

**SPATIAL TRACKING USING AN ELECTRO-OPTIC
NUTATOR AND A SINGLE-MODE OPTICAL FIBER**

by
Todd Eric Knibbe

Submitted to the
Department of Electrical Engineering and Computer Science
in Partial Fulfillment of the Requirements
for the Degrees of

BACHELOR OF SCIENCE

and

MASTER OF SCIENCE

and

ELECTRICAL ENGINEER

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 1993

© Todd E. Knibbe 1993. All rights reserved.

The author hereby grants to MIT permission to reproduce and to
distribute copies of this thesis document in whole or in part.

Signature of
Author _____

Department of Electrical Engineering and Computer Science
December 1, 1992

Certified
by _____

Professor James K. Roberge
Thesis Supervisor

Certified
by _____

Eric A. Swanson
MIT Lincoln Laboratory

Accepted
by _____

Professor Campbell L. Searle
Chairman, Department Committee on Graduate Students

ARCHIVES

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

SPATIAL TRACKING USING AN ELECTRO-OPTIC NUTATOR AND A SINGLE-MODE OPTICAL FIBER

by

Todd Eric Knibbe

Submitted to the Department of Electrical Engineering
and Computer Science in Partial Fulfillment of the Requirements
for the Degrees of Bachelor of Science, Master of Science,
and Electrical Engineer

ABSTRACT

The potential advantages of using single-mode optical fibers within free-space laser communication systems are well known. Of particular importance within an optical heterodyne receiver is the spatial tracking subsystem. The use of fiber optics in these systems can alleviate some of the problems associated with the conventional bulk optics approach. The main drawback is the need to find a method of coupling the received signal field into a single-mode optical fiber in the presence of environmental disturbances. Because optical beamwidths are small (μrad) and typical disturbances can be large (mrad), this requirement makes the design of the spatial tracking subsystem an important and difficult part of a free-space optical communication system. Previous systems have utilized mechanical techniques for nutating the tip of the receiving fiber to derive tracking information. A new technique, based on using a fixed fiber and an electro-optic beam deflector, is proposed. Design considerations and experimental results are presented.

We believe that a heterodyne receiver architecture using an electro-optic nutator (EON) and a single-mode optical fiber can offer several improvements over existing mechanical nutation techniques, at the expense of a small additional coupling loss (~ 0.5 dB). The EON is small, lightweight, rugged, and has no moving parts. It can operate over a broad range of frequencies and can be driven by a simple op-amp, requiring low power. We have demonstrated its performance in a laboratory heterodyne receiver set-up. The measured open-loop noise performance was found to be in good agreement with predictions. Our active noise cancellation scheme also extended the useful automatic gain control (AGC) range of the receiver down into the low receiver i.f. signal-to-noise (SNR) region by greater than 15 dB over an approach which does not use the method. Lastly, we were able to demonstrate the EON's use in a two-axis, 1 kHz open-loop crossover frequency, spatial tracking feedback loop. Two-axis tracking was maintained over a receiver SNR range of 30 dB. Overall, our nutating receiver offered good performance over its entire expected operating range, with the use of fiber optics providing additional modularity, scalability, and compatibility with the large existing fiber technology base.

Thesis Supervisor: Prof. James K. Roberge
Title: Professor, Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science

TABLE of CONTENTS

1. Introduction.....	6
1.1 Laser Communications Technology Overview	
1.2 Advantages of Fibers in Free-Space Lasercom	
1.3 Nutating Receiver	
1.4 Thesis Organization Overview	
2. Nutation Concept.....	14
2.1 Mechanical Nutation	
2.2 Electro-Optic Nutation	
2.3 Analysis of the Electro-Optic Nutator	
2.4 Alignment Prescription for the Electro-Optic Nutator	
2.5 Performance of the Electro-Optic Nutator	
2.6 Deflection Sensitivity of the Electro-Optic Nutator	
2.7 The Driving Electronics for the Electro-Optic Nutator	
3. Noise Equivalent Angle.....	50
3.1 Derivation of the Receiver NESD	
3.2 NESD Measurement Procedure	
3.3 NESD Measurements	
4. Automatic Gain Control.....	74
4.1 Probing the Tracking Loop at Phase Crossover	
4.2 AGC by Noise Power Cancellation	
4.3 Implementation of AGC by Noise Power Cancellation	
4.4 AGC Measured Performance	
5. Spatial Tracking.....	100
5.1 EON Spatial Tracking Synchronous Demodulator	
5.2 Open Loop Spatial Tracking Measurements	
5.3 Feedback Compensation	
5.4 Two-Axis Spatial Tracking	
6. Conclusions.....	117
References.....	118
Appendices.....	121
Acknowledgments.....	135

LIST of FIGURES and TABLES

1. Introduction

Figure 1.1.1: Lasercom technology program overview.-----	8
Figure 1.1.2: Optical lasercom pointing requirement.-----	8
Figure 1.2.1: Strawman optical prescription.-----	10
Figure 1.3.1: Nested servo tracking loop design.-----	11
Figure 1.4.1: EON receiver block diagram.-----	13

2. Nutation Concept

Figure 2.0.1: Nutating receiver fiber coupling profile.-----	15
Figure 2.1.1: The mechanical fiber nutator.-----	17
Figure 2.1.2: Mechanical fiber nutator frequency response.-----	18
Figure 2.1.3: Mechanical nutation receiver.-----	19
Figure 2.2.1: Deflection of an optical beam by an optical wedge.-----	20
Figure 2.2.2: A medium whose index of refraction varies...-----	21
Figure 2.3.1: Basic alignment prescription for a static LiTaO ₃ ...-----	31
Figure 2.4.1: A contour plot of the quadrupole potential.-----	33
Figure 2.4.2: The 'best' electrode configuration on finite sized...-----	34
Figure 2.4.3: Front face view of LiTaO ₃ deflector.-----	35
Figure 2.4.4: Boundary conditions imposed by our LiTaO ₃ -----	36
Figure 2.4.5: Inside the electro-optic nutator.-----	39
Figure 2.5.1: The electro-optic nutator.-----	41
Table 2.5.1: Electro-Optic Nutator Measured Performance-----	41
Figure 2.6.1: Measuring the EON deflection sensitivity.-----	44
Figure 2.7.1: The EON driving electronics.-----	46

3. Noise Equivalent Angle

Figure 3.0.1: A simple model for a spatial tracking loop.-----	50
Figure 3.1.1: A typical intersatellite link.-----	52
Figure 3.1.2: Use of the Fraunhofer far-field diffraction equation...-----	54
Figure 3.1.3: The front-end of our receiver using the EON.-----	56
Figure 3.1.4: Two dimensional decomposition of the tilt...-----	57
Figure 3.2.1: Optical bench layout schematic.-----	64
Figure 3.2.2: Frequency shifted beam intensity versus...-----	66
Figure 3.3.1: Mode-mismatch loss and discriminator gain.-----	71
Figure 3.3.2: Nutating receiver Noise-Equivalent-Angle.-----	73

4. Automatic Gain Control

Figure 4.0.1: Nutating Receiver Block Diagram.-----	75
Figure 4.1.1: AGC by probing at phase crossover.-----	77
Figure 4.2.1: Automatic gain control using a signal power...-----	81
Figure 4.2.2: Model for an optical communication receiver front-end.--	81

Figure 4.2.3: Heterodyne signal spectrum.-----	88
Figure 4.3.1: Schematic for the nutating receiver AGC block.-----	93
Figure 4.4.1: AGC measured performance.-----	99
Figure 4.4.2: Change in the AGC circuit to remove the low...-----	99

5. Spatial Tracking

Figure 5.0.1: Nutating receiver block diagram.-----	101
Figure 5.1.1: Nutating receiver synchronous demodulator block...-----	102
Figure 5.1.2: Nutating receiver demodulator schematic.-----	103
Figure 5.2.1: General scanning FSM.-----	108
Figure 5.2.2: Azimuth spatial tracking open-loop response.-----	108
Figure 5.3.1: Compensator schematic.-----	111
Figure 5.3.2: Compensator block diagram.-----	113
Figure 5.4.1: A simple model for a spatial tracking loop (neg...-----	114
Figure 5.4.2: Azimuth closed-loop rejection.-----	115

6. Conclusions

References

Appendices

Figure A1.1: Nutating receiver block diagram.-----	122
Figure A1.2: Expanded nutating receiver block diagram.-----	122
Figure A1.3: Optical bench layout schematic.-----	123
Figure A1.4: Schematic for the nutating receiver AGC block.-----	124
Figure A1.5: Nutating receiver demodulator schematic.-----	126
Figure A1.6: Oscillator schematic.-----	128
Figure A1.7: Compensator schematic.-----	129
Figure A1.8: Current driver schematic.-----	130

Acknowledgments

1. Introduction

The potential advantages of using single-mode optical fibers within free-space laser communication systems has been recognized. Of particular importance within an optical heterodyne receiver is the spatial tracking subsystem. The use of fiber optics in these systems can alleviate some of the problems associated with the conventional bulk optics approach. The main drawback is the need to find a method of coupling the received signal field into a single-mode optical fiber in the presence of environmental disturbances. Because optical beamwidths are small (μrad) and typical disturbances can be large (mrad), this requirement makes the design of the spatial tracking subsystem an important and difficult part of a free-space optical communication system. Previous systems have utilized mechanical techniques for nutating the tip of the receiving fiber to derive tracking information. A new technique, based on using a fixed fiber and an electro-optic beam deflector, is proposed. Design considerations and experimental results are presented.

1.1. Laser Communications Technology Overview

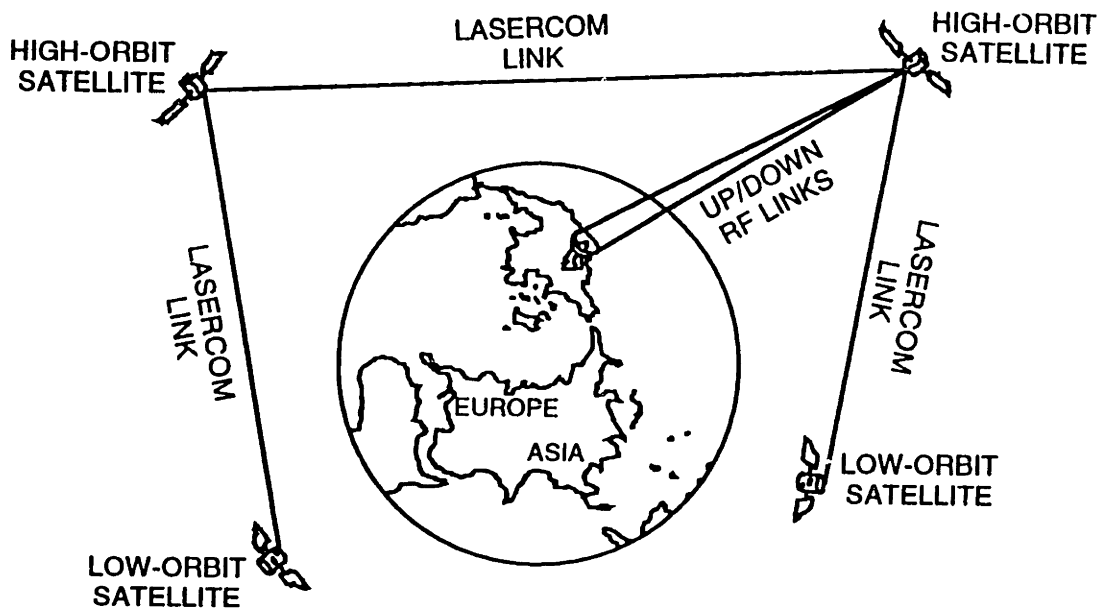
High capacity intersatellite links will allow for more efficient operation of future satellite systems. Fig. 1.1.1 shows a cartoon overview of a laser communications crosslink¹. By establishing a communications network, via intersatellite links, between communication, data collection, and surveillance

¹ This figure is a black and white print of a color viewgraph which has been used in many technical presentations by the Optical Communication Technology group at Lincoln Laboratory.

satellites, the need for an expensive world-wide ground tracking network is averted. Both microwave (i.e. 60 GHz) and optical (i.e. $3.6 \cdot 10^{14}$ Hz @ 830 nm) intersatellite links are candidates for such a system. However, at data rates above approximately 100 Mbps, microwave frequencies require large antenna apertures and transmitting powers that are difficult to integrate on existing host spacecraft [1]. Optical links use narrow transmit beams to produce a high gain communication system, yielding a comparatively modest size and weight package.

This advantage does not come without a price, however, as Fig. 1.1.2 shows². Because of the high antenna gain involved, optical beamwidths are very narrow ($\sim\mu\text{rad}$). As shown, a typical optical communication link from a geosynchronous (GEO) satellite to a low earth orbit (LEO) satellite has only a 150m footprint where the signal power can be received. Also because of these narrow beamwidths, a point ahead angle must be implemented to intercept the sending satellite with a return communication link. Even a well designed optical platform can have angular disturbances which are many beamwidths ($\sim\text{mrad}$). Typically, spatial tracking must be maintained to within a fraction of a beamwidth, making the design of the spatial tracking subsystem for a free-space optical communication link a challenging task. A number of papers have been published on the design of these systems [1-21].

² *ibid.*



- HIGH SPEED (1 Mbps - 5 Gbps)
- MODEST SIZE, WEIGHT AND POWER PACKAGE

Figure 1.1.1: Lasercom technology program overview.

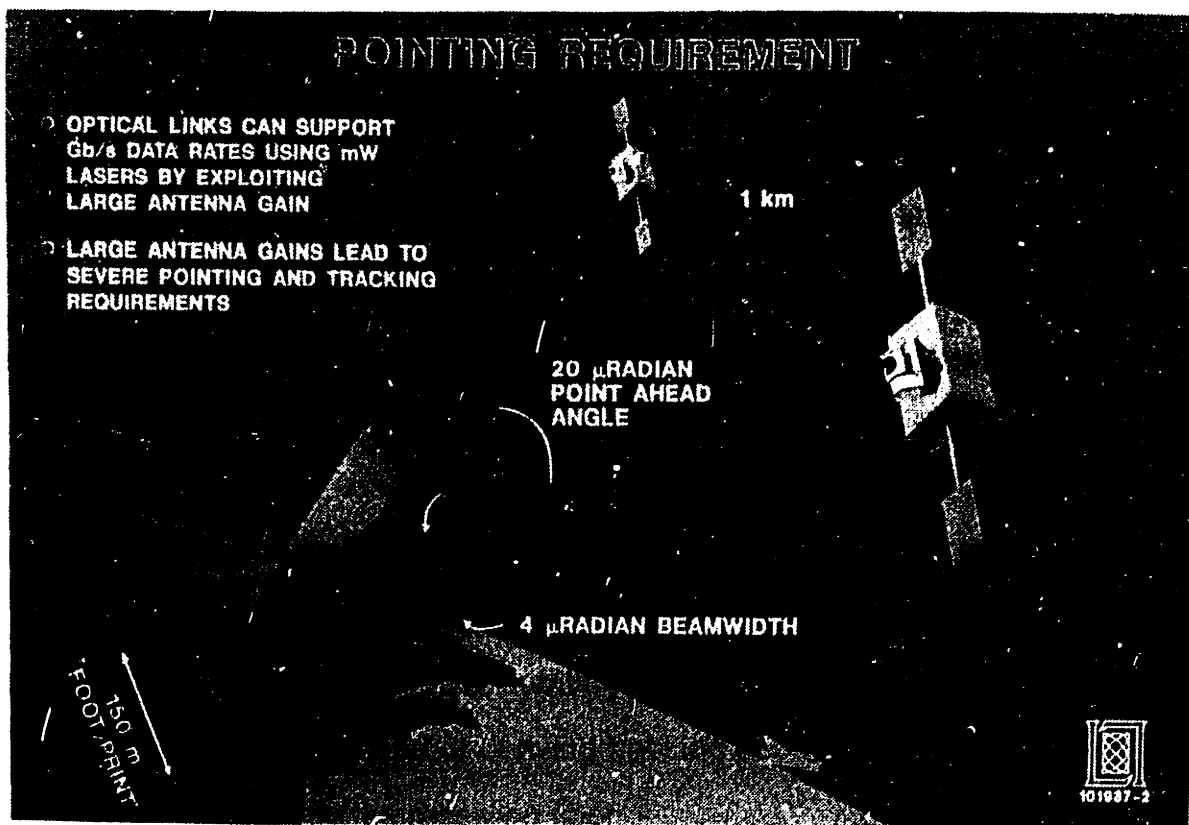


Figure 1.1.2: Optical lasercom pointing requirement.

1.2. Advantages of Fibers in Free-Space Lasercom

One of the main drawbacks of current, free-space optical communication designs is their extensive use of bulk optics to couple the incoming signal field from the entrance antenna aperture to the optical detector [1,4-8,11,14-18]. Because these designs require strict control over the relative position of component optical elements, they are all very sensitive to thermal and mechanical fluctuations of the optical platform, as well as to vibrational disturbances of the host spacecraft. Fig. 1.2.1 is a strawman optical prescription showing how the use of fiber optics in these systems can alleviate some of the associated problems. By developing a means of coupling the incoming optical power into a single-mode optical fiber after it has been collected by the receiving antenna, many optical components can be remotely located from the front-end optical bench. This adds increased flexibility to the mechanical design, which can lead to further reductions in the size, weight, and stability requirements of the optical communication module as a whole. As shown in the figure, the system is modular, compatible with the existing fiber technology base, and allows for easy upgrades of subsystems within the communications package as future technology becomes available.

The main disadvantage of a fiber-based system is the difficulty in effectively coupling the received signal field into the receiving single-mode fiber (spatial tracking), and the difficulty coupling from the fiber to free-space (spatial pointing). Also present are difficulties associated with coupling the transmitter laser into a fiber with sufficient isolation. The nutating optical receiver, which is the focus of this thesis, provides a means of deriving tracking information while coupling the received signal field into the receiving fiber.

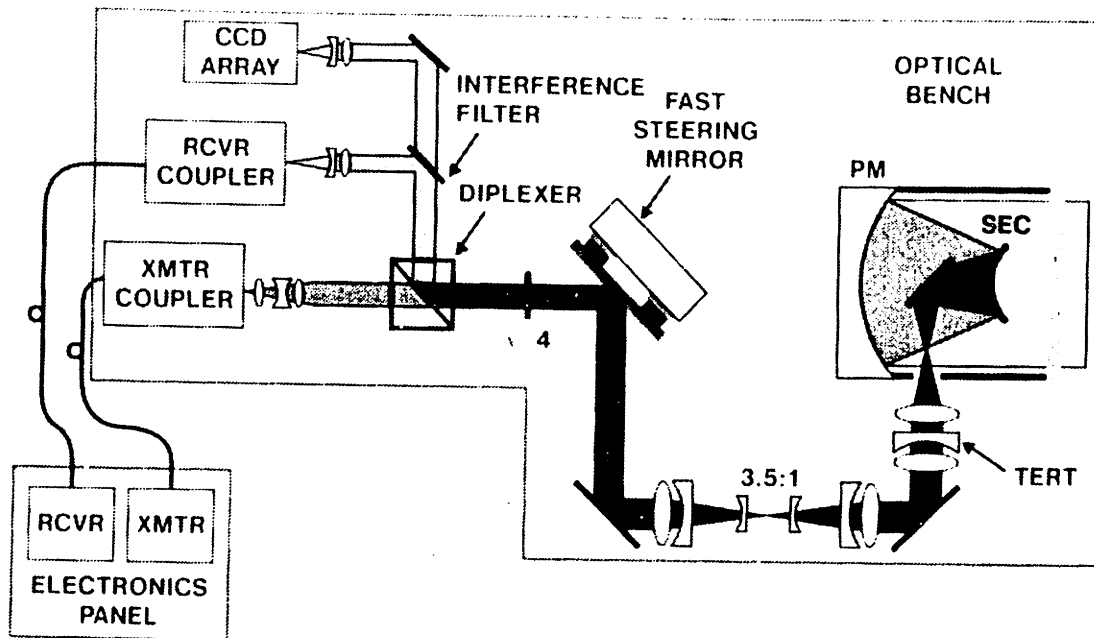


Figure 1.2.1: Strawman optical prescription.

1.3. Nutating Receiver

Figure 1.3.1 shows a block diagram design of a typical spatial tracking loop. In a spatial tracking loop there is a tracking detector (a quad-cell, the electro-optic nutator, a charge coupled device, etc.) which produces spatial tracking error estimates via an error discrimination process. One example is a quadrant detector which uses sum and difference channels to produce azimuth and elevation error estimates. These estimates are then sent through feedback loop compensation to high bandwidth steering optics, typically a fast-steering mirror (FSM), to correct for the disturbances. Because of its large bandwidth (~ 1-5 kHz), a FSM has a small angular range. To provide the necessary low frequency

rejection, a nested servo loop design is employed. A coarse tracking loop, consisting of a coarse pointing mirror (CPM) or a gimbaled telescope, is used to null out the low frequency, large amplitude disturbances and to keep the received signal centered in FSM's operating range. This nested loop design can achieve a dynamic range more than $10^6:1$, low frequency disturbance rejection in excess of 140 dB, and -3 dB closed-loop tracking bandwidths greater than 2.4 kHz [4].

In an optical communication system, the boresight, acquisition, and pointing systems are closely tied to the operation of the tracking system. This document will not discuss these interactions however, assuming that initial signal acquisition has already been accomplished and active tracking has begun. This thesis will also focus on the high bandwidth inner tracking loop, evaluating the theoretical and measured performance of our electro-optic nutator (EON) in a two-axis, high bandwidth spatial tracking loop.

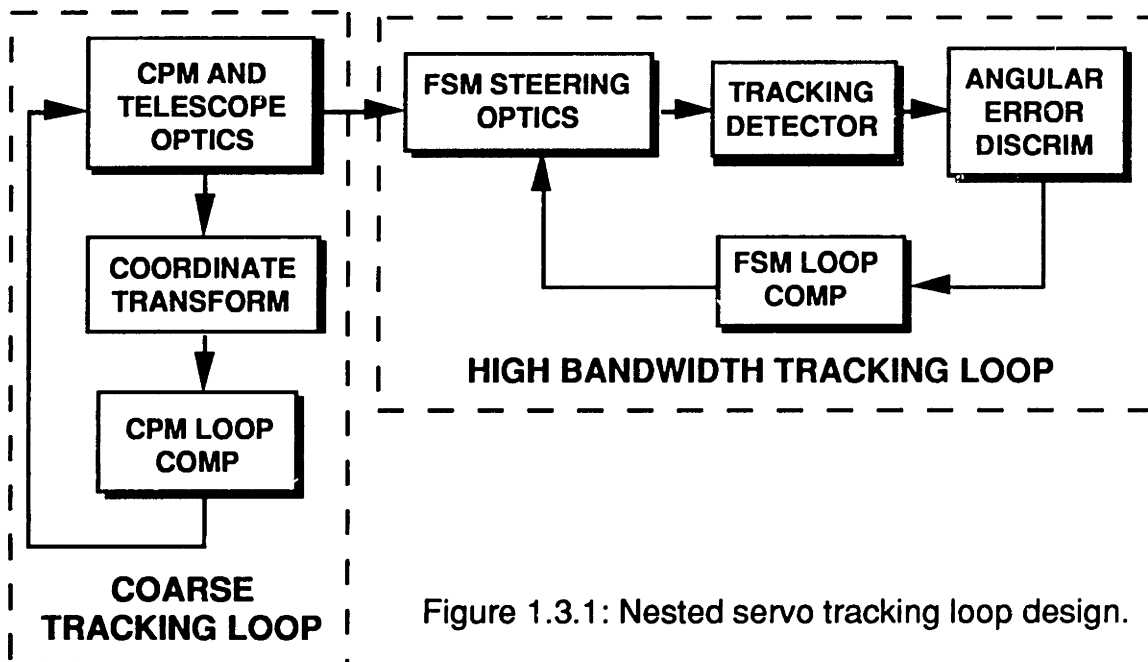


Figure 1.3.1: Nested servo tracking loop design.

1.4. Thesis Organization Overview

A block diagram of the receiver we studied is shown in Fig. 1.4.1. Although a coherent (heterodyne) detection architecture was evaluated, the results are easily extended to a direct detection system. Our receiver has the signal and local-oscillator (LO) beams coupled into one of the inputs of a fiber coupler. In the path of the signal beam, right before it is coupled into the fiber coupler input, is an electro-optic nutator (EON) which nutates the received signal field to create tracking information. After the signal and LO fields mix within the fiber coupler, they are detected by a photodiode, amplified, and bandpass filtered around the heterodyne operating frequency. After the filter, this i.f. signal is sent through an automatic gain control (AGC) stage, designed to maintain constant system gain in the presence of changing received signal power level, and sent to the demodulator. The demodulator produces azimuth and elevation spatial tracking error estimates which are then sent, via loop compensation and a power amplifier, to the steering mirrors, completing the receiver spatial tracking loop.

This thesis presents a fiber-based free-space optical communication receiver, design considerations, theoretical predictions, and experimental verification in a two-axis spatial tracking system. After this overview, Section 2 discusses the nutation concept and the EON used to derive spatial tracking error estimates. Section 3 derives the theoretical noise performance bound for the receiver and compares the prediction to the measured open-loop noise performance. Section 4 motivates the need for an AGC stage and evaluates the performance of the method employed in our receiver. The rejection of the spatial tracking loop is discussed in Section 5 and experimental results of the nutating receiver in a two-axis spatial tracking feedback loop are presented. Finally,

Section 6 presents our conclusions, followed by references, appendices, and acknowledgments.

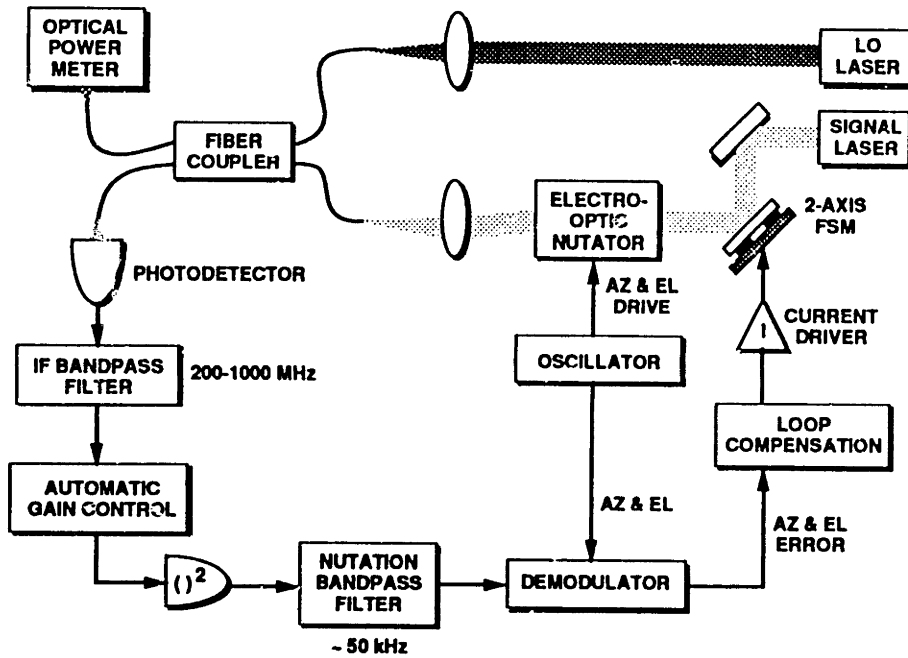


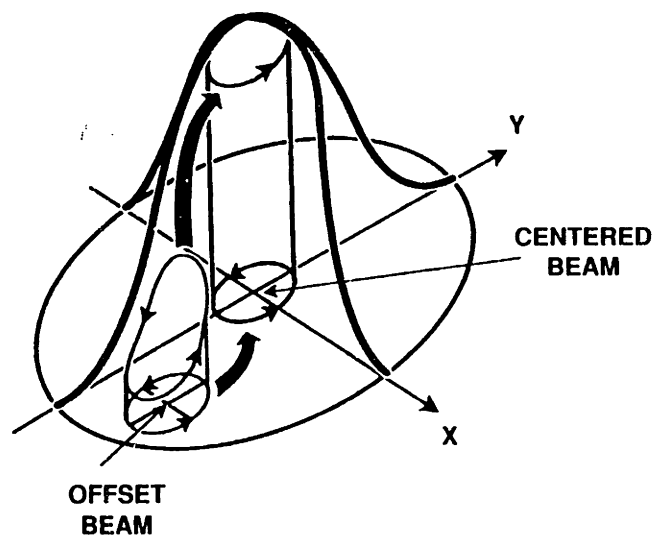
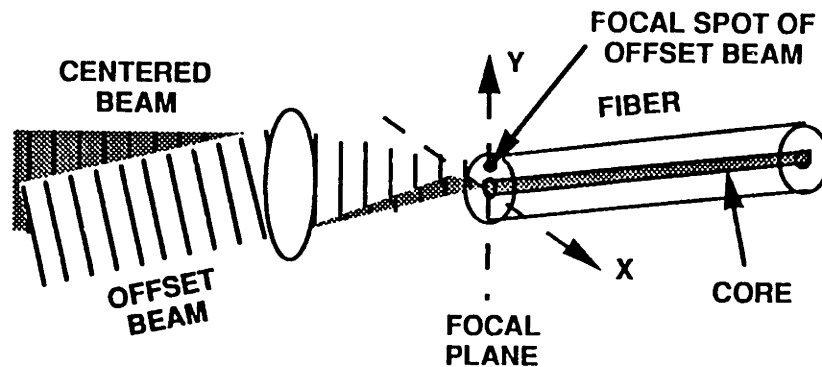
Figure 1.4.1: EON receiver block diagram.

2. Nutation Concept

The receiver spatial tracking loop must effectively couple power from the received planar (far-field) signal field into a single-mode fiber in the presence of environmental disturbances. In order to compensate for these spatial environmental disturbances, a method of estimating the spatial tracking error must be found. One method of deriving tracking information is to add a small nutation¹ signal to the received signal [2,3,9,20,21]. The additional spatial tracking error induced by this known nutation can be used to estimate the unknown tracking error. This estimated error is then minimized by other components in the receiver.

The nutation concept is pictorially outlined in Fig. 2.0.1. The fiber coupling profile shows the signal field coupling efficiency into the receiving single-mode optical fiber, as a function of the position of the signal field focal spot in the plane of the fiber. This profile is derived by performing an overlap integral between the planar signal field and the Gaussian mode profile of the fiber. For this discussion, only the shape of the profile is important; an exact analysis is presented in Sec. 3. If the received signal field is focused onto the center of the fiber core (the origin of the graph), then the maximum possible amount of signal energy is coupled into the fiber. Spatial tracking errors cause the focal point to move off-center and decrease the coupling efficiency. The spatial tracking system is designed to keep the focal spot of the signal field as close to the fiber core center as possible.

¹ In this context, nutation refers to a small amplitude circular scan motion imposed on the signal beam as it enters the receiver. Strictly speaking, this usage of the word does not match any of its definitions. However, this term has become accepted as a word used to describe the method used to derive tracking information that this section highlights; a concept which is essentially the same as used in a conical scan receiver for RF systems.



175974-4

Figure 2.0.1: Nutating receiver fiber coupling profile. Top: A focusing lens focuses an aligned signal field (centered beam) onto the core of the receiving single-mode fiber. A tilted signal field (offset beam) will produce an offset focal spot in the fiber focal plane. Bottom: The nutation scan for a perfectly tracked, nominally centered, signal beam traces out a constant power contour of the fiber coupling profile. For an offset beam, the amount of power coupled into the fiber varies during a nutation scan. These power variations contain the spatial tracking, error information.

Spatial tracking information is produced by imparting a small nutation scan on either the fiber or the received signal field. Two examples of the coupling performance, with a circular nutation signal added, are shown in Fig. 2.0.1. If the focal spot of the signal field, without nutation, is nominally centered on the fiber core, the nutation scan traces out a constant power contour of the fiber coupling profile, and the amount of coupled power remains constant. On the other hand, a nominally off-center signal focal spot travels up and down the coupling profile, coupling a varying amount of power. These fluctuations in received signal power contain the tracking error information.

There are two equivalent ways of imparting this nutation scan onto the coupled signal field; one can either nutate the receiving single-mode optical fiber, or the received signal field itself. Mechanical techniques for nutating the tip of the receiving fiber have been investigated [3,9,20,21]. The next subsection presents a brief introduction to a mechanical nutation tracker developed at Lincoln Laboratory, as a benchmark for comparison to the electro-optic nutator (EON). The nutation concept is then extended to the EON in sections 2.2 -- 2.4 and its performance is evaluated in sections 2.5 -- 2.7.

2.1. Mechanical Nutation

The key element in a mechanical nutating receiver is the resonant fiber coupler, shown in Fig. 2.1.1. Although nutating fiber devices have been built before [22], designing one for space based applications is difficult [3,9,20,21]. The nutation depth (amount of deflection in each axis of the focal plane) required, as well as the high nutation frequency (~10 kHz) necessary to achieve the desired tracking loop bandwidth make the design difficult. In addition, there are the usual requirements for space based systems: small size, lightweight, stable, rugged, low power, etc.

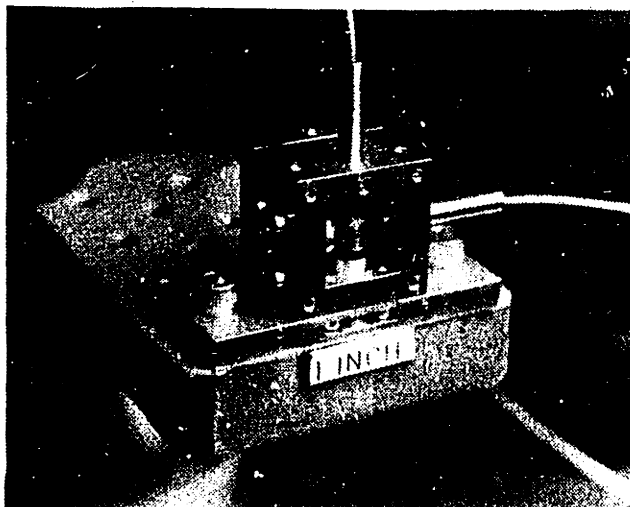


Figure 2.1.1²: The mechanical fiber nutator. The tip of the receiving fiber is mechanically nutated by two "voice coil" actuators driving the mounting flexure at its resonance.

The device shown has two orthogonally mounted "voice coil" actuators to mechanically nutate the tip of the receiving optical fiber (as in Fig. 2.0.1). The fiber is mounted in the center of a mechanical flexure to which the actuators are

² Figures 2.1.1 and 2.1.2 are reproduced with the permission of the author [19].

attached. The mechanical design is carefully tuned so that the resonant frequency of the flexure when loaded by the actuators is the desired nutation frequency. This resonant structure design provides a necessary 20 dB force increase over a non-resonant structure (Fig. 2.1.2).

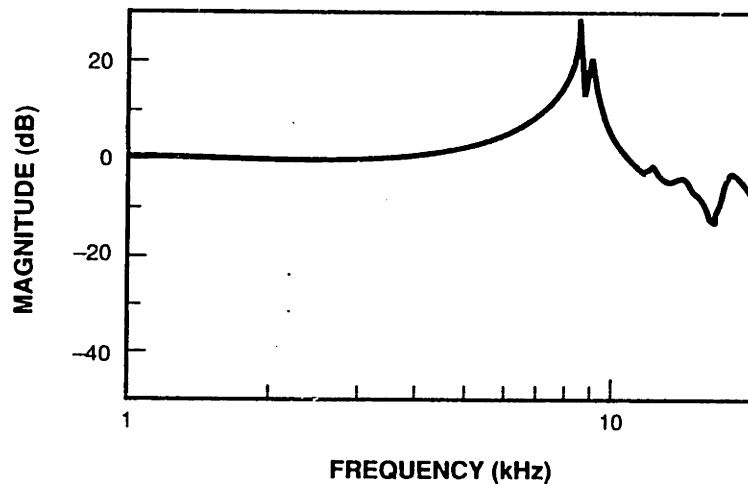


Figure 2.1.2: Mechanical fiber nutator frequency response.

The mechanical nutator clearly has a limited operating frequency range (essentially single frequency operation at the designed mechanical resonance). It is also difficult to match the resonances of the azimuth and elevation actuators. Nevertheless, the device has qualitatively met the requirements for a space based tracking system and an engineering model has been in operation in our laboratory for several years [3,9,20]. The architecture of the spatial tracking subsystem for this engineering model, shown in Fig. 2.1.3, is essentially the same as the architecture used with the electro-optic nutator (Fig. 1.4.1). The electro-optic nutator (EON) is presented as a possible improvement to the mechanical system.

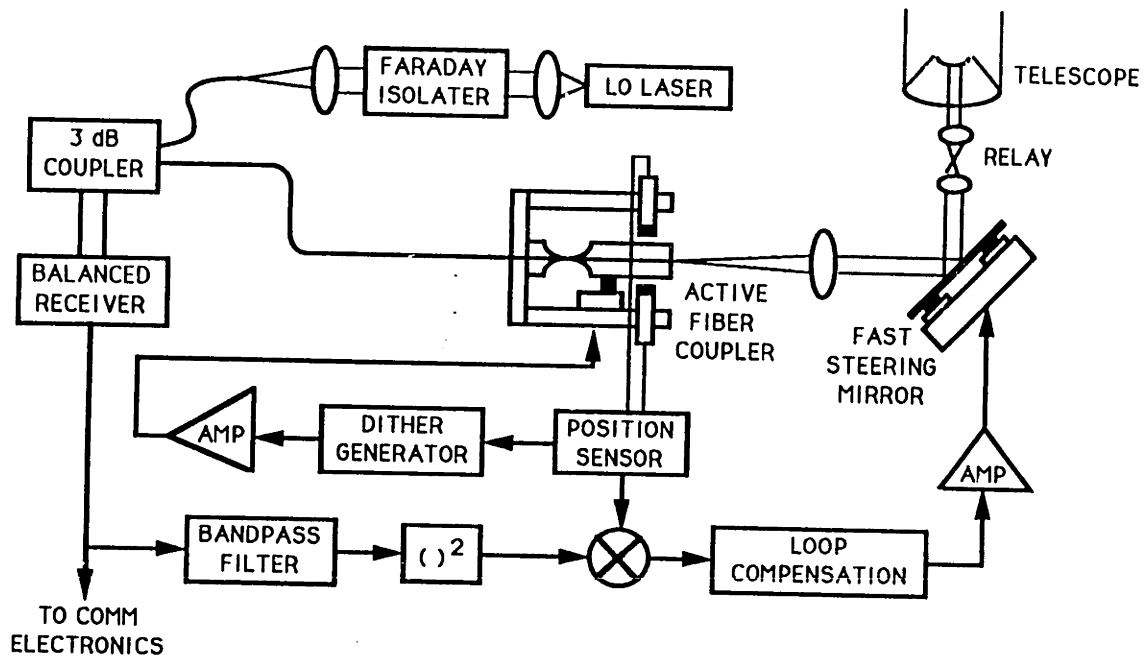


Fig. 2.1.3: Mechanical nutation receiver.

2.2. Electro-Optic Nutation

The mechanical nutator relies on a mechanically driven resonance to produce a nutation scan of the receiving fiber tip. The electro-optic nutator (EON) relies on the complimentary approach to the same basic nutation concept. Instead of mechanically nutating the receiving fiber, the EON electro-optically nutates the received signal field. By changing the signal field's direction of propagation (tilting it) just before it reaches the focusing lens, a nutation scan in the focal plane of the fiber can be achieved (see Fig. 2.0.1).

One simple way to tilt the incoming signal beam is to place an optical wedge in the path of the signal field just before the fiber focusing lens. Consider the

optical wedge shown in Fig. 2.2.1. An incident optical beam propagates with an angle α with respect to the wedge normal. It will emerge with an angle β , which can be found using Snell's Law (assuming $n_{air} \cong 1$):

$$\beta = \sin^{-1}(n_1 \sin \alpha) \quad (2.2.1)$$

where α , from simple geometric considerations, is also the wedge angle. The angle with respect to the original optic axis - θ , that the beam emerges from the wedge, is simply $\beta - \alpha$. If we use the small angle approximation ($\sin \alpha \approx \alpha$) and $n_1 = n_{glass} = 1.5$, then the beam emerges with its direction of propagation tilted by

$$\theta = \frac{\alpha}{2} \quad (2.2.2)$$

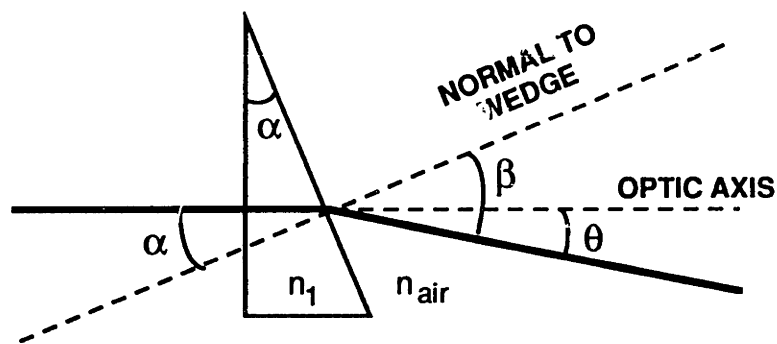


Figure 2.2.1: Deflection of a optical beam by an optical wedge. An optical beam incident at an angle α (the wedge angle) emerges from an optical wedge, with index of refraction $n_1 = 1.5 > n_{air}$, at an angle $\theta = \alpha / 2$.

Rotating the wedge transmits a circularly scanned signal beam to the fiber focusing lens, which will produce a circular nutation scan in the focal plane of the receiving single-mode fiber (Fig. 2.0.1). Nevertheless, mechanically rotating the wedge has the same drawbacks as the mechanical fiber nutator. What is necessary is a device that can perform the same function as a rapidly spinning wedge, without any moving parts: an electrically controlled optical wedge.

Towards this end, consider an object whose index of refraction varies linearly with position as shown in Fig. 2.2.2. A light ray which enters this medium at the bottom (ray B) 'sees' an index of refraction $n_z(x=0) = n$, whereas a light ray that enters at the top (ray A) 'sees' an index of refraction $n_z(x=D) = n + \Delta n$. Since the ray's velocity of propagation in the medium is $c/n_z(x)$, the transit time through the object for ray A is $T_A = \frac{L}{c}(n + \Delta n)$ and for ray B is $T_B = \frac{L}{c}n$. The effective difference in path length between the two light rays, Δy , is the product of the velocity of propagation in the medium and the difference in their transit times through it:

$$\Delta y = \frac{c}{n} [T_A - T_B] = L \frac{\Delta n}{n} \quad (2.2.3)$$

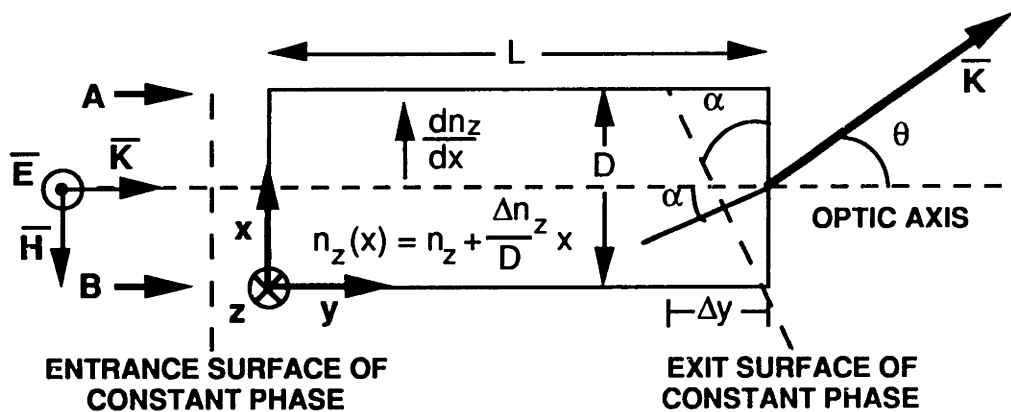


Figure 2.2.2: A medium whose index of refraction varies linearly with position functions like an optical wedge by changing a transmitted beam's direction of propagation. \vec{E} and \vec{H} are the electric and magnetic fields and \vec{K} is the direction of propagation of the optical beam.

If we assume a normally incident planar signal field (as shown), the surfaces of constant phase for the field at the entrance to the medium will be perpendicular to \hat{y} . Because of the index gradient, as the beam passes through

the medium, the surfaces of constant phase at the far end will be tilted by an angle

$$\alpha = \tan^{-1}\left(\frac{\Delta y}{D}\right) = \tan^{-1}\left(\frac{L\Delta n_z}{Dn_z}\right) = \frac{L}{n_z} \frac{dn_z}{dx} \quad (2.2.4)$$

Since the beam's direction of propagation is perpendicular to these surfaces, it is also tilted by the same angle. Using Snell's Law, Eqn. 2.2.3, and small angle approximations, the beam's exit angle at the boundary of the medium is

$$\theta_y = \sin^{-1}\left[\frac{n_z}{n_{air}} \sin\{\alpha\}\right] = \frac{L\Delta n_z}{D} = L \frac{dn_z}{dx} \quad (2.2.5)$$

Since a medium's index of refraction along a particular direction determines the speed at which a beam with polarization along this direction will travel through the medium, we can conclude that a medium whose index of refraction varies linearly along a certain direction will act like a wedge for optical beams with this polarization (compare Eqns. 2.2.2 and 2.2.5). An electro-optic crystal can produce the necessary refractive index gradient via an externally applied electric field. By placing two electro-optic crystals in series in an integrated device, rotated 90° with respect to each other, two dimensional deflections of the beam can be achieved. This is the basic design of our electro-optic nutator (EON).

2.3. Analysis of the Electro-Optic Nutator

In order to rigorously understand how an electro-optic crystal can act like an optical wedge, we need to begin with the description of traveling electromagnetic waves in crystal media. From this description, an expression known as the optical indicatrix is derived. The optical indicatrix is a mathematical abstraction which makes finding the preferred directions of propagation in an electro-optic crystal easier. Finally, the optical indicatrix is used to aid in the design and alignment of the electro-optic nutator (EON).

In an isotropic medium, the dielectric permittivity ϵ is independent of the direction of propagation. If the permittivity changes with direction, then the medium is termed anisotropic. Since a given crystal's geometry is made up of a periodic array of atoms (or molecules), with various degrees of symmetry, one might expect that the dielectric permittivity of a crystal depends upon the direction of propagation through it. Instead of the simple scalar relationship between the electric displacement (\vec{D}) and the electric field (\vec{E}) of a traveling wave in isotropic media³ -- $\vec{D} = \epsilon\vec{E}$, the dielectric response of an anisotropic (crystal) medium is described by means of its dielectric permittivity tensor:

$$\begin{bmatrix} D_x \\ D_y \\ D_z \end{bmatrix} = \begin{bmatrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{21} & \epsilon_{22} & \epsilon_{23} \\ \epsilon_{31} & \epsilon_{32} & \epsilon_{33} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} \quad (2.3.1)$$

where the ϵ_{ij} component of the tensor relates the electric displacement along the i axis -- D_i to the component of the traveling wave's electric field along the j axis -- E_j . It is clear that the index of refraction ($n_{ij} \equiv \sqrt{\epsilon_{ij}/\epsilon_0}$) seen by a wave

³ This relationship is only true if dispersion is neglected, which is correct if the medium has very low transmission loss at the frequencies of interest. Our EON is essentially lossless at optical frequencies.

traveling through this medium depends on both its direction of propagation and its polarization.

If we assume the medium is homogeneous, nonabsorbing (low loss), and magnetically isotropic, then the dielectric tensor for an anisotropic medium, given by Eqn. 2.3.1, is symmetric [24,25] and therefore has only six independent components. Because of the anisotropic nature the medium, the polarization of an arbitrary plane wave may change as it passes through the medium. However, along a given direction of propagation there exist two eigenpolarizations which will not change when traveling through the medium. These two eigenpolarizations can be determined from Eqn. 2.3.1 and Maxwell's Equations.

Remarkably enough, it is always possible to choose a coordinate system such that the off diagonal elements of the dielectric tensor vanish⁴, leaving just the three diagonal terms [32]:

$$\epsilon = \begin{bmatrix} \epsilon_{11} & 0 & 0 \\ 0 & \epsilon_{22} & 0 \\ 0 & 0 & \epsilon_{33} \end{bmatrix} \quad (2.3.2)$$

This coordinate system, where $\epsilon_{11} = \epsilon_x$, $\epsilon_{22} = \epsilon_y$, and $\epsilon_{33} = \epsilon_z$, is called the principal coordinate system of the crystal. Because there are no longer any off diagonal terms in the dielectric tensor when expressed in the crystal's principal coordinate system, the determination of the eigenpolarizations is trivial. For example, a plane wave traveling along the z axis of the principal coordinate system has its eigenpolarizations along the x and y axes of the principal coordinate system.

The two eigenpolarizations for a given direction of propagation within a crystal can also be found from the index ellipsoid or optical indicatrix⁵,

⁴ Obtained by diagonalizing the permittivity tensor of Eqn. 2.3.1.

⁵ A formal proof of the equivalence of the indicatrix to solving Maxwell's Equations can be found in Yariv [24] and Nye [25].

$\sum_{i=1}^3 \sum_{j=1}^3 \frac{1}{n_{ij}} x_i x_j = 1$. In the principal coordinate system, the optical indicatrix reduces

to:

$$\frac{x^2}{n_{11}^2} + \frac{y^2}{n_{22}^2} + \frac{z^2}{n_{33}^2} = 1 \quad (2.3.3)$$

where the notational correspondences $n_{11} \rightarrow n_x$, $n_{22} \rightarrow n_y$, $n_{33} \rightarrow n_z$ have been used. This equation represents a general ellipsoid, with its major axes parallel to the principal crystal coordinate directions \hat{x} , \hat{y} , and \hat{z} , with lengths of $2n_{11}$, $2n_{22}$, and $2n_{33}$ respectively, where $n_{ii} \equiv \sqrt{\epsilon_{ii}/\epsilon_o}$ are the principal indices of refraction. Physically, the equation represents surfaces of constant energy density of a wave traveling through the crystal medium; i.e. the same wave, but with any possible direction. The geometric construction performed when solving the index ellipsoid is analogous to fixing some of the electromagnetic wave's parameters (namely requiring only transverse fields) and allowing others (the propagation direction) to vary consistent with Maxwell's equations.

The optical indicatrix is mainly used to determine the two eigen-indices of refraction, and thus the corresponding eigenpolarizations, for an arbitrary direction of propagation within the crystal⁶. This is accomplished by the following geometric construction: Determine the ellipse generated by the intersection of a plane through the origin that is normal to the direction of propagation with the index ellipsoid of Eqn. 2.3.3. The major and minor axes of this ellipse are parallel

⁶ The optical indicatrix actually determines the two mandated directions of the electric displacement vector (\vec{D}) associated with an arbitrary direction of propagation within the crystal. These two directions, \hat{D}_1 and \hat{D}_2 , along with the direction of propagation, \hat{K} , form an orthogonal triad which can be used as a set of basis vectors to describe the traveling plane wave. In the principal coordinate system, the two eigenpolarizations, \hat{E}_1 and \hat{E}_2 , are parallel to the two displacement vectors. Since we are ultimately concerned with the eigenpolarizations of a given direction of propagation, we will henceforth be operating strictly in the crystal principal coordinate system.

to the directions of the two eigenpolarizations corresponding to the given direction of propagation.

Since the dielectric permittivity tensor for an anisotropic crystal depends upon the atomic structure of the crystal, any distortion of the crystal lattice will change the tensor's components. When an external electric field is applied to a crystal, a slight deformation of the crystal lattice is produced due to the redistribution of the internal bond charges. This deformation induces small changes in some of the dielectric tensor's components, and thus changes the corresponding indices of refraction in the optical indicatrix. Each of the changes in indices of refraction -- Δn_{ij} , can be expressed as a power series expansion:

$$\Delta n_{ij} = \sum_{k=1}^3 r_{ijk} \bar{E}_k + \sum_{k=1}^3 \sum_{l=1}^3 R_{ijkl} E_k E_l + \dots \quad (2.3.4)$$

where r_{ijk} are the linear (Pockel's) electro-optic coefficients, R_{ijkl} are the quadratic (Kerr) electro-optic coefficients, and E_1, E_2, E_3 represent the applied E_x, E_y, E_z fields respectively. In crystals which possess inversion symmetry⁷ the linear coefficients must all be zero, leaving only the second, and higher, order terms. In many crystals that have non-zero linear electro-optic terms, the higher order terms are insignificant and thus can be neglected.

Considering only the linear electro-optic effect, the expression for the optical indicatrix in the presence of an applied field becomes:

$$\sum_{i=1}^3 \sum_{j=1}^3 \left(\frac{1}{n_{ij}} + \Delta n_{ij} \right) x_i x_j = \sum_{i=1}^3 \sum_{j=1}^3 \left(\frac{1}{n_{ij}} + \sum_{k=1}^3 r_{ijk} E_k \right) x_i x_j = 1 \quad (2.3.5)$$

As was mentioned previously, physical arguments can be used to show that the dielectric permittivity tensor -- ϵ (and hence the index tensor -- $n_{ij}(\bar{E})$) is

⁷ A crystal possessing inversion symmetry has a center of symmetry. Reversing the applied field direction changes the sign of the applied field without changing the physical appearance of the crystal. The linear electro-optic effect cannot exist since this implies $\Delta n_{ij} = -\Delta n_{ij}$, which is only true if $\Delta n_{ij} = 0$.

symmetric; this implies $n_{ij} = n_{ji}$ and $r_{ijk} = r_{jik}$. Because of this symmetry, we can introduce a set of contracted indices to simplify the electro-optic tensor, reducing the number of independent components from 27 to 18, as follows:

$$\begin{aligned}
1 &= (11) \rightarrow r_{1k} = r_{11k} \\
2 &= (22) \rightarrow r_{2k} = r_{22k} \\
3 &= (33) \rightarrow r_{3k} = r_{33k} \\
4 &= (23) = (32) \rightarrow r_{4k} = r_{23k} = r_{32k} \\
5 &= (13) = (31) \rightarrow r_{5k} = r_{13k} = r_{31k} \\
6 &= (12) = (21) \rightarrow r_{6k} = r_{12k} = r_{21k}
\end{aligned} \tag{2.3.6}$$

These contracted indices simplify the expression for the optical indicatrix:

$$\begin{aligned}
&x^2 \left[\frac{1}{n_1^2} + \sum_{k=1}^3 r_{1k} E_k \right] + y^2 \left[\frac{1}{n_2^2} + \sum_{k=1}^3 r_{2k} E_k \right] + z^2 \left[\frac{1}{n_3^2} + \sum_{k=1}^3 r_{3k} E_k \right] \\
&+ 2yz \left[\frac{1}{n_4^2} + \sum_{k=1}^3 r_{4k} E_k \right] + 2xz \left[\frac{1}{n_5^2} + \sum_{k=1}^3 r_{5k} E_k \right] + 2xy \left[\frac{1}{n_6^2} + \sum_{k=1}^3 r_{6k} E_k \right] = 1
\end{aligned} \tag{2.3.7}$$

where x, y and z are coordinates within the crystal geometry, n_1 -- n_6 are the indices of refraction, and r_{1k} -- r_{6k} are the changes in those indices produced by an electric field applied to the crystal (the linear electro-optic coefficients).

The general expression for the optical indicatrix (Eqn. 2.3.7) can always be reduced to the form of Eqn. 2.3.3:

$$\frac{x^2}{n_1^2} + \frac{y^2}{n_2^2} + \frac{z^2}{n_3^2} = 1 \tag{2.3.8}$$

In the absence of an applied field, x, y and z are the principal crystal coordinates and n_1, n_2, n_3 are the indices along those directions. Eqn 2.3.8 implies $(1/n_4^2) = (1/n_5^2) = (1/n_6^2) = 0$ in the principal coordinate system; so the cross terms of the indicatrix only appear when the crystal is subjected to an applied electric field. In the presence of applied fields Eqn. 2.3.7 can also be reduced to the form of Eqn. 2.3.8, with new coordinate axes x', y' and z' which are related to the crystal's principal axes by a coordinate transformation (in fact, a rotation).

The electro-optic crystal used in our electro-optic nutator (EON) is Lithium-Tantalate (LiTaO₃). LiTaO₃ is an anisotropic uniaxial crystal⁸ belonging to the crystal symmetry group 3m [24,25,29]. Considering only the linear electro-optic effect, the electro-optic tensor for LiTaO₃ in contracted notation is⁹:

$$r = \begin{bmatrix} 0 & -r_{22} & r_{13} \\ 0 & r_{22} & r_{13} \\ 0 & 0 & r_{33} \\ 0 & r_{51} & 0 \\ r_{51} & 0 & 0 \\ -r_{22} & 0 & 0 \end{bmatrix} \quad (2.3.9)$$

LiTaO₃ electro-optic tensor

The structure of the tensor shown above, its zero, equal, and opposite elements, can be determined entirely from the crystal geometry [25]. Each of the non-zero terms r_{ij} will produce a change in the crystal index of refraction along the i axis when an external field is applied along the j axis. Of the tensor's non-zero terms, r_{33} is the largest ($33 \cdot 10^{-12} \text{ m/V}$), and therefore is the term to exploit. Actually, as this derivation will soon show, it is the largest $n^3 r_{ij}$ combination which is used for deflection, but this will still involve r_{33} . Clearly, we expect a refractive index change along the crystal's extraordinary (z) axis when a field is applied along this axis.

By using Eqn. 2.3.9 for the electro-optic coefficients in Eqn. 2.3.7, the general optical indicatrix for LiTaO₃ in the principal coordinate system is found to be:

⁸ A uniaxial crystal has two different indices of refraction along the crystal principal axes, ordinary (n_o) and extraordinary (n_e), in the absence of an applied electric field. Two of the principal axes possess the ordinary index of refraction and the third has the extraordinary. By convention the extraordinary axis is chosen to be the z axis, so Eqn. 2.3.8 becomes $x^2(1/n_o^2) + y^2(1/n_o^2) + z^2(1/n_e^2) = 1$.

⁹ Because of its contracted form, this matrix does not conform to the usual conventions of matrix arithmetic; it must first be expanded before being used algebraically in a matrix equation.

$$1 = \left(\frac{1}{n_o^2} - r_{22}E_y + r_{13}E_z \right) x^2 + \left(\frac{1}{n_o^2} + r_{22}E_y + r_{13}E_z \right) y^2 + \left(\frac{1}{n_e^2} + r_{33}E_z \right) z^2 + 2r_{51}E_y yz + 2r_{51}E_x xz - 2r_{22}E_x xy \quad (2.3.10)$$

If we choose to propagate through the crystal along the y axis and only apply fields normal to the direction of propagation, Eqn. 2.3.10 reduces to:

$$1 = \left(\frac{1}{n_o^2} + r_{13}E_z \right) x^2 + \left(\frac{1}{n_o^2} + r_{13}E_z \right) y^2 + \left(\frac{1}{n_e^2} + r_{33}E_z \right) z^2 + 2r_{51}E_x xz - 2r_{22}E_x xy \quad (2.3.11)$$

This is the general expression for the optical indicatrix of an electromagnetic wave propagating along a LiTaO₃ crystal's y -axis, with external fields applied normal to the direction of propagation (E_x and E_z).

Our objective is to find a coordinate transformation which will reduce this quadratic form to its principal coordinate system (like Eqn. 2.3.8). An electromagnetic wave propagating along \hat{y} will only have polarization components along \hat{x} and \hat{z} . Consequently, the ellipsoid represented by Eqn. 2.3.11 can be reduced to an ellipse in the xz plane by dropping the y^2 term and evaluating the expression at $y = 0$:

$$1 = \left(\frac{1}{n_o^2} + r_{13}E_z \right) x^2 + \left(\frac{1}{n_e^2} + r_{33}E_z \right) z^2 + 2r_{51}E_x xz \quad (2.3.12)$$

The new reduced coordinate system (x', y' and z') will be the major axes of the ellipse described by Eqn. 2.3.12 and the crystal's original y -axis ($y' = y$).

As mentioned earlier, one can determine x' and z' from the diagonalization of the coefficient matrix formed by Eqn. 2.3.12. We begin by finding the eigenvalues (γ) of this matrix:

$$\det \begin{bmatrix} \left(\frac{1}{n_o^2} + r_{13}E_z \right) - \gamma & r_{51}E_x \\ r_{51}E_x & \left(\frac{1}{n_e^2} + r_{33}E_z \right) - \gamma \end{bmatrix} = 0 \quad (2.3.13)$$

which are simply the diagonal elements if the cross terms can be neglected ($r_{51}E_x \ll 1/n_o^2$ and $r_{51}E_x \ll 1/n_e^2$). The corresponding eigenvectors, representing the major and minor axes of the ellipse, are just the original principal crystal x and z axes respectively, $\vec{\Gamma}_1 = \hat{x}$, $\vec{\Gamma}_2 = \hat{z}$. Ignoring the variation in the crystal's refractive index along the direction of propagation (\hat{y}), the final expression for the optical indicatrix due to the applied transverse fields is:

$$1 = \left(\frac{1}{n_o^2} + r_{13}E_z \right) x^2 + \left(\frac{1}{n_o^2} \right) y^2 + \left(\frac{1}{n_e^2} + r_{33}E_z \right) z^2 \quad (2.3.14)$$

where the indices of refraction along each direction are simply the coefficients of the three terms.

Since we wish to exploit the strongest electro-optic coefficient, r_{33} , we need to have the signal field impinging on the crystal have its polarization only along \hat{z} . Any polarization along \hat{x} will also be electro-optically deflected, due to the r_{13} component, but this effect is weaker and undesired. The basic geometry is shown in Fig. 2.3.1. Any voltage applied to the electrodes will produce a uniform electric field along \hat{z} throughout the bulk of the crystal (ignoring fringing effects at the ends). This voltage will uniformly change the crystal's index of refraction along \hat{z} through the r_{33} electro-optic component. To make the LiTaO₃ crystal perform like an optical wedge, the change in the index must be linear in x along \hat{z} (refer back to Sec. 2.2, Eqn. 2.2.5).

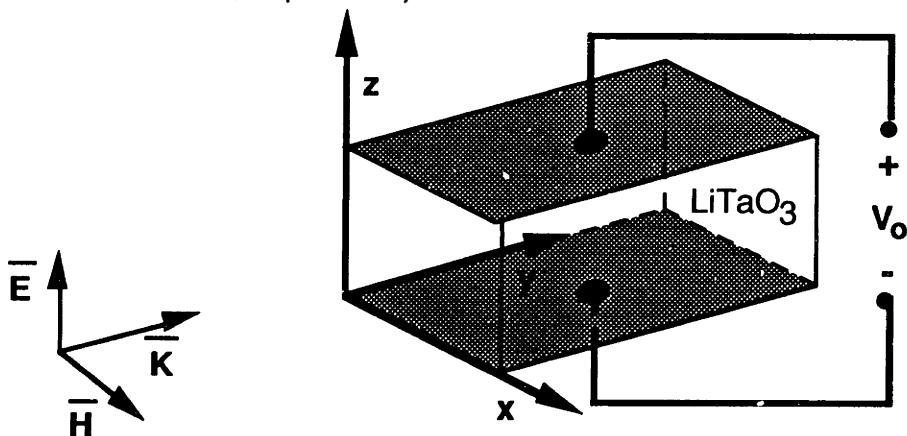


Figure 2.3.1: Basic alignment prescription for static LiTaO₃ electro-optic deflector. In the crystal principal coordinate system, the optical signal field propagates along \hat{y} with polarization along \hat{z} .

The crystal's index of refraction along \hat{z} (from Eqn. 2.3.14) is:

$$\frac{1}{n_z^2} = \frac{1}{n_e^2} + r_{33}E_x \rightarrow n_z^2 = n_e^2 \left(\frac{1}{1 + n_e^2 r_{33} E_x} \right) \quad (2.3.15)$$

By using the Taylor series expansion of $1/(1+x) \approx 1-x$ about $x=0$ (assuming $n_e^2 r_{33} E_x \ll 1$) this expression becomes $n_z = n_e \sqrt{1 - r_{33} E_x n_e^2}$. Using another Taylor series expansion of $\sqrt{1-x} \approx 1 - x/2$ about $x=0$ (again assuming $n_e^2 r_{33} E_x \ll 1$) the expression for the crystal's index of refraction along \hat{z} becomes

$$n_z = n_e - \frac{n_e^3}{2} r_{33} E_x \quad (2.3.16)$$

In order to make this relationship conform to the desired medium property (Fig. 2.2.2), we need the derivative of this index with respect to x to be a constant. Since all of the terms except E_x are measured constants, our task now becomes finding a means of establishing an electric field inside the LiTaO₃ crystal which is linearly proportional to the x position along \hat{z} ($d/dx[E_x] = K$). Equivalently, we must find some electrode configuration to place on the crystal surfaces which will create the desired field inside the crystal.

2.4. Alignment Prescription for the Electro-Optic Nutator

From electromagnetic theory, the electrostatic potential inside a region of constant permittivity must satisfy Poisson's Equation,

$$\nabla^2\Phi = -\frac{\rho}{\epsilon} \quad (2.4.1)$$

where Φ is the potential, ρ is the charge density, and ϵ is the dielectric permittivity of the region. Since we are only allowed to place electrodes (constrained potentials) on the outer surfaces of the LiTaO₃ crystal, there are no charges in the bulk of the material; consequently Poisson's Equation reduces to Laplace's Equation, $\nabla^2\Phi = 0$. Solutions to Laplace's Equation in the three basic coordinate systems (Cartesian, cylindrical and spherical) can be found in most electromagnetic fields textbooks [28,34].

Recall that we were looking for an electric field whose derivative along a perpendicular direction is a constant. One solution of interest has the form of:

$$\Phi = \underbrace{\frac{C}{2}r^2 \sin(2\theta)}_{\text{cylindrical coordinates}} = Cr^2 \sin(\theta)\cos(\theta) = \underbrace{Cxz}_{\text{cartesian coordinates}} \quad (2.4.2)$$

where r and θ are cylindrical coordinates, x and z are Cartesian coordinates, and C is a constant determined from boundary conditions. This solution is the potential distribution of an electric quadrupole (alternate equidistant plus and minus charges in each of the four quadrants of a Cartesian plane) and is plotted in Fig. 2.4.1. Since the electric field is the negative gradient of the potential, $\vec{E} = -\nabla\Phi$, the quadrupole electric field:

$$\vec{E} = -C(z\hat{x} + x\hat{z}) \quad (2.4.3)$$

is linearly proportional to the z position in \hat{x} and linearly proportional to the

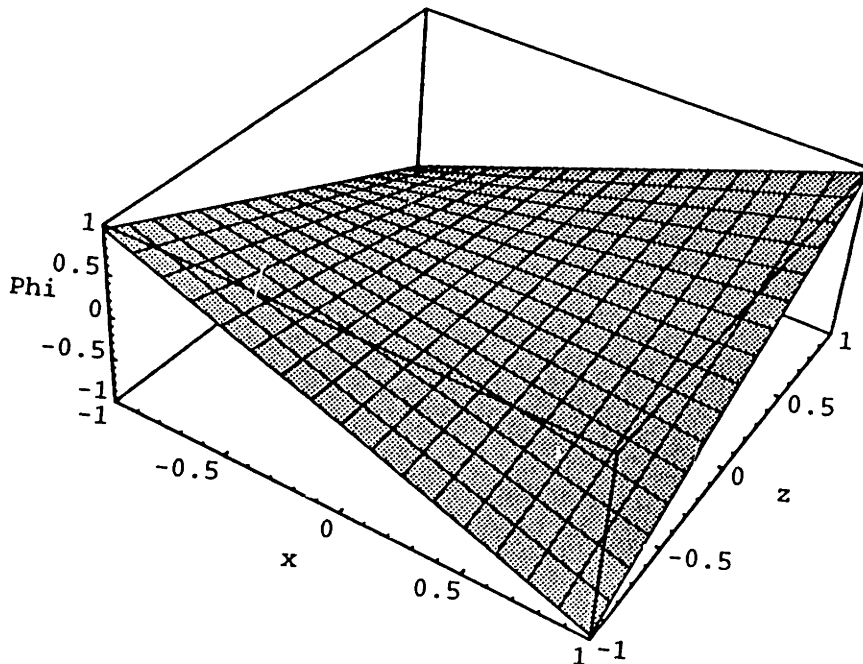
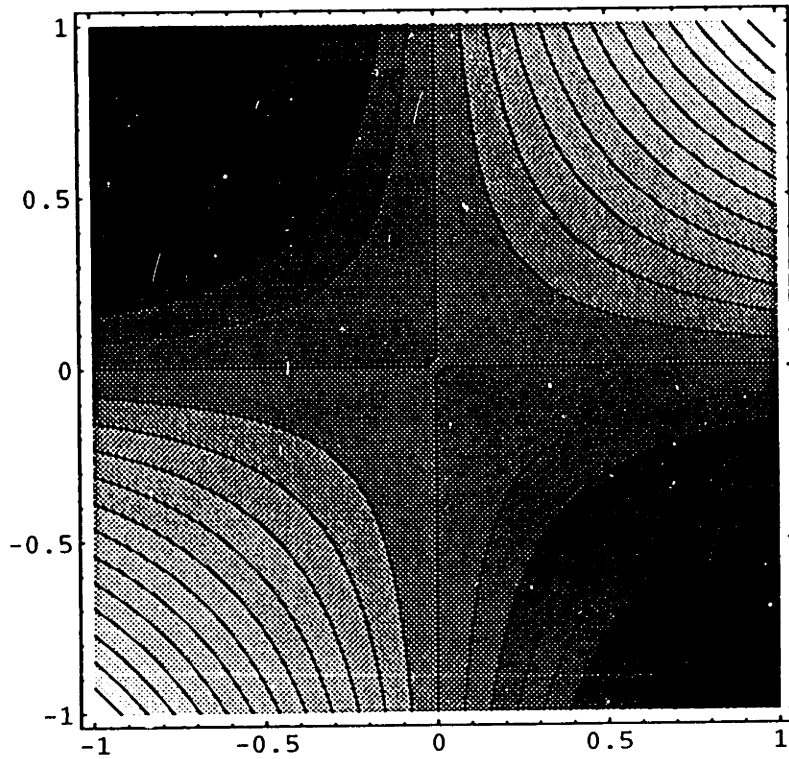


Figure 2.4.1: Top: A contour plot of the quadrupole potential $\Phi = xz$. Constant potential contours are hyperbolas. Bottom: A three-dimensional plot of $\Phi = xz$. The electric field produced by this potential, $\vec{E} = -z\hat{x} - x\hat{z}$, is linearly proportional to the xz position in the coordinate plane.

x position in \hat{z} . By re-examining Eqns. 2.2.5 and 2.3.16, one can observe that this is the desired property for the crystal's electric field. It will produce a deflection in the xy plane of the crystal's principal coordinate system.

Now that we have found the desired potential, the only obstacle that remains is finding an external electrode configuration which will induce this potential inside our LiTaO₃ crystal. One way to generate the proper field is to build electrodes, which are equipotentials, that match the surfaces of constant potential shown in Fig. 2.4.1. On finite crystal boundaries, this electrode arrangement is shown in Fig. 2.4.2. Unfortunately, it is practically impossible to cut a crystal into this shape.

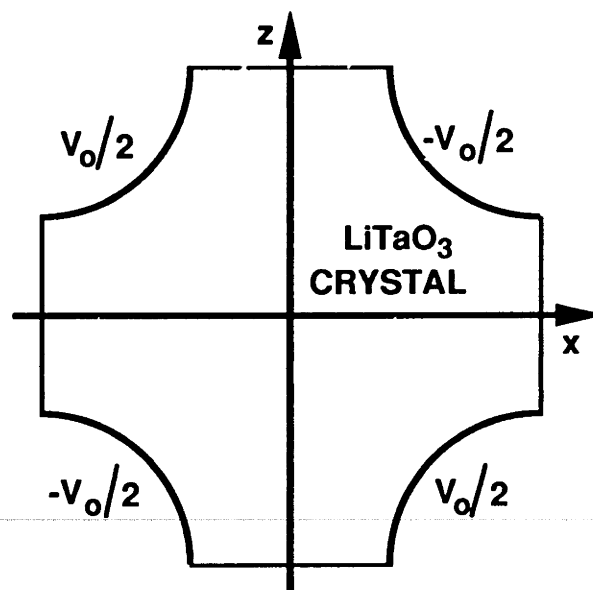


Figure 2.4.2: The 'best' electrode configuration on finite sized crystal boundaries to produce a quadrupole electric field inside the crystal medium.

A practical approximation to this electrode arrangement relies on cutting the crystal in the shape of a regular octagon.¹⁰ The octagonal electrode

¹⁰ A regular octagon has sides of equal length, which implies that all of its interior angles are equal as well, with value = $3\pi/4$.

configuration of Fig. 2.4.3 will produce a reasonable approximation to the electric quadrupole near the origin. By comparing the 'best' electrode arrangement of Fig. 2.4.2 with the employed approximation of Fig. 2.4.3, it is clear that the octagonal approximation hinges on the accuracy of the four linear electrodes as approximations of the four hyperbolic electrodes in the 'best' configuration. It should also be noted that deviations from an ideal quadrupole electric field are not as important as one might initially expect because we employ a synchronous demodulation scheme to derive angular error estimates. Imperfections in the field will reduce the overall gain of the spatial tracking loop, but they will not affect its basic operation.

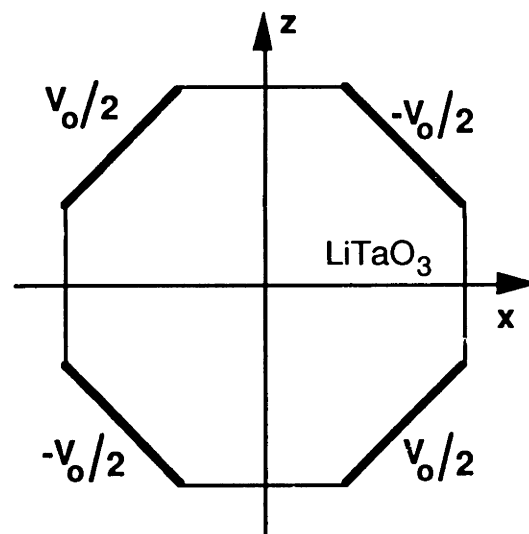


Figure 2.4.3: Front face view of LiTaO₃ deflector. Alternate crystal faces have electrodes. A potential $V_o/2$ is applied to two opposite electrodes and the other two are at $-V_o/2$. This produces a quadrupole field which is linear in z position in \hat{x} and in x position in \hat{z} .

Fig. 2.4.3 shows the final geometry for our LiTaO₃ deflector in the crystal principal coordinate system, looking into the front face of the crystal. Alternate faces of the regular octagon crystal have electrodes attached to them. The two

electrodes in the first and third quadrants have a potential $-V_o/2$ applied to them and the two in the second and fourth quadrants have $V_o/2$ applied to them. Referring to the equation for the field inside the crystal (Eqn. 2.4.3), one can see that this configuration has the desired property discussed in Sec. 2.2.

The flat electrodes of the octagonal crystal shape are an approximation to the quadrupole potential of Eqn. 2.4.2. To determine the constant in this equation one must consider the boundary conditions imposed by the crystal geometry of Fig. 2.4.3, as shown in Fig. 2.4.4:

$$\begin{aligned} \Phi &= -\frac{V_o}{2} \text{ when } (x,z) = \left(\frac{\sqrt{2}}{2} R_o, \frac{\sqrt{2}}{2} R_o \right), \left(-\frac{\sqrt{2}}{2} R_o, -\frac{\sqrt{2}}{2} R_o \right) \\ \Phi &= +\frac{V_o}{2} \text{ when } (x,z) = \left(-\frac{\sqrt{2}}{2} R_o, \frac{\sqrt{2}}{2} R_o \right), \left(\frac{\sqrt{2}}{2} R_o, -\frac{\sqrt{2}}{2} R_o \right) \end{aligned} \quad (2.4.4)$$

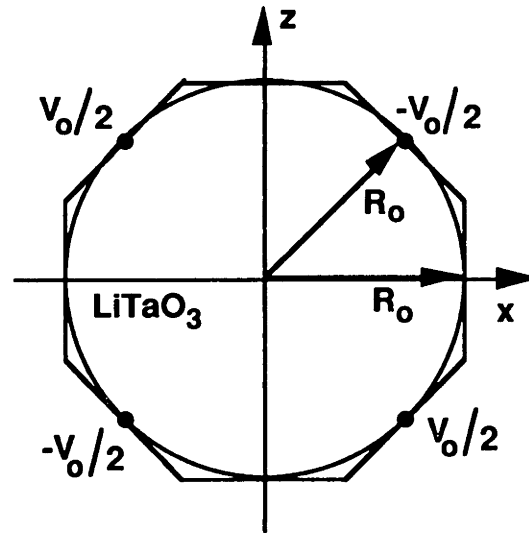


Figure 2.4.4: Boundary conditions imposed by our LiTaO₃ crystal electrode configuration.

After substituting these boundary conditions into the quadrupole potential, the potential becomes:

$$\Phi = -\frac{V_o}{R_o^2} xz \quad (2.4.5)$$

where V_o is the voltage applied to the electrodes of Fig. 2.4.3 and R_o is one-half of the inner diameter of the LiTaO₃ crystal.

Now that we have an expression for the potential inside the LiTaO₃ deflector, all that remains is the determination of the amount of deflection expected. To do this we must take the electric field produced by the quadrupole potential

$$\vec{E} = -\nabla\Phi = \frac{V_o}{R_o^2}(z\hat{x} + x\hat{z}) \quad (2.4.6)$$

and substitute the z component of this field into the expression for the crystal's index of refraction along this direction derived in Sec. 2.3 (Eqn. 2.3.16):

$$n_z = n_e - \frac{n_e^3}{2}r_{33}E_z = n_e - \frac{n_e^3}{2}r_{33}\frac{V_o}{R_o^2}x \quad (2.4.7)$$

Finally this expression is substituted into the basic expression for the external deflection angle of a medium with a linear index gradient derived in Sec. 2.2 (Eqn. 2.2.5):

$$\theta_y = L\frac{dn_z}{dx} = \left(\frac{-Ln_e^3r_{33}}{2R_o^2}\right)V_o \quad (2.4.8)$$

To determine expected deflection, one must substitute into Eqn. 2.4.8 the values for our LiTaO₃ crystals; the crystal length -- $L = 8 \cdot 10^{-3}m$, the extraordinary index of refraction for LiTaO₃ -- $n_e = 2.18$, the appropriate electro-optic tensor coefficient -- $r_{33} = 33 \cdot 10^{-12}m/V$, and one-half of the inner radius of the octagonal crystals -- $R_o = 2.75 \cdot 10^{-3}m$. These values yield an expected deflection of $\theta_y = 0.18\mu rad/V$ for the crystals in our electro-optic nutator (EON). This value compares very well with our measured deflection sensitivity of $\theta_y = 0.17\mu rad/V$.

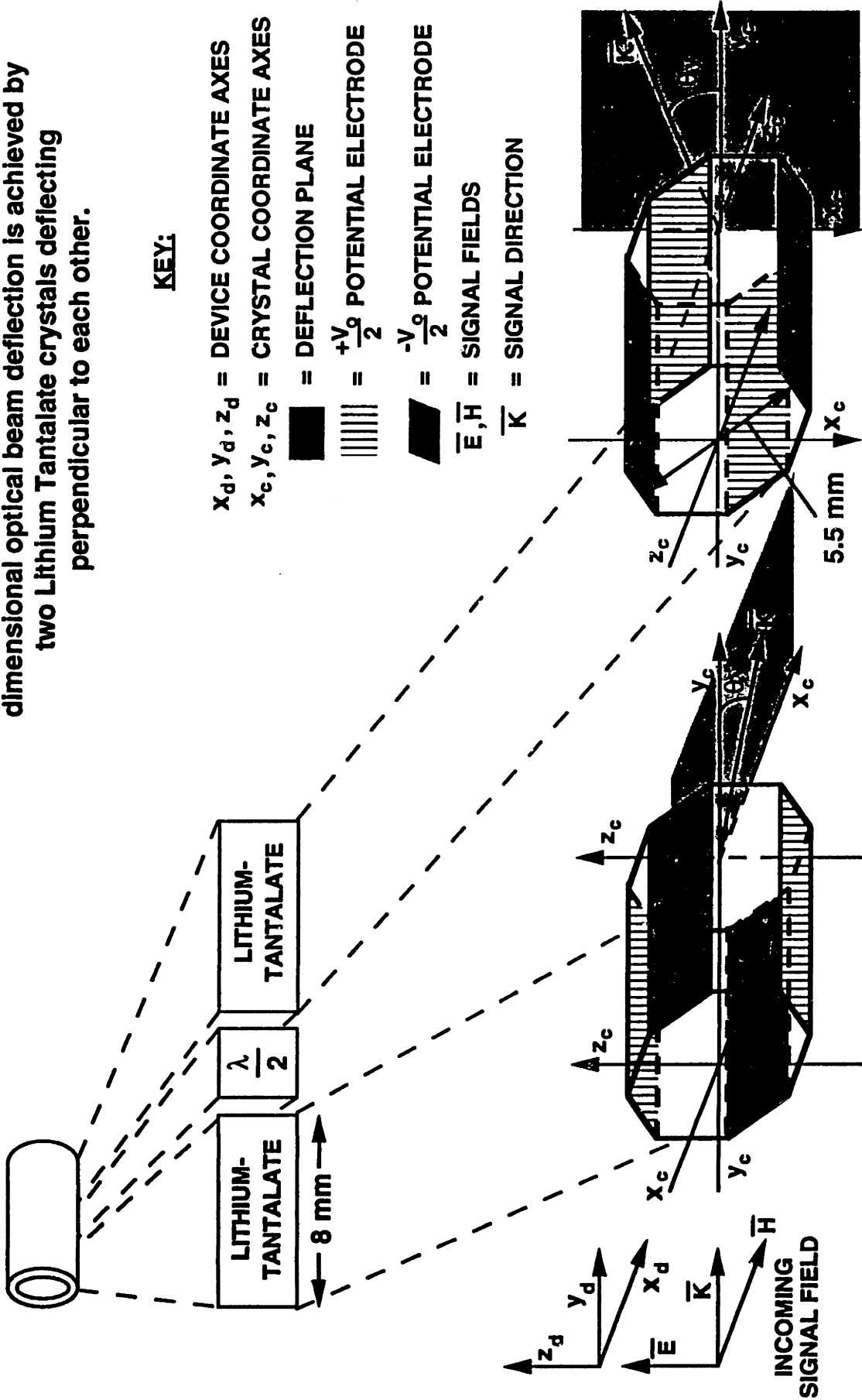
To achieve the desired two dimensional deflections, a second LiTaO₃ crystal is necessary. The first crystal is aligned to deflect along θ_y (like Fig. 2.4.3) and the second is used to deflect perpendicular to it. Since each LiTaO₃ beam deflector will only deflect a transmitted beam whose polarization is along the axis

of deflection, a $\lambda/2$ plate is placed between the two crystals to rotate the signal field's initial z polarization to the z crystal axis of the second deflector. An expanded view of this geometry, the internal structure of the EON, is shown in Fig. 2.4.5. Inside the device are two LiTaO₃ crystals, each 8mm long and 5.5mm inside diameter, with a $\lambda/2$ plate between them.




The first LiTaO₃ crystal deflector conforms to the proper geometry of Fig. 2.4.3 and the proper signal field alignment of Fig. 2.3.1. The incoming signal field is propagating down the first crystal's y axis (y_c) with its polarization along the first crystal's z axis (z_c). The transmitted beam emerges deflected by an angle θ_y in the $x_c y_c$ plane of the first crystal, which is also the $x_d y_d$ plane of the EON. The polarization of this transmitted field is then rotated 90 degrees by a $\lambda/2$ plate before passing into the second crystal. The second crystal is aligned like the first one, so that the beam is still propagating along the y crystal axis (y_c) and is polarized along the z crystal axis (z_c). The beam transmitted through the second crystal is also deflected by an angle θ_y in the $x_c y_c$ plane of the second crystal, which is the $y_d z_d$ plane of the device. Therefore, after passing through both crystals, the original signal field is deflected in two dimensions. By sending this deflected beam into a focusing lens and applying a sine wave signal to one crystal and a cosine signal to the other, a circular nutation scan is imparted on the received signal field (refer back to Fig. 2.0.1).

The astute observer may notice that the signal beam's direction after passing through the first crystal of the EON is no longer exactly along the optic (y_d) axis of the device. Thus, when the polarization of this beam is rotated 90 degrees by the $\lambda/2$ plate, its polarization will not be strictly along the z axis (z_c) of the second crystal. There will be some component of this beam's polarization which is aligned with the second crystal's y axis (y_c) and it will be deflected by the r_{13} electro-optic component (refer to Eqn. 2.3.11). This is an undesired effect, but it

Figure 2.4.5: Inside the electro-optic nutator. Two dimensional optical beam deflection is achieved by two Lithium Tantalate crystals deflecting perpendicular to each other.



KEY:

- X_d, Y_d, Z_d = DEVICE COORDINATE AXES
- X_c, Y_c, Z_c = CRYSTAL COORDINATE AXES
-  = DEFLECTION PLANE
-  = $+\frac{V_0}{2}$ POTENTIAL ELECTRODE
-  = $-\frac{V_0}{2}$ POTENTIAL ELECTRODE
- \vec{E}, \vec{H} = SIGNAL FIELDS
- \vec{K} = SIGNAL DIRECTION

is negligible due to the fact that θ_y from the first crystal is very small (μrad) and for LiTaO_3 $r_{13}/r_{33} \sim 1/4.4$. The component of beam's polarization that is along the y_c axis will be $E \sin(\theta_y) \cong E\theta_y$ (small angle approximation); an error on the order of 10^{-6} .

2.5. Performance of the Electro-Optic Nutator

A photograph of the electro-optic nutator (EON) is shown in Fig. 2.5.1 and its measured performance as used in the receiver of Fig. 1.4.1 is summarized in Table 2.5.1. When compared to mechanical nutation techniques, the EON has several advantages. The EON is small, rugged, lightweight, and requires low driving power. Also, because a mechanical nutator is a mechanically resonant device, it can only be driven at its resonant frequency. Further, it is difficult to precisely align the azimuth and elevation resonances in a mechanical nutator, resulting in additional complications. The EON does not suffer from these limitations. Although we operated the EON at a driving frequency of 46.8 kHz in our receiver, its bandwidth was measured¹¹ to be more than 100 kHz. Another possible advantage is the expected increased long term reliability of a non-mechanical system over a mechanically resonant one.

¹¹ The bandwidth of the EON is determined by the high frequency response of the linear electro-optic coefficients. At high frequencies the magnitude of these coefficients decreases as a secondary effect begins to counteract some of the electro-optic effect [24,25].

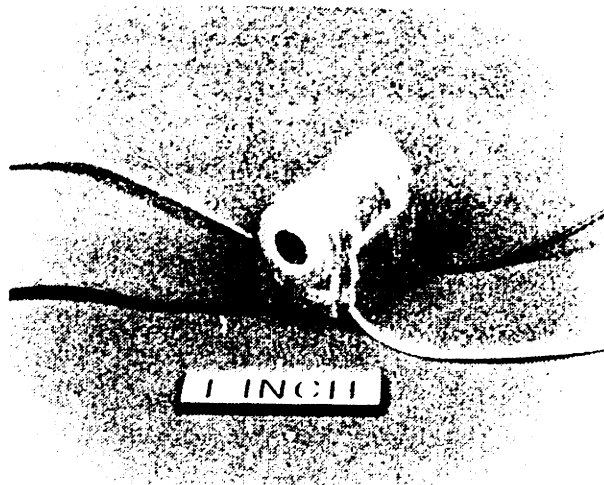


Figure 2.5.1: The electro-optic nutator.

Table 2.5.1: Electro-Optic Nutator Measured Performance

ANTI-REFLECTION COATINGS	830 nm
DEFLECTION SENSITIVITY	0.17 μ rad / V
RESPONSE BANDWIDTH	~ 100 kHz
CLEAR APERTURE	5 mm
WAVEFRONT DISTORTION AT 830nm	λ / 15 rms
WAVEFRONT DISTORTION LOSS	-1 dB
OPTICAL TRANSMISSION	93 % (-0.3 dB)
TRANSFORMER ELECTRICAL DRIVE POWER	77 mW
SIZE	0.8" x 0.5" x 0.6"

At this point, it is also worth noting that the EON's operating wavelength (830 nm) as well as its required electrical drive power (77 mW), specified in Table 2.5.1 are not fundamental limitations of the device itself. The operating wavelength of the EON is primarily determined by the anti-reflection coatings on the crystal surfaces and the half wave plate inside the device. Our measured value for electrical drive power is at the primary side of our transformer drive electronics (will be discussed in Sec. 2.7). This value is also not a fundamental limit of the device because it includes magnetic losses in the transformer drive.

The advantages of the EON do not come without a price, however. There are two unavoidable coupling losses associated with placing the EON into the signal field path: insertion loss and wavefront distortion. Insertion loss is the signal field attenuation caused by passing through the device. Since the inherent absorption of the LiTaO₃ crystals is negligible (lossless), this loss is dominated by reflections at the crystal boundaries due to imperfect anti-reflection (AR) coatings on their surfaces. The wavefront distortion ($\lambda / 15$ rms), measured by a WYKO Corp. Ladite interferometer, produces a loss which is equal to the Strehl ratio of the beam emerging from the device (0.80 Strehl \rightarrow -1 dB) and is a result of poor crystal quality. We believe a future generation EON, with narrow band AR coatings (yielding < 0.1 dB transmission loss) and crystals selected for best wave front quality (giving < 0.4 dB Strehl loss) can reduce the total coupling loss due to the insertion of the EON into the signal field path to less than 0.5 dB.

Two of the performance specifications in Table 2.5.1 merit further consideration. The justification of the EON deflection sensitivity, given as single numerical value ($0.17 \mu\text{rad}/V$) in the table, is explained in the next subsection. This is followed by a discussion of the driving electronics and the determination of the electrical drive power required by the EON.

2.6. Deflection Sensitivity of the Electro-Optic Nutator

The angular deflection provided by a single LiTaO₃ crystal was derived in Sec. 2.4 and was shown to be (Eqn. 2.4.8 repeated):

$$\theta_{ext} = \left(\frac{-Ln_e^3 r_{33}}{2R_o^2} \right) V_o \quad (2.6.1)$$

where, as before, L is the crystal length, n_e (2.180) is LiTaO₃ crystal's extraordinary index of refraction, r_{33} ($33 \cdot 10^{-12} \text{ m/V}$) is the strongest electro-optic coefficient of LiTaO₃, R_o is one-half the inner diameter of the crystal, and V_o is the voltage applied to the crystal's electrodes.

Inside the EON there are two of these crystals, deflecting perpendicular to each other. Each crystal will deflect the signal field an amount proportional to the voltage applied across it. If we define the deflection sensitivity given by Eqn. 2.6.1 as a gain K , then the angular deflection along each axis is simply proportional to the voltage applied across the crystal which deflects along that axis by K :

$$\theta_{xy} = KV_x \text{ and } \theta_{yz} = KV_z \quad (2.6.2)$$

where θ_{xy} is the deflection in the $x_d y_d$ plane of the EON and θ_{yz} is the deflection in the $y_d z_d$ plane of the EON in Fig. 2.4.5.

Since a circular nutation scan is desired, one only has to apply a sine wave voltage to one of the EON crystals and a cosine voltage, of the same amplitude and frequency, to the other. By converting to spherical coordinates in the $x_d z_d$ plane, or simply by inspection, one can see that this yields a circular nutation scan:

$$\begin{aligned}
 x_d &= KV_o \cos(\omega t) , z_d = KV_o \sin(\omega t) \\
 r &= \sqrt{x_d^2 + z_d^2} = KV_o \sqrt{\sin^2(\omega t) + \cos^2(\omega t)} = KV_o \\
 \theta &= \tan^{-1}[z_d/x_d] = \tan^{-1}[\tan(\omega t)] = \omega t
 \end{aligned}
 \tag{2.6.3}$$

where V_o is the amplitude of the sinewave and ω is its angular frequency. The radius of the circular nutation scan, r , is linearly proportional to the amplitude of the sinewave signal applied. The proportionality constant, K , was measured to be $0.17 \mu\text{rad}/V$ for our device. Thus the signal beam is circularly nutated a radius of $0.17 \mu\text{rad}$ for every volt applied to the LiTaO₃ crystals of the EON. This derivation highlights the fact that a circular nutation scan is easily achieved with the EON.

The procedure used to measure the gain of the EON is straightforward (Fig. 2.6.1). A beam is sent through the EON down an optical bench a distance d

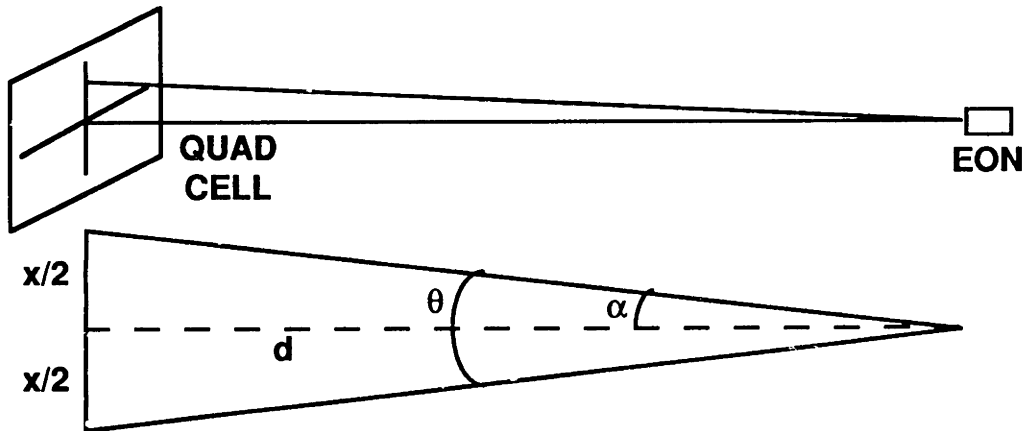


FIG 2.6.1: Measuring the EON deflection sensitivity.

onto the center of a quad-cell detector¹². The quad cell is displaced a known amount and the resulting voltage produced is used to calibrate the gain of the detector ($K_{quad} = \Delta V/\theta = d\Delta V/x$). After re-centering the beam on the quad-cell, a

¹² A quad-cell detector will produce error signals (voltages) which are proportional to the linear offset of the focused beam position from the center of the quad cell.

known voltage is applied across one of the crystals of the EON and the resulting voltage produced by the quad cell is recorded. The gain, K [radians / V], of the crystal to which the voltage was applied is

$$K = \frac{V_{quad}}{V_{eon}} \frac{1}{K_{quad}} \left[\frac{\text{radians}}{V} \right] \quad (2.6.4)$$

where V_{quad} is the voltage on the quad-cell detector, V_{eon} is the voltage applied to the EON crystal, and K_{quad} is the measured gain of the quad-cell detector. Provided that the crystals are the same size, their gains should be very similar. The gains of the two crystals within our EON were quite close ($K_{az} = 0.176 \mu\text{rad}/V$ and $K_{at} = 0.172 \mu\text{rad}/V$; equal within experimental error) and only the rounded average value is given in Table 2.5.1.

2.7. The Driving Electronics for the Electro-Optic Nutator

The measured deflection sensitivity of the EON, $0.17 \mu\text{rad}/V$ (Table 2.5.1), requires a driving voltage of 350V to achieve our desired nutation depth of 0.36 beamwidths.¹³ The EON's measured impedance is mainly capacitive ($9 - 10 pF$) throughout its entire operating bandwidth. Because we only wish to operate the EON at a single frequency, the circular nutation frequency, the driving electronics can be simply two resonant transformers, one for each deflecting crystal of the EON. There are some subtleties which must be kept in mind in the design of these transformers, which the following analysis will highlight.

We begin with a basic model of the situation:

¹³ This is the maximum nutation depth allowed while still maintaining the communication loss < 2 dB; discussed in more detail in Sec. 3.3.

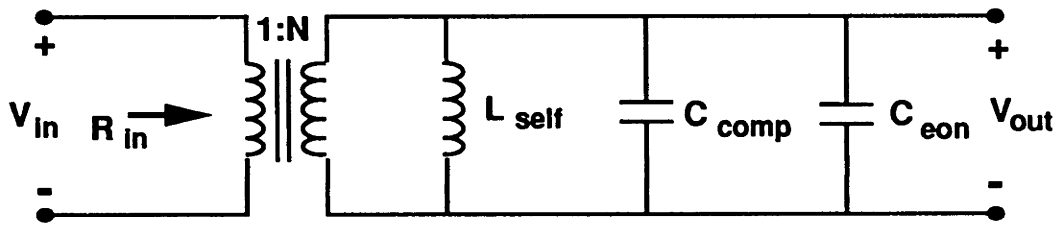


Figure 2.7.1: The EON driving electronics.

In the model shown above, L_{self} is the self-inductance of the transformer secondary windings, C_{comp} is a compensating capacitor, C_{eon} is the measured capacitance of the EON at the nutation frequency, and N is the turns ratio of the transformer. The electrical power required to drive the EON will be determined by the impedance, R_{in} , looking into the primary side of the transformer, which is simply:

$$R_{in} = \left(\frac{1}{N^2} \right) \frac{L_{self} j\omega}{1 - L_{self} (C_{comp} + C_{eon}) \omega^2} \quad (2.7.1)$$

By tuning the resonance of this impedance, $\omega = \sqrt{1/L_{self} (C_{comp} + C_{eon})}$, to the desired nutation frequency by adjusting the value of C_{comp} , R_{in} is maximized¹⁴.

To complete the design, one has to choose a transformer turns ratio N sufficient to step up standard analog voltages ($\pm 12V$ maximum) to the necessary driving voltage for the EON (~ 350 V for the device specified in Table 2.5.1). When designing the transformers for our driving electronics, we found a trade-off between the measured input resistance and the saturation of the magnetic core material. If the number of primary turns was too small, the magnetic core material was saturated. If the number of primary turns was too big, the measured input resistance was not high enough, due to the increase in L_{self} with increasing

¹⁴ According to Eqn. 2.7.1, R_{in} becomes infinite as $\omega \rightarrow \omega_{res}$. This is because the simple model of Fig. 2.7.1 is not a complete model of the situation. It does not include other parasitic circuit elements, such as the resistance of the transformer windings, which will limit the maximum value of R_{in} and give a finite Q to this resonance. The main point is that R_{in} is maximized at this resonance and including the other elements only needlessly complicates the analysis.

secondary windings. An iterative design compromise between these two constraints was eventually arrived at.

When starting from scratch, the easiest way to begin is to find the minimum number of primary turns necessary to avoid saturating the transformer core. This value is found by calculating the number of ampere-turns produced by the driving application. The number of ampere-turns is the product of the peak current sent into the transformer primary windings multiplied by the number of primary turns -- $i_{peak}n$. The peak current is the peak current sent into the transformer self inductance, found by integrating the volt-seconds applied to the primary side of the transformer, $i_{peak} = (1/L_{self}) \int_{\pi/2\omega}^{\pi/\omega} B \sin(\omega t) dt$.¹⁵ This self-inductance, L_{self} , is simply $L_{self} = A_L n^2$, where A_L is the core inductance of the magnetic material used¹⁶ and n is the number of primary turns. Putting all of these relations together yields an expression for the number of amperes turns due to an applied sinewave voltage of amplitude B :

$$\text{amp - turns} = \frac{B}{2\omega A_L n} < 10 \quad (2.7.2)$$

where the given constant 10 is the maximum number of ampere-turns that our ferrite material could handle (given in the data sheets) [36]. This expression can be solved for n to determine the minimum number of primary turns required; for our design $n > 7$.

¹⁵ I have assumed that a sinewave of amplitude B is applied (for a circular nutation scan). The limits of integration, only 1/4 of a period, might be cause for some speculation at first because one might expect to be integrating over 1/2 a period of the sinewave. For a sinewave drive in the steady state, the transformer core must have the volt-seconds applied to it be symmetric about the zero point, therefore only 1/4 of a period matters. This is because the positive part of the sinewave cycle must begin at the maximum negative volt-seconds. The first 1/2 of the positive hump of the sinewave merely brings the core back to zero volt-seconds, and the second 1/2 of the positive hump (1/4 of a period) takes it up to its maximum positive volt-seconds value.

¹⁶ This should be found from the data sheets describing the magnetic material used for the transformer core. For our material, Ferroxcube 3B9, 2213A600 ferrite, this value was $.6 \times 10^{-6}$ H / turns². The expression for L_{self} is easily modified if this value is given in other units.

After choosing a number of primary turns greater than the value specified by solving Eqn. 2.7.2, one needs to check for resistive heating in the windings and determine if R_{in} from Eqn. 2.7.1 is large enough. To check for resistive heating in the core, determine the length of wire used in both windings of the transformer and compute the power dissipated in each set of windings (I^2R), and check if this value is acceptable¹⁷. Since the simple circuit model of Fig. 2.7.1 did not include any of the elements that ultimately limit the maximum value of R_{in} , the only practical way to determine if R_{in} is large enough is to pick a transformer design that satisfies the other two constraints and measure R_{in} after the transformer is wound. If the measured value of R_{in} is not large enough, then one should iterate the design procedure.

The two transformers used to drive our EON were made from Ferroxcube 3B9, 2213A600 ferrite material pot-cores, each with 15 turns on the primary side, and 600 turns on the secondary side, yielding a turns ratio of 40:1. When the transformer resonances were tuned, by selecting C_{comp} , to our driving frequency¹⁸ of 46.8 kHz, the measured impedance of the EON, R_{in} , seen from the primary side of the transformer drive, was $> 1 \text{ k}\Omega$, an impedance which can be driven by any common op-amp. This driving point impedance stays above $500 \text{ }\Omega$ for a frequency range¹⁹ of $36.8 \text{ kHz} < f < 56.8 \text{ kHz}$; much larger than the expected drift in a typical oscillator circuit. The turns ratio of 40:1 made the voltage required to achieve our desired nutation depth at the primary side of the transformer 8.8 V, for a total power drain of only 77 mW which is easily supplied by the OP-42 driving op-amps used.

¹⁷ This value is probably not much of a concern because even if the entire 77 mW input power in our system was dissipated in the windings (which it clearly is not), there still wouldn't be any significant resistive heating.

¹⁸ The driving frequency is determined by the center of the bandpass filter in the nutation demodulation electronics (discussed in Sec. 5.1).

¹⁹ This is near the lower limit of the driving capability of many common op-amps (including the OP-42 op-amps actually used).

There are three main advantages to choosing this driving circuitry. First is the simplicity of the circuitry itself; only four components, two transformers and two compensating capacitors, are required. Second, and perhaps most important, is the fact that the transformer drive can supply the necessary ~ 350 V to the EON from any common op-amp. Finally, by stepping up to the high voltages required by the EON right at the device itself, assuming the transformers are mounted close to the EON, the high field intensity is localized to the vicinity of the EON, making any necessary shielding easier and improving system reliability.

3. Noise Equivalent Angle

As long as the tracking loop is operating within its linear region, it can be represented by the simple model shown in Fig. 3.0.1. In this simple model, the total tracking error σ_t^2 consists of two parts, σ_{uc}^2 -- the uncompensated tracking error (disturbances that the tracking loop cannot track out, due to its finite bandwidth), and the noise-equivalent-angle -- *NEA* (the tracking error induced by noise):

$$\sigma_t^2 = \sigma_{uc}^2 + NEA^2 \quad (3.0.1)$$

System performance specifications dictate that the total tracking error should be less than approximately 0.05 beamwidths.

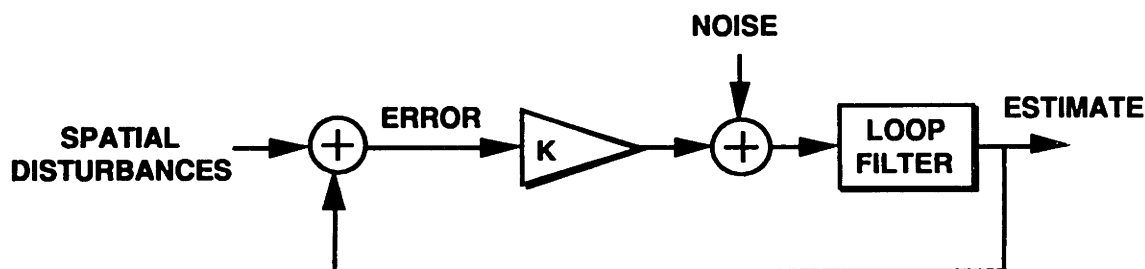


Figure 3.0.1: A simple model for a spatial tracking loop.

The total tracking error can further be expressed in terms of the noise-equivalent spectral density (*NESD*), the closed-loop transfer function of the tracking loop ($H(f)$), and the power spectrum of the environmental angular disturbances ($S_\theta(f)$):

$$\sigma_t^2 = \int_{-\infty}^{\infty} |1 - H(f)|^2 S_\theta(f) df + \frac{NESD}{2} \int_{-\infty}^{\infty} |H(f)|^2 df \quad (3.0.2)$$

From the above relation, it is evident that the closed-loop rejection of the tracking loop, $|R(f)| = |1 - H(f)|$, must be sufficient to track out the power spectrum of the angular disturbances to within the system specification for σ_{uc}^2 . The closed-loop

bandwidth of the tracking loop cannot be increased without bound, however, because this increases the noise induced tracking error through the second term of Eqn. 3.0.2. In a well designed tracking loop, an optimization on both σ_{μ}^2 and NEA is performed to determine the optimum closed-loop transfer function ($H(f)$). Irregardless of any choice of $H(f)$, the noise induced tracking error should be reduced by getting the system $NESD$ as close to its theoretical limit as possible.

This section will consider the $NESD$ of the nutating receiver shown in Fig. 1.4.1. It begins by deriving the receiver theoretical noise performance in Sec. 3.1. The measurement procedure used to test this bound is discussed in Sec. 3.2. Lastly, the measured results are compared to the theoretical predictions in Sec. 3.3.

3.1. Derivation of the Receiver NESD

In order to derive the theoretical noise performance for our nutating receiver of Fig. 1.4.1, one must consider the intended application. The receiver is designed as part of an intersatellite communication link within a larger global satellite network. As a result, this subsection will follow a communication signal through a typical intersatellite link.

Consider the situation shown in Fig. 3.1.1, which is simply Fig. 1.1.2 reproduced here for convenience. This figure illustrates an intersatellite link between a geosynchronous orbit (GEO) satellite to a low-earth orbit (LEO) satellite. As was mentioned in the introduction, the boresight, acquisition, and

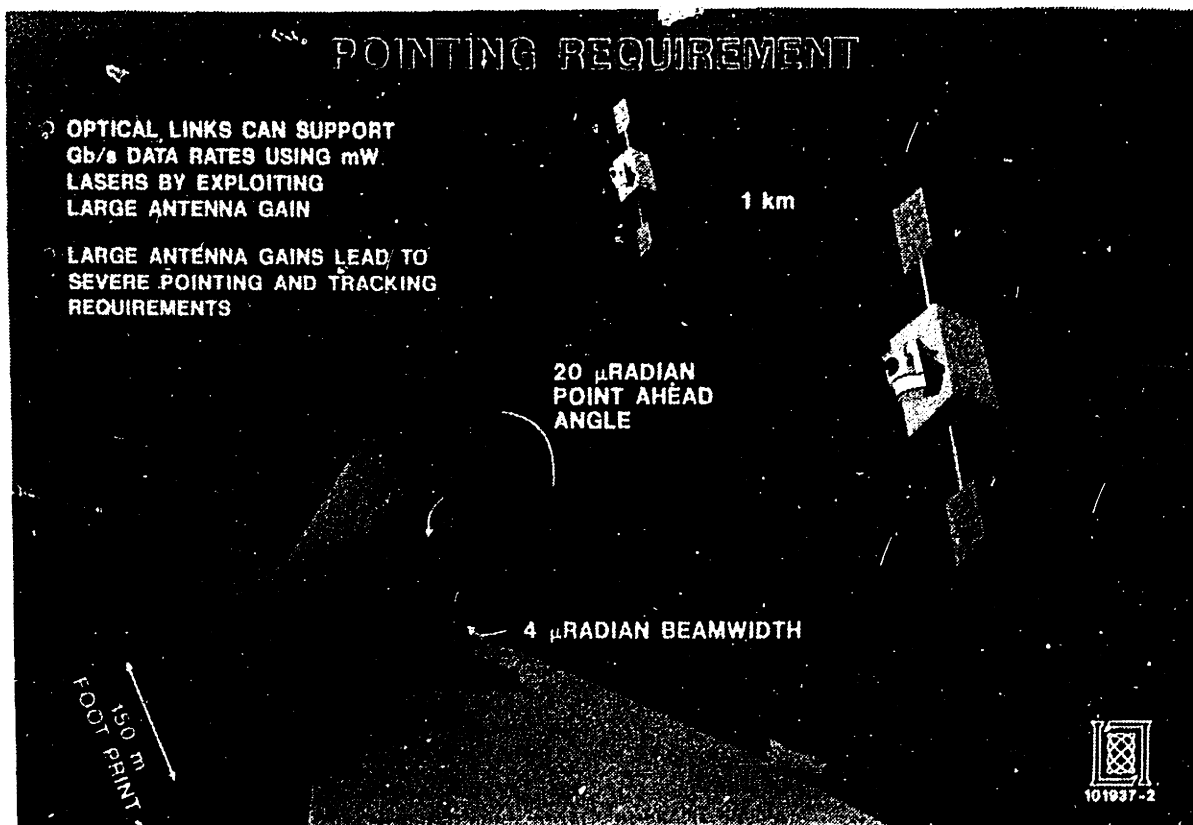


Figure 3.1.1: A typical intersatellite link.

pointing subsystems are closely tied to the operation of the tracking system in an optical communication system. However, these interactions are not the focus of this document, to examine the performance of the electro-optic nutator (EON) in a high-bandwidth spatial tracking loop, and henceforth we shall assume that initial signal acquisition has already been accomplished and active tracking has begun.

We begin by assuming that the sending satellite transmits a uniform intensity signal through a circular aperture antenna towards the receiving satellite¹:

$$U(r_1) = \sqrt{\frac{4P_s}{\pi d^2}} \text{circ}\left(\frac{r_1}{d/2}\right) \quad (3.1.1)$$

In the above expression for the signal electric field distribution -- $U(r_1)$, P_s represents the total transmitted signal power, d is the antenna aperture diameter, and r_1 is the radius component in the plane of the transmitting antenna. The $\text{circ}(x)$ function is a function whose magnitude is unity when $|x| \leq 1$ and zero otherwise. The constant term in Eqn. 3.1.1 was chosen to normalize the transmitted signal power to P_s , ($\int_0^{d/2} \int_0^{2\pi} |U(r_1)|^2 r_1 dr_1 d\theta_1 = P_s$).

The receiving satellite receives the far-field intensity distribution of Eqn. 3.1.1. To determine this distribution we employ the far-field Fraunhofer diffraction transformation [31]:

$$U(x_o, y_o) = \frac{\exp(jkz) \exp\left[j\frac{k}{2z}(x_o^2 + y_o^2)\right]}{j\lambda z} \iint_{\text{aperture}} U(x_1, y_1) \exp\left[-j\frac{2\pi}{\lambda z}(x_o x_1 + y_o y_1)\right] dx_1 dy_1 \quad (3.1.2)$$

The use of this expression is illustrated in Fig. 3.1.2. In the figure and Eqn. 3.1.2, λ is the signal wavelength, k is the wave-number ($2\pi/\lambda$), z is the propagation

¹ For a different type of antenna or transmitted signal field, Eqn. 3.1.1 must be modified accordingly and the calculation proceeds analogously.

distance, (x_1, y_1) are the coordinates in the transmitting aperture plane, and (x_o, y_o) are the coordinates in the observation plane. Converting the Fraunhofer diffraction equation to polar coordinates yields²:

$$U(r_o) = \frac{\exp[jkz] \exp\left[j \frac{kr_o^2}{2z}\right]}{j\lambda z} 2\pi \int_0^{\tilde{r}} r_1 dr_1 U(r_1) J_0\left(\frac{2\pi}{\lambda z} r_o r_1\right) \quad (3.1.3)$$

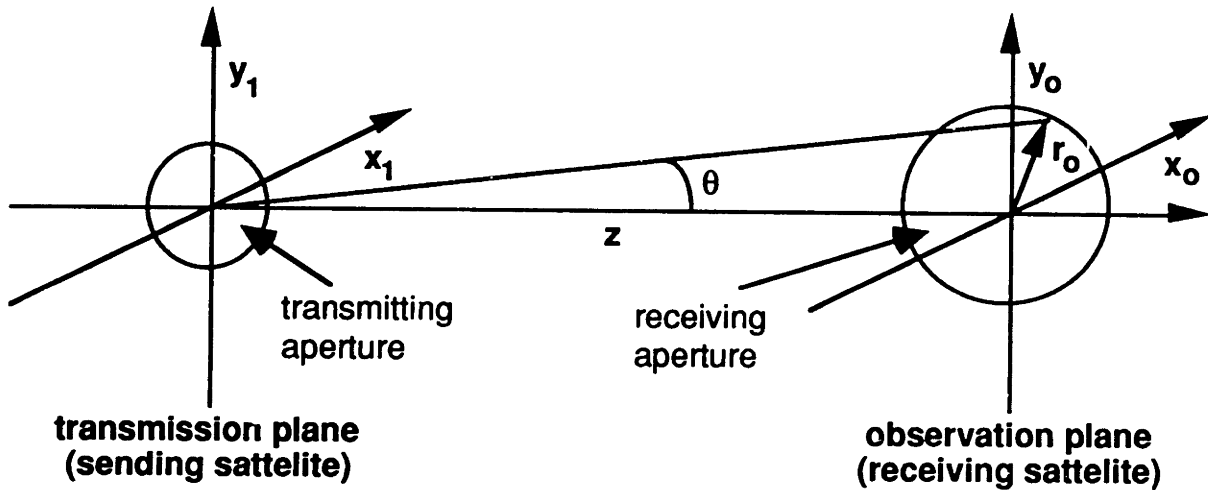


Figure 3.1.2: Use of the Fraunhofer far-field diffraction equation to transform a signal field given in an aperture coordinate plane (x_1, y_1) , to an observation plane (x_o, y_o) .

The previous expression assumes a radially symmetric (no θ dependence) signal field, where r_1 and r_o are the radius coordinates in the transmitting and observation planes respectively and $J_0(x)$ is the zero-order Bessel function of the first kind, defined as follows:

² In order to use the Fraunhofer diffraction equation, the following condition must be satisfied:

$$z \gg \frac{k(r_1)_{\max}^2}{2} \quad (F3.2)$$

In our receiver $k = 2\pi/\lambda = 7.57 \cdot 10^6 \text{ m}^{-1}$ and $r_1 \cong 0.3\text{m}$, which says that $z \gg 340\text{km}$. For typical intersatellite communication distances (i.e. $z = 22,000\text{km}$ for GEO \rightarrow LEO), this requirement is satisfied.

$$J_0(a) = \frac{1}{2\pi} \int_0^{2\pi} \exp[-ja \cos(\theta - \phi)] d\theta \quad (3.1.4)$$

Recognizing

$$2\pi \int_0^\infty U(r_1) r_1 dr_1 J_0(2\pi r_1 \rho) = B\{U(r_1)\} \quad (3.1.5)$$

as the Fourier-Bessel transform, one can see that the Fraunhofer diffraction pattern is simply the Fourier-Bessel transform of the transmitted signal field with and additional phase term:

$$U(r_o) = \frac{\exp[jkz]}{j\lambda z} \exp\left[j\frac{kr_o^2}{2z}\right] B\{U(r_1)\}_{\rho=r_o/\lambda z} \quad (3.1.6)$$

Applying the relationship to our transmitted signal field (Eqn. 3.1.1) yields:

$$U(r_o) = \exp[jkz] \exp\left[j\frac{kr_o^2}{2z}\right] \sqrt{\frac{4P_s}{\pi d^2}} \left(\frac{kd^2}{j8z}\right) \left[2\frac{J_1(kdr_o/2z)}{kdr_o/2z}\right] \quad (3.1.7)$$

Since the previous expression represents the signal field distribution in the observation (far-field) plane, the observed intensity pattern is:

$$I(r_o) = |U(r_o)|^2 = \frac{4P_s}{\pi d^2} \left(\frac{kd^2}{8z}\right) \left[2\frac{J_1(kdr_o/2z)}{kdr_o/2z}\right]^2 \quad (3.1.8)$$

After a little algebraic manipulation and the substitution $\theta = r_o/z$ (refer to Fig. 3.1.2) this intensity becomes:

$$I(r_o) = P_s \left(\frac{\pi}{4z^2}\right) \frac{1}{(\lambda/d)^2} \left[2\frac{J_1\left(\frac{\pi\theta}{\lambda/d}\right)}{\left(\frac{\pi\theta}{\lambda/d}\right)}\right]^2 \quad (3.1.9)$$

This equation is easily recognized as the Airy pattern, named after G. B. Airy who first derived it.

The radius to the first zero of Eqn. 3.1.9 is given by:

$$1.22 = \frac{\theta}{\lambda/d} \Rightarrow r_o = \frac{1.22\lambda z}{d} \quad (3.1.10)$$

For our receiver, $\lambda = 830nm$ (near infrared), $z \cong 22,000km$ (GEO to LEO), and $d \cong 0.3m$ (1ft) corresponds to a radius of $r_o = 75m$. The diameter of this full-width main lobe is shown in Fig. 3.1.1 as the 150m foot print cast on the receiving satellite by the sending satellite. Since the antenna on the receiving satellite will hog out an approximately one-foot diameter spot of this far-field pattern, we can assume that the received signal field has a uniform intensity (planar). If the original signal field differs from Eqn. 3.1.1, it should be corrected accordingly and the preceding calculations carried out with the new $U(r_1)$ to determine if the received signal field can be approximated as planar.

Now that we have shown that the incoming signal field is planar, one must consider what happens once inside the receiver itself. The front-end of our receiver using the electro-optic nutator (EON) is shown in Fig. 3.1.3. The incoming signal field is collected and focused by a low-bandwidth, large angular range telescopic instrument, such as a gimbaled telescope. This instrument forms a critical part of a nested servo tracking loop, consisting of a coarse tracking loop and a high-bandwidth tracking loop, as diagrammed in Fig. 1.3.1.

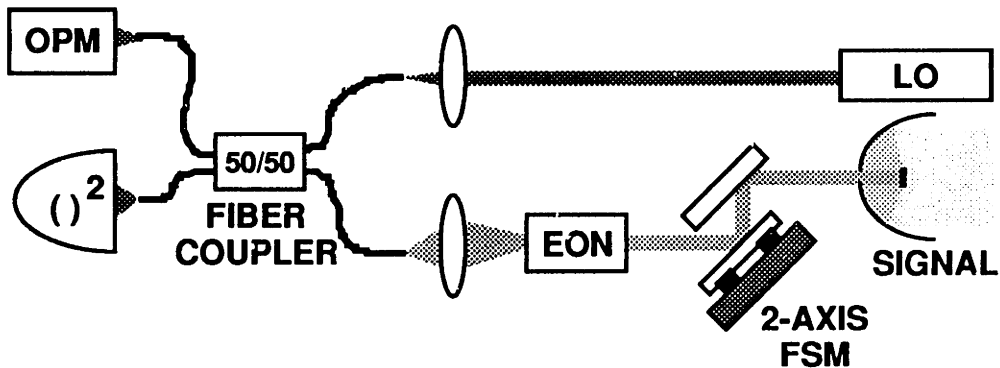


Figure 3.1.3: The front-end of our receiver using the EON.

After the far-field signal is collected and collimated by the receiving satellite's telescope it can be expressed as:

$$S(r, \theta) = \sqrt{\frac{4P_{sig}}{\pi d^2}} \text{circ}\left(\frac{r}{d/2}\right) \exp[-jkr(A \cos \theta + E \sin \theta)] \quad (3.1.11)$$

The above expression represents a normalized ($\int_0^{2\pi} \int_0^\infty |S(r, \theta)|^2 r dr d\theta = P_{sig}$) planar signal field, where P_{sig} is the total received signal power, d is the diameter of the collecting antenna's (circular) aperture, $k = 2\pi/\lambda$ is the signal wave-number, and r and θ are polar coordinates. The direction of propagation of this received signal field can, in general, be tilted in any direction. This tilt can be decomposed into two components, an azimuth (A) and elevation (E) tilt component, as shown

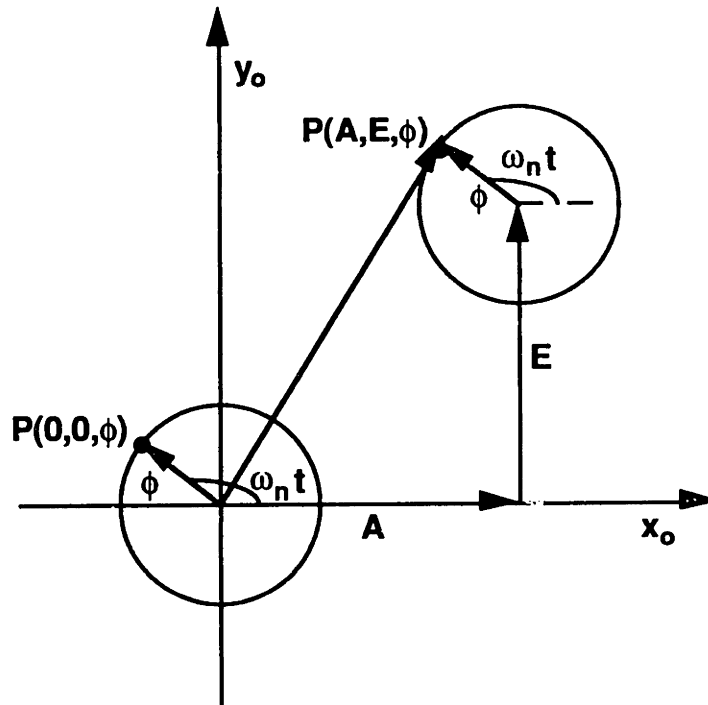


Figure 3.1.4: Two dimensional decomposition of the tilt in the received planar signal field's direction of propagation. A typical point, $P(A, E, \phi)$, is shown with azimuth and elevation tilt (A and E) and added nutation scan (nutation depth ϕ). Also shown is the same point once perfectly tracked.

in Fig. 3.1.4. Once this field is nutated by the EON, it can be represented as:

$$S_n(r, \theta) = \sqrt{\frac{4P_{sig}}{\pi d^2}} \text{circ}\left(\frac{r}{d/2}\right) \exp\left[-jkr\left\{\left[A - \phi \cos(\omega_n t)\right] \cos \theta + \left[E - \phi \sin(\omega_n t)\right] \sin \theta\right\}\right] \quad (3.1.12)$$

where ϕ is the nutation depth, ω_n is the nutation frequency, and t is time.

As shown in Fig. 3.1.3, this nutated signal field is then focused into one input of a 3 dB fiber coupler, with a strong local-oscillator (LO) beam coupled into the other, forming the basis of a heterodyne detection scheme. To complete the characterization of the receiver front-end, one must consider the coupling profile of the nutated signal field into the receiving optical fiber. It is well known that the mode profile of an optical fiber can be well approximated as a Gaussian [29]. To derive the coupling profile one can either take the expression for our nutated signal field (Eqn. 3.1.12) through the lens and overlap it with the fiber mode profile, or one can propagate the fiber mode profile back through the focusing lens and overlap it with the nutated signal field. We will take, the hopefully simpler, second approach.

The Gaussian fiber mode profile will transform to another appropriately scaled Gaussian when brought through the focusing lens. To show this, it is easiest to assume this result and begin with a normal Gaussian signal field impinging on the focusing lens of Fig. 3.1.3:

$$U(r) = \sqrt{\frac{2}{\pi\sigma}} \exp\left[-\left(\frac{r}{\sigma}\right)^2\right] \quad (3.1.13)$$

where r is the radius component in the incident plane of the focusing lens and σ is the full-width $1/e$ radius of the received signal field. The constant term in front of Eqn. 3.1.13 was chosen so that the total signal power is normalized to unity ($\int_0^{2\pi} \int_0^\infty |U(r)|^2 r dr d\theta = 1$). From this relationship, one can see that σ is also the full-width $1/e^2$ radius of the received power.

To transform the previous expression to the focal plane of the lens -- $U(r_f)$, we use a thin lens approximation and the transformation [31]:

$$U(r_f) = \frac{\exp\left[j\frac{kr_f^2}{2f}\right]}{j\lambda f} \exp[jkf] B\{U(r_f)\} \Big|_{\rho=r/\lambda f} \quad (3.1.14)$$

where λ is the signal wavelength, k is the wave-number, f is the focal length of the lens, r and r_f are the radius coordinates in the incident and focus planes of the lens respectively, and $B\{U(r_f)\} \Big|_{\rho=r/\lambda f}$ is the Fourier-Bessel transform of $U(r_f)$ evaluated at $\rho = r/\lambda f$. Performing this transformation on the Gaussian signal field of Eqn. 3.1.13 yields:

$$U(r_f) = \sqrt{\frac{2}{\pi\sigma_f^2}} \exp\left[-\left(\frac{r_f}{\sigma_f}\right)^2\right] \quad \text{where } \sigma_f = \frac{\lambda f}{\pi\sigma} \quad (3.1.15)$$

where we have ignored an additional phase term. Clearly, an incident Gaussian signal field transforms to another scaled Gaussian in the focal plane of the lens with a new full-width $1/e$ radius -- σ_f which is related to the incident full-width $1/e$ radius -- σ by a coordinate transformation.

Since we have an expression for the received signal field (Eqn. 3.1.12) and the fiber mode profile (Eqn. 3.1.13) in the incident plane of the focusing lens of Fig. 3.1.3, we can find the coupling profile by performing an overlap integral on these two fields:

$$C(A, E, \phi) = \int_0^{2\pi} d\theta \int_0^\infty U(r) S_n(r, \theta) r dr \quad (3.1.16)$$

After a bit of manipulation (shown in Appendix 2), this integral equals:

$$C(A, E, \phi) = \sqrt{2P_{sig}} \frac{d}{\sigma_0} \int_0^1 dx \exp\left[-\left(\frac{dx}{2\sigma}\right)^2\right] J_0\left[\frac{\pi x \sqrt{(A - \phi \cos(\omega_n t))^2 + (E - \phi \sin(\omega_n t))^2}}{\lambda/d}\right] \quad (3.1.17)$$

where all quantities are as previously defined and $\sigma = \lambda f / \pi \sigma_f$. Eqn. 3.1.17 is the coupling mode profile for the nutating receiver of Fig. 1.4.1, assuming a planar received signal field.

To obtain simplified expressions describing communication and tracking performance we will linearize $C(A, E, \phi)$. In particular, for small tracking errors we can take a Taylor's series expansion of $C(A, E, \phi)$; the details are in Appendix 2:

$$C(A, E, \phi) \approx C(0, 0, \phi) + A \left. \frac{\partial C(A, E, \phi)}{\partial A} \right|_{A=E=0} + E \left. \frac{\partial C(A, E, \phi)}{\partial E} \right|_{A=E=0} \quad (3.1.18)$$

Taking advantage of the symmetry between the azimuth and elevation channels, Eqn. 3.1.18 can be expressed as:

$$C(A, E, \phi) \approx \sqrt{P_{sig}} \left[m(\phi) + \sqrt{2} AK(\phi) \cos(\omega_n t) + \sqrt{2} EK(\phi) \sin(\omega_n t) \right]$$

$$\text{where } m(\phi) = \sqrt{2} \frac{d}{\sigma_0} \int_0^1 x dx \exp \left[- \left(\frac{dx}{2\sigma} \right)^2 \right] J_0 \left[\frac{\pi x \phi}{\lambda/d} \right] \quad (3.1.19)$$

$$K(\phi) = \frac{\pi \frac{d}{\sigma_0} \int_0^1 x^2 dx \exp \left[- \left(\frac{dx}{2\sigma} \right)^2 \right] J_1 \left[\frac{\pi x \phi}{\lambda/d} \right]}{\lambda/d}$$

Assuming the polarization of the signal field coupled into the optical fiber of Fig. 3.1.3 is aligned to the polarization of the local-oscillator (LO) and that the mode profiles of the signal and LO are the same, they will combine inside the optical coupler and mix at the photodetector surface. For this heterodyne detection scheme, an expression describing the normalized signal out of the balanced receiver is given by:

$$r = C(A, E, \phi) \cos(\omega_y t) + w(t) \quad (3.1.20)$$

$$S_w(f) = \frac{h\nu}{2\eta}$$

where ω_y is the heterodyne i.f. frequency, $w(t)$ is the additive uncorrelated white Gaussian shot noise, and $S_w(f)$ is the double-sided spectral density of $w(t)$,

where h is Planck's constant, ν is the optical frequency, and η is the quantum efficiency of the photodetector.

The output of the balanced receiver is normally split into two paths; one to the communication system and one to the tracking system. For small tracking errors, where the linearity of the Taylor's series expansion of Eqn. 3.1.17 is still valid, the communication performance will depend mostly on $m(\phi)$. The tracking performance is dependent on the processing used to derive angular error estimates. In an ideal nutating receiver, angular tracking error estimates would be derived from coherent demodulation at both the heterodyne i.f. frequency (ω_{if}) and the nutation frequency (ω_n). However, in most heterodyne detection schemes this is not practical due to the low received signal powers ($P_{sig} \sim \text{nW}$). One approach, used in our receiver of Fig. 1.4.1, is to incoherently detect the output of the balanced receiver using a square-law detector.

The output of our receiver is given approximately by (Eqn. 3.1.19 substituted into Eqn. 3.1.20):

$$r = \sqrt{P_{sig}} \left[m(\phi) + \sqrt{2}AK(\phi)\cos(\omega_n t) + \sqrt{2}EK(\phi)\sin(\omega_n t) \right] \cos(\omega_{if} t) + w(t) \quad (3.1.21)$$

As shown in Fig. 1.4.1, if this signal is bandpass filtered around ω_{if} , square-law detected, and bandpass filtered around ω_n , the resulting signal becomes:

$$s = P_{sig} \left[m^2(\phi) + 2m(\phi)K(\phi)A\sqrt{2}\cos(\omega_n t) + 2m(\phi)K(\phi)E\sqrt{2}\sin(\omega_n t) \right] + n(t)$$

$$S_n(f) = 4m^2(\phi)\frac{h\nu}{2\eta} + 4B_{if}\left(\frac{h\nu}{2\eta}\right)^2 \quad (3.1.22)$$

where B_{if} is the noise-equivalent bandwidth (NEB) of the i.f. bandpass filter and $S_n(f)$ is the double-sided spectral density of $n(t)$.

To derive estimates of the azimuth and elevation spatial tracking errors, s is demodulated and normalized giving:

$$\begin{aligned} \left[\frac{s\sqrt{2} \cos(\omega_n t)}{2P_{sig} m(\phi) K(\phi)} \right] &\approx A + n'(t) \\ \left[\frac{s\sqrt{2} \sin(\omega_n t)}{2P_{sig} m(\phi) K(\phi)} \right] &\approx E + n'(t) \end{aligned} \quad (3.1.23)$$

The resulting (single-sided) noise equivalent spectral density (NESD) of $n'(t)$ is given by:

$$NESD = \frac{1}{\frac{\eta P_{sig} K^2(\phi)}{h\nu}} \left[1 + \frac{\frac{B_f}{2}}{\frac{\eta P_{sig} m^2(\phi)}{h\nu}} \right] \quad (3.1.24)$$

In the preceding *NESD* expression, $K(\phi)$ is a measure of the discriminator gain and $m(\phi)$ represents the loss in received signal power due to nutation. The noise induced tracking error, or the noise equivalent angle (NEA) of Fig. 3.0.1 can be found from the NESD by:

$$NEA = \sqrt{NESD \cdot B_{track}} \quad (3.1.25)$$

where B_{track} is the NEB of the spatial tracking loop.

3.2. NESD Measurement Procedure

If the receiver of Fig. 1.4.1 achieves quantum limited operation, the expression for the system *NESD* can be recast in terms of the receiver carrier-to-noise density ratio (*CNDR*) as follows:

$$NESD = \frac{1}{\left(\frac{K^2(\phi)}{m^2(\phi)}\right) CNDR} \left[1 + \frac{1}{2} \frac{B_y}{CNDR} \right] \quad (3.2.1)$$

In the above expression, the receiver *CNDR* is the ratio of the received signal power to the spectral noise density measured after the photodetector and front-end amplification:

$$CNDR = \frac{\eta P_{sig} m^2(\phi)}{h\nu} \quad (3.2.2)$$

In the above *CNDR* expression, $\eta P_{sig} m^2(\phi)$ is the received signal power (including the detector quantum efficiency -- η and nutation loss factor -- $m^2(\phi)$) and $h\nu$ is the spectral noise density (at the receiver operating frequency ν).

To measure the system noise performance (assuming quantum limited operation) the *NESD* of the spatial tracking loop is measured as a function of the receiver *CNDR* and compared to the values predicted by Eqn. 3.2.1. With the tracking loop opened, this is essentially the procedure we followed to measure the noise performance of our receiver. A couple of key elements pertinent to this measurement are shown in the optical bench layout of Fig. 3.2.1; their importance is explained next.

Although they are shown emanating from separate sources, both the signal and local-oscillator (LO) beams were derived from a single Hitachi 30mW semiconductor laser diode in the receiver of Fig. 1.4.1. As shown in Fig. 3.2.1,

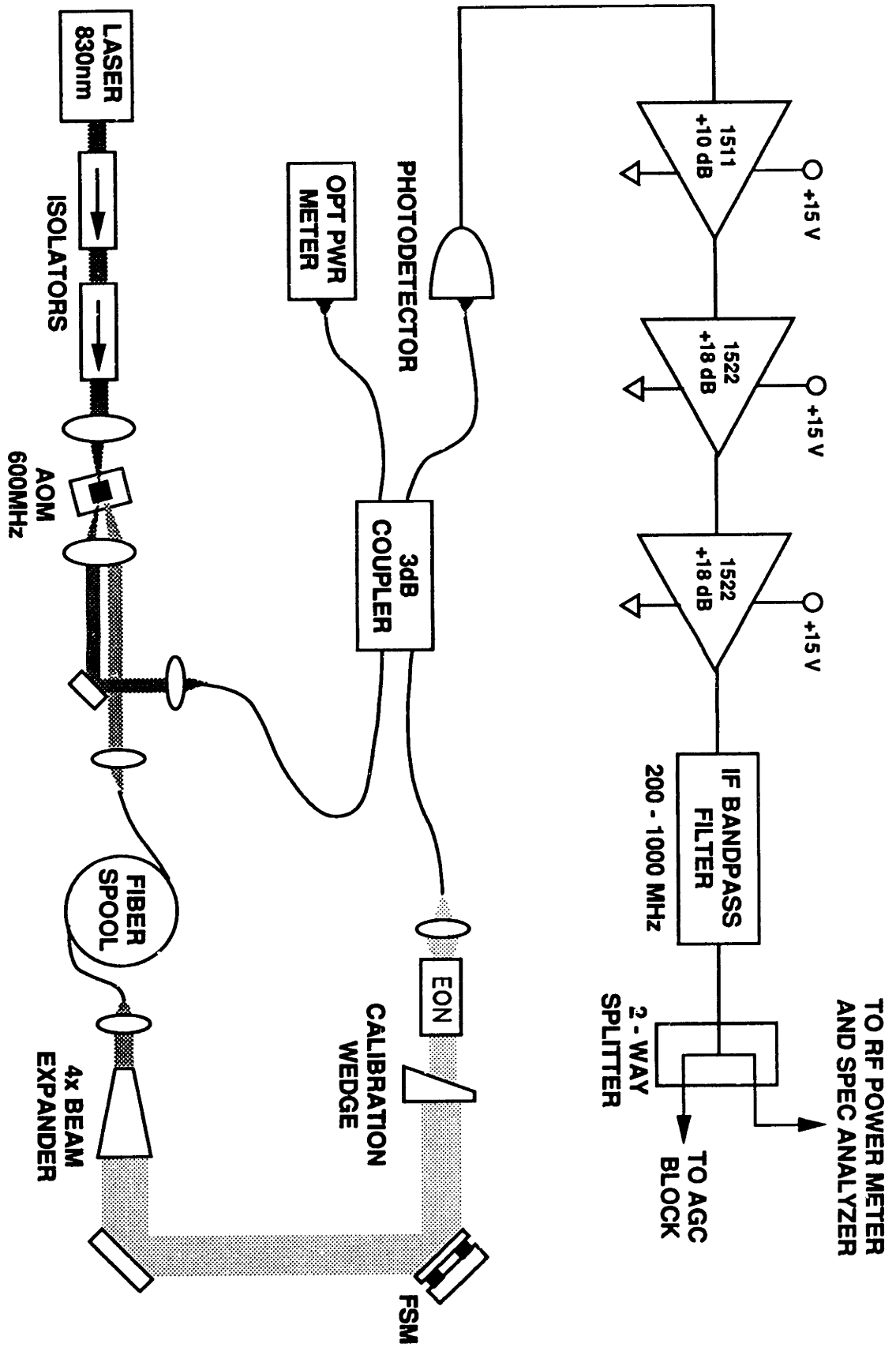


FIG 3.2.1: Optical bench layout schematic.

the output from this laser was collimated, sent through two optical isolators (with >70 dB isolation) and focused down onto an acousto-optic modulator (AOM) operating at 600 MHz, our heterodyne frequency. Inside the AOM, the incoming beam interacts with an acoustic standing wave in a Lithium Niobate crystal and an amount of optical power proportional to the acoustic signal power is frequency shifted by 600 MHz through the acousto-optic effect. This frequency shifted beam emerges from the crystal at a different angle from the rest of the incident beam, which passes through the crystal unaffected. A plot of the frequency shifted beam's intensity versus the applied acoustical signal power for the AOM of Fig. 3.2.1 is shown in Fig. 3.2.2. A detailed discussion of the acousto-optic effect is beyond the scope of this discussion; interested readers are referred to [24,25,29].

The inclusion of this device serves two purposes; it simplifies the optical setup because only a single laser source is required, as opposed to frequency locking two independent laser sources, and it provides a means of setting the receiver *CNDR* (Eqn. 3.2.2). With an appropriately strong acoustic signal applied to the AOM it is possible to diffract most of the incident light intensity into the frequency shifted beam (78% diffraction efficiency at 0.8W acoustic power for our AOM). However, in the power range in which the AOM was operated in our receiver, the frequency shifted beam was always weaker than the unaffected beam (~0.4 mW maximum frequency shifted intensity versus ~8.0 mW unshifted intensity). As a result, the frequency shifted beam was used as the weak signal beam, while the stronger, unaffected beam was used as the LO in our receiver.

The other important element used for noise performance measurements was the calibration wedge inserted in the signal beam path between the fast steering mirror (FSM) and the EON. This wedge tilted the signal field a known amount, allowing for the calibration of the angular error discriminator gain.

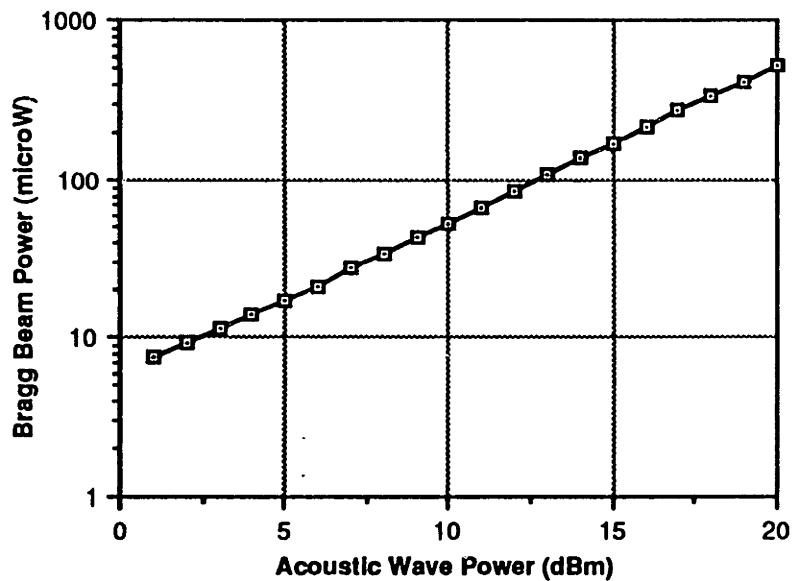


Figure 3.2.2: Frequency shifted beam intensity versus acoustical signal power for the AOM of Fig. 3.2.1.

There are two other elements in the optical bench set-up which are not shown in the receiver block diagram; the optical fiber spool and the beam expander. These two elements are included to make our laboratory receiver more like a free-space laser communications link. First, the frequency shifted signal beam from the AOM is focused into a 1 km length spool of polarization preserving York HB800 fiber, a length sufficient to decorrelate the frequency noise in the signal and LO beams to the natural 10 MHz i.f. linewidth, and its output is recollimated. This collimated output, which can be well approximated as Gaussian (verified to first order by measurement), was then sent through a 4x beam expander to make it appear as a plane wave at its center. This expanded beam was then bounced off the FSM and sent into the EON described in the previous section. The EON has a clear aperture of 5mm, which will only transmit the approximately planar (far-field) center portion of the expanded Gaussian signal field.

An automated measurement program was written to measure the receiver *NESD* as a function of receiver *CNDR*. It began by calibrating the system at an acoustic power of 1 mW (0 dBm) applied to the AOM. At this power level, a receiver signal-to-noise ratio (SNR) measurement was taken. First, a shutter blocked the signal beam and at least 10 (usually 20) measurements of the noise density in the receiver were taken with the LO beam unblocked by the RF power meter of Fig. 3.2.1. These measurements were averaged and recorded as the receiver noise power density³ -- P_N . The signal beam was then unblocked and a measure of the signal and noise power density -- P_{S+N} was taken, again with several averages. From these two measurements, the receiver reference *CNDR* was found as:

$$CNDR_{ref} = \frac{P_{S+N} - P_N}{P_N} B_{if} \quad (3.2.3)$$

where B_{if} is the NEB of the heterodyne i.f. bandpass filter (736.5 MHz). The multiplication by B_{if} is necessary to convert the measured signal-to-noise ratio (SNR) -- $P_{S+N} - P_N / P_N$ to the receiver carrier-to-noise density ratio (CNDR)⁴.

From this reference receiver *CNDR* measurement -- $CNDR_{ref}$, the AOM acoustic wave power necessary to set the receiver *CNDR* to a specific value ($CNDR_{set}$) was determined by using the following relationship:

$$AOP[dBm] = CNDR_{set}[dB] - CNDR_{ref}[dBm] + AOP_{ref}[dBm] \quad (3.2.4)$$

where *AOP* is the required AOM acoustic signal power and AOP_{ref} is the reference acoustic signal power (set to 0 dBm). To convince oneself that this relationship is valid, reconsider the linear relationship of the frequency shifted (signal) beam power to the applied acoustical signal power for our AOM shown in Fig. 3.2.2. Clearly, if the 600 MHz acoustical signal power is increased by 1

³ The receiver noise power density dominates the noise in the RF power meter.

⁴ For more information, refer to Sec. 4.2.

dB from a known reference point, AOP_{ref} , then the receiver $CNDR$ ($CNDR_{sat}$), which is proportional to the ratio of the signal to the noise power density, will be increased by 1 dB also.

After the calibration, the receiver $NESD$ is measured at evenly spaced intervals between 70 and 110 dB receiver $CNDR$. This measurement is a two step process. First, an averaged measurement of the $NESD$ in the power spectrum of the azimuth and elevation tracking error signals is taken by a Hewlett Packard Dynamic Signal Analyzer⁵. As long as the power spectrum of the tracking error signals is flat, this measurement can be taken anywhere. The power spectrum of the error signals in our receiver was flat beyond 10 kHz, so the measurement⁶ was taken at 1 kHz (chosen somewhat arbitrarily)⁷. Conveniently, the analyzer can take this measurement directly, in units of V^2/Hz .

After each $NESD$ measurement is taken, the calibration wedge is mechanically spun at a constant rate in the signal beam's path. This produces a circular deflection pattern that is visible in the power spectrum of the azimuth and elevation error signals. This circular scan imparted on the signal beam by the spinning calibration wedge is similar in effect to the nutation signal imposed on the signal beam by the EON (as suggested in Sec. 2), but at a much lower frequency (~20 Hz). This produces a sinusoidal voltage in the tracking error signals at the wedge spinning frequency. Because the amount of beam deflection caused by the spinning wedge is known, a measurement of the discriminator gain for each of the tracking loops can be found from this voltage:

⁵ Recall that the tracking error estimates from the demodulator are sent into the Dynamic Signal Analyzer.

⁶ In this measurement, the noise floor of the instrument was always at least 10 dB below the $NESD$ of the error signals throughout the entire measured receiver $CNDR$ range (70 - 110 dB).

⁷ Remember that the spatial tracking loops are open during this measurement; so the compensator, which is designed to give a 1 kHz open-loop crossover frequency, is not included in the loop. Also, the signal beam is centered on the fiber before this entire measurement procedure begins.

$$K \left[\frac{V}{BW} \right] = \frac{2V_{wedge}[V]BW[\mu\text{rad}/BW]}{\theta_{wedge}[\mu\text{rad}]} \quad (3.2.5)$$

where V_{wedge} is the measured voltage produced in the error signals power spectrum by the wedge, BW is the signal field beamwidth ($\lambda/d = 166[\mu\text{rad}/BW]$ for our receiver), and θ_{wedge} is the full-width (circular) deflection angle of the wedge.

With the discriminator gain value from the calibration wedge (Eqn. 3.2.5) and the *NESD* measurement taken by the Dynamic Signal Analyzer, a calibrated measure of the receiver *NESD* can be found:

$$\frac{NESD[V^2/\text{Hz}]}{K^2[V/BW]^2} = NESD[BW^2/\text{Hz}] \quad (3.2.6)$$

This *NESD* measurement is more useful than the previous measurement in units of V^2/Hz because the noise induced tracking error is now calibrated as a function of receiver signal beamwidth. From this *NESD* value, the noise induced tracking error (*NEA*) can be found in terms of beamwidths by using the relationship:

$$NEA[BW] = \sqrt{NESD[BW^2/\text{Hz}] \cdot B_{track}[\text{Hz}]} \quad (3.2.7)$$

where B_{track} is the noise-equivalent-bandwidth (*NEB*) of the receiver spatial tracking loop in Fig. 1.4.1.

The receiver *CNDR* calibration performed by Eqn. 3.2.4 assumed that our receiver was achieving quantum limited operation. As a final step in the receiver *NESD* measurement, the accuracy of this calibration is tested. A measurement of the actual receiver *CNDR* was taken at each measurement point and compared to the "set" (assumed) *CNDR* value given by Eqn. 3.2.4. Accuracy to within 0.5 dB throughout the entire measured *CNDR* range of 70 - 110 dB was consistently achieved.

After all of the measurement points are taken, all that remains is to plot the measured *NESD* values compared to those predicted by Eqn. 3.2.1. This measurement procedure was performed for our receiver shown in Fig. 1.4.1 and the results are discussed in the next subsection.

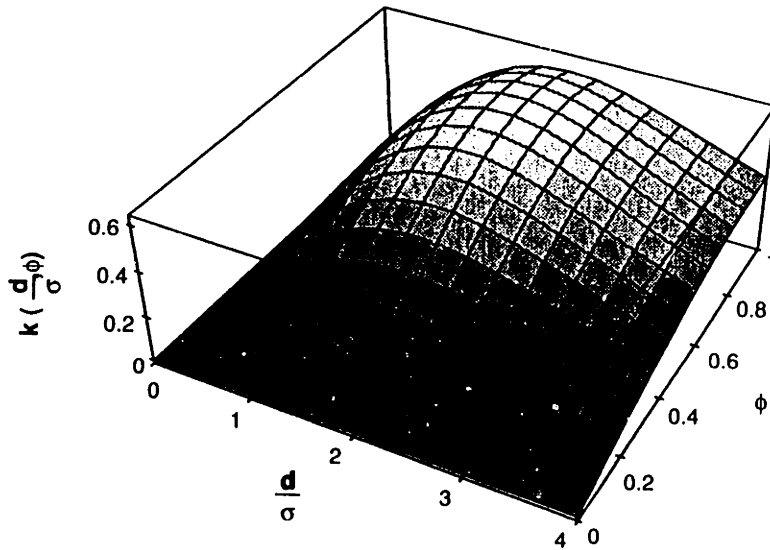
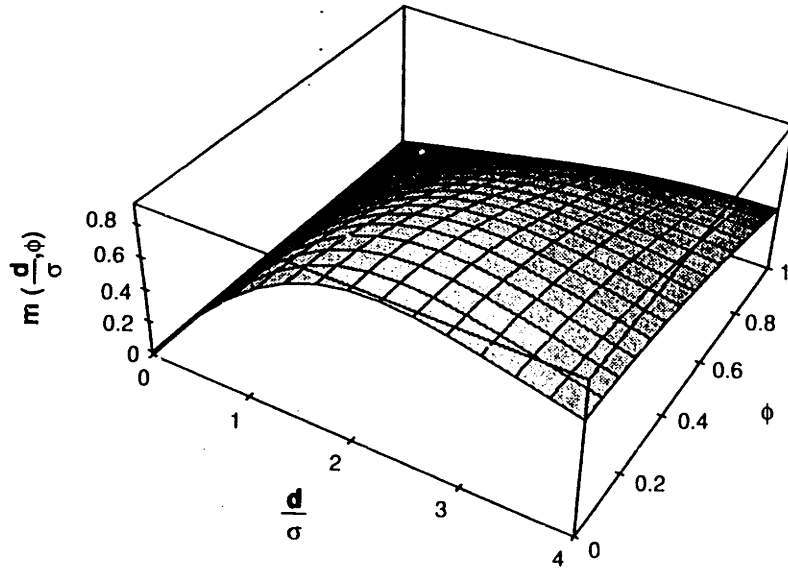
3.3. NESD Measurements

To measure the *NESD* for our receiver, one must first choose an operating value for the mode-mismatch loss -- $m(d/\sigma, \phi)$, and the discriminator gain -- $K(d/\sigma, \phi)$. Both of these factors are plotted in Fig. 3.3.1. From the figure, the minimum mode-mismatch loss⁸ of -0.9 dB ($m(d/\sigma, \phi) = 0.812$) occurs with no nutation ($\phi = 0$) and $d/\sigma = 2.2$. This loss is due to the fundamental mode mismatch between the received planar signal field and the Gaussian fiber mode profile. Nutating the signal field further increases this loss. For nutation depths < 0.5 beamwidths ($\phi < 0.5$), the optimum value of d/σ remains near 2.2. The discriminator gain profile -- $K(d/\sigma, \phi)$ generally increases with increasing nutation depth, reaching its maximum value when $d/\sigma = 1.92$ and $\phi = 0.78$.

When choosing an operating point for these two values within a nutating receiver, a trade-off between communication and tracking performance must be considered. The communication performance of the nutating receiver improves with increasing amounts of received signal power⁹. Because $m(d/\sigma, \phi)$ represents the amount of received signal power, communication performance is maximized by nutating as little as possible. On the other hand, increasing the

⁸ In the upcoming expressions, remember that d is the diameter of the received signal field (the 5mm aperture diameter of our EON) and σ is the effective 1/e radius of the signal field after passing through the focusing lens (both from Eqn. 3.1.12).

⁹ A detailed discussion of the communication performance for a nutating receiver is beyond the scope of this text; interested readers are referred to [3,9,20].



193936-4

Figure 3.3.1: Mode-mismatch loss -- $m(d/\sigma, \phi)$ and discriminator gain -- $K(d/\sigma, \phi)$ for the nutating receiver.

nutaton depth ϕ increases the tracking error sensor, discriminator gain $K(d/\sigma, \phi)$, which improves the receiver spatial tracking performance. We chose our nutaton depth of $\phi = 0.36$ beamwidths as the maximum nutaton depth allowed while still maintaining a communication loss of less than 2 dB ($m(2.2, 0.36) = 0.79$). Our measured d/σ was 1.8, slightly less than 2.2, yielding $m(1.8, 0.36) = 0.76$ and $K(1.8, 0.36) = 0.43$ for our receiver of Fig. 1.4.1.

The *NESD* performance for our receiver, shown for the azimuth channel in Fig. 3.3.2, is in good agreement with theory (Eqn. 3.1.24) from 70 to about 95dB receiver *CNDR*, where the measured performance levels off. The slight shift to the right at low receiver *CNDR* is probably a result of the wavefront distortion in the EON. The measured *NESD* levels off around $-47\text{dB BW}/\sqrt{\text{Hz}}$, corresponding to a noise-equivalent-angle of 0.0012 beamwidths (Eqn. 3.2.7 with a typical 1 kHz crossover tracking loop NEB of 3.5 kHz). This floor is due to post detection noise in the demodulation electronics, but is of no consequence due to its small value¹⁰. It should be noted that this plot is typical of our best results; many things can make the system perform worse. For instance, a deterioration of the coupling into the signal fiber causes a degradation in performance.

Before leaving this discussion, it should be noted that the noise performance of the elevation spatial tracking loop was consistently ~ 1 dB worse than that of the azimuth tracking loop. However, this difference is attributable to the slight ellipticity in the mode profile of the polarization preserving fibers used in our receiver, because increasing the gain of the elevation channel by an amount equal to compensate for the ellipticity corrected the problem.

¹⁰ We experimentally determined that our receiver is limited by noise in the demodulation electronics at high receiver *CNDR* values. Since no particular effort was made to design low noise electronics in our laboratory receiver, their inclusion could possibly extend the receiver's upper noise performance bound further. However, this is not necessary because the total receiver tracking error (Eqn. 3.0.1) is dominated by the uncompensated tracking error in this high receiver *CNDR* regime.

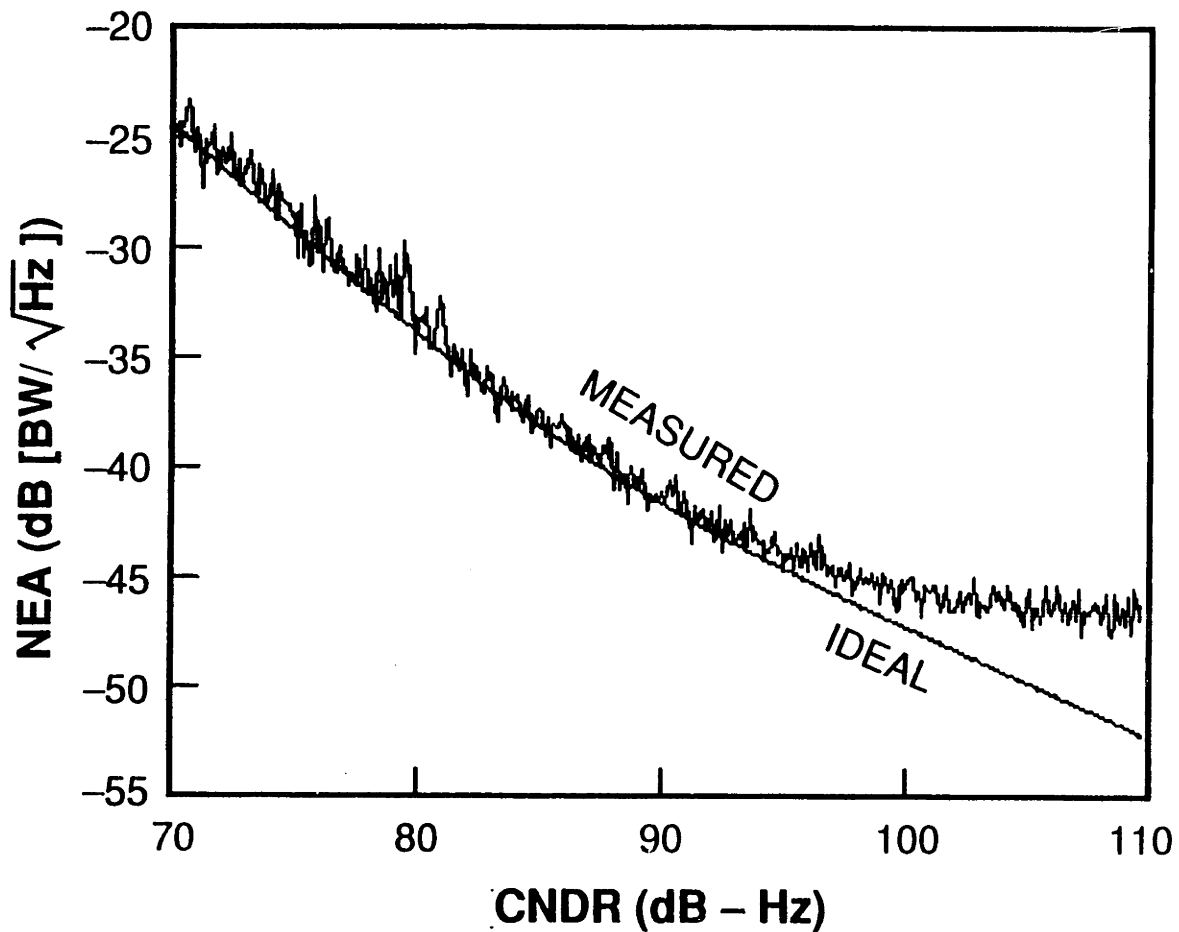


Figure 3.3.2: Nutating receiver Noise-Equivalent-Angle. The receiver measured noise performance agrees well with theory until ~95 dB receiver carrier-to-noise density ratio (CNDR). The noise floor at high receiver CNDR is due to post detection noise in the demodulation electronics and is not a limit of the nutating receiver.

4. Automatic Gain Control

The signal present at the photodetector in our receiver of Fig 1.4.1 (reproduced as Fig. 4.0.1) was found in Sec. 3 (Eqn. 3.1.21) to be:

$$r = \sqrt{P_{sig}} \left[m(\phi) + \sqrt{2}AK(\phi)\cos(\omega_n t) + \sqrt{2}EK(\phi)\sin(\omega_n t) \right] \cos(\omega_y t) + w(t) \quad (4.0.1)$$

where ω_n is the nutation frequency, ω_y is the heterodyne operating frequency, A and E are the azimuth and elevation tracking errors respectively, $m(\phi)$ is the coupling loss due to nutation, $K(\phi)$ is the discriminator gain, and $w(t)$ is the noise in the received signal. The received signal power is typically too low to synchronously demodulate this signal at the heterodyne frequency. Instead, this signal is bandpass filtered around ω_y and square law detected before being sent to the demodulator. Ignoring the double frequency terms and a small dc term proportional to the tracking errors, the resulting signal becomes (see Eqn. 3.1.22):

$$s = P_{sig} \left[m^2(\phi) + 2m(\phi)K(\phi)A\sqrt{2}\cos(\omega_n t) + 2m(\phi)K(\phi)E\sqrt{2}\sin(\omega_n t) \right] + n(t) \quad (4.0.2)$$

To derive an estimate of the azimuth or elevation tracking error s is bandpass filtered around the nutation frequency and synchronously demodulated with the nutation signals driving the EON. This process yields the noise performance shown in Sec. 3 and spatial tracking error estimates which are used in the receiver spatial tracking loop discussed in Sec. 5.

For the azimuth spatial tracking loop (similarly in elevation), synchronous demodulation of s with $\sqrt{2}\cos(\omega_n t)$ and low-pass filtering the result (done by the tracking loop compensation) to remove the double frequency terms gives:

$$\hat{A} = 4m(\phi)K(\phi)P_{sig}A \quad (4.0.3)$$

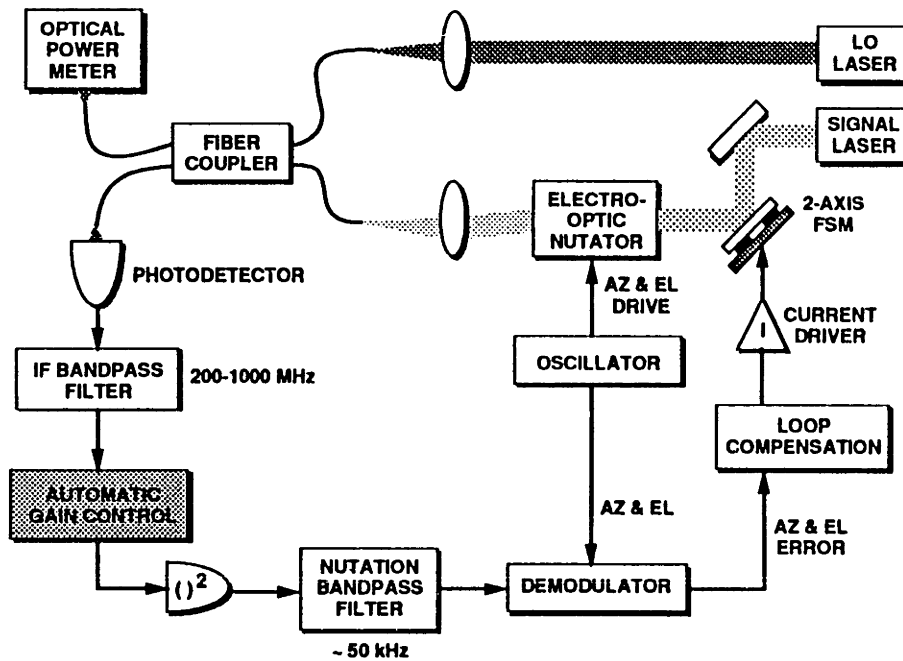


Figure 4.0.1: Nutating Receiver Block Diagram

as an estimate of the azimuth spatial tracking error (A). This estimate is proportional to the received signal power -- P_{sig} . While a dependence upon received signal power does not impede the basic operation of the servo tracking loop, it merely makes the feedback loop respond faster or slower to disturbances, it can affect the stability of the feedback path by modulating the tracking closed-loop gain. An increase in received signal power greater than the originally designed gain margin of the tracking loop would force the feedback loop into instability. Conversely, decreases in received signal power, and corresponding decreases in tracking loop gain, while not normally able to cause instability in a well-designed servo loop (except in conditionally stable systems), will compromise the bandwidth of the tracking loop. Decreasing the tracking loop bandwidth may reduce the system tracking rejection to a point where a loss of lock on the signal beam can occur.

Clearly, some form of automatic gain control (AGC) is necessary to normalize out the fluctuations in received signal power. In fact the existence of an AGC method is tacitly assumed in order to extract the azimuth, tracking error estimate from Eqn. 4.0.3. Returning to our general receiver block diagram (Fig. 4.0.1), we need to analyze the requirements for the design of the AGC (shaded) block of the figure. This section will begin with a discussion of a previously tried AGC method relying upon probing the tracking loop at its phase crossover. Then the AGC method which was used in the receiver of Fig. 4.0.1 will be discussed and its performance evaluated.

4.1. Probing the Tracking Loop at Phase Crossover

Similar to the nutation concept used for spatial tracking, this AGC method uses a small signal injected into the feedback tracking loop to obtain an estimate of the closed-loop gain. A block diagram highlighting the essential features of this method is shown in Fig. 4.1.1. The figure shows the application of the phase crossover method to a typical one axis tracking loop. This tracking loop consists of a power amplifier, a fast-steering mirror (FSM), a quad cell detector, and feedback loop compensation. Although this procedure for AGC was tested with the one-axis, quad-cell tracking loop shown in the figure, the concept is readily applied to a two axis nutation tracker with only minor modifications. Attached to the tracking loop is the AGC control loop, which is fed into the variable gain element (marked with a "G") for gain normalization.

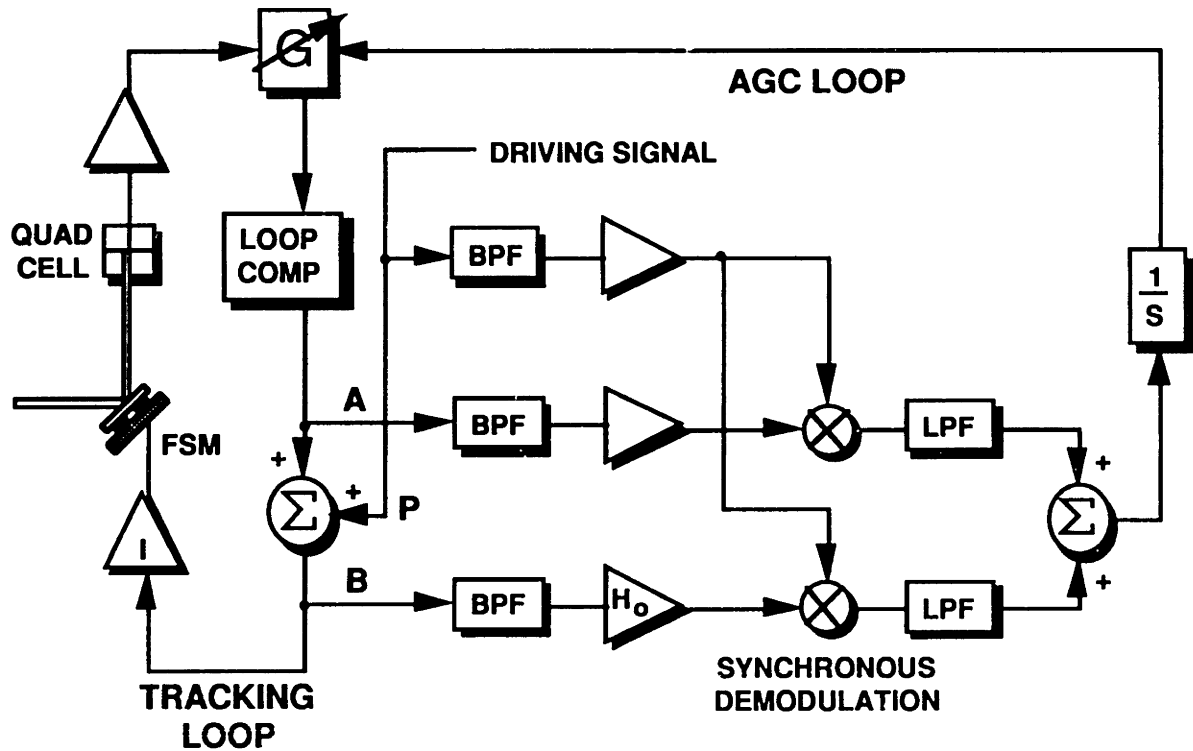


Figure 4.1.1: AGC by probing at phase crossover.

The AGC function is performed by injecting a small amplitude sinusoidal disturbance signal into the tracking loop at a unity gain summation point¹ (point P of the figure). If we probe the tracking loop immediately before the injected signal (point A), the transfer function observed is:

$$\frac{A}{P} = \frac{GH}{1+GH} \quad (4.1.1)$$

where G is the present gain of the variable gain element and H is the combined transfer function of the rest of the tracking loop. The signal present just after the injected signal (point B) is:

$$\frac{B}{P} = \frac{1}{1+GH} \quad (4.1.2)$$

¹ This unity gain summation point can be created by placing a simple summing op-amp into the electrical signal path.

These two signals can be used to obtain an estimate of the closed-loop gain of the tracking loop provided that the injected disturbance is at the phase crossover frequency of the tracking loop.

If the driving signal (injected at point P) is at the phase crossover frequency of the tracking loop, the observed signals at points A and B will differ by a sign and a scale factor H_o (they will be 180° out of phase). This is because the phase of the combined loop dynamics, denoted by the parameter H in Eqns. 4.1.1 and 4.1.2, is -180° (equivalent to a sign change) at the loop phase crossover frequency. As shown in the figure, these two signals are then bandpass filtered, amplified, synchronously demodulated, and their low frequency (essentially dc) components are summed together. The dc signal which is present after the summation is performed is proportional to:

$$dc \propto \frac{H_o - GH}{1 - GH} \quad (4.1.3)$$

where H_o is the nominal (desired) gain of the tracking loop and the two negative signs are due to the 180° of phase shift from H at the phase crossover frequency.

To see how this signal can provide AGC, consider fluctuations in received signal power. If the received signal power increases from its nominal value, the magnitude of the combined loop dynamics will increase and the numerator of Eqn. 4.1.3 will become negative because the condition $|GH| > H_o$ is satisfied. Conversely, if the received signal power decreases, the numerator will become positive because the alternate condition, $|GH| < H_o$, is met. For the unlikely case where the received signal power remains constant at its nominal value, the dc signal given by Eqn. 4.1.3 should be zero. If this signal is sent to an integrator, initially charged to a specific voltage corresponding to H_o (the nominal gain), then the voltage present after the integrator can be used as a control voltage for

AGC. This is exactly the approach shown in Fig. 4.1.1, where the control voltage is used as one input to a voltage controlled gain element² which ultimately gives the AGC function.

Unfortunately, this system did not offer very good AGC performance. From Eqn. 4.1.3, the error signal applied to the integrator of Fig. 4.1.1 controlling the variable gain element is nonlinear. This non linearity itself is undesirable, but not entirely detrimental. The system was, however, too sensitive to sudden (step) changes in received optical intensity. A relatively small step change³ in received optical signal power would force the voltage on the integrator negative, changing the sign of the tracking loop feedback and cause the tracking loop to lose lock. As a result of this and several other problems, this method of gain normalization, although interesting, was deemed too unstable to be useful.

² The gain element used was a multiplier. The voltage on the integrator was initialized to unity so that the gain of the tracking loop was unaffected when the received signal power was the desired nominal amount.

³ I readily acknowledge that this is a vague statement, but since this AGC method was rejected quantitative results are not necessary. This method is only presented here as interesting in the fact that it was able to work at all.

4.2. AGC by Noise Power Cancellation

The automatic gain control (AGC) method employed in the receiver of Fig. 4.0.1 is shown in Fig. 4.2.1. Rather than a gain control method in the strictest sense, this AGC method is actually a signal power estimator. This estimate of the received signal power is then used to maintain the received signal power at a constant level, via a voltage controlled attenuator. This subsection will determine the theoretical performance of this AGC method.

We begin with a small signal model for the front-end of a typical optical communication receiver (Fig. 4.2.2). The small signal model for the photodiode has four components; a current source producing a small signal proportional to the received signal power ($I \propto P_{sig}$), the junction capacitance of the photodiode p-n junction (C_j), the shunt resistance of the p-n junction (R_j), and the series resistance of the photodiode (R_s). Any front-end amplification can be modeled as a high impedance amplifier with a small-signal input resistance (R_{in}) and output resistance (R_{out}).

In any photodiode of practical use as an optical detector, the shunt resistance of the diode junction is large compared to its series resistance and the input resistance of the front-end amplifier ($R_j \gg R_s, R_{in}$). Also the junction series resistance is typically negligible compared to the input resistance of the front-end amplifier ($R_s \ll R_{in}$). Ignoring the limitations on the high frequency response of the photodiode due to the finite transit time of charge carriers across the junction depletion layer and the finite diffusion time of charge carriers in the p and n regions, the main limitation upon the high frequency response of the photodiode is the time constant formed by the diode junction capacitance and the input resistance of the amplifier:

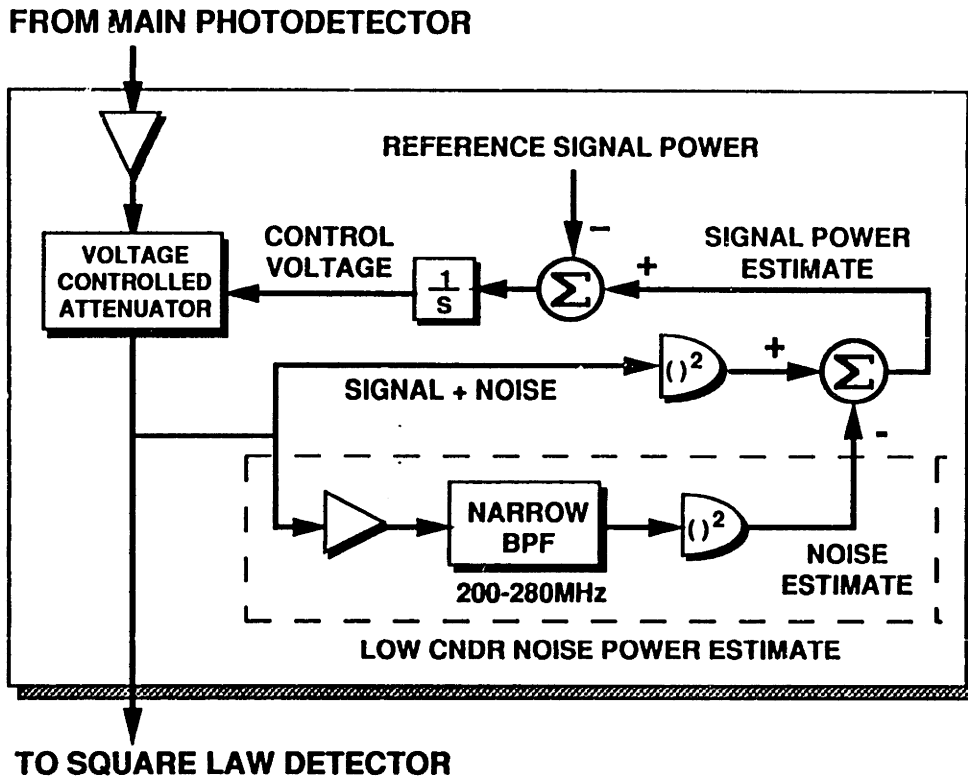


Figure 4.2.1: Automatic gain control using a signal power estimate obtained by low receiver i.f. signal-to-noise ratio (SNR) noise power cancellation.

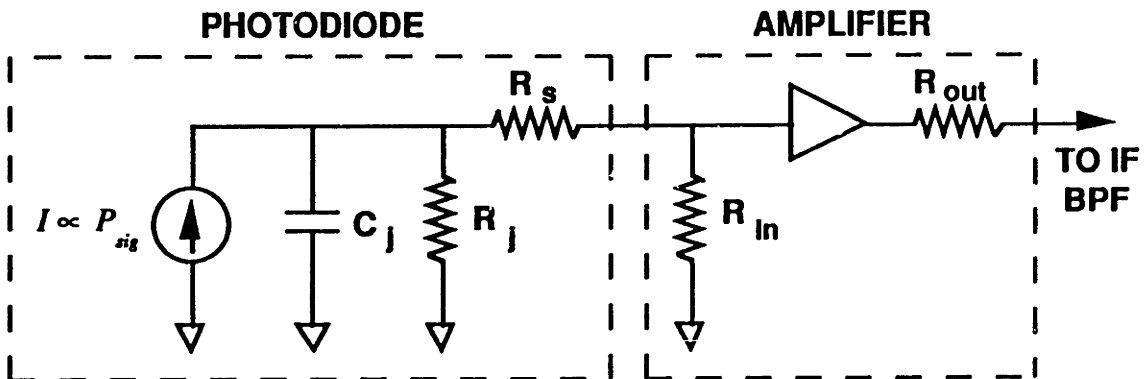


Figure 4.2.2: Model for an optical communication receiver front-end.

$$\omega_{\max} \equiv \frac{1}{R_{in} C_j} \quad (4.2.1)$$

Consider the current produced by the photodetector in the model of Fig. 4.0.1. In a heterodyne receiver⁴, the signal field is mixed with a strong local-oscillator field at the photodetector. Assuming that the spatial modes of the signal and local-oscillator fields are aligned, the photodetector current (the current source of Fig. 4.2.2) will be proportional to the optical power impinging on the device:

$$I = \frac{e\eta}{h\nu} \left[P_{lo} + m(\phi)P_{sig} + 2m(\phi)\sqrt{P_{lo}P_{sig}} \cos(\omega t) \right] \quad (4.2.2)$$

In the above expression, P_{lo} is the local-oscillator (LO) power, P_{sig} is the signal power, $m(\phi)$ is the nutation loss factor (discussed in Sec. 3.2), η is the photodetector quantum efficiency, e is the electron charge, h is Planck's constant, and ν is the LO frequency. The quantity $e\eta/h\nu$ is the responsivity of the photodiode in units of A/W.

There are three main sources of noise in the photodetection scheme of Fig. 4.2.2, shot noise due to the discrete nature of the photodetection process, Johnson noise in the load resistor⁵, and excess noise created by the front-end amplification. If the signal after the front-end amplifier of Fig. 4.2.2 is passed through a bandpass filter centered around the heterodyne frequency, with a noise-equivalent-bandwidth (NEB)⁶ of B_{if} hertz, then the total shot noise power

⁴ Although this derivation will proceed assuming heterodyne detection is used, the results are easily extended to a direct detection system. In a direct detection receiver, the derivation proceeds analogously, beginning with $I = \frac{e\eta}{h\nu} P_{sig}$ in Eqn. 4.2.2.

⁵ Johnson noise is present in all of the resistors of Fig. 4.2.2, but remember we are considering the case where the load resistor can be approximated as just the input resistance of the front-end amplifier -- $R_j \gg R_{in} \gg R_s$ (Eqn. 4.2.1).

⁶ The noise-equivalent-bandwidth of a filter is the effective bandwidth of the passband when noise is concerned. It can be found through the integration

can be found using the spectral density function, $S(f)\Delta f = 2e\bar{I}\Delta f$ [23,29]. This shot noise power is due to the dc components of the diode signal current:

$$\overline{i_{NS1}^2} = 2e \left[(P_{lo} + m^2(\phi)P_{sig}) \frac{e\eta}{h\nu} \right] B_{if} \quad (4.2.3)$$

and to the photodetector dark current⁷ -- i_d , and the background signal intensity on the detector -- P_b :

$$\overline{i_{NS2}^2} = 2e \left[i_d + P_b \frac{e\eta}{h\nu} \right] B_{if} \quad (4.2.4)$$

The total shot noise power is the sum of these two equations:

$$\overline{i_{NS}^2} = 2e \left[i_d + (P_{lo} + m^2(\phi)P_{sig} + P_b) \frac{e\eta}{h\nu} \right] B_{if} \quad (4.2.5)$$

The second major component of noise in the photodetection process is Johnson noise in the load resistor (R_{in} of Fig. 4.2.2). This noise power can be represented⁸ by either placing a voltage source in series with the load resistor, with a mean-square voltage amplitude of $\overline{v_N^2}(f) = 4kTR_{load}\Delta f$, or by placing a current source in parallel with the load resistor having a mean-square amplitude of $\overline{i_N^2}(f) = 4kT\Delta f/R_{load}$. For convenience and consistency with the other noise sources in the model, we choose the noise source due to Johnson noise as a current source, its mean-squared amplitude being

$$\Delta f_{NEB} = \frac{1}{A_{v_o}^2} \int_0^{\infty} |A_v^2(f)| df \quad (F4.6)$$

where $A_v(f)$ is the voltage frequency response and A_{v_o} is the maximum midband voltage gain of the filter being measured [35].

⁷ Photodetector dark current is the current which flows in the photodetector in the absence of any optical signal; the reverse saturation current in the diode current expression.

⁸ A derivation of this relationship can be found in Yariv [29]. A derivation from the statistical properties of a Poisson process, of which photodetection is an example, can also be found in Papoulis and Yariv [23,29].

$$\overline{i_{NR}^2} = \frac{4kT_o}{R_{in}} B_{if} \quad (4.2.6)$$

where k is Boltzman's constant, T_o is the temperature in Kelvin degrees, R_{in} is the input resistance of the front-end amplifier of Fig. 4.2.2, and B_{if} is the NEB of the heterodyne bandpass filter.

Following the example of Yariv⁹ we can include the noise power due to any front-end amplification in the term for the load resistor Johnson noise. This is done by choosing a temperature in Eqn. 4.2.6 sufficient to account for both the noise in the load resistor and the front-end amplifier. The excess noise injected into the signal path by the front-end amplifier is characterized by its noise figure -- F . An effective noise temperature is chosen from this excess noise figure by using the relationship:

$$T_A = T_o(F - 1) \quad (4.2.7)$$

where the noise figure, if expressed in terms of decibels, is converted to a number, and the background temperature of $T_o = 290K$ is a matter of convention. After using this relationship, the Johnson noise power of Eqn. 4.2.6 becomes:

$$\overline{i_{NR}^2} = \frac{4kT_A}{R_{in}} B_{if} \quad (4.2.8)$$

The signal power due to the heterodyne signal is the average value of the component at the heterodyne frequency in the photocurrent expression (Eqn. 4.2.2):

$$\overline{i_{SIG}^2} = 2P_{lo}P_{sig} \left(\frac{e\eta}{h\nu} m(\phi) \right)^2 \quad (4.2.9)$$

We now can form the signal-to-noise ratio (SNR) at the input to the AGC block of Fig. 4.0.1 (shown in detail in Fig. 4.2.1), by taking the ratio of Eqn. 4.2.9 to Eqns. 4.2.8 and 4.2.5:

⁹ Yariv [29], p. 431.

$$SNR_{dct} = \frac{2P_{lo}P_{sig}(e\eta m(\phi)/h\nu)^2}{\left[2e\left\{i_d + (P_{lo} + m^2(\phi)P_{sig} + P_b)(e\eta/h\nu)\right\} + 4kT_A/R_{in}\right]B_{if}} \quad (4.2.10)$$

From this SNR one can see why a heterodyne detection system is important; the local-oscillator (LO) acts as an optical amplifier, boosting the signal power. If the power in the LO field is strong enough, the shot noise term from the LO will dominate the other system noise sources at the photodetector and the system detection performance is pushed to its quantum limit¹⁰:

$$SNR_Q = \frac{\eta m^2(\phi)P_{sig}}{h\nu B_{if}} = CNDR \cdot \frac{1}{B_{if}} \quad (4.2.11)$$

As shown above, the expression for the receiver SNR in quantum limited operation is also equal to the receiver carrier-to-noise density ratio (CNDR) divided by the bandwidth of the receiver (limited by its i.f. bandpass filter bandwidth), where $h\nu$ is the spectral noise density (at the receiver operating frequency ν) and $\eta P_{sig} m^2(\phi)$ is the detected carrier power (including the losses due to detector quantum efficiency -- η and nutation -- $m^2(\phi)$). When the spectral noise density is multiplied by the operating bandwidth, the total system noise -- $h\nu B_{if}$ is found. The ratio of these two quantities forms the receiver SNR at the photodetector.

¹⁰ To see this evaluate Eqn. 4.2.11 when SNR_Q is equal to unity. Since $m^2(\phi)P_{sig}/h\nu$ [*photons/sec*] is the rate of signal photon incidence on the photodiode, when it's divided by the bandwidth of the i.f. bandpass filter, B_{if} [*1/sec*], assuming the detector quantum efficiency, η , is unity, the resulting expression $1 = m^2(\phi)P_{sig}/h\nu B_{if}$ [$1 = \textit{photons}$] says that there is a single photon every $1/B_{if}$ seconds. In other words, the system is able to make a detection for every incident photon. This is clearly the limit of the system's sensitivity, because we must have at least one photon (and corresponding electron-hole pair) to make a detection. Further, this relationship also tells us that a system with a bandwidth of B_{if} Hertz cannot detect an event which is shorter than $1/B_{if}$ seconds, where $1/B_{if}$ seconds is known as the resolution or integration time of the system.

The SNR of Eqn. 4.2.10 is present at the input to the automatic gain control block of Fig. 4.2.1. The AGC block is essentially a signal power estimator which is used to maintain the received signal power at some desired reference level. AGC is performed by taking an estimate of the noise in the heterodyne signal spectrum (the LOW I.F. signal-to-noise ratio -- SNR NOISE POWER ESTIMATE branch of the figure) and subtracting this from the total power spectrum, containing both the signal and noise power. This signal power estimate is compared to a desired reference signal power, and the error signal is sent to an integrator. The voltage on the integrator is then sent to a voltage controlled attenuator which ultimately provides the AGC function.

The total average power (the DC component) present at the input to the AGC block is the sum of the signal and noise components (Eqns. 4.2.9, 4.2.8, and 4.2.5):

$$\begin{aligned} \overline{I^2} = & 2P_{lo}P_{sig}\left(\frac{e\eta}{h\nu}m(\phi)\right)^2 \\ & + \left[2e\left\{i_d + (P_{lo} + m^2(\phi)P_{sig} + P_b)\left(\frac{e\eta}{h\nu}\right)\right\} + \frac{4kT_A}{R_{in}}\right]B_{if} \end{aligned} \quad (4.2.12)$$

When the receiver i.f. signal-to-noise ratio (SNR) is high (Eqn. 4.2.10 $SNR \gg 1$), this power is a good estimate of the received signal power, P_{sig} :

$$\overline{I^2} \cong 2P_{lo}P_{sig}\left(\frac{e\eta}{h\nu}m(\phi)\right)^2 \propto P_{sig} \quad \text{if } SNR \gg 1 \quad (4.2.13)$$

This signal corresponds to the upper branch (labeled SIGNAL + NOISE) of the signal power estimate formed by the AGC block. As long as this signal power is within the square-law region of the square-law detector ($(\)^2$), it will produce a voltage that is linearly proportional to the received signal power¹¹. One must

¹¹ This is why we are keeping track of the signal and noise power bookkeeping in terms of the average value of the square of the photodetector current. We have tacitly assumed secondary square-law detection.

insure that the entire expected range of received signal power falls within the secondary detector's square-law region to ensure the validity of Eqns. 4.2.3 -- 4.2.12.

As the receiver i.f. SNR approaches unity (Eqn. 4.2.10 $SNR \rightarrow 1$), the total received power is no longer an accurate estimate of the received signal power because of the contribution due to noise (Eqns. 4.2.5 and 4.2.8). If one were to continue to use the total power as an estimate of the received signal power, the performance of the AGC block would degrade. This is easily understood if we consider the case where the receiver i.f. SNR is poor (Eqn. 4.2.10 $SNR \ll 1$); in this case:

$$\bar{I}^2 \equiv \left[2e \left\{ i_d + (P_{lo} + m^2(\phi)P_{sig} + P_b) \left(\frac{e\eta}{h\nu} \right) \right\} + \frac{4kT_A}{R_{in}} \right] B_{if} \quad \text{if } SNR \ll 1 \quad (4.2.14)$$

which is essentially constant because $P_{sig} \ll P_{lo}$. This constant signal power estimate would limit the maximum gain of the AGC block, right at the time when additional gain is most necessary. As a result, a better estimate of the received signal power at low receiver i.f. SNR must be derived.

As shown in Fig. 4.2.1, a better signal power estimate is achieved if one first attempts to cancel the mean noise power in the received signal (the LOW I.F. SNR NOISE POWER ESTIMATE branch of the figure). The basic idea is to sample a small part of the mean noise density in the received heterodyne signal spectrum and use this as an estimate of the entire mean spectrum noise. This process is shown in Fig. 4.2.3, where the entire heterodyne signal spectrum is from 200-1000 MHz (the passband of our i.f. bandpass filter in Fig. 4.0.1) and the portion from 200-280 MHz is used as an estimate of the total noise power. The passband of the narrow bandpass filter must be located away from any expected heterodyne signal energy. Thus, the bandwidth of the i.f. filter must be slightly increased to accommodate this AGC control system. In our design, the

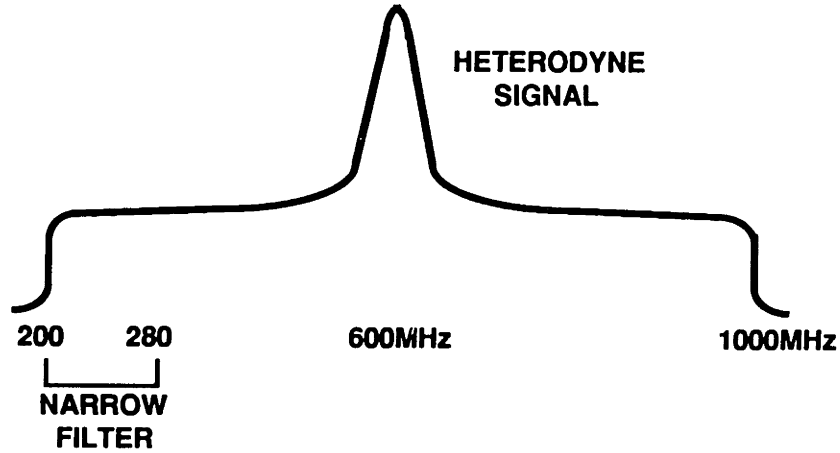


Figure 4.2.3: Heterodyne signal spectrum.

bandwidth was increased by 10 % (80 MHz of the 800 MHz i.f. passband was used for AGC). If one is designing a new system from scratch, a similar 10 % increase in necessary i.f. bandwidth is a reasonable first estimate.

The noise power estimate formed by the low CNDR branch of the AGC block is (similar to Eqns. 4.2.5 and 4.2.8):

$$\overline{I_{NE}^2} = \left[2e \left\{ i_d + (P_{lo} + m^2(\phi)P_{sig} + P_b) \left(\frac{e\eta}{h\nu} \right) \right\} + \frac{4kT_A}{R_{in}} \right] KB_n \quad (4.2.15)$$

where K is the gain of the amplifier in the narrow filter channel, B_n is the NEB of the narrow bandpass filter, and the other terms are as explained previously. The gain of the amplifier in front of the narrow filter is set so that the noise power in both channels is equal when no signal power ($P_{sig} = 0$) is striking the photodetector:

$$\begin{aligned} \left[2eP_{lo} \left(\frac{e\eta}{h\nu} \right) + \frac{4k(T_A + F_N T_o)}{R_{in}} \right] KB_n = \\ \left[2eP_{lo} \left(\frac{e\eta}{h\nu} \right) + \frac{4kT_A}{R_{in}} \right] B_{if} \end{aligned} \quad (4.2.16)$$

$$\therefore K = \left(\frac{B_{if}}{B_n} \right) \left[\frac{2eP_{lo}(e\eta/h\nu) + 4kT_A/R_{in}}{2eP_{lo}(e\eta/h\nu) + 4k(T_A + F_N T_o)/R_{in}} \right]$$

In the above relation, T_A is the effective noise temperature for the wide channel of the AGC block (only filtered by the i.f. BPF of Fig. 4.0.1) given by Eqn. 4.2.7, F_N is the noise figure for the additional amplifier in the low CNDR path of the AGC block in Fig. 4.2.1, T_o is the nominal temperature (290K by convention), and all other quantities are as previously described. Two approximations are made in Eqn. 4.2.16; the background signal power is negligible compared to the local-oscillator power ($P_{lo} \gg P_b$) and the shot noise term due to the local-oscillator dominates the photodetector dark current ($P_{lo}(e\eta/h\nu) \gg i_d$). This result makes good sense because if the additional noise added into the narrow channel by the extra amplifier is neglected (the LO shot noise is still dominant), then the total noise in each channel is proportional to the bandwidth of the channel ($K \rightarrow B_f/B_n$ as $P_{lo} \rightarrow \infty$), as the spectral density function implies.

Finally, the noise power estimate from the narrow filter (Eqn. 4.2.15 with K as given by Eqn. 4.2.16) is subtracted from the signal and noise power in the heterodyne spectrum (Eqn. 4.2.12), giving a better signal power estimate. The same approximations used to derive K are used, namely $P_{lo} \gg P_b$ and $P_{lo}(e\eta/h\nu) \gg i_d$, only now we are illuminating the diode with some signal power ($P_{sig} \neq 0$):

$$\begin{aligned} \overline{\hat{I}_{sig}^2} &= \overline{I^2} - \overline{I_{NE}^2} \\ &= \underbrace{2P_{lo}P_{sig} \left(\frac{e\eta}{h\nu} \right)^2}_{\text{signal power}} + \frac{2eP_{sig}(e\eta/h\nu)B_f(4kF_N T_o/R_{in})}{\underbrace{2eP_{lo}(e\eta/h\nu) + [4k(T_A + F_N T_o)/R_{in}]}_{\text{noise term}}} \end{aligned} \quad (4.2.17)$$

There are two limiting cases worth consideration. If the receiver is performing at its quantum limit (LO shot noise limited -- Eqn. 4.2.11), then the

approximation $2eP_{lo}(e\eta/h\nu) \gg 4kT_A/R_{in}$ is valid¹² and the signal power estimate becomes:

$$\hat{I}_{sig}^2 = \underbrace{2P_{lo}P_{sig}\left(\frac{e\eta}{h\nu}\right)^2}_{\text{signal power}} + \underbrace{\frac{P_{sig}}{P_{lo}}B_{if}\left(\frac{4kF_N T_o}{R_{in}}\right)}_{\text{noise term}} \quad \begin{array}{l} \text{quantum} \\ \text{limited} \end{array} \quad (4.2.18)$$

On the other hand, if the effective Johnson noise is the dominant noise term at the photodetector ($4kT_A/R_{in} \gg 2eP_{lo}(e\eta/h\nu)$ in Eqn. 4.2.10), the signal power estimate¹³ is:

$$\hat{I}_{sig}^2 = \underbrace{2P_{lo}P_{sig}\left(\frac{e\eta}{h\nu}\right)^2}_{\text{signal power}} + \underbrace{2eP_{sig}\left(\frac{e\eta}{h\nu}\right)B_{if}\left(\frac{F_N}{F_N + F_W - 1}\right)}_{\text{noise term}} \quad \begin{array}{l} \text{Johnson} \\ \text{noise} \\ \text{limited} \end{array} \quad (4.2.19)$$

where Eqn. 4.2.7 is used with F_W as the noise factor for the photodiode front-end amplification to simplify the expression. If neither of these approximations are true, then the full expression given by Eqn. 4.2.17 must be used.

The essential difference between the two AGC methods is evident when Eqn. 4.2.12 is compared to Eqn. 4.2.17. A simple AGC scheme uses the entire i.f. spectrum as an estimate of the received signal power, and corresponds to using only the upper (SIGNAL + NOISE) branch of Fig. 4.2.1 for AGC. In this configuration the AGC block will keep the total average, received power constant (\bar{I}^2 of Eqn. 4.2.12). Our low receiver, i.f. SNR, noise power cancellation method cancels the mean noise density from the i.f. spectrum (via the NOISE ESTIMATE

¹² I am assuming the other approximations we have been using, namely $P_{lo} \gg P_{sig}, P_b$ and $P_{lo}(e\eta/h\nu) \gg i_d$, are still valid. Combining all of these approximations yields the equivalent condition that the term due to the local-oscillator power in the first bracket of noise terms in Eqns. 4.2.12 and 4.2.15 is dominant; mathematically:

$$2e\left\{i_d + (P_{lo} + P_{sig} + P_b)(e\eta/h\nu)\right\} \cong 2eP_{lo}(e\eta/h\nu) \quad (F.4.12)$$

Another implicit assumption used to reduce Eqn. 4.2.17 is that the LO shot noise term is also able to dominate the additional noise added by the amplifier in the low CNDR signal path ($F_N T_o$).

¹³ Once again, assumption F.4.15 is also assumed valid.

branch of Fig. 4.2.1) to form a better estimate of the received signal power. It is this improved signal power estimate which is then kept constant by the AGC block of Fig. 4.2.1 ($\overline{\hat{I}_{sig}^2} = \overline{I^2} - \overline{I_{NE}^2}$ of Eqn. 4.2.17). It should be noted that this AGC method is only able to cancel the average noise in the received i.f. signal spectrum. Once this has been done, the accuracy of the improved signal power estimate is limited by the rms power fluctuations in the i.f. signal spectrum.

Further observation of the method used to estimate the receiver noise suggests a modification to the AGC block of Fig. 4.2.1 which should provide the best possible performance that this method can offer. If the filter in the noise estimate branch had the same NEB as in the signal channel, the residual dc error term from Eqn. 4.2.17 is eliminated. In this case, as with the previous system, performance is limited by the rms power fluctuations in the received signal, because the method only subtracts the average noise power. However, the fluctuations in each channel should be more similar than the present case, precisely because both channels are operating with the same NEB, offering better performance. The drawback to this improvement would be the increased bandwidth necessary at the receiver front-end; in fact, its size would have to double. Because the configuration shown in Fig. 4.2.1 worked satisfactorily, this modification went untried as of this writing. A hint at its success was evident in the fact that we received better noise performance when the passband of our narrow sampling filter was increased from 40 MHz to 80 MHz during testing.

4.3. Implementation of AGC by Noise Power Cancellation

The AGC method described conceptually in the previous subsection was tested in the receiver of Fig. 4.0.1. The realization of this AGC method, shown in block diagram form in Fig. 4.2.1, was implemented by the circuitry shown in Fig. 4.3.1. This subsection will briefly describe this circuitry and the next subsection will discuss the performance of this AGC method in our receiver.

As shown in Fig. 4.2.3, this AGC method begins by taking an estimate of the noise power in the heterodyne signal spectrum. This estimate is taken by the second branch of the four-way RF power splitter at the top of Fig. 4.3.1A, and corresponds to the low I.F. SNR branch of Fig. 4.2.1. From the figure, the total received signal power is passed through a narrow bandpass filter whose passband (200 - 280 MHz)¹⁴ does not pass any of the heterodyne signal energy. The output of this filter is square-law detected and forms a noise power estimate. This noise power estimate is then effectively amplified by 10 dB with respect to the power in the total heterodyne spectrum¹⁵, which is attenuated and square-law detected by a matched detector¹⁶ in the top branch from the power splitter. Since the noise power in the signal spectrum is proportional to the bandwidth of the

¹⁴ Unfortunately, due to an oversight on my part, the NEB of this filter was never measured.

¹⁵ The power passed by the narrow filter is amplified by 5 dB, the net effect of the -15 dB of attenuation and the +20 dB gain of the 1522 Advantek RF power amplifier. The same effect could be achieved by a +5 dB RF power amplifier (which we did not have). The total power in the received signal is attenuated by -5 dB. The net difference in power level between the wide and narrow channels is 10 dB.

¹⁶ The responses of the two square-law detectors in the wide and narrow channels of the AGC block need to be matched because this was implicitly assumed in the derivation of the AGC performance.

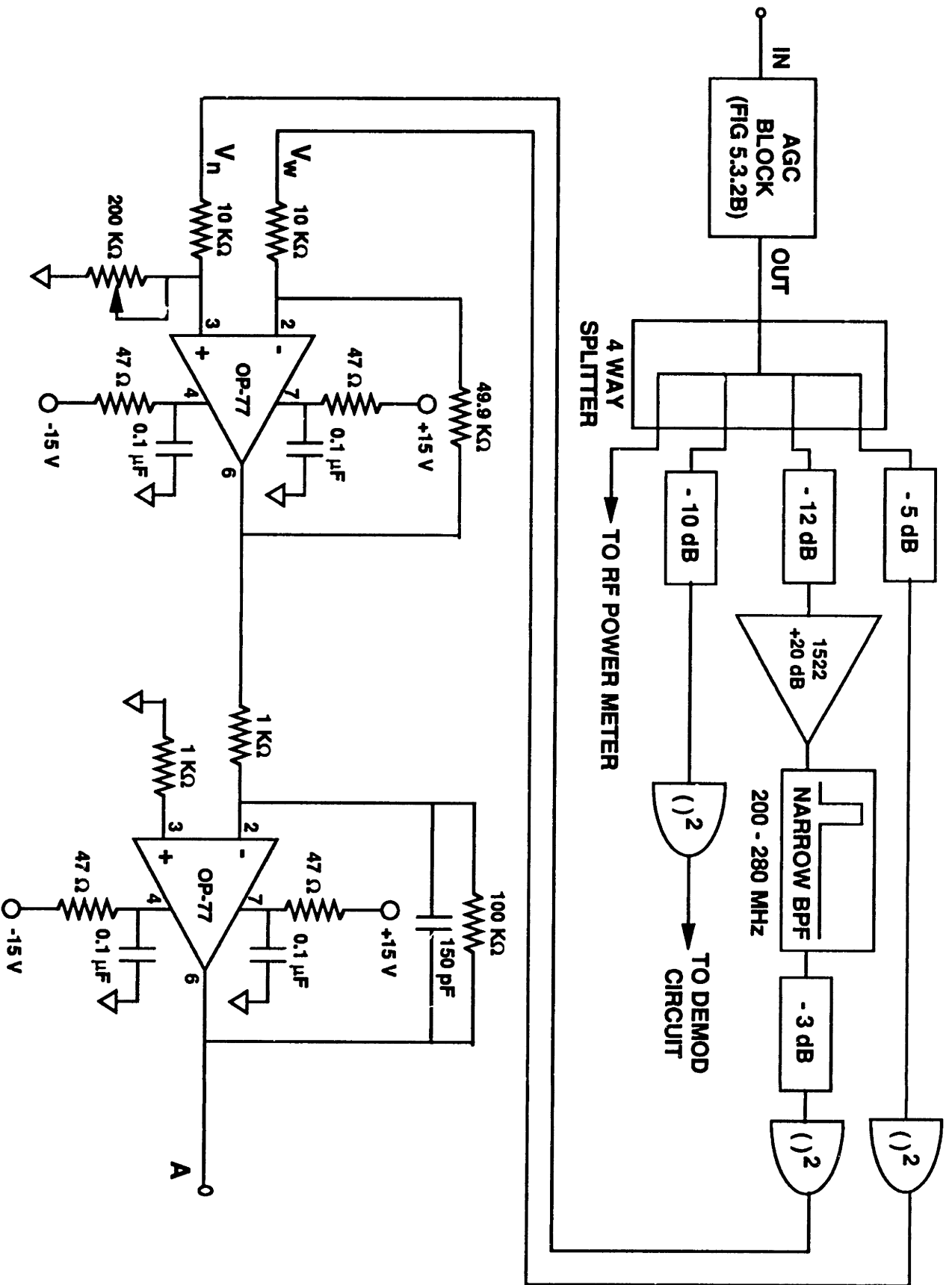


Figure 4.3.1A: Schematic for the nutating receiver AGC block.

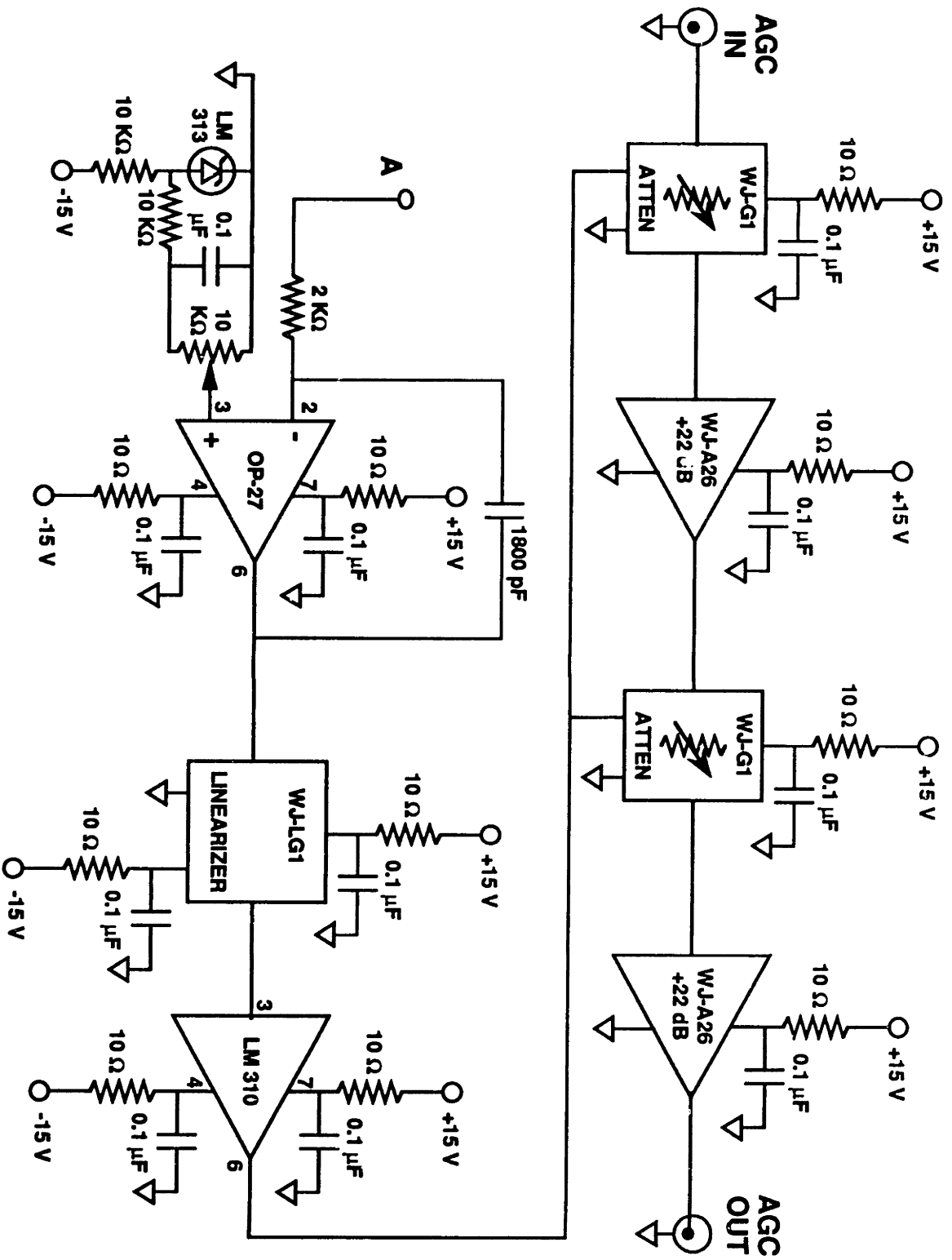


Figure 4.3.1B: Schematic for the nutating receiver AGC block.

filter used (refer back to Eqns. 4.2.5 and 4.2.8), a 10 dB amplification of the noise power in the narrow filter channel is expected because the bandwidth of the narrow filter is about one-tenth that of the wide filter (NEB = 736.5 MHz)¹⁷. The amount of attenuation in each channel was also set so that the entire expected operating range was within the square-law region of the detectors¹⁸.

The noise power estimate is then subtracted from the total received signal power by the first OP-77 operational amplifier (op-amp) of Fig. 4.3.1A (the summation joint of Fig. 4.2.1). The transfer function of this op-amp is:

$$V_{out} = -\left(\frac{R_2}{R_1}\right)V_w + \left(\frac{R_4}{R_3 + R_4}\right)\left(1 + \frac{R_2}{R_1}\right)V_N \quad (4.3.1)$$

where $R_1 = R_3 = 10K\Omega$, $R_2 = 49.9K\Omega$, R_4 is a $200K\Omega$ variable resistor, V_w is the voltage from the wide channel detector, V_N is the voltage from the narrow channel detector, and V_{out} is the output voltage of the op-amp (pin #6). After substituting in these values and accounting for the gain of the second OP-77 op-amp¹⁹, the transfer function becomes:

$$\hat{V}_{sig} = 500V_w - (0 - 571)V_N \quad (4.3.2)$$

where \hat{V}_{sig} is now the voltage at point A (pin #6 of the second op-amp). The value of the variable resistor, R_4 , is set so that the noise power in both channels is equal with no heterodyne signal present (as described by Eqn. 4.2.16).

¹⁷ The i.f. filter bandwidth was chosen to allow for a Frequency Shift Keyed (FSK) modulation spectrum.

¹⁸ Recall that a square-law detector will produce an output voltage which is linearly proportional to the square of the field intensity (power) striking the detector. The derivation assumes square-law secondary detection.

¹⁹ The transfer function of the second OP-77 amplifier is not strictly a gain of 100. The full transfer function:

$$\frac{V_{out}}{V_{in}} = 100 \left(1 / \frac{s}{10600 \cdot 2\pi} + 1 \right) \quad (F4.19)$$

has a pole at 10.6 KHz which provides high frequency roll-off to stabilize the op-amp itself. Within the operating bandwidth of the AGC control loop, however, this op-amp can be treated as a gain of 100.

The improved signal power estimate of Eqn. 4.3.2 needs to be compared to a desired reference signal power. This is done by the OP-27 op-amp of Fig. 4.3.1B. If the voltage at the non-inverting input (pin #3) terminal, V_+ , is set to the desired reference level (by adjusting the $10K\Omega$ potentiometer at this pin), then the voltage out of this op-amp (pin #6) becomes:

$$V_{out} = \frac{V_+ - \hat{V}_{sig}}{RCs} + V_+ \quad (4.3.3)$$

where $R = 2K\Omega$ and $C = 1800pF$. This is the control voltage used for gain control (as in Fig. 4.2.1). It consists of a desired reference signal level, added to the integrated difference between the actual signal level and the reference level.

This control signal is then sent, via a WJ-LG1 linearizer and a LM310 current driver, to the two WJ-G1 voltage controlled attenuators, which ultimately provide the gain control function. The WJ-LG1 linearizer is specially designed to compensate for the non-linear attenuation curve of the WJ-G1 voltage controlled attenuators. The series combination of the linearizer and the voltage controlled attenuators provides an attenuation which is linearly proportional to the control voltage from the integrator, making the response of the control loop linear.

This completes the AGC control loop which began with the noise power estimate described previously. The gain of the receiver is controlled by the combination of the two WJ-A26 RF amplifiers and the two WJ-G1 voltage controlled attenuators. The amplifiers provide a static gain of +44 dB and the attenuators can provide up to -60 dB of attenuation, for a gain range of -16 dB to 44 dB for the AGC block as a whole.

4.4. AGC Measured Performance

The method used to measure the open-loop noise performance of the receiver in Fig. 4.0.1 provides a measure of the functionality of the AGC control loop as well. By referring back to Sec. 3.2, recall that an optical wedge was inserted in the signal beam path just before the electro-optic nutator of Fig. 1.4.1 to calibrate the tracking error discriminator gain to an optical beamwidth (Eqn. 3.2.6). The induced voltage in the power spectrum of the tracking error signals produced by the spinning wedge was measured after the AGC feedback loop. If the AGC control loop is functioning properly, one should expect that the discriminator gain would remain constant as the receiver CNDR varies. After all, this is the purpose of the AGC control block, to maintain the signal power sent to the receiver demodulation electronics at some desired reference level. Therefore, by observing the measured voltage produced by the calibration wedge, a direct reflection on the performance of the AGC control loop is achieved.

The AGC electronics module described in the previous subsection is a modification of an earlier AGC block which did not use the low receiver CNDR noise cancellation technique. Fig. 4.3.1B, combined with an additional op-amp and power splitter, shown in Fig. 4.4.2, is a fully functional AGC block in itself. In this scheme, the total received power is sampled, with some filtering, and sent as the signal power estimate to the integrator (Eqn. 4.3.3). This AGC method keeps the entire received power, $\overline{I^2}$ from Eqn. 4.2.12, constant. From the previous analysis, we expect this AGC method to degrade as the receiver i.f. SNR approaches unity (Eqn. 4.2.10 $SNR_{det} \rightarrow 1$).

A plot of this AGC performance, compared to the performance of the low receiver CNDR noise cancellation technique is shown in Fig. 4.4.1. As expected, the performance of the simple AGC method degrades quickly. The normalized loop gain is equal to one-half its nominal value when the receiver CNDR is 89 dB-Hz, corresponding to an i.f. SNR of unity in our system. This stands to reason because if the entire received power is used as a signal power estimate, then this estimate will be twice as big as it should be when the i.f. SNR is unity. As a result, the normalized tracking loop-gain should be one-half its nominal value, as observed.

On the other hand, the noise cancellation AGC method performs well down to 77 dB CNDR (corresponding to an i.f. signal-to-noise ratio of < -10 dB in our system), extending the AGC range by more than 15 dB. Around 77 dB CNDR, the AGC loop is not attenuating the received signal power at all, and any further decrease in received signal power is not corrected. This gentle roll-off in AGC performance at low receiver CNDR is desirable because as the receiver NEA (Eqn. 3.1.25) approaches significant fractions of a beamwidth, it is desirable to lower the tracking loop NEB. Typically the gain decrease margin of the tracking loop is sufficient to accommodate this.

Before leaving the discussion of the receiver AGC block, we should mention the consideration of the bandwidth for the AGC control loop itself. The bandwidth of this loop is primarily determined by the RC integrator gain constant (the OP-27 op-amp of Fig. 4.3.1B whose transfer function was given by Eqn. 4.3.3). The values in our AGC block were chosen to yield an AGC loop bandwidth of 100 Hz (the -3 dB rejection point). This bandwidth was chosen somewhat arbitrarily, and an analysis of the necessary bandwidth for an AGC block in an optical heterodyne receiver needs to be performed in the future.

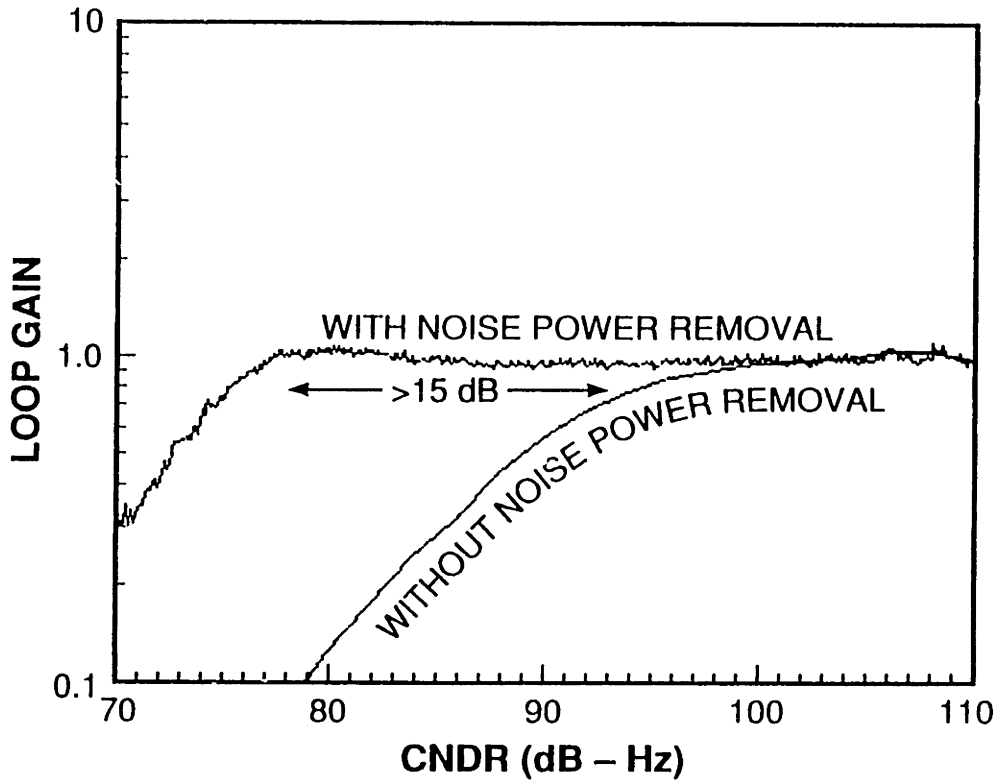


Figure 4.4.1: AGC measured performance.

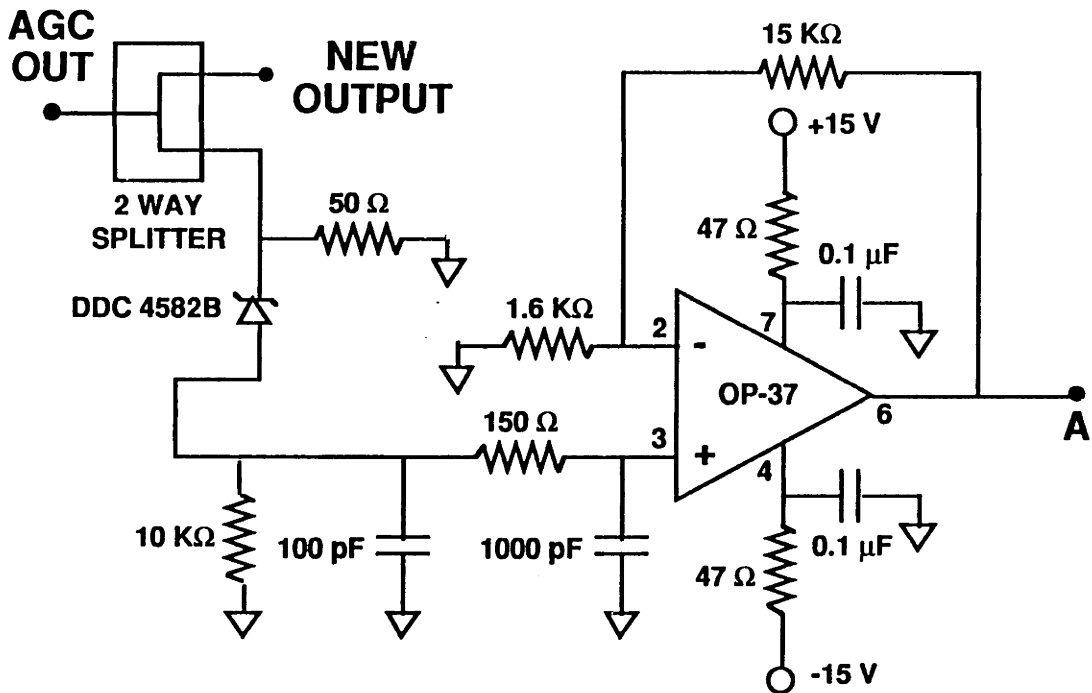


Figure 4.4.2: Change in the AGC circuit of Figure 4.3.1B to remove the low receiver CNDR noise cancellation scheme.

5. Spatial Tracking

The uncompensated spatial tracking error is determined by the rejection of the tracking loop. This can be seen by expressing σ_{uc}^2 from the total tracking error of Eqn. 3.0.1 in terms of the rejection of the spatial tracking loop -- $R(f)$, and the power spectrum of the environmental angular disturbances -- $S_\theta(f)$:

$$\sigma_{uc}^2 = \int_{-\infty}^{\infty} |R(f)|^2 S_\theta(f) df \quad (5.0.1)$$

To provide adequate tracking performance, the rejection of the tracking loop must be sufficient to reduce σ_{uc}^2 below a specified level.

This section will finish the discussion of the receiver of Fig. 1.4.1 (reproduced as Fig. 5.0.1) by closing the two-axis spatial tracking feedback loop. Sec. 5.1 describes the nutator demodulator. Sec. 5.2 completes specification of the open-loop transfer function for the tracking loop. Sec 5.3 discusses the design issues associated with the compensator used to close and stabilize the tracking loop. Lastly, Sec. 5.4 discusses the two axis closed-loop spatial tracking measurements.

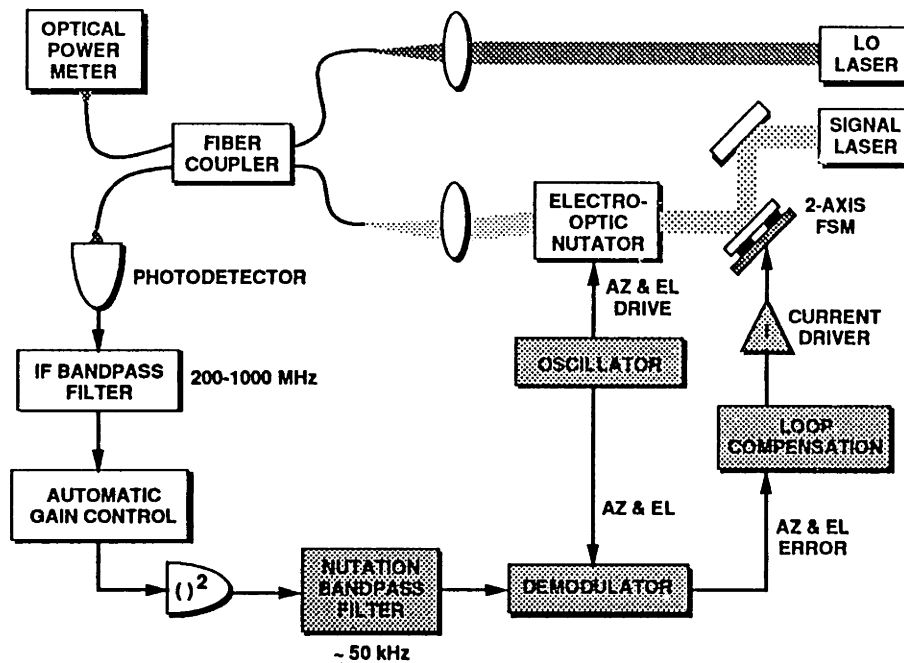


Figure 5.0.1: Nutating receiver block diagram.

5.1. EON Spatial Tracking Synchronous Demodulator

The heterodyne signal from the AGC block described in the previous section is sent to the synchronous demodulator. This signal comes through the third leg of the four-way splitter of Fig. 4.3.1A and is square-law detected. The detector converts the impinging signal power to an analog voltage which is sent into the demodulator input. A block diagram of the demodulator, whose functionality is discussed in this section, is shown in Fig. 5.1.1. A complete circuit diagram of the implementation of this demodulator is shown in Fig. 5.1.2.

From the block diagram, the heterodyne signal is sent through some front-end amplification. This amplification is performed by the four OP-37 op-amps of

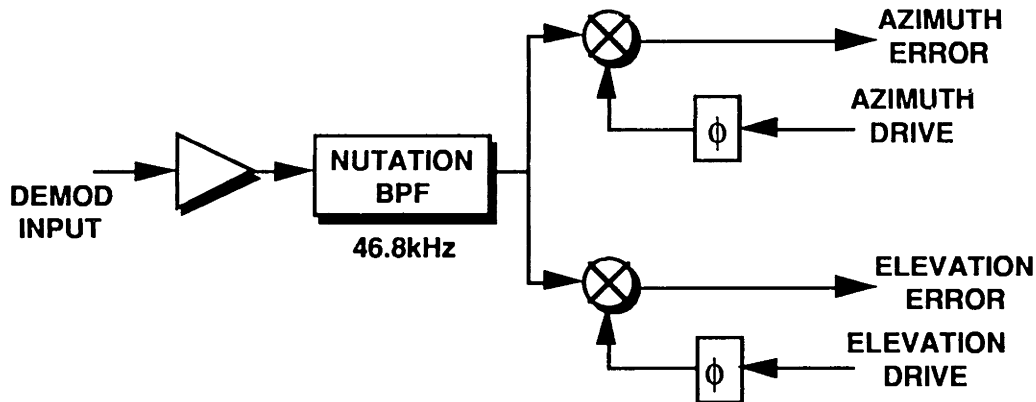


Figure 5.1.1: Nutating receiver synchronous demodulator block diagram.

Fig. 5.1.2A. The first amplifier has a transfer function¹ of:

$$\frac{V_{out}}{V_{in}} = -\frac{R_2}{R_1} \quad (5.1.1)$$

With the values from the circuit diagram, $R_1 = 499\Omega$ and $R_2 = 4.99K\Omega$, this amplifier provides a gain of ten. The value for the input resistors, $R_1 = 499\Omega$, was chosen to make the input impedance of the amplifier:

$$R_{input} = 2R_1 \quad (5.1.2)$$

match the load impedance² of the square-law detector, $R_{load} = R_{input} = 1K\Omega$.

¹ To see that this amplifier has the same gain as a simple inverting op-amp, write the KCL relations at the V^+ and V^- nodes (pins #2 & 3):

$$\begin{aligned} (V_{in}^+ - V^-)/R_1 &= (V^- - V_{out})/R_2 && \text{KCL pin \# 2} \\ (V_{in}^- - V^+)/R_1 &= V_+/R_2 && \text{KCL pin \# 3} \end{aligned} \quad (F5.1)$$

where R_1 is the resistance of the two input resistors (assumed equal) and R_2 is the feedback resistor and the one going to ground from the non-inverting input (also assumed equal). After a little bit of algebra and substituting the relationship, $V_{in}^+ - V_{in}^- = V_{in}$, we arrive at the same gain as the simple inverting connection given in Eqn. 5.1.1.

² This input resistance is found by writing two KVL relations at the input terminals to the op-amp:

$$\begin{aligned} V^- &= V_{in}^+ - I_{in}R_1 && \text{KVL pin \# 3} \\ V^+ &= V_{in}^- + I_{in}R_1 && \text{KVL pin \# 2} \end{aligned} \quad (F5.2)$$

Combining these by using the standard op-amp approximation, $V^- \approx V^+$, gives Eqn. 5.1.2. Our square law detector required a 1 k Ω load impedance for proper operation.

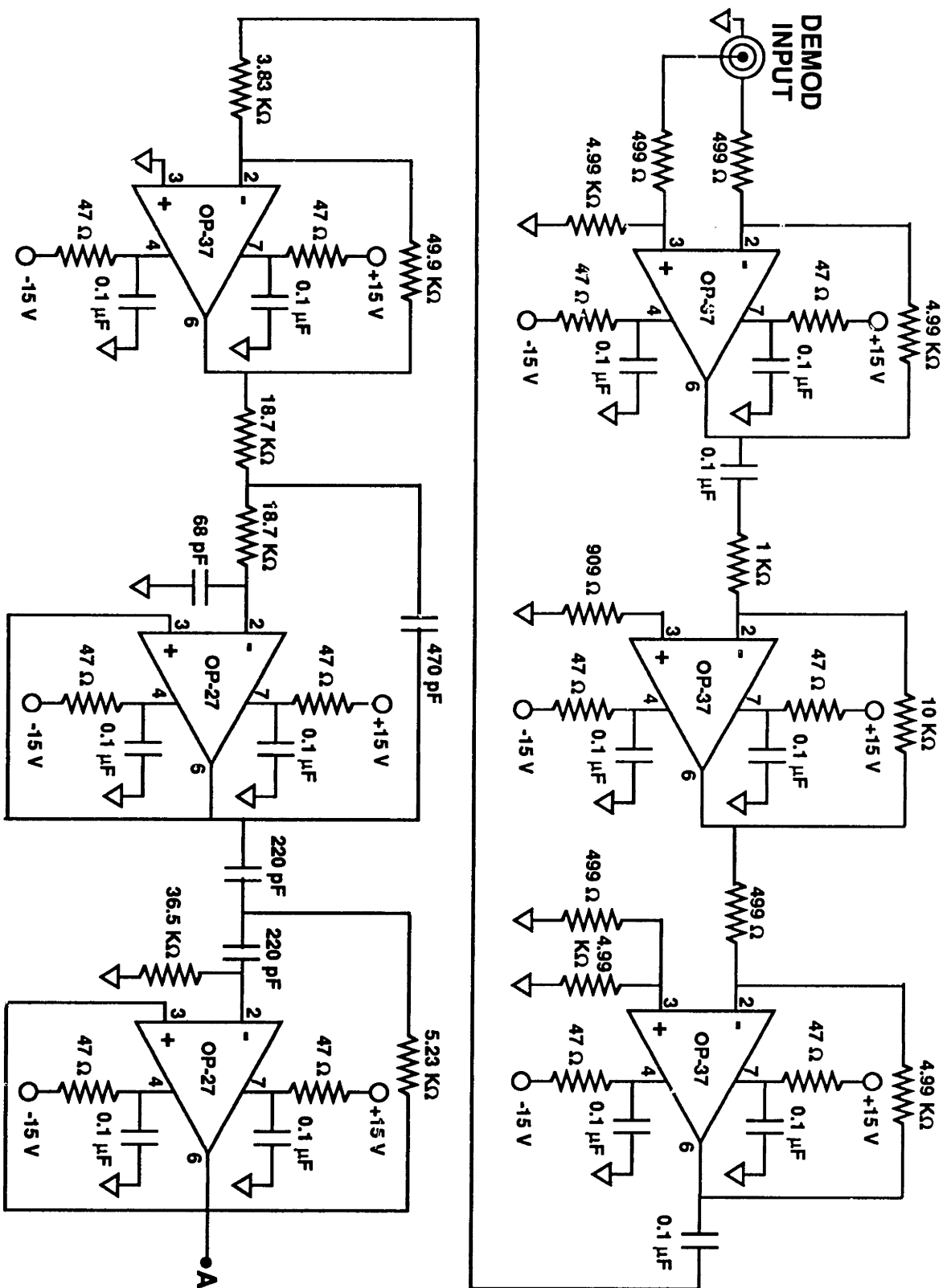


Figure 5.1.2A: Nutating receiver demodulator schematic

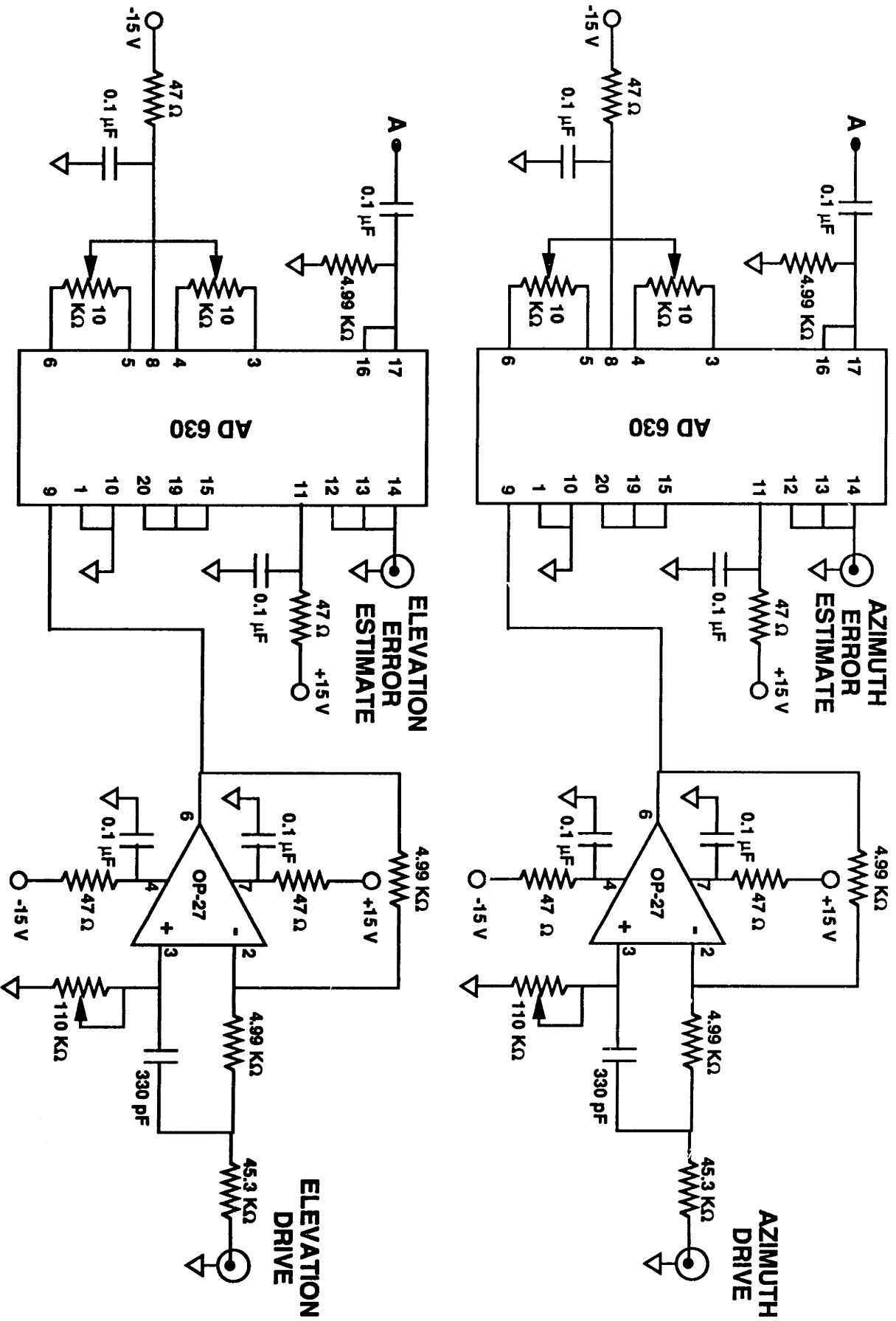


Figure 5.1.2B: Nutating receiver demodulator schematic.

The next three OP-37 op-amps provide a composite transfer function³ of:

$$\frac{V_{out}}{V_{in}} = -\frac{R_4 R_6 R_8}{R_3 R_5 R_7} \quad (5.1.3)$$

where $R_3 = 1K\Omega$, $R_4 = 10K\Omega$, $R_5 = 499\Omega$, $R_6 = 4.99K\Omega$, $R_7 = 3.83K\Omega$, and $R_8 = 49.9K\Omega$ from the circuit diagram. After plugging in these values, the gain becomes -1300. As a result, the entire front-end gain is -13000.

The next two OP-27 op-amps are the nutation bandpass filter from the block diagram. The first is a standard Sallen-Key lowpass [33] filter:

$$\frac{V_{out}}{V_{in}} = \frac{1}{R_1 R_2 C_1 C_2 s^2 + (R_1 + R_2) C_1 s + 1} \quad (5.1.4)$$

Recognizing the standard form for a quadratic, $\frac{s^2}{\omega_o^2} + \frac{2\zeta}{\omega_o} s + 1$, we can solve for the

resonant frequency:

$$\omega_o = \frac{1}{\sqrt{R_1 R_2 C_1 C_2}} = 47600 \cdot 2\pi \left[\frac{\text{rad}}{\text{sec}} \right] \quad (5.1.5)$$

and the damping ratio⁴:

$$\zeta = \frac{(R_1 + R_2)}{2\sqrt{R_1 R_2}} \sqrt{\frac{C_1}{C_2}} = 0.38 \quad (5.1.6)$$

of this filter. The second op-amp is a standard Sallen-Key highpass [33] filter:

$$\frac{V_{out}}{V_{in}} = \frac{R_1 R_2 C_1 C_2 s^2}{R_1 R_2 C_1 C_2 s^2 + (C_1 + C_2) R_2 s + 1} \quad (5.1.7)$$

Once again using the standard quadratic form, the resonant frequency is:

$$\omega_o = \frac{1}{\sqrt{R_1 R_2 C_1 C_2}} = 52400 \cdot 2\pi \left[\frac{\text{rad}}{\text{sec}} \right] \quad (5.1.8)$$

³ These three op-amps are simply three cascaded inverting op-amps. The $0.1\mu F$ capacitors in-between them are for AC coupling. They each will introduce a low frequency zero that will not affect the basic operation of the demodulator.

⁴ The correlation between the coefficients and the values in the circuit diagram of Fig. 5.1.2A is as follows; $R_1 = R_2 = 18.7K\Omega$, $C_1 = 68pF$, and $C_2 = 470pF$.

and the damping ratio⁵ becomes:

$$\zeta = \frac{(C_1 + C_2)}{2\sqrt{C_1 C_2}} \sqrt{\frac{R_2}{R_1}} = 0.38 \quad (5.1.9)$$

The composite transfer function of this lowpass and highpass filter cascade is the two pole nutation bandpass filter⁶ with a center frequency of 46.8 kHz.

The bandpassed heterodyne signal is then sent to the two AD630 demodulators of Fig. 5.1.2B. The azimuth and elevation drive signals from the oscillator (Fig. 5.0.1) are sent through phase shifters, by the two OP-27 amplifiers, before being used for demodulation. The transfer function of each of the two phase shifters is:

$$\frac{V_{out}}{V_{in}} = \left(\frac{RCs - 1}{RCs + 1} \right) \quad (5.1.10)$$

This transfer function is an all-pass filter whose phase shift is adjusted by changing the value of the potentiometer -- R, to null out crosstalk between the azimuth and elevation channels in the demodulation process. Since the determination of the exact phase shift that the nutation signal receives from the drive at the EON to the demodulators is a difficult task, the phase shifters are necessary to align the phase of the drive signals to the nutation signal for proper demodulation⁷. This demodulator produces the spatial tracking error estimates given by Eqn. 3.1.23.

⁵ This time, the correspondence between the coefficients and the values in Fig. 5.1.2A is $C_1 = C_2 = 220 \text{ pF}$, $R_1 = 36.5 \text{ K}\Omega$, and $R_2 = 5.23 \text{ K}\Omega$.

⁶ Due to a small error in my design calculations, the measured center frequency of this filter, 46.8 kHz, differed from the designed and simulated center frequency of 50 kHz.

⁷ The two 10 k Ω potentiometers on each of the demodulators are used for fine nulling of the differential and common mode offsets of the demodulators.

5.2. Open Loop Spatial Tracking Measurements

The other two elements involved in the open-loop tracking system are the 46.8 kHz nutation oscillator, the transimpedance amplifier, and the FSM of Fig. 5.0.1. The circuit diagrams for the nutation oscillator and the transimpedance amplifier (designed by James K. Roberge) are shown in the appendix as Figs. A1.6 and A1.8 respectively⁸. The two-axis FSM in the receiver is actually two one-axis galvanometric beam steerers made by General Scanning, Inc. A photograph of one of these FSMs is shown in Fig. 5.2.1.

A measurement of the open loop transfer function for the azimuth tracking loop is shown in Fig. 5.2.2. This measurement was taken by a Hewlett Packard dynamic signal analyzer by injecting a small disturbance directly into the spatial tracking loop at the input to the current driver for the azimuth tracking FSM. The resulting azimuth spatial tracking error estimate was measured at the output of the demodulation electronics (Fig. 5.1.2B). As a result, the measured response is the open-loop response for the azimuth spatial tracking loop. The measured response for the elevation tracking loop was almost identical and is not shown, to avoid redundancy.

This open-loop transfer function is dominated by the dynamics of the General Scanning FSM. This device is essentially a second order resonant device with a very lightly damped (high Q) primary resonance around 110 Hz. It also has another high Q, secondary resonance at 10 kHz which is not as big as the

⁸ Since the design of the oscillator circuit is not a primary focus of this section and the current driver was a device previously used in the mechanical nutator [3,9,20], a detailed discussion of their operation will not be presented here. Their circuit diagrams are included in the appendix for completeness.

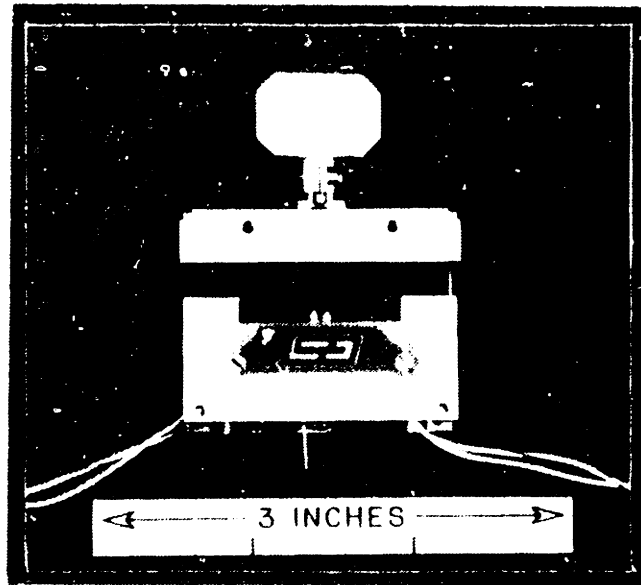


Figure 5.2.1: General Scanning FSM.

X=997.12 Hz
 Ya=10.7134 dB
 FREQ RESP
 Yb=158.799 Deg
 FREQ RESP

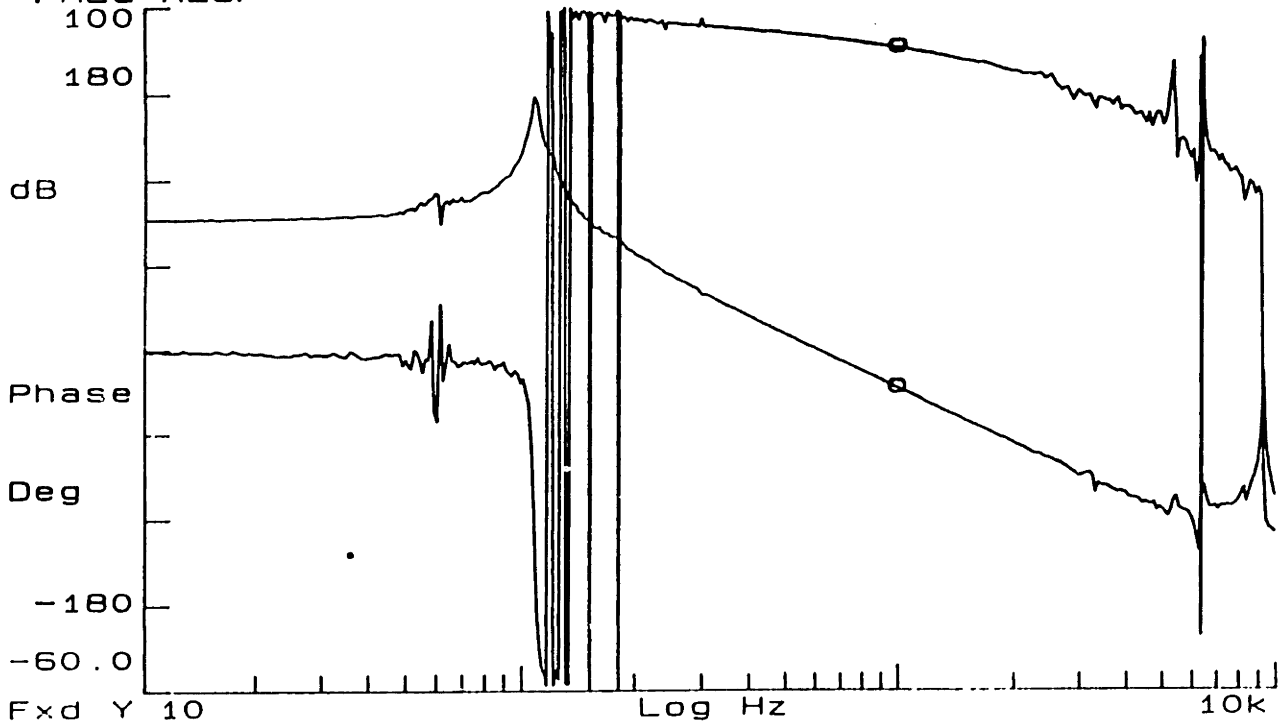


Figure 5.2.2: Azimuth spatial tracking open-loop response.

primary resonance. A 1 kHz open-loop crossover frequency⁹, a good generic benchmark for spatial tracking systems, was targeted. At the expected crossover frequency of 1 kHz, the measured open-loop response (Fig. 5.2.2) has -180° of phase shift from the primary FSM resonance, plus an additional -22° of phase shift from other sources¹⁰. The stability of the secondary resonance with a 1 kHz crossover frequency is also a driving concern.

The large bandwidth of the EON opens up the possibility of a larger bandwidth tracking loop. Since one needs to nutate at ~ 10 times the desired tracking loop bandwidth, the mechanical nutator is limited to a 1 kHz tracking loop crossover frequency due to its ~ 10 kHz nutation frequency. The EON, whose bandwidth was measured to be in excess of 100 kHz, can support a tracking loop bandwidth of up to 10 kHz. In our laboratory model, however, we were still limited to a 1 kHz open-loop crossover frequency for the tracking loop due to stability concerns of the secondary resonance of the FSMs used in our receiver. Higher bandwidth mirrors, as well as a two-axis FSM, do exist, but were unavailable at this time.

⁹ The open-loop crossover frequency is the frequency at which the gain of the open-loop transfer function is unity.

¹⁰ This extra negative phase shift comes from three sources; the demodulation electronics nutation bandpass filter, the secondary mirror resonance, and magnetic losses in the FSM.

5.3. Feedback Compensation

Closing the tracking loop completes the nutating receiver block diagram of Fig. 5.0.1. For the azimuth tracking loop, the tracking error estimate from the demodulation electronics is sent through feedback compensation and a transimpedance driving amplifier to a one-axis FSM, closing the tracking loop which began with the demodulation of the nutation signal placed upon the received heterodyne signal by the EON (Sec. 2). The elevation tracking loop is sent along an identical path to a second FSM, providing the second axis of spatial tracking. Together, the two one-axis FSMs perform the function of a single two-axis FSM¹¹. As long as the tracking errors are small, the case assumed in the derivation of Eqn. 4.0.3 (which took a linear expansion of the discriminator profile), each tracking loop will remain in its linear region, and they will operate independently. If this is not the case, the higher order terms in the Taylor's expansion of the discriminator profile will cause cross-coupling between the two tracking loops, leading to a loss of lock.

A circuit diagram for the compensator used in the tracking loops of Fig. 5.0.1 is shown in Fig. 5.3.1. This compensator begins with an inverting op-amp (#1) and two integrators (#2 & #3) feeding into a summing op-amp (#4). By writing the KCL relationships at the inverting and non-inverting inputs to the summation integrator, we can solve for the transfer function of this op-amp. After a bit of algebra this relationship becomes:

$$V_4 = -\left(\frac{R_3}{R_1}\right)V_1 - \left(\frac{R_3}{R_2}\right)V_3 + \left(\frac{R_5}{R_4 + R_5}\right)\left(\frac{R_1R_2 + R_2R_3 + R_1R_3}{R_1R_2}\right)V_2 \quad (5.3.1)$$

¹¹ A two-axis FSM exists, but was unavailable at the time of this study.

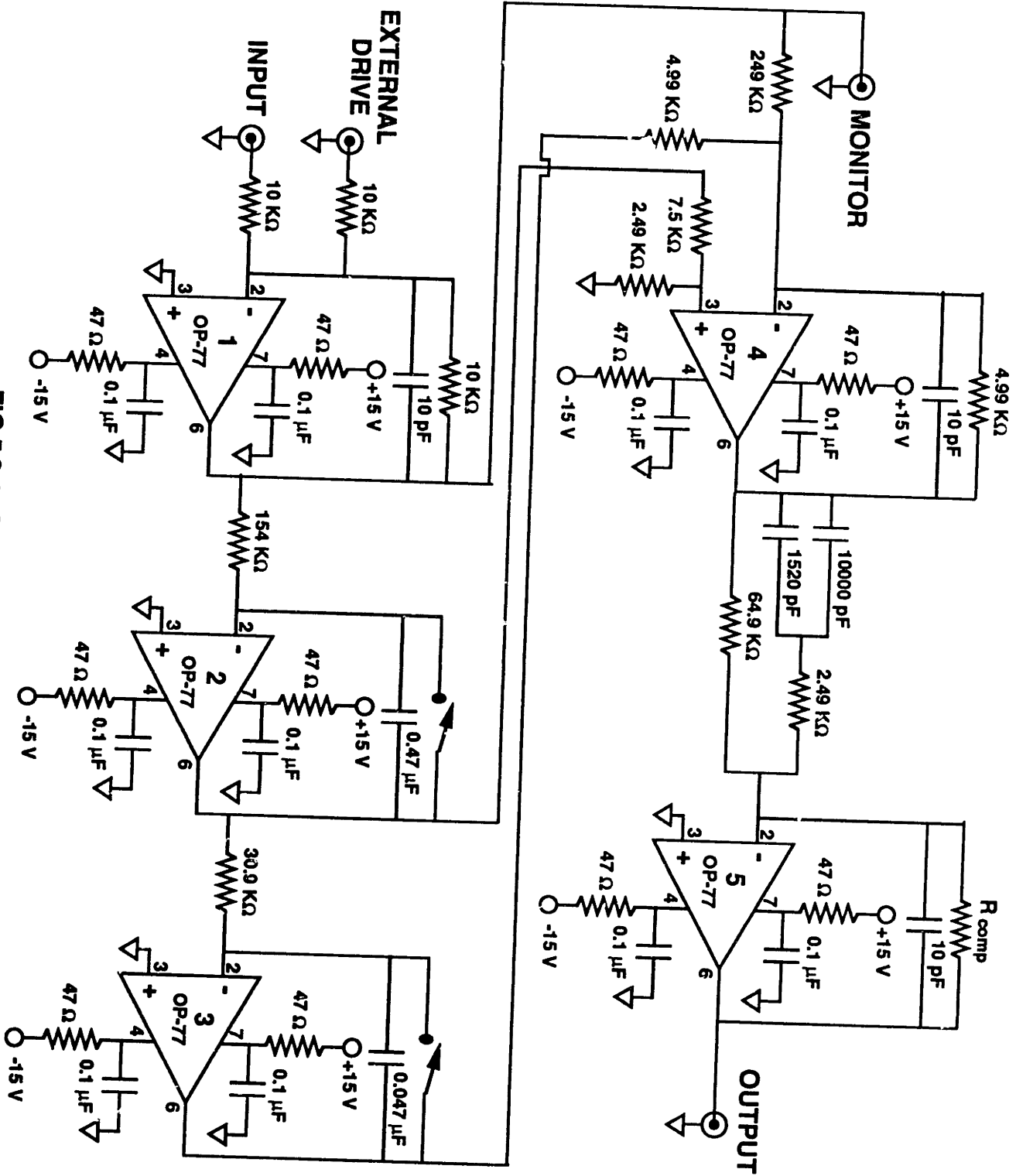


FIG 5.3.1: Compensator schematic.

where V_n is the voltage at the output of the n-th op-amp (pin #6), $R_1 = R_5 = 2.49K\Omega$, $R_2 = R_3 = 4.99K\Omega$, and $R_4 = 7.5K\Omega$. After substituting in these values, the transfer function given by Eqn. 5.3.1 becomes¹²:

$$V_4 = -2V_1 + V_2 - V_3 \quad (5.3.2)$$

The transfer function of the fifth op-amp is found by a similar procedure and is as follows:

$$V_5 = -\frac{R_{comp}}{R_2} \left[\frac{(R_1 + R_2)C_{eq}s + 1}{R_1C_{eq}s + 1} \right] V_4 \quad (5.3.3)$$

where R_{comp} is a compensating resistor used to fine tune the closed-loop gain, $R_1 = 2.49K\Omega$, $R_2 = 64.9K\Omega$, and $C_{eq} = 11500pF$ (the parallel combination of the $0.01\mu F$ and $1500pF$ capacitors). After performing these substitutions, the transfer function becomes:

$$\frac{V_5}{V_4} = \frac{R_{comp}}{64900} \left[\frac{\frac{s}{2\pi \cdot 200} + 1}{\frac{s}{2\pi \cdot 5600} + 1} \right] \quad (5.3.4)$$

Including the transfer function of the input summation op-amp (#1) and the two integrators (#2 & #3), we can form a block diagram of this compensator, Fig. 5.3.2, where the large numbers next to any block correspond to the op-amp numbers on the circuit diagram of Fig. 5.3.1. By observation of the block diagram, the composite transfer function for the compensator is as follows:

$$C(s) = \left[\frac{(0.02)(2\pi \cdot 110)^2}{s^2} \right] \left[\frac{s^2}{(2\pi \cdot 110)^2} + \frac{2(0.32)}{(2\pi \cdot 110)}s + 1 \right] \left[\frac{\frac{s}{(2\pi \cdot 200)} + 1}{\frac{s}{(2\pi \cdot 5600)} + 1} \right] \left[\frac{-R_{comp}}{64900} \right] \quad (5.3.5)$$

¹² This equation ignores the high frequency pole introduced by the 10 pF capacitor feeding back across op-amp #4. This pole, located at around a megahertz, does not affect the operation of the compensator (in a 1 kHz tracking loop). It is included to stabilize the high frequency performance of the op-amp. interested readers are referred to [33].

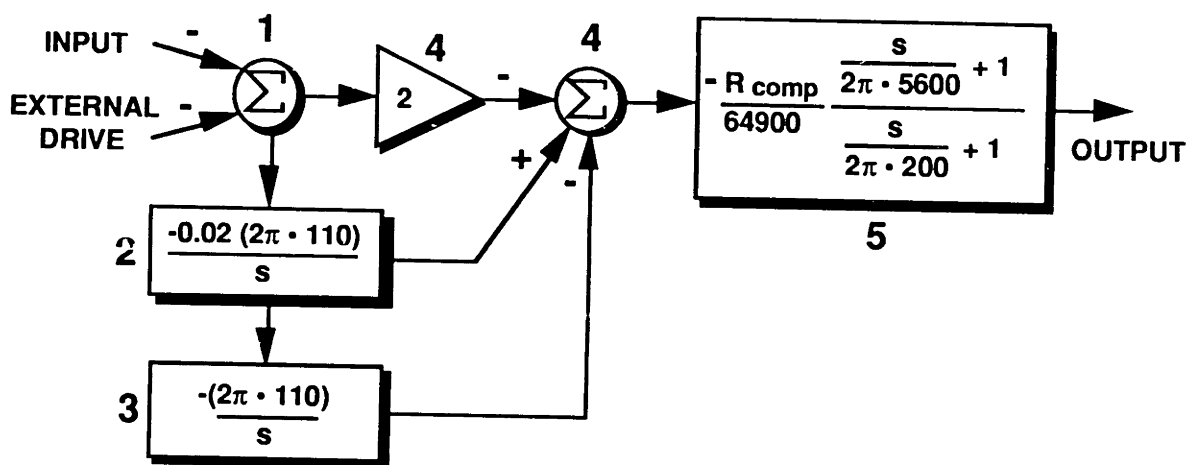


Figure 5.3.2: Compensator block diagram.

The compensator has four parts, two integrators, a complex pair of zeros, a lead network, and a gain term. The two integrators add a smooth roll-off to the open-loop characteristics and provide good rejection at low frequencies¹³. The complex pair of zeros at 110 Hz, with a damping ratio of 0.32, are present to slightly diminish the high Q of the FSM primary resonance by partial cancellation. Because this resonance is very lightly damped, exact cancellation is neither possible nor desirable. The lead network comprised of the zero at 200 Hz and pole at 5600 Hz, provides 68.6° of positive phase shift at the designed open-loop crossover frequency of 1 kHz. Another important consideration in the design of this compensator was the phase stability of the FSM secondary resonance¹⁴ at 10 kHz. In our design, the secondary resonance, while not amplitude stable, was phase stable, with a compensator tuned angle of $\sim 0^\circ$ at this resonance. The resulting compensated azimuth and elevation tracking loops had a derived open-

¹³ They also give zero steady-state error to step and ramp disturbance inputs.

¹⁴ This compensator is a modification to an earlier design used in the mechanical fiber nutating receiver of Fig. 2.1.3 and designed by James K. Roberge. The main difference between this design and the original design is the removal of a complex pair of poles at 4 kHz. This additional pair of poles, as well as a different location for the lead network, stabilized the mechanical nutator tracking loop to an open-loop crossover frequency of 1 kHz [2,8,19].

loop crossover frequency¹⁵, measured with the tracking loops closed, of 1 kHz, with 40° of phase margin and 9dB of gain margin.

5.4. Two-Axis Spatial Tracking

If we observe the basic tracking loop shown in Figure 5.4.1, a model which neglects the injected disturbances due to noise, we can determine the basic figures of merit for a feedback tracking loop. From the figure, the loop will produce an ESTIMATE of the spatial tracking errors (the environmental DISTurbance spectrum of Eqn. 5.0.1) that will differ from the actual disturbances by some amount (the ERROR). Since $H(f)$ represents the combined loop dynamics of the tracking loop (its open-loop response), the closed-loop response

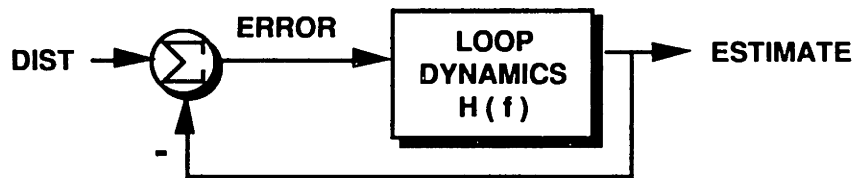


Figure 5.4.1: A simple model for a spatial tracking loop (neglecting noise).

of the tracking loop is simply:

$$\frac{\text{ESTIMATE}}{\text{DIST}} = \frac{H(f)}{1 + H(f)} \quad (5.4.1)$$

As long as $|H(f)| \gg 1$, the tracking loop will produce an accurate estimate of the injected disturbances. A more commonly used benchmark for a spatial tracking

¹⁵ The open-loop response of the tracking loop, $H(f)$, can be derived from the closed-loop response, $H(f)/(1+H(f))$, by algebraic manipulations. Conveniently, the Hewlett Packard Dynamic Signal Analyzer used to measure the closed-loop response can perform these manipulations on the measured and stored data.

loop is its rejection; the ratio of the uncompensated tracking error to the injected disturbances:

$$R(f) = \frac{\text{ERROR}}{\text{DIST}} = \frac{1}{1 + H(f)} \tag{5.4.2}$$

The rejection of the azimuth tracking loop, measured with both tracking loops locked, is shown in Fig. 5.4.2. The performance of the elevation tracking loop

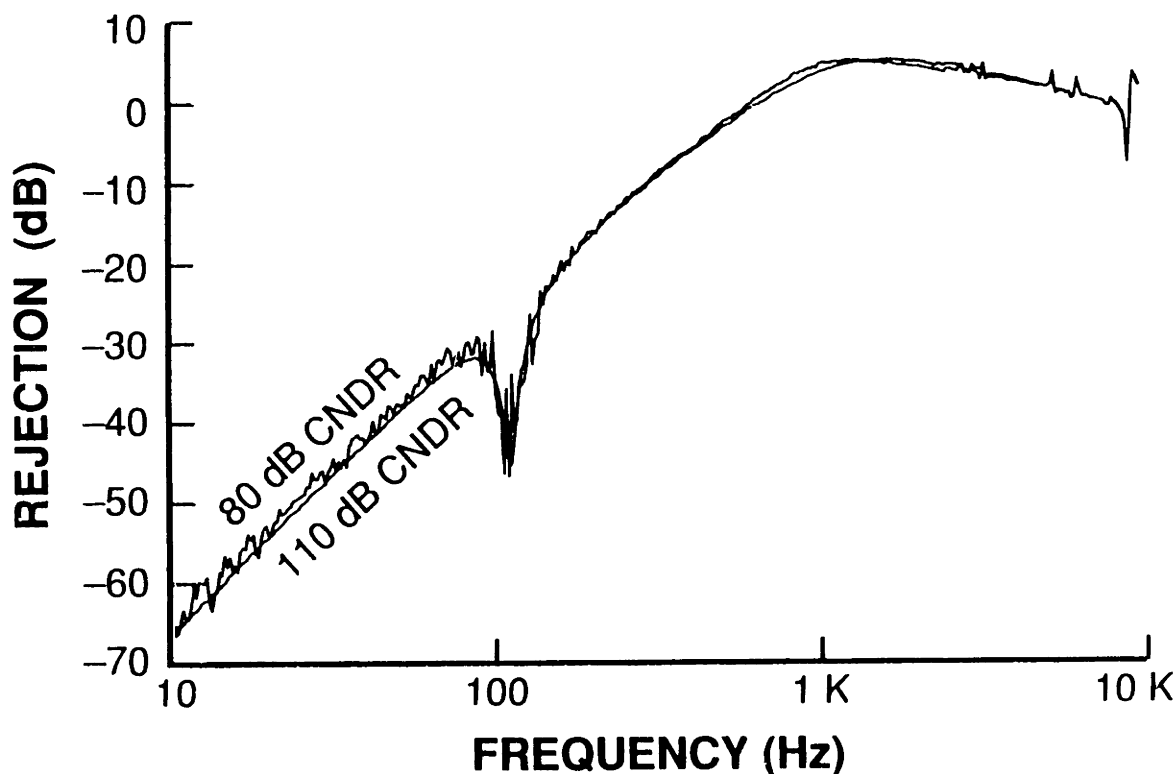


Figure 5.4.2: Azimuth closed-loop rejection.

was almost identical. The measured -3 dB bandwidth of the closed-loop transfer function¹⁶ (not shown) was ~2 kHz. The figure shows the tracking system rejection at the highest and lowest receiver CNDR values (80 dB and 110 dB) where two-axis tracking was able to maintain lock. The fact that the system

¹⁶ This typically corresponds to a loop NEB of ~3.5 kHz, which was not measured.

rejection did not change as receiver CNDR was varied by 30 dB, a factor of 1000 change in received signal power, is further evidence that the AGC control loop was functioning properly. The measured rejection is -65 dB at 10 Hz and continues its two-pole roll-off at lower frequencies, leveling off at ~-90 dB rejection around 1 Hz. Useful tracking rejection (< -3 dB) is provided up to 500 Hz.

6. Conclusions

We believe that a heterodyne receiver architecture using an electro-optic nutator (EON) and a single-mode optical fiber can offer several improvements over existing mechanical nutation techniques, at the expense of a small additional coupling loss (1.3 dB measured and 0.5 dB as a goal). The EON is small, lightweight, rugged, and has no moving parts. It can operate over a broad range of frequencies and can be driven by a simple op-amp, requiring low power (77 mW in our design). We have demonstrated its performance in a laboratory heterodyne receiver set-up. The measured open-loop noise performance was found to be in good agreement with predictions. Our active noise cancellation scheme also extended the useful automatic gain control (AGC) range of the receiver down into the low receiver i.f. signal-to-noise (SNR) region by greater than 15 dB over an approach which does not use the method. Lastly, we were able to demonstrate the EON's use in a two-axis, 1 kHz open-loop crossover frequency, spatial tracking feedback loop. Two-axis tracking was maintained over a receiver CNDR range of 30 dB, corresponding to a factor of 1000 change in received signal power. Overall, our nutating receiver was able to offer good performance over its entire expected operating range.

REFERENCES

1. V.W.S. Chan, "Coherent optical space communication system architecture and technology issues," SPIE, vol. 295, Control and Communication Technology in Laser Systems, Aug 25-26, San Diego, CA, 1981.
2. T.E. Knibbe, E.A. Swanson and J.K. Roberge, "Spatial tracking using an electro-optic nutator and a single-mode optical fiber," SPIE vol 1635, Free-Space Laser Communications Technologies IV, Jan 23-24, Los Angeles, CA, 1992.
3. E.A. Swanson and R.S. Bondurant, "Using fiber optics to simplify free-space lasercom systems," SPIE Free-Space Laser Comm. Tech. II, vol 1218, pp. 70-82, 1990.
4. E.A. Swanson and J.K. Roberge, "Design considerations and experimental results for direct detection spatial tracking systems," Opt. Eng., 28/18, 6, 659-666 (1989).
5. J.E. Kaufmann and V.W.S. Chan, "Coherent optical intersatellite crosslink systems," IEEE, Sept. 1988.
6. J.E. Kaufmann and L.L. Jeromin, "Optical heterodyne intersatellite links using semiconductor lasers," IEEE Global Telecommunications Conf. Record, paper 28.4, Atlanta, GA, Nov. 26-29, 961-968 (1984).
7. L.G. Kazovsky, "Theory of tracking accuracy of laser systems," Opt. Eng. vol. 22, no. 3, 339-347 (1983).
8. K. Wallmeroth, et al., "Towards a coherent optical free-space communications system," Elect. Lett., vol. 26, no. 9, 26 Apr. 1990, pp. 572-573.
9. E.A. Swanson and R. S. Bondurant, "Fiber-based receiver for free-space coherent optical communications systems," OFC 1989 Technical Digest, Optical Fiber Communication Conference, Houston, Texas, February 6-9 (1989).
10. S.B. Alexander, "Design of Wide-Band Optical Heterodyne Balanced Mixer Receiver," IEEE J. Lightwave Tech. LT-5, 523-537 (1987).
11. J.E. Kaufmann and E.A. Swanson, "Laser intersatellite transmission experiment spatial acquisition, tracking, and pointing subsystem," MIT Lincoln Laboratory, Lexington, Mass., Project Report SC-80, (12 September 1989).
12. K.J. Held and J.D. Barry, "Precision pointing and tracking between satellite-borne optical systems," Opt. Eng., vol 27, No. 4, 325-333 (1988).

13. V.W.S. Chan, L.J. Jeromin, and J.E. Kaufmann, "Heterodyne LASERCOM systems using GaAs lasers for ISL applications," IEEE International Conference on Communications, ICC '83, June 19-22, Boston, MA., 1983.
14. E.A. Swanson et al., "Optical spatial tracking using coherent detection in the pupil-plane," J. of Appl. Opt., 28, 18, 3918-28 (1989).
15. R.S. Bondurant et al., "An opto-mechanical subsystem for space based coherent optical communication," in High Data Rate Atmospheric and Space Communication, T. Hauptman (ed.), Proc. SPIE 996, 92-100 (1988).
16. E.A. Swanson and V.W.S. Chan, "Heterodyne spatial tracking system for optical space communication," IEEE Trans. Commun. COM-34(2), 118-126 (1986).
17. T.S. Wei and R.M. Gagliardi, "Direct detection vs heterodyning in optical beam tracking," in Phase Conjugation and Beam Combining and Diagnostics, R.A. Fisher and I. Abramowitz, eds., Proc. SPIE 739, 189-196 (1987).
18. K. Goto, K. Araki, and K. Yasukawa, "Effects of antenna pointing errors on the design of heterodyne optical intersatellite communications," Elect. and Comm. in Japan, Part 1, vol. 74, no. 3, pp. 107-115, 1991.
19. D. Harres, "Analog dividers for acquisition and tracking normalization," SPIE Vol 1218, Free-Space Laser Communications Technologies II, 1990.
20. E.A. Swanson and R.S. Bondurant, "A space-based optical communication system utilizing fiber optics," MIT Lincoln Laboratory, Lexington, Mass, Technical Report TR-865 (ESD-TR-89-185) (9 November 1989).
21. U.A. Johann, H. Sontag and K. Pribil, "A novel optical fiber based conical scan tracking device," in Optical Space Communication II, J. Franz, ed., Proc. SPIE 1522, 243-251, 1991.
22. M. Johnson, "High-Precision Multiposition Singlemode Fiber Switch using Magnetic Levitation," in Fiber Optics, L.R. Baker, ed., SPIE Vol 734, 154-158, 1987.
23. A. Papoulis, *Probability, Random Variables, and Stochastic Processes - 3rd Edition* . New York: McGraw-Hill, 1991.
24. A. Yariv and P. Yeh, *Optical Waves in Crystals* . New York: John Wiley & Sons, 1984.
25. J. F. Nye, *Physical Properties of Crystals* . Oxford: The Clarendon Press, 1960.
26. W.K. Pratt, *Laser Communication Systems* . New York: John Wiley & Sons, 1969.

27. Katzman, M. (ed.), *Laser Satellite Communication* . Englewood Cliffs, NJ: Prentice-Hall, 1987.
28. M. Zahn, *Electromagnetic Field Theory - A Problem Solving Approach* . Malabar, Florida: Robert E. Kreiger Pub. Co., 1979.
29. A. Yariv, *Optical Electronics - 4th Edition* . Chicago: Holt, Rinehart and Winston, 1991.
30. L. C. Shen and J. A. Kong, *Applied Electromagnetism - 2nd Edition* . Boston: PWS Engineering, 1987.
31. J.W. Goodman, *Introduction to Fourier Optics* . New York: McGraw-Hill, 1968.
32. E. Merzbacher, *Quantum Mechanics - 2nd Edition* . New York: John Wiley & Sons, 1970.
33. J.K. Roberge, *Operational Amplifiers: Theory and Practice* . New York: John Wiley & Sons, 1975.
34. H.A. Haus and J.R. Melcher, *Electromagnetic Fields and Energy* . New Jersey: Prentice-Hall, 1989.
35. C.D. Motchenbacher and F.C. Fitchen, *Low Noise Electronic Design* . New York: John Wiley and Sons, 1973.
36. Philips Corp., Discrete Products Div. *Ferrite Materials and Components Catalog - 8th ed.*
37. Gradshteyn, I.S. and Ryzhik, I.M., *Table of Integrals, Series, and Products*. New York: Academic Press, 1980.
38. Beyer, W.H., *CRC Standard Mathematical Tables - 27th Edition*. Florida: CRC Press, Inc., 1984.

Appendix 1

For convenience, this section of the appendix has the compiled block/circuit diagrams for the electro-optic nutator (EON). Most of these figures have been presented and explained in the body of the text. The following figures, listed with their text figure number (if any), are included:

Figure A1.1 (1.4.1, 5.0.1, and 6.0.2): Nutating receiver block diagram.

Figure A1.2: Expanded nutating receiver block diagram.

Figure A1.3 (3.2.1): Optical bench layout schematic.

Figure A1.4 (4.3.1): Schematic for the nutating receiver AGC block.

Figure A1.5 (5.1.2): Nutating receiver demodulator schematic.

Figure A1.6: Oscillator schematic.

Figure A1.7 (5.3.1): Compensator schematic.

Figure A1.8: Current driver. (Designed by James K. Roberge - included for completeness).

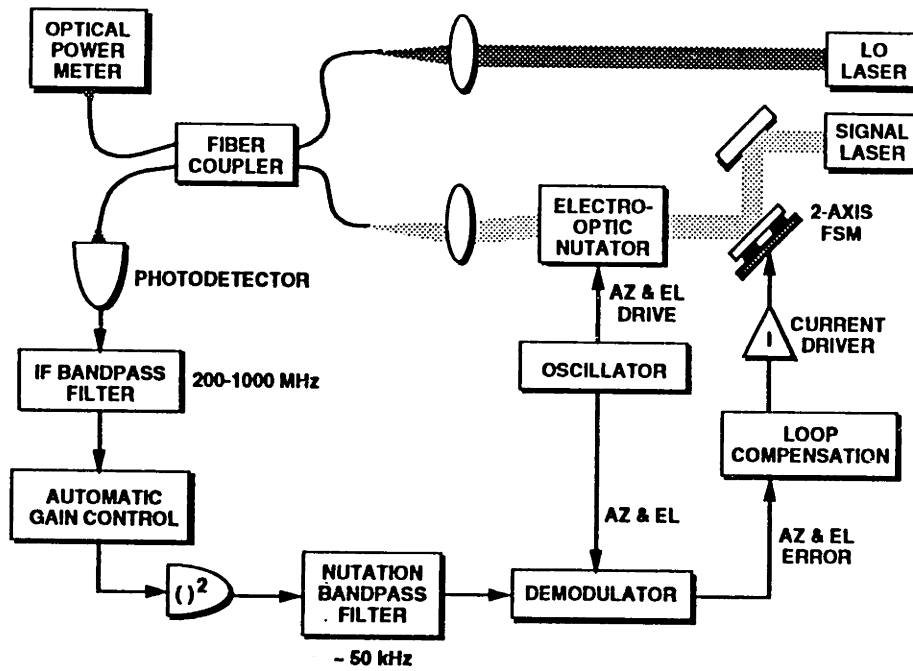


Figure A1.1: Nutating receiver block diagram.

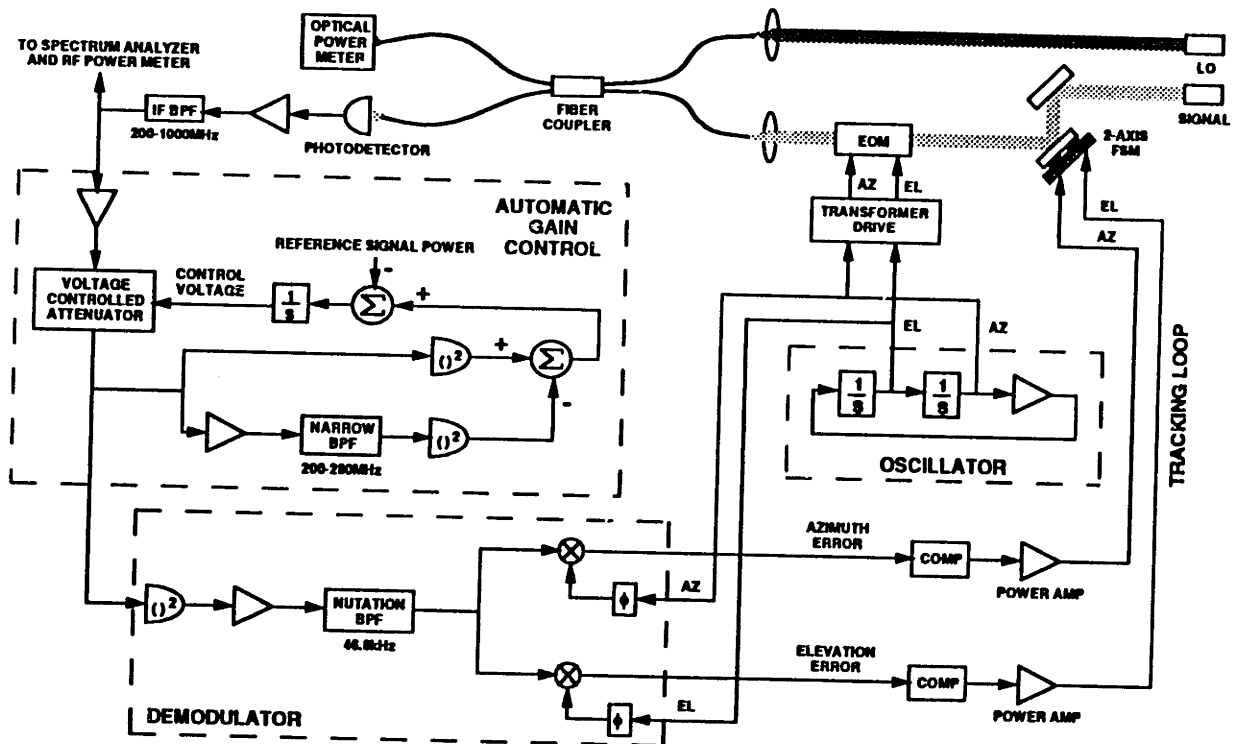


Figure A1.2: Expanded nutating receiver block diagram.

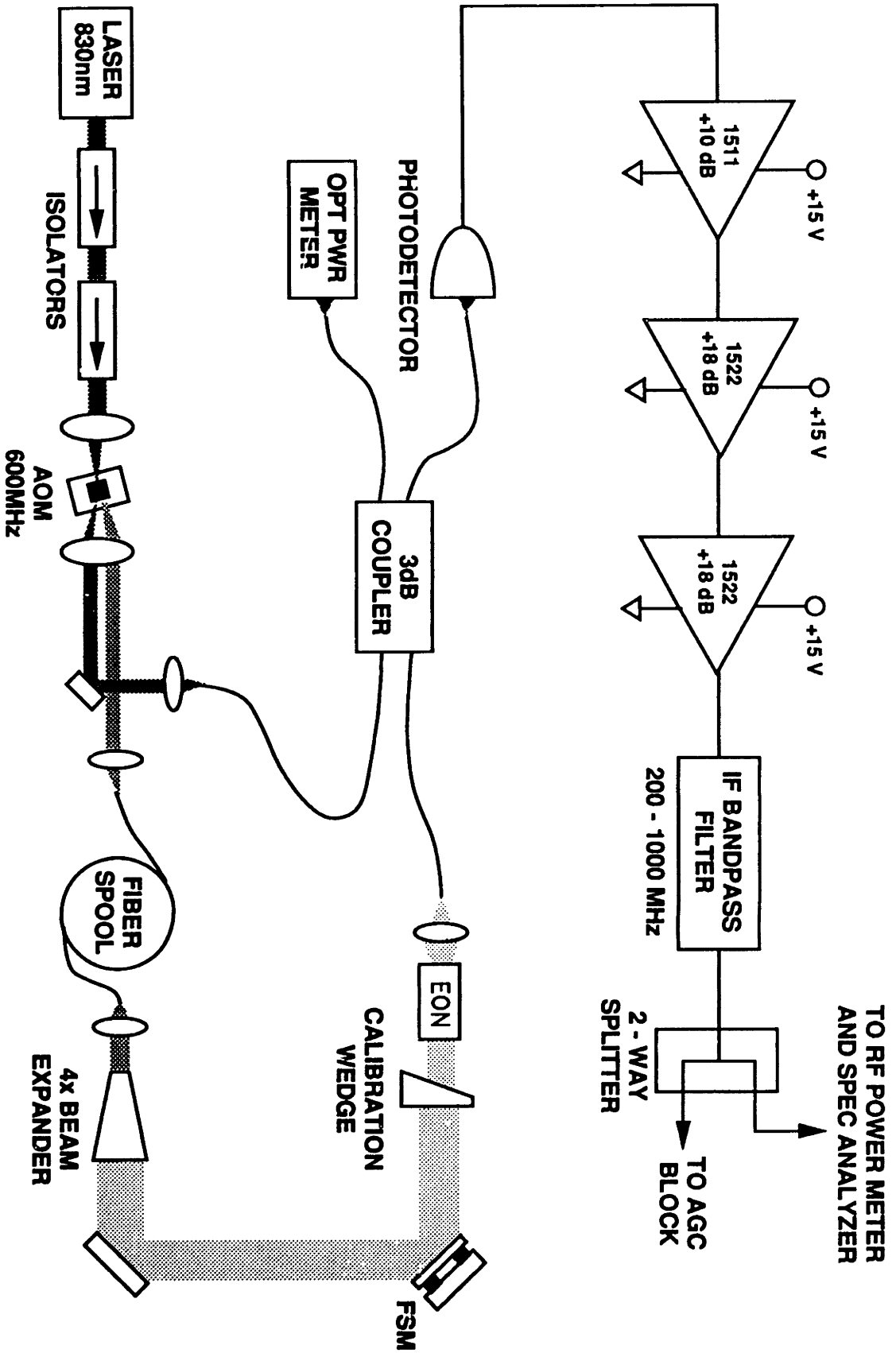


FIG A1.3: Optical bench layout schematic

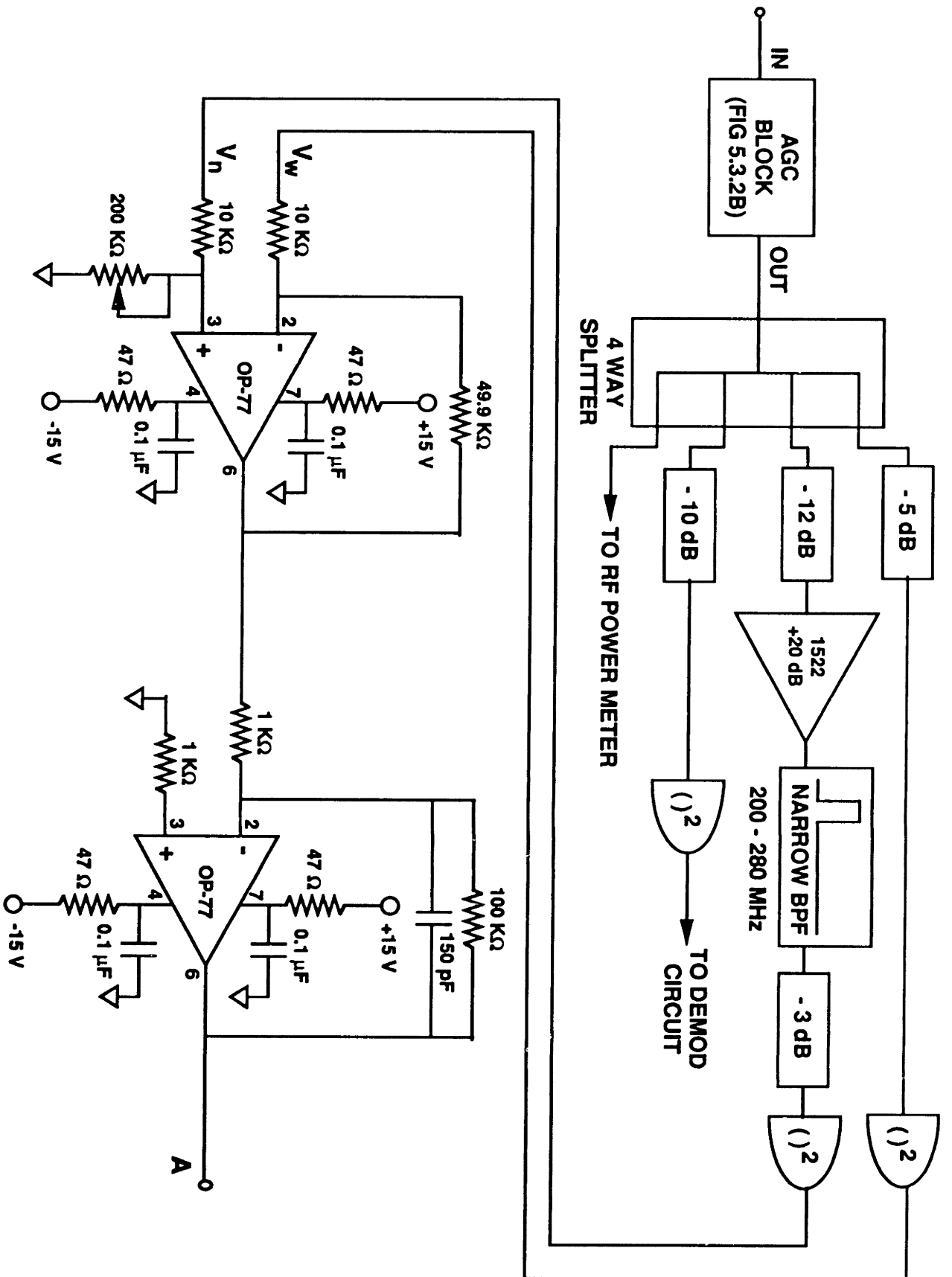


Figure A1.4A: Schematic for the rotating receiver AGC block.

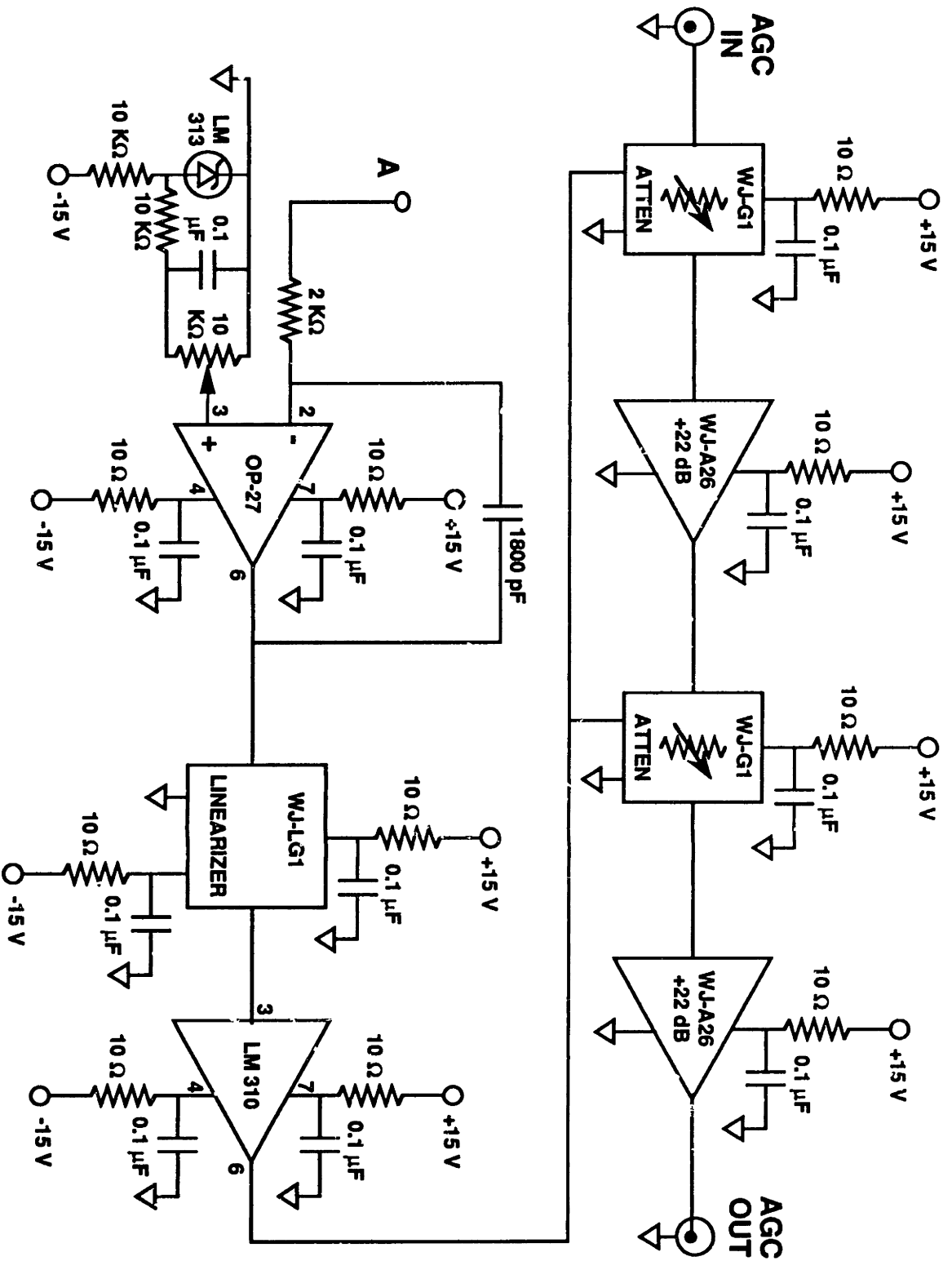


Figure A1.4B: Schematic for the nutating receiver AGC block.

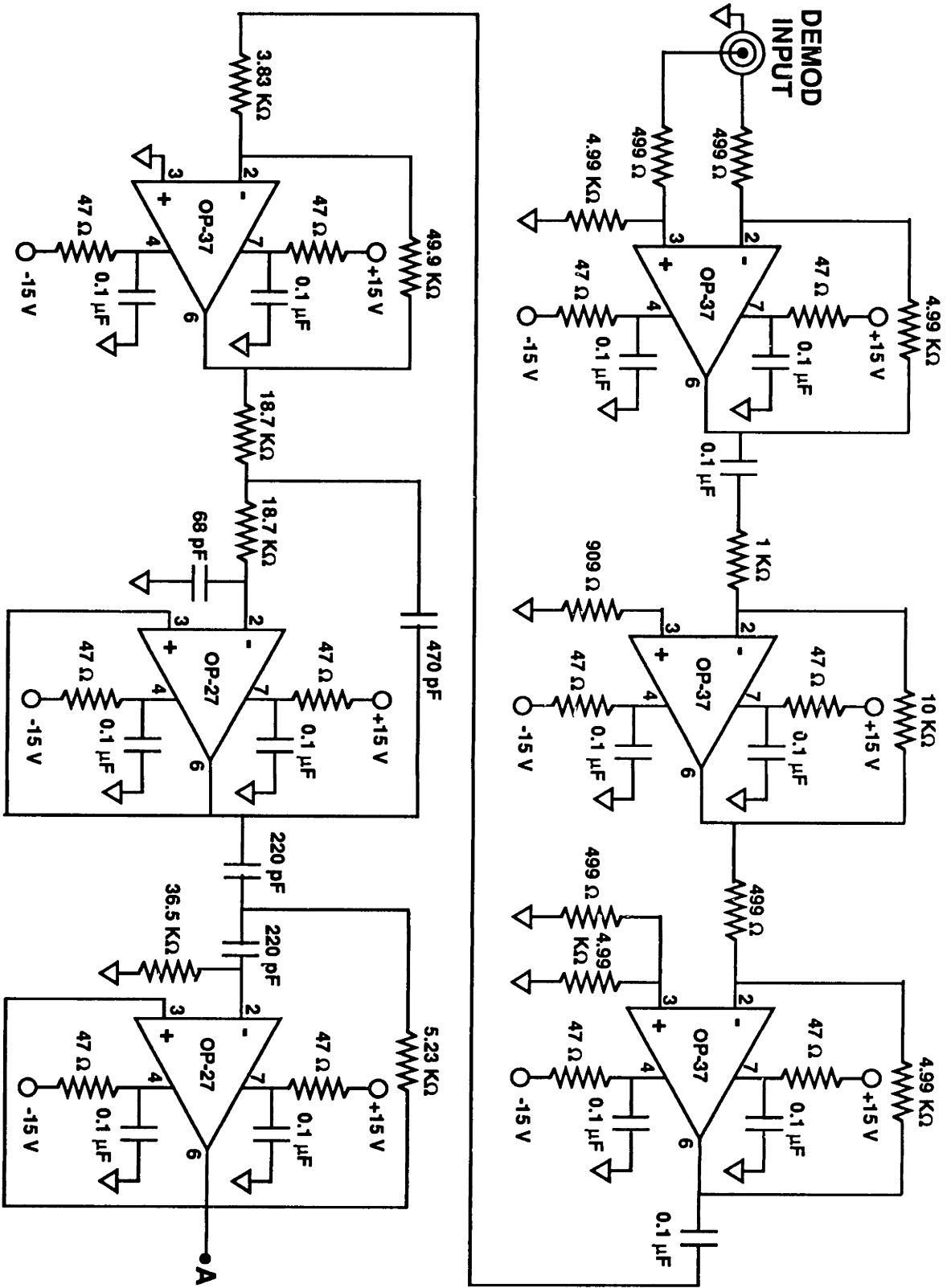


Figure A1.5A: Nutating receiver demodulator schematic

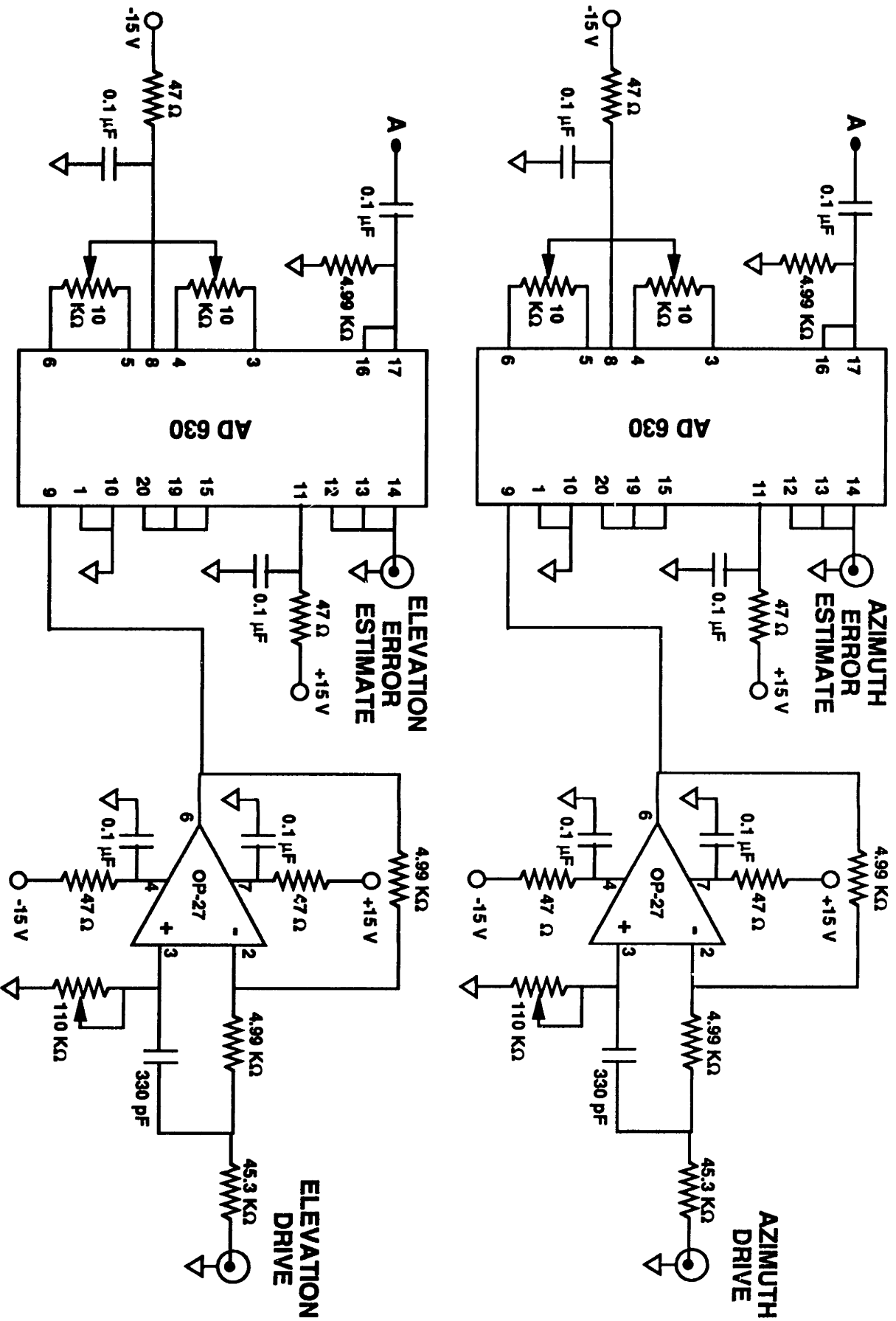


Figure A1.5B: Nutating receiver demodulator schematic.

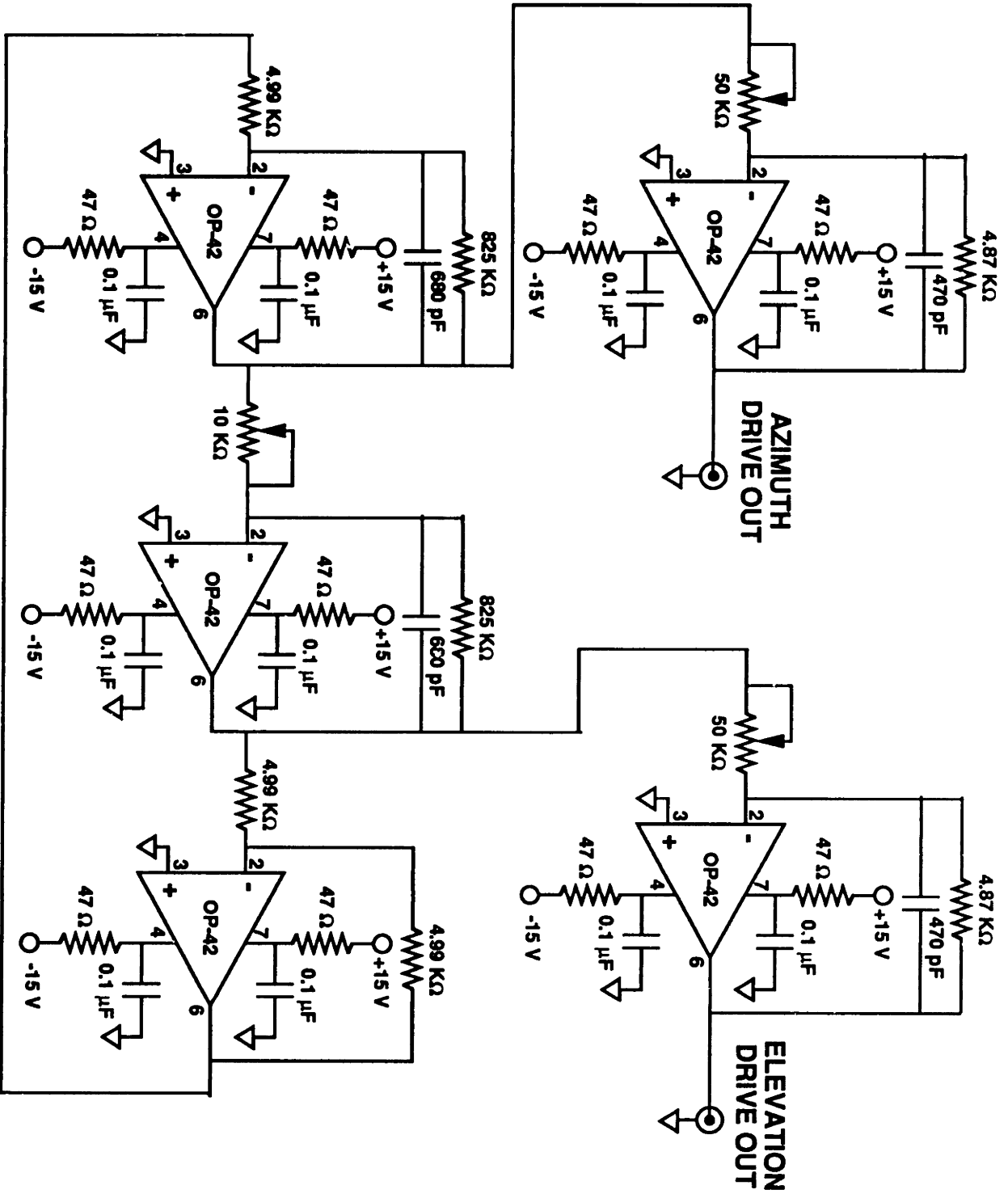


Figure A1.6 Oscillator schematic.

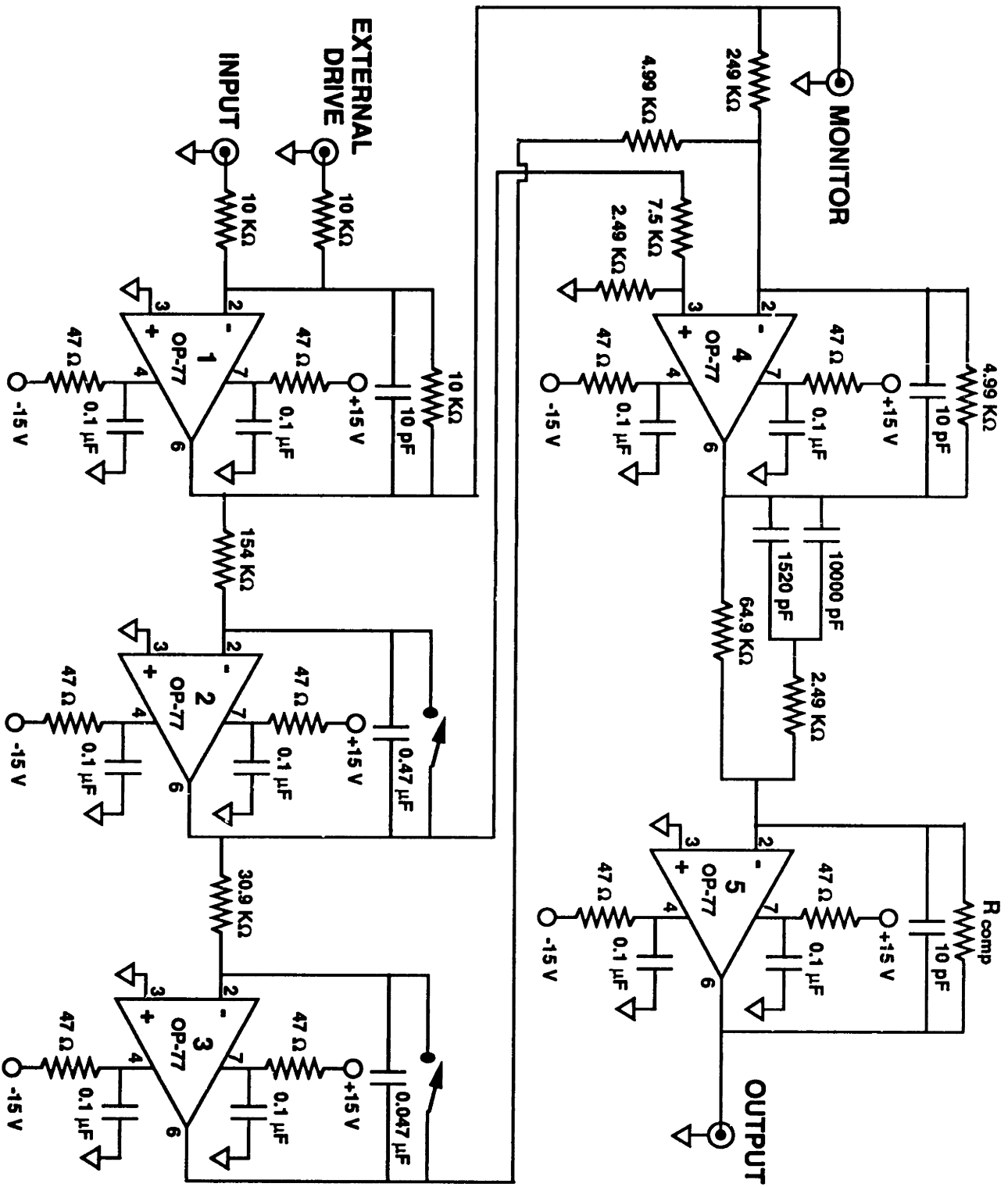
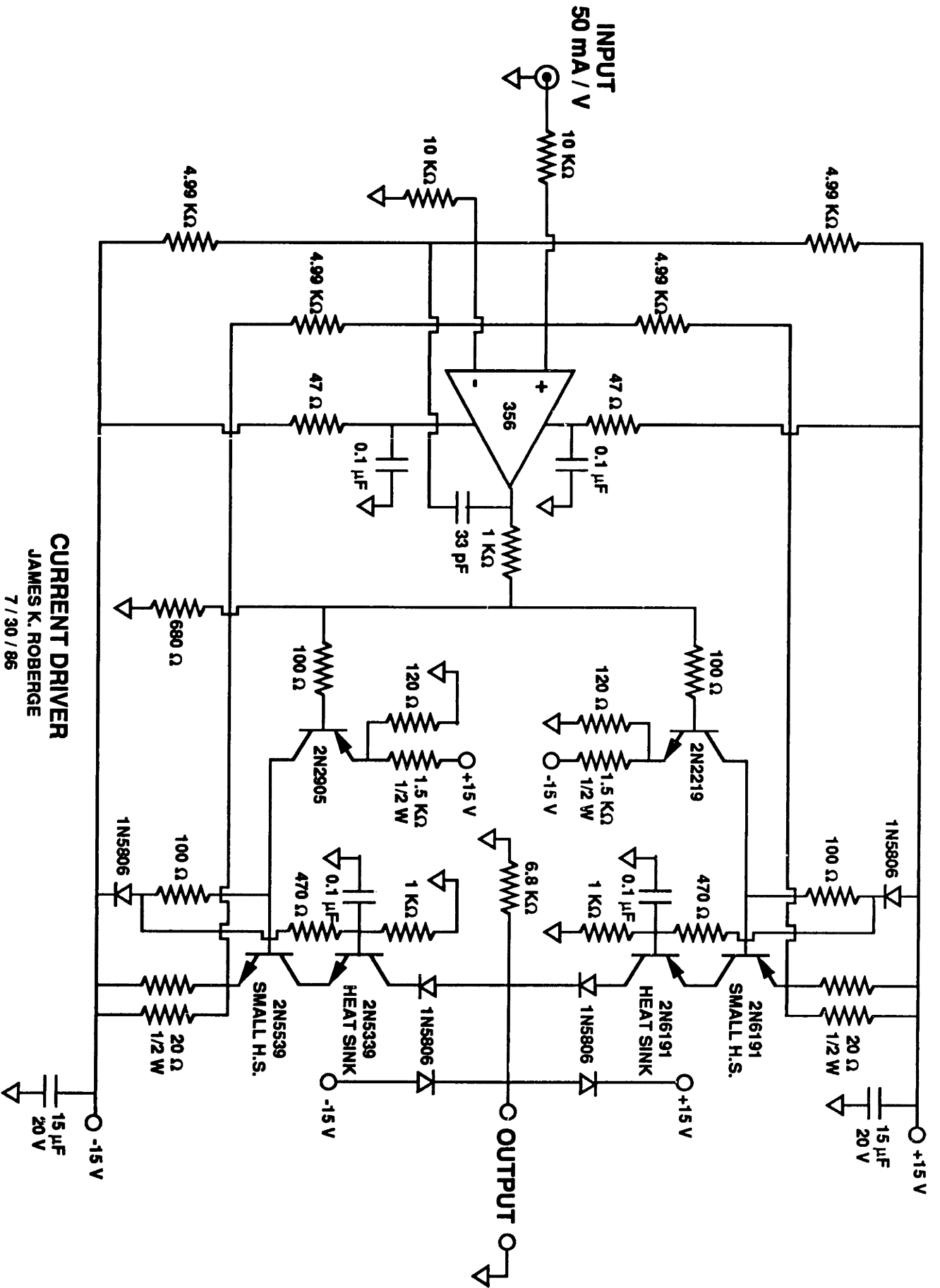


FIG A1.7: Compensator schematic.



CURRENT DRIVER
 JAMES K. ROBERGE
 7 / 30 / 86

Figure A1.8: Current driver schematic.

Appendix 2

This appendix shows some of the more obscure mathematical manipulations used to derive the receiver NESD (Sec. 3.1). Each part outlines the transformations used between successive equations, as noted, in the main text.

We begin with the overlap integral used to find the fiber coupling profile:

$$C(A, E, \phi) = \int_0^{2\pi} d\theta \int_0^{\infty} U(r) S_n(r, \theta) r dr \quad (3.1.16)$$

Into this equation we substitute the fiber mode profile (in the incident plane of the focusing lens of Fig. 3.1.3):

$$U(r) = \sqrt{\frac{2}{\pi\sigma}} \exp\left[-\left(\frac{r}{\sigma}\right)^2\right] \quad \text{where} \quad \sigma = \frac{\lambda f}{\pi\sigma_f} \quad (3.1.13)$$

and the nutated received signal field:

$$S_n(r, \theta) = \sqrt{\frac{4P_{sig}}{\pi d^2}} \text{circ}\left(\frac{r}{d/2}\right) \exp\left[-jkr\left\{\left[A - \phi \cos(\omega_n t)\right] \cos \theta + \left[E - \phi \sin(\omega_n t)\right] \sin \theta\right\}\right] \quad (3.1.12)$$

which gives:

$$C(A, E, \phi) = \frac{2\sqrt{2P_{sig}}}{\pi d\sigma} \int_0^{2\pi} d\theta \int_0^{d/2} r dr \exp\left[-\left(\frac{r}{\sigma}\right)^2\right] \exp\left[-jkr(\theta_x \cos \theta + \theta_y \sin \theta)\right] \quad (A2.1)$$

where $\theta_x = A - \phi \cos(\omega_n t)$
 $\theta_y = E - \phi \sin(\omega_n t)$
 $\sigma = \lambda f / \pi\sigma_f$

In order to get Eqn. A2.1 into a form which can be integrated we need to employ the following general integration property [p. 213 ref 37]:

$$\int_0^{2\pi} f[p \cos(x) + q \sin(x)] dx = 2 \int_0^{\pi} f[\sqrt{p^2 + q^2} \cos(x)] dx \quad (\text{A2.2})$$

where $f[x]$ is integrable on the interval $(-\sqrt{p^2 + q^2}, \sqrt{p^2 + q^2})$

Applying this property to Eqn. A2.1 yields:

$$\frac{4\sqrt{2P_{sig}}}{\pi d \sigma} \int_0^{d/2} r \exp\left[-\left(\frac{r}{\sigma}\right)^2\right] dr \int_0^{\pi} \exp\left[-jkr(\sqrt{\theta_x^2 + \theta_y^2} \cos(\theta))\right] d\theta \quad (\text{A2.3})$$

which can be integrated using [p.482 ref 37]:

$$\int_0^{\pi} \exp[j\beta \cos(x)] \cos(nx) dx = j^n \pi J_n(\beta) \quad (\text{A2.4})$$

to produce:

$$\frac{4\sqrt{2P_{sig}}}{d \sigma} \int_0^{d/2} r \exp\left[-\left(\frac{r}{\sigma}\right)^2\right] J_0(-kr\sqrt{\theta_x^2 + \theta_y^2}) dr \quad (\text{A2.5})$$

Finally, the coupling profile is found from Eqn. A2.5 by performing the following substitution:

$$r = \frac{d}{2} x \quad ; \quad dr = \frac{d}{2} dx \quad (\text{A2.6})$$

Once this is done, we receive the coupling profile shown in the text:

$$C(A, E, \phi) = \sqrt{2P_{sig}} \frac{d}{\sigma} \int_0^1 dx \exp\left[-\left(\frac{dx}{2\sigma}\right)^2\right] J_0\left[\frac{\pi x \sqrt{(A - \phi \cos(\omega_n t))^2 + (E - \phi \sin(\omega_n t))^2}}{\lambda/d}\right] \quad (\text{3.1.17})$$

We need to take a Taylor's series expansion of the coupling profile (Eqn. 3.1.17 shown above):

$$C(A, E, \phi) \approx C(0, 0, \phi) + A \left. \frac{\partial C(A, E, \phi)}{\partial A} \right|_{A=E=0} + E \left. \frac{\partial C(A, E, \phi)}{\partial E} \right|_{A=E=0} \quad (\text{3.1.18})$$

First, $C(0,0,\phi)$ is easily found by substituting $A = E = 0$ into Eqn. 3.1.17:

$$C(0,0,\phi) = \sqrt{2P_{sig}} \frac{d}{\sigma_0} \int_0^1 x dx \exp\left[-\left(\frac{dx}{2\sigma}\right)^2\right] J_0\left[\frac{\pi x \phi}{\lambda/d} \sqrt{\cos^2(\omega_n t) + \sin^2(\omega_n t)}\right] = \sqrt{P_{sig}} m(\phi) \quad (\text{A2.7})$$

To find $\partial C(A,E,\phi)/\partial A$ and $\partial C(A,E,\phi)/\partial E$ we need to use the generalized chain rule [p. 227 ref 38]:

$$\frac{d}{dx}[f[f(x)]] = \frac{d}{df(x)}[f[f(x)]] \frac{df(x)}{dx} \quad (\text{A2.8})$$

In our case A2.8 says:

$$\frac{d}{dA}[J_0[f(A)]] = -J_1[f(A)] \frac{df(A)}{dA} \quad (\text{A2.9})$$

because [p. 968 ref 37]:

$$\frac{d}{dz} J_0(z) = -J_1(z) \quad (\text{A2.10})$$

Applying A2.9 to $C(A,E,\phi)$ yields:

$$\begin{aligned} \left. \frac{\partial C(A,E,\phi)}{\partial A} \right|_{A=E=0} &= \frac{\sqrt{2P_{sig}} \pi \frac{d}{\sigma_0} \int_0^1 x^2 dx \exp\left[-\left(\frac{dx}{2\sigma}\right)^2\right] J_1\left[\frac{\pi x \phi}{\lambda/d}\right] \cos(\omega_n t) A}{\lambda/d} \\ &= \sqrt{2P_{sig}} AK(\phi) \cos(\omega_n t) \end{aligned} \quad (\text{A2.11})$$

and, similarly:

$$\begin{aligned} \left. \frac{\partial C(A,E,\phi)}{\partial E} \right|_{A=E=0} &= \frac{\sqrt{2P_{sig}} \pi \frac{d}{\sigma_0} \int_0^1 x^2 dx \exp\left[-\left(\frac{dx}{2\sigma}\right)^2\right] J_1\left[\frac{\pi x \phi}{\lambda/d}\right] \sin(\omega_n t) E}{\lambda/d} \\ &= \sqrt{2P_{sig}} EK(\phi) \sin(\omega_n t) \end{aligned} \quad (\text{A2.12})$$

Finally, because of the inherent symmetry in these expressions, the linearized coupling profile shown in the text results:

$$C(A, E, \phi) = \sqrt{P_{sig}} [m(\phi) + \sqrt{2}AK(\phi)\cos(\omega_n t) + \sqrt{2}EK(\phi)\sin(\omega_n t)]$$

$$\text{where } m(\phi) = \sqrt{2} \frac{d}{\sigma_0} \int_0^1 x dx \exp\left[-\left(\frac{dx}{2\sigma}\right)^2\right] J_0\left[\frac{\pi x \phi}{\lambda/d}\right] \quad (3.1.19)$$

$$K(\phi) = \frac{\pi \frac{d}{\sigma_0} \int_0^1 x^2 dx \exp\left[-\left(\frac{dx}{2\sigma}\right)^2\right] J_1\left[\frac{\pi x \phi}{\lambda/d}\right]}{\lambda/d}$$

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my parents, Tjeerd O. and Constance K. Knibbe, for without them none of this would have been possible. Thanks to my thesis supervisors, Professor James K. Roberge at MIT and Eric A. Swanson of the Optical Communications Technology Group (67) at MIT Lincoln Laboratory. I am grateful to both the 6A Internship Program and MIT Lincoln Laboratory for the opportunities and resources provided to me. Thanks also go to many other members of the Optical Communications Group at Lincoln; special mention to Jeffrey C. Livas for his helpful hints throughout my thesis research and his detailed readings and constructive criticisms of the rough drafts of this manuscript. Thanks to Richard Koch from Conoptics, Inc. for his help with the derivation of the EON deflection sensitivity. Thanks to Brian J. Lucia who constructed most of the electronic circuits and Ron E. Sprague who machined many of the mechanical components used in this research.

Special thanks to the Optical Communications Group at MIT Lincoln Laboratory for the opportunity to present this research at the SPIE Free-Space Laser Communications Technology IV Conference. Thanks to many members of the group for their constructive criticisms of the dry runs for my presentation.

This work was sponsored by the Department of the Air Force.