

LINEARIZED AND HIGH FREQUENCY
ELECTROOPTIC MODULATORS

Thesis by

Uri Cummings

In Partial Fulfillment of the Requirements for the
degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2005

(Defended April 27, 2005)

© 2005

Uri Cummings

All Rights Reserved

To the Memory of

Francine Sachs Cummings (1940 - 1995)

My mother,

who encouraged me every day,

always expressed spirited love,

and wanted more for me than for herself.

ACKNOWLEDGEMENTS

I thank Professor William B. Bridges, my graduate advisor. Bill taught me as an undergraduate, and then he invited me into his group for graduate studies. Bill took an active interest in my (and all of his students') professional development and personal well-being. Whether on my behalf, or that of so many others, he sets a constant example of dedication and integrity. It was a pleasure to work for him. He has an incredible scientific intuition, a sense of humor, and a rich set of experiences that he shared from Caltech, Hughes, Berkeley, and many other places. After I took an extended leave to start Fulcrum Microsystems, Bill remained very supportive of my finishing my thesis. He spent a lot of time with me and invited me to his homes in Nevada City and Sierra Madre. No person in my professional life has had a larger impact.

I thank Linda McManus, Bill's spouse, who was also very eager to see me finish. She invited me to dinner many times and showed a keen interest in my well-being as well as that of all of Bill's other students.

This research was done under contract F30602-96-C-0020 from the U.S. Air Force Rome Laboratory and earlier under contract F30602-92-C-0005. I thank the late Brian Henderickson and the late Norman Bernstein who supported us both financially and morally.

I thank all the people who helped in our laboratory. Our technician, Reynold Johnson, helped me with numerous experiments. His craft is first-rate, as is his humor. Lee Burrows, Bill's other student, did a disproportionate amount of the laboratory setup.

I thank our collaborators at Hughes Research Labs. Jim Schaffner worked closely with us, collaborating on modulator development and our peer publications. Joe Pikulski fabricated our modulators.

I thank Axel Scherer for the generous use of his laboratory, and for help with difficult photolithographic tasks.

I thank Charlie Cox at MIT Lincoln Labs who is a leader in the field and visited us many times and shared his insights.

I thank and remember Ed Poser. Ed taught an insightful class, and did so with great wit. He gave me a Summer Undergraduate Research Fellowship (SURF) after just that one class. I was a transfer student, still learning my way around Caltech. On the first day of the SURF, Ed lost his life in a tragic accident. Bill offered to take me on, and gave me a new SURF project that eventually led to my graduate studies in his group.

I thank Carolyn Ash who managed the SURF program. She matched contributions to support my SURF, and she showed interest in the details of research.

I am happy with the formative years I spent at Caltech, 1992-1999. In particular I took many wonderful classes. Some professors taught so well that it seemed almost more important to take a class if taught by a particular professor than for the subject matter, as long as the class contributed to the department course requirements! Some, but not all of those professors, and in no particular order, are P.P. Vaidyanathan, David Rutledge, David Middlebrook, Alain Martin, R. J. McEliece, Amnon Yariv, Ed Posner, Joel Franklin, and of course Bill Bridges.

I thank my colleagues at Fulcrum for being patient with me, when I had to reschedule work with them in order to work on my thesis, and for generally being supportive, in particular, Bob Nunn, Alain Gravel, Mike Zeile, and Andrew Lines.

Many fellow students made Caltech a wonderful place, whether for intellectual excitement, social activities, or just silliness (pranks). I thank Andrew Lines, Ted Turocy, and Mika Nystrom, a core group of my friends.

My family has been very supportive of me. I thank my father, Robert, my brother Theo, and my Aunt, Marguerite, and all of my father's extended family. I thank and remember my mother Francine.

ABSTRACT

An analysis is performed of many standard and linearized electrooptic modulators known in the industry. The transfer functions of these modulators are evaluated under a consistent set of performance figures of merit, which are gain and spur-free dynamic range, using a canonical set of optical link parameters. The tolerance of the needed precision of the parameters of the linearization mechanisms of all of these modulators is compared over the entire interesting range of noise bandwidth.

A computer program was written to analyze the frequency dependence of any modulator transfer function under any set of functional inputs. The program is used to illustrate and compare the frequency dependence of the figures of merit of all of the modulators for which a d-c analysis was performed. Further analysis looks at the effect of greater noise-bandwidth and recovering the frequency-dependent degradation of gain and dynamic range through re-phasing techniques. The gain of directional couplers is analyzed in-depth.

Two novel modulator schemes are produced. The first uses reflective wave techniques to retime the electrical and optical waves half way through the modulator. The second uses fabrication geometry and properties of the linearization technique to make a more robust modulator (applicable to three of the modulators analyzed).

A 94 GHz antenna-coupled directional coupler modulator was initially demonstrated using an old modulator chip from Finbar Sheehy. A peculiar bug with the chip was uncovered. And a new modulator experiment was constructed and many aspects of the experimental apparatus were optimized. Though the revised experiment ultimately did not yield modulation side bands, it did couple a 94 GHz microwave signal into the optical waveguide, and many interesting challenges of high frequency electrooptic modulator fabrication were evaluated and improved upon.

TABLE OF CONTENTS

Acknowledgements	iv
Abstract	vi
Table of Contents	vii
List of Figures	ix
List of Tables	xi
Chapter I: Introduction	1
Thesis Organization	3
Chapter II: Electrooptic Modulator Fundamentals	5
The Electrooptic Effect	6
Electrooptic Modulation in Bulk Crystals	7
Waveguide Phase Modulators	10
Intensity Modulators Based on Mach-Zehnder Interferometers	11
Waveguide Electrooptic Modulators Based on Directional Couplers	14
Fidelity in Electrooptic Modulators	17
Frequency and Bandwidth Limitations in Electrooptic Modulators	19
Increasing Frequency Response of E-O Modulators by Velocity Matching	21
Chapter III: Distortion in Modulation	
Abstract	25
Link Model	26
Analysis with One and Two Tones	27
Linearized Modulators	33
Super-Octave Modulators	33
The Ideal Limiter	34
The Dual Parallel Mach-Zehnder Modulator	36
The Cascade Mach-Zehnder Modulator	37
The Directional Coupler with Two Passive Biases	39
The Y-Fed Directional Coupler	40
Sub-Octave Modulators	42
The Sub-Octave Dual Series Mach-Zehnder Modulator	43
The Sub-Octave Directional Coupler Modulator	43
Comparison of SoDCM and DSMZM	44
Others Modulators	45
Performance of Electrooptic Modulators	46
Ideal Performance	46
Noise Scaling	52
Parameter Sensitivity	54
Multi-Tone Analysis	62
Triple Beat Calculations for Electrooptic Modulators	64

Chapter IV: Bandwidth of Linearized Electrooptic Modulators	
Abstract	69
Finite Transit Time Computation.....	69
Background.....	69
Wave Velocity Mismatch.....	70
Mathematical Description of the Modulator Analysis Program.....	71
Non-Linearized Modulators: The Standard MZM, DCM, and YFed-DCM.....	74
Y-Fed Directional Coupler Modulator.....	78
Broad-Band Linearized Modulators	80
Frequency Independence of the DPMZ.....	81
Frequency Dependence of the CMZM.....	82
Frequency Dependence of the YFDCM.....	82
Frequency Dependence of the DCM2P	83
Sub-octave Linearized Modulators.....	84
The Effects of Noise Bandwidth	86
Re-optimization of Sub-octave Modulators for Band-Pass Applications	92
Bandwidth Recovery with Periodic Rephasing.....	95
Bandpass Features of Directional Coupler-Based Modulators.....	103
The gain of a Y-Fed Directional Coupler.....	104
The Gain of a Standard Directional Coupler Modulator	109
 Chapter V: Modulator Improvements	
Abstract	112
Reflective Traveling Wave Modulators	113
Traveling Wave Reflective Modulators.....	114
Application to Useful Modulators	116
Dynamic Range of Reflective Modulators	117
Conclusion.....	119
Sensitivity Splits	120
Sensitivity Split in the Dual Parallel Mach-Zehnder	120
Sensitivity Split in other Mach-Zehnder based linearized modulators	125
 Chapter VI: Millimeter-wave Directional Coupler Modulators	
Abstract	127
An Introductory Note.....	128
Background Work up to the Fall of 1994.....	134
Modulator Fabrication Work Beginning in the Fall of 1994	135
94 GHZ Directional Coupler Modulator Fabrication	137
R-F Waveguide and Coupling Measurements	141
94 GHz Measurements	150
 Chapter VII: Suggestions for Future Work	
Theoretical Analysis.....	155
Experimental Work	157

LIST OF FIGURES

<i>Number</i>	<i>Page</i>
Figure 2 - 1: a) Bulk crystal modulator in transverse configuration.....	9
Figure 2 - 2: Phase modulator based on in-diffused waveguides in an electrooptic crystal.	12
Figure 2 - 3: Schematic cross-section view of a phase modulator.....	13
Figure 2 -4: Mach-Zehnder modulator based in indiffused waveguides.....	14
Figure 2 - 5: Mach-Zehnder modulator transfer function.	15
Figure 2 - 6: Directional coupler modulator based on in-diffused waveguides.....	16
Figure 2 - 7: Fields in a directional coupler modulator looking down the waveguides.....	17
Figure 2 - 8: Transfer functions of both arms of a directional coupler modulator.....	18
Figure 2 - 9: Schematic of antenna-coupled phase modulator.....	23
Figure 3 - 1: Canonical model of an optical link for evaluating external modulators.....	27
Figure 3 - 2: The signal (gain) and third-order intermodulation product.	32
Figure 3 - 3: A “Darwin chart” of many proposed linearized modulators.....	34
Figure 3 - 4: Five superoctave modulators.	35
Figure 3 - 5: Two sub-octave modulators.	45
Figure 3 - 6: Power sweep of signal, second harmonic, and intermodulation distortion.	47
Figure 3 - 7: Power sweep of DCM2P.	49
Figure 3 - 8: Power sweep of the Signal, Second harmonic and intermodulation distortion....	50
Figure 3 - 9: Dynamic range as a function of signal bandwidth.	54
Figure 3 - 10: The dynamic range of the simple MZM.	56
Figure 3 - 11: The dynamic range of the SoDCM.....	57
Figure 3 - 12: Lines of equal dynamic range reduction with bias error in a MZM.	58
Figure 3 - 13: Time domain samples of a 30 tones in a Mach-Zehnder modulator.	65
Figure 3 - 14: Power sweeps of 30 tones in a MZM.	66
Figure 3 - 15: 30 tone CTB in a linearized modulator.	67
Figure 4 - 1: The convergence of the calculation of the gain of a simple MZM.....	74
Figure 4 - 2: Gain and dynamic range of a standard Mach-Zehnder modulator.	77
Figure 4 - 3: Gain and dynamic range of a simple directional coupler modulator.....	78
Figure 4 - 4: Non-linearized Y-fed DCM gain and dynamic range.....	79
Figure 4 - 5: Dynamic range comparison of four broadband linearized modulators.....	80
Figure 4 - 6: Dynamic range comparison of four broadband linearized modulators.....	81
Figure 4 - 7: Dynamic range comparison of two sub-octave linearized modulators.....	85
Figure 4 - 8: Dynamic range comparison with 1 kHz signal bandwidth.....	88
Figure 4 - 9: Dynamic range comparison with 1 kHz signal bandwidth.....	89
Figure 4 - 10: Dynamic range comparison with 1 MHz signal bandwidth.....	90
Figure 4 - 11: Dynamic range comparison with 1 MHz signal bandwidth.....	91
Figure 4 - 12: Dynamic range comparison with 1 GHz signal bandwidth.....	92
Figure 4 - 13: Re-optimization of the SoDCM by slightly adjusting the bias.	93
Figure 4 - 14: Re-optimization of the SoDSMZM.....	94

Figure 4 - 15: The dynamic range of the SoDCM re-optimized	95
Figure 4 - 16: Dynamic range versus frequency for the DCM.	96
Figure 4 - 17: Dynamic range versus frequency for the DCM2P.....	98
Figure 4 - 18: Dynamic range of rephased YFDCM ($\theta=1.43\pi$).....	99
Figure 4 - 19: Dynamic range of a rephased CMZM.....	100
Figure 4 - 20: Dynamic range of the SoDSMZM.	101
Figure 4 - 21: Dynamic range versus frequency for the SDCM.....	102
Figure 4 - 22: The gain of a standard MZM with multiple electrode segments.	103
Figure 4 - 23: The gain of a Y-fed DCM.	105
Figure 4 - 24: The transfer function of Y-fed DCMs.....	106
Figure 4 - 25: The gain of a Y-fed DCM	108
Figure 4 - 26: The gain of DCM.....	110
Figure 4 - 27: Gain of the DCM.....	111
Figure 5 - 1: Phase modulator in reflective traveling wave configuration.	114
Figure 5 - 2: MZM and DCM reflection modulators	117
Figure 5 - 3: DSMZM reflection modulators.....	118
Figure 5 - 4: The optimal optical split as a function of the RF split.	121
Figure 5 - 5: The optimal dynamic range.....	122
Figure 5 - 6: A DPMZ with a precise optical split.	123
Figure 5 - 7: The waveguide an electrode structure of a sensitivity split.....	126
Figure 6 - 1 Actual electrode mask used in FTS-1 design.....	129
Figure 6 - 2: D-c transfer function of Sheehy's DCM.....	130
Figure 6 - 3: (a) Curve fit to experimental data.	131
Figure 6 - 4: Oscilloscope Polaroid of the operation of directional coupler.....	133
Figure 6 - 5: "Mouse-bite" defects (a) on LINC-1 directional coupler.....	135
Figure 6 - 6: Electrode mask UVC-el-1	136
Figure 6 - 7: Photograph of the modulator mounting structure.	139
Figure 6 - 8: Modulator chip mounted on the mechanical carrier stage	140
Figure 6 - 9: Blown up view of the modulator chip and millimeter-wave slab guide.....	141
Figure 6 - 10: Slab dielectric waveguide coupling scheme for millimeter-wave modulators....	142
Figure 6 - 11: Photograph of a "back-to-back" coupling structure.	143
Figure 6 - 12: Plot of dielectric constant correction factor.....	144
Figure 6 - 13: a) A layer of Stycast between two polypropylene feeds.....	146
Figure 6 - 14: A slab layer of Stycast 5 between two polypropylene feeds.....	148
Figure 6 - 15: Photograph of the overall modulator measurement setup.	150

LIST OF TABLES

<i>Number</i>	<i>Page</i>
3-1: Parameters of the canonical link defined for this study	28
3-2: Gain, 1-Hz DR, 4-MHz DR, and noise figure	51
3-3: Parameter tolerance of standard and linearized modulators	59
6-1: R-f feed optimization results	149

Chapter 1

INTRODUCTION

In the last few decades the field of optics has become very important. Especially with the invention of optical fiber, technologies for the manipulation and transport of coherent light offer great advantages in a wide range of applications. Optical systems are used to send information from as long a distance as across oceans, to as short a distance as the machine room or even the back-plane in an equipment chassis. There are commercial fiber-optic products in telecom voice & data, computer networking, cable TV, and scientific measurement. In the defense industry fiber-optic systems are used in similar applications, and in addition, they are used in antenna remoting and some advanced weaponry. Fiber-optics promises significant bandwidth, above a terahertz, and low transport loss, 0.1 dB/km, all in a “light” cable.

A basic fiber-optic data link consists of an electrical amplifier, a laser, a means of modulating the optical signal with the electrical signal, a length of fiber, an optical detector, and an electrical post-amplifier. The “intrinsic link” is the laser, modulator, fiber, and detector. Link engineers select a set of optical components that when combined form a fiber-optic link to optimally meet the system requirements. In component research and engineering, the aim is to improve the figures of merit of particular components of the fiber-optic link. In this thesis electrooptic modulators are evaluated in the context of a normalized intrinsic link. The application scope is 1-100 GHz analog links used for short range military applications, such as antenna remoting. However, many of the results apply to links more broadly in other applications.

The electrooptic modulator is often the limiting component of an analog fiber-optic link. The important figures of merit for the analog link are: gain, noise figure, dynamic range, and modulation bandwidth. The gain of the link is defined as the electrical r-f output power of the photodetector divided by the input r-f drive power of the modulated signal. Note that it is

possible for even an intrinsic link to have net positive power gain if sufficient optical laser power is used, as shown by Cox in Ref. [1.1]. The noise figure is the noise power out of the link divided by GkT_0B (link gain, Boltzmann's constant, room temperature taken as 290K, and noise bandwidth), usually expressed in dB. Dynamic range is the peak power ratio of the signal to the dominant distortion term, also usually expressed in dB. Bandwidth (sometimes termed "noise bandwidth") is the electrical frequency range passed by the link. It often is not possible to simultaneously optimize all of the figures of merit, so trade-off analysis is employed to evaluate modulators for different link configurations.

There are two different approaches to optical modulation in the fiber-optic link, *direct modulation of the laser* and *external modulation of the laser output*. Many links employ direct modulation, in which the modulation voltage is applied to the power supply of the laser. While simple, direct modulation is undesirable in noise-sensitive applications because it creates chirp, phase modulation that can be converted to amplitude modulation distortion in the receiver. In high fidelity links, the modulation function is usually separated from the laser into an external component. There are three types of external modulators in common use today: electro-absorption modulators based on the semiconductors gallium arsenide or indium phosphide, electrooptic modulators based on polarized polymers, and electrooptic modulators based electrooptic crystals such as lithium niobate, and other materials. The theoretical calculations in this thesis apply to all forms of high fidelity modulators, while the laboratory work focuses on high frequency LiNbO_3 electrooptic modulators.

The research undertaken for this thesis treats two important aspects of electrooptic modulation: The first aspect is an analytical treatment of various schemes to improve the linearity of electrooptic modulators for use in high dynamic range analog microwave fiber-optic links, with particular attention to the bandwidth limitations imposed by the linearization scheme employed. The second aspect studied is an experimental demonstration with directional coupler modulators of the antenna-coupling scheme used by Sheehy, et al in Ref. [1.2] for Mach-Zehnder modulators.

Thesis Organization

Chapter 1 introduces the thesis topic.

Chapter 2 gives a brief tutorial overview of the different kinds of lithium niobate electrooptic modulators treated in this research. The Mach-Zehnder modulator (MZM) is by far the most common configuration used today, but the directional-coupler modulator (DCM) also offers some attractive characteristics. The linearity and frequency limitations of both these modulators are introduced.

Chapter 3 addresses distortion and its reduction in various electrooptic modulator configurations. This work follows the work done in Ref. [3.8] but adds additional modulators and additional comparisons. A consistent link model is defined against which to compare many different modulators and linearization schemes. Gain, noise figure, and dynamic range are defined and calculated numerically for all of the different modulators, as appropriate for a single analog signal. The modulators are further compared using composite triple beat and composite second-order distortion, the figure of merit for CATV, where multiple channels are transmitted over the link. Finally, the modulators are compared for their sensitivity to manufacturing and bias parameters.

Chapter 4 addresses the bandwidth limitations of different modulators and linearization schemes. A numerical modulator simulation computer program is defined. The program calculates the gain and dynamic range for the modulators defined in Chapter 2. The question of whether a linearization scheme can be reoptimized for a particular frequency within the band-pass for high frequency operation is answered. The effect of multiple electrode segments, “rephasing on the average,” on dynamic range is analyzed. Much of this work was published in Ref. [4.2], but there are some extensions to this work that appear originally here.

Chapter 5 introduces two modulator scheme improvements. First it describes reflective-wave modulators, which use reflections at the crystal surface for the optical wave and an open circuit in the traveling wave electrode at a specific location to gain the effects of twice the crystal real estate and a one-time rephrasing of electrode segments. This is “velocity matching

on the average” with a two segment electrode. The technique is broadly applicable to different modulators. It is covered in U.S. Patent 5,076,655. The second modulator improvement uses “sensitivity splits.” This technique applies to two well-known linearized modulators. It significantly reduces the complexity of the modulator design by building asymmetry into the modulator, and then taking advantage of the fact that the modulators only need one precisely controlled degree of freedom, whereas the published version of the prototype modulator structure calls out for two precisely controlled degrees of freedom.

Chapter 6 covers the laboratory work done on high-frequency antenna-coupled external modulators. A 100 GHz directional coupler modulator is demonstrated. This experiment follows on work done by Sheehy in Refs. [1.2] and [1.3]. A set of directional coupler modulators with different electrical lengths and different modulator electrode lengths (physical length) was fabricated and characterized. The r-f feed structure for antenna-coupled modulators was optimized. A successful demonstration of the basic operation of this device was made very early in the work, and further improvements were carried out. Unfortunately, this portion of the study had to be abandoned (due to events beyond our control) before the improved device could be fully demonstrated.

Chapter 7 covers suggestions for future work. This is broken into a laboratory component and a theoretical component.

Chapter 2

ELECTROOPTIC MODULATOR FUNDAMENTALS

The Electrooptic Effect

Electrooptic modulators of the variety studied in this thesis exploit the Pockel's effect. In crystals such as LiNbO_3 , the index of refraction varies linearly with the applied electric field, with the constant of proportionality depending on the orientation of the electric field with respect to the crystallographic axes. The resulting change in the optical path length of the light passing through the crystal is proportional to the voltage applied to a set of electrodes. Over a fixed physical length of crystal, this converts to a phase shift of the light passing through it.

The effect of an applied electric field on the propagation of light is defined by the change of the indices of refraction in the tensor equation $\Delta(1/n^2) = \mathbf{R}\mathbf{E}$. $\Delta(1/n^2)$ gives the change in the index of refraction for the elements of the tensor Δ , \mathbf{R} is the tensor of electrooptic coefficients, and \mathbf{E} is the vector of the applied electric field. (See, for example Ref. [2.1]). There are six relevant indices of refraction from the index ellipsoid, three terms for the three principle directions (x, y, and z), and three cross terms. An electric field has x, y, and z components, so in general this effect is described with a 6 x 3 tensor, \mathbf{R} . Each component $\Delta(1/n^2)_i$ for $1 < i < 6$ is the dot product of the i-th row of \mathbf{R} with the electric field vector. In non-electrooptic crystals (centrosymmetric), all of the terms r_{ij} are zero. Most electrooptic crystals have some degree of symmetry leading to many of the r_{ij} terms equaling zero. The electrooptic effect is generally very small, but the crystal lithium niobate (LiNbO_3) possesses relatively large values of these electrooptic coefficients, and has other desirable properties (hardness, non-hydroscopic, etc.). As well, LiNbO_3 is a “3m” symmetric crystal and has 8 non-zero terms, of which there are 4 distinct absolute values, some terms are equal, and some are equal and opposite. For optical wavelengths of interest in this study¹ the electrooptic constants are (in picometers per volt):

- $r_{13} = r_{23} = 8.6 \text{ pm/V}$,

¹ These values are taken from Ref. [2.1] for the 0.633 μm . That reference does not give the constants for 1.3 μm or 1.5 μm . The same values for r_{33} and r_{13} are listed in Ref. [2.3] but in that reference they are given for 1.5 μm . Accordingly, the values are largely constant over 0.6 μm to 1.5 μm wavelengths. They are about 10-20% smaller at 3.3 μm . It is also assumed that the modulating frequency is above 10 MHz. At lower frequencies, some of these constants can change considerably due to mechanical effects.

- $r_{22} = -r_{12} = -r_{61} = 3.4 \text{ pm/V}$,
- $r_{33} = 30.8 \text{ pm/V}$,
- $r_{51} = r_{42} = 28 \text{ pm/V}$.

The crystals are cut and used in an orientation to maximize the field projected onto the direction with the largest electrooptic effect. Clearly, it would be inefficient to rely on the r_{13} and r_{22} terms in LiNbO_3 . While the formalism to describe the electrooptic effect is fairly complex, in practice there is only one scalar quantity used in the analysis of electrooptic modulators based on LiNbO_3 , that is r_{33} (30.8pm/V), which describes the effect of an applied electric field in the z-direction on an optical wave polarized in the z-direction. For z-cut LiNbO_3 the z-axis is normal to the crystal and the optical wave is “TM” polarized and for the x-cut LiNbO_3 , the x-axis is normal to the crystal and the optical field is “TE” polarized.

Electrooptic Modulation in Bulk Crystals

The basic mechanism for all electrooptic (E-O) modulators treated here is a voltage-dependent phase shift. The optical and electrical fields must be oriented with respect to the crystal in such a way that a good phase shift is obtained. The first E-O phase shifters were built using bulk crystals, and used either a longitudinal or transverse orientation of the electric field with respect to the direction of propagation of the optical wave. The longitudinal method is not particularly desirable because the electric field direction and the optical path length are parallel. Thus, it increases the length of the crystal to increase the accumulated phase shift. It also decreases the electric field for the same voltage applied to the crystal ends. The net effect is no improvement in phase shift with increased crystal length. Furthermore, the electrodes needed to apply the electric field in this direction typically interfere with the optical pathway. The transverse mode of operation is more desirable, in that the field is applied normal to the direction of propagation, as shown in Figure 2 - 1a. The optical phase shift (often termed “retardation”) can now be increased by increasing the length of the crystal without changing the strength of the applied electric field. A second advantage is that the electrodes now do

not interfere with the optical beam. Even so, it takes a lot of voltage to make an appreciable phase shift using bulk crystals. The electrooptic induced change in electrical path length is given by $\Gamma = (2\pi f/c)(L_m/d)n_o^3 r_i V$ in the transverse configuration, where L_m is the length of the crystal, d is the spacing between the electrodes, and r is the appropriate r_i for the crystal orientation. The half-wave voltage, or $V\pi$, is the voltage at which the optical beam has experienced an additional 180° phase shift. For a $1.3 \mu\text{m}$ beam in LiNbO_3 in which the r_{33} electrooptic coefficient used is about 4 kV assuming $(\ell/d) = 1$. If the crystal is four times longer than it is high, then $V\pi$ is 1 kV. The voltages are even higher for crystals with smaller values of electrooptic coefficients, such as ammonium dihydrogen phosphate (ADP) or potassium dihydrogen phosphate (KDP) or their deuterated analogs.

One might think that simply increasing (L_m/d) would solve the problem. This is not the case. As a practical matter, it becomes increasingly difficult to fabricate long, thin crystals. However, a more fundamental limitation occurs because of diffraction of the propagating light. If the transverse dimension d is made very small, the light propagating through the crystal will spread more rapidly, and hit the sides of the crystal, thus being absorbed. The optimum strategy for a light beam with a Gaussian transverse distribution is to place a Gaussian focus or beam waist half-way through the crystal, so that the propagation path is symmetrical. With this strategy, there is still an optimum (l/d) . This is illustrated following the derivation for a Gaussian beam in a lens waveguide from Ref. [2.1]. Assuming in Figure 2 - 1b that the lens has a focal length f , $2w_0$ is the beam waist diameter at the thin lens, $2w_1$ is the beam waist diameter at the minimum in the crystal, and θ is the beam diffraction angle of the far field, then

$$\begin{aligned}\theta &= \frac{\lambda}{\pi w_0 n}, \\ z_0 &= \frac{\pi w_0^2 n}{\lambda}, \\ L_1 &= \frac{f}{1 + \left(\frac{f}{z_0}\right)^2}.\end{aligned}\tag{2.1}$$

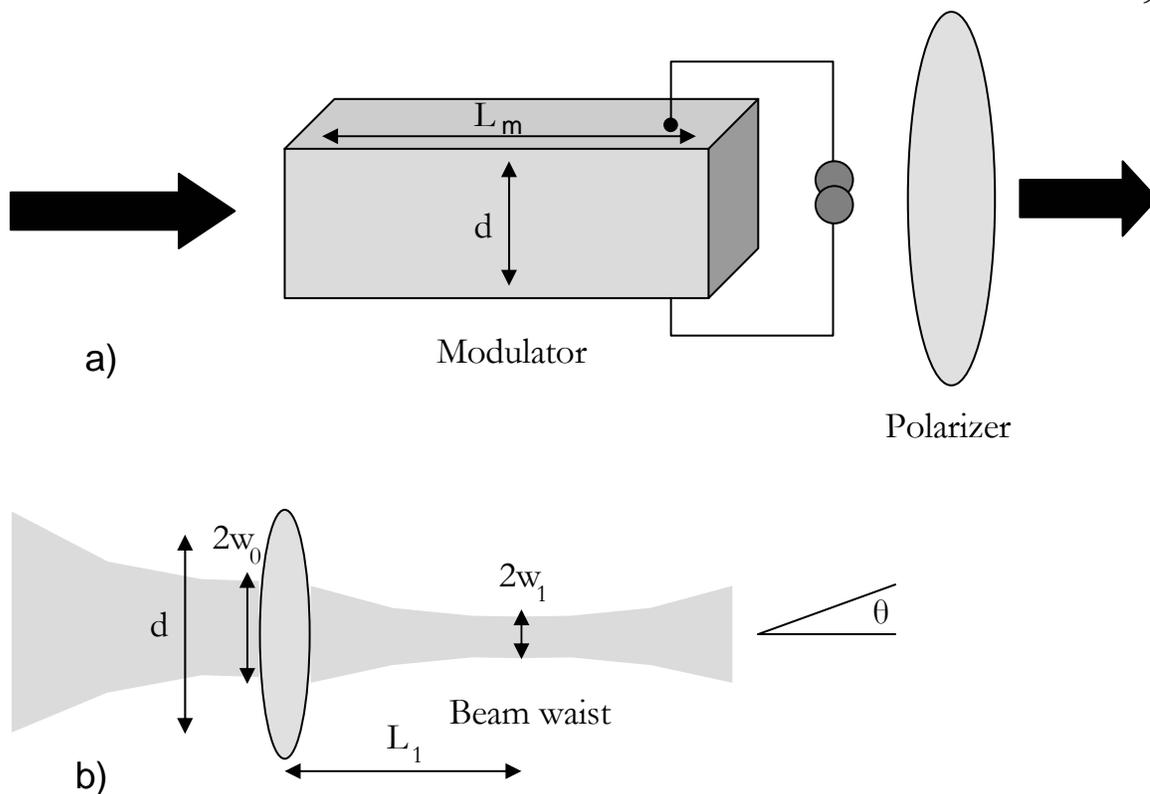


Figure 2 - 1: a) Bulk crystal modulator in transverse configuration, b) Illustration of Gaussian beam waists in a bulk crystal modulator.

The diffraction angle θ illustrates the dependence of the initial beam waist. A second useful parameter is the confocal beam parameter z_0 . It is not the case that the distance L_1 simply depends on the focal length of the lens. If $f \ll z_0$ then $L_1 \approx f$, if $f = z_0$ then $L_1 = f/2 = z_0/2$ (the maximum value for L_1), and if $f \gg z_0$ then $L_1 \approx z_0^2/f$. So to maximize the distance from the lens to the beam waist $2w_1$, a focal length is chosen based on the fundamental parameter z_0 . A few numerical examples illustrate this effect. The modulator dimension “ d ” must be larger than the input beam waist $2w_0$ to avoid losing optical energy at the input. Suppose $d=4w_0$, $n=2.2$, $\lambda=1.3\mu\text{m}$, and $L_1=10\text{ mm}$. Then since $d = 4 (\lambda L_1/\pi n)^{1/2}$ we have $(L_1/d) = 0.25*(\pi n L_1/\lambda)^{1/2} = 58$. The ratio of (L_1/d) can be arbitrarily large with a crystal of arbitrary length, but it is desirable to make the required half-wave voltage as low as possible with a

manufacturable crystal. This can be achieved with the introduction of waveguides in the crystal.

Waveguide Phase Modulators

The ability to fabricate optical waveguides in electrooptic crystals removes this limitation on (L/d) . The waveguide eliminates diffraction, so that the crystal length may be made as long as required to achieve the desired sensitivity, limited only by optical loss in the crystal. And the transverse dimension d is no longer limited by the fragility of crystal fabrication; the bulk of the crystal mechanically supports the *in situ* wavelength-scale waveguide. This technique has reduced the $V\pi$ by three orders of magnitude or more over the best bulk crystal modulators, making these devices suitable for a wide range of applications.

To make an optical waveguide, photolithography is used to define a strip mask on the crystal surface. In titanium in-diffused waveguides, a thin metal layer of a few hundred angstroms, by several micrometers in width, is deposited through the mask and then in-diffused into the crystal by heating it at very high temperature, around 1050° C. The titanium atoms enter the crystal lattice slightly increasing the index of refraction in the crystal by about $\Delta n=0.01$, as described in Ref. [2.3],² creating a dielectric waveguide about 4 μm deep, centered about 2 μm below the crystal surface, and about 6 μm wide for single mode operation at an optical wavelength of 1310 nm. The electric field is applied perpendicular to the waveguide, exciting the transverse mode of modulation. The electrodes are also defined by photolithography; they are located on the surface of the crystal, and separated by the thickness of the waveguide. This reduces the required voltage to about 10 V for a half-wave phase shift over a modulator length of about 10mm. By increasing the length and optimizing the electrode geometry, researchers have reduced the required voltage to about 2 V in modern LiNbO_3 modulators, yielding $(l/d) = 2000$.

² It is not well known whether the Ti is exchanged with the Nb, or it goes interstitially. In Ref. [2.11], Schmidt suggests that either case is plausible. He reports single guide fabrication in LiNbO_3 for the transition metals Ti, V, and Ni are a representative set.

A phase modulator is the simplest modulator to build. Figure 2 - 2 shows a waveguide-based lithium niobate phase modulator in the “x-cut” orientation. There is a single optical waveguide parallel to the y-axis of the crystal. The electrodes are fabricated as strips of metal (usually gold) on top of the crystal. The gap between the waveguides is sized and positioned to maximize the electric field component along the direction of the crystal with the strongest electrooptic coefficient. In case of “x-cut” lithium niobate, light propagates along the y-direction, the z-direction is normal to the substrate, and the x-direction is along the surface, perpendicular to the direction of propagation. The electric field lines penetrate down into the crystal, continue perpendicular to the waveguide in the z-direction of the crystal, and then go back up to the opposite electrode. The fields are illustrated explicitly for both x-cut and z-cut modulators in Figure 2 - 3.

Intensity Modulators Based on Mach-Zehnder Interferometers

Phase modulation is not easy to use in a photonic link since photodetectors are sensitive only to intensity. Somehow, phase modulation must be converted to intensity modulation. It is possible to use optical phase modulation, and then demodulate it by coherently converting the modulated signal to a lower (radio) frequency, then using an electronic phase detection system. Such systems have been demonstrated. However, far and away the more popular approach is to convert the optical phase modulation to intensity modulation by using optical interference. Figure 2 - 4 shows one such configuration to do this, a Mach-Zehnder modulator (MZM), based on the famous Mach-Zehnder optical interferometer, and adapted to optical waveguide components. A Mach-Zehnder modulator is typically built with two, phase modulators and two “Y-branches.” The light enters a single waveguide and is then split equally into two paths through the two phase modulating sections. The three-piece electrode structure produces the same strength electric field across both optical waveguides but with opposite directions of the field in the two, phase modulators. This yields equal and opposite phase modulation in each waveguide, known as a “push-pull” configuration. In the second Y-branch, the two optical waves are combined coherently (interfered), which yields a cosine-squared transfer function of intensity with applied voltage. This transfer function is shown in Figure 2 - 5. The voltage required to introduce enough phase shift to produce destructive interference so that the net

output is zero is termed “ $V\pi$.” For analog applications, the modulator is typically d-c biased at $V\pi/2$, so that the optical output is 50% with no modulation applied. Applying the modulation voltage in addition to this bias then results in intensity modulation about the most linear portion of the cosine-squared curve. That gives 3-dB in optical loss from the modulator transfer function in addition to the mode-matching losses at the crystal interfaces and any propagation losses through the crystal. A useful property of simple interference modulators like the MZM, with a cosine-squared transfer function, is that when biased at the half-wave point, all even-order harmonics of the modulating signal are zero, since all even-order derivatives of the cosine-squared function are zero at that point. The distortion produced by the modulator is thus dominated by the odd-order intermodulation products and odd-order harmonics from the (non-zero) odd-order derivatives of the transfer function, which have

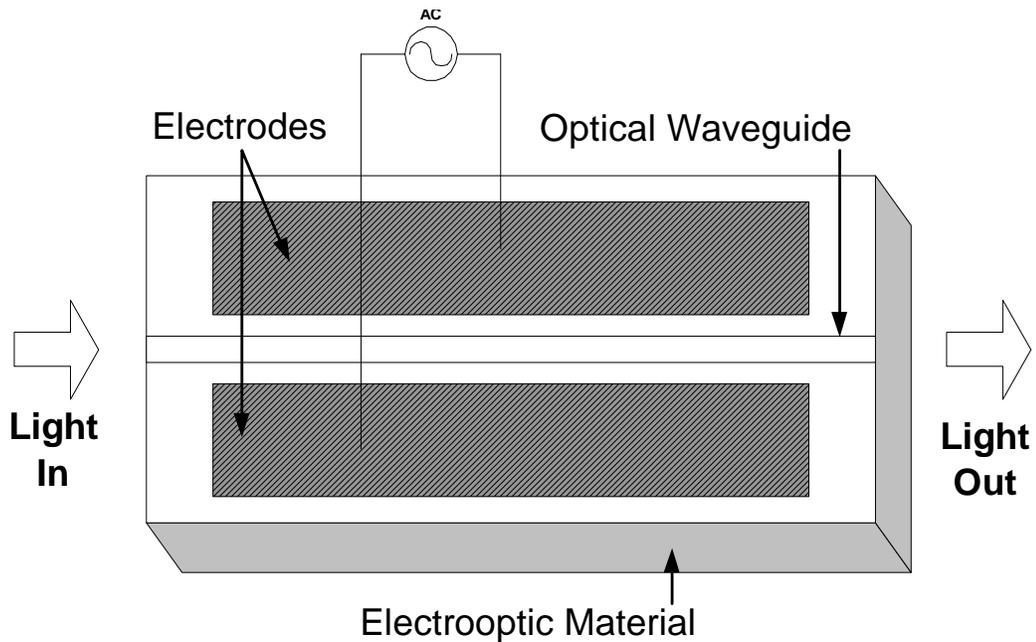


Figure 2 - 2: Phase Modulator based on in-diffused waveguides in an electrooptic crystal.

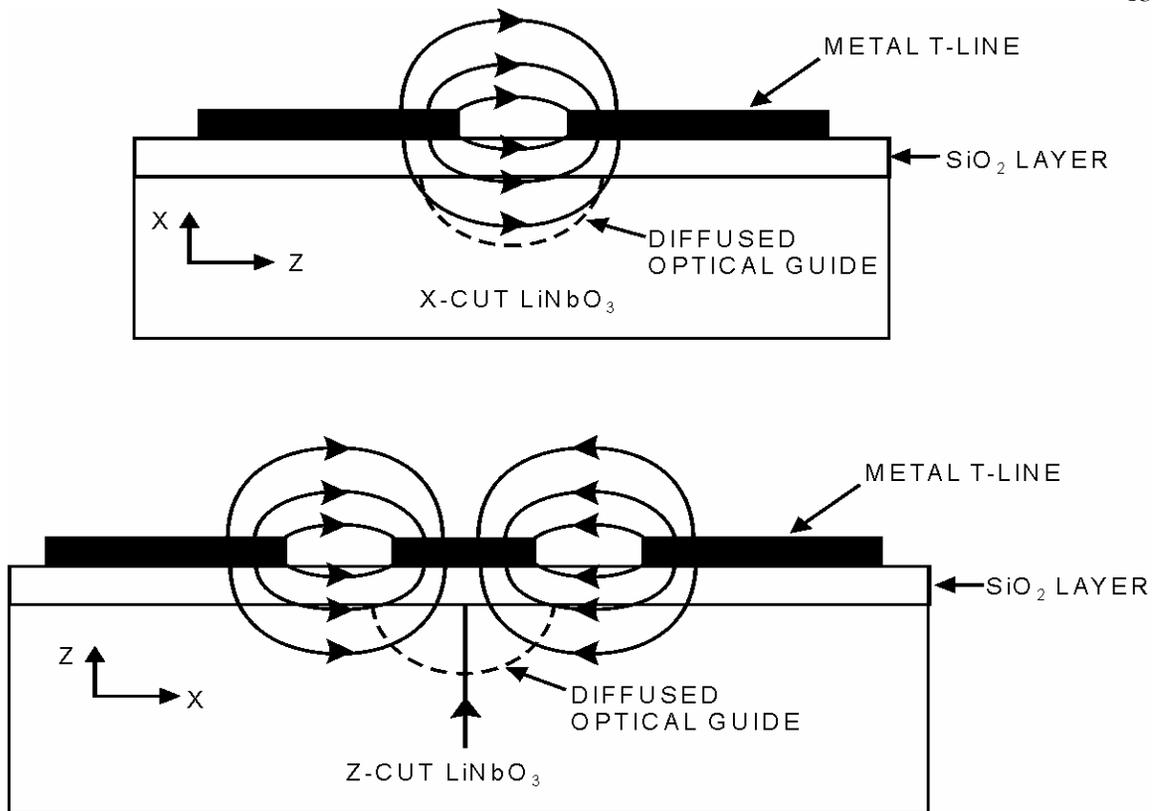


Figure 2 - 3: Schematic cross-section view of a phase modulator in an x-cut lithium niobate crystal (top schematic) and a z-cut lithium niobate crystal (bottom schematic).

their maximum values at this bias point. The third-order intermodulation is the strongest of these and is usually the limiting distortion.

The MZM is relatively insensitive to small fabrication non-idealities in the Y-splits and phase lengths, making it the most commonly used external optical modulator in the industry. Unequal intensity splits or combinations in the Y-branches cause incomplete cancellation, or reduced extinction ratio in the output, not an important parameter in small-signal analog systems (although important in large-signal or digital modulation systems). Unequal path lengths simply require shifting the bias voltage to the proper quadrature point.

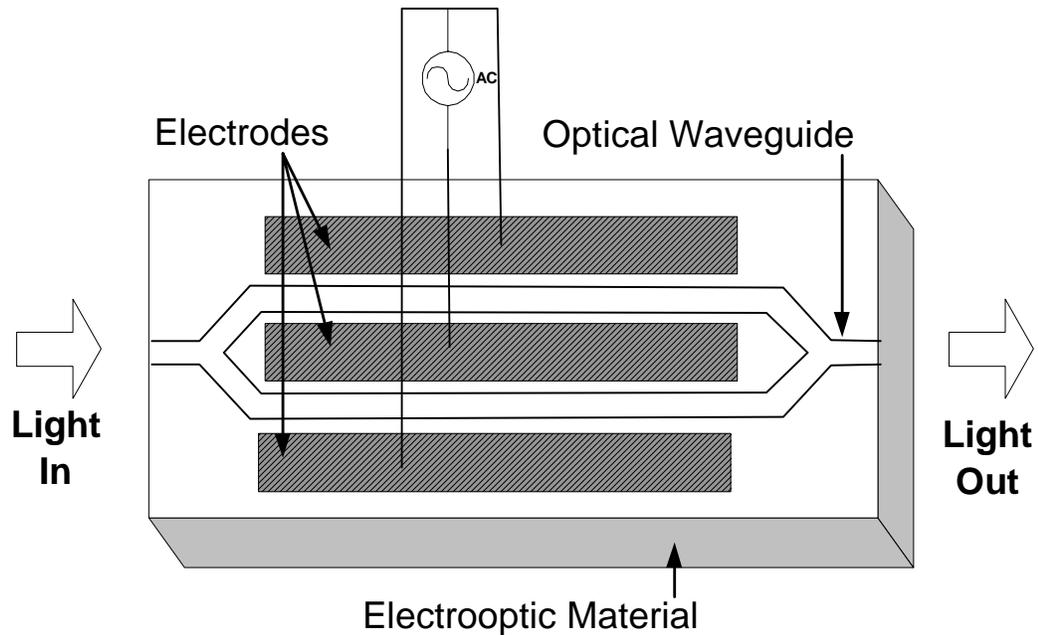


Figure 2 - 4: Mach-Zehnder modulator based in indiffused waveguides in an electrooptic crystal.

Waveguide Electrooptic Modulators Based on Directional Couplers

Another basic modulator configuration is based on an optical directional coupler, with waveguides integrated into the E-O crystal surface, as in the MZM. The directional coupler modulator (DCM) utilizes distributed interference to convert phase modulation to intensity modulation, and has a mathematically richer transfer function than the MZM, and for that reason is an interesting topic of research. The derivatives of the transfer function do not exactly “line up” with each other, as they do for the cosine-squared transfer function of the MZM. One novel result is that the third-order intermodulation product and the signal do not null at the same voltage, which makes it a good candidate for linearization with “sub-octave” modulation; that is, modulation of less than an octave bandwidth, so that the second harmonic distortion is out of band and therefore does not limit the dynamic range. Figure 2 - 6 shows a

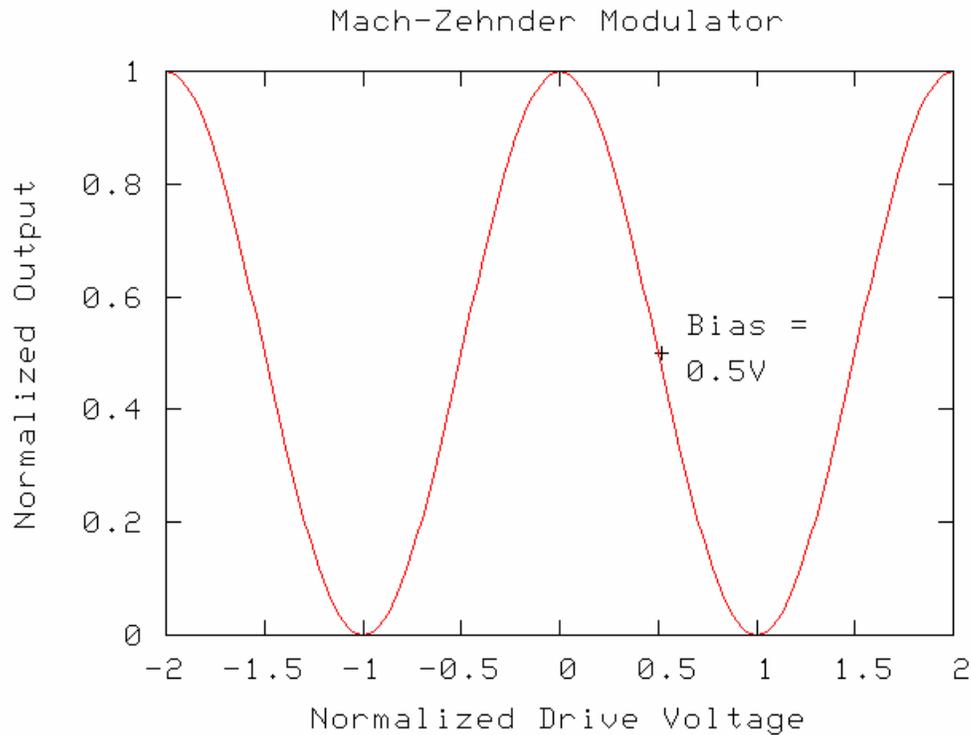


Figure 2 - 5: Mach-Zehnder Modulator transfer function: normalized output intensity as a function of drive voltage (normalized to $V\pi$).

directional coupler modulator. There are two waveguides fabricated in the crystal, and they are spaced about $6 \mu\text{m}$ apart, comparable to the dimensions of the waveguides themselves. This spacing is small enough that the evanescent tails of the modes in each waveguide can penetrate the other waveguide, and thus waves from each guide couple into the adjacent guide. If the two optical waveguides have identical phase propagation constants, a complete transfer of the wave will occur in some length of guide (termed the “transfer length”), and then transfer back again and so forth down the guide. If the two guides do not have the same propagation constants, then this transfer will be incomplete. By adding electrodes as shown, the effective refractive index of the guides can be changed. The electrodes are arranged so that the electric fields are in opposite directions through the two guides, and a significant differential shift can be made in the propagation constants, which results in a significant change in the transfer of

energy from one guide to the other. The fields in the directional coupler are shown in Figure 2 - 7.

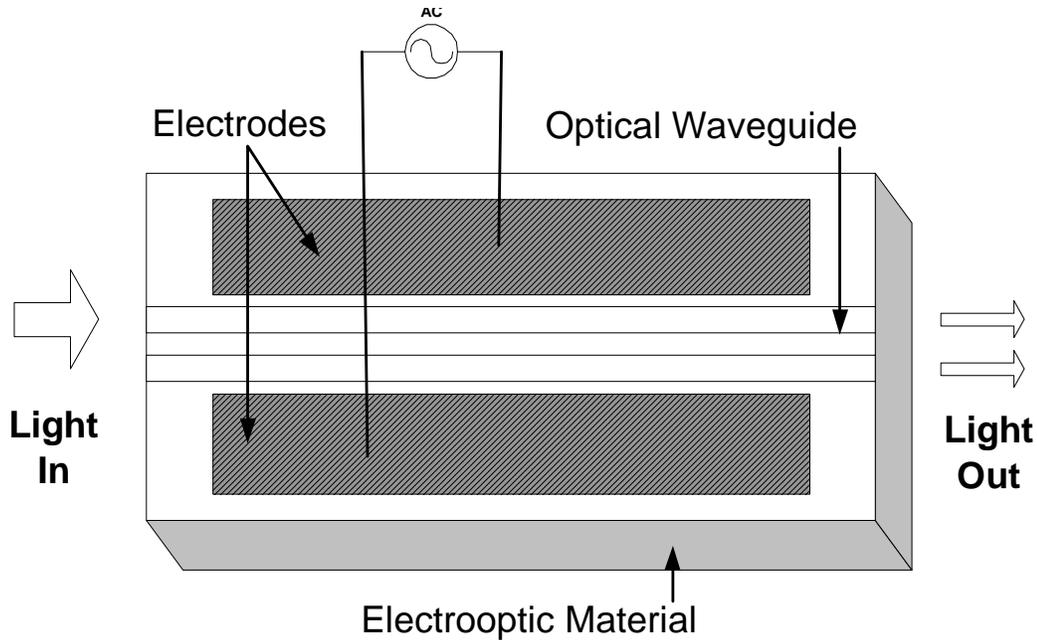


Figure 2 - 6: Directional Coupler Modulator based on indiffused waveguides in an electrooptic crystal. The z-direction of the crystal is used to create the modulation. Not shown in the figure, the waveguides are typically bent through anti-symmetric “S” curves between the edge of the crystal and the active area of the modulator to provide more lateral separation of the two guides for ease in coupling light in and out.

With two optical inputs and two optical outputs, the DCM has more possible configurations than the MZM, which has only one input and one output. A typical DCM configuration might have all the optical power applied to one input guide, and be fabricated with a length that would result in a complete transfer to the other guide. The electrodes would be biased so that the transfer is reduced to only about 50%, as in the MZM, and then the modulating voltage would be added to the bias to increase or decrease the output. It turns out that modulation sensitivity is very similar to an MZM of the same length, but the transfer function has important differences. Figure 2 - 8 shows the transfer function of a directional coupler modulator.

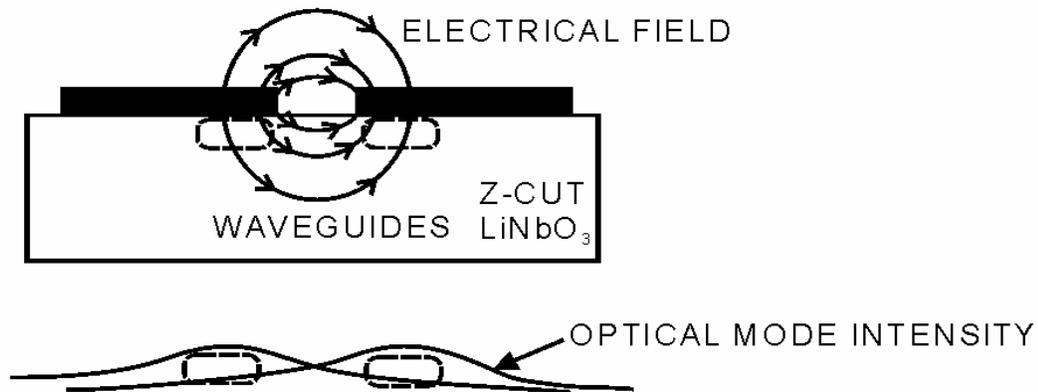


Figure 2 - 7: Fields in a directional coupler modulator looking down the waveguides; the top view is of the crystal and electrodes, the bottom view is of the overlap of the optical field intensity between the two waveguides.

The coupling is created by the same physical mechanism as in r-f and electronic directional couplers; and the governing equations are very similar. The fields of one waveguide partially penetrate the other, and there is a transfer of optical energy. In a modulator, the applied electrical field linearly changes the strength with which light couples from one guide to the next. In approximate form, this device creates a $\sin^2(x)/x^2$ transfer function. Unlike the Mach-Zehnder, it does not have the property that all even-order distortion terms cancel at a single bias point, and it is not as robust to fabrication errors, particularly those affecting the optical coupling, such as the guide-to-guide spacing and the size of the “tails.” However, it can be employed in many configurations, the subject of future chapters of this thesis, which make it an interesting alternative to the straightforward Mach-Zehnder modulator.

Fidelity in Electrooptic Modulators

In analog applications, it is desirable to make a photonic link as high fidelity as possible. While there are a number of parameters that determine link fidelity, one inherent limiting factor is the non-linearity of the modulator’s transfer function. The modulators described in Figure 2 - 4

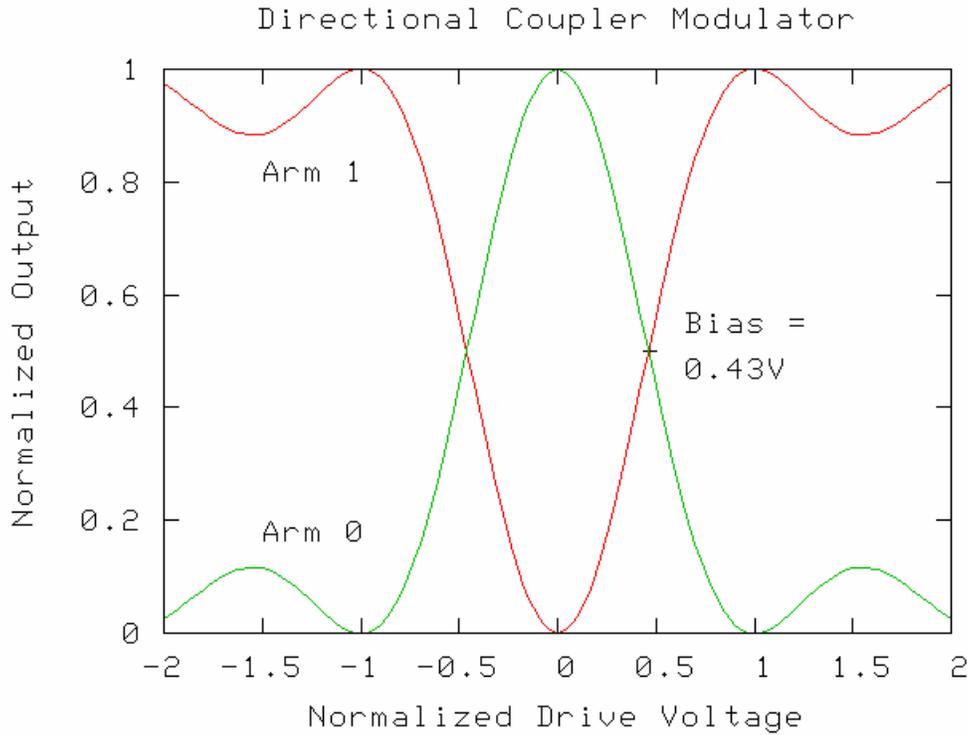


Figure 2 - 8: Transfer functions of both arms of a directional coupler modulator. The bias voltage is 0.43 V_s .

and Figure 2 - 6 are not linearized, and are often referred to as “normal” modulators in this thesis. They introduce distortion products that limit the fidelity of the photonic link. In general, it is the third-order intermodulation products that limit the fidelity since they fall very near the signal in frequency.³ A linearized modulator employs some scheme that nulls the third-order intermodulation product, so that the limiting distortion is due to the fifth-order intermodulation products, which occur at still higher signal levels. There are many such schemes to achieve linearization; they are analyzed in detail this thesis. If greater than octave bandwidth is required, then whatever scheme is used to null the third-order products must also null the even-order products (particularly the second harmonic) as well. It is generally a harder task to accomplish both simultaneously, but that is the challenge in broad-bandwidth modulators.

Frequency and Bandwidth Limitations in Electrooptic

Modulators

Frequency response (and bandwidth in the case of baseband modulators) is a key issue for external electrooptic modulators. The modulators shown in Figure 2 - 4 and Figure 2 - 6 use lumped capacitance electrodes. Typically, an inductance L and a load resistor R_L is placed in parallel with this capacitance to make an RLC circuit, with a 3-dB point given by $\Delta f = 1/(2\pi R_L C_{\text{mod}})$. Assuming the modulator has a 50 pF capacitance and the load resistance is 50 Ω , then the 3-dB point is 64 MHz away from the center frequency. This will have a bandwidth if $f = 2/(2\pi R_L C_{\text{mod}})$, which is 128 MHz, but now centered at the desired operating frequency. Modern communication systems require bandwidths in the GHz range, so the lumped capacitance electrode structures simply are not practical. A different implementation is required.⁴

Even for applications where the bandwidth of a resonant circuit is sufficient, there is a fundamental limitation to the frequency of operation of an electrooptic modulator that comes from the finite transit time of the modulated beam. Essentially, the phase of the modulating voltage changes while the light is traversing the length of the crystal so that, effectively, the same voltage is not applied over the entire length. This results in a net reduction from the peak retardation of

$$r = \frac{1 - e^{-i\omega_M \tau_d}}{i\omega_M \tau_d}, \quad (2.2)$$

where τ_d from (2.2) is the transit time of light through the crystal, $\tau_d = n\ell/c$, and ω_M is the radial frequency of operation (see, for example Ref. [2.1]). The modulation amplitude reduction factor is a sinc function of $\omega_M \tau_d$. Taking the absolute value of $|r| = 0.9$, gives a

³ For two signal components at f_1 and f_2 , the third-order products are at $2f_1 - f_2$ and $2f_2 - f_1$, both only $|f_2 - f_1|$ away from a signal component.

⁴ Lumped capacitance electrodes are often used for d-c biasing. In this case, they are a separate electrode structure from the modulating electrodes, and have no effect on the maximum frequency of operation.

maximum frequency of $c/(4n\ell)$, or equivalently $1/(4t_d)$. In LiNbO_3 , $n=2.15$, and for a 10 mm modulator, this is 3.49 GHz.

The solution to both the narrow bandwidth RC response of the electrode structure and the finite transit time of the optical beam is to use traveling wave electrodes. In a traveling wave configuration, the electrodes become a transmission line. The line is excited at the input end, the modulating electrical field propagates down the electrodes parallel to the propagation of the optical wave, and then is terminated in matching impedance at the end of the electrodes. The electrodes no longer form an RC filter; instead they become a transmission line with distributed inductance and capacitance. The issues of loss, reflections, and radiation must now be considered instead.

The finite transit time problem is improved but not totally solved by the transmission line electrodes because there is generally a different velocity between the optical and r-f waves. The reduction factor expression now becomes

$$r = \frac{e^{-i\omega_M \tau_d (1 - n_m / n_o)} - 1}{i\omega_M \tau_d (1 - n_m / n_o)}. \quad (2.3)$$

Taking the optical velocity to be c/n_o and the microwave velocity to be c/n_m , this gives an improvement of the frequency at which $|r| = 0.9$ of $|1/(1 - n_m/n_o)|$ over the previous value $c/4n_o\ell$. When $n_o = n_m$ then the optical wave and the r-f wave are traveling with the same phase velocity and there is no velocity mismatch. When $n_m = 0$ (infinite microwave phase velocity) then the reduction factor is the same as the lumped electrode. For lithium niobate the r-f phase velocity is slower than the optical phase velocity for waves traveling in parallel. For example, in LiNbO_3 for fields in the z-direction, $n_e = 2.15$. For the microwave velocity, the electric permittivity is 28. However, half of the electric field between the electrodes is in the air so it is appropriate to use the average permittivity (recall Figure 2 - 3). The effective microwave index of refraction is taken as

$$n_m = \epsilon_m^{1/2} = \left(\frac{\epsilon_r + 1}{2} \right)^{1/2}. \quad (2.4)$$

Here ϵ_r is the permittivity for the crystallographic axis of propagation. For propagation in a lithium niobate crystal in which the electric field is equally split between the z-axis and the x-axis, $\epsilon_r = (\epsilon_x \epsilon_z)^{1/2} = (28 \cdot 43)^{1/2} = 34.7$. And thus $n_m = 4.2$. In a lithium niobate traveling wave modulator where both fields travel in parallel and are polarized in the same direction, the frequency at which $|r| = 0.9$ is $c/4n\ell(1-n/n_m) = 7.15$ GHz for a 10 mm modulation section, an improvement of more than 2x over the “lumped element” electrodes, but clearly still frequency limited.

Increasing Frequency Response of E-O Modulators by Velocity Matching

Velocity-matching techniques are often used to improve the frequency response of the traveling wave modulator. In Ref. [2.3] many velocity matching techniques are described and referenced. In waveguide-based modulators the fields propagate down coplanar electrodes on the surface; some field lines penetrate into the crystal (these are useful for modulation) and other field lines are in the air. While Eqn. (2.4) assumes an equal split, it is possible to create thick enough electrodes so that more than half of the electric field is in air, resulting in $n_m = f(n_{\text{air}}, n_{\text{mb}}) = n_o$. This is done by up-plating the coplanar waveguide. Unfortunately, this matching technique also decreases the modulator sensitivity to the electric field (increases $V\pi$) since the field drawn up into the air between the electrodes no longer passes through the optical waveguides. In addition to up-plating the electrodes, a thick SiO_2 buffering layer can be deposited between the crystal and the electrodes. The microwave index of refraction of SiO_2 is less than that of lithium niobate so the electric field that passes through the SiO_2 layer also reduces the effective refractive index of the transmission line electrode structure. A number of successful high frequency modulator experiments have been made exploiting the velocity matching technique beyond 40 GHz, for example, velocity matching techniques were used to go above 30 GHz in Ref. [2.4] and truly exceptional work was done to get 100 GHz velocity matched modulators in ref. [2.5].

This kind of velocity matching is a limited solution because it leads to loss in sensitivity, which must be made up by increasing the length of the modulator and, at such high frequencies,

electrode loss becomes a serious compounding problem. In essence, the electrode loss will limit the “length” of the modulator no matter how long it is made physically. An alternative technique that addresses the velocity mismatch problem is called “velocity matching on the average” and has been demonstrated in several forms. See, for example Ref. [2.6]. A velocity-mismatched electrode may be divided into sections, where each section is typically made shorter than the maximum length of a useful modulating electrode at the target frequency, as determined by electrode loss and velocity mismatch. The modulating electric field is then arranged to be “re-phased” at the beginning of each modulator section, so that it never gets too far out of step with the modulated optical wave; it is “velocity matched on the average.”

One method of accomplishing this was the “phase reversal” modulator demonstrated by Alferness, et al [2.7] in which the electrode sections were made 180° long in phase (at some desired modulation frequency), then connected to the next electrode section with an electrical transposition, and so forth down the entire structure. This assured “phase matching on the average,” but only at one frequency, where the sections are 180° in phase delay. This is not true velocity matching or true-time-delay matching. The bandwidth of the modulator decreased as more sections were added.

Another technique was proposed and patented by Schaffner and Bridges, Ref. [2.8], in which the modulating signal is divided by a multi-branched transmission-line feed that has varying lengths that just compensate for the optical delay down the waveguide under the electrode structure. Each section of the modulator is then fed at its entrance end with the proper delay to equal the optical delay to that section. This is a true velocity matching scheme and has the advantage of not narrowing the bandwidth, but it does have the disadvantage of dividing the modulating signal by the number of paths, all of which have transmission loss.

A third technique is the “antenna-coupled” technique proposed and patented by Bridges in Ref. [2.9], in which small antennas are put on the surface of the modulator and connected to the inputs of each section. The modulator is then illuminated by a plane wave at the correct angle so that the modulating wave received at each antenna is in synchronism with the optical wave entering that section, as shown schematically in Figure 2 - 9. This has the advantage of low-loss free-space transmission to the segment inputs, but now is limited to band-pass

operation by the bandwidth of the on-surface antennas and the wave radiating structure (but

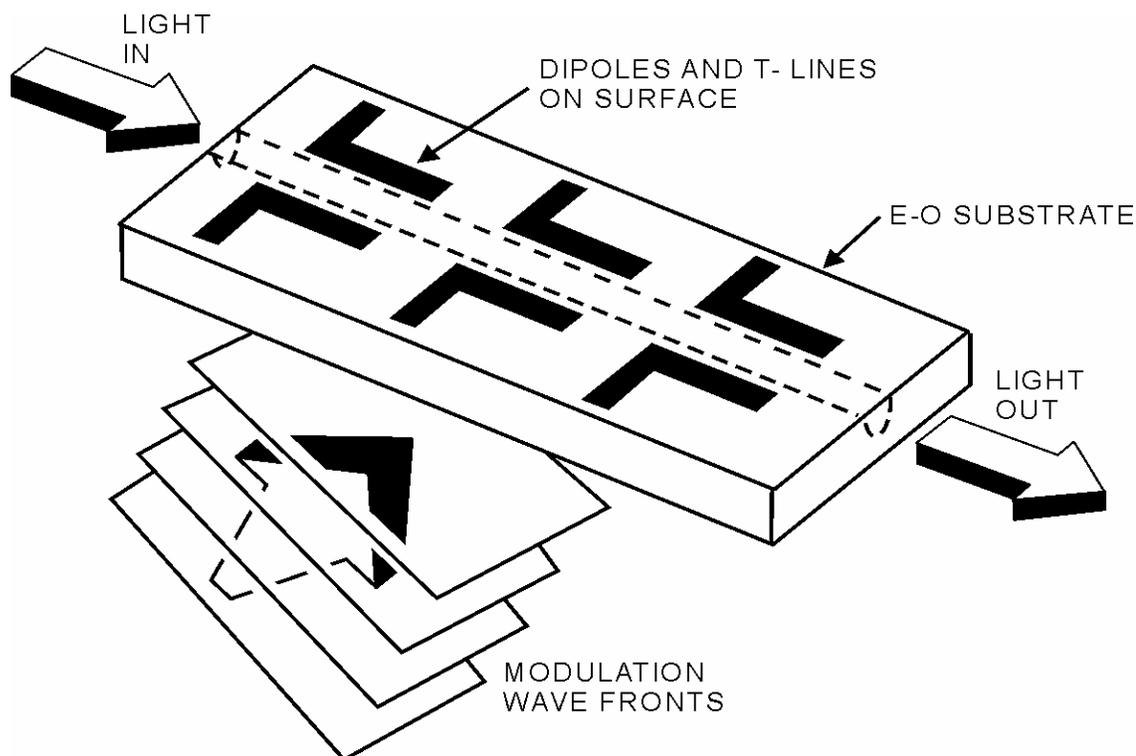


Figure 2 - 9: Schematic of antenna-coupled phase modulator. The microwave signal is received by the antenna elements in synchronism with the arriving optical wave that travels through the waveguide. Selection of the appropriate angle of incidence of the microwave yields a velocity matched modulator “on the average,” discussed in Ref. [2.6].

not further narrowed by cascading, as in the phase reversal modulator). The electrical signal power is also reduced by the number of sections used, but this is partially offset by the lower exponential attenuation in the modulating electrodes, as shown by Sheehy in Ref. [1.2]. The antenna-coupled structure has been demonstrated with MZMs again in Ref. [1.3] and with DCMs in Ref. [2.10], which is detailed in the last chapter of this thesis. The issue of velocity mismatch limiting the gain of a simple modulator is well-known. However, how velocity mismatch affects the various linearization schemes used with modulators was unknown before

24

the research treated in this thesis was undertaken. This behavior is a major topic of this thesis, and the results turn out to be quite interesting.

Chapter 3

DISTORTION IN MODULATION

Abstract

This chapter addresses distortion in electrooptic intensity modulation. A simple low frequency link model with a set of canonical link parameters is used, establishing a reference by which to compare modulation schemes. The gain, two-tone dynamic range, noise figure, and multi-tone dynamic range are defined. Seven different linearization schemes are compared with a numerical calculation of these quantities. The bandwidth scaling of two-tone dynamic range is investigated and there is a systematic analysis of the sensitivity of the key parameters used in linearization schemes. From this analysis, engineering conclusions are drawn of where it is appropriate to use linearized modulators.

Link Model

Figure 3 - 1 shows a simple intensity modulated photonic link. The intrinsic link consists of a low noise laser, an electrooptic modulator, a length of fiber, and a square-law detector. Electronic pre- and post-amplifiers are almost always included in actual practice, but with the exception of their signal bandwidth parameter, they are left out of the calculation. Since there is modulator non-linearity, the performance calculation must be made numerically, and specific numerical values for the component parameters must also be used. It is desirable to establish a reference link with which to compare modulation schemes; the numerical parameters from Ref. [3.8] are used here to compare with the results of their work, and these parameters are given in Table 3-1. Note that these are “garden variety” values, not “best ever” results obtained only in research laboratories. In fact, they are likely not as good as in the latest commercial fiber-optic links (available in 2005), but they suffice for a comparison of linearized modulator performance from one modulator to another. This link is as simple as possible while still capturing the properties of the modulator. Non-linear effects in the fiber and the detector are ignored. While it would be appropriate to add such effects as Brouillin scattering, noise from optical amplification, and photodetector non-linearity to accurately model a link, these effects are unnecessary for the comparison of linearized modulation schemes and are left out.

Table 3-1 shows the parameters associated with this link. The laser has an optical power, P_L , of 100 mW and a relative intensity noise (RIN) of -165 dB/Hz. These parameters are no longer state of the art; now many links have optical powers of 200 - 300 mW and a RIN of -175 dB/Hz, but the parameters are left as stated for compatibility with the literature. The modulator sensitivity, $V\pi$ for Mach-Zehnder modulators (MZM) and V_s for directional coupler modulators (DCM), is the voltage required to switch the output from unity to zero. The velocity mismatch and modulator length parameters are not required for the zero-frequency analysis in this chapter, but they are included here to be used in the following chapter. They are necessary for the finite transit time calculation of modulator performance. The optical loss is a lumped optical loss equal to the sum of modal mismatch loss at the modulator-fiber interface, the net loss in the modulator itself, and losses in the fiber.

These parameters normalize the input r-f drive, the output optical intensity and the noise level of the electrooptic modulator. Given this, and the d-c transfer function of each modulator, a standard Fourier analysis is applied to obtain gain, dynamic range, and noise figure for any electrooptic modulator. While some of these figures of merit are analytically calculable for some modulators (i.e., for the simple MZM, the cosine-squared function can be expanded in a Bessel series), it is more systematic to design a numerical calculation and use it for every modulator.

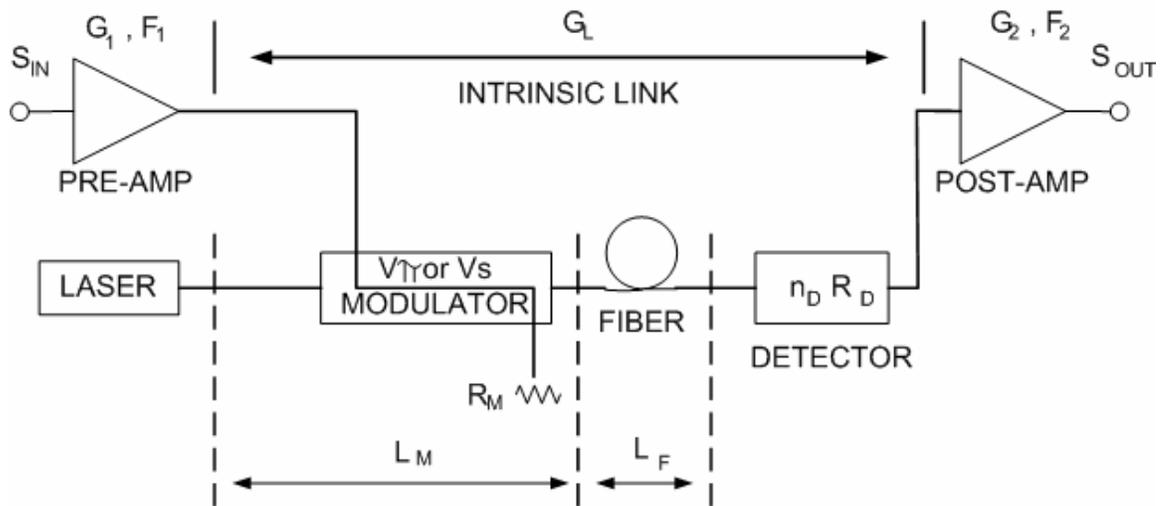


Figure 3 - 1: Canonical model of an optical link for evaluating external modulators.

Analysis with One and Two Tones

A computer program was written for the analysis in this thesis. The program calculates the time domain output of a modulator, taking as its input the modulator transfer function and the input r-f excitation signal. It works for any predefined transfer function and input signal. The transfer functions for all of the modulators analyzed in this thesis have been defined in the program. The input signal functions are sinusoidal, and have been developed to capture all of the effects described in this and later chapters: distortion, frequency effects, loss, reflections, and rephrasing structures. This program will be described in more detail in Chapter 4.

Table 3-1: Canonical Link Parameters

Parameter	Symbol	Value	Units
Laser Power	P_L	100	mW
Laser Noise	RIN	-165	dBm/Hz
Modulator Sensitivity	$V\pi, V_S$	10	V
Modulator Impedance	R_M	50	Ω
Velocity Mismatch ($n_{\text{micro}}-n_{\text{opt}}$)	Δn	1.8	-
Modulator Length	ℓ_M	10	mm
Optical Loss	L_O	-10	dB
Detector Responsivity	η_D	0.7	A/W
Detector Load	R_D	50	Ω
Signal Bandwidth	BW	1 or 10^6	Hz
Maximum Photocurrent	$P_L L_O \eta_D$	7	mA

Fourier analysis is used to calculate distortion, gain, and dynamic range of the modulators. An input voltage of one, two, or multiple sinusoidal tones is applied to the modulator as an input to the optical transfer function. The transfer function relates the output complex optical amplitude to the input complex optical amplitude by a modulator-specific and voltage dependent function, $H(V) = A_O/A_I$. The output wave form from the detector is the squared absolute value of the optical amplitude. Fourier analysis of the normalized photodetector output current wave form, $|A_O|^2$, yields the frequency components used to compute the

modulator figures of merit. Note that the calculations ignore any deviation from ideal square law behavior of the photodetector, even though such deviations have been measured and are, in fact, important for links with high average currents on the photodetector, of example, see Ref [3.1]. For a clear comparison of the properties of low-distortion modulators, other components in the optical link are assumed to be ideal.

Small signal gain is determined from single tone analysis. Let $P(p_{IN},t)$ be the electrical signal power after the detector, given the modulator r-f drive power p_{IN} . Let $P^{\sim}(p_{IN},f)$ be the Fourier transform of $P(p_{IN},t)$. The gain is

$$\mathit{Gain}_{dB} = 10\{Log[P^{\sim}(p_{in}, f)] - Log(p_{in})\}. \quad (3.1)$$

The small signal gain is obtained by evaluating Eqn. (3.1) at sufficiently small p_{IN} such that the log-log plot of $P^{\sim}(p_{IN},f)$ is linear with slope one. In practice, we take $p_{IN} = -100$ dBm to determine small-signal gain, which is about the geometric mean of p_{SAT} (the power that drives the modulator voltage to about $V\pi$ and the precision of double precision floating point numbers.⁵ The frequency of the tone is set to an integer multiple of 2π so that the gain and any harmonics will fall precisely on the discrete samples of the FFT, making the calculation fast and accurate. This frequency does not correspond to a physical frequency, and thus cannot be used to analyze the roll-off of the modulator until the finite transit time of the optical and electrical signals, and frequency-dependent r-f loss is introduced into the analysis (Chapter 4). The phase in a single-tone analysis has no effect in an intensity modulated link. Single-tone analysis also gives harmonic distortion levels.

Two-tone analysis captures the signal, harmonic distortion, and intermodulation distortion. By convention, the amplitudes of the two tones are set equal, which gives the worst case distortion. The frequencies are chosen to be consecutive integer multiples of 2π . It makes no difference if they are more largely spaced, provided they remain integer multiples of 2π . The input drive is represented functionally as

⁵ A detailed analysis of numerical convergence is given in Chapter 4.

$$v(t) = 0.5\sqrt{P_{IN}R_M} (\sin(2\pi f_1 t) + \sin(2\pi f_2 t)). \quad (3.2)$$

The relative phase of the two tones makes no difference in the calculation. In the time-domain, the two tones form a harmonic oscillation at the mean frequency modulated by an envelope created by the beat of two signals. Changing the relative phase of the two signals only modifies the phase of the envelope oscillation relative to a universal reference frame; it does not change the shape of the time-domain signal, and consequently the relative phase of the tones has no effect on the dynamic range.

The spurious-free dynamic range, $SFDR_{dB}$, is the power interval that spans the input power level at which the signal is just distinguishable from the link noise to the input power level at which the strongest distortion term becomes distinguishable from the noise. The calculation of $SFDR_{dB}$ is

$$D_{dB}(P_{IN}) = 10 \max\{\log[P^{\sim}(P_{IN}, 2f_1 - f_2)], \log[P^{\sim}(P_{IN}, 2f_1)]\} - noise_{dB}, \quad (3.3)$$

$$noise_{dB} = 10 \log(GkT + RIN \cdot I_{DC}^2 R_D + 2eI_{DC}R_D + kT), \quad (3.4)$$

$$p_0 = \min(P_{IN} \mid_{D_{dB}(P_{IN})=0}), \quad (3.5)$$

$$SFDR_{dB} = 10 \log[P^{\sim}(p_0, f_1)] - noise_{dB}. \quad (3.6)$$

This definition may seem overly complicated compared to the English definition just given, however, it allows for multiple roots in the intersection of the intermodulation curve with the noise level, something which commonly occurs in linearized modulators. $D_{dB}(P_{IN})$ in Eqn (3.3) is the maximum of the intermodulation and second harmonic distortion terms minus the noise level in dB. Eqn. (3.4) expresses the noise as the sum of the small signal gain times Boltzmann's constant times temperature plus the laser RIN proportional to the square of the current times the detector load plus the shot noise $2eI_{DC}R_D$ plus kT from the terminating resistor, where I_{DC} is the average photocurrent out of the detector. The only significant sources of noise in the simple link model are the shot noise and the contribution from RIN.

Of the roots of $D_{\text{dB}}(p_{\text{IN}})$, p_0 is the root that occurs at the lowest power level, in the event that $D_{\text{dB}}(p_{\text{IN}})$ intersects the noise level at more than one value of p_{IN} . SFDR_{dB} is the difference between p_0 and the input power level at which the signal intersects the noise level. Since the log-log plot of the signal has slope one, this interval is equivalent to the signal power [dB] minus the noise power [dB] at the r-f drive power at which the distortion power equals the noise level. Figure 3 - 2 shows the dynamic range calculation for a simple Mach-Zehnder modulator. In simple modulators, the log-log plots of the distortion terms are linear, intersecting the noise floor only once. In linearized modulators, the distortion terms may be nulled at some discrete power level or levels. Thus, they may cross the noise level more than once. It is then necessary to find all of the roots of R_{dB} , the distortion minus the noise, and use the root representing the lowest RF drive power in the definition of dynamic range.

While dynamic range generally refers to all harmonics and intermodulation products, in non-linearized modulators, there are two dominant distortion terms, the second harmonic, $P^{\sim}(p_{\text{IN}}, 2f_0)$, and the third-order intermodulation product, $P^{\sim}(p_{\text{IN}}, 2f_0 - f_i)$. This thesis distinguishes two categories of linearized modulators, defined by their operating bandwidths. If f_{UPPER} and f_{LOWER} are the upper and lower frequency edges of the operating band, then a *sub-octave* modulator has $f_{\text{UPPER}} < 2f_{\text{LOWER}}$ and a *super-octave* modulator has $f_{\text{UPPER}} \geq 2f_{\text{LOWER}}$. Dynamic range is treated differently for each case. Eqn. (3.3) applies to broadband, or super-octave modulators. That is, $D_{\text{dB}}(p_{\text{IN}})$ is the maximum of the second harmonic and the intermodulation product. In narrow band or sub-octave modulators, $D_{\text{dB}}(p_{\text{IN}})$ contains only the third-order intermodulation product, since the second harmonic falls outside the sub-octave bandwidth.

In multi-tone analysis there are N equally spaced tones of equal amplitude. This is used, for example, to describe the carriers of TV signals in a multi-channel cable TV system. Unlike single and double tone analysis, the relative phases of the tones are now a critical factor in the analysis. If all of the phases are locked and equal, the signal is a pulse in the time-domain,

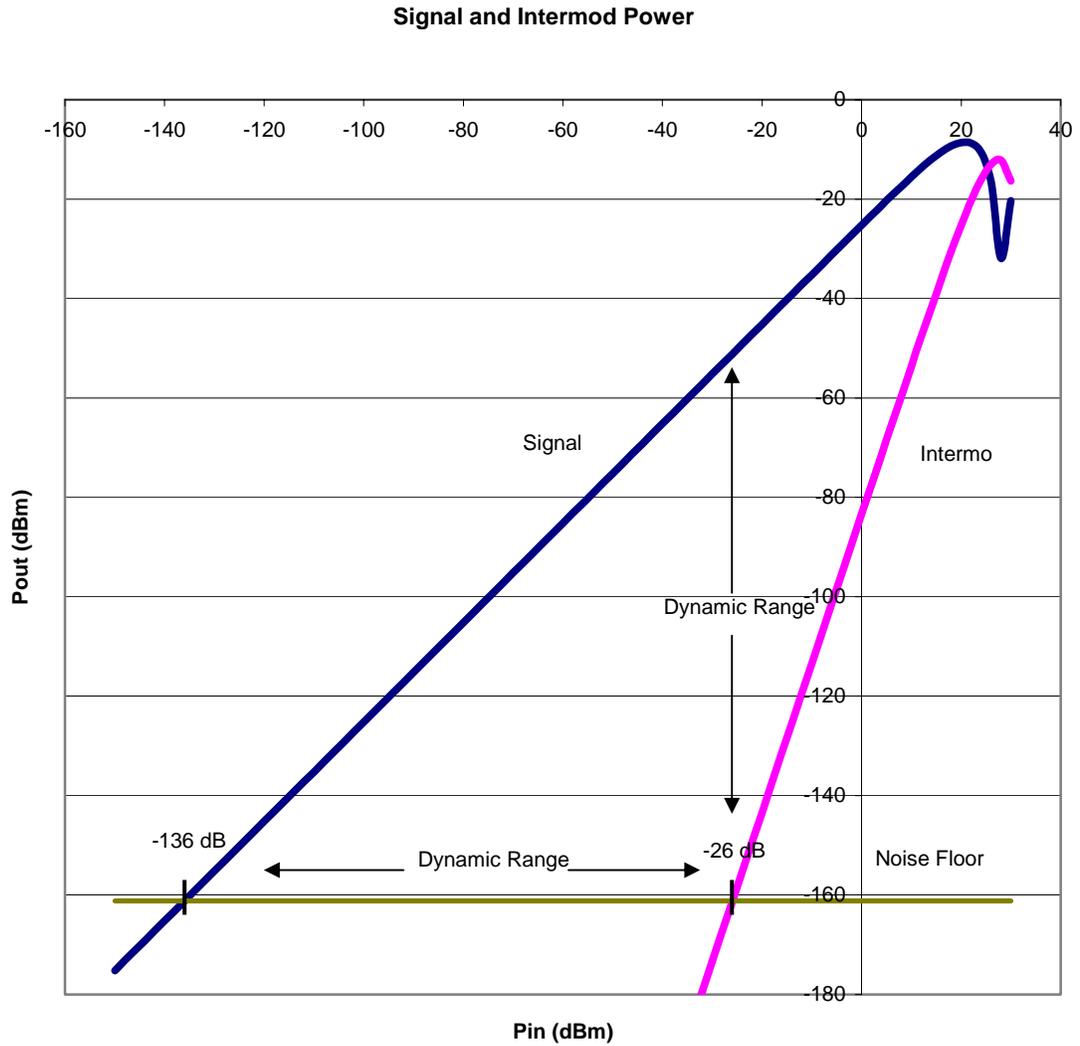


Figure 3 - 2: The signal (gain) and third-order intermodulation product of a link using a Mach-Zehnder modulator (MZM) and the parameters given in Table 3-1. Since there is no second harmonic distortion in a normally biased MZM, D_{dB} equals the intermodulation product shown. D_{dB} crosses the noise floor only once at an input power level of $p_0 = -26$ dBm. Since the signal has slope one, the dynamic range may be found from the difference between the signal and p_0 either horizontally or vertically in the plot, giving the familiar dynamic range triangle.

approximating a true delta function in the limit of large N . And for any N , this gives the maximum signal excursion for a given input power level, making it the worst case for distortion. A correlated but non-equal set of phases yields the best case, the Newman condition. The most common relationship is an uncorrelated normal distribution of phases, which yields very good results statistically. The multi-tone case will be discussed in more detail in the section on Multi-Tone Analysis at the end of this chapter.

Linearized Modulators

Many linearization schemes have been proposed in the literature, most of which consist of some combination of MZMs or DCMs and may bias the modulators away from the half-wave levels. In fact, there are so many different variations that it has prompted one researcher, G. Betts, to make a “Darwin Chart,” from Ref. [3.2] reprinted here as Figure 3 - 3 with modifications. The chart depicts the evolution of various linearization schemes from their early ancestors, the simple MZM and DCM. The calculations prepared for this thesis cover 24 different modulator transfer functions, although only six key modulators (seven including an unrealizable “reference” modulator) are analyzed in-depth in this chapter.

Super-Octave Modulators

A superoctave modulator is any modulator in which the passband is more than an octave wide (i.e., as defined above, the upper cutoff frequency is more than two times the lower cutoff frequency). This analysis addresses five broadband linearized modulators:

1. An ideal limiter, which is not physically realizable (LMTR),
2. Two Mach-Zehnder modulators in parallel optically (DPMZ),
3. Two Mach-Zehnder modulators in cascade optically (CMZM),
4. A directional coupler modulator with two passive bias sections in cascade (DCM2P),
5. A directional coupler modulator fed in both input arms, optical signals with equal amplitude and phase, with a specific electrical length (YFDCM).

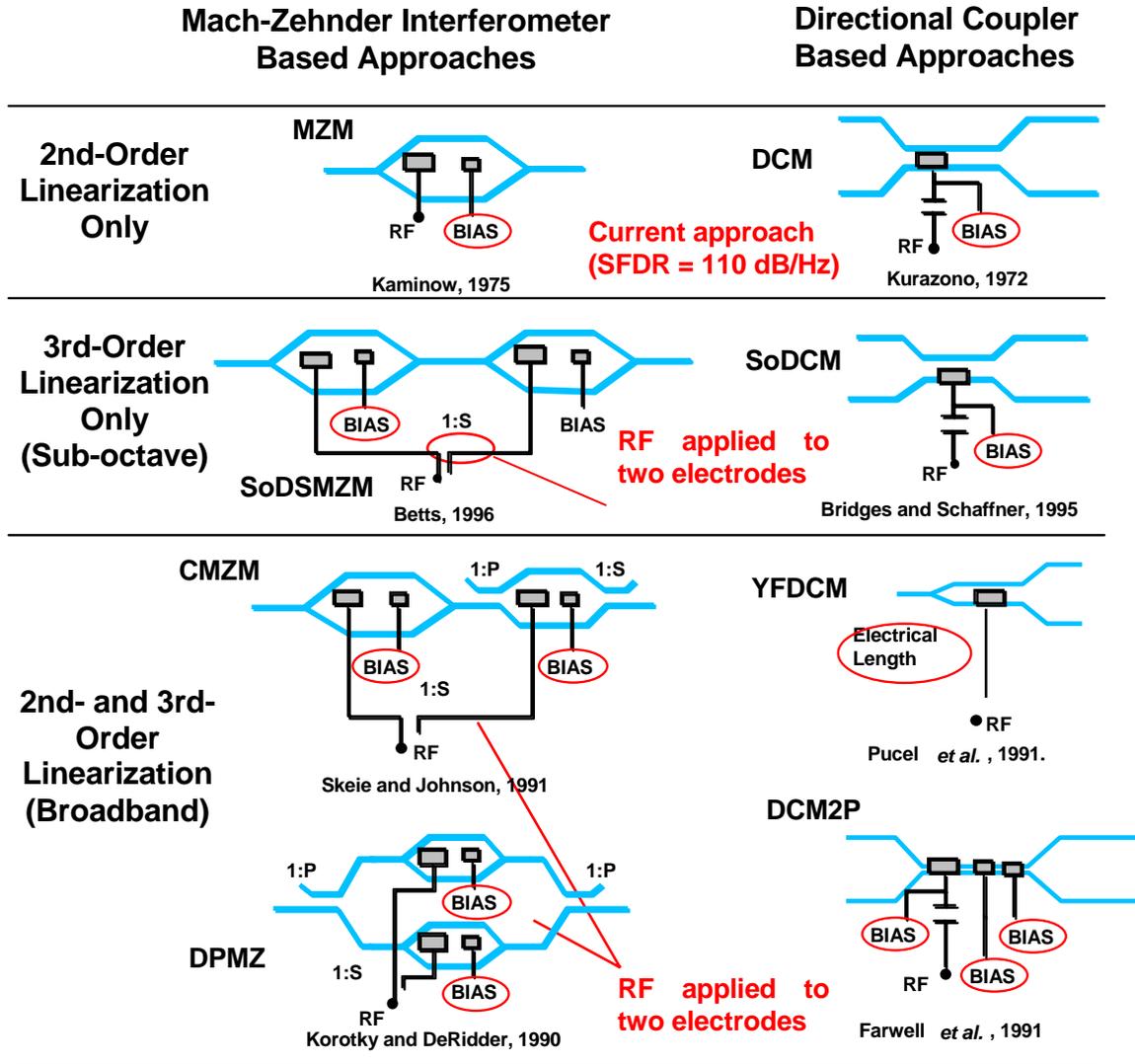


Figure 3 - 3: a “Darwin chart” of many proposed linearized modulators, reprinted from reference [Betts] and revised. Standard modulators MZM and DCM, sub-octave modulators SDSMZ and SDCM, and broadband modulators CMZM, YFDCM, DPMZ, and DCM2P. Note the DSMZ is shown with an r-f split, it doesn’t need one for linearization, however using one results in a lower shot noise link.

Figure 3 - 4 shows the d-c transfer function for these five super-octave linearized modulators. The input voltage is normalized to $V\pi$ or V_s as appropriate.

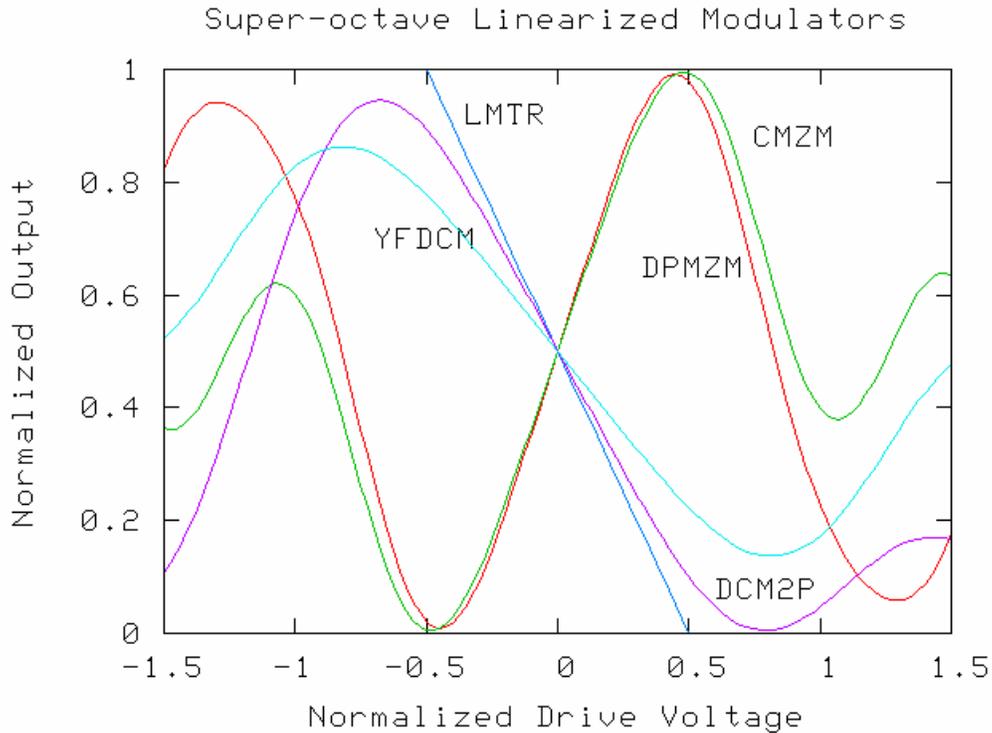


Figure 3 - 4: Five superoctave modulators: The ideal limiter LMTR; the directional coupler with two passive sections DCM2P; the dual parallel Mach-Zehnder DPMZM, the cascade Mach-Zehnder CMZM, and the Y-Fed directional coupler YFDCM. Slopes may be either positive or negative about zero with no affect on linearity.

The Ideal Limiter

A limiter is an idealization of an intensity modulator. Its output is a scaled copy of the input voltage, and it is hard limited at zero and one, the normalized maximum and minimum output. Although the transfer function is trivial around the bias point, the limiter is a useful analytical tool in system analysis. It sets an upper bound on the dynamic range improvement gained by linearizing a modulator, the difference between the system noise level and the power at which the laser is modulated from off to full intensity. Any higher power level causes clipping

distortion. In sub-carrier multiplexing, clipping is a major concern since the phases of the many carriers may randomly align and drive the modulator into its clipping regime (for real modulators this is an approximation).

The normalized d-c transfer function (switching voltage and optical output both normalized to unity, with zero bias voltage at the operating point) for an ideal limiter is

$$H = \min\left(\max\left(0, \frac{1}{2} - V_n\right), 1\right). \quad (3.7)$$

The Dual Parallel Mach-Zehnder Modulator

The DPMZ is described in Refs. [3.3] and [3.4] and shown schematically in Figure 3 - 3. The optical and electrical input signals are split unequally between two Mach-Zehnder modulators, the output optical powers are combined incoherently in two photodetectors, and the electrical signals from the photodetectors are added. Both of the MZMs remain biased at their half-wave voltage, thus no even-order harmonic distortion is generated in either modulator. The relative r-f and optical levels are chosen such that the third-order intermodulation distortion from each modulator is exactly the same, but the desired signals are different. The relative phases of the modulation signals are chosen so that the IMD products exactly cancel. While the signals do subtract, they do not completely cancel. For example the bias points on the cosine-squared transfer function are chosen to have opposite slopes. The signal level from this modulator may be a couple of dB below that of a single MZM, but the dynamic range is greatly improved. For example, one modulator has a large portion of the optical power, a small portion of the r-f power, and is biased at $0.5 V\pi$. The other modulator has the remainder of the r-f and optical power, and is biased at $-0.5 V\pi$.

The normalized d-c transfer function of a Dual Parallel Mach-Zehnder modulator (DPMZ) is

$$H = P \sin^2(\sqrt{S}(V - 0.5)) + (1 - P) \sin^2(\sqrt{1 - S}(V + 0.5)). \quad (3.8)$$

The parameters, S and P of Eqn. (3.8), are the r-f and optical power splits, respectively. An optimal point for these parameters is $\{0.1267, 0.9474\}$. However, there is a fairly large range of values for which some linearization is obtained provided that the splits are not too close to

zero (no cancellation) or too close to 0.5 (where signals would cancel along with the intermodulation distortion). The optimal points form a line segment in $\{S, P\}$, as will be discussed in Chapter 5.

This modulator has many alternative implementations in addition to the one shown in Figure 3 - 3. First, the link may have two separate fibers, in which case, the photodetector currents are summed as shown in the figure. Second, a 90° optical polarization rotator, may be inserted after one of the MZMs, putting the signals in polarization quadrature. The two signals are then combined in a single fiber, and detected in a single photodetector. Third, the two modulators can be driven by independent lasers at two slightly different wavelengths, and their outputs combined in a single fiber and detected by a single photodetector. The alternate implementations are chosen to reduce complexity and cost, but result in some performance compromises. For example, the polarization rotation scheme suffers from polarization cross talk in the fiber. It is interesting to note that the first experimental demonstration of a “Dual MZM” was made by Johnson at MIT Lincoln Laboratories (Ref. [3.3]), using only one MZM, but propagating both TE and TM optical modes. The sensitivity of TM and TE waves to the modulation voltage in lithium niobate is about 1:0.33, which is not the optimum of $(0.1267)^{0.5} = 0.356$, but is pretty close. The optical splitting ration was provided by rotation of the input polarization plane. Since the TE and TM waves are orthogonally polarized, they add incoherently on a single photodetector.

The Cascade Mach-Zehnder Modulator

The dual series Mach-Zehnder (CMZM) is described in by Skeie in Ref [3.5]. It is similar to the DPMZ in that there are two MZMs, biased at their half-wave voltage, and there is an r-f split between the two modulators, which may be provided by an external broadband r-f power splitter, or by splitting the r-f power equally and making the two modulators of different lengths. However, unlike the DPMZ the two modulators are in cascade, so there is no optical power split and the signals are coherently combined in the modulator, a major advantage over the DPMZ. Note that a close inspection of the schematic shown in Figure 3 - 3 shows an ideal passive coupler to implement the $1:s^2$ split, the actual derivation of the modulator

properties developed in Ref. [3.5] defines the modulator with two ideal “combiners,” one after each MZM. The transfer function for the DSMZ is

$$M_t = M_c(\gamma) \times M_p(x) \times M_c(\gamma) \times M_p(kx), \quad (3.9)$$

$$M_c(\zeta) = \begin{pmatrix} \cos \zeta & -j \sin \zeta \\ -j \sin \zeta & \cos \zeta \end{pmatrix}, \quad (3.10)$$

$$M_p(x) = \begin{pmatrix} e^{jx} & 0 \\ 0 & e^{-jx} \end{pmatrix}. \quad (3.11)$$

In Eqns. (3.9) - (3.11), ζ is the coupling parameter in the ideal coupler and x and kx are the phase shifts in the MZM arms. The r-f power split is represented by the k parameter, and the modulation signal is applied to both phase shifters. The coupling parameter ζ is controlled by a bias voltage. Skeie gives the results of $k=-0.5$ and $\zeta = 27$ degrees. These results were verified in our calculations.

Unfortunately, the ideal coupler needs a carefully controlled coupling angle and would likely need to be controlled by an active directional coupler (that is a directional coupler section with a tunable voltage) rather than simply depending on fabrication precision for the correct value of ζ . A directional coupler reduces to the form in (3.10) when it has zero voltage, but with non-zero voltage there is both a power split and a phase split, defined below in Eqn. (3.12). If the coupling function is performed by active modulator bias sections, then it requires two sections with different voltages to control both the power split and phase split independently. That is, it requires both a DCM and a phase shifter, or two DCM sections with different bias voltages. The CMZM requires two ideal couplers and thus four passive sections in series optically, making it challenging to manufacture on a finite-sized lithium niobate wafer.⁶ As shown in Figure 3 - 6, the transfer function of the CMZM is very similar to that of the DPMZM over a range of $\{-0.5, 0.5\}$ volts (normalized to $V\pi$), however they differ outside of that range.

The Directional Coupler with Two Passive Biases

A DCM followed by two unmodulated but dc-biased electrode sections on the same underlying directional coupler makes a broadband linearized modulator (DCM2P), described by Farwell in Ref. [3.6], and shown schematically in Figure 3 - 3. The electrode sections have electrical lengths $\{0.5\pi, 0.25\pi, 0.25\pi\}$, where the last two sections of length 0.25π have an electrical bias but no r-f drive. The two passive sections with different bias voltages linearize the directional coupler modulator. The lengths are not unique, but were chosen to be able to analyze a specific example. The first and principal directional coupler modulator section is fed with the full optical input to one arm only and biased where its 2nd harmonic is nulled, no different from an ordinary directional coupler modulator. The outputs from the two arms of this DCM section are complementary. The two passive sections create a relative power split and phase shift between these two modulated signals. As the two signals are coupled together through the passive sections, the third-order intermodulation product is nulled, but only with the correct power split and relative phase shift, introduced by the two passive sections. This requires two degrees of freedom and hence there are two passive sections.⁷

The optical amplitude of a “simple” directional coupler modulator is given by

$$B(\theta, V_N) = A(\theta, V_N) \times A_{IN} = \begin{pmatrix} m_{11} & -jm_{12} \\ -jm_{12} & m_{11}^* \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad (3.12)$$

$$m_{11} = \cos(\theta\sqrt{1+3V_N^2}) + j \frac{\sqrt{3}V_N}{\sqrt{1+3V_N^2}} \sin(\theta\sqrt{1+3V_N^2}), \quad (3.13)$$

$$m_{12} = \frac{1}{\sqrt{1+3V_N^2}} \sin(\theta\sqrt{1+3V_N^2}). \quad (3.14)$$

⁶ In Ref. [3.5] the difference between an ideal coupler and a DCM is not addressed. As a double check of the need to keep the phases aligned, we ran the optimization routines on the DSMZ with one DCM electrode after each modulator phase shifter, and we were unable to obtain a linearized modulator.

⁷ A DCM followed by one passive biased coupler section was proposed by Lam and Tangonen Ref. [3.7]. See also Bridges and Schaffner, Ref. [3.8], for further discussion. With only one additional degree of freedom it is not possible to simultaneously zero both second harmonic and third-order IMD. Thus this modulator falls in the sub-octave category discussed in the next section.

$B(\theta, V_N)$ is a 2x1 vector of complex amplitudes corresponding to the outputs of the two waveguides in a directional coupler. The parameter m_{11} is the “through arm,” that is, the amplitude of the optical wave that exits the same waveguide that it entered, and m_{12} is the “cross term,” that is, the amplitude of the optical wave that couples to the opposite waveguide from the one it entered. The modulation voltage is normalized to the switching voltage V_s by $V_N = V/V_s$. The electrical length θ is equal the coupling coefficient of the modulator times the physical length $\kappa\ell$. $A(\theta, V_N)$ is a unitary matrix; upon close inspection it is clear that the magnitude squared of the through-arm plus the magnitude squared of the cross term equals one for either input. The transfer function for a DCM is given by

$$\begin{aligned} h_1 &= |m_{11} \cdot a_1 - jm_{12} \cdot a_2|^2, \\ h_2 &= |m_{11}^* \cdot a_2 - jm_{12} \cdot a_1|^2. \end{aligned} \quad (3.15)$$

The transfer function for a DCM2P is given by

$$\begin{aligned} h_1 &= \left| (1 \ 0) \times A(\theta_3, v_{b3}) \times A(\theta_2, v_{b2}) \times A(\theta_1, v_{RF} + v_{b3}) \times A_{IN} \right|^2, \\ h_2 &= \left| (0 \ 1) \times A(\theta_3, v_{b3}) \times A(\theta_2, v_{b2}) \times A(\theta_1, v_{RF} + v_{b3}) \times A_{IN} \right|^2. \end{aligned} \quad (3.16)$$

The reference bias values for a DCM2P in normalized voltage for Eqn. (3.16) is $\{0.50529, 0.73803, 0.77000\}$, which is the normalized bias voltage for the r-f, and 2nd and 3rd passive bias section, as in Ref. [3.8]. This modulator has the advantage of no r-f split and no optical split. Furthermore, the only precise control needed is the bias voltage of the three directional coupler sections. Bias voltages are the easiest modulator parameter to fine tune. However, as will be shown Chapter 4, it is unusually difficult to preserve this linearization in the presence of velocity mismatch.

The Y-fed Directional Coupler

The Y-fed directional coupler modulator (YFDCM) has a Y-junction similar to a Mach-Zehnder modulator, which feeds both inputs of a directional coupler modulator with signals of equal optical amplitude and phase, as shown schematically in Figure 3 - 3. This modulator was first proposed in Ref. [3.9] in 1986. The physical symmetry of this configuration forces the

transfer function to be anti-symmetric. Similar to the MZM (and different from the DCM), all even-order harmonics are zero at the operating point, $V=0$. He proposed an electrical length of 0.707π , which is an innovative implementation of a 1x2 switch in that allows a complete transfer of optical power, and does not require any bias as would be required in a MZM or DCM. However, to use this structure as a linear modulator the electrical length should be $\pi/2$.

There are many variations of this modulator created by changing the electrical length, adding a passive (and unbiased) section and changing the electrode topology. As a matter of convention in this thesis, “standard YFDCM” is used to describe the Y-fed DCM with arbitrary electrical length, that may or may not be linearized. YFDCM is taken to be a specific version of the linearized modulator based on Thaniyavarn’s topology but with electrical length 1.43π . There are other ways of linearizing the standard YFDCM. First, it was found by Pucel, Ref. [3.10], that making the modulator a specific electrical length L_A , followed by an additional passive coupling length L_p (that is, without the modulation applied, and no bias) would cause the cubic distortion term to null, without upsetting the null of the second harmonic (after all, this is still an anti-symmetric device). Pucel gives the linearized values as $L_A=0.6\pi$ (active length) and $L_p=0.6121\pi$ (passive length). Second, in Ref. [3.11] Tavlykaev and Ramaswamy found that taking the YFDCM (as proposed by Thaniyavarn) and breaking the electrode and reversing the phase on the second section also creates a linearized modulator. In other words a modulation voltage V is applied over L_A and $-V$ is applied over L_B . They give values of $L_A = 2.05\pi$ and $L_B = 0.5025\pi$. They also demonstrate this modulator experimentally in Ref. [3.12].⁸ Third, in Ref. [3.13] it is shown that the standard YFDCM is linearized by simply increasing the electrical length of the modulated coupler to 1.4297π . In the author’s analysis program, it is shown that for both the standard DCM and the standard YFDCM, there is always a range of voltages as a function of electrical length, in which the third-order intermodulation distortion is nulled without a corresponding signal null. In the standard DCM this voltage can never be

⁸ Tavlykaev and Ramaswamy use a different definition for dynamic range, which may be adequate for CATV applications, but is not the definition taken in this study. They look at how close to the switching voltage the modulator transfer function can get before the error between the modulator and an ideal limiter is $\pm 1\%$. Their experimental results were positive under this definition. However, the author of this study has not found that this modulator configuration gives good results for SFDR as defined in Eqn. (3.6).

zero for any electrical length, as there is no signal gain at that point. However, for the standard YFDCM, this voltage is zero at the electrical length 1.4297π . This is the only linearized YFDCM treated in detail in this thesis.

The d-c transfer function of the YFDCM is found from Eqn. (3.15) taking one-half of the through-arm amplitude and one half of the cross-term amplitude of the DCM Eqn. (3.12), that is the terms a_1 and a_2 of vector A_{IN} are each equal to $\frac{1}{2}$

$$A = \frac{1}{2} \cos(\theta\sqrt{1+3V_n^2}) + \frac{j}{2} \times \frac{\sqrt{3V_n-1}}{\sqrt{1+3V_n^2}} \sin(\theta\sqrt{1+3V_n^2}), \quad (3.17)$$

$$H = |A|^2 = \frac{1}{2} - \frac{\sqrt{3V_n}}{1+3V_n^2} \sin^2(\theta\sqrt{1+3V_n^2}). \quad (3.18)$$

The YFDCM is particularly sensitive to the optical phase in the Y-splitter. An optical phase shift of 180° in the input to the two arms completely cancels the signal and produces only second harmonic. To keep the second harmonic below the third-order intermodulation requires a phase error of less than one optical degree depending on the operating signal bandwidth, Ref. [3.13].

It is also helpful to represent the transfer functions in terms of the physical directional coupler parameters; length ℓ , electrically induced coupling δ , coupling coefficient κ , see Ref. [3.14]:

$$H = \frac{1}{2} - \frac{\delta\kappa}{\delta^2 + \kappa^2} \sin^2\left(\ell\sqrt{\delta^2 + \kappa^2}\right). \quad (3.19)$$

Sub-Octave Modulators

Sub-octave modulators produce significant second harmonic distortion, and are thus suitable for use over only a single octave of modulation frequency. All harmonic distortion is then out of band, and only the intermodulation terms contribute to the nonlinear distortion. These modulators are generally biased into a highly quadratic region of the modulator transfer function. Specifically, the intermodulation distortion is nulled at the expense of second harmonic distortion, which is in excess of that of a standard DCM and MZM, but since the

extra harmonic distortion is out of band, it does not interfere with the signal. Sub-octave modulators are inherently simpler devices than super-octave modulators, since they need only null one type of distortion. Both a Mach-Zehnder and a directional coupler based sub-octave modulator can be made with one precision parameter (bias voltage) and no r-f split.

The Sub-octave Dual Series Mach-Zehnder Modulator

The Sub-octave modulator, the Dual Series Mach-Zehnder (SoDSMZ) is described by Betts in Refs. [3.15] and [3.16]. It consists of two simple MZMs in series optically, but in parallel electrically. There is no required r-f split between the two modulators, and they both have the same bias.⁹ One monolithic electrode structure may cover two Mach-Zehnder optical interferometers. Thus its fabrication is as simple as modifying the waveguide design from a standard Mach-Zehnder by placing a Y-junction combiner followed by a splitter in the center of the modulator. The gain is reduced from a normal Mach-Zehnder by the fact that the modulator is effectively half as long (accommodating two modulators in series) and from the residual signal cancellation that is a necessary byproduct of the intermodulation cancellation.

The d-c transfer function for this modulator is

$$H = \cos^4(V). \quad (3.20)$$

A normalized bias voltage of $V_b = 0.581$ nulls the third-order intermodulation distortion ($V=V_b+V_{rf}$).

The Sub-octave Directional Coupler Modulator

The Sub-octave Directional Coupler (SoDCM) only differs from the normal directional coupler in its bias. The d-c transfer function is given by Eqns. (3.12) - (3.14) just as a simple DCM. The name SoDCM is used to indicate the sub-octave usage of the DCM. The optimal bias voltage is $V_b = 0.795 V_s$. The usage of a DCM as a sub-octave modulator was first proposed by Bridges and Schaffner in Ref. [3.8]. This is the most simple linearization scheme

⁹ However, in some implementation an r-f split is added (which adds complexity) and that results in a reduced average photo-current level and subsequently a reduced shot noise. This analysis treats the DSMZM with no r-f split.

conceptually, in that it is the only “simple” modulator that has a native linearization mode obtained through a precision voltage bias.¹⁰

Comparison of SoDCM and DSMZM

Figure 3 - 5 shows the transfer function for the sub-octave Dual Series Mach-Zehnder Modulator (SoDSMZ) and the simple directional-coupler-based sub-octave modulators (SoDCM). These transfer curves are translated by the bias voltage such that the abscissa value is zero for no r-f drive and the voltage is normalized to either $V\pi$ or V_s . The operating bias points are shown by a “+” sign on the graph. A salient feature of these curves is the normalized optical output at the target bias, which in both cases is quite small, unlike the broadband and ordinary electrooptic modulators that are biased at or near the half power level. Even without looking at the frequency spectrum it is clear from the curvature about the biases that these modulators obtain their intermodulation suppression at the expense of large even-order harmonic distortion.

There is a substantial d-c optical attenuation at these biases. The d-c output of the two modulators is 7% and 14% of the input optical level, a reduction of 11.5 dB¹¹ and 8.5 dB, compared to only 3 dB for broadband modulators. However in shot-noise-limited links, this results in improved link performance, since reducing the d-c optical level reduces the system noise level as well. The small signal gain depends on the slope of the transfer function at the bias point. In both cases moving from about the middle optical level to the sub-octave bias point reduces the small signal gain less than it reduces the shot noise resulting from the d-c optical level. The sub-octave modulators can have a higher dynamic range than the super-octave modulators for this reason. Of course, in a laser RIN-limited optical link, the performance is degraded by the loss of signal gain.

¹⁰ That is, assuming it is just as easy to fabricate a directional coupler as an MZM. The state of the art is such that manufacturers have had more success at making consistent, repeatable, MZMs than DCMs.

¹¹ This reduction is only for the through arm of the DCM. The output of the cross arm has a complementary transfer function and thus a much larger zero signal optical power, resulting in significantly more shot noise, and a degraded dynamic range.

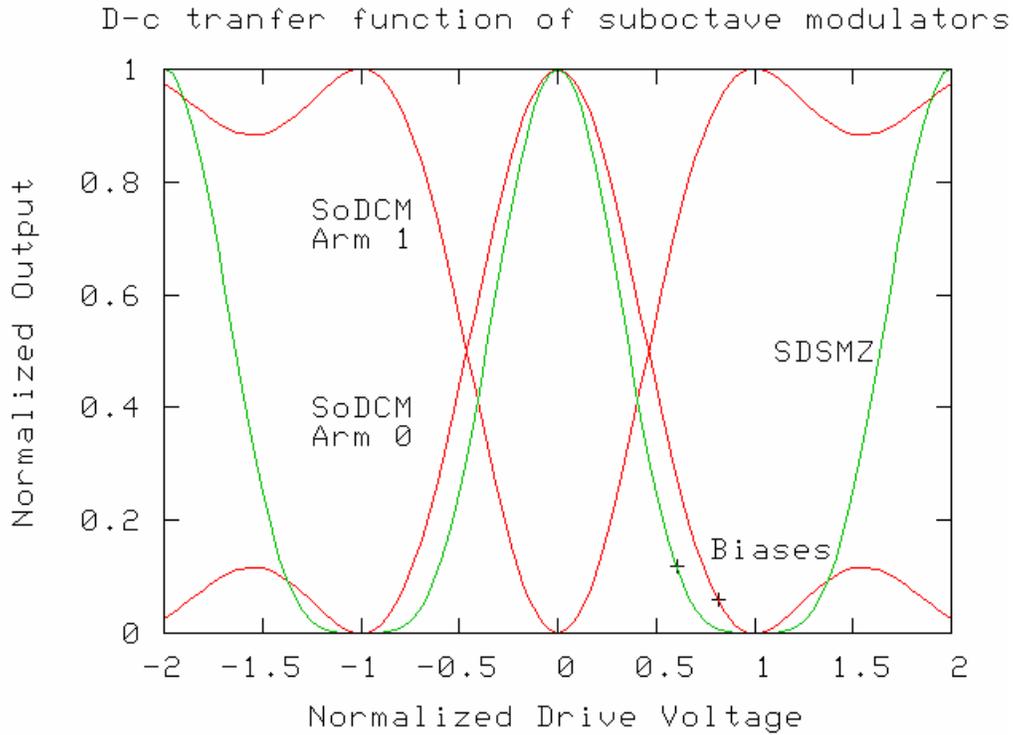


Figure 3 - 5: Two sub-octave modulators: the Dual Series Mach-Zehnder (DSMZM) and the suboctave directional coupler (SoDCM). For reference both output arms of the DCM are shown; however, one should only use the arm with the low output photocurrent. The biases are shown with hash marks at $V=0.79$ Vs for the SoDCM and $V=0.58$ $V\pi$ for the DSMZM.

Other Modulators

This analysis has addressed two normal electrooptic modulators and five linearized electrooptic modulators (not counting the unrealizable ideal limiter). The number can grow substantially by introducing some minor alterations to the modulators. In some cases, modulators have been left out because there is not a strong interest to build them. For example, a triple parallel Mach-Zehnder has been proposed to null the fifth-order intermodulation distortion product; see Ref. [3.8]. In other cases, the modifications are not fundamental, such as replacing a Y-splitter with a passive directional coupler. When minor variations are taken into account, the modulators count goes from 7 to over 20.

Performance of Electrooptic Modulators

This section compares the performance of the linearized modulators described in the previous section.

Ideal Performance

The ideal performance is characterized by Eqn. (3.1) for gain and Eqn. (3.6) for the two-tone dynamic range. The calculation makes a number of assumptions. First, the d-c transfer functions given in the previous section are approximations of the real behavior of the modulator. The analysis ignores all electrooptic effects but the first order Pockel's effect and any other non-linearity in the materials present in the modulator. (Though not well-understood, there is evidence for "bias drift" in lithium niobate electrooptic modulators due to ion migration in the material under applied dc fields. How this contributes to r-f modulation non-linearity is not understood at all). The directional coupler equation is an approximation from coupled mode theory. From the engineering point of view, the equations assume the waveguides are purely single mode, and that the optical fields are linearly polarized along the correct axis for the given modulator. It is assumed that there is no radiation, loss, or reflection, and that the directional couplers and Y-junctions act as perfect couplers and combiners and the devices are perfectly symmetric. Furthermore, it is also assumed that the detectors exhibit a perfect square law behavior, even though deviations are significant at high power levels, as noted before [3.1].

Figure 3 - 6 shows the signals, noise, and nonlinear distortion terms of the five superoctave modulators defined above, the LMTR, DPMZ, CMZM, YFDCM, and the DCM2P. The modulator linearization parameters were adjusted to maximize the dynamic range for the noise level shown, -161 dBm, appropriate to a 1 Hz noise bandwidth for a shot-noise limited link. For the link parameters given in Table 3.1, both thermal noise and laser RIN noise were essentially negligible. It is important to note that the optimized linearization parameters will change with noise level due to the complicated behavior of the distortion curves with modulation drive power; they are not straight lines of slope 3, as they are for simple modulators. The intermodulation distortion curves, when optimized, are approximately slope 5 above the noise level and slope 3 below the noise level (ignoring the null at a specific input

power). Changing the noise bandwidth will simply translate the noise level upward in dB, but the linearization parameters should be readjusted for maximum dynamic range at the desired noise level. Note that the noise level for all of these modulators is the same, independent of photocurrent, since all are biased at 50% transmission. Comparisons with other modulators may have to account for changed levels of average photocurrent, and thus shot noise level, for example, in comparing a simple DCM biased at its “normal” bias to eliminate second harmonic or at its “linearized” bias to null the third-order intermodulation distortion; the two conditions produce different shot noise levels for the two different biases.

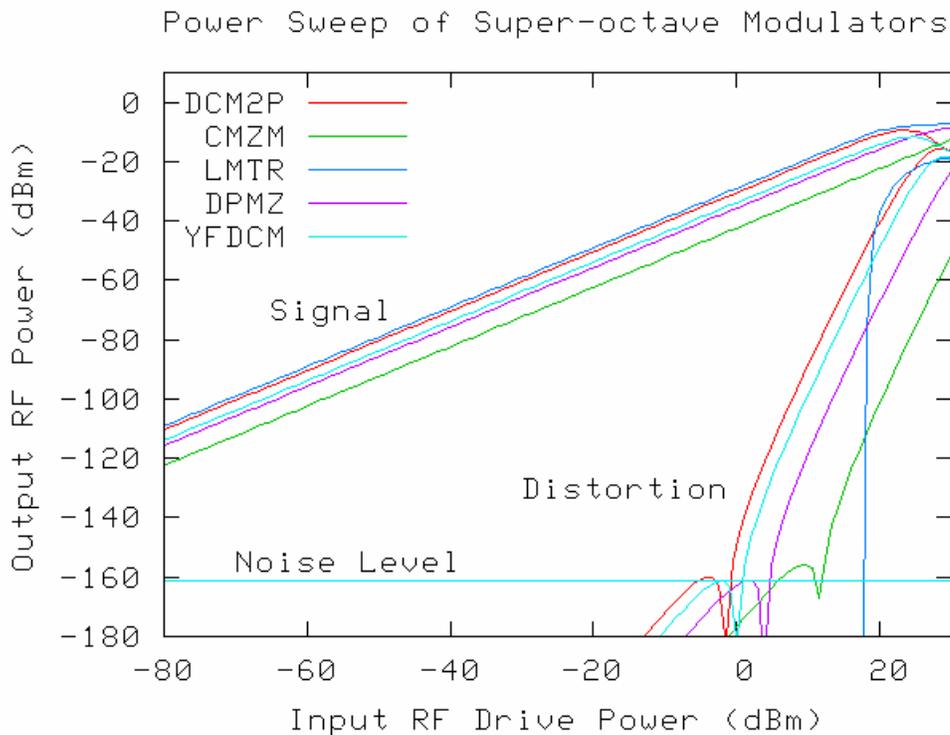


Figure 3 - 6: Power sweep of signal, second harmonic, and intermodulation distortion of five super-octave modulators, DPMZ, CMZM, DCM2P, YFDCM, and DCM2P and the ideal limiter reference assuming a signal bandwidth of 1-Hz. Note that the ordinate and abscissa dB scales are not the same size, so the “slope 1” signal does not appear at 45°

The modulators are shown together in Figure 3 - 6 to compare their different characteristics. The limiter has the largest gain, not surprisingly the ideal linearized modulator is the most efficient since it is not using any part of the signal to cancel distortion (note the gain is still worse than the standard modulators). Impressively, the DCM2P has only a slightly lower gain. The CMZM and DPMZ have inferior gains, each about 6 dB lower than the DCM2P. The second harmonic is not shown on this figure because in all cases but the DCM2P, there is no even-order harmonic distortion. The intermodulation distortion for these modulators is shown as the representative distortion. The maximum distortion-free r-f input level for the CMZM and DPMZ is higher than the DCM2P, but this does not translate to a higher dynamic range since it exactly makes up for the lower output signal level. The distortion of the limiter comes in at a near vertical rise at an input level of 18 dBm. This is purely clipping distortion.

Figure 3 - 7 shows the power sweep of the DCM2P independently from the other super-octave modulators. In this figure, both the second harmonic and the intermodulation distortion appear in the plot window. Like the intermodulation distortion, the second harmonic is only nulled at a single finite power level. This property is the cause for the difficulty in using this modulator that will be illustrated in detail in Chapter 4. Note that the two distortion terms appear to have the same shape in power levels above the null. However, below the null, the second harmonic takes on “slope 2” and the third-order intermodulation term takes on “slope 3.”

Figure 3 - 8 shows a power sweep of the signal and distortion terms for the two sub-octave linearized modulators: the SoDCM and the SoDSMZM. The signal level for the SoDCM is a little lower than the DCM2P (from Figure 3 - 7). And that of the SoDSMZM is a little higher than the CMZM (again from Figure 3 - 7). The second harmonics are plotted here, to illustrate how large they are, but this distortion is out of band, and therefore does not affect performance. In the case of the SoDCM, there are two output arms. The signals are the same but the noise is not. The noise level is -170.8 dBm for the good arm of the SoDCM (-158.5 dBm for the bad arm of the SoDCM) and -167.6 dBm for the SoDSMZM. The plot assumes the output arm with the lower average photocurrent level is the one chosen. The chosen noise levels are lower than the super-octave modulators and that accounts for a modest increase in dynamic range of the sub-octave modulators.

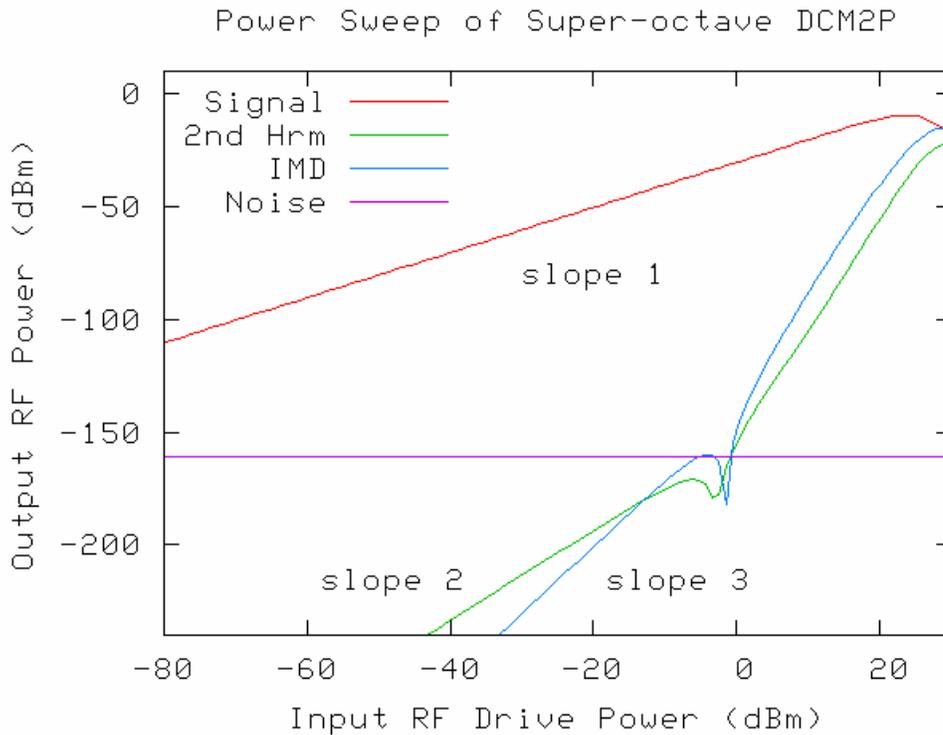


Figure 3 - 7: Power sweep of DCM2P showing both third-order intermodulation product and second harmonic distortion.

All of the curves of the distortion in the linearized modulators of Figure 3 - 6 and Figure 3 - 8 display a characteristic named the “kissing condition.” The distortion is nulled at a finite power level. The location of this null as a function of input power is controlled by the actual biases of the linearization mechanism in each modulator. At power levels starting just below the null, the distortion comes back to a local maximum and then tapers off at slope 3 for IMD (or slope 2 in for the DCM2P where the second harmonic is also nulled at a finite power level) as the power is decreased further. In optimizing the bias values, the null moves up to higher power levels, until that local maximum just “kisses” the noise floor. Unfortunately, this condition is very sensitive to errors in bias and frequency effects (studied in Chapter 4). Once a small error is introduced, and that local maximum rises above the noise floor, then the dynamic range is reduced discontinuously by a few dB. Note this effect, as it explains the

peculiar discontinuous drop in dynamic range that occurs in some of the calculations later in this study.

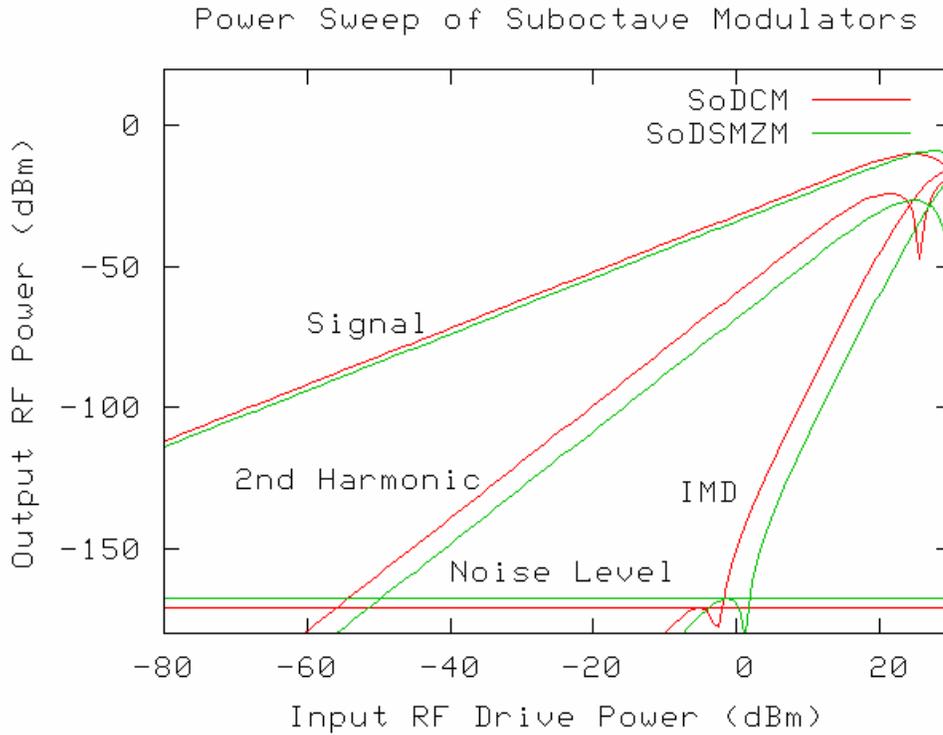


Figure 3 - 8: Power sweep of the signal, second harmonic and intermodulation distortion of two sub-octave modulators, SoDCM and SoDSMZM compared to a “noise floor” assuming a 1-Hz signal. Note the noise floors are different because the average photo current is different.

Table 3-2 summarizes the performance of standard, super-octave, and sub-octave modulators.¹² The salient feature of the table is how similar modulators within a class are and how different they are from the other classes (not counting the unrealizable limiter). The standard MZM and DCM are included for comparison; they both have nearly the same gain and dynamic range. Note that the difference between the gain of the DCM2P and that of both the DSMZ and DPMZ is 6 dB. Given the similarity in performance of the linearized

¹² Note that the noise figure for the DSMZ can be greatly reduced by the introduction of an r-f split, as reported in Ref. [3.15].

modulators, it is far more important to judge a modulator by its ease of manufacture, operation, and reliability, than by the performance calculations of idealized modulators.

Table 3-2: Modulator Figures of Merit

Modulator	Gain (dB @ 1 Hz)	Dynamic Range (dB @ 1 Hz)	Dynamic Range (dB @ 4 MHz)	Noise Figure (dB @ 1 Hz)
Standard Modulators				
MZM	-25.2	110.9	66.8	37.96
DCM	-24.8	110.4	66.4	37.93
Super-octave Modulators				
Limiter	-29.1	150.1	84.8	41.88
DPMZ	-36.0	130.1	77.4	48.57
DSMZ	-36.4	130.5	77.3	55.21
DCM2P	-30.5	130.3	77.2	43.27
YFDCM	-33.9	128.1	74.5	46.67
Sub-octave Modulators				
SDSMZ	-34.0	131.4	78.6	40.37
SDCM	-31.8	134.8	82.2	35.05

Noise Scaling

Dynamic Range is a linear function of the logarithm of the noise bandwidth with a slope that depends on the degree of linearization, provided that the intermodulation is given by a simple slope on the log-log plot, which is not actually the case, as seen in Figure 3 - 6 and Figure 3 - 8, however it is a good approximation for the analysis of noise scaling. This relationship is valid up until the r-f drive saturates the modulator. The dynamic range of any modulator scales as

$$D(\Delta F) = D \Big|_{\Delta F=1Hz} - \frac{n-1}{n} 10 \log(\Delta F), \quad (3.21)$$

where ΔF is the noise bandwidth and the parameter n is the order of the dominant distortion term. In a standard modulator $n=3$, so that the dynamic range varies as the $2/3$ power of the bandwidth, since there is no second harmonic and the modulator is limited by the third-order intermodulation term. Recall Figure 3 - 2 in which the signal is a line with slope 1, and the third-order intermodulation product is a line with slope 3. The “noise floor” is a third, horizontal line that forms a triangle with the signal and the intermodulation line. The length of the base of this triangle is the dynamic range. The vertical position of the noise line is proportional to the $\log(\Delta F)$, so from simple geometry, it is clear that the dynamic range goes as $(\Delta F)^{2/3}$. In both the super-octave and sub-octave linearized modulators described above, the dominant distortion is the fifth-order intermodulation term, where $n=5$ and the slope is $4/5$.¹³ (More complicated realizable linearization schemes can result in even steeper slopes for intermodulation, for example, the triple MZM modulator discussed in Ref. [3.8], which has slope of 7 for distortion). In the ideal limiter, n is nearly infinite and so is the slope of $D(\Delta F)$, at the onset of clipping.

In some systems, the noise bandwidth can be reduced to match the signal bandwidth, but in others it cannot. For example, if the signal is wide band or if one does not know what frequency the signal will have (e.g., signal surveillance) then it is necessary to design for a large

¹³ This is an approximation. The actual noise curve is not a line with slope 5 because there is a finite null. However, sufficiently far away the null, in the direction of greater signal power, the noise curve converges to a line of slope 5. Interestingly, sufficiently far away from the null in the direction of lower signal power, the noise curve converges to a line of slope 3.

system bandwidth. When the absolute dynamic range is critical, and the signal is narrow-band, then reducing the noise bandwidth is a clear system objective. The higher degree of linearization (larger n), the more the dynamic range increases per reduction in noise bandwidth. However, some applications have fixed noise bandwidth such as the 6 MHz AM-VSB (vestigial sideband) channels in CATV. When the noise bandwidth isn't a design parameter, this scaling rule influences the decision to use a linearized modulator. For example, Table 3-2 shows that the ideal limiter gives 15.3 dB above the SoDCM, the best third-order linearized modulator in the 1 Hz bandwidth, but only 2.6 dB above this modulator in a 4 MHz bandwidth application. Assuming that one could make both an effectively ideal limiter and a 1 Hz system at finite but high cost, it might make sense to employ the ideal limiter in the 1 Hz system, but probably not in the 4 MHz system. A more realistic example, consider a 4 MHz super-octave system versus a 4 GHz system. The linearized modulator increases the dynamic range by 10 dB in the 4 MHz system but only 3 dB in the 4 GHz system.

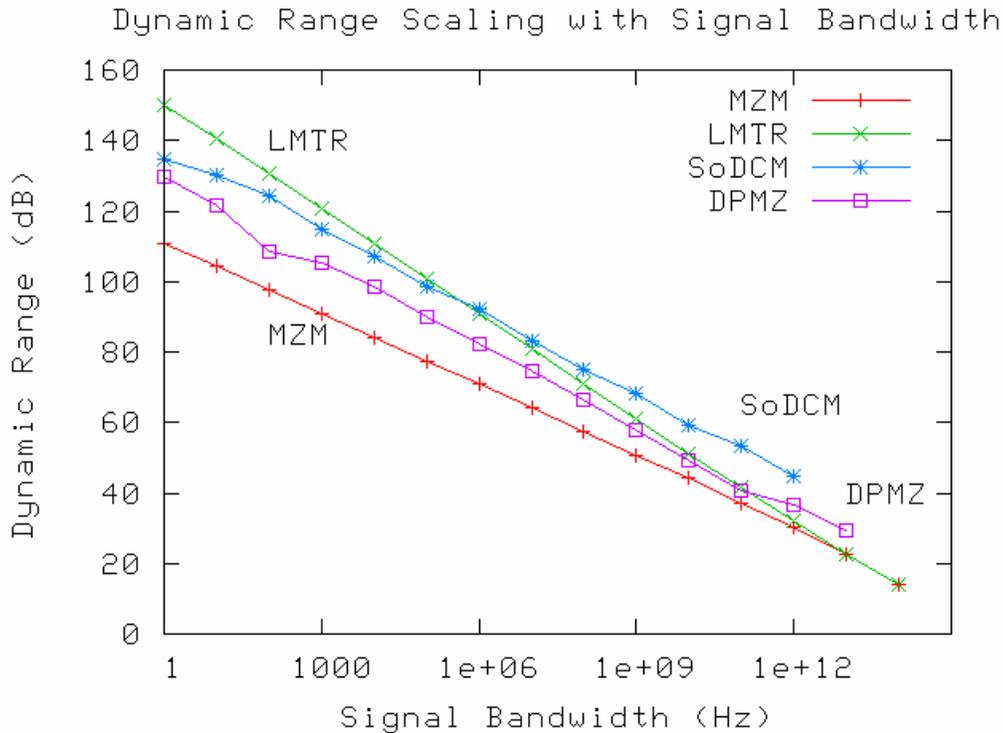


Figure 3 - 9: Dynamic range as a function of signal bandwidth ignoring finite transit time and losses of the MZM, LMTR, SoDCM, and DPMZ.

Figure 3 - 9 illustrates the signal bandwidth scaling of the various types of modulators. The MZM is the canonical non-linearized reference limited by third-order intermodulation, and the LMTR is the theoretical perfectly linearized modulator limited only by clipping. These curves converge to a single point at which standard, third-order intermodulation linearized and “perfect” modulators become indistinguishable.¹⁴ To first-order, this convergence point is universal for links with identical canonical parameters, equivalent gains, and quiescent photocurrent levels. The SoDCM and DPMZ are plotted within this window. The SoDCM was chosen for its superior shot-noise performance, and the DPMZ because it is the most widely known super-octave modulator. There is some clumpiness in the SoDCM and DPMZ curves, due to the difficulty of root finding in the dynamic range equations. Interestingly, the SoDCM outperforms the LMTR at high signal bandwidth because it has a lower photocurrent level and hence lower shot noise level in the DCM through-arm, and it retains sufficient gain.

Given these link parameters and a specified system bandwidth, an engineer then has three (possibly four) different specs for dynamic range depending only on whether he chooses a modulator on the curve with slope $1/2$, $2/3$, or $4/5$ (or possibly $6/7$). Having chosen the class of modulator, choosing the actual modulator is an economic decision. Much of the linearized modulator literature ranks modulators by their 1 Hz theoretical dynamic range; however the dynamic range differences for practical noise bandwidths will be considerably smaller, and the choice will depend more on cost and complexity. It is important to identify the modulator's class, and then rank them by reliability and manufacturability.

Parameter Sensitivity

The idealized performance of electrooptic modulators given in the literature and summarized above assumes optimal control over fabrication and operating parameters. Standard modulators have a bias voltage (either built in or applied during operation); and linearized

¹⁴ The curves above 100 GHz are purely for theoretical illustration, in that LiNbO₃ is electrically lossy at 1 THz. Other physical assumptions may also break down at frequencies above 100 GHz.

modulators may have two additional bias voltages, or an optical split or coupling ratio and an r-f split in addition to regular modulator biases. The tolerance for standard modulators is tight, and for linearized modulators at low noise bandwidth the requirements may be so severe that doubt is cast on whether the modulators could be successfully manufactured. The tolerance becomes less severe with increasing noise bandwidth, however.

Figure 3 - 10 shows dynamic range of a standard Mach-Zehnder modulator as a function of the error of its normalized bias voltage $0.5 V\pi$. There are eleven curves corresponding to decade increments of noise bandwidth from 1 Hz to 10 GHz. For each order of magnitude decrease in noise bandwidth, there is a $6 \text{ and } 2/3$ dB increase in dynamic range at zero error in bias, agreeing with the predictions of noise scaling. Every curve consists of three discontinuous sections. The central region is the intermodulation-product limited dynamic range. This region is terminated on either side by two corners beyond which the dynamic range drops off dramatically. This is the transition between intermodulation limited dynamic range and second-harmonic limited dynamic range. The unrealistic 1 Hz bandwidth curve can only tolerate a $0.001 V\pi$ error before the dynamic range falls off, and by an error of $0.008 V\pi$ the dynamic range has lost 10 dB. The modulator becomes less sensitive at higher noise levels. At a realistic system bandwidth of 1 MHz, the dynamic range is flat out to an error of $0.0085 V\pi$ and doesn't lose 10 dB until the error in voltage is $0.09 V\pi$. For this reason, in the literature many results are now reported with 1 MHz bandwidths instead of the traditional 1 Hz bandwidth. The curves also illustrate why active feedback control is used on the bias voltage to minimize second harmonic content.

Figure 3 - 11 shows the same sweeps for a sub-octave directional coupler modulator. The dynamic range is purely a function of the intermodulation distortion. The parameter tolerance is tighter than that of simple Mach-Zehnder. The 1 Hz curve loses a dB at a tolerance of about $\pm 0.00005 V_s$ and it loses 10 dB at $\pm 0.00075 V_s$. Near the linearized optimum the drop off is very rapid, but by a 10% error the 1 Hz curve converges to around 110 dB. This is the dynamic range of a standard directional coupler. By spoiling the bias the modulator moves from the bandwidth scaling curve of slope $4/5$ to that of slope $2/3$, just as the Mach-Zehnder went from $2/3$ to $1/2$. In other words, The SoDCM goes from being fifth-order IMD limited

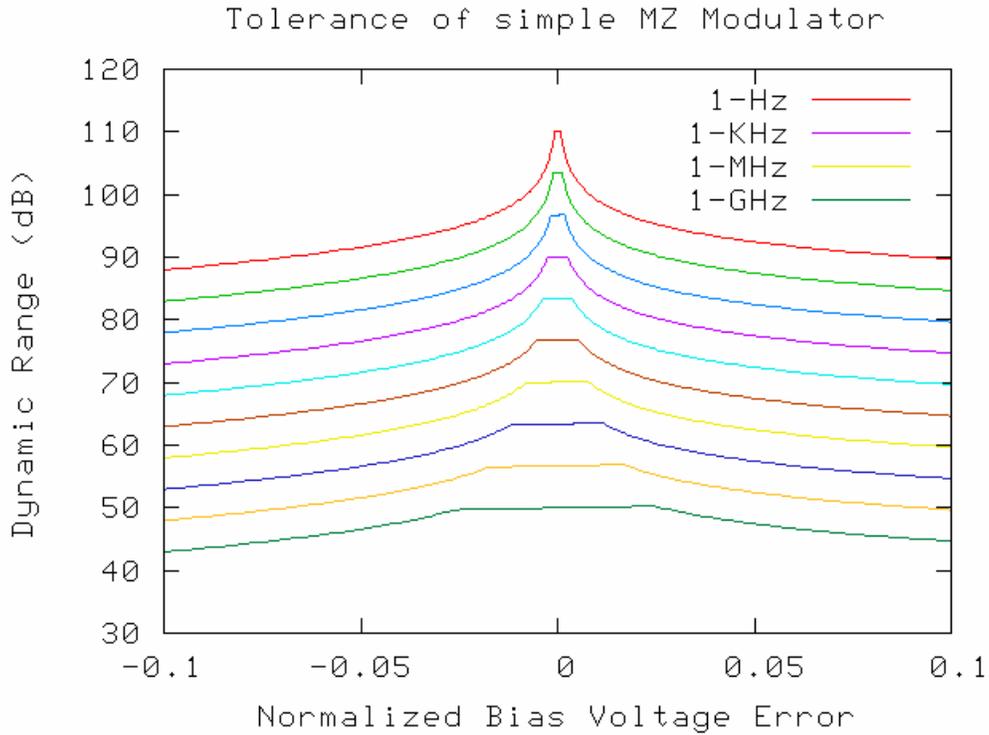


Figure 3 - 10: the dynamic range of the simple MZM as a function of the bias voltage swept over $\pm 10\%$ of $V\pi$ for 10 orders of magnitude of signal bandwidth from 1 Hz to 1 GHz.

to third-order IMD limited, just as the MZM went from third-order IMD limited to second harmonic limited. The SoDCM does not fall to the $1/2$ line since its second harmonic distortion falls out of band. The points at zero bias error are related across the 11 scans by a $4/5$ scaling law while the skirts of the curves are related by a $2/3$ law.

Figure 3 - 12 shows the logarithm of the tolerance as a function of the logarithm of signal bandwidth of a MZM. The curves are points of equal reduction from the optimal dynamic range at that noise bandwidth. There are ten curves from 1 dB degradation through 10 dB. It is remarkable that the relationship is almost perfectly linear. Only on the 10 dB curve near a saturated bandwidth is it not linear to the eye, and that is at a 25% error in bias. It turns out that every modulator addressed in this study demonstrates this linear relationship.

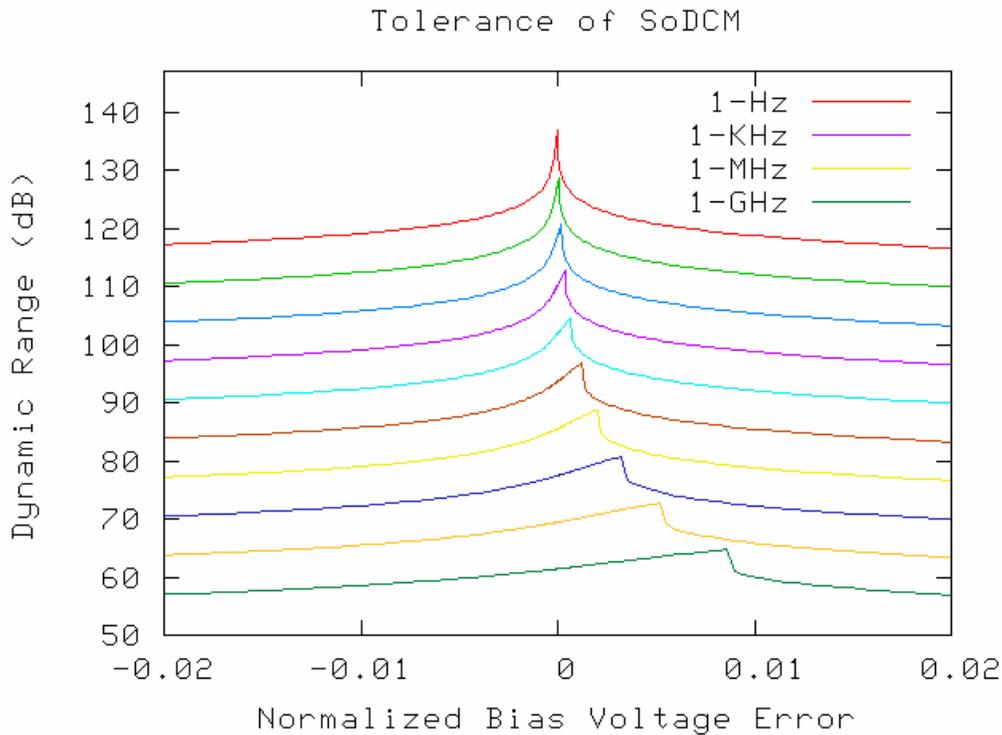


Figure 3 - 11: The dynamic range of the SoDCM swept over $\pm 2\%$ of the optimal 1-Hz bias voltage $0.79 V_s$ for 10 orders of magnitude of signal bandwidth, 1 Hz to 1 GHz. At larger noise BW the discontinuous drop in dynamic range is a result of the “kissing condition” described above.

Comparing the linearized modulators treated in this study presented some problems in that the linearization is accomplished by varying different physical parameters of the modulator in the different cases. However, if all the parameters are sensibly normalized such a comparison is valid. Voltages in Mach-Zehnder and directional coupler based modulators are normalized by $V\pi$ and V_s , respectively. Optical and r-f splits (or optical coupling ratios) are normalized so that unity represents a complete transfer. The electrical length of directional couplers is normalized to unity for a full transfer. This is unconventional; normally a full transfer is 0.5π . It was done for comparison to the other parameters that are all normalized to unity. Additionally, most of the modulators have multiple parameters. For simplicity, every parameter but the one being swept was held at its optimal value. In fact similar errors in

multiple parameters are found to compound the error, but only subtly. That is, if multiple parameters are swept at the same time, the resulting performance degradation is a little worse than if just one parameter is swept, but it is not substantial, and in the interest of keeping the dimensionality of the problem space low, only one parameter is swept at a time. The author looked at a number of cases, but did not do so exhaustively; this was enough to see that the interaction of errors in multiple bias parameters goes like the “sum of the squares.”

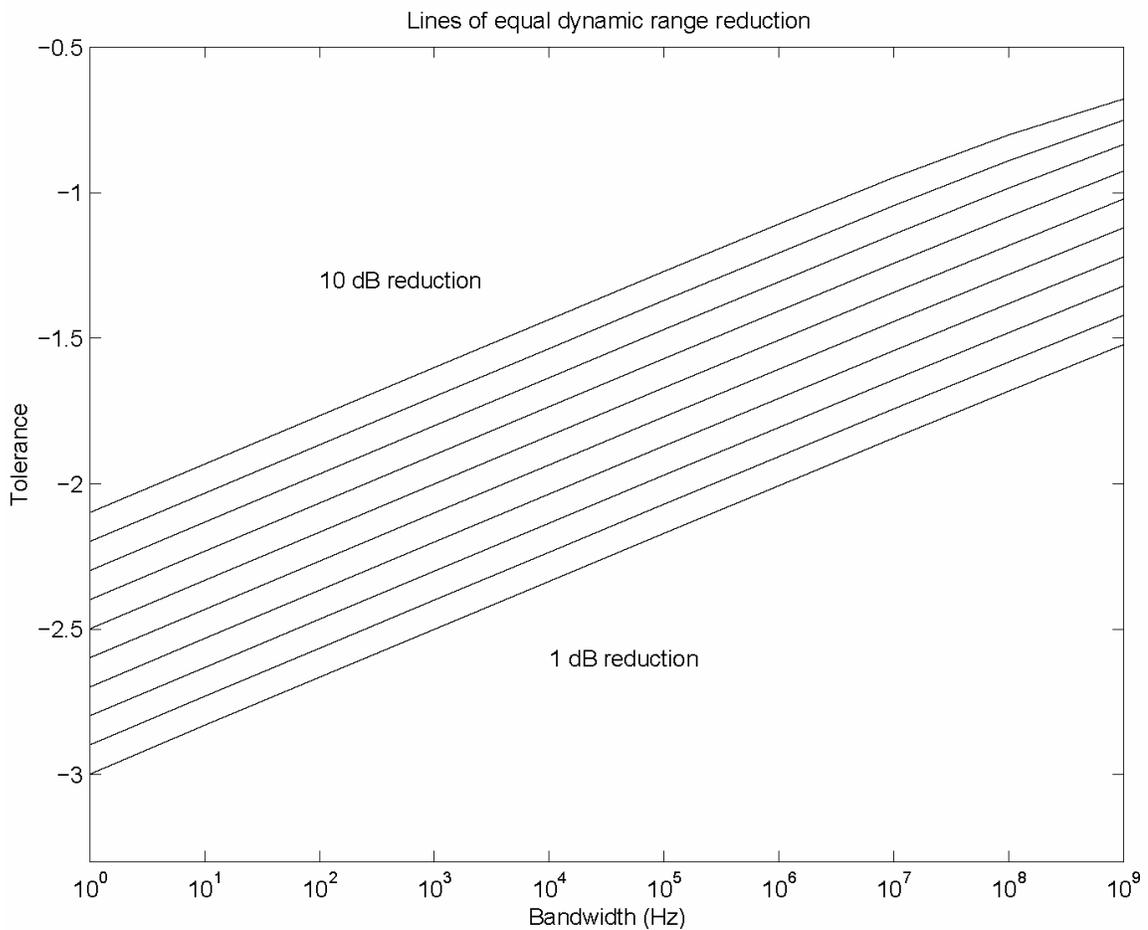


Figure 3 - 12: Lines of equal dynamic range reduction with bias error in a MZM.

Table 3-3: Modulator Parameter Sensitivity

Modulator	Parameter	Slope	Intercept	Intercept	Intercept
			I(1-dB)	I(3-dB)	I(10-dB)
Standard Modulators (no distortion cancellation)					
Simple MZ	V_{BIAS}	0.164	-2.99	-2.97	-2.08
Simple DCM	V_{BIAS}	0.164	-3.02	-2.72	-2.12
Sub-octave Modulators (third-order IMD cancellation, second harmonic out of band)					
SDCM	V_{BIAS}	0.199	-4.64	-4.20	-3.26
SDSMZ	V_{BIAS}	0.198	-4.28	-3.82	-2.86
Broadband Modulators (third-order IMD cancellation, no second harmonic)					
DPMZ	RF Split	0.209	-4.35	-3.89	-3.08
DPMZ	Optical Split	0.213	-4.53	-4.07	-3.02
DSMZ	RF Split	0.214	-3.98	-3.51	-2.57
DSMZ	Coupler Angle	0.209	-4.28	-3.84	-2.97
YFDCM	Electrical Length	0.201	-3.87	-3.41	-2.34
Broadband Modulators – DCM2P (third-order IMD & second harmonic cancellation)					
DCM2P	V_{BIAS-1}	0.282	-5.11	-4.85	-4.08
DCM2P	V_{BIAS-2}	0.299	-4.80	-4.60	-3.91

Calculations similar to those shown in Figure 3 - 10 and Figure 3 - 11 were performed for all eleven unique modulator parameters, and the data were reduced to create similar log-bandwidth-log-tolerance plots such as shown in Figure 3 - 12. Below a log-tolerance of about -0.5 all curves were linear, giving them a characteristic slope and 1 Hz intercept. Table 3-3 summarizes the dynamic range versus parameter errors for all modulator parameters. It shows the tolerances at the 1 dB, 3 dB, and 10 dB degradation levels. All of the log-bandwidth-log-tolerance curves have nearly identical slopes for a given modulator parameter. The values given in the slope column are an average of the slopes of the ten curves (1 dB to 10 dB). The 1 dB, 3 dB, and 10 dB columns contain the value of the log-tolerance at 1 Hz signal bandwidth only. Other noise bandwidths are calculated with the slope and the 1 Hz intercept. The tolerance for any signal bandwidth to experience an n dB dynamic range reduction (ignoring transit time and r-f losses) can be found for any parameter in the table with the relationship

$$\log(tol_{n_{dB}}) = slope \times \log(BW) + \frac{10-n}{9} I_{1_{dB}} + \frac{n-1}{9} I_{10_{dB}},$$

$$1 \leq n \leq 10,$$

$$BW \ll 10^{10} \text{ Hz.} \tag{3.22}$$

The table has three categories of modulator parameters, those in standard modulators, sub-octave linearized modulators and super-octave linearized modulators (except the DCM2P), and the DCM2P, broken out separately. Parameters of similar slope are grouped together. The standard MZM and DCM have the same slope and nearly the same intercepts. A 1 dB reduction of dynamic range at 1 Hz (109 dB absolute dynamic range) would require a normalized accuracy of 0.001, or for the canonical parameters ($V\pi$ or $V_s = 10$ V) about a 10 millivolt accuracy. A more realistic 1 MHz signal on the 3 dB line could withstand an 18% error, or a 180 millivolt error in the reference link. The sub-octave and super-octave modulators (except the DCM2P) require a (1 Hz, 1 dB) accuracy of 30 - 100 parts in a million, and a (1 MHz, 3 dB) accuracy of 1 - 6 parts in one thousand. The DCM2P has a questionably attainable (1 Hz, 1 dB) accuracy of 10 parts in one million, but it is only a little worse than the other linearized modulators at (1 MHz, 3 dB).

There are many different electrooptic intensity modulators, yet the tolerances of these parameters are remarkably similar. There are only three salient numbers in the table. They are the average slopes of categories one, two, and three. In fact, it makes little difference whether the modulator is a DCM or a MZM, whether it uses one technique for linearization or another; the distinction of the categories is more fundamental.¹⁵ It is only the type of distortion cancellation that determines the slope of the log-tolerance-log-bandwidth lines. In the first category (standard modulators), the second harmonic is nulled to the level of the uncanceled third-order distortion term. This yields a slope of 0.164. In the second category the third-order IMD is nulled and the second harmonic plays no part in the linearizing mechanism. The second harmonic is out of band in the sub-octave modulators and in the DPMZ, DSMZ, and YFDCM, the harmonic cancellation occurs by a different and more robust mechanism than the third-order cancellation, e.g., device symmetry. This yields a slope of about 0.21. In the DCM2P the same linearization mechanism nulls both the second harmonic and the third-order distortion. The critical sensitivity of the DCM2P results from using the sensitive linearization technique to null the second harmonic distortion to the level of the nulled third-order IMD. There is about 50 dB of second harmonic distortion suppression in the DCM2P at a 1 Hz bandwidth. Thus in comparison to the SoDCM of Figure 3 - 11, the DCM2P is on the 4/5 slope scaling curve, but when there is an error in the bias it does not drop down to the 2/3 slope curve, but instead it falls all the way to the 1/2 slope curve. This effect is unique to the DCM2P of the modulators covered in this study.

Important engineering lessons can be drawn from the analysis of ideal performance and parameter tolerance. Table 3-3 shows that in dynamic range performance there are basically two types of modulators, standard and third-order linearized modulators. Among the linearized modulators the only noticeable performance difference comes from the average photocurrent level of the modulator. This gives a couple of extra dB of dynamic range to the sub-octave modulators. Table 3-3 also shows that the parameter sensitivity of these two types (with the exception of the DCM2P at low signal bandwidths) is nearly identical over all the parameters within each type. This study covers eight different modulators, and if some subtle

¹⁵ The different techniques and types of modulators may have significantly different weights in terms of manufacturability; this point refers only to the idealized mathematical models.

permutations were included, it could number above 20. Yet, the calculated ideal performance and parameter sensitivity is incredibly similar across all modulators within each class. The conclusion is that modulators should be chosen by their manufacturability and reliability.

Multi-Tone Analysis

The effects and consequences of more than two simultaneous signals being applied to a non-linear system are substantially more complex and require a different evaluation scheme than simply “dynamic range.” There are many ways to have “more than two signals.” In this dissertation, the special case of N signals, equally spaced and of equal amplitude, is treated since this situation is encountered in cable television (CATV), as well as other sub-carrier multiplexed systems.

As stated briefly in the section “Analysis with One and Two Tones,” if N ($N > 2$) equally spaced, equal amplitude tones are applied to a non-linear system, then, unlike the one- and two-tone analysis, the relative phases of the tones is important. If all of the sinusoids are in phase at one point in time, then they add to N times the amplitude. As time progresses, they run out of phase, and the amplitude decreases. However, they will again add to N times the amplitude at a time $T = 1/F_D$ later, where F_D is the frequency interval between tones. The result is a train of pulses with amplitude N and width $\sim 1/N$. As $N \rightarrow \infty$ the shape of the wave approaches a delta function. Even for finite N (consider the carriers of 100 TV channels on a typical cable TV system) the peak amplitude is quite high compared to that estimated from the total average power. Clearly, this is an undesirable situation for transmission through a non-linear system.

Two other cases are common in the literature. According to Ref. [3.18], and subject to their assumptions, the best case performance for sub-carrier systems is given by distributing the phases $\varphi_k = (\pi/N)k^2$ for $k = 1, 2, \dots, N$. This is an extension of the Newman phases. In practical sub-carrier multiplexed communication systems, the phases are neither correlated in the best case, nor correlated in the worst case. The practical case is where the N tones have a random distribution of phases, and the amplitude can be treated statistically.

Three or more tones create multiple noise terms at each frequency location, and create a type of intermodulation distortion, previously untreated in this work, the so-called triple beats. In two-tone analysis, intermodulation distortion arises from one tone beating with the second harmonic of the other tone, $f_2 - 2f_1$, occurring at $f_2 + \Delta f$ and $f_1 - 2f_2$ occurring at $f_1 - \Delta f$ assuming the tones are spaced by Δf . There are only two distortion terms. In multi-tone analysis there are many more distortion terms, some of which are triple beats. Some of these terms appear at the frequency of the fundamental tones. For example, the triple beat, $f_1 + f_2 - f_3$ occurs at f_2 , as does the two-frequency intermodulation term $2f_3 - f_4$.¹⁶ A thorough treatment of this phenomenon is given in Ref. [3.19], from which the definitions in this study are taken. The number of distortion terms at each frequency grows dramatically with the number of tones in the sub-carrier multiplexed system. The single frequency with the most distortion terms is the center frequency of all the tones.

With this proliferation of distortion terms, it becomes advantageous to construct a single composite term from multiple terms that fall on one frequency, as in ref [3.19]. The **composite triple beat** (CTB) is the sum of the powers of all intermodulation distortion terms at any given frequency. The product of the *equivalent* triple beat with the power of a single triple beat gives the analytic value of the CTB assuming the phases of the tones are randomly distributed. The equivalent triple beat is the total number of individual triple beats plus one-fourth the total number of two-frequency intermodulation beats, since the latter are found to be one-fourth the strength of the former, also from ref [3.18]. The **composite second-order** (CSO) is the sum of the powers of all the second harmonic distortion terms. For example, $(2f_2)$ and $(f_1 + f_3)$ are both second order distortion for the CSO at f_4 . The analytical method of calculation is the same as that of the CTB except that all terms are equally weighted.

In this work, the numerical calculation of the composite terms is similar to the method of experimental measurement. In the laboratory, a series of N frequency synthesizers are incoherently combined, fed into a non-linear transducer, and the transducer output is analyzed on a spectrum analyzer. The synthesizer corresponding to the frequency of interest is turned

off. The distortion suppression (CTB + CSO) is the difference between the power levels at the off-signal frequency and an adjacent channel. There are two implicit assumptions: first, all on-tones have the same power level, and second, the on-signal is sufficiently larger than its own composite distortion, that the presence of this distortion has no effect on the measurement. The numerical calculation takes the exact same approach. Since this method obtains the total composite distortion at any frequency, a distinction between CSO and CTB cannot be made, without carefully choosing the location of the signal band in the simulation.

Triple Beat Calculations for Electrooptic Modulators

Figure 3 - 13 shows the time domain of the samples used to calculate the composite triple beat for the MZ and DPMZ modulators. Thirty tones of equal amplitude and equal frequency spacing are chosen. The power level was selected to match the desired operating range to optimize the composite triple beat suppression, under the Newman phase distribution. The correlated phases are also shown but at the same power level (if the phases were really correlated, then the desired operating level would be about 7.5 dB lower). This was chosen to illustrate the pulse train that occurs with correlated phases. An arbitrary number of samples was chosen (1024), and the samples cover one complete period. Notice that for the correlated phases, at the beginning and the end of the sample train, there is a peak 3 - 4 times greater than the peak amplitude in the Newman phase case, but that for most of the time, the correlated phases has a lower peak than the Newman phases. It is desirable to minimize the maximum signal excursion from the bias point. Only deterministic pulse trains are analyzed. It is assumed that an uncorrelated pulse train would appear slightly better than the Newman condition for most, but not all, random seeds!

Figure 3 - 14 shows the power sweep of a standard Mach-Zehnder modulator with the canonical link parameters from Table 3-1. In this case the noise bandwidth was taken to be 6 MHz. The assumption is that each channel is spaced by 6 MHz and filtered to its spacing. And that the noise bandwidth from the other 29 channels does not contribute. This gives a

¹⁶ Here “two-tones” refers to the fact that this distortion term arises from only two of the N tones. Intermodulation distortion of the type $(2f_1 \pm f_2)$ will be referred to as two-frequency, while intermodulation of the form $(f_1 \pm f_2 \pm f_3)$ will be referred to as triple beat.

noise level of -93.4 dBm. The signal from the two-tone test is also shown, that is slope 1. For the input power levels of interest, the signal strength for all cases (two-tone, Newman phases CTB and correlated phases CTB) have the same signal level. When the correlated phases CTB is at the same level of the signal this assumption breaks down. The three distortion terms have the same slope since they are all based on third-order products. The CTB crosses the noise level at a lower power level.

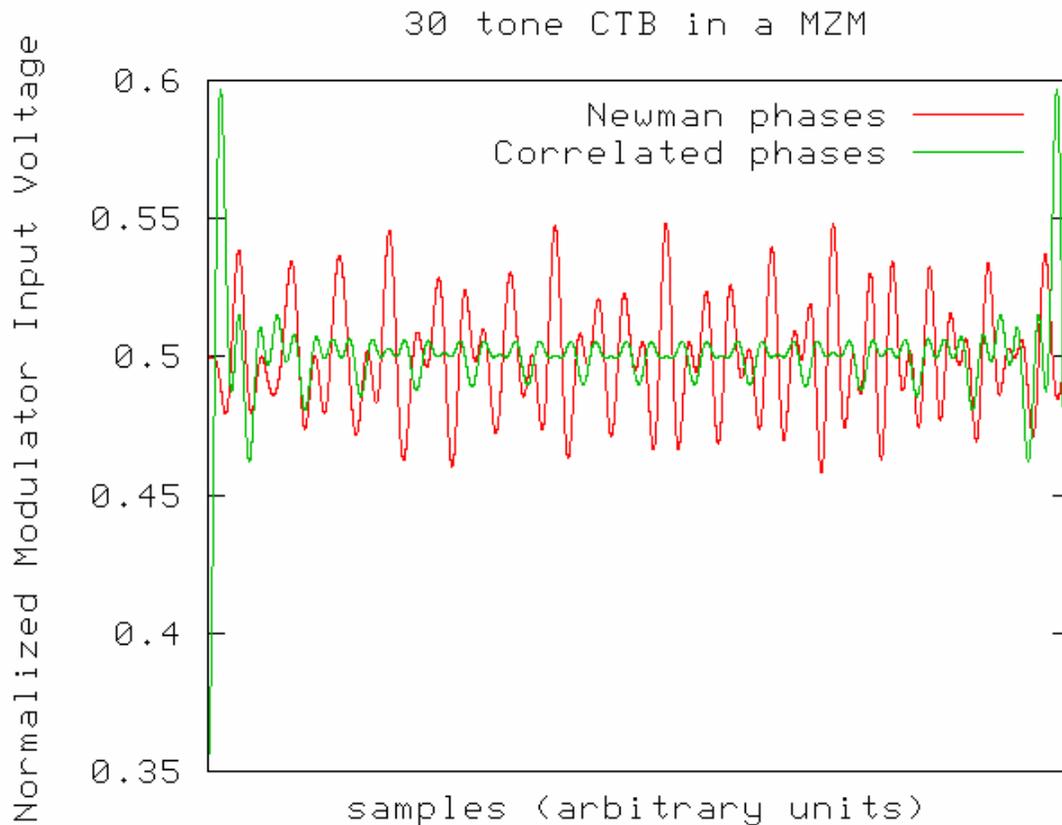


Figure 3 - 13: Time-domain samples of 30 tones in a Mach-Zehnder modulator; when the phases are correlated, a pulse is formed, when the phases are related by the Newman condition, the peaks are nearly optimally averaged.

CATV channels generally need 52 dB of suppression, see Ref. [3.20]. The two-tone dynamic range is 64.7 dB, nearly 13 dB more than required. However, the Newman CTB dynamic

range is 53.7 dB, and the correlated CTB dynamic range is 46.7 dB. Thus this simple link model is barely good enough for 30 channels. Note that modern CATV experiments that use a simple MZM external modulator may have better link parameters than are assumed in Table 3-1.

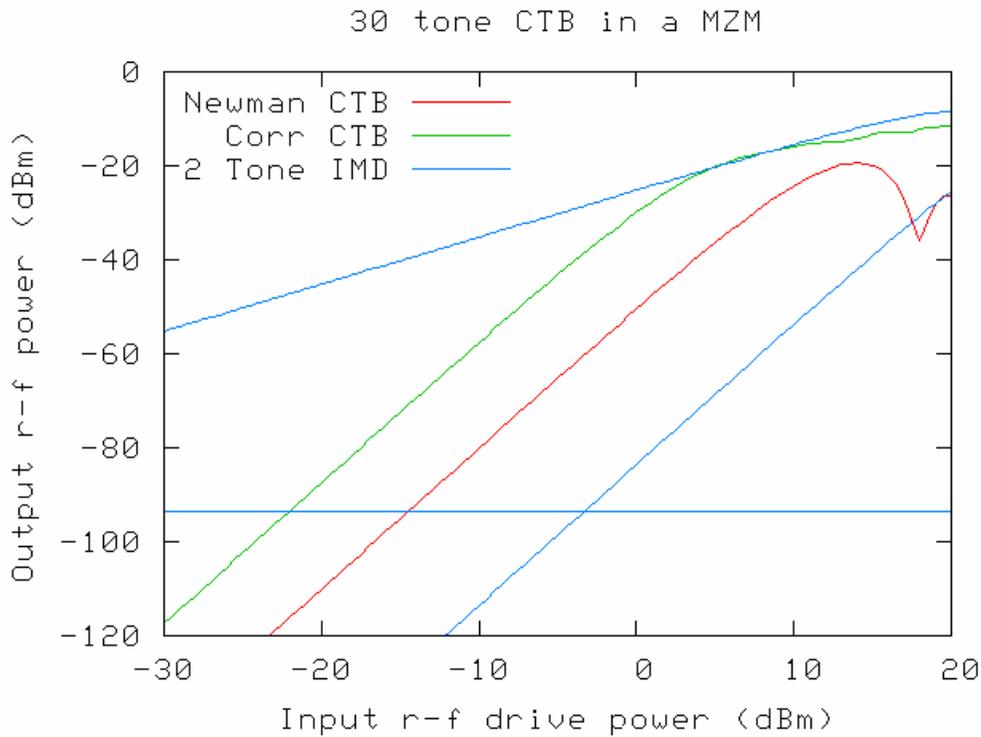


Figure 3 - 14: Power sweeps of 30 tones in a MZM. The two-tone results are shown for comparison, including two-tone SFDR, noise level, and signal level. The Newman CTB dynamic range is about 11 dB worse than the two-tone SFDR, and the correlated CTB is 7.5 dB worse still.

Figure 3 - 15 shows the same 30 channel measurement but with a linearized modulator: the DPMZM. The DPMZM successfully nulls both the Newman phase CTB and correlated phase CTB. The two-tone dynamic range is 75.4 dB at the 6 MHz noise level. The Newman CTB dynamic range is 62.4 dB and the correlated phase CTB is 54 dB. Of course, the correlated phase case is not generally analyzed, because it is so bad, and, if designed for,

systems are generally highly overbuilt. An interesting negative effect is that the difference between the two-tone case and the Newman phase dynamic range is 13 dB, while it was only 11 dB for the Mach-Zehnder modulator. There is not an obvious answer for this, though it is clear from the plot that the shape of the two curves does not match as closely as it did in the standard Mach-Zehnder modulator plot.

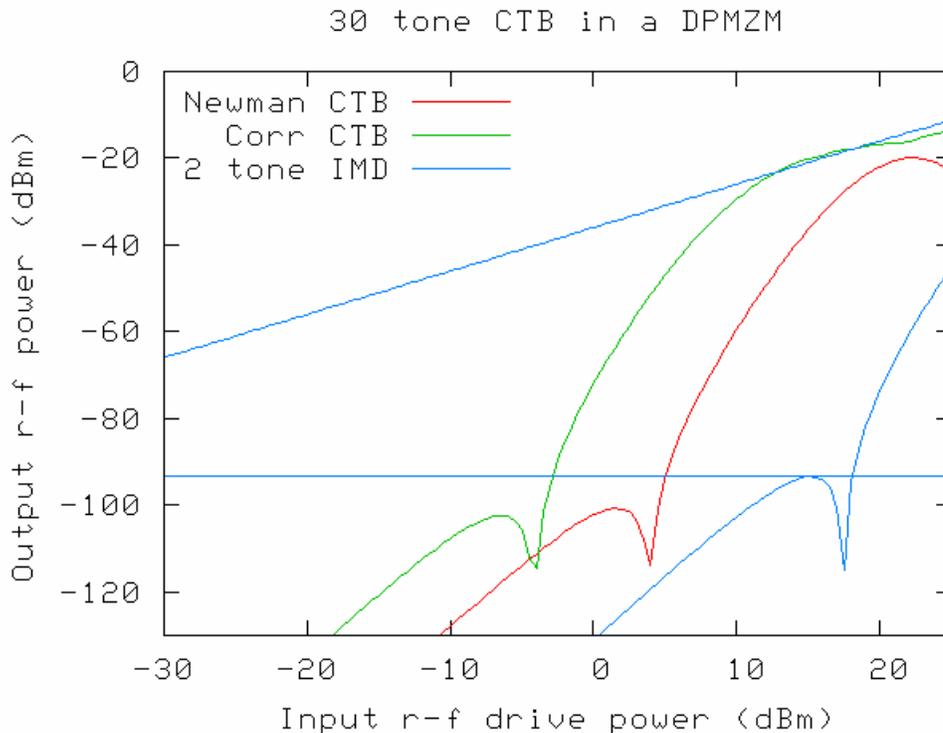


Figure 3 - 15: 30 tone CTB in a linearized modulator.
The modulator choice is the DPMZM.

This leads to the basic question: Is 9 dB worth it? Calculations were run for 90, 120, 150, 180, 210, 240, 270, and 300 channels with the DPMZ (not plotted). The 300 channel calculations give 54 dB dynamic range for the Newman phase relationship. The increase in the number of channels leads to an equivalent reduction in the dynamic range, in this range of operation. Therefore, linearized modulators offer great potential for increasing the bandwidth of CATV, or other sub-carrier systems.

The author speculates that a very high bandwidth system could be built using sub-octave modulators. Recall from noise scaling that as the system bandwidth increases, the ability to reduce shot noise becomes a more significant contribution to the reduction in dynamic range. Many sub-carrier systems are in the 1 - 10 MHz per channel range. At this noise bandwidth, a well-optimized sub-octave modulator would yield a significantly higher dynamic range. However the trick would be in building a sub-carrier system that is sub-octave. For example consider a system of 1000 5 MHz channels in which the lowest channel starts at 5 GHz. The author did not get to an in-depth analysis of the sub-octave modulators in sub-carrier systems. It is left as a subject for future work.

BANDWIDTH OF LINEARIZED ELECTROOPTIC MODULATORS

Abstract

Many schemes have been proposed to make high dynamic range analog r-f photonic links by linearizing the transfer function of the link's modulator, as described in the previous chapter. This chapter studies the degrading effects of finite transit time and optical and electrical velocity dispersion on such linearization schemes. The majority of the linearization techniques, but not all, experience a reduction in dynamic range with increasing frequency far more severe than the corresponding reduction in gain. However, much of the lost dynamic range in some of these modulators may be regained by segmenting and rephasing the r-f transmission line. Additionally, this chapter addresses some novel properties of the gain function of directional coupler-based modulators.

Finite Transit Time Computation

A major aspect of this thesis is the development and application of a general computational tool for the analysis of all electrooptic modulators. This tool was designed to cover the basic building blocks of DCM and MZMs, and linearized modulators composed of multiple modulator blocks in various configurations. It models arbitrary, but functionally defined, r-f inputs. It is used for frequency-dependent analysis of all modulation figures of merit.

Background

The general problem of the frequency-dependent transfer function of a directional coupler modulator has no analytic solution. The underlying mathematical complexity comes from wave dispersion and finite transit time. The electrical wave, which creates a modulation envelope in the optical wave, travels at a different velocity than the optical wave, and the coupling interaction of the optical waves in the modulator may be distributed over a distance of many wavelengths of the modulation frequency in high frequency applications. The Mach-Zehnder is simpler (and an analytic representation exists) because the interaction between the electrical and optical wave accumulates a phase change over the modulator length, and then the interaction between the two optical waves occurs effectively instantly as the optical waves are interfered in a Y-junction. There have been some attempts at approximating the solution to the directional coupler, see for example Ref. [3.10]. However, it is also possible to solve this problem numerically as shown in Ref. [3.11].

At high frequency, this problem is solved numerically without ever developing an analytic transfer function that is accurate at the frequency of interest. A modulator may be represented as a cascade of short modulator sections where the output of the n -th modulator section is the input of the $(n+1)$ -th modulator section. The applied voltage to the $(n+1)$ -th modulator section is the voltage that was applied to the previous section, but at a time that is adjusted for the difference in velocity between the optical and electrical waves, $\Delta t = \Delta x / (v_O - v_M)$. The overall output is found by matrix multiplication of all of the transfer function sections, each of which has a corrected optical input and electrical input from the previous stage. By dividing the modulator into enough sections, an accurate output can always be calculated. Each section is sufficiently short that the interaction is small compared to a modulation wavelength, and

thus the transfer function section approximates the low frequency or d-c behavior that is known.

Farwell used this method to investigate the frequency response of the distortion terms in the DCM2P and a Quad Mach-Zehnder modulator, designed to null the third-order and fifth-order IMD. Following Farwell's solution scheme, a computer program in the C language was written and used to investigate several linearized modulators. Though the basic computation scheme is the same as Farwell's, this analysis investigates many more modulators, both super-octave and sub-octave, and many other physical effects than Ref. [4.1], such as larger signal bandwidths, electrode re-phasing, band-pass dynamic range re-optimization, and band-pass high frequency gain of the DCM and YFDCM. Some of the results presented here were published by Cummings and Bridges in Ref. [4.2] in 1998. Additional results developed since then are also presented here. It has been found that some (but not all) linearized modulators display a frequency limitation that is much more severe than what would normally be associated with the frequency dependence of the gain of the modulator.

Wave Velocity Mismatch

In the program, it is assumed that there is traveling wave propagation along a microwave transmission line with velocity c/n_m , and this wave modulates an optical signal in the optical waveguide that is traveling at c/n_o . From the perspective of the microwave wave front, it appears as if the optical wave is moving relatively at a velocity of $c/(n_m - n_o)$. If there were no difference in wave velocity, it would appear to the microwave wave front as if the optical wave were stationary, and then the microwave would have the entire interaction distance to build up the modulation. However since the two waves travel at different velocities, over a long enough interaction length (relative to the modulation wavelength), the microwave modulates the optical wave, and then unmodulates it, and then remodulates it, etc. In phase modulators and MZMs, the total accumulated phase change along this interaction length can be found with a simple integral, despite the wave dispersion. But in directional couplers, where the optical wave is also coupling into an adjacent waveguide along this interaction length, no exact analytical solution exists to describe the propagation of the two optical waves and the r-f wave.

A modulator transfer function is represented by a matrix with terms that describe the output the complex optical amplitude as a function of all of the input optical amplitudes of that waveguide section. Complicated modulators may be represented by a matrix cascade for the series sections. Additional modulator functions such as Y-branches, phase shifts, and combiners all have simple matrix representations.

Mathematical description of the modulator analysis program

Let $A_{IN}(t)$ and $A_{OUT}(t)$ be the input and output complex vectors of the optical wave in the overall modulator. For a system with K parallel waveguides, $A_{IN}(t)$ and $A_{OUT}(t)$ are $K \times 1$ vectors. However, here, all of the modulators analyzed are at most two waveguide systems so $A_{IN}(t)$ and $A_{OUT}(t)$ are 2×1 vectors. Let $H(V, \ell)$ be the overall transfer function matrix, where V is the normalized modulator drive voltage and ℓ is the active length of the overall modulator. If the optical and electrical waves travel at the same velocity, or if the operating frequency is so low that V is effectively constant over ℓ , then the output after a transit time of τ is given by

$$A_{OUT}(t) = H[V(t - \tau), \ell] \times A_{IN}(t - \tau). \quad (3.23)$$

Even if there is a significant difference in velocity of the two waves as they travel the distance ℓ , over a short enough section of the guide Δx , the change in the complex optical amplitudes may be described with the d-c transfer function. This is the basis for the frequency-dependent calculation. Let x be the coordinate along the optical waveguide. Then for a short section

$$\lim_{\Delta x \rightarrow 0} A(x - \Delta x, t - \Delta t) \times H[V(x - \Delta x, t - \Delta t), \Delta x] \rightarrow A(x, t). \quad (3.24)$$

The optical and electrical signals are now functions of two variables, x and t . In Eqn. (3.24) the elapsed time equals the incremental length divided by the velocity along the electrical transmission electrodes, $\Delta t = \Delta x / v$.

Let there be N equally spaced increments of x , and M equally spaced increments of t , and let $x_0 = 0$ and $x_N = \ell$. The modulator is divided into N sections over which each of the optical and electrical fields are approximately constant. The finite product of the resulting N unitary d-c transfer matrices gives the overall transfer function matrix from $A_{IN}(t)$ to $A_{OUT}(t)$. Let $A(x_0, t)$

$= A_{IN}(t)$, and $A(x_N, t) = A_{OUT}(t)$, and note that $A_{OUT}(t) = 0$ for $t < x_N/v$. The approximate modulator transfer function is

$$A_{OUT}(t_i) \approx \left\{ \prod_{k=0}^{N-1} H \left[V \left(x_k, t_i - \frac{x_{N-1-k}}{v} \right), \frac{l}{N} \right] \right\} \times A_{IN}(t_{i-M}). \quad (3.25)$$

The function $V(x, t)$ representing the two tone test is now

$$V(x_k, t_i) = p_{IN} \left[\cos \left(\frac{\gamma}{f_1} \left(\frac{i}{M} - \frac{k}{N} \right) \right) + \cos \left(\frac{\gamma}{f_2} \left(\frac{i}{M} - \frac{k}{N} \right) \right) \right]. \quad (3.26)$$

The parameter γ depends on the operating frequency f_0 , the difference between the effective optical and electrical indices Δn , and the active length ℓ . It is important to note that the calculated results depend solely on γ and not on Δn , ℓ , and f_0 independently. Thus the results of different lengths or relative wave velocities at different frequencies will be the same if γ is the same. The results are universal in the sense that they apply to more than the ‘‘worst case’’ velocity mismatch, which is $\Delta n = 1.8$ for lithium niobate modulators, with simple coplanar strip line transmission lines. Any change in Δn , f_0 , or ℓ can be accounted for by a simple rescaling of the frequency axis for the gain and dynamic range plots.

Eqn. (3.25) is general and is the basis for the frequency dependent computations in the model. However, when the active region of a modulator consists of only simple phase shifts, as it does for the Mach-Zehnder modulator, a further simplification may be made. The transfer function is just a diagonal matrix of exponentials. Instead of multiplying exponentials, their arguments are summed. This is equivalent to integrating the location variable out of the voltage function. That is

$$A_{OUT}(t) = H \left[\int_0^l V \left(x, t - \frac{l-x}{v} \right) dx, l \right] \times A_{IN} \left(t - \frac{l}{v} \right), \quad (3.27)$$

$$A_{OUT}(t_j) \approx H \left[\sum_0^{N-1} V \left(x_k, t_i - \frac{x_{N-1-k}}{v} \right), l \right] \times A_{IN}(t_{i-M}). \quad (3.28)$$

The approximation introduced in Eqn. (3.28) is a substitution of a summation for an integral; the relationship between A_{OUT} and H shown in Eqn. (3.27) is exact. It may seem odd to use an approximation for a function for which a trivial analytic solution exists, that is the integral of Eqn. (3.26). However, this is done to mirror the calculation technique for directional couplers and to support the modeling of modulator voltage functions which may not have an analytic integral representation.

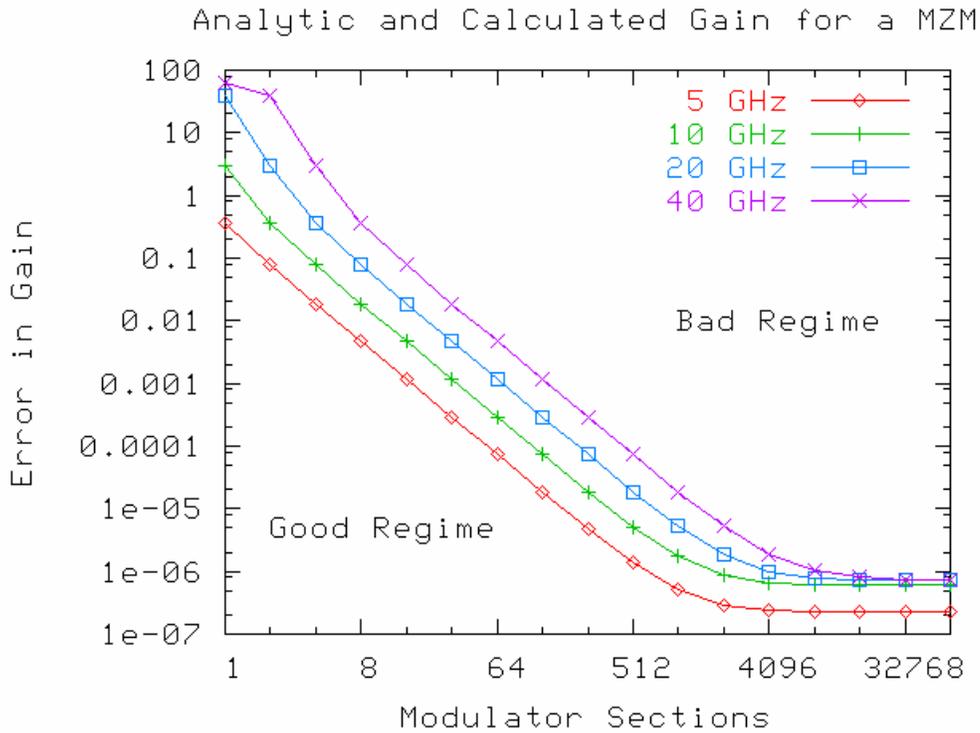


Figure 4 - 1: The convergence of the calculation of the gain of a simple MZM to its analytic value with an increasing number of modulator sections. The error is the magnitude of the calculated gain (not in dB) minus the analytical value normalized by the d-c analytical value for the gain. The curves of the log-log plot are for 5, 10, 20, and 40 GHz.

In the C program, the number of temporal increments M are restricted to powers of 2, so that a radix-2 FFT algorithm may be used for the spectral analysis.¹⁷ Figure 4 - 1 shows the convergence of the gain of a simple Mach-Zehnder Modulator compared to the analytic value as a function of the number of spatial sections N at 5, 10, 20, and 40 GHz. The error is normalized, and thus the y-axis value “one” corresponds to a 3 dB error and “0.1” corresponds to a 0.4 dB error in the calculated gain. The curves in Figure 4 - 1 are linear until the error is 6 - 7 orders of magnitude below the d-c value of the gain, that is adding modulator sections linearly reduces the error until the convergence saturates at a normalized error of about 10^{-6} . The r-f drive power for the small signal gain calculation was chosen to be -100 dBm. Whether in the calculation for gain or dynamic range, the numbers have components that differ by 6 - 7 orders of magnitude. Since these components occupy the same mantissa, there is a loss in accuracy not recovered by the floating decimal point. Double precision numbers must be used to attain a satisfactory accuracy.

It is interesting to note that the calculated points form a horizontal line across the curves in the linear regime. This indicates that doubling γ (by doubling the frequency for instance) exactly requires a doubling of the modulator sections to achieve the same error. Efficient code allows the calculation hundreds of frequency points with a 128-point FFT and a comparable number of spatial increments in seconds on a 100 MHz x86 desk-top machine (as used when this research was begun).

Non-Linearized Modulators: The Standard MZM, DCM, and Y-Fed DCM

The most common electrooptic modulator is the simple Mach-Zehnder interferometer modulator configuration (MZM) as described in Chapter 2. Its analytic solution is a cosine-squared transfer function, and the gain is a sinc function of the frequency-length-index product γ . When biased at the half-wave voltage, $V=0.5 V\pi$, it attains its maximum linearity

¹⁷ It is customary to use the FFT algorithm which is $O(N\log(N))$, instead of the DFT algorithm, which is $O(N^2)$. However, it should be noted that the algorithm to compute the modulator output is $O(N^2)$, so the time spent in the FFT algorithm is inconsequential if the number of spatial increments is similar to the number of temporal increments.

and dynamic range. All even-order harmonics are identically zero, so the intermodulation distortion products solely determine the dynamic range, even in a super-octave system. The signal decays with frequency, but the intermodulation product decays identically, so that the dynamic range is independent of frequency. The range of r-f drive power (in dB), over which there are no spurs above the noise, shifts with a change in frequency, but it does not expand or contract. Given its analytical simplicity and widespread use, the MZM is used first to evaluate the accuracy of the numerical calculation, and then for a comparison with the results for linearized modulators.

Figure 4 - 2 shows the calculations for gain and dynamic range of an analog r-f link using a simple MZM with traveling wave electrodes and the link parameters given in Table 3.1. The length $\ell = 10\text{mm}$ and the maximum velocity mismatch represented by $\Delta n = 1.8$ (non-velocity matched lithium niobate) are used. The gain has the form $\sin(\pi\gamma)^2/(\pi\gamma)^2$ with zeros at multiples of 16.2 GHz (where $\gamma = 1$ for the canonical parameters from Table 3.1), and a low frequency link gain of -25.5 dB (also appropriate for the link parameters). The dynamic range is flat across frequency because the signal and the intermod experience the same roll-off as a function of frequency. The null in the second harmonic is not disturbed by the velocity mismatch.

Figure 4 - 3 shows the analogous calculations for a simple directional coupler modulator with the optical input fed into one arm, and a bias voltage of $0.43 V_S$. The low frequency gain is -24.8 dB, 0.7 dB better than Mach-Zehnder.¹⁸ The first null of the gain of the DCM occurs at 26 GHz compared to 16.2 GHz for the MZM, even though both modulators have the same frequency-length-mismatch product γ . The first lobe of the gain curve does not correspond to the sinc function of the Mach-Zehnder. However, subsequent nulls of the DCM are periodic with a 16.2 GHz period, just as in the MZM, resulting from the underlying γ of the DCM.

¹⁸ The comparison of the gain bandwidth product between the MZM and the DCM assumes that $V\pi = V_S$; that is, the modulators have the same normalization “on-off” voltages. While these voltages should be similar in the same manufacturing process, they should not be identical because the electrode geometries and crystal orientations of the two modulator types differ.

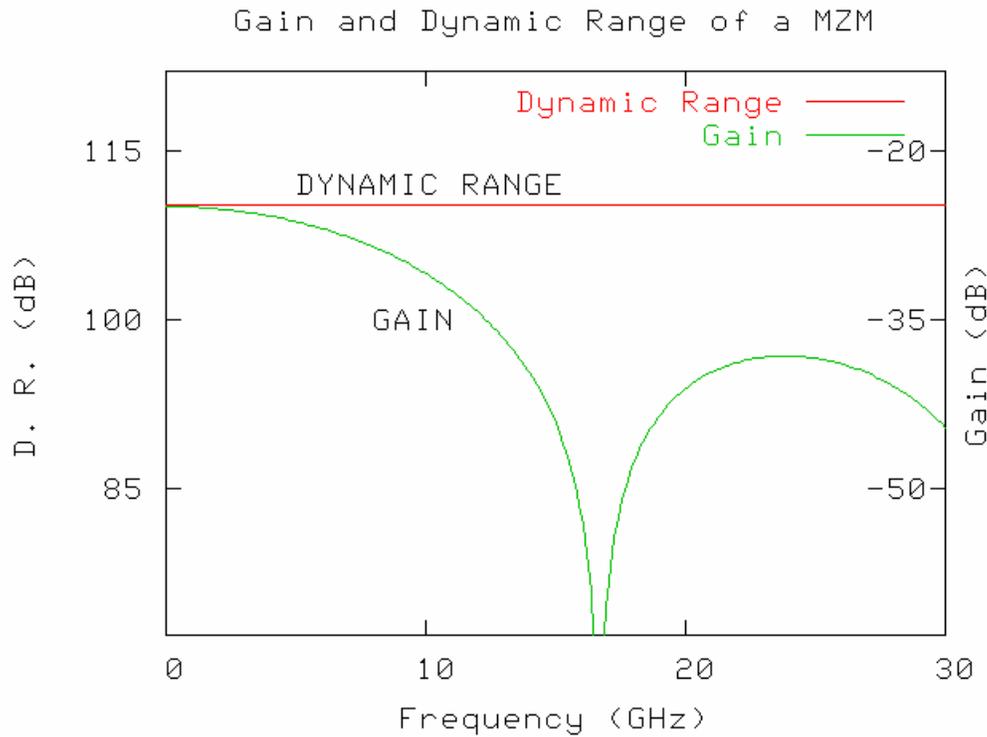


Figure 4 - 2: Gain and dynamic range of a standard Mach-Zehnder modulator from d-c to 30 GHz using the canonical parameters of Table 3.1. The dynamic range uses the left hand vertical axis and the gain uses the right hand vertical axis.

The frequency at which the gain has fallen by 3 dB is 40% higher than that of the Mach-Zehnder modulator with the same index-length product, γ . The dynamic range of the DCM compares unfavorably to that of the MZM at higher frequencies, which comes from the appearance of the second harmonic. At low frequencies it is similar to that of the MZM, and it is approximately flat with frequency. However, there is a kink in the curve at 1.8 GHz (left vertical arrow), after which the dynamic range decays rapidly with frequency. Unlike the MZM, where all even-order derivatives of the transfer function are identically zero when the modulator is biased at $0.5 V\pi$, in the directional coupler only the second derivative is zero at its optimal bias, $0.43 V_s$. Distortion migrates from non-zero fourth, sixth, eighth, ... order derivatives to the frequency at which the second harmonic occurs. It does so rapidly with increasing modulation frequency, and it equals the third-order intermodulation product at 1.8

GHz. At higher frequencies, distortion from these non-zero even derivatives limits the dynamic range. The dynamic range that would result from third-order intermodulation distortion with the second harmonic ignored is also shown with the label “D. R. (Intermod).” $0.43 V_S$ is not the correct bias point to optimize this modulator in a sub-octave system. However, the dynamic range based on the intermod only is shown here to illustrate the effect of the second harmonic on the DCM.

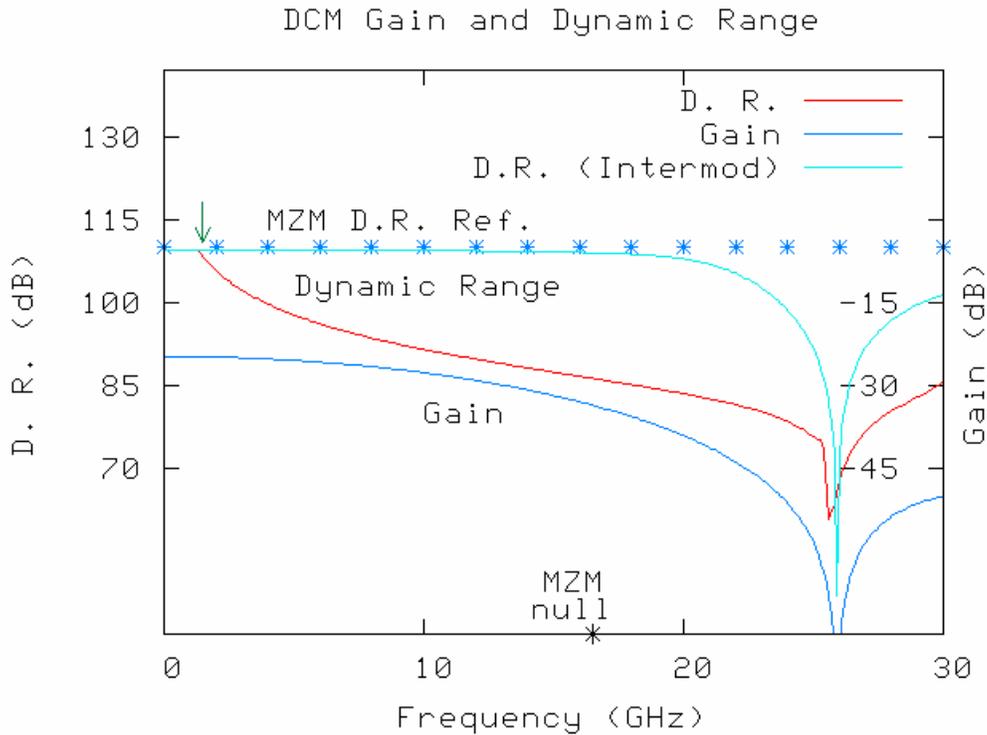


Figure 4 - 3: Gain and dynamic range of a simple directional coupler modulator with traveling wave electrodes biased at the $0.43 V_S$. “D.R.” is the dynamic range, which is second harmonic limited. “D.R. (Intermod)” would be the D.R. if there were no second harmonic.

Y-Fed Directional Coupler Modulator

The dynamic range of a standard YFDCM is superior to that of the DCM. This modulator does not have third-order suppression because the electrical length is 0.5π ¹⁹ instead of 1.43π

¹⁹ In Ref. [3.9] Thaniyavarn proposed an electrical length of $(1/\sqrt{2})\pi$. This is the correct electrical length for a complete transfer for a YFDCM (necessary for a 1x2 switch). The plot uses 0.5π for a consistent comparison to the DCM and MZM.

for the linearized YFDCM. Unlike the ordinary DCM it is always limited by third-order intermodulation. The transfer function is anti-symmetric around the d-c operating point, just like the MZM. There is a slight roll-off in the dynamic range frequency response, but compared to the MZM the roll-off is not appreciable until the MZM is already close to its gain null. From these results, the best non-linearized modulator for super-octave applications is a Y-Fed DCM with electrical length 0.5π . The transfer function of a Y-fed DCM is anti-symmetric around the $V_b = 0.0$ operating point for any electrical length. The anti-symmetry property makes all higher even-order harmonics zero, just as in the case of the MZM biased at the quadrature point. In this instance the different coupling mechanisms in the MZM and standard YFDCM makes less of a difference than distinction between the DCM and the standard YFDCM, namely, that only the second harmonic is nulled in the DCM at the super-octave operating point, $V = 0.43 V_s$.

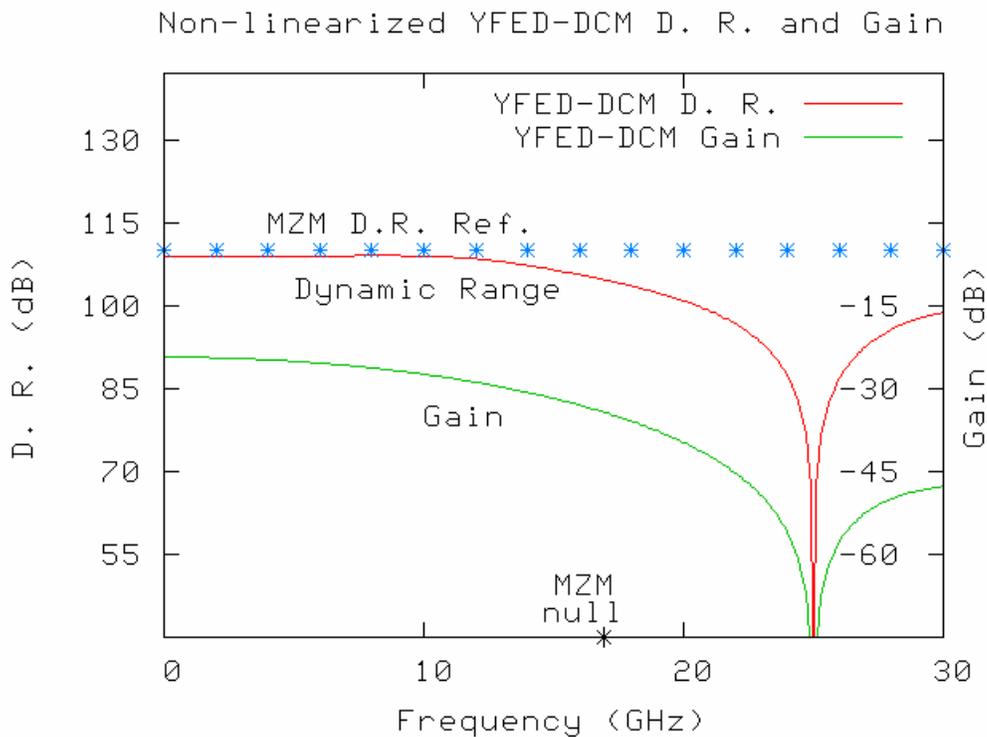


Figure 4 - 4: Non-linearized Y-fed DCM gain and dynamic range with the same modulation parameters as Figure 4 - 2 and Figure 4 - 3.

Broadband Linearized Modulators

The frequency dependence of the dynamic range was calculated for the four linearized super-octave, modulators: the dual parallel Mach-Zehnder (DPMZM), the directional coupler with two passive sections (DCM2P), the cascade Mach-Zehnder modulator (CMZM), and the Y-fed directional coupler modulator with electrical length 1.43π (YFDCM). Figure 4 - 5 shows the dynamic range as a function of frequency for 0 to 8 GHz for these four modulators with the canonical link parameters used in Chapter 3. Figure 4 - 6 is the same calculation repeated to show more detail in the range 0 to 250 MHz where the dynamic range changes dramatically for three of the four.

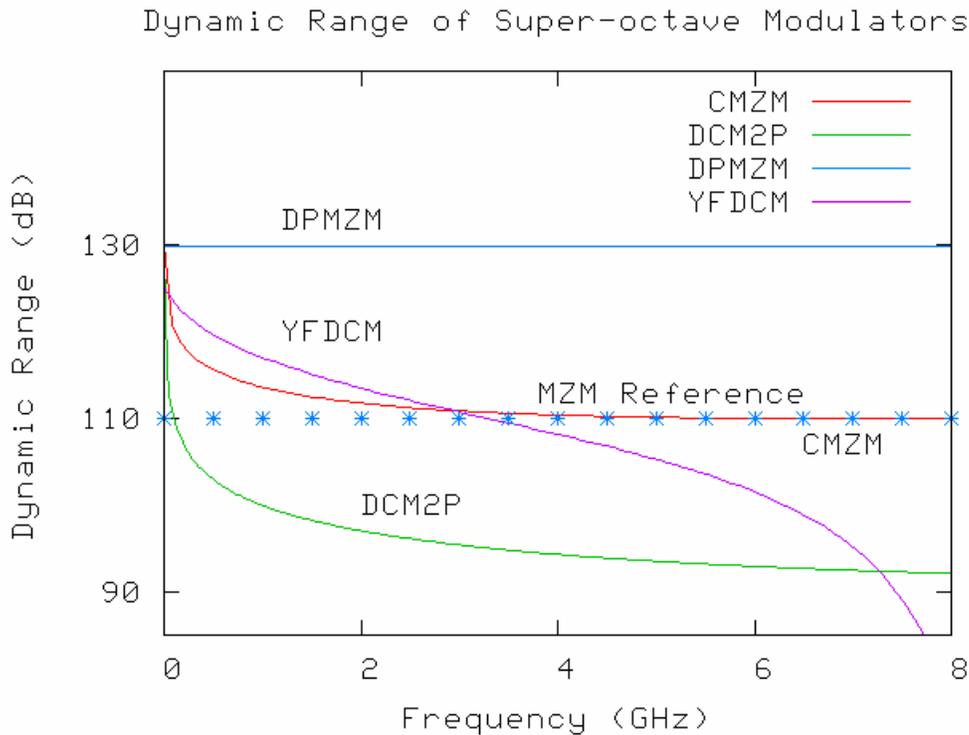


Figure 4 - 5: Dynamic range comparison of four broadband linearized modulators: DPMZ, DCM2P, DSMZM, and YFDCM-1.43. The DPMZ is independent of frequency, the DSMZM and YFDCM are moderately susceptible to frequency, and the DCM2P's response drops precipitously with frequency.

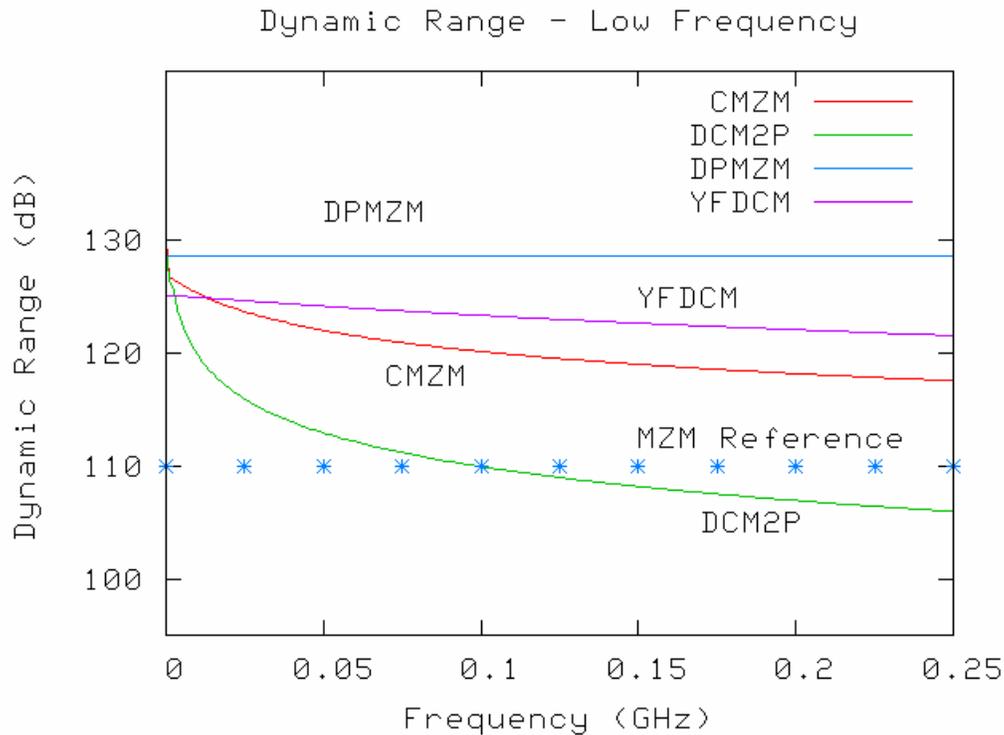


Figure 4 - 6: Dynamic range comparison of four broadband linearized modulators: DPMZM, DCM2P, CMZM, and YFDCM, shown here from 0 to 0.25 GHz.

Frequency Independence of the DPMZ

The DPMZM has two identical, single Mach-Zehnder modulators in parallel optically and electrically but with unequal levels of optical and r-f power driving the two modulators. Both modulators are biased at the voltage, $0.5 V_{\pi}$, but with opposite slopes, so the modulators are 180° out of phase. Most of the optical power and a small fraction of the r-f power drive one modulator. The remainder of optical power and the r-f power drive the other modulator, creating relatively larger distortion products than in the first modulator. The two signals are combined incoherently in the photodetector. The r-f and optical splits are adjusted so that the distortion terms cancel exactly, but the signals do not. Since these two paths are in parallel, the effects of velocity mismatch apply equally to each MZM. Thus the distortion terms still precisely cancel regardless of frequency, and this is the result shown in Figure 4 - 5 and Figure 4 - 6. While the DPMZM is robust to velocity mismatch, in practice it is hard to make

broadband. A precise r-f power split must be maintained over the desired frequency band. If the split varies with frequency, the dynamic range will decrease at all but the narrow frequency at which the r-f split is correct.

Frequency Dependence of the CMZM

The Cascade Mach-Zehnder modulator was introduced as a series analogy to the DPMZM, of which the principle advantage is that there is only one optical path to the photodetector. Unfortunately the series nature of the modulator breaks the symmetry that protects the DPMZM from frequency-dependent dynamic range degradation. The dynamic range is reduced 3 dB by about 10 MHz, and 10 dB in 105 MHz, as shown in Figure 4 - 6, an expanded version of Figure 4 - 5. If $\Delta n = 0.18$ through velocity matching techniques, for example, then the 3 and 10 dB frequency points would increase by a factor of 10. Since velocity mismatch only effects the linearization mechanism which is nulling the third-order intermodulation product, the CMZM remains above (or equal to) the Mach-Zehnder modulator reference line at all frequencies. Similar to the DPMZM, this modulator is hard to build in that it requires an accurate broad-band split in the r-f drive power.

Frequency Dependence of the YFDCM

The linearized YFDCM exhibits the most gradual dynamic range roll-off of any of the broadband modulators except the DPMZM. Although the YFDCM has an anti-symmetrical transfer function about its operating point, there is no symmetry between the signal and the distortion cancellation mechanisms. In Figure 4 - 5, the YFDCM maintains 10 dB of linearization above the MZM at about 450 MHz. Interestingly, its dynamic range falls below the MZM reference line at about 3.3 GHz and nulls at approximately 8 GHz. This comes from the gain null in the YFDCM. As shown in the discussion at the end of this chapter, standard Y-fed DCMs become band-pass modulators with electrical lengths above $\theta = \pi$. The main passband of the modulator begins after 8 GHz, where unfortunately the dynamic range is the standard 110 dB.

The YFDCM modulator presents a real fabrication challenge. Although it is a simple structure with no precision biases, the linearization comes from careful control of the electrical length of the modulating electrodes. It may be significantly more difficult to control the electrical length

to adequate precision than it is to control a bias voltage. While the physical length of the electrodes can be made precisely by photolithography, the electrical length also depends on the coupling constant between the optical waveguides, κ , and this in turn depends on the waveguide in-diffusion process. The successful fabrication of directional couplers depends on highly reproducible values of κ , and this has to date been very difficult to do, as reported in Ref. [3.13].

Frequency Dependence of the DCM2P

The linearized directional coupler with two passive sections can be adjusted to provide high dynamic range with both intermodulation and second harmonic reduction in the absence of velocity mismatch. However, it suffers severely from velocity mismatch as shown in Figure 4 - 5 and Figure 4 - 6. By 80 MHz it is no better than the ordinary MZM and by 8 GHz it is 18 dB worse than the ordinary MZM. From the data, the DCM2P starts losing its dynamic range advantage at very low frequency, and loses 10 dB by about 10 MHz.

Unlike the DPMZM, in the DCM2P the mismatch between the r-f drive and the modulated signal velocities upsets the critical distortion cancellation conditions. Of the other three broadband linearized modulators in this analysis, the linearization mechanism only nulls the third-order intermodulation distortion. The second harmonic distortion, and all other even-order harmonic distortion terms, are nulled as a result of the anti-symmetric transfer function of the underlying simple modulator (MZM and standard YFDCM) and the linearization scheme does not interfere with even-order harmonic elimination. Thus the frequency dependent degradation in linearization only applies to the third-order intermodulation cancellation. In the DCM2P, both second-order harmonic and third-order intermodulation distortion terms are nulled by the same mechanism, and so frequency dependent degradation applies to both forms of distortion. As discussed in Chapter 3, second harmonic cancellation requires much stricter tolerances than intermodulation cancellation to the same power level for narrow signal bandwidths. Thus when the cancellation is upset in the DCM2P the effects are more drastic than the other linearized modulators. This makes the DCM2P uniquely bad. One could argue that with the velocity matching techniques discussed in Chapter 2, it would be possible to achieve a 100 MHz 3 dB point (1 Hz signal bandwidth) in this modulator, but

that would require $\Delta n < 0.05$. It is doubtful that it should ever be the modulator of choice even for a link with such a narrow signal bandwidth.

While the DCM2P does have the advantage that linearization is achieved with precision bias control instead of precision fabrication, the voltage on the bias electrodes must be maintained to a very high accuracy. This requires active bias stabilization. The original experiments on this modulator were performed at audio frequencies by Farwell in his Ph.D. dissertation, Ref. [4.3] where the narrow bandwidth effects would not be noticed. Subsequently, measurements at 1 and 2 GHz were reported by Schaffner in Ref. [4.4]; however, these were single-frequency measurements, with the bias values hand-optimized for the operating frequency. No bandwidth measurements around 1 and 2 GHz were made in that study.

Sub-octave Linearized Modulators

In sub-octave applications, the second harmonic may be ignored; the third-order IMD alone determines the dynamic range. Two sub-octave modulators are analyzed in this section: (1) the sub-octave dual series Mach-Zehnder modulator (SoDSMZM), as described in Refs. [3.15] and [3.16], which has two MZMs in series optically, and the same bias level on each Mach-Zehnder modulator, and a single r-f electrode covering both modulators;²⁰ and (2) a sub-octave directional coupler modulator (SoDCM), i.e., a directional coupler biased at $0.79 V_s$ instead of $0.43 V_s$ as described in Ref. [3.8]. Unlike the MZM, in the simple directional coupler modulator, the third-order intermodulation distortion and the signal null at different voltages, $0.79 V_s$ for the IMD and $1.0 V_s$ for the signal. Thus no extra electrode sections (as in the DCM2P) are needed to make a sub-octave directional coupler modulator.

Figure 4 - 7 shows the dynamic range as a function of frequency for the SoDSMZM and SoDCM compared to the standard MZM reference (horizontal dotted line). For both modulators, the dynamic range at zero frequency is superior to that of the four super-octave modulators in this analysis. As discussed in Chapter 3, there is no fundamental difference in

²⁰ There are other cascaded Mach-Zehnders proposed in the literature. One can reduce the noise figure, increasing the dynamic range, with a linearization point that requires different bias levels for each Mach-Zehnder modulator; also described in Ref. [3.15]. One can also split the r-f to reduce the loss and accumulated phase error on the second modulator.

the amount of non-linear distortion among modulators that null the third-order distortion terms. However, the average photocurrent and hence the shot-noise is very small in the SoDSMZM and in one arm of the SoDCM. And the resulting low shot noise accounts for the improved dynamic range. This is dramatically illustrated by the inferior dynamic range of the SoDCM if the other output arm is used, in which case the noise level is 12.31 dB higher. Even though there is still a frequency roll-off, it should be stressed that the SoDCM is a good candidate sub-octave modulator given its simplicity. It is just a standard directional coupler modulator evaluated at a different bias voltage. It does not require two modulators, there is no r-f or optical split, and the electrical length θ is not particularly critical around 0.5π ; any errors in θ can be compensated by adjusting V_B .

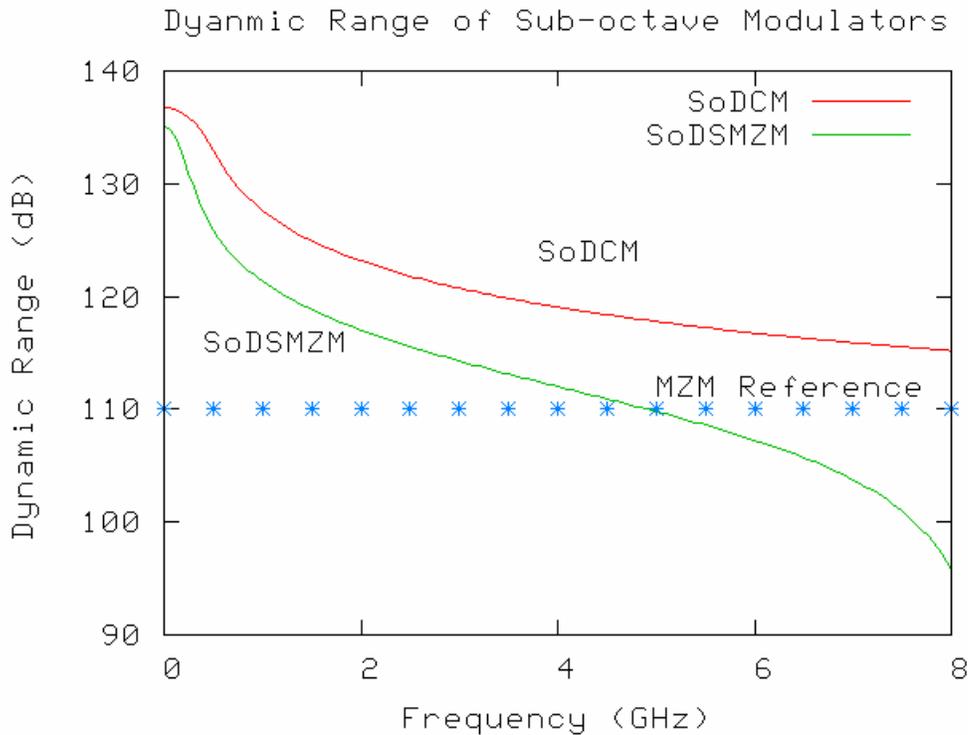


Figure 4 - 7: Dynamic range comparison of two sub-octave linearized modulators: SoDSMZM and SoDCM with a 1 Hz noise bandwidth.

The SoDSMZM dynamic range is a couple of dB lower than that of the SoDCM at zero frequency, and the dynamic range rolls off more rapidly. The initial difference comes from the higher average photocurrent level in the SoDSMZM. The roll-off is more rapid due to the gain null at 8 GHz of the SoDSMZM.²¹ Notice that the roll-off is less rapid than in the case of the most analogous super-octave linearized modulator, the CMZM. In the CMZM there is an r-f split between the two modulators. It is inferred from these plots that the distortion cancellation is more rapidly disturbed in the CMZM from the addition of the error in the r-f split caused by the velocity mismatch between the r-f and optical waves.

The dynamic ranges of these sub-octave modulators suffer from velocity mismatch, but generally less severely than super-octave modulators, with the exception of the DPMZM. The sub-octave modulators are much simpler than super-octave modulators. Sub-octave modulators should be considered for applications in which system design permits their use.

The Effects of Noise Bandwidth

In Chapter 3 the effects of noise bandwidth were analyzed in two respects: the total dynamic range of a link and the tolerance needed in the modulator parameters to be within 1 dB, 3 dB, and 10 dB of that optimal dynamic range. Scaling laws were shown for each of the modulators depending on the underlying non-linear distortion mechanism, e.g. second-harmonic or third-order intermodulation distortion. The velocity mismatch-frequency-length product parameter γ is analogous to any of the other modulator parameters that govern linearization; for instance, if the band edges of operation $\pm\Delta\gamma$ are considered to be the “tolerance” of γ around the design frequency.²² Every modulator has a frequency “tolerance” that is linear (in log-log space) over many orders of magnitude similar to the other parameter tolerance plots shown in Chapter 3. At higher noise bandwidths, the dynamic range is reduced by the ordinary scaling laws, but the tolerance requirement is similarly reduced. Since the two-tone analysis in

²¹ The SoDSMZM has a gain null at 8.1 GHz instead of 16.2 GHz because it was assumed that the modulator is 2x as long, in order to keep the same $V\pi$ as the DPMZ and MZM.

²² It is only a good analogy if allowances are made that different modulators have different inherent sensitivities to frequency. The DPMZM is theoretically independent of frequency if one ignores the frequency-dependence of the required r-f power splitter, and among the frequency-dependent modulators there are varying degrees of sensitivity.

preceding plots was done for 1 Hz signal bandwidths, it suggests greater frequency sensitivity than occurs in links with realistic noise bandwidths.

Figure 4 - 8 through Figure 4 - 12 show the dynamic range of photonic links using all of the modulators covered in this study at three signal bandwidths, 1 kHz, 1 MHz, and 1 GHz. The linearization parameters have been readjusted to maximize the dynamic range at zero frequency for each noise bandwidth case. Together with the previous figures showing a 1 Hz bandwidth, nine orders of magnitude in intervals of three orders of magnitude are shown. For traveling wave modulators with the canonical link parameters of Table 3-1, most applications would be limited to within a few GHz.²³ Together the plots show the evolution from 1 Hz to a 1 GHz, but generally a narrower range occurs in typical systems. For this reason and for the sake of brevity only the plots for a 1 kHz and 1 MHz noise bandwidth are described in detail.

A photonic link has both an absolute bandwidth range, and a noise bandwidth that may be different parameters. For example, in antenna remoting, the link must pass some large range of frequencies from the antenna to a remote receiver, but the electronic receiver searching for signals over this large range need only have a noise bandwidth large enough to accommodate the desired signal, assuming that signal is filtered in front of the modulator. This distinction is important because the modulator must have good gain over the entire bandwidth range, but the system dynamic range and the sensitivity to linearization parameters depends on the noise that gets through the signal filters after the modulator. In another example, consider a future CATV system with 200 channels. Channels are about 4.5 MHz bandwidth on sub-carriers separated by 6 MHz, starting from baseband, 200 channels requires a frequency range of 1.2 GHz. If the desired channel is demultiplexed after the modulator, then the noise bandwidth that effects the dynamic range sensitivity is 4.5 MHz bandwidth after filtering. A modulator for such a link would require flat gain and dynamic range out to 1.2 GHz (super-octave) but the dynamic range and sensitivity to frequency, as well as the modulation parameters discussed in Chapter 3, would be calculated with the 4.5 MHz noise bandwidth.

²³ Again, recall that these curves are all for $\Delta n = 1.8$. Changing ℓ or Δn will require rescaling the abscissa. Good velocity matching, for example $\Delta n=0.18$ instead of $\Delta n = 1.8$, with a loss of a factor of two in sensitivity requiring a new length of 2ℓ , will multiply the abscissa values by 5.

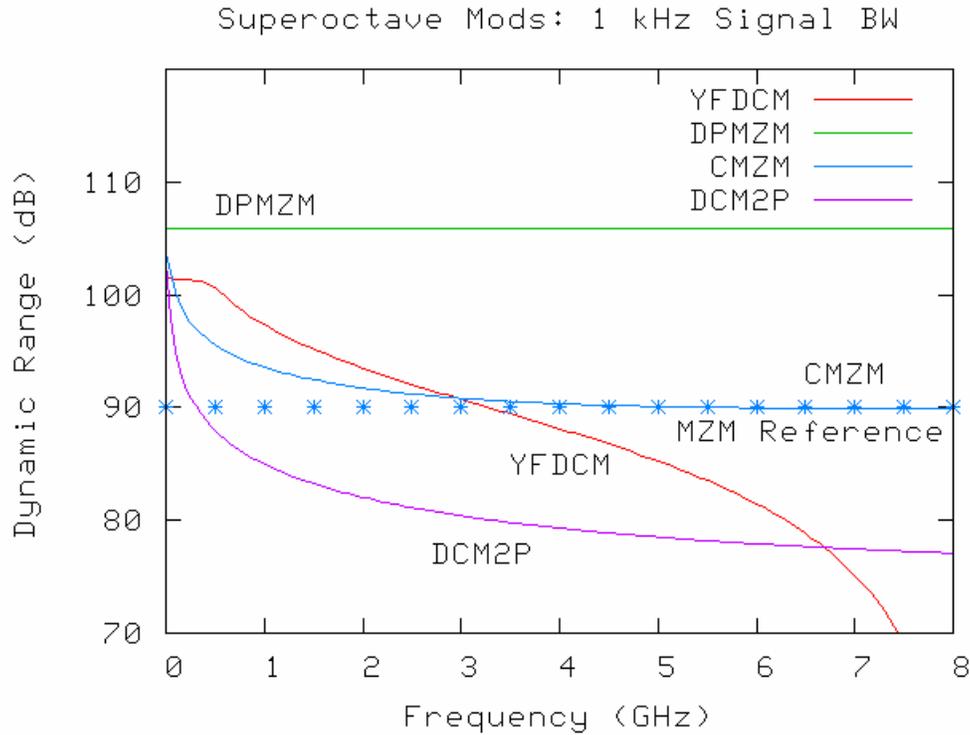


Figure 4 - 8: Dynamic range comparison with 1 kHz signal bandwidth of four super-octave modulator configuration: DPMZM, CMZM, DCM2P, and YFDCM. The modulator parameters have been adjusted to give the maximum dynamic range for a 1 kHz noise bandwidth.

Figure 4 - 10 shows a dynamic range sweep from 0 to 8 GHz in which the signal bandwidth is 1 MHz. The results are similar to the 1 Hz case but the roll-off in dynamic range with frequency is less severe. The DCM2P is still the least robust to frequency variation. However, it becomes less sensitive to velocity mismatch more rapidly with noise bandwidth than any other modulator. In the 1 MHz signal bandwidth calculation, the DCM2P loses 3 dB at about 150 MHz and 10 dB at 1 GHz; by comparison, the 10 dB point at 1 Hz was 10 MHz. This comparison isn't completely fair, because at 1 MHz there is less total linearization advantage; the DCM2P is only about 10 dB above the MZM at d-c, versus 20 dB in the 1 Hz case. Recall in Table 3-3, the slope of the log-tolerance-log-bandwidth scaling line equals about 5 for the DCM2P, the highest of any modulator linearization parameter (of any modulator covered in this study). It is this high because the linearization mechanism linearizes the second harmonic

as well as the third-order distortion, and the effect is that an increase in noise bandwidth quickly increases the tolerance. We have drawn an analogy between the tolerances of the

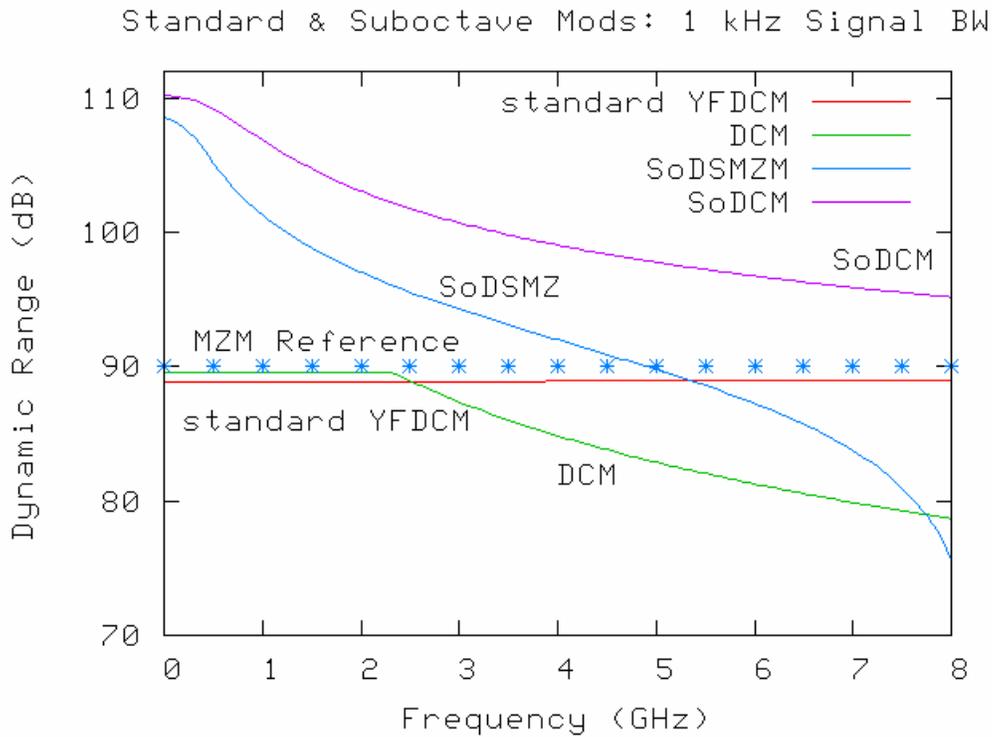


Figure 4 - 9: Dynamic range comparison with 1 kHz signal bandwidth of two sub-octave modulators: SoDCM and SoDSMZ; and the two non-linearized directional coupler modulator variants: the DCM and the standard YFDCM.

linearization parameters and the range of frequency operation, $\Delta\gamma$. The tolerance scaling in this analogy holds, and thus while the DCM2P seems hopeless at 1 Hz, it might have limited applications in links with 1 MHz or greater signal bandwidths. The YFDCM and the CMZM both show less sensitivity to frequency, though the improvement is not as dramatic as in the DCM2P because in the YFDCM and the CMZM the linearization scheme only nulls the third-order IMD and not the second harmonic distortion. While these modulators are improved, their dynamic range bandwidth is still significantly smaller than their bandwidth limitation from signal gain. Note also, in the case of the YFDCM, the relatively low frequency at which the gain nulls appears to contribute to the rapid roll-off in dynamic range.

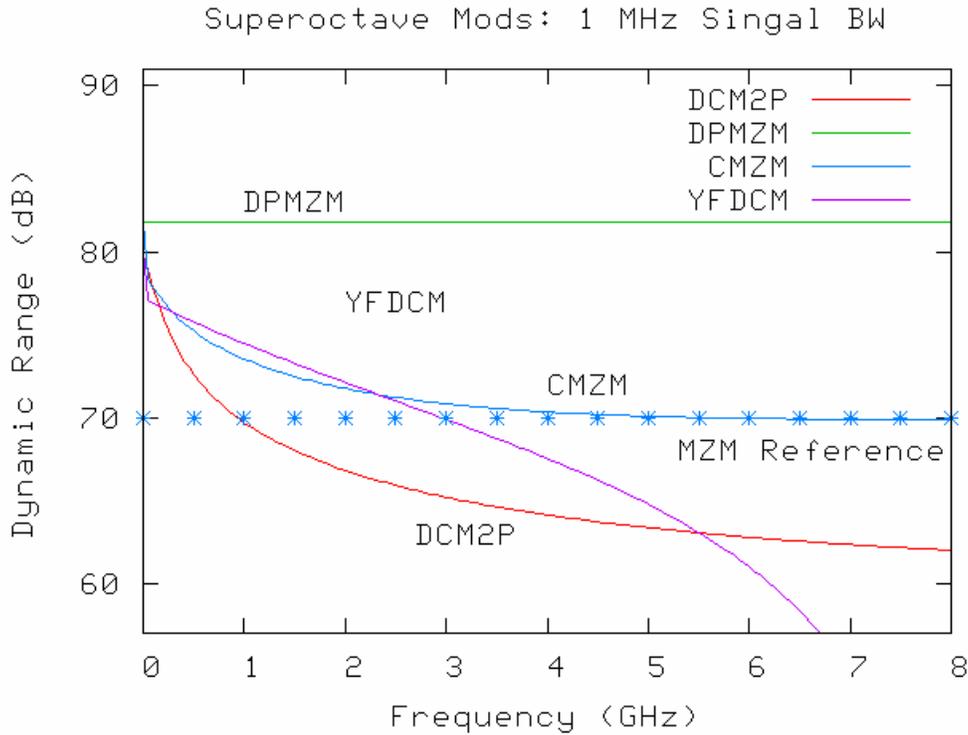


Figure 4 - 10: Dynamic range comparison with 1 MHz signal bandwidth of four super-octave modulator configuration: DPMZM, CMZM, DCM2P, and YFDCM.

Figure 4 - 11 shows a dynamic range sweep from 0 to 8 GHz in which the signal bandwidth is 1 MHz, the same sweep as in Figure 4 - 10; however, now the standard DCM and sub-octave modulators are shown. The break in the dynamic range of the standard DCM, where the second harmonic overtakes the third-order intermodulation distortion, occurs at 4.2 GHz instead of 1.8 GHz and the dynamic range has not reached the natural second harmonic level (the slope = 1/2 line from chapter 3) by 8 GHz. The sub-octave modulators, SoDCM and SoDSMZM, exhibit flatter frequency performance. The SoDCM only loses 10 dB over 8 GHz at which time it still has a 6 dB advantage over the MZM reference. Since this is a sub-octave modulator, it would be optimized for its target band-pass frequency range. Thus in a 1 GHz system centered at say 2 GHz it would be fairly independent of frequency. The figure highlights the performance of the SoDCM. At d-c its dynamic range of 88.8 dB is about 19 dB better than the MZM. This is nearly the improvement that the super-octave modulators get at d-c for 1 Hz signal bandwidth!

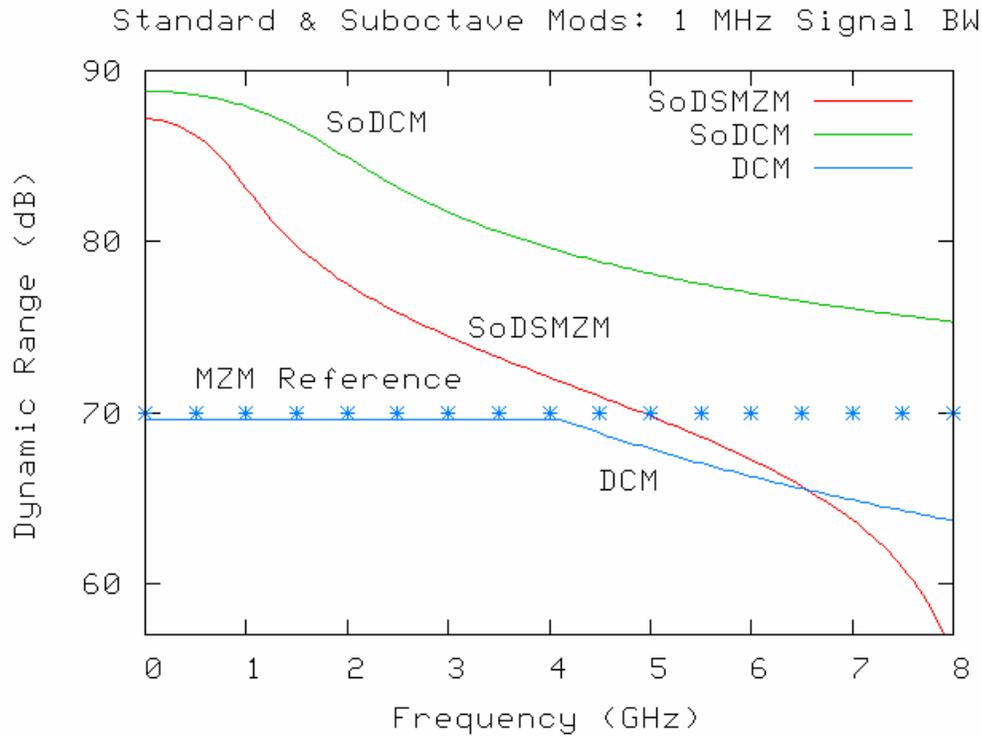


Figure 4 - 11: Dynamic range comparison with 1 MHz signal bandwidth of two sub-octave modulators: SoDCM and DSMZM; and the standard DCM.

For completeness, Figure 4 - 12 shows the dynamic ranges of the modulators in Figure 4 - 10 and Figure 4 - 11, but this time with a 1 GHz bandwidth. All of the effects illustrated in this figure are predicted from the descriptions of Figure 4 - 10 and Figure 4 - 11.

These figures help illustrate a point made in Chapter 3 on the benefit of sub-octave modulation. It is clear that larger signal bandwidths reduce the link dynamic range by the scaling law of the modulator, hence the MZM reference line is lower in each plot of higher noise bandwidths. It is also the case, that since linearized modulators scale with a $(4/5)$ slope, and standard modulators scale with a $(2/3)$ slope, that the dynamic range advantage one gains by using a linearized modulator in place of a standard modulator decreases with larger noise bandwidths. The difference between the levels of the MZM reference line and that of the

DPMZM illustrate this. It is 20 dB at 1 Hz but only 6 dB at 1 GHz. However, the increased dynamic range that comes from the reduced photocurrent level of sub-octave modulators is independent of noise bandwidth. For the SoDCM, it is 6 dB at 1 Hz yielding a total 26 dB above the MZM reference, and 6 dB at 1 GHz yielding a total of 12 dB above the MZM reference. At high enough noise bandwidths this becomes a major fraction of the total linearization, and more than half at 1 GHz. The larger the noise bandwidth, the more it pays to use a sub-octave modulator where possible.

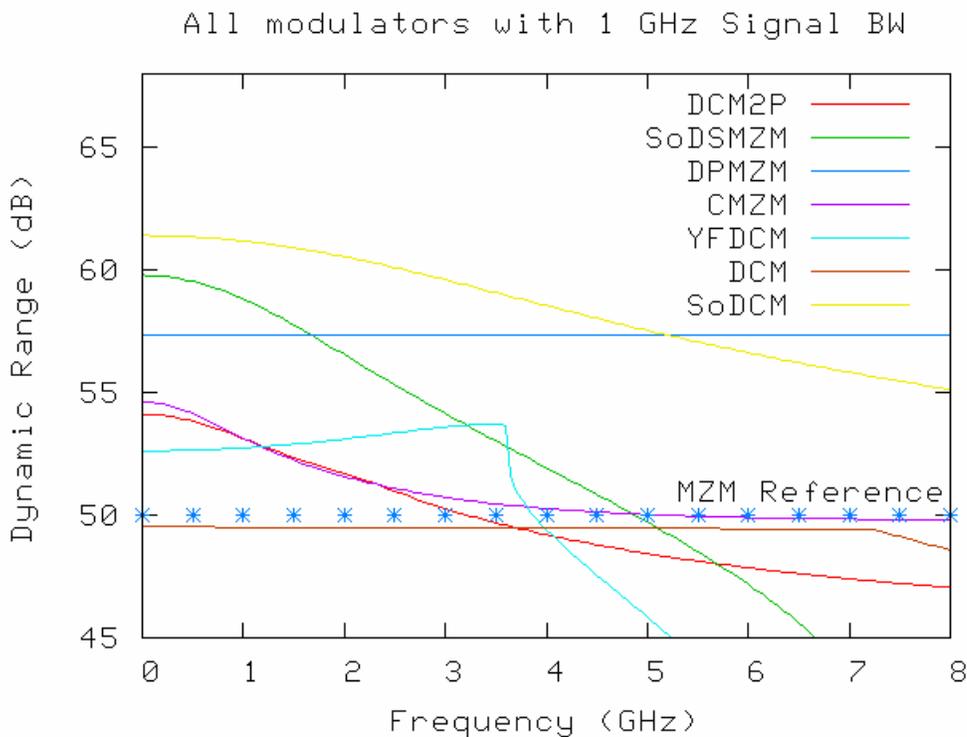


Figure 4 - 12: Dynamic range comparison with 1 GHz signal bandwidth of four super-octave modulator configuration: DPMZM, CMZM, DCM2P, and YFDCM; and the sub-octave modulators: SoDSMZM and SoDCM; and the standard DCM.

Re-optimization of Sub-octave Modulators for Band-Pass Applications

Figure 4 - 13 shows that the bias voltage of the SoDCM controls the frequency of the dynamic range optimum. Small adjustments in the bias around the d-c value of $0.79 V_S$ allow the full recovery of the low-frequency dynamic range optimum at any center frequency.

Unfortunately, this results in a relatively narrow operating bandwidth around the re-optimization frequency. And the bandwidth decreases as the center frequency is increased. But, recall that the actual dynamic range bandwidth also depends on the noise bandwidth as discussed in Chapter 3. For Figure 4 - 13 and Figure 4 - 14, the noise bandwidth is 1 Hz. For practical values of noise bandwidth, the peaks shown will be lower in value, but broader in frequency.

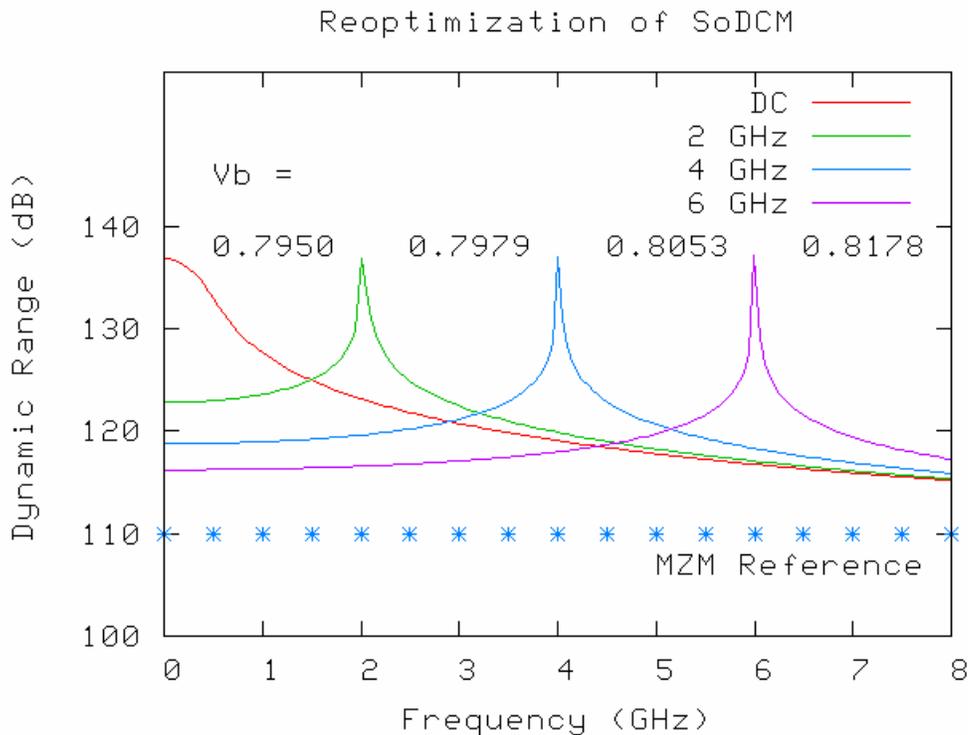


Figure 4 - 13: Re-optimization of the SoDCM by slightly adjusting the bias voltage around $0.79 V_S$ to fully recover the low frequency dynamic range optimum, but only over a narrow bandwidth.

Figure 4 - 14 shows the re-optimization of the SoDSMZM at the same frequencies as was shown in Figure 4 - 13 for the SoDCM. It is also possible to re-optimize the SoDSMZM at frequencies above zero. However, notice that in moving from zero-frequency to 2 GHz operation in the SoDSMZM the optimal bias changes by 2.2% while in the case of the SoDCM it only changed by 0.36%. The percentage change in optimal bias voltage is even

more pronounced moving from 2 GHz to 4 GHz, 10.5% versus 0.93%. The 6 GHz curve looks like it was omitted, however it just does not optimize well. And the bias voltage that the optimization routine found is about a 30% change from the 4 GHz curve and is a smaller value even though the progression had been larger bias values from one curve to the next. The reason is that the transfer function of the SoDSMZM goes through a null at 8 GHz. While the MZM nulls at 16 GHz, the SoDSMZM was modeled in the canonical link as a modulator of twice the length (20 mm) since it has two r-f modulating sections in series.

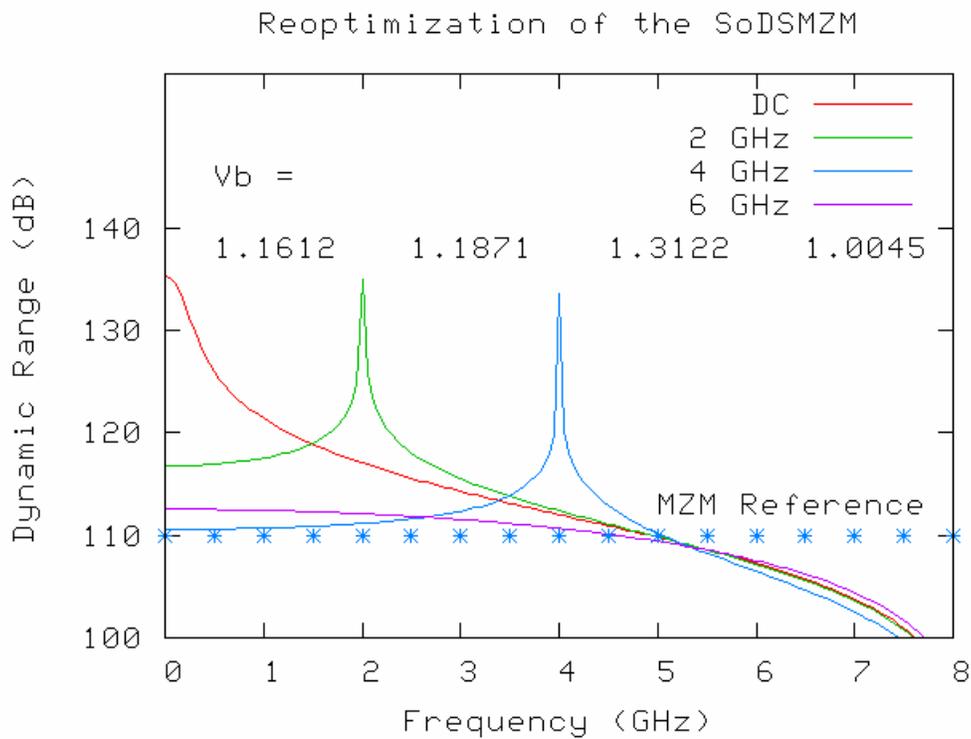


Figure 4 - 14: Re-optimization of the SoDSMZM. Adjusting the bias voltage fully recovers the low frequency dynamic range optimum, but only over a narrow bandwidth.

Figure 4 - 15 shows the curves of Figure 4 - 13 recalculated with a 1 MHz signal bandwidth. The bandwidth of the linearization is greater than in the 1 Hz case, but the re-optimization is still narrow band. The breaks on the left side of the curves of 2, 4, and 6 GHz come from a discontinuity in the equation for dynamic range. This effect arises from the “kissing condition” and was described in Chapter 3.

There are two interesting points from both of these modulators, which also apply to the super-octave modulators. First, as any modulator approaches its null, the required change in bias voltage to re-optimize increases. This results in the re-optimization having a narrower bandwidth with each band-pass region of increasing frequency. Second, the critical control of bias voltage will doubtless require the use of pilot tones to minimize the second harmonic or IMD at low signal bandwidths. However, the results of Figure 4 - 13 through Figure 4 - 15 indicate that these pilot tones will have to be within the band of interest, not at low frequencies, and such a feedback circuit is more challenging to engineer.

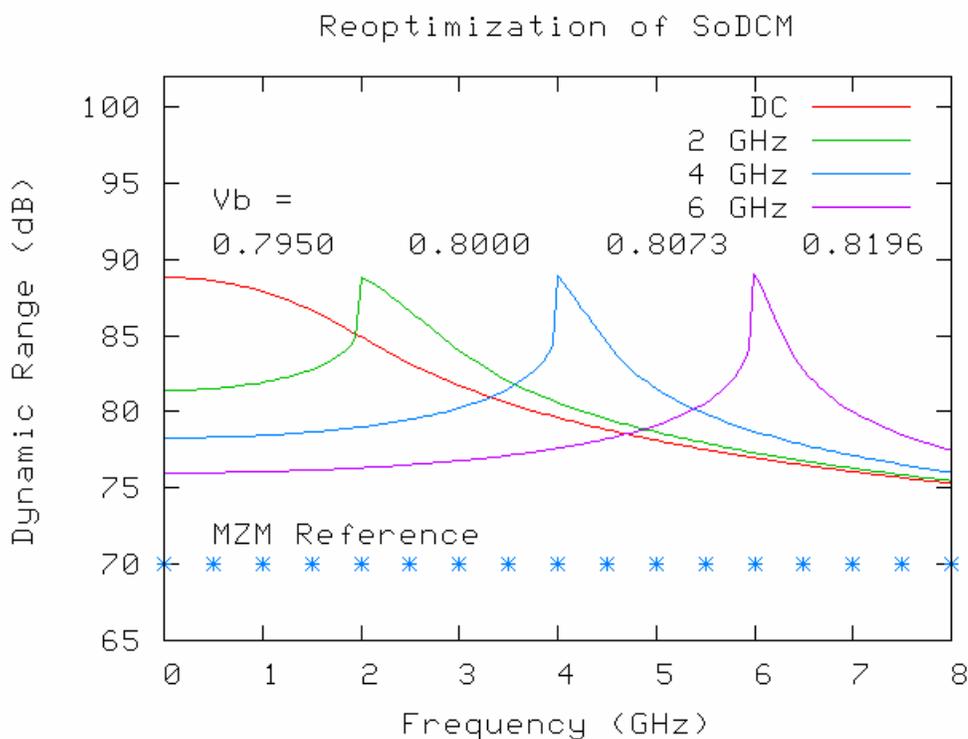


Figure 4 - 15: the dynamic range of the SoDCM re-optimized at 2, 4, and 6 GHz, this time assuming a 1-MHz signal bandwidth.

Bandwidth Recovery with Periodic Rephasing

The most direct way of increasing the bandwidth of linearized modulation is to arrange for velocity matching by one or another transmission fabrication technique described in Chapter

2. However, there are limits to the accuracy with which the match can be accomplished, and they all sacrifice modulator sensitivity. An alternative method of obtaining increased bandwidth is to periodically correct the phase delay in the modulator transmission line electrodes. This is done by breaking the transmission line into a number of segments and rephasing (more accurately, “re-timing”) the signal at the beginning of each segment. This is velocity-matching “on the average.” The technique has been used successfully in a number of forms as described in Chapter 2; see also Ref. [2.3]. The numerical solver written for this study is capable of calculating such periodically re-phased modulators. The modulator is incrementally velocity mismatched for a few matrices and then re-phased for the next set of matrices, and so on. This calculation assumes that the re-phasing is done as a correction in time delay independent of frequency. Some schemes, for example the phase reversal technique of Ref. [2.7], are quite narrow band because they are correctly re-phased at only one frequency. Others, such as the corporate feed and the antenna segmenting technique, also summarized in Ref. [2.6], which use geometric “true time delays” are inherently broadband.

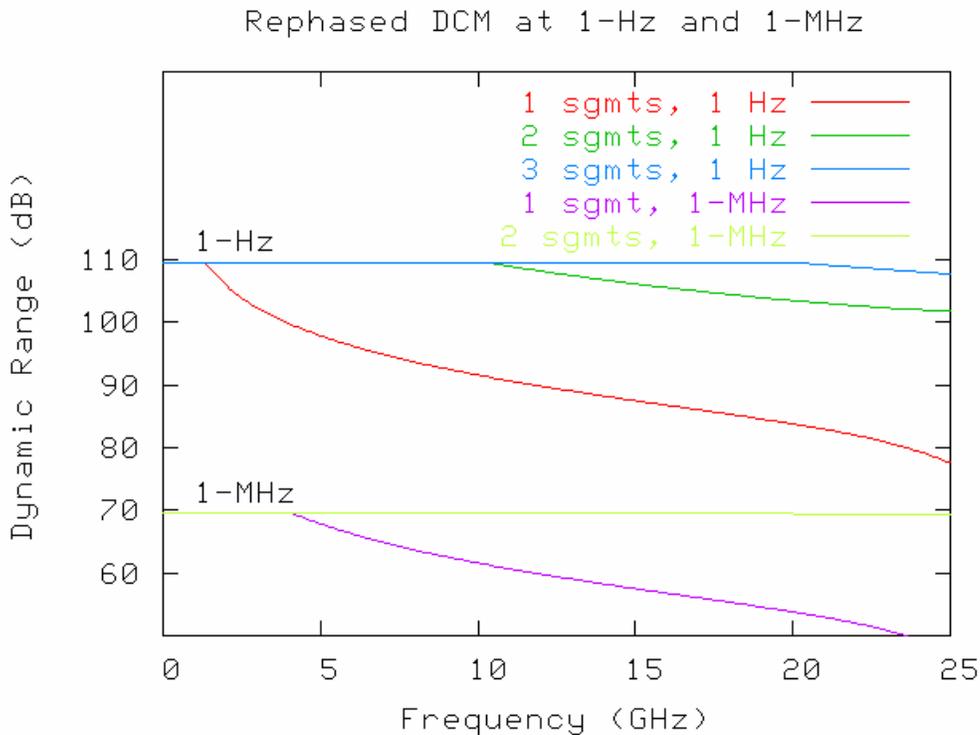


Figure 4 - 16: Dynamic range versus frequency for the simple DCM with multiple electrode segments, at 1 Hz and 1 MHz. The

breaks between occur where the second-harmonic distortion becomes more significant than the IMD.

Figure 4 - 16 shows the results of this calculation for the standard DCM link with the parameters previously used, but with the modulator's transmission line having one, two, three, and four segments. The curve representing a single segment repeats the result in Figure 4 - 3 (no rephasing) for reference. In the 1 Hz case, with only two segments (one rephasing), the bandwidth over which the dynamic range is flat improves vastly, and using four segments gives an essentially flat dynamic range. Note that the null that appears at 26 GHz for the one-segment modulator (not shown in the figure), occurs at 104 GHz, a frequency four times higher, in the four-segment case. The 1 MHz case is also plotted, and all other plots of re-phased modulator electrodes will be done at a noise bandwidth of 1 MHz.²⁴ In the 1 MHz case, a single rephrasing gives essentially flat dynamic range. Of course, it would still be preferable to use a standard MZM for non-linearized modulators in broad-band links since it has no second harmonic and thus does not require re-phasing to eliminate it. That is, unless rephrasing is also desired to increase the gain bandwidth of the MZM and the DCM.

The results for the DCM2P super-octave modulator remain discouraging, even with re-phasing. Figure 4 - 17 shows the dynamic range for one to eight re-phased electrode segments. This plot spans an 8 GHz range, which is enough to capture the interesting effects of the bandwidth of linearized modulation and its recovery. Recall that previous plots of the DCM2P, done with a 1 Hz noise bandwidth, and no recovery mechanism, showed the loss of linearization in the MHz range instead of the GHz range. This plot also assumes ideal directional coupler modulator sections, perfect bias control, and no electrode loss. Given the modulator's sensitivity to its two passive bias voltages, it is unlikely that this modulator is suitable for microwave applications of relatively narrow noise bandwidth. As described above, the sensitivity to velocity mismatch and linearization parameters become less severe with increasing signal bandwidth faster than any other modulator.

²⁴ In general, re-phasing adds considerable device complexity (an exception can be made in some cases for reflection modulators, see Chapter 4). And so, during the evaluation of these schemes, the realistic 1-MHz signal bandwidth is used to avoid the conclusion that an excessive amount of rephrasing is necessary.

There is an interesting and unexplained effect in the DCM2P. The dynamic range recovery is not monotonic in the number of re-phasing sections. In particular, three is slightly better than four and seven is significantly better than eight. The author does not have a good physical argument for this effect. The modulation is applied to a single section in the DCM2P, which makes these results counterintuitive. The DCM2P is already a complicated structure without re-phasing. This mystery is left as an open item for future work.

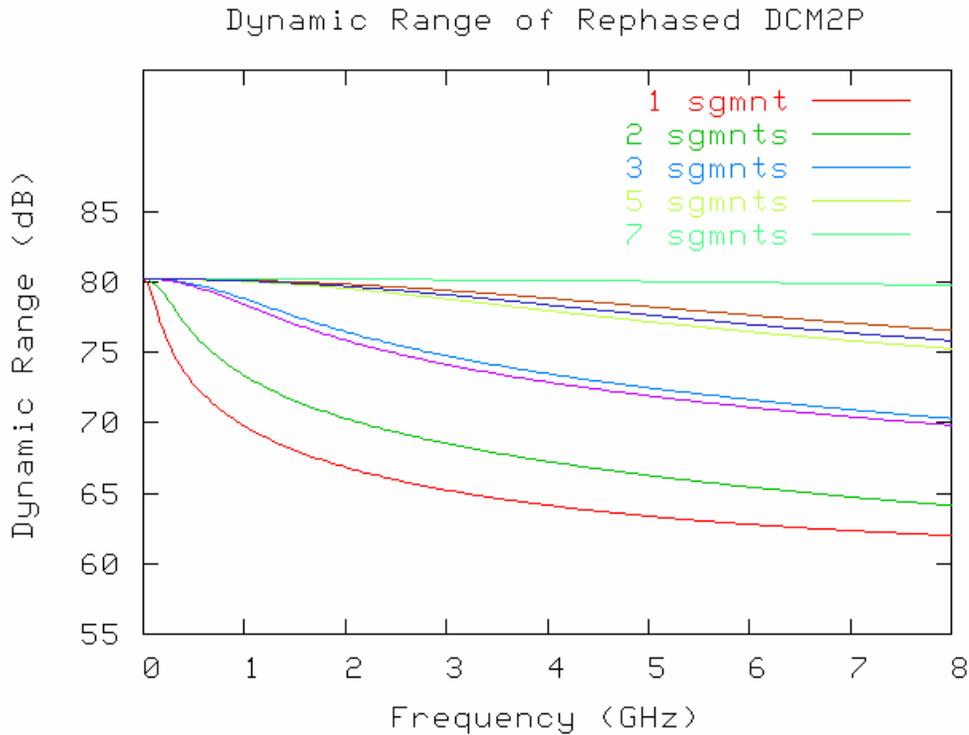


Figure 4 - 17: dynamic range versus frequency for the DCM2P with one to eight electrode segments, calculated with a 1 MHz noise bandwidth. While rephasing yields improvements, the results remain inferior to other modulators.

Figure 4 - 18 shows the effects of rephasing on the YFDCM. The curve for three and four segments show a sustained improvement in the frequency dependence of the dynamic range. The potential for further improvements by re-phasing are discussed further in at the end of this chapter in the section on band-pass operation of DCMs.

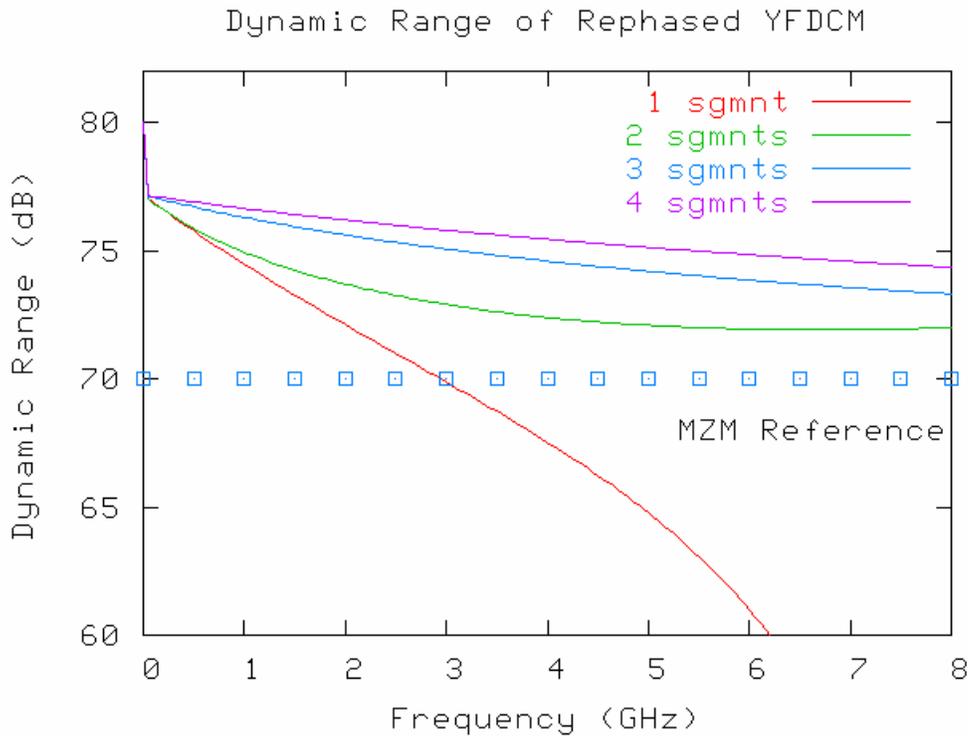


Figure 4 - 18: Dynamic Range of Re-phased YFDCM ($\theta=1.43\pi$) for one to four electrode segments, computed with a 1 MHz noise bandwidth.

An interesting result of the application of rephasing is seen in both of the serial MZM modulators: the CMZM and the SoDSMZM. The frequency-dependence of the dynamic range is shown in Figure 4 - 19 and Figure 4 - 20. In both figures, the curve labeled “1 segment,” refers to one set of traveling wave electrodes spanning the two Mach-Zehnder modulators (with an unrealistic zero distance assumed between the two modulators). This is simply a repeat of the curves in Figure 4 - 10 and Figure 4 - 11. The interesting point is that with only one re-phasing (2 segments) the modulators exhibit frequency independent performance. Breaking the electrodes into three segments yields an improvement over one segment, but not as good as two or any even number of segments, all of which yield frequency independent dynamic range. In this case, the improvement is not monotonic with re-phasing, the optimum is for the r-f to be re-phased once at the second modulator.

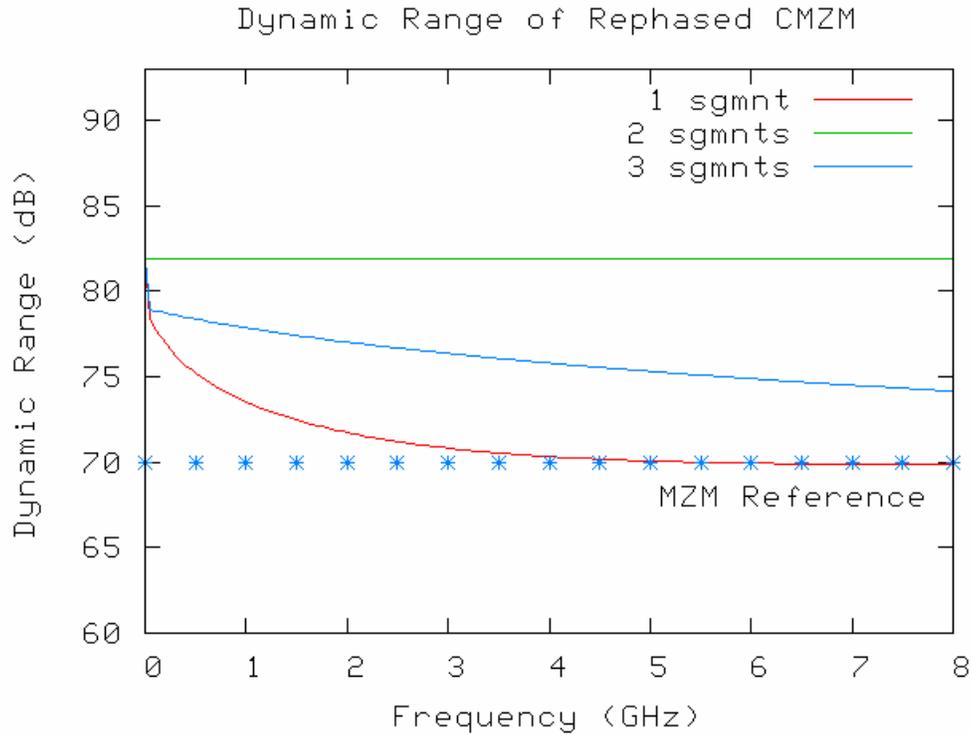


Figure 4 - 19: Dynamic range of a re-phased CMZM with one, two, and three electrode segments calculated at 1 MHz signal bandwidth. The CMZM with two segments (or any even number) has a dynamic range that is independent of frequency.

This phenomenon results from the symmetry of the modulator. The degree of modulation of a Mach-Zehnder with finite transit time is described exactly by Eqn. (3.27). This describes both a lumped electrode with finite transit time and a traveling wave modulator with velocity mismatch between the r-f and optical modulation wave. The integral gives the accumulated phase shift over the length of the modulator. The differential phase shift in each arm of the Mach-Zehnder is then converted to modulation by a lumped interference mechanism, the second Y-branch. If two MZMs are in series, each modulator is the same length, and the r-f wave and optical modulation envelope are re-phased at the second modulator, then the accumulated phase shift is the same on each modulator. Thus the only effect frequency can have is on the sensitivity of the overall modulator, it cannot effect the linearization of the modulator. The effects of frequency simplify to a measure of the phase error at the second modulator. Thus in the case of three segments, the phase error is non-zero at the second

modulator because the modulator is re-phased at $1/3$ and $2/3$ of its total effective length. There is less phase error than in the case of one segment but more than the case of two segments. While the performance does not monotonically improve with the number of segments, it does monotonically improve with a reduction in phase error at the second modulator. An additional point: the CMZM is more sensitive than the SoDSMZM to a given amount of phase error as demonstrated by the more rapid roll-off in the one and three segment cases. As stated above this is a result of the added complexity in the r-f split of the CMZM over the SoDSMZM.

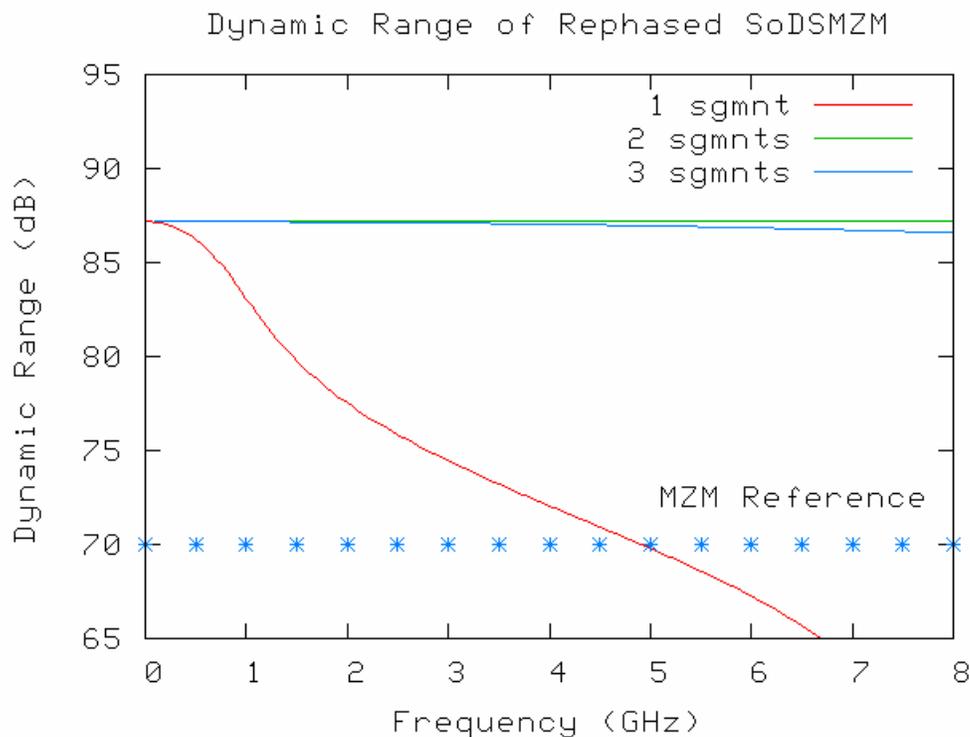


Figure 4 - 20: Dynamic range of the SoDSMZM with one, two, and three electrode segments, calculated with a 1 MHz signal bandwidth. Note that two segments (or any even number) give frequency independent dynamic range.

A similar dramatic improvement is obtained in the SoDCM, biased at the $0.79 V_s$ point, as shown in Figure 4 - 21. The curve from Figure 4 - 11 is shown for reference along with curves for two, three, and four electrode segments. The curve for one segment initially shows a deep

roll-off in dynamic range and then a more gradual roll-off, with only 5 dB of dynamic range improvement over the standard MZM remaining at 8 GHz. With just two segments, the roll-off is made gradual over the whole range. With four segments, there is 14 dB of dynamic range improvement remaining at 8 GHz. This figure assumes a 1 MHz noise bandwidth. As a comparison to super-octave modulators, even though the two electrode-segment CMZM (shown in Figure 4 - 19) has no roll-off, the dynamic range of the two electrode-segment SoDCM is better (shown in Figure 4 - 21) up to 6.5 GHz.

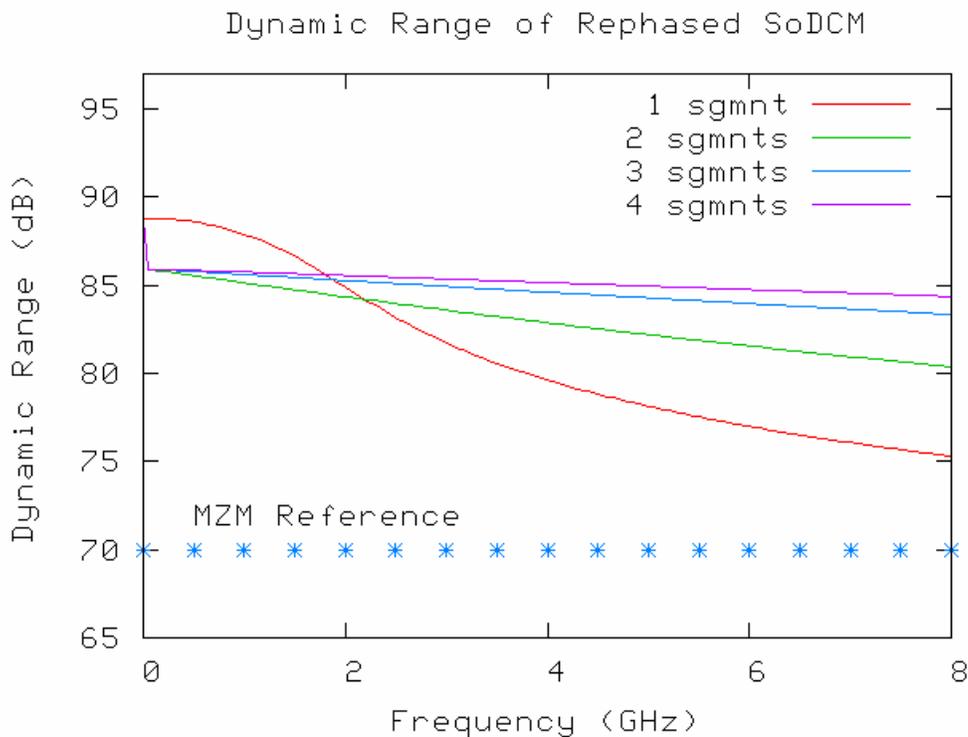


Figure 4 - 21: Dynamic range versus frequency for the SDCM with 1, 2, 3, and 4 electrode segments, calculated with a 1 MHz signal bandwidth. Note the initial drop-off of curves 2 - 4 dB is a result of the “kissing condition” described in Chapter 3.

Segmenting the r-f electrode has a low-frequency gain penalty. While flattening the gain vs. bandwidth curve, it also lowers the absolute gain at low frequencies. The reduction occurs from splitting the r-f power into the re-phased paths. When the r-f drive power is split N ways to feed the N electrodes, the voltage at the input to each electrode section is reduced by \sqrt{N} , as

described by Sheehy in Ref. [1.3]. Figure 4 - 22 shows the calculated gain of a simple MZM with one, two, three, and four electrode segments. Above some cross-over frequency each curve has a better absolute gain than the curve with one fewer electrode segments.

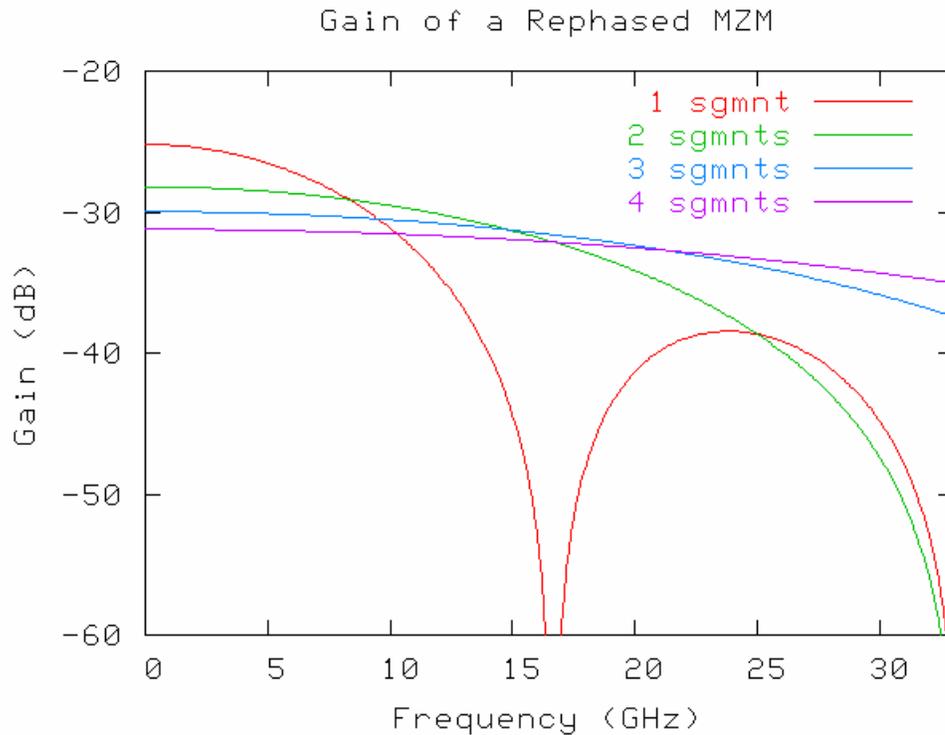


Figure 4 - 22: The gain of a standard MZM with multiple electrode segments.

Bandpass Features of Directional Coupler-Based Modulators

The frequency-dependent gain of modulators using directional couplers such as the standard DCM and the standard YFDCM is a mathematically rich function, but not analytically solvable, unlike the gain function of a MZM. An approximation for a small velocity mismatch exists in the literature, Ref. [3.10], but it is not accurate in our regime of interest and in fact does not even satisfy conservation of energy. Using the numerical calculation method developed for this study, we have discovered interesting properties of the gain function of a simple directional coupler and a Y-fed directional coupler modulator. This section treats some

of these interesting properties, particularly the possibility of band-pass operation of DCM-based modulators at very high frequencies with reasonable bandwidths, even though velocity mismatched.

The gain functions of a MZM and a directional coupler-based modulator have the functional dependencies

$$G_{MZM} = A(V_B) \frac{\sin^2(\pi\gamma)}{(\pi\gamma)^2}, \quad (3.29)$$

$$G_{DCM} = F(V_B, \gamma, \theta). \quad (3.30)$$

The function $A(V_B)$ in the MZM gain expression contains all of the link parameters of Table 3-1 that establish the gain at zero frequency. The bias voltage dependence is in A , indicating that while an off-biased MZM has a non-optimal gain, the voltage does not effect the frequency profile of the gain in any way. In a directional coupler the bias voltage and the electrical length both affect the frequency profile of the gain function, and this dependency must be explored numerically, which also means that an explicit modulator must be chosen. The standard DCM and the standard Y-fed DCM were both studied. Although the Y-fed modulator has a more complicated zero frequency transfer function, it actually has a simpler frequency profile and a simpler functional form since $V_B=0$ is the desired bias condition. Thus, the Y-fed DCM is addressed first.

The Gain of a Y-Fed Directional Coupler

Figure 4 - 23 shows the gain of a Y-fed directional coupler modulators with electrical lengths 0.5π , 1.5π , 3.5π , and 5.5π . The standard value is $(1/\sqrt{2})*\pi$ for the Y-fed DCM proposed in Ref. [3.9] and that of the linearized modulator is 1.43π , proposed in Ref. [3.11] are sufficiently similar to the curves of electrical length 0.5π and 1.5π . This figure shows that the Y-fed DCM is essentially a band-pass modulator for the lengths 1.5π and longer. Given acceptable conductor losses, r-f reflections, and radiation, we could operate this modulator at any higher frequency by simply increasing the electrical length with no need to velocity match the modulator. Moreover, the passbands for length 1.5π , 3.5π , etc. are nearly twice as wide as a

low-pass modulator (0.5π). The modulator for $\theta=0.5\pi$ has a few extra dB in gain, but after $\theta=\pi$ there is no further reduction in the optimum passband gain, for larger electrical lengths, which occur at higher frequencies. This represents the transition from a low θ modulator to a high θ modulator.

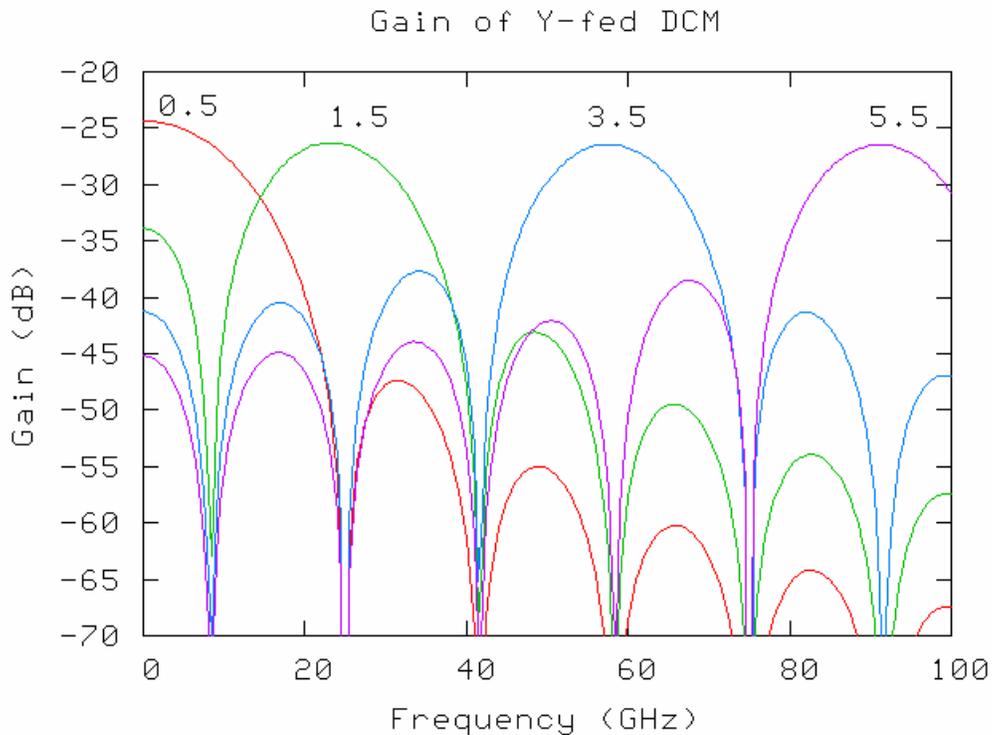


Figure 4 - 23: The gain of a Y-fed DCM for $V_B=0$ and the electrical length, $\theta=(2n+1)\pi/2$ for $n=0, 1, 2$, and 3.

Notice that the zero frequency values for the different sweeps are decaying with θ even though for $\theta = (n+1)\pi/2$ (n integer) the DC value is a local optimum. This is counterintuitive since there is an underlying sinusoid that has the form $\sin(\ell(\Delta\beta^2 + \kappa^2)^{1/2})$ in all directional couplers (recall that κ is the inherent coupling coefficient between the two coupled waves of a DCM and $\Delta\beta$ is the electrooptically induced term for the differential coupling). If we increase θ , ℓ , or κ why should that reduce the sensitivity? This effect is further illustrated in Figure 4 - 24. Notice that the transfer functions become less sensitive at their bias voltage $V=0$ for increasing electrical lengths. Further note that every curve with $\theta > \pi/2$, the voltage at which the peak

transfer occurs is greater, and that the greater the electrical length, the smaller the zero voltage sensitivity, the larger the voltage at which it regains its peak sensitivity.

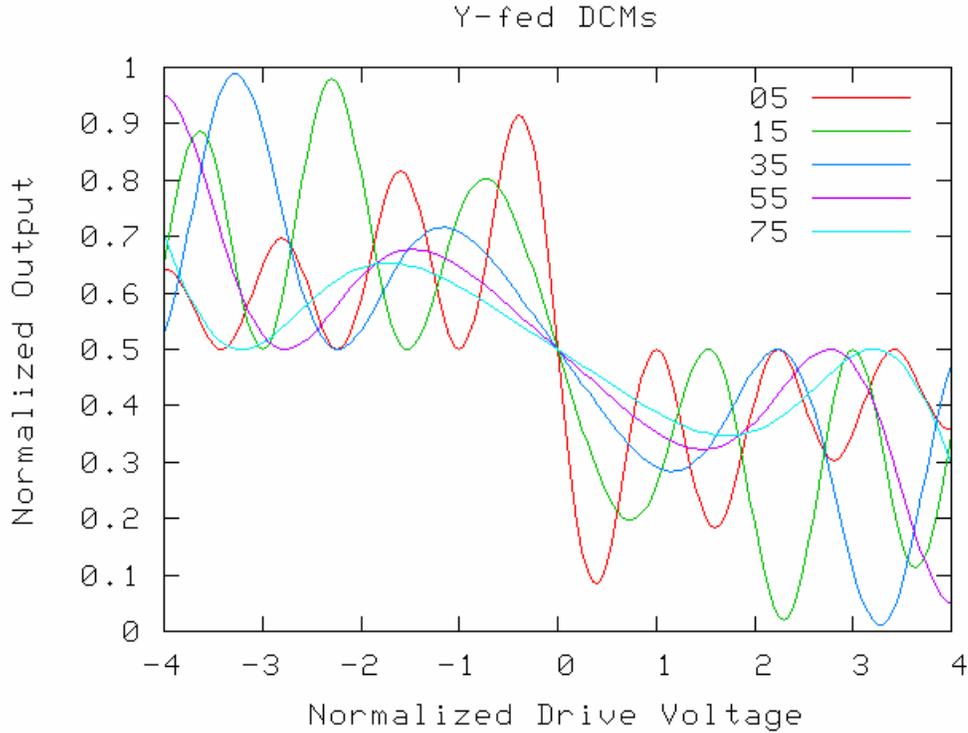


Figure 4 - 24: The transfer function of Y-fed DCMs with electrical length $\theta = (2n+1) \pi/2$, for $n=0, 1, 2, 3,$ and 4 . This figure illustrates the reduction of sensitivity at $V=0$ for higher electrical lengths, but notice that for any higher electrical length, there is still a voltage at which the sensitivity is only about 3 dB down from the that of the modulator with $\theta = \pi/2$.

There is a very simple first-order approximation for $(\gamma-\theta)$ small of the gain function for Y-Fed directional coupler modulators. This expression works best for the second and higher lobes of Figure 4 - 23, $\theta \geq 3\pi/2$. The approximation is

$$Gain_{yF} \approx A \frac{\sin^2(\pi(\gamma-\theta))}{\pi^2(\gamma-\theta)^2}. \quad (3.31)$$

A is a scalar value representing the zero-frequency link gain. This expression is accurate around the passband; it gives the location and width with high accuracy and almost gives the precise shape. It is sufficiently accurate that the difference from the numerical method results would not likely be verifiable in an experimental setting. It is inaccurate in the side lobes, especially many side lobes away from the passband. It shows the gain of the Y-fed DCM has approximately the same frequency profile as the MZM, a sinc-squared function, provided that γ in the MZM gain expression is replaced by $(\gamma-\theta)$. The electrical length θ is a controlled fabrication parameter over a wide range,²⁵ making the DCM a promising candidate for a high-frequency band-pass modulator.

It is desirable to have an expression of the Y-fed DCM gain that also correctly models the behavior for $\theta \leq \pi$. The following expression contains a multiplicative correction for the side lobes and the first curve with $\theta = \pi/2$

$$Gain_{YF} \approx Ae^{\frac{-2}{\theta\pi}(\gamma-\theta)} \frac{\sin^2(\pi(\gamma-\theta))}{\pi^2(\gamma-\theta)^2}. \quad (3.32)$$

This expression is very accurate. It predicts all of the curves in Figure 4 - 23. It should be used for building band-pass modulators based on the Y-fed DCM. Expressions (3.31) and (3.32) were “discovered” by the author, but not derived. Their accuracy was empirically ascertained.

Figure 4 - 25 shows a Y-fed directional coupler with electrical lengths between 0.5π and 1.5π . The modulator is biased at its point of anti-symmetry $V_B=0$. These curves show many interesting effects of directional couplers and the effects of velocity mismatch. First, notice that the peak gain at $f = 0$ is actually achieved for $\theta = 0.4\pi$, not $\theta = 0.5\pi$. Additionally, for lengths that are not integral multiples of 0.5π , the gain does not null completely although it can fall by at least 25 dB. This occurs because there are really two optical paths in a directional

²⁵ It is not easy to fabricate a specific electrical length that is repeatable. However, it is controllable over a wide range. That is by reducing the spacing of the two waveguides in a directional coupler while keeping the length the same, the electrical length goes up very fast (exponentially), and one could make a modulator from a wide range of electrical lengths. However, it is hard to keep the electrical length controlled to within one part in ten.

coupler and if their null positions do not exactly coincide there will be residual modulation. Second, notice the position of the nulls. A standard MZM with the same length and velocity mismatch nulls at 16.2 GHz. The standard Y-fed DCM nulls at 25 GHz, an improvement of almost 40% over the Mach-Zehnder at the 3 dB point. This is a one-time effect; the nulls after the first are separated by about 16 GHz, resulting from the underlying $\text{sinc}(\pi\gamma)$ variation of the modulator. The curve of $\theta = 1.5\pi$ is the same curve in both Figure 4 - 23 and Figure 4 - 25, showing the passband property in more detail in the second plot.

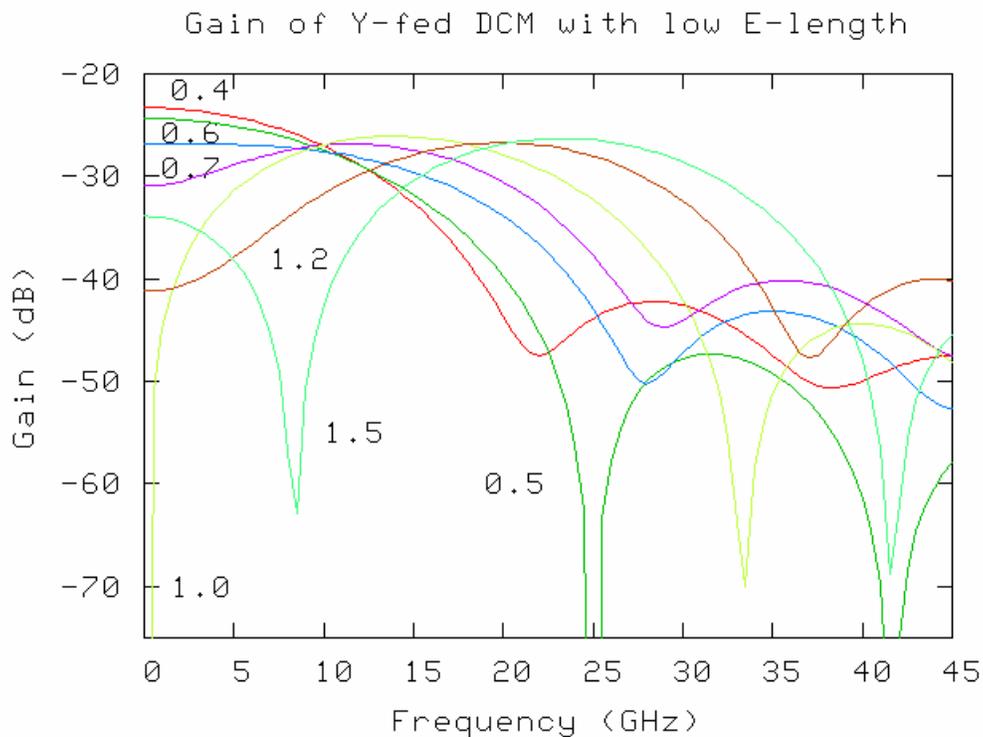


Figure 4 - 25: The gain of a Y-fed DCM with bias voltage $V_B=0$ and electrical length $\theta=0.4\pi, 0.5\pi, 0.6\pi, 0.7\pi, 1.0\pi, 1.2\pi,$ and 1.5π

The electrical length may be chosen to make a flatter modulator response over a wide band even in the presence of loss. The 0.6π curve is flat, losing only 1 dB out to 10 GHz, unlike the response for an electrical length of 0.5π or the response for a MZM. An electrical length of $0.6\pi - 1.0\pi$ demonstrates a variable amount of peaking. The zero frequency value may be 3 - 10 dB below the optimum, but some of this is regained such that at around 10 - 15 GHz the

response is only a couple of dB below the optimum. Given this novel property and a known value for frequency-dependent conductor losses, the electrical length may be chosen to compensate for the electrode losses, providing a flatter response over a wide band. Equalizing the response of a traditional link can require complex and expensive microwave networks. There is a genuine opportunity with this modulator to inexpensively and compactly make a broadband link with a flat frequency response.

The Gain of a Standard Directional Coupler Modulator

The calculations shown in Figure 4 - 26 and Figure 4 - 27 are a repeat of the calculations shown in Figure 4 - 23 and Figure 4 - 25, but for a standard directional coupler modulator (all the optical input into one arm of the directional coupler) in place of a Y-fed DCM. This calculation is a little less straightforward than that for the Y-fed DCM in that the bias voltage is not zero, and the appropriate bias voltage depends on the electrical length. The bias voltage used for each modulator length was chosen to maximize the zero frequency gain, giving a consistent operating point to compare the various electrical lengths. However, the result is a slightly different operating point than one would ordinarily choose to null the second harmonic, which would also be different for each electrical length.

Figure 4 - 26 shows similar behavior to the Y-fed DCM, but contains a few key differences that make it a less suitable candidate for a band-pass modulator.²⁶ First, notice that although the zero frequency gain response decreases with increasing modulator length, it does not fall as rapidly as that of the Y-fed DCM. In the Y-fed DCM it was shown the transfer function goes to zero gain at infinite electrical length about zero bias voltage, however the DCM has no such property. There is a secondary lobe at higher frequencies as also occurred in the case of the y-fed DCM, but now the secondary lobe in gain reaches a lower maximum than the corresponding value at zero frequency, unlike the YFDCM. This lobe shifts in frequency with the form and location, $\sin^2(\pi(\gamma-\theta))/\pi^2(\gamma-\theta)^2$, and becomes lower in peak gain the longer the

²⁶ “Band-pass” here does not mean sub-octave, but means operation in the higher frequency lobes of the gain transmission function of the velocity mismatched case; in this case the optimum gain response is shifted in frequency from zero to some higher frequency by increasing the electrical length. The DCM is a great candidate for sub-octave modulators.

modulator. Thus, the “long” standard DCM seems to be a less attractive bandpass modulator than the Y-fed DCM.

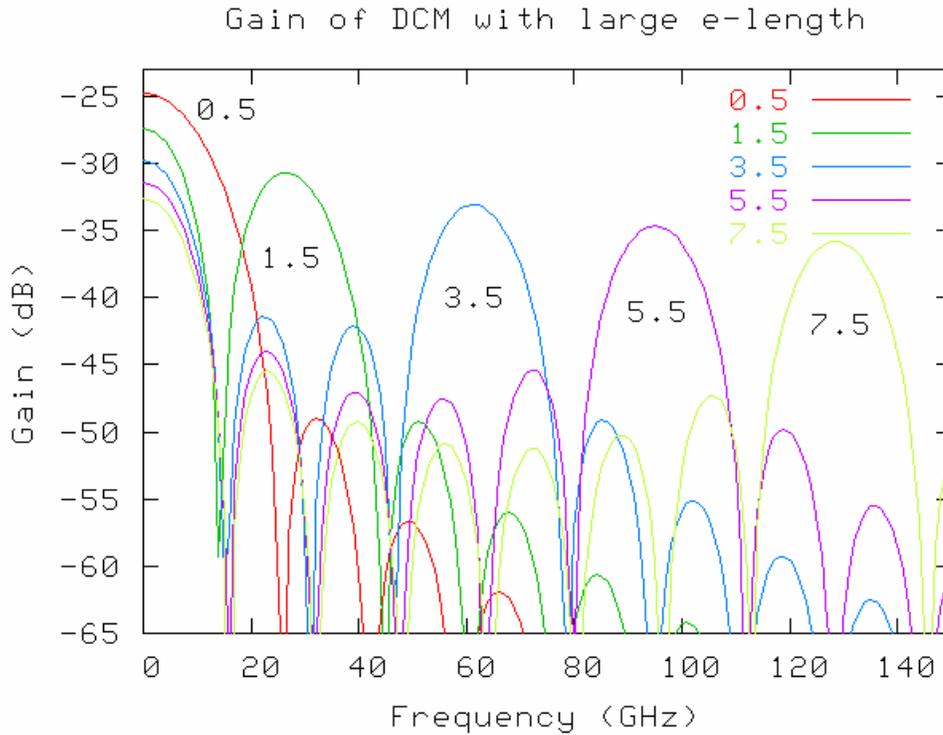


Figure 4 - 26: The gain of DCM with the electrical length, $\theta=(2n+1)\pi/2$ for $n=0, 1, 2, 3,$ and 4 , and bias voltage chosen to optimize the gain at $f=0$. Choosing a bias to optimize the gain at the maximum passband frequency makes about 0.5 dB improvement for all curves.

Figure 4 - 27 shows the DCM with normalized electrical lengths between 0.5π and 1.5π . This behavior is similar to that of Figure 4 - 25 except that the gain penalty is more severe. For a flat or peaking response, there is little reason to use this modulator when the Y-fed version preserves near optimal gain. An electrical length of 1.3π is quite interesting in that it does not null until 40 GHz, but at a penalty of being down 16 dB from the optimum gain. This is over twice the frequency of the Mach-Zehnder null, and has not yet become a band-pass modulator.

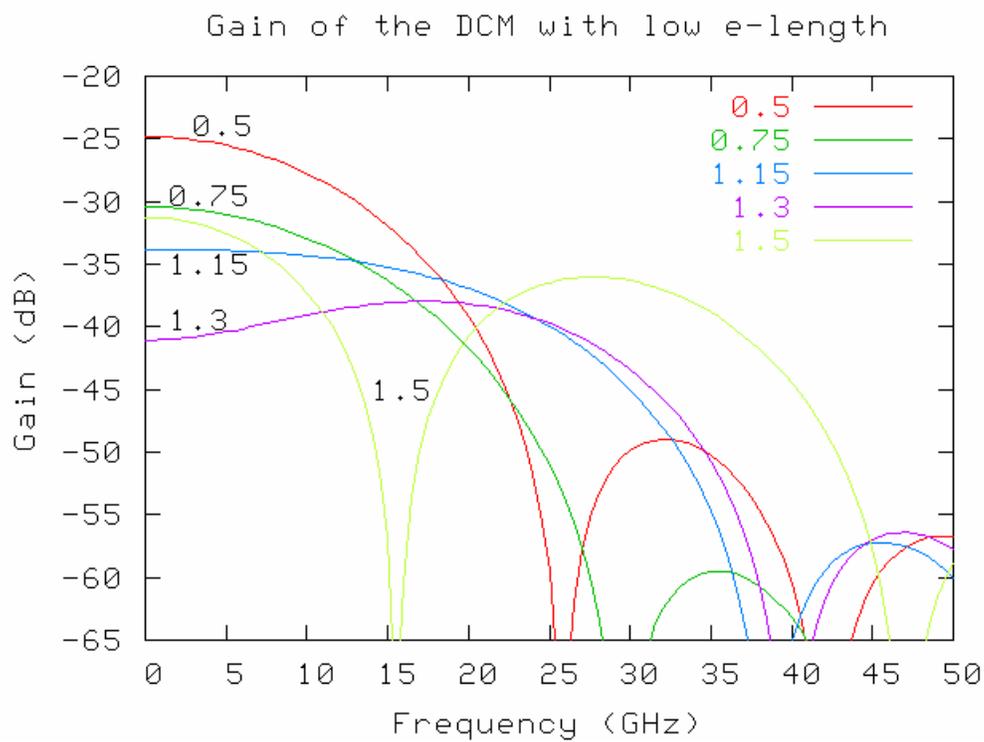


Figure 4 - 27: Gain of the DCM with curves representing multiple electrical lengths between 0.5π and 1.5π .

MODULATOR IMPROVEMENTS

Abstract

Many schemes have been proposed to make high dynamic range analog r-f photonic links by linearizing the transfer function of the link's modulator. However there is not yet a clear winning modulator that is highly robust to velocity mismatch and at the same time easy to manufacture and maintain sensitive linearization parameters. This chapter covers two novel techniques used to enhance modulator performance. First it covers reflective-wave modulators, a broadly applicable technique that effectively gives one modulator rephasing without ever splitting the r-f. Then a technique is described to create an effective r-f split without actually making one, giving a “sensitivity split” useful in making the DPMZM and SoDSMZ.

Reflective Traveling Wave Modulators

Research into high frequency modulation focuses on equalizing the velocity of the electrical and optical waves, as they travel through the lithium niobate modulator structure. One can either speed up the electrical wave or slow down the optical wave. Most research focuses on speeding up the electrical wave. There are two standard methods for this. One matches the inherent velocities as described in Chapter 2. A large buffering layer of a low index insulator (usually SiO_2) is placed between the electrooptic crystal and the electrode, and the electrodes are made with a large vertical height. Both actions lift the fields out of the crystal and reduce the microwave index of refraction, which increases the electrical wave's velocity. This technique has the disadvantage of reducing the field penetration to the crystal, lowering the sensitivity. It is also difficult to grow a large uniform buffering layer. However, velocity matching is useful, and it may also be combined with the method described below. The rephasing concept was presented in the previous chapter, and a different method of re-phasing will be described in the next chapter.

Recall that the rephasing technique allows the electrical wave to “catch up” to the optical wave at discrete points. There is velocity mismatch over each segment, but this time delay is not allowed to accumulate from one segment to the next. The modulator is said to be velocity matched “on the average.” This scheme works as long as each segment is small enough. There are two patents on this work. U.S. Patent 5,076,655. Ref. [2.9], covers the use of an antenna array to drive the segments. The array is phased at the correct reception angle to velocity match on the average. U.S. Patent 5,291,565, Ref. [2.8], describes a coplanar waveguide splitter with different electrical path lengths to achieve the correct average velocity. These schemes have two disadvantages. First, splitting the electrical wave N ways reduces the voltage on the electrode by $1/\sqrt{N}$. Second, the r-f electrode is more complicated than in a traditional modulator, leading to higher losses and difficulty in fabrication. In addition, the antenna-coupled system is inherently a passband system, limited by the bandwidth of the antenna elements and the r-f radiating system.

Current schemes do not try to slow down the optical wave. Creating an optical path delay by putting bends in the optical waveguide would be unacceptably lossy. Waveguides in

electrooptic crystals are weakly-guiding, and they leak light when they have only slight bends. However, reflective modulators can have optical paths that are longer than the electrical paths without introducing bends. This is the basis for the following scheme, which resulted in U.S. Patent 5,886,807; see Ref. [5.1]

Traveling Wave Reflective Modulators

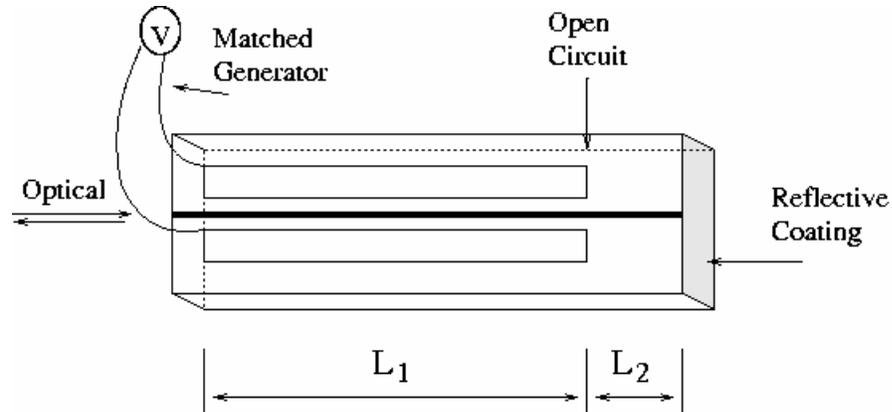


Figure 5 - 1: Phase modulator in reflective traveling wave configuration.

In a reflective modulator, the light enters one side of the crystal, traverses the modulator, encounters a reflective coating at the end of the crystal, and then traverses the modulator backwards, leaving at the same place that it entered. These modulators are normally made with lumped constant electrodes. Figure 5 - 1 shows a phase modulator in the proposed reflective traveling wave configuration. The optical wave and the electrical wave travel together over a distance L_1 . The electrical wave then encounters an open-circuit transmission-line termination. This creates a reflected r-f wave with the same magnitude and phase as the incident r-f wave. The optical wave continues to the end of the crystal where it encounters a reflective coating. The coating reflects the optical wave with unit magnitude and without introducing any phase shift into the optical modulation signal. The optical wave rejoins the electrical wave after an additional distance of $2L_2$. The key is to pick values for L_1 and L_2 such that the waves are re-phased at the beginning of the return trip.

Given L_1 we can solve for L_2

$$\frac{\text{Distance}_1}{\text{Velocity}_1} = \frac{\text{Distance}_2}{\text{Velocity}_2},$$

$$L_1 \frac{n_m}{c} = (2L_2 + L_1) \frac{n_o}{c}, \quad (5.1)$$

$$L_2 = \frac{1}{2} \left(\frac{n_m}{n_o} - 1 \right) L_1.$$

For LiNbO₃, $n_m = 4.0$ and $n_o = 2.2$, thus $L_2 = 0.41 L_1$. The modulator must be 41% longer than the r-f electrode. This extra real estate can be usefully employed. Most modulators have at least one d-c bias electrode. There is nothing wrong with putting the bias electrode between the r-f electrode and the reflective end of the modulator provided there is no capacitive coupling of r-f into the bias electrode.

This modulator is now similar to a phase modulator with length $2L_1$ and two electrode segments with the electrical and optical wave re-phased at the second segment. However, since it does not actually split the r-f power, the electrical wave is $\sqrt{2}$ larger than it would be in the modulator with two physical segments. It uses half the real estate and has twice the gain (+3 dB) of the analogous phase modulator.

This technique may be combined with velocity matching techniques to yield a higher frequency modulator. Velocity matching techniques reduce the inherent velocity dispersion in the crystal. Partial velocity matching may be used on a traveling wave reflective modulator to achieve higher bandwidths. It is often simple to reduce the refractive index difference somewhat, but hard to null it exactly. For example, with great care, velocity matched simple MZMs and phase modulators, Ref. [5.2], have been made with 3 dB bandwidth of 40 GHz. From [5.3], this is not easy to do. In any event, as $n_m \rightarrow n_o$, $L_2 \rightarrow 0$.

There is now the problem of separating optical input and output ports. This will require an optical directional coupler, which introduces 6 dB optical loss, or an optical circulator. The circulator introduces a second problem. To use a circulator, a polarization shift of $\lambda/4$ is placed between the circulator and the modulator. The light from the laser is vertically polarized, it is then circularly polarized in the modulator, and then on the return path after a

second trip through the polarizer, it is horizontally polarized. But circular polarization in the modulator is non-optimal because the TE configuration has three times the electrooptic sensitivity as the TM configuration. However, this is the method that has been successfully demonstrated experimentally in the literature. Reflective wave modulators were introduced in Ref. [5.4]. They have been used in the hunt to reduce the $V\pi$, and the first MZM with a $V\pi < 2V$ was reported in Ref. [5.5] and summarized in Ref. [5.7]. These researchers have also demonstrated a reflective modulator with $V\pi < 0.9 V$, and an intrinsic r-f link with gain; see Ref. [5.6].²⁷ Clearly, the problem of separating the input light and the output light is surmountable. As an interesting confirmation of Ref. [5.3], the velocity matching parameters of [5.5] were $\Delta n = 0.06$, not nearly as good as the velocity matching in Ref. [5.2]. It appears quite reasonable for high frequency applications to use good velocity matching ($\Delta n \leq 0.1$) in combination with the technique described here, instead of requiring excellent velocity matching ($\Delta n \leq 0.02$), a hard to obtain goal.

Application to Useful Modulators

Figure 5 - 2 shows MZM and DCM reflection modulators using this “velocity matching” scheme. The MZM has similar attributes to the phase modulator of Figure 5 - 1. The DCM has an added feature: the designer can choose the electrical length to be either $\pi/2$ or π . If it is the former, then all of the optical power will come out of the cross waveguide. If it is the latter case, then all of the power will come out of the through waveguide. Depending on the system, one or the other may be desirable. In some remote applications, it is desirable to have only one fiber and not to have the laser near the modulator. In this case the designer would choose the MZM or a DCM with the electrical length = π ; however, in other applications of DCMs the designer could choose electrical length = $\pi/2$ which would not require an optical circulator. Note also that the bias electrodes conveniently fit in the space allocated for delaying the optical wave; that is, L_2 . For the simple MZM and DCM, the bias does not need to be separate from the r-f electrode, but making it so simplifies the r-f connection.

²⁷ An “intrinsic link with gain” means that the r-f output/r-f input > 1 , without any r-f amplifier in the path. The source of the gain comes from the combination of a high power laser, a low $V\pi$ modulator, and minimizes the losses in the system. It is a tour de force.

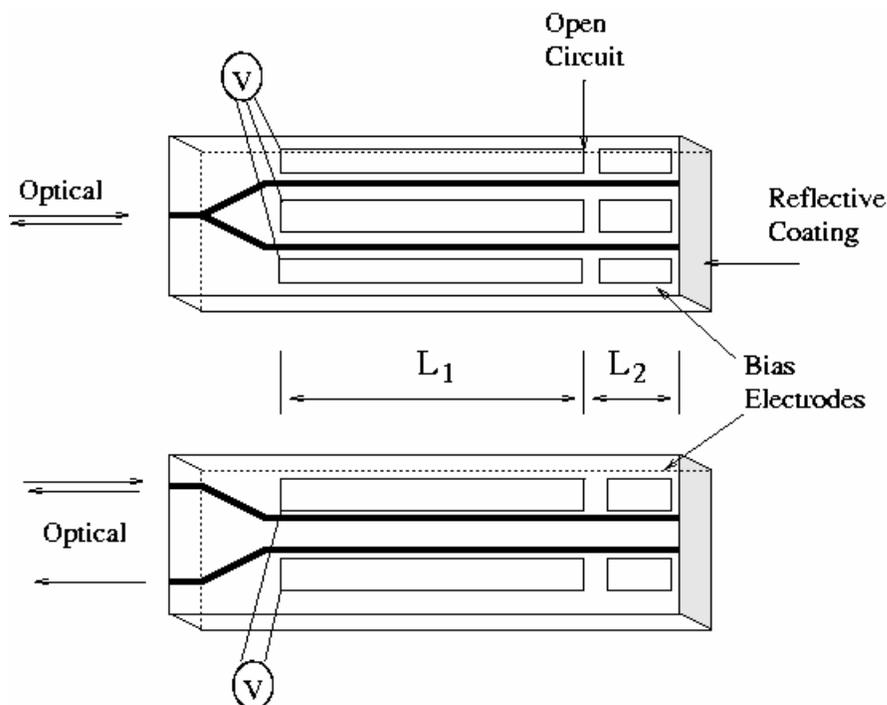


Figure 5 - 2: MZM and DCM reflection modulators, using bias electrodes in the space of L_2 .

The reflective scheme is especially useful in linearized electrooptic modulators. The linearized dual series Mach-Zehnder (DSMZM) is shown in Figure 5 - 3. A DCM2P cannot be built with a reflection modulator because it is not symmetric about the centerline of the crystal. However, one could imagine many configurations of one or two DCMs + passive sections that are symmetric about the centerline. The author speculates that one or more of those configurations is linearizable. The SoDSMZM consists of two Mach-Zehnders in series. This modulator is covered in detail in chapters 3 and 4. It is possible to linearize it with the same bias on both Mach-Zehnders modulators. This is critical for its use as a reflective modulator because the reflection images the second Mach-Zehnder.

Dynamic Range of Reflective Modulators

The re-phased reflected wave is especially useful in linearized electrooptic modulators. These modulators use sensitive cancellation techniques which are disturbed by only a slight amount of velocity mismatch. A linearization scheme may yield an extra 20 dB of dynamic range at

low frequencies, but the fidelity is quickly lost. The reduction of dynamic range from velocity mismatch happens much more quickly than the reduction in gain for most modulators, even though the reduction in gain is normally considered the limiting factor, in the literature.

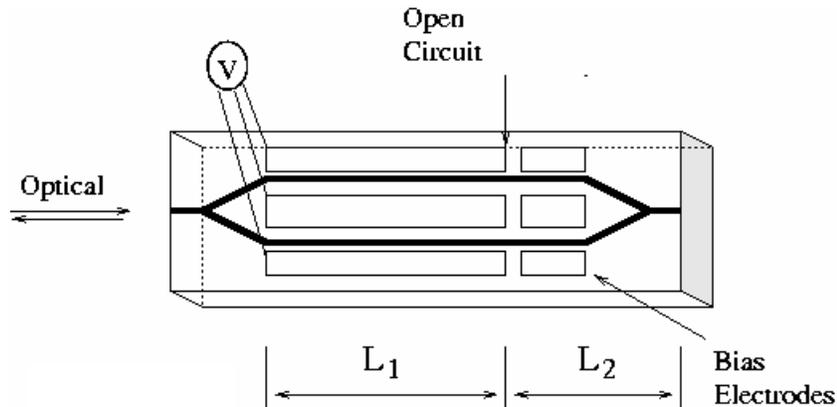


Figure 5 - 3: DSMZM reflection modulators.

The CMZM and SoDSMZM are perhaps the most noteworthy of these cases. Recall Figure 4 - 19 and Figure 4 - 20, from the previous chapter, which show the calculated dynamic range for a typical lithium niobate CMZM and SoDSMZM modulator. When a singly segmented traveling wave electrode structure is used for either of these linearized modulators, there is an initial 20 and 24 dB improvement, respectively, in the dynamic range. But the dynamic range degrades dramatically with increasing frequency due to velocity mismatch. For both modulators, the reduction in dynamic range is a function of the phase error at the second MZM only. When the electrical and optical waves are re-phased once at the second MZM, velocity mismatch does not spoil the enhanced dynamic range at all. The traveling wave reflective modulator has this condition automatically. This modulator is only limited by gain. The 3 dB gain point in the calculation is about 8 GHz for the SoDSMZM. Many linearized modulators were analyzed in Chapter 4, most exhibit a similar rapid roll-off when there is no re-phasing of the electrical and optical waves, but only series CMZM and SoDSMZM are flat with a single re-phasing at the second Mach-Zehnder.

This reflection method enables an efficient implementation of the precision r-f split in the CMZM modulator. The phase reversal method has the drawback that the r-f split is negative, that is, the two modulators are 180 degrees out of phase. By simply using a short circuit instead of an open circuit at the reflection end of the modulator, the proper phase reversal is achieved.

Conclusion

Electrooptic modulators can be made in a reflective traveling wave configuration. The input end of the electrical transmission line is matched to the signal source and the far end is terminated in an open circuit. The optical waveguide extends past the end of the transmission line by a fixed distance which the electrode length and the indices of refraction determine. A reflective coating is put on the end of the crystal. This scheme has three main advantages. First, it uses only half of the real estate of the crystal. Second, there is one re-phasing of the electrical and optical waves without introducing any added geometric complexity over a simple traveling wave structure. And third, the re-phasing comes without splitting the r-f power, which results in a 3 dB gain over a modulator with two separate electrode segments. Higher bandwidth performance may be attained by combining this scheme with some amount of conventional velocity matching. When applied to linearized modulators, specifically the CMZM and the SoDSMZM, the inherent re-phasing can preserve the enhanced dynamic range. Without the re-phasing, there would be a rapid degradation in dynamic range caused by the velocity mismatch, which is more rapid than the degradation in gain, the original reason researchers started looking into reflective-wave modulators.

Sensitivity Splits

It was argued in Chapter 3 that it is more important to evaluate modulators by the number and type of linearization parameters than only by their peak dynamic range. Bias voltage parameters are more easily controlled than coupler electrical lengths and optical and r-f levels. From the d-c analyses in that chapter, it was concluded that the best linearization schemes use exclusively voltage parameters. In Chapter 4, this conclusion is qualified by the observation that modulation schemes have vastly different responses to velocity mismatch. The dual parallel Mach-Zehnder modulator and, to a lesser degree the Cascade Mach-Zehnder modulator, become attractive. Furthermore, though the sub-octave dual series Mach-Zehnder modulator is assumed not to have an r-f split, in the literature a lower noise figure, and higher dynamic range, SoDSMZM can be achieved by splitting the r-f. In this section a method for achieving a robust r-f split is described.

Sensitivity Split in the Dual Parallel Mach-Zehnder

The DPMZM has two linearization parameters: an r-f and an optical split. These parameters must achieve a tolerance given in Table 3-3, which is prohibitively challenging with the normal implementations of the splits: an unbalanced Y-junction for the optical split and a microwave directional coupler for the r-f split. Even if adequate accuracy were attained in the microwave split at one frequency, it is also challenging to maintain that accuracy over the microwave signal bandwidth. Figure 5 - 4 and Figure 5 - 5 show properties that can be exploited to achieve a practical implementation of the DPMZM modulation scheme.

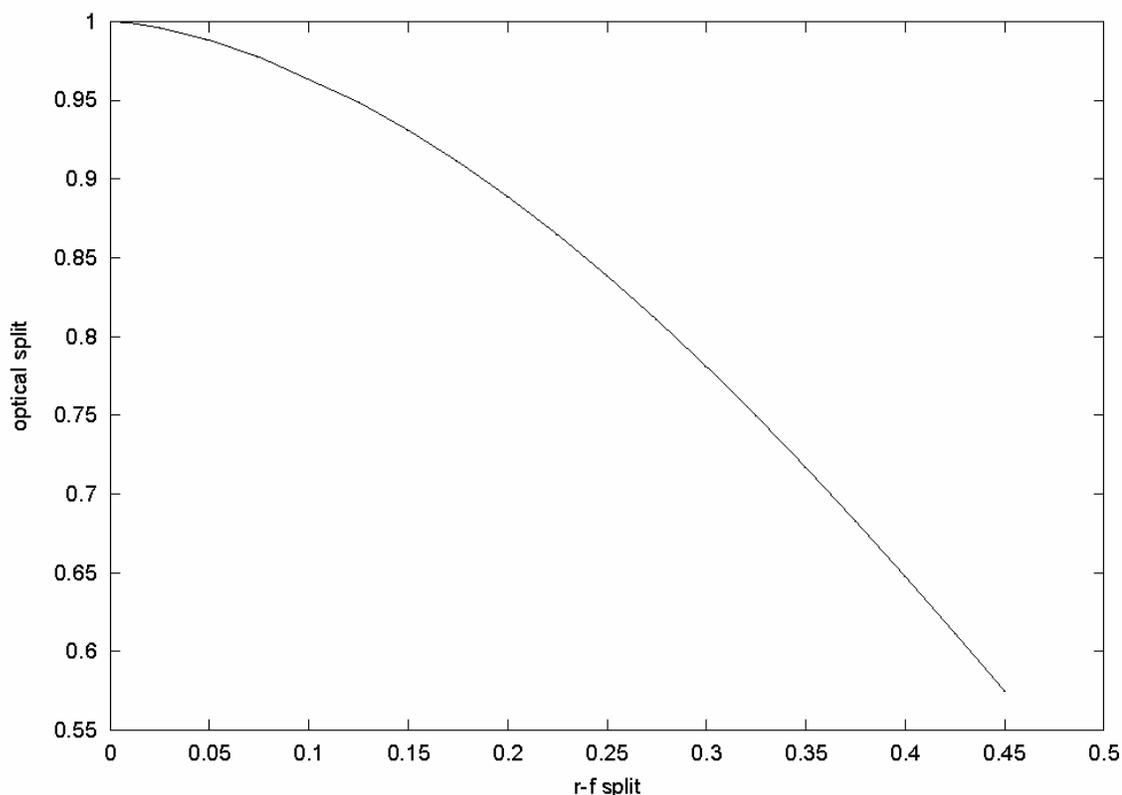


Figure 5 - 4: The optimal optical split as a function of the RF split.

The original publication of the DPMZM in Ref. [3.3] and subsequent summary publications such as Ref. [3.4] refer to either one optimal r-f and optical split, or a set of distinct optimal points. However, only one degree of freedom is required to achieve the linearization. The basic MZMs in the linearization scheme are biased at their point of anti-symmetry, generating no even-order harmonic distortion. The two precise signal levels are only used in nulling the third-order intermodulation distortion term. There is a redundancy in the degrees of freedom of the modulator. Figure 5 - 4 demonstrates this by plotting that optimal optical split as a function of the r-f split. This forms a curve of points that correspond to a local optimum in dynamic range. Holding either r-f or the optical split constant, and optimizing the other split leads to a point on this curve. Figure 5 - 5 shows the value of the optimum versus the r-f split. There is a global optimum, however, the curve is nearly flat over a wide range. If the r-f split is

imprecise but at least uneven, the linearization mechanism still works given precise control of the optical split. And this is the idea behind a sensitivity split.

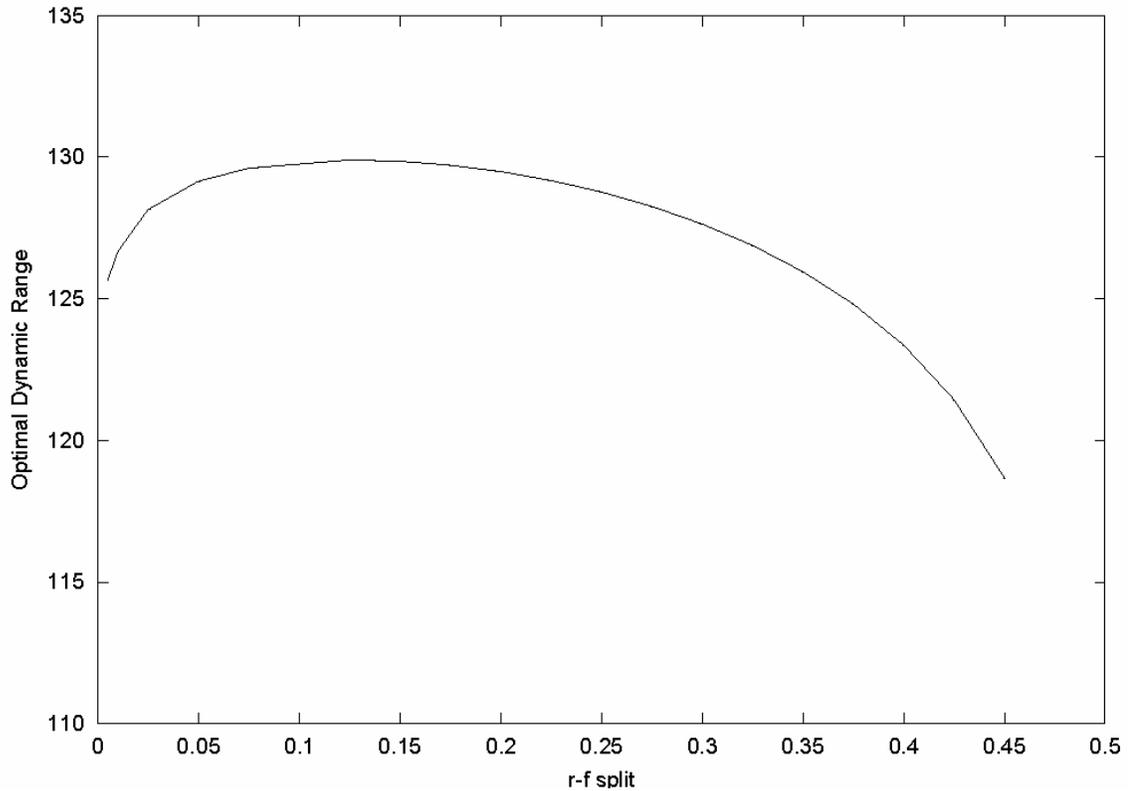


Figure 5 - 5: The optimal dynamic range as a function of the RF split of the DPMZ; at every value of the r-f split the optical split has been re-optimized

In the implementation of a DPMZM, the r-f path consists of one transmission line and the r-f power is never split. Figure 5 - 6 shows such a modulator schematically. 1) shows a top-down view of the optical waveguides of the modulator, 2) shows the waveguides again, but with a push-pull electrode structure on top of them, 3) and 4) show end views of the waveguides and electrodes for the two different Mach-Zehnders labeled B and C. Circles represent the waveguides and rectangles show the electrodes above them. The key difference between 3) and 4) is that in 3) the waveguides of MZM B are centered in the region of the crystal that has the strongest electric fields, while in 4) the waveguides of MZM C are off-center just enough to build in a distinct sensitivity split between the two Mach-Zehnder modulators. That is, even

though there is one uniform electrode, the effect of that electrode on the modulation in the two Mach-Zehnders is different.

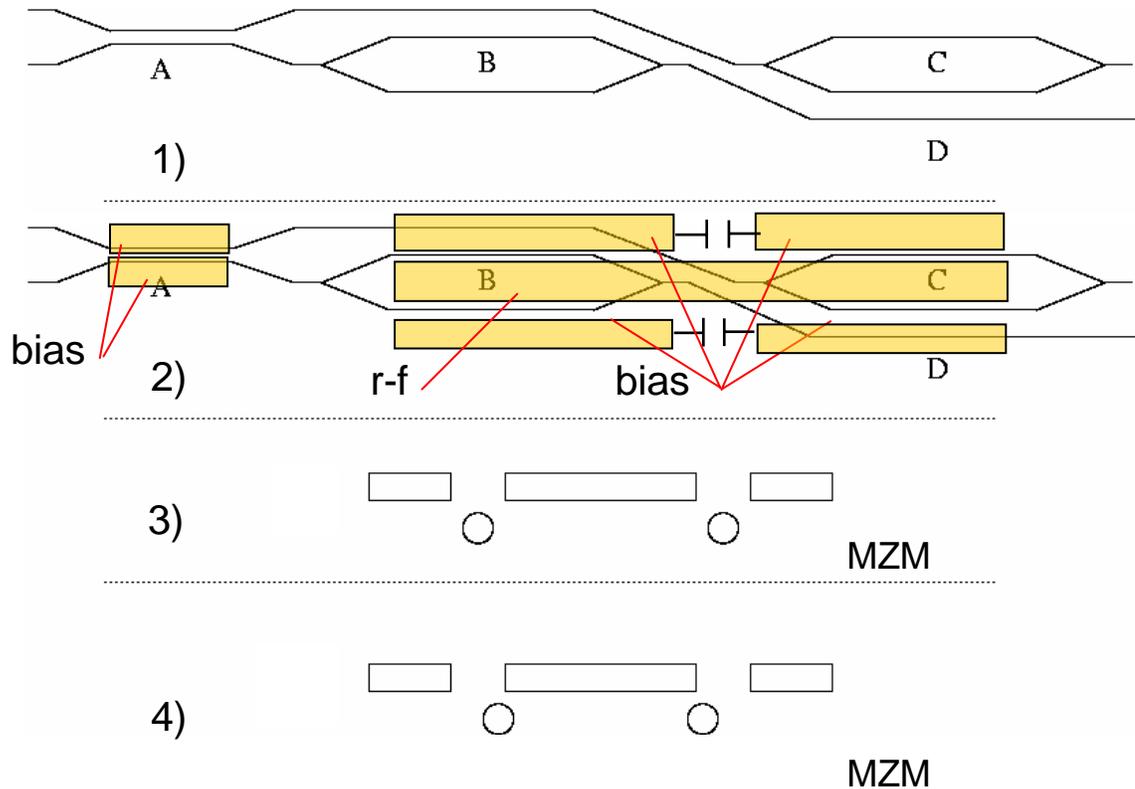


Figure 5 - 6: A DPMZ with a precise optical split and an imprecise “sensitivity” split implemented by the structure of the waveguides, so the RF is never physically split.

Block A is a passive directional coupler with DC electrodes only. The optical split between the path to the MZMs B and C is set approximately by the choice of coupler length and the applied voltage is used only for fine tuning. The coupler controls the optical output to the precision of the bias voltage, which may be set after fabrication and even dynamically during operation. The directional coupler introduces a phase shift between the output and the photodetector. If two distinct fibers, two distinct laser wavelengths, or two distinct polarization states are used, then this phase shift is irrelevant because the currents of the two signals are summed in the photodetector incoherently. However, one proposal in the literature, Ref. [2.3], uses one laser, one polarization, and one fiber, and combines the two

MZM outputs in a passive directional coupler. This configuration requires an additional phase shifter, used to keep the two signals in phase quadrature so that they don't interfere. If the single output fiber approach is used with the sensitivity split shown here, then the phase change introduced by the coupler section A will need to be counteracted in addition to the precision phase shift to guarantee that the two signals are in quadrature. This could be located at block D. Or alternatively, the directional coupler at block A could consist of two bias electrodes, which with different voltages is enough to give a precise power split and phase split, thus eliminating a need for a separate phase shift region at block D.

MZMs B and C appear in series spatially, but they are parallel, each modulator is fed from a different optical path. The r-f flows through one continuous transmission line element that passes over both B and C. This creates a time delay between the two modulation signals, but the path length from the input to the output of this chip is the same for both the optical signal that traverses modulator B and the one that traverses C. While it is beneficial to avoid physically splitting the r-f, the total roll-off, governed by the entire length of the traveling wave electrode structure, is now twice as long as the individual MZMs. For frequency regimes in which microwave loss is significant, other approaches (such as the antenna segmented approach discussed in Chapter 2) may need to be combined with this approach.

Figure 5 - 5 shows that the r-f power split must be uneven. The schematic diagrams B) and C) from Figure 5 - 6 show how this unevenness is obtained. The waveguides in MZM C are intentionally placed non-optimally so that the sensitivity of the MZM is reduced because the modulation field within the optical waveguide is no longer optimally aligned with the direction for maximum modulation of the polarized optical wave in the x-cut crystal. If the waveguides are not directly centered in the waveguide gap, the sensitivity will be reduced, and then the two modulators will have different sensitivities, creating an effective r-f split, even though the r-f power has never been split. Obviously this positioning cannot be manufactured to yield a precise split. However any split between 1:20 and 1:4 yields near optimal dynamic range provided the optical split is precise and configurable after fabrication. The end result is a realizable modulator for which the dynamic range performance is robust to frequency.

Sensitivity Split in other Mach-Zehnder based linearized modulators

Sensitivity splits are easily employed in the cascade and series Mach-Zehnder modulators. Since the two MZMs are in series, there is no need to alter the layout of the MZMs. This technique could be used to construct a CMZM that requires an r-f split, or it could be used to construct a SoDSMZM in which an r-f split was used to get to a lower total noise figure, and thus better dynamic range, than what can be achieved with the SoDSMZM analyzed in Chapter 3 and Chapter 4 (that has no r-f split). In fact, in Ref. [2.3], G. Betts indicates that this r-f split does not need to be precise.

Figure 5 - 7 shows the sensitivity split concept applied to a SoDSMZM. In this modulator there is one optical path that traverses both Mach-Zehnders. The r-f is split by the off-alignment of the waveguides as depicted in schematic 3) versus 2). This effective r-f split need not be a precision split. However, different biases must be applied to the Mach-Zehnders, unlike the version of the DSMZM studied in Chapter 3, Chapter 4, and in the reflective-wave modulators of this chapter. The different biases are implemented by segmenting the bias electrode as shown in schematic 1). It is not necessary to segment the r-f electrode.

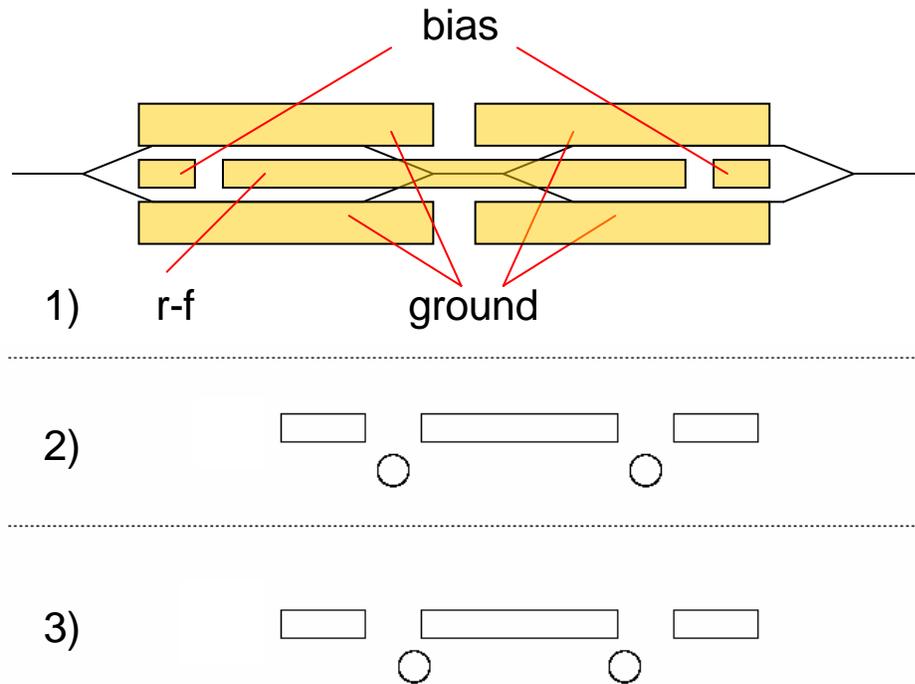


Figure 5 - 7: The waveguide an electrode structure of a sensitivity split to implement a SoDSMZM in which the two separate drive strength is applied to each Mach-Zehnder to reduce the noise figure of the modulator.

MILLIMETER-WAVE DIRECTIONAL COUPLER MODULATORS

Abstract

The laboratory work of this thesis demonstrates the 94 GHz directional coupler modulator for the first time, and makes improvements in the efficiency of the components that go into the link experiment. The work picked up where Finbar Sheehy, a previous Ph.D. student of Prof. Bridges left off, Ref [1.3]. An already built, but non-functional, DCM chip was analyzed by the author in a Caltech SURF project and then in a Caltech Senior Thesis. This chip was used to successfully demonstrate operation at 94 GHz DCM. Another DCM chip was built with the same design, and an entirely new DCM was designed and built. The r-f feed system was redesigned and built. New carrier components to house the modulator were built, and the experiment was moved from all free space components to a mixture of fiber and free space components. New measurements were taken with the complete system, but successful operation was not obtained. The ultimate reason is speculated to be a processing issue in the fabrication of the modulator chip. The project was abandoned because new chips were no longer obtainable from our source, the Hughes Research Laboratories.

An Introductory Note

The research described in this chapter is chronologically the first undertaken in the author's graduate studies. It was an extension of a Summer Undergraduate Research Fellowship (SURF) project in the summer of 1993 and a senior thesis in Electrical Engineering in the academic year 1993 - 94. This extension was almost purely experimental, and was an attempt to improve the antenna-coupled directional coupler modulator successfully demonstrated during the author's senior year Refs. [6.1] and [6.2]. Many improvements were incorporated into a new modulator chip and millimeter-wave excitation structure during the years 1994 - 1999, and these are presented here for the record. But an overall demonstration of the improved modulator was not realized, and the project had to be abandoned due to events largely outside the author's control, as described below.

Background Work up to the Fall of 1994

Directional coupler modulators (DCM) suffer a roll-off in sensitivity with increasing frequency due to velocity mismatch, just as Mach-Zehnder modulators (MZM) do, as previously described in Chapter 4. However, the -3 dB frequency in modulation sensitivity is about 40% higher for a simple DCM than a simple MZM of the same length and velocity mismatch, thus making it an attractive candidate for high frequency applications. But a *linearized* modulator based on a DCM is much more adversely affected by velocity mismatch than a *linearized* modulator based on a MZM, as previously discussed in Chapter 4. Thus it was important to see if the antenna-coupling scheme to counter the adverse effects of velocity mismatch would work as well with DCMs as with MZMs. This demonstration was initiated by Finbar Sheehy, a previous graduate student in the group, who designed a DCM with an array of simple dipole antennas on the surface of z-cut lithium niobate as part of his Ph.D. research. The design is given in Sheehy's thesis and the supporting contract final report, Refs. [6.3] and [6.4].

The mask used to make Sheehy's modulator is shown in Figure 6 - 1. There were 12 modulators on each 3" diameter wafer. A string of millimeter-wave dipoles connected to short modulating electrode sections occupies the center section of each modulator. At either end of this string of dipoles there is a short d-c bias electrode section to bias the modulator to the

proper operating point. These separate bias sections were used in lieu of making d-c connections to the dipoles themselves.

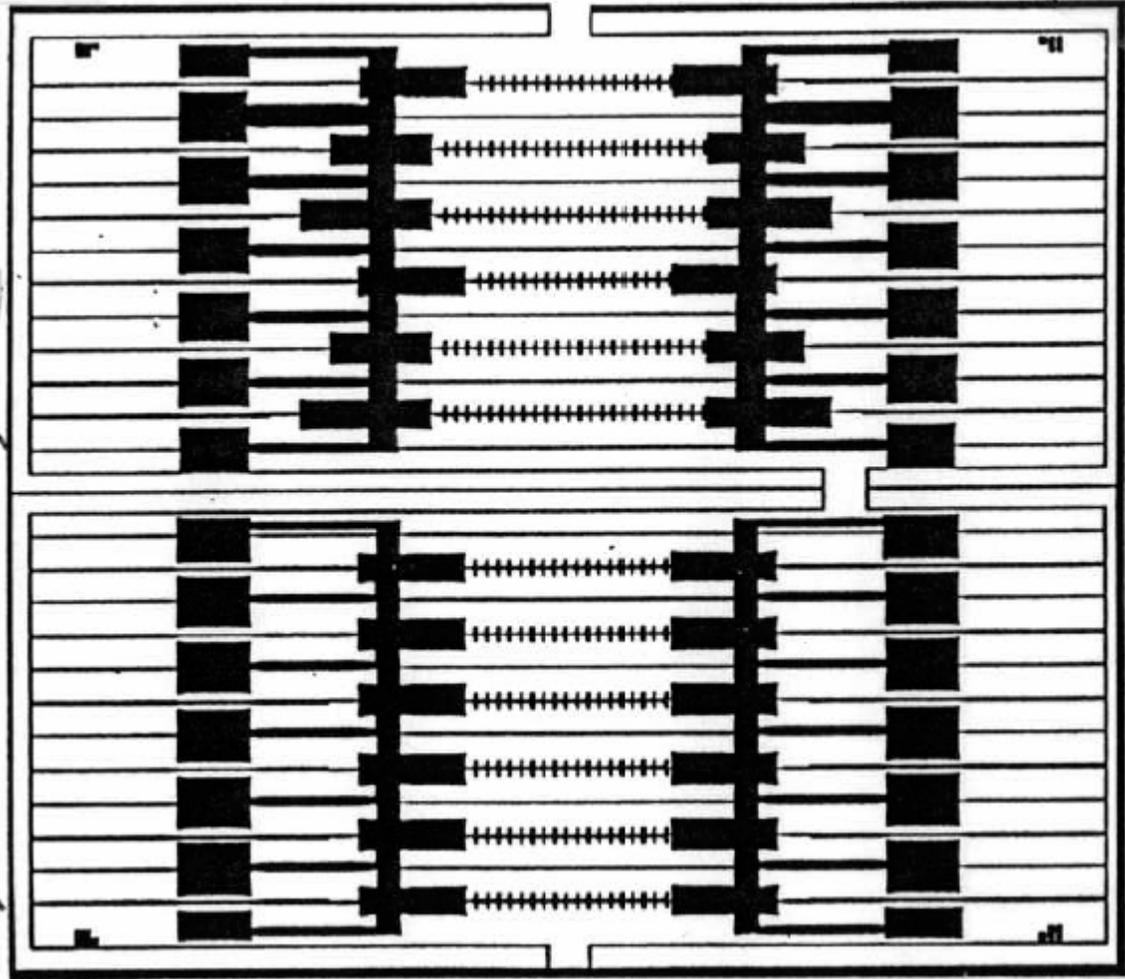


Figure 6 - 1 Actual electrode mask used in FTS-1 design, to make 12 modulators on a 3" diameter LiNbO_3 wafer. The gaps between electrode pairs do not show at this scale. The particular modulator used in the transfer function measurements is the fourth from top edge. This mask was designed by Finbar Sheehy.

Sheehy's modulator design was fabricated at Hughes Research Laboratories (now HRL, LLC) under an informal arrangement to assist the work at Caltech. Sheehy set up the DCM chip in the laboratory and attempted to make it work. While some asymmetrical modulation

sidebands could be observed when the modulator was operated at 94 GHz, he could make no sense of the transfer function of the modulator. This was the situation when he finished his Ph.D. work, graduating in 1993.

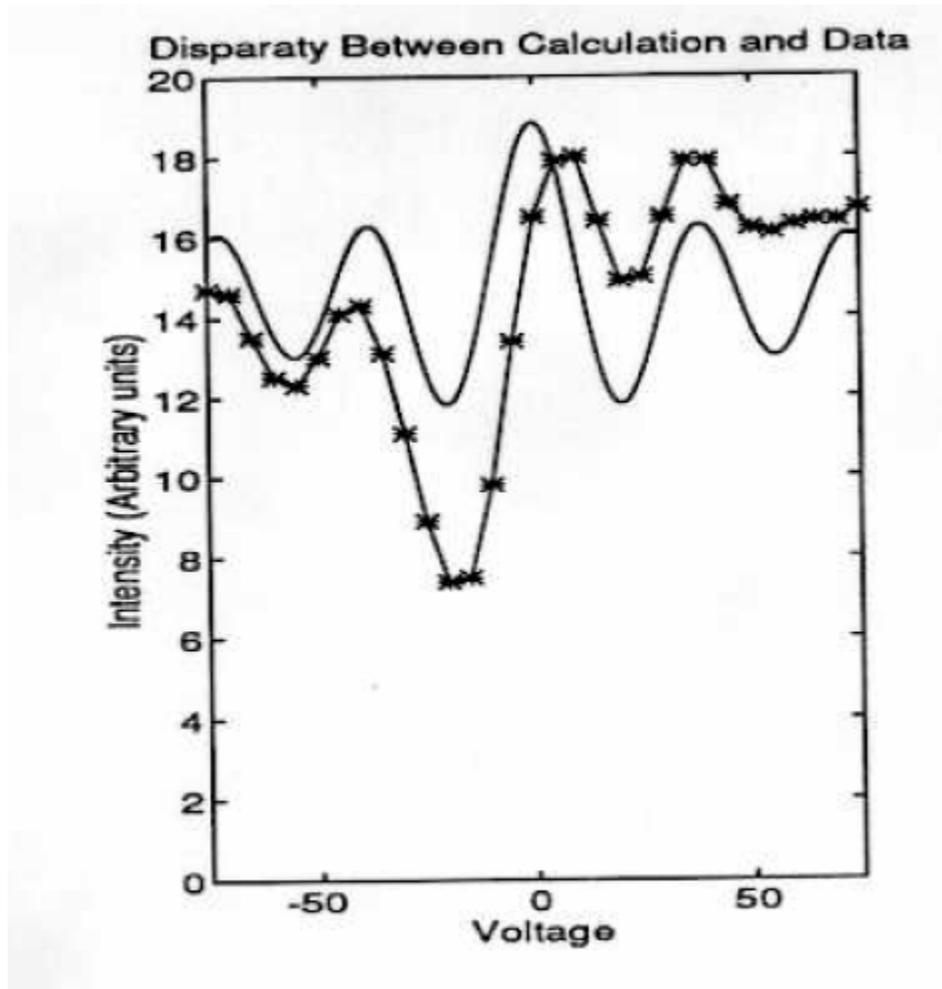


Figure 6 - 2: D-c transfer function of Sheehy's DCM. The curve with data points is the experimental measurement and the curve without data points is the best-fit for the theory for a modulator of the design length, assuming perfect directional coupler symmetry.

The problem of understanding what was happening with this modulator was given to the author, then an undergraduate student, as a Summer Undergraduate Research Fellowship (SURF) summer project. Upon measuring the d-c transfer of the bias electrodes, it became clear that there must have been some asymmetry in the underlying optical directional coupler.

Figure 6 - 2 compares a sample theoretical bias electrode curve for a symmetrical directional coupler modulator (solid curve) with the measured transfer curve. Note that the theoretical curve is an even function around zero bias. Since all such curves must be even functions, it is clear that no set of modulator parameters could result in a fit to the experimental data. However, assuming that the underlying optical directional coupler has a built-in asymmetry in the $\Delta\beta_s$, the voltage induced optical coupling constants, then a good fit to the experimental data could be obtained, as illustrated in Figure 6 - 3. Although the origin of the built-in asymmetry was not understood at the time, measuring its value allowed the modulator to be biased at the correct operating point

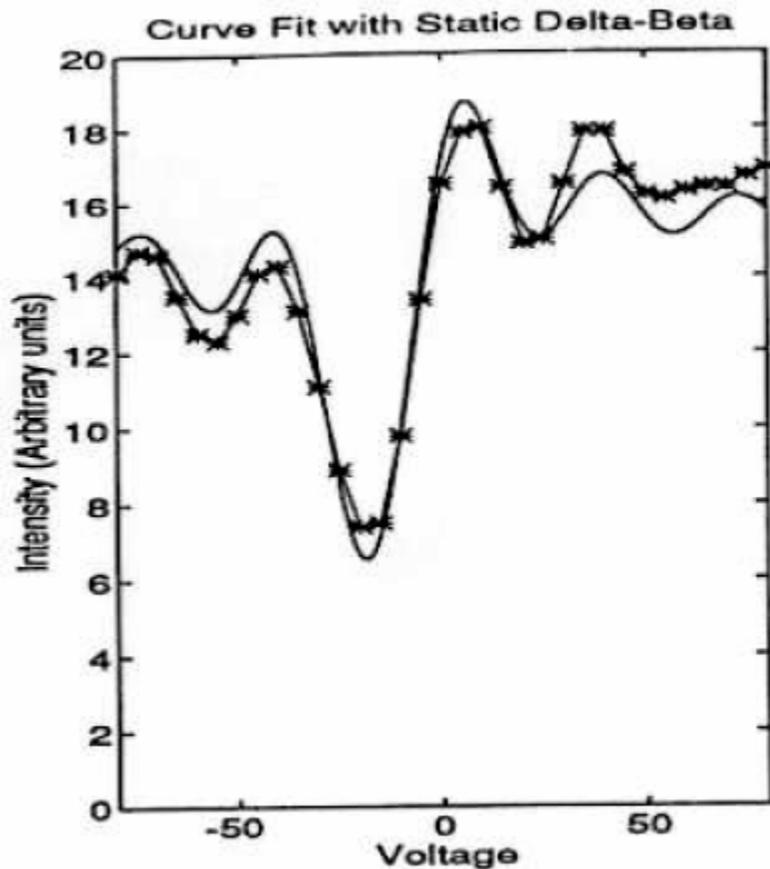


Figure 6 - 3: (a) Curve fit to experimental data assuming a static asymmetry $\Delta\beta_s$ of 1.7 radian/meter, fitting data of $-70 \text{ V} < V < +70 \text{ V}$. (b), the same as (a), but using data $-80 \text{ V} < V < +20 \text{ V}$.

The asymmetric transfer function is modeled based on Eqns. (3.12) and (3.14). The cross term m_{12} is rewritten in terms of the physical parameters, $\Delta\beta$ and κ , and $\Delta\beta$ now consists of a static term and an electrooptically induced term, $\Delta\beta = \Delta\beta_s + \Delta\beta_e$. Taking the magnitude squared of the complex amplitude to get the intensity

$$I_{out} = \frac{I_0}{1 + \left(\frac{\Delta\beta_s + \Delta\beta_e}{\kappa}\right)^2} \sin^2 \left(\theta \sqrt{1 + \left(\frac{\Delta\beta_s + \Delta\beta_e}{\kappa}\right)^2} \right). \quad (6.1)$$

Demonstrating this modulator at 94 GHz was undertaken as a senior thesis project during the academic year 1993 - 94. The experimental set up was the same as that used by Sheehy, shown in conceptually in Figure 2 - 9 schematically and described later in Figure 6 - 10.

The modulation depth was deduced by taking the ratio of the sideband amplitude to the carrier amplitude displayed on the scanning Fabry-Perot output. The millimeter-wave power was measured with a calibrated W-band bolometer head. The resulting performance of 0.2 [m²/Watt] was reported at the PSAA-IV conference in 1994 [Ref. 6.1]. This modulation sensitivity corresponds to a V_π of about 40 volts for a conventional 50 ohm modulator, which is not all that good for a conventional low-frequency, 9 mm active length modulator, but it was good enough to be interesting for a 94 GHz modulator (not commercially available in 1994, nor now). Unfortunately, this particular modulator chip also had a very bad optical transmission, about -20 dB optical loss.

Figure 6 - 4 shows the actual detected side bands from a picture of the oscilloscope taken during the experimental demonstration. There are two sidebands in view, and they are indicated by hand-drawn arrows shown in the scan of the Polaroid. The measurement technique (using a Fabry-Perot interferometer) is not the most sensitive, and it is difficult to see the sidebands above the skirt of the optical carrier. However the resulting modulation is good considering the frequency of operation. From these results, an important next step is improving the efficiency of the system to demonstrate a stronger modulation signal.

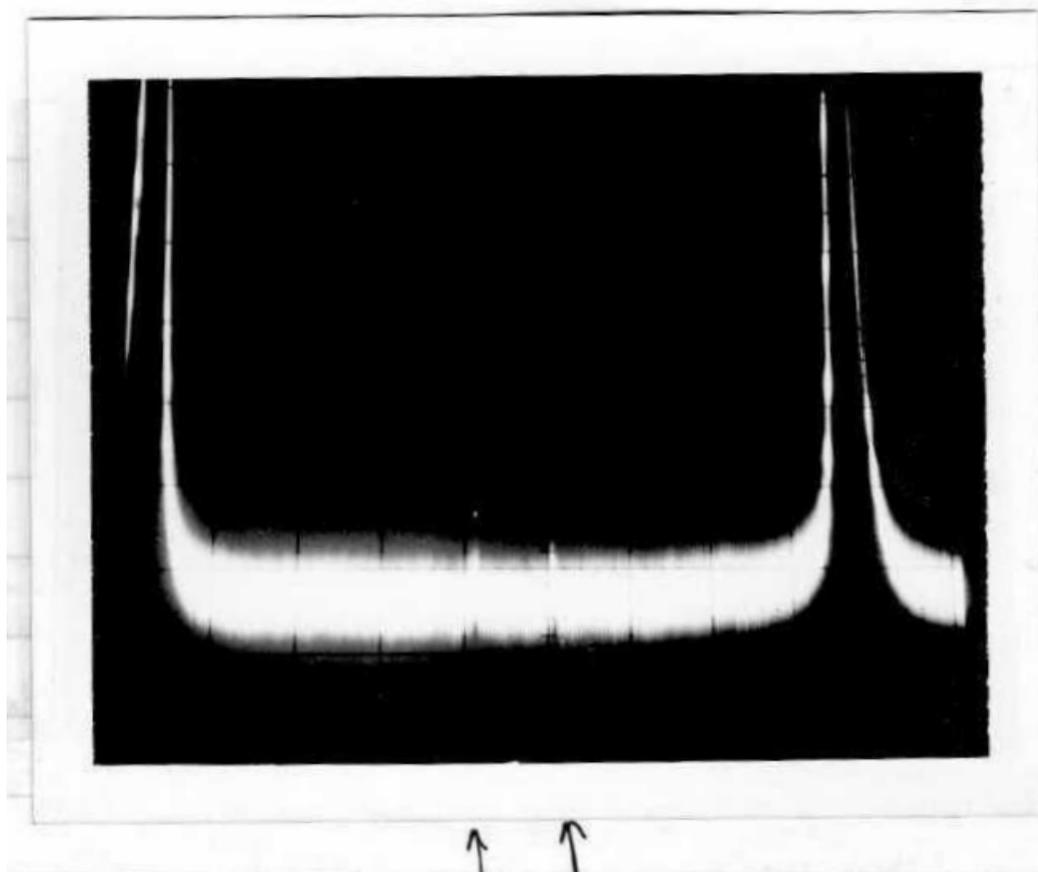


Figure 6 - 4: Oscilloscope Polaroid of the operation of directional coupler antenna-coupled modulator at 94 GHz. The hand drawn arrows point to the two side bands that are just distinguishable from the skirt

The origin of the built-in $\Delta\beta$ or the high optical loss was unknown at the time the modulator was successfully demonstrated. However, after the PSAA meeting, while photocopying the waveguide masks used by Hughes to fabricate the modulator, “mouse-bite” defects were discovered on the mask: Chips of the mask metal had broken off along the edges of the waveguide regions. These chips resembled a “bite,” as if a microscopic mouse had bit into the metal coating, hence the name. Such defects are known in the mask-making industry, and are attributed to various causes. Two such causes are (1) a dirty glass substrate causing adhesion problems when the metal is deposited, and (2) electrostatic discharge when the mask is used over a lithium niobate (pyroelectric) wafer. It turns out that Hughes was experiencing “mouse

bite” defects on a number of masks just at the time they were discovered on the particular waveguide mask used in Sheehy’s DCM. There was an ongoing argument between Hughes and the mask maker, Align-Rite, with Hughes citing cause (1) when the mask was made at Align-Rite, and Align-Rite asserting cause (2) when the mask was used at Hughes. This argument was never settled, but for the Sheehy modulators, the damage was already done. Upon close inspection of the modulator under a microscope, it was verified that the “mouse-bite” defects in the mask had been faithfully transferred to Sheehy’s modulator waveguides, resulting in a significant “bulge” in the waveguide width at every defect. A photograph of these defects, along with a schematic drawing to make them clearer is shown in Figure 6 - 5.

This was the explanation for both the deduced $\Delta\beta_s$ and the high optical loss: The two waveguides clearly had different defects and thus different $\Delta\beta_s$, and each defect scattered light out of the waveguide into the surrounding substrate (excess loss).

Modulator Fabrication Work Beginning in the Fall of 1994

A new directional coupler waveguide mask, identical to the original one designed by Sheehy was fabricated on z-cut lithium niobate (but without the “mouse bites”), and received in mid 1994. New modulator chips were fabricated with this mask by Hughes and delivered to Caltech. These chips exhibited reasonable optical loss (-8 dB), and the transfer functions of the bias sections were now symmetrical about zero bias (as a DCM with $\Delta\beta = 0$ should be). However, the transfer functions exhibited very little change in transmission with voltage applied to the bias sections. After some investigation, it turned out that the original directional coupler design parameters determined by Dr. Cel Gaeta at Hughes Research Labs (unpublished) were incorrect. Following Gaeta’s data, Sheehy had used the wrong dimensions in his directional coupler modulator design. The waveguides in the coupler were too far apart, resulting in too little coupling. The bias sections of the coupler could not adjust the modulator to the proper condition for the antenna-coupled sections to work. In fact, the only reason that the results reported in Ref. [6.1] were obtained was due to the static $\Delta\beta_s$ and the light scattering introduced by the mouse bites! The new modulator, without the mouse bites, was too insensitive to bias properly.

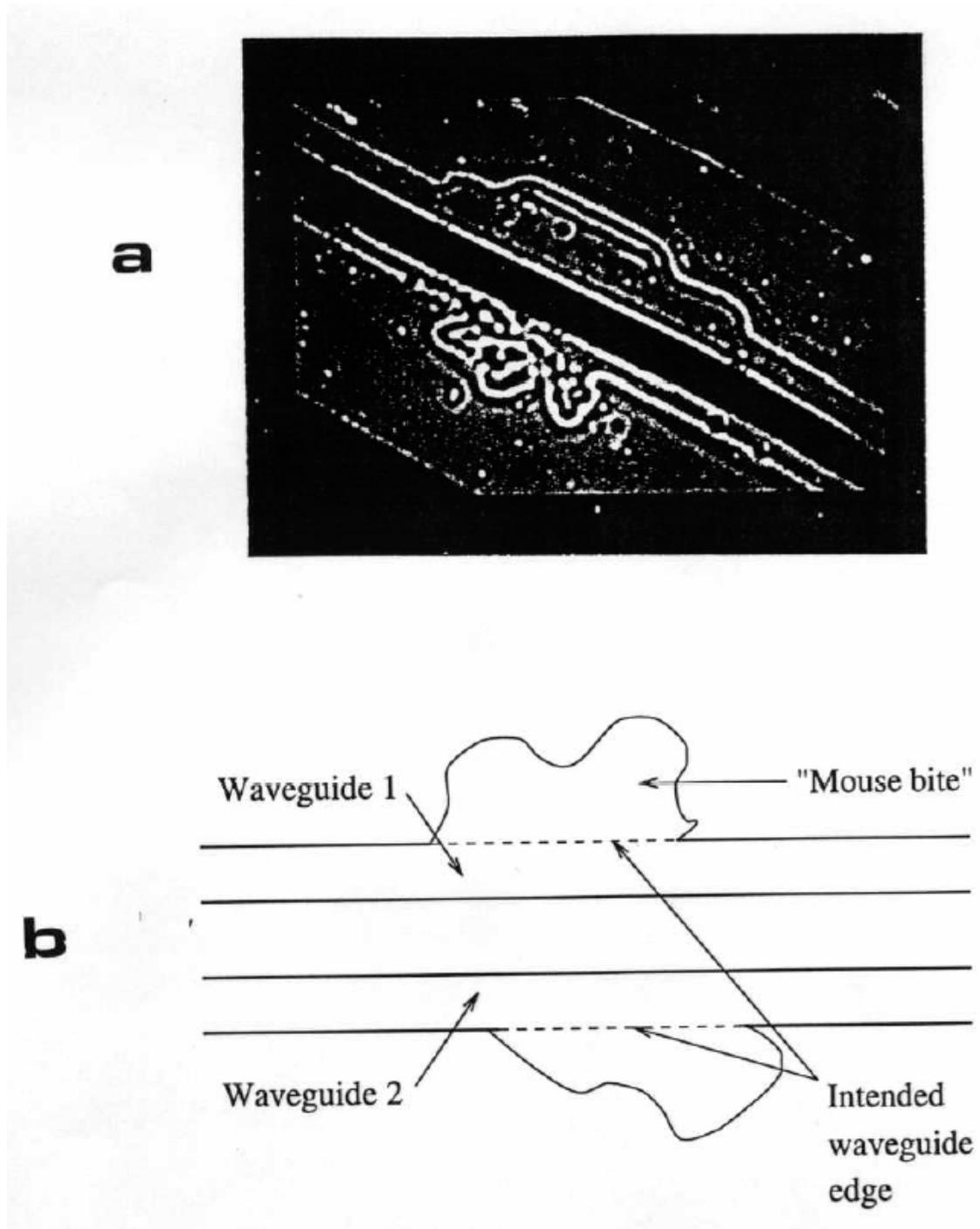


Figure 6 - 5: "Mouse-bite" defects (a) on LINC-1 directional coupler waveguide mask used to make Sheehy's DCM, and (b) a schematic representation of these defects.

Fortunately, this error had also been discovered at Hughes, and Gaeta was already in the process of making new measurements of the directional coupler properties. Using Gaeta's new directional coupler parameters, a new waveguide and electrode masks were designed. The new electrode mask layout (Denoted UVC-el-1) is shown in Figure 6 - 6. Eleven different antenna-coupled DCMs are on the mask, along with some simple directional coupler test sections. The modulators have various combinations of antenna sections and d-c bias sections (heavy lines). The gap between the bias electrodes is too small ($6\ \mu\text{m}$) to be resolved on this figure. The masks were fabricated by Align-Rite, but without the mouse bites, and sent to Hughes Research Laboratories for modulator fabrication.

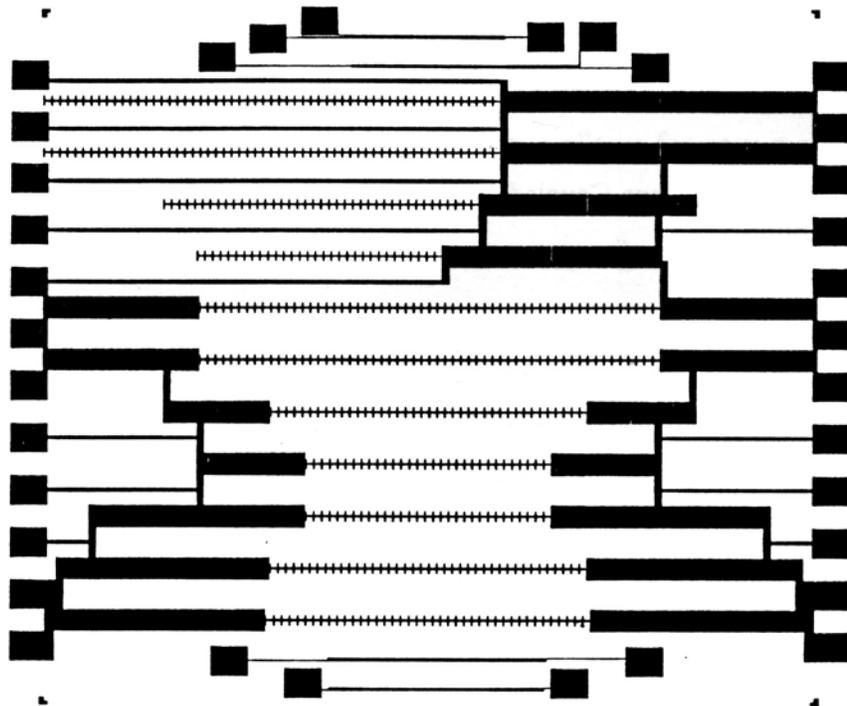


Figure 6 - 6: Electrode mask UVC-el-1 for 11 different antenna-coupled directional coupler modulators, Ref. [6.3]

The antenna-plus-transmission-line-segments are arrayed in the center of each modulator. The large black rectangles at the ends are the bias sections, which are connected to contact pads at the right and left edges of the mask. Various combinations of bias electrode length and

position, and antenna array length and position were included. The shorter horizontal traces at the top and bottom of the pattern are simple transmission lines for conventional DCMs.

94 GHZ Directional Coupler Modulator Fabrication

Work on this new modulator began early in 1996, when we received new contractual funding from Rome Laboratories after an 18 month gap in funding. HRL had fabricated the directional coupler titanium in-diffused waveguides and applied a silicon dioxide buffer layer. This wafer was returned to Caltech in May to have the electrodes applied. Since this was the author's first experience with making the electrodes, and the Hughes process could not be duplicated, a number of preliminary experiments were followed to gain some experience with the procedures. Prof. Axel Scherer generously agreed to help with his laboratory facilities for this processing. Prof. Scherer uses negative photo resist while HRL uses positive resist and a "lift off" gold process (which is why the mask needed to be copied and reversed, the process by which the "mouse bites" were originally discovered). Tests of spinning-on and evaluating negative resist on glass slides were followed to build up an understanding for the characteristics of negative resist. Good results could be obtained, but that the exposure time and level were much more critical than with the positive resist that had previously been used (at Caltech and HRL).

Using negative resist, a uniform gold layer is evaporated over the entire surface of the lithium niobate wafer. Then the photo resist is spun on, dried, and exposed with the electrode mask. The photo resist is developed and washed away everywhere except over the gold where the modulator electrode pattern is desired. Then the rest of the gold is etched away in the exposed regions with a reactive ion etching (RIE) process.

The use of negative photo resist created some difficulty in the fabrication process that may be relevant to the ultimate unsuccessful results obtained for this modulator. With the negative resist process, the edges of the metallic regions were less well-defined than when positive resist was used. This became a concern in the region between the two electrodes forming the

traveling wave transmission line and the region between the halves of the bias electrodes. These gaps are only a few microns. After the first etching step, some conductivity between these electrodes was measured, indicating there was still some gold bridging the gap. This required spinning on more resist, exposing the mask again (with the difficulties of realigning it to exactly the same position, developing, and continuing the etching process. The consequence of over-etching would be the removal of the traveling wave electrodes themselves, so it was important to get the exact total etching time correct. The time was judged to be correct if there were high d-c resistance between the two electrodes. However, the resulting structure likely had continuously varying thickness “filets” across the gap. This may have greatly upset the distribution of transverse electric field between the electrodes.

Since a simple evaporated gold coating without a chromium underlayer was used (in order to allow the gold to be more easily etched off), the result was a very fragile gold electrode pattern. To protect this fragile layer, a silicon dioxide overcoating was applied. The modulator substrates were returned to HRL for this coating, since HRL was willing to continue helping the experiment with that processing step, and the author was not yet experienced in using the silicon dioxide sputtering apparatus in Prof. Scherer’s lab. When the wafer was received back from HRL, another photo resist layer was spun on and exposed with a mask that gave electrical access to the bias pads in the underlying gold electrode layer. Using this mask, contact holes were etched through the silicon dioxide protective layer. The modulators on the wafer were checked for dc short circuits. None were found. The wafer was returned to HRL to be sawed into two 1” by 2” substrates, which were then polish-cut on the ends.

A mechanical support structure for the modulator substrate was designed and given to Caltech’s central machine shop for fabrication, as shown in Figure 6 - 7. The substrate is mounted on a carrier oriented in a vertical plane. Both sides of the substrate are accessible so that the d-c bias voltages may be applied with probes on the front side and the millimeter-wave illumination applied with a slab waveguide from the back side. A vertical steel plate surrounds the carrier so that the magnetic-based probe holders may be positioned on it. A close-up of the modulator mounted on its carrier is shown in Figure 6 - 8, and a further enlargement is shown in Figure 6 - 9. In this figure, the dipole-plus-transmission-line-segments are seen more clearly against the end of the lithium niobate slab waveguide

contacting the modulator chip from behind (white rectangle). The large rectangles at either end of the dipole array are the bias electrodes. (The gap in between the bias electrodes and the transmission line electrodes is too small to be seen, even in this blow-up.)

Optical transmission measurements were initially made at 633 nm, and the transmission appeared to be good (although multimode, since the waveguides were designed for 1.3 micrometer wavelength). At first, GRIN-lens-terminated fibers were used to do the optical input and output coupling at 1.3 micrometers.

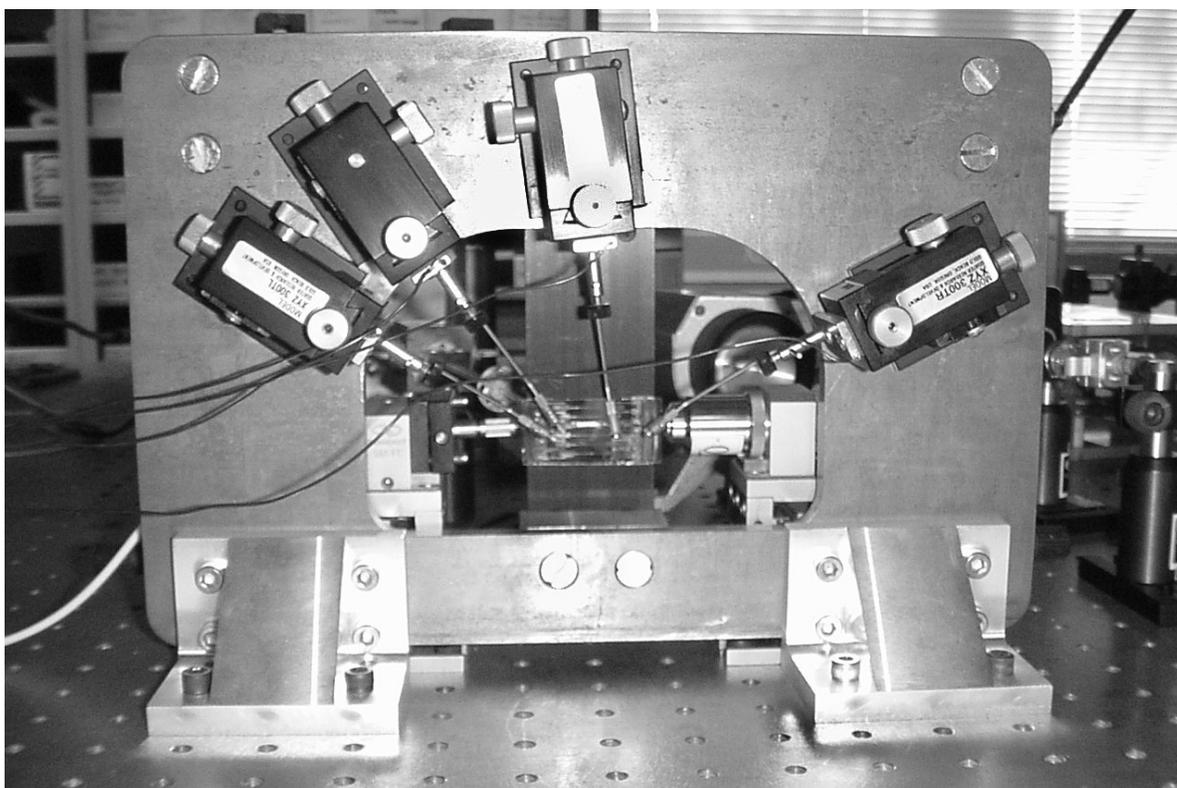


Figure 6 - 7: Photograph of the modulator mounting structure. The modulator is installed and the probes are positioned to apply d-c biases. The r-f feed is applied to the back side.

Between the experimental work on the first chip (with the “mouse bites”) and the third chip, shown in Figure 6 - 6, the entire laboratory was converted to single-mode, polarization-

preserving fiber with FP/APC connectors, so that every source, detector, and input/output fiber could be interchanged easily with minimum realignment. This work was done by graduate student Lee Burrows. While most of the changeover went smoothly, the input/output coupling fibers proved very troublesome. The measurements of the GRIN-lens-terminated fibers proved them to be inadequate in their optical quality; they were mechanically unusable. Also, a great deal of difficulty was experienced in using commercial fiber chucks from Newport Corporation (non-axial position and drift in alignment), and eventually these were abandoned in favor of a Caltech-designed and fabricated fiber holder. After a lot of trials, it was found that butt-coupled fibers in the Caltech chucks would work the best as the input source for all of the labs' experiments.

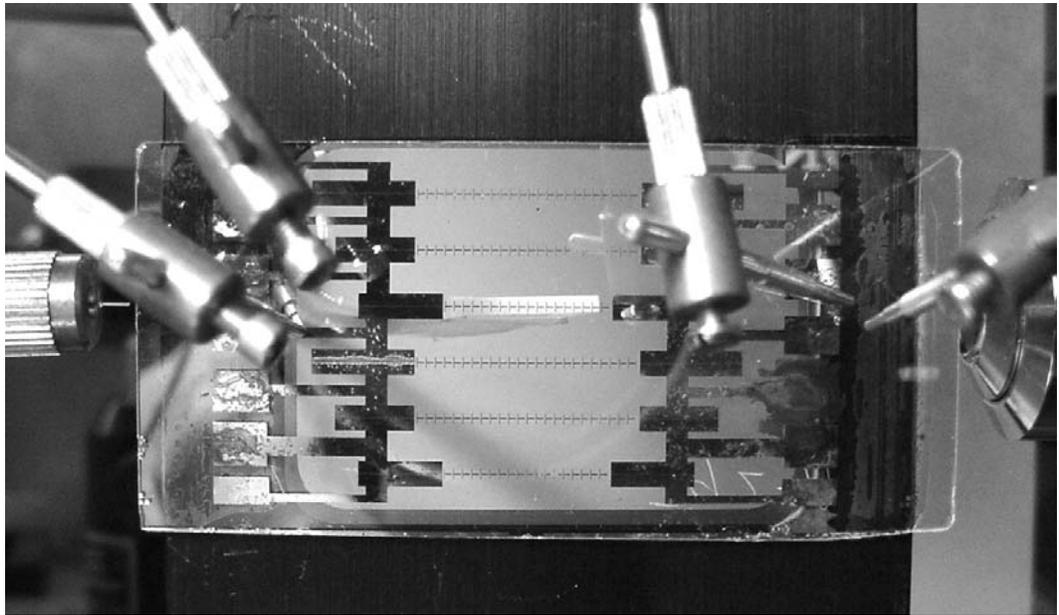


Figure 6 - 8: Modulator chip mounted on the mechanical carrier stage. The r-f feed can be seen positioned to apply the r-f signal from the backside on the third modulator from the top.

The modulator was set up, and transmission measurements were made at 1.3 micrometers with the new fiber-coupled components. The transmission loss averaged about -8 dB for the several modulators on the substrate. While not “commercial quality,” it is much better than

the -20 dB of the chip with the “mouse bite” waveguides. The d-c electrode sections were biased and exhibited the proper transfer function form, indicating that Gaeta’s new values for the coupling parameters were correct and had been used correctly in the design of the modulator. Bias was applied for several days to assure that drift would not be a problem with these chips. The bias points appeared to be sufficiently stable for the measurements.

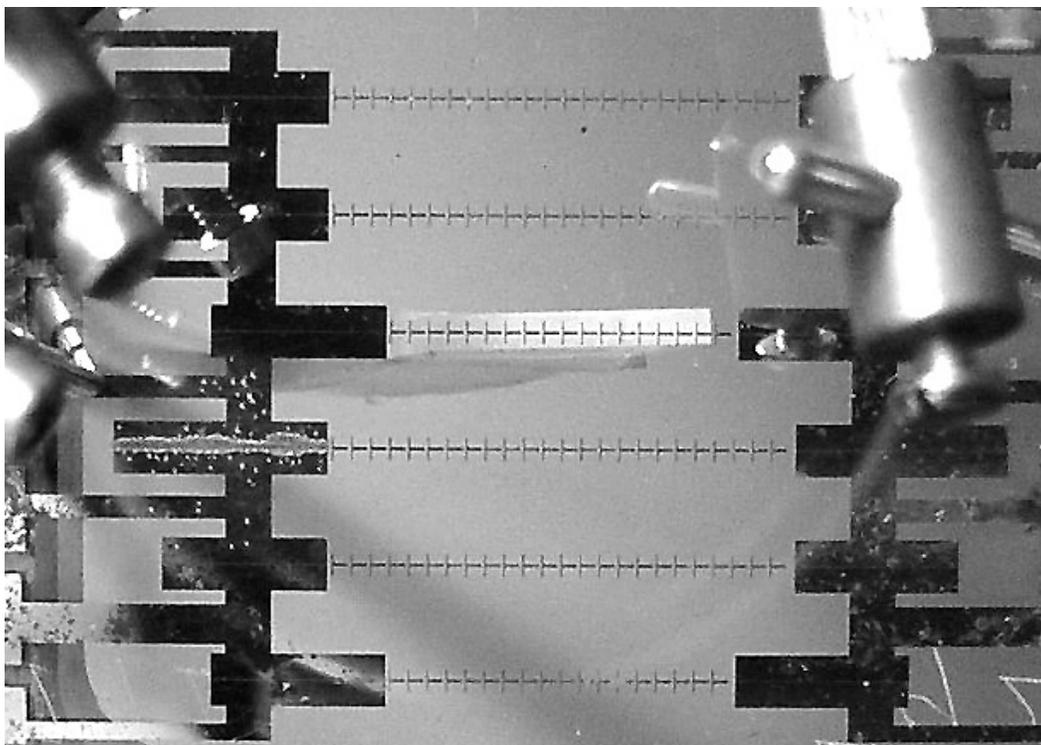


Figure 6 - 9: Blown up view of the modulator chip and millimeter-wave slab guide.

R-F Waveguide and Coupling Measurements

A 94 GHz Varian VRC-2113B23 klystron purchased in the late 1970s was tested, using a power supply borrowed from Prof. David Rutledge’s group. The 94 GHz tube still measured approximately 80 mW output over its 90.75 to 97.75 GHz tuning range.

The millimeter-wave coupling structure used to excite the antennas on the modulator utilizes a tapered, low dielectric constant dielectric slab waveguide “expansion fan” to increase the width

of the slab waveguide from the 0.100 inch width (2.54 mm) of WR-10 waveguide out to the length of the antenna array on the modulator, approximately 10 mm. The wide end of the fan is butt-coupled to a lithium niobate slab waveguide ($\epsilon_r = 28$) through a “quarter-wave” section of intermediate dielectric constant to minimize reflections. The lithium niobate waveguide is cut in a wedge of the correct angle for velocity matching, and is glued to the backside of the modulator chip. This coupling scheme is shown schematically in Figure 6 - 10. It was used successfully by Sheehy in Refs. [1.2] and [1.3], and by Cummings in Refs. [6.1] and [6.2], but it was not truly optimized for those measurements.

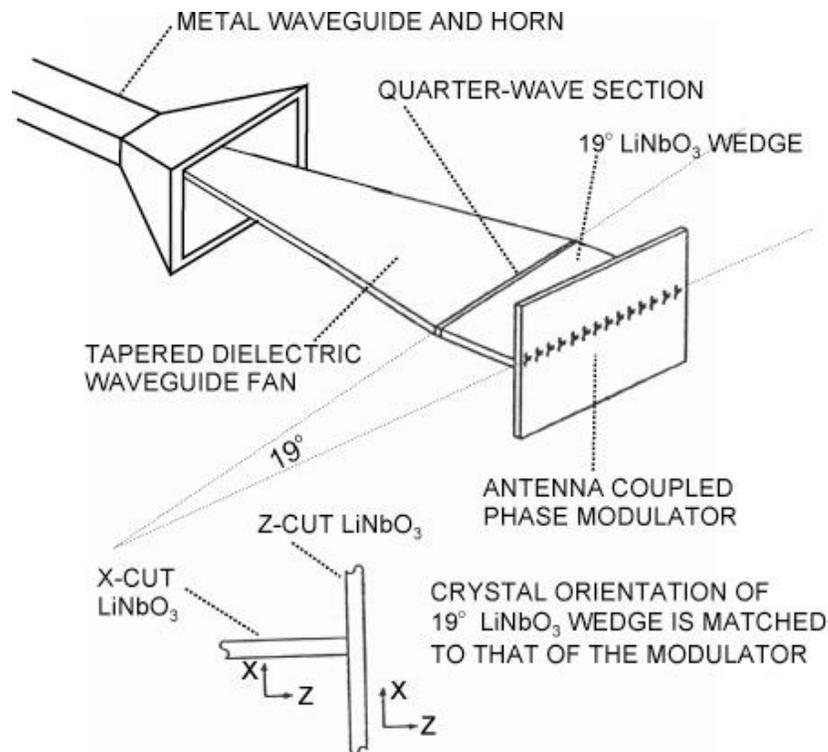


Figure 6 - 10: Slab dielectric waveguide coupling scheme for millimeter-wave modulators.

To optimize the coupling of the millimeter-wave modulation into the modulator, reflections at the junctions between the various slab waveguide sections must be minimized. This must be done by trial and error, since no accurate theory exists for the metal-waveguide-to-slab wedge junction, or for slab-to-slab junctions using a “quarter-wave” matching section. To carry out

this optimization, transmission measurements of the two such structures, coupled “back-to-back” were made with variously sized matching layers. It can be assumed that the transmission loss from the WR-10 waveguide to the modulator would be approximately half the loss (in dB) of the overall transmission. A photograph of a back-to-back slab waveguide arrangement is shown in Figure 6 - 11. Two polypropylene tapered slab waveguides are coupled to a uniform section of lithium niobate through “quarter-wave” thicknesses of an intermediate dielectric constant material, Stycast.[®]

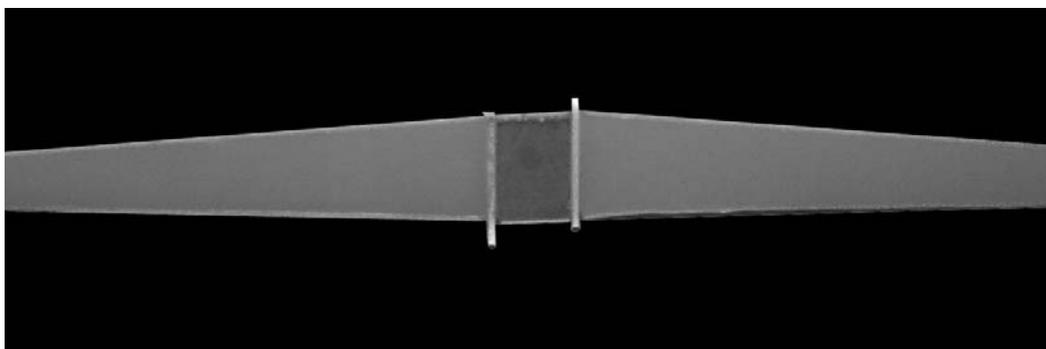


Figure 6 - 11: Photograph of a “back-to-back” coupling structure.

In the measurement setup, the fanned-out sections of slab waveguide were inserted into W-band horns. The overall transmission was measured and compared to the signal measured with the metal waveguide connected “flange-to-flange.” To remove the effects of detector non-linearity, a precision attenuator was used to keep the detector signal level the same for both the transmission through the entire structure and the transmission through just flange-to-flange WR-10.

Many combinations of fan dimensions, and matching section dielectric constants and thicknesses were evaluated. Polypropylene ($\epsilon_r = 2.2$) 1 mm thick was originally used for the expansion fans. The 1 mm thickness is a good match for the 1.25 mm height of WR-10 waveguide and the 0.656 mm width of the dipoles on the lithium niobate substrate. It also

gives a reasonably well-confined mode on the slab. In the case of infinite plane-wave transmission through a dielectric interface, a “quarter-wave” matching layer should have a dielectric constant equal to the geometric mean of polypropylene and lithium niobate, $(2.2 \times 28)^{1/2} = 7.8$, and should have a thickness of $0.25 \times (3.2 \text{ mm}/\sqrt{7.8}) = 0.28 \text{ mm}$ (or an odd multiple thereof). Of course, in the slab waveguide coupling structure, the fields are not fully confined to the dielectric material. The effective dielectric constant of the guide will be reduced in all cases. Simple slab waveguide theory, such as in Ref. [6.5] for example, shows how to calculate the change in effective dielectric constant. For the polypropylene guide 1 mm thick at 94 GHz, the effective dielectric constant is $0.60 \times 2.2 = 1.32$. For the lithium niobate guide 1 mm thick at 94 GHz, it is $0.916 \times 28 = 25.6$. For the intermediate dielectric constant

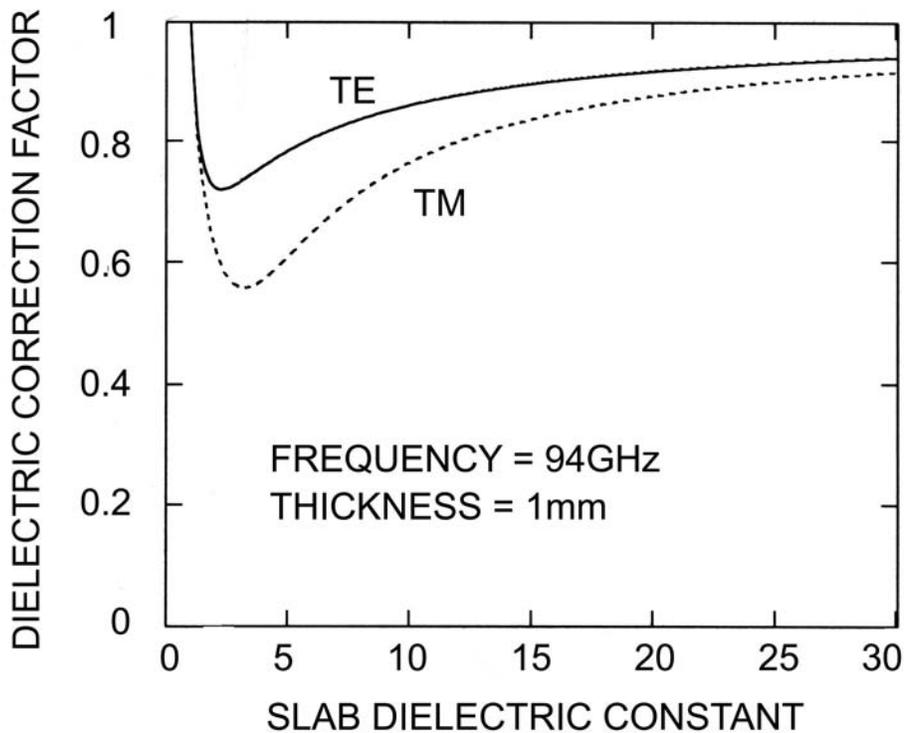


Figure 6 - 12: Plot of dielectric constant correction factor as a function of dielectric constant for a 1 mm thick slab waveguide at 94 GHz.

material in the vicinity of 8, the correction factor is about 0.72. These factors are for TM_0 propagation in the slab, as is correct for this setup. For TE_0 propagation, the correction values are somewhat larger. Figure 6 - 12 shows the dielectric constant correction factor for both TE_0 and TM_0 modes as a function of the slab dielectric constant for 1 mm thick slab waveguides at 94 GHz.

One might think that knowing these correction factors for the effective dielectric constant, the “quarter-wave” matching formula could be recalculated for an ideal match. Unfortunately, in addition to the change in propagation velocity and wave impedance, there will be a mismatch in the spatial distribution of the modes in the three regions, so that there will be a “discontinuity” reflection, even if the mode velocities and wave impedances are correctly arranged for “quarter-wave” matching. At the time of this work, no reasonable theory exists for these discontinuities. For these reasons, a trial and error approach was followed, using the materials at hand rather than ordering special values of dielectric constant.

An additional problem arose as the back-to-back measurements proceeded, that of mechanical instability. Polypropylene is a flexible plastic material, and it proved difficult to maintain a flat, rigid structure while making the measurements. Sheehy also had this problem, but decided to simply “live with it.” This time, in the interest of using a more rigid material, fused silica (1 mm thick) was chosen for the expansion fans. Fused silica has a dielectric constant of $\epsilon_r = 4.0$ at millimeter-wave frequencies, so that the optimum (infinite plane wave approximation) matching dielectric constant would be $(4 \times 28)^{1/2} = 10.6$, and a quarter wave would be $0.25 \times (3.2 \text{ mm}/\sqrt{10.6}) = 0.25 \text{ mm}$ thick. The correction factor for a 1 mm thick slab at 94 GHz is 0.572, so the effective dielectric constant is 2.3. For a matching slab 1 mm thick at 94 GHz with dielectric constant about 10, the correction factor is 0.765 for the TM_0 mode. However, the transmission losses were a few dB higher than in the guides based on polypropylene, so the fused silica slab was abandoned for the flexible but low loss polypropylene.

It is difficult to make actual quarter-wave layers because they are so thin. The dielectric material was Emerson and Cuming Stycast[®] artificial dielectric. This material is made from a high dielectric constant powder suspended in a low loss plastic matrix. The dielectric constant

is controlled by the fraction of the powder used. It is not a strong material, and thin sections break easily. The most successful method found to make thin layers was to cut as thin a piece as possible with a saw, then glue it to the dielectric fan and lap it to the desired thickness with sandpaper on glass. Even so, the material would generally break off during attempts to make to make sections actually one quarter wavelength in thickness. Typically, odd numbers of quarter wavelengths were used to gain some material strength. However, the absolute accuracy required is the same.

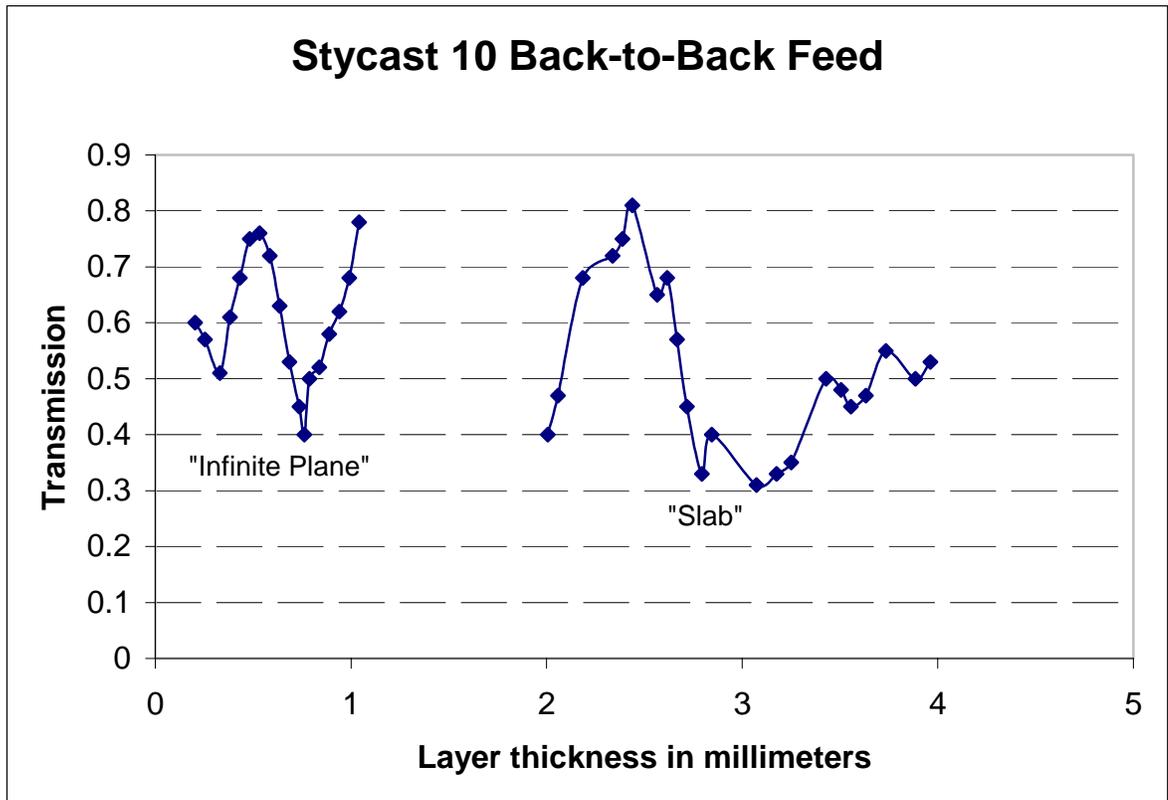


Figure 6 - 13: a) A layer of Stycast between two polypropylene feeds, the rectangle is larger than the polypropylene. b) A slab layer of Stycast between two polypropylene feeds; the dimensions are the same as the polypropylene.

Figure 6 - 13 shows two types of matching layers with a dielectric with $\epsilon_r = 10$ between the two polypropylene feeds (no lithium niobate section). The first experiment recreated the

matching layer used in the original Sheehy experiment. This layer is an “infinite plane layer” in that the width and height dimensions of the dielectric material are a few millimeters greater transversely than the polypropylene waveguide. Thus the fields are entirely in the dielectric material. In this back-to-back feed, odd multiples of $1/4 \lambda$ minimize transmission and even multiples maximizes transmission. Unlike a quarter-wave matching layer, the back-to-back experiment goes from low dielectric to high dielectric back to low dielectric. In the experiment, a wavelength appears to be about 1 mm since there are maxima at 0.5 mm and 1.0 mm and minima at 0.25 mm and 0.75 mm. This corresponds to theory, $\lambda = c/f = 3 \cdot 10^{11} / (94 \cdot 10^9 \cdot \sqrt{10}) = 1.01$ mm. However the maximum and minimum transmission is not simply calculable. It stands to reason, that the “infinite matching layer” would have a lower peak transmission than a slab guide matching layer because the modes are less well-matched transversely. The second plot in Figure 6 - 13 shows a similar experiment, but with a slab matching layer of Stycast 10 material. Notice that the peak transmission, 81% at 2.43 mm, is higher than the peak transmission of the infinite plane matching layer, 76% at 0.53 mm, even though there are considerable losses in Stycast 10 material, and a layer of 2.43 mm should have less transmission than a layer of 0.53 mm. The peak transmission of the slab should occur at 2 wavelengths (that is an integer number of $1/2$ wavelengths in the range of 2 – 3 mm). With the correction factor for TM_0 we would expect a length of 2.29 mm (the corrected guide wavelength is 1.14 mm). The experimental peak is at 2.43 mm. We should not expect exact values of $\lambda/4$, however, since there are discontinuity effects at both ends of the matching layer that will produce reflections in addition to those at the surfaces themselves. The shape and smoothness of the graph is non-ideal. This most likely comes from the difficulty of controlling the dimensions in the highly non-precise “hand-sanding” method of building these layers.

Given that Stycast 10 is lossy, and the slab layer is difficult to “hand-sand,” the experiment was duplicated for a slab layer of Stycast 5 material. This is shown in Figure 6 - 14. Theory with the approximate correction factor predicts a guide wavelength of 1.61 mm. This would give a maxima at 0.81 mm, 1.61 mm, and 2.42 mm etc, and a minima at 1.21 mm, 2.01 mm, and 2.82 mm, and so forth. From the data, it appears that there is a periodicity of twice that. However, there is a lot variation. The experiment was not refined until the data corresponded well to

theory, given the challenge of creating precision dimensions with this material at the frequency of interest. However, note that the maximum transmission at 94% is pretty good. Since Stycast 5 and Stycast 10 are both reasonable candidates for a matching material to lithium niobate, both materials were used to produce the r-f feed, but since the Stycast 5 material appears to yield a higher transmission rate, it was favored.

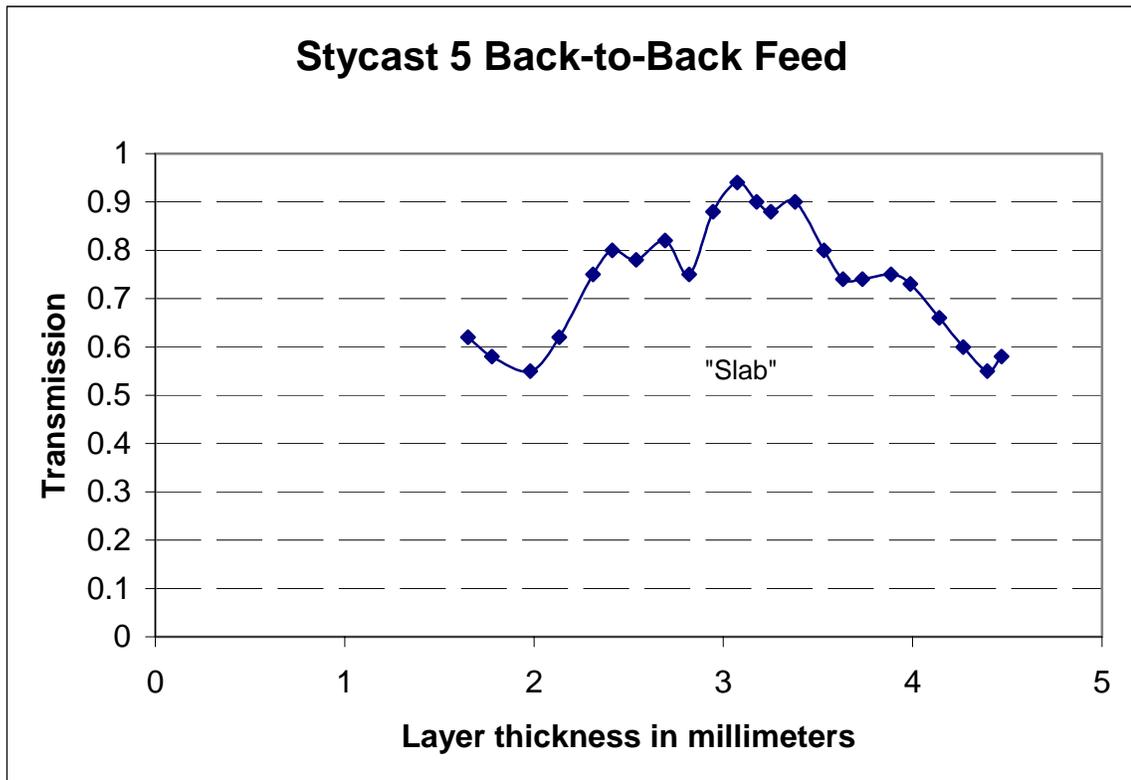


Figure 6 - 14: A slab layer of Stycast 5 between two polypropylene feeds.

The best results obtained for matching from WR-10 to a polypropylene dielectric fan, through a matching layer to a lithium niobate section and back again were flange-to-flange losses of 11.6 to 12.2 dB. Note that an equivalent length (about 40 cm) of WR-10 waveguide would have a loss of about 1.5 dB. Two different matching layer configurations gave the same range of losses:

(1) 2.0 mm of $\epsilon_r = 5$ material ($5/4 \lambda$).

(2) 2.6 mm of $\epsilon_r = 10$ material ($9/4 \lambda$).

These were the smallest pieces that could reasonably be used with the equipment at hand.

An attempt was made to reduce losses that originate as radiation from the dielectric discontinuities by layering copper foil on each side of the slab waveguide assembly. Transmission losses increased about 3 dB with the copper foil in place. Actually, this is the expected value from theory, since the currents induced in the copper walls add loss compared to the negligible loss in the exterior dielectric (air) when they are absent. While it is unclear how much loss came from radiation and reflections from the discontinuities in the slab guide, the measurements indicate that the losses in the two copper walls are 3 dB worse.

The foil-clad waveguide has the merit that the modes can be calculated exactly, and theoretical values of dielectric constant and length can be obtained for the proper match. And the transverse distributions of the TM_0 mode are forced to match by the confining metal conductors. However, the additional loss observed experimentally seemed to argue against using the foil-clad waveguide, so that approach was abandoned.

Table 6-1: Feed Transmission Measurements

Feed Description	Transmission
Topology in Sheehy experiment: Polypropylene, Stycast 10 infinite plane matching layer, lithium niobate	0.15
Polypropylene, optimal width Stycast 10 slab layer, lithium niobate	0.22
Polypropylene, optimal width Stycast 5 slab layer, lithium niobate	0.38

Table 6-1 shows the key results of the back-to-back feed measurements. It appears that a 3x improvement over the original feed construction was achieved. However, construction of the feed is so mechanically unstable and non-reproducible that such a conclusion really cannot be claimed definitively. In any event, the combination of fan and matching layers given above would exhibit about 2 - 3 dB of loss (estimate) between the WR-10 waveguide and the bottom surface of the modulator.

94 GHz Measurements

The overall setup for the actual modulation measurements is shown in Figure 6 - 15. A 94 GHz klystron is on the left, mounted in a large finned block. A cooling fan blows on the fins. A ferrite isolator is used to eliminate reflections back into the klystron. Following the isolator, a -20 dB cross-guide directional coupler is used to sample the klystron output for the precision absorption wave meter. The main signal continues through WR-10 waveguide to a millimeter-wave horn and the dielectric slab fan (obscured behind the modulator support structure in this photograph).

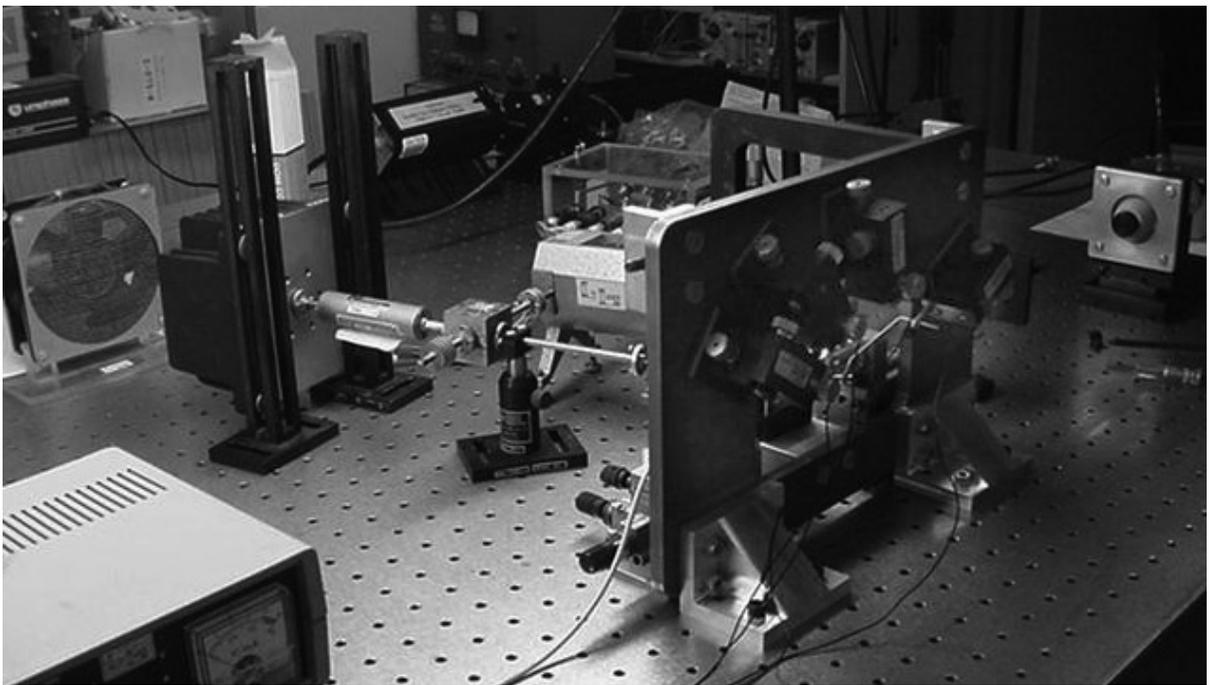


Figure 6 - 15: Photograph of the overall modulator measurement setup.

The optical power to the modulator at 1.3 microns was input via a butt-coupled fiber mounted in a precision 3-axis motion. The modulator output was imaged with a microscope objective lens onto a scanning Fabry-Perot (SFP) interferometer (right edge of Figure 6 - 15) used to detect modulation sidebands. (Recall that there was no available photodetector with 94 GHz response. Instead, the modulator is evaluated by looking for sidebands in the optical spectrum). The free spectral range of the SFP was 2.00 GHz, so there would be the order of 50 times aliasing in observing the sidebands. However, the frequency of the klystron was measured with an Alpha Industries model W551 wave meter, with absolute accuracy of 0.2%, (about 200 MHz at 94 GHz) so it is fairly discernable where the klystron-induced sidebands should occur within the 2.00 GHz display range. The measurement technique was essentially the same as the measurement that had proven successful before, Ref. [6.1], except now the mechanical structure was much more stable, and the new modulator chip exhibited lower optical loss and a reasonable transfer function compared to the “mouse bite” modulator of Ref. [6.1].

Unfortunately, the expected modulation sidebands were not observed. Many variations on the measurement were made (described below), but none produced sidebands visible on the oscilloscope trace of the SFP output. In the course of these measurements, a curious phenomenon was observed, one for which no explanation, only speculation, is reported, and that is also described below.

One difference from the old measurement was the source laser. The experiment used an ATX model 1.3AMPU 1.3 micron Nd-YAG laser, one capable of about 150 mW output. The old experiment used an Amoco Laser Company (later, ATX) model ALC 1320-EHS laser that produced only 5 mW maximum output power, the only laser available for the experiment within the contract budget at the time. In an ordinary photonic link, the link gain is proportional to the laser power, which motivated the switch to the higher power laser. However, in the detection scheme, the measurement is of the relative height of the sideband and the carrier, as detected on the SFP display. This ratio is independent of laser power, as long as the modulation sideband is visible above the oscilloscope noise level. However, the sideband must also be somewhat above the “skirt” of the carrier as displayed on the

oscilloscope screen. (The “skirt” results from the finite width of the SFP filter. To minimize the interference from the skirts of the carrier, it is best to tune the modulation frequency to a value halfway between the aliased “carriers” on the SFP. For a 2.00 GHz SFP free spectral range, this means the modulation frequency should be chosen at odd numbers of GHz: 91.00, 93.00, 95.00, or 97.00 GHz for the tuning range of the klystron. Operation at 92.00, 94.00, or 96.00 GHz would result in modulation sidebands that fall exactly on the aliased carriers on the display and would be completely undetectable.

An unfortunate disadvantage of the new high power laser was that it seldom operated with a single optical mode, typically oscillating at three or four optical frequencies simultaneously. (It was supposed to be a truly single-frequency laser, and often was when it was new. But by the time measurements were taken on this modulator, the components within the laser had changed so that it never operated at a single frequency). As a consequence, the SFP 2.00 GHz display was cluttered with several aliased “carriers” from all the frequencies output from the laser. Each of these carriers should display sidebands, but with multiple modes, it became more difficult to find a modulation frequency that would place any sidebands at the low points among all the aliased carrier skirts.

After some frustrating days with the high power laser, the experiment was re-instrumented with the 5 mW output original laser. This made the oscilloscope display less cluttered, but also failed to produce any visible sidebands. And, of course, the risk was now that the sidebands were actually below the oscilloscope noise level.

The next attempt at improving the measurement involved building a new higher-gain, lower-noise amplifier for the photodetector. This assumes that the measurement was limited by the photodetector amplifier noise, not the built-in oscilloscope amplifier noise. The design proved a bit tricky to do. The eventual configuration used a special op-amp as a trans-admittance amplifier for the photodetector, one with very low equivalent current noise rather than the more usual low equivalent voltage noise. This is done because the photodiode source looks like a current source rather than a voltage source. The op-amp that was selected was the Analog Devices 795, the lowest current-noise op-amp that the author could find. The feedback resistor in the first stage was 20 M Ω , so the output of the amplifier is 20 volts per

microamp of input current. With this amplifier, the noise level viewed on the SFP output on the oscilloscope was amplifier-noise-limited. However, even with this improved amplifier, there were still no discernable modulation sidebands.

In an attempt to further improve the sensitivity of the setup, the SFP amplifier output was fed to a “lock-in amplifier” (audio phase-synchronous detector) as well as the oscilloscope. The klystron power supply had built-in square-wave modulation capability at 1000 Hz, and this was enabled. Thus the DCM was now being fed a W-band signal that is 100% on-off modulated at 1000 Hz. With this arrangement, a very narrow equivalent bandwidth of the synchronous detection process could be used. Of course, the scans of the SFP must be done very slowly to accommodate the long integration time of the lock-in. Typically, this required scanning the SFP manually, observing the deflection of the meter on the lock-in. Even using this technique, still no modulation sidebands were found at locations in the SFP frequency scan where they were predicted them to be from the precision W-band frequency meter. From the “heroic” detection improvements on the original measurements of the “mouse-bite” modulator, the author is forced to conclude that the present modulator chip is at least three orders of magnitude less sensitive than the old modulator, for reasons not understood, but which are due to faulty chip fabrication in some way.

It was in the course of making the synchronous detection measurements that a new and, so far, unexplained phenomenon was discovered. There were repeatable synchronously detected outputs from the SFP detector/amplifier combination when the SFP was tuned to the steep slope of an aliased carrier signal!

The behavior was as if the applied millimeter-wave modulation were frequency-modulating the optical signal at 1000 Hz, and then “slope detection” of that FM was used by tuning the optical filter of the SFP to have its steep slope tuned to the optical carrier frequency. It was then checked that this phenomenon was, indeed, coupled into the modulator via the millimeter-wave slab waveguide, rather than by stray pickup in the bench wiring. It was.

The only hypothesis at this point is that the modulator chip itself is somehow “detecting” the W-band signal, and using the “detected” output (a 1000 Hz square wave) to produce phase

modulation (which would be indistinguishable from frequency modulation in this experiment). It may be possible that the process of evaporating and etching the electrodes several times on this one chip (using the negative resist process in Prof. Scherer's lab, rather than the positive resist "gold lift-off" method that HRL used on the original modulator) had inadvertently left a thin conductive layer bridging the transmission line electrodes, which acted as a "thin-film bolometer" that "shorts out" the millimeter-wave signal, but heats and cools rapidly enough to phase modulate the underlying optical waveguides. Admittedly, this explanation is grasping at straws. The only way to check for sure is to make another modulator chip by the original processes used at HRL (or by successfully developing the negative resist process) and find successful results. Unfortunately, by the time these measurements were made, HRL had decided to go out of the modulator fabrication business entirely, so there was no opportunity to have them make additional modulator chips. Since the contract support from Rome Labs had also run out, there was no opportunity to make additional modulator chips at Caltech. At this point, it was decided to abandon the "improved antenna-coupled directional coupler modulator" demonstration, and concentrate of the linearized modulator work described in chapter 3 - 5.

SUGGESTIONS FOR FUTURE WORK

There are many frontiers left in the development of high-frequency, high-fidelity analog modulators, both theoretical and experimental. Some of these are listed below.

Theoretical Analysis

This thesis focused on the basics of intensity modulators, and two-tone analysis of a modulator in a simple link model. It merely touched on CTB and CSO analysis. There are significant additional questions to ask along those lines, for example Ref. [1.1] suggests that no clean relationship exists between two-tone analysis to determine the SRDR and the CTB and CSO analysis in multi-channel systems.

One simple theoretical expansion of the techniques developed for this thesis research is the addition of finite r-f loss to the traveling wave electrodes. Actually this capability already exists in the modulator analysis program, but was not actually used in the calculation. The effects of such losses on the linearization process would be especially interesting. Similarly, one could include transmission line non-uniformity, such as tapered coupling, dispersion, etc. to see what effects these would have on the modulator characteristics and linearization techniques.

This thesis (and much of the research and development in the wider modulator community) has dealt only with intensity modulators, since intensity-modulated signals are the easiest to detect. A simple photodiode turns the modulated intensity into a current replicating the original modulation signal. However, it is well-known that true amplitude modulation, where the optical field amplitude rather than the intensity follows the information signal, generally produces better dynamic range performance. Detection typically requires an optical heterodyne system to translate the modulated optical electric field into a modulated r-f electric field, where it can be further translated to the original information signal. It is this added complexity that has kept optical amplitude modulation from being widely used. It turns out that all the modulators considered in this thesis, including the linearized modulators, have analogous performance as amplitude modulators, with minor modification in optical structure and operating biases. Future research should include an extension of the performance analysis to amplitude modulated optical links.

Experimental Work

On the experimental side of high-fidelity optical links, one is struck by how few experimental demonstrations have been made compared to the number of theoretical publications that have appeared. Certainly, a fruitful area of development would be simply to confirm experimentally the predictions of this thesis and other published schemes. It is clear that to do so, one must have one's own modulator chip fabrication facility, one that is capable of precise parameter control, such as the coupling constant in lithium niobate waveguide directional couplers. Without such careful control, it would be difficult to proceed. However, with such capability, it would be interesting to explore several areas experimentally as listed below.

The basic idea of making an antenna-coupled millimeter-wave modulator based on a directional coupler optical structure was proven very early on, albeit with a defective optical directional coupler structure (the “mouse bites”), Ref. [6.3]. This program was aimed at improving and optimizing the structure (mainly by removing these serious defects that were discovered in the original design). Unfortunately, we were unable to demonstrate our improvements, due to limited chip fabrication availability.

In the meantime, a new directional coupler structure has come to light that would be a much better choice for an antenna-coupled structure, the Y-fed DCM, Ref. [3.9]. This structure was described in detail in Chapter 4 of this thesis, since it is also a good candidate for linearization. The Y-fed DCM is simply a DCM with equal optical powers fed to the two input arms, rather than all of the optical input fed into one arm. Due to this symmetrical feed, the transfer function changes dramatically. As discussed in Chapter 3, “zero” is automatically the correct bias for this modulator. This means that an antenna-coupled modulator need only the antenna-plus-transmission-line-segments to couple the r-f into the modulator. No extra bias electrodes are required (or desired, for that matter). This structure is a “natural” for antenna-coupled devices. Any future demonstration of antenna-coupled directional coupler modulators should use this much simpler configuration.

Another experimental area would be to confirm the predictions in this thesis for the improved bandwidth of the reflection modulators. These are relatively simple to make in the MZM versions, and should offer a significant improvement at low cost.

Once the directional coupler fabrication precision is established, an experimental demonstration of the “band-pass” characteristics of the DCM and YFDCM would be in order.

Finally it should be noted that the accuracy of photolithography defining the antenna structures is not even challenged at 100 GHz. It would be interesting to try modulating light at higher frequencies using the antenna-coupled schemes up to the point where lithium niobate has significant absorption, 0.5 THz.

One could also develop variations on the antenna elements themselves, such as “Vee” antennas with gain replacing simple dipoles, as suggested in Sheehy’s thesis (but not demonstrated) could be investigated experimentally.

BIBLIOGRAPHY

CH-1

- [1.1] Charles Cox and William S. C. Chang, "*Figures of merit and performance analysis of photonic microwave links*," RF Photonic Technology in Optical Fiber Links, edited by William S. C. Chang, Chapter 1, First Ed., Cambridge University Press, 2002.
- [1.2] F. T. Sheehy, W.B. Bridges, and J. H. Schaffner, "*94 GHz Antenna-coupled LiNbO₃ Electrooptic Modulators*," IEEE Photonics Technology Letters, vol. 5, pp 307-310, March 1993.
- [1.3] F. T. Sheehy, "*Antenna-Coupled millimeter-wave Electrooptic Modulators and Linearized Electrooptic Modulators*," Ph.D. Dissertation, California Institute of Technology, June 1993.

CH-2

- [2.1] Yariv, Amnon. *Optical Electronics*, 4th ed. Orlando, FL: HRW, Saunders, 1991.
- [2.2] Simon Ramo, John R. Whinnery, and Theodore Van Duzer, *Fields and Waves in Communication Electronics*, Second Ed. New York, NY: John Wiley & Sons, 1984.
- [2.3] Gary E. Betts, "*LiNbO₃ modulators and high performance analog links*," RF Photonic Technology in Optical Fiber Links, edited by William S. C. Chang, Chapter 4, First Ed., Cambridge University Press, 2002.
- [2.4] M. Seino, N. Mekada, T. Namiki, and H. Nakajima, "*33-GHz-cm broadband Ti:LiNbO₃ Mach-Zehnder modulator*," Tech. Dig. ECOC '89 paper ThB22-5, 1989.
- [2.5] K. Noguchi, O. Mitomi, and H. Miyazawa, "*Millimeter-wave Ti:LiNbO₃ optical modulators*," J. Lightwave Technol., 16, 615-19, 1998.
- [2.6] William B. Bridges, "*Antenna-coupled millimeter-wave electrooptical modulators*," RF Photonic Technology in Optical Fiber Links, edited by William S. C. Chang, Chapter 11, First Ed., Cambridge University Press, 2002.
- [2.7] R. C. Alferness, S. K. Korotky, and E. A. J. Marcatili, "*Velocity matching techniques for integrated optic traveling wave switch/modulators*," IEEE J. Quantum Electron., QE-20, 301-9, 1984.
- [2.8] William B. Bridges and James H. Schaffner, "*Broad band, low power electrooptic modulator apparatus and method with segmented electrodes*," USPTO 5,291,565.
- [2.9] William B. Bridges, "*Antenna-fed electrooptic modulator*," USPTO 5,076,655.

- [2.10] W. B. Bridges, U. V. Cummings, F. T. Sheehy, and J. H. Schaffner, "Wave-Coupled LiNbO_3 Directional Coupler Modulators," Paper 4.3, Photonic Systems for Antenna Applications (PSAA-IV), Monterey, CA, 18 - 21 January 1994.
- [2.11] R. V. Schmidt and I. P. Kaminow, "Metal-diffused optical waveguides in LiNbO_3 ," Appl. Phys. Lett., Vol. 25, No. 8, 1974.

CH-3

- [3.1] R.R. Hayes and D. L. Persechini, "Nonlinearity of PIN photodetectors," IEEE Photon. Technol. Lett., Vol. 5, No. 1, pp. 70 - 72, Jan 1993.
- [3.2] Charles H. Cox, *Analog Optical Links: Theory and Practice*, Cambridge Studies in Modern Optics, 2004
- [3.3] L.M. Johnson, H. V. Roussel, "Reduction of Intermodulation distortion in interferometric Optical Modulators," Optics Letters, Vol. 13, No. 10, October 1988.
- [3.4] Steven K. Korotky, René M. De Ridder, "Dual Parallel Modulation Schemes for Low-Distortion Analog Optical Transmission," IEEE Journal on Selected Areas in Communications, Vol. 8, No. 7, September 1990.
- [3.5] Halvor Skeie and Richard V. Johnson, "Linearization of Electrooptic Modulators by a Cascade Coupling of Phase Modulating Electrodes," SPIE Vol. 1583 Integrated Optical Circuits, 1991.
- [3.6] Mark L. Farwell, Zong-Qi Lin, Ed Wooten, and William S. C. Chang, "An Electrooptic Intensity Modulator with Improved Linearity," IEEE Photonics Technology Letters, Vol. 3, No. 9, September 1991.
- [3.7] J. F. Lam and G. L. Tangonan, "Optical modulation system with enhanced linearization properties," IEEE Photon. Technol. Lett., Vol. 3, No. 12, pp. 1102 - 1104, Dec. 1991.
- [3.8] William B. Bridges, James H. Schaffner, "Distortion in Linearized Electrooptic Modulators," IEEE Transactions on Microwave Theory and Techniques, Vol. 43, No. 9, September 1995.
- [3.9] S. Thanjavarn, "Modified 1×2 Directional Coupler Waveguide Modulator," Electronics Letters, Vol. 22, No. 18, 28 August 1986.
- [3.10] Benoit Pucel, Luc Rivière, and Jean Le Bihan, "New Model for the Active Directional Coupler," J. Lightwave Technol., Vol. 14, No. 6, June 1996.
- [3.11] R. F. Tavlykaev and R.V. Ramaswamy, "Highly linear Y-fed directional coupler modulator with low intermodulation distortion," J. Lightwave Technol., Vol. 17, pp. 282 - 291, Feb. 1999.
- [3.12] Tetsuya Kishino, Robert F. Tavlykaev, and Ramu V. Ramaswamy, "A Y-Fed Directional Coupler Modulator with a Highly Linear Transfer Curve," IEEE Photon. Technol. Lett., Vol. 12, No. 11, November 2000.
- [3.13] W. Ng, L. Luh, D. Yap, W. Bridges, M. Mokhtari, J Jensen, C. Fields and Y.M. So, "Photonic Analog-to-Digital Conversion Technology," PACT Final Report, DARPA Contract No. F30602-99-C-0022.

- [3.14] W.H. Louisell, *Coupled Modes and Parametric Electronics*, John Wiley, New York, 1960.
- [3.15] G. E. Betts, "Linearized modulator for suboctave-bandpass optical analog links," IEEE Trans. Microwave Theory Tech., Vol. 42, No. 12, pp. 2642 - 2649, Dec. 1994.
- [3.16] Betts, G. E., O'Donnell, F.J. and Ray, K.G., "Sub-octave-bandwidth analog link using linearized reflective modulator." PSAA-5 Proceedings, pp. 269-299, 1995.
- [3.17] James H. Schaffner and William B. Bridges, "Intermodulation distortion in high dynamic range microwave fiber-optic links with linearized modulators," J. Lightwave Technol., Vol. 11, No. 1, January 1993.
- [3.18] Gary D. Alley, Yen Long Kuo, "Optimal Control of Intermodulation Distortion in Hybrid Fiber-Coaxial CATV Systems," IEEE Transactions on Microwave Theory and Techniques, Vol. 45. No. 12, pp 2183-96, December 1997.
- [3.19] Bert Arnold, "Third-order Intermodulation Products in a CATV System," IEEE Transactions on Cable Television, Vol. CATV-2, No. 2, pp 67-80, April 1977.
- [3.20] A. A. M. Saleh, "Fundamental Limit on Number of Channels in Subcarrier-Multiplexed Lightwave CATV Systems," Electron. Lett., Vol. 25, No. 12, pp 776-777, June 8, 1989.

CH-4

- [4.1] M.L. Farwell and W.S.C. Chang, "Simulating the response of coupled channel and interferometric modulator designs," Journal of Lightwave Technology, Vol. 13, October 1995.
- [4.2] Uri V. Cummings and William Bridges, "Bandwidth of Linearized Electrooptic Modulators," Journal of Lightwave Technology, Vol. 16, No. 8, August 1998.
- [4.3] Mark L. Farwell, "Improved Linear Dynamic Range in Externally Modulated Fiber-optic Links," Ph.D. Dissertation, UCSD, June 1994.
- [4.4] J. H. Schaffner, J. F. Lam, C. J. Gaeta, G. L. Tangonan, R. L. Joyce, M. L. Farwell, and W. S. C. Chang, "Spur free dynamic range measurements of a fiber-optic link with traveling wave linearized directional coupler modulators," IEEE Photon. Technol. Lett., Vol. 6, No. 2, pp. 273-275, Feb. 1994.
- [4.5] G. K. Gopalakrishnan, W. K. Burns, R. W. McElhanon, C. H. Bulmer, and A. S. Greenblatt, "Performance modeling of broadband LiNbO₃ traveling wave optical intensity modulators," Journal of Lightwave Technology, Vol. 12, pp. 1807-1819, 1994.

CH-5

- [5.1] Uri Cummings, "Traveling-Wave Reflective Electrooptic Modulator," USPTO 5,886,807, Mar. 23, 1999.

- [5.2] K. Noguchi, O. Mitomi, K. Kawano, and M. Yanagibashi, "Highly efficient 40-GHz bandwidth Ti:LiNbO₃ optical modulator employing ridge structure," IEEE Photon. Technol. Lett., **5**, 52-4, 1993.
- [5.3] R.R. Hayes, (HRL) private communication 2003.
- [5.4] R. H. Buckley and J. F. Sonderegger, "A novel single-fiber antenna remoting link using a reflective external modulator," in Proc. Photonics Systems for Antenna Applications-91, Monterey, CA, 1991.
- [5.5] W. K. Burns, M. M. Howerton, R. P. Moeller, A. S. Greenblatt, and R. W. McElhanon, "Broad-Band Reflection Traveling-Wave LiNbO₃ Modulator," IEEE Photonics Technology Letters, Vol. 10, No. 6, June 1998.
- [5.6] W. K. Burns, M. M. Howerton, R. P. Moeller, "Broadband, unamplified optical link with r-f gain using a LiNbO₃ modulator," IEEE Photon. Technol. Lett., **11**, 1656-8, 1999.
- [5.7] W. K. Burns and M. M. Howerton, "Broadband traveling wave modulators in LiNbO₃," RF Photonic Technology in Optical Fiber Links, edited by William S. C. Chang, Chapter 5, First Ed., Cambridge University Press, 2002.

CH-6

- [6.1] W. B. Bridges, U. V. Cummings, F. T. Sheehy, and J. H. Schaffner, "Wave-Coupled LiNbO₃ Directional Coupler Modulator at 94 GHz," Paper 4.3, Photonic Systems for Antenna Applications-IV, Monterey, CA, 18 - 21 January 1994.
- [6.2] Uri V. Cummings, Senior Thesis, Submitted to Caltech in 1994. Unpublished.
- [6.3] William B. Bridges, Lee J. Burrows, Uri V. Cummings, Reynold E. Johnson, and Finbar T. Sheehy, "60 and 94 GHz Wave-Coupled Electrooptic Modulators, RL-TR-96-188," Final Technical Report, Rome Laboratory, Rome, New York, September 1996.
- [6.4] William B. Bridges, Lee J. Burrows, Uri V. Cummings, and Reynold Johnson, "Wave-Coupled Millimeter-Wave Electrooptic Techniques," Final Technical Report, CITLL-2000-01, Rome Laboratory, Rome, New York, December 2000.
- [6.5] H. G. Unger, *Planar Optical Waveguides and Fibers*, Oxford University Press, 1977, Chapter 2.

