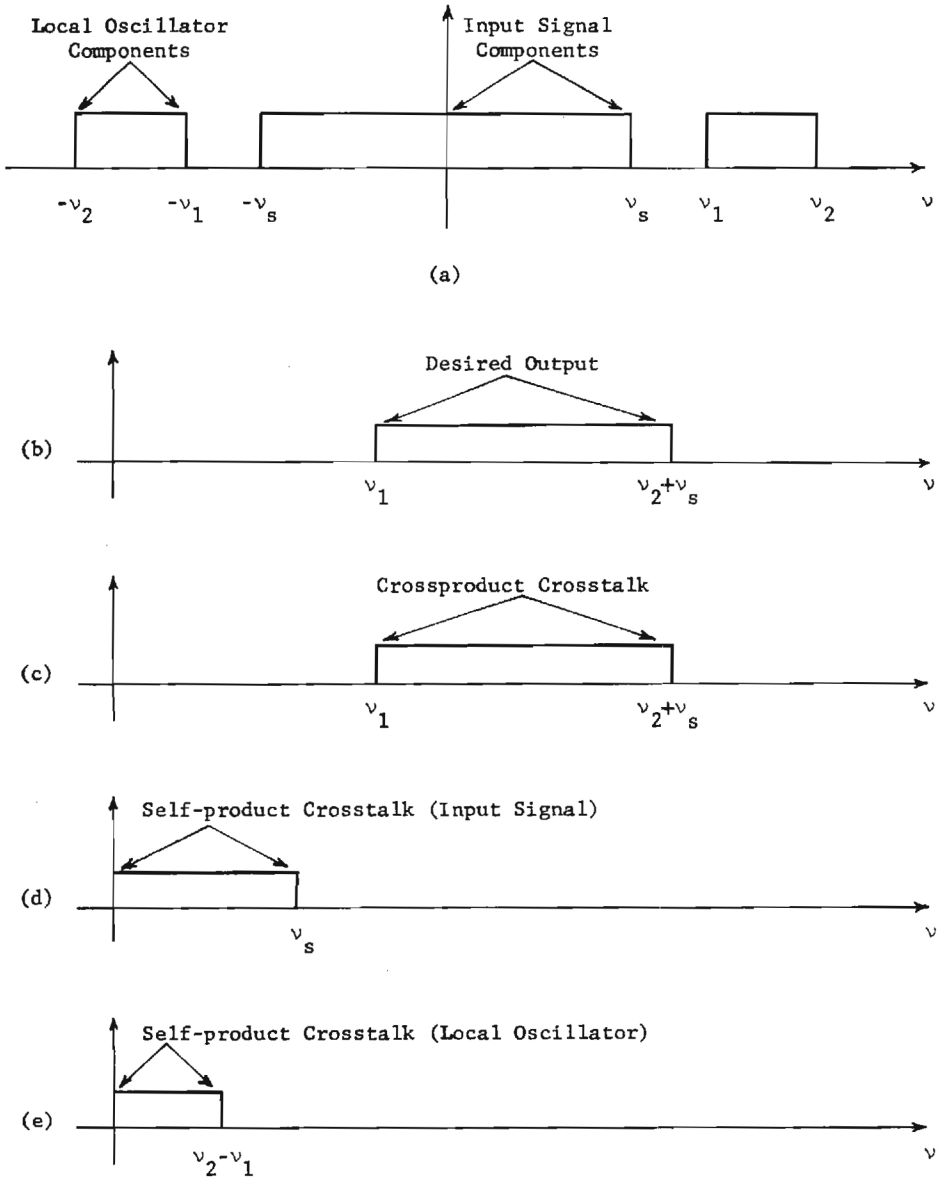


Figure 5-11. Potential Frequency Ranges of the System Output Terms when Positive Input Signal Components Mix with Negative Local Oscillator Components

indicated in Fig. 5-11(b), between ν_1 and $\nu_2+\nu_s$. The range of the cross-product crosstalk term, Fig. 5-11(c), is the same as the desired channel signal, while the self product terms appear in the ranges shown in Figs. 5-11(d) and (e). Thus, in this processing scheme, the self-product crosstalk terms can be completely removed from the system output by highpass filtering, subject only to the requirement that $\nu_2 < 2\nu_1$. (This requirement insures that the mixing of two local oscillator components will not fall in the range of the desired channel output.)

If the self-product crosstalk terms are removed by post-detection filtering, then the total channel output reduces to

$$\hat{g}_n(t) = C_a^2 A_0 \hat{f}_n(t)\hat{r}_n^*(t) + \sum_{\substack{m=1 \\ m\neq n}}^{N} C_a^2 A_{nm} \hat{f}_n(t)\hat{r}_m^*(t). \qquad (5\text{-}19)$$

Using a processing technique we refer to as staggered channel processing, we can further reduce the number of crosstalk terms in the channel output by post detection filtering. We note first that the desired output term in any channel can be formed either by mixing a positive-frequency input signal component with a negative-frequency local oscillator component or by mixing the corresponding negative-frequency input signal component with the corresponding positive-frequency local oscillator component. If the input signal and local oscillator are both real-valued signals, then either mixing will produce the same term in the real output signal. In the staggered channel processing technique the mixing used in the various channels alternates as shown in Fig. 5-12. In the n^{th} channel, mixing occurs between a positive-frequency signal component and a negative frequency local oscillator component. In the closest neighboring channels, n±1, the oppo-

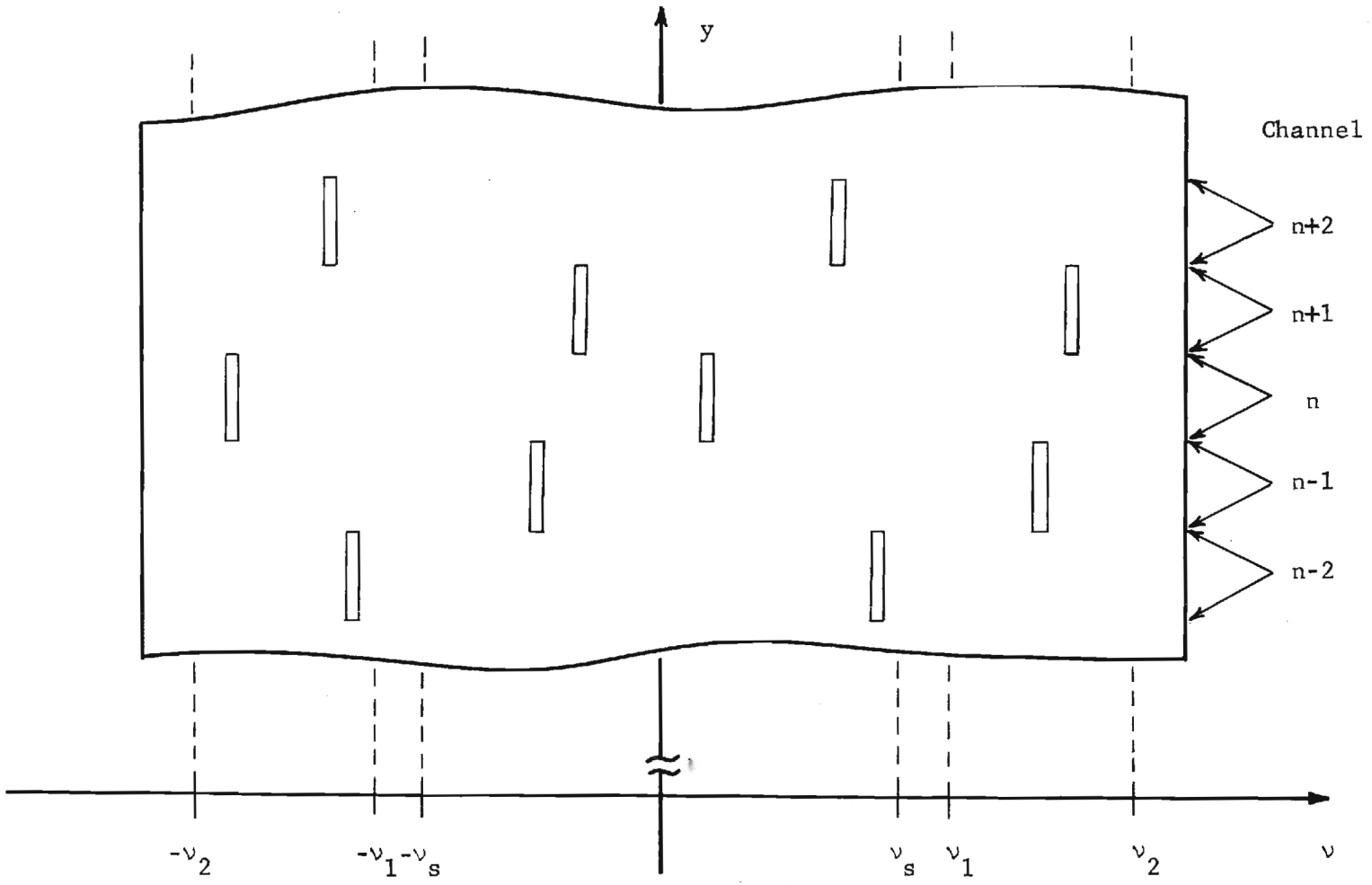


Figure 5-12. Mask Used in Staggered Channel Processing.

site mixing method is used; in the second closest channels, n 2, the same mixing as in channel n is used, and so on. This staggering of channel mixing has a significant effect on the potential frequency ranges of the crosstalk terms as shown in Fig. 5-13. The input signal component and the local oscillator component of the n^{th} channel mix to produce the desired output term in the range ν_1 to $\nu_2+\nu_s$. The local oscillator components from the two closest channels n±1 also mix with the input signal component in the n^{th} channel, but they are both positive-frequency components. The cross-product crosstalk terms produced by this mixing are therefore in the range $\nu_1-\nu_s$ to ν_2. The same is true for any odd-ordered cross-product crosstalk term, i.e. the cross-product term from any channel n±1, n±3, ... The even-ordered cross-product crosstalk terms (from channels n±2, n±4, ...) are produced by the same type of mixing as the desired output - positive-frequency input signal component with negative-frequency local oscillator component. These terms therefore have the same potential frequency range as the desired channel output, ν_1 to $\nu_2+\nu_s$. The self-product terms are also affected by the staggering of channels, as indicated in the figure. The odd-ordered self-product terms can have significant high frequency content, whereas even-ordered self-product terms remain in the relatively isolated low-frequency regions.

In Fig. 5-13, we can observe some of the advantages of the staggered channel processing technique. First, if we employ a local oscillator signal such that $\nu_1 > 2\nu_s$ and $\nu_2+\nu_s < 2\nu_1$, the range of the desired output will be separate from all self-product crosstalk terms. The self-product terms can then be removed by bandpass filtering the processor output. Next, we note that most of the potential range of the odd-ordered cross product

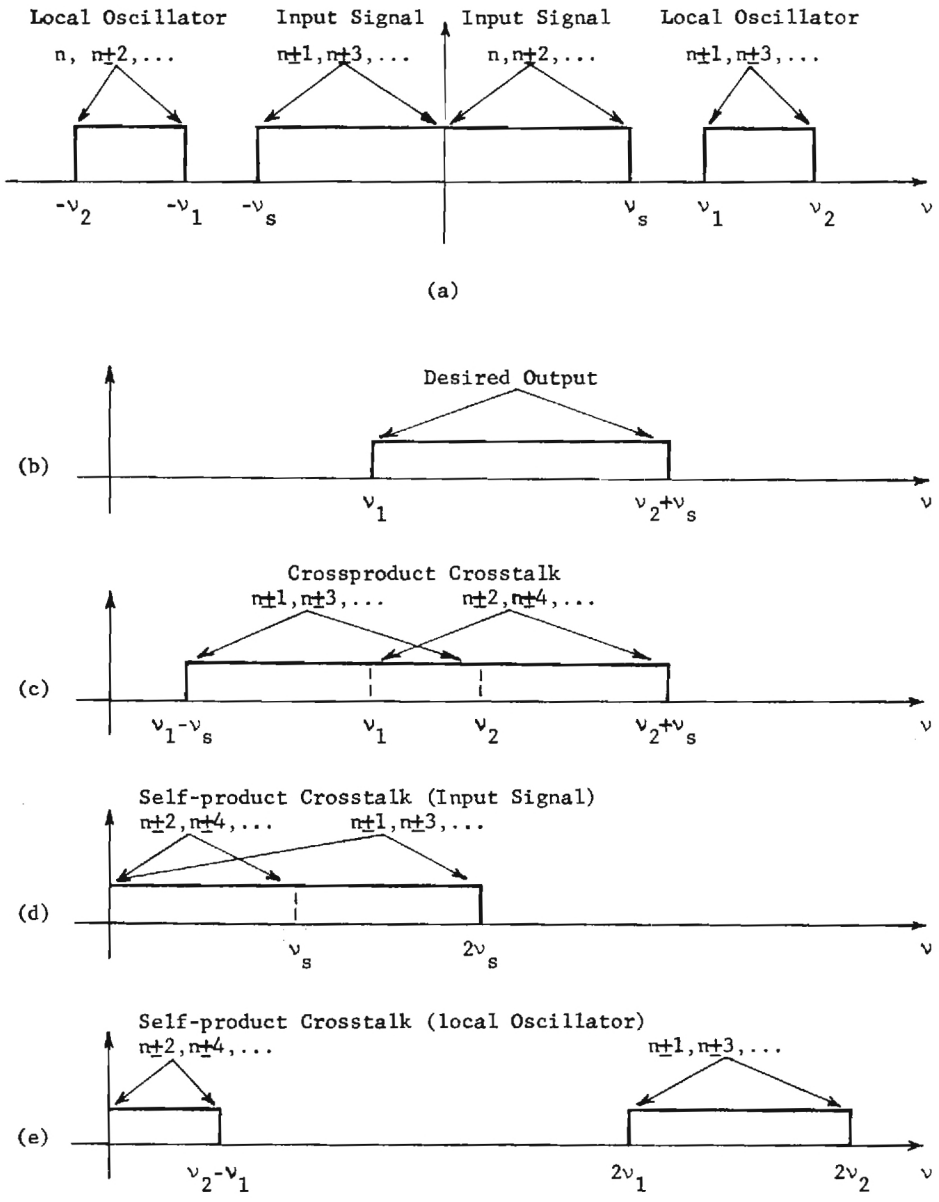


Figure 5-13. Potential Frequency Ranges of the System Output Terms using Staggered Channel Mixing.

terms are outside the range of the desired signal components. A proper choice of optical mapping parameter variations from channel to channel can insure that all significant odd-ordered cross product terms appear in that portion of the potential range. As we show below, the amplitude of the crosstalk terms falls off rather rapidly with increasing distance between channels. Thus, this choice does not impose any severe restriction on the processing capabilities of the system. Using the staggered channel approach we are therefore able to remove all self-product crosstalk and all significant odd-ordered cross product crosstalk terms in the channel output by bandpass filtering the system output. The channel output can then be written as

$$\hat{g}_n(t) = C_a^2 A_0 \hat{f}_n(t)\hat{r}_n^*(t) + \sum_{\substack{m=1 \\ m\neq n,n\pm1,n\pm3,\ldots}}^{N} C_a^2 A_{nm} \hat{f}_n(t)\hat{r}_m^*(t). \qquad (5\text{-}20)$$

As we mentioned above, the crosstalk terms can be considered to be signal dependent noise in the channel output. The presence of these terms in a channel output can also be viewed as an indication of the dependence of the operation in that channel on surrounding channels. The ratio of the power in the desired channel output signal to the power of the channel crosstalk signals can therefore be viewed as a signal-to-noise ratio measure for the channel and as a measure of channel independence. Larger signal-to-crosstalk noise ratios mean greater independence of the channel operation from surrounding channels; the operation of the practical system more closely resembles that of the ideal system shown in Fig. 5-5.

The power of the desired signal in the channel output is given by

$$P_s = \langle \left| C_a^2 A_0 \hat{f}_n(t)\hat{r}_n^*(t) \right|^2 \rangle = C_a^4 A_0^2 \langle \left| \hat{f}_n(t)\hat{r}_n^*(t) \right|^2 \rangle, \qquad (5\text{-}21)$$

where $\langle \cdot \rangle$ indicates time average. The power of the crosstalk terms in a non-staggered channel processing system would be

$$P_c = \left\langle \left| \sum_{\substack{m=1 \\ m\neq n}}^{N} C_a^2 A_{nm} \hat{f}_n(t) \hat{r}_m(t) \right|^2 \right\rangle$$

$$= \sum_{\substack{m=1 \\ m\neq n}}^{N} C_a^4 A_{nm}^2 \left\langle \left| \hat{f}_n(t) \hat{r}^*_{m(t)} \right|^2 \right\rangle + \sum_{\substack{m=1 \\ m\neq n}}^{N} \sum_{\substack{k=1 \\ k\neq n \\ m\neq k}}^{N} C_a^4 A_{nm} A_{nk} \left\langle \left| \hat{f}_n(t) \right|^2 \hat{r}^*_m(t) \hat{r}_k(t) \right\rangle . \tag{5-22}$$

The terms $r_m(t)$ represent different narrowband components of the local oscillator signal. Assuming sufficient frequency separation between the different components, the time averages of the terms in the second summation will be zero, i.e.

$$\left\langle \left| \hat{f}_n(t) \right|^2 \hat{r}^*_m(t) \hat{r}_k(t) \right\rangle = 0 \quad \text{for all } m\neq k.$$

The crosstalk signal power in the n^{th} channel is then simply

$$P_c = \sum_{\substack{m=1 \\ m\neq n}}^{N} C_a^4 A_{nm}^2 \left\langle \left| \hat{f}_n(t) \hat{r}^*_m(t) \right|^2 \right\rangle , \tag{5-23}$$

and the signal-to-crosstalk noise ratio is

$$S/N = \frac{P_s}{P_c} = \frac{A_0^2 \left\langle \left| \hat{f}_n(t) \hat{r}^*_m(t) \right|^2 \right\rangle}{\sum_{\substack{m=1 \\ m\neq n}}^{N} A_{nm}^2 \left\langle \left| \hat{f}_n(t) \hat{r}^*_m(t) \right|^2 \right\rangle} . \tag{5-24}$$

This signal-to-noise ratio is obviously signal dependent, particularly on the various oscillator levels. However, we can obtain a reasonable measure

of channel performance assuming all local oscillator levels to be equal. Then all time averages in Eq. (5-24) will be equal and

$$S/N = \frac{A_0^2}{\sum_{\substack{m=1 \\ m\neq n}}^{N} A_{nm}^2}, \tag{5-25}$$

for the non-staggered channel processor. Using the staggered channel processing technique we can eliminate all odd-ordered crosstalk terms in the signal, resulting in an improved signal-to-crosstalk noise ratio

$$S/N = \frac{A_0^2}{\sum_{\substack{m=1 \\ m\neq n,n\pm1,n\pm3,\ldots}}^{N} A_{nm}^2}. \tag{5-26}$$

To determine the order of magnitude of the signal-to-crosstalk noise ratio we may encounter in a practical system, we consider two representative special cases. In the first case, we assume the ideal channels are tightly packed, one next to the other. Channel separation is then equal to channel width Δy. These channels are imaged with a non-ideal system whose point spread function is $p(y) = \text{sinc}\,(y/d)$, where d is the minimum resolution width of the imaging system. The signal amplitude factor A_0 is then

$$A_0 = \int_{-\infty}^{\infty} \left[\text{rect}\left(\frac{y-n\Delta y}{\Delta y}\right) * \text{sinc}\left(\frac{y}{d}\right)\right]^2 dy, \tag{5-27}$$

and the crosstalk amplitude factors are

$$A_{nm} = \int_{-\infty}^{\infty} \left[\text{rect}\left(\frac{y-n\Delta y}{\Delta y}\right) * \text{sinc}\left(\frac{y}{d}\right)\right]\left[\text{rect}\left(\frac{y-m\Delta y}{\Delta y}\right) * \text{sinc}\left(\frac{y}{d}\right)\right] dy. \tag{5-28}$$

The integral expressions of Eqs. (5-27) and (5-28) have been evaluated numerically for different values of Δy ranging from d to 10d. The results of this evaluation are shown in Fig. 5-14. The variations of A_0 with Δy are shown in Fig. 5-14(a), while variations of the amplitude factors $A_{n(n\pm1)}$ through $A_{n(n\pm3)}$ are shown in Fig. 5-14(b) and (c). (By a substitution of variables in Eq. (5-28), it can be shown that $A_{n(n-k)}=A_{n(n+k)}$ for any k.) We note immediately from Figs. 5-14(b) and (c) that the crosstalk amplitudes drop off rapidly for higher order terms. Also, all crosstalk amplitudes fall off rapidly compared to A_0 as the channel width increases. The sidelobe structure of the point spread function introduces the cyclic variations in the crosstalk amplitudes, with most high order amplitude factors taking both positive and negative values as Δy varies.

The signal-to-crosstalk noise ratios for both the staggered channel and non-staggered channel processing technique were evaluated numerically including the crosstalk contribution from the 40 nearest neighboring channels. The variations of these signal-to-noise ratios with channel width are shown in Fig. 5-15. As expected, in the non-staggered case the S/N increases slowly and uniformly with increasing channel width. The S/N for the staggered channel processor is significantly higher than the non-staggered case for almost all channel widths. The major cause of this significant increase is the elimination by channel staggering of the crosstalk signals from the two nearest neighboring channels. The staggered channel S/N oscillates through a large range of values with variations in channel width maintaining a steadily increasing average. The channel widths can be chosen at the peaks of this curve enabling channels of small width to operate with a good degree of independence from other channels.

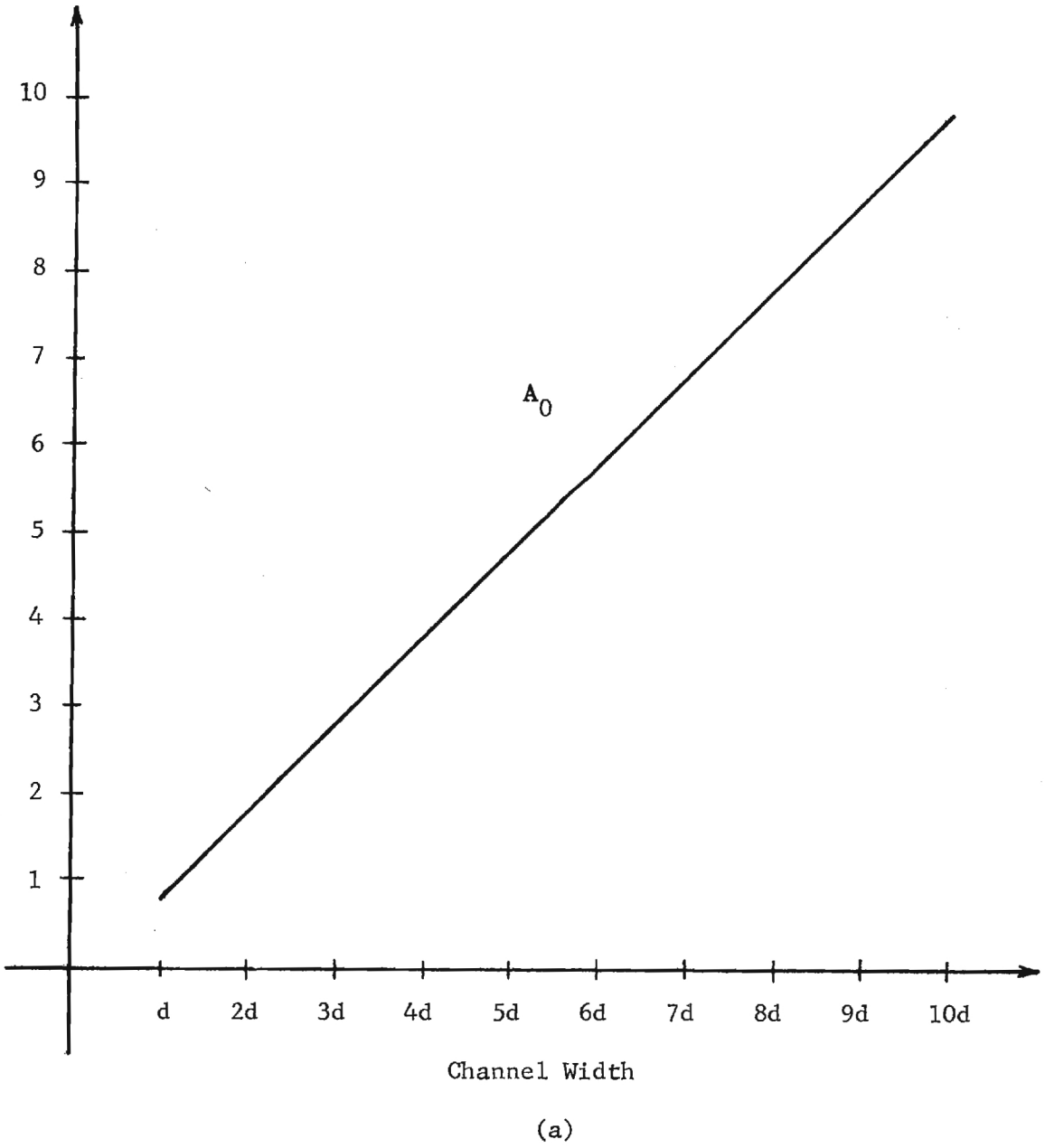
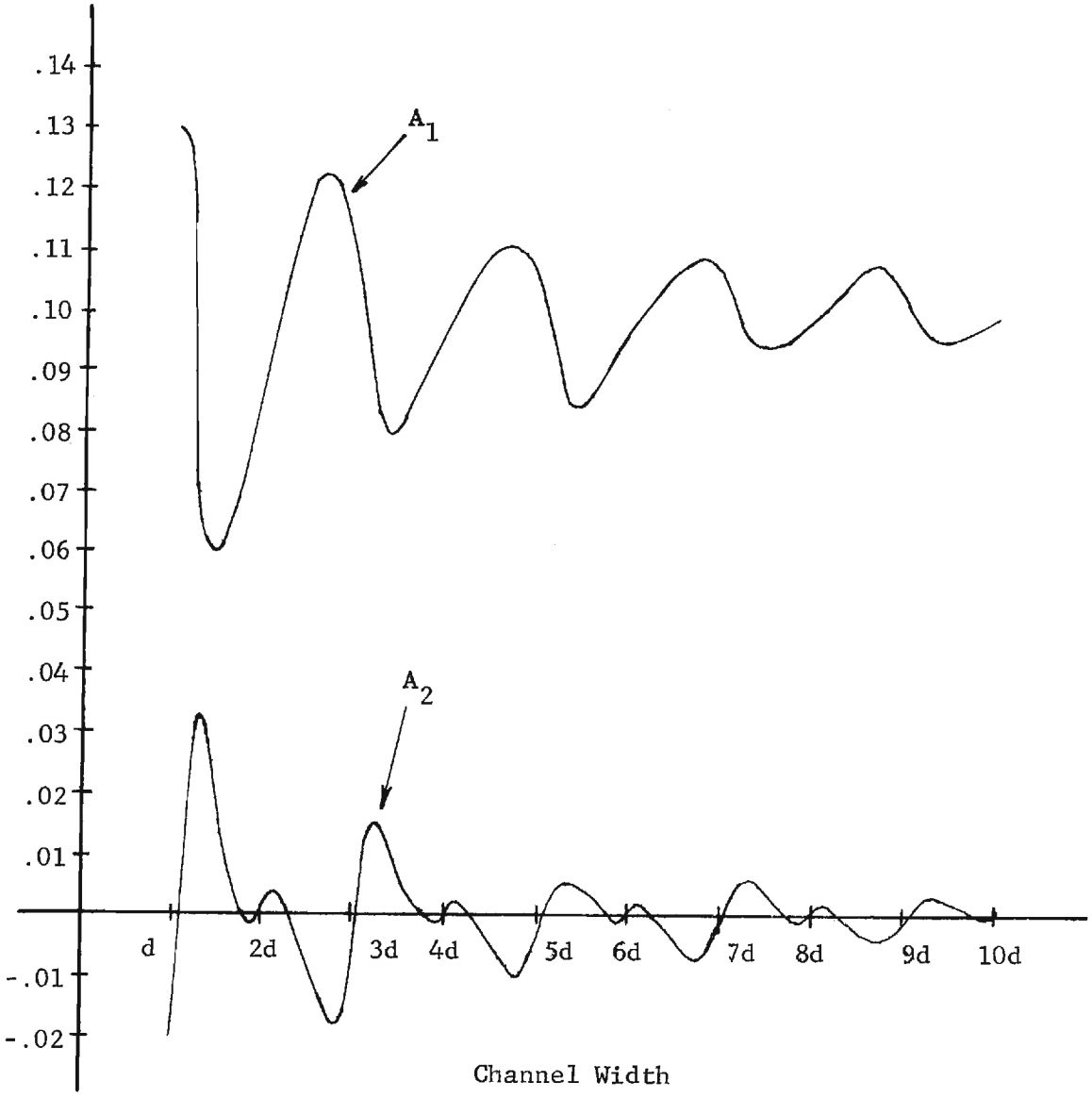


Figure 5-14. Variations in Signal and Crosstalk Amplitude Factors with Channel Width: (a) Signal Amplitude Factor.



(b)

Figure 5-14(cont.) Variations in Signal and Crosstalk Amplitude Factors with Channel Width; (b) Crosstalk Amplitude Factors A_1 and A_2.

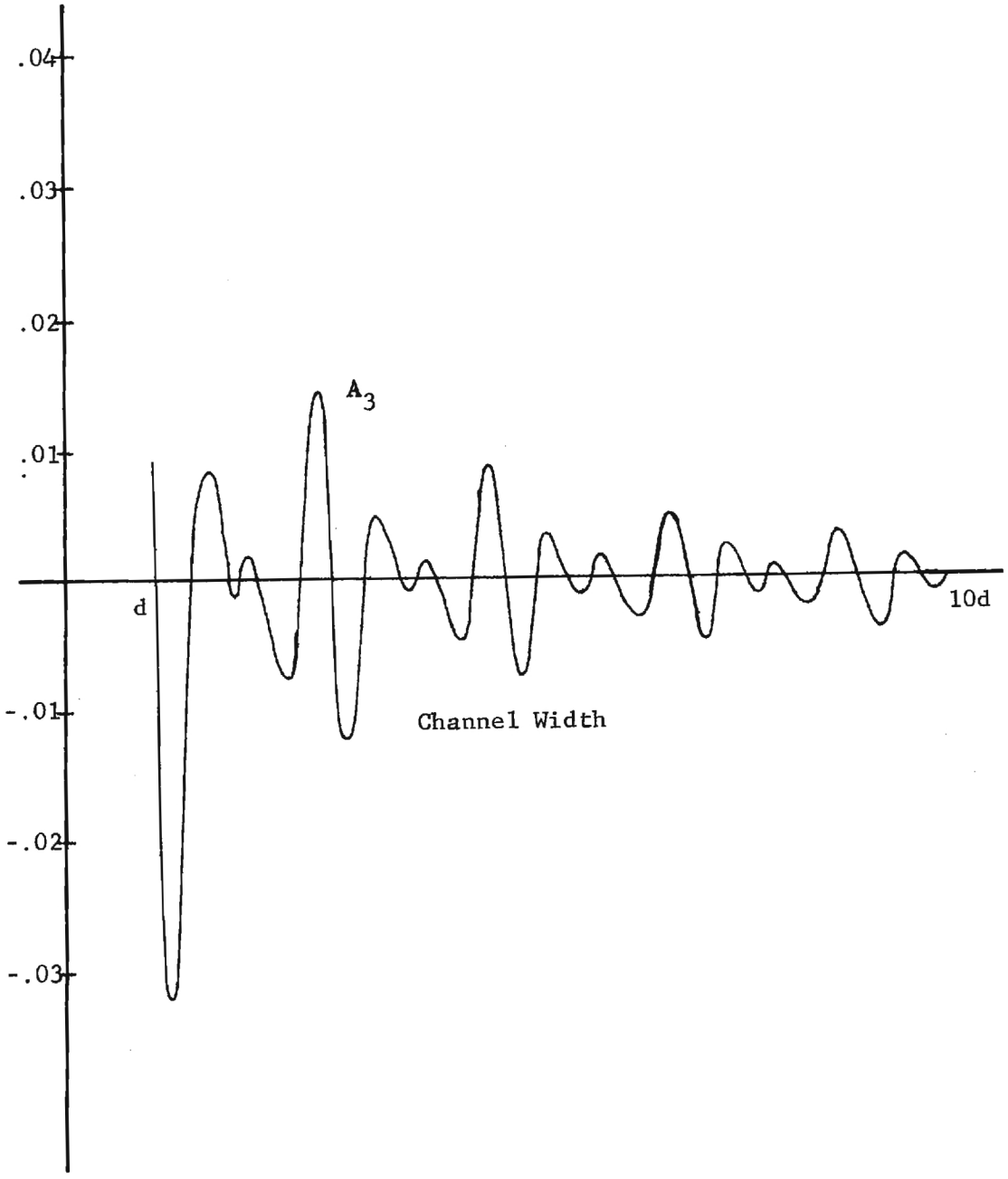


Figure 5-14(cont.) Variations in Signal and Crosstalk Amplitude Factors with Channel Width; (c) Crosstalk Amplitude Factor A_3.

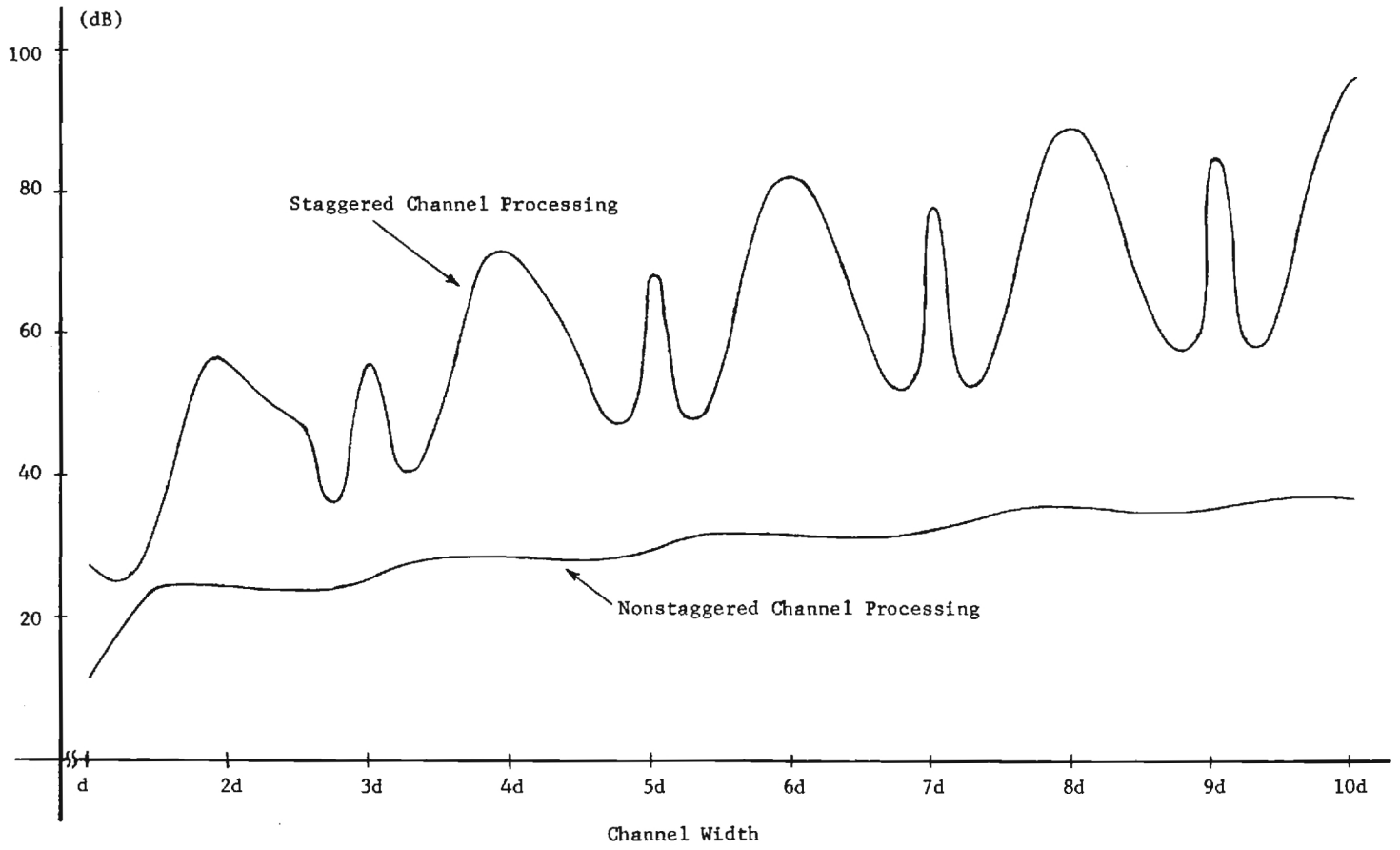


Figure 5-15. Variations in Signal to Crosstalk Power Ratio with Channel Width.

The total number of channels available in the system is dependent on the channel width. The maximum number of optical resolvable channels, i.e. those separated by d, is equal to the space-bandwidth product (SW) of the imaging system. From the curves of Fig. 5-15, we can determine the channel width, Δy, required to give a desired degree of channel independence (a desired S/N). The maximum number of channels available in the system with that degree of independence is then

$$N_{max} = SW(d/\Delta y). \tag{5-29}$$

It is clear from Fig. 5-15 that using the staggered channel processing technique makes available a significantly greater number of channels with given S/N (especially if we operate at the peaks of the S/N curve) than using the non-staggered approach. For example, if the SW for the imaging system is 1000, then by choosing $\Delta y=2d$ we can operate 500 separate channels with the staggered channel technique, all with a S/N of 55 dB. Although not shown on the curve, an S/N of 55 dB with the non-staggered processing technique would require $\Delta y \simeq 20$, indicating that only 50 channels could be operated with that degree of independence.

As a second example, we consider the variations in signal-to-noise ratio as the channel spacing is increased while holding the channel width constant. We assume rectangular channel cross-sections with channel width equal to the minimum resolution width, d, and the spacing, s, varying from d to 10 d. The signal amplitude factor remains constant, depending only on the channel width. The variations of the crosstalk amplitude factors $A_{n(n\pm1)}$ through $A_{n(n\pm3)}$ with increasing channel spacing are shown in Fig.

5-16. The variations in signal-to-crosstalk noise ratio with channel spacing in both the staggered and the non-staggered processors are shown in Fig. 5-17. Again, the staggered channel approach provides a larger signal-to-noise ratio, but the curve also indicates a significant peak in S/N in the non-staggered processor when channel spacing is an integer multiple (2 or above) of the channel width. These peaks can again be attributed to the sidelobe structure of the sinc-function pointspread. The maximum number of channels in this case depends on the channel spacing with

$$N_{max} = SW(d/s). \tag{5-30}$$

The curves of Fig. 5-17 indicate that, although the channel S/N in the staggered channel processor is larger, the non-staggered processor can be operated with a large number of channels, all with reasonably high S/N, by choosing the channel spacing at integer multiples of channel width.

As we have mentioned above, the oscillations and peaks that are observed in the channel S/N as channel width and spacing change are a result of the sidelobe structure in the pointspread function of the imaging system. We would not observe these oscillations if the pointspread function had little or no sidelobe structure as with a Gaussian function. For a Gaussian pointspread function, the variations of the channel S/N would follow steadily increasing curves approximately along the average of the curves for the sinc-function pointspread. Apodized apertures are often used in imaging systems to modify image appearance by the reduction of sidelobe structure in the imaging pointspread function. However, in the heterodyne processing system, non-apodized imaging is clearly preferable

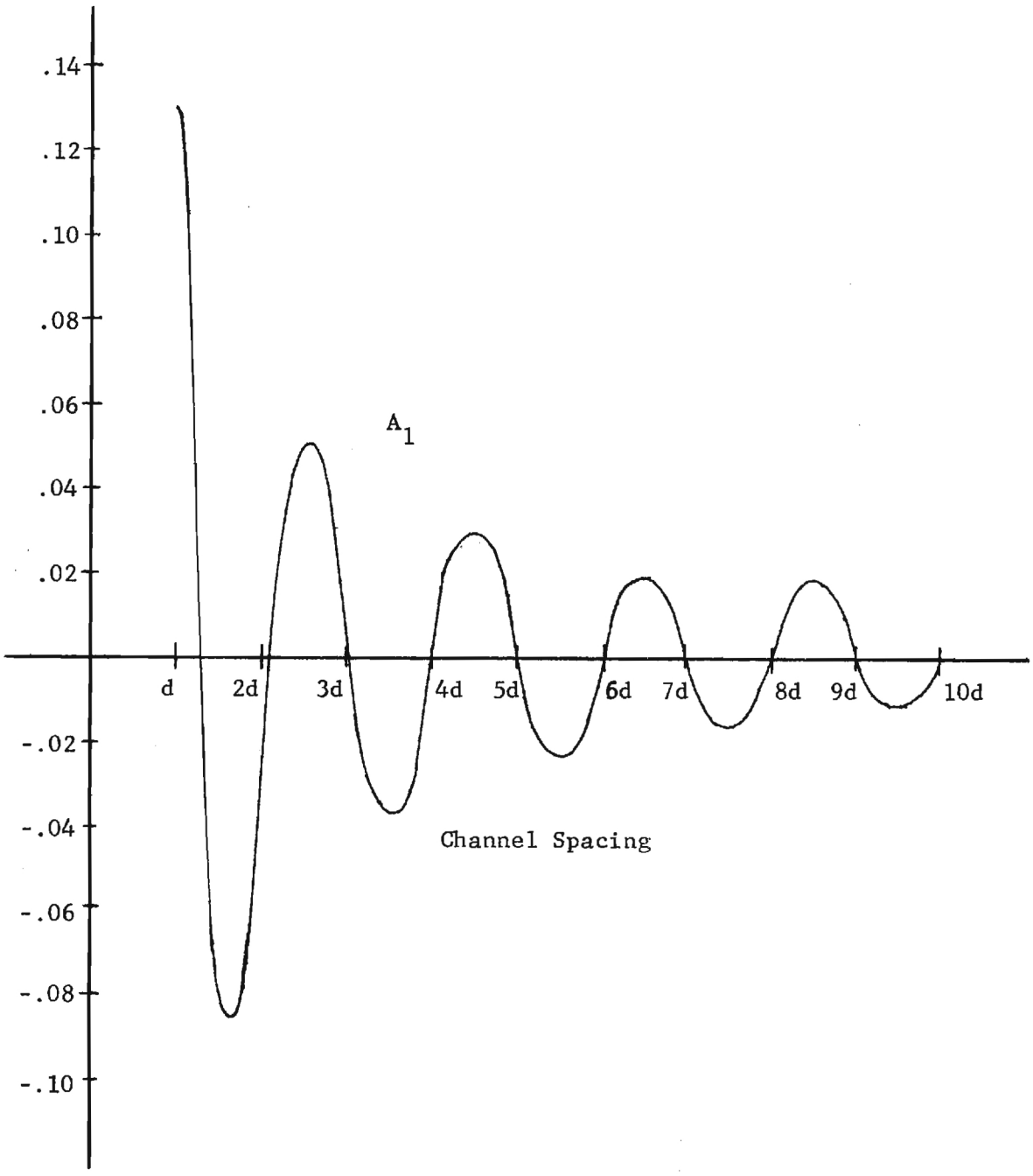


Figure 5-16. Variations in Crosstalk Amplitude Factors with Increasing Channel Spacing, Channel Width = d ; (a) Crosstalk Amplitude Factor A_1.

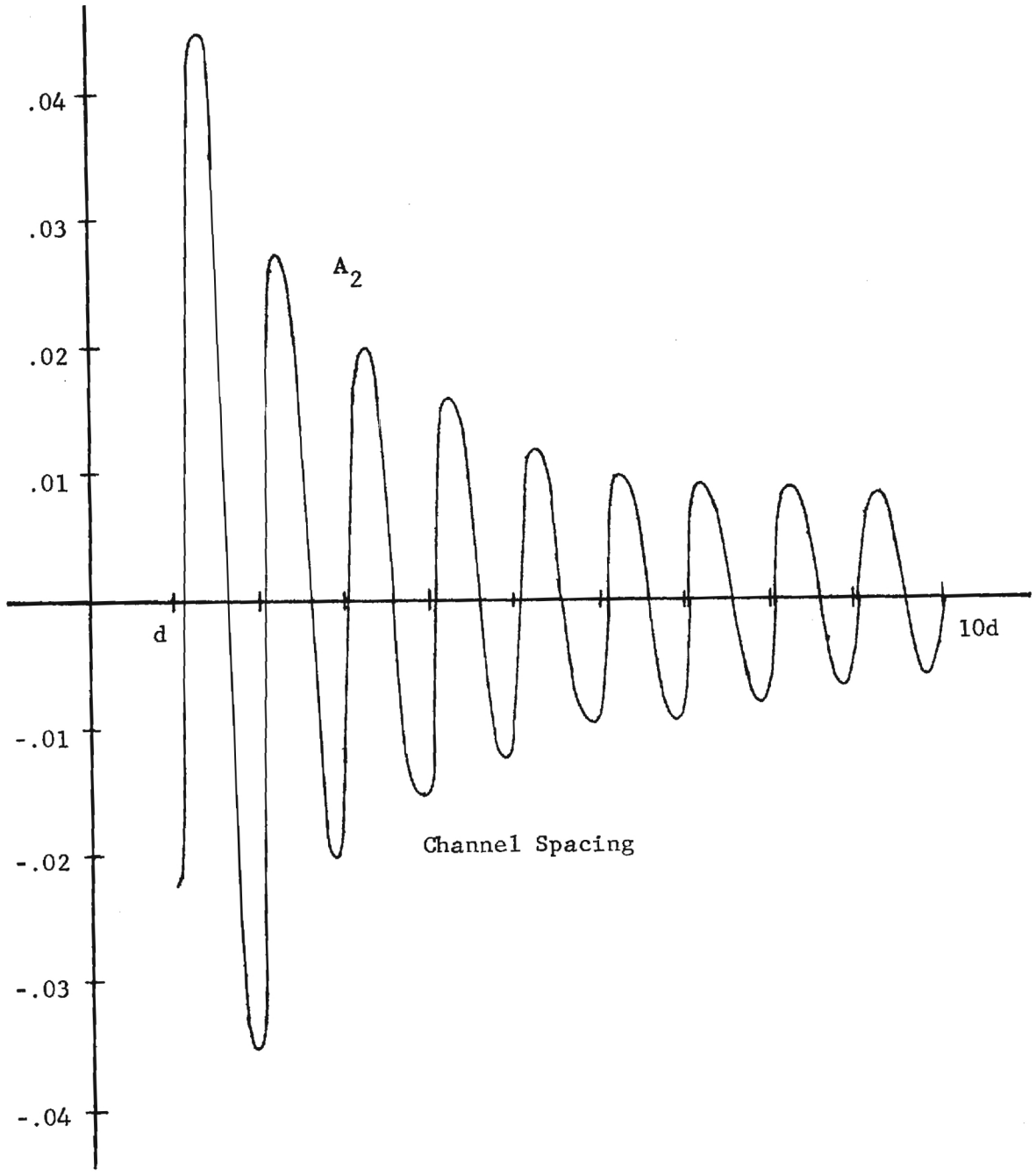


Figure 5-16(cont.) Variations in Crosstalk Amplitude Factors with Increasing Channel Spacing; (b) Crosstalk Amplitude Factor A_2.

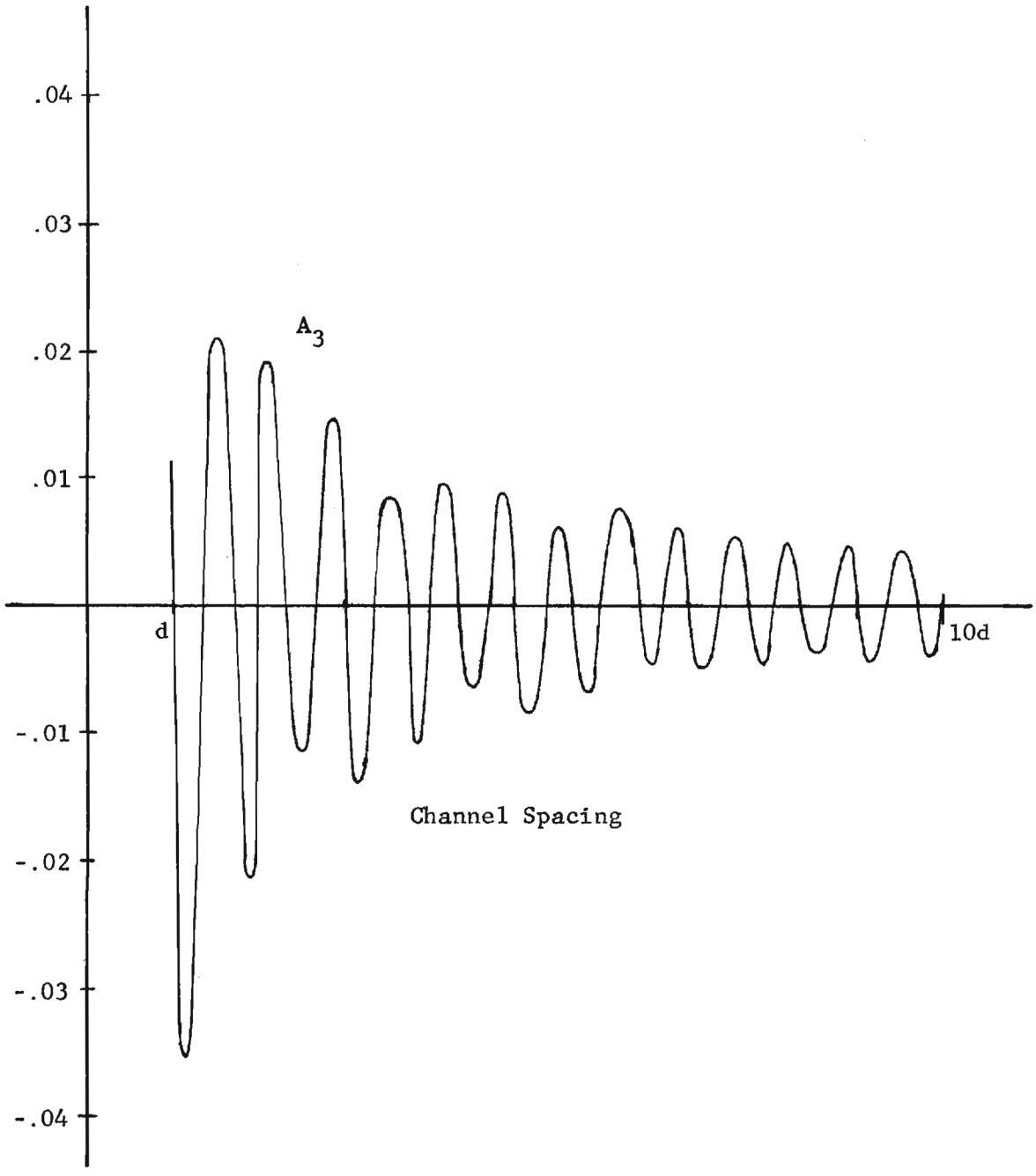


Figure 5-16(cont.) Variations in Crosstalk Amplitude Factors with Increasing Channel Spacing; (c) Crosstalk Amplitude Factor A_3.

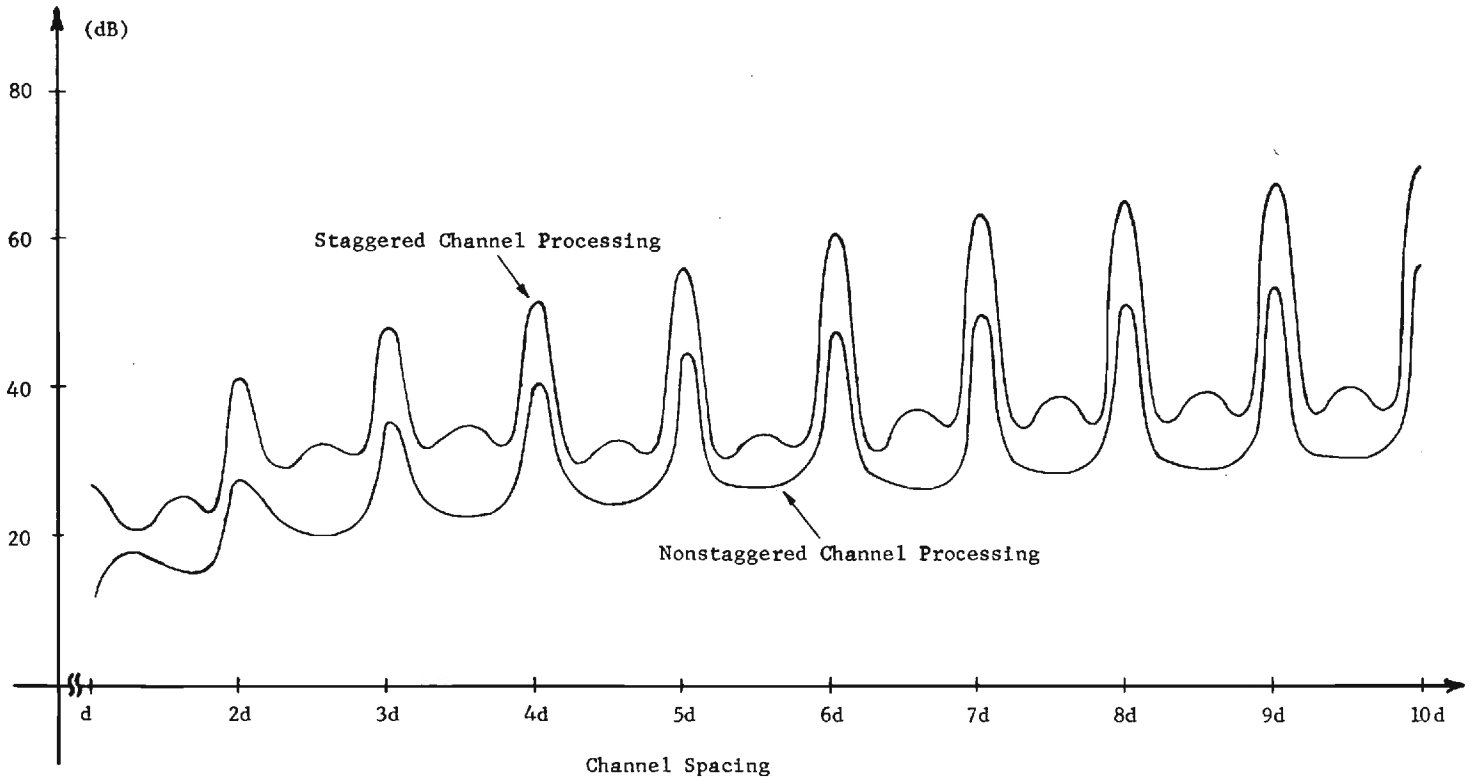


Figure 5-17. Variations in Signal to Crosstalk Power Ratio with Channel Spacing.

for it enables us to take advantage of the peaks in the channel signal-to-crosstalk noise ratio.

Experimental Implementation - Demonstration of Basic System Capabilities

A general channelized version of the frequency-variant heterodyne processing system has been implemented experimentally, and the basic operational characteristics of the system have been demonstrated. A block diagram of the experimental system and the supporting electronics is shown in Fig. 5-18. The optical configuration used in the processor was the frequency-multiplexed, common optical system with a moving film sound track serving to input the combined input signal/local oscillator. A film transport system was built from a surplus magnetic tape drive powered by a synchronous motor to provide control over film transport velocity. Included in the supporting electronics was a photomultiplier tube to detect the processor output signal. An additional photodetector was used to detect a portion of the light illuminating the sound track. The processor output signal was normalized by the output of this second detector in an analog dividing circuit for reasons we discuss below. An amplifier and electronic bandpass filter were then used to boost signal level and eliminate undesired portions of the processor output. An analog multiplier, a tunable oscillator, and a second bandpass filter were used if post-detection heterodyning was necessary and the final system output was recorded on magnetic tape.

The density modulated sound tracks used in the experimental demonstrations were recorded using the film transport system and a commercially available sound track recording unit (Back Auricon Modulite). A slow speed, blue sensitive film (Kodak 7476 Microfilm) was used for the sound

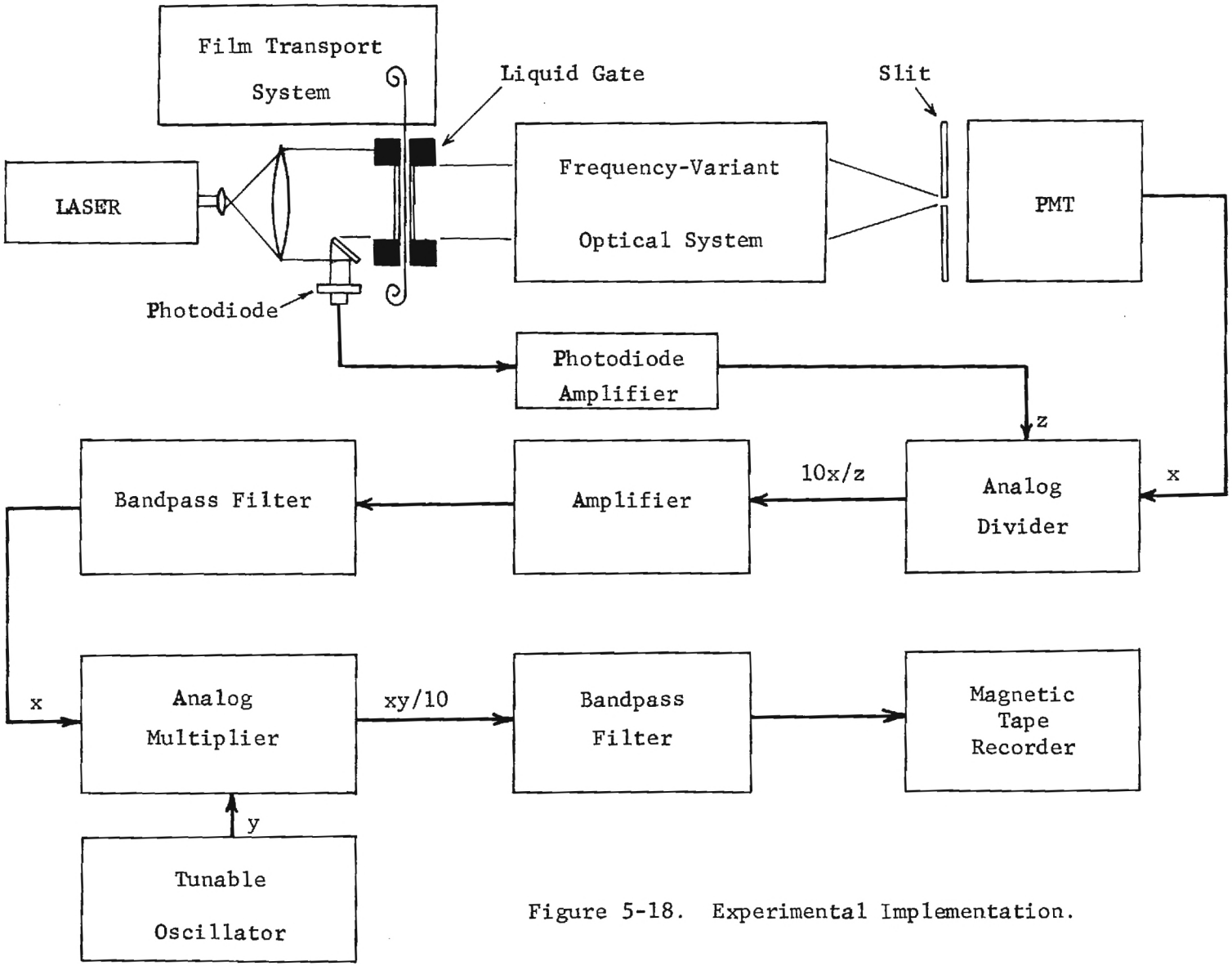


Figure 5-18. Experimental Implementation.

tracks to enable handling in relatively bright safe light conditions. The slow speed of the film required a slow transport velocity for sufficient film exposure. Driving the synchronous motor of the transport system at 100 Hz provided a transport velocity of 6.2 cm/sec, which was found to be adequate for proper exposure. Curves showing the developed film amplitude transmittance vs. applied galvonometer voltage (the t-v characteristic) for several values of lamp filament voltage are shown in Fig. 5-19. In order to properly record a signal in this fashion it was necessary that we record on a linear portion of the t-v characteristic.

The recording unit was originally designed to record voice band signals in a 35mm motion picture camera. The film recording velocity in such a camera is approximately 60 cm/sec (≃24 frames/second x 24.8 cm/frame). The much slower recording velocity used in these experiments reduced the bandwidth of the recorded signal because of the finite width of the exposing line source. A curve showing the recorded signal level vs. frequency for constant amplitude input is shown in Fig. 5-20. The recorded signal levels were measured directly from the level of the light diffracted by the sound track recording. From this curve, it is clear that the combined input signal/local oscillator had to be restricted to the range 0 to 1500 Hz.

In this system, a wideband noise signal was present in the PMT output signal, with or without system input, as a result of noise in the laser illumination source. (For the Helium-Neon laser used in these experiments, the beam intensity noise in the range 0 to 100 KHz was approximately 3% of the rms beam intensity.) With a system input, the laser noise appeared in the PMT output signal as a multiplicative noise term. In order to reduce

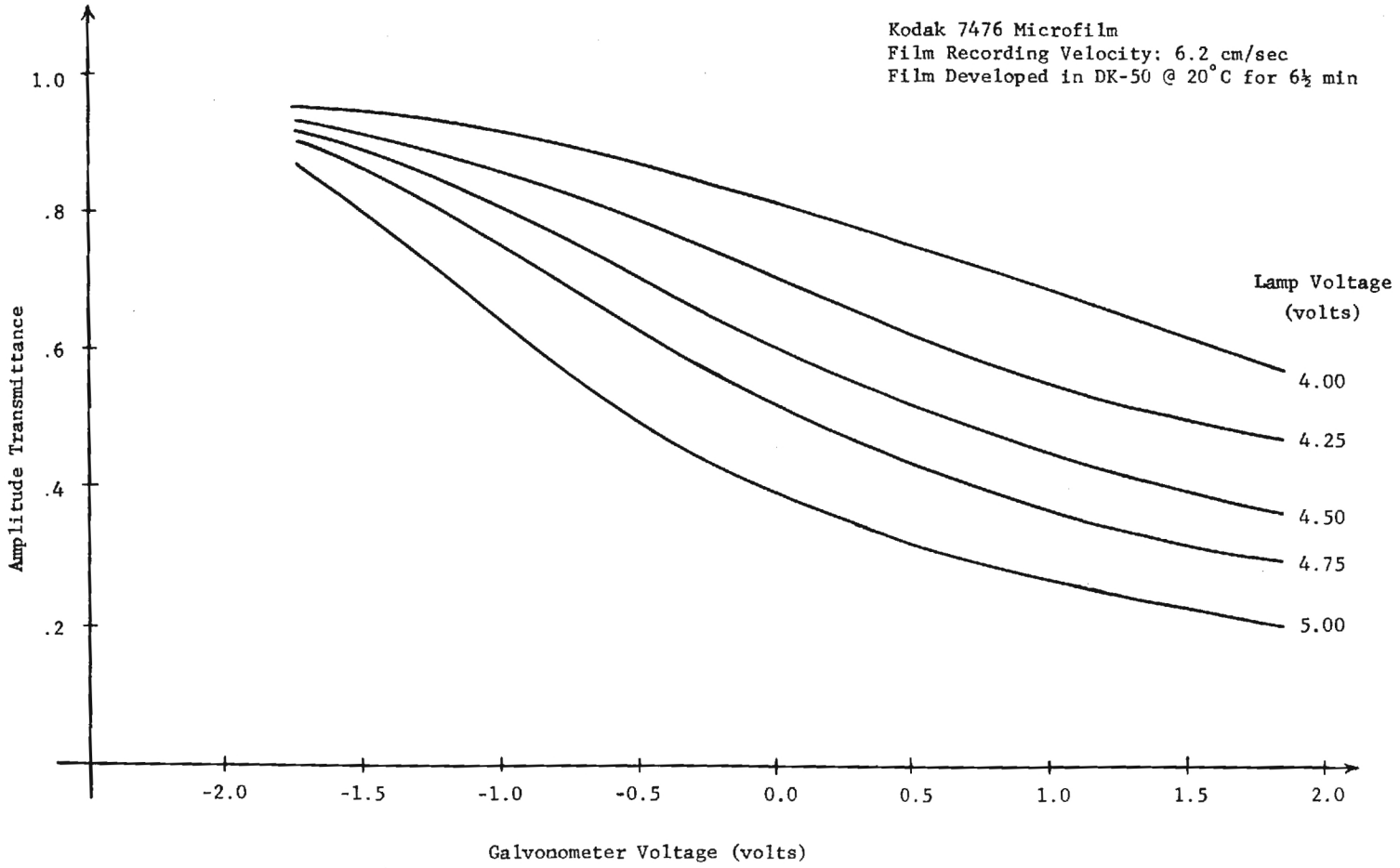


Figure 5-19. t-v Characteristic for Recording Film Soundtracks.

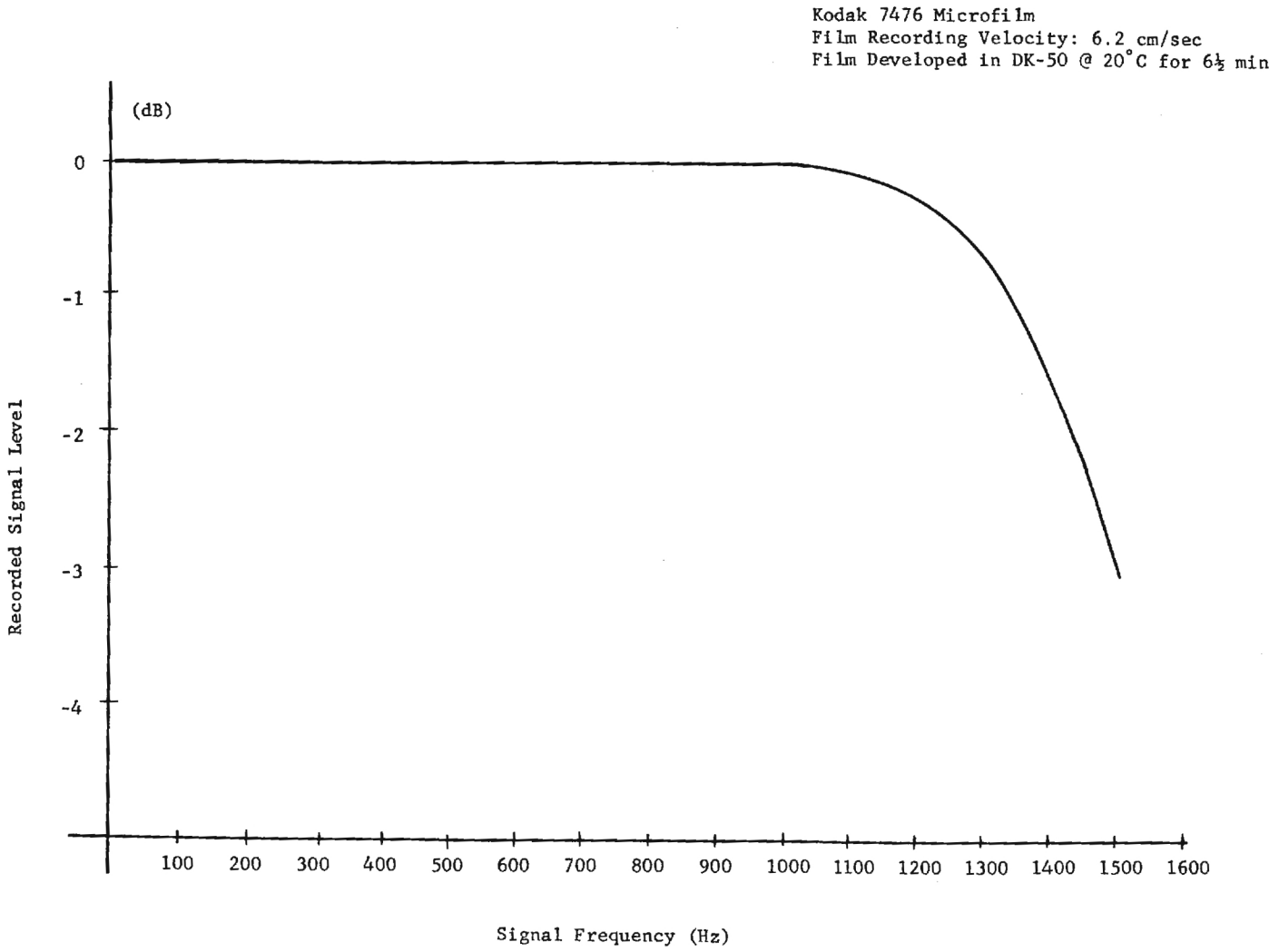


Figure 5-20. Recorded Signal Level Variations with Signal Frequency.

substantially this term, a portion of the unmodulated plane wave illuminating the sound track was picked off and detected by a photodiode. The system output from the PMT was normalized by the photodiode output signal, and the laser noise was thereby virtually eliminated.

There were two major practical problems encountered in the experimental implementation of the processing system that were not completely resolved in the course of this research. These problems are not inherent to the processing technique itself; rather, they were a result of the use of film as an input medium.

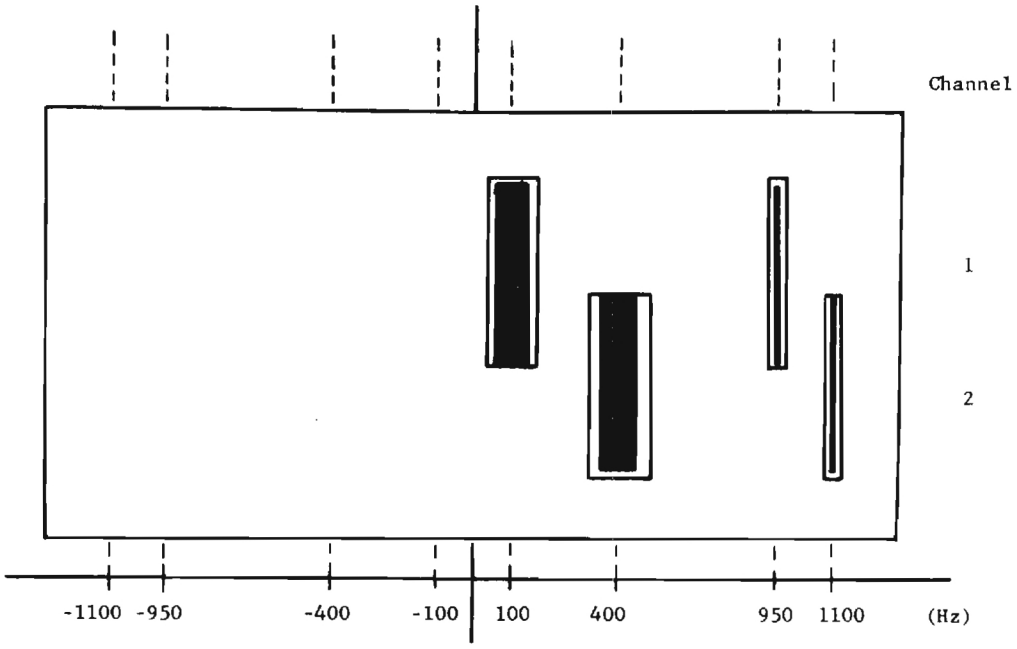
First, there was a problem with noise in the system output induced by variations in emulsion thickness of the film sound tracks. Emulsion thickness variations cause the optical path length through the film to change from point to point along the film strip. These variations caused noise-like variations in the phase of the sound track amplitude transmittance, introducing a substantial noise signal on top of the input signal and local oscillator. It was found that, without compensation, the emulsion-induced noise level was high enough to effectively obliterate the input signal in a coherent playback. Most, but not all, of this noise was eliminated by the use of a liquid gate to compensate for the emulsion thickness variations. In the liquid gate, the film strip was placed between two optically flat pieces of glass and the remaining air space was filled with xylene, a liquid whose index of refraction approximately equals that of the film emulsion. The optical path length through the film and liquid gate was then approximately constant from point to point, eliminating most of the phase variations in the sound track amplitude transmittance.

The second major problem with the film sound track input was the

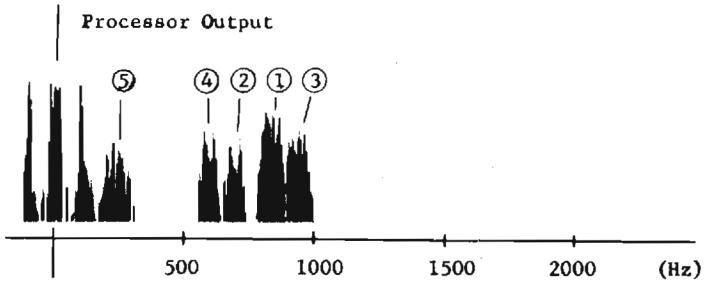
low diffraction efficiency of the recorded film strips. The maximum signal variations applied to the recorder galvonometer were restricted in order to record the signal on a linear portion of the film t-v characteristic. Because of this restriction, the dynamic range of the recorded transmittance variations was small, and, as a result, only a small portion of the light illuminating the film strips was diffracted by the recorded signal. The light levels of the signal and local oscillator distributions in the system were therefore extremely low. The residual effects of the emulsion thickness variations not totally eliminated by the liquid gate introduced noise signals at levels often approaching that of the desired input signal and local oscillator. In addition, with such low light levels, the thermal noise of the PMT (shot noise and dark current) became a substantial portion of the detector output. In order to maintain the light levels in individual channels at levels that could be detected by the PMT without being lost in the thermal noise, the channel widths in the y-direction had to be fairly wide - between 1mm and 2mm. However, the sound tracks recorded were narrow in the y-direction, only about 4mm wide. Thus, due to the low light levels, only 2 or 3 channels could be successfully operated in this implementation, and these remained quite noisy.

Despite the low number of channels available in this experimental implementation, the basic capabilities of the frequency-variant processor in mapping signal components have been demonstrated. In the first series of experiments the frequency ranges of the various terms in a channel output (i.e.,the desired output, the self-product terms, and the cross product terms) were verified for the three different methods of mixing signal components: (1) positive frequency signal components with positive fre-

quency local oscillator components, (2) positive frequency signal components with negative frequency local oscillator components, and (3) staggered channel mixing. The results from these three different methods of mixing are shown in Figs. 5-21, 5-22 and 5-23, respectively. In each experiment, the input signal consisted of two narrowband noise-like components entered at 100 Hz and at 400 Hz and both with bandwidth of approximately 100 Hz; the local oscillator consisted of two sinusoids at 950 Hz and at 1100 Hz. In part (a) of each figure the channelized transform plane mask used to perform the mixing is shown superposed with the frequency-multiplexed transform distributions. Because of the low light levels in the processing system, the effect of the image smoothing is simulated by overlapping the channels in the transform plane mask. In each case, the input signal was processed with the mask shown and the system output recorded on magnetic tape without filtering or electronic heterodyning. The recorded output signal was then analyzed using an electronic spectrum analyzer. (Signals were analyzed on a Kay Sonograph operated in a sectioning mode.) The output of the electronic spectrum analyzer is shown in part (b) of each figure. In the first two methods of mixing, we see that the desired output terms and the cross product crosstalk terms in both channels are closely spaced. (The local oscillator frequencies were chosen such that crosstalk terms would not appear at the same frequency as the desired output terms to facilitate the illustration. With a different choice of local oscillators, these terms could appear at the same frequencies as the desired terms and would not therefore be removable by filtering.) However, in the staggered channel mixing, the desired output terms are well separated from the other components of



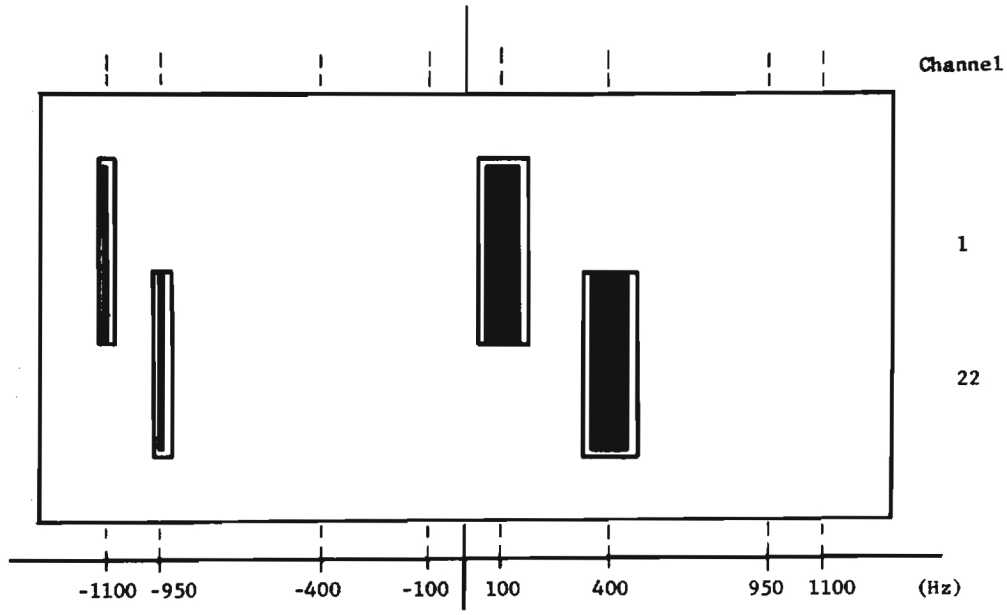
(a)



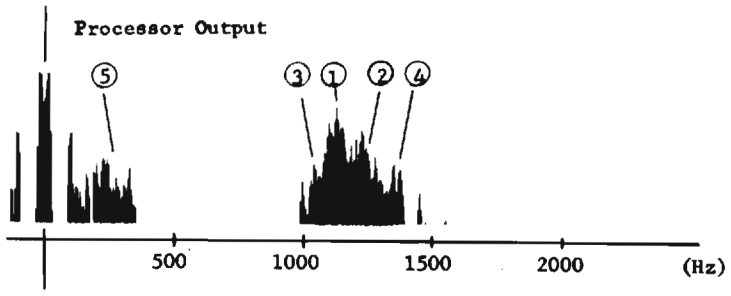
(b)

Desired Output { Channel 1 ①, Channel 2 ②
Crossproduct { Channel 1 ③
Crosstalk { Channel 2 ④
Self-product Crosstalk ⑤

Figure 5-21. Experiment No. 1: Positive-Positive Mixing; (a) Transform Plane Mask, (b) Processor Output.



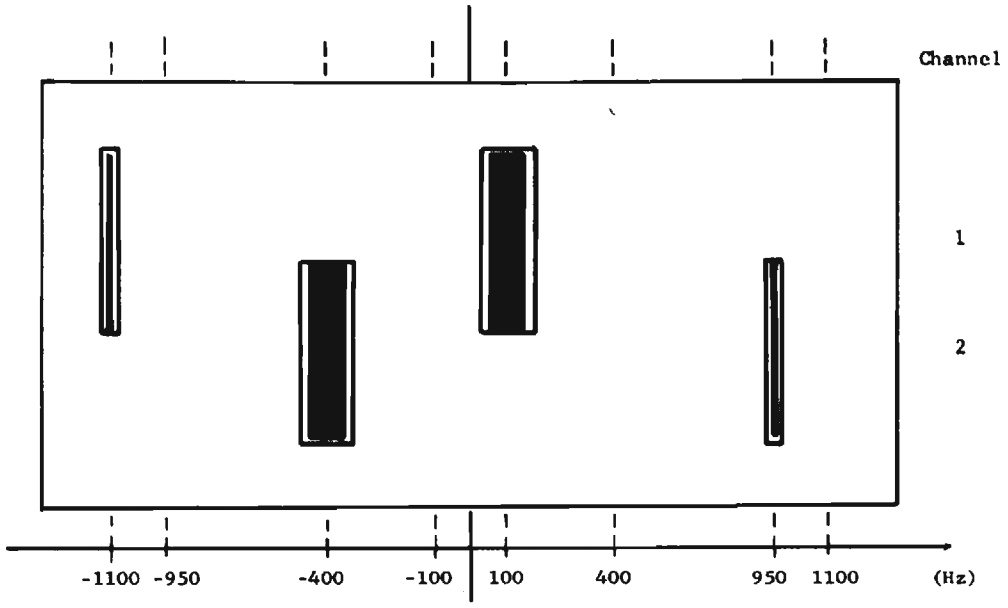
(a)



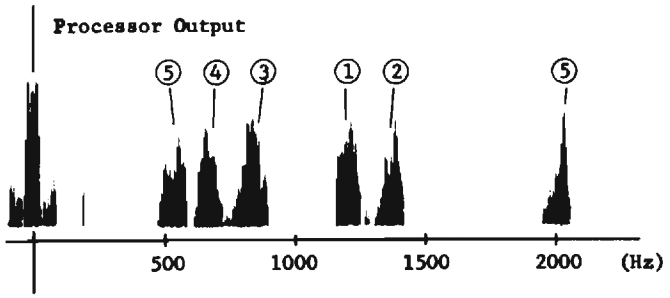
(b)

Desired Output { Channel 1 ①, Channel 2 ②
Crossproduct Crosstalk { Channel 1 ③, Channel 2 ④
Self-product Crosstalk ⑤

Figure 5-22. Experiment No. 2: Positive-Negative Mixing; (a) Transform Plane Mask, (b) Processor Output.



(a)



(b)

Desired Output	Channel 1	①
	Channel 2	②
Crossproduct Crosstalk	Channel 1	③
	Channel 2	④
Self-product Crosstalk		⑤

Figure 5-23. Experiment No. 3: Staggered Channel Mixing; (a) Transform Plane Mask, (b) Processor Output.

the system output.

A second series of experiments were performed to demonstrate the system's capabilities in rearranging signal components in the frequency domain with frequency-dependent resolution. To demonstrate this aspect of the system the input signal was chosen to be two frequency modulated sinusoids, one at 300 Hz, the other at 500 Hz. The frequency modulation was narrowband (modulation index =0.5) with the modulating signals also sinusoids. To distinguish the two signals, the carrier at 300 Hz was modulated with a 75 Hz sinusoid while the carrier at 500 Hz was modulated with a 50 Hz sinusoid. The two frequency-modulated sinusoids were recorded on a film strip with a local oscillator consisting of sinusoids at a number of different frequencies. The results of processing this signal are shown in Figs. 5-24 through 5-29. Again, part (a) of each figure shows the channelized transform plane mask used to process the signal superposed with the transform distribution of the signal and local oscillator. In this series of experiments, staggered channel mixing was used and the crosstalk and self product terms were eliminated by bandpass filtering. The signal was then returned to the baseband region by electronic heterodyning and recorded. Electronic spectrum analysis was performed on the recorded signals with the result shown in part (b) of each figure. In Fig. 5-24, we show the results of a simple playback of the input signal showing the two carrier signals and the sidebands resulting from the frequency modulation. In the next three figures, Figs. 5-25 through 5-27, we show the results of a slight bandwidth compression with frequency dependent resolution. In each case, the signal is processed so that the 300 Hz carrier appears in the output at 300 Hz while the 500 Hz carrier is downshifted to 450 Hz.

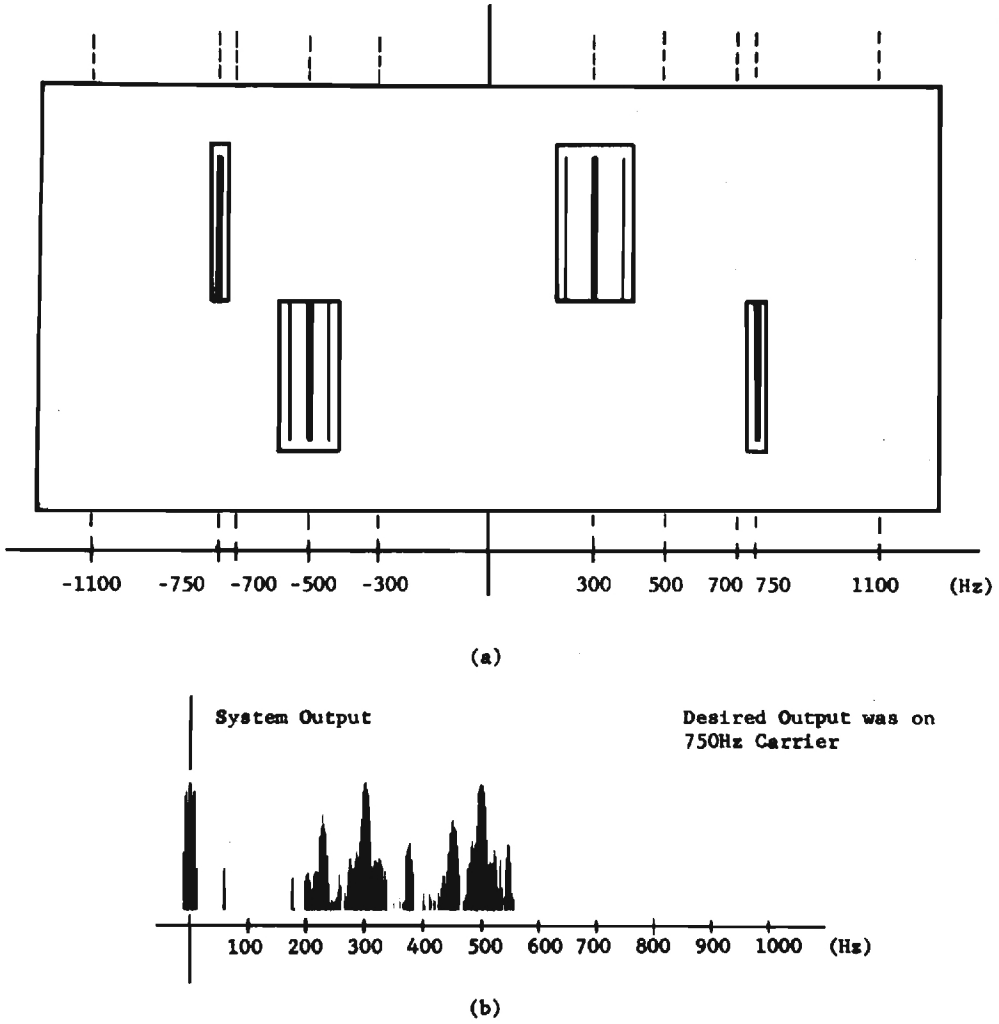


Figure 5-24. Experiment No. 4: Playback of Recorded Signal; (a) Transform Plane Mask, (b) System Output.

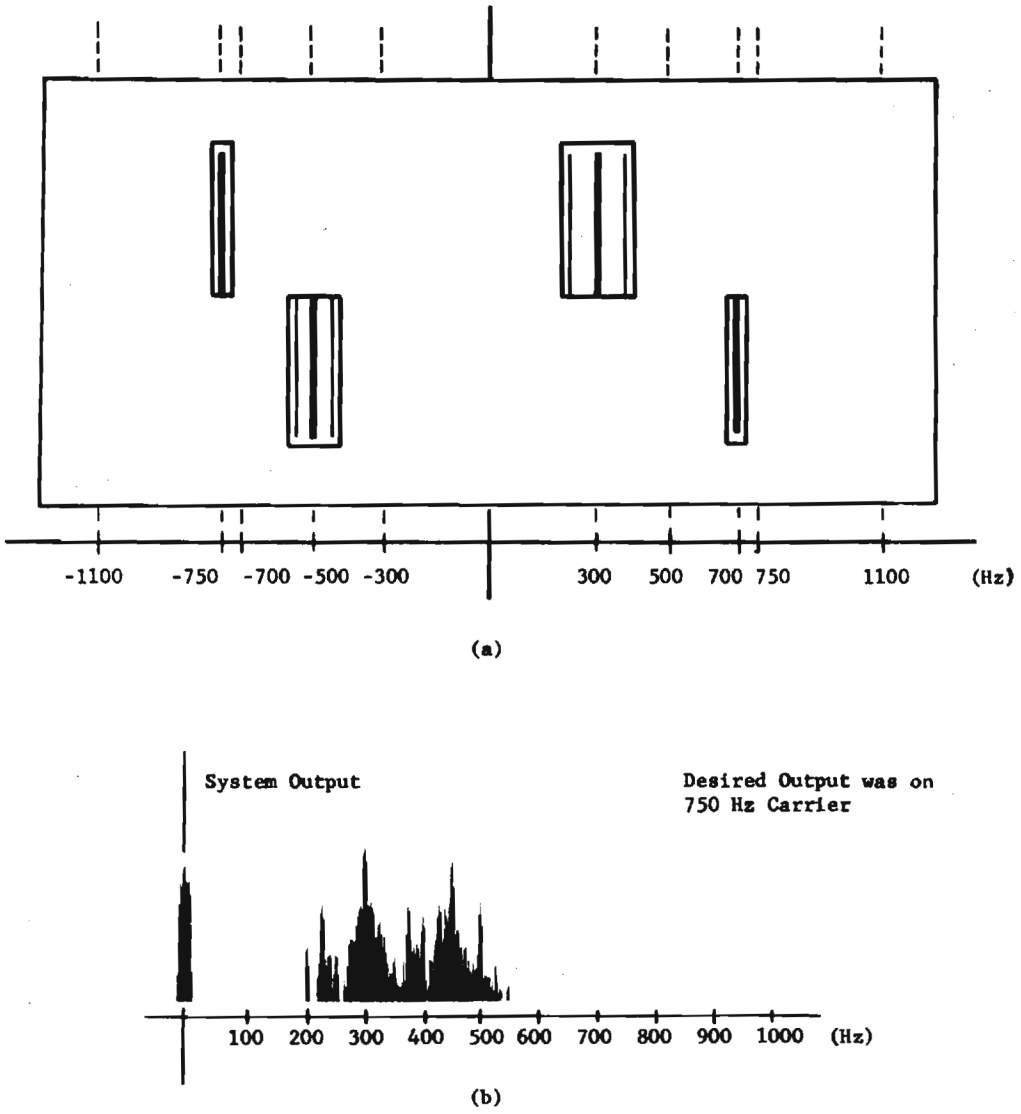


Figure 5-25. Experiment No. 5: Bandwidth Compression; (a) Transform Plane Mask, (b) System Output.

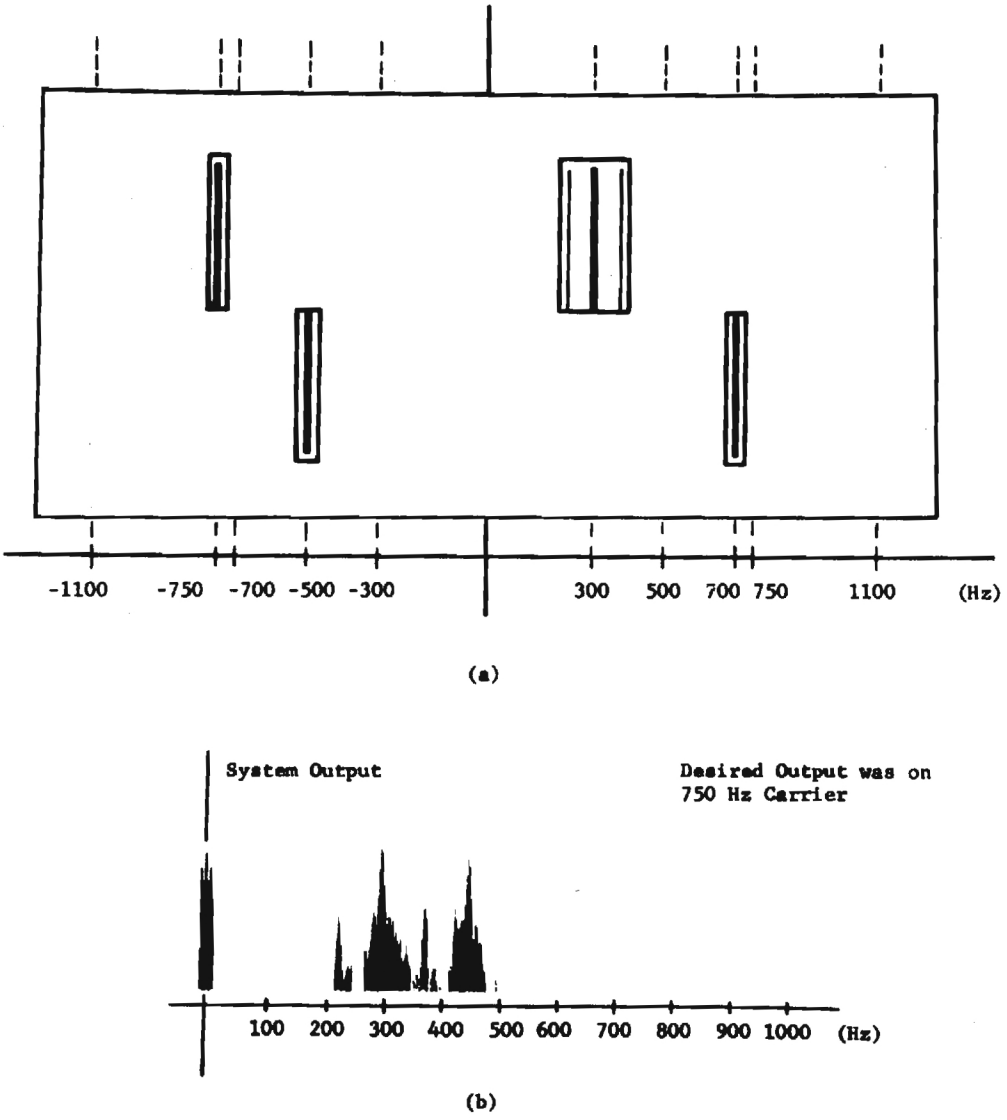
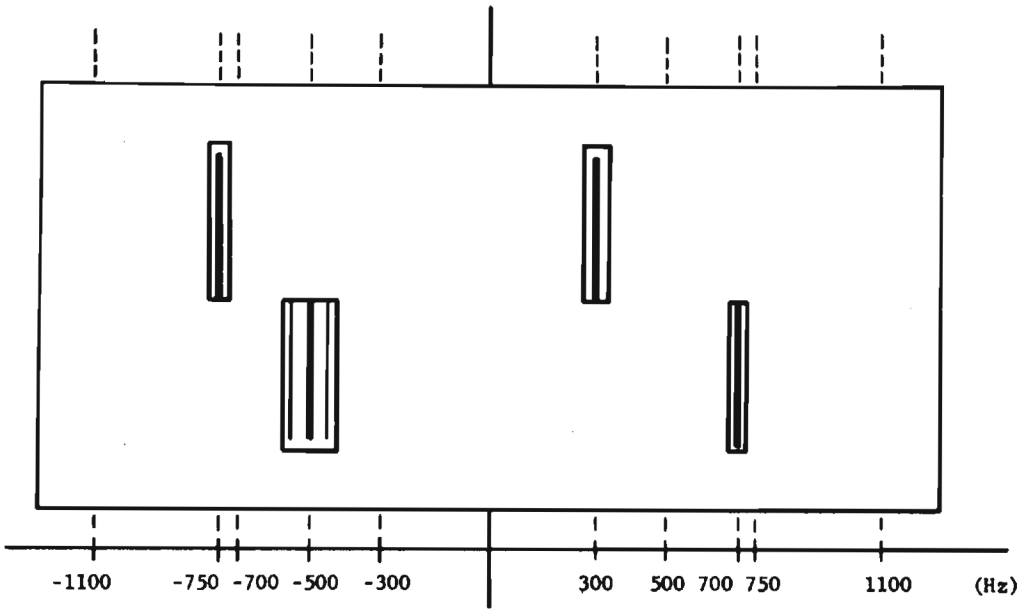
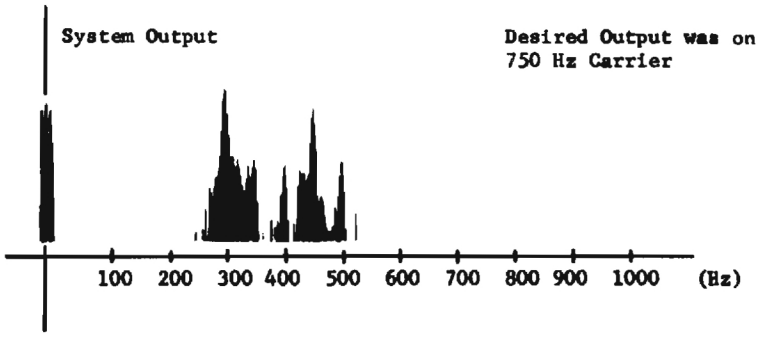


Figure 5-26. Experiment No. 6: Bandwidth Compression with Unequal Channel Resolution; (a) Transform Plane Mask, (b) System Output.



(a)



(b)

Figure 5-27. Experiment No. 7: Bandwidth Compression with Unequal Channel Resolution; (a) Transform Plane Mask, (b) System Output.

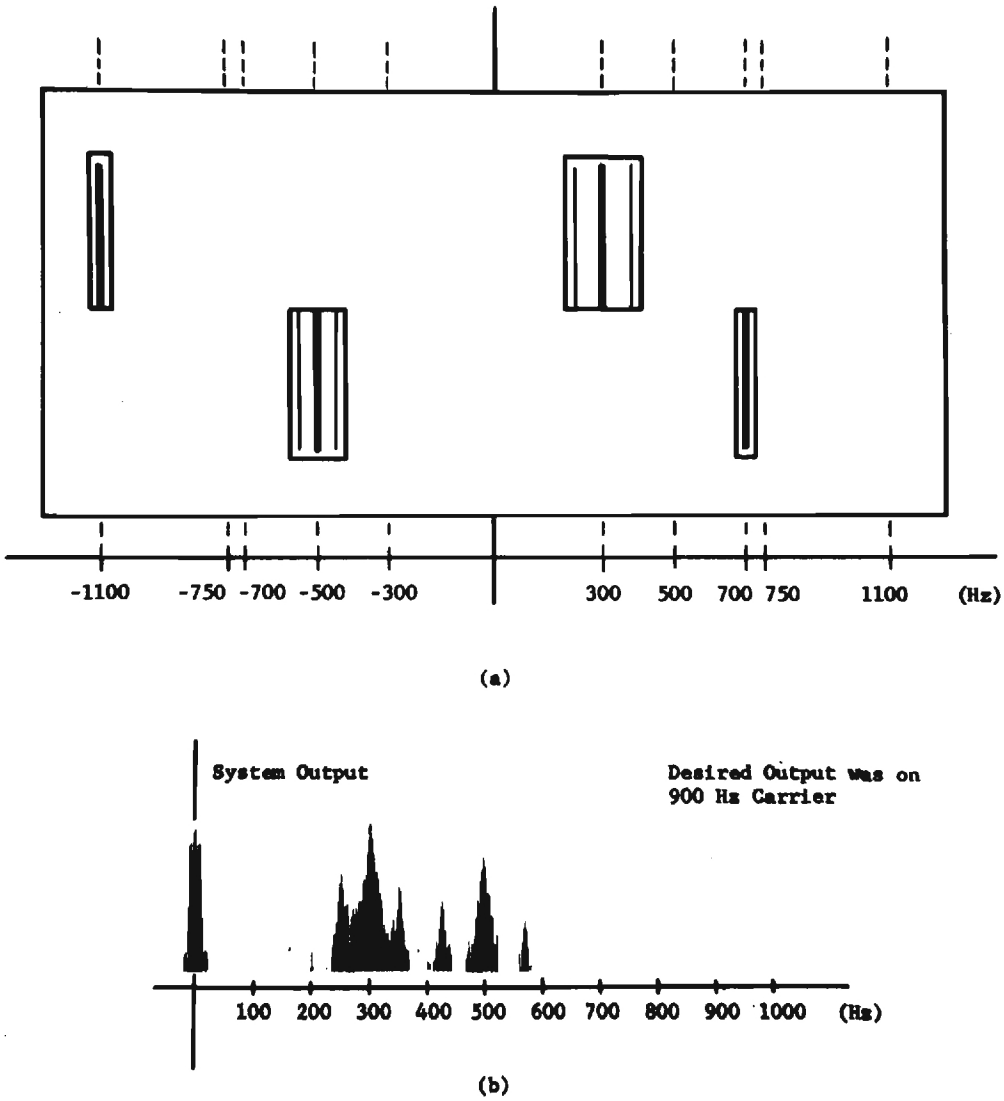


Figure 5-28. Experiment No. 8: Spectral Rearrangement; (a) Transform Plane Mask, (b) System Output.

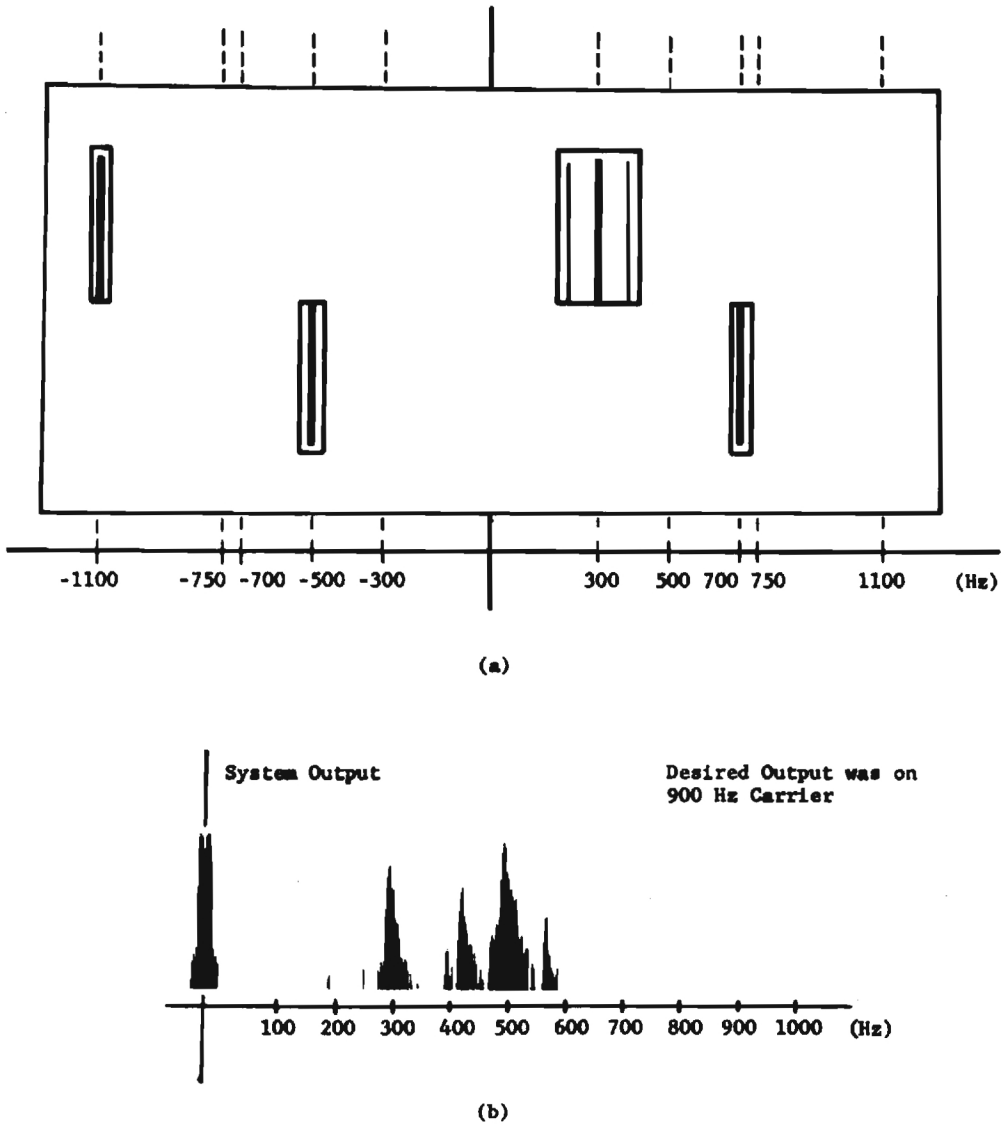


Figure 5-29. Experiment No. 9: Spectral Rearrangement with Unequal Channel Resolution; (a) Transform Plane Mask, (b) System Output.

In the experiment illustrated in Fig. 5-25, the frequency resolution of both channels was the same and sufficient for the FM sidebands of each carrier to be processed. In the experiments shown in Figs. 5-26 and 5-27, the frequency resolution in the channels was different with the FM sidebands of the higher frequency carrier removed from the output in the first case, and those of the lower frequency carrier removed from the output in the second case. In the next two experiments, shown in Fig. 5-28 and 5-29, the two carrier signals are switched in frequency, i.e. the 300 Hz carrier appears in the output at 500 Hz and the 500 Hz carrier appears at 300 Hz. This is evident in the output signal spectrum shown in Fig. 5-28(b) where now the 500 Hz carrier is modulated by the 75 Hz signal and the 300 Hz carrier is modulated by the 50 Hz signal. The results of the experiment shown in Fig. 5-29 again show the ability to process the signal with frequency dependent resolution.

Finally, in the results shown in Figs. 5-30, 5-31, and 5-32, we illustrate the general characteristics of the channelized processor in both time and frequency. In the first figure, Fig. 5-30, we show the time-frequency diagram (a sonograph) obtained from the direct playback of the recorded signal. As shown, the input is a swept frequency tone, sweeping from 200 Hz to 700 Hz in 2 seconds and the dropping back to 200 to begin another sweep. In parts (a) of Figs. 5-31 and 5-32, we show the transform plane masks used in the two frequency mapping operations performed on this input signal. In both cases, the frequency range 200 Hz to 700 Hz is divided equally among three separate channels. In the experiment illustrated in Fig. 5-31, the local oscillators for the channels are chosen to perform a bandwidth compression. As seen in the output signal sonogram, shown in

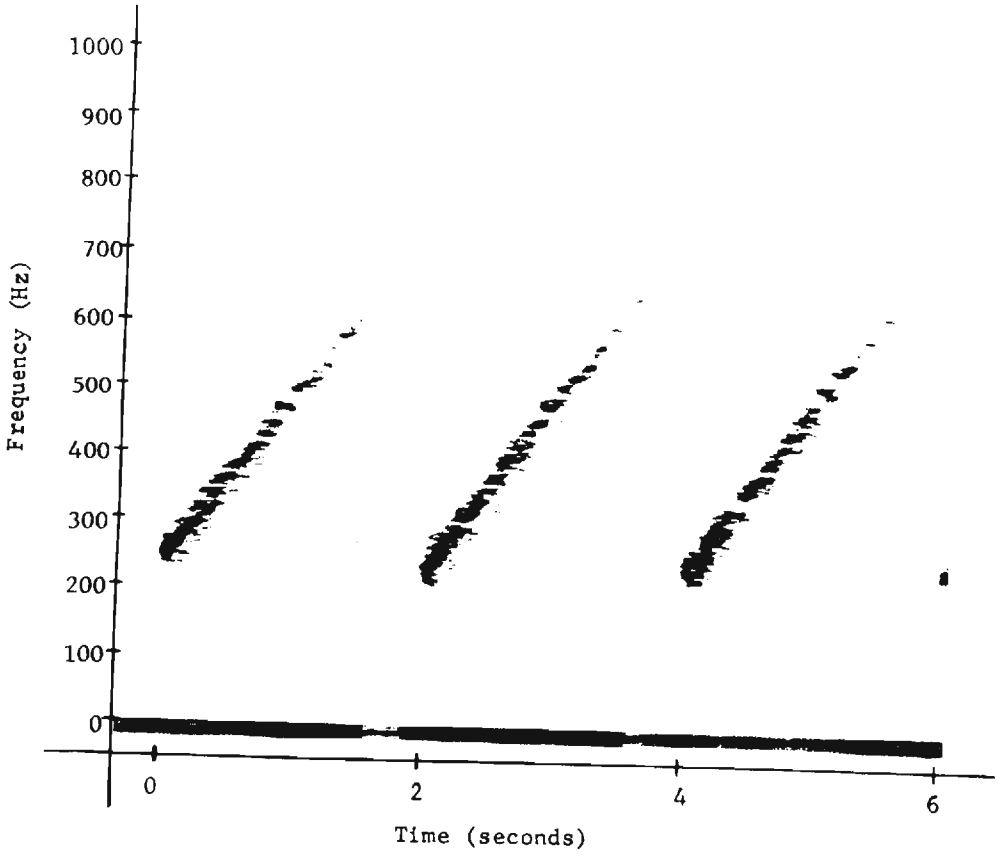
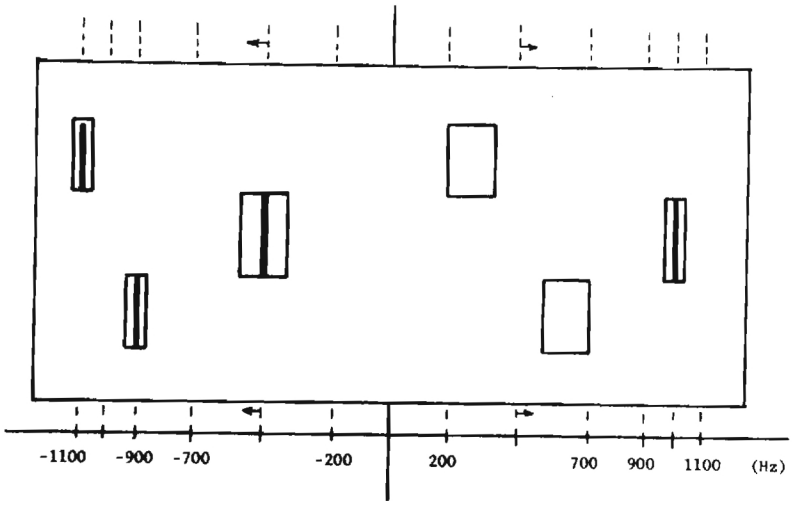
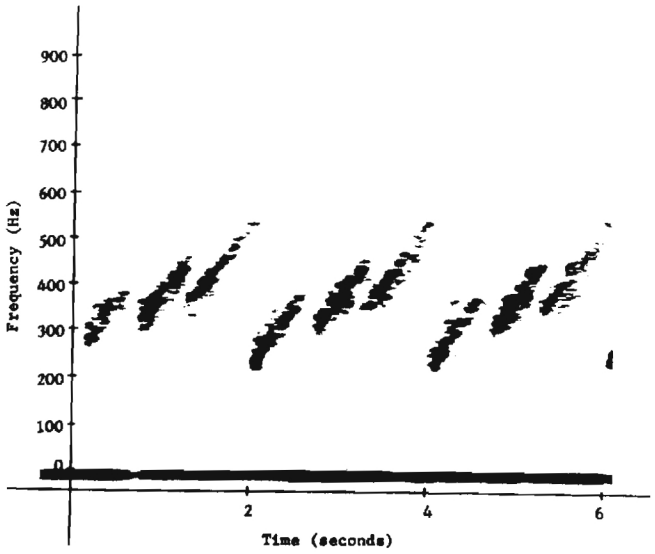


Figure 5-30. Time-Frequency Diagram for Swept Frequency Tone.

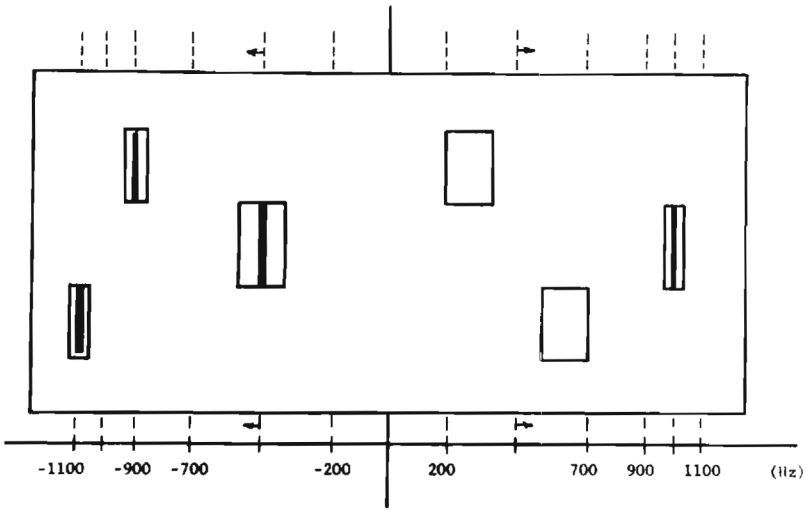


(a)

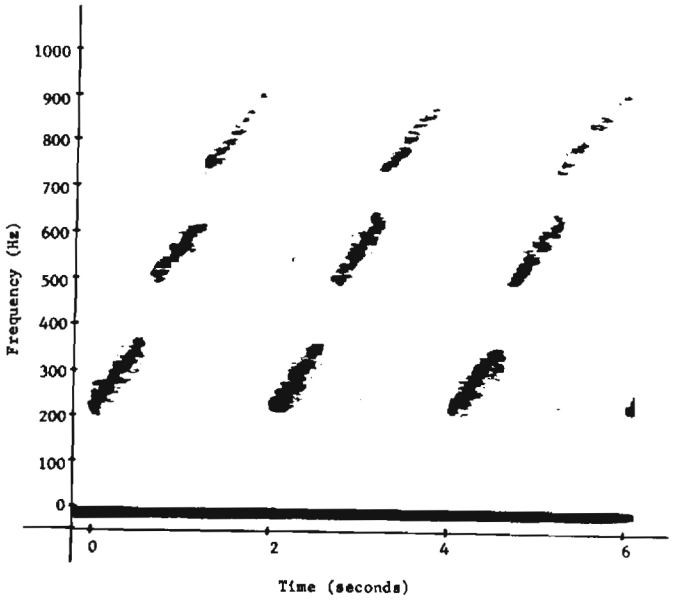


(b)

Figure 5-31. Experiment No. 10: Bandwidth Compression of Swept Frequency Tone; (a) Transform Plane Mask, (b) Time-Frequency Diagram of Compressed Signal.



(a)



(b)

Figure 5-32. Experiment No. 11: Bandwidth Expansion of Swept Frequency Tone; (a) Transform Plane Mask (b) Time-Frequency Diagram of Expanded Signal.

part (b) of this figure, the channelized compression results in an output consisting of three shorter sweeping tones in the range 200 Hz to 500 Hz. In Fig. 5-32, we illustrate the results of a bandwidth expansion where, again because of the channelization, the resulting output signal consists of three short sweeping segments, now in the range from 200 Hz to 900 Hz.

CHAPTER VI

CONCLUSION

In this work, we have described new capabilities of coherent optical systems in processing 1-D information. These new capabilities were seen to be a direct result of exploiting the second degree of freedom available in an optical system when processing 1-D signals.

The essential features of the new processing techniques were demonstrated in the general mapping of spatial frequencies performed by the frequency-variant spectrum analyzer. The ability to control position and resolution of frequency components within the spectral display has made possible log-frequency and constant-Q spectrum analyzes with optical systems. A general description of the space-bandwidth product applicable to the frequency-variant system has also been developed to enable the determination of system information handling capacity with arbitrary mapping characteristics.

A discussion of the two fundamental implementations of general 1-D linear operations was offered to clarify the methods by which these general operations can be performed in optical systems.

The basic theory of operation of a general frequency-variant optical heterodyne system for signal waveform processing has been developed, and the method required for proper interpretation of the system's operation in both time and frequency was described. A practical implementation of the heterodyne processing system employing channelized mapping characteristics has been described and its basic operational capabilities have been demon-

strated. In the development of the practical implementation, special attention was focused on non-interferometric processor configurations and on processing techniques designed to insure channel independence.

There are a number of topics for future research opened by this work, primarily in the area of signal waveform processing with the frequency-variant optical heterodyne system.

A more precise analytical description of the general nonchannelized system operation in terms of time-frequency components would be highly desirable if such a system is to be designed and implemented. The possibility of an analytic description of this type raises questions concerning the validity and the accuracy of signal representations with elementary components whose time-frequency resolution characteristics change with frequency. The work of Markel and Carey suggests that such a representation is valid, at least in a qualitative sense, for speech signals [56].

From another point of view, the limitations on the maximum number of channels in a channelized system suggests that finite vertical imaging resolution in the optical system might force a channelized structure on the general nonchannelized system. Using shift-variant sampling techniques [39,58], a channelized optical mapping may be found, equivalent to the nonchannelized optical mapping of both the input signal and the local oscillator. The general heterodyne mapping could then be modeled as that occurring in the channelized system, probably with the introduction of channel crosstalk to properly describe the system operation.

In an experimental implementation, considerable improvement on the diffraction efficiency of film sound tracks could be accomplished by computer controlled recording to use the entire dynamic range of the film.

This would certainly reduce the problem of low light levels, but the residual effects of film emulsion thickness variations would continue to introduce significant noise in the heterodyne processor output. It is therefore recommended that acoustooptic input devices be used in any future implementation. This presents a problem for the processing of audio frequency signals as acoustooptic devices normally operate at radio frequencies. However, it may be possible to employ speed-up processing or recently developed time-integration techniques [59] for near real-time processing of audio signals.

Assuming that audio frequency signals can be effectively processed in the optical processor, a number of interesting psychophysical tests are suggested in the area of speech or music processing. Many audio signal processing systems (e.g. speech vocoders) compromise between time and frequency resolution by processing the signal with fixed time windows approximately 30 msec long. The human ear, however, operates with what Gabor described as an adjustable time constant, at times displaying time resolution approaching 5 msec, and at others displaying frequency resolution approaching 4 Hz or better (requiring an effective time-window of 250 msec). By matching the processing system's resolution characteristics with those of the ear, the ear should be largely unaware of any processing taking place. With the frequency-variant processing system it may prove possible to obtain a reasonable match of the ear's resolution characteristics by providing, simultaneously, different degrees of resolution in different frequency regions, e.g. good frequency resolution in the low frequency range of most first forments in speech and good time resolution in the high frequency range of many speech or music cues.

Finally, there are potential applications of the frequency-variant signal waveform processor in the shift-variant processing of 2-D information. The techniques of general 1-D linear operations for 1-D spatial signals are not directly extendable to 2-D signal processing because these techniques rely on the second dimension to perform the processing operation (direct extension would require a 4-D optical system). However, Rhodes has recently described a method by which a time-varying 2-D image is encoded on a temporal signal waveform where each element of the image modulates a different temporal frequency carrier [60]. Application of the frequency-variant processing capabilities of the heterodyne processor to this encoded signal then translates into the shift-(space-)variant processing of the 2-D image.

BIBLIOGRAPHY

1. P. M. Duffieux, L'intégrale de Fourier et ses applications á l'optique, (Faculté des Sciences de Besancon, Rennes, 1946).

2. O. H. Schade, "Electro-Optical Characteristics of Television Systems," RCA Review, Vol. 9, No. 1-4 (1948).

3. P. Elias, D. S. Grey, and D. Z. Robinson, "Fourier Treatment of Optical Processes," J. Opt. Soc. Am., Vol. 42, pp. 127-134 (1952).

4. J. E. Rhodes, Jr., "Analysis and Synthesis of Optical Images," Am. J. Phys., Vol. 21, pp. 337-343 (1953).

5. A. Maréchal and P. Croce, "Un filtre de frequences spatiales pour l'amelioration du contrast des images optiques," C. R. Acad. Sci., Vol. 127, p. 607 (1953).

6. E. L. O'Neill, "Spatial Filtering in Optics," IRE Trans. Inform. Theory, Vol. IT-2, pp. 56-65 (1956).

7. P. Elias, "Optics and Communication Theory," J. Opt. Soc. Am., Vol. 43, pp. 229-232 (1953).

8. J. W. Cooley and J. W. Tukey, "An Algorithm for the Machine Computation of Complex Fourier Series," Math. Comput., Vol. 19, pp. 297-301 (1965).

9. C. Braccini and A. V. Oppenheim, "Unequal Bandwidth Spectral Analysis Using Digital Frequency Warping," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-22, pp. 236-244 (1974).

10. O. Bryngdahl, "Geometrical Transformations in Optics," J. Opt. Soc. Am., Vol. 64, pp. 1092-1099 (1974).

11. F. P. Carlson and R. E. Francois, Jr., "Bandlimiting Considerations in Geometrical Distortion Systems," Proc. SPIE, Vol. 83, pp. 137-145 (1976).

12. F. P. Carlson and R. E. Francois, Jr., "Generalized Linear Processors for Coherent Optical Computers," Proc. IEEE, Vol. 65, pp. 10-18 (1977).

13. A. Papoulis, The Fourier Integral and Its Application, (McGraw-Hill, New York, 1962), p. 144.

14. R. Adler, "Interactions Between Light and Sound," IEEE Spectrum, pp. 42-54, May (1967).

15. J. W. Goodman, Introduction to Fourier Optics, (McGraw-Hill, New York, 1968).

16. K. Preston, Jr., Coherent Optical Computers, (McGraw-Hill, New York, 1972).

17. L. J. Cutrona, E. N. Leith, C. J. Palermo, and L. J. Porcello, "Optical Data Processing and Filtering Systems," IRE Trans. Inform. Theory, Vol. IT-6, pp. 386-400 (1960).

18. C. E. Thomas, "Optical Spectrum Analysis of Large Bandwidth Signals," Applied Optics, Vol. 5, p. 1782 (1966).

19. B. V. Markevitch, "Optical Processing of Wideband Signals," in Five Ampex Corporation Papers Presented at the Third Wideband Analog Recording Symposium, Rome Air Development Center, Griffiss AFB, New York, September, 1969.

20. W. T. Rhodes and W. R. Limburg, "Coherent and Non-Coherent Optical Processing of Analog Signals," in Proceedings of the 1972 Electro-Optical Systems Design Conference, New York, pp. 314-320, September, 1972.

21. F. T. S. Yu, "Synthesis of an Optical Sound Spectrograph," J. Acousto. Soc. Am., Vol. 51, p. 433-438 (1972).

22. J. L. Flanagan, Speech Analysis, Synthesis, and Processing, (Springer-Verlag, New York, 1972).

23. A. Papoulis, Systems and Transforms with Applications in Optics, (McGraw-Hill, New York, 1968), p. 95.

24. W. T. Rhodes, "Log-Frequency Variable Resolution Optical Spectrum Analysis Using Holographic Mapping Techniques," Optics Communications, Vol. 18, pp. 492-495 (1976).

25. G. M. Jenkens and D. G. Watts, Spectrum Analysis and Its Applications, (Holden-Day Company, San Francisco, 1969).

26. A. A. Karkevich, Spectra and Analysis, translated from the Russian, (Consultants Bureau, New York, 1960).

27. A. Papoulis, "Apodization for Optimum Imaging of Smooth Objects," J. Opt. Soc. Am., Vol. 62, p. 1423 (1972).

28. A. Papoulis, "Minimum-Bias Windows for High-Resolution Spectral Estimates," IEEE Trans. Inform. Theory, Vol. IT-9, p. 9 (1973).

29. H. Stark and B. Dimitriadis, "Minimum-Bias Spectral Estimation with a Coherent Optical Spectrum Analyzer," J. Opt. Soc. Am., Vol. 65, pp. 425-431 (1975).

30. L. J. Cutrona, "Recent Developments in Coherent Optical Technology," in Optical and Electro-Optical Information Processing, edited by J. T. Tippett, D. A. Kerkowitz, L. C. Clapp, C. J. Koester, and A. Vanderburgh, Jr., (MIT Press, Cambridge, 1965), pp. 83-124.

31. W. T. Rhodes and J. M. Florence, "Two-Dimensional Optical Processing of One-Dimensional Signals," J. Opt. Soc. Am., vol. 65, p. 1178 (1975).

32. W. T. Rhodes and J. M. Florence, "Two-Dimensional Optical Processing of One-Dimensional Signals," Technical Report No. GIT-EE-OIPL-75-1, Georgia Institute of Technology, July 1975.

33. W. T. Rhodes and J. M. Florence, "Frequency Variant Optical Signal Analysis," Applied Optics, vol. 15, pp. 3073-3079 (1976).

34. J. W. Goodman, P. Kellman, and E. W. Hansen, "Linear Space-Variant Optical Processing of 1-D Signals," Applied Optics, vol. 16, pp. 733-738 (1977).

35. R. J. Marks II, J. F. Walkup, M. O. Hagler, and T. F. Krile, "Space-Variant Processing of 1-D Signals," Applied Optics, vol. 16, pp. 739-745 (1977).

36. R. J. Marks II, Space-Variant Coherent Optical Processing, Ph.D. Thesis, Texas Tech University, Lubbock, Texas, 1977.

37. P. Kellman and J. W. Goodman, "Coherent Optical Implementation of 1-D Mellin Transforms," Applied Optics, vol. 16, pp. 2609-2610 (1977).

38. R. J. Marks II, J. F. Walkup, and T. F. Krile, "Ambiguity Function Display: An Improved Coherent Processor," Applied Optics, vol. 16, pp. 746-750 (1977).

39. T. Kailath, "Channel Characterization: Time-Variant Dispersive Channels," in Lectures on Communications Systems Theory, edited by E. J. Baghdady, (McGraw-Hill, New York, 1960).

40. R. N. Bracewell, The Fourier Transform and Its Application, (McGraw-Hill, New York, 1965), p. 113.

41. D. Casasent and D. Psaltis, "Scale Invariant Optical Transforms," Optical Engineering, vol. 15, pp. 258-261 (1976).

42. D. Casasent and D. Psaltis, "New Optical Transforms for Pattern Recognition," Proc. IEEE, vol. 65, pp. 77-84 (1977).

43. D. Casasent and F. Casasayas, "Optical Processing of Pulsed Doppler and Linear FM Step Radar Data," Applied Optics, vol. 14, pp. 1364-1372 (1975).

44. D. Casasent, "Spatial Light Modulators," Proc. IEEE, vol. 65, pp. 143-157 (1977).

45. D. Gabor, "Theory of Communication," J. of the IEE, vol. 93, pp. 429-457, London (1946).

46. F. Vilbig, "An Apparatus for Speech Compression and Expansion and for Replaying Visible Speech Records," J. Acoust. Soc. Am., vol. 22, pp. 754-761 (1960).

47. F. S. Cooper, "Spectrum Analysis," J. Acoust. Soc. Am., vol. 22, p. 761 (1950).

48. F. S. Cooper, A. M. Liberman, and J. M. Borst, "The Inter-Conversion of Audible and Visible Patterns as a Basis for Research in the Perception of Speech," Proc. Nat. Acad. Sci. U.S., vol. 37, p. 318 (1951).

49. R. Whitman, A. Korpel, and S. Lotsoff, "Application of Acoustic Bragg Diffraction to Optical Processing Techniques," presented at the Symposium on Modern Optics, Polytechnic Institute of Brooklyn, March 1967.

50. A. B. Larsen, "A Heterodyne Scanning System for Hologram Transmission," BSTJ, vol. 48, pp. 2507-2527 (1969).

51. A. Macovski, "Efficient Holography Using Temporal Modulation," Ph.D. Thesis, Stanford University, 1968 (University Microfilms Order No. 69-258).

52. W. H. Stevenson, "Optical Frequency Shifting by Means of a Rotating Diffraction Grating," Applied Optics, vol. 9, pp. 649-652 (1970).

53. G. L. Rogers, Noncoherent Optical Processing, (John Wiley & Sons, New York, 1977) Chapter 8.

54. S. Duker, Editor, Time Compressed Speech, (The Scarecrow Press, Metuchen, New Jersey, 1974), vol. 1.

55. R. M. Lerner, "The Representation of Signals," IEEE Trans. on Circuit Theory, Special Supplement, p. 197, May (1959).

56. J. D. Markel and B. Carey, "Digital Voice Processing with a Wave-function Representation of Speech," 1970 Fall Joint Computer Conference, AFIPS Proc., pp. 387-397, Houston, November 1970.

57. T. A. Giordano, H. B. Rothman, and H. Hollien, "Helium Speech Unscramblers - A Critical Review of the State of the Art," IEEE Trans. Audio and Electroacoust., vol. AU-21, pp. 436-444 (1973).

58. R. J. Marks II, J. F. Walkup and M. O. Hagler, "A Sampling Theorem for Space-Variant Systems," J. Opt. Soc. Am., vol. 66, pp. 918-921 (1976).

59. P. Kellman, "Detector Integration Acousto-Optic Signal Processing," Proceedings of the 1978 International Optical Computing Conference, London, S. Horvitz, Editor (IEEE Press, 1979), p. 91.

60. W. T. Rhodes, "Image Processing Using Space-to-Temporal Frequency Conversion," Proc. SPIE, vol. 128, pp. 322-331 (1977).

VITA

James Martin Florence, the son of Milan George and Rachel Stovall Florence, was born in Kansas City, Kansas on December 1, 1949. He was married to Maureen Sue Blount of Dallas, Texas on August 29, 1969.

He graduated from Warren T. White High School, Dallas, Texas in 1968. In 1972, he received a BSEE degree with High Honors from the University of Texas at Arlington. He received the MSEE degree from the Georgia Institute of Technology in 1973, at which time he entered the doctoral program.

From 1972 to 1973, Mr. Florence held a Schlumberger Fellowship at the Georgia Institute of Technology. From 1973 to 1978 he was a Graduate Research Assistant in the School of Electrical Engineering, Georgia Institute of Technology.

He is a member of Tau Beta Pi, Eta Kappa Nu, the Institute of Electrical and Electronic Engineers, and the Optical Society of America.