SPATIAL TRACKING FOR OPTICAL HETERODYNE COMMUNICATION

by

Eric A. Swanson

Submitted to the Department of
Electrical Engineering and Computer Science
in partial fulfillment of the requirements
for the degree of

Master of Science in Electrical Engineering

at the

Massachusetts Institute of Technology

August 9, 1984

Signature of Author _____
Department of Electrical Engineering and Computer Science
August 9, 1984

Certified by _____
Vincent W.S.Chan
Thesis Supervisor

Accepted by _____
Arthur C. Smith
Chairman, Department Committee on Graduate Students

SPATIAL TRACKING FOR OPTICAL HETERODYNE COMMUNICATION

by

Eric A. Swanson

Submitted to the Department of
Electrical Engineering and Computer Science on
August 14, 1984, in partial fulfillment of the requirements
for the Degree of Master of Science in
Electrical Engineering

ABSTRACT

Optical heterodyne techniques are used to meet the stringent angular tracking requirements necessary for an optical communication system to operate properly. The spatial tracking system is assumed to operate in its linear region, use semiconductor lasers, and use a squaring loop to combat the frequency noise due to the lasers (the results are directly applicable to other types of lasers). The arriving optical field is assumed to be a plane wave. The results reveal a near optimum detector array and a near optimum L.O. field distribution. Also included are some preliminary experimental results. The experimental work is confined to tracking in one dimension while the theoretical work is more general. The experimental work shows good agreement with theory. The work was carried out at M.I.T. Lincoln Laboratory.

Thesis Supervisor: Dr. Vincent W. S. Chan
Title: Assistant Group Leader
Communication Technology
M.I.T. Lincoln Laboratory

ACKNOWLEDGEMENTS

I wish to thank John Kaufmann for taking the time from his busy schedule to provide support in both the theoretical and experimental work. His ideas, observations, and suggestions were a valuable contribution and are deeply appreciated.

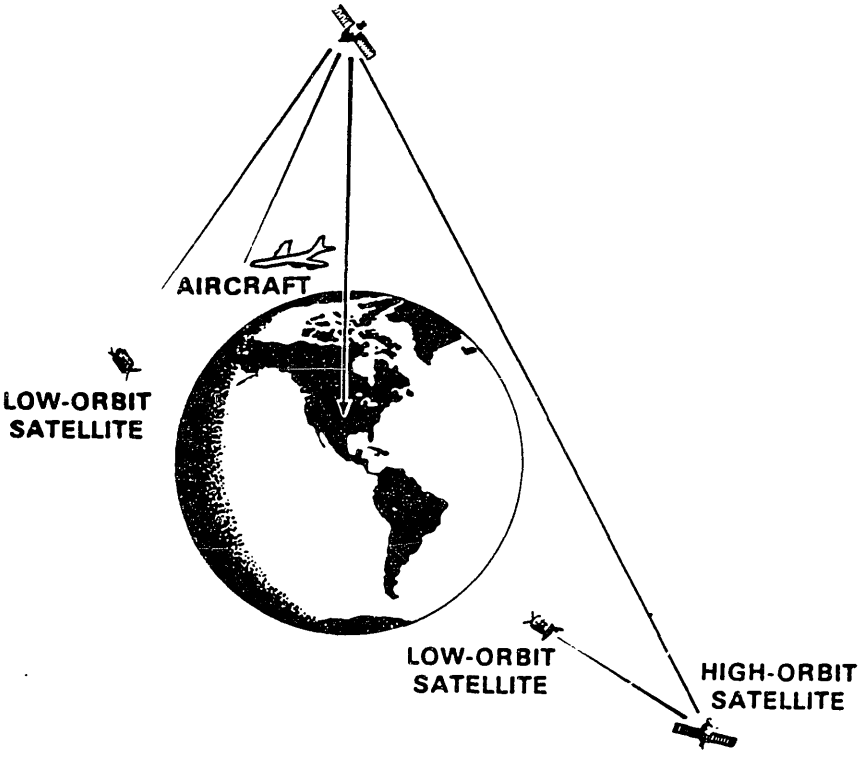
I would also like to thank Vincent Chan for his support and for providing the valuable opportunity to work in the Lincoln Laboratory, LASERCOM group. I would also like to thank the people in the LASERCOM group for their help throughout this work.

TABLE OF CONTENTS

To my Mother and to Lynne.

1. INTRODUCTION

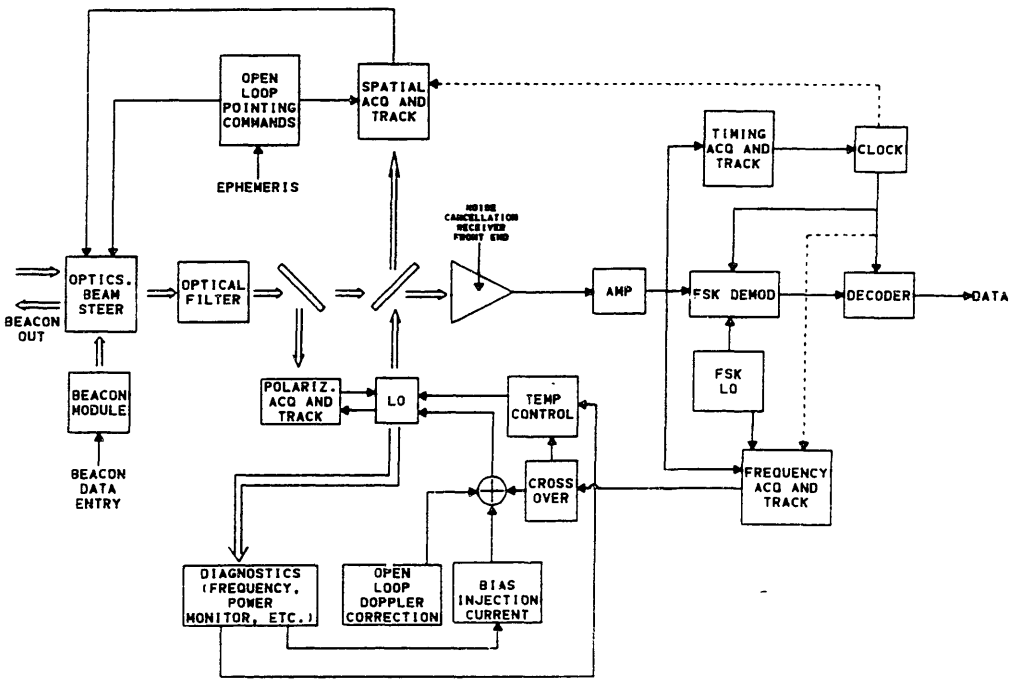
The development of Inter–Satellite Links (ISL) is a new and interesting area in satellite communications systems. Figure 1 shows an example of a satellite system using ISL's. Optical ISL's provide an attractive alternative to microwave technology [1,2]. For example: at optical frequencies relatively small apertures can provide extremely high antenna gain; optical frequencies provide a large bandwidth for high data rates; and the effects of spectral crowding are reduced.



Inter–satellite links.
Figure 1

Optical heterodyne communication systems at visible and IR wavelengths are an emerging technology potentially offering superior performance over the traditional and

technologically simpler direct detection systems [1,2]. For instance, in the heterodyne–detection receiver (Figure 2) at high local oscillator power, thermal noise and detector dark current noise are masked out by the quantum shot noise and near quantum–

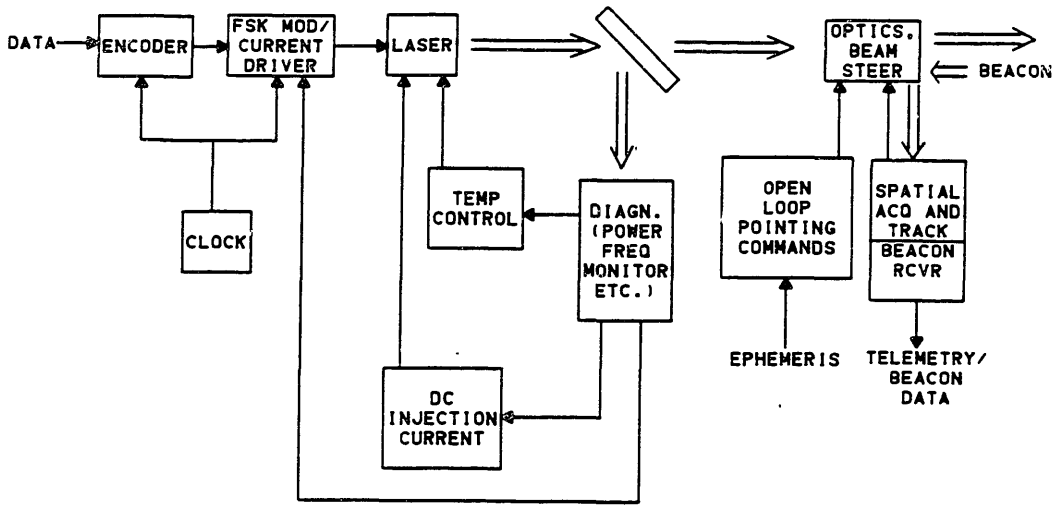


Heterodyne detection receiver.
Figure 2

limited performance can be achieved. Furthermore, the effects of background radiation are reduced, since any received optical field must heterodyne with the local oscillator in order to appear in the i.f. spectrum. In the direct detection receiver, performance is limited by detector noise, thermal noise, and background radiation. [1]

---

[1] When the sun is in the receiver field of view, direct detection receiver performance can be severely degraded.



Optical transmitter.
Figure 3

Figure 3 illustrates the basic features of an optical transmitter. This block diagram (and this thesis) assumes the use of GaAs diode lasers. GaAs diode lasers are a preferred choice for many reasons: for instance, the output optical field has a frequency proportional to the injection current which allows for various types of frequency modulation [1,2,3]; GaAs diode lasers have a high prime-to-output optical power conversion efficiency which is useful in ISL applications where limited prime power is available; and GaAs diode lasers are very compact. The actual lasing region occupies a volume on the order of 1 μm X 10μm X 100μm. This is useful in ISL applications, since both weight and size are critical issues in satellite design. However, frequency noise and excess intensity noise are the undesireable consequence of using GaAlAs lasers. To track out the phase noise requires a larger amount of signal power than is presently available in an ISL [4]. Therefore coherent demodulation is not possible. It is known that if the low frequency components of the frequency noise are

tracked out then the center frequency can be stabilized and heterodyne detection is possible [5].

In order for an optical heterodyne communication system to meet its performance predictions many technology issues still need to be solved. One of these issues is spatial tracking.[2] It is estimated that the spatial tracking system will need to hold the rms tracking error to less than .2 beamwidths, which corresponds to a received power loss of .5 dB [6]. For a wavelength of .8 μm and an aperture of 20 cm, this corresponds to an rms angular error of less than 1 μrad [7]. Most of the satellite's motion can be accurately predicted by the laws of physics, and is therefore deterministic motion. In contrast to this deterministic component, there is a random component that can only be estimated. The job of the spatial tracking system is to minimize the estimation error.

The spatial tracking system will be treated as two distinct subsystems in this thesis: one associated with the deterministic component of the satellite motion (often called the coarse tracking system) and one associated with the residual component (often called the file tracking system or image motion compensator).[3] The deterministic subsystem can be considered an open loop system that uses telemetry commands, past data, and a computer to track out the deterministic (or average) motion and calculate point–ahead angles.[4] The residual tracking subsystem, the focus of this thesis, is a closed loop system that operates on the difference between the actual position and the

---

[2] This tracking is in terms of angular tracking not absolute position.

[3] There is actually often a third subsystem that monitors the state of the mechanical tracking devices to prevent them from going out of range and saturating.

[4] Because of the large distances involved in ISL's (10^7 m), large satellite velocities (10^4 m/s), and narrow beamwidths (10 μrad) transmitting satellites must transmit ahead of the receiving satellite in order for its message to be received.

estimated (open loop) position, so as to minimize this difference.

The performance of the residual tracking system will obviously be influenced by the performance of other systems. For example, the polarization tracker, frequency tracking system, and the spatial tracking systems (of the other platform) all affect the performance of the residual tracker. However, this thesis will assume that most of these systems do a perfect job and that their quantities are available to the spatial tracking system. Assuming spatial acquisition has occurred and certain other quantities are known, the residual tracking system can be modeled as a distinct and separate system that tries to estimate the angle of arrival of an optical field. [5] Because of the long link distances involved (10^7 m), this arriving optical field can be considered a plane wave.

The intent of this thesis is to reveal some fundamental issues ,both theoretical and experimental, concerning the application of heterodyne techniques to the residual spatial tracking system. [6]

The spatial tracking system design will be mainly treated as an angular estimation problem. The estimator will be the minimum mean–square estimator based on a linearized observation equation (linearized via a small estimation error assumption) and assumed model for the angular process. Hence, implicitly assumed is that the tracking loop operates in its linear region. The following assumptions will also be made: 1. This system is distinct and separate from all other systems, and important quantities such as polarization are known. 2. Center frequency tracking will be carried out by a separate system. 3. The object of the spatial tracker is to track the incidence

---

[5] For a treatment of spatial acquisition see [8].

[6] For a treatment of direct detection spatial tracking systems see [9].

angle of an incoming optical field.

Some of the results that will be presented are: choice of detector array; optimum and near optimum L.O. patterns; bounds on the spatial tracking estimation error; and finally some experimental results will be presented and compared with theoretical predictions.

2. MODELING OF THE SPATIAL TRACKING SYSTEM

2.1. INTRODUCTION

This chapter is mainly devoted to characterize the first and second moment of the photo–currents. As mentioned in Chapter 1, the spatial tracking system design will be based on the first and second moments of a linearized observation equation. This observation equation will be simply related to the moments of the photo–currents and hence it is important to have an accurate (but tractable) description of the photo–currents.

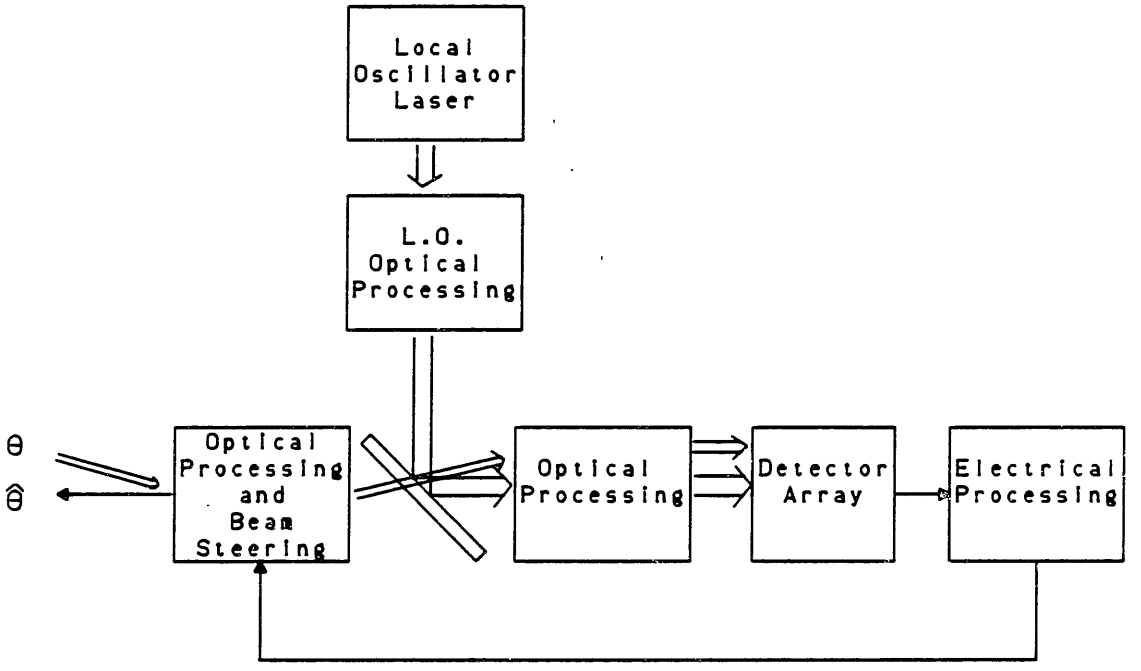
The next section deals with the system structure. There are actually many very different ways in which heterodyne detection can be accomplished, and as seen in section 2.2. the heterodyne detection in the conventional sense is chosen. There are also some further constraints imposed on the system discussed in the section. Section 2.3. gives a brief overview of the photo–detection process and gives general expressions for the first and second moments of the photo–currents. Section 2.4. deals with the field statistics as pertains to GaAlAs lasers and the spatial tracking system. These statistics along with the expressions in Section 2.3. can be used to accuratly describe the first and second moment of the photo currents. However, this accurate description is very cumbersome to deal with and many of the terms are insignificant. Section 2.5 will deal with determining which terms are dominant. It will also linearize the problem by making a small estimation error assumption and concluded with an simplified expressions for the photo–currents that are consistent in a mean–square sense.

2.2. SYSTEM STRUCTURE

A block diagram of the spatial tracking system structure is shown in Figure 4. [7] The optical processing and beamsteering unit receives an electrical control signal (representing the estimate of the angle of the arriving optical field) and steers the received optical field (assumed a plane wave) in accordance with this control. In addition, it performs optical processing such as angle multiplication and beam contraction. The actual structure and limitations of the optical processing units will not be discussed in this thesis. The Local Oscillator (L.O.) optical processing performs similar tasks and, in addition, spatially tailors the L.O. field to achieve optimum angle discrimination. The beam combiner combines the L.O. field and error signal field and directs them towards additional optical processing. It is assumed that focal place processing of the signal field is used. The two beams are sent onto a photo–detector array. The outputs of this array are sent to an electrical processor. This electrical processor operates on the data so as to estimate $\theta(t)$ (or equivalently to minimize $\Delta\theta(t)$). It then sends this estimate to the beam–steering unit. This completes the loop. It should be noted that $\theta(t)$ is actually a two dimensional signal, since tracking must occur in azimuth and in elevation.

---

[7] This obviously does not model all the possible optical heterodyne tracking systems but is sufficient for the scope of this thesis.



Optical heterodyne spatial tracking system structure.
Figure 4

2.3. THE PHOTO–DETECTION PROCESS

As mentioned in the introduction, the spatial tracking system design will be based mainly on the first and second moment of a linearized observation equation. This section will be concerned mainly with describing the first and second moment characterization of the photo–current generated by an optical field impinging on a photo–diode array.

When an optical signal illuminates a photo–detector surface, electrons are emitted and therefore a current is generated. The emission of electrons is a stochastic process. The statistics of this process can be related, in part, to quantum mechanical effects. [8] The emissions process when analyzed from a semi–classical point is known to be a conditional Poisson process [7]. Formally, the probability that k electrons will be emitted over the time interval t is

$$\Pr[N(t_0+t)-N(t_0)=k] = \int_0^{\infty} P_{m_v}(m)\,[m^k e^{-m}/k!]\,dm \qquad (1)$$

where
$$mv = 0 \quad t < t_0$$
$$= \alpha_A \int_{t_0}^{t_0+t} I(\rho,r)\,d\rho\,dr \quad t \geq t_0$$
$$= \int_{t_0}^{t} \lambda(\rho)\,d\rho \quad t \geq t_0$$
$$\lambda(t) = \alpha_A \int I(t,r)\,dr \quad \alpha=\eta/h\nu$$

I(t,r) is the magnitude of the Poynting vector, h is Plank's constant, ν is the optical frequency, and η is the quantum efficiency of the photo detector. $\lambda(t)$ is referred to as the count rate parameter and is a stochastic process whenever the intensity is a stochastic process. If the statistics of the intensity profile are known, then equation 1

---

[8] Namely, the Heisenberg Uncertainty Principle.

is a complete description of the emissions process.

The emissions process is not sufficient to describe the detector output currents. The electrons must travel from the point where they are emitted to the output terminals of the detector. This imposes many constraints. The one considered here is bandwidth. If the bandwidth of the detectors is very large relative to the fluctuations in the field intensity, or if the detector is to be cascaded with a filter with much narrower bandwidth, then the travel time of the electrons can be ignored. This will be assumed.

It is straight forward, but tedious, to show from the above that the mean and auto–correlation function of the detector output currents can be represented by

$$y_i(t) = \text{the current out of the } i^{th} \text{ detector} \tag{2}$$

$$m_y(t) = e\, m_\lambda(t)$$

$$K_y(t,s) = e^2\, m_\lambda(t)\, \delta(t-s) + e^2\, K_\lambda(t,s)$$

$$R_y(t,s) = e^2\, m_\lambda(t)\, \delta(t-s) + e^2\, R_\lambda(t,s)$$

$$R_{y_i y_j}(t,s) = e^2\, R_{\lambda_i \lambda_j}(t,s)$$

$$\lambda_i(t) = \alpha \int_{A_i} I(t,r)\, dr \qquad \alpha = \eta/h\nu$$

$m_y(t)$ represents the mean of the photo–current. e is the electron charge. $K_y(t,s)$ is the auto–covariance function of the photo–current and $R_y(t,s)$ is the auto–correlation function.

The first term in $K_y(t,s)$ represents quantum shot noise and is fundamental. The second term in $K_y(t,s)$ is due to the partial coherence of the light incident on the detector. In general this coherence can be due to both temporal and spatial fluctuations of the optical fields.

As a final comment, it can be shown that for high L.O. power, the filtered shot noise process will approach a Gaussian process [7,10]. Thus the shot noise term can be considered white Gaussian noise of spectral height $e^2 m_\lambda(t)$.

2.4. FIELD STATISTICS

2.4.1. Assumptions

The spatial tracking model consists of a planar array of detectors illuminated by a local oscillator and a signal field. The L.O. is assumed to be spatially coherent and have a symmetric intensity profile about the x axis and the y axis. Its amplitude can be divided into a deterministic component $U_{l_d}(\sigma)$, and a zero mean stochastic component $U_{l_r}(\sigma)$. The amplitude will be assumed to be a separable process – in space and time. Further, the the spatial term will be assumed totally coherent. The quantity L(t) will be used to denote the temporal randomness of the amplitude of the L.O. The intensity of the stochastic component is assumed small compared to the deterministic component. Also the L.O. is modeled as having a deterministic center frequency ω_l and a zero mean stochastic phase $\phi_l(t)$. Furthermore, the dependency between the amplitude and the phase processes is ignored in this thesis. It is believed that this is a valid approximation for the subject of interest. The signal is modeled similarly except that it is displaced spatially by a position vector (two dimensional) stochastic process $\epsilon(t)$. In terms of Figure 4, $\epsilon(t)$ is related to $\Delta\theta(t)$ for small $\epsilon(t)$, by

$$\epsilon(t) \approx f\,\Delta\theta(t) \tag{3}$$

where f is the effective focal length of the optical system. $\epsilon(t)$ represents the spatial displacement of the centroid of the signal field from the center of the detector array. The intensity of the signal field is assumed small compared to the intensity of the L.O. field and the object is to minimize $\epsilon(t)$ in a mean square sense. Thus, we have the following model[9]

---

[9] σ and ρ will represent either one dimensional time coordinates or two dimensional

$$u_l(t,\sigma) = \mathrm{Re}[\, U_l(t,\sigma)\, e^{-j(\omega t + \phi_l(t) + \phi_l(t))} \,] \tag{4}$$

$$U_l(t,\sigma) = U_{l_o}(\sigma) + U_{l_r}(t,\sigma)$$

$$= (1 + L(t))\, U_{l_o}(\sigma)$$

$$u_s(t,\sigma - \epsilon(t)) = \mathrm{Re}[\, U_s(t,\sigma - \epsilon(t))\, e^{-j(\omega t + \phi_s(t) + \phi_s(t))} \,] \tag{5}$$

$$U_s(t,\sigma - \epsilon(t)) = U_{s_o}(\sigma) + U_{s_r}(t,\sigma - \epsilon(t))$$

$$= (1 + S(t))\, U_{s_o}(\sigma - \epsilon(t))$$

2.4.2. Intensity Noise Characterization

L(t) (and S(t)) will be modeled as a zero mean stationary Gaussian process with independent, identically distributed, real and imaginary parts. This model has been experimentally shown to yield the correct behavior for the single detector case [11,12]. Thus,

$$L(t) = L_r(t) + iL_i(t) \tag{6}$$

L(t) will give rise to additional noise in the photo–currents and mathematically appears in the photo–currents via the second term in the covariance expressions of equation 2. It arises from the excess intensity noise of the laser. It is known that this excess intensity noise can be approximated as spectrally white in the region of 1 MHz to 1 GHz. It is assumed that the i.f. bandwidth of the spatial tracking system is well within this range and thus the excess intensity noise will be considered white Gaussian noise. Therefore, neglecting the d.c. components,

---

space coordinates, and should be clear from the context.

$$E[\,|U_1(t,\sigma)|^2\,|U_1^*(s,\rho)|^2\,] = N/2\,|U_{L_o}(\sigma)|^2\,|U_{L_o}(\rho)|^2\,\delta(t-s) \qquad (7)$$

When $\epsilon(t)$ is identically zero the results for the signal field are identical.

As a final comment, the excess noise can be comparable to the quantum shot noise [12]. Thus, in some instances excess noise can have a detrimental effect on the performance of the spatial tracking system. However, it will turn out that the electrical processor to be considered later on will subtract the detector currents from adjacent quadrants in order to estimate the signal field arrival angle. It is straight forward to show from equations 2, 4, and 7 that the excess noise can be (almost) totally cancelled by either delaying the L.O. field a half wavelength in the second and fourth quadrants relative to the first and third quadrants, or by using two identical detector arrays with one positioned on each side of the beamsplitter. [10] The effect of excess noise has been included in case it is decided not to implement a cancellation scheme.

2.4.3. Frequency Noise Characterization

As mentioned in Chapter 1, frequency noise will be present in the photo–currents and, as a result, coherent demodulation of the photo–currents is not possible. As a result some sort of noncoherent detection will be used in the tracking loop. This frequency noise will have a large impact on the spatial tracking system and it will be shown that the performance of the spatial tracking system will be dependent on how much frequency noise there is. This subsection describes the frequency noise spectral density.

---

[10] The latter method is more robust and is discussed in [13].

The frequency[11] noise of the detector currents will be a function of the difference of the L.O. frequency noise and the signal frequency noise.

$$\dot{\phi}(t) = \dot{\phi}_{l}(t) - \dot{\phi}_{s}(t) \tag{8}$$

where $\dot{\phi}(t) = d\phi(t)/dt$

The frequency noise is assumed to be a stationary white Gaussian process. The actual spectrum has been shown to be stationary white Gaussian plus a dominant Gaussian 1/f component at low frequencies (<100kHz) [14]. Using a phase-lock loop to track the phase of the I.F. signal is difficult because of the moderate signal powers [2]. Furthermore, because the white spectrum extends to very high frequencies, bandwidth limitations dictate that the best one can do is to track out the 1/f component to stabilize the center frequency [15]. It will be assumed that this frequency tracking information will be available to the spatial tracking system. Thus, all that will be considered is the white component of the frequency noise. It should be noted that if frequency tracking is not implemented, the lasers will drift in frequency and the i.f. spectrum will eventually drift out of the i.f. filter bandwidth. [12] Therefore, it is essential that frequency tracking be used to stabilize the i.f. spectrum center frequency.

Assuming that the frequency noise is white and Gaussian, it is straight forward to show that [16] if

$$S_{\dot{\phi}}(f)=L \quad (\text{rad/sec})^2/\text{Hz} \tag{9}$$

Then

---

[11] The frequency noise will be in terms of radian frequency.

[12] Laser frequency is very sensitive to temperature (Ghz/ °C) and current drifts (Ghz/mA).

$$E[\, e^{-j(\omega(t-s)+\phi(t)-\phi(s))} \,] = e^{-j\omega(t-s)}\, e^{-L|t-s|/2} \qquad (10)$$

Therefore the i.f. signal has a Lorentzian spectrum. This Lorentzian lineshape has a full–width, half–maximum spectral linewidth of $L/2\,\pi$. It will be shown that the spatial tracking performance is directly dependent on the parameter $L/2\,\pi$ for low signal–to–noise ratio.

2.5. APPROXIMATION OF THE MEAN AND COVARIANCE

To find the first and second moments for the detector currents, the field expressions in equations 4 and 5 are substituted into the moment equation 2. When designing the electrical processor (Figure 4) it is not necessary to take into account all these terms. This section is concerned with determining what terms are dominant and finding a new (effective) mean and covariance equation based on these dominant terms. Also a small estimation error assumption will be made. This is the first step in the linearization of the observation equation.

Using equations 4 and 5 in equation 2, and the fact that the signal power is assumed much smaller than the L.O. power over each detector surface and that the power in the random component of the intensity is small compared to the power in the deterministic component

$$m_\lambda(t) = \alpha \int_{A_i} |U_{l_0}(\sigma)|^2 \, d\sigma \tag{11}$$

Now to find $R_{\lambda_i\lambda_j}(t,s)$. Since the power in the signal is small compared to the power in the L.O. over each detector, this is straight forward.

$$\begin{aligned} R_{\lambda_i\lambda_j}(t,s) = \alpha^2 E[\int_{A_i} \int_{A_j} (& |U_l(t,\sigma)|^2 \, |U_l(s,\rho)|^2 \\ & + U_l(t,\sigma)\, U_s^*(t,\sigma - \epsilon(t))\; U_l^*(s,\rho)\, U_s(s,\rho - \epsilon(t))\, e^{-j(\omega(t-s)+\phi(t)-\phi(s))} \\ & + (c.c)) \; d\sigma \, d\rho \;] \end{aligned} \tag{12}$$

where c.c. denotes complex conjugate of previous term

The above equation is, in general, nonlinear in $\epsilon(t)$. If the tracker is working well, $\epsilon(t)$ will be small. Under this assumption $U_s(t,\sigma - \epsilon(t))$ can be linearized by taking

the first two terms of the Taylor series expansion about $\epsilon(t)$.

$$U_s(t,\sigma-\epsilon(t)) \;\alpha\; U_s(t,\sigma) - \epsilon_x(t)[\partial U_s(t,\sigma)/\partial Ux] - \epsilon_y(t)[\partial U_s(t,\sigma)/\partial y] \qquad (13)$$

After substituting equations 4, 5, and 13 into equation 12, five types of terms appear and are listed in Appendix 7.1. These terms are categorized as follows: 1) Terms that involve only the L.O. intensity noise; 2) Terms that involve the L.O. phase noise and the signal phase noise; 3) Terms that involve the L.O. phase noise, the signal phase noise, and the spatial error angle; 4) Terms that involve the L.O. phase noise, the signal phase noise, and the excess intensity noise (of either the L.O., signal, or both); 5) Terms that involve the L.O. phase noise, the signal phase noise, the excess intensity noise (of either the L.O., signal, or both), and the error angle. It is assumed that the spatial derivative of the signal is well behaved and that the power in the L.O. is still much larger than the its power over all detector surfaces. Under these conditions, it will be assumed that terms 4 and 5 are negligible compared to terms 2 and 3, and will therefore be dropped. This is an important assumption and may not always be valid. Eventually it may be necessary to modify this assumption. Thus, we assume the excess intensity noise enters the covariance equation only additively.

It is convenient to normalize each detector current by

$$y_{n_i}(t) = (\sqrt{2}\;\alpha e \int_{A_i} |\, U_{i_o}(\sigma)\,|^2\, d\sigma)^{-1/2}\; y_i(t) \qquad (14)$$

and also to define A_i, B_i, and m_i as

$$A_i = \int_{\Omega_i} U_{s_o}^*(\sigma) U_{L_o}(\sigma)\, d\sigma \;/\; Ps^{.5} \qquad (15)$$

$$B_i = \left(\int_{\Omega_i} |U_{L_o}(\sigma)|^2\, d\sigma \right)^{.5}$$

$$m_{i_x} = \int_{\Omega_i} [\partial U_{s_o}^*(\sigma)/\partial x]\, U_{L_o}(\sigma)\, d\sigma \;/\; (A_i\, Ps^{.5})$$

$$m_{i_y} = \int_{\Omega_i} [\partial U_{s_o}^*(\sigma)/\partial y]\, U_{L_o}(\sigma)\, d\sigma \;/\; (A_i\, Ps^{.5})$$

As mentioned in Chapter 1, it is assumed that focal plane processing is to used. Therefore, the optical field distributions over the detector array can be found from the Fraunhoffer diffraction equation [7]. It is further assumed that the constant quadratic phase factor (arising from the Fraunhoffer equation) in the L.O. field is matched to the quadratic phase factor in the signal field at the focal place. It can be shown that if this condition is not met, performance will be degraded. Therefore the quantities in equation 15 are real.

To summarize this section, a set of effective detector currents that are consistent with the mean and covariance equations of the actual normalized detector currents can be represented by[13]

---

[13] It is assumed that ϵ^2 is small compared to 1.

$$y_{n_i}(t) = r_i(t) + n_i(t) + w_i(t) \tag{16}$$

$n_i(t)$ and $w_j(t)$ are independent white Gaussian processes

$$r_i(t) = \sqrt{2Ps}\ A_i/B_i\ (1 - m_{i_x} \epsilon_x(t) - m_{i_y} \epsilon_y(t))\ \cos(\omega_0 t + \phi(t))$$

$$= \sqrt{2Ps}\ S_i(\epsilon(t))\ \cos(\omega_0 t + \phi(t))$$

$$R_{w_{ij}}(t,s) = No/2\ \delta(t-s)\ \delta_{ij} \quad No = h\nu/\eta$$

$$R_{n_{ij}}(t,s) = NB_iB_j/2\ \delta(t-s)$$

$$R_{cc}(t,s) = E[\ \sqrt{2Ps}\ \cos(\omega_0 t + \phi(t))\ \sqrt{2Ps}\ \cos(\omega_0 s + \phi(s))\ ]$$

$$= Ps\ e^{-L|t-s|/2}\ \cos(\omega_0(t-s))$$

$$S_{cc}(f) = 2Ps\ (L/2)\ /\ [(L/2)^2 + (\omega - \omega_0)^2]$$

$n_i(t)$ represents the excess intensity noise and $w_i(t)$ represents the quantum shot noise. R stands for auto–correlation function and S for spectral density. ω_0 represents the i.f. center frequency.

3. PERFORMANCE OF THE SPATIAL TRACKING SYSTEM

3.1. AN ESTIMATION ERROR LOWER BOUND

This section will be concerned with describing a lower bound on the mean–square estimation error of the spatial tracking system. In order to attain quantitative results, it is necessary to assume some model for the angular process. It will be assumed in the theoretical analysis presented here that the angular process that is to be tracked is a two–dimensional uncorrelated Wiener process. [14] Three idealizations will be made and it is because of this that the bound to be presented is considered a lower bound instead of the actual estimation error. These assumptions are: 1. That the beam steering unit has infinite bandwidth and infinite precision; 2. That the loop operates in its linear region; 3. That there is no phase noise in the i.f. signal. The detector array contains $4N^2$ elements where N^2 is the number of elements in each quadrant. Figure 5 shows the indexing that will be used. [15]

Under the above assumptions, the model for the normalized observation equation is

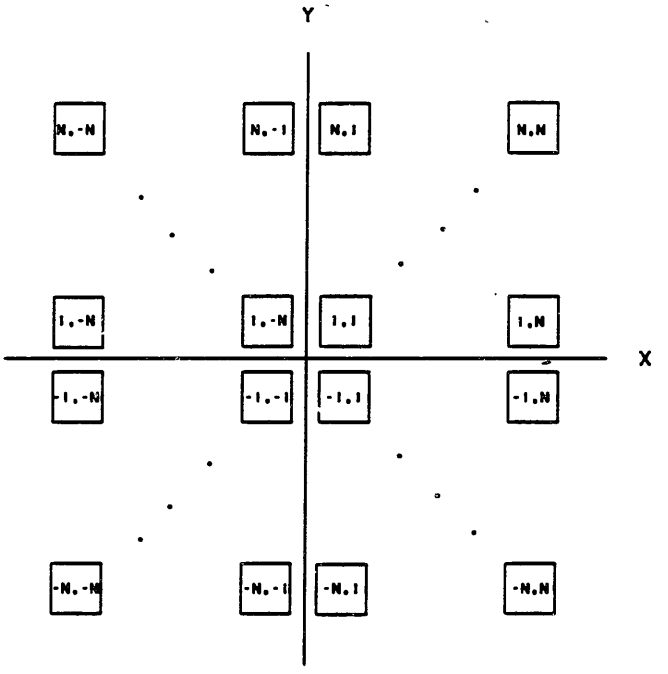
$$\text{O.E.} \qquad y_{n_{ij}}(t) = \sqrt{2Ps}\, A_{ij}/B_{ij}\, (1 - m_{x_{ij}} f\theta_x(t) - m_{y_{ij}} f\theta_y(t)) \cos(\omega_0 t) + w_{ij}(t) + n_{ij}(t) \qquad (17)$$

By the field symmetry assumptions

---

[14] The choice of a Wiener process is somewhat arbitrary; however, the procedure can be (and has been) straight forwardly applied to any angular process.

[15] It will be convenient here to introduce two subscripts to identify each detector element instead of one.



4 N^2 detector array.
Figure 5

$$A_{ij} = A_{i-j} = A_{-i-j} = A_{-ij} \tag{18}$$

$$B_{ij} = B_{i-j} = B_{-i-j} = B_{-ij}$$

$$m_{x_{ij}} = -m_{x_{i-j}} = -m_{x_{-i-j}} = m_{x_{-ij}}$$

$$m_{y_{ij}} = m_{y_{i-j}} = -m_{x_{-i-j}} = -m_{y_{-ij}}$$

Finding the optimum, linear, steady–state, mean–square estimation error associated with a linear observation equation that can be represented in state space reduces to solving a matrix Riccati equation [17]. In general, there is no closed form solution to the Riccati equation. However, for the present model there is. The state equation for the angular process is

S.E. $$\dot{\theta}(t) = U(t) \tag{19}$$

$$\theta(t) = \begin{vmatrix} \theta_x(t) \\ \theta_y(t) \end{vmatrix}$$

$$E[\, U^T(t)U(s)\,] = \sigma^2 I_2\, \delta(t-s)$$

σ^2 has units rad^2 / sec

Using the symmetry properties of the fields it is straight forward but tedious to show that the steady–state solution to the Riccati equation, or equivalently, the steady–state, mean–square error is

$$\xi_{xx} = \sigma / (8\ Ps/N_0\ F_x)^{1/2} \tag{20}$$

$$\xi_{yy} = \sigma / (8\ Ps/N_0\ F_y)^{1/2}$$

$$\xi_{xy} = 0$$

where $$\xi_{xy} = E[\, (\theta_x(t) - \hat{\theta}_x(t))(\theta_y(t) - \hat{\theta}_y(t))\,]$$

$$F_x = \sum_{ij=1}^{N} ((A_{ij}/B_{ij})m_{x_{ij}})^2$$

$$F_y = \sum_{ij=1}^{N} ((A_{ij}/B_{ij})m_{y_{ij}})^2$$

It is interesting to see that the excess intensity noise has no effect on the performance of the spatial tracking system This is because the electrical processor subtracts detector currents from adjacent quadrants in order to form its estimate of the error and the excess intensity noise in each detector comes from the same sample function.

It should be noted that although a Wiener angular process was considered, the quantities

$$4\ Ps/N_0\ F_x \tag{21}$$

and

$$4\ Ps/N_0\ F_y$$

are fundamental to the spatial tracking system. It can be shown that these quantities appear for any angular process obeying a stochastic differential equation. They represent the effective number of photons--per--second received by the spatial tracker. Although the mean--square error for other angular processes may have a different dependence on these quantities, it is reasonable to expect that for any angular process the mean--square error is non--decreasing as the quantities in equation 21 are increased.

A lower bound for equation 20 is provided by the Schwartz--inequality. That is

$$((A_{ij}/B_{ij})m_{x_{ij}})^2 \leq \int_{\Omega_{ij}} |\partial U_s(\sigma)/\partial x|^2 \, d\sigma \tag{22}$$

Therefore

$$\xi_{xx} \geq \sigma / (2\, Ps/N_0\, f^2 \int_{\Omega} |\partial U_s(\sigma)/\partial x|^2 \, d\sigma/Ps)^{1/2} \tag{23}$$

$$\xi_{yy} \geq \sigma / (2\, Ps/N_0\, f^2 \int_{\Omega} |\partial U_s(\sigma)/\partial y|^2 \, d\sigma/Ps)^{1/2}$$

where Ω = The detector array region

Equation 23 is essentially a Cramer--Rao bound on the spatial tracking error for a Wiener process.

In general there is no L.O. pattern that will achieve both lower bounds simultaneously. However, for tracking in one dimension the lower bound can be achieved and the optimum L.O. pattern is independent of the number of detectors and the detector layout (as long as equation 13 remains valid). This L.O. pattern is easy to find and is given by

$$U_l(\sigma) = k\, (\partial U_s(\sigma)/\partial x) \tag{24}$$

where k is a normalization constant

For tracking in two dimensions the optimum L.O. pattern is much more difficult to find. In principle, equation 20 and equation 15, along with some cost function, can be used to find the optimum L.O. pattern for a given number of detectors and a given detector layout. It is straight forward to solve this problem for quadrant detector (N=1) and the following cost function

$$C_t = \xi_{xx}^2 + \xi_{yy}^2 \tag{25}$$

The result agrees with intuition (an overview of the derivation is given in Appendix 7.2.)

$$U_l(\sigma) = k\,((\partial U_s(\sigma)/\partial x) + (\partial U_s(\sigma)/\partial y)) \tag{26}$$

where k is a normalization constant

Due to the symmetry of the signal field equation 26 is well approximated by the radial derivative for many fields of interest.

For the more general case one thing is clear: the L.O. pattern should have no components orthogonal to both the signal derivative in x and the signal derivative in y. Therefore the optimum L.O. pattern is of the form

$$U_l(\sigma) = \sum_{ij=-N}^{N} l_{xij} S_{xij}(\sigma) + l_{yij} S_{yij}(\sigma) \tag{27}$$

where

$S_{xij}(\sigma) = \partial U_s(\sigma)/\partial x$ over Ω_{ij}

$S_{yij}(\sigma) = \partial U_s(\sigma)/\partial y$ over Ω_{ij}

l_{xij} and l_{yij} are weighting coefficients

Assuming the signal field pattern over the detector array is an Airy–disc pattern and assuming infinite detector area, equation 23 reduces to

$$\xi_n \geq \sigma_n / (\pi^2/2\ Ps/N_0)^{1/2} \tag{28}$$

where

$$\xi_n = \xi / (\lambda/d)^2$$

$$\sigma_n = \sigma / (\lambda/d)$$

Assuming the above, and further that the detector array is an infinite area quadrant detector and the L.O. is given as in equation 26, then equation 20 reduces to

$$\xi_n = \sigma_n / (\pi(\pi+2)/4\ Ps/N_0)^{1/2} \tag{29}$$

The loss in equation 29 compared to the unrealizable bound in 28 is less than .5 dB. Assuming as above except that the L.O. pattern is the radial derivative of the Airy–disc, then equation 20 reduces to

$$\xi_n = \sigma_n / (4\ Ps/N_0)^{1/2} \tag{30}$$

$$(A/B)m = (d/\lambda)/2^{1/2}$$

The difference between equation 29 and 30 is a small fraction of a dB.

It is useful to compare the above result with the best performance that can be attained with a plane wave L.O. pattern, Airy–disc signal field, and an annules quadrant detector. The normalized variance and optimum inner and outer radius are given by

$$\xi_n = 1.21\ (\sigma_n / (4\ Ps/N_0)^{1/2}) \tag{31}$$

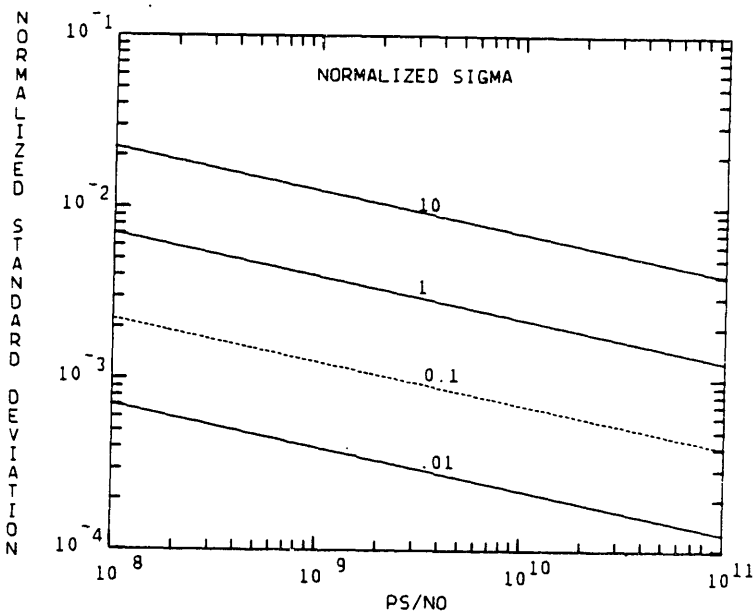
$$R_{in} = .183\ \lambda f/d$$

$$R_{out} = 1.34\ \lambda f/d$$

There is less than 1 dB loss in (realizable) performance in using in using a plane wave L.O. It should be noted that R_{in} can be extended to zero with little loss in

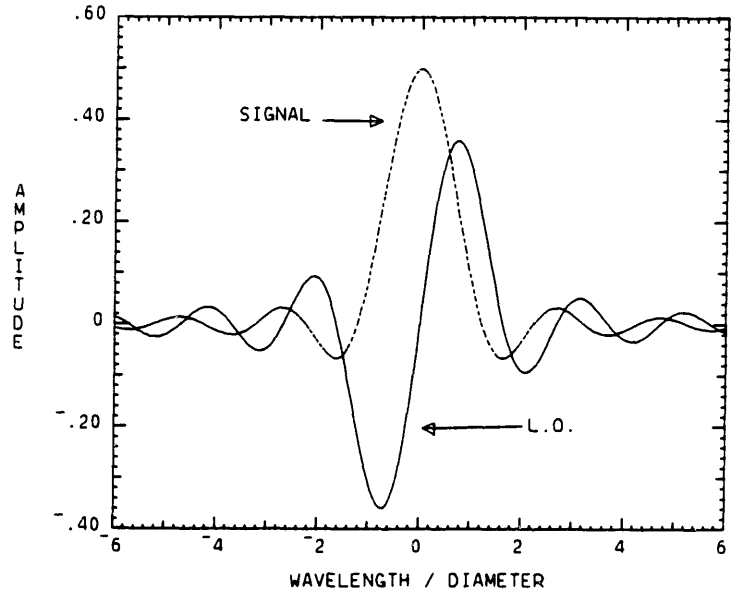
performance.

These results reveal that the there is not much room for improvement by considering detector arrays other than the quadrant detector or L.O. patterns other than the radial derivative of the signal field. Figure 6 shows equation 30 plotted vs Ps/N_o for four values of σ_n. Ps/N_o is an important parameter, it is the optical signal power times the quantum efficiency divided by the photon energy. It represents the average number of photons-per-second incident onto the detectors times the quantum efficiency of the detectors. Of course, not all this signal power is detected since the L.O. field is not matched to the signal field.



Tracking error for a Wiener process (coherent demodulation).
Figure 6

Figure 7 shows the Airy-disc signal pattern and the radial derivative of the Airy-disc L.O. pattern (both normalized). The optimum L.O. pattern can be interpreted as follows: Any L.O. power concentrated in areas where the signal field is relatively flat



Signal field and L.O. field distributions.
Figure 7

(or very weak) will be detrimental because (for small estimation errors) there is no angular information in these areas and appling L.O. power there will only introduce unnecessary shot noise. The reason that there is no angular information in these areas is because the electrical processor subtracts currents from adjacent quadrants and the signal power that is (almost) always common to all quadrants will disappear in the subtraction. This interpretation is an important guideline in choosing what L.O. pattern is to be used in the actual system implementation.[16] A second interpretation of the optimal L.O. pattern is more easily understood for tracking in one dimension. The optimum L.O. pattern is given by equation 24. This is simply a scaled version of the information part of the signal field, and since the signal field is symmetric (except for a minus sign) the optimum L.O. pattern can be seen to be an "optical matched filter".

---

[16] Since creating the exact radial derivative might require to much hardware to warrant its use a suboptimal L.O. pattern might be chosen.

For two dimensional tracking, the L.O. field is a compromise between the matched filter desired by the azimuth tracking and the matched filter desired by the elevation tracking this can be seen in equation 26.

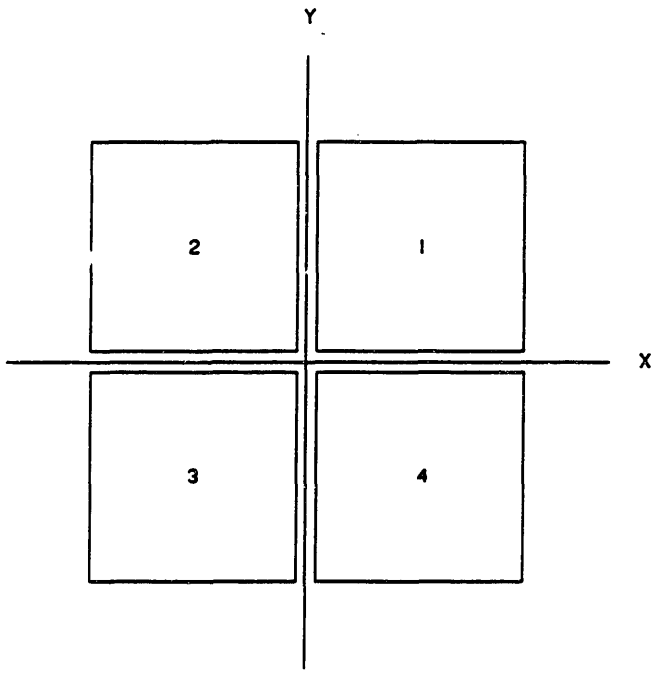
This section has dealt with the issue of heterodyne spatial tracking in the absence of frequency noise, and finite bandwidth and precision considerations of the beam–steering devices. When these effects are added to the observation equation, the analysis of the spatial tracker is much more complicated. However, equations 20 and 23 still represent bounds on the performance. Furthermore, for high signal–to–noise ratio the performance of the spatial tracker should approach the bound in equation 20 (barring beam steerer bandwidth and precision problems).

3.2. TRACKING PERFORMANCE USING A SQUARING LOOP

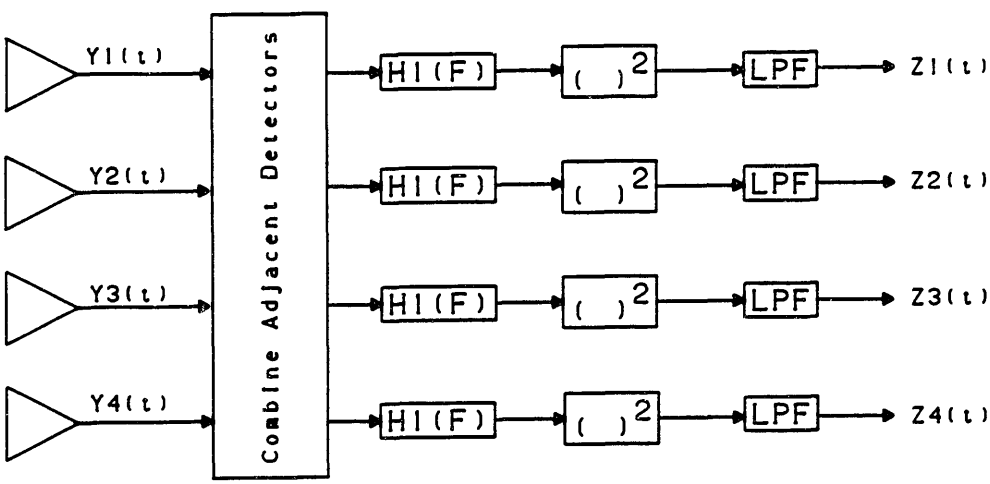
This section will deal with the important issue of phase noise in the observation equation. When phase noise is present, the optimum mean–square estimator is, in general, very difficult to find. The normal procedure would be to use a phase–lock loop to track out the phase noise (even this is not optimum). However, with a moderate data rate system there will not be enough signal power to effectively track out the phase noise as mentioned earlier. To combat the phase noise it will be assumed that the spatial tracking loop will employ the use of squaring devices.

It will be shown that in the range of interest, the squaring loop estimation error will be very close to the lower bounds of section 3.1. and hence their use is justified. The important assumption that the beamsteering devices have infinite bandwidth and infinite precision is still in effect. So in a sense the results are really lower bounds. As in section 3.1. it will be assumed that the angular process to be tracked is a Wiener process. Also, in light of the results of section 3.1. (and in light of implementation considerations) the detector array will be a quadrant detector. Figure 8 shows the indexing that will be used.

The squaring–loop has the property that when the tracker is working well, its output is in the form of signal plus noise and therefore finding the optimum, linear, steady–state estimator (based on the squaring–loop output) and its associated estimation error is relatively easy. The squaring–loop to be used is shown in Figure 9. Note that the outputs of adjacent detectors are combined before squaring and subsequent processing. It can be shown that this yields superior performance over the case where each detector output is first squared and then processed. Intuitively, combining before squaring is superior because the signal components add



Quadrant detector array.
Figure 8



Squaring filter.
Figure 9

constructively, whereas the noise processes will not. It is assumed, for simplicity, that the bandwidth of $H_1(f)$ is large compared to $L/2\pi$ so that $r_i(t)$ is passed through $H_1(f)$ undisturbed. Essentially $H_1(f)$ is used to bandlimit the noise before squaring. It is also assumed that $H_1(f)$ is symmetric about its center frequency. The sole purpose of the low pass filter is to remove the double frequency terms.

Let $H_1(f)$ be represented in terms of its low pass part by

$$H_1(f) = H_0(f-f_0) - H_0(f+f_0) \tag{32}$$

Under the above assumptions the normalized output of each detector is given by equations 16. By the symmetry assumptions

$$A_1 = A_2 = A_3 = A_4 \tag{33}$$

$$B_1 = B_2 = B_3 = B_4$$

$$m_{x_1} = -m_{x_2} = -m_{x_3} = m_{x_4}$$

$$m_{y_1} = m_{y_2} = -m_{y_3} = -m_{y_4}$$

With the above assumptions the input to the squarer can be represented as

$$r_i(t) + w_i(t) + n_i(t) \tag{34}$$

where

$$w_i(t) = \sqrt{2}(w_{ci}(t)\cos(\omega_0 t) - w_{si}(t)\sin(\omega_0 t))$$

$$n_i(t) = \sqrt{2}(n_{ci}(t)\cos(\omega_0 t) - n_{si}(t)\sin(\omega_0 t))$$

$$S_{w_{ci}}(f) = S_{w_{si}}(f) = N_0/2\,|H_0(f)|^2 = S_{w_e}(f)$$

$$E[w_{ci}(t)w_{cj}(s)] = 0 \quad i \neq j$$

$w_{ci}(t)$ and $w_{sj}(s)$ are independent processes

$$S_{n_{cij}}(f) = S_{n_{sij}}(f) = NB^2/2\,|H_0(f)|^2 = S_{n_e}(f)$$

$n_{ci}(t)$ and $n_{sj}(s)$ are independent processes

$$E[w_{ci}^2(t)w_{ci}^2(s)] = (R_{w_e}(0))^2 + 2(R_{w_e}(\tau))^2$$

$$E[n_{ci}^2(t)n_{cj}^2(s)] = (R_{n_e}(0))^2 + 2(R_{n_e}(\tau))^2$$

$$E[w_{ci}^2(t)w_{cj}(s)] = 0$$

$$E[n_{ci}^2(t)n_{cj}(s)] = 0$$

Using the above, $z_1(t)$ can be shown to be

$$\begin{aligned} z_1(t) = \; & 4(\sqrt{Ps}(A/B))^2(1 - m_x\epsilon_x(t))^2 \; + \\ & (w_{c1}(t) + w_{c4}(t) + n_{c1}(t) + n_{c4}(t))^2 \; + \\ & (w_{s1}(t) + w_{s4}(t) + n_{s1}(t) + n_{s4}(t))^2 \; + \\ & 4\sqrt{Ps}(A/B)(1-m_x\epsilon_x(t))[(w_{c1}(t)+w_{c4}(t)+n_{c1}(t)+n_{c4}(t))\cos(\phi(t)+\theta) \; + \\ & \qquad (w_{s1}(t)+w_{s4}(t)+n_{s1}(t)+n_{s4}(t))\sin(\phi(t)+\theta)] \end{aligned} \tag{35}$$

The Taylor series expansion in equation 13 is valid only when the loop error is small. Therefore the following inequality should hold

$$1 >> m_x\epsilon_x \tag{36}$$

Under this assumption, the $\epsilon(t)$ term in the noise can be ignored and the $\epsilon^2(t)$ term in

the signal can be ignored.[17] Therefore equation 35 reduces to

$$z_1(t) = 4(\sqrt{Ps}\ A/B)^2\ (1 - 2m_x\epsilon_x(t)) + m_1(t) \tag{37}$$

where $m_1(t) = (w_{c1}(t) + w_{c4}(t) + n_{c1}(t) + n_{c4}(t))^2 +$

$(w_{s1}(t) + w_{s4}(t) + n_{s1}(t) + n_{s4}(t))^2 +$

$4\sqrt{Ps}(A/B)[(w_{c1}(t)+w_{c4}(t)+n_{c1}(t)+n_{c4}(t))\cos(\phi(t)+\theta) +$

$(w_{s1}(t)+w_{s4}(t)+n_{s1}(t)+n_{s4}(t))\sin(\phi(t)+\theta)]$

Similar expressions can be found for $z_2(t)$, $z_3(t)$, and $z_4(t)$.

The auto–correlation $m_1(t)$ and the cross–correlation of $m_1(t)$ with the other outputs can be found to be

$$R_{m_{11}}(\tau) = 16(R_{w_c}(\tau))^2 + 64(R_{n_c}(\tau))^2 + 64R_{w_c}(\tau)R_{n_c}(\tau) + \tag{38}$$

$$16(R_{w_c}(0))^2 + 64(R_{n_c}(0))^2 + 64R_{w_c}(0)R_{n_c}(0) +$$

$$16(\sqrt{Ps}\ A/B)^2(2(R_{w_c}(\tau) + 4R_{n_c}(\tau))\ e^{-L|\tau|/2}$$

$$R_{m_{12}}(\tau) = 64(R_{n_c}(\tau))^2 +$$

$$16(R_{w_c}(0))^2 + 64(R_{n_c}(0))^2 + 64R_{w_c}(0)R_{n_c}(0) +$$

$$16(\sqrt{Ps}\ A/B)^2(4R_{n_c}(\tau))\ e^{-L|\tau|/2}$$

$$R_{m_{13}}(\tau) = 4(R_{w_c}(\tau))^2 + 64(R_{n_c}(\tau))^2 + 32R_{w_c}(\tau)R_{n_c}(\tau) +$$

$$16(R_{w_c}(0)^2 + 64(R_{n_c}(0))^2 + 64R_{w_c}(0)R_{n_c}(0) +$$

$$16(\sqrt{Ps}\ A/B)^2(R_{w_c}(\tau) + 4R_{n_c}(\tau))\ e^{-L|\tau|/2}$$

$$R_{m_{14}}(\tau) = R_{m_{13}}(\tau)$$

Similar results for the auto–correlation and cross–correlation of the other outputs can be found.

---

[17] Actually the $\epsilon^2(t)$ term will be subtracted out by the estimator and the latter assumption is not needed.

From the results thus far, a linear observation equation can be found. It is

$$\text{O.E.} \quad Z(t) = 8(\sqrt{P_s}A/B)^2 f \begin{vmatrix} -m_x & 0 \\ m_x & 0 \\ 0 & -m_y \\ 0 & m_y \end{vmatrix} \begin{vmatrix} \theta_x \\ \theta_y \end{vmatrix} + M(t) + \text{D.C.} \tag{39}$$

where $$\text{D.C.} = (4(\sqrt{P_s}\,A/B)^2 + 4R_{w_o}(0) + 8R_{n_o}(0))\,1$$

$$R_M(\tau) = \begin{vmatrix} 4 & 0 & 1 & 1 \\ 0 & 4 & 1 & 1 \\ 1 & 1 & 4 & 0 \\ 1 & 1 & 0 & 4 \end{vmatrix} 4(R_{w_o}(\tau))^2 +$$

$$\begin{vmatrix} 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 1 & 1 & 0 & 2 \end{vmatrix} 16(\sqrt{P_s}A/B)^2(R_{w_o}(\tau)e^{-L|\tau|/2} + 32R_{w_o}(\tau)R_{n_o}(\tau)) +$$

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{vmatrix} 64(R_{n_o}^2(\tau) + R_{n_o}(\tau)e^{-L|\tau|/2})$$

$$S_{w_o}(f) = N_o/2\,|H_o(f)|^2$$

$$S_{n_o}(f) = NB^2/2\,|H_o(f)|^2$$

It can be shown that the estimator will subtract out the D.C. term in Z(t). More importantly, it can be shown that the estimator will subtract out the third term in $R_M(\tau)$. The reason for this is that the third term in $R_M(\tau)$ is composed entirely of excess–intensity noise and phase noise in the detector outputs. These noises are common to all the detectors and can be considered noise from the same sample function. Since the estimator will be taking the difference of terms in Z(t), it is no wonder that the common terms will be subtracted out.

The state equation for the angular process is given by equation 19.

Because the observation noise is not white, finding the optimum, linear, mean-square estimator and its associated estimation error requires some modifications to the normal Kalman filter algorithm. These modifications are straight forward [18]. Basically, they involve creating a matrix whitening filter and, in general, require the use of differentiators. However, it can be shown that under the present assumptions, [18] taking into account the bandlimited spectrum of the observation noise yields little or no better performance than just assuming the observation noise is white, equal to its value around zero frequency. [19] This will be assumed.

Before solving for the estimator structure it is necessary to assume some sort of filter for $H_0(f)$. For simplicity $H_0(f)$ will be chosen to be a single-pole filter with double-sided bandwidth B_0. Extension to other types of filters can easily be made. Therefore

$$H_0(f) = \pi B_0/(\pi B_0 + j2\pi f) \tag{40}$$

Approximating $R_M(\tau)$ as white noise with spectral height equal to the actual spectral height around zero frequency yields

---

[18] The most important assumption is that the bandwidth of $H_0(f)$ is large compared to the Lorentzian line-width.

[19] This is definitely the case when $H_0(f)$ is assumed to be a single pole filter. I believe it is also the case for other filters.

$$R_M(\tau) = \left[\begin{vmatrix} 4 & 0 & 1 & 1 \\ 0 & 4 & 1 & 1 \\ 1 & 1 & 4 & 0 \\ 1 & 1 & 0 & 4 \end{vmatrix} \pi B_0 (N_0/2)^2 + \right. \tag{41}$$

$$\left. \begin{vmatrix} 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 1 & 1 & 0 & 2 \end{vmatrix} (8PsN_0(B_0/(B_0+(L/2\pi)) + 2\pi B_0 N_0 NB^2) \right] \delta(\tau)$$

It is straight forward to solve for the steady–state solution of the Riccati equation, or equivalently, the steady–state, mean–square error. The answer is

$$\xi_{xx} = (\sigma / ((8Ps/N_0)^{1/2} f (A/B) m_x)) S_L \tag{42}$$

$$\xi_{yy} = (\sigma / ((8Ps/N_0)^{1/2} f (A/B) m_x)) S_L$$

$$\xi_{xy} = 0$$

where $$\xi_{xy} = E[(\theta_x(t) - \hat{\theta}_x(t))(\theta_y(t) - \hat{\theta}_y(t))]$$

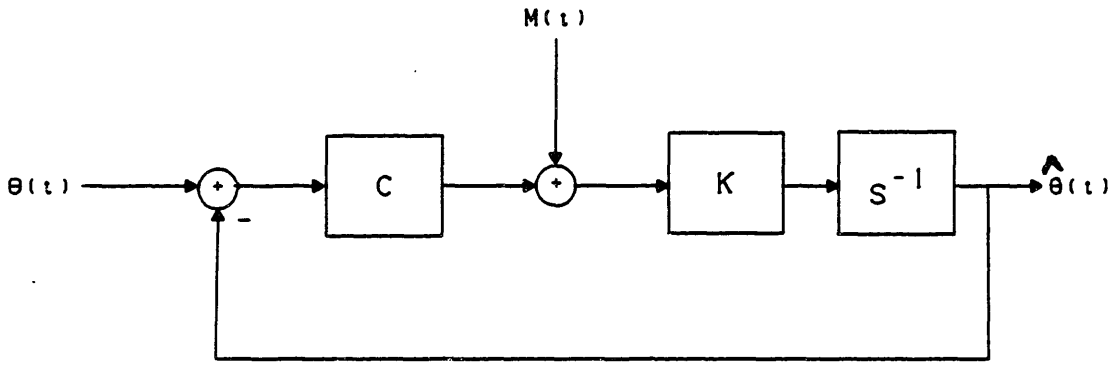
$$S_L = \sqrt{\frac{B_o}{B_o + \frac{L}{2\pi}}} \left[1 + \frac{\frac{\pi}{2}(B_o + \frac{L}{2\pi})}{8 \frac{Ps}{N_0}\left(\frac{A}{B}\right)^2} + \frac{2\pi(B_o + \frac{L}{2\pi})}{8 \frac{Ps}{NB}^2 \left(\frac{A}{B}\right)^2} \right]^{\frac{1}{2}}$$

The estimator structure (when the loop is operating in its linear region) is shown in Figure 10. C and K are given by

$$C = 8(\sqrt{Ps}\, A/B)^2 \begin{vmatrix} -m_x & 0 \\ m_x & 0 \\ 0 & -m_y \\ 0 & m_y \end{vmatrix} \tag{43}$$

$$K = k \begin{vmatrix} -1 & 1 & -0 & 0 \\ 0 & 0 & -1 & 1 \end{vmatrix}$$

$$k = (\sigma/(2N_0)) / ((8Ps/N_0)^{1/2} A/B\, S_L)$$



Linearized tracking loop.
Figure 10

Since equation 13 is valid only for $B >> L/2\pi$, it is easily seen that for high signal–to–noise ratio S_L approaches one and the estimation error approaches the lower bound in mean–square estimation error given is Section 3.1. Hence, the squaring–loop is asymptotically optimum. S_L represents the loss in performance over the ideal case where there is no phase noise, S_L will be termed the squaring loss.

A lower bound for equation 42 is provided by the Schwartz inequality. That is

$$((A_i/B_i)m_{x_i})^2 \leq \int_{\Omega_i} |\partial U_s(\sigma)/\partial x|^2 \, d\sigma \qquad (44)$$

$$(A_i/B_i)^2 \leq \int_{\Omega_i} |U_s(\sigma)|^2 \, ds$$

Ω_i = Area over one detector

For high signal–to–noise ratio (S_L=1) the optimum L.O. pattern for the cost function given in equation 25 is given by 26.

I have not solved for the optimum L.O. pattern for the more general case. However, one thing is clear: the L.O. pattern should be linearly dependent on the signal derivative in x the signal derivative in y, and the signal field itself. Therefore

$$U_l(\sigma) = \sum_{i=1}^{4} l_{xi}S_{xi}(\sigma) + l_{yi}S_{yi}(\sigma) + l_{si}S_{si}(\sigma) \tag{45}$$

where

$$S_{xi}(\sigma) = \partial U_s(\sigma)/\partial x \quad \text{over } \Omega_i$$

$$S_{yi}(\sigma) = \partial U_s(\sigma)/\partial y \quad \text{over } \Omega_i$$

$$S_{si}(\sigma) = U_s(\sigma) \quad \text{over } \Omega_i$$

l_{xi}, l_{yi}, and l_{si} represent weighting coefficient

Assuming the signal field over the detector array is an Airy–disc pattern and assuming infinite detector area, the quantities in equation 44 reduce to

$$(A/B)m \leq (\pi/4)(d/\lambda f) \tag{46}$$

$$(A/B) \leq 1/2$$

Assuming high signal–to–noise ratio the optimum L.O. field is given in equation 26. However, assuming infinite detector area and an Airy–disc signal field, it was shown in Section 3.1. that the radial derivative of the Airy–disc closely approximates the optimum L.O. field. Using this approximation

$$(A/B)m = (1/\sqrt{2})(d/\lambda f) \tag{47}$$

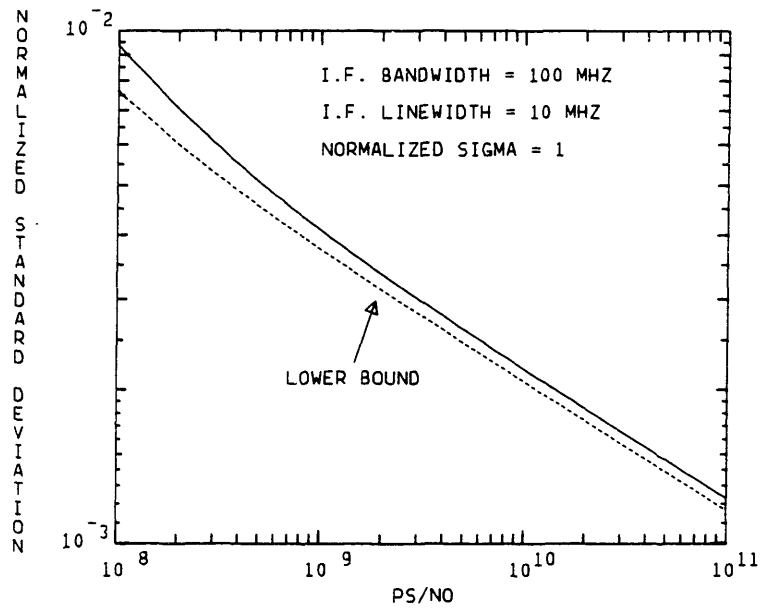
$$(A/B) = \sqrt{8}/3\pi$$

Let the normalized variance in the tracking error and the normalized sigma of the Wiener process be represented as

$$\xi_n = \xi/(\lambda/d)^2 \tag{48}$$

$$\sigma_n = \sigma/(\lambda/d)$$

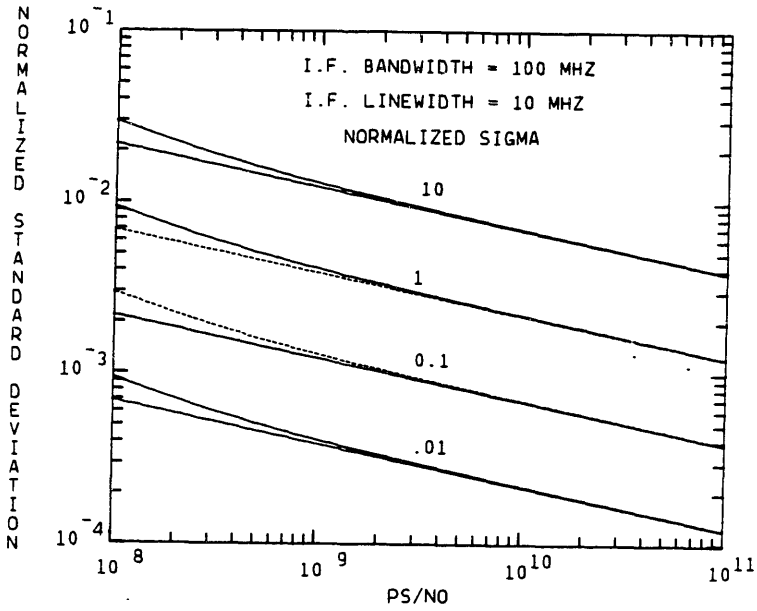
The normalized standard deviation is plotted versus Ps/N_0 for σ_n=1, B_0=100MHz, $L/2\pi$=10MHz in Figure 11. The Lower curve represents the lower bound provided by equation 46 the upper curve represents the optimum high signal–to–noise case provided by equation 47. This figure reveals that there is not much room for improvement by



Comparison of tracking performance with lower bound.
Figure 11

considering other types of L.O. patterns. Figure 12 shows the normalized standard deviation plotted versus Ps/N_0 for four values of σ_n. This plot assumes: B_0=100MHz, $L/2\pi$=10MHz, infinite detector area; the signal field over the detector array is an Airy–disc; the L.O. field is the radial derivative of the Airy–disc (equation 47); and no excess intensity noise. Also included is the absolute lower bound (for the case where there is no phase noise and the same assumptions as above) provided by Section 3.1.

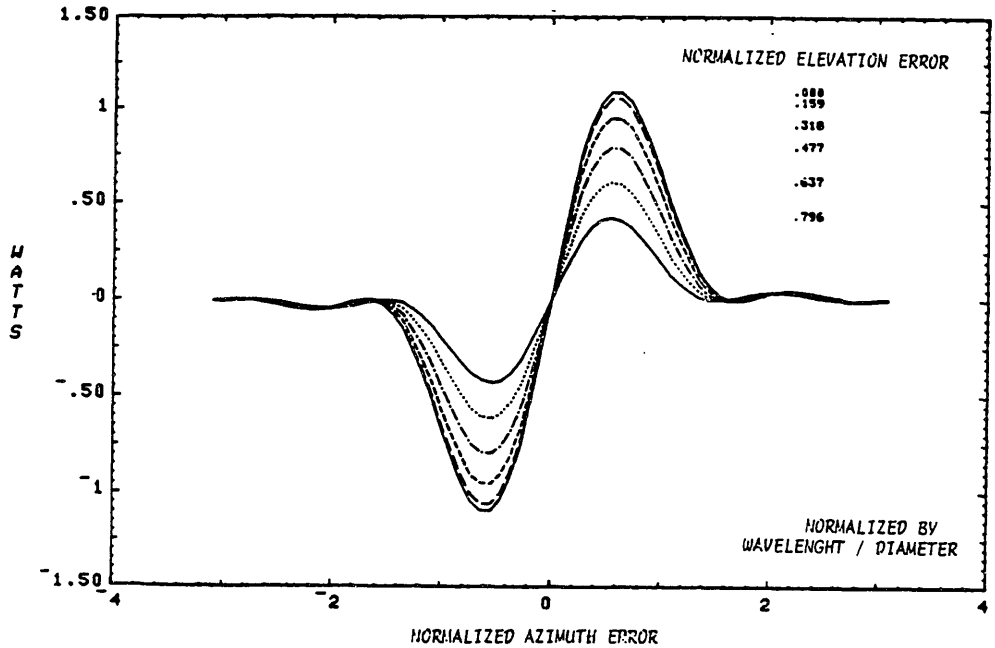
This figure shows the performance is near optimum over the range of signal powers of



Comparison of coherent demodulation with squaring loop.
Figure 12

interest.

This section has dealt with the issue of heterodyne spatial tracking using a squaring loop and assuming that the loop operates in its linear region and that the beam–steering device has infinite bandwidth and precision. If the error of the loop is very small compared to the linear range of the angle discriminator, the former assumption is valid. When the loop error approaches the linear dynamic range of the angle discriminator the probability of loss of lock increases and the analysis presented in this section must be modified since loop operation becomes non–linear. Figure 13 shows the angle–discriminator profile assuming an Airy–disc signal pattern and the radial derivative of the Airy–disc L.O. pattern. It can be inferred from this figure when the loop can be expected to operate in its linear region. A more rigorous analysis of loss of lock can be presented using Fokker–Plank techniques but will not be presented



Angle discriminator profile.
Figure 13

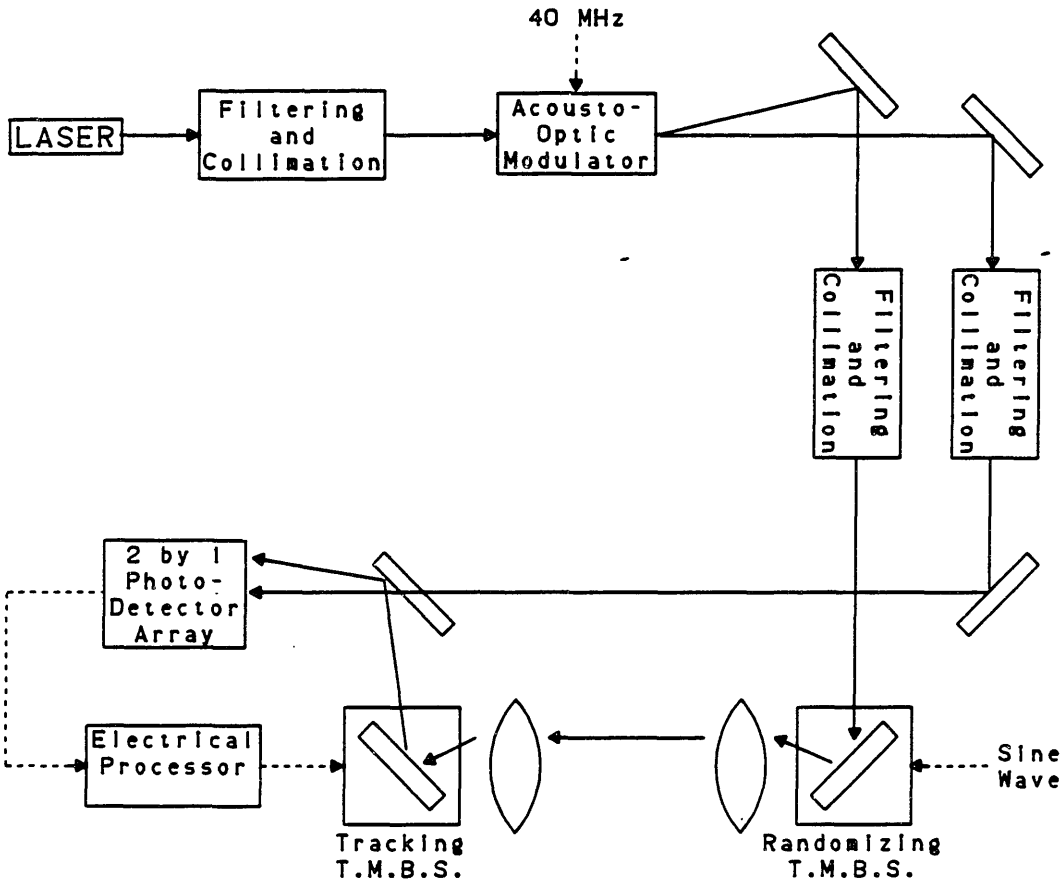
here. Beam–steerer bandwidth and precision considerations will be discussed later. However, as far as this section is concerned the bandwidth of the beam–steerer must be large enough so that it does not appreciably distort the performance of the loop integrators.

4. EXPERIMENTAL RESULTS

4.1. INTRODUCTION

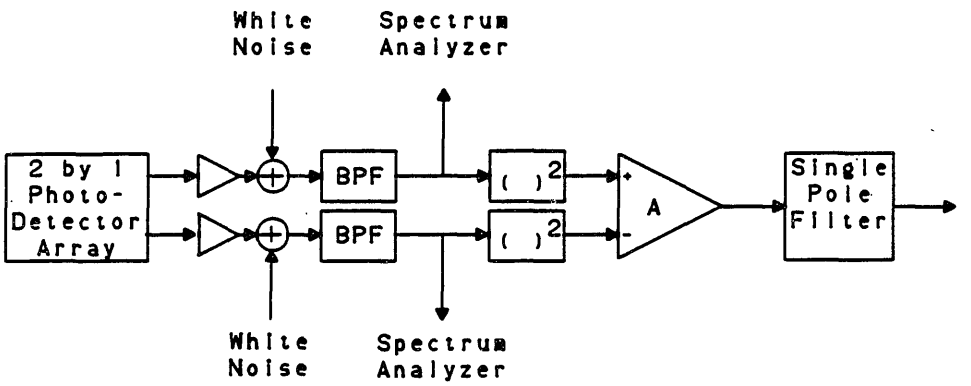
This chapter discusses the experimental results. The main intent of the experimental work was to display some of the fundamentals of heterodyne spatial tracking. Some of the results are partly qualitative. Qualitative not because there is not any meaningful quantitative results that are directly applicable to optical ISL's, but because a full experimental simulation of optimal, optical heterodyne spatial tracking is beyond the scope of a masters thesis. The main issues that were not dealt with formally are: creating a truly optimal L.O. pattern; and dealing with the actual phase noise spectrum. No attempt was made to create and track a Wiener process as discussed in Chapter 3. Instead, sinusoidal angular spectrums were used.

4.2. EXPERIMENTAL SETUP



Experimental setup.
Figure 14

Figures 14 and 15 show the experimental setup in block diagram form. The laser used was a Hitachi CSP-1400 GaAlAs semiconductor laser diode. The light from the laser was collected using a microscope objective and spatially filtered using a pin hole. The beam was then sent into an acousto-optic (AO) modulator. The output of this AO modulator contains two beams separated in frequency by the acoustic driving frequency which was 40 MHz. Due to the fact that the two beams were derived from the same input beam, they are very highly correlated and the normal Lorentzian lineshape will not be observed. This is one of the main differences mentioned earlier.



Electrical processor.
Figure 15

Since the subsequent electronics were not made to take advantage of this coherence, the resulting experimental data is meaningful and can be treated as in section 3.2. for the case where $B_o >>> L/2\pi$.

The two beams were again spatially filtered, collimated, and their diameters adjusted as follows:[20] The local oscillator beam diameter was chosed so as to yield near uniform illumination of the 2X1 photo–detector array; diffraction effects were observed on the signal field and the distance to the first null was chosen to be slightly less that the detector width.[21] The signal field was then sent onto a TMBS. This TMBS was used to randomize the incidence angle of the signal field entering the lens assembly. Due to the fact that the incoming optical field had a smaller diameter

---

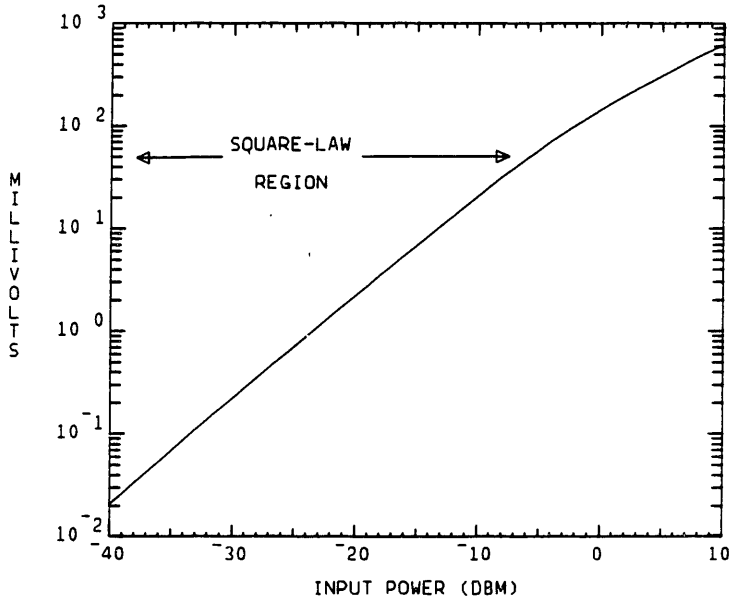
[20] Good optical quality was insured using an optical wedge.

[21] Some motivation for this choice is given in section 3.1.

than the effective aperture of the lens assembly, the focal lengths shown were chosen.[22] This lens setup, for obvious reasons, would not be valid in an actual system implementation. Basically this lens assembly redirects the incident beam onto the tracking mirror. The signal field and L.O. were then combined through a beam spliter and sent on to a 2X1 photo–detector array. The signals were then amplified. At this point a 10 dB coupler was inserted to allow for noise injection to set the SNR. The noise sources consisted of back biased avalanche diodes followed by variable attenuators. The signals were then bandpass filtered. At the output of the bandpass filter another 10 dB coupler was inserted. This coupler was used to monitor the i.f. spectrum. The signals were then square–law detected. The square–law detectors were full wave Schottky diodes followed by a low pass filter. A typical power to voltage transfer curve is shown in Figure 16. The outputs of these square–law detectors were then differentially amplified and low pass filtered with a single pole filter.

---

[22] No angular multiplication was used here. However, in an actual system, angular multiplication would be used to get out of the scanner noise.

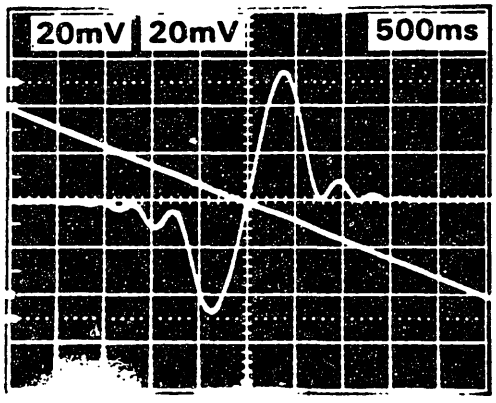


Square-law detector response.
Figure 16

4.3. INITIAL CHARACTERIZATION OF THE TRACKING LOOP.

It was important to begin characterization of the spatial tracking loop with little or no noise present. This was the first step in determining if the actual closed loop behavior could be predicted from the theoretical open loop behavior.

The loop can be characterized in terms of two parameters; the open loop gain and the open loop bandwidth. The open loop gain can be broken up into two different gains: a gain due to the individual component gains and a gain to the optical fields. The latter gain will be referred to as the angle discriminator constant Kd. Figure 17



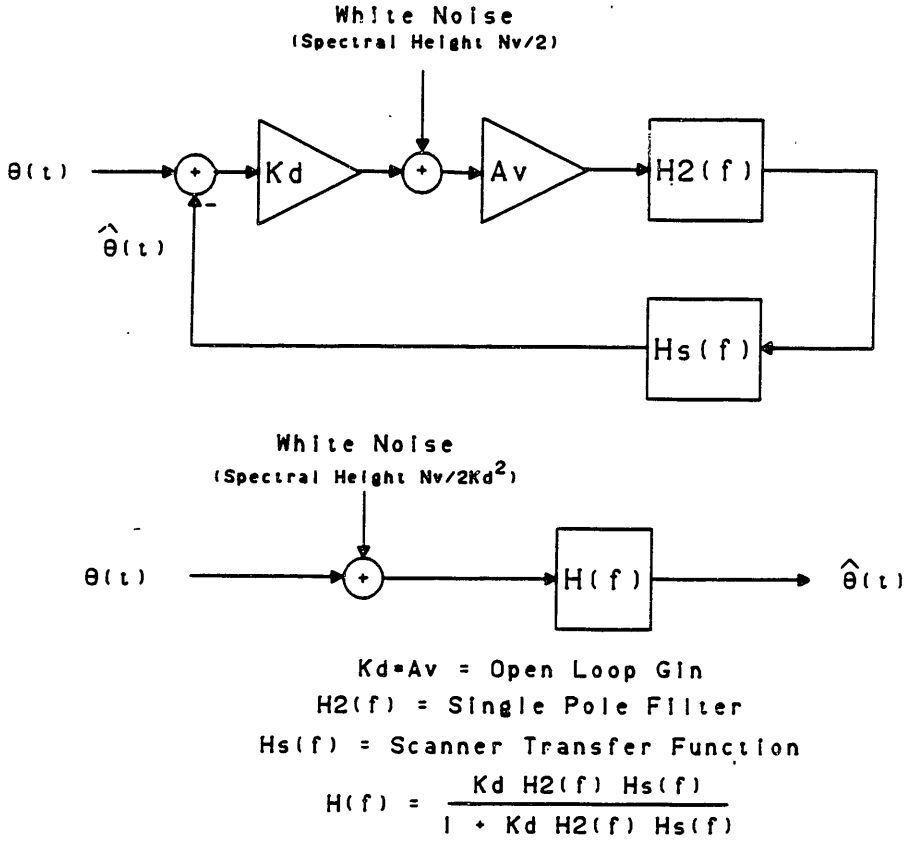
$$K_D = \frac{150\ \mu V}{\mu rad}$$

DYNAMIC RANGE = 500 μrad
(peak to peak)

Angle discriminator profile.
Figure 17

shows a typical plot of the angle discriminator profile. The dynamic range (1 dB deviation from linearity) is approximately 500 μradians and the gain is 150

μV/μradian.[23] For a plane wave signal field and a plane wave L.O. field the angle discriminator profile should be a linear term times a sinc function squared. The argument of the sinc function is the angular error divided by the diffraction limited field of view, λ/d, (this is easy to derive). The measured discriminator profile agrees well with the theoretical prediction. Attempt was made throughout the experimental work to keep approximately the same gain profile. The dynamic range of the discriminator will put a limit on how large an error the loop can tolerate before it enters a non-linear region and begins to break lock, (in the present case it is about $\pm$ 250 μradians or about half the diffraction limited field of view).



Linearized tracking loop.
Figure 18

A linearized model for the tracking loop is shown in Figure 18. $H_s(f)$ represents the scanner transfer function (Appendix 7.3.), $H_2(f)$ represents the low pass filter, and

---

[23] Actually this number includes component gains too and therefore is not truly representative of Kd.

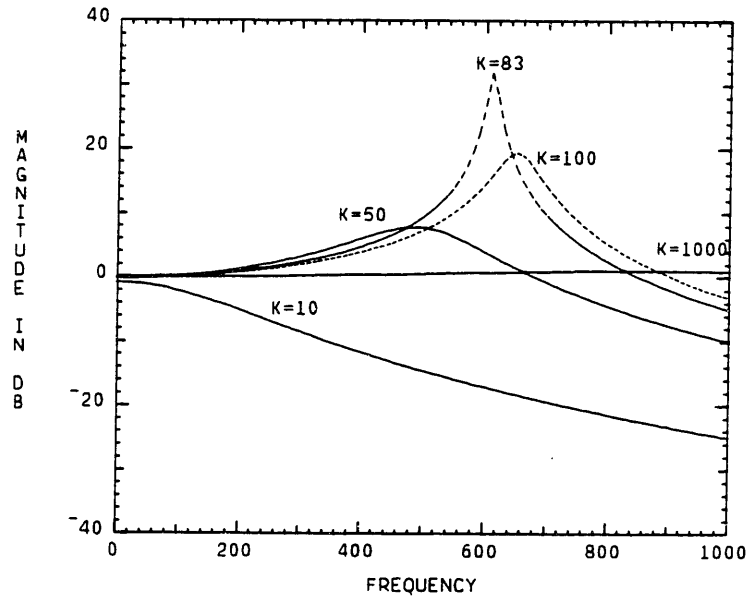
Av represents the voltage gain. H(f) is the closed loop transfer function. An expression for the noise density can be derived as in Section 3.2. and is given the the next section. It was decided to use sinusoidial angular deviations to test the loop behavior. An expression for the loop error (in the linear region) is easily derived using the techniques of Chapter 3. and is given by

$$\sigma^2 = |1 - H(fo)|^2 As + \frac{Nv}{(K_d)^2} \int_o^\infty |H(f)|^2 df \qquad (49)$$

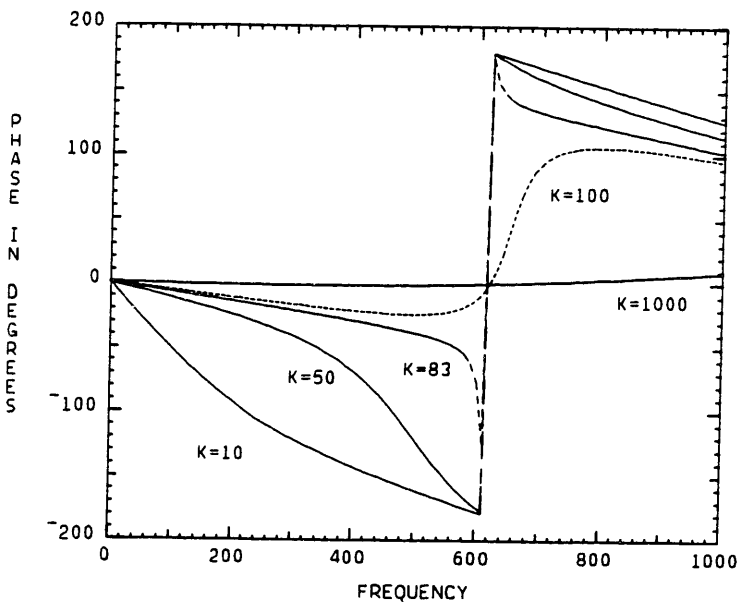
σ^2 = Error Variance
H(f) = Closed Loop Transfer Function
fo = Sinusodial frequency
As = Sinusoid Power
Nv = Noise Spectral Density
K_d = Angle Discriminator Gain

A typical plot of the closed loop transfer function is shown in Figures 19 and 20. Note that the loop is not useable (in this instance) for an open loop gain greater than about 75 due to the oscillation at approximately 600 Hz. The actual point of the instability will be dependent of the cutoff of $H_2(f)$. If the scanner were perfect, i.e. $H_s(f)=1$, then this instability would never occur and most of the loop behavior could be predicted in terms of the gain bandwidth product. However, because the scanner is imperfect and since the loop filter is restricted to a single pole filter, this instability will limit the maximum open loop gain that can be applied to the loop. As a side note, in order to achieve a large open loop gain it is necessary to use a very low cutoff for $H_2(f)$. This is because the scanner contributes a phase that is almost linear in frequency and $H_2(f)$ contributes a phase that reaches a maximum and a magnitude that rolls off linearly.

Due to the scanner noise (and non-linearity) it was necessary to ensure that the error in equation 49 was larger than the scanner noise in order to verify the loop operation. It was chosen to use an amplitude of 2.8 mradians (normalizes to 5.6-see

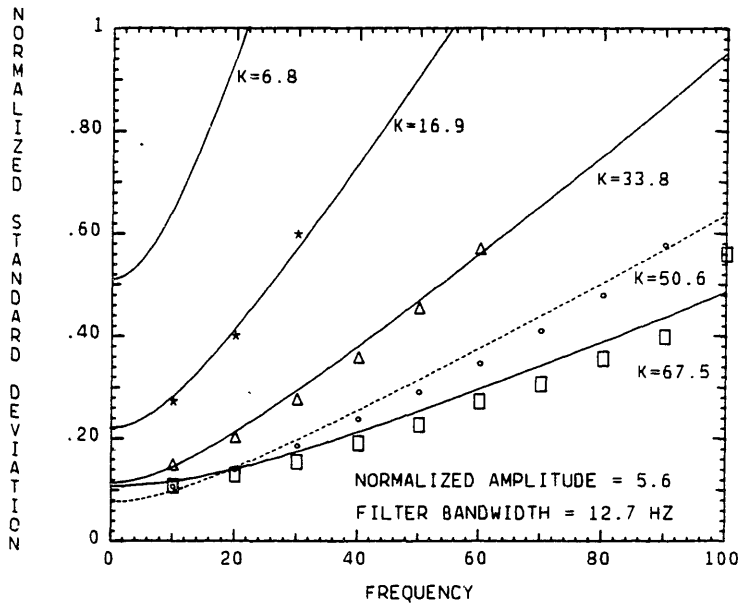


Magnitude of closed loop transfer function.
Figure 19



Phase of closed loop transfer function.
Figure 20

next paragraph) for the sinusoids throughout the experimental work. The error was measured by assuming that the position outputs on the scanner drive units were a true indication of the mirrors' position.



Spatial tracking error versus frequency.
Figure 21

Figure 21 show some typical results. The solid curves represent the theoretical prediction from equation 49 and the symbols represent the actual measured error. The theoretical and experimental tracking errors are normalized by the signal field wavelength divided by the signal field diameter (λ/d = 500μrad). The actual signal field diameter is not known very accuratly and was estimated to be 1.6 mm the uncertainty in the normalization is estimated to be about 10 percent.

Each curve represents a different open loop gain. The minimum gain was determined where the rms error was larger than the linear range of the angle discriminator. The maximum gain was determined by the onset of oscillations caused by the non–ideal scanner. The abrupt stop in the symbols is due to loss of lock. As mentioned in the begining of this section these are results for high signal–to–noise ratio. Although noise effects were included in the theoretical prediction they are dominated by the first term in equation 49. [24]

In summary, except for a possible normalization problem, it appears that the spatial tracking system is well modeled by the present theory for the frequency range of interest and high signal–to–noise.

---

[24] Actually in Figure 21 some of the effects due to noise can be seen where the K=50.6 curve crosses the K=67.5 curve.

4.4. CHARACTERIZATION OF THE LOOP WITH NOISE PRESENT

This section deals with characterization of the spatial tracking loop when noise is present. An important consideration for the spatial tracking system is the tracking error versus the number of photons–per–second (optical signal power) incident onto the spatial tracking photo–detector array. To find the actual optical signal power is not a simple task due to the fact that there are some unknown losses in the detectors and, more importantly, some components of the signal field never heterodyne with the L.O. and therefore will not appear to in the i.f. spectrum. The difficulty is related to the fact that in an optimum system, it is not desirable to match the signal field with the L.O. field. Hence a simple ratio of the i.f. signal power to noise spectral density gives only partial information in characterizing the incident optical signal power. However, to avoid having to deal with determining the optical field distributions exactly, this ratio will be used as an indicator of the signal power. As seen from section 3.1., the ratio of i.f. signal power to noise spectral density is actually [25]

$$(A/B)^2 \, P_s/N_0 \tag{50}$$

The only unknown quantity in equation 49 is the ratio $Nv/(Kd)^2$. Under the assumptions and methods of Chapter 3 this quantity can be shown to be

$$\left(\frac{Nv}{K_d}\right)^2 = \frac{1}{2\left(\frac{Ps}{N_0}\right)\left(\frac{Am}{B}\right)^2}\left[1 + \frac{B.W.}{2\left(\frac{Ps}{N_0}\right)\left(\frac{A}{B}\right)^2}\right] \tag{51}$$

B.W. = Effective I.F. Bandwidth

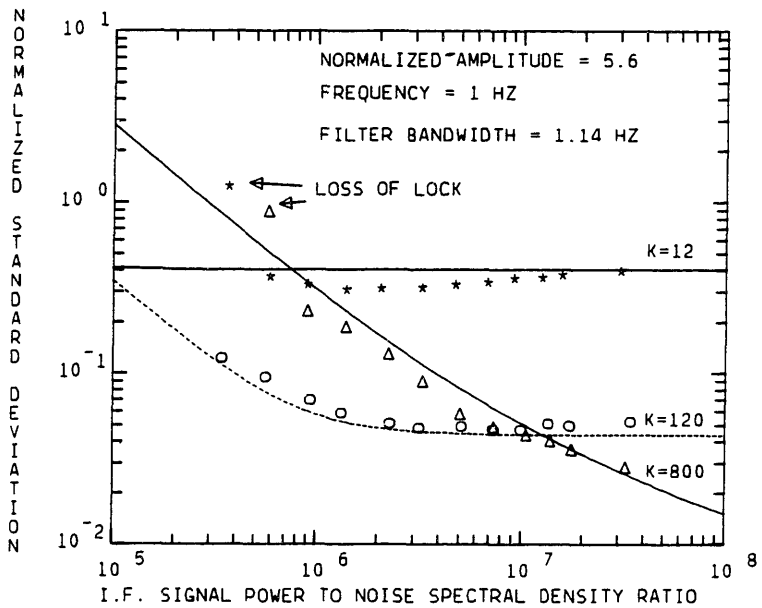
The first term is actually the effective number of photons–per–second (tracking is in

---

[25] For the experimental setup A/B is approximately .3.

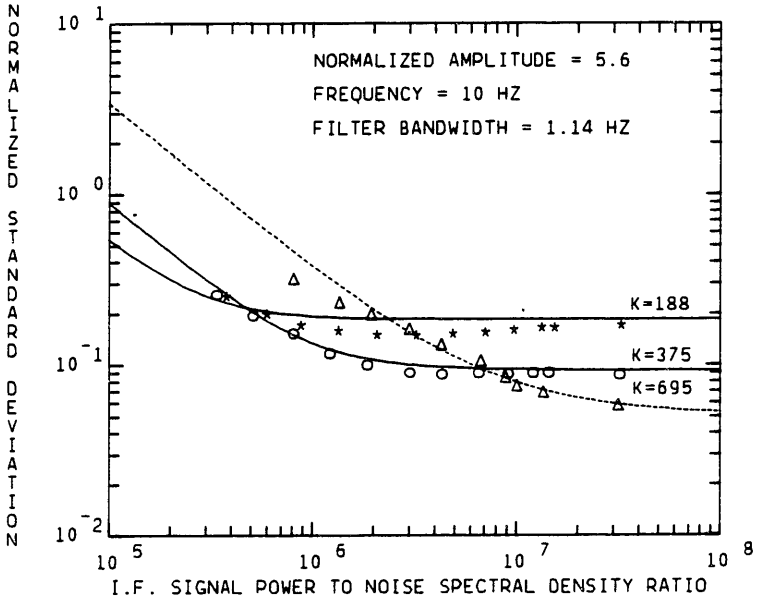
one dimension instead of two which accounts for the discrepancy in the factor of two) and the second term in the brackets is the squaring loss as defined in section 3.2. (here there is no excess noise and $B_0 >>> L/2\pi$).

Figures 22, 23, and 24 show some typical results of the experimental work. The

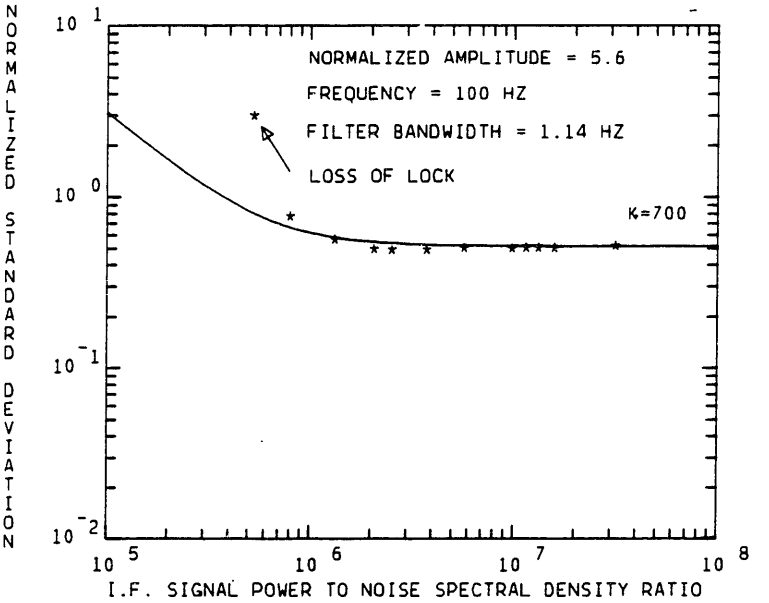


Spatial tracking error vs the received number of photons–per–second (1 Hz).
Figure 22

results are for 1 Hz, 10 Hz, and 100 Hz sinusoidal angular spectrum respectively. The open loop bandwidth was chosen as 1.14 Hz. The solid lines represent the theoretical predictions and the symbols represent the experimental data points. This data was taken by fixing the open loop gain and bandwidth to the desired value and then injecting noise to change the (apparent) number of received photons–per–second. This value was measured from the i.f. ports. As shown in Figure 15, the i.f. bandwidth was 4.4 MHz and thus, the plots range from a signal–to–noise ratio of approximately –10 dB to 10 dB. The agreement is well within the experimental uncertainty. All errors are within about 1 dB. For the 1 Hz sinusoid and the open loop gain of 800, the tracking



Spatial tracking error vs the received number of photons-per-second (10 hz).
Figure 23



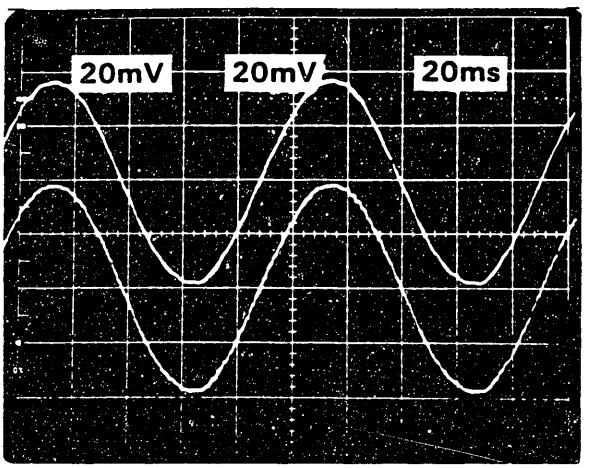
Spatial tracking error vs the received number of photons-per-second (100 hz).
Figure 24

accuracy for large signal power approached the scanner resolution limit. The discrepancy between the theoretical and experimental results are mainly due to the following: errors between the theoretical and actual scanner transfer function; measurement errors in the parameters, especially the square–law detector constant; and that the square–law detectors are not perfect squaring devices.

Figure 25 shows the input sine wave (top trace) and the spatial tracking loop's estimate (bottom trace) for a high signal–to–noise (8 dB) and low signal–to–noise (–8 dB) respectively.

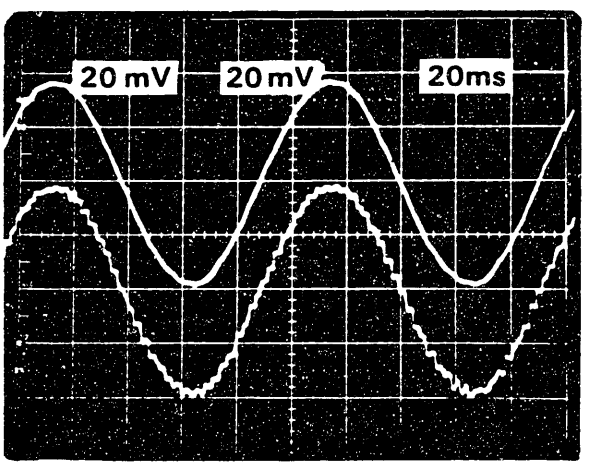
The tracking error at 100 Hz was almost to the limit of breaking lock. As seen in Figure 24, the deterministic error was much larger than the noise induced error. However, there was another error term that was also dominated by the deterministic error. This was error due to nonlinearity in the loop. This distortion is due to the fact that the loop error was approaching the dynamic range of the angle discriminator (some of the nonlinearity could also be due to saturation of the torque transducer constant). This error can be seen in Figure 26 for high and low SNR.

In summary, the effects of noise have been experimentally observed and theoretically predicted. Nonlinear effects have also been observed. However, they have not been worked directly into the theoretical model.



BW = 1.14 Hz
K = 584
F = 10 Hz
AMP = 2.8 mrad

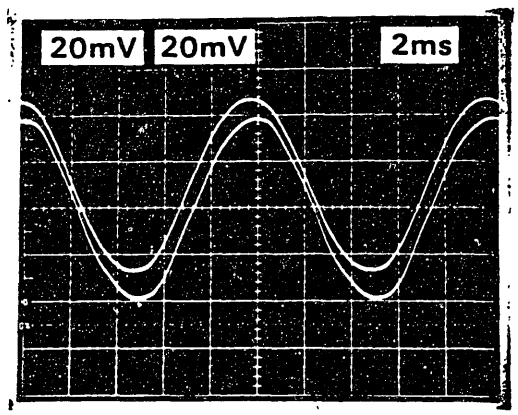
SNR = 8 dB



BW = 1.14 Hz
K = 584
F = 10 Hz
AMP = 2.8 mrad

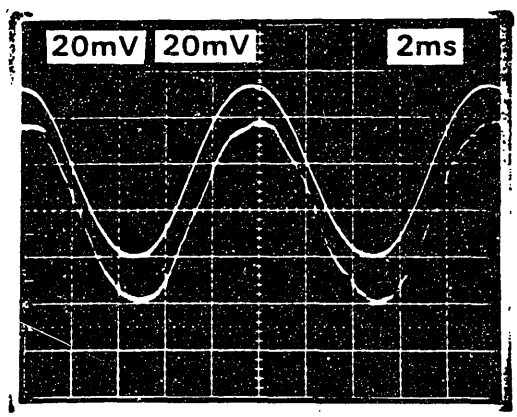
SNR = -8 dB

Comparison of input angle and estimated angle.
Figure 25



BW = 1.14 Hz
K = 200
F = 100 Hz
AMP = 2.8 mrad

SNR = 8 dB



BW = 1.14 Hz
K = 200
F = 100 Hz
AMP = 2.8 mrad

SNR = -8 dB

Observed nonlinearity in the loop.
Figure 26

5. CONCLUSION

A linearized model for the residual spatial tracking system using heterodyne techniques has been developed and analyzed. Theoretical performance estimates and bounds were developed based on the assumption that the angular process was a Wiener process and as stated previously the method used can be extended to any angular process. These results indicated that a quadrant detector is essentially the optimum detector and that the radial derivative of the Airy–disc is close to the optimum L.O. distribution (less than .5 dB potential loss); however, little loss is incurred by using a plane wave distribution (less than 1 dB). The important issue of phase noise in the observation equation was dealt with and it was concluded that a squaring loop could successfully combat the phase noise for particular parameters chosen. Although the theoretical results assumed the loop operated in its linear region, the biggest assumption was that the beamsteering devices had infinite precision and infinite bandwidth. The experimental work revealed that this assumption is by no means valid. This, however, does not violate the above theoretical results, but rather states that elaborate compensation networks will be needed in the final design and that the ultimate precision of the spatial tracking system may be effected by the noise in the beam steering devices.

The experimental work showed good agreement between measurement and theoretical prediction. The torque motor beam steerers (TMBS) used revealed their limitations some expected and some not. Among those not expected was non–linearity in the torque transducer constant. It should be noted that there are TMBS made that are superior to the ones used in the experiment.

And finally, no direct comparison between direct detection and heterodyne

detection spatial tracking was made. However, it is easily shown that at best with no background radiation or thermal noise direct detection is 3dB better than heterodyne detection. It is expected that background radiation and thermal noise is well in excess of this 3 dB margin and hence heterodyne detection is superior.

6. FUTURE WORK

The room for future work is extensive. To start with a through investigation into the different types of beam steererings devices in needed. Only after a head to head comparison (including in some cases experimental verification for the specs) can the appropriate beam steering device be chosen. It is expected that for a low orbit to high orbit ISL the beam steering devices chosen will be different. In particular the low orbit satellite will need large dynamic range and it is expected that this can only be accomplished with a TMBS[26] The high orbit satellite has comparatively small dynamic range and perhaps a piezo–electric beamsteering device would provide increased bandwidth and resolution.

Another important area to be pursued is characterization of the angular process to be tracked. This is a difficult this task due to the fact that it is dependent on so many things. In order to characterize the angular process the open loop tracking system (coarse tracker) must first be specified. Also the needs of the acquisition system will influence the spatial tracking system design. Most importantly, the particular satellite stabilization system, orbit, and beam steering system influence the angular process tremendously.[27] Obviously exact characterization of the angular process is impossible and approximations must be made. The first step towards an bound might be a ramp whose slew rate is larger than the maximum of the sum of: the angular rate of change of the satellite to be tracked, the coarse spatial tracking system maximum angular rate of change, and the maximum angular rate of change of the satellite attitude during

---

[26] Probably similar to the one used in the experimental work except with a tighter spring constant for larger bandwidth and flexure–pivots for repeatability.

[27] It is for these reasons the characterization of an angular process was not attempted in this thesis.

firing of the stabilization or maneuvering rockets. The next step might be to have a ramp as above plus a Wiener process where the variance parameter of the Wiener process will characterize the orbital perturbations due to various gravitational non-uniformities.

This thesis has assumed space to space communication. This is a very simple channel compared to earth to space, or even plane to plane communication where the earths atmosphere complicates the spatial tracking issue. It is believed that many of the methods in this thesis are applicable to this problem. For instance the optimum L.O. pattern can be found in a similarly except that the field incident on the entrance aperture will not necessarily be a plane wave. To account for this a Karhunen-Loeve (or similar) expansion can be used and it is expected that the L.O. pattern should be chosen as the information part of the eigenfunction with the largest eigenvalue. The angular process can be bounded as above except there will be another additive term due to the atmosphere. This additive term may be very complicated and it is necessary to accurately model it so as to take its dynamics into account in the Kalman filter. [27]

In a sense this section has speculated about what future work should be done. No doubt many new and more important matters will appear.

---

[27] In fact it is possible for more than one spot to appear in the focal place at a given time.

7. APPENDIX

7.1. EXPANSION OF COVARIANCE EQUATION

When the Taylor series expansion of the signal field (over the detector array), equation 13, is substituted into equation 12 the resulting terms can be grouped into five categories as discussed in Section 2.5. This section lists those terms and briefly comments on them.

Term 1

$$\int_{A_i}\int_{A_j} |U_{l_o}(t,\sigma)|^2 \, |U_{l_o}(s,\rho)|^2 \, d\sigma \, d\rho \tag{52}$$

Term 2

$$\int_{A_i}\int_{A_j} U_{l_o}(\sigma)\, U_{s_o}^{*}(\sigma)\, U_{l_o}^{*}(\rho)\, U_{s_o}(\rho)\, e^{-j(\omega(t-s)+\phi(t)-\phi(s))}\, d\sigma\, d\rho + (c.c) \tag{53}$$

Term 3

$$\int_{A_i}\int_{A_j} U_{l_o}(\sigma)\, (\epsilon_x(t)\, \partial U_{s_o}^{*}(\rho)/\partial x + \epsilon_y(t)\, \partial U_{s_o}^{*}(\rho)/\partial y)\, U_{l_o}^{*}(\rho)\, U_{s_o}(\sigma)\, e^{-j(\omega(t-s)+\phi(t)-\phi(s))}\, d\sigma\, d\rho + \int_{A_i}\int_{A_j} U_{l_o}(\sigma)\, U_{s_o}^{*}(\rho)\, U_{l_o}^{*}(\rho)\, (\epsilon_x(t)\, \partial U_{s_o}(\sigma)/\partial x + \epsilon_y(t)\, \partial U_{s_o}(\sigma)/\partial y)\, e^{-j(\omega(t-s)+\phi(t)-\phi(s))}\, d\sigma\, d\rho + (c.c) \tag{54}$$

Term 4

$$[1 - (1+L(t))(1+L^{*}(s))(1+S(t))(1+S^{*}(s))] \text{ X Term 2} \tag{55}$$

$$\text{Term 5} \qquad [1 - (1+L(t))(1+L^{*}(s))(1+S(t))(1+S^{*}(s))] \text{ X Term 3 } + O\epsilon^{2}(t) \qquad (56)$$

It can be shown that for the application of interest term 2 dominates term 4 and term 3 dominates term 5. Therefore, it would appear that the approximations in section 2.4 are valid. However, eventually if the spatial tracking error becomes so small, term 4 may mask out term 3. It is projected that this won't happen untill a rms error of much less that .2 beamwidths is reached. It should also be noted that for coherent demodulation this is not a problem since currents from adjacent quadrants are compared – hence, the noise is canceled. For noncoherent square–law detection currents are also compared but the cancellation is not total due to cross terms.

7.2. AN OPTIMAL L.O. PATTERN

This section gives a motivation for the optimum L.O. pattern for a quadrant detector array as stated in Section 3.1. equation 26. The assumption of focal plane processing is essential. A motivation for the optimum L.O. pattern for tracking in one dimension can be similarly found; this pattern is optimum for any angular process. A proof of this is straight forward and is analogous to the standard proof of matched filters given in communication texts. The optimum L.O. pattern for tracking in two dimensions is somewhat dependent on the particular angular process. However, for processes with equal variances in azimuth and elevation the optimum L.O. pattern for the cost function given in equation 25 can be shown to be more general. [29]

The most important observation begins with the realization that the optimum L.O. pattern must be as given in equation 27. This is rather obvious. If the L.O. has components orthogonal to the signal field derivative in x and the derivative in y over any detector element, then clearly it is suboptimum. This is because, this orthogonal component will not heterodyne with the information part of the signal field but will contribute additional shot noise; hence, performance can be improved by getting rid of it. Thus all that remains is to specify the four weighting coefficients.

By symmetry the four l_{xij} (one in each quadrant) should be the same and similarly for the l_{yij} weightings. Therefore, only two coefficients need to be specified. Let these coefficients be called lx and ly. Furthermore, let

---

[29] If the variances are not equal then the additional requirement of a cost function with normalizes the weightings is needed.

$$Px = \int_{\Omega} (S_x(\sigma))^2 \, d\sigma \tag{57}$$

$$Py = \int_{\Omega} (S_y(\sigma))^2 \, d\sigma$$

$$Pxy = \int_{\Omega} (S_x(\sigma)\, S_y(\sigma)) \, d\sigma$$

Ω = the area over one quadrant of the detector

It is straight forward to show that for the cost function given in equation 25 reduces to

$$C_T = \left[\frac{(LxPx + Ly(Pxy/Py))^{\frac{1}{2}}}{Lx^2 + Ly^2} \right]^{-1} + \left[\frac{(Lx(Pxy/Px) + LyPy)^{\frac{1}{2}}}{Lx^2 + Ly^2} \right]^{-1} \tag{58}$$

By symmetry Px equals Py. Taking the partial derivative of C_t with respect to l_x and l_y, setting the resulting two equations to zero, and simultaneously solving for a minimum yields the desired result; $l_x = l_y$.

7.3. THE OPTICAL SCANNERS

7.3.1. Introduction

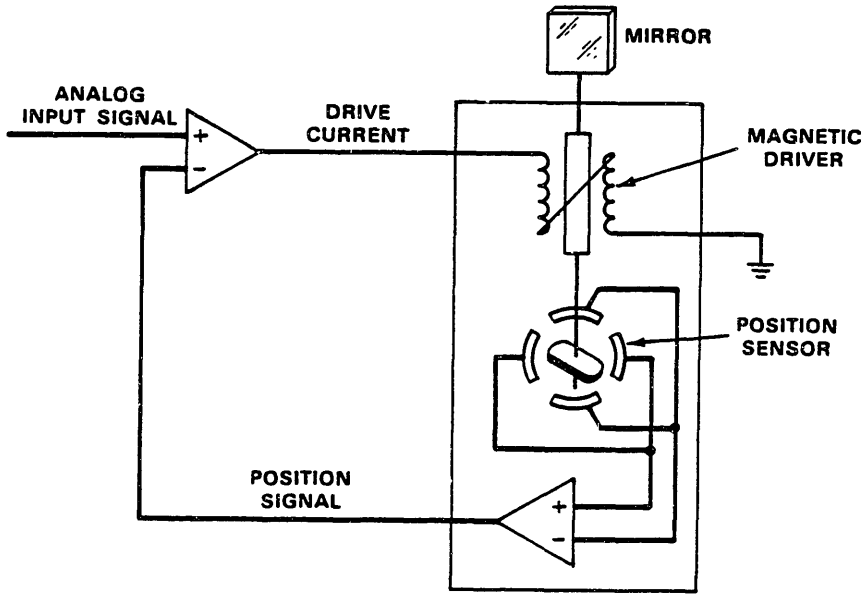
This section describes the beamsteering devices used in the experimental work. The optical beam steering devices used were galvanometer optical scanners made by General Scanning Inc. These devices are commonly called Torque–Motor–Beam–Steerers (TMBS). In particular the G–100PDC optical scanner and its corresponding CX–651C drive unit were used. The G–100PDC optical scanners were chosen because it was desired to have fast, accurate scanning and wide dynamic range for the beam–steering devices. General Scanning indicated that these devices should be able to: follow square waves up to 300 Hz; be able to resolve 5 μrad; and have a scanning range of 40 degrees. [30]

The G–100PDC is a closed loop device that used a base–band current to create a magnetic field that applies a torque to a rotor to which a mirror is attached. A counter torque is (mainly) provided by a spring mechanism. The torque produced by the drive current rotates the mirror until it is balanced by the restoring torque of the spring. The drive current is derived from the difference between an (external) input voltage and a signal from a position sensor on the G–100PDC. The position sensor voltage is created from a "stable" 2 MHz oscillator that drives a capacitive–position transducer. This transducer consists of 4 electrodes (one in each quadrant) surrounding the electrically grounded rotor. Diametrically opposed electrodes are connected. The resulting two currents are rectified and put into a differential amplifier. The result is a voltage whose magnitude should be proportional to the

---

[30] This is dynamic range is much larger than necessary.

angle of the mirror. A highly simplified block diagram of the above is shown in Figure 27.



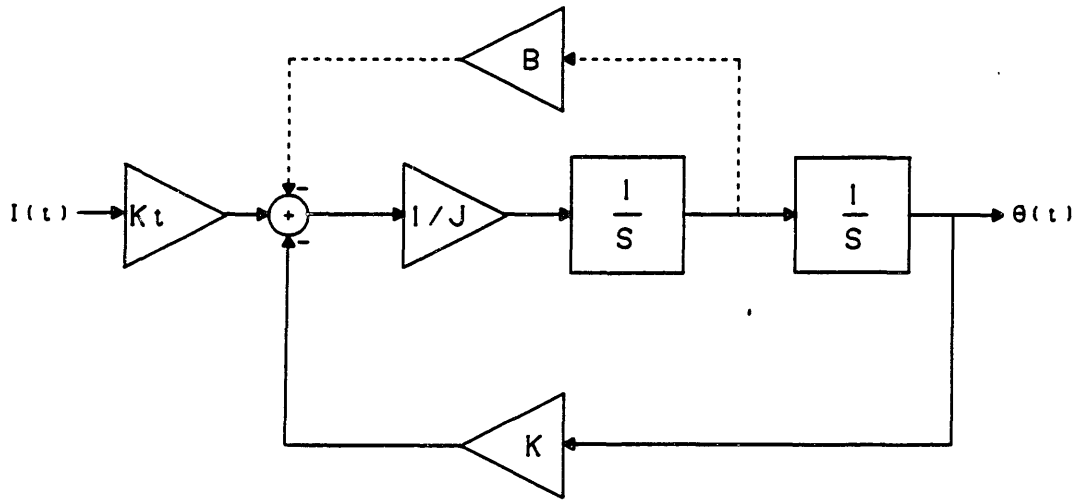
Galvanometer optical scanner.
Figure 27

7.3.2. The Beam–Steering Transfer Function

A transfer function for the drive current to angular deflection can be derived from the previous section and the model in Figure 28. Due to the permanent magnets in the scanner when a drive current is applied, the rotor experiences a torque. The magnitude of this torque is given by the torque transducer constant, Kt. From Newton's laws, this torque (if not balanced by a counter torque) will cause angular acceleration of the rotor. The magnitude of this angular acceleration is inversly proportional to the moment of inertia of the combined rotor–mirror assembly, J. The integral of the angular acceleration is, of course, the angular velocity. [31] The integral of the angular

---

[31] There is a damping term due to the angular velocity. However, this will be



Kt=Torque Transducer Constant
J=Moment of Inertia
K=Spring Constant
B=Damping

Drive current to angle transfer function $H_{\theta I}(s)$.
Figure 28

velocity is the position of the rotor. Due to the spring, any angular deviation will produce a restoring torque. The magnitude of this restoring torque is given by the spring constant, K. This completes the the loop.

The transfer function can now be written down directly and is

$$H_{\theta I}(s) = (Kt/K) \,/\, ((s/\omega_n)^2 - 1) \qquad (59)$$

where $$\omega_n = (K/J)^{.5}$$

ω_n represents the natural frequency and is equal to the spring constant divided by the moment of inertia. This quantity puts a limit on the useful frequency range of the

---

insignificant compared to the artificial damping from the scanner drive unit and therefore will be ignored.

TMBS.[32] For a given spring constant, the bigger the mirror the smaller the useful frequency range.[33]

Similarly a voltage to angle transfer function can be found. This transfer function is almost the current to angle transfer function cascaded with a series L–R circuit. This transfer function is

$$H_{\theta V}(s) = (Kt/KR) \,/\, ((\tau_\omega/\omega_n^2)s^3 + (1/\omega_n^2)s^2 + (\tau_\omega + (Kt^2/KR)s + 1)) \qquad (60)$$

where

$$\tau_\omega = L/R$$

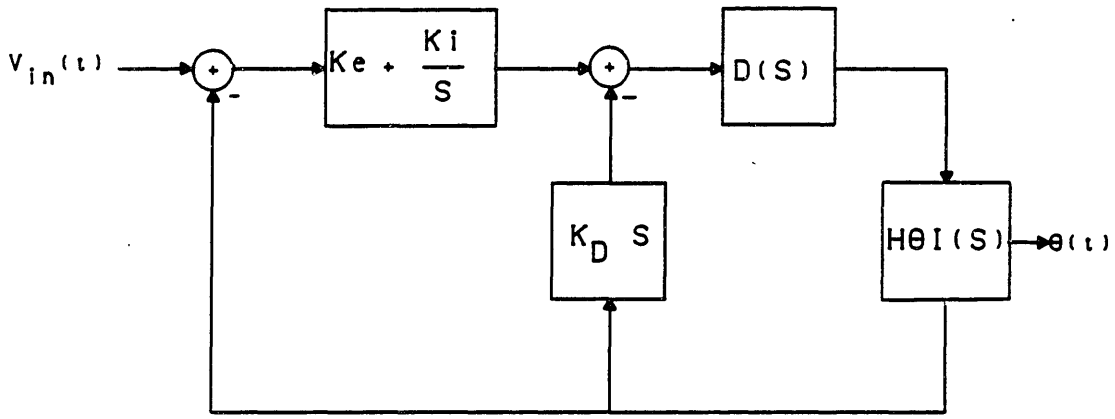
$$\omega_n = (K/J)^{.5}$$

The CCX–651C drive unit used integral plus proportional feedback along with damping (derived from the position sensor output) to control the mirror deflection. A block diagram of the complete system is shown in Figure 29. A closed loop transfer function from the above diagram can easily be found. Once the constants are identified (or measured) this transfer function can be plotted. Figures 30 and 31 show the theoretical transfer function and the measured transfer function. The general behavior is in quite good agreement. The discrepancies are due to the fact that many of the constants in were not measured but were either estimated or taken from the General Scanning brochure. Since many of these constants in the brochure are only listed to one significant figure and the transfer function is very sensitive to these same constants it is not surprising that there is some discrepancy.

---

[32] Even in a close loop configuration the scanner can not be driven significantly beyond this quantity.

[33] For the G–100PDC scanners and the mirrors used (7mmX7mm) the resonant frequency is about 350 Hz.



Ke=Error Gain
K_D=Damping Gain
Ki=Integrator Gain
D(S)=Current Dividing
HθI(S)=Current to Angle Transfer Function

Closed loop scanner transfer function.
Figure 29

7.3.3. Accuracy

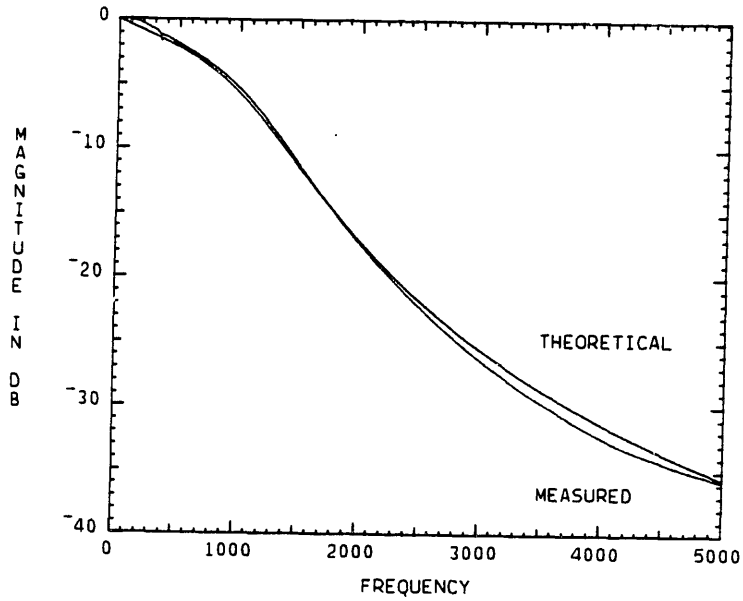
The accuracy the G–100PDC scanners is very difficult thing to quantify. There are many interrelated factors that effect the "accuracy". This section will try to deal with some of them.[34]

7.3.3.1. Linearity

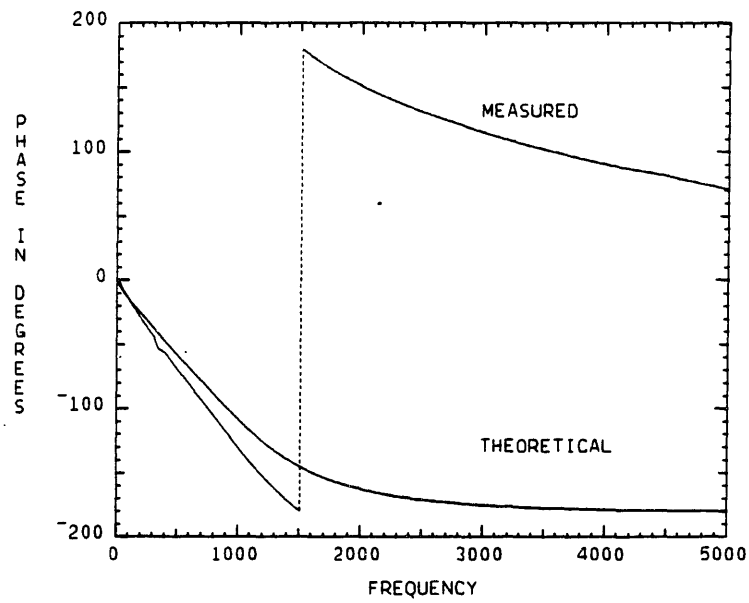
Linearity in the torque transducer constant and the position sensor are the primary concerns. The linearity of the position sensor is very good (.3% peak–to–peak).

---

[34] The definitions to be used are chosen to be consistent to those used by General Scanning.



Magnitude of scanner transfer function.
Figure 30



Phase of scanner transfer function.
Figure 31

However, when dealing with μradian accuracy it gets to the point where ticks and scratches from the machinery creating the scanner set the limit to position sensor linearity. The more important consideration is the torque transducer constant linearity. General Scanning does not give any accuracy bounds on this quantity. However, Figure 32 and 33 show a plot of the closed-loop transfer function as a function of drive signal amplitude. These figures were generated by sweeping a sinusoid of constant amplitude over the frequency range of interest and assuming that drive unit's output was a true indication of the mirrors position. [35] 1 V corresponds to 4 degrees of optical scanning. The result is that the scanner transfer function is nonlinear for very small excursions. If the rms angular deviation that the scanners are tracking is greater than approximately 2mrad, then the scanner will operate linearly. If not, the non-linearity will have to be dealt with. General Scanning claimed they had not come across this problem before. I believe the cause of this is due to breakout friction (or possibly a poor choice of spring constant) and could be alleviated by not using the normal suspension (ball bearings) for the rotor. Instead, the use of flexure pivots could be used. Flexure pivots have the property of having the highest possible scan repeatability. However, this is at the expense of linearity. Hence a more complicated drive unit is needed.

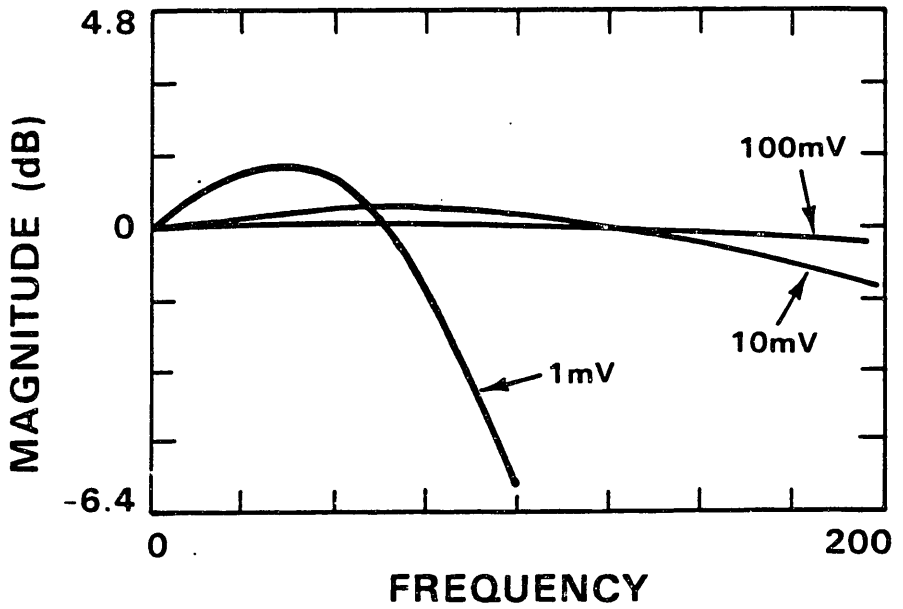
There is another type of nonlinearity in the torque transducer constant. There is a maximum level of torque that can be applied to the scanners. [36] This level is due to a combination of magnetic saturation and a structural limit of the scanner itself. [37]

---

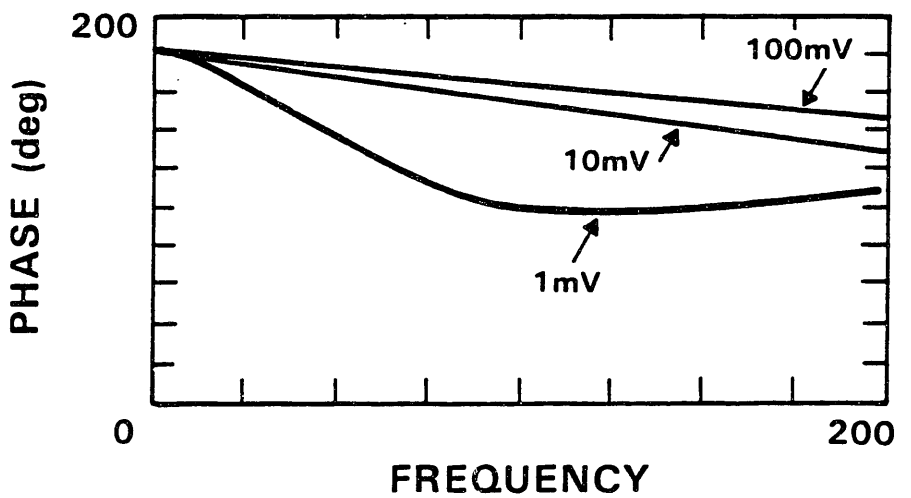
[35] This plot does not indicate harmonic content since the output was filtered in a narrow bandwidth about the input frequency.

[36] This level is on the order of .012 Nm for the G-100PDC scanners.

[37] There is protection circuitry in the drive unit to prevent over torque.



Non–linearity in the optical scanners (magnitude).
Figure 32



Non–linearity in the optical scanners (phase).
Figure 33

Assuming perfect scanners the maximum required torque to produce a particular scanning signal is easy to find from section 7.3.2. For sinusoids and triangular waves the quantities are listed below

$$T = (2\pi)^2 J \theta_m f^2 \qquad (61)$$

$$T = 8 J \theta_m (f/t)$$

where f = Frequency of the scan signal.

θ_m = Amplitude of the scan signal (in beam rotation).

t = Time taken for triangular wave to switch from positive ramp to negative ramp.

In a closed loop configuration for large rates of change saturation of the torque transducer can be a limiting factor.

One more related topic is that of deformation of the mirror. It can be shown that the mirror (or any surface) will deform when it is subject to angular acceleration. For the present application the deformation will be proportional to the applied torque. To minimize this deformation requires using very rigid mirrors, but in some cases this may be inconsistent with the requirements of large bandwidth. According to General Scanning, the deformation should be a small fraction of a wavelength and is therefore inconsequential.

7.3.3.2. Jitter and Wobble

Jitter is defined as the standard deviation from average scan angle, measured in the direction of the scan. General Scanning quotes this number from 5 to 10 μradians. However, the drive unit for the TMBS has 60 Hz harmonic leakage. It is estimated that this leakage is large enough such that the estimated Jitter is actually 10 to 30 μradians. This number was arrived at by observing the position output spectrum and

also by observing the error in the spatial tracking loop for very high signal–to–noise ratio and very slow scan signals.

Wobble is defined as the standard deviation from average scan angle, measured in the direction perpendicular to the scan. Indications by General Scanning are that the Wobble should be less than one half of the Jitter. [38]

---

[38] The estimated error in the Jitter actually includes some of the effects of the Wobble.

8. REFERENCES

1 V.W.S. Chan, "Coherent optical space communications system architecture and technology issues", SPIE Vol. 295 Control and communication Technology in Laser Systems, 1981.

2 V.W.S Chan, L.L. Jeromin, and J.E. Kaufmann, "Heterodyne Lasercom Systems Using GaAs Lasers for ISL Applications", internal memo, M.I.T. Lincoln Laboratory.

3 T. Kimura and Y. Yamamoto, "Progress of coherent optical fiber communication systems", Optical and Quantum Electronics, Vol. 15, 1983.

4 J.E. Kaufmann, "Phase and Frequency Tracking Considerations for Heterodyne Optical Communications", Proceedings International Telecommunications Conference, San Diego, CA, September, 1982.

5 V.W.S. Chan, S.D. Lowney, and J.E. Kaufmann, Frequency Tracking for Heterodyne Optical Communications Using Semiconductor Lasers, CLEO '83 Proceedings, Baltimore, MD, 1983.

6 J.E. Kaufmann, "Heterodyne Lasercom Link Budget", Internal Memo, M.I.T. Lincoln Laboratory, February 17, 1984.

7 R.M. Gagliardi and S. Karp, "Optical Communications", John Wiley, 1976.

8 P. Van Hove and V.W.S Chan, "Spatial Acquisition Algorithms and Systems for Optical ISL", internal memo, M.I.T. Lincoln Laboratory.

9 Chen-Wu Chen, "Optical Tracking Systems", Doctor of Science Thesis, Department of Systems Science and Mathematics, Washington University, December, 1982.

10 A. Papoulis, "Probability, Random Variables, and Stochastic Processes", McGraw-Hill, 1965.

11 V.W.S. Chan, "Statistical Model for Heterodyne Detection with Partially Coherent Lasers-part I and part II", internal memo, M.I.T. Lincoln Laboratory, July 1982.

12 T.K. Yee, G.L. Abbas, and V.W.S. Chan, "Study of Intensity Noise Statistics and Power Spectra of GaAlAs Laser Diodes at Various Output Powers and Frequencies", internal memo, M.I.T. Lincoln Laboratory (submitted to IEEE J–Q.E. July 1983).

13 G.L.. Abbas, V.W.S. Chan, and T.K. Yee, "Local–oscillator excess–noise suppression for homodyne and heterodyne detection", Optical Letters, Vol. 8, No. 8, August 1983.

14 F.G. Walther and J.E. Kaufmann, "Characterization of GaAlAs Laser Diode Frequency Noise", internal memo, M.I.T. Lincoln Laboratory.

15 F.G. Walther, S.D. Lowney, and J.E. Kaufmann, "Frequency tracking for Heterodyne Optical Communications", internal memo, M.I.T. Lincoln Laboratory.

16 J.E. Kaufmann, "Relationship of Laser Lineshape to its Frequency Noise Spectrum", internal memo, M.I.T. Lincoln Laboratory, July 1982.

17 H.L. Van Trees, "Detection, Estimation, and Modulation Theory – Part I", John Wiley, 1968.

18 A.E. Bryson, JR. and D.E. Johansen, "Linear Filtering for Time–Varying Systems Using Measurements Containing Colored Noise", IEEE Transactions on Automatic Control, AC–10, No. 1 (January 1965)