TIME AND SPACE INTEGRATING ACOUSTO-OPTIC SIGNAL PROCESSING

Thesis by

Kelvin H. Wagner

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1987

Submitted May 14, 1987

Acknowledgements

First and foremost, I must acknowledge the many contributions to this work, and to my own intellectual development, made by my research advisor, Dr. Demetri Psaltis. His penetrating insight into the physics, technology, and relevance of our research efforts has helped to place the problems we are working on into an appropriate scientific context. Without a doubt, knowing Dr. Psaltis has been the highlight of my educational career, and his inspiration and enthusiasm will be greatly missed when I leave Caltech.

The many contributions of my fellow students in Dr. Psaltis' research group have made my graduate studies very enjoyable and productive. I express many thanks to Dr. Michael Haney, for several years of collaborative experimental research efforts into optical SAR processing. The invaluable insights of Robert T. Weverka, and the many enjoyable discussion we had regarding optical systems were a large influence on my research. The many interesting discussions with the rest of the research group made working on optical processing at Caltech both stimulating and informative. I especially want to recognize the numerous valuable contributions to my research made by Dr. Eung Paek, Dr. Gabriel Sirat, Dr. Young Shi, Dr. Santosh Vekatesh, John Hong, Fai Mok, Jeff Yu and Nabeel Riza.

Finally, I must acknowledge the financial and material support provided by JPL and NASA of our research into SAR processing, and specifically I would like to thank Tom Bicknell and Bill Farr at JPL. I also acknowledge the support provided by JPL, NASA and the SETI science working group of our research into folded spectrum processing, and the motivation and interest expressed by Dr. Sam Gulkis and Dr. Jim Lesh of JPL. My research was partially sponsored by a fellowship provided by the Army Research Office, and I would like to thank Dr. R. Gunther at ARO for providing this support.

To Gail and my Parents

Abstract

One dimensional acousto-optic signal processing techniques are examined from the systems and functional viewpoint, and are then used as building blocks to synthesize multidimensional time and space integrating architectures. Time and space integrating signal processing systems are capable of performing 2-dimensional linear transformations upon images or matrices, by sequentially entering rows of the image with a travelling wave acousto-optic Bragg cell. The travelling rows are frozen by a pulsed laser diode, and the stationary diffracted fields are spatially processed by an optical system. The successively transformed rows are sequentially multiplied by a time varying reference wavefront, and accumulated on a time integrating CCD detector array to complete the two dimensional processing. Long 1-dimensional signals can also be linearly transformed by a time and space integrating system, by using a similar strategy upon a folded, or rastered, version of the high time bandwidth product signal. Small pieces of the long signal are slid into the system with an acousto-optic devices, and are spatially transformed over the device aperture. Then, successively transformed portions of the long signal are multiplied by a reference, and appropriately delayed and accumulated on a 2-D CCD in order to perform multichannel time integrations in the orthogonal dimension. The desired high time bandwidth one dimensional linear transformation is represented in the folded coordinate space of the 2-dimensional output detector.

The operational characteristics of the principal active devices used in these time and space integrating systems are examined from the viewpoint of the system architect. The effects of the devices on the overall system operation are discussed, and device designs intended for application in a time and space integrating system operating environment are proposed.

The final chapter is a detailed theoretical and experimental investigation into the particular operating characteristics of systems designed to perform a folded spectrum analysis of very high time bandwidth signals. This spectrum analysis problem has a shift variant transformation kernel, which can be broken down into a succession of smaller temporal and spatial sub transformations. The 1-dimensional space integrating spectrum analysis operation performed by a lens is used to produce a coarse spectral channelization of the input signal, displayed as a one dimensional spatial profile. Each resolvable spectral channel is fine frequency analyzed by temporal integration, producing a resulting intensity variation of each channel in the orthogonal direction, thereby forming a folded representation of the desired high time bandwidth spectrum analysis. The information which is needed to perform the fine frequency analysis is carried on the optical phase, so interferometric techniques are employed in order to detect the phase and transform it to an optical intensity modulation. Various bias terms are produced on the detector by the interferometric detection operation, and techniques for removing the unwanted bias are investigated. These include spatial carrier encoding of the interferometric terms combined with bandpass filtering, and direct bias subtraction techniques.

Table of Contents

List of Figures

Chapter 2

Chapter 3

Tables

Commonly Used Abbreviations

1-D	One Dimensional
2-D	Two Dimensional
A/D	Analog to Digital Converter
AM	Amplitude Modulation
AO	Acousto-Optics
AOD	Acousto-Optic Device or Deflector, Bragg Cell
AOM	Acousto-Optic Modulator
CCD	Charge Coupled Device
CW	Continuous Wave Signal
D/A	Digital to Analog Converter
DFT	Discrete Fourier Transform
DLO	Distributed Local Oscillator
DR	Dynamic Range
FFT	Fast Fourier Transform
FH	Frequency Hopped
FM	Frequency Modulation
FT	Fourier Transform
LD	Laser Diode
LED	Light Emitting Diode
LFM	Linear Frequency Modulation or Chirp
LO	Local Oscillator
MTF	Modulation Transfer Function of an imaging detector array
PD	Photo Detector or Photo Diode
PG	Processing Gain=SNR_{out}/SNR_{in}
PN	Pseudo Noise Code generated by a feedback shift register
PRF	Pulse Repetition Frequency

PRI	Pulse Repetition Interval, PRI=1/PRF
PSF	Point Spread Function = Fourier transform of the MTF
RF	Radio Frequency
SAR	Synthetic Aperture Radar
SB	Space Bandwidth Product
SETI	Search for Extra-Terrestrial Intelligence
SI	Spatial Integration
SLM	Spatial Light Modulator
SNR	Signal to Noise Ratio
TB	Time Bandwidth Product
TDI	Time Delay and Integrate or Shift and Add operation of a CCD
TI	Temporal Integration
TV	Television
TSI	Time and Space Integration

CHAPTER 1

INTRODUCTION

Optical signal processing systems are parallel analog computers which are useful when other signal processing techniques are somehow inappropriate or inadequate. Typical performance advantages of optical systems over the digital alternatives are in size, weight, or power requirements, which are important considerations for real time on board applications. Optical techniques also have very wide bandwidth capabilities, and they have a very large degree of parallelism and interconnectivity, which makes optical techniques appear attractive when compared with other analog signal processing techniques. These advantages can only be realized in certain classes of signal processing transformations which can be realized with available optical devices, and in this thesis I will consider the optical implementation of fixed linear transformations of images and high time bandwidth sequences.

Although digital signal processing techniques are becoming more sophisticated, and parallel hardware implementations are increasing the computational capabilities of these systems, they are still unable to cope with some of the demanding scenarios involving real time high resolution processing of high bandwidth data. For low frequency signal processing tasks, such as speech processing and audio analysis, digitization and real time processing in a computer are readily achieved. Image processing applications involving small kernel transformations are readily performed with pipelined digital processors. However, for wide band radar and communication applications, the digitization operation itself presents difficulties, and processing multi GHz bandwidth signals in real time with digital techniques may stretch the capabilities of conventional silicon digital technology beyond the limits. It is for these reasons that acousto-optic signal processors have been investigated for many years.

The one dimensional acousto-optic signal processing systems discussed in Chapter 2 are at a stage of maturity where they are beginning to be fielded in operational systems as key subsystems. These applications typically involve high instantaneous

bandwidths in multisignal environments, with moderate resolution capabilities. An alternative approach examined in this thesis is to use these one dimensional systems as components of more complex two dimensional systems, where extremely high resolution is the requirement and moderate bandwidths and processing speeds are sufficient. This outlook has become practical because of the simultaneous maturing of three technologies; laser diode optical sources, acousto-optic (AO) travelling wave modulators, and two dimensional charge coupled device (CCD) detector arrays. The small size of laser diodes when compared with gas lasers makes the development of extremely small, powerful two dimensional optical processors technically feasible. Although AO devices can have bandwidths of several GHz, limitations due to the other components limits TSI architectures to bandwidths in the 100s of MHz range. The processing speed is primarily limited by the CCD detector array readout time, which is currently in the 10s of frames per second, but could be speeded up dramatically by parallel readout structures. An essential feature of the processors under consideration here, is that signals flow through the processor dynamically, and it is the spatial as well as the temporal variations of the optical modulations which are used to perform the two dimensional processing operations. When real time two dimensional spatial light modulators (SLMs) reach an equivalent level of development, with resolution and dynamic range characteristics approaching that of film and data throughputs approaching that of AO devices, then more conventional static space integrating optical processors will replace those discussed in this thesis.

The optical signal processing system designer should have a top down, as well as a bottom up perspective, in order to satisfy both the constraints upon the signal processing system in its application environment, as well as obeying the physical limits intrinsic to the device technology which is utilized. In this thesis the application requirements are not considered in any detail, because the systems under investigation are still at the stage of laboratory experiments. It is hoped that this document will provide a running start to development engineers interested in building two dimensional time and space integrating systems for particular systems applications, especially for high resolution spectrum analysis applications. For this reason a number of experimental procedures, details, and difficulties which were encountered, are discussed at a greater length than might be considered necessary for a thesis.

1.1 Thesis Outline

In Chapter 2, I begin with a review of the simple model for the operation of an acousto-optic (AO) Bragg cell as a travelling wave optical modulator. Spatial integration is defined, and the classic examples of image spectrum analysis and matrix vector multiplication are used to illustrate the principles. The most successful application of acousto-optic technology, the AO spectrum analyzer, is reviewed, and its interferometric implementation is presented. The shift invariant operations of correlation and convolution are defined, and several different topologies for one dimensional acousto-optic space integrating correlators are reviewed. Temporal integration is described and time integrating AO correlators are reviewed and contrasted with space integrating correlators, and pulsed source time integrating matrix vector multipliers are presented. The one dimensional time delay and integration (TDI) correlators that make use of the scrolling capability of charge coupled device optical detectors are discussed. Finally, the chirp transform algorithm used in convolutional approaches to spectrum analysis is reviewed, and its characteristics and limitations are discussed. These one dimensional systems provide the necessary building blocks for the two dimensional time and space integrating systems discussed in the remainder of the thesis.

A general framework for combining the one dimensional AO space integrating systems with parallel arrays of one dimensional time integrating processors to create multidimensional time and space integrating (TSI) signal processors is presented in Chapter 3. These hybrid TSI systems can operate on data with high dimensionality, and the two cases of image transformation, and very high time bandwidth signal processing are considered. The set of linear operations which are achievable with TSI systems includes the shift invariant systems, and the separable and cascadable systems, but not the most general space variant transformation. This is not too disconcerting since conventional optical processors can not perform this fully general transformation, and digital computers are hard pressed as well, because of practical limitations on storage and processing time. The problem of synthetic aperture radar (SAR) processing is introduced, and the kernel structure is shown to map well into the processing capabilities of a TSI system. The general performance limitations of the hybrid time and space integrating technique are illustrated for

the SAR processor, and the issues of signal demodulation, and bias build up are addressed, and the advantage of additive interferometers is illustrated.

The active device characteristics are the primary limitations to the TSI technique, and an understanding of device operation can result in more practical system design, which utilizes the available features of existing devices, without expecting unreasonable performance from any of the components. Laser diodes (LDs) are used as the optical source in the experiments presented in this thesis, because of the ease with which they can be directly modulated at fast rates, or pulsed with narrow pulses. The power and coherence limitations of these lasers provide constraints on the processor design and on the optical efficiency which must be obtained from the system. Acousto-optic devices (AODs) are the most mature optical modulator technology currently available, and are used throughout this thesis to sequentially enter real time data into the optical processors. Ideally, these devices are simply travelling wave modulators, but an exploration of the physics of the acousto-optic interaction leads to a deeper understanding of a number of deleterious effects that can degrade system performance if not properly controlled. The charge coupled device (CCD) detector arrays are used to accumulate and store optically generated image profiles during the period of temporal integration and to transduce these optical signals back into electronic format. The ease with which a two dimensional CCD can be operated in a linearly scrolling mode represents another flexibility for the system designer that can simplify the two dimensional processor, and improve its performance by incoherently averaging coherent optical artifacts. It was found that the detector array used in this thesis represented the limit on system resolution and noise performance, but a state of the art device could have characteristics that matched the performance limits of other parts of the system.

The problem of extremely high resolution spectral analysis of high time bandwidth signals was selected as the test bed problem for the system concepts being investigated in this thesis. One dimensional systems are simply not big enough to contain all of the signal information that must be Fourier transformed by the spectrum analyzer, so the signals are entered into the system a small section at a time. Individual sections are spatially Fourier transformed by a lens onto an array of time integrating fine frequency spectrum analyzers. The fine frequency processors pro-

vide an orthogonal spectral analysis of the interferometrically detected sequences produced within each resolvable spot of the spatial transformation. This results in a folded format of the high resolution spectrum on the two dimensional output detector array, even though the entire input signal was never simultaneously present within the window of the input AOD. This is possible because the succession of partial products are properly weighted, delayed, and sequentially accumulated in order to produce the complete transformation. Several techniques are examined which perform this same signal processing task, in order to compare the different techniques, and to select the optimal processor topology for this application. A distributed local oscillator (DLO) interferometer combined with a scrolling TDI CCD chirp transform processor is analyzed theoretically, and found to have a good sidelobe suppression property due to the continuous temporal and spatial integrations. Several pulsed source interferometers are examined theoretically and were demonstrated experimentally. The pulsed source is used to alias a number of high frequency components to baseband for subsequent temporal integration, and the time between pulses is used to slide new reference information into the reference AOD. This technique is used to perform multichannel discrete Fourier transformations (DFTs) upon each resolvable spot at the output of an interferometric space integrating spectrum analyzer. A simpler system that uses a TDI CCD detector array to perform time integrating chirp transforms on each resolvable channel of the space integrating spectrum analyzer was also tested. The bias terms can be removed by recombining the signal and reference beams at an angle in order to introduce an interferometric spatial carrier which can be used to band pass filter the output, and reject the bias. Alternatively, the bias can be eliminated with a bias subtraction technique, and successive frame bias subtraction is demonstrated to test this concept.

The conclusion of the thesis contrasts the different approaches to spectrum analysis that were investigated, discusses the limitations of the TSI technique when applied to spectrum analysis, and makes suggestions for future research directions in TSI processing.

CHAPTER 2

ACOUSTO-OPTIC SIGNAL PROCESSING

Optical signal processing systems consist of sources, modulators, detectors and associated optics for interconnection and communication between elements. The device characteristics and their effects on system performance will be treated in detail in Chapter 4. In this chapter the idealized behavior of an acousto-optic beam deflector is introduced, and the basic 1-dimensional architectures in which they can be used are catalogued.

2.1 Simple Acousto-Optic Device Model

An acousto-optic Bragg cell is a device used to modulate an optical beam in both space and time through its interaction with a travelling acoustic wave[1]. A simple model of the device geometry to be discussed in this section is presented in Figure 2.1.1. The acoustic wave is generated from an RF electronic signal by a piezoelectric transducer bonded to a transparent photoelastic medium in order to launch a bulk acoustic wave into the medium. The acoustic wave is a travelling wave replica of the signal applied to the device, and induces a corresponding travelling wave index perturbation through the photoelastic effect. This results in a phase modulation of an optical field passing through the acoustic beam, which diffracts a travelling wave optical replica of the applied electronic signal. The finite size of the crystaline photoelastic medium limits the extent of this interaction, and the acoustic delay time from the transducer to the termination end of the device is a fundamental device characteristic. The product of the device delay T_a with the modulation bandwidth B gives the number of Rayleigh resolvable independent

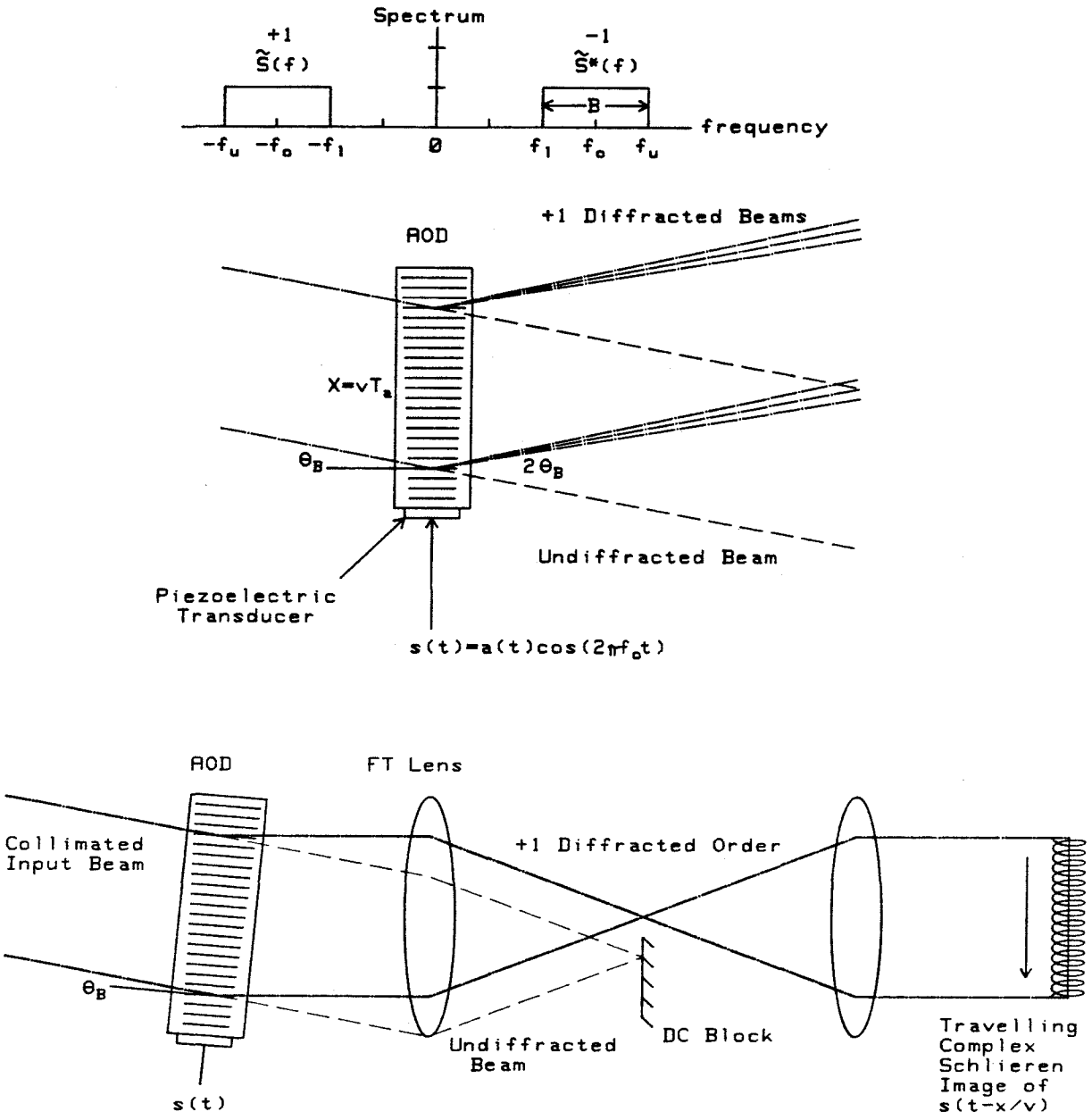


Figure 2.1.1. Simple model of an acousto-optic travelling wave modulator, and Schlieren imaging system to block the undiffracted light.

analog samples that can be represented at one time in the device, and typically this number is approximately $T_a B \approx 10^3$. All AO devices to be discussed in this thesis are operated in the Bragg regime, which requires incidence at the Bragg angle, where a long interaction length between the acoustic and optical waves produces a volume diffraction effect allowing only one sideband of the diffracted beam to be produced. In the remainder of this section I will present a relatively simple analytical model of the behavior of an acousto-optic Bragg cell. In section 4.2 a much more detailed analysis will be presented.

A bandlimited signal $s(t)$ at center frequency f_0 is applied to the Bragg cell. In order to avoid the nonlinear generation of second harmonic intermodulation terms, it is usually required that the signal be less than octave bandwidth, where the upper frequency limit f_u is less than twice the lowest frequency component f_l. Thus the bandwidth of the signal must be less than $B = f_u - f_l \leq 2f_0/3$. The input signal $s(t)$ has a Fourier expansion which is given by

$$s(t) = \int S(f) e^{i2\pi ft} df. \tag{2.1.1}$$

This signal is amplified and applied to the piezoelectric transducer, which has a resonant octave bandwidth centered at f_0. This launches an acoustic wave replica of the portion of the input signal within the transducers bandwidth into the photoelastic medium. The acoustic wave propagates at a velocity v_a across the length of the crystal X, and is represented by a finite aperture travelling wave in the x direction, with the origin at the transducer.

$$\begin{aligned} u(x,t) &= \text{rect}\left[\frac{x - X/2}{X}\right] s(t - x/v) \\ &= \text{rect}\left[\frac{x - X/2}{X}\right] \int s(t') \delta(t - x/v - t') dt' \end{aligned} \tag{2.1.2}$$

The acoustic wave produces a travelling wave volume index perturbation through the photoelastic effect that produces a phase modulation that results in an effective

optical transmittance modulation that may be approximated as

$$t(x,t) = \text{rect}\left[\frac{x - X/2}{X}\right] e^{i\epsilon s(t-x/v_a)} \approx \text{rect}\left[\frac{x - X/2}{X}\right] [1 + i\epsilon s(t - x/v_a) + ...]$$
$$= \text{rect}\left[\frac{x - X/2}{X}\right] \left[1 + \frac{i\epsilon}{2}\tilde{s}(t - x/v_a) + \frac{i\epsilon}{2}\tilde{s}^*(t - x/v_a)\right]. \quad (2.1.3)$$

Where ϵ is the modulation index and relates to the acoustic amplitude applied to the device and the photoelastic interaction efficiency, and in the case of small diffraction efficiencies the higher order expansion of the phase modulation can be neglected. The phase modulation is expressed as upper and lower sidebands, which are complex conjugate pairs resulting from the analytic signal expansion of the real input signal $s(t) = [\tilde{s}(t) + \tilde{s}^*(t)]/2$. The spectrum of the signal is similarly broken up into the corresponding purely negative frequency sideband $\tilde{S}(f) = \int_{-\infty}^{0} s(t)e^{i2\pi ft}df$, and the positive frequency sideband $\tilde{S}^*(f) = \int_{0}^{\infty} s(t)e^{i2\pi ft}df$. When the optical field is incident at the negative Bragg angle on a thick index perturbation so that the x component of its direction of propagation is counterpropagating to the acoustic wave, then the last term will produce negligible diffraction, and the diffracted optical field is said to be upshifted. Alternatively, when the incident optical field is at the positive Bragg angle so that its x component of propagation is copropagating with the acoustic wave, then the first signal term will not produce any diffraction and the diffracted optical field will be downshifted. The purely negative frequency sideband is responsible for the doppler upshifting interaction. This is because the analytic representation chosen for the incident optical field uses the negative temporal frequency as the representation of the optical carrier, so a doppler upshift requires the increase in the magnitude of the negative frequency carrier, by the addition of a negative frequency.

The optical source is assumed to be a continuous wave laser emitting coherent light of wavelength λ with temporal frequency $\nu = c/\lambda$. This light is spatially filtered and collimated producing a gaussian apodized approximation to a plane

wave, which will be modelled as an ideal plane wave for this simple initial model. The collimated light is tilted so that it is incident at the negative midband Bragg angle $\theta = -|\theta_B| = -\sin^{-1}(\lambda/2\Lambda_0) \approx -\lambda/2\Lambda_0$, in order to maximize the bandwidth and coupling efficiency into the upshifted single sideband of the first order diffracted wave. In this expression $\Lambda = v_a/f_0$ is the midband acoustic wavelength corresponding to the midband frequency of f_0. The input optical wave referenced to the direction of acoustic propagation at midband is given by

$$a(x,z,t) = \mathrm{Re}[Ae^{-i2\pi[\nu t - \sin\theta_B x/\lambda + \cos\theta_B z/\lambda]}]. \tag{2.1.4}$$

We will usually drop the explicit notation indicating the real part of the coherent optical field, and it should be noted that the use of the negative temporal sideband representation has sign consequences in other definitions, such as the Fourier transform operation performed by a lens, and the meaning of doppler upshifting and downshifting. The emerging field from the acousto-optic device is given by the product of the incident optical field with a single sideband of the transmittance of the device for a field incident at the Bragg angle.

$$\begin{aligned} b(x,z,t) &= a(x,z,t)t_+(x,t) \\ &\approx A\,\mathrm{rect}\left[\frac{x - X/2}{X}\right] e^{-i2\pi[\nu t - \sin\theta_B x/\lambda + \cos\theta_B z/\lambda]} \left[1 + \frac{i\epsilon}{2}\tilde{s}(t - x/v_a)\right] \end{aligned} \tag{2.1.5}$$

This consists of the undiffracted term which continues to propagate at the input angle $-\theta_B$, and the diffracted modulation term propagating at an angle centered around the angle $+\theta_B$. In reality the undiffracted field must be slightly attenuated by a negative image of the diffracted field so that its intensity profile will be equal to the incident intensity minus the diffracted intensity in order to conserve optical power, however this can effect be ignored for small diffraction efficiencies. The angle of diffraction as a function of frequency is most easily found for the case of a single sinusoidal input $s(t) = \cos(2\pi ft) = [e^{i2\pi ft} + e^{-i2\pi ft}]/2$, in which case the diffracted

field is given by

$$d_f(x,z,t) = A\text{rect}\left[\frac{x-X/2}{X}\right] e^{-i2\pi[\nu t - f_0 x/2v_a + \cos\theta_B z/\lambda]} \frac{i\epsilon}{2} e^{-i2\pi f(t-x/v_a)}$$
$$= A\frac{i\epsilon}{2}\text{rect}\left[\frac{x-X/2}{X}\right] e^{-i2\pi[(\nu+f)t + (f-f_0/2)x/v_a + \cos\theta_B z/\lambda]}. \qquad (2.1.6)$$

Thus for an input at the midband frequency $f = f_0$ the angle of diffraction is seen to be $\theta = +\theta_B = \sin^{-1}(\lambda f_0/2v_a) \approx \lambda f_0/2v_a$. For other input frequencies f the angular spatial frequency of diffraction is seen to be linearly proportional to the input temporal frequency, $\alpha_x = \lambda(f - f_0/2)/v_a$. We will ignore the change in wavelength of the diffracted light, since it is less than 1 part in 10^6. The output field is also seen to be doppler upshifted by the frequency f, which is also only 1 part in 10^6, but the frequency shift is of utmost importance when interferometric detection is utilized. In the case of a broadband input made up of a number of spectral components the diffracted signal is given by the superposition of that due to each component individually. Usually we drop the propagation dimension z from explicit notation in the diffracted field for simplicity.

$$d(x,t) = A\frac{i\epsilon}{2}\text{rect}\left[\frac{x-X/2}{X}\right] e^{-i2\pi[\nu t - f_0 x/2v_a]} \int_{-\infty}^{0} \tilde{S}(f) e^{i2\pi f(t-x/v_a)} df$$
$$= A\frac{i\epsilon}{2}\text{rect}\left[\frac{x-X/2}{X}\right] \int \tilde{S}(f) e^{i2\pi[(f-\nu)t - (f-f_0/2)x/v_a]} df$$
$$= A\frac{i\epsilon}{2}\text{rect}\left[\frac{x-X/2}{X}\right] e^{-i2\pi[\nu t - f_0 x/2v_a]} \tilde{s}(t - x/v_a) \qquad (2.1.7)$$

This is the basic result of the simple model of a Bragg cell. The output field is a travelling wave representation of a single sideband of the octave bandwidth input signal. Temporally it is doppler shifted in an exact replication of the input for as long as the laser remains on. Spatially it is diffracted at an angle proportional to the input, but it is also windowed by the finite aperture of the Bragg cell. The simplest representation is obtained by dropping the multiplicative constants and the explicit optical carrier frequency ν, and the optical axis is aligned with the input beam. The

resulting Bragg cell model is given by the simplified expression for the diffracted field

$$d(x,t) = \text{rect}\left[\frac{x - X/2}{X}\right] \tilde{s}(t - x/v_a) = w(x)\tilde{s}(t - x/v_a) \qquad (2.1.8)$$

This expression will be used in all of the following sections as the simplest approximate representation of the diffracted wave. In most cases however the Gaussian spatial profile of the laser beam, the frequency dependent attenuation of the acoustic wave, and various phase distortions combine to make a hybrid window function that can be much more complicated than the simple rectangular aperture of the photoelastic crystal. This window function is essentially the spatial aperture of the induced polarization field, which is responsible for radiating the diffracted wave.

$$w(x) = \text{rect}\left[\frac{x - X/2}{X}\right] e^{-(x-x_0-X/2)^2/\sigma^2} e^{-\alpha(f)x/2} \phi(x) \qquad (2.1.9)$$

In this expression σ is the $1/e$ width of the input Gaussian optical field, which may be offset by x_0 from the center of the AOD to partially compensate for the acoustic attenuation. The exponential $\alpha(f) = \alpha_0 f^2$ is a common frequency dependent attenuation behavior of the acoustic power, so half of this quantity is used to represent the exponential decay of the diffracted optical field. The term $\phi(x)$ is a phase response due to acoustic diffraction and optical imperfections. Often we will utilize this window function in a coordinate system referenced to the center of the AOD where $x' = x - X/2$.

The diffracted field from a Bragg cell is proportional to a windowed travelling wave representation of the applied voltage signal in the case of small diffraction efficiencies. However in many cases we are interested in producing a travelling wave intensity modulated optical representation of the applied signal. This can be accomplished in several different manners which will be briefly described here, and explained in more detail from the point of view of the coupled mode theory in section 4.2. When the diffraction efficiency is less than 10% the diffracted amplitude

is proportional to the applied voltage, as shown in the voltage in to amplitude out relationship plotted in Figure 2.1.2a. This means that in the small diffraction efficiency regime the diffracted intensity is proportional to the applied electrical power, which is illustrated in Figure 2.1.2b. We can therefore modulate the applied power in order to produce an equivalently modulated intensity. An alternative interpretation is that we can achieve a linear intensity modulation, by amplitude modulating a midband RF carrier with the square root of the desired signal, as illustrated in Figure 2.1.2d. The square root operation compensates for the square law of optical field amplitude to intensity conversion, but it may be quite difficult to perform on wideband signals without introducing spurious components.

Alternatively we can utilize an interferometric technique to produce the desired signal, in which a modulated signal is added to a coherent local oscillator reference so that the diffracted optical fields from the two signals linearly combine and the resulting field amplitude is modulus squared by the optical detection mechanism. If we represent the reference as a constant amplitude sinusoid, $\tilde{r}(t) = r_0 e^{-i2\pi f_r t}$, and the signal as a phase and amplitude modulated carrier $\tilde{s}(t) = a(t)e^{-i[2\pi f t + \varphi(t)]}$, then the interferometrically detected image plane intensity, after blocking the undiffracted beam in the Fourier plane, is given by

$$\begin{aligned} I(x,t) &= |w(x)[\tilde{s}(t - x/v_a) + \tilde{r}(t - x/v_a)]|^2 \qquad (2.1.10) \\ &= w^2(X)\Big[|a(t - x/v_a)|^2 + |r_0|^2 \\ &\quad + 2r_0 a(t - \frac{x}{v_a}) \cos\Big(2\pi[(f - f_r)(t - \frac{x}{v_a}) + \varphi(t - \frac{x}{v_a})]\Big)\Big]. \end{aligned}$$

This has turned the phase and amplitude modulated field into an intensity modulated detected waveform on a temporal carrier of frequency $(f - f_r)$, and on a spatial carrier of spatial frequency $(f - f_r)/v_a$ with a constant bias terms $|r_0|^2$, and a signal dependent bias $|a(t - x/v_a)|^2$. This is most useful when the applied signal is a purely phase modulated signal so that the signal dependent bias term is

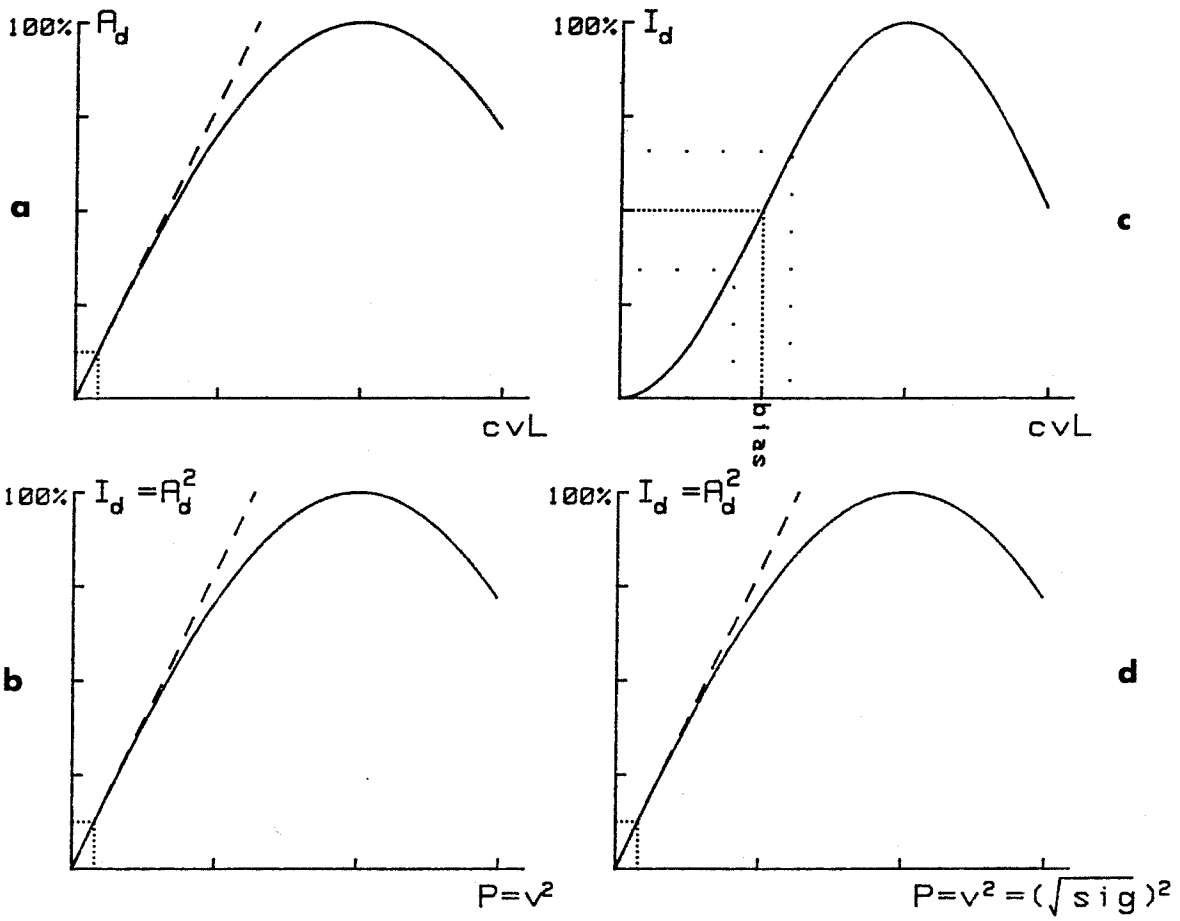


Figure 2.1.2. Input output relations of a Bragg cell optical modulator. a) Diffracted amplitude vs. voltage, b) diffracted intensity vs. power, c) diffracted intensity vs. voltage, and d) diffracted intensity vs. square root of signal.

constant. It is possible to use the undiffracted light as the interferometric reference, but this will result in very high temporal and spatial carrier frequencies, and a poor modulation depth of the heterodyned product term, since the undiffracted beam is much stronger than the diffracted beam in the weak coupling regime. The sum term of the interferometric detection operation, which is due to the implicit presence of the other sidebands that are dropped by the use of the complex notation, is not presented in the interferometric detection equation because it rides on a temporal carrier of $2\nu \approx 10^{15}$Hz, so it is not directly observable, and this term averages to zero over a detector integration time constant.

Another approach to linear intensity modulation utilizes the strong coupling results of the coupled mode theory of the Bragg interaction presented in section 4.2. This analysis shows that the diffracted amplitude builds up linearly until it is strong enough to begin the process of rediffraction back into the undiffracted beam, which is beginning to be depleted. The diffracted amplitude, normalized by the input amplitude, as a function of the applied voltage v, which is plotted in Figure 2.1.2a, is analytically expressed as

$$\frac{A_d(v)}{A_i(0)} = -i\frac{cv}{|cv|}\sin(cvL). \tag{2.1.11}$$

Where L is the interaction length, or the transducer length, and c is the coupling constant per unit applied voltage, and is proportional to the appropriate photoelastic constant, and the piezoelectric coupling efficiency. From this equation we can see that for small diffraction efficiencies the diffracted field is linearly proportional to the applied voltage, and is within .2% of linearity as long as the diffraction efficiency is less than 10%. The resulting intensity diffraction efficiency is the ratio of input intensity to diffracted intensity, and is plotted in Figure 2.1.2c.

$$\frac{I_d(v)}{I_i} = \sin^2|cv|L \approx (|cv|L)^2 \tag{2.1.12}$$

This shows that the diffracted intensity is proportional to the voltage squared, or

acoustic power, for small arguments of the sine function. The plot of this nonlinear intensity diffraction efficiency relationship indicates that another region of linear intensity diffraction versus applied voltage should be achievable by applying an RF bias power to the transducer in order to operate around the linear portion of the $\sin^2(cvL)$ curve[22], which occurs around $cvL = \pi/4$. The problems with this mode of operation are that only small relative intensity variations around this operating point are allowed, so a large bias is present with small modulation depth, and the acoustic power density becomes so large that acoustic nonlinearities begin to produce large nonlinear intermodulation products.

In the following general discussions of acousto-optic signal processing architectures, it is often necessary to utilize one of the intensity diffraction modes briefly presented here. For simplicity I will often assume that it is possible to operate a Bragg cell in a manner in which the diffracted intensity is proportional to the applied signal without going into specifics as to which method that we shall utilize. From the above discussion a generic form of Bragg cell intensity diffraction will be given by

$$I(x,t) = w^2(x)[s(t - x/v_a) + \text{bias}] \tag{2.1.13}$$

Often it will be convenient to ignore the bias term in simple system analysis, but it is important to realize that it is present, and sometimes signal dependent. The presence of the optical bias term will invariably degrade overall system performance, especially detector dynamic range, and special techniques must be employed to remove the optical bias from the final output signal.

2.2 Spatial Integration

Spatial Integration is the collection of light from a region of space onto an output plane location. It can be characterized by an integration over the spatial aperture defining the given region. Often parallel space integrating channels of processing are performed that separate different spatial or spectral frequencies to different output locations. Examples of this type of system would be Fourier analysis of images, acousto-optic spectrum analyzers, and optical spectrum analyzers. The region over which spatial integration is performed can be a 2-dimensional aperture, but in acousto-optic signal processing we employ 1-dimensional spatial integrations using cylindrical optics.

A classic space integrating processor is the two dimensional Fourier transformation performed by a spherical lens[2,3], which is shown in Figure 2.2.1. This powerful operation is the basis of many space integrating optical signal processing systems. A mask $g(x,y)$ is illuminated by a collimated coherent wave of wavelength λ. The mask is placed at the front focal plane of a lens of focal length F, and at the back focal plane the field observed is given by the scaled Fourier transform of $g(x,y)$.

$$G(\hat{x},\hat{y}) = \int_A g(x,y) e^{-i\frac{2\pi}{\lambda F}(x\hat{x}+y\hat{y})} dxdy \qquad (2.2.1)$$

The actual spatial size of the transform depends on the illuminating wavelength λ, and the lens focal length F. Often we will express the transform in terms of the normalized spatial frequency variables $u = \hat{x}/\lambda F$, and $v = \hat{y}/\lambda F$. Each position in the output plane is the integrated product of the transparency with the space variant Fourier kernel corresponding to that output position.

Another classic space integrating processor is the Vander Lugt correlator[4], which uses a Fourier domain multiplication in order to implement convolutions and correlations. The key to this approach is the ability to represent complex image Fourier transforms with an off axis holographic recording process. To make a

matched filter of the image $g(x,y)$, an off axis coherent reference beam is interfered with the spatial Fourier transform $G(\hat{x},\hat{y})$ on a holographic recording medium.

$$I(u,v) = \left|G(u,v) + re^{i\alpha u}\right|^2 \tag{2.2.2}$$
$$= |G(u,v)|^2 + |r|^2 + G(u,v)r^*e^{-i\alpha u} + G^*(u,v)re^{i\alpha u}$$

The final term is the Fourier domain representation of the matched filter for the image $g(x-x_0,y-y_0)$. This hologram is developed and carefully repositioned in the Fourier plane of the optical processor illustrated in Figure 2.2.1, and a test scene containing a shifted and corrupted version of the reference scene is placed at the processor input. The hologram is illuminated with the field $G(u,v)e^{i2\pi(ux_0+vy_0)} + N(u,v)$, and the matched filter term represents the phase conjugate of the desired input, so that a plane wave is diffracted by the hologram, at an angle proportional to the test image displacement. The final Fourier lens retransforms this plane wave back to the correlation domain where a bright spot is produced corresponding to the correlation peak at the appropriately shifted output coordinate.

A simpler type of spatial integration is simply the accumulation of light intensity across an aperture A. There are three primary techniques for performing this simpler operation. A detector matched to the aperture A can be placed directly behind the aperture, and photoelectric current generated across the detector area is summed electronically by the detector current collection mechanism. Alternatively Parseval's theorem can be utilized in order to attempt to decrease the detector size. It states that a power integration in the space domain is equal to a power integration in the spatial frequency domain, or at any intervening Fresnel transform plane along the propagation direction of a lossless optical system. Thus a smaller detector can be placed in the back focal plane of a lens, where it detects the optical intensity at each point in the transform plane and electrically sums the photocurrent across the detector surface. The detector must be larger than the highest spatial frequency

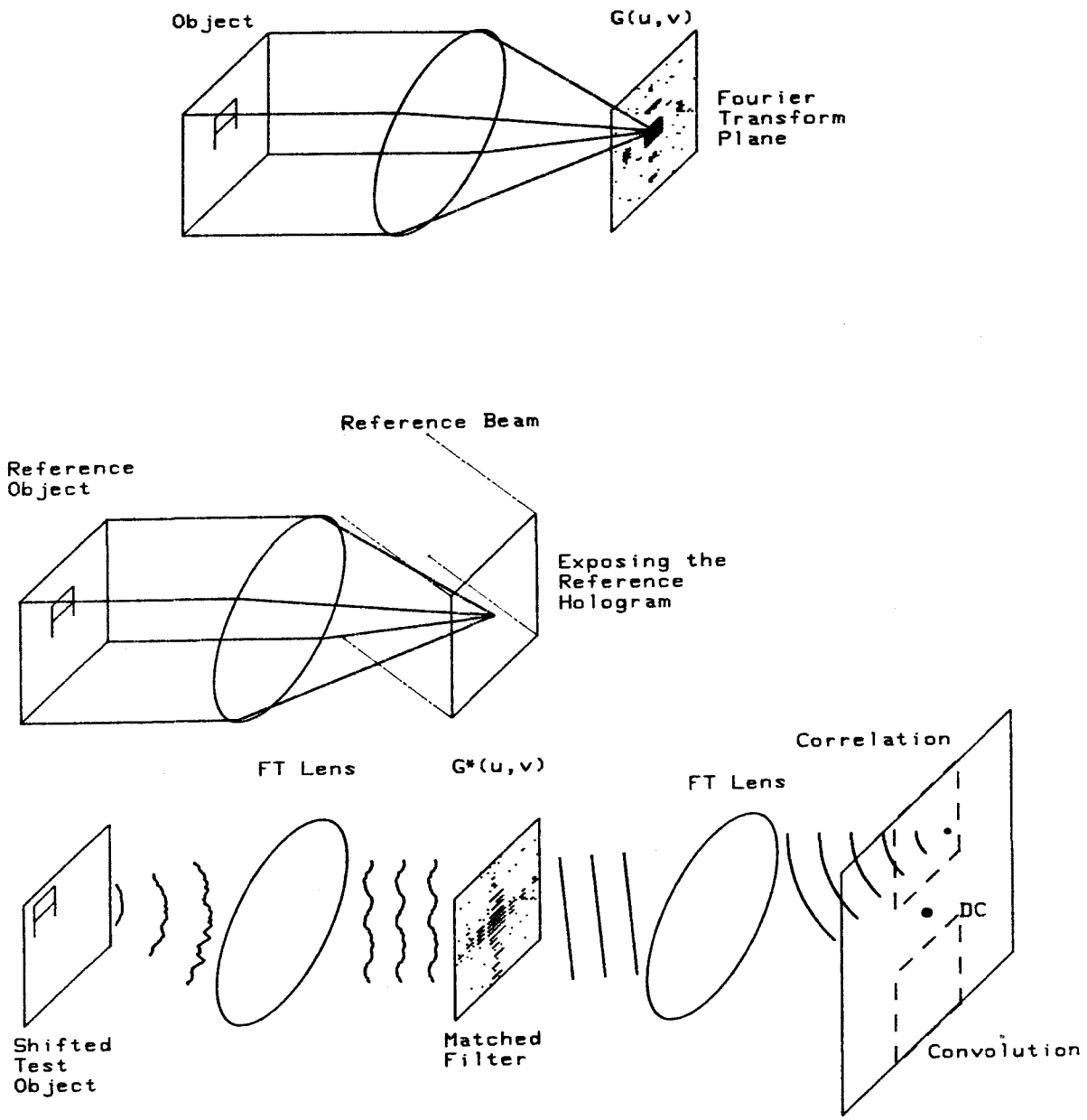


Figure 2.2.1. Image Fourier transformation as performed by a spherical lens, and Vander Lugt holographic matched spatial filter correlator.

component of the image in order to accumulate all of the light in the Fourier domain. Another approach to image integration results from the realization that the Fourier kernel at the position $(0,0)$ is equal to one, so the DC spot in the Fourier plane is proportional to the space integration of an image. The size of the DC spot sampling aperture must be smaller than the scaled image bandwidth, in order to avoid transform roll off within the aperture, which will result in an incorrect estimate of the DC intensity. These three techniques for space integrating light across an aperture can be used interchangeably as appropriate in various systems.

The inner product operation is the basis of almost all optical processing systems, and the space integrating inner product is a powerful parallel signal processing operation. An optical beam that is modulated by an image $f(x,y)$, and is then imaged onto a transparency with transmittance function $g(x,y)$, will multiply pixel by pixel the corresponding image values. By space integrating this image product, the space integrated inner product between the two images can be calculated, $\int\int f(x,y)g(x,y)dxdy$. One dimensional optical modulators such as acousto-optic devices or laser diode arrays can be used to calculate vector inner products, with a one dimensional space integrating optical system. A number of vector inner products can be calculated in parallel in the vector matrix multiplier system shown in Figure 2.2.2[5-8]. The input vector $\vec{v}$ is used to spatially modulate an optical beam in the x dimension, which is imaged in x and smeared in y, broadcasting in parallel to a number of spatially multiplexed rows of a two dimensional matrix mask $T(x,y)$. The light transmitted by the mask is the product $\vec{v}(x)T(x,y)$, and this is space integrated along x and imaged along y to from a spatially multiplexed array of vector inner products.

$$\vec{u}(y) = \int T(x,y)\vec{v}(x)dx \qquad (2.2.3)$$

$$u_m = \sum_n T_{mn}v_n$$

This system can compute an arbitrary linear transformation of a spatially modulated one dimensional optical vector, thereby implementing a space integrating matrix vector product in the time it takes light to propagate the length of the system.

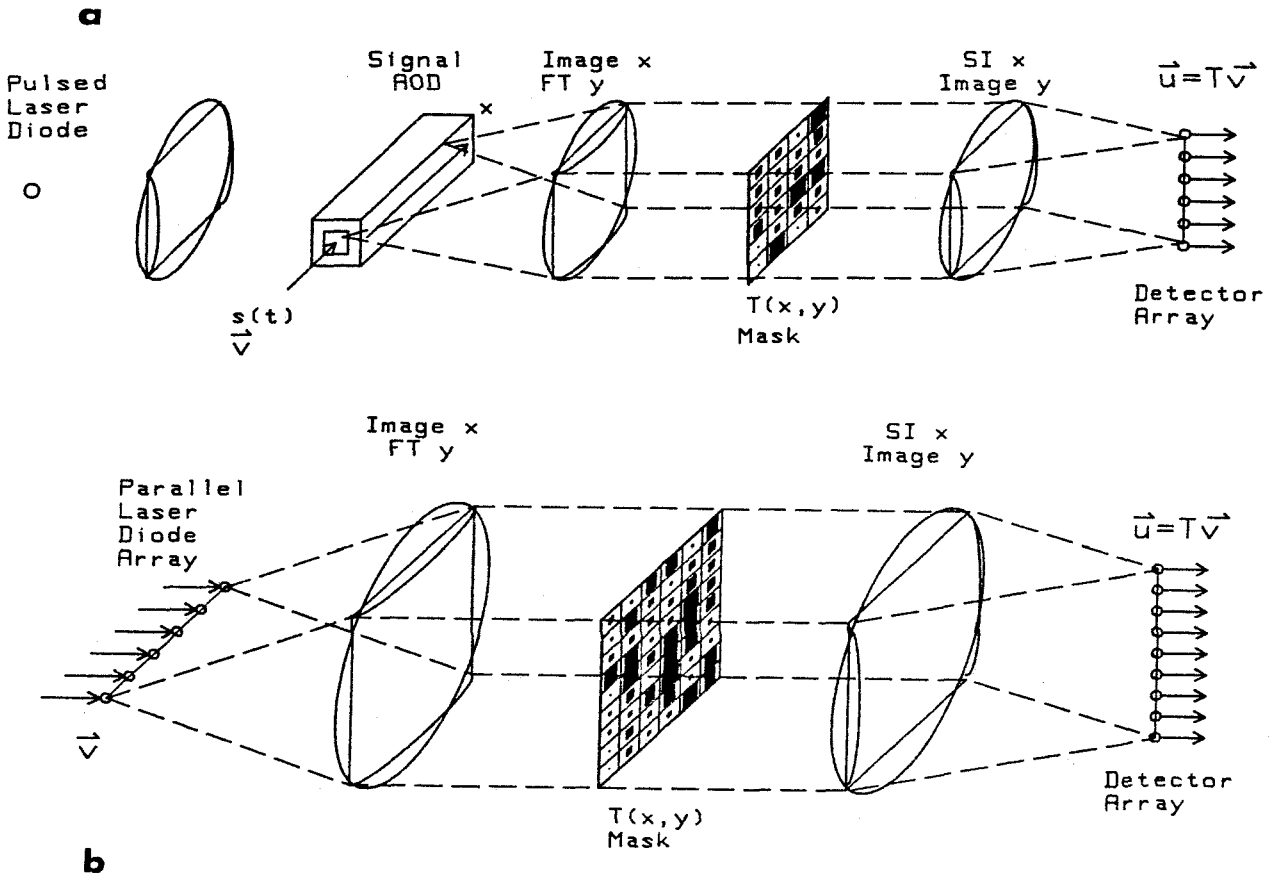


Figure 2.2.2. Space integrating matrix vector multiplier a) using a serial AOD to enter the vector, b) using a parallel laser diode array to enter the vector.

2.2.1 Acousto-optic spectrum analyzer

The space integrating acousto-optic spectrum analyzer is the simplest and most successful application of acousto-optic technology for signal processing[9,10]. It is also a fundamental building block for 2-dimensional signal processing applications as discussed in this thesis, so a thorough review of its operation is essential. These systems are usually operated as power spectrum analyzers, but they can also be operated interferometrically to produce a heterodyned output that allows the measurement of both amplitude and phase of the detected frequency components. The heterodyne receivers can extend the input dynamic range of the system, given a fixed output detector dynamic range[11]. There are two major types of AO spectrum analyzers, the video detector filter banks, and the incoherent radiometers, and the difference is determined by the characteristics of the output detector array. Parallel wideband output detector arrays are typically used with under 100 channels as pulsed tone detectors, when time of arrival is an important parameter to be measured. Integrating photodetector arrays with on the order of 1000 detectors that are serially read out are used when high spectral resolution and large integration times are desired. The most important attributes of the acousto-optic spectrum analyzers are the capability for wide instantaneous bandwidth, the large dynamic range of the AO device, the large number of spectral samples, and the simplicity of this compact, low power system.

The basic acousto-optic spectrum analyzer is shown in Figure 2.2.1.1. It consists of an optical source, collimating lens, the AOD, RF amplifier and electronics, The Fourier transforming lens, and the detector array with its associated electronic readout circuitry. A bandlimited signal $s_1(t)$, centered at frequency f_1, is mixed with a local oscillator with frequency $f_{lo} = f_0 \pm f_1$, producing a signal $s_0(t)$ with a frequency component at the AOD center frequency f_0, and a spurious term centered at the frequency $f_s = 2f_1 \pm f_0$ that can be removed by electronic filtering or by the

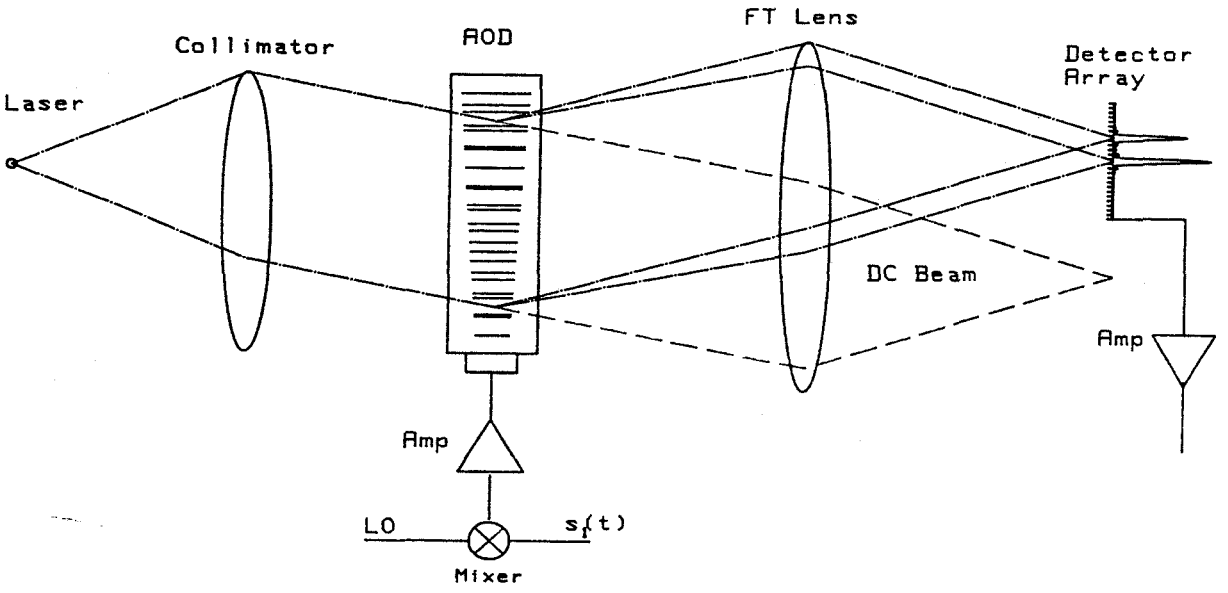


Figure 2.2.1.1. Acousto-optic spectrum analyzer system, illustrating a two tone input.

acousto-optic device frequency response. The octave bandwidth signal $s(t)$ resulting from this passband filtering operation is amplified and applied to the transducer of the AOD.

The emerging field from the acousto-optic device is given by the product of the incident optical field and the transmittance of the device, as expressed in Equation (2.1.5).

$$\begin{aligned} b(x,t) &= a(x,t)t_+(x,t) \\ &= w(x)e^{-i2\pi[\nu t-\sin\theta_B x/\lambda]}\left[1+\tfrac{i}{2}\epsilon\tilde{s}\left(t-\frac{x-X/2}{v_a}\right)\right] \end{aligned} \tag{2.2.1.1}$$

In this expression the origin of the coordinate system has been shifted to the center of the Bragg cell. This field consists of the undiffracted term propagating at the angle $-\theta_B$, and the diffracted modulation term propagating at an angle centered around $+\theta_B$. The optical field propagates through a distance F, and is incident on a Fourier transforming lens of focal length F, so that at a distance F behind the lens the field observed will be the spatial Fourier transform of the field emerging from the aperture of the AOD.

$$\begin{aligned} B(\hat{x},t) &= \int b(x,t)e^{-i2\pi x\hat{x}/\lambda F}dx \\ &= \int w(x)\left[1+\tfrac{i}{2}\epsilon\tilde{s}\left(t-\frac{x-X/2}{v_a}\right)\right]e^{-i2\pi[\nu t-\sin\theta_B x/\lambda]}e^{-i2\pi x\hat{x}/\lambda F}dx \end{aligned} \tag{2.2.1.2}$$

The function $w(x)$ of Equation (2.1.9) is a hybrid window function including effects of AOD surface imperfections, optical apodization, and the finite aperture of the device. The spatial Fourier transform of the window function gives the impulse response of the system, and is the fundamental limit on spectral resolution obtainable with this system.

$$W(u) = \int w(x)e^{-i2\pi ux}dx \tag{2.2.1.3}$$

For an unapodized device the window function is just the rectangular aperture of the crystal, and the number of Rayleigh resolvable (19% dip between adjacent peaks)

spots observable in the Fourier plane is given by the time bandwidth product of the device $T_a B$.

The undiffracted term will result in a bright DC spot $W\left(\frac{\hat{x}}{\lambda F} + \frac{\sin\theta_B}{\lambda}\right)$, which carries no information about the signal and it will be blocked. It is sometimes useful to tilt the Fourier transform lens so that it is aligned with the undiffracted wave, and in the Fourier plane the DC spot will appear at the position defined as $x = 0$. This also removes the angular spatial frequency term $e^{i2\pi \sin\theta_B x/\lambda}$ due to the minus Bragg angle incidence, and simplifies the resulting equations. In reality this introduces unwanted off axis aberrations into the Fourier transform of the diffracted field if the Bragg angle is large, and often the lens is actually aligned with the midband diffracted field component. An expression for the term of interest due to the diffracted wave at the Fourier plane is found by Fourier expanding the input analytic signal, and reversing the order of integration.

$$\begin{aligned} B_1(\hat{x},t) &= \tfrac{i}{2}\epsilon e^{-i2\pi\nu t}\int w(x)\tilde{s}\left(t-\frac{x-X/2}{v_a}\right)e^{-i2\pi x\hat{x}/\lambda F}dx \\ &= \tfrac{i}{2}\epsilon e^{-i2\pi\nu t}\int w(x)\left[\int_{-\infty}^{0} H_{AOD}(f)\tilde{S}(f)e^{i2\pi f(t-\frac{x-X/2}{v_a})}df\right]e^{-i2\pi x\hat{x}/\lambda F}dx \\ &= de^{-i2\pi\nu t}\int \widetilde{H}_{AOD}(f)\tilde{S}(f)e^{i2\pi ft}W\left(\frac{\hat{x}}{\lambda F}+\frac{f}{v_a}\right)df \end{aligned} \qquad (2.2.1.4)$$

The constant $d = \frac{i}{2}\epsilon e^{i\pi f T_a}$ has absorbed all of the constants and phase factors, and the single sideband frequency response of the acousto-optic device is given by $\widetilde{H}_{AOD}(f)$. Thus at the focal plane of the lens the diffracted field from the AOD produces a spatial representation of the Fourier spectrum of the applied signal weighted by the device frequency response, and spatially blurred by the convolution with the resolution limiting window function transform. By utilizing the normalized spatial frequency variable $u = \hat{x}/\lambda F$ we can represent the field at the back focal plane of a Bragg cell spectrum analyzer with the compact notation

$$B_1(u,t) = e^{-i2\pi\nu t}\left[d'\widetilde{H}_{AOD}(uv_a)\tilde{S}(uv_a)e^{i2\pi v_a u t}\right] * W(uv_a) \qquad (2.2.1.5)$$

The modified constant $d' = dv_a$. The term in brackets would be the output for an ideal, infinitely long, perfect Bragg cell spectrum analyzer, and the $*$ represents convolution with the blur function $W(uv_a)$, which accounts for the finite length and apodization of the acousto-optic interaction. Let us examine the case of a single sinusoid of frequency f' input to the spectrum analyzer, so that $s(t) = 2a\cos(2\pi f't) = ae^{-i2\pi f't} + c.c.$, and the temporal Fourier spectrum is given by $S(f) = a\delta(f + f') + a\delta(f - f')$. When this spectrum is substituted into the above equation, the negative sideband is selected by the single sided frequency response of the AOD, and the resulting spatial velocity scaled, delta function spectrum multiplies the next term, known as the distributed local oscillator or DLO. The DLO is a temporal oscillation whose frequency varies linearly in space, across the Fourier plane, or alternatively it is a plane wave, that is pivoted on the DC spot, and whose angular spatial frequency varies linearly with time. The delta function spectrum selects out of the DLO the exact negative temporal frequency of the input signal, but the spatial convolution with the blur function produce a spatial blur spot $d'aW(uv_a - f')$ with a phase term that is uniformly oscillating at the frequency $-f'$. This is considered a doppler upshifting interaction, even though we are using the negative frequency sideband analytic notation, because we are representing the optical field as a negative temporal frequency $e^{-i2\pi\nu t}$, thus the upshifted signal is represented as $e^{-i2\pi(\nu+f')t}$.

Incoherent detection in the Fourier plane produces an instantaneous measurement of the power spectrum of the signal that is currently within the AOD aperture. The width of the intensity blur spot in the Fourier plane, $W^2(u - f'/v_a)$, determines the spectral resolution of the space integrating spectrum analyzer. An integrating detector array, such as a charge coupled device (CCD) or a photodiode array, is used to accumulate instantaneous power spectra for a time T_0. Thus the time integrated output of this type of acousto-optic integrating radiometer is given by the

convolution of the signal spectrum with the AOD blur spot modulus squared.

$$I(u) = \int_{t'}^{t'+T_0} |B_1(u,t)|^2 dt = \int_{T_0} \left| d \int H_{AOD}(f)\tilde{S}(f) e^{i2\pi ft} W\left(u + \frac{f}{v_a}\right) df \right|^2 dt$$
$$= \int_{T_0} |\tilde{S}'(uv_a) * W(v_a u)|^2 dt \qquad (2.2.1.6)$$

Where I have abbreviated the AOD frequency response as a primed spectrum, $\tilde{S}'(f) = H_{AOD}(f)\tilde{S}(f)$. Thus we can see that the space integrating acousto-optic power spectrum analyzer actually utilizes a combination of spatial and temporal integration. In this case the spatial integration is used for spectral channelization and the temporal integration is used to incoherently average a number of instantaneous power spectra in order to improve the detectability of a signal buried in noise[10,12]. In succeeding chapters of this thesis we will see that the temporal integration can alternatively be used for further coherent signal processing, such as correlations, matrix multiplications, or spectral analysis of each spatially integrated output channel.

The interferometric spectrum analyzer is the basis for the folded spectrum architectures presented in Chapter 5. The simplest case of interferometric detection involves the addition of a simple plane wave reference beam to the output of a Bragg cell spectrum analyzer, and is illustrated in Figure 2.2.1.2. The instantaneous output will have a linear spatial dependence of the doppler induced temporal frequency of the detected output intensity that is given by

$$I(u,t) = |B_1(u,t) + r_0 e^{-i2\pi\nu t}|^2 \qquad (2.2.1.7)$$
$$= |\tilde{S}'(uv_a) * W(uv_a)|^2 + |r_0|^2 + 2r_0|\tilde{S}'(uv_a)|\cos[2\pi(v_a ut + \Omega'(uv_a))] * W(uv_a)$$

Where I have expanded the spectrum in terms of its magnitude and phase, $\tilde{S}'(f) = |\tilde{S}'(f)|e^{i2\pi\Omega'(f)}$. The first term is a signal dependent bias term, identical to the instantaneous output of the power spectrum analyzer. The second term is a uniform bias due to the reference beam. The final term reproduces the amplitude and phase

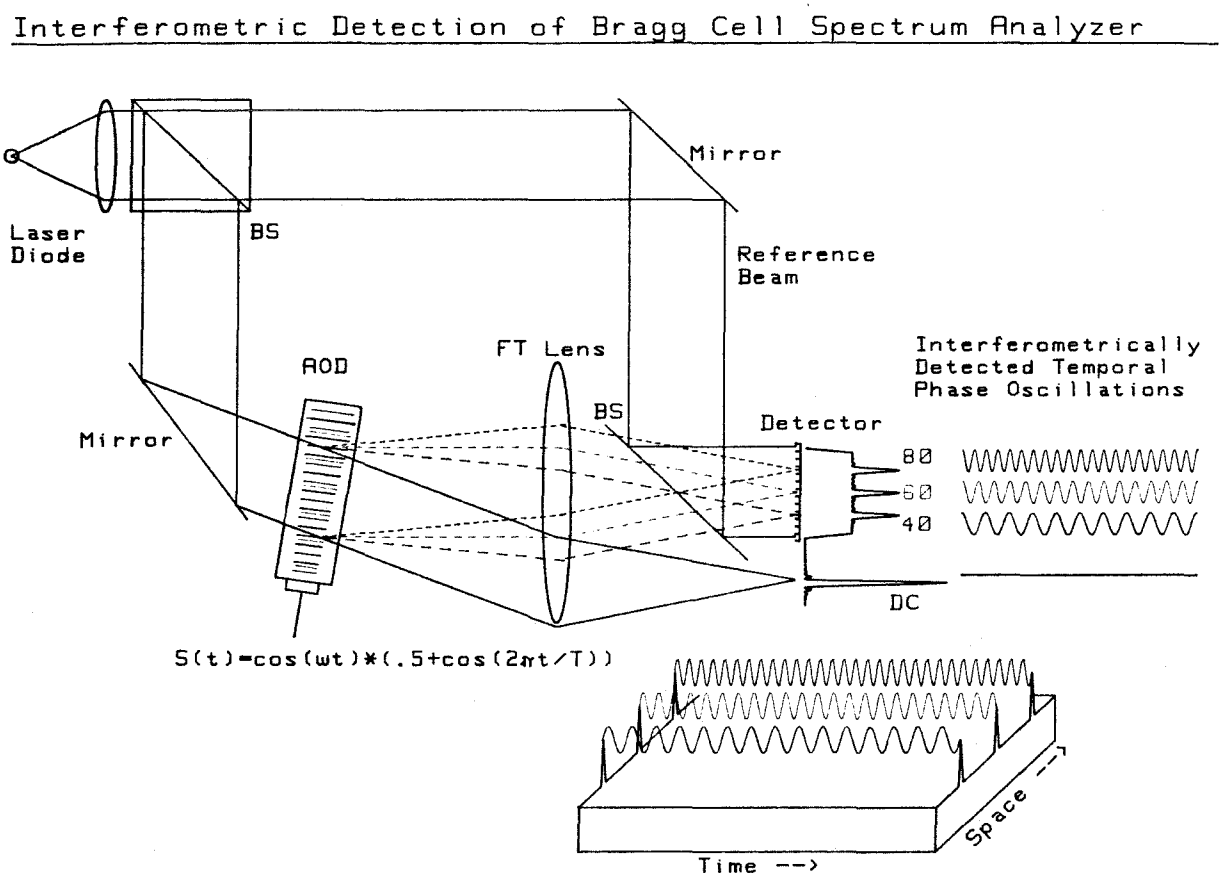


Figure 2.2.1.2. Interferometric detection of the output of a Bragg cell spectrum analyzer for three input tones, showing the temporal variations of the interferometrically detected oscillations.

of the complex weighted input spectrum as temporal modulations of the detected intensity, spatially blurred due to the the input apodization. For the case of a single sinusoidal input tone at a frequency f', with amplitude $|a|$, phase Ω, represented as $s(t) = |a|\cos(2\pi f't + \Omega)$, the heterodyne detected instantaneous intensity output is

$$I_{f'}(u,t) = |aW(u - f/v_a)e^{-i[2\pi f't+\Omega]}e^{-i2\pi\nu t} + r_0 e^{-i2\pi\nu t}|^2 \qquad (2.2.1.8)$$
$$= |a|^2 W^2(u - f'/v_a) + |r_0|^2 + 2|a|r_0 W(u - f'/v_a)\cos(2\pi f't + \Omega)$$

Thus we can see that the temporal modulation of the last term reproduces the input sinusoid in frequency, amplitude, and phase, at the spatial position corresponding to its frequency, and spatially blurred by the resolution limiting apodization function. We will come back to this result several times in the succeeding chapters.

2.2.2 Correlators and Convolvers

The mathematical definition of a one dimensional correlation is given by the equation

$$r_{fh}(\tau) = \int_A f(t)h^*(t+\tau)dt = [f \star h](\tau) \qquad (2.2.2.1)$$

It is seen to involve the point by point multiplication of a function $f(t)$ with a shifted version of the complex conjugate of the function $h(t)$, followed by the integration of all the products over the aperture A to obtain the output function as a function of the shift ordinate τ. Acousto-optic devices have the capability of rapidly shifting and point by point multiplying two functions, and a lens can readily accumulate the light across an aperture, resulting in a space integrating (SI) family of acousto-optic correlators. The correlation integral is a measure of the similarity of shifted versions of the two functions f and h, and it becomes strongly peaked at the shift ordinate τ if $f(t) = h(t+\tau)$. This property makes the correlation integral very useful in measuring the delay of a transmitted waveform $f(t)$, which propagates an unknown distance through a homogenous medium, reflects off an object, and is returned to the transmitter where a delayed version of the transmitted signal is detected. The range delay resolution of a correlator is given by one half the signal propagation velocity divided by the bandwidth of the transmitted waveform $\delta r = v/2B$, where the factor of one half is due to the round trip of the transmitted waveform. Because of the high bandwidths achievable with acousto-optic systems, they have seen wide application in radar and communication systems requiring high bandwidth correlators. In this section we will present a catalog of basic 1-dimensional space integrating correlator architectures, from which in later sections of this thesis we will select the most appropriate building blocks of more complicated multidimensional systems.

A similar operation to the correlation is the convolution, which can be mathe-

matically described as

$$c_{fh}(\tau) = \int_A f(t)h(\tau - t)dt = [f * h](\tau) \qquad (2.2.2.2)$$

The convolution operation involves the point by point multiplication of a signal f(t) with a time reversed and τ delayed signal $h(t)$, followed by the integration of all the products within the aperture A in order to form the output $g = f * h$ at all shift ordinates τ. Because of the similarities between correlations and convolutions, a system that can perform correlations of real signals can also perform convolutions by reversing one of the signals. For this reason we will present the acousto-optic architectures in this section as correlators, but all of them can be used as convolvers as well, by temporally or spatially reversing one of the signals, and in some cases using a different diffracted order from the AODs to represent complex conjugation.

The convolution theorem is a mathematical identity which has a great deal of physical significance in terms of Fourier plane implementation of optical correlators and convolvers. It states that the product of the Fourier transform of two signals is equal to the Fourier transform of their convolution

$$C_{fh}(\omega) = F(\omega)H(\omega) = \int c_{fh}(\tau)e^{-j\omega\tau}d\tau = \int\int_A f(t)h(\tau - t)dte^{-j\omega\tau}d\tau \qquad (2.2.2.3)$$

This means that if we can optically multiply the Fourier transform of two functions and inverse transform the product, then we can perform convolutions. There is a similar Fourier domain relationship for correlations that is given by

$$R_{fh}(\omega) = F(\omega)H^*(\omega) \qquad (2.2.2.4)$$

So we can also perform correlations with multiplications in the Fourier domain.

2.2.2.1 Fixed mask space integrating correlator

The basic space integrating correlator[13–15], illustrated in Figure 2.2.2.1, uses photographic film, or a 1-dimensional spatial light modulator (SLM), to store the

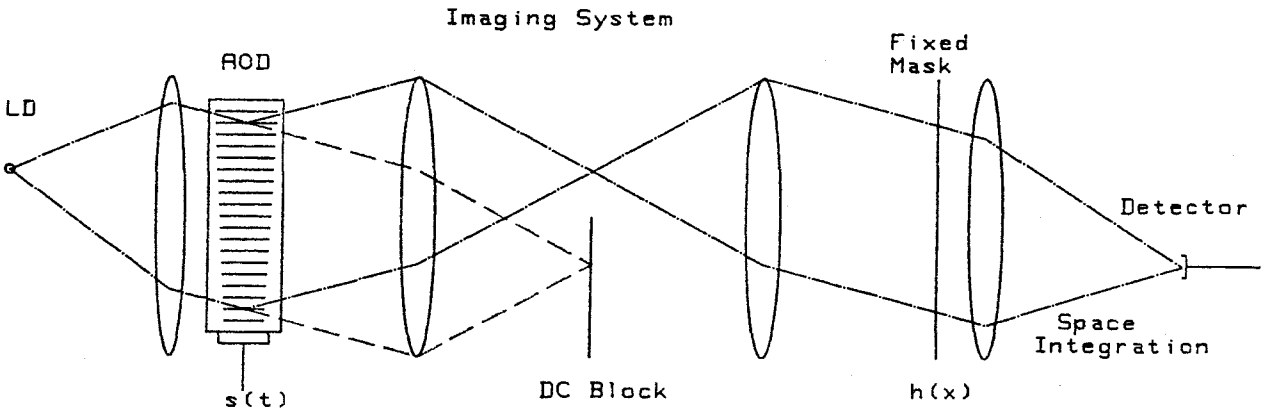


Figure 2.2.2.1. Fixed mask space integrating correlator.

reference function as a spatial representation of the film intensity transmittance $h(x)$. The diffracted intensity $s(t+x/v)$ from the AOD is imaged onto the reference transparency, and the resulting product of the transmittances of the shifting AOD and film needs to be integrated. The integration can be performed with any one of the three methods of spatial integration; image plane integration, Fourier plane integration, or Fourier plane zeroth order moment sampling. In the technique of image plane integration a large photodetector is placed directly behind the film and detects the time varying product intensity distribution at each point across the detector surface. The detector electronically integrates the current generated across its surface of width A, in order to produce a temporal representation of the real correlation integral.

$$I(t) = \int_A s(t + x/v)h(x)dx = r_{sh}(t) \qquad (2.2.2.4)$$

The problem with this technique is that the large area photodetector required will usually have too large a capacitance, resulting in a lower photodetector bandwidth than the bandwidth of the AOD. In order to minimize the size of the photodetector it is necessary to optimally compact the product intensity transmitted through the film. This is accomplished at the focal plane of a lens, where the Fourier transform of the product field is formed when coherent illumination is utilized. Parseval's identity then states that the power integration in the Fourier domain is equal to the power integration in the space domain.

$$I(t) = \int |\int_A [s(t + x/v)h(x)]^{\frac{1}{2}} e^{-j2\pi ux}dx|^2 du = \int_A s(t + x/v)h(x)dx = r_{sh}(t) \qquad (2.2.2.5)$$

This technique can be used with a much smaller photodetector which can usually have sufficient bandwidth. However, if even higher bandwidth is required and sufficient light intensity is available, then the DC component of the Fourier plane can be measured by placing a pinhole or a very small detector at the axis of the optical

system in the focal plane of the lens. If the detector width is smaller than the maximum spatial frequency of the product field $\Delta u < f_{max}\lambda F$, then the detected temporal signal will be proportional to the correlation integral.

$$I(t) = \int_{\frac{-\Delta u}{2}}^{\frac{\Delta u}{2}} |\int_A [s(t+x/v)h(x)]^{\frac{1}{2}} e^{-j2\pi ux} dx|^2 du \propto \int_A s(t+x/v)h(x)dx = r_{sh}(t) \tag{2.2.2.6}$$

Any one of these three space integration techniques can be utilized for several of the succeeding space integrating correlators that utilize a time domain output, but will not be explicitly elucidated for each case.

2.2.2.2 Programmable, counterpropagating space integrating correlator

This architecture uses a time reversed temporal reference function as a programmable input to the correlator[13–15], rather than the fixed mask of the previous system. This system is illustrated in Figure 2.2.2.2, where it appears that the two AODs are propagating in the same direction, but because of the image reversal of the imaging system the two AODs are actually counterpropagating. The space integrated product output is a function of time that is given by

$$I(t) = \int_A f(t+x/v)g(-(t-x/v))dx = \int_{A+vt} f(2t+x'/v)g(x'/v)dx' = r_{fg}(2t) \tag{2.2.2.7}$$

This type of system is seen to time compress the output correlation function by a factor of two.

2.2.2.3 Space integrating correlator with time compressed input

Rather than using a time reversed reference function, we can employ a time compressed reference, and we can use a spatially demagnifying system to match the scale of the spatial signal representations[16]. This is especially useful when a time compressed reference, such as that produced by the previous system, needs to be utilized in a programmable correlator. Such a system is shown in Figure

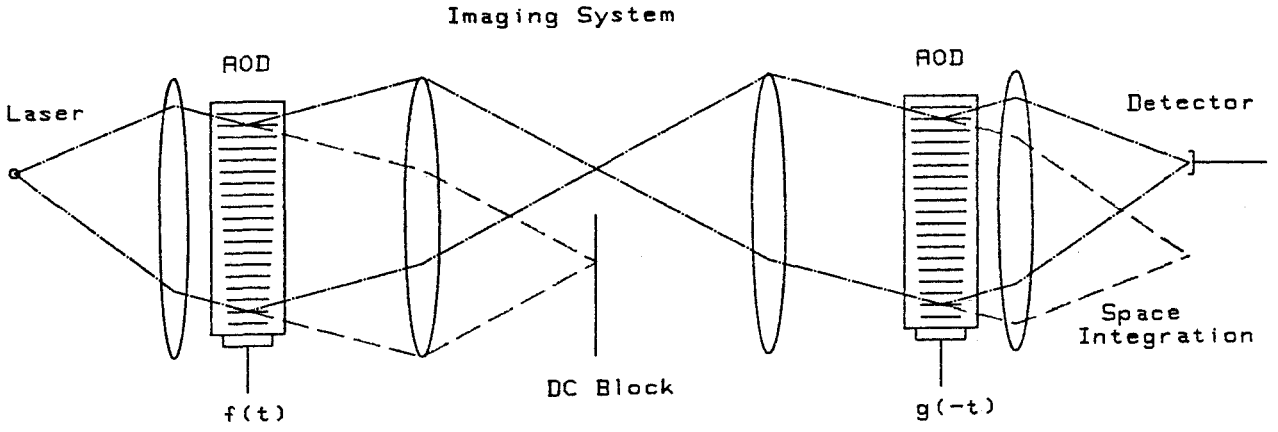


Figure 2.2.2.2. Programmable, counterpropagating space integrating correlator.

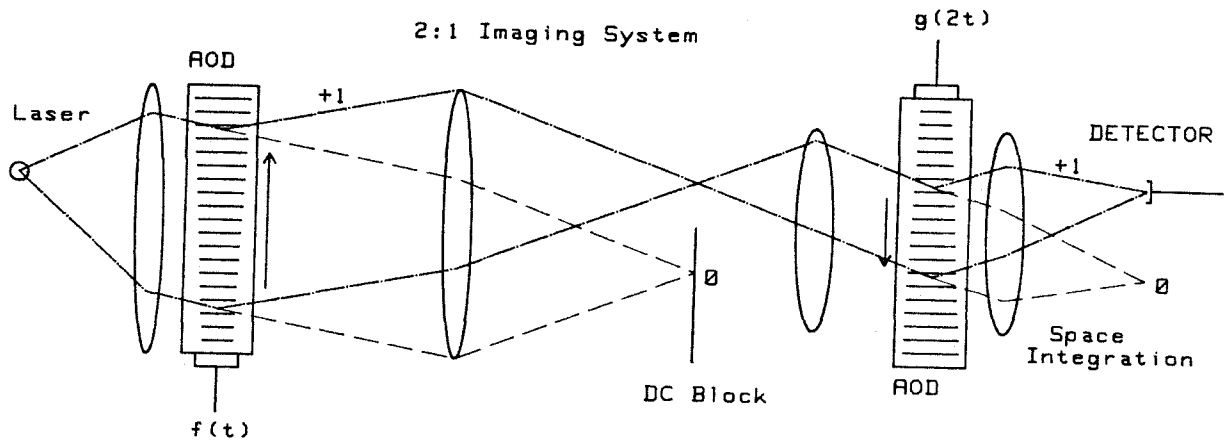


Figure 2.2.2.3. Space integrating correlator with time compressed input.

2.2.2.3, where the first AOD is demagnified by a factor of two onto the second AOD, consequently its spatial velocity is half the acoustic velocity. The signal applied to the second AOD is time compressed by a factor of two, and its acoustic wave is copropagating with the demagnified image of the first AOD, so the two spatial representations have the same scale factor at this plane. The acoustic wave in the second AOD propagates at twice the apparent velocity of the image of the first AOD, producing the relative shift needed in the correlation integral. The system output is not time compressed, since the relative motion between the two copropagating different velocity signals in the coordinate system fixed with the demagnified signal is not counterpropagating.

$$I(t) = \int_A f(t - 2x/v)g(2(t - x/v))dx = \int_{A-vt/2} f(2x'/v)g(t + 2x'/v)dx' = r_{fg}(t) \tag{2.2.2.8}$$

This type of technique can also be utilized when AODs with different acoustic velocities need to be utilized to implement a programmable correlator.

2.2.2.4 Holographic filter space integrating correlator

The classic Vander Lugt type of correlator[4] can be implemented in 1-dimension with travelling wave optical modulators by using a pulsed laser source to freeze the acoustic motion, and this type of system is illustrated in Figure 2.2.2.4a. In this type of system we must form a holographic reference of the Fourier transform of the signal to be correlated. This is accomplished by interfering the Fourier transform of a transparency $g(x)$ with an off axis reference plane wave, incident at an angle $\theta = \sin^{-1} k_x\lambda$, and recording the interference pattern on a piece of film. The resulting transmittance function of the developed film is given by the expression

$$\begin{aligned} T(u) &= |\int g(x)e^{-i2\pi ux}dx + r_0 e^{i2\pi k_x x}|^2 \\ &= |G(u)|^2 + |r_0|^2 + G(u)r_0^* e^{-i2\pi k_x x} + G^*(u)r_0 e^{i2\pi k_x x} \end{aligned} \tag{2.2.2.9}$$

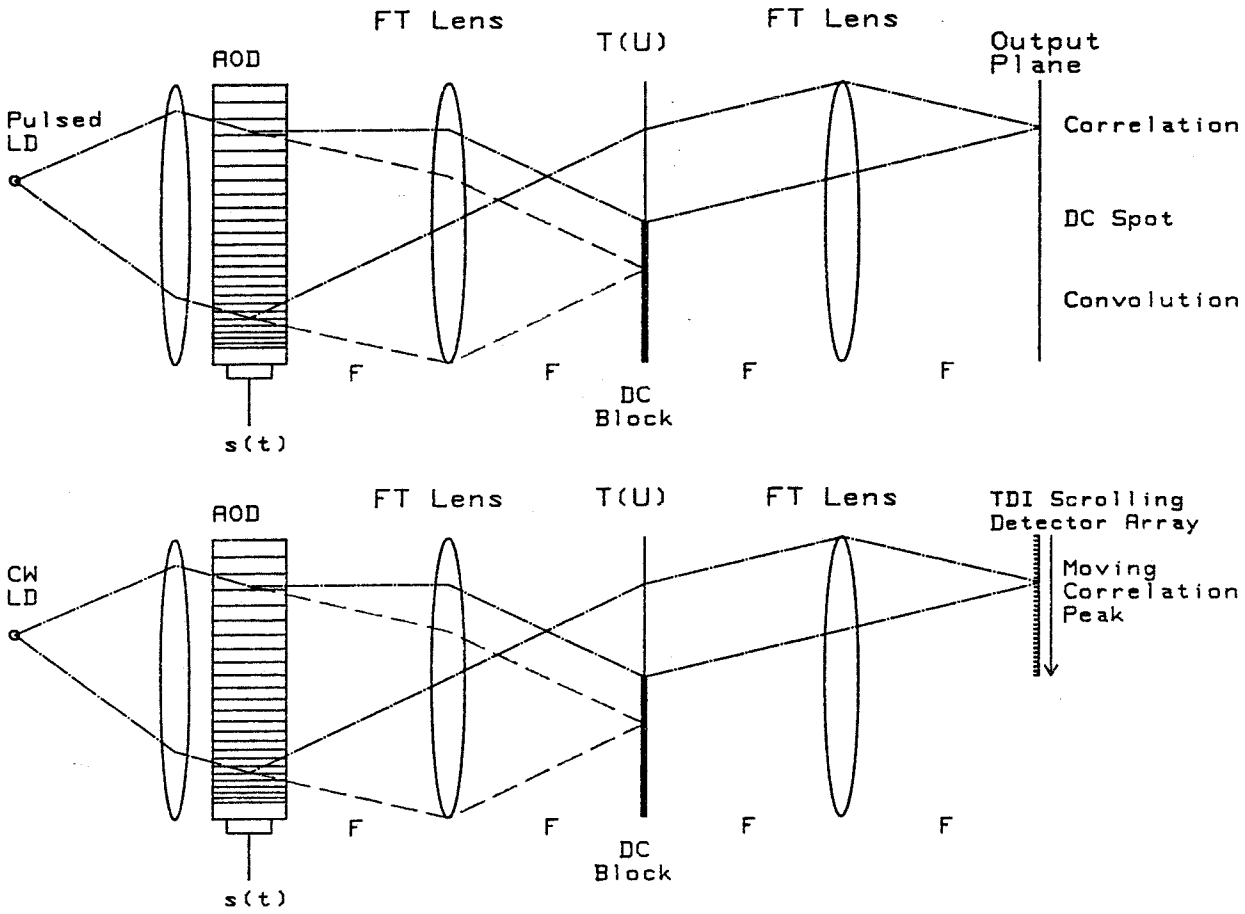


Figure 2.2.2.4. Vander Lugt holographic filter space integrating correlator.

This consists of two low spatial frequency components, and the reference signal transform on a spatial carrier $-k_x$, and the conjugate of the reference signal transform on a spatial carrier k_x. The film is developed and realigned in the Fourier plane of an AOD, that is illuminated by a pulsed laser. The signal applied to the AOD transducer is $s(t)$, and the diffracted signal is given by the product of the incident illumination with the travelling wave single sideband amplitude modulation of the AOD.

$$a(x,t) = p(t)w(x)\tilde{s}(t+x/v) = \delta(t)w(x)\tilde{s}(t+x/v) = w(x)\tilde{s}(x/v) \qquad (2.2.2.10)$$

The laser pulse $p(t)$ has been approximated by a delta function which has the effect of freezing the acoustic motion. A spatial representation of the input signal is then produced. As long as the laser pulse is narrow enough so that the highest spatial frequency of $s(x/v)$ moves by much less than a single fringe in the AOD, then this delta function approximation is valid. This results in the condition that the laser pulse width $\tau < 2/B$, where B is the bandwidth of the AOD. The temporally sampled light that is diffracted by the AOD is Fourier transformed and is incident upon the reference hologram. The transmitted field after the hologram is given by the product

$$\begin{aligned} b(u) &= T(u)\int a(x,t)e^{-i2\pi ux}dx = T(u)[d'S(uv) * W(uv)] \qquad (2.2.2.11) \\ &= [|G(u)|^2 + |r_0|^2 + G(u)r_0^* e^{-i2\pi k_x x} + G^*(u)r_0 e^{i2\pi k_x x}]vS(uv) * W(uv) \end{aligned}$$

This field is Fourier transformed by the lens following the hologram, producing both the correlation and convolution of the reference signal with the windowed input signal, at different locations in the output plane.

$$\begin{aligned} b(x') &= [|g(x)|^2 + |r_0|^2] * [w(x)s(x/v)] + r_0^* g(x) * [w(x)s(x/v)] * \delta(x' - k_x\lambda F) \\ &\quad + r_0 g(x) \star [w(x)s(x/v)] * \delta(x' + k_x\lambda F) \qquad (2.2.2.12) \end{aligned}$$

The first term represents an on axis spot bearing no useful information. The second term is centered at the position $x' = k_x\lambda F$, and is a spatial representation of the

convolution $c_{gs}(x'-k_x\lambda F)$. The last term is located at the position $x' = -k_x\lambda F$, and is a spatial representation of the correlation function $r_{gs}(x' + k_x\lambda F)$. By adjusting the angle of the reference beam to a large enough angle, we can guarantee that none of these terms overlap, and we can center an output detector array at the spatial location corresponding to the desired operation of correlation or convolution.

This type of correlator has a severe optical inefficiency problem because of the short amount of time during which the optical source is on, and the multiplicative diffraction from both the AOD and the hologram throws away most of the available light. The source must be pulsed in order to freeze the motion of the correlation peak as it travels across the output plane with a velocity equal to the optically demagnified acoustic velocity. An alternative is to use a detector array that operates in a scrolling mode that travels in synchronism with the travelling correlation peak[17], as shown in Figure 2.2.2.4b. In this case the laser can be turned on for as long as the detector integration coordinate frame remains in synchronism with the travelling correlation peak. Integrating in a moving coordinate system is a type of time integration that will be explored further in section 2.3. In this application it helps to increase the optical throughput of the system, and it averages out stationary coherent artifacts that may be present, thereby improving the fidelity of the detected correlation.

2.2.2.5 Joint transform filter space integrating correlator

The joint transform correlator is similar to the holographic correlator, except it is programmable in real time through the use of an optically addressable spatial light modulator (SLM), or a nonlinear optical medium, which is operated as a Fourier domain multiplier. A simplified schematic of an acousto-optic joint transform correlator is shown in Figure 2.2.2.5. Once again a pulsed laser source is utilized, but in this architecture it is split into two beams which illuminate two Bragg cells at the

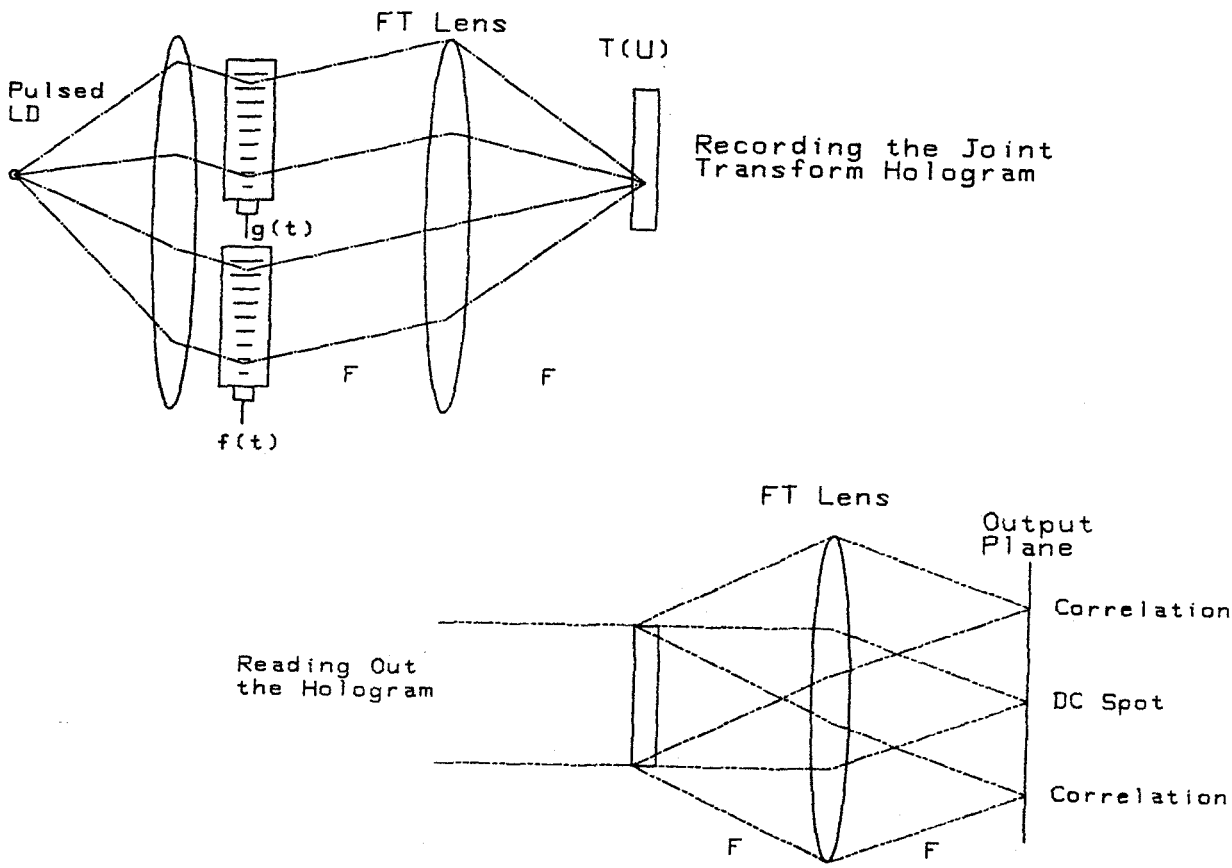


Figure 2.2.2.5. Joint transform space integrating correlator, showing hologram exposure, and readout operations.

same Bragg angle, shifted off-axis by an amount x_0, generating two upshifted (or two downshifted) diffracted beams. These two spatial representations of the input signals are Fourier transformed by the same large lens, so that at the Fourier plane of the lens the field amplitude is

$$\begin{aligned} a(u,t) &= p(t)[d'e^{i2\pi vut}F(vu)]e^{-i2\pi[\nu t - x_0 u]} \\ &\quad + p(t)[d'e^{i2\pi vut}G(vu)]e^{-i2\pi[\nu t + x_0 u]} \\ &= d'F(vu)e^{i2\pi x_0 u} + d'G(vu)e^{-i2\pi x_0 u} \end{aligned} \tag{2.2.2.13}$$

The pulsed source, $p(t) = \delta(t)$, samples the distributed local oscillator phase term at time t=0, thereby removing the temporal oscillations. The time duration of the pulse must be short enough so that locally channelized distributed local oscillator terms produce stationary fringes, and this requires that the pulse width is less than the AOD aperture $\tau < A/2$. At the Fourier plane we insert a thin nonlinear optical material in which an absorption or index modulation can be produced that is proportional to the incident intensity.

$$\begin{aligned} t(u) \propto I(u) &= |a(u,0)|^2 = |d'F(vu)|^2 + |d'G(vu)|^2 \\ &\quad + [d'F(vu)][d'G(vu)]^* e^{i2\pi 2x_0 u} \\ &\quad + [d'F(vu)]^*[d'G(vu)]e^{-i2\pi 2x_0 u} \end{aligned} \tag{2.2.2.14}$$

The transmittance of the recording medium is assumed to be proportional to the detected intensity. By illuminating the recorded interference pattern with another source, with an appropriate wavelength and illumination angle, we can read out the product terms, and Fourier transform them with a lens.

$$\begin{aligned} A(x) &= \int t(u)e^{-i2\pi xu}du = f(x) \star f(x) + g(x) \star g(x) \\ &\quad + f(x - 2x_0) \star g^*(x) + f^*(x) \star g(x + 2x_0) \end{aligned} \tag{2.2.2.15}$$

This system produces the autocorrelations of the two inputs as on axis terms, and the cross correlation appears off axis at the locations $\pm 2x_0$. The pulsed source used

to freeze the acoustic motion in the AODs, also must expose the optical recording medium, which places a high peak optical power requirement on the laser, but not as high as the Vander Lugt correlator of the previous section, because of the longer allowed pulse width. It is also possible implement this type of system with a single AOD, and temporally multiplex the two applied signals, so that they are not overlapping in the Bragg cell.

2.2.2.6 Space integrating chirp correlators

In radar signal processing linear frequency modulations (LFM), or chirps, are often used for ranging applications because of the ease with which high time bandwidth chirps can be generated, and there good correlation and cross correlation properties. The propagation of an optical field through free space can be described as a linear system with a quadratic phase impulse response. This can be utilized in several different fashions to produce compact chirp pulse compression systems[15,18–21], and one version that produces a spatial representation of the output chirp correlation is illustrated in Figure 2.2.2.6. A pulsed laser source is used to illuminate an AOD to which a linear FM signal with unknown delay has been applied. The delay t_0 must be within the time delay aperture of the AOD. The applied signal is given by

$$s(t) = a\text{rect}\left[\frac{t-t_0}{T}\right] e^{-i2\pi[f_c(t-t_0)+\frac{b}{2}(t-t_0)^2]} \qquad (2.2.2.16)$$

The instantaneous frequency of the applied linear FM (or chirp) signal is given by the derivative of the phase factor, and is seen to be $f_c + b(t - t_0)$, which is characterized by the chirp rate b, usually measured in MHz/μsec (10^{-12}sec^{-2}), and the center frequency f_c. The total bandwidth of the linear FM is given by the chirp rate b times the duration T, so the time bandwidth product of the chirp waveform is $TB = bT^2$. The diffracted light from the AOD is given by the usual travelling wave modulation expression, illuminated by a laser pulsed at time $t = 0$ with a

pulse of width less than the chirp inverse bandwidth $\tau < 1/B$ so that the focussed chirp will not appreciably smear during the laser pulse.

$$\begin{aligned} a(x,t) &= p(t)w(x)\tilde{s}(t+x/v) \\ &= \delta(t)w(x)a\text{rect}\left[\frac{t-t_0-x/v}{T}\right]e^{-i2\pi[f_c(t-t_0-x/v)+\frac{b}{2}(t-t_0-x/v)^2]} \\ a(x,0) &= w(x)a\text{rect}\left[\frac{-t_0-x/v}{T}\right]e^{i2\pi[f_c(t_0+x/v)+\frac{b}{2}(t_0+x/v)^2]} \end{aligned} \tag{2.2.2.17}$$

The temporally frozen diffracted signal is allowed to propagate through a distance z of free space, which results in a linear transformation of the diffracted signal that can be expressed as

$$\begin{aligned} b(x') &= \frac{i}{\sqrt{\lambda z}}\int a(x,t)e^{i\frac{\pi}{\lambda z}(x'-x)^2}dx \\ &= \frac{i}{\sqrt{\lambda z}}\int w(x)a\text{rect}\left[\frac{-t_0-x/v}{T}\right]e^{i2\pi[f_c(t_0+x/v)+\frac{b}{2}(t_0+x/v)^2]}e^{i\frac{\pi}{\lambda z}(x'^2-2xx'+x^2)}dx \end{aligned} \tag{2.2.2.18}$$

When the condition $-b/v^2 = 1/\lambda z_0$ is satisfied, then the spatial quadratic phase terms in x will cancel and the diffracted wave comes to a sharp focus at that plane. For an upshifting acousto-optic interaction as illustrated here we would need to use a down chirp with a negative b in order to obtain a real focussed spot in front of the AOD with positive z_0. If we use an up chirp then the diffracted light will appear to be emanating from a point a distance z_0 behind the AOD. We can use a down shifting acousto-optic interaction to obtain a real focus of an upchirp in front of the AOD, or we can image the virtual source behind the AOD with an auxiliary lens. This is equivalent to having performed an autocorrelation of the applied chirp signal, which is displayed as a function of the output spatial variable. The detected intensity at the plane $z = z_0 = -v^2/b\lambda$ is given by

$$\begin{aligned} I(x') = |b(x')|^2 &= \frac{1}{\lambda z_0}|ae^{i\frac{\pi x'^2}{\lambda z_0}}\int w(x)\text{rect}\left[\tfrac{x+vt_0}{-vT}\right]e^{i2\pi[f_c(t_0+\frac{x}{v})+\frac{b}{2}(t_0^2+2t_0\frac{x}{v})]}e^{-i\frac{2\pi}{\lambda z_0}xx'}dx|^2 \\ &= \frac{1}{\lambda z_0}|a|^2|W(x'/\lambda z_0) * (-vT)\text{sinc}[-vT(x'/\lambda z_0 - f_c/v - bt_0/v)]|^2 \\ &= BT|a|^2|W(x'/\lambda z_0) * \text{sinc}[B(x'/v - f_c/b - t_0)]|^2 \end{aligned} \tag{2.2.2.19}$$

Thus the output plane contains a blurred sinc of width given by the velocity scaled inverse chirp bandwidth $\Delta x = v/B$, at a position that is proportional to its center frequency divided by the chirp rate, plus its unknown delay. At this position all of the energy of the chirp has been concentrated to its focal spot, and a large peak with an intensity that is the chirp time bandwidth product TB times as large as the average diffracted intensity just after the AOD.

When the laser is operated in a continuous mode, then the chirp focus will scan across the focal plane at the acoustic velocity. Placing a tiny high bandwidth output detector in the focal plane will produce a time domain output of the chirp correlation. Alternatively, a synchronously scanning CCD detector array could be used to accumulate energy from the focussed chirp as it scans across the output plane. This will help to average out coherent artifacts and spatial noise, and will lead to the same system resolution as long as the scrolling CCD and the travelling chirp focus remain in perfect synchronism.

A travelling wave chirp lens as described here can also be used as an acousto-optic scanner by modulating the input laser source as the chirp focus scans across an output detector array. In this case the time integrated charge profile detected in the chirp focus plane is given by the convolution of the laser diode temporal modulation $f(t)$ with the chirp focus blur spot, which is valid as long as the chirp is fully within the AOD.

$$
\begin{aligned}
I_s(x') &= \int f(t)\frac{1}{\lambda z_0}|ae^{i\frac{\pi}{\lambda z_0}x'^2}\int w(x)\mathrm{rect}\left[\frac{v(t-t_0)-x}{vT}\right] \\
&\qquad e^{-i2\pi[f_c(t-t_0-x/v)+\frac{b}{2}((t-t_0)^2-2(t-t_0)x/v)]}e^{-i\frac{2\pi}{\lambda z_0}xx'}dx|^2dt \\
&= \int f(t)BT|a|^2|W(x'/\lambda z_0) * \mathrm{sinc}[B(x'/v - f_c/b + (t-t_0))]|^2dt \\
&\approx f(x'/v - f_c/b - t_0)BT|a|^2 * |W(x'/\lambda z_0) * \mathrm{sinc}[Bx'/v]|^2 \qquad (2.2.2.20)
\end{aligned}
$$

This result shows that to within the resolution limit imposed by the chirp focus, the integrated intensity profile at the chirp focus output plane is a spatially scanned image of the temporal laser modulation function.

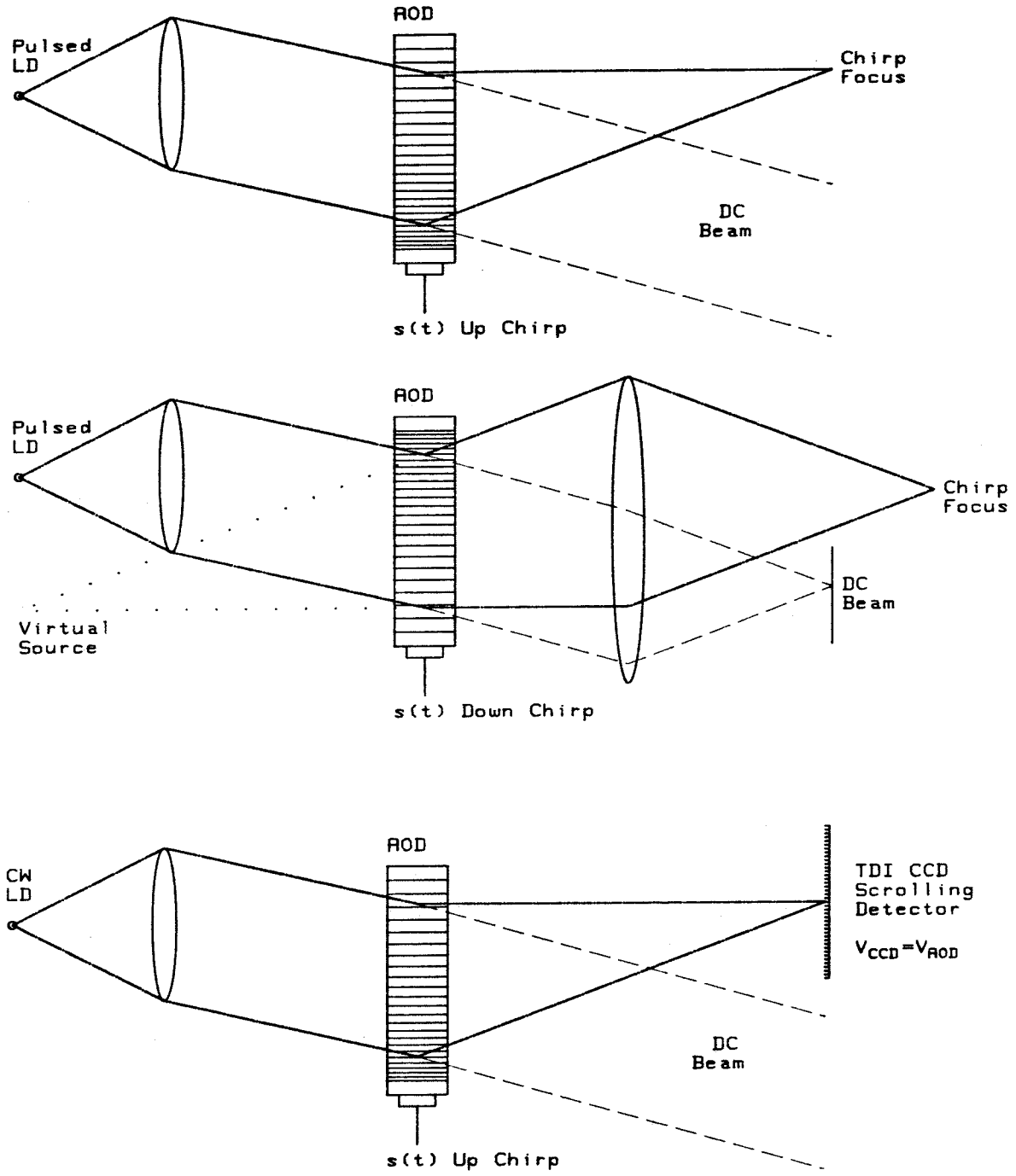


Figure 2.2.2.6. Free space chirp correlator for an up chirp, imaging of the virtual focus for a down chirp, and detecting in a scrolling coordinate frame with CW illumination.

2.3 Temporal Integration

Temporal integration[22–29] (TI) in an acousto-optic signal processor is accomplished through the sequential accumulation of photogenerated charge in an integrating detector array. It can be analytically represented as a finite time integration, or for temporally sampled data, as a summation over the sequential data samples. In the systems we will be discussing, parallel spatial channels of temporally integrated data are simultaneously formed on a detector array. These can represent different shift ordinates of correlation integrals, just as they did for the space integrating systems of the last section.

Typically charge coupled devices (CCDs) are used for temporal integration, since they have a high quantum efficiency as optical detectors, low read out noise, and an inherently integrating potential well structure. Alternatively, photodiode arrays can be utilized as integrating detectors by including an external integration mechanism, such as an external capacitor or CCD, or integrating in the junction capacitance itself. In a sense any optical detection operation involves an implicit temporal integration over a time of the order of the inverse detector bandwidth, but this is not the perspective which will be considered here.

There are three primary types of temporal integration; data formatting, noncoherent averaging for processing gain, and signal processing. Temporal integration for data formatting is employed in the space integrating architectures utilizing pulsed sources, where an implicit time integration turns the product of the δ-function source with the travelling wave modulator into a purely spatial modulation, as well as in the chirp scanner of the previous section. Temporal integration in a moving coordinate frame was used to detect a moving correlation peak produced by the holographic space integrating AO correlator in section 2.2.2.4 in order to improve the light efficiency and average out coherent noise. Temporal integration for noncoherent processing gain was utilized in section 2.2, in order to improve the

signal to noise ratio of an acousto optic power spectrum analyzer[12]. In this section we will present optical architectures for linear signal processing tasks that use temporal integration for the explicit integration variable of linear transformations.

Temporal integration is invariably accompanied by an undesirable bias buildup on the detector array, due not only to thermally generated dark current accumulated in the integrating potential well, but also due to the optical intensity bias that must be employed in order to represent bipolar signals with inherently unipolar optical intensity modulation. This is the major drawback of time integration as compared to coherent spatial integration, and appropriate bias removal techniques must be employed. On the other hand time integration has the advantage that the integration interval can be very long, while for space integrating techniques the integration time is limited to the acoustic delay time of the AO medium. The limit on integration time for TI techniques is given by the rate of bias buildup, so that the detector dynamic range is not saturated by the bias. If still longer integration times are desired then bias removal can be followed by further digital integration, giving an effectively infinite integration capability.

2.3.1 Time integrating AO correlators

The basic TI correlator is shown in Figure 2.3.1, and it consists of a modulated laser source, and an AOD which is imaged onto the integrating detector array. The output intensity of a laser diode can easily be modulated as a function of time by applying across the diode terminals a modulated voltage at a large enough DC bias so that the modulation signal remains within the linear regime. In this case the output intensity from a temporally modulated laser diode contains a modulation term on an optical bias, and it is given by $I_{LD} = A + v(t)$, where $v(t)$ can be bipolar. The emitted light from the laser diode can be collimated and aligned incident to an AOD at the Bragg angle, as shown in Figure 2.3.1. The diffracted intensity is given by the

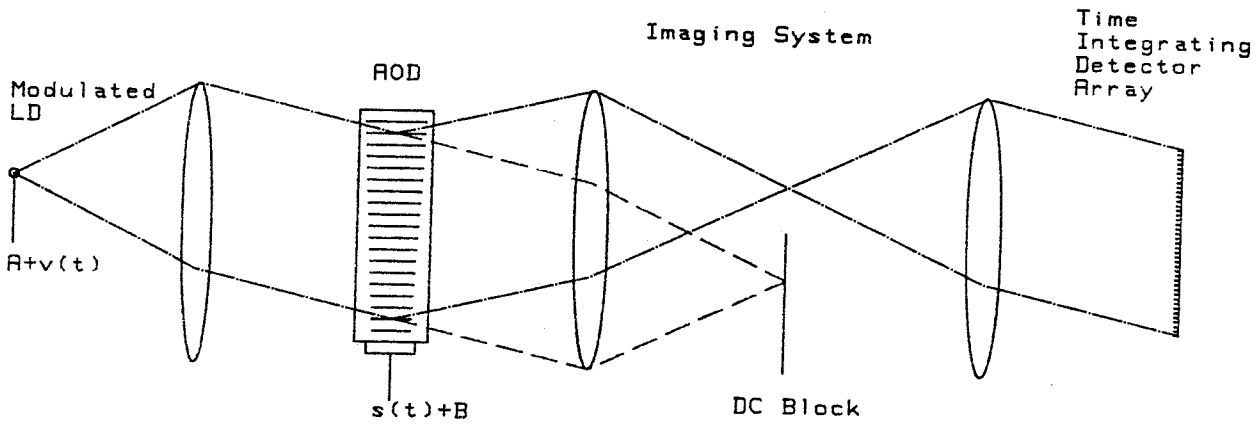


Figure 2.3.1. Time integrating AO correlator.

product of the laser diode intensity modulation with the AOD intensity diffraction expression given in Eq. 2.1.13. This output intensity is Schlieren imaged onto a time integrating detector array, which has a width equivalent to the (demagnified) image of the AOD. The apparent acoustic velocity in the image plane is given by the actual acoustic velocity times the demagnification factor, $v = v_a m_x$. The time integrated intensity that will be accumulated on the discrete detector array will be spatially sampled, so that after an integration period T_0 the output is given by

$$\begin{aligned} I(x_n) &= \int_{t'}^{t'+T_0} (A + v(t))\, w^2(x)\, [s(t - x/v) + B]\, dt \\ &= w^2(x_n) \left[T_0 AB + \int_{T_0} As(t - x_n/v)dt + \int_{T_0} Bv(t)dt + \int_{T_0} v(t)s(t - x_n/v)dt \right] \\ &= w^2(x_n) \left[T_0 AB + A\bar{s} + B\bar{v} + r_{vs}(x_n/v) \right] \end{aligned} \qquad (2.3.1.1)$$

For zero mean signals, $\bar{s} = 0$, and $\bar{v} = 0$, the signal dependent bias terms disappears and the resulting time integrated signal consists of an apodization function times a constant bias plus a finite time correlation between the signal applied to the laser diode $v(t)$, and the intensity diffracted by the Bragg cell $s(t)$. Depending on the acousto-optic intensity modulation technique which is used, the bias term B may contain a time dependent signal power bias such as that given in Eq. 2.1.10, which will not integrate to zero, and will result in a deleterious fluctuating bias level.

The most notable difference between time integrating and space integrating correlators is the presence of the bias terms, which can use up the available dynamic range of the detector array. The constant bias is relatively easy to remove with electronic post processing circuitry, but the signal dependent bias causes much more difficulty. The most common technique to remove the bias is spatial carrier demodulation[24], in which the correlation term is placed on a spatial carrier of a high enough frequency so that it can be unambiguously separated from the bias terms with an electronic bandpass filter on the serial output samples from the photodetector array. In this case both input signals are placed on temporal carriers,

so that the laser diode intensity modulation is given by

$$I_{LD}(t) = A[1 + m_1 v(t)\cos(2\pi f_0 t)], \tag{2.3.1.2}$$

and the AOD modulation is given by

$$\begin{aligned} I_{AOD}(t - x/v) &= [B^{\frac{1}{2}}(1 + m_2 s(t - x/v_a)e^{i2\pi f_0(t-x/v_a)}]^2 \\ &= B[1 + m_2^2 s^2(t - x/v_a) + 2m_2 s(t - x/v_a)\cos(2\pi f_0(t - x/v_a)]. \end{aligned} \tag{2.3.1.3}$$

The time integrated output contains the correlation term on a spatial carrier.

$$\begin{aligned} I(x_n) = AB\Big[T_0 &+ m_2^2 \int_{T_0} s^2(t - x_n/v)dt \\ &+ 2m_1 m_2 \cos(2\pi f_0 x_n/v) \int_{T_0} v(t)s(t - x_n/v)dt\Big] \end{aligned} \tag{2.3.1.4}$$

The other three terms contain a temporally oscillating cosine, which time averages to zero. This discretely sampled spatial signal representation is turned into a serial data stream by the readout mechanism incorporated in the photodetector array, that reads out one pixel of width Δx each pixel sample time Δt.

$$\begin{aligned} I_0(t_k) &= \sum_n I(x_n)\delta(n\Delta x - k\Delta t) \\ &= AB\left[T_0 + m_2^2 \int_{T_0} s^2(t - t_k v_d/v)dt + 2m_1 m_2 \cos(-2\pi f_0 t_k v_d/v) r_{vs}(t_k v_d/v)\right] \end{aligned} \tag{2.3.1.5}$$

Where the serial readout is at an effective spatial velocity $v_d = \Delta x/\Delta t$, which converts the sampled spatial carrier to a sampled temporal carrier of frequency $f_0 v_d/v$. The correlation signal $r_{vs}(t_k v_d/v)$ has a temporal bandwidth equal to the overlap bandwidth of the two signal $v(t)$ and $s(t)$, scaled by the velocity ratio $v_d/v = v_d/(v_a m_x)$ which is usually much less than one, since AODs have much higher acoustic velocity than photodetector arrays have readout velocity. In the case of the autocorrelation of two signals with two sided bandwidth B, the condition $f_0 > 3B$ will guarantee no spectral overlap between the bias terms and the correlation term so that a temporal bandpass filter with bandwidth Bv_d/v at center frequency $f_0 v_d/v$ can be used to remove the bias terms.

It is interesting to notice that an implicit spatial integration has been included in this equation to accomplish the discrete pixel sampling, that converts the continuous spatial modulation into a discrete photodetector output. Given a pixel response function that is independent of position and given by $h(x)$, and with a pixel spacing of Δx, then the conversion from the continuous spatial intensity to the discrete photodetector representation is accomplished with a spatial sampling operation.

$$\begin{aligned} I(x_n) &= \int h(x - n\Delta x) I(x) dx \qquad (2.3.1.6) \\ &= \int h(x - n\Delta x) w^2(x) \left[\int_{t'}^{t'+T_0} (A + v(t)) \left[s(t - x/v_a) + B \right] dt \right] dx \end{aligned}$$

This type of spatial integration due to pixel sampling is always implicititly present whenever discrete photodetector arrays are utilized. Notice that in order to properly sample the time integrated spatial intensity profile the maximum spatial frequency of the intensity at the photodetector plane must be less than the Nyquist limit of $1/2\Delta x$. This places a restriction on the spatial bandwidth of the intensity modulation of the AOD that can be satisfied when the number of photodetector pixels that the AOD is imaged onto is greater than twice the time bandwidth product of the Bragg cell.

An alternative approach to time acousto-optic integrating correlation utilizes two counterpropagating Bragg cells as shown in Figure 2.3.2. The reverse imaging of the intervening optical system results in the images of the intensity diffracted from the two AODs to be counterpropagating, even though in the figure they appear to be aligned. The relative motion between the signals in the two AODs allows us to perform a correlation integral by time integrating the doubly diffracted intensity.

$$\begin{aligned} I(x) &= \int_{t'}^{t'+T_0} [A + f(t + x/v)][B + g(t - x/v)] dt \\ &= ABT_0 + A \int_{T_0} g(t - x/v) dt + B \int_{T_0} f(t + x/v) dt + \int_{T_0} f(\tau + 2x/v) g(\tau) d\tau \\ &= ABT_0 + A\bar{g} + B\bar{f} + r_{fg}(2x/v) \qquad (2.3.1.7) \end{aligned}$$

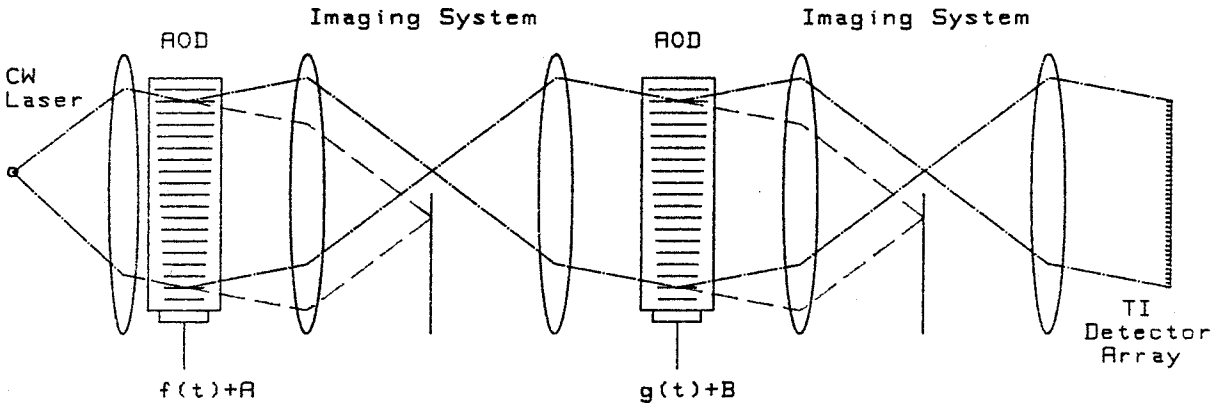


Figure 2.3.2. Counterpropagating AOD time integrating correlator.

In this case the spatial scale factor of the detected correlation integral is twice that of the modulated source correlator of the previous section. Once again there are several bias terms which will simplify if the signals are zero mean so that $\bar{f} = 0, \quad \bar{g} = 0$. The bias terms can be removed with a carrier demodulation approach as in the previous section. An alternative interferometric, or additive implementation, of this type of system has a higher optical throughput, since multiplicative diffractions from the AOD are avoided, and is often the preferable approach.

2.3.2 Pulsed source time integrating optical processors

Space invariant linear transformations such as the correlation and convolution operations are very useful, but some applications require more general space variant transformations such as vector matrix multiplications. A pulsed laser source can be utilized in a number of ways to implement these more general linear transformations on discrete data sequences using acousto-optic technology. The correlators and convolvers that have been considered are a special case of a general one dimensional linear transformation, in which the transformation kernel, or matrix, is Toeplitz, which means that entries along the diagonals are constant. The general space variant linear transformation of one dimensional data is given by a vector matrix product.

$$y_m = \sum_{n=1}^{N} A_{mn} x_n \qquad (2.3.2.1)$$

We can implement this transformation by sliding successive columns of the matrix A_{mn} into an AOD and freezing the acoustic motion by pulsing the laser source for a short interval and modulating its amplitude by successive samples of the vector x_n. This involves a raster transformation of the 2-dimensional matrix A_{mn} into a long 1-dimensional temporal signal composed of successive columns of the matrix, each of the appropriate duration to fill the portion of the AOD aperture that is being

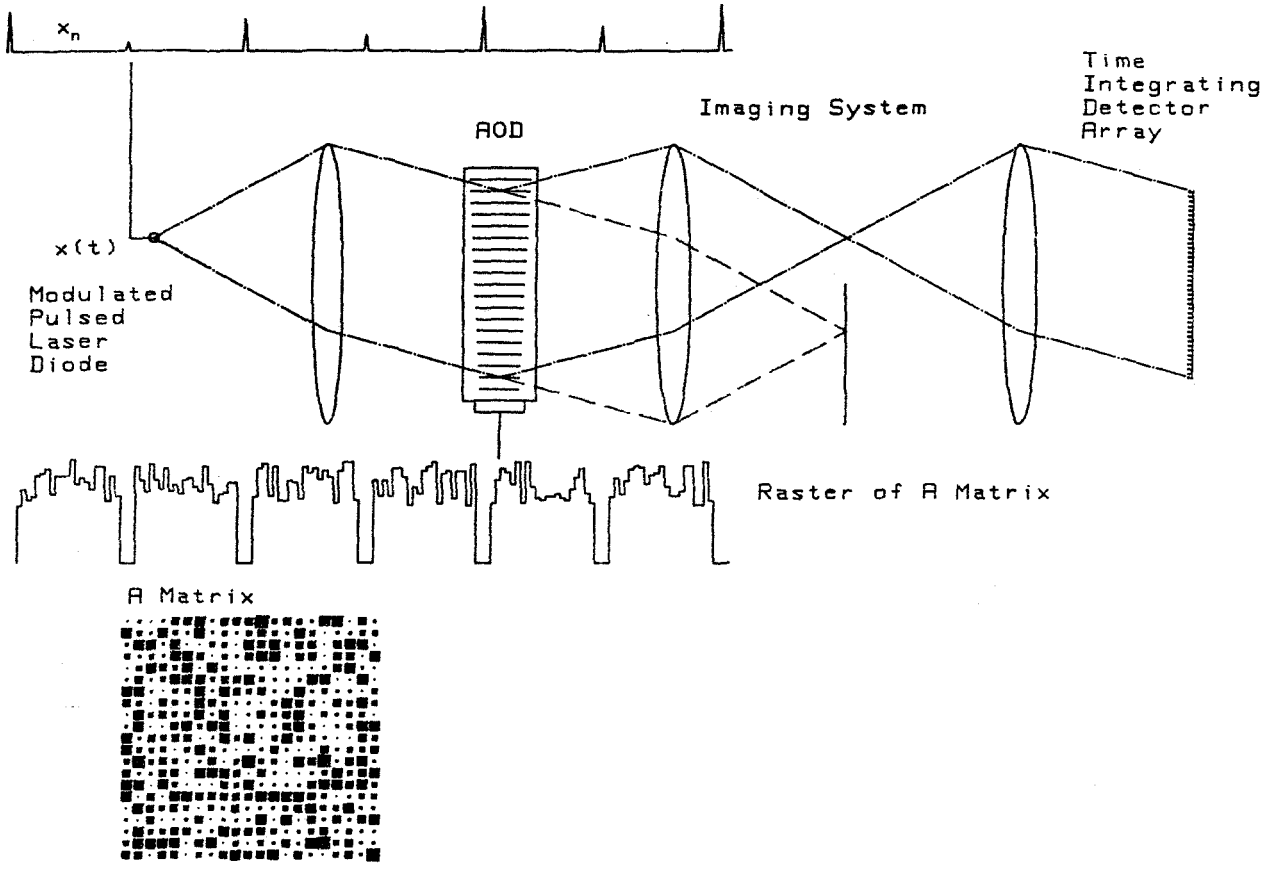


Figure 2.3.3. Time integrating pulsed source matrix vector multiplier.

imaged onto the time integrating detector array.

$$a(t) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \text{rect}\left[\frac{t - nT - m\tau}{\tau}\right] A_{mn} \qquad (2.3.2.2)$$

In this expression the time between row samples is given by τ, and the time between adjacent columns must be greater than the total time required for each column $T \geq M\tau$. The modulation of the pulsed laser diode is given by the expression

$$x(t) = \sum_{n=0}^{N-1} \text{rect}\left[\frac{t - n'T - t_0}{\Delta t}\right] x_{n'} \approx \sum_{n'=0}^{N-1} \delta(t - n'T - t_0) x_{n'} \qquad (2.3.2.3)$$

The width of the laser diode pulses $\Delta t << \tau$, must be small enough so that it essentially freezes the acoustic motion of the matrix columns, and the pulses are delayed by t_0 to facilitate the alignment of each column within the aperture of the AOD which is imaged onto the detector array. The signals a(t) and x(t) are applied to the AOD and LD, respectively, in the system illustrated in Figure 2.3.3, which is an identical architecture with the time integrating correlator shown in Figure 2.3.1. The diffracted intensity from the Bragg cell is imaged onto M time integrating detectors, spaced by $\Delta x = m_x v_a \tau = v\tau$, where m_x is the demagnification of the imaging system. The signal accumulated on the detector array after N laser diode pulses is given by the time integration of the product of the source modulation with the AOD diffracted intensity.

$$\begin{aligned} y_m(x = m\Delta x) &= \int_0^{NT} x(t) a(t - x/v) \text{rect}\left[\frac{x - X/2}{X}\right] dt \\ &= \sum_{n'=0}^{N-1} x_{n'} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \text{rect}\left[\frac{n'T + t_0 - x/v - nT - m\tau}{\tau}\right] A_{mn} \text{rect}\left[\frac{x - X/2}{X}\right] \\ &= \sum_{n=0}^{N-1} \text{rect}\left[\frac{X/2 - m\Delta x - x}{\Delta x}\right] A_{mn} x_n \\ &\propto \sum_{n=0}^{N-1} A_{mn} x_n \end{aligned} \qquad (2.3.2.4)$$

The portion of the AOD aperture that is utilized is $X = M\Delta x$, and the delay $t_0 = X/2v$, allows us to overlap the rectangle functions in space which enforces the

requirement that $n' = n$. The integrated output of the mth detector at position $m\Delta x$ is proportional to the mth element of the matrix vector product. In this expression the bias effects associated with time integrating intensity based processing have been ignored for analytical simplicity, but they will be present in actual implementations, so bias removal must again be employed.

2.3.3 Time delay and integration

Another approach to performing correlation with optical time integration is to utilize a modulated source with a stationary reference, and a sliding detector plane. This approach is called time delay and integration (TDI)[27–29], or shift and add, and is a complimentary technique to the normal modulated source time integrating correlator where the reference function in the AOD is a travelling wave representation, and the detector coordinate frame is stationary. However, they require different technologies for implementation, the TI correlator of section 2.3.1 require AODs in order to spatially shift the reference function with respect to the fixed detector, while the TDI approach requires a shifting time integrating detector array in order to transfer the photogenerated signals with respect to the fixed reference transparency. Because of limits on the TDI rate of currently available CCD detector arrays, the TDI approach should be used for low bandwidth correlation applications, but high speed GaAs CCDs may eliminate this limitation.

One possible implementation of a sliding detector frame is to translate a photosensitive medium such as film at a constant velocity in the detector plane, however a much more elegant implementation would involve no mechanically moving parts. The main reason that the TDI approach is important is the ease with which a CCD detector array can be operated in the shift and add mode, by simply modifying the applied electronic clock waveforms. In the TDI mode of operation the CCD array detects the incident optical intensity profile for a time Δt, then shifts the discrete

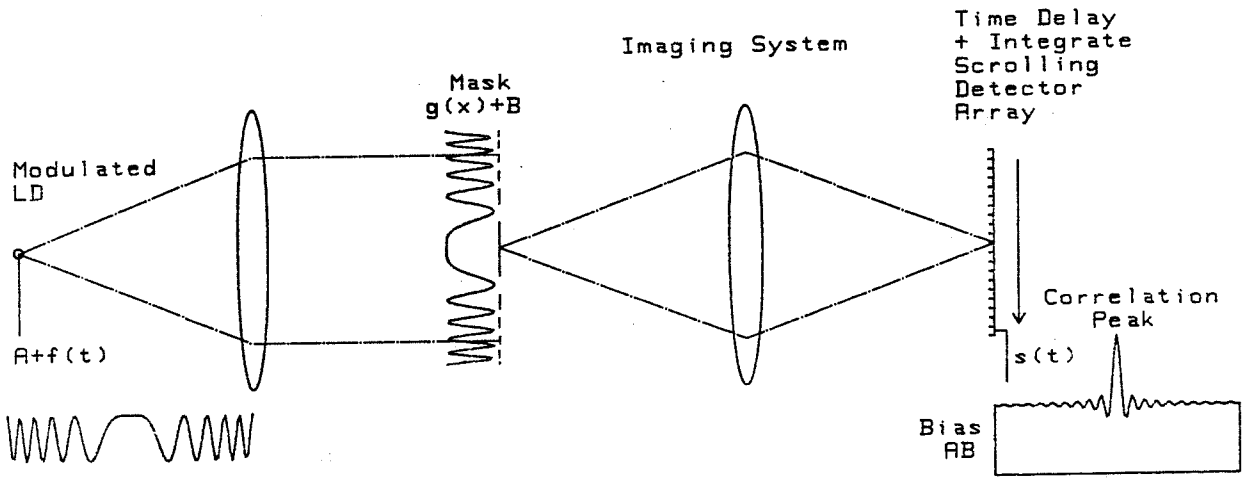


Figure 2.3.4. Time delay and integrate (TDI) Correlator based on scrolling CCD.

photogenerated charge profile along the detector length a distance Δx. New photogenerated electrons are added to those already accumulated in the previous cycles at successively shifted spatial locations. This process of shifting the accumulated charge, and adding the new photogenerated charge, continues until a potential well reaches the edge of the detector chip, where it is amplified and read off chip. Analytically the TDI operation can be described as a time integrating detection in a discretely shifting coordinate frame, for an arbitrary time and space varying optical signal $I(x,t)$.

$$s_m(t = m\Delta t) = \sum_{n=m-N}^{m-1} \int_{n\Delta t}^{(n+1)\Delta t} \int [h(x - (m-n)\Delta x) I(x,t) dx]\, dt \qquad (2.3.3.1)$$

The function $h(x)$ is the spatial response function of a pixel centered at position x, and for low enough spatial frequencies incident on the detector array it acts like a sampling impulse. The mth output, which occurs at time $t = m\Delta t$, is seen to consist of time integrated photogenerated charge from a succession of pixel positions which move linearly in space from the first pixel at position $x = N\Delta x$ at time $t = (m - N)\Delta t$, to the final integration position centered at position $x = \Delta x$ at time $t = (m-1)\Delta t$. If we assume that the pixel spatial response acts like a sampling impulse, and the time integration between shifts simply averages a slowly varying temporal dependence of each pixels incident intensity, thereby replacing t by n, then we can simplify the expression for a TDI CCD.

$$s_m(t = m\Delta t) = \sum_{n=m-N}^{m-1} \int \delta(x - (m-n)\Delta x) I(x,n) dx = \sum_{n=m-N}^{m-1} I[(m-n)\Delta x, n] \qquad (2.3.3.2)$$

We can implement a TDI correlator by replacing the arbitrary time and space varying intensity profile incident on the shift and add detector, with the product of a temporally modulated laser diode with a fixed transparency that is imaged onto the detector. This system is illustrated in Figure 2.3.4, where the intensity emitted

by the pulsed laser diode is given by

$$f(t) = \sum_{n=0}^{N-1} \text{rect}\left[\frac{t - n\Delta t}{\tau}\right] f_n \approx \sum_{n=0}^{N-1} \delta(t - n\Delta t) f_n \qquad (2.3.3.3)$$

The reference transparency is either a fixed computer generated 1-dimensional photographic transparency, or a spatial light modulator with a programmable spatial reference function. The mask intensity transmittance $g(x)$, is sampled by the detector at location $k\Delta x$ as $g(k\Delta x)$. The product of the emitted laser diode intensity and the mask transmittance is imaged onto the TDI CCD detector array. The signal produced by the TDI detector can be expressed as the sampled convolution of the most recent N samples of f_n with the reference function $g(x)$.

$$\begin{aligned} s_m(t = m\Delta t) &= \sum_{n=m-N}^{m-1} \int \delta(x - (m-n)\Delta x) f(t) g(x) dx \\ &= \sum_{n=m-N}^{m-1} f_n g((m-n)\Delta x) = c_{fg}(m\Delta t) \end{aligned} \qquad (2.3.3.4)$$

Spatially reversing the mask $g(x)$, or reversing the shift direction of the TDI detector array results in the correlation of the most recent N samples of the unipolar intensity modulated signal f with the reference signal $g(x)$. In order to implement bipolar signal processing operations it is again necessary to place the two signals on appropriate biases. This will, as usual, introduce deleterious bias terms into the resulting correlation that will have to be removed. The laser diode modulation can be placed on a temporal carrier, and the reference transparency can be placed on a spatial carrier, in order to facilitate a frequency domain approach to bias removal. Complex correlations can be performed once both signals are placed on carriers by simultaneously phase and amplitude modulating the respective temporal and spatial carriers. However, placing the signals on carriers will cut down the processing gain obtainable with an N stage TDI correlator, by requiring that the spatial modulation bandwidth be only 1/4 or less of the available number of CCD pixels in order to leave at least 4 pixels per cycle of the spatial carrier.

2.4 Chirp Transform Algorithm

The chirp transform algorithm[9,24–26,31] is the basis for a number of analog optical spectral analysis systems. It is a technique which allows the implementation of the shift variant Fourier kernel with shift invariant linear systems such as the correlators and convolvers described in the previous sections. We will make extensive use of this algorithm in Chapter 5 in order to implement high resolution spectral analysis systems and an introductory explanation of the functioning of the algorithm will be given here. There are several different forms of the chirp transform algorithm, both continuous and discrete implementations may be realized, space integrating[30] and time integrating[9,24–26] systems are commonly implemented, and both symmetric and asymmetric representations are possible. All these different realizations arise from the observation that the exponent of the shift variant Fourier kernel, $e^{-i2\pi ft}$, can be decomposed as follows into shift invariant operators.

$$-2ft = (t-f)^2 - f^2 - t^2 = f^2 + t^2 - (t+f)^2 = \frac{(t-f)^2 - (t+f)^2}{2} \qquad (2.4.1)$$

The first version of the expanded exponent will result in a convolutional expression of the algorithm and the second will result in a correlation approach. The two are equivalent because of the even symmetry of the quadratic. The final identity results in the symmetric counterpropagating implementation of the chirp algorithm, and does not require a chirp postmultiplication. Any of these expressions can be substituted into the definition of the Fourier transform to yield one of the common representations of the chirp transform algorithm.

$$\begin{aligned} S(f) &= \int s(t)e^{-i2\pi ft}dt \qquad (2.4.2)\\ &= e^{i\pi f^2}\int [s(t)e^{i\pi t^2}]e^{-i\pi(t+f)^2}dt = e^{i\pi f^2}\left[[s(t)e^{i\pi t^2}] \star e^{i\pi t^2}\right](f)\\ &= e^{-i\pi f^2}\int [s(t)e^{-i\pi t^2}]e^{i\pi(f-t)^2}dt = e^{-i\pi f^2}\left[[s(t)e^{-i\pi t^2}] * e^{i\pi t^2}\right](f)\\ &= \int s(t)e^{i\pi(t-f)^2/2}e^{-i\pi(t+f)^2/2}dt \end{aligned}$$

From the first form of this Equation the algorithm for computing the Fourier transform of the signal $s(t)$, is seen to consist of a premultiplication of the signal with a quadratic phase function or chirp in the time domain $e^{i\pi t^2}$, followed by a correlation with an identical chirp (or a convolution with a conjugate chirp), followed by a postmultiplication with a chirp in the frequency domain $e^{i\pi f^2}$. Often we are only interested in the power spectrum of the signal s(t) in which case we can neglect the final chirp postmultiplication, since this represents only a complex phase factor correction in the frequency domain.

$$|S(f)|^2 = \left| \int [s(t)e^{i\pi t^2}] e^{-i\pi(t+f)^2} dt \right|^2 \tag{2.4.3}$$

So the simplified chirp transform algorithm for the computation of the power spectrum consists of a chirp premultiply, followed by a chirp convolution with the resulting spectral representation modulus squared.

The algorithm can best be understood by examining the case of a single sinusoidal tone as the input signal, $s(t) = ae^{i2\pi f_0 t}$. When this CW tone is premultiplied by the chirp we can combine the exponential terms by completing the square.

$$\begin{aligned} S(f) =& e^{i\pi f^2} \left[ae^{i2\pi f_0 t} e^{i\pi t^2} \star e^{i\pi t^2} \right](f) \\ =& e^{i\pi f^2} \left[ae^{-i\pi f_0^2} e^{i\pi(t+f_0)^2} \star e^{i\pi t^2} \right](f) \\ =& a\delta(f - f_0) \end{aligned} \tag{2.4.4}$$

So the complex exponential single tone input acts to shift the position of the premultiplied chirp by an amount proportional to the frequency f_0, and the correlation operation measures the amount of that shift, resulting in a peak at the position corresponding to the input frequency. The quadratic phase factor postmultiplication term is needed to correct for the phase factor introduced in completing the square. The correlation (or convolution) of two chirps of infinite extent results in the delta function in the spectral domain, just as an infinite extent Fourier transform does.

Thus the chirp algorithm can be seen to exactly compute the Fourier transform of a single tone input, and since any input signal can be decomposed into a linear superposition of sine waves, the linearity of the system will give the appropriate Fourier transformation of an arbitrary signal.

A system implementing the chirp transform algorithm will integrate for a finite time T, and use finite bandwidth chirps, which will limit the spectral resolution that is achievable. Define a finite time chirp of bandwidth $B = bT$ that starts at time $t = 0$ at frequency f_l, and sweeps linearly to frequency $f_u = f_l + B$ in time T.

$$c(t) = \text{rect}\left[\tfrac{t-T/2}{T}\right] e^{i2\pi[f_l t + \frac{b}{2}t^2]} \tag{2.4.5}$$

This linear FM waveform has chirp acceleration parameter $b = B/T$, usually expressed in MHz/μsec, and center frequency $f_c = f_l + B/2$. By substituting this signal into the chirp transform algorithm and neglecting the chirp postmultiplication, we obtain the finite chirp transform expression.

$$\begin{aligned}
\hat{S}(b\tau) &= \int_0^T s(t)\text{rect}\left[\tfrac{t-T/2}{T}\right] e^{i2\pi[f_l t + \frac{b}{2}t^2]} \text{rect}\left[\tfrac{(t+\tau)-T/2}{T}\right] e^{-i2\pi[f_l(t+\tau) + \frac{b}{2}(t+\tau)^2]} dt \\
&= e^{-i2\pi[\tau f_l + \frac{b}{2}\tau^2]} \int \text{rect}\left[\tfrac{t-(T-\tau)/2}{T-|\tau|}\right] s(t) e^{-i2\pi b\tau t} dt \qquad (2.4.6) \\
&= e^{-i2\pi[\tau f_l + \frac{b}{2}\tau^2]} \int S(f)(T-|\tau|)\text{sinc}[(T-|\tau|)(b\tau - f)] e^{-i\pi(b\tau - f)(T-\tau)} df \; |\tau| < T
\end{aligned}$$

As the overlap of the finite extent chirps decreases, the gain of the transform decreases linearly and the resolution decreases linearly, until there is no chirp overlap and the output becomes zero. The output chirp transform is in a scaled coordinate system, with frequency variable $f = b\tau$ taking on non-zero values for the region $|\tau| < T$, which implies that the two sided analysis bandwidth is twice the chirp bandwidth $B_A = 2bT$, but the analysis bandwidth within the region of gain and resolution loss of a factor of two is equal to the chirp bandwidth. The range of spectral analysis is centered on the difference of the center frequencies of the premultiply chirp and the reference chirp, which in this case is 0, so both positive

and negative frequency components may be discerned. By using a longer duration, wider bandwidth chirp as the reference function, the chirp overlap within the integration time will become trapezoidal instead of triangular, and the resolution and gain will remain constant in the region of constant chirp overlap. As in any spectral analysis system the technique of apodization can be used to decrease the sidelobe levels of the sinc function at the expense of a loss in resolution, and this can be simply accomplished in this system by weighting the premultiply chirp or the reference chirp with a smooth function such as a Gaussian. For a single tone CW input at frequency f_0 with amplitude a, the complex spectrum is given by $S(f) = a\delta(f - f_0)$, so the output of the chirp spectrum analyzer is given by

$$\hat{S}(b\tau) \approx e^{-i2\pi f_c \tau} e^{i\pi f_0 (T-\tau)} a(T - |\tau|)\mathrm{sinc}[(T - |\tau|)(b\tau - f_0)] \tag{2.4.7}$$

So the system impulse response has a gain and resolution that linearly improve towards the central frequency, and the sinc function impulse is on a carrier of frequency $f_c - f_0$. Interestingly with this realization the explicit quadratic postmultiply chirp is canceled by the phase factor due to the motion of the centroid of the product of the two rect functions.

In some of the implementations to be considered the signals will be real instead of complex, and the chirp transform will implement a two sided cosine transform instead of a Fourier transform.

$$\begin{aligned}\hat{S}(b\tau) &= \int s(t)p(t)\mathrm{rect}\left[\tfrac{t}{T_1}\right]\cos[2\pi(f_c t + \tfrac{b}{2}t^2)] \\ &\qquad \mathrm{rect}\left[\tfrac{t+\tau}{T_2}\right]\cos[2\pi(f_c'(t+\tau) + \tfrac{b}{2}(t+\tau)^2)]dt \qquad (2.4.8)\\ &= \int s(t)p(t)\mathrm{rect}\left[\tfrac{t}{T_1}\right]\mathrm{rect}\left[\tfrac{t+\tau}{T_2}\right]\frac{1}{2}\Big[\cos[2\pi(f_c'\tau + (f_c' - f_c)t + \tfrac{b}{2}(2t\tau + \tau^2))] \\ &\quad + \cos[2\pi(f_c'\tau + (f_c' + f_c)t + \tfrac{b}{2}(2t^2 + 2t\tau + \tau^2))]\Big]dt\end{aligned}$$

In this generalized expression for a finite real chirp transform, an apodization function $p(t)$ has been included, and the durations of the premultiply chirp, T_1, and

the reference chirp, T_2, in general may be different. The chirps are represented symmetrically about the respective center frequencies, which may be different if analysis of a spectral region away from baseband is desired. When real chirps are used in place of the earlier complex notation we see that both sum and difference terms are generated. The difference term results in the desired cosine transform of the apodized signal, but the sum term retains a quadratic phase factor that upon integration will produce a term similar to a Fresnel cosine integral. For high time bandwidth chirps this will produce a small chirping ripple contribution due to each signal frequency component, containing as much power as the desired peak, which can cause serious degradation of the system performance in a multitone signal environment. When the reference chirps are on a high enough carrier frequency this term can be ignored unless the signal contains a matched chirp component, in which case a spurious peaked integral could arise.

When an interferometric implementation is used to generate the multiplication between the premultiplied chirp and the shifting reference chirp, then only the desired difference term is generated, and the quadratic sum term is absent. An example of an interferometric chirp transform processor based on the TDI correlator presented in section 2.3 is shown in Figure 2.4.1. In this system the signal is applied to an AO modulator in one arm of an interferometer, while the time domain chirp reference is applied to a modulator in the other arm. An auxiliary wavefront curvature is applied to the reference arm with a movable lens so that the spatial interference pattern generated on the detector array is a chirp whose curvature matches that of the temporal reference chirp, with the TDI velocity as the conversion factor. Interferometric detection accomplishes the chirp premultiplication, while time integration in the TDI coordinate frame accomplishes the chirp convolution of the matched chirps in order to produce a spectral scan at the detector output. In Chapter 5 this type of system is successfully applied to the folded spectrum systems

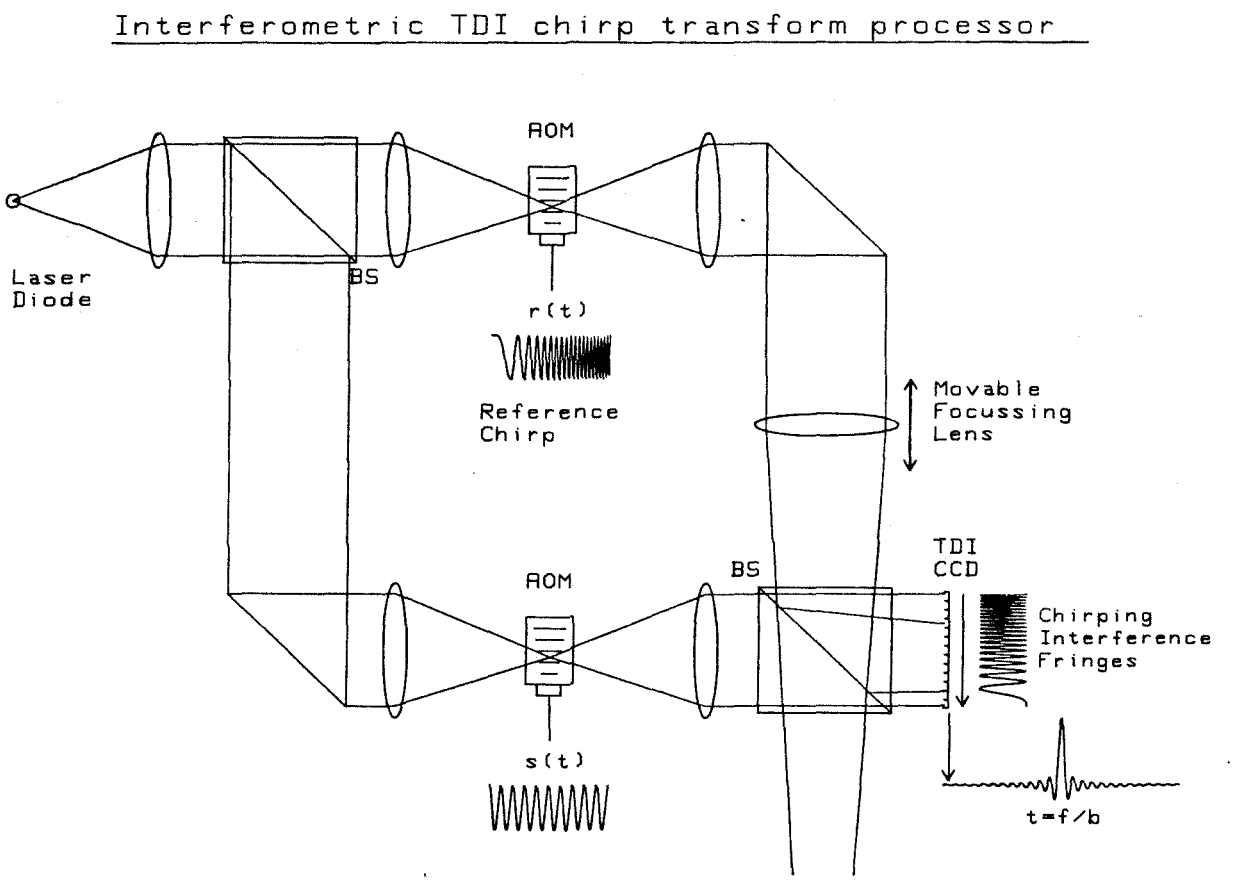


Figure 2.4.1. Interferometric TDI chirp transform spectral analysis system.

that are the major topic of this thesis. In any of the additive implementations of the interferometric chirp transform system the output does not contain the unwanted sum term of the cosine arguments and the desired output can be found from the first term in Equation 2.4.8.

$$\begin{aligned}
\hat{S}(b\tau) &= \int \left| \int S(f) e^{i2\pi ft} df\, p(t) \mathrm{rect}\left[\tfrac{t}{T_1}\right] e^{i2\pi[f_c t + \frac{b}{2}t^2]} + \mathrm{rect}\left[\tfrac{t+\tau}{T_2}\right] e^{i2\pi[f_c'(t+\tau) + \frac{b}{2}(t+\tau)^2]} \right|^2 dt \\
&= \int \left| s(t)p(t)\mathrm{rect}\left[\tfrac{t}{T_1}\right] \right|^2 + \left| \mathrm{rect}\left[\tfrac{t+\tau}{T_2}\right] \right|^2 dt + \int |S(f)| \int p(t)\mathrm{rect}\left[\tfrac{t}{T_1}\right] \mathrm{rect}\left[\tfrac{t+\tau}{T_2}\right] \\
&\qquad \cos[2\pi(f_c'\tau + (f_c' - f_c)t + \tfrac{b}{2}(2t\tau + \tau^2) - ft) - \Omega(f)] dt df \qquad (2.4.9) \\
&= \int \left| s(t)p(t)\mathrm{rect}\left[\tfrac{t}{T_1}\right] \right|^2 + \left| \mathrm{rect}\left[\tfrac{t+\tau}{T_2}\right] \right|^2 dt \\
&\quad + \cos[2\pi(f_c'\tau + \tfrac{b}{2}\tau^2)] \int |S(f)| \int p(t)\mathrm{rect}\left[\tfrac{t}{T_1}\right] \mathrm{rect}\left[\tfrac{t+\tau}{T_2}\right] \\
&\qquad \cos[2\pi(b\tau t + \delta f_c - f)t - \Omega(f)] dt df \\
&\quad + \sin[2\pi(f_c'\tau + \tfrac{b}{2}\tau^2)] \int |S(f)| \int p(t)\mathrm{rect}\left[\tfrac{t}{T_1}\right] \mathrm{rect}\left[\tfrac{t+\tau}{T_2}\right] \\
&\qquad \sin[2\pi(b\tau + \delta f_c - f)t - \Omega(f)] dt df
\end{aligned}$$

The first term is a bias term due to the interferometric detection, while the interferometric term can be split into its quadrature components. The cosine and sine transform quadrature components produce spectral scans centered at $\delta f = f_c' - f_c$, with a scaled output frequency variable $f = b\tau$. The quadrature components ride on chirping carriers that can be demodulated in quadrature, or envelope detection can be used to measure spectral amplitude. In this case the interferometric component which generates the difference frequency will be twice as strong as that in the multiplicative approach. This is because the interferometric implementation results in a fringe which slides along with the output variable producing a constructive contribution at all integration positions, while the multiplicative approach produces chirp patterns which beat on and off with the integration variable and contribute constructively only half the time.

In the special case when both chirps are centered at DC, $f_c = f_c' = 0$, we obtain

an interferometric difference term which contains the two quadrature components with chirping carriers centered at DC. Neglecting the bias terms from Eqn. 2.4.9 and performing the baseband chirp transform we obtain the following expression.

$$
\begin{aligned}
\hat{S}(b\tau) \approx & \cos[\pi b\tau^2] \int \tilde{s}^*(t)p(t)\mathrm{rect}\left[\tfrac{t}{T_1}\right]\mathrm{rect}\left[\tfrac{t+\tau}{T_2}\right]\cos[2\pi b\tau t]dt \qquad (2.4.10)\\
& + \sin[\pi b\tau^2] \int \tilde{s}^*(t)p(t)\mathrm{rect}\left[\tfrac{t}{T_1}\right]\mathrm{rect}\left[\tfrac{t+\tau}{T_2}\right]\sin[2\pi b\tau t]dt \\
& = \cos[\pi b\tau^2] \int |S(f)| * P(f) \int_{\tau h(\tau-\frac{T_2-T_1}{2})-\frac{T_1}{2}}^{\tau h(\frac{T_2-T_1}{2}-\tau)+\frac{T_1}{2}} \cos[2\pi(b\tau - f)t - \Omega(f)]dt \\
& + \sin[\pi b\tau^2] \int |S(f)| * P(f) \int_{\tau h(\tau-\frac{T_2-T_1}{2})-\frac{T_1}{2}}^{\tau h(\frac{T_2-T_1}{2}-\tau)+\frac{T_1}{2}} \sin[2\pi(b\tau - f)t - \Omega(f)]dt \\
& = \cos[\pi b\tau^2] \int |S(f)| * P(f)\mathrm{trap}_{\Delta T}[\tfrac{\tau}{T_1}]\mathrm{sinc}\left[(b\tau - f)\mathrm{trap}_{\Delta T}[\tfrac{\tau}{T_1}]\right] \\
& \quad \cos[\pi(b\tau - f)\mathrm{ledge}_{\Delta T}[\tfrac{\tau}{T_1}] - \Omega(f)] \\
& - \sin[\pi b\tau^2] \int |S(f)| * P(f)\mathrm{trap}_{\Delta T}[\tfrac{\tau}{T_1}]\mathrm{sinc}\left[(b\tau - f)\mathrm{trap}_{\Delta T}[\tfrac{\tau}{T_1}]\right] \\
& \quad \sin[\pi(b\tau - f)\mathrm{ledge}_{\Delta T}[\tfrac{\tau}{T_1}] - \Omega(f)]
\end{aligned}
$$

The limits of integration are given by the overlap of the two rectangle functions, and are specified in terms of the heavyside function $h(t) = \begin{cases} 1 & t > 0 \\ 0 & t \leq 0 \end{cases}$. The convolution of the two rectangular window functions results in a trapezoidal resolution and gain of the spectral estimate, and an additional phase term due to the motion of the centroid of the product of the two rectangle functions. The trapezoidal function $\mathrm{trap}_{\Delta T}[\frac{\tau}{T_1}]$ has a flat top of width $\Delta T = |T_2 - T_1|$ at an amplitude of T_1, with $\pm$ unit slope sides extending a distance T_1 from the flat top, and this trapezoidal function becomes triangular when $\Delta T = 0$, $\mathrm{trap}_0[\frac{\tau}{T_1}] = (T_1 - |\tau|)\mathrm{rect}[\frac{\tau}{2T_1}]$. The additional phase factor function $\mathrm{ledge}_{\Delta T}[\frac{\tau}{T_1}]$ is zero over the central region $|\tau| < \frac{\Delta T}{2}$, and has unity slope extending for a distance T_1 outside this region, hence the name "ledge" and it becomes simply $\tau\mathrm{rect}[\frac{\tau}{2T_1}]$ when $\Delta T = 0$. Over the region when the chirp overlap is constant, $|\tau| < \frac{\Delta T}{2} = |\frac{T_2-T_1}{2}|$, the resolution and gain are constant, and no additional phase factors are introduced due to the motion of the centroid

of the region of chirp overlap. Thus in this region the computation of the chirp transform is ideal, and the quadratic phase factor can be canceled out with the appropriate chirp postmultiply circuitry. In the special case when the reference chirp has twice the duration of the premultiply chirp $T_2 = 2T_1$, then the analysis bandwidth over the region of constant resolution is given by the premultiply chirp bandwidth $B_1 = bT_1$. The resolution with no apodization is given by the sinc function width, which in the region of constant resolution is given by the inverse integration time $\Delta f = \Delta\tau/b = 1/T_1$. The number of Rayleigh resolvable frequency elements within the region of constant resolution is given by the ratio of bandwidth to resolution, which is equivalent to the premultiply chirp time-bandwidth product, $B_1/\Delta f = bT_1^2 = B_1T_1$.

When the duration of the two chirps are the same, $T_1 = T_2$ then this equation becomes somewhat simpler, and it can be recognized as just the two quadrature components corresponding to Eq. 2.4.6.

$$\begin{aligned}\hat{S}(b\tau) = &\cos[\pi b\tau^2]\int |S(f)| * P(f)(T_1 - |\tau|)\mathrm{sinc}\Big[(b\tau - f)(T_1 - |\tau|)\Big] \\ &\cos[\pi(b\tau - f)\tau - \Omega(f)]df \\ &- \sin[\pi b\tau^2]\int |S(f)| * P(f)(T_1 - |\tau|)\mathrm{sinc}\Big[(b\tau - f)(T_1 - |\tau|)\Big] \\ &\sin[\pi(b\tau - f)\tau - \Omega(f)] \qquad (2.4.11)\end{aligned}$$

This equation is valid over the region $-T < \tau < T$, and in this region the gain and resolution of the spectral scan are triangular. In theory these two quadrature components could be electronically demodulated, but practically it would be difficult to maintain enough chirp stability to accomplish this in an analog system.

In Figure 2.4.2 and 2.4.3 are illustrated some of the important characteristics of this algorithm for both the multiplicative and interferometric implementations. In the first figure a baseband double sided reference chirp with time bandwidth

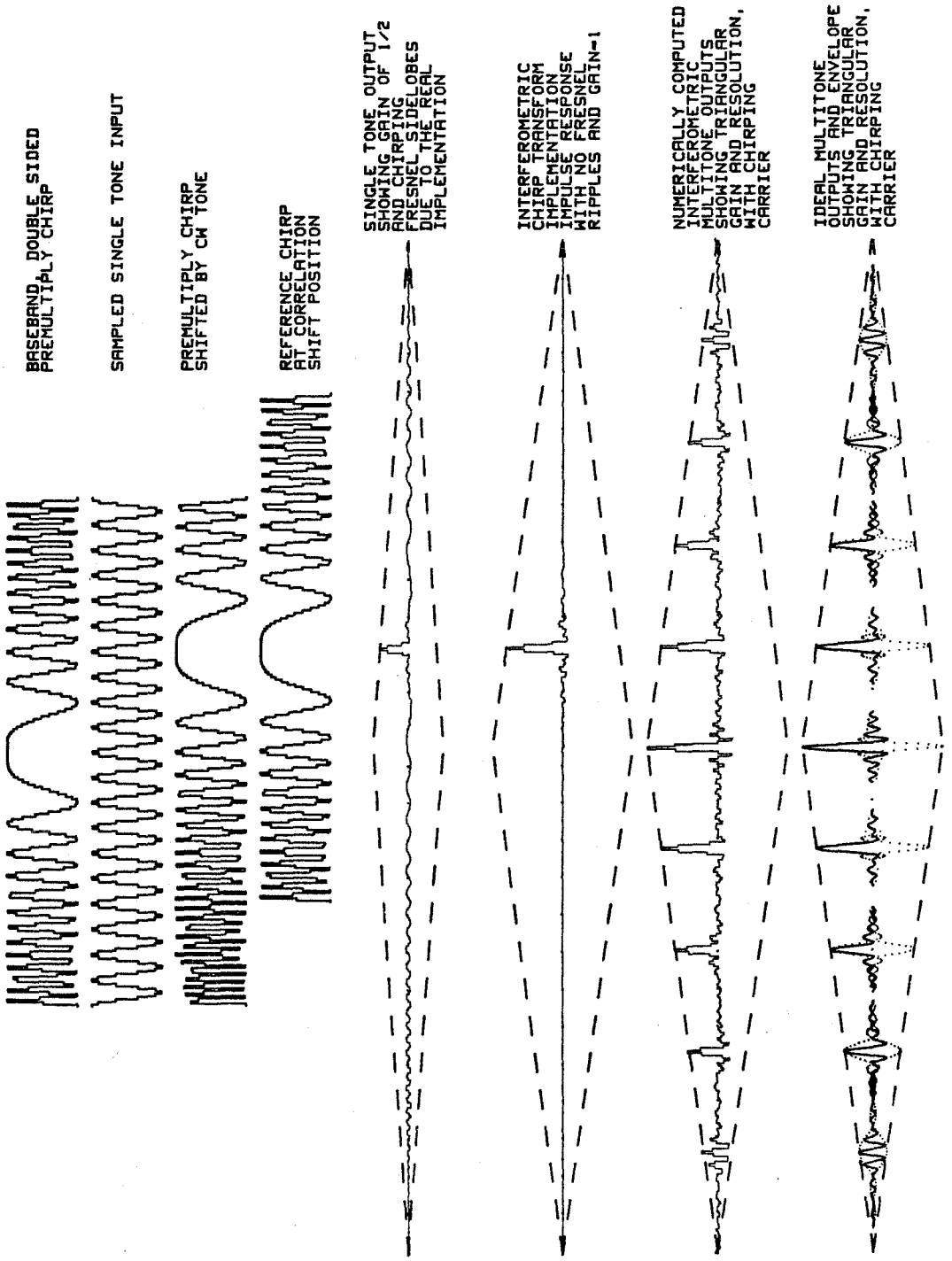


Figure 2.4.2. Baseband chirp transform algorithm.

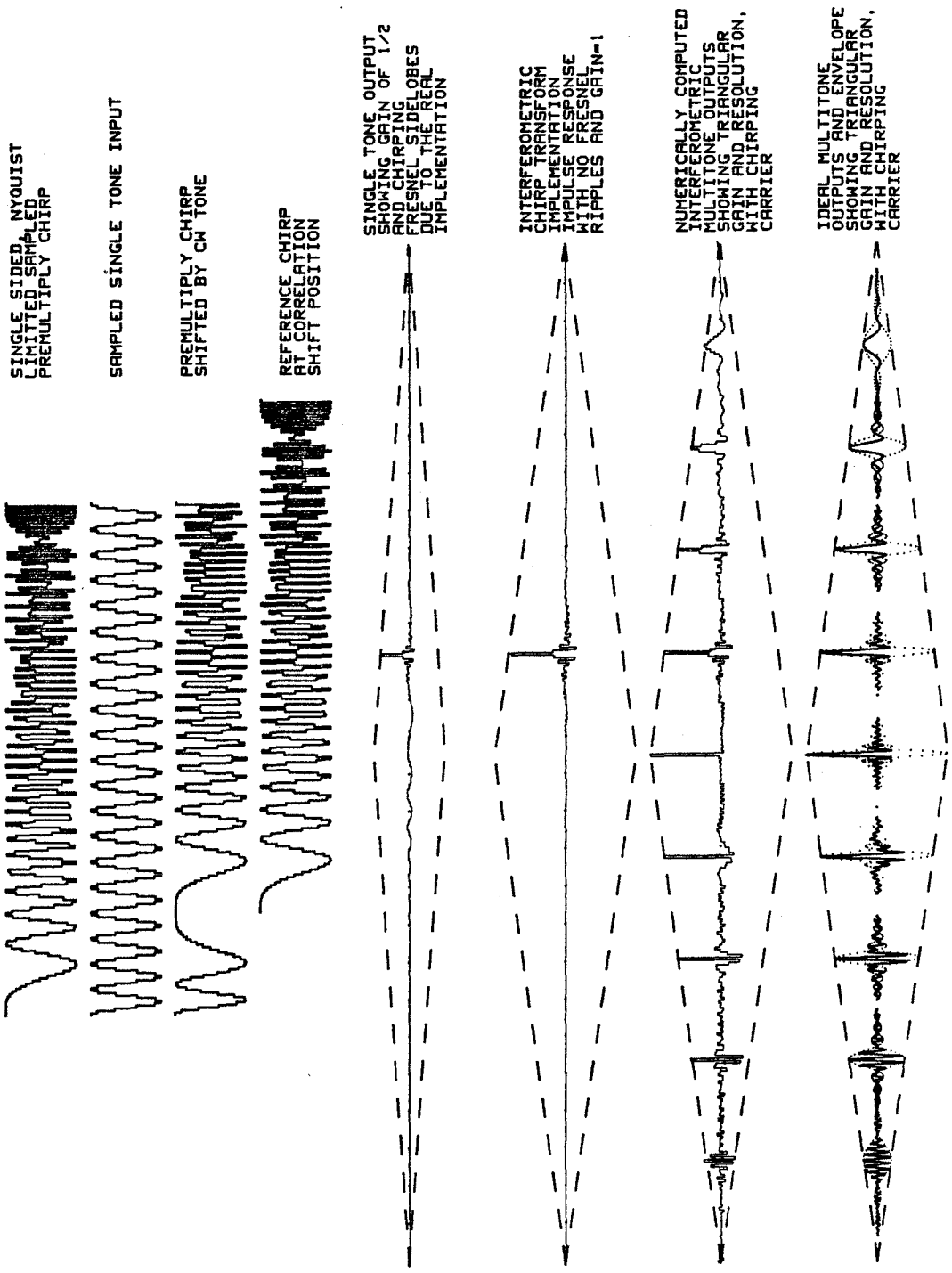


Figure 2.4.3. Single sided Nyquist limited chirp transform algorithm.

product TB=100 is multiplied by a constant frequency complex sinusoid producing a shifted chirp. If the sinusoid is real, then the chirp is shifted in both positive and negative directions, and two peaks will be generated. The reference chirp slides by the shifted chirp and at the position where they are maximally aligned a peak is generated. In the real implementation of this baseband correlation integral a peak whose gain is one half times the triangular chirp overlap is produced, and the rest of the power is miss focused into the chirping Fresnel sidelobes. These chirping sidelobes have half the chirp rate of the input chirps, with twice the duration, so they have the same baseband bandwidth as the peak, and therefore can not be separated from the peak. In a multisignal environment the sidelobes can constructively add up to produce false peaks, or can destructively add up to cancel a true peak. Since there is equal power in the peak as in the chirping sidelobes the signal to noise ratio (SNR) of this type of real baseband chirp transform processor is always 1:1, so this is not an acceptable spectrum analysis technique. However, when an interferometric correlation technique is utilized, then the chirping sidelobes are eliminated, and all the power is concentrated into the desired peak. In the numerical simulation of a multitone environment the respective peaks are clearly visible, and have the expected triangular gain, and the resolution appears to decrease away from the DC analysis position as expected. The carrier frequency is given by the highest frequency of the overlapping chirps, which in the case of baseband chirps is symmetric for both positive and negative frequencies. The theoretical results from Eq. 2.4.7 are plotted at the bottom of the figure with the corresponding peak envelopes averaged over different phases (dotted), and they can be seen to be in excellent agreement with the numerical simulations. In the following figure the case of a single sided Nyquist limited chirp with a TB=100 is plotted, and some important differences become apparent. This chirp has the same chirp rate and time bandwidth as the previous figure, but is on a carrier of 4 pixels per cycle. When

a Nyquist limited up chirp is down shifted then it becomes more and more like a double sided chirp as is illustrated. When it is up shifted then part of the chirp is under sampled and aliases to produce a down chirp, but this does not correlate with the shifting reference up chirp and the algorithm still works. When a real chirp correlation is performed the Fresnel sidelobe term is significantly suppressed from the baseband case, although the gain is still only one half. The interferometric chirp transform produces twice the gain of the real case, and shows no Fresnel sidelobes at all, and so is the preferable implementation. The numerically computed multitone case shows the triangular gain and resolution as expected, and in this single sided case the chirping carrier is readily apparent. Although it may appear that this single sided chirp transform has better spectral resolution than the double sided case, this is misleading because this transform comes out on a higher frequency carrier, and actually the phase averaged resolutions are identical. The theoretical plots of impulse response and phase averaged resolution agree quite well with the numerical solutions, except for sampling effects associated with the large carrier frequencies.

The chirp transform algorithm can be implemented with discrete systems, and in this context it is sometimes referred to as the chirp z-transform[31]. In order to perform a discrete Fourier transform (DFT) upon a sequence $f(n)$, we can perform a similar substitution as in the continuous case.

$$F(m) = \sum_{n=0}^{N-1} f(n) e^{-i2\pi nm/N} = e^{-i\pi m^2/N} \sum_{n=0}^{N-1} [f(n) e^{-\pi n^2/N}] e^{i\pi (n-m)^2/N} \qquad (2.4.12)$$

Thus a DFT can be accomplished by premultiplying a sampled data sequence with a sampled chirp, then correlating (or convolving) with a sampled chirp with the opposite curvature, and finally postmultiplying with a matched chirp in the sampled spectral domain. The same considerations apply as in the continuous case with respect to finite duration transforms, windowing, apodization and real implementations. The additional constraint of the Nyquist limit applied to all signals within the system limits the achievable analysis bandwidth to the Nyquist limit.

References

[1]. N.J. Berg and J.N. Lee Eds., *Acousto-optic signal processing*, Marcel Decker, New York (1983).

[2]. J.W. Goodman, *Introduction to Fourier Optics* McGraw-Hill, San Francisco (1968).

[3]. K. Preston, Use of the Fourier transforming properties of lenses for signal spectrum analysis, in Optical and Electrooptical Information Processing, J.T. Tippet et al. Eds. MIT Press, Cambridge Mass. (1965).

[4]. A. Vander Lugt, Signal detection by complex spatial filtering, IEEE Trans. on Inf. Th., IT-10, p. 139 (1964).

[5]. L.J. Cutrona, Recent developments in coherent optical technology, in Optical and Electrooptical Information Processing, J.T. Tippet et al. Eds. MIT Press, Cambridge Mass. (1965).

[6]. J.W. Goodman, A.R. Dias, and L.M. Woody, Fully parallel high-speed incoherent optical method for performing discrete Fourier transforms, Optics Letters, vol. 2(1), p. 1 (1978).

[7]. J.W. Goodman, P. Kellman and E.W. Hansen, Linear space-variant optical processing of 1-D signals, Appl. Opt., vol 16(3), p. 733 (1977).

[8]. D. Psaltis, E. Paek and J. Hong, Acoustooptic implementation of the Hopfield model, abstract in JOSA, vol. 2(13), p. 48 (1985).

[9]. T. Turpin, Spectrum analysis using optical processing, Proc IEEE, vol. 69(1), p. 79 (1981).

[10]. P. Kellman, H. Shaver, and J.W. Murray, Integrating acousto-optic channelized receivers, Proc. IEEE, vol. 69(1), p. 93 (1981).

[11]. A. Vander Lugt, Interferometric spectrum analyzer, Appl. Opt., vol. 20(16), p. 2770 (1981).

[12]. D. Psaltis and V.K. Kumar, Acoustooptic spectral estimation: a statistical

analysis, Appl. Opt., vol. 20(4), p. 601 (1981).

[13]. W. T. Rhodes, Acousto-optic signal processing: convolution and correlation, Proc. IEEE, vol. 69(1), p. 65 (1981).

[14]. R. Sprague, A review of acousto-optic signal correlators, Opt. Eng., vol. 16(5), p. 467 (1977).

[15]. W.T. Maloney, Acoustooptical approaches to radar signal processing, IEEE Spectrum,

[16]. D. Psaltis and J. Hong, Adaptive acoustooptic filter, Appl. Opt., vol. 23(9), p. 3435 (1984).

[17]. D. Casasent et al., Acoustooptic matched filter correlator, Appl. Opt., vol. 21(13), p. 2356 (1982).

[18]. J.S. Gerig and H. Montague, A simple optical filter for chirp radar, Proc. IEEE, vol. 52, p. 1753 (1964).

[19]. D.H. McMahon, Wideband pulse compression via Brillouin scattering in the Bragg limit, Proc. IEEE, vol. 55(9), p. 1602 (1967).

[20]. M.B. Shultz et al., Optical pulse compression using Bragg scattering by ultrasonic waves, Appl. Phys. Lett., vol. 11(7), p. 237 (1967).

[21]. J.H. Collins et al., Pulse compression by Bragg diffraction of light with microwave sound, Appl. Phys. Lett., vol. 11(7), p. 240 (1967).

[22]. R.A. Sprague and C.L. Koliopoulos, Time integrating acoustooptic correlator, Appl. Opt., vol. 15(1), p. 89 (1976).

[23]. T.M. Turpin, Time integrating optical processors, SPIE 154-29, p. 196 (1978).

[24]. P. Kellman, Time Integrating optical signal processing, PhD. Thesis, Stanford Univ. (1979).

[25]. P. Kellman, Time Integrating optical signal processing, Opt. Eng., vol. 19(3), p.370 (1980).

[26]. P.S. Guilfoyle, Joint transform time-integrating acousto-optic correlator for

chirp spectrum analysis

[27]. K. Bromley et al., Electro-optic signal processing module, SPIE 180-07, p. 107 (1979).

[28]. K. Bromley, An optical incoherent correlator, Appl. Opt., vol. 21(1), p. 35 (1974).

[29]. M. Copeland et al., An optical CCD convolver, IEEE Trans. Elec. Dev., ED-23(2), p. 152 (1976).

[30]. D. Hotz, Bulk acousto-optic signal processor for continuous Fourier transforms, Appl. Opt., vol. 23(10), p. 1613 (1984).

[31]. L.R. Rabiner et al., The chirp z-transform algorithm, IEEE Trans on Audio and Electroacoustics, vol. AU-17(2), p. 86 (1969).

CHAPTER 3

TIME AND SPACE INTEGRATING SIGNAL PROCESSING

Time integrating and space integrating signal processing techniques, as described separately in the previous chapter, can be combined in order to yield more powerful multidimensional signal processing modules with some of the best characteristics of both methods. The primary motivation for investigating these hybrid techniques is the commercial availability of high quality laser diodes, acousto-optic devices and two dimensional charge coupled device photodetector arrays which can be effectively utilized to synthesize efficient signal processing modules with powerful computational abilities. The technique of temporal and spatial integration (TSI) allows the real time computation of certain classes of 2-dimensional linear transformations, without requiring the use of real-time 2-D spatial light modulators, as would be required by a purely space integrating system. The drawback of the TSI technique is the signal dependent bias that is always associated with temporal integration, which builds up within each space integrated channel of the detector array. Although this problem is not as severe as the uniform signal dependent bias term that floods the entire detector in a 2-D time integrating processor, its nonuniformity and signal dependence makes it a serious drawback of the TSI technique. However, the bias terms can be effectively eliminated through the use of spatial carrier encoding, or with an auxiliary bias subtraction CCD.

Time and space integrating (TSI) optical processing is a hybrid technique for performing signal accumulation along two axis of a signal space. Space integration (SI) is accomplished with a lens system or a free space propagation, that collects all of the light in a particular optical mode onto an array of corresponding photodetectors, thereby channelizing the modes at spatially multiplexed locations. Time

integration (TI) involves the sequential accumulation of time varying optical waveforms in an array of parallel photodetectors, which accumulate the photogenerated charge in the detector capacitance. There are always a number of implicit spatial integrations performed in any AO signal processing system, such as Fourier plane filtering to remove the undiffracted beam, and integration over the spatial aperture of the sampling pixels. Similarly, there are a number of types of implicit temporal integrations that are unavoidable in any AO signal processing system, such as the time integration implied by the finite frequency response of a photodetector. Temporal integration of the output of a space integrating system can be performed on a 1-D photodetector array in order to improve the SNR of the Fourier transform of a signal input to a Bragg cell power spectrum analyzer[1], as discussed in section 2.1. Temporal integration can be performed in a coordinate frame moving synchronously with the output of a space integrating acousto-optic Vander Lugt correlator or a free space chirp correlator in order to average out coherent artifacts[2], as discussed in sections 2.2.4 and 2.2.6. Alternatively, temporal integration can be performed on a 2-D photodetector array to improve the detection statistics of a space integrated time independent wavefront, such as in a Vander Lugt filter, which is usually referred to as exposure control rather than TSI processing. Another type of system uses space integration to decrease the dimensionality of an image via a projection transformation on sequential rows applied to a Bragg cell, then performs a time integrating correlation of the projected image with a projected reference in order to implement a simple recognition without using two dimensional transformations[3]. However, in this thesis I will only consider TSI processing systems where the spatial integration performs a channelization onto a two dimensional array of photodetector columns, while an explicit time integrating inner product signal processing operation is performed at each photodetector site along each detector column. The time integration at each pixel location is performed on a dynamically varying signal wavefront incident on that detector column, which is beat against a family of refer-

ence wavefronts. If the slow variations of the signal incident on a detector column matches one of the references of that column, then a DC component is generated, thereby producing a time integrated peak at the corresponding pixel, and localizing the signal in the dimension orthogonal to the space integration. Thus the time integration performs a further signal channelization operation into a 2-D output space of signal parameters, which is represented in the two spatial dimensions of the photodetector array. This is therefore a multidimensional signal processing technique, with integrations performed over two variables, time and space, projecting the input signal onto the two dimensional parameterization represented by the detector array. The coherent processing operation performed by the multichannel TI transformation sacrifices a potential noncoherent processing gain that is obtainable through incoherent averaging, in a tradeoff for increased resolution or signal parameterization.

In the context of this thesis, multidimensional signal processing can have several different meanings. It can refer to an image processing problem, where the input signal is a 2-dimensional function, and so is the output signal. Raster formatted image processing problems are really 2-D processing problems with 2-D outputs, but the serial raster format of the input and output allows a TSI optical system to treat the data as 1-D data streams with appropriate serial to parallel, and parallel to serial accumulation and temporary memory for partial products included within the processor. Multidimensional processing can also refer to a 1-dimensional input signal in a processing problem that produces an intrinsically two dimensional output space, such as the various time-frequency representations[4–7]. However these operations involve only one explicit integration over the input variable, and in the context considered here it is the dimensionality of the integration rather than the output dimensionality that I am referring to when I discuss multidimensional signal processing. Various subspace projections of the time-frequency representations can be performed with multidimensional integrations, as in the time averaged in-

stantaneous power spectrum analysis performed by a Bragg cell spectrum analyzer, but in this case the input and output functions are only one dimensional, and the subspace projection is really just an averaging operation. It is also possible to describe a very high time-bandwidth 1-dimensional signal processing operation in terms of an artificially constructed 2-dimensional output space, that can represent the desired long 1-dimensional output in a folded fashion, and more effectively use the space bandwidth capabilities of available optical devices[5.1–5.16]. In this case multiple integrations are required over the short time scale of the signal, and over the long term slow variations of the input signal, so even though a 1-dimensional signal processing operation is being performed it is being performed in a folded multidimensional space. The Synthetic Aperture Radar (SAR) processing problem can be viewed as either a long 1-D input or a raster scanned 2-D input, with a 2-D image output, and orthogonal integrations over the fast local variations (range) and slow long term variations (azimuth).

The beating of the temporal dependence of the space integrated signal against a family of time varying reference functions can be accomplished with two different methods, called additive and multiplicative. An additive system is constructed as an optical interferometer, with the space integrating signal transformation system included in one arm, and the generation of the family of reference functions needed for the time integrating inner product accumulations accomplished separately in the other arm. The multiplication between the signal and reference wavefronts is accomplished as a cross term in the interferometric detection of the recombined wavefronts. I will use s to represent a generic signal wavefront, and r to represent the reference wavefronts, with both coherent wavefronts functions of time and space, and mutual coherence will be assumed. The detected output intensity is given by the modulus squared of the sum of the incident signal and reference fields.

$$I = |s + r|^2 = |s|^2 + |r|^2 + 2\Re[sr^*] \tag{3.1}$$

The detected intensity contains the desired product which results in the beating between the signal and complex conjugate of the reference, thereby cancelling the temporal variations of the input signal at a particular reference location and allowing the production of a time integrated peak.

In a multiplicative approach to system implementation, the AODs operate as intensity modulators, by including a local oscillator reference with the signal, and the output of the first cell is multiplicatively rediffracted by the following cells. The multiplication of the space integrated signal with the reference AODs accomplishes the necessary heterodyning of the slow variations of the signal with the slow variations of the references.

$$I = |s+a|^2|r+b|^2 = |(s+a)(r+b)|^2 = \tag{3.2}$$
$$= (s^2+a^2)2\Re[rb^*]) + (r^2+b^2)2\Re[sa^*] + (s^2+a^2)(r^2+b^2)$$
$$+ 2\Re[sra^*b^*] + 2\Re[sr^*a^*b]$$

In this expression a and b are the local oscillator reference signals added to the signal and reference waveforms respectively, and in a sense they play the role of the beamsplitters of an interferometer. The final term is the desired beat between the signal and the complex conjugate of the reference. The first two terms will disappear if the signal and reference are zero mean signals. Both approaches introduce additional unwanted bias terms, both signal dependent and signal independent. The multiplicative approach has the distinct advantage that all of the optical components are in line, and if a component vibrates then all the wavefronts that pass through it are similarly affected, so that no additional modulations appear upon detection. This makes the multiplicative approach an inherently self referencing approach with intrinsic stability and suppression of unwanted modulations often associated with interferometers. On the other hand interferometers are notoriously unstable, with air currents and minute component vibrations introducing severe oscillations of the output fringe pattern, greatly decreasing the practicality, especially

in on board applications. However, there is a serious drawback associated with the in line multiplicative approach, and that is the low optical throughput caused by the multiplicative diffractions from the inefficient acousto-optic devices. When the intensity diffraction efficiency of the AODs is given by $\epsilon \approx .1$, then the light efficiency of a two cell multiplicative architecture is given by $\epsilon_1 \epsilon_2 \approx .01$, while the light efficiency of an additive interferometer is $\frac{1}{2}(\epsilon_1 + \epsilon_2) \approx .05$, which is significantly more light. When pulsed optical sources are utilized, as they are in most of the architectures considered in this thesis, then the extra light efficiency of the additive approach becomes very attractive. Another reason that makes the separated path approach to multidimensional TSI processing attractive is the decoupling of the optical design constraints that the in line architecture imposes on the various lens focal length ratios. A more fundamental reason to avoid the multiplicative approach arises in the consideration of complex signal processing operation, where an extra conjugate sideband is generated in the in line processors which is avoided completely in the interferometric approach. This term, which is the second to last term in Eq. 3.2 will produce additional spurious sidelobes, and additional considerations must be invoked to minimize these terms. Appropriate carrier encoding can be utilized to eliminate the unwanted conjugate sideband at the expense of using up some of the detector arrays space bandwidth, but since some sort of bias removal technique, such as carrier encoding, must be employed in any case this is not too severe of a restriction upon the multiplicative implementations. In order to avoid detailed consideration of this problem I will usually consider interferometers as the preferred implementation of a given TSI architecture, but it is usually possible to rearrange the system into a more stable in line multiplicative processor if coherent high power lasers are available.

3.1 Time and Space Integrating Image Processing

The linear transformation is the basis of all the signal processing tasks to be considered in this thesis, and it is the structure of the kernel that determines the classification and nature of a given task. The most general 2-D linear system is represented as a linear transformation from the two dimensional input space with variables x, y, to the two dimensional space of output variables $\hat{x}, \hat{y}$ through the space variant kernel h.

$$g(\hat{x}, \hat{y}) = \int \int_A f(x, y) h(x, y, \hat{x}, \hat{y}) dx dy \tag{3.1.1}$$

The kernel $h(x, y, \hat{x}, \hat{y})$ is a function of 4 independent variables, which exceeds the number of available dimensions (2 spatial and 1 temporal) that the optical system has available to internally represent information. Additionally, the number of degrees of freedom in the optical system is insufficient to represent all of the samples of this completely general kernel. For example, if each input and output variable has a space bandwidth product of 10^3, then the kernel contains 10^{12} independent analog samples, while a spatial mask can contain only about 10^6 samples, and a Bragg cell can only accommodate 10^3 samples at a time about 10^4 completely independent times during the period of temporal integration, far short of the required number. For these reasons it is not convenient to implement this general 2-D space variant system in either a purely space integrating or in a TSI architecture, but several important special cases can be accommodated within the framework of a TSI processing system. The special cases of 2-D linear transformations that can be implemented without using multiplexed devices, and obeying the physical restrictions inherent to a TSI system are the shift invariant (or convolutional) systems, the separable linear systems, and the cascadable linear systems. If we include multichannel Bragg cells, parallel laser diode arrays, or multiplexed holographic arrays we can broaden the class of problems computable with TSI techniques to include

certain space variant systems[8] by decreasing certain dimensionalities of the system or by choosing impulse response from among a finite library[9–12] of separable, or shift invariant, or scaled impulse responses.

For an image processing application using serially addressed AODs as the input light modulators, it is necessary to represent the image as a serial raster waveform, which is the standard method of transmitting video and graphics information. The video raster transforms the 500×500 video image $f(x,y)$ into a very long serial waveform $s(t)$ consisting of 500 image lines, with intervening synchronization signals. Ignoring the interleaving, which is often utilized to minimize flicker in commercial broadcast applications, the serial raster waveform is calculated from the image intensity distribution by the raster transformation.

$$s(t) = \sum_{n=1}^{N} f(v_s(t-nT), n\Delta y)\text{rect}\left[\frac{t-nT+\frac{T_0}{2}}{T_0}\right] \qquad (3.1.2)$$

In this equation the nth scan line is scanned over its full width X at a velocity $v_s = X/T_0$ from time $t = nT$ to $t = nT + T_0$, with $T > T_0$, and this scanning operation is repeated sequentially for each of the N lines in the image. Since the time bandwidth product (TB) of an AOD is typically 10^3, we can contain only one video line at a time within the aperture of an AOD. Typical video line times of $T_0 = 52\mu$sec, and $T = 63\mu$sec, fit well with available TeO_2 slow shear device apertures of $T_A = 68\mu$sec, although video bandwidths of 5 MHz under utilize the available 50 MHz AOD bandwidths. When the serial raster representation is mixed up to the transducer center frequency ω_0, and applied to an AOD travelling wave modulator, it propagates across the aperture, so that at sequential times $t = nT+t_0$ successive video lines are fully contained within the delay line length. At these times the laser can be pulsed synchronously with the raster scanning of the video image, with a very narrow pulse that effectively freezes the acoustic motion in the AOD.

$$a(t) = \sum_{n=1}^{N} \delta(t-nT-t_0) \qquad (3.1.3)$$

Each pulse will transduce successive video lines on the carrier into the optical system as spatial modulations of the coherent wavefront diffracted by the AOD. The transformation of the serial image raster into spatial modulations of successive laser pulses can be described as a space-time raster conversion.

$$\begin{aligned} s(x,n) &= \sum_{n=1}^{N} \delta(t - nT - t_0) s(t - x/v - A/2) e^{i\omega_0(t-x/v)} \text{rect}\,[x/A] \\ &= f(xv_s/v_a, nT) e^{i\omega_0(nT+t_0-x/v)} \end{aligned} \tag{3.1.4}$$

The position and size of the AOD aperture $A = Xv_s/v_a$ should be adjusted with respect to the timing of the laser diode pulses, t_0, so that exactly one video line is contained within the aperture without any synchronization signals on either side. The 2-D image is rescaled into the spatial coordinates of the optical system by the ratio of raster scan velocity to acoustic velocity, and the orthogonal dimension of the image raster is represented as variations of the wavefront diffracted on successive pulses. The carrier causes a spatial tilt of the coordinate system that can be ignored by aligning the lens with the diffracted wave, and the temporal variation from pulse to pulse can be canceled by making the carrier a harmonic of the video line rate or by interfering with the same frequency carrier from the reference AOD.

The nature of this transformation to a space-time raster suggests the processing strategy that should be utilized. First, each spatially modulated wavefront should be spatially transformed on each pulse, then the succession of pulses must be accumulated appropriately in order to perform the transformation over the y (or n) dimension of the raster scanned image. The time and space integrating technique is a natural consequence of the serial raster format, and its transformation into a space-time raster by an AOD illuminated by a pulsed source. The convenience with which the serial raster image data are entered into the AOD, formatted for processing by the travelling wave modulator, and formatted for output by the raster scan mechanism of the output detector array, allows the intervening TSI image processor

to act as a real time system, with a latency of only 1 image frame. Thus the image pours smoothly into the processor, and as soon as the it is contained within the processor, the desired computation has been completed, and the transformed image data can be poured smoothly out of the output detector array. It is this smooth flow of data into and out of the processor in a standard video format that makes the TSI image processing technique convenient and practical for real time image processing applications.

A generic interferometric time and space integrating optical image processing system is shown schematically in Figure 3.1.1. It consists of a space integrating multichannel 1-D linear transformation system, which is interfered with a time and space varying reference wavefront, and the resulting succession of interference patterns are time integrated on the detector array. The time and space varying reference wavefront can be generated by spatially transforming the multiplicative output from an optional point phase modulator and a pair of crossed Bragg cells, either of which is also optional. None of the systems examined in this thesis requires the full generality of this system, which some TSI systems in the literature require[13], because we include the extra flexibility of operating the detector array in a mode where it accumulates photogenerated charge in either a stationary or a scrolling pixel coordinate system.

The space integrating arm is described by an optically implemented multichannel linear transformation equation which is illuminated by the temporally modulated coherent laser source. The collimated optical wavefront with time varying intensity $a(t)$ is focussed in y into the signal AOD cell producing a narrow slit illumination at the Bragg angle, $\delta(y)e^{i2\pi x \sin\theta_b/\lambda}$, that is modulated by the travelling acoustic wave, $f(t - x/v)$, and this is spatially transformed in x by the possibly multichannel linear system $h(x, \hat{x}, \hat{y})$.

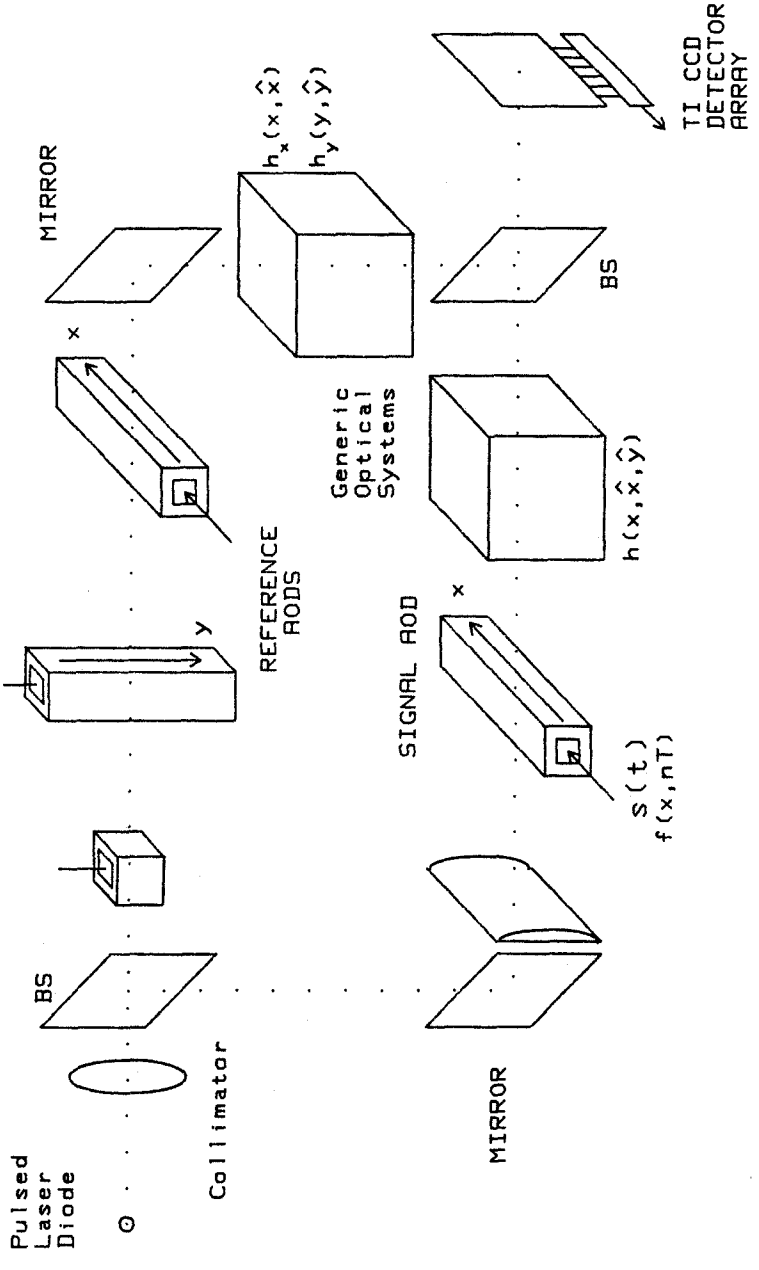


Figure 3.1.1. Generic TSI interferometer.

$$s(\hat{x},\hat{y},t) = a(t)\int s(t-x/v_a)h(x,\hat{x},\hat{y})dx \qquad (3.1.5a)$$

$$s(\hat{x},\hat{y},nT) = \int_A f(xv_s/v_a,nT)h(x,\hat{x},\hat{y})dx \qquad (3.1.5b)$$

In general it is very difficult to implement this fully general spatial transformation for arbitrary kernels h without a multiplexed hologram, and in the following discussion we will indicate the useful special cases of this general transformation that can be conveniently implemented, with conventional optical components. The two special cases of primary interest are the multichannel shift invariant 1-D correlator defined by the spatial transformation kernel $h(x-\hat{x},\hat{y})$, and the separable system defined by the kernel $h_x(x,\hat{x})h_y(0,\hat{y})$.

The time integrating reference arm of the interferometer consists of an optional phase modulator, and a pair of crossed Bragg cells.In its full generality the equations describing this reference waveform is quite complex and is not readily implementable for general kernels.

$$r(\hat{x},\hat{y},t) = a(t)\left[\int b(t-x/v)h_1(x,\hat{x})dx\right]\left[\int d(t-y/v)h_2(y,\hat{y})dy\right] \qquad (3.1.6)$$

The transformation of the AODs onto the output system is represented as independently transformed separable kernels, although in some cases slightly more general transformations are used. A slow parametric variation on these transformation kernels by the pulse number can be implemented by moving an optical component during the time of integration. When one or the other AOD is not included, then the processor will be assumed to produce a uniform profile in that dimension. There are three primary simplifications of this generalized reference wavefront that will be considered in detail in this thesis.

The first important special case that is of interest is when we eliminate the AOD in the reference arm that is parallel to the signal AOD, which results in orthogonally

propagating signal and reference acoustic waves, allowing us to process information in both the x and y dimensions.

$$r_2(\hat{x},\hat{y},t) = a(t)p(\hat{x};n)\int d(t-y/v)h_2(y,\hat{y};n)dy \tag{3.1.7}$$

This equation includes a $\hat{x}$ dependent spatial profile, for instance a focusing wavefront produced by a cylindrical wavefront, that can be slowly varied from pulse to pulse, by moving the lens during the integration period. Similarly the transformation kernel h_2 is shown as being parametrically dependent on the pulse number, but this extra generality is rarely used.

Another special case is when the orthogonal reference AOD is not needed, because the orthogonal multichannel time integration is performed in a y dimensional moving coordinate system on the detector. In this case the reference wavefront is given by a similar expression with a colinearly propagating reference AOD.

$$r_1(\hat{x},\hat{y},t) = a(t)q(\hat{y};n)\int b(t-x/v)h_1(x,\hat{x};n)dx \tag{3.1.8}$$

Only certain classes of kernels h_1 and h_2 can be implemented for either of these spatial transformation of the reference acoustic signal.

Finally, the simplest interferometric system eliminates both crossed AODs in the reference arm, so the reference wavefront is simply phase modulated in time, and may have a slow parametric change in angle of incidence or curvature with pulse number.

$$r_0(\hat{x},\hat{y},t) = a(t)h_0(\hat{x},\hat{y};n) \tag{3.1.9}$$

In order to perform a useful two dimensional signal processing operation with this type of reference wavefront it is again necessary to time integrate in a coordinate system on the detector that is moving in the y dimension, orthogonal to the space integration. They kernels h_0, h_1 and h_2 may have a slow variation of a simple parameter with the pulse number n, as may the orthogonal spatial wavefronts $p(\hat{x})$

and $q(\hat{y})$ but the systems analyzed in this thesis will used fixed components so this slow variation will be ignored.

There are three types of transformation kernels between the reference Bragg cells and the output detector which are commonly utilized, but more complicated transformations can be performed as well by including generalized optical elements, such as tilted lenses or holograms. The simplest spatial kernel is just an imaging transformation of the acoustic field with some magnification, while smearing in the orthogonal dimension in order to form multiple copies of the diffracted image, and this can be performed on either a horizontal or vertical reference AOD.

$$h_r(x,\hat{x},y,\hat{y}) = \begin{cases} \delta(x-m\hat{x})e^{i2\pi y\hat{y}} \\ \delta(y-m\hat{y})e^{i2\pi x\hat{x}} \end{cases} \tag{3.1.10}$$

Fourier transforming the Bragg cell onto the output detector array, while smearing the orthogonal dimension is simply accomplished with a spherical lens that performs a 2-D Fourier transformation of the diffracted wavefront illuminated by a cylindrically focussed wave.

$$h_r(x,\hat{x},y,\hat{y}) = e^{i2\pi x\hat{x}}e^{i2\pi y\hat{y}} \tag{3.1.11}$$

A more general transformation that is also used is the chirp transformation which produces wavefronts with particular curvatures in order to produce chirping fringe profiles.

$$h_r(x,\hat{x},y,\hat{y}) = \begin{cases} e^{i\pi b(x-\hat{x})^2}e^{i2\pi y\hat{y}} \\ e^{i2\pi x\hat{x}}e^{i\pi b(y-\hat{y})^2} \end{cases} \tag{3.1.12}$$

It is possible to implement more general reference wavefront transformations, and one example is the space variant kernel required for the SAR processor discussed in section 3.3. In that case the reference kernel is a quadratically curved phase front in y, where the curvature is inversely proportional to the $\hat{x}$ coordinate, as can be obtained with a tilted cylindrical lens. It is a fortuitous circumstance of the SAR processing problem that the space variant convolutional kernel can be obtained with a simple scale change of a conveniently produced optical modulation.

The crossed Bragg cell processor based on the reference wavefront given in Equation 3.1.7 uses a time integrating CCD to perform the multidimensional transformation. The interferometric output of the processor produces a heterodyne multiplication between the signal and reference arms, as well as some interferometric bias terms. The bias terms are necessary to represent a bipolar or complex signal on a detector array that accumulates charge proportional to the incident optical flux, which is a purely unipolar quantity. The output intensity as a function of space and time that is produced on the detector array will contain the product between the transformed signal and one form of the transformed reference.

$$\begin{aligned} I(\hat{x},\hat{y},t) &= a(t)\left|s(\hat{x},\hat{y},t)+r_2(\hat{x},\hat{y},t)\right|^2 \\ &= a(t)2\Re\left(\left[\int s(t-x/v)\delta(y)h(x,\hat{x},\hat{y})dx\right]^*\int d(t-y/v)h_2(y,\hat{y};n)dy\right) \\ &\quad +\left|s(\hat{x},\hat{y},t)\right|^2+\left|r_2(\hat{x},\hat{y},t)\right|^2 \end{aligned} \tag{3.1.13}$$

The temporal accumulation of this optical intensity profile on a two dimensional matrix of photosensors over the full sequence of N optical pulses completes the required processing. However, the unwanted bias terms must be removed in the electronic post processing, and for this reason the signal and reference beams are recombined at an angle in order to introduce an interferometric spatial carrier.

$$\begin{aligned} I(\hat{x},\hat{y}) &= \int_{NT} a(t)\left|s(\hat{x},\hat{y},t)+r_2(\hat{x},\hat{y},t)e^{i2\pi(\alpha\hat{x}+\beta\hat{y})}\right|^2 dt \\ &= \int_{NT} a(t)\Bigg[\left|s(\hat{x},\hat{y},t)\right|^2+\left|r_2(\hat{x},\hat{y},t)\right|^2 \\ &\quad 2\Re\left(\left[\int s(t-x/v)h(x,\hat{x},\hat{y})dx\right]^*\int d(t-y/v)h_2(y,\hat{y})dy e^{i2\pi(\alpha\hat{x}+\beta\hat{y})}\right)dt\Bigg] \\ &= \sum_{n=1}^{N} 2\Re\Bigg[\left(\int f(x',nT)\delta(y)h(x',\hat{x},\hat{y})dx'\right)^* a(nT) \\ &\quad \int d(y',nT)h_2(y',\hat{y})dy' e^{i2\pi(\alpha\hat{x}+\beta\hat{y})}\Bigg]+\left|s(\hat{x},\hat{y},nT)\right|^2+\left|r_2(\hat{x},\hat{y},nT)\right|^2 \end{aligned} \tag{3.1.14}$$

The pulsing of the laser diode effectively freezes the acoustic motion of each raster

line of both the input and the reference, transforming the long sequential rasters into a space time raster representation within the optical system. Thus we can represent the processing operation as a sequential accumulation of the product of orthogonal space time rasters which have been spatially transformed by the optical system. The spatial carrier is usually chosen to be only in one direction either horizontally or vertically, so that either $\alpha = 0, \beta \neq 0$, or $\beta = 0, \alpha \neq 0$. This equation encompasses most two dimensional TSI image processing operations that can be performed with a stationary detector array, but it is so general it is difficult to see how any particular operation can be performed. In the subsequent sections I will illustrate particular implementations for special cases of this general equation.

The time and space varying optical intensity profile of Eq. 3.1.13 can alternatively be time integrated in a moving detector coordinate frame, where the motion is a rigid body motion of the photosensor matrix described by a change of variables to a coordinate system that is stationary with the accumulating pixels.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta(t) & \sin\theta(t) \\ -\sin\theta(t) & \cos\theta(t) \end{bmatrix} \begin{bmatrix} \hat{x} + x_0 \\ \hat{y} + y_0 \end{bmatrix} + \begin{bmatrix} w(t) \\ z(t) \end{bmatrix} \tag{3.1.15}$$

This describes a rotation to a time varying angle $\theta(t)$ about a coordinate center (x_0, y_0) in the $\hat{x}, \hat{y}$ plane, with an auxiliary linear motion $(w(t), z(t))$. The full generality of this rigid coordinate rotation is not used very often except an extremely clever implementation suggested by Haney[39], to compensate for higher order range migration effects in a synthetic aperture radar processor. In this thesis I will examine a much simpler class of time integrating processors where the detector is either fixed, or operating in a uniform one dimensional scrolling mode, orthogonal to the space integration, which is referred to as the time delay and integrate (TDI) mode of operation.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \hat{x} \\ \hat{y} + [Vt] \end{bmatrix} \tag{3.1.16}$$

The scrolling detector can be implemented by mechanically scanning a piece of film

in the output focal plane of the processor, but a much more practical and reliable approach is to utilize a self scanning solid state CCD detector array. The CCD shifts by one pixel of width Δy in the y direction upon the application of clock pulses with period T forming an effective velocity $V = \Delta y/T$ in a discretely stepping fashion, represented by the greatest integer notation [Vt] in Eq. 3.1.16. In general when this scrolling detector mode is utilized, then an AOD is not needed in the y dimension, so an x propagating reference AOD will be assumed, and the reference wavefront transformation given in Equation 3.1.8 will be used. On a scrolling detector array the window of time integration for an individual pixel begins when the pixel enters the top of the CCD array and continues as the pixel propagates down the array until it leaves the bottom of the array where it is read out.

$$
\begin{aligned}
s(\hat{x}, y'_{t=mT}) &= \sum_{n=m-M}^{m} \int \left[\int_{(n-\frac{1}{2})T}^{(n+\frac{1}{2})T} I(\hat{x}, \hat{y}, t)\, dt \right] \delta(\hat{y} - y' - n\Delta y)\, d\hat{y} \qquad (3.1.17) \\
&= \sum_{n=m-M}^{m} \left| s(\hat{x}, y' + n\Delta y, nT) + r(\hat{x}, y' + n\Delta y, nT) e^{i2\pi(\alpha\hat{x} + \beta(y' + n\Delta y))} \right|^2 \\
&= \sum_{n=m-M}^{m} 2\Re \left[\left(\int f(x', nT) h(x', \hat{x}, y' + n\Delta y)\, dx' \right)^* \right. \\
&\qquad \left. a(nT) \int d(x', nT) h_1(x', \hat{x})\, dx' e^{i2\pi(\alpha\hat{x} + \beta(y' + n\Delta y))} \right] \\
&\quad + \left| \sum_{n=m-M}^{m} s(\hat{x}, y' + n\Delta y, nT) \right|^2 + \left| \sum_{n=m-M}^{m} r(\hat{x}, y' + n\Delta y, nT) \right|^2
\end{aligned}
$$

This equation represents the most general transformation that can be performed on an input raster image in an interferometric TSI processing system that utilizes a scrolling detector array. The resulting two dimensional linearly transformed output image scrolls out of the bottom of the detector array, and is converted back into a conventional raster by the high speed parallel to serial CCD horizontal raster read out register. This equation contains the desired transformed image on a spatial carrier, plus a nonuniform signal dependent bias and a reference beam bias. If the spatial carrier is in the horizontal dimension, ($\alpha \neq 0, \beta = 0$), then it can be tuned independently from the processing, by rotating the recombining beamsplitter.

However, when the spatial carrier is in the vertical dimension, ($\alpha = 0, \beta \neq 0$), then it plays a role in the time integration in the detector moving coordinate system, and varying the spatial carrier frequency will effect the processing operation. To produce a peak on a scrolling detector array illuminated by a travelling fringe interference profile, it is necessary that the velocity of the fringe pattern is equal to the detector shifting velocity. The fringe velocity is given by the spatial period divided by the temporal oscillation period, $v_f = \Lambda/\tau_0 = f_0/\beta$, so the spatial frequency along the scrolling direction plays an important role in the processing operation.

3.1.1 Shift invariant 2-D linear system

The simplest and most widely used two dimensional linear system is the space invariant linear system, with the correlation being the most common signal processing operation. The kernel becomes 2-dimensional, with an argument that only depends on the relative coordinate positions in the input and output spaces, $h(x, y, \hat{x}, \hat{y}) = h(x - \hat{x}, y - \hat{y})$ for convolutions, and $h(x, y, \hat{x}, \hat{y}) = h^*(x + \hat{x}, y + \hat{y})$ for correlations. This type of correlator system can be implemented with a multichannel space integrating correlator followed by a shift and add detector that accumulates the partial correlations in the appropriate format necessary to form a full 2-D correlation[14–16]. The operation of the system is described with reference to Figure 3.1.2, which shows the pulsed laser diode illuminating a multichannel 1-D Vander Lugt correlator, the output of which is imaged onto a scrolling CCD detector array. The correlator contains a holographic filter that is composed of a vertically multiplexed array of 1-D Fourier transforms of horizontal slices through the reference image. Thus, with each pulse of the laser diode the field incident on the CCD due to the space integration is the multichannel correlation with the bank of reference filters.

$$s(\hat{x}, \hat{y}, n) = \int f(x, nT) h^*(x + \hat{x}, \hat{y}) dx \tag{3.1.18}$$

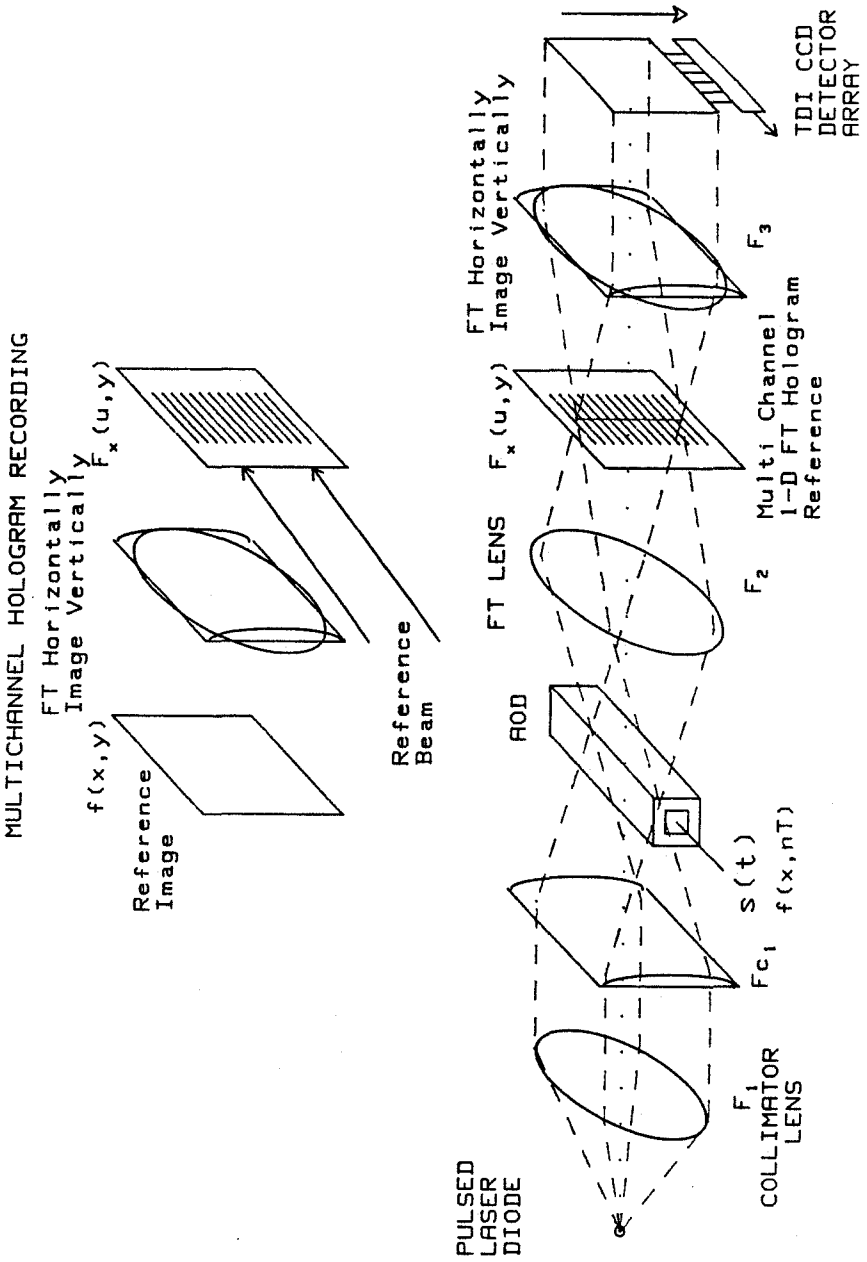


Figure 3.1.2. a) Forming the multichannel 1-D Vander Lugt Hologram, b) the TSI image correlator.

If we apply a video signal to the AOD that matches the image stored in the multichannel reference hologram, a 1-D correlation peak will be generated at the top of the CCD on the first laser pulse since the first line of the video raster matches the top slice stored in the hologram. On the nth laser pulse a 1-D correlation peak will appear in the nth row of the CCD because the nth video raster line matches the nth slice stored in the hologram. These 1-D correlation peaks will appear to travel linearly down a column of the CCD from pulse to pulse, and to perform the 2-D correlation they must be accumulated in this travelling coordinate frame in order to complete the processing. To accomplish this the multichannel correlation pattern detected on each pulse is vertically translated by 1 pixel, and summed with that detected on the previous pulses. This system computes a nonlinear type of correlation integral, because the detector is sensitive to intensity, which is the modulus squared of the 1-D correlation peaks computed on each pulse.

$$I(\hat{x}, y'_{t=mT}) = \sum_{n=M-m}^{m} \int \delta(\hat{y} - y' - n\Delta y) \left| \int f(x, nT) h^*(x + \hat{x}, \hat{y}) dx \right|^2$$
$$= \sum_{n=M-m}^{m} \left| \int f(x, nT) h^*(x + \hat{x}, y' + n\Delta y) dx \right|^2 \qquad (3.1.19)$$

This type of nonlinear system also has recognition capabilities that can be superior to that of a linear correlator in certain types of operating conditions[15].

The linear correlation can also be computed by using a uniform interferometric reference wave, doppler shifted by an AOD to the same center frequency as the signal AOD, and tilted to provide a spatial carrier for bias demodulation.

$$I(\hat{x}, y'_{t=mT}) = \sum_{n=M-m}^{m} \int \delta(\hat{y} - y' - n\Delta y) \left[2\Re \left(r^* e^{i2\pi[\alpha x + \beta y]} \int f(x, nT) h^*(x + \hat{x}, \hat{y}) dx \right) \right.$$
$$\left. + |r|^2 + \left| \int f(x, nT) h^*(x + \hat{x}, \hat{y}) dx \right|^2 \right] d\hat{y} \qquad (3.1.20)$$

For an input image which is matched to that of the filter a single correlation peak will be generated, modulated by a spatial carrier and riding on the signal dependent nonlinear correlation, all residing on a relatively uniform bias plateau. The

appropriate carrier demodulation can separate the interferometric term from the bias terms leaving a discrete representation of the desired correlation integral, that pours out of the bottom of the scrolling detector array in a conventional raster format.

$$I(\hat{x}, y'_{t=mT}) = \sum_{n=M-m}^{m} 2r \left[\int f(x, nT) h^*(x + \hat{x}, y' + n\Delta y) dx \right] \qquad (3.1.21)$$

Where the CCD has M vertical pixels in which the sliding window integration is performed, and the output correlation appears as a raster waveform at the bottom of the CCD from time $t = 0$ to $t = M + N$. The system has effectively computed the correlation (or convolution) of the input signal with the image stored in the multichannel reference hologram in real time. For a video frame rate of 30 frames per second the computational power of such a signal processing module in raw analog multiplies per second is approximately $30N^4 = 30(500)^4 \approx 2 \times 10^{12}$. An equivalent digital processor would utilized an FFT based processing algorithm, and consequently would only need to perform $30 \times 2N^2 \log N^2 \approx 3 \times 10^8$ multiplies per second, which represents a significant computational capability. An even higher computation rate should be achievable as the detector technology matures, with a potential of correlating 1000×1000 images at a frame rate of greater than 100Hz.

3.1.2 Separable 2-D linear system

The separable linear system is described by a kernel that can be broken up into the product of systems operating in the x and y dimensions independently, so the kernel function is given by $h(x, y, \hat{x}, \hat{y}) = h_x(x, \hat{x}) h_y(y, \hat{y})$. Each kernel function is simply a two dimensional representation with 10^6 independent analog entries, so that it can be represented within the available temporal or spatial degrees of freedom in a TSI optical system. In this case the integration over either one of the kernel functions can be done first, completely independently of the other kernel function, producing an intermediate data representation that must then be linearly

transformed along the orthogonal dimension in an identical fashion for each channel of the system.

$$g(\hat{x},\hat{y}) = \int_Y \left[\int_X f(x,y)h_x(x,\hat{x})dx\right] h_y(y,\hat{y})dy = \int_Y f_x(\hat{x},y)h_y(y,\hat{y})dy \quad (3.1.22)$$
$$= \int_X \left[\int_Y f(x,y)h_y(y,\hat{y})dy\right] h_x(x,\hat{x})dx = \int_X f_y(x,\hat{y})h_x(x,\hat{x})dx$$

An alternate interpretation of this type of system for the case of discrete data is as a triple matrix product processor[18], where the two dimensional input and output functions, and the two parts of the separable kernel, are interpreted as matrices. In this case the processing operation performed by the system can be written as a triple matrix product.

$$[G] = [B][F][D] \quad (3.1.23)$$
$$g_{ij} = \sum_{k=1}^{K} b_{ik}\left(\sum_{l=1}^{L} f_{kl}d_{lj}\right) \qquad i = 1,\ldots,I \qquad j = 1,\ldots,J$$
$$= \sum_{l=1}^{L}\left(\sum_{k=1}^{K} b_{ik}f_{kl}\right) d_{lj}$$

The input K by L matrix $[F]$ is right multiplied by the L by J matrix $[D]$ and left multiplied by the I by K matrix $[B]$, in order to produce the I by J output matrix $[G]$.

A time and space integrating interferometer that can perform the separable kernel image processing operation[17], or the triple matrix product, is shown in Figure 3.1.3. It consists of a space integrating matrix vector multiplier[2.5–2.8] in one arm, and an orthogonal outer product processor in the other arm, which interfere on the detector array in order to produce a time integrated output. The space integrating matrix vector multiplication is accomplished by first imaging the signal AOD onto the mask in the x dimension, while smearing the light uniformly in the orthogonal dimension. Each pixel in the AOD multiplies every element in a vertical row of the mask on each laser pulse, then the inner products between the entire AOD and the horizontal columns of the mask are spatially accumulated in parallel on

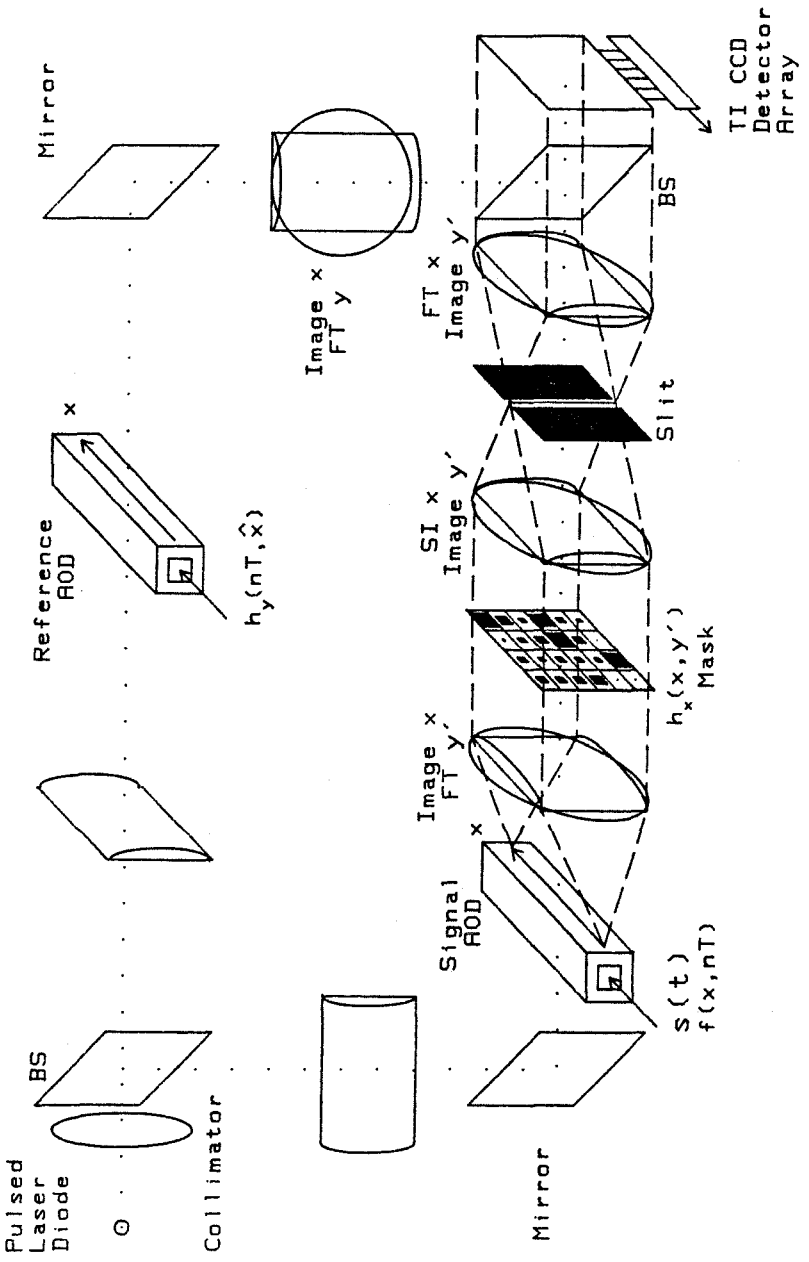


Figure 3.1.3. Interferometric TSI separable kernel processor, or triple matrix multiplier.

every pulse producing the matrix vector multiplication. The spatial accumulation of the inner products is accomplished with a horizontal Fourier transformation system that vertically images the parallel array of inner products. Thus on each pulse, before the sampling slit we have the Fourier transforms of the products of the mask horizontal columns with the information diffracted by the Bragg cell on that pulse. The desired inner products are formed by sampling the 0th moment of these transforms with a vertical slit whose width is less than the scaled bandwidth B_x of the image and reference.

$$g(\hat{y}, nT) = \int f(x, nT) h_x(x, \hat{y}) e^{i2\pi xu} dx \Big|_{u<B_x} \tag{3.1.24}$$

The slit limits the light efficiency of this system, and so should be made as wide as possible, while still narrow enough to produce diffracted light that is essentially uniform over the horizontal aperture of the detector array after it is recollimated horizontally by the following lens system. With these restrictions the light incident on the detector array on each pulse from the signal arm of the interferometer is uniform horizontally, and represents a family of inner products vertically.

$$\begin{aligned} g(\hat{x}, \hat{y}, nT) &= \int \left[\int f(x, nT) h_x(x, \hat{y}) e^{i2\pi xu} dx \right] \text{rect}\left[\frac{u}{B}\right] e^{i2\pi u\hat{x}} du \\ &\approx \int f(x, nT) h_x(x, \hat{y}) dx \end{aligned} \tag{3.1.25}$$

The other arm of the interferometer is used to sequentially enter raster lines of the orthogonal transformation kernel, producing a reference space-time raster, which is imaged horizontally and uniformly smeared vertically.

$$r(\hat{x}, \hat{y}, nT) = h_y(nT, \hat{x}) \tag{3.1.26}$$

The interferometric product between the signal and reference arms must be time integrated on a stationary detector in order to complete the 2-D processing operation.

$$I(\hat{x}, \hat{y}) = \sum_{n=1}^{N} |g(\hat{x}, \hat{y}, nT) + r(\hat{x}, \hat{y}, nT) e^{i2\pi[\alpha x + \beta y]}|^2 \tag{3.1.27}$$

$$= \sum_{n=1}^{N} 2\Re\left[\int f(x,nT)h_x(x,\hat{y})dx \quad h_y^*(nT,\hat{x})e^{-i2\pi[\alpha x+\beta y]}\right]$$
$$+ \sum_{n=1}^{N} |\int f(x,nT)h_x(x,\hat{y})dx|^2 + \sum_{n=1}^{N} |h_y(nT,\hat{x})|^2$$

The output consists of the separably transformed image modulated by the spatial carrier, which is riding on the signal dependent bias, plus a bias due to the kernel applied to the reference AOD. The appropriate carrier demodulation circuitry can be used to separate the desired linear transformation from the bias. The effective computational power of this architecture is much larger than the shift invariant system, because for arbitrary kernels an accelerated digital algorithm may not exist. Thus for 30 frames per second this system can compute approximately $30N^4 = 30(500)^4 \approx 2 \times 10^{12}$ analog multiplies per second, which is a computation rate that is unobtainable with other technologies.

A special case of the separable kernel linear transformation is the 2-D image Fourier transform.

$$h_x(x,\hat{x})h_y(y,\hat{y}) = e^{i2\pi x\hat{x}}e^{i2\pi y\hat{y}} \tag{3.1.28}$$

In this case the separable kernel functions are each space variant, and could be computed with the general purpose processor described above, but it is more fruitful to design a special purpose processor that makes use of the ability of a lens to perform a Fourier transformation[13,16,19,20]. A TSI interferometer that can perform image Fourier transformations is shown in Figure 3.1.4, but many other topologies are also possible, several of which are discussed in a different context in Chapter 5, where they are referred to as folded spectrum processors. The signal arm is just a Bragg cell spectrum analyzer that spatially Fourier transforms in x each horizontal line of the video raster, smearing and collimating vertically at the same time. A synchronous raster representation of a discrete Fourier transform (DFT) matrix is applied to the orthogonal reference AOD, which is imaged onto the CCD in y, and uniformly smeared and collimated in x, in order to interfere with the signal beam.

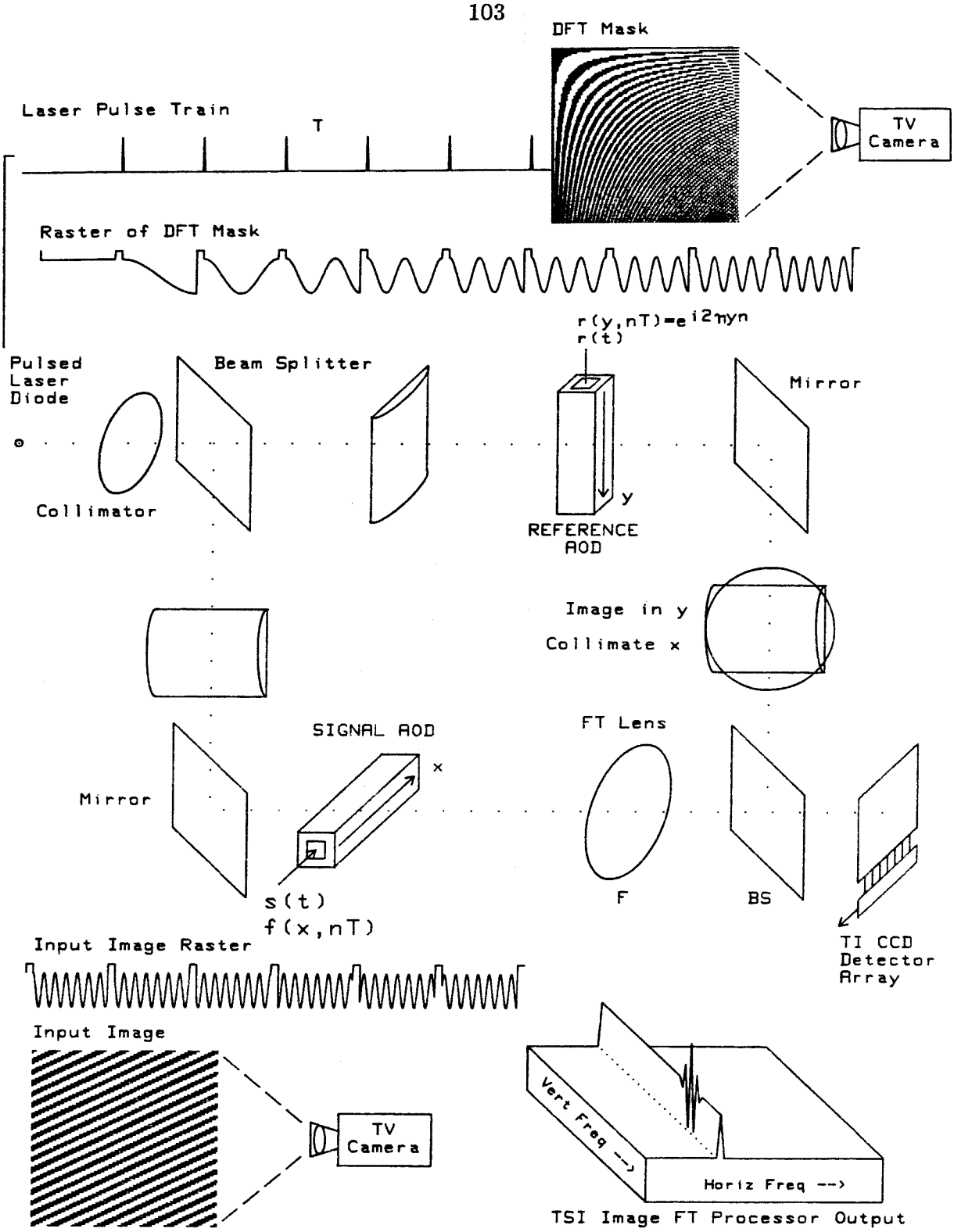


Figure 3.1.4. Image Fourier transformation TSI interferometer, showing the DFT mask that is entered into the reference AOD, a single spatial frequency input image, and the processor output.

The operation computed by this system is the desired 2-D Fourier transform of the raster image, plus the usual interferometric bias terms associated with a TSI processing system.

$$\begin{aligned} I(\hat{x},\hat{y}) &= \int_{NT} a(t) \left| \int_A s(t+x/v) e^{i2\pi x\hat{x}/\lambda F} dx + \int r(t+y/v) e^{i2\pi[\alpha x+\beta y]} \delta(y-\hat{y}) dy \right|^2 \\ &= \sum_{n=1}^{N} \left| \int_A f(x,nT) e^{i2\pi x\hat{x}/\lambda F} dx + re^{i2\pi n\hat{y}} e^{i2\pi[\alpha x+\beta y]} \right|^2 \\ &= \sum_{n=1}^{N} 2\Re \left[\int_A f(x,nT) e^{i2\pi x\hat{x}/\lambda F} dx \quad re^{i2\pi n\hat{y}} e^{i2\pi[\alpha x+\beta y]} \right] \\ &\quad + \left| \sum_{n=1}^{N} \int_A f(x,nT) e^{i2\pi x\hat{x}/\lambda F} dx \right|^2 + N|r|^2 \end{aligned} \tag{3.1.29}$$

The spatial carrier can be placed in either dimension, whichever is the most convenient, or it can be eliminated entirely and an auxiliary detector can be used to compute the bias, which can then be subtracted, as long as only the real part of the image is to be computed. When an image consisting of a single spatial frequency fringe pattern is applied to the processor, the output will consist of a uniform bias, a signal dependent bias ridge at the position corresponding to the horizontal spatial frequency, with a carrier modulated impulse at the position along the ridge that corresponds to the vertical spatial frequency of the image. Demodulation of the interferometric term on the spatial carrier will produce the desired image Fourier transformation.

Another separable case that can be accommodated in a TSI system is when the dimensionality of the output space is much smaller than the input space, as in the generation of image moments. In this case the output space is a discrete set of about 100 image inner products indexed by the variables p and q, and described by the moment generating equation.

$$m_{pq} = \int \int_A x^p y^q f(x,y) dx dy \tag{3.1.30}$$

In this case the separable kernel is given by $h(x,y;p,q) = x^p y^q$, and simple topologies

of a moment generating optical processor can be constructed using purely space integrating techniques, or with slightly more complicated TSI architectures[21].

3.1.3 Cascadable 2-D linear system

In a cascadable system the kernel is almost separable, but there is a parametric dependence of the kernel h_2 on the output variable of the h_1 linear transformation, so that $h(x,y,\hat{x},\hat{y}) = h_1(x,\hat{x})h_2(y,\hat{y},\hat{x})$. The integration must be performed first over the h_1 kernel, producing an intermediate data set that has been spatially channelized in $\hat{x}$. This must subsequently be linearly transformed in the orthogonal dimension over h_2, where the transformation kernel parametrically depends on the output coordinate $\hat{x}$of the h_1 kernel.

$$g(\hat{x},\hat{y}) = \int \left[\int f(x,y)h_1(x,\hat{x})dx\right] h_2(y,\hat{y};\hat{x})dy = \int f_x(\hat{x},y)h_2(y,\hat{y};\hat{x})dy \quad (3.1.31)$$

It is actually quite involved to perform this general transformation, although it might be possible with multiplexed holograms or multichannel Bragg cells, but some special cases can be conveniently implemented.

One special case is when the second linear transformation is shift invariant.The kernel functions for each channel produced by the first transformation can be arbitrarily dependent on the output coordinate $\hat{x}$, so that $h_2(y,\hat{y};\hat{x}) = h_2(y-\hat{y};\hat{x})$This type of system is implementable because the transformation kernels are both two dimensional functions that can be represented with 2-D spatial transparencies.A multiplicative system that performs this processing operation is shown in Figure 3.1.5. The operation of this system is as a space integrating matrix vector multiplier of the output of the signal AOD when it is illuminated by the pulsed source.

$$a(\hat{x},nT) = \int f(y,nT)h_1(y,\hat{x})dy \quad (3.1.32)$$

The sequence of output components produced by the matrix vector multiplication on successive laser diode pulses addresses a multichannel time integrating convolver,

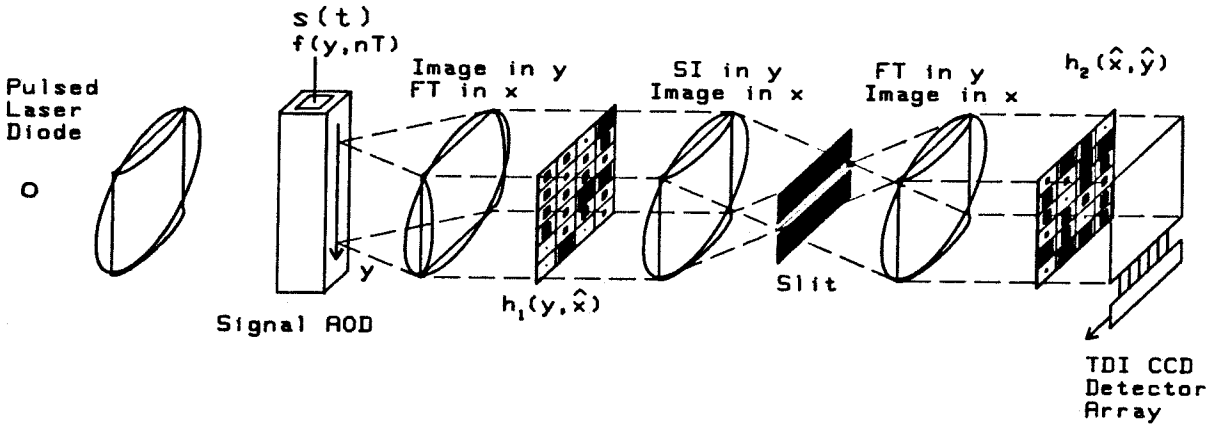


Figure 3.1.5. TSI system for performing cascadable linear transformations on an image.

so that each output from the matrix vector multiplier is convolved (or correlated) against its own reference function in parallel columns of a scrolling detector array.

$$g(\hat{x}, y'_{t=mT}) = \sum_{n=M-m}^{m} \int \delta(\hat{y} - y' - n\Delta y) \left[\int f(y, nT) h_1(y, \hat{x}) dy \right] h_2(\hat{x}, \hat{y}) d\hat{y}$$
$$= \sum_{n=M-m}^{m} \left[\int f(y, nT) h_1(y, \hat{x}) dy \right] h_2(\hat{x}, y' - n\Delta y) \qquad (3.1.33)$$

The reference functions can be stored as parallel columns of a reference mask as illustrated, or they can be produced interferometrically from a multichannel 1-D hologram in the reference arm of an interferometer, the diffracted component of which is imaged onto the scrolling output detector. An alternative topology that can be used to implement this transformation is a bank of 1-D holographic correlators like that used in the image correlator, which is interfered with a space time raster reference like in the separable processor, and then time integrated on a detector.

A somewhat more involved transformation is possible when the second kernel can be decomposed into the product of a shift invariant kernel with a shift variant kernel, $h_2(y, \hat{y}; \hat{x}) = h_A(y - \hat{y}; \hat{x}) h_B(y, \hat{x})$. Once again this is possible because both parts of this decomposed kernel are two dimensional functions, part A will be represented as a spatial mask $h_A(\hat{x}, \hat{y})$, and part B will be represented as a space time raster input through an auxiliary Bragg cell in the reference arm of an interferometer $h_B(nT, \hat{x})$.

Another special case is when the parametric dependence of the channel coordinate $\hat{x}$ on the kernel function is as a scale factor only, in which case $h_2(y, \hat{y}; \hat{x}) = h_2(y, \hat{y} \times \hat{x})$, or $h_2(y, \hat{y}; \hat{x}) = h_2(y \times \hat{x}, \hat{y})$. This system can be implemented with the interferometer used for separable system with a slight modification in which the reference AOD is imaged with different magnifications onto each row of the CCD.

The synthetic aperture radar (SAR) problem very conveniently has both of these properties, as well as shift invariance of the first kernel, so that $h(x, y, \hat{x}, \hat{y}) \approx h_1(x - \hat{x}) h_2((y - \hat{y})/\hat{x})$. In fact in SAR the impulse response of the shift invariant

kernels are complex quadratic phase functions, which are easily implemented in an optical system, because it is the kernel of free space for spatial integration, and a chirping interference pattern can be generated with a lens as the spatial impulse response for a scrolling time integrating correlation. The SAR application will be discussed further in section 3.3.

3.2 Time and Space Integrating Processing of Long 1-D Signals

An alternative interpretation of a TSI image processing system can be arrived at by considering a raster transformation of both input and output dimensions. The space variant linear system of Eq. 3.1.1 can be represented in discrete notation as a rank 4 tensor transformation of a rank 2 matrix producing an output rank 2 matrix.

$$G(i,j) = \sum_{m=1}^{M} \sum_{n=1}^{N} F(m,n) H(m,n,i,j) \tag{3.2.1}$$

In this equation we can associate a particular element of the input matrix $F(m_0, n_0)$, (or pixel of the input image), with a space variant impulse response matrix (or image) of the transformation tensor $H(m_0, n_0, i, j)$. To transform this equation into a very long 1-D system we unfold these tensors with the following discrete raster transformation, sometimes called the stacking operator, without intervening gaps between the rows.

$$\begin{aligned} m + nM &\rightarrow k \qquad k = 1, \ldots, NM \\ i + jI &\rightarrow l \qquad l = 1, \ldots, IJ \end{aligned} \tag{3.2.2}$$

This transformation turns the 2-dimensional space variant linear system into a huge matrix-vector multiplication.

$$g(k) = \sum_{k=1}^{MN} H(k,l) f(l) \tag{3.2.3}$$

In this equation we can associate a particular element of the input vector $f(l_0)$ with its corresponding row in the transformation matrix $H(k, l_0)$. Since the input and output vectors are of the same dimensionality as an image, but folded up into a long vector, $MN \approx IJ \approx 10^6$, then the matrix is a huge transformation matrix of size $MN \times IJ \approx 10^6 \times 10^6$. Although the optical system has enough dimensions to represent a matrix it does not have the resolution, or enough internal degrees of freedom, to represent this huge transformation matrix, so we can not implement this equation upon our folded vectors for arbitrary transformation matrices.

The corresponding continuous representation of the space variant linear system is given by an integral transformation.

$$g(\tau) = \int f(t)h(t,\tau)dt \tag{3.2.4}$$

This general 1-D transformation is beyond the capabilities of a TSI optical system for general transformation kernels of dimensionality $10^6 \times 10^6$, but two special cases are implementable. The first is when the transformation is shift invariant, as in the correlation integral, because in this case there are only $NM \approx 10^6$ intrinsic degrees of freedom in the transformation kernel.

$$g(\tau) = \int f(t)h^*(t+\tau)dt \tag{3.2.5}$$

The kernel $h(\tau)$ is the impulse response for all the input delays t, and its 10^6 independent samples can be represented with a $10^3 \times 10^3$ spatial mask. In section 3.2.1, the techniques for folding a long correlation into the space and time representation of a TSI processing system will be explored.

The other type of 1-D processing system that can be incorporated within the constraints of a folded TSI processing system is when the space variant transformation kernel has an algorithmically decomposable substructure that can be implemented as a separable sequence of sub transformations. The structure of this separated transformation kernel implies that there is a short range and long range decomposability to the transformation. If we consider the transformation matrix to be broken up into blocks of size $M \times I$, and index these blocks by indices n and j, then we can write any element as the product of a short range function of m and i with a long range function of m and n. This factorization of the kernel makes it only depend on the $MI \approx 10^6$ degrees of freedom of the short range function, and the $NJ \approx 10^6$ degrees of freedom of the long range function. Thus there are enough degrees of freedom within the TSI optical system to represent this type of factored

kernel, where the short range (i.e. high bandwidth) terms will be represented spatially, and the long range (i.e. low bandwidth) factor will be represented as a space time raster reference used for time integration. The decomposed transformation allows a TSI implementation as a sequential computation with a succession of space integrated transformations being accumulated by time integration. This is the basis of the folded spectrum operation which is the main topic investigated in this thesis, and reported in depth in Chapter 5.

3.2.1 Folded TSI correlator for long 1-D signals

The 1-D shift invariant transformation of a very large time bandwidth (TB) input vector can be implemented with a delay line whose time bandwidth is as large as the input vector. The delay line can be the travelling wave input modulator, as in a time integrating or space integrating correlators, or a scrolling detector, as in the TDI correlator. However, delay lines with a TB of 10^6 are not available, so it is necessary to fold the problem into two dimensions. With the long 1-D problem folded into two dimensions it is necessary to avoid edge effects due to the raster structure that has been arbitrarily imposed upon the signal in order to represent the high TB waveform. One way that this can be accomplished is by using a reference function which has redundantly sampled the cyclically shifted 1-D signal as it is being rastered into two dimensions[22]. This is explained with reference to Figure 3.2.1, which shows a binary signal of length 63 that has an unknown cyclic shift from the reference template. The input signal with an unknown delay is rastered into a 9 by 7 two dimensional format which will be represented in space and time (or pulse number) in a TSI optical system. The 2-D reference template is a redundant raster of the cyclically shifted reference waveform, where each line is twice as long as the spacing between lines, and there are twice as many lines in the reference template as in the raster of the unknown signal. This formatting allows the signal raster to

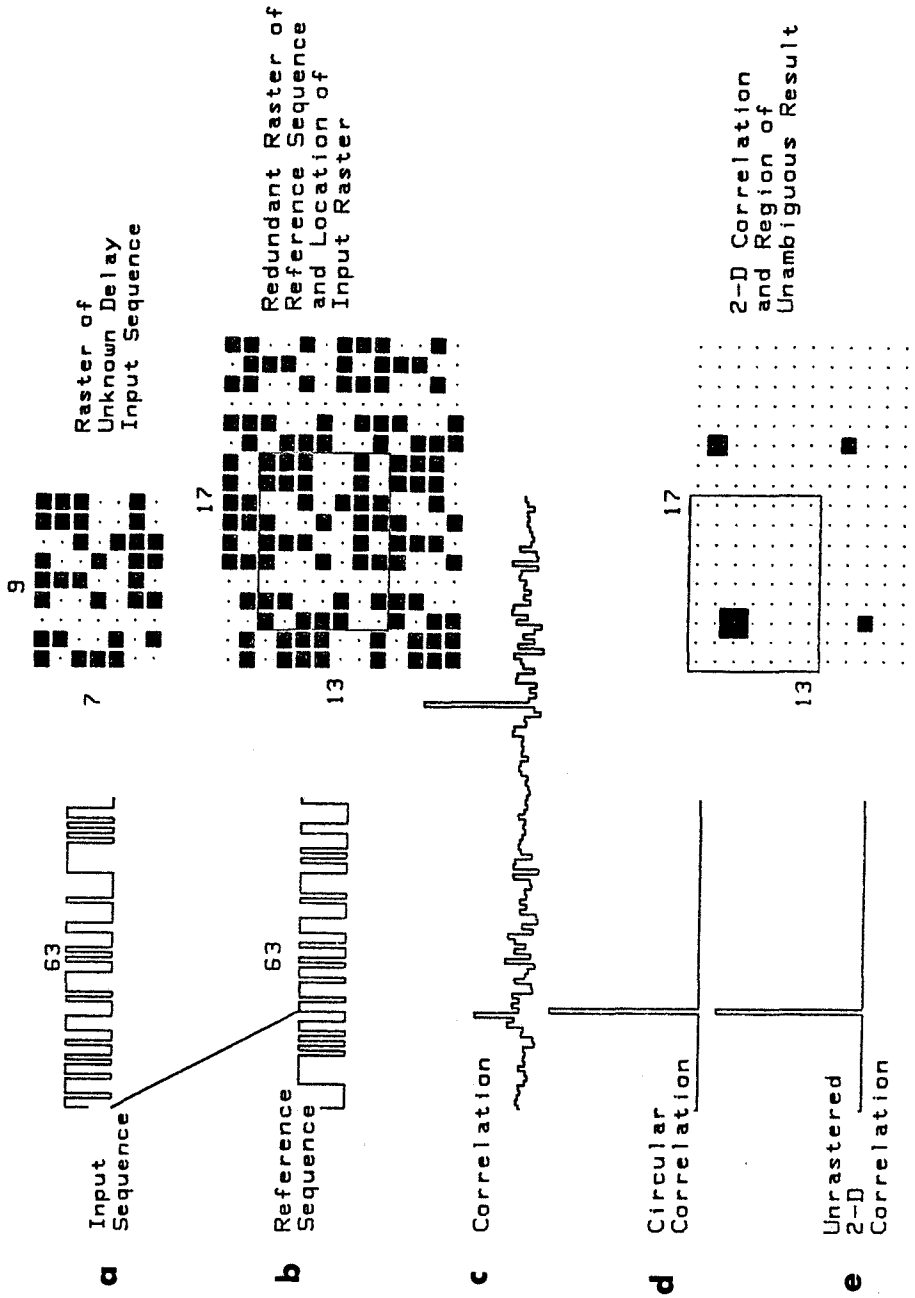


Figure 3.2.1. Redundant raster for a folded circular correlator, a) delayed input signal and its raster, b) reference and redundant raster, c) 1-D correlation, d) 1-D circular correlation, e) 2-D correlation and its unrastered version.

slide around within the boundaries of the reference template, always overlapping with the necessary information to form an appropriately delayed inner product, without encountering the boundaries of the reference template. Performing a two dimensional correlation between the raster of the input and the redundant raster of the reference will produce a correlation peak at the folded position that corresponds to the unknown delay of the input signal. The unambiguous portion of the output correlation plane, corresponding to the extent the input raster can shift within the reference raster without encountering the boundaries, can be unrastered to reproduce the desired high TB correlation.

It should be mentioned that this correlation operation as described is a circular correlation, which wraps back around on itself, to form a correlation output as long as the input. This is the desired operation for many of the pseudo-noise (PN) phase shift key (PSK) codes used in spread spectrum systems, because the maximal length codes have a nearly ideal sidelobe behavior in a circular correlation sense. The maximal length codes have a period 2^r-1, and are produced by an r stage linear feedback shift register whose input is generated from its current state by an exclusive OR of a combination of its taps which represent a primitive generator polynomial over the Galois field $G(2^r)$. These bipolar binary codes have the property that all sidelobes of the circular autocorrelation are between zero and one, while essentially all of the signal energy has been collapsed into a correlation peak of magnitude 2^r-1, while any noise included in the received signal remain uncompressed. As in digital FFT based circular correlators, when a conventional correlation is desired, it can be performed by padding the inputs with zeroes. In the folded correlator approach this means that the redundantly rastered reference template does not cyclically shift, so the reference template lines are twice as long as the space between lines, but there is only one more line in the reference template as there is in the input signal raster.

One version of the folded TSI correlator of very long signals is very similar

to the image correlator presented in section 3.2.1, except that when forming the multichannel reference hologram a double image of the rastered reference sequence is used. The space time raster of a long 1-D signal is given by a similar transformation to the raster transformation used in the previous section, and this is accomplished by sliding a long signal into an AOD and illuminating with a repetitively pulsed source.

$$\begin{aligned} f(x,nT) &= \int s(t+x/v)\delta(t-nT)dt \quad \text{rect}\left[\frac{x}{X}\right] \\ &= s(nT+x/v)\text{rect}\left[\frac{x}{X}\right] \end{aligned} \tag{3.2.1.1}$$

When a signal is given by a delayed version of a reference signal, $s(t) = r(t-t_0)$, with the delay expressed as an integral number of raster line times $T = X/v$, plus a fine offset τ, so that the overall delay is $t_0 = mT+\tau$, then the raster transformation is written as a serpentine delayed version of the reference signal.

$$\begin{aligned} g(x,nT) &= \int r(t-mT-\tau+x/v)\delta(t-nT)dt \quad \text{rect}\left[\frac{x}{X}\right] \\ &= r(nT-\tau+x/v)\text{rect}\left[\frac{x}{X}\right] \qquad\qquad m<n<m+N \end{aligned} \tag{3.2.1.2}$$

This consists of a space time raster beginning with the delayed pulse number m which consists of lines which are all spatially shifted in x by a fine delay $v\tau$. This means that the first line at time mT only contains a portion $v(T-\tau)$ of the input waveform, and the last line at time $(N+m)T$ contains a portion $v\tau$ of the long sequence, while all the $N-1$ lines in between fill the spatial aperture with a portion $vT = X$ of the input waveform.

The redundant spatial raster of the reference waveform consists of lines of length $2T$, each overlapping with the preceeding and the following line by T. Stated alternatively, it consists of two identical adjacent rasters of line length T, one shifted vertically by one line.

$$r(x,y) = \sum_{n=0}^{N} r(nT+x/v)\text{rect}\left[\frac{x}{2X}\right] * \delta(y-n\Delta y) \tag{3.2.1.3}$$

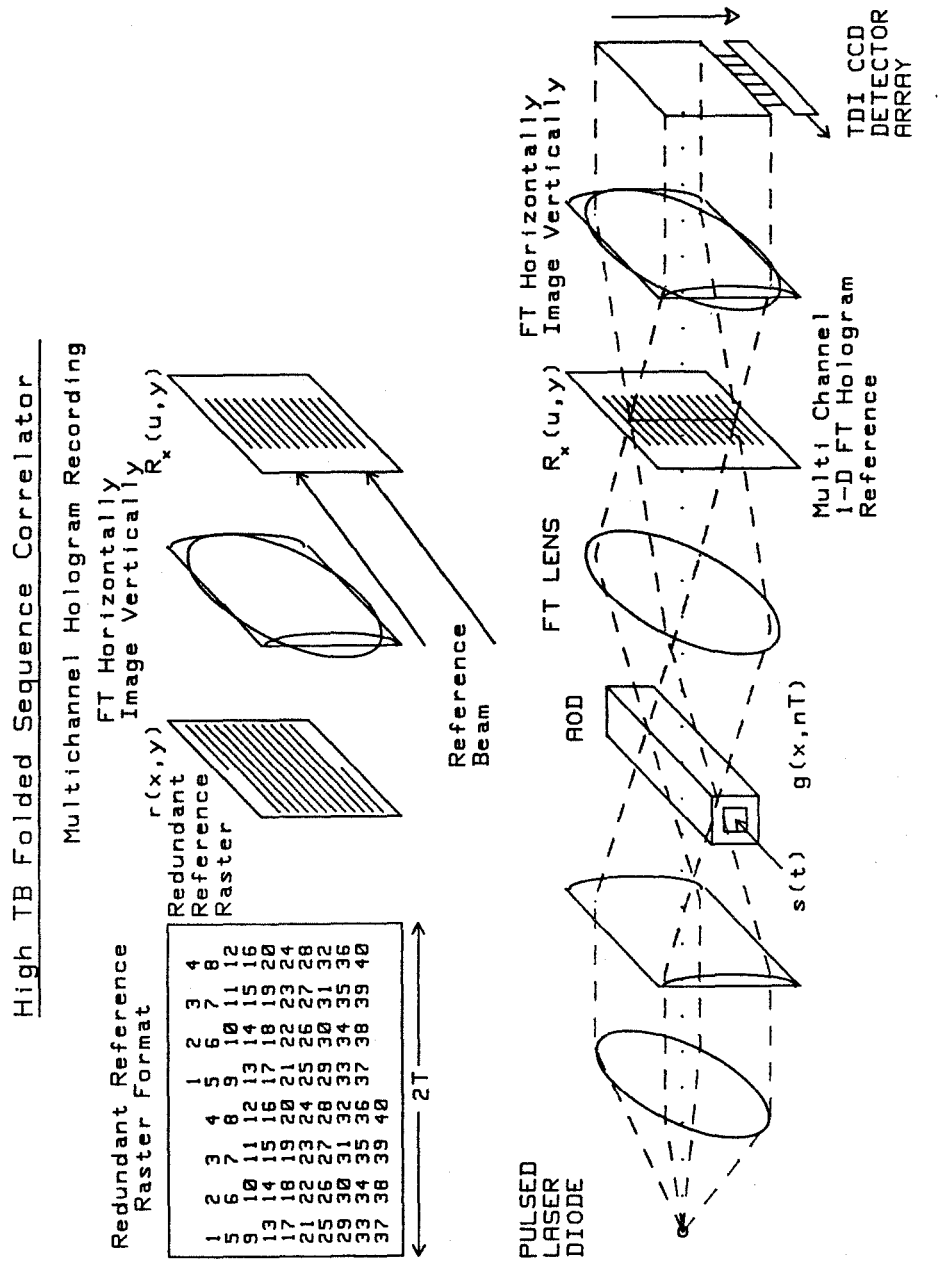


Figure 3.2.2. TSI folded correlator for very high time bandwidth signals.

$$= \sum_{n=1}^{N} r(nT + x/v)\text{rect}\left[\left[\frac{x}{X} - \frac{1}{2}\right] * \delta(y - n\Delta y) + \text{rect}\left[\frac{x}{X} + \frac{1}{2}\right] * \delta(y - (n-1)\Delta y)\right]$$

In this expression $v = X/T$ is the effective scan velocity of the time to space conversion in the fast dimension of the raster, the spacing between adjacent raster lines vertically is given by Δy, and the height of the raster lines is neglected for simplicity. For a long reference signal of duration NT, there are a total of $N+1$ raster lines of length $2T$ each, overlapping and separated by T, but the first and last lines are only half occupied. This pattern can be formed in real time by modulating a laser beam incident upon a dual travelling chirp lens Bragg cell, and vertical deflector which address a 2-D SLM, if a programmable joint transform implementation is desired. I will assume that a conventional Vander Lugt holographic correlator is to be built with a fixed reference function, so the two dimensional redundant reference template can be rastered off line with a computer controlled film recorder. The resulting mask is Fourier transformed vertically and imaged horizontally to form a multichannel 1-D hologram, which will act as the multichannel reference in the TSI correlator.

$$\begin{aligned} R(u,y) &= \sum_{n=0}^{N} \left| \int r(nT + x/v)\text{rect}\left[\frac{x}{2X}\right] e^{i2\pi ux} dx + ae^{i2\pi\alpha u} \right|^2 * \delta(y - n\Delta y) \\ &= \sum_{n=0}^{N} \Big[ae^{-i2\pi nTu} vR^*(vu) * 2X\text{sinc}[2Xu]e^{i2\pi\alpha u} \\ &\quad + a^* e^{i2\pi nTu} vR(vu) * 2X\text{sinc}[2Xu]e^{-i2\pi\alpha u} \\ &\quad + |a|^2 + |vR(vu) * 2X\text{sinc}[2Xu]|^2 \Big] * \delta(y - n\Delta y) \\ &= \sum_{n=0}^{N} \left[avR_n^*(vu)e^{i2\pi\alpha u} + a^* R_n(vu)e^{-i2\pi\alpha u} + |a|^2 + |vR_n(vu)|^2 \right] * \delta(y - n\Delta y) \end{aligned} \tag{3.2.1.4}$$

In this expression $R_n(f)$ is the temporal Fourier transform of the information of duration $2T$ in the nth raster line, while $R(f)$ is the Fourier transform of the

entire signal, and it is the first term containing the matched filters R_n^* that is of interest. When this multichannel 1-D holographic Vander Lugt filter is placed in the multichannel 1-D Fourier plane of the TSI folded correlator shown in Figure 3.2.2, then the TDI CCD detector will accumulate the full correlation of the high TB signal.

With each pulse of the laser a delayed and shifted version of part of the long signal is transduced into the processor, and this portion of the signal is Fourier transformed horizontally and uniformly expanded vertically to be incident upon each line of the multichannel hologram reference. The wavefront transmitted by the multichannel hologram on the nth laser pulse contains one line that strongly correlates with all of the light diffracted by the AOD for any delay τ. Adjacent lines may also match part of the illuminating signal, but when these partial correlation peaks are inverse transformed they will produce peaks outside of the TDI detector plane width, so these terms can be neglected. All of the other diffracted terms from the multichannel hologram contribute to the normal sidelobes associated with the waveform $r(t)$.

$$
\begin{aligned}
T(u,y,nT) =& R(u,y)\int r((n-m)T-\tau+x/v)\mathrm{rect}\left[\frac{x}{X}\right]e^{i2\pi ux}dx \qquad (3.2.1.5)\\
=& \sum_{n'=0}^{N}\left[ae^{-i2\pi n'Tu}vR_{n'}^*(vu) * 2X\mathrm{sinc}[2Xu]e^{i2\pi\alpha u}\right] * \delta(y-n'\Delta y)\\
& \quad vR_{n-m}(vu)e^{i2\pi u((n-m)T-\tau)} * \mathrm{sinc}[Xu]\\
=& e^{-i2\pi u(\tau-\alpha)}a|vR_n(vu)|^2 * \delta(y-(n-m)\Delta y) * X\mathrm{sinc}[Xu] * 2X\mathrm{sinc}[2Xu]\\
& + \sum_{n'\neq n-m} av^2e^{-i2\pi u[n'-(n-m))T+\tau-\alpha]}R_n'(vu)R_{n-m}(vu)\\
& * \delta(y-n'\Delta y) * X\mathrm{sinc}[Xu] * 2X\mathrm{sinc}[2Xu]
\end{aligned}
$$

The first term represents a diffracted plane wave coming out of the $(n-m)$th row of the multichannel hologram, tilted at a spatial frequency of τ, on the nth pulse of the laser. The other term is the Fourier domain representation of all of the partial

cross correlations of individual reference raster lines with the part of the reference sequence diffracted on the nth laser pulse, and these will be accumulated in order to produce the correct correlation sidelobes. Only the term which will produce a focussed correlation peak on the detector will be included in the following expression, but the other terms diffracted by the hologram will build up the appropriate sidelobes of the long 1-D correlation as well. The inverse 1-D Fourier transform of this diffracted plane wave will produce a correlation peak on the CCD detector array in the $(n-m)$th row and in the $(v\tau/\Delta x)$th column, where Δx is the horizontal pixel separation.

$$t(x,y,nT) = \delta(x-\tau) * \delta(y-(n-m)\Delta y) * \sqrt{TB}\mathrm{sinc}[Bx/v] \qquad (3.2.1.6)$$

From pulse to pulse this represents a correlation peak whose spatial width is given by the velocity scaled bandwidth of each segment, which is assumed constant, which then travels linearly down the detector array from pulse to pulse. Detecting on a TDI CCD which is synchronized with the laser pulsing, and shifting vertically a distance of Δy after each laser pulse will accumulate all of the correlation peaks in a single scrolling pixel. The CCD needs to have a horizontal width equal to the image of the time $T = X/v$ in the AOD, and it must be shifted to the position corresponding to the matched filter diffracted term from the hologram. The intensity integrated in the moving coordinate frame of the TDI CCD is the noncoherent sum of the intensity integrated at each delayed detector position.

$$\begin{aligned} I(x,t=kT) &= \sum_{n=k-N}^{k} \int |t(x,y,nT)|^2 \delta(y-y'-nT)dt \\ &= NTB\mathrm{sinc}^2[B(x/v-\tau)]\delta_{km} \end{aligned} \qquad (3.2.1.7)$$

This represents a nonlinear type of pseudo correlation of the very long delayed input sequence with the reference sequence stored in the hologram, and it has the full processing gain of the high time bandwidth sequence. This operation has an

infinite range delay search because the TDI CCD continuously scrolls out the next slice of the folded pseudo correlation and the horizontal readout register unfolds this back into a high TB waveform.

When a linear correlation is desired, then a plane wave reference can be added and the linear correlation can be placed on a spatial carrier, so that carrier demodulation can be used to separate the linear correlation from the nonlinear correlation and from the reference beam bias.

$$I(x,t=kT) = \sum_{n=k-N}^{k} \int |t(x,y,nT) + re^{i2\pi\alpha x}|^2 \delta(y-y'-nT)dt \quad (3.2.1.8)$$
$$= NTB\mathrm{sinc}^2[B(x/v-\tau)]\delta_{km} + |r|^2$$
$$+ rN\sqrt{TB}\mathrm{sinc}[B(x/v-\tau)]\delta_{km}\cos[(2\pi\alpha x)+\phi]$$

The last term can be demodulated in order to obtain the linear correlation, as well as the phase ϕ of the returned signal. The time delay of the return is determined by the position of the unfolded two dimensional correlation peak, and this unfolding can be accomplished by the raster read out mechanism incorporated within the TDI CCD. The time delay of the return is measured as the horizontal fine delay measurement in the TDI column containing the correlation peak at position $x = v\tau$, plus the coarse delay measurement which is determined by when this peak scrolls out of that column, $t = mT$. Thus the overall delay of the returned waveform is determined to be $t_0 = mT+\tau$, which is the correct delay of the input waveform. This delay measurement can be determined from the interferometrically detected linear correlation, or from the incoherently accumulated nonlinear correlation, because both will produce a peak at the same output location. The optimum reference beam amplitude is equal to the correlated signal beam at the detector plane, $r = \sqrt{TB}$, and in this case the linear correlation has the full processing gain of the high time bandwidth waveform.

The incoherently accumulated pseudo correlation, consists of segments which

are linearly correlated, then magnitude squared and accumulated with subsequently delayed and computed squared correlations. This type of nonlinear correlation will destroy the beautiful sidelobe properties of conventional PN sequence codes. However, a code that is created from a number of maximal length PN codes of the same length stacked end to end, and then rastered into a 2-D format with one maximal length code per raster line may have excellent nonlinear correlation properties in this type of TSI processor operated in the simpler incoherent detection mode.

Alternatively, a folded image plane style of space integrating correlator can be constructed by using a parallel output wideband detector array connected by high bandwidth delay lines to appropriately accumulate partial correlations[23]. However, since this type of detector and accumulating delay line structure is not commercially available, an architecture that folds the problem into two spatial dimensions and uses standard 2-D CCD technology might be preferable. This can be accomplished by using a travelling lens chirp scanner to scan the array of partial correlations onto a 2-D TDI CCD detector array. The chirp lens repeats with a period T that is equal to the image plane mask width scaled by the AO velocity, and that is synchronized with the orthogonal shifting of the TDI CCD array. This effectively turns the temporal output of a space integrating image plane correlator into a time integrated spatial variation of light for each traversal of the acoustic lens. This type of system is illustrated in Figure 3.2.3. It has an infinite range delay search, and a processing gain equal to the 2-D space bandwidth of the mask which represents the long sequence in a folded raster format. The mask does not need to be redundantly rastered as long as the chirp scan time is accurately matched to the TDI shift time, and the scanner fly back time is negligible. The image plane approach has a distinct advantage over the holographically implemented folded sequence correlator that the mask is in the image plane, so it can be updated in real time by using a

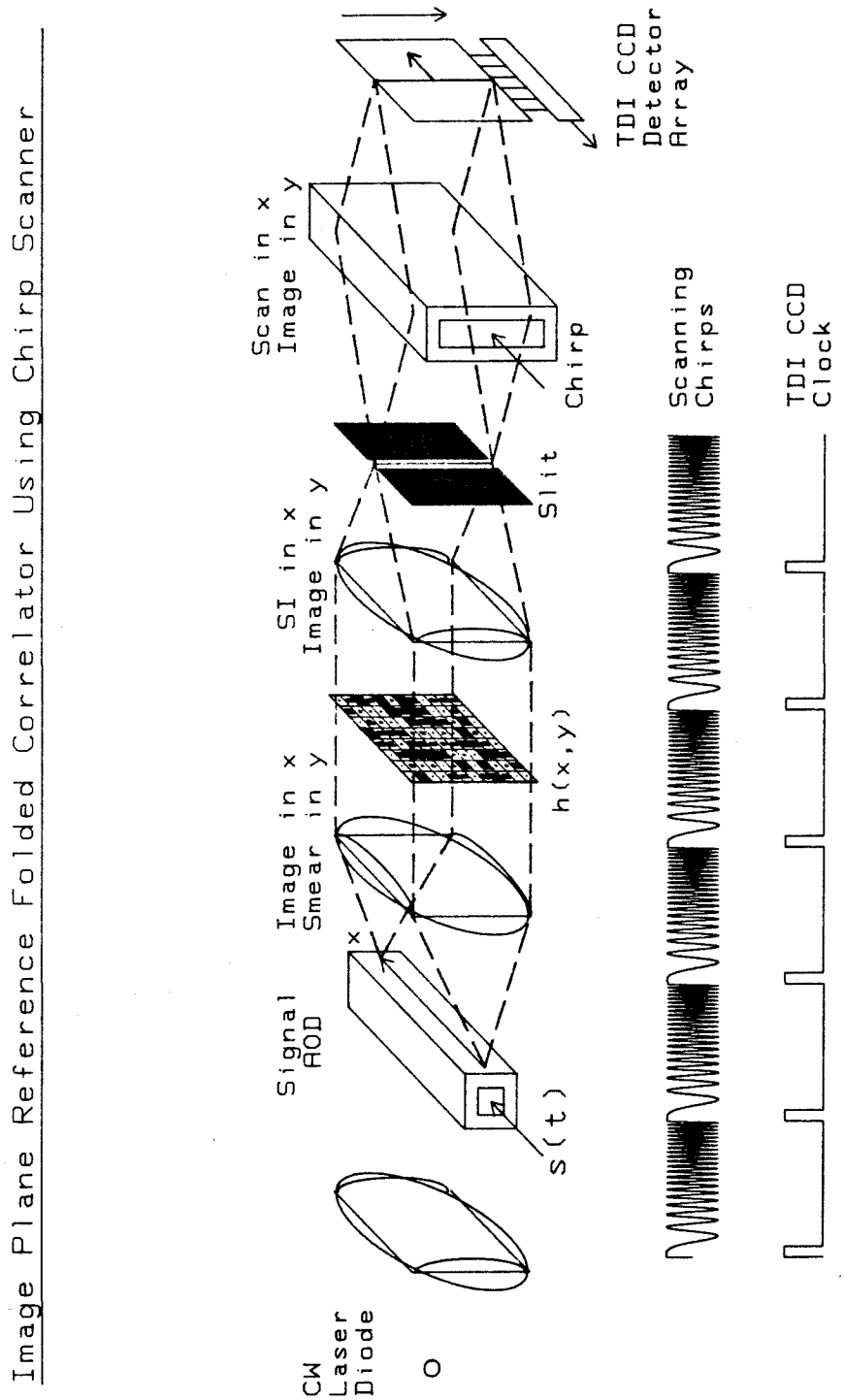


Figure 3.2.3. Image plane folded sequence TSI correlator.

2-D SLM. Once again this is a nonlinear type of pseudo correlator because coherent integration are performed over the width of one line of the mask, and incoherent summations of these partial correlations are performed by the CCD. When linear correlations are desired, a reference beam can be added, but instead of illuminating the CCD with a uniform reference, a line reference should be used to illuminate the scanned sampling slit, which can be accomplished by illuminating the sampling slit with a focussed reference beam. This will decrease the bias due to the reference, and better match the reference beam amplitude to the space integrated signal beam amplitude at the peak.

The sequence correlators described in this section have essentially the same computational power as the image correlator described previously. Raster images almost always occur at 30 frames per second, so there is very little motivation to speed up the processor, even though Bragg cells with multi GHz bandwidths, and μsec access time are commercially available. On the other hand, high time bandwidth sequences with GHz bandwidths and msec duration could be processed with the folded sequence correlator, if a fast enough CCD detector was avaliable, and if enough power could be obtained from the pulsed laser diodes. Parallel output CCD detector arrays have been fabricated that operate at 2000 frames per second[24], and GaAs CCDs may go even faster. The laser pulse width must be cut down to the nsec time scale if GHz bandwidth signals are to be processed, and pulse repetition frequencies will be in the MHz range, which will probably be difficult to achieve with coherent pulses, but the partially coherent summation could probably be achieved. The power output that will be needed will be about 100 times as large as the image processor, which only requires about 100μW of total power if 30% diffraction efficiency can be achieved in the hologram. This corresponds to 100mW peak power pulses with a duty factor of 1:1000, and this becomes 10W peak power for the high bandwidth sequence correlator, and an average power of 10mW. These peak and

average power capabilities are currently available in multimode laser diodes, so it should be possible to construct this type of high bandwidth sequence correlator out of existing hardware. The processing power that could be achieved would be on the order of 6×10^{13} raw analog multiplies per second, with an FFT equivalent processing power of 5×10^{9} multiplies per second.

3.2.2 TSI processing of long frequency hopped signals

The TSI frequency hopped (FH) code processor[25] is really a special case of the folded sequence correlator presented in the previous section, with the 2-D TDI detector array replaced by a wide aperture 1-D TDI detector array, and the bank of holographic matched filters replaced with a 2-D spatial mask with one aperture open in each row, so that it only has coarse delay resolution. This has the advantage that the mask can be changed in real time in order to alter the received code, but this flexibility is at the cost of a much decreased processing gain. The transmitted frequency hopped code with constant time between hops is represented as a sum over each non overlapping constant frequency segment.

$$s(t) = \sum_{n=1}^{N} \text{rect}\left[\frac{t-(n+\frac{1}{2})T}{T_1}\right] e^{i2\pi[f_n t+\phi_n]} \qquad (3.2.2.1)$$

Because of limits in the hopping rate of conventional signal generators T is typically between 1 and 10μsec or more, and there may be some dead time between tones, so that $T \geq T_1$. The sequence of frequencies f_n is chosen to have a low correlation with other sequences being used by other transmitters known to be operating in the spectral band of interest. The sequence of starting phases, ϕ_n, of each tone segment are typically random, with a uniform probability distribution between 0 and 2π, because the wideband frequency hopping signal generators are not phase coherent, and for this reason an incoherent correlation is appropriate.

A TSI correlator designed to detect a particular sequence of tones with an arbitrary delay is shown in Figure 3.2.4, and the correlator can be programmed to

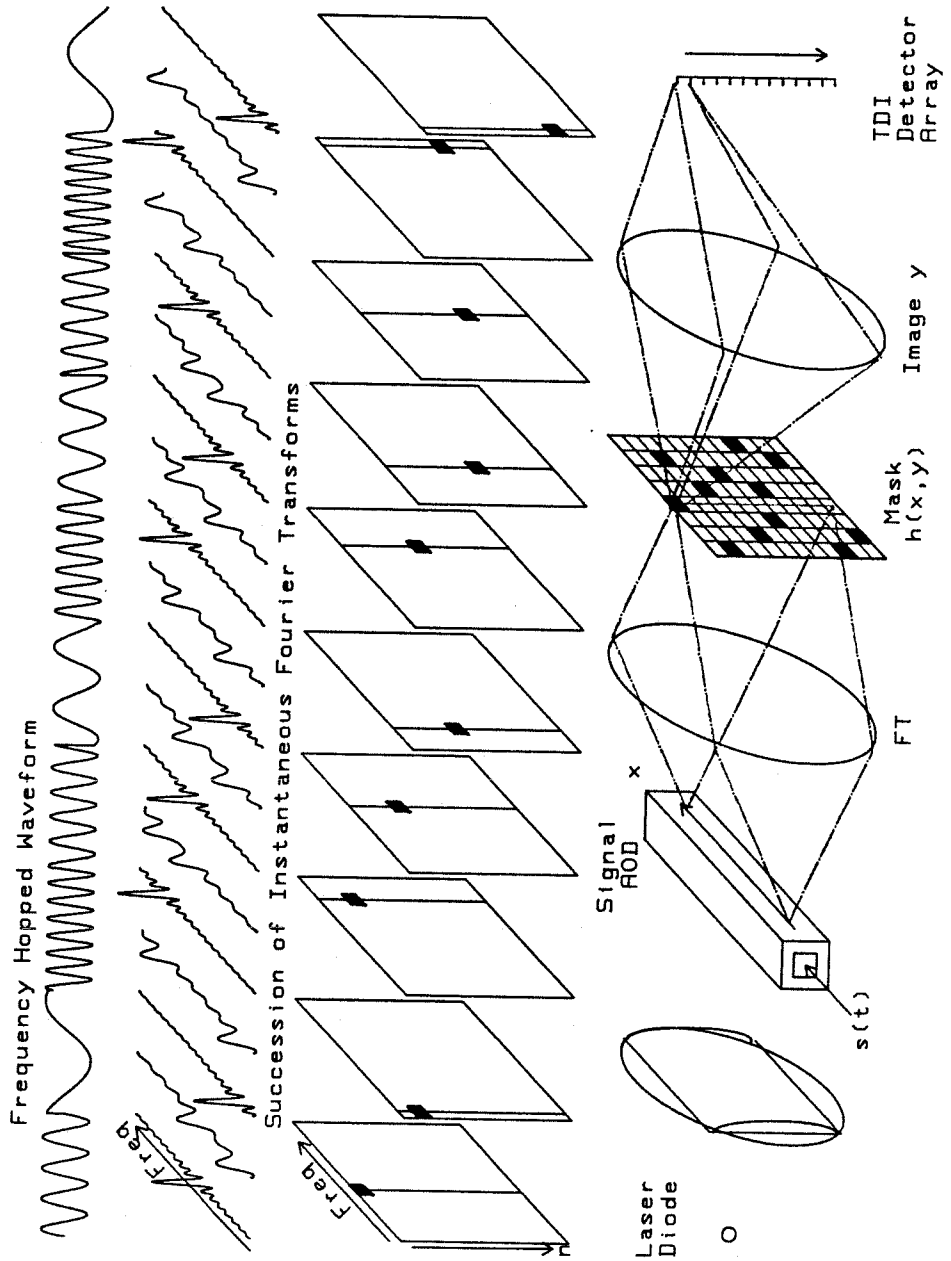


Figure 3.2.4. TSI frequency hopped processor, a) frequency hopped waveform, b) succession of instantaneous Fourier transforms, c) spatial pattern passing through the mask for successive frequency components, d) optical processor architecture.

detect another FH code sequence by modifying the spatial mask. The mask can be described as a sum over the rows of one aperture per row at the expected frequency for the time slot corresponding to that row.

$$H(u,y) = \sum_{j=1}^{N} \text{rect}\left[\frac{uv - f_j}{\Delta f}\right] \text{rect}\left[\frac{y - j\Delta y}{\Delta y}\right] \qquad (3.2.2.2)$$

The operation of this TSI system is easily explained by considering the succession of space integrated spectra incident on the mask as the code slides through the AOD, which has an aperture approximately equal to the duration of a frequency hopped tone. As the first frequency slides into the AO cell, a vertical slit of light appears at position f_1, which shrinks in width and grows in amplitude as the AO aperture is filled. This lets a beam of light through the mask aperture at position f_1 in the first row, and this is imaged onto the first row of the 1-D TDI detector array. As the second frequency segment enters the AO cell, a slit of light appears on the mask at the column corresponding to f_2, letting light through the aperture in the second row. The CCD is vertically shifted by one mask pixel at the same rate as the expected frequency hopping rate, so that the beam of light transmitted by the mask in the second row adds to that already produced at the previous integration location. As the integrating potential wells continue to shift down the detector array the one that is most closely synchronized with the hopping code incoherently accumulates a large peak, while the other pixels remain empty. It is easy to show that the time integrated power that passes through a mask aperture due to a single frequency segment is a constant, say A, which is proportional to the received power and the tone duration. However, since the TDI detector array shift time is not necessarily aligned with the incoming frequency hops, it is possible for this time integrated power to be divided among the adjacent pixels, thereby losing up to 6dB of processing gain, but for simplicity synchronization will be assumed, so that the

delay will be assumed to be $t_0 = pT$.

$$\begin{aligned} I(u,y,kT) &= \int_{kT}^{(k+1)T} \left| \int_A s(t-t_0+x/v) e^{i2\pi xu} dx \right|^2 H(u, y-k\Delta y) dt \\ &= A \sum_{n=1}^{N} \delta(uv - f_n)\delta(k-(p+n)) \\ &\quad \sum_{j=1}^{N} \delta(uv-f_j)\delta(y-j\Delta y) * \text{rect}\left[\frac{uv}{\Delta f}\right] \text{rect}\left[\frac{y}{\Delta y}\right] \\ &= A\delta(uv - f_{k-p})\delta(y-(k-p)\Delta y) * \text{rect}\left[\frac{uv}{\Delta f}\right] \text{rect}\left[\frac{y}{\Delta y}\right] \end{aligned} \tag{3.2.2.3}$$

During the time interval from kT to $(k+1)T$ a slit of light is incident upon the f_{k-p}th column of the mask, and a small windowed beam of light passes through the aperture in the $(k-p)$th row of the mask, and this is illustrated in Figure 3.2.4 by the sequential time slices shown passing through the mask. By detecting in a vertically scrolling coordinate frame behind the mask, N successive frequency hopped pulses can be accumulated in a travelling potential well. The TDI detector needs to have pixels as wide as the horizontal extent of the mask in the Fourier plane, or an anamorphic demagnification of the Fourier plane onto the detector must be used. This is analytically described as an integration over the spatial frequency plane variable u.

$$\begin{aligned} I(y'_{t=(m+N)T}) &= \sum_{k=m}^{m+N} \int \delta(y-y'+k\Delta y)\delta(y-(k-p)\Delta y)\text{rect}\left[\frac{y}{\Delta y}\right] dy \\ &\quad A\int \delta(uv-f_{k-p}) * \text{rect}\left[\frac{uv}{\Delta f}\right] du \\ &\approx NA\text{rect}\left[\frac{y'-p\Delta y}{\Delta y}\right] \end{aligned} \tag{3.2.2.4}$$

This indicates that all of the energy in the FH waveform has been incoherently accumulated into a peak at the pixel that scrolls out of the bottom of the detector array at time $(p+N)T$, thereby measuring the delay of the received waveform. The sidelobe components that pass through other mask apertures, and fall on TDI CCD pixels with different shift ordinates will contribute to the appropriate sidelobes of

the FH correlation function. Once again this is actually a nonlinear correlation operation, which has excellent delay measurement capabilities, but is not identical with the linear correlation.

3.2.3 TSI separable kernel transformation of long 1-D signals

The high resolution output variable of a long 1-D linear transformation can be represented in a folded coordinate space, $\tau = \tau_x + N\tau_y$ with about $N \approx 10^3$ positions in each folded dimension, in order to represent the full 10^6 TB of the output. The linear transformation equation transforming the high TB input signal to the folded output space has too many degrees of freedom to represent within the constraints of an optical system, unless the transformation kernel is separable in the folded output variables.

$$g(\tau_x, \tau_y) = \int f(t)h(t, \tau_x, \tau_y)dt = \int f(t)h_1(t, \tau_x)h_2(t, \tau_y)dt \qquad (3.2.3)$$

The transformation kernels h_1 and h_2 have $10^6 \times 10^3$ independent entries, and therefore they still overflow the information handling capacity of the optical system. Anticipating the folded time space raster representation that will be used on the input variable t, the input space can also be folded, and if the kernels in x and y are each independent of the other folded input variable, then the problem is separable in the folded coordinate system.

$$\begin{aligned} g(\tau_x, \tau_y) &= \sum_{n=1}^{N} \int f(x, nT)h_1(x, nT, \tau_x)h_2(x, nT, \tau_y)dx \\ &= \sum_{n=1}^{N} \int f(x, nT)h_x(x, \tau_x)h_y(nT, \tau_y)dx \end{aligned} \qquad (3.2.4)$$

The transformation kernels h_x and h_y are two dimensional functions with $10^3 \times 10^3$ degrees of freedom, so they can be internally represented in the optical system, h_x as a spatial transformation and h_y as a time space raster.

The classic example of such a transformation kernel is the Fourier transform, and it is this type of decomposability of the Fourier kernel which results in the

fast Fourier transform (FFT) algorithm used in digital signal processing. The FFT is based on successively dividing the DFT in half $\log_2 N$ times, while the folded spectrum is based on dividing the DFT into two successive orthogonal arrays of $\sqrt{N}$ transforms, each with dimensionality of $\sqrt{n}$. When the output frequency ordinate is represented in a radix base equal to the square root of the overall resolution $f = \sum_{j=0}^{1} f_j(\sqrt{N})^j = f_0 + f_1\sqrt{N}$, then the Fourier transform can be derived as orthogonal sub transformations in a folded coordinate space, where f_0 is a fine frequency variable and f_1 is a coarse frequency variable.

$$e^{i2\pi ft} = e^{i2\pi \sum_{j=0}^{1} f_j(\sqrt{N})^j} = \prod_{j=0}^{1} e^{(i2\pi f_j(\sqrt{N})^j)t} = e^{i2\pi f_0 t} e^{i2\pi f_1(\sqrt{N})t} \qquad (3.2.7)$$

The Fourier kernel is thus decomposable as a fine frequency phase factor times a coarse frequency phase factor, and this allows a separable implementation of a high resolution Fourier transformation. A long integration of a high TB input can be broken up into the delayed summation of a number of smaller integrations. In the TSI approach the summation is accomplished with a time integration over successive pulses, while the smaller integration is accomplished with a repetitively computed spatial integration over the AOD aperture. In terms of the computation of the Fourier transform of a long input signal we can factor the Fourier kernel to express this transformation as a sequential computation.

$$S(f) = \int_0^{NT} s(t)e^{-i2\pi ft}dt = \sum_{n=0}^{N-1} e^{i2\pi fnT} \int_0^T s(x+nT)e^{-i2\pi fx}dx$$

The input variable t has been decomposed into a succession of N high bandwidth segments, and correspondingly we can separate the output variable f into a coarse frequency component l/T, and a fine frequency component k/NT, so that $f = l/T + k/NT$.

$$\begin{aligned} S\left(\frac{l}{T}, \frac{k}{NT}\right) &= \sum_{n=0}^{N-1} e^{-i2\pi ln} e^{-i2\pi kn/N} \int_0^T s(x+nT)e^{-i2\pi lx/T} e^{-i2\pi kx/NT}dx \\ &\approx \sum_{n=0}^{N-1} e^{-i2\pi kn/N} \int_0^T s(x,nT)e^{-i2\pi lx/T}dx \end{aligned}$$

$$= \sum_{n=0}^{N-1} e^{-i2\pi kn/N} S_n(l/T)$$

The phase factor $e^{-i2\pi ln} = 1$ for all n and l, so it can be dropped. The additional phase factor $e^{-i2\pi kx/NT}$ has been dropped, because it simply provides a skew to the output coarse and fine raster. A succession of coarse spectra can be formed via space integration and these complex spectra can be transformed in the orthogonal dimension with a temporal accumulation over a succession of pulses forming a folded representation of the high TB output spectra.

The various decompositions of the Fourier kernel which are possible allow a large number of possible implementations of spectral analysis systems, including the Fourier transform by a lens, the DFT, the FFT, and the chirp transform. When designing a hybrid TSI folded spectrum system of a very high TB waveform there are a large number of possible systems that can be designed, and a number of these are analyzed both theoretically and experimentally in detail in Chapter 5. In this section I will briefly outline the difference between a TSI folded spectrum processor and a TSI image Fourier transformation system. The image Fourier transform system presented in section 3.1.2 and shown in Figure 3.1.4 can also be used to perform folded spectrum analysis of a very long signal with only slight modifications. The optics can remain identical and the same DFT reference waveform space time raster can be used for fine frequency analysis via temporal integration. The system is illuminated by a pulsed source of period T, which freezes the acoustic motion of the columns of the DFT matrix in the reference AOD, and simultaneously transduces successive delayed portions of the input signal into the system to be spatially Fourier transformed by a lens. This description sounds identical to the operation of the image Fourier transformation system, but there are critical differences. In the image transformation system the two orthogonal transformations are completely independent, and the various resolution limitations do not affect the other dimension.

While in the folded spectrum system it is very important that the resolution of the space integrating coarse analyzer be less than or equal to the analysis bandwidth of the sampled time integrating fine frequency DFT processor. The maximum obtainable bandwidth of the fine frequency processor is given by the Nyquist sampling rate of the laser $f_N = 1/2T$ and the coarse resolution is limited by the signal AOD aperture $f_c = 1/T_A$, so there must be more than two laser pulses during the acoustic transit time of the signal AOD. Therefore, the space time raster is oversampled and adjacent lines contain overlapping pieces of the long input signal. Another difference between a folded spectrum processor and an image Fourier transformation system arises when considering the succession of starting phases transduced from pulse to pulse in the input space time raster. In the folded spectrum processor with a single tone input the phase advance from pulse to pulse and across a line is always the same, so that in the space time raster only certain fringe angles are possible; while in an image transform system with a tilted and skewed sinusoidal fringe input the phase advance in the horizontal dimension is completely independent from that in the vertical dimension, so that tilted fringes at any angle are possible inputs. These issues will be discussed at great length in Chapter 5.

3.3 Synthetic Aperture Radar

Synthetic Aperture Radar (SAR) has been the most successful application area of coherent optical processing techniques from the inception, and remains so to this day[26–31]. It also represents an almost ideal match with the capabilities of the TSI processing technique. The design and performance of TSI SAR processing systems has been reported in great detail in the thesis by my colleague[32], as well as in several publications[33–39], so it will only be briefly reviewed in this section. The importance of the conjugate sideband term present in multiplicative processors will be illustrated, and interferometers designed to perform SAR processing will be presented that eliminate the spurious sidelobes associated with this unwanted misfocusing term.

The SAR problem arises from the desire to obtain a high resolution radar image of the ground in both range and cross range dimensions from a moving aircraft platform without carrying a huge real aperture antenna. Radar systems can obtain high resolution ranging information through pulse compression techniques by using time delay measurements of reflected waveforms. Ideally, a narrow RF pulse would be broadcast, and the resulting range delay measurements could be made to the precision of the pulse width times half the speed of light (due to the round trip propagation), as long as the reflected wave produced a detected signal well above the noise floor. However, the peak power limitations of RF transmitters, combined with the $1/R^4$ decay of the reflected amplitude with range R, leads to a pulse coding approach to RF ranging. Typically the narrow pulse is applied to a surface acoustic wave (SAW) dispersive delay line producing a spread out chirp pulse of bandwidth B, and duration T_c, that can be broadcast with low peak power. The received pulse with an unknown range delay can be correlated against the broadcast reference pulse using an AO correlator which will achieve a processing gain and compression ratio of T_cB. The optical matched filter recompresses the received signal to an effective

pulse of width $1/B$, leading to a slant range resolution $\delta_r = c/2B$. Bandwidths of 50 MHz can easily be obtained with this type of technique, thereby obtaining range information with a resolution of 3 m. A scanning real aperture antenna of width D_a has a 3dB angular beamwidth (in radians) $\theta = .88\lambda_r/D_a$, so the diffraction limited cross range ground resolution is $\delta \approx .88R\lambda_r/D_a$. To obtain the same resolution in the range and cross range dimension for typical ranges of 10 Km and radar frequencies of 1 GHz, requires that the antenna width must be impractically large $D_a = R\lambda_r/\delta_r \approx$ 1Km. The synthetic aperture technique allows the use of a small moving antenna in order to obtain high resolution in the azimuth (or cross range) dimension. This is accomplished through the use of a special purpose signal processor that coherently accumulates radar returns recorded at a sequence of locations as the aircraft traverses a distance equal to the required real aperture. A small aperture radar will illuminate a wide antenna footprint on the ground of 4dB width in the along track dimension $L = R\lambda_r/D_a$, so that a succession of pulses made as the aircraft traverses a distance L will all illuminate a given point scatterer on the ground. This allows the signal processor to synthesize a synthetic aperture of width L and achieve an azimuthal ground resolution of $\delta_a = D_a/2$, that is twice as fine as a real aperture of width L. This azimuthal resolution is independent of range, because the beamwidth, or synthesized processing aperture, grows with range, thereby compensating for the range dependent azimuthal resolution of the synthesized processing aperture. The inverted dependence of the azimuth resolution on the radar antenna width can only be achieved when adequately sampled data from the full synthetic aperture for each range are processed coherently retaining all phase relationships, and when the aircraft follows a simple known trajectory, such as a straight line at constant velocity.

There are several equivalent viewpoints which can be used to explain the operation of a SAR imaging system[26–29], including the range doppler principle, the

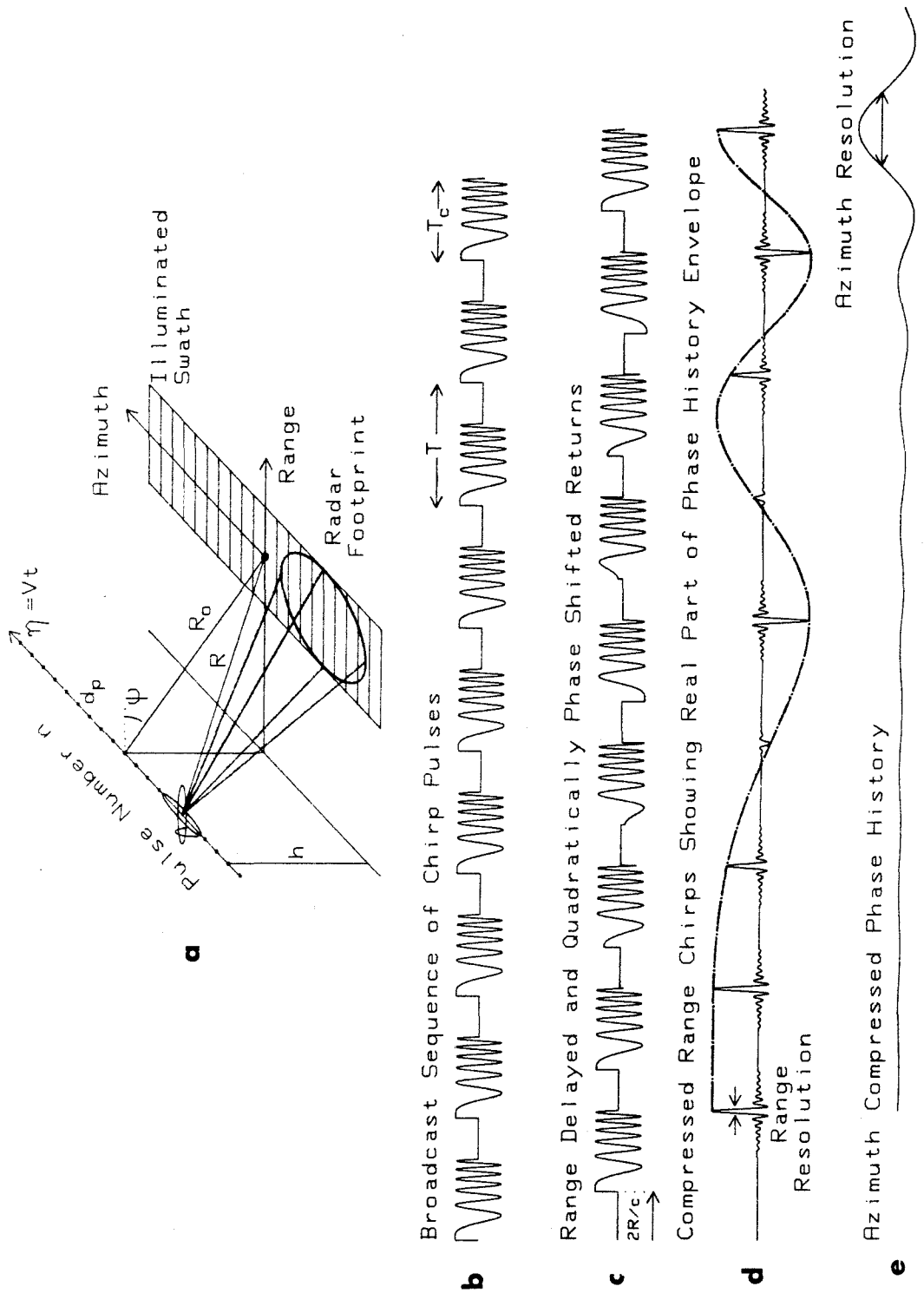


Figure 3.3.1. a) Synthetic Aperture Radar (SAR) geometry, b) sequence of broadcast chirp pulses, c) received range delayed and phase shifted chirp pulses, d) range compressed returns showing real part of the azimuthal phase history, e) azimuth compressed phase history.

synthetic aperture technique, the azimuthal phase history approach, the matched filter or cross correlator viewpoint, and the holographic viewpoint. The azimuthal phase history approach will be briefly discussed with reference to Figures 3.3.1 and 3.3.2. An airplane at a height h is travelling at a velocity V in the η direction, carrying a small antenna with its boresight mounted orthogonal to the velocity vector (side looking), pointed downward at a slant angle ψ. The radar emits a series of chirp pulses of duration T_c at a nominal pulse repetition frequency (PRF) $f_p = 1/T$ ($T_c << T$), that is slaved to the aircrafts velocity, so that the pulses are actually transmitted from discrete locations, separated by a constant distance $d_p = VT$. The broadcast sequence of chirp pulses is given by the reference signal

$$s(t) = \sum_n \text{rect}\left[\frac{t - nT}{T_c}\right] e^{i2\pi\nu_r t} e^{i\pi b(t-nT)^2} \tag{3.3.1}$$

In this expression ν_r is the radar carrier frequency (typically 1-10 GHz), and $b = B/T_c$ is the chirp acceleration parameter, usually expressed in MHz/μsec. The signal that is reflected from a single point scatterer of reflectivity σ, on the ground at location (ξ_0, η_0), where ξ is the ground range coordinate and η is the azimuth position, is given by a delayed version of the transmitted signal, apodized by the radar footprint as it sweeps across the stationary scatterer.

$$S(t) = \sigma \tilde{A}(t, \xi_0, \eta_0) s(t - 2R/c) \frac{1}{R^4} \tag{3.3.2}$$

The radar footprint $\tilde{A}(\xi, \eta) = A(\xi \cos\psi, \eta) = A(R_0, \eta)$ is given in the far field by the scaled Fourier transform of the antenna aperture of width D_a and height D_r, so that $\xi = \lambda_r R/D_a$ and $\eta = \lambda_r R/D_r$, stretched in range by the cosine of the slant angle ψ. The $1/R^4$ range dependent gain of the radar is due to the $1/R^2$ beam spreading of the transmitted and reflected waves, and is compensated with a t^4 time varying amplification of the returned signal on each pulse. The target range R is a function of time t, or pulse number nT, and reaches a minimum R_0 when the aircraft position is equal to the targets azimuthal coordinate $\eta = Vt = \eta_0$.

The azimuthal phase history of a point scatterer can be derived by considering the aircraft to be essentially stationary during each broadcast and reception interval, separated spatially by d_p, so the reflection from a single point scatterer on the ground is simply a delayed version of each transmitted pulse. The delay is given by the instantaneous range at the time of the pulse broadcast $R(t) \approx R(nT)$, and the range is typically much larger than the footprint width on the ground, allowing a small angle approximation of the square root.

$$R(t) = \sqrt{h^2 + \xi_0^2 + (\eta - \eta_0)^2} = \sqrt{R_0^2 + (Vt - \eta_0)^2}$$
$$R(nT) \approx R_0 + \frac{(VnT - \eta_0)^2}{2R_0} = R_0 + \frac{(nd_p - \eta_0)^2}{2R_0} \tag{3.3.3}$$

The received signal due to a single point scatterer consists of a sequence of chirp pulses which have been delayed by the large constant delay $R_0 = \sqrt{h^2 + \xi_0^2}$, plus a small quadratic delay term, the offset of which is dependent on the targets azimuthal position η_0. For SAR processing with a reasonably small synthetic aperture $L_s = Nd_p$ the higher order effect of range walk and range curvature can be ignored in the side looking geometry as long as the target range migration is less than the range resolution, $(Nd_p)^2/8R_0 < c/2B$. In this case the fine delay term is only measurable on the phase of the carrier on successive returns, and this phase delay is approximately quadratic in the pulse number.

$$S(t) = \sigma \sum_n \text{rect}\left[\frac{t-2R(nT)/c-nT}{T_c}\right] e^{i\pi b(t-2R(nT)/c-nT)^2} e^{i2\pi\nu_r(t-2R(nT)/c)} \tilde{A}(nT, \xi_0, \eta_0)$$
$$\approx \sigma e^{i2\pi\nu_r t} \sum_n \text{rect}\left[\frac{t-2R_0/c-nT}{T_c}\right] e^{i\pi b(t-2R_0/c-nT)^2} e^{i2\pi\nu_r(nd_p-\eta_0)^2/R_0 c} A(nT, R_0, \eta_0) \tag{3.3.4}$$

This returned signal must be processed in order to recreate the image of the point scatterer on the ground, at position (R_0, η_0), where R_0 is the slant range and η_0 is the azimuth position. The approach used in the TSI processor is based on sequential processing of the range and azimuth dimensions. This can be accomplished by

first correlating against the transmitted chirp to obtain a sequence of range compressed impulses, detecting the phase of each range compressed impulse, and then correlating against the appropriate range dependent phase history for each range bin, as shown in Figure 3.3.1. This successive range and multichannel azimuth correlation approach is the basis for modern digital SAR processors[40], and it is the same principle which is employed in the cascaded TSI SAR processor architecture.

The original approach to processing SAR data utilized a film recorder to stack the successive radar returns into a raster format $r(t, nT)$, and then a 2-dimensional space integrating coherent optical processor was used to focus these microwave holograms into the desired image simultaneously in both range and azimuth[26–31]. In this approach the range focal length of the film is a constant, $F_R = v_R^2/b\lambda_o$, where v_R is the range scanning velocity and λ_o is the optical wavelength. However, the focal length in the orthogonal azimuth dimension, or pulse number dimension, nT, is proportional to the range, $F_A = Rv_A^2 c/2V^2\lambda_o\nu_r$, where v_A is the azimuthal scanning velocity, resulting in the range-azimuth coupling phenomenon of the SAR image formation operation. This coupling can be understood by imagining concentric spherical shells around each target of separation $\lambda_r/2 = c/2\nu_r$, which represent isophase surfaces of the round trip radar returns, and considering the intersections of the aircraft's trajectory with these successive shells. The closer target the shells have a larger curvature, and the straight line constant velocity aircraft trajectory intersects more shells per unit time, indicating a higher rate of phase change from pulse to pulse than for a target with a larger range. The phase history of the returned pulses is approximately a quadratic function of the pulse number, and the phase curvature is scaled inversely proportional to the target range, as is illustrated in Figure 3.3.2. This produces a linear dependence of azimuth focal length on range, since focal length, F, is inversely proportional to the quadratic phase factor of an equivalent lens, $e^{i\frac{\pi}{\lambda_o F}x^2}$. When the radar detects the returned pulses

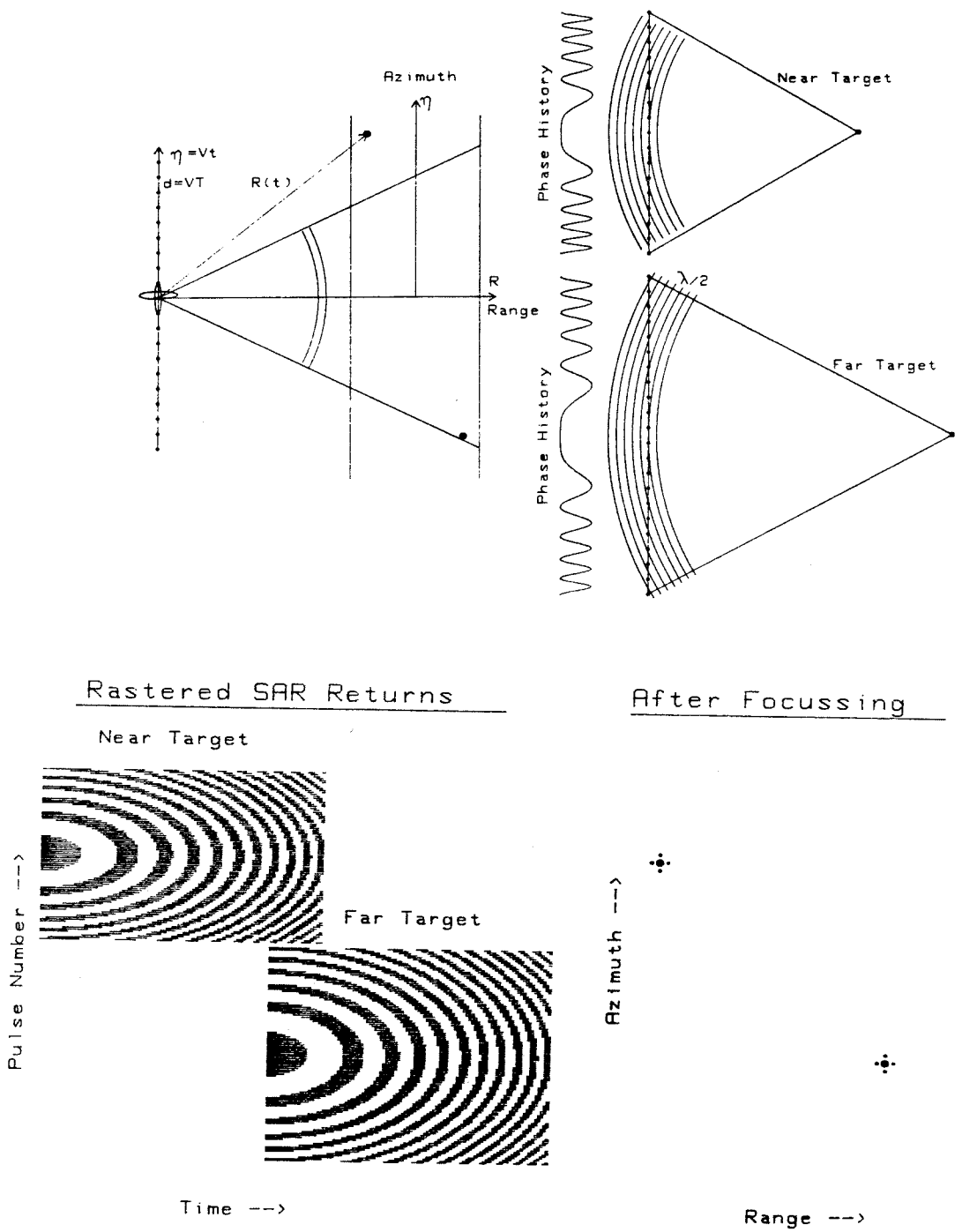


Figure 3.3.2 Top view of a SAR imaging geometry, azimuthal phase histories of near and far targets, isophase contours surrounding a target intersecting with the aircraft trajectory, the two dimensional range dependent anamorphic zone plates produced by a SAR film recorder, and the corresponding focussed image.

on the microwave carrier frequency, they are mixed with a coherent local oscillator down to a near baseband frequency, in order to allow raster scanning onto a photographic medium. However, the chirps are not mixed with the original microwave carrier clear down to baseband, because in this case the conjugate sideband term (or virtual image) diffracted by the hologram can not be filtered from the desired image in the Fourier plane of the optical processor. A spatial carrier frequency is used in the recording process, so that the microwave records represent an off-axis optical hologram, thereby allowing the convenient spatial filtering of the unwanted conjugate sideband diffraction of the real hologram.

3.3.1 TSI processing of synthetic aperture radar data

The SAR processing problem is ideal for casting into the framework of a real time TSI optical signal processing system. This is because the incoming SAR echoes are in the form of a very long 1-dimensional signal consisting of high bandwidth segments each of which must be independently range focussed, with slow azimuth variations from pulse to pulse, which must be independently correlated against the appropriate range dependent impulse responses. As described above, the processing is often performed in a rastered coordinate space, which is the conventional form of TSI image processors, but in the SAR case the raster scanning of the microwave hologram is naturally accomplished by the motion of the aircraft rather than through the scanning operation performed by a video camera. In the TSI approach, the SAR data must first be range focussed via a space integration before the time integrating azimuth processing can be attempted. This is because the required azimuth filter is range dependent, so a cascadable type of processor must be utilized, which is the most powerful of the TSI processor architectures. The output of the range correlation must be displayed spatially, so that it can address a multichannel CCD correlator array, which suggests a pulsed laser space integrating

AO correlator, or a time integrating correlator architecture[41]. However, because of the severity of the bias buildup associated with 2-D time integrating techniques, only space integrating range correlation techniques will be considered here. The SI correlator could be a free space chirp correlator for chirp pulses, or a Vander Lugt holographic correlator could be used for arbitrary pulse codings. The SAR echoes usually comes in high bandwidth segments with large dead times between radar pulses, so the AO processor can take in very high bandwidth radar data, and produce much lower bandwidth focussed imagery at the output of the two dimensional TDI CCD detector array. The continuous scrolling nature of the TDI CCD array, is combined with the strip mapping capabilities of a SAR system to produce continuous swaths of focussed radar imagery for the duration of the aircraft's flight, with a latency of only one frame delay, which is very convenient for on board real time image preview applications.

The TSI processor originally designed for real-time SAR processing consists of a pulsed source, free space, range chirp correlator, and an orthogonal, range dependent multichannel TDI azimuthal phase history correlator array[33–36], as shown in Figure 3.3.3. Each radar return is applied to the AOD transducer and the laser is synchronously pulsed with a narrow pulse, thereby freezing the range chirps within the AOD aperture. The diffracted light propagates through a distance $z_c = -v_a^2/\lambda_o b$, where v_a is the acoustic velocity, which brings the range chirps into sharp focus. This free space propagation accomplishes a complex space integrating chirp correlation on each successively entered radar return, and the cylindrical lens collimates these uniformly expanded focussed chirps in the orthogonal dimension. At the chirp focus plane a mask is inserted which contains the appropriate range dependent azimuthal phase histories as appropriately scaled y dependent transmittance variations, so each focussed range chirp illuminates a single column of the mask which contains the appropriate azimuth chirp. From pulse to pulse the range

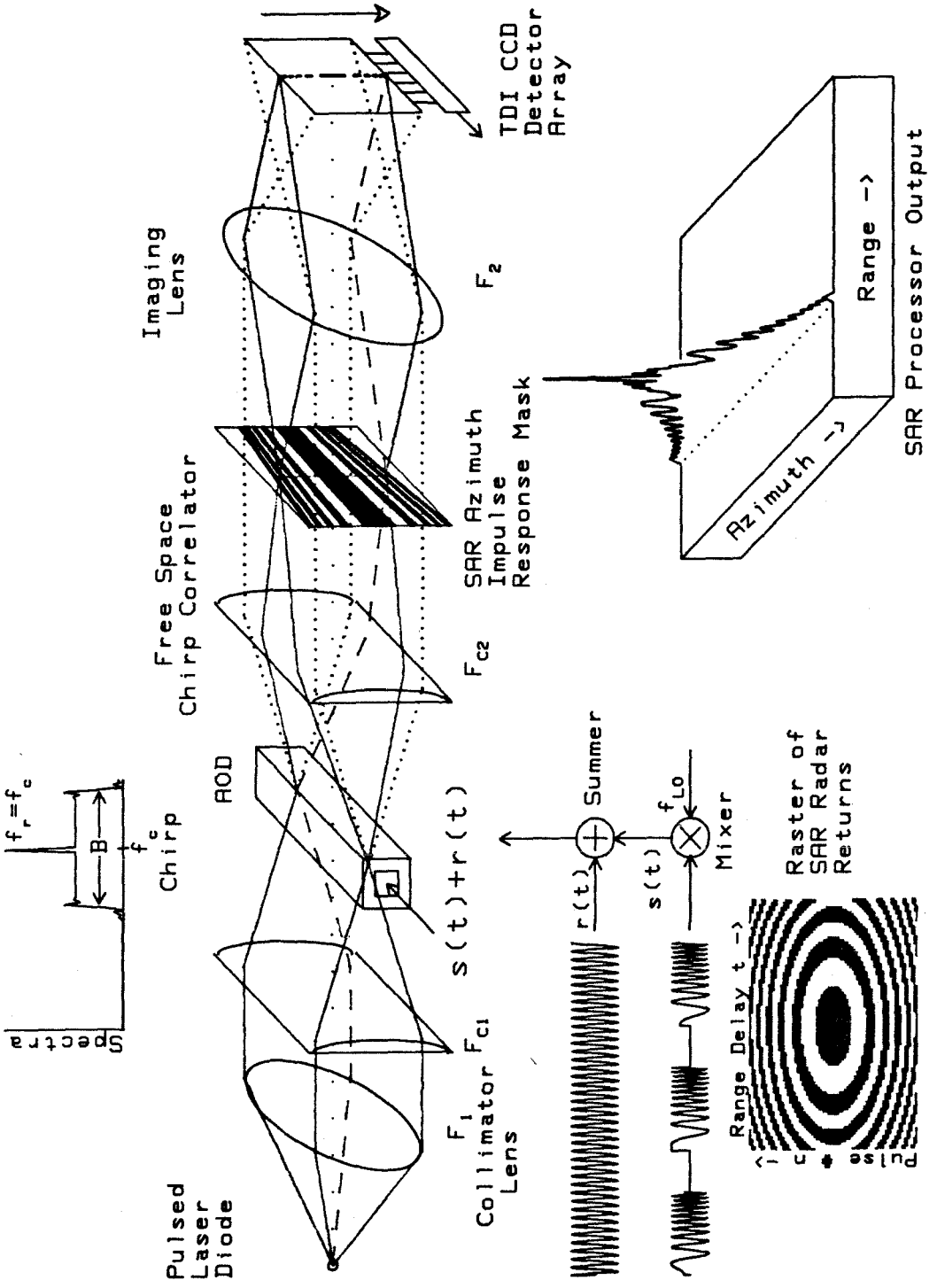


Figure 3.3.3 Schematic representation of the TSI AO/CCD SAR processor.

delay is essentially constant, so the focussed range chirp illuminates the same column of the mask. In order to detect the azimuth phase histories it is necessary to use an interferometric technique, which can be simply accomplished by adding a synchronous CW local oscillator to the radar return before applying to the AOD. For a single target the time and space varying intensity pattern observed at the chirp focus plane will be the interference between the plane wave reference and the range focussed chirp. From pulse to pulse the quadratically phase modulated sequence of radar returns will produce an intensity modulated sampled chirp with a range dependent temporal chirp rate, and the entire vertical slit of range focussed light will uniformly beat on and off accordingly. This temporally modulated slit of light will multiply by the spatially varying mask transmittance function corresponding to that range bin. Imaging this time and space dependent intensity pattern onto the CCD and detecting in a scrolling coordinate frame will produce a real correlation between the time dependent interferometrically detected azimuth history of a given point target with the mask reference function in that column. This multichannel TDI correlation completes the azimuthal focussing operation of a point scatterer, resulting in a well focussed SAR image in both range and azimuth, plus additional bias terms due to the interferometric detection operation, and a spurious azimuth misfocussed term due to the multiplicative implementation. The radar returns from a large number of point scatterers can be processed in parallel because of the linearity of the SAR system, however the cumulative bias and conjugate sideband due to the TSI optical processor must be removed. The unwanted terms can be removed with a carrier demodulation scheme, with the carrier in either the range or azimuth dimension, at the expense of a corresponding loss of resolution in the dimension chosen for the carrier. Alternatively, an additive TSI interferometer can be constructed which eliminates the conjugate sideband term, even without a spatial carrier, thereby increasing the obtainable image resolution. This allows an auxil-

iary CCD to be used to compute the signal dependent bias and directly subtract it, producing the real part of the desired SAR image or a special purpose bias removal CCD structure can be employed.

The analysis of the TSI SAR processor begins with the results of the space integrating free space chirp correlator of section 2.2.2.6, and combines this with a multichannel version of the TDI correlator array presented in Section 2.3.3. The pulse width of the laser must be narrow enough so that the focussed chirp scans less than a CCD pixel during the pulse duration, and so that any internal fringe structure moves by much less than half a period during the pulse duration. These conditions can be met by requiring that the laser pulse width τ_0 is less than the inverse bandwidth of the highest frequency components interferometrically detected at the chirp focus, $\tau_0 < \Delta B = |f_r - (f_c \pm B/2)|$, where f_r is the reference frequency, and f_c is the chirp carrier frequency in the AOD. For a single point scatterer at position (ξ_0, η_0), the optical field observed on the nth laser pulse at the plane of the chirp focus will be the sum of the plane wave reference and the focussed chirp.

$$a(x,y,t) = p(nT - t_0)q(y)\frac{d}{\sqrt{\lambda_o z_c}}\int_A \left[\tilde{S}(t - x'/v_a) + ge^{i2\pi f_r(t-x'/v_a)}\right] e^{i\frac{\pi}{\lambda_o z_c}(x'-x)^2}dx$$

$$a(x,y,nT) \propto dq(y)\left[\sigma A(nT)\sqrt{T_c B}\,\mathrm{sinc}\left[B\left(\frac{x}{v_a} - \frac{f_c}{b} - \frac{2R_0}{c}\right)\right]\right. \tag{3.3.5}$$

$$\left. e^{i2\pi\nu_r\frac{(nd_p-\eta_0)^2}{R_0 c}} e^{i2\pi f_c(nT-x/v_a)} + ge^{i2\pi f_r(nT-x/v_a)}\right]$$

In this expression $p(nT - t_0)$ is the nth laser diode pulse of width τ_0, and normalized amplitude, which has been delayed from the nth radar pulse by a time t_0 to allow the radar return to fully occupy the AOD aperture A. The orthogonal profile of the laser, $q(y)$, is approximately constant after recollimation by the cylindrical lens, and d has absorbed all of the constants and arbitrary phase factors. The gain g represents the relative amplitude of the reference to the radar return, and for the optimum modulation depth for a single target $g = \sigma\sqrt{T_c B}$, so that at the chirp focus the reference and chirp have the same amplitude. The diffracted

plane wave due to the reference tone is assumed to remain a plane wave after the Fresnel diffraction, and edge effect ripples are neglected. The carrier frequency $f_c = \nu_r - f_{LO}$ is the resulting center frequency of the range chirp after mixing down from the microwave carrier frequency ν_r, and this carrier must be coherent with the transmitted chirp, and near the AOD center frequency. The coherent reference local oscillator frequency is chosen to be displaced from the chirp carrier by a harmonic of the laser PRF, plus a small offset, $f_r = f_c + k/T + f_a$, so that from pulse to pulse the relative phases of the two carriers varies at an adequately sampled temporal frequency $f_a < 4/T$. The intensity incident on the reference mask placed in the chirp focus plane is given by the modulus squared of the incident field, and the integrated intensity over the pulse width is proportional to the pulse duration τ_0.

$$I(x,y,nT) = \int_{(n-\frac{1}{2})T}^{(n+\frac{1}{2})T} |a(x,y,t)|^2 dt \tag{3.3.6}$$

$$\tau_0 d^2 q^2(y) \Big[\sigma^2 A^2(nT) T_c B \mathrm{sinc}^2[B(x/v_a - f_c/b - 2R_0/c)] + g^2$$

$$+ g\sigma A(nT)\sqrt{T_c B}\,\mathrm{sinc}[B(x/v_a - f_c/b - 2R_0/c)]$$

$$2\cos\left[2\pi\left(\nu_r \frac{(nd_p - \eta_0)^2}{R_0 c} + (f_c - f_r)x/v_a + f_a nT\right)\right]\Big]$$

The first two terms are the usual bias terms associated with interferometric detection in a TSI system consisting of a signal dependent bias ridge and an approximately uniform bias. The focussed range chirp has the expected width $1/B$ scaled into the spatial output coordinate by the acoustic velocity v_a, and the intensity ridge has the usual chirp processing gain $T_c B$, while the interferometric term has an amplitude $g\sqrt{T_c B}$. The last term is the desired term which contains the azimuthal phase history as a time dependent intensity modulation of the amplitude of the range focussed sinc function, on a sampled temporal carrier f_a, and a spatial carrier in the range dimension of spatial frequency $(f_c - f_r)/v_a$. Only one of these two carriers needs to be utilized in order to appropriately represent the complex azimuthal information in the resulting unipolar intensity distribution. However, if

both carriers are eliminated then the unwanted spurious sidelobe term will produce a large chirping contribution which can not be removed, that will seriously degrade the processor SNR in a multitarget environment, as is typical in SAR imaging applications.

When the range carrier is utilized, then a spatial mask is placed in the chirp focus plane that also has a range carrier in the intensity transmittance of the mask, as illustrated in Figure 3.3.4, and f_a is set equal to zero.

$$T(x,y) = \frac{1}{2} + \frac{1}{2}\cos\left[2\pi\left(f_x x + \frac{y^2}{b'x + x_0}\right)\right] \tag{3.3.7}$$

The factor b' represents the effect of range azimuth coupling scaled into the spatial representation of the processor coordinate space, and will be calculated in order to produce the desired azimuth focussing operation at all range bins. The offset factor x_0 represents the appropriate midrange coordinate corresponding to the center of the CCD, and the CCD range coordinate x is symmetric about this offset position, varying from $-W/2$ to $+W/2$. The CCD is composed of K columns of pixels, so the overall CCD width is $W = K\Delta x$, where Δx is the width of an individual pixel in range, which has a corresponding range resolution $\delta_r = 2\Delta x v_a/c$. The vertical spatial coordinate y will be centered, corresponding to a side looking offset position, and it extends from $-H/2$ to $+H/2$, where the CCD height is equal to $H = N\Delta y$. The spatial carrier in the x dimension must be chosen large enough so that the sum frequency spatial carrier generated on the CCD is large enough to allow a bandpass filter to separate the image modulation on this carrier from the bias terms and the difference term. This can be accomplished when the mask spatial carrier is greater than or equal to the spatial carrier produced by the interference of the radar returns and the local oscillator, and the local oscillator is outside the bandwidth of the chirps, $f_r - f_c > B/2$, and for convenience one might choose $f_x = (f_c - f_r)/v_a > B/2v_a$. This condition necessarily implies that the range focussed sinc illuminates at least one full cycle of the mask range carrier, and

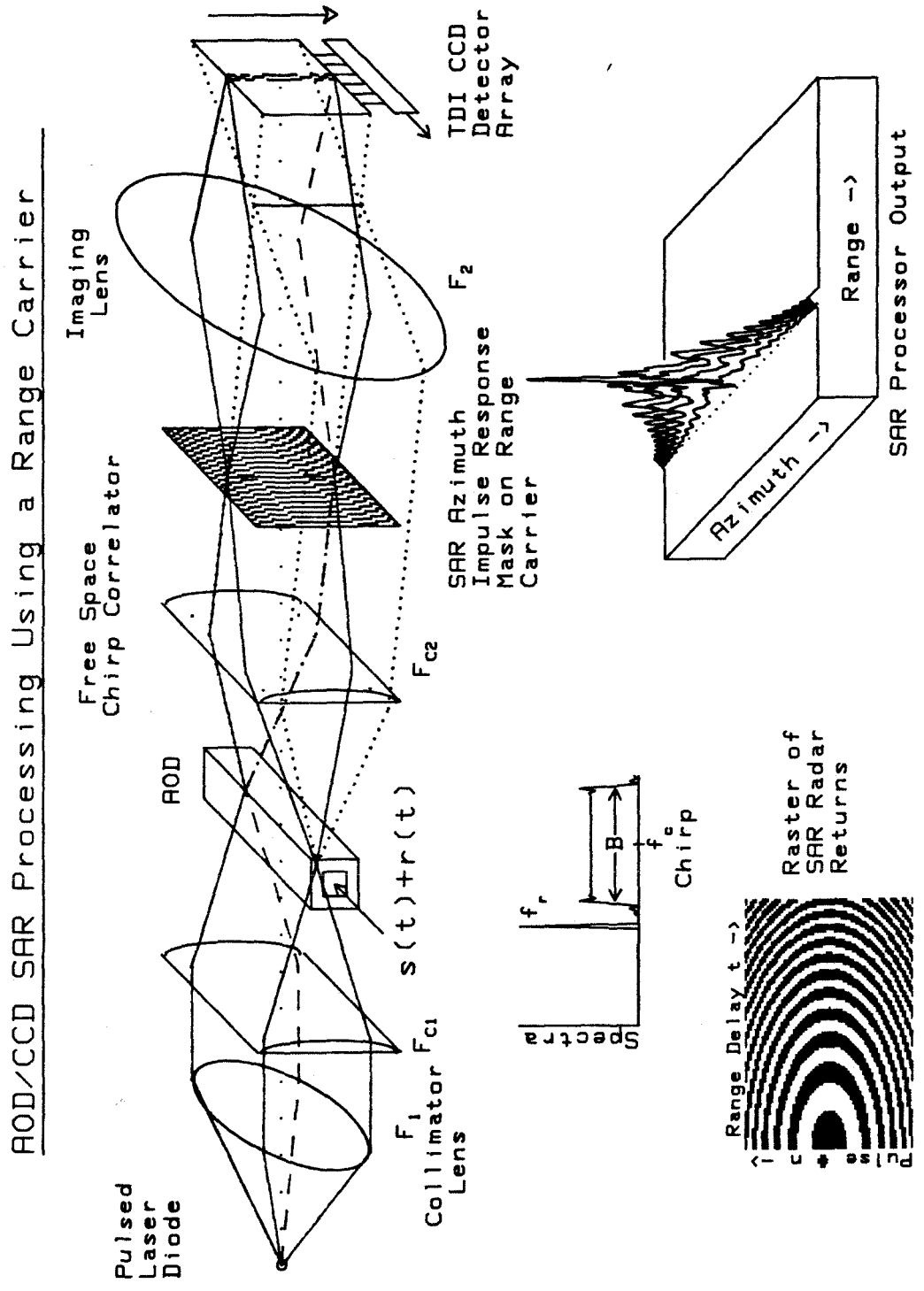


Figure 3.3.4. multiplicative TSI AO/CCD processor architecture utilizing a range carrier.

contains at least one full cycle of the interferometrically produced range carrier. It is then necessary to adequately sample the frequency doubled carrier produced on the CCD, which requires at least 2, and hopefully 4, pixels of width Δx per carrier cycle. The range focussed sinc profile will contain 2 full cycles under the main lobe, therefore requiring 8 CCD pixels per range resolution element, so that $v_a/B > 8\Delta x$. The two full cycles of the range carrier under the sinc profile are necessary to adequately separate the spatial frequency of the interferometrically detected term from the incoherent sinc^2 ridge. The range carrier decreases the number of resolvable range bins from the number of CCD pixels in x by a factor of 8, giving $K/8$ independent range bins, with effective slant range resolution at the 4dB point of $\delta_r = 8\Delta x v_a/c$.

Alternatively, the mask can be fabricated with a spatial carrier in the y dimension, which can be used to produce an azimuth carrier, and the corresponding architecture is schematically illustrated in Figure 3.3.5. This azimuth carrier approach is at the expense of decreasing the processing gain and resolution in the azimuth dimension, and it requires the radar to massively oversample the azimuth bandwidth. This requires a higher PRF or the use of a wider antenna, or operation at a closer range than might otherwise be desired, but the processor can easily accommodate PRFs up to 15KHz, so this is not a serious disadvantage. The mask on an azimuth carrier is given by an equation similar to Eqn. 3.3.7, with the x carrier replaced by the y carrier.

$$T(x,y) = \frac{1}{2} + \frac{1}{2}\cos\left[2\pi\left(f_y y + \frac{y^2}{b'x + x_0}\right)\right] \tag{3.3.8}$$

In this case the interferometric range carrier is eliminated by setting $|f_c - f_r| \ll v_a/W$, so that no spatial fringe pattern is detected in the range dimension over the width W of the CCD. The azimuthal carrier frequency must be larger than any frequency within the processing aperture, thereby requiring that the mask is single sided, and the more stringent constraint that the temporal carrier is greater than

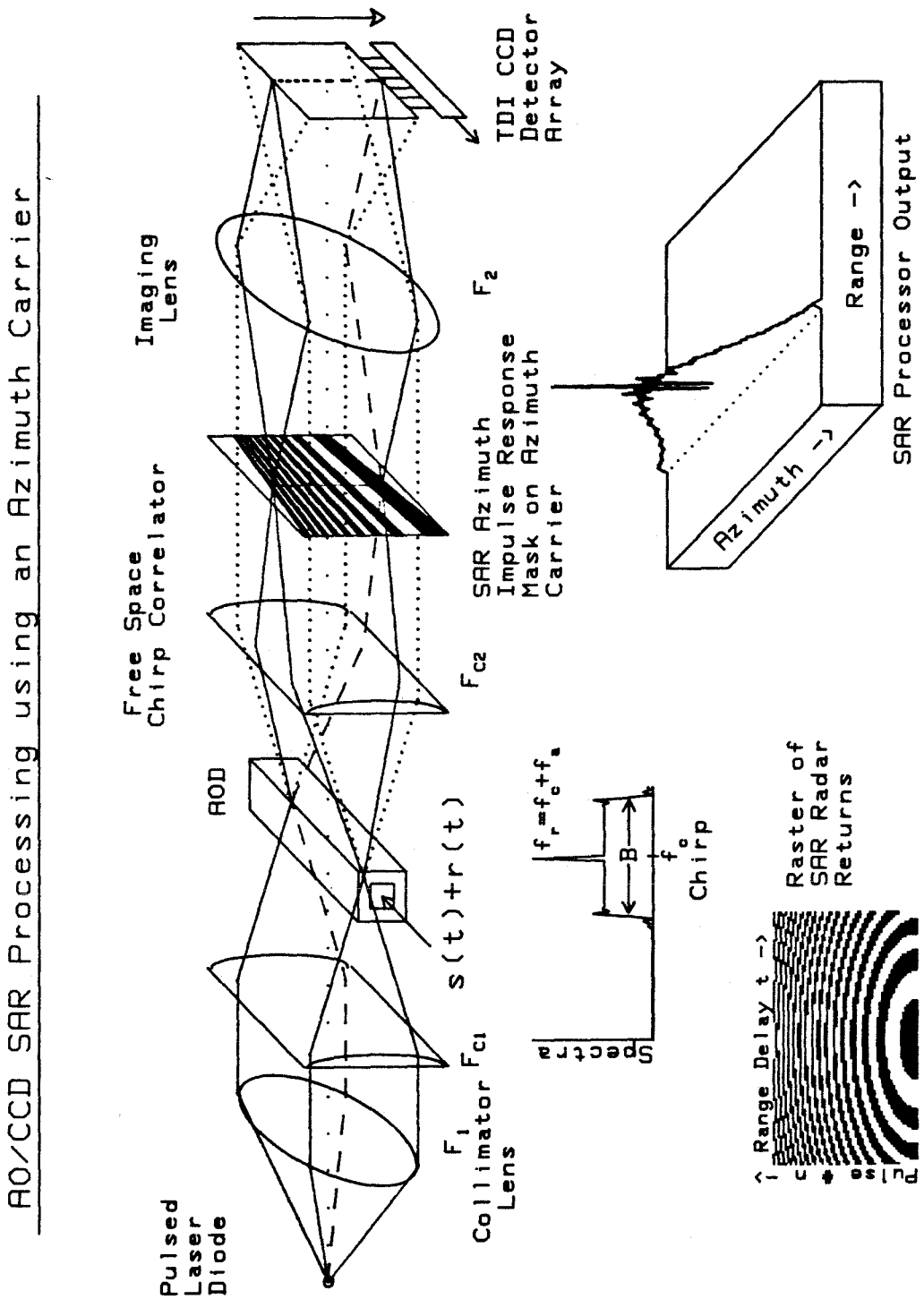


Figure 3.3.5. multiplicative TSI AO/CCD processor architecture utilizing an azimuth carrier.

1.5 times the signal bandwidth will result in an octave bandwidth mask, which will more effectively eliminate the contribution of the unwanted conjugate side-band term. For a synthesized aperture $L_s = Nd_p \leq L$ containing N radar pulses, corresponding to N TDI CCD stages, the azimuth reference frequency should be greater than $VL_s/R_0\lambda_r < f_a$. Using an azimuth carrier frequency higher than the highest azimuth frequency produced within the main lobe of the radar footprint, $f_a > f_{max} \approx V/D_a$, is an even better limit that eliminates any frequency fold over effects. Simultaneously, the azimuth carrier frequency must be less than half the Nyquist frequency of the azimuthal sampling rate, so that the upshifted azimuth information is adequately sampled in the processor coordinate space, $f_a < 1/4T$. The lowest frequency of the carrier modulated azimuth information should be well above DC, and the highest frequency should be well below the Nyquist limit to minimize unwanted within band sidelobes due to the real correlation operation.

In the following analysis, both cases of range and azimuth carrier will be treated simultaneously, using a reference mask that contains both carriers, however, only one carrier would be used in practice. The product of the incident intensity with the mask transmittance is imaged onto a TDI CCD shifting by a pixel of height Δy each l laser pulses, where l is the number of samples which are incoherently averaged, or presummed, before azimuth processing is performed. The effective velocity of the TDI CCD is therefore given by $v_{CCD} = \Delta y/lT$, and this scale factor must be used to match the temporal carrier, f_a, with the y spatial carrier of the mask through the equation $f_y = f_a/v_{CCD} = f_a lT/\Delta y$. Clearly this spatial carrier must be adequately sampled by the CCD pixels of height Δy in y, and this implies that $f_y < 1/4\Delta y$, which is equivalent to the previously stated temporal sampling limit imposed upon f_a. For either choice of carrier direction, the detection in the moving coordinate frame of the TDI CCD will contain a term that performs the desired complex azimuthal phase history correlation for each range bin of pixel width Δx

on the CCD. Ignoring the presumming capability by setting $l = 1$, the operation of the multichannel space variant TDI azimuth correlator array can be described with the usual shift and add notation.

$$I(x,mT) = \sum_{n=m-N}^{m} T\big(x,(m-n)\Delta y\big) I\big(x,(m-n)\Delta y, nT\big) \tag{3.3.9}$$

$$= \tau_0 \frac{d^2}{2} \sigma^2 T_c B \mathrm{sinc}^2[B(x/v_a - f_c/b - 2R_0/c)] \sum_{n=m-N}^{m} A^2(nT) q^2((m-n)\Delta y)$$

$$+ N\tau_0 \frac{d^2}{2} g^2 \bar{q}^2 + \tau_0 dg\sigma \sqrt{T_c B}\,\mathrm{sinc}[B(x/v_a - f_c/b - 2R_0/c)]$$

$$\sum_{n=m-N}^{m} A(nT) q^2((m-n)\Delta y) \cos\left[2\pi\left(f_x x + f_a nT + \frac{(nd_p - \eta_0)^2}{R_0\lambda_r}\right)\right]$$

$$\cos\left[2\pi\left(f_x x + f_a(m-n)T + \frac{((m-n)\Delta y)^2}{b'x + x_0}\right)\right]$$

The bias in the mask produces the first two terms which represent the TDI averaging of the uniform reference beam intensity, and the TDI convolution between the squared antenna footprint and the laser beam apodization in the y dimension, which will result in a slowly time varying bias within the coarse frequency bin corresponding to the point scatterer at range R_0. The final term represents the desired azimuthal phase history autocorrelation that must be performed on each range bin, and it is apodized by both the range dependent antenna footprint and the laser beam profile in y. The product of the cosinusoidal interferometrically generated intensity pattern with the cosinusoidal mask transmittance will generate sum and difference terms, the sum term appears on a frequency doubled range carrier, or on an m dependent azimuthal carrier, and is the desired azimuthal focussing term, while the difference term represents a spurious misfocussing contribution. The dimension that is chosen to represent the carrier frequency will suffer a decrease in the available resolution capabilities of the detector array. Filtering in the range dimension must remove the contribution due to the tightly focussed incoherent ridge, while the bias in azimuth is slowly varying, so that more carrier cycles will be needed within a blur spot if the range carrier is utilized.

The scaled mask range-azimuth coupling factor needed to fabricate the reference mask can be computed by setting the coefficients of n^2 in the spatial and temporal cosines equal, or by setting the number of fringes over the processing aperture equal to the number of fringes over the CCD height for all ranges. The range swath processed over the CCD width is given by $\Delta R = cW/2v_a$, and this is centered about the gross range delay at the center of the AOD and CCD, R_c, which is determined by the time delay $t_0 = T_c/2 + 2R_c/c + A/2v_a$ between each radar pulse and the corresponding laser pulse. In terms of these variables the constants which specify the mask range azimuth coupling become quite simple.

$$\begin{aligned} b' &= \left(\frac{H}{L_s}\right)^2 \frac{\lambda_r \Delta R}{W} = \left(\frac{\Delta y}{d_p}\right)^2 \frac{\lambda_r \Delta R}{W} = \frac{v_{CCD} c^2}{2V^2 \nu_r v_a} \\ x_0 &= \left(\frac{H}{L_s}\right)^2 \lambda_r R_c = \left(\frac{\Delta y}{d_p}\right)^2 \lambda_r R_c \end{aligned} \tag{3.3.10}$$

This completely specifies the functional form of the image of the mask on the CCD in terms of the CCD height H and width W, the processed synthetic aperture width $L_S = Nd_p$, the midrange R_c, and the range processing width ΔR. However other scale factors, such as imaging system magnification, and appropriate pixel size scale factors and sampling effects on the mask and CCD must be accounted for in an experimental implementation.

The output of the SAR processor will consist of a number of bias terms near baseband, the conjugate sideband azimuth misfocussed term, and the desired range and azimuth focussed image riding on a high frequency spatial carrier. The appropriate carrier demodulation will separate the desired term from the other terms, however they will use up some of the available resolution and dynamic range of the CCD, especially in a multi target environment. For the point target under consideration here, the processor output can be greatly simplified by considering only the desired term which is modulated by the spatial carrier, and assuming that the other terms have been removed by electronic bandpass filtering of the CCD output, and

neglecting the apodizations due to the radar footprint, the laser beam profile in y, and additional CCD blurrings due to resolution limitations.

$$
\begin{aligned}
I(x,mT) =& \frac{\tau_0 dg}{2}\sigma\sqrt{T_c B}\mathrm{sinc}[B(x/v_a - 2(R_0 - R_c)/c)] \\
& \sum_{n=m-N}^{m} \cos\left[2\pi\left(2f_x x + f_a T m - \frac{d_p^2}{R_0\lambda_r}\left[m^2 - \frac{\eta_0^2}{d_p^2} - 2n\left(m - \frac{\eta_0}{d_p}\right)\right]\right)\right] \\
\approx& \frac{\tau_0 dg}{2}\sigma\sqrt{T_c B}\mathrm{sinc}[B(x/v_a - 2(R_0 - R_c)/c)] N\mathrm{sinc}\left[N\frac{2d_p^2}{R_0\lambda_r}(m - \eta_0/d_p)\right] \\
& \cos\left[2\pi\left(2f_x x + f_a T m - \frac{d_p^2}{R_0\lambda_r}\left(m^2 - \frac{\eta_0^2}{d_p^2}\right)\right)\right] \qquad (3.3.11)
\end{aligned}
$$

This demonstrates that the TSI AO/CCD SAR processor has the capability to achieved a good focussing operation in both range and azimuth dimensions, for all range bins. The azimuth resolution that has been achieved is linearly proportional to the range, and in ground coordinates η the resolution is $\delta_a = R_0\lambda_r/2Nd_p = R_0\lambda_r/2L_s$, which is equivalent to a Rayleigh resolved real aperture equal to twice the range dependent processed aperture width. The reason that the resolution gets worse for farther ranges, and is not constant with range as was mentioned previously, is because only N pulses are being coherently accumulated, and for farther ranges the footprint is wider and more pulses should be included within the synthesized aperture. The neglected apodization effects will degrade the achievable resolution below this idealized limit, but will also help to minimize the sidelobe amplitudes. Similarly, the interferometrically detected range focussed sinc function has the expected range resolution width given by $\delta_r = c/2B$, which is scaled into the spatial processor space by the acoustic velocity. The two dimensional peak is on a carrier in either the range or azimuth dimensions, and the additional quadratic carrier in azimuth has a curvature which goes to zero at the position of the peak, and contains less than a fringe over the width of the azimuth peak, so it can be neglected.

The amplitude of the detected scatterer at $x = 2R_0v_a/c, m = \eta_0/d_p$ is proportional to the laser power and optical efficiency d, the target reflectivity σ, the laser pulse width τ_0, the reference beam amplitude g, the number of TDI gain stages N, and the square root of the range chirp time bandwidth product $\sqrt{T_cB}$. However, the important parameters that will affect the resulting image quality are the modulation depth and resulting SNR that is achieved on a CCD with a limited dynamic range. The modulation depth requires knowledge of the amplitude of the various bias terms within a given range bin. The unwanted terms which use up the available detector dynamic range include the reference beam bias, the signal dependent range focussed intensity ridge due to each scatterer, and the conjugate sideband azimuth misfocussed term, all of which can be removed by bandpass filtering. The modulation depth can be defined as the peak to peak interferometric signal swing divided by twice the bias, which is equivalent to $m = (I_{max} - I_{min})/(I_{max} + I_{min})$, and is also equivalent to the signal to bias ratio (SBR).

$$\begin{aligned} m &= \frac{\text{signal swing}}{2 \times \text{bias}} = \frac{s}{b} = SBR \\ &= \frac{2g\sigma\sqrt{T_cB}N}{Ng^2 + N\sigma^2T_cB} \end{aligned} \tag{3.3.12}$$

It is interesting to note that the number of TDI gain stages N cancels in this expression, so the SBR can not be improved by using larger CCD arrays, because both signal and bias terms build up linearly with the number of TDI stages. For a single target the modulation depth can be optimized by setting $g = \sigma\sqrt{T_cB}$, and the resulting optimum modulation depth is $m_{opt} = 1$. In this case the radar image can theoretically achieve an SNR equal to the detector dynamic range, but limits on the pulsed laser coherence will probably reduce this considerably.

In a distributed multitarget environment, the modulation depth of each target is decreased by the bias contributions of all other targets within that range bin that are simultaneously illuminated by the antenna footprint. For M_k targets within the

illuminated footprint in the kth range bin, with average reflectivity σ, the resulting modulation depth of an average target is decreased.

$$m = \frac{2g\sigma\sqrt{T_c B}}{g^2 + M_k \sigma^2 T_c B} \tag{3.3.13}$$

In this case the optimum reference beam amplitude is found to be increased by the square root of the number of illuminated targets within a range bin, $g = \sigma\sqrt{T_c B}\sqrt{M_k}$, so it is range dependent and can not be optimized simultaneously for all range bins unless the reference beam amplitude is appropriately modulated. The resulting optimum modulation depth is decreased to $m_{opt} = 1/\sqrt{M_k}$. The compromise gain should be set equal to the expected number of scatterers within the various range bins, which has an expected value of $\bar{M}$ at R_c. Since the number of illuminated scatterers grows linearly with range as the footprint widens, $M_k \propto k$, a square root increase of g with time would be a sensible choice, $g(t - nT) = \sigma\sqrt{T_c B}\sqrt{\bar{M}}\sqrt{2t/c}$, and at R_c the reference beam and expected signal power should occupy about one quarter of the CCD full well each. The achievable dynamic range is found by setting the maximum signal plus bias equal to the CCD full well capacity, and the resulting SAR image dynamic range is given by the CCD dynamic range, DR_0, times the fraction of that dynamic range that the signal swing occupies.

$$DR = DR_0 \frac{2s}{s+b} = DR_0 2\frac{s/b}{1+s/b} = 2DR_0 \frac{1/\sqrt{M_k}}{1+1/\sqrt{M_k}} = \frac{2DR_0}{\sqrt{M_k}+1} \tag{3.3.14}$$

The image dynamic range decreases from the CCD dynamic range inversely proportional to one plus the square root of the number of illuminated targets within a range bin, so for 100 targets within a given range bin the image dynamic range is decreased by a factor of 11. A typical CCD might have a dynamic range on the order of $10^3 : 1$, and typically at least 100 targets would be illuminated within each range bin, resulting in an image amplitude dynamic range of 91:1 or 39dB, which should be acceptable for many real time applications. It is important to realize that in this type of interferometer the image dynamic range does not fall

inversely proportional to the number of targets within a range bin, which would be disastrous for more than 10 targets. Even more important is the fact that the dynamic range does not depend on the number of targets within the entire image, as it does with purely time integrating approaches[40], since the antenna footprint can easily illuminate 10^5 or more targets. Since M_k is the number of illuminated targets within a range bin and not the number within the processing aperture of a range bin, the processor should be designed to process the entire radar footprint width at midrange, in order to avoid excess bias build up.

3.3.2 Additive approaches to TSI SAR processing

The unwanted conjugate sideband term that results from the multiplicative implementation of the AO/CCD SAR processor can be eliminated by using an additive approach. In this configuration the spatial carrier demodulation can be replaced with a bias subtraction operation, so that the full resolution capabilities of the CCD detector can be achieved in both dimensions. The additive SAR interferometer can be configured as a Mach-Zehnder interferometer as shown in Figure 3.3.6, with one TDI CCD on each beam splitter output port. The appropriate azimuth reference functions are introduced interferometrically by introducing a quadratically curved vertical phase in the reference arm, with a cylindrical lens. However, because of the range azimuth coupling the cylindrical lens will need to be tilted about its y axis, or alternatively a conical lens could be used. The interference between a range focussed chirp that has been collimated in y with the quadratically curved reference wavefront will produce chirping fringe profiles along the columns of the CCD. The lens tilt will introduce an inverse linear dependence on the chirp rate from column to column, so adjustment of the tilt angle will produce the appropriate range azimuth coupling for a given radar geometry. The spatial intensity variation detected within each column acts as the azimuthal impulse response with an

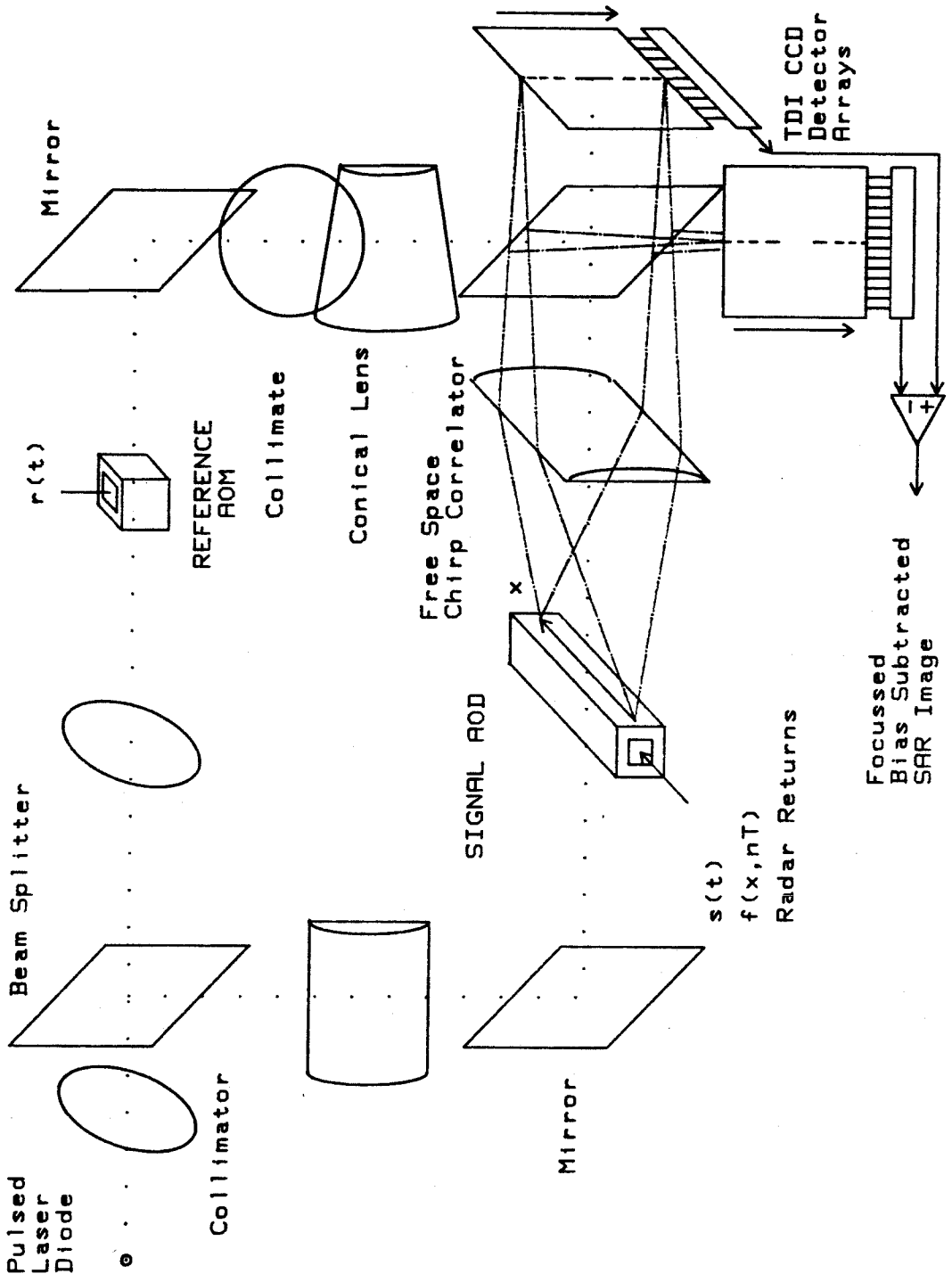
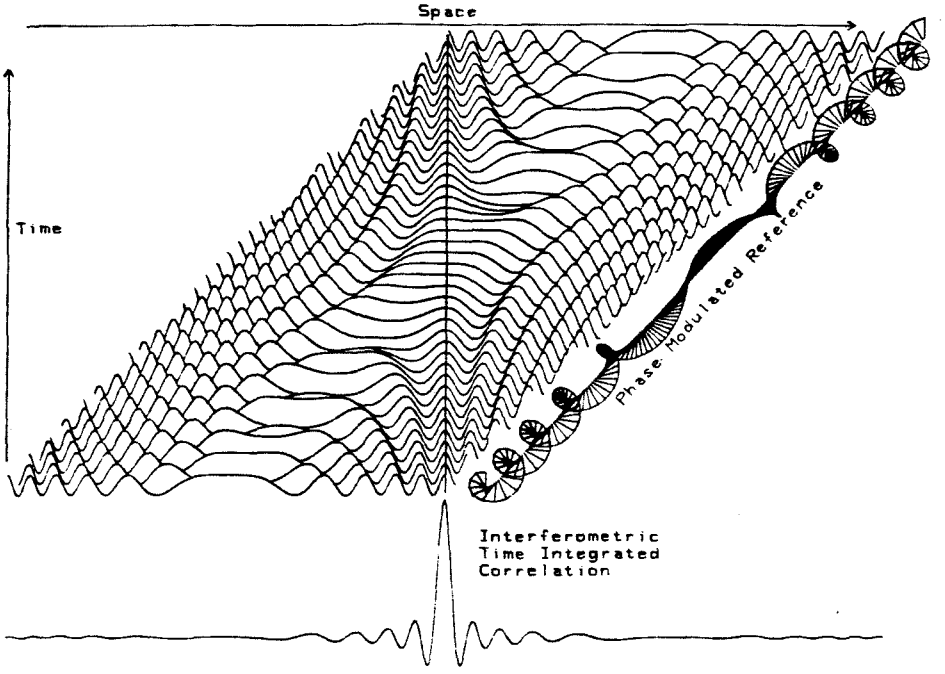


Figure 3.3.6. Mach-Zehnder architecture for AO/CCD SAR processing, with two out of phase TDI CCDs used for bias subtraction.

a)



b)

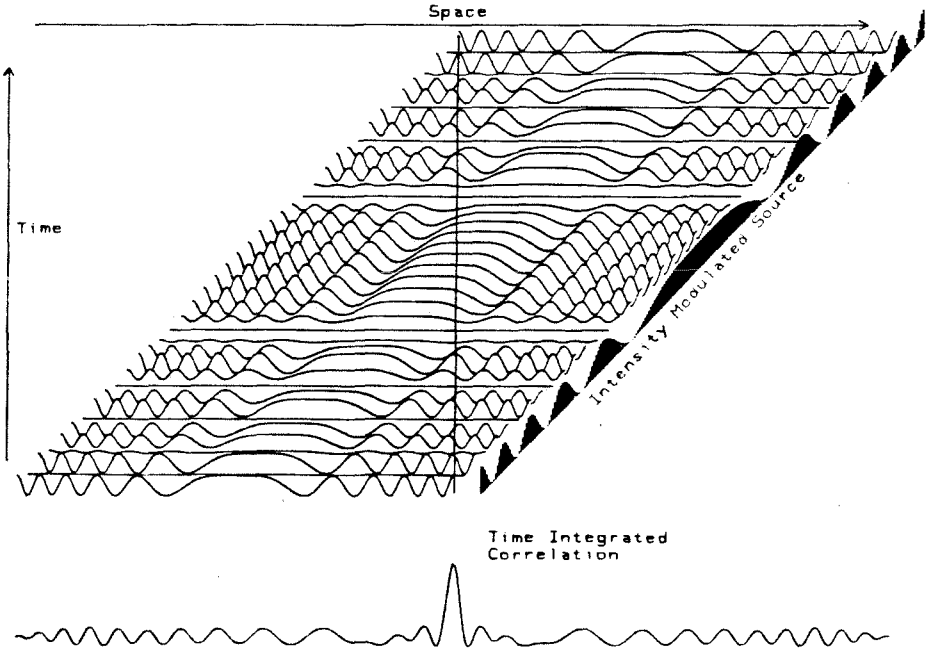


Figure 3.3.7 Time and space dependent intensity profiles observed on a column of the TDI CCD, and the resulting time integrated chirp correlations, a) interferometric case, b) multiplicative approach.

important difference, the absence of the conjugate sideband. This single sideband modulation results in the travelling of the chirping fringes seen on the CCD from pulse to pulse, rather than the blinking on and off of the stationary mask that was obtained in the multiplicative implementation. The pulse to pulse phase shift of the radar returns produces a corresponding shift of the vertically chirping interference profile, as shown in Figure 3.3.7a, and when the spatial and temporal chirp rates are appropriately matched, then one of the fringes will shift by a single pixel from pulse to pulse, producing a large peak on the synchronously scanning TDI detector. The fundamental difference with the multiplicative approach is illustrated in Figure 3.3.7b, where from pulse to pulse the stationary mask is seen to blink on and off, which can be decomposed into the interference of two oppositely travelling chirping fringe profiles. In the multiplicative approach half of the incident energy is in the desired sideband and produces a peak with half the gain of the interferometric approach, but the other half is the wrong sideband which produces fringes travelling in the wrong direction, and this results in the chirping azimuth misfocussed conjugate sideband. A spatial carrier must be used to eliminate this sideband, because a bias subtraction CCD can not compute this term.

When the phase shift upon reflection from the beamsplitter is 90°, as it should be theoretically for an ideal beamsplitter, then the two CCDs shown in Figure 3.3.6, will produce out of phase images that can be directly subtracted in order to remove all the bias contributions. For any phase shift produced by a lossless dielectric beamsplitter on reflection and transmission the two output fields must be exactly 180° out of phase in order to conserve power and obey Stokes relation $r't^* + r^*t' = 0$. In this expression r and r' are the amplitude reflection coefficients and t and t' are the correspoding amplitude transmittances of the beam splitter, and the two interference patterns must be exactly out of phase since $r't^* = -r^*t'$. Thus in the interferometric case the complex processing operation can be accomplished

without the use of a spatial carrier by subtracting the two carefully aligned detector outputs. The full resolution capabilities of the CCD detector array can be achieved in both dimensions with this technique, although the detector area is effectively doubled as well.

$$D(x,m) = I_1(x,m) - I_2(x,m) = \sum_{n=m-N}^{m} |S+iR|^2 - |iS+R|^2 = \sum_{n=m-N}^{m} 4\Im(SR^*) \tag{3.3.15}$$

This bias subtraction processing scheme would result in the real (imaginary) part of the desired SAR image, so if true complex processing is desired a carrier could be included as well, or an auxiliary imaginary (real) part processor could be constructed. The phase shifts of the returns is a uniform random variable from 0 to 2π, so each target will be multiplied by a random gain term equal to the sin of its phase. Only a few targets will disappear completely, and typically the amplitude will decrease by $1/\sqrt{2}$. In the interferometric implementation the gain of the signal term is twice that of the multiplicative approach, and the combination of the out of phase outputs of the two CCDs results in an additional increase by two, so that the resulting image dynamic range is 4 times as large. A carrier can be included to represent the complex information in conjunction with the two out of phase CCDs. In this case the carrier is not used to separate the spectra of the image from the bias and sideband terms, because the bias is directly subtracted. This means that the number of carrier pixels per resolution element can be minimized to approximately 3 or 4, rather than 6 or 8 as required before, giving an additional resolution improvement of the interferometric implementation.

The disadvantage of the Mach-Zehnder structure is its sensitivity to component vibrations, and for this reason an in-line approach is usually preferable. However, the multiplicative in line approach suffers from the problem of the conjugate sideband contribution, so an additive in-line architecture as shown in Figure 3.3.8 can be used instead. In this approach the reference frequency, which is near a harmonic

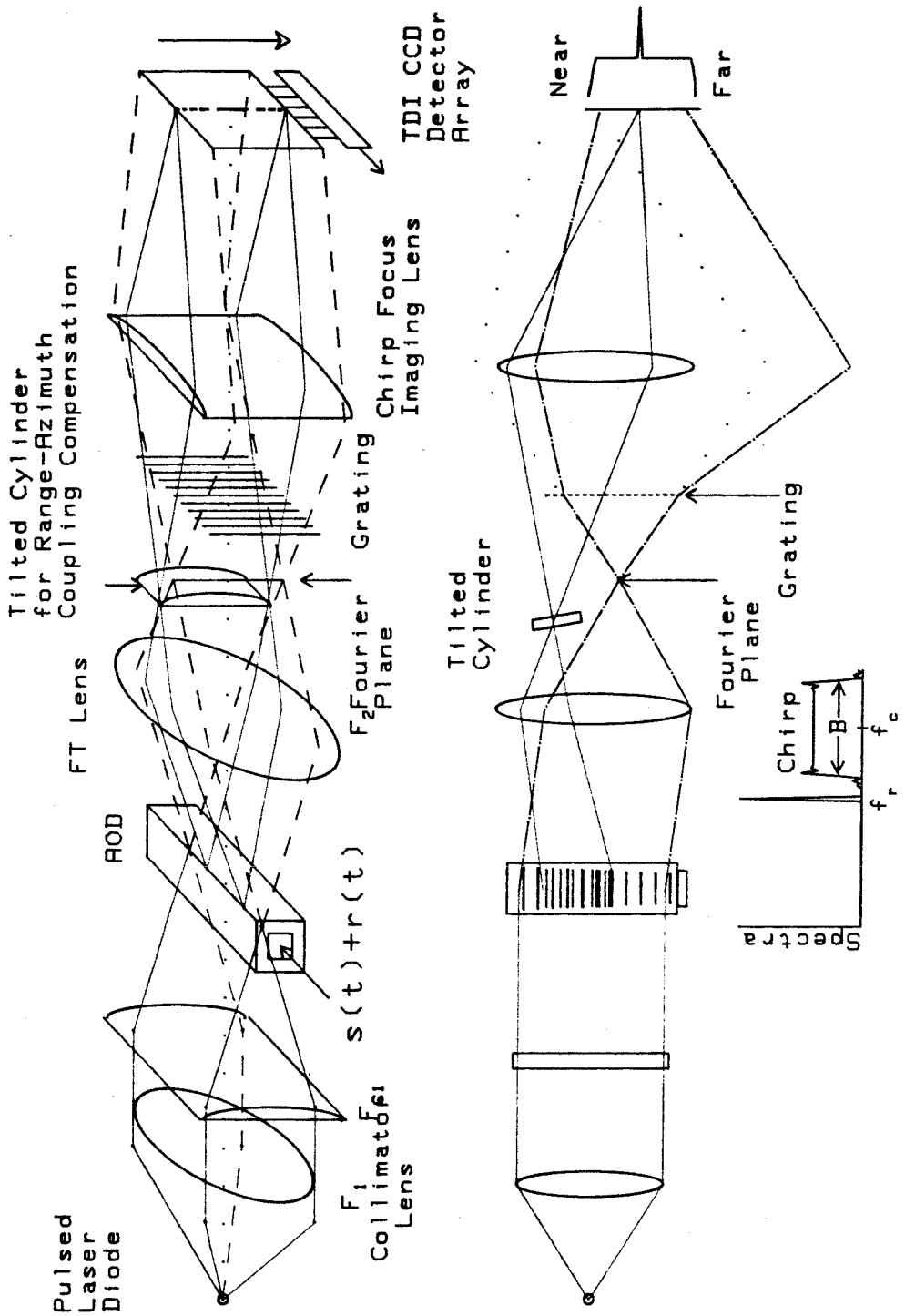


Figure 3.3.8. In line approach to additive AO/CCD SAR processor.

of the laser PRF, is well outside of the chirp bandwidth, and a long focal length chirp is used, so that at the chirp focal plane the reference and focussed chirps are non overlapping. This leads to an inefficient utilization of the available AOD octave bandwidth, B_a, given by $B_c/B_a < 1.5T_c/T_a$, so that the achievable number of independent range resolution elements is decreased from the AOD time bandwidth by a factor of $1.5(T_c/T_a - T_c^2/T_a^2)$. In this case the optimum chirp duration, T_c, is half the AOD width, T_a, and the chirp bandwidth is $.75B_a$, thereby using only 3/8 of the available AOD time bandwidth, but this may be acceptable since AODs are available with much higher time bandwidths than CCDs have pixels. A tilted cylinder, or a conical lens, with power in y is inserted in the chirp focal plane in the region of chirp focus, but not in the region of the reference beam, thereby introducing an appropriate range dependent quadratic phase factor between the focussed radar chirps and the reference. Imaging of the chirp focus plane onto the CCD will not produce any interference in this case because the reference and focussed chirps are not overlapping spatially. The reference beam can be diffracted to overlap with the range focussed signal beam by placing a grating with the appropriate high spatial frequency in the Fourier plane of the lens which is used to image the chirp focus plane onto the TDI CCD. This grating will diffract the reference beam to be collinear with the chirp midband frequency f_c, thereby eliminating the range carrier, and allowing individual range pixels to correspond to single range resolution elements. The signal dependent bias terms can be computed with an auxiliary TDI CCD that picks off the focussed chirps before they are recombined with the reference, and this can be subtracted from the processed output. This leaves only the approximately uniform reference beam bias, which can be eliminated in real time from each processed azimuth swath with a premeasured and stored 1-D reference, since it is constant.

References

[1]. D. Psaltis and B.V.K Vijaya Kumar, Acoustooptic spectral estimation: a statistical analysis, Appl. Opt., vol. 20(4), p. 601 (1981).

[2]. D. Casasent, G. Silbershatz, and B.V.K. Kumar, Acoustooptic matched filter correlator, Appl. Opt., vol. 21(13), p. 2356 (1982).

[3]. D. Casasent et al., Some 2D image processing operations with 1D acousto-optic devices, to appear in SPIE vol. 752, (1987).

[4]. R. Athale et al., Acousto-optic processors for real-time generation of time frequency representations, Opt. Lett., vol. 8(3), p. 166 (1983).

[5]. J. D. Cohen, Ambiguity processor architecture using one dimensional acousto-optic transducers, SPIE vol. 180, p. 134 (1979).

[6]. K. Said and D. Cooper, Crosspath real-time optical correlator and ambiguity-function processor, Proc IEE vol. 120,p. 423 (1973).

[7]. P. Tamura et al., Real time optical computation of the ambiguity function, SPIE vol. 241, p. 104 (1980).

[8]. J.W. Goodman, Linear space-variant optical data processing, in *Optical Information Processing*, S.H. Lee Ed., Topics in Appl. Phys., vol. 48, ch. 6, Springer-Verlag, Berlin (1981).

[9]. B. K. Jenkins, Optical logic systems: Implementations and architectural implications, PhD Thesis, USC (1984).

[10]. D. Casasent, Frequency-multiplexed acoustooptic architectures and applications, Appl. Opt., vol. 24(6), p. 856 (1985).

[11]. J. D. Cohen, Frequency division multiplexing optical processors, SPIE vol. 341, p. 165 (1982).

[12]. D. Psaltis and D. Casasent, Wavelength diversity: an extra dimension in optical processing architectures, SPIE vol. 232-40, (1980).

[13]. T. Bader, Acoustooptic spectrum analysis: a high performance hybrid tech-

nique, Appl. Opt., vol. 18(10), p. 1668 (1979).

[14]. E.G. Paek, C.H. Park, F. Mok, and D. Psaltis, Acoustooptic image correlators,SPIE vol. 638-05, (1986).

[15]. D. Psaltis, E.G. Paek, and S. Venkatesh, Acousto-optic/CCD image processor, Proc. Int Opt. Comp.Conf. IEEE cat 83CH1880-4, (1983).

[16]. D. Psaltis, Two dimensional optical processing using one dimensional input devices, IEEE Proc., vol. 72(7), p. 962 (1984).

[17]. D. Psaltis, Optical image correlation using acoustooptic and charge-coupled devices, Appl. Opt., vol. 21(3), p. 491 (1982).

[18]. R. Athale and J. Lee, Optical systems for efficient triple-matrix-product processing, Opt. Lett., vol. 8(11), p. 590 (1983).

[19]. D. Psaltis, Image processing using acoustooptics, Proc. EOSD Conf., Anaheim Ca., (1981).

[20]. D. Psaltis and D. Casasent, Time and space integrating spectrum analyzer, Appl. Opt., vol. 18(19), p. 3203 (1979).

[21]. K. Wagner and D. Psaltis, Real-time computation of moments with acoustooptics, SPIE vol. 352-19, (1982).

[22]. D. Casasent and D. Kessler, Optical signal processing using long coded waveforms, Opt. Comm., vol. 17(3), p. 242 (1976).

[23]. G. Silbershatz and D. Casasent, Hybrid time and space integrating processors for spread spectrum applications, Appl. Opt., vol. 22(14), p. 2095 (1983).

[24]. T.H. Lee et al., A solid-state image sensor for image recording at 2000 frames per second, IEEE trans. on Elect. Dev., vol. ED-19(9), p. 1469 (1982).

[25]. D. Psaltis and D. Casasent, Spread spectrum time- and space-integrating optical processor, Appl. Opt., vol. 19(9), p. 1546 (1980).

Synthetic Aperture Radar

[26]. E. N. Leith, Quasi-holographic techniques in the microwave region, IEEE Proc. vol. 59(9), p. 1305 (1971).

[27]. L. J. Cutrona, Synthetic Aperture Radar, ch. 23 in *Radar Handbook*, M. I. Skolnik Ed., Mcgraw-Hill, New York (1970).

[28]. E. N. Leith, Synthetic Aperture Radar, in *Optical Data Processing*, D. Casasent Ed., Springer-Verlag, Berlin (1978).

[29]. F. Ulaby, R. Moore, and A. Fung, *Microwave Remote Sensing*, ch. 9, Addison-Wesley, Reading Ma., (1982).

[30]. L. J. Cutrona et al., On the application of coherent optical processing techniques to synthetic-aperture radar, IEEE Proc, vol. 54(8), p. 1026 (1966).

[31]. L. J. Porcello, Optical processing operations in synthetic-aperture radar systems, SPIE vol. 128, p. 108 (1977).

[32]. M. W. Haney, Acousto-optical time-and-space integrating processors for real-time synthetic aperture radar imaging, PhD. Thesis, Caltech (1986).

[33]. D. Psaltis and K. Wagner, Generation of synthetic aperture radar images using acousto-optics, SPIE vol. 271-13, (1981).

[34]. D. Psaltis and K. Wagner, Real-time optical sythetic aperture radar (SAR) processor, Opt. Eng., vol. 21(5), p. 822 (1982).

[35]. D. Psaltis, K. Wagner, and M. Haney, Synthetic aperture radar imaging using acousto-optics and charge-coupled devices, SPIE vol. 352-16, (1982).

[36]. D. Psaltis, M. Haney, and K. Wagner, Real-time synthetic aperture radar processing, Proc. NASA conf. on Optical Information Processing, (1983).

[37]. M. Haney, K. Wagner, and D. Psaltis, Programmable real-time acousto-optic /CCD SAR processor, SPIE vol. 495-25, (1984).

[38]. M. Haney and D. Psaltis, Acousto-optic techniques for real time SAR imaging, SPIE vol. 545-26, (1985).

[39]. M. Haney and D. Psaltis, Real-time acousto-optic spot-light mode SAR processor, OSA Topical Meeting on Optical Computing, Tech. Dig. vol. 11, p. 159 (1987).

[40]. D. A. Ausherman, Digital versus optical techniques in synthetic aperture radar (SAR) data processing, Opt. Eng., vol. 19(2), p. 157 (1980).

[41]. C.D. Daniel, Concepts and techniques for real-time optical synthetic aperture radar data processing, IEE Proc., vol. 133(1), p. 7 (1986).

CHAPTER 4

DEVICE PERFORMANCE FOR TSI SIGNAL PROCESSING

The performance characteristics of the active devices utilized in a Time and Space Integrating (TSI) signal processing system determine the accuracy and practicality with which a desired computational task can be performed. The principal active devices which affect the system operation are the light source, the optical modulators, and the photodetectors, although the quality of the passive devices, such as lenses and beamsplitters, may also affect system performance. The most compact and efficient coherent light source available is the semiconductor laser diode (LD), and this was selected to illuminate the signal processing systems studied in this thesis because of the ability to directly modulate and pulse the optical source. Two dimensional optical modulators exist, but are not yet as fully developed as the 1-D travelling wave acousto-optic devices (AOD) which were used to enter data into the TSI systems discussed here. Solid state silicon charge coupled device (CCD) detector arrays have emerged at the leading edge of VLSI technology as the largest integrated circuits in production, and they represent an excellent technology for performing 2-D time integrating detection in optical signal processing systems.

With ideal devices, the signal processing operation can easily be accomplished with almost arbitrary precision and fidelity, limited only by shot noise, and the achievable resolution and computational speed are phenomenal. However, the fundamental device limitations must be considered by the system architect while designing an optical signal processing system, so that the performance capabilities of each component are maximally utilized, and unreasonable performance expectations are not required. The engineering performance limits of commercially available de-

vices will provide another limit on the capabilities of any practical optical signal processing system that must be considered in the system design phase. The possibility of building custom devices for the particular system should only be considered if standard devices are inappropriate. The first step should be to examine the system architecture from the point of view of obtaining the full performance capabilities of existing devices, and selecting the simplest architecture that makes the best use of available hardware.

Commercially available devices have reached a high level of sophistication, and the appropriate utilization of these state of the art electro-optic components can lead to optical signal processing with significant performance advantageous over other analog processing techniques, or the digital signal processing counterparts. These advantages can include size, weight, power, computational speed and cost, although there will also be disadvantages associated with the analog nature of the information representation, and the sensitivity to alignment errors associated with optical systems. Highly coherent laser diodes with modulation bandwidths up to 10GHz and single mode optical power outputs of 10-100mW are available, and much higher power multimode devices exist as well[6,7]. Acousto-optic deflectors are available with over 2000 resolvable spots, dynamic range as large as 80dB, and bandwidths on the order of 50MHz[32], although bandwidths of several GHz are also commercially available, although with smaller time bandwidth products and smaller dynamic range[33]. Two dimensional scientific CCD detectors with 2048×2048 pixels are available as wafer scale integrated chips with dynamic ranges on the order of 10^4:1, and readout rates in the 10s of megapixels per second, though not simultaneously with the high SNR operation[45]. Each device is made with its own tradeoffs and optimized regime of operation, and numerous manufactures produce a plethora of designs, from inexpensive consumer products not intended for scientific applications to extremely expensive state of the art devices. The optical signal pro-

cessing systems designer must select from the available and affordable devices the appropriate selection of components to build the desired system, with the required performance characteristics. The interrelationships between the device characteristics will affect the system performance, and must be considered from an interacting system viewpoint, rather than looking at each device in its optimum region of operation. Similarly the requirements of the architecture and of the optical design will affect the choice of devices and the operating characteristics of the devices.

As an example of these tradeoffs consider the most common commercially available laser diodes, which are made from GaAs/GaAlAs, and consequently have an operating wavelength in the range of 780-900nm. However in this spectral region the resolution of silicon CCD imagers is significantly decreased due to the migration of deep carriers generated by the deeply penetrating near IR light. This can be overcome by using a visible gas laser, or by using a thinned backside illuminated CCD detector array. The gas laser is bulky, inefficient and not as easily modulated as the laser diode, and thinned CCDs are fragile and expensive. Another consideration is the diffraction efficiency of commercially available acousto-optic deflectors at this near IR wavelength. Because of the dispersion of the optical rotary power, as well as the wavelength dependence of the photoelastic coefficients, commercially available TeO_2 deflectors operate much less efficiently at the laser diode wavelengths than at the visible wavelengths for which they are designed. Interacting considerations of this nature indicate the importance of system and device tradeoffs.

In this chapter, I will review the operation and performance of laser diodes, acousto-optic devices and charge coupled device detector arrays. The performance characteristics that can have significant effects on TSI signal processing systems are reviewed from the device perspective in order to assist in the interpretation of the experimental results presented in Chapter 5. Some speculations as to custom device design intended to optimize TSI signal processing system performance are included to illustrate the relationship between systems characteristics and device tradeoffs.

4.1 Laser Diode Optical Sources

The laser is the "power supply" of an optical signal processing system, as well as the carrier of the information within the system. When the laser power output is directly modulated it also represents an additional signal port through which data and control signals can be entered into the system. Laser diodes represent the most efficient coherent optical source currently available, boasting a differential quantum efficiency of up to 70%, and an overall electrical to optical conversion efficiency of over 30%. The small size and efficient operation of laser diodes makes them extremely attractive for on board systems applications, where size and power considerations are important. Even more significant is the ability to directly modulate the light power output of a laser diode with an applied voltage waveform at bandwidths of up to 1 GHz or more. This allows the simple implementation of modulated source time integrating correlator architectures and triple product processors. Laser diodes can also be pulsed with narrow pulses in order to freeze the acoustic motion in an AOD and implement space integrating holographic correlators and time integrating matrix multipliers. This pulsed mode of operation was the primary mode utilized in this thesis, and this section will briefly review the operating performance of a pulsed laser diode and its effect on a multidimensional TSI signal processing system.

A laser diode is a forward biased p-n junction in a direct gap semiconductor inside an optical resonator[1-3]. When a high injection condition is established at the junction, a population inversion is set up at the junction boundary between the electrons in the conduction band and holes in the valence band. In the active region where the population inversion is maintained, spontaneous emission of photons will take place due to the recombination of electrons and holes. These spontaneously emitted photons, with energy approximately equal to the semiconductor bandgap, will be isotropically radiated, while some fraction will be confined by the optical

resonator. The resonator can be formed with external mirrors, but is usually formed by cleaving the semiconductor diode along parallel crystal planes, and relying on the index mismatch between air and the active layer to form a low finesse optical cavity. Those photons which are within the angular aperture confined by the mirror facets will have a probability of reflecting back and forth within the cavity, where they can be absorbed by inducing band to band transitions in the semiconductor. Since the direct semiconductor is in an inverted condition the photons in a cavity mode will interact with the inverted carrier population and produce coherent photons via stimulated emissions, thereby causing electron hole pairs to recombine. When the rate of stimulated emission is greater than the absorption rate then the inverted semiconductor will have gain for a photon propagating within the cavity. A lasing threshold condition is obtained when the cavity gain coefficient is equal to the loss coefficient plus the mirror losses. Below this threshold the emitted light is dominated by the incoherent spontaneous generation term and the emission spectra has a width given by the gain profile of the semiconductor. When the p-n diode is pumped at an injection current density larger than the lasing threshold current density, then the emission spectra becomes dominated by the cavity mode structure, and the emitted radiation becomes coherent. At high enough injection levels, small cavity devices can operate in a single mode, giving extremely long coherence length, which is ideal for coherent optical processing, and it is this mode of operation which is most important for TSI interferometers.

The diode laser used in the experiments reported in Chapter 5 was a GaAlAs double hetero structure single mode device manufacture by Hitachi (HLP 1600)[3], that had a rated CW output power of 15mW, but could be pulsed to much higher levels. This diode had about .2mW of spontaneous emission at its 49mA threshold, and an electrical to optical conversion efficiency above threshold of .32mW/mA. The double hetero structure, which is illustrated in Figure 4.1.1a, is used in these

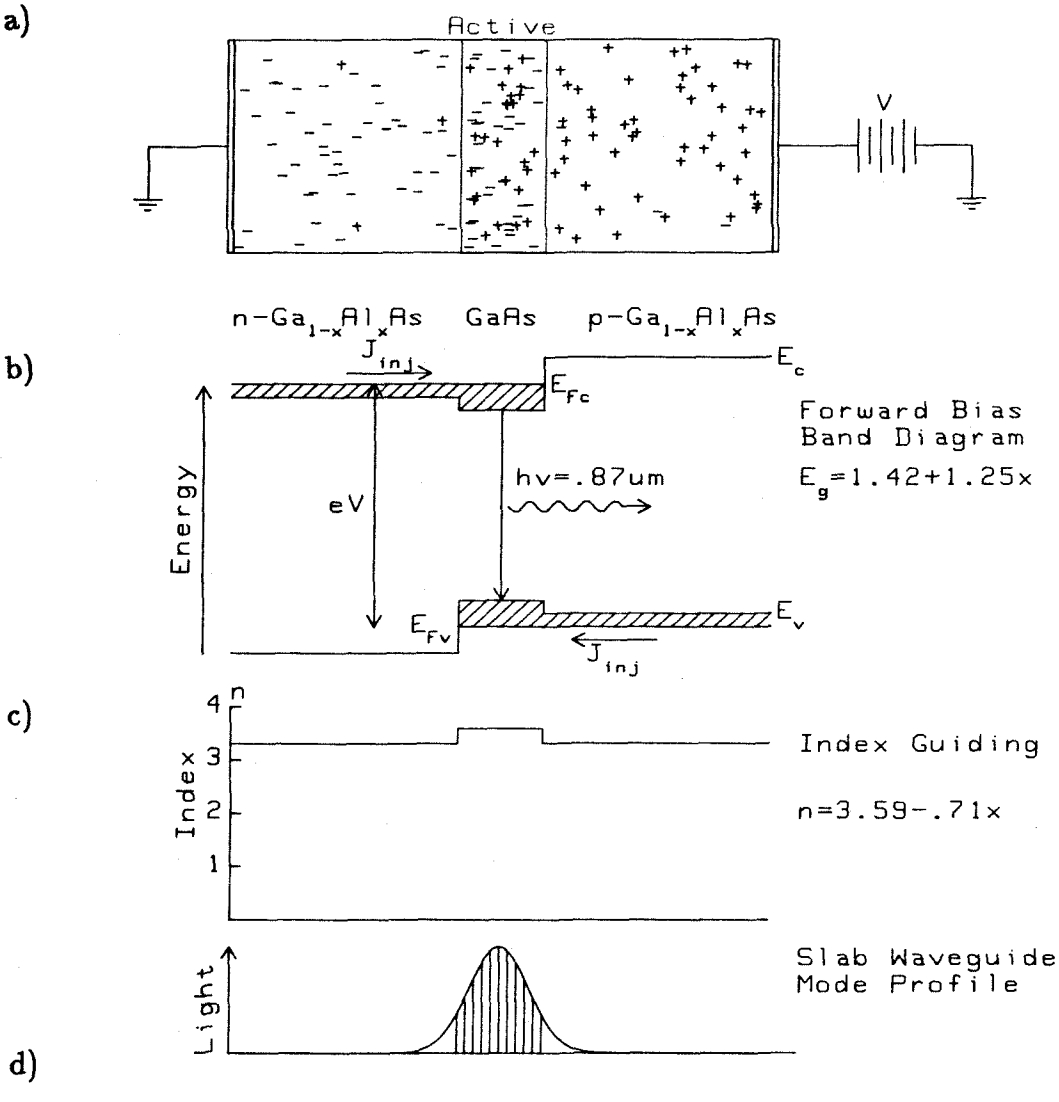


Figure 4.1.1. a) Double hetero structure GaAlAs laser diode, b) Band gap of the diode at a lasing forward bias, c) Index profile across the diode, d) Resulting optical mode profile. (after [1])

devices because it results in a small threshold current density by confining the inverted carriers in the same region as the optical waveguide. Carrier confinement is accomplished by sandwiching a thin ($.1\mu m$-$.5\mu m$) active layer of GaAs between a p-type layer of GaAlAs and an n-type layer of GaAlAs. The bandgap of $Ga_{1-x}Al_xAs$ is a linear function of the composition given by $E_g(eV) \approx 1.424+1.247x$, for $x < .37$, so the emission wavelength in microns will be approximately $\lambda \approx 1.24\mu m/E_g(eV) = 1.24/(1.42 + 1.25x)$. When the narrow band gap GaAs is sandwiched between the wide gap ternary GaAlAs p and n layers, intervening conduction band and valence band discontinuities will result, with most of the discontinuity in the conduction band. When this diode is forward biased under high injection conditions the heterojunction band discontinuities effectively trap the electrons and holes in an overlapping spatial volume where a highly inverted condition is maintained, as shown in Figure 4.1.1b. This active layer of GaAs will have a high gain coefficient which is proportional to the inversion density. The emitted optical radiation will be confined within the active region because the refractive index is also a function of the composition, $n = 3.59-.71x$, thereby producing a dielectric waveguiding step index profile at the heterojunction boundaries as shown in Figure 4.1.1c. This index guiding can also be accompanied by gain guiding or thermal guiding, where the active layer index dependence on the injected carrier density or on the junction temperature can be used to confine the optical mode. The mode is effectively confined to the region of gain as shown in Figure 4.1.1d, so that only small evanescent tails extend into the GaAlAs region, which is essentially transparent at the lasing wavelength because of the wider bandgap. This simultaneous waveguiding and inverted carrier population confinement of the narrow GaAs layer results in a highly efficient device. In the lateral dimension the light can be confined by burying the heterostructure within a lateral region of GaAlAs producing a built in lateral waveguide. Alternatively, only a thin strip is injected by a metal contact along the length of the laser, producing

a gain guided strip within the thin GaAs layer. Typically light is emitted equally from both end facets, but only one direction is utilized by the optical system, and the other facets emission represents wasted optical power.

The transverse modal properties of the laser are determined by the size and shape of the active and confinement regions. When a very thin layer of GaAs is used then the electric field of the emitted laser light is linearly polarized in the plane of the junction. Moreover, a small cavity can be made to oscillate in only the fundamental TEM_{00} mode, and this was the case for the Hitachi laser, which had an effective near field laser profile of $1.3\mu m$ by $4.\mu m$, measure at the 1/e point of the emitted Gaussian beams. This very small optical spot size radiates as an anamorphically expanding Gaussian beam with half power full width divergence angles of 10° in the plane of the active layer, and 30° in the transverse plane of the heterojunction, as illustrated in Figure 4.1.2. This beam can be well collimated since it emerges from such a well defined point source, unless the horizontal and vertical beam waists are at different positions within the cavity. However, it will produce an elliptical beam profile unless special collimating optics are employed.

The longitudinal modal properties of the laser determine the coherence and spectral purity of the emitted light. In an optical interferometer the coherence determines the visibility of the fringes detected at the output which determines the modulation depth of the desired information with respect to the incoherent bias. The Fabry-Perot cavity resonance condition determines the allowed frequencies of oscillation of the laser. The allowed oscillation wavelengths for a cavity of length L and index n are determined by the requirement that an integral number N of half cycles fit within the resonator, $nL = \lambda N/2$. At sub threshold drive currents the laser is almost totally incoherent, with a spectral width as large as the GaAs gain profile, and may not exhibit a modal structure. As the drive current is raised above threshold a number of modes within the gain profile begin to lase, and they com-

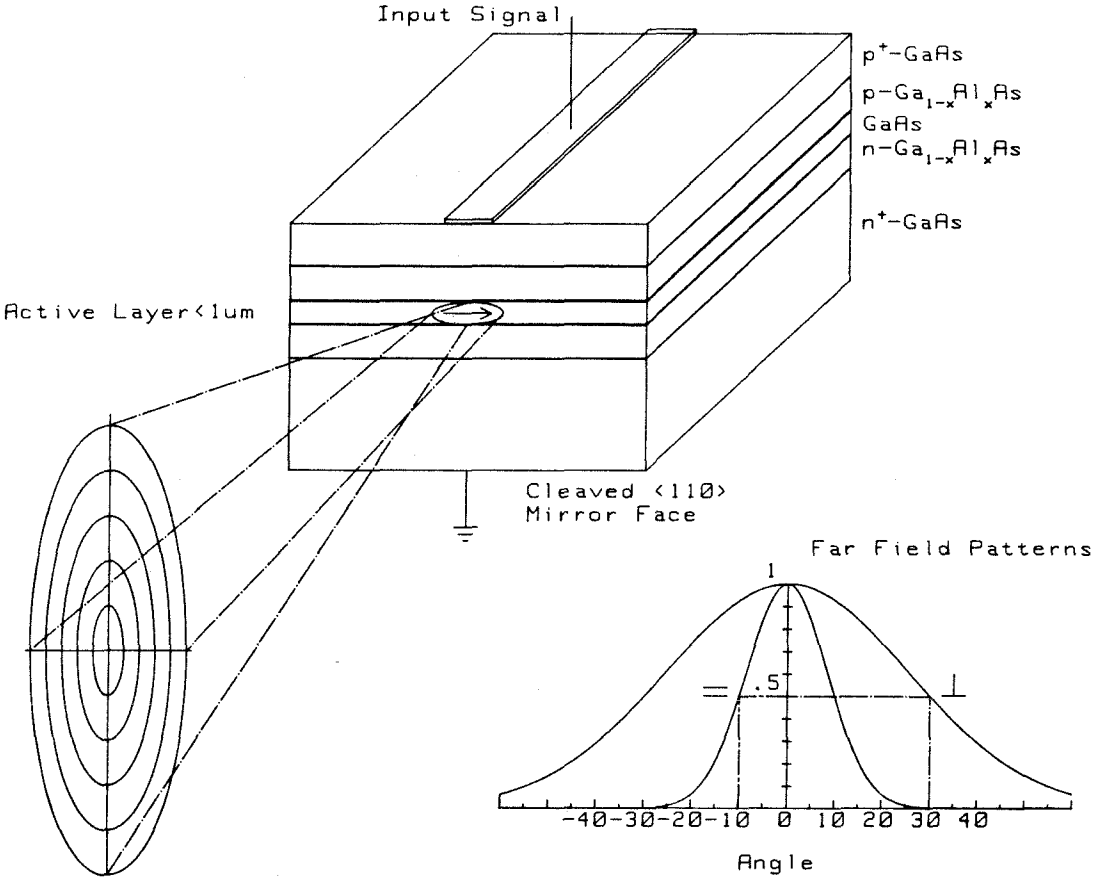


Figure 4.1.2. Diode structure with resulting anamorphically diverging beam profile, and far field cross sections parallel and perpendicular to the junction plane.

pete for the available population of inverted carriers. The Hitachi laser diodes are designed to operate as single mode devices at higher drive currents, and for output powers over a few mW one of the longitudinal modes wins the modal competition. The winning mode steals the gain from all the other competing modes and contains more than a thousand times the power of any other mode. The spectral width of a single oscillating mode can be as low as $\Delta\nu$=50MHz, out of a center frequency of $\nu = c/\lambda = 3.6 \times 10^{14}$, giving a spectral purity of almost one part in 10^7. The corresponding coherence length $l_c = c\Delta\nu$ can be as long as 10m, making the laser diode one of the most coherent lasers available. However when these lasers are operated in a pulsed mode rather than with DC drive current then the laser coherence suffers drastically. For the first few nanoseconds of a pulse several modes will compete until one wins and steals the gain from the others, which results in an essentially incoherent output until single mode operation is achieved[4]. The winning mode is usually the most favored mode that is near the peak of the emission spectrum, but several modes can show a non negligible probability of becoming the dominant single mode, so that on successive pulses different modes may dominate. The winning mode will experience a small frequency drift during a long pulse due to junction heating effects, until an equilibrium condition is reached. If the frequency drift is large enough the adjacent mode will become favored, and the lasing mode will hop to the adjacent oscillation frequency[3]. These frequency drifting and hopping effects will introduce fringe shifts in the interferometrically detected output from an interferometer, which become more severe for longer path length differences. These effects can seriously affect a TSI interferometer where thousands of laser pulses need to be coherently time integrated on a detector array as spatial fringe patterns[5]. The probability of interpulse modal hopping can be minimized by biasing the laser just below threshold, and adjusting the rise time and pulse amplitude. The incoherent emission during the first few nanoseconds of the pulse should be made

a small fraction of the pulse duration, so the bias introduced during this time is only a small percentage of the coherently detected fringes during the remainder of the pulse. Conversely the pulse length should be short enough to avoid effects of junction heating during the pulse, and to freeze the acoustic motion and travelling fringes caused by the doppler shifted AO diffraction. These considerations lead to a pulse with a 50-100 nsec width, and a 30-50 nsec rise time on a bias just below the laser threshold, in order to maximize the laser coherence.

The power obtainable with a single mode diode laser is typically from 10-30mW, but in pulse operation peaked powers on the order of 100mW may be obtainable. However, the coherence may suffer when the laser is pulsed at higher output intensities than its rated CW output power. The typical current input versus optical power output relationship of a laser diode is plotted in Figure 4.1.3, which shows a strong thresholding behavior. Below the lasing threshold only incoherent spontaneous emission is generated and the spectral output is broad, like a light emitting diode (LED). Above the lasing threshold the coherent laser output dominates and the spectral output becomes dominated by a set of Fabry-Perot resonance peaks. Well above threshold a single mode can win the modal competition for the available gain and the laser output becomes highly coherent with a very narrow spectral peak. The linearity of the current power relationship is very good over two orders of magnitude with a second harmonic distortion as low as 45dB below the fundamental[3].

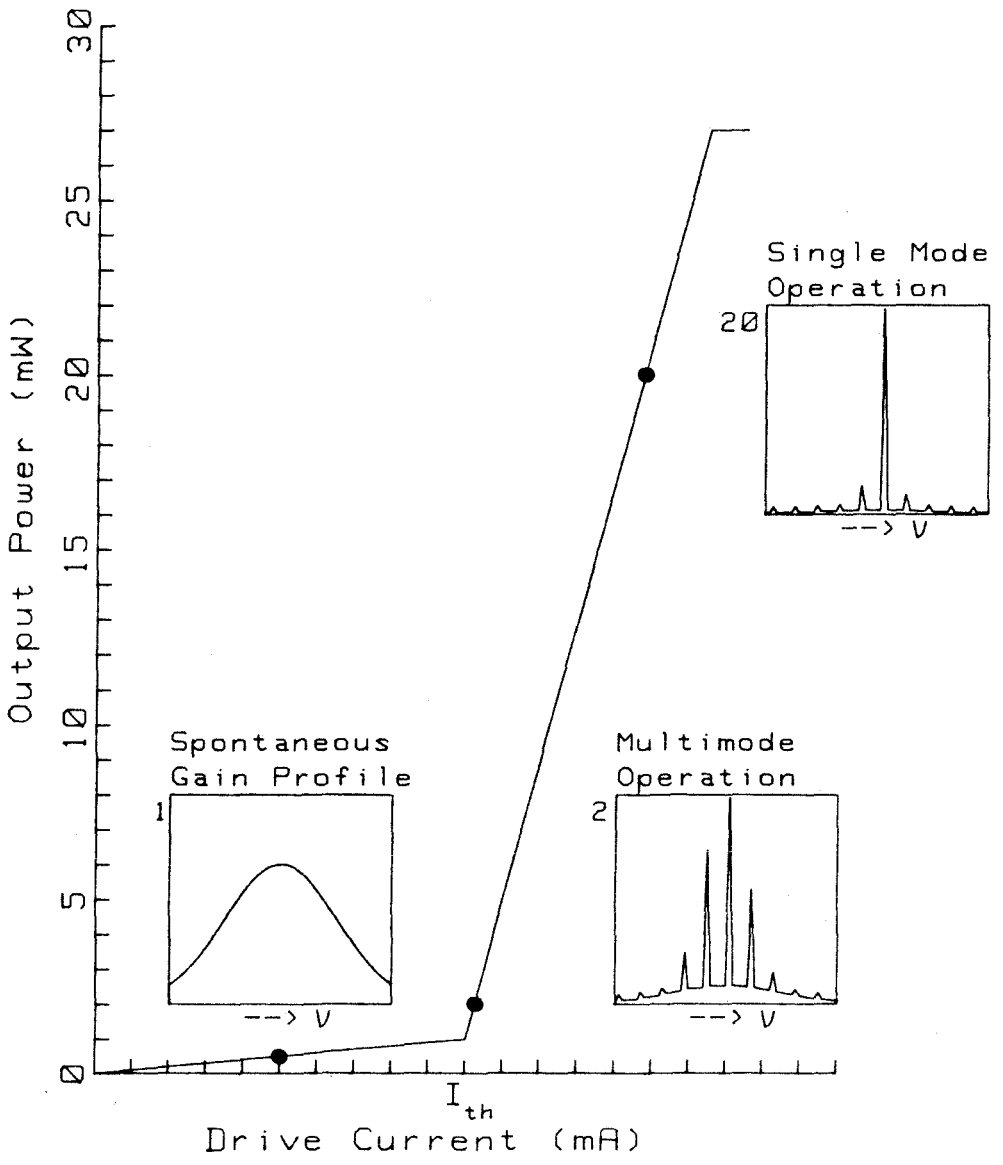


Figure 4.1.3. Laser diode optical power output versus current input, and the expected spectra at various different operating points.

4.2 Acousto-Optic Device Model

In this section I will review some of the essential operating characteristics of acousto-optic devices[8] (AODs) that are used in signal processing systems. These devices are based on the periodic modulation of the optical index of refraction caused by an acoustic wave propagating in a transparent material. An optical wave passing through the region containing the acoustic wave will experience a periodic phase modulation that can produce a corrugation of the optical wavefront which will result in a scattering of the optical wave into diffracted orders. The acousto-optic interaction is somewhat different from other types of optical volume diffraction effects because the perturbation is required to be a superposition of propagating eigenmodes of the acoustic wave equation. This is especially relevant in the case of highly anisotropic acoustics, such as in the slow shear mode of paratellurite (TeO_2), which was used as a birefringent AO deflector in the systems reported in Chapter 5. An understanding of the nonideal behavior of this type of device is important at the stage of system design, in order to minimize unwanted effects, and in order to understand apparently anomalous behavior of the experimental arrangement. However, from the systems analysis point of view the issues discussed in this section are far too involved to be treated analytically at the level of the full signal processing system operation.

The acousto-optic interaction results in the phase modulation of an optical beam by an acoustic wave in a photoelastic medium. In order to properly describe this interaction I will represent the acoustic and optical waves separately, and then introduce the coupling between them. This interaction is mediated by the 4th rank photoelastic tensor which describes the dielectric impermeability tensor perturbation caused by the propagating strain wave. In practice, devices are designed to make use of only a single element of the coupling tensor, and the problem is greatly simplified.

4.2.1 Acoustic eigenmodes

The acoustic wave is modelled as a time and space varying particle displacement vector field $\vec{u}(\vec{x},t)$, which at each crystal lattice site describes the particle displacement from its equilibrium position[9,10]. A sinusoidal displacement wave of amplitude W, radian frequency Ω, wave vector $|\vec{K}| = 2\pi/\Lambda$, acoustic wavelength Λ, propagating with a phase velocity $v_a(\vec{s}) = \Omega/|K|$ in the direction defined by the unit vector $\vec{s} = \vec{K}/|K|$, and with unit polarization $\hat{U}$ is described by the particle displacement field

$$\vec{u}(\vec{x},t) = W\hat{U}\cos(\Omega t - \vec{K}\cdot\vec{x}) = W\hat{U}\cos\left[\Omega(t - \frac{\vec{s}\cdot\vec{x}}{v_a(\vec{s})})\right] \tag{4.2.1}$$

The displacement gradient matrix is the spatial derivative of the displacement field, and its components are given by

$$Q_{ij}(\vec{x},t) = \left[\frac{\partial u_i(\vec{x},t)}{\partial x_j}\right] \tag{4.2.2}$$

The symmetric part of the displacement gradient matrix is known as the linearized Strain tensor and its components are given by

$$\mathbf{S}_{ij} = \frac{1}{2}\left[\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right] \tag{4.2.3}$$

The Stress tensor $\mathbf{T}_{ij}$, which is symmetric in non feroic materials, is related to the Strain tensor through the 4th rank elastic stiffness tensor in a generalization of Hooke's law.

$$\mathbf{T}_{ij} = \mathbf{c}_{ijkl}\mathbf{S}_{kl} \tag{4.2.4}$$

Where the Einstein summation convention over repeated indices is implied, and i,j,k,l may take on any of the three spatial directions x_1,x_2,x_3 or equivalently x,y,z. The elastic coefficients possess certain symmetries because of the symmetry of $\mathbf{S}$ and $\mathbf{T}$, so that $\mathbf{c}_{ijkl} = \mathbf{c}_{jikl} = \mathbf{c}_{ijlk} = \mathbf{c}_{jilk}$, and energy arguments show that $\mathbf{c}_{ijkl} = \mathbf{c}_{klij}$. The acoustic field equations show that energy oscillates between the

stress energy and the strain energy in a fashion that is analogous to the electromagnetic oscillation between electric and magnetic energy. The dynamical equation of motion for a vibrating medium relates the restoring force as given by the divergence of the Stress tensor with the mass times acceleration of the displacement field.

$$\vec{F} = \nabla \cdot \mathbf{T} = \rho_m \frac{\partial^2 \vec{u}}{\partial t^2} \tag{4.2.5a}$$

$$F_i = \frac{\partial}{\partial x_j} \mathbf{T}_{ij} = \rho_m \frac{\partial^2 u_i}{\partial t^2} \tag{4.2.5b}$$

In this equation ρ_m is the scalar equilibrium mass density of the medium. Substitution of the Stress-Strain relationship, and the definition of Stress in terms of particle displacements into the dynamical equation of motion results in the differential equation governing the propagation of particle displacement fields.

$$F_i = \mathbf{c}_{ijkl} \frac{\partial^2 u_k}{\partial x_l \partial x_j} = \rho_m \frac{\partial^2 u_i}{\partial t^2} \tag{4.2.6}$$

Substitution of the assumed plane wave of Equation 4.2.1 into this equation results in the equation for the allowed modes of propagation.

$$\mathbf{c}_{ijkl} s_j s_l k^2 U_k = \Gamma_{ik}(\vec{s}) k^2 U_k = \rho_m \Omega^2 U_i \tag{4.2.7a}$$

$$\left[\mathbf{c}_{ijkl} s_j s_l / v_a^2 - \rho_m \delta_{ik} \right] U_k = 0 \tag{4.2.7b}$$

This system is only compatible for all waves if the determinant of the system is zero, which results in the dispersion relationship in terms of the Christoffel matrix $\Gamma_{ik}(\vec{s}) = \mathbf{c}_{ijkl} s_j s_l$ as a function of the propagation direction $\vec{s}$, where s_x, s_y, s_z are the appropriate direction cosines.

$$\det \left| \frac{\Gamma_{ik}(\vec{s})}{v_a^2} - \rho_m \delta_{ik} \right| = 0 \tag{4.2.8}$$

This equation has three solutions for the acoustic slowness, or inverse velocity $1/v_a(\vec{s}) = \vec{K}/\Omega$ for each direction $\vec{s}$, forming three equivalent frequency scaled surfaces in $\vec{K}$ space, known as acoustic momentum space. The corresponding eigenpolarizations $\hat{U}(\vec{s})$ for each direction of propagation correspond to the longitudinal (or quasilongitudinal) and two shear (or quasishear) solutions.

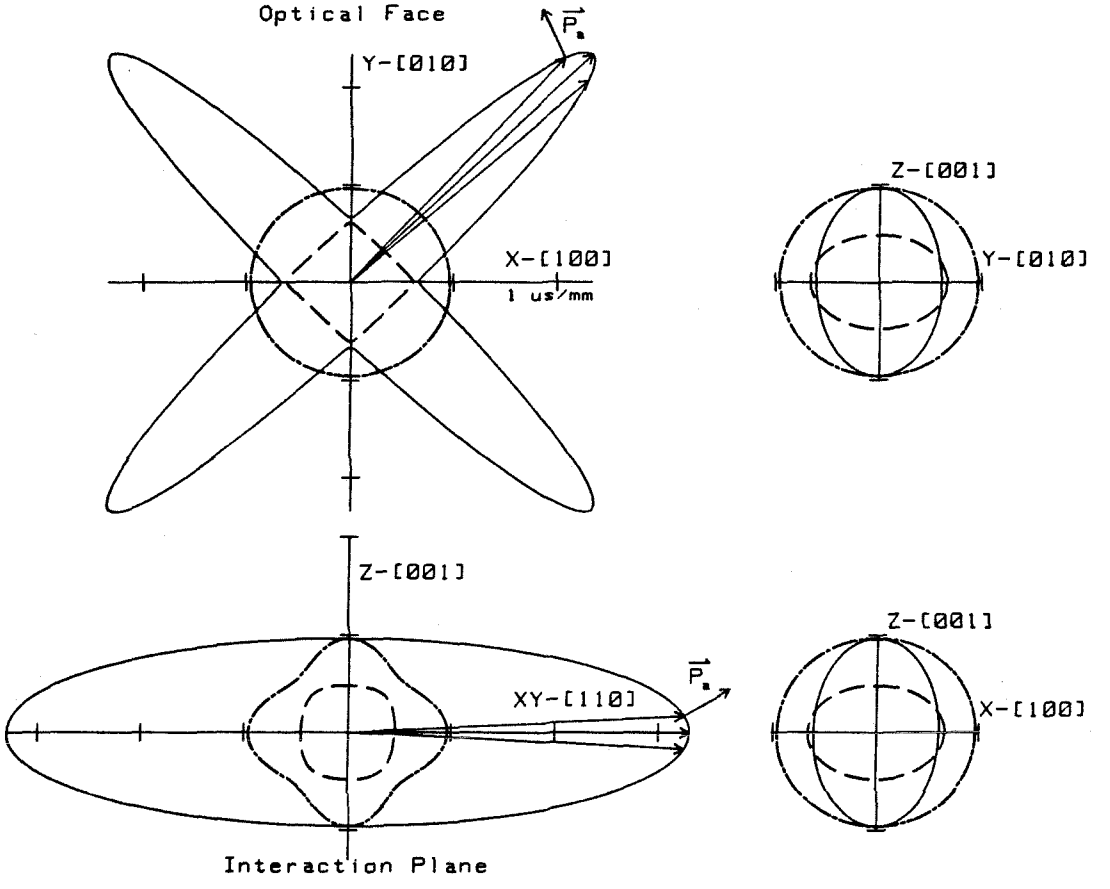


Figure 4.2.1. Cross sections of the acoustic slowness surfaces for the case of TeO_2.

An example of cross sections through this surface, calculated via an explicit solution of the Christoffel equation are shown in Figure 4.2.1 for the important case of paratellurite TeO_2[11]. The anomalously slow shear mode is recognized by the large lobes extending in the [110] direction, and this is the acoustic mode utilized in the Bragg cells used in this work. The acoustic velocity along this axis is .62mm/μsec, and the radius of curvature in the x-y plane is about 1/44, and in the xy-z plane that orthogonally cuts through the slow shear lobe the radius of curvature is about 1/12. The acoustic phase velocity surface is found by inverting the slowness surface radially about the unit sphere. It is important to realize that the acoustic Poynting's vector, which describes the direction of energy flow, is orthogonal to the slowness surface, thus in the regions of high curvature around the [110] direction slightly off axis acoustic components will not only propagate faster, they also will rapidly walk off from the main beam.

4.2.2 Optical eigenmodes

Optical propagation through a homogenous, lossless anisotropic medium can be described in terms of Maxwell's equations[12]. Faraday's law gives the relation between an induced electric field and a time varying magnetic field. Ampère's law describes the creation of a magnetic field due to a dielectric flux, a conductivity current, and a source current, however for the case of interest to acousto-optics, no currents are present. Similarly, we will assume no free electric charges, and of course no free magnetic monopoles.

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{4.2.9}$$

$$\nabla \times \vec{H} = \frac{\partial \vec{D}}{\partial t} + \vec{J}_c + \vec{J}_s = \frac{\partial \vec{D}}{\partial t} \tag{4.2.10}$$

$$\nabla \cdot \vec{D} = \rho_e = 0 \tag{4.2.11}$$

$$\nabla \cdot \vec{B} = 0 \tag{4.2.12}$$

In an optically anisotropic medium the displacement vector $\vec{D}$ and electric field $\vec{E}$ are not necessarily parallel, and are related by the Hermitian second rank permittivity tensor ϵ. In magnetically isotropic material the magnetic vector $\vec{H}$ is related to the magnetic induction $\vec{B}$ by the scalar permeability μ. The resulting constitutive relationships describe the effect of material media on the propagation of electromagnetic waves, and allows the unique solution to Maxwell's equations with a given set of boundary conditions.

$$\vec{D} = \epsilon\vec{E} = \epsilon_0\vec{E} + \vec{P} \tag{4.2.13a}$$

$$D_i = \epsilon_{ij}E_j = \epsilon_0(\delta_{ij} + \chi_{ij})E_j = \epsilon_0 E_i + \chi_{ij}E_j = \epsilon_0 E_i + P_i \tag{4.2.13b}$$

$$\vec{B} = \mu\vec{H} = \mu_0\vec{H} + \vec{M} \tag{4.2.14a}$$

$$B_i = \mu H_i \tag{4.2.14b}$$

The permittivity tensor is expressed as the free space permittivity ϵ_0 plus a material dependent susceptibility tensor, $\epsilon_{ij} = \epsilon_0(1 + \chi_{ij})$, and the presence of the matter is seen to induce a polarization vector $\vec{P}$ which is related to $\vec{E}$ through the linear susceptibility tensor χ_{ij}.

In order to derive the optical eigenmodes we will assume an electromagnetic plane wave with angular frequency ω, and wave vector $|\vec{k}| = 2\pi/\lambda$, propagating in the direction of the unit vector $\vec{s} = \vec{k}/|\vec{k}|$, with a phase velocity $v_p = c/n = 1/\sqrt{\mu\epsilon}$. The refractive index $n = \sqrt{\epsilon/\epsilon_0}$ is a function of the direction of propagation, the polarization of the wave, and the frequency if the material is dispersive, and it is the allowed eigen-velocities v_p and eigen-polarizations $\vec{E}_0$ which are to be determined.

$$\vec{E}(\vec{x},t) = \vec{E}_0 e^{-i(\omega t - \vec{k}\cdot\vec{x})} \tag{4.2.15a}$$

$$\vec{H}(\vec{x},t) = \vec{H}_0 e^{-i(\omega t - \vec{k}\cdot\vec{x})} \tag{4.2.15b}$$

We can now substitute these assumed solutions into Maxwell's equations to obtain

$$\vec{k}\times\vec{E} = \omega\mu\vec{H} \tag{4.2.16a}$$

$$\vec{k} \times \vec{H} = -\omega\epsilon\vec{E} \tag{4.2.16b}$$

Substituting the first equation for $\vec{H}$ into the second equation yields an equation for $\vec{E}$.

$$\vec{k} \times (\vec{k} \times \vec{E}) + \omega^2\mu\epsilon\vec{E} = 0 \tag{4.2.17a}$$

$$\left[k_i k_j - \delta_{ij}k^2 + \omega\mu\epsilon_{ij}\right] E_j = 0 \tag{4.2.17b}$$

In the absence of optical activity, the symmetry of the permittivity tensor ϵ allows us to rotate to a principal dielectric coordinate system where ϵ_{ij} is purely diagonal. In order for a nontrivial solution to exist the determinant of the matrix in brackets in equation (4.2.17b) must be zero.

$$\det\left[k_i k_j - \delta_{ij}k^2 + \omega\mu\epsilon_{ij}\right] = 0 \tag{4.2.18}$$

This is the equation for the optical normal surface in $\vec{k}$ space, referred to as optical momentum space. It is analogous to the acoustic momentum surface, and the inverse optical phase velocity $v_p^{-1} = n/c = k/\omega$ in a particular direction, is proportional to the radius of the momentum surface in that direction divided by the optical angular frequency. For each direction of propagation there are two possible eigen-phase-velocities, with corresponding orthogonal, eigen-polarizations as solutions to equation (4.2.17). The two surfaces intersect in degenerate directions known as the optical axes, and there may be up to four such intersections in biaxial media. For uniaxial materials there are only two such intersections along a single line, giving a single optical axis. For the case of TeO_2 optical activity breaks the degeneracy along the optical axis, and results in a slight splitting of the two eigen-velocities along the optical axis, as shown in Figure 4.2.2. The eigen-polarizations are circular along the optical axis, and become elliptical as the direction of propagation moves away from the optical axis.

The impermeability tensor $\eta_{ij} = \epsilon_0(\epsilon^{-1})_{ij}$ is the inverse of the dielectric tensor ϵ and is given by the relation $(\epsilon^{-1})_{kj}\epsilon_{ik} = \delta_{ij}$. As a second rank symmetric tensor

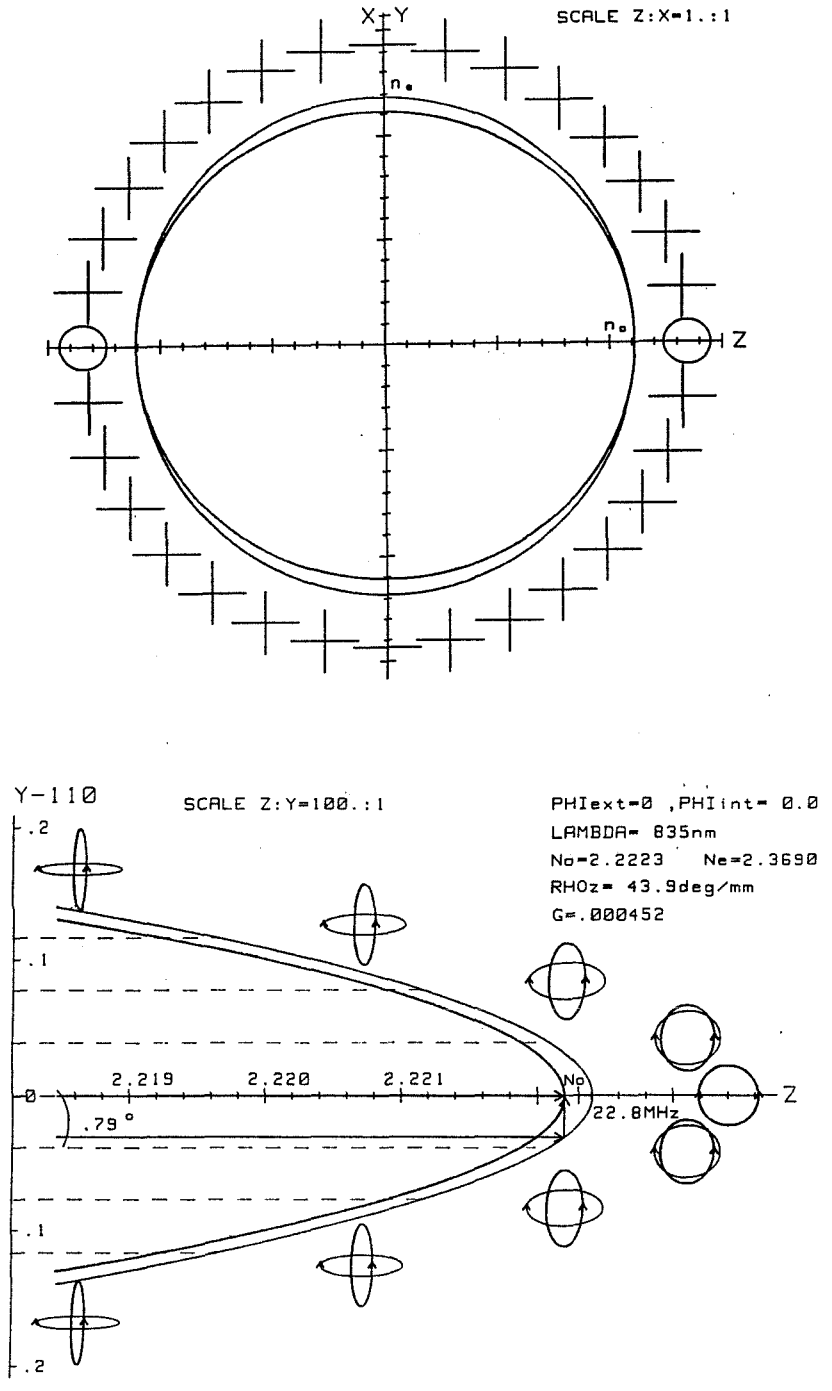


Figure 4.2.2. The optical normal surface for TeO_2 at a wavelength of λ =835 nm, and a magnified version of the normal surface near the optical axis in the presence of optical activity showing the resulting elliptical eigen-polarizations.

it describes a quadratic surface known as the index ellipsoid

$$(\epsilon^{-1})_{ij}x_i x_j = \left(\frac{1}{n^2}\right)_{ij} x_i x_j = 1. \tag{4.2.19}$$

In the principal coordinate system this equation reduces to the familiar representation of a general ellipsoid.

$$\frac{x_1^2}{n_1^2} + \frac{x_2^2}{n_2^2} + \frac{x_3^2}{n_3^2} = 1 \tag{4.2.20}$$

This surface is a convenient geometric representation for finding the optical eigenmodes of the displacement vector $\vec{D}$ for a given direction of propagation. These eigenmodes are found by finding the principal axes of the ellipse normal to the propagation direction. Associated with each eigenmode is a corresponding index of refraction, equal to the ellipse radius along each principal axis.

4.2.3 Photoelastic coupling

The photoelastic effect is usually described in terms of an elastically induced perturbation of the impermeability tensor mediated by a fourth rank elastooptic tensor, which has nonzero components in all materials[10,13].

$$\Delta\eta_{ij} = \Delta\left(\frac{1}{n^2}\right)_{ij} = \mathrm{p}_{ijkl}\mathrm{S}_{kl} \tag{4.2.21}$$

The symmetry of the index ellipsoid in i and j, and the strain tensor in k and l result in the symmetry relations for the strain-optic tensor $\mathrm{p}_{ijkl} = \mathrm{p}_{jikl} = \mathrm{p}_{ijlk} = \mathrm{p}_{jilk}$. Because the index perturbations due to the photoelastic effect are small we can use the relationship $dn = -\frac{1}{2}n^3 d\left(\frac{1}{n^2}\right)$ to write

$$(\Delta n)_{ij} = -\frac{n^3}{2}\mathrm{p}_{ijkl}\mathrm{S}_{kl} \tag{4.2.22}$$

Thus the perturbation of the optical index is proportional to the magnitude of the applied strain, and as long as the appropriate photoelastic tensor coefficient is nonzero there will be a resulting phase modulation of a properly polarized optical wave passing through the medium.

An acoustic eigenmode plane wave as in Eq. (4.2.1), will induce a strain wave of the form

$$\mathbf{S}(\vec{x},t) = \frac{1}{2}W\hat{S}_{mn}\cos(\Omega t - \vec{K}\cdot\vec{x}) \tag{4.2.23}$$

Where $\hat{S}_{mn}$ is a unit strain tensor for the given mode. This will induce a periodic travelling wave volume dielectric tensor perturbation that will couple the ith polarization component of the input mode with the jth polarization component of the output mode

$$\Delta\epsilon_{ij}(\vec{x},t) = \frac{W}{2}\epsilon_{ik}(\mathrm{p}_{klmn}\hat{S}_{mn})\epsilon_{lj}\cos(\Omega t - \vec{K}\cdot\vec{x}) \tag{4.2.24}$$

For an isotropic medium this corresponds to a perturbation of the index of refraction given by

$$\Delta n_{ij}(\vec{x},t) = -\frac{n^3}{2}W\mathrm{p}_{ijmn}\hat{S}_{mn}\cos(\Omega t - \vec{K}\cdot\vec{x}) = \delta n_0\cos(\Omega t - \vec{K}\cdot\vec{x}). \tag{4.2.25}$$

This index grating of amplitude $\delta n_0 = -\frac{n^3}{2}W\mathrm{p}_{ijmn}\hat{S}_{mn}$ can diffract an incident optical plane wave, of polarization i into a diffracted beam with polarization j, if the appropriate photoelastic tensor element is nonzero, and as long as both energy and momentum are conserved. This results in two conservation equations for the incident and diffracted optical waves.

$$\vec{k}_d = \vec{k}_i \pm m\vec{K}_a \qquad m = 0,1,2,...N \qquad \text{(momentum conservation)} \tag{4.2.26}$$

$$\omega_d = \omega_i \pm m\Omega_a \qquad m = 0,1,2,...N \qquad \text{(energy conservation)} \tag{4.2.27}$$

The plus and minus signs in these equations correspond to the annihilation or creation of m phonons, respectively.

With near normal incidence, many orders are simultaneously created and we are said to be in the Raman-Nath regime, which is illustrated in Figure 4.2.3a. When the width of the acoustic wave in the direction of light propagation greatly exceeds the characteristic length $L_0 = \frac{n\Lambda_0^2\cos\theta}{\lambda}$, indicating that an optical beam encounters several acoustic wavefronts as it traverses the acoustic wave at an angle θ, then the

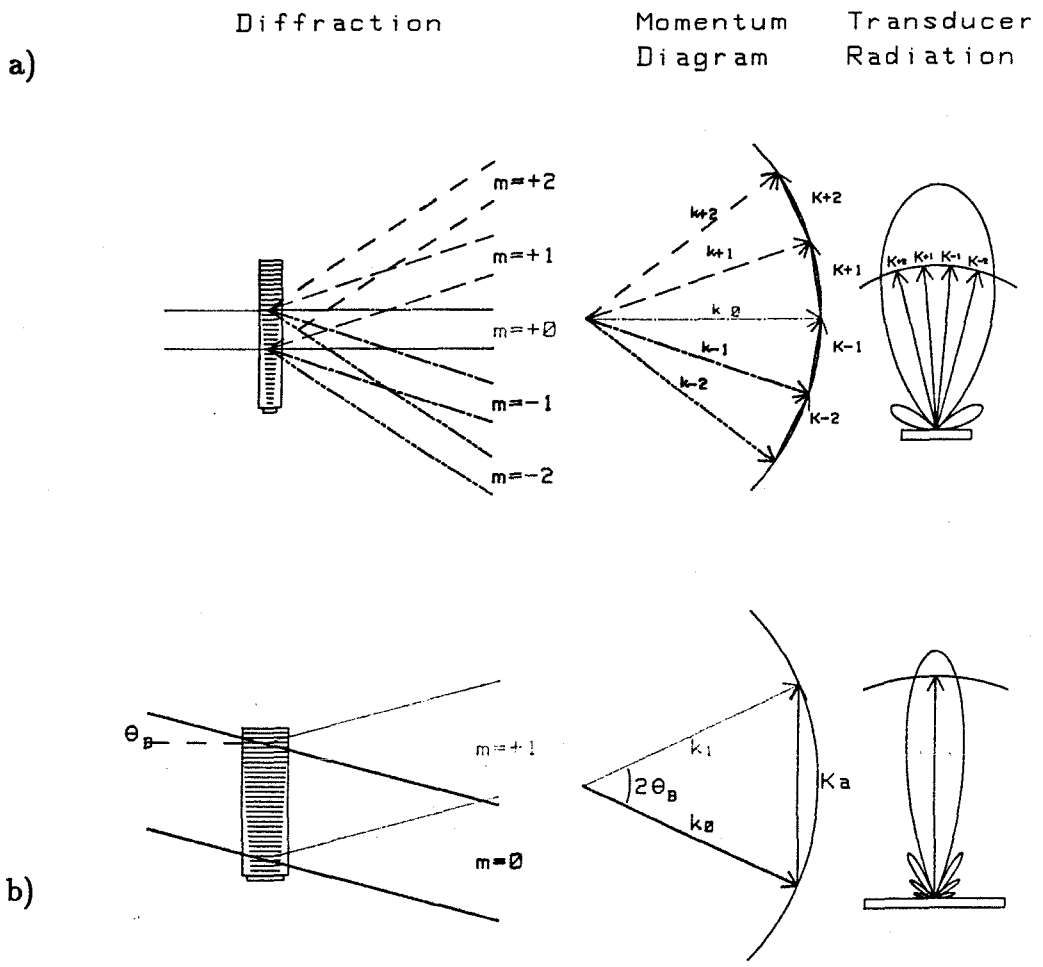
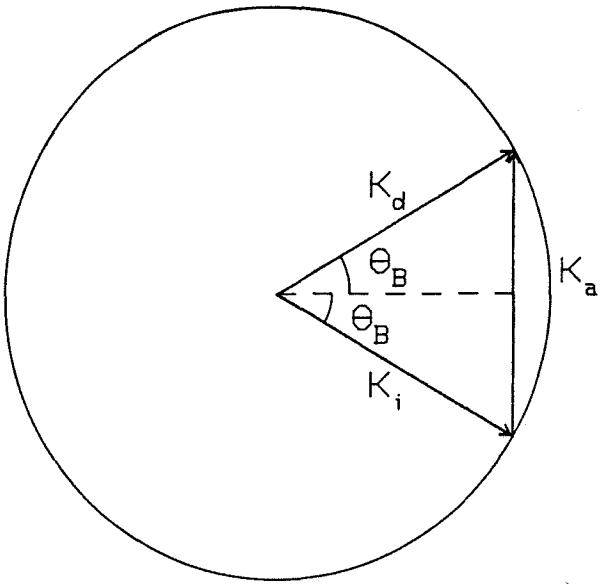


Figure 4.2.3. a) The Raman-Nath regime of acousto-optic diffraction, b) The Bragg regime, with incidence at the Bragg angle.

a)



b)

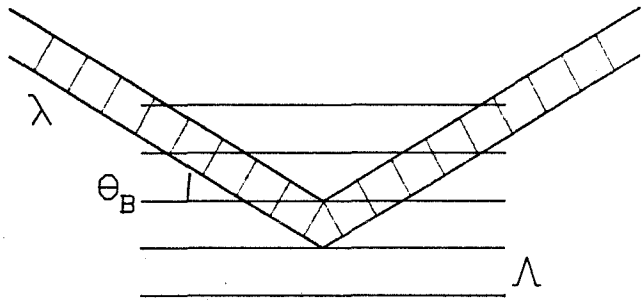


Figure 4.2.4. a) Momentum matching conditions for isotropic Bragg diffraction in the interaction plane, b) coherent reflection off of individual acoustic wavefronts.

interaction is in the Bragg regime. When the incident light is at the Bragg angle with respect to the acoustic wave, then only one diffracted order is allowed with either $m = \pm 1$, which is illustrated in Figure 4.2.3b. This describes a quantum mechanical particle scattering interaction between three particles, the incident photon, the diffracted photon, and the acoustic phonon. In order to conserve energy the diffracted photon is doppler shifted by the moving phase grating by an amount exactly equal to the frequency of the acoustic phonon. This frequency shift can only be observed interferometrically, and it plays a key role in the spectrum analysis systems analyzed in Chapter 5. The momentum matching equation (4.2.26) describes a closed triangle in momentum space whose vertices, corresponding to the incident and diffracted wave, must be allowed optical eigenmodes. This is illustrated in Figure 4.2.4a for the isotropic case, such as would be observed in fused silica, in which case the closed triangle is isosceles and $|\vec{k}_i| = |\vec{k}_d|$, because the change in the optical frequency and energy due to the acoustic wave is less than one part in 10^6. The angles between the acoustic plane wavefronts and the incident and diffracted optical beams are the same, and are called the Bragg angle θ_B. This direction corresponds to constructive interference between different portions of the diffracted beam. It can be derived by modelling each acoustic wavefront as a partially reflecting mirror, and requiring the reflected portion from adjacent wavefronts to add up in phase as is illustrated in Figure 4.2.4b. The resulting Bragg angle is given by

$$\theta_B = \sin^{-1}\frac{|K|}{2|k|} = \sin^{-1}\frac{\lambda}{2n\Lambda} = \sin^{-1}\frac{\lambda f}{2v_a n} \approx \frac{\lambda f}{2v_a n} \qquad (4.2.28)$$

The angular deviation of the diffracted light is seen to be proportional to the frequency of the applied acoustic wave in the small angle approximation, which is the basis of many acousto-optic systems.

In an anisotropic medium the optical and acoustic momentum vectors can be a function of direction and polarization. The acousto-optic interaction can result in a change of the state of polarization of the diffracted optical wave as well as

the direction of propagation. The technologically most relevant case is when the incident optical polarization is in the slow mode, and an acoustic mode with the proper photoelastic tensor coefficient is used to switch the state of polarization to the fast mode[14]. This case is illustrated in Figure 4.2.5a for a positive uniaxial crystal, where the angle of incidence θ_1 has a corresponding extraordinary index $n_s(\theta_1) = n_e(\theta_1)$, and the angle of diffraction θ_2 has a corresponding ordinary index $n_f(\theta_2) = n_o$. This type of birefringent diffraction can result in an increased bandwidth for a given efficiency by using tangentially degenerate phase matching, where the midband acoustic momentum vector is parallel with the nearest part of the inner optical momentum surface. For an acoustic beam rotated by an angle β, the angles of incidence $\theta_i = \theta_2 - \beta$, and diffraction $\theta_d = \theta_1 + \beta$, measured with respect to the normal to the acoustic beam, are the appropriate coordinates for describing the interaction. The momentum matching condition requires the acoustic momentum to have a magnitude which closes the triangle in momentum space.

$$|K_a| = \frac{2\pi}{\Lambda} = k_s(\theta_i)\sin\theta_i + k_f(\theta_d)\sin\theta_d = \frac{2\pi n_s(\theta_1)}{\lambda}\sin\theta_i + \frac{2\pi n_f(\theta_2)}{\lambda}\sin\theta_d \quad (4.2.29)$$

The optical momentum wave vectors must also be conserved orthogonal to the rotated acoustic beam which leads to an expression for the full angle through which the light is diffracted.

$$\sin(\theta_i + \theta_d) = \sin(\theta_1 + \theta_2) = \frac{\lambda_0}{n_s(\theta_1)\Lambda}\cos(\theta_d) = \frac{\lambda_0}{n_f(\theta_2)\Lambda}\cos(\theta_i) \qquad (4.2.30)$$

For an acoustic plane wave propagating along the optical symmetry axis with $\beta = 0$, the expressions for the required angle of incidence and diffraction of a polarization switching anisotropic diffraction mechanism are modified from the isotropic or polarization preserving relationship, yielding the Dixon equations[14].

$$\begin{aligned}
\sin\theta_i &= \frac{\lambda}{2n_s\Lambda}\left(1 + \frac{\Lambda^2}{\lambda^2}(n_s^2 - n_f^2)\right) = \frac{\lambda f}{2n_s v_a}\left(1 + \frac{v_a^2}{f^2\lambda^2}(n_s^2 - n_f^2)\right) \\
\sin\theta_d &= \frac{\lambda}{2n_f\Lambda}\left(1 + \frac{\Lambda^2}{\lambda^2}(n_f^2 - n_s^2)\right) = \frac{\lambda f}{2n_f v_a}\left(1 + \frac{v_a^2}{f^2\lambda^2}(n_f^2 - n_s^2)\right)
\end{aligned} \qquad (4.2.31)$$

These relations are plotted for the same birefringence indicated in the anisotropic Bragg matching diagram in Figure 4.2.5. The broad minimum of θ_i can be used as a fixed incidence angle which yields a wide range of diffraction angles for different input frequencies. When the slow and fast modes are degenerate, $n_s(\theta_1) = n_f(\theta_2)$, then this is seen to reduce to the familiar Bragg diffraction condition of isotropic media where $\theta_i = \theta_d$.

When the acoustics are also anisotropic, then v_a varies with the acoustic beam angle β, and the Bragg matching condition is most easily solved for graphically, as illustrated in Figure 4.2.6. In this figure the diffracted optical momentum vector is found at the intersection of the optical normal surface with the perturbing acoustic normal surface of the appropriate scale for the input frequency which is centered on the input optical momentum vector eigenmode. The acoustic momentum surface can be approximated with a taylor series expansion in the vicinity of the crystal symmetry axis, and to second order the acoustic momentum vector can be represented as an asymmetric paraboloid.

$$|K_a(\beta_y, \beta_z)| = (1 + a_y\beta_y + a_z\beta_z - b_y\beta_y^2 - b_z\beta_z^2)\frac{2\pi f}{v_0} \qquad (4.2.32)$$

In this expression β_y is the off axis angle in the height direction, β_z is the off axis angle in the interaction plane, and v_0 is the acoustic velocity along the symmetry axis. The linear terms in β_y and β_z represent on axis acoustic walkoff effects, and are identically zero if the transducer face is accurately aligned normal to an acoustic symmetry axis. The intersection of this paraboloid with the optical normal surface gives the locus of exactly phase matched momentum vectors of the diffracted optical field, and this is illustrated in the random dot stereo pair shown in Figure 4.2.7a. In the x-z interaction plane we can derive a relationship between the incident and

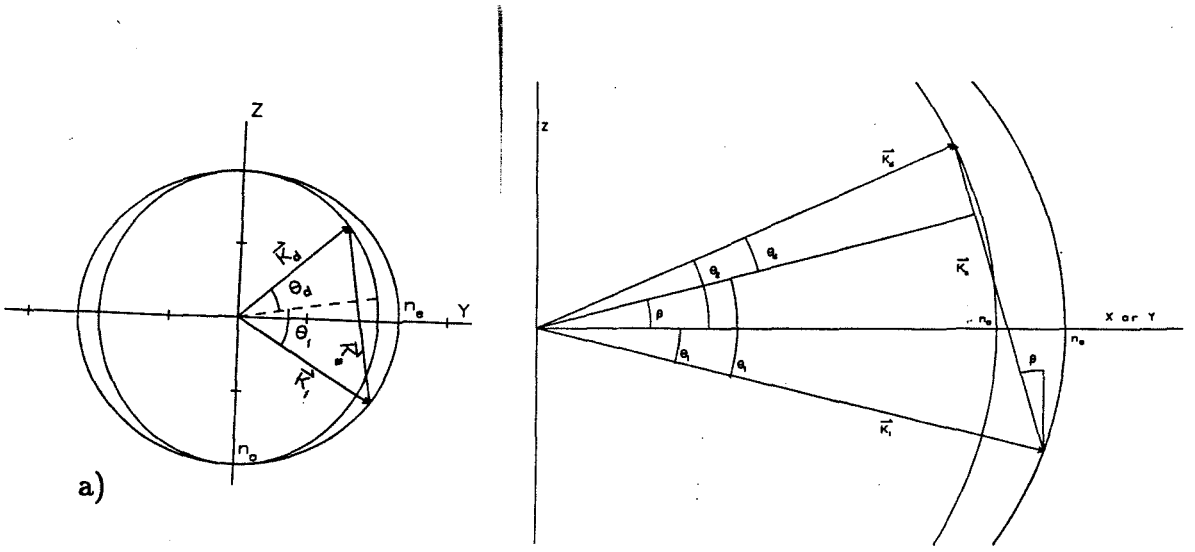


Figure 4.2.5. a) Anisotropic phase matching for positive uniaxial crystal with 15% birefringence, b) the corresponding perfectly phase matched incidence and diffracted angles.

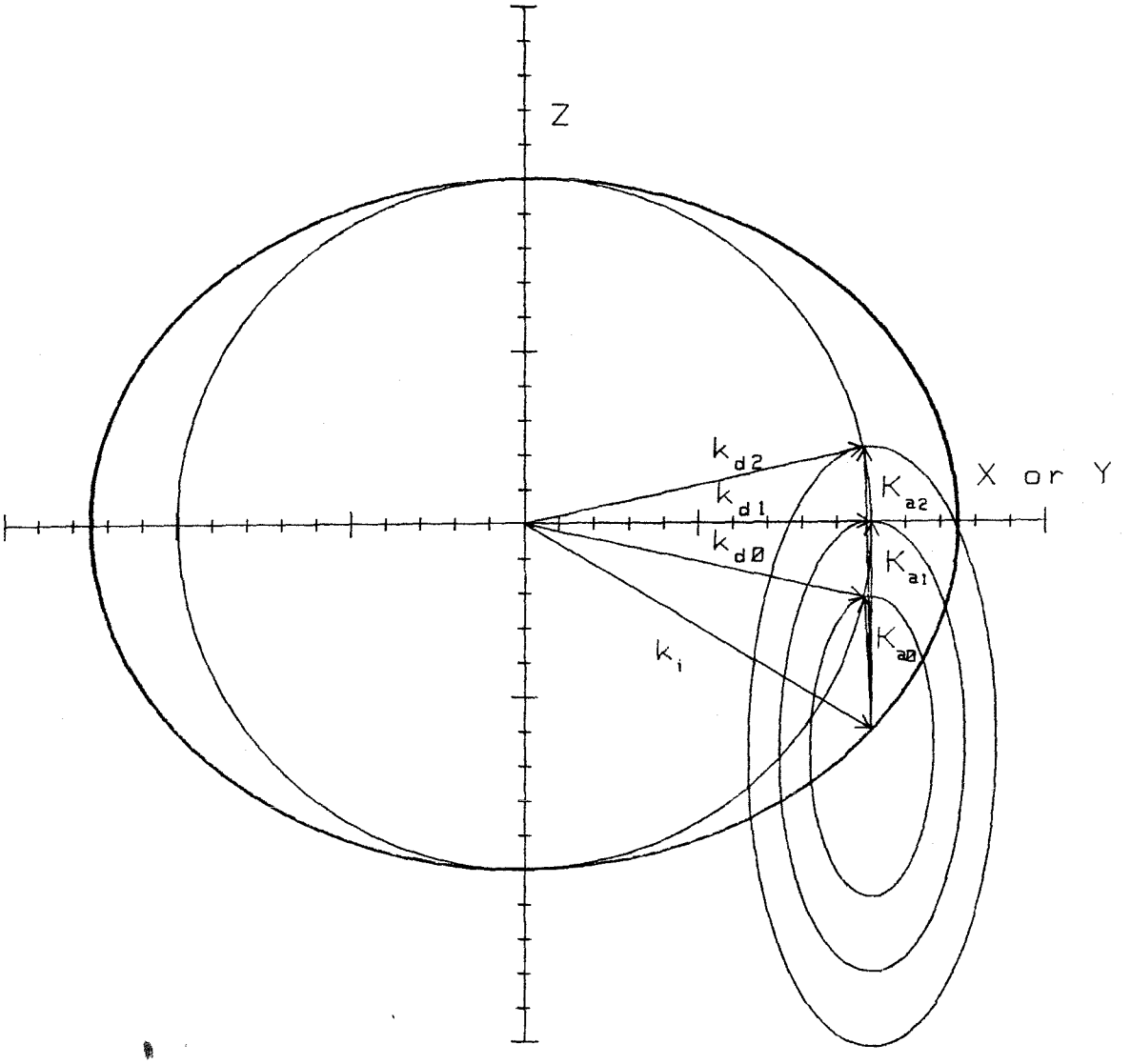


Figure 4.2.6. a) Tangential degenerate birefringent phase matching for positive uniaxial crystal when both acoustics and optics are anisotropic.

diffracted optical angles with the acoustic off axis angle β_z.

$$n_f \sin(\theta_1) - n_s \sin(\theta_2) = n_f(\theta_2)\sin(\theta_d - \beta_z) - n_s(\theta_1)\sin(\theta_i - \beta_z) = (1 - b_z\beta_z^2)\frac{\lambda f}{v_0} \tag{4.2.33}$$

This relationship coupled with Eqs. 4.2.29-30 allows us to uniquely solve for the diffraction angle in arbitrarily anisotropic media. However, from the systems point of view it usually suffices to assume a linear relationship between input frequency and diffraction angle. In tellurium dioxide the acoustic anisotropy of the anomalously slow shear [110] mode is so large that these effects must be considered in attempting to explain non ideal AOD operation, such as frequency plane scale distortion and frequency plane blur[15]. An illustration of the normal surface intersection for the case of the TeO_2 slow shear mode with a greatly exaggerated frequency is shown as a random dot stereo pair in Figure 4.2.7b. The Schaeffer-Bergman diffraction pattern is given by the momentum conserving interactions in a uniformly excited acoustic medium, which are just the intersection of the diffracted wave vector momentum surface with the perturbing acoustic momentum surface, giving the locus of phase matched diffraction allowed eigenmodes.

4.2.4 1-Dimensional coupled mode solutions

I will solve the 1-dimensional acousto-optic coupled mode equation assuming a perfectly phase matched interaction, while assuming that the modal amplitudes are only a function of the z coordinate, even though this is not completely rigorous. For an incident optical plane wave with polarization $\hat{E}_0$ and incident amplitude $A_0(0)$, that is diffracted by the acoustic wave into a plane wave with polarization $\hat{E}_1$ whose amplitude grows with interaction distance z within the AO medium, the total field can be written in the interaction regime $0 < z < L$, as

$$\vec{E}(\vec{x},t) = A_0(z)\hat{E}_0 e^{-i(\omega_0 t - k_{x0}x - k_{z0}z)} + A_1(z)\hat{E}_1 e^{-i(\omega_1 t - k_{x1}x - k_{z1}z)} \tag{4.2.34}$$

In the presence of the dielectric perturbation given by Eq. 4.2.24 the two modes

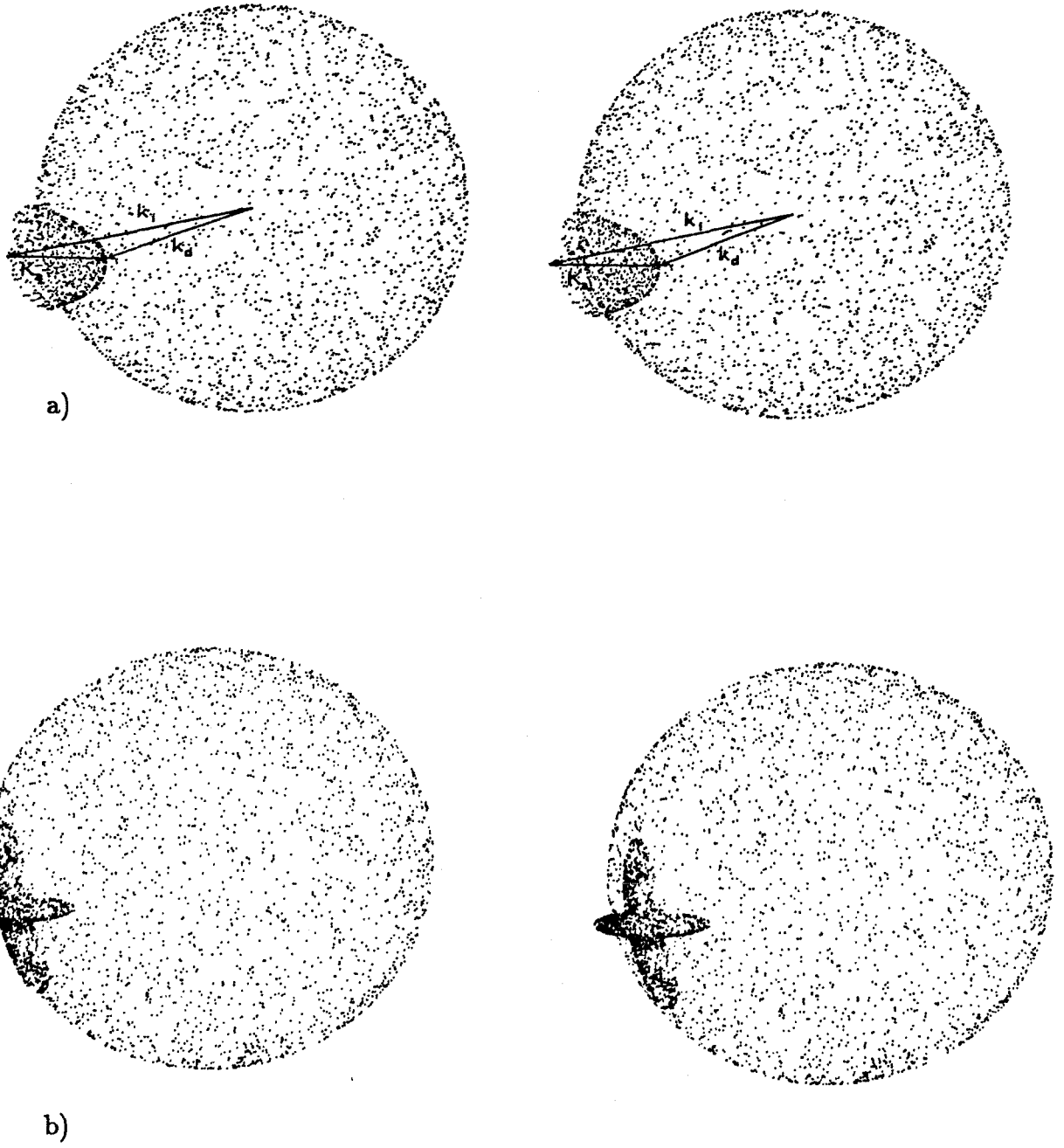


Figure 4.2.7. a) 3-D random dot stereo pair normal surface intersection for an arbitrary acoustic paraboloid. b) stereo pair of the normal surface intersection for the slow shear acoustic mode of TeO_2 with a greatly exaggerated acoustic frequency.

are coupled, provided that momentum and energy are conserved. The total field is a solution of the inhomogenous wave equation

$$\nabla^2 \vec{E}(\vec{x},t) - \mu(\epsilon + \delta\epsilon(\vec{x},t))\frac{\partial \vec{E}(\vec{x},t)}{\partial t^2} = 0 \tag{4.2.35}$$

The individual fields with constant amplitudes are solutions of the unperturbed homogenous wave equation which allows us to substitute the total field and the acoustically induced dielectric perturbation in the wave equation, and obtain for the one dimensional interaction configuration

$$\sum_{m=0,1} \hat{E}_m e^{-i(\omega_m t - k_{xm}x - k_{zm}z)} \left[\frac{\partial^2}{\partial z^2} - 2ik_{zm}\frac{\partial}{\partial z}\right] A_m(z) \tag{4.2.36}$$
$$+ \frac{W}{2}\epsilon_{ik}(\mathrm{p}_{klmn}\hat{S}_{mn})\epsilon_{lj}\left[e^{-i(\Omega t - K_x x - K_z z)} + \mathrm{c.c.}\right]\mu \sum_{l=0,1} \omega_l^2 A_l(z)\hat{E}_l e^{-i(\omega_l t - k_{xl}x - k_{zl}z)} = 0$$

The adiabatic condition allows us to assume that the modal amplitudes $A_m(z)$ are slowly changing in z compared to an optical wavelength so that we can neglect the second order spatial derivative. We now require that the coefficients of $\hat{E}_m e^{-i(\omega_m t - k_{xm}x - k_{zm}z)}$ vanish for $m = 0,1$, in order to obtain a phase synchronous transfer of power between the modes. This requires that $k_{x1}x - k_{z1}z = k_{x0}x - k_{z0}z \pm K_x x - K_z z$ and $\omega_1 = \omega_0 \pm \Omega$, which are the familiar momentum and energy conservation conditions.

For Bragg interactions only a single order has a totally phase synchronous interaction with the input fields so we only get coupling to a single mode. This results in the well known coupled mode equations for perfectly phase matched interactions.

$$\begin{aligned} \frac{dA_0(z)}{dz} &= -i\kappa A_1, \\ \frac{dA_1(z)}{dz} &= -i\kappa^* A_0. \end{aligned} \tag{4.2.37}$$

Where the coupling constant is found from the incident and diffracted polarization vectors left and right projected onto the dielectric perturbation tensor.

$$\kappa = \frac{\omega}{4}\hat{E}_0^* \cdot \Delta\epsilon \hat{E}_1 = \frac{\omega}{4}|E_{0i}^* \cdot \frac{W}{2}\epsilon_{ik}(\mathrm{p}_{klmn}\hat{S}_{mn})\epsilon_{lj}E_{1j}| \tag{4.2.38}$$

The solution of the coupled mode equation with the implied boundary condition of $A_1(0) = 0$, yields the perfectly coupled mode solution for $z > 0$

$$\begin{aligned} A_0(z) &= A_0(0)\cos|\kappa|z, \\ A_1(z) &= -i\frac{\kappa}{|\kappa|}A_0(0)\sin|\kappa|z. \end{aligned} \tag{4.2.39}$$

So the optical field is seen to oscillate back and forth between the incident and diffracted field each distance $2L_m = \pi/|\kappa|$, as the beams propagate in the z direction. Complete power transfer occurs from the input to the diffracted mode after a distance of z propagation given by L_m. The diffraction efficiency is the ratio of the incident optical intensity to the intensity transferred to the diffracted beam in a distance L, and it is given by

$$\frac{I_2}{I_1} = \frac{|A_1(L)|^2}{|A_0(0)|^2} = \sin^2|\kappa|L \tag{4.2.40}$$

Thus given the amplitude W of the acoustic mode $\hat{S}_{kl}$, the interaction length L, the input polarization $\hat{E}_0$, the output polarization $\hat{E}_1$, and the material tensors ϵ_{ij} and p_{ijkl}, we can find the percentage of light diffracted into the output beam.

4.2.5 Acousto-optic devices

An acousto-optic device is constructed by bonding an acousto-electric transducer onto a photoelastic medium, so that acoustic waves can be launched into the medium, and is illustrated in Figure 4.2.8. The transducer is usually a piezoelectric crystal of thickness t_0, metalized on both faces so that an electric field can be applied transversely in the k direction. This induces a strain through the third rank piezoelectric tensor $\mathbf{d}$, which can only exist in non-centrosymmetric materials.

$$\mathbf{S}_{ij}(t) = \mathbf{d}_{ijk}E_k = \mathbf{d}_{ijk}v_k(t)/t_0 \tag{4.2.41}$$

The appropriate choice of transducer crystal cut and orientation are used to produce the desired polarization of the acoustic wave which is launched into the photoelastic medium. The time dependent strain within the transducer is coupled into the

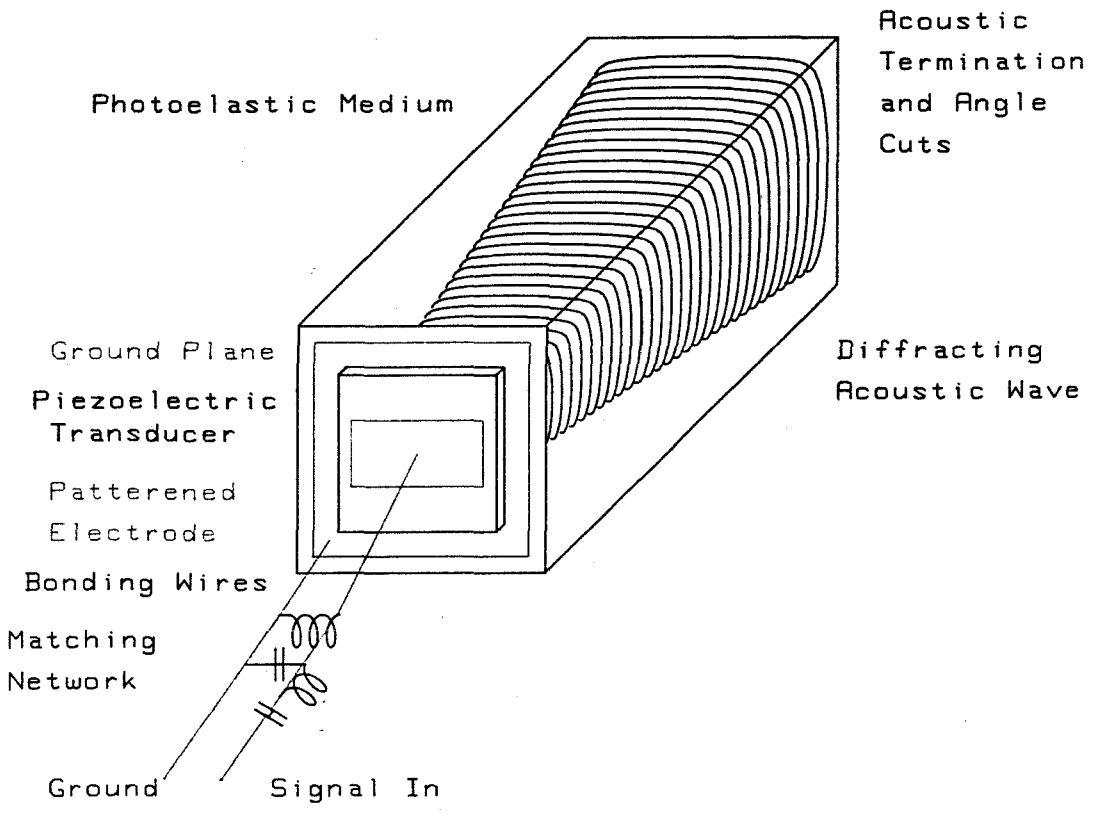


Figure 4.2.8. Schematic diagram of an acousto-optic device.

photoelastic medium with a frequency dependent efficiency dictated by the acoustic impedance matching of the transducer and bonding layers. The frequency dependence of the electro-acoustic conversion $R(f)$, is due to the electrical matching network, the transducers resonant bandwidth and the mechanical coupling efficiency bandshape. By applying a sinusoidally varying electric field to the transducer, that is within its acoustic resonant bandwidth, we can launch a propagating acoustic wave into the photoelastic medium. Since the transducer has a finite spatial aperture the harmonic acoustic field will have a spatial angular divergence. The angular spectrum of the transducer is given by the Fourier transform of its aperture $p(y,z)$, scaled by the appropriate acoustic wavelength. The angular divergence of the transducer in the interaction dimension coupled with the phase matching condition is what determines the acousto-optic bandshape of the device. For an isotropic device with a simple uniform rectangular transducer of length L and height H the transducer radiation pattern acoustic angular spectrum in the interaction dimension is given by the simple 1-D transform $\text{sinc}^2(L\beta_z/\Lambda)$. The resulting isotropic acousto-optic bandshape can be derived from the viewpoint of phase mismatched interaction, or from the viewpoint of selecting the appropriate angular component from the transducer radiation field in order to obtain perfect phase matching, and these different interpretations are schematically illustrated in Figure 4.2.9a. The isotropic acousto-optic bandshape is given in terms of the perfectly matched on axis frequency f_m, normalized by the center frequency f_0, and the normalized interaction length, as a function of the normalized frequency $F = f/f_0$.

$$W(F) = \text{sinc}\left[\frac{L}{2L_0}F(F_m - F)\right] \tag{4.2.42}$$

Thus the transducer length L determines the acousto-optic bandwidth, and it determines the effective interaction length as well, thereby affecting the diffraction efficiency.

In the tangentially degenerate approach to birefringent phase matching[16], the

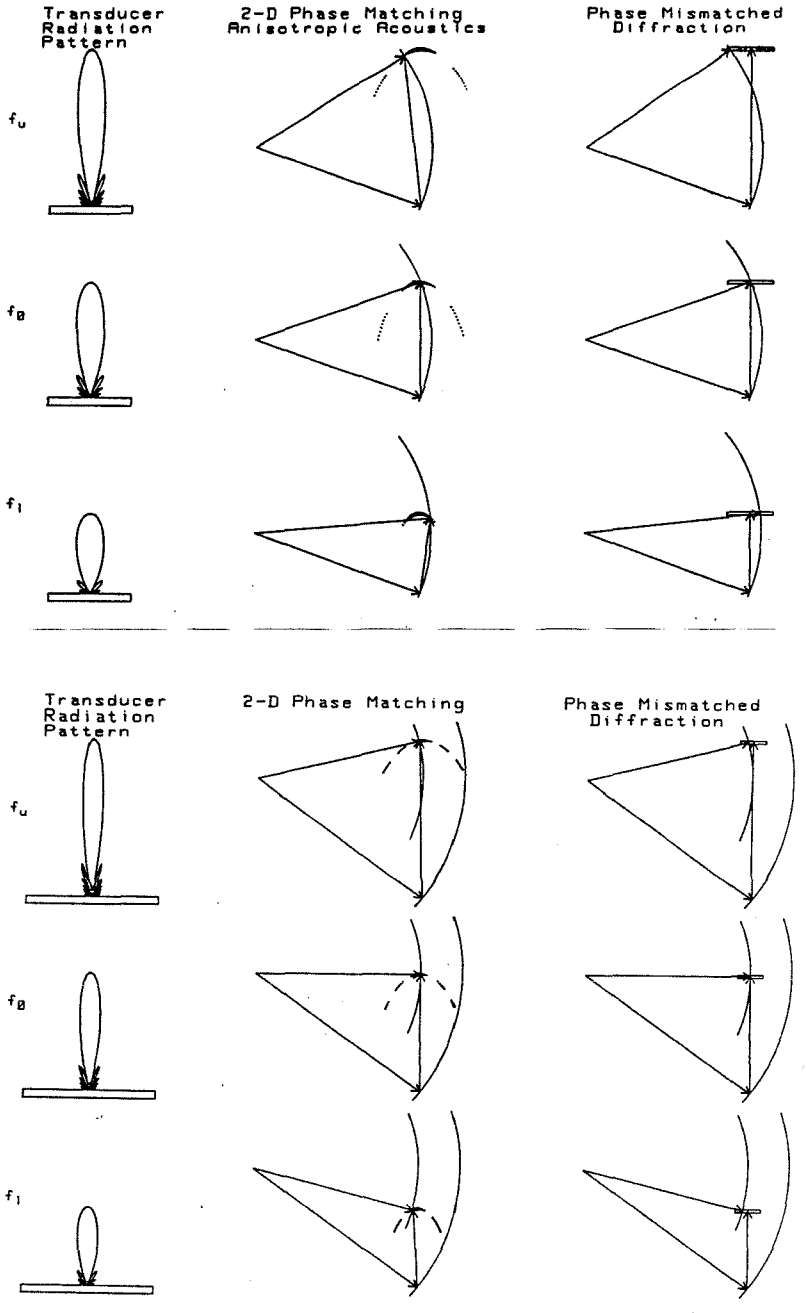


Figure 4.2.9. Transducer angular radiation pattern, 2-dimensional phase matching, and phase mismatched diffraction, for the lower, midband and upper frequencies of a) an isotropic device and b) a tangentially degenerate phase matched device.

acoustic wave vector is tangent to the locus of fast diffracted wave vectors at the symmetric center frequency f_1, and the peak of the transducer angular spectrum intersects the diffracted wave vector surface at two frequencies $f_m^{\pm} = f_0 \pm \Delta f/2$. The resulting birefringent acousto-optic bandshape is broadened, and is given in terms of normalized frequency variables.

$$W(F) = \text{sinc}\left[\frac{L}{2L_0}(F - F_1)^2 - \frac{\Delta F}{2}\right] \tag{4.2.42}$$

When the symmetrical frequency is at midband and phase matched $\Delta f = 0$, then $f_m = f_0 = f_1$, and the tangentially matched acousto-optic bandshape has a simple quadratic frequency dependent phase mismatch term $(F-1)^2$. The decreased transducer angular bandwidth needed by a device operating in the tangentially degenerate phase matching regime is illustrated in Figure 4.2.9b. The transducer can be longer and a larger interaction length results in a greater diffraction efficiency, and the bandshape becomes symmetrical.

The transducer height H determines the degree of collimation in the orthogonal dimension which determines the diffraction limited usable aperture time. For the isotropic acoustic case the transducer radiation pattern remains within the near field in the height dimension for a distance $D = H^2/\Lambda$, which is equal to the distance that the transducer angular spectrum at the 4dB half width intersects the transducer geometric shadow. In the anisotropic acoustic case the near field in the height dimension is modified by the acoustic curvature factor giving a near field distance $D' = H^2/\Lambda(\frac{1}{2} - b_y)$. The transducer area LH relates to the acoustic power density through the equation $P_a = \frac{1}{2}\rho_m v_a^2 |S|^2 LH$. and thus determines the optical diffraction efficiency in terms of the electrical power input to the transducer. A tradeoff between the diffraction efficiency and the interaction bandwidth can be tailored by an appropriate choice of the transducer length to height ratio L/H.

A Bragg cell or acousto-optic deflector (AOD) is the basic electrical to optical transducer used in the systems studied in this thesis. From the system point of view

this device is modelled as a 1-D travelling wave modulator with a finite aperture window. Thus when a temporal signal $s(t)$ is applied to the device, the resulting transmission function of an idealized device is represented as

$$t(x,y) = \alpha s(t - x/v_a)w(x) \tag{4.2.43}$$

The acoustic velocity v_a is taken to be the nominal velocity along the symmetry axis normal to the transducer. The window function $w(x)$ is the product of the device finite aperture window with the acoustic attenuation, and the input optical beam gaussian profile is often included as well, thereby making the window function a hybrid window of the induced polarization field which is the product of input field and dielectric perturbation. This device model is sufficient for simple system calculations, but it ignores the nonideal behavior of a real device as discussed in this section. In considering the proper use of a device in a system, it is necessary to minimize the effects of its nonideal behavior, such as its polarization response[17], off axis diffraction effects[15], and intermodulation nonlinearities[18]. A more complete description of a Bragg cell response[19] is based on a superposition of its response at each frequency across the device bandwidth. When we apply a broadband signal $s(t)$ to the transducer each of its Fourier components launches a spatially diffracting harmonic travelling wave into the photoelastic medium. The temporal Fourier decomposition of a single sideband of the input signal is given by

$$\tilde{s}(t) = \int \tilde{S}(f)e^{i2\pi ft}df \tag{4.2.44}$$

The device response to an individual Fourier component is a frequency dependent acoustic diffraction integrated along the direction of optical propagation. Thus we can write a more general description of the diffracted field from the device for an input plane wave at the Bragg angle.

$$a_d(x,y,z,t) = e^{-i2\pi[\nu t + k_z z]}i\int t_f(x,y)\tilde{S}(f)e^{i2\pi ft}dt \tag{4.2.45}$$

The function $t_f(x,y)$ is the frequency dependent amplitude transmission of the Bragg cell. It can be obtained from the 3-D acoustic Sommerfeld diffraction integral, integrated along the direction of optical propagation to form a 2-D projection of the acoustic pressure density field[20,21,22].

The momentum space representation of the amplitude transmission of the device is given by the spatial Fourier transform of $t_f(x,y)$, and it is exactly what would be observed at the Fourier plane of an AO spectrum analyzer illuminated by a plane wave[23], and it more clearly illustrates the dependence of the device behavior on the physical mechanisms discussed previously in this section.

$$T_f(k_x,k_y) = \int t_f(x,y)e^{-i2\pi(k_x x+k_y y)}dxdy = \eta R(f)W(f)S(k_x,k_y) \qquad (4.2.46)$$

The constant $\eta = \Delta n \frac{\pi L}{\lambda \cos\theta_0}$ is just the amplitude diffraction efficiency derived from the coupled mode model and the transducer geometry, with a skewed interaction length $L/\cos(\theta_0)$. The acousto-electric conversion efficiency $R(f)$ is the frequency and phase response of the tuning network, the transducer and the coupling to the photoelastic medium. The acousto-optic bandshape $W(f)$ is determined by the type of diffraction , either normal or birefringent, and upon the direction of propagation of the input wave and the transducer orientation and shape. The spatial frequency angular spectrum of the device, $S(k_x,k_y)$, is the result of the phase matching conditions applied to the travelling wave anisotropic acoustic propagation.

For the ideal travelling wave modulator device model of equation 4.2.42, the frequency dependent spatial frequency spectrum takes an especially simple form.

$$T_f(x,y) = \eta e^{i\pi A[k_x - \frac{f}{v_a}]}\text{sinc}\left[A\left(k_x - \frac{f}{v_a}\right)\right]\delta(k_y) \qquad (4.2.47)$$

This equation represents a single diffracted spot in the Fourier plane, whose frequency resolution is limited by the device aperture in the acoustic propagation dimension. The phase factor represents the plane wave acoustic phase delay accumulated in the propagation to the center of the device where the optical axis

is located. The term $\delta(k_y)$ represents the Fourier transform of an infinitely high transducer, and in general it will be replaced by a 2-dimensional Fourier transform of the transducer aperture, $P(\frac{\beta_y}{\Lambda}, \frac{\beta_z}{\Lambda})$. For an input optical field propagating in the x-z plane the transducer angular spectrum is sampled along the $k_z = 0$ slice, equivalently projecting the 3-D diffracting acoustic wavefront by an integration of the acousto-optic modulation along the path of optical propagation, via the projection slice theorem.

For an anisotropic medium, the acoustic propagation has additional curvature factors which enter into the phase matching conditions, and the acoustic propagation equations yielding a more complex spatial frequency angular spectrum.

$$\begin{aligned} S_{k_z}(k_x, k_y) &= \int_0^A e^{ix\frac{\Lambda_0}{2}[K_y^2(b_y-\frac{1}{2})+K_z^2(b_z-\frac{1}{2})]} P(K_y, K_z)dx \\ &= e^{i\pi A\left[k_x-\frac{f}{v_a}+\frac{\Lambda_0}{2}[K_y^2(b_y-\frac{1}{2})+K_z^2(b_z-\frac{1}{2})\right]} \\ &\quad \text{sinc}\left[A\left(k_x-\frac{f}{v_a}+\frac{\Lambda_0}{2}[K_y^2(b_y-\frac{1}{2})+K_z^2(b_z-\frac{1}{2})]\right)\right] P(K_y, K_z) \end{aligned} \tag{4.2.48}$$

The curvature factors $(b_y-\frac{1}{2})$ and $(b_z-\frac{1}{2})$ are the anisotropic curvature factors of the energy flow or acoustic Poynting vectors, and are derived as the normal vector of the momentum surface. This represents a curved locus in the Fourier domain weighted by the finite aperture sinc in the propagation dimension and the transducer angular spectrum in the height dimension. When a collimated plane wave is incident on the device at a specific $\vec{k}$ then a slice out of this spatial frequency spectrum is sampled by the integrating optical propagation.

In the experiments performed in this thesis a more complicated mode of operation is utilized, where the optics is focussed into the center of the acoustic column. In this case the individual slices of the spatial frequency spectrum will generate diffraction components due to each plane wave component of the input beam which must be summed up to find the full diffracted wavefront. An additional complication in anisotropic media is the fact that the acousto-optic bandshape function

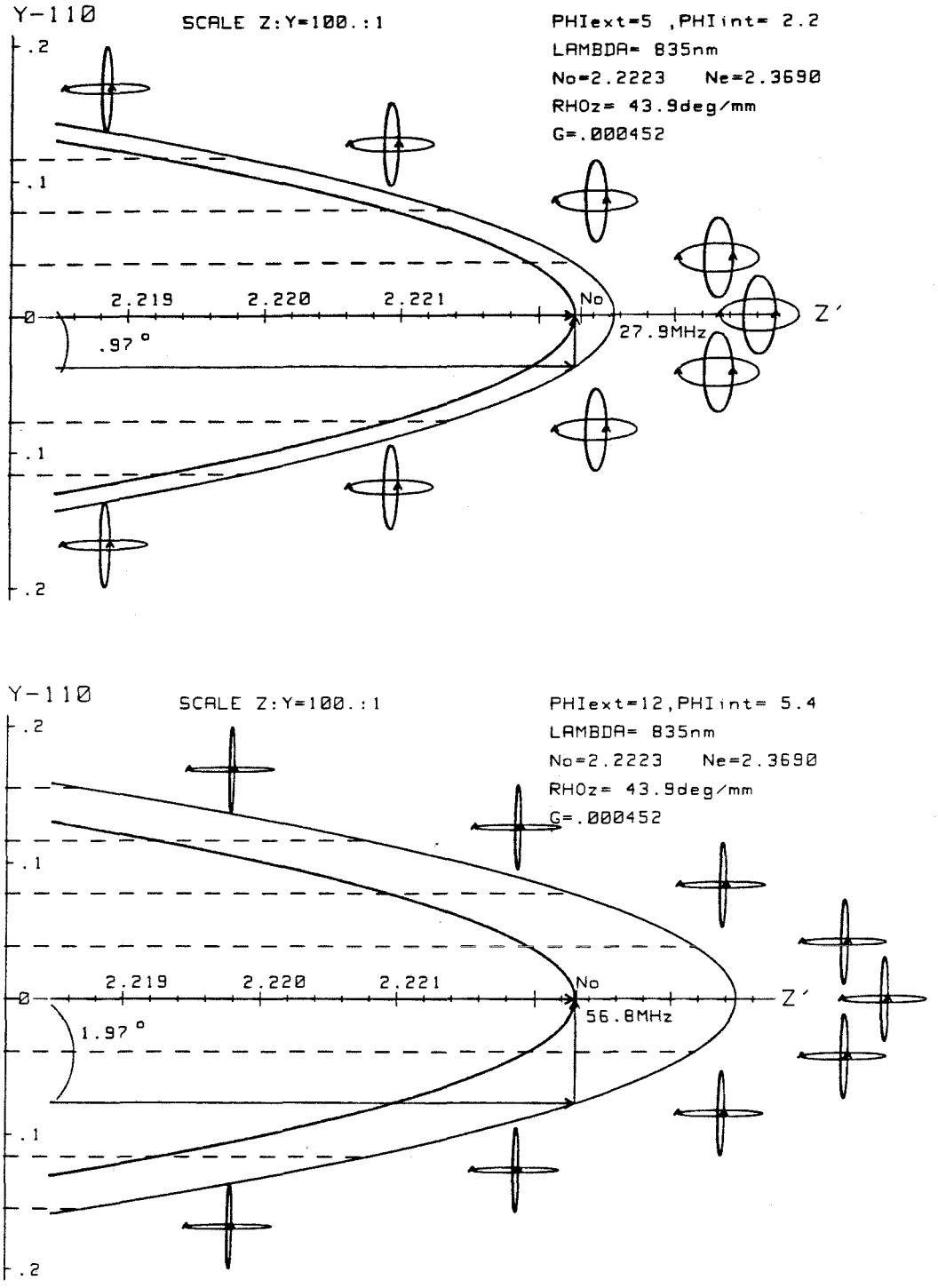


Figure 4.2.10. Cross sections through the optical momentum surfaces for TeO_2 at external off axis propagation angles of 5 and 12 degrees, showing the polarization eigenstates at 1 degree increments and the optically rotated tangential degenerate frequencies.

is also a function of the input optical propagation direction $W_{\bar{k}}(f)$, and as the off axis angle increases the frequency corresponding to the tangential degeneracy also increases. In optically active media such as TeO_2[24,25] the polarization states of the diffracted optical wave also depend on the off axis propagation angle because the birefringence begins to dominate the degeneracy splitting of the optical rotary power. These effects are illustrated for off axis propagation angles of 5 and 12 degrees in TeO_2 in Figure 4.2.10., with the appropriate birefringence and rotary powers interpolated for 835 nm wavelength light[26], and these figures should be compared with the on axis optical normal surface plotted in Figure 4.2.2. A final complication is the dependence of the effective photoelastic constant and the coupling efficiency on the angle of propagation through the interaction medium. To minimize these off axis diffraction effects a limited range of off axis angles should be used to illuminate a slow shear TeO_2 Bragg cell.

4.2.6 Acousto-optic modulators

Another important device is the acousto-optic point modulator or AOM, which is similarly constructed by bonding an acousto-electric transducer to a photo-elastic medium. The principal difference between an AOM and an AOD is that the AOM is used with a tightly focussed optical spot incident on the center of the acoustic column and the AOD is illuminated with a collimated wave. This is because the modulator must have a quick access time in order to achieve wide modulation bandwidth, which is the inverse of the acoustic propagation time across the beam waist of the focussed spot[27]. In order to be able to modulate each spatial frequency component of the focussed optical field there must be a corresponding phase matched frequency component of the diffracting acoustic wavefront, thereby dictating a short transducer length L. This leads to a tradeoff between diffraction efficiency and access time. The Bragg angle centered focussing incident beam and the diverging

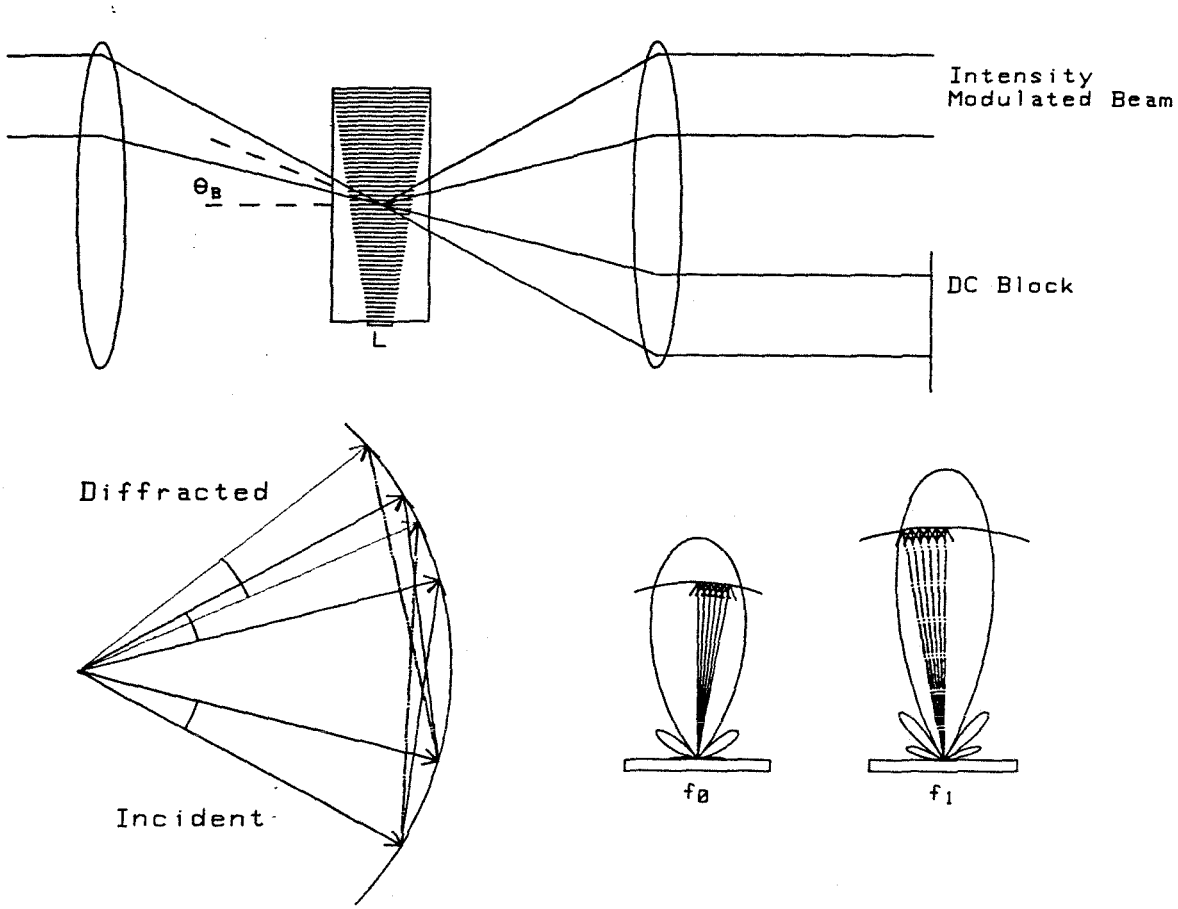


Figure 4.2.11. An acousto-optic modulator operating in the Bragg regime.

diffracted wave must be separated in the Fourier plane for single sideband suppressed carrier modulation, where the diffracted intensity is proportional to the envelope modulation of the acoustic carrier. This leads to a requirement that the carrier frequency be several times higher than the access time limited modulation bandwidth, which in turn requires that several acoustic phase fronts be contained within the optical beam waist. The intensity modulation of the diffracted beam by the envelope of the acoustic signal can be considered to be a mixing of the carrier with the sidebands which are responsible for the amplitude modulation. When the modulation frequency increases the diffraction angle of the sidebands increases so that less of the diffracted cones from the sidebands overlap with the carrier. Optical heterodyning is optimized when the interfering components are colinearly propagating, so as the modulation frequency increases the modulation depth decreases, which is an alternative derivation of the modulator bandwidth limitation. This type of modulator arrangement is illustrated in Figure 4.2.11, and the overlap of the diffracted waves in momentum space is illustrated for a frequency near the modulator 3dB bandwidth limitation.

An alternative modulation geometry that results in a phase modulated optical wave is obtained when the optical beam waist is much smaller than an acoustic wavelength, and is illustrated in Figure 4.2.12. In this case the optical beam sees at any one time an instantaneous acoustic density which homogenously slows down the optical wave uniformly across the aperture of the focussed beam. As an example of this regime of operation, consider $LiNbO_3$ which has a longitudinal acoustic velocity of 6.57 mm/μsec. This indicates that at an acoustic frequency of 50 MHz the acoustic wavelength will be 131.4 μm, and the optical wave can easily be focussed to a much smaller 10-30 μm spot. This type of phase modulation requires the production of a large number of Bessel function weighted temporal sidebands, which are a characteristic of Raman-Nath regime AO diffraction, and since the optical

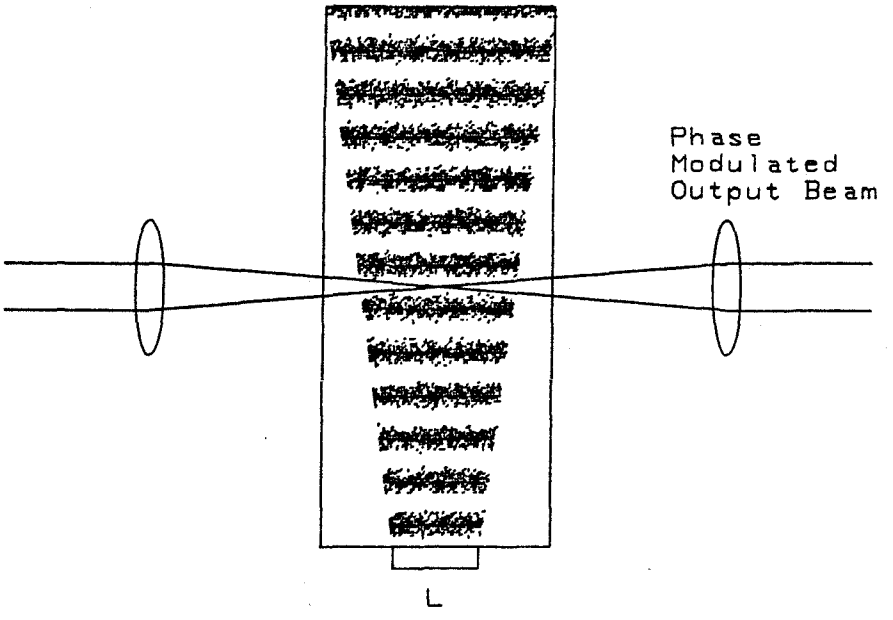


Figure 4.2.12. An acousto-optic undulation phase modulator.

wave passes through less than one acoustic wavefront it is no longer a Bragg regime modulator. This type of acoustic undulation modulator is similar to an electrooptic modulator, and the phase modulated signal can be considered to be produced by the interference of the appropriately weighted temporal sidebands, which are essentially colinear spatially. A short transducer is again required in order to contain within its angular bandwidth appropriate components needed to diffract different optical momentum components several times in order to produce the appropriately weighted sidebands. An acoustic phase modulator of this nature is only useful in the context of an interferometric system, because the intensity is unmodulated by the passage of an acoustic wave, and the diffraction angle is far less than the optical momentum width.

A Bragg cell can be used as a modulator as well as a deflector, but because of the long transducers that are utilized with narrow acoustic angular spectrums a limit is placed on the amount of focussing of the incident optical wave that is allowed in order to have a phase matched acoustic component for each momentum component of the input optical field. This places a limit on how small of an aperture that can be illuminated and therefore limits the modulation bandwidth, or inverse access time. In the experiments where I used an AOD as a phase modulator the incident light was collimated and incident at the Bragg angle, and it was pulsed with a very narrow pulse several thousand times a second. In this case access time is not a constraint, and the Bragg cell can be loaded with a phase shifted sinusoidal tone on each pulse. The diffracted phase modulated and doppler shifted plane wave is easily separated from the undiffracted wave in the Fourier plane since the AOD is deep within the Bragg regime and illuminated with a plane wave.

4.2.7 Systems considerations

It is possible to apply the system considerations examined in this thesis to the device design in order to obtain an optimized device for a particular systems application. Although in general it may not be an economically viable option to have a special purpose device custom made for a particular systems research project, it is another degree of freedom that should be considered when systems issues cannot be resolved with existing commercial devices. As an example I will consider a couple of specialized device designs that might be considered in the context of some of the systems discussed in this thesis.

A number of the processing systems that I have investigated involve the interferometric detection of acoustically phase modulated data diffracted from a Bragg cell. This can be accomplished with split path interferometers, or it can be accomplished with a self referencing in line approach, where a coherent local oscillator is added to the signal and the sum is applied to the transducer of a travelling wave Bragg cell. Each of these signal components produces a diffracted field which propagate through the optical system and interfere on the detector. In the systems I have examined the reference wave must be used to illuminate a large number of channels, while the signal is focussed into each channel by the optical system, thereby achieving a large processing gain. This means that the reference wave should have an amplitude N times larger than that of an individual signal component, where N is the number of channels that the signal is separated into, and also the processing gain of the signal channelization operation. In order to obtain sufficient optical intensity when a pulsed source is used for illumination it is necessary to achieve a high diffraction efficiency. With only one signal present the strong reference will push the device up into the nonlinear regime of the coupled mode theory. In this case acoustic nonlinearities and acousto-optic multiple diffractions will produce unwanted intermodulations between the signal and reference which may degrade system performance.

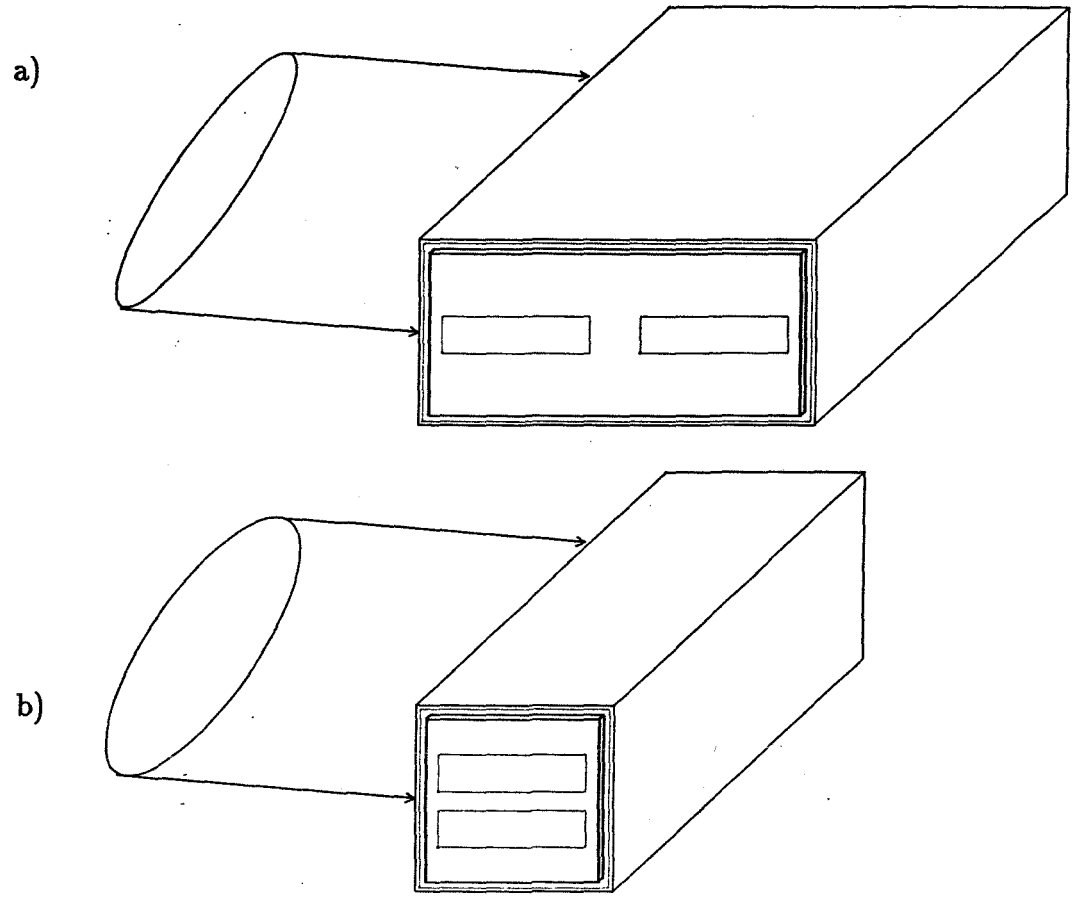


Figure 4.2.13. Dual transducer Bragg cells for self referencing systems, a) in interaction dimension, b) in height dimension.

One possible solution[28] could be the use of the double transducer device pictured in Figure 4.2.13a, where the signal is applied to the first transducer, and the reference is applied to the second transducer. Both diffracted waves will be generated and will interfere as before, but because the acoustic columns are separated, the acoustic nonlinearities between the signal and reference will not be generated. Both diffracted wavefronts are generated from within the same crystal so the system will still have the same interferometric stability of the self referencing approach. This approach will not eliminate the acousto-optic multiple diffraction effects so an alternative vertically stacked double transducer device pictured in Figure 4.2.13b might be used if the additional vertical interferometric carrier can be tolerated or utilized in the system operation.

In the synthetic aperture radar (SAR) processor discussed in section 3.3, it is necessary to perform a chirp correlation on a TDI CCD, which is accomplished by multiplying a time dependent chirp by a spatial chirp and detecting in a moving coordinate frame. When range curvature is considered the locus of focussed range energy should be a curved locus whose radius of curvature decreases with the target range. The spatial profile along the curved locus of focussed energy should be a chirp, with a range dependent spatial chirp rate. An interesting approach to the generation of these spatial patterns is to turn the temporal chirps at various range delays into different temporal frequencies by mixing the radar return with a properly matched long duration reference chirp of the same chirp rate[29]. These sinusoidal tones can be applied to a Bragg cell spectrum analyzer that is illuminated with a CW plane wave rather than a focussed pulsed source. A 2-D Fourier transform of the diffracted field will produce spatially curved loci modulated along the vertical dimension with the Fourier transform of the acoustic transducer aperture in the height dimension, which is illustrated in Figure 4.2.14 for the case of a vertically chirped transducer. Usually when a rectangular transducer is utilized this vertical

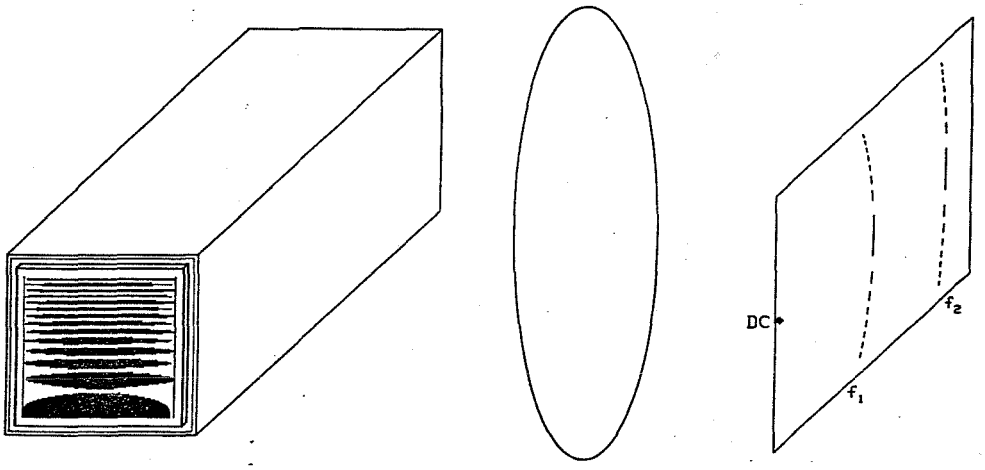


Figure 4.2.14. Fourier plane diffraction loci that will result from the use of a vertically chirped transducer.

modulation is a simple $\text{sinc}^2[H\beta_y/\Lambda]$ function of the transducers height. To obtain a chirp spatial modulation along these loci in the Fourier plane requires an inverse synthesis problem that may not be solvable with the constraint of the transducer being a uniform zero-one function determined by the transducer patterning. One simple possibility is to use the fact that the Fourier transform of a chirp is a chirp of inverse curvature, and pattern the transducer vertically with a hard clipped version of the transform of the desired spatial chirp, as is illustrated in Figure 4.2.15a. However, the Fourier transform of a hard clipped chirp is a poor reproduction of the desired spatial chirp modulation. The projection slice theorem applied to the transducer in the direction of optical propagation through the acoustic wave, allows the synthesis of a more accurate representation of a spatial chirp by sinusoidally modulating the interaction length along the vertically chirping transducer, as is illustrated in Figure 4.2.15b. A final alternative is to attempt to implement a bipolar spatial chirp transducer by using a split electrode phased array [30]chirping transducer with an interaction length modulation as shown in Figure 4.2.15c. In this case the signal is applied across the two split electrodes, and the ground plane remains floating at an intermediate potential, thereby producing a push-pull acoustic wavefront. This type of complicated transducer design simplifies some aspects of the processing system, but perhaps at an unreasonable expense of device complexity, and it introduces other problems at the systems level as well. For example, the spatial chirps in the Fourier plane all have the same vertical chirp rates because of the cancelling effects of the decrease of the acoustic diffraction angles with higher frequencies coupled with the increased length of the acoustic momentum vector. In a SAR processing geometry where range curvature is significant, the range azimuth coupling which makes the azimuth chirp rate a function of the range will also be significant. This additional problem can be overcome at the systems level or at the device level but will not be considered here. An alternative is to consider this

transducer geometry in the context of folded spectrum processing, where a vertical slit modulated by a chirp is the desired spatial impulse response for each coarse frequency tone. This can be obtained if a self collimating acoustic mode[31] is utilized with $b_z = \frac{1}{2}$, so that the phase matched diffraction locus in the Fourier plane has zero curvature.

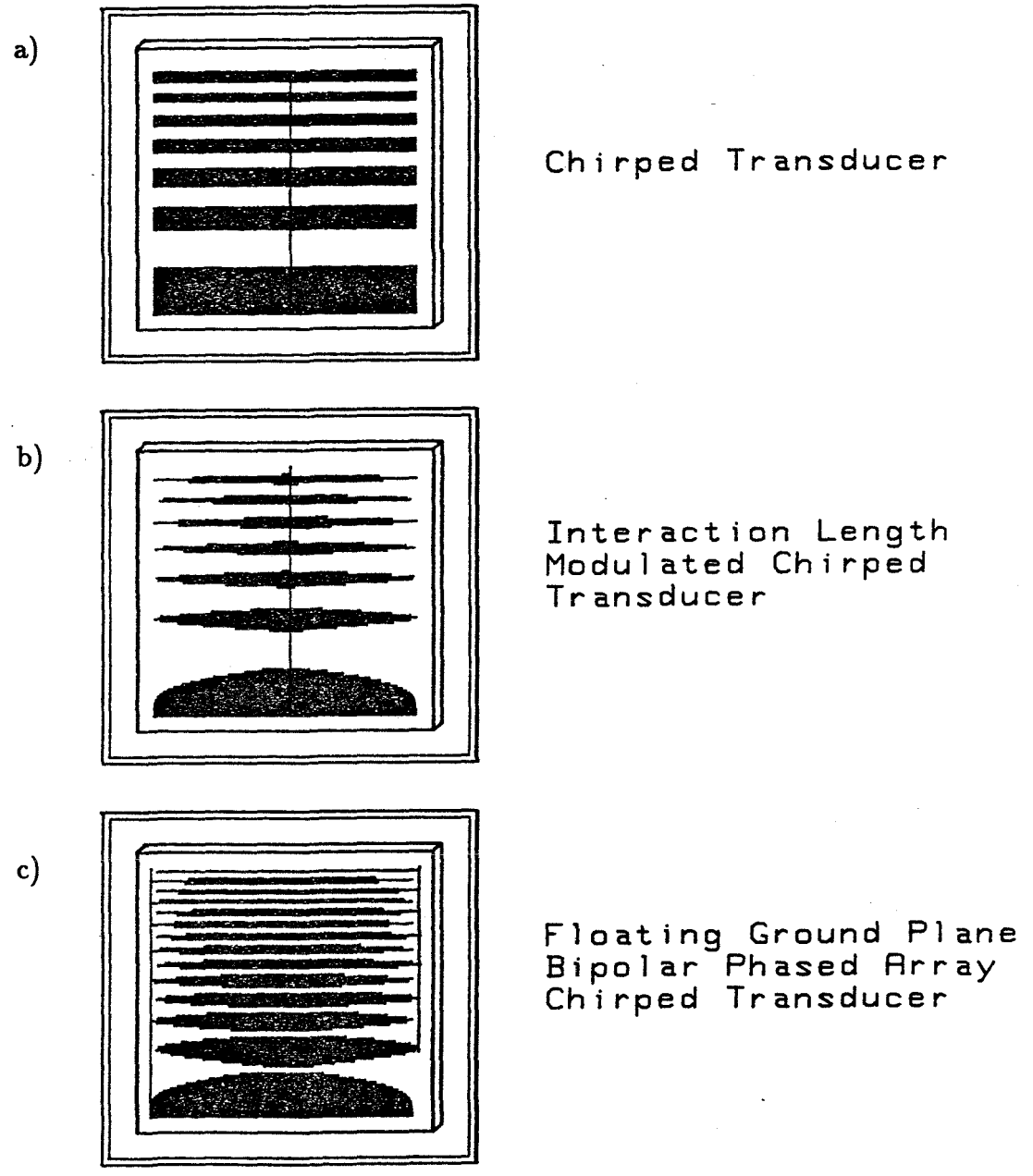


Figure 4.2.15. Chirped transducers producing quadratically diffracting acoustics, a) one-zero chirp transducer, b) interaction length modulated chirp transducer, c) bipolar vertical phased array chirp transducer.

4.3 Charge Coupled Device Detector Arrays

A charge coupled device (CCD)[34–45] detector array is composed of a matrix of photosensors, with an associated charge transfer readout structure. The photosensors can either be photodiodes, which can be gated into the charge transfer readout register, or the photosensing can take place within the photosensitive CCD readout array itself. The CCD consists of a closely spaced array of metal-oxide-semiconductor (MOS) capacitors, usually constructed out of doped polysilicon gates, deposited on top of a silicon dioxide insulator which is grown on the surface of a silicon substrate. When a bias voltage of the proper polarity is applied to the gate of a MOS capacitor it repels the nearby majority carriers in the bulk semiconductor and creates a depletion region which acts as a potential well that can accumulate mobile minority carriers. In the context of an optical photodetector, the accumulation of minority carriers is due to the photogeneration of electron-hole pairs in the substrate, followed by the collection of the mobile minority carriers by the depleted potential well. After an integration period the potential well fills up with an amount of photogenerated carriers proportional to the locally sampled incident intensity. The sampled photogenerated charge representation of the incident image needs to be transferred to an output node, where the accumulated signal can be read out, and this transfer process is accomplished through the use of charge coupling. Charge coupling takes place when two MOS capacitors are placed in close proximity, or preferably with overlapping edges, so that the accumulated carriers within the joint potential well attempt to minimize the configurational energy of the system by flowing to the region of highest potential, corresponding to the deepest well. Charge packets can be transferred along a linear array of overlapping MOS capacitors with very little signal degradation by the application of appropriately phased clocking waveforms, as long as a barrier is always kept between adjacent packets during the transfer process. This requires three distinct clocking phases

per unit cell, two to accomplish the transfer, and the third to prevent merging of adjacent potential wells. However by forming fixed potential steps, directionality can be built into the charge transfer process with fewer clocking phases. At the final stage of the array an output node is used that can sense the size of the charge packet within the potential well, and this signal is amplified and read off chip as a voltage or current proportional to the amount of light which was accumulated by each potential well. This is usually followed by some off chip signal processing to clean up the signal, removing clock feedthrough and correlated noise, and this signal is sampled and quantized by a digital to analog converter.

The simplest type of charge coupled devices are formed out of surface channel MOS capacitors, where the potential well is formed at the Si-SiO_2 interface. The energy band diagram corresponding to a p-substrate surface channel MOS capacitor, biased with a positive voltage in order to repel holes and form a depletion region at the surface, is illustrated in Figure 4.3.1. At the top of the figure the parabolic band bending in the depletion region is illustrated for an empty potential well, and in the bottom of the figure the reduced band bending of a well partially filled with accumulated electrons is shown. The depletion region width collapses, and the interface potential decreases as more charge is stored in the potential well, until the well is completely full at which point no more charge can be accumulated. The full well capacity per unit area is approximately given by the capacitance per unit area times the gate voltage divided by the electronic charge, and for a typical $100\mu m^2$ pixel the capacity is approximately $2x10^6$ electrons. An overlapping electrode surface channel CCD structure is schematically illustrated in Figure 4.3.2, which shows an input diode coupled through a pair of four electrode unit cells with two phase clocking, coupled through a final gate into an output diode. The potential wells at the surface under the electrodes at successive times in the clocking operation are illustrated under the cross sectional diagram, along with the two phase clocking

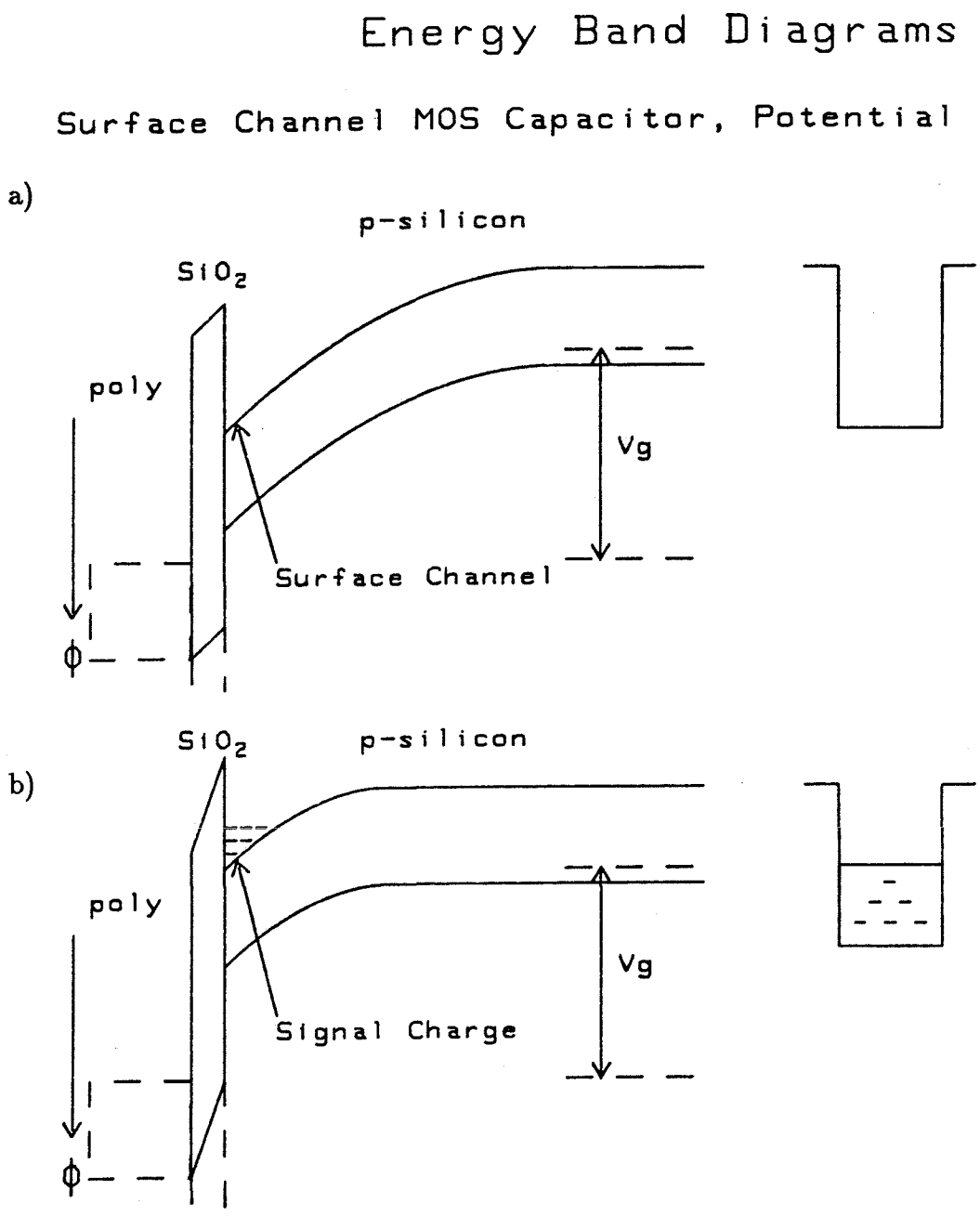


Figure 4.3.1. Surface channel MOS capacitor energy band diagram a) empty potential well, b) partially full well.

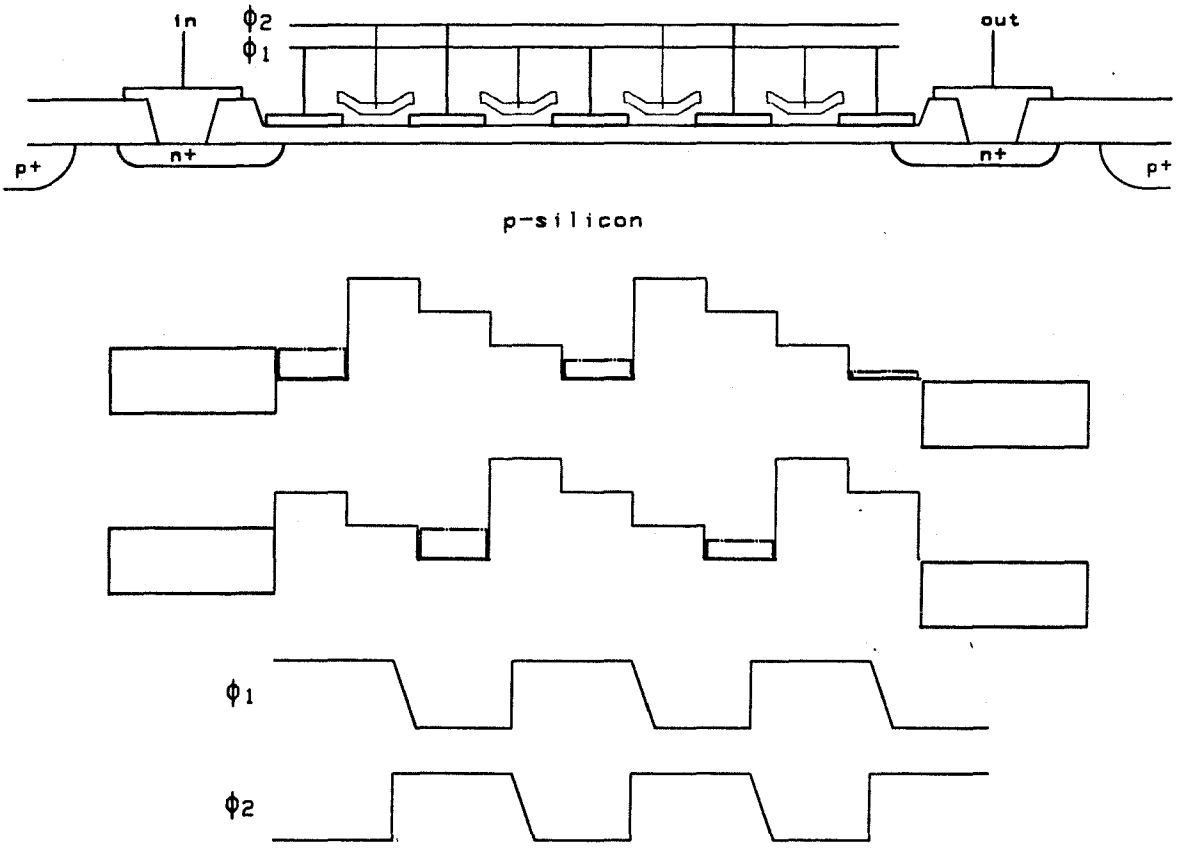


Figure 4.3.2. Cross section through a two phase surface channel CCD, the surface potential at two times during the transfer process, and the clocking waveforms.

waveform. The clock waveforms have slow fall times in order to allow stored charge under one electrode time to diffuse to the next phase. There are three mechanisms of charge transfer in CCDs, the first 99% of the charge is transferred by the self repulsion of the electrons, diffusion is responsible for the last bit of charge transfer, and it can be assisted by fringing field drift. The main problem with surface channel devices is the fast interface states at the Si-SiO_2 surface, which trap electrons from a potential well as it is travelling by, and release it into a subsequent potential well that passes by. This introduces a trailing decay of a passing impulse, that can introduce serious impulse blurring when a large number of transfers are used, even though surface channel transfer efficiency is usually about .9999 at a transfer rate of several MHz.

The buried channel (or bulk channel) charge coupled device (BCCD) was introduced to overcome the limitations of surface channel devices[36]. In a BCCD a surface implanted layer of the opposite type to the substrate is used to form a junction a few microns deep, and shift the potential maximum away from the surface. The corresponding energy band diagram for a buried channel MOS capacitor is illustrated in Figure 4.3.3 for the case of a p-type substrate with an n-type surface layer with 4 times the bulk doping density. The parabolic band bending in the depleted device is due to the integration of the fixed doping charge in the solution of Poisson's equation, and the energy band curvature, or second derivative, changes sign and magnitude at the junction. As the potential well fills up the mobile electrons cancel the fixed space charge and produce a flat band region near the bottom of the potential well. A cross section through a two level polysilicon overlapping gate buried channel CCD is shown in Figure 4.3.4, along with the bulk channel potential energy profiles, and the two phase clocking waveforms. The potential energy distribution in the buried channel is more graded because each position feels the influence of the neighboring electrodes, rather than being pinned at the surface potential of the

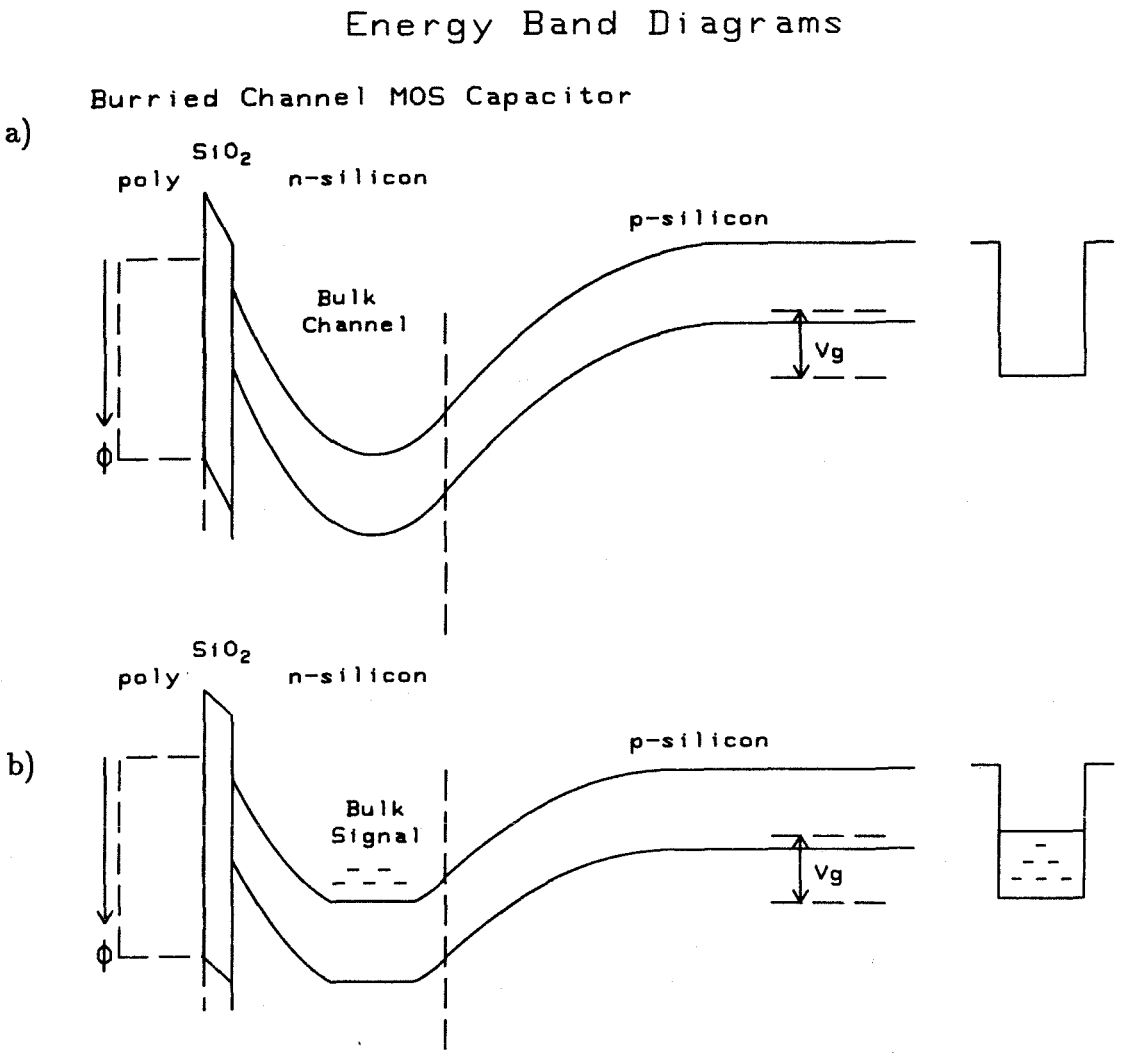


Figure 4.3.3. Buried channel MOS capacitor energy band diagram a) empty potential well, b) partially full well.

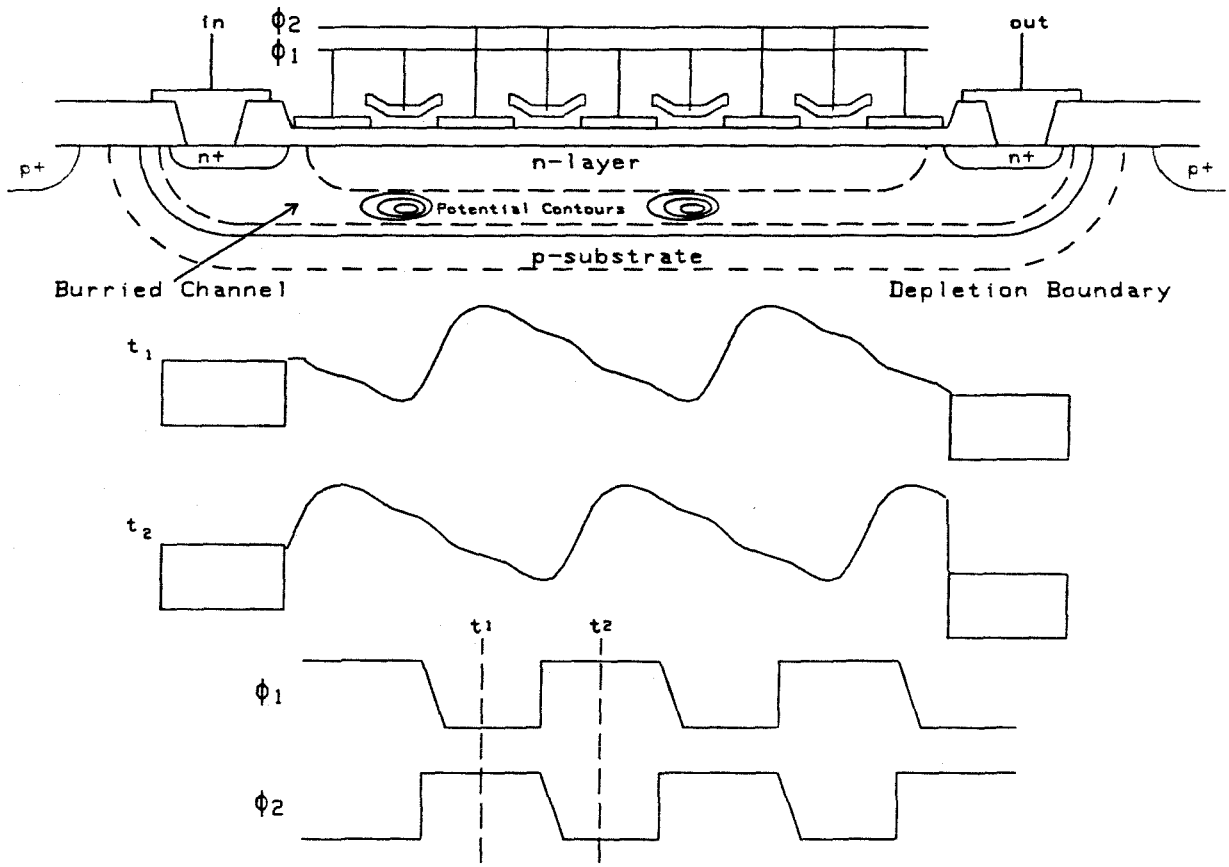


Figure 4.3.4. Cross section through a two phase buried channel CCD, the bulk potential at two times during the transfer process, and the clocking waveforms.

closest electrode. This results in much larger fringing fields, given by the potential gradient, which assist in the charge transfer, greatly increasing the achievable transfer speed and improving the charge transfer efficiency to about .99999 at several MHz clocking frequency. It is interesting to note that in a buried channel device the deepest potential minimum is under the electrode with the thicker oxide layer, and for this reason the clock wires have been ordered differently from Figure 4.3.2, in order to achieve the same direction of charge transfer. In buried channel devices it is very important to use overlapping electrodes, because interelectrode gaps lead to charge collecting potential valleys, while in surface channel devices these gaps lead to charge confining potential hills. Since charge is stored in the bulk the surface trapping states do not degrade this transfer efficiency, but bulk leakage does cause thermally generated carriers to fill up the potential wells. The charge storage capacity of a buried channel device is typically less than half that of an equal area surface channel device.

The output stage of a CCD must sense and amplify the individual charge packets while introducing as little distortion and noise as possible. A floating gate above a CCD channel, schematically shown in Figure 4.3.5a, will sense a potential induced by a charge packet travelling underneath it, and this can be used to nondestructively measure the signal charge contained in a potential well. The signal charge induces a proportional voltage on the floating gate through a capacitive coupling that is amplified by modulating the current through an on chip MOSFET. The area of the floating gate is proportional to the induced image charge, and this fixed multiplicative gain factor combined with the nondestructive sensing inherent to the floating gate can be used in a CCD tapped delay line filter known as the split electrode filter. The floating diffusion output circuit shown in Figure 4.3.5b is another common charge sensing technique. It is operated by resetting the implanted diffusion to a reference potential and then transferring the charge packet through a final out-

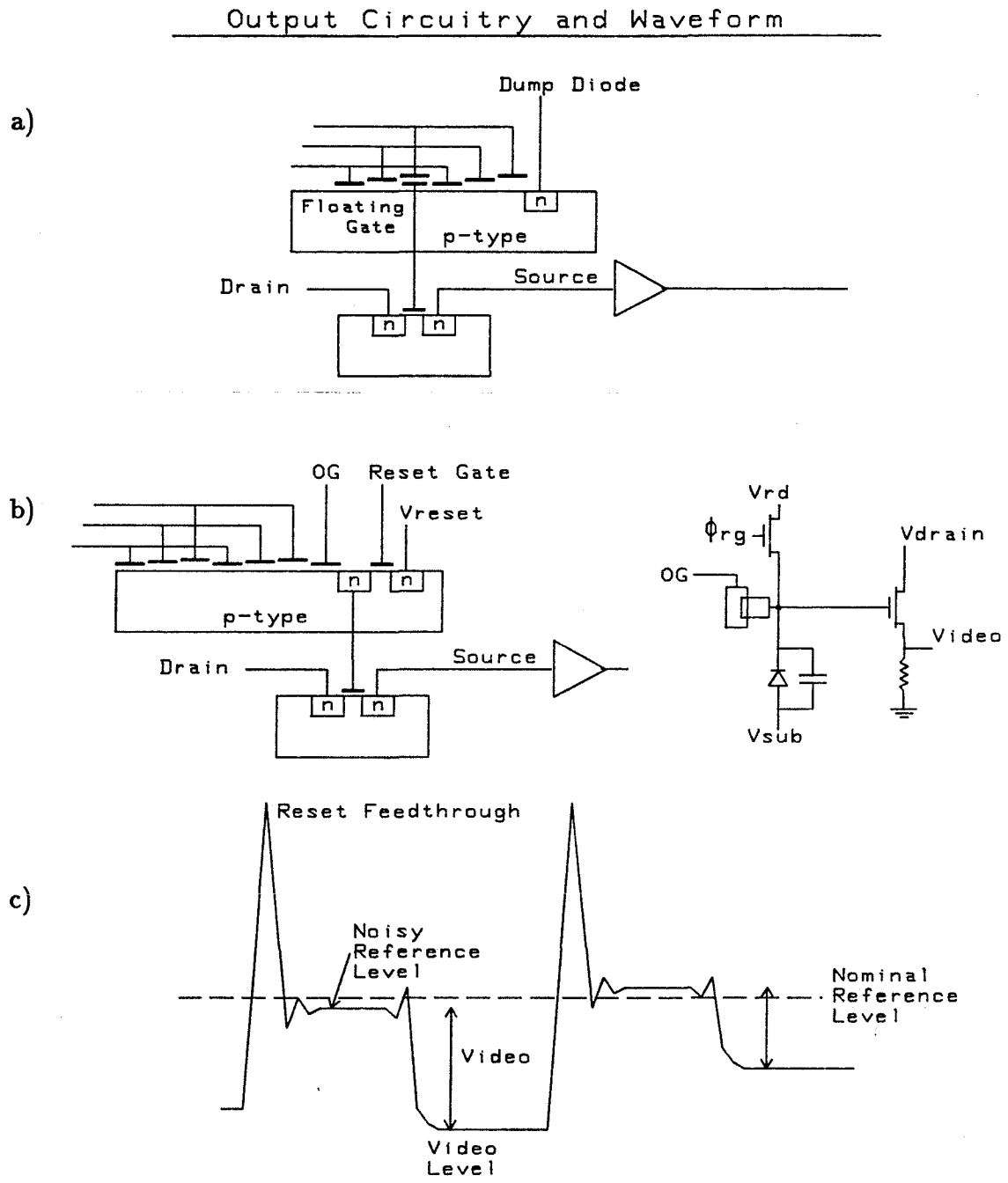


Figure 4.3.5. a) Floating gate schematic representation, b) floating diffusion output circuitry and its equivalent circuit, and c) a schematic of the output waveform that is generated by the floating diffusion circuit.

put gate onto the floating diffusion. The charged up floating diffusion capacitance is connected to the gate of a MOSFET source follower amplifier, which acts as a charge to current transducer, producing a proportional current which can be read off chip as a voltage across a resistor or into a low noise preamplifier. The output waveform is shown in Figure 4.3.5c, and a few features are worth noting. The reset feedthrough is often very large and must be removed by appropriate sampling, and the feedthrough due to the output gate can be minimized by not clocking it and simply leaving it at an intermediate potential which allows charge transfer across the final stage into the floating diffusion. Finally, the reference level is noisy due to the thermal noise in the reset transistor, so the video should be measured as the difference between the reference level and the video level.

Photodetection in a MOS capacitor occurs by the absorbtion of an incident photon with energy larger than the silicon bandgap within the bulk of the semiconductor creating an electron-hole pair. The majority carriers are swept away by the applied field to recombine within the bulk, while the mobile minority carriers that are generated within the depletion region or that wander into it are collected within the potential well. The electrode structure on top of the capacitor can reflect and absorb some of the incident energy, and to minimize this loss doped polysilicon electrodes should be used rather than aluminum. However, aluminum can be used as an effective light shield where photosensitivity is unwanted. The sensitivity of the detection process increases with the diffusion length of the minority carriers, and it depends on the size and depth of the potential well. The most important metric for this detection process is the quantum efficiency, which is the number of electrons collected by the potential well for each incident photon, and typically it can exceed 70%. In a CCD imaging array it is desired that photogenerated carriers are collected by the nearest potential well, in order to avoid optical cross talk. However when near infra red illumination is used, such as that provided by GaAs laser

diodes, then the absorption length is long enough that some of the incident photons penetrate deeply into the bulk of the substrate. In this case the mobile minority carriers can random walk several pixels away from the absorption site until they are collected by a potential well. This is most severe when the minority carrier lifetime is long resulting in a long diffusion length, so there is a tradeoff between sensitivity and optical crosstalk.

4.3.1 Sensor organization

Two dimensional CCD imaging arrays[37,38] are organized in order to produce a serial raster readout of the photogenerated charge packets. This is accomplished by horizontally stacking N_x 1-dimensional CCD detector arrays with N_y pixels each, and shifting all the columns simultaneously into a fast horizontal readout register at the bottom of the array. The horizontal CCD quickly shifts each video line to the output node, where each charge packet is collected, amplified and read off chip, thereby clearing out the horizontal register so another line of video can be shifted down in order to raster out the entire video array. There are two classes of 2-D CCD image sensor organization, frame transfer and interline transfer, which are schematically illustrated in Figure 4.3.6. In a frame transfer CCD photodetection occurs within the CCD transfer channel itself, while in an interline transfer imaging array photodetection takes place in an auxiliary photodiode structure, and the opaque CCD array is only used to transport the signal charge to a low noise output structure. Frame transfer arrays consist of an imaging section, which is used to integrates up an image, and a readout section that is used to store and slowly shift out the image produced by the previous field, and between fields the image subsection is rapidly shifted into the readout section. Interline transfer arrays consist of an array of photodiodes or MOS photosensors which accumulate a signal until they are gated in parallel into a raster organized CCD transfer register array, so that the

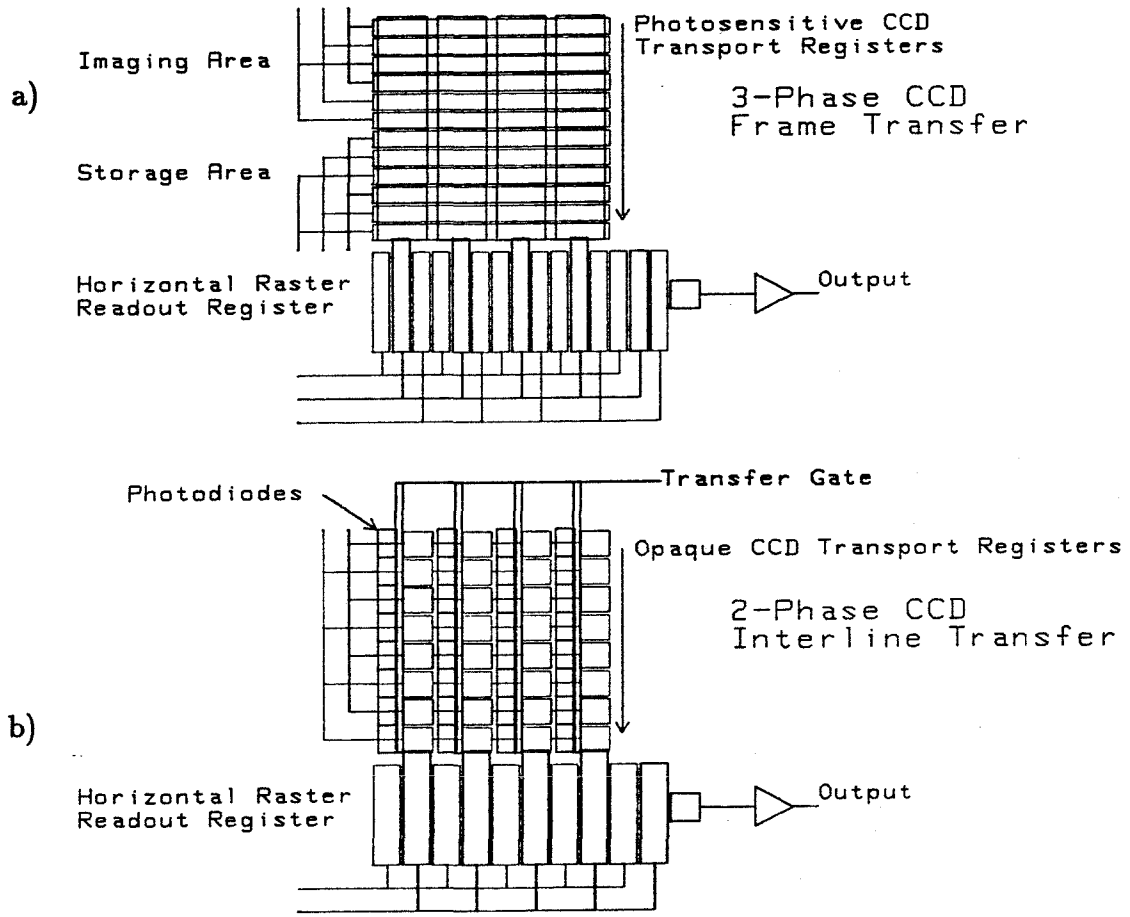


Figure 4.3.6. a) Frame transfer organization for a 3 phase CCD image sensor with separate image and storage CCDs, b) interline transfer organization of a 2 phase CCD image sensor with separated photodetectors and transfer registers.

accumulated charge packets can be read to the output node while the next image is being accumulated in the photosensors. Both types of arrays can be operated in an interlaced format where alternating fields are accumulated under different phases of the individual pixels, so that the achieved resolution is doubled vertically. An essential difference of these interleaving techniques for signal processing applications is that the frame transfer CCD photodetectors integrate for a single field time, typically 1/60 of a second, while the interline transfer photodiodes integrate for offset and half overlapping full frame times, typically 1/30 of a second. The vertical columns need to be isolated from each other and this is usually accomplished by including an implanted diffusion between columns that is biased in order to drain away any excess charge that overflows from one column, and to keep deep carriers from diffusing horizontally.

One final variant of a two dimensional imaging array that was very important for the signal processing work investigated in this thesis is the time delay and integrate (TDI) scrolling array. This mode of operation was intended for use when a moving image was to be detected by a stationary CCD, such as in satellite surveillance applications. In this case the CCD can be electronically scanned to remain in synchronism with the moving image so that the photogenerated signal builds up with the number of TDI stages, while the dark current and trapping noise add incoherently, giving an improvement of the SNR that goes as the square root of the number of TDI stages. In this case the number of TDI stages is given by the number of pixels times the number of clock phases per pixel, so that it is advantageous to have many phases per pixel[39], as is illustrated in the six phase TDI array shown in Figure 4.3.7. In the signal processing applications examined in Chapter 5, the scrolling detector is operated with time and space varying input optical fields so that instead of incoherently improving the SNR by averaging a moving image across the array, a coherent signal correlation is performed. In this

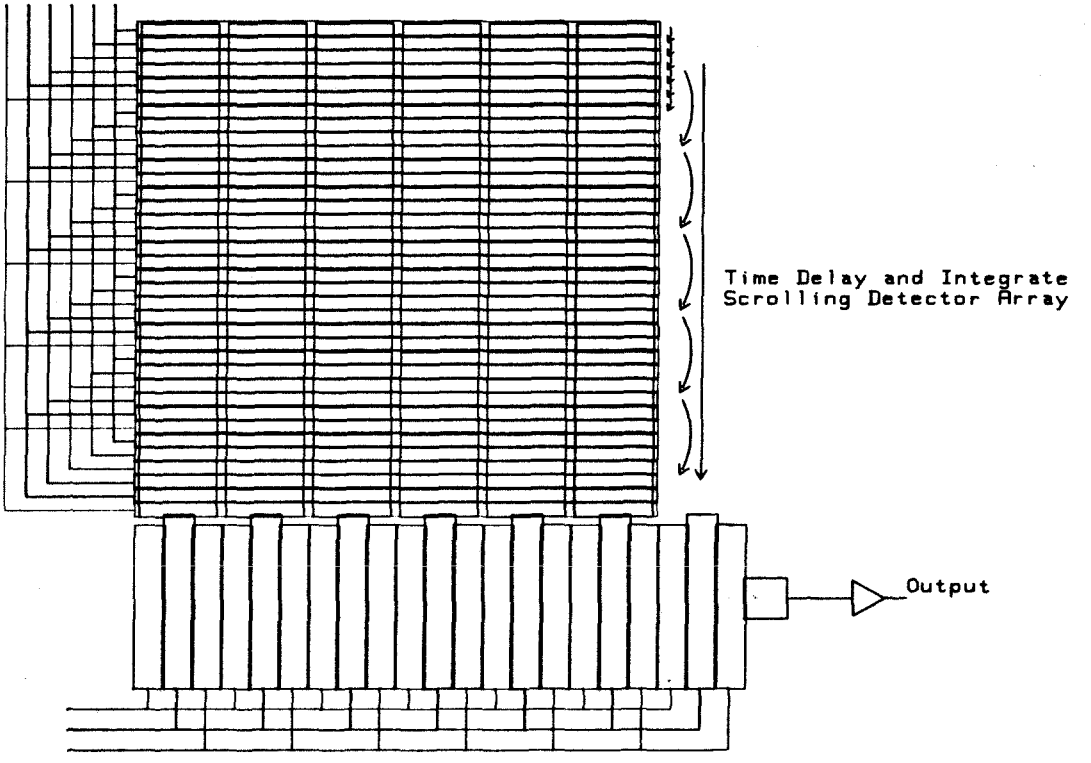


Figure 4.3.7. Six phase Time Delay and Integrate (TDI) sensor organization.

mode of operation, also called shift and add, an optical intensity profile is spatially sampled and time integrated between TDI shifts, then this charge profile is vertically shifted by one pixel and subsequent intensity profiles are integrated and added to that previously accumulated. This type of operation can be performed in a frame transfer architecture by simply hooking up the image and storage registers and clocking them together, forming a scrolling detector array coordinate frame. Integration can be performed at each clock phase as well if desired to increase the processing gain, but this can lead to an undersampling of the correlation function. The TDI mode of operation can also be performed in an interline transfer device by holding the transfer gate in the open position at all times, so that as the CCD readout register potential wells travel by the adjacent photodetectors the charge accumulated since the last shift time is poured from the photodetector into the CCD. If the CCD is shifted by full pixel steps then the two photodetectors per CCD pixel will simply add the accumulated charge, and will act as a single larger photodetector. However, if the dwell time is the same for each photodetector within a pixel, then undersampling of the integrated correlation function will result, because the origin of the integration times of the pixels associated with the two fields are offset.

4.3.2 Imaging array performance

The performance of a CCD imaging array is determined by several factors, the most important for signal processing applications are the signal to noise ratio (SNR) of the detected image, the achievable resolution as determined by the modulation transfer function (MTF), the optical sensitivity as determined by the quantum efficiency (QE), the photoresponse linearity, and the achievable speed of operation. For monochromatic illumination, such as that provided by a laser, it is only the sensitivity, resolution and linearity at the given wavelength of operation that is of concern, so wide spectral response is not important. The speed of operation should

be determined by the signal processing system that is being implemented, and the array should be clocked as slow as possible while still satisfying the constraints of the desired processing operation in order to minimize the noise bandwidth and clock feed through.

There are a large number of mechanisms that can create noise in a CCD image sensor system, and these will be briefly summarized. Since the CCD is detecting an optical field, the generation of photoelectrons will be governed by the Poisson statistics of the photon source, which is usually approximated with a standard deviation which is equal to the square root of the mean. There are several noise sources associated with the semiconductor trapping states, affecting the transfer and integration processes in CCDs, and these will present an ultimate limitation on CCD image sensors, but are not as large as other noise sources at room temperature. The thermal generation of electron hole pairs at recombination-generation centers within a diffusion length of the volume of the depleted potential well will produce a dark current that will add a bias to each signal well. This dark signal will produce a spatially fixed pattern across the image whose magnitude will grow linearly with integration time, and a statistical variation about the expected value of each pixels dark charge that is given by the square root of the mean dark charge. The expected value of the fixed pattern noise can be stored in a frame buffer and subtracted off with a field flattening technique that can significantly improve the image quality by removing the spatial variation of the dark signature, although this technique can not remove the variance of the thermal noise. At room temperature integration in CCD image sensors must be limited to about 1 second or less to avoid using up the available charge storage capacity with accumulated dark current, but a cooled device can be used to integrate for much longer periods. Finally there are unavoidable noise sources associated with charge detection and readout from the chip and the following signal amplification and quantization. After each charge

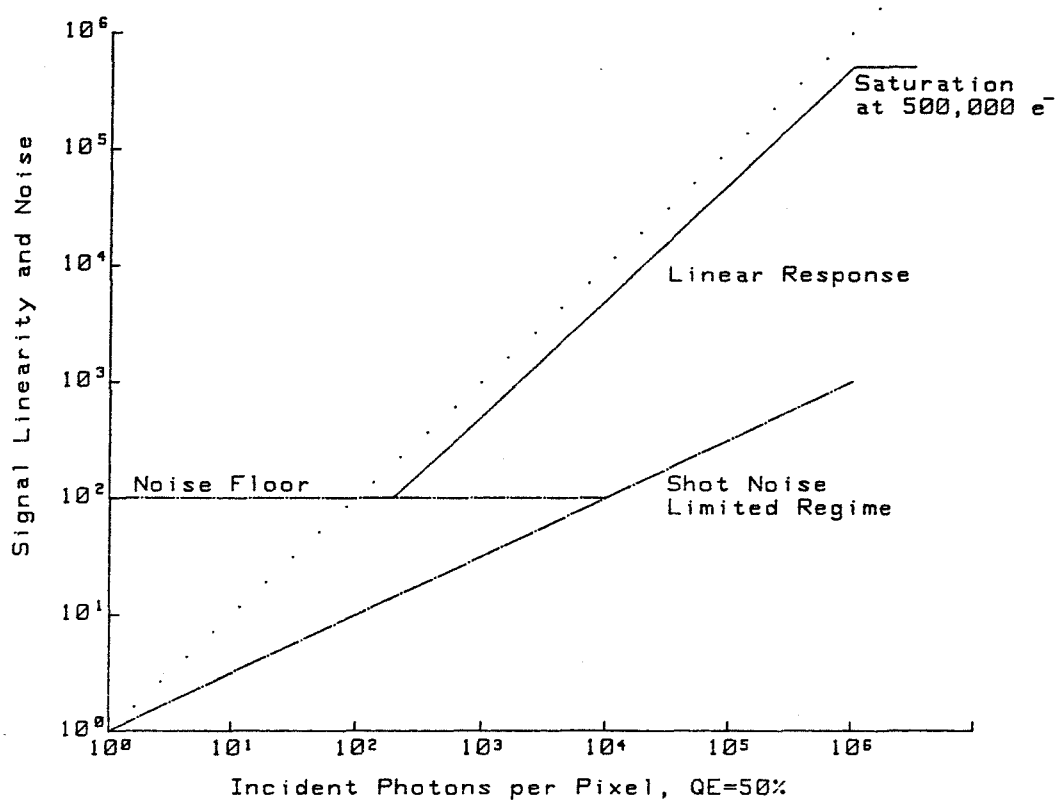


Figure 4.3.8. Plot of typical signal linearity up to saturation, and noise response of a CCD detector.

detection operation by an output floating diffusion, the floating diffusion must be reset to a reference potential by a reset switch, and this introduces a noise charge on the output node of capacitance C that is given by kTC, if the switch is left open long enough ($t > RC$). This noise charge remains on the floating diffusion when the next charge packet is clocked onto the output node and adds a noise variance to the amplified signal charge. However, this noise source can be removed with a technique known as correlated double sampling[40], by sampling the output voltage after the floating diffusion is reset to the desired bias potential, and subtracting this value from the sample obtained after the desired charge packet is transferred to the output node. In cooled, slow scan CCD cameras noise levels as low as 18 electrons per pixel have been reported[41], and noise of only 5 electrons should be achievable, yielding an SNR on the order of 10^4 : 1 or larger. However, in room temperature cameras operating at video rates 1000 electrons per pixel would be more typical, yielding an SNR of 500:1. An illustration of the signal dependent noise typically seen in a CCD detector is shown in Figure 4.3.8, along with the highly linear response of the detector up to its saturation limit. In this example a noise floor of 100 electrons limits the SNR at low light levels, but at higher signal levels the noise response is essentially shot noise limited. This is ignoring the fixed pattern noise associated with the thermal dark current and pixel to pixel sensitivity variations caused by fabrication variations which are typically the most noticeable noise sources, but these additive and multiplicative effects can be removed with field flattening techniques.

The unambiguous spatial resolution of a CCD image sensor is fundamentally limited by the Nyquist sampling rate imposed by the periodicity of the discrete pixel structure. When spatial frequencies higher than one half of the sampling detector periodicity are incident upon the CCD, then they can beat with the detector spatial frequency producing frequency components reflected about the Nyquist

frequency, resulting in low frequency beats referred to as aliasing or Moiré. The resolution capabilities of a spatially sampled detector are primarily determined by the sensitivity profile of a detector element, but other factors decrease the obtained resolution. The overall resolution degradation can be described in terms of the point spread function (PSF), which describes the resulting detected spot for an infinitesimal input optical blur spot. The Fourier transform of the PSF is called the modulation transfer function (MTF), and describes the spatial frequency dependent ratio of the modulation contrast of a detected sinusoid with the contrast produced at zero spatial frequency. The integration MTF degradation due to a rectangular pixel of width W and periodicity $1/\Delta x$ is simply given by the Fourier transform of the sensitivity profile.

$$\mathrm{MTF}_{\mathrm{integration}}(f) = \mathrm{sinc}[fW] \tag{4.3.1}$$

Thus for a flat spatial frequency response, infinitesimal width pixels would be desired, but in this case sensitivity drops to zero. In an interline transfer device the integration sensitivity profile is a rectangular aperture as large as the individual photodetectors. In the horizontal dimension the fill factor is only as large as the ratio of photodetector width to transfer channel width, which is typically equal to one half. The detector spacing for a single field in the vertical dimension of an interline transfer device is twice the detector width, but when the two fields are interleaved the resulting detector spacing is equal to its width. In a frame transfer or TDI device the sensing pixels in the vertical transfer direction have a sensitivity profile height equal to the pixel spacing, which can be approximated as a rectangular aperture or as a trapezoid with a flat top equal to the number of phase which are on during integration, with linearly decreasing sensitivity over the adjacent phases. When separate fields integrated under offset phases are interleaved to form a standard video frame, the resulting sampling pixels are twice as wide as their spacing producing a faster roll off of the high frequency response with respect

to the interleaved sampling frequency. In the horizontal dimension the sensitivity profile is as wide as the transfer channels, with trapezoidal wings as wide as the channel stopping implants between the channels. The integration MTF due to a trapezoidal pixel with pitch W, flat sensitivity width $W - w$, and wing width w is degraded from the rectangular sensitivity profile by an extra roll off term.

$$\mathrm{MTF}_{\mathrm{integration}}(f) = \mathrm{sinc}[fW]\mathrm{sinc}[f(W-w)] \tag{4.3.2}$$

The integration modulation transfer function is shown in Figure 4.3.9a for the rectangular sensitivity profiles of an interleaved interline transfer device and for the trapezoidal sensitivity profiles of an interleaved frame transfer device in both horizontal and vertical dimensions. The spatial frequency response is better in the vertical direction by the interleave factor for both frame transfer and interline transfer devices. However, the use of a noninterleaved CCD imager, sometimes referred to as a scientific imager, makes the interpretation of the MTF and of the other operating characteristics considerably simpler for signal processing applications.

The charge transfer inefficiency ϵ introduces an exponential tail point spread blurring, where the tail length is a linear function of the number of transfers undergone by an individual pixel, thereby yielding a space variant resolution with exponential tails in the horizontal and vertical dimensions. Thus for a pixel that performs its image integration at position (n,m), the resulting MTF degradation due to charge transfer losses is given by the Fourier transform of the space variant pixel point spread function.

$$\mathrm{MTF}(n,m)_{CTI}(f_x,f_y) = e^{-np_x\epsilon_x[1-\cos(2\pi f_x\Delta x)]}e^{-np_y\epsilon_y[1-\cos(2\pi f_y\Delta y)]} \tag{4.3.3}$$

Where p_y and p_x are the number of phases in the vertical columns and the fast horizontal readout register, Δy and Δx are the corresponding electrode lengths, and the charge transfer efficiencies in these two dimensions are given by ϵ_y and ϵ_x, respectively. Typically the fast horizontal readout register has a much worse

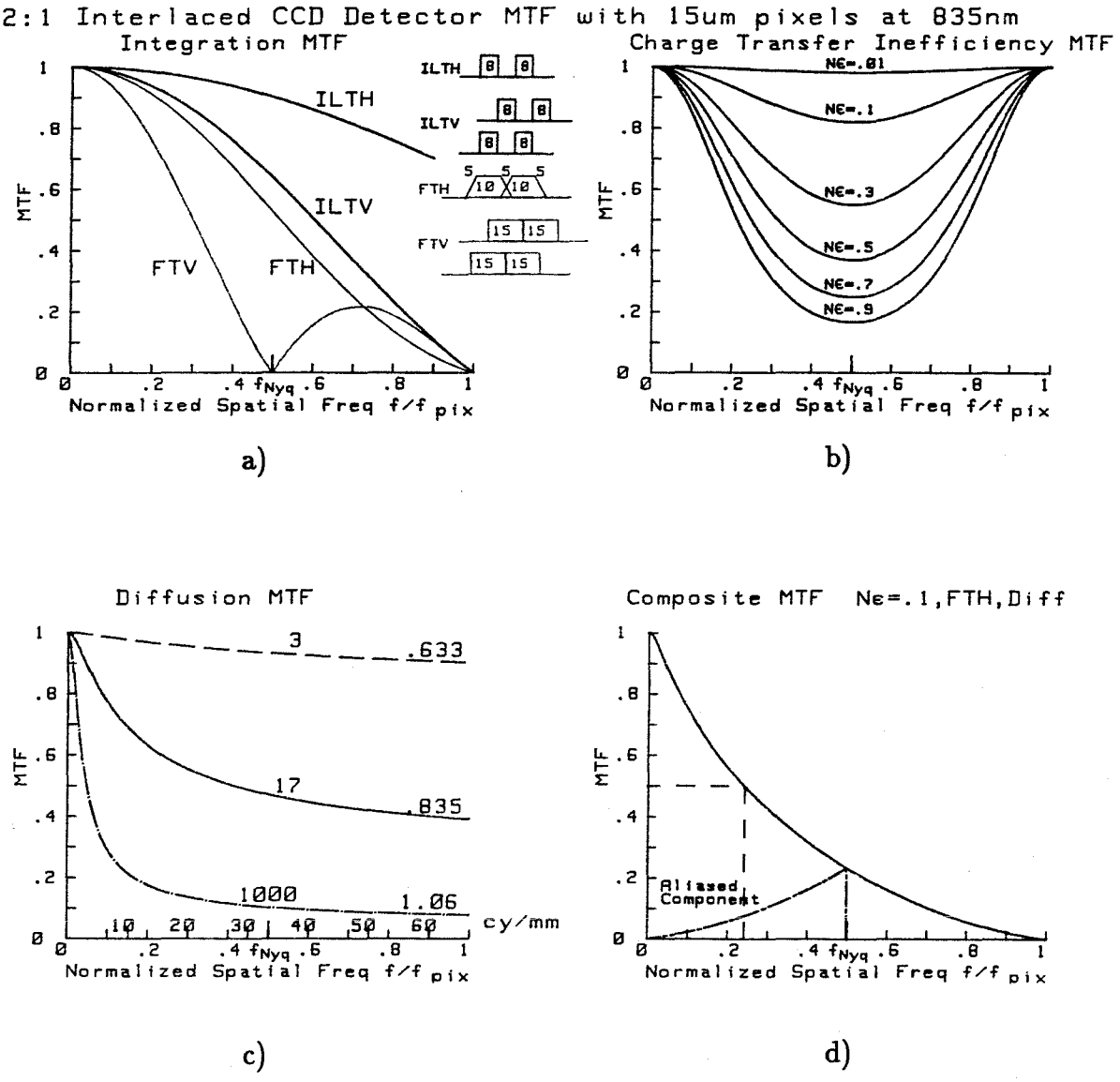


Figure 4.3.9. CCD MTF effects. a)integration MTF for rectangular interline transfer and trapezoidal frame transfer sensitivity profiles for a vertically interlaced detector, b) MTF loss due to charge transfer inefficiency, c) diffusion degradation of MTF, d) example of a composite MTF for a horizontal trapezoidal element with transfer loss $N\epsilon = 1$, and deep carrier diffusion at 835nm.

transfer efficiency than the slow vertical registers, so we can neglect the second term. The charge transfer inefficiency is illustrated in Figure 4.3.9, for various products of number of shifts N with the charge transfer inefficiency ϵ, and it is seen to be symmetric about the Nyquist rate.

The bandwidth limitation of the readout amplifier and post processing signal chain, $B(f)$, provides a degradation of the horizontal MTF without affecting the vertical spatial frequency response.

$$\mathrm{MTF}_{\mathrm{amp}}(f_x) = B(f_x t_0/\Delta x) \tag{4.3.4}$$

Where t_0 is the fast horizontal parallel to serial converter readout pixel shift time, and this expression simply scales the temporal signal chain bandwidth into the spatial dimension via the effective readout velocity.

A serious MTF degradation encountered in the experiments reported in this thesis was due to the diffusion of the carriers generated by penetrating infrared photons which were absorbed deep in the silicon bulk. For a depletion layer thickness given by L_D, a diffusion length in the silicon L_0, and an optical absorbtion coefficient at the wavelength of interest α, the resulting diffusion MTF is given by[42]

$$\mathrm{MTF}_{\mathrm{diffusion}}(f) = \frac{1 - e^{-\alpha L_D/(1+\alpha L_0/\sqrt{1+(2\pi f L_0)^2})}}{1 - e^{-\alpha L_D/(1+\alpha L_0)}}. \tag{4.3.5}$$

For the case of illumination at 835nm the absorbtion coefficient is approximately 17 μm at 300°K[43], the diffusion length is at most $L_0 \leq 100\mu$m, and the depletion width is typically $L_D \approx 5\mu$m, and this case is illustrated in Figure 4.3.9c for three common laser wavelengths. The overall MTF of the detector is given by the product of the individual spatial frequency responses, and an example of a composite MTF is shown in Figure 4.3.9d. In this example the 50% contrast is obtained at a sampled spatial frequency of approximately 4 pixels per cycle, which is much lower than that predicted from the integration MTF alone.

4.3.3 Bias removal detector structures

A major obstacle to performing useful signal processing functions with time and space integrating techniques discussed in this thesis is due to using up the available CCD dynamic range with optical bias. This problem is not as severe in the TSI approach to multidimensional signal processing as in purely time integrating approaches because the signal dependent contribution to the bias is spatially channelized by the space integrating processing, and does not flood the entire detector. However, each space integrated channel will end up having a different bias component that must be removed in order to obtain the desired bias free multidimensional signal. Typically 90% of the available charge handling capacity will be used up with the unwanted optical bias, leaving only 10% of the CCD dynamic range to represent the desired signal, and in fact this is the least linear region of the CCD dynamic range as well. It is interesting to consider performing this bias removal operation in a special CCD structure, and I will suggest a few possible approaches to performing bias removal on the CCD detector array.

The first is the simplest and most conventional, it involves interferometrically placing the desired information on a spatial carrier in either the horizontal or vertical dimension and performing a carrier demodulation on chip in order to separate the signal from the bias. The carrier demodulation operation can be accomplished by using a bipolar CCD tapped delay line based on a nondestructive sensing of the signal charge within the output channel by split electrode floating gates placed over the last few output pixels. The split ratio is modulated sinusoidally at the expected carrier frequency and the induced image charges from the two halves are separately accumulated and differentially subtracted in order to perform the bandpass filtering operation. This is called a split electrode transversal filter, and it can be implemented in either the horizontal or vertical directions of the video output, as illustrated in Figure 4.3.10 and 4.3.11, respectively. If a horizontal carrier

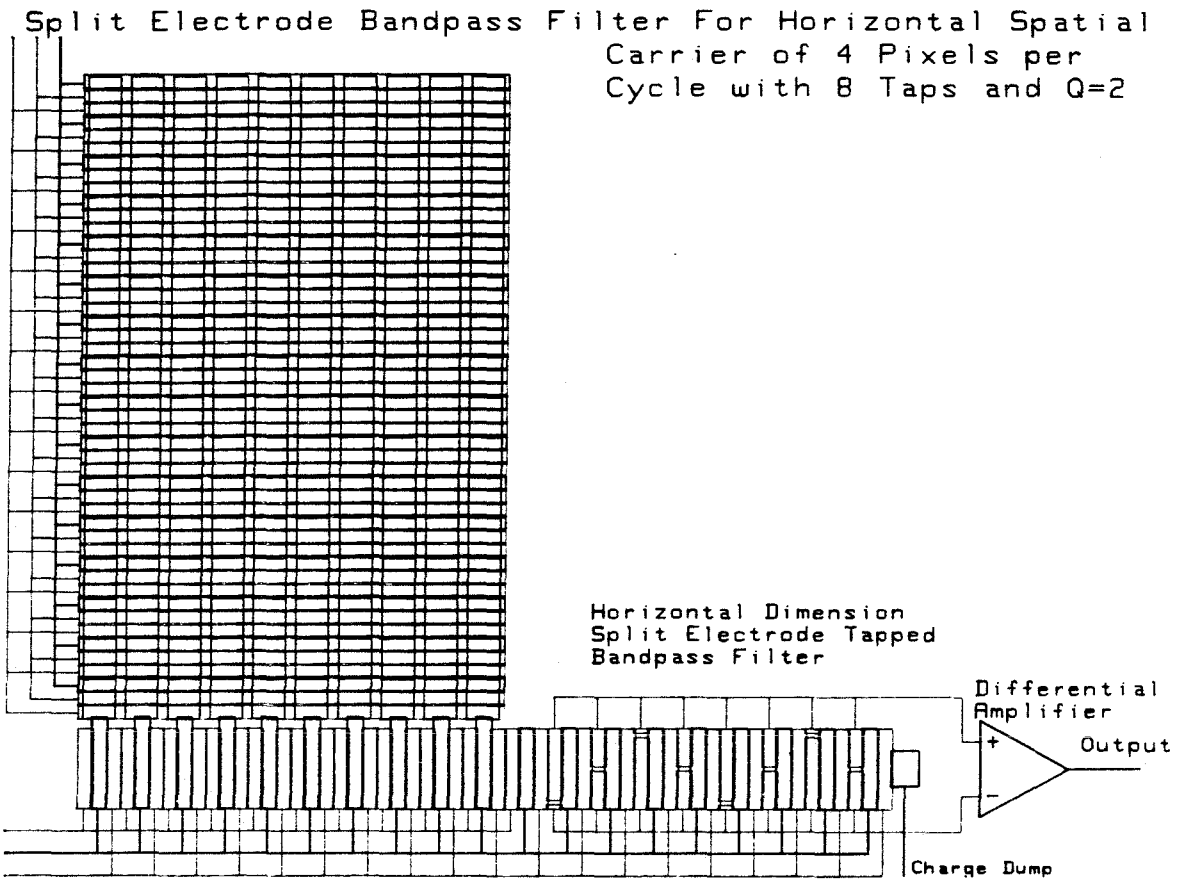


Figure 4.3.10. CCD imaging array with on chip horizontal bandpass filter implemented with a split electrode filter.

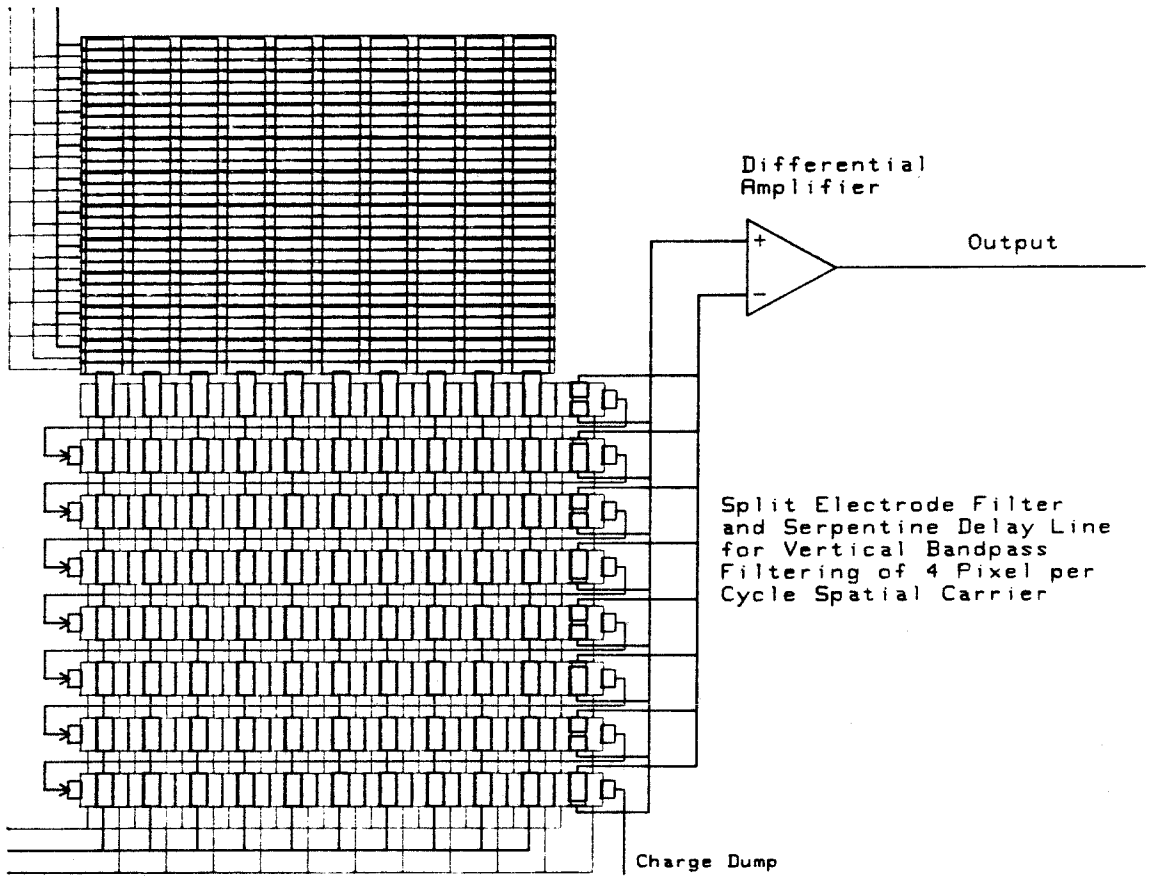


Figure 4.3.11. CCD imaging array with on chip vertical bandpass filter implemented with a serpentine delay line and split electrode filter.

is used, a split electrode transversal filter with the desired bandpass characteristics is simply put at the end of the horizontal readout register. In order to implement a vertical bandpass filtering operation it is necessary to store several horizontal video slices on chip. This is accomplished with a serpentine CCD delay line that circulates several horizontal video lines with the proper delay and formatting so that vertical slices remain synchronously stacked so they can be bandpassed in real time and appear at the chip output in a conventional video format. Programmable bandpass filtering can be achieved as well, but at a cost of increased filter complexity, which is probably unwarranted since the spatial carrier frequency can be optically adjusted to match the bandpass characteristic of a fixed frequency split electrode filter. The problem with these bandpass filtering techniques is that they do not increase the available CCD dynamic range, because the bias still uses up the charge handling capacity of the detector pixels.

A more promising approach would be to remove the bias as it is building up, leaving only the signal to occupy the available charge handling capacity of the CCD pixels. This can be accomplished with a fill and spill metering well which is used to bleed off any unwanted excess charge out of a detector pixel[44]. In order to determine how much bias needs to be skimmed off each column adaptively as the signal and bias accumulate, it is necessary to have an incoherent radiometer at the top of each column that detects the bias alone. A conceptual layout of an optically adaptive bias subtraction CCD is shown in Figure 4.3.12a, and the operation of the fill and spill metering well which is used for bias subtraction is illustrated in Figure 4.3.12b. The potential difference between the metering well and the overflow gate determines the amount of subtracted charge, and is controlled by an amplified output from the incoherent radiometer. The charge packet in the CCD channel is either accumulated in or transferred to the phase adjacent to the fill and spill bias subtraction structure. Then the channel potential is lowered, or the metering

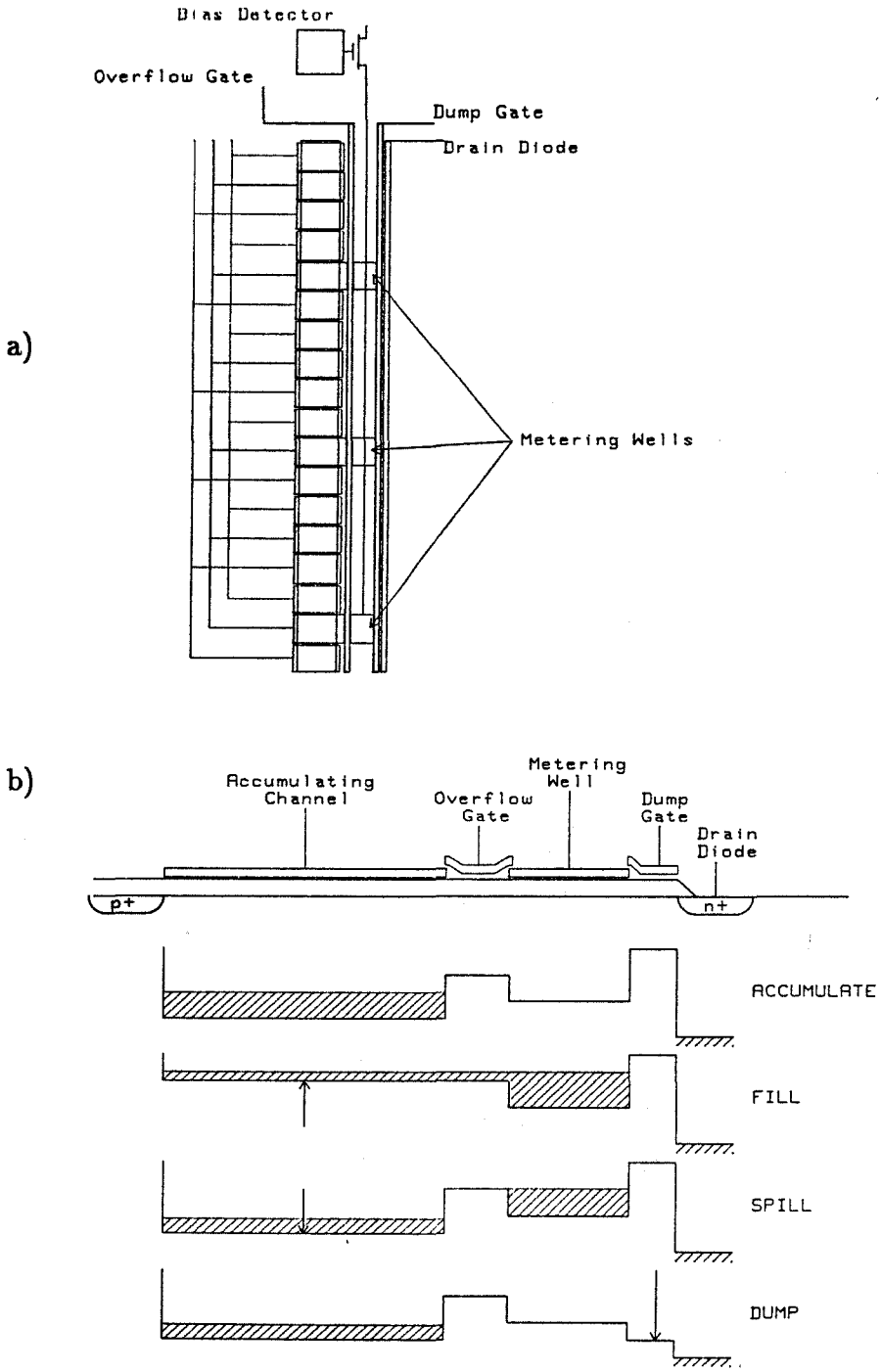


Figure 4.3.12. a) Conceptual layout of a bias subtraction CCD, b) Cross section through a fill and spill structure, which can be used to skim off a desired bias charge from the signal packet in the channel.

well and overflow gate are raised in tandem, so that the charge distributes between the metering well and the channel storage well, thereby filling the metering well. The voltage on the channel is then raised back to the starting potential, or the metering well and overflow gate are lowered to their initial potential, so that all the excess charge spills back into the channel, while a precisely measured charge packet remains in the metering well. The size of the skimmed charge packet in electrons is approximately given by the potential difference between the metering well and overflow gate times the metering well capacitance. The skimmed charge packet is dumped into the drain diode by raising the potential of the dump gate sufficiently to remove the potential barrier between the metering well and the higher potential drain diode. Thus a precisely measured bias charge has been skimmed off of the charge packet accumulated in the CCD channel leaving only the signal charge, and enough bias to represent a bipolar signal. After integrating up more signal and bias the fill and spill bias subtraction operation can be repeated as necessary.

Each time this bias subtraction operation is performed a thermal noise charge with mean square deviation of $\sigma^2_{Thermal} = kTC$, and a shot noise contribution with mean square deviation equal to the subtracted charge $\sigma^2_{shot} = CV = Q$ are included in the skimmed charge. Thus the total noise root mean square deviation after N bias subtraction operations will build up as $\sigma = \sqrt{N}\sqrt{kTC + CV}$, so the number of bias subtraction operations should be minimized. This dictates the appropriate strategy for optimizing the performance of a bias subtraction CCD in a low signal to bias ratio (SBR) environment, and this process is illustrated in Figure 4.3.13. The bias component should be allowed to build up to 3/4 of the CCD saturation charge, at which point a bias equal to half the CCD linear charge capacity should be subtracted, leaving a signal plus bias charge of approximately 1/4 the CCD linear dynamic range. The charge packet is then allowed to accumulate an additional signal plus bias, and when the bias component is equal to half the CCD dynamic

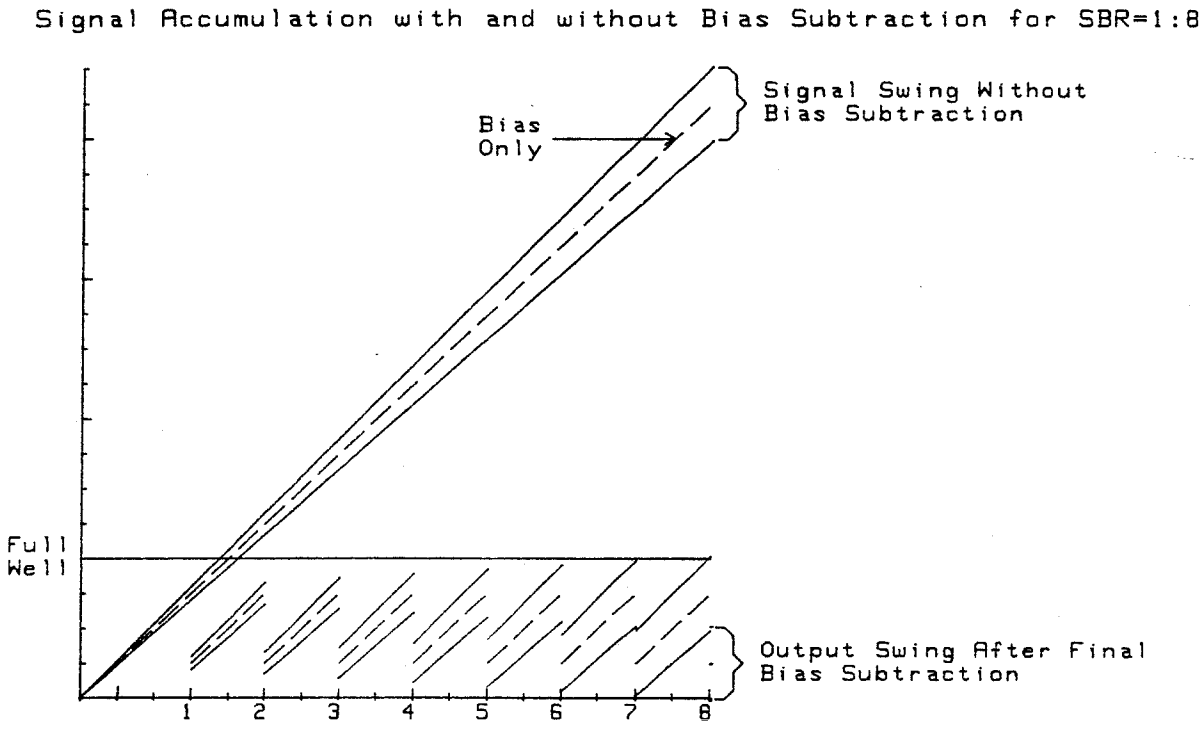


Figure 4.3.13. Signal accumulation as a function of time in a low signal to bias ratio environment with and without a repetitive bias subtraction operation.

range it is subtracted, and this process continues for the full integration period. After the final bias subtraction, a bias charge of 1/4 of the CCD dynamic range remains in order to represent the bipolar signal swing of 1/2 the CCD dynamic range, for a resulting contrast ratio of 1:1 when the signal is read off chip. This implies that the number of bias subtraction operations that are utilized during the time integration should be made equal to one over the SBR detected by the CCD in order to optimize the signal to noise ratio and contrast of the final output. The degradation of the SNR with multiple bias subtraction operations is minimized by using a small metering well with a large applied voltage, and a resulting capacity equal to half the channel capacity can be easily achieved for a metering well of half the channel area. In this case the shot noise contribution will dominate the thermal switching noise introduced by the bias skimming operation, and the resulting SNR after N=1/SBR bias subtraction operations will be given by the CCD dynamic range times $1/\sqrt{2N} = \sqrt{SBR/2}$. For small signal to bias ratios this is a significant improvement over the SNR obtained with a conventional CCD operating without bias subtraction, which is approximately given by the CCD dynamic range times the SBR. Thus for SBRs below .1 it is advantageous to use a fill and spill bias subtraction structure to repeatedly skim off excess bias charge, but for SBRs larger than this value it is only a marginal improvement in system dynamic range. An intriguing result reported by Kosonocky and Sauer[44] is the saturation of the shot noise introduced by repetitive bias subtraction operations in a recirculating delay line. If this effect can be reproduced in a bias subtracting optical detector operated in a low SBR environment, then significantly improved SNR performance would result for the signal processing systems considered in this thesis.

Since the SBR will vary from column to column of the detector array, each column will need its own bias detector and a compromise will be required as to how often to subtract the accumulating bias. In an imaging array every pixel needs

to have its own bias subtraction circuitry, which in an interline transfer device will require a bias subtraction node per photodetector, while in a frame transfer device only one metering well structure is required per integrating potential well. In a TDI array the bias subtraction nodes can be sparsely positioned along each TDI column, since each travelling potential well will pass each subtraction node as it traverses the column. In the coherent TDI processors considered in Chapter 5, the multitarget SBR decreases with the square root of the number of equal power targets within a single column, $1/\sqrt{m}$, and in a dense signal environment the maximum number of targets is given by the number of pixels in a column divided by the width of a resolution spot. In this type of coherent TDI processing a bias subtraction node should be placed as often as the square root of the total number of signal components within a TDI column. Unavoidable degradation of the optical coherence, and thermal dark current generation will decrease the signal to bias ratio even further, while many columns will have fewer targets than the maximum, so spacing the bias subtraction stages by the square root of the number of pixels within a column can be considered a compromise. This requires operating each bias subtraction node upon each TDI shift, alternatively the bias subtraction circuitry can be spaced more densely, and operated less frequently as required by the SBR of the signal environment. The extra circuitry between columns required by the fill and spill bias subtraction structure increases the horizontal spacing between photodetectors, thereby decreasing the fill factor, the achievable resolution and the optical sensitivity. This can be minimized to about 50% or less by folding the bias subtraction circuitry to be parallel with the CCD columns, rather than perpendicular. At an increase of clocking waveform complexity, the metering well and overflow gate can be folded into the transfer channel of a 4 or more phase CCD, further decreasing the spacing between detector columns.

References

Laser Diodes

[1]. A. Yariv, *Optical Electronics, 3rd Ed.*, Chap. 15, Holt, Rinehart and Winston, New York, 1985.

[2]. H.C. Casey and M.B. Panish, *Heterostructure Lasers*, Academic Press, New York, 1978.

[3]. Hitachi Laser Diode applications manual.

[4]. M. Haney and D. Psaltis, Measurement of the temporal coherence properties of pulsed single-mode laser diodes, Applied Optics, vol. 24(13), p. 1926 (1985).

[5]. M. Haney, Acousto-optical time-and-space integrating processors for real time synthetic aperture radar imaging, Caltech PhD Thesis, (1986).

[6]. Ortel Corporation preliminary data sheet, SL-1010 single mode laser diode, 12mW output power CW, 840nm wavelength, 10GHz bandwidth, 20mA threshold, 50mA operating current, Alhambra Ca.

[7]. Spectra-Diode Labs preliminary data sheet, Model SDL-4450 multi mode laser diode, 500mW output power CW (2.6 watts CW reported), 780-850nm wavelength, 2nm spectral width, 500psec rise time, .9A threshold, 1.6A operating current, San Jose Ca.

Acousto-Optics

[8]. I. C. Chang, Acousto-optic interactions - a review: I. Acousto-optic devices and applications, IEEE trans. Sonics and Ultrasonics, vol. SU-23(1), p. 2 (1976).

[9]. B. A. Auld, *Acoustic Fields and Waves in Solids*, John Wiley & Sons, New York, 1973.

[10]. J. Sapriel, *Acousto-optics*, John Wiley & Sons, 1979.

[11]. N. Uchida and Y. Ohmachi, Elastic and photoelastic properties of TeO_2 single crystals, Journal of Applied Phys., vol. 40(12), p. 4692 (1969).

[12]. A. Yariv and P. Yeh, *Optical Waves in Crystals*, John Wiley & Sons, New York, 1984.

[13]. A. Korpel, Acousto-optics, in Applied Solid State Science, vol. 3, R. Wolfe, Ed, Academic Press, New York, ch. 2, p. 73 (1972).

[14]. R. W. Dixon, Acoustic diffraction of light in anisotropic media, IEEE Journal of Quantum Electronics, vol. QE-3(2), p. 85 (1967).

[15]. D. L. Hecht, Three dimensional acoustooptic dispersion effects in acoustooptic devices for optical information processing, IEEE Ultrasonics symp., p. 463. (1983)

[16]. D. L. Hecht, Spectrum analysis using acousto-optic devices, Opt. Eng., vol. 16(5), p. 461 (1977).

[17]. S. Wofford, G. Petrie and D. L. Hecht, Polarization effects in shear wave telluriom dioxide acousto-optic devices, Proc. SPIE 202, p. 180 (1979).

[18]. D. L. Hecht, Multifrequency acousto-optic diffraction, IEEE trans. Sonics and Ultrasonics, vol. SU-24(1), p. 7 (1977).

[19]. I. C. Chang and D. L. Hecht, Characteristics of acousto-optic devices for signal processors, Opt. Eng. vol. 21(1), p. 76 (1982).

[20]. M. G. Cohen, Optical study of ultrasonic diffraction and focussing in anisotropic media, Jounal of Applied Phys., vol. 38(10), p. 3821 (1967).

[21]. B. D. Cook, E. Cavanagh and H. D. Dardy, A numerical procedure for calculating the integrated acoustooptic effect, IEEE trans. Sonics and Ultrasonics, vol. SU-27(4), p. 202 (1980).

[22]. L. Bademian, Parallel-channel acousto-optic modulation, Opt. Eng., vol. 25(2), p. 303 (1986).

[23]. A. VanderLugt, Bragg cell diffraction patterns, Applied Optics, vol. 21(6), p. 1092 (1982).

[24]. A. W. Warner, D.L. White and W. A. Bonner, Acousto-optic light deflectors

using optical activity in paratellurite, Journal of Applied Phys., vol. 43(11), p. 4489 (1972).

[25]. A Fukumoto et al., Polarization considerations in the operation of a two dimensional TeO_2 abnormal Bragg deflector, Appl. Opt., vol. 14(4), p. 812 (1975)

[26]. N. Uchida, Optical properties of single-crystal paratellurite (TeO_2), Phys. Rev. B, vol. 4(10), p. 3736 (1971).

[27]. E. I. Gordon, A review of acoustooptic deflection and modulation devices, Proc. IEEE, vol. 54(10), p. 1391 (1966).

[28]. D. Psaltis, private communication.

[29]. M. Haney, private communication.

[30]. R. T. Weverka, private communication.

[31]. D.L. Hecht and G.W. Petrie, Acousto-optic diffraction from acoustic anisotropic shear modes in GaP, IEEE Ultrasonics symp., (IEEE cat. 80CH1602-2), p. 474, (1980).

[32]. Crystal Technolgy Data Sheets, Model 4075-4, TeO_2 slow shear AOD, B=50 MHz, T_a=100μsec, T_aB = 5000 attenuation limited, 200%/RF watt efficiency, Palo Alto Ca.

[33]. Brimrose corp. product guide, RAST-1 GaP longitudinal AOD, B=1 GHz, T_a=.6μsec, T_aB = 600, 8% /RF watt efficiency; JR-100 $LiNbO_3$ longitudinal AOD, B=.7 GHz, T_a=4μsec, T_aB = 2800, 2% /RF watt efficiency, Baltimore Md.

Charge Coupled Devices

[34]. C. H. Sèquin and M. F. Tompsett, *Charge Transfer Devices*, suppl. 8 Advances in Electronics and Electron Physics, Academic Press, New York, (1975).

[35]. R. Melen and D. Buss, Eds., *Charge Coupled Devices: Technology and Applications*, IEEE Press Reprint Series, New York, (1977).

[36]. R. H. Walden et al., The Buried Channel Charge Coupled Device, Bell Syst. Tech. Jour. 51, p. 1635 (1972).

[37]. D. F. Barbe, Imaging Devices Using the Charge Coupled Concept, IEEE Proc. vol. 63(1), p. 38 (1975).

[38]. J.A. Hall, Arrays and charge coupled devices, chap. 8 in *Applied Optics and Optical Engineering, vol. VIII*, R. Shannon and J. Wyant Eds., Academic Press, New York (1980).

[39]. W. C. Bradley and A. A. Ibrahim, a 10,240 pixel focal plane with five butted 2096 element TDI CCDs, SPIE vol. 175, p. 72 (1979).

[40]. M. H. White et al., Characterization of Surface Channel CCD Image Arrays at Low Light Levels, IEEE SC-9, p. 1 (1974).

[41]. M. M. Blouke et al., 800x800 Charge Coupled Device Image Sensor, Opt. Eng. vol. 22(5), p. 607 (1983).

[42]. D. H. Seib, Carrier Diffusion Degradation of Modulation Transfer Function in Charge Coupled Imagers, IEEE trans. ED-21(3), p. 210 (1974).

[43]. R. H. Dyck, Design, Fabrication and Performance of CCD Imagers, in *VLSI Electronics Microstructure Science, vol. 3*, N. G. Einspruch Ed., Academic Press, New York (1982).

[44]. W. F. Kosonocky and D. J. Sauer, Low-Loss Charge-Coupled Device, RCA Review, vol. 40, p. 241 (1979).

[45]. Tektronix Charge Coupled Device Preliminary Specifications, TK2048M, 2048 $\times$ 2048 pixels, 27μm$\times$27μm pixel size, 6μm channel stop, ¿80dB dynamic range (cooled, slow scan), 10MHz output bandwidth (C amplifier), 7×10^5 e^- full well capacity, .99999 charge transfer efficiency, 10nA/cm^2 dark current, Tektronix Inc. , Beaverton Ore.

CHAPTER 5

MULTIDIMENSIONAL SPECTRAL ANALYSIS

The ability of a lens to produce the two dimensional Fourier transformation of a coherent wavefront at its back focal plane was recognized as one of the most promising capabilities of optical processing systems long ago. Real time operation requires the use of a rapidly programmable two dimensional spatial light modulator (SLM) as the input transducer, and although they have been under development for many years, there still is no real time SLM with resolution comparable to film. The true potential of a 2-D optical Fourier analysis system can not be realized until a high resolution, wide dynamic range and rapidly programmable SLM becomes available. In this chapter I will explain how 1-D acousto-optic devices can be utilized to perform 2-D Fourier transforms using the technique of time and space integration, without the need for 2-D SLMs. These systems were designed to perform a high resolution Fourier transform of a very long high time bandwidth input signal in a folded output coordinate representation. With small modifications these systems could be used to perform image Fourier transformations as discussed in section 3.1.2. Although slightly different design considerations apply in the image Fourier transform case, the general principles of time and space integration remain the same. However, the specific modifications necessary for image Fourier analysis will not be further elucidated in this chapter.

The concept of the space integrating (SI) folded spectrum[1-5] was introduced by Thomas[1] as an approach to wideband high resolution optical spectral analysis. He showed that the 2-D optical Fourier transform of a raster scanned signal was an orthogonally raster scanned version of the 1-D spectrum of the original high time

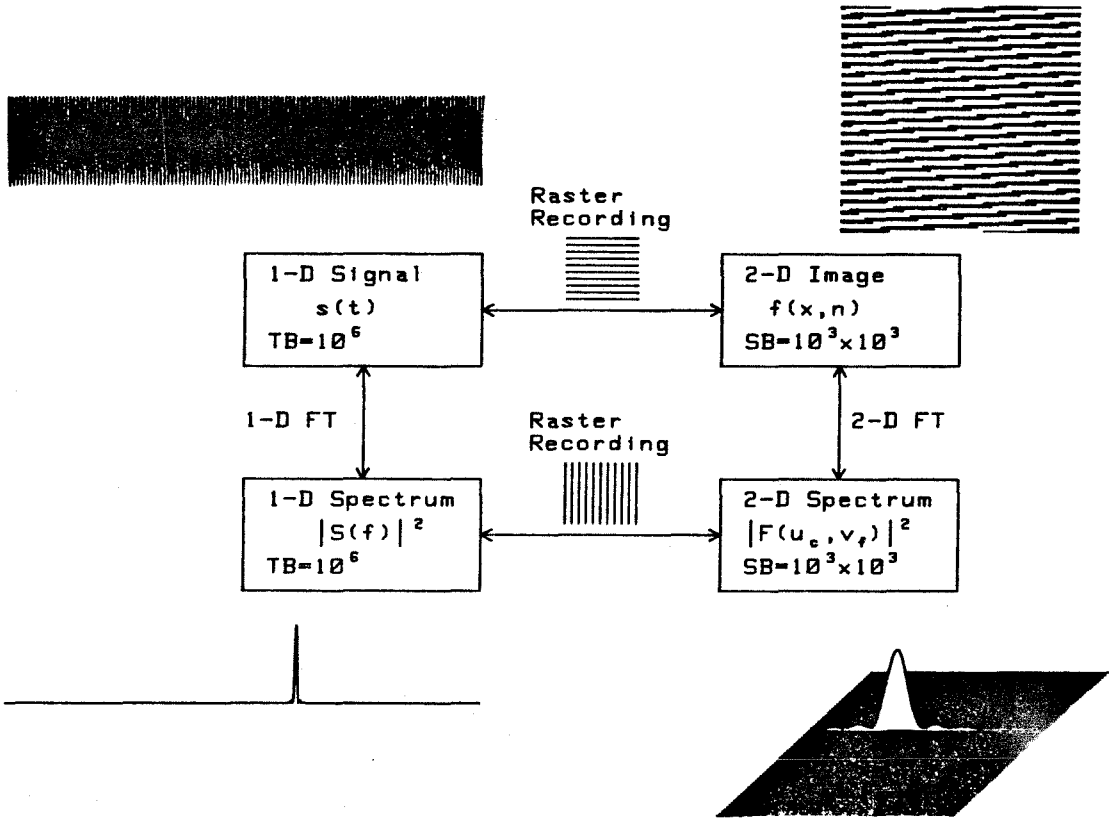


Figure 5.1. Diagram illustrating the relation between a long 1-D signal, its 2-D raster, the 2-D folded spectrum, and the high resolution 1-D Fourier transform.

bandwidth signal. This indirect approach to spectral analysis of very long signals is illustrated schematically in Figure 5.1. The folded spectrum appears ideal for an optical signal processing system since it effectively utilizes both available spatial dimensions in order to perform a high resolution spectrum analysis operation that would be impossible in a 1-D system due to the limitation in space-bandwidth products of 1-D devices. As an example, if we had a signal with time-bandwidth of 10^6, that we wished to write on a one dimensional optical modulator with 1 micron resolution, then the spatial representation of the signal would be 1 meter long, and the optics required to illuminate and Fourier transform the signal would be impractical. However, if we represent the signal in a two dimensional raster format, then even with 25 micron resolution the signal can be compactly represented on a 2-D SLM of less than two inches on a side. The folded spectrum processing operation can be performed in real time by raster scanning a very long input $s(t)$ onto a 2-D SLM, and using the 2-D space integrating Fourier transformation performed by a spherical lens[5,6]. This system is illustrated schematically in Figure 5.2. After the complete 2-D raster is formed, coherent light is passed through the SLM and the resulting amplitude distribution $f(x,y)$ is Fourier transformed by the lens. The power spectrum $|F(u_c,v_f)|^2$ is detected by a 2-D charge coupled device (CCD) detector array at the back focal plane of the lens. The orthogonally oriented raster readout mechanism of the CCD can be used to turn the 2-D power spectrum back into a high resolution 1-D power spectrum $|S(f)|^2$, at a rate of up to 60 such transformations each second, which is the current limit on both SLM and CCD frame rates.

The separability of the 2-D Fourier kernel, $e^{i2\pi(xf_x+yf_y)} = e^{i2\pi xf_x}e^{i2\pi yf_y}$, allows an alternative implementation of the folded spectrum operation as a cascaded system of multichannel 1-D transformations. This approach allows the use of acousto-optic devices to enter the signal to be transformed into the optical sys-

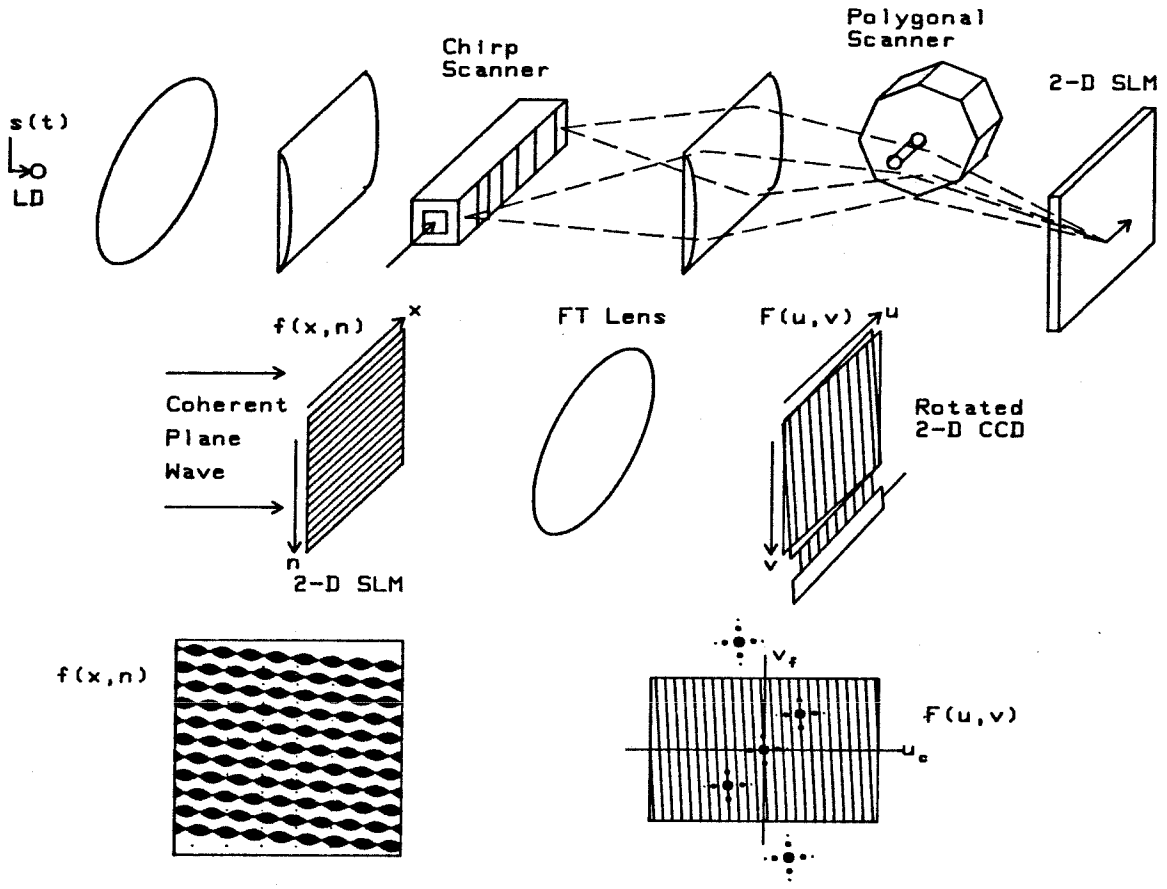


Figure 5.2. Space integrating folded spectrum processing system requiring a real time 2-D SLM.

tem, and thereby avoids the need for high quality, real time 2-D SLMs. Turpin[7] and Kellman[8,9] have suggested such an approach utilizing acousto-optic devices (AODs) in an interferometric time integrating (TI) configuration. In this approach a reference 2-D array of local oscillators, with coarse and fine frequencies in the x and y dimensions, are produced by orthogonally oriented AODs to which fast and slow chirps are applied respectively. This 2-D reference distributed local oscillator array is interfered with a plane wave which is phase modulated in time with the signal information. When a frequency component of the signal matches that of a particular local oscillator, then an interferometrically generated baseband heterodyned beat frequency will be produced, and a time integrated peak will be created at that location on the output detector array. Since this approach involves time integration of an interferometrically detected signal, a bias will be generated due to the incoherently detected local oscillator reference, and a signal dependent bias will occur throughout the detector array due to the power in the phase modulated signal beam. Each frequency component of the signal will produce an incoherently detected bias term at each detector pixel, so that in a multitone signal environment a large signal dependent bias term will be generated thereby using up the detector dynamic range that is available for the interferometric signal detection. The interferometrically detected signal can be separated from the bias terms by placing it on a spatial carrier, through a change of the incidence angle of the signal and reference beams, and spatially bandpass filtering the detector output. However the dynamic range of the demodulated amplitude spectrum will be limited by the signal to bias ratio on the detector, which can fall as one over the square root of the number of equal power tones within the analysis bandwidth.

A hybrid approach to folded spectrum processing that utilizes a 1-D space integrating Fourier transform performed by an AO spectrum analyzer in one dimension, followed by a multichannel interferometric time integrating fine resolution spectral

analysis in the orthogonal direction can combine the best features of 2-D SI and 2-D TI approaches to folded spectrum processing. Since an acousto-optic device is used to enter the data into the system, a 2-D real time SLM is not required. A standard space integrating acousto-optic spectrum analyzer (Section 2.1) is used to perform the coarse spectral analysis in one dimension without introducing a signal dependent bias at all other coarse frequency bins. A time integrating approach to fine frequency spectral analysis must be performed in parallel on each coarse spectral bin, using either the chirp transform algorithm (Section 2.4) or the discrete Fourier transform (DFT). In a dense multisignal environment this type of hybrid system will have a less severe signal dependent bias buildup problem than the 2-D TI approach, because only signals within each coarse frequency channel will contribute any bias to that channel.

The space integrating AO spectrum analyzer has on the order of TB=1000 spatially resolvable coarse frequency bins at the 1-D output Fourier plane. The doppler frequency shift associated with acousto-optic Bragg diffraction produces a linear spatial dependence of the temporal frequency of the output spectrum, meaning that each spatially channelized Fourier component has a phase history of oscillation given by the frequency causing the diffraction. Time integrating processing on a CCD detector array is a baseband operation resulting in the accumulation of DC difference frequency components, so that the coarse frequency components of the SI spectrum analyzer must be removed within each coarse frequency bin, in order to facilitate further fine frequency analysis. There are two generic techniques for accomplishing this conversion to baseband, interference with a 1-D distributed local oscillator (DLO) reference wave, or illumination with a repetitively pulsed light source. Bader[10–12] proposed a hybrid time and space integrating (TSI) system utilizing a DLO reference wave that was produced by applying a repetitive wideband chirp to an auxiliary collinear AOD, and interfering the Fourier spectra of the signal and reference. An orthogonal AOD and a modulator were utilized to im-

plement the time integrating chirp transform algorithm required for fine frequency analysis of each coarse frequency bin in his original system. In section 5.1 I present a modified version of this type of DLO interferometer that utilizes a time delay and integrate (TDI) CCD in order to perform an interferometric chirp transform fine frequency analysis, thereby eliminating an AOD from the system, and improving the immunity to coherent artifacts. Psaltis[13–16] proposed an alternative technique for removing the coarse part of the temporal frequency of oscillation present at the Fourier plane of an interferometric AO spectrum analyzer which utilizes a repetitively pulsed light source, and this was first demonstrated in the work reported in this thesis[17]. This narrow pulse train massively under samples the interferometrically produced temporal oscillations, thereby aliasing the signal to baseband and removing the coarse frequency part of the sampled temporal oscillations present within each coarse frequency bin. The use of a pulsed light source increases the flexibility in the design of the time integrating fine frequency processing, and allows the implementation of the space variant DFT operation, which is presented in section 5.2. A slightly simpler system, presented in section 5.3, uses the chirp transform algorithm on a TDI CCD to perform the fine frequency analysis of each coarse frequency bin.

In this chapter I examine in depth the variations and tradeoffs between these TSI folded spectrum processing schemes, and I present some experimental results of pulsed source folded spectrum processing. Three different pulsed source TSI systems were experimentally demonstrated with slight readjustments of the same basic crossed Bragg cell interferometer. These were the time integrating DFT fine frequency processor, the TDI sliding DFT fine frequency processor, and the interferometric TDI chirp transform fine frequency processor. All of these techniques were able to perform the multidimensional spectral analysis operation, but there were various tradeoffs between the different approaches which will be illustrated in the following discussion.

5.1 Distributed Local Oscillator Interferometer

The folded spectrum interferometer proposed by Bader[10–12] required 3 Bragg cells and a modulator. One is used for the signal arm spectrum analyzer, one for producing the distributed local oscillator (DLO), and the third cell and the modulator were used to perform the chirp transform in the orthogonal dimension upon each resolvable frequency bin produced by the interferometric space integrating spectrum analyzer. In this section I propose and analyze a modified DLO interferometer that uses a TDI CCD to perform the multichannel chirp transform fine frequency spectral analysis, and this technique eliminates one of the Bragg cells from Baders system. The major advantage of the TDI approach to performing the chirp transform is that it averages out coherent artifacts and CCD fixed pattern variations in the dimension of the scrolling detection. The primary disadvantage of this technique is that the fine frequency analysis bandwidth is limited to the TDI transfer rate, which for commercially available CCDs is only 15.7 kHz. This is below the required rate, which is equal to the resolution of the space integrating spectrum analyzer, and is only appropriate for low bandwidth (15MHz), high resolution systems (60Hz). However, special purpose parallel output TDI CCDs[4.39] have the potential of clock rates approaching 100 kHz, allowing the implementation of wider bandwidth (100MHz), medium resolution (500Hz) systems.

The folded spectrum interferometer to be considered in this section consists of a signal arm which acts as a space integrating acousto-optic spectrum analyzer, which is interfered with a reference arm spectrum analyzer. The reference arm produces the distributed local oscillator as well as the necessary spatial and temporal modulations which are needed in order to implement the time integrating chirp transform algorithm on a multichannel TDI CCD in the orthogonal direction. There are two generic approaches to implementing the chirp transform algorithm in this architecture, an interferometric approach to chirp multiplication, or an in-

tensity modulated approach. The interferometric, or additive, approach appears to be more elegant, and avoids the sum term generation that results in undesirable Fresnel ripple sidelobes, so it will be considered in the detailed system analysis. The intensity modulated, or multiplicative, approach to TDI chirp transform fine frequency analysis will result in a more stable and compact, in line architecture, but the chirp transform analysis must be performed on a carrier in order to eliminate the unwanted conjugate sideband, so the systems analysis bandwidth is decreased.

A schematic diagram of the interferometer configuration considered in this section is shown in Figure 5.1.1. It consists of a signal arm space integrating spectrum analyzer, which is interfered with a reference arm space integrating spectrum analyzer which also includes additional temporal and spatial phase modulation. These additional phase modulation elements are used to perform the orthogonal multichannel chirp transform fine frequency processing with the technique illustrated in Figure 2.4.1. The 2-D time and space varying interference pattern is detected on a two dimensional TDI CCD array to complete the fine frequency analysis of each coarse frequency bin.

A coherent laser beam of wavelength λ is spatially filtered and collimated, in order to produce a collimated plane wave with separable beam profile $g(x,y) = q(x)p(y)$, where $q(x)$ and $p(y)$ can be taken to be Gaussian apodizations. The laser beam is incident upon a beamsplitter, and a fraction $\sqrt{\xi}$ of the incident field is split off to the signal arm, while a fraction $\sqrt{(1-\xi)}$ of the incident field continues on to the reference arm. In the signal arm the plane wave is reflected off a mirror and is made incident upon a cylindrical lens with focussing power in the y direction, and focal length F_{c1}. At the focal line of this cylindrical lens, a distance F_{c1} beyond, is placed the signal AOD. The input signal, $s(t)$, which is to be Fourier analyzed by the folded spectrum processor is applied to the transducer of this AOD. As usual the AOD is modeled as a windowed, travelling wave single side band modulator.

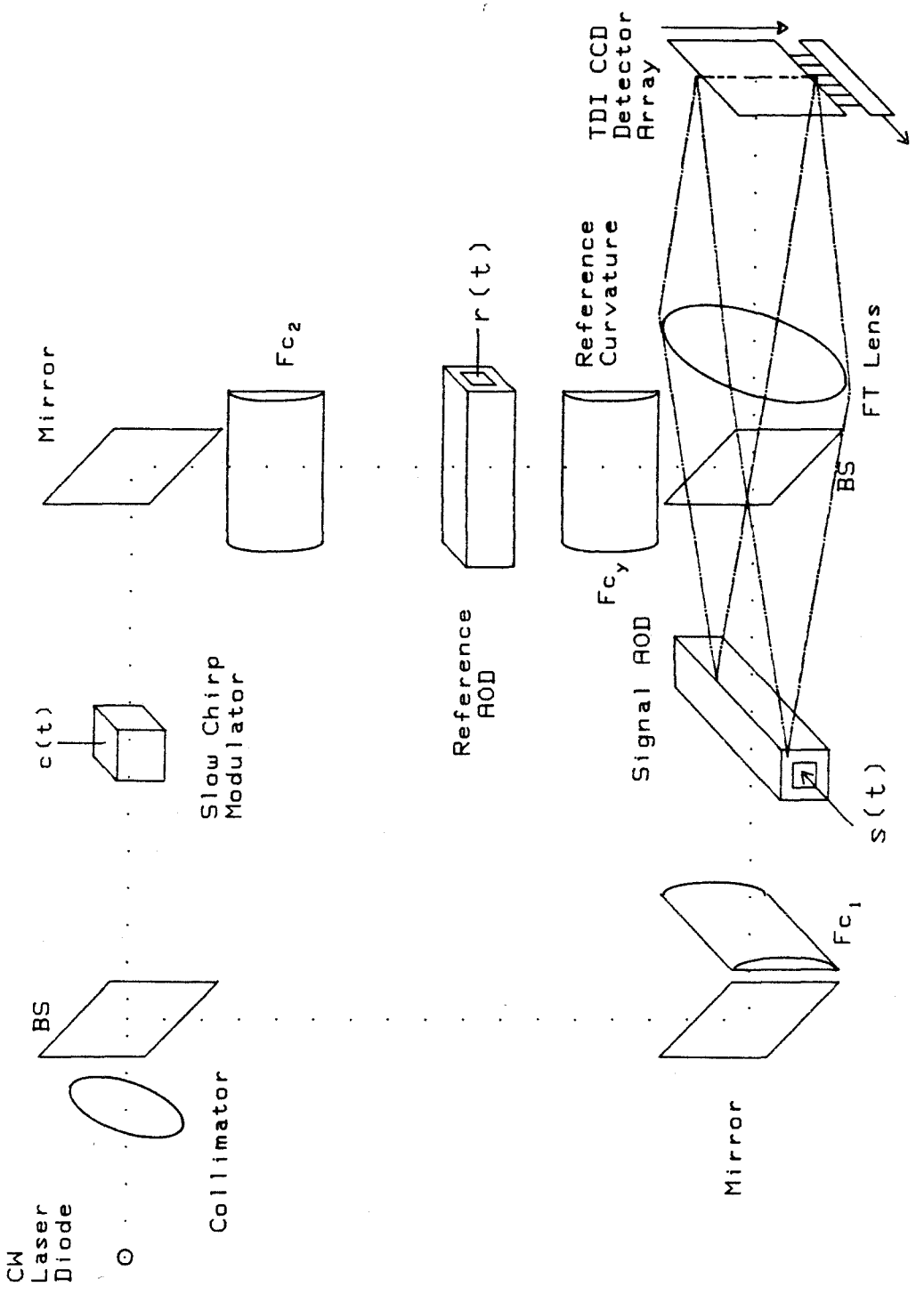


Figure 5.1.1. Schematic diagram of the DLO interferometer, which consists of a signal and reference arm spectrum analyzers, and an orthogonal multichannel TDI chirp transform CCD

The diffracted field from the AOD is Fourier transformed by a spherical lens of focal length F, placed a distance F behind the AOD, so that at the output plane a distance F behind the transform lens the observed field is recollimated and uniform in the y dimension with magnification $m_y = F/F_{c1}$. Thus from Equation 2.2.1.4 the output field of the signal arm can be represented by

$$U(\hat{x},\hat{y},t) = \sqrt{\xi}p(m_y\hat{y})d_s e^{-i2\pi\nu t}\int_{-f_u}^{-f_l}\tilde{S}(f)e^{i2\pi ft}W\left(\frac{\hat{x}}{\lambda F}+\frac{f}{v_a}\right)df \qquad (5.1.1)$$

In this expression $W(u)$ is the Fourier transform of the AOD window profile $w(x) = q(x)\text{rect}(\text{x/X})$, where X is the AOD length, and the frequency response of the AOD is omitted for simplicity, although a finite integration bandwidth from the lower frequency f_l to the upper frequency f_u is indicated. The AOD is operated in the doppler up shifting Bragg regime, so only the negative frequency analytic sideband is used in this expression. The constant d_s has absorbed all the constants and phase factors due to optical losses, diffraction efficiency and arbitrary path lengths. The output field produced by this space integrating spectrum analyzer has a wide bandwidth of spatially varying frequency of the phase oscillation that must be removed before further fine frequency analysis can be performed on the fine frequency information contained within each blur spot. The reference arm also contains a space integrating spectrum analyzer, and both arms produce an identical spatial variation of the phase oscillation frequency, so the interference of these two arms can remove these wideband oscillations, and convert the fine frequency information to baseband, as long as the appropriate repetitive signal is applied to the reference Bragg cell.

5.1.1 Distributed local oscillator reference

In this section the operation of a distributed local oscillator is analyzed, and the required reference signal is derived, for the application of converting the spatially varying oscillation frequencies produced by the signal arm spectrum analyzer

to baseband. The reference arm of the interferometer is designed to generate a distributed local oscillator (DLO), that is composed of a uniform array of evenly spaced blur spots in the Fourier plane $W(u)$, each oscillating at one distinct temporal frequency, which linearly increases as a function of space. Since an acousto-optic spectrum analyzer displays a spatially blurred representation of the ideal Fourier spectrum, with the correct temporal modulation of each spectral component, then the DLO spectrum can be produced in the Fourier plane of the reference arm by applying a broadband signal to the reference AOD that has a sampled comb spectrum and uniform power over its full bandwidth. A sampled spectrum implies that the temporal signal is repetitive, and if the spectral samples are spaced by $1/T$ then the broadband signal must be repeated with a period T. Any broadband signal of duration T that has uniform power over the desired analysis bandwidth B can be used to generate a reference DLO by repetitively applying it to the reference AOD. Different choices for the repetitive signal will result in different phase relationships of the spectral samples and it would be desirable to minimize this effect. The simplest possible broadband repetitive signal is the bandlimited impulse train, or comb function, of finite duration.

$$r(t) = d(t) * \sum_{n=0}^{N-1} \delta(t - nT) = d(t) * \left[\mathrm{comb}(t/T)\mathrm{rect}\left(\frac{t}{NT} - \frac{1}{2}\right)\right] \qquad (5.1.2)$$

Where $d(t)$ is a narrow impulse response with bandwidth B, on a temporal carrier centered at the AOD center frequency f_0. In practice a single cycle gated carrier could be used as a simple reference function, $d(t) = \mathrm{rect}(t/\tau)e^{-i2\pi f_0 t}$, where $\tau \approx 1/f_0$. When this signal is applied to an acousto-optic spectrum analyzer, each impulse will travel across the AOD aperture and act like a travelling point source of spatial extent $v_a\tau$, and separation v_aT. In the Fourier plane of an AO spectrum analyzer, one would see a plane wave due to each travelling impulse, whose angular spatial frequency increases linearly as a function of time until the impulse leaves the AOD aperture. These rocking plane waves, pivoted at the DC spot, are the

essence of a distributed local oscillator (DLO), which is designed to interferometrically remove the spatially varying doppler induced temporal frequency of the signal spectrum across the Fourier plane.

In the reference arm of the interferometer the light transmitted by the beam splitter is focussed into the reference AOD by a cylindrical lens of focal length F_{c2}. The diffracted output field is Fourier transformed by an identical lens to that in the signal arm, or the same lens can be used, with focal length F, placed a distance F beyond the AOD, so that the field is recollimated in y with magnification $m_y' = F/F_{c2}$. The field seen at the Fourier plane of the reference arm acousto-optic spectrum analyzer when the finite impulse train $r(t)$ is input would be given by

$$A(u,\hat{y},t) = \sqrt{1-\xi}\,p(m_y'\hat{y})d_r e^{-i2\pi\nu t}\Big[\tau\mathrm{sinc}(\tau(uv_a - f_0)) \qquad (5.1.3)$$
$$[T\mathrm{comb}(uv_aT) * NT\mathrm{sinc}(NTuv_a)e^{-i\pi NTuv_a}]e^{i2\pi uv_a t}\Big] * W(uv_a)$$

The first sinc function is due to the width of the reference impulses, and is approximately constant, to within 3dB, over an analysis bandwidth $B = 1/\tau$. The comb function multiplies by the DLO term $e^{i2\pi uv_a t}$, and each of the comb impulses selects at a spatial position $u = n/v_aT$ selects a linearly proportional oscillation frequency $f = n/T$, and these are spatially blurred by the convolution with $W(uv_a)$. This array of blur spots each oscillating with its own temporal frequency which is linearly proportional to its position is the essential nature of a DLO. If the impulse train were of infinite duration then this would be of the form of an ideal sampled distributed local oscillator, an array of evenly spaced blur spots or "beads", the phase of each blur spot oscillating in time with a temporal frequency proportional to its central position, on a carrier at the optical frequency ν. The finite extent of the impulse train results in a slight broadening of the frequency sampling, which in turn is due to the repetitive nature of the reference waveform, but as long as $NT \gg T_a$ then this additional frequency blurring will have almost no effect on the

size of the spatial blur spots. The problem with using the impulse train to generate a spectrally sampled, spatially distributed local oscillator in the Fourier plane is its inefficient utilization of the available light that is incident on the reference AOD, since it only utilizes a fraction τ/T of the available diffracted light.

From the form of the equation of the reference signal it is clear that any broadband signal can be used for $d(t)$, however the phase in the Fourier plane will no longer be the simple linear expression as given above[18]. It might be possible to compensate for phase variations of an arbitrary broadband signal with a holographic technique, if it was necessary. However, an additional holographic optical element will also decrease the efficiency of the utilization of the available optical power because of the limited diffraction efficiency of the hologram. The linear frequency modulated (LFM), or chirp signal, has a nearly flat power spectrum over the sweep bandwidth, and it has the distinct advantage that the resulting quadratic phase curvature can be compensated quite simply with a cylindrical lens. The repetitive wideband chirp waveform can easily be generated by applying the repetitive pulse train to a surface acoustic wave (SAW) dispersive delay line. For the case of the repetitive chirp the DLO reference function would be given by

$$r(t) = \text{rect}\left(\frac{t}{T_1}\right) e^{-i2\pi f_c t} e^{-i\pi b t^2} * \sum_{n=0}^{N-1} \delta(t - nT). \tag{5.1.4}$$

The bandwidth of each chirp segment is given by the chirp rate times the time duration, $B = bT_1$. In this expression the broadband signal which is repeated N times can be identified as a finite aperture chirp, with duration T_1, not necessarily the same as the chirp separations T, so that $d(t) = \text{rect}\left(\frac{t}{T_1}\right) e^{-i2\pi f_c t} e^{-i\pi b t^2}$. The Fourier transform of a high time bandwidth chirp, where $BT_1 = bT_1^2$ is very large, can be computed in an approximate fashion using the convolution theorem.

$$\begin{aligned} D(f) &= \int \text{rect}\left(\frac{t}{T_1}\right) e^{-i2\pi f_c t} e^{-i\pi b t^2} e^{-i2\pi f t} dt \\ &= \int T_1 \text{sinc}[T_1 f'] \sqrt{\tfrac{i}{b}} e^{i\pi (f + f_c - f')^2 / b} df' \end{aligned}$$

$$\approx\sqrt{\tfrac{i}{b}}e^{i\pi(f+f_c)^2/b}\int T_1\mathrm{sinc}[T_1 f']e^{-i2\pi(f+f_c)f'/b}df'$$
$$\approx\sqrt{\tfrac{i}{b}}e^{i\pi(f+f_c)^2/b}\mathrm{rect}\left(\frac{f+f_c}{bT_1}\right) \qquad (5.1.5)$$

In this expression the quadratic term f'^2, of the integration variable was ignored, because over the width of one lobe of the sinc function its contribution is negligible, $\pi f'^2/b = \pi/bT_1^2 = \pi/BT_1 \ll 1$. Because of the finite system bandwidth, the approximately flat bandlimited spectral envelope will have Fresnel ripples near its edges. The chirp has spectral power over a bandwidth $B = bT_1$, centered at $f = -f_c$, and with a quadratic phase curvature that is centered at the midband frequency. When the repetitive chirp signal is applied to a Bragg cell spectrum analyzer the resulting field distribution in the Fourier plane will be a spatially sampled DLO with a quadratic phase curvature.

$$A(u,\hat{y},t) = d_r\sqrt{1-\xi}p(m_y'\hat{y})e^{-i2\pi\nu t}\Big[\mathrm{rect}\left(\frac{uv_a-f_c}{B}\right)\sqrt{\tfrac{i}{b}}e^{-i\pi(uv_a-f_c)^2/b} \qquad (5.1.6)$$
$$[T\mathrm{comb}(uv_aT) * NT\mathrm{sinc}(NTuv_a)e^{-i\pi NTuv_a}]e^{i2\pi uv_at}\Big] * W(uv_a)$$

This is similar to the ideal sampled distributed local oscillator given in Equation 5.1.3, except for the quadratic spatial phase curvature term due to the focussing of the diffracted wavefront. Since the chirp duration T_1 can be made equal to or even greater than the chirp separation T (resulting in overlapping chirps), this reference waveform will have a diffracted light utilization efficiency of T_1/T, that is much greater than that of the repetitive pulse waveform. The quadratic phase curvature of the diffracted light can be approximately compensated by placing a cylindrical lens with the conjugate phase curvature in the Fourier plane. The required focal length of this lens is the distance in front of the Fourier plane at which the moving chirp focuses. The phase compensator lens will then recollimate the light diverging from the moving chirp focus, recreating a plane wave whose angular spatial frequency

increases linearly as the chirp slides through the AOD. The distance in front of the Fourier plane that the chirp comes into focus in terms of the focal length of the Fourier transform lens F, will be $f_c = b\lambda F^2/v_a^2$. If a down chirp is used, $b < 0$ then the doppler up shifting interaction will result in a diffracted field from each chirp that is an expanding spherical wavefront, emanating from a virtual source a distance $z_0 = v_a^2/b\lambda$ behind the AOD, and this virtual source will be reimaged beyond the Fourier plane, so the required chirp collimator will be a negative cylindrical lens. In order to produce a diffraction from the chirp signal that will focus before the Fourier plane, thereby allowing the utilization of a positive cylindrical chirp collimator, an up chirp can be used, $b > 0$, with a doppler up shifting interaction, or a down chirp with a doppler down shifting interaction can be used. In any of these cases, the resulting field of the distributed local oscillator at the Fourier plane just beyond the cylindrical lens will have its quadratic spatial phase curvature term almost exactly canceled, as long as the appropriate cylindrical lens is properly aligned with the center frequency diffracted component.

$$\begin{aligned} B(u,\hat{y},t) =& A(u,\hat{y},t)e^{-i\pi(x-x_0)^2/\lambda f_c} \\ =& d_r\sqrt{1-\xi}p(m_y'\hat{y})e^{-i2\pi\nu t}\sqrt{\tfrac{i}{b}}\Big[\text{rect}\left(\frac{uv_a - f_c}{B}\right) \\ & [T\text{comb}(uv_aT) * NT\text{sinc}(NTuv_a)e^{-i\pi NTuv_a}]e^{i2\pi uv_at}\Big] * W(uv_a) \end{aligned} \tag{5.1.7}$$

This distributed local oscillator reference wave is nearly identical to that given in Equation 5.1.3, except that the envelope modulation of the spectral beads is flatter over the bandwidth B for the chirped reference. Actually, it is not necessary to cancel out the quadratic phase curvature for power spectrum calculation, since it is the temporal oscillations of the DLO beads that is used for the interferometric conversion to baseband, while the phase relations of adjacent beads is irrelevant, and it is only indicated here for analytical simplicity. An amplitude spectrum with the correct phase response at all frequencies would require that the quadratic

phase curvature be cancelled out, and this could be done with the chirp collimator lens, or alternatively in the electronic post processing circuitry. The effect of the broadening of the spectral samples that is due to the finite duration of the repetitive chirp waveform can be ignored, by including this finite duration as a term in the slow modulation introduced to perform the chirp transform algorithm, or as a laser power modulation.

It is interesting to consider how the Bragg cell with temporal aperture T_a, can give rise to a spectrally sampled spatially distributed local oscillator with temporal frequency definition of the individual blur spots $1/NT \ll 1/T_a$. Clearly, during the time it takes the first chirp segment to propagate across the AOD aperture, the spectral narrowing of the frequency content of an individual spectral bead can be no smaller than the inverse acoustic aperture. But as successive chirps slide across the AOD aperture, the assumed coherence between the chirp segments induces the creation of better defined spectral harmonics by including more cycles in the temporally defining phase history of each bead. This interpretation of the generation of accurately defined temporal frequencies of oscillation of the spectral beads obeys the principle of causality, and leads to an interpretation of the definition of the spectral resolution that evolves as a function of the integration time to the maximum resolution $1/NT$, at which point integration ceases. Since this is the point at which analysis begins, the spectral harmonics giving rise to the temporal oscillations within each DLO bead can be considered to be well defined for our purposes. The continuous evolution of the well defined spectral content of the temporal phase oscillations within each spectral bead requires accurate coherence of the segments of the repetitive chirp waveform. If the successive chirp segment start times, or beginning phase, or chirp rate parameter, or chirp center frequency, deviates from the ideal theoretical value, in a statistical sense, with a variance that is relatively larger than one part in the overall number of spectral measurements to be per-

formed, then spectral broadening of the fine frequency analyzers will occur. Thus, if one million spectral measurements are to be performed by the folded spectrum processor, then the jitter time of the impulse driving the SAW dispersive delay line must be less than one millionth of the entire processing window, and the frequency stability of the chirp carrier frequency must be better than one part in a million. Additional slow temporal modulations will be added to the reference wavefront in order to complete the fine frequency analysis, and the spectral coherence of these additional low bandwidth phase modulations is required to be better than the the inverse of the overall integration time, or additional fine frequency spectral blurring will occur.

5.1.2 Spatial and temporal modulations for TDI chirp transformation

In order to implement multidimensional spectral analysis one needs to perform a fine frequency analysis of each coarse frequency bin of the signal arm after the fine frequency information has been beat to baseband through the interference with the DLO. In order to accomplish this the system needs to perform a chirp transformation on each coarse frequency bin by premultiplying each baseband signal with a slowly varying temporal chirp and then convolving with a properly matched chirp. Both of these operations can be accomplished interferometrically in the Mach-Zehnder configuration being considered in this section. The chirp premultiplication can be accomplished by phase modulating one of the arms of the interferometer with respect to the other with a quadratic phase function. This could be accomplished by placing an AO modulator in one arm of the interferometer and applying a very slow quadratic phase modulation at the device center frequency. However, the additional doppler frequency shift due to the acoustic carrier frequency in the modulated arm would have to be compensated for by an additional modulator in the other arm with an unmodulated carrier applied, or by arranging that the sum of the carrier

frequencies in one arm of the interferometer is the same as the carrier frequency in the other arm. Alternatively, a baseband phase modulator, such as an electro-optic device, could be placed in one arm of the interferometer, and a baseband quadratic phase modulation could be applied across its electrodes. An interesting variant of this would be to parabolically move one of the mirrors in the interferometer during the integration period. This could be accomplished by mounting the mirror on a piezoelectrically movable translation stage. The maximum mirror velocity would be one quarter of a wavelength per half the period of the maximum bandwidth within each coarse frequency blur spot, $v_{max} = \lambda/2T$, which is achievable with high resolution over a limited travel range with commercially available piezoelectric micrometers. Another possibility would be to slowly phase modulate the carrier used in one of the AODs with a quadratic phase shift, thereby using the AOD as a long access time modulator. The problem with this approach is that the maximum analysis frequency of the chirp algorithm is the maximum temporal frequency of the chirp, which would be the Nyquist limited frequency resulting from the acoustic transit time limitation, $1/2T_a$. But the analysis bandwidth required is given by the DLO blur spot separation of $1/T$, which can be covered with a two sided chirp modulation requiring a maximum frequency of $1/2T$. This seems to indicate that as long as $T > T_a$ that this modulation technique would be acceptable, however as we shall see in the following analysis, it is required that $T < T_a$, in order to properly separate and define the DLO spectral components spatially. Without being more specific as to the preferred implementation of the slow temporal phase modulation technique which is to be utilized, we shall consider a slow quadratic temporal phase modulation to be applied to the reference arm of the interferometer, in order to implement the chirp premultiplication on each coarse frequency bin of the interferometer.

The chirp convolution can be performed using a TDI CCD shift register array,

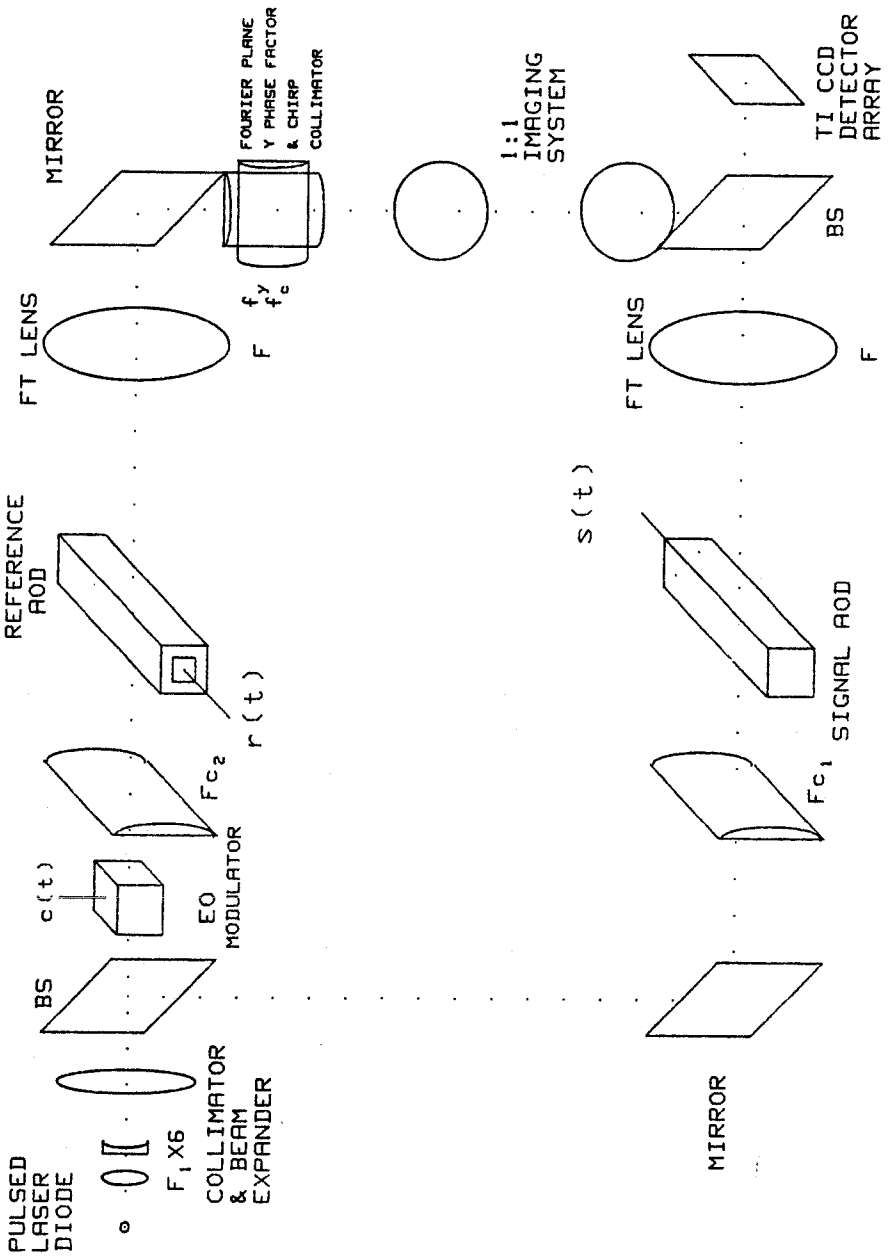


Figure 5.1.2. Optical layout of the TDI chirp transform DLO interferometer.

by multiplying the time dependent interferometric beat incident on each column of the detector array with a matched spatial chirp in the shifting direction. The spatial chirp can be produced with an interferometric technique by interfering a plane wave from one arm of the interferometer with a cylindrically focussing wavefront from the other arm of the interferometer. Each coarse spectral component is produced by the focusing of the Fourier transform lenses in the x direction, and the TDI operation will occur in the y dimension. Thus, a relative quadratic curvature must be introduced between the wavefronts in the y dimension only, and this can be simply implemented with a cylindrical lens with its focussing power in the y dimension, appropriately placed in one arm of the interferometer. A better approach to this interferometric chirp generation in the y dimension might be to place the cylinder focussing into the reference AOD at a larger distance than its focal length from the acoustic column, but with the constraint that all of the unfocussed optical energy passes through the acoustic column. In this approach a slightly detuned telescope is used so that the following Fourier transform lens does not simultaneously recollimate in the y dimension, leaving a residual quadratic curvature in y as is desired. For simplicity, I shall consider a cylindrical lens to be placed in the Fourier plane of the reference arm of the interferometer, and to have a focal length in the y direction of f_y. Thus the complete distributed local oscillator reference wave which is produced following the chirp recollimation and y dimensional focusing, with the baseband temporal quadratic phase modulation will be

$$\begin{aligned} C(u,y,t) =& B(u,t)e^{-i\pi(y-y_0)^2/\lambda f_y}e^{-i\pi\beta(t-t_0)^2}\text{rect}\left(\frac{t-NT/2}{NT}\right) \qquad (5.1.8)\\ =& d_r\sqrt{1-\xi}p(m'_y y)\sqrt{\tfrac{i}{b}}e^{-i\pi(y-y_0)^2/\lambda f_y}e^{-i\pi\beta(t-t_0)^2}e^{-i2\pi\nu t}\text{rect}\left(\frac{t-NT/2}{NT}\right)\\ & \left[\text{rect}\left(\frac{uv_a+f_c}{B}\right)T\text{comb}(uv_aT)e^{i2\pi uv_a t}\right] * W(uv_a) \end{aligned}$$

In this expression, a spatial shift y_0 of the origin of the spatial phase modulation has been included , and a matched temporal shift t_0 of the origin of the temporal phase

modulation has also been included. The spatial shift y_0 of the quadratic spatial phase factor can be introduced by shifting the lens which gives rise to this quadratic curvature term in the y dimension, or by appropriately rotating the recombining beamsplitter. The slow temporal phase modulation has a quadratic chirp rate β, that must be matched to the interferometrically produced spatial phase modulation, with a space to time conversion factor given by the shifting velocity of the TDI CCD. This entire field must be recombined collinearly and aligned spatially with the signal arm of the interferometer. Because of the proximity of the chirp collimator lens and y focussing lens to the Fourier plane, this plane will have to be 1:1 imaged to the output plane with a telescopic imaging system, in order to avoid introducing any auxiliary phase curvature.

In order to obtain proper interferometric detection, both AODs must doppler shift the light in the same sense, either up shifted or down shifted. If the two AODs are copropagating then the corresponding Fourier planes are aligned with the same spatial orientation, and will have the same scale factors as long as the two Fourier lenses have identical focal lengths. However if a two lens telescopic 1:1 imaging system is used to image the reference Fourier plane onto the CCD output plane, then the image inversion of the imaging system will spatially reverse the orientation of the reference, so that its high frequency terms align with the signal low frequency terms and vice versa. This can be corrected by using a 4 lens 1:1 double telescopic imaging system to reinvert the reference Fourier plane, or alternatively the reference AOD can be used with the acoustics propagating in the opposite direction, but still utilizing the corresponding diffracted order as the signal AOD. The complete optical layout of the Mach-Zehnder interferometer configuration is shown in Figure 5.1.2, where I have shown the implementation with counterpropagating AODs and an auxiliary electro-optic baseband phase modulator. A simpler optical architecture can be constructed by introducing the relative spatial quadratic curvature with the

alternate technique of moving the cylinder focussing into the reference AOD at an appropriate distance $F + F_{c2} + \Delta z$ behind the reference Fourier transforming sphere. The chirp collimator lens can be eliminated if phase information does not need to be extracted from the folded spectrum. Thus, no lenses need to be placed in the reference Fourier plane, so the auxiliary telescopic imaging system can be eliminated, and the reference and signal Fourier planes can be directly combined. This simplified architecture is shown in Figure 5.1.3, and the analysis proceeds almost identically for either implementation.

5.1.3 Interference of DLO reference and signal Fourier spectra

The two arms of the interferometer are collinearly combined and aligned spatially using a dielectric beam splitting prism, to produce an output amplitude given by the weighted sum of the respective Fourier plane amplitudes. The final beamsplitter reflects a portion $\sqrt{\chi}$ of the incident field, and transmits $\sqrt{1-\chi}$, so the output field is given by the appropriately weighted superposition of signal Fourier plane given by Equation 5.1.1 and the reference Fourier plane given by Equation 5.1.8.

$$
\begin{aligned}
E(u,y,t) &= \left[\sqrt{1-\chi}U(u,y,t) + \sqrt{\chi}C(u,y,t)\right] \text{rect}\left(\frac{t-NT/2}{NT}\right) \qquad (5.1.9)\\
&= \text{rect}\left(\frac{t-NT/2}{NT}\right)\left[d_s\sqrt{\xi}\sqrt{1-\chi}p(m_y y)e^{-i2\pi\nu t}\int \tilde{S}(f)e^{i2\pi ft}W\left(u+\tfrac{f}{v_a}\right)df\right.\\
&+ d_r\sqrt{1-\xi}\sqrt{\chi}p(m'_y y)\sqrt{\tfrac{i}{b}}e^{-i\pi(y-y_0)^2/\lambda f_y}e^{-i\pi\beta(t-t_0)^2}e^{-i2\pi\nu t}\\
&\left.\left[\text{rect}\left(\frac{uv_a+f_c}{B}\right)T\text{comb}(uv_aT)e^{i2\pi uv_a t}\right] * W(uv_a)\right]
\end{aligned}
$$

In this expression an infinite duration perfectly coherent repetitive chirp reference waveform has been assumed, and the finite integration time has been included as an explicit temporal gating window of the light. Thus the intensity incident on the

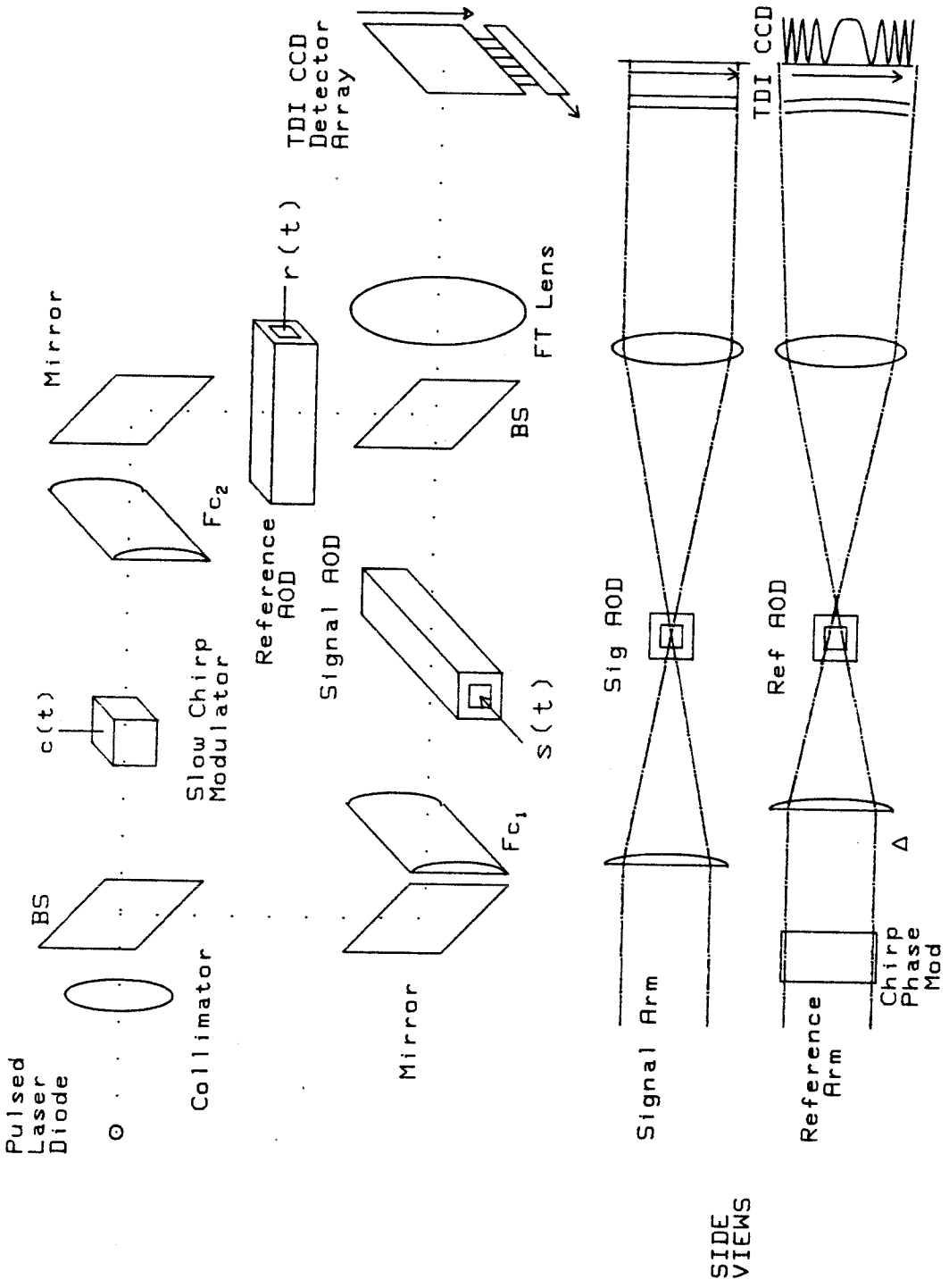


Figure 5.1.3. Simplified optical layout of the TDI chirp transform DLO interferometer.

2-D TDI CCD array at time t is given by the modulus squared of the incident field.

$$
\begin{aligned}
I(u,y,t) =& \text{rect}\left(\frac{t-NT/2}{NT}\right)\Big[(1-\chi)|U(u,y,t)|^2+\chi|C(u,y,t)|^2 \\
&+2\text{Re}\Big[\sqrt{1-\chi}U(u,y,t)[\sqrt{\chi}C(u,y,t)]^*\Big]\Big] \qquad (5.1.10)\\
=& \text{rect}\left(\frac{t-NT/2}{NT}\right)\Big[d_s^2\xi(1-\chi)p^2(m_y y)|\tilde{S}(uv_a)*W(uv_a)|^2 \\
&+d_r^2(1-\xi)\chi p^2(m'_y y)\frac{1}{b}\left|\left[\text{rect}\left(\frac{uv_a+f_c}{B}\right)T\text{comb}(uv_aT)\right]*W(uv_a)\right|^2\Big] \\
&+\eta\left|\left(d_s d_r\sqrt{\xi}\sqrt{1-\chi}\sqrt{\tfrac{i}{b}}\sqrt{1-\xi}\sqrt{\chi}\right)\right|p(m'_y y)p(m_y y)\text{rect}\left(\frac{t-NT/2}{NT}\right) \\
&\sum_{n'=\frac{f_c-B/2}{T}}^{\frac{f_c+B/2}{T}}\int|\tilde{S}(f)|W(u-f/v_a)W(u-n'/Tv_a) \\
&2T\cos\left[2\pi\left(\frac{(y-y_0)^2}{2\lambda f_y}+\frac{\beta}{2}(t-t_0)^2+\frac{n't}{T}-ft\right)+\phi\right]df
\end{aligned}
$$

The constant $\eta \leq 1$ is included to indicate a possible degradation of the interferometrically produced temporal and spatial beat terms from the theoretical optimum, due to degradations caused by the imperfect laser coherence. The last term contains all of the desired modulation, while the first two terms are merely bias terms. Temporal beat terms will be generated between the spatially displayed signal spectrum and the discrete DLO frequency components that it overlaps, and these are the inputs to the TDI chirp transform fine frequency processors. Each spectral component will overlap with a reference DLO bead, and this will interferometrically generate the resulting temporal beats that are needed for further fine frequency analysis processing, and the appropriate alignment of the signal and Fourier spectra will guarantee that these temporal beats are near baseband.

In order to understand the interferometric production of baseband heterodyne beat frequencies, we will examine the case of a single tone of frequency f' and a corresponding single sided Fourier spectra $\tilde{S}(f) = \frac{a}{2}\delta(f+f')$. This frequency can conveniently be broken up into a sum of coarse and fine frequency components,

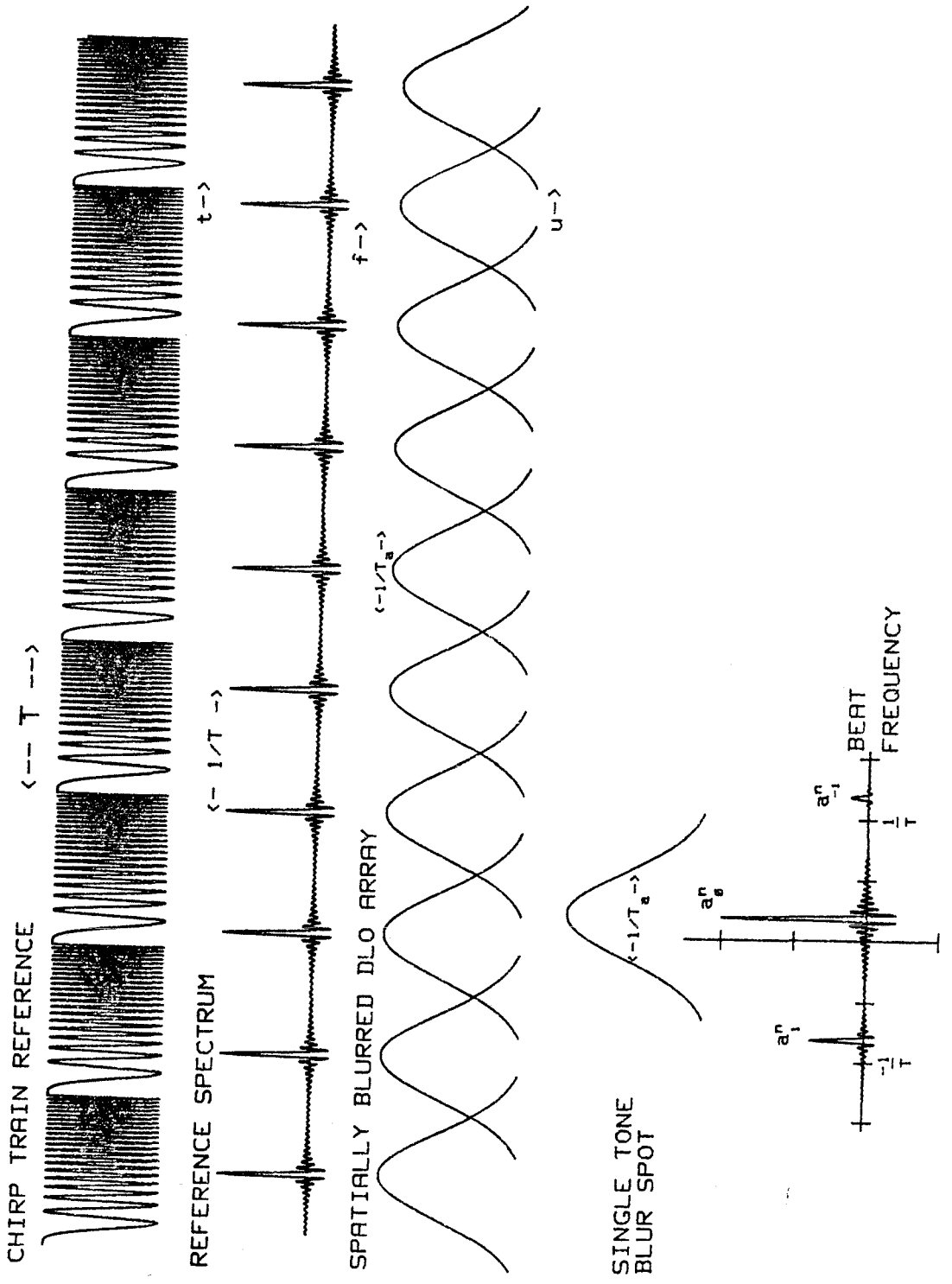


Figure 5.1.4. Repetitive chirp reference waveform, its Fourier spectra, the output of the reference Bragg cell spectrum analyzer, the single tone output of the signal spectrum analyzer and the interferometrically produced beat frequencies.

where the coarse component is a harmonic of the DLO reference repetition frequency and the fine frequency is less than this rate $f' = j/T + \delta_f$, with $j = [f'T]$ an integer, and $|\delta_f| = |f' \mathrm{mod} \frac{1}{T}| < 1/2T$. This signal will produce a single blur spot in the Fourier plane of the signal AOD, at a position $uv_a = f'$, and with an up shifted temporal oscillation frequency of f'. This is illustrated schematically in Figure 5.1.4, where it can be seen that the signal blur spot primarily overlaps with the DLO blur spot with an oscillation frequency of j/T, but the sidelobes of the signal blur spot also overlaps with the tails of the adjacent DLO beads, producing additional temporal modulation components. With proper apodization of the AODs, the sidelobes can be decreased enough so that we only need to consider the temporal beat frequencies produced by the immediately adjacent DLO beads, oscillating at frequencies $(j-1)/T$ and $(j+1)/T$. The relative strengths of the various temporal oscillation components will be proportional to the spatial overlap integrals of the field amplitudes detected within a single pixel, and the following abbreviated notation for the strength of the overlap integral detected on the nth pixel due to the lth harmonic of the reference waveform interfering with the blur spot due to the input tone at frequency f' will be used.

$$a_l^n(f') = \int h_u(u - n\Delta u) W(u - j/v_a T + \delta_f/v_a) W(u - l/v_a T) du \qquad (5.1.11)$$

where the function $h_u(u - n\Delta u)$ is the 1-D aperture profile of the pixel at position $n\Delta u$, and the overlap integral is only significantly large for $l = j-1, j, j+1$. The DLO transform in the Fourier plane should be scaled so that the reference beads are spatially separated by exactly one pixel width, and the detector array should be positioned so that a DLO bead is centered on top of each pixel, thereby maximizing the desired term a_j^n, which would occur at pixel $n = j$. This can be accomplished by relating the period of the DLO reference waveform to the pixel separation, $1/T = \Delta u v_a = \Delta x v_a / \lambda F$. Under these conditions the pixel integrated spatial overlap integral only depends on the frequency difference between the DLO

bead and the signal spectral component, and the pixel weighted overlap integral incident upon the nth pixel can be notated by $a_{l'}^{n}(\delta_f - \frac{l}{T})$ with $l' = -1, 0, 1$. If the AOD is not apodized, then $W(u) = X\text{sinc}(uX)$, which has high sidelobe levels falling only as $1/\pi l'$ for the l'th sidelobe, and this would produce significant temporal beats several lobes away from the main lobe. Apodization with a gaussian illumination profile will greatly decrease the sidelobe levels at the cost of a slight increase of main lobe width. In order to equally detect arbitrarily positioned frequencies, with respect to the discrete DLO blur spot locations it is important to space the blur spots appropriately. If they are spaced too far apart so that $1/v_aT > \text{width}\{W(uv_a)\}$, then a signal that falls between two DLO beads will produce weak temporal beats. To avoid this, the spacing of the DLO blur spots can be arranged such that the half amplitude of adjacent spatial blur spots are aligned, which will result in a minimum ripple of the reference profile, and allow nearly constant detectivity for arbitrary input frequencies. However, for typical apodizations the main lobe width to the first null will be several times wider than the half amplitude width, resulting in significant overlap of several adjacent DLO beads.

Another approach, known as single sided processing, has the advantage that the chirp transform will be on a carrier, and allows for demodulation of the signal from the bias. This can be accomplished by providing for a $u_0 = 1/2v_aT$ relative shift between the signal and reference spectra. Thus, a weak DC beat frequency will be generated when a signal blur spot falls half way between two DLO beads, and a strong beat frequency of $1/2T$ will be generated when the signal blur spot falls exactly on a DLO bead. The maximum frequency of $1/T$ will be generated when the signal blur spot falls half way in between two DLO beads on the other side of the peak of the central bead. This single sided fine frequency analysis approach has a a spectral analysis range from 0 to $1/T$, rather than from $-1/2T$ to $1/2T$. The chirping carrier resulting from the nonsymmetric chirp algorithm

requires chirp post multiply correction factors, which make this approach to carrier encoding and demodulation overly complicated. The time integrating suppression of high frequency temporal beats to be discussed, will also have more of a suppression on the temporal beats generated by the single sided processing approach. For these reasons the single sided approach will not be considered in detail any further.

The corresponding beat frequencies generated will be given by the difference of the signal frequency with that of the appropriate DLO beads. The three frequencies generated in this example will be $f'_{j-1} = \delta_f + 1/T$, $f'_j = \delta_f$, and $f'_{j+1} = \delta_f - 1/T$. These frequencies will be primarily detected at pixel $n = j$, but they will also be detected on the adjacent pixels $n = j-1$ and n=$j+1$. In the bottom of Figure 5.1.4 the temporal spectrum of the beat frequencies generated by the interference with the three DLO beads are illustrated, with the appropriate amplitudes. Because of the finite duration of the repetitive chirp reference, or the finite integration time of the detector, there is a small broadening of the beat oscillation frequency, and this leads to the fundamental limit on system fine frequency resolution. Notice that the two undesired sidebands will always have temporal oscillation frequencies greater than $1/2T$ and separated from the fine frequency baseband component δ_f by $1/T$. For the example of a single tone, the resulting modulation due to the final interferometric term in Equation 5.1.10 will be given by the sum of the three DLO beat frequencies.

$$
\begin{aligned}
I_i(u,y,t) =& c_i p(m'_y y) p(m_y y) \mathrm{rect}\left(\frac{t - NT/2}{NT}\right) \\
& \sum_{l'=-1}^{+1} aTW(uv_a - (j/T + \delta_f))W(uv_a - (j + l')/T) \\
& \cos\left[2\pi\left(\frac{(y-y_0)^2}{2\lambda f_y} + \frac{\beta}{2}(t-t_0)^2 + \frac{(j+l')t}{T} - \left(\frac{j}{T} + \delta_f\right)t\right) + \phi\right]
\end{aligned}
\tag{5.1.12}
$$

The constant $c_i = \eta 2\mathrm{Re}(d_s d_r \sqrt{\xi}\sqrt{1-\chi}\sqrt{\frac{i}{b}}\sqrt{1-\xi}\sqrt{\chi})$ has absorbed all of the multiplicative constants, and the arbitrary phase term ϕ, may be adjusted to zero by

minute motions of the mirrors or beamsplitters. The temporal modulation produced within a single detector pixel will be given by the weighted sum of the three beat terms being considered, integrated over the spatial aperture of that pixels responsivity, which can be represented with our abbreviated notation.

$$I_i(n,y,t) = c_i p(m'_y y) p(m_y y) \text{rect}\left(\frac{t - NT/2}{NT}\right) \tag{5.1.13}$$
$$aT \sum_{l'=-1}^{+1} a_{l'}^n (\delta_f - \frac{l'}{T}) \cos\left[2\pi\left(\frac{(y-y_0)^2}{2\lambda f_y} + \frac{\beta}{2}(t-t_0)^2 - \left(\delta_f - \frac{l'}{T}\right)t\right) + \phi\right]$$

Each pixel time integrates the incident intensity for a period equal to the time between spatial shifts of the TDI detector, $\Delta t \leq T/2$, which will adequately sample the desired baseband component, but will under sample the two sideband contributions. The sidebands are offset in frequency by $1/T$, and will have an extra temporal oscillation during the local integration period, which will average to zero leaving a smaller amount of time to integrate a nonzero component. The temporal chirp must have a maximum frequency which is within the Nyquist limited frequency response of the TDI detector sampling $|f_{max}| \leq \Delta t/2$. The temporal chirp sweeps from $-\beta t_0$ to $\beta(NT - t_0)$, which will be bandlimited to the desired bandwidth if it is generated from a phase stepped reference adequately sampled at the same frequency as the CCD local time integration between TDI shifts. With this sampled chirp reference approach, the reference function phase will be constant during a single local pixel integration time, and it will only change from integration site to integration site. Thus the local time integration only acts upon the three temporal oscillation frequencies generated by the interference with the DLO beads.

$$I_i(n,y,k) = \int_{(k-\frac{1}{2})\Delta t}^{(k+\frac{1}{2})\Delta t} I_i(n,y,t)dt \tag{5.1.14}$$
$$= c_i p(m'_y y) p(m_y y) \left[\cos\left[2\pi\left(\frac{(y-y_0)^2}{2\lambda f_y} + \frac{\beta}{2}(k\Delta t - t_0)^2\right)\right]\right.$$
$$\left.\int_{(k-\frac{1}{2})\Delta t}^{(k+\frac{1}{2})\Delta t} aT \sum_{l'=-1}^{+1} a_{l'}^n (\delta_f - \frac{l'}{T}) \cos\left[2\pi(\delta_f - \frac{l'}{T})t + \phi\right] dt\right.$$

$$
\begin{aligned}
&-\sin\left[2\pi\left(\frac{(y-y_0)^2}{2\lambda f_y}+\frac{\beta}{2}(k\Delta t-t_0)^2\right)\right] \\
&\quad\int_{(k-\frac{1}{2})\Delta t}^{(k+\frac{1}{2})\Delta t} aT\sum_{l'=-1}^{+1} a_{l'}^{n}(\delta_f-\frac{l'}{T})\sin\left[2\pi(\delta_f-\frac{l'}{T})t-\phi\right]dt\Bigg] \\
&=c_i p(m_y' y)p(m_y y)\left[\cos\left[2\pi\left(\frac{(y-y_0)^2}{2\lambda f_y}+\frac{\beta}{2}(k\Delta t-t_0)^2\right)\right]\right. \\
&\quad aT\sum_{l'=-1}^{+1} a_{l'}^{n}(\delta_f-\frac{l'}{T})\mathrm{sinc}[(\delta_f-\frac{l'}{T})\Delta t]\Delta t\cos\left[2\pi(\delta_f-\frac{l'}{T})k\Delta t+\phi\right] \\
&-\sin\left[2\pi\left(\frac{(y-y_0)^2}{2\lambda f_y}+\frac{\beta}{2}(k\Delta t-t_0)^2\right)\right] \\
&\quad aT\sum_{l'=-1}^{+1} a_{l'}^{n}(\delta_f-\frac{l'}{T})\mathrm{sinc}[(\delta_f-\frac{l'}{T})\Delta t]\Delta t\sin\left[2\pi(\delta_f-\frac{l'}{T})k\Delta t-\phi\right]\Bigg] \\
&=c_i p(m_y' y)p(m_y y)aT\sum_{l'=-1}^{+1} a_{l'}^{n}(\delta_f-\frac{l'}{T})\Delta t\mathrm{sinc}[(\delta_f-\frac{l'}{T})\Delta t] \\
&\quad\cos\left[2\pi\left(\frac{(y-y_0)^2}{2\lambda f_y}+\frac{\beta}{2}(k\Delta t-t_0)^2-(\delta_f-\frac{l'}{T})k\Delta t\right)+\phi\right]
\end{aligned}
$$

The effective strength of the undesired sidebands is further suppressed by the sinc function response of the temporal local averaging integration, giving effective amplitudes for the three beat frequencies $b_l^n(f_l) = a_l^n(f_l)\Delta t\mathrm{sinc}[f_l\Delta t]$. Since this is now a sampled signal representation with sample time Δt, any frequency component higher than the Nyquist rate $1/2\Delta t$ will be aliased. The desired baseband components should always be sampled sufficiently, since the TDI sampling rate is required to be faster than the laser pulsing rate, $f_s = 1/\Delta t \geq 1/T$ for double sided processing ($f_s = 1/\Delta t \geq 2/T$ for single sided processing), but the undesired sideband beats can fold down into the spectral band of interest. The resulting aliased frequency will always be within the range $0 < f_l'^a < f_s/2$, and is given by $f_l'^a = \frac{f_s}{2} - |f_l' \mathrm{mod} f_s - \frac{f_s}{2}|$. Interestingly, if the TDI shift frequency is chosen to be equal to the reference repetition rate $f_s = 1/T$, then the sidebands will alias to the desired baseband frequency $f_l'^a = \delta_f = f_0'$. In this case the resulting spectral peaks due to the sidebands will be at the same fine frequency position as the desired main lobe, or they would be at the same fine frequency ordinate in the adjacent

coarse frequency pixel bins. This would produce separable sidelobe responses in the appearance of a sinc function cross, like that due to the 2-D Fourier transform of a rectangular aperture. The detected baseband beat tones are now sampled in discrete time and can be processed by the TDI CCD.

5.1.4 TDI detector output of the DLO interferometer

The detector operates in the TDI mode in which it time integrates in a sliding window coordinate frame, moving in discrete steps with an approximately continuous velocity $v = \Delta y/\Delta t$. Each pixel with 2-D spatial response profile $h(x,y) = h_x(x)h_y(y)$, time integrates the incident intensity for a period $\Delta t \leq T$ ($\Delta t \leq T/2$ for single sided processing), then shifts the accumulated photogenerated charge profile by one pixel of width Δy in the y direction. Integration commences at the new location and adds to the previously accumulated charge. As a potential well traverses the full aperture $M\Delta y = H$ of the CCD, it integrates up three components; the time averaged reference beam intensity, the time varying average of the signal beam intensity, and the processed interferometrically detected component which produces a fine frequency scan within each coarse frequency bin. This operation is described by the multichannel TDI equation.

$$D(n, t = m\Delta t) = \sum_{k=m-M}^{m} \int_{(k-\frac{1}{2}\Delta t}^{(k+\frac{1}{2})\Delta t} \int\int h_x(\lambda F(u - n\Delta u)) h_y(y - (m-k)\Delta y) \, I(u,y,t)\,du\,dy\,dt \tag{5.1.15}$$

This equation can be simplified by assuming that the CCD pixel response function acts like a spatially sampling impulse $H(x,y) = \delta(x,y)$. However it should be realized that the degradation of the CCD modulation transfer function (MTF), given by the Fourier transform of the separable pixel response will apodize the chirp transform in the y dimension, and may decrease the achievable resolution of the space integrating spectrum analysis in the u (or x) dimension. The effect

of the pixel spatial response weighted overlap integral of the DLO beads with the signal blur spot will be accounted for separately as in Equation 5.1.11. The above analysis of the temporal sampling performed by time integration at each stationary pixel positions, with widths equal to the DLO bead separation can be used, in order to obtain the simplified expression for the processing performed by the TDI CCD on the sampled incident intensity.

$$D(n,t=m\Delta t)=\sum_{k=m-M}^{m} I(n,(m-k)\Delta y,k\Delta t) \tag{5.1.16}$$

The result of this processing operation will be to perform a convolution between the time varying portion of the incident locally integrated intensity and the y spatial intensity profile detected by the CCD. Since the spatial intensity is a chirp, and the temporal variation is a temporal chirp multiplied by a superposition of sinusoidal tones, the resulting output from each TDI column will be a chirp transform fine frequency scan of the signals within that coarse frequency bin, plus bias terms. This can be shown by plugging the locally integrated incident intensity for a single tone into the simplified multichannel TDI equation.

$$\begin{aligned}
&D(n,t=m\Delta t)=\sum_{k=m-M}^{m}\operatorname{rect}\left(\frac{k\Delta t-NT/2}{NT}\right)\\
&\Big[c_s^2p^2(m_y(m-k)\Delta y)\int h(u-n\Delta u)|aW(uv_a-f')|^2du\\
&+c_r^2p^2(m_y'(m-k)\Delta y)\int h(u-n\Delta u)\left|\left[\operatorname{rect}\left(\frac{uv_a+f_c}{B}\right)T\operatorname{comb}(uv_aT)\right]*W(uv_a)\right|^2du\\
&+c_i\bar{p}((m-k)\Delta y)aT\sum_{l'=-1}^{+1}a_{l'}^n(\delta_f-\frac{l'}{T})\Delta t\operatorname{sinc}[(\delta_f-\frac{l'}{T})\Delta t]\\
&\cos\Big[2\pi\left(\frac{((m-k)\Delta y-y_0)^2}{2\lambda f_y}+\frac{\beta}{2}(k\Delta t-t_0)^2-(\delta_f-\frac{l'}{T})k\Delta t\right)+\phi\Big]\Big]
\end{aligned} \tag{5.1.17}$$

The constant $c_s^2=d_s^2\xi(1-\chi)$ represents the signal power at the CCD, while $c_r^2=d_r^2(1-\xi)\chi/b$ represents the reference beam power, and the amplitude of the interferometric modulation is proportional to $c_i=2\eta c_sc_r$. The first term is the

sliding window weighted average of the pixel detected space integrated coarse spectrum of the single tone signal scrolling out of each TDI column. It can vary slowly in time, bandwidth limited by the full integration time $M\Delta t$, as the average power within each coarse frequency bin changes. The second term is the pixel integrated reference intensity, and should be approximately uniform at all coarse frequency TDI columns, but since it is gated on and off it will vary in amplitude as it comes out of each TDI column as that column scrolls downward. In the case of no laser beam apodization in the y dimension its dependence on m would be trapezoidal, which would become triangular if M=N, and the inclusion of the magnified aperture profile $p^2(m'_y y)$ would round the corners of the trapezoid. Both bias terms can be made approximately constant in m by performing repetitive chirp transforms without a break in between, thereby keeping the reference and signal powers constant, but this may lead to overlapping of successively computed spectral terms. The final term is the term of interest, which represents an apodized, spectrally weighted, chirp transform operating on all of the fine frequency components which are interferometrically generated within each coarse frequency TDI column, and the effects of pixel spatial sampling in u and local time integration have been included. The hybrid apodization function, $\bar{p}$, includes the product between the magnified signal and reference beam aperture profiles, with a further weighing of the interferometric spatial chirp modulation depth by the MTF of the CCD, which is given by the Fourier transform of the pixel response function in y, given by $H_y(f_y)$. Because the spatial chirp profile has a linear dependence of the spatial frequency on spatial position, the effect of the MTF is to weight the chirp profile with the scaled CCD MTF centered at the chirp offset position k_0 at which the spatial frequency goes to zero, with the appropriate chirp rate spatial frequency to space conversion factor, so that $\bar{p}(k\Delta y) = p(m'_y k\Delta y)p(m_y k\Delta y)H_y((k-k_0)\frac{\Delta y}{\lambda f_y})$. The chirp algorithm also results in additional trapezoidal or triangular spectral weighing due to variations

in the chirp overlap, as the temporal chirp is shifted by multiplication with fine frequency sinusoidal tone signal terms. The resolution of the chirp algorithm will also decrease as the matching chirp overlap decreases, and the carrier frequency varies as the center frequency of the matching chirp overlap region changes linearly with frequency, resulting in a quadratic phase factor.

The interferometrically generated TDI chirp transform component can be analytically represented using a similar analysis to that in section 2.4 on the chirp transform algorithm. The temporal quadratic chirp rate β, must be set equal to the spatial quadratic chirp rate $b_y = 1/\lambda f_y$, as temporally detected in the TDI coordinate frame. This will result in the cancellation of the quadratic terms in k, and requires that $\Delta y^2/\lambda f_y = -\beta \Delta t^2$. The minus sign in this expression means that a correlation between matched chirp segments is being performed rather than a convolution of mirrored chirps, which are equivalent for two sided chirps. The offset of the spatial and temporal chirps should be set to cancel in order to maximize the chirp overlap at midband, and this leads to the requirement $y_0 \Delta y/\lambda f_y = \beta \Delta t t_0$. Additionally, the arbitrary phase factors can be aligned if desired, in order to cancel one more term by requiring $y_0^2/\lambda f_y = -\beta t_0^2$. In order to simplify this equation I will assume that $T = \Delta t$, and $N = M$, so the chirp transform will have a triangular gain and resolution response, and the double sided, baseband signal representation will have the sidebands aliased to the main lobe fine frequency. The effect of the optical amplitude profile along the TDI direction, $\bar{p}(k\Delta y)$, will be to suppress the sidelobes and decrease the achievable spectral resolution, via a convolution with the transform of the spatial apodizing profile, and this effect will be left out in order to simplify the final result. The interferometric term for a single frame of processing, can be analyzed over the region $|m| < M$, where the output is nonzero, for the case of a single tone, by performing the interferometric TDI chirp transform summation

upon the final term of Equation 5.1.17.

$$
\begin{aligned}
D_i(n,m) =& aT \sum_{k=m-M}^{m} c_i \bar{p}((m-k)\Delta y)\mathrm{rect}\left(\frac{k\Delta t - NT/2}{NT}\right) \sum_{l=-1}^{+1} b_l^n(\delta_f - \tfrac{l}{T}) \\
& \cos\left[2\pi\left(\frac{((m-k)\Delta y - y_0)^2}{2\lambda f_y} + \frac{\beta}{2}(k\Delta t - t_0)^2 - (\delta_f - \frac{l}{T})k\Delta t\right) + \phi\right] \\
=& c_i aT \sum_{l=-1}^{+1} b_l^n(\delta_f - \tfrac{l}{T}) \sum_{k=-\infty}^{\infty} \mathrm{rect}\left(\frac{k - NT/2\Delta t}{NT/\Delta t}\right) \mathrm{rect}\left(\frac{k + (m - \frac{M}{2})}{M}\right) \\
& \cos\left[2\pi\left(m^2\beta\Delta t^2/2 + m\beta\Delta t t_0 - mk\beta\Delta t^2 - (\delta_f - \tfrac{l}{T})k\Delta t\right) + \phi\right] \\
=& c_i aT \sum_{l=-1}^{+1} b_l^n(\delta_f - \tfrac{l}{T}) \sum_k \mathrm{rect}\left(\frac{k - (M-m)/2}{M - |m|}\right) \\
& \frac{1}{2}\left[e^{-i\left[2\pi\left(mk\beta T^2 + (\delta_f - \frac{l}{T})kT - m^2\beta T^2/2 - m\beta T t_0\right) + \phi\right]} + c.c.\right] \\
=& c_i aT \sum_{l=-1}^{+1} b_l^n(\delta_f - \tfrac{l}{T}) \frac{1}{2}\left[e^{i[2\pi(m^2\beta T^2/2 + m\beta T t_0) - \phi]}\right. \\
& \left.(M - |m|)\mathrm{sinc}[(M-|m|)(m\beta T^2 + \delta_f T)]e^{-i\pi[(m\beta T^2 + \delta_f)(M-m)]} + c.c.\right] \\
=& c_i aT \sum_{l=-1}^{+1} b_l^n(\delta_f - \tfrac{l}{T})(M - |m|)\mathrm{sinc}[(M-|m|)T(m\beta T + \delta_f)] \\
& \cos\left[2\pi\left(m^2\beta\Delta t^2 + m(\beta\Delta t(t_0 - M\Delta t/2) - \delta_f/2) + \delta_f M/2\right) - \phi\right] \\
=& c_i aT \sum_{l=-1}^{+1} b_l^n(\delta_f - \tfrac{l}{T})(M - |m|)\mathrm{sinc}[(M-|m|)T(m\beta T + \delta_f)]\cos[\varphi] \\
& \qquad\qquad |m| < M \qquad (5.1.18)
\end{aligned}
$$

The impulse response of the unapodized TDI chirp transform is a sinc function located at a position proportional to the fine frequency $m = -\delta_f/\beta T$, with a gain and resolution width which are triangular, and maximized for $m = 0$, placed on a chirping carrier, with a complicated phase factor in y represented by φ. The idealized spectral resolution is given by the sinc lobe width in frequency $\delta f = 1/(M - |m|)T$, which is an inverse triangular function with maximum resolution given by the inverse of the integration time seen by a potential well fully traversing the CCD aperture. The spectral gain $M - |m|$ is proportional to the number of TDI sites at which integration takes place. The triangular spectral gain can be corrected after bias removal by an inverse triangular amplification, but the resulting signal

to noise ratio will be best at $m = 0$, and will degrade away from this position. The potential well that is at the top of the CCD as the processing begins will coherently integrate signal as it propagates down the TDI column, and will achieve the best possible resolution. The potential wells that begin the processing operation in the middle of the CCD will achieve a gain and resolution proportional to the time during which they are within the integration window of the CCD. Similarly potential wells that enter the top of the CCD after processing has begun will only integrate signal for the time that processing continues after they have entered the CCD aperture, and not the full time that it takes to traverse the CCD, unless $NT > M\Delta t$. In theory the complicated chirping carrier phase term φ can be removed with electronic post processing, but in practice ideal phase alignment would be difficult. A much better approach to carrier encoding for the purpose of bias removal, would be to introduce a spatial carrier in the coarse dimension with a rotation of the recombining beam splitter, followed by carrier demodulation in the coarse dimension. This requires an increase in the number of pixels in the coarse dimension in order to adequately represent several carrier cycles within each coarse frequency resolvable spot, and the individual pixels are then much smaller than the blur spot. For this reason a CCD with very high resolution orthogonal to the TDI direction would be very appropriate, such as the Itek OSD 2048x96 space surveillance TDI array with 96 TDI gain stages[4.39]. Another possibility for removing the bias would be to use an auxiliary bias subtraction CCD, in which case a more nearly square array would be appropriate.

In order to visualize the coherently detected and TDI chirp transformed CCD output, the pixel detected, locally integrated response for the single tone being considered will be expanded, using Equation 5.1.11, and including the sidelobe

suppression sinc derived in Equation 5.1.14.

$$b_l^n(\delta_f - \tfrac{l}{T}) = \Delta t \text{sinc}[(\delta_f - \tfrac{l}{T})\Delta t] \int h_u(u - n\Delta u) W(u - \frac{j}{v_a T} + \frac{\delta_f}{v_a}) W(u - \frac{(j+l)}{v_a T}) du \tag{5.1.19}$$

This equation shows that the additional sidelobe suppression due to the local temporal integration between TDI shifts suppresses the higher frequency interferometric beats which are due to the overlap of the signal blur spot with the adjacent reference DLO beads on the adjacent pixel in x. The choice of $T = \Delta t$ causes the aliased fine frequency sidelobes to have the same TDI sampled apparent frequency as the main peak so that the sidelobes at adjacent pixels will produce chirp transformed outputs at the same fine frequency locus as the main lobe. Within the $n = j$th column, the overlap with adjacent DLO beads will produce spectral beats that will be aliased to the same frequency as the desired baseband beat by the TDI sampling as long as the DLO reference waveform repeats each TDI shift time. In this case the resulting chirp transformed main peak will be at the same location as the weaker adjacent bead overlap contributions.

The system impulse response to a single tone can be represented as the sum of the signal bias and the reference beam bias plus the interferometric TDI processed folded spectrum contribution.

$$\begin{aligned} D(n,m) =& c_s^2 T(M - |m|) \int h(u - n\Delta u) |aW(uv_a - f')|^2 du \\ &+ c_r^2 T(M - |m|) \text{rect}\left(\frac{n\Delta u v_a + f_c}{B}\right) W^2 \\ &+ c_i a T \sum_{l=-1}^{+1} \Delta t \text{sinc}[(\delta_f - \tfrac{l}{T})\Delta t] \int h_u(u - n\Delta u) W(u - \tfrac{j}{v_a T} + \tfrac{\delta_f}{v_a}) W(u - \tfrac{(j+l)}{v_a T}) du \\ &(M - |m|) \text{sinc}\Big[(M - |m|)T(m\beta T + \delta_f)\Big] \cos[\varphi + 2\pi\alpha_x n\Delta x] \end{aligned} \tag{5.1.20}$$

All of the simplifying assumptions have been used in this representation, $M = N$, $T = \Delta t$, and $1/T = \Delta u v_a = \Delta x v_a / \lambda F$, and the additional spatial carrier in the coarse dimension has been explicitly included. The apodization of the

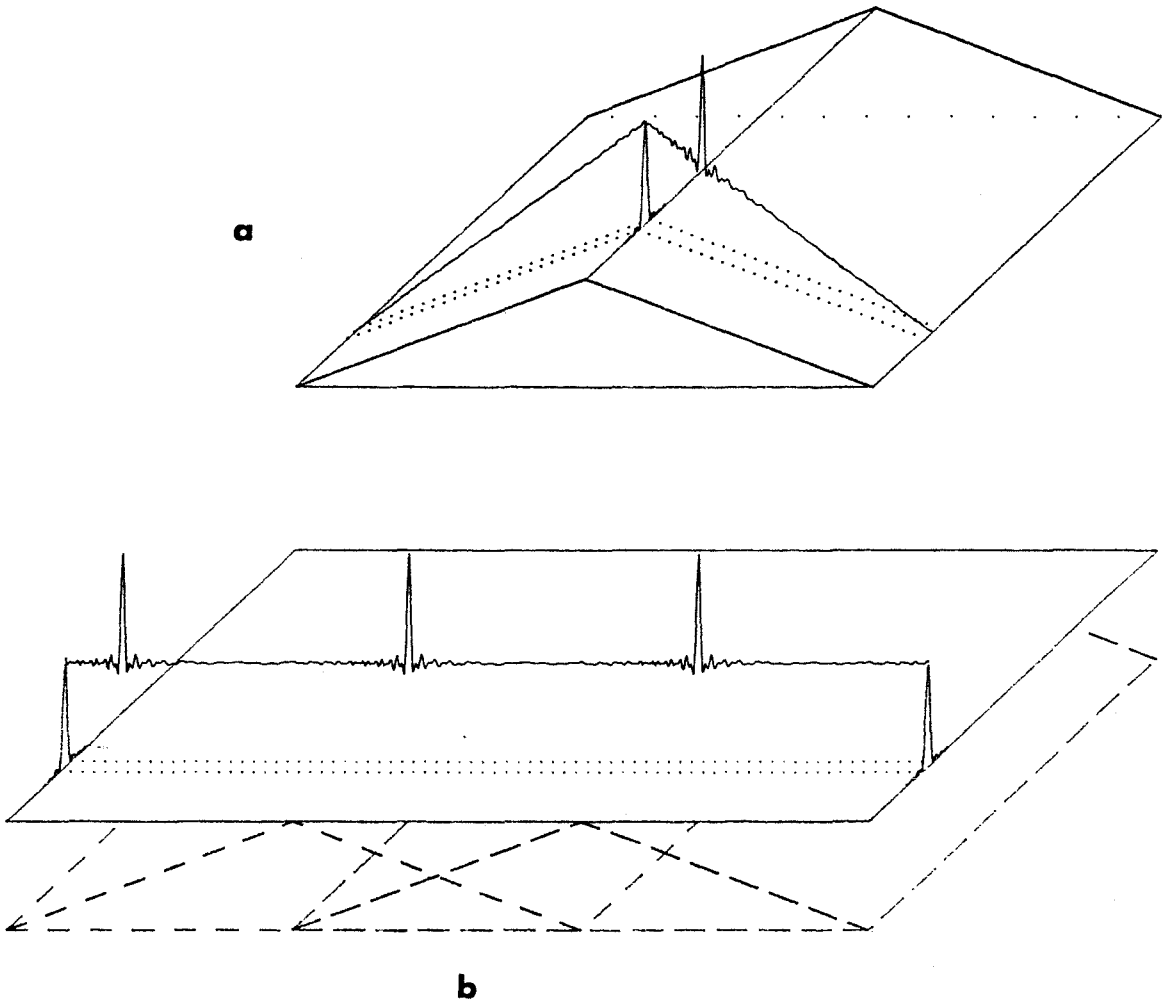


Figure 5.1.5. a) Idealized impulse response of the TDI chirp transform DLO interferometer, with triangular bias variation. b) Overlap of successively computed spectra, which eliminates bias variation.

TDI chirp transform is ignored in order to examine the idealized system performance. The power in the reference beam integrated over individual coarse pixel columns is assumed constant for all pixels, and is abbreviated by $W^2 = \int h(u-n\Delta u)|T\text{comb}(uv_aT) * W(uv_a)|^2 du$. Since the laser is assumed to be gated on and off for a period NT which is the same as the TDI integration time $M\Delta t$, the bias terms have triangular weighing. When successive folded spectrum operations are performed back to back then the bias variation will be removed, but the tails of adjacent fine frequency scans will extend into the adjacent frames of computation. With apodization of the chirp transform, then the bias variation is no longer exactly triangular, and successive overlapping transforms will not produce a perfectly flat bias contribution, and the bias will have small undulations which would repeat with each transform segment. The idealized impulse response for this system is shown in Figure 5.1.5a, it consists of the triangular bias due to the reference and signal beam power, plus the interferometrically detected, TDI chirp transformed folded spectrum information which rides along the incoherently detected coarse frequency ridge. The position of the peak along the TDI column represents the fine frequency of the single tone input and the position of the space integrated ridge represents its coarse frequency. The overlap of successively computed folded spectrum calculations results in the removal of the triangular bias variation, as is illustrated schematically in Figure 5.1.5b, but also causes the fine frequency scans along each coarse frequency bin to overlap. The carrier placed in the coarse frequency dimension is schematically illustrated in the idealized isometric projection of the theoretical impulse response shown in Figure 5.1.6. The lines plotted along the coarse frequency ridge are seen to be at the Nyquist limit for the coarse frequency carrier, so that the scan at the center of the coarse frequency ridge contains a constructive peak, while the adjacent line contains a destructive peak, and so on. In reality, many more cycles at a lower sampled frequency than the Nyquist frequency would have to be utilized, in order

to separate the spatial frequency spectrum of the incoherent coarse ridge from the coherently detected peak on a carrier in the coarse dimension, for effective demodulation of the signal from the bias. The additional sidelobe suppression factor due to the time integration between TDI shifts, of the temporal beats generated by the interference with the adjacent DLO beads is manifested as a rapid decrease of the interferometric sidelobes in the coarse dimension, falling almost as fast as the power in the incoherent ridge. When apodization on the coarse dimension space integrating spectrum analyzers is included, as in Figure 5.1.6b, the sidelobes in the coarse dimension fall so fast they are practically invisible on the scale of this plot, and the main lobe is widened, which helps separate the spatial frequency spectrum of the incoherent ridge from that of the peak on a carrier.

If we assume that the signal spectrum is composed of a number of CW tones at discrete frequencies f_q with amplitudes a_q, then the signal spectrum can be represented as the sum $\tilde{S}(f) = \sum_q a_q \delta(f + f_q)$. In order to decipher the processing operation that is occurring, we should represent the signal spectrum as a double sum over the coarse and fine frequency components.

$$\tilde{S}(f) = \sum_j \sum_k^N a_{jk} \delta(f + j/T + k/NT) \tag{5.1.21}$$

The folded spectrum processor is designed to reproduced the representation a_{jk}. This spectral representation can be inserted into the above equation to obtain the output of the multichannel TDI CCD, and this output will be given by the linear superposition of the input frequencies.

$$\begin{aligned} D(n,m) =& c_s^2 T(M - |m|) \sum_j \sum_k^N |a_{jk}|^2 \int h(u - n\Delta u) |aW(uv_a - \tfrac{j+k/N}{T})|^2 du \\ &+ c_r^2 T(M - |m|) \text{rect}\left(\frac{n\Delta u v_a + f_c}{B}\right) W^2 \\ &+ \sum_j \sum_k^N a_{jk} c_i T \sum_{l=-1}^{+1} \Delta t \text{sinc}[(\tfrac{k}{NT} - \tfrac{l}{T})\Delta t] \end{aligned} \tag{5.1.22}$$

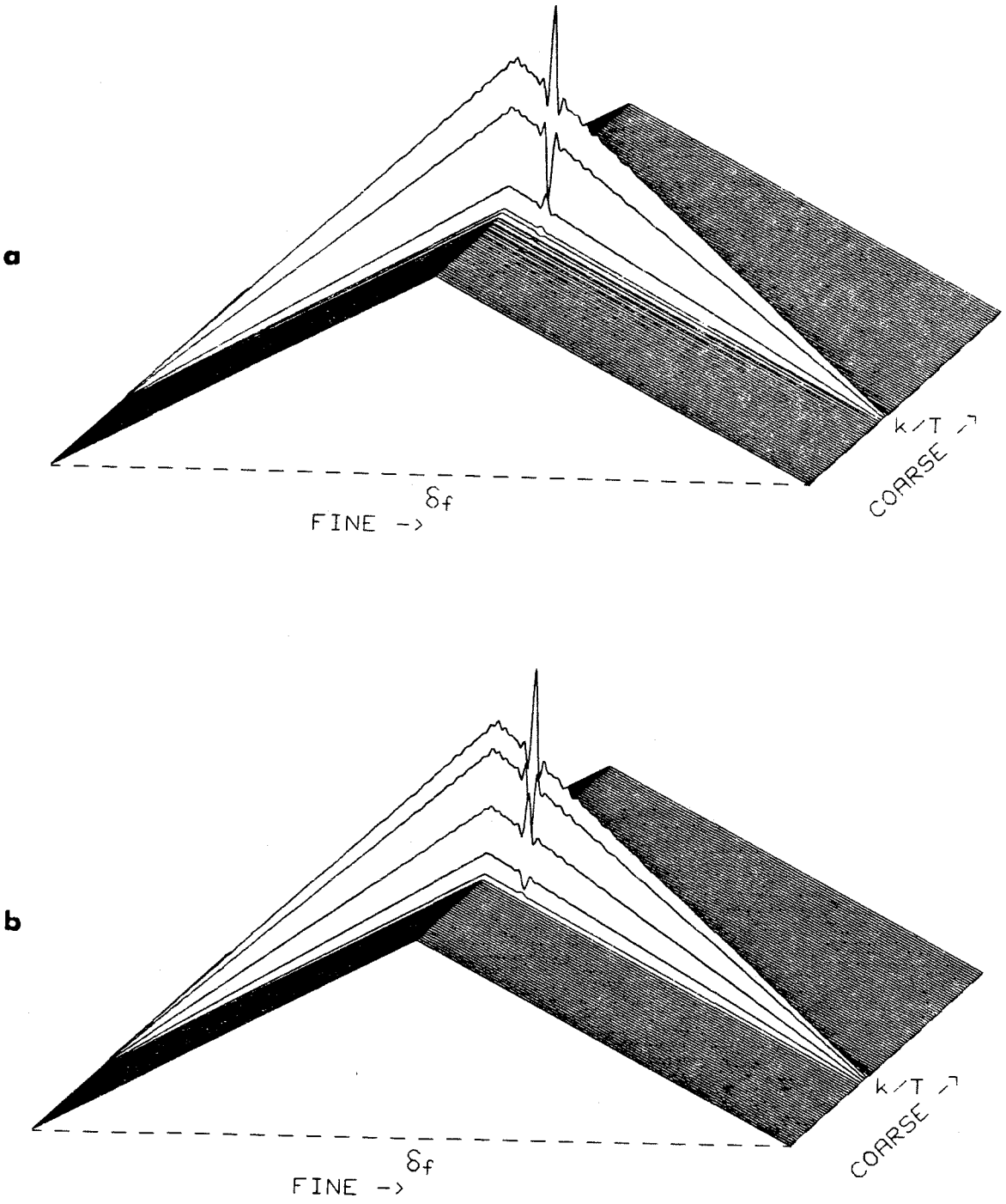


Figure 5.1.6. a) Isometric projection of folded spectrum output showing carrier frequency in coarse dimension. b) Isometric projection with apodization of Bragg cells included.

$$\int h_u(u - n\Delta u)W(u - \frac{j}{v_a T} + \frac{k}{v_a NT})W(u - \frac{(j+l)}{v_a T})du$$
$$(M - |m|)\mathrm{sinc}\Big[(M - |m|)T(m\beta T + \tfrac{k}{NT})\Big]\cos[\varphi + 2\pi\alpha_x n\Delta x]$$

Each spectral component will generate a chirp transformed peak on a chirping carrier at a position in the m dimension corresponding to its fine frequency $m = k/\beta NT^2$, along a triangular ridge locating its coarse frequency in the n dimension at the position $n\Delta u = j/v_a T$. The space integrated position of the incoherently detected bias ridge within each coarse bin will not vary significantly with the fine frequency, and for this reason its strength will be proportional to the sum of the squares of the elements in that column of the folded spectrum matrix $\bar{a}_j^2 = \sum_k |a_{jk}|^2$, which is the same as a standard space integrating AO power spectrum analyzer output. If a very strong single tone jammer or a strong narrowband modulated signal is applied to the time and space integrating folded spectrum processor, then it will be spatially channelized, and it will saturate one or two coarse frequency bins, preventing fine frequency analysis within those channels, but other fine frequency processors will still operate unaffected by the strong signals.

The modulation amplitude of each detected peak is a triangular gain function of the fine frequency offset times the spectral amplitude of that frequency component, however the modulation depth of each peak depends on the total power within that coarse frequency bin. That is because modulation depth, or signal to bias ratio (SBR), is defined as coherently detected fringe depth over twice the bias level, and optimizing the system requires that the reference beam power within a pixel, $c_r^2 = d_r^2(1 - \xi)\chi/b$, should be equal to the expected signal beam power, $\bar{a}_j^2 c_s^2 = \bar{a}_j^2 d_s^2 \xi(1 - \chi)$. The modulation depth of the frequency component in the jth coarse frequency bin with a fine frequency k/NT, depends on the optical efficiencies c_i, c_r^2, c_s^2, the coherence factor η, the signal amplitudes a_{jk} the power in that column $\bar{a}_j^2$,

and the time integrated overlap integral b_l^n defined in Eq. 5.1.19.

$$m_{jk} = \eta \frac{2c_s c_r a_{jk}}{c_s^2 \bar{a}_j^2 + c_r^2} b_0^n(k/NT) \tag{5.1.23}$$

In order to optimize the modulation depth and the resulting demodulated folded spectrum dynamic range the reference beam power should be made equal to the expected signal beam power. When a single tone of the expected average power falls in its own coarse frequency bin, $c_s^2 a_j^2 = c_r^2$, then the modulation depth will be the amplitude of the coherence function, η, times the sidelobe suppression factor, which for low offset frequencies will be negligible. However, when K equally strong tones fall within the same coarse frequency bin with a total power equal to the reference beam then the modulation depth of each tone will only be $\eta/\sqrt{K}$ times the sidelobe suppression factors for each tone. This is illustrated schematically in Figure 5.1.7, neglecting the sidelobe suppression factors, where 4 coarse frequency bins contain equivalent energy, but K varies from 1 at the back of the plot, through 8 at the front, by powers of 2 with randomly positioned fine frequencies within each coarse frequency bin. Several observations can be made from these plots, the modulation depth decreases as the power within a coarse frequency bin is divided up among more tones, and the triangular weighing function decreases the amplitude of large fine frequency offset tones. The overlap of successively computed TDI folded spectra results in an overlapping of the fine frequency scans, and decreases the effective modulation depth of tones with a large fine frequency offset.

The fact that the modulation depth of each spectral component varies as $1/\sqrt{K}$ rather than $1/K$ is clearly important in dense multitone environments where K is large. This can be explained in terms of the coherent detection of K fine frequencies with $1/K$ of the total power of that coarse frequency bin in each fine frequency component. The amplitude of each term is $1/\sqrt{K}$ times the amplitude that would be observed if $K = 1$, and the interferometric detection produces a modulation proportional to this amplitude, and not each tones power. The maximum that K

can ever be is given by the number of resolvable fine frequencies per coarse frequency bin, which is bounded above by $K < N = M$, so the modulation depth of the worst case (white noise) would be $\eta/\sqrt{N}$, again neglecting the sidelobe suppression. This places a lower limit on the acceptable CCD dynamic range that could be utilized in a white noise environment, by requiring that the worst case modulation amplitude be several times larger than the detector noise. When the signal power in a coarse frequency bin is not equal to the reference power, optimum modulation depth will not be achieved. As a compromise the gain on the reference beam Bragg cell should be set to incoherently fill up one quarter of the CCD dynamic range, and the gain on the signal Bragg cell should be adjusted so the coarse frequency bin of the most interest also fills up about one quarter of the CCD dynamic range. In this case the perfectly coherent modulation due to a single tone in the optimized channel would swing from 0 to the CCD saturation level, stronger tones would saturate, and weaker tones would swing by less.

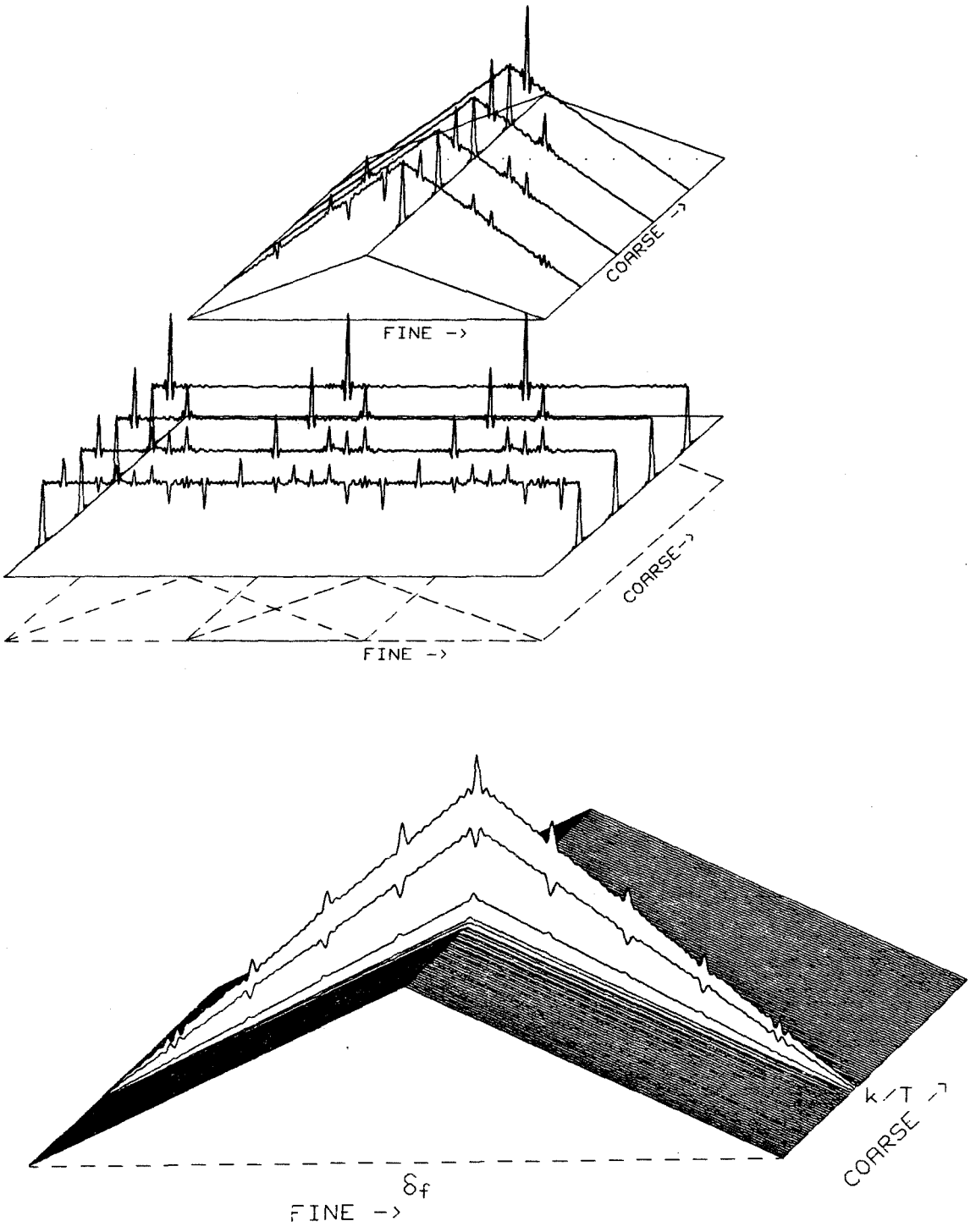


Figure 5.1.7. Idealized multitone spectra, showing that modulation depth decreases as more tones are added to a coarse frequency bin.

5.2 Pulsed Source Discrete Fourier Transform Folded Spectrum Processors

In this and the following sections, several possible implementations of a time and space integrating folded spectrum processors that utilize a pulsed laser diode as the optical source will be considered. The architectures in this section use crossed Bragg cells and a time integrating CCD in order to implement the desired two dimensional folded spectrum processing operation. The pulsed laser is an alternative approach to using a reference distributed local oscillator (DLO), which interferometrically beats the output of a space integrating acousto-optic spectrum analyzer to baseband. Instead, the sampling action of the pulsed laser effectively aliases interferometrically detected tones to baseband, thereby removing the coarse frequency content, leaving only the fine frequency for subsequent time integrating processing. Fine frequency analysis of each coarse frequency bin can be performed with a number of possible techniques; a multichannel DFT can be implemented on a starring CCD array, a multichannel TDI chirp transform can be performed as in the previous section, or a hybrid approach can be utilized to implement a TDI DFT operation. The discrete Fourier transformation (DFT) is a space variant operation that is implemented with a matrix vector multiplication, and the utilization of a pulsed source gives the system the extra versatility needed in order to perform this more general operation, as elucidated briefly in section 2.3.2.

In a more generalized context, it is possible to use the type of system described in this section to perform parallel time integrating matrix-vector multiplications on the field amplitudes of the outputs of a pulsed source space integrating processor. The space integrating processors could be a Vander Lugt correlator, or a free space chirp correlator, or a Fourier analyzer as discussed in this section. The system described in this section could thus be thought of as a special case of a frequency multiplexed time integrating outer product matrix-matrix multiplier, where in this case the matrix is a DFT.

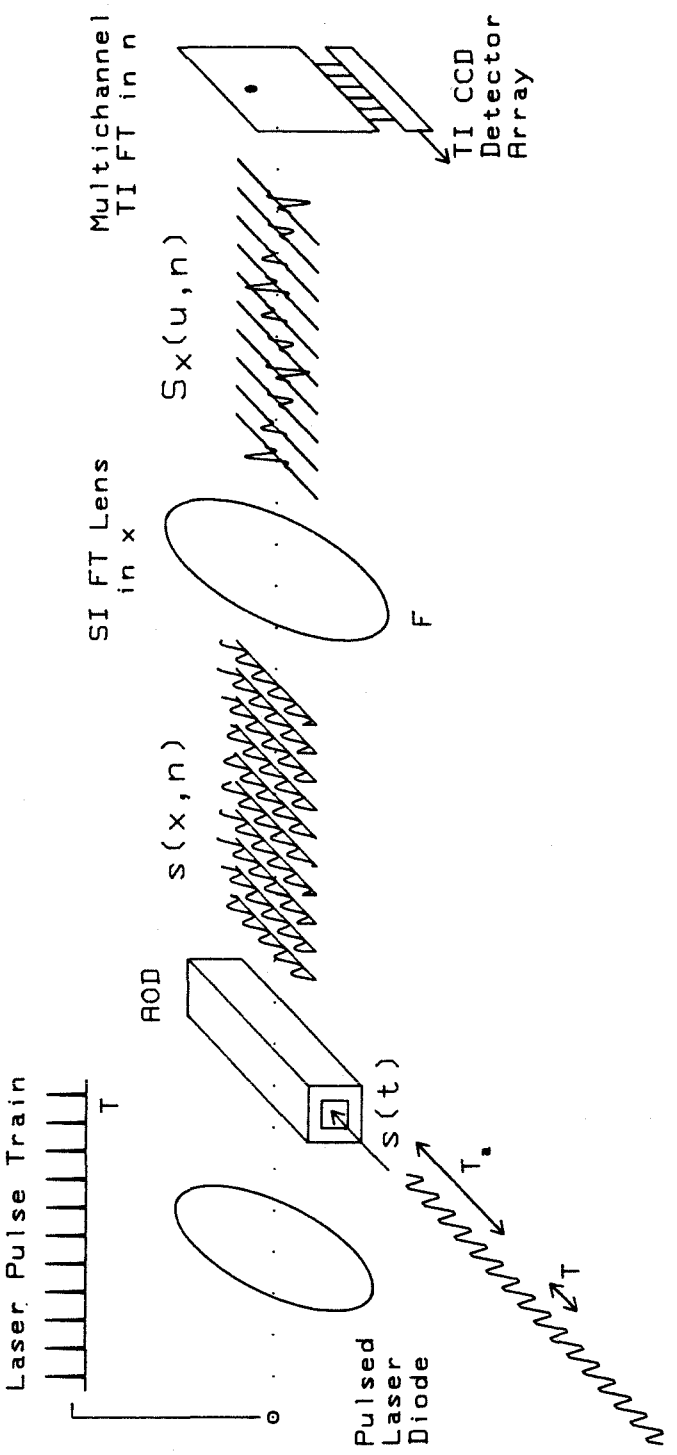


Figure 5.2.1. Schematic representation of the Time and Space Integrating (TSI) folded spectrum processor illuminated by a pulsed source which produces the space time raster, which is spatially Fourier transformed by a lens, and then temporally Fourier transformed by a CCD processor.

The action of the pulsed laser is to convert a very long signal applied to the AOD into a raster like format in space and time, rather than the more familiar 2-D spatial raster. Since the limited time bandwidth product of the AOD will not allow the entire signal to be within the aperture simultaneously, successive overlapping portions of the long signal will slide into the AOD and the acoustic motion will be frozen by pulsing the illuminating laser with a very narrow pulse. For each laser pulse the AOD transduces part of the input signal into a 1-D spatial modulation of a coherent optical wavefront, and on successive pulses it will transduce succeeding portions of the input signal into the optical system. Thus, the long 1-D signal has been transformed into a 2-D raster of space (x), and discrete time or pulse number (n). The transformation performed by the AOD illuminated with a repetitively pulsed laser produces a raster of space and time, where the spatial dimension is continuous, and time is discrete.

$$s(x,t) = \int s(t - x/v_a)\mathrm{rect}(x/A)\delta(t - nT)dt$$
$$s(x,nT) = s(nT - x/v_a)\mathrm{rect}(x/A) \qquad (5.2.1)$$

In this expression A is the AOD width, T is the laser pulse repetition interval, and n indexes the pulse number. If the AOD temporal window is greater than the laser pulse repetition interval $T_a = A/v_a > T$, then the raster is oversampled, which means that successive raster lines contain overlapping portions of the signal. The desired 1-D Fourier transform can be performed on the input signal $s(t)$, with a 2-D transform on its oversampled space-time raster representation $s(x,n)$. Because of the separability of the multidimensional Fourier kernel $e^{-i2\pi(ux+yv)} = e^{-i2\pi ux}e^{-i2\pi vy}$, a 2-D transformation can be performed as N_y 1-D transforms along the x axis followed by N_x 1-D transforms along the orthogonal y axis. This cascaded system can be implemented by performing a spatial Fourier transform in x with a lens on the diffracted wavefront for each of the N_y laser

diode pulses, producing an intermediate data set $S_x(u,n)$, which is also a space-time raster. This is followed by an array of N_x parallel time integrating narrowband Fourier transform modules multiplexed in x, which utilize the second spatial dimension (y) to perform the desired accumulation of partial products, thereby Fourier transforming $S_x(u,n)$ across n. After a succession of N_y laser diode pulses the full folded spectrum is produced in a 2-D format on a TI CCD detector array. A simple analytical representation of the idealized action of this system is as a cascade of a 1-D continuous Fourier transformation, and a 1-D discrete Fourier transform.

$$S(u,m) = \sum_{n=0}^{N-1} \left[\int_A s(x,n) e^{-i2\pi ux} dx \right] e^{-i2\pi mn/N} \tag{5.2.2}$$

This system is schematically illustrated in Figure 5.2.1, which illustrates the conversion of a very long 1-D signal into the space-time raster which is spatially Fourier transformed by a lens, and then temporally Fourier transformed by an array of narrowband time integrating spectrum analyzers.

5.2.1 Interferometric spectrum analyzer illuminated by a pulsed source

In order to model the detailed operation of this system, the operation of an interferometric space integrating acousto-optic spectrum analyzer that is illuminated with a narrow pulse train laser source will be considered. This type of 1-D interferometric spectrum analyzer is shown schematically in Figure 5.2.2. It is critical that the pulsing laser is an accurate subharmonic of the systems coherent reference clock signal, which is accomplished by digitally dividing a coherent reference frequency f_r by a convenient integer L'. The laser pulses are of width τ, separation $T = f_r/L'$, and amplitude c, so that the average laser power is decreased by a factor τ/T from that of a continuous source of the same amplitude. The pulse width must be very short, less than the inverse of the highest temporal frequency which will be interferometrically generated. Therefore the laser pulse width must

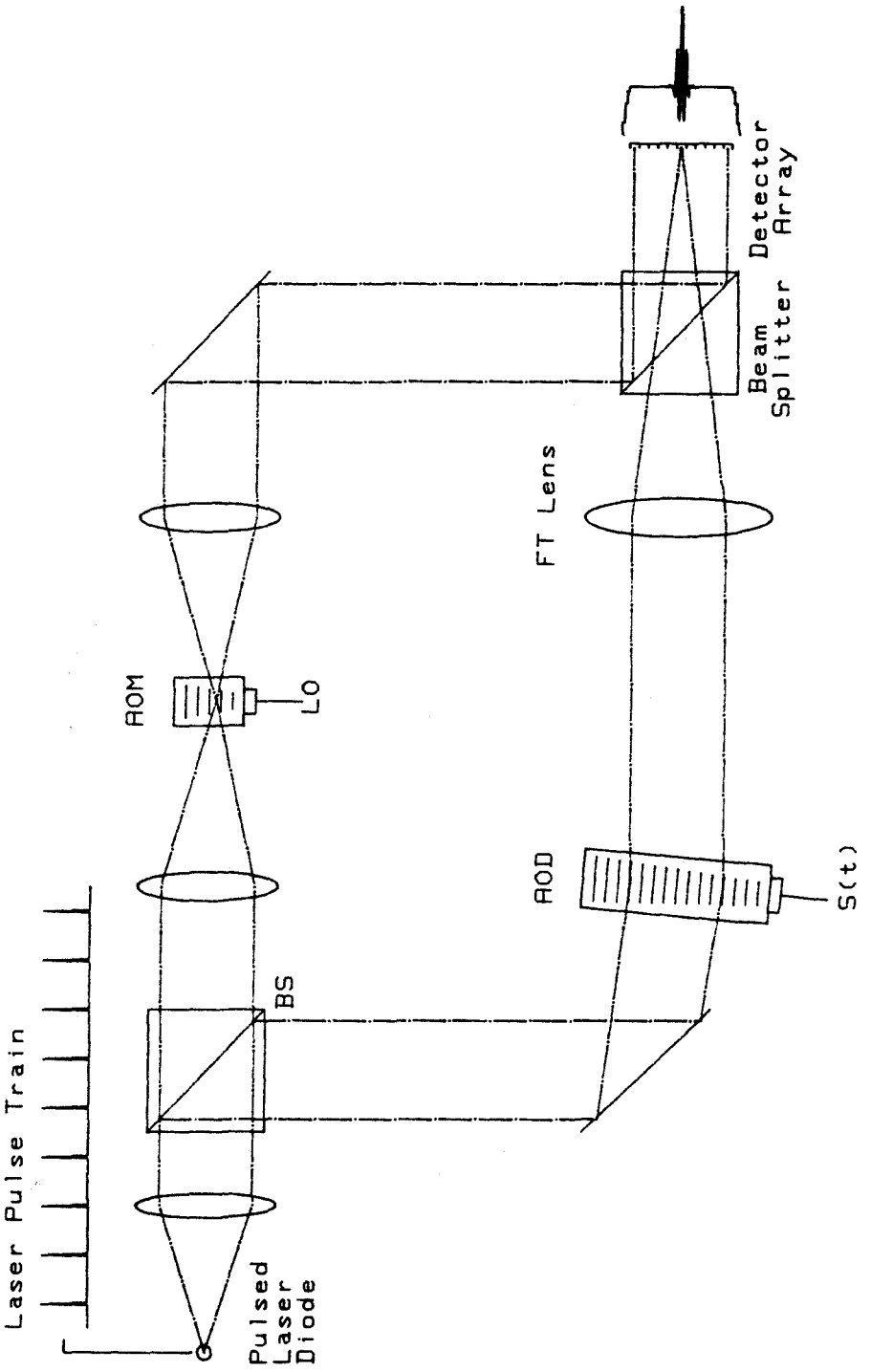


Figure 5.2.2. An interferometric AO spectrum analyzer illuminated by a repetitively pulsed source.

be less than half the inverse of the AOD bandwidth which is to be utilized $\tau < 2/B$, but of course it must be much greater than the optical period of oscillation for the optical monochromaticity condition to be valid. The pulse train can be represented in the time domain, or a Fourier identity can be used to represent it as a sum of its frequency harmonics.

$$g(t) = c\,\mathrm{rect}\left[\frac{t}{\tau}\right] * \sum_n \delta(t - nT) = c\sum_m T e^{i2\pi mt/T}\tau \mathrm{sinc}[m\tau/T] \tag{5.2.3}$$

This pulsed laser waveform is used to illuminate an AOD at the Bragg angle, and is split off to act as the reference beam of a generic interferometric spectrum analyzer system. When the reference beam is a plane wave of amplitude r_0, frequency $\nu' = \nu + f_0$ and normalized spatial frequency $\bar{\alpha}_u = \lambda F(\sin\theta/\lambda) = F\sin\theta$, where θ is the incidence angle of the reference wave with respect to the undiffracted signal beam, then the output of an interferometric spectrum analyzer is found by modifying equation 2.2.1.7.

$$\begin{aligned} I(u,t) =& g(t)\left|e^{-i2\pi\nu t}d\int \tilde{S}(uv_a)e^{i2\pi ft}W\left(u+\frac{f}{v_a}\right)df + r_0 e^{-i2\pi[\nu' t+\bar{\alpha}_u u]}\right|^2 \\ =& g(t)\Bigg[|d\tilde{S}(uv_a) * W(uv_a)|^2 + |r_0|^2 \\ &+ 2r_0 d\int |\tilde{S}(f)|\cos[2\pi(ft+\Omega(f)+\bar{\alpha}_u u + f_0 t)]W\left(u+\frac{f}{v_a}\right)df\Bigg] \\ =& I_s(u,t) + I_r(u,t) + I_i(u,t) \end{aligned} \tag{5.2.4}$$

In this equation the standard representation of the field due to the first order diffraction term at the back focal plane of a Bragg cell spectrum analyzer is interfered with a plane wave reference beam at a shifted frequency and angle, which results in additional spatial and temporal carriers. The signal applied to the AOD has a within band single sided negative frequency complex spectrum given by $\tilde{S}(f) = |\tilde{S}(f)|e^{i\Omega(f)}$, which consists of both magnitude and phase, and will result in an upshifted doppler interaction. The first term is the signal dependent bias, and the second term is the uniform reference beam bias, while the final term contains

the interferometrically produced spatial and temporal beats of interest. Both the amplitude and phase of the spectrum of the applied signal are reproduced as temporal and spatial modulations of each resolvable frequency bin, although there is a spatial blurring due to the AOD apodization. At each location in the output plane all the temporal harmonics of the laser diode pulse train mix with the interferometrically produced temporal beats resulting in a convolution of the frequency spectra. Time integration of this pattern for a period T results in a low pass filtering operation that selects out the baseband component at each position. As one moves along the output plane it would be noticed that the temporal frequency due to the interferometric detection increases linearly with normalized output variable u, and the harmonic of the pulsed laser that is responsible for producing a time integrated baseband component would jump to the next harmonic each distance $u = 1/v_a T$.

$$
\begin{aligned}
I_i(u,n) &= \int_{(n-\frac{1}{2})T}^{(n+\frac{1}{2})T} I_i(u,t)dt \qquad (5.2.5)\\
&= \int_{(n-\frac{1}{2})T}^{(n+\frac{1}{2})T} 2r_0 d \int |\tilde{S}(f)| \cos[2\pi((f+f_0)t + \Omega(f) + \bar{\alpha}_u u)] W\left(u + \frac{f}{v_a}\right) df \\
&\qquad c\sum_m T e^{i2\pi mt/T} \tau \mathrm{sinc}[m\tau/T]dt \\
&= c\sum_m \tau \mathrm{sinc}[m\tau/T] \int \mathrm{sinc}[(f+f_0-m/T)T] \\
&\qquad 2r_0 d|\tilde{S}(f)| \cos[2\pi(fnT + \Omega(f) + \bar{\alpha}_u u + f_0 nT - mn)] W\left(u + \frac{f}{v_a}\right) df
\end{aligned}
$$

Within the bandwidth due to the first sinc, which is imposed by the laser pulse width τ, the spectral components closest to a harmonic of the pulse repetition frequency will be selected by the second sinc function as the dominant time integrated temporal oscillation within any blur spot. This becomes clear when the case of a single sinusoidal tone at a frequency f', amplitude $|a|$ and phase Ω, with single sided upshifting spectrum $\tilde{S}(f) = |a|e^{i\Omega}\delta(f+f')$ is examined. It is convenient to consider the input temporal frequency to be broken up into a coarse and fine part by choosing the closest harmonic of the laser diode pulse repetition frequency as the

coarse term, and the offset from this frequency as the bandlimited fine resolution component. In this case, $f' = k/T + \delta_f$, where the coarse harmonic $k = \left[\frac{f'}{\frac{1}{T}}\right] = [f'T]$ is the greatest integer function (or the rounded integer function when double sided representation is to be used), and the baseband offset $0 < \delta_f = f' \mathrm{mod} \frac{1}{T} < 1/T$ is always less than the harmonic separation (or $|\delta_f| < 1/2T$ for double sided representation). For this case of single tone input, the time and space varying output of the interferometric spectrum analyzer illuminated by sampling impulses of the pulsed laser is

$$\begin{aligned} I_{f'}(u,t) =& g(t)\Big[|ad|^2 W^2(u - f'/v_a) + |r_0|^2 \\ &+ 2|a|r_0 dW(u - f'/v_a)\cos[2\pi((f_0 - f')t + \Omega + \bar{\alpha}_u u]\Big] \qquad (5.2.6) \\ =& \mathrm{crect}\left[\frac{t}{\tau}\right] * \sum_n \delta(t - nT)\Big[|ad|^2 W^2(u - (k/T + \delta_f)/v_a) + |r_0|^2 \\ &+ 2|a|r_0 dW(u - (k/T + \delta_f)/v_a)\cos[2\pi((f_0 - k/T - \delta_f)t + \Omega + \bar{\alpha}_u u)]\Big]. \end{aligned}$$

The temporal modulation of the interferometric term thus reproduces the input sinusoid in amplitude and phase, and places this on spatial and temporal carriers which can be set to zero if desired. The fine frequency component δ_f will not move the blur spot $W(u)$ by a significant amount and can be dropped from within the argument of the space integrated blur spot $W(u)$. The coarse frequency component will be sampled by the laser diode pulses at times $t = nT$, resulting in an argument within the cosine function of $2\pi nk$, with n and k both integers, so that this term can be dropped leaving only the fine frequency temporal modulation. This is the key to the use of a repetitively pulsed laser source in this type of interferometer where it must act as an aliasing baseband converter. In this manner the distributed local oscillator reference used in the previous section can be replaced by the harmonic spectrum of the pulsed laser diode, because the pulsed laser accomplishes the same type of baseband conversion. However, there is an important difference between the two techniques, because the DLO spatially channelizes the available harmonics,

while all the harmonics of the pulsed laser are available everywhere for baseband conversion, so the sidelobes can choose the closest harmonic, and are not suppressed as they were with the DLO technique.

It often is desirable to use another AOD to shift the reference beam frequency by a harmonic of the reference coherent oscillator, so that it is also an integral multiple of the laser diode pulse repetition frequency $f_0 = \nu - \nu' = Lf_r = LL'/T$, where L and L' are both integers. In this way the temporal carrier term f_0 can also be dropped from within the argument of the cosine, because it will always be sampled by the laser pulses at integral multiples of 2π. The reference frequency should be chosen to be near the center of the signal AOD bandwidth in order to minimize the effect of the laser pulse width, which results in the laser pulse temporal spectrum weighing $\tau \mathrm{sinc}[m\tau/T]$, thereby allowing the utilization of the longest possible laser pulse $\tau < 2/B$. Thus for the nth laser diode pulse at time $t = nT$ the sampled output can be approximated by assuming an infinitesimal laser pulse width of amplitude $c\tau$, while explicitly including the amplitude of the laser diode spectral harmonic closest to the interferometrically generated temporal frequency.

$$
\begin{aligned}
I_{f'}(u,nT) &= \int_{(n-\frac{1}{2})T}^{(n+\frac{1}{2})T} I_{f'}(u,t)dt \qquad (5.2.7)\\
&\approx c\tau\Big[|a|^2W^2(u-k/Tv_a) + |r_0|^2 \\
&\quad + 2|a|r_0W(u-k/Tv_a)\cos[2\pi(\delta_f nT + \Omega - \bar{\alpha}_u u)]\mathrm{sinc}[(LL'-k)\tau/T]\Big]
\end{aligned}
$$

Thus for a single sinusoidal input the interferometric output of the space integrating spectrum analyzer illuminated with an impulse train can be seen to consist of a temporally sampled, interferometrically detected blur spot located at the coarse frequency locus k/T, and oscillating in sampled time at the fine frequency δ_f, with phase Ω, and amplitude proportional to $|a|$. The carrier fringe pattern at normalized spatial frequency $\bar{\alpha}_u$ will be seen to be moving from pulse to pulse at δ_f fringes per second under the envelope of $W(u - f'/v_a)$ if it has more than one spatial

oscillation within the width of the blur spot. Otherwise, the interferometrically detected blur spot will be seen to constructively and destructively beat on and off if $\bar{\alpha}_u \approx 0$ at a sampled rate of δ_f oscillations per second. The weighing of the sampled interferometrically generated term by the pulsed laser diode temporal spectrum sinc function can be kept to within 3dB by requiring that the maximum temporal frequency of the interferometrically generated temporal oscillations be less than the inverse of the laser pulse duration.

The effect of the sampling pulse train that is utilized in order to eliminate the coarse frequency component is usually referred to as aliasing, and is considered highly undesirable in sampling data aquisition systems. It indicates that the sampling rate is too low to adequately sample the given signal bandwidth, and it results in different parts of the signal spectrum folding over on top of each other, thereby producing invalid signal components. However, in this case the coarse frequency channelization of the SI spectrum analyzer effectively separates in space various frequency components that would otherwise alias with each other. It is still important to make sure that the sampling rate is fast enough to avoid aliasing of the bandlimited signals within each blur spot $W(u)$. The half power bandwidth of the signal information within each blur spot is increased over the acoustic transit time limited resolution of $1/T_a$ by the effects of apodization to approximately $2/T_a$, and the bandwidth to the first zeroes of the sinc function is $2/T_a$, so an appropriate blur spot bandwidth can be assumed to be approximately $2/T_a$. In order to adequately sample these bandlimited signal components the Nyquist criterion indicates that the sampling rate should be at least twice as fast as the highest frequency signal component that is present, thus the minimum sampling rate implemented by the pulsed laser should be at least $1/T > 4/T_a$. This indicates that the laser diode should be pulsed at least 4 times as an acoustic signal is in transit across the laser beam illuminating the AOD aperture, in order to adequately sample the fine

frequency information within each coarse frequency bin, and produce the appropriately oversampled space-time raster. If a double sided representation of the blur spot frequency content is used, then both positive and negative half bandwidths of $1/T_a$ will be within an apodized blur spot, which can be adequately sampled with a sampling rate of $1/T > 2/T_a$, indicating that the laser is pulsed twice per acoustic transit time, but positive and negative frequencies may be indiscernible in this case. Analysis of 2-D space integrating folded spectrum systems usually leads to the condition that the oversampling factor should equal 2, meaning that the duration of a raster line is twice the separation between lines. However, if an apodization were included in the fast dimension of a 2-D space integrating system in order to decrease the sidelobes in that dimension, then more oversampling would be required, just as in the time and space integrating implementation.

5.2.2 Time integrating interferometric DFT processor

In order to complete the folded spectrum processing the required operation that needs to be performed on the signal produced by the interferometric spectrum analyzer illuminated by a pulsed source is a narrowband time integrating discrete Fourier transformation (DFT) on each temporally sampled coarse frequency bin. The DFT is the discrete version of the Fourier transform as is appropriate for sampled data streams such as the interferometric term in Equation 5.2.7, which is a sampled fine frequency temporal oscillation. The TI DFT needs to be a baseband processor with analysis bandwidth equal to the blur spot width $2/T_a$, and it should process the full Nyquist limited bandwidth of half the laser pulsing frequency. The resolution of the time integrating fine frequency processor should ideally be equal to the inverse of the total processing time $1/NT$. The DFT kernel is space variant,

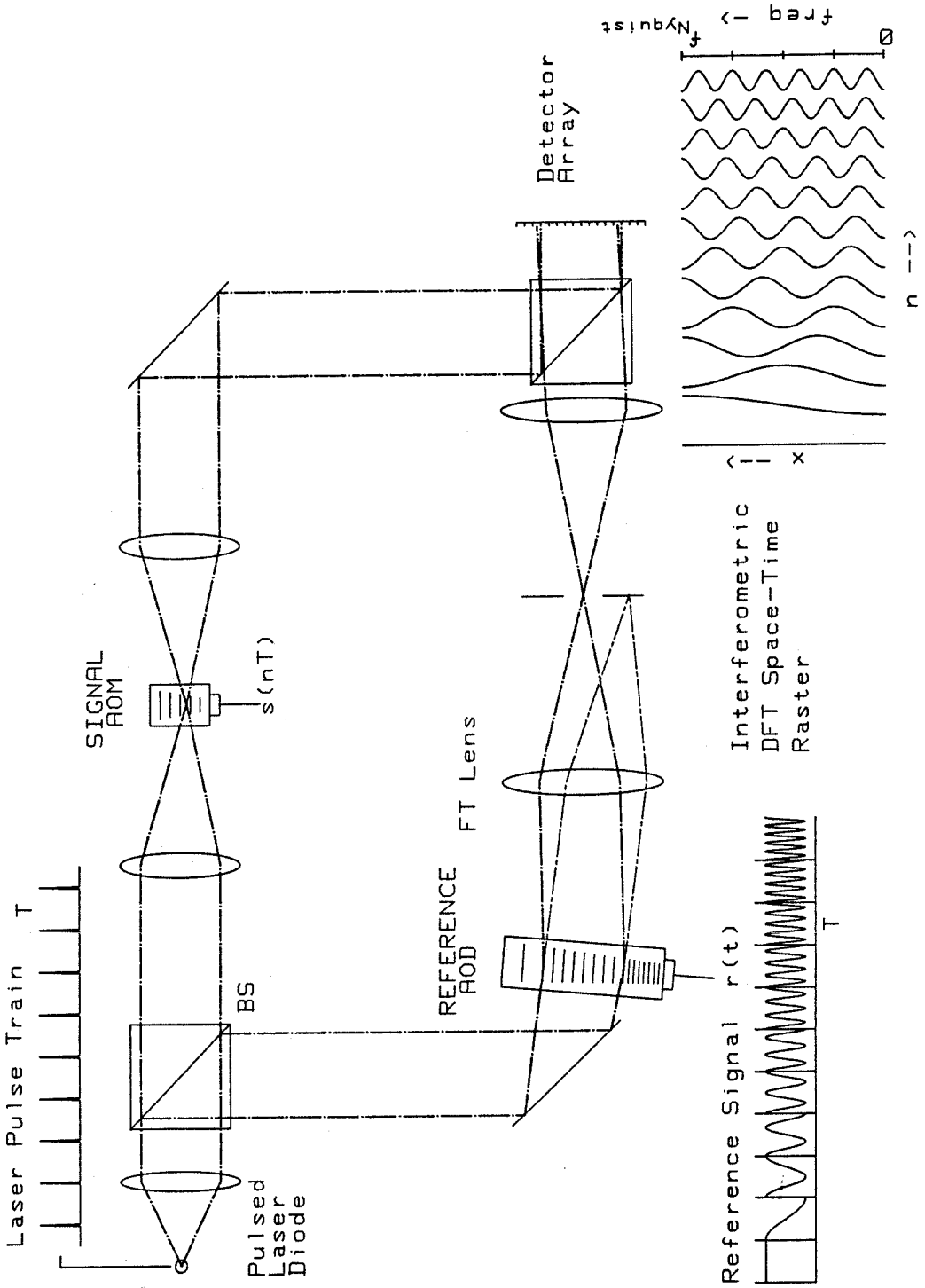


Figure 5.2.3. Time integrating interferometric DFT processor that uses the time between laser pulses to slide new matrix columns into the AOD.

and it can be implemented with a vector-matrix multiplication. The input vector $s(n)$ is multiplied by the DFT matrix producing the output vector $S(m)$.

$$S(m) = \sum_{n=0}^{N-1} s(n) e^{-i2\pi nm/N} \tag{5.2.8}$$

The pulsed source time integrating matrix vector multiplier described in section 2.3.2 can be utilized to implement the DFT operation since it utilizes the time between laser pulses to slide new columns of the DFT matrix into an AOD. A 1-D time integrating interferometric implementation of the DFT operation that utilizes a pulsed laser source is shown schematically in Figure 5.2.3. The light from a repetitively pulsed laser diode is split into two by a beamsplitter. One arm illuminates an acousto-optic modulator that has a narrowband signal $s(t)$ applied to its transducer, and the diffracted field is sampled by the narrow pulse of the laser diode producing a sampled temporal modulation $s(nT)$. In the other arm, a reference AOD with aperture T, equal to the laser pulse repetition interval, is schlieren imaged onto the output CCD array where it is interferometrically combined with the spatially uniform signal beam. With each laser pulse, a new column of the DFT matrix is applied to the reference AOD transducer, and propagates across the aperture to align with the CCD, at which time the acoustic motion is frozen by the laser pulse thereby imaging that column onto the CCD. Since the successive columns of the DFT matrix represent sinusoids of successively higher frequencies, the reference waveform is a stepped frequency chirp, with the constraint that each new frequency phase reset to zero phase with each new frequency segment. The reference signal starts at frequency f_l and frequency steps by Δf each time T.

$$r(t) = \sum_{n=0}^{N-1} g_n \text{rect}\left[\frac{t-(n+\frac{1}{2})T}{T}\right] \cos[2\pi(f_l + n\Delta f)(t-nT)] \tag{5.2.9}$$

Thus the reference waveform is analogous to a raster scanned version of the DFT matrix, except that along the columns it can use a continuous representation, which will be sampled later by the spatially discrete CCD pixel structure. An apodizing

function may be explicitly included in the reference signal by modulating the amplitude of the sequence of reference segments by some broad smooth function such as a Gaussian, $g_n = e^{-(n-N/2)^2/\sigma^2}$, or the pulsed laser power can be modulated with this function. Usually the uncontrollable analog apodizations, to be discussed later, will introduce enough sidelobe suppression, so that additional apodization will not be needed and g_n will be set to a constant for all n. This signal propagates in the y direction at the acoustic velocity v_a, and each rect window aligns with the AOD and CCD apertures at the times of the laser diode pulses as given in Equation 5.2.3. The spatial modulation diffracted by the AOD on the nth pulse is windowed by an aperture which will exactly match that imaged onto the CCD.

$$\begin{aligned} r(y,t) =& \text{rect}\left[\frac{y}{v_a T}\right] \tilde{r}(t-y/v)\delta(t-nT) * \sqrt{c\tau}\,\text{rect}\left[\frac{t}{\tau}\right] e^{-i2\pi\nu t} \\ r(y,n) \approx& \sqrt{c\tau}\,\text{rect}\left[\frac{y}{v_a T}\right] e^{i2\pi(f_l+n\Delta f)y/v_a} e^{-i2\pi\nu t} \end{aligned} \tag{5.2.10}$$

Thus on each laser pulse a finite aperture plane wave with an angular spatial frequency that is discretely stepped linearly with the pulse number n, is diffracted by the AOD and imaged onto the CCD. When this interferes with the collimated diffracted signal from the AOM a sinusoidal fringe pattern will be created, whose spatial frequency will increase linearly with the pulse number. The interferometrically detected intensity pattern on the nth laser pulse is found by combining weighted signal and reference beams. The signal beam is given an offset angle θ_y with respect to the starting frequency of the reference by rotating the beam splitter, which results in an additional spatial carrier of spatial frequency $\alpha_y = \sin\theta_y/\lambda$.

$$\begin{aligned} I(y,n) =& c\tau \left| a\tilde{s}(nT) e^{i2\pi(f_l/v-\alpha_y)y} e^{-i2\pi\nu t} + b\,\text{rect}\left[\frac{y}{vT}\right] e^{i2\pi(f_l+n\Delta f)y/v} e^{-i2\pi\nu t} \right|^2 \\ =& c\tau \left[|as(nT)|^2 + |b|^2 + 2ab|\tilde{s}(nT)| \cos[2\pi(n\Delta f/v + \alpha_y)y + \Omega(nT)] \right] \end{aligned} \tag{5.2.11}$$

The stepped spatial frequency fringe pattern has an amplitude and phase given by the sampled signal waveform. The apparent spatial velocity of the AOD at the CCD

plane is given by the magnified acoustic velocity $v = mv_a$. By time integrating over N successive laser pulses a discrete summation of the interferometrically detected, photogenerated charge distribution due to each pulse is obtained.

$$I(y) = c\tau \left[N|b|^2 + \sum_{n=0}^{N-1} |as(nT)|^2 + 2ab \sum_{n=0}^{N-1} |\tilde{s}(nT)| \cos[2\pi(n\Delta f/v + \alpha_y)y + \Omega(nT)] \right] \quad (5.2.12)$$

This can be recognized as a sum of bias terms plus the discrete cosine transform of the complex sequence $\tilde{s}(nT)$, on a spatial carrier at the midband spatial frequency $\gamma = \alpha_y + (N-1)\Delta f/2v_a$. The carrier allows the utilization of a bandpass filter operation to remove the bias terms, and the full complex DFT can be reconstructed by measuring the amplitude and phase angle of the modulated carrier for each resolvable frequency component. The achievable resolution of this operation is $1/NT$, and the best single sided analysis bandwidth is given by the Nyquist limit $1/2T$, which is achieved when an additional half cycle of the reference waveform is added with each frequency step $\Delta fT = 1/2$. Thus $N/2$ independent frequency measurements can be made per sideband of processing. In order to accomplish double sided processing the time of phase resetting should be in the middle of each frequency stepped sinusoid, and the frequency step interval should be doubled so that $\Delta fT = 1$, if double sided Nyquist limited analysis bandwidth is desired. This can best be visualized by examining the temporal history of each position on the integrating CCD, and noting that the reference induces an interferometric, sampled temporal oscillation whose frequency depends linearly on position along the detector array. Interferometric detection with a single tone of the signal will mix the signal and reference frequencies producing a low frequency difference and an unobservable sum frequency at twice the optical carrier. A peak will only be integrated up when the signal and reference phases oscillate in synchronism so that the difference frequency produces less than half an oscillation during the entire period of integration over all N laser pulses. Thus each integration position can be considered to be a heterodyned integrating

radiometer, with a heterodyne center frequency that changes linearly with position along the output photodetector array. At the detector position corresponding to the phase resetting of the stepped frequency sinusoid, the interferometrically generated oscillation frequency induced by the reference is at DC and a peak will be integrated up if the signal $s(nT)$ contains a DC component with no temporal phase oscillation. At the Nyquist limited high frequency end of the detector array an extra half cycle is added to the reference waveform each segment, so the reference generates a temporal oscillation which beats from high to low on alternate laser pulses. When the signal contains a matched Nyquist frequency term, then a DC component difference frequency will be generated and a peak will be produced by the time integrating detector at the high frequency edge of the photodetector array. Adding an extra half cycle of the reference each frequency step produces the desired alternating oscillation at the high frequency position at the edge of the detector, but a larger frequency increment Δf will result in aliasing due to the laser sampling and will not produce a larger analysis bandwidth.

The CCD is composed of an array of evenly spaced, discrete, time integrating pixels that must have sufficient resolution to adequately sample the spatial fringe structure under the DFT transform, and each of the integrated spatial sinusoidal patterns. The maximum spatial frequency that is incident on the CCD is given by $\gamma_{max} = (N-1)\Delta f/v + \alpha_y$, and the minimum spatial frequency is α_y, so the DFT carrier frequency is given by the midband spatial frequency γ. For a pixel separation Δy, the number of pixels per cycle of the maximum spatial frequency sinusoid is $1/\gamma_{max}\Delta y$, and the minimum number of pixels required to sample this spatial frequency is limited by the Nyquist rate to at least 2 pixels. The number of pixels per cycle of the DFT carrier fringe structure is $1/\gamma\Delta y$, and should be adjusted to about 4 to 8 sampling pixels per cycle in order to facilitate phase measurements. The total number of CCD pixels required is the product of the

DFT space bandwidth product with the number of pixels per frequency resolution element $N_{CCD} = N\Delta fT/\gamma\Delta y > 4N\Delta fT$, which must be at least one cycle and hopefully more in order to improve the frequency definition required by the bandpass filtering operation in order to adequately reject the bias. When double sided Nyquist limited resolution is desired, then $\Delta fT = 1$, so the number of CCD pixels required is given by the number of reference DFT matrix columns that are illuminated by the laser pulses, times the number of pixels per frequency resolution element, so that $N_{CCD} > 4N$.

5.2.3 Pulsed source TSI folded spectrum processor architectures

The characteristics of the interferometric TI DFT processor are complementary to that of the SI AO spectrum analyzer illuminated by a pulsed source. The TI DFT processor is a narrow band processor with analysis bandwidth $\Delta f < 1/2T$ per sideband, and high resolution of $1/NT$. It requires pulsed source illumination in order to accomplish the sequential loading of successive DFT columns into the travelling wave reference AOD aperture. The pulsed SI AO spectrum analyzer is a broadband processor with bandwidth equal to the AOD bandwidth B or inverse laser pulse width whichever is smaller, and apodized resolution of approximately $2/T_A$. It converts the interferometrically produced fine frequency modulations within each coarse spectral bin to baseband by the aliasing due to the sampling laser pulse train. The complete folded spectrum operation can be obtained by performing the interferometric time integrating DFT on each coarse frequency resolution element produced by the SI AO spectrum analyzer. Since the TI DFT processor is a purely 1-D system, an array of such processors can be multiplexed in the orthogonal dimension to the SI spectrum analyzer by using a 2-D CCD detector array. Each of the TI DFT processors in the multiplexed array needs to perform the same narrowband spectrum analysis operation, and each coarse frequency signal is brought

down to baseband by the aliased sampling of the pulsed laser diode, so they all can use the same reference AOD.

There are several ways that an optical system can be configured in order to perform the desired 2-D acousto-optic folded spectrum operation. The additive Mach-Zehnder interferometer shown in Figure 5.2.4 will be concentrated on, because this was the preferred implementation for the experimental demonstration of these system concepts. It consists of a pulsed source space integrating acousto-optic spectrum analyzer to perform coarse frequency channelization in the x dimension of one arm of the interferometer, and a spatially multiplexed array of time integrating DFT processors in the other arm. The DFT processors perform fine frequency analysis in the y dimension upon the interferometrically generated, baseband, sampled, fine frequency information within each coarse frequency bin. The separated path interferometer was chosen because of the added freedom that this allows in the choice of lens parameters to perform the various imaging and Fourier transformation operations that are necessary. The disadvantages of this separated path Mach-Zehnder configuration are its sensitivity to air motions and component vibrations, and the necessity of precisely aligning the path lengths of the two arms of the interferometer to obtain optimum pulsed laser coherence.

An alternative interferometric implementation is an in line interferometer shown in two optical designs in Figure 5.2.5. In the top implementation the pulsed laser is cylindrically focussed into the signal AOD first and the output field is Fourier transformed by spherical lens F_2, which also recollimates in y. The diffracted component in the Fourier plane is allowed to pass beside the orthogonal reference AOD acoustic column, while the undiffracted component from the signal AOD is diffracted by the reference AOD. The output field from the reference AOD, consisting of both the Fourier transform of the signal AOD and the diffraction from the reference AOD is imaged onto the detector array with the two lens imaging system, while

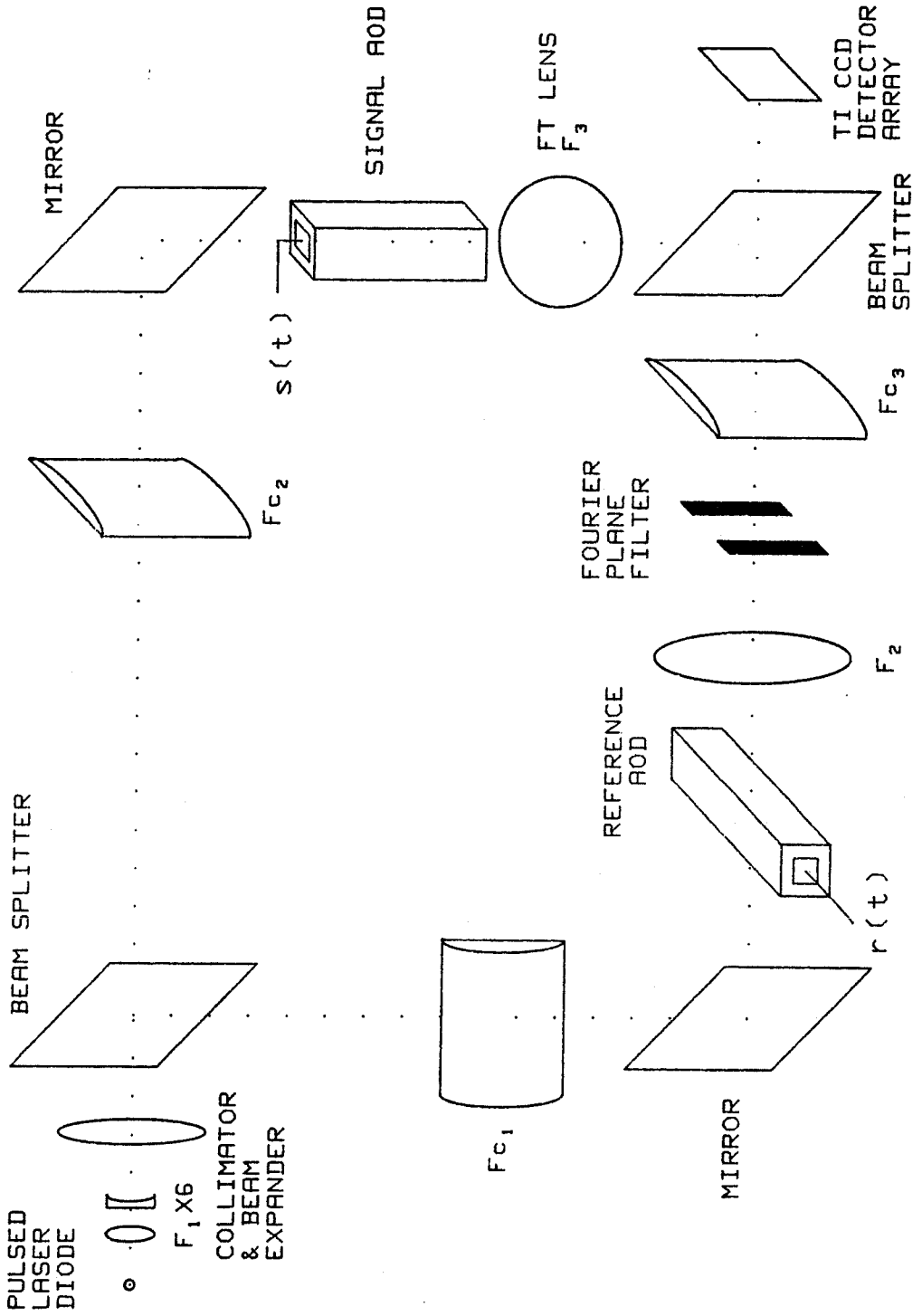


Figure 5.2.4. Mach-Zehnder interferometer configuration for performing the time and space integrating folded spectrum processing operation.

the undiffracted term is blocked in the Fourier plane of lens F_3. In the bottom implementation the collimated pulsed laser is first focussed into the vertical reference AOD, which is Fourier transformed into the horizontal signal AOD. The undiffracted component is diffracted by the signal AOD and both singly diffracted components are Fourier transformed onto the detector, thereby reimaging the reference AOD vertically and simultaneously transforming the signal horizontally and collimating vertically. In either system the interferometric product between the signal and reference beams will produce a spatial carrier along a diagonal, which is inconvenient for practical spatial carrier demodulation. However these in line interferometers have the advantage that the optical path lengths are automatically closely matched, and component vibrations introduce smaller unwanted modulations since both beams traverse the same optical path.

Another option for the implementation of these system is to use a multiplicative in line approach using the same system and optics as either system shown in Figure 5.2.5, with a slight optical modification, and additional reference signals added to the signal and reference waveforms. This approach would involve moving the second AOD so that the Fourier transform of the diffracted order from the first AOD is multiplicatively rediffracted by the second AOD, and the doubly diffracted term is imaged onto the detector, while the undiffracted and singly diffracted terms are blocked. In order to perform the desired interferometric detection operations of both signal and reference, an additional coherent reference tone must be added to the reference waveform, and a DLO reference signal must be added to the signal in order to accomplish interferometric detection within any coarse frequency bin. The DLO is not intended to beat each coarse frequency bin to baseband, which is accomplished instead by the laser pulse induced aliasing. The DLO field at each coarse frequency bin is needed to act as a reference beam for interferometric detection, but since its frequency changes with position the sampling laser pulse could

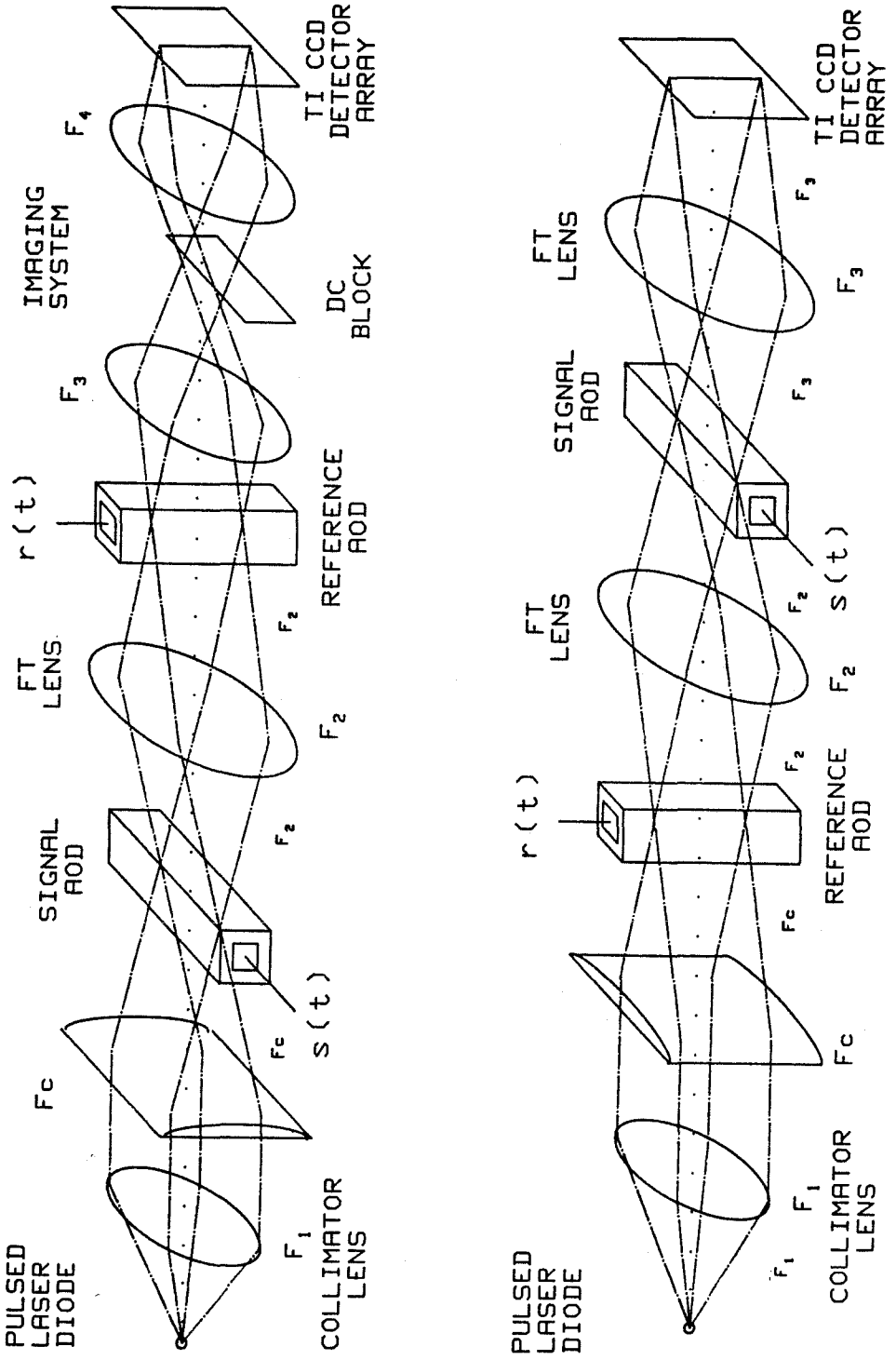


Figure 5.2.5. In line approaches to TSI folded spectrum processing that can be configured as either additive or multiplicative systems.

be lengthened without decreasing the coarse analysis bandwidth. This approach suffers from the light inefficiency of the multiple diffraction operation, but this may be offset by the allowance of longer duration pulses, and it limits the signal bandwidth to the scaled height of the reference AOD transducer. The advantage of the multiplicative approach is its almost complete insensitivity to component vibrations, low coherence length requirements, and the ability to introduce a carrier frequency in the x or y dimension separately. However, the multiplicative approach is essentially an intensity modulated system and introduces an additional conjugate sideband contribution that will produce strong sidelobes that will be devastating in a multitone environment, unless a carrier modulated encoding technique is used to eliminate these unwanted contributions along with the bias.

One other optical architecture that should be considered because of its optical simplicity and ease of electronic reference generation is shown in Figure 5.2.6. In this folded spectrum system a dual axis acousto-optic device with wide aperture transducers is illuminated at the Bragg angle in both x and y by a collimated plane wave from a repetitively pulsed laser diode. The signal is applied to the x transducer of the AOD, and a reference waveform consisting of set of narrow pulses of a slightly different repetition frequency from the laser pulse train is applied to the orthogonal y transducer. With each laser pulse, separated by a period T, the wide aperture reference pulses of separation $T + \Delta T$ will have moved farther down the dual axis deflector than the previous pulse by a distance $v_a \Delta T$, and it will be frozen at that position by the narrow laser pulse. The multiply diffracted field from the dual axis AOD is Fourier transformed in both x and y by the following spherical lens and detected on the 2-D detector array. With each laser pulse a portion of the signal will have been transduced into the optical system in the x dimension, but the y position of the multiplicatively diffracted wave will be a linear function of the laser pulse number. At the detector plane the signal will be Fourier transformed producing

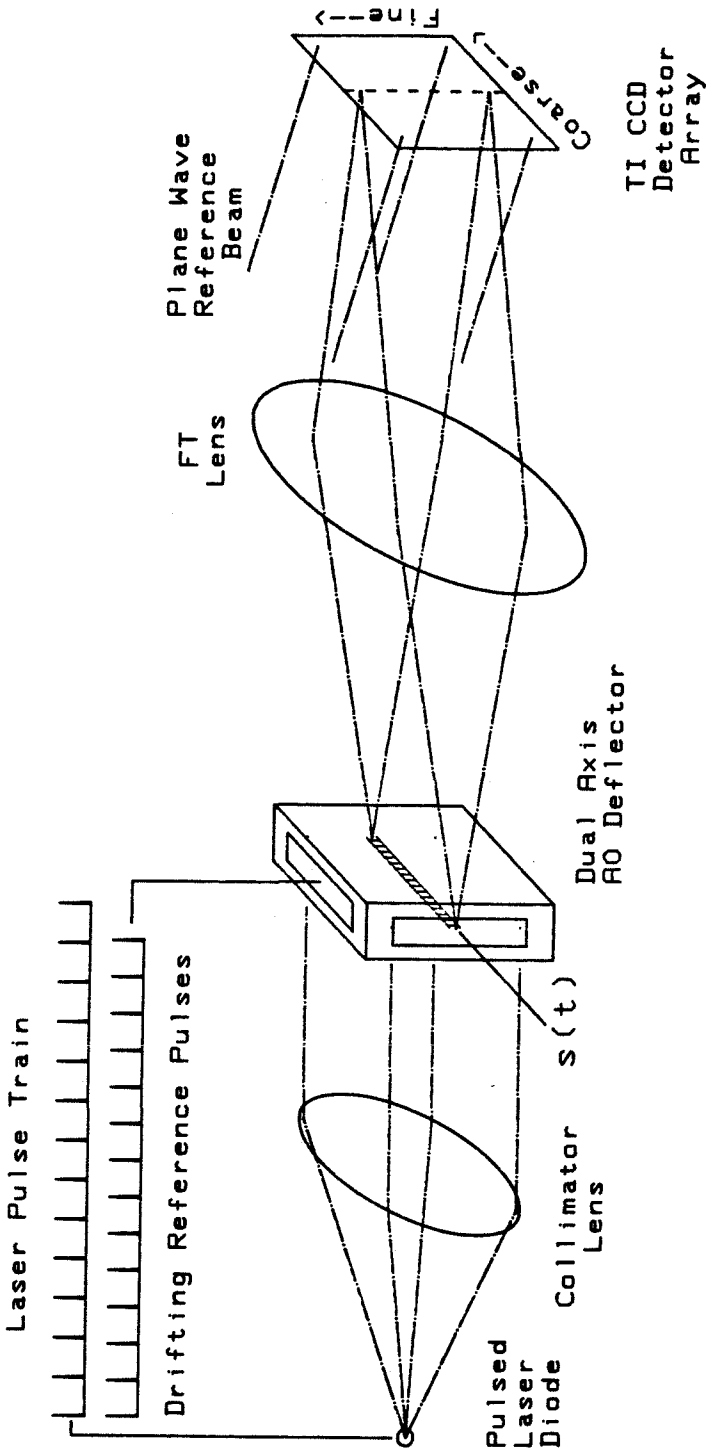


Figure 5.2.6. Single AOD in the Fourier plane approach to TSI folded spectrum processing that can also be configured as either an additive or a multiplicative systems.

a coarse frequency ridge whose angular spatial frequency is a linear function of the pulse number, as is desired in a DFT type of fine frequency spectrum analysis system. Interferometric detection and time integration will produce the desired folded spectrum, and this can be accomplished with an auxiliary reference wave or by adding extra reference signals to the two AOD transducers. The low light efficiency of this system make it somewhat impractical, although the light efficiency can be improved by using a chirp reference that is recollimated by an additional cylindrical lens. This system is mentioned primarily because it shows the connection between 2-D space integrating folded spectrum and TSI folded spectrum based on the DFT. This is made clear by thinking of the interferometric time integral of the image diffracted by this dual axis modulator, which would be a spatial raster of the input signal. Therefore the time integral of the interferometric Fourier transform of the multiply diffracted wave will be the same as the Fourier transform of the spatial raster, which will result in the folded spectrum.

5.2.4 Mach-Zehnder interferometer for TSI folded spectrum processor

The description and analysis of the additive Mach-Zehnder interferometer designed to perform the folded spectrum operation is made with reference to Figure 5.2.4. This figure approximately corresponds with the experimental apparatus which was used to demonstrate folded spectrum processing, and contains two orthogonal AODs referred to as signal and reference corresponding to the applied electrical waveform. The signal AOD is oriented vertically and the reference AOD is oriented horizontally, however for analytic convenience and continuity the signal AOD will always be referred to as x propagating and the reference AOD as y propagating. The timing of the system is controlled by a stable coherent crystal oscillator reference of frequency f_r. The oscillator frequency is digitally divided by an integer L' in order to produce a stable pulse train of frequency $f_{LD} = f_r/L' = 1/T$ to be

used as the laser diode trigger. The laser is biased just below threshold, and pulsed with waveform $g(t)$, with a pulse of width τ each period T, while the bias, pulse amplitude c, and rising and falling edge durations of the pulse are adjusted to optimize the coherence both between and within the laser pulses. The laser diode emits light of frequency ν and wavelength λ in the near infrared, and it is critical that the laser wavelength does not vary during or between pulses. The anamorphic Gaussian beam profile emitted by the laser diode is collimated by a short focal length lens $F_1 = 14.8mm$, and beam expanded by 6 times in order to produce a large aperture well collimated coherent plane wave with separable Gaussian beam amplitude profile $q(x)p(y)$. The collimated beam is incident upon a beam splitter at a 45° angle, sending a fraction $\sqrt{1-\xi}$ of the incident field into the signal arm, and a fraction $\sqrt{\xi}$ to the reference arm. In the space integrating signal arm a cylindrical lens with curvature in y and focal length $F_{c2} = 300mm$, focuses the light horizontally as it reflects off the mirror to a narrow vertical slit that is incident upon the vertical signal AOD at the midband Bragg angle. The signal $s(t)$ applied to the AOD transducer produces a travelling wave propagating in the x (vertical) dimension, which produces a diffracted component that is spatially Fourier transformed by the following spherical lens with focal length $F_3 = 250mm$. The separation between the cylinder and sphere is given by the sum of the respective focal lengths, so that the spherical lens also recollimates the light in the y dimension with a magnification given by the focal length ratio $M_y = -F_3/F_{c2} = 5/6$. The diffracted field is incident upon another beam splitter that reflects a portion $\sqrt{\chi}$ towards the time integrating CCD detector array that is placed in the focal plane of the Fourier transforming lens. The output field can be represented in the standard form in terms of the single sided, negative frequency, within band analytic signal spectrum $\tilde{S}(f)$, and the spatial Fourier transform of the hybrid blur function $w(x) = \text{rect}[x/A]q(x)$.

$$U(u,y,t) = d_s\sqrt{1-\xi}\sqrt{\chi}p(M_y y)g(t)\int \tilde{S}(f)e^{i2\pi ft}W(u+f/v_a)df e^{-i2\pi\nu t}$$

$$=c_s p(M_y y) g(t) e^{-i2\pi\nu t}[\tilde{S}(uv_a)e^{i2\pi v_a u t}] * W(uv_a) \qquad (5.2.13)$$

The constant $c_s = d_s\sqrt{1-\xi}\sqrt{\chi}$ has absorbed all of the multiplicative constants and arbitrary phase factors have been set to zero. Each frequency component of $s(t)$ produces a horizontal slit of light at the output plane during each laser pulse, with a field that is weighted along its length by $p(M_y y)$ and with a spatial width given by the blur spot $W(u)$.

In the time integrating reference arm of the interferometer a cylindrical lens with vertical focussing power and focal length $F_{c1} = 300mm$ collapses the light as it reflects off the mirror, and this is incident on the reference AOD as a narrow horizontal slit at the Bragg angle. Only a portion of the AOD corresponding to the length of one reference segment of duration T needs to be uniformly illuminated by the horizontal beam profile $p(y)$. The signal AOD should be completely illuminated over the entire width $A = v_a T_a$, while the reference AOD only requires illumination over the width $X = v_a' T$, and the laser will be pulsed several times during the full acoustic transit time because $T < T_a$. For this reason the laser should be oriented with its wider angle of diffraction in the vertical dimension in order to illuminate more of the signal AOD aperture. The additional complication of different acoustic velocities will not be considered here, but it represents another possible degree of freedom in system design which might be required if a circular beam profile was to be optimally utilized. The stepped frequency reference signal is applied to the reference AOD transducer, and launches an acoustic wave in the y (horizontal) dimension with velocity v_a. The AOD is positioned so that one segment of the reference will be imaged onto the CCD horizontally, and so the focusing horizontal slit is optimally aligned with the acoustic column. The diffracted reference wavefront is spatially Fourier transformed by the spherical lens with focal length $F_2 = 120mm$, and this lens is separated from the focusing cylinder by the sum of the focal lengths, so the light is recollimated vertically with magnification $m_x = -F_2/F_{c1}$. At a distance

F_2 behind the spherical lens will be the transform plane of the AOD, and in this plane a vertical slit of diffracted light will be observed to hop horizontally by a distance $\delta y = \Delta f \lambda F_2 / v_a$ with each laser diode pulse. An adjustable aperture slit will be placed in the Fourier plane to facilitate DC blocking and reference band pass filtering. The filtered light is retransformed in the y dimension by the following cylinder of focal length $F_{c3} = 150mm$, that is separated from F_2 by the sum of the focal lengths, to form an inverting telescopic imaging system in y with magnification $m_y = -F_{c3}/F_2 = 1.25$. This auto collimating imaging system turns the finite aperture slit plane waves diffracted by the AOD back into plane waves in y while simultaneously collimating them in x. The magnification of this imaging system is chosen in order to image the portion of the AOD corresponding to one reference segment $X = v_a T$, onto the full available height of the CCD $H = vT$, so $m_y = -H/X$. The final beam splitter passes a fraction $\sqrt{1-\chi}$ of the reference field, reflecting the rest and is used to carefully align the signal and reference wavefronts. The alignment procedure to be discussed in the experimental section ensures that the image plane of the reference AOD coincides with the Fourier plane of the signal AOD, but this is not really necessary since plane waves diffracted by the reference AOD will be plane waves at all locations beyond the retransforming cylinder. The laser diode is pulsed synchronously with the reference waveform so that each segment is imaged onto the CCD, and the pulse width τ is narrow enough that it essentially freezes the acoustic motion, producing a spatial image of each reference segment on the CCD during each pulse.

$$\begin{aligned} r(x,y,t) &= d_r \sqrt{\xi}\sqrt{1-\chi}\, q(m_x x) p(m_y y) \tilde{r}(t - m_y y/v) g(t) \\ r(x,y,n) &= c_r q(m_x x) p(m_y y) e^{i2\pi(f_l + n\Delta f) m_y y / v_a} \end{aligned} \tag{5.2.14}$$

A constant $c_r = d_r \sqrt{\xi}\sqrt{1-\chi}$ has been defined that absorbs the appropriate multiplicative terms, and the phase factors are arbitrarily set to zero. The signal and

reference beams are recombined with an adjustable angle between them, and the interferometric products produced on each laser pulse are accumulated on a time integrating CCD detector array.

The optical field incident on the CCD during the nth pulse is the sum of the reference wave, which acts as a linearly stepped spatial frequency modulation in the y dimension, with the spatially Fourier transformed signal that is currently within the signal AOD aperture. The signal beam recombines with the reference beam at an angle $\theta_y = \sin^{-1}[\alpha_y \lambda]$ controlled by the beam splitter rotation, that is referenced to the magnified starting spatial frequency diffracted by the reference AOD, resulting in a spatial carrier in the y dimension with spatial frequency α_y. The photogenerated charge profile detected by the CCD on each laser diode pulse is the time integral of the modulus square of the field which is incident on the CCD of aperture H during that pulse.

$$
\begin{aligned}
I(x,y,t) = \int_{(n-\frac{1}{2})T}^{(n+\frac{1}{2})T} & \operatorname{crect}\left[\frac{t}{\tau}\right] * \sum_n \delta(t-nT)\operatorname{rect}\left[\tfrac{y}{H}\right] \Big| c_r q(m_x x) p(m_y y) \qquad (5.2.15)\\
& \operatorname{rect}\left[\tfrac{t-(n+\frac{1}{2})T-m_y y/v_a}{T}\right] e^{-i2\pi(f_l+n\Delta f)[t-(n+\frac{1}{2})T-m_y y/v_a]} e^{-i2\pi\nu t} \\
& + c_s p(M_y y) e^{i2\pi(m_y f_l/v_a-\alpha)y} e^{-i2\pi\nu t} \int \tilde{S}(f) e^{i2\pi f t} W\left(\frac{x}{\lambda F}+\frac{f}{v_a}\right) df \Big|^2 dt \\
= & c\tau c_r^2 q^2(m_x x) p^2(m_y y) + c\tau c_s^2 p^2(M_y y) |\tilde{S}(v_a x/\lambda F) * W(v_a x/\lambda F)|^2 \\
& + \eta 2 c_s c_r q(m_x x) p(m_y y) p(M_y y) \int_{(n-\frac{1}{2})T}^{(n+\frac{1}{2})T} c \sum_m \tau \operatorname{sinc}[m\tau/T] \\
& \int \cos\left[2\pi\left[ft - \frac{mt}{T} + (f_l+n\Delta f)(t-(n+\tfrac{1}{2})T) - \left(\frac{n\Delta f m_y}{v_a} - \alpha_y\right) y\right] + \Omega(f)\right] \\
& |\tilde{S}(f)| W\left(\frac{x}{\lambda F}+\frac{f}{v_a}\right) df\, dt \\
= & c\tau c_r^2 q^2(m_x x) p^2(m_y y) + c\tau c_s^2 p^2(M_y y) |\tilde{S}(v_a x/\lambda F) * W(v_a x/\lambda F)|^2 \\
& + \eta c_s c_r q(m_x x) p(m_y y) p(M_y y) c \sum_m \tau \operatorname{sinc}[m\tau/T] \\
& \int \operatorname{sinc}[T(f+(f_l+n\Delta f) - m/T)] |\tilde{S}(f)| W\left(\frac{x}{\lambda F}+\frac{f}{v_a}\right)
\end{aligned}
$$

$$2\cos\left[2\pi[(f-\tfrac{m}{T})nT-(f_l+n\Delta f)\tfrac{T}{2}-(\frac{n\Delta f}{v}-\alpha_y)y+\Omega(f)]\right]df$$

As usual this pattern consists of an approximately uniform bias, a signal dependent bias ridge and the desired interferometrically detected term, which in this case consists of the coarse frequency channelized spectral information multiplied by a spatial cosine in the orthogonal direction. The modulation depth of the interferometric term is decreased from perfect coherence by the factor η, which includes effects due to path differences, vibrations, and pulsed laser diode intra- and interpulse mode hopping. The reference window function is synchronously timed with the laser diode pulses and the reference AOD is translated so that the magnified image of the desired reference segment exactly aligns with the CCD aperture at the time of the laser diode pulse. The amplitude and phase of each component of the signal spectrum are reproduced as modulations of the amplitude and phase of the detected fringe pattern in the y dimension on each pulse. The important information is in the terms within the argument of the cosine that exhibit sampled temporal oscillation dependent on the pulse index n. The first term in the argument of the cosine $(f-m/T)nT$ is the aliased to baseband component due to each input frequency, and is due to the periodic sampling by the laser diode of the interferometrically detected signal as it continuously drifts in and out of phase with the reference waveform at each position in space. The next term $(f_l+n\Delta f)T/2$ contains an arbitrary phase factor due to the starting frequency phase at the middle of the first reference segment that can be dropped along with an oscillation that is due to the definition of the reference segments with centroids at $nT/2$. The final term $(n\Delta f/v-\alpha_y)y$ contains the spatial fringe pattern that increases its spatial frequency with each laser pulse, on a carrier that remains fixed, and it is this term that results in the desired fine frequency processing. It is very similar to a distributed local oscillator, in the sense that the sampled temporal frequency $\Delta f y/v$

is a linear function of space, and the spatial frequency $n\Delta f/v$ is a linear function of the pulse index. This temporally sampled distributed local oscillator has a sampled frequency of temporal oscillation that increases from $f_{dlo} = 0$ at the bottom of the CCD, to the maximum frequency represented, $f_{dlo} = \Delta f$ at the top of the CCD, that is given by $f_{dlo}(y) = \Delta f(T/2 + y/v) = \Delta f(\frac{1}{2} + y/H)$. At the bottom of the CCD $y = -H/2$, and the sampled phase oscillation $n\Delta fT/2$ cancels with the spatial term $n\Delta fy/v$ leaving no oscillations induced by the reference signal. The equation can be simplified by changing to an alternative spatial coordinate system on the CCD $y' = y + H/2$, so the sampled temporal oscillation will then be given as a function of the new spatial variable as $f_{dlo}(y') = \Delta fy'/H$. Each input frequency will produce within its coarse blur spot a position at which its sampled fine frequency of oscillation matches closely with that of the sampled distributed local oscillator, and it is at this position of stationary phase that a peak will be integrated up, thereby locating the fine frequency of oscillation.

The first sinc function in the above equation describes the limited bandwidth of analysis due to the limitation imposed by the finite laser pulse duration τ. The second sinc function selects the laser diode spectral harmonic which is closest to the difference frequency of the input signal and the reference frequency during the nth segment of the reference waveform, as the primary term responsible for producing a baseband time integrated contribution. Thus the particular pulsed laser spectral harmonic which is responsible for producing the detected contribution is a function of the stepped reference frequency, and since the first sinc function weights the spectral harmonics, a modulation depth of the cosinusoidal fringes will be observed that is a sinc function of width $1/\tau$ of the difference frequency between signal and instantaneous reference. As the difference frequency between a particular reference segment and the input frequency varies away from a multiple of the laser diode pulse repetition frequency, the second sinc function introduces another amplitude

modulation of the fringe pattern. However since the sinc functions at the multiples of the laser diode pulse repetition frequency have a width equal to the laser diode pulse rate the nearest sinc will never have an amplitude less than $\mathrm{sinc}(.5) = .64$. When the reference frequency step increment is given by $\Delta f = K/K'T$, with $K < K'$ and both integers with no common divisor, then the difference frequency will drift through the equally spaced array of sincs in a manner like a vernier, that will repeat each K' laser pulses thereby inducing an amplitude modulation of the interferometrically detected fringe pattern with a principal frequency components at $1/K'T$ and its harmonics. When the input signal has a fine frequency of oscillation that is a multiple of one of these harmonics, then a peak can be integrated up at all locations within an interferometrically detected coarse frequency bin, and if a spatial carrier is being utilized then this carrier will run the length of the ridge. This problem can be eliminated by making the reference frequency step increment equal to the laser diode pulse repetition frequency. In this case as the fine frequency of oscillation of each input temporal frequency moves away from an exact multiple of the laser diode PRF the second sinc function produces an amplitude modulation of the cosinusoidal fringes that is the same for all the laser pulses. This will produce a weighing of the transform that will suppress the large fine frequency terms by as much as .64 when the fine frequency offset is half the laser pulse repetition rate. The major effect of this sinc function will be to select the appropriate laser harmonic needed to convert the interferometrically generated difference frequency between the signal and reference to baseband, so that $f + f_l + n\Delta f - \frac{m}{T} = (f + f_l + n\Delta f) \mathrm{mod} \frac{1}{T}$. When the frequency step size is chosen equal to the laser diode pulse repetition frequency, and the first stepped frequency to be an integral multiple of this frequency as well then the second sinc function causes the first terms in the cosine argument to take the especially simple form $f \mathrm{mod} \frac{1}{T}$, and it introduces a constant amplitude modulation of the fine frequency $\mathrm{sinc}[T(f \mathrm{mod} \frac{1}{T})]$.

The importance of the phase resetting of the reference waveform is demonstrated by the cancellation of the quadratic term in n, within the argument of the cosine that occurs when nT is substituted for t as a result of the time integration over the nth pulse. If the reference waveform was not phase reset, then the coefficient of the quadratic term would be required to always be an integral multiple of 2π, so that it could be dropped from within the argument of the cosine, and this leads to the condition ΔfT =integer. However, for Nyquist limited analysis an extra half cycle of reference should be added per sideband with each frequency step, so that $\Delta fT = 1/2$ per sideband. Thus if phase resetting the stepped frequency reference waveform is not used, then the processor is limited to performing Nyquist limited double sided processing that utilizes phase continuous frequency stepping. This type of phase continuous frequency stepping is very difficult to perform over the bandwidths necessary for the reference signal. Many frequency hop techniques require a dead time during frequency switching, so phase continuity is impossible, but all that is really required is the ability to phase reset each new reference tone coherently with the laser diode trigger. For this reason phase resetting of the reference waveform with each frequency step should be performed. This eliminates any conditions on the frequency step size, and allows the time integrated DFT to be arbitrarily scaled by changing Δf, which is very convenient for implementing spectral zooms of the coarse frequency bins, and testing the folded spectrum analysis technique with lower bandwidth reference waveforms.

During the period τ that the laser diode is emitting light, the spatial fringe pattern can move slightly on the detector array in the y dimension, and the time integrated charge profile which is accumulated is of the form of a convolution between the laser pulse shape $g(t)$, and the cosinusoidal spatial fringe profile in y which is moving at a magnified velocity $v = v_a/m_y$. It is required that τ be chosen small enough, so that during the duration of the laser diode pulse the highest spatial

frequency incident on the detector for all n, will move by only a small portion of a spatial fringe. The highest spatial frequency will occur when the reference beam and the signal beam are at the largest incidence angle, and this can occur for either $n = 0$ or $n = N - 1$, depending on the beam splitter rotation angle. If the signal beam is aligned with the midband reference beam angle in y, then there will be no spatial carrier, and a two sided transformation will be implemented, with analysis of both positive and negative fine frequencies, and the maximum spatial frequency will occur at both $n = 0$ and $n = N - 1$. If the signal beam is aligned near the ending reference frequency, then the interferometric fringe pattern will start at a high spatial frequency, and from pulse to pulse the spatial frequency will decrease towards zero, and a negative fine frequency spectral analysis will be implemented. When the signal beam is aligned closer to the beginning reference beam angle, then the fringe pattern will begin at a low spatial frequency, and will step to higher spatial frequencies on each pulse to a maximum of $\gamma_{max} = (N-1)\Delta f/v + \alpha_y$, and a positive fine frequency spectral analysis will result. The motion of this fringe pattern during the laser diode pulse should be less than one half of a fringe of the maximum spatial frequency, $v\tau < 1/2\gamma_{max}$, so that $\tau < m_y/2v_a\gamma_{max}$. If no additional spatial carrier is introduced, $\alpha_y = 0$, then this limits the pulse duration to $\tau < 1/2N\Delta f$, so the laser is on for a maximum of half the inverse bandwidth of the reference waveform. When the beamsplitter is rotated in order to place the DFT on a spatial carrier then the pulse duration must be even less than the inverse reference waveform bandwidth. Recalling that the spatial pixel sampling is limited by the Nyquist rate to at least 2 pixels of width Δy per cycle of the maximum spatial frequency γ_{max}, one arrives at an alternate condition. This states that the pulse duration must be less than the time it takes for the image of the reference acoustic wave to propagate through a distance of one pixel $\tau < \Delta y/v$. This is an auxiliary condition to the previously arrived at relation for the laser diode pulse

duration, where it was determined that in order to have spectral harmonics of sufficient amplitude over the full signal bandwidth to act as local oscillators it is required that the laser pulse is shorter than the inverse of half the signal AOD bandwidth, $\tau < 2/B_{sig}$. Normally, it is desired that the highest resolution possible is obtained, which requires long integration and wide bandwidth of the reference waveform, so if the full available octave bandwidths of both signal and reference AODs are used, and the transform is placed on a high frequency spatial carrier to facilitate bias removal, then the auxiliary condition on the pulse duration limited by the imaged reference acoustic wave pixel transit time, is more stringent. Notice that the time integration of these moving fringe patterns will suppress the effective modulation depth of the higher spatial frequency fringes, and at the limiting duration of τ indicated above the highest spatial frequency fringe will move by half a fringe during its period of integration, and thus produce no modulation, thereby introducing a strong apodization of the time integrating discrete Fourier transformation. The reference AOD will have a Bragg regime frequency response that will also apodize the spectral computation. The CCD modulation transfer function (MTF) will also act to suppress the higher spatial frequencies of the interference pattern thereby introducing a further apodization of the DFT. For baseband processing all of these apodization functions will be smooth symmetric functions and will help to decrease the sidelobes at the expense of decreasing the spectral resolution. However, when the DFT is placed on a spatial carrier by rotating the beamsplitter then the MTF and time integration of moving fringes will produce more modulation suppression of the high spatial frequency components, and this will no longer be symmetric with respect to the center frequency and will not suppress the resulting sidelobes as much as it broadens the main lobe. The AOD Bragg frequency response can be made nearly symmetric by rotating the Bragg cell in the interaction plane, or the band shape can even be made purposefully asymmetric, in order to partially compensate

for the MTF asymmetry, with a further rotation of the AOD.

The analysis of the processing is completed by summing over the laser pulses during the period of integration. This requires the representation of the pulse train in the time domain, but the effect of the spectral harmonic bandwidth of the laser diode due to the pulse width should be retained. The contribution due to the two bias terms will be simply N times their contribution during a single pulse for highly coherent single tones. For pulsed signals, the signal spectrum changes during the integration time but the signal bandwidth becomes so wide that fine frequency analysis is unnecessary. For this reason it will be assumed that the signal spectrum is constant from laser pulse to laser pulse, and consists of a superposition of stable single tones. The interferometric folded spectrum term of interest can be calculated with a grating lobes summation of the interferometrically detected intensity profile on each pulse.

$$
\begin{aligned}
I_i(x,y') =& \eta c_s c_r q(m_x x) p(m_y y) p(M_y y) \mathrm{rect}\left[\tfrac{y'}{H}-\tfrac{1}{2}\right] \\
& \int |\tilde{S}(f)| W\left(\frac{x}{\lambda F}+\frac{f}{v_a}\right) df \\
& \sum_n c\tau \mathrm{sinc}[(f+f_l+n\Delta f)\tau] MTF(n\Delta f/v+\alpha) \\
& 2\cos\left[2\pi[([f \bmod \frac{1}{T}]T - y'\Delta f/v)n + \alpha_y y' + \Omega(f)]\right] \\
\approx & \eta c_s c_r q(m_x x) p(m_y y) p(M_y y) \mathrm{rect}\left[\tfrac{y'}{H}-\tfrac{1}{2}\right] \\
& \int |\tilde{S}(f)| c\tau \mathrm{sinc}[(f+f_c)\tau] W\left(\frac{x}{\lambda F}+\frac{f}{v_a}\right) N \frac{\sin[2\pi NT(f \bmod \frac{1}{T} - \Delta f \frac{y'}{H})]}{\sin[\pi T(f \bmod \frac{1}{T} - \Delta f \frac{y'}{H})]} \\
& 2\cos[2\pi(\alpha_y y' + \tfrac{(N-1)}{2} T(f - \Delta f \tfrac{y}{H})) + \Omega(f)] df \\
\approx & \eta 2 c_s c_r q(m_x x) p(m_y y) p(M_y y) \mathrm{rect}\left[\tfrac{y'}{H}-\tfrac{1}{2}\right] \int |\tilde{S}(f)| c\tau \mathrm{sinc}[(f+f_c)\tau] \\
& W\left(\frac{x}{\lambda F}+\frac{f}{v_a}\right) N \mathrm{sinc}[NT(f \bmod \frac{1}{T} - \Delta f \tfrac{y'}{H})] \cos[2\pi \alpha_y y' + \Omega(f)] df
\end{aligned}
\tag{5.2.16}
$$

The grating lobes sum results in a fine frequency localization of each frequency component along the y axis of each coarse frequency ridge at the position of fringe

generation, apertured in x by the AOD amplitude blur spot and in y by the grating lobes sinc function. The grating lobes occur periodically in y' for any input frequency but are windowed by the magnified AOD aperture which is imaged onto the detector of height H. Over the detector height only one peak will be generated for any single input frequency, and the bandwidth over the height of the detector is seen to be Δf. The apertured grating lobes can be conveniently replaced by a sinc function for large N, and the phase factors which are due to the definition of the shifted reference window segments can be dropped. The spectral resolution of the time integrated DFT is given by the inverse of the integration time $1/NT$, but it will be degraded from this by the effects of apodization. If an input frequency increases by an exact harmonic of the laser diode pulse repetition frequency, then the fine resolution sinc in the y' dimension is at the exact same position. Since the bandwidth of the fine frequency analysis is given by Δf, then it is clear that the space integrating coarse frequency channelizer should have a resolution of at least Δf, so that two input signal with a frequency difference of exactly Δf would be discernible. This indicates that the acoustic transit time of the signal AOD should be at least $T_a = 2/\Delta f$ if apodization effects are to be included.

5.2.5 System impulse response and demodulation

The system output for a single tone input can be best visualized by breaking up the input frequency into a sum of its coarse and fine parts, $f' = k/T + \delta_f$. The coarse part is chosen to be the nearest harmonic of the laser diode pulse repetition frequency (PRF), and the fine part is the offset from this frequency. The analytic single sided negative frequency spectrum of the input tone that will result in a doppler upshifted diffraction is given by $\tilde{S}(f) = |a|e^{i\Omega}\delta(f + f')$. In this case the impulse response of the folded spectrum processor can be found by assuming that the fine part will not make the space integrated blur spot move by a significant

amount, and the coarse part is aliased away by the sampling action of the pulsed laser diode. For full generality, the possibility of a spatial carrier in the coarse (x) dimension will be included as well.

$$
\begin{aligned}
I_{f'}(x,y') =& Nc\tau c_r^2 q^2(m_x x)p^2(m_y y) + Nc\tau c_s^2 p^2(M_y y)\left|aW\left(\frac{x}{\lambda F}-\frac{f'}{v_a}\right)\right|^2 \\
&+ 2\eta c_s c_r q(m_x x)p(m_y y)p(M_y y)\int |a|\delta(f+f')c\tau \mathrm{sinc}[(f+f_c)\tau] \\
&W\left(\frac{x}{\lambda F}+\frac{f'}{v_a}\right) N\mathrm{sinc}[NT(f-\Delta f\tfrac{y'}{H})]df\cos[2\pi\alpha_y y' + \alpha_x x + \Omega] \\
=& Nc\tau c_r^2 q^2(m_x x)p^2(m_y y) + Nc\tau c_s^2 p^2(M_y y)\left|aW\left(\frac{x}{\lambda F}-\frac{f'}{v_a}\right)\right|^2 \\
&+ 2\eta c_s c_r q(m_x x)p(m_y y)p(M_y y)|a|c\tau\mathrm{sinc}[(f'-f_c)\tau]W\left(\frac{x}{\lambda F}-\frac{k}{v_a T}\right) \\
&N\mathrm{sinc}[NT(-\delta_f-\Delta f\tfrac{y}{H})]\cos[2\pi\alpha_y y' + \alpha_x x + \Omega] \qquad (5.2.17)
\end{aligned}
$$

This pattern consists of a uniform bias, a signal dependent space integrated coarse frequency bias ridge, and the desired interferometrically detected folded spectrum impulse at position $x = k\lambda F/v_a T, \quad y' = -\delta_f H/\Delta f$, and riding on a spatial carrier in the y' dimension of spatial frequency α_y. and in the x dimension of spatial frequency α_x. The amplitude and phase of the interferometrically produced spatial carrier under the envelope of the integrated peak at the coarse and fine locus are proportional to that of the input tone with the addition of an arbitrary phase factor.

Usually, the spatial carrier will be introduced in only one dimension, however, the dimension that is chosen to represent the spatial carrier will require an increase in the corresponding number of detector pixels. The theoretical impulse response is shown in Figure 5.2.7 and 5.2.8, for the two cases of carrier in the coarse dimension, and carrier in the fine dimension respectively, with apodizations of the two orthogonal transformations neglected. Notice that there is a tradeoff, whichever dimension is chosen to represent the spatial carrier must suffer a decrease of spatial resolution in order to represent the carrier fringes within the limit imposed by the detector space bandwidth product. When the carrier is placed in the fine frequency

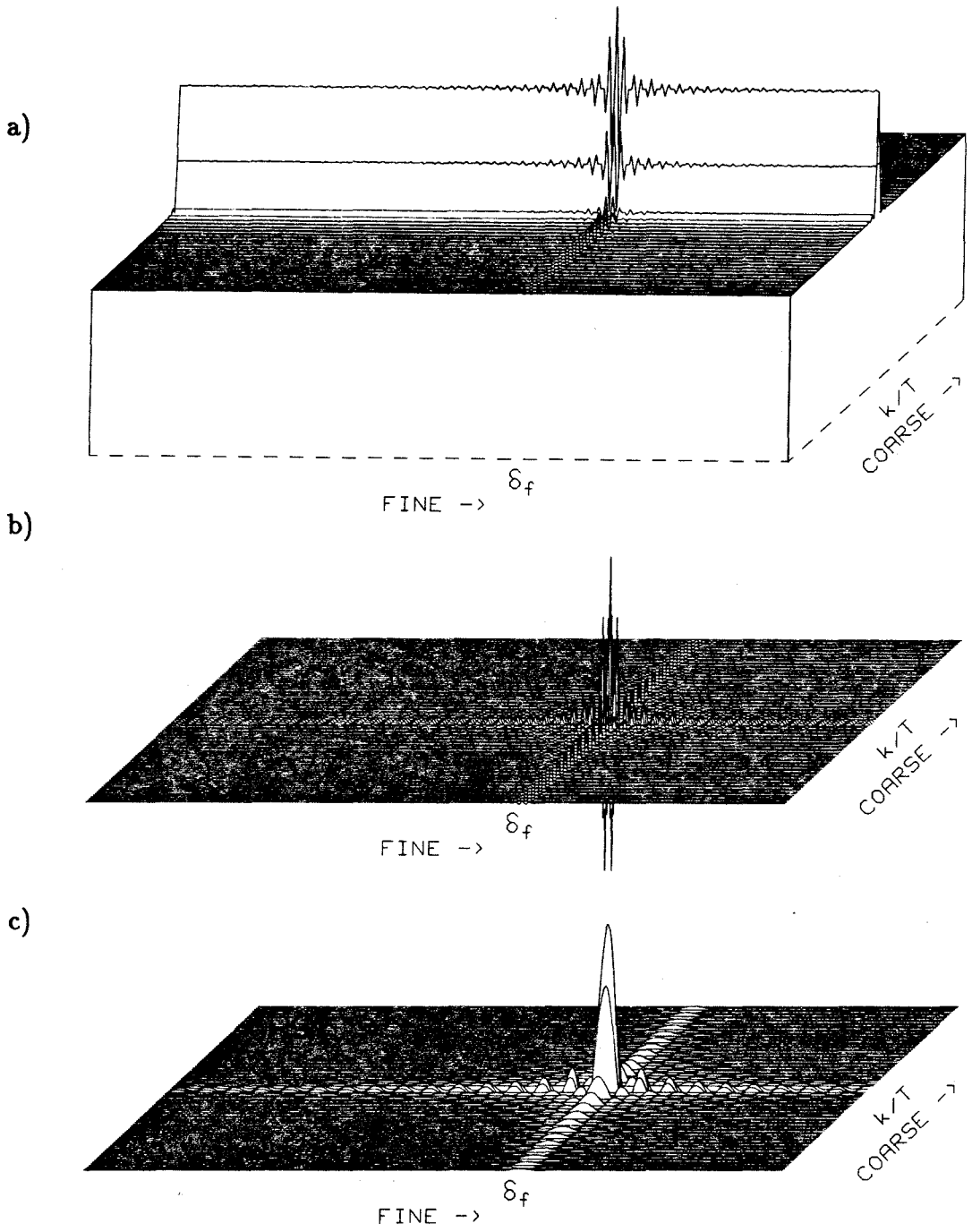


Figure 5.2.7. Perspective view of the theoretical impulse response when the carrier frequency is placed in the fine dimension with the apodization neglected. a) output of CCD, b) after bandpass filtering, c) after rectification and low pass filtering.

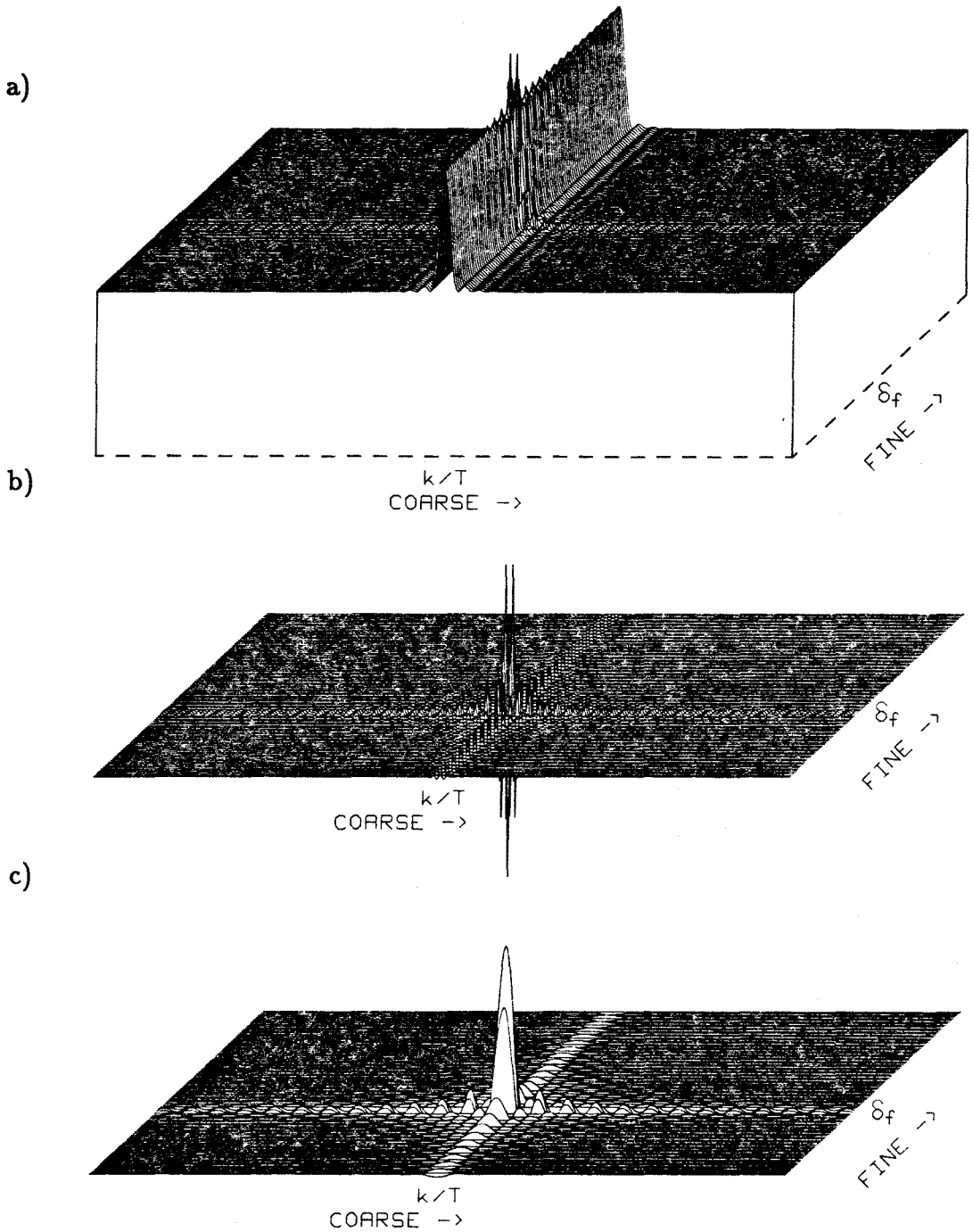


Figure 5.2.8. Theoretical impulse response when the carrier frequency is placed in the coarse dimension with the apodization and the feed through of the bias ridge neglected. a) output of CCD, b) after bandpass filtering, c) after envelope detection.

dimension, as in Figure 5.2.7 (a)-(c), the fine dimension impulse response becomes very broad in order to contain enough carrier fringes to define its center frequency accurately, and the coarse dimension can be very tightly focussed on the detector array. When the carrier is placed in the coarse dimension, as in Figure 5.2.8 (a)-(c), The impulse response in that dimension must be broadened in order to represent several carrier cycles within the width of the coarse resolution blur spot, but the fine dimension is allowed to focus down to a very narrow impulse response. These tradeoffs do not necessarily imply anything about the actual spectral resolution achieved in either coarse or fine dimension, they simply refer to spatial scaling factors at the detector plane, and efficiency of the utilization of the available CCD space bandwidth product (ie number of pixels) in the two dimension. For instance if a CCD were available with 2000×500 pixels, then whichever dimension, either coarse or fine, was chosen to contain the spatial carrier information would clearly be the direction that utilized the 2000 pixels available in the high resolution direction of the CCD. By choosing a spatial carrier frequency of 4 pixels per cycle and a resolution spot of 8 pixels in the dimension containing the spatial carrier, and a resolution spot width of 2 pixels in the orthogonal dimension, then the number of independently resolvable frequencies that could be detected in this example would be 250 in each dimension for a total of 250^2=62,500. The demodulated outputs after bandpass filtering and after envelope detection are also shown in these figures in order to demonstrate the final impulse response of the entire signal processing system. If phase information was to be obtained it would be measured after bandpass filtering the detector output. After envelope detection, the phase information is removed and direct spectral amplitude measurements can be performed. If power spectrum information is desired the spectral amplitude coefficients can be squared.

The impulse response is given by a product of sinc functions, and the sidelobes in both dimensions only fall off as 1 over π times the inverse of the number of

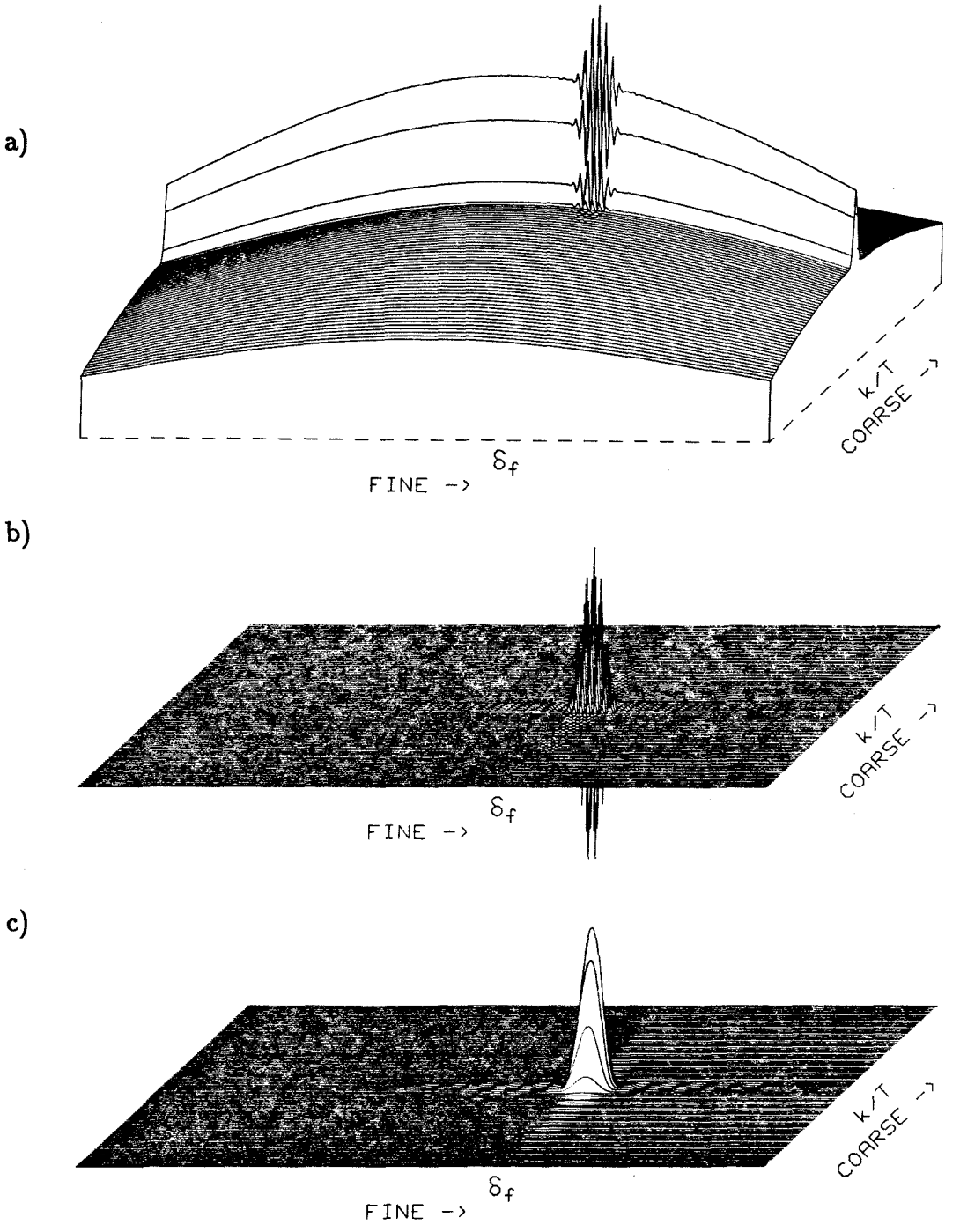


Figure 5.2.9. Theoretical impulse response when the carrier frequency is placed in the fine dimension including the sidelobe suppression due to apodization. a) output of CCD, b) after bandpass filtering and rectification, c) after low pass filtering.

lobes away from the main lobe. The sidelobes extend clear to the edge of the region plotted, and would produce spurious responses at coarse and fine frequency positions where no input tone exists. The importance of apodization in both dimensions is demonstrated by the sequence of more realistic impulse responses plotted in Figure 5.2.9. Here, the Gaussian profile of the laser beams results in an overall spatial weighing of the output, and this also results in an apodization of the space integrating coarse spectral analysis. The time integrating DFT performed on each coarse frequency bins is also apodized by the reference AOD frequency response, the CCD MTF, the integration of moving fringes, the sequence of weights of the nearest laser spectral harmonics as the reference frequency is stepped, and an auxiliary weighing function that can be applied to the successive columns of the raster scanned DFT reference signal. The apodizations broaden the main lobes in both the coarse and fine dimensions, and thereby decrease the available spectral resolution in the fine dimension and increase the required analysis bandwidth of each DFT processor to the achieved resolution of the apodized coarse frequency analyzer. But this tradeoff between resolution and sidelobe level is clearly worthwhile in order to avoid spurious responses due to the sidelobes. The maximum amount of apodization that should ever be used would decrease the sidelobe level of a strong CW tone that almost saturates the CCD dynamic range to the level of the noise.

The spatially continuous optical field will be detected on a discrete pixel detector array with pixel profile $h(x,y)$, and pixel separation Δx in the coarse dimension and Δy in the fine direction.

$$\begin{aligned} D(n,m) &= \sum_{n'=0}^{N-1} \left[\int\int h(x-n\Delta x, y-m\Delta y) I(x,y,n') dxdy \right] \\ &= \int\int h(x-n\Delta x, y-m\Delta y) I(x,y) dxdy \\ &= \int\int I(f_x,f_y) H(f_x,f_y) e^{i2\pi[f_x x + f_y y]} df_x df_y \end{aligned} \qquad (5.2.18)$$

In this case the only effect of the discrete sampling is essentially a weighing of the

spatial frequency spectrum of the incident intensity on each pulse by the Fourier transform of the pixel response function, $H(f_x, f_y)$, known as the CCD modulation transfer function (MTF). This will result in an additional apodization of the DFT transformation given by the vertical component of the CCD MTF. The sidelobe levels will be reduced and the main lobe will be broadened, thereby decreasing the obtainable fine frequency resolution, and broadening the coarse frequency impulse response. This spatial frequency weighing can equivalently be performed after the time integrating accumulation of laser pulses since both operations are commutable linear operators, assuming that the CCD is in the linear response regime. The pixel response function consists of the rectangular aperture convolved with the wavelength and temperature dependent Gaussian diffusion of carriers, convolved with the exponential tails in both dimensions due to the CCD readout mechanism charge transfer inefficiency. The horizontal component of the CCD MTF will degrade the space integrating coarse frequency resolution, which will increase the required fine frequency analysis bandwidth. Of course it is required that any incident intensity profile must obey the Nyquist condition for spatial sampling by the CCD in order to avoid aliasing effects on the CCD.

The temporal output data stream will be bandpass filtered along one of the axis (or possibly both) in order to remove the bias terms. The ability to reject the bias terms depends on the filter design, and on the spectrum of the various bias terms in the dimension of the filtering operation. Primarily two filtering methods should be considered, tapped delay lines and active bandpass filters. When the fringe pattern is in the horizontal dimension of the CCD, the output is a video bandwidth sampled data stream that requires high frequency filtering such as an active filter or multistage LC filter can provide. Alternatively when the fringe pattern is in the vertical dimension of the CCD the data streams of all the vertical columns are interleaved and must be stored in video delay lines or a computer so that they

can be processed by a discrete tapped delay line filter or processed by a digital convolution in a frame buffer memory. In order to choose the optimum dimension in which to insert the spatial carrier, the two cases of carrier demodulation in coarse and fine will be considered. The spatial frequency spectrums of the bias and signal terms of the two cases are plotted in Figure 5.2.10, along with the appropriate filter response and tap coefficients for an example demodulation operation in each of the two cases.

When the spatial carrier is placed in the coarse dimension, $\alpha_y = 0$, it rides on top of the incoherently detected coarse frequency ridge, which because it is tightly focussed will have a wide spatial bandwidth. Ignoring the effects of apodization, the incoherently detected coarse frequency ridge will have a spatial profile that will be an adequately sampled version of the square of the Fourier transform of the AOD aperture, $W^2(x) = \text{sinc}^2[xA/\lambda F]$, so that its spatial spectrum will be triangular $|\lambda F/A| - f_x$. Needless to say, the AOD apodization will narrow this spectrum while it broadens the blur spot. The reference beam bias should be approximately uniform and have a very narrow spatial frequency spectrum given by $Q * Q(f_x/m_x)$, but multiple reflection may introduce high frequency fringe patterns, and the CCD noise will probably be spatially white and thus have a flat spectrum clear out to the Nyquist frequency. The interferometric fringe pattern of spatial frequency α_x in the x dimension is under a coherently detected envelope given by the Fourier transform of the idealized AOD aperture, and will have an amplitude profile $W(x) = \text{sinc}[xX/\lambda F]\cos[2\pi\alpha_x x]$ with higher sidelobes than the incoherent ridge. The sampled version of this high spatial frequency pattern will have its modulation amplitude decreased by the MTF of the CCD, and the resulting spatial spectrum will be a shifted rectangular bandshape, $\text{rect}[(f_x - \alpha_x)\lambda F/X]$. Clearly the spatial spectra of the signal cannot overlap with any of the bias terms, and its overlap with multiple reflection coherent artifact modulations and CCD noise spectra should be

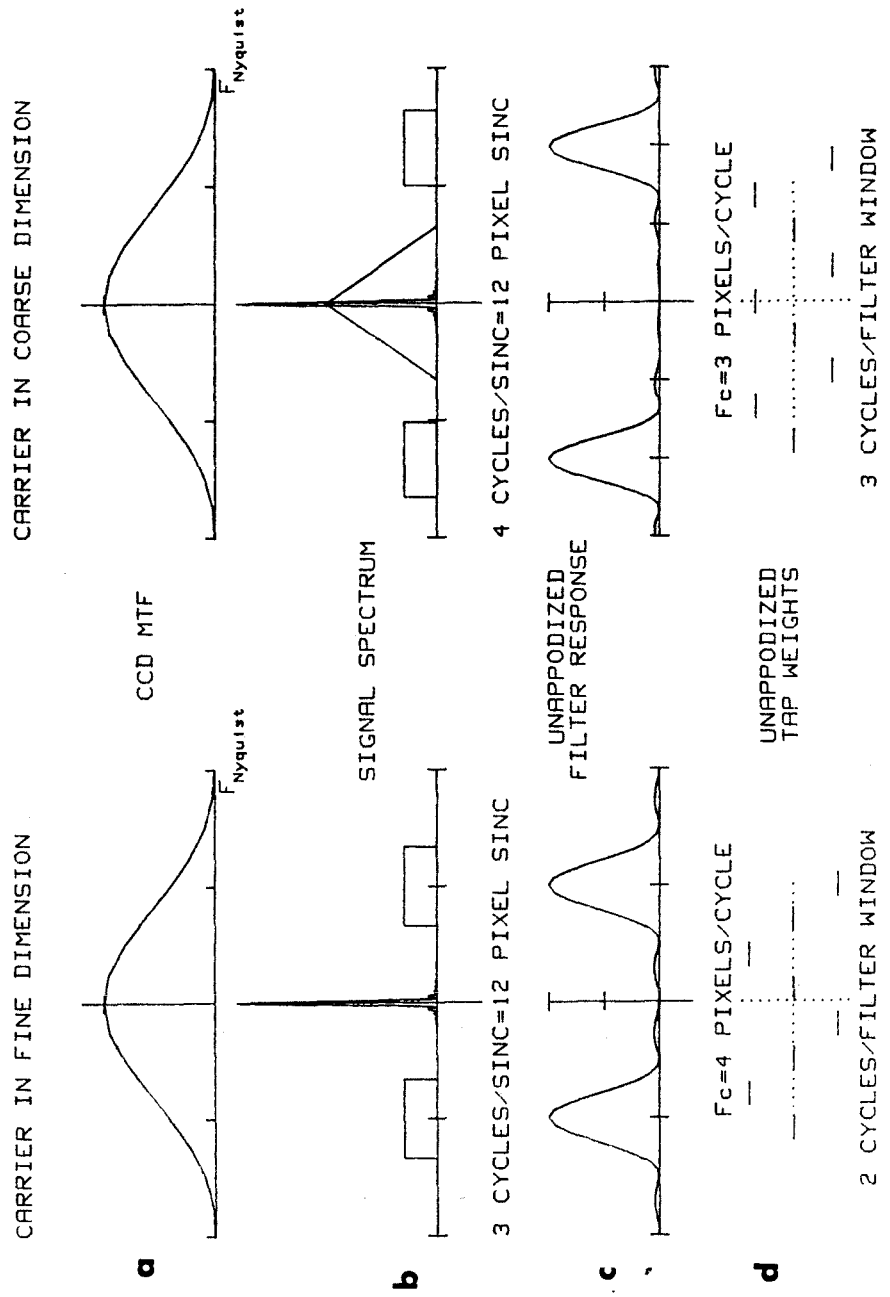


Figure 5.2.10. Carrier demodulation for the two cases of carrier in coarse and fine. a) MTF of the CCD, b) Spatial frequency spectrum in that dimension of the CCD, c) unapodized bandpass filter frequency response, and)d taps that are used to implement the filter.

minimized if possible. A more restrictive constraint is that the integrated power of the bias terms that passes through the sidelobes of the bandpass filter should be less than the integrated power of the noise that passes through the filter. Because the spectrum of the incoherent coarse ridge is so wide, filter apodization may be required to reduce the sidelobe amplitude feedthrough of the bias ridge, which will further broaden the bandpass filter and increase the filter length thereby broadening the filtered coarse analyzed fringe pattern of width $W(x/\lambda F)$, by the convolution with the filter impulse response.

Alternatively, the interferometrically produced spatial carrier can be placed along the coarse frequency ridge in the fine dimension ($\alpha_x = 0$), with a spatial frequency in y of α_y. In this case the reference beam bias will be approximately uniform, and will have a very narrow spectrum along the y axis given by $P * P(f_y/m_y)$, but it could also contain high spatial frequency terms due to coherent noise artifacts, and the CCD noise spectrum in y will very likely be flat. Similarly, the signal beam bias along the ridge should be essentially uniform and should have a very narrow spectrum given by $P * P(f_y/M_y)$, but it can also contain high frequency coherent artifacts along the ridge. The interferometrically detected fringe pattern will have an idealized spatial spectrum, ignoring any apodization effects, that is given by $\mathrm{rect}[(f_y - \alpha_y)H/\Delta fNT]$, and the modulation of the high spatial frequency terms will be attenuated by the CCD MTF. The width of the bias spectral components are much less in this case, so the condition that the spectra of the signal and bias terms due not overlap can be more easily met with a lower spatial frequency carrier than when the carrier is in the coarse dimension. The bandpass filter needs to be designed so that very little of the bias terms near zero spatial frequency will pass through, but the number of filter taps should be minimized in order to avoid broadening the fine resolution peak by the convolution with the filter impulse response. The DC response of a tapped delay line filter can always be made zero

by simply having the average of the filter tap coefficients to be zero, and an unapodized filter with an integral number of half cycles of the proper phase will have a sinc zero crossing at 0 spatial frequency. The Q of the filter is determined by the frequency width over the center frequency which, for an unapodized rectangular filter window, is given by the number of cycles within the filter impulse response. The filter bandwidth should be matched to the bandwidth of the sinc riding on a carrier, and optimally the filter should be a matched filter to the expected impulse response. Thus if the expected impulse response is a sinc function with 3 full cycles under the main lobe, then the filter should be designed as a sinc function apodized filter with 3 cycles under the main lobe. In this case the signal spectrum and filter spectrum are matched rectangles and the resulting convolution will not widen the filtered impulse response and the spectral resolution will be the same as the width of the initial sinc. However, this requires a very long filter to accurately represent the infinite sinc sidelobes, and the added processing requirements of the output bandpass filtering is proportional to the filter length. In the case of an apodized DFT the sinc function impulse response is convolved with the transform of the apodization window, thereby reducing the sidelobes and broadening the main lobe. A subsequent bandpass filtering operation will further broaden the impulse response in the filtering direction by approximately the sum of the filter width and impulse response width, thereby decreasing the obtainable spectral resolution. For this reason there is a trade off between the obtainable resolution after spatial carrier demodulation, and the rejection of the unwanted bias terms.

5.2.6 Optical design of the DFT interferometer

In this section and the following sections I will describe the experimental verification of the ability of a time and space integrating folded spectrum processor to accurately measure the coarse and fine frequency of an input tone. This begins with a theoretical optical design, which is followed by the modifications necessary in order to utilize commercially available optical components. A careful alignment procedure must be followed in order to obtain a close approximation to the desired system behavior. Once the system is accurately aligned initial interferometric tests were performed in order to measure the coherence and stability of the optical system and reference waveforms. Finally, the full folded spectrum processing operation was performed with a variety of test waveforms and bias removal techniques in order to characterize the system operation.

Each of the active components in the system has a physical size and relative scale parameter which dictates certain aspects of the optical design. In order to illustrate the design procedure of a time and space integrating processor, I will briefly discuss the important design choices which were necessary to compensate for device geometry in the interferometer tested in these experiments. The active optical devices which dictate the design parameters are the laser diode, the reference and signal acousto-optic devices, and the 2-dimensional CCD video detector array.

The laser diode used in these experiments was a Hitachi HLP-1600 single mode device with approximately 20mW of CW output power. Because the laser will be pulsed with a very low duty factor, being on for only about 50 nsec every 10 μsec, the very limited optical power (about 10^9 photons per pulse) emitted by this device must be efficiently utilized by the processor. The laser has a near field spot size of about d_x =1.3 μm by d_y =4 μm, with the plane of polarization in the elongated dimension. The far field output beam has an anamorphic divergence at the full angle half power points of θ_x =30 degrees by θ_y =10 degrees. When this beam

is collimated by a spherical lens of focal length F_{col}, then the output elliptical Gaussian beam profile will have a half power size given by $D_x = 2F_{col}\tan\frac{\theta_x}{2}$ by $D_y = 2F_{col}\tan\frac{\theta_y}{2}$. The light utilization efficiency will be limited by the f-number of the lens, which is the ratio of lens aperture to focal length, and in general smaller lenses will have a larger f-number. For this reason a short focal length collimator, with F_{col} =14.8 mm, and aperture of 8 mm, was combined with a 6 times beam expander, with an aperture of 48 mm, in order to produce a collimated output beam profile with an aperture at the half power points of the elliptical Gaussian beam given by D_x' =48 mm by D_y' =15 mm. Notice that the half power profile is actually the .7 amplitude profile, which contains slightly more than half the optical power within the aperture of the beam expanding telescope.

The AODs used in these experiments were slow shear TeO_2 anisotropic polarization switching devices, with a nominal acoustic velocity of .62 mm/μsec, and a spatial aperture of about 42 mm, which yields an acoustic delay time of 68 μsec, although less than that is actually utilizable due to edge effects. Since these devices are highly anisotropic, and will be used with a focussing optical beam in the transducer height dimension, the angular spread of the focussing beam should be minimized in order to reduce the extent of off-axis diffraction effects. This dictates that long focal length cylinders should be utilized, and the longest readily available focal length of 300 mm was chosen for F_{c1} and F_{c2}. The signal AOD needs to be fully illuminated, while the reference AOD only needs to be uniformly illuminated over the spatial aperture of one reference segment, which should be about 1/4 of the full aperture of the signal AOD. Clearly, one should use the large axis of the elliptical Gaussian beam to illuminate the full aperture of the signal AOD, and use the narrow axis of the elliptical beam to illuminate the smaller portion of the reference AOD that must be illuminated. The unapodized resolution of a space integrating AO spectrum analyzer implemented with the full aperture of the signal Bragg

cell would be the inverse aperture time, giving a resolution of 15 kHz. However the acoustic attenuation and laser beam apodizations will decrease the obtainable resolution quite severely. The field uniformity of the collimated incident Gaussian profile over the 42 mm aperture of the signal AOD is about .77 at the edges, and the additional resolution loss due to the optical nonuniformity will be only 10%. The acoustic attenuation across the aperture will introduce a 4.4 dB attenuation at the center frequency of 60 MHz, which will introduce another amplitude profile variation from one edge of the cell to the other of about .72, further broadening the space integrated blur spot.

An even more important consideration is the spatial scale factor of the space integrated Fourier transform resolvable spot size on the CCD detector array, and the additional spatial blurring due to the resolution limitation of the detector array. The CCD used in these experiments was a Sony XC-37 industrial black and white video camera with 384 horizontal by 491 vertical pixels. The individual pixel separation in the horizontal dimension is 23 μm, and in the vertical dimension the pixel separation is 13.4 μm. This yields an overall array size of 8.8 mm in the horizontal dimension by 6.6 mm in the vertical dimension, for a conventional television aspect ratio of 4/3. Notice that the horizontal dimension is the fast readout dimension of the raster output format with a pixel output frequency of 7.4 MHz, and the vertical dimension is a double interlaced scan at a video line rate of 15.7354 kHz. There are two possible orientations of the CCD in the processor output plane, with the space integrated coarse folded spectrum axis in either the horizontal or vertical CCD dimension. The horizontal dimension has fewer pixels, yielding fewer individually addressable coarse frequency analysis channels, but the increased spatial separation of the individual pixels will help to decrease the number of pixels within the space integrated blur spot of the signal AOD. Ideally, the coarse blur spot size $\lambda F_3/A$ should be about the same as the pixel separation Δx when the spatial carrier is

placed in the fine dimension. However, this will require a very short focal length lens, which will have too small an aperture to integrate the entire diffracted field without introducing additional spatial blurring due to lens aberrations, which are larger on high f-number lenses. The coarse pixel resolution is given by $\Delta x v_a/\lambda F_3$, which should be equal to the fine frequency analysis bandwidth, while the overall bandwidth is given by the number of pixels in the coarse dimension times the frequency width of each pixel.

For convenience in the definition of coarse and fine frequency, the laser diode PRF was chosen to be exactly 100 kHz in the initial set of experiments. The reference segment is therefore going to be approximately 10 μsec in duration, which will have a spatial extent in the reference AOD of 6.2 mm. The narrow axis of the elliptical Gaussian beam profile will have a 95% amplitude uniformity over this spatial aperture. This will be imaged onto the CCD with a sphere cylinder telescopic lens pair with a magnification $m_y = -F_{c3}/F_2$. The aperture profile in the orthogonal dimension of the reference wavefront will be a demagnified image of the wide axis of the elliptical Gaussian beam profile with magnification given by $m_x = -F_2/F_{c2}$. In order to obtain an optimum light efficiency utilization, a very large demagnification of the wide beam orthogonal to the reference AOD should be used in order to collect it all on the detector, but this must be traded off against the desire for a uniform amplitude profile of the reference wave at all coarse frequency bins. Similarly, in the signal arm the demagnification of the narrow axis of the elliptical Gaussian beam profile orthogonal to the direction of the signal AOD must be limited by the desire for uniform amplitude weighing of the signal beam along each coarse frequency ridge.

All of these conditions and tradeoffs are mollified by the commercial availability of the dictated lens ratios, and the additional condition that the path lengths of the two arms should be approximately equal. The final design parameters which were

Lens	Focal Length	Aperture
Collimator	F_{col}=14.8 mm	8 mm
Beam Expander	x6	48 mm
Cylinder 1	F_{c1}=300 mm in y	60 x 50 mm
Cylinder 2	F_{c2}=300 mm in x	50 x 60 mm
Signal Fourier Transform lens	F_3=250 mm	50 mm
Reference Transform lens	F_2=120 mm	48 mm
Reference imaging cylinder	F_{c3}=150 mm in y	60 x 50 mm

Table 5.2.1. Lens values chosen for optical components in the experimental demonstration of the folded spectrum interferometer.

chosen for these experiments are shown in Table 5.2.1. The compromise lens values were not fully optimal, but they served to demonstrate the principles involved.

The major limitation of the experiments described in this chapter is that the space integrated coarse spectral resolution is much wider than the fine frequency analysis bandwidth. There are primarily two reasons for this limitation, difficulty of high bandwidth reference waveform generation and excessive blurring of the coarse frequency analysis. The reference waveform is a phase reset stepped frequency raster scan of the DFT matrix, which must have excellent coherence throughout its entire duration. The reference waveform should have a bandwidth that fully utilizes the AOD bandwidth, and its temporal extent is N times the duration of each segment, which results in a time bandwidth of about $N^2 \approx 10^6$. Two methods of reference waveform generation were examined, analog generation with a voltage controlled oscillator (VCO), and digital reference generation with a video frame buffer. The excessive blurring of the space integrated coarse frequency resolution

spot was caused by optical aberrations, as well as photo generated charge diffusion due to the deep penetration of the infrared photons in the CCD detector. These limitations will be discussed in more detail as the experimental results are presented.

5.2.7 Alignment procedure of the crossed Bragg cell interferometer

A highly involved optical alignment procedure was required in order to obtain even reasonable fringe modulation in the active crossed Bragg cell interferometer described here. The invisible nature of the infrared radiation from the laser diode made these alignments much more involved than an equivalent visible alignment procedure would be. This procedure involved the iterative realignment of all optical components in order to have three separate alignment measurements in as good agreement as possible. These three conditions were the alignment of all back reflections from the multiple surface lens elements on the optical axis, the transmission of a pencil beam without any angular deviation, and most important, the interferometric testing of the intermediate wavefront through the interferometer. Except for the collimator, the lenses closest to the CCD were aligned first, and once a component was aligned it was never moved again. The interferometer configuration that was arrived at was remarkably stable when the optical table was isolated on air, and needed very little realignment from day to day. The interferometer alignment easily survived a noticeable shaking due to an earthquake.

It was very important that each optical component had all of the necessary degrees of freedom of its motion, and the critical motions needed to be on smooth micrometer driven stages. Spherical lenses required 3 translational degrees of freedom, with the focussing mechanism being the most critical, and a gimbal motion to adjust the two tilt axes. Cylindrical lenses required an accurate focussing method, and a translation only in the dimension of lens curvature, with a critical in plane

rotation, and both axes of tilt. The mirrors required in plane translational motion, with both axes of tilt. The acousto-optic devices require all three dimensions of translational motion, combined with a Bragg angle adjustment and optical tilt plane adjustment, and an acoustic rotation would have been useful but was not utilized. The CCD detector array had all translational degrees of freedom, and it had an important rotational motion, as well as a vertical tilt. The laser diode had all 6 degrees of freedom, and was coupled to the collimator and beam expander on a single plate that could be tilted and translated in both horizontal and vertical directions to direct the collimated beam without affecting the collimation. The beam splitters were the most critical of all the components, they were aligned with a translation stage running perpendicular to the split face, and they had all three rotational degrees of freedom arranged with an in plane rotation, and a critical tilt orthogonal to the split face with an unnecessary and relatively insensitive rotation about the splitting face.

The alignment procedure begins with the collimation of the laser diode by the collimator beam expander, as interferometrically tested with an optical flat. These components were mounted together on a tiltable and rotatable plate so that the beam could be directed parallel to the optical table and parallel to an alignment guide in the first leg of the interferometer. An adjustable iris was placed immediately beyond the collimator output aperture and centered on the beam, in order to produce a narrow pencil beam for component alignment. The first beam splitter cube, with a 2 inch aperture, was inserted and aligned so the back reflection from the cube face retroreflected, and the beam split off was made parallel to the table and to an alignment guide orthogonal to that in the transmission leg. The mirrors were inserted in both arms and aligned at 45 degree angles producing output beams parallel to the table, and parallel to the alignment guides in each arm, thereby forming a nearly rectangular pair of beams. The recombining 1 inch beam

splitter cube was inserted where the beams crossed, with the splitting face at a 45 degree angle in order to recombine the beams as parallel wavefronts. Using a small aperture pencil beam the back reflections from the cube faces were trace back up the table, and the mirrors and beam splitters were iteratively realigned until the back reflections traced back up the same exact path in both arms, while the output beams were interferometrically measured to be parallel, thereby guaranteeing that the beams now propagated in a perfectly orthogonal rectangle. The CCD detector array was centered in one arm of the output beam, and the other arm was magnified and observed with an IR viewer. The recombined output beams were very nearly parallel, and the interferometrically produced output seen on the CCD is shown in Figure 5.2.11a, where less than half a wave of nonparallelism is observed. The worst problem encountered in these Mach-Zehnder interferometer experiments was the unwanted honeycomb pattern which is observed in this photograph, which is probably due to multiple reflections off the cube beam splitter faces. The cube beam splitters used in these experiments were antireflection coated on the cube faces to a specification of less than 1% intensity reflectivity per face, and the multiple reflection artifacts are due to a pair of double pass reflections that reflect off at least 2 cube faces and the splitting face at least once, and then interfere with the main beam. Thus the expected amplitude of each of the multiple reflections should be about 1% as strong as the main beam, so the coherently produced multiple reflection artifacts should have a modulation depth of about 4%, however the observed modulation depth was closer to 10%, and this anomalously large modulation depth introduced severe bias variations. This might be reduced by using a larger recombining cube beamsplitter so that the path length difference between the multiple reflections and the main beam was longer than the pulsed laser coherence length, but this would not eliminate the much weaker coherent artifact introduced by the interference of different multiple reflection paths with each other. Another problem

is that cube beamsplitters should only be used with plane waves, and in this TSI Mach-Zehnder interferometer a space integrated focussing slit passes through the beam splitter introducing unwanted cylindrical aberrations. The best solution to all of these problems would be to use a wedge beamsplitter with one of the faces antireflection coated, or to use a very thin pellicle beam splitter. The unwanted vibrational modes associated with pellicle beamsplitters might not cause serious problems because of the short integration times involved with a single frame of processing.

When the interferometer is fully aligned, the reference arm will reproduce a plane wave, while the signal arm will produce a focussing slit. Because of this the reference arm can be used to align the signal arm but not vice versa, so the reference arm should be aligned first. Optical systems are always aligned with the component farthest from the source being aligned first. The cylindrical lens F_{c3}, with power in y, was inserted and focussed onto the CCD, and rotated so the focal slit from the cylinder was parallel with the sampling CCD detector pixels. Next the spherical lens F_2 was inserted before the cylinder by a distance equal to the sum of their focal lengths which was interferometrically verified by observing that the output beam is quadratic in the x dimension and perfectly collimated in y, as is shown in Figure 5.2.11b. The cylindrical lens F_{c1}, with focussing power in the x dimension, is inserted before the mirror in the reference arm, at a distance before the sphere given by the sum of the respective focal lengths. The focussing beam from this cylinder bounces off the mirror as it collapses down to a slit where the reference AOD will be positioned, and then expands back out, to be recollimated in x by the following sphere. The rotation of this cylinder with respect to the orthogonal cylinder is quite critical in order to remove all residual quadratic phase factors. After the alignment of this cylinder, the reference arm reproduces an anamorphically magnified plane wave, that interferes with the plane wave in the unmodified signal

arm, and produces the fringeless pattern shown in Figure 5.2.11c, which indicates a wavefront aberration of less than half a wave over the CCD aperture.

The next component that should be aligned is the y propagating reference beam Bragg cell, at the focal slit of the first cylinder, but this presents certain problems, because the AO diffraction bends the angle of the output beam, and doppler shifts it so it will no longer interfere with the plane wave in the signal arm which is being used as an alignment reference. Thus both AODs must be aligned, and then the diffracted components must be interfered to measure the residual wavefront aberrations after acousto-optic diffraction. The reference AOD is inserted at the position of the focus of the first cylinder and aligned at an angle of θ_B with respect to the optical axis in order to produce a strong +1 diffraction order. In the Fourier plane of F_2 a very narrow slit is positioned at the DC focal slit, to mark the center of the previously aligned optical components, and the position of the +1 diffraction order is carefully marked in the Fourier plane. Then the mirror before the AOD is rotated through an angle of θ_B, so that the DC beam is rotated by $2\theta_B$, and strikes the AOD at the angle $-\theta_B$, and is focussed onto the same spot in the Fourier plane that the +1 diffraction order was previously. At exactly the right mirror rotation angle the -1 diffracted order will pass through the narrow slit positioned in the center of the Fourier plane, thereby guaranteeing that the -1 diffracted order will be aligned with the following optical components. The only problem with this approach is that in the space between the mirror and the AOD, the optical beam shifts along the direction of acoustic propagation by several centimeters. This must be compensated for by translating the mirror back towards the laser by an appropriate distance so that a deflected pencil beam once again strikes the center of the CCD. This arm is now fully aligned and is shown in cross section in the top view of the interferometer shown in Figure 5.2.12., where the bend due to the AOD diffraction, and the compensating mirror tilt and translation are indicated to scale.

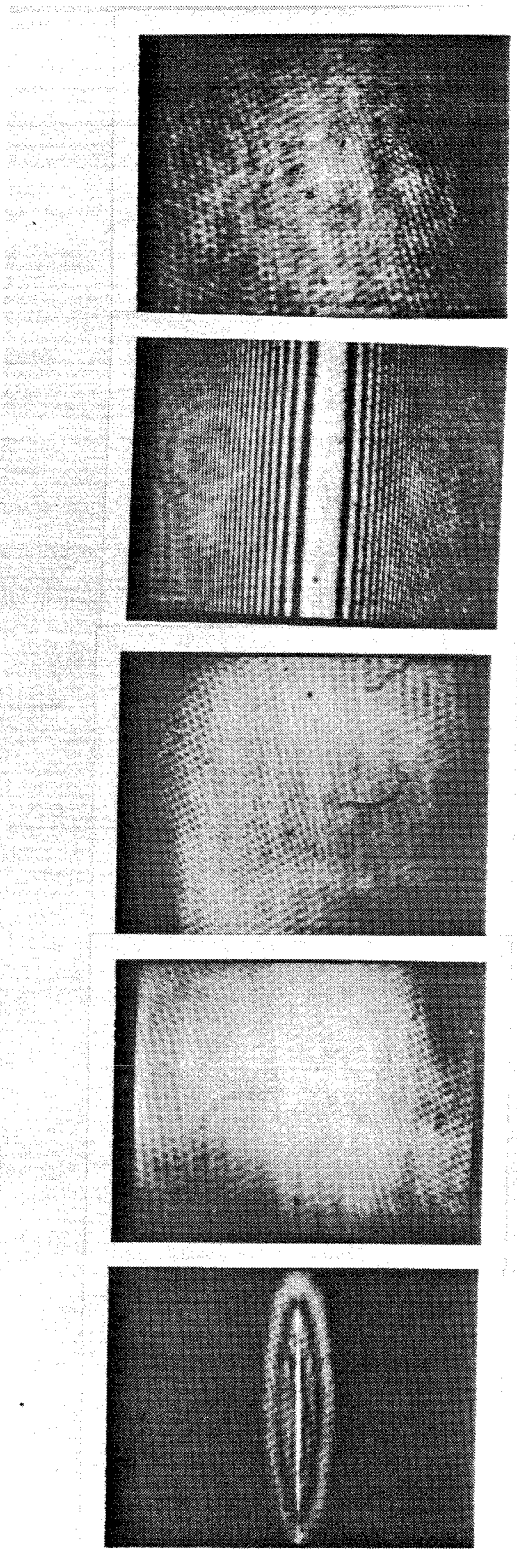


Figure 5.2.11. The interferometrically produced pattern detected on the CCD at various stages of the alignment procedure (see text).

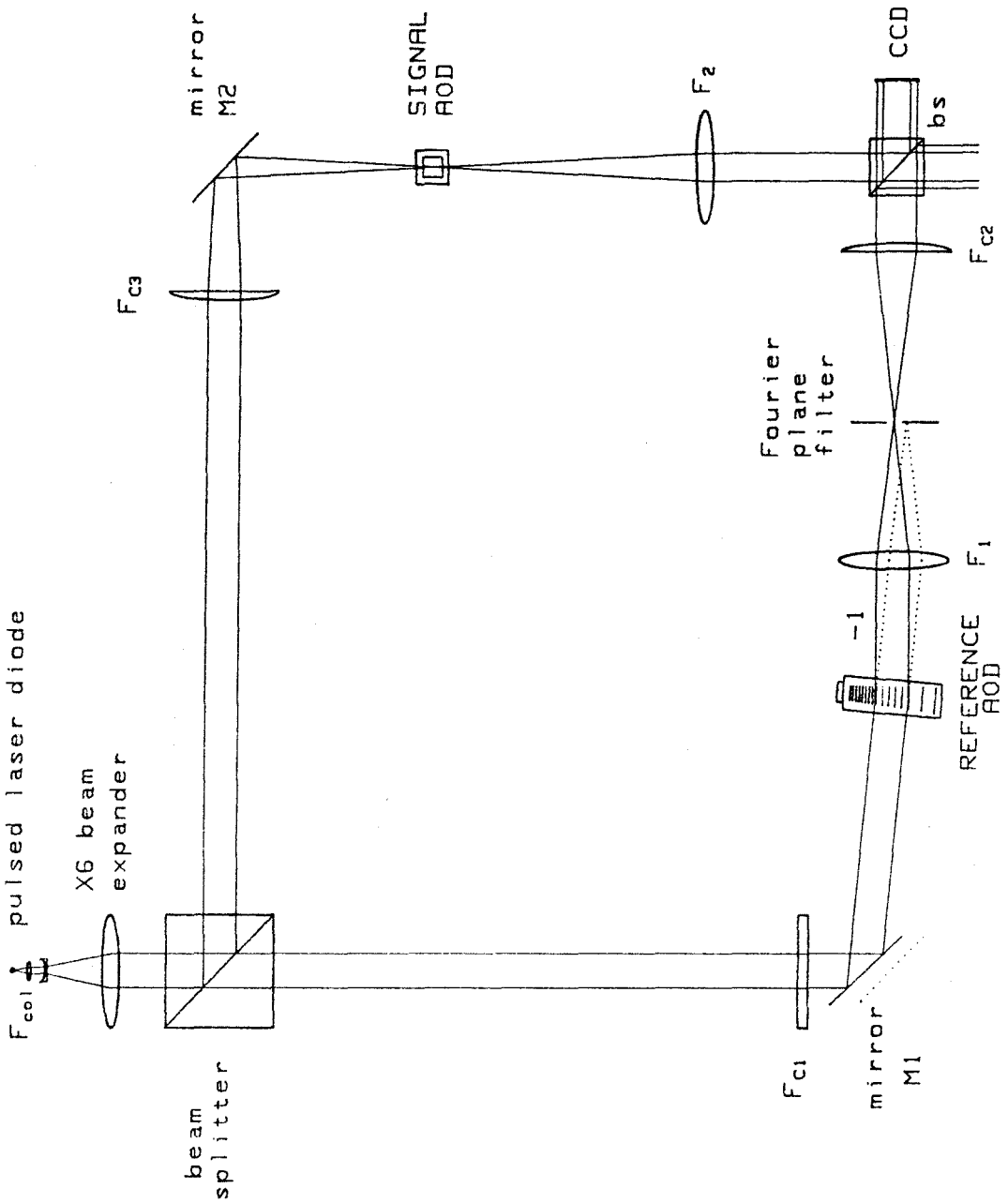


Figure 5.2.12. Top view of the TSI DFT interferometer as aligned, approximately to scale.

A very similar procedure must be performed for the signal AOD, using the input plane wave, with the additional complication that the acousto-optic diffraction bend is out of the plane of the interferometer. The appropriate position for the signal AOD is found by inserting the Fourier lens so that it optimally focusses on the CCD and then placing the focussing cylinder behind the sphere by the sum of their focal lengths and interferometrically aligning them into a telescopic system. At the focal plane of the cylinder the x propagating signal AOD is inserted, and aligned to produce a strong +1 diffraction order, at which point the FT sphere and the cylinder are removed from the signal arm. A small aperture pencil beam was used to illuminate the AOD, so that the +1 diffracted order was deflected down towards the table. Then the mirror before the AOD was tilted out of the plane, by an angle $2\theta_B$, so that the -1 order was deflected perfectly parallel with the table, and the DC beam propagated down towards the table at the same angle the +1 order did before the mirror was tilted. The problem is that in the intervening space between the mirror and the signal AOD the tilted incident beam walks down towards the table, and the diffracted -1 order comes out of the AOD too low. This cannot be corrected as simply as in the reference arm, because the shift is out of the plane of the interferometer. In order to correct for this shift, the first beamsplitter must be tilted up so the reflected beam walks up, before it hits the mirror, by the same amount that it will walk down after reflecting off the tilted mirror. However the two adjustments are coupled so that the output beam is no longer parallel to the table, and the tilts of the mirror and beam splitter must be performed iteratively, until the -1 diffracted order at midband is diffracted exactly parallel to the table at the same height as the original DC beam. At this point the laser aperture can be widened, and with the same midband frequency of 60 MHz applied to both AODs the interference pattern after AO diffraction can be recorded on the CCD, and is shown in Figure 5.2.11d, where less than half a wave of aberrations are visible. By

adding an extra Hertz to one of the signals the interference pattern is made to blink on and off at a 1 Hz rate. The coherence was found to be somewhat limited by the crossed Bragg cell alignment procedure, in that the reference beam path length was shortened by about 2 cm, while the signal beam path was slightly lengthened. To compensate for this path length difference a glass optical flat was added to the reference arm of the interferometer before the first cylinder, effectively lengthening the optical path in that arm. Perfect path length matching was not achieved because the thickness of the optical flat was not large enough, and the residual path length difference introduced a degradation of the coherence of the two arms that limited the achievable modulation depth to approximately 40%.

The signal Fourier transforming sphere can now be reinserted, in order to focus the diffracted plane wave from the signal AOD down to a small spot on the output detector array, and the spot size should be minimized. Finally, the cylinder which bounces a focussing cylindrical wavefront off the tilted mirror and produces an optical slit in the signal AOD can be replaced where it was before, however it no longer focusses into the signal AOD properly. The focal slit is rotated so that it is no longer parallel with the x propagating acoustic wave in the signal AOD. This is because the entire optical system between the beam splitter and the AOD have been rotated into a new coordinate system by the tilting of the mirror and beam splitter, and the cylinder must be aligned with this new coordinate system. This requires that the cylinder be rotated so the focal slit is parallel with the direction of acoustic propagation in the signal AOD, and that the top of the cylinder be tilted back towards the beamsplitter to make the cylinder orthogonal to the incident wave. The cylinder should be interferometrically aligned so that the output produces uniform slits of light which interferometrically beat on and off in unison, and the rotation of this cylinder must be aligned with those in the reference arm and with the CCD. The interferometrically produced output of the completely aligned interferometer is

shown in Figure 5.2.11e. The signal arm is now fully aligned, and a 3-D projection which attempts to show the various tilts of the components and the paths of the optical beams is shown in Figure 5.2.13.

Two techniques are suggested in order to simplify the alignment procedure of a crossed Bragg cell interferometer. The first is to design the interferometer configuration in order to minimize the distance between the mirror and Bragg cell in each arm, thereby minimizing the spatial shift caused by tilting the mirror. This approach will be limited by the space required for the mechanical mounting of the optical components. An alternative that will eliminate the required rotation of the cylinder in the signal arm would be to place the cylinder in the signal arm after the tilted mirror, and simultaneously to minimize the distance between the mirror and Bragg cell in the reference arm, in order to minimize the resulting path length difference between the two arms.

A picture of the fully aligned interferometer is shown in Figure 5.2.14. The laser diode and collimator assembly are mounted on a post at the right edge of the photograph, and the CCD detector array is seen at the opposite corner of the interferometer at the left edge. The large dielectric beamsplitter cube reflects about half the collimated light towards the signal arm and transmits half to the reference arm. In the signal arm the first cylinder, with horizontal power, focusses down the light, which bounces off the mirror and into the vertically propagating signal AOD. The diffracted wave from this AOD is horizontally recollimated, and Fourier transformed vertically, by the following spherical lens, to produce a diffracted horizontal slit of light for each coarse frequency component. The diffracted light is reflected off the the small output beam splitter cube and onto the CCD detector array. In the reference arm the collimated beam passes through the path length compensating optical flat, is focussed down in the vertical dimension, by the cylindrical lens, bounces off the mirror to illuminate the reference AOD with a horizontal slit. The

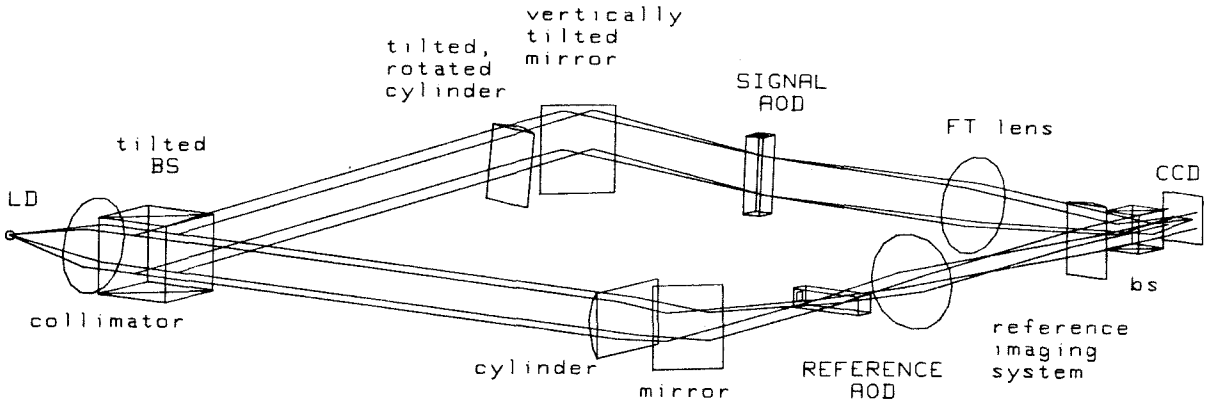


Figure 5.2.13. Perspective view of the interferometer, showing the paths of the optical beams with the out of plane tilt of the signal arm needed to compensate for the AO diffraction angle.

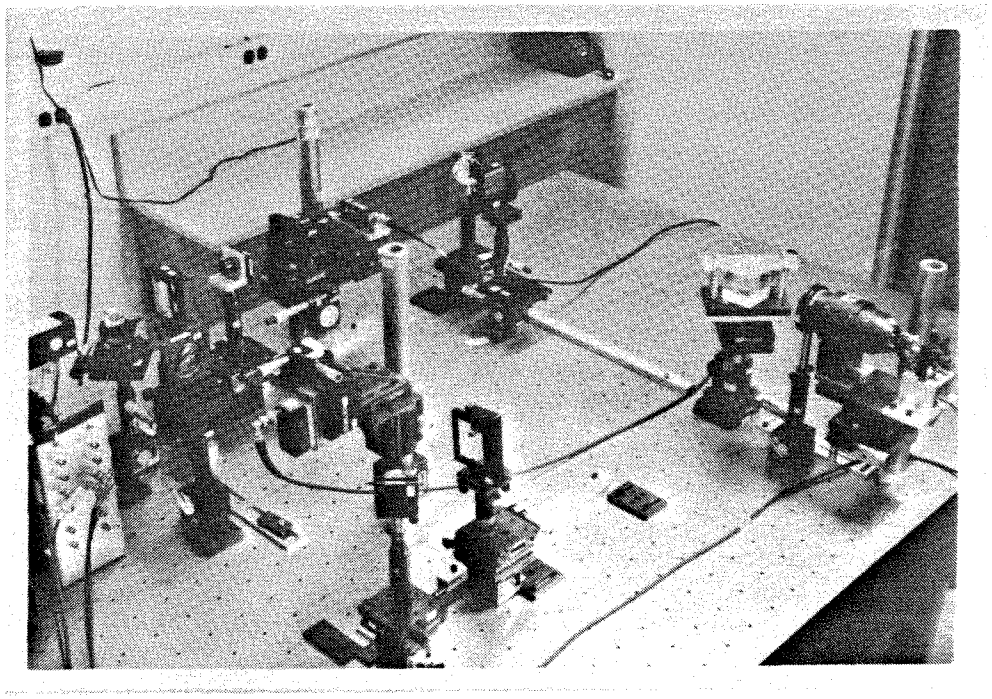


Figure 5.2.14 Photograph of the crossed Bragg cell interferometer.

diffracted light is Fourier transformed by the following sphere, which also recollimates vertically, to produce a slit of light in the Fourier plane for each reference segment. An adjustable slit acts as a Fourier plane filter to block the DC, and bandpass the diffracted light from the reference AOD. The following cylinder horizontally reimages the AOD reference segments through the beamsplitter onto the output CCD array, while leaving them uniformly expanded in the vertical dimension. The recombined space integrated signal and imaged reference are interfered and time integrated on the CCD detector array.

5.2.8 Characterization of the DFT interferometer

In order to characterize the coherence, stability and resolution of the interferometer some tests were performed before the system was operated as a folded spectrum processor. By placing a variable aperture immediately before the signal AOD the effect of the wide aperture space integration of the -1 order diffracted optical wavefront could be measured. The spatial resolution as a function of the aperture width is plotted in Figure 5.2.15. For very narrow signal AOD apertures the resolution decreases as expected, but as the aperture is widened beyond about 6mm, the detected spot size remains constant because the resolution limit due to the extended pixel response function and limited output bandwidth of the CCD in the horizontal dimension is larger than the spatial blur spot. The coarse frequency resolution and bandwidth of the space integrating power spectrum analyzer with a 6mm aperture are illustrated in Figure 5.2.16, In the top of the figure the entire CCD output is shown as the input frequency is varied, while in the bottom a magnified portion of this trace is examined. The full power spectrum analyzer bandwidth over the 384 CCD pixels is seen from the figure to be 25 MHz, giving a scale factor of 65 kHz per horizontal pixel, compared with an expected theoretical value of $\Delta x v_a/\lambda F$ =68 kHz per pixel. From the magnified image of the pixel detected blur spot, the full

width at half maximum (FWHM) can be measured to be a little more than 2 pixels, giving a detected coarse resolution of approximately 150 kHz.

The interferometric modulation depth was found to be a strong function of the illuminated aperture of the signal AOD, and this dependence is plotted as a function of aperture width in Figure 5.2.17. For an aperture width larger than 10mm the modulation depth flattened out at an unacceptably low level of about 15% . At an aperture size of 6mm the modulation depth was about 35% and the resolution was as good as it was going to be, so this was chosen to be the aperture width used in the following experiments. There are several possible explanations that can contribute to this dependence of modulation depth on aperture size, these include acoustic diffraction effects, poor laser beam collimation quality, sidelobe integration of the focussed spot by CCD pixel response, and path length differences affecting the laser coherence. To investigate these possibilities the 6mm aperture was translated along the signal AOD, normalizing the diffracted power at each stage, to obtain the plot of modulation depth versus aperture position shown in Figure 5.2.18. This plot shows an unexpected strong dependence of the modulation depth on the position of the aperture, with a peak near the center of the collimated laser beam profile, where the 6mm aperture should be placed for all further experiments. The observed oscillatory behavior could be due to an acoustic phenomenon, or to poor beam collimation. To determine which was the probable cause, the aperture was placed at the optimized position and the AOD was translated vertically behind the fixed aperture and the interferometric modulation depth and diffracted power were measured, and are plotted in Figure 5.2.19. The diffracted power can be seen to slowly roll off by about 40% as the aperture is translated the length of the AOD as is expected due to acoustic attenuation, and diffraction decreasing the acoustic power density as one moves away from the transducer. However, the modulation depth actually increases gradually as one moves away from the transducer, indicating that

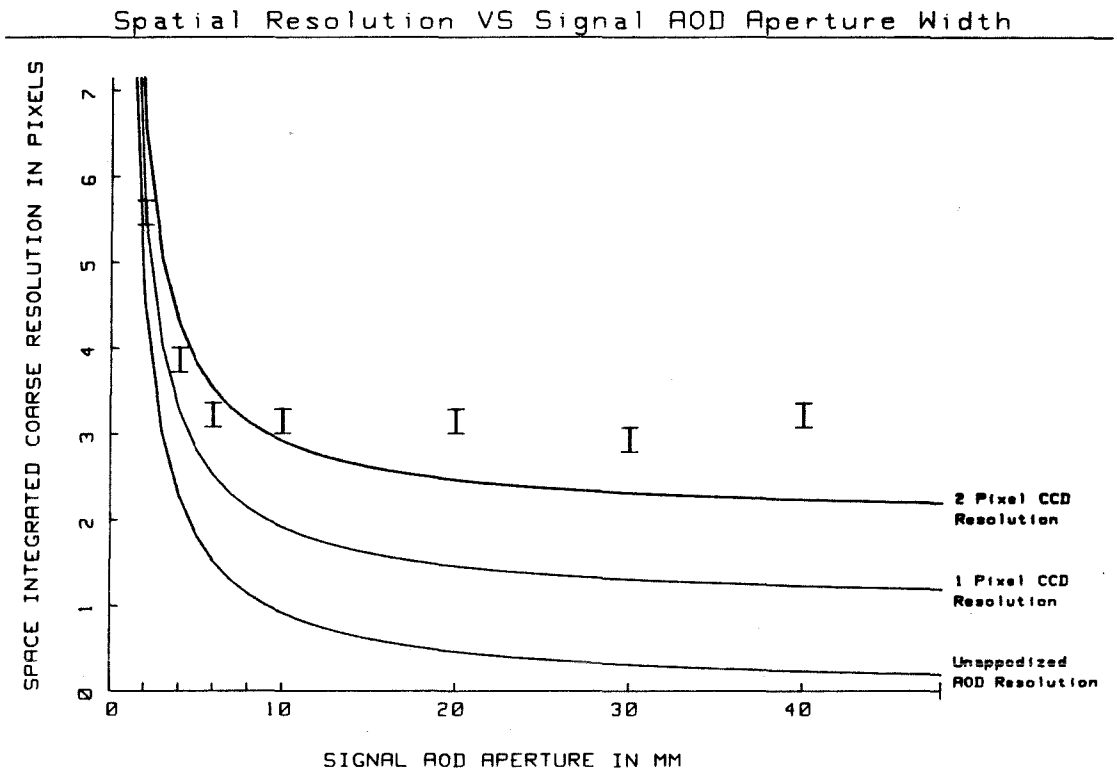


Figure 5.2.15. Spatial resolution of space integrating spectrum analyzer as a function of AOD aperture width, and the approximate effect of CCD pixel blurring.

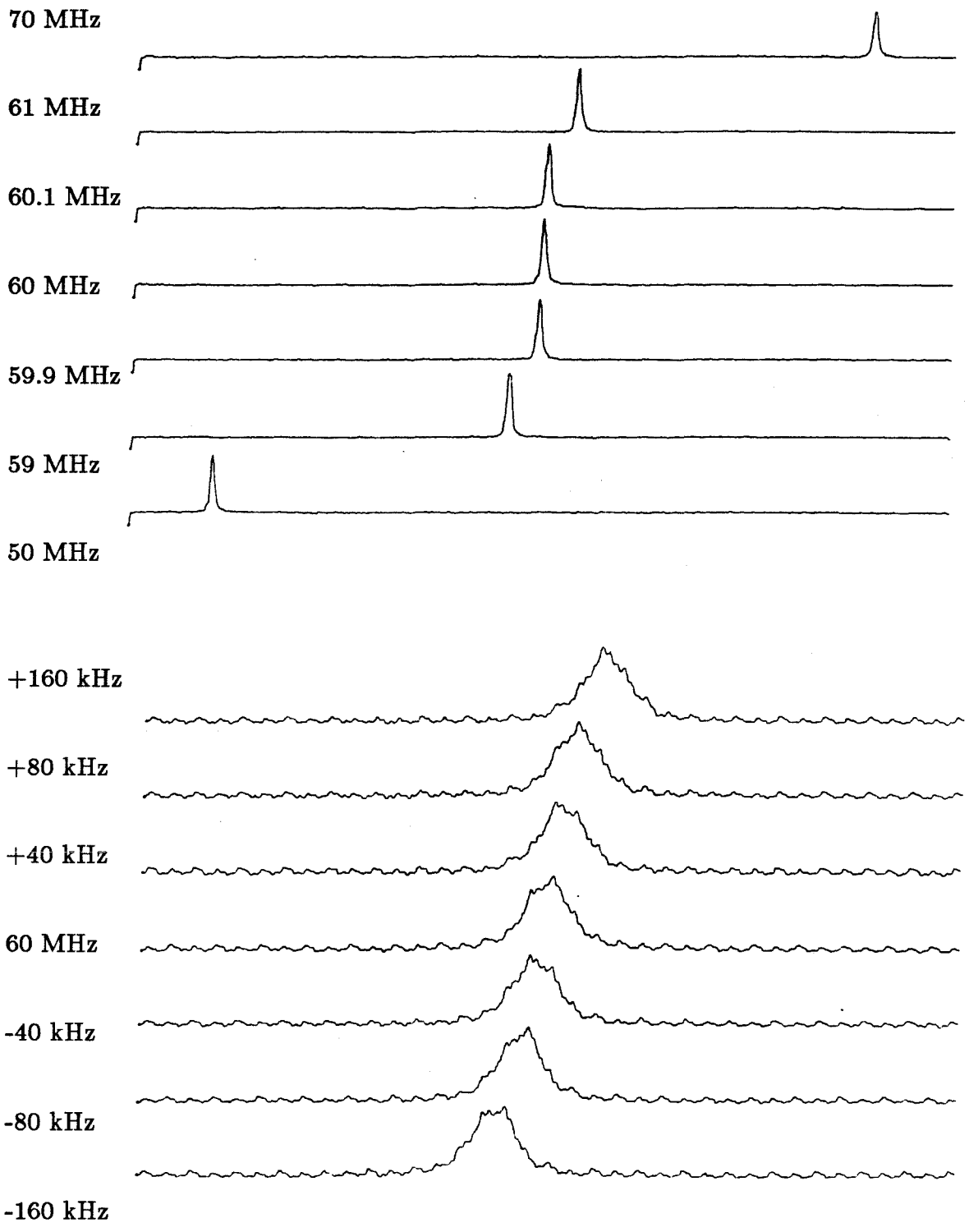


Figure 5.2.16. Space integrated blur spot for various frequencies across the processor bandwidth, and magnified pixel response for small frequency variations.

the acoustic diffraction actually helps to smooth out possible phase ripples near the transducer producing a more uniform acoustic beam near the termination end of the device. Thus the modulation depth variation seen in Figure 5.2.18, as the aperture was translated was probably due to poor laser beam collimation quality. This can be understood by considering the space integration across the aperture when a half wave plate is inserted over half of the aperture. The diffracted field from the two halves of the aperture will always be out of phase, and if the amplitudes are the same the interferometric modulations will essentially cancel. This is similar to what will occur in the presence of a one wavelength phase curvature over the utilized AOD aperture, and this is why in interferometric space integrating processors the laser beam collimation is critical. I was able to obtain good collimation over the extent of the CCD, but a few waves of curvature remained over the full 48mm collimator aperture, which is why I could not spatially integrate across this full aperture. Another effect that limited the modulation depth of the space integrated ridge was due to the collection of the out of phase sidelobes, along with the main peak within the extended pixel response function. The first sidelobes will destructively contribute about 1/3 the contribution of the main lobe. By doubling the resolution of the unapodized space integrated peak from when its width to the first zero is equal to the extended pixel response function, one would expect a decrease of the interferometric modulation depth by one third. This is one reason why proper apodization of the space integrated signal AOD is more important in an interferometric system than in a power spectrum analyzer. With the small 6mm aperture which was used in these experiments the signal AOD is essentially unapodized, and the coarse frequency impulse response will be an extended pixel response sampled sinc function.

The space integrated diffracted wavefront also suffers from the introduction of different optical path lengths caused by the acousto-optic diffraction geometry

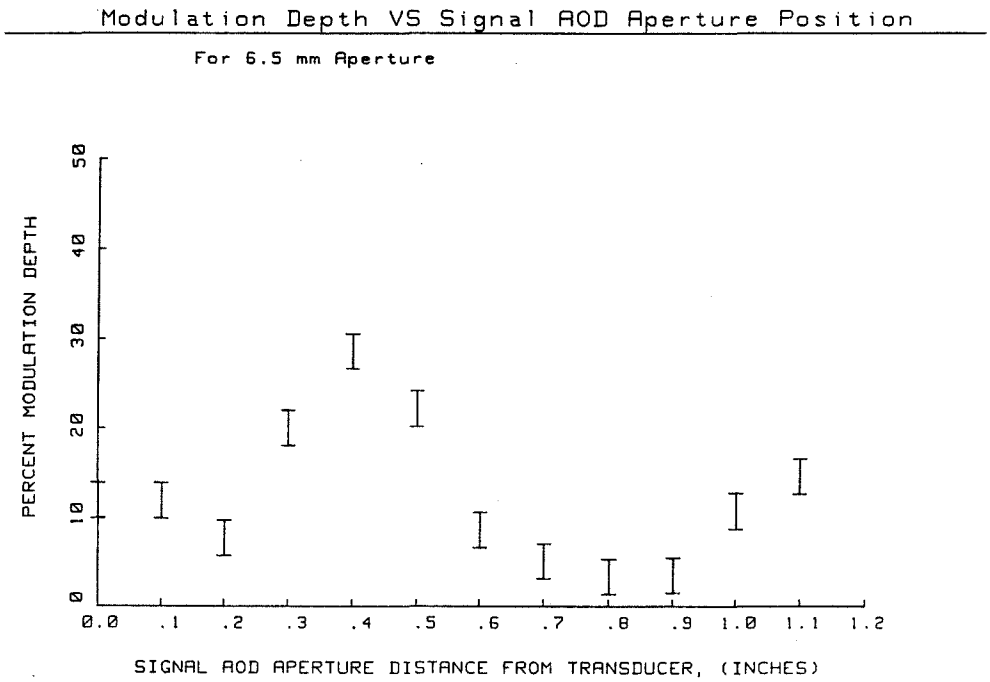


Figure 5.2.17. Modulation depth as a function of signal AOD aperture width.

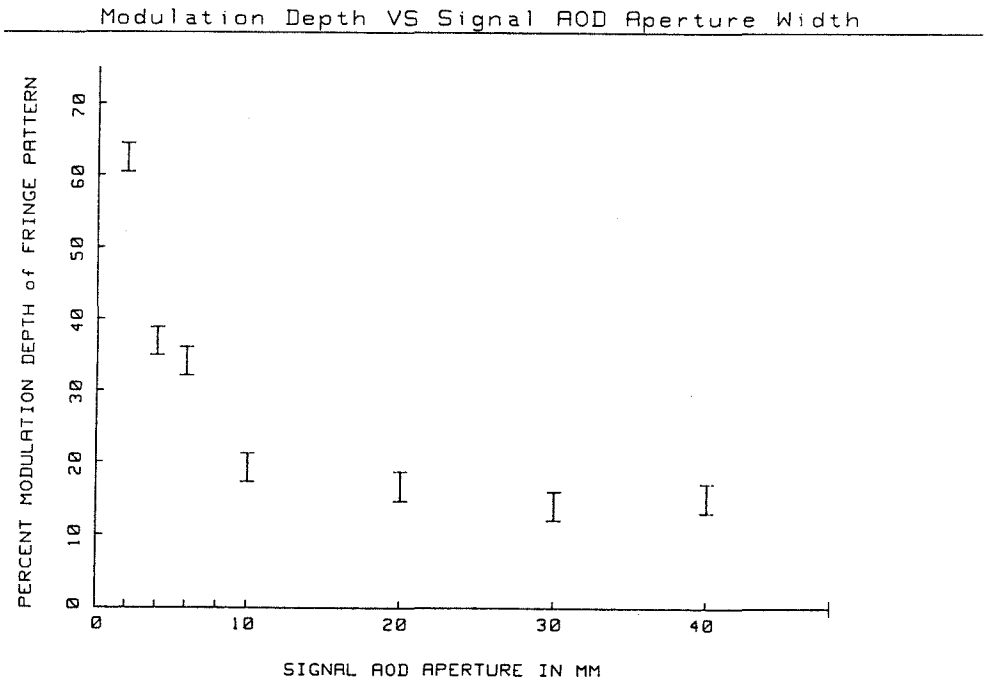


Figure 5.2.18. Modulation depth as a function of the aperture position along the signal AOD for a 6.5mm aperture.

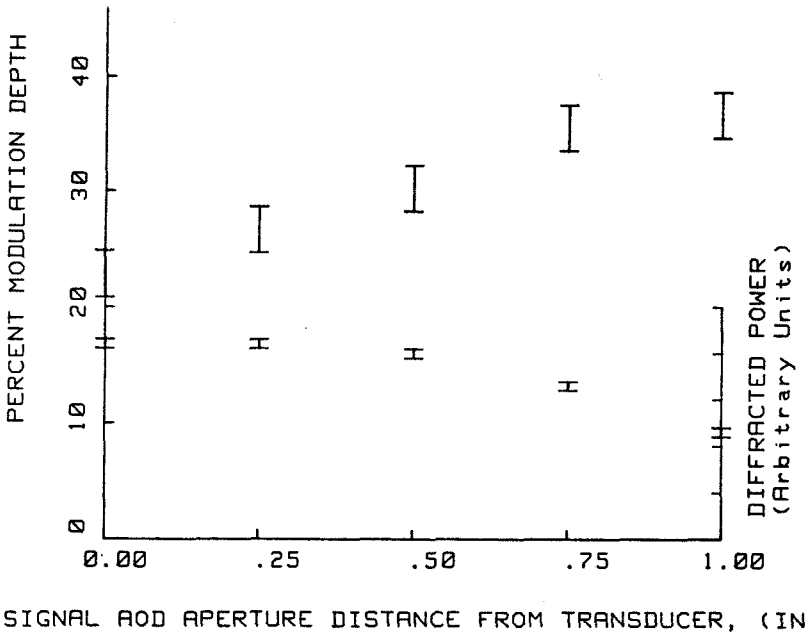


Figure 5.2.19. Modulation depth and diffracted power as the signal AOD is translated behind the fixed 6mm aperture at the center of the collimated beam.

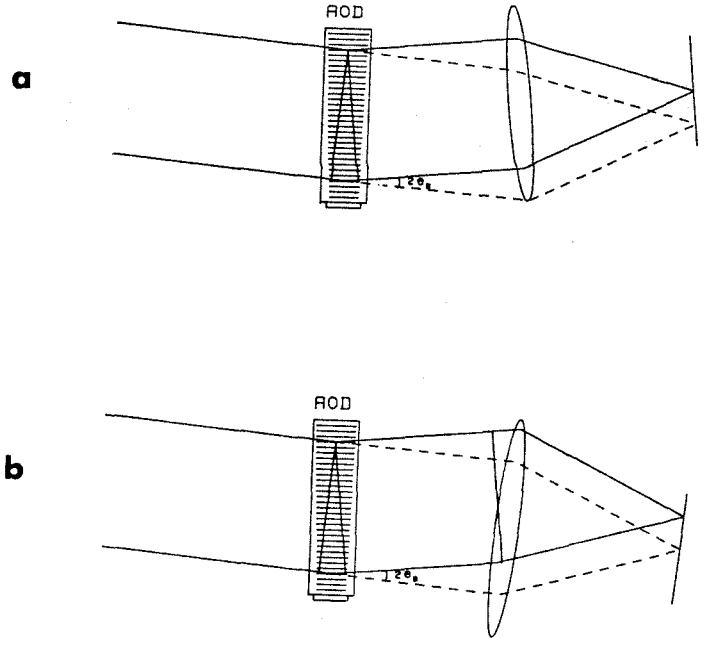


Figure 5.2.20. a) Schematic illustration showing path length difference due to acoustic diffraction through angle $2\theta_B$, when lens is aligned with diffracted wavefront. b) Path length difference is compensated when the lens is aligned with the undiffracted wave, thereby decreasing temporal coherence requirements.

which are then accumulated by the lens, and thereby decreasing the coherence of the blur spot. The total path length difference from one side of the aperture to the other is illustrated in Figure 5.2.20a, and can be seen to be $2A\sin\theta$, for an aperture A. This is a consequence of the Fourier transform lens being realigned with the diffracted wavefront in order to minimize the aberrations produced by non paraxial rays. If instead the lens remained aligned with the undiffracted incident beam and was used off axis by the diffracted wave, then the path length differences caused by the tilt caused by the AOD diffraction is compensated at midband by the effective tilt of the Fourier transform lens as seen by the diffracted wave. This alternative arrangement of the space integrating Fourier transform lens is illustrated in Figure 5.2.20b. It will result in larger off axis lens aberration effects, but the increased coherence of the focussed spot may make this tradeoff worthwhile, since limited coherence was one of the largest problems encountered in these experiments.

The system resolution and CCD modulation transfer function (MTF) were measured in the fine frequency dimension with two different methods. In order to measure the CCD MTF in the interferometric system, 60 MHz CW tones were applied to both Bragg cells and the beam splitter was rotated, thereby producing various spatial frequency modulations on the CCD. An extra .1 Hz was added to one of the 60 MHz tones and the modulation depth of the resulting moving fringe pattern was measured at its maximum. This is plotted in Figure 5.2.21, and it is nearly flat clear out to the Nyquist frequency limit of the detector. This plot does not display the expected sinc^2 function MTF dependence because each measurement was spatially shifted in order to be detected with the optimum modulation depth by the CCD, thus this is really a best case MTF. When the spatial frequency of the incident intensity profile is equal to the Nyquist frequency or half this frequency then the worst case modulation depth can be zero. Another important system characterization is the modulation transfer function detected by the CCD when

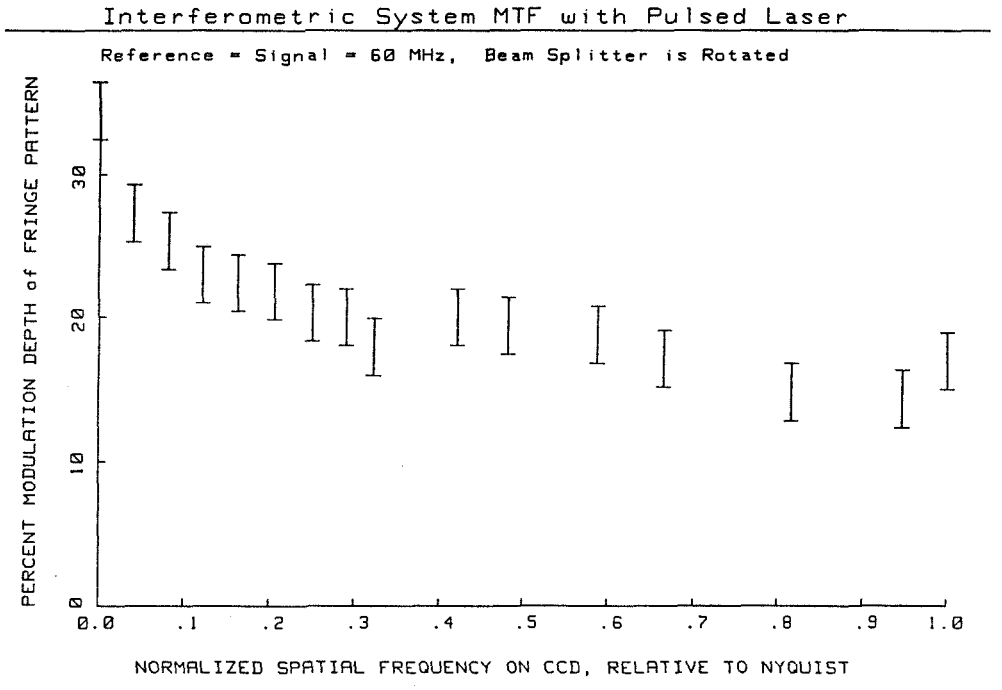
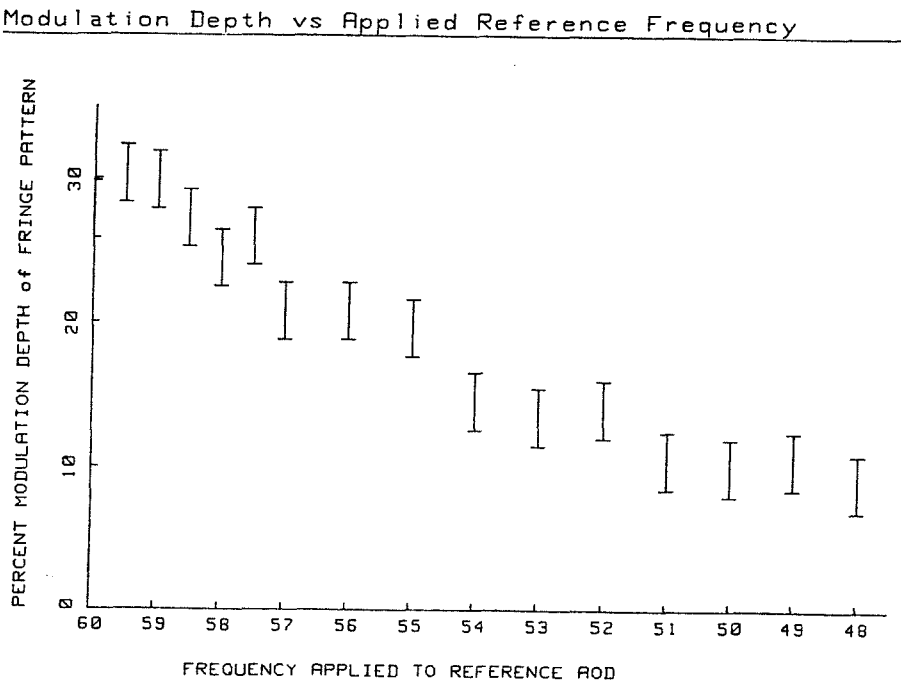


Figure 5.2.21. Interferometer MTF with pulsed laser illumination as beam splitter is rotated.

a)



b)

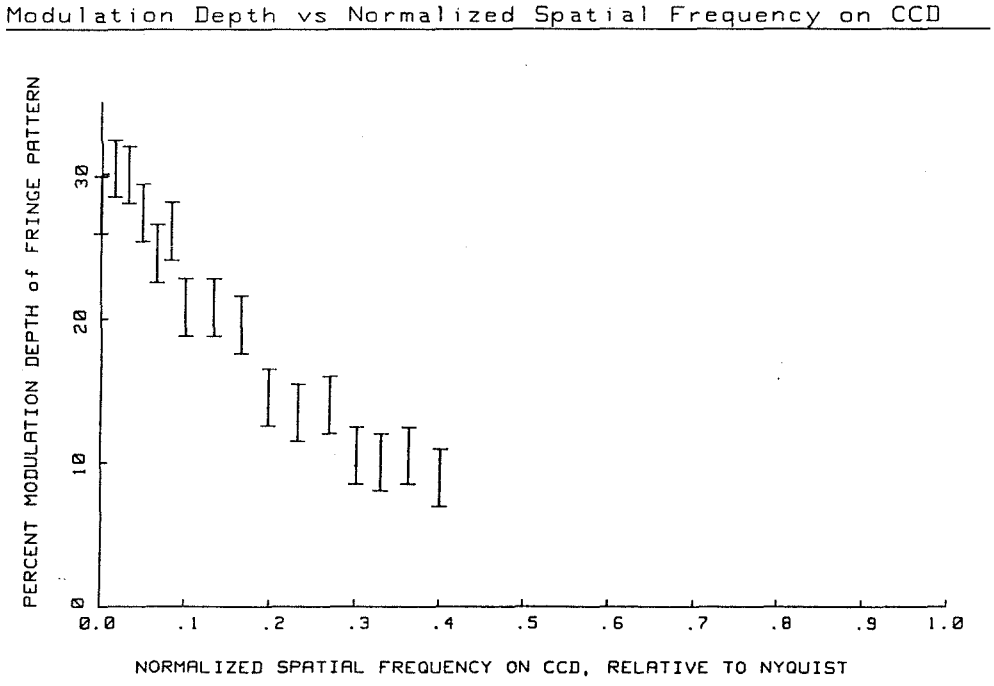


Figure 5.2.22. a) Modulation depth of fringe pattern versus reference frequency with pulsed source illumination. b) Modulation depth as reference frequency is varied plotted versus normalized spatial frequency detected on the CCD.

different reference frequencies are applied to the reference AO with pulsed source illumination. The modulation depth as a function of applied frequency is plotted in Figure 5.2.22a, and it is seen to roll off as the reference frequency deviates from the 60 MHz signal frequency because of the integration during the laser pulse of the moving interference fringes. The half modulation offset frequency is about 6 MHz, which indicates that the effective pulse width was about 80nsec for this experiment. The modulation depth data can also be plotted as a function of normalized CCD detector frequency as is shown in Figure 5.2.22b, and the roll off can be seen to be much faster than the CCD MTF alone. Since this frequency response is an apodization function of the time integrating DFT, and will strongly affect the modulation depth as well as the achievable resolution, it is important to minimize the laser pulse width, while still having enough light to fully utilize the available CCD dynamic range.

A set of interferometrically produced spatial fringe patterns produced by applying 60 MHz to the signal AOD and varying the frequency applied to the reference AOD by increments of 80 times the video line rate is shown in Figure 5.2.23. The linearity of the RF frequency applied to the reference AOD versus the interferometrically produced spatial frequency detected on the CCD was very good as can be seen in Figure 5.2.24, with a variance of only .5%. The theoretical value of this slope should be $Hm_y/v_a = 8.5$, and the experimentally determined value of 8.15 (cycles/CCD height)/MHz indicates that either the magnification was not exactly 1.25, or the assumed acoustic velocity of v_a=.62 mm/μsec is too low. Since the acoustic velocity is a strong function of interaction angle, increasing by an order of magnitude as the acoustic direction deviates from the <110> slow shear axis, a larger than assumed acoustic velocity is the most likely explanation of the slight discrepancy.

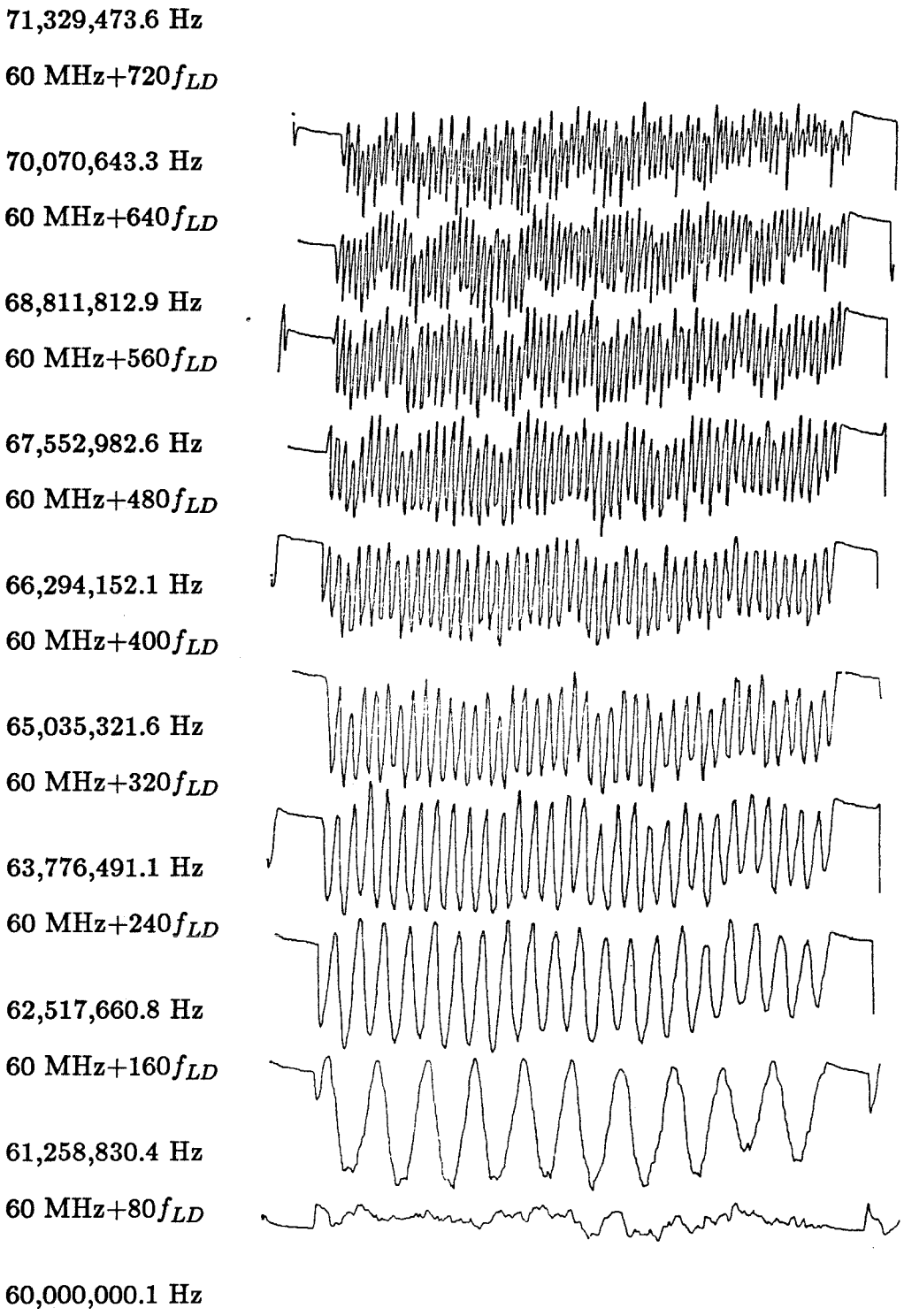


Figure 5.2.23. Interferometrically produced spatial fringe patterns detected on the CCD detector for reference frequencies incremented by 80 times the video line rate.

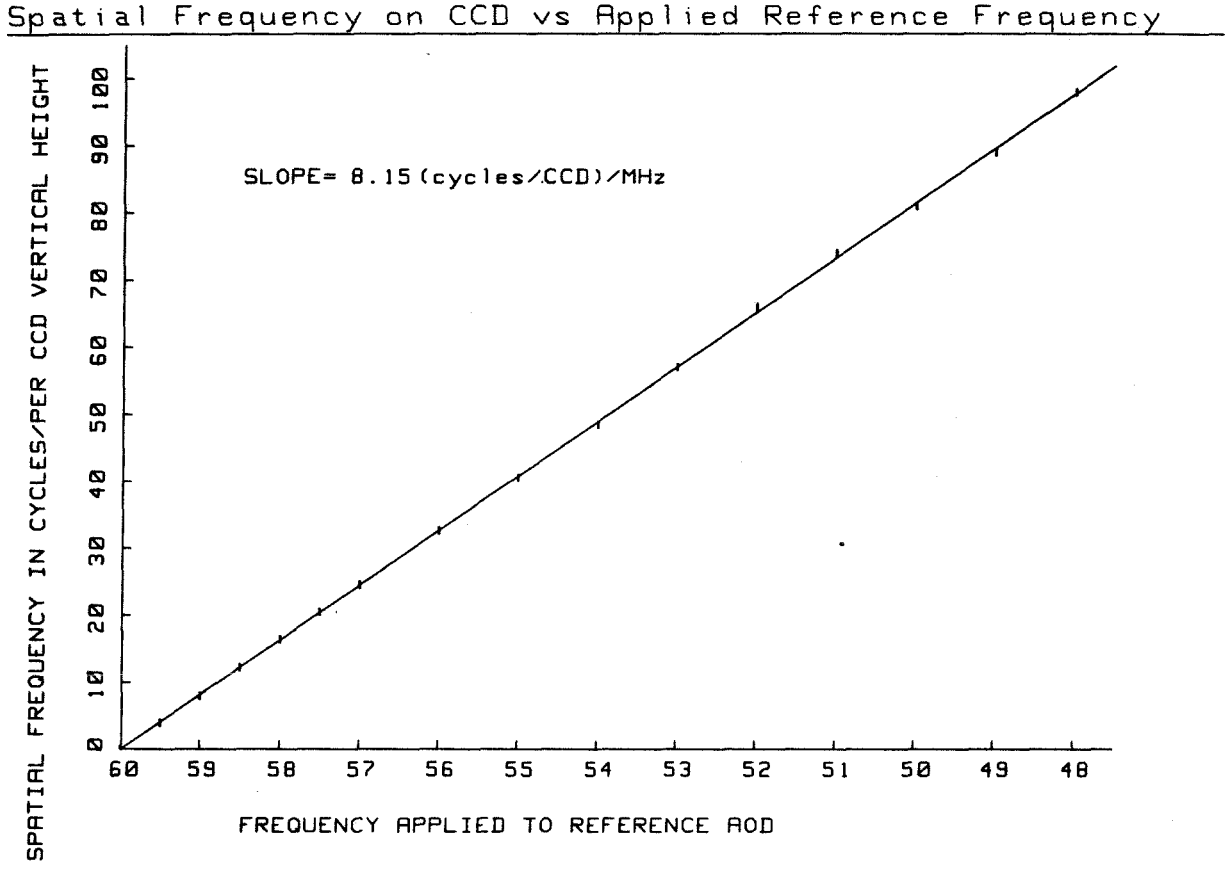


Figure 5.2.24. Linearity of the spatial frequency produced on the CCD versus the applied reference frequency.

5.2.9 Analog reference signal DFT experimental results

The first approach to implementing the folded spectrum processor relied on the analog generation of the reference waveform. This was based on a the use of a phase resettable voltage controlled oscillator (VCO) chip, that could have its phase reset to zero under external control. The coherent clock 10 MHz reference signal was digitally divided by 100 to make the 100 kHz laser diode trigger clock. This was applied to a 10 bit digital counter whose 10 outputs were applied to a 10 bit digital to analog (D/A) converter, producing a staircase stepped ramp, with 1024 segment of 10 μsec each, for a total duration of 10.24 msec. Upon reaching the maximum counter state, the counter was disabled until it received a field synchronization signal from the CCD detector, and the laser diode was only pulsed for the 1024 pulses while the counter was active. The synchronized ramp and video vertical sync pulse are shown in Figure 5.2.25. The staircase waveform was applied to the frequency control input of the phase resettable VCO, and the phase is reset by a 400 nsec wide laser diode trigger pulse. Thus the circuit produces a stepped frequency, phase reset reference waveform with reference segments of duration 9.6 μsec, separated by 10 μsec. This circuit could produce oscillation frequencies up to 85 MHz, with bandwidths as large as 1/3 of the center frequency, but when high frequencies were utilized the waveform coherence suffered drastically. Because of this limitation the VCO was operated at lower frequencies, usually sweeping from 20 MHz to 30 MHz, and mixed up to the AOD center frequency of 60 MHz with an auxiliary 30 MHz reference. By adding a reference segment to a coherent local oscillator (LO) an amplitude modulated beat can be generated that can be used to display the coherence between the VCO output and the coherent LO. When this amplitude modulated reference waveform is mixed to the AOD center frequency, applied to the AOD transducer and imaged onto the CCD an equivalent spatially magnified, detector sampled fringe pattern is generated on the CCD. An example

of a reference segments and its corresponding image is shown in Figure 5.2.26. This is a noninterferometric test of the uniformity of the fringes produced at all coarse frequency bins, and the image detected by the CCD shows that the fringes are uniform orthogonal to the reference AOD.

In order to test the operation of the time integrating DFT processor with the reference waveform generated by the VCO, the time integrating arm of the processor was operated independently by adding the signal to the frequency stepping reference, mixing to the AOD center frequency and applying to the transducer. The spatial carrier frequency that the DFT rides on is given by the difference between the reference function center frequency and the coarse frequency component of the input signal, scaled by the optically magnified acoustic velocity. The CCD time integrates two successive DFT transformation fields per frame, then reads them out as a sampled video waveform, turning the spatial carrier back into a temporal carrier with the CCD spatial readout velocity as the conversion factor. The optical system was designed with a magnification of 1.25 so that the image of the 9.6 μsec reference waveform, which has a spatial extent of about 6 mm, would be imaged onto 7.4 mm of the 8.8 mm detector width, thereby utilizing 324 of the 384 available horizontal resolution elements. In this experiment I chose to detect the time integrated transform along the horizontal dimension of the CCD, in order to perform real time bandpass filtering of each video raster line read out of the CCD. A set of typical output video horizontal line traces for different fine frequency inputs is shown in Figure 5.2.27. The performance of the active bandpass filter of the fast horizontal video was not satisfactory, because of the large amount of ringing produced by the video timing signals and by the time integrated peaks themselves. However, it did demonstrate that real time carrier demodulation could be performed on video rate signals using active filter technology, effectively removing the bias, albeit at an expense of a huge decrease of resolution due to the filter ringing. The bias pedestal

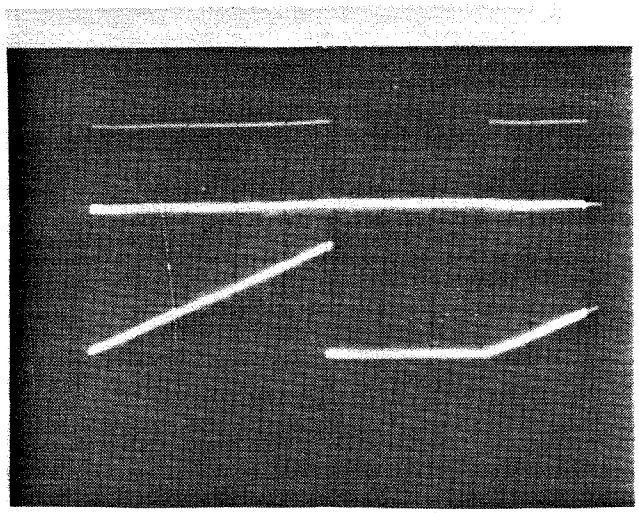


Figure 5.2.25. Video synchronized staircase ramp, and region where laser diode is pulsed 1024 times each video frame.

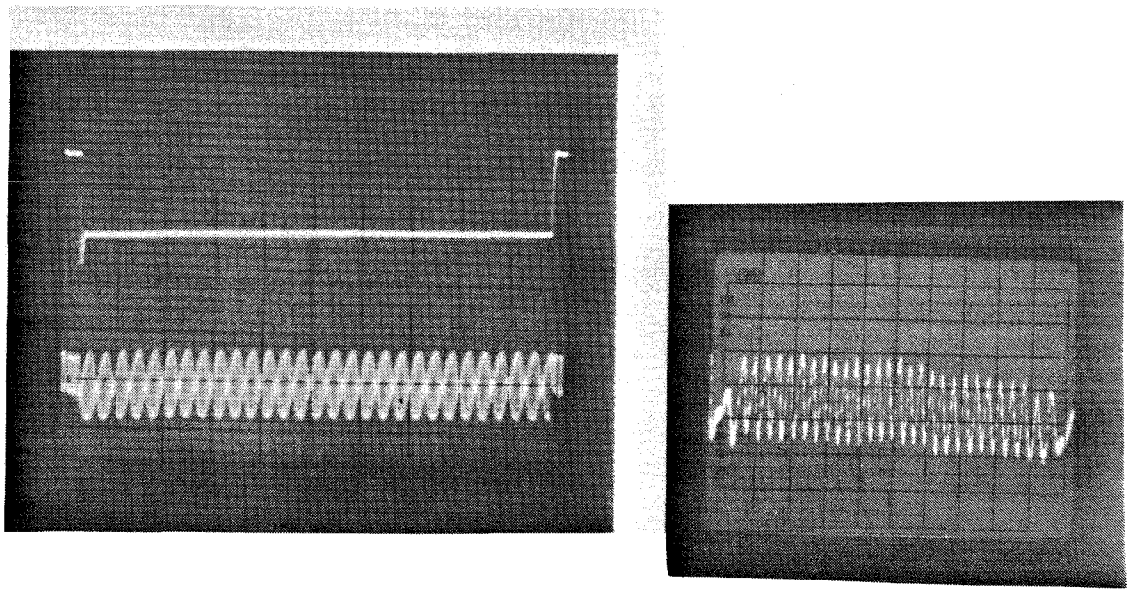


Figure 5.2.26. An example 10μsec reference segment produced by the VCO and its image on the CCD.

of the video output drops away near the edge of the output traces, indicating the region of processor analysis corresponding to the image of the reference segments. As the fine frequency is varied the time integrated peak on a carrier indeed moves from the left to the right as desired, and the linearity of the peak position for various input fine frequencies is shown in Figure 5.2.28. In this experiment the peak position deviates from the actual linear relationship by less than 4% . At a fine frequency of 120 Hz the constructive peak, has good modulation depth, and has moved away from the phase reset edge position, which represents 0 fine frequency, by about a half width. At a fine frequency of 240 Hz the peak, destructive in this photograph, has moved over by about its width, indicating that a resolution of 120 Hz has been achieved. At a frequency of 600Hz the peak has moved over farther, but appears to be smaller. In the bottom trace the peak has moved more than half way across the CCD, but has broadened considerably, and decreased in amplitude. As the fine frequency is increased the peak amplitude decreases, and the width increases markedly, thereby decreasing the resolution. This appears to be due to the analog nature of the reference waveform generation, which introduces a lack of reference function coherence.

Two types of reference waveform coherence degradations were encountered, the first was due to noise and interference on the frequency control input of the VCO, and the second was due to nonlinearities of the staircase produced by the D to A and of the VCO voltage to frequency conversion. This can be understood by considering each pixel of the CCD as a time integrator across the successive reference segments. At the phase reset position at the edge of the CCD, the phase is set to 0 for all frequencies and the processor integrates up a narrow peak for a DC fine frequency input. At some position T_0 in the middle of the reference segment, the reference phase is given by the integral of the frequency from the phase reset position to the position T_0, and if the VCO frequency has a statistical variation during each

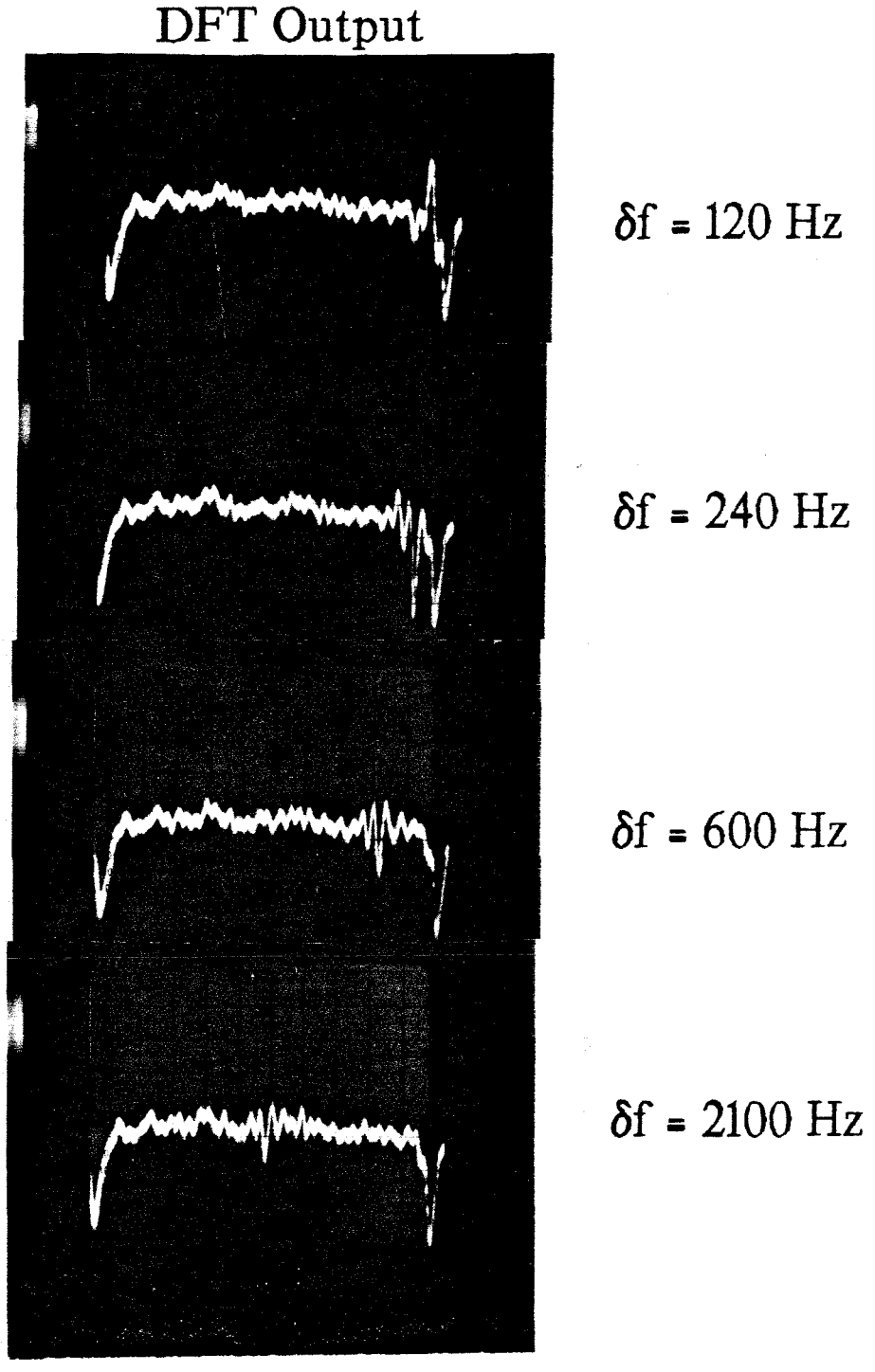


Figure 5.2.27 A typical set of horizontal fine frequency video output traces for various input fine frequencies, note broadening of peak for higher frequencies.

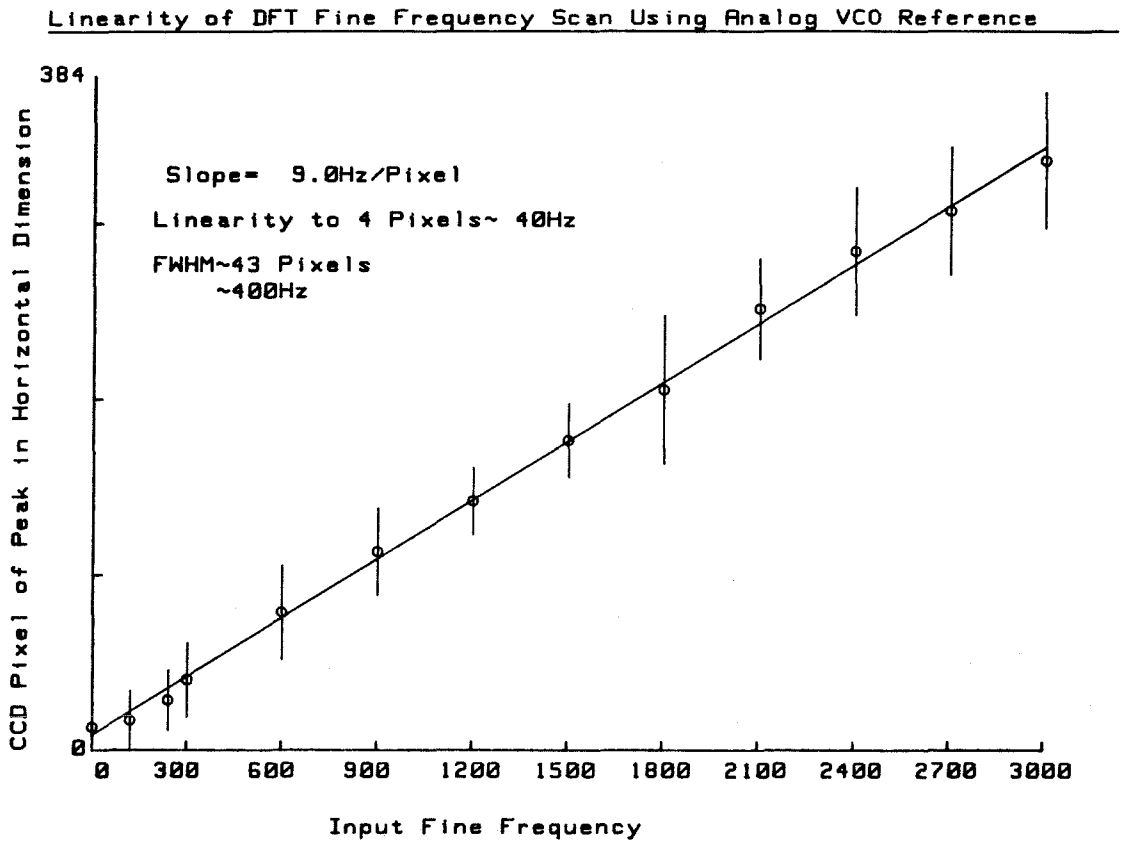


Figure 5.2.28. Linearity of peak position versus input fine frequency, width of bars equals FWHM after bandpass filtering.

reference segment, and from segment to segment, then the variance of the phase will integrate up, becoming larger for larger T_0. Thus the processor is integrating across rows of the DFT matrix where the phase is a statistical function with a variance that increases for the higher frequency rows. The nonlinearity of the frequency sweep compounds this problem by including a small chirping across the rows, so the kernel implemented by the system becomes $e^{i2\pi m(n+\alpha n^2+\cdots+\sigma)}$, where the quadratic nonlinearity has a coefficient $\alpha \ll 1$, and the variance of the frequency of the reference segments is σ. For small m the nonlinear and statistical terms will introduce small phase variations that can be ignored if they are much less than π during the time of integration. But for large m these phase variations will significantly affect the performance of the processing operation. Because of these limitations of the analog reference generation method I decided to use a digital technique in order to generate the reference waveform, and hopefully avoid the problems with reference stability and coherence.

5.2.10 DFT processing using a video frame buffer for reference generation

The desired reference waveform should consist of between 1000 and 3000 constant frequency segments, each with a duration of 10 μsec containing about 1000 samples read out at a data rate of 100 MHz, yielding a Nyquist limited analog bandwidth of 50 MHz. Each reference segment should contain 1000 samples, and the entire reference waveform should contain at least 10^6 samples, and have an overall time bandwidth product of at least $5x10^5$. These are very demanding specifications, and cannot be readily achieved with commercially available hardware. For this reason I chose to use a standard digital video frame buffer memory with 512x512 resolution and 10 MHz sample rate as our digital reference generator, in order to test the system concepts being investigated in this thesis. The limited 5MHz bandwidth of this frame buffer, and its video line rate format, will result

in a limitation of the analysis bandwidth of the time integrating DFT processor that is well below the coarse frequency resolution of the space integrating spectrum analyzer. By using a larger area CCD, or changing the magnification of the optics one could increase the time integrating analysis bandwidth, but not enough to equal the coarse frequency resolution. The experimental folded spectrum data to be presented in this section will thus be only a small part of the full folded spectrum, but this should be enough to demonstrate the principles involved and to explore the limitations of this signal processing technique.

In order to use the video frame buffer as the reference function generator, it is essential that the laser diode be timed to pulse in synchronism with the video scan lines. This can be accomplished by using the 15.7354 kHz composite blanking video timing signal as the laser diode trigger pulse. In this case the coarse and fine frequency components are computed modulo this video line rate, or more accurately speaking the coarse and fine frequency offsets from the carrier frequency $f_0 = 60$ MHz are computed modulo the video line rate. It is also convenient to have the time integrating CCD in frame synchronism with the video frame buffer, so that a single frame of processing is an integration over the full reference function. This can all be accomplished by synchronizing all of the system components as indicated in Figure 5.2.29. The carrier frequency for the reference function need not be synchronized with the video signals, but it should be coherent enough to avoid introducing unwanted phase modulations.

An idealized DFT mask produced by the frame buffer is shown in Figure 5.2.30, the 2-D format of the reference function is raster scanned to produce a baseband modulated video waveform. This baseband video waveform needs to be mixed up to the AOD center frequency, and the mixing operation is illustrated in Figure 5.2.31. The mixer has a low frequency cutoff, and suffers from ringing produced by the video sinc signals, and it produces a double sided modulation with some unwanted

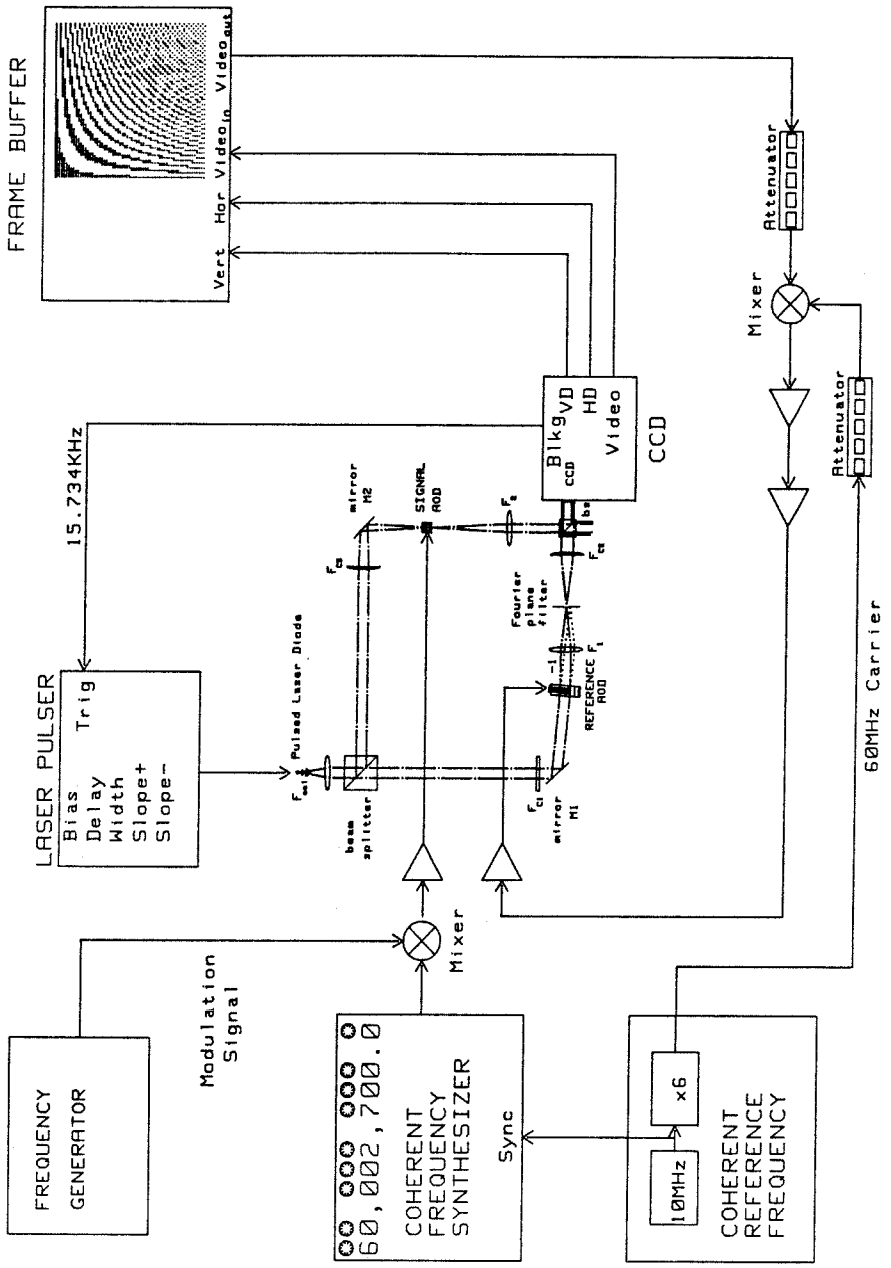


Figure 5.2.29. Synchronization needed between the video frame buffer, the CCD camera, the pulsing laser diode, the coherent clock and the coherent signal generator.

a)

b)

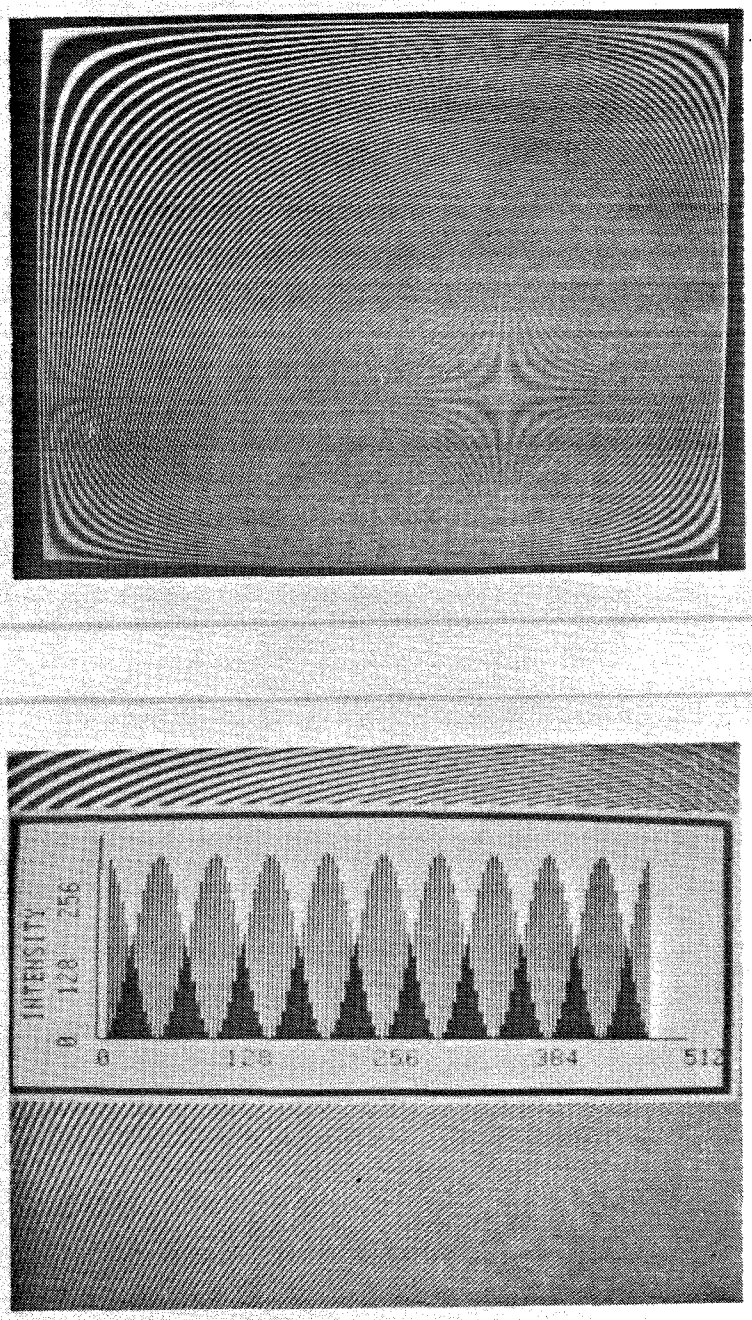


Figure 5.2.30. Idealized DFT mask produced by the frame buffer, with no scale factors or interleaving taken into account, and a cross section through a single horizontal line.

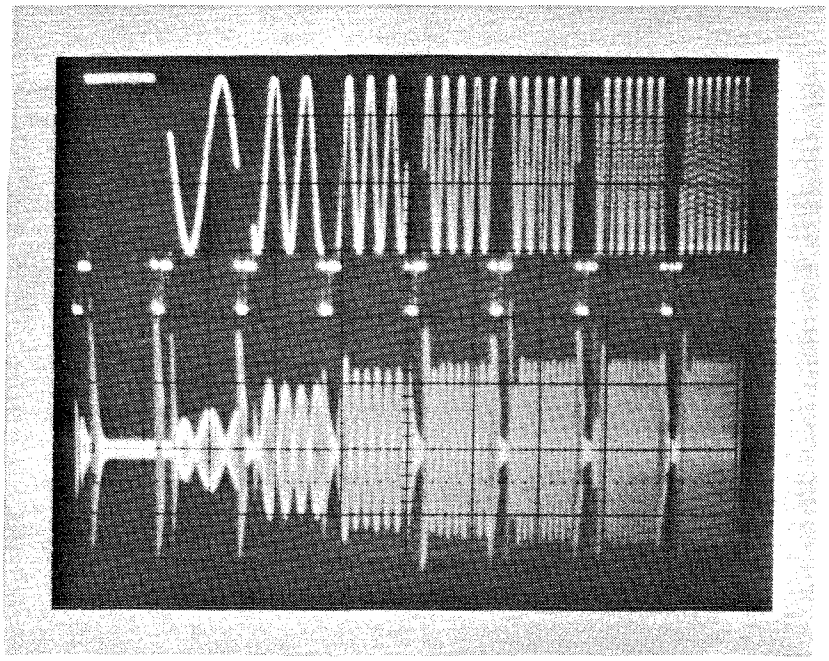


Figure 5.2.31. An example of the first few lines out of the frame buffer, and the resulting signal after mixing with 60 MHz.

a)

b)

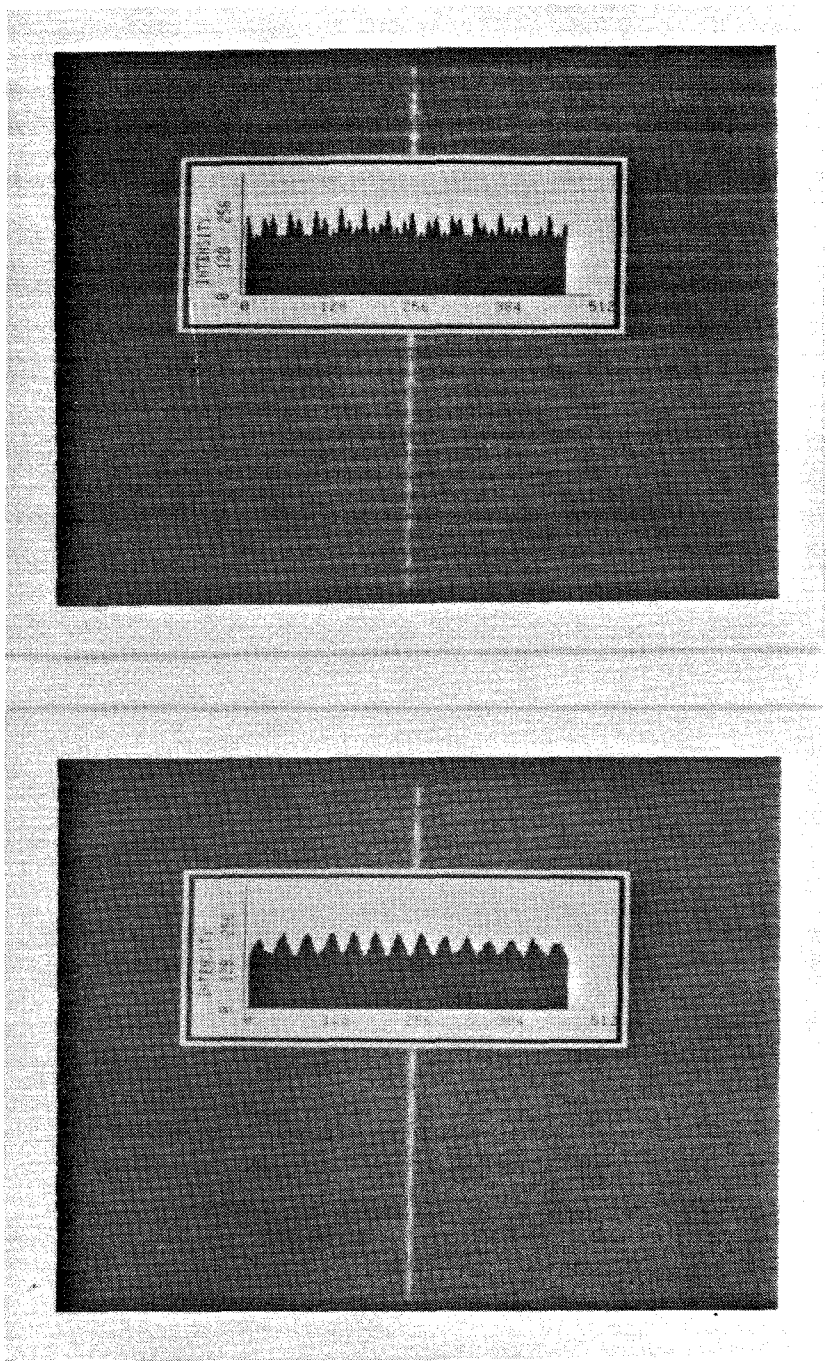


Figure 5.2.32. a) Video image and cross section through ridge produced when Fourier plane filter was removed allowing both sidebands and carrier to interfere on CCD. b) Video image and cross section through interferometric ridge with proper reference Fourier filtering.

carrier feedthrough. These effects can be eliminated by careful positioning of the reference arm Fourier plane filter, in order to completely block one sideband and the carrier. The importance of this Fourier plane filtering of the video amplitude modulated RF carrier is illustrated in Figure 5.2.32a, where the effect of allowing both sidebands to fall on the detector can be seen to produce unwanted modulations due to the reference arm alone, as well as additional interferometrically produced spatial beats along the ridge due to the carrier feedthrough and unwanted sideband. The properly filtered interferometrically produced sinusoidal modulation is illustrated in Figure 5.2.32b, where the reference beam alone produces no unwanted modulations, and the modulation along the ridge is an excellent approximation to a single spatial frequency. It was found that the carrier feedthrough was strong enough that even the first few spatial sidelobes had to be filtered, thereby losing signal bandwidth of the reference function lines that had a frequency near DC. It was interesting to note that each successive sidelobe caused a different spatial frequency interferometric fringe pattern to be produced on the detector, all moving at the same number of spatial fringes per second, which means the spatial fringe velocity was different for each sidelobe. This is because the sidelobes were at shifted positions in the reference Fourier plane, as if they were due to different reference frequencies, but they all oscillated at the same sampled fine frequency. Because of this effect, apodization of the reference AOD should be used in order to decrease the sidelobe amplitude, and improve the spatial reference bandpass filtering operation.

The details of the interleaved fields must be appropriately compensated for in the video frame buffer reference function programming. A video frame consists of two successive video fields, referred to as odd and even fields, which are interleaved by the video monitor in order to form a conventional TV image. However, they will occur sequentially in the time-space raster format, with a video blanking interval between them, during which the laser diode will not be pulsed. This transformation

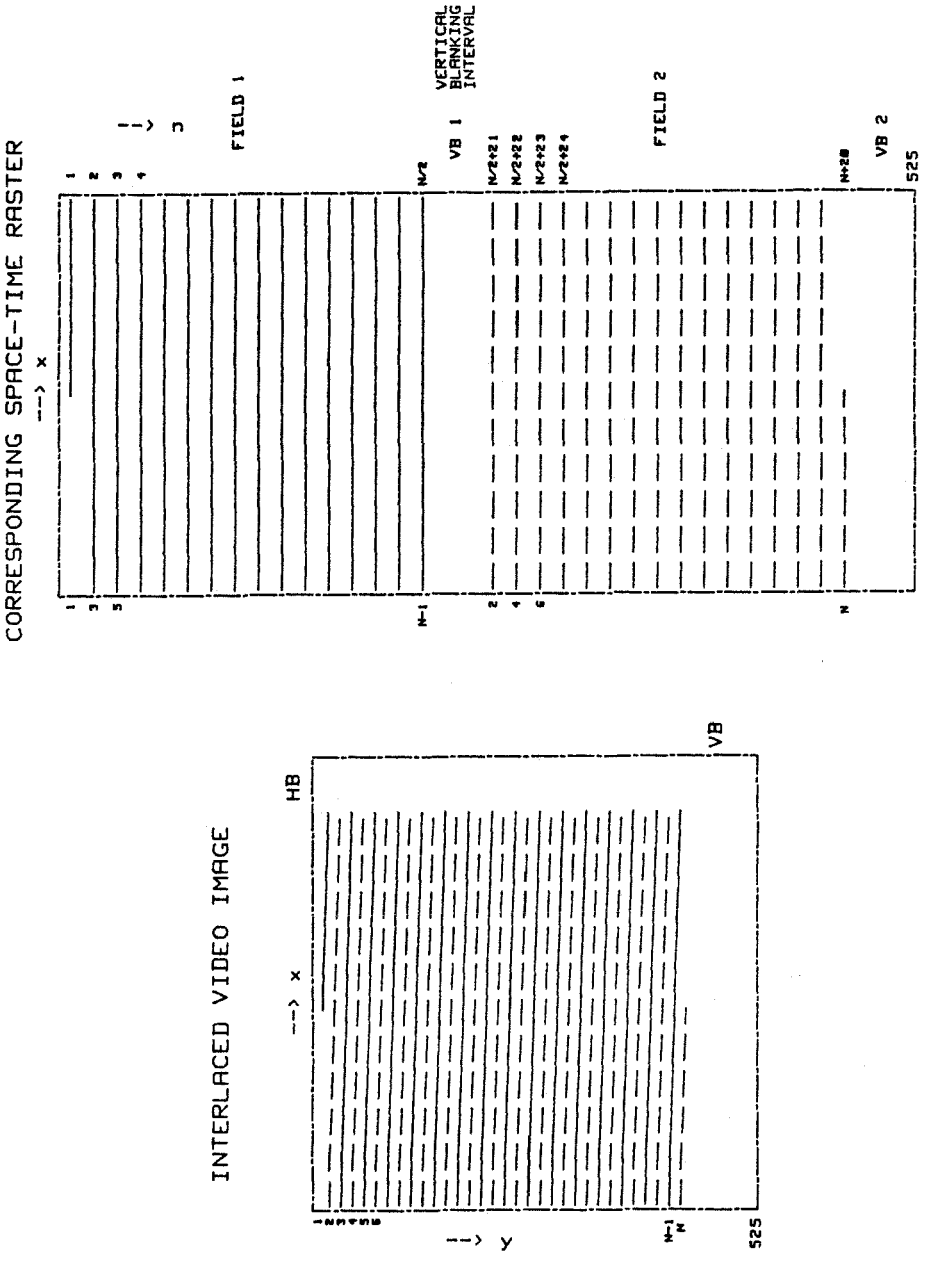
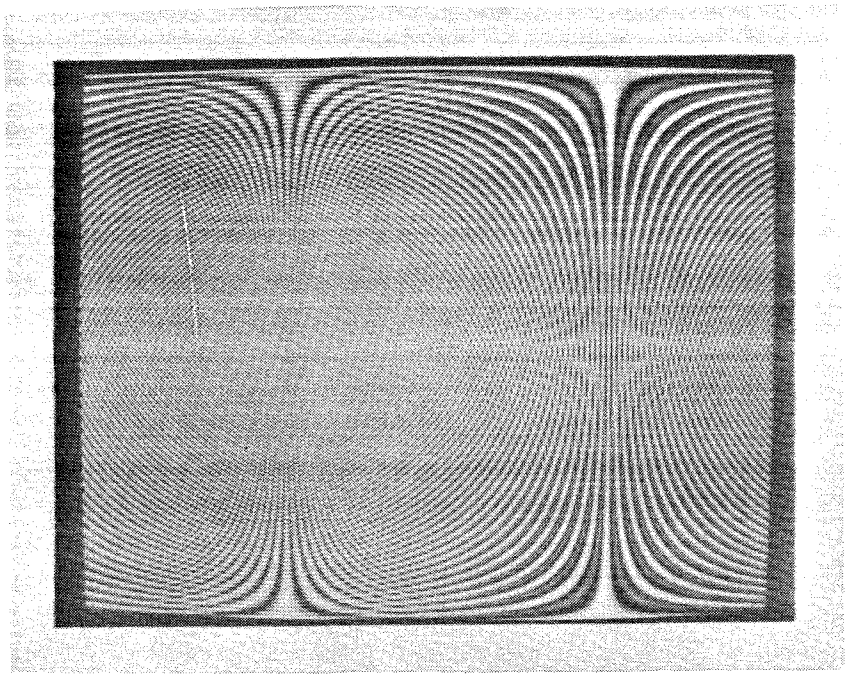


Figure 5.2.33. Transformation between interleaved field video format, and sequential field space time raster with vertical blanking dead time between fields.

a)



b)

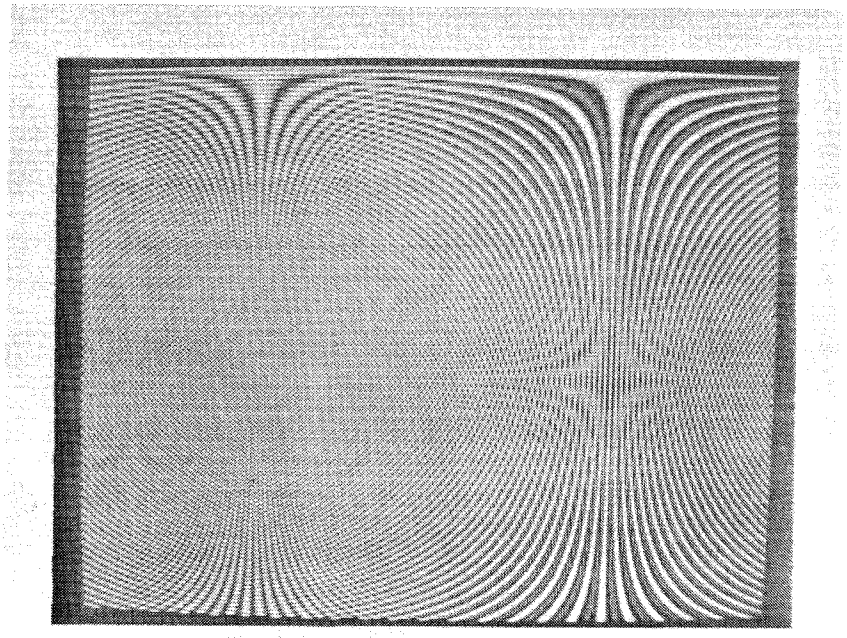


Figure 5.2.34. Video images of DFT reference masks used in these experiments a) vertical blanking interval ignored. b) appropriate scaling to compensate for dead time during vertical blanking interval.

from standard interlaced video format to sequential field time space raster format is illustrated in Figure 5.2.33. The representation of the DFT kernel in the frame buffer should be a properly sampled DFT matrix in the de-interleaved format of the space time raster, so when it is displayed on an interleaved monitor adjacent scan lines from the two fields will not appear continuous. The video format consists of 525 scan lines per frame, which are divided up between the two fields, with 262 lines in field 1, and 263 lines in field 2, which are out of phase by one half of a video line time from those in the first field. The laser diode is pulsed at a correspondingly half line phase shifted time during field two, thereby spatially aligning the scan lines during the two fields, but sampling the temporal interferometric beats at a phase shifted time, which can alter the results for very high fine frequency modulations. Typically the last 20 horizontal scan lines of each field contain no information and this blanking time is used for vertical retrace. This dead time must be compensated for with appropriate phase shifts of the DFT rows during the vertical blanking interval even though the laser diode is not pulsed during this time, because the sampled fine frequency temporal oscillations continue to phase advance during this blanking time. An example of a DFT mask that ignores the vertical blanking interval, and one that compensates for this blanking time are shown in Figure 5.2.34, as interleaved images taken off of a video monitor. The DFT operation that is implemented during an entire field of integration has the center 4% of the rows blanked out by the vertical blanking interval, but this will not affect the processor operation very much. An additional complication arises due to the nature of video time integration, where adjacent fields of video output, separated by 1/60 second, have performed half overlapping 1/30 second time integrations, so the frame buffer reference should be continuous between frames as well as between fields. Each scan line contains 52 μsec of video data, and 12 μsec of horizontal blanking before the next line begins, however in the experiments reported here I only utilized about

8.5 μsec of the available video line, because of the acoustic velocity and optical magnification scaling factors. All of these scale factors must be included in the digital program used to generate the space time reference raster. Additionally, the frequencies near DC should be avoided to facilitate Fourier plane carrier feedthrough filtering without filtering the reference function information, and the frequency near the Nyquist frequency of the horizontal frame buffer sampling should be avoided to minimize the effect of frequency fold over that occurs about the Nyquist frequency. All of these considerations are properly taken care of in the DFT reference function shown in Figure 5.2.35, where the DC and Nyquist frequency regions have been translated into the middle of the vertical blanking interval. The 8.5 μsec portion of this mask that is imaged onto the CCD is shown in its interlaced format and in its space time raster format, and a vertical cross section is shown that demonstrates the phase continuity between fields and between frames.

A useful method of testing the reference waveform was to apply it to the signal AOD, while the CCD was operated in the TDI mode in the direction orthogonal to the signal AOD, with a very narrow slit placed in front of the CCD. This produces a video output which is a mapping of the instantaneous frequency produced on each laser pulse versus pulse number, as a 2-D format, as illustrated in Figure 5.2.36, where the instantaneous frequency is clearly seen to be a linear function of time, or pulse number, as is desired. The successive video frames are displayed in an interleaved format, because the TDI mode of operation produces one TDI scan per field and the monitor interleaves these fields. If this frequency versus pulse number representation were deinterleaved, and stacked on top of each other with the appropriate vertical blanking interval between them, a linear relationship between frequency and pulse number would be observed throughout the entire frame. The simple space integrated spectrum of the fully optimized reference function shown in Figure 5.2.35 is shown in Figure 5.2.37, where the DC component and the Nyquist

a)

b)

c)

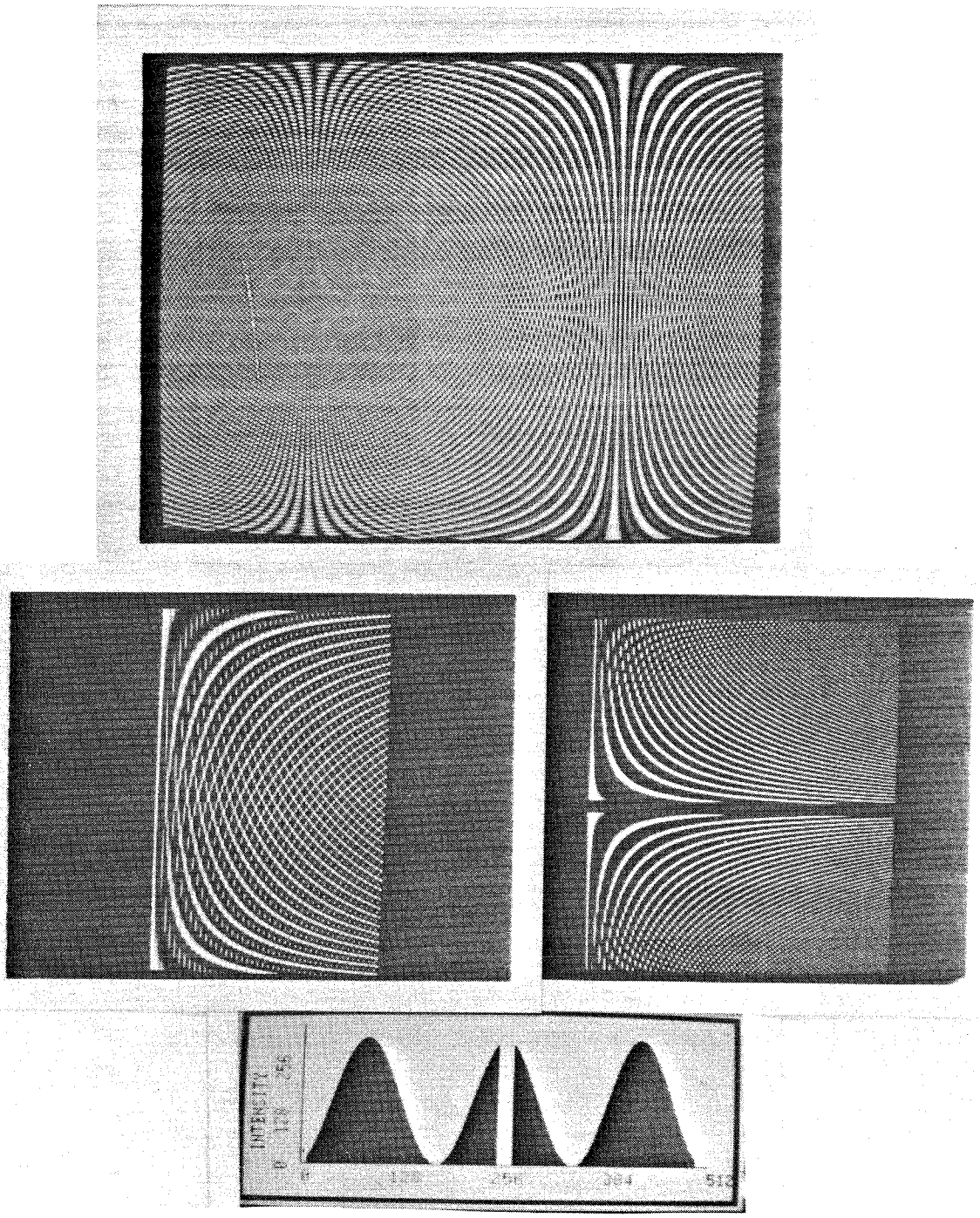


Figure 5.2.35. a) Fully optimized DFT reference mask displayed on interleaved video monitor with single sided scaling and frequencies near DC and Nyquist translated into the vertical blanking interval. b) The 84 pixel section that is imaged onto the CCD in its interlaced format and in its space time raster format, c) a vertical cross section showing continuity between fields and frames.

frequency component are seen to be suppressed, but the above Nyquist fold over terms are apparent. Both sidebands are clearly evident, and the symmetric Bragg frequency response roll off is also visible, proper reference arm Fourier plane filtering must be used to completely remove one sideband and the above Nyquist terms, and this is facilitated by the lack of scan lines at these frequencies.

Folded spectrum data produced by the time integrating CCD were taken using the same video frame buffer that was used to generate the reference waveform. As the reference waveform was read out, the pixels accumulated on the previous frame of integration were shifted out of the CCD, and these analog values were digitally sampled and stored in the frame buffer. This facilitated digital post processing operations that needed to be performed on the time and space integrated folded spectrum data in order to remove the unwanted bias terms. Two generic approaches to bias removal were investigated, subtraction of constructive and destructive frames, and spatial carrier demodulation.

In the subtraction of constructive and destructive frames approach, the beam splitter was rotated in order to remove the spatial carrier producing an interferometrically detected peak which beat up and down without producing a travelling fringe structure under its envelope. It was necessary to observe the data, and grab a frame with the frame buffer as the phase of the input tone drifted at its super fine frequency between producing a constructive peak and a destructive one. This was only possible with a single tone, because multiple tones would achieve maximum constructive and destructive peak size at arbitrary times. If two CCD detector arrays were used on both output ports of the recombining dielectric beam splitter cube, then a real time bias subtraction technique could be implemented. This would require that the phase shift upon reflection from the multilayer dielectric surface would be 90 degrees, and this is often a good approximation to the actual phase shift behavior. In this case the relative phase of the interferometrically produced

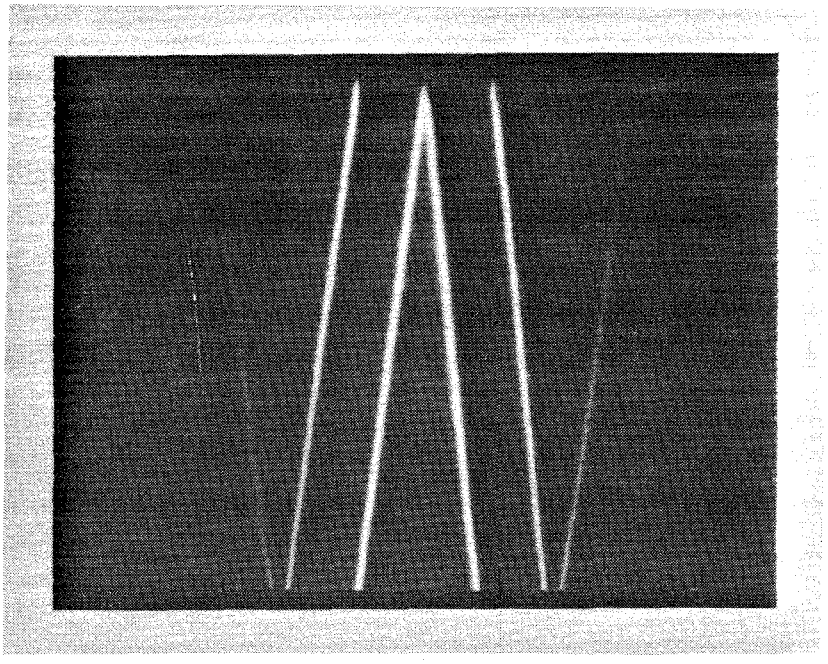


Figure 5.2.36. Frequency versus line number display on interlaced monitor of the DFT reference mask shown in Figure 5.2.35.

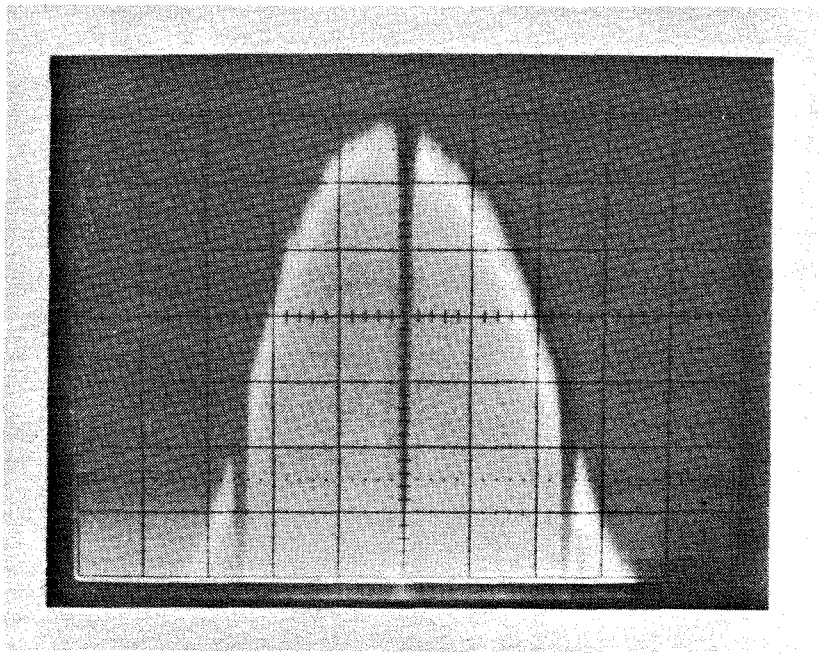


Figure 5.2.37 Simple space integrated power spectrum of the DFT reference waveform shown in Figure 5.2.35.

modulation would be 180 degrees, so that the video outputs from two synchronized CCDs could be directly subtracted to produce a bias free folded spectrum. It was in order to simulate this type of two CCD bias subtraction operation that the successive frame bias subtraction technique was investigated, not because successive frame subtraction is a viable system approach to bias removal in a multitone signal environment. The video reference used in these experiments was a double sided mask, shown in Figure 5.2.34, that had the phase reset position in the middle of the CCD, so that any single tone input would produce two fine frequency peaks symmetrically about this central position, and positive and negative fine frequencies were indistinguishable. This mask was chosen in order to be able to measure the phase uniformity of the fine frequency processing at different spatial positions, if one of the two symmetric peaks was significantly smaller than the other then that indicated that they were out of phase, and hence a poor phase uniformity and phase stability could be inferred. These foded spectrum images were obtained by grabbing first a constructive peak pair image and storing it in the computer memory, then a destructive peak was grabbed, and the two were subtracted. All of the bias terms should be constant between the two images, and should cancel, while the interferometrically produced folded spectrum information will be doubled in amplitude, while the uncorrelated CCD noise will grow by $\sqrt{2}$. Since the peak in these experiments was quite wide, and the remaining CCD noise was spatially white, the signal to noise ratio could be significantly improved by low passing the difference image with a sampled Gaussian convolution, thereby limiting the within band noise power, without decreasing the peak height or increasing its width too severely. The measured signal to noise ratio improved from 8:1 after bias subtraction to better than 15:1 after the 3 by 3 Gaussian convolution.

Several examples of bias subtracted folded spectrums are shown in the sequence of photographs in Figure 5.2.38, where both fine frequencies and coarse frequencies

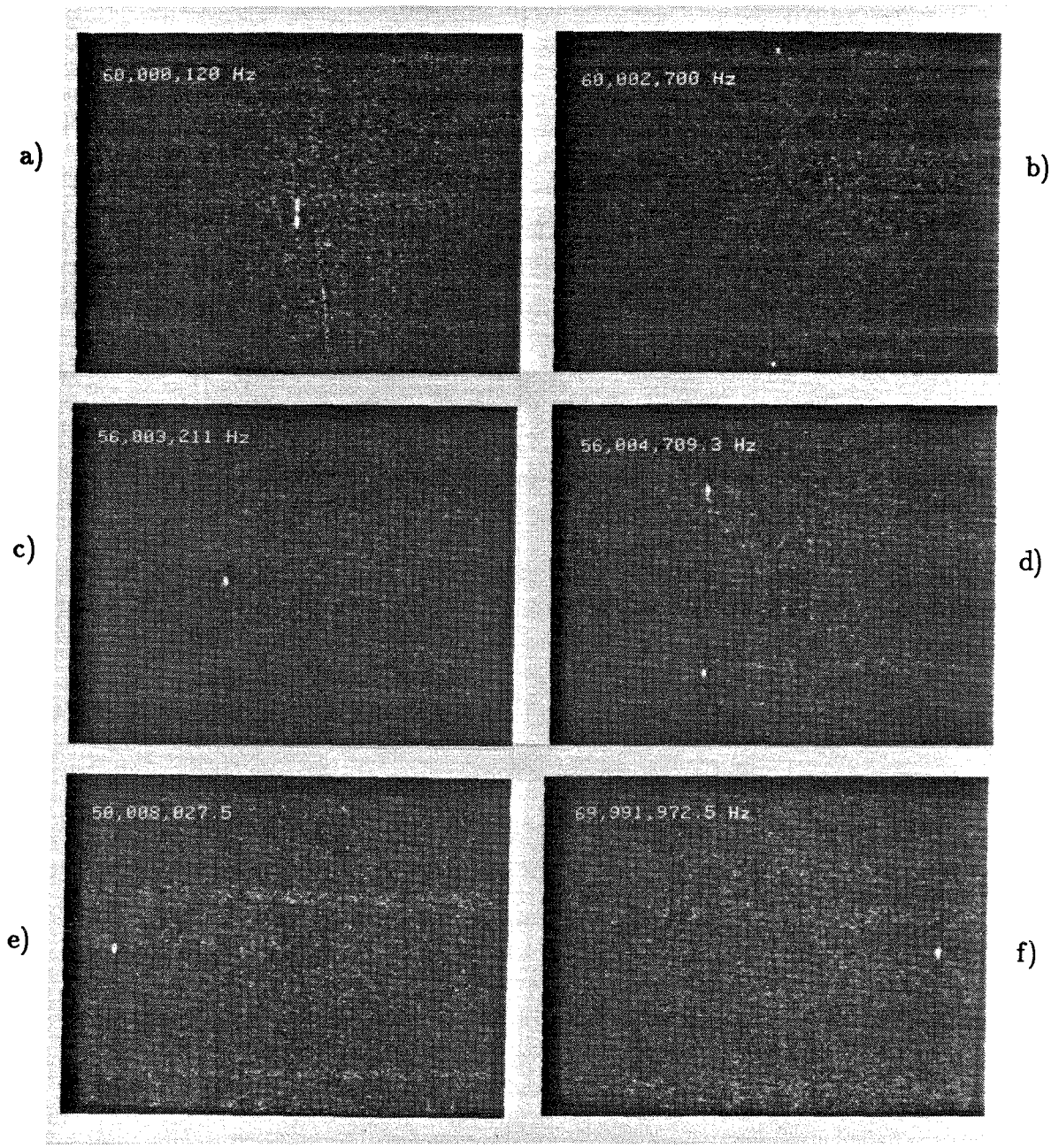
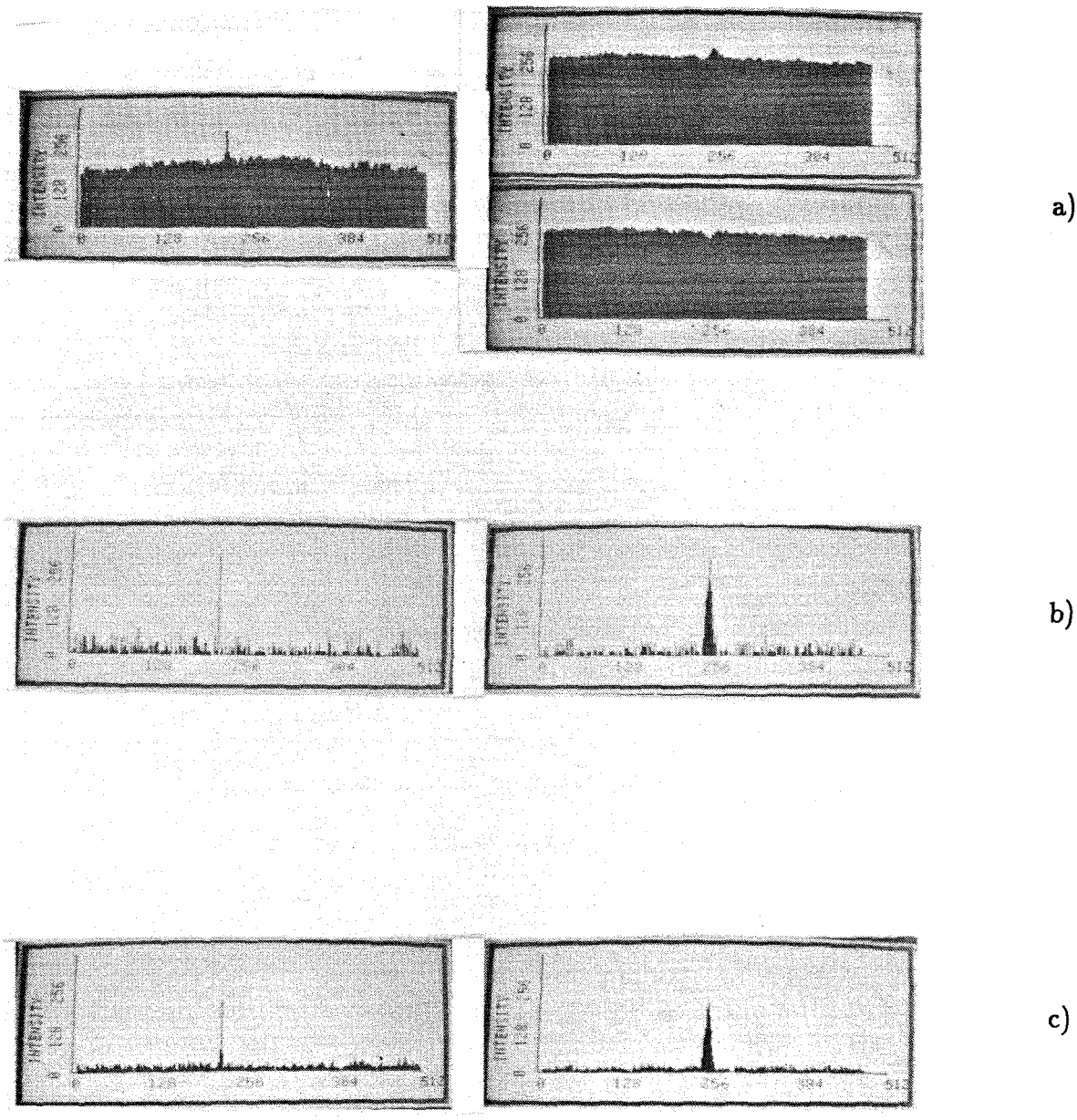


Figure 5.2.38 Double sided bias subtracted folded spectrum video images for input frequencies a) 60,000,120 Hz b) 60,002,700 Hz c) 56,003,211 Hz d) 56,004,709.3 Hz e) 50,008,027.5 Hz f) 69,991,972.5 Hz.

are varied. The fine frequency resolution can be inferred from the input tone at 60,000,120 Hz to be about 120 Hz, since the resulting peak pairs fall approximately on each others first zeroes, and the full width at half max of the peak was 11 vertical pixels which is also equal to 120 Hz. The fine frequency analysis bandwidth is 5400 Hz over the 484 vertical pixels, which can be seen by noting that the peak pairs fall on the edges of the CCD for an input frequency of 60,002,700 Hz. The sampled and interferometrically detected coarse frequency analysis bandwidth is better than $\pm$10 MHz, although the laser pulse width had to be cut down to 40 nsec, and the signals applied to the AODs had to have a compensating amplitude increased in order to obtain the data at 50,008,027.5 Hz and 69,991,972.5 Hz. These frequencies are displaced from 60 MHz by exactly 635 f_{LD}, so they produce a single peak at the 0 fine frequency coordinate. Examples of cross sections through the coarse and fine frequency axis before and after bias subtraction, and after Gaussian convolution are shown in Figure 5.2.39. From these cross sections it is clear that the bias has been successfully removed, and the coarse and fine frequency components of the single tone inputs have been measured by the time and space integrating folded spectrum technique with successive frame bias subtraction.

Cross sections through the time integrated interferometrically produced folded spectrum ridge are shown in Figure 5.2.40, for fine frequencies varying over the fine frequency analysis bandwidth for the double sided reference DFT mask of Figure 5.2.34, with the 0 fine frequency peak at the CCD edge. This reference DFT sweeps down during one field, analyzing negative frequencies, and it sweeps up during the other field analyzing positive frequencies, and this has the disadvantage that the 0 fine frequency component is contributed to by both fields and is therefore twice as strong and half as wide. When the reference mask of Figure 5.2.35 is utilized an entire frame of processing produces a single sided coherent time integration over half the bandwidth per sideband of the previous example. Cross sections through

SI coarse analysis cross section TI fine analysis cross section



a)

b)

c)

Figure 5.2.39. Cross sections through the coarse and fine axis of the folded spectrum data a) as constructive and destructive frames come out of the CCD, b) after the bias is subtracted and the resulting image amplified by 8, c) and after the 3x3 Gaussian convolution.

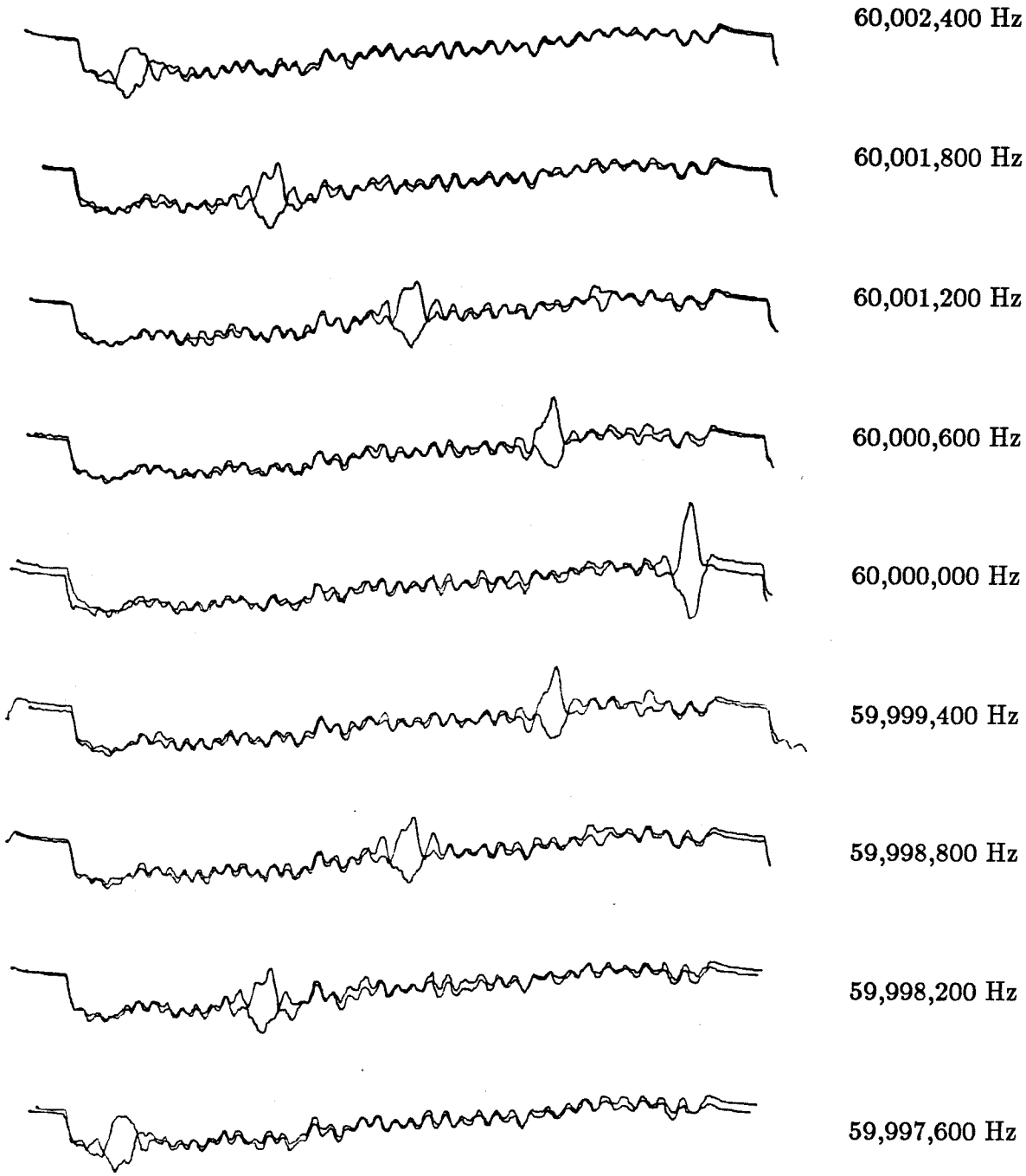


Figure 5.2.40. Cross sections through the coarse frequency ridge for various positive and negative fine frequencies with the double sided reference of Figure 5.2.34.

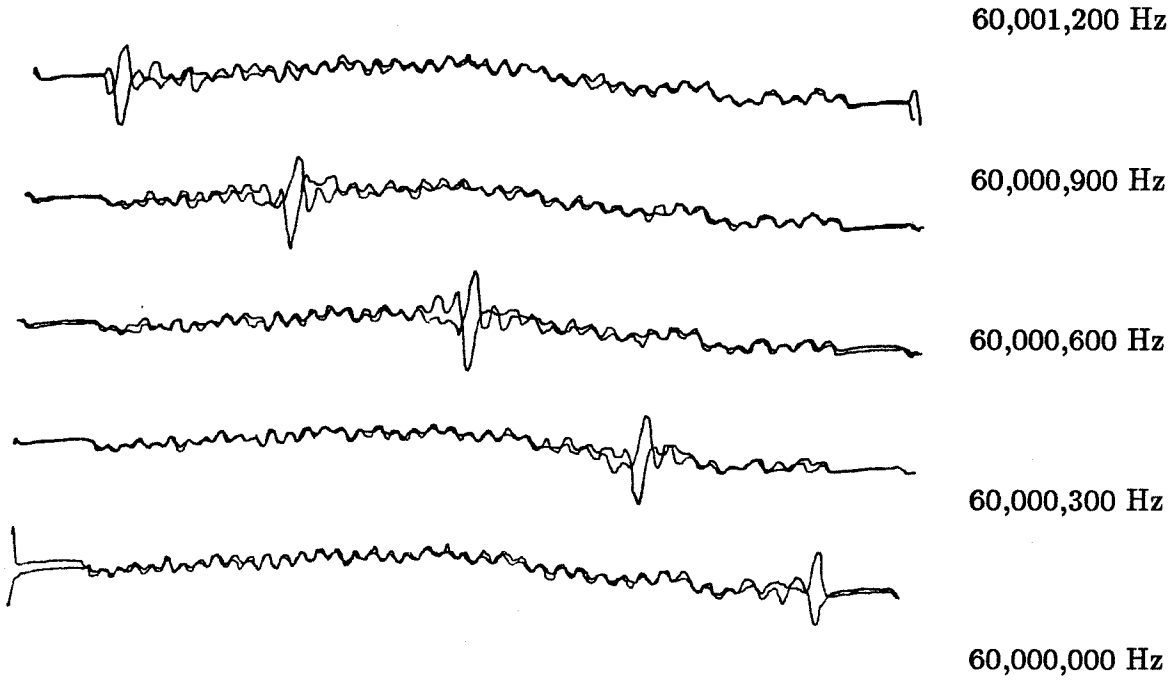


Figure 5.2.41. Fine frequency cross section of the single sided time integrating DFT processing with the reference mask of Figure 5.2.35.

the coarse frequency ridge for the single sideband of analysis are shown in Figure 5.2.41, and the resolution and gain are found to be constant over the fine frequency analysis bandwidth. In this case negative fine frequency inputs produced no time integrated peaks, however if the reference Fourier plane filter was translated to the other side of the carrier, then negative fine frequency processing occurred. When the reference Fourier plane filter included both sidebands, then both positive and negative fine frequency processing occurred, but the reference arm alone acted as a DFT processor to produce a time integrated ridge at the 0 fine frequency position in all of the coarse frequency bins.

Spatial carrier encoding of the time and space integrated folded spectrum data is accomplished by rotating the recombining beamsplitter in the plane of the interferometer, in order to produce a fringe pattern along the coarse frequency bias ridge at the position of the fine frequency peak. Cross sections through the coarse frequency ridge for various fine frequencies are shown in Figure 5.2.42, for two different beam splitter rotation angles, showing how the time integrating fine frequency analysis can be placed on an arbitrary spatial carrier frequency. When a DFT reference waveform was applied to the reference AOD, the beam splitter was adjusted to produce a fringe pattern of 4 pixels per cycle at the fine frequency locus. This spatial carrier frequency was chosen because a short and simple tapped filter with 7 delay stages could be used to filter the image, and produce a bias suppressed image. The zero mean unapodized filter which was convolved with the video image stored in the frame buffer in order to remove the bias was oriented along the coarse frequency bias ridge in the vertical dimension of the CCD, and had tap weight coefficients given by [1 0 -1 0 1 0 -1]. A typical high frequency spatial carrier which was interferometrically produced when 60 MHz was applied to both AODs is shown in Figure 5.2.43a, as a video image. When a cross section along the ridge of every other pixel is plotted as in Figure 5.2.43b, the spatial carrier can be seen to be

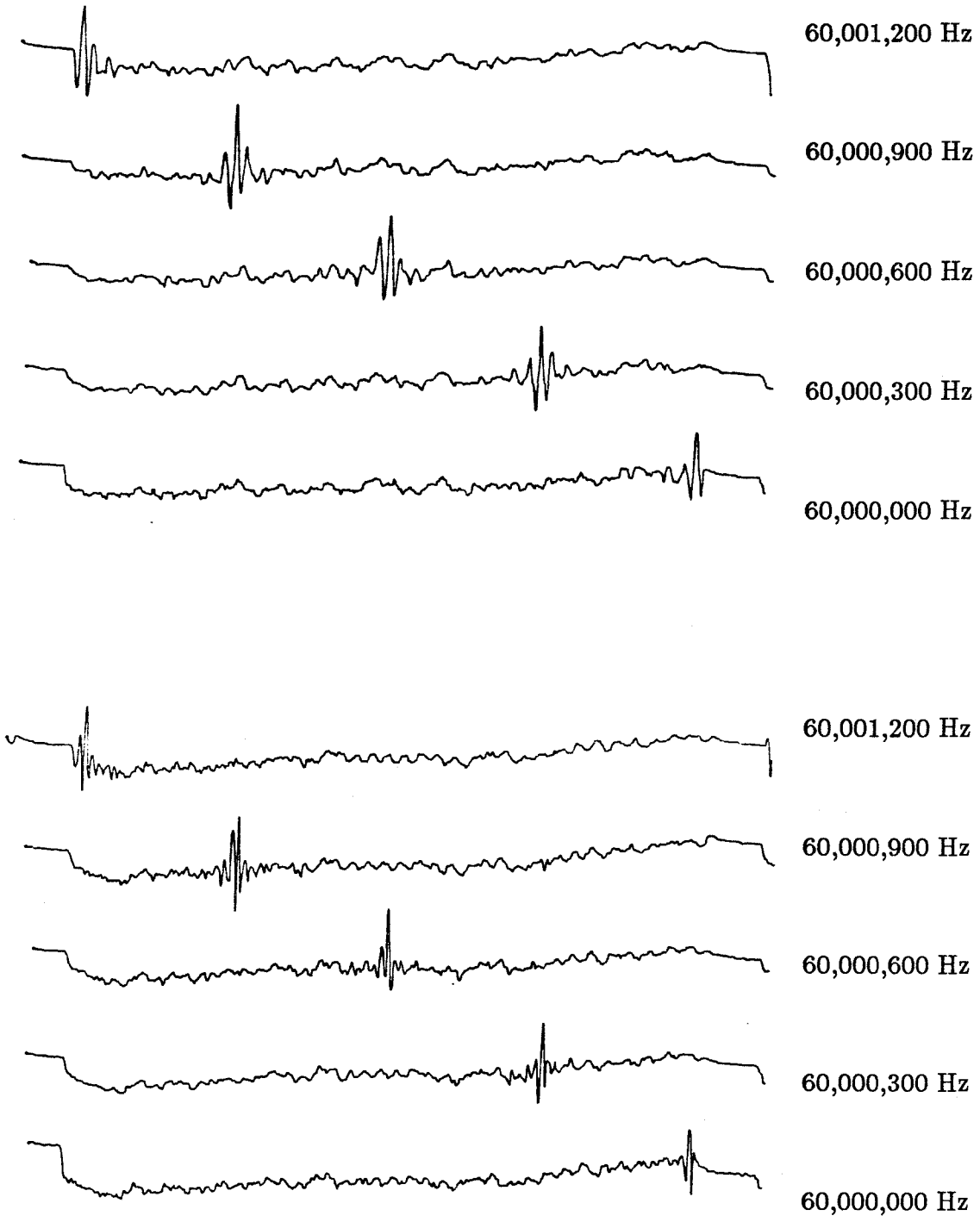


Figure 5.2.42 Fine frequency cross sections for various input frequencies for two different spatial carrier frequencies corresponding to different beam splitter rotations.

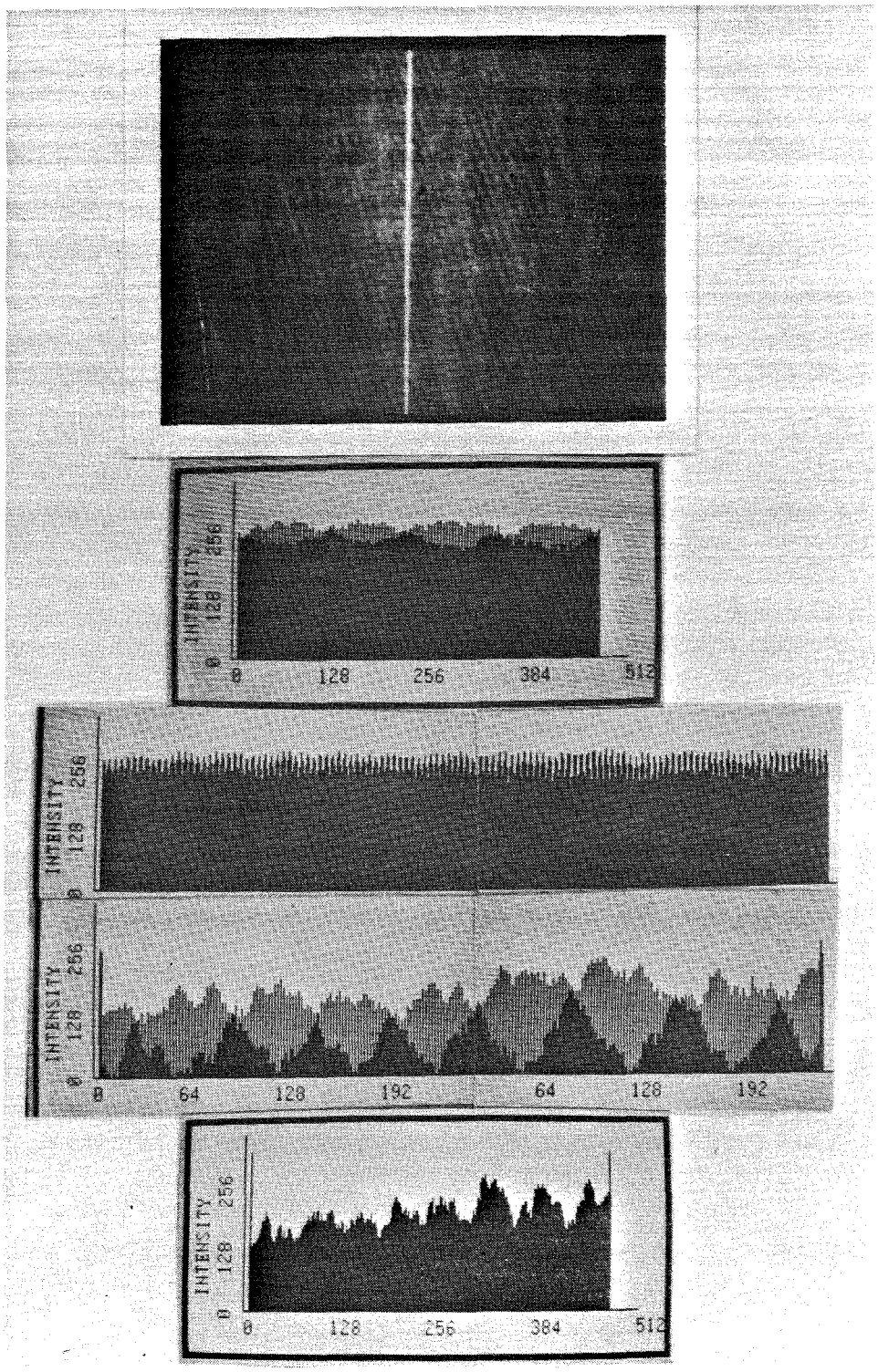


Figure 5.2.43. a) Interferometrically produced high frequency spatial carrier as a video image, b) as a cross section of every other pixel, c) as a cross section through every pixel, d) after band pass filtering and rectification, e) after low pass filtering.

approximately 4 pixels per cycle. A cross section of every pixel is shown in Figure 5.2.43c, and the resulting cross section after band pass filtering is shown in Figure 5.2.43d, which can be seen to have removed the bias and rectified the filtered carrier. Low pass filtering of the rectified and band pass filtered image will demodulate the entire ridge on a carrier with a radiometric accuracy that will depend on the relative phase of the carrier with respect to the pixel structure that can vary by $\sqrt{2}$, and this is shown in the low passed output shown in Figure 5.2.43e.

A typical single tone image grabbed by the frame buffer is shown in Figure 5.2.44a, where a small region containing a high frequency fringe pattern can not be seen along the bias ridge, because of the low contrast of the reproduction. A cross section through the bias ridge shows the localization of the time integrated peak on a carrier in Figure 5.2.44c before filtering and after filtering and rectification is shown in Figure 5.2.44d, and the corresponding image produced after the filtering operation is shown in Figure 5.2.44b. Typically about 4 or 5 cycles of the 4 pixel per cycle carrier were produced under the full width of the time integrated peak, so the spatial frequency spectrum of the peak was narrow enough to be easily separated from the bias. The bipolar filtering operation of the unipolar video image in the digital frame buffer produces a zero mean bipolar bias removed output image that was then rectified by a digital absolute value operation, which was simply accomplished by dropping the sign bit. This effectively doubles the carrier frequency of the filtered image to 2 pixels per cycle, and produces a baseband component of the rectified peak, which allows a succeeding low pass Gaussian convolution to produce a fully demodulated folded spectrum amplitude image. A cross section along the coarse ridge produced after low pass filtering the band passed and rectified image is shown in Figure 5.2.44e, and this has removed the frequency doubled carrier that results from the rectification operation. As the phase of the input tone drifts with respect to the system clock frequency, the phase of the carrier linearly drifts

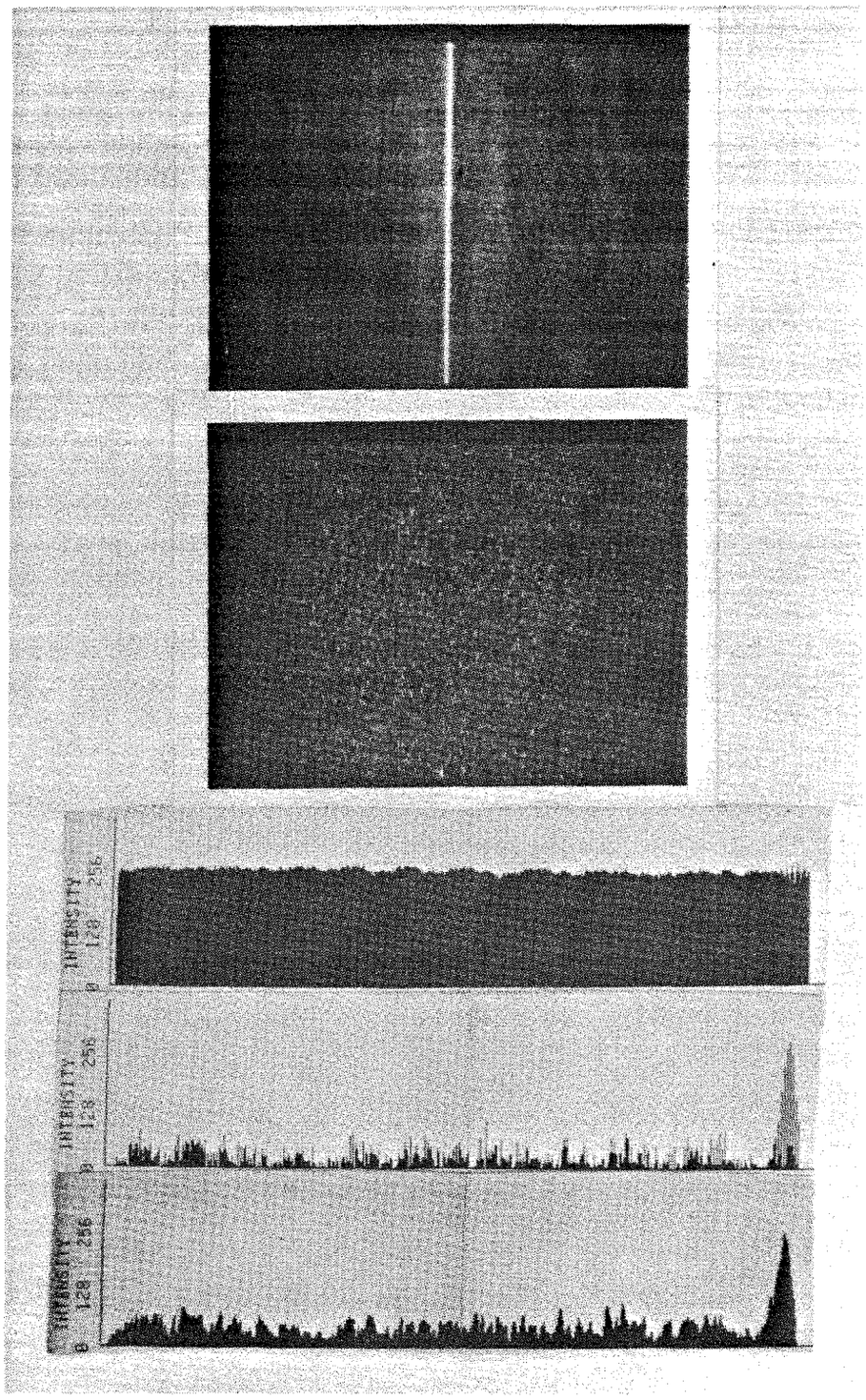


Figure 5.2.44 Single tone folded spectrum a) video image, b) video display after carrier demodulation, c) cross section along coarse frequency ridge, d) cross section after carrier demodulation and rectification, e) cross section after Gaussian convolution .

through the region of the time integrated peak. However, the phase information is lost after the absolute value rectification operation, so if phase measurements are to be performed they must be performed before the rectification operation. A significant advantage of spatial carrier encoding is that the amplitude of the signal is measured by the filtering, rectification, low pass, demodulation scheme, as opposed to the amplitude of a projection of the signal onto an arbitrary axis, such as the real axis, which is obtained in any bias subtraction scheme. As the phase of the spatial carrier changes with respect to the sampling pixel structure the amplitude passed by the discrete filter can vary by about a factor of $\sqrt{2}$, so radiometric amplitude measurements will require a compensation for this effect. The SNR after carrier demodulation was typically 8:1, and after the following Gaussian filtering the peak height decreased significantly, but the SNR improved to 16:1 for a dynamic range of the fully demodulated amplitude spectrum of 24dB. The resolution was typically better than 60 Hz in these experiments, which is within a factor of two of the theoretical maximum for an unapodized transformation.

The performance of the folded spectrum processor in the presence of multiple tones was examined through the use of amplitude modulated signals, and some examples of fully demodulated multitone spectra are illustrated in Figure 5.2.45. These spectra had a typical SNR of 6:1, so the probability of a noise sample exceeding the peak height was on the order of e^{-6}, but with close to 200,000 pixels this meant that hundreds of noise samples were larger than the desired peak. For this reason, at least 4 frames were averaged, thereby doubling the SNR and decreasing the probability that any noise samples would exceed the peak to a negligible level, and allowing good photographic presentation of the resulting multitone spectra. When the amplitude modulation was less than the coarse resolution, as in part b), then the bias ridges added incoherently, and the modulation depth and SNR of each peak along the ridge fell. Typically, the modulation depth was adjusted so the side-

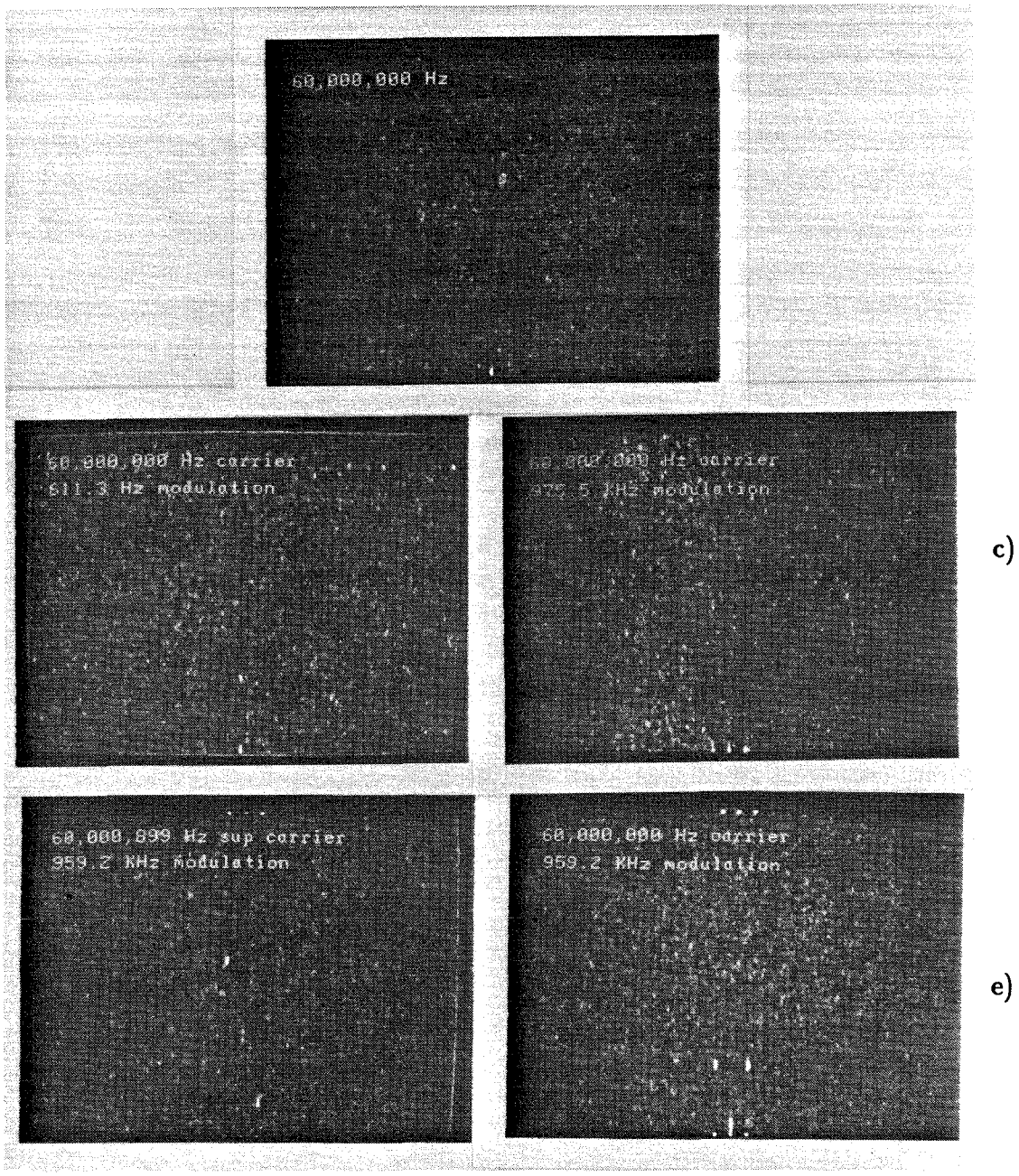


Figure 5.2.45. Demodulated multitone spectra that have been frame averaged. a) 60 MHz, b) 60 MHz 611.3 Hz AM, c) 60 MHz 975.5 kHz AM, d)60,000,900 Hz suppressed 959.2 kHz AM, e) 60 MHz 959.2 kHz AM.

bands had equal power to the carrier, so that when low frequency modulations were used the SNR fell as the one plus the square root of the number of equal power tones within a coarse frequency bin, which for amplitude modulations would be $\eta/(1+\sqrt{3})$. As the modulation frequency was increased the coarse frequency bias ridge split into two separated ridges, and the interferometrically produced fringe patterns within each ridge produced larger modulation depths. When the modulation frequency is an exact multiple of the laser diode PRF, as in parts c), then the sideband peaks appear horizontally displaced from the main peak, since they have the same fine frequency. For an arbitrary modulation frequency the sideband peaks are displaced symmetrically in both the horizontal and vertical dimensions from the carrier frequency folded spectrum peak, as shown in part d), since the upper and lower sidebands are separated in frequency by the same coarse and fine frequency components from the carrier. In a double sided analysis system, positive and negative fine frequencies produce peaks at the same fine frequency position producing a reflected folded spectrum as shown in part e). When a very strong square wave modulated carrier was applied to the signal AOD, the first harmonic would saturate the CCD detector in that coarse frequency vertical column, but successful time integrating fine frequency analysis was still performed on the weaker higher order harmonics. This is one of the significant advantages of the TSI technique over purely time integrating approaches, in which a single strong frequency component can saturate the entire detector. The size of the demodulated amplitude spectrum peaks was a monotonic function of the input spectral amplitude, but the relationship was not as linear as would be desired for typical radiometric applications. There were several reasons for the lack of linearity, these included nonlinear compression in the CCD, nonlinear response of the reference AOD, nonlinearities in the reference mixer, and sampled filter response variation. The nonlinear compression of the CCD could be avoided by only utilizing the region of linear dynamic

range, which in a high quality CCD can approach 10^4:1, but in these experiments in order to minimize the effects of the CCD noise and to appropriately use the digitizer quantization window it was necessary to work near the compression point of the CCD. In order to fully utilize the available optical power of the pulsed laser it was necessary to drive the reference AOD as strongly as possible, which resulted in nonlinear compression of the reference waveform by the $\sin^2$ transfer function of the large modulation depth AO diffraction. The signal AOD was not operated with a high diffraction efficiency since the diffracted light was space integrated onto a single column of the CCD, rather than being imaged over the entire aperture like the reference AOD. The phase of the spatial carrier with respect to the CCD sampling pixel structure introduces an amplitude variation upon bandpass filtering, as mentioned above. Another amplitude variation was introduced due to the super fine frequency of the input tone, which results in a motion of the carrier fringes during time integration, thereby maximally suppressing frequencies with a super fine component of 15 Hz, although this suppression depended on the relative phase of the spatial carrier with the pixel structure.

The main results of these experiments are summarized in Table 5.2.2 along with the potential characteristics of a fully optimized state of the art TSI DFT based folded spectrum processor, and the system characteristics which will limit the processor performance. In the experiments reported in this section I have demonstrated the feasibility of this approach to time and space integrating folded spectrum processing, based on a pulsed source and a DFT fine frequency analysis. A practical system would require a much better CCD detector array with larger dynamic range and better spatial resolution. The CCD performance was the primary limitation on system operation, and further characterization of the limitations of this processing scheme would require a CCD detector optimized for scientific purposes, rather than the industrial video camera used to demonstrate this folded spectrum system.

Parameter	Prototype	Potential	Limiting Factors
B=Bandwidth	24 MHz	250 MHz	pulsed laser power
N_c=Coarse Bins	384 pixels	2048 pixels	CCD Size, AOD TB
Coarse Resolution	68 kHz/pixel 150 kHz FWHM	B/1000	$1/T_a$ CCD MTF
N_f=Fine Bins	50	1000 $N\Delta f$	CCD size, carrier freq., Reference BW
Fine Resolution	60 Hz	1Hz	CCD Integration Time
SBR=Signal to Bias Ratio	.4 single tone	.9 single tone	$\eta/\sqrt{m}$ laser coherence, path lengths, multitone limited
DR_0=CCD Dynamic Range	40:1	10^4:1	Readout rate, Output Circuitry
DR=Spectrum Dynamic Range	24:1 after filter rect.,and low pass	$>10^2$:1	$DR_0\eta/(1+\sqrt{m})$ multitone limited
N=Number of Spectral Bins	10^4	10^6	CCD size, CCD MTF, Reference BW
Spectral Sensitivity	-40 dB	-60 dB	$N/\sqrt{m}$ TSI Processing Gain

Table 5.2.2. TSI folded spectrum processor performance and the potential for this technique, with some of the important interrelated limiting factors.

5.2.11 TDI approach to DFT processing

A major performance limitation of the TSI DFT processor is the presence of a large modulation amplitude coherent artifact due to multiple reflections in the cube beamsplitters, which appears in the form of a honeycomb ripple produced on the staring CCD. The bias subtraction technique will remove this pattern since it is constant from frame to frame. Alternatively, the spatial carrier demodulation will eliminate most of the unwanted ripple because it is at a much higher spatial frequency than the interferometric spatial carrier, although there will be some residual leakage through the sidelobes of the filters frequency response. However, the spatial frequency content of other coherent artifacts, such as dust or scratches on the optical components, can be within the bandpass of the bias removal filter producing large amplitudes of unwanted feedthrough into the demodulated output. For this reason a DFT folded spectrum processing technique, that incoherently averages these coherent artifacts was experimentally tested. This approach relies on the ability of a TDI CCD to sequentially integrate photogenerated charge at each of the photodetector sites within a TDI column, and average the coherent artifacts at each position to produce a coherently processed output with the suppressed spatial noise characteristic of an incoherently illuminated system. One such approach would be to utilize a TDI chirp transform of each coarse frequency bin of the pulsed source interferometer, as proposed for the DLO interferometer, and this approach is examined in section 5.3. In this section I will consider an alternative system which computes a DFT in a TDI translating coordinate frame, and serves to illustrate the connection between the DFT approach to spectrum analysis and the chirp transform approach.

The DFT operation as analyzed in section 5.2.2 requires the time integration of the phase reset sinusoidal columns of the raster scanned DFT matrix on a staring CCD detector array. The key to this processing operation is that the DFT rows

which are summed over as the laser is pulsed are imaged onto stationary pixels of the detector. In this manner the phase reset reference position is always integrated within the same pixel, producing a coherent radiometer for a sampled fine frequency at DC. An alternative which is easily accomplished with the flexibility provided by a digital video frame buffer acting as the reference function generator, is to have the columns of the DFT reference translate linearly along a TDI column of the CCD from pulse to pulse of the laser. When the TDI CCD is synchronously scanned with this image of the spatially moving reference, the signal detected in the moving coordinate frame will again appear as a phase reset DFT raster. At the phase reset position a coherent radiometer for a sampled fine frequency at DC will be synthesized by the time integration within the moving pixel. At the other TDI time delays, corresponding to images of different rows of the DFT mask in the TDI sliding window coordinate frame, an array of equally spaced coherent radiometers will be synthesized. An idealized TDI DFT mask, which is a tilted version of the idealized DFT reference of Figure 5.2.30, is shown in Figure 5.2.46. The phase reset position runs along the central diagonal of the reference mask, and any parallel diagonal in this mask will contain a sampled single tone, just like vertical slices out of the original idealized DFT reference mask. However, there is an important difference between the two approaches to computing the DFT. In the stationary integration approach the length of each row which is time integrated across is the same, giving equal resolution and processing gain at all fine frequencies. In the sliding window TDI DFT the integration is across the diagonals of the square mask, so the main diagonal is longest, while other diagonals linearly decrease in length away from the main diagonal. The resolution and gain linearly decreases away from the main diagonal, just as in the chirp transform algorithm. A quadratic phase factor is also produced by this sliding window processing operation, because the centroid of the integration period is a linear function of the output frequency variable. This effect

also produces an additional quadratic phase factor in the chirp transform algorithm.

A TDI processing time on the video format detector array used in the experiments lasted for one field readout time, with coherent integration over 245 successive integration stages at a 15,735.4 kHz line rate for a total integration time of 15.6 msec. Thus the maximum resolution that is obtainable with this mode of fine frequency processing is 65 Hz for a pixel that begins the integration at the top of the TDI column and coherently accumulates charge at each position until it reaches the bottom of the array. The super fine frequency in this case is measured modulo the field repetition frequency which is 60 Hz. The next field of integration produces a completely independent processing operation, and the time integrated peak produced during successive fields will have a relative phase given by the superfine frequency component which causes a phase drift between those fields. The video frame buffer grabs two successive fields of processing, and interleaves them in its memory. If the same tilted DFT reference is used during the two independent fields of processing, and the super fine frequency drift between fields is eliminated then a peak will be produced by both fields of processing along the coarse frequency ridge at the position corresponding to the fine frequency. When the super fine frequency is 30 Hz, the phase drift between fields is π, so one field would produce a constructive peak, and the following field would produce a destructive peak, so upon interleaving these peaks in the frame buffer memory an apparent Nyquist frequency carrier is produced. The envelope of this fake carrier frequency is given by the difference of the constructive and destructive sinc patterns, rather than a sinc function envelope as a true carrier encoding would produce. The sinc function is itself on a true carrier which is at a spatial frequency that is a linear function of the output frequency variable, and this carrier has a relative phase during the two successive processing fields that is given by the super fine frequency phase drift. The coherently processed information can be demodulated by bandpass filtering the interleaved data

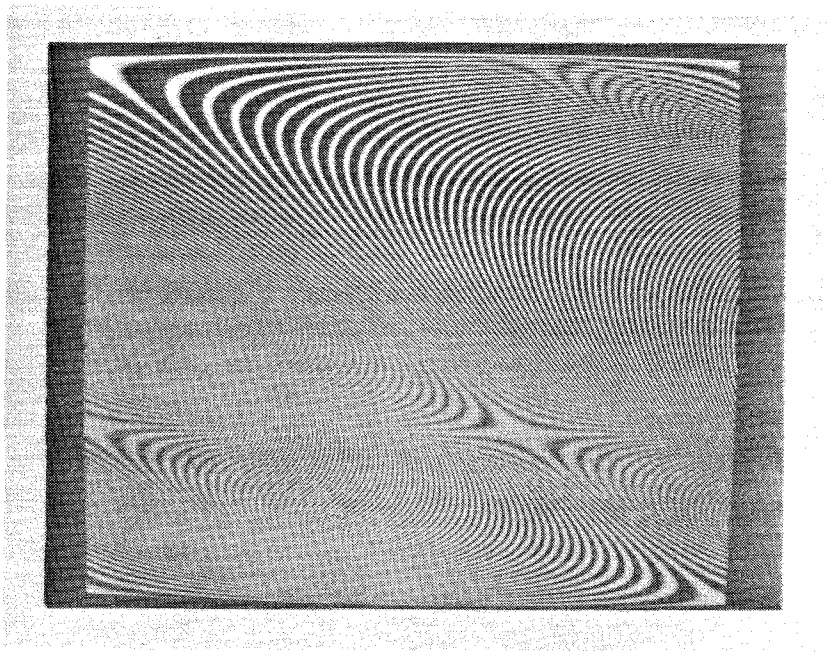


Figure 5.2.46. An idealized tilted DFT mask for performing fine frequency spectral analysis on a TDI CCD.

in the frame buffer with a filter along the fine frequency TDI dimension with zero mean tap coefficients given by [1 -1 1 -1 1 -1]. Rather than thinking of this operation as a carrier demodulation, it should be considered a successive TDI field bias subtraction with spatial redundancy averaging, which is facilitated by the interleaving of the successively processed fields in the digital frame buffer. This is because this operation is really a projection of the input super fine frequency phase drift onto the real axis, like a bias subtraction technique, and it is not a magnitude detection technique like a true carrier demodulation. If the phase drift is zero no carrier will be produced upon interleaving the two TDI fields, and the bandpass filtering will not produce a peak. The two successive TDI fields of processing can have independent DFT reference functions, and by making these two DFT masks π out of phase with each other, then a projection of the super fine frequency induced phase drift onto the orthogonal imaginary axis will result. In this case a super fine frequency of 0 Hz will result in a fake carrier when the two fields are interleaved, and no carrier will be produced when the super fine frequency is 30 Hz. This mode of bias subtraction was convenient to implement with the experimental apparatus that was available, and it helped to characterize the TDI DFT processing, but it is not a viable approach to system implementation.

To demonstrate this type of TDI DFT processing it was necessary to compensate for the acoustic velocity and reference magnification in the tilting of the DFT reference mask. The appropriately tilted TDI DFT reference mask is shown in Figure 5.2.47, along with an expanded version of the 84 pixel wide portion which was imaged onto the CCD during processing. It is interesting to note that in the digital frame buffer, this mask is poorly sampled in the tilted coordinate frame of integration on the TDI CCD, as demonstrated by the ripples in the main diagonal (DC) slice shown in Figure 5.2.47e. However, the reference AOD diffracts a plane wave component for each shifted DFT column, which accomplishes an almost ideal inter-

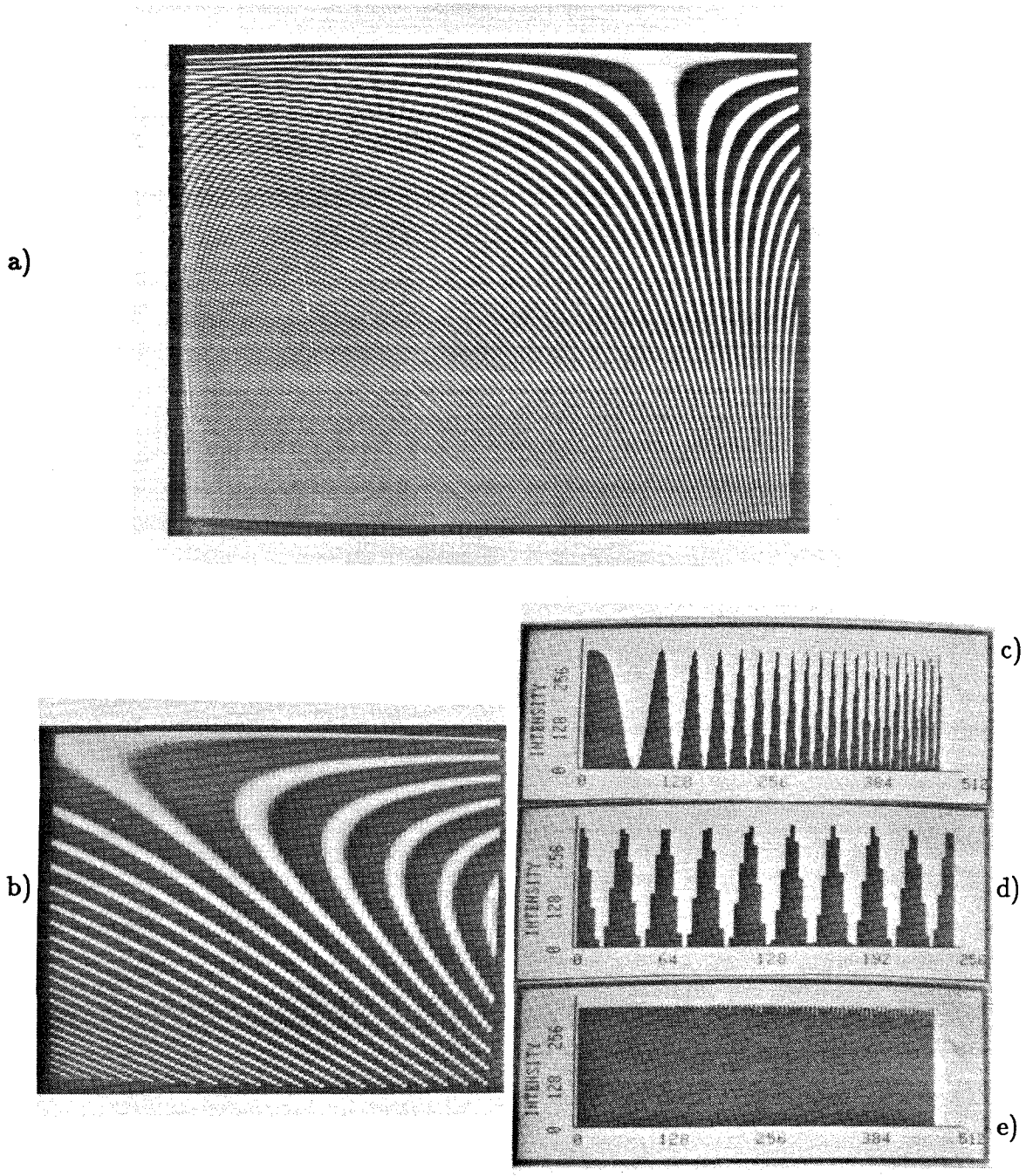


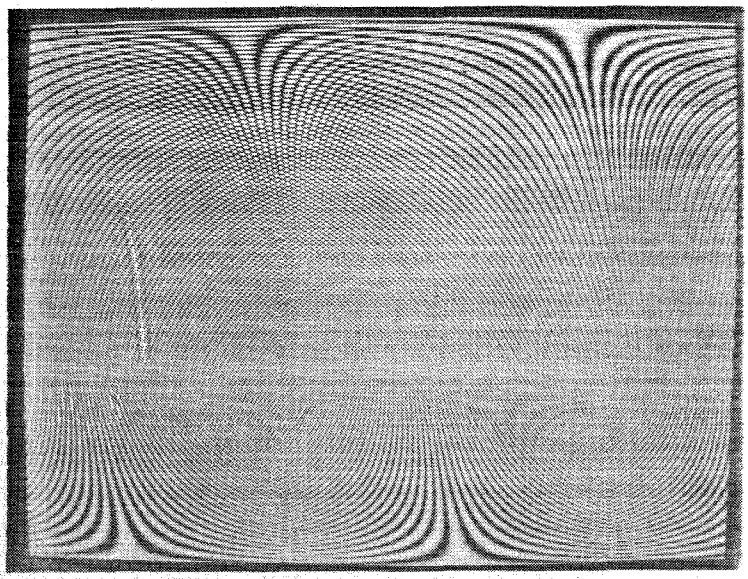
Figure 5.2.47. a) Properly scaled and tilted DFT reference mask, b) the magnified portion which was image onto the TDI CCD c) vertical, d) horizontal, e) and main diagonal cross sections.

polation operation upon interferometric detection in the sliding coordinate frame of the TDI CCD. The actual processing operation is determined by the interferometric detection of the tilted DFT reference mask on the TDI CCD. It was found that as the recombining beamsplitter was rotated in order to vary the center frequency of the chirping carrier, the center frequency of the processing operation also shifted. This is in distinct contrast to the staring approach to DFT processing where for a single tone input the time integrated fine frequency peak remained in a fixed position as the beam splitter was rotated, and only the carrier frequency under that peak varied. In the TDI approach to tilted DFT processing as the beamsplitter was rotated the position of a time integrated peak translated across the TDI window, and the carrier frequency simultaneously changed. This can be understood by considering the pixel at the top of the CCD at the start of a TDI processing field corresponding to a sliding window integration across the phase reset position at the tilted rectangular mask main diagonal. We would expect this pixel to correspond to the time integrating radiometer for a fine frequency at DC, but by rotating the beam splitter the interferometric detection can introduce extra spatial fringes along the TDI dimension that will be encountered as spatial phase variations as the pixel translates across the CCD aperture. If we rotate the beamsplitter sufficiently to produce an extra interferometrically produced spatial carrier frequency of 1 cycle per each 10 TDI stages, then the center frequency of processing will be shifted by 1/10 the TDI line rate, and the output due to this center frequency will be on a spatial carrier of 10 pixels per cycle. This is because the sampled fine frequency of interferometric oscillation which is necessary to shift the fringes detected along the TDI column in order to compensate the additional spatial variation encountered as the pixel traverses the column is equal to the number of additional fringes seen by the moving pixel per unit of time. The undesired spatial intensity variations of the signal and reference beams produced by the multiple reflections off the beam split-

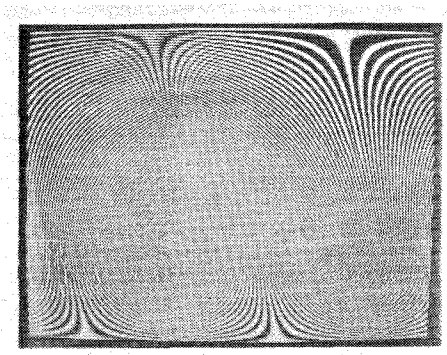
ter and other optical artifacts will be averaged out by the incoherent summation over a TDI column. However, the multiple reflections in the reference arm consist of tilted plane waves that can interfere with the signal arm at an offset angle from the directly transmitted beam and produce a shifted spectral scan artifact, similar to the shift introduced by rotating the beam splitter. This was not observed in our system because the dynamic range of the CCD was too small to detect small shifted sidelobe contributions, but in a high SNR system these multiple reflections should be minimized in order to avoid this effect.

The calculated tilt of the DFT reference mask was confirmed experimentally by applying a mask consisting of equivalently tilted sinusoidal tones to the reference AOD, and switching the CCD into the TDI mode of operation. The interferometrically produced magnified fringes were reproduced on the CCD output with much better fidelity than when this type of imaging was performed in the staring mode, because the coherent artifacts are greatly suppressed in the TDI dimension. When the tilted DFT reference function with out of phase interleaved fields shown in Figure 5.2.48 was applied to the reference AOD, and the CCD was operated in the TDI mode, out of phase folded spectrum processing was performed on successive fields. An example of the video image produced by this folded spectrum technique is shown in Figure 5.2.49a for an input frequency of 60,000,600 Hz. The smooth quality of the image in the vertical TDI direction is readily apparent from this photograph, but some spatial variations of the bias are still visible in the coarse frequency dimension as vertical streaks. The processed folded spectrum peak is not visible in this video image, because of the low contrast of the reproduction, but a cross section through every other pixel of the coarse frequency ridge reveals the presence of the peak, as shown in Figure 5.2.49c. The low contrast is partly due to an artifact of the digitization and direct analog measurements indicated a modulation depth of better than 30%. A magnified cross section which displays each pixel

a)



b)



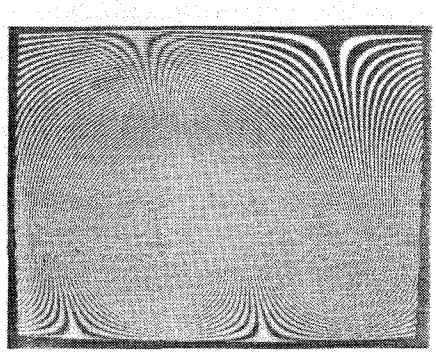


Figure 5.2.48. a) Reference TDI DFT mask with out of phase interleaved fields, at a tilt angle of 84 horizontal frame buffer pixels over the frame buffer vertical height of 240 pixels per field. b) Deinterleaved into the out of phase successive field space time raster format.

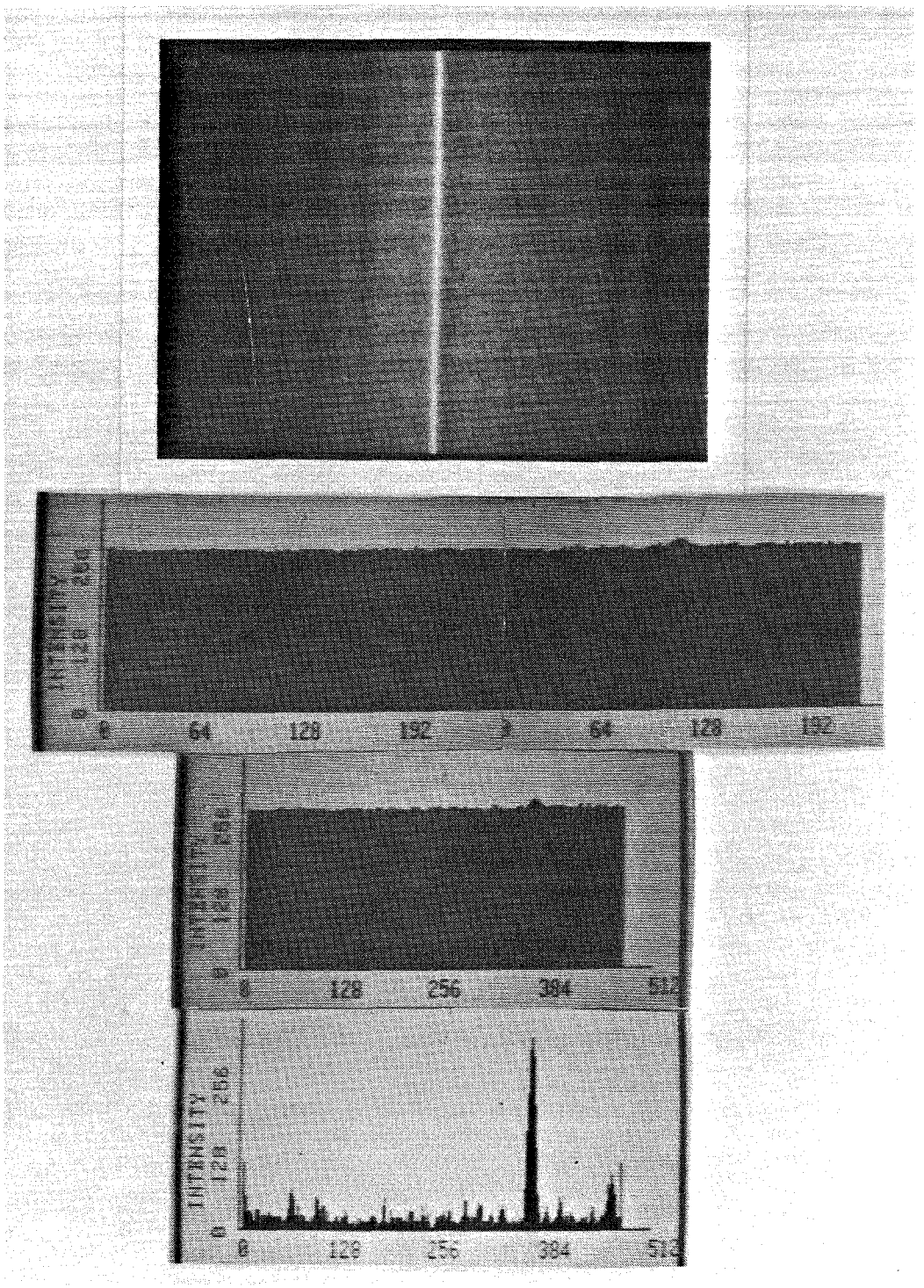


Figure 5.2.49. a) Video image produced by the TDI DFT processor for an input frequency of 60,000,600 Hz. b) Cross section through ridge showing reinterleaved out of phase peaks, c) cross section through constructive field only, d) cross section after demodulation.

is shown in Figure 5.2.49b, and the presence of an impulse on a carrier of every other pixel is visible. This image can easily be demodulated in the video frame buffer by convolving with the bipolar impulse response [1 -1 1 -1 1 -1] in the TDI dimension, and dropping the sign bit of the result, which effectively rectifies the bandpass filtered image. The rectification operation doubles the frequency of the fake Nyquist frequency carrier caused by the interleaving of the out of phase fields, and produces a demodulated peak at baseband because twice the Nyquist frequency folds back to zero, completing the necessary demodulation operation. The resulting folded spectrum had an SNR on the order of 12:1, and this could be improved by using a 3 by 3 Gaussian convolution, which resulted in an SNR of 24:1. The fully demodulated peak is shown in cross section in Figure 5.2.49d, and a large peak is clearly visible.

Some examples of fully demodulated single tone TDI DFT folded spectrum images are shown in Figure 5.2.50. The sequence of images shows a single tone moving up in frequency by 600 Hz increments from 59,999,400 Hz to 60,001,800 Hz. For the tone with a negative fine frequency of -600 Hz the peak appears near the top of the image with high resolution. At exactly 60 MHz the peak lies near the bottom edge of the CCD, and any tone with fine frequencies from -150 to -350 would occur during the vertical blanking interval and would not produce visible peaks. As the fine frequency is increased the peak moves up the CCD, but a considerable amount of widening, and a decrease of amplitude is apparent for a fine frequency of +1800 Hz. The final image shows a peak corresponding to a +600 Hz offset from the 100th laser diode harmonic, and it can be seen to be at the same fine frequency position as 60 MHz +600Hz, but it has moved over in the coarse dimension as expected, although by less than 100 pixels as would be appropriate in a properly scaled folded spectrum system.

Multitone spectra were created by amplitude modulating a carrier with a sinu-

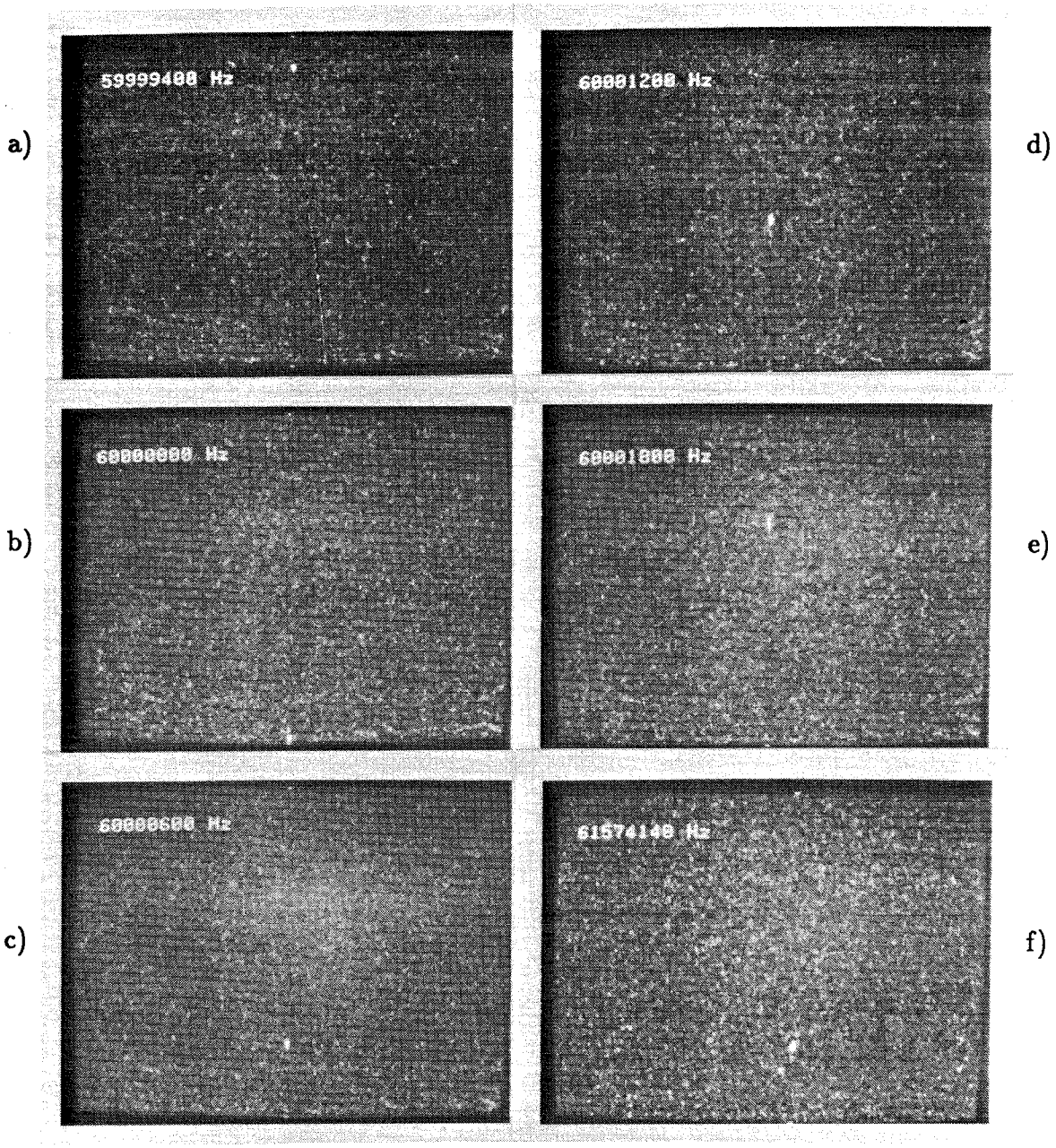


Figure 5.2.50. Some fully demodulated single tone folded spectra. a) 59,999,400 Hz b) 60,000,000Hz c) 60,000,600Hz d) 60,001,200Hz e) 60,001,800Hz f) 61,574,140Hz.

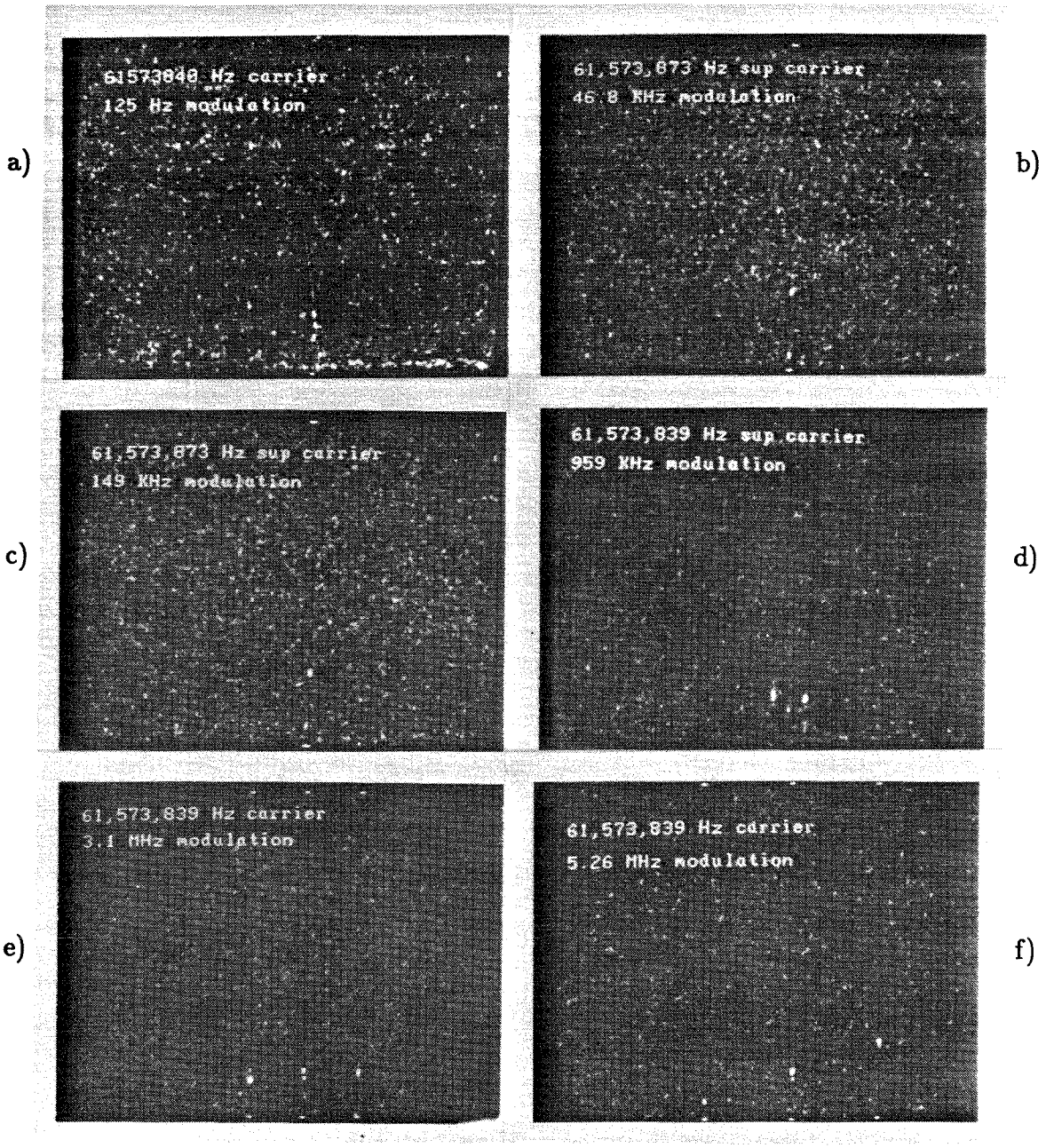


Figure 5.2.51. Some demodulated amplitude modulated multitone folded spectrum images. Carrier is at the 100th LD harmonic +300 Hz, modulation frequencies a) 125 Hz b) 46.8 kHz c) 149 kHz d) 959 kHz e) 3.1 MHz f) 5.26 MHz.

soidal modulation of various frequencies. This creates two sidebands of equal amplitude and displacement from the carrier, which appear symmetrically displaced from the position of the carrier peak in both coarse and fine dimensions. Some examples of demodulated multitone folded spectrum images made with this modulated carrier technique are shown in Figure 5.2.51, for a carrier frequency of 60MHz plus 100 times the laser pulse frequency plus 300 Hz fine frequency offset. When the modulation frequency is much less than the coarse resolution, then the modulation depth of each peak within that coarse frequency ridge is decreased by $1/\sqrt{m}$, where m is the number of equal power tones within that coarse frequency bin. The resulting spectrum will have a lower SNR, and the maximum dynamic range that is achievable is decreased from the CCD dynamic range by $\eta/(1+\sqrt{m})$. This is illustrated in part a) where a modulation frequency of 125 Hz is used to split the carrier peak into three peaks stacked vertically, and the resulting folded spectrum is of low SNR (approximately 6:1). When the modulation frequency is increased to 46.8 kHz and the carrier is suppressed, the 3 pixel wide coarse ridge produced on the CCD does not widen, but the resulting fringe patterns occur only on the edges of the ridge, so that the demodulated peaks are in different coarse frequency columns. This seems to indicate that the actual optical resolution incident on the CCD is on the order of a single pixel wide, and it is CCD blurring that results in the 3 pixel wide coarse frequency blur spot which is detected, and for this reason the modulation depth is not decreased in this case. When the modulation frequency is increased to 149 kHz the coarse ridge splits into two ridges, each producing an independent peak symmetrically displaced from the position of the suppressed carrier. When a large modulation frequency is adjusted to be an even multiple of the laser diode pulsing frequency then the modulation sidebands are displaced horizontally, as shown in parts d) and e). For a modulation frequency of approximately 5.26 MHz, three peaks are clearly lined up in a tilted straight line, corresponding to the

carrier and two sidebands.

The linearity of the TDI DFT fine frequency analysis operation is plotted in Figure 5.2.52. The spectral analysis is linear to almost one pixel, although repeatability is to only 4 or 5 pixels. The full width at half maximum (FWHM) is indicated by the vertical bars, and it is seen to grow with frequency offset, as expected. Typical full widths at half maximum for these peaks are about 6 pixels out of a single field in the fine frequency dimension, but this almost doubles for the +1800 Hz offset case. The resolution corresponding to this 6 pixel width is approximately 69Hz, which is almost equal to the theoretical limit on the unapodized resolution that is obtainable with this system.

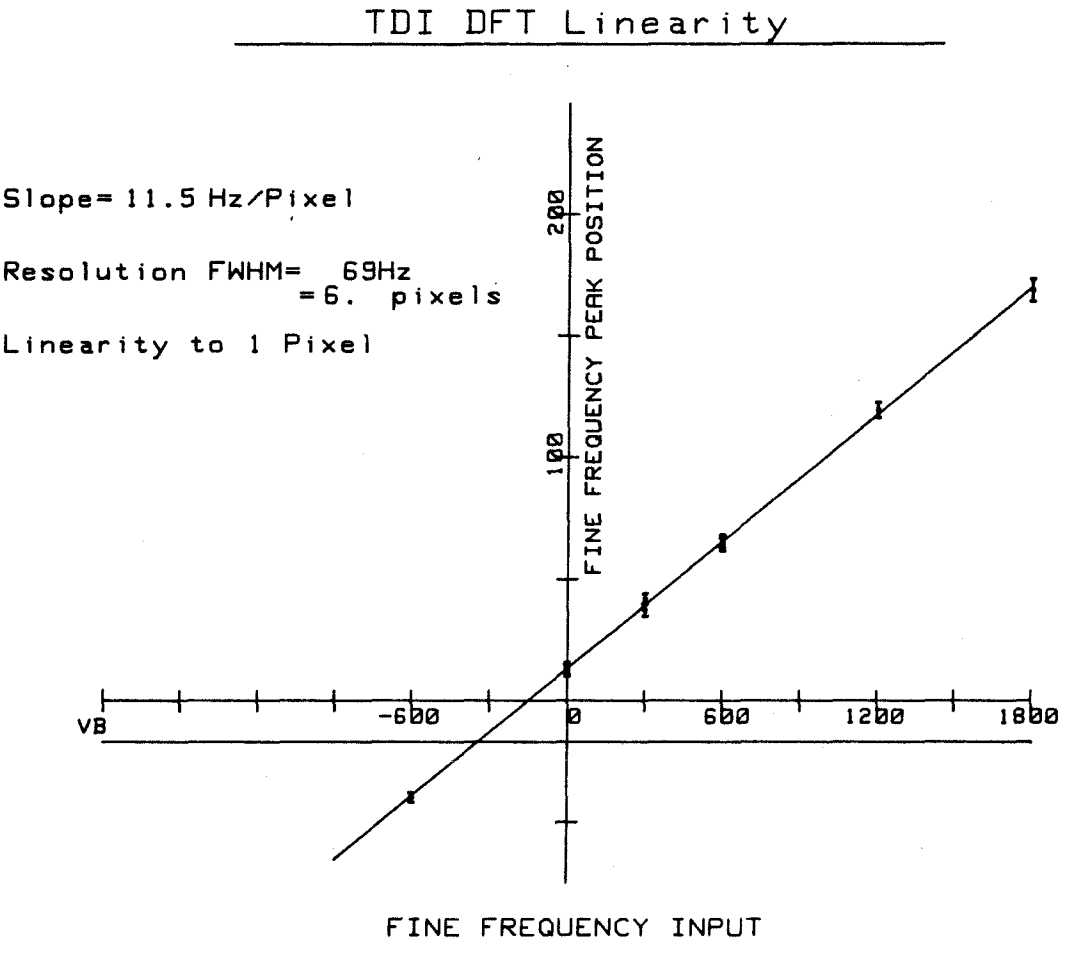


Figure 5.2.52. Linearity of the TDI DFT spectral analysis.

5.3 Pulsed Source Chirp Transform Interferometer

The chirp transform algorithm can be used to perform the fine frequency spectral analysis of a pulsed source interferometer, in a manner similar to that proposed for the distributed local oscillator interferometer. This has the advantage over the DFT based processor analyzed in the previous section that the reference waveform is an easy to generate low time bandwidth product, phase modulated chirp, rather than the complex, extremely high TB, stepped frequency raster scan of the DFT matrix. However, because of the chirping carrier the chirp transform algorithm suffers from the difficulty of the spatial carrier demodulation that must be performed. This problem can be overcome in 2-dimensional processing by placing the spatial carrier in the coarse frequency dimension, which also allows the utilization of the widest possible bandwidth two sided chirp transform. This is important because readily available CCD detectors can be made to TDI at a video line rate of 15.7354 kHz, giving a Nyquist limited analysis bandwidth per sideband of 7.867 kHz. With special effort these scan rates can be doubled yielding a two sided analysis bandwidth of 31.5 kHz, which implies that the space integrating AO spectrum analyzer must have a resolution of about 30 kHz, and because of the spatial apodization effects this will require an acoustic transit time of about 65 μsec, which is near the limit imposed by acoustic attenuation in TeO_2. An unapodized 65μsec Bragg cell could be combined with a video rate double sided chirp transform array in order to build a prototype of the system ideas presented in this section. Another approach to TDI chirp transform demodulation is the interlaced video field carrier demodulation which is easily performed in a digital frame buffer. This approach relies on super fine frequency drift from video field to the next video field of operation, and successive TDI processing fields are subtracted to eliminate the bias.

The Mach Zehnder architecture for performing interferometric TDI chirp transforms on each space integrated coarse frequency bin is very similar to the additive

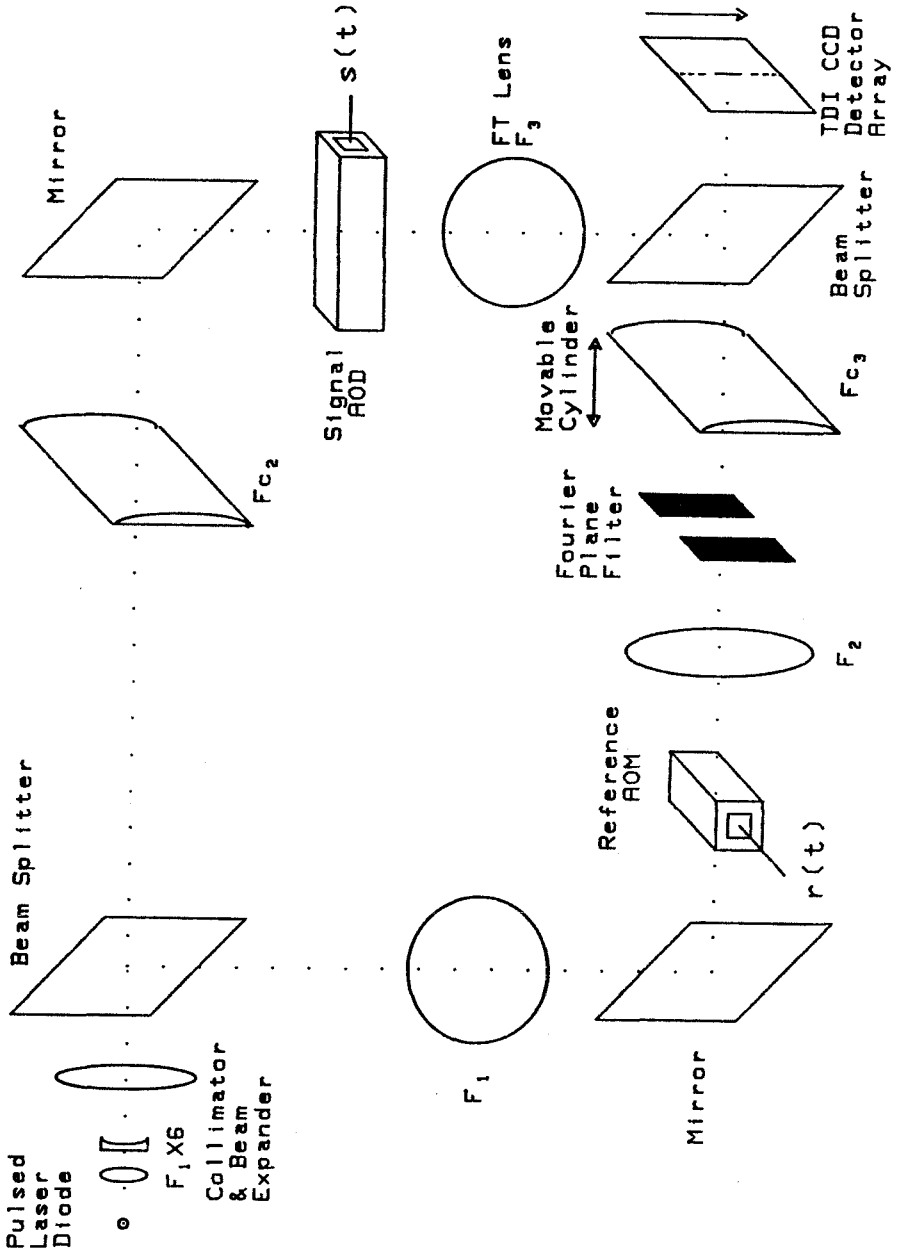


Figure 5.3.1. Schematic diagram of the pulsed source TSI folded spectrum processor that uses an interferometric implementation of the TDI chirp transform algorithm to perform fine frequency analysis on each coarse frequency bin.

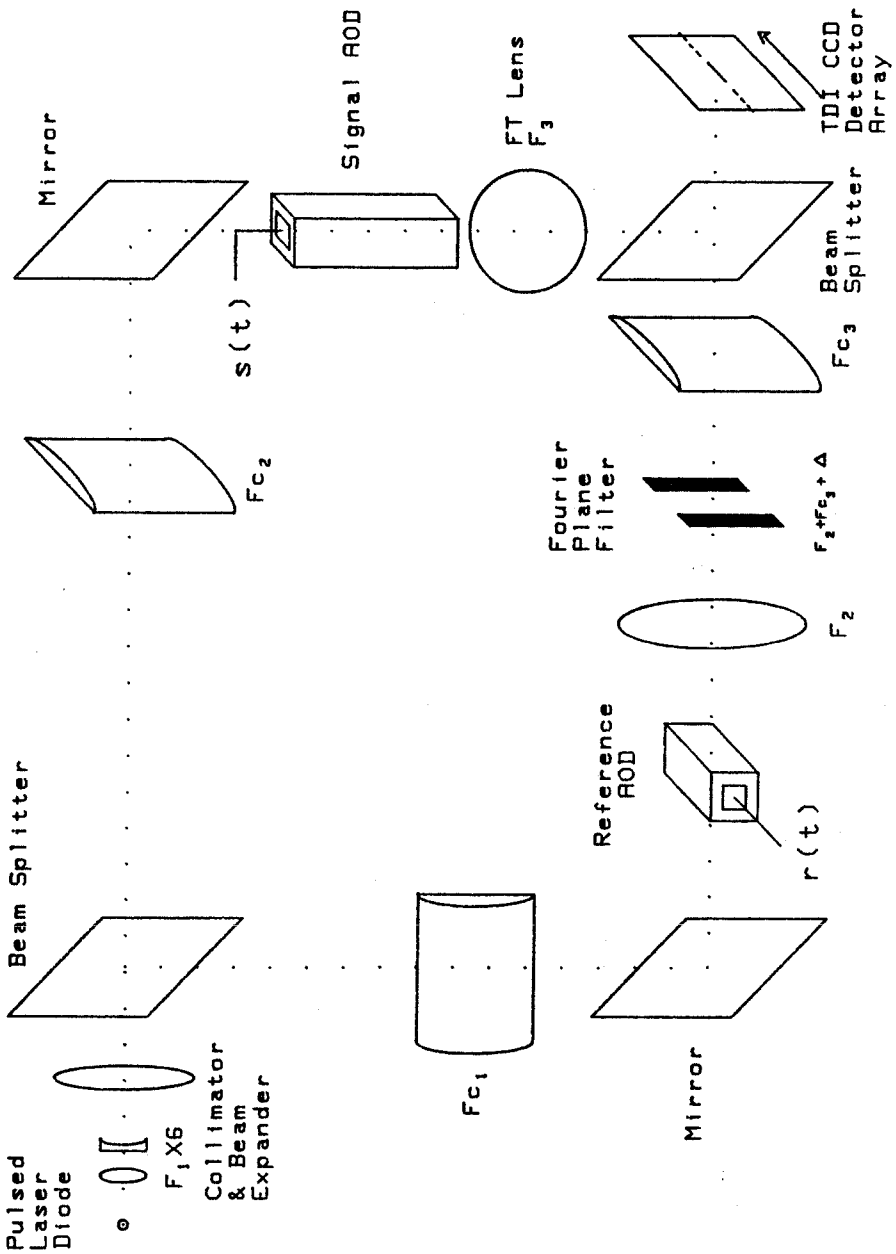


Figure 5.3.2. The modification of the TDI DFT system that was used to perform experiments on the TDI chirp transform approach to folded spectrum processing.

architecture used in the last section for performing the DFT, except for the realignment of a cylindrical lens, in order to produce a chirping interferometric fringe pattern along each coarse frequency ridge, and the replacement of the orthogonal reference AOD with a collinear acousto-optic modulator (AOM). The AOM is operated as a phase modulator with a focussing input optic field, which means that the focal spot in the modulator must be smaller than an acoustic wavelength, or a Fourier plane filter must be used to block the DC if it is spatially resolvable. The phase modulated wavefront emerging from the AOM must be recollimated in the x dimension, and a quadratic spatial curvature must be introduced in y in order to produce the desired spatial chirp interference pattern along each coarse frequency ridge produced by the signal arm. The quadratic spatial curvature in the y dimension can either be negative (focussing), or positive (expanding), so that by collimating the diffracted field from the AOM in the x dimension with a cylinder an expanding cylindrical wave is produced that could be used as the reference beam. However, in this case the quadratic spatial profile in y is then only a function of the distance to the AOM, and is not easily programmable, and the light utilization efficiency is not good if the expanding wave illuminates a region much larger than the CCD. For these reasons it is desirable to collimate the output of the AOM, and place a movable focussing cylinder with curvature in y beyond the Fourier plane to reimage the point source of the AOM as a slit at a plane beyond the CCD. A schematic diagram of the architectural implementation of this type of interferometer system is shown in Figure 5.3.1. Actually the reference AOD of the DFT processor shown in Figure 5.2.3 from the previous section can be used as a wide aperture modulator without any realignment, and an existing cylinder can be moved so that it deviates from the telescopic condition and leaves a residual quadratic curvature orthogonal to the dimension of space integration. This type of system is illustrated in Figure 5.3.2, and it is seen to be very similar to the system used to perform the

TSI DFT approach to folded spectrum processing, and in fact the same system was used to perform experimental verification of both system concepts.

For interferometric chirp transform analysis the reference waveform is a phase modulated chirp, with a very low bandwidth which is matched to the CCD TDI rate. Since the AOM is illuminated by a pulsed laser source, the phase modulated chirp used as a reference can be produced with a discretely phase stepped waveform that jumps to a new phase delay obeying a quadratic phase profile each CCD shift time. Since phase values are measured modulo 2π, the quadratic phase history can be phase unwrapped and reset to 0 each time it reaches 2π. Over small bandwidths phase shifts can be interpreted as time delays, so the actual reference waveform is a discretely stepped, phase unwrapped, quadratically time delayed carrier at the AOM center frequency.

$$r(t) = \sum_{k=0}^{K-1} \text{rect}\left[\frac{t-kT}{T}\right] e^{-i2\pi f_0(t-\frac{b}{2}(k-k_0)^2 \text{mod} t_0)} \qquad (5.3.1)$$
$$= \sum_{k=0}^{K-1} \text{rect}\left[\frac{t-kT}{T}\right] e^{-i2\pi[f_0 t-\frac{b}{2}(k-k_0)^2]}$$

The important parameter in this equation is the sampled quadratic chirp rate b, which must be made equal to the spatial chirp rate of the interference fringe profile detected on the CCD, scaled by the TDI shifting velocity. The phase unwrapping is represented as the modulus (or residue) of the quadratic phase delay divided by the carrier frequency period $t_0 = 1/f_0$, yielding a time delay between 0 and t_0, that is multiplied by the carrier frequency to yield a phase delay between 0 and 2π.

Spatially the reference wavefront is quadratically curved along the dimension orthogonal to the acoustic propagation in the signal AOD in order to produce an interferometric spatial chirp profile along each coarse frequency ridge. The wavefront diffracted from the reference Bragg cell on the kth laser pulse is collimated along the x dimension and after passing through the non telescopic sphere cylinder pair, retains a residual spatial curvature along the y dimension, so the incident reference

field at the CCD plane on the kth laser pulse can be represented as follows.

$$r(x,y,k) = c_r e^{-i2\pi\nu t} p(m_x x) q'(y) e^{-i2\pi[f_0 kT - \frac{b}{2}(k-k_0)^2]} e^{i2\pi[\alpha_x x + \frac{\beta}{2}(y-y_0)^2]} \qquad (5.3.2)$$

The spatial quadratic curvature at the detector plane β is determined by how far the separation between the sphere and cylinder, $z = F_2 + F_{c3} + \Delta$, deviates from the telescopic condition, $z_0 = F_2 + F_{c3}$. This can easily be found by solving for how far behind the detector array the image of the Fourier plane comes to a focus, $z_b = \Delta + F_{c3}^2 \Delta$, which yields a spatial curvature $\beta = 1/\lambda z_b = \Delta/\lambda(F_{c3}^2 + \Delta^2)$ at the detector plane. The offset factor of the quadratic spatial curvature, y_0, can be adjusted with the beamsplitter in plane rotation, and is used to match the offset of the time domain quadratically phase modulated reference, k_0. For double sided baseband processing, both the space and time chirps are made symmetric within their respective windows, the space chirp is centered on the TDI CCD aperture, and the time chirp is centered within the temporal processing aperture. The spatial frequency of the reference wave in the x dimension $\alpha_x = \sin\theta_x/\lambda$ is adjusted by the beam splitter tilt angle θ_x, and is designed to place a high frequency spatial carrier in the x dimension in order to allow the separation of the signal from the bias terms. The aperture function of the reference wave in the x dimension is the magnified collimated laser profile, while in the y dimension it is given by a diffracted version of the the orthogonal laser profile, but if the Fresnel diffraction regions near the edge of the lens apertures are avoided it can be assumed to be a smooth Gaussian weighing function.

As usual we will interferometrically combine the output of a space integrating AO spectrum analyzer with the reference wavefront to obtain a detected intensity on the CCD with each pulse that is the modulus squared of the weighted fields from the two arms, and this will be time integrated over the duration of each pulse. The input to the spectrum analyzer is a signal $s(t)$ with a single sided complex spectrum $\tilde{S}(f)$ that can be broken up into its magnitude $|\tilde{S}(f)|$ and phase $\Omega(f)$.

$$I(x,y,k) = \int_{(k-\frac{1}{2})T}^{(k+\frac{1}{2})T} \text{crect}\left[\frac{t}{\tau}\right] * \sum_k \delta(t-kT) \tag{5.3.3}$$

$$\left| c_r e^{-i2\pi\nu t} p(m_x x) q(y) e^{-i2\pi[f_0 t - \frac{b}{2}(k-k_0)^2]} e^{i2\pi[\alpha_x x + \frac{\beta}{2}(y-y_0)^2]} \right.$$

$$\left. + c_s e^{-i2\pi\nu t} p(M_y y) \int \tilde{S}(f) e^{i2\pi f t} W\left(\frac{x}{\lambda F} + \frac{f}{v_a}\right) df \right|^2$$

$$= c\tau c_r^2 p^2(m_x x) q'^2(y) + c\tau c_s^2 p^2(M_y y) \left| \int \tilde{S}(f) W\left(\frac{x}{\lambda F} + \frac{f}{v_a}\right) df \right|^2$$

$$+ 2\eta c c_s c_r p(m_x x) q'(y) p(M_y y) \int \tau \text{sinc}[(f+f_0)\tau] |\tilde{S}(f)| W\left(\frac{x}{\lambda F} + \frac{f}{v_a}\right)$$

$$\cos\left[2\pi\left((f+f_0)kT - b(k-k_0)^2/2 - \alpha_x x + \beta(y-y_0)^2/2\right) + \Omega(f)\right] \; df$$

As before the interferometric term is weighted by the amplitude of the spectral harmonic of the pulsed laser diode waveform at the difference frequency of the signal and reference. The fringe pattern detected on the CCD has a quadratic spatial profile in y, and has a quadratic temporal variation plus the difference frequency of signal and reference sampled from pulse to pulse. The difference frequency sampling will yield an aliased fine frequency variation that will be localized by the fine frequency processing TDI chirp algorithm. The coarse frequency information can be determined by the space integrated ridge position as before. The TDI chirp algorithm is completed by detecting on a shift and add CCD, that spatially shifts in the y dimension by one pixel of width Δy with each laser pulse, yielding an effective velocity $v_{CCD} = \Delta y/T$. The spatial and temporal chirp matching requirements will be met by "focussing" the time integrating chirp transform. This yields the chirp rate matching condition $bT^2 = \beta \Delta y^2$, or $b/\beta = v_{CCD}^2$, the chirp offset matching condition $bk_0 T = -\beta y_0 \Delta y$, and the arbitrary phase alignment $bk_0^2 = \beta y_0^2$. Ignoring the spatial weighing functions and the apodization effects of the CCD MTF, we can arrive at the simplified output of the TDI chirp transform system.

$$
\begin{aligned}
D(n,m) &= \sum_{k=m-M}^{m} \int\int h_x(x-n\Delta x)h_y(y-(m-k)\Delta y)I(x,y,kT)dxdy \\
&= \sum_k \mathrm{rect}\left(\frac{k-(M-m)/2}{M-|m|}\right)\Bigg[c\tau c_r^2 p^2(m_x n\delta x)q'^2((m-k)\Delta y) \\
&+ c\tau c_s^2 p^2(M_y(m-k)\Delta y)\left|\int \tilde{S}(f)W\left(\frac{n\Delta x}{\lambda F}+\frac{f}{v_a}\right)df\right|^2 \\
&+ 2\eta c c_s c_r p(m_x n\Delta x)q'((m-k)\Delta y)p(M_y(m-k)\Delta y) \\
&\int \tau\mathrm{sinc}[(f+f_0)\tau]|\tilde{S}(f)|W\left(\frac{n\Delta x}{\lambda F}+\frac{f}{v_a}\right) \\
&\cos\Big[2\pi\Big((f+f_0)kT - b(kT-k_0)^2/2 + \beta((m-k)\Delta y - y_0)^2/2) \\
&- \alpha_x n\Delta x\Big) + \Omega(f)\Big] \quad df\Bigg] \\
\approx &(M-|m|)c\tau c_r^2 + (M-|m|)c\tau c_s^2\left|\int \tilde{S}(f)W\left(\frac{n\Delta x}{\lambda F}+\frac{f}{v_a}\right)df\right|^2 \\
&+ 2\eta c c_s c_r \int \tau\mathrm{sinc}[(f+f_0)\tau]|\tilde{S}(f)|W\left(\frac{n\Delta x}{\lambda F}+\frac{f}{v_a}\right) \\
&(M-|m|)\mathrm{sinc}\Big[(M-|m|)T[mbT+(f+f_0)\mathrm{mod}\tfrac{1}{T}]\Big] \\
&\cos\Big[2\pi\Big(mbT(k_0+TM)-(M-m)[(f+f_0)\mathrm{mod}\tfrac{1}{T}] \\
&- \alpha_x n\Delta x - bT^2m^2/2\Big) + \Omega(f)\Big] \quad df
\end{aligned}
\tag{5.3.4}
$$

The impulse response consists of the triangularly weighted in m reference beam bias, the triangularly weighted signal dependent coarse frequency analyzed ridge, and the desired interferometric term. The interferometrically detected term consists of the spatially channelized amplitude spectrum in n, with the triangularly weighted spectral scan of each coarse frequency column with the usual chirp transform triangular resolution and complicated phase factors, placed on a carrier in the n dimension. This theoretical pulsed source TDI chirp transform impulse response pattern is shown in isometric projection in Figure 5.3.3, and the triangular profiles are readily apparent, and the peak riding on the coarse analyzed ridge can be seen to be on a high frequency carrier running across the ridge in the coarse dimension. The triangular bias variation can be removed by performing successive folded spec-

trum transforms without a delay between them, thereby overlapping the triangular ridges and bias terms to produce a uniform background in the m dimension, but at the expense of overlapping the successively computed spectral scans.

The principal differences between the pulsed source chirp transform folded spectrum system and the pulsed source DFT system are the triangular gain and resolution effects, and the inconvenient chirping phase factors in m. The principal difference between this pulsed source chirp transform interferometer and the DLO chirp transform interferometer is the spectral weighing by the laser pulse train harmonic spectrum rather than the DLO reference spectrum, and the absence of the coarse axis sidelobe suppression factor due to the time integration of the temporal oscillations caused by interference with the adjacent DLO beads. This is because the DLO interferometer spatially channelizes the local oscillators which are used for time integrating detection, while in the pulsed source processor the harmonics of the laser diode are available as interferometric baseband converters at all spatial positions. So in the DLO interferometer the sidelobes of a coarse analyzed frequency component beat with adjacent DLO beads to produce moving fringes that time integrate to zero, while in the pulsed source processor the sidelobes choose the laser diode spectral harmonic closest to the signal frequency in order to produce an unsuppressed sidelobe contribution. The high sidelobe level means that coarse frequency apodization of the signal Bragg cell is more important in the pulsed source family of processors than in the DLO interferometers. In Figure 5.3.4a the theoretical impulse response for a coarse apodized pulsed source TDI chirp transform folded spectrum processor is illustrated, and in the coarse dimension the additional sidelobe suppression is obvious, but it is not as great as in the DLO approach. The coarse frequency ridge has been widened by the apodization, which allows more cycles of the spatial carrier to be contained within the peak amplitude profile, which helps to better define the spatial frequency of the carrier, allowing improved bias

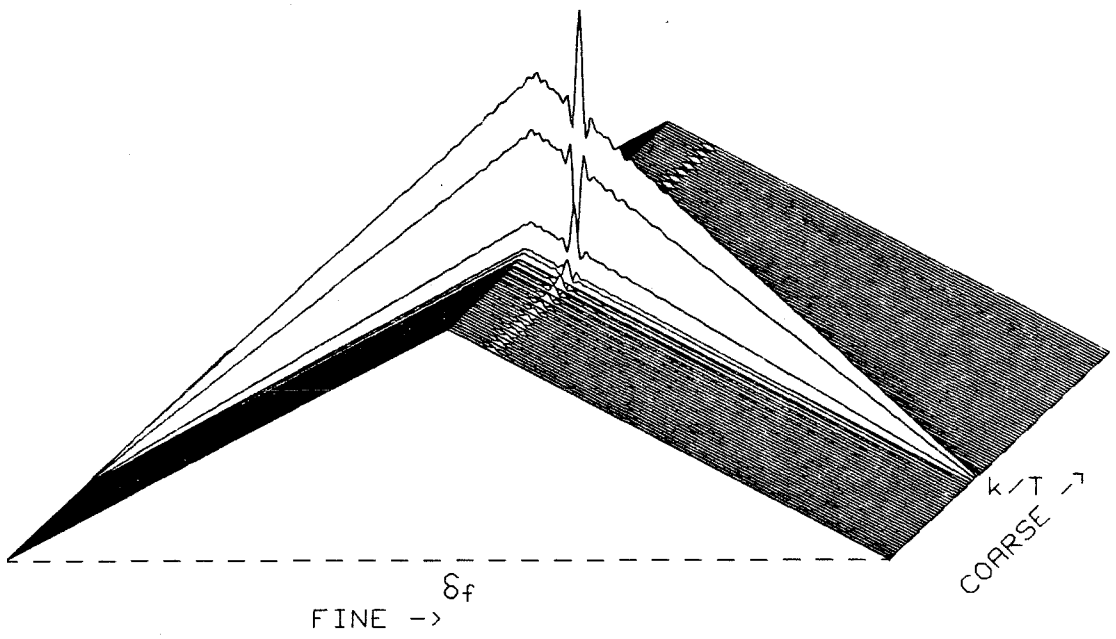


Figure 5.3.3. Theoretical impulse response of the TDI chirp transform folded spectrum processor.

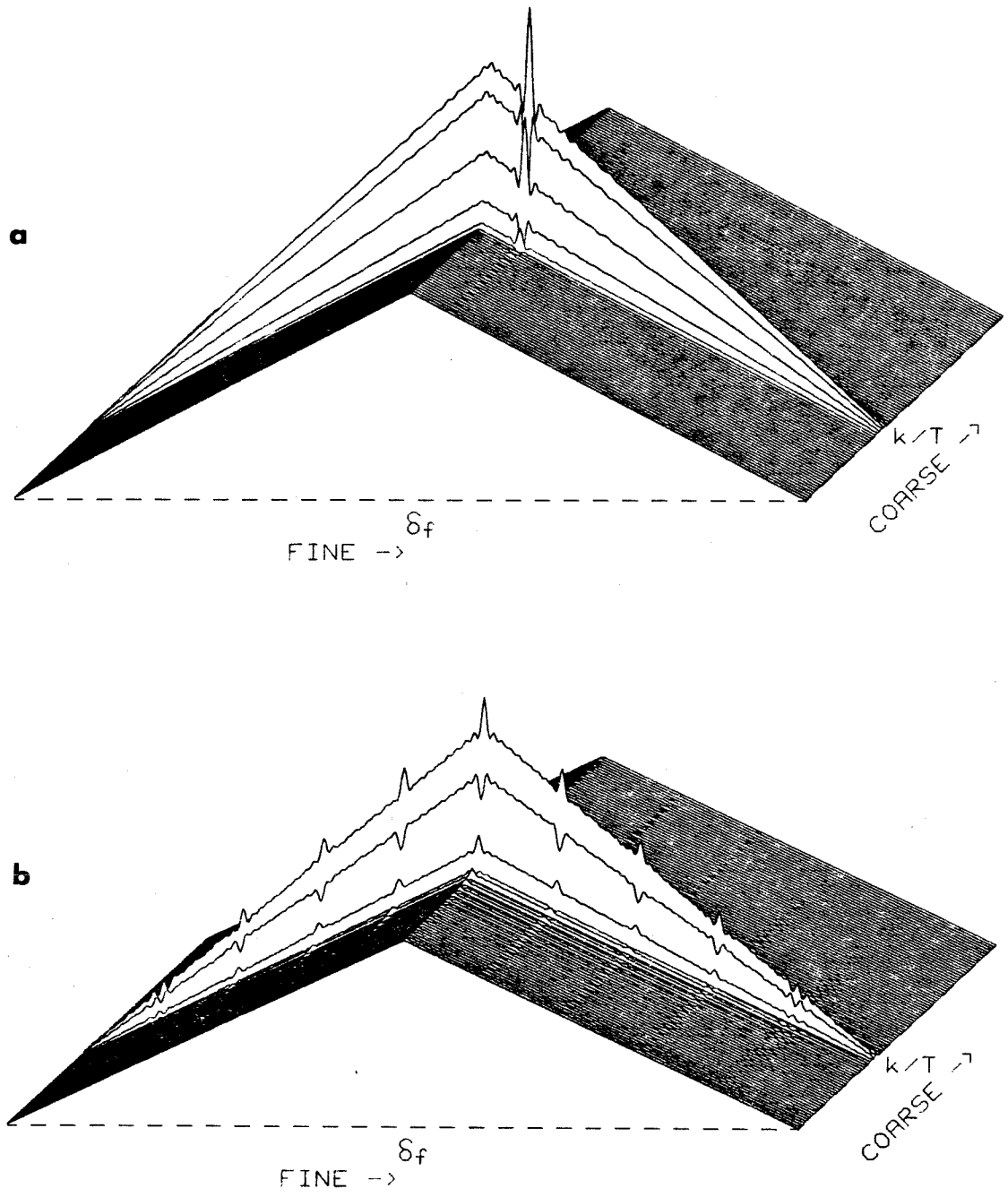


Figure 5.3.4. a)Impulse response for the coarse frequency apodized TDI chirp transform folded spectrum processor. b) multitone folded spectrum for nine equal power tones in one coarse frequency bin.

rejection in the carrier demodulation circuitry. In Figure 5.3.4b a theoretical multitone spectra is illustrated, for the case when the nine spectral components are all equally spaced in one coarse bin and they all have equal power, and the total power is the same as in part a. The triangular gain and resolution effects normally associated with the chirp algorithm are apparent, and the modulation depth decrease of each tone over the single tone case is apparent. The modulation depth decreases as $1/\sqrt{m}$, where m is the number of equal power tones, as in the previously discussed interferometers, and not as $1/m$, which would be a disastrous decrease in a white noise environment ($m \approx M$).

5.3.1 Experimental realization of the chirp transform interferometer

The experimental demonstration of the chirp transform interferometer required moving the cylinder, which was used for reference AOD reimaging, forward on its translation stage in order to interferometrically produce a chirping fringe profile along each coarse frequency ridge. The limited range of travel of the translational stage only produced a few fringes along the ridge, and by rotating the beam splitter the DC position of the chirping fringe could be positioned at the edge of the CCD. The resulting spatial impulse response is shown in Figure 5.3.5d, which was produced by applying 60 Mhz to both AODs and operating the CCD in the normal staring TI mode. A matched temporal impulse response was produced by applying the appropriately quadratically phase modulated 60MHz carrier to the reference AOD, which acts as a wide aperture modulator for these low frequency phase modulations. The phase unwrapped, and sampled quadratic phase modulation is clocked by the TDI CCD horizontal reset signal, and is shown in Figure 5.3.5b, along with an electronically produced interferometrically detected chirp profile. The phase modulated signal actually applied to the reference AOD had almost no amplitude modulation, and is shown in Figure 5.3.5c. The resulting optically produced tem-

a)

b)

c)

d)

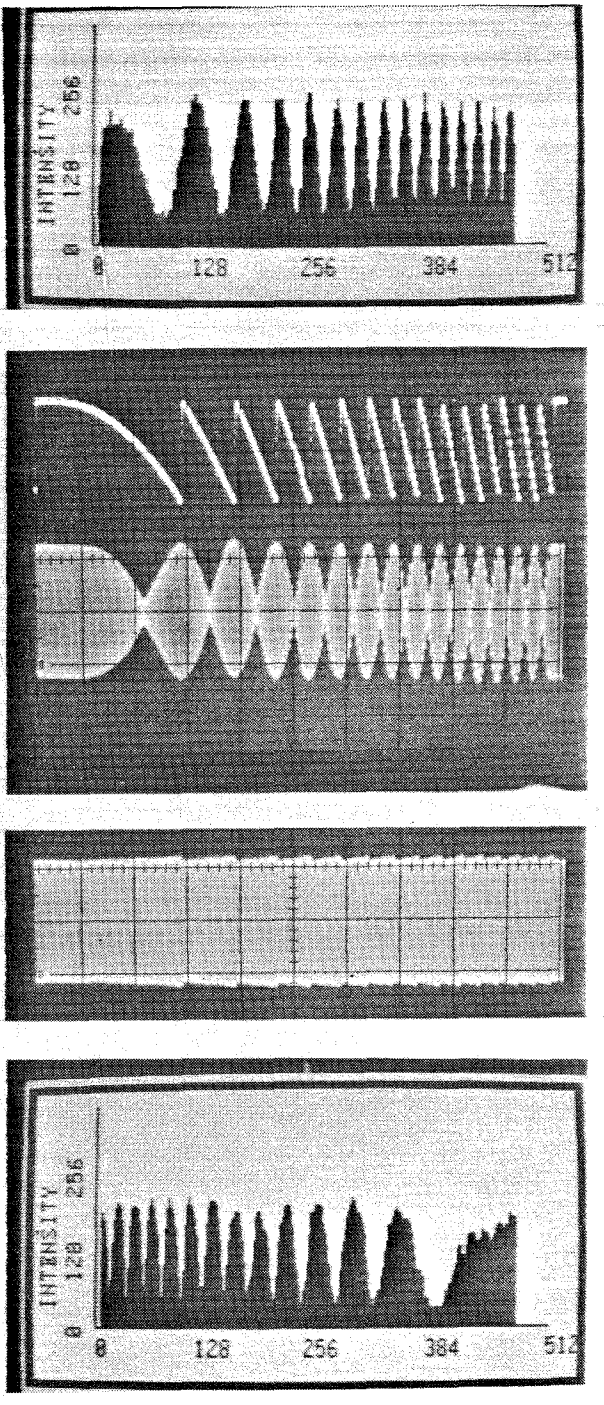
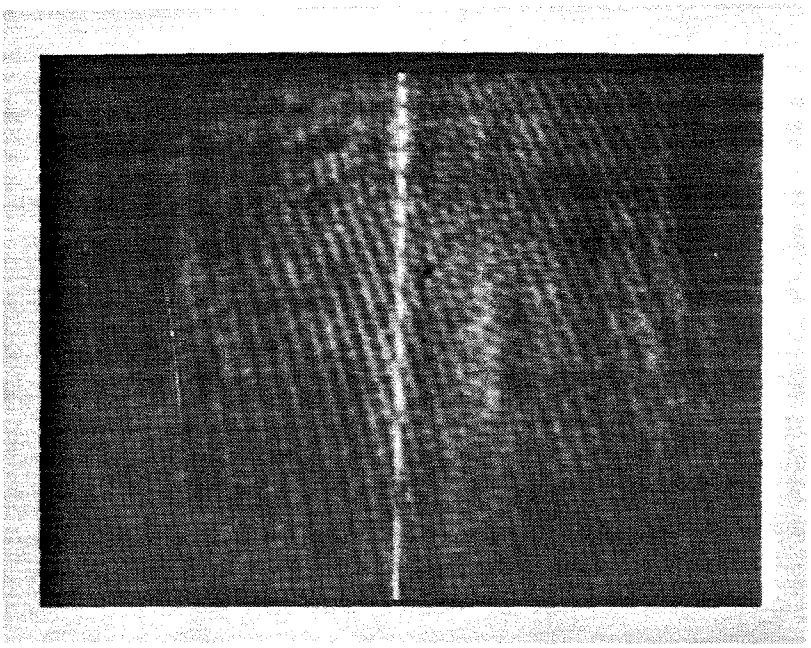


Figure 5.3.5. a) Interferometric temporal impulse response detected on TDI CCD with slit. b) Phase unwrapped quadratic phase delay of reference and electronically detected temporal impulse response. c) Actually phase modulated signal applied to reference AOD. d) Optical interferometric spatial impulse response.

a)



b)

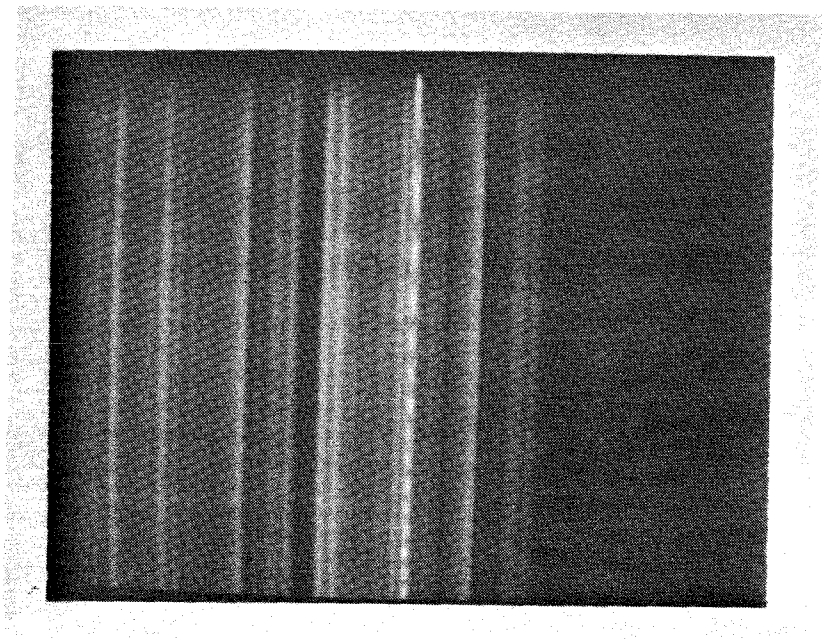


Figure 5.3.6. a) Image of the spatial impulse response, b) image of the interferometric temporal impulse response detected on a TDI CCD with slit illumination.

poral impulse response is shown in Figure 5.3.5a, which was measured by applying 60 MHz to the signal AOD and operating the CCD in the TDI mode with a single pixel wide slit placed in front of the detector array. The corresponding images of the interferometric spatial and temporal impulse responses are shown in Figure 5.3.6. The temporal impulse response (14 fringes) is seen to be reasonably well matched to the spatial impulse response (13 fringes), although reversed in direction as is appropriate for a chirp convolution.

Folded spectrum processing was accomplished by applying the phase modulated 60 MHz tone to the reference AOD while applying a single tone to the signal AOD and operating the CCD in the TDI mode. A set of overlapped constructive and destructive cross sections through a coarse frequency ridge for various fine frequencies are shown in Figure 5.3.7. The resolution is best for the tone that produces a peak at the edge of the CCD, corresponding to a pixel that coherently integrated over the full TDI aperture of the detector. The resolution decreases as the input fine frequency moves away from a harmonic of the laser diode PRF, and the resulting spectral peak widens. The carrier frequency is a linear function of the fine frequency, which results in a chirping carrier under the spectral peaks, which approaches zero for a fine frequency near +1200 Hz. The peaks at ±1200 Hz exhibit extended sidelobe response which probably indicates imperfect chirp matching or chirp nonlinearities. The TDI CCD is continuously shifting, even during the vertical blanking interval when no output is produced, so the spectral scan between a fine frequency of 0 Hz and +200 Hz is blanked out. The foldover of the positive frequencies to the successive field is best illustrated by plotting a cross section of a coarse frequency ridge for an entire frame as shown in Figure 5.3.8. The triangularly varying bias of two successive fields of computation overlap to produce an approximately uniform bias, and the spectral scans of two successive fields also overlap. However, there is little difficulty in distinguishing to which field of processing a spectral peak is

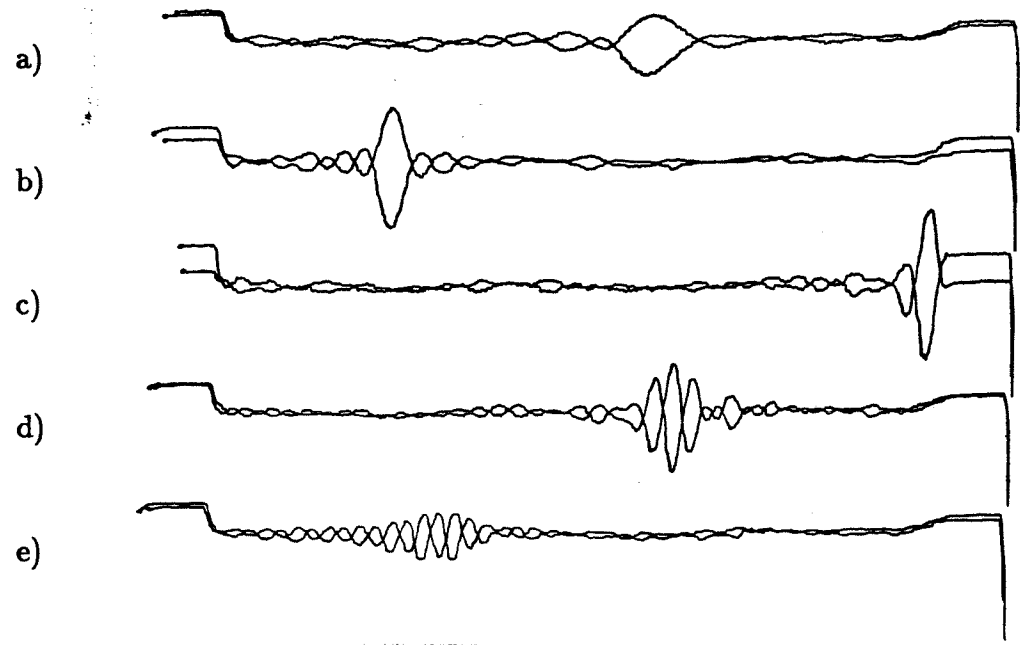


Figure 5.3.7. Cross sections through a single field TDI coarse frequency ridge for various applied fine frequencies. a) 60,001,200 Hz b) 60,000,600 Hz c) 60,000,000 Hz d) 59,999,400 Hz e) 59,998,800 Hz.

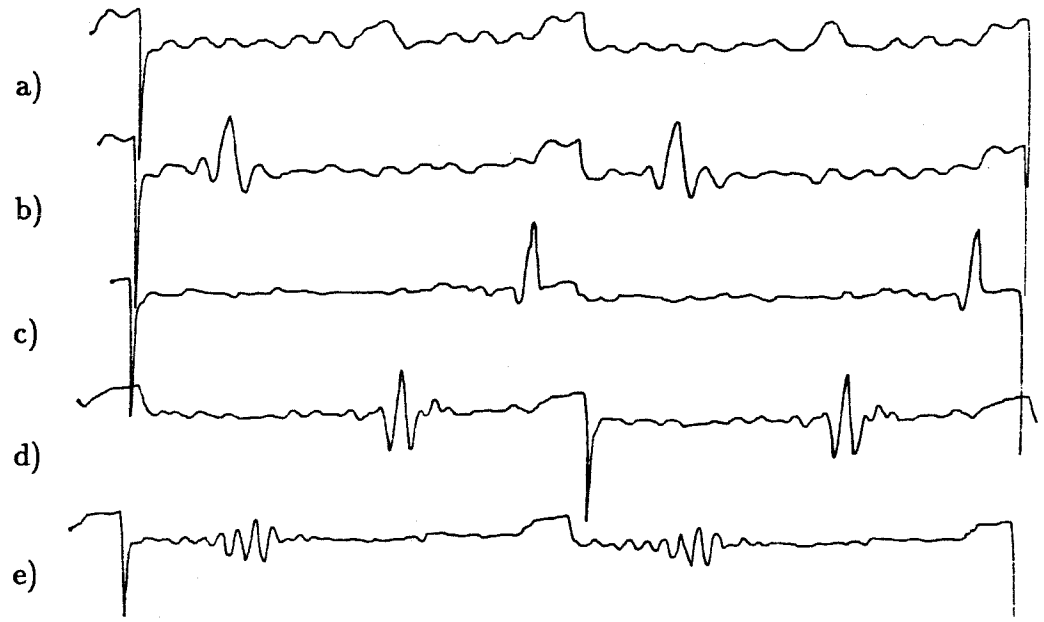


Figure 5.3.8. Cross sections through a full frame TDI coarse frequency ridge for various applied fine frequencies, showing foldover of positive frequencies to the following field. a) 60,001,200 Hz b) 60,000,600 Hz c) 60,000,000 Hz d) 59,999,400 Hz e) 59,998,800 Hz.

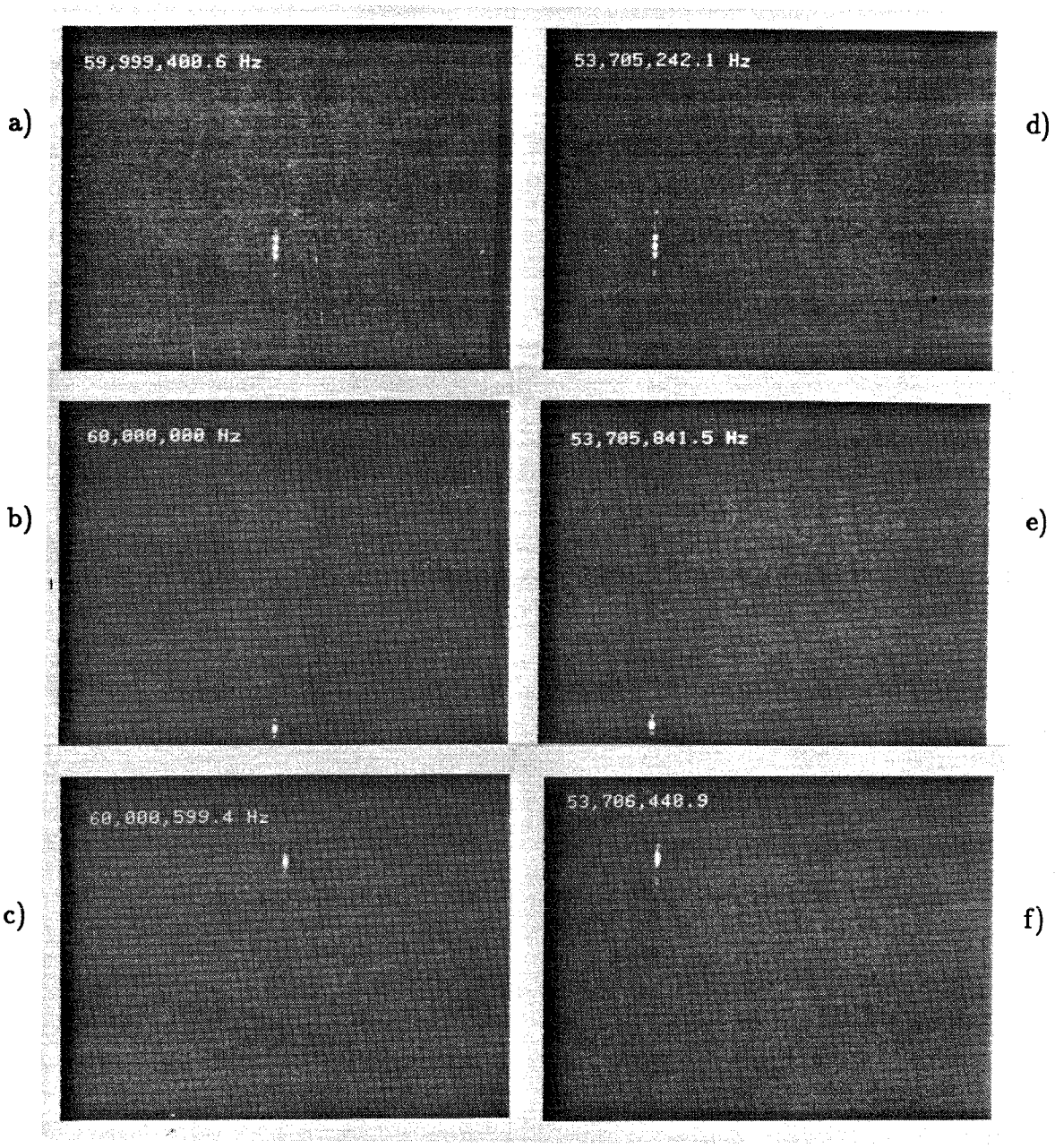


Figure 5.3.9. Single tone successive frame bias subtracted images produced by the TDI chirp transform folded spectrum processor. a) 59,999,400.6Hz, b) 60,000,000Hz c) 60,000,599.4 Hz, d) 53,705,242.1 Hz e) 53,705,841.5Hz f) 53,706,440.9 Hz.

due, because of the chirping carrier. Single tone successive frame bias subtraction demodulation was accomplished in the digital frame grabber by manually grabbing constructive and destructive frames, and subtracting the results. Several single tone images are presented in Figure 5.3.9, and these images provide the best example of multidimensional spectral analysis presented in this thesis. This is primarily because successive frame bias subtraction works much better than carrier demodulation, but a real system can not use this type of bias removal in a multitarget environment. The chirping carrier and triangular resolution associated with the chirp algorithm are clearly visible, and independent of the coarse frequency. The SNR of these images was typically better than 30:1, on a CCD which resulted from peaks with about than 40% modulation depth before digitization, on a CCD with an SNR of 40:1, so the expected SNR should be 32:1. The use of a pair of high SNR CCDs operating out of phase on the two output ports from the recombining beam splitter, with real time frame subtraction could produce high quality folded spectrum processing with this architecture.

In order to demonstrate the full frequency analysis bandwidth capabilities of the video rate TDI chirp transform fine frequency analysis technique it was necessary to insert an auxiliary cylindrical lens in the reference arm. The additional focussing power was needed to increase the space bandwidth product of the interferometric spatial impulse response to that imposed by the vertical Nyquist limit. The two sided quadratic phase modulation which was applied to the 60 MHz reference, can be visualized with electronic interferometry and the resulting chirp and the corresponding phase unwrapped quadratic phase delay function, are shown in Figure 5.3.10a. When the purely phase modulated carrier is applied to the reference AOD, with an unmodulated carrier in the signal AOD, then the optically produced temporal impulse response can be measured with a TDI CCD with a slit illumination, and the result is shown in Figure 5.3.10b. The reference cylinder was translated

a)

b)

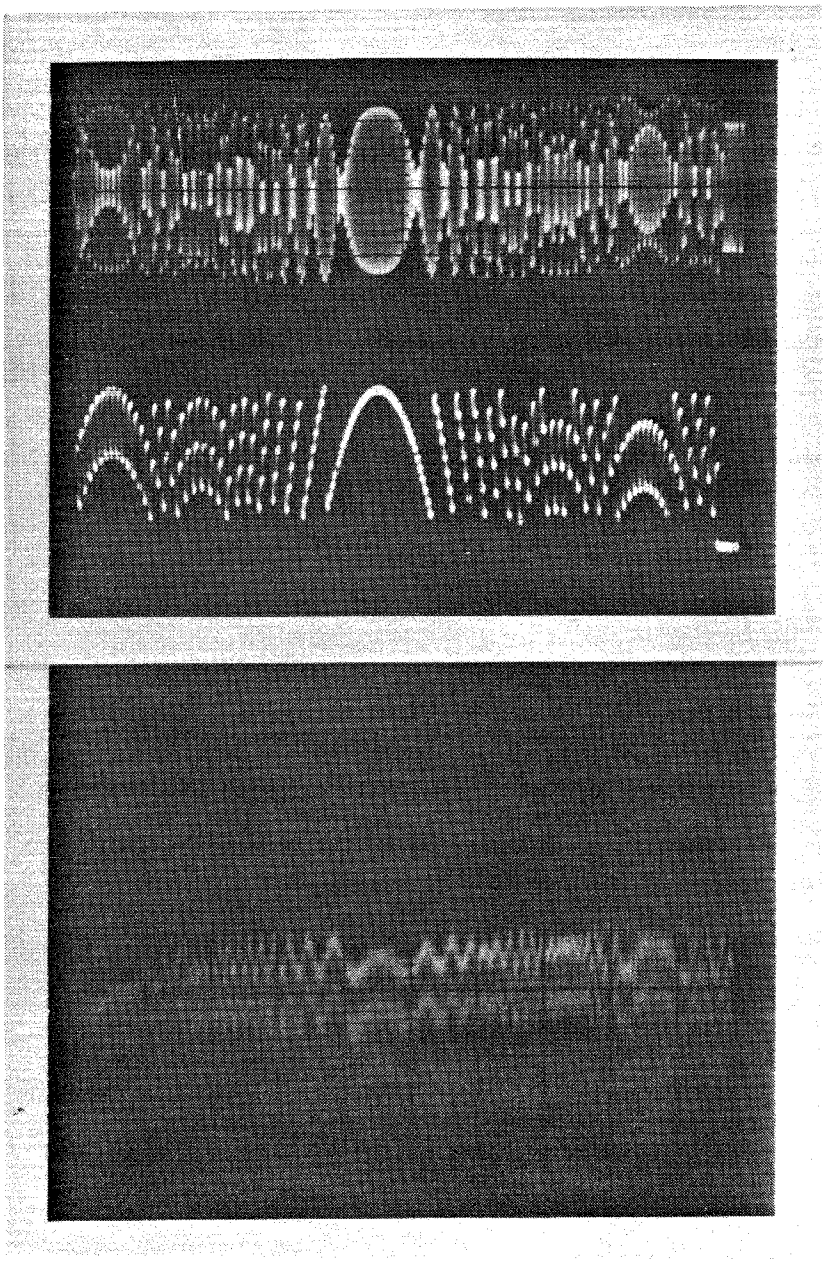


Figure 5.3.10. a) Two sided, phase unwrapped quadratic phase delay of reference and electronically detected temporal impulse response. b) Optical interferometric temporal impulse response detected on TDI CCD with slit.

a)

b)

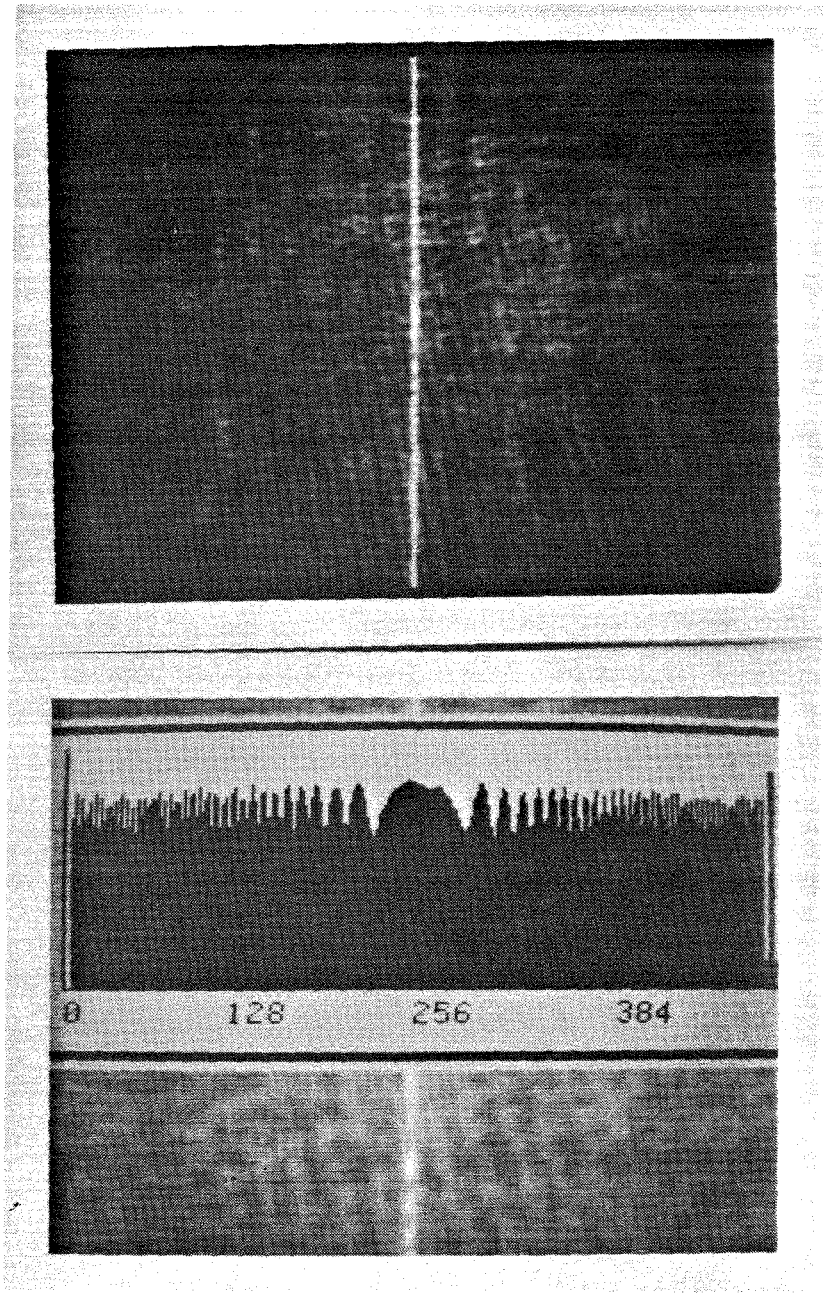


Figure 5.3.11. a) Image of optically produced spatial impulse response. b) cross section through optical interferometric spatial impulse response.

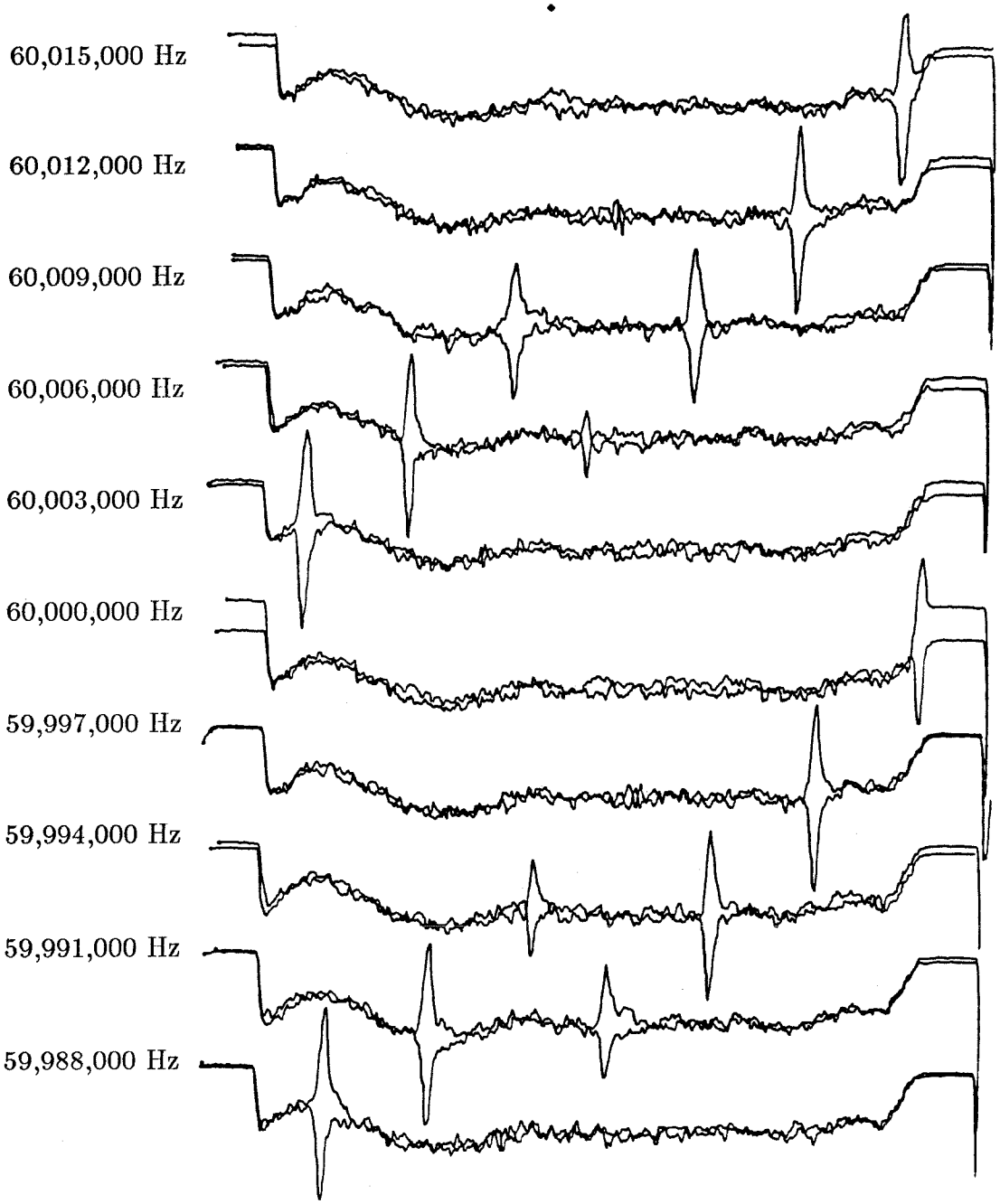


Figure 5.3.12. Cross sections through a single field TDI coarse frequency ridge for various applied fine frequencies for Nyquist limited chirp transform fine frequency analysis.

along the optical path in order to produce an interferometric spatial impulse response that was well matched to the temporal impulse response. An image of the interferometric profile detected on a staring CCD, and a cross section of the interferometric modulation along the bias ridge are shown in Figure 5.3.11. Both the spatial and temporal impulse responses exceed the Nyquist frequency and are thus under sampled at the high frequency ends of the chirps, but this had a negligible effect on the processing operation because of the small extent of the under sampled region, and the MTF suppression of these high frequencies. The actual position of the reference cylinder was optimized by "focussing" the time integrated chirp transform in order to obtain the highest and narrowest peak possible. Overlapping constructive and destructive cross sections through the coarse frequency ridge for various fine frequencies are displayed in Figure 5.3.12, which illustrates that successful two sided chirp processing over the full video line rate has been demonstrated. The bottom three traces show a spectral peak with a corresponding positive fine frequency due to the next lower laser diode harmonic. The spectral peak is handed off to the next laser harmonic in the middle 7 traces, and finally the peak is handed off to the next higher laser diode harmonic in the top three traces. Ideally each time the fine frequency changes by a laser diode harmonic the coarse frequency ridge would move over to the next pixel in order to produce an unambiguous folded spectrum, however this experiment did not have sufficient coarse frequency resolution to accomplish this goal. Frequencies close to half way in between two laser diode harmonics produced two peaks, one corresponding to a negative fine frequency due to the higher harmonic, and one corresponding to a positive fine frequency due to the lower harmonic. The linearity of the chirp transform fine frequency analysis is plotted in Figure 5.3.13, and is found to be linear to within a slope variance of less than a pixel, with a best resolution of about 260 Hz, corresponding to a FWHM of the chirp transform impulse response of between 3 and 4 pixels. A pair of phase av-

eraged impulse responses separated by 60 Hz are plotted in Figure 5.3.14, and they appear to be shifted by almost 1 pixel and they have the stated 3 to 4 pixel width, although they have an asymmetrically extended tail which is probably due to CCD charge transfer inefficiency. Within the two sided Nyquist limited analysis bandwidth of ± 7867 Hz, there are about 62 independent fine frequency measurements being performed, although 7 of these occur during the vertical blanking interval and do not produce an output peak within the video frame. The folded spectrum processor computes about 10^4 independent spectral observations at a 60 Hz rate, although the 24 MHz analysis bandwidth divided by the 260 Hz resolution achieved indicates that 10^5 spectral observations should be achievable if the coarse frequency resolution of the CCD could be improved. This processor thereby performs analog multiplication and additions at a rate of 6×10^9 operations per second, which has an FFT equivalent (N log N) processing rate of 7×10^6 operations per second. These processing rates can be greatly increased by using a higher resolution CCD detector array that can compute many more parallel spectral samples, and the utility of such a spectral analysis module will be greatly increased by using a CCD with a higher dynamic range.

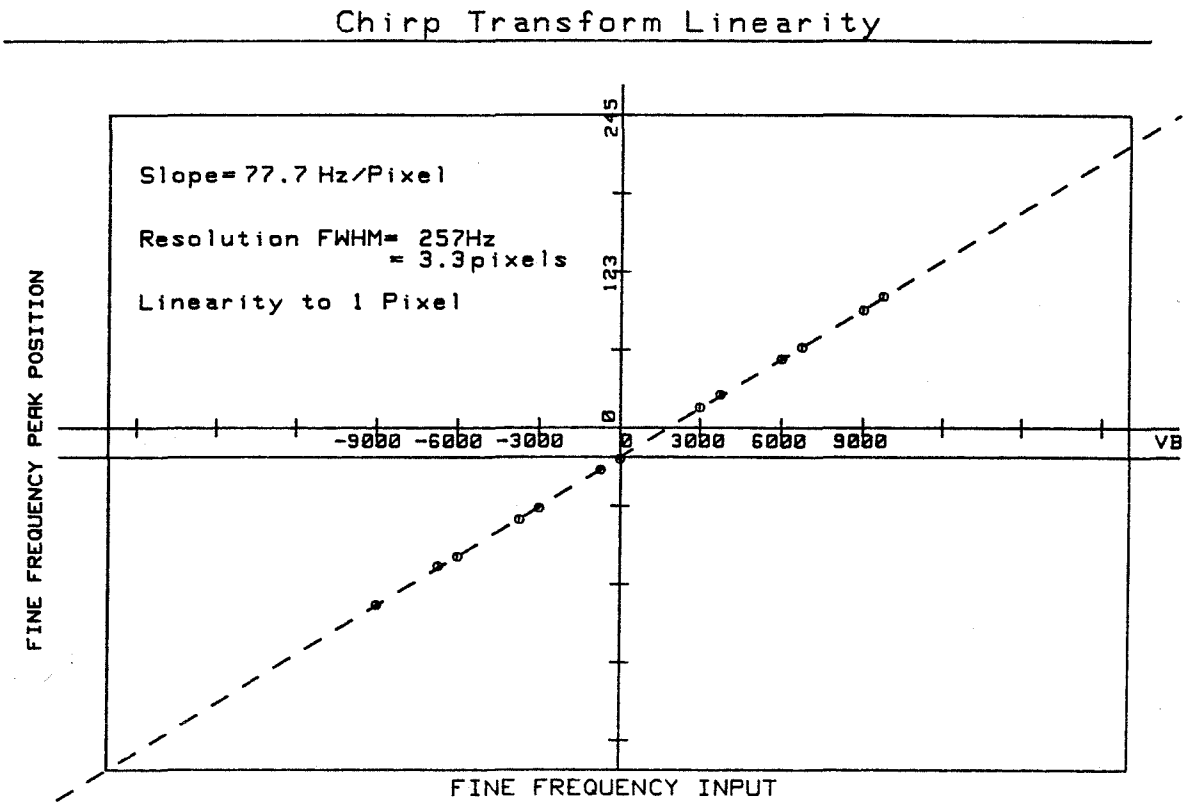


Figure 5.3.13. Linearity of the high space bandwidth chirp transform processor.

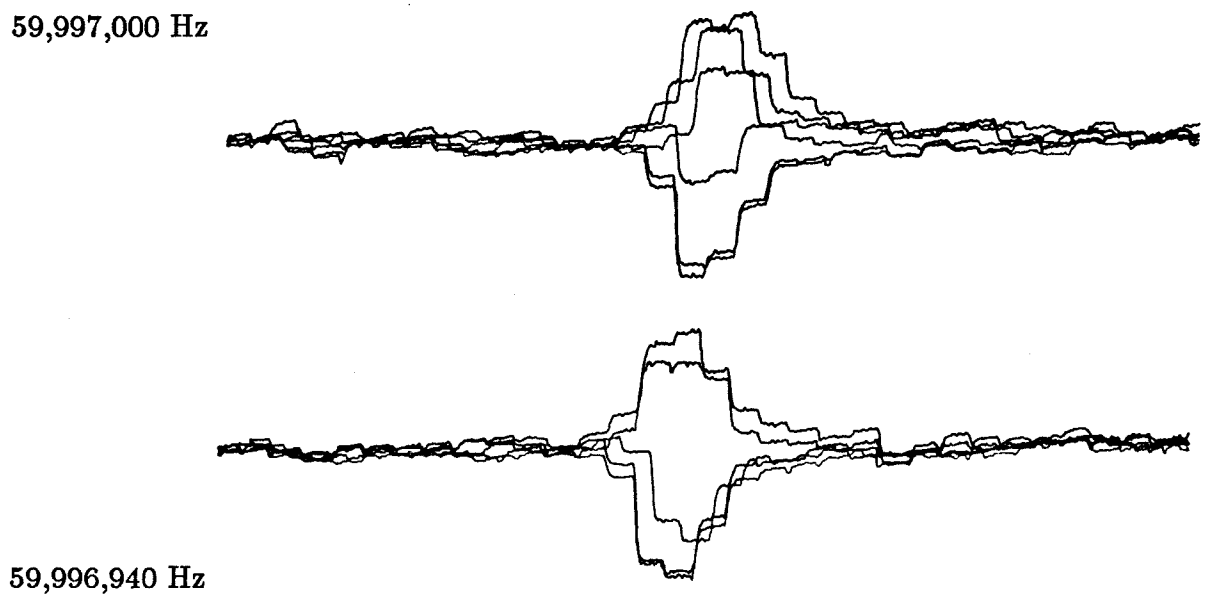


Figure 5.3.14. Phase averaged impulse responses of the TDI chirp transform processor for input tones separated by the integration time dictated fundamental resolution limit of 60Hz.

References

[1]. C. E. Thomas, Optical spectrum analysis of large space bandwidth signals, Appl. Opt., vol. 5, p. 1782 (1966).

[2]. D. Casasent, Optical Signal Processing, chap. 8 in *Optical Data Processing*, D. Casasent Ed., Topics in Appl. Phys. vol. 23, Springer-Verlag, Berlin (1978).

[3]. G. Lebreton, Power spectrum of raster scanned signals, Optica Acta, vol. 29(4), p. 413 (1982).

[4]. W.T. Rhodes, The falling raster in optical signal processing, SPIE vol. 373, (1981).

[5]. J.L. Anderson, H.B. Brown, and B.V. Markevitch, Wideband real-time Fourier analyzer using folded spectrum techniques, SPIE vol. 180-21, p. 128 (1979).

[6]. L. N. Flores, and D. L. Hecht, Acousto-optic signal processors, SPIE vol. 128, p. 118 (1977).

[7]. T. M. Turpin, Spectrum analysis using optical processing, Proc IEEE, vol. 69(1), p. 79 (1982).

[8]. P. Kellman, Time integrating optical signal processing, Ph.D. dissertation, Stanford Univ., Stanford, Ca, (1979).

[9]. P. Kellman, Time integrating optical signal processing, Opt. Eng., vol. 19(3), p. 370 (1980).

[10]. T. R. Bader ,Coherent optical hybrid techniques for spectrum analysis, SPIE vol. 185, p. 140 (1979).

[11]. T. R. Bader, Acoustooptic spectrum analysis: a high performance hybrid technique, Appl. Opt., vol. 18(10), p. 1668 (1979).

[12]. T.R. Bader, Coherent hybrid optical processors, SPIE vol. 232, p. 1 (1980).

[13]. D. Psaltis, and D. Casasent, Time- and space-integrating spectrum analyzer, Appl. Opt., vol. 18(19), p. 3203 (1979).

[14]. D. Psaltis, Image processing using acousto-optics, Proc. EOSD conf., Anaheim Ca. (1981).

[15]. D. Psaltis and D. Casasent, general formulation for optical signal processing architectures, Opt. Eng., vol. 19(2), p. 193 (1980).

[16]. D. Psaltis, Two dimensional optical processing using one dimensional input devices, Proc. IEEE, vol. 72(7), p. 962 (1984).

[17]. K. Wagner and D. Psaltis, Time and space integrating acousto-optic folded spectrum processing for SETI, SPIE vol. 564-31, (1985).

[18]. A Vander Lugt and A. M. Bardos, Spatial and temporal spectra of periodic functions for spectrum analysis, Appl. Opt., vol. 23(23), p. 4269 (1984).

CHAPTER 6

CONCLUSION

The multidimensional acousto-optical signal processing technique of temporal and spatial integration (TSI) has been investigated in this thesis, both theoretically and experimentally. Spatial variations of the diffracted wavefront from an acousto-optic device (AOD) are transformed with a multichannel optical processing systems, and the temporal variations of successively transformed wavefronts are used for time integrating processing in the orthogonal dimension of a 2-D detector array. This technique allows the utilization of highly mature one dimensional AOD technology to enter the data and spatially transform small segments, and two dimensional charge coupled devices (CCDs) in order to detect, time integrate, and read out the processed results. This approach can be used produce compact yet powerful signal processing systems for very high time bandwidth waveforms, which may find application niches in real time, on board systems, where power consumption, size, and weight are hard limited by the operating environment.

The capabilities of TSI architectures are well matched to image processing problems since the serial raster of the images produced by a video camera can be applied to the transducer of an AOD, and the processed image appears in real time at the processor output as another serial raster. Individual image raster lines are entered sequentially into the AOD, the acoustic motion is frozen by a pulsed laser, and the diffracted wavefront due to a single line is spatially transformed by a multichannel optical system. The succession of 1-D space integrated video lines are appropriately, multiplied, delayed, and accumulated on a 2-D CCD, in order to implement a time integrating transformation in the orthogonal dimension upon each spatially resolved channel produced by the space integrating system. TSI systems can be

configured to perform image correlations, separable linear transformations such as the Fourier transform, and some types of cascadable transformations.

Alternatively, very long 1-dimensional waveforms can be processed with the TSI technique by sliding small sections of the long signal into the AOD, spatially transforming these individual components, and temporally transforming the succession of space integrated portions in the orthogonal dimension of the 2-D detector array. This displays the long 1-D output in the 2-D coordinate space of the output detector, which turns it back into a long 1-D signal through the raster readout mechanism. Very high time bandwidth signals can be correlated, or separable transformations can be performed, such as the folded spectrum transformation.

Synthetic aperture radar (SAR) imaging is a cascadable two dimensional transformation with a range dependent azimuth kernel whose processing requirements are an excellent match for the TSI technique. Range focussing of high bandwidth chirp waveforms can be simply accomplished with free space propagation, and azimuth focussing can be accomplished on a TDI CCD multichannel correlator array, in which the azimuth kernel is range dependent. By using the TSI technique the high bandwidth capabilities of AODs can increase the achievable range resolution of real time SAR systems, and the video rate TDI capability of CCD detectors can increase the PRF rates that are obtainable with real time SAR processing. The complex kernel of the SAR transformation can be implemented with carrier encoding techniques, or an interferometer can be configured with an auxiliary bias subtraction CCD that performs the real part of the complex transformation.

Several different TSI architectures were considered for the folded spectrum pr ocessing of very high time bandwidth input signals. Two classes of architectures were considered in detail, the distributed local oscillator (DLO) interferometers, and the pulsed source processors. The DLO interferometers use CW visible light, and a pair of colinear signal and repetitive reference AODs whose spatial Fourier transforms are interfered in order to remove the position dependent coarse frequency

component of the signal spectrum. An additional curvature is introduced between the signal and reference wavefronts and the travelling temporal and spatial interferometric beats are detected on a 2-D TDI CCD detector array which performs the fine frequency analysis through a multichannel implementation of the chirp transform algorithm. The pulsed source processors use a pair of crossed AODs illuminated by a repetitively pulsed IR laser diode, and the transform of one AOD interferes with the image of another on the output detector array. The pulsed laser aliases the coarse frequency components of the signal spectrum to baseband, where they interfere with the reference for subsequent time integrating fine frequency analysis. The fine frequency analysis is performed with the aid of the orthogonal reference AOD, and the extra flexibility provided by the use of a pulsed source allows a wide range of different fine frequency analysis systems to be implemented.

The DLO interferometer has several potential advantages, since it uses CW visible light, there will be no problem with optical efficiency, and CCD resolution will be improved by operation in the visible regime. The sidelobe suppression which occurs because of the continuous time integration is an attractive feature not present in the pulsed source architectures. This is because the local oscillators used for interferometric baseband conversion are spatially channelized, so the sidelobes interfere with adjacent DLO beads, and produce moving fringes which time integrate to zero. While in the pulsed source processors the spectral harmonics of the laser diode which are used for baseband conversion are all available at all coarse frequency bins, so the sidelobes can chose the appropriate component which will alias to baseband and produce stationary fringes. The alignment of a pair of colinear AODs needed for the DLO interferometer should be easier than the alignment of a pair of orthogonal AODs needed for the pulsed source DFT processor, because all of the optical components can remain in the plane. However, the appropriate overlapping of the spectra may be quite sensitive for the DLO interferometer.

The pulsed source processors are light starved because of the low duty factor of

the laser, and the limited coherence available with pulsed laser diodes is a severe drawback, as is the resolution loss in the CCD due to the deep penetration at the IR wavelength. The time integrating fine frequency analysis of each coarse frequency bin can be performed with a TDI chirp transformation as in the DLO interferometers, or a discrete Fourier transformation (DFT) can be utilized. This is possible since the laser is pulsed, which allows entire columns of a space variant reference matrix to be slid into the reference AOD in the interval between pulses. This allows the implementation of a parallel array of time integrating DFT fine frequency spectrum analyzers operating on the temporal phase history within each resolvable coarse frequency bin produced by the space integrating spectrum analyzer. This DFT approach avoids the generation of the chirping carrier normally associated with the chirp transform, allowing simple carrier encoding, and it achieves equal gain and resolution at all frequency bins, unlike the triangular dependences of the chirp algorithm. However, the sequentially entered rows of the DFT matrix form an extremely high time bandwidth (TB) reference which must be entered into the processor, using an auxiliary high bandwidth digital frame buffer, rather than the simple low bandwidth chirp needed for TDI chirp transform processing. The TDI approach to fine frequency analysis has the distinct advantage of suppressing spatial noise due to coherent optical artifacts by incoherently averaging across a TDI detector column. This inherent advantage of the TDI approach led to the development of a new TDI DFT approach to fine frequency analysis. However, this new technique suffers from the same chirping carriers and variable gain and resolution as the conventional chirp transform, and it has a high TB reference waveform. So a main advantage of the DFT was lost by performing it in a scrolling coordinate frame, and the only gain was the incoherent averaging within a TDI column, which is more easily performed with a chirp transform approach. The simplest system with the widest analysis bandwidth that was implemented was based on a multichannel interferometric implementation of the TDI chirp transform fine frequency

analysis of the outputs of a pulsed source acousto-optic spectrum analyzer.

The crossed Bragg cell interferometer that was built for the application of folded spectrum processing successfully demonstrated the capabilities of these one dimensional devices to perform very high time bandwidth processing operations in the folded coordinate space of the two dimensional output detector array. The performance of the systems was inadequate for systems applications primarily because of two major factors, problems with the coherence of the laser diode, and the limitations of the CCD detector array. Coherence limitations of the laser diode can be minimized by appropriate pulse shaping, and by carefully equalizing the path lengths in the interferometer, or by using an in line approach. Beam collimation is critically important for obtaining highly coherent space integrated transformations across wide AOD apertures, which is necessary to obtain a complete frequency coverage of the folded spectrum. A technique for minimizing coherence requirements caused by Bragg diffraction was suggested, by aligning the transform lens with the undiffracted beam, thereby compensating for the path length differences caused by the diffraction. The CCD which was used in these experiments was an inexpensive industrial video camera, and it is expected that a scientific CCD will improve the obtainable SNR by more than an order of magnitude, and the resolution capabilities can be improved by almost an order of magnitude in each dimension. The resolution of silicon CCDs is significantly worse in the near IR than it is in the visible because of migration of deep carriers produced by the penetrating IR radiation. For this reason a visible pulsed source would be highly desirable, which could be accomplished with a cavity dumped gas laser, however the small size and low power advantage of laser diodes would be lost, along with the possibility of on board applications.

6.1 Future Research and Limitations of TSI

Optical signal processing techniques have been reasonably successful in a few extremely demanding signal and image processing tasks, such as synthetic aperture radar, and wideband spectrum analysis. As an example, a 1 GHz Bragg cell can perform a 1000 point SI Fourier transform each μsec, thereby computing 10^{12} analog multiplies per second with an interferometric dynamic range of 60dB, while obtaining an FFT equivalent processing rate of 10^{10} multiply-additions per second, which is quite respectable for such a simple system. However, modern digital signal processing techniques, using parallel, pipelined array processors and accelerated algorithms, have achieved a performance level that is beginning to rival the high throughputs of optical techniques, and with much higher fidelity, accuracy, repeatability and flexibility. For this reason it is necessary to go well beyond the current level of data throughputs and processing power obtainable with 1-D acousto-optic signal processing systems.

The approach taken in this thesis in order to enhance the processing power obtainable with these 1-D AO devices is to raise the dimensionality of the problems on which they can operate. Thus, video images, very high time bandwidth signal waveforms, and radar holograms were chosen as appropriate large dimensionality problems to be tackled by real time acousto-optic signal processing systems. However, practical limitations, such as available optical power, coherence, and CCD readout rates, decrease the speed with which data can be entered into and read out of these systems. An optimistic estimate of the performance capabilities of this type of system suggest that 100 MHz bandwidth data can be entered 1000 points at a time each 10μsec for a total of 10^6 points in 10msec. With the techniques presented in this thesis, these data can be Fourier transformed and read out on a parallel output CCD within this 10msec time window, thereby increasing the processing power to 10^{14} analog multiplies per second, but decreasing the multitone dynamic range to under 40dB, and the FFT equivalent processing rate drops to 2×10^9 multiply-

adds per second. Although the analog computation rate has increased considerably with the multidimensional system, the digital equivalent processing power and the dynamic range have decreased making this less competitive with digital techniques than the simple AO spectrum analyzer. The reason for the decrease in processing power is the existence of an accelerated digital algorithm (the FFT) for spectrum analysis, which can also be applied to convolutions and correlations. The high resolution spectrum analysis problem discussed in this thesis has been an attempt to discover the strengths and weaknesses of the TSI technique, and not an attempt to compete with digital signal processing alternatives in terms of raw processing power.

The most competitive application of TSI optical signal processing technology is for problems that do not have accelerated digital approaches. For instance the pulsed source DFT folded spectrum system has the capability of performing an arbitrary time integrating matrix vector product on the output of each channel of the spectrum analyzer, and this could result in a digital equivalent processing rate of 10^{11} multiplies per second. Another example of a more competitive TSI system might be the separable space variant kernel image transformation, or triple matrix multiplier, which can achieve a processing power of 10^{14} multiply-adds per second, which would require a processing rate of 2×10^{11} multiplies per second to implement digitally. The TSI technique makes use of all the information that is available in an acousto-optic system by processing it both spatially and temporally, so if optical signal processing is going to be competitive, then it will probably be through the hybrid technique of TSI. It is hoped that some type of TSI optical processing system can find a practical niche in real time signal processing applications, and that a significant performance advantage can be demonstrated, although this may only be possible in size, weight, or power, constrained applications.