TR-DA2180

STUDY OF EFFICIENT TRANSMISSION AND RECEPTION OF IMAGE-TYPE DATA USING MILLIMETER WAVES

By

C. Louis Cuccia

Prepared For The

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
Electronic Research Center
Contract Number NAS 12-2186
B. J. Goldstein, Technical Monitor

AUG 1971
RECEIVED
NASA STI FACILITY
INPUT BRANCH

PHILCO-FORD CORPORATION
Western Development Laboratories Division
Palo Alto, California

FACILITY FORM 602

71-34156	
(ACCESSION NUMBER)	(THRU)
542	63
(PAGES)	(CODE)
CR-121755	07
(NASA CR OR TMX OR AD NUMBER)	(CATEGORY)

PHILCO Ford

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151

This report was prepared as an account of Government-sponsored work. Neither the United States, nor the National Aeronautics and Space Administration (NASA), nor any person acting on behalf of NASA:

A.) Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately-owned rights; or

B.) Assumes any liabilities with respect to the use of, or for damages resulting from the use of, any information, apparatus, method or process disclosed in this report.

As used above, "person acting on behalf of NASA" includes any employee or contractor of NASA, or employee of such contractor, to the extent that such employee or contractor of NASA or employee of such contractor prepares, disseminates, or provides access to any information pursuant to his employment or contract with NASA, or his employment with such contractor.

Requests for copies of this report should be referred to

National Aeronautics and Space Administration
Scientific and Technical Information Facility
P. O. Box 33
College Park, Md. 20740

PRECEDING PAGE BLANK NOT FILMED

FINAL REPORT

STUDY OF EFFICIENT TRANSMISSION AND RECEPTION OF IMAGE-TYPE DATA USING MILLIMETER WAVES

By

C. Louis Cuccia

Prepared For The

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Contract Number NAS 12-2186

June 15, 1970

Under The Technical Management Of

B. J. Goldstein, Technical Monitor

NASA Electronic Research Center
Cambridge, Massachusetts

This report describes high data rate millimeter wave systems for the transfer of image type data. The recommended system approaches use delta modulation as the digital communication technique, and quadriphase modulation as the modulated carrier technique. This report has been written in two volumes. Volume 1 deals with all pertinent system and circuit concepts involved while Volume 2 addresses the principle technologies and technologies involved.

PRECEDING PAGE BLANK NOT FILMED

PRINCIPAL CONTRIBUTORS TO THIS REPORT

Wilfred B. Whalley:	Sensor Classification and Noise
Robert R. Jordan:	Sensor MTF
Dr. James J. Spilker:	Communication Theory and A/D Conversion Technology
Dr. D. Thomas Magill:	Quadriphase Modulation and Demodulation Analysis
Dr. Richard T. Sherman:	QPSK Demodulation in the Presence of Noise
William K. S. Leong:	Computer Analysis of Delta Modulation
C. Louis Cuccia:	High Speed A/D Conversion, Phase Modulation, and Millimeter Wave Systems
Don G. Middlebrook:	Link Analysis and Budgets
Dr. Fred Dietrich:	Millimeter Wave Antennas
J. Jay Jones:	Digital System Analysis
K. Ward:	Matched Filter Analysis
Dr. F. Natali:	Matched Filter Sampling Time Analysis
Richard Davies:	Comparison of Modulation Techniques
Oddvar Johanneson:	Phase Shift Modulators

Report Organized and Written by C. Louis Cuccia

Program Manager:
C. Louis Cuccia, Manager
RF Systems Department

Report Approved by:
James J. Spilker, Jr.,
Advanced Information Systems Activity

PRECEDING PAGE BLANK NOT FILMED

SUMMARY PREFACE

This report investigates and extends the present low video-rate image-transfer technology to high data rates using millimeter waves as the transmission medium. The investigations include:

- Types of sensors capable of frequency components up to 40-60 GHz

  Candidates: Image orthicons
  Vidicons
  Isocons
  Image dissectors
  Spin cameras with photoelectric cells

- Types of A/D conversion techniques capable of developing bit rates from 40 to 400 Mbps

  Candidates: Delta modulation
  Cyclic conversion
  Phase domain conversion
  CRT A/D converters

- Optimum types of modulation

  Candidates: Analog FM
  Phase shift keying
  Frequency shift keying

- Types of modulator and demodulator techniques capable of handling bit rates from 40 to 400 Mbps

  Candidates: Modulators
  - Switched-circulator types

Demodulators

- Differential phase shift keying demodulators
- Coherent phase shift keying demodulators

- Frequency ranges for supporting high data rate digital transmission

  Candidates: 35 GHz
  60 GHz
  75 GHz
  90 GHz

- Antenna and receivers to provide millimeter wave links with margin above maximum bit error rate

  Candidates: Optimum acquisition and tracking procedures
  Antenna size and gain
  Antenna construction techniques
  Receiver techniques

TR-DA2180

STUDY OF EFFICIENT TRANSMISSION AND RECEPTION OF IMAGE-TYPE DATA USING MILLIMETER WAVES

FINAL REPORT

VOLUME 1

SYSTEM REQUIREMENTS AND CIRCUITS

PHILCO Ford

SPACE & RE-ENTRY SYSTEMS DIVISION
Philco-Ford Corporation

TABLE OF CONTENTS

VOLUME 1

Section No.		Page
1	The Communication Aspects of Image Transmission High Data Rate	1-1
1.1	Objectives	1-1
1.2	Sensor Satellite to Relay Satellite Systems	1-4
1.3	System Variants and Communication Parameters	1-4
1.4	Basic Communication Subsystems	1-7
1.5	Basic Report Conclusions	1-9
2	Communication System Aspects of High-Data Rate Imagery Transmission	2-1
2.1	Introduction	2-1
2.1.1	Typical Image-Data Transmission Links	2-1
2.2	Multiple Sensor Image Data Transmission System	2-3
2.2.1	Recommended Modulation Technique	2-3
2.3	Comparison of Analog and Digital Modulation Techniques	2-3
2.3.1	Advantages of Digital Links	2-6
2.3.1.1	Data Regeneration	2-6
2.3.1.2	Multipath Rejection	2-6
2.3.1.3	Less Susceptibility to Channel Distortion	2-6
2.3.2	Quantitative Comparison of Digital and Analog Links for Transmission of Image Data	2-7
2.3.2.1	Signal Power Requirements of Analog and Digital Links	2-8
2.3.2.2	Non-Ideal Digital Information Transmission	2-10
2.3.2.3	Bandwidths for FM, AM, and PSK	2-12
2.4	Comparison of Modulation Techniques at 40 to 400 Mbps Rates	2-19
2.5	Candidate Transmission Bands for High Data Rate Image Transmission	2-21

TABLE OF CONTENTS (continued)

VOLUME 1

Section No.		Page
3	Image Sensor Technology	3-1
3.1	The Data Characteristics of Image Sensors	3-1
3.1.1	General Sensor Requirements	3-3
3.1.1.1	Wavelengths of Interest	3-3
3.1.1.2	Resolution	3-4
3.1.1.3	Illumination Levels	3-4
3.2	Candidate Photo Sensor Imaging Devices for Space Applications	3-4
3.2.1	Functional Classes of Image Tubes	3-5
3.2.2	Camera Tube Types	3-5
3.2.3	Self-Scanned Mosaic Image Sensors	3-9
3.2.4	Mosaic Resolution	3-9
3.3	Microwave Radiometer and Side-Looking Radar Sensors	3-11
3.4	Video Bandwidth of Image Sensors	3-15
3.4.1	Bandwidths of Parallel Operating Image Sensors	3-17
3.4.2	Wide Band Video Amplifiers	3-18
3.4.3	Wide Band On-Board Image Processing Techniques	3-18
3.5	Signal to Noise Considerations	3-22
3.5.1	Basic Noise Contributors	3-22
3.5.1.1	Signal to Noise Ratio and Illumination Level	3-22
3.5.2	Noise and Sensitivity	3-22
3.5.2.1	Noise Levels in Vidicons	3-23
3.5.2.2	Noise Levels in Image Orthicons	3-23
3.5.3	Basic Signal to Noise Relationships	3-23
3.5.4	Practical Signal to Noise Considerations	3-25
3.6	Image Sensor Systems	3-25
3.6.1	Sensor System Combinations	3-26
3.6.2	Image Data Bandwidths for Highest Resolutions	3-26
3.6.3	Spin-Scan Image Cameras	3-29

TABLE OF CONTENTS (continued)

VOLUME 1

Section No.		Page
3.6.4	Multi-Spectral Scanner (MSS) for Earth Resources Technology Satellite	3-31
3.6.5	An Illustration of a Multiple Sensor System - the ERTS-A Spacecraft	3-37
3.6.6	The Interface Between the Sensor System and the A/D Converter	3-37
4	Digital Transmission of Imagery Data	4-1
4.1	Introduction	4-1
4.2	Digital Code Equivalent of Analog Signals	4-1
4.3	Data Rate Required for a Given Baseband Bandwidth	4-3
4.4	Allowable Bit Error Rate	4-8
4.5	Types of Ultra High Speed A/D Converters	4-8
4.6	Delta Modulator/Demodulator Implementation	4-10
4.6.1	The Two-Bit Delta Modulator - Practical Circuit Aspects	4-12
4.6.1.1	Role of the Integrator	4-14
4.6.1.2	Variable Step-Size Delta Modulator	4-14
4.6.2	Basic Delta Modulator System	4-19
4.6.2.1	Digital Synchronization	4-24
4.6.2.2	PN Test Generator	4-25
4.6.3	Delta Demodulator Configuration	4-27
4.6.3.1	PN Test Signal Comparator	4-27
4.6.3.2	Allowable BER	4-27
4.7	Multiplexing and Demultiplexing Digital Bit Streams	4-30
4.7.1	Interleaving Digital Pulses	4-30
4.7.2	Time Sharing of A/D Converters	4-30
4.7.3	Multiplexed Delta Modulator and Demodulator Systems	4-31
4.7.3.1	Data Bit Rate from Multiple Sensor Systems	4-35
4.7.4	Multiplexed Delta Modulator Implementation	4-36

TABLE OF CONTENTS (continued)

VOLUME 1

Section No. Page

4.7.5 Multiplexed Delta Demodulators 4-38
4.8 Image Quality in Digital Communications Systems 4-40
4.8.1 Some Measured Results 4-41
4.8.1.1 Signal-to-Noise Ratios 4-42
4.8.1.2 Noise and Distortion Optimization 4-43
4.8.1.3 Phase Characteristics 4-44
4.8.1.4 Dynamic Range 4-44
4.8.1.5 Recovery from Data Link Errors 4-45
4.8.1.6 Test Voltage Waveforms 4-46
4.9 Comparison of Multiplexing and Modulation Techniques 4-46
4.9.1 SSSC/FDM/FM 4-47
4.9.2 DSB/FDM/FM and VSB/FDM/FM 4-49
4.9.3 TDM/PAM/FM 4-49
4.9.4 TDM/PCM/PSK 4-49
4.9.5 TDM/DM/PSK 4-51
4.9.6 Recommended Technique for Multiplexing 4-51

5 Modulator/Demodulators for PSK Digital Data Systems 5-1
5.1 Phase Modulation Techniques for High Speed Digital Data 5-1
5.2 PSK Systems at 40-400 Mbps Data Rates 5-1
5.2.1 BPSK and QPSK Mod/Demod System 5-2
5.3 Carrier Synchronization in Biphase and Quadriphase Signaling Systems 5-5
5.4 Quadriphase Modulators - Circuit Aspects 5-7
5.4.1 Circulator-Type Diode Phase Shifters 5-7
5.4.1.1 Diode Switching Modes 5-10
5.4.1.2 Diode RF Admittance 5-15
5.4.1.3 PSK Switching Waveforms at High Bit Rates 5-16
5.4.2 3-dB Hybrid Biphase Modulator 5-22

TABLE OF CONTENTS (continued)

VOLUME 1

Section No.		Page
5.4.3	Quadriphase Modulator Using Hybrid/Combiner	5-22
5.4.4	Modulator RF Driver	5-25
5.4.5	Modulating Diode Driver	5-27
5.4.6	Differential Encoding at Diode Driver Input	5-27
5.5	Differentially Coherent Demodulator	5-34
5.5.1	Demodulator Configuration	5-34
5.5.2	Differentially Coherent Detection Operation	5-38
5.5.2.1	Data Filter Configurations	5-41
5.5.2.2	Bit Synchronizer	5-43
5.6	Coherent and Differentially Coherent Phase Detection	5-45
5.6.1	PSK (Coherent) Demodulation - Carrier Reconstruction	5-47
5.6.1.1	Times-Four Quadriphase Carrier Reconstruction Loop	5-51
5.6.2	Carrier Reconstruction Loop Bandwidth Design	5-52
5.6.2.1	Implementation Errors of Carrier Reconstruction Loop	5-56
5.6.2.2	Carrier Reconstruction in the Presence of a Noisy Reference	5-58
5.6.2.3	Effects of Oscillator Noise on Quadriphase Performance	5-64
5.6.2.3.1	Total Loop Phase Error	5-64
5.6.2.3.2	Probability of Bit Error for QPSK	5-67
5.6.2.3.3	Probability of Bit Error for BPSK	5-72
5.6.2.3.4	Calculation of σ^2 FM	5-74
5.6.3	Matched Filter Data Detection	5-79
5.6.3.1	Data Waveforms	5-85
5.6.3.2	Data Detection Matched Filters for High Rate Modems	5-88
5.6.3.3	Matched Filter Principle	5-89
5.6.3.4	Loss Effects for Realizable Data Filters	5-90
5.6.3.4.1	Energy Loss Due to Bandlimiting	5-91

TABLE OF CONTENTS (continued)

VOLUME 1

Section No.		Page
5.6.3.4.2	Mismatch Loss Due to the Data Filter	5-93
5.6.3.4.3	Data Detection Filter E/n_o Loss Due to Sample Timing Error	5-100
5.6.3.5	An Approximation to a Matched Filter for a Typical Modulator Output Waveform	5-104
5.6.3.6	Degradation Due to Decision Threshold Uncertainty	5-106
5.6.3.7	Filter Distortion and Intersymbol Interference Effects on QPSK	5-109
5.6.4	Bit Synchronizer	5-113
5.6.4.1	Bit Synchronizer Circuit Implementation	5-114
5.6.4.2	Bit Synchronizer Error in Modulator Switching Time	5-115
6	Acquisition - Search - Track	6-1
6.1	Introduction	6-1
6.2	Acquisition Problem	6-1
6.2.1	Spatial Acquisition Considerations	6-2
6.2.1.1	Position Uncertainty	6-2
6.2.2	Short Term Vehicle Stability	6-4
6.2.3	Attitude Control Subsystem	6-7
6.2.4	Spatial Acquisition Accuracy	6-8
6.3	Spatial Search	6-10
6.3.1	Search Sequence	6-10
6.3.2	Signal Detection	6-11
6.3.3	Considerations of Signal Power Level	6-13
6.3.4	SNR and Frequency Sweep Rate	6-13
6.4	Tracking	6-15
6.4.1	Tracking Receiver Configurations	6-17
6.4.1.1	Wideband Tracking	6-17
6.4.1.2	Narrowband Tracking	6-18
6.5	Propagation Time Error	6-20
6.6	Total Error	6-20

TABLE OF CONTENTS (continued)

VOLUME 1

Section No.		Page
7	The Millimeter Wave Link	7-1
7.1	The Millimeter Wave Link in Image Transmission	7-1
7.2	Link Budget in Millimeter Wave Space Communications	7-1
7.2.1	Link Budget Analyses for Sensor Satellite to Relay Satellite	7-3
7.2.2	Link Budget Nomographs	7-5
7.3	Millimeter Wave Transmitters	7-5
7.3.1	Millimeter Wave Transmitters for High Data Rates	7-13
7.3.1.1	Group-Delay Distortion in Millimeter Wave Amplifiers	7-16
7.3.1.2	AM-PM Conversion Effects	7-17
7.4	RF Receiver	7-24
7.4.1	Low Noise Millimeter Wave Amplifiers	7-24
7.4.2	Millimeter Wave Mixer Noise Figures	7-26
7.4.2.1	Mixer IF Amplifiers	7-26
8	Data Recording Technology	8-1
8.1	Wideband Recording	8-1
8.2	Mission Considerations in Data Recording	8-1
8.2.1	System Lifetime	8-2

LIST OF ILLUSTRATIONS

VOLUME 1

Figure No.		Page
1-1	Synchronous Relay Satellite Application	1-5
1-2	Communication Parameters of Image Data System	1-6
1-3	Subsystems of a Satellite Image Data Relay System	1-7
2-1	Sensor Satellite to Earth Terminal via Satellite Relay	2-2
2-2	Multiple Sensor Image Information Transmission System	2-4
2-3	Basic Elements of Analog and Digital Data Links	2-5
2-4	Comparison of Analog and Digital Modulation Based on Required Signal Power	2-9
2-5	Theoretical Coherent PSK Performance and Typical System Degradation	2-13
2-6	RF Bandwidths Required for Several Modulation Methods	2-14
3-1	Basic Image-Device to User System	3-2
3-2	Sensitivity and Dynamic Range Characteristics of Camera Tubes	3-8
3-3 (a)	Block Diagram of Self-Scanned Image Sensor	3-10
3-3 (b)	A Proposed Method for Fabricating a Simple Two-Capacitor Photodiode Array	3-10
3-4	Block Diagram Shows Microwave Radiometric Imager Developed by Aerojet-General Corp. for NASA for Nimbus Meteorological Satellite Application	3-12
3-5	Block Diagram of Lunar Orbiter V Side-Looking Radar Experiment	3-13
3-6	Basic Scanning Beam Image Device Parameters	3-16
3-7	Transmission Bit Rate vs Vidicons Employed in Parallel R/O Configuration	3-19
3-8 (a)	Comparison of (A) A Typical Vidicon Photo-Conductive Camera Tube with (B) An Array of Discrete Photo-Sensitive Elements	3-20
3-8 (b)	FET Transistor Video Amplifier of Image Information	3-20
3-8 (c)	Video Amplifier Response for 60 MHz Vidicon	3-20

LIST OF ILLUSTRATIONS (continued)

VOLUME 1

Figure No.		Page
3-9	Film Rate Vs Video Bandwidth	3-21
3-10	Schematic of Multicolor Spin-Scan	3-30
3-11	Spin Scan Block Diagram (Hughes-ESSA System)	3-30
3-12	Multi-Spectral Scanner System	3-34
4-1	Circuit System Performing the Coding and Decoding Functions in High Speed Digital Communications	4-2
4-2	The Technique of Converting an Analog Signal into a Digital Signal by Sampling, Quantizing, and Coding - after D. F. Hoth	4-4
4-3	Digital Data Rate vs Sensor Bandwidth (2.8 samples/ video cycle)	4-7
4-4	Two-Bit Delta Modulator, Basic Operation	4-13
4-5	Response of the Two-Bit Modulator to a Step Input	4-15
4-6	Delta Modulation Using Double Integration with Prediction	4-16
4-7	Characteristics of Exponential Integration	4-17
4-8	Spilker - Luby Variable Slope A/D Converter	4-20
4-9	Increase in A/D Conversion Slope Due to Edge Effects	4-20
4-10	Delta Modulator Block Diagram	4-21
4-11	PN Test Generator	4-26
4-12	Delta Demodulator Block Diagram	4-29
4-13	System for Multiplexing Individual Delta Modulators	4-32
4-14	Dual System for Multiplexing Different Sets of Sensors	4-33
4-15	Multiplexed Sensor Data by Switched Integrators	4-34
4-16	Digital Data Processor Delta Modulator	4-37
4-17	Digital Data Processor Multiple Channel Delta Demodulator	4-39
5-1	Basic Modulator-Demodulator System	5-3
5-2	High Data Rate PSK Modulation Techniques	5-4
5-3	Types of PSK Modulators	5-8

LIST OF ILLUSTRATIONS (continued)

VOLUME 1

Figure No.		Page
5-4 (a)	Biphase Modulator - Single Pole Single Throw Diode Mode	5-11
5-4 (b)	Biphase Modulator - Variable Reactance Diode	5-11
5-5 (a)	Switching Diode Curve	5-13
5-5 (b)	Diode Equivalent Circuits	5-13
5-6	Philco-Ford Double Resonance Switch Equivalent Circuits and Biphase Modulator Application	5-14
5-7	Silicon Epitaxial, Resonant Diode Admittance	5-17
5-8	Gallium Arsenide Diode Admittance	5-17
5-9	Spectrum of 100 Mbps Biphase Signal	5-18
5-10	RF Amplitude at Biphase Modulated Signal	5-19
5-11	Trapezoidal Pulse Waveform	5-21
5-12	Mismatched Filter Losses Versus Rise-Time-to-Pulse-Duration Ratio; (a) Worst Case, and (b) Average	5-23
5-13	One Type of Quadriphase Modulator Using Signal Combiner and Biphase Modulator	5-24
5-14 (a)	Crystal Oscillator Source	5-26
5-14 (b)	Avalanche Diode Source	5-26
5-15	Modulating Diode Driver	5-28
5-16	Differential Encoder and Decoder	5-30
5-17	Timing Diagram for Differential Encoding and Decoding	5-31
5-18	Quadriphase Differential Encoder	5-33
5-19	Differentially Coherent Biphase Demodulator	5-35
5-20	Waveforms of a Differentially Coherent Biphase Demodulator	5-35
5-21	Differential Phase Detector and Timing Recovery Circuit	5-37
5-22	Differentially Coherent Quadriphase Demodulator	5-39
5-23	Data Decision Circuitry	5-42
5-24	Bit Synchronizer	5-44

LIST OF ILLUSTRATIONS (continued)

VOLUME 1

Figure No.		Page
5-25	Comparison of Theoretical Performance of Quadriphase PSK with Coherent and Differentially Coherent Detection	5-46
5-26	Carrier Reconstruction Loop	5-48
5-27	Coherent Biphase Demodulator	5-50
5-28	Average Degradation as a Function of Phase Error	5-57
5-29	Physical Implementation of Carrier Reconstruction	5-60
5-30	Phasor Diagram of Received Quadriphase Signal	5-60
5-31	Probability of Bit Error vs E_b/N_o	5-63
5-32	Error Probability of System Which Uses a Squaring Loop	5-65
5-33	System Block Diagram	5-66
5-34	Quadriphase BER Performance with Noisy Oscillator ($\beta = 1$)	5-68
5-35	Quadriphase BER Performance with Noisy Oscillator ($\beta = 5$)	5-69
5-36	QPSK Performance Degradation (from Ideal) vs E_b/N_o for Various Values of β and σ^2 FM	5-70
5-37	Biphase Error Rate Performance as a Function of E/N_o and σ^2 FM ($\beta = 5$)	5-71
5-38	Biphase BER Degradation from Ideal ($\beta = 5$)	5-73
5-39	Phase Noise For Various Oscillators	5-75
5-40	Straight Line Approximation to Curves of Figure A-1	5-76
5-41	Matched Filter CKT	5-80
5-42	Matched Filters	5-82
5-43	Matched Filter Detector	5-83
5-44	Matched Filter Sample Dump Timing Diagram	5-84
5-45	Basic Bit Synchronizer	5-86
5-46	Statistical Probability of Bit Error as a Function of the Ratio Between Peak Signal and RMS Noise (with the Noise Bandwidth Referenced to One-Half the Bit Rate)	5-87

LIST OF ILLUSTRATIONS (continued)

VOLUME 1

Figure No.		Page
5-47	Model of Data System Bit Energy Loss and Data Filter Loss and Data Filter Mismatch Loss	5-92
5-48	Mismatched Filter Loss Compared with Matched Filter (Rectangular Pulse Input)	5-94
5-49	Comparison of Mismatched Filters with Intersymbol Interference	5-96
5-50	SNR Degradation vs Modulator Rise Time for No Bandlimiting	5-97
5-51	System Degradation vs Receiver Bandwidth Limiting for Modulator Rise Time = 0	5-98
5-52	System Degradation vs Receiver Bandwidth Limiting Modulator Rise Time is 20% and Receiver Predetection Filter is 5-Pole Chebychev .1 dB Ripple	5-99
5-53	Effects of Variation in Sampling Time	5-101
5-54	Effects of Variation in Sampling Time	5-102
5-55	Data Filter E/N_o Degradation for Constant Bit Error Probability	5-103
5-56	Response of Integrate and Dump Filter to a Trapezoidal Waveform	5-105
5-57	Zone for Correction Reception in MPSK Signalling	5-107
5-58	Areas of Applicability of Logic Devices for Symbol Decision in MPSK Data Detection	5-108
5-59	QPSK Lowpass System Model Showing Transmission Filter and Data Filters (Quadrature Channel Bit Detection)	5-111
5-60	2-Pole Butterworth Data Filter Detection of Coherent QPSK and BPSK	5-111
5-61	Single-Pole RC Data Filter Detection of Coherent QPSK and BPSK	5-111
5-62 (a)	Bandwidth-Limiting Degradation of QPSK and BPSK Signals	5-112
5-62 (b)	Mistuned-Boradband Filtering	5-112
5-62 (c)	Phase Distortion with I&D Detection	5-112
5-62 (d)	Cascade Filter Degradation of QPSK and BPSK Signals	5-112
5-63	Bit Synchronizer Timing	5-116
5-64	Effect of Transmission Channel Filter on Biphase PSK Bit Error Rate	5-117

LIST OF ILLUSTRATIONS (continued)

VOLUME 1

Figure No.		Page
6-1	Satellite Configuration	6-3
6-2	Angular Uncertainty vs Position Uncertainty and Altitude	6-5
6-3	Angular Uncertainty vs Position Uncertainty and Synchronous Satellite Separation; d	6-6
6-5	Dwell Time versus Signal/Noise, $P_d = 0.98$	6-12
6-6	Loop SNR versus $\Delta\dot{\omega}$ max/ωn^2	6-14
6-7	Tracking System, Block Diagram	6-16
6-8	Simplified Wideband Tracking Receiver	6-19
6-9	Phase-Lock Angle Tracking Receiver, Block Diagram	6-19
7-1	Basic Millimeter Wave Data Transmission Link	7-2
7-2	Antenna Gain Versus Antenna Diameter	7-7
7-3	Space Loss versus Distance at Millimeter Waves	7-8
7-4	Link Budget Nomograph at 35 GHz	7-9
7-5	Link Budget Nomograph at 60 GHz	7-10
7-6	Link Budget Nomograph at 75 GHz	7-11
7-7	Link Budget Nomograph at 94 GHz	7-12
7-8	Maximum Power vs Frequency for Semiconductor Diodes	7-14
7-9	Frontier of Microwave and Millimeter Wave Power Generation - After J. Osepchuck	7-15
7-10	TWT Drive Power as a Function of Normalized Time for a) a 0° Phase Shift, b) a 90° Phase Shift, and c) an 180° Phase Shift	7-18
7-11	TWT Output Power as a Function of Normalized Time for a) a 0° Phase Shift, b) a 90° Phase Shift, and c) an 180° Phase Shift	7-19
7-12	Phase Error Advance as a Function of Normalized Time for: a) 180° Shift Typical TWTA, b) 180° Shift Good TWTA, c) 90° Shift Typical TWTA, and d) 90° Shift Good TWTA	7-20

LIST OF TABLES

VOLUME 1

Table No.		Page
2-1	Signal to Noise Ratio for Analog and Digital Links	2-11
2-2	Comparison of Analog and Digital Link	2-17
2-3	Digital and Analog Data Link Parameters	2-18
2-4	Power and Bandwidth Comparison in Digital Communications	2-20
2-5	Principal Communication Bands	2-22
3-1	Summary of Representative Camera Tube Performance	3-7
3-2	Imaging Sensors	3-27
3-3	Present Spin Scan System Performance	3-32
3-4	Multi-Spectral Scanner System Summary	3-33
3-5	Agricultural Spectral Signatures	3-36
4-1	Electrical vs Optical Analogs	4-6
4-2	Integrator Characteristics	4-18
4-3	Quantizer to Bit-Stream Conversion	4-24
6-1	Acquisition Error Budget	6-9
6-2	Tracking Error Budget	6-21
7-1	Baseline Link Budgets	7-6
7-2	Cross-Talk Sign	7-22
7-3	Millimeter Wave Low Noise Amplifiers	7-25
8-1	Wideband Recording Current Magnetic Tape Technology	8-3
8-2	Current Magnetic Tape Technology	8-3
8-3	Current Film Recording Technology	8-4

SECTION 1

THE COMMUNICATION ASPECTS OF IMAGE TRANSMISSION HIGH DATA RATE

1.1 OBJECTIVES

The objective of this contract is to evaluate suitable signal processing and modulation techniques, and to assess current and projected technology problems and needs associated with the transmission and reception of image type data via millimeter wave relay satellites.

The effort under this contract was directed to the communication, signal handling and technological problems associated with the transmission, reception and processing of image type data originating in a surveying satellite and to be transmitted via high speed digital links using millimeter wave relay satellites. A considerable effort was directed to corresponding methods of implementation, related problems, and technology. For the purpose of this study, Philco-Ford was guided by the following:

a. The data should originate in an Earth orbiting satellite to be transmitted via one or more relay satellites at synchronous orbits.

b. The sources of information should consist of but not be limited to the following:

   1. Linear arrays of detectors with no memory
   2. TV or infrared type imagers
   3. Side looking radars
   4. Radiometers
   5. A collection of low data rate sources

   This data may be generated separately or simultaneously.

c. The millimeter wave relay satellites should operate in the 60 to 100 GHz range, with power requirements less than 100 watts.

d. The baseband information is in the range of 4 to 40 MHz. The high data rate requirements are in the range of 40 to 400 MHz.

e. A bit stream format should be used, comparable to 32 levels or 5 bit quantization.

f. Modulation/coding schemes were restricted to systems with channel to data rate ratios of 6.4 and preferably less.

Philco-Ford furnished the necessary personnel, facilities, services and materials to accomplish the work set forth below:

Item 1 - a. Assess millimeter wave technology for space borne use including power sources, detectors, components, circuit techniques, and wideband capability.

- b. Assess high data rate transmitting and receiving technology, including the processign and recording of wideband information.

Item 2 - Determine and compare suitable candidate modulation/coding schemes for the above data rates and classes of data including use of PSK, FSK, M'ary communication and biorthogonal coding. Determine in each case suitable techniques and problems associated with oscillator stability, timing, synchronization, including data format, frame and bit synchronization.

Item 3 - In relation to Items 1 and 2, determine for candidate systems possible methods of implementation with the aid of block and circuit diagrams and evaluate critically with respect to performance, complexity, current and projected technology.

Item 4 - Assess and evaluate the problems and possible techniques for pointing acquisition and tracking between an earth orbiting vehicle and synchronous satellites whose high gain antennas are on the order of 0.2 degrees and less.

Item 5 - Recommend suitable configurations for transmitting and receiving the above classes of data via high data rate information links. Define problem areas requiring further considerations, and needed technological developments to realize suitable system configurations.

This report presents the results of such investigations into the technologies of (1) image sensor systems, (2) analog-to-digital converters, (3) high speed digital communications, (4) antenna acquisition and tracking, and (5) millimeter wave communications.

The report has been prepared in two volumes.

- Volume 1 describes the system aspects of image data transfer
- Volume 2 describes the technology basis of the image data transfer systems

The data and systems described in this report reflect actual experience on the part of Philco-Ford, including the transmission of video information from Vietnam to Washington utilizing delta modulation A/D conversion of a scanned representation of photographic data developed by a slow scan vidicon. A double link from Vietnam to Hawaii by one IDCSP satellite, and then to Washington from Hawaii via another IDCSP satellite was accomplished using digital communications to represent and transmit pictorial data without deterioration of image information. In addition, an in-house-operated 200 Mbps BPSK and delta modulation system using a millimeter wave transmission link has provided a test vehicle to determine the various problems of high speed image data transmission.

1.2 SENSOR SATELLITE TO RELAY SATELLITE SYSTEMS

A conceptual diagram of an image data transmission system using sensor satellites and synchronous relay satellites is shown in Figure 1-1.

The basic circuits of the transmission system include the sensors which produce the image data; this image data is converted into a digital data stream, the phase modulators and the millimeter wave amplifiers provide for RF transmission of this data stream.

The link includes the path from the sensor satellite to the relay satellite, and from the relay satellite to a ground terminal. More than one sensor may be involved and the broad problem is one of achieving sensor data of proper bit error rate and pictorial quality.

This report primarily addresses the sensor satellite to relay satellite system, but the overall system of transferring imagery data from a sensor complex to a data display or recording system can also be applied to a space to ground system.

1.3 SYSTEM VARIANTS AND COMMUNICATION PARAMETERS

Figure 1-2 shows a partial list of the various system variants and parameters which must be evaluated in determining the effectiveness of the system for translating image data from the sensors to an ultimate data receiver or storage system.

These parameters are listed for the various stages of the system to which they relate, and in general, provide the data of the study tasks listed in paragraph 1.1.

Despite the varied and general nature of the various parameters, in combination they completely define the capability of data relay satellite system to translate image data produced by a sensor satellite to a ground terminal.

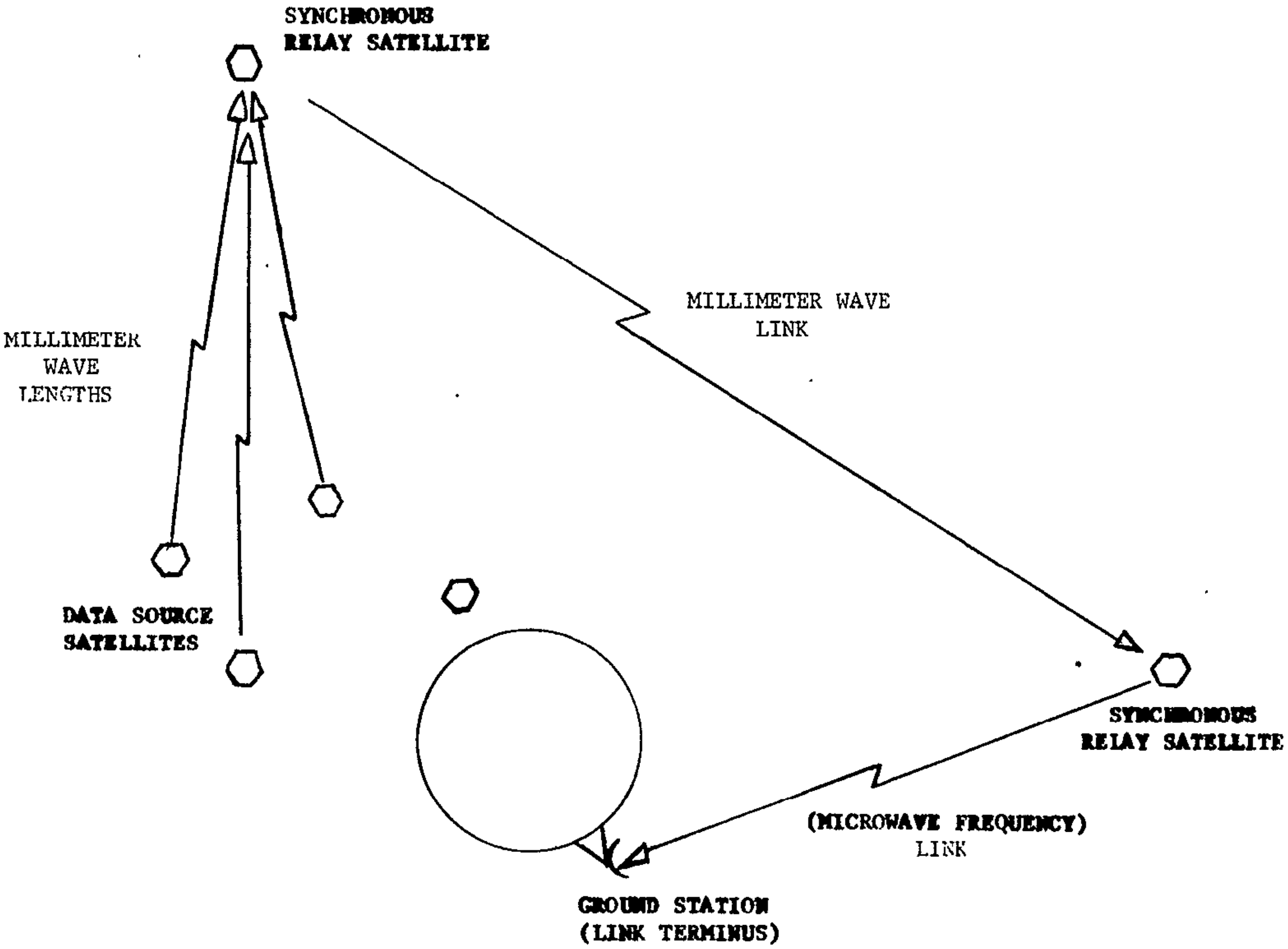

SYNCHRONOUS RELAY SATELLITE
MILLIMETER WAVE LINK
MILLIMETER WAVE LENGTHS
DATA SOURCE SATELLITES
SYNCHRONOUS RELAY SATELLITE
(MICROWAVE FREQUENCY) LINK
GROUND STATION (LINK TERMINUS)

Figure 1-1 Synchronous Relay Satellite Application

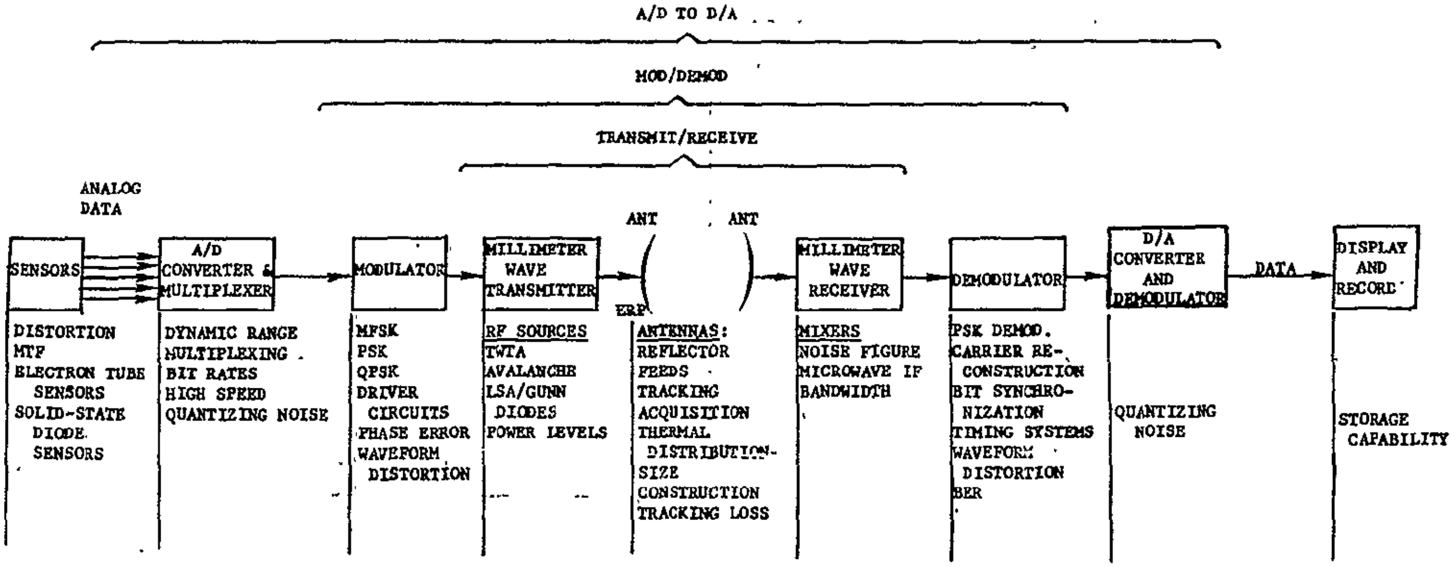

A/D TO D/A
MOD/DEMOD
TRANSMIT/RECEIVE
ANALOG DATA
SENSORS
A/D CONVERTER & MULTIPLEXER
MODULATOR
MILLIMETER WAVE TRANSMITTER
ANT
ANT
ERP
MILLIMETER WAVE RECEIVER
DEMODULATOR
D/A CONVERTER AND DEMODULATOR
DATA
DISPLAY AND RECORD
DISTORTION
MTF
ELECTRON TUBE SENSORS
SOLID-STATE DIODE SENSORS
DYNAMIC RANGE
MULTIPLEXING
BIT RATES
HIGH SPEED
QUANTIZING NOISE
MFSK
PSK
QPSK
DRIVER CIRCUITS
PHASE ERROR
WAVEFORM DISTORTION
RF SOURCES
TWTA
AVALANCHE
LSA/GUNN DIODES
POWER LEVELS
ANTENNAS:
REFLECTOR
FEEDS
TRACKING
ACQUISITION
THERMAL DISTRIBUTION
SIZE
CONSTRUCTION
TRACKING LOSS
MIXERS
NOISE FIGURE
MICROWAVE IF
BANDWIDTH
PSK DEMOD.
CARRIER RE-CONSTRUCTION
BIT SYNCHRONIZATION
TIMING SYSTEMS
WAVEFORM DISTORTION
BER
QUANTIZING NOISE
STORAGE CAPABILITY

Figure 1-2 Communication parameters of Image Data System

1.4 BASIC COMMUNICATION SUBSYSTEMS

The image data communication system of Figure 1-2 can be reduced to a set of subsystems which are shown in Figure 1-3. These subsystems in general define and identify the tasks and basic topics of investigation of the contract.

- The basic communication aspect involves the choice of modulation to be used. Section 2 in both volumes addresses this key consideration.

- Case I identifies the subsystem comprised of the sensor complex which translates optical or radiometric image information into an analog electrical signal which is communicated to a display and recording system. The transmission path for the analog data has the characteristics of an electrical transfer medium which transfers all sensor data to the display and recording system without degradation. See Section 3 in both volumes.

- Case II identifies the analog-to-digital (A/D) and digital-to-analog (D/A) converters which convert the analog information to digital bit streams and back to analog information at bit rates of 40, 100, 200 and 400 Mbps. Here, only performance of the A/D and D/A converters is of importance, with the criterion of performance being the bit error rate of the digital system, and the introduction of noise and impairments into the recovered analog information which is applied to the display and recording system. See Section 4 in both volumes.

- Case III highlights the modulator and demodulator system which converts the output of the A/D converter into a digitally modulated carrier. This digitally modulated carrier is demodulated and a digital bit stream is developed which reproduces the bit stream output of

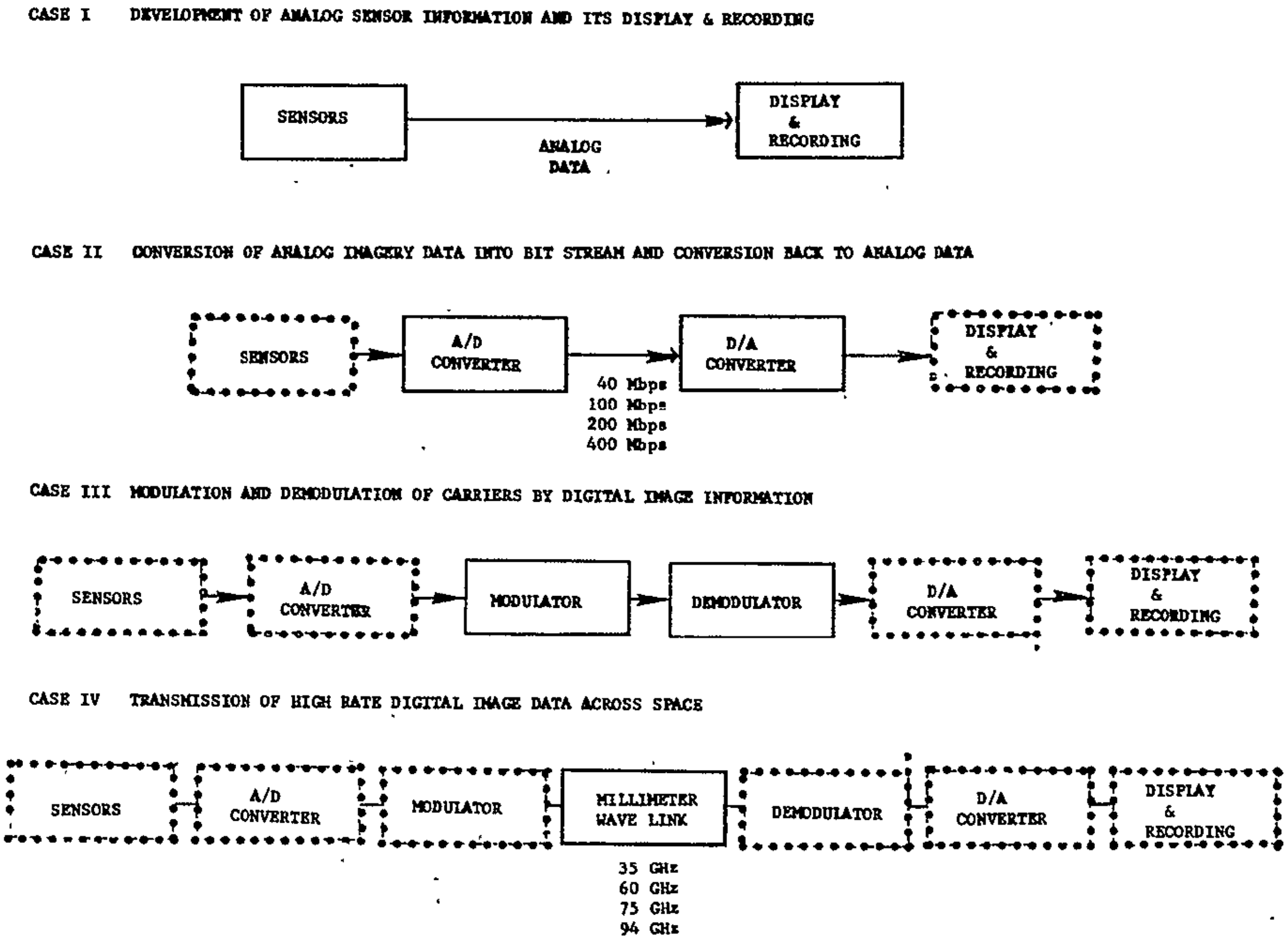

CASE I DEVELOPMENT OF ANALOG SENSOR INFORMATION AND ITS DISPLAY & RECORDING
SENSORS
ANALOG DATA
DISPLAY & RECORDING
CASE II CONVERSION OF ANALOG IMAGERY DATA INTO BIT STREAM AND CONVERSION BACK TO ANALOG DATA
SENSORS
A/D CONVERTER
D/A CONVERTER
DISPLAY & RECORDING
40 Mbps
100 Mbps
200 Mbps
400 Mbps
CASE III MODULATION AND DEMODULATION OF CARRIERS BY DIGITAL IMAGE INFORMATION
SENSORS
A/D CONVERTER
MODULATOR
DEMODULATOR
D/A CONVERTER
DISPLAY & RECORDING
CASE IV TRANSMISSION OF HIGH RATE DIGITAL IMAGE DATA ACROSS SPACE
SENSORS
A/D CONVERTER
MODULATOR
MILLIMETER WAVE LINK
DEMODULATOR
D/A CONVERTER
DISPLAY & RECORDING
35 GHz
60 GHz
75 GHz
94 GHz

Figure 1-3 Subsystems of a Satellite Image Data Relay System

the A/D converter with noise and enhancement of the probability of bit error introduced by the modulator/demodulator system. See Section 5 in both volumes.

- Case IV refers to the millimeter wave link which must transmit the modulated carrier across space and receive it at a terminal such as a data relay satellite. The received digitally modulated carrier with the noise and impairments introduced by the millimeter wave link is then applied to the demodulator. See Sections 6 and 7 in both volumes.

The various sections to follow of this report will follow the philospphy of the various systems as shown in Figure 1-3 and identify the various technologies in each subsystem which relate to the data rates and link frequencies involved.

1.5 BASIC REPORT CONCLUSIONS

The conclusions of the report are summarized as follows in accordance with the objectives listed above.

This report presents more than a "block diagram" description of applicable technology and systems; rather it arrives at specific recommendations of system and circuit configurations relating to delta modulation and quadriphase-modulation signalling.

The basic conclusions are listed in specific and concise terms. They reflect the directions which have been pursued by the various investigations of the report and are supported by the detail and analysis contained in the various sections of both volumes.

The basic conclusions are:

- Optimum types of modulation

  Digital modulation over analog modulation:

    Improved power efficiency and range; improved ability to cope with link distortion and system noise and to interface with computer systems.

- QPSK and BPSK (quadriphase and biphase phase shift keying)

  Quadriphase modulation is the most efficient means of communication from the standpoint of bandwidth; the required bandwidth is equal to the bit rate.

- Optimum Frequency ranges in link

  Millimeter wave frequencies in 35 to 95 GHz range to most efficiently accomodate the 400 MHz bandwidth required of 400 Mbps data rates and to take advantage of small weight and sizes of very high gain antennas.

- High-speed high definition sensors

  TV scanning types of sensors including vidicons, image orthicons, isocons, image dissectors and spin cameras.

    Species of scanning sensors can operate with limiting resolutions to 5000 lines and actual scans of 2000 lines, can store information, can detect a wide variety of visible and infrared optical spectra.

    Signal to noise output of wide band sensor systems as determined by illumination level, dark current, and noise figure of video amplifier.

    Present vidicon technology operates up to 60 MHz bandwidth.

- Optimum A/D conversion techniques for developing bit rates from 40 to 400 Mbps

  Fastest A/D conversion techniques involve phase domain cyclic conversion to develop up to 5 bit words at 1200 Mbps.

  Optimum fast A/D conversion techniques for image transmission with high redundant pictorial content - delta modulation.

    Optimum from the standpoint of circuit simplicity - 2-bit delta modulation at 2.8 bits/sample

    Optimum from the standpoint of image detail - 3-bit delta modulation.

    200 Mbps delta modulation can A/D convert and time multiplex the analog outputs of fast scanning-rate sensors.

    In a link limited data link, 50 MHz baseband bandwidth can be represented by 280 Mbps 2-bit delta modulation.

- Optimum types of PSK modulation techniques for 40-400 Mbps bit rates

  Diode - switched circulator type QPSK modulator for highest speed, simplest circuit, and best quadrature relationships between the four phase vectors.

  Parameters limiting speed and bit rates

    Diode switching speed

    Loss of higher harmonics of switching waveform leading to truncated wave forms.

    Detection loss up to 2 dB resulting from waveform degradation.

- Optimum types of PSK demodulator techniques for 40-400 Mbps bit rates

  Differentially coherent demodulator

    Simplest demodulation system

    For error rates of 10^{-4}, signal to noise ratio required is 0.75 dB higher than required of coherent system.

    Ideally adapted to millimeter waves used by BTL in 50 GHz link.

  Coherent demodulator

    Uses carrier reconstruction, bit synchronization, phase locked oscillator, matched filter. Optimum performance.

    Analysis presented for conditions of noisy reference carrier, and intersymbol interferences due to narrow band filter use.

- Link optimum at 60-90 GHz for high data rate transmission across space

  Uses carbon epoxy light weight antenna to give 60 dB gain with 6 to 8 ft. antenna diameter.

  Link analysis shows QPSK transmission to 500 Mbps in 60 to 90 GHz range possible with margin at BER = 10^{-5} when used with 40-100 watt TWT (transmit) and 10-12 dB mixer (receive) for worst case distances of 24,000 miles.

SECTION 2

COMMUNICATION SYSTEM ASPECTS OF HIGH-DATA RATE IMAGERY TRANSMISSION

2.1 INTRODUCTION

The transmission of image information across space at data rates from 40 to 40 megabits per second requires (1) special modulation techniques which are especially suited for high data rate transmission with low error characteristics, and (2) frequency ranges in which such transmission can optimally occur.

This section will compare analog and digital modulation techniques which can be used for the transmission of image data, and introduce the communication technologies which are related to the sensors, analog-to-digital conversion, modulator/demodulator and millimeter wave links. Section 2 of Volume 2 provides an analysis of the digital modulation techniques described here.

2.1.1 Typical Image-Data Transmission Links

Figure 2-1 shows a candidate image data transmission system, using delta modulation techniques and phase-shift-keying modulation techniques, showing the development of sensor information and the development of a modulated carrier from this information. This modulated carrier is amplified in a traveling wave tube amplifier (TWTA) and propagated across space to a relay satellite using millimeter waves as the propagation medium.

The relay satellite can provide either of two following operational modes:

(a) Frequency translation of the received millimeter wave carrier for propagation to an earth terminal or to another relay satellite.

(b) Data demodulation of the received carrier and modulation of a new carrier for re-propagation to the receiving terminal on each.

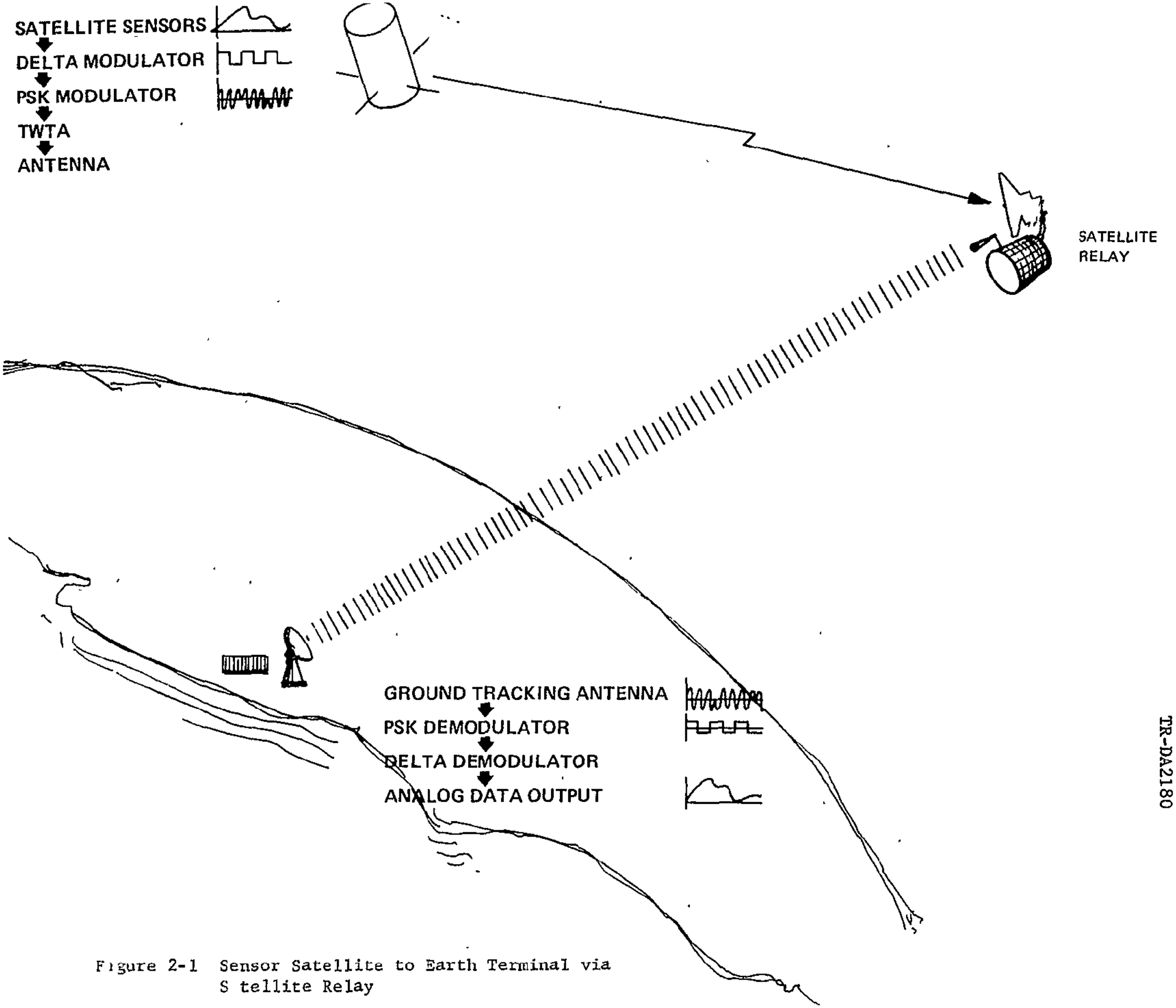

SATELLITE SENSORS
DELTA MODULATOR
PSK MODULATOR
TWTA
ANTENNA
SATELLITE RELAY
GROUND TRACKING ANTENNA
PSK DEMODULATOR
DELTA DEMODULATOR
ANALOG DATA OUTPUT

Figure 2-1 Sensor Satellite to Earth Terminal via Satellite Relay

2.2 MULTIPLE SENSOR IMAGE DATA TRANSMISSION SYSTEM

Figure 2-2 shows a multiple sensor image data transmission system which illustrates the numerous alternative parameters which are legitimate candidates for various system embodiments.

As shown, a variety of sensors of the different scanning types can be multiplexed in time to form a data bit stream having a total bit rate up to 400 Mbps. This multiplexed data bit stream is then modulated onto a carrier using analog or digital modulation techniques, and then transmitted in the millimeter wave frequency range at the typical frequencies of 35, 60, 75, and 94 GHz.

The propagated millimeter wave modulated carrier is then frequency translated to a lower frequency typically in Ku-band for transmission via a space-to-earth link. A demodulator recovers the multiplexed bit stream, and a data demultiplexer clocked by locally reconstructed synchronizing information produces analog sensor data waveforms.

2.2.1 Recommended Modulation Technique

Candidate modulation techniques which will be discussed in the next section will include both analog and digital techniques. However, phase shift keying in combination with delta modulation analog-to-digital conversion will be shown to more optimally meet the transmission needs of wide-band width image sensors considered singly or in sets which are mutliplexed into a common data channel.

2.3 COMPARISON OF ANALOG AND DIGITAL MODULATION TECHNIQUES

Analog video signals can be transmitted from their point of collection (sensor aircraft) to a data recovery point (surface terminal) using either an analog or a digital data link. The basic elements of these two types of data links are shown in Fig. 2-3.

Analog data links are more commonly used at present because of the availability of equipment. However, for high data rate image transmission applications, digital data links have important advantages over analog

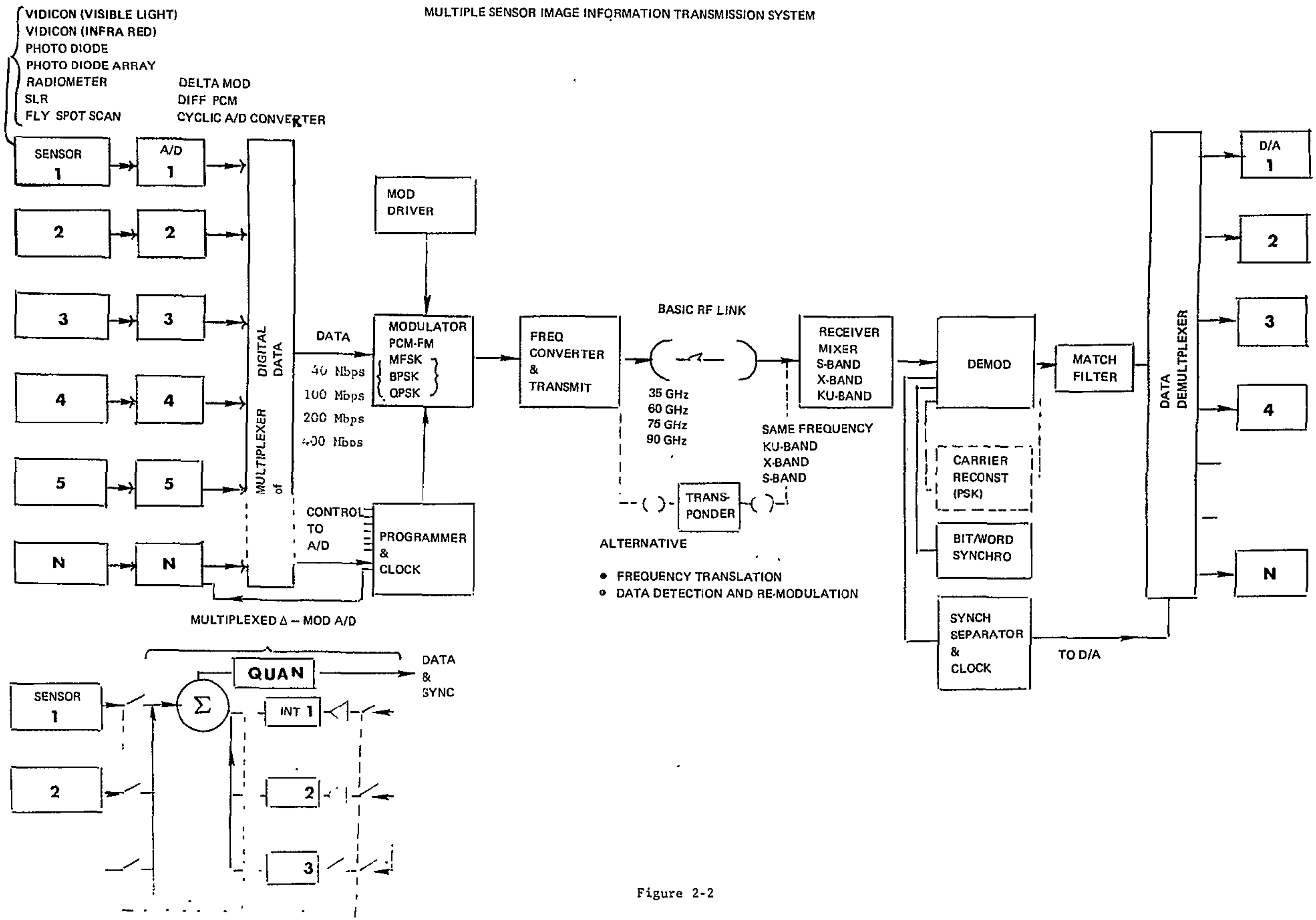
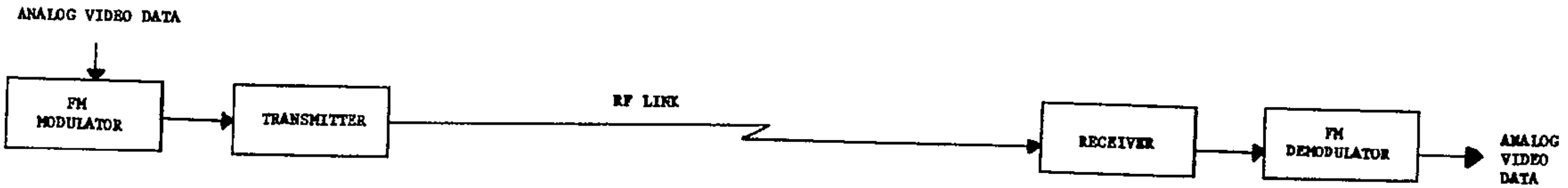
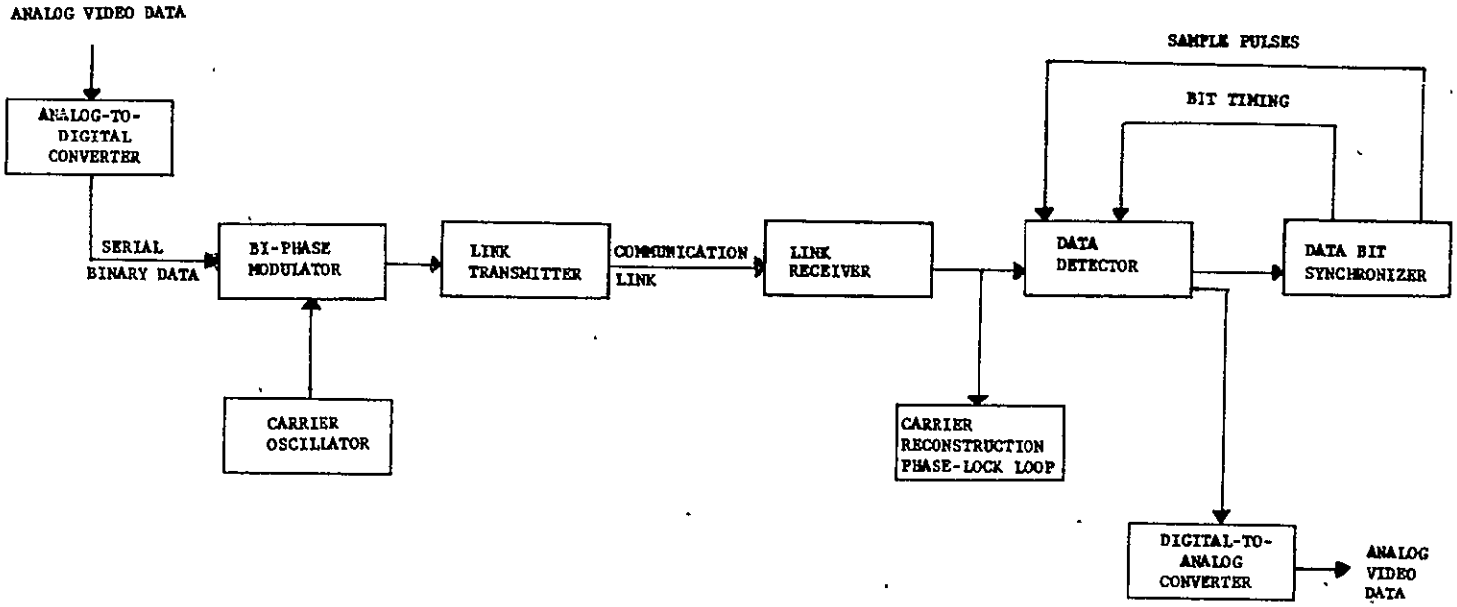
MULTIPLE SENSOR IMAGE INFORMATION TRANSMISSION SYSTEM
VIDICON (VISIBLE LIGHT)
VIDICON (INFRA RED)
PHOTO DIODE
PHOTO DIODE ARRAY
RADIOMETER
SLR
FLY SPOT SCAN
DELTA MOD
DIFF PCM
CYCLIC A/D CONVERTER
SENSOR 1
A/D 1
2
3
4
5
N
MULTIPLEXER of DIGITAL DATA
DATA
40 Mbps
100 Mbps
200 Mbps
400 Mbps
CONTROL TO A/D
MOD DRIVER
MODULATOR
PCM-FM
MFSK
BPSK
QPSK
PROGRAMMER & CLOCK
FREQ CONVERTER & TRANSMIT
BASIC RF LINK
35 GHz
60 GHz
75 GHz
90 GHz
TRANS-PONDER
ALTERNATIVE
• FREQUENCY TRANSLATION
• DATA DETECTION AND RE-MODULATION
SAME FREQUENCY
KU-BAND
X-BAND
S-BAND
RECEIVER
MIXER
S-BAND
X-BAND
KU-BAND
DEMOD
CARRIER RECONST (PSK)
BIT/WORD SYNCHRO
SYNCH SEPARATOR & CLOCK
TO D/A
MATCH FILTER
DATA DEMULTPLEXER
D/A 1
2
3
4
N
MULTIPLEXED Δ – MOD A/D
SENSOR 1
2
QUAN
Σ
INT 1
2
3
DATA & SYNC

Figure 2-2


ANALOG VIDEO DATA
FM MODULATOR
TRANSMITTER
RF LINK
RECEIVER
FM DEMODULATOR
ANALOG VIDEO DATA

(a) Analog FM Link


ANALOG VIDEO DATA
ANALOG-TO-DIGITAL CONVERTER
SERIAL BINARY DATA
BI-PHASE MODULATOR
CARRIER OSCILLATOR
LINK TRANSMITTER
COMMUNICATION LINK
LINK RECEIVER
CARRIER RECONSTRUCTION PHASE-LOCK LOOP
DATA DETECTOR
SAMPLE PULSES
BIT TIMING
DATA BIT SYNCHRONIZER
DIGITAL-TO-ANALOG CONVERTER
ANALOG VIDEO DATA

(b) Digital PSK Link

Figure 2-3 Basic Elements of Analog and Digital Data Links

links and are less complex and less expensive because of the rapid developments in integrated digital circuits.

2.3.1 Advantages of Digital Links

Some of the advantages of digital data links relative to analog modulation are given below.

2.3.1.1 Data Regeneration

In some data link applications, it becomes necessary to relay the data through one or more active repeaters. In an analog link, the noise as well as the signal is amplified and retransmitted. Digital data links permit regeneration of the digital data, thus removing the noise. Erroneous data bits will, of course, be present in the regenerated data, but the cumulative bit errors will cause far less degradation than cumulative noise in each receiver.

2.3.1.2 Multipath Rejection

Multiple signal paths can cause degradation of both analog and digital links by causing: (1) signal fading and hence noisy reception, and (2) spurious output signals (ghosting). For analog and digital links, the effect of fading can be countered by increasing the link margin, e.g., more transmitter power. Spurious signals cannot be removed in analog links by increased power since the multiple paths will also be enhanced. In FM reception, good limiter-discriminator design that results in a high capture ratio will discriminate against spurious signals. In a digital data link, the spurious signals are essentially removed when the link margin is sufficient to reduce bit errors caused by fading to an acceptable value.

2.3.1.3 Less Susceptibility to Channel Distortion

It is of critical importance to the transmission of video information to preserve resolution and phase linearity. As long as a specified bit error rate is maintained, a digital data link can cause no video data degradation as a result of transmission link noise or data link non-linearities such as phase distortion caused by microwave component mismatches. Such is not the case in an analog modulation data link where amplitude and phase non-linearities as well as noise and multipath signals throughout the complete data link must be carefully controlled to limit video data degradation.

2.3.2 Quantitative Comparison of Digital and Analog Links for Transmission of Image Data

The comparison of output quality for digital and analog links is difficult because the noise effects are quite different. In analog links, additive noise in the data link receiver appears in the output video as random perturbations on the video waveform. For pictorial data the noise perturbations will degrade the resolution of the original picture. A criterion commonly used for adequate quality pictorial data is a video SNR of 30 dB. This SNR corresponds to 39 dB peak-to-peak signal-to-rms noise ratio for a sinusoidal video signal.

There are two types of noise to contend with in digital data links - additive receiver noise and quantization noise. Additive receiver noise does not appear directly in the video output. It affects quality by causing erroneous bit errors in the data stream. The bit errors in turn cause erroneous shifts in the grey scale of the pictorial output.

For delta modulation, the grey scale will shift by an amount equal to the step size and persist for the duration of the integrator time constant, on the order of 10 bits. The visual effect is a short grey streak in the output picture.

For PCM, the grey scale will shift by an amount corresponding to the PCM bit which was in error. An error in the most significant bit will cause, for example, a shift from black to white; whereas an error in the least significant bit will cause a shift of only one level of grey. The effect on the output picture is a spotting, or "salt and pepper", effect.

For either delta modulation or PCM, the additive receiver noise does not alter the picture resolution unless the bit errors become excessive. Evaluation of a considerable amount of pictorial data indicates that a BER of 10^{-4} or better results in a negligible effect on output quality. Adequate pictorial data transmission is still possible at BER's as high as 10^{-3}.

The second noise effect in digital links is quantization noise. Quantization noise appears in the video output as discrete steps in the picture rather than producing an exact replica of the continuum of values in the original video waveform.

For PCM the quantization noise is directly related to the number of bits per sample by the following:

$$\text{Quantization SNR} = L^2 = 2^{2n}$$

where L = number of quantization levels

n = number of PCM bits per sample

For 5-bit PCM the ratio of rms signal to rms quantization noise is 30.1 dB.

For delta modulation, the quantization noise is determined by the size of the delta step. In 2-bit delta modulation, the size of the small step determines the quantization noise. (The large step size is chosen to provide the required transient response.) Choosing a small step equal to 1/100 of the video dynamic range results in a quantization SNR of 40 dB.

2.3.2.1 Signal Power Requirements of Analog and Digital Links

The signal power required for analog and digital links can be determined on a relative basis by using comparable criteria for output quality and Fig. 2-4. For a digital link, a BER of 10^{-4} is specified, while an

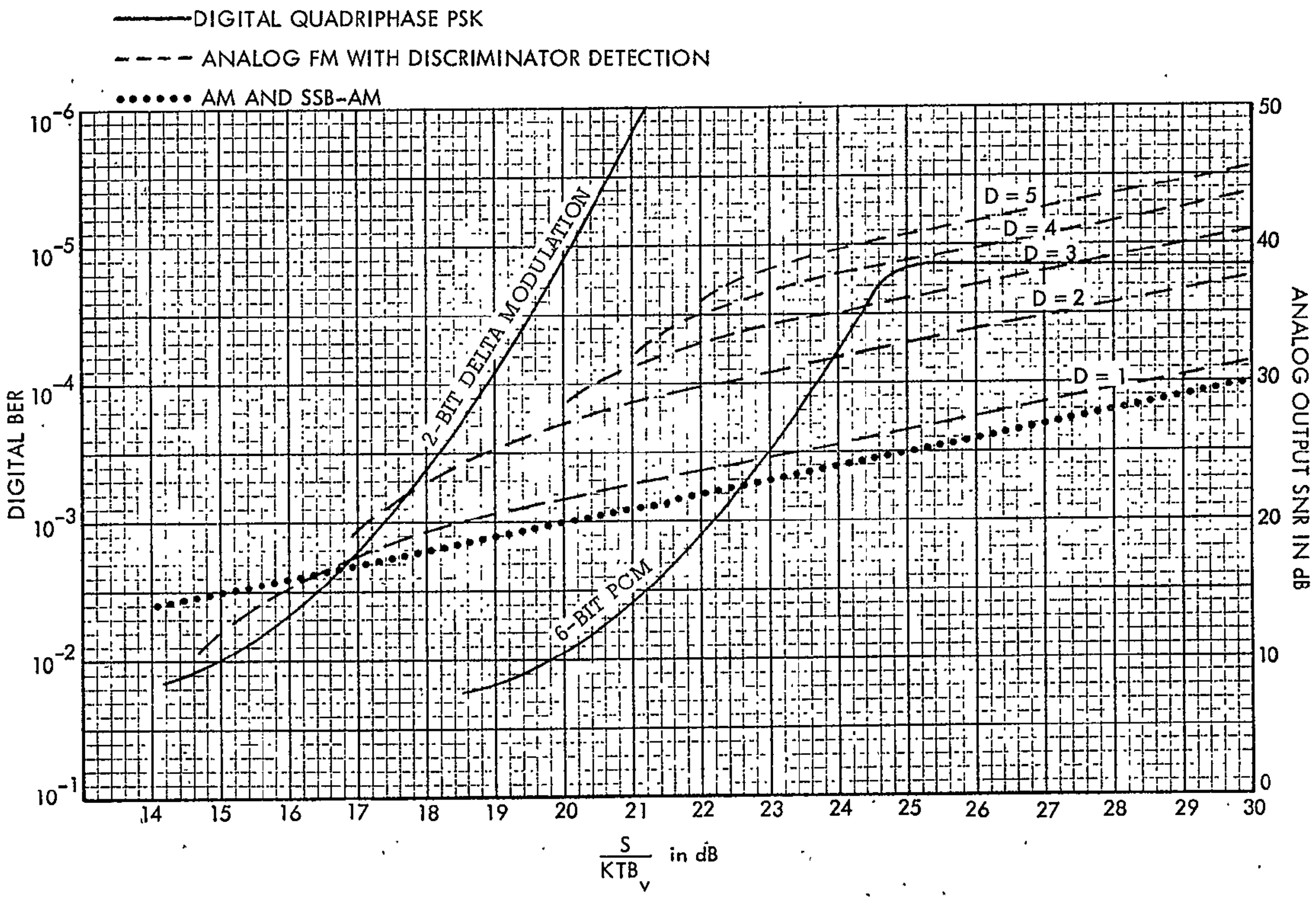


Figure 2-4 Comparison of Analog and Digital Modulation Based on Required Signal Power

analog link output SNR of 30 dB (39 dB peak-to-peak, signal-to-rms noise) will be required.

The formulas used to generate the curves shown in Fig. 2-4 are listed in Table 2-1.

The abscissa used in Fig. 2-4 is the signal-to-noise ratio relative to a bandwidth equal to the video bandwidth B_v. This is an artificial quantity because the input SNR is actually measured in the RF bandwidth B_{RF} rather than the video bandwidth. S/KTB_v was chosen rather than S/KTB_{RF} because B_v is independent of the type of modulation, thus permitting a normalized measure of input SNR for all types of modulation.

S/KTB_v in dB can be converted to the more familiar S/KTB_{RF} by subtracting $10 \log (B_{RF}/B_v)$ for the modulation method of interest.

2.3.2.2 Non-Ideal Digital Information Transmission

Digital information quality is characterized by the error rate as a function of the signal energy per infomation bit (E) divided by the noise power density (N_o). In equation form the signal-to-noise density ratio (S/N_o) may be expressed as

$$S/N_o = (E/N_o)b$$

where b = binary data rate (bit rate)

E/N_o is specified for a particular error rate

From this equation for S/N_o, it is a simple matter to translate an error rate into the value of S/N_o required in the communication link.

A factor must be added to S/N_o to account for non-ideal performance in a practical system. The true system degradation is a complex function of many parameters such as modulator rise time, power amplifier AM-to-PM conversion, phase and amplitude nonlinearities, and many other factors. Detailed system analyses at Philco-Ford have indicated that, for systems

Table 2-1 a

Signal to Noise Ratio for Analog and Digital Links

Analog Data Links (AM)	**Analog Data Links (Wideband FM)**	**Digital Data Links (Biphase or Quadriphase PSK)**
$SNR_{out} = \begin{cases} \frac{S}{KTB_v} = \frac{2S}{KTB_{RF}}, & \text{AM suppressed carrier} \\ \frac{S}{KTBv}, & \text{SSB-AM} \end{cases}$ where $B_{RF} = 2B_v$ for AM $= B_v$ for SSB-AM J. J. Downing, Modulation Systems and Noise, Prentice-Hall, New Jersey, 1964, P. 99.	$SNR_{out} = \frac{3}{2} D^2 \left(\frac{B_{RF}}{B_v}\right)\left(\frac{S}{KTB_{RF}}\right)$ $= \frac{3}{2} D^2 \left(\frac{S}{KTB_v}\right)$ where $B_{RF} = 2\,(D + 1)\,B_v$ (Carson's Rule) D = deviation ratio $\left(\frac{S}{KTB_{RF}}\right)$ threshold $= 10$ $\left(\frac{S}{KTB_v}\right)$ threshold $= 20\,(D + 1)$	$BER = \frac{1}{\sqrt{2\pi}} \int_{\sqrt{\frac{2E_b}{N_o}}}^{\infty} e^{-x^2/2}\, dx$ $\frac{S}{KTB_v} = a\left(\frac{E_b}{N_o}\right)\left(\frac{R_b}{B_v}\right) = 2\left(\frac{E_b}{N_o}\right)\left(\frac{5.6\,B_v}{B_v}\right) = 11.2\left(\frac{E_b}{N_o}\right)$ where $S = E_b R_b$ = signal power $KTB_v = N_o B_v$ = noise power in the video bandwidth, B_v R_b = data bit rate $= 5.6\,B_v$ for two-bit Delta Modulation $= 16.8\,B_v$ for six-bit PCM $a = 2$ = ratio of required E_b/N_o for practical demodulation to theoretical E_b/N_o (equipment margin) A. J. Viterbi, Principles of Coherent Communication, McGraw-Hill, New York, 1966, P. 191.

TABLE 2-1 b

MODULATION TECHNIQUE	PROBABILITY OF BIT ERROR	BANDWIDTH
GENERAL RELATIONSHIPS	$\Pr\{\text{bit error}\} = \frac{2^{k-1}}{2^k-1} \Pr\{\text{symbol error}\}$ $= \frac{M}{2(M-1)} \Pr\{\text{symbol error}\}.$	
MFSK (NONCOHERENT)	$\Pr\{\text{symbol error}\} = \frac{\exp\left(-\frac{E}{N_o}\right)}{M} \sum_{j=2}^{M} (-1)^j \binom{M}{j} \exp\left(\frac{E}{jN_o}\right)$	$Bw = \frac{M}{T} = \frac{M}{T_b \log_2 M} = \frac{2^k}{kT_b}.$
PSK (COHERENT)	$\Pr\{\text{bit error}\} = \Phi\sqrt{2E/N_o}$	The transmitted power spectrum of PSK is of the form $\left[\frac{\sin \pi T_b(f - f_{carrier})}{\pi T_b(f - f_{carrier})}\right]^2.$ The bandwidth between the principal nulls is $2/T_b$ cps.
PSK (DIFFERENTIALLY COHERENT)	$\Pr\{\text{bit error}\} = \frac{1}{2} \exp\left(-\frac{E}{N_o}\right).$	$Bw = \frac{2}{T_b}.$
QPSK	$\Pr\{\text{symbol error}\} = 2\Phi\left(\sqrt{\frac{E}{N_o}}\right) - \Phi^2\left(\sqrt{\frac{E}{N_o}}\right)$ $= 2\Phi\left(\sqrt{\frac{2E_b}{N_o}}\right) - \Phi^2\left(\sqrt{\frac{2E_b}{N_o}}\right).$ The Pr{bit error} is given by $\Pr\{\text{bit error}\} = \Phi\left(\sqrt{\frac{E}{N_o}}\right) = \Phi\left(\sqrt{\frac{2E_b}{N_o}}\right).$	for M = 4, i.e., QPSK, the required bandwidth is $1/T_b$
MPSK	$\Pr\{\text{symbol error}\} \doteq 2\Phi\left(\sqrt{\frac{2E}{N_o}} \sin(\pi/M)\right).$ For $M = 4$ and $\sqrt{E/N_o} \gg 1$:	$Bw = \frac{2}{T_b \log_2 M}.$
M-ARY CODED PSK (BI-ORTHOGONAL CODING)	$\Pr\{\text{symbol error}\} = 1 - \int_{-\sqrt{2E/N_o}}^{\infty} \frac{\exp\left(-\frac{x^2}{2}\right)}{\sqrt{2\pi}}$ $\cdot\left[\int_{-x-\sqrt{2E/N_o}}^{x+\sqrt{2E/N_o}} \frac{\exp\left(-\frac{y^2}{2}\right)}{\sqrt{2\pi}}\, dy\right]^{(M/2)-1} dx$	$Bw = \frac{2n}{kT_b}.$
PCM/FM (DISCRIMINATOR DETECTION)	$\Pr\{\text{bit error}\} = \Phi\left(1.1\sqrt{\frac{E}{N_o}}\right).$	$Bw \doteq \frac{1}{T}.$

M = number of symbols in a signaling alphabet

$k = \log_2 M$ = number of bits of information in an M-ary symbol word

n = number of binary digits in a binary code word

E = symbol energy in watt-seconds

E_b = energy per bit in watt-seconds

T = transmitted symbol duration in seconds

T_b = bit duration in seconds = $(\text{information rate})^{-1}$

N_o = one-sided noise spectral density in watts/cps

$$\Phi(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^{\infty} (-\tfrac{1}{2} y^2)\, dy$$

Bw = required transmission bandwidth in cps.

similar to the relay application, the degradation can be expected to be about 1.5 dB. Therefore, upon conversion to decibels

$$S/N_o = (E/N_o)b + 1.5 \text{ dB}$$

and the error rate curve of Figure 2-5 may be modified accordingly.

Many tradeoffs exist in apportioning the error rate in a multisegment communication link such as the relay system, especially when regeneration of the digital data is feasible. An error rate of 1×10^{-5} errors per bit of information transmitted is assumed for the relay system based upon the more detailed studies of similar systems at Philco-Ford. One must appreciate that a change of plus or minus one order of magnitude of error rate (1×10^{-4} to 1×10^{-6}) constitutes a ± 1 dB change in the required E/N_o, which may certainly be accommodated in any detailed system design.

For an error rate of 1×10^{-5}, $E/N_o = 9.5$ dB
and

$$S/N_o = 9.5 \text{ dB} + 1.5 \text{ dB} + 10 \log (b) = 11.0 \text{ dB} + 10 \log (b)$$

This is the required S/N_o used in the link calculations.

2.3.2.3 Bandwidths for FM, AM and PSK

The minimum bandwidth and transmitter power for a digital link are achieved using 2-bit Delta modulation for the analog-to-digital conversion and quadriphase PSK for the carrier modulation. A comparison of bandwidths for PSK, FM, and AM is given in Figure 2-6.

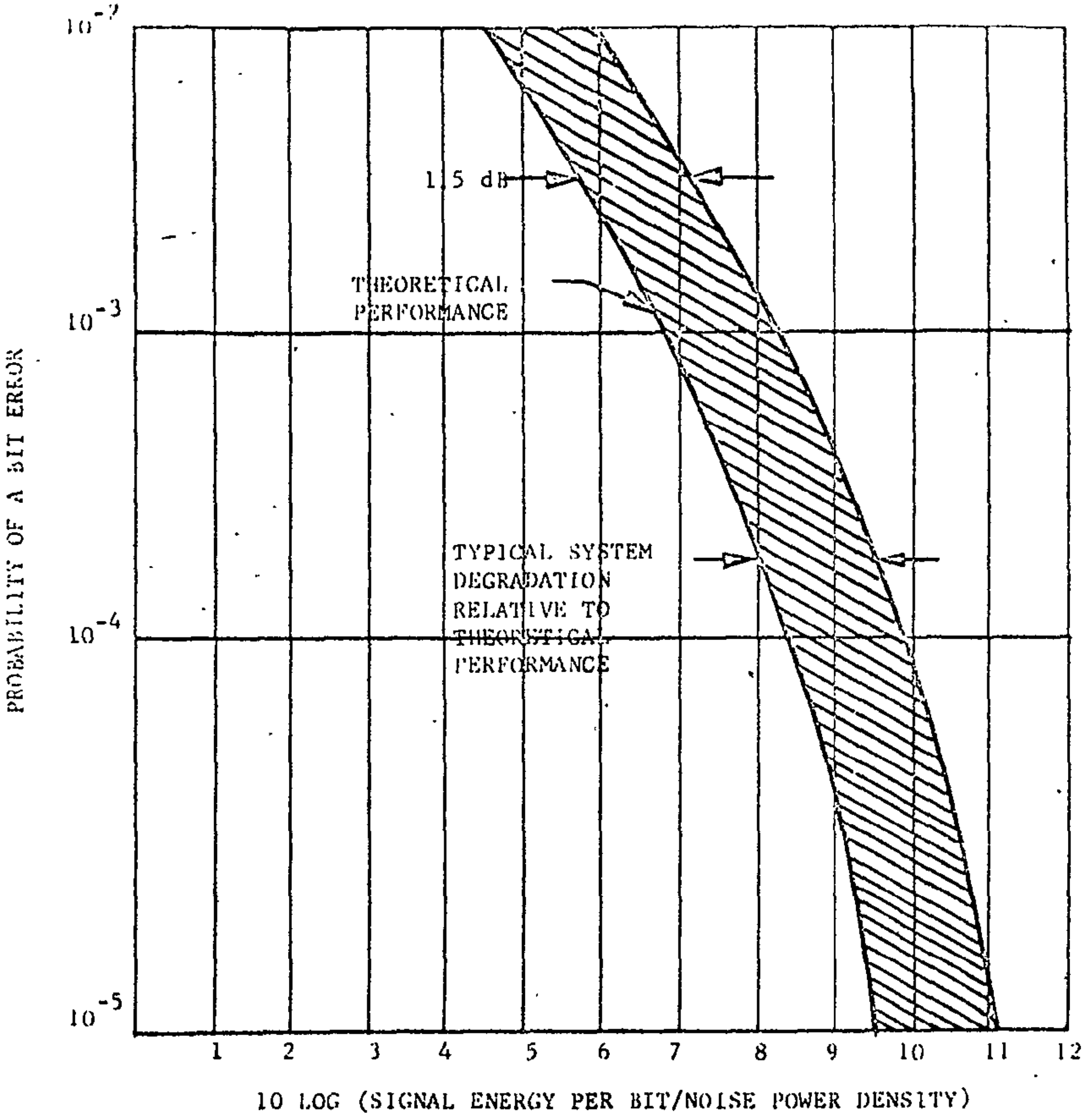
The bandwidth formulas used to generate the curves of Figure 2-6 are:

Digital Links

$$B_{RF} = \begin{cases} 2\,R_b = 11.2B_v & \text{for biphase PSK} \\ R_b = 5.6B_v & \text{for quadriphase PSK} \\ \frac{2}{3}\,R_b = 3.73B_v & \text{for octaphase PSK} \end{cases}$$

where R_b = digital data bit rate

B_v = sensor video bandwidth



1-527A

Figure 2-5 Theoretical Coherent PSK Performance and Typical System Degradation

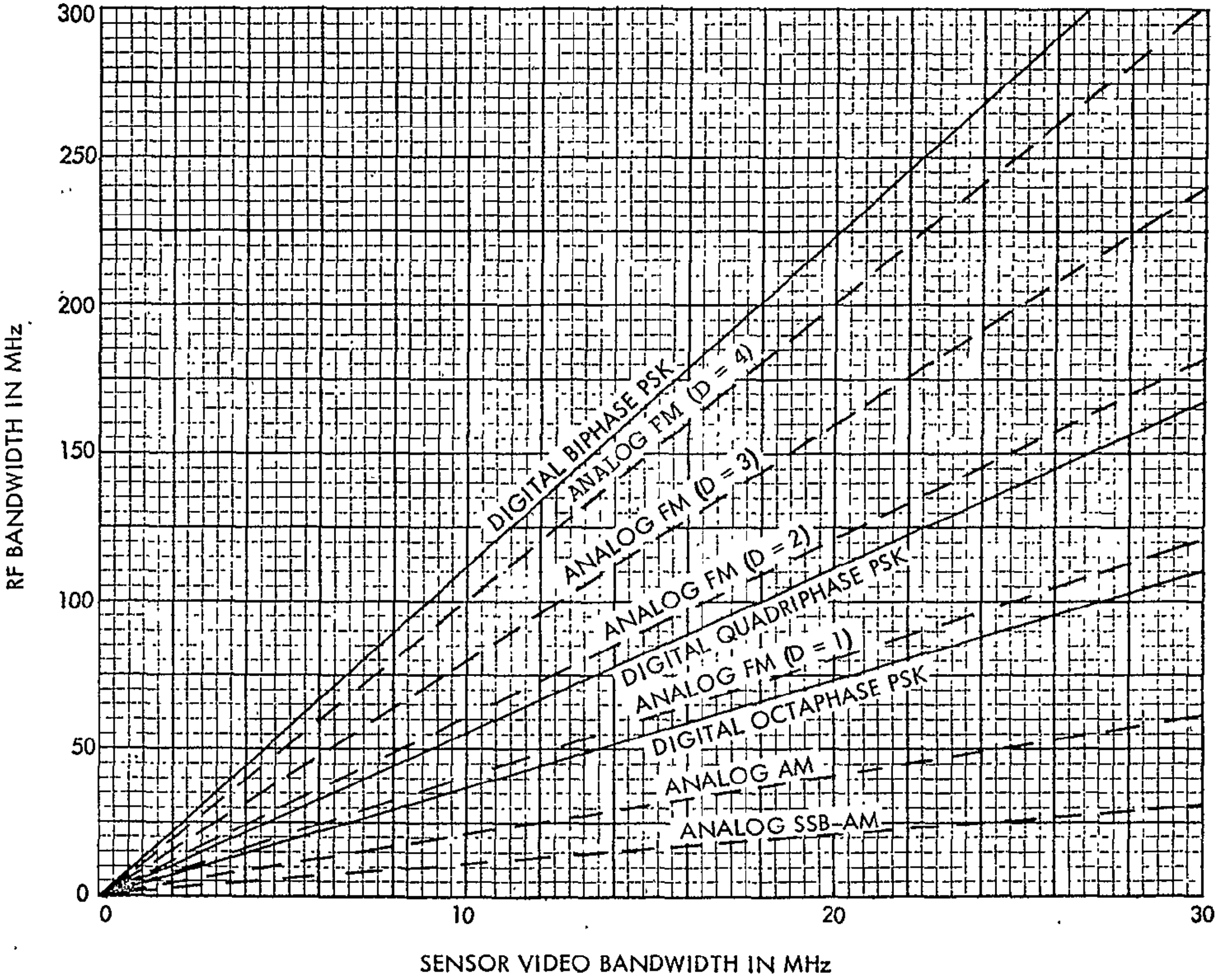


Figure 2-6 RF Bandwidths Required for Several Modulation Methods

Analog Links

$$B_{RF} = \begin{cases} 2\,(D + 1)B_v & \text{for FM} \\ 2B_v & \text{for AM} \\ B_v & \text{for SSB-AM} \end{cases}$$

where D = FM deviation ratio

B_v = sensor video bandwidth

In the above expressions for digital data links, it is assumed that 2-bit delta modulation is used for analog-to-digital conversion with 2.8 samples per cycle. It will be noted that octaphase PSK permits slightly less RF bandwidth than the proposed quadriphase PSK. However, it is more susceptible to noise than either biphase or quadriphase and requires considerably more demodulator complexity.

For the FM analog data link, an RF bandwidth equal to the Carson's rule bandwidth is assumed. To minimize video phase distortion caused by the required sharp cut-off filters, the RF bandwidth for FM video is sometimes chosen to be wider than the Carson's rule bandwidth by factors ranging from 1.4 to 2. On the other hand, RF bandwidths narrower than the Carson's rule bandwidth are possible if a careful trade-off is made between allowable video distortion and required bandwidth restriction.

AM, especially single-sideband (SSB) AM, permits the narrowest RF bandwidth. However, these modulation methods are more susceptible to noise, multipath and channel distortion than either FM or digital PSK and should, therefore, be considered only for applications where extremely wide sensor bandwidths are required.

The use of a threshold extension FM demodulator permits a trade-off between RF bandwidth and transmitter power. Without threshold extension the analog link requires 3.2 dB more power and 7 percent more bandwidth than the best digital link. Increasing the deviation ratio to 4 and allowing for 3.5 dB of threshold extension provides an analog link requiring 2.8 dB less power than digital but 1.8 times more RF bandwidth.

Not included in this comparison is the incorporation of error correction coding in the digital link. Error correction, like threshold extension in the analog links, trades increased RF bandwidth for reduced transmitter power. Coding will reduce the required power by 2 dB to 3 dB at the expense of doubling the RF bandwidth.

A summary of digital and analog data link parameters is shown in Table 2-2 and Table 2-3.

TABLE 2-2

COMPARISON OF ANALOG AND DIGITAL LINK

	Analog Link	Digital Link
Power Efficiency	---	2 x better
Range (without relay)	---	1.4 x better
Range (with relay)	Noise is Commulative	Data Regeneration Removes Noise
RF Bandwidth	Comparable	Comparable
Video Bandwidth (state-of-the-art)	50 MHz	35 MHz (200 Mbps)
Video Bandwidth (under development)	---	70-100 MHz (400-600 Mbps)
Multipath	Fading and Ghosting	Increased Error Rate
Link Distortion	Directly Effects Data	No Direct Effect on Data
Security	None	Encryption
Anti-jam	None	Applicable to Spread Spectrum with Reduced Data Rate
Error Correction	None	Coding or Off-Line Processing
Data Processing	Limited	Computer Processing of Selected Data
Size, Weight, Power	Comparable	Comparable

TABLE 2-3

DIGITAL AND ANALOG DATA LINK PARAMETERS

Link Type	$\frac{S}{KTB_v}$ (dB)	B_{RF}/B_v	Relative Power
Biphase PSK, 2-Bit Delta Modulation	19.0	11.2	Reference
Quadriphase PSK, 2-Bit Delta Modulation	19.0	5.6	0 dB
Biphase PSK, 6-Bit PCM	23.8	33.6	4.8 dB more
Quadriphase PSK, 6-Bit PCM	23.8	16.8	4.8 dB more
FM (D = 2) With Discriminator Detection	22.2	6.0	3.2 dB more
FM (D = 3) With Extended Threshold	19.0	8.0	0 dB
FM (D = 4) With Extended Threshold	16.5	10.0	2.5 dB less
AM, Suppressed Carrier	30.0	2.0	11 dB more
SSB-AM	30.0	1.0	11 dB more

2.4 COMPARISON OF MODULATION TECHNIQUES AT 40 TO 400 Mbps RATES

As the bit rate approaches 500 megabits per second, as shown in Table 2-4 the required bandwidths for transmission exceed gigahertz bandwidths except quadriphase modulation* which requires the same C/kT as binary PSK but only one-half the bandwidth.

Bit error rates for quadriphase modulation which represent phase shift keying at 0°, 90°, 180°, 270° are compared in Figure 2-5 with two other modulation techniques; MFSK (multiple frequency shift keying) and M-ary biorthogonally coded PSK provide maximum communication capacity while quadriphase modulation QPSK provides bandwidth efficiency as well as good power efficiency.

Figure 2-5 shows that, for the same signal-to-noise ratio, standard quadriphase modulation has a higher power requirement than either MFSK or M-ary biorthogonally coded PSK. However, at half-gigahertz rates, MFSK may be impractical due to the complex switching of matched filters required in the MFSK demodulator in what are referred to as the BPID circuits (bandpass integrate-and-dump); on the other hand, error coding which requires at least a doubling of the bit rate and is a key to M-ary biorthogonally coded PSK may make the use of coding techniques impractical until gigahertz-bit-stream technology is as fully developed as is, say, 50-megabit bit stream technology at the present time.

Accordingly, quadriphase modulation, which will be described later in this paper, is the key candidate for gigahertz-rate modulation for microwave and millimeter wave frequencies.

---

*8-phase, 16-phase modulations require even less bandwidth but have less power efficiency.

TABLE 2-4

POWER AND BANDWIDTH COMPARISON IN DIGITAL COMMUNICATIONS

Digital Modulation Methods	Required Power $BER=10^{-5}$	Required Bandwidth	Required Bandwidth (MHz)			
			40 MBS	100 MBS	200 MBS	400 MBS
16-ary Coded PSK	0 dB(reference)	4 x bit rate	160	400	800	1600
16-ary MFSK	0.8 dB more	4.5 x bit rate	180	450	1000	1800
Quadriphase	2.5 dB more	1 x bit rate	40	100	200	400
Coherent PSK	2.5 dB more	2 x bit rate	80	200	400	800
PCM/FM	4.6 dB more	2 x bit rate	80	200	400	800
PSK	6.0 dB more	3 x bit rate	120	300	600	1200

The bit error rate relates each type of modulation to the E_b/N_o where E_b is the energy per bit and N_o is the noise power density, KT_s, where K is Boltzman's constant and T_s is the system noise temperature.

The terminal sensitivity factor, C/N_o, and E_b/N_o are related by the bit rate, R_b, as

$$\frac{C}{KT} = \frac{E_b}{N_o} \cdot R_b \ .$$

Thus, as the bit rate increases to, say a half-gigahertz, additional carrier power (C) will be required to maintain a value of 10 Log E_b/N_o corresponding to a particular bit error rate.

2.5 CANDIDATE TRANSMISSION BANDS FOR HIGH DATA RATE IMAGE TRANSMISSION

As indicated in Table 2-5, the principal communication bands below 10 GHz have bandwidths limited to 500 MHz(by international agreement). Such bands can barely accommodate the 400 MHz bandwidth required for quadriphase modulation and, certainly, 16-ary MFSK would use almost a quarter of the presently occupied microwave frequency band.

It is evident that higher frequency bands from Ku-band to the optical frequencies will be required for 400 Mbps digital carriers. Of particular interest are the bands at: 15, 35, and 94 MHz where minimum atmospheric absorption windows occur; 60 GHz and 120 GHz for exo-atmospheric spacecraft-to-spacecraft communication when privacy or low RFI is required; and the wavelengths of 1.06 and 10.6 microns of the helium-neon and CO_2 lasers, respectively, to take advantage of the high-speed modulator techniques available at 1.06 microns, or the high efficiency/high power space-to-earth capability of the 10.6 micron infrared laser.

Frequency ranges in the 30 to 100 GHz frequency range are the principal candidates for the transfer of imagery information from a sensor spacecraft to a relay spacecraft. As will be discussed in Section 7, the technology

TABLE 2-5

PRINCIPAL COMMUNICATION BANDS

Frequency Band Center	Bandwidth Capability	Use	Bandwidth Required For 400 Mbps QPSK
2.25 GHz	100 MHz	SGLS, Apollo	400
4.0 GHz	500 MHz	COMSAT	400
7.5 GHz	500 MHz	Military SATCOM	400
15.3 GHz	500-1000 MHz	ATS-E/Domestic TV/ Military SATCOM	400
35.0 GHz	1 - 2 GHz	Space-Ground Communications	400
60.0 GHz	1 - 5 GHz	Spacecraft to Spacecraft Communications	400
94.0 GHz	1 - 5 GHz	Radiometry	400
120.0 GHz	1 - 5 GHz	Spacecraft to Spacecraft Communications	400
1.06 Microns	1 - 50 GHz	Light-guide PCM Helium/ Neonlaser	400
10.6 Microns	1 - 50 GHz	Infrared-Space/Earth Communications	400

of millimeter wave power amplification, propagation and reception is now at a stage which makes space links at these frequencies capable of being designed with a high level of reliability and product assurance.

SECTION 3

IMAGE SENSOR TECHNOLOGY

3.1 THE DATA CHARACTERISTICS OF IMAGE SENSORS

Image sensors perform the function, either singly or in groups, of transducing impinging pictorial data radiation in the visible or infrared bands into an analog electrical signal which can then be applied to a recording or display device. The basic image system is shown in Figure 3-1.

Three basic parameters are critical to image systems:

a. MTF: MTF is the modulation transfer function which determines the limiting resolution of the image device.

b. Video Bandwidth: Video bandwidth is the total electrical bandwidth of the image sensor plus video amplifiers, which determines the image information bandwidth which is a function of such parameters as the number of scanning lines, the frame rate, and the time duration an image sensor "looks" at a particular segment of pictorial data. Bandwidth is a separate parameter from MTF and is valid only in an image sensor system having an MTF capable of supplying an information bandwidth which can fill the video bandwidth.

c. Signal-to-Noise: The output of the image sensor circuit, including video amplifiers will include a noise level which will be a function of both sensor characteristics and currents, received light level, and video amplifier noise figure.

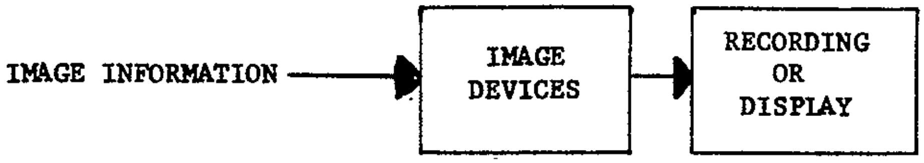


IMAGE INFORMATION
IMAGE DEVICES
RECORDING OR DISPLAY

- MTF
- VIDEO BANDWIDTH
- SIGNAL-TO-NOISE RATIO

Figure 3-1 Basic Image-Device to User System

The total art of image sensing is vast and beyond the scope of this study; however, the three parameters identified above are critical to determining the bandwidth and the signal-to-noise ratio of the analog pictorial information which is converted to digital information in the analog-to-digital converter. The image data transmission system of Figure 3-1 will be discussed in detail in the following paragraphs relative to the types of sensors used and their various characteristics.

3.1.1 General Sensor Requirements

Sensor requirements have been developed in the areas of geology, agriculture, forestry, geography, meterology, and oceanography and can be determined for the most part by the following factors:

- Illumination Level
- Wavelength Region and Bandwidth
- Limiting Resolution
- Scene Characteristics
- Vehicle Altitude
- Sensor System Noise Level

3.1.1.1 Wavelengths of Interest

The various wavelength regions of interest to be processed by the sensors may be divided into three basic categories:

- Visible - 0.4 to 0.7 µ
- Near Infrared - 0.7 to 5 µ
- Intermediate and Long Wavelength Infrared - 5 to 20 µ

The visible spectrum is dominated by diffusely scattered solar radiation. The near infrared spectrum is a transition region from dominance by solar effects to self emission by the earth. The spectrum region beyond 5 µ is essentially terrestrial emission with only indirect solar effects. A sensor designed for use in one of these spectral regions may have a bandwidth considerably narrower than that of the whole region. This will, of course, have an impact on S/N calculations for each particular sensor.

3.1.1.2 Resolution

In order to establish a baseline for sensor resolution, reference is made to the NAS/NRC panels on the Useful Applications of Earth-Oriented Satellites. In the panels on agriculture, geology, et. al., a basic ground resolution of 100 feet was specified for a wide variety of applications. In this case the satellite would be at an altitude of 300 to 500 nautical miles. On the other hand, the "high resolution" radiometric experiment for ATS F&G has a resolution of 6 nautical miles at synchronous altitude. The common factor in these experiments is angular resolution, which is about 0.1 milliradians in each case.

3.1.1.3 Illumination Levels

The full range of illumination levels cannot be treated in the scope of this report. However, as a reference baseline, the earth's illumination at noon is around 10,000 ft.-candles; this can fall to 500 ft.-candles at other times of day and for dark patches of the earth's surface, with an average reflectance at the earth's surface of 10% and an additional 10% reduction in illumination due to spacecraft optics. The actual foot-candles of illumination reaching the sensor will be around 1% of the earth's illumination.

3.2 CANDIDATE PHOTO SENSOR IMAGING DEVICES FOR SPACE APPLICATIONS

Image devices, which are currently available for space-borne image sensor applications, comprise a large number of different types and species which includes the following:

a. Scanning-beam image tubes:

- Vidicons
- Iricons (infrared vidicons)
- Image orthicons
- Image intensifier orthicons
- Isocons
- Image dissectors
- Scotoscopes

b. Mosaics or clusters of solid state or electron beam photosensitive cells:

- Self-scanned solid state image sensor
- Spin-camera
- Radiometric sensors
- Scanning radiometers

3.2.1 Functional Classes of Image Tubes

There are two functional classes of image tubes. One, termed an image intensifier, produces an illuminated image on its output surface which is brighter than the image on its input photosensitive surface. The other class, termed a camera tube, also requires an optical input image on its photosensitive surface, but the output is a video signal obtained by sequentially scanning its image by an electron beam. The output of the camera tube may be amplified in a separate image intensifier tube or the image intensifier function can be incorporated inside the single image tube. Electron multiplication can be utilized at each element of the image, or it can be applied within the tube to the video signal obtained from scanning.

3.2.2 Camera Tube Types

The candidate camera tubes are the vidicon, the image orthicon, and the image intensifier orthicon (which is an image orthicon preceded by one or more stages of image intensification and the Iricon (Infra Red Vidicon). The first three devices can be available for operation in any part of the spectral region from the near ultraviolet to the infrared. The Iricon is also available for operation into the far infrared.

Camera tubes involve scanning beams and have the feature of storage of information from frame to frame. Storage is not provided by image dissectors or solid state photo-sensor clusters or mosaics. Representative features of various camera tubes are listed in Table 3-1.

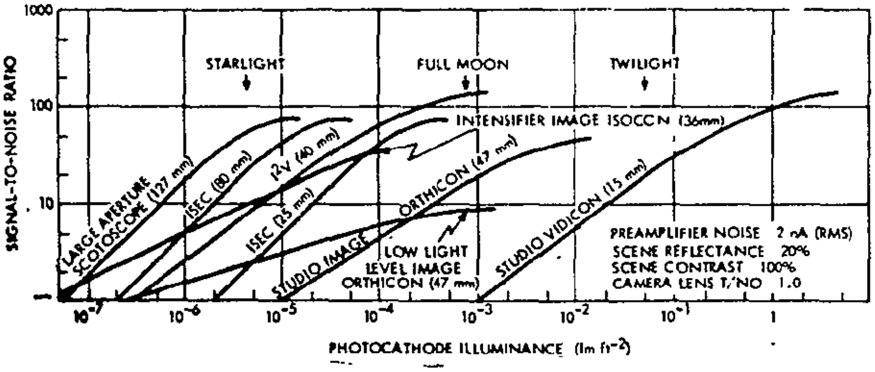
Figure 3-2 gives representative data on various camera tubes including combinations with intensifiers. This table does not necessarily represent the entire current state-of-the-art, and the data are not complete on any one type of tube. However, it does provide a means of comparison between the various camera tubes, particularly with respect to signal-to-noise ratio, and photocathode illuminance. Figure 3-2 indicates the relation between signal-to-noise ratio and scene illumination, assuming a lens speed of T/1.0 and a bandwidth of 4.5 MHz.

TABLE 3-1

SUMMARY OF REPRESENTATIVE CAMERA TUBE PERFORMANCE

CAMERA TUBE TYPE AND SIZE	FACE PLATE ILLUMINATION (OPERATING POINT) FOOT - CANDLES	% AMPLITUDE RESPONSE AT 400 T.V. LINES	LIMITING RESOLUTION T.V. LINES	TYPICAL db SIGNAL TO NOISE RATIO	
Image Orthicon				4.5 MC	9 MC
4-1/2"	2.0×10^{-4}	75	800	55	50
3"	5.0×10^{-4}	50	600	45	40
Return Beam Vidicon					
4-1/2"*	1×10^{-3}	3000 lines 50% response	4000		45 (50 MHz)
Vidicon					
1	0.1 (Average Sensitivity)	60	1000	55	50
1-1/2"	0.1 (Average Sensitivity)	60	1200	58	54
	2.0 (Minimum Lag				
1-1/2" High Sensitivity	0.05 (Average Sensitivity)	45	1000	30	25
3" FPS	0.1	2000 lines 50% response	3000		45 (45 MHz)
Intensifier Image Orthicon					
3"	2×10^{-5} (S-20)	30	450 center 350 edge	30	25

* Experimental O. H. Schade (RCA)



Sensitivity and Dynamic Range Characteristics of Camera Tubes.

Figure 3-2

3.2.3 Self-Scanned Mosaic Image Sensors

Self-scanned image sensors due to P. Weimer[1] et al, produce a video signal without the help of an electron beam. Figure 3-3 shows the principal parts of a solid-state image sensor. It consists of an array of photosensitive elements, each located at the intersection of mutually perpendicular address strips, which are connected to scan generators and video coupling circuits. Sequential scan pulses to the address strips permit an image to be scanned, and a video signal is produced similar to that generated by a television camera tube.

For image detail comparable to broadcast television, the array must contain hundreds of thousands of picture elements. Although these objectives have not yet been reached, recent progress indicates that self-scanned arrays can eventually replace camera tubes in some applications.

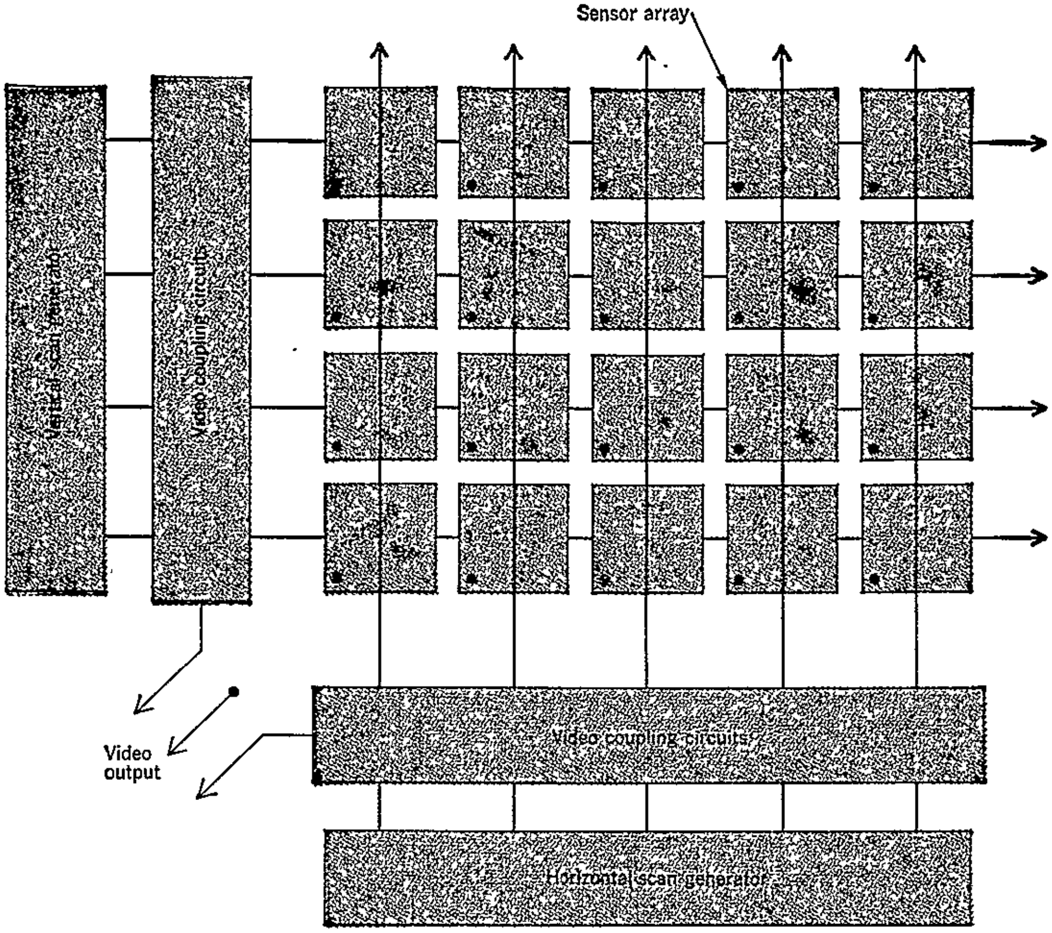
Fabrication of experimental image sensors has utilized silicon technologies that are highly refined. This approach can be justified by the enormous versatility of silicon, which provides, in addition to its integrated circuits, at least four different types of photosensitive elements. These include p-n junction photodiodes, phototransistors, photoconductors, and photovoltaic cells. The intrinsic spectral response of silicon ranges from the visible to the near infrared and it can be extended into the far infrared by impurity photoconduction. Even more important to the sensor application is the fact that silicon junctions can be made with sufficiently high resistance to allow integration of light by charge storage for periods exceeding the normal television scanning periods.

3.2.4 Mosaic Resolution

The image-forming photosensor mosaic is a large device, requiring a panel edge length of about 0.8 inch for each 100 resolution elements along an edge. A panel of a few hundred lines resolution, then, would

---

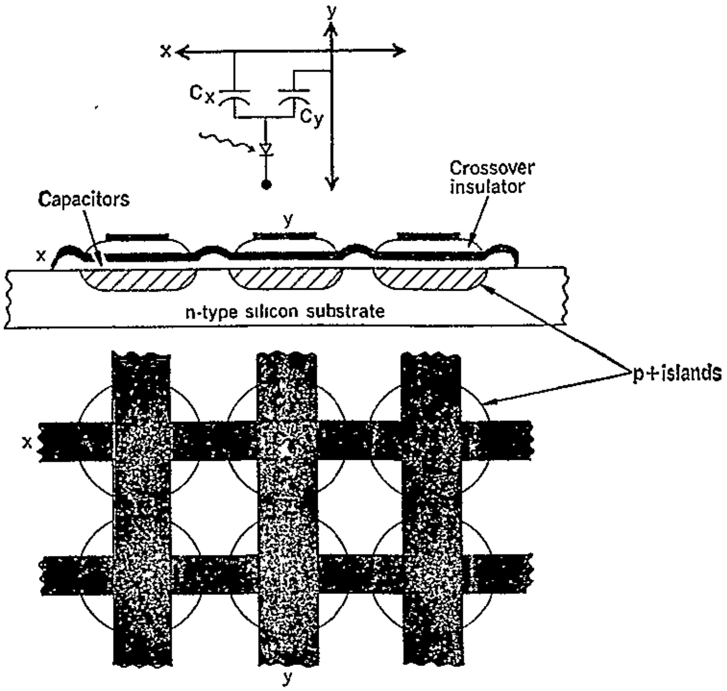
[1] P.K. Weimer et. al, "Multielement Self-Scanned Mosaie Sensors", IEEE Spectrum, March 1969.


Sensor array
Vertical scan generator
Video coupling circuits
Video output
Video coupling circuits
Horizontal scan generator

Block diagram of self-scanned image sensor, showing three alternative methods of coupling out the video signal.

(a)

A proposed method for fabricating a simple two-capacitor photodiode array.


y
x
C_x
C_y
Capacitors
y
Crossover insulator
x
n-type silicon substrate
p+islands
x
y

(b)

Figure 3-3

measure several inches on an edge. Further, it has a longtime constant, requiring about one second to read out each 10^4 elements (100 x 100 elements square).

Multi-element photosensor mosaics with resolution capabilities approaching 600 to 1200 lines - that is, devices which could have from 360,000 to over 1,000,000 resolution elements and which could seriously compete with vidicons and image orthicons. Such mosaics are beyond the present state-of-the-art which is presently at the 350 element row/350 element line but the Weimer technique shows promise of some day making such a system useful for selected applications.

3.3 MICROWAVE RADIOMETER AND SIDE-LOOKING RADAR SENSORS

Two types of imaging sensors which do not use camera tubes or photosensitive elements are the microwave radiometer sensor and the side-looking radar.

The microwave radiometer provides thermal or temperature information, and is widely used in instrumentation for measuring noise temperature, for millimeter wave. The theory of the microwave or millimeter wave radiometer - particularly of the Dicke type - is well known and is not repeated here. However, such radiometer devices are capable of resolving temperature levels with high accuracy at temperature levels representative of the earth; this is accomplished by comparing the received noise temperature power against a reference noise source, and using a super heterodyne driving an integrator to develop a temperature - representative voltage.

Figure 3-4 shows the block diagram of a microwave radiometer sensor used in Nimbus satellite applications, using a 19.35 GHz receiver, and using the cold of space as a reference temperature source.

Another type of non-camera imaging sensor is the side-looking radar, which is now used in terrestrial applications with pulse outputs representing image data bandwidths in excess of 100 MHz.

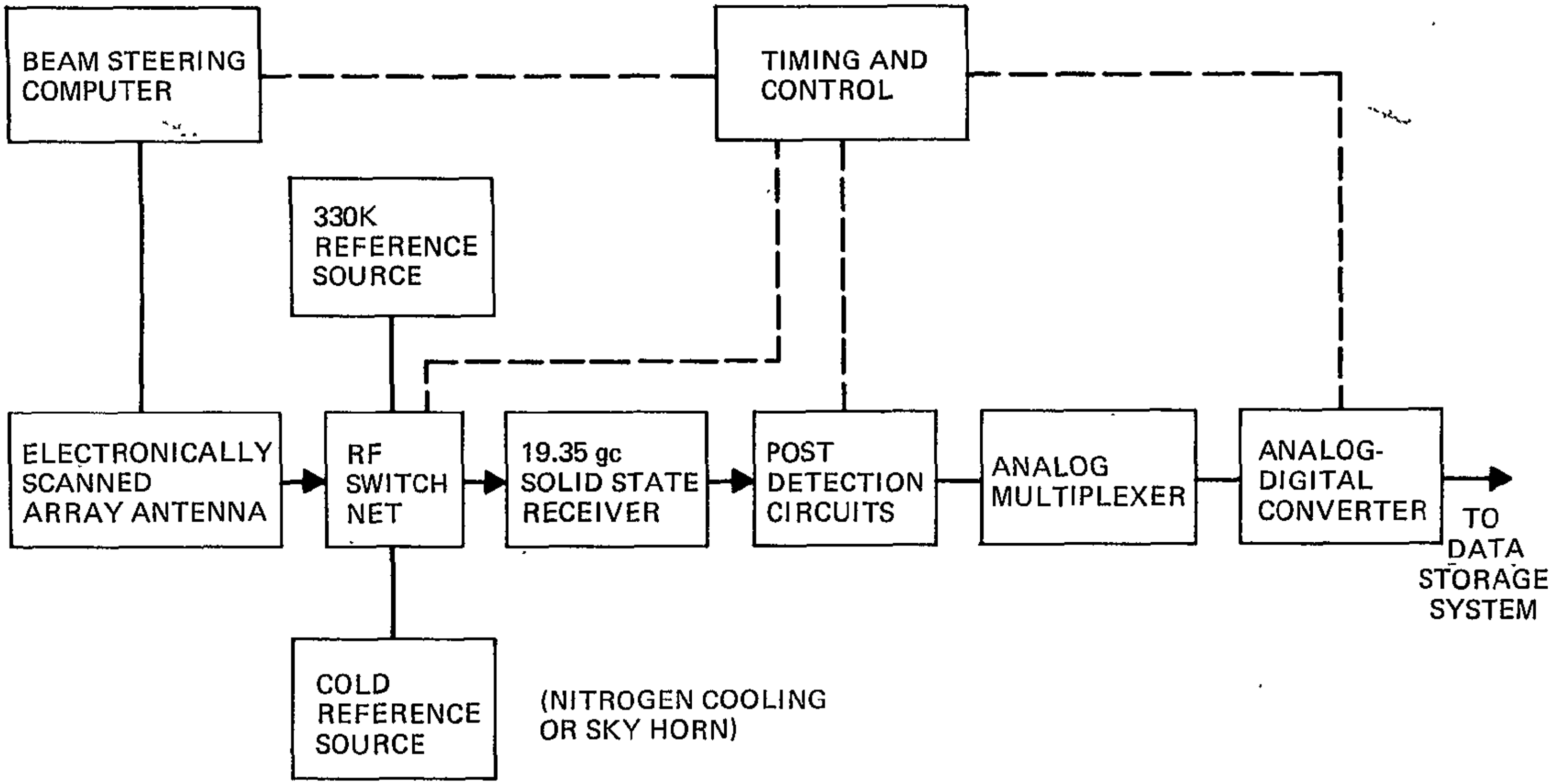
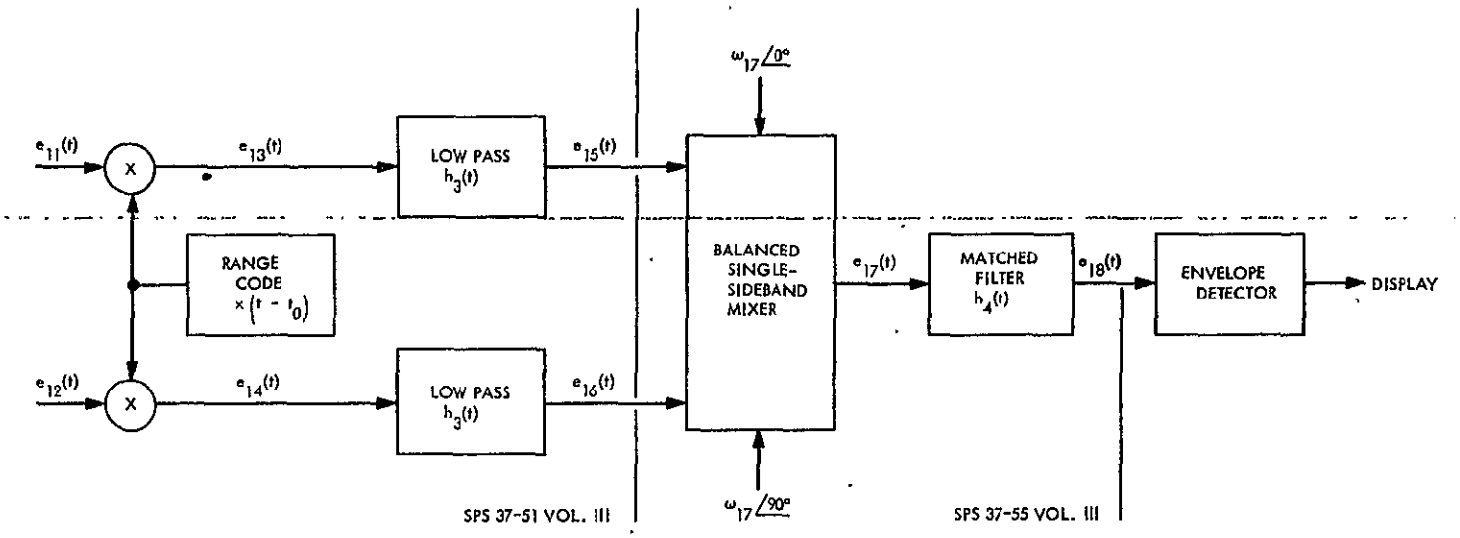

BEAM STEERING COMPUTER
TIMING AND CONTROL
330K REFERENCE SOURCE
ELECTRONICALLY SCANNED ARRAY ANTENNA
RF SWITCH NET
19.35 gc SOLID STATE RECEIVER
POST DETECTION CIRCUITS
ANALOG MULTIPLEXER
ANALOG-DIGITAL CONVERTER
TO DATA STORAGE SYSTEM
COLD REFERENCE SOURCE
(NITROGEN COOLING OR SKY HORN)

Figure 3-4 Block diagram shows microwave radiometric imager developed by Aerojet-General Corp. for NASA for Nimbus meteorological satellite application. A 19.35 gc solid state superheterodyne receiver and phased array antenna are used.


e11(t)
X
e13(t)
LOW PASS h3(t)
e15(t)
ω17∠0°
BALANCED SINGLE-SIDEBAND MIXER
e17(t)
MATCHED FILTER h4(t)
e18(t)
ENVELOPE DETECTOR
DISPLAY
RANGE CODE x(t - t0)
e12(t)
X
e14(t)
LOW PASS h3(t)
e16(t)
ω17∠90°
SPS 37-51 VOL. III
SPS 37-55 VOL. III

Data-processing block diagram for *Lunar Orbiter V* side-looking radar

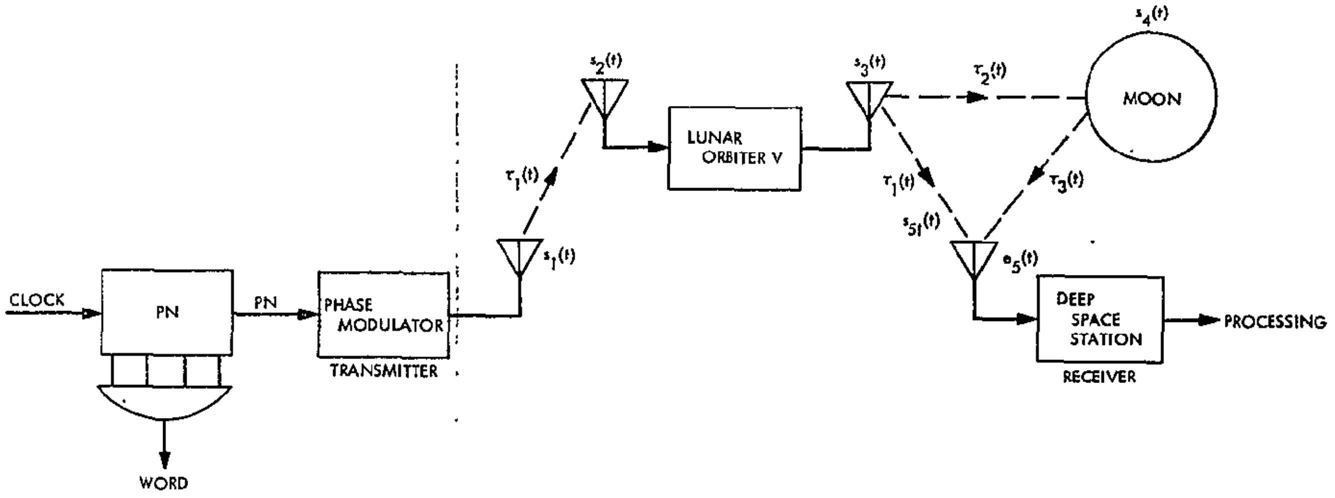

s4(t)
s2(t)
s3(t)
τ2(t)
MOON
LUNAR ORBITER V
τ1(t)
τ1(t)
τ3(t)
s51(t)
s1(t)
e5(t)
CLOCK
PN
PN
PHASE MODULATOR
TRANSMITTER
DEEP SPACE STATION
RECEIVER
PROCESSING
WORD

Figure 3-5 **Block diagram of *Lunar Orbiter V* side-looking radar experiment**

NASA has considered the use of side-looking radar for surface imaging or mapping radar systems applicable to lunar and planetary missions.

One approach has involved the use of an existing S-band ranging transponder with a high-gain antenna as a side-looking radar. An experiment was performed on January 24, 1968, using the S-band ranging transponder and the high-gain antenna of the Lunar Orbiter V spacecraft in flight as shown in Figure 3-5. A description of this experiment was contained in SPS 37-51, Vol. III, pp. 314-327.

The phase behavior of the reflected signal was related to the physical geometry of the spacecraft-surface positions. Also, the effects of earth-surface-spacecraft relative motion were included.

Side-looking radar equipments in orbiting spacecraft are capable of image detail and image resolution comparable to or in excess of the best camera type sensors. However, they require very high transmitter power, and also require processing and correlation with respect to time of the various pulse returns to develop a video signal indicative of the physical area of the earth being scanned. This can be accomplished by on-board spacecraft processing or can be accomplished by transmitting the unprocessed pulse stream as a data stream to a remote or central terminal.

3.4 VIDEO BANDWIDTH OF IMAGE SENSORS

The bandwidth of an image sensor - either a scan tube, a mosaic, or a rotating scanner can be derived from a consideration of the number of scanning lines, the size of the picture target, and the time element over which an image or frame is produced. This bandwidth must be consistent with the limiting resolution of the imaging sensor.

The video bandwidth is in general a function of the number of picture elements defined by the mosaic or by the scanning tube, i.e., number of horizontal scanning lines times the equivalent number of lines in vertical direction.

The image is scanned by the image sensor and is then converted into an analog video signal, the bandwidth of which is determined by the relation:

$$\text{Video bandwidth in Hertz} = \frac{L_H\, L_V}{2\ keT}$$

where

L_H = horizontal resolution in TV lines

L_V = vertical resolution in TV lines

k = vertical Kell factor

e = scan efficiency

T = time to scan one frame in seconds

If a vidicon for example has a resolution of 1200 lines horizontally and vertically, then for a Kell factor of 0.7, a scan efficiency of 0.95, and a frame time of 16 seconds, the video bandwidth is approximately 68 kHz.

Figure 3-6 relates the number of horizontal lines with the video bandwidth for the vertical Kell factor of 0.7 and the scan efficiency factor of 0.95 used above.

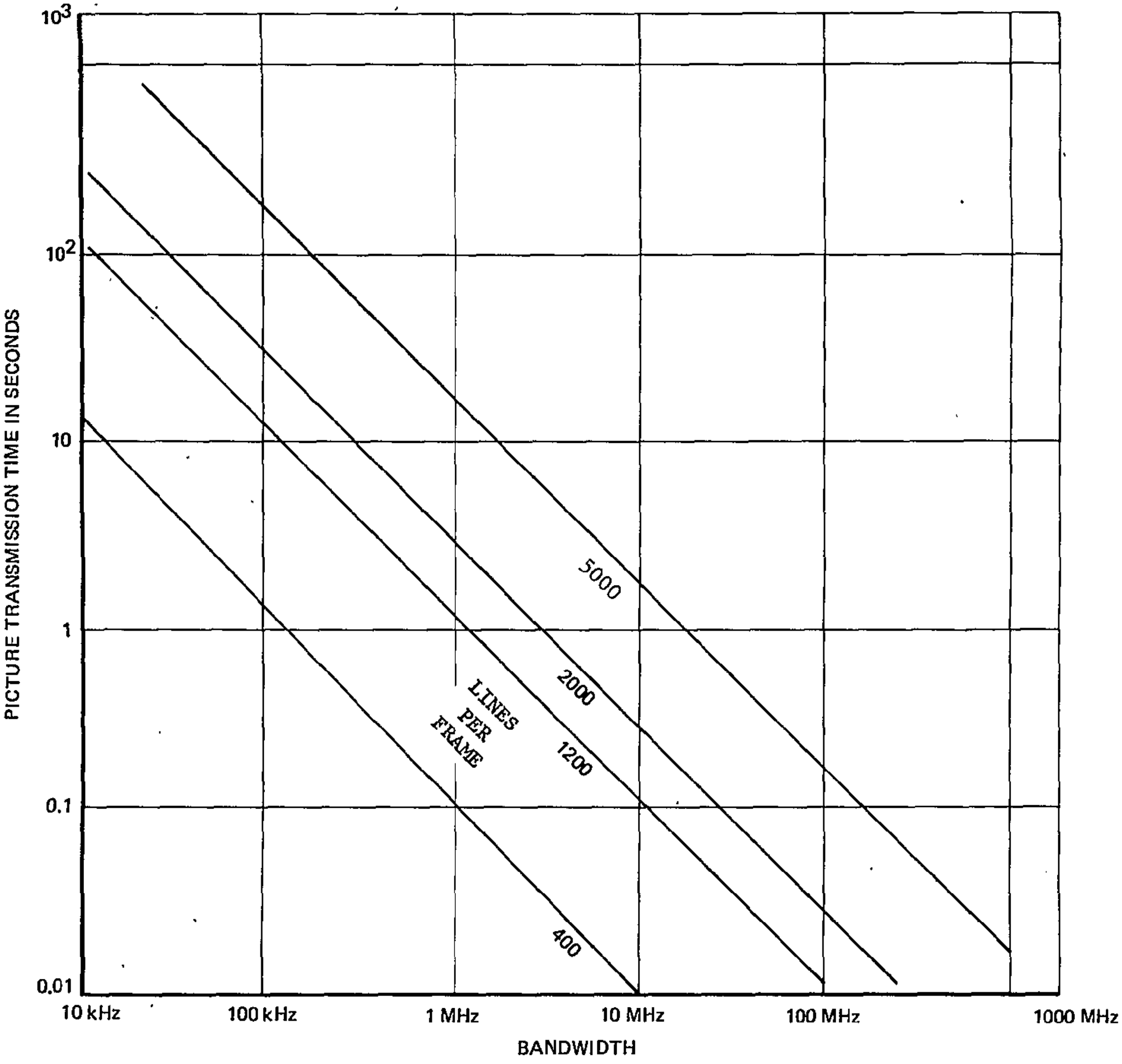


Figure 3-6 Basic Scanning Beam Image Device Parameters

3.4.1 Bandwidths of Parallel Operating Image Sensors

Consider the total video bandwidth provided by N tubes operating in parallel in a readout (R/O) mode.

Let K = number of picture elements per tube

N = number of tubes in parallel R/O mode

t_{RO} = readout time of each tube

The video bandwidth is then given by

$$BW = \frac{KN}{2t_{RO}} \quad \text{Hz}$$

Figure 3-19 represents the total video bandwidth as a function of the number of vidicons employed in any parallel R/O configuration. Curve A represents data for the 5140 vidicon tube (R/O time of .09 secs, 1.68×10^6 picture elements per tube).

This curve was derived using the following considerations which apply to the 5140 vidicon:

a. Recognize that the output of the 5140 vidicons during R/O is a set of parallel analog video signals which can be multiplexed into a single video signal.

b. The 5140 vidicons are read out in parallel during .09 seconds.

c. Each tube contains 1125 horizontal lines and 1500 vertical lines, or 1.6875×10^6 discrete picture elements.

d. At a readout time of .09 seconds, this converts to 18.75×10^6 picture elements per second.

e. Let four vidicons be read out in parallel. There is a total of 75×10^6 picture elements per second being presented in parallel to the multiplexer.

f. Apply the Nyquist sampling theorem, i.e., given a transmission channel with BW = Δ f Hz, the maximum number of picture elements that can be transmitted per second is 2 Δ f.

$$\text{Thus,} \quad 2\Delta f = 75 \times 10^6$$

$$\text{or,} \quad \Delta f = 37.5 \times 10^6 \text{ Hz,}$$

so that BW is 37.5 MHz as shown in Figure 3-7.

3.4.2 Wide Band Video Amplifiers

The overall bandwidth of the image sensor is determined by the bandwidth of the analog signal produced by the sensor across its load resistance (R_L) as shown in Figure 3-8a, and the video amplifier which amplifies this analog signal.

Since the driving impedance presented by the sensor and load resistor is in general high - in excess of 50,000 ohms, and into the megohms, the FET low noise transistor (Figure 3-8) is ideally suited to match to this high impedance, and a gain curve is provided with high frequency emphasis such as the characteristic response of the Schade amplifier shown in Figure 3-8 which was developed for a 60 megahertz video system.

3.4.3 Wide Band On-Board Image Processing Techniques

One of the most potent techniques for obtaining highest resolution and therefore widest bandwidth video information is to take actual photographs or film aboard the spacecraft and to then scan the developed film with flying spot scanner or laser scanner techniques. Figure 3-9 relates film rate, lines per millimeter, and film width to video bandwidth showing the 30 to 50 MHz video bandwidth possible at normal processing techniques and standard scanner resolutions.

G. Schade, "A Solid State Low Noise Preamplifier and Picture Tube Drive Amplifier for a 60 MHz Video System", RCA Review, March 1968.

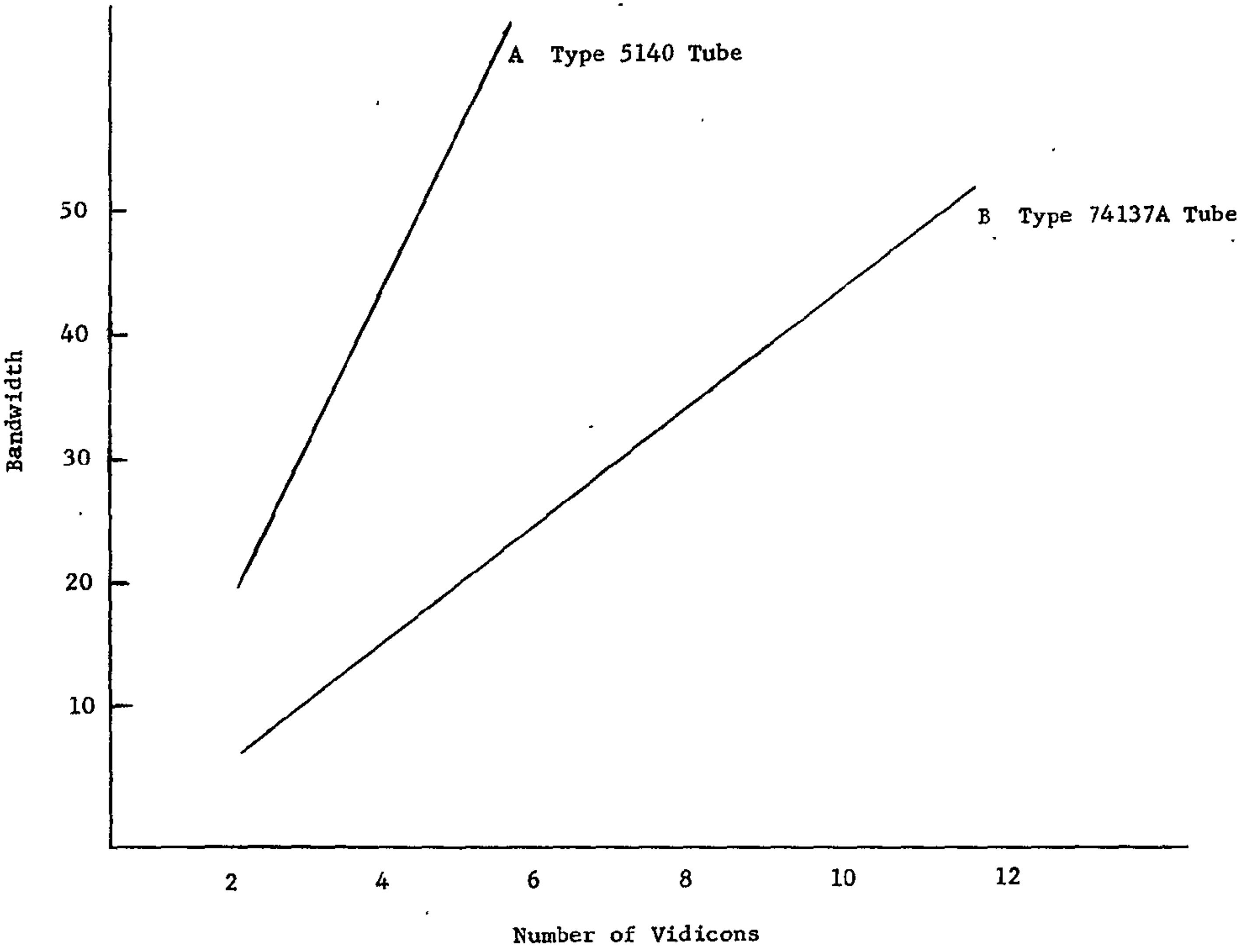
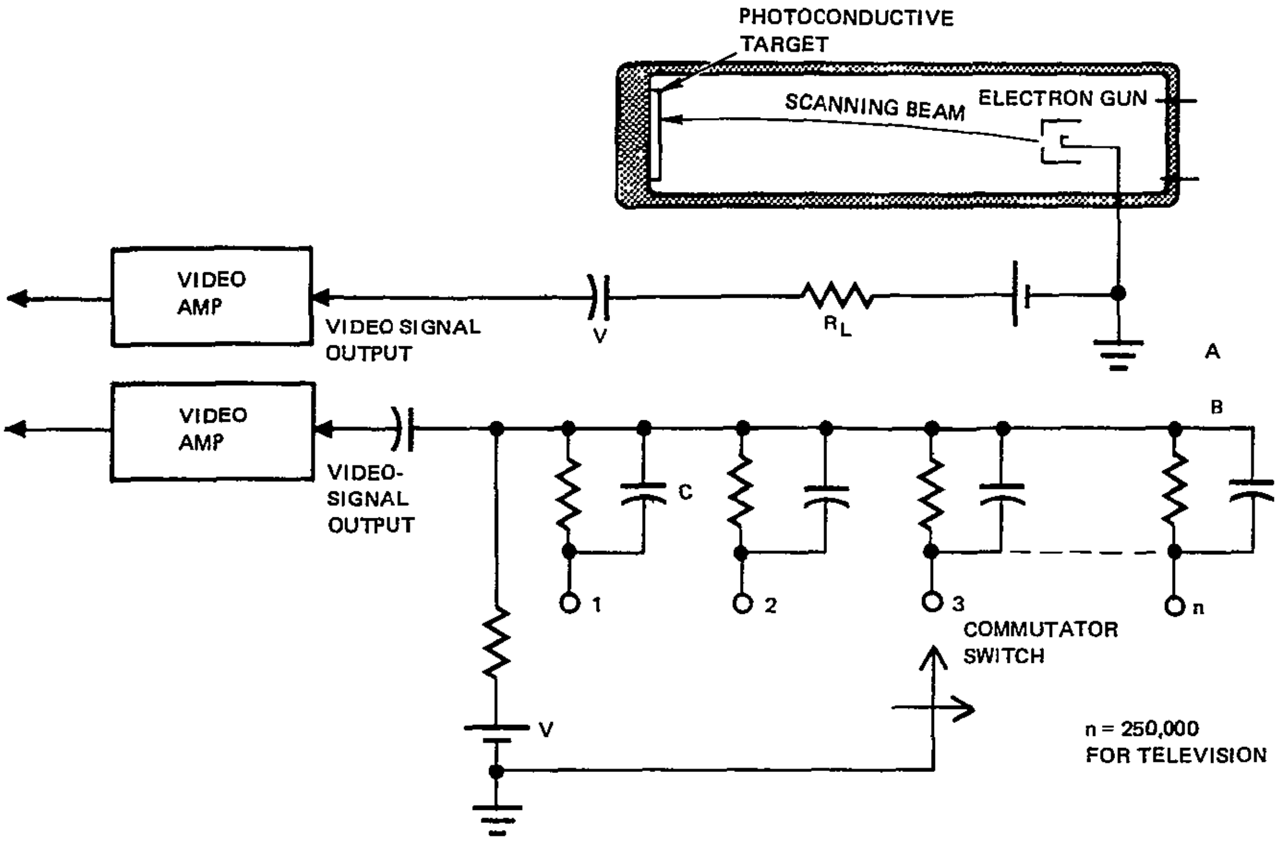
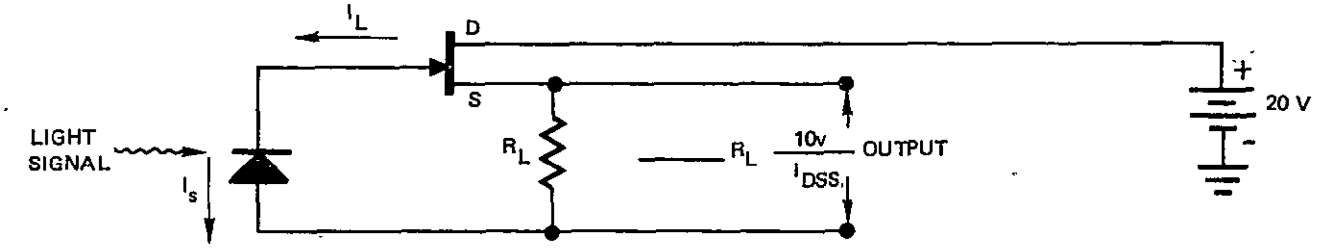


Figure 3-7 Transmission Bit Rate vs Vidicons Employed in Parallel R/O Configuration


PHOTOCONDUCTIVE TARGET
SCANNING BEAM
ELECTRON GUN
VIDEO AMP
VIDEO SIGNAL OUTPUT
V
R_L
A
VIDEO AMP
VIDEO-SIGNAL OUTPUT
B
C
1
2
3
n
COMMUTATOR SWITCH
V
n = 250,000 FOR TELEVISION

(a) COMPARISON OF (A) A TYPICAL VIDICON PHOTOCONDUCTIVE CAMERA TUBE WITH (B) AN ARRAY OF DISCRETE PHOTOSENSITIVE ELEMENTS.


I_L
D
S
LIGHT SIGNAL
I_s
R_L
R_L 10v / I_DSS
OUTPUT
+
20 V
-

(b)

FET Transistor Video Amplifier of Image Information

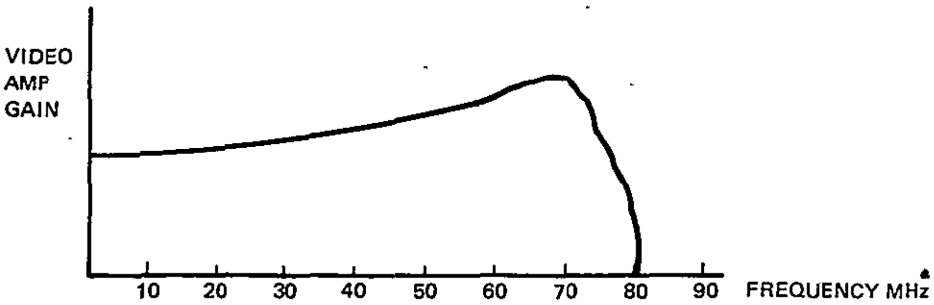


Figure 3-8 (c) Video Amplifier Response for 60 MHz Vidicon

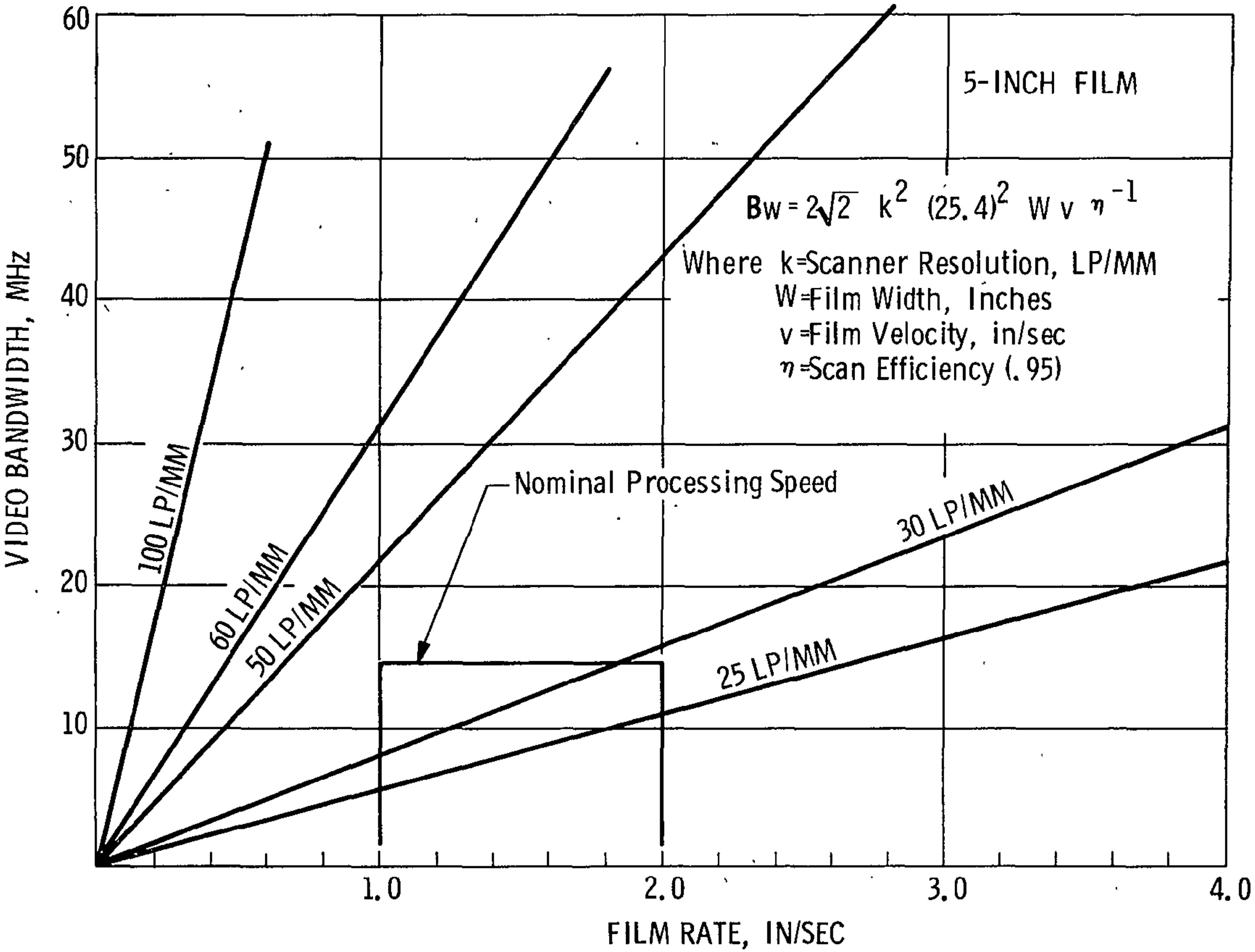


Figure 3-9 Film Rate Vs Video Bandwidth

Because of the optimum light levels used in the scanning process, signals provided by flying spot scanners or laser scanners are also capable of the highest signal to noise ratios.

3.5 SIGNAL TO NOISE CONSIDERATIONS

The sensor and video amplifier of Figure 3-8 provides an output signal voltage indicating the incident illumination level of the sensor, to which must be included the noise which occurs with the output signal at the video amplifier output.

3.5.1 Basic Noise Contributors

The basic noise contributors of the sensor system are the noise of the sensor - due to the shot-noise of the electron beam in the case of a camera tube type of sensor, the target noise, the thermal Johnson noise (20 eKTR) produced by the load resistor, and the noise contributions of the video amplifier.

Of particular interest in this noise budget is the so-called dark current of the sensor, usually a small fraction of a microampere which represents the principal noise current output of the sensor in the absence of any illumination.

3.5.1.1 Signal to Noise Ratio and Illumination Level

It is a particular and almost unique feature of an image sensor - whether photo conductive or photo emissive that the signal to noise ratio depends primarily upon the incident illumination level, increasing from zero to very large values as illumination of the target increases from near zero through the useful illumination range of the sensor.

3.5.2 Noise and Sensitivity

In any electronic sensor device, sensitivity is determined by minimum allowable signal-to-noise ratio. Signal level is usually determined by the image illuminance on the photo-conductor target and the exposure time, since the latter determines the total number of light quanta available to raise electrons or holes into the conduction band. Noise level is then determined by the RMS combination of thermal noise, generation-recombination noise and current noise.

Vidicons, image orthicons, and image intensifier orthicons are capable of televising scenes with scene illumination ranging from direct sunlight ($\sim 10^4$ foot-candles) to completely overcast starlight ($\sim 10^{-5}$ foot-candles). The vidicon covers those scenes where photocathode illumination can be arranged to be 10^{-2} foot candles or more.

The image orthicon covers photocathode illumination levels to below 10^{-6} foot-candles, while the image intensifier orthicon extends the range to about 10^{-8} foot-candles. The range of operation of these devices overlap one another to some extent.

3.5.2.1 Noise Levels in Vidicons

At scenes of either low or high illumination, the lowest illumination detectable by vidicon imaging systems is set by noise in the preamplifier. Noise from the statistical fluctuations in the electron readout beam is below the level of the lowest-noise preamps. Thus, vidicon noise is essentially dependent only on electrical bandwidth, not on magnitude of signal current.

3.5.2.2 Noise Levels in Image Orthicons

Image orthicons use low noise electron multipliers and the beam current (to within 10 percent) is practically independent of signal level. Internal noise in image orthicons is therefore essentially independent of signal level and will be largely determined by the noise contributions of the video amplifier.

3.5.3 Basic Signal to Noise Relationships

The basic relationships between brightness, contrast and resolution in an ideal video sensor are given by Weimer[1] in the following equation:

$$\frac{BC^2}{1^2} = 5.3 \frac{k^2}{D^2 t\theta}$$

---

[1] P. K. Weimer, " Television Camera Tubes; A Research Review", Advances in Electronics and Electron Physics", Acadamic Press, New York, 1960.

B = scene brightness (foot-lamberts)

C = percent contrast ($C = \frac{\Delta B}{B} \times 100$)

k = threshold signal to noise ratio

D = diameter of lens aperture (in.)

l = limiting resolution in TV lines = N/L 50.6

t = storage time (1/30 sec for commercial TV)

N = number of TV lines

θ = quantum yield of the primary photo process

L = length of TV line in inches

This relationship was obtained by considering the accuracy with which one can count the total number of photons hitting an elemental area of the photocathode in time t.

This equation shows that objects with small detail and low contrast will require more light to be sensed. The maximum resolution of a video sensor increases as the square root of the scene illumination and directly as the contrast. It is possible to increase the sensor resolution capability by increasing the storage time or by using materials that have higher quantum yields. The ideal sensor is approached by present high sensitivity video sensors at low light levels.

The signal to noise ratio of an ideal tube can be expressed as:

$$\frac{BC^2}{l} = k_1 k^2$$

where k_1 is equal to 5.3 K^2/D^2 $t\theta$ for uniform illumination.

According to this equation, an increase in brightness or a decrease in resolution will increase the signal to noise ratio of the sensor.

3.5.4 Practical Signal to Noise Considerations

In image sensor systems it is not possible to provide meaningful signal-to-noise analyses and guide lines to meet all illumination conditions, types of sensors, and types of video amplifiers.

In general, the level of dark current of the sensor must be determined and the working illumination level range compatible with the sensor characteristics will determine the signal level. The number of lines or picture elements and the frame repetition rate will then determine the video bandwidth, which in turn will require the selection of a type of video amplifier whose noise characteristics will largely determine or limit the noise output of the system.

In image intensifier orthicon systems working at very low incident illumination levels - of the order of 1×10^{-5} foot-candles, a typical signal-to-noise ratio in a 4.5 MHz bandwidth can be as low as 3:1.

Vidicons and image orthicons operating at higher illumination levels can realize signal-to-noise ratios from 10:1 to ratios of many hundreds.

O. Schade's 60 MHz video system, using a wideband high transconductance dual gate MOS video amplifier and operating with visible resolving power up to 3000 lines reported SNR of 40 dB with a signal level output of 1-volt and a noise of 10 millivolts rms.

3.6 IMAGE SENSOR SYSTEMS

Modern image sensor systems which are capable of providing images of fine detail of earth features from a vantage point of many thousands of miles in space, are capable of providing analog output signals with bandwidths into the hundreds of megacycles. These sensor systems will in general include a number of individual sensors - thermal, visible spectrum, and so on.

Table 3-2 lists the principle characteristics of the various image sensors which have been listed in Table 3-1 and which have been described in Section 3 of Volume 2. As indicated, signal to noise ratios in excess of 40 dB are available with most sensors operated at light levels well above dark current noise with image storage devices such as the FPS vidicon or the return-beam vidicon capable of limiting resolutions up to 4000 TV lines.

3.6.1 Sensor System Combinations

Combinations of image sensors, to cover various ranges of the visible and infrared spectrum can be operated in a single spacecraft, using multiplexing systems such as are described in the next section, to produce a single digital signal. The total analog bandwidth of the multiple sensor system must be greater than the sum of the individual sensor analog bandwidths. The digitized data signal will have a total bandwidth which corresponds to this analog bandwidth.

3.6.2 Image Data Bandwidths for Highest Resolutions

Modern high resolution imaging tubes and devices are a far cry from the image tubes developed to serve the commercial television industry. The return tube vidicon, developed by O. Schade at RCA and the FPS vidicon developed by Schlesinger at G.E. have made available to new spacecraft imaging camera tubes with analog video bandwidths up to 80 MHz bandwidths; one such tube operated at a high frame rate - would fill the available digital data bandwidths (400 MHz) to be discussed in the next section. Groups of lower resolution tubes, on the other hand, with individual lower bandwidths in the 15-20 MHz range, can be digitized and multiplexed to provide a combined comparable digital bandwidth, and most new high data imaging systems will use sensor devices in groups to provide such features as multispectral operation.

TABLE 3-2

IMAGING SENSORS

0.4 to 0.8 micron

	SENSOR TYPES	GENERAL USE	LIMITATIONS	ADVANTAGES	TYPICAL RESOLUTION
	Image Storage				
1.	Vidicon (Including FP.S type)	Moderate Light Level	Lag	Simple or Stable	(TV Lines at MTF = 10%) 1" = 1000 (FPS) 3" = 3000 (FPS)
2.	Plumbicon (Pbo target)	Moderate Light Level		Simple, Stable and Negligible Lag	1 1/2" = 800
3.	Return beam vidicon	Low-Light Level with high resolution	Greater lag (lower I beam than #1 or 2)	Higher Resolution	4 1/2" = 4000
4.	Image Orthicon	Intermediate to low light level	More complex than 1 or 2.	Higher Sensitivity	4 1/2" = 800 3" = 600
5.	Image Isocon	Low Noise, Low Light Level	More complex than 1 or 2.	Low Light level negligible lag and high dynamic range (2000:1)	4 1/2" = 800
6.	Any one of above plus one or more stages of image intensification.	Higher sensitivity (typically 100 to 300 gain per stage)	High voltages and lower resolution (due to presently available image intensifier)	Higher Sensitivity	400 to 500
	Storage Type Infrared Image Sensors				
	Iricon	I-R Sensing		Simple & Stable	3" = 800
	Non-Storage Infrared Sensors				
	Point Source Diodes or Photo Transistors (PbSe or Pbs)	Rotating Array Assemblies	Non-storage		Determined by Number of Sensors

TABLE 3-2

IMAGING SENSORS (continued)

0.4 to 0.8 micron SENSOR TYPES Image Storage	Sensitivity Target LIGHT LEVEL Ft.Candles	TYPICAL DB SIGNAL TO NOISE RATIO	S/N INFLUENCED PRIMARILY BY:
1. Vidicon (including FPS type)	0.1	50 (9 MHz) 45 (45 MHz)	Video Preamplifier Video Preamplifier
2. Plumbicon (Pbo target)	0.1	55 (9 MHz)	Video Preamplifier
3. Return Beam Vidicon	10^{-3}	45 (50 MHz)	Beam Current and Tube Design
4. Image Orthicon	10^{-4} 10^{-4}	40 (10 MHz) 35 (8 MHz)	Beam Current and Tube Design Beam Current and Tube Design
5. Image Isocon	10^{-6}	40 (10 MHz)	Beam Current and Tube Design
6. Any One of Above Plus One or More Stages of Image Intensification	Determined by Number of Stages		----
Storage Type Infrared Image Sensors			
Iricon	Equiv. 0.1	35 (9 MHz)	Video Preamp, and Target Cooling
Non-Storage Infrared Sensors			
Point Source Diodes or Photo Transistors (PbSe or Pbs)	---	---	Video Preamplifier

3.6.3 Spin-Scan Image Cameras*

Spin-scan imaging is used in such systems at ATS, ERTS, and will be used in SMS. Such cameras or image systems use the spinning of the satellite to scan the surface of the earth and direct the equivalent of a scanning line to each of a group of point sensors.

One type of spin-scan camera uses a geometry similar to Figure 3-10, which provides cross track scanning by use of a high duty cycle "rocking mirror" located in front of the telescope collector. The image produced at the primary image plane of the telescope is relayed by use of fiber optic bundles to detectors where conversion to an electrical signal is accomplished. Optical filters are used to select the optical pass-band corresponding to each spectral band. Six detectors are employed in each spectral band to permit a slower scanning motion of the "rocking mirror" system. A small mirror, located on a 45° angle with respect to the optical axis, is used for calibration of the scanner. This calibration is accomplished as the spacecraft comes into the sun view from the dark side of the earth. This mirror drops in place in front of the viewing part of the scanner and requires a clear vertical view in the easterly direction between the earth "limb" and the horizontal. The horizontal view must be adequate to catch the sun for tolerances resulting from seasonal variations and the launch window.

In the ATS multi-color spin-scan cloud camera system, the spacecraft rotation is 100 rpm, the lines per frame is 2407 lines with a frame time of 24 minutes (2.4 minutes retrace time).

A dwell time of 9.5 microseconds is provided for an instantaneous field to scan a point source. Photo-multipliers using S-11 and S-12 photocathodes are used.

---

* S. D. Dorfman, J. R. Castaldo, and W. L. Exner, "Some Experimental Results in Earth Resources Data Transmission", AIAA Paper 70-328, AIAA 3rd Communications Satellite Systems Conference, Los Angeles, California, April 6-8, 1970.

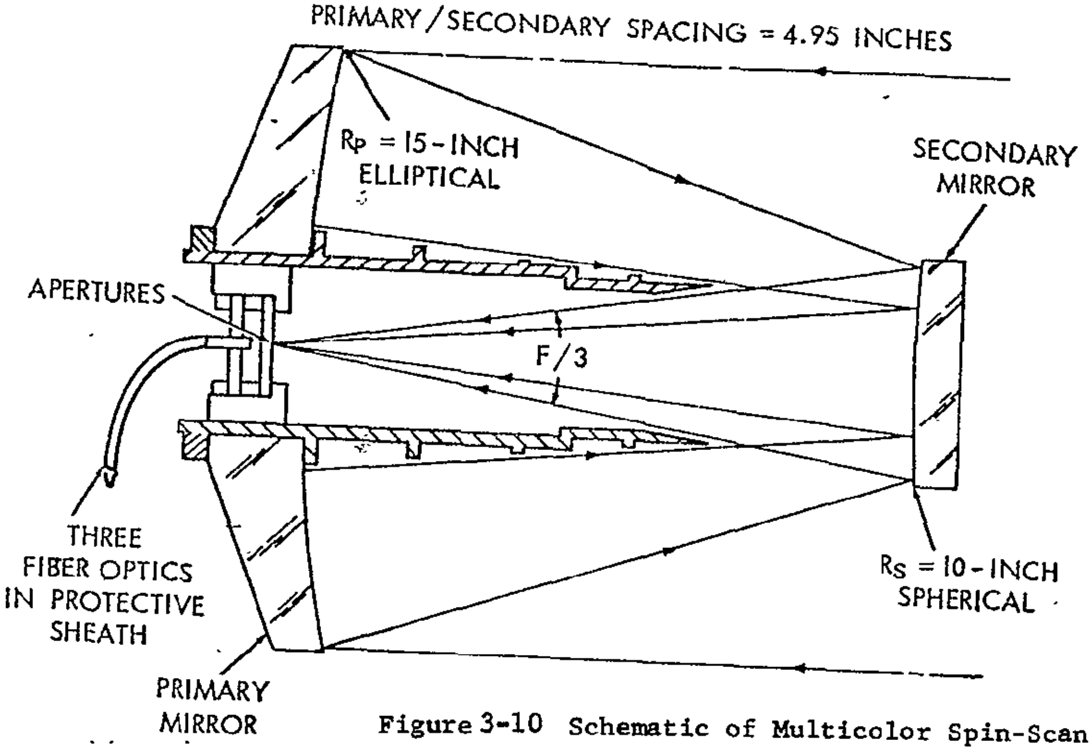

PRIMARY/SECONDARY SPACING = 4.95 INCHES
RP = 15-INCH ELLIPTICAL
SECONDARY MIRROR
APERTURES
F/3
THREE FIBER OPTICS IN PROTECTIVE SHEATH
RS = 10-INCH SPHERICAL
PRIMARY MIRROR

Figure 3-10 Schematic of Multicolor Spin-Scan

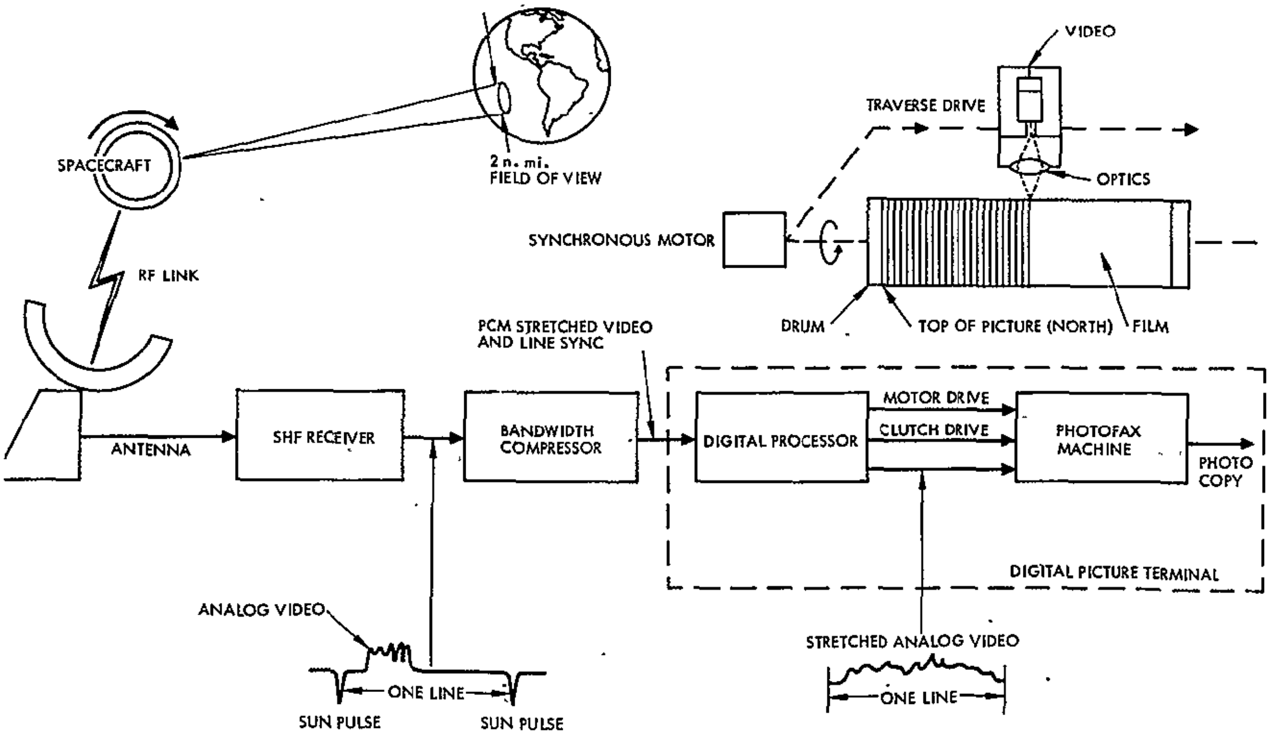

SPACECRAFT
2 n. mi. FIELD OF VIEW
RF LINK
VIDEO
TRAVERSE DRIVE
OPTICS
SYNCHRONOUS MOTOR
DRUM
TOP OF PICTURE (NORTH)
FILM
PCM STRETCHED VIDEO AND LINE SYNC
ANTENNA
SHF RECEIVER
BANDWIDTH COMPRESSOR
DIGITAL PROCESSOR
MOTOR DRIVE
CLUTCH DRIVE
PHOTOFAX MACHINE
PHOTO COPY
DIGITAL PICTURE TERMINAL
ANALOG VIDEO
ONE LINE
SUN PULSE
SUN PULSE
STRETCHED ANALOG VIDEO
ONE LINE

Figure 3-11 Spin Scan Block Diagram (Hughes-ESSA System)

To record the earth-image as transmitted from the Spin Scan Cloud Camera, a precise line-start timing reference is required. Since no such reference is needed by the camera itself, none is generated in the spacecraft; however, if the ground equipment is to derive a suitable timing signal it must have available some sort of spin angle reference. This is provided by transmitting, along with the camera video, the output of one of the spacecraft sun sensors, which produces one pulse each spin revolution as its plane of view crosses the sun. (Figure 3-11)

Table 3-3 lists many of the operating characteristics of the spin-scan camera system. Some advantages of the spin camera system, as set forth by S. D. Dorfman et. al. are listed as follows:

Compatible with spin stabilization
Wide coverage and high resolution
Wide dynamic range
Simple and reliable space hardware
Excellent scan precision and stability
Simple, on-axis optics
Adaptable to IR

3.6.4 Multi-Spectral Scanner (MSS) for Earth Resources Technology Satellite

The MSS for ERTS differs from the spin scan camera for ATS in that satellite motion is utilized to generate one dimension of the picture and an oscillating mirror provides the other dimension of picture scan. The satellite is stabilized and earth oriented in a 500 n.mi. sun-synchronous orbit. Pictures are generated in a 100 n.mi. swath.

A schematic of the MSS camera is shown in Figure 3-12. The 45-degree scan mirror oscillates at 15.2 Hz and scans 24 or 26 detector elements in four or five spectral bands to generate an MSS image. The first ERTS mission will utilize four bands, and the second ERTS will add the IR channels. A summary of the characteristics of the MSS is given in Table 3-4.

TABLE 3-3

PRESENT SPIN SCAN SYSTEM PERFORMANCE

Resolution	0.1 mrad (2 miles)
Dynamic range	> 400:1
Scan lines	2400
Earth coverage	Full earth
Scan rate	100 lines/minute
Picture time	24 minutes
Retrace time	2 minutes
I.F.O.V. dwell time	9.6 microseconds
Video bandwidth	100 kHz
Spectral response	0.48 to 0.63 micron

TABLE 3-4

MULTI-SPECTRAL SCANNER SYSTEM SUMMARY

	Scanner
Optics	9-inch Cassegrain type with 4-inch secondary mirror, f/3.3
Scanning method	Flat mirror oscillating ± 2.9 degrees at 15.2 Hz
Instantaneous field of view	0.077 mr
Spectral bands and detectors	1) 0.5 to 0.6 micrometer, photomultiplier tubes 2) 0.6 to 0.7 micrometer, photomultiplier tubes 3) 0.7 to 0.8 micrometer, photomultiplier tubes 4) 0.8 to 1.1 micrometer, silicon photodiodes 5) 10.4 to 12.6 micrometers, intrinsic IR detector
Number of lines scanned/band	Six for bands 1, 2, 3, and 4; two for band 5
Limiting resolution from 494 n.mi.	116 feet for bands 1, 2, 3, and 4; 348 feet for band 5
Temperature for IR detector	$\sim 77^{\circ}$ to 100°K
	Multiplexer
Number of channels	26
Bandwidth	15 megabits/sec
Accuracy	6 bits
Initiation of sampling	Slit pulse
Number of samples/scan line	4096

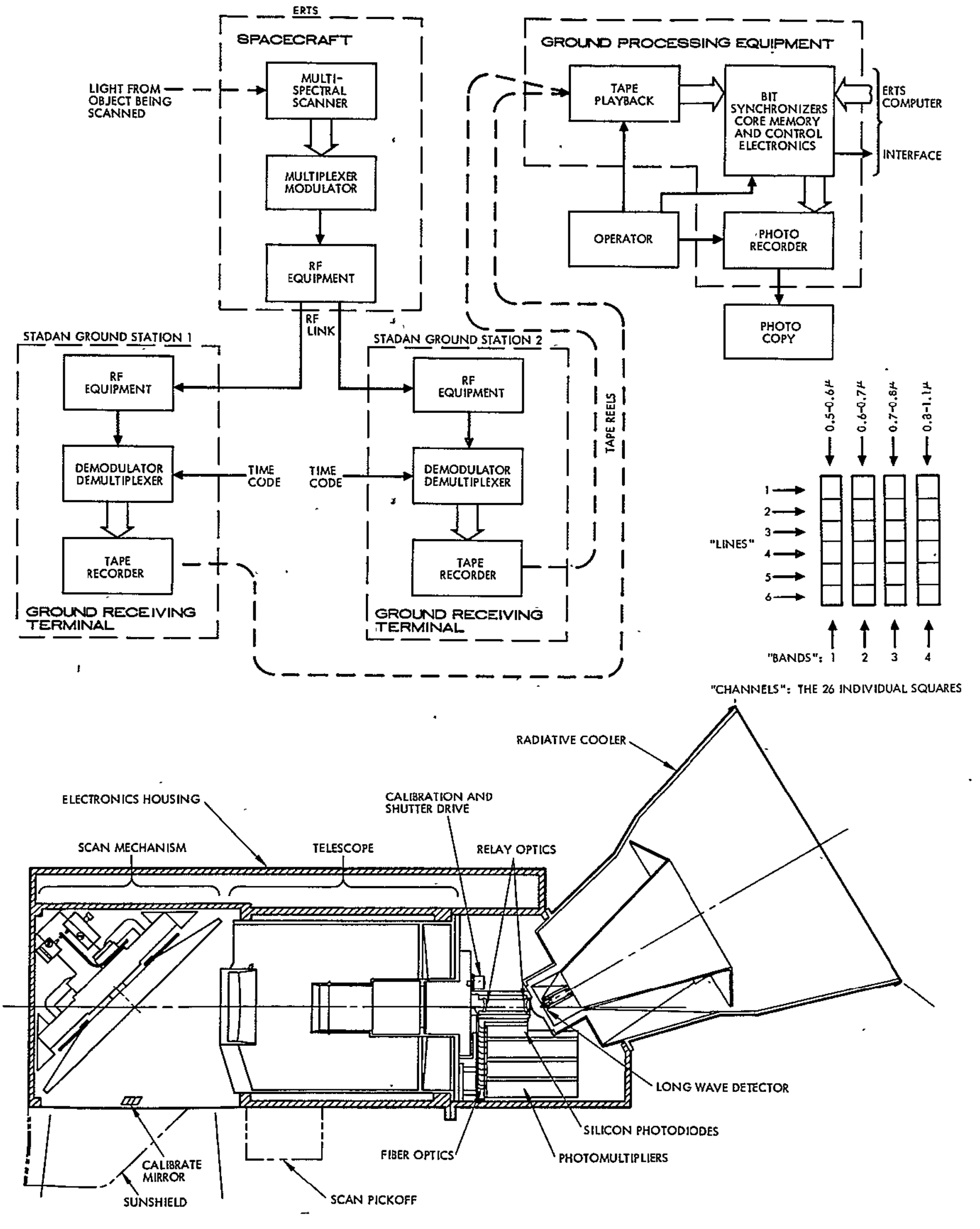

ERTS
SPACECRAFT
LIGHT FROM OBJECT BEING SCANNED
MULTI-SPECTRAL SCANNER
MULTIPLEXER MODULATOR
RF EQUIPMENT
RF LINK
STADAN GROUND STATION 1
RF EQUIPMENT
DEMODULATOR DEMULTIPLEXER
TIME CODE
TAPE RECORDER
GROUND RECEIVING TERMINAL
STADAN GROUND STATION 2
RF EQUIPMENT
TIME CODE
DEMODULATOR DEMULTIPLEXER
TAPE RECORDER
GROUND RECEIVING TERMINAL
TAPE REELS
GROUND PROCESSING EQUIPMENT
TAPE PLAYBACK
BIT SYNCHRONIZERS CORE MEMORY AND CONTROL ELECTRONICS
ERTS COMPUTER
INTERFACE
OPERATOR
PHOTO RECORDER
PHOTO COPY
0.5-0.6μ
0.6-0.7μ
0.7-0.8μ
0.8-1.1μ
1
2
3
4
5
6
"LINES"
"BANDS": 1 2 3 4
"CHANNELS": THE 26 INDIVIDUAL SQUARES
RADIATIVE COOLER
ELECTRONICS HOUSING
CALIBRATION AND SHUTTER DRIVE
SCAN MECHANISM
TELESCOPE
RELAY OPTICS
LONG WAVE DETECTOR
SILICON PHOTODIODES
PHOTOMULTIPLIERS
FIBER OPTICS
CALIBRATE MIRROR
SUNSHIELD
SCAN PICKOFF

Figure 3-12 Multi-Spectral Scanner System

Different vegetation transmits, absorbs, reflects and re-emits radiant energy in various portions of the electromagnetic spectrum.

Reflectance is high and thermal emission is low in the photographic infrared band (0.7 - 1.0 microns). In the thermal infrared band (8 - 14 microns), the opposite relationship prevails. A mix of reflectance and emission are important for identifying vegetation or objects between these bands.

The same plant also reflects and re-emits radiant energy as a function of its health. Hence, agriculturists are able to develop signatures of distressed plants, and when they are discovered ground observers can be sent to investigate.

These signatures will be used to detect crop problems before they are readily visible on the ground. The condition and acreage of an agricultural crop can be quickly estimated and used to make national and world estimates of food supplies early in the crop cycle.

Wavelengths to cover appropriate regions for agricultural spectral signatures have been determined by joint research of the Agriculture Dept. and NASA and are tabulated in Table 3-5 . They are 0.5 - 0.6 microns, 0.6 - 0.7 microns, 0.7 - 0.8 microns and 0.8 - 1.1 microns. The first three bands are in the visible spectrum and the last is in the near infrared region.

Present plans call for these four bands on ERTS-A and an additional band, 10.4 - 12.6 microns to be put on ERTS-B.

The spacecraft, covering the earth in 100-mi. swaths from 496-mi. polar orbit, will photograph the earth in 18 days, thereby affording 20 opportunities for complete coverage in its lifetime of a year. Ground features will be resolved to dimensions of about 200 ft. with the multi-spectral scanner, compared with resolution of approximately 340-475 ft. for the return beam vidicon.

TABLE 3-5
AGRICULTURAL SPECTRAL SIGNATURES

	Band 1 0.5=0.6 Microns	Band 2 0.6=0.7 Microns	Band 3 0.7=0.8 Microns	Band 4 0.8=1.1 Microns	Band 5 10.4=12.6 Microns
Wheat - low fertilizer......	3.44	2.27	3.56	8.95	
Wheat - high fertilizer.....	3.69	2.58	3.67	9.29	
Water........................	3.75	2.24	1.20	1.89	
Barley, healthy.............	3.96	4.07	4.47	9.29	
Barley, Mildewed............	4.42	4.07	5.16	11.60	
Oats........................	4.02	2.75	3.50	9.64	
Oats........................	3.21	2.20	3.27	9.46	
Pine........................	3.96	3.35	5.62	16.17	312/307
Poplar......................	3.90	2.96	4.13	13.08	
Soybean - high H_2O..........	3.29	2.78	4.11	8.67	
Soybean - low H_2O...........	3.35	2.60	3.92	11.01	
Sand........................	5.19	4.32	3.46	6.71	
Loam - 20% H_2O..............	4.21	4.02	3.38	7.57	
Loam - 1% H_2O...............	6.70	6.79	6.10	14.01	
Cultivated land.............	3.27	2.39	1.58		
Clay........................	14.34	14.40	11.99		
Gneiss......................	7.02	6.54	5.37	10.70	
Ice.........................	18.30	16.10	12.20	11.00	225/283
Snow........................	19.10	15.00	10.90	9.20	265
Loose soil..................	7.40	6.91	5.68		301
Fallow......................					302
Soil........................					315
Cotton......................					301
Sorghum.....................					307
Corn........................					307.5
River.......................					293
Pine - diseased.............					312
Pine - healthy..............					307
Sycamore....................	4.19	5.22	8.14	20.13	
Sycamore....................	4.07	4.59	7.92	18.24	
Oceans......................					300
Oceans......................					264
Grass meadows...............					289.5

The output of the scanner is made up of the individual outputs of 24 photo sensors, six per spectral band. These are digitized on board the spacecraft, multiplexed into a single 15 megabits/sec bit stream, and transmitted to the ground.

At the receiving stations the bit stream is demultiplexed and recorded on tape. At the central processing station the tapes are played back and the information for each scan (24 channels) is stored in core memories. The information is then read out of the memories, converted to analog signals and applied to the photo recorder properly sequenced to produce the four spectral images.

3.6.5 An Illustration of a Multiple Sensor System - the ERTS-A Spacecraft

The ERTS-A spacecraft which is designed for agricultural applications will carry two sensor packages, a four-channel multispectral sensor (MSS) and three return beam vidicon (RBV) cameras.

The multispectral sensor, accommodating four spectral bands from ultraviolet to far infrared, provides the color registration and resolution needed.

3.6.6 The Interface Between the Sensor System and the A/D Converter

The camera and mosaic image sensors described in this section have the characteristics of wide dynamic range, wide bandwidths, and MTF characteristic curves whose amplitudes decrease with frequency.

When the sensor output is handled by conventional analog or digital PCM techniques, the problem of dynamic range will result in requirements of many levels of quantization on the part of the PCM format, or will require the handling of an analog signal having wide range of voltage excursion in addition to very high frequency components. Otherwise the resulting display of information will switch only between black and white and few intermediate shades of grey giving an image of degraded detail.

This problem has been approached in low data rate systems by using 6-bit PCM representing 64 quantization levels with excellent reproduction capabilities. In the case of the Mars Mariner image data transmissions, computer programmed image "sharpening" at the data terminus has added to picture contrast and resolution.

When real time high data rate transmission of high resolution images with large dynamic range (baseband bandwidths up to 50 megahertz) is required, analog techniques become impractical and PCM signals with accuracy in excess of 3-bits require enormous bandwidths.

Here techniques relating to bandwidth compression and to the rejection of redundant data become very useful and practical, provided that they are consistent with primary edge-resolution or contouring requirements. For most applications, delta modulation can represent both the large bandwidth and dynamic ranges and will be discussed in the Sections 5 of both volumes.

SECTION 4

DIGITAL TRANSMISSION OF IMAGERY DATA

4.1 INTRODUCTION

In this section, the use of delta modulation for the conversion of the analog video data information to a digital bit stream, and its reconversion back to analog data will be discussed. The technology of developing digital bit streams with bit rates from 40 to 400 MBPS is very different from the technology of developing low data rate streams, and this unusual characteristic of the video signal - having a high redundancy data content coupled with high speed image-edge effects-places severe constraints on applicable techniques.

High data rate A/D conversion approaches are identified and discussed in Section 4 of Volume 1; however, emphasis will be placed here on delta modulation and its application to imagery data transmission. The final report will specifically deal with the technology of 200 to 400 MBPS bit streams.

4.2 DIGITAL CODE EQUIVALENT OF ANALOG SIGNALS

In digital communications, the analog input signal must be converted to the pulse code which is the digital equivalent of the analog signal.

Figure 4-1 shows the general circuit used in digital communications. The output of the A/D converter, which develops the bit stream from the analog signal, is translated by a communication link to a digital-to-analog converter (decoder) where the output is the original analog signal.

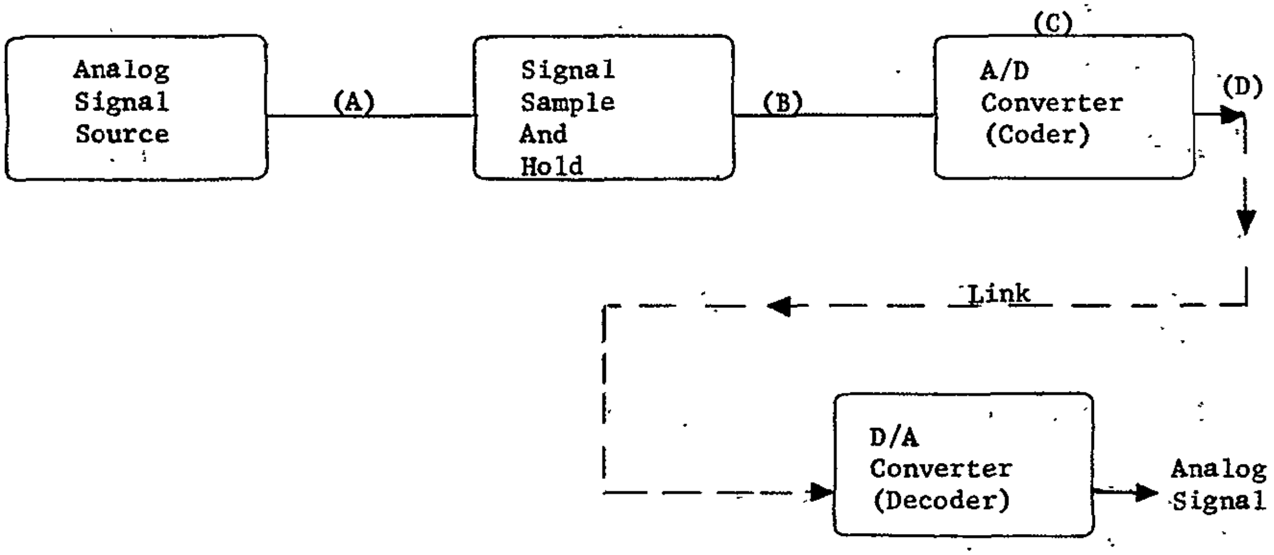

Analog
Signal
Source
(A)
Signal
Sample
And
Hold
(B)
(C)
A/D
Converter
(Coder)
(D)
Link
D/A
Converter
(Decoder)
Analog
Signal

Figure 4.1 Circuit System Performing the Coding and Decoding Functions in High Speed Digital Communications

Figure 4-2 shows the principal steps in the development of the digital bit stream, following the circuit points identified in Figure 4-1. The analog signal is first sampled to establish the amplitudes at key points in the analog waveform.

The amplitudes of the samples are then converted to the particular code used by an analog-to-digital converter. In the converter, the sample amplitudes are measured against a quantized scale. If a binary code is used, two bits correspond to 4 amplitude levels, 3 bits correspond to 8 amplitude levels, 4 bits correspond to 16 amplitude levels, 5 bits to 32, 6 bits to 64, 7 bits to 128 and so on. Thus, in each sampling period, a series of bits are developed which represent the sequence or code representing that amplitude. In the case of delta modulation, which will be discussed in depth in this section, the code is developed from difference information rather than the gross scale measurement.

The codes shown in Figure 4-2 are for a three-bit PCM code with a fourth (first) bit used to indicate whether a sample is positive or negative. The zeros or ones of the coded bit stream must then be converted to the frequency in phase levels of the modulated carrier as will be discussed in the next section.

4.3 DATA RATE REQUIRED FOR A GIVEN BASEBAND BANDWIDTH

When voice information is to be transmitted by a digital bit stream equivalent, the nyquist rate which equals twice the highest significant frequency in the analog signal is coded according to the required bit accuracy; i.e., if the highest significant frequency is F, then the bit rate for N bits per sample is

$$\text{Bit rate in Mbps} = 2.F.N$$

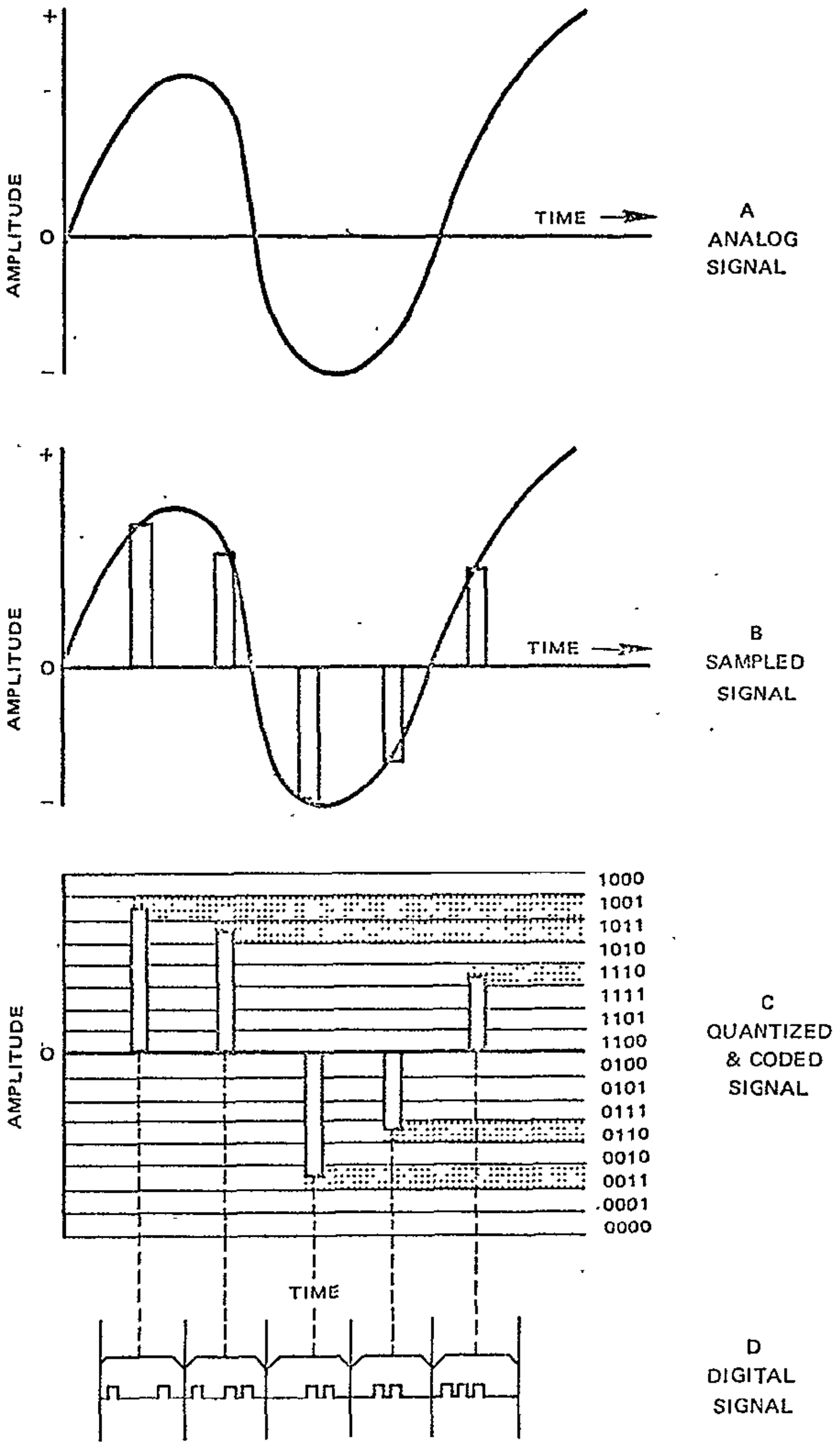

+
0
−
AMPLITUDE
TIME
A
ANALOG
SIGNAL
+
0
−
AMPLITUDE
TIME
B
SAMPLED
SIGNAL
AMPLITUDE
0
1000
1001
1011
1010
1110
1111
1101
1100
0100
0101
0111
0110
0010
0011
0001
0000
C
QUANTIZED
& CODED
SIGNAL
TIME
D
DIGITAL
SIGNAL

Figure 4-2 The Technique of Converting an Analog Signal into a Digital Signal by Sampling, Quantizing, and Coding - after D. F. Hoth

For image information, however, more than two samples per cycle are used -- as high as five. This requires trade-offs to be established between increased sample rates leading to an increase in detail, versus the reduction in signal-to-noise ratio due to reduced energy-per-bit due to the reduced duration time of each bit.

For systems where the data link capacity is to be the limitation to resolution, a sample rate of 2.8 samples per cycle of baseband data is required. For a 2-bit digital system, the bit rate will therefore be 5.6 times the highest baseband frequency which is desired to be resolved. A data link limited system will give the maximum quantity of information picture elements for a given data bit rate limitation.

In systems that are intended to be sensor bandwidth limited, experience has indicated that a minimum sample rate of 3.6 samples per baseband cycle are required. A greater sample rate than 3.6 samples per cycle will not yield detectably increased performance and should not be exceeded if some control over data link efficiency is required.

An actual system might choose some compromise between the two extremes, depending upon the nature of the mission and the input data characteristics. Various combinations of video baseband bandwidth and data link bit rates for an optimized 2-bit delta modulator are listed in Table 4-1.

The current Philco-Ford two-bit delta modulator has a maximum sampling rate of 100 megasamples per second (200 Mbps data rate). This sampling rate limitation occurs because of the integrator loop propagation delay using current flight applicable circuit components. Using this maximum capability, the link limited baseband bandwidth is 36 MHz.

Figure 4-3 shows curves of digital data rate as a function of sensor video bandwidth for 2-bit delta modulation, 4-bit PCM and 6-bit PCM based on 2.8 samples per cycle. The 4-bit and 6-bit curves are equally applicable to delta modulation.

TABLE 4-1

Electrical vs Optical Analogs

Baseband Bandwidth	Link-Limited Data Link Bit Rate *	Sensor-Limited Data Link Bit Rate **
25 MHz	140 Mbps	180 Mbps
36 MHz	200 Mbps	260 Mbps
50 MHz	280 Mbps	360 Mbps
100 MHz	560 Mbps	720 Mbps

* 2.8 samples per cycle of baseband data.

** 3.6 samples per cycle of baseband data.

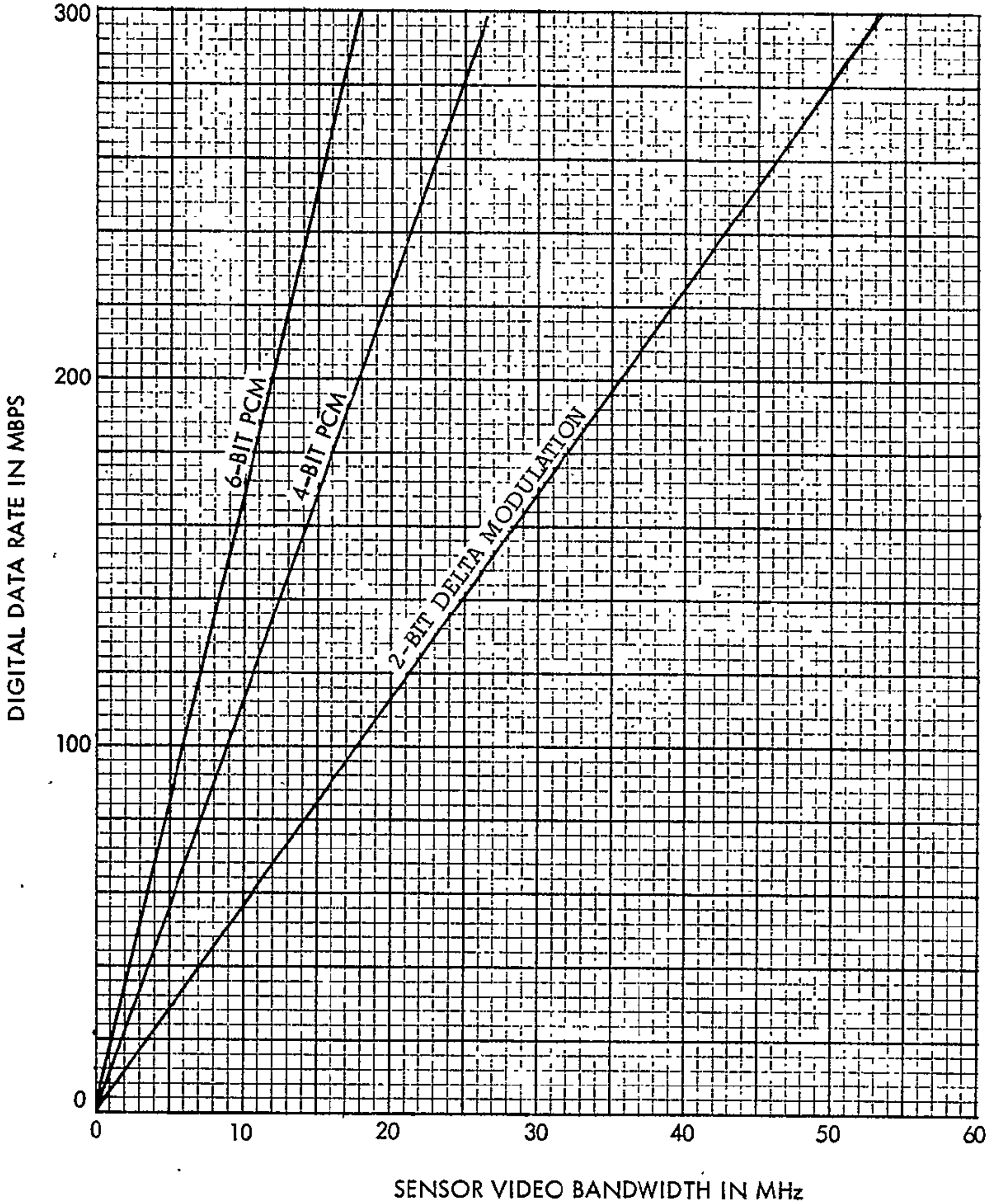


Figure 4-3 Digital Data Rate vs. Sensor Bandwidth
(2.8 samples/video cycle)

4.4 ALLOWABLE BIT ERROR RATE

Noise in the data link receiver will cause errors in the digital data input to the delta demodulator. Data bit errors affect the delta demodulator by causing incorrect step changes into the integrator and a resulting shift in gray scale of the output video. This shift in the gray scale will persist for the memory time of the integrator, which is typically ten samples in duration. The effect on the video output is a gray streak on line in width and ten samples in length.

Since the resolution of the video output is not affected by occasional digital errors, the data subsystem can tolerate a relatively large number of errors. Experience with operational transmission of video data using delta modulation indicates that useful data can be obtained with bit error rates as high as 10^{-2}. A threshold value of BER based on subjective evaluation of actual video information seems to be about 10^{-3}. A BER of 10^{-4} or lower gives excellent results.

4.5 TYPES OF ULTRA HIGH SPEED A/D CONVERTERS

A/D converters which provide conversion at rates less than 1,000,000 per second can use many candidate circuits, including parallel and series converter techniques, successive approximation circuitry.

In the 40 to 400 Mbps rate range, high speed components and a minimum number of successive logic levels leading to each bit must be used.

Problems of accuracy arise from the circuit problems inherent at these high rates, assuming that the individual logic devices and switching elements have switching times and delay intervals consistent with such speed.

- At these bit rates, every wire and connection becomes a transmission line resulting in requirement for termination to avoid reflectors.
- In low speed A/D converters, static accuracy is important. At ultra high speed, dynamic accuracy is most important.
- At ultra high speed bit rates, the threshold devices must slow and settle to the correct value in the specified period.
- Noise becomes a problem in ultra high speed A/D converters due to each wire acting as an antenna into which is cross coupled clock signals, data frequencies from other lines, and crosstalk.

At the 40-400 Mbps A/D converter speeds, few types of circuits can meet the requirement of speed, low noise and with accuracy. The following paragraphs will discuss the basic aspects of several candidate ultra high speed A/D converter techniques which can operate at these speeds; i.e.

- Delta modulation method of feedback encoding (redundancy removal)
- Cyclic conversion encoder
- Phase-plane conversion encoding
- Threshold encoders
- Scanning tube-encoder.

Delta modulation technology is particularly suited to the transmission of high speed sensor data, and will be discussed at length in this section; however all above methods and techniques will be discussed in the paragraphs to follow:

4.6 Delta Modulator/Demodulator Implementation

For pictorial video information, the most important parameters are good transient response or linear phase response, and low slope loading distortion. **Intermodulation distortion caused by non-linear amplitude response is generally** not critical. On the contrary, certain types of amplitude nonlinearities such as gamma correctors or contract enhancers may improve the output image quality.

Pictorial video transmission systems can also operate at signal-to-noise ratios considerably less than acceptable for quality audio transmission. This fact becomes more obvious when considering that very few end-to-end photographic or imaging systems can exceed the discrimination of more then 16 shades of grey. For this type of information, a transmission system with a signal-to-noise ratio sufficient to determine 100 discrete levels or shades of grey clearly has an excess of signal-to-noise ratio.

The channel capacity may be better utilized by increasing the baseband data rate. Lower signal-to-noise ratios can also be tolerated when the quantizating noise has the property that the peak noise amplitudes are discretely limited. For example, a digital system whose peak noise amplitude is less than a shade of grey will produce a better result (assuming contouring is not present) than an analog system with the same RMS signal-to-noise ratio, because the peak analog noise amplitudes will often exceed the visual detectability level.

For a multiple bit delta modulator, the optimum operation generally occurs when the slope response limitation of the delta modulator just meets or exceeds the maximum slew rate of input data. When this situation occurs, the system signal-to-noise ratio is maximized without losing the capability of resolving fine detail after a rapid grey-scale transistion. Care must be maintained not to over enhance the input data slew rate artifically by the use of techniques such as aperture correction.

The use of aperture correction can be used to produce nearly any desired modulation transfer function at a given spacial frequency -- at the expense, however, of signal-to-noise ratio. If an excessive amount of aperture correction is used, resulting in a poor high frequency signal-to-noise ratio, slope loading distortion will result unless the slew rate of the delta modulator increases to meet the input data slew rate.

An increase in the delta modulator slew rate always trades off an increase in quantization noise on a first order approximation. The actual noise increase may be an overall noise increase or an increase in the loop instability, which leads to excessive delta modulator signature. The optimum method chosen to increase the slew rate again depends upon input data characteristics.

Unfortunately, while aperture correction decreases primarily the high frequency signal-to-noise ratio, an increase in delta modulator slew rate decreases the signal-to-noise ratio across the baseband spectrum, resulting in an overall poorer system transfer characteristic when excessive aperture correction is used. An optimum system in which a delta modulator is incorporated places a significant amount of the system aperture compensation after the delta demodulator.

In summary, an optimum choice of delta modulator parameters must include consideration of the following characteristics of the input data.

- Data type and critical distortions
- Modulation transfer function
- Input data slew rate
- Input data noise distribution across total baseband spectrum
- Statistical peak amplitude distribution

2-bit delta modulation provides the optimum choice to meet these considerations, even though 3-bit delta modulation offers improvement in accuracy as pointed out earlier in this section.

4.6.1 The Two-Bit Delta Modulator - Practical Circuit Aspects

The two-bit delta modulator employs three levels of quantization of the error between the integrator and the input. Small differences cause a small step correction, which determines the basic peak-to-peak noise, as in the case of the one-bit modulator. However, when larger errors occur, the application of large steps to the integrator greatly improve the slew rate relative to the one-bit modulator. The two-bit delta modulator does not, of course, match the true PCM analog-to-digital converter in the ability to slew the entire peak-to-peak video in one sample. However, it is fortunate that virtually all practical video systems have high frequency responses which are significantly attenuated. For example, limiting resolution in many systems is determined at approximately 2 percent response. Although a typical two-bit delta modulator requires 10 samples to slew the operating peak-to-peak video range, it will not be the resolution limiting factor in most video systems.

The operation of the delta modulator is illustrated in Figure 4-4. The differential amplifier senses the difference between the integrator voltage and the analog input. Its output level is applied in parallel to the three "quantizers" which, together, constitute a 2-bit parallel analog-to-digital converter.

Each of the quantizers is identical except for the threshold level voltage and the charge increment output to the integrator. In each, the input voltage is quantized as above or below the threshold voltage and time-sampled to set a flip-flop accordingly.

The flip-flop outputs are logically combined (now shown) for transmission to the delta demodulator. They also control the charge switches which increment or decrement the integrator. The delta demodulator employs "quantizers" which do not need the sharp threshold circuit and have much less rigorous requirements for narrowness in the sample gate aperture time interval.

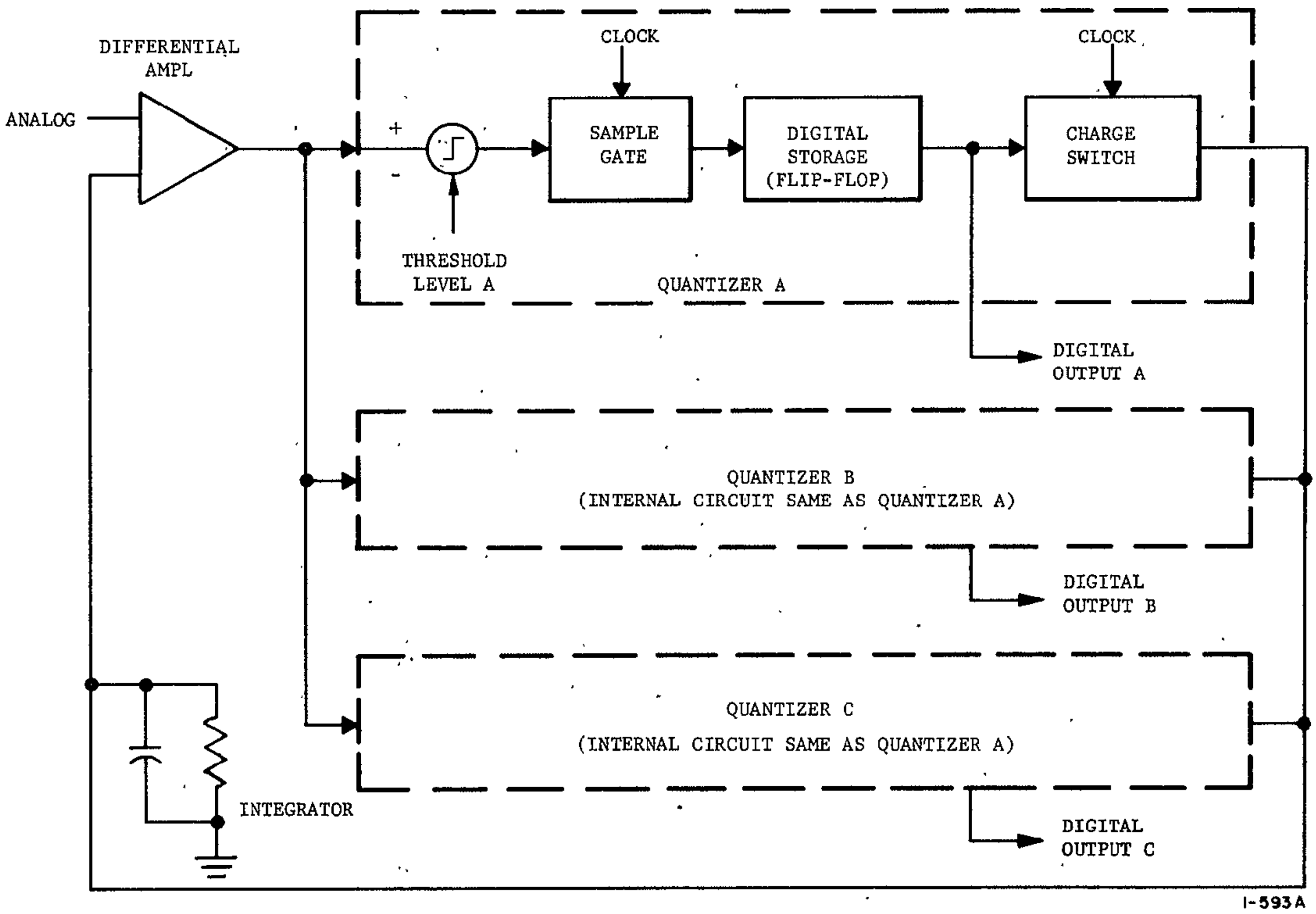

DIFFERENTIAL
AMPL
ANALOG
+
-
CLOCK
SAMPLE
GATE
DIGITAL
STORAGE
(FLIP-FLOP)
CLOCK
CHARGE
SWITCH
THRESHOLD
LEVEL A
QUANTIZER A
DIGITAL
OUTPUT A
QUANTIZER B
(INTERNAL CIRCUIT SAME AS QUANTIZER A)
DIGITAL
OUTPUT B
QUANTIZER C
(INTERNAL CIRCUIT SAME AS QUANTIZER A)
INTEGRATOR
DIGITAL
OUTPUT C
I-593A

Figure 4-4 Two-Bit Delta Modulator, Basic Operation

In Figure 4-4 typical timing waveforms are shown with a hypothetical integrator response to a step input. The scale shown represents the greatly magnified detail of an infinitely sharp edge. As previously explained, this is not representative of typical video sources. (Note that, because of the integrator decay, the pattern of steps required to maintain the static level after the edge contains more positive (+S) than negative steps (-S).)

4.6.1.1 Role of the Integrator

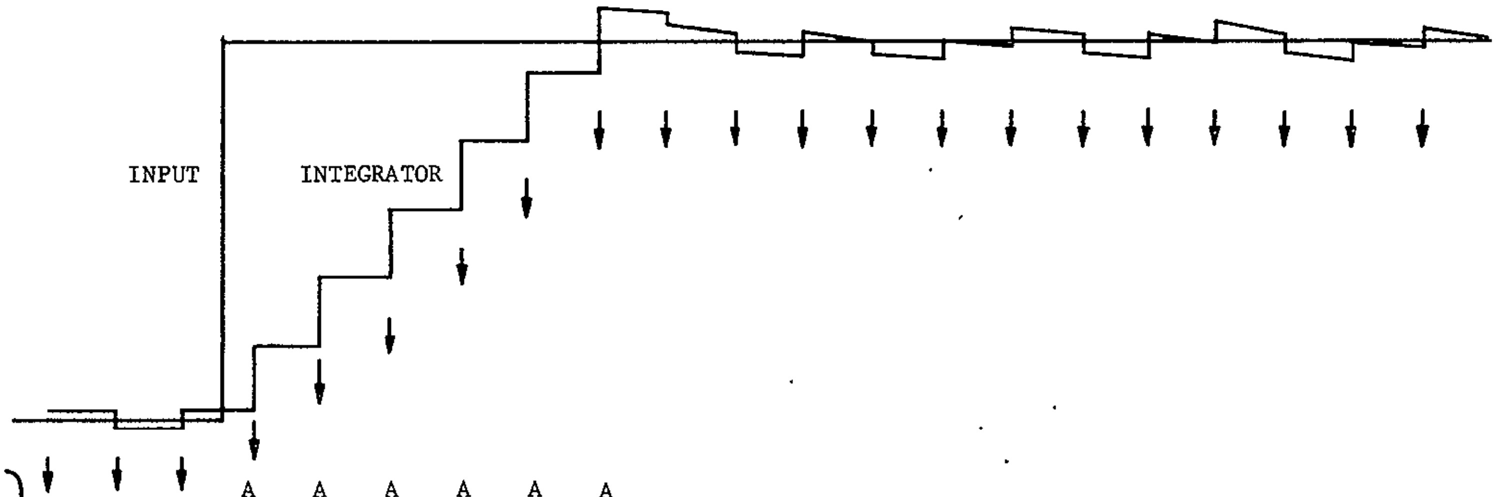
In Figures 4-5 through 4-7 and Table 4-2, integrator characteristics are illustrated and related to transmission errors, transient response, and advantages versus disadvantages.

Prediction techniques are in general, costly from the standpoint of speed since such techniques involve one or more sequences of logic levels. Experimental integrator characteristics have the best overall characteristics relative to noise immunity and adaptability to optimizing image data.

4.6.1.2 Variable Step-Size Delta Modulator

One type of variable slope delta modulator which has a relatively simple implementation is shown in Figure 4-8. The step size is increased exponentially with a +++ or --- pattern of the last 3 bits, remains unchanged for the first transition after the first change from a repetition sequence, i.e., a pattern +-- or -++, and decreases otherwise.

This type of variable slope delta modulator uses a shift register in series with the integrator and has a much improved transient response and dynamic range as compared to the linear slope system. Distortion for sinusoidal inputs is much improved also. Minimum step sizes can be selected to give SNR = +35 dB for both types at low modulation frequencies $f_{clock}/f_{sinusoid} > 200$. However $f_{clock}/f_{sinusoid} = 100$ gives SNR of +28 dB for variable slope as compared to only +4 dB for linear slope.



THRESHOLD LEVELS EXCEEDED	B C	C	B C	A B C	A B C	A B C	A B C	A B C	A B C	C	C	B C	C	B C	B C	C	B C	B C	C	B C	B C
STEP SIZE S = SMALL L = LARGE	+S	-S	+S	+L	+L	+L	+L	+L	+L	-S	-S	+S	-S	+S	+S	-S	+S	+S	-S	+S	+S
DIGITAL OUTPUT TO DELTA DEMODULATOR	0,1	1,0	0,1	1,1	1,1	1,1	1,1	1,1	1,1	1,0	1,0	0,1	1,0	0,1	0,1	1,0	0,1	0,1	1,0	0,1	0,1

I-602A

Figure 4-5 Response of the Two-Bit Delta Modulator to a Step Input

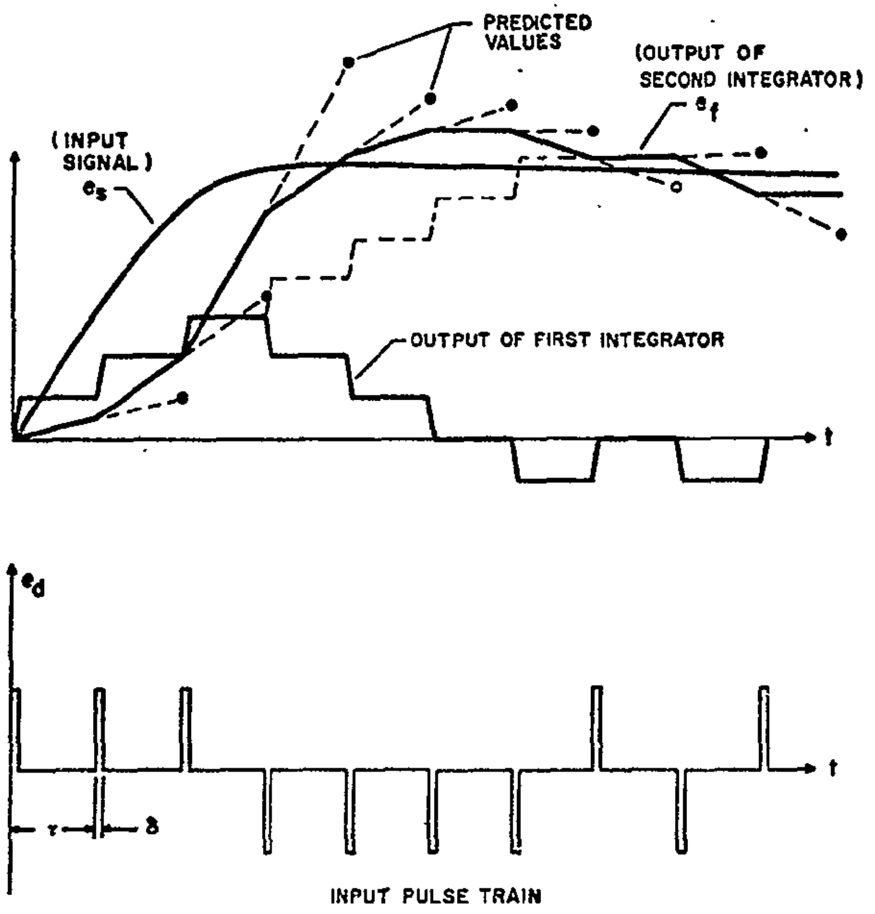
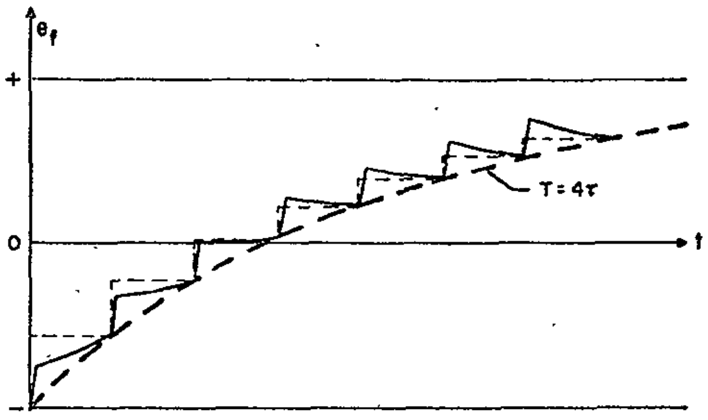
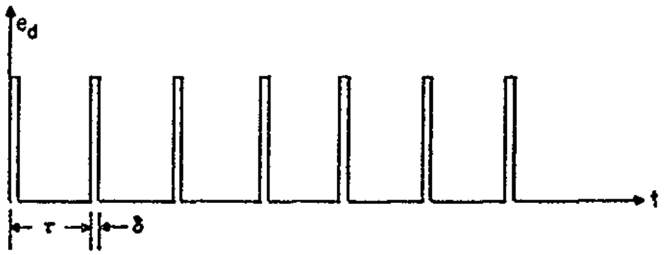


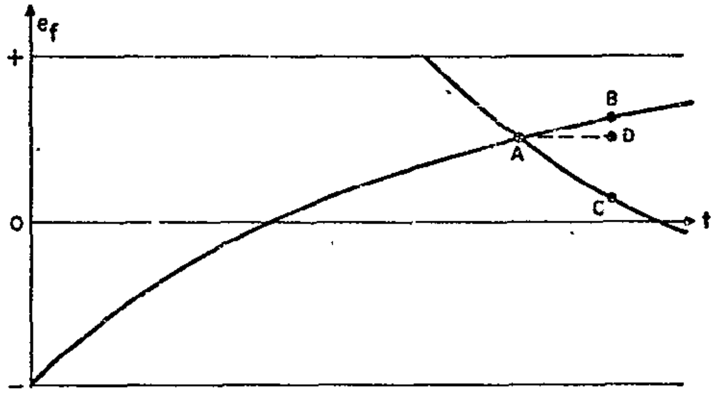
Figure 4-6 DELTA MODULATION USING DOUBLE INTEGRATION WITH PREDICTION



a. OUTPUT OF INTEGRATOR



b. INPUT PULSE TRAIN



c. CHARGING PATH AS A FUNCTION OF POSITION AND DIRECTION

Figure 4-7 CHARACTERISTICS OF EXPONENTIAL INTEGRATION

TABLE 4-2

INTEGRATOR CHARACTERISTICS

Type	Increment	Number of Reproducible Levels	Effects of Transmission Errors	Length of Time Constant	Transient Response	Implementation Considerations	Advantages	Disadvantages
Single Linear	Fixed steps	Limited by step size	Entire Scan line	Long	Limited by no. of steps to cover dynamic range of input signal to delta system.	Attentuation integrator network. Gain required		Contouring Poor Noise
Double Linear	Fixed Slopes	Continuous	Entire Scan line	Long	Limited by size of steps in out put of first integrator	Same	Reduced Contouring Better rise time	Overshoot Poor Noise immunity
Double Linear with Prediction	Fixed slopes	Continuous	Entire scan line	Long	Same; also limited by prediction interval used	Same	Reduced overshoot. Good rise time	Poor Noise immunity
Exponential	Variable steps	Continuous	Few elements Time of recovery directly proportional to time constant	Short	Limited by time constant	Relatively easy to implement	Reduced contouring. Good noise immunity Makes use of probability characteristics of image data	High amplitude oscillation about DC

1K-DA218U

One alternative approach is to examine the last 3 or 4 output bits of the delta modulator and to control the step size as follows:

	Output Sequence	Step Size
3 bit storage	+ + +	+2
	+ + -	+1
	+ - x	+1
4 bit storage	+ + + +	+3
	+ + + -	+2
	+ + - x	+1
	+ - x x	+1

The symbol x stands for "don't care." Negative steps correspond to inverted output sequences. The bit at the farthest left is the last bit transmitted.

The step response of the encoders can be compared with the linear slope M as shown in Figure 4-9. As can be seen the 3 stage delta modulator responds in 7 steps or half the time required of the linear slope delta modulator. The 4-bit storage M responds in only 5 steps.

4.6.2 Basic Delta Modulator System

A Delta Modulator will receive analog video and will produce non-return-to-zero (NRZ) digitized output. It will include a self-contained sampling clock, a means to test the video response of the Delta Modulator and Demodulator upon command, and a means to test the link error rate by a commanded digital pseudo noise (PN) generator. Provision can optionally be made for generating precise digitally transmitted video sync actuated by the video source sensor.

A block diagram of an illustrative Delta Modulator system is given in Figure 4-10. The basic Delta Modulator function is contained in the lower half of the diagram, encompassing the video input amplifier to the serial-to-parallel converter.

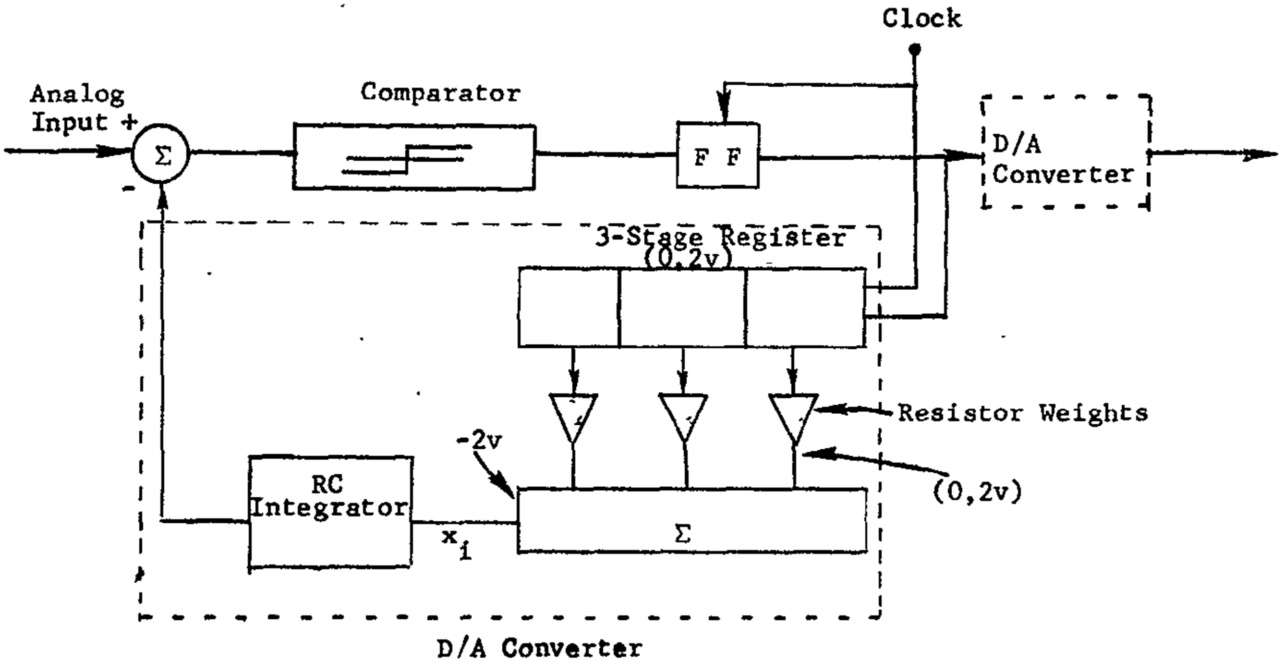

Clock
Analog Input +
-
Σ
Comparator
F F
D/A Converter
3-Stage Register
(0,2v)
Resistor Weights
(0,2v)
-2v
RC Integrator
x_i
Σ
D/A Converter

Figure 4-8 Spilker - Luby Variable Slope A/D Converter

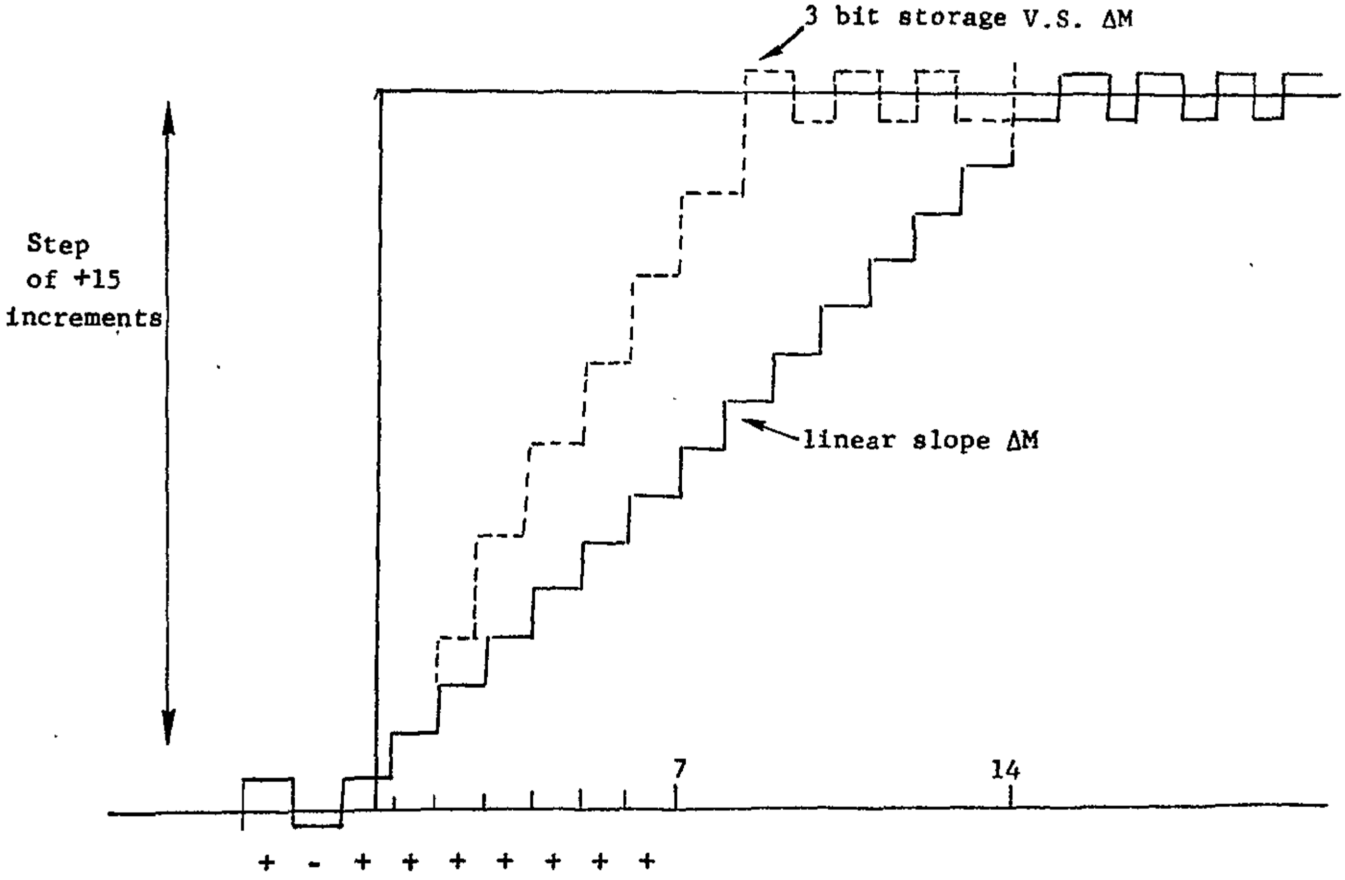


Figure 4-9 Increase in A/D Conversion Slope Due to Edge Effects

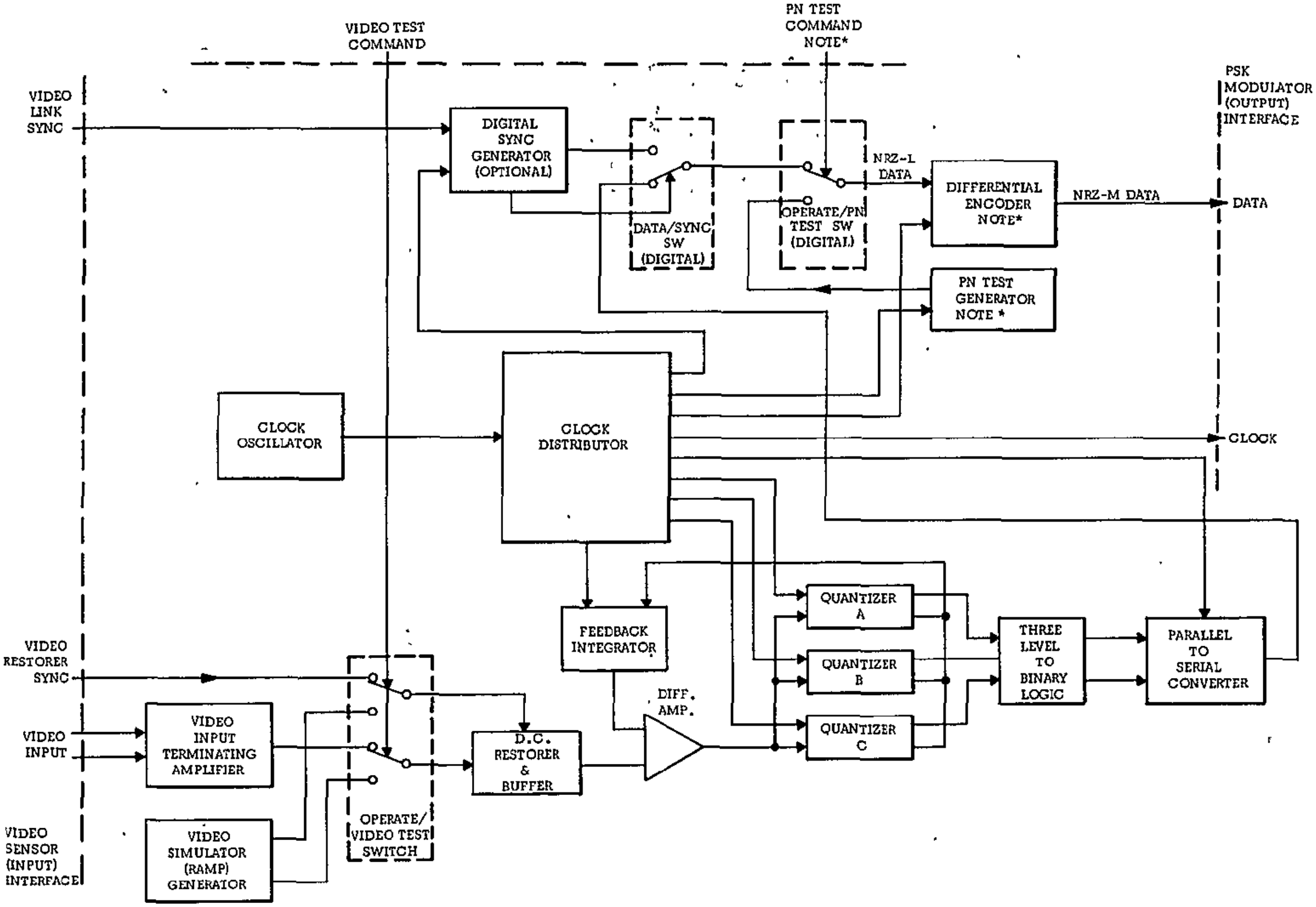

VIDEO TEST COMMAND
PN TEST COMMAND NOTE*
VIDEO LINK SYNC
DIGITAL SYNC GENERATOR (OPTIONAL)
DATA/SYNC SW (DIGITAL)
OPERATE/PN TEST SW (DIGITAL)
NRZ-L DATA
DIFFERENTIAL ENCODER NOTE*
NRZ-M DATA
PSK MODULATOR (OUTPUT) INTERFACE
DATA
PN TEST GENERATOR NOTE *
CLOCK OSCILLATOR
CLOCK DISTRIBUTOR
CLOCK
FEEDBACK INTEGRATOR
QUANTIZER A
QUANTIZER B
QUANTIZER C
THREE LEVEL TO BINARY LOGIC
PARALLEL TO SERIAL CONVERTER
VIDEO RESTORER SYNC
VIDEO INPUT
VIDEO INPUT TERMINATING AMPLIFIER
OPERATE/ VIDEO TEST SWITCH
D.C. RESTORER & BUFFER
DIFF. AMP.
VIDEO SENSOR (INPUT) INTERFACE
VIDEO SIMULATOR (RAMP) GENERATOR

Figure 4-10 Delta Modulator Block Diagram

The interface with the video sensor is accomplished at the video input terminating amplifier. The video input line is assumed to be a floating-shield, unbalanced (single-ended) coaxial cable. Triaxial cable may also be employed where the outer shield is a.c. bypassed to the outer case ground. This is especially recommended if UHF or lower frequency radio transmitters or any other on-board pulse or continuous interference sources exist. The input is nominally terminated in the cable impedance to insure flat video response up to the 25 MHz maximum. If a long video cable is anticipated, (e.g., over 1 1/2 meters for 25 MHz), then the use of either additional source (sensor) output termination or a precision, adjustable, cable termination at the Delta Modulator is recommended.

The video input level is nominally 1.0 volt peak-to-peak, single-ended. The shield is d.c. isolated to avoid spacecraft d.c. ground loops. Differential input amplification will reject low frequency common mode signals (i.e., ground noise) by at least 60 dB. Higher frequency common mode signals (above 100 KHz) will be bypassed from the input shield to local ground. The use of lower signal levels (e.g., as low as 120 millivolts peak-to-peak) is readily feasible in order to conserve driving power providing that the signal to common mode noise can be kept above unity within the video passband. Such an interface is potentially critical, but it has been shown that it is entirely manageable.

The input video is assumed to contain a periodic, stable reference level for restoring d.c. levels. The duty cycle of this reference level may be made as small as 0.1 percent or a minimum of 200 nanosec, whichever is larger, providing that an initialization cycle is utilized prior to data transmission. This initialization cycle must contain a number of reference level periods transmitted in the normal cycle to permit the d.c. restorer to stabilize. Three hundred to five hundred such periods, or an accumulated total equal to the repetition time plus 10 usec, (whichever is larger) is suggested as a guide.

A more rapid response requirement would increase complexity and require special attention to freedom from noise in the video source reference levels. These allowances do not reflect the requirements of possible d.c. restoring in the receiver video system inasmuch as the Delta Demodulator provides inherent d.c. video response. Restoring d.c. in the receiver is not recommended unless careful consideration has been given to the influence of the link error rate on the required integration (i.e., time lag) in restoring, which requires both a minimum duty cycle and a minimum initializing cycle. These requirements generally supersede those listed immediately above for transmitter restoring.

The video simulator (ramp) generator allows for a commanded operational check of virtually the entire Delta Modem (including the Demodulator) and the satellite-to-ground link. The ramp waveform is chosen as the simplest means of test, requiring only a few components and less than 50 milliwatts of power. As shown, the ramp generates its own restoring sync pulse. However, it may be desirable also to synchronize the ramp to the sensor restorer sync in order to simplify evaluation of the received signal. This presumes that the ramp is not used for standby transmission while the video sensor is shut off.

The differential amplifier, the three quantizers (A, B, and C), and the feedback integrator constitute the feedback loop i.e., the heart of the Delta Modulator.

Circuit technology is significantly different in the feedback loop for the 40 Mbps version as compared to the 200 Mbps version. In the former, almost all functions can be implemented with monolithic silicon devices in stock packages. Digital logic, for example, is one of several high-speed TTL families, e.g., "SUHL II."

Signal wiring must, of course, be reasonably short, and considerable attention is required to ground plane and power plane bypassing. In the 200 Mbps version, custom hybrids (thick film and chip) employ state-of-the-art, high gain-bandwidth devices in a reliable, all-silicon design. Signal interconnections are direct, point-to-point. Digital logic is a high-speed ECL, e.g., MECL III. Power dissipation is higher consistent with the current levels needed for rapid charge and discharge of circuit capacitances. These considerations require careful interlocking of electrical and packaging design in a manner akin to microwave circuit packaging.

The clock oscillator can be a simple crystal oscillator operating in the 25 MHz region followed by frequency multiplication to the operating bit rate. The stability requirements of this oscillator are readily met without an oven by proven temperature stabilizing techniques. The resulting long-term drift is overshadowed by the doppler uncertainty, and the short-term jitter is insignificant within the power bandwidth product available for the receiver clock tracking look.

The clock distributor provides both bit rate and sample rate (= 1/2 bit rate) clock to various functions. It is composed of a binary frequency divider and parallel output amplifiers, i.e., digital gates.

The logic levels from the three quantizers are converted to two binary data streams emerging at the sampling rate. The conversion occurs as listed as follows in Table 4-3.

TABLE 4-3

Quantizer to Bit-Stream Conversion

QUANTIZER THRESHOLD EXCEEDED	INTEGRATOR STEP SIZE	DIGITAL OUTPUTS D1, D2
A, B, and C	Pos. Large	1, 1
B, and C	Pos. Small	1, 0
C	Neg. Small	0, 1
None	Neg. Large	0, 0

The two digital outputs D1, D2 are arranged in serial sequence for transmission. Both the three-level-to-binary and the parallel-to-serial conversions are handled digitally by high-speed ECL logic.

4.6.2.1 Digital Synchronization

Digital synchronization is preferred as a convenience (optional) to replace an analog sync tone burst transmission. It offers better link efficiency by reducing the time required to transmit sync and greatly lessens time jitter in the received sync at the expense of another interface, and a small amount of logic.

The digital sync generator operates by generating a high-speed sync word such as five repetitions of a seven-bit Barker code at the bit rate when a video sync signal is received from the sensor. This sync word is inserted in the data stream, presumably during a video dead time (e.g., within the reference signal; that is, the d.c. restorer sync period). Assuming that the video sync follows a stable cycle period and falls within, for example, a ten bit interval ("window") over several consecutive cycles, then straightforward digital detection techniques in the receiver will give fairly rapid lock-up and negligible probability of false sync and loss of sync at any usable bit error rate. The actual performance parameters achievable depend, of course, on a numerical knowledge of the problem, i.e., the cycle period and the maximum tracking rate required.

To conserve power, the Barker code generator and cycle of five counter are deactivated between sync periods. The average power is thus reduced considerably. For a nominal peak power of four watts and a conservative maximum of 5 percent sync time, the average power is 200 mw.

4.6.2.2 PN Test Generator

The PN test generator can be utilized in the high-speed model of the Delta Modulator in order to certify acceptable bit error rates during operational pre-use tests. The PN generator provides a fixed maximum length sequence from 11 storage elements, i.e., a 2047 bit sequence. It is actuated by command as required.

The PN test generator utilizes 11 high-speed (ECL) flip-flops, and digital gates in a standard regenerative configuration. See Figure 4-11.

3 4

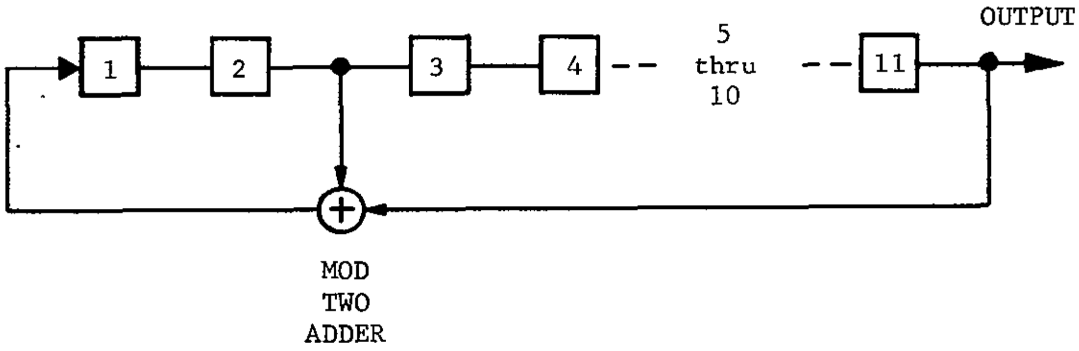

OUTPUT
1
2
3
4
5
thru
10
11
MOD
TWO
ADDER

Figure 4-11 PN Test Generator

4.6.3 Delta Demodulator Configuration

The essential function of the Delta demodulator is to duplicate the integrator voltage of the Delta modulator; this duplicated voltage then becomes the video output. The Delta modulator performs this function by rearranging the incoming digital data (from the differential decoder output) into the same three-level signals that existed in the Delta modulator flip-flops. This is done by converting serial back to parallel, and parallel data back to three-level which is the inverse of the modulator. High-speed ECL logic is again utilized.

The Delta demodulator quantizer duplicates the flip-flop and charge switches utilized in the Delta modulator quantizer. Therefore, the similar delta demodulator integrator receives similar charge increments and thus has an output slaved to the delta modulator.

A block diagram of the delta demodulator is shown in Figure 4-12.

4.6.3.1 PN Test Signal Comparator

The delta demodulator will include a pseudo-noise (PN) sequence generator identical to that in the spacecraft with the exception that the generator can be rapidly locked in sequence synchronization with the received PN sequence, then used as a reference for comparison on a bit-by-bit basis with the received sequence in a modulo-two adder. The output of the mod-2 adder is ANDed with the bit clock. When a received bit is in error, a clock pulse will be gated-out to an external digital counter.

4.6.3.2 Allowable BER

Noise in the data link receiver will cause errors in the digital data input to the delta demodulator. Data bit errors affect the delta demodulator by causing incorrect step changes into the integrator and a resulting shift in gray scale of the output video. This shift in the gray scale will persist for the memory time of the integrator, which is typically ten samples in duration. The effect on the video output is a gray streak one line in width and ten samples in length.

Since the resolution of the video output is not affected by occasional digital errors, the data subsystem can tolerate a relatively large number of errors. Experience with operational transmission of video data using delta modulation indicates that useful data can be obtained with bit error rates as high as 10^{-2}. A threshold value of BER based on subjective evaluation of actual video information seems to be about 10^{-3}. A BER of 10^{-4} or lower gives excellent results.

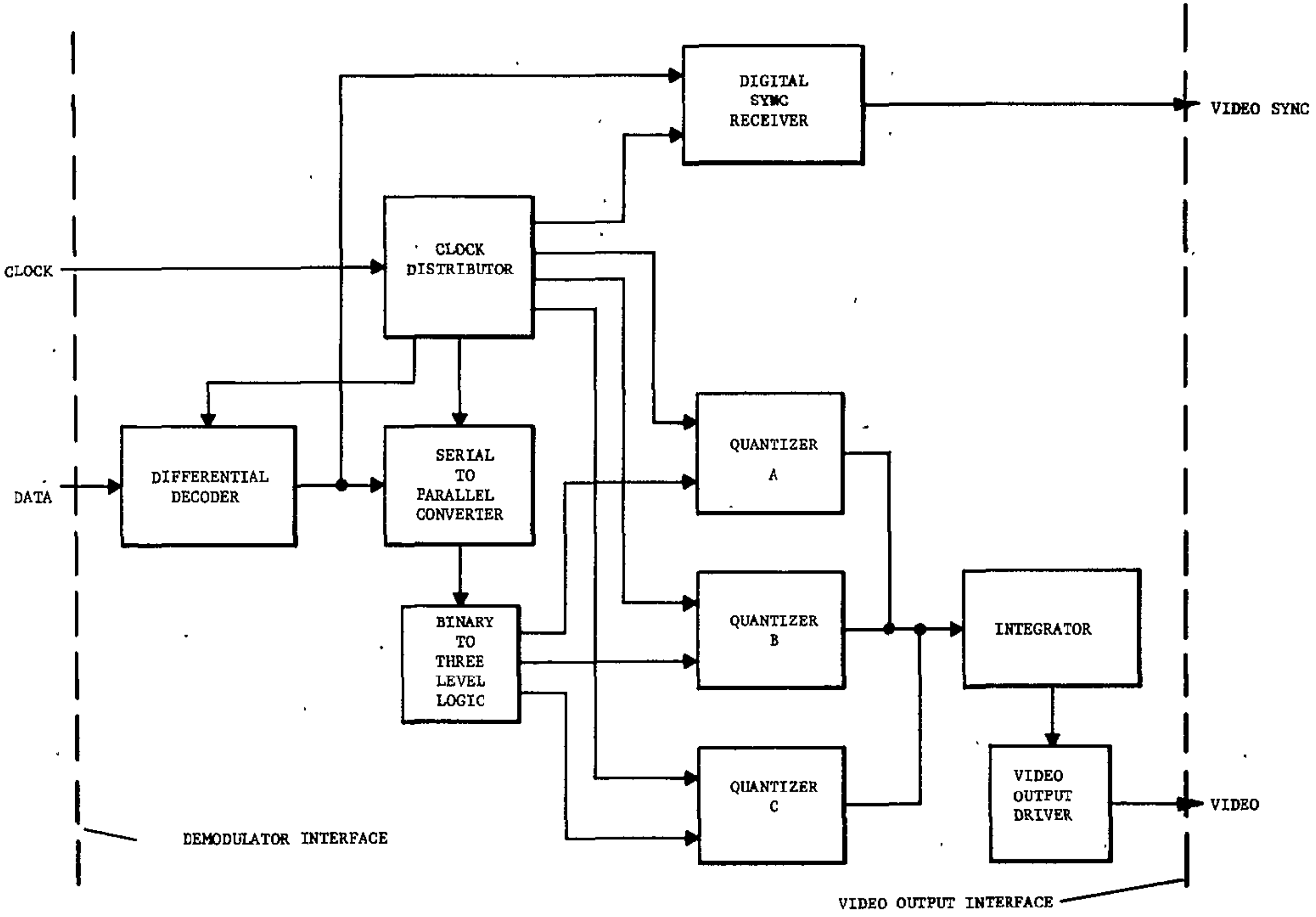

DIGITAL SYNC RECEIVER
VIDEO SYNC
CLOCK
CLOCK DISTRIBUTOR
DATA
DIFFERENTIAL DECODER
SERIAL TO PARALLEL CONVERTER
BINARY TO THREE LEVEL LOGIC
QUANTIZER A
QUANTIZER B
QUANTIZER C
INTEGRATOR
VIDEO OUTPUT DRIVER
VIDEO
DEMODULATOR INTERFACE
VIDEO OUTPUT INTERFACE

Figure 4-12 Delta Demodulator Block Diagram

4.7 MULTIPLEXING AND DEMULTIPLEXING DIGITAL BIT STREAMS

When more than one sensor provides information to be transmitted across the image data link, multiplexing and demultiplexing of bit streams must be used. The following discusses the technology involved.

4.7.1 Interleaving Digital Pulses

Bell Telephone Laboratories has designed an experimental high-speed pulse code modulation system to transmit 224 million bits per second over a coaxial cable. This is accomplished by developing a series of bit streams representing various information signals and various timing and synchronizing pulses. The pulse streams of these individually coded signals can be interleaved to form a 224-Mb/s bit stream.

In general, when a number of pulse streams are to be combined into a single stream, the rates of the streams may differ and the transmission times of each from source to multiplexing point may vary with time due to temperature changes and other factors. One way of synchronizing such a system is by locking all of the clocks at the sources to a single master reference clock and using elastic buffer stores to absorb time delay variations. This is practical as long as the sources are close to the multiplexing point.

4.7.2 Time Sharing of A/D Converters

Time-sharing of a/d converters can be employed only if the over-all system will permit reductions in accuracy and speed.

Only when the a/d converter is much more complex than the circuitry needed to multiplex one channel will time sharing offer a reduction in hardware. But only when the error and reduction in conversion rate introduced by the multiplexing circuitry can be tolerated is time sharing desirable.

Every time an analog signal is processed by some circuit, no matter how simple, an error is introduced. Such is the case with time-sharing, where analog signals are connected sequentially in time to the a/d converter. Time-sharing deteriorates the over-all conversion accuracy: errors due to time-sharing generally increase with the number of signals being multiplexed.

There is another penalty paid for time-sharing; namely, conversion rate. When "n" analog signals are sequentially converted in time, each signal is converted only at a conversion rate of 1/n. Using the example of 16 inputs again, and assuming that the maximum conversion rate is 16,000 per second, when each of the 16 inputs can be converted only at a rate of 1000 per second. Obviously, any economy in hardware through time-sharing can be achieved only by sacrificing conversion speed.

4.7.3 Multiplexed Delta Modulator and Demodulator Systems

Two methods of digitizing and multiplexing satellite sensor data are considered. In the first method shown in Figure 4-13, each sensor analog video output is fed to a separate delta modulator, which can be a 1, 2, 3 or 4 bit unit, depending upon the nature of the analog data it must quantize and encode. The output of each delta modulator is fed to a multiplexor which generates a serial bit stream. An advantage of this technique is each delta modulator is specially designed to handle its own specific analog input. This approach lends itself to a multiple sensor system configuration.

The second method shown in Figure 4-14 involves the feeding of analog outputs from identical sensors to a multiplexing delta modulator. In this approach, each sensor input together with its related feedback integrator is sequentially switched so as to form, during the sample time, a complete delta modulator unit. Thus, the summing unit and quantizer is common to all inputs.

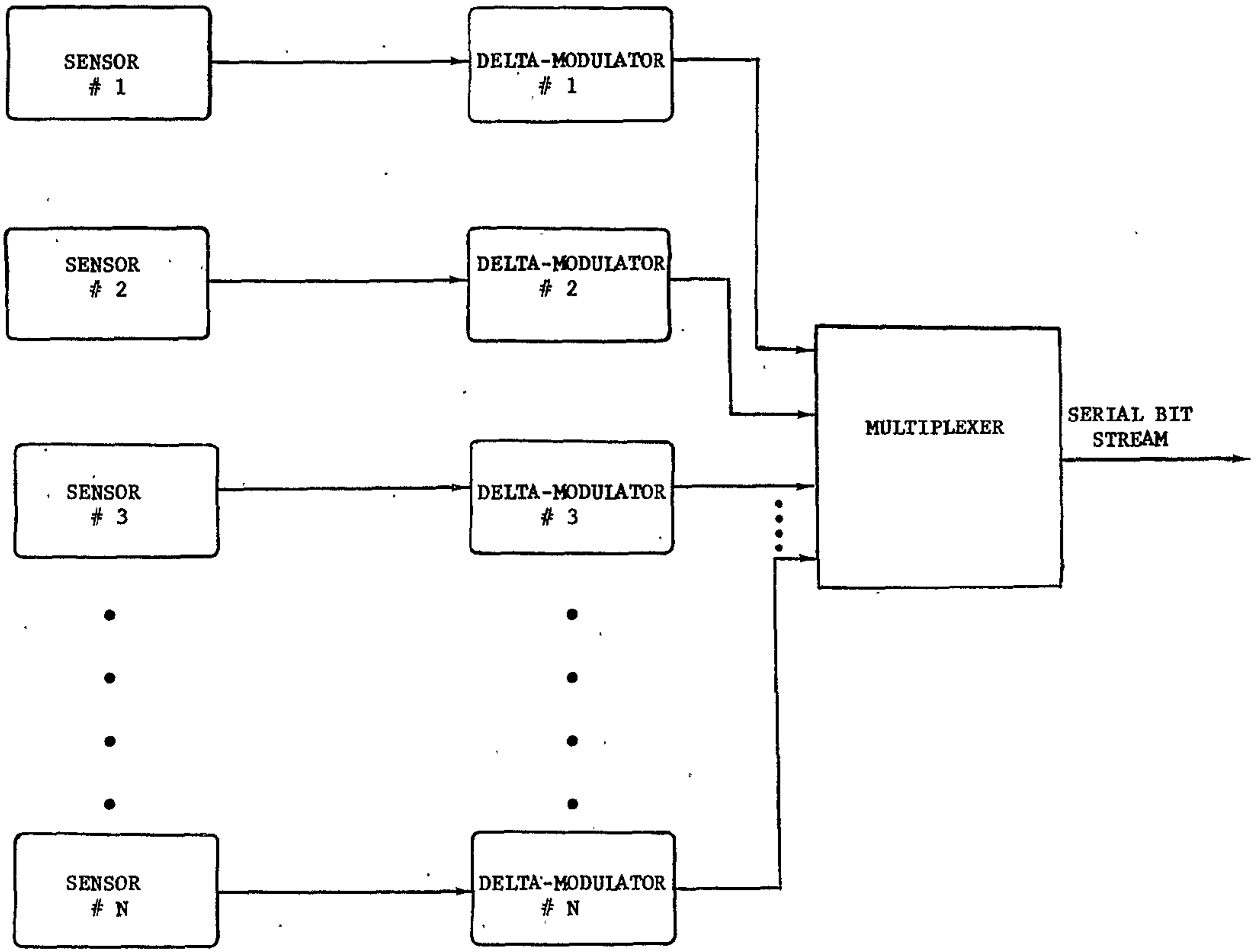

SENSOR # 1
DELTA-MODULATOR # 1
SENSOR # 2
DELTA-MODULATOR # 2
SENSOR # 3
DELTA-MODULATOR # 3
SENSOR # N
DELTA-MODULATOR # N
MULTIPLEXER
SERIAL BIT STREAM

Figure 4-13 System for Multiplexing Individual Delta Modulators

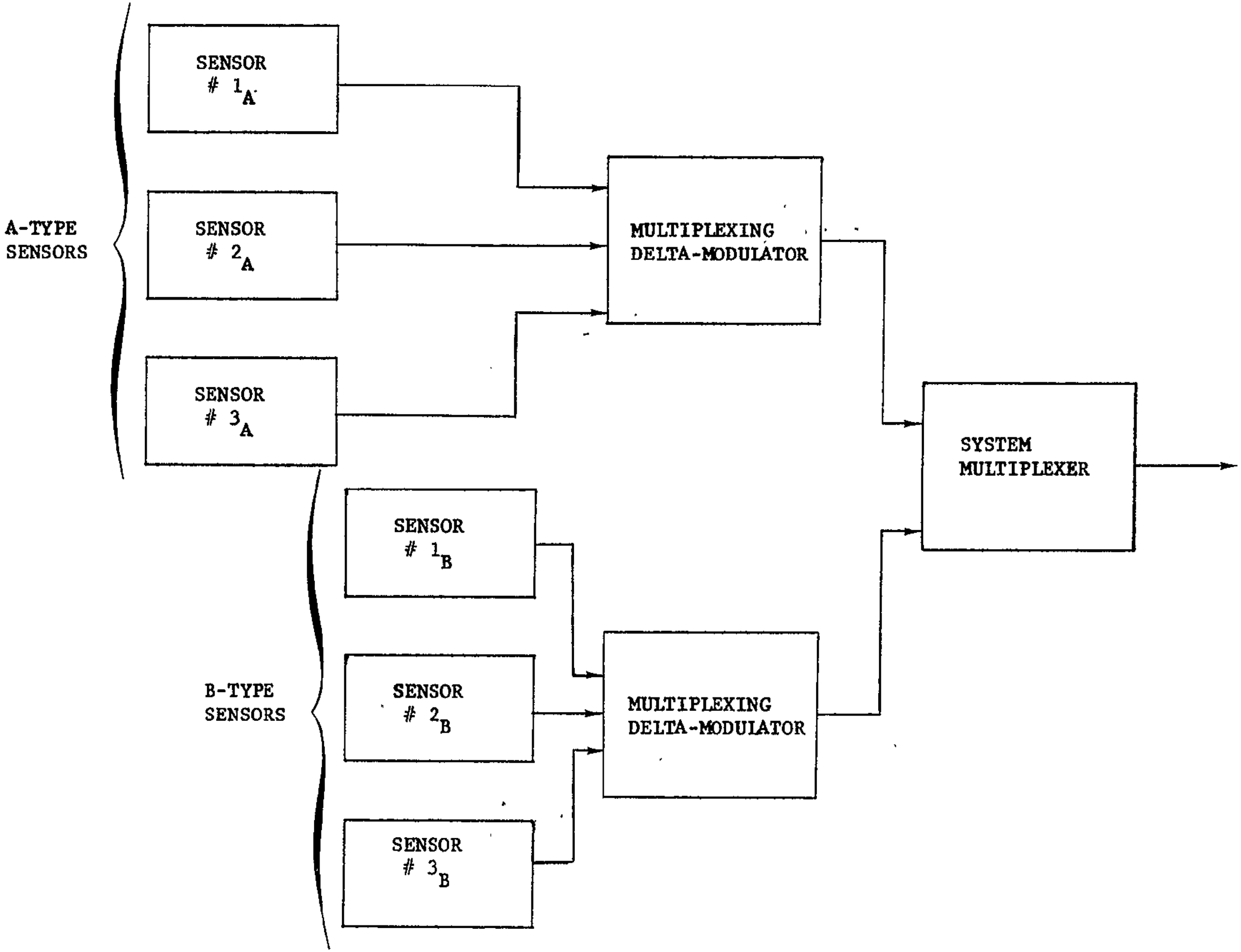

A-TYPE SENSORS
SENSOR # 1_A
SENSOR # 2_A
SENSOR # 3_A
MULTIPLEXING DELTA-MODULATOR
SYSTEM MULTIPLEXER
B-TYPE SENSORS
SENSOR # 1_B
SENSOR # 2_B
SENSOR # 3_B
MULTIPLEXING DELTA-MODULATOR

Figure 4-14 Dual System for Multiplexing Different Sets of Sensors

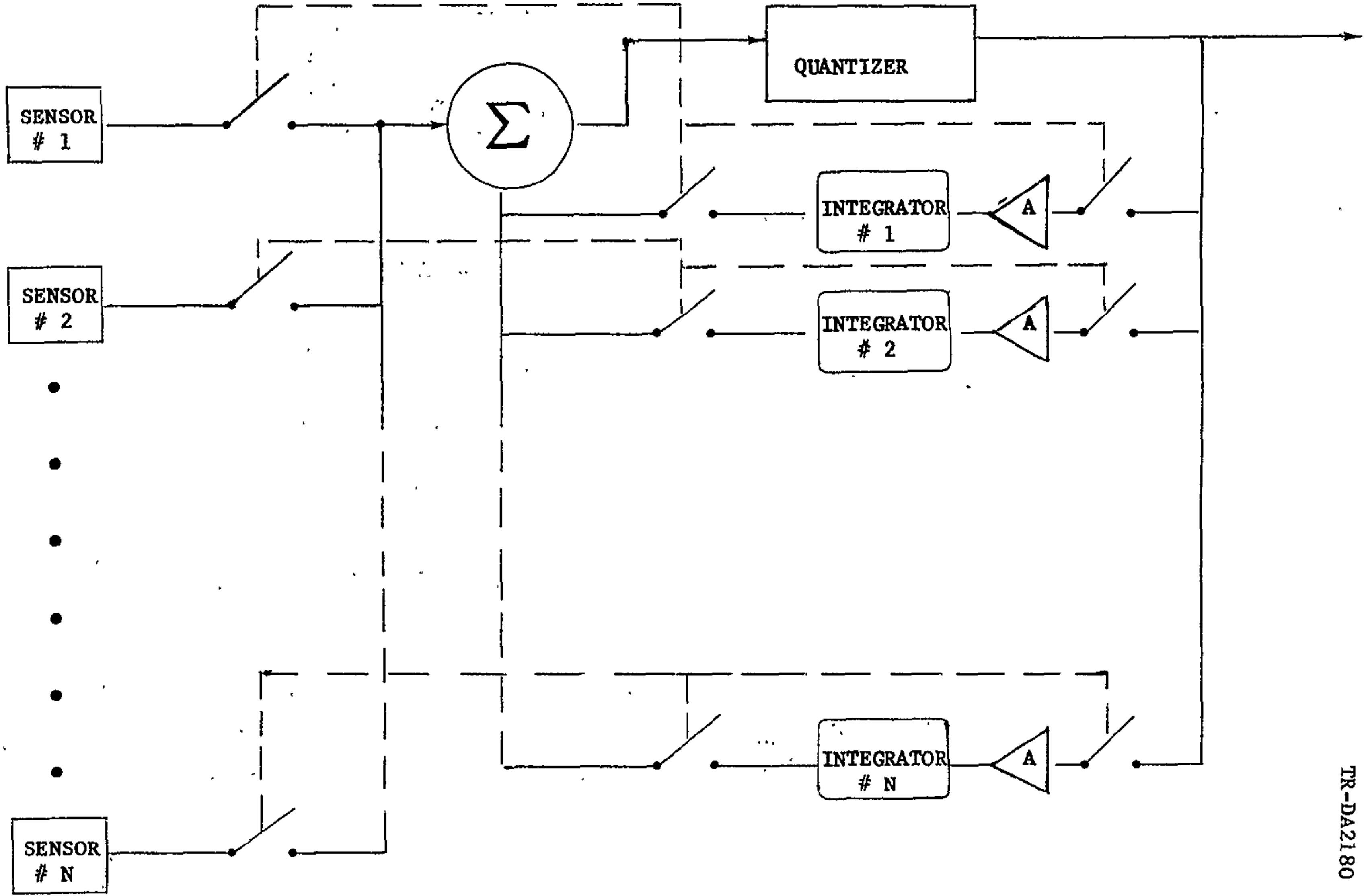

SENSOR # 1
SENSOR # 2
SENSOR # N
Σ
QUANTIZER
INTEGRATOR # 1
INTEGRATOR # 2
INTEGRATOR # N
A
A
A

Figure 4-15 Multiplexed Sensor Data by Switched Integrators

A logical variation on this second method would be the use of, say, two multiplexing delta modulators to handle two different groups of sensors, the output of these delta modulators in turn being multiplexed to form one serial bit stream. Refer to Figure 4-15.

The major constraint on the satellite sensor system configuration is the maximum data rate attainable by delta modulation. This upper bound is, of course, a function of current and projected developments in high speed digital circuitry. At this point in time a 200 MBPS data bit rate is deemed highly practical.

For an imaging type sensor system (vidicon, image orthicon, etc.) the major system constraints are dictated by the desired coverage and resolution. These two factors determine the kind and numbers of sensors employed. Readout time is determined by mission requirements and by individual sensor characteristics, i.e., the frequency with which a scene must be viewed and the physics of the imaging tube used.

4.7.3.1 Data Bit Rate From Multiple Sensor Systems. Of prime importance in the readout of the on-board sensor system is the relatively high data bit rate generated when delta modulation is employed. To derive this bit rate, it is necessary to first define the required analog video BW. This is done as follows:

a. Recognize that the sensor output during readout is an analog video signal.

b. For multiple sensors, define the R/O method, i.e., parallel or serial.

c. Define the number of data elements, a function of tube face format.

d. Define the minimum R/O time, determined by operational considerations and sensor physical characteristics.

e. Apply the Nyquist sampling theorem to determine the required bandwidth.

f. Transform the BW into data bit rate by multiplying by the quantizing code and sampling rate.

To derive an expression for data bit rate consider the parallel readout of N identical vidicons having P picture elements each, in T seconds. The resulting analog video will be applied to a multiplexing n-bit delta modulator, sampling at S samples per cycle. The data bit rate is then found to be:

$$Q = \text{data bit rate} = \frac{Npns}{2T}$$

Suppose we apply the foregoing to the readout of 8 type 74137A RCA vidicons having a 5000 x 5000 TV line format. Readout time is 3.63 seconds. The analog data will be applied to a 2-bit delta modulator having a sampling rate of 2.8 samples per cycle. Thus,

$$Q = \frac{(8) \times (5000 \times 5000)\ (2)\ (2.8)}{(2)\ \ (3.63)} = 154 \text{ MBPS}$$

Clearly, serial R/O would result in a Q of 19.25 MBPS, so that, depending upon system data R/O requirements, the transmission bit rate can vary from 19.25 to 154 MBPS.

4.7.4 Multiplexed Delta Modulator Implementation

The multiple video signals generated by the various sensors are encoded by delta modulators and digitally multiplexed into dual serial binary digital data streams. The implementation of this equipment is described here.

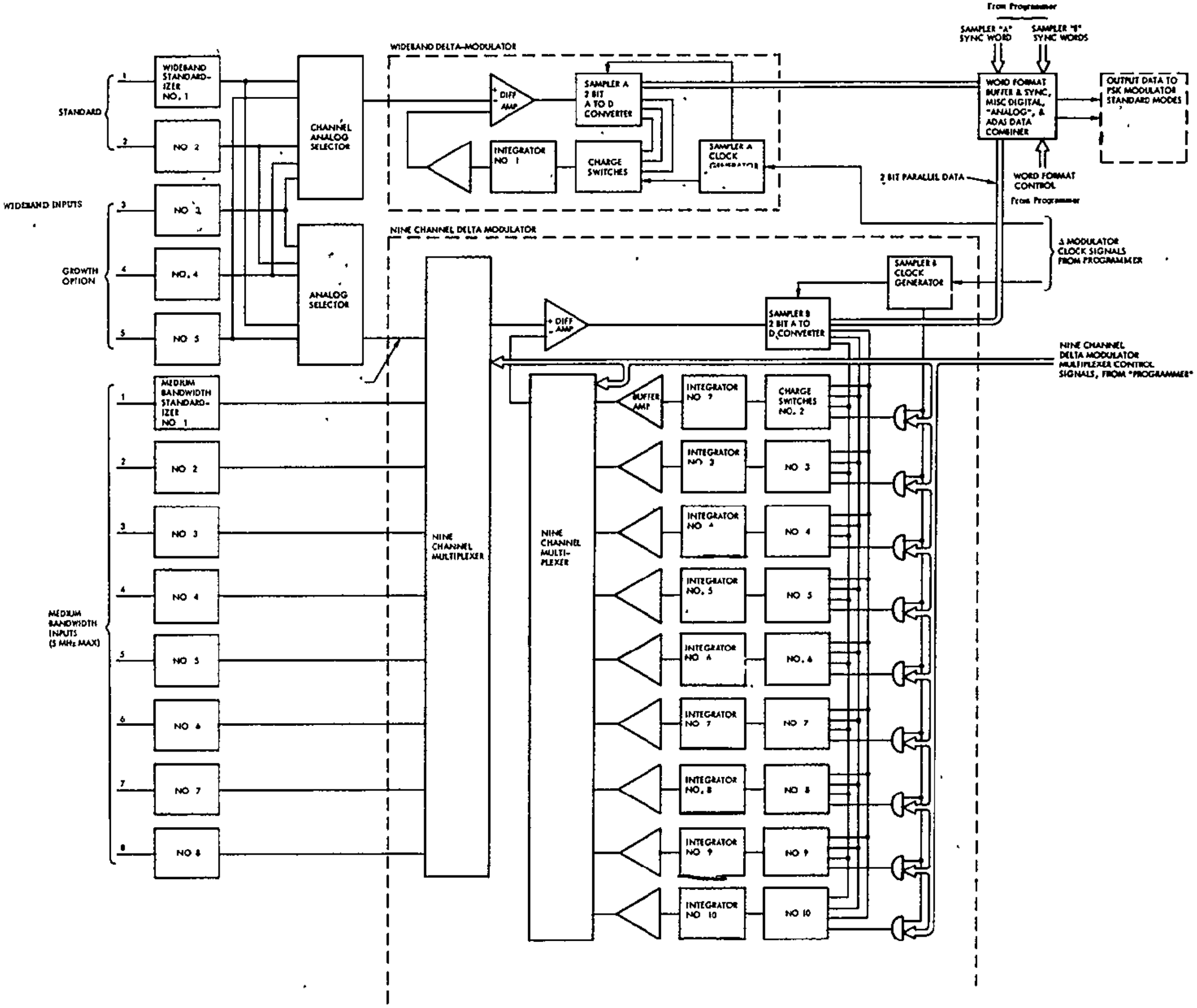

WIDEBAND INPUTS
STANDARD
GROWTH OPTION
1
2
3
4
5
WIDEBAND STANDARDIZER NO. 1
NO 2
NO 3
NO. 4
NO 5
CHANNEL ANALOG SELECTOR
ANALOG SELECTOR
WIDEBAND DELTA-MODULATOR
DIFF AMP
SAMPLER A 2 BIT A TO D CONVERTER
INTEGRATOR NO 1
CHARGE SWITCHES
SAMPLER A CLOCK GENERATOR
From Programmer
SAMPLER "A" SYNC WORD
SAMPLER "B" SYNC WORDS
WORD FORMAT BUFFER & SYNC, MISC DIGITAL, "ANALOG", & ADAS DATA COMBINER
OUTPUT DATA TO PSK MODULATOR STANDARD MODES
2 BIT PARALLEL DATA
WORD FORMAT CONTROL
From Programmer
Δ MODULATOR CLOCK SIGNALS FROM PROGRAMMER
NINE CHANNEL DELTA MODULATOR
SAMPLER B CLOCK GENERATOR
DIFF AMP
SAMPLER B 2 BIT A TO D CONVERTER
NINE CHANNEL DELTA MODULATOR MULTIPLEXER CONTROL SIGNALS, FROM "PROGRAMMER"
MEDIUM BANDWIDTH INPUTS (5 MHz MAX)
1
2
3
4
5
6
7
8
MEDIUM BANDWIDTH STANDARDIZER NO 1
NO 2
NO 3
NO 4
NO 5
NO 6
NO 7
NO 8
NINE CHANNEL MULTIPLEXER
NINE CHANNEL MULTIPLEXER
BUFFER AMP
INTEGRATOR NO 2
INTEGRATOR NO 3
INTEGRATOR NO 4
INTEGRATOR NO. 5
INTEGRATOR NO 6
INTEGRATOR NO 7
INTEGRATOR NO. 8
INTEGRATOR NO 9
INTEGRATOR NO 10
CHARGE SWITCHES NO. 2
NO 3
NO 4
NO 5
NO. 6
NO 7
NO 8
NO 9
NO 10

Figure 4-16 Digital Data Processor Delta Modulator

In Figure 4-16 an illustrative multiplexed delta modulator has ten video inputs provided and utilized in selected patterns as determined by the Programmer mode control. Two of these inputs are for wideband (i.e., up to 15 MHz) video and eight are for medium bandwidth (i.e., up to 5 MHz). A single bandwidth channel (i.e., up to 40 MHz) is provided at the maximum data rate. A growth option of three more wideband inputs is recommended for future sensor options in conjunction with other growth potential to be described.

To handle simultaneously the multiple video inputs (up to 9 in number), the delta modulator is split into two parallel sections. The Sampler A section encodes only one channel at up to 216 Mbps. The Sampler B section handles 1 to 9 video signals sequentially by multiplexing the nine inputs and the D/A feedback to and from the nine integrators. Control of this multiplexing is mode-determined and formatted in the programmer. The split delta modulator approach is necessitated because the internal loop feedback cycle would have been impractically short for a mass production design converter within current art at the required rates with the added handicap of the multiplexing switching delay. However, the use of two independent converters permits a wide variety of sensor channels to be flexibly accommodated.

4.7.5 Multiplexed Delta Demodulators

A multiple-channel delta demodulator for up to 10 channels is shown in Figure 4-17. The digital-to-analog converter receives the component data streams and a digital clock. It also receives digital control lines from the programmer part of the surface DDP and utilizes this control to demultiplex the nine channels. The deformatted and demultiplexed data is directed to integrator channels.

The ten integrators are programmed individually with time constants which are the same as the corresponding integrators in the satellite delta modulator. The ten video outputs are amplified and presented to ten video coaxial output connectors.

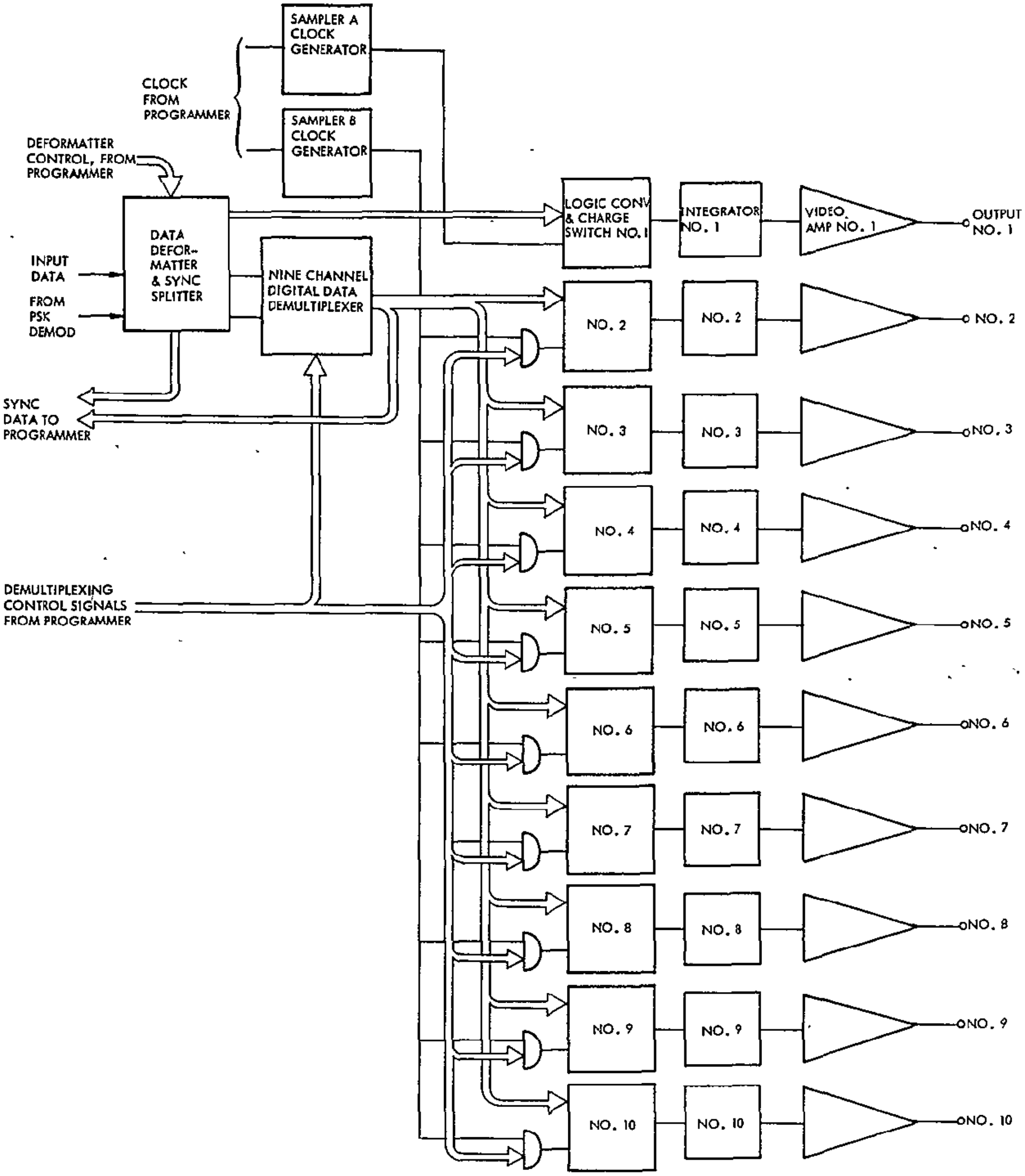

SAMPLER A CLOCK GENERATOR
SAMPLER B CLOCK GENERATOR
CLOCK FROM PROGRAMMER
DEFORMATTER CONTROL, FROM PROGRAMMER
DATA DEFOR-MATTER & SYNC SPLITTER
INPUT DATA FROM PSK DEMOD
NINE CHANNEL DIGITAL DATA DEMULTIPLEXER
SYNC DATA TO PROGRAMMER
DEMULTIPLEXING CONTROL SIGNALS FROM PROGRAMMER
LOGIC CONV & CHARGE SWITCH NO. 1
INTEGRATOR NO. 1
VIDEO AMP NO. 1
OUTPUT NO. 1
NO. 2
NO. 3
NO. 4
NO. 5
NO. 6
NO. 7
NO. 8
NO. 9
NO. 10

Figure 4-17 Digital Data Processor
Multiple Channel Delta Demodulator

4.8 IMAGE QUALITY IN DIGITAL COMMUNICATIONS SYSTEMS

The image quality of the proposed digital data link can be comparable to that of an analog link. Experimental results and data from Philco-Ford operational programs such as Project Compass Link, have indicated that output imagery is undistinguishable from imagery generated with typical sensor-recorder combinations connected in the back-to-back analog mode provided that the channel sensitivity and received energy per bit ratio are proper.

The noise characteristics (quantizing errors) of the analog-to-digital converter are sufficiently low such that more than 18 shades of gray can be discernible on the output imagery if the sensor and recording film are capable of this level of performance.

The amplitude linearity should be better than 5%. Any deviation from an ideal linear approximation will then be a smooth monotonic function that will be undetectable in the output imagery. The overall system baseband signal phase shift is determined by the digital-to-analog converter and the demodulator. Linear phase characteristics are critical to video data and the net phase characteristics are adjusted so that transient response overshoot will be less than 5%, and that no detectable image smearing results.

Data link errors above a given bit error rate will, in general, cause undetectable signature in the imagery. Certain multiple errors will cause visible signature that will have the appearance of a short "comet"-like streak along the direction of scan that rapidly decreases in intensity, returning to the correct intensity level in approximately 20 to 50 data samples. As a first order approximation, the decay may be characterized by an exponential function. At data link error rates which are better than 10^{-3}, the output imagery is essentially unaffected by data link errors. The additional link margin should then be provided to insure error rates of at least 10^{-5}.

The common, analog-link, multipath effect of "ghosting" and phase error which cause ringing, "white after black" and general image smearing and loss of detail are eliminated from the digital data. The effect of multipath will be an increase in the data link error rate, which will cause additive noise to be visible in the recorded imagery. Smearing, ringing and ghosting will be absent.

In summary, the performance of an imagery digital link will be superior to an analog link in the presence of multipath reception and the imagery recorded at the link terminus can be generally indistinguishable from that of a hard-wired, back-to-back analog link provided that the bit error rate of from 10^{-4} to 10^{-5} is maintained.

4.8.1 Some Measured Results

This paragraph will discuss various results of measurements made by Philco-Ford on practical delta modulators optimized for transmission of image sensor data.

The most important performance characteristic of any imaging transmission system is its ability to reproduce faithfully a useful facsimile of the input data. The overall primary system measures of performance are generally parameters borrowed from photographic technology such as resolution, image noise, gray-scale rendition and fidelity, and geometric linearity. Other important characteristics which are subjectively important, but difficult to quantitatively measure and interrelate with the primary system parameters, are: smearing and ringing caused by non-linear phase response, contouring, ghosting, saturation, sync jitter, and effects caused by the sampling process and resultant aliasing error inherent in any line scan system, such as spurious resolution and Moire' effects. The primary source of difficulty is the fact that the evaluating mechanism is the human eye and brain. The difficulty arises when we attempt to use familiar and convenient one-dimensional parameters such as linearity and signal-to-noise ratio as a basis of comparison between totally different and non-linear systems. At present, no precisely reliable method has yet been devised for predicting the "usefulness" of a recorded image from knowledge of the simple single-dimensional parameters. Fortunately, experimental data, including data verified in operational systems, is available to evaluate the performance of the delta modulator. The final test of any image transmission system is the viewing.

4.8.1.1 Signal-to-Noise Ratios

The signal-to-noise ratio of a delta modulator is primarily a function of large and small step size, integrator characteristics, reference threshold levels, input signal dc component, and the input signal structure itself. The determination of the SNR is most easily accomplished by graphical analyses or actual measurements when hardware exists.

The measured peak-to-peak signal-to-noise ratio for a carefully optimized delta modulator is 36 dB for the uncorrelated quantization noise. The total noise consists of uncorrelated quantization noise, as well as a lower frequency noise that is correlated over intensity values and over a resultant area in the output image.

Correlated noise may cause a delta modulator signature effect if the optimization required to prevent excessive slope-overloading distortion results in detectable noise. The effect is usually not visible in a properly optimized system. When the effect is visible, its effect on image usefulness is usually minimal, since the "noise" is correlated over common intensity areas and the eye is very effective in seeing through the signature.

A typical value of SNR (when total noise is considered) for an optimized delta modulator is 31 dB peak-to-peak signal-to-rms noise. It is important to note that this value of SNR which would cause a visible "salt and pepper" effect in analog transmission due to the gaussian nature of the noise is undetectable in the delta modulation system, since the "rms" noise is not random and the peak noise amplitude is fixed and never approaches a shade of gray difference in the dynamic range of the video.

4.8.1.2 Noise and Distortion Optimization

For reconnaissance sensor information which will be displayed as intensity values in a 2-dimensional spacial coordinate system such as pictorial video, SLAR, IR scanners, and laser illuminators, the most important parameters are good transient response or linear phase response, low slope loading distortion, and an output video signal-to-noise ratio sufficiently high so as not to degrade the information quality. For many of these sensors, the limiting system signal-to-noise ratio is set by the ultimate display and recording medium (typically photographic film) rather than the sensor.

Some sensors might have video signal-to-noise ratios of better than 50 dB, while most recording film is limited to 15 to 20 shades of gray. This fact allows this general class of pictorial video transmission systems to operate at signal-to-noise ratios considerably less than acceptable for quality audio transmission. For this type of information, a transmission system with signal-to-noise ratio sufficient to determine 100 discrete levels or shades of gray clearly has an excess of signal-to-noise ratio. The channel capacity may be better utilized by increasing the baseband data rate.

Lower signal-to-noise ratios can also be tolerated when the quantizating noise has the property that the peak noise amplitudes are discretely limited. For example, a digital system whose peak noise amplitude is less than a shade of gray will produce a better result (assuming contouring is not present) than an analog system with the same RMS signal-to-noise ratio, because the peak analog noise amplitudes will often exceed the visual detectability level.

Intermodulation distortion caused by non-linear amplitude response is generally not critical. On the contrary, certain types of amplitude nonlinearities such as gamma correctors or contrast enhancers may improve the output image quality.

For a multiple bit delta modulator, the optimum operation generally occurs when the slope response limitation of the delta modulator just meets or exceeds the maximum slew rate of input data. When this situation occurs, the system signal-to-noise ratio is maximized without losing the capability of resolving fine detail after a rapid gray-scale transition. Care must be maintained not to over-enhance the input data high-frequency content and slew rate artifically by the use of techniques such as aperture correction.

4.8.1.3 Phase Characteristics

A properly optimized delta modulator with balanced and matched integrators has ideal linear phase characteristics for video transmission. Overshoot, ringing and linearity can be balanced to null. Overshoot specifications of a few percent are required, not to allow for non-linear phase effects, but for worst-case, slope-limiting quantization error overshoot. The phase linearity balance can deliberately be set to cause some edge enhancement overshoot which is often a desirable improvement in the output imagery. The integrator balance is routinely held within a fraction of a percent when linear phase is required.

4.8.1.4 Dynamic Range

One of the major advantages of a delta modulation system is that it theoretically can have an infinite dynamic range as contrasted to conventional PCM system or power-limited analog systems. Practical circuit limitations obviously limit the dynamic range of a realizable delta modulator to the saturation voltages of various amplifiers and components. Practical dynamic ranges can easily exceed the specifications and ranges of the image sensors.

The wide dynamic range capability of the delta modulator and resultant freedom from amplitude overloading distortion are obtained at the expense of a slope limitation factor which causes a slope-overloading distortion if the slew rate of the input exceeds the slew rate of the delta modulator.

An optimized delta modulator design adjusts the input data slew rate for optimum noise considerations and then adjusts the slewing capabilities of the delta modulator to just exceed the processed data slew rate.

4.8.1.5 Recovery from Data Link Errors

A delta modulator with a "leaky" integrator has the inherent capability to recover from errors. Properly optimized delta modulators typically recover in an approximately exponential manner to less than 10 percent of the error step amplitude in 20 to 70 data sample periods.

Most single bit errors are just on the threshold of detectability for delta modulators. Multiple bit errors will cause detectable errors where maximum step difference amplitude is usually limited to 1 or 2 shades of gray. Multiple bit error recovery time constants are essentially the same as those for single bit errors.

The use of aperture correction can be used to produce nearly any desired modulation transfer function at a given spacial frequency; however, this is at the expense of signal-to-noise ratio. If an excessive amount of aperture correction is used, resulting in a poor high frequency signal-to-noise ratio, slope-loading distortion will result unless the slew rate of the delta modulator is increased to meet the input data slew rate. An increase in the delta modulator slew rate always trades off an increase in quantization noise as a first-order approximation. The actual noise increase may be an overall noise increase or an increase in the loop instability, which leads to excessive delta modulator signature. The optimum method chosen to increase the slew rate depends upon input data characteristics.

Unfortunately, while aperture correction decreases primarily the high frequency signal-to-noise ratio, an increase in delta modulator slew rate decreases the signal-to-noise ratio across the baseband spectrum, resulting in an overall poorer system transfer characteristic when excessive aperture correction is used. An optimum system in which a delta modulator is incorporated may place a significant amount of the system aperture compensation after the delta demodulator.

4.8.1.6 Test Voltage Waveforms

Test waveforms can be used to show the response of an optimized delta modulator. These waveforms include swept frequency square wave, sine wave, and triangular wave inputs.

An important consideration of the apparent MTF which limits the response at high frequencies is that the limitation is not a linear modulation transfer function, but rather a non-linear slope-loading limitation. The effect of the limitation is reduced or eliminated as the amplitude is reduced. This is also the case for real video information which can be characterized by an integrated spectrum and exponential probability density function due to finite system apertures. Thus, the true value of the modulation transfer function is unity for all frequencies of interest, as long as slope loading does not occur. An optimized delta modulator does not slope-limit or reduce the effect to a minimum.

4.9 COMPARISON OF MULTIPLEXING AND MODULATION TECHNIQUES

Various multiplexing/modulation techniques have been considered and evaluated for multiple sensor transmissions:

a. SSSC/FDM/FM*
b. DSB/FDM/FM
c. VSB/FDM/FM
d. TDM/PAM/FM
e. TDM/PCM/PSK
f. TDM/DM/PSK

---

* SSSC - Single side band suppressed carrier, DSB - Double side band; VSB - Vestigial side band; TDM - Time Division Multiplex; FDM - Frequency division multiplex; PAM - Pulse amplitude modulation; PCM - Pulse code modulation; DM - Delta modulation

Each of these will be evaluated to provide a comparison of the performance of each technique.

4.9.1 SSSC/FDM/FM

This technique could develop a single side band suppressed carrier subcarrier representation of each sensor channel, then frequency multiplex all sensor channel subcarriers and then use the composite signal to frequency modulate* the transmitted carrier.

Considering the SSC/FDM/FM multiplexing technique in general there are a number of other difficulties which are apparent.

One of these is the extreme difficulty of obtaining good low frequency amplitude and phase response to near dc, since the carrier must be suppressed. This can be alleviated by using vestigial or double sideband multiplexing and this will be discussed later. Another difficulty will be the reduction of inter-channel crosstalk to a level below thermal noise.

---

*For FM modulation the signal/noise ratio SNR at the output of the demodulator is given by

$$SNR = \left(\frac{C}{N_o}\right) \left[\frac{3\ (\Delta f)^2}{f(f_2^3 - f_1^3)}\right]$$

where F = peak deviation

f_2 = upper frequency

f_1 = low frequency

and C/N_o = carrier/noise power density at the demodulator input.

It can be shown that for linear channels, i.e., SNR proportional to signal level, the SNR required for channels carrying visible-light data is 59 dB (peak-to-peak to RMS) and that for channels carrying infrared data is 54 dB (peak-to-peak RMS).

The next problem which will be mentioned is intermodulation distortion noise due to linearity and group delay deviations in the IF/RF sections of the downlink. It may not be too difficult to keep this noise to a low level however its contribution must be recognized and minimized.

Implementation of SSSC/FDM necessitates the provision of highly stable frequency sources both in spacecraft and on the ground since frequency differences between these are transferred directly to the video channel spectra.

4.9.2 DSB/FDM/FM and VSB/FDM/FM

Both of these multiplexing techniques offer the advantage (over SSSC/FDM) of improved low frequency response and either could be made to meet the amplitude and phase requirements.

Both require a greater baseband bandwidth than SSSC/FDM and therefore the permissible peak channel deviations and consequently the attainable SNR's must be less than for SSSC/FDM.

4.9.3 TDM/PAM/FM

This technique for multiplexing the multiple sensor channels would require the channels to be sequentially sampled and then to allow these samples to frequency modulate the downlink carrier.

The resulting spectrum will be more than twice that of the SSSC/FDM. It is therefore obvious that the attainable SNR will be very much less than for SSSC/FDM.

4.9.4 TDM/PCM/PSK

As a logical progression from TDM/PAM multiplexing, the merits of digitization by PCM by means of analog/digital conversion becomes apparent.

One of the major benefits of digital transmission is the independence of the channel SNR on the link carrier/noise (for error rates $< 1 \times 10^{-5}$) and the very reduced effect of transmission deviations on channel SNR*. Furthermore, as a consequence of digitizing the analog signals, it is possible to employ non-linear digitization without introducing non-linear distortion.

There are other advantages to this technique also. Firstly, the low frequency response of each channel extends down to dc thus fulfilling another important requirement. Secondly, the channel SNR is essentially independent of C/N and thirdly, PCM/QPSK is less adversely affected by IF/RF linearity and group delay deviations than analog FM.

4.9.5 TDM/DM/PSK

The merits of TDM/DM/PSK over TDM/PCM/PSK are derived from the fact that for most image data transfer applications, 2-bit and 3-bit delta modulators (DM) provide substantially the same accuracy as 6-bit PCM while requiring much less bandwidth (see Figure 4-3). Also circuit complexity and power requirements for delta modulation are significantly less than that required for 6-bit PCM.

However, the disadvantages of DM over PCM are the loss of the d.c. component of the image, and the absence of the full digital equivalent of any particular element of the image - as is provided by PCM which provides a bit word which represents each sample.

---

* In PCM, the SNR due to quantization noise only is

$$\text{SNR} = \frac{3}{2} \times 2^{2N} \left(\frac{\text{RMS}}{\text{RMS}}\right)$$

where N = No. of bits

and, for 5 bit PCM; for example, which represents an extremely favorable case,

$$\text{SNR} = \frac{3}{2} \times 2^{10} = 32 \text{ dB} \left(\frac{\text{RMS}}{\text{RMS}}\right)$$

$$= 41 \text{ dB} \left(\frac{\text{P-P}}{\text{RMS}}\right)$$

4.9.6 Recommended Technique for Multiplexing

Reviewing the foregoing analysis it is apparent that only three techniques seem worthy of further consideration. These are SSSC/FDM/FM, TDM/PCM/QPSK and TDM/DM/QPSK. Each, at the present time, has problems associated with its implementation. SSSC/FDM/FM will meet worst SNR requirements, unless it can be demonstrated that analog companding is feasible and will provide the necessary improvement, and the low frequency response and inter-channel crosstalk requirements will present a severe problem to the FDM designer. On the other hand, TDM/PCM/QPSK can provide acceptable performance but will require A/D converters operating at very high bit rates. Here the high bit rate capability coupled with circuit simplicity of TDM/DM/QPSK is definitely an advantage.

Section 5

MODULATOR/DEMODULATORS FOR PSK DIGITAL DATA SYSTEMS

5.1 PHASE MODULATION TECHNIQUES FOR HIGH SPEED DIGITAL DATA

For digital bit stream data rates up to 5 Mbps, many modulation techniques are available, to provide both the modulated RF carrier which represents the bit stream and also demodulate the bit stream from the modulated carrier.

The data rates of 40 to 400 Mbps which are being addressed in this study are most efficiently served by phase-shift-keying techniques in order to

- Conserve bandwidth
- Utilize optimum receiver techniques and data regeneration
- Avoid the requirement of transmitting carrier synchronization data.

The features involving conservation of bandwidth were discussed in Section 2 where it was shown for example, that the same bit rates, quadriphase modulation has half the bandwidth of biphase shift keying, and less than a quarter of the bandwidth of MFSK.

The circuit features mentioned above will be illustrated by the actual systems to be described in this section.

5.2 PSK Systems at 40-400 Mbps Data Rates

When the data rates exceed 40 megabits in any type of modulation system, certain types of problems must be specifically solved on assure minimizing bit error rate.

- The bit waveform in which is rectangular at low bit rates, may degrade to trapezoidal or sinusoidal waveforms as the bit rate increases to, in particular, greater than 100 Mbps.

- Unbalance in the phases of the phase-shift-keyed components can occur.

- Transmission at a very low energy per bit makes the system receiver particularly susceptable to noise and phase jitter particularly in the areas of carrier reconstruction and bit synchronization at the receiver.

- Switching waveforms and transients must be accounted for to assure bit integrity.

- The transmission system will appear as a filter with determinable filter characteristics. Due to the speed of the bit stream modulation, system ranging and intersymbol interference can occur which can degrade transmission.

These and other characteristics will be discussed in the paragraphs to follow.

5.2.1 BPSK and QPSK MOD/DEMOD System

The BPSK and QPSK modulators are now capable of biphase operation at data rates of 100 Mbps to 400 Mbps, respectively. With improvements now occurring in the switching times of the modulator driver, it is expected that rates up to 400 Mbps will soon be possible with BPSK.

Figure 5-1 shows the basic system using PSK techniques to transmit digitized sensor data. Note that timing information is also developed by the demodulator.

Quadriphase QPSK is recommended rather than biphase for rates above 200 Mbps, because quadriphase requires half the RF bandwidth as shown in Fig. 5-2 and will tolerate twice the switching time as biphase for the same data bit rate with no increase in transmitter power.

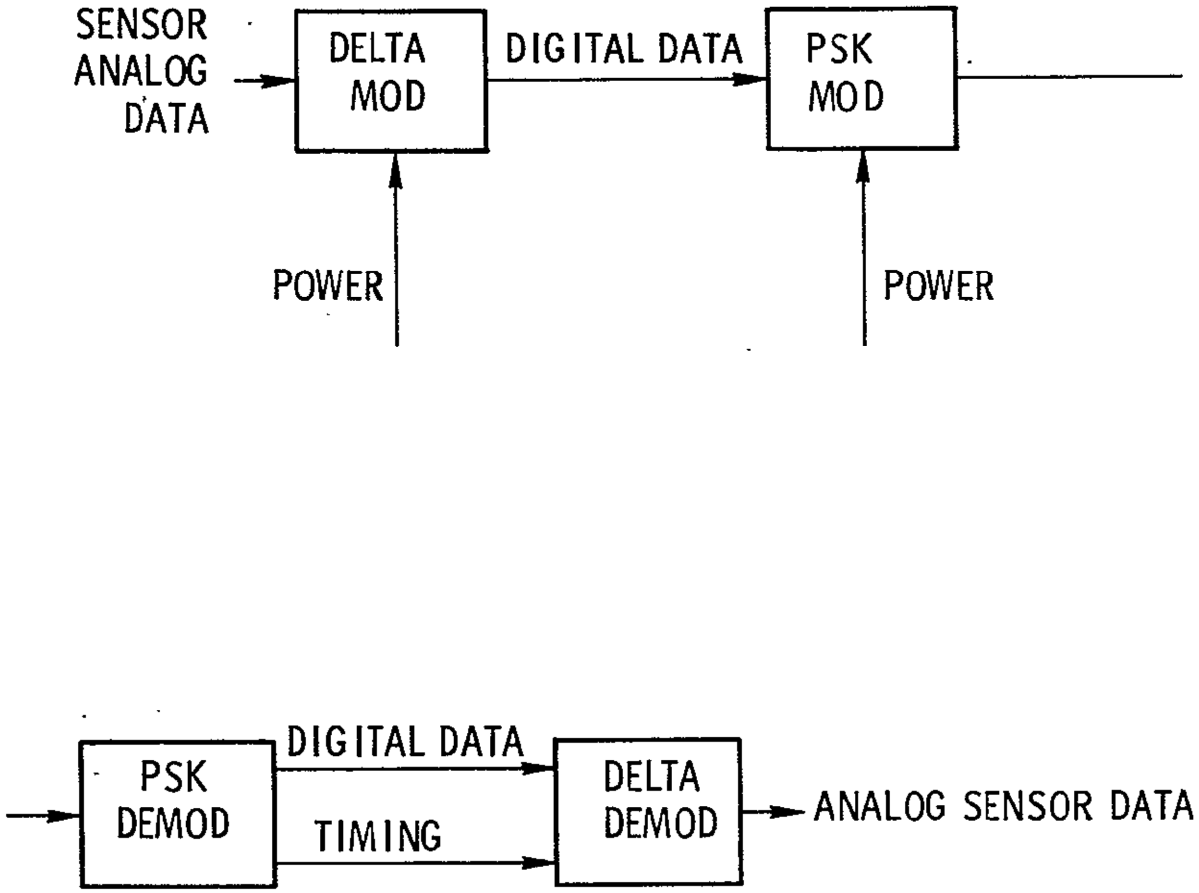

SENSOR ANALOG DATA
DELTA MOD
DIGITAL DATA
PSK MOD
POWER
POWER
PSK DEMOD
DIGITAL DATA
TIMING
DELTA DEMOD
ANALOG SENSOR DATA

Figure 5-1 Basic Modulator-Demodulator System

HIGH DATA RATE PSK MODULATION TECHNIQUES

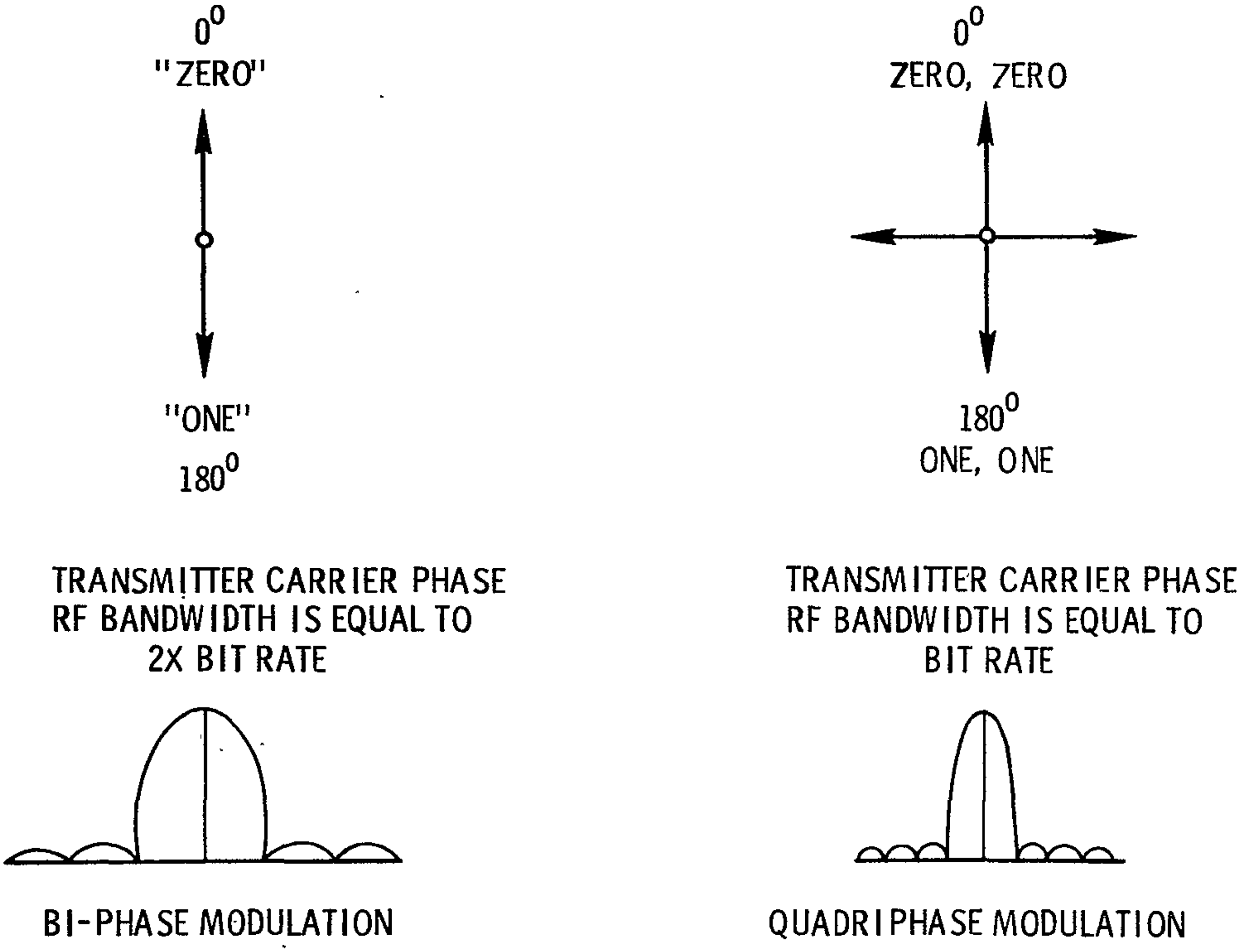

0^0
"ZERO"
"ONE"
180^0
TRANSMITTER CARRIER PHASE
RF BANDWIDTH IS EQUAL TO
2X BIT RATE
BI-PHASE MODULATION
0^0
ZERO, ZERO
180^0
ONE, ONE
TRANSMITTER CARRIER PHASE
RF BANDWIDTH IS EQUAL TO
BIT RATE
QUADRIPHASE MODULATION

Figure 5-2

A quadriphase modulator is essentially two biphase modulators in phase quadrature. The quadriphase demodulator is more complicated than the biphase modulator since it requires maintaining a proper phase balance and phase accuracy between the two biphase modulated carriers, requires that an amplitude balance be maintained, and requires that crosstalk be minimal.

5.3 CARRIER SYNCHRONIZATION IN BIPHASE AND QUADRIPHASE SIGNALING SYSTEMS

The normal PSK spectrum does not include a component at the carrier frequency. Therefore, special efforts must be made to recover the carrier from the incoming signal. Three principal methods of carrier recovery are listed as follows:

a) Multiplication and filtering. Multiplying the frequency and phase of the received PSK signal by n (the number of phase positions) yields one spectral line at n times the carrier frequency. One may either filter out the spectral line or recover it with a phase lock loop (PLL). The resultant wave form may then be divided by n to yield the original carrier in one of n stable phase positions. The narrower the bandwidth of the filter or the loop bandwidth of the PLL, the smaller the phase jitter of the recovered carrier; however, the recovery time is proportionately longer. This technique for recovering the carrier is normally used in continuous PSK operation where there is no demand for fast recovery times.

b) Phase Lock Loop with Remodulation. In this approach the received PSK signal is modulated inversely by the baseband signal, which is demodulated by using a VCO output signal as a reference. The transmitter side modulation of the carrier is therefore cancelled and the carrier itself is received in one of n stable phase positions at the output of the remodulator and fed back to the input of the VCO. In this case also the phase jitter and recovery time are inversely dependent on the loop bandwidth. This method also yields faster recovery times because of the sawtooth shape of the phase detector characteristic.

c) Phase regenerative PLL. The PSK input is first demodulated by the output of a VCO used as a reference carrier. At the same time, this VCO output is PSK-modulated with the received baseband signal. The phase difference between the input PSK signal and the regenerated PSK signal is then fed back to the VCO. The regenerated PSK signal can also be derived from the input PSK signal by direct PSK regeneration without demodulation. Considering jitter and recovery time, this type of demodulation and carrier recovery has nearly the same characteristics as the PLL with remodulation.

5.4 QUADRIPHASE MODULATORS - CIRCUIT ASPECTS

There are several methods for obtaining phase modulation with switching diodes at microwave frequencies. They are characterized by the microwave components used:

a. Magic T
b. Hybrid (Rat-Race)
c. 3-dB Coupler
d. Circulator

The first three require two balanced diodes and are thus more complicated than the circulator type where only one diode is used.

5.4.1 Circulator-Type Diode Phase Shifters

Two types of diode phase shifters capable of phase modulating a microwave or millimeter wave carrier are shown in Figure 5-3 representing techniques now in use to produce bit rates in the hundreds of megabits and capable of gigahertz modulation rates when such rates are required.

The biphase modulator of Figure 5-3 diode-switches between two different lengths of line at the second port of the circulator while the circuit of Figure 5-3 combines two such biphase modulators in a series mode to provide quadriphase modulation.

In the biphase modulator of Figure 5-3, the CW signal is supplied to Port 1 of the circulator. At Port 2, most of the power is reflected and proceeds to Port 3, which is the modulator output. The reflecting impedance at Port 2 is determined by the modulating diode and its circuitry, and the phase angle of the impedance is a function of the diode bias delivered by the driver. High and equal voltage standing wave ratios (VSWR) are obtained for both phase angles 0° and 180°. This means that the output power at Port 3 is the same for both phase conditions and AM-free, biphase modulation with high

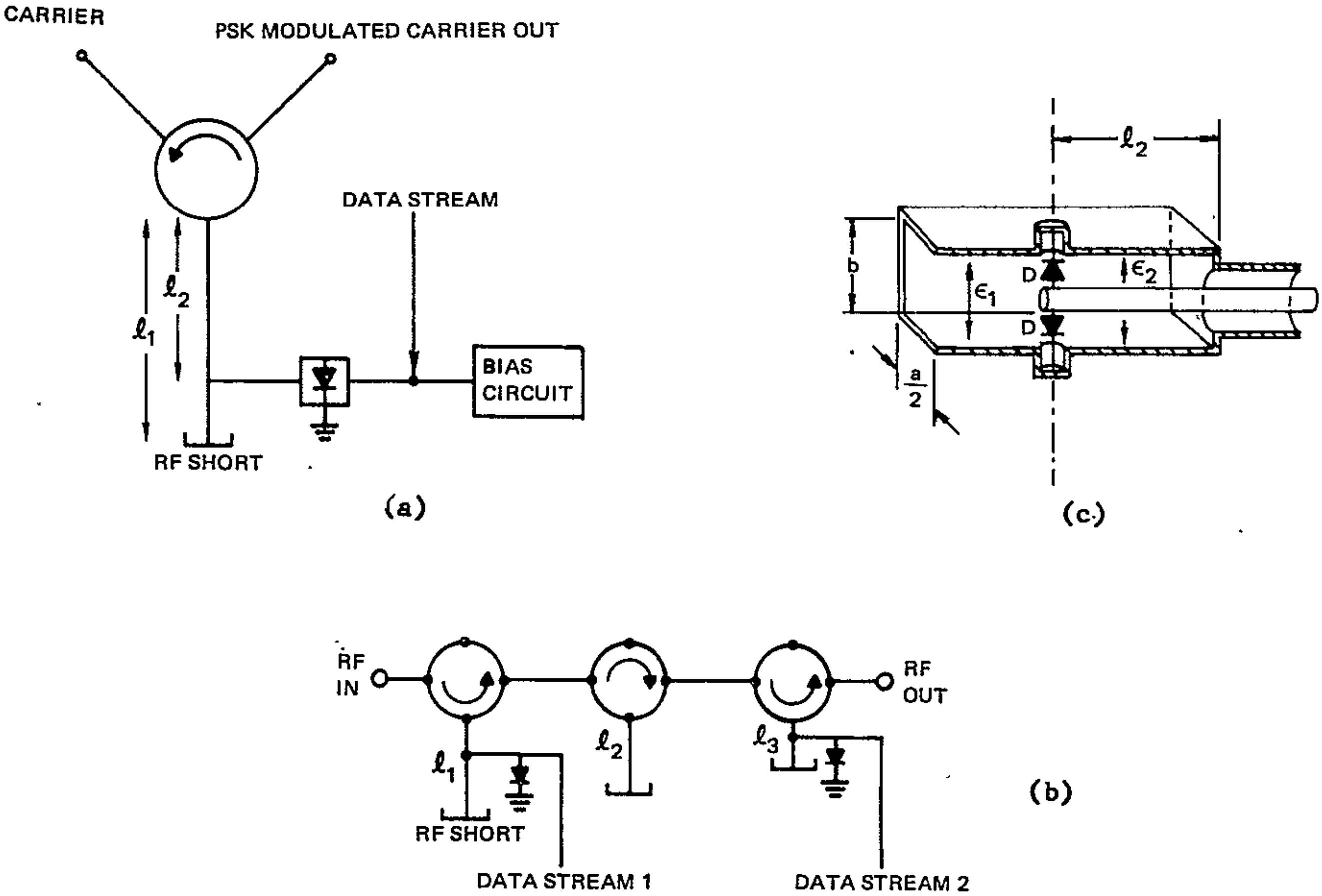

CARRIER
PSK MODULATED CARRIER OUT
DATA STREAM
ℓ_2
ℓ_1
BIAS CIRCUIT
RF SHORT
(a)
ℓ_2
b
ϵ_1
ϵ_2
D
D
$\frac{a}{2}$
(c)
RF IN
RF OUT
ℓ_1
ℓ_2
ℓ_3
RF SHORT
DATA STREAM 1
DATA STREAM 2
(b)

Figure 5-3 Types of PSK Modulators

efficiency is obtained. With rise times at the receiver output of less than 2 nanoseconds, the link can be operated with good efficiency up to about 200 Mbps, which corresponds to 400 Mbps for quadriphase modulation.

The diode acts as a controlled short, and when shorted, reflects the carrier energy, thereby decreasing the line length by one-half wavelength, or 180°. The diode conduction, and hence the reflection property is controlled by a digital input, and the resulting waveform at the circulator output is a biphase modulated carrier.

Since modulation of the carrier is accomplished external to, and independent of the RF power source, the device most important to the direct biphase transmitter is the waveguide switch system. To achieve efficient biphase modulation, the switch system must deliver a true 180° phase difference with a minimally small amplitude unbalance between phase positions.

For an ideal circulator, the input and output voltage ratio, which becomes the reflection coefficient Γ at Port 2 is given as[1]

$$\Gamma = \frac{E_{out}}{E_{in}} = \frac{Y_{oc} - Y_t}{Y_{oc} + Y_t} = |\Gamma| \angle \theta(v)$$

where

$$\theta(v) = -2 \tan^{-1} B(v)/Y_{oc} \quad \text{if} \quad Y_t = jB_1$$

and

Y_{oc} is the circulator characteristic admittance,

$|\Gamma|$ becomes unity when the load admittance Y_t is purely susceptive. Therefore, no amplitude modulation is introduced and the phase shift $\theta(v)$ is an arc tangent function of the susceptive load, which is a function of the modulating signal voltage v.

[1] C. S. Kim, C. W. Lee, and J. R. Baner "Varacter S-Band Direct Phase Modulator", IEEE, Journal & Solid State Circuits, Sept. 1966.

If the load is simply an ideal varactor diode, then

$$Y_t = jB(v) = j\omega C(v)$$

where

$$C(v) = \frac{C_b}{(1 - v/V_b)^n}$$

V_b = dc bias voltage + junction contact potential,

C_b = average capacitance at V_b, and

n = 1/3 and 1/2 for the graded and abrupt junctions, respectively.

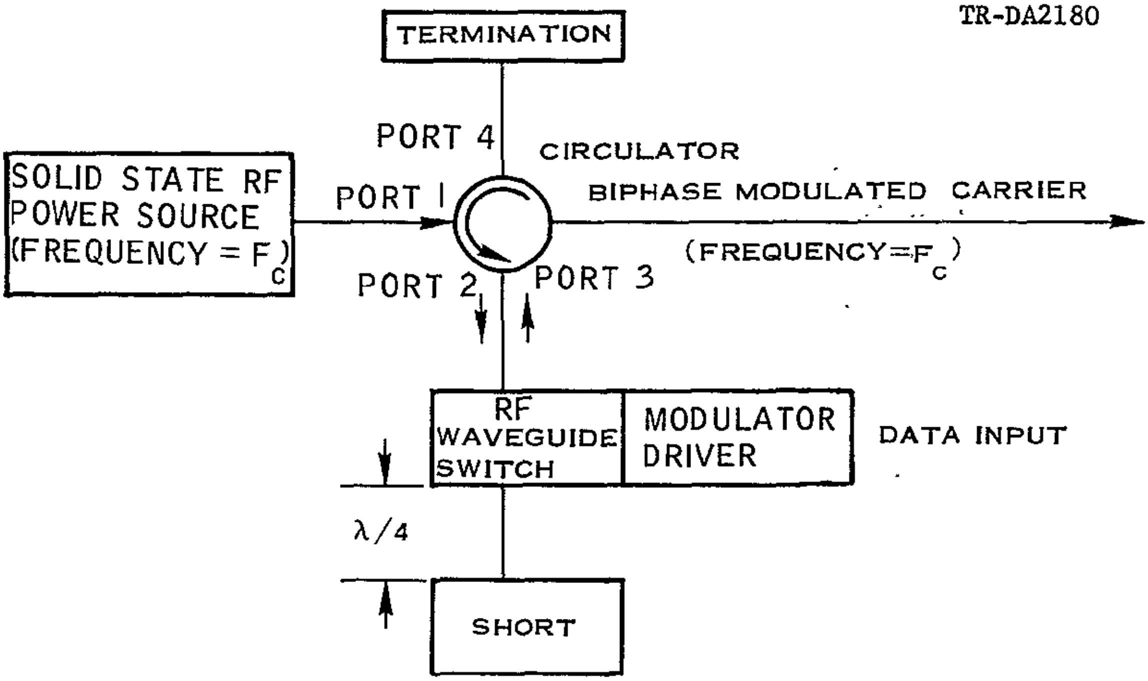
The phase function $\theta(v)$ is thus a complex nonlinear function of v.

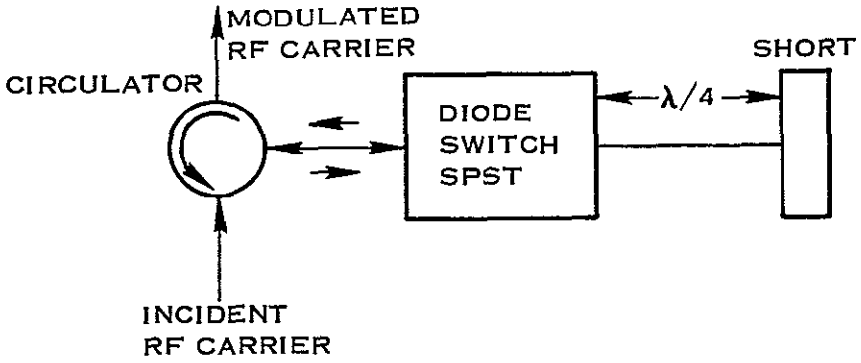
5.4.1.1 Diode Switching Modes

Diode switching for biphase modulation may be implemented in either of the two fundamental circuits of Figure 5-4.

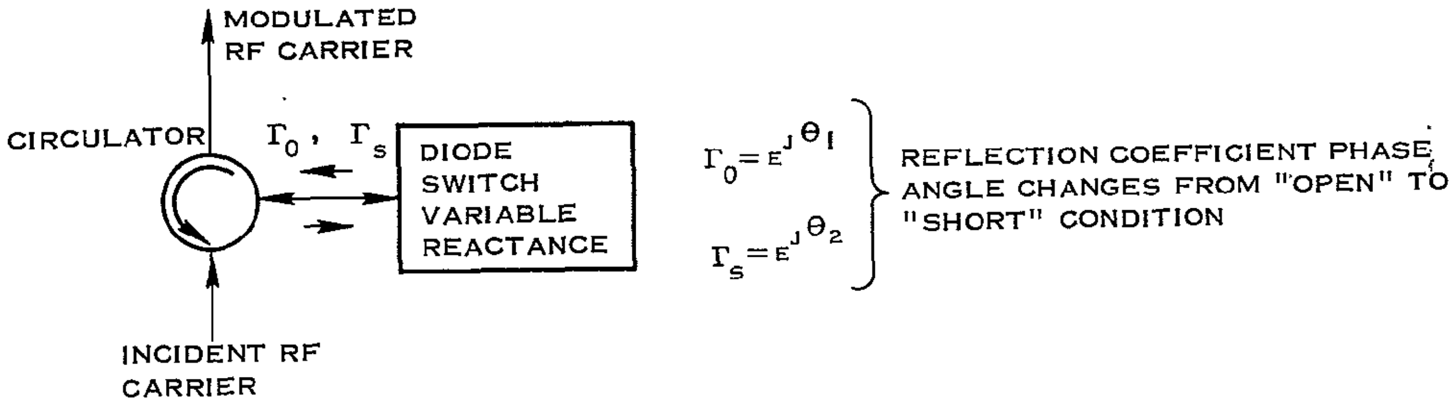
In Circuit A, the diode is utilized as a single-pole-single-throw switch (SPST). In the closed position, incident RF energy is reflected and in the open position, RF energy is transmitted to the short and reflected so that the differential phase shift is 180°. In the design, the phase shift is determined by the length of waveguide and the diode may be optimized for switching speed, power capability, and RF amplitude balance between states. From a device standpoint, common switching diodes do not present sufficient impedance magnitude changes to allow efficient biphase modulation with this approach; however, resonant switching diodes are especially suited to this application, as will be explained.

In Circuit B, the diode is employed as a lossless device which presents a variable reactance when the diode bias is changed from the forward to reverse condition or vice-versa. In this approach, the phase shift is


TERMINATION
PORT 4
CIRCULATOR
SOLID STATE RF POWER SOURCE (FREQUENCY = F$_C$)
PORT 1
BIPHASE MODULATED CARRIER
(FREQUENCY=F$_C$)
PORT 2
PORT 3
RF WAVEGUIDE SWITCH
MODULATOR DRIVER
DATA INPUT
λ/4
SHORT


MODULATED RF CARRIER
CIRCULATOR
DIODE SWITCH SPST
λ/4
SHORT
INCIDENT RF CARRIER

A Biphase Modulator - Single Pole Single Throw Diode Mode


MODULATED RF CARRIER
CIRCULATOR
Γ_0, Γ_s
DIODE SWITCH VARIABLE REACTANCE
$\Gamma_0 = \epsilon^{j\theta_1}$
$\Gamma_s = \epsilon^{j\theta_2}$
REFLECTION COEFFICIENT PHASE ANGLE CHANGES FROM "OPEN" TO "SHORT" CONDITION
INCIDENT RF CARRIER

B Biphase Modulator - Variable Reactance Diode

Figure 5-4

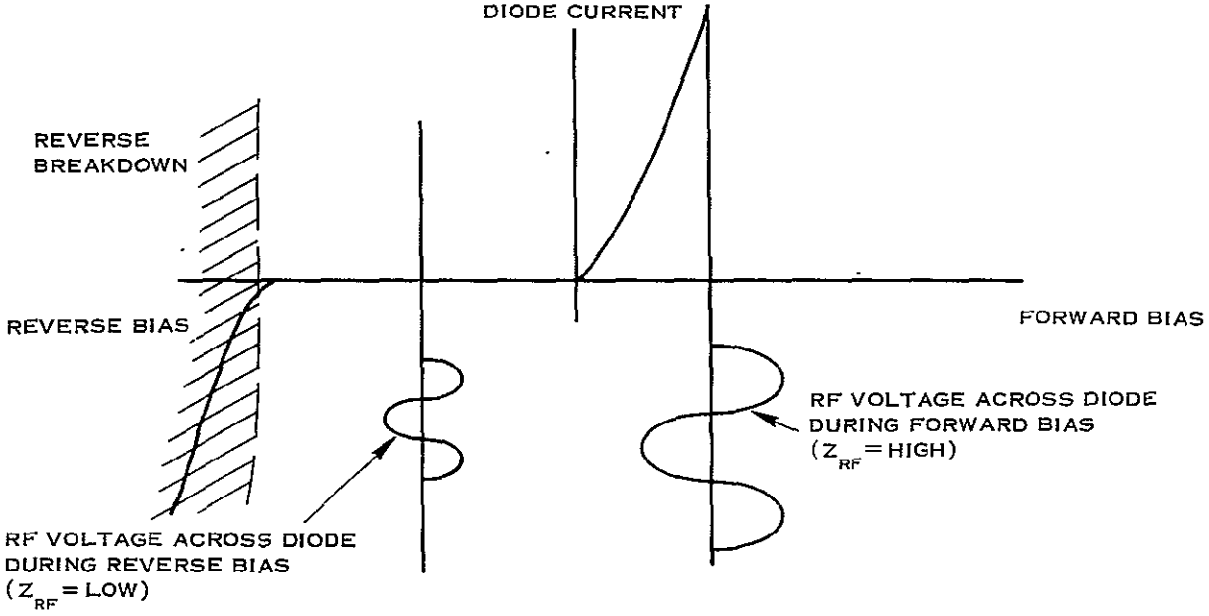
accomplished by a diode design which changes the phase angle of the equivalent reactance of the diode when switched from forward to reverse bias. To achieve effective biphase modulation, the magnitudes of the impedances must be well-controlled so as to minimize amplitude modulation, and in general, external circuit techniques must be employed to provide a reasonable compromise in the device performance.

The SPST diode switch technique has the significant advantage of external phase adjustment which has little effect on amplitude imbalance, power capability, and switching time. In effect, the diode may be optimized for these other properties, a fact which leads to a superior biphase modulator design.

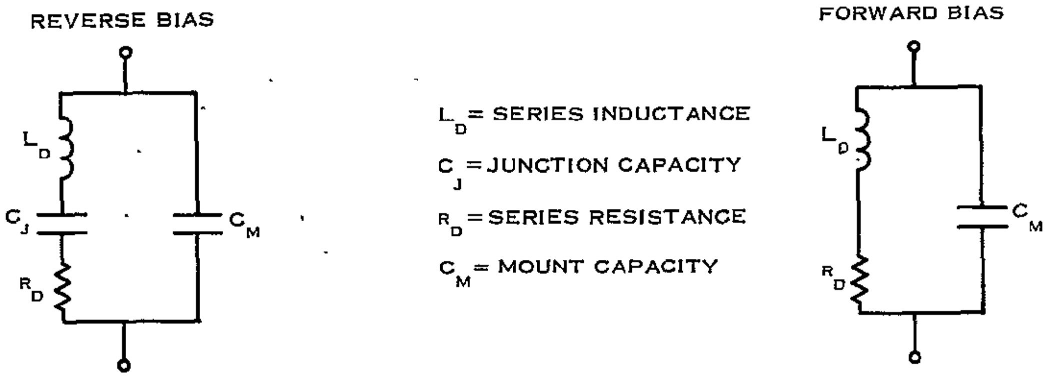
The key to a successful biphase modulator design utilizing the SPST switch is the double resonance switch. Consider the diode waveform and equivalent circuits in Figure 5-5. When reverse-biased (bias = V_r) the diode equivalent circuit consists of the series elements L_d (series inductance), C_j (junction capacity), and r_d (series resistance) in parallel with the mount and package capacity C_m. Forward bias (V_f) has the effect of increasing C_j so that C_j has a very small reactance.

In most switch diode packages, the parallel and series resonances of the switch are quite different in frequency, and for a waveguide mount, the diode is driven into conduction to provide a short to the RF signal and into reverse bias to provide an RF open circuit. In the latter case, the energy propagates through the diode and is reflected from the waveguide short, thus, producing a voltage across the reversed bias diode of twice the incident voltage. With such a configuration, the incident power is severely limited to a value less than one-half the voltage corresponding to the reverse bias voltage.

The double resonance diode is designed so that the series and parallel resonant frequencies coincide. Equivalent circuits and switch actions of a biphase modulator are shown for the switch in Figure 5-6. In



A Switching Diode Curve


REVERSE BIAS
FORWARD BIAS
L_D
C_J
R_D
C_M
L_D = SERIES INDUCTANCE
C_J = JUNCTION CAPACITY
R_D = SERIES RESISTANCE
C_M = MOUNT CAPACITY

B Diode Equivalent Circuits

Figure 5-5 Diode Waveform and Equivalent Circuits

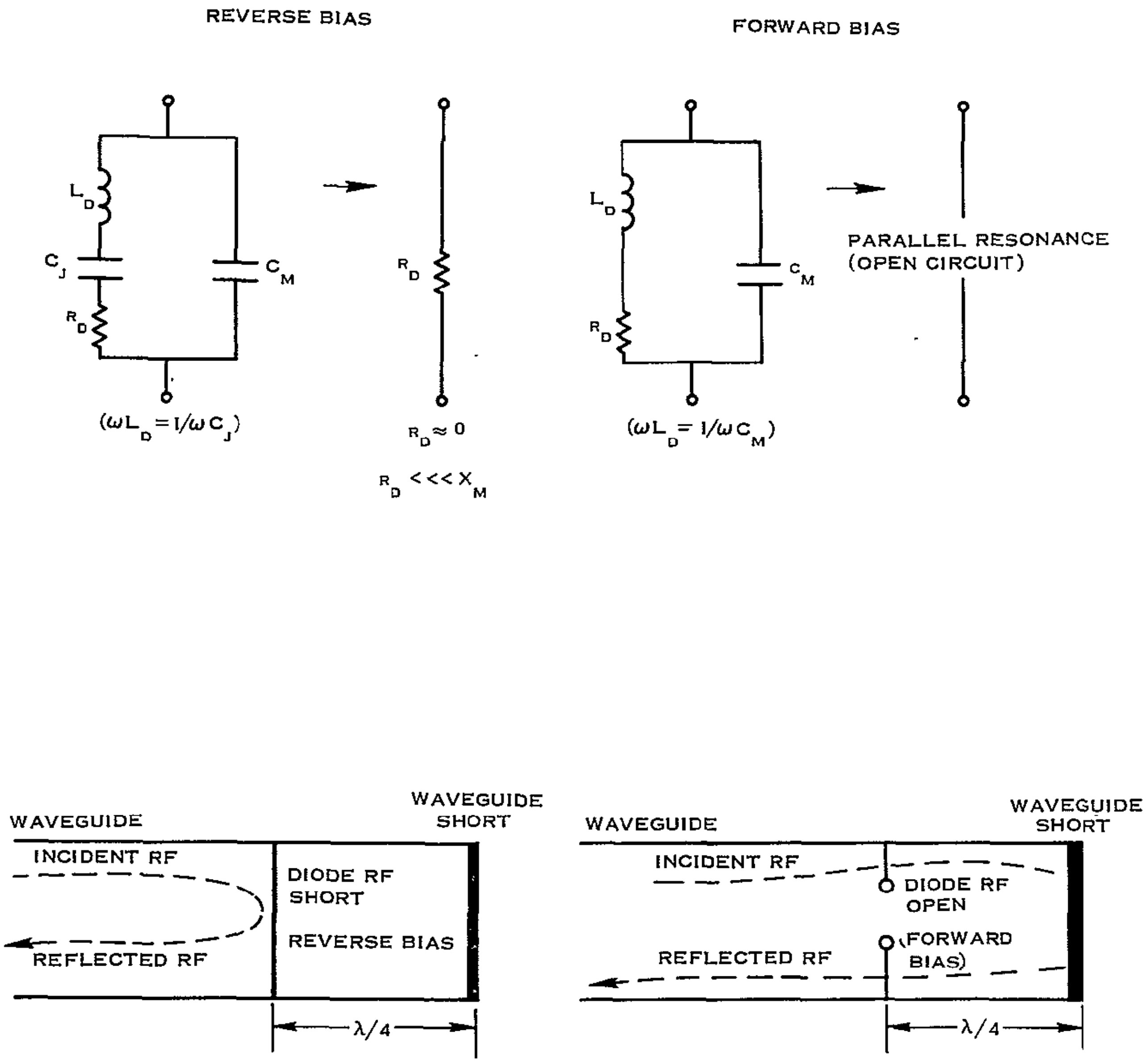

REVERSE BIAS
FORWARD BIAS
L_D
C_J
C_M
R_D
$(\omega L_D = 1/\omega C_J)$
R_D
$R_D \approx 0$
$R_D <<< X_M$
L_D
C_M
R_D
$(\omega L_D = 1/\omega C_M)$
PARALLEL RESONANCE
(OPEN CIRCUIT)
WAVEGUIDE
WAVEGUIDE SHORT
INCIDENT RF
DIODE RF SHORT
REVERSE BIAS
REFLECTED RF
$\lambda/4$
WAVEGUIDE
WAVEGUIDE SHORT
INCIDENT RF
DIODE RF OPEN
(FORWARD BIAS)
REFLECTED RF
$\lambda/4$

Figure 5-6 Philco-Ford Double Resonance Switch Equivalent Circuits and Biphase Modulator Application

reverse bias, the series resonance produces an RF short at the diode plane in the waveguide, while forward bias produces an open circuit, and the RF energy is transmitted to the waveguide short, resulting in a phase difference of 180° relative to the reversed-bias condition. For this type of operation the diode in forward bias passes the RF energy and the reflected voltage (twice the incident value) appears across the diode in the forward bias. The incident power limit is then set by the diode dissipation and currently diode switches operate at levels in excess of two watts up to 26 GHz. Higher levels may be reached with multiple diodes.

Risetimes are currently measured in the range of 1.5 to 2 nanoseconds for the transition from 0° to 180° which corresponds to a data bit risetime of 0.75 to 1.0 nanoseconds.

5.4.1.2 Diode RF Admittance

The input admittance of the diode/RF circuitry can be measured with the slotted line/VSWR meter as function of diode bias and plotted on a Smith chart for both silicon epitaxial and gallium arsenide diodes as illustrated on Figure 5-7 and Figure 5-8. It can be seen that the VSWR is equally high (10 or more) for backward ("0") and forward ("1") diode bias, and that the admittances are separated by 180°. This indicates that AM-free biphase modulation would result in normal operation.

However during the bias transition (risetime of the modulating squarewave) the admittance behavior is quite different. The epitaxial silicon diode is nearly resonant at the operating frequency for both bias conditions; series resonant for backward and parallel resonant for forward bias, and the admittance follows a nearly straight line that passes close to the point 1.0 (matched). The non-resonant gallium arsenide diode admittance closely follows a circle.

Due to these two admittance plots, the modulator output waveforms for phase and RF amplitude with a modulating waveform have a risetime which can be predicted as shown somewhat idealized on Figure 5-7 and on Figure 5-8. As mentioned above, the admittance of the resonant diode modulator, at intermediate bias values, will pass near the matched point so the reflected (output) voltage will have a deep dip with a duration approximately equal to the rise time of the modulating voltage. The output phase on the other hand shifts rapidly near the point 1.0. The gallium arsenide diode modulator has an almost constant VSWR during the transition, so the output voltage stays nearly constant, but the output phase follows the modulating voltage and will have approximately the rise time of this voltage.

5.4.1.3 PSK Switching Waveforms At High Bit Rates

As the bit rate increases, the rectangularity of the Bpsk waveform degrades resulting in a transmission loss which will be discussed here. This is a principal problem encountered for data rates above 40 Mbps, and results from bandwidth limitation in the number of Fourier components making up the waveform.

Consider, for example, the spectrum of a 50 MHz 2 nanosecond rise time modulating voltage from a solid state driver, as shown on Figure 5-9. The curve of Figure 5-10 uses only 5 sets of sidebands and illustrates the ripple and increase of the rise time occuring at these rates.

Filter action on the modulator rise time can introduce an effective energy loss by four mechanisms as analyzed by D. T. Magill. First, filtering FM signals cause an AM envelope to occur. Second, the filtering distorts the rectangular pulse shape such that the ground receiver detector is not properly matched to the now trapezoidal (approximately) pulse. Thus, a mis-matched filter detection loss occurs. Third the creation of an envelope (prior to the TWT) causes signal distortion and reduced signal detectability through the AM-PM conversion mechanism. The fourth and more mundane filter degradation is insertion loss.

SILICON EPITAXIAL, RESONANT DIODE
ADMITTANCE

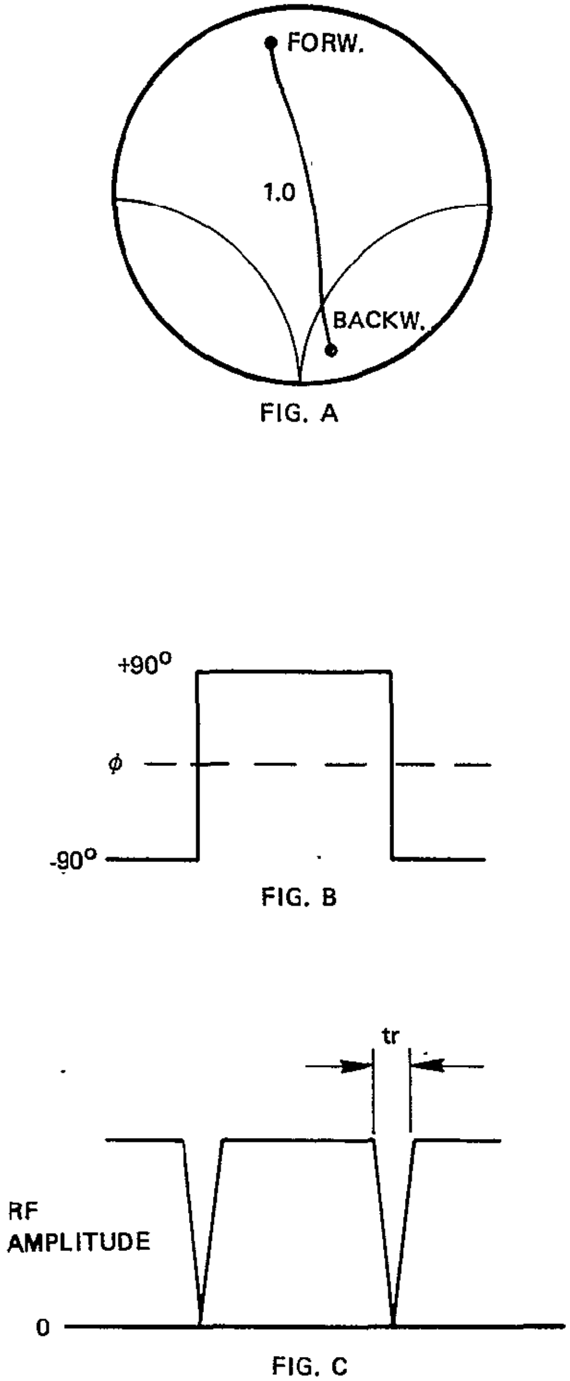

FORW.
1.0
BACKW.
FIG. A
+90°
ϕ
-90°
FIG. B
tr
RF
AMPLITUDE
0
FIG. C

Figure 5-7

GALLIUM ARSENIDE DIODE
ADMITTANCE

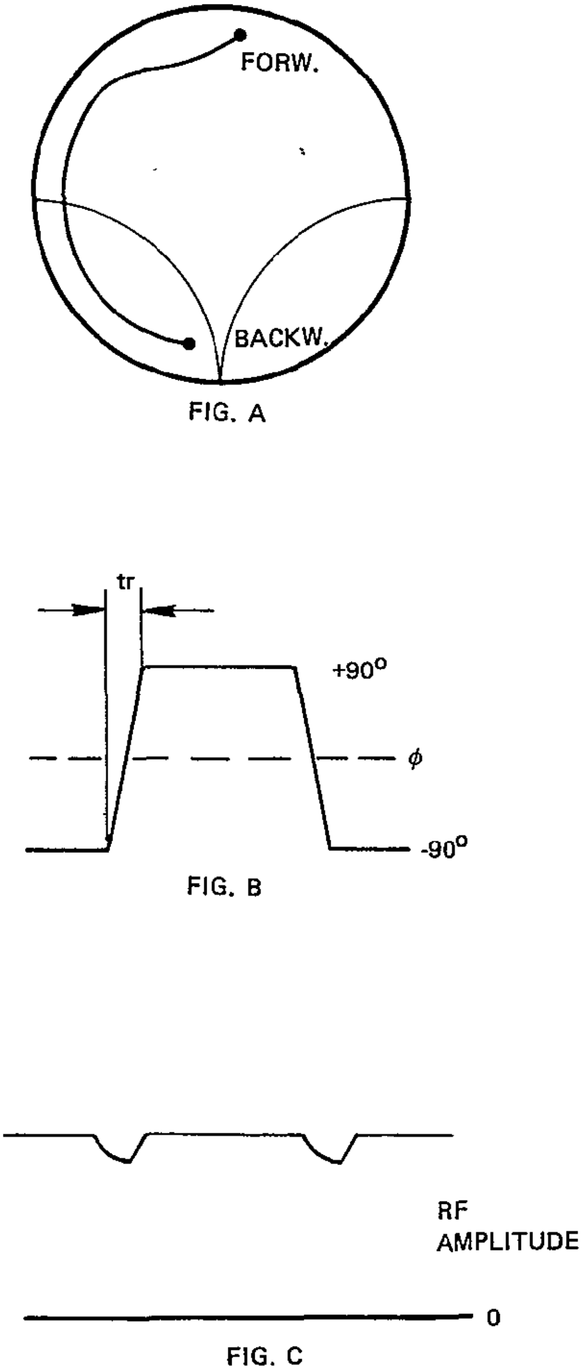

FORW.
BACKW.
FIG. A
tr
+90°
ϕ
-90°
FIG. B
RF
AMPLITUDE
0
FIG. C

Figure 5-8

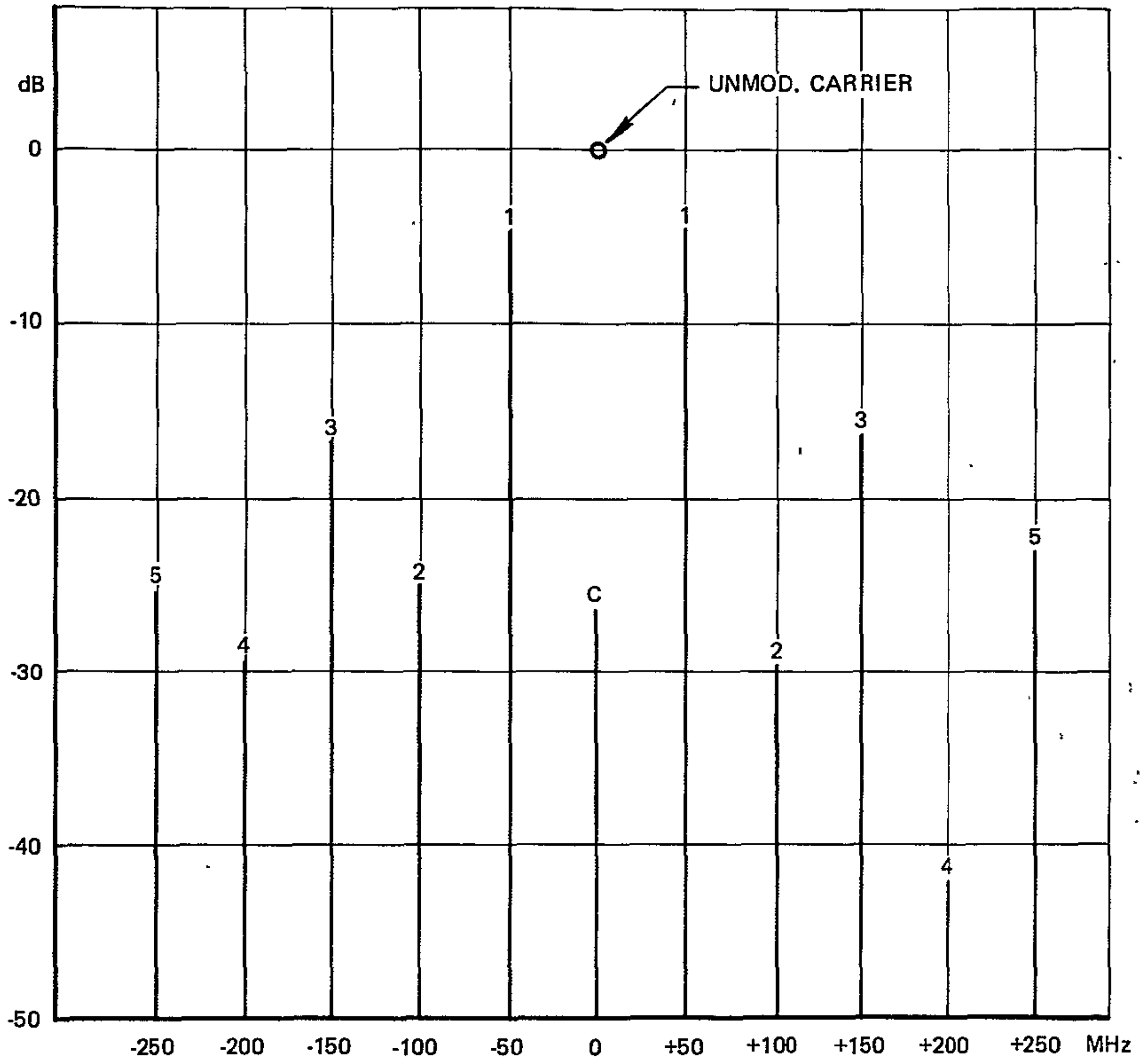


Figure 5-9 Spectrum of 100 Mbps Biphase Signal

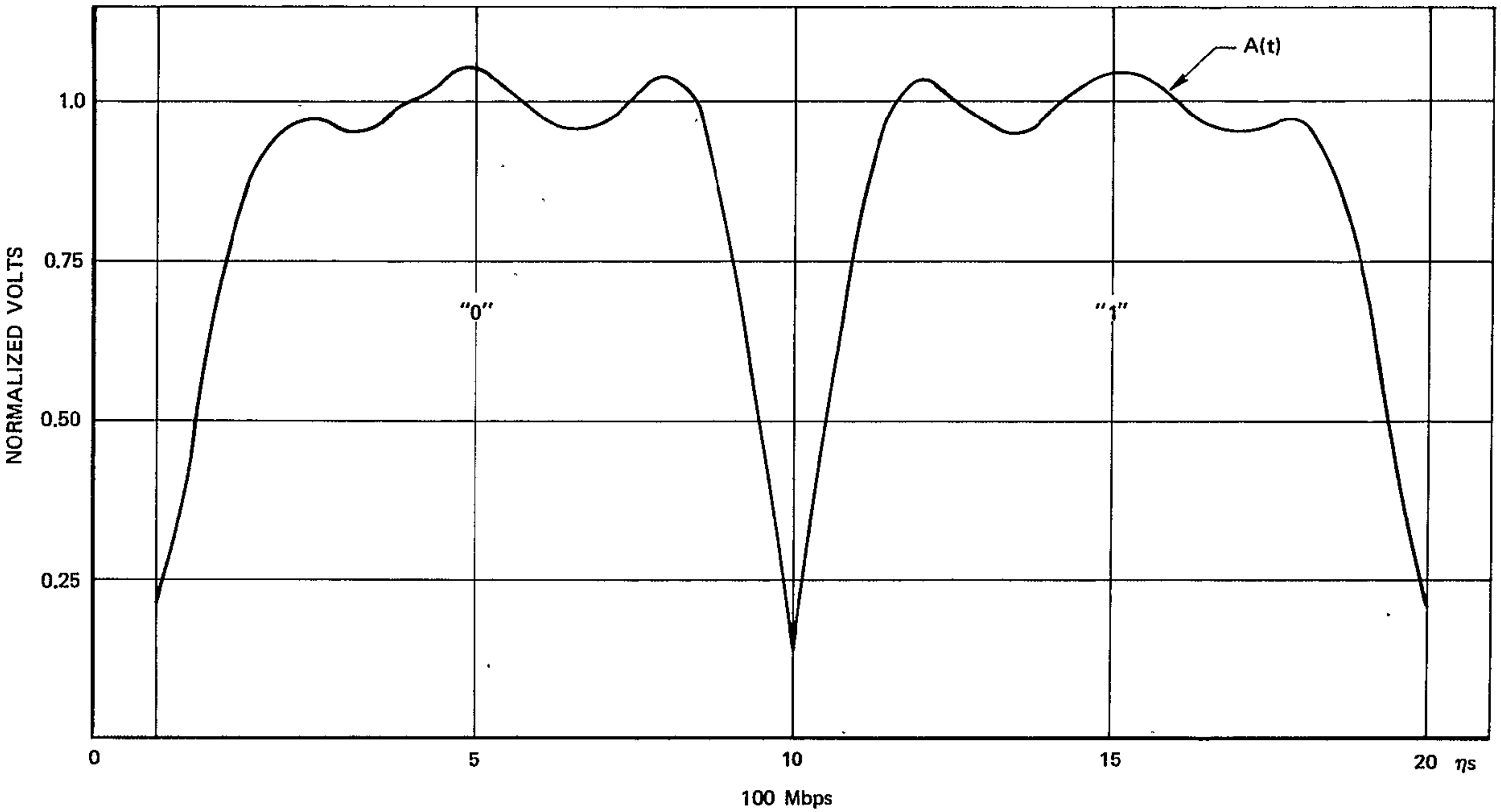


Figure 5-10 RF Amplitude at Biphase Modulated Signal

The filter degradation analyses assume a trapezoidal pulse shown in Figure 5-11. This model represents a reasonable compromise between accuracy and analytical tractability.

The existence of an AM envelope implies that the average radiated power is less than the peak radiated power. For the case of a trapezoidal (+1, -1) amplitude pulse, this degradation, known as the envelope loss, is readily calculated to be

$$D_e = 1 - \left[\frac{2}{3}\right]\left[\frac{t_r}{T}\right]$$

where t_r is the rise time (-100% to +100%) of the trapezoidal pulse and T is the pulse duration (100 nanoseconds).

Note that this loss is not as great as when dealing with (1,0). A bipolar pulse has, with respect to envelope, one-half the rise time of a uni-polar pulse.

The above equation is based on a worst-case analysis, i.e., a phase transition occurs each 100 nanoseconds. With a quadri-phase signal, a complete envelope null (180° phase reversal) occurs only one-fourth of the time. Another one-fourth of the time no change and no null occurs. One-half the time a 3-dB dip occurs. If the dip precedes the TWTA, i.e., non-post-TWTA filtering, it is essentially removed by the TWTA saturation characteristic and the average envelope loss is

$$\overline{D_e} = 1 - \left[\frac{1}{6}\right]\left[\frac{t_r}{T}\right]$$

The worst case mis-matched filter correlation loss can be shown to be given by

$$D_{WC} = \left\{1 - \frac{t_r}{2T}\right\}^2$$

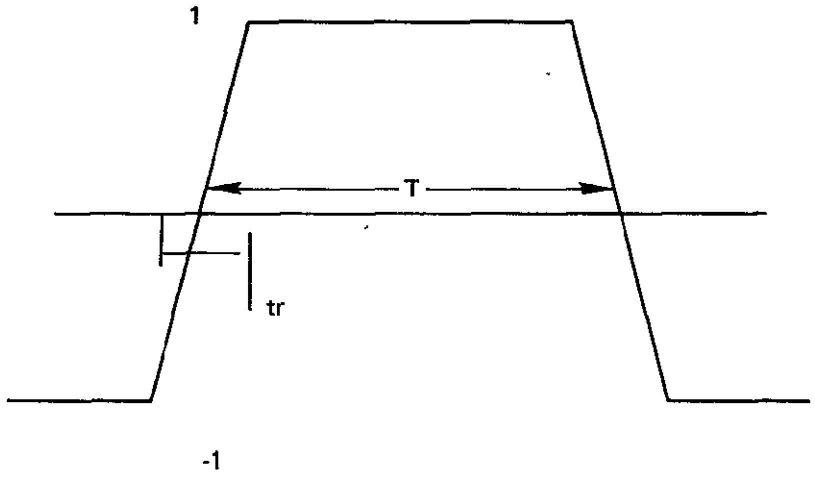


Figure 5-11 Trapezoidal Pulse Waveform

and the average loss by the lower curve (b) of Figure 5-12 which is generated from

$$\overline{P_e} = \left[\frac{1}{4}\right] P_e \left(D_{WC}\rho_o\right) + \left[\frac{1}{4}\right] P_e \left(D_{BC}\rho_o\right) + \left|\frac{1}{2}\right| P_e\left(D_{TC}\ \rho_o\right) = P_e \left(D_{AC}\rho_o\right)$$

where D_{BC}, D_{TC}, and D_{AC} are the degradations in the best case, typical case, and average case, respectively, and ρ_o is the nominal E/N_o. Figure 5-12 illustrates these losses. The additional loss occurs since the receiver detector, being matched to a rectangular pulse, accepts too much thermal noise power.

5.4.2 3-dB Hybrid Biphase Modulator

A biphase modulator can be configured using a 3-dB hybrid with two ports terminated by a pair of varactor diodes, each resonated by an inductance at the carrier frequency when back-biased. Each diode represented could actually be a parallel combination of several diodes. Each diode ensemble reflects most of the indirect rf energy, when switched on, and this reflected power is summed in the other arm of the hybrid coupler.

5.4.3 Quadriphase Modulator Using Hybrid/Combiner

Figure 5-13 shows a method of combining two biphase modulators, using a 3-dB hybrid and in-phase signal combiner to produce a quadriphase modulated carrier. In this circuit, the carrier is split into quadrature components to drive each of biphase modulators whose outputs are then combined to form a quadriphase modulated carrier wherein the 0°, 180° phases represent one serial bit stream, and the 90°, 270° phases represent another bit stream.

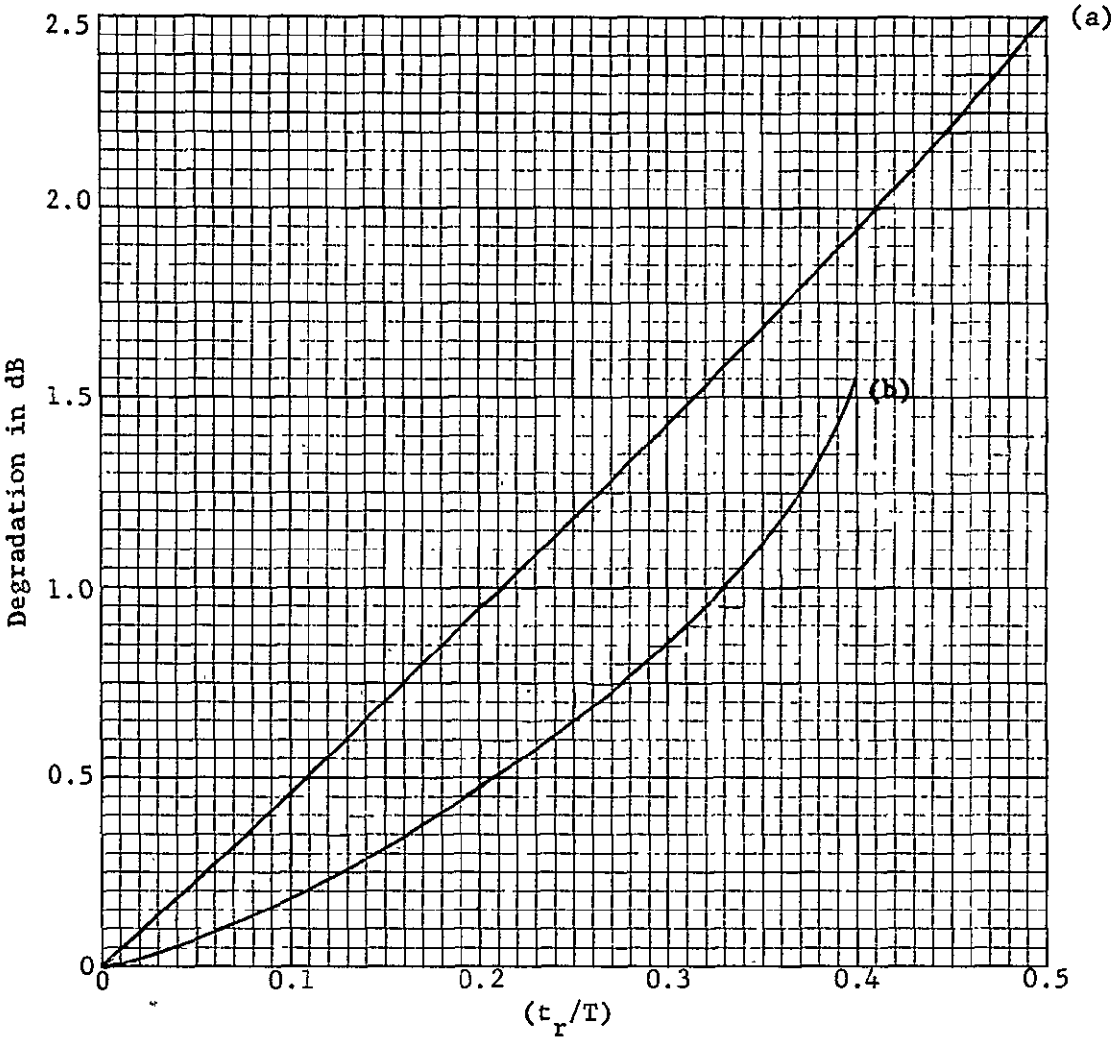


Fig. 5-12 Mis-matched filter losses versus rise-time-to-pulse-duration ratio; a) worst case, and b) average.

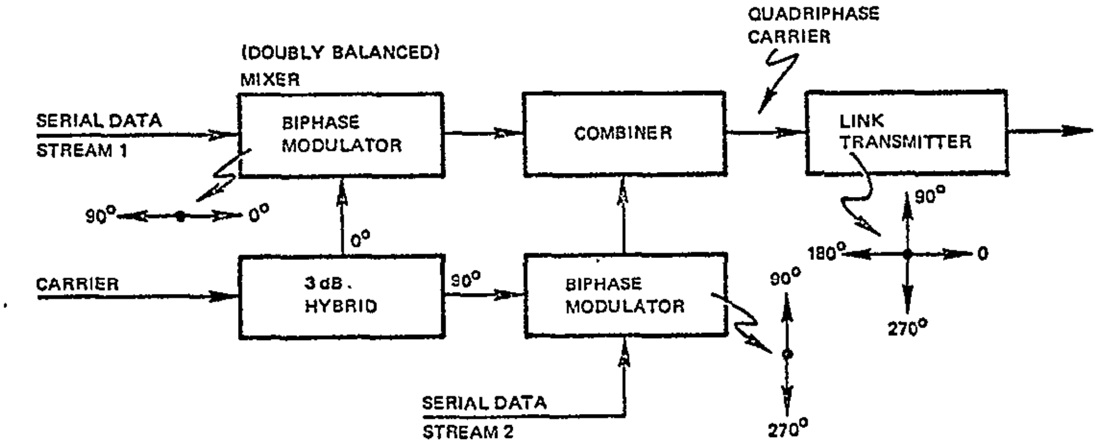

QUADRIPHASE CARRIER
(DOUBLY BALANCED) MIXER
SERIAL DATA STREAM 1
BIPHASE MODULATOR
COMBINER
LINK TRANSMITTER
90°
0°
0°
CARRIER
3 dB. HYBRID
90°
BIPHASE MODULATOR
90°
180°
0
270°
90°
SERIAL DATA STREAM 2
270°

Fig. 5-13 One Type of Quadriphase Modulator Using Signal Combiner and Biphase Modulator

5.4.4 Modulator RF Driver

The configuration of the oscillator driven of the biphase modulator will be determined by the type of demodulator used in the receiver. If the receiver is of the coherent type (phase-lock), the oscillator must have very low phase jitter and a crystal oscillator-amplifier-multiplier is mandatory, either to be used direct or at lower power for injection locking of an avalanche diode. If the receiver uses a differentially coherent demodulator, noise frequencies up to several MHz will be well correlated thus leaving the phase detector output practically noisefree.

This points to the possible use of a free-running avalanche diode as the oscillator driver of the biphase modulator. The noise distribution of this type of oscillator is (nearly) given by the 1/f rule where f is the frequency displacement from the oscillator frequency. Therefore the noise is considerably reduced before frequency range of poor cross-correlation in the phase detector is reached.

The use of a free-running avalanche diode oscillator replacing the crystal oscillator-amplifier-multiplier string was tested in an operating 7.5 GHz link whereas the transmitter was modulated with 100 MHz square-wave (200 MB/s) having a rise time of about 2 nanoseconds. Figure 5-14a shows the phase detector output with the crystal oscillator source in the transmitter.

Next the crystal oscillator source was replaced with the avalanche diode oscillator, and the phase detector output was then as shown on Figure 5-14b. Any increase in "fuzziness" caused by noise from the diode oscillator was insignificant.

NOT REPRODUCIBLE

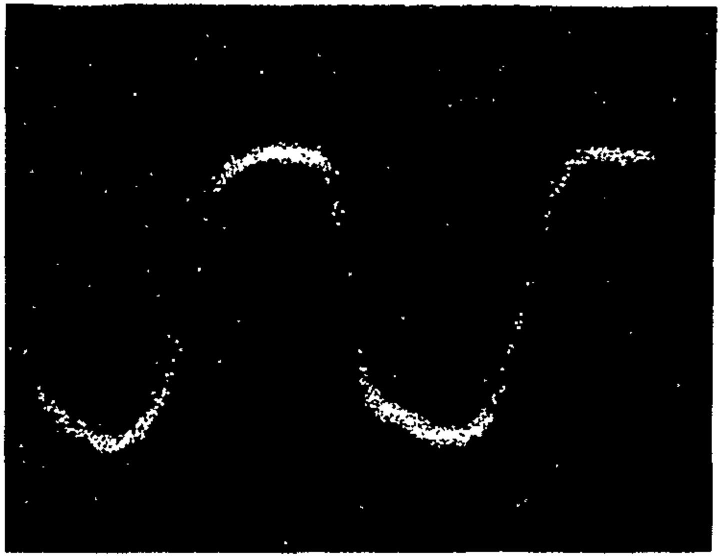


Figure 5-14 a
Crystal Oscillator Source

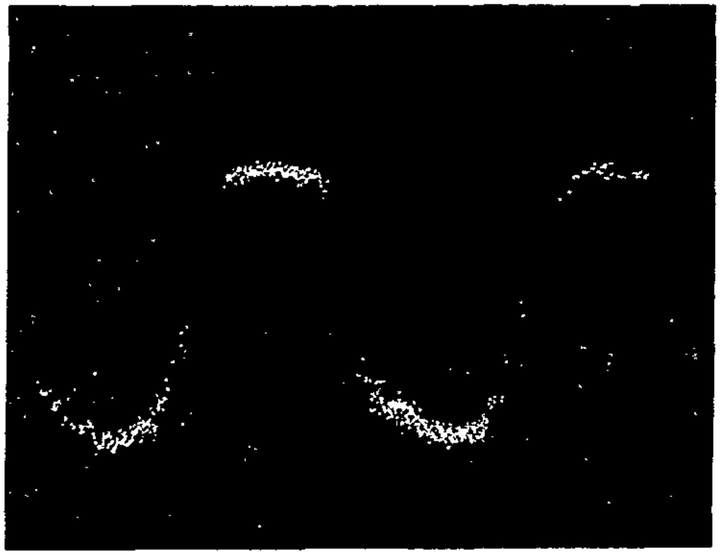


Figure 5-14 b
Avalanche Diode Source

Figure 5-14 Receiver Phase Detector Output
Modulation 100 MHz Squarewave, 200 MB/s
Hor. 2 nanosec./cm

5.4.5 Modulating Diode Driver

The impedance looking into the modulating diode and its circuitry has a VSWR of 1.5:1 for both phase conditions, so the normalized reflected voltage is .871, accounting for the high efficiency with a loss of only 1.2 dB. With about 1.2 dB loss in the circulator, the modulator loss amounts to 2.4 dB, or an efficiency of 57.5 percent.

The driver consists of three emitter-coupled squaring amplifier driver stages and the power output stage. All stages are non-saturating to avoid storage effects. The output voltage levels are controlled accurately by "swing" stabilizers. Figure 5-15 shows a block diagram for the driver.

5.4.6 Differential Encoding at Diode Driver Input

Differential encoding is employed to resolve the inherent data polarity ambiguity concomittant with carrier recovery in a biphase PSK detection system. In differential encoding, a digital "1" input causes a data output transition, i.e., either 1 to 0 to 0 to 1 and a digital "0" input causes no change. Thus, data polarity reversals are irrelevant. The implementation of both differential encoding and decoding is quite simple; see Figure 5-16. A timing diagram illustrating the operation is shown in Figure 5-17.

The data ambiguity may also be resolved by the use of any fixed word of known polarity, such as the digital sync word. In the simplest version, a data polarity reversing gate is placed in the receiver and is controlled by the polarity of the received digital word.

The digital data is differentially encoded at the input to the diode driver of the modulator in order to resolve the 4-phase ambiguity at the demodulator. The data is encoded so that each of the four possible data words corresponds to a multiple of 90^o phase shift in the modulator, according to the following rule:

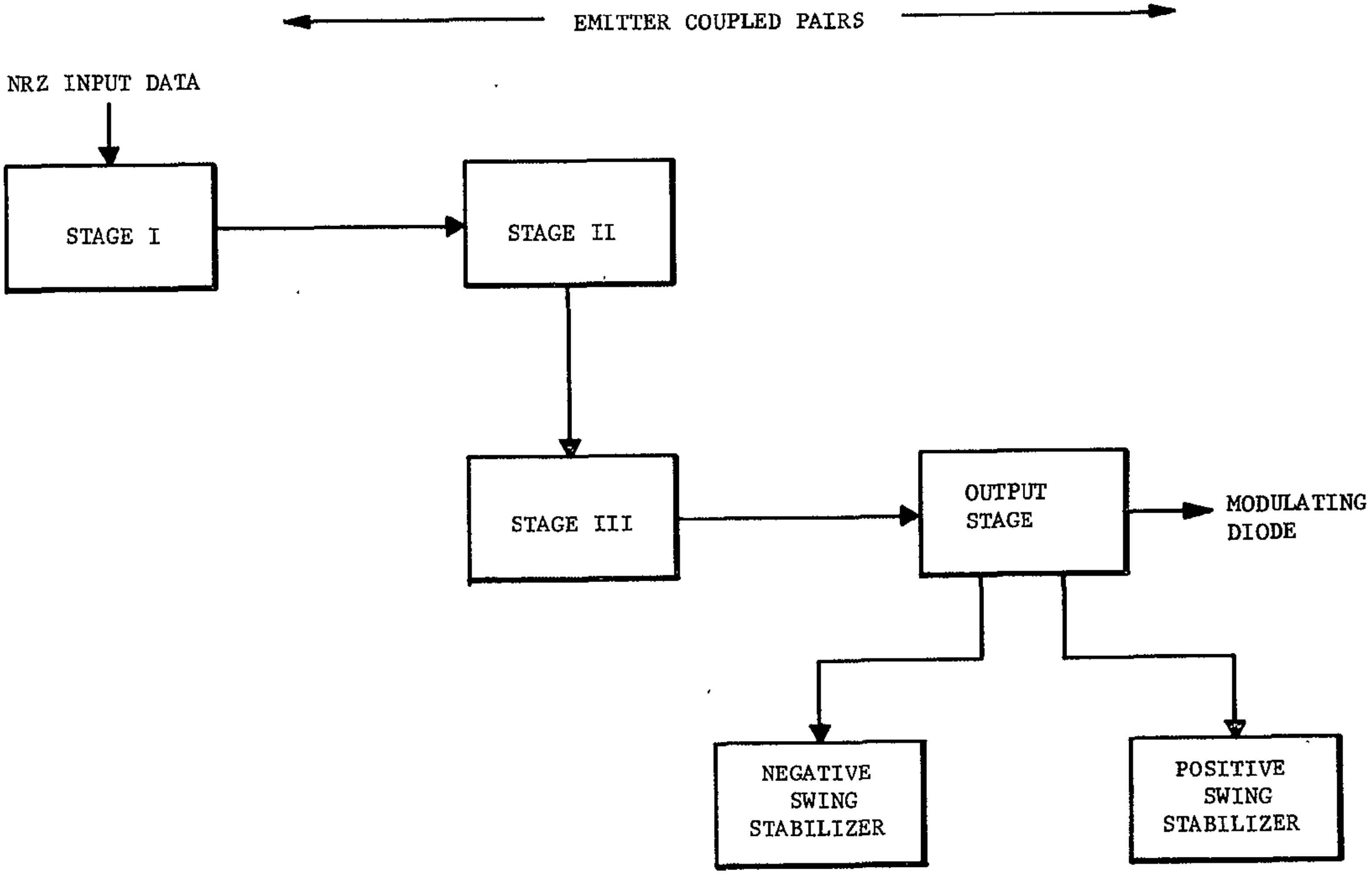

EMITTER COUPLED PAIRS
NRZ INPUT DATA
STAGE I
STAGE II
STAGE III
OUTPUT STAGE
MODULATING DIODE
NEGATIVE SWING STABILIZER
POSITIVE SWING STABILIZER

Figure 5-15 Modulating Diode Driver

Data Word	Modulator Phase Shift
00	0°
01	90°
11	180°
10	270°

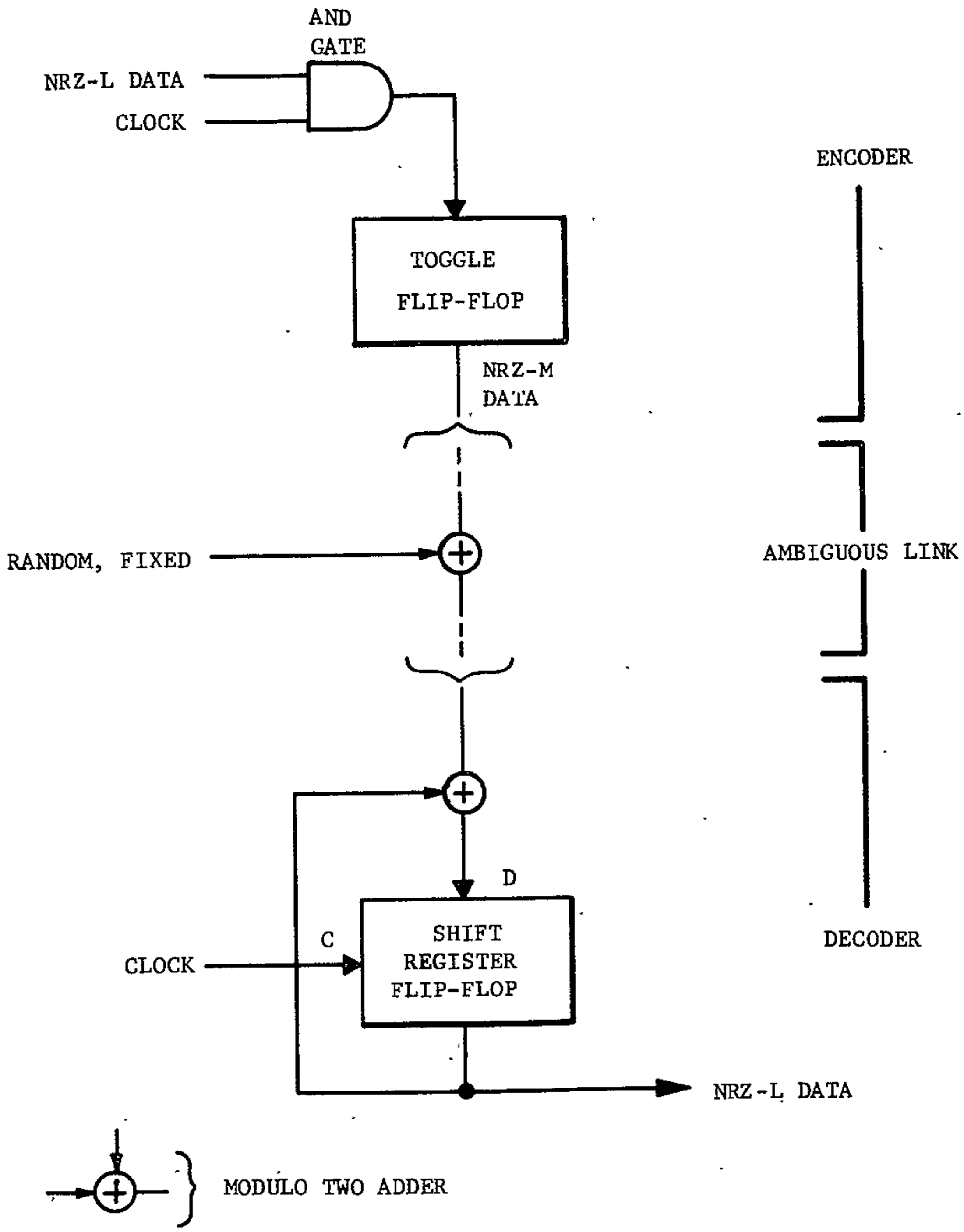
The present output of the differential encoder depends upon the present data word, as well as the previous output of the encoder. Note that the data word determines a shift in the modulator phase. The actual input to the modulator, however, determines a specified phase of the carrier according to the following rule:

Modulator Input*	Carrier Phase
00	0°
01	90°
11	180°
10	270°

*Differential decoder output

A truth table illustrating the operation of the differential encoder is:

		Previous Carrier Phase			
		0° (00)	90° (01)	180° (11)	270° (10)
Present	00	00	01	11	10
input	01	01	11	10	00
data	11	11	10	00	01
words	10	10	00	01	11


AND
GATE
NRZ-L DATA
CLOCK
ENCODER
TOGGLE
FLIP-FLOP
NRZ-M
DATA
RANDOM, FIXED
AMBIGUOUS LINK
D
C
CLOCK
SHIFT
REGISTER
FLIP-FLOP
DECODER
NRZ-L DATA
MODULO TWO ADDER

I-594A

Figure 5-16 Differential Encoder and Decoder

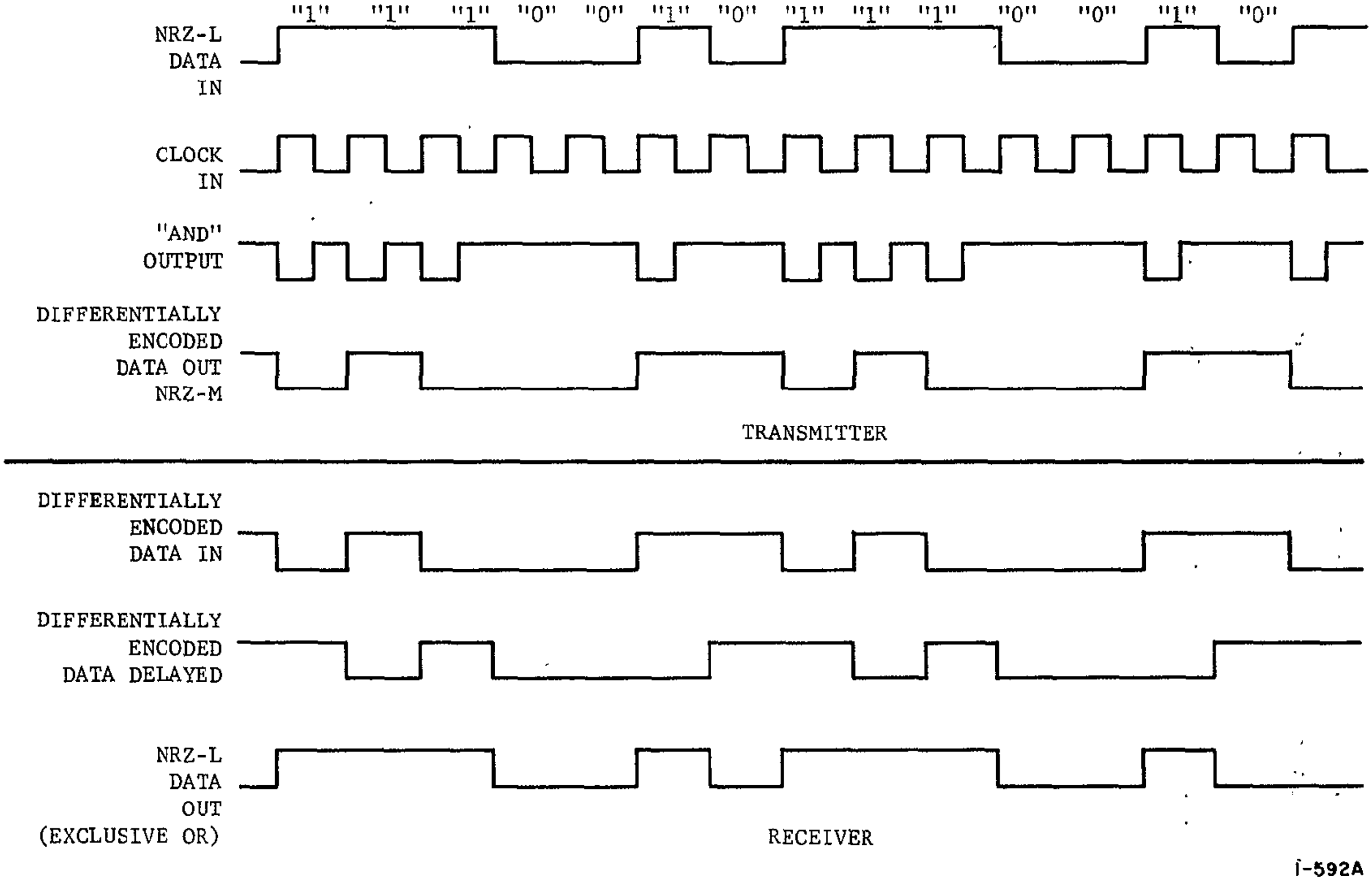
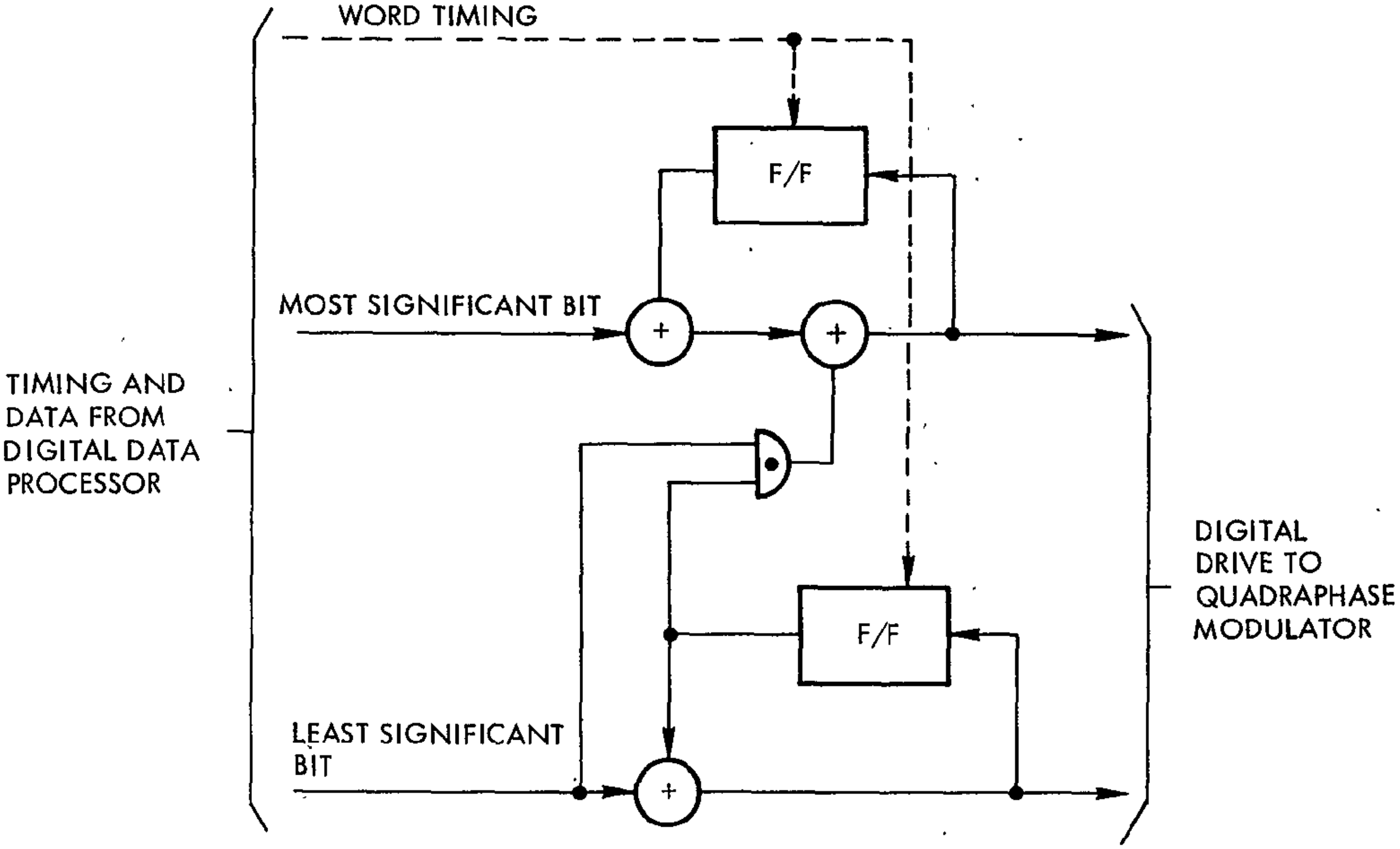

"1" "1" "1" "0" "0" "1" "0" "1" "1" "1" "0" "0" "1" "0"
NRZ-L DATA IN
CLOCK IN
"AND" OUTPUT
DIFFERENTIALLY ENCODED DATA OUT NRZ-M
TRANSMITTER
DIFFERENTIALLY ENCODED DATA IN
DIFFERENTIALLY ENCODED DATA DELAYED
NRZ-L DATA OUT (EXCLUSIVE OR)
RECEIVER
I-592A

Figure 5-17 **Timing** Diagram for Differential Encoding and Decoding

The entries within the truth table are the possible differential encoder outputs (modulator inputs) for the various conditions of the data words and previous carrier phase.

The circuit implementation of the differential encoder is rather simple and is shown in Figure 5-18.


WORD TIMING
F/F
MOST SIGNIFICANT BIT
+
+
TIMING AND DATA FROM DIGITAL DATA PROCESSOR
F/F
LEAST SIGNIFICANT BIT
+
DIGITAL DRIVE TO QUADRAPHASE MODULATOR

(⊕ INDICATES MODULE-2 ADDER)

Figure 5-18 Quadriphase Differential Encoder

5.5 DIFFERENTIALLY COHERENT DEMODULATOR

Where optimum sensitivity, i.e. minimum Energy per Bit/Noise Density for a given error rate must be achieved, a fully coherent demodulator is used. This type of detector utilizes a phase lock loop; it requires the solution of problems regarding acquisition, pull-in range, tracking range, etc. These problems all stem from the fact that a coherent reference signal must be generated or reconstructed from the incoming carrier for use in the phase detector. This reference signal is produced by a voltage controlled crystal oscillator- amplifier- multiplier string. An ideal biphase modulated signal has no carrier, a condition that is approximately true in practical operation, so a coherent CW signal is produced by feeding the signal through a frequency doubler. This signal and the data detector reference, also fed through a doubler, represent the inputs to the loop phase detector.

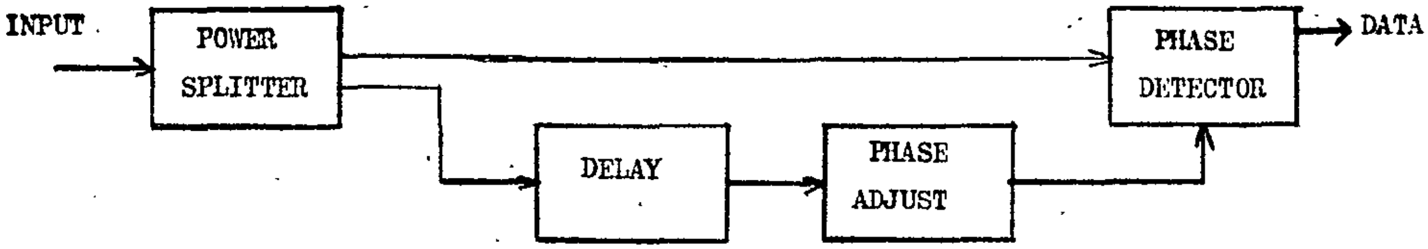
The differentially coherent demodulator does not need a reconstructed reference signal, and is very simple compared to the coherent type. It therefore has important advantages regarding volume, weight, power consumption and reliability factors when used in satellite electronic equipment.

The trade-off for its simplicity is a lower sensitivity. For an error rate of 10^{-4}, the E_b / N_o is about 9.15 DB or about 0.75 DB higher than for the coherent demodulator.

The theoretical performance comparison for coherent and differentially coherent detection of quadriphase PSK is given in Figure 5-19. In practice, the achievable performance with actual equipment will run 2.5 to 3.5 DB below theoretical, with more than half of this degradation attributed to energy loss resulting from modulator switching time and transmitter bandwidth restriction filtering.

5.5.1 Demodulator Configuration

As seen on Figure 5-20, the PSK modulated carrier is split in two parts. One part is fed directly to the reference port of the phase detector. The other part goes through a delay line and a phase adjuster to the signal


INPUT
POWER SPLITTER
DELAY
PHASE ADJUST
PHASE DETECTOR
DATA

DIFFERENTIALLY COHERENT BIPHASE DEMODULATOR.

Figure 5-19

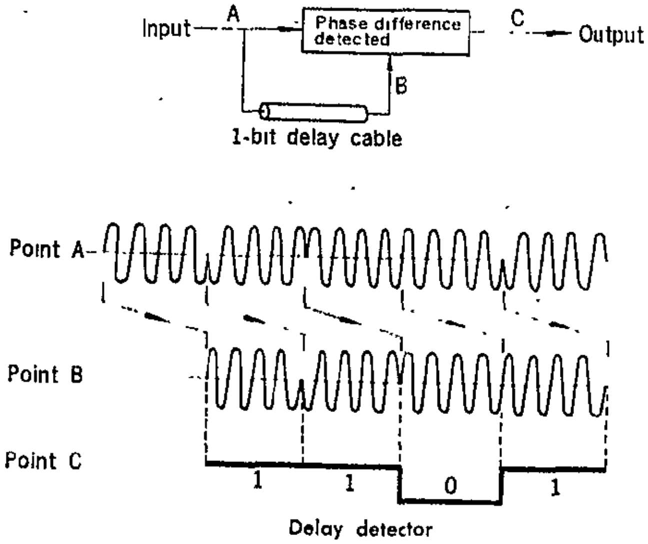

Input
A
Phase difference detected
C
Output
B
1-bit delay cable
Point A
Point B
Point C
1
1
0
1
Delay detector

Figure 5-20 Waveforms of a Differentially Coherent Biphase Demodulator

port. The total delay is made equal to one symbol period, which for 150 Mbps is 6.66 nanoseconds. The phase detector thus operates with the phase from one symbol and that of the previous symbol that now has arrived via the delay line.

As an example, let one received symbol have the phase 0°. If the next symbol also has 0°, both ports will have the same phase during this second symbol period. The phase adjuster has been set so that the detector output is maximum positive for this condition. Both symbols could as well have had the phase 180°, and the port inputs would still be in phase during the second symbol period; i.e., the output would still be positive. Thus, the detector output stays positive for a series of symbols having the same phase.

If, on the other hand, the first symbol with phase 0° is followed by one with phase 180°, during the second symbol period, the port inputs will be antiphase, and the output will be maximum negative. This output will remain negative as long as there is an 180° phase shift from symbol to symbol. The detector output is therefore "coded" as follows, as illustrated in Figure 5-20.

No phase change : positive output (0)
180° phase change : negative output (1)

The input signal to the transmitter modulator is differentially encoded; therefore, with differentially decoding in the demodulator as just described, the phase detector yields the uncoded data without ambiguity.

Figure 5-21 shows the complete differential phase detection and timing recovery circuit used by W. Hubbard,[1] of Bell Telephone Laboratories, for a 300 Mbps link.

The couplers, delay lines, and diode mounts are microwave printed circuits; the filters and combining Tee are coaxial. The differential phase detector and the timing recovery circuit are combined (share a common delay line) in order to save space and cost.

---

[1] W. Hubbard et. al, "A Solid State Regenerative Repeater For Guided Millimeter-Wave Communication Systems", BSTJ, November 1967.

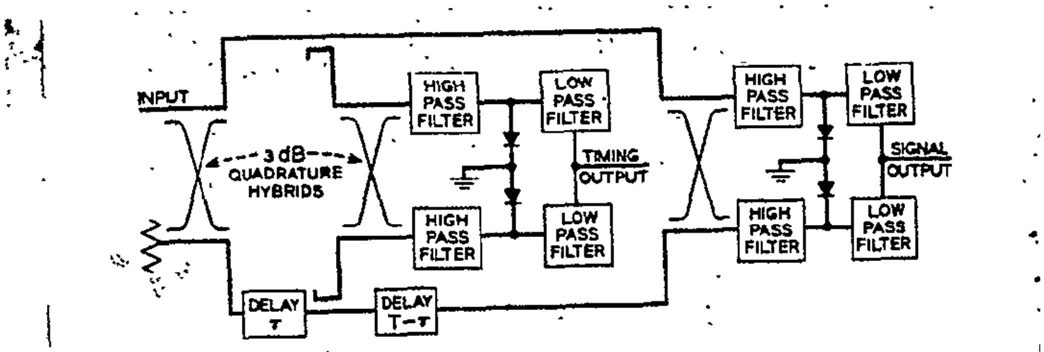

INPUT
3 dB QUADRATURE HYBRIDS
HIGH PASS FILTER
LOW PASS FILTER
TIMING OUTPUT
HIGH PASS FILTER
LOW PASS FILTER
DELAY τ
DELAY $T-\tau$
HIGH PASS FILTER
LOW PASS FILTER
SIGNAL OUTPUT
HIGH PASS FILTER
LOW PASS FILTER

Fig. 5-21 [1]Differential phase detector and timing recovery circuit.

A straightforward analysis shows that the output voltage of this circuit is given by

$$V(t) = \cos\left\{\omega_0\tau + \int_{t-\tau}^{t} \omega(t')\,dt'\right\}$$

for an input DPSK signal given by

$$S(t) = \sqrt{2}\cos\left\{\omega_0 t + \int_0^{t} \omega(t')\,dt'\right\},$$

with

$$|\omega(t)| = |\omega(t+nT)|, \qquad \int_{(n-\frac{1}{2})T}^{(n+\frac{1}{2})T} \omega(t')\,dt' = \pm\pi/2.$$

If $\omega_0\tau$ is chosen to be an odd multiple of $\pi/2$ and τ is chosen equal to T, the reciprocal of the bit rate, at the sampling instants $[t = (n+\frac{1}{2})T]$ the output is given by

$$V(t) = \cos\{(m+\tfrac{1}{2})\pi \pm \pi/2\} = \pm 1.$$

Thus, under these conditions the device is the desired differential-phase detector for this signal.

5.5.2 Differentially Coherent Detection Operation

Figure 5-22 shows the simple configuration of the differentially coherent demodulator. Its operation is described as follows.

The input signal is split into two parts. One part goes through a power splitter and phase adjusters to the "Reference" ports of the two data detectors. The other part goes through a delay line and a power splitter to the "Signal" ports. The delay at the signal ports is made (very nearly) equal to one symbol period. With quadriphase modulation (2 bits/symbol) and at 100 Mbps one symbol period equals 20 nanoseconds. The phase detectors thus operate with the phase of one symbol at the reference port and the phase of the previous symbol, which now has arrived via the delay line, at the signal port. In other words, the phase detector outputs indicate the difference in phase between two consecutive symbols.

The operation of Phase Detector #1 will first be considered. With an IF frequency of 1000 MHz, for example, one period (cycle) lasts one nanosecond.

The required delay for 200 Mbps is:

$$\frac{2}{200 \times 10^6} = 10 \text{ nanoseconds}$$

This is equivalent to:

$$\frac{10}{1} = 10 \text{ cycles at } 1000 \text{ MHz}$$

To get the proper phase relation between the phases of Reference 1 and Signal Phase A, the Signal Phase A must be delayed the equivalent of:

9-7/8 or 9.875 cycles of Reference 1

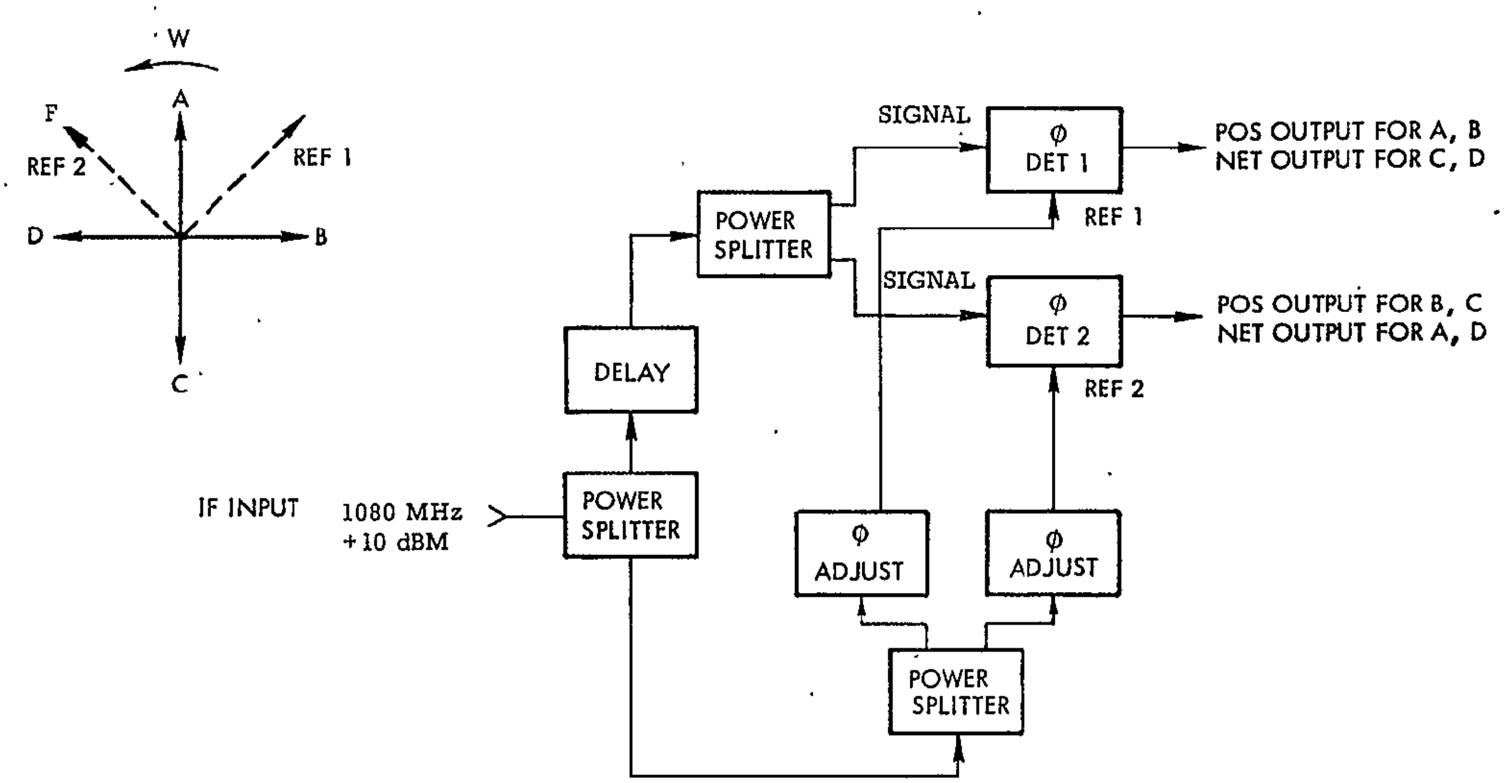

W
A
F
REF 2
REF 1
D
B
C
IF INPUT
1080 MHz
+10 dBM
POWER SPLITTER
DELAY
POWER SPLITTER
SIGNAL
φ DET 1
REF 1
SIGNAL
φ DET 2
REF 2
POS OUTPUT FOR A, B
NET OUTPUT FOR C, D
POS OUTPUT FOR B, C
NET OUTPUT FOR A, D
φ ADJUST
φ ADJUST
POWER SPLITTER

Figure 5-22 Differentially Coherent Quadriphase Demodulator

Thus the actual delay will be:

$$1.0 \times 9.875 = 9.875 \text{ nanoseconds}$$

The "timing error" will be about 125 picoseconds.

For Phase Detector #2, the Signal Phase A must be delayed the equivalent of:

10-1/8 cycles

and the actual delay is:

$$1.0 \times 10.125 = 10.125 \text{ nanoseconds}$$

so here the timing error is also 125 picoseconds. This is still only a small fraction of the rise time and will affect the integrate and dump value very little.

The 4 phase positions are related to the phase detector outputs as shown on Figure 5-22. If symbol after symbol has the same phase, the phase detector outputs will indicate Phase A. The next symbol may go to the three other positions: $+90^{\circ}$, 180°, -90°, and it is easy to show the two detector outputs for these phase changes:

Phase change for next symbol	None	$+90^{\circ}$	180°	-90°
Output Ø Det. #1	+	+	-	-
Output Ø Det. #2	+	-	-	+
Indicated Phase	A	B	C	D

The transmitter is differentially encoded according to the top and bottom lines in this table. If the receiver is to indicate a B phase, the transmitter goes $+90^{\circ}$ from whichever position it happens to be in. If Phase C should be indicated, the transmitter goes 180°, and so on. So by differentially encoding the transmitter and with differentially coherent demodulation, there will be no ambiguity in the data stream output of the demodulator.

5.5.2.1 Data Filter Configurations

The optimum data filters at the output of the phase detectors are integrate-and-dump circuits. Each symbol is integrated over one symbol duration prior to sampling to maximize the SNR at the decision instant. The energy is then dumped to prevent interference with the following symbol. The configuration of the data filters and output logic are shown in Figure 5-23.

The data filter will be a single-stage RC with a time constant approximately 10 times the symbol duration. The time constant will be switched for each of the three data rates. The energy on the integrator capacitor will be dumped to prevent intersymbol interference. To allow sufficient energy discharge time, each data channel will use two time-shared filters.

The decision threshold is set at zero volts so that a hard decision is made as to whether the integrated phase detector output is positive or negative at the end of each symbol. Referring to Figure 5-23, the decision threshold outputs (x and y) will have the following values:

x = positive for A or B
negative for C or D

y = positive for B or C
negative for A or D

A, B, C, and D refer to the original data symbols 00, 01, 10, and 11, respectively.

The logic resolves x and y into a single received symbol and outputs a 2-bit parallel word according to the following rules:

$A = 00 = x\,\bar{y}$

$B = 01 = x\,y$

$C = 10 = \bar{x}\,y$

$D = 11 = \bar{x}\,\bar{y}$

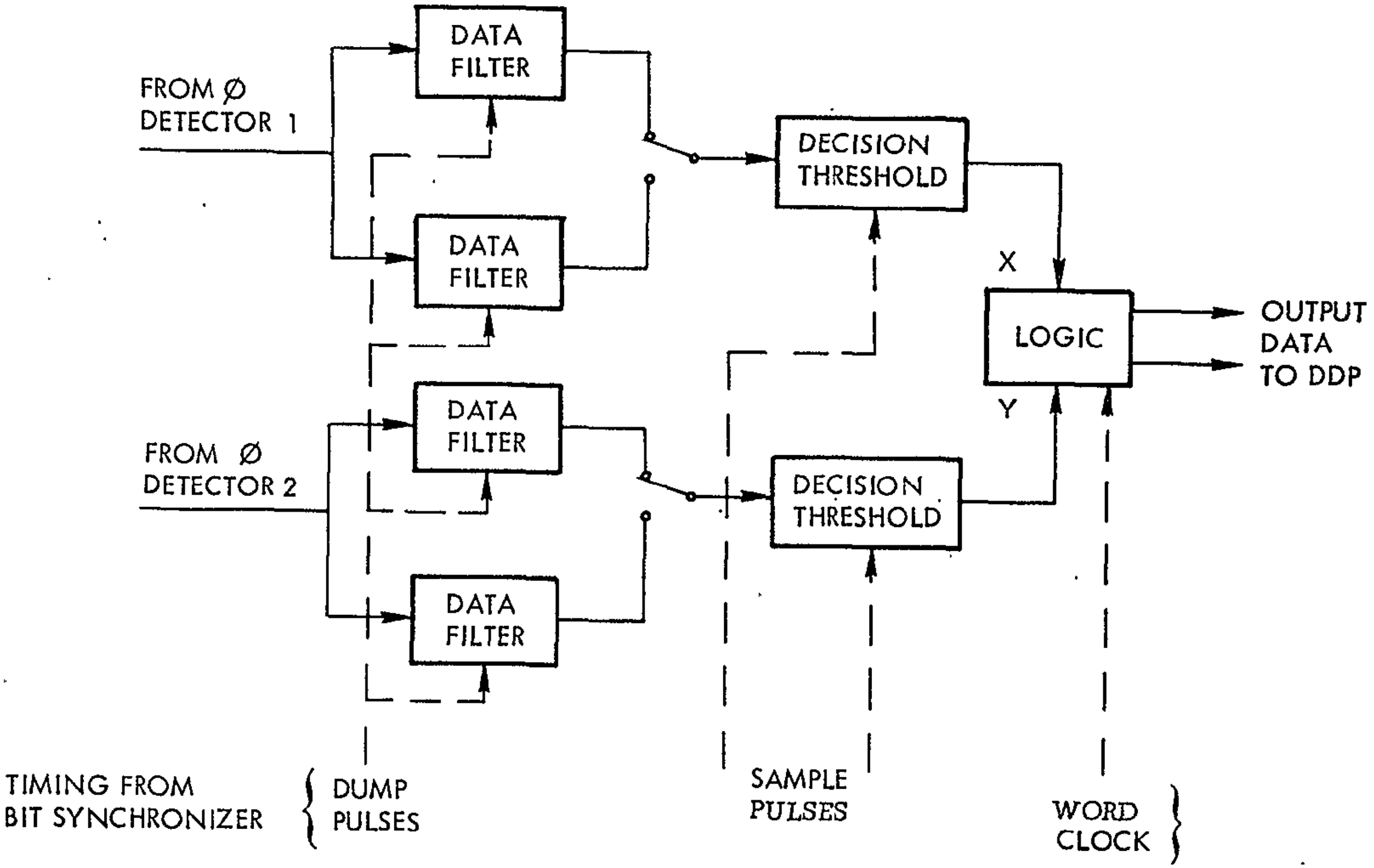

FROM Ø
DETECTOR 1
DATA FILTER
DATA FILTER
DECISION THRESHOLD
X
LOGIC
OUTPUT
DATA
TO DDP
Y
FROM Ø
DETECTOR 2
DATA FILTER
DATA FILTER
DECISION THRESHOLD
TIMING FROM
BIT SYNCHRONIZER
DUMP
PULSES
SAMPLE
PULSES
WORD
CLOCK

Figure 5-23 - Data Decision Circuitry

5.5.2.2 Bit Synchronizer

The bit synchronizer shown in Figure 5-24 will generate a coherent data clock from the output of the quadriphase detectors. The data clock will be used as a basis for sampling of the data filter outputs, as well as provide a digital clock required by the digital data processor.

The bit synchronizer operates by phase locking a crystal oscillator to the data transitions at the output of the quadriphase detectors. A pulse is generated for each transition by multiplying the detector output by itself delayed by 1/2 symbol. One such transition detector is employed for each of the two quadriphase detectors in order to maximize the number of transitions. The transition detectors are summed in a linear summer and the resulting output is used as the input to a conventional phase-locked loop. The bandwidth of the phase-locked loop is chosen to minimize timing jitter from additive noise and multipath.

Since the bit synchronizer locks to the quadriphase data bit stream with a clock rate equal to the symbol rate (1/2 the output bit rate), a divider is used in the phase-locked loop which allows the VCXO to operate at the data bit rate. Also, since the phase-lock loop locks with a clock in quadrature with the actual symbol clock, a second divider is used which, when toggled from the inverted bit clock, will produce a symbol clock of the proper phase.

To change bit rates, the delay is switched to 1/2 a data symbol and the VCXO output is switched to the proper frequency.

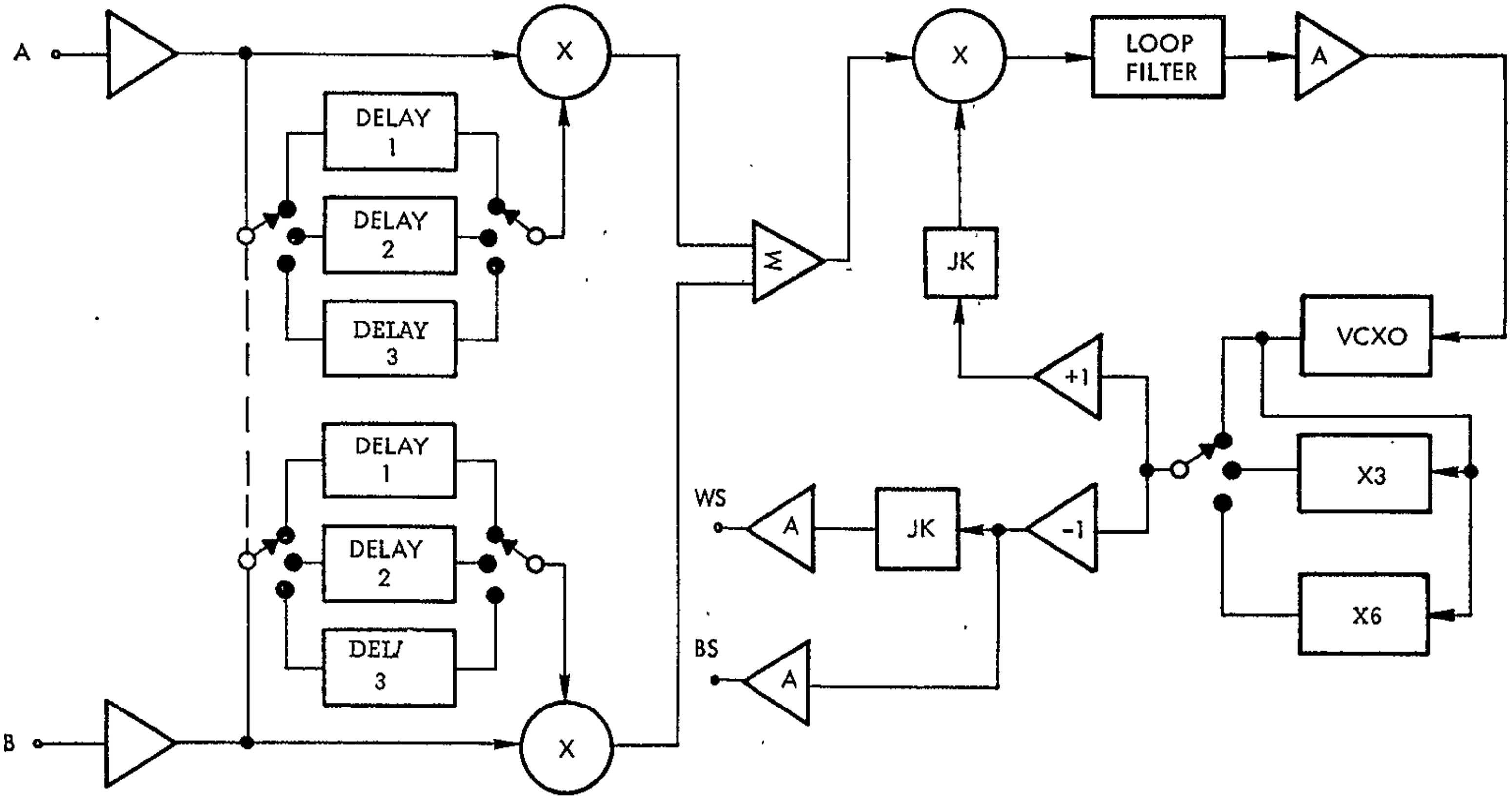

A
B
DELAY 1
DELAY 2
DELAY 3
DELAY 1
DELAY 2
DELAY 3
X
X
Σ
X
LOOP FILTER
A
JK
VCXO
+1
-1
X3
X6
WS
A
JK
BS
A

Figure 5-24 Bit Synchronizer

5.6 Coherent and Differentially Coherent Phase Detection

The PSK signal can be demodulated either by coherent or differentially coherent phase detection. Coherent detection involves the generation of a phase coherent reference carrier at the demodulator using a phase-locked loop. Coherent detection requires approximately 2.4 dB less power than differentially coherent detection but involves added circuit complexity. Further, the phase-locked carrier reconstruction loop will be affected by pull-in range, acquisition time, maintenance of carrier lock and phase jitter due to additive noise and multipath. Good phase-lock loop design can account for all of these link characteristics.

As discussed in Section 5.5, differentially coherent detection uses each preceding received RF symbol as a phase reference for the next symbol. Thus, carrier phase shifts from symbol to symbol are detected without the necessity for establishing a fixed phase-locked reference.

The theoretical performance comparison for coherent and differentially coherent detection of quadriphase PSK is given in Figure 5-25. In practice, the achievable performance with actual equipment will run 2.5 to 3.5 dB below theoretical, with more than half of this degradation attributed to energy loss resulting from modulator switching time and transmitter bandwidth restriction filtering.

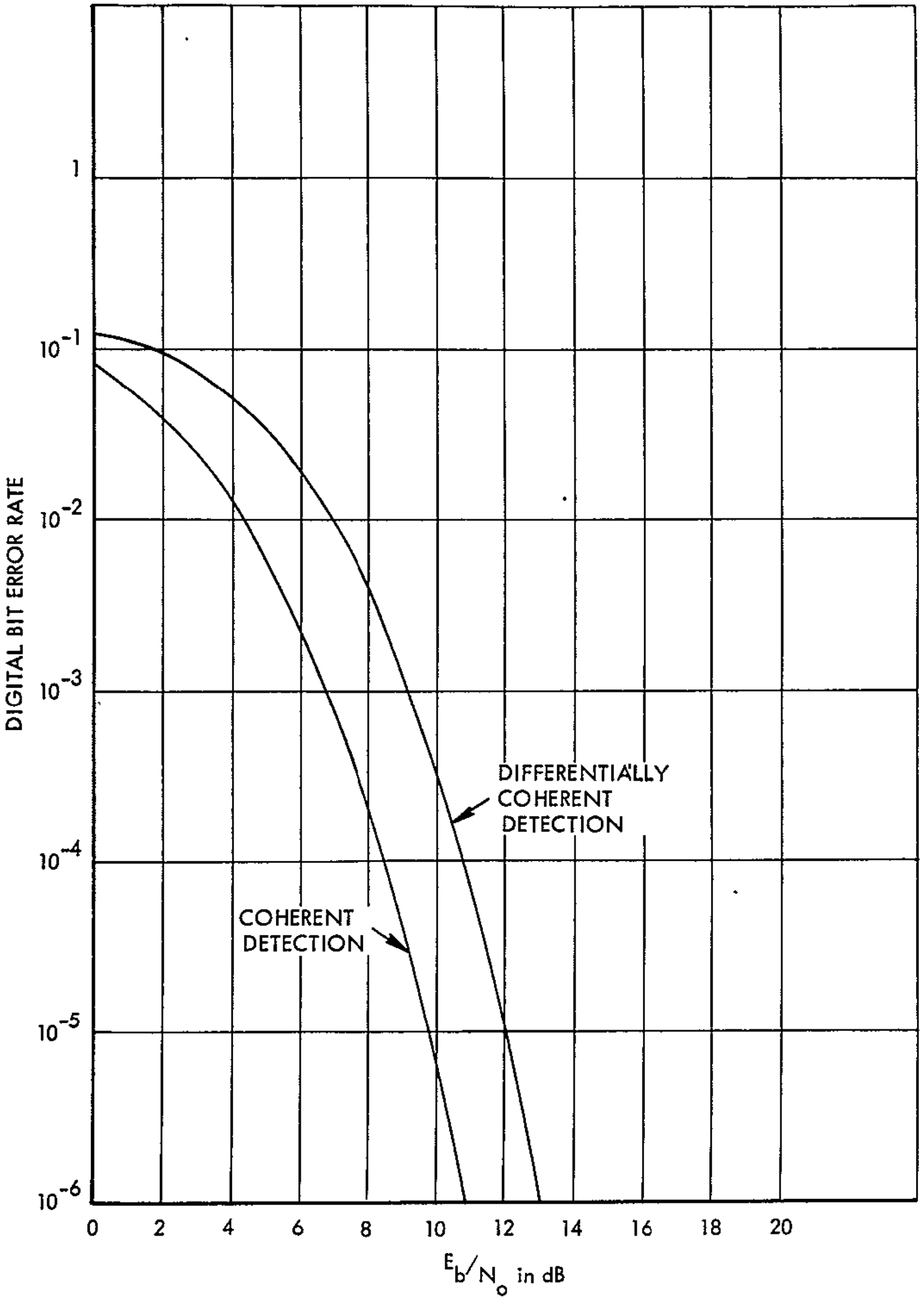


Fig. 5-25 Comparison of Theoretical Performance of Quadriphase PSK with Coherent and Differentially Coherent Detection

5.6.1 PSK (Coherent) Demodulation - Carrier Reconstruction

PSK refers to phase shift keying by a binary message. The phase of the carrier is shifted by 180° each time the binary data changes. Coherent detection requires that the receiver have knowledge of the phase of the carrier with an uncertainty of exactly 180°. A phase-locked carrier reconstruction loop is used at the receiver to generate the coherent phase reference.

Two types of carrier reconstruction loops are used for binary PSK: (1) the Costas loop or I-Q loop and (2) the frequency doubling loop. For quadriphase modulation, the frequency quadrupling loop can be used.

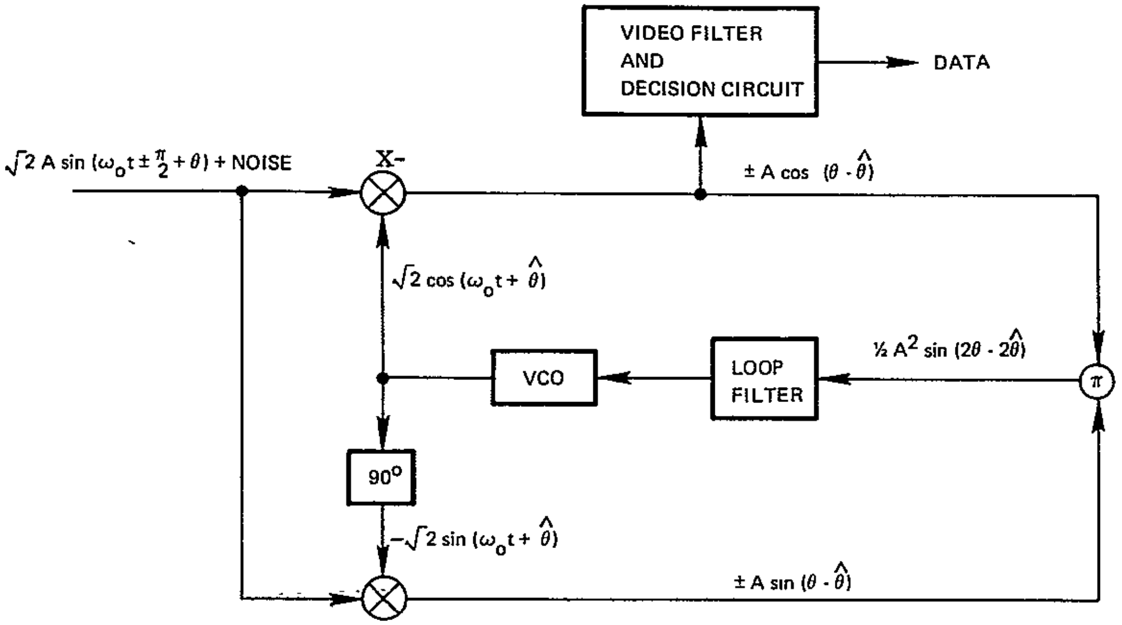
The Costas loop generates a reference carrier signal locked in phase to either of the two possible received signal phases by mating both an in-phase and phase-quadrature phase comparison as shown in Figure 5-26a. In the frequency doubling loop a coherent reference is generated at twice the received center frequency. The modulation is removed by the frequency doubling. The operation of the doubling loop is shown in Figure 5-26b.

The error probability for coherent PSK has been listed in Vol. 2, Sec. 2; that is

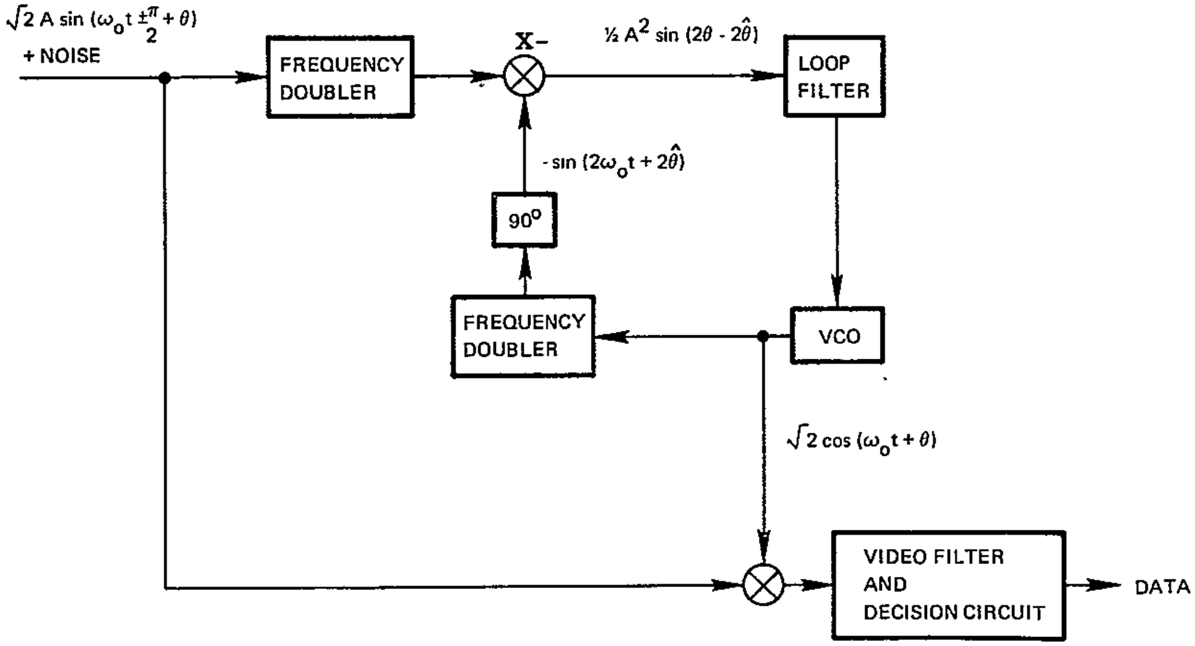
$$\mathrm{BER}_{\mathrm{PSK}} = \phi\left[\sqrt{2E_b/N_o}\right] .$$

With regard to terminology, it might be pointed out that PSK and PCM/PM are identical when ±90° phase deviation is used in PCM/PM. Phase reversal keying (PRK) and PSK are also identical. PSK is also commonly referred to simply as bi-phase modulation.

For complete phase reversal keying; i.e., ±90° phase deviation, there is no explicit carrier component in the signal. For applications in which a residual carrier is required such as for Doppler tracking, a phase deviation of less than ±90° can be used. For PSK with $\pm\theta$ radians deviation ($\theta \leq \frac{\pi}{2}$), the ratio of the carrier component to the total signal power is $\cos^2\theta$. The BER is given by:


VIDEO FILTER AND DECISION CIRCUIT
DATA
$\sqrt{2}\,A \sin(\omega_0 t \pm \frac{\pi}{2} + \theta)$ + NOISE
X-
$\pm A \cos(\theta - \hat{\theta})$
$\sqrt{2} \cos(\omega_0 t + \hat{\theta})$
VCO
LOOP FILTER
$\frac{1}{2} A^2 \sin(2\theta - 2\hat{\theta})$
π
90°
$-\sqrt{2} \sin(\omega_0 t + \hat{\theta})$
$\pm A \sin(\theta - \hat{\theta})$

(a) Costas Loop PSK Detector


$\sqrt{2}\,A \sin(\omega_0 t \pm \frac{\pi}{2} + \theta)$
+ NOISE
FREQUENCY DOUBLER
X-
$\frac{1}{2} A^2 \sin(2\theta - 2\hat{\theta})$
LOOP FILTER
$-\sin(2\omega_0 t + 2\hat{\theta})$
90°
FREQUENCY DOUBLER
VCO
$\sqrt{2} \cos(\omega_0 t + \theta)$
VIDEO FILTER AND DECISION CIRCUIT
DATA

(b) Frequency Doubling Phase Lock PSK Detector

Figure 5-26 Carrier Reconstruction Loop

$$BER_{PSK} = \varphi\left(\mathrm{Sin}\ \theta \sqrt{2E_b/N_o}\right),$$

for $\pm\theta$ radians deviation. Typically one radian deviation is used, so that 1.5 dB more power is required to achieve the same BER possible with $\theta = \pi/2$.

Figure 5-27 shows the block diagram of a coherent biphase demodulator. The function of the frequency doubler carrier reconstruction loop is described as follows: Carrier reconstruction is performed by a voltage-controlled oscillator/frequency miltiplier operating in a phase-locked loop. As can be seen in Figure 5-27, the "Signal" for the loop detector is obtained by feeding part of the biphase modulated signal through a frequency doubler. (When the input signal to the doubler "flips" 180°, The doubler output phase change is 360°, i.e., it does not move at all.) The output from the doubler is a CW signal with twice the frequency and a fixed phase relation relative to the (actually nearly missing) input carrier. An extra doubler fed from the synthesized reference signal for the loop detector completes the loop.

The loop calculations are performed in the conventional manner with parameters such as tracking range, pull-in range, tracking rate, etc. The VCXO must have as clean a spectrum as practically possible (low phase jitter), and must have good temperature and long-term stability so as not to require too large a part of the loop detector characteristic. These requirements, of course, hold true also for the transmitter oscillator.

The input level to the demodulator is usually at least 10 mW and the data detector output will be about ± 75 mV with + 2 dBM signal and + 8 dBM reference. The loop detector operates with about -6 dBM signal and 0 dBM reference.

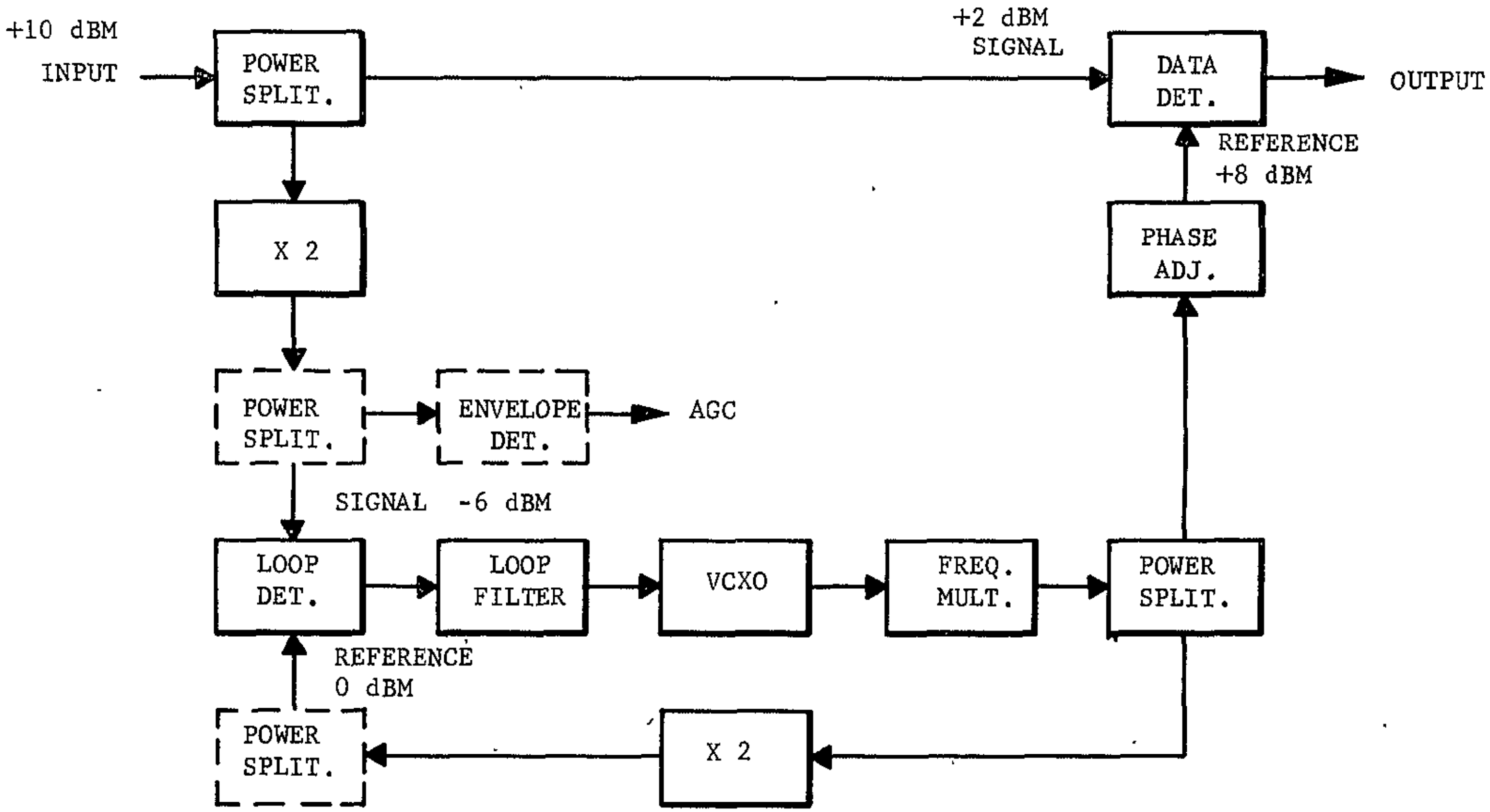

+10 dBM
INPUT
POWER SPLIT.
+2 dBM SIGNAL
DATA DET.
OUTPUT
REFERENCE +8 dBM
X 2
PHASE ADJ.
POWER SPLIT.
ENVELOPE DET.
AGC
SIGNAL -6 dBM
LOOP DET.
LOOP FILTER
VCXO
FREQ. MULT.
POWER SPLIT.
REFERENCE 0 dBM
POWER SPLIT.
X 2

Figure 5-27 Coherent Biphase Demodulator

5.6.1.1 Times-Four Quadriphase Carrier Reconstruction Loop

The times-four carrier reconstruction loop uses frequency quadruplers instead of frequency doublers in the phase lock PSK detector of Figure 5-26b.

The data modulation is removed by the fourth harmonic operation since any multiple of 90° when multiplied by four yields 0° (modulo 360°). Limiting the RF bandwidth prior to the X4 operation decreases the time (due to finite rise times) that the signal is at a multiple of 90° and this decreases the recovered carrier power. However, decreasing the bandwidth at the X4 input significantly improves the signal-to-noise ratio and decreases the small signal suppression effect. Analyses and experiment have confirmed that filtering (3 dB bandwidth) at the first spectral nulls represents a reasonable compromise between these effects.

More severe filtering than that proposed above can create a data sideband lock problem if periodic data is transmitted for test purposes. The bandwidth limiting converts the digital phase modulation to amplitude modulation which passes through the X4 device unaffected and creating sidebands at the repetition rate of the test word. While this problem is not usually severe it can be troublesome particularly for unattended operations that encounter periodic words.

The output of the X4 device should be reasonably narrow so as to increase the efficiency of the X4 device. However, the bandwidth must be great enough to encompass the frequency uncertainty (at 4 f_0). In fact, since it is impossible to perfectly match the two X4 devices (one in the signal path and one in the reference path), it is desirable to make the filters sufficiently broad that the frequency uncertainty produces relatively little phase shift (e.g., 5° or 10°). Thus, with this approach one can readily expect the filter to track to within 0.5 to 1 degree producing degradation of less than 0.1 dB.

The X4 device can be achieved either with a single stage or two doubling stages. Theoretically, identical performance is achieved. However, the latter approach tends to be favored since it keeps the desired signals at relatively high levels at all points and thus avoids problems with spurious signals causing interference. When the two doubling stages approach is taken, the signal at the output of the first X2 device is biphase modulated. Thus, the filter bandwidth at this point must be no narrower than at the input.

The two X4 devices should be made as nearly identical as possible so as to increase the likelihood of accurate phase tracking with frequency uncertainty and temperature drift. Poor tracking will cause a phase offset error at the phase detectors and degrade performance. This effect is discussed in a subsequent section on implementation degradations.

The harmonic devices should be of the full wave type rather than snap action diodes. The former type offer more stable phase shift with temperature changes and signal level changes than the latter.

5.6.2 Carrier Reconstruction Loop Bandwidth Design

A second-order loop with nominal dumping coefficient of $\zeta = 0.707$ is standard. The permissible closed-loop bandwidth is bounded from above by thermal noise and loop threshold considerations and from below by dynamic tracking errors and acquisition time requirements. Some typical calculations are performed below.

The closed-loop signal-to-noise ratio can be shown to be given by

$$SNR_{loop} = SNR_{RF} + (W/B)dB - 12\text{ dB} - \Psi\text{ dB}$$

where SNR_{RF} is the RF signal-to-ratio prior to the X4 operation, W is the RF bandwidth, B is the closed-loop bandwidth of the phase-lock-loop, and Ψ is a small signal suppression degradation effect.

The RF signal-to-noise ratio is related to the energy-per-bit-to-noise spectral density ratio by

$$SNR_{RF} = (E_b/N_o)\ (R/W)$$

where R is the data rate in megabits per second (Mbps). For a worst case bound on performance let us assume that error rates as high as 1% are acceptable and that the RF bandwidth W is set at 2.5 times the data rate corresponding to very broad filtering - more typically the filter bandwidth might be set equal to the data rate. In this worst case the loop signal-to-noise ratio can be shown to be equal to

$$SNR_{loop} = 17.5 \text{ dB} + (W/B) \text{ dB}$$

Assuming a very conservative loop threshold of 12.5 dB this implies that B = 0.001 W. Taking a 300 Mbps example we find that B = 750 kHz. If the loop bandwidth is decreased to 75 kHz the minimum loop signal-to-noise ratio will be increased to 22.5 dB which offers extremely strong lock.

More typically one might choose to set W = R where R is the data rate and to operate at a 10^{-4} error rate. In this case E_b/N_o is increased by 3.5 dB, and SNR_{RF} is increased by 7.5 dB. As a result, the small signal suppression is decreased by 4.5 dB. The net increase in signal-power-to-noise density ratio prior to the phase-lock loop is only 8 dB since the noise is concentrated in 0.4 of the previous (worst case) bandwidth. As a result, the loop signal-to-noise ratio is 8 dB higher for the typical case than for the worst case. Thus, the nominal loop signal-to-noise ratios are 20.5 and 30.5 dB for the wide and narrow loops, respectively. The worst case values will be used in the remaining analysis for the purposes of establishing a conservative design.

Let us now consider the loop bandwidth requirements, or rather the capabilities of the two bandwidths of 75 and 750 kHz. The acquisition time is given by

$$T_{acq} = T_F + T_\phi$$

where $T_F \approx 4\ (\Delta\ f)^2/B^3$ and $T_\phi \approx 4/B$, correspond to frequency and phase acquisition, respectively. The resulting acquisition times are 343 milliseconds and 348 microseconds for the 75 and 750 kHz bandwidths, respectively.

The loop gain required to achieve a pull-in range of ± 6 MHz (at 4 f_o) is given by

$$k \approx 4\pi^3\ \Delta\ f^2/B \approx 2020\ \Delta\ f^2/B$$

which corresponds to loop gains at 5.82×10^9 and 5.82×10^{10} for 75 and 750 kHz loop bandwidths, respectively. These loop gains are quite high and can probably be best achieved through the use of LC rather than the crystal VCO's.

The loop bandwidth is also set by the maximum transient error, due to a parabola of phase, that is, the rate of change of doppler. In this section the peak phase error tolerable (at 4 f_o) is assumed to be 10^o which yields a maximum degradation of 0.2 dB and which exists only during the period of maximum acceleration. The peak transient error (in degrees) is given by

$$E_D = 1.12\ a/B^2$$

where $\phi\ (t) = at^2$ describes the path length change measured in degrees. Thus, $a < 8.90\ B^2$ or expressed as doppler rates at f_o the maximum tolerable doppler rates are 35 MHz/sec and 3.5 GHz/sec for the 75 kHz and 750 kHz bandwidths, respectively. It is anticipated that the narrow bandwidths suffice for most applications.

The loop bandwidth must also be wide enough to accommodate the link phase jitter and incidental FM due to imperfect up and down conversion operations. Typically, phase jitter and incidental FM are only problems for very low data rate systems where loop bandwidths must be on the order of 100 Hz or lower. Thus, no difficulty is anticipated with these phenomena since the minimum loop bandwidth proposed is three orders of magnitude broader. Previous analyses on quadriphase modems have established the feasibility of 100 Hz loop bandwidths at X-band. Certainly, three orders of magnitude loop bandwidth increase can accommodate one order of magnitude change in the RF frequency.

5.6.2.1 Implementation Errors of Carrier Reconstruction Loop

This part summarizes a portion of the extensive analyses of signalling performed with Philco-Ford quadriphase modems.

The prime source of system degradation, due to phase-lock loop performance is static phase error. Figure 5-28 plots the error rate degradation (in dB) as a function of this phase error and also the non-orthongonality phase error δ. There are several sources of phase error most of which can be calibrated out. Phase shift due to frequency uncertainty and phase error due to the finite loop signal-to-noise ratio are the fundamental error sources that cannot be readily calibrated.

The static phase error (at 4 f_o) due to frequency uncertainty is given by $\phi = \Delta f/k$ where Δf is the frequency uncertainty (at 4 f_o) and k is the loop gain. For the loop gains proposed to achieve a pull-in range of ± 6 MHz the phase errors are found to be 0.4 and 0.04°, for the narrow and wide loops, respectively. Before these phase errors can be used with Figure 5-28, they must be divided by four to convert them to the phase detection frequency (f_o). As may be seen from Figure 5-28, the loss due to this effect is a negligible few hundreds of a decibel.

The effect of finite loop signal-to-noise ratio can be roughly estimated from Figure 5-28 also. The loop phase jitter (at 4 f_o) measured in radians is given by

$$\sigma = 1/\sqrt{\mathrm{SNR}_{\mathrm{loop}}}$$

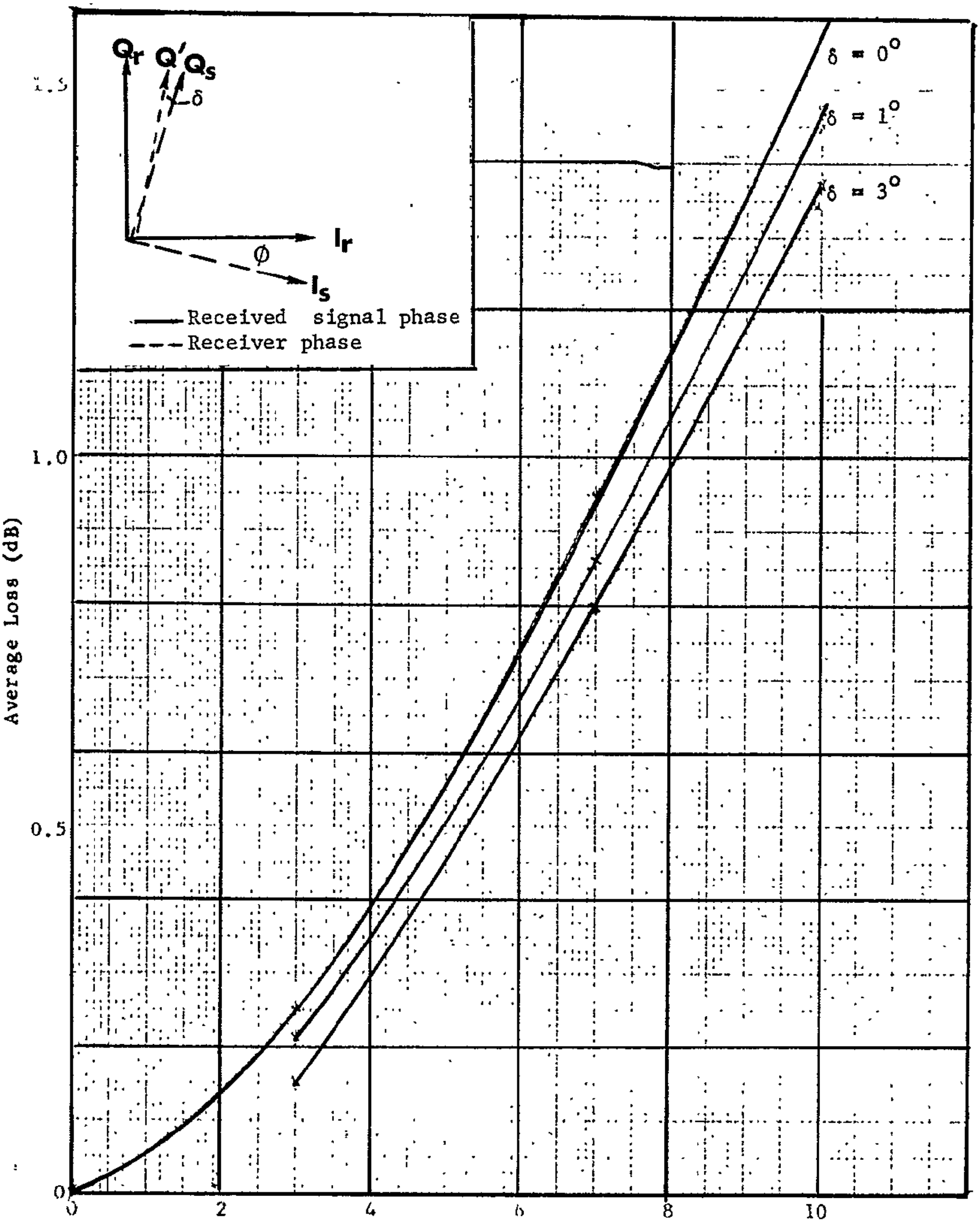


Figure 5-28 Average Degradation as a Function of Phase Error

For the wide and narrow loops we have 0.237 radians and 0.075 radians, respectively. Referred to the fundamental frequency, these errors yield degradations of 0.31 and 0.06 dB, respectively. Clearly, the narrow loop performs better but it should be cautioned that even the broad loop will perform better than the above worst case calculations indicate.

5.6.2.2 Carrier Reconstruction in the Presence of a Noisy Reference

In order to coherently detect a quadriphase-modulated signal, it is necessary to either reconstruct the carrier or to make use of a transmitted pilot carrier. This memo is an analysis of the carrier reconstruction section of a quadriphase demodulator.

Basically, the quadriphase (QPSK) signal may be described as

$$2S \cos\left[w_o t + \varphi(t) + \theta\right]$$

where S is the average power of the signal,

ω_o is the carrier radian frequency,

$\varphi(t)$ is the modulation (i.e., $\varphi(t)$ takes on one of the values $\pm \pi/4$, $\pm 3\pi/4$ radians for a period of T seconds),

and θ is a random phase angle uniformly distributed over $(0,2\pi)$.

Passing this signal through a nonlinear device and filtering so as to recover the fourth harmonic yields

$$K \cos\left(4\omega_o t + \pi + 4\,\theta\right)$$

where we see that the modulation has been removed and in its place a constant phase term is left. To reconstruct the carrier, one can divide down by four and subtract out $\pi/4$ radians and obtain $K_1 \cos(\omega_o t + \theta)$ which may be used as the coherent reference.

In a physical system the implementation is generally as indicated in Figure 5-29. The QPSK signal and additive gaussian noise with one-sided spectral density of N_o watts/Hz are passed through a bandpass filter sufficiently wide to pass at least the main lobe of the QPSK signal. The phase-locked loop (PLL) tracks the signal component at $4\omega_o$ from the $(\cdot)^4$ device and the reconstructed carrier is obtained from the VCO output at ω_o.

This analysis consists of deriving the signal-to-noise ratio at the input of the PLL and, knowing this, the average error rate for the detected QPSK signal.

Since each of the possible phases of the transmitted quadriphase signal are equally likely, we need consider only one of them in formulating the probability of making an error in the detection process. Furthermore, since the PLL is tracking a carrier at $4f_o$ we must divide its output signal by 4 yielding a reference signal at f_o. This divide-by-four operation has the advantage of also reducing the phase difference φ by 4. Figure 5-30 shows a phasor diagram of a received signal originally transmitted as

$$\sqrt{E} \cos (\omega_o t + \frac{\pi}{4} + \theta).$$

where E is the symbol energy

The reference signal used in coherently detecting this signal is of the form

$$K \cos(\omega_o t + \theta + \frac{\varphi}{4})$$

The four equally likely choices for symbol in each time period T results in

$$\log_2 4 = 2 \text{ bits of information/symbol}$$

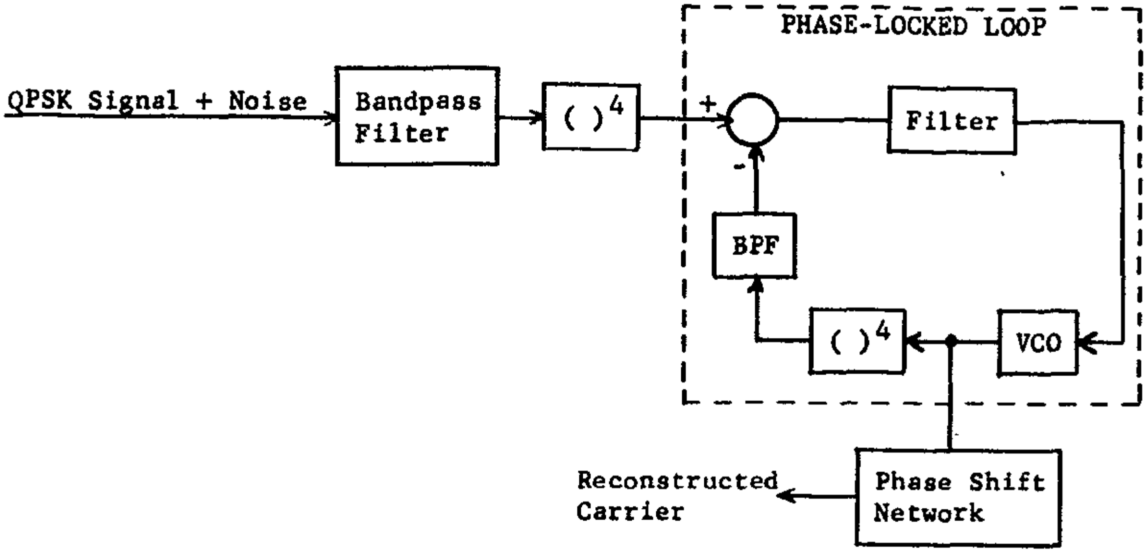

PHASE-LOCKED LOOP
QPSK Signal + Noise
Bandpass Filter
( )^4
+
-
Filter
BPF
( )^4
VCO
Reconstructed Carrier
Phase Shift Network

Fig. 5-29 Physical implementation of carrier reconstruction.

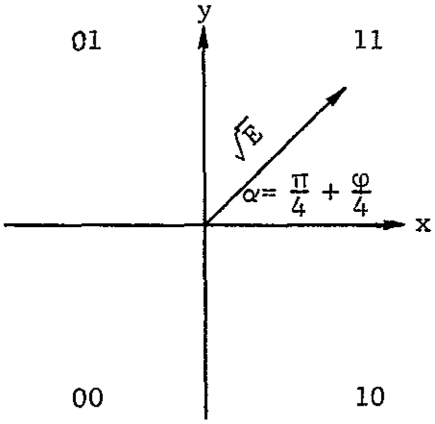

y
01
11
$\sqrt{E}$
$\alpha = \frac{\pi}{4} + \frac{\varphi}{4}$
x
00
10

Fig.5-30 Phasor diagram of received quadriphase signal.

The assignment of the symbol to each pair (x, y) of binary digits transmitted may be arbitrary and we will choose to do it as indicated in Figure 5-30. Here we see that the assignments are:

x	y	$\varphi(t)$
0	0	$\frac{-3\pi}{4}$
0	1	$\frac{+3\pi}{4}$
1	0	$\frac{-\pi}{4}$
1	1	$\frac{\pi}{4}$

Since the transmitted symbol may be considered the vector addition of the x and y components we will assume that the noise is also composed of x and y components. Each component of noise (n_x, n_y) will be assumed gaussian with zero mean and one-sided noise density N_o watts/Hertz. Further, we assume the two components to be statistically independent.

A hard decision as to which two bits were transmitted is made by determining in which quadrant the received signal vector lies. For the case shown in Figure 5-30, one bit will be in error if the received vector lies in quadrant II or IV and two bits will be in error if it lies in quadrant III. No error will result if it lies in the first quadrant. The conditional probability of a bit error when two bits (one symbol) are transmitted given an angle φ is

$$\text{Prob(bit error per 2 bits}/\varphi) = 1/2 \left\{ \text{erfc} \left(\sqrt{\frac{E}{N_o}} \cos \alpha \right) + \text{erfc} \left(\sqrt{\frac{E}{N_o}} \sin \alpha \right) \right\}$$

where $\text{erfc}\,(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-t^2}\, dt$

Using the ideas of conditional probability,

$$\overline{\text{Prob(bit error/bit transmitted)}} = \overline{\text{Prob(bit error}|\varphi)\text{Prob}(\varphi)}$$

$$= \int_{-\pi}^{\pi} \frac{\exp(\text{SNR} \cos \varphi)}{8\pi I_o(\text{SNR})} \left[\text{erfc} \left(\sqrt{\frac{E}{N_o}} \cos \alpha \right) + \text{erfc} \left(\sqrt{\frac{E}{N_o}} \sin \alpha \right) \right] d\varphi$$

where again, $\alpha = \frac{\pi+\varphi}{4}$

and E = symbol energy
= 2 bit energy = $2E_b$

The above equation cannot be evaluated in closed form, however; a computer program was written to numerically integrate it with various values of SNR and E_b/N_o.

For the sake of simplicity, SNR_L was taken to be of the form

$$SNR_L = \beta \; E_b/N_o$$

where β = constant

Curves of bit error probability vs input E_b/N_o with several values of β are shown in Figure 5-31. It will be noted that for $\beta = 5$ there is at most a degradation in performance of 0.2 dB (i.e. the transmitted power must be increased by at most 0.2 dB to provide the error performance achieved by a noiseless reference). The portions of the curves drawn broken are areas where the PLL is not in lock and should not be used for system operating points.

In physical systems one would want to keep β as large as possible to reduce performance degradation. This is usually done by lowering B_L. A compromise must be made here, however, because reducing B_L increases the acquisition time of the loop. If E_b/N_o is small, it is to be expected that the PLL will lose lock and be required to reacquire before coherent reception is possible. One then uses the largest B_L possible consistent with low system degradation.

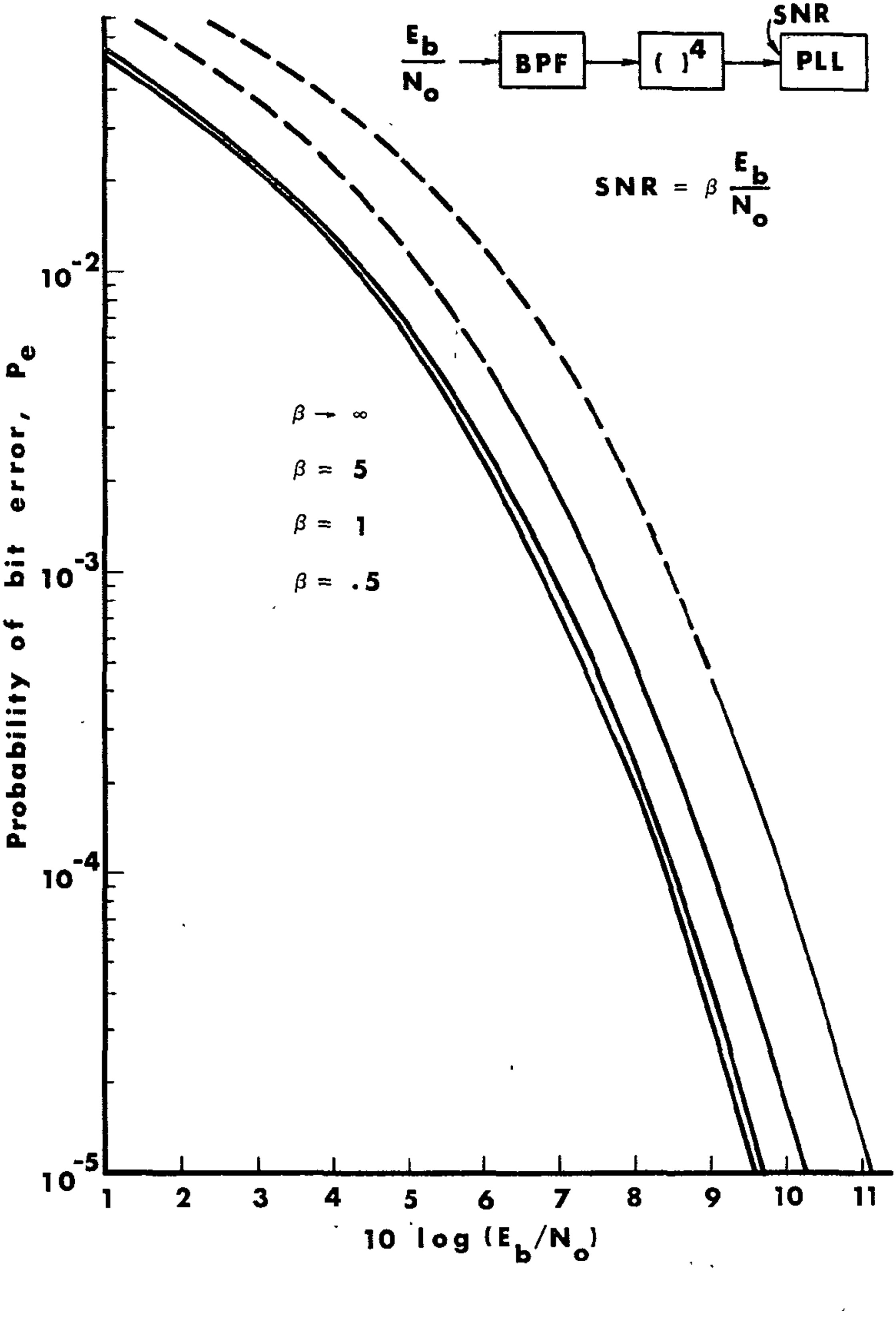


Figure 5-31 Probability of Bit Error vs E_b/N_o

Figure 5-32 shows a similar curve, due to Lindsey,[1] for BPSK.

5.6.2.3 Effects of Oscillator Noise on Quadriphase Performance

When operating at frequencies in the tens of gigahertz it is extremely difficult to obtain highly stable local oscillators. The problem is further compounded when a requirement is made for tens of milliwatts of power to be generated efficiently by the oscillators (as in a transmitter up-converter).

The avalanche diode oscillator is capable of these power levels, but has poor frequency stability (i.e., the FM noise spectrum of this oscillator has large components over much of the range fc ± 20 MHz [4)]. This FM noise spectrum produces a phase error in the tracking loop (see Figure 5-33) which degrades the loop performance and hence increases the average probability of bit error.

It is the purpose of this note to determine the degradation of quadriphase and biphase error rate performance due to the noisy oscillator. It should be noted that although Figure 5-33 shows the receiver L.O. to be the noisy one, in actual fact the noisy oscillator is in the transmitter up-converter. However, since we are not concerned with non-linearities in the transmission path we can place this noise source as it is shown in the figure and obtain the same degrading effect.

5.6.2.3.1 Total Loop Phase Error

The total mean square phase error of the phase locked loop, σ^2_T, is composed of two parts - that due to the FM noise of the oscillator, σ^2_{FM}, and that due to thermal noise in the receiver, σ^2_N. A properly designed PLL is operated near phase lock and in this condition may be considered a linear device. A rule of thumb[2] for this condition is $\sigma^2_T < \frac{1}{4}$.

---

[1] Lindsey, W. C., "Phase-shift-keyed Signal Detection with Noisy Reference Signals," IEEE Trans. on Aerospace and Electronic Systems, Vol. AFS-2, pp. 393-401; July, 1966.

[2] Loc. Cit., pg 5-78

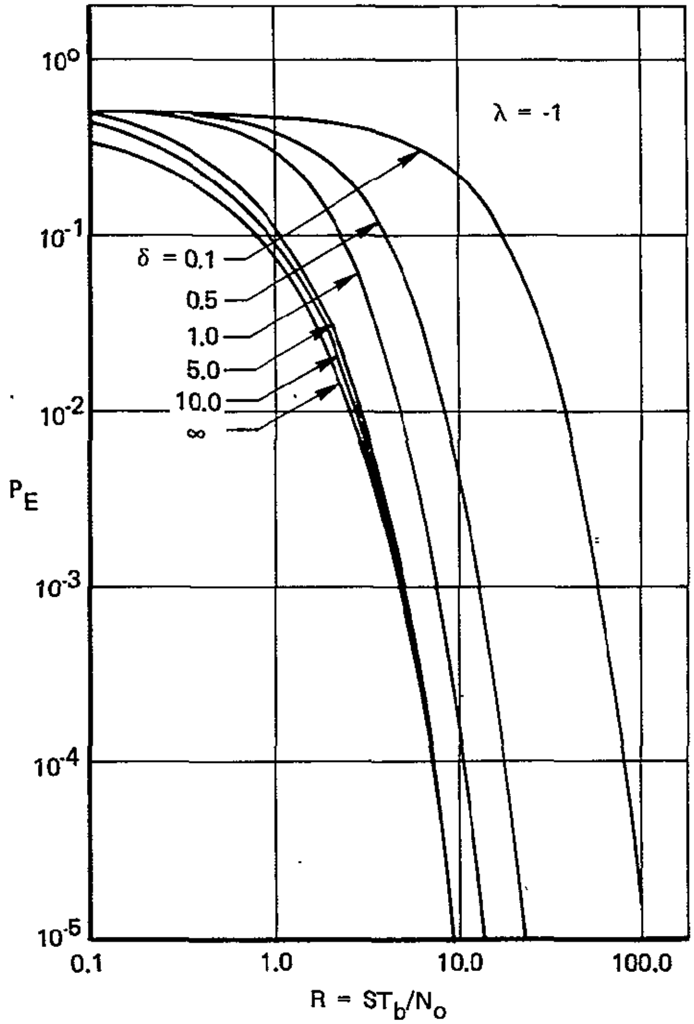
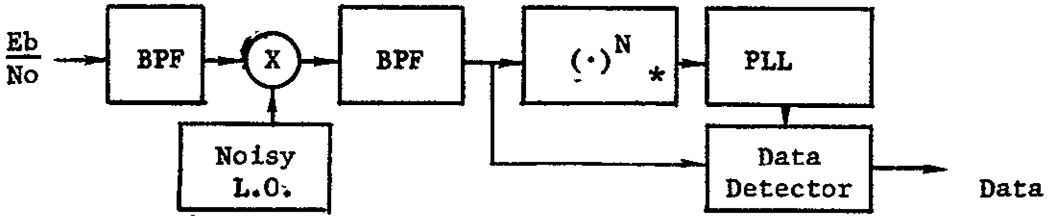


Figure 5-32 Error Probability of System Which Uses a Squaring Loop


Eb
No
BPF
X
BPF
(·)N *
PLL
Noisy L.O.
Data Detector
Data

* N = 2,4

Figure 5-33 System Block Diagram

Since the FM noise and thermal noise are clearly independent and the PLL is assumed linear, we may invoke the superposition theorem and show that the mean square phase error components are additive.

$$\sigma^2_T = \sigma^2_{FM} + \sigma^2_N \qquad (1)$$

The signal-to-noise ratio in the PLL, SNR_L, is

$$SNR_L = \frac{1}{\sigma^2_T} = \frac{1}{\sigma^2_{FM} + \sigma^2_N} \qquad (2)$$

A calculation of σ^2_{FM} for a specific avalanche diode oscillator is performed in the Appendix. Included is the mathematical model and equations for determining σ^2_{FM} for all oscillator types which fit the fairly broad characteristics of the model.

5.6.2.3.2 Probability of Bit Error for QPSK

Reference [2] has derived the probability of bit error vs Eb/No for various values of the parameter β ($SNR_L = \beta$ Eb/No) for the quadriphase case where $\sigma^2_{FM} = 0$. For notational convenience we will call

$$\frac{1}{\sigma^2_N} = SNR^*_L \qquad (3)$$

Eq (2) can now be written as

$$SNR_L = \frac{1}{\sigma^2_T} = \frac{SNR^*_L}{1 + SNR_L \times \sigma^2_{FM}} = \frac{\beta \ Eb/No}{1 + \beta \ Eb/No \ \sigma^2_{FM}} \qquad (4)$$

Substituting this modified value of SNR_L into Eq (42) of [2] yields

(2) Loc. Cit. pg 5-78

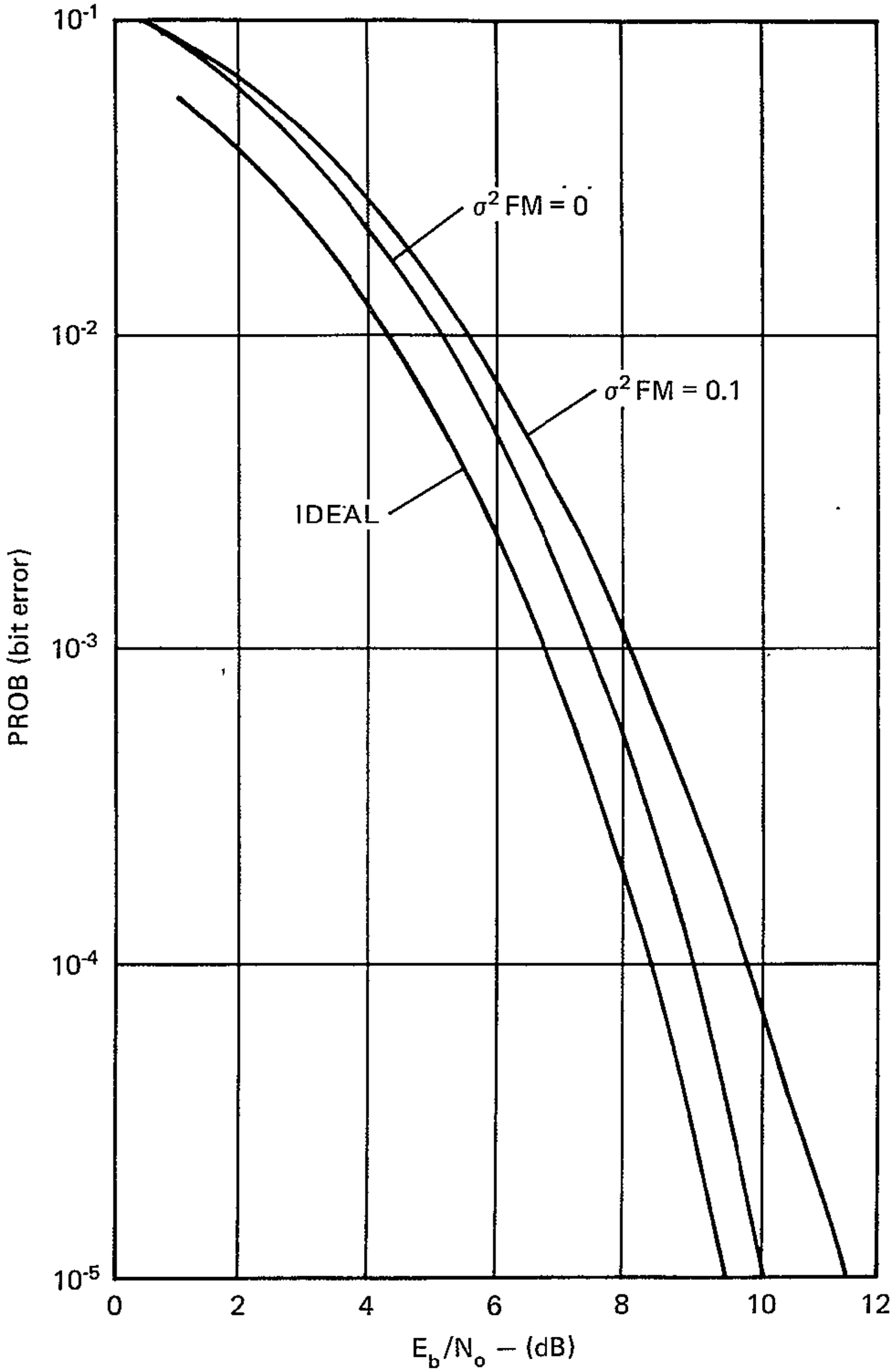


Figure 5-34 Quadriphase BER Performance with Noisy Oscillator - (β=1)

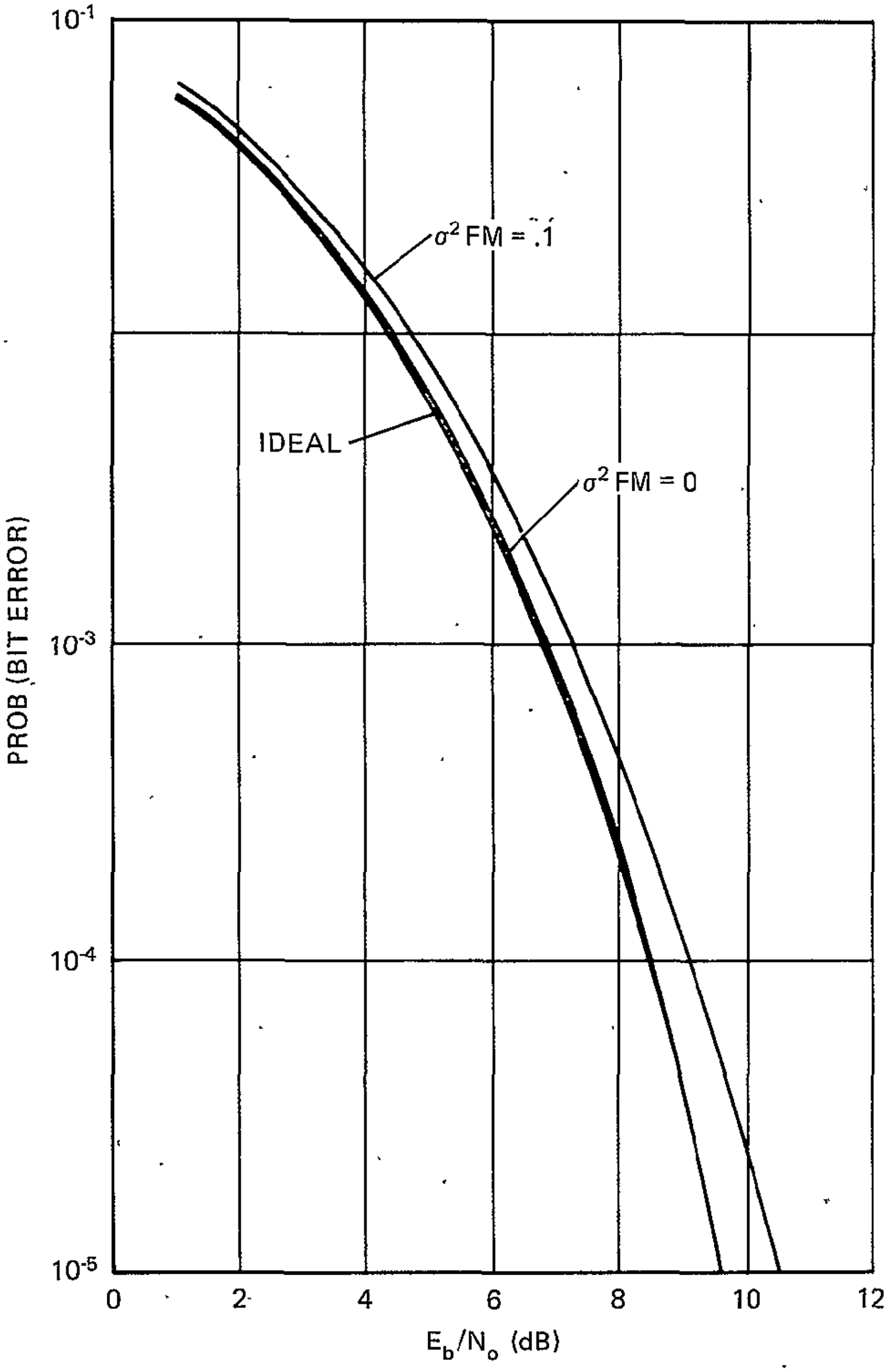


Figure 5-35 Quadriphase BER Performance with Noisy Oscillator (β=5)

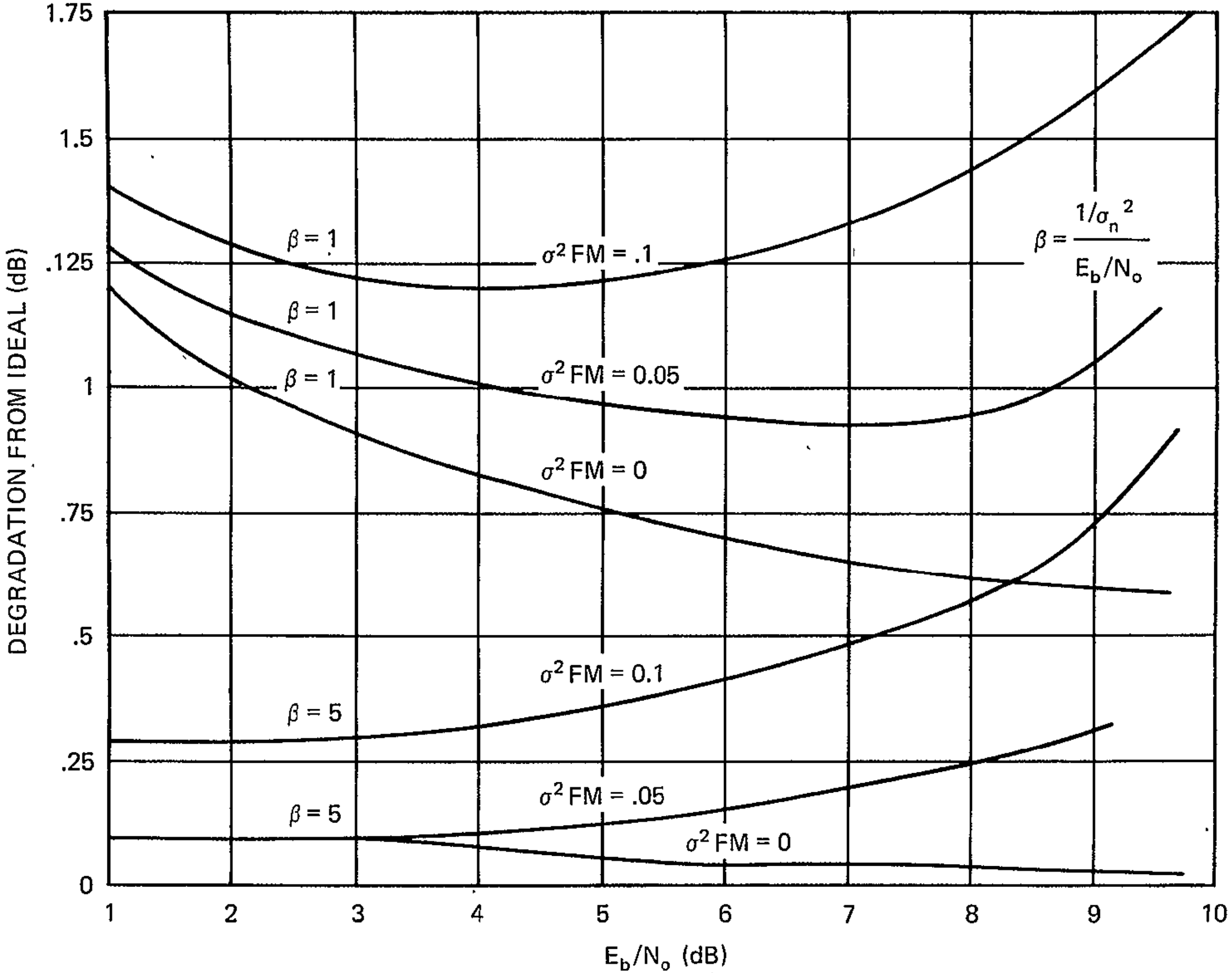


Figure 5-36 QPSK Performance Degradation (From Ideal) vs Eb/No for Various Values of β and σ^2 FM

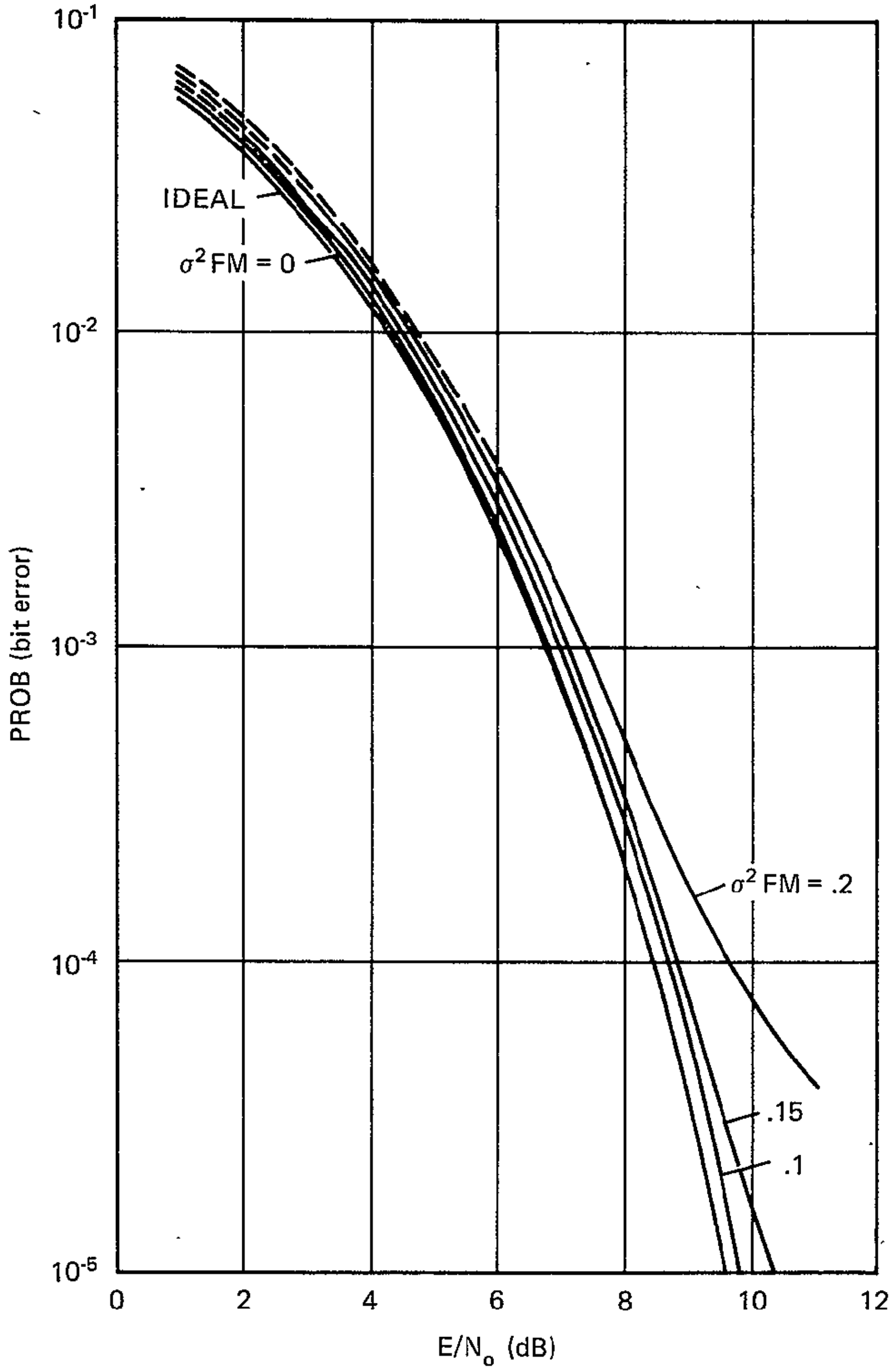


Figure 5-37 Biphase Error Rate Performance as a Function of E/No and σ^2 FM. (β=5)

$$P\ (\text{bit error}) = \int_{-\pi}^{\pi} \frac{\exp\ (SNR_L \cos \phi)}{8\pi\ Io\ (SNR_L)} \left[\text{erfc}\ \left(\sqrt{\frac{2Eb}{No}} \cos\alpha \right) + \text{erfc}\ \left(\sqrt{\frac{2\ Eb}{No}}\ \text{Sin}\ \alpha \right) \right] d\ \phi \qquad (5)$$

where $\alpha = \frac{\pi + \phi}{4}$

$Io\ (.) \triangleq$ Modified Bessel Function

The integration is carried out numerically using the Gaussian method with the results shown in Figures 5-34 and 5-35. Figure 5-36 shows the degradation in graphical from for β equal to 1 and 5. For finite values of β it is seen that the degradation caused by $\sigma^2_{FM} = 0.1$ increases with E_b/No. The FM noise causes the probability of bit error to asymptotically approach some non-zero value as $E_b/No \rightarrow \infty$.

5.6.2.3.3 Probability of Bit Error for BPSK

In an analogous manner we may show that the average probability of bit error for the biphase case is

$$P(\text{bit error}) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\exp(SNR_L \cos\phi)}{I_o\ (SNR_L)}\ \text{erfc}\ \sqrt{\frac{E}{N_o}} \cos \frac{\phi}{2}\, d\phi \qquad (6)$$

This expression was numerically integrated and the results are plotted as a function of E/No in Figure 5-37 with σ^2_{FM} as a parameter. Degradation in performance is plotted vs. E/No in Figure 5-38.

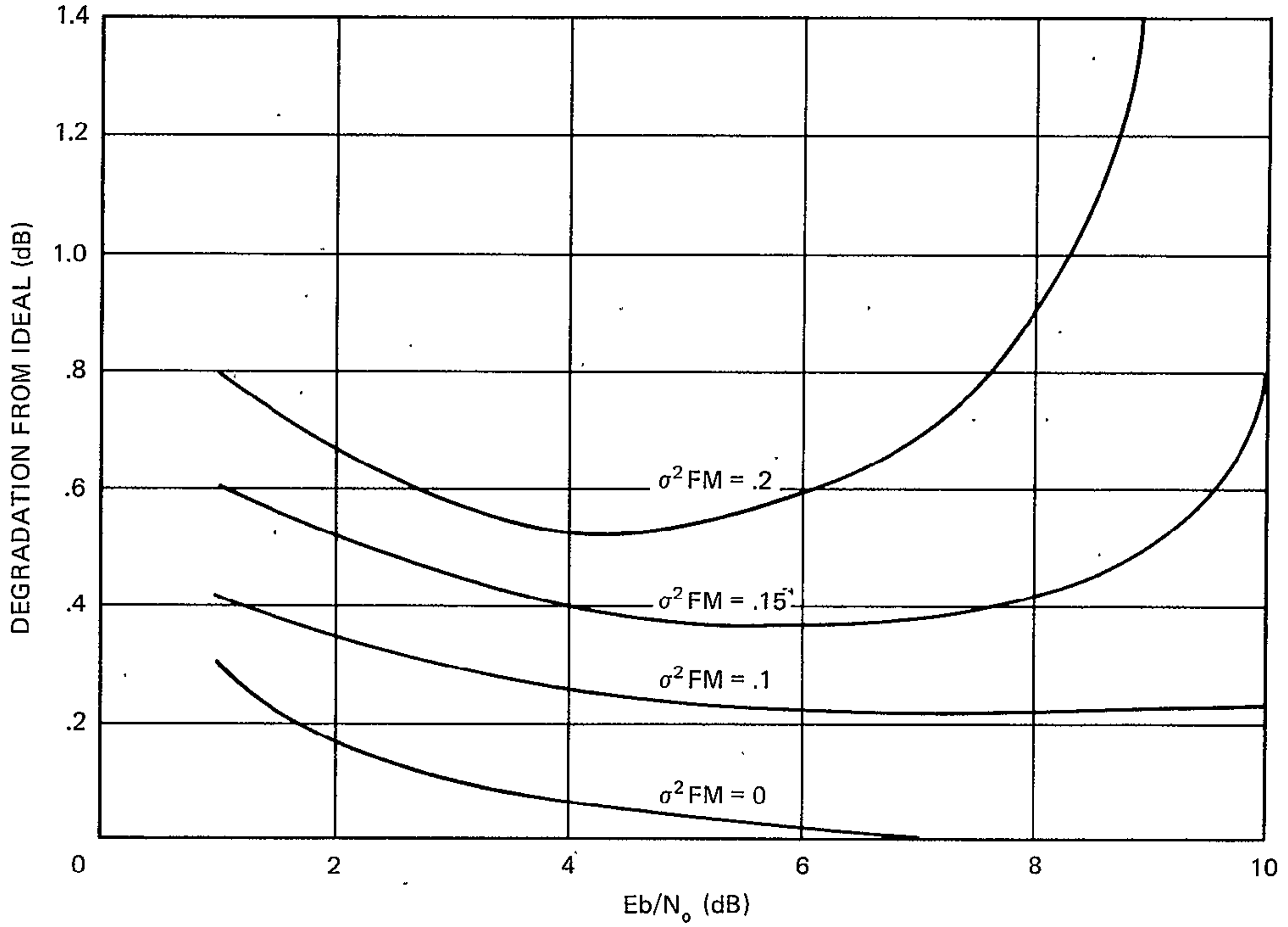


Figure 5-38 Biphase BER Degradation from Ideal (β=5)

5.6.2.3.4 Calculation of σ^2 FM

Without too much difficulty it is possible to show[1] that the mean square phase error of the PLL due to the FM noise of the local oscillator is

$$\sigma^2_{FM} = \int_0^\infty \frac{1}{2} \left(\frac{\Delta f}{f_m}\right)^2 \frac{(f_m/f_n)^4}{1 + (f_m/f_n)^4} \, df_m \qquad (A-1)$$

where Δf is the frequency deviation (peak) (Hz) at f_m of the FM spectrum f_m is the frequency offset from carrier frequency (Hz) f_n is related to the PLL bandwidth (f_n = .3x (loop bandwidth) (assumes second order loop with .7 damping factor)

To use (A-1) one must know the FM spectrum of the LO (scaled to account for the x4 or x2 operation prior to entering the PLL) and the equivalent noise bandwidth of the PLL.

A typical application may have an avalanche diode oscillator (ADO) as the LO. As a free running oscillator the ADO has a very noisy output. To reduce this FM noise it may be injection-locked to a more stable source of considerably less power [4] like a crystal oscillator-harmonic generation chain. In general, once the ADO is locked to the low power source its output FM spectrum close to the carrier approaches that of the low power source and as f_m increases the noise spectrum tends asymptoically to that of the free running ADO. A typical example is given in Figure 5-39.

(4) Lo. Cit, pg 5-78

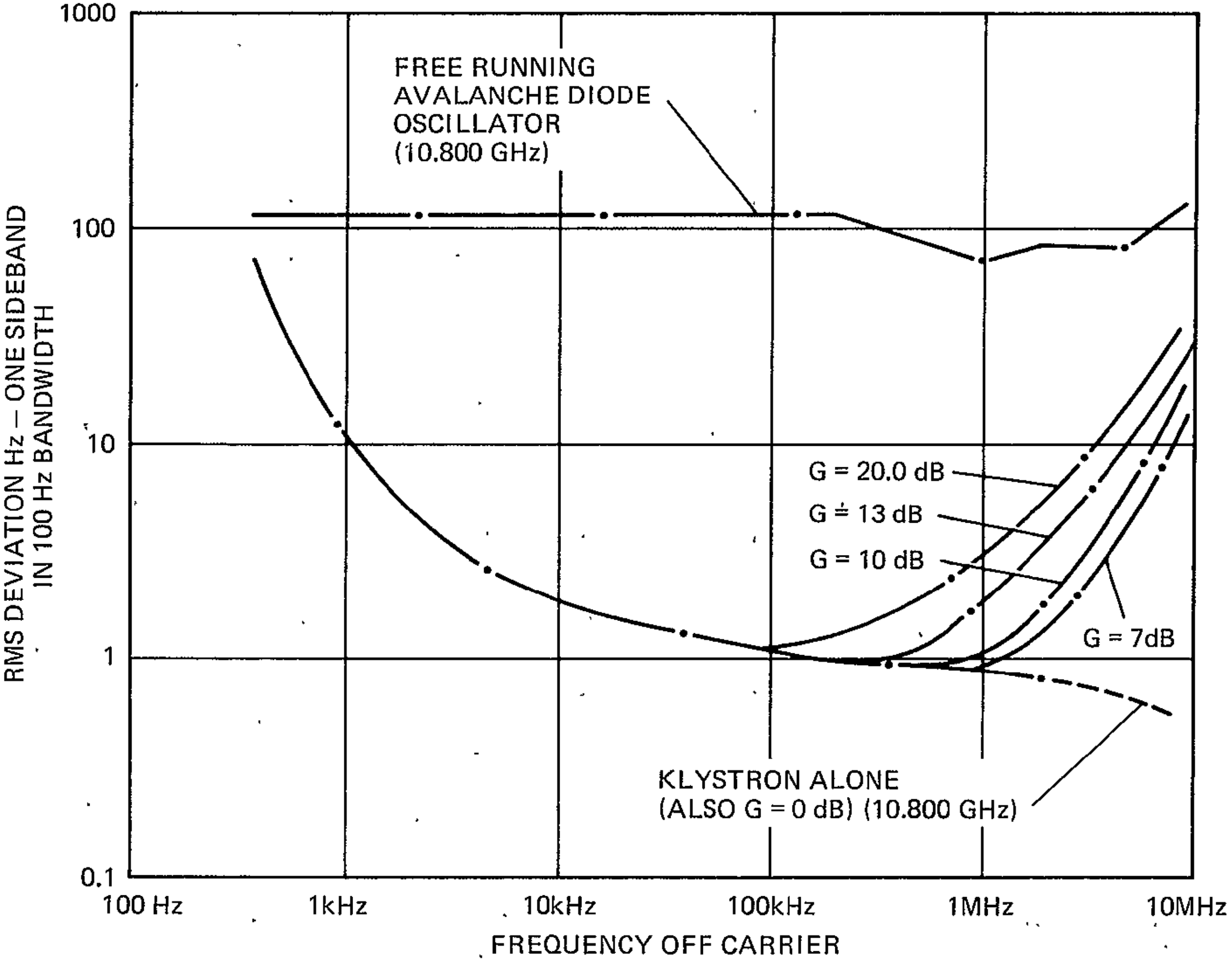


Figure 5-39 Phase Noise for Various Oscillators

A generalized worst case solution of (A-1) can be obtained by drawing straight line segments above the actual spectral components as shown in Figure 5-40. The integral in (A-1) is now evaluated as

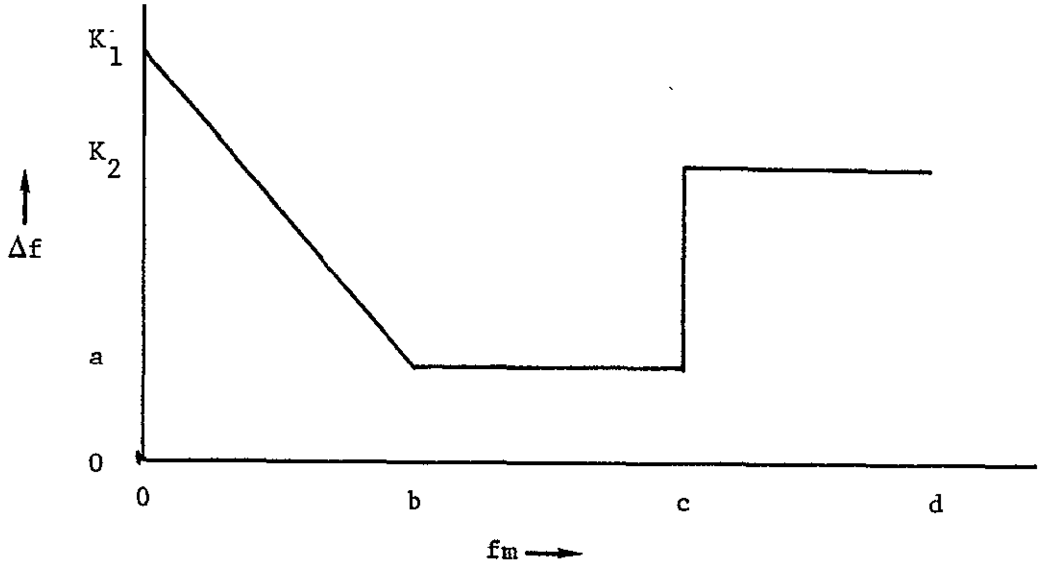


Figure 5-40 Straight Line Approximation to Curves of Figure A-1.

$$\sigma^2_{FM} = 1/2 \int_o^\infty \left(\frac{\Delta f}{f_m}\right)^2 \frac{(f_m/f_n)^4}{1 + (f_m/f_n)} 4 \; df_m$$

$$= 1/2 \int_o^b \left[K_1 - \left(\frac{K_1 - a}{b}\right) f_m\right]^2 f_m^{-2} \frac{(f_m/f_n)^4}{1 + (f_m/f_n)} 4 \; df_m \qquad \text{(A-2)}$$

$$+ 1/2 \int_b^c \left(\frac{a}{f_m}\right)^2 \frac{(f_m/f_n)^4}{1 + (f_m/f_n)^4} df_m + 1/2 \int_c^d \left(\frac{K_2}{f_m}\right)^2 \frac{(f_m/f_n)^4}{1+(f_m/f_n)^4} df_m$$

For f_m in the range (0, b) the integral in (A-2) would be difficult to evaluate in closed from so we obtain the maximum value of the product $f \cdot f_m$ in this interval which is

$$\Delta f \cdot f_m \Big|_{max} = \frac{K_1^2 \; b}{4 \; (K_1 - a)} \approx \frac{K_1 b}{4} = R_1 \text{ if } a \ll K_1 \qquad \text{(A-3)}$$

This gives an upper bound on the first integral in A-2 of

$$\int \left[K_1 - \frac{(K_1 - a)}{b} f_m \right]^2 f_m^{-2} \frac{(f_m/f_n)^4}{1 + (f_m/f_n)^4} df_m$$

$$\leq \int \left(\frac{R_1}{f_m} 2 \right)^2 \frac{(f_m/f_n)^4}{1+(f_m/f_n)^4} df_m \qquad \text{(A-4)}$$

Using the change of variable

$$X = f_m/f_n \qquad \text{(A-5)}$$

it is now possible to obtain an upper bound on (A-2) in terms of a sum of the following two integral types:

$$\int_{x1}^{x2} \frac{dx}{1+x4} = \left\{ \frac{\sqrt{2}}{8} \ln \left[\frac{x^2 + \sqrt{2}x + 1}{x^2 - \sqrt{2}x + 1} \right] + \frac{\sqrt{2}}{4} \tan^{-1} \left(\frac{\sqrt{2}x}{1-x^2} \right) \right\} \Big|_{x1}^{x2}$$

and (A-6)

$$\int_{x1}^{x2} \frac{x^2 dx}{1+x4} = \left\{ -\frac{\sqrt{2}}{8} \ln \left[\frac{x^2+\sqrt{2}x + 1}{x^2-\sqrt{2}x + 1} \right] + \frac{\sqrt{2}}{4} \tan^{-1} \left(\frac{\sqrt{2}x}{1-x^2} \right) \right\} \Big|_{x1}^{x2}$$

Combining (A-2) through A-6) we obtain

$$\sigma^2_{FM} \leq \frac{\sqrt{2}}{16} fn \left\{ \left(\frac{R_1^{\,2}}{f_n^{\,2}} + a^2 \right) \ln \left[\frac{\beta^2 + \sqrt{2} \; \beta+1}{\beta^2 - \sqrt{2} \; \beta+1} \right] +2 \left(\frac{R_1^{\,2}}{f_n^{\,2}} - a^2 \right) \tan^{-1} \left(\frac{\sqrt{2} \; \beta}{1-\beta^2} \right) \right.$$

$$+ \left[K_2^{\,2} - a^2 \right] \left[\ln \left(\frac{\gamma^2+\sqrt{2} \; \gamma+1}{\gamma^2-\sqrt{2} \; \gamma+1} \right) -2 \tan^{-1} \left(\frac{\sqrt{2}\gamma}{1-\gamma^2} \right) \right] \qquad \text{(A-7)}$$

$$\left. + K_2^{\,2} \left[-\ln \left(\frac{\delta^2 + \sqrt{2}^{\delta} + 1}{\delta^2 - \sqrt{2}^{\delta} + 1} \right) +2 \tan^{-1} \left(\frac{\sqrt{2}\delta}{1-\delta^2} \right) \right] \right\}$$

where $\beta' \triangleq b/f_n$,

$\gamma \triangleq c/f_n$,

and $\delta \triangleq d/f_n$,

For free running oscillators which are relatively stable the FM noise spectrum may be considered to be that of Figure 5-40 with $K_2 = 0$ and $c, d \longrightarrow \infty$. Equation (A-7) still holds for this case if the parameter modifications are used.

---

1. Viterbi, A. J., Principles of Coherent Communication, New York, McGraw-Hill; 1966.

2. Sherman, R. J. "Quadriphase-Shift-Keyed Signal Detection with Noisy Reference Signal", Philco-Ford Corp. Communication Sciences Department Tech Memo #160, Aug. 1969. Also published under same title in IEEE Eascon '69 Convention Record pp 45-52.

3. Kurokawa, K. "Noise in Synchronized Oscillators", IEEE Trans. Microwave Theory and Techniques. Vol MTT 6-16, No. 4, April 1968 pp 234-240.

4. Hines, M. E., et al, "FM Noise Suppression of an Injection Phase-Locked Oscillator", IEEE Trans. Microwave Theory and Techniques. Vol. MTT-16, No. 9, Sept. 1968, pp 638-742.

5. Shaw, R. C. and H. L. Stover, "Phase-Locked Avalanche Diode Oscillators", Proc. IEEE, Vol. 54, No. 4, April 1966 pp 710-711.

6. Ashley, J. R., and F. M. Palka, "Measured FM Noise Reduction by Injection Phase Locking", Proc. IEEE Vol. 58, No. 1, January, 1970. pp 155-157.

5.6.3 Matched Filter Data Detection

A matched data detector of the matched filter variety is required at the coherent demodulator output in order to realize fully the advantages of coherent PSK modulation.

The matched filter, a technique widely discussed in the literature, is a circuit wherein the receiver input signal is "matched" to the filter transfer function characteristics in such a manner as to maximize the filter output signal-to-noise ratio, thereby enhancing the decision process. It can be shown that the matched filter produces the optimum performance independent of the information bandwidth. This is true to the point of intersymbol interference caused by insufficient bandwidth.

At low bit rates, the matched filter employs a linear integrate-and-dump technique shown in Figure 5-41, wherein the coherent phase detector signal plus noise output is integrated over a bit period, the integrator output is sampled and the filter memory is "dumped" or reset to zero in readiness for the next data bit.

In order to achieve efficient matched filter operation, it is necessary to sample the filter at precisely the end of a bit interval and then to dump the filter energy in order to produce a zero memory at the start of the next bit period. The sampling command is generated in the proposed design by a bit clock which is accurately phased to the incoming data. Dumping is accomplished by using two matched filters where one filter integrates over a bit period while the previous bit information is dumped from the other filter. The integrate/dump roles of the filters reverse on each bit, and the timing system is used to select the proper filter for each decision period. If the integration period is not in phase with bit period, there will be crosstalk between adjacent bits.

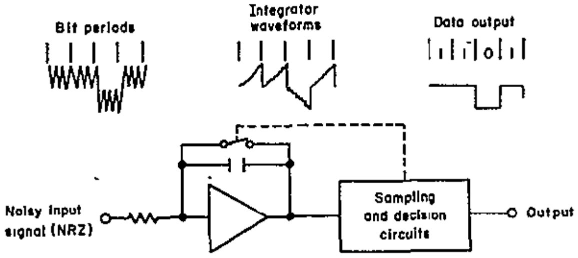

Bit periods
Integrator waveforms
Data output
1 1 1 0 1
Noisy input signal (NRZ)
Sampling and decision circuits
Output

Figure 5-41 Matched Filter CKT

As the bit rate increases to above 100 Mbps, the matched filter design and operation can become the determining factor of system operation. At low frequency rates, the matched filter is an RC circuit or filter circuit followed by a dump circuit which short circuits a capacitor comprising the output impedance of the matched filter. At higher frequency rates, the problems of "moving" capacitance stored charge by the dump mechanism becomes difficult and new types of filters, shown in Figure 5-42 such as transversal filters, or 2-pole Butterworth filters, using series decision-switched gates can be used. Such gates are associated with very wide band operational amplifiers which must have bandwidths from DC to a frequency high enough to pass all principal components. In all cases no energy dumping is used so that some intersymbol interference will exist. For each filter type there exists an optimum choice of a 3-dB cutoff frequency, f_c, relative to the symbol duration, T.

A block diagram of a matched filter system following a coherent PSK demodulator is shown in Figure 5-43 .

As illustrated in Figure 5-43 alternate matched filters are used in order to provide a reasonable period for "dumping" the integrating capacitors. The output selector alternately selects the filter which is integrating and sampling. The sampling and dump commands are generated from the data clock. Timing signals are shown in Figure 5-44 for the commands.

To establish bit timing, for the sampling of the output of the matched filters, the output of the coherent demodulator, as shown in Figure 5-43 is conditioned and filtered to compensate for variations in amplitude, baseline offset and bandwidth. Passing the signal through a nonlinear element (such as a zero-slicer and/or rectifier) develops a frequency component at bit rate. The phase of this component is compared with the phase of pulses from the voltage-controlled oscillator (VCO). The error output drives the VCO to maintain lock with the input signal. Appropriate loop compensation (filtering and amplification) permits

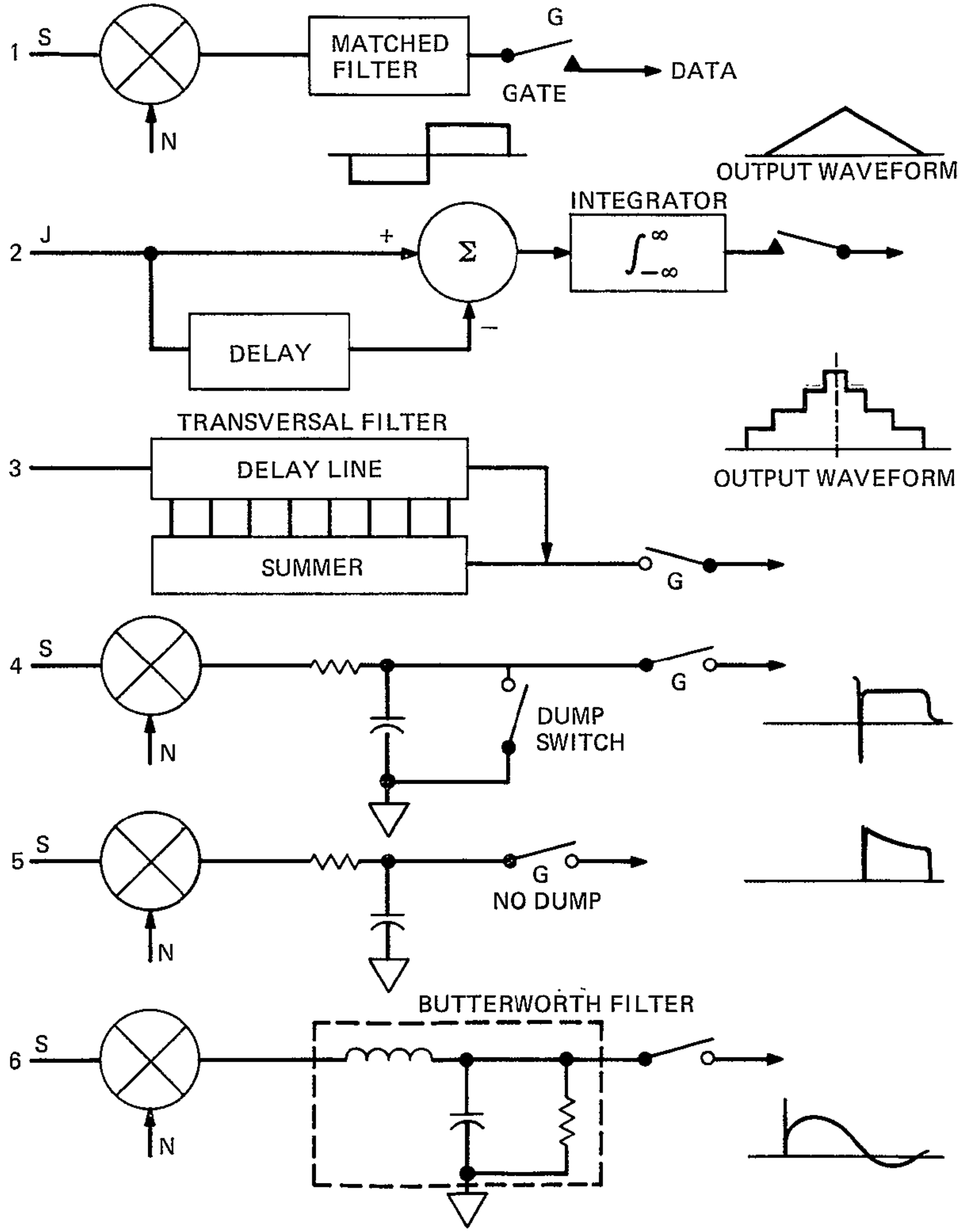

1
S
N
MATCHED FILTER
G
GATE
DATA
OUTPUT WAVEFORM
2
J
+
Σ
−
DELAY
INTEGRATOR
$\int_{-\infty}^{\infty}$
3
TRANSVERSAL FILTER
DELAY LINE
SUMMER
G
OUTPUT WAVEFORM
4
S
N
DUMP SWITCH
G
5
S
N
G
NO DUMP
6
S
N
BUTTERWORTH FILTER

Figure 5-42 Matched Filters

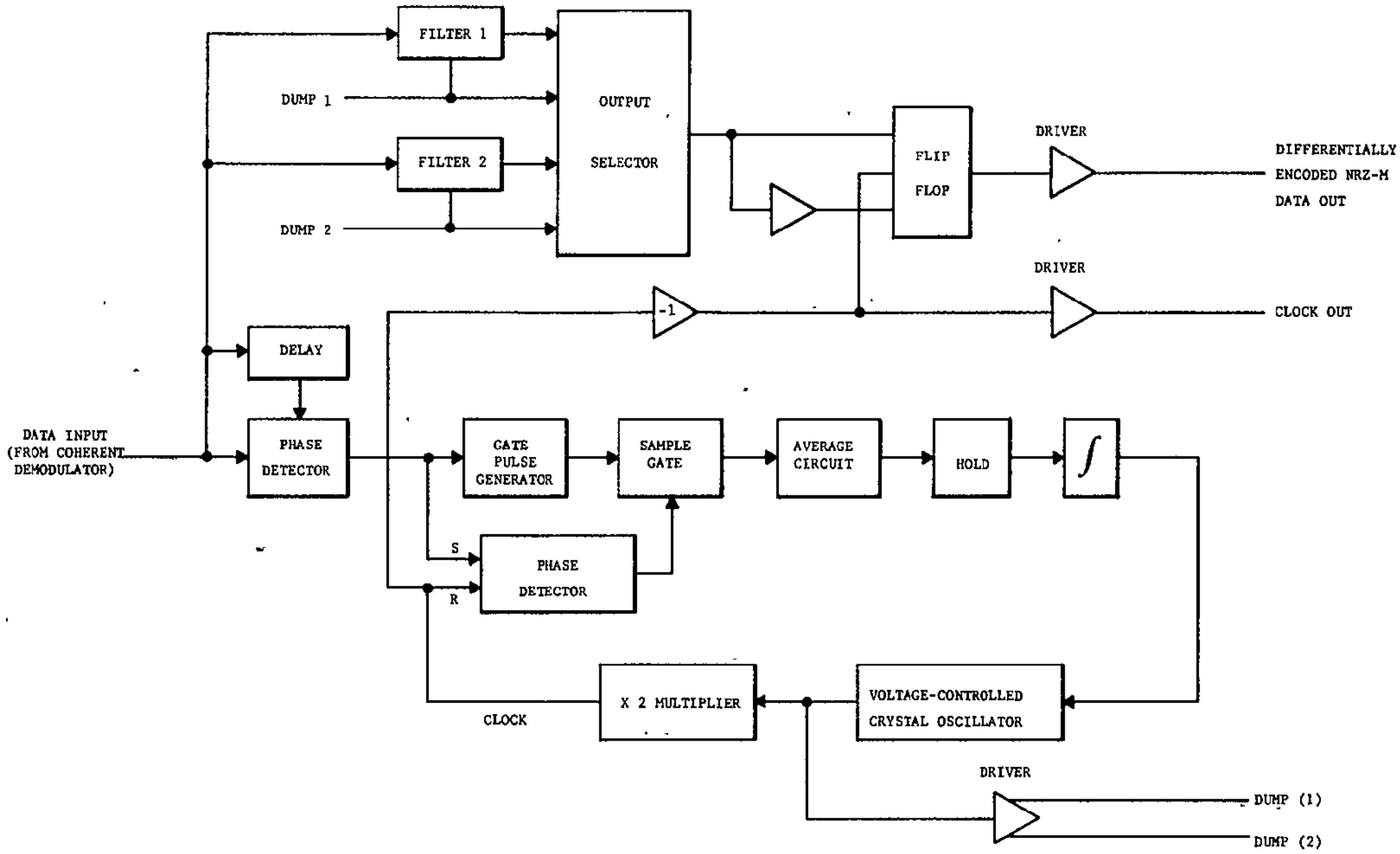

FILTER 1
DUMP 1
FILTER 2
DUMP 2
OUTPUT SELECTOR
FLIP FLOP
DRIVER
DIFFERENTIALLY ENCODED NRZ-M DATA OUT
DRIVER
-1
CLOCK OUT
DELAY
DATA INPUT (FROM COHERENT DEMODULATOR)
PHASE DETECTOR
GATE PULSE GENERATOR
SAMPLE GATE
AVERAGE CIRCUIT
HOLD
∫
S
R
PHASE DETECTOR
X 2 MULTIPLIER
VOLTAGE-CONTROLLED CRYSTAL OSCILLATOR
CLOCK
DRIVER
DUMP (1)
DUMP (2)

Figure 5-43 Matched Filter Detector

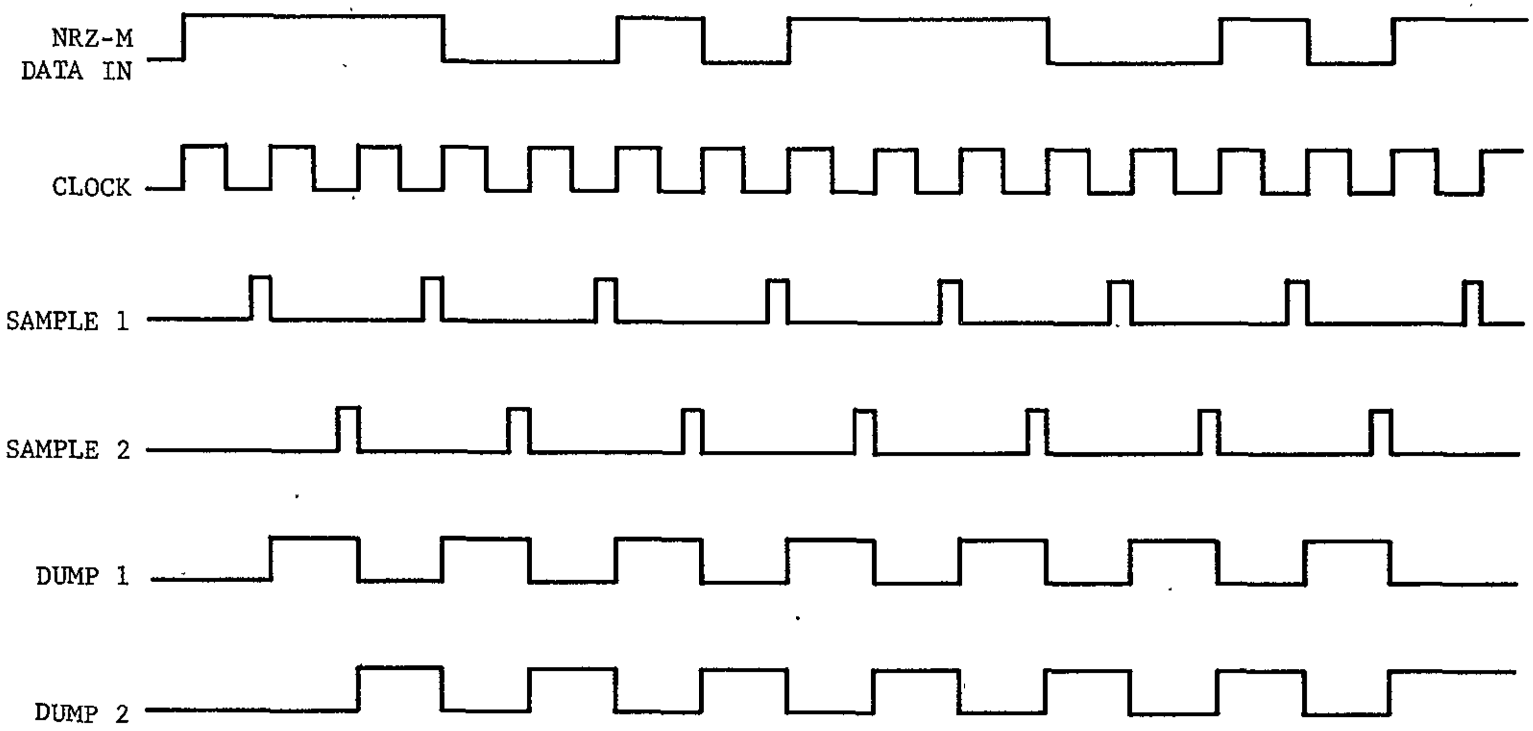

NRZ-M
DATA IN
CLOCK
SAMPLE 1
SAMPLE 2
DUMP 1
DUMP 2

Figure 5-44 Matched Filter Sample Dump Timing Diagram

the loop to acquire phase-lock, allows optimum signal tracking under adverse conditions, and maintains phase-lock even in the event of temporary loss of data. Acquisition and tracking ranges of ± 10% of the nominal bit rates are typical.

5.6.3.1 Data Waveforms

The non-return-to-zero (NRZ) waveform developed at the output of the matched filter and shown in Figure 5-45 , uses the bandwidth efficiently, since the entire bit period contains signal information. Timing information is indicated by signal transitions of binary ones and zeros, or zero-crossings of the waveform's average value. The lack of near-dc response in the signal path will cause shifts in the signal's baseline for period of few zero crossings or low transition density. If this baseline-shifting is not compensated, it will seriously affect synchronization and regeneration of bits.

Bi-phase waveform compensates for the lack of dc response at the expense of increasing the total bandwidth required. A zero-crossing is assured for each bit period, and the effect of transition density upon the average signal value is minimized. The entire bit period contains signal information.

Return-to-zero (RZ) waveform transition density depends upon the quantity of binary "ones" transmitted. This method of signal coding wastes 50% of its bandwidth - because one-half of its bit period does not contain signal information - resulting in a theoretical 3-dB loss in bit detection performance compared to NRZ and bi-phase, which are more efficient.

Figure 5-46 compares bit error rate versus signal to noise ratio for the three waveforms described above.

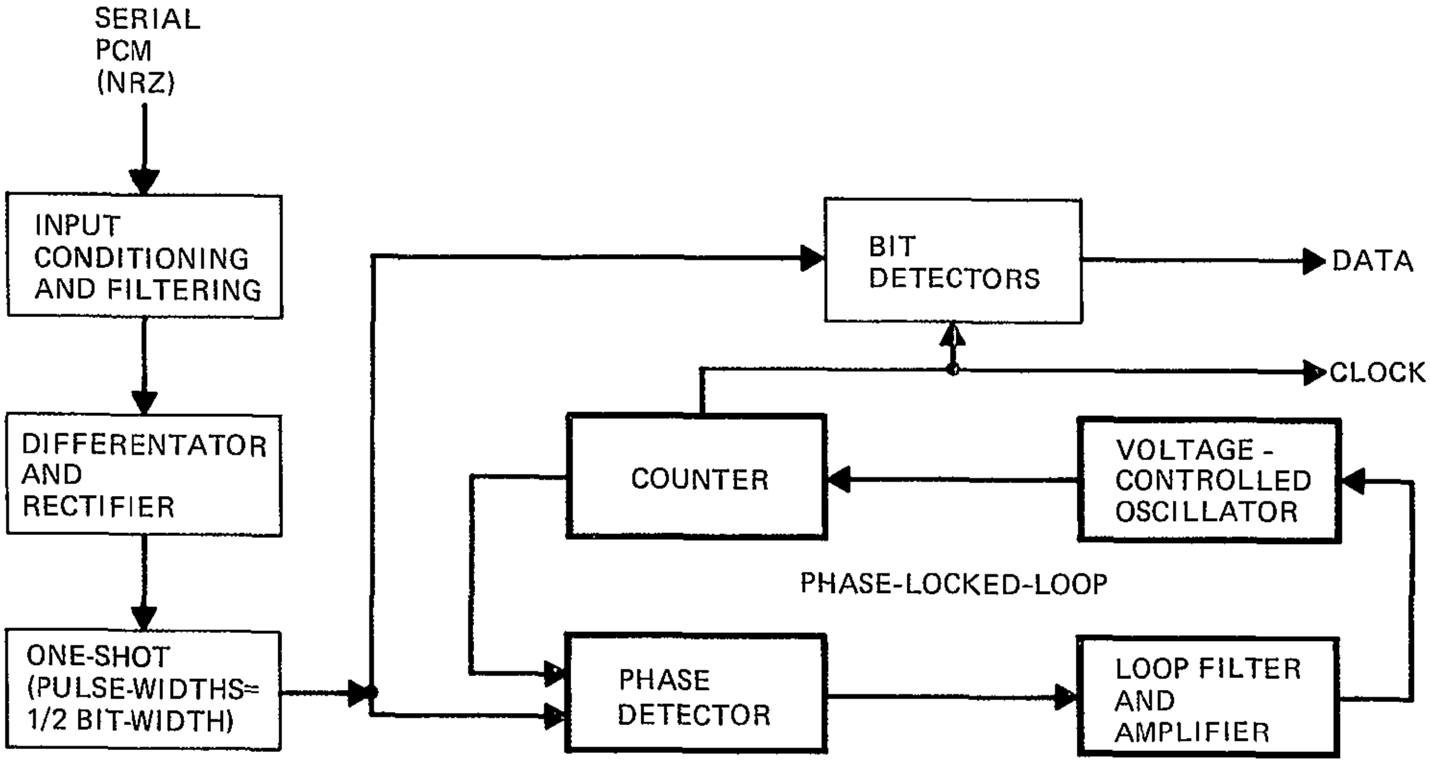
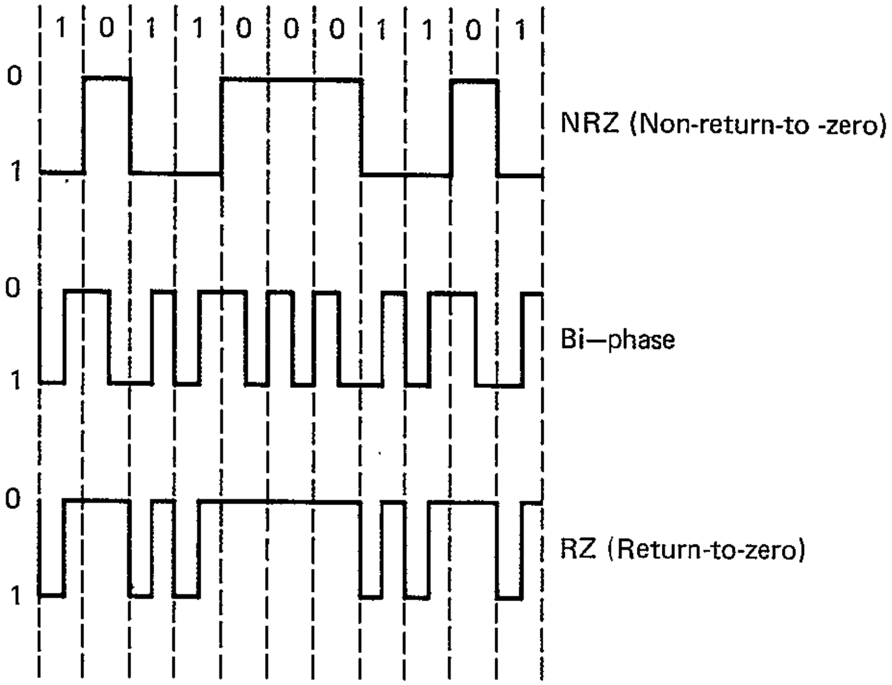

SERIAL
PCM
(NRZ)
INPUT
CONDITIONING
AND FILTERING
DIFFERENTATOR
AND
RECTIFIER
ONE-SHOT
(PULSE-WIDTHS≈
1/2 BIT-WIDTH)
BIT
DETECTORS
DATA
CLOCK
COUNTER
VOLTAGE -
CONTROLLED
OSCILLATOR
PHASE-LOCKED-LOOP
PHASE
DETECTOR
LOOP FILTER
AND
AMPLIFIER

Figure 5-45 Basic Bit Synchronizer


1 0 1 1 0 0 0 1 1 0 1
0
1
NRZ (Non-return-to -zero)
0
1
Bi—phase
0
1
RZ (Return-to-zero)

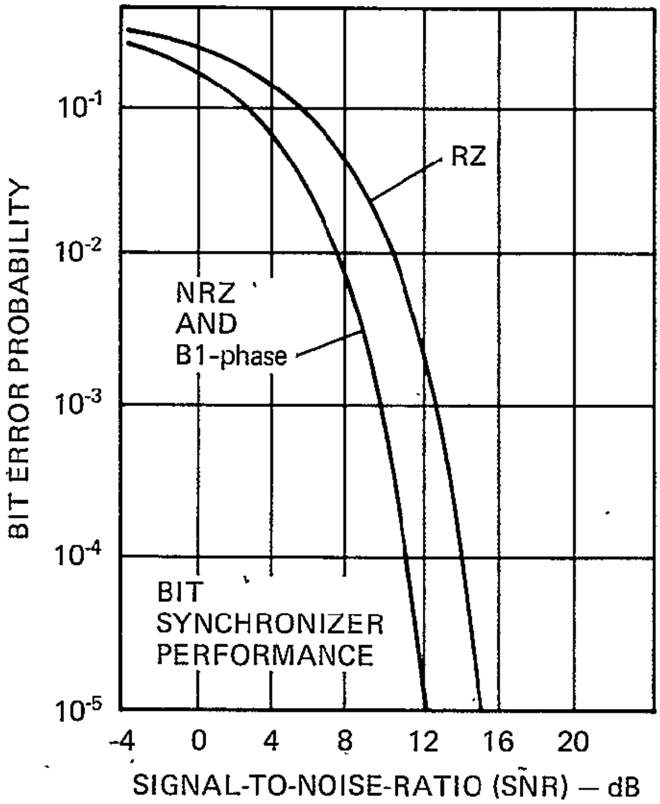


Figure 5-46 Statistical probability of bit error as a function of the ratio between peak signal and rms noise (with the noise bandwidth referenced to one-half the bit rate). Because of secondary signal perturbations such as ac offset, bit jitter and transition density, the actual performance of a bit synchronizer falls within 1 dB of the theoretical over its bit rate spectrum.

5.6.3.2 Data Detection Matched Filters for High Rate Modems

In the design of digital data demodulators for data transmission systems with bit rates of 40 megabits/second and above, special care must be given to the selection of the data detection matched filter.

At lower bit rates, modulator and demodulator filter bandwidths can be very wide providing a nearly square phase modulation characteristic and a nearly ideal integrate-and-dump detector characteristic. Hence, the data detection filter will be a good match to the phase-detector output time waveform if no intermediate bandwidth reduction occurs. However, at the higher bit rates several factors prevent the above ideal modulation waveform and data filter from being achieved. These factors include:

a. Non-zero rise time of the PSK modulator

b. Bandwidth restriction in the modulator output circuit

c. Bandwidth restriction in the receiver prior to the data detection filter

d. Non-zero integrator gate turn-on time and non-zero sample time uncertainty in the data detection filter

e. Appreciable filter dump time

f. Non-zero decision threshold width of the bit decision circuit of the data detection filter

In addition, other factors peculiar to high bit rate modems which cause power (E/N_o) degradation and must be considered when choosing the most practical configurations of modulator and data detection filter are the following:

g. Envelope modulation by the PSK modulator

h. AM to PM conversion power loss in a hard-limiting TWT amplifier

In view of the problem that exists at high bit rates where rectangular pulses can no longer be expected into the data filter and that the integrate-and-dump filter can no longer be ideally realized, it is clear that a review of matched filter principles is needed. This section will define the matched filter for a given waveform and discuss ways of approximating the matched filter for a high bit rate waveform with realizable filters. The requirements for the filter can be stated simply as follows: its design must minimize the sum of mismatch losses arising through factors (a) through (f) and incidental losses (g) and (h) listed above.

5.6.3.3 Matched Filter Principle

A filter is "matched" to the complex envelope of an input signal, whose spectrum function is U_{in} (w), if, and only if, the filter's frequency function is given by

$$H(w) = \frac{U^*_{in}(w)\ e^{-jwT_s}}{N_o(w)}$$

"Matching" occurs in that

a. The filter amplitude variation vs frequency is proportional to the signal components at those frequencies

b. The filter phase is the negative of the signal phase (denoted by the conjugate of U_{in} (w)) so that all signal components add vectorially to a maximum

c. Signal components are weighted inversely as the noise density (N_o (w)) at each point in the band if the noise density is not uniform

The "matched filter" H (w) then maximizes the output SNR at the sampling instant T_s.

For a flat noise spectrum, the matched filter may be described by its time domain properties by taking the Fourier transform of the above expression. The result is the impulse response of the filter:

$$h(t) = u^*_{in}(T_s - t)$$

This expression shows that the matched filter for the signal time waveform $u_{in}(t)$ is that network whose impulse response is an exact copy (but time-reversed and delayed) of the complex envelope waveform of the signal. It can be shown that, for white noise, the maximum achievable SNR with a linear filter is $\gamma = E/n_o$. In addition, the output SNR of the matched filter depends only on the ratio of the energy in the signal pulse to the noise power spectral density, E/n_o, and is independent of any detail of the waveform envelope or bandwidth used in the signal.

5.6.3.4 Loss Effects for Realizable Data Filters

These principles reveal that there exists a distinct hypothetical, but not necessarily realizable, matched filter for any waveform and that this matched filter produces a SNR = E/n_o at the proper sample time. Noting that E is the energy under the modulation pulse waveform, the loss effects and their causes can be listed as follows:

a. Reduction of area under the magnitude squared of the modulation pulse which reduces E. This is caused by bandwidth restriction in the modulator (including equivalent modulator rise time effects) and receiver prior to the data detection filter.

b. Loss due to mismatch between a physically realizable filter and the received waveform. This is caused by there not existing a realizable matched filter for every waveform, by timing errors in the filter (gate turn-on time, sample time and dump time), and by decision threshold level uncertainty.

c. Intersymbol interference. This results if the data filter is not dumped after sampling. A trade-off exists between mismatch loss and intersymbol interference loss (a more lossy filter has less intersymbol interference).

d. AM to FM conversion power loss in a hard-limiting transmitter amplifier. This results if there is envelope modulation due to operation of the PSK modulator.

The system containing the above sources of energy loss and mismatch loss is modeled in Figure 5-47 . Loss effect (d) above is not included since this does not affect the data filter design but only modulator design. Envelope modulation at the phase switch may be largely avoided by employing variable-reactance diode switching which provides a constant VSWR during the switching transient.*

The model above shows the two possible locations of the bandwidth limiting filters (BPF), before and after the addition of noise. The BPF in the transmitter represents the equivalent modulator low pass function which causes non-zero modulator rise time and any physical bandlimiting built into the transmitter. Bandlimiting at the receiver differs in that it also shapes the noise spectrum. The energy loss due to net bandlimiting is represented by k_1. Data filter mismatch loss is represented by k_2. Additional losses are associated with sample timing error τ/T and decision device threshold level uncertainty. These three general losses are treated separately in the material that follows.

5.6.3.4.1 Energy Loss Due to Bandlimiting

This loss is rather ficticious from a system point of view, if the bandlimiting takes place in the modulator, for the following reason. There will be less energy in a pulse of unit height due to slow rise time. This energy loss may be associated with the reduction of the amount of power transmitted in the outer sidebands - the higher order lobes of the sinc (x) function. The power taken from these outer sidebands appears, of course,

---

* See the Interim Engineering Report, pages 5-24 through 5-36.

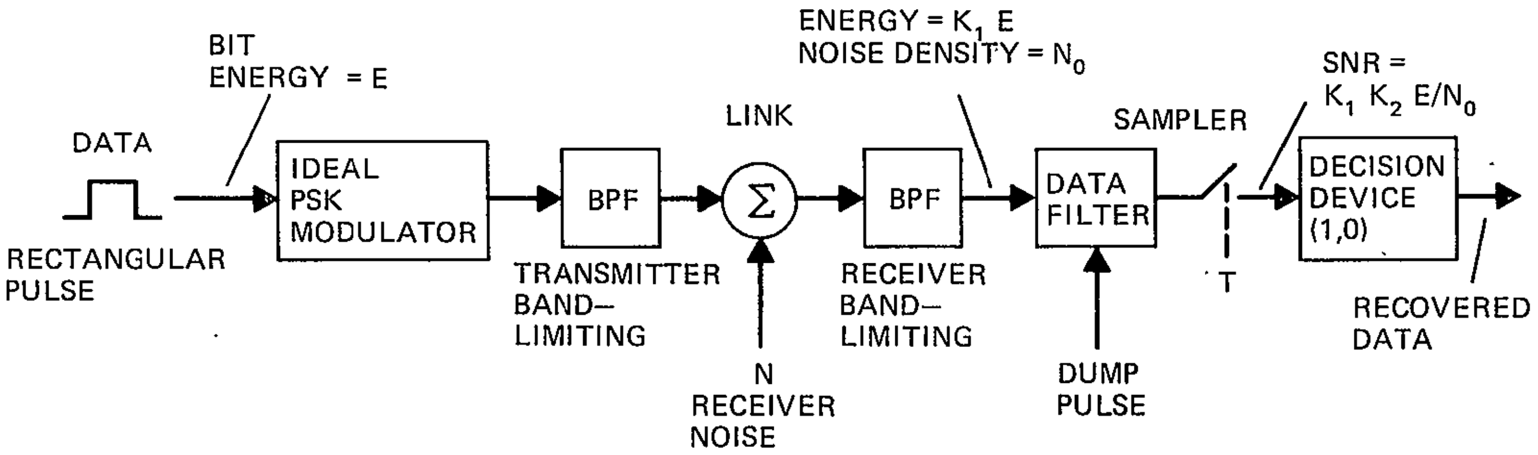

BIT
ENERGY = E
DATA
RECTANGULAR
PULSE
IDEAL
PSK
MODULATOR
BPF
TRANSMITTER
BAND–
LIMITING
LINK
Σ
N
RECEIVER
NOISE
ENERGY = K1 E
NOISE DENSITY = N0
BPF
RECEIVER
BAND–
LIMITING
DATA
FILTER
DUMP
PULSE
SAMPLER
T
SNR =
K1 K2 E/N0
DECISION
DEVICE
(1,0)
RECOVERED
DATA

Figure 5-47 Model of Data System Bit Energy Loss and Data Filter Loss and Data Filter Mismatch Loss. Energy loss factor k_1 is due to spectrum truncation which results in a rounding of the data pulse and a reduction of the area under the square of the pulse magnitude profile. Mismatch loss k_2 is an additional energy loss due to the filter not being matched to the data filter input waveform and hence not extracting the total energy k_1E from the waveform pulse.

in the more central spectral components because the transmitter is a constant power device. Hence the energy loss does not represent an actual power loss. The energy loss is reflected in the E actually received by the data filter and the loss is given by

$$k_1 = 2\, t_r/(3T)$$

where t_r is the rise time of a trapezoidal wave. This is a worst case figure corresponding to integration starting and sampling taking place at the beginning of a transition. If integration begins and sampling takes place at the 50% rise time points (which is more likely because of bit synchronizer timing) the value of k_1 is $5\, t_r/(12T)$.

It should be noted that in the quantitative comparison of data filters below, each filter's performance is computed for a square pulse in. The effect of bandlimiting loss k_1 in E is therefore lumped with mismatch loss k_2. The total is then labeled "mismatch loss".

5.6.3.4.2 Mismatch Loss Due to the Data Filter

As pointed out above, the matched filter for the general pulse waveform is not necessarily realizable. The output SNR of the mismatched data filter is $\gamma = k_2 E_1/n_o$ where E_1 is the energy under the data filter input pulse. Comparison of filters is carried out for a square pulse, hence $\gamma = k_1 k_2 E/n_o$ where E is the energy of a square pulse. Therefore, in the following discussion "mismatch loss" is taken to be $k_1 k_2$. If the actual value of k_2 is desired, the waveform shape would be needed to evaluate and remove k_1.

The performance of several mismatched filters has been evaluated for no intersymbol interference (filters are dumped) and a square pulse input by Jones*. The results are summarized in Figure 5-48 . Note that the hypothetical gaussian filter out-performs the others for BT = .75. Among the realizable filters the 2-pole Butterworth results in 0.6 dB loss with respect to the matched filter (ideal integrate-and-dump for a square pulse). These results are independent of filter input E/n_o.

---

* Jones, J. J., "SNR Degradation and Intersymbol Interference with Mismatched Filters", Philco-Ford SRS Communication Sciences Dept. TM 141, March 15, 1968.

58

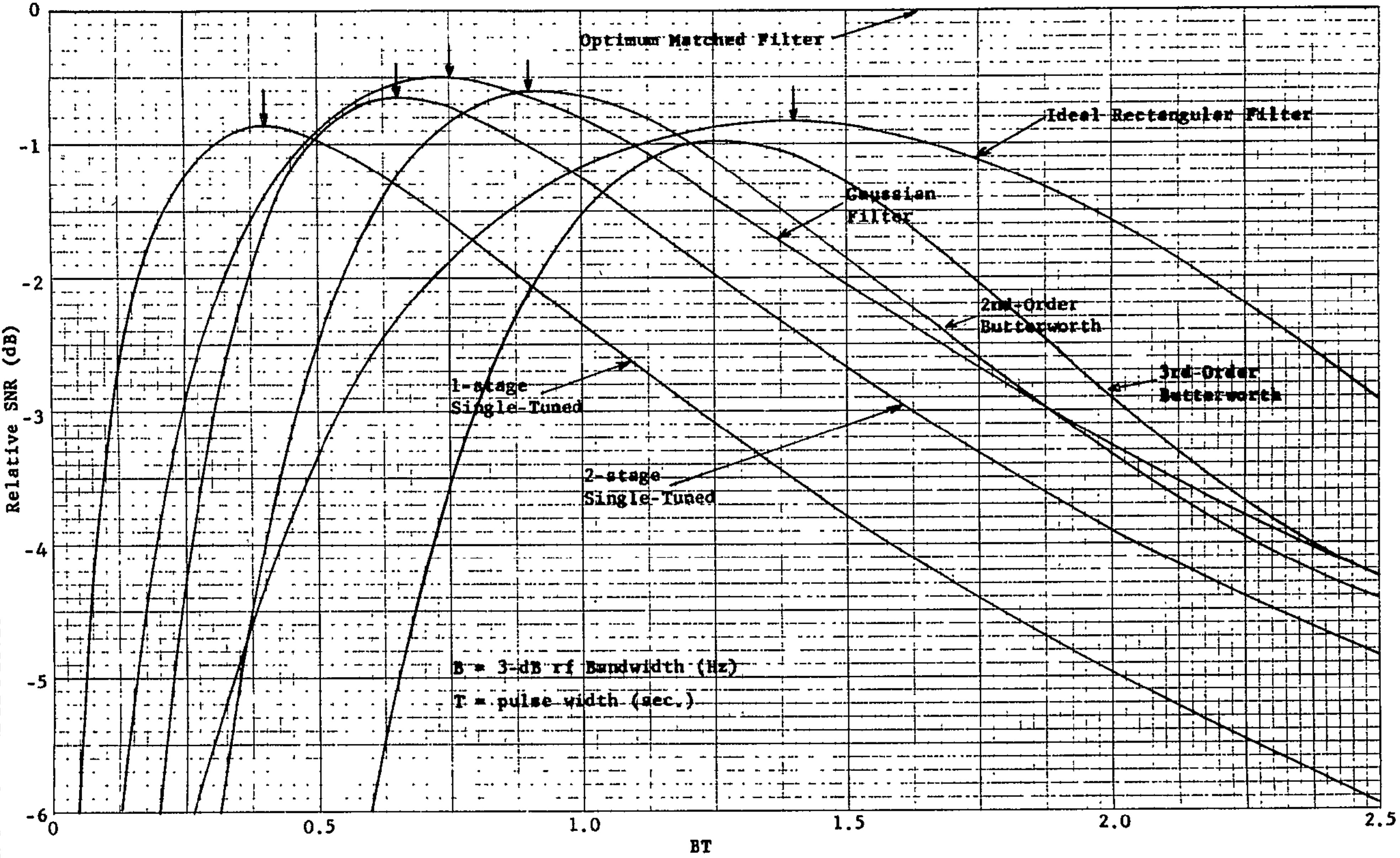


Figure 5-48 Mismatched Filter Loss Compared with Matched Filter (Rectangular Pulse Input)

The performance of several filters has also been evaluated in presence of intersymbol interference by Natali[1]. These results are a function of filter input E/n_o, and are presented in Figure 5-49, for $E/n_o = 6$ dB. Note that BT is defined as in the previous figure. The 2-pole Butterworth again has the least loss, best performance occuring for BT = 1.0. No dumping takes place in these filters.

The second-order Butterworth data filter has been selected due to the above results for further investigation as an approximation to a matched filter for a bandlimited waveform into the data filter. The performance of the ideal integrate and dump filter and the Butterworth filter with BT = 1 have been evaluated[2] for the model of Figure 5-50. The effect of various modulator rise times t_r and receiver bandwidths B_R for a specific bandpass filter characteristic were determined. Figure shows the effect of modulator rise time alone. Note that the Butterworth filter gives less degradation for rise time greater than 24%. Figure 5-51 shows the effect of receiver bandlimiting alone. Again the Butterworth out-performs the integrate-and-dump for receiver bandwidth less than 2/T. The combined effect of 10% modulator rise time and receiver bandlimiting is shown in Figure 5-52. Here the Butterworth degradation is less for $B_r < 2.5/T$. It is clear then that the Butterworth filter will give nearly equal or better performance than the ideal integrate-and-dump filter if significant modulator rise time (or equivalent transmitter bandlimiting) or receiver bandlimiting exists. Since this filter may be more readily implemented and since transmitter bandlimiting is unavoidable at high bit rates (> 100 Mbps), then the second-order Butterworth is preferred over the integrate-and-dump filter. Additionally, the Butterworth filter without dumping might be considered where the additional degradation due to intersymbol interference could be accepted.

---

[1] Private communication from F. D. Natali, Philco-Ford.

[2] Jones, J. J., "Filter Distortion and Intersymbol Interference Effects on Coherent BPSK and QPSK", Philco-Ford SRS Communication Sciences Dept. Technical Memo (TM) to be published.

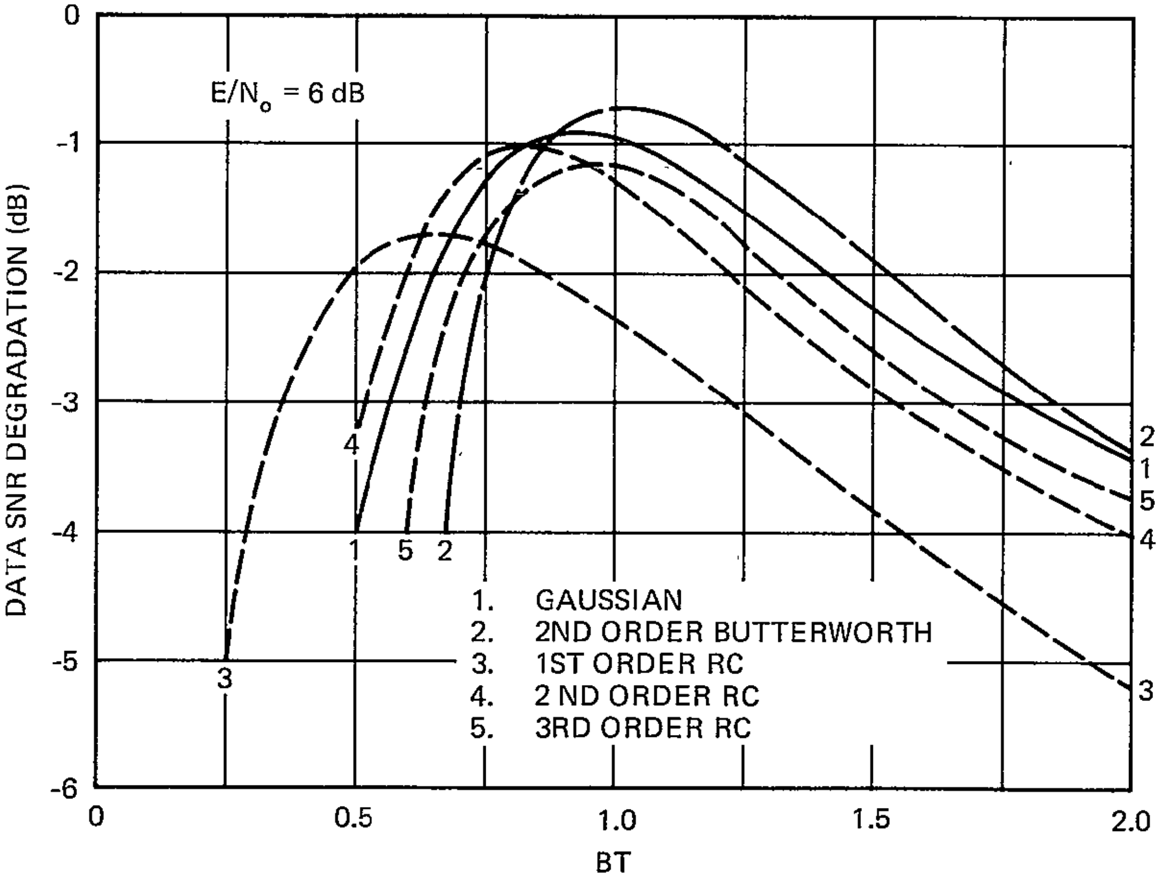


Figure 5-49 Comparison of Mismatched Filters with Intersymbol Interference

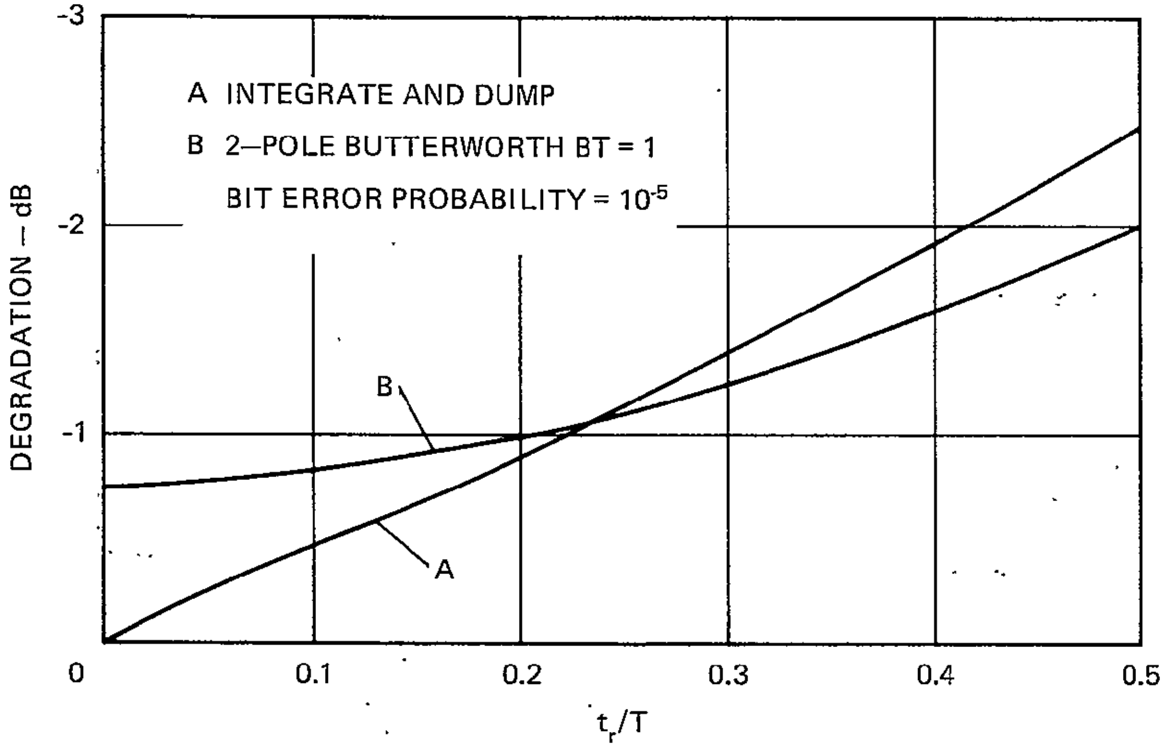


Figure 5-50 SNR Degradation vs Modulator Rise Time for No Bandlimiting

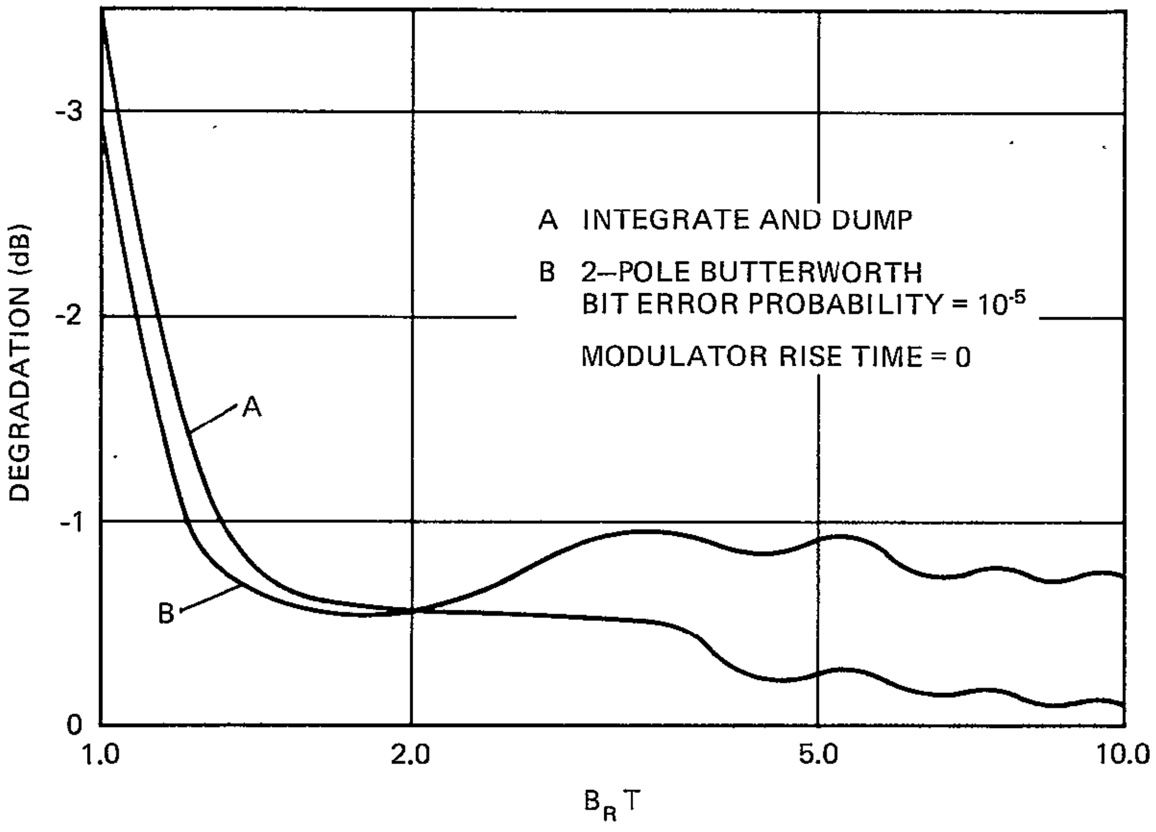


Figure 5-51 System Degradation vs Receiver Bandwidth Limiting for Modulator Rise Time = 0

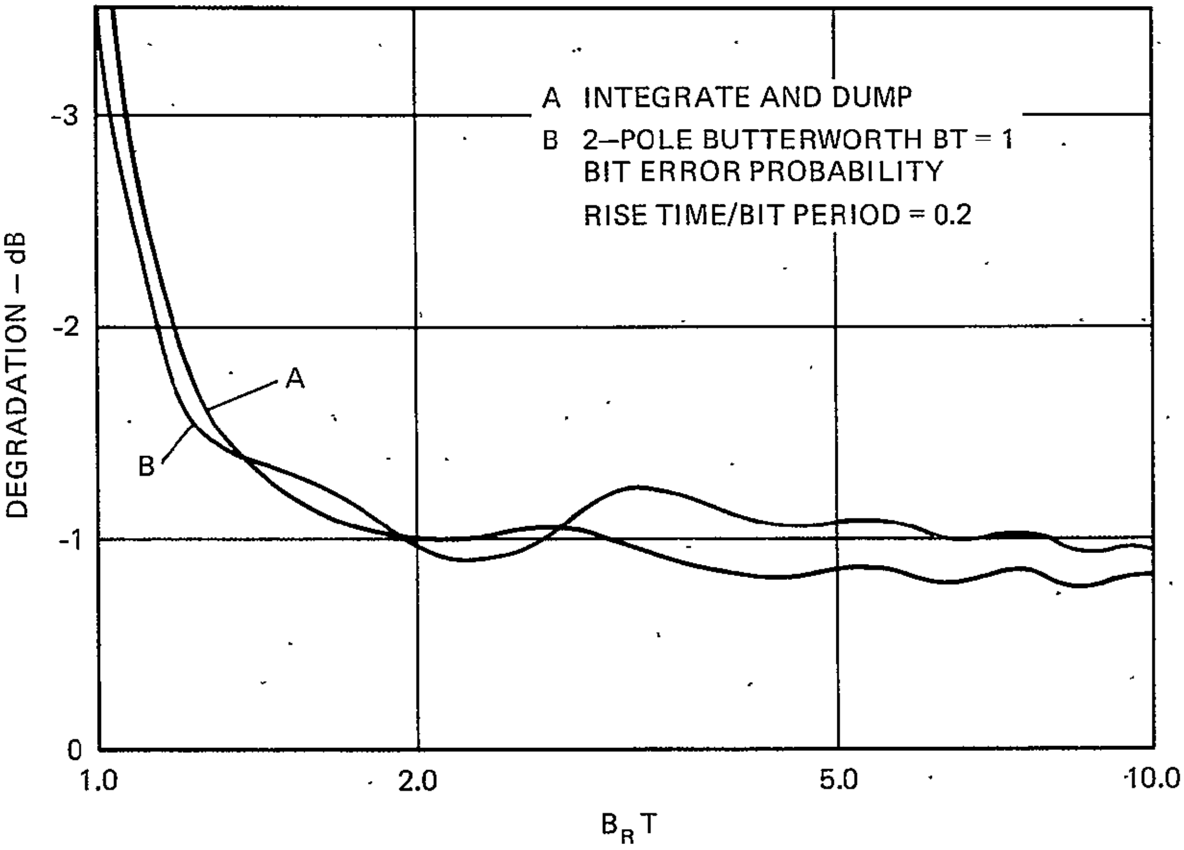


Figure 5-52 System Degradation vs Receiver Bandwidth Limiting. Modulator Rise Time is 20% and Receiver Predetection Filter is 5-Pole Chebychev .1 dB Ripple

5.6.3.4.3 Data Detection Filter E/n_o Loss Due to Sample Timing Error

Data filter output SNR degradation has been evaluated by computer for the case of the non-dumped filter. Computations of degradation were run for both constant E/n_o and constant bit error probability. Figures 5-53 and 5-54 gives the degradation for a square pulse into the data filter due to a static offset in sample timing[1]. Input E/n_o was constant at 9 dB, which corresponds to about P_e (bit) = 10^{-5}.

Three suboptimum data filters are compared to the ideal integrate-and-dump filter. The suboptimum filters are not dumped. The resulting intersymbol interference was computed including the effect of the six preceding symbols.

Note that in the non-dumped filters, the optimum sampling time is between zero and eight percent after the beginning of the next bit. The second-order Butterworth filter out-performs the other suboptimum filters tested. It is evident that, in presence of ± 10% sample time error, the Butterworth filter might be preferable to the integrate-and-dump filter.

Figure 5-55 presents results for data filter output SNR degradation computed for a constant bit error probability of 10^{-5}.[2] These results compare the undumped second-order Butterworth data filter with the ideal integrate-and-dump filter. The effects of modulator rise time and predetection bandlimiting are presented for the Butterworth filter. Intersymbol interference is computed including only the effect of the previous bit. In these results optimum sampling time is 0 to 3% late.

These curves are more sensitive to sample timing error than the curves for constant E/n_o. Note that the curve for predetection bandlimiting results in performance equal to the no bandlimiting case. This is because the Butterworth filter fits the bandlimited signal spectrum better than the unlimited case. Curve #4 is for 20% modulator rise time. Modulator rise time was approximated by a single real pole with time constant equal to the desired rise time.

---

[1] Private communication from F. D. Natali, Philco-Ford.

[2] Jones, J. J.

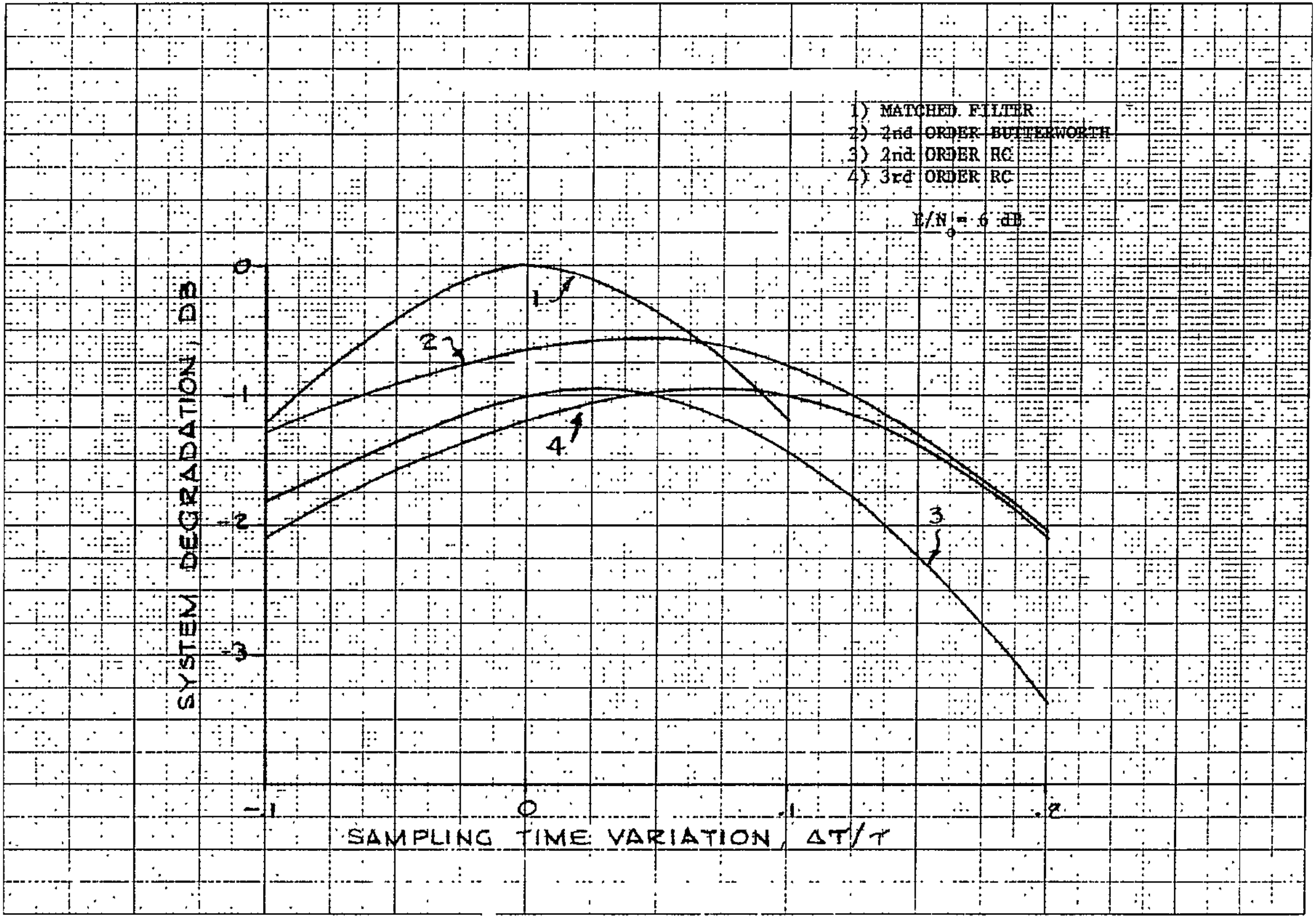


Figure 5-53 Effects of Variation in Sampling Time

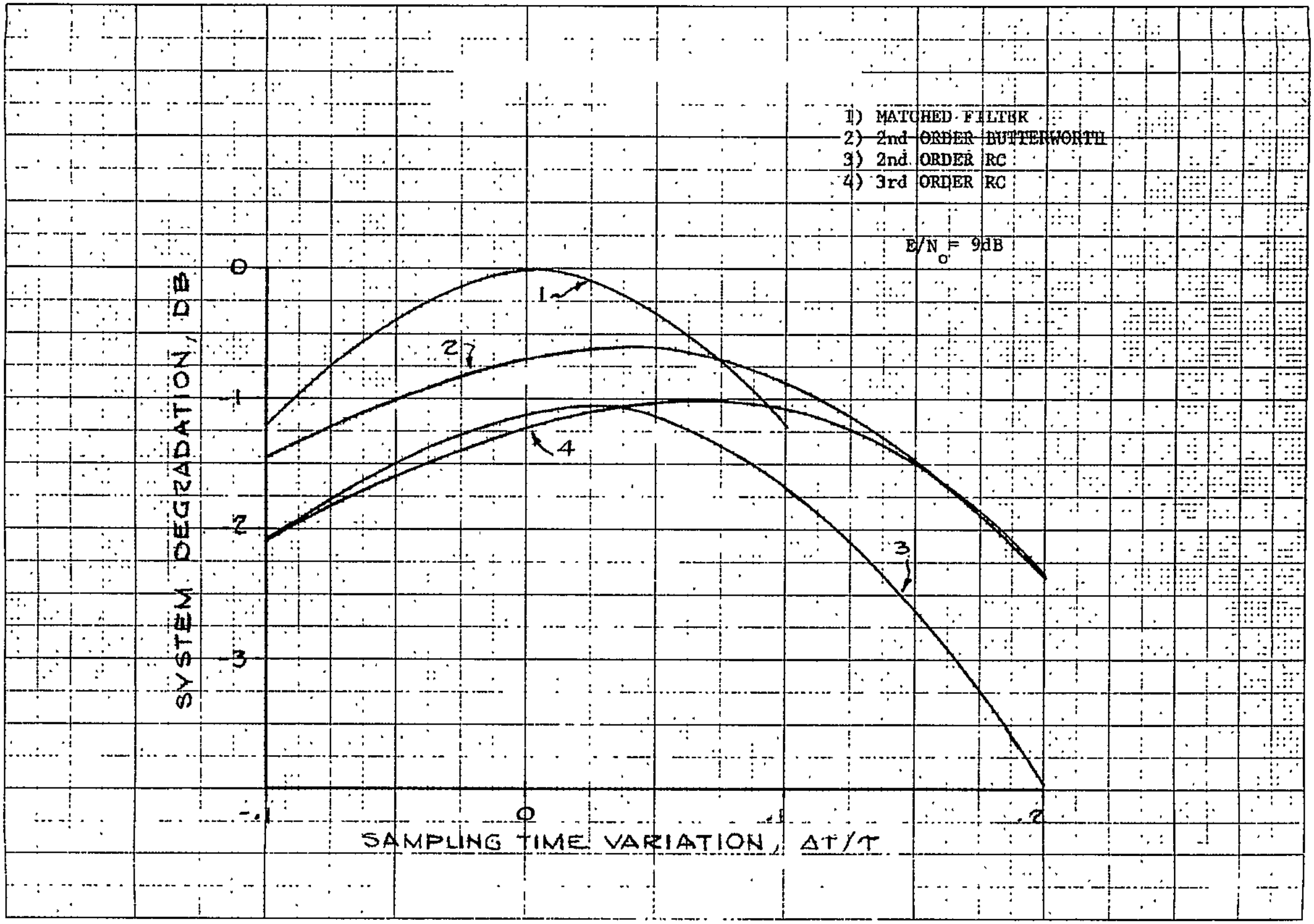


Figure 5-54 Effects of Variation in Sampling Time

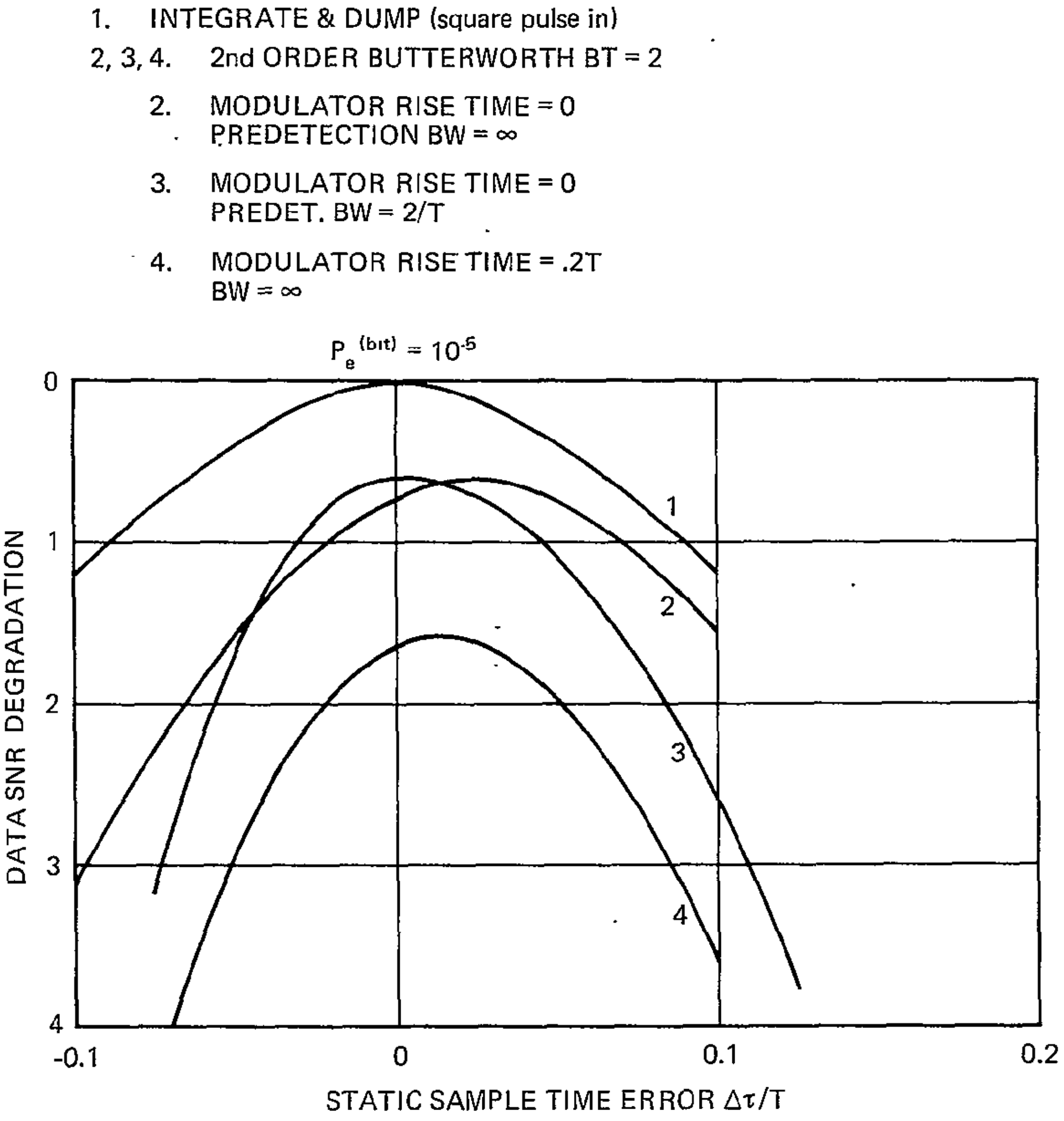


Figure 5-55 Data Filter E/No Degradation for Constant Bit Error Probability

5.6.3.5 An Approximation to a Matched Filter for a Typical Modulator Output Waveform

Consider the approximately trapezoidal time waveform u(t) which results from the modulator non-zero switching-time characteristic. As pointed out above, a data filter which is matched to this time waveform has an inverse time impulse response h $(T - \tau)$ which is equal to the waveform's time function u (t). Although exact matched filters do not necessarily exist for the general time function, some matched filters can be closely approximated with practical circuits. Such is the case here. The matched filter for the trapezoidal wave can be closely approximated by the lossy integrate-and-dump filter whose impulse response is the second-order time function.

$$h\,(t) = \begin{cases} e^{-\beta t} \sin \beta t, & 0 < t \leq T \\ 0 \qquad\quad , & T < t < 0 \end{cases}$$

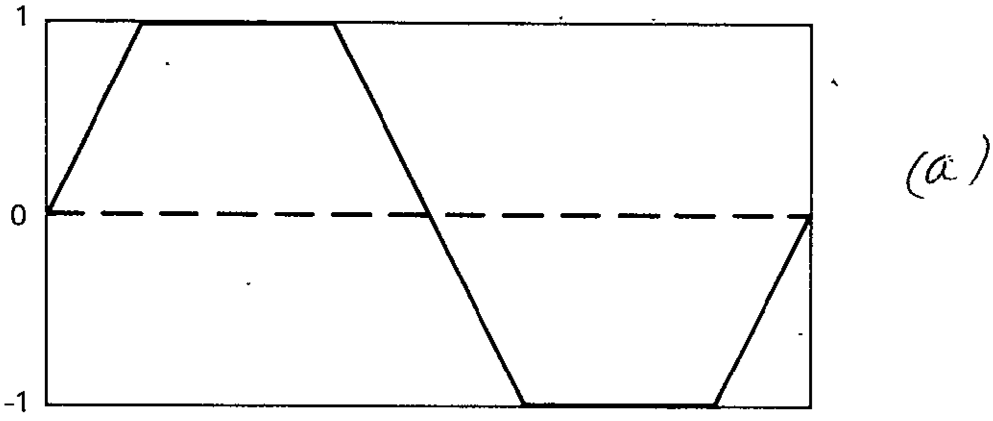
The trapezoidal modulator output waveform (not further band-restricted in the transmitter or receiver) and the lossy I&D impulse response are shown in Figure 5-56 for comparison. It is evident from the figure 5-56 that the impulse response of the second-order filter approximates the shape of the trapezoidal wave. Therefore, it can be expected that the degradation of this filter with respect to the (hypothetical) matched filter for the trapezoidal wave will be small.

The above intuitive argument is supported by computation of performance of the above realizable, second-order filter with a rectangular wave input. This filter is a two-pole Butterworth and for BT = $(\sqrt{2}/\pi)$ $\beta T = .9$ the degradation with respect to the matched filter for the rectangular wave is 0.6 dB*. Note that intersymbol interference effects are not included in the computation, hence the filter is assumed to be dumped at the end of each bit period. The conclusion from this evidence is that since the filter impulse response more nearly approximates the trapezoidal wave than the square wave then the mismatch loss should be less than 0.6 dB for an input trapezoidal wave.

---

*Jones, J. J., "SNR Degradation and Intersymbol Interference with Mismatched Filters", Philco-Ford SRS Communication Sciences Department TM 141, March 15, 1968.

(a) Trapezoidal Wave
(b) Impulse Response of
a 2-Pole Butterworth (Dumped)



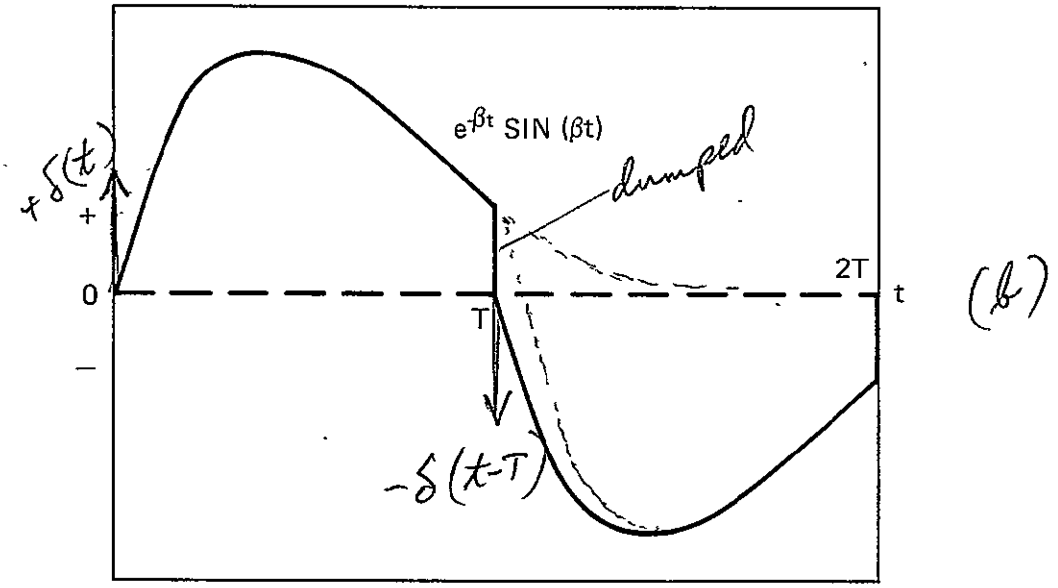


Figure 5-56

~~Response of Integrate and Dump Filter to a Trapezoi~~dal ~~Waveform~~

5.6.3.6 Degradation Due to Decision Threshold Uncertainty

The implementation problems of MPSK for $M = 2, 4, 8$ and 16 are now considered. The use of MPSK as a bandwidth-conserving technique has been discussed previously. An additional advantage of MPSK for $M > 2$ is that signalling rate decreases with 1/M. This means that data filter circuits will be more easily realized with lumped-constant circuits. The increase in required E/n_o for $M > 4$ should be kept in mind, however.

The increased demands upon the phase stability of the modulator and demodulator is evident when one considers the phase coding diagram, Figure 5-57. As shown, a phase error of π/M radians is sufficient to cause a bit error. Phase stability of the reference and narrowness of the phase decision threshold become more important as M increases.

In a present quadriphase system[1] phase stability of 70 MHz references are readily maintained to approximately 0.1^o accuracy by using temperature and parameter stable delay lines. Similar techniques should provide 1^o accuracy at 700 MHz, the higher frequency being required to carry 100 to 200 MHz signalling rates. The more significant error may be the angular width of the threshold uncertainty zone, which, for a given switching circuit, will increase with the signalling rate. The dependence of threshold phase uncertainty upon decision device time uncertainty is plotted vs signalling rate in Figure 5-58. Included are upper bounds on phase error for 1.0 dB loss[2]. Note that for 4 PSK and data rates above 200 Mbps (symbol rate = 100 MHz), tunnel diode logic (or techniques with similarly small timing uncertainty) should be used in the decision circuit. Decision circuit timing uncertainty becomes much more critical for 16 PSK, where tunnel-diode logic should be used for 80 Mbps (symbol rate = 20 MHz) or greater.

---

[1] Magill, D. T., "Quadriphase Communication System Design: Part I Theory, Part II Design Parameter Considerations", Philco-Ford SRS TM 135; December, 1967.

[2] Prabhu, V. K., "Error Rate Considerations for Digital Phase-Modulation Systems", IEEE Trans. on Comm. Tech., Vol. COM-17, No. 1, pp 33-42; February, 1969.

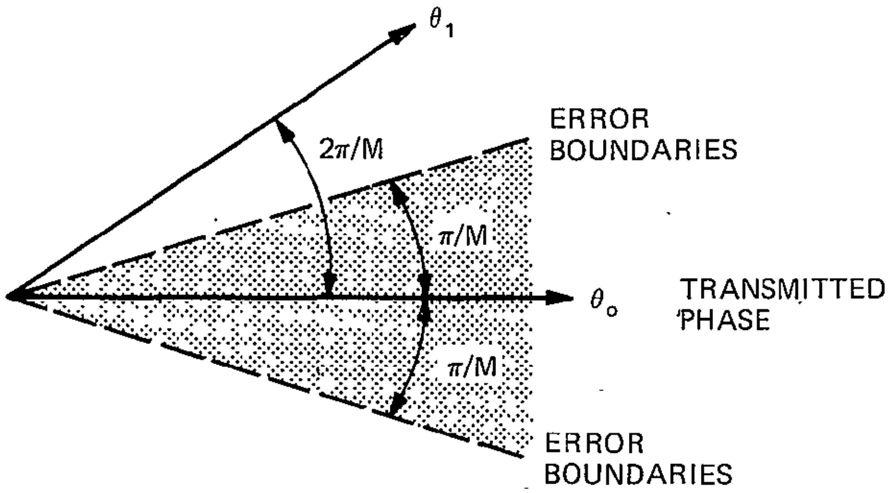

θ1
2π/M
ERROR
BOUNDARIES
π/M
θo
TRANSMITTED
PHASE
π/M
ERROR
BOUNDARIES

Figure 5-57 Zone for Correction Reception in MPSK Signalling

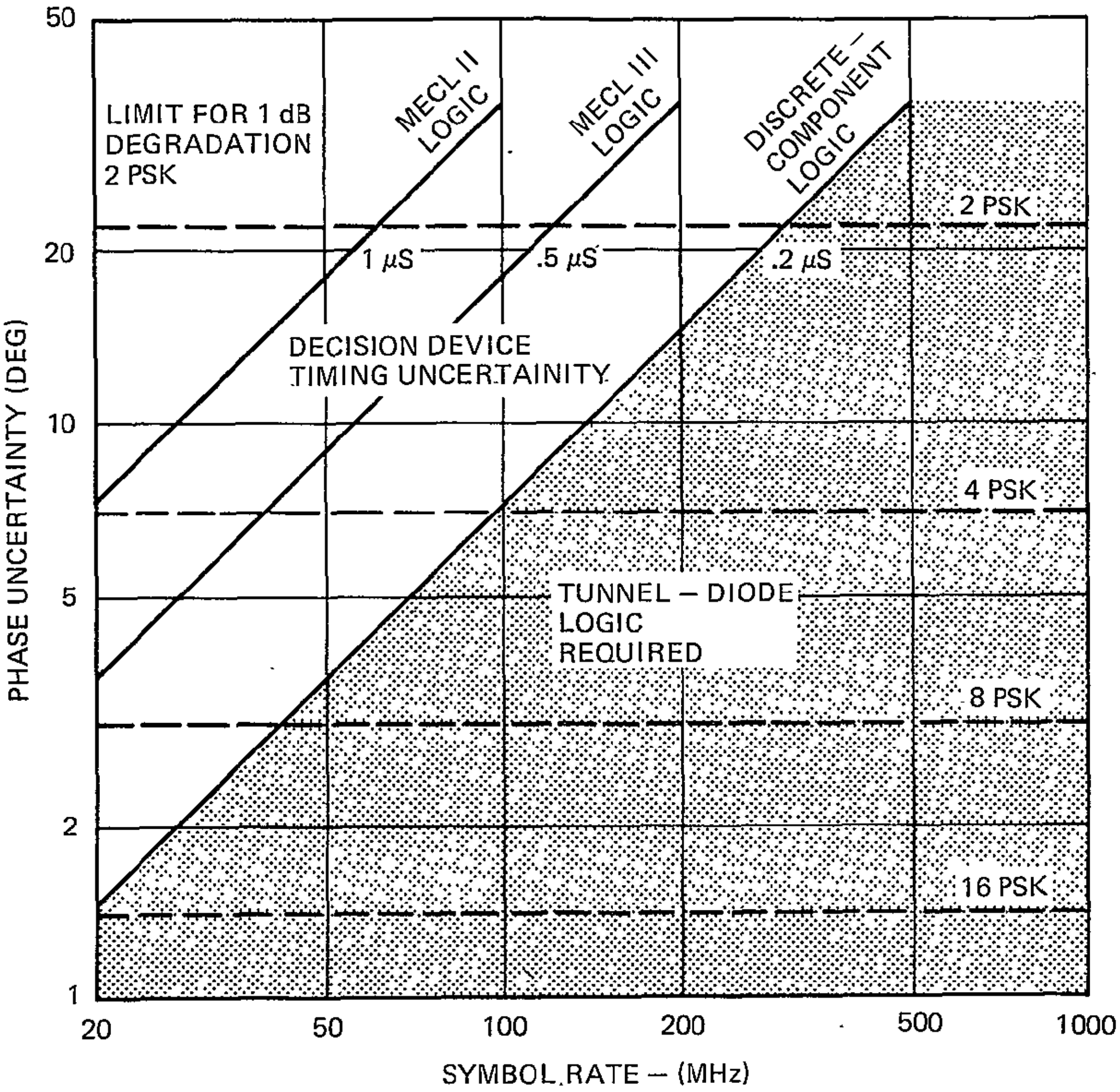


Figure 5-58 Areas of Applicability of Logic Devices for Symbol Decision in MPSK Data Detection

5.6.3.7 Filter Distortion and Intersymbol Interference Effects on QPSK

Distortion and intersymbol interference in a sequence of pulses in PSK transmissions, transmitted at intervals of T seconds, arise from both the transmission filter and from the data filters. System performance degradation is measured by the additional amount of energy contrast ratio, E/N_o, needed to maintain a given average bit error probability with a specified filtering condition relative to the undistorted ideal case. This is particularly true at the high bit rates discussed in this report.

Consideration of intersymbol interference from K previous data symbols results in a set of 4^K (2^K for BPSK) possible signal vectors defined as samples of the data filter output at time, $T = \tau$. The squared length of each vector is a signal-to-noise ratio.

The signal vector angles, together with the phase detector reference, determine that portion of the length which is effective in making bit decisions. The effective portion of each vector length, perpendicular to the decision boundary, is the projection of the length on a line colinear with the phase reference vector shifted $\pm \pi/4$ (no shift for BPSK) with choice of sign associated with a particular quadrature channel. Because of symmetry, the average performance is the same for both channels and is obtained by averaging the resultant individual bit error probabilities over the entire set of 4^K (or 2^K) possibilities.

For QPSK, the average bit error probability is

$$P_b = \frac{4^{-K}}{2} \sum_{\text{all } i_K} \text{erfc} \left[\frac{v\,(T + \tau \mid i_1, \ldots i_K)}{\sqrt{2N}} \cos\left(\theta_{i_1, \ldots i_K} - \varphi_{\text{ref}} \mp \frac{\pi}{4}\right) \right]$$

where the individual signal vector angles are given by

$$\theta_{i_1, \ldots i_K} = \arctan \frac{\text{Im}\ v\,(T + \tau \mid i_1, \ldots i_K)}{\text{Re}\ v\,(T + \tau \mid i_1, \ldots i_K)}$$

and the coherent phase reference angle is taken to be the weighted average:

$$\varphi_{ref} = \frac{\sum\limits_{\text{all } i_K} \left| v\,(T + \tau \right| i_1, \ldots i_K)\right|^2 \theta_{i_1,\ldots i_K}}{\sum\limits_{\text{all } i_k} \left| v\,(T + \tau \right| i_1, \ldots\, i_K)\right|^2}$$

The delay, τ, in the decision sampling time is taken to be the group delay of the transmission filter measured at the signal carrier frequency offset, Δ, from the transmission filter bandcenter. This is a near-optimum sampling time. (See Figure 5-59.)

Figures 5-60 and 5-61 show the effect of mismatched data filter detection only (no transmission filter) as a function of BT product. In each case, the optimum BT product is slightly dependent on E/N_o but BT = 1 appears to be a good choice for the 2-pole Butterworth filter and BT = 0.75 for the single-pole RC filter.

In Figure 5-62a, Chebychev transmission filter bandwidth-limiting degradation is shown versus symbol-rate-to-filter-bandwidth ratio for both integrate-and-dump and Butterworth data filters. With symmetrical filtering and the same symbol rate, 1/T, QPSK and BPSK have identical performance. To avoid excessive degradation, the 3-dB RF transmission filter bandwidth should be at least twice the symbol rate ($BT \geq 2$). For severe bandwidth limiting, the Butterworth data filter is more closely matched to the distorted signal than is the integrate-and-dump data filter. Degradations shown can be reduced by the energy truncation loss curve if transmission filtering occurs before power amplification.

The effect of broadband, mistuned transmission filtering is shown in Figure 5-62b as a function of signal carrier displacement, Δ, from filter bandcenter. Low loss requires the signal carrier offset to be no closer than the symbol rate to the filter 3-dB bandedge. For

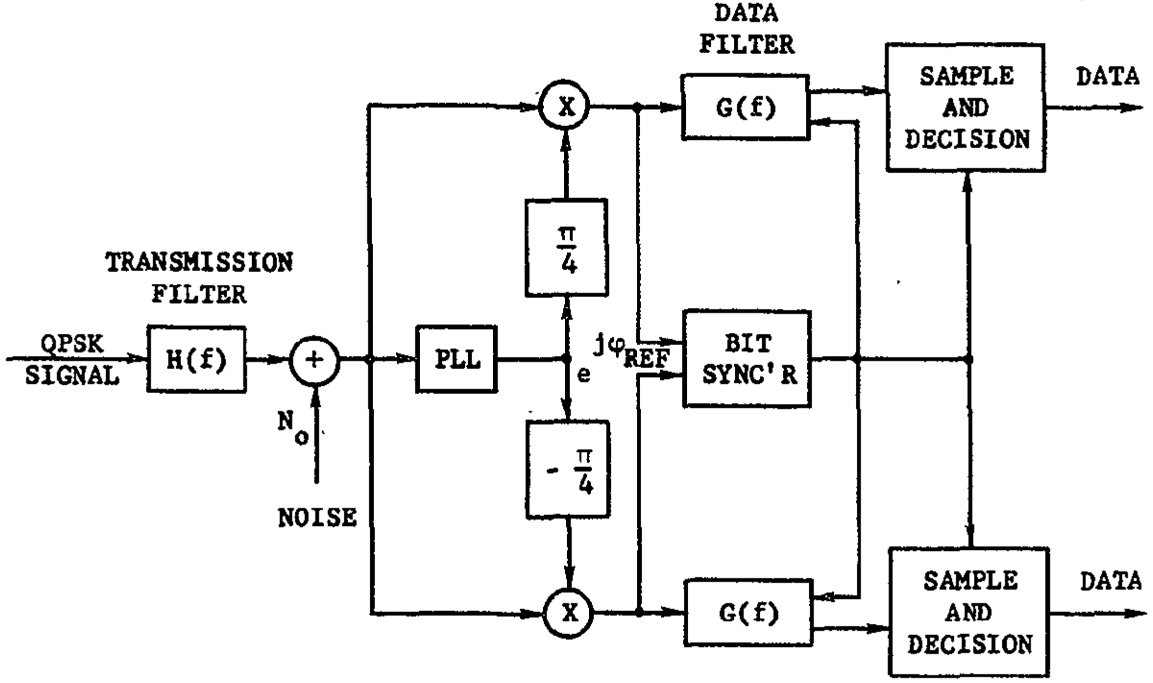

DATA
FILTER
TRANSMISSION
FILTER
QPSK
SIGNAL
H(f)
+
N_o
NOISE
PLL
X
$\frac{\pi}{4}$
$e^{j\varphi_{REF}}$
$-\frac{\pi}{4}$
X
G(f)
BIT
SYNC'R
G(f)
SAMPLE
AND
DECISION
DATA
SAMPLE
AND
DECISION
DATA

Fig. 5-59 **QPSK LOWPASS SYSTEM MODEL SHOWING TRANSMISSION FILTER AND DATA FILTERS (QUADRATURE CHANNEL BIT DETECTION)**

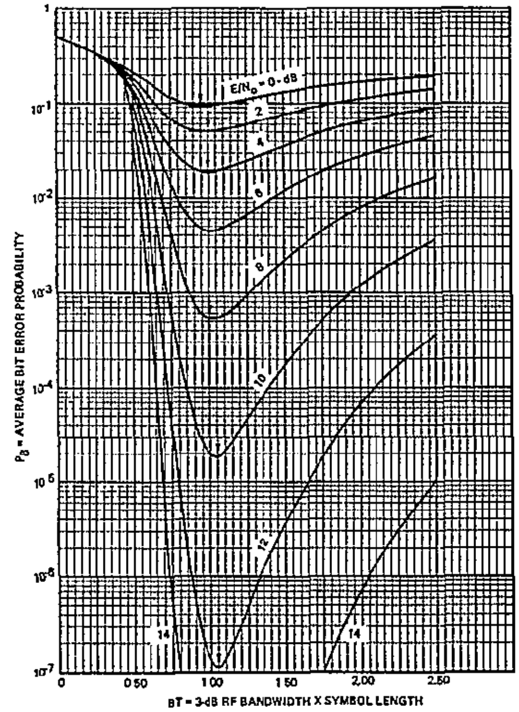


Fig. 5-60 2-Pole Butterworth Data Filter Detection of Coherent QPSK and BPSK

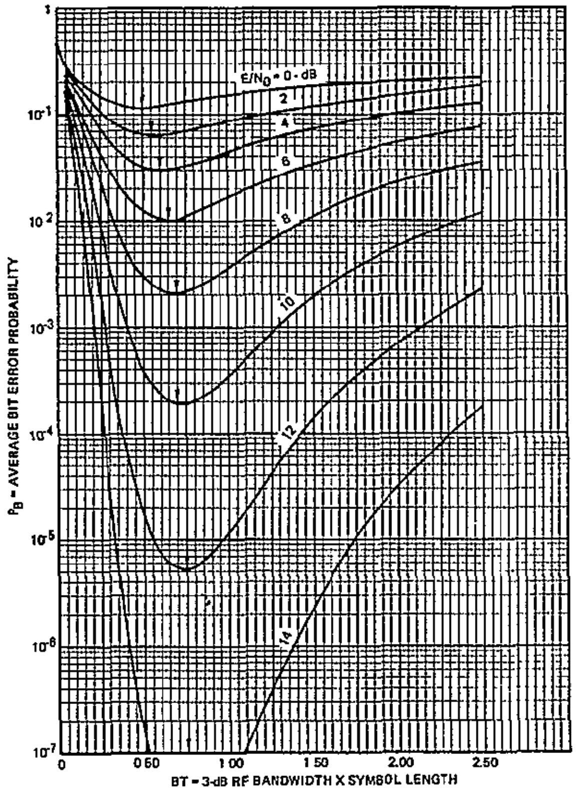
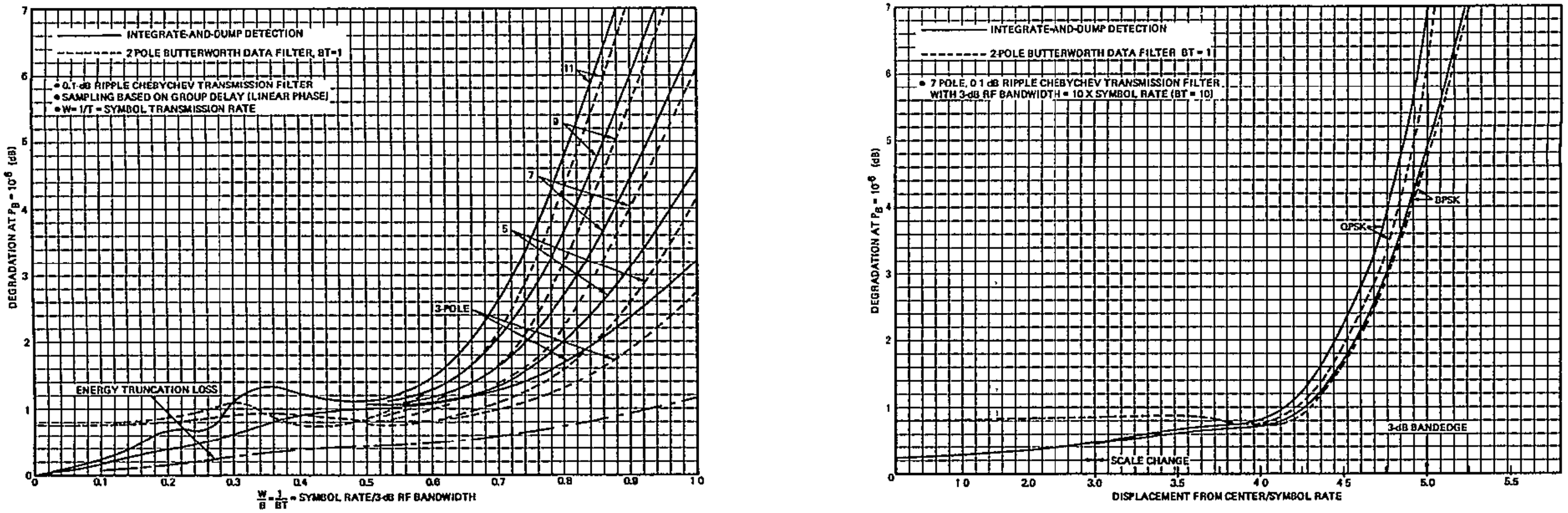
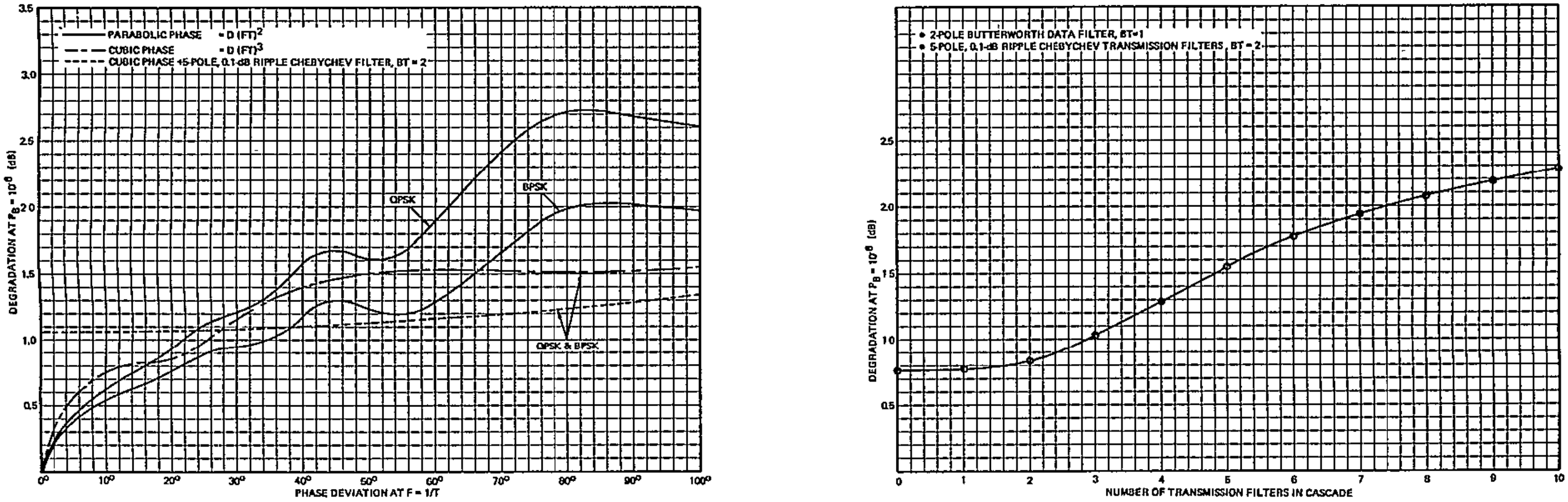


Fig. 5-61 Single-Pole RC Data Filter Detection of Coherent QPSK and BPSK



(a) Bandwidth-Limiting Degradation of QPSK and BPSK Signals

(b) Mistuned-Broadband Filtering



(c) Phase Distortion with I&D Detection

(d) Cascade Filter Degradation of QPSK and BPSK Signals

Figure 5-62

a mistuned or unsymmetrical filter condition, QPSK suffers a larger degradation than does BPSK at the same symbol rate and energy/bit.

Figure 5-62c illustrates the degradation due to a pure phase distortion with the phase deviation measured at f equal to the symbol rate. Parabolic phase distortion corresponds to a mistuned condition whereas cubic phase distortion corresponds to symmetrical filtering. The introduction of a bandlimiting transmission filter significantly reduces the sensitivity to phase distortion.

The degradation for a number of identical transmission filters in cascade is shown in Figure 5-62d. The loss grows slowly with the number of filters in cascade, especially from 1 to 2 filters. As observed also in Figure 5-62d. an initial bandwidth limiting reduces the sensitivity to further distortions.

5.6.4 Bit Synchronizer

The bit synchronizer (see 5.5.2.2) generates a coherent data clock from the output of the PSK detectors. The data clock is used as a basis for sampling of the data filter outputs.

The bit synchronizer phase locks a crystal oscillator to the data transitions at the output of the PSK detectors.

There are several types of bit synchronizers which are adaptable to coherent-data-clock generation from a high speed PSK demodulator.

- Differentiated Bit Synchronizer: In this system, the detected data stream is differentiated and squared to generate in series of impulses or toggle signals representing the transitions of the data stream.

- Early - Late Bit Synchronizer: In this system the data stream is applied to a pair of integrate and dump circuits which are driven by a VCO. The two integrate and dump circuits are circuits are clocked differently to dump early, and dump late respectively. The outputs of the two integrate and dump circuits are squared and applied to a comparator and low pass filter which adjusts the timing of the VCO.
- Digital Bit Synchronizer: In this system each bit period is scanned by a series of high speed pulses and the change in pulse polarity developed from pulses occuring on each side of a zero-crossing transition is used to time an oscillator.

While the bit (and word) synchronizer would appear to be peripheral and of secondary importance, actually, it is a key circuit for generating a properly timed clock and will determine the effectiveness of the demodulator process.

5.6.4.1 Bit Synchronizer Circuit Implementation

The sampled data phase-locked loop phase locks a bit clock to any non-return-to-zero (NRZ) waveform in which data transitions occur. A transition detector provides output pulses coincident with all transitions in the NRZ data. This data transition is used to generate a gating or sampling pulse of one (1) bit period duration and also sets a "set-reset" flip flop used as a phase detector. The sampling pulse is used to gate the output of the phase detector into an averaging circuit and a first-order hold circuit. The output of the hold circuit represents the error signal. The error is then integrated and the integrator output used to control the frequency of a voltage-controlled oscillator (VCO). The clock thus generated resets the phase detector on its negative-going transition, thereby closing the phase-locked loop.

A bit synchronizer timing diagram is shown on Figure 5-63, the output of the phase detector is averaged over only one bit period following each detected data transition. Assuming that the negative-going clock transition occurs in the center of the data bit, the output of the phase detector will be positive for one-half of the bit period and negative for the other half, averaging to zero over a full bit period, and negative for the other half, averaging to zero over a full bit period. If a phase error exists, the negative clock transition will not occur in the precise center of the bit and the average will not be zero; hence, an error signal will result.

5.6.4.2 Bit Synchronizer Error In Modulator Switching Time

Non-zero modulator switching time results in less transmitted symbol energy and hence degraded BER at the demodulator. The rise time of practical modulators is on the order of one nanosecond so that at data rates of 100 Mbps or greater the modulator switching time will produce measurable degradation.

The degradation produced by 10% rise time can be determined by comparing the curves labeled "0" and "1" in Figure 5-64. For example, at a BER of 10^{-4}, the 10% modulator rise time causes 0.6 dB degradation.

It can be noted by generating other curves for other rise times that the degradations introduced by the modulator rise time and the transmission filters are not independent. The transmission filters have less effect on a signal generated by a modulator having non-zero rise time. For example, at a BER of 10^{-4}, the rise time degradation is 0.6 dB and the degradation from only the filter is 1.1 dB. When the two effects are combined, the net degradation is 1.3 dB, rather than 1.7 dB, the sum of the two degradation sources taken independently.

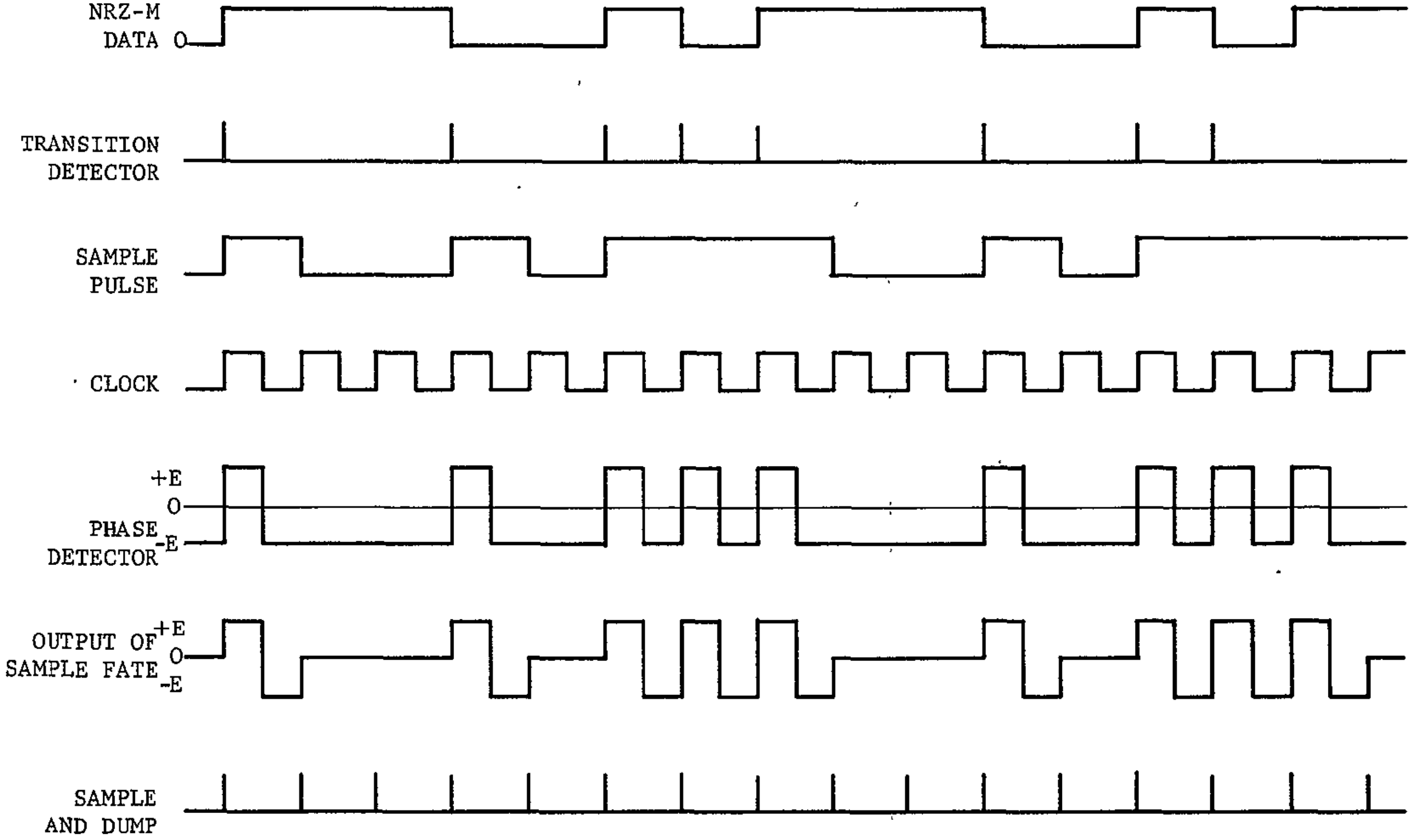

NRZ-M
DATA 0
TRANSITION
DETECTOR
SAMPLE
PULSE
CLOCK
+E
0
PHASE
DETECTOR
-E
OUTPUT OF
SAMPLE FATE
+E
0
-E
SAMPLE
AND DUMP

Figure 5-63 Bit Synchronizer Timing

I-598A

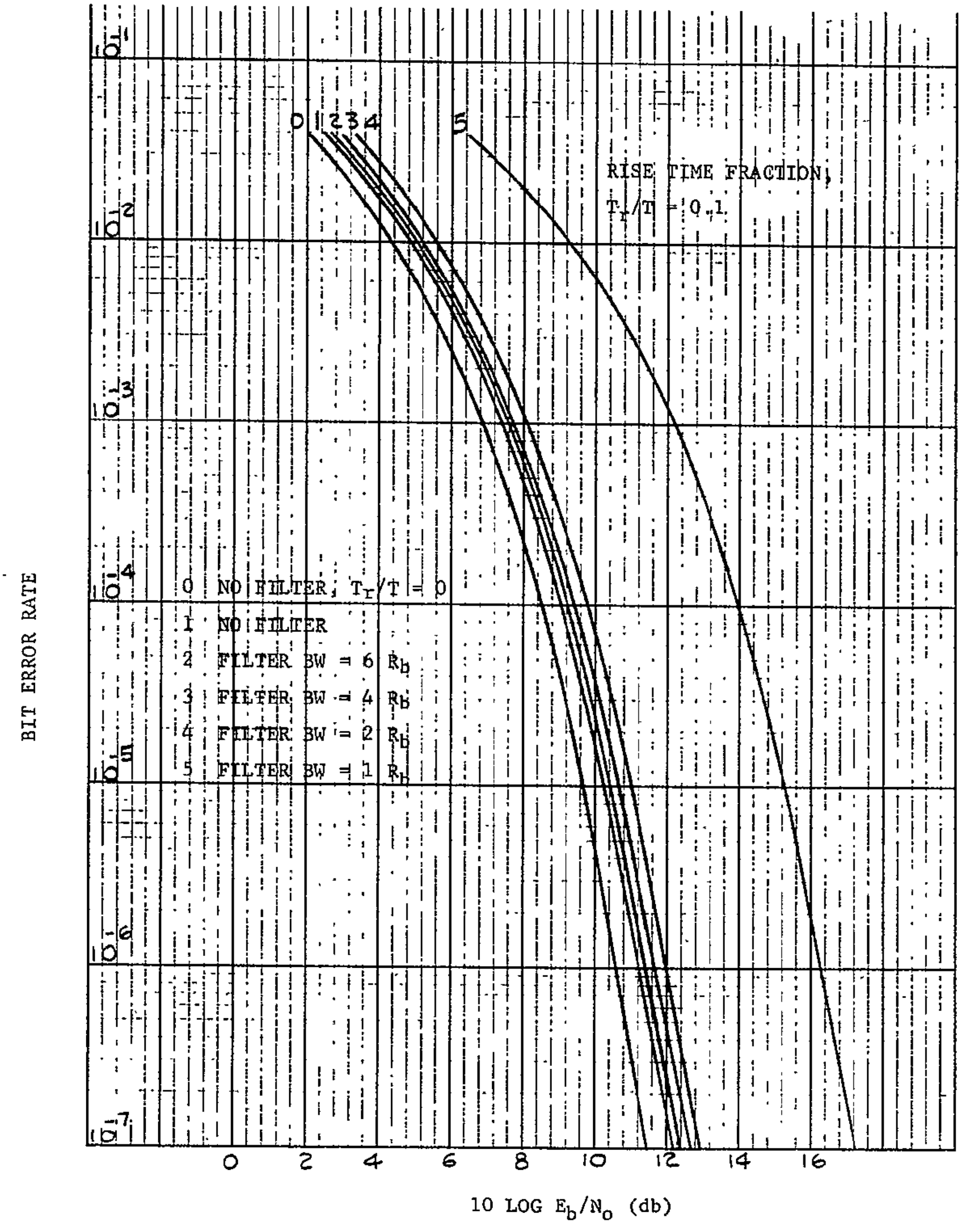


Figure 5-64 Effect of Transmission Channel Filter on Biphase PSK Bit Error Rate

In all cases no energy dumping is used so that some intersymbol interference will exist. For each filter type there exists an optimum choice of a 3-dB cutoff frequency, f_c, relative to the symbol duration, T.

The degradation relative to a perfectly sampled matched-filter is shown in Figures 5-53 and 5-54 for E/N_o equal to 6-dB and 9-dB, respectively, for several non-optimal data filter types. The second order Butterworth provides the least degradation. In the analysis of the various filters, the intersymbol interference out to six succeeding symbols was included.

As the bit rate increases to above 100 Mbps, the matched filter design and operation can become the determining factor of system operation. At low frequency rates, the matched filter is usually an RC circuit or filter circuit followed by a dump circuit which short circuits a capacitor comprising the output impedance of the matched filter. At higher frequency rates, the problems of "moving" capacitance stored charge by the dump mechanism becomes difficult and new types of filters, such as transversal filters, or 3-pole Butterworth filters, using series decision-switched gates can be used. Such gates are associated with very wide band operational amplifiers which must have bandwidths from DC to a frequency high enough to pass all principal components.

The problem then becomes one of addressing both switching rates and very wide bandwidths, and this technology will be addressed in the final report.

An additional technology which will also influence the quality of the recovered bits is the gate waveform which will deteriorate at the high bit rates; this too will be discussed in the final report.

SECTION 6

ACQUISITION - SEARCH - TRACK

6.1 INTRODUCTION

Before an RF transmission link can be established between a sensor satellite and a relay satellite, one satellite must locate, and then acquire and track the other satellite.

Any acquisition and tracking system will be a function of: (1) the capability of the system to acquire in a specified time interval, and (2) the ability of the system to maintain a minimum tracking loss resulting from beam misalignment.

Acquisition and tracking of a data relay satellite by a sensor satellite, using millimeter waves, will encounter the unique problem of requiring very high gain antennas (60 dB gain) for both satellites to provide the necessary link margin (Section 7). This results in very narrow bandwidths for both antennas and acquisition may never be accomplished. This can be overcome, however, by the sensor satellite using a separate low gain wide beamwidth antenna for the initial acquisition phase and then switching to the high gain antenna when acquisition is achieved. The following paragraphs will address the general problems* of search, acquisition, and track, assuming that this narrow-beam/broad-beam capability is present.

6.2 ACQUISITION PROBLEM

To acquire a satellite in space requires precise knowledge of satellite positions versus time. Each satellite is required to be commanded to point its antenna in a given direction at a given time. If the satellite to be acquired appears in an antenna beamwidth, acquisition is accomplished. If the sought after vehicle does not appear in the beamwidth, then a spatial search is required over the uncertainty region.

---

* The calculations to be used are discussed in Section 6, Volume 2.

The establishment of a low altitude sensor-to-synchronous-relay link is a critical acquisition problem because large relative motion of the two vehicles can exist with the sensor moving in and out of view of the relay satellite.

6.2.1 Spatial Acquisition Considerations

Acquisition beamwidth refers to the antenna beamwidth that is useable for acquisition. Normally, beamwidth is given at the 3-dB points - here 0.2 degrees or 3.5 milliradians. Useable beamwidth, however, may be defined by the 10-dB points, or lower, providing there is sufficient signal power to trigger the acquisition device if the signal is received at the beam edge.

An acquisition beamwidth must be defined based on received signal power. The larger this beamwidth can be made the better, since it allows for larger uncertainties in satellite location or antenna pointing angle. For purposes of this discussion the 3-dB beamwidth will be used.

6.2.1.1 Position Uncertainty

A typical relay satellite configuration is illustrated in Figure 6-1. A position uncertainty exists between the relay satellite and the sensor satellite.

Satellite uncertainty volume can be defined, based on accuracies of ground tracking antennas, after the orbit is established. This uncertainty volume is then translated into angle uncertainty of one vehicle with respect to another.

A relay satellite will establish communication with a low altitude satellite shortly after the latter comes into view and will retain the link as long as possible or until a second relay satellite can acquire it. The distance between the two satellites can vary

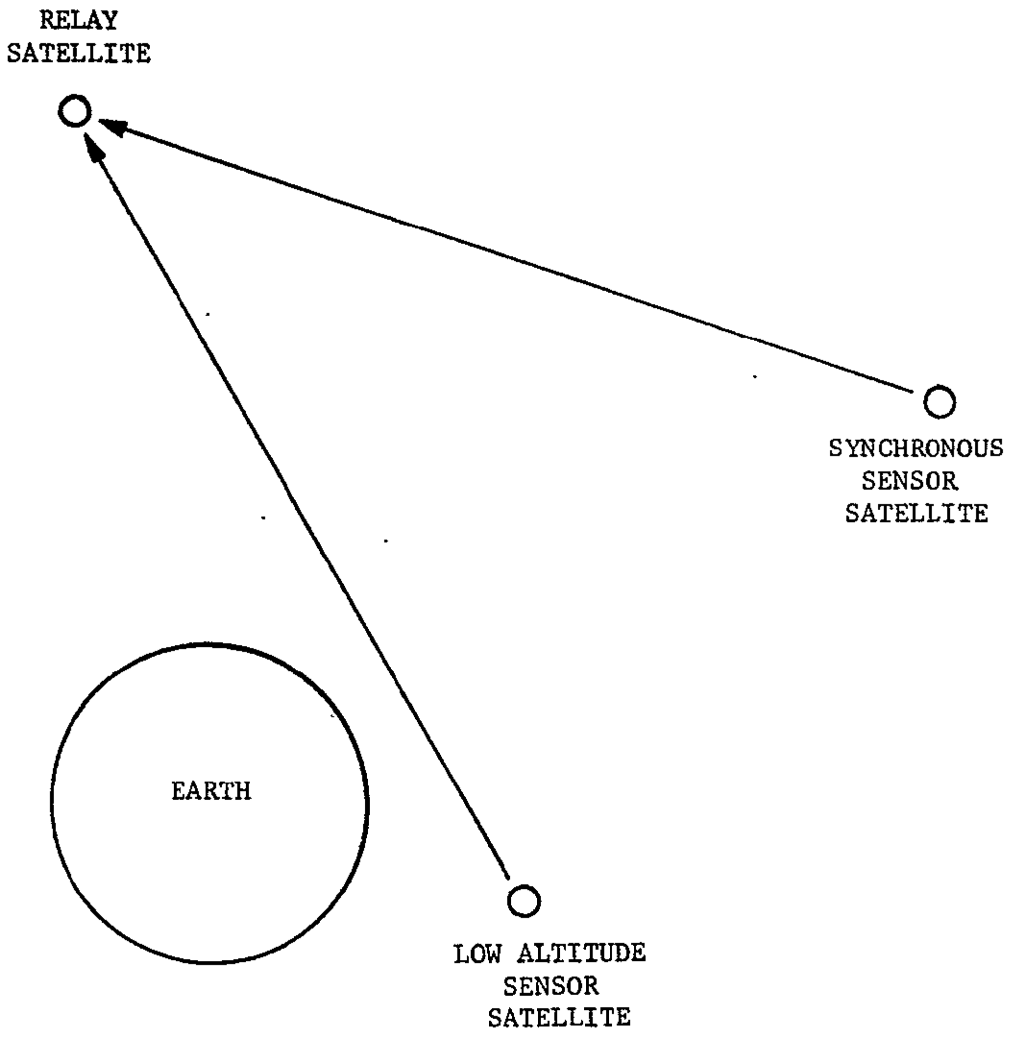

RELAY SATELLITE
SYNCHRONOUS SENSOR SATELLITE
EARTH
LOW ALTITUDE SENSOR SATELLITE

Figure 6-1 Satellite Configuration

considerably during this time interval. For a given position uncertainty, the angular uncertainty will be minimum at the end points and maximum at the mid-point since the range is minimum. Normally, acquisition will not occur at the mid-point; but, for calculation purposes, i.e., worst case angular error, this will be assumed to be the case.

At the point of closest approach the distance between the relay and low altitude vehicle will be:

$$d = 19{,}327 - A \text{ in nautical miles}$$

where A = low altitude satellite height above earth

Angular uncertainty is then

$$\theta_1 = \tan^{-1}\left(\frac{p}{d}\right)$$

where p = position uncertainty in nautical miles

Figure 6-2 is a plot of position uncertainty and altitude, A, versus angular uncertainty. Since both the low altitude vehicle and the relay will have a position uncertainty, the total worst case angular error will be the sum of the two.

Figure 6-3 is a plot of position uncertainty versus angular uncertainty for vehicle separation of 30,000 n.m. and 40,000 n.m.

6.2.2 Short Term Vehicle Stability

Three attitude control methods determine short term stability.

a. Three-Axis Gas Control: The three axis gas control cycles between preset limits about all three axes. If the satellite is heavy and the thrust small, the short term stability can be held to less than 1°/min. (0.00003 rad/sec).

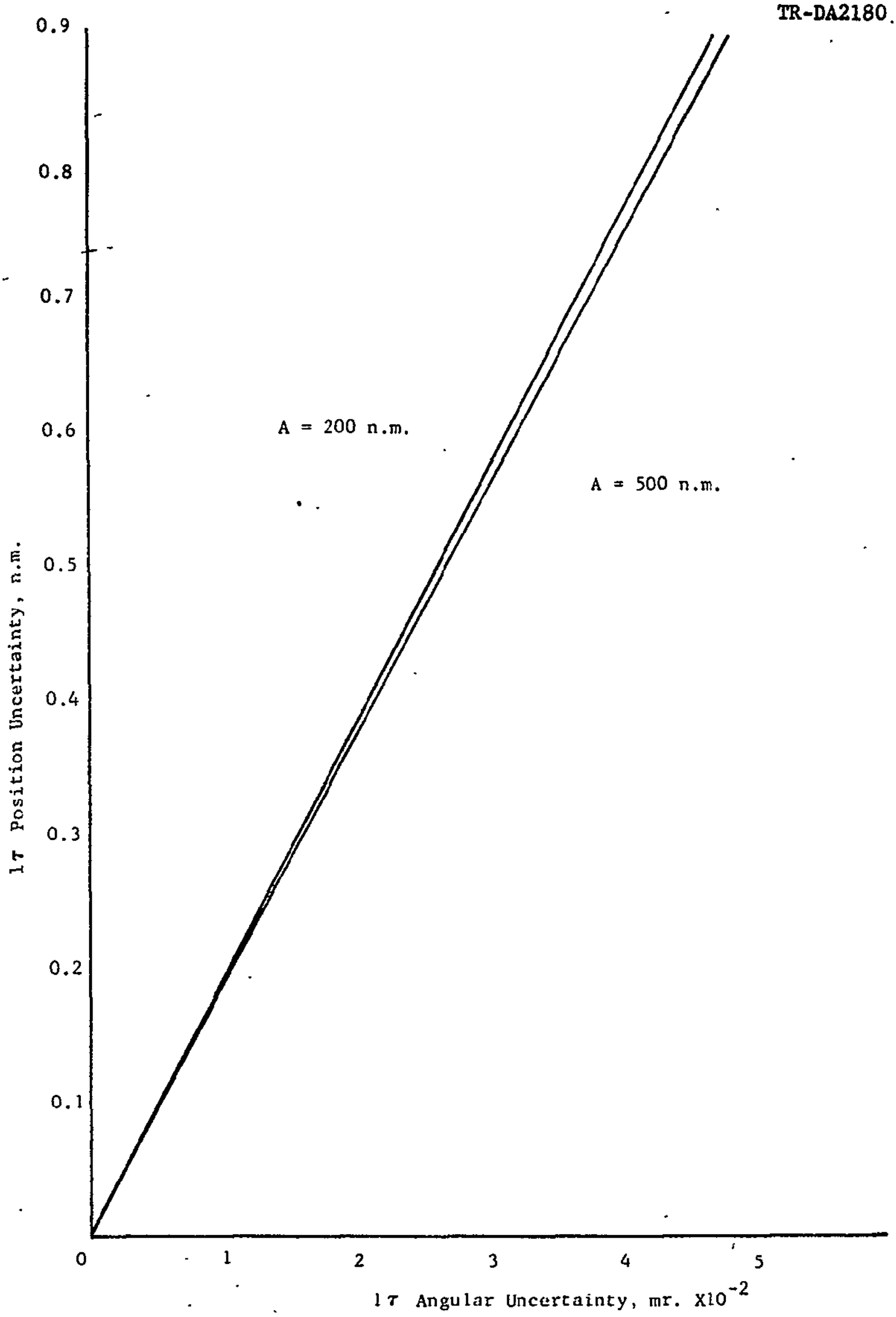


Figure 6-2 Angular Uncertainty vs Position Uncertainty and Altitude

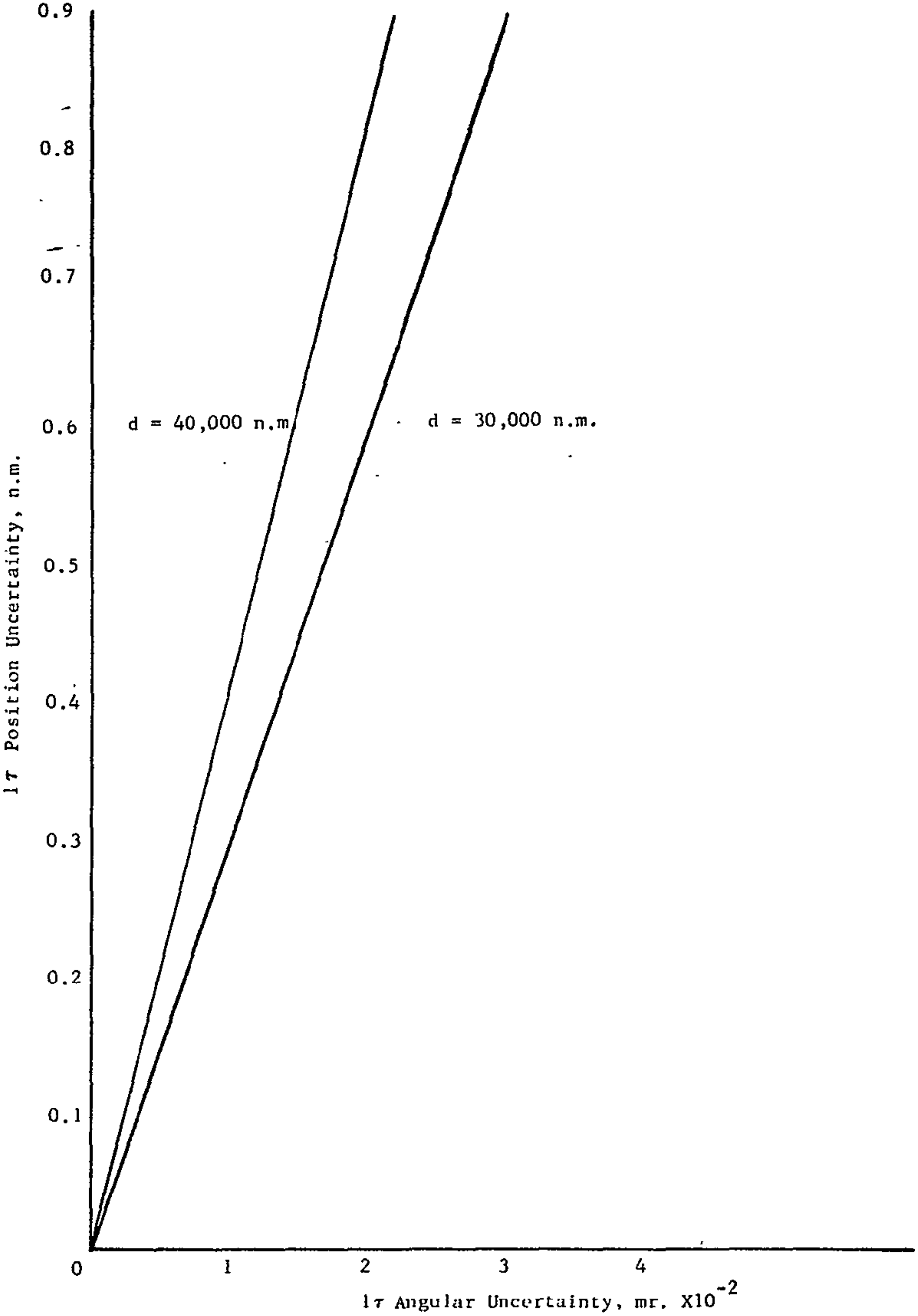


Figure 6-3 Angular Uncertainty vs Position Uncertainty and Synchronous Satellite Separation; d

b. Variable Speed Flywheel: Variable speed flywheels provide very good short term stability. Using gas control for gross corrections only, flywheels can provide a short term stability on the order of 3×10^{-6} rad/sec when the gas control is not thrusting.

c. Spin Stabilization: The spin stabilized satellite provides the best short term stability, of the order of 0.1×10^{-7} rad/sec except when thrusting, to correct precession or orbit parameters.

6.2.3 Attitude Control Subsystem

Attitude control subsystems require earth, sun, and star tracking sensors:

a. Earth Sensors: Earth sensors are the least accurate of the three types of sensors. Most earth sensors are designed to sense the energy in the 14 to 16 micron CO_2 band as the altitude of the measured CO_2 is more stable than the water vapor band. The absolute one reading accuracy of a good earth sensor is approximately 0.00175 radians. After averaging multiple readings an accuracy of 0.00087 radians can be obtained.

b. Sun Sensors: Sun sensors are more accurate than earth sensors. The accuracy is limited to 0.00002 radians. This limitation is due to solar flares and sun spots. Obviously, the sun sensor is inoperative when it is in the earth's shadow.

c. Star Trackers: Star trackers are the most accurate of the sensors. The accuracy is greater than 5×10^{-6} radians and is limited by the drive brain and encoder accuracy.

Star trackers are large and heavy, and require a sophisticated thermal design. A sixteen bit encoder seems reasonable related to size, weight, and power (12-bit encoders are available) and gives a resolution of 0.1 mr. However, due to a possible ambiguity in the least significant bit, an absolute accuracy of 0.2 mr is used.

6.2.4 Spatial Acquisition Accuracy

Acquisition accuracy is essentially determined by the type of attitude control employed in the spacecraft. The advantage of the more sophisticated star tracker attitude control system can be illustrated here through a calculation of overall acquisition accuracy to show that a spatial search is not required to acquire the signal. A spin stabilized satellite with a despun antenna is assumed.

Total acquisition error, in Table 6-1, is obtained by addition of the individual contributors. Combining errors in this fashion may yield a pessimistic result since all sources must generate errors in the proper sense so that the resultant is the arithmetic sum.

Errors are tabulated as either a bias or noise error:

- Bias Error: An error which is non-varying over short periods of time but which cannot be compensated.
- Noise Error: A rapidly varying (jitter) error.

Total error must be less than 1.7 milliradians for a 3.5 mr, 3-dB beamwidth if acquisition is to be accomplished without spatial search.

It is seen that total error will, in fact, permit the antenna to be directed with sufficient accuracy so that a spatial search is not necessary to acquire the vehicle.

TABLE 6-1

ACQUISITION ERROR BUDGET

ERROR SOURCE	ERROR	
	Bias (mr)	Noise (mr)
1. Position (see angular uncertainty vs position uncertainty Figure 6-2)	0.04	
2. Spacecraft		
a. Attitude		
1. Short Term Stability		0.1
2. Attitude Control	0.2	
b. Gimbal System		
1. Gimbal Alignment	0.5	
2. Axes Orthogonality	0.5	
c. Position Reference		
1. Encoder Accuracy and Resolution	0.2	
d. Antenna		
1. Boresight Shift	0.2	
Total Error (worst case)	1.64 mr	0.1 mr

Since the tabulated errors are slanted toward the high side and usable beamwidth is taken as 3-dB instead of, perhaps, 10-dB, a margin actually exists in the above calculation.

The point to be made, here, is that acquisition is feasible without spatial search. The outcome would have been reversed had an earth sensor been employed instead of a star tracker as the spacecraft attitude reference.

6.3 SPATIAL SEARCH

A spatial search by two highly directive antennas on two separate spacecrafts to locate each other is a complicated procedure. Each vehicle is faced with the problem of locating the other in a given time interval. If both vehicles initiate search independently at the same time, one may be searching one end of the search sector while the other is at the opposite end with the result that the search is prolonged and, possibly, the two never acquire each other. A search sequence must be established, therefore, that permits full coverage of the sector in the allocated time and in such a manner that assures essentially 100% probability of successful acquisition.

6.3.1 Search Sequence

For spatial search to be successful, complete cooperation is required between two satellites. At a given time, one satellite positions its antenna in a predetermined direction and holds this position long enough to permit the second satellite to sweep the complete search sector. At the end of one search period, the first satellite repositions its antenna for the second search and so on until acquisition is accomplished. The above sequence assumes the first satellite to be radiating a signal and the second to be silent.

Acquisition by the second satellite immediately actuates its transmitter so that a signal is radiated to permit the first satellite to acquire.

The time required to complete the sweep depends on sector width, spacecraft antenna slew limitations, and receiver detection capability. The search may be in the form of a spiral around center point or in the form of a raster scan.

The number of times the acquisition cycle must be repeated depends on the uncertainty region and on signal level. If the region is sufficiently small and the signal power sufficiently high to provide substantial margin, then perhaps only one complete search will be required to produce successful acquisition.

6.3.2 Signal Detection

The second phase of the acquisition problem consists of detecting the signal when it appears in the antenna beam and indicating its presence to the antenna servo/drive so that the antenna can be stopped and autotrack initiated.

One method used in detecting the presence of the input signal is shown in Figure 6-4, consisting of a simple amplitude detector, filter and threshold device.

Signal presence is indicated when the integrator output exceeds the preset bias level. Bias level is set by signal and noise levels expected in actual operation. Other factors to be considered are the probability of detection and probability of false alarm. A graph is presented in Figure 6-5 showing dwell time and false alarm time versus signal to noise ratio.

Dwell time is a property of the integrator which imposes restraints on antenna search speed and/or beamwidth such that the signal is present in the beam for a minimum time period. The probability of detection, Pd, is taken as 0.98. Trade-offs must be made between signal-to-noise ratio, dwell time and false alarm time in the course of system design.

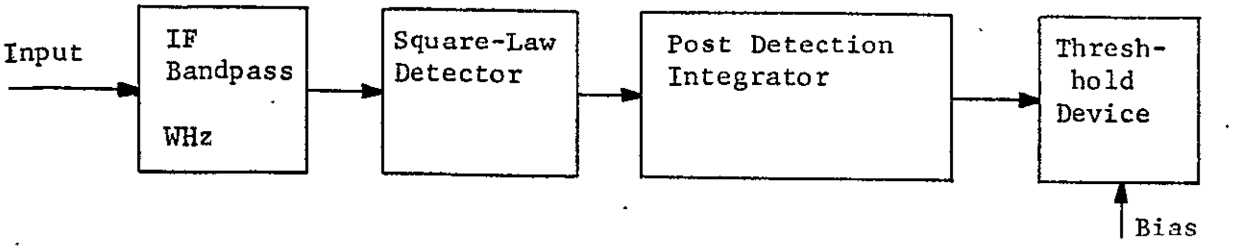

Input
IF Bandpass
WHz
Square-Law Detector
Post Detection Integrator
Thresh-hold Device
Bias

Figure 6-4 Simplified Signal Detection Circuit

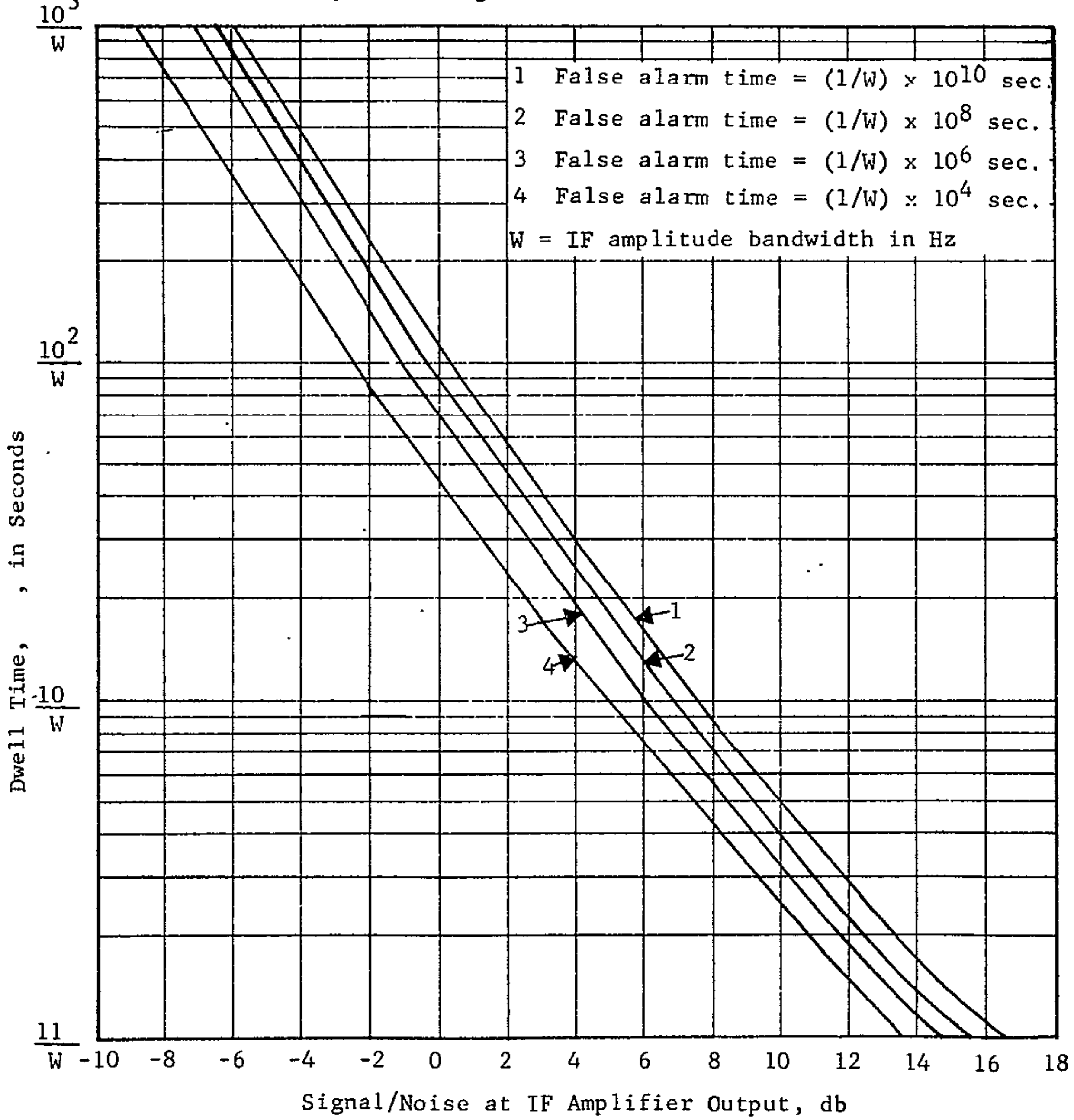


Figure 6-5 Dwell Time versus Signal/Noise, $P_d = 0.98$

6.3.3 Considerations of Signal Power Level

In situations where signal power is not sufficient to permit use of wideband tracking techniques, narrowband frequency or phase-lock techniques are utilized. Receiver complexity is increased in the process but performance is improved.

Signal detection in a phase-lock receiver is accomplished in frequency as well as in phase during the acquisition sequence by sweeping the VCO (voltage controlled oscillator) across the frequency uncertainty region.

Another variable is therefore added to the acquisition problem, i.e., frequency search rate is required to be compatible with the spatial search rate. This does not, however, impose major problems on the system, since the desired frequency sweep rate can be obtained rather simply in the receiver by providing a separate acquisition loop bandwidth, if necessary.

6.3.4 SNR and Frequency Sweep Rate

For the phase-lock receiver, signal detection is indicated when the phase-lock loop locks to the incoming signal. An in-lock indication removes the frequency search in the receiver and spatial search at the antenna. The transition from the search phase to tracking is accomplished instantaneously when the loop locks. The only constraint is that the antenna search then be stopped.

Maximum frequency sweep rate is dependent on loop parameters and on loop signal-to-noise ratio. An empirical equation that provides a 90% probability of acquisition is given as follows:

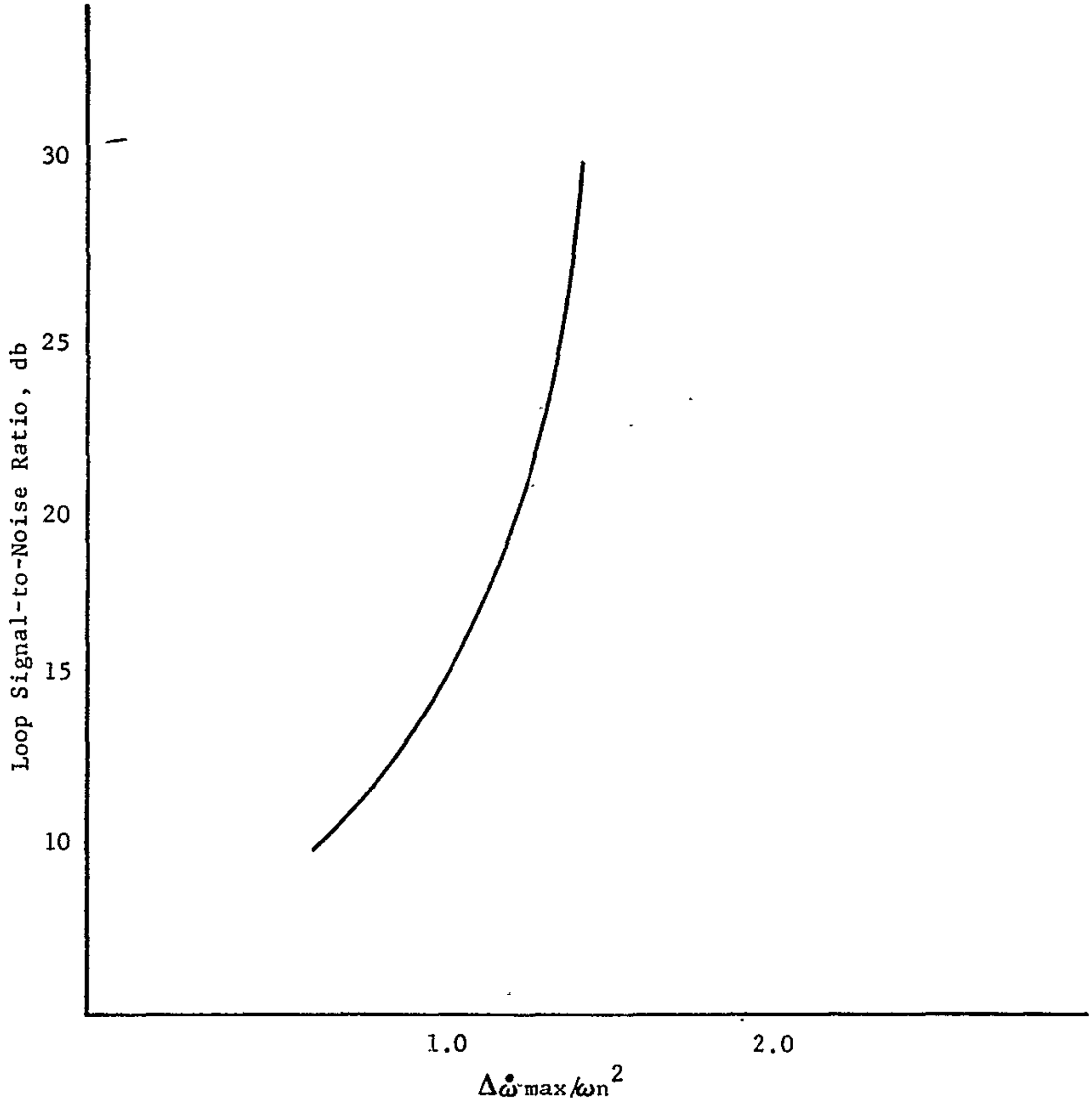


Figure 6-6 Loop SNR versus $\Delta\dot{\omega}\text{max}/\omega n^2$

$$\Delta \dot{\omega} \max = \left[1 - \frac{(\mathrm{SNR})^{-1/2} \quad K \quad \omega_o^2}{1 + d} \right] = \text{maximum sweep rate}$$

SNR = Loop signal-to-noise ratio

K = Factor due to limiter

ω_o = Loop natural frequency

d = Factor depending on damping < 0.05 for damping factor of 0.7

Requiring $\Delta \dot{\omega} < \Delta \dot{\omega}$ max will increase the probability of signal detection.

A plot of $\Delta \dot{\omega} \max/\omega_o^2$ versus loop SNR is given in Figure 6-6. Limiter bandwidth is assumed to be ten times the threshold loop bandwidth for the calculation.

6.4 TRACKING

Once the signal from the satellite is acquired, the antenna must be placed in an autotrack mode to enable the receiver to constantly receive the incoming signal and to maintain the antenna pointed properly for the transmitted signal. The block diagram of the tracking system is given in Figure 6-7.

A feed or equivalent generates a sum and difference channel in two orthogonal axes. Tracking information is contained in the difference channel amplitude which is processed by the receiver to generate error voltages in each axis for the antenna servos.

One of the primary concerns of the tracking system is that it track the incoming signal sufficiently close to the antenna beam peak so that the gain loss is negligible.

A criteria is established that tracking error produce 0.5 dB or less degradation in antenna gain.

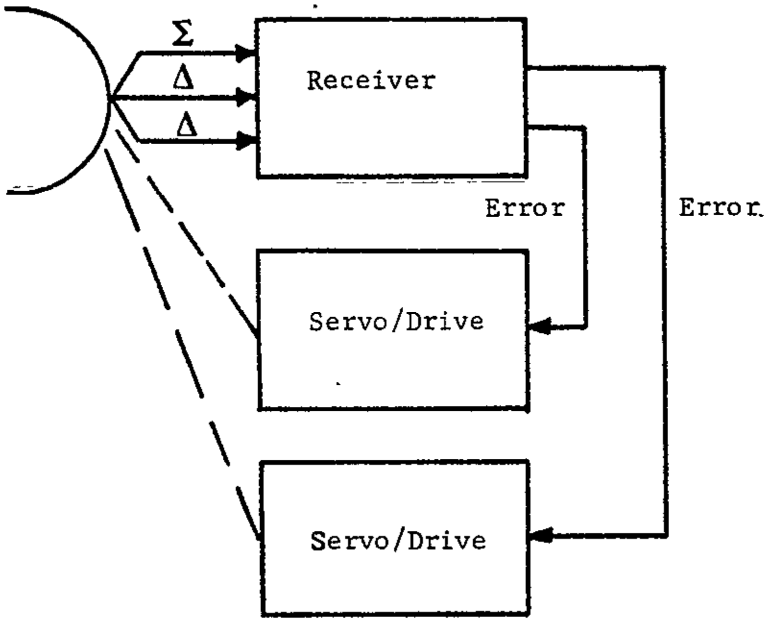

Σ
Δ
Δ
Receiver
Error
Error
Servo/Drive
Servo/Drive

Figure 6-7 Tracking System, Block Diagram

6.4.1 Tracking Receiver Configurations

Frequency acquisition (location of the incoming signal in frequency) is accomplished concurrently or after spatial acquisition is complete. In the absence of a spatial search, frequency acquisition occurs after the antenna is directed at the signal source. When a spatial search is required, then the two events must occur simultaneously. The complexity of frequency acquisition depends on the type of receiver employed.

A wideband receiver covering the total frequency uncertainty region (oscillator drift and doppler effects) will require very little added capability to perform the function. A narrowband receiver, however, will require a search capability to cover the frequency uncertainty range.

A coherent tracking receiver will provide:

a. Autotrack capability for a two-axis tracking antenna utilizing simultaneous-lobing tracking techniques.

b. Capability of demodulating telemetry data that is carried on the tracking signal.

6.4.1.1 Wideband Tracking

One form of tracking receiver is the cross-correlation receiver.

Error signals for the servos are obtained by forming the product of the sum channel signal with each of the difference channel signals to produce, after filtering, a dc error voltage with amplitude proportional to the amplitude of the difference channel signal and polarity indicating the direction or sense of the error.

The receiver bandwidth must be broad enough to encompass frequency variations in the input signal due to oscillator drift and doppler shift. At the frequencies of interest the doppler shift is on the order of a megacycle when tracking the sensor vehicle; therefore, the receiver bandwidth must be of the same order of magnitude. Noise in the error channels introduces tracking jitter which must also be limited to avoid tracking error.

6.4.1.2 Narrowband Tracking

The wideband receiver shown in Figure 6-8 can be modified into a narrowband configuration by incorporating automatic frequency control (AFC) to maintain the signal centered in the passband. The width of the passband can be reduced with corresponding decrease in noise power and increased signal-to-noise ratio.

Another form of the narrowband receiver is the phase lock receiver. A block diagram is given in Figure 6-9. In this case, phase lock is established on the sum channel signal to maintain the signal in each channel centered in the passband. Since the loop tracks the carrier with zero frequency error, the IF apertures can be made quite small; a 6-dB signal-to-noise ratio in a 1 MHz bandwidth will require 30-dB more signal power than in a 1 kHz phase lock bandwidth assuming the noise density is the same in each case.

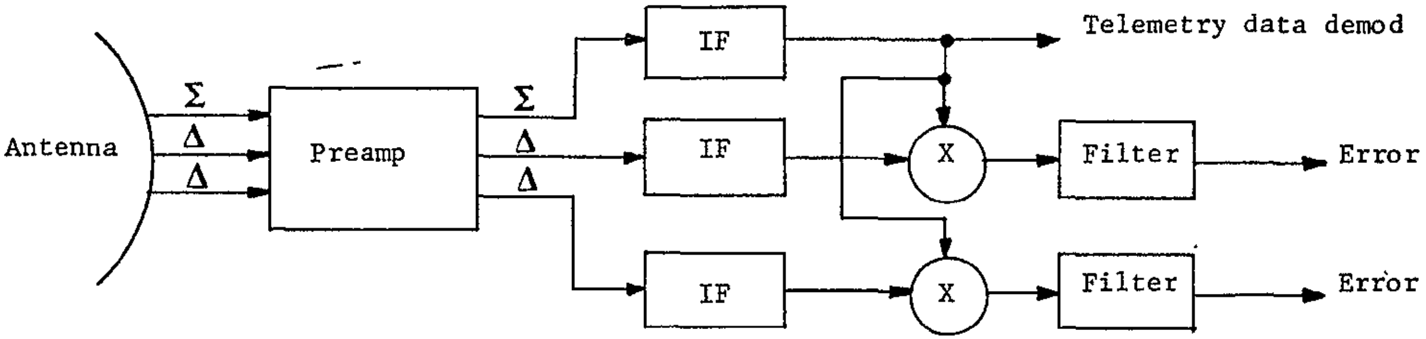

Antenna
Σ
Δ
Δ
Preamp
Σ
Δ
Δ
IF
IF
IF
Telemetry data demod
X
X
Filter
Filter
Error
Error

Fig. 6-8 Simplified Wideband Tracking Receiver

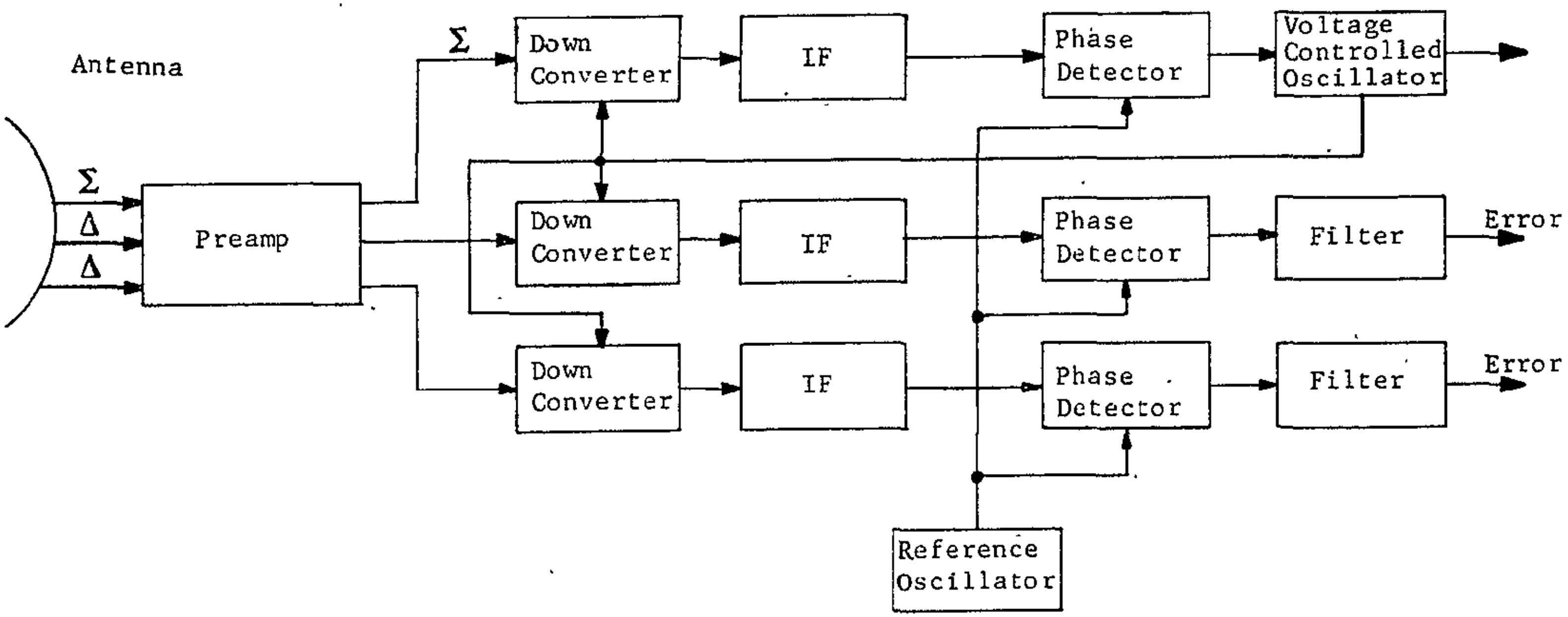

Antenna
Σ
Δ
Δ
Preamp
Σ
Down Converter
Down Converter
Down Converter
IF
IF
IF
Phase Detector
Phase Detector
Phase Detector
Voltage Controlled Oscillator
Filter
Filter
Error
Error
Reference Oscillator

Fig. 6-9 Phase-Lock Angle Tracking Receiver, Block Diagram

6.5 PROPAGATION TIME ERROR

When two satellites are tracking and are in communication with each other, each satellite tracks the incoming signal and radiates its signal to the other through the same antenna. Thus, as Satellite B receives the signal from Satellite A the antenna (Satellite B) will be pointed slightly behind Satellite A. This will cause the transmitted signal (B to A) to be propagated from a point, on the beam, which is off the beam peak and hence, cause an error. The magnitude of the propagation error can be evaluated as follows for a distance of 19,027 nm between Satellite A and Satellite B.

$$\text{Time delay } \tau = \frac{19{,}027 \text{ nm}}{c}$$

where c = velocity of light. The Satellite A will move, d, with velocity, v, in time, τ,

thus, the angle error, θ can be derived from

$$\tan \theta \approx \theta = \frac{v\tau}{19{,}027} = 0.025 \text{ mr}$$

where v = 4.1 n.m./sec.

The worst case error will occur when the velocity vector of one satellite is perpendicular to a line connecting the two satellites.

6.6 TOTAL ERROR

Total tracking error can now be computed by combining the bias and noise errors obtained in the above paragraphs and tabulated as in Table 6-2 for a tracking error budget.

For the antenna beamwidth of 3.5 mr, a degradation of 0.5 dB in gain corresponds to an angle of 0.66 mr. Hence, the total error must be less than 0.66 mr to meet desired tracking accuracy.

TABLE 6-2

TRACKING ERROR BUDGET

Source	(Milliradian) Bias Peak	(Milliradian) Noise 1 σ
Thermal Noise (15 dB SNR)		0.02
Phase Shift (Pre = 6°) (Post = 10°)	0.02	
Amplitude Unbalance (0.5 dB)	0.07	
Beamwidth Unbalance (1.1:1)	0.04	
Temperature	0.10	
Servo Lag	0.10	
Servo Bias	0.10	
Propagation Time Delay	0.03	
Total	0.46	0.02

Bias errors above are summed by straight addition. It should be noted that errors combined in this fashion yield a very pessimistic result as the sense of each error must be such that they combine to produce a composite error in a given direction. Such an occurance is highly improbable and the total error will actually be much less.

The above worst case combined error is seen to be well within the allocated value of 0.66 mr.

SECTION 7

THE MILLIMETER WAVE LINK

7.1 THE MILLIMETER WAVE LINK IN IMAGE TRANSMISSION

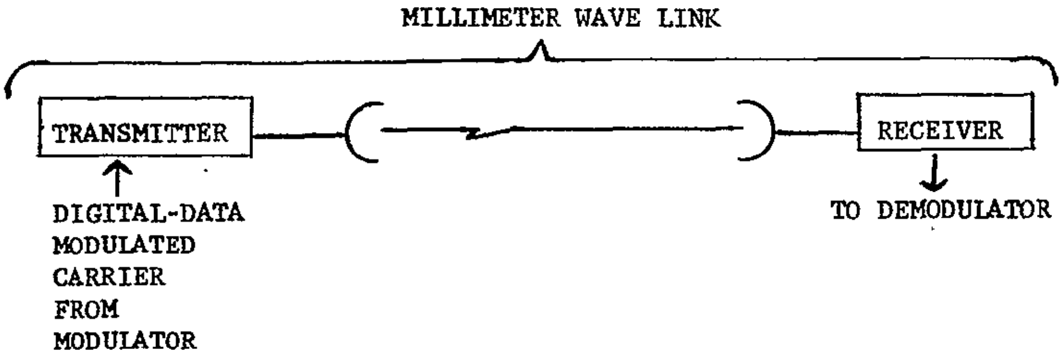
The millimeter wave link is the medium which is used to transmit image information across space from a transmitting terminal to a receiving terminal as shown in Figure 7-1. This link will determine the amount of margin required to transmit a given information data rate at a particular frequency over a given distance. Of prime inportance are:

- Transmitter Power
- Transmitter Antenna Gain
- Space Loss
- Receiving Antenna Gain
- Receiving System Noise Figure

These various parameters will be addressed with particular reference to the particular problems incurred in transmitting data rates of 40 to 400 Mbps at the millimeter wave frequencies from 35 to 90 GHz.

7.2 LINK BUDGET IN MILLIMETER WAVE SPACE COMMUNICATIONS

The link budgets of a spacecraft to spacecraft link, or spacecraft to ground link are vital for determining how much data can be transmitted from one point to another with given antenna sizes, transmitting terminal ERP and receiving terminal gain and system temperature. The following paragraphs discuss how a millimeter wave link budget is computed, and nomographs which relate data rates from 40 to 400 Mbps, frequencies of 35, 60, 75, and 94 GHz for various antenna sizes and receiver noise figures over worst case distances.


MILLIMETER WAVE LINK
TRANSMITTER
RECEIVER
DIGITAL-DATA MODULATED CARRIER FROM MODULATOR
TO DEMODULATOR

PRINCIPAL LINK PARAMETERS

- Transmitter Power
- Transmitter Antenna Gain
- Space Loss
- Pointing Error
- Receiving Antenna Gain
- Receiver Noise Figure

Figure 7-1 Basic Millimeter Wave Data Transmission Link

7.2.1 Link Budget Analyses for Sensor Satellite to Relay Satellite

Consider the specific conditions defining the transmission between a sensor satellite and a relay satellite.

Freq = 60 GHz
Data Rate = 100 - 400 Mbs
Range = 24,500 N Mi
Data Error Rate = 1×10^{-5} errors/bit

The link budget is then determined as follows:

1. Tx Pwr (tradeoff parameter) P^T dBm

2. Insertion loss of Tx components -2 dB
   (Includes filter combiner, polarizer, diplexer, waveguide RF coupling)

3. Transmit antenna gain (tradeoff parameter) - See Figure 7-2

$$\text{Gain} = \frac{4\pi A \eta}{\lambda^2}$$

$$\text{where } A = \frac{\pi D^2}{4}$$

D^T = Tx Antenna Aperture dia. (feet)
η = Antenna eff. ($\eta = 0.75$)
λ = Wavelength (feet)

$$\text{Gain} = \frac{\pi^2 \eta D^2}{\lambda^2} \qquad \eta = \frac{984}{f\ (\text{mc})}$$

$$\text{Gain (dB)} = 10 \log \frac{\pi^2 \eta}{\lambda^2} + 20 \log D^T$$

$$\text{Gain (dB)} = 10 \log \frac{(9.9)(.75)}{2.69 \times 10^{-4}} + 20 \log D^T$$

$$= 10 \log (2.71)(10^4) + 20 \log D^T$$

$$\text{Gain (dB)} = 44.3 + 20 \log D^T$$

4. Space Loss, see Figure 7-3.

$L = 36.6 + 20 \log F + 20 \log d$	F = MHz
$L = 36.6 + 20 \log 60 \times 10^3 + 20 \log 24500 \times 1.15$	d = miles
$L = 36.6 + 95.6 + 89$	
$L = -221.2$	-221.2 dB

5.	Tracking Loss (both)	-0.5 dB
6.	Polarization Loss (both)	-0.3 dB
7.	Oxygen Absorption	-0.1 dB
8.	Receive Antenna Gain (tradeoff)	44.3 dB $+ 20 \operatorname{Log} D^R$
9.	Insertion Loss of Receiver Components (Polarizer diplexer, filter separator and waveguide)	-1.5 dB
10.	Link Margin	-3 dB

11. $S = P^T + 44.3 + 20 \log D^T -2 -221.2 -.5 -0.3 -0.1 + 44.3 + 20 \log D^R -1.5 -3.0$

$S = P^T + 20 \log D^T D^R -140$ dB Power (S) at Receiver Input

12. Noise Density (N_o) at receiver input.

$N_o = KT_s = K$ to F where T_s is the system temperature,

T_o is the reference temperature = 290°K

and F is the equivalent system noise figure.

$N_o = KT_o F$, $KT_o = -174$ dBm

N_o (dB) = -174 dBm $+ 10 \log F$

13. Received Signal to Noise Density Ratio (S/N_o)

$$S/N_o = \frac{P^T + 20 \log D^T D^R -140 \text{ dB}}{-174 + 10 \log F}$$

$$S/N_o = PT + 20 \log D^T D^R + 34 \text{ dB} -10 \log F$$

14. Link Tradeoff Equation

Received S/N_o = Required S/N_o

$$P^T + 20 \log D^T D^R + 34 \text{ dB} - F = 11 + 10 \log b$$

$$P^T + 20 \log D^T D^R + 34 - F = 11 + 65$$

$$P^T + 20 \log D^T D^R - F = 42 \text{ dBm}$$

$$P^T + 20 \log D^T D^R = 42 + F \text{ (dB)}$$

15. Plot the system equation for the ranges of $P^T D^T D^R$ and F that are feasible in millimeter wave communications. See Figures 5-5 and 5-6.

7.2.2 Link Budget Nomographs

The link budget analysis approach of the preceeding paragraph has been used to develop the nomographs of Figures 7-4 through 7-7 based on the baseline link budget computations of Table 7-1. These nomographs are computed to permit relating satellite transmitter power required to transmit various data rates at 35, 60, 75 and 94 GHz, for various transmitting antenna sizes, to the receiver antenna size and receiver noise figure.

The various secondary impairments of the link are also included, as identified in Table 7-1.

7.3 MILLIMETER WAVE TRANSMITTERS

A critical parameter in developing Effective Radiated Power (ERP) from the transmitting satellite is the amount of transmitter power which can be developed in the millimeter wave frequencies from 35 to 90 GHz.

TABLE 7-1

BASELINE LINK BUDGETS

1. Transmission Frequency - f	Hz	35 G	60 G	75 G	94 G
2. Data Rate	B_{ps}	100 Mbps	100 Mbps	100 Mbps	100 Mbps
3. Baseband Bandwidth - B_B	Hz				
4. Receiver IF Bandwidth - B_{if}	Hz				
5. Transmitted Signal Bandwidth - B_T	Hz				
6. RF Channel Bandwidth - B_{CHAN}	Hz				
7. Type Modulation		Biphase or QPSK with Diff. Encod.			
8. Type Detector					
9. Transmitter Power Output	Watts	13	13	13	13
10. Transmitter Power Output-Av.Unmod.Carrier	dbW	11.2	11.2	11.2	11.2
11. Line and Feed Losses	-db	-2.0	-2.0	-2.0	
12. Peak Antenna Gain 10 Log $\left(\frac{\pi\ df}{c}\right)^2 N$	db	53.4	58.1	60.0	62.0
Antenna Dimensions η = 70%	ft(m)	5'	5'	5'	5'
Coverage	Deg.				
13. ERP/Channel (Sum of 10, 11, 12)	DbW	62.6	67.3	69.2	71.2
14. Propagation Loss (37.81 + 20 Log $d_{n.mi}$ + 20 Log f_{MHz})	-db	215.5	-220.2	-222.1	-224.1
15. Polarization and Multipath Losses	-db	-0.3	-0.3	-0.3	-0.3
16. Atmos. Att. and Misc. Losses	-db	-	-	-	-
17. Xmit. Antenna Pointing Loss $\left[\text{Loss}_{dB} = 12\left(\frac{\theta D.f}{70.C}\right)^2\right]$	-db	-0.5	-0.5	-0.5	-0.5
18. Rcv. Antenna Pointing Loss	-db	-0.5	-0.5	-0.5	-0.5
19. Up Link Noise Contribution	-db	-	-	-	-
20. Transmission Losses (14 + ... +17	dB	-216.8	-221.5	-223.4	-225.4
21. Rcv. Ant. Gain (Pk)	dB	6.7' 56.0	3.9' 56.0	3.1' 56.0	2.5' 56.0
22. Rcv. System Line & Feed Losses	dB	-2.0	-2.0	-2.0	-2.0
23. Rcv. System Temp. Factor	^oK-db	NF=9dB -33.1	-33.1	-33.1	-33.1
24. Receiving System G/T	dB	+20.9	+20.9	+20.9	+20.9
25. Boltzmann's Constant	dBW/Hz/^oK	228.6	228.6	228.6	228.6
26. Net C/No (13 + 20 + 24 + 25)	dB-Hz	95.3	95.3	95.3	95.3
27. 10 Log Data Rate	bits/sec	80.0	80.0	80.0	80.0
28. Equip. Variance from Theoretical	dB	2.5	2.5	2.5	2.5
29. E_b/n/No (26 - 27 - 28)	dB-Hz/bits/sec	12.8	12.8	12.8	12.8
30. Required E_b/n/No for BEP=10^{-5}	dB-Hz/bits/sec	9.8	9.8	9.8	9.8
31. NET LINK MARGIN (29 - 30)	dB	3.0	3.0	3.0	3.0

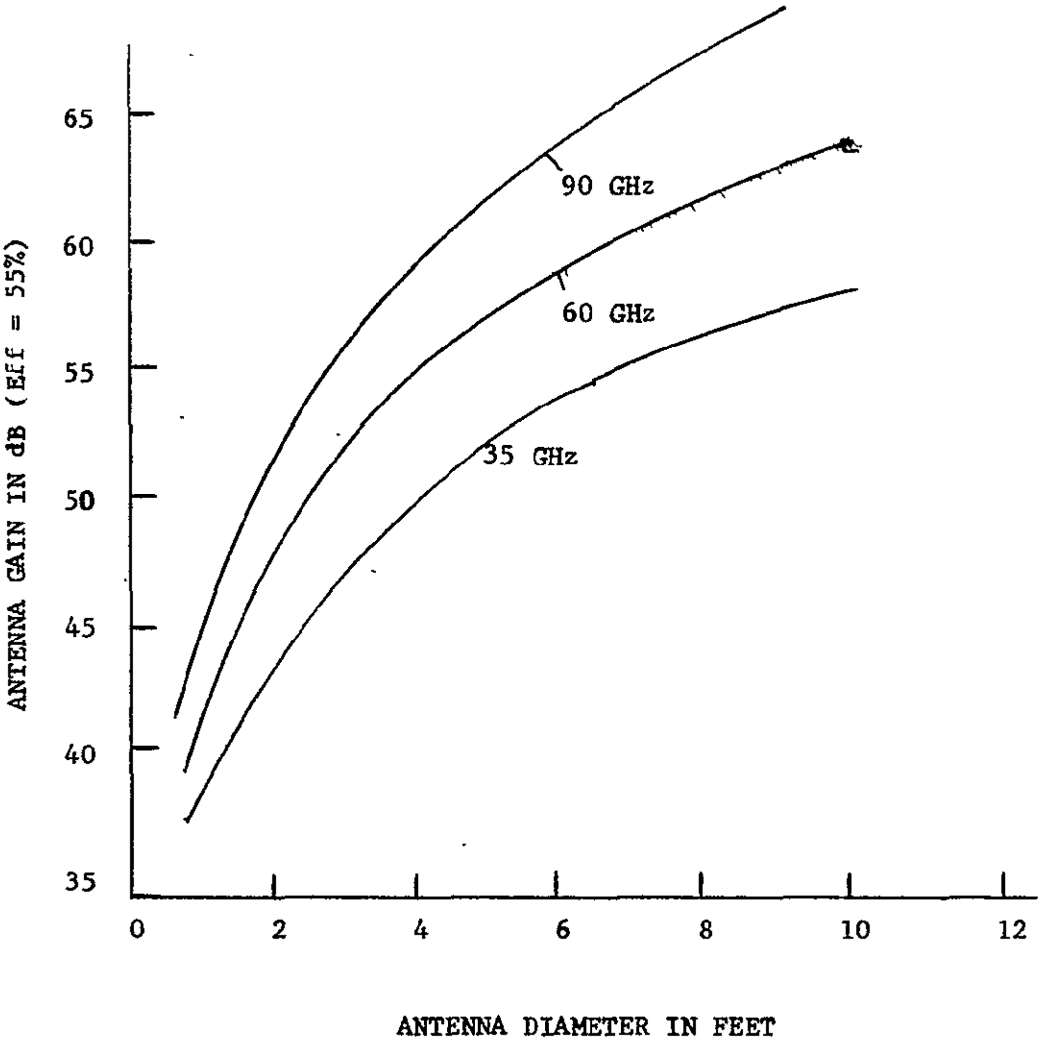


Figure 7-2 Antenna Gain versus Antenna Diameter

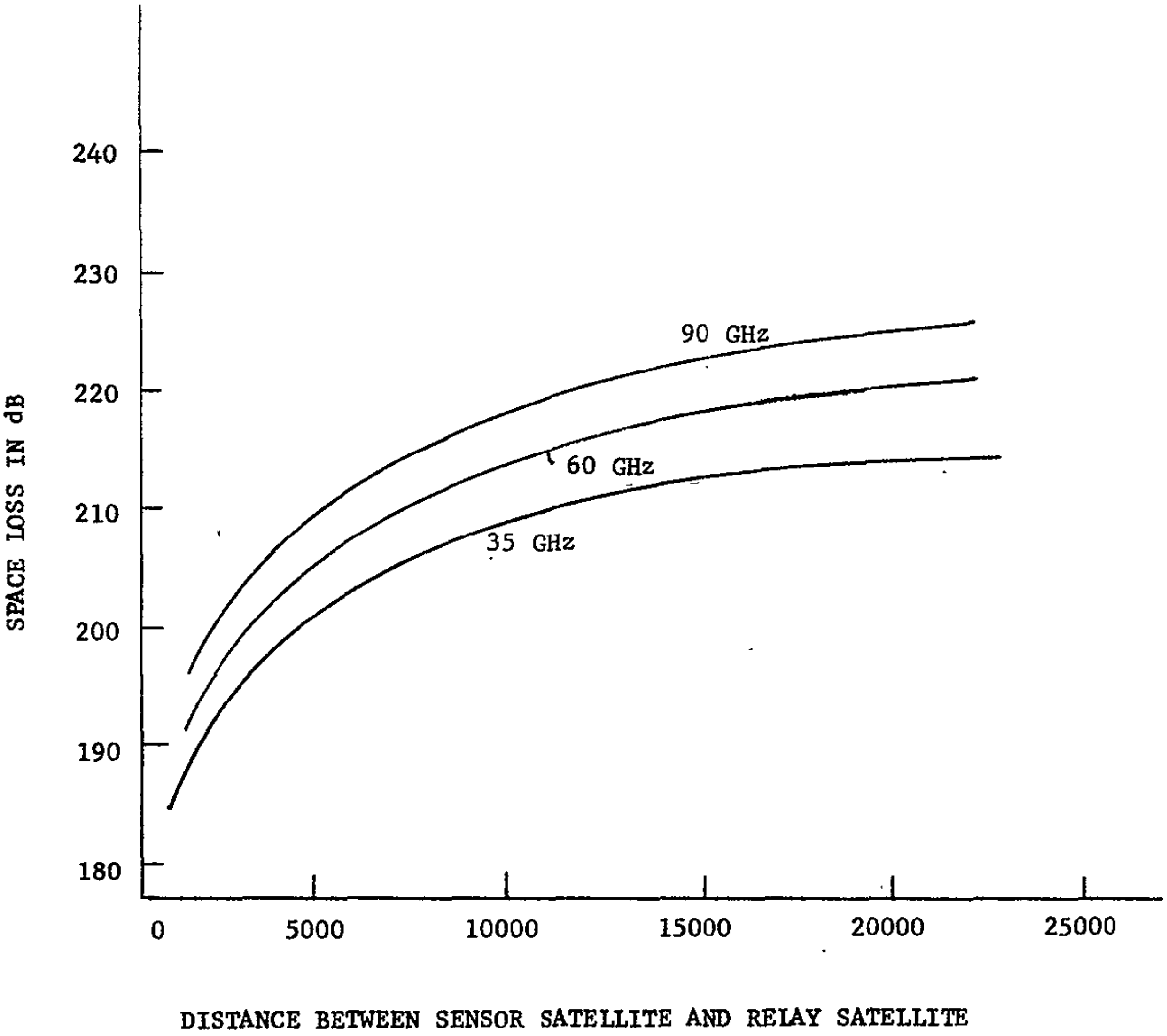


Figure 7-3 Space Loss versus Distance at Millimeter Waves

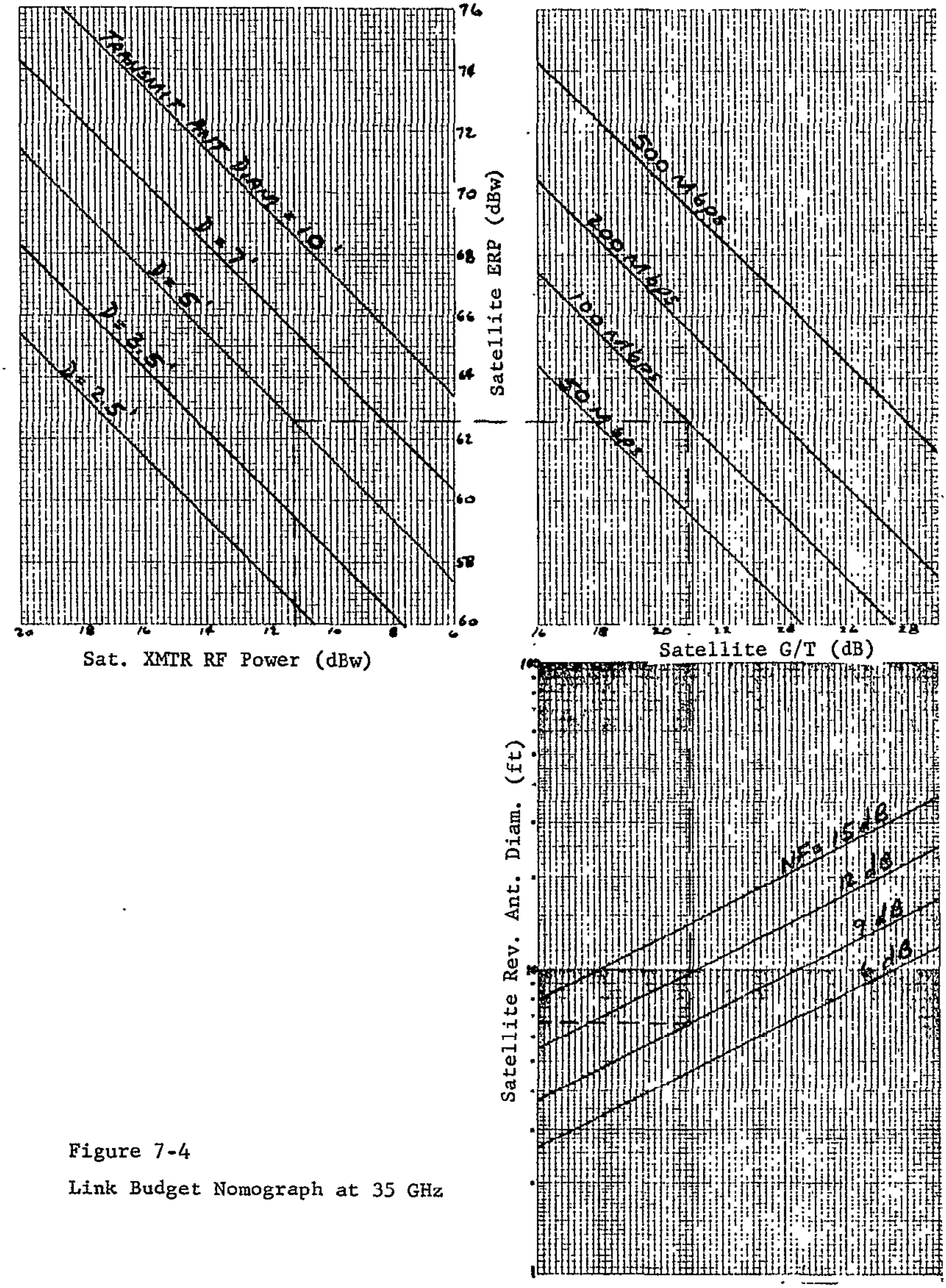


Figure 7-4

Link Budget Nomograph at 35 GHz

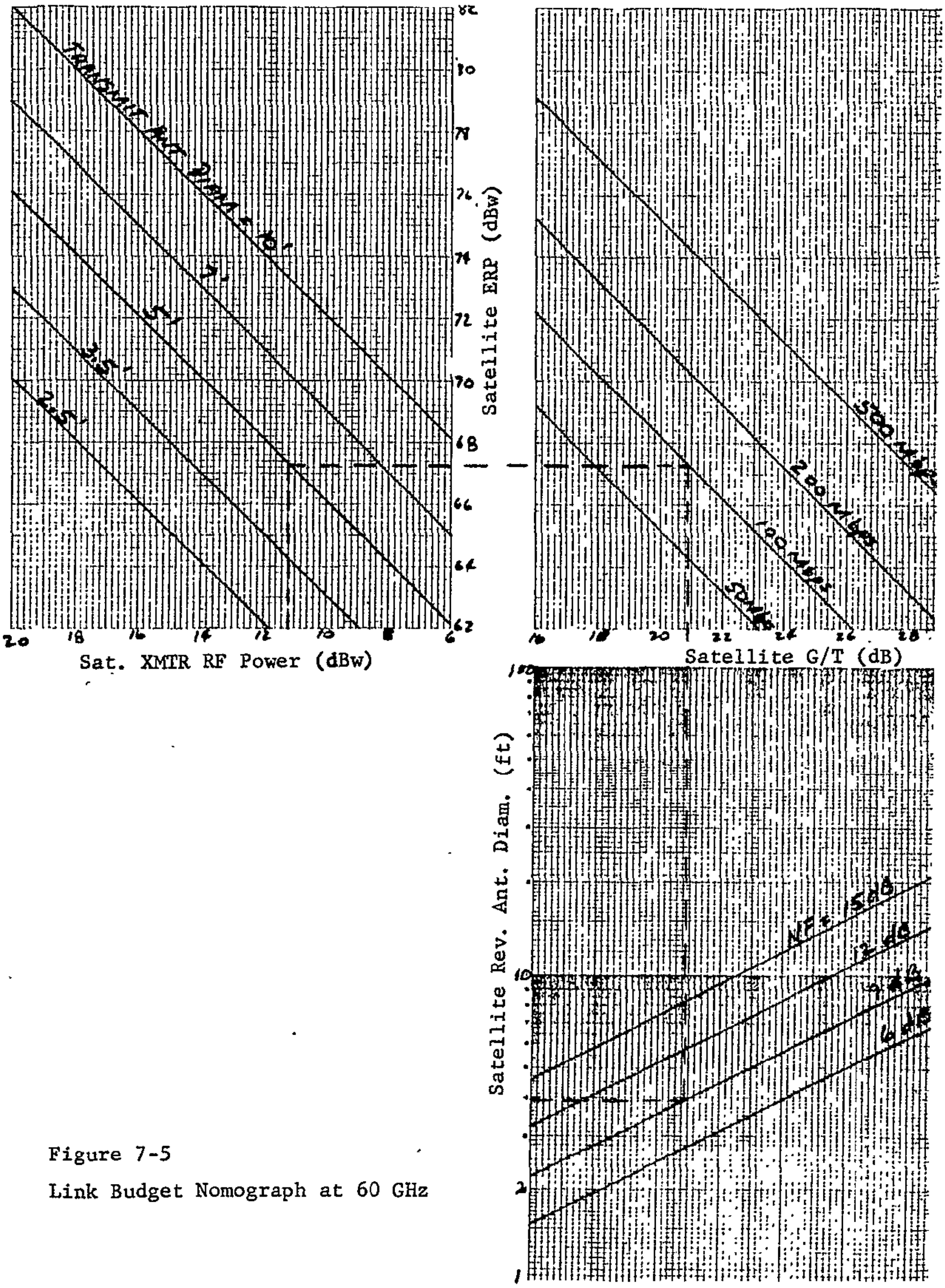


Figure 7-5

Link Budget Nomograph at 60 GHz

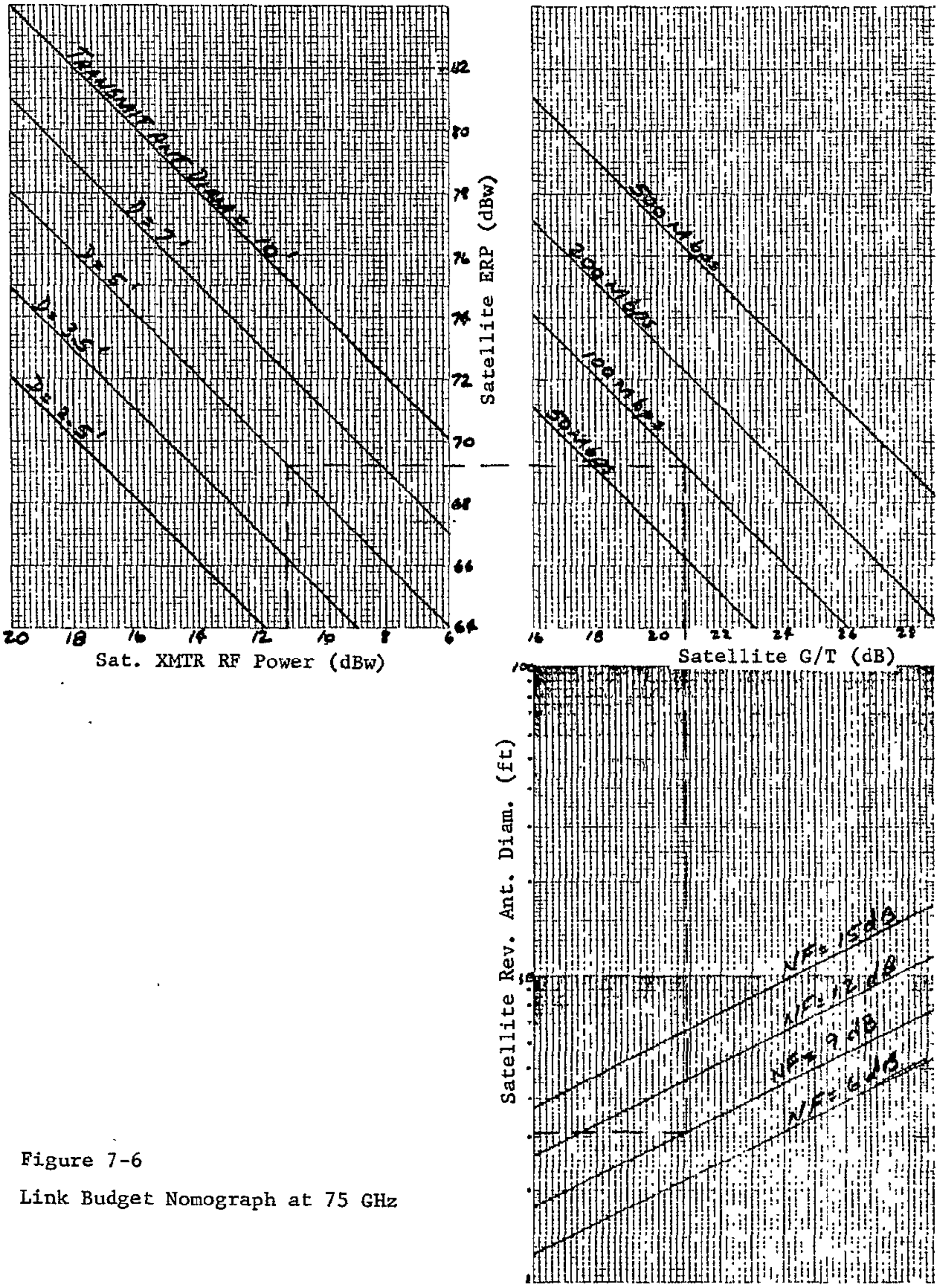


Figure 7-6

Link Budget Nomograph at 75 GHz

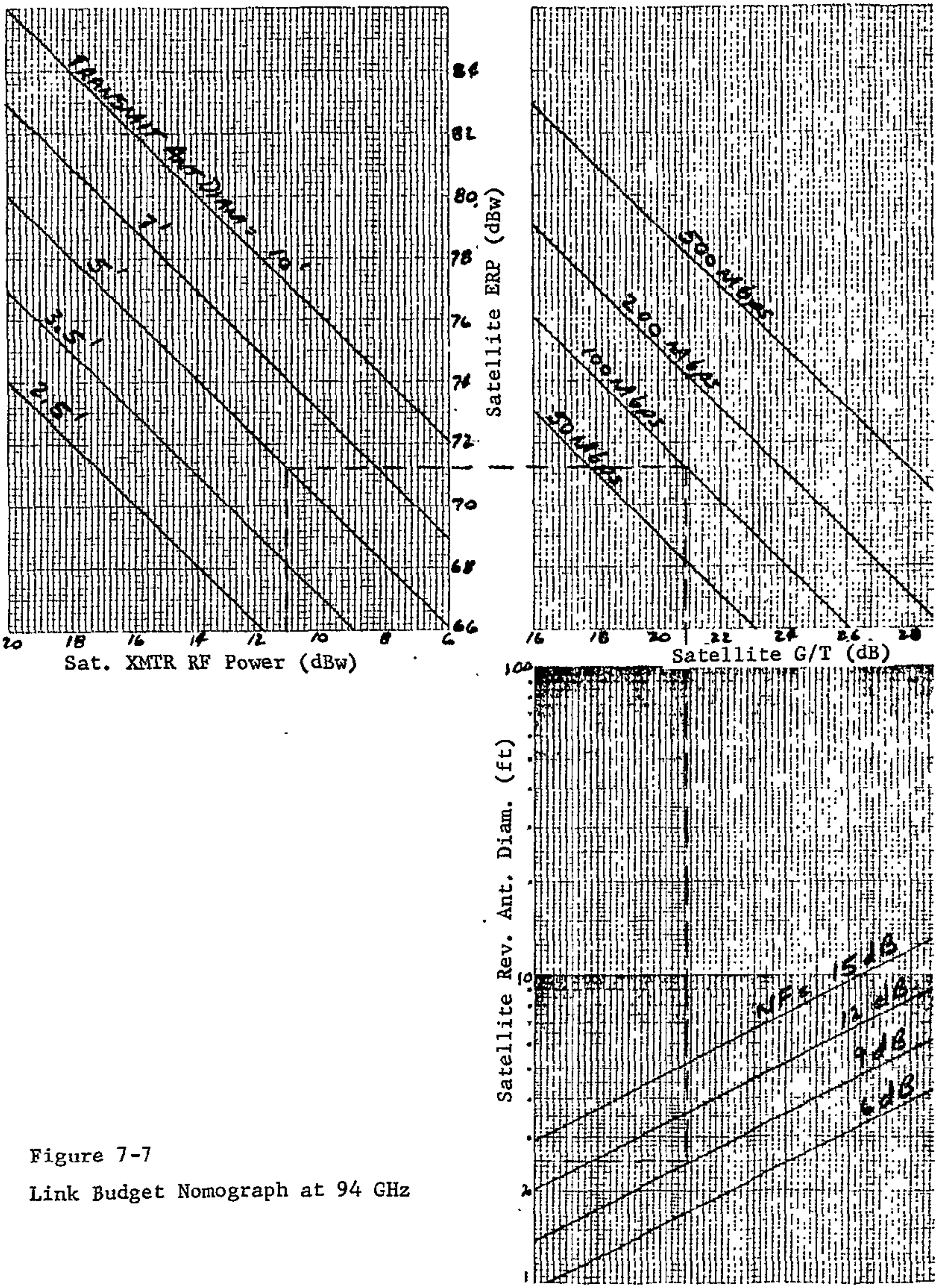


Figure 7-7

Link Budget Nomograph at 94 GHz

Figures 7-8 and 7-9 identify the transmitter power potential in this frequency range for both solid state and electron tube devices, showing that for transmitter powers up to 1-watt, the solid state device has the greatest potential, while for power outputs (and high gain) from 20 watts into the kilowatts, electron tube devices such as traveling wave tubes have the best capability. At a time point, Circa January 1970, a solid state transmitter is capable of developing in the vicinity of 100-150 mw, while electron tube technology can provide power levels at the 20 to 100 watt saturated power levels.

7.3.1 Millimeter Wave Transmitters For High Data Rates

Electron tube amplifiers such as TWT have the features of high gain (>35 dB) and flat gain and minimum phase distortion suitable for high data rate PSK transmission and can serve as pre-amplifiers and modulator drivers operating at the 1-10 milliwatt level. Such tubes are scaled from microwave frequency electron tubes which have an enviable record of long-life high-efficiency operation in space.

Solid state devices can perform as low power transmitters or as TWT drivers and can have many configurations using semiconductor diodes of the Gunn, LSA, or avalanche varieties, i.e.,

- Diode oscillator plus diode phase shifter
- Diode oscillator driving a diode modulator and a diode amplifier
- Varactor diode parametric up-converter from a lower frequency
- Harmonically related lower frequency diode oscillator driving a data modulator followed by a parametric multiplier to the transmit frequency

Each of these techniques must provide gain, modulator and power generation such that channel requirements relative to gain flatness, AM-to-PM conversion, and phase and group-delay distortion are met at the data rates used.

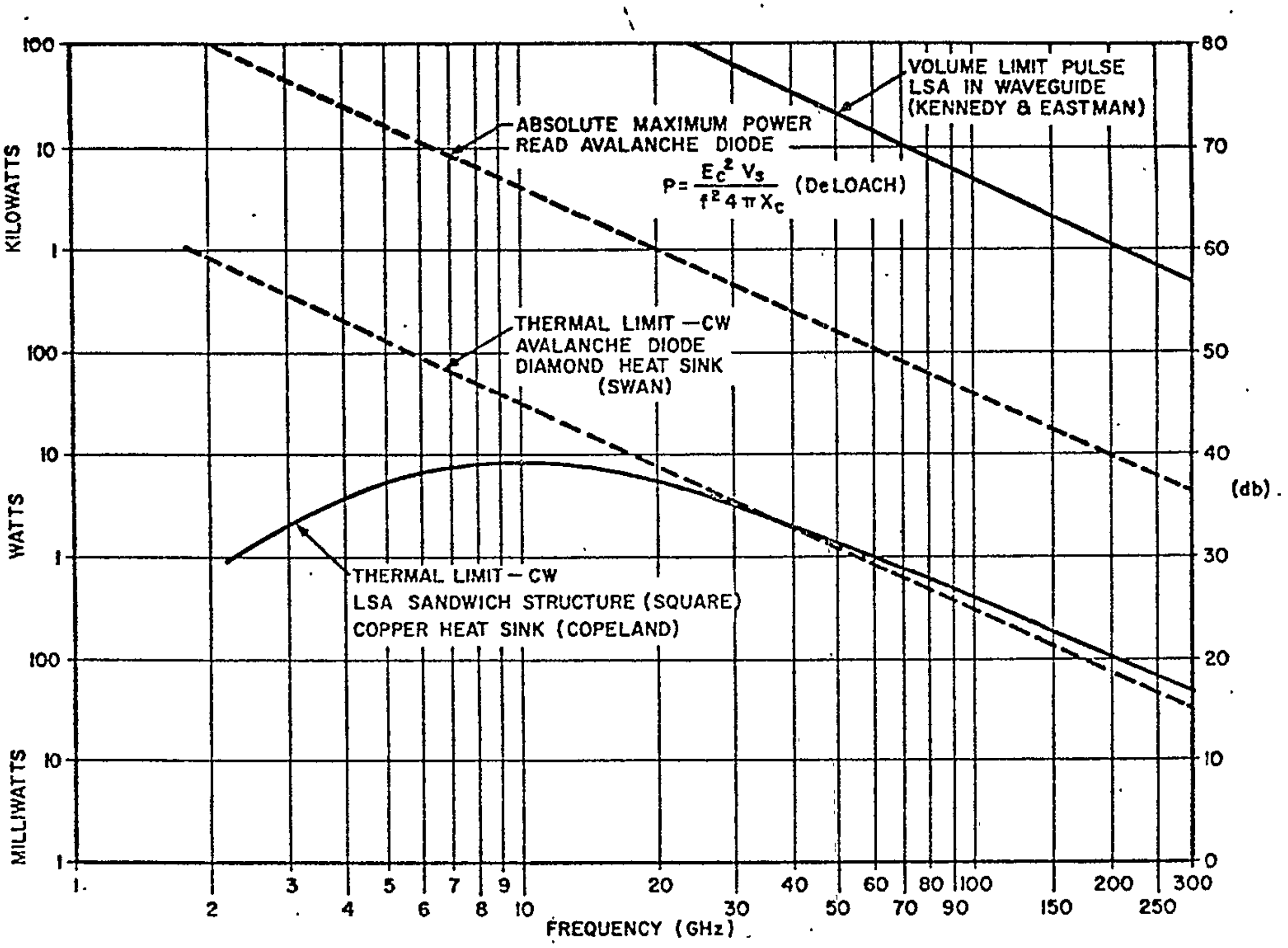


Figure 7-8 Maximum Power vs Frequency for Semiconductor Diodes

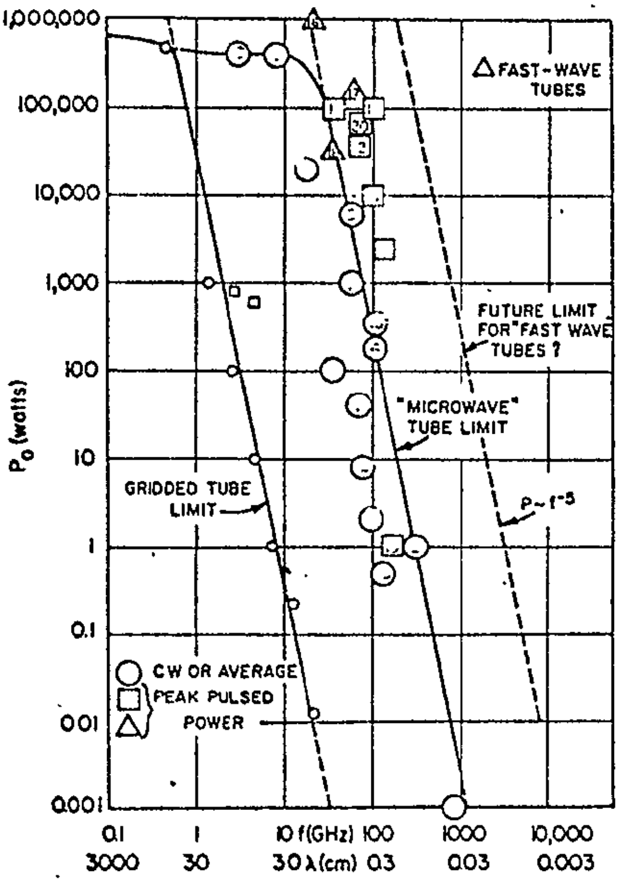


Figure 7-9 Frontier of Microwave and Millimeter-wave Power Generation - After J. Osepchuck

7.3.1.1 Group-Delay Distortion in Millimeter Wave Amplifiers

A unique problem which exists in millimeter wave amplifiers and transmission systems is one of determining group delay distortion in devices in which the time of passage is extremely short.

At microwave frequencies, linear and parabolic group delay distortion is measured in nanoseconds. In a 60 GHz amplifier, however, the total time of passage may be less than 40 wavelength and total group delay may be less than 800 picoseconds, with the group delay distortion then being measured in picosecond increments. Techniques of measurement must be developed for these particular frequency ranges to serve these short transit times.

In the present 4/6 GHz COMSAT and ATS receive/transmit bands and the 7/8 GHz DCS receive/transmit bands, group delay and phase linearity specifications for the receive or transmit chains are already tight over 500 MHz bands, requiring the following.

o Group Delay (COMSAT):	Linear: Any 40 MHz band, better than 0.1 nsec/MHz Parabolic: Better than 0.05 nsec/MHz Ripple: Better than 0.6 nsec peak-to-peak
o DSCS Phase Linearity:	± 0.1 radian over any 40 MHz band ± 0.25 radian over 125 MHz band

Similar and more stringent specifications must be specified for millimeter wave tube systems designed to handle bit rates above 50 Mbps.

Group delay distortion in the TWT amplifier, or any link component, is a critical link parameter upon which the efficiency of the transmission depends. It is implicitly discussed in terms of intersymbol interference due to the filter-characteristics of the component in 5.7.3.1 of Section 5, and will be specifically addressed in the Final Report as a key limiting characteristic in the transmission of image type data.

7.3.1.2 AM-PM Conversion Effects

In this part, the AM-PM conversion loss mechanism of the traveling wave tube amplifier is detailed. Data of many microwave power TWT were analyzed with respect to degrees of phase advance with respect to mid-pulse saturated phase as a function of input drive level. A piece-wise linear (2 lines) saturating TWTA power-out versus power-in characteristic was assumed with peak drive level 2 dB above that necessary to provide saturation.

The worst-case total system loss was determined for a) TWTA input drive level, b) TWTA output power, and c) phase error as a function of time. These variables are plotted in Figures 7-10, 7-11 and 7-12 respectively, as a function of normalized time (t/t_{r1}) where t_{r1} is the rise-time prior to the TWTA.

The total worst-case theoretical AM-PM loss can be determined by evaluating the degradation in the output voltage of an ideal matched-filter at the sampling instant. This sampled voltage is given by

$$V_{xo} = \frac{1}{T}\int_0^T \left\{V_x(t)\cos\theta(t) + V_y(t)\sin\theta(t)\right\} dt$$

where $V_x(t)$ and $V_y(t)$ are independent bandwidth-distorted (+1, -1) pulse waveforms and $\theta(t)$ is the phase error function. If there were no AM-PM distortion, $\theta(t) = 0$. An equation analogous to the above equation holds for the quadrature channel.

It suffices to consider only the in-phase channel,

$$V_{xo} = \left[1 - \frac{t_r}{T}\right] + \int_0^{t_r/2} \left\{V_x(t)\cos\theta(t) + V_y(t)\sin\theta(t)\right\} dt + \int_{T-(t_r/2)}^{T} \left\{V_x(t)\cos\theta(t) + V_y(t)\sin\theta(t)\right\} dt \quad (2\text{-}7)$$

and the degradation in dB is given by

$$D_T = -20\log_{10}(V_{xo})$$

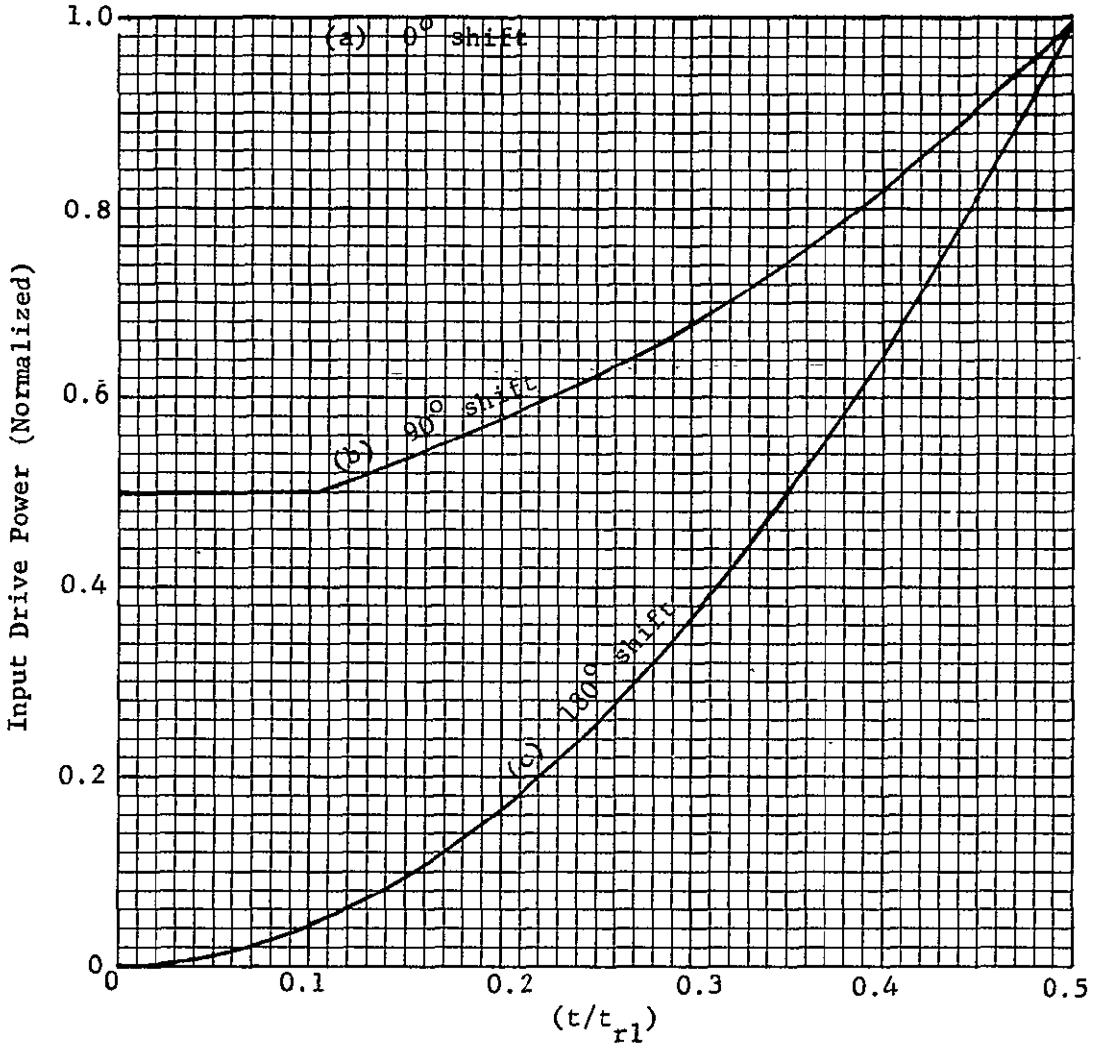


Fig. 7-10 TWTA drive power as a function of normalized time for: a) a 0° phase shift, b) a 90° phase shift, and c) an 180° phase shift.

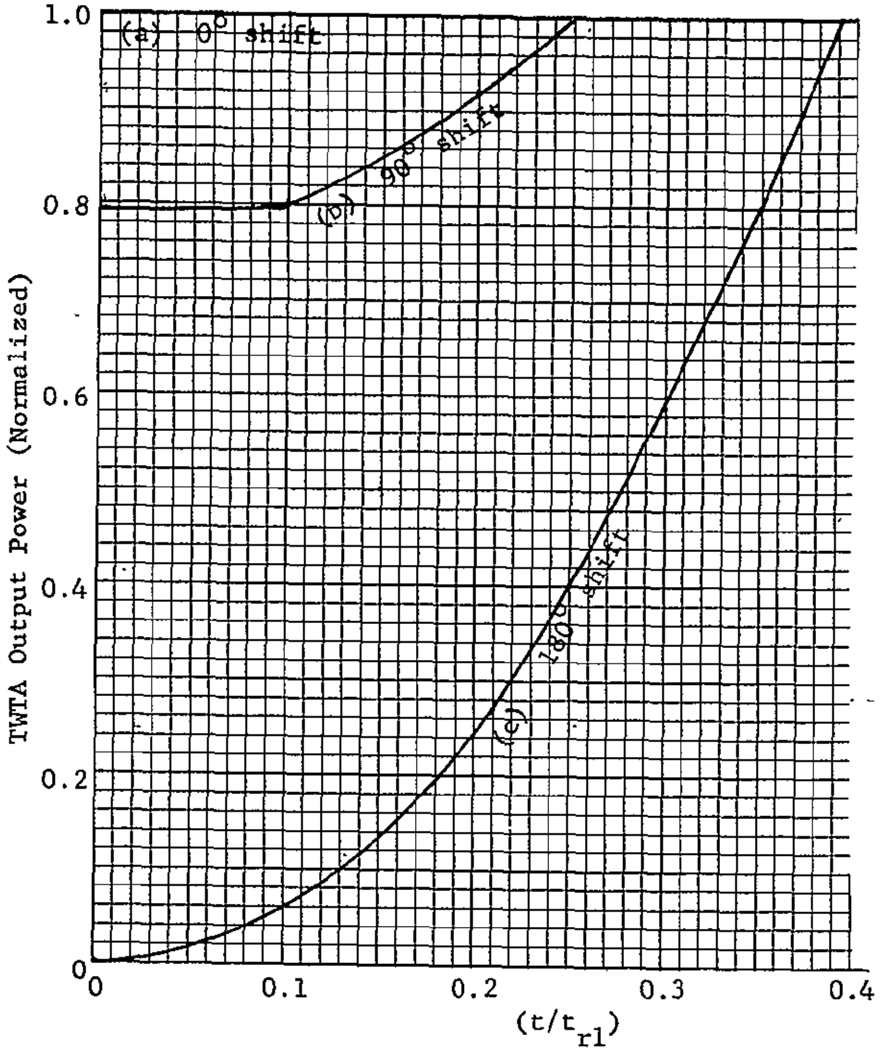


Fig. 7-11 TWTA output power as a function of normalized time for: a) a 0° phase shift, b) a 90° phase shift, and c) an 180° phase shift.

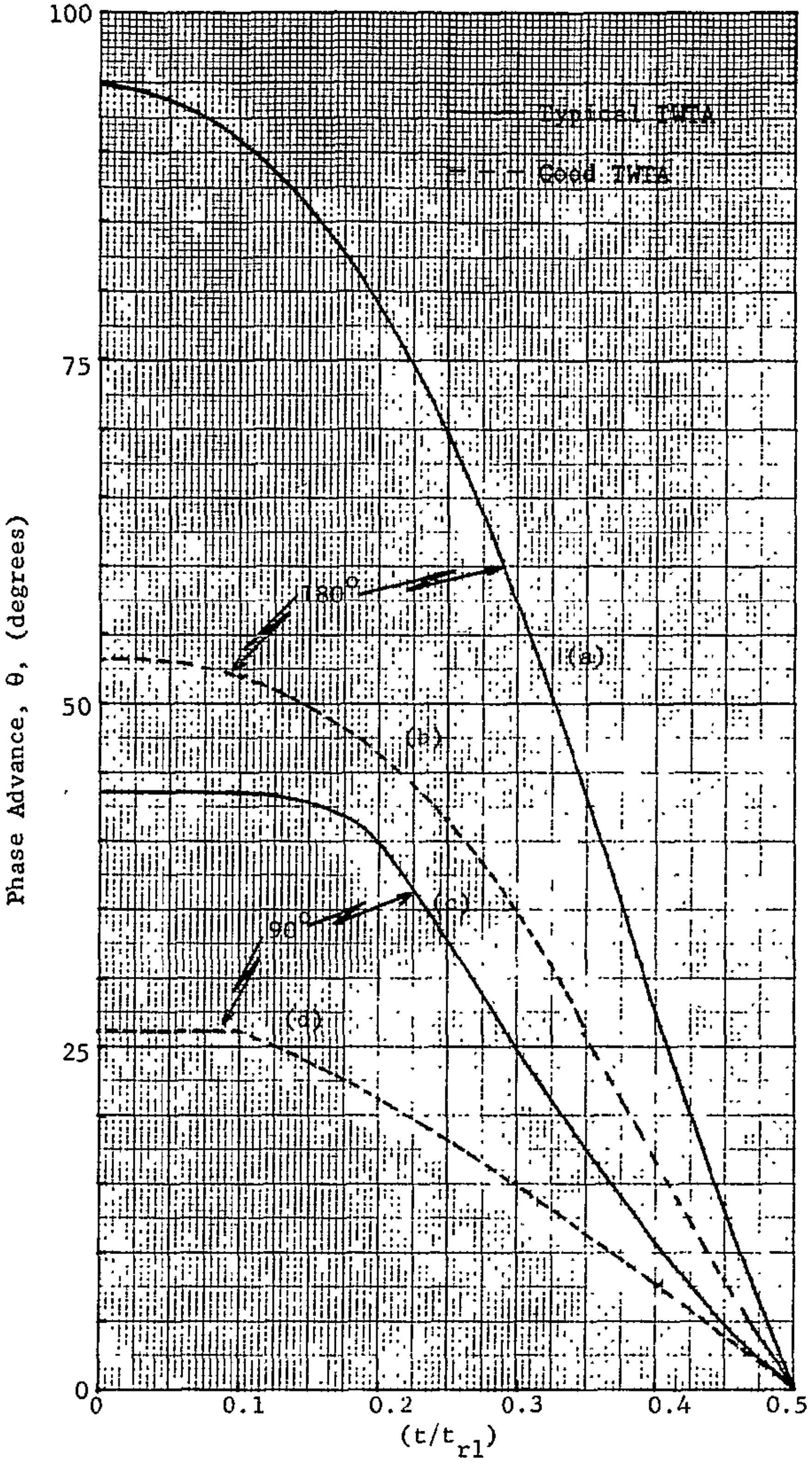


Fig. 7-12 Phase error advance as a function of normalized time for: a) 180° shift typical TWTA, b) 180° shift good TWTA, c) 90° shift typical TWTA, and d) 90° shift good TWTA.

Inspection of the equation for V_{xo} indicates that the worst-case degradation occurs when the integrals in this equation are minimized. This occurs when $V_x(t)$ and $V_y(t)$ have opposite signs permitting destructive cross-talk between the two channels.

There are 64 transition patterns for the in-phase (or x) channel when leading and trailing edge and initial phase possibilities are included. Table 7-2 which was generated from phasor diagrams, presents the sign of cross-talk from the quadrature channel for leading and lagging pulse edges, respectively. For example, an entry of 00 corresponds to no degradation while an entry of -+ corresponds to destructive cross-talk on the leading edge and constructive cross-talk on the trailing edge.

The potential worst-case consists of a 90° shift (x change) on both the leading and lagging edges. However, Table 7-2 shows that it is not possible to get destructive cross-talk on both leading and lagging edges in this case. Consequently, the worst-case is given by a 90° phase shift (x change) on the leading edge and a 180° phase shift on the lagging edge (or vice versa). For this worst case, it is seen that the net contribution of the integrals is essentially zero and the worst case degradation is given by

$$D_{wc} = -20 \log_{10} \left[1 - \frac{t_r}{T} \right]$$

A modulator rise time of less than 10 nanoseconds yields a worst-case degradation under 1 dB. A rise time of 5 nanoseconds yields a 0.45 dB worst-case loss. The rise time of the modulator is defined for bi-polar pulses and not from the zero crossing to full magnitude.

The average theoretical degradation due to modulator rise time could be evaluated by averaging the probabilities of error for each of the cases entered in Table 7-2. The probability of error for each case (note that several entries appear more than once) can be evaluated

Table 7-2 Cross-Talk Sign

Transition		Initial Phase Position			
Leading	Lagging	0	1	2	3
0	0	00	00	00	00
0	90y	0-	0+	0-	0+
0	90x	0-	0+	0-	0+
0	180	0-	0+	0-	0+
180	0	-0	+0	-0	+0
180	90y	--	++	--	++
180	90x	--	++	--	++
180	180	--	++	--	++
90y	0	-0	+0	-0	+0
90y	90y	-+	+-	-+	+-
90y	90x	-+	+-	-+	+-
90y	180	-+	+-	-+	+-
90x	0	-0	+0	-0	+0
90x	90y	-+	+-	-+	+-
90x	90x	-+	+-	-+	+-
90x	180	-+	+-	-+	+-

Entries are signs of leading and lagging edge cross-talk, respectively.

from the degradations determined from the table. The worst-case degradation only occurs twice out of 64 chances and there are only six cases of destructive cross-talk interference on both leading and lagging edges. Thus, the cases giving the worst losses are reduced in effectiveness by a factor of 6/64. Consequently, the average degradation is more strongly influenced by the more average cases such as no transition on leading edge and a 90° (in-phase channel change) on the lagging edge (or vice versa).

Basing the estimate of the average bit error rate loss on this typical case, the average degradation is found to be

$$D_{ave} = -20 \log_{10} \left[1 - \frac{t_r}{2T} \right]$$

7.4 RF RECEIVER

The RF receiver, with the antenna and feed, determines the final sensitivity or G/T of the receiving system. The receiver noise temperature, in combination with the antenna noise temperature and the feed loss temperature determine the system noise temperature T upon which the G/T is based.

7.4.1 Low Noise Millimeter Wave Amplifiers

Low noise amplifier technology has been highly developed at microwave frequencies due to COMSAT and Unified S-band requirements. However, comparable developments have not occurred at millimeter waves, and only a 70 GHz maser due to Hughes and Kremenack[1] and a tunnel diode due to Burrus and Trambarulo[2] at BTL have provided noise figures to compare with the millimeter wave mixer which has been widely used for more than a decade in radiometer applications.

Table 7-3 lists the principal candidate low noise amplifiers for millimeter wave amplifications. As shown, the maser requires both cryogenic cooling and a high magnetic field and is not a suitable candidate for space applications.

Standard parametric amplifier techniques and design, based on lower frequency paramps, require varactor diodes with cut-off frequencies well above 1000 GHz, and paramp power at frequencies in excess of 150 GHz. Such varactors are beyond the state of the art, and long-life reliable pump power generation above 150 GHz in the space environment is unrealistic at this time. A possible candidate is the lower-frequency pumped (40-90 GHz) paramp of which there are several species and which can produce noise figures in the 5-10 dB range.

---

1. W. Hughes and C. Kremenack, Proc IEEE, pg. 856, May 1963.

2. C. Burrus and R. Trambarulo, Proc IRE, pg. 1075, June 1961.

TABLE 7-3

MILLIMETER WAVE LOW NOISE AMPLIFIERS

	Maser	Ultra Low Noise Paramp	Low Noise Paramp (Conventional)	Medium Low Noise Paramp Idler Freq Above Signal Freq	Medium Low Noise Paramp Idler Freq Below Signal Freq	Mixer
Semi-conductor Require-ments	iron doped rutile in 4^oK environ.	$f_c \approx 2400$ GHz $f_p \approx 360$ GHz	$f_c \approx 1400$ GHz $f_p \approx 180$ GHz	$f_c \approx$ 300-600 GHz $f_p \approx$ 130-160 GHz	$f_c \approx$ 300-600 GHz $f_p \approx$ 75-100 GHz	$f_c > 500$ GHz L.O.$\approx$50-55 GHz
Special Reqmts	magnetic field of 5000 gauss 120 GHz f_p	extremely high pump frequency beyond state of art	varactor and pump extreme state of art	waveguide mounted varactor chip	waveguide mounted varactor chip	low noise IF
Typical Noise Figures at 300^oK	--	1.5 dB	3 dB	4-6 dB	7-9 dB	10-13 dB
Key Aspects	noise figure capability 1/2 dB requires spaceborne refrigeration	requires major technological advance	requires very high pump frequency major reliability barrier	requires high pump frequency is in research stage	highly feasible needs moderate pump development simple circuit	low L.O. power wide-band highest reliability

f_c = varactor cut-off frequency measured at minus 0.5 volt bias

f_p = pump frequency

The mixer, with a low noise post amplifier, is the key candidate for 10 dB noise figure operation and has the advantage of low local oscillator power, simplicity, and high reliability.

7.4.2 Millimeter Wave Mixer Noise Figures

Currently the status of millimeter wave mixers is such that mixer noise figures as low as approximately 10 dB are achievable. Conversion losses run from 8 dB to 10 dB and excess noise ratio is approximately 1.4 for silicon Schottky-barrier balanced mixers. The actual achievable overall receiver noise figure will be dependent upon the IF amplifier noise figure. Solid state S-band IF amplifiers have noise figures in the range of 4.5 dB to 6 dB depending upon bandwidth. Thus, with a 4.5 dB IF, and 8 dB mixer conversion loss and an excess noise ratio of 1.4, the overall receiver noise figure is 13 dB.

Recent research at MIT by Dr. D. H. Steinbrecher has led to a modified design theory for balanced mixer diodes which is especially suited to the millimeter wave region. Steinbrecher maintains that the optimum design criteria for a balanced mixer is to operate the diodes as switches, e.g., design for a minimum switch time and to reactively terminate the image frequency and higher harmonics. The effects of the approach are to reduce the excess noise ratio to unity since the diodes are driven relatively hard to theoretically reduce the conversion loss to zero dB. A V-band mixer has been constructed utilizing this technique and has produced a 6 dB noise figure, on a highly experimental basis.

7.4.2.1 Mixer IF Amplifiers

The IF frequency for a data relay (or ground receiver) can be located in a wide choice of frequency bands, in the 1.5 to 1.5 GHz band. At frequencies from 1.5 to 4 GHz low noise figures are achieved with reliable transistor amplifiers. Above 4.0 GHz tunnel diode amplifiers (TDA's) provide lower noise figures than do transistors.

- Uncooled IF parametric post amplifiers can now be considered for wide band amplification, providing lower noise figures than the transistor amplifier or bandwidths up to 500 MHz at 26 Hz and 26 Hz at 20 GHz are now possible, and new integrated circuit techniques and the use of Gunn or avalanche diode oscillators make such amplifiers attractive post amplifiers.

SECTION 8

DATA RECORDING TECHNOLOGY

8.1 WIDEBAND RECORDING

The status of wideband recorder technology presented in Tables 8-1, 8-2, and 8-3 is of interest to sensor satellite imagery transmission from the dual aspects of ground station data recording and on-board satellite recording.

Satellite recording may be required either for back-up or store-and-forward in the event that the sensor satellite is not in view of a relay satellite or ground station at the time of data acquisition. The application of wideband recorder technology for satellite borne recorders is dictated by the mission requirements, specifically the orbit network in which the satellite recorder is to function.

A typical satellite store-and-forward system requirement would most probably be: "The store-and-forward system shall record satellite sensor data while the satellite is over remote areas and shall play back the data on command. The system shall record data at rates compatible with the satellite sensor systems with minimum degradation of signal quality caused by the record/playback cycle. The system shall operate on a duty cycle of 5 minutes read-in per revolution and shall be capable of storing up to 6 revolutions worth of data. The system shall be designed for an on-orbit operational lifetime of 3 years."

8.2 MISSION CONSIDERATIONS IN DATA RECORDING

The data in the tables are presented to illustrate how satellite operations tend to constrain or dictate store-and-forward system configurations. These mission dictated constraints, when evaluated with the current state-of-the-art recording technology, indicate the areas where further development is necessary.

The mission network described in this section is a basic model that utilizes one sensor satellite at a 400 mile altitude and a relay satellite or ground readout terminal for receiving the satellite data. The mission example described below considers the store-and-forward of sensor data from the sensor satellite to a ground terminal. Another possible mission configuration is store-and-forward from the sensor satellite to the relay satellite. Since the readout to a ground station generally involves a shorter readout interval, that case is considered here to establish typical mission parameters.

The readout rate of a store-and-forward system should be high enough to play back all the stored data during a pass in which the relay satellite or ground station is visible. For the assumed orbit altitude of 400 miles and one readout station, an average readout time of 22 minutes per day is available, but the time is divided into nominally 2 contact periods of 11 minutes each. The duration of the station contact and the recorder playback rate then establishes the amount of data that can be readout. High performance recorders normally record and play back at the same speed which establishes the amount of data that can be accumulated on the store-and-forward recorder and completely readout during any given station contact period.

8.2.1 System Lifetime

Current satellite technology for this class of satellite gives expected lifetimes on the order of 2 to 3 years. This lifetime is also needed to provide an economic data gathering satellite considering the satellite, booster and launch base costs.

For scaling and technique selection purposes, it is assumed that a nominal duty cycle for a store-and-forward system would record for 5 minutes every revolution, and readout all recorded data during every available station contact. A linear recording speed of 15 ips was selected as typical.

TABLE 8-1

WIDEBAND RECORDING
CURRENT MAGNETIC TAPE TECHNOLOGY

Longitudinal Recording	
Analog Response	1.4 MHz at 120 ips per track
Digital Rate	2.0 Mbps at 120 ips
Number of Tracks	7 to 9 maximum on 1/2 inch wide tape
Time Base Error	± 0.5 microseconds at 120 ips
Dynamic Skew	± 0.5 microseconds at 120 ips
Tape Capacity	4800' of 1 mil tape on 10 inch reel
Record/Playback Time	8 minutes at 120 ips
Bit Error Rate	1 x 10^{-5} on specially tested tape
SNR	Nominally 25 dB RMS/RMS
Head Life	1000 hour nominal

TABLE 8-2

CURRENT MAGNETIC TAPE TECHNOLOGY

Rotary Head Recording	
Analog Response	6 to 8 MHz at 1600 in/sec head to tape speed
Digital Rate	5 to 10 Mbps NRZ data
Number of Channels	1 or 2
Tape Width	2 inch wide
Tape Speed-Longitudinal	15 ips (1 channel recording)
Tape Capacity	4500' of 1 mil tape on 10 inch reel
Record/Playback Time	1 hour (1 channel)
Bit Error Rate	1 x 10^{-5} on specially tested tape
SNR	Nominally 25 dB RMS/RMS
Time Base Error	± 15 nanoseconds peak-to-peak
Head Life	200 hour nominal

TABLE 8-3

CURRENT FILM RECORDING TECHNOLOGY

Laser Electron Beam CRT Line Scan	These systems are characterized by scanning photographic film on a line by line basis with the beam intensity varied to expose the film. The film must be processed before the data can be read out. Current development status is in the recording of photographic imagery using wide video bandwidths. The use of these techniques for digital recording presents no problems.
Response	25 to 100 MHz video bandwidth up to 100/200 Mbps digital
Film	16mm to 5 inch wide
Access Time	Within 30 minutes after recording
Storage	Permanent data - high packing density
Development Status (Ground Only)	Analog Systems - in operation Digital Systems - require modulator development

TR-DA2180

STUDY OF EFFICIENT TRANSMISSION AND RECEPTION OF IMAGE-TYPE DATA USING MILLIMETER WAVES

FINAL REPORT

VOLUME 2

SYSTEM TECHNOLOGY AND ANALYSIS

PHILCO Ford

SPACE & RE-ENTRY SYSTEMS DIVISION
Philco-Ford Corporation

TABLE OF CONTENTS

VOLUME 2

Section No.		Page
1	Nanosecond Switching Technology	1-1
1.1	The Role of Nanosecond Switching Technology in Digital Communications	1-1
1.2	Pulse Rise Times	1-2
1.3	Transistors	1-4
1.4	Tunnel Diodes	1-8
1.5	Snap Back Diodes	1-8
1.6	Phase Locked Oscillators	1-11
1.7	Balanced Modulators	1-11
1.8	Gunn Diode Switching Circuits	1-13
1.9	Microwave-Rate Storage and Coding Circuits	1-13
2	Digital Modulation Techniques	2-1
2.1	Digital Modulation Techniques	2-1
2.2	Comparison of Channel Efficiency of Different Modulation Methods on the Basis of BER Versus E_b/N_o	2-1
2.2.1	FSK and MFSK (Non-Coherent)	2-1
2.2.1.1	PCM/FM (Discriminator Detection)	2-3
2.2.1.2	PSK (Coherent)	2-3
2.2.1.3	PSK (Differentially Coherent)	2-4
2.2.1.4	MPSK (Coherent)	2-4
2.2.1.5	M-ary Coded PSK (Orthogonal Coding)	2-5
2.2.1.6	M-ary Coded PSK (Bi-Orthogonal Coding)	2-5
2.2.2	Comparison Curves and Discussion	2-5
2.2.3	Comparison of Digital Modulation Techniques on the Basis of Required Bandwidth	2-7
2.2.3.1	FSK and MFSK (Non-Coherent)	2-7
2.2.3.2	PCM/FM	2-9
2.2.3.3	PSK (Coherent)	2-9
2.2.3.4	PSK (Differentially Coherent)	2-9

TABLE OF CONTENTS (continued)

VOLUME 2

Section No. Page

2.2.3.5 MPSK (Coherent) 2-9

2.2.3.6 M-ary Coded PSK 2-10

2.2.4 Bit Error Rate Degradation of BPSK and QPSK Due to Filtering 2-10

2.3 Comparison of Analog and Digital Modulation Techniques 2-11

3 Sensors and Sensor MTF 3-1

3.1 Sensors and Sensor Resolution 3-1

3.2.1 Vidicons 3-1

3.2.2 The Image Orthicon 3-3

3.2.2.1 Intensifier Image Orthicon 3-7

3.2.2.2 Isocon 3-8

3.2.2.3 The Image Dissector 3-8

3.2.3 Solid State Sensor Imagery 3-9

3.2.3.1 Classes of Radiation Detectors 3-10

3.2.3.2 Photoconductors 3-10

3.2.3.3 The Photo Transistor 3-11

3.2.3.4 Photo-Emissive Cells 3-13

3.2.3.5 Mosaics 3-13

3.3 Modulation Transfer Function 3-14

3.3.1 Modulation Transfer Function and Limiting Resolution 3-14

3.3.2 Resolution Notation 3-17

3.3.3 Convolution Analysis Basis of MTF 3-17

3.3.4 The Concept of Spatial Frequency 3-20

3.3.5 Use of the Modulation Transfer Function for a System 3-22

3.3.5.1 MTF and Scene Statistics 3-23

3.3.5.2 Lens MTF 3-26

TABLE OF CONTENTS (continued)

VOLUME 2

Section No.		Page
3.3.5.3	Fiber Optics MTF	3-26
3.3.5.4	Vidicon MTF	3-27
3.3.5.5	Step Response of a Scanning Sensor to a Knife Edge Target	3-29
3.3.5.6	Contribution of Image Motion to MTF	3-35
3.3.5.7	Cathode Ray Tube MTF	3-37
3.3.5.8	Image Dissector	3-38
3.3.5.9	MTF of Side Looking Radars	3-42
4	High Speed A/D Converter Technology	4-1
4.1	High Speed A/D Converters	4-1
4.2	Types of High Speed A/D Converters	4-1
4.2.1	Folding Converter A/D Conversion	4-1
4.2.1.1	Fisher High Speed Folding Encoder	4-2
4.2.2	Phase Domain Encoders	4-4
4.2.3	Fast Shift Register Encoders	4-10
4.2.4	Cathode-Ray-Tube A/D Converters	4-14
4.3	Redundancy-Removal Encoders	4-18
4.3.1	Introduction to Delta Modulation	4-20
4.3.2	One-Bit Delta Modulation	4-21
4.4	The Accuracy of Delta Modulation	4-25
4.4.1	N-Bit Delta Modulator	4-25
4.4.1.1	The 2-Bit Delta Modulator	4-27
4.4.1.2	The 4-Bit Delta Modulator	4-30
4.5	Quantizing Noise in A/D Conversion	4-37
4.5.1	Quantizer Representation	4-38
4.5.2	Linearization of the Quantizer and the Recovery of the Unquantized Signal Statistics	4-44
4.5.3	Linearization of the Quantizer and the Reconstruction of the Unquantized Information Signal - A Heuristic Discussion	4-44

TABLE OF CONTENTS (continued)

VOLUME 2

Section No.		Page
5	Theory of Quadriphase Modulation	5-1
5.1	Phase Modulation Techniques for High Speed Digital Data	5-1
5.2	Basic Quadriphase Communication System Theory	5-2
5.2.1	Advantages of Quadriphase Signalling	5-2
5.2.2	Problems with Quadriphase	5-4
5.2.3	Two Fundamental Approaches to Multi-Phase Detection	5-4
5.3	Quadriphase Modulators - Functional Description	5-5
5.3.1	Functional Description of Quadriphase Modulator	5-5
5.3.2	Sensitivity Analysis	5-9
5.3.2.1	(I,Q) Channel Amplitude Imbalance	5-9
5.3.2.1.1	Unbalanced Power Loss	5-9
5.3.2.1.2	(1,0) Pulse Amplitude Imbalance (Each Channel)	5-10
5.3.3	Pulse Width Asymmetry	5-14
5.4	Coherent BPSK and QPSK Demodulation and Detection	5-18
5.4.1	Functional Description of BPSK and QPSK Receivers	5-18
5.4.1.1	Quadriphase Demodulator	5-18
5.4.1.2	Matched Filter Detector and Differential Decoder	5-20
6	Search and Tracking Calculations	6-1
6.1	Introduction	6-1
6.2	Search Calculations	6-1
6.3	Phase-Lock Loop Bandwidth	6-6
6.4	Time for Frequency Search	6-6
6.5	Frequency Search Rate	6-7
6.6	Tracking Errors	6-7
6.6.1	Loop Noise Bandwidth	6-7
6.6.2	IF and Loop Bandwidth vs Doppler Shift	6-8
6.6.3	Tracking Accuracy	6-8

TABLE OF CONTENTS (continued)

VOLUME 2

Section No.		Page
6.6.4	Tracking Error Due to Thermal Noise	6-12
6.6.5	Tracking Error Due to Pre-Comparator and Post-Comparator Differential Phase Shift	6-12
6.6.6	Tracking and Pointing Error Due to Amplitude Unbalance	6-18
6.6.7	Tracking and Pointing Error Due to Unequal Individual Beamwidth	6-21
6.6.8	Tracking Error Due to Antenna Temperature	6-23
6.6.9	Servo Lag Errors	6-26
6.6.10	Servo Bias Error	6-26
7	Millimeter Wave Antennas	7-1
7.1	The High-Gain Millimeter Wave Antenna	7-1
7.1.1	Antenna Sizes	7-1
7.2	High-Efficiency Reflector Design	7-3
7.3	Gain Versus Antenna Surface Distortion	7-6
7.4	Baseline Material Design	7-8
7.4.1	Analytical Approach	7-11
7.4.2	Thermal Model	7-12
7.4.3	Antenna Pattern Analysis	7-13
7.4.4	Worst Case Thermal Environment Definition	7-15
7.5	Tracking Feeds in High Gain Millimeter Wave Antennas	7-21

LIST OF ILLUSTRATIONS

VOLUME 2

Figure No.		Page
1-1	Switching Times and Voltage Levels of Modern Switching Devices	1-3
1-2	Comparison of Different Types of Transistor Logic Circuits	1-5
1-3	Simplified Transistor Switching Circuit and Switching Characteristics	1-6
1-4	Tunnel Diode Balanced Pair (GOTO Pair)	1-9
1-5 (a)	Simplified Snap-Diode Pulse Generator	1-10
1-5 (b)	Typical Waveforms of Pulse Generator	1-10
1-6	Parametric Phase Locked Oscillator Circuits (a) One Possible lumped Parameter Variable Capacitance PLO Circuits, (b) One Possible Microwave PLO (in microstrip), (c) Symbolic Representation of PLO - after L. Onyshkevych, et.al.	1-12
1-7	JK Flip-Flop Shift Register Showing Clock Distribution System Simulated by Delay Lines	1-14
1-8	Circuit Diagram of a One-gigahertz Shift Register Stage Due to B. Sear. Tunnel Diodes are Used.	1-16
2-1	Binary Error Probabilities for MFSK	2-6
2-2	A Comparison of M-ary and Binary Signaling Techniques	2-8
2-3	Bandwidth Limiting Degradation of Biphase PSK Signals	2-12
2-4	Effect of Transmission Channel Filter on Biphase PSK Bit Error Rate	2-13
2-5	Effect of Transmission Channel Filter on Quadriphase PSK (QPSK)	2-14
3-1 (a)	Image Orthicon	3-2
3-1 (b)	Vidicon Camera Tube	3-2
3-1 (c)	Image Dissector	3-2
3-2	Vidicon Image Pickup	3-4
3-3	Absolute Spectral Responsitivities of Several Vidicons	3-5

LIST OF ILLUSTRATIONS (continued)

VOLUME 2

Figure No.		Page
3-4	Intrinsic Photoconductivity Response of Typical Photoconductors vs Wavelength, Where "Intrinsic" Means Response of the Material Itself Without Incorporated Impurities	3-12
3-5	The Geometry of a Scanning Beam Traversing a Scanning Line	3-16
3-6	Optical Linear System Model	3-24
3-7	Spatial Frequency in Cycles/Foot Versus MTF	3-25
3-8	Plot of MTF, 7 Micron Fiber Optics	3-28
3-9	Modulation Transfer Function for Laser Scanner on Axis with Diffraction Limited Optics (Symmetrical Aperture) and with a Gaussian Point Spread Function	3-30
3-10	Normalized Step Response for a Gaussian Spot Scanning a Knife Edge	3-34
3-11 (a)	Assumed f(X) for Image Motion MTF Derivation	3-36
3-11 (b)	Fourier Transform of $y = f(X)$	3-36
3-12	Plot of MTF, High Resolution CRT	3-39
3-13	MTF of an Image Dissector with a Rectangular Aperture of Width a	3-41
3-14	Normalized Amplitude Spectrum for Coherent Side Looking Radar	3-45
4-1 (a)	Output Versus Input Wave Forms of the Stages of the Folding Coder	4-3
4-1 (b)	N-Stage Serial Type of Folding Coder for Analog-to-Digital Conversion	4-3
4-2	Details of One Stage of the Fisher Microwave Analog/Digital Converter to Gray Code Including Microwave Circuit Equivalent of One Stage for 5-bit 1200 Megabit Transmission	4-5
4-3	Encoder Logic	4-7
4-4	Waveform Diagram of u (θ_n) and $F(\pi)$	4-11
4-5	Block Diagram of Binary Encoder	4-12

LIST OF ILLUSTRATIONS (continued)

VOLUME 2

Figure No.		Page
4-6	Circuit Diagram of a One-Gigahertz Shift Register Stage Due to B. Sear. Tunnel Diodes are Used.	4-13
4-7	Diffuse Threshold Decoder/Encoder (N = 2) Interleaved to Order 2	4-15
4-8	Cathode-Ray Tube A/D Converter Produces a Digital Output as the Electron Beam is Moved Across the Aperture Plate by the Analog Voltage	4-17
4-9	One-Bit Delta Modulation System	4-22
4-10	Multibit Delta Modulation System	4-23
4-11	Tunnel Diode Delta Modulator for Video Information after Balder and Kramer	4-24
4-12	Delta Modulator Model	4-26
4-13	Typical 2-Bit Exponential Quantizer Characteristic (4 levels). An Infinite Time Constant for the Integrator is Assumed.	4-28
4-14 (a)	Response of 2-Bit Delta Modulator to a Ramp Function	4-29
4-14 (b)	Response of 2-Bit Delta Modulator to a Ramp Function	4-29
4-15	Typical 4-Bit Exponential Quantizer Characteristic (16 levels). An Infinite Time Constant for the Integrator is Assumed. The Negative Portion (incompletely shown) is Obtained by Reflecting about the Origin.	4-31
4-16	Response of 4-Bit Delta Modulator to a Ramp Function Width of Ramp Portion = 1.25 x Sampling Period	4-32
4-17	Mean Percent Error of Output of 3-Bit Delta Modulator vs Median Slope of Quantizer Characteristic	4-35
4-18	Mean Percent Error of Output of 3-Bit Delta Modulator vs Median Slope of Quantizer Characteristic	4-36
4-19	Block Diagram of Quantizer	4-39
4-20	Quantizer Input-Output Characteristic	4-39
4-21	Quantizer Noise as a Function of the Input	4-39

LIST OF ILLUSTRATIONS (continued)

VOLUME 2

Figure No.		Page
5-1	Quadriphase Phasor Diagram	5-3
5-2	Quarternary Truth Table for Two Inputs (Binary)	5-3
5-3	Quasi-Coherent Demodulator and Differential Decoder Approach for Bi-Phase Signals	5-3
5-4	Differential Demodulation and Detection Approach for Bi-Phase Signals	5-3
5-5	Block Diagram of Quadri-Phase Modulator	5-6
5-6	Phasor Diagram for $\phi_e(t-1)$ as a Function of Variables x_1'' and y_1''	5-6
5-7	Block Diagram of Quarternary Differential Encoder	5-8
5-8	Degradation as a Function of Channel Power Imbalance Factor	5-11
5-9	Phasor Diagram Illustrating (Exaggerated for Convenience) Effect of (1,0) Pulse Amplitude Imbalance	5-11
5-10	Average Power Loss Due to (1,0) Pulse Amplitude Imbalance	5-13
5-11	Phasor Diagram Illustrating (Exaggerated for Convenience) (1,0) Imbalance Effect with AM-PM Conversion Distortion Shown	5-13
5-12	Average Total Steady State Loss Due to (1,0) Pulse Amplitude Imbalance - Includes Average Power Loss, AM-PM Distortion Loss, Compensation.	5-15
5-13	Average Loss Due to (1,0) Phase Imbalance Under Worst-Case Assumptions	5-17
5-14	Block Diagram of Receiver	5-19
5-15	Block Diagram of One Possible Quadri-Phase Demodulator	5-19
5-16	Block Diagram of Matched Filter Detectors	5-21
5-17	Block Diagram of Quarternary Differential Decoder	5-23
6-1	Antenna Pointing Uncertainty Geometry	6-2
6-2	Search Sector	6-2

LIST OF ILLUSTRATIONS (continued)

VOLUME 2

Figure No. Page

6-3 Basic Acquisition Geometry 6-4

6-4 Acquisition Geometry, Sensor Satellite in View 6-4

6-5 IF Bandwidth Required Due to Doppler Shift versus Carrier Frequency 6-9

6-6 Maximum Doppler Rate versus Frequency for 300 n.m. and 500 n.m. Sensor Satellite Earth Orbit 6-10

6-7 Loop Bandwidth versus Maximum Doppler Rate 6-11

6-8 Tracking Error versus Signal-to-Noise Ratio 6-13

6-9 Simplified Tracking Antenna/Receiver Configuration 6-14

6-10 Tracking Error versus Post-Comparator Differential Phase Shift and Pre-Comparator Differential Phase Shift 6-19

6-11 Error Voltage versus Angle γ

6-12 Amplitude Unbalance versus Boresight Shift 6-22

6-13 Tracking Error due to Unequal Individual Beamwidths Sum Channel Antenna Beamwidth = 3.4 mr 6-24

7-1 Gain and Beamwidth of a Parabolic Antenna (Efficiency 55%) 7-2

7-2 Comparison of the Illumination Characteristics of a Standard and High Efficiency Cassegrain Antenna System 7-4

7-3 Gain Loss vs Uniform Dish Distortion 7-7

7-4 Maximum Operating Frequency of Parabolic Antenna vs Reflector Error 7-9

7-5 Comparison of Candidate Antenna Structural Approaches 7-10

7-6 Maximum Temperature Difference (ΔT) Versus Thermal Conductivity (k) of Face Sheet Material for Constant Face Sheet Thickness (t) 7-14

7-7 (a) Calculated Radiation Pattern for 8' Dish at 60 GHz - No Sun 7-16

7-7 (b) Calculated Radiation Pattern for 8' Dish at 60 GHz - Worst Case Sun, Graphite/Epoxy 7-17

LIST OF ILLUSTRATIONS (continued)

VOLUME 2

Figure No.		Page
7-7 (c)	Calculated Radiation Pattern for 8' Dish at 60 GHz - Worst Case Sun, Invar	7-18
7-7 (d)	Calculated Radiation Pattern for 8' Dish at 60 GHz - Worst Case Sun, Beryllium	7-19
7-8	Temperature Difference and Gradient Versus Sun Angle	7-20
7-9	Antenna Feed System Block Diagram	7-22

LIST OF TABLES

VOLUME 2

Table No.		Page
2-1	Comparison of Digital Technique	2-2
3-1	Electrical Optical Analogs	3-14
3-2	Gaussian Shaped Modulation Transfer Function	3-31
4-1	Optimized Values for L(1), G for 2-Bit Delta Modulator	4-30
4-2	4-Bit Delta Modulator Response for $A=1$, $N_1=50$	4-30
4-3	Number of Cases for Which AMNER 2.6% vs G	4-37
4-4	Comparison of the First- and Second-Order Statistics of $M(t_k)$ and $G(t_k)$ when the Characteristic Function of $x(t)$ Satisfies (1) or has a Gaussian Shape	4-42
4-5	Evaluation of $E\ M(t_k)$ when x is a Uniformly Distributed Random Variable with Mean m and Variance $\frac{\delta^2}{3}$	4-42
4-6	Evaluation of $\left[E\ M(t_k)\right]$ when $f(x) = \frac{\alpha}{2} e^{-\alpha(x-m)}$	4-43
5-1	Transmitter Truth Table	5-7
5-2	Truth Table for Differential Encoder Output Variables x_2'' and y_2''	5-9
5-3	Differential Decoder Truth Table	5-23
7-1	Efficiency Comparison	7-5
7-2	Average Beam Efficiency of Feed Horns	7-23
7-3	Tracking Feed Characteristics	7-23

SECTION 1

NANOSECOND SWITCHING TECHNOLOGY

1.1 THE ROLE OF NANOSECOND SWITCHING TECHNOLOGY IN DIGITAL COMMUNICATIONS

The processing of analog information into digital bit streams and vice versa will, in general, depend upon the technology of switching circuits and fast rise time pulse devices, providing a sequence of switching or gating operations controlled by a timing signal known as the clock. A sequence of "logic levels" which imply the sequence of decisions or actions are necessary to develop the digital signal.

Memory, too, is one of the requirements of an analog-to-digital system, particularly if storage or counting is required, and bi-stable circuits having either of two discrete states - such as in a multivibrator - such as on-off or present or not-present must be used. These states can be described in terms of voltage levels such as are used for low data rate logic circuits; either of two phases of a carrier; either of two frequencies; the presence or the absence of a carrier; or the direction of a magnetic field.

A series of "logic levels" or circuit actions is required in any processing of digital information, and for data rates in the 40 to 400 Mbps range, nanosecond and subnanosecond pulse or switching technology is required.

Four hundred Mbps data rate requires two nanosecond switching to handle the first harmonics of switching pulses at this rate, and subnanosecond switching to include higher harmonics of the pulse waveforms. As many as eight to ten logic levels may be required to convert analog information; it is, therefore, necessary that the switching, gating and storage elements of microwave digital circuitry be capable of switching at much higher rates corresponding to speeds from 100 to 300 picoseconds. The technology of super-high rate digital communication, therefore, must rely on microwave components capable of such speeds.

1.2 PULSE RISE TIMES

Figure 1-1 relates typical candidate switching devices to rise time and switching voltage.

For fastest rise-time pulses, the favorite device is a tunnel diode. TD generators available today have rise times of about 20 ps. This is a best-guess figure, for the rise time of the 28 ps sampler used to observe it increases the observed rise time to about 35 ps (rise times combine approximately in rms fashion).

Exactly how fast a tunnel diode can switch is open to some question. The fastest unpackaged tunnel diodes run around 7 to 10 ps, but the effect of the parasitic elements associated with them and their enivronment raises that figure to about 15 ps.

The speed of the voltage step generated by a tunnel diode is an intrinsic property of the diode itself, so does not depend on the speed of the input trigger waveform. Less fortunate is the intrinsic maximum amplitude of that step, which is limited to about 500 mV.

Another type of diode that can switch in the subnanosecond region is the step-recovery (also called Boff or snap-off) diode. These can give 100 to 500 ps rise time and output of several volts.

Subnanosecond rise time pulses can be generated with a charged line and a tunnel diode or step recovery diode switch. Approximately 100 ps rise time pulses of up to kilovolt amplitude can be generated by charging a length of transmission line, then shorting it with a switch.

Transistors, too can be used to generate fast rise time pulses provided that the transistor is not switched into saturation, and can be clamped to prevent overshoot on turn-on or turn-off. Transistors in integrated circuits can now be switched with nanosecond rise times and are presently the principle devices of modern high speed digital communications.

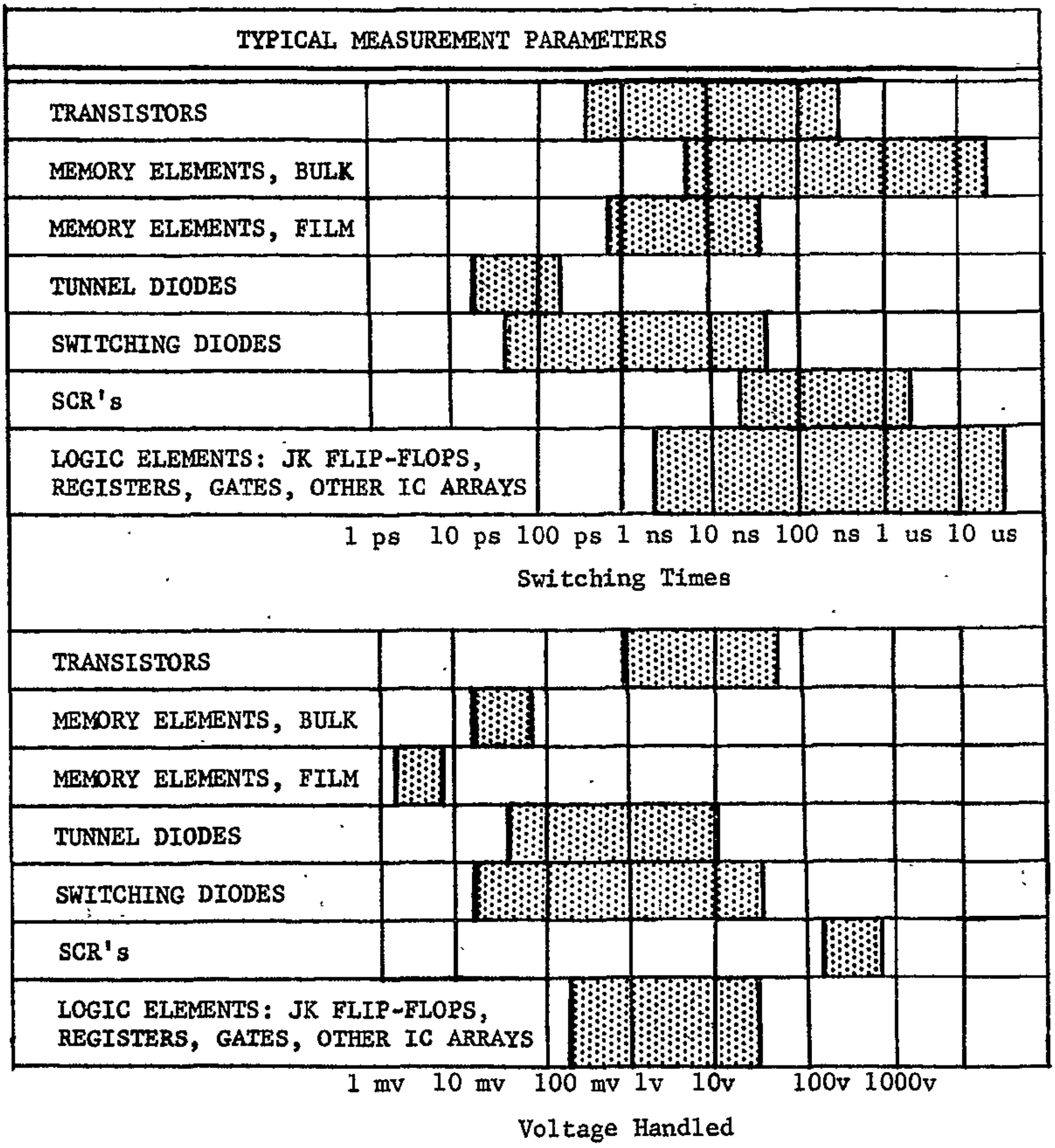


Figure 1-1 Switching Times and Voltage Levels of Modern Switching Devices

Regardless of type of switch, care must be taken with the overall switch circuit otherwise the fast rise time pulse will be degraded. To illustrate, consider a tunnel diode generator with specified rise time of 20 ps. By adding a fast sampler so that the pulse can be viewed on a sampling scope typically will increase the observed rise time to 35 ps.

The shorter the distance and the fewer the fittings or connections, the less altered will be the pulse characteristics.

The principle devices mentioned above will be discussed in the following paragraphs.

1.3 TRANSISTORS

Figure 1-2 shows a range of different transistor logic circuits which are in present use. Transistors can now be designed for f_T's well into the microwave frequencies and high logic or switching rates are possible up to 1.5 nanoseconds with the current mode logic - also known as emitter-coupled logic.

The most commonly used arrangement in computers, TTL or transistor-transistor logic is not a good candidate for gigahertz-rate systems because of low propagation delays of around 10 nsec due to these elements being driven between saturated states thereby requiring time for charge to drain or change. Current mode logic, on the other hand, uses transistors in linear operation and the present speeds now involve one nanosecond delays.

As the speed increases, the transistor equivalent circuit becomes complex and the propagation and delay intervals will be limited by the reactances, stray capacitors, and inductive wiring of the circuit.

The rise time t_r of a typical basic transistor switch shown in Figure 1-3 can be found using the following parameters:

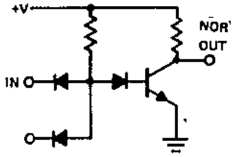
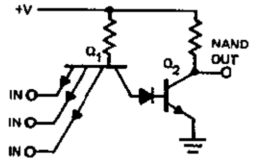
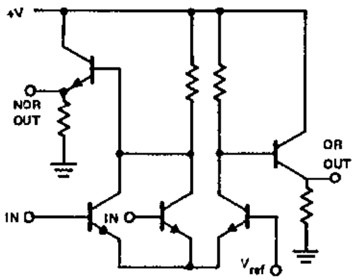
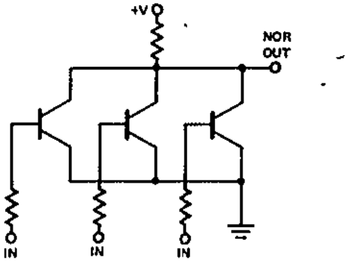
DIODE-TRANSISTOR LOGIC	TRANSISTOR-TRANSISTOR LOGIC	CURRENT-MODE LOGIC	DIRECT-COUPLED TRANSISTOR LOGIC	
				
Accepts large number of inputs; low cost; easy to fabricate; high noise margin	Slightly better speed and power dissipation than diode-transistor logic	Highest speed logic form - non-saturation, complementary outputs available	Relatively fast; low-cost; easy to fabricate	Advantages
Large signal swing, relatively slow speed	Low noise margin; inverse gain of Q limits number of outputs	Needs reference voltage; high power dissipation	Unequal splitting of currents in circuit can be problem; low noise margin	Disadvantages

Figure 1-2 Comparison of Different Types of Transistor Logic Circuits

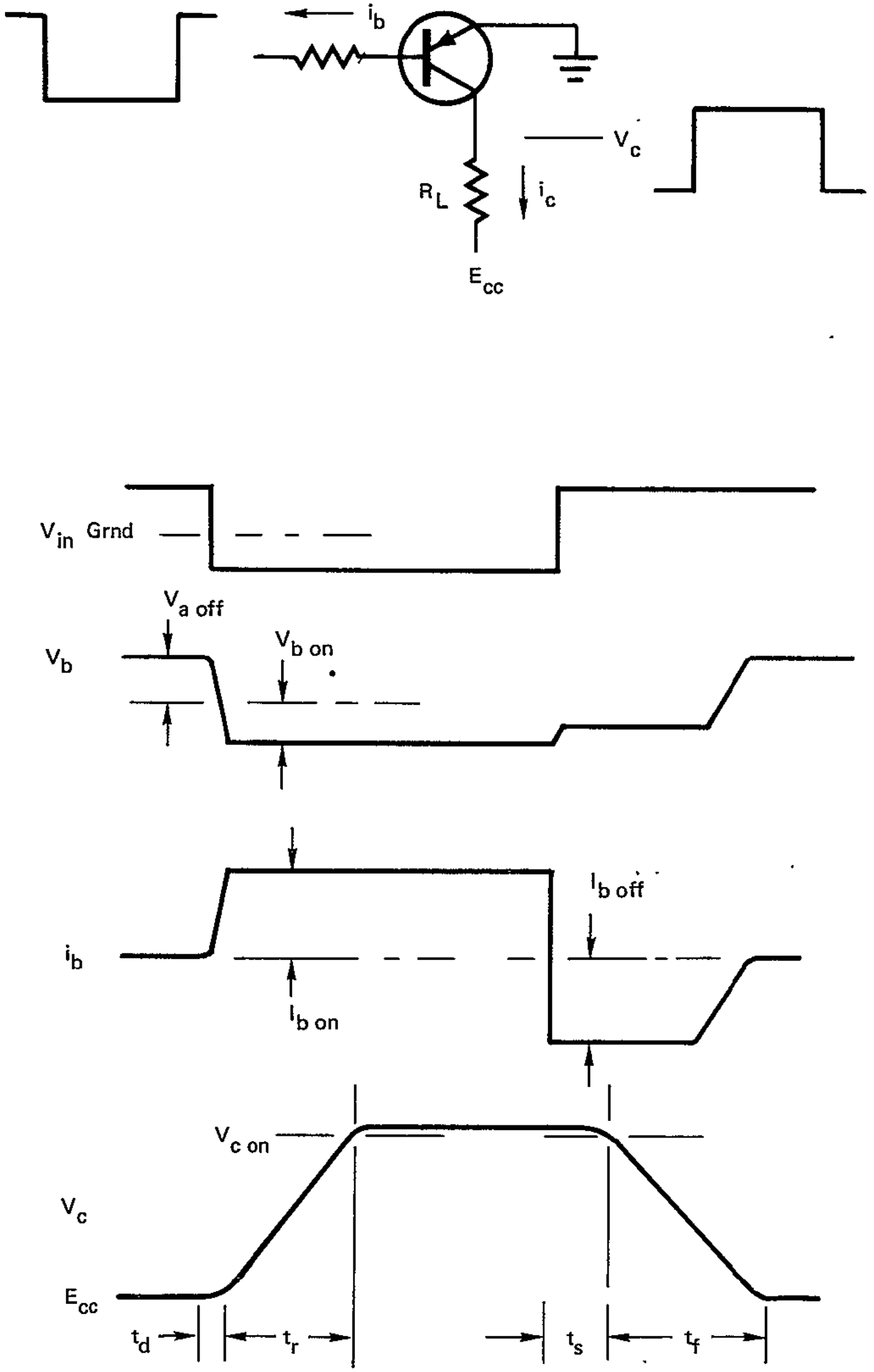

i_b
V_c
R_L
i_c
E_{cc}
V_{in} Grnd
$V_{a\ off}$
V_b
$V_{b\ on}$
i_b
$I_{b\ off}$
$I_{b\ on}$
$V_{c\ on}$
V_c
E_{cc}
t_d
t_r
t_s
t_f

Figure 1-3 Simplified Transistor Switching Circuit and Switching Characteristics

ω_T = $2\pi f_T$ where f_T = gain-bandwidth product

I_c = Collector current

I_b = Base current

C_c = Collector capacitance

R_L = Collector load resistor

β_{s0} = Current gain at the edge of saturation

According to J. Kuno[1],

$$t_r = \left(\frac{I_c}{I_{b\ on}}\right)\left[\frac{\dfrac{1}{\omega_T(I_c,\ V_{c\ on})} + \bar{C}_c R_L}{1 - \dfrac{1}{2\beta_{s0}\left(\frac{I_c}{2}\right)}\left(\dfrac{I_c}{I_{b\ on}}\right)}\right]$$

Similarly, Kuno has described an equation expressing the fall time t_f. However, in this case the reverse driving current does not drive the transistor in the same way as the forward driving current. Thus the expression for the voltage fall time must be modified to take $R_L C_c$ fall time into consideration. Then the fall time is expressed by

$$t_f = \left(\frac{I_c}{I_{b\ off}}\right)\left\{\frac{\dfrac{1}{\omega_T(I_c,\ V_{c\ on})}}{1 + \dfrac{1}{2\beta_{s0}\left(\frac{I_c}{2}\right)}\left(\dfrac{I_c}{I_{b\ off}}\right)}\right\} + 2.3\, R_L \bar{C}_c$$

where the last term is $R_c \bar{C}_c \ln \frac{I_c}{0.1 I_c} = 2.3 R_L \bar{C}_c$; that is, the fall time if I_c falls to zero very quickly.

---

[1] J. Kuno, "Rise and Fall Time Calculations of Junction Transistors", IEEE Transactions on Electron Devices, pp 151 - 153, April 1964.

1.4 TUNNEL DIODES

Tunnel diodes have had wide application to digital logic circuits.[1,2,3] As shown in Figure 1-4, the shape of the tunnel diode current-voltage curve is ideal for switching between two states when used with the proper resistive loadline; a balanced tunnel diode pair referred to as the Goto-pair can be operated as a flip-flop.

Tunnel diodes are now available with resistive cutoff frequencies 30 GHz from companies such as KMC Semiconductors and are packed in PILL-type cartridges which have minimum series inductance.

Tunnel diodes are also ideally adapted to integrated circuits techniques. In general, the 100-300 picosecond switching speeds required of one-half gigahertz-rate bit streams are well within the capability of the present-day tunnel diode.

1.5 SNAP BACK DIODES

Snap-back or step recovery diodes have voltage waveforms,[4,5] as shown in Figure 1-5a. In the snap-back region, excellent pulses can be developed. The charge storage rates of a snap-back diode have been shown by B. Sear to generate rise times of less than 100 picoseconds at repetition rates exceeding 200 MHz and future rates have been predicted in excess of one gigahertz.

---

[1] R. C. Sims, E. R. Beck and V. C. Kamm, "A Survey of Tunnel Diode Digital Techniques", Proc. IRE, pp 136-146; January, 1961.

[2] E. Goto, et. al., "Esaki Dilde High Speed Logic Circuits", IRE Trans. on Electronic Computers, pp 25-29; March, 1960.

[3] V. F. Chow, "Tunnel Diode Digital Circuitry", Digest of 1960 International Solid-State Circuits Conference, pp 32-33.

[4] J. Moll, S. K. Rakaner, and R. Shen, "P-n Junction Charge Storage Diodes", Proc. IRE, pp 43-53; January, 1962.

[5] B. E. Sear, "Charge Controlled Nanosecond Logic Circuitry", Proc. IEEE, pp 1215-1227; September, 1963.

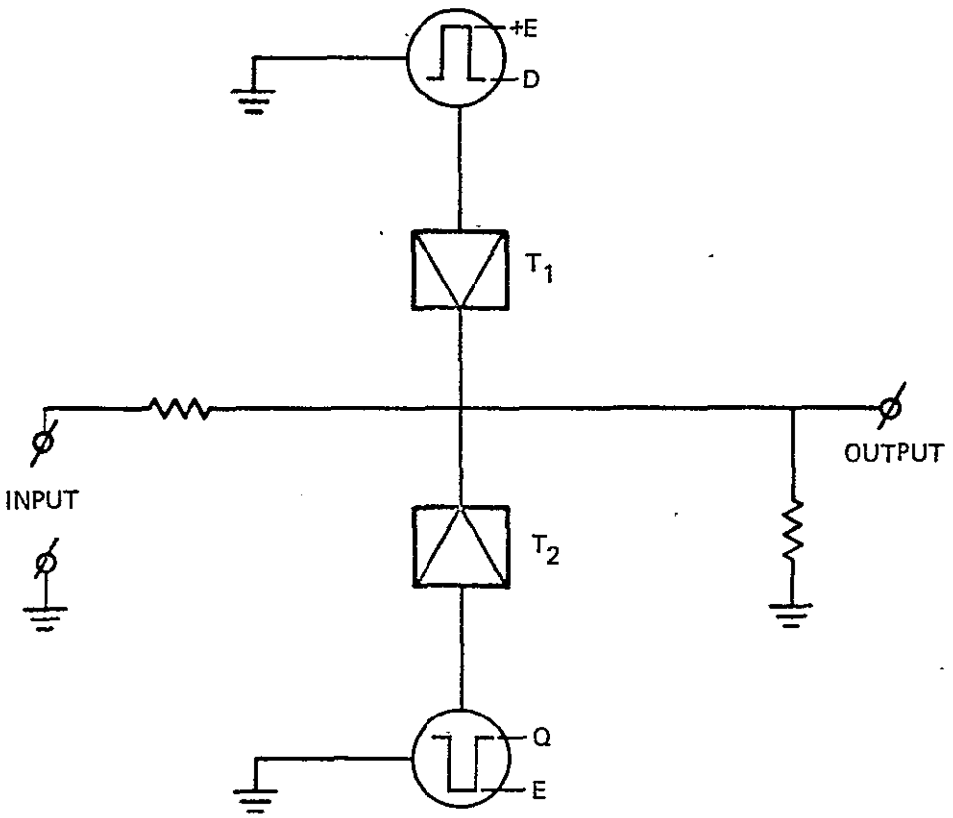

+E
D
T1
INPUT
OUTPUT
T2
Q
E

Figure 1-4 Tunnel Diode Balanced Pair (GOTO Pair)

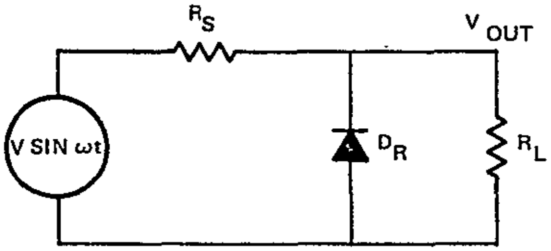


Figure 1-5a Simplified Snap-Diode Pulse Generator

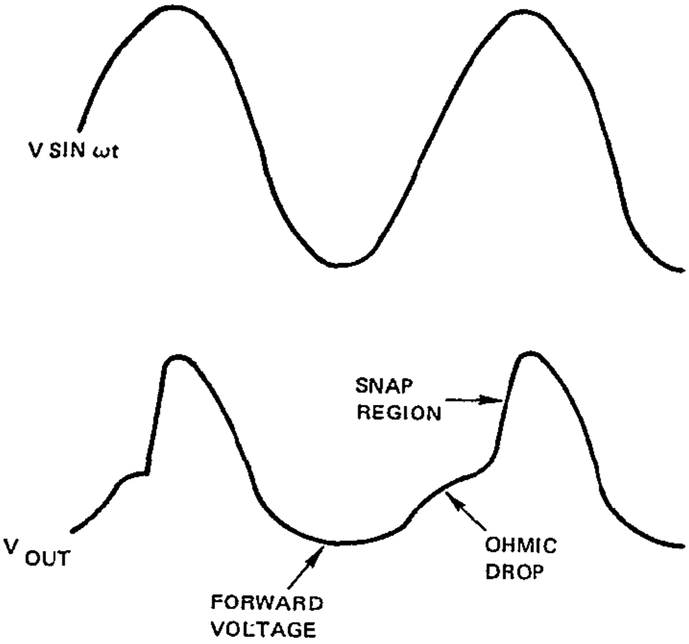


Figure 1-5b Typical Waveforms of Pulse Generator

1.6 PHASE LOCKED OSCILLATORS

The phase locked oscillator (PLO) was the first solid-state two-state device considered for high speed computers. The basic parametric phase-locked oscillator, due to Sterzer and Beam[1] is shown in Figure 1-6. It consists of a tank circuit which is forced to provide parametric oscillations by a pump operating at twice pump frequency. Sustained oscillations can occur in either of two possible phases 180° apart. The circuit can be caused to change phase states by "Locking" the oscillations to the phase state of a locking signal.

Computer logic systems using phase-locked oscillators were built in 1958-1959 using clock rates above one gigahertz, performing logic at rates as high as 1.85 GHz. However, the advent of the tunnel diode and the inherent complexity of a PLO computer caused interest in this technology to wane.

1.7 BALANCED MODULATORS

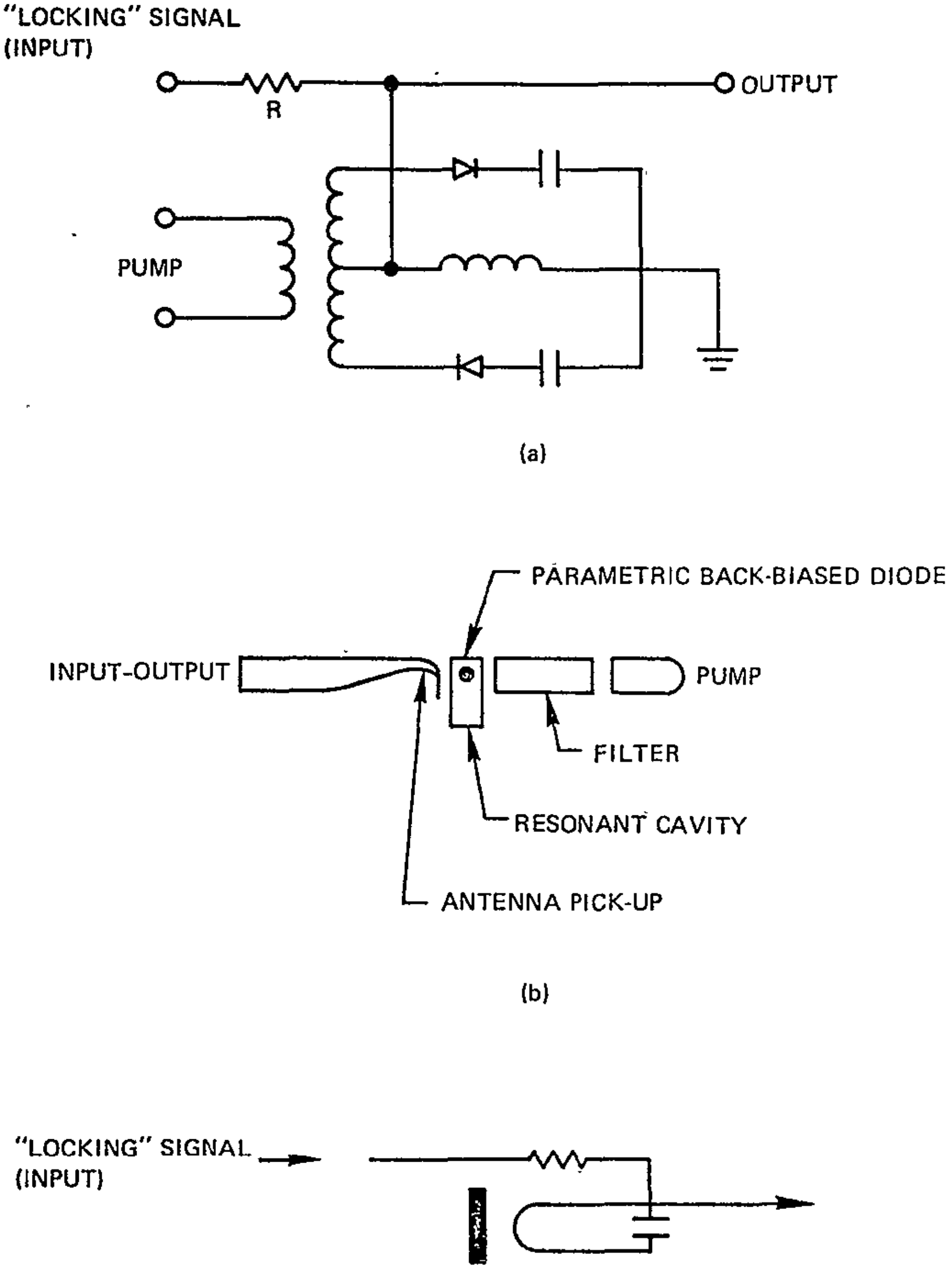
The use of the balanced modulator as a logic element was suggested in 1959 by Eckhardt and Sterzer[2]. Such circuits are now in use in microwave frequency applications and are readily adaptable for logic or switching functions at gigahertz rates.

Balanced modulator circuits can be used to develop most logic functions. An AND circuit, for example, is formed by two modulators in series; the modulator transmits an rf signal only if a baseband signal is applied to it.

---

[1] F. Sterzer and W. Beam, "Parametric Subharmonic Oscillators", Digest of 1959 Solid-Stages Circuits Conference; February, 1959.

[2] W. Eckhardt and F. Sterzer, "A Modulation - Demodulation Scheme for Ultra-High-Speed Computing and Wideband Amplification", Digest of 1960 International Solid-Stage Circuits Conference; February, 1960.


"LOCKING" SIGNAL
(INPUT)
R
OUTPUT
PUMP
PARAMETRIC BACK-BIASED DIODE
INPUT-OUTPUT
PUMP
FILTER
RESONANT CAVITY
ANTENNA PICK-UP
"LOCKING" SIGNAL
(INPUT)
PUMP

(a)

(b)

(c)

Figure 1-6 .—Parametric phase-locked oscillator circuits (a) one possible lumped-parameter variable-capacitance PLO circuits, (b) one possible microwave PLO (in microstrip), (c) symbolic representation of PLO. - after L. Onyshkevych, et. al.

1.8 GUNN DIODE SWITCHING CIRCUITS

Two similar Gunn diodes can be connected in a Goto pair arrangement, with an input and an output resistor connected to the common node. The bias voltages are adjusted so that there is sufficient current through the diodes, in the absence of high-field domains, to cause the internal fields of both devices to be just above the threshold value. For devices with a large low-field to high-field mobility ratio and a low-contact resistance, the above bias conditions will insure that when a domain is propagating in one device, the internal field in the other device will be below the sustaining value.

If the above constraints are satisfied, the polarity of the current in the input resistor at the time of the nucleation of a domain will determine in which device a domain will form. Once a domain has formed, no other domains will nucleate (for reasonable levels of input current) during the transit time of the domain. When the domain collapses at the anode, the polarity of the input current will again determine the device in which the new domain will nucleate.

2-GHz operation has been demonstrated at Bell Telephone Laboratories by the reversal of the mode for a single Gunn domain transit time by a negative input spike of 1.4 volts (node potential) and 250 ps duration.

1.9 MICROWAVE-RATE STORAGE AND CODING CIRCUITS

Microwave-rate logic elements must be combined in circuit complexes which provide the primary functions of storage, A/D conversion and D/A conversion.

The principle type of storage circuit is the shift register whose basic circuit at microwave frequencies is shown in Figure 1-7. As indicated, the shift register is a series of flip-flops connected by a clock or timing circuit.

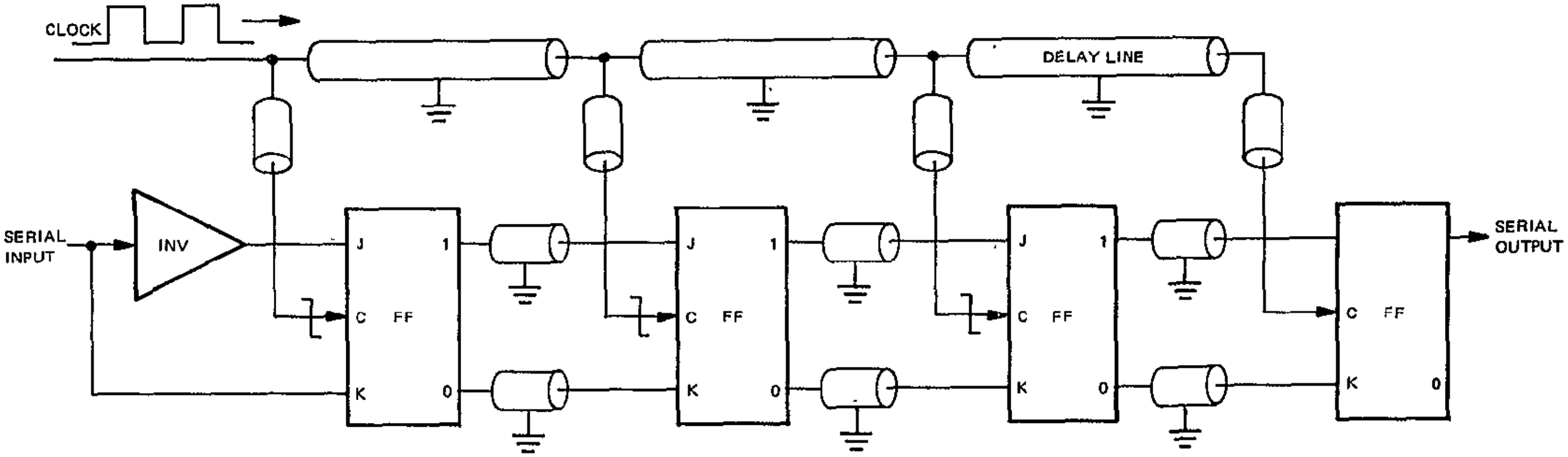

CLOCK
DELAY LINE
SERIAL INPUT
INV
J
1
C
FF
K
0
SERIAL OUTPUT

Figure 1-7 JK Flip-Flop Shift Register Showing Clock Distribution System Simulated by Delay Lines

A series of bits is applied to the input in synchronization with a clock which actuates all flip-flops simultaneously so that a given series of bits moves or progresses through the shift register and occupies the shift register - one bit per flip-flop at a rate determined by the clock. Once the shift register is occupied, the serial input is removed. Application of the clock again to the flip-flop will read out the stored bits as a serial bit stream output. If during the readout the clock is very fast, the bit stream will comprise a bit burst. For gigahertz data rates, the shift register should be capable of moving bits from flip-flop to flip-flop in time intervals equal to or less than a nanosecond; the serial burst reading can be at a much lower rate.

The shift register circuit of Figure 1-7 shows the clock routing circuit and the interconnects of the flip-flop circuits in the form of delay lines. Assuming fast flip-flop propagation times using, for example, the tunnel diode shift registers shown in Figure 1-8 due to B Sear[1] or using CML circuits, the clock routing, interconnect and bit stream distribution circuits become the limiting circuit of the system.

If the speed goal is to be at least 4 to 5 logic levels per nanosecond, relative problems of packing density, wiring propagation delays and cross-talk must be solved. If a velocity factor of about 0.5 is assumed, an interconnecting transmission line of 1 1/2 inch in length will represent a delay of one logic level since light travels around one foot in a nanosecond in such circuits.

---

[1] B. Sear, "Research for High-Speed Analog-to-Digital Conversion Techniques," WPAFB Technical Documentary Report No. AL TDR-64-154.

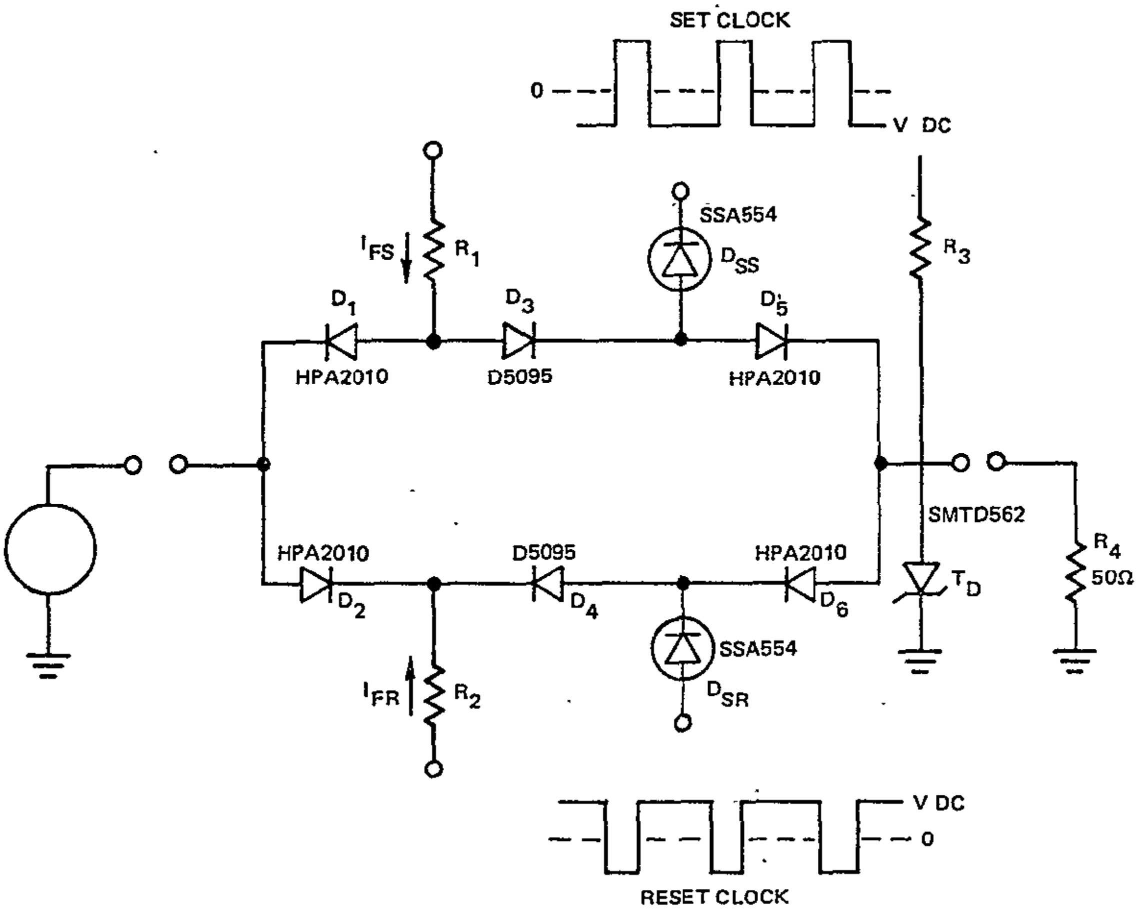

SET CLOCK
0
V DC
I_{FS}
R_1
SSA554
D_{SS}
R_3
D_1
D_3
D_5
HPA2010
D5095
HPA2010
SMTD562
HPA2010
D5095
HPA2010
D_2
D_4
D_6
T_D
R_4
50Ω
SSA554
I_{FR}
R_2
D_{SR}
V DC
0
RESET CLOCK

Figure 1-8 Circuit Diagram of a One-gigahertz Shift Register Stage Due to B. Sear. Tunnel Diodes are Used.

O. G. Gabbard[1] has investigated the problems of signal distribution in high-speed circuits recommending a stable ground reference in microstrip with signal interconnects designed such that ground return currents can flow immediately below the signal interconnect, thus providing low impedance and minimizing cross-talk when the interconnect is used by several signals. Fairly large transient currents and Ldi/dt voltages can exist during switching to drive wiring and circuit input capacity. Reducing Ldi/dt effects is one of the most important considerations for achieving good performance at high speed.

As the rise times become shorter and shorter, as will be the case in gigahertz rate circuits, the lengths of unterminated inter-connects must be very short in the digital circuit complex, otherwise reflections and ringing effects will degrade circuit operation.

---

[1] O. G. Gabbard, "High-Speed Digital Logic for Satellite Communications", Electro-Technology, pp 59-65; April, 1969.

SECTION 2

DIGITAL MODULATION TECHNIQUES

2.1 DIGITAL MODULATION TECHNIQUES

In an ideal system, transmission of digitized data can be accomplished using conventional binary techniques such as FSK or PSK, or the more efficient M-ary signaling techniques. Binary or M-ary signaling techniques can be combined with forward error control encoding to detect or correct transmission errors.

Various binary and M-ary signaling techniques which are candidates for high speed data transmission are compared on the basis of bit error rate (BER) versus energy-per-bit-to-noise-power-density ratio (E_b/N_o), and required RF bandwidth in Table 2-1.

2.2 COMPARISON OF CHANNEL EFFICIENCY OF DIFFERENT MODULATION METHODS ON THE BASIS OF BER VERSUS E_b/N_o

A comparative evaluation of the channel efficiency of digital modulation methods is based on the average probability of bit error rate (BER) as a function of the energy-per-bit-to-noise-spectral-density ratio, E_b/N_o.

2.2.1 FSK and MFSK (Non-Coherent)

Frequency shift keying is a digital modulation method in which each transmitted data symbol is represented by one of M possible frequencies. For FSK there are two possible frequencies each representing one data bit. For MFSK, M frequencies are used and each of the M-ary symbols represent $\log_2 M$ bits of data.

Non-coherent detection is normally used in FSK and MFSK systems to avoid the requirement for generating a coherent phase reference for each frequency. The optimum non-coherent receiver consists of a bank of matched

TABLE 2-1

COMPARISON OF DIGITAL TECHNIQUE

MODULATION TECHNIQUE	PROBABILITY OF BIT ERROR	BANDWIDTH
GENERAL RELATIONSHIPS	$\Pr\{\text{bit error}\} = \frac{2^{k-1}}{2^k - 1} \Pr\{\text{symbol error}\}$ $= \frac{M}{2(M-1)} \Pr\{\text{symbol error}\}.$	
MFSK (NONCOHERENT)	$\Pr\{\text{symbol error}\} = \frac{\exp\left(-\frac{E}{N_o}\right)}{M} \sum_{j=2}^{M} (-1)^j \binom{M}{j} \exp\left(\frac{E}{jN_o}\right)$	$Bw = \frac{M}{T} = \frac{M}{T_b \log_2 M} = \frac{2^k}{kT_b}$.
PSK (COHERENT)	$\Pr\{\text{bit error}\} = \Phi \sqrt{2E/N_o}$	The transmitted power spectrum of PSK is of the form $\left[\frac{\sin \pi T_b(f - f_{carrier})}{\pi T_b(f - f_{carrier})}\right]^2$. The bandwidth between the principal nulls is $2/T_b$ cps.
PSK (DIFFERENTIALLY COHERENT)	$\Pr\{\text{bit error}\} = \frac{1}{2} \exp\left(-\frac{E}{N_o}\right).$	$Bw = \frac{2}{T_b}$.
QPSK	$\Pr\{\text{symbol error}\} = 2\Phi\left(\sqrt{\frac{E}{N_o}}\right) - \Phi^2\left(\sqrt{\frac{E}{N_o}}\right)$ $= 2\Phi\left(\sqrt{\frac{2E_b}{N_o}}\right) - \Phi^2\left(\sqrt{\frac{2E_b}{N_o}}\right).$ The Pr{bit error} is given by $\Pr\{\text{bit error}\} = \Phi\left(\sqrt{\frac{E}{N_o}}\right) = \Phi\left(\sqrt{\frac{2E_b}{N_o}}\right).$	for $M = 4$, i.e., QPSK, the required bandwidth is $1/T_b$
MPSK	$\Pr\{\text{symbol error}\} \doteq 2\Phi\left(\sqrt{\frac{2E}{N_o}} \sin(\pi/M)\right).$ For $M \cdot 4$ and $\sqrt{E/N_o} \gg 1$:	$Bw = \frac{2}{T_b \log_2 M}$.
M-ARY CODED PSK (BI-ORTHOGONAL CODING)	$\Pr\{\text{symbol error}\} = 1 - \int_{-\sqrt{2E/N_o}}^{\infty} \frac{\exp\left(-\frac{x^2}{2}\right)}{\sqrt{2\pi}}$ $\cdot\left[\int_{-x-\sqrt{2E/N_o}}^{x+\sqrt{2E/N_o}} \frac{\exp\left(-\frac{y^2}{2}\right)}{\sqrt{2\pi}} dy\right]^{(M/2)-1} dx$	$Bw = \frac{2n}{kT_b}$.
PCM/FM (DISCRIMINATOR DETECTION)	$\Pr\{\text{bit error}\} = \Phi\left(1.1\sqrt{\frac{E}{N_o}}\right).$	$Bw \doteq \frac{1}{T}$.

M = number of symbols in a signaling alphabet

$k = \log_2 M$ = number of bits of information in an M-ary symbol word

n = number of binary digits in a binary code word

E = symbol energy in watt-seconds

E_b = energy per bit in watt-seconds

T = transmitted symbol duration in seconds

T_b = bit duration in seconds = $(\text{information rate})^{-1}$

N_o = one-sided noise spectral density in watts/cps

$$\Phi(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^{\infty} (-\tfrac{1}{2} y^2)\, dy$$

Bw = required transmission bandwidth in cps.

filters feeding a greatest-of comparator. The matched filters are high-Q bandpass integrate and dump filters followed by envelope detectors. Each filter is tuned to a particular symbol frequency. The energy in each filter is dumped immediately after each symbol decision to prevent intersymbol interference caused by residual energy.

In general FSK and MFSK are listed here to provide reference against phase shift keying techniques. The difficulties in building practical band pass integrate and dump filters for data rates above 10 Mbps, and the sheer impracticality of building such filters for handling the multiple frequency bands for data rates above 100 Mbps make such a modulation technique a poor candidate for high speed imagery data transmission.

2.2.1.1 PCM/FM (Discriminator detection)

Frequency modulation (FM) of a carrier by binary pulse-code-modulation (PCM) is widely used for telemetry and data transmission. In this technique, the frequency of the carrier oscillator is deviated to one of two possible discrete amounts corresponding to the binary data. The receiver consists of a frequency discriminator followed by a video matched-filter and a decision threshold. The post-discriminator matched-filter intgrates over a bit duration after which it is sampled and its energy is dumped. The sample is compared with the decision threshold and a choice made as to which binary symbol was received. PCM/FM is somewhat similar to FSK (MFSK with $M = 2$) with the distinction that FSK consists of gating on one of two independent oscillators with the binary data. A further distinction is made here in that MFSK (or FSK) is received with M noncoherent matched-filters, whereas, a discriminator and post-discriminator matched-filter are assumed for PCM/FM reception.

2.2.1.2 PSK (Coherent)

PSK refers to phase shift keying by a binary message. The phase of the carrier is shifted by 180° each time the binary data changes. Coherent detection requires that the receiver have knowledge of the phase of the carrier with an uncertainty of exactly 180°. This phase information must be derived at the receiver and complicates the receiver implementation.

PSK and PCM/FM are identical when ± 90° phase deviation is used in PCM/FM. Phase reversal keying (PRK) and PSK are also identical.

In the transmitted power spectrum of PSK, the bandwidth between the principal nulls is $2/T_b$ Hz. However, it is common practice to band-limit the spectrum to $1/T_b$ Hz. This corresponds to the -4 db bandwidth. The required transmission bandwidth for PSK is thus taken to be $Bw = 1/T_b$.

2.2.1.3 PSK (Differentially Coherent)

In this technique, the binary data is detected by comparing the carrier phase of a received bit with the carrier phase of the preceding bit. The receiver, therefore, can take advantage of the differential phase coherence of successive bits to detect a phase change without requiring knowledge of the absolute carrier phase.

Since the detection process involves a noisy phase reference (the preceding bit) the error probability for differentially coherent PSK is higher than for PSK using phase coherent detection.

2.2.1.4 MPSK (Coherent)

In multiple phase shift keying (MPSK), the signaling alphabet consists of M equally spaced carrier phase shifts. For example, for $M = 4$, the possible phase shifts are 0°, 90°, 180°, and 270°.

The Pr (symbol error) given above for MPSK is an upper bound that becomes quite accurate for large M. MPSK for $M = 4$ is referred to as 4-phase PSK or quadriphase modulation.

2.2.1.5 M-ary Coded PSK (Orthogonal Coding)

The performance of any binary signaling technique can be improved by encoding the data into binary code words. This coding can take many forms but can be categorized by its intended purpose. Coding schemes are available for correcting or detecting errors in the reception process while other schemes are used to minimize detection errors. The techniques described in this section utilize coding to minimize detection errors.

Just as in MFSK, the signaling performance of PSK can be improved if the binary data, k bits at a time, is encoded into an alphabet of M orthogonal symbols. In M-ary coded PSK with orthogonal coding, the signaling alphabet consists of $M = 2^k$ mutually orthogonal binary words each consisting of n binary digits $(n>k)$. Code words can be readily generated which have $n = M$, if the system can accomodate the additional speed required.

2.2.1.6 M-ary Coded PSK (Bi-Orthogonal Coding)

A bi-orthogonal signaling alphabet for M-ary coded PSK can be constructed by simply taking the complement or negative of each code word in an orthogonal alphabet. Thus, the number of available symbols is doubled and the average symbol correlation coefficient is $-1/(M-1)$ instead of zero as in the case of orthogonal coding. Thus, the performance using bi-orthogonal coding should be better than that of orthogonal coding, though the improvement becomes negligible for large M.

2.2.2 Comparison Curves and Discussion

The performance of MFSK based on probability of error is plotted and compared with other signaling techniques in Fig. 2-3. The probability of bit error for MFSK for various values of M is plotted as a function of E_b/N_o. For a specified error rate, the reduction in required energy per bit is quite significant for a moderate increase in M. Beyond $M = 16$, the improvement is less dramatic while the complexity of the system increases considerably and emphasizes the point made earlier; i.e. that

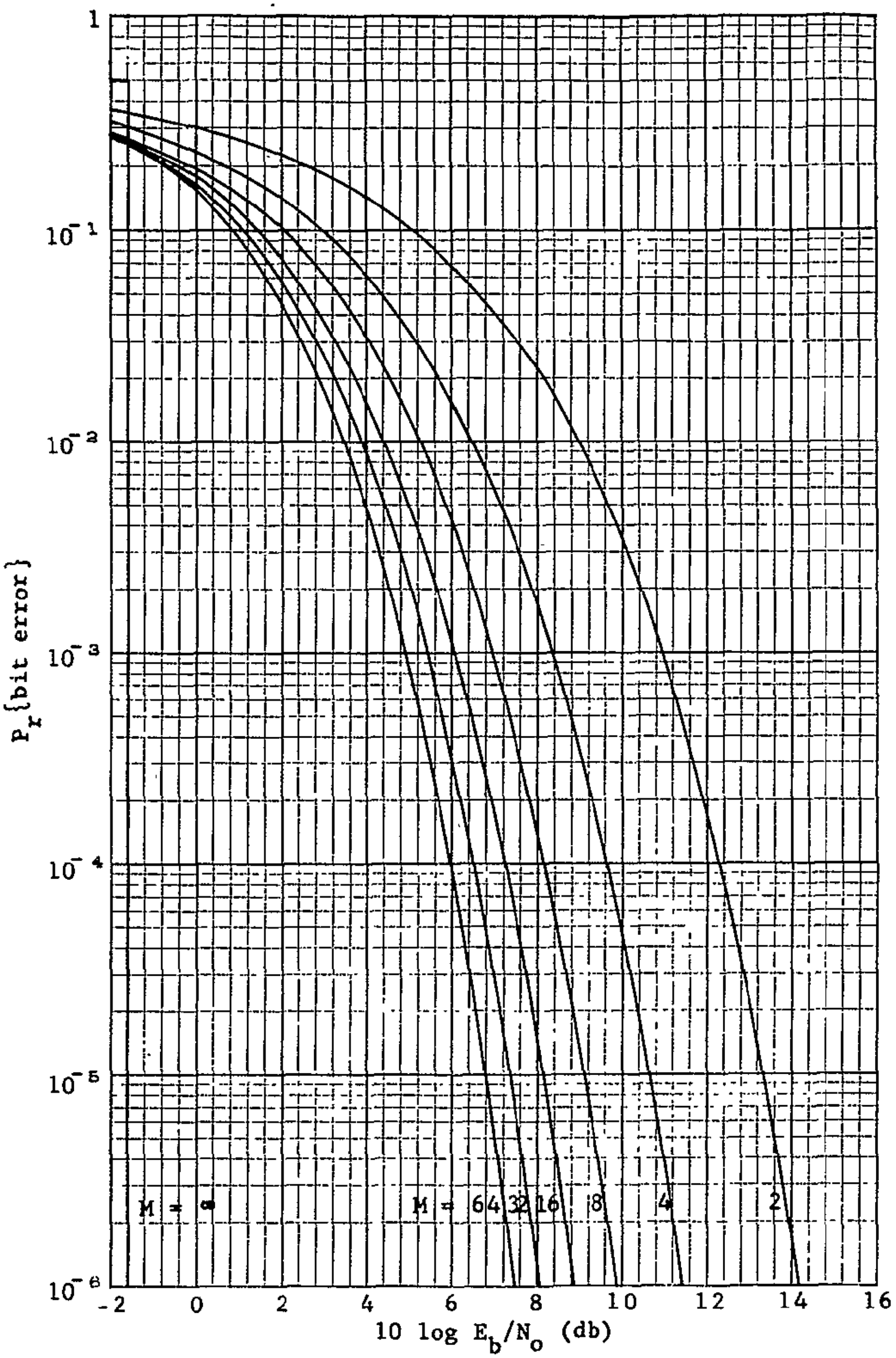


Fig. 2-1 Binary error probabilities for MFSK

MFSK is impractical for the data rates used in this study.

Fig. 2-4 compares 16-ary MFSK and QPSK with respect to other techniques.

2.2.3 Comparison of Digital Modulation Techniques on the Basis of Required Bandwidth

Transmission bandwidth becomes an important criterion when channel bandwidth is limited or when transmitter and receiver components are significantly bandwidth limited.

For data relay satellite links using millimeter wave carriers, power efficiency may prove to be more important than bandwidth efficiency. However, any practical link will experience some bandwidth limiting due to component band-limiting and intentional filtering to avoid interference with other links.

The bandwidth characteristics of each of the modulation methods covered in paragraphs 2.2.2 above will be considered. The bandwidths are expressed as functions of R_b, the equivalent data rate over each link in bits per second, and are listed in Table 2-1.

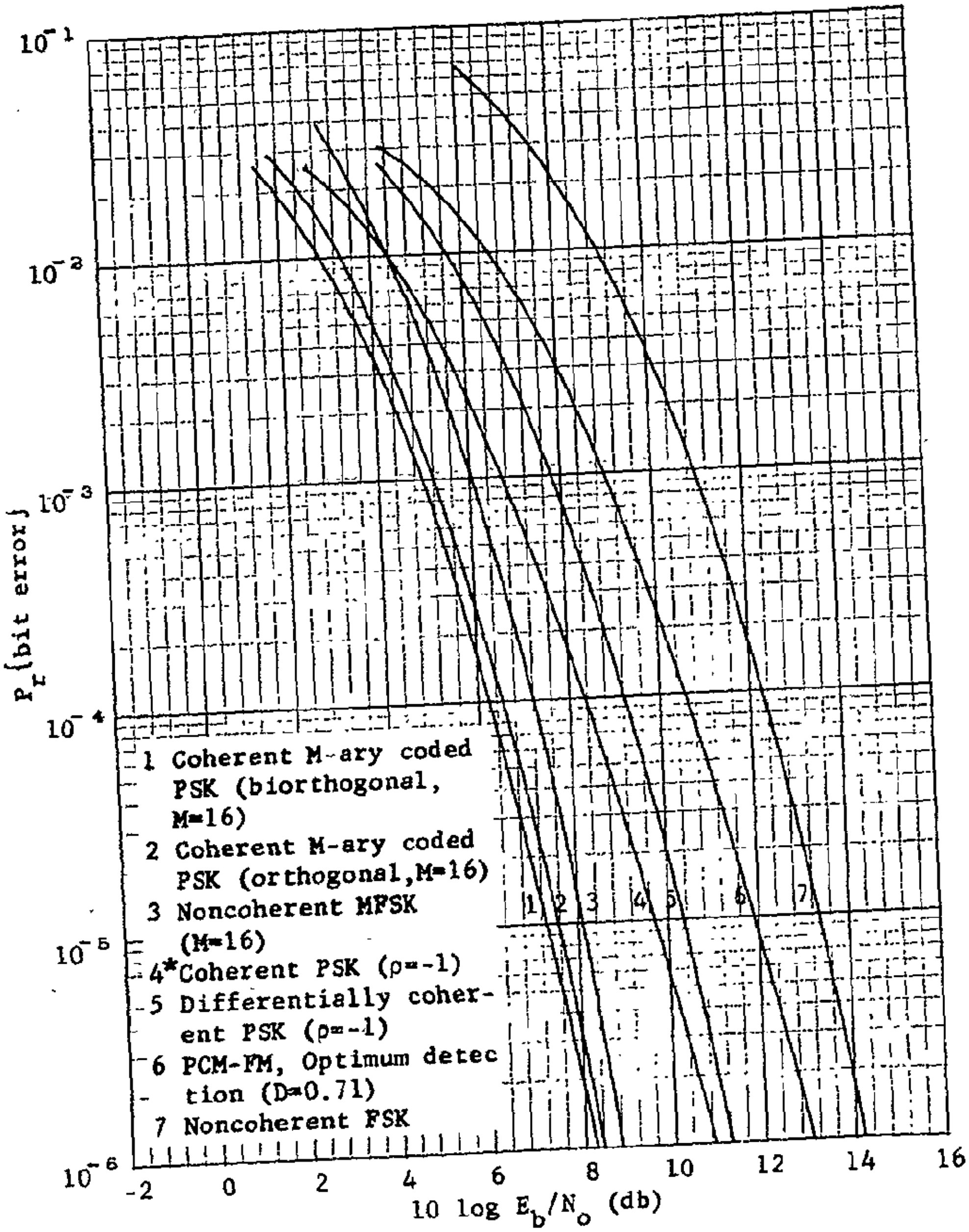
2.2.3.1 FSK and MFSK (Non-Coherent)

The frequency spacing between each of the M frequencies in an MFSK signal is chosen as T^{-1} (reciprocal of the signalling rate). This choice of spacing provides an orthogonal set of signal waveforms for the minimization of BER as well as minimizing transmission bandwidth and symbol cross-talk.

For binary FSK, $M = 2$ and $k = 1$ so that:

$$BW_{FSK} = 2R_b.$$

The spacing of FSK signalling frequencies is sometimes increased (with a corresponding increase in bandwidth) to permit the utilization of simpler receiver data filters without causing symbol crosstalk.



* Also for coherent quadri-phase

Fig. 2-2 A comparison of M-ary and binary signaling techniques

2.2.3.2 PCM/FM

The optimum peak-to-peak deviation for coherently detected PCM/FM is 0.707 R_b. However, almost all practical PCM/FM links employ discriminator detection. Depending upon the amount of premodulation filtering and the type of post-discriminator data filter, the peak-to-peak deviation is normally chosen between $0.75R_b$ and $0.9R_b$. Bennett and Rice determined that the required transmission bandwidth for a peak-to-peak deviation of 0.8 R_b is approximately equal to R_b. Thus:

$$BW_{PCM/FM} = R_b.$$

2.2.3.3 PSK (Coherent)

The unfiltered output power spectrum from an ideal biphase PSK modulator is a sin X/X spectrum with $2/T_b$ between principal nulls.

This ideal spectrum will be modified in an actual data link because of modulator pulse response (switching time), transmitter and receiver filtering, and ionospheric phase distortion. Transmitter filtering is required to reduce the interference by the signal sidelobes with adjacent channels and other services. The filtering, however, will introduce demodulator degradation because of loss of the energy in the removed sidelobes and phase distortion introduced by the filter. Therefore, the choice of bandwidth restrictive filtering must include an analysis of the resulting demodulator degradation.

2.2.3.4 PSK (Differentially Coherent)

If the same signaling rate is used for both coherent and differentially coherent PSK the required transmission bandwidths are identical; i.e., $2R_b$.

2.2.3.5 MPSK (Coherent)

In MPSK, each transmitted symbol carries $\log_2 M$ data bits. The actual signaling rate is, therefore, lower than the data rate R_b by this factor of $\log_2 M$. Thus, the required transmission bandwidth for MPSK is:

$$BW_{MPSK} = \frac{BW_{PSK}}{\log_2 M} = \frac{2R_b}{\log_2 M}$$

The principal communications advantage of MPSK over PSK is the reduced transmission bandwidth. Beyond M=4 (QPSK), however, the BER performance is reduced for increasing M. Thus MPSK trades power efficiency for bandwidth efficiency.

2.2.3.6 M-ary Coded PSK

The transmission bandwidth for coded PSK is n/K times the bandwidth required by biphase PSK having the same information data rate R_b since each data bit is essentially encoded into n/K transmitted bits. For orthogonally coded PSK, n=M and k>1 so that:

$$BW_{ORTH.\ CODED\ PSK} = \frac{n2R_b}{k} = \frac{2MR_b}{\log_2 M} .$$

For bi-orthogonally coded PSK, n=M/2 and k>1 so that:

$$BW_{BI\text{-}ORTH.\ CODED\ PSK} = \frac{MR_b}{\log_2 M} .$$

2.2.4 Bit Error Rate Degradation of BPSK and QPSK Due to Filtering

A computer program has been developed at Philco-Ford that determines the degradation introduced by filtering. A five-pole, 0.1 dB ripple Chebychev filter provides good, low-loss, side lobe rejection. More poles in the filter make the filter larger without significantly increasing sidelobe rejection; fewer poles do not give adequate sidelobe rejection.

The filter can simulate the entire channel through which the data is transmitted as long as limiters are not used.

Other filter types considered were Bessel filters (because of their linear phase characteristics) and Butterworth filters (because of their simplicity). The Chebychev filters were selected because of their excellent skirt selectivity, low ripple and reasonably good phase response. Odd-order filters were found to be preferable to even-order for both biphase and quadriphase because odd-order filters cause less quadratic phase distortion.

Fig. 2-3 is a plot of the degradation introduced in PSK signalling by the filter for 0.1 dB Chebychev filters having 3, 5, 7, and 9 poles.

In Fig. 2-4 the BER degradation for BPSK is indicated for various filter bandwidths. The filter type is a five-pole, 0.1 dB Chebychev filter. The bandwidths indicated are 3 dB bandwidths. Note that at a BER of 10^{-4} only about 1 dB degradation is introduced by filtering at the first spectral nulls ($BW = 2R_b$). Filtering at the bit rate R_b, introduces 4.4 dB degradation. It is concluded from this analysis that the practical transmission bandwidth for PSK is:

$$BW_{PSK} = 2R_b.$$

The signalling rate for QPSK is one-half the signalling rate (bit rate) for biphase PSK. Thus, all filter bandwidths can be reduced by a factor of two for QPSK. Fig. 2-5 is a replot of Fig. 2-4 for QPSK and narrower filter bandwidths.

2.3 COMPARISON OF ANALOG AND DIGITAL MODULATION TECHNIQUES

Analog video signals can be transmitted from their point of collection (sensor aircraft) to a data recovery point (surface terminal) using either an analog or a digital data link. The basic elements of these two types of data links are shown in Fig. 2-5.

Analog data links are more commonly used at present because of the availability of equipment. However, for high data rate image transmission applications, digital data links have important advantages over analog

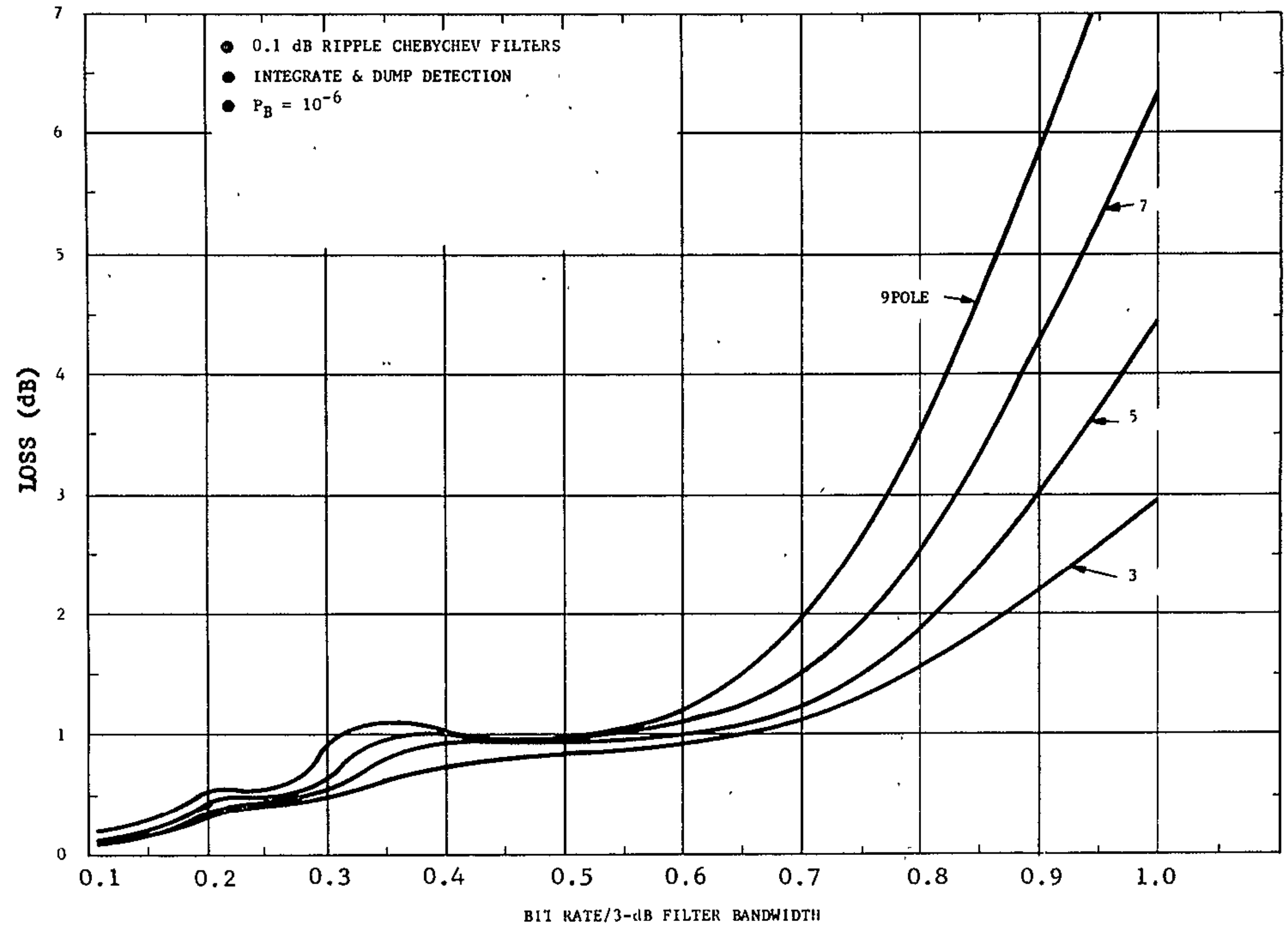
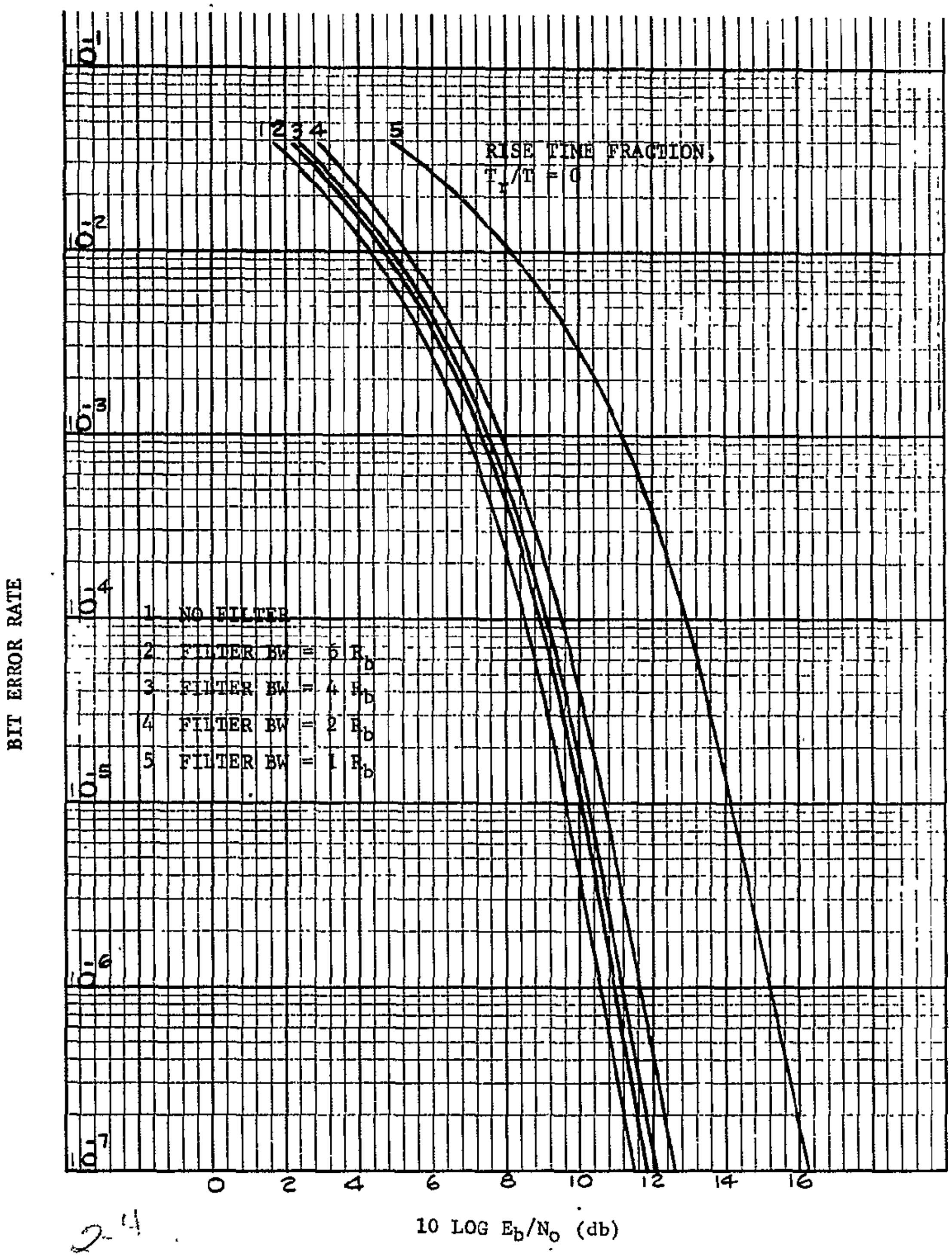


Figure 2-3 Bandwidth Limiting Degradation of Biphase PSK Signals



2-4

Figure 2-4 Effect of Transmission Channel Filter on Biphase PSK Bit Error Rate

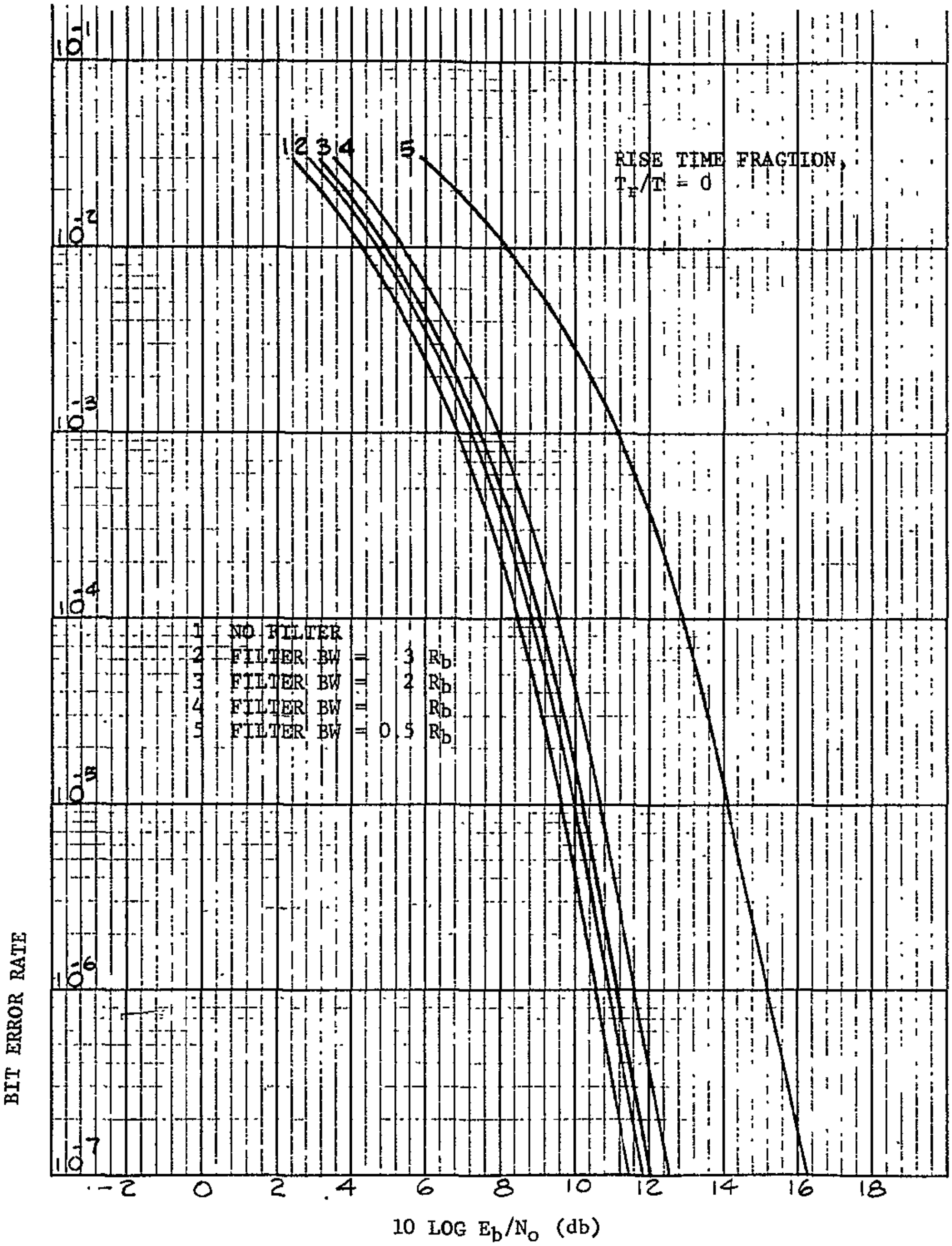


Figure 2-5 Effect of Transmission Channel Filter on Quadriphase PSK (QPSK)

SECTION 3

SENSORS AND SENSOR MTF

3.1 SENSORS AND SENSOR RESOLUTION

Any image-data system will be dependent on one or more sensors to provide image information, and upon the MTF (modulation transfer function) to determine the limiting resolution and therefore the video bandwidth capabilities of the sensors.

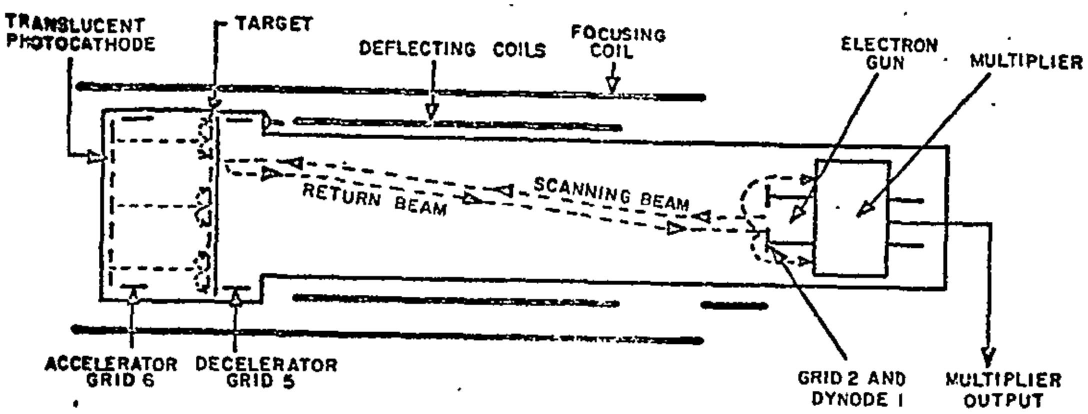
3.2. TYPES OF SENSORS

The various types of sensors, which are candidates for spacecraft use, and which were discussed in Section 3 of Volume 1 will be discussed in the following paragraphs.

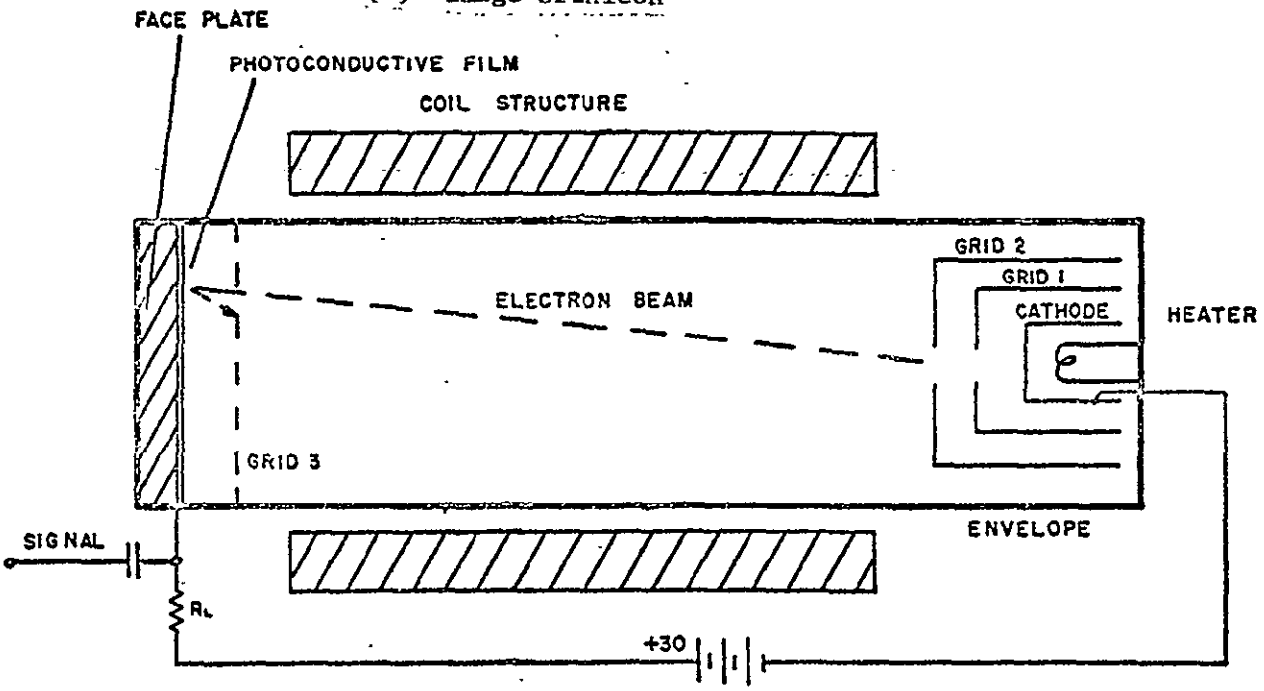
3.2.1 Vidicons

The vidicon (Figure 3-1a) is a scanning image tube operating on the photoconductive principle.

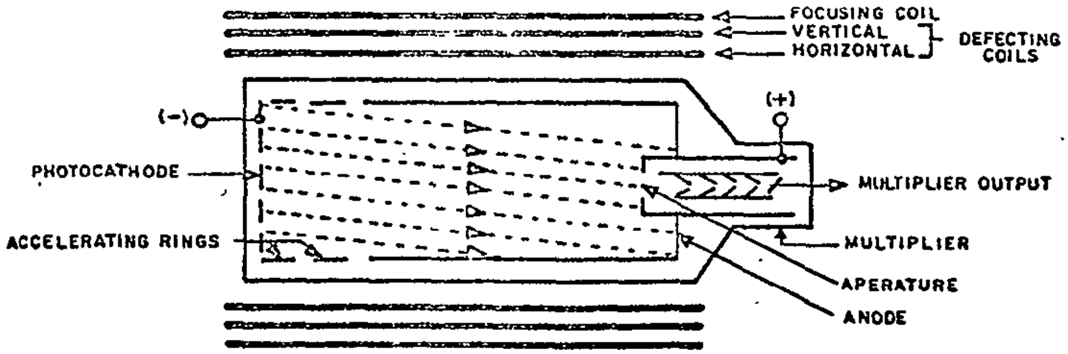
A transparent conducting film is deposited on the inside of a glass face plate which is scanned by an electron beam. A thin layer of the photoconductive material is deposited over the conducting film. The local variations in resistance of the photoconductor corresponds to the image intensity distribution on the face plate. The electron beam sweeps the photoconductor and deposits a charge so that the instantaneous voltage across the photoconductor is equal to the cathode-signal plate voltage. The charge leaks through the photoconductor at a rate proportional to the photo-resistance at that point. The following electron scan deposits charges to replace the leakage. The signal voltage is developed across a load resistor which is connected to the conducting film. This signal voltage is proportional to the deposited charge.


TRANSLUCENT PHOTOCATHODE
TARGET
DEFLECTING COILS
FOCUSING COIL
ELECTRON GUN
MULTIPLIER
RETURN BEAM
SCANNING BEAM
ACCELERATOR GRID 6
DECELERATOR GRID 5
GRID 2 AND DYNODE 1
MULTIPLIER OUTPUT

(a) Image Orthicon


FACE PLATE
PHOTOCONDUCTIVE FILM
COIL STRUCTURE
GRID 2
GRID 1
CATHODE
HEATER
ELECTRON BEAM
GRID 3
ENVELOPE
SIGNAL
RL
+30

(b) Vidicon Camera Tube


FOCUSING COIL
VERTICAL
HORIZONTAL
DEFECTING COILS
(-)
(+)
PHOTOCATHODE
ACCELERATING RINGS
MULTIPLIER OUTPUT
MULTIPLIER
APERATURE
ANODE

(c) Image Dissector

Figure 3-1

The vidicon acts as a constant current generator for the output resistor. The signal current does not increase linearly with illumination. The signal current is usually less than one micro-ampere so that the principal source of noise is the video amplifier.

For the visible spectrum vidicons the minimum tube face illumination is of the order of 0.1 to 0.01 lumen ft^{-2}. With 400 line resolution of a 0.3 inch diameter face plate, the minimum signal power per element is of the order of 10^{-12} to 10^{-13} watts.

The illumination range is approximately three decades from noise level to saturation. Fast image motions blur the electrical image since the time constant of the photoconductor is usually longer than the conventional frame rate of 1/30 second. The signal stability of the vidicon is better than that of the image orthicon because of the simpler electron tube construction with fewer elements.

Figure 3-4 shows the scanning philosophy of the vidicon; Figure 3-5 shows the wavelengths served by various vidicons using different photoconductive surfaces.

The infrared vidicon - also known as the Iricon is substantially identical to the vidicon with the exception of the use of a lead sulphide target which extends the wavelength range of the imagery beyond 10 microns. This type of tube is particularly useful for space applications which address the earth since more than 60% of the direct radiation from the earth takes place in the 0.6 to 10 radiation band.

3.2.2 The Image Orthicon

The image orthicon is analogous to a vidicon. As shown in Figure 3-3b it uses a photoemissive target rather than a photoconductive target and includes a built-in image intensifier and very low-noise signal preamplifier. Therefore, the image orthicon is a more sensitive device than the vidicon, while the image intensifier orthicon is still more sensitive. The image orthicon pays a penalty in physical size

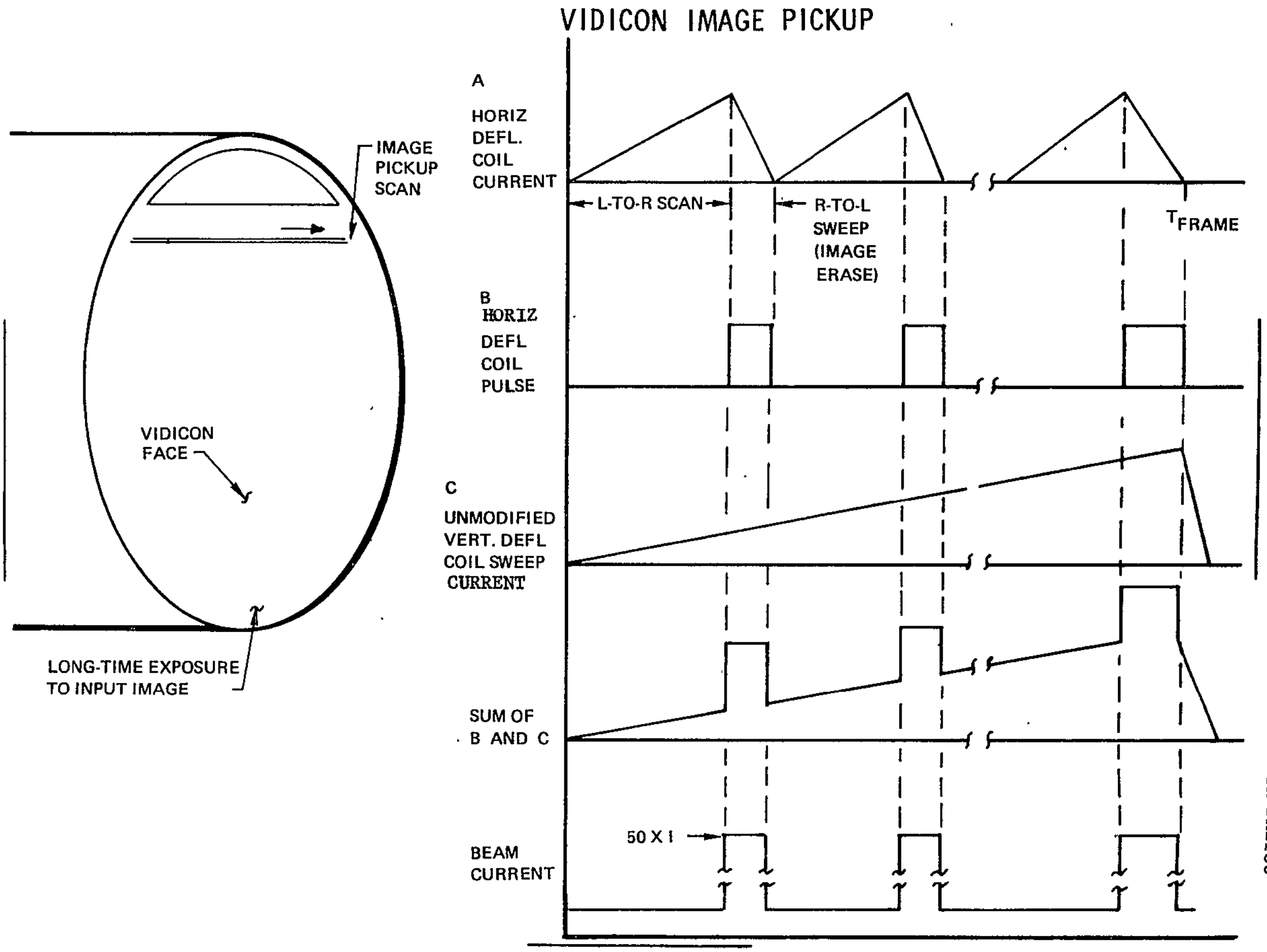

VIDICON IMAGE PICKUP
IMAGE PICKUP SCAN
VIDICON FACE
LONG-TIME EXPOSURE TO INPUT IMAGE
A
HORIZ DEFL. COIL CURRENT
L-TO-R SCAN
R-TO-L SWEEP (IMAGE ERASE)
T_{FRAME}
B
HORIZ DEFL COIL PULSE
C
UNMODIFIED VERT. DEFL COIL SWEEP CURRENT
SUM OF B AND C
50 X I
BEAM CURRENT

Figure 3-2

TR-DA2180

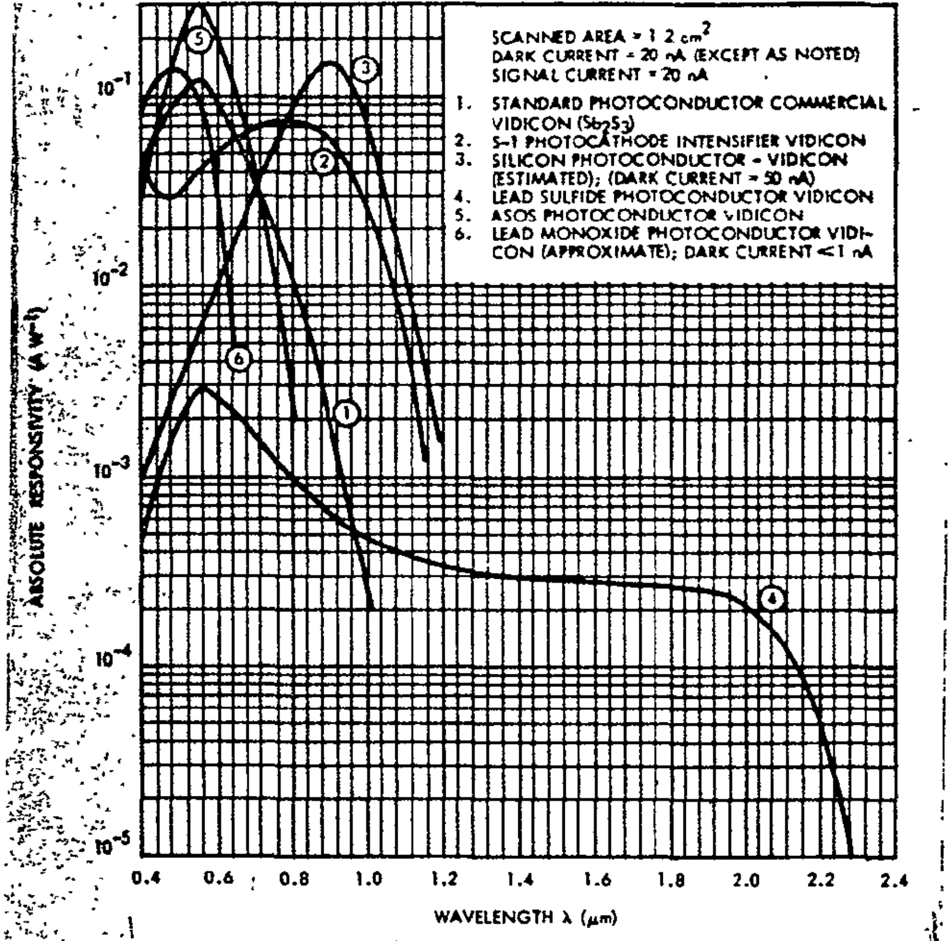


Figure 3-3 Absolute Spectral Responsivities of Several Vidicons

and in the added complexity of the necessary electrical circuitry to maintain its focussing, image intensifying, beam deflecting, and amplifying elements.

Although the operation of the image orthicon is fundamentally different from that of the vidicon, both devices are functionally alike. Each has an image transducing section where a light image is transformed into an electrical image in the form of a voltage pattern. Also, each has a readout section, consisting of a scanning electron beam. These electrons interact with the target which contains the stored electrical image, producing an electronic signal which may be amplified and then displayed.

The image orthicon contains its own amplifier for this purpose (an electron multiplier), while the vidicon must utilize an external preamplifier. The image orthicon also incorporates part of a single-stage image intensifier in its image transducing section while the vidicon may not.

The image orthicon can perform the same functions as the image dissector; the internal process is quite different however. A charge distribution corresponding to the image intensity distribution is accumulated during the frame interval. A four hundred line resolution picture would have 160,000 elements to sample each frame. The integration interval is then 160,000 times as long as the sample interval. The integration procedure should theoretically improve the minimum detectable signal over that of the image dissector by about the same ratio as the frame period to the element sample period; however, the beam current noise and target ohmic leakage decreases the expected improvement in detectivity from the integration procedure.

To compare the image orthicon to other detectors, the minimum detectable irradiance must be estimated. One of the more sensitive image orthicons has a minimum illumination level corresponding to 1×10^{-5} lumens/ft^2 incident on the S-10 photocathode. The S-10 conversion of lumens to watts at the spectral peak is 400 lumens = 1 watt at 0.45 micron.

The equivalent radiance is then approximately 2.7×10^{-11} watts cm^{-2} on the photocathode. The typical thermonic emission from the S-10 photocathode is 10^{-13} to 10^{-14} amperes cm^{-2}, equivalent to a radiance level of 6×10^{-12} to 6×10^{-13} watts cm^{-2}. As mentioned above, the image orthicon is limited by the beam current and factors other than the photocathode thermonic emission. At 400 line resolution each element has an area of 6.2×10^{-5} cm^{2}. The irradiance from a point source when focused by an optical system would have to produce a power level of 1.7×10^{-15} watts per element to match the minimum radiance level of the sample image orthicon.

3.2.2.1 Intensifier Image Orthicon

As a means of providing more gain prior to the target, an additional image section is added to the image orthicon in the intensifier image orthicon. Electrons from the photocathode of the intensifier image section are accelerated onto the first screen. The screen consists of glass coated with fluorescent material on one side and photocathode on the other side. The spectral response of the photocathode matches the response of the fluorescent material. The accelerated electrons hitting the phosphor produce fluorescence which in turn starts photoemission from the photocathode. Emitted electrons from this photocathode are accelerated towards the target and stored there. The scanning and multiplier section of this tube are similar to the one of the image orthicon.

The additional gain obtained from the intensifier section is enough to raise the stored signal level to a point where the fundamental noise from the primary photocathode can exceed the beam noise. Although such a video signal obtained at very low light levels is inherently noisy and limited in resolution, the intensifier image orthicon approaches the performance of the ideal tube.

3.2.2.2 Isocon

One of the serious disadvantages in the use of the Image Orthicon is the low signal to noise ratio of the tube, especially at low light levels. The low ratio is primarily due to the poor percentage of modulation (35%) of the return beam.

In the Isocon tube, which is similar to the Image Orthicon, the diffusely scattered electrons produced by the target instead of specularly reflected electrons dispersed by bright areas are collected by the multiplier section. This type of technique can give modulation that approaches 100%.

3.2.2.3 The Image Dissector

The image dissector is a photomultiplier with a scanning system for deflecting the electrons emitted from the photosurface. The deflection system systematically scans the electron image of the light pattern on the photocathode past a mechanical aperture, and transforms a two-dimensional light pattern into a time-dependent electronic signal.

The scene is imaged on the photocathode. The photo-emissive cathode emits electrons whose local density varies directly with the number and energy of the incident quanta. The aperture intercepts only a small section of the electron image. The multiplier structure then amplifies the signal by secondary emission. The vertical and horizontal deflection coils bend the focus field and scan the electron image over the aperture (Figure 3-3c).

Some image dissectors are capable of resolving more than 1,000 television lines per frame corresponding to more than one million elements sampled each frame. However, the electron flow during the sampling interval becomes very small and difficult to amplify and detect. Under the above conditions, an electron flow of 10^6 electrons per second would yield only one electron per sampling interval, corresponding to approximately 3×10^{-12} watts incident per element of the photocathode.

3.2.3 Solid State Sensor Imagery

Sensitive solid state discrete-point sensors now exist for use in the visible, near infrared and long wavelength infrared spectral regions. Photo-conductive diodes, photo-transistors, and photo-emissive diodes (photo-electric cells) are included.

In general, these photo-conductive and photo-emissive cells can be considered as discrete element equivalents of portions of the targets of vidicons and image orthicons; a vidicon can be considered to be an LSI (large scale integration) version of the photo-conductive cell with the scanning electron beam considered as a switching or commutating mechanism. In like fashion, an image orthicon can be considered as an LSI version of the photo-emissive cell using a scanning-electron beam commutator.

Actually, the use of discrete cells of either type for imagery information poses significant constraints on the design of imagery apparatus which requires that a scanning function be included to allow photo-interrogation of all sections of a surface being photographed. Either of two types of apparatus are then considered:

• Self-Scanned Mosaic:	This represents an array of elements which are electrically scanned to produce an output signal comprising an image representation of a viewed surface.
• Spin Camera System:	The spin camera uses the spinning motion of a satellite to direct narrow angle scanning-line views of a surface to one or more point sensors. The overall scanning which changes the scanning-line view from scan to scan to produce a raster is accomplished by optical means under electronic command.

The basic elements of these two systems will be discussed in the next paragraphs.

3.2.3.1 Classes of Radiation Detectors

Radiation detectors can be divided into two classes, thermal detectors and quantum detectors. The thermal detector operates by the measurement of some characteristic that changes when heat is absorbed or lost. The simplest example would be the common mercury thermometer. In the quantum detectors, the absorbed quantum directly changes the electrical characteristics of the detector. This may be the emission of an electron from the surface of a photo-emitter or photo-cathode or the production of a free electron or hole in a semi-conductor crystal.

3.2.3.2 Photoconductors

The photoconductive detectors are made from the semi-conductor elements and compounds and are of the same technology as in vidicon targets. The detectors may be crystalline solids or amorphous solids. A photon can release an electron from the valence band and raise it to the conduction band. The migration of the electron and the "hole" (missing electron) through the material under the influence of an electric field constitutes an electric current. A small fraction of impurities in the crystal structure can aid the sensitivity of the detector. The impurity or donor element has a different valence. The crystal structure then has either a hole or electron available at the impurity site that is loosely bound to the atom. The energy of a photon can ionize the site and produce a migrating hole or electron.

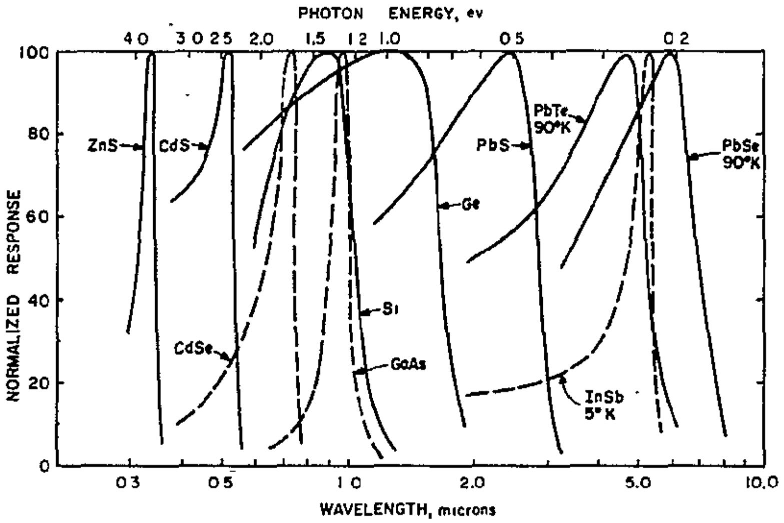
Although photoconductivity was first discovered in 1873 by Willoughby Smith while investigating selenium as a resistor in underwater cables, actual progress in understanding and material development has awaited the recent leap forward in all solid-state physics which was activated by the development of the transistor. Photoconducting cells were made only of selenium, copper oxide, or thallous sulfide up to 1940; today, commercial cells use none of these, but are based almost exclusively on germanium, silicon, cadmium sulfide, cadmium selenide, lead sulfide, lead selenide, or lead telluride, all resulting from recent developments. It was not until the end of World War II that detection of infrared radiation by the lead-sulfide-type photoconductors provided superior to former thermoelectric or bolometric methods. Figure 3-6 shows the intrinsic photoconductivity response as a function of incident wavelength for a number of typical photoconductors.

In many cases, photoconducting cells today perform the same functions previously performed by photoemissive cells, with the advantages of decreased size and cost, and increased ease of operation. The solid-state photoconductor cell bears the same relationship to the photo-emissive vacuum phototube as the solid-state diode bears to the thermionic-emissive vacuum diode.

3.2.3.3 The Photo Transistor

The photo transistor, by virtue of its current gain, offers the possibility for higher sensitivity than does a simple photodiode. Its reverse-biased collector junction acts as the photosensor and the emitter junction provides the switching function. The collector junction must have a much larger area than the emitter to provide adequate storage capacitance with a small switch capacitance.

To a first approximation, the total charge that flows through the emitter-collector circuit during the sampling pulse is proportional to the accumulated charge on the base multiplied by the current gain of the transistor. This relationship, however, does not hold at all light levels and nonlinear transfer characteristics and sluggish transient response may limit the application of the photo transistor for low-light performance.



—*Intrinsic photoconductivity response of typical photoconductors vs. wavelength, where "intrinsic" means response of the material itself without incorporated impurities. Impurities extend response to longer wavelength than the intrinsic response. Intrinsic responses extend from the ultraviolet (<0.4μ) through the visible (0.4 to 0.7μ) into the infrared (>0.7μ).*

Figure 3-4

3.2.3.4 Photo-emissive Cells

The photo-electric cell or photo-emissive cell consists of a photo-cathode in a vacuum tube whose output is passed through one or more secondary emitter amplifiers - called dynodes. Such tubes as photo-emissive tubes can address any part of the visible, UV or IR spectrum depending on the photo-cathode material used, and the use of many dynodes (up to 12) makes such devices among the most sensitive and quietest (from the noise standpoint) sensor devices available.

3.2.3.5 Mosaics

Mosaics of large numbers of elements are very difficult to fabricate. The main difficulty arises from making electrical contact by wire to each individual cell. The fabrication of the cells themselves, by contrast, is quite simple. These are deposited as thin films on a suitable substrate. The present state-of-the-art in cell size is about 0.0015 inch, with a separation from its nearest neighbors of about 0.001 inch. Variation in sensitivity of ± 10 percent are typical from cell to cell. The largest mosaic fabricated to date is in the neighborhood of 125,000 elements.

The problem of fabrication is compounded by the problem of readout of the mosaic. The use of an individual preamplifier for each cell with sequential sampling of the outputs is attractive as it allows the use of a narrow bandwidth for each individual detector. Generally, the output of the cells or individual preamplifiers is multiplexed sequentially, into one, or several preamplifiers. This technique appears to be developed to the point where readout of 10^6 elements in less than a second poses no problem. The cell wiring problem in large mosaics is so formidable that some manufacturers are going to electron beam readout techniques, thus eliminating the need for individual cells, and turning the mosaic into a continuous detector surface. Detectors are usually photo-voltaic for this mode of readout and can incorporate a limited amount of image storage capability.

3.3 MODULATION TRANSFER FUNCTION

Optical system performance is quantitatively expressed in terms of a modulation transfer function. The concepts and mathematical formulism used in evaluating this function, or optical system index of performance, have their origins in linear filter analysis of communications theory. Mathematical techniques and theories of the transform calculus are utilized, especially those of Fourier integral theory. Where communications theory makes use of the Laplace single-ended transform to relate functions of time and frequency, linear optical systems theory applies the double-ended Fourier transform to relate functions of space and frequency. Approximate optical analogs of communications theory concepts are presented below in Table 3-1.

TABLE 3-1

Electrical Optical Analogs

	Communications Theory Terms	Optical Analog
(a)	Voltage as a function of time	Brightness as a function of linear dimension
(b)	Rotational Frequency in Hz	Spatial frequency in cycles per MM
(c)	Unit Impulse	Point Source
(d)	Impulse Response	Point Spread Function

3.3.1 Modulation Transfer Function and Limiting Resolution

The Modulation Transfer Function, or MTF, of a sensor is analagous to the amplitude transfer function of linear electrical networks. It gives the amplitude response of the sensor as a function of spatial frequency. The MTF can be generated experimentally by scanning a test target with a sinusoidally weighted grid of alternating black and white bars normal to the scan line. The magnitude of the voltage at the sensor output would then be plotted as a function of scan position. If the distance between the black and white bars (period of the spatial frequency test input) is large relative to the scanner spot size then maximum amplitude response should appear at the sensor output. As the spacing of the bars becomes closer (increasing spatial frequency) the sensor output response will decrease as shown in Figure 3-5(a). The MTF of a sensor is, therefore, a direct measure of the response bandwidth of the sensor.

The MTF is normally expressed as a normalized (zero to one) response plotted as a function of spatial frequency. The most common units of spatial frequency are line pairs per millimeter (lp/mm). Multiplying the spatial frequency by scan velocity gives the sensor output response in electrical frequency.

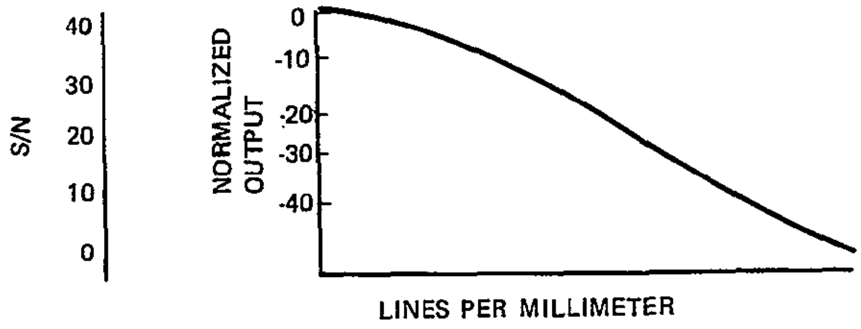
The concept of limiting resolution is illustrated in Figure 3-5. As the lines per millimeter increase, that is, the number of scanning lines increase, under condition of constant illumination, then as the output decreases, the output signal level will approach the noise level which is independent of the line pairs per millimeter and the output of the sensor will become unusable.

Thus a vidicon which has a useful scanning rate of, say, 1700 lines, may require a limiting resolution of 4000 lines (converted from lines per millimeter by multiplying lp/m by the scan distance in millimeters) to assure adequate resolution with proper signal to noise ratio at the light levels being observed.

Consider, for example, the limiting resolution spatial frequency at the center-field resolution of the vidicon. This spatial frequency is

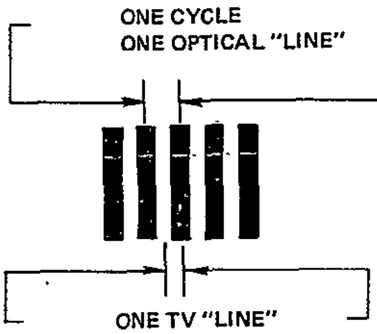
$$f_{\text{limiting}} = N/50.8L.$$

$$= \frac{1500}{50.8 \times 1.12} = 26.4 \text{ LP/MM}$$



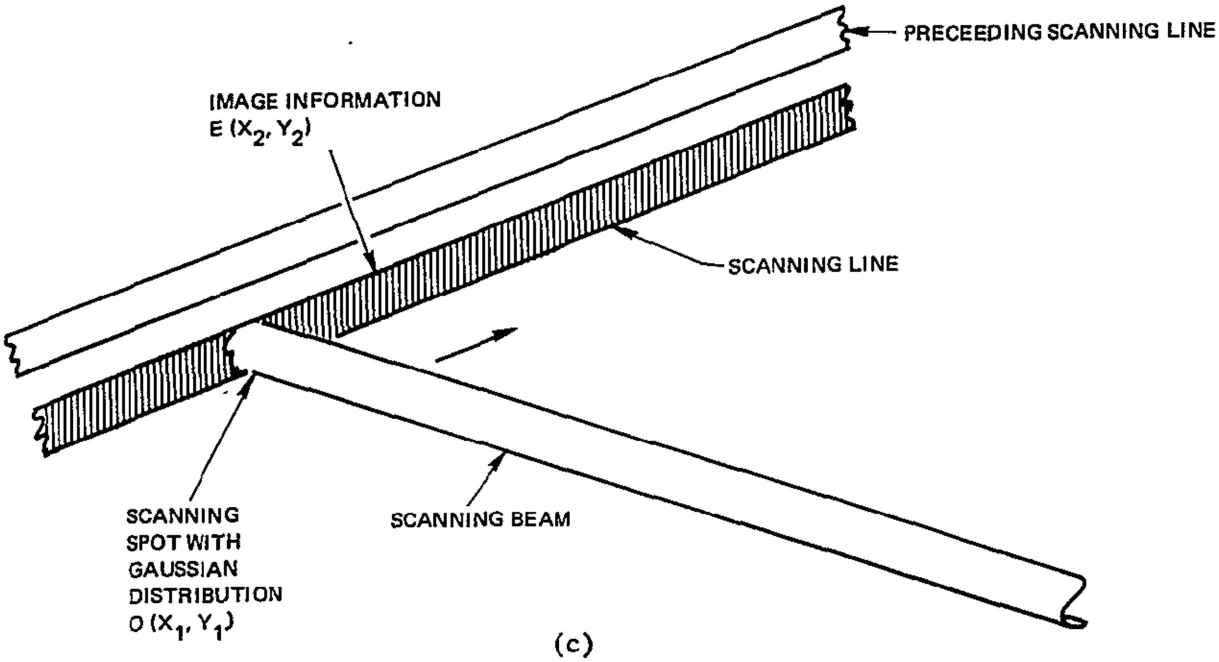
(a)

Signal to Noise Ratio versus Resolution


ONE CYCLE
ONE OPTICAL "LINE"
ONE TV "LINE"

(b)

RELATIONSHIP OF OPTICAL AND TELEVISION PARLANCE.


PRECEEDING SCANNING LINE
IMAGE INFORMATION
E (X2, Y2)
SCANNING LINE
SCANNING SPOT WITH GAUSSIAN DISTRIBUTION
O (X1, Y1)
SCANNING BEAM

(c)

Typical Scanning Beam/Scan Line Geometry

Figure 3-5 The Geometry of a Scanning Beam Traversing a Scanning Line

From typical 5140 vidicon characteristics, the tube response at 26.4 LP/MM is 0.52.

3.3.2 Resolution Notation

There may be a difference of a factor of two between two individuals' interpretation of a resolution. Figure 3-5 illustrates the distinction. The source of this misinterpretation is avoided if the term "line-pairs per millimeter" or, preferably, "cycles per millimeter" is used.

There is a divergence of definition of the term "lines per millimeter" between those concerned principally with optics and those concerned principally with electro-optical (for example, television).

The optical definition developed from consideration of resolution in terms of two close point images, then of the resolution of two close lines, and finally of the resolution of equally spaced lines; therefore, to the optical scientist, a "line" is the spacing between two black (or white) lines, separated by a space of equal width. In television, the principal concern is the scan line on the camera or kinescope tube, and resolving two black (or white) lines separated by a space of equal width requires two scan lines on the tube. Thus, to imagery applications a "line" is the distance between the black line and the intermediate space.

3.3.3 Convolution Analysis Basis of MTF

As shown in Figure 3-5(b), E represents the intensity distribution in the image plane for a point source anywhere in the object plane, and is a result of applying the scalar wave (Kirchhoff) approach to Fraunhofer diffraction.

Any object imaged by an optical system such as a scanning spot will have a distribution of intensity which can be expressed by some function $0(x_1y_1)$ where x_1 and y_2 are the coordinates in the object plane. The image will have an intensity distribution given by multiplying the value of $0(x_1, y_1)$ for each point in the scene by $E(x_1, y_1)$ and summing the result.

The image intensity function can be expressed as:

$$I(x', y') = \iint_E 0\ (x_1, y_1)\ E\ (x_1-x', y_1-y')\ dx_1\ dy_1$$

This equation represents an intensity function derived by convolving $E(x_1, y_1)$ with $0(x_1, y_1)$, and can be recognized as a convolution or "faultung" integral.

The convolution of E with an object intensity distribution function 0 yields an image intensity function I. Symbolically,[1] then,

$$I = E * 0$$

E describes how an aperture (or optical system) responds at the image (or system output) to a point source. It describes how an aperture responds to an arbitrary object intensity distribution in terms of its response to a point source, The parallel with linear systems analysis is strikingly evident. In linear system analysis, the system response function is the response to a unit impulse (Dirac delta) excitation. The response to some arbitrary excitation in the convolution of the system impulse response (also called "Green's" function in mathematical physics) and the given excitation function (Borel's theorem).

Let $F\left|\ \right|$ represent the Fourier transform of a function in the bracket.

---

[1] H. P. Hau, "Outline of Fourier Analysis", Unitech Outlines, 1967.

If $\quad F \left[O\,(x_1,\ y_1) \right] = O\,(u,\ v)$

and $\quad F \left[E\,(x_1,\ y_1) \right] = (u,\ v)$

then, the following convolution theorem holds:

$$F \left[O(x_1,\ y_1) * E(x_1,\ y_1) \right] = O(u,\ v)\ \ E(u,\ v)$$

i.e., the Fourier transform of the convolution of two functions equals the simple product of the Fourier Transform of the functions. In the previous section $E\ (x_1,\ y_1)$ was defined as the optical transfer function. If only the modulus, $\left| E\ (u\ v) \right|$, is considered, $\left| E\ (u,\ v) \right|$ is called the "modulation transfer function." The Fourier transform of functions of space are transformed into functions of frequency. As an example, the Fourier transformed object intensity function, $F \left[O(x_1,\ y_1) \right]$ is called the spatial intensity spectrum and represents the object as being composed of a set of sinusoidally varying intensity patterns of different magnitudes and positions, superimposed upon each other. Analogously, the Fourier transform of the image intensity function, $F \left[E(x_1,\ y_1) \right]$, is its spatial intensity spectrum expressed in lines per millimeter.

The Fourier transform techniques permit a simple derivation of the transfer function. Recalling that $U\ (a,\ b) = F \left[T\ (x,\ y) \right]$, i.e., the complex amplitude function is the Fourier transform of the aperture function, and that $E\ (x_1,\ y_1) = UU^*$, i.e., the point spread function is the square of the modulus of the amplitude function, then:

$$E\ (x_1\ y_1) = F \left[T(x_1 y) \right] \quad F^* \left[T(x_1 y) \right]$$

or

$$E\ (x_1, y_1) = T(x_1 y) \ * \ T^*(x_1 y)$$

by applying the Fourier convolution theorem. This means that the optical transfer function is the convolution of the aperture function with its complex conjugate, i.e., the transfer function is the auto correlation function of the aperture (pupil) function.

3.3.4 The Concept of Spatial Frequency

For the sake of simplicity and without loss of generality, a normalized, one-dimensional Fourier transform of a point spread function E(x) can be written as:

$$H(\omega) = \frac{\int_{-\infty}^{\infty} E(x)\, e^{-i\omega x} dx}{\int_{-\infty}^{\infty} E(x) dx}$$

where ω now has the dimension of radian per unit distance. Since $\omega = 2\pi\lambda$, then the MTF for the point spread function is given by

$$H(\lambda) = \left| H \frac{\omega}{2\pi} \right|$$

Now, in an evaluation of above, the limits on the integral are generally finite, explicitly determined by the spread function boundary condition. Suppose there exists an idealized symmetrical point spread function of unity amplitude between $x = -d/2$ and $x = +d/2$, zero elsewhere. Then,

$$H(\lambda) = \frac{2\int_{0}^{d/2} (1)\, e^{-i2\pi\lambda} dx}{2\int_{0}^{d/2} (1)\, dx}$$

$$= \frac{\operatorname{Sin} \pi\lambda d}{\pi\lambda d}$$

and $\left| H(\lambda) \right| = \left| \frac{\operatorname{Sin} \pi\lambda d}{\pi\lambda d} \right| = \text{MTF}$

Note that the zeros of $H(\lambda)$ occurs for $\lambda = n/d$, i.e., for integral multiple of 1/d.

The limiting spatial frequency is given by

$$k_{\text{Limit}} = \frac{1}{d}$$

It can be shown that an approximate relation for the resolving power of a lens is given by:

$$d/f = \frac{1}{D}$$

Where d = minimum separation resolvable in the image
f = focal length
D = aperture
λ = wavelength

Combining the two above expressions yields

$$k_{\text{Limit}} = \frac{1}{d} = \frac{1}{\frac{f\lambda}{D}} = \frac{D}{f\lambda} = \frac{1}{F\lambda}$$

where F is lens F-number.

This last relationship implies that the faster a lens (smaller the f-number) the higher the limiting spatial frequency, or the better the resolution, which is generally true.

3.3.5 Use of the Modulation Transfer Function for a System

For the usually-encountered case of non-coherent illumination, the MTF described earlier is the item of interest. In linear system analysis if there is a chain of n successive independent operations performed on some excitation, each operation represented by its representative transfer function, then the overall system transfer function can be represented by the product of all "n" individual transfer functions.

In optical systems analysis using the transfer function approach, the overall MTF for the final image can be given by

$$Hi(\lambda) = H_o(\lambda) \prod_{n=1}^{n} T_n(\lambda)$$

MTF's have been computed for

- Film
- Image tubes
- Human eye
- Atmospheric turbulance
- Image motion
- Atmospheric transmission
- Cathode Ray Tubes

The object or target itself can be treated as if it has its own MTF. Transfer functions are usually presented in log-log coordinates, normalized.

The techniques of linear system analysis can be applied to derive an overall optical system MTF or "System Index of Performance." Using the transfer function approach allows the overall MTF for the final image to be calculated by taking the product of the MTF's of the individual system components. Figure 3.10 which depicts the linear system model applicable to the total image system is being discussed in this report.

The various MTF's developed for the imagery system of Figure 3-6 are discussed as follows:

3.3.5.1 MTF and Scene Statistics

A very critical factor in determining optimum data encoding is the MTF (modulation transfer function) of the system coupled with valid descriptions of scene statistics. This information will determine the electrical spectrum of the output signal. The most valid scene statistics are the probability distributions of radiance difference between adjacent regions as a function of the size of the regions. This data is more appropriate than the scene power spectrum since the latter represents an average, and the processor must be able to handle data at the limits of the distribution curve in order to make optimum use of the information gathering capabilities of the sensor.

In the visible and near infrared it is rather easy to construct situations where the earth radiance will change over its full potential range in a distance of 100 feet.

In the intermediate infrared, where the dynamic range may represent temperature extremes, in the range of 250 to 350°K, a radiance difference caused by a 100°K temperature gradient in 100 feet is not very likely. Maximum changes of 25°K would be more representative of the tail of the distribution curve.

In general, it is observed that the output signal characteristics are primarily determined by the sensor MTF, as shown in Figure 3-7. For a rectangular field stop in the image plane, this is simply the sin x/x function which has its first zero at $\frac{1}{D}$ cycles/ft on the ground, where D is the subtense of the stop or detector element.

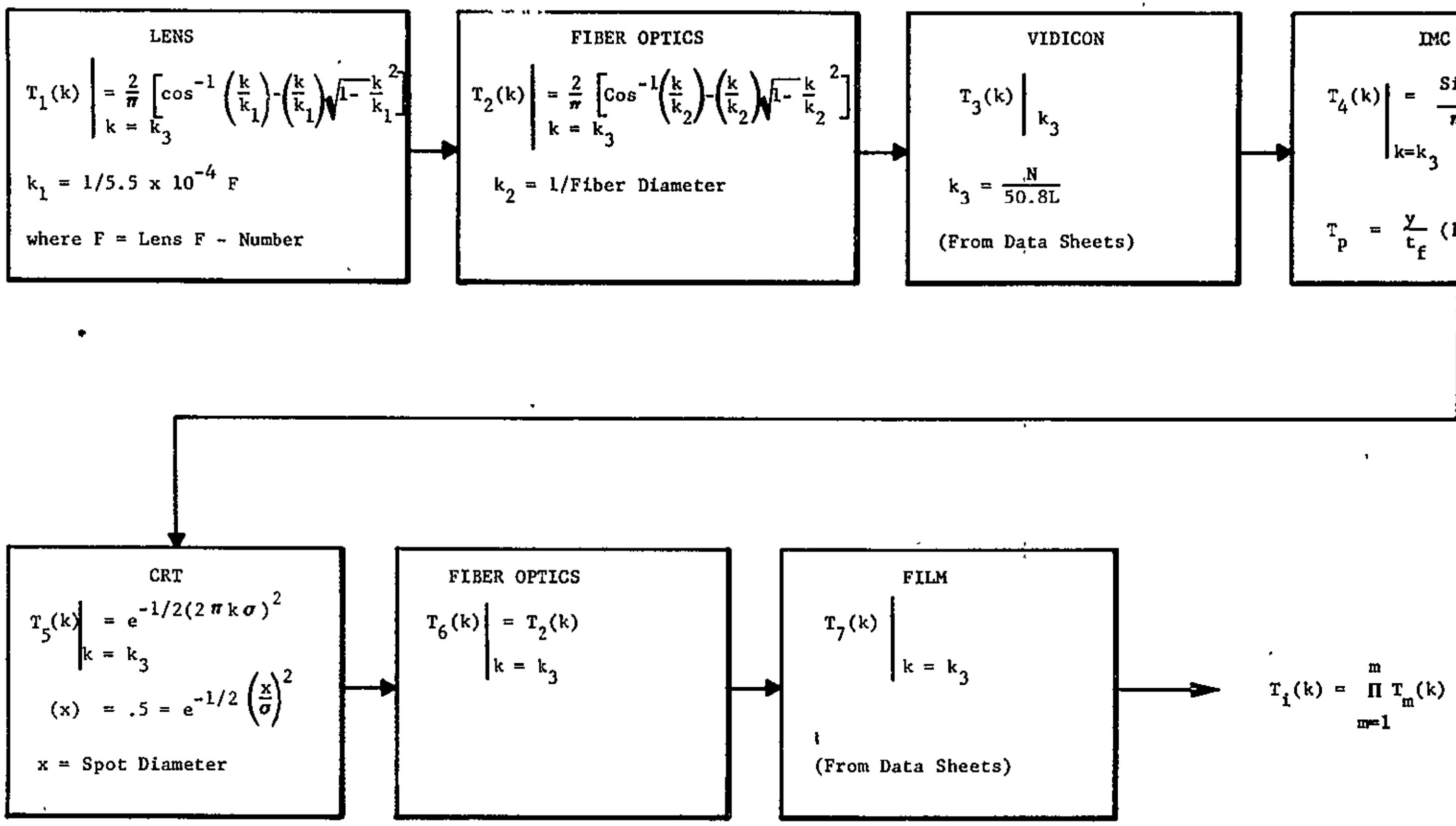

LENS
$T_1(k)\Big|_{k=k_3} = \frac{2}{\pi}\left[\cos^{-1}\left(\frac{k}{k_1}\right)-\left(\frac{k}{k_1}\right)\sqrt{1-\frac{k}{k_1}^{2}}\right]$
$k_1 = 1/5.5 \times 10^{-4} F$
where F = Lens F - Number
FIBER OPTICS
$T_2(k)\Big|_{k=k_3} = \frac{2}{\pi}\left[\mathrm{Cos}^{-1}\left(\frac{k}{k_2}\right)-\left(\frac{k}{k_2}\right)\sqrt{1-\frac{k}{k_2}^{2}}\right]$
k_2 = 1/Fiber Diameter
VIDICON
$T_3(k)\Big|_{k_3}$
$k_3 = \frac{N}{50.8L}$
(From Data Sheets)
IMC
$T_4(k)\Big|_{k=k_3} = \frac{\mathrm{Sin}\,\pi k T_p}{\pi k T_p}$
$T_p = \frac{y}{t_f}(1 - \mathrm{IME})$
CRT
$T_5(k)\Big|_{k=k_3} = e^{-1/2(2\pi k\sigma)^2}$
$(x) = .5 = e^{-1/2\left(\frac{x}{\sigma}\right)^2}$
x = Spot Diameter
FIBER OPTICS
$T_6(k)\Big|_{k=k_3} = T_2(k)$
FILM
$T_7(k)\Big|_{k=k_3}$
(From Data Sheets)
$T_i(k) = \prod_{m=1}^{m} T_m(k)$

Figure 3-6 Optical Linear System Model

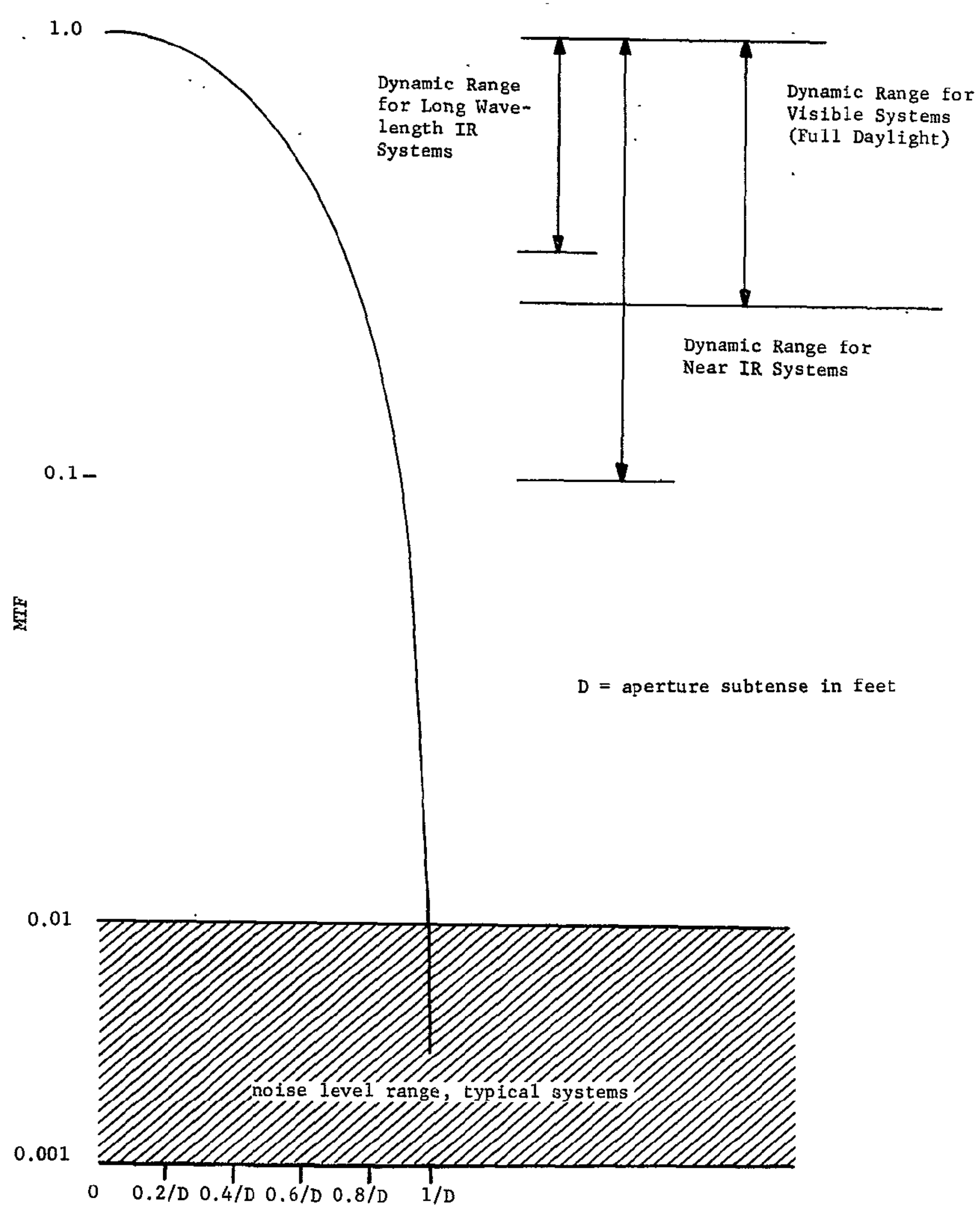


Figure 3-7 Spatial Frequency in Cycles/Foot Versus MTF

A circular stop would of course be represented by a Bessel function with a similar shape. When these aperture functions are convolved with diffraction and image motion effects, further attenuation occurs. If it is assumed that there will be at least 4 samples/dwell time, and that the resolution is within the diffraction limit, then the sin x/x function is still representative (although somewhat optimistically) of the system MTF.

3.3.5.2 Lens MTF

$T_1(k)$ is derived by taking the convolution of a circular aperture with itself. The lens cutoff frequency is the reciprocal of 5500 Å times the lens F-number. Lens F-number, or relative aperture, is a trade-off between a physically realistic aperture size and the vidicon tube minimum exposure requirement which is the image plane illumination multiplied by the exposure time.

3.3.5.3 Fiber Optics MTF

The MTF's for a 7-micron fiber optics array are computed by considering a 7-micron fiber to be the equivalent of a circular aperture (excluding transmission losses) having a uniform intensity distribution. The MTF for this aperture is the convolution of the aperture with itself, with a limiting (cut-off) frequency of 1/.007 MM or 143 LP/MM. The MTF in terms of spatial frequencies is found to be:

$$T(k) = \frac{2}{\pi} \left[\cos^{-1}\left(\frac{k}{k_o}\right) - \left(\frac{k}{k_o}\right) \sqrt{1 - \left(\frac{k}{k_o}\right)^2} \right]$$

where k is a varying spatial frequency, and k_o is the limiting frequency.

Consider two spatial frequencies of interest, namely 26.4 LP/MM and 49.2 LP/MM. Evaluation of the above equations at these k's yields MTF's of .76 at 26.4 LP/MM and .57 at 49.2 LP/MM. A rough plot of T (k) is presented on Figure 3-8. The equation can also be applied to calculate the response of a "perfect" lens. The k_o for the lens is gotten from

$$k_o = \frac{1}{5.5 \times 10^{-4} F} \qquad \text{where F is the lens F-number.}$$

3.3.5.4 Vidicon MTF

A value for T (k) is most readily acquired from the data sheets for the selected vidicon. The MTF for an image tube can be calculated by taking the Fourier transform of the spread function of the tube. The spread function generally depends on the size of the scanning beam and on the construction and characteristics of the sensitive surface. The scanning spot is assumed to have a Gaussian shape so that the spread function is of the form:

$$S(x) = e^{-1/2 \left[\frac{x}{\sigma}\right]^2}$$

The Fourier transform is thus

$$G(k) = e^{-1/2(x/\sigma)^2} \cdot e^{-ikx} dx$$

Letting U = x + ik

$$G(k) = e^{-1/2 (2\pi k \sigma)^2} \int_{-\infty}^{\infty} e^{-1/2 \left(\frac{u}{\sigma}\right)^2} du$$

$$= e^{-1/2 (2\pi k \sigma)^2} \int_{-\infty}^{\infty} S(u)\, du$$

But $$T(k) = \frac{G(k)}{\int_{-\infty}^{\infty} S(u)\, du}$$

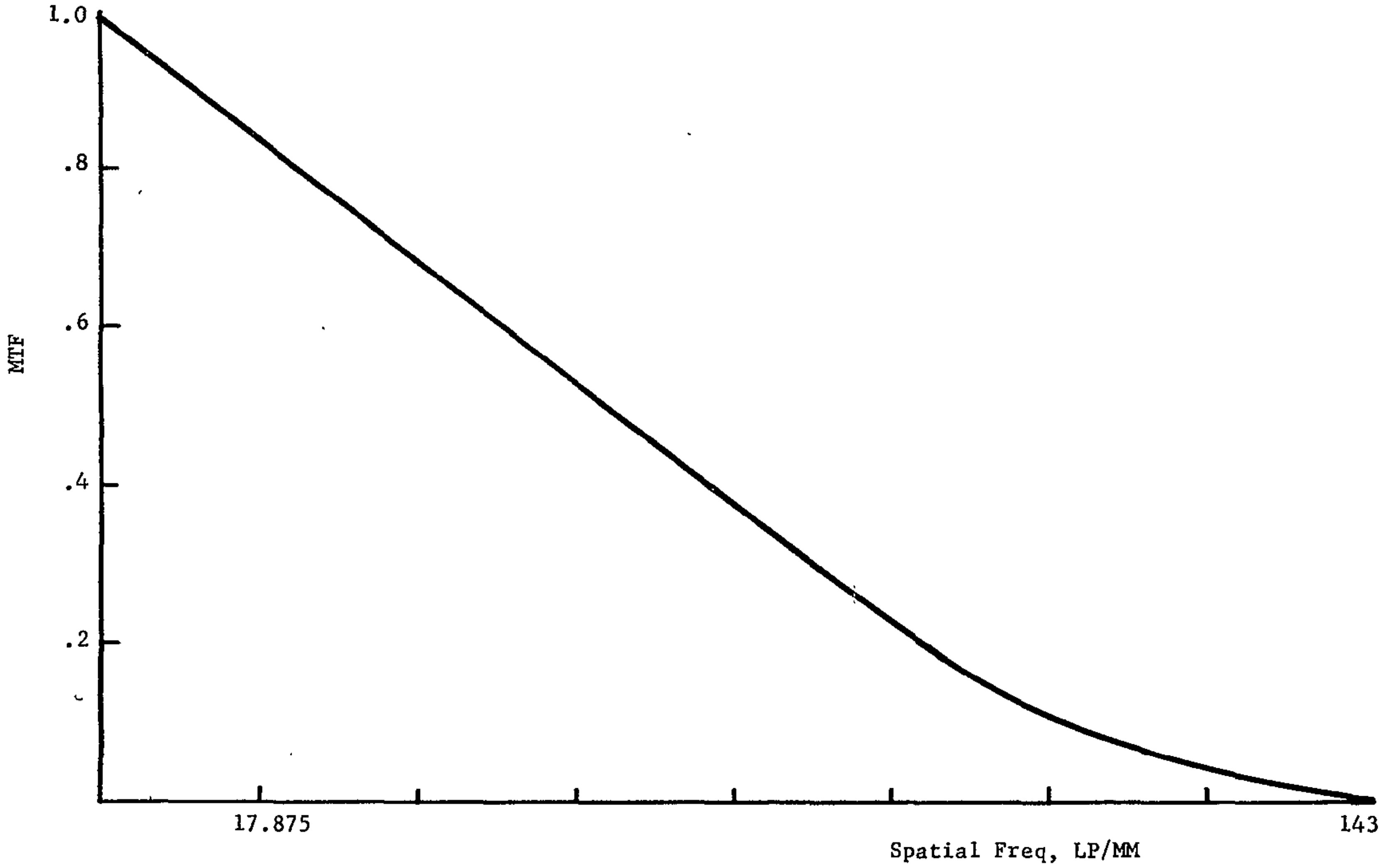


Figure 3-8 Plot of MTF, 7 Micron Fiber Optics

which is the MTF for a point spread function.

Thus, $T(k) = e^{-2 (\pi k \sigma)^2}$

This function is plotted in Figure 3-9 along with the MTF for the diffraction limited spot.

Schade has made an extensive study of point image definition and has evaluated the effect of various image intensity distributions on image definition. He has pointed out that a succession of aperturing processes seems always to cause the final image intensity to tend toward a Gaussian distribution. He has also pointed out that for an imaging process involving a series of apertures each having a Gaussian aperture transmittance (and thus Gaussian sine wave response functions) the overall system sine wave response function is Gaussian. Finally, his numerical calculations with different apertures have shown that round apertures with $\cos^2$ and Gaussian transmittance functions have essentially the same aperture response. This indicates that any symmetric, round aperture transmittance function having a smoothly varying transmittance over the central area of the aperture but falling off exponentially from the aperture center, will have essentially the same response function.

Table 3-2 presents a useful set of computations for a vidicon or image orthical system with a Gaussian-shaped MTF. The unit length has been arbitrarily chosen to contain 100 cycles of the spatial frequency which has a 1% sine wave response. Other columns give the sine- and square-wave response. N_e is the equivalent rectangular passband.

3.3.5.5 Step Response of a Scanning Sensor to a Knife Edge Target

The transient response of scanning sensors to a knife edge target is of interest since this defines the maximum possible slew rate of the sensor output. This information is of particular importance when the sensor output is to be processed by a system element that is capable

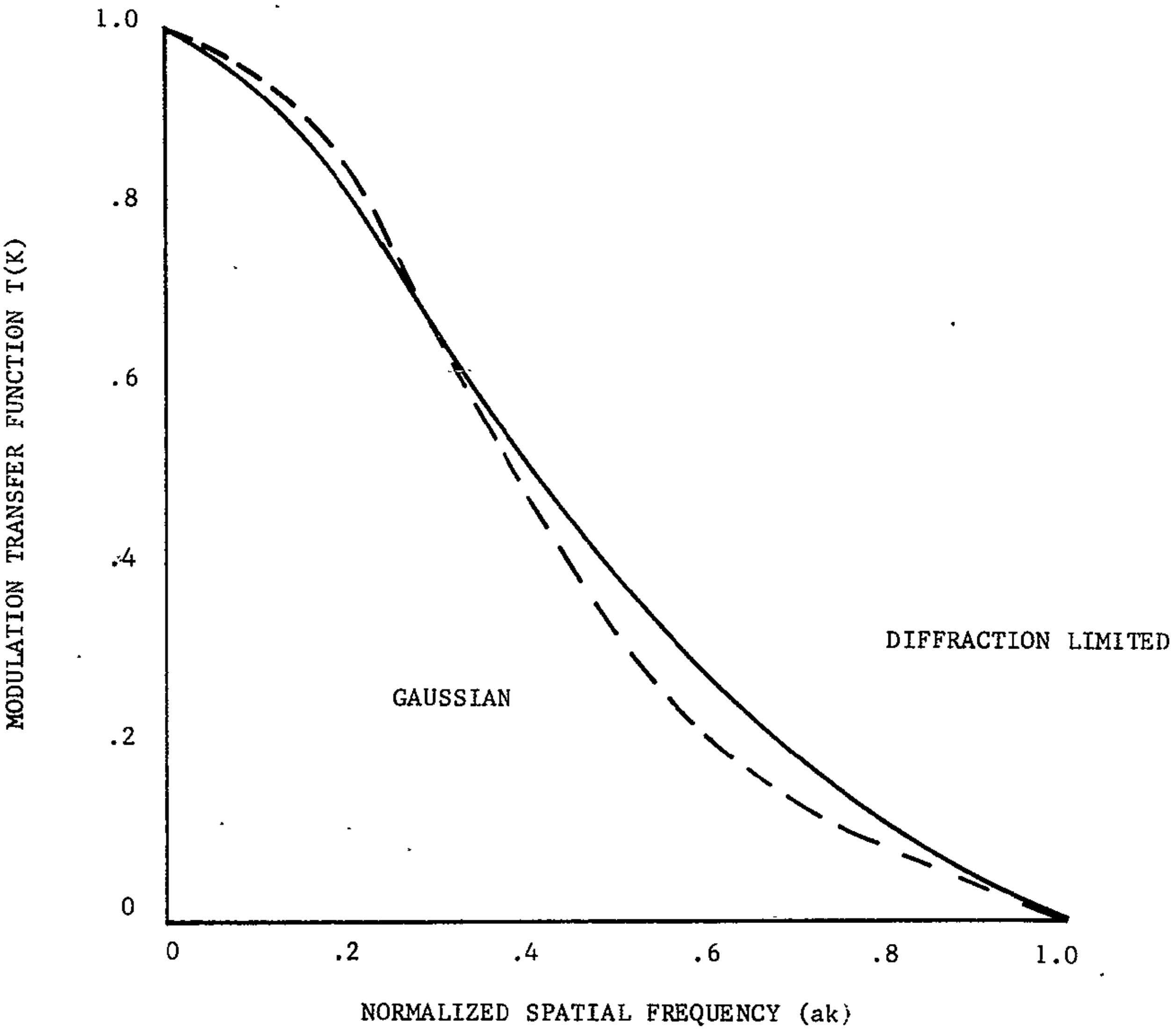


Figure 3-9 Modulation Transfer Function for Laser Scanner on Axis with Diffraction Limited Optics (Symmetrical Aperture) and with a Gaussian Point Spread Function

TABLE 3-2

GAUSSIAN SHAPED MODULATION TRANSFER FUNCTION

NORMALIZED SPATIAL FREQUENCY N $\left(\frac{\text{LINE PAIRS}}{\text{UNIT LENGTH}}\right)$	MTF (SINE WAVE) (RESPONSE) %
0	100
2	99.6
5	98.5
10	95.2
20	83.1
29.1	67.5
30	66.0
40	48.0
50	31.2
60	19.0
70	10.5
80	5.2
80.5	5.0
83.5	3.0
92.4	2.0
100	1.0

$$*(\text{MTF}) = \exp\left(-4.59 \times 10^{-4} N^2\right)$$

of producing slope overloading distortion. All differential PCM and delta modulation encoders are slope limited, when adjusted for optimum noise performance. In general, a delta modulator or differential PCM encoder is optimized and gives minimum quantization error consistent with no slope overloading distortion, when the slew rate of the encoder is adjusted to just equal or exceed the maximum output data slew rate. (This rate should be determined by the overall system modulation transfer function including the MTF of the source if known.)

High frequency circuits, such as very wideband video amplifiers, also can produce slope overloading distortion when the input data slew rate exceeds the capability of the circuit. The slew rate of many amplifiers is limited by the ability of the output stage maximum current to charge its internal shunt capacitance and the external load capacitance. In feedback application, the resultant nonlinear phase shift can cause the system to become unstable.

High voltage wideband amplifiers such as those used for driving wideband (20-50 MHz) optical modulators, which are typical in laser recorders, are particularly sensitive to the maximum data slew rate. Accurate knowledge of maximum data slew rate can result in considerable design simplifications.

The step response to a knife edge may be determined for a known point spread function, s (x, y), by scanning the function with a knife edge.

Consider a Gaussian shaped spot with the two dimensional point spread function:

$$s(x, y) = \frac{1}{2 \pi \sigma^2} \exp \left[- \frac{(x^2 + y^2)}{2 \sigma^2} \right]$$

Let the knife edge be oriented vertically so that it is opaque for $x < o$. Then the response as a function of spot center position x, is:

$$R\ (x_1) = \int_{-x_1}^{\infty} \int_{-\infty}^{\infty} s\ (x,\ y)\ ds\ dy$$

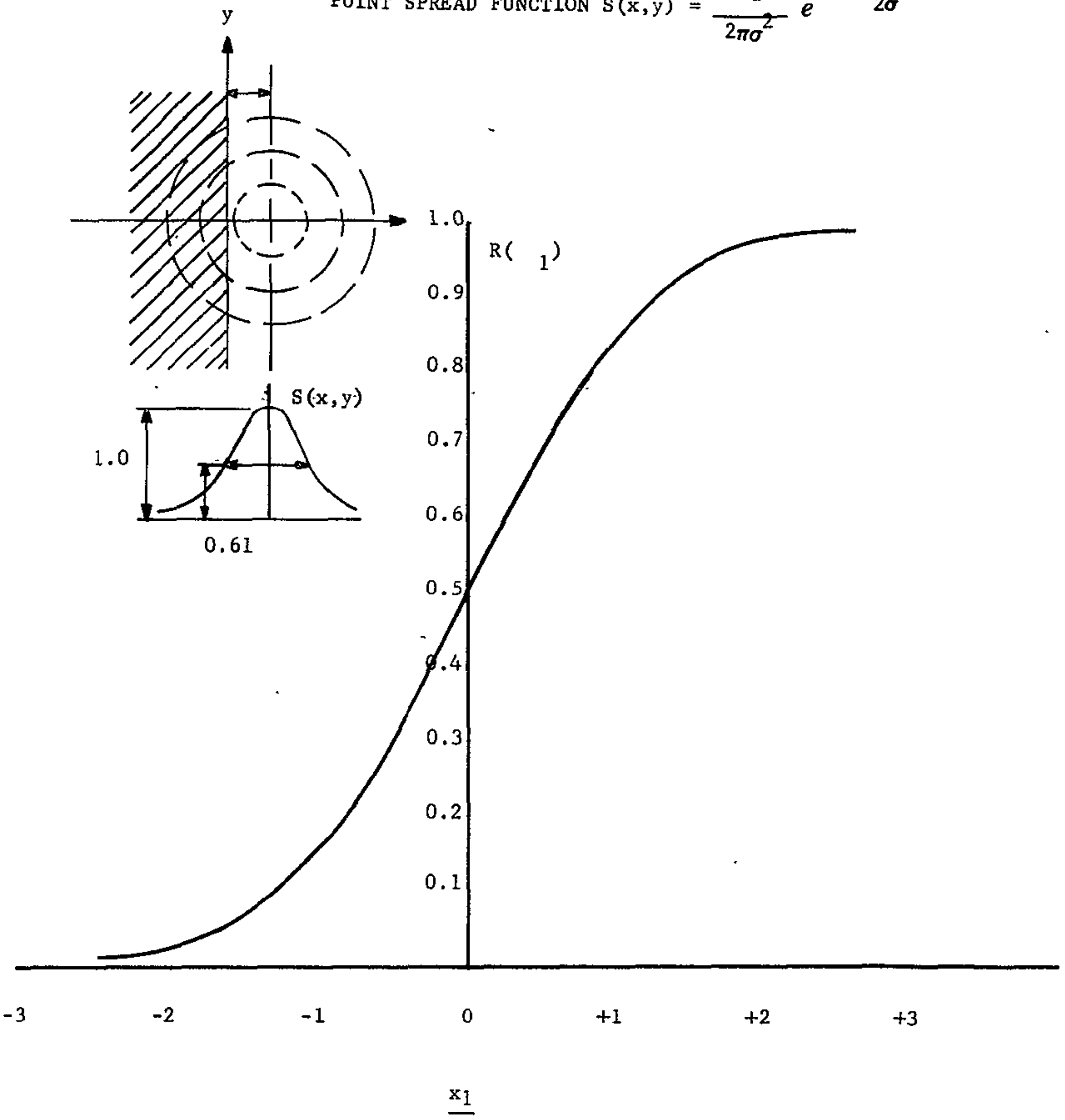
$$= \tfrac{1}{2} \left[1 - \mathrm{erf}\ \left(\frac{x_1}{\sigma \sqrt{2}} \right) \right] \quad \text{for } x_1 < o$$

$$= 1 \left[1 + \mathrm{erf}\ \left(\frac{x_1}{\sigma \sqrt{2}} \right) \right] \quad \text{for } x_1 > o\ ,$$

where $\mathrm{erf}\ \left(\frac{x_1}{\sigma \sqrt{2}} \right) = \frac{1}{\sigma \sqrt{2}} \int_{-x_1}^{x_1} \exp \left[-\tfrac{1}{2} \left(\frac{x}{\sigma} \right)^2 \right] dx.$

This step response function is plotted in Figure 3-10 as a function of the normalized beam location. The 10%-to-90% rise time is approximately 2.6σ or 1.3 times the spot diameter $2\ \sigma$.

POINT SPREAD FUNCTION $S(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}}$



X_1 = DISPLACEMENT OF BEAM CENTER FROM KNIFE EDGE

Figure 3-10 Normalized Step Response for a Gaussian Spot Scanning a Knife Edge

3.3.5.6 Contribution of Image Motion to MTF

To calculate the contribution to the overall MTF caused by image motion during frame time, the following data are required:

a. Speed of longitudinal image plane dimension during frame time.

b. Distance traveled by longitudinal image plane dimension during exposure time.

c. Normalized one-dimensional Fourier transform of a pulse of amplitude A and space dimension equal to uncompensated distance traveled by longitudinal image plane dimension during exposure time. Consider a swath width in the focal plane is with a speed of approximately 90 MM per second. During a .01 second exposure time this length will move .9 MM. Assuming a 99% effective image motion compensation, the effective image motion will be approximately .009 MM. This .009 MM is the width T_p of the pulse shown in Figure 3-11a used as the model for determining image motion contribution to the optics system response.

$T_4(k)$ is derived by taking the Fourier transform of a pulse of amplitude A and duration T_p where:

$$T_p = \frac{y}{t_{frame}} (1 - IME)$$

y = longitudinal dimension of image plane,

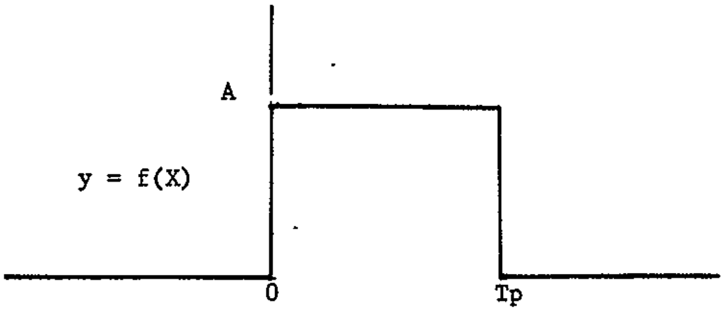
t_{frame} = frame time, seconds

and IME = image motion compensation efficiency, %.

The result is the function

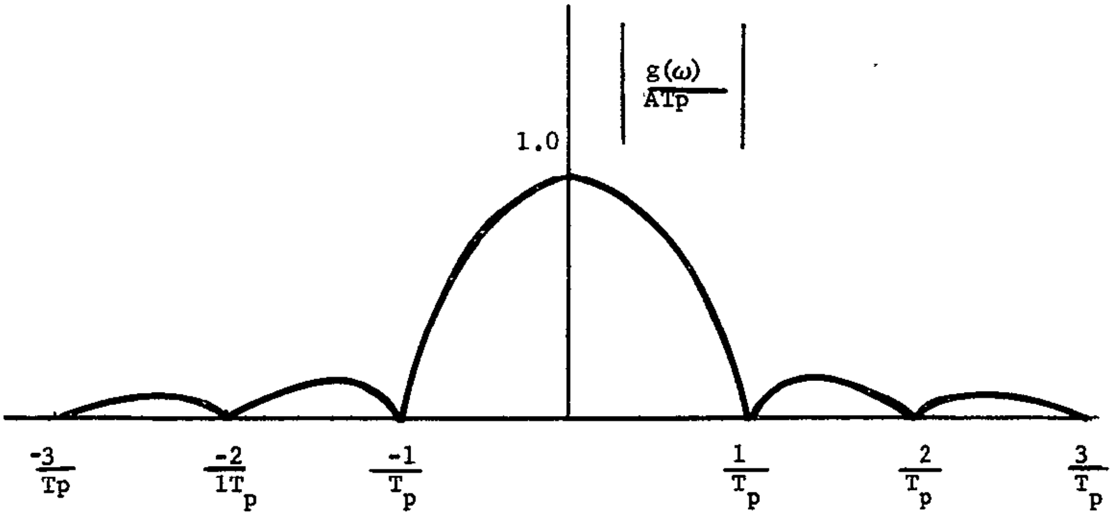
$$\frac{g(k)}{AT_p} = \frac{\text{Sin}\,(\pi\, kT_p)}{\pi\, kT_p}$$

where k is the limiting spatial frequency of the associated vidicon,



Assumed f(X) For Image Motion MTF Derivation

Figure 3-11 a



Fourier Transform of y = f(X)

Figure 3-11 b

I-609A

But $\omega = 2\pi f$., so that the foregoing equation becomes

$$\frac{g(f)}{aT_p} = \frac{\text{Sin}\ (\pi f T_p)}{\pi f T_p}$$

A plot of this last equation is depicted on Figure 3-11b. The first zero of the above function occurs at

$$\pi f T_p = \pi$$

$$\text{or} \quad f = \frac{1}{T_p}$$

Let $T_p = .009$ MM so that $f = 1/T_p = 111$ LP/MM

The response of 26.4 LP/MM is computed from above, i.e.,

$$\frac{g(f)}{AT_p} = \frac{\text{Sin}\ (3.14 \times 26.4 \times .009)}{3.14 \times 26.4 \times .009} = 0.91$$

3.3.5.7 Cathode Ray Tube MTF

The derivation of the MTF for a CRT depends on the definition of the spot diameter. If, as is generally done, the spot is assumed to have a Gaussian distribution, then we define the given diameter, D, as the spot width between the half-amplitude points. However, the MTF calculation is based on the one-sigma (.7) points. Therefore, the sigma width value must be calculated. A high resolution CRT having a spot diameter of .0006" or .01524 MM is assumed. Using

$$\sigma\ (X) = e^{-1/2} \left(\frac{X}{\sigma}\right)^2$$

the spread function defining spot luminance,

where $X = .01524/2 = .00762$

and $\sigma\ (X) = .5$

$$\text{Then} \qquad .5 = e - 1/2 \left(\frac{10^{-4} x \;\; .580644}{\sigma \; 2}\right)$$

$$\text{or} \qquad \text{Log } .5 = -1/2 \left(\frac{10^{-4} x \;\; .580644}{\sigma \; 2}\right)$$

$$\text{so that } \sigma^2 = \frac{10^{-4} x \;\; .580644}{\sigma \; 2}$$

$$\sigma = .00552 \text{ MM}$$

and 2σ = .011 MM, which being the width between the .7 points on a Gaussian curve, is less than .01524 MM, defined as the width between the .5 points.

This value is now substituted into

$$T(k) = e^{-1/2} \, (2 \, \pi \, k \, \sigma)^2$$

which is the MTF expression for this Gaussian spread function.

$$\text{Hence, } T(k) = e^{-1/2} (\pi \, k .011)^2$$

$$= e^{-.0012 \, k^2}$$

This expression is now to be evaluated at the two illustrative frequencies of 26.4 LP/MM and 49.2 LP/MM, i.e.,

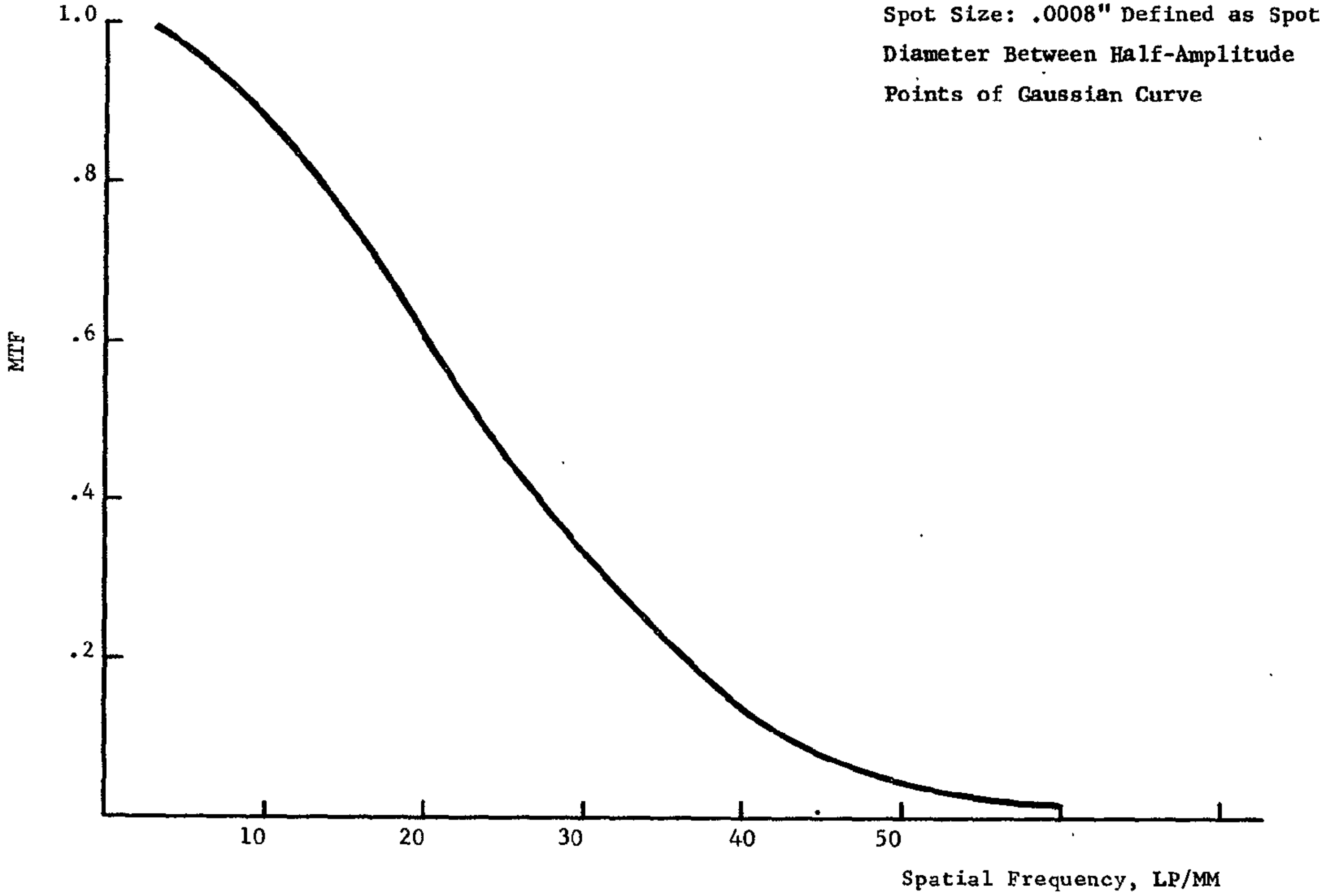
$$T(26.4) = e^{-.0012 \, (26.4)^2} \approx e^{-.836} \approx .43$$

$$T(49.2) = e^{-.0012 \, (49.2)^2} = e^{-2.9} \approx .06$$

A plot of T(k) is presented on Figure 3-12.

3.3.5.8 Image Dissector

The major exception to the Gaussian shaped aperture in TV sensors is the Image Dissector. In the Image Dissector the aperture is a mechanical hole in a metal disc and theoretically can be of any shape. Practical resolving aperture diameters range from 0.0005" to 0.060". Circular or square apertures are normally used. Image Dissectors are capable of very high resolution at low scan speeds (at high speeds the aperture



Spot Size: .0008" Defined as Spot Diameter Between Half-Amplitude Points of Gaussian Curve

Figure 3-12 Plot of MTF, High Resolution CRT

must be made larger to emit more photoelectrons to retain adequate signal-to-noise ratio). At low speeds limiting resolution of 3200 TV lines/inch (63 lp/mm) can be obtained.

The point spread function for an image dissector with a rectangular aperture is defined as follows:

$$s(x) = \begin{cases} 1/a \text{ for } x \text{ between } -2/2 \text{ and } a/2 \\ 0 \text{ for all other } x \end{cases}$$

where a = aperture width

Since the aperture is symmetrical the MTF can be determined by evaluation of:

$$T(k) = \frac{\left| \int_{-\infty}^{\infty} s(x) \cos 2\pi kx \, dx \right|}{\left| \int_{-\infty}^{\infty} s(x) \, dx \right|}$$

For the rectangular spread function, the integral in the numerator becomes:

$$2/a \int_{0}^{a/2} \cos 2\pi kx \, dx$$

$$= \frac{\sin \pi ka}{\pi ka}$$

The integral in the denominator becomes by inspection:

$$\int_{-\infty}^{\infty} x(x) \, dx = 1$$

The MTF is, therefore:

$$T(k) = \frac{\sin \pi ka}{\pi ka}$$

which is plotted in Figure 3-13.

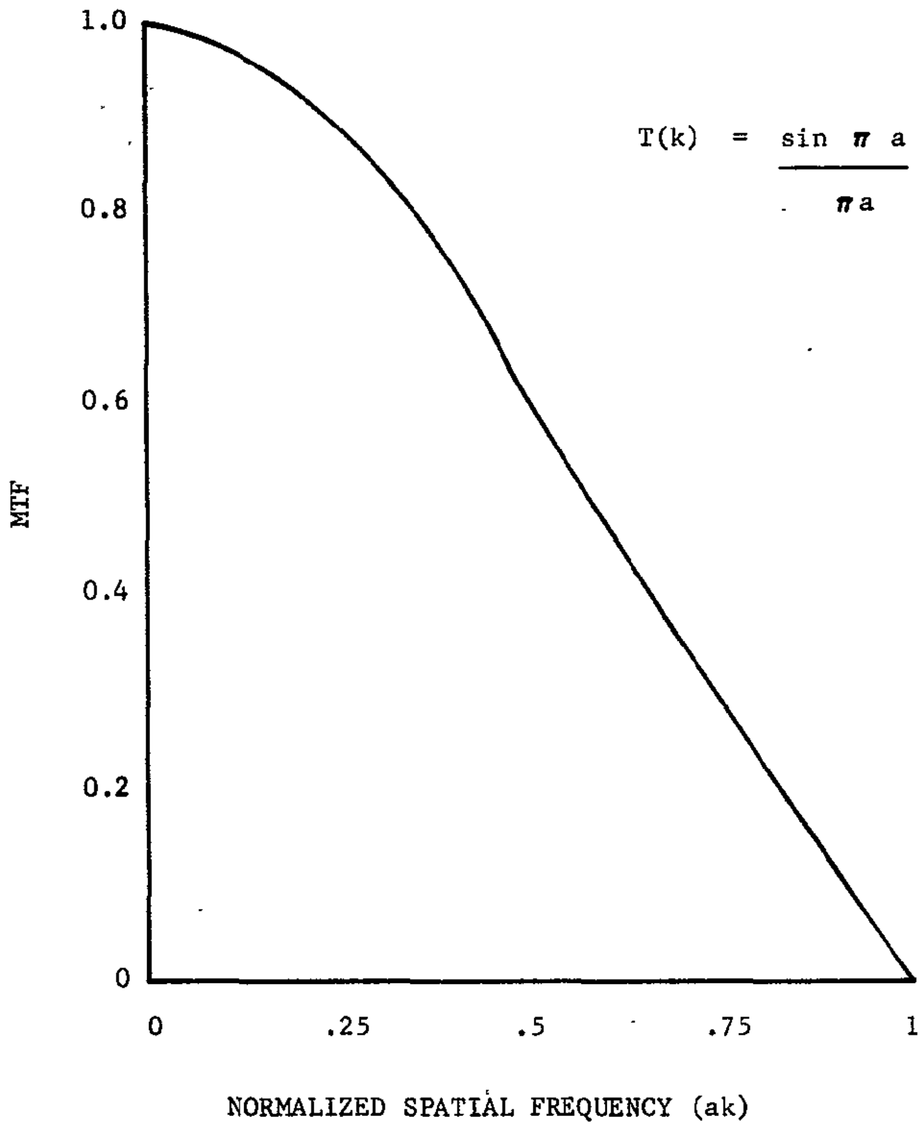


Figure 3-13 MTF OF AN IMAGE DISSECTOR WITH A RECTANGULAR APERTURE OF WIDTH a

3.3.5.9 MTF of Side Looking Radars

Side-looking radars* fall into two general classes: noncoherent and coherent. The noncoherent radars are essentially conventional pulsed radars with the look angle normal to the vehicle track. Range resolution is dependent upon the pulse width while azimuthal resolution is proportional to antenna beamwidth.

Coherent side-looking radars maintain phase coherence between the transmitted and received RF signals. This coherence makes possible improved azimuthal resolution using synthetic aperture processing. The synthetic aperture technique utilizes the motion of the vehicle and the doppler history of the target returns to compress the azimuthal resolution down to a value nearly equal to the real antenna along-track dimension. The range resolution is determined by the effective pulse width of the radar. Either conventional short pulses can be used or pulse compression can be employed to reduce transmitter peak power.

A radar transmitting ideal rectangular pulses of width T has an MTF expressed by:

$$T(k) = \left| \sin \pi kT / \pi kT \right|$$

This MTF is identical to the MTF given earlier for the image dissector and plotted in Figure 3-13. The pulse width T corresponds to the image dissector aperture width a.

The above MTF is normalized for some maximum target reflectance. For a lower target reflection coefficient the magnitude of the radar receiver output will be reduced. Thus, the effective MTF for a radar is scaled by the target reflectance. The position of the target in range and azimuth will also scale the MTF. Nonetheless, the MTF is a useful parameter for determining the relative frequency content in the return signal.

---

*Side looking radar has been discussed in Section 3 of Volume 1. There, it is used as a surface imaging or mapping technique (JPL Space Programs Summary 37-55, Vol. III, pp 78-89) and is discussed here in terms of its MTF to further extend the range of this report.

In most practical radars the pulse is not an ideal rectangle. Practical limitations on rise time and filtering will modify the MTF. Also, the side lobe response in the ideal sin x/x signal spectrum is undesirable and is usually minimized by careful pulse shaping. The MTF for an actual radar may very well look somewhat different from that shown above. However, the sin x/x model and possibly the Gaussian model are quite suitable for use in determining A-to-D converter techniques.

If conventional short pulses are used in the synthetic aperture radar then the radar MTF is essentially the same as that given above for the noncoherent radar.

However, most synthetic aperture radars use pulse compression or chirp to allow longer transmitted pulse width and, hence, lower peak power. Pulse compression also permits two-dimensional matched-filter receiver processing using an optical correlation.

The MTF for a pulse compression radar measured in the receiver prior to compression or de-chirping is quite different from the MTF for conventional short pulse radar.

For linear FM pulse compression, which is commonly known as chirp, a linearity swept RF modulation is transmitted on each pulse, where T is the transmitted pulse duration, f_c is the RF center frequency and f_1 is the frequency deviation of the linear sweep. The product f_1 T is called the pulse compression ratio since the compressed pulse after receiver processing will be approximately T/ (f_1 T).

The amplitude spectrum for a chirp radar can be shown to have an envelope described by:

$$\left| P(f) \right| = \sqrt{\frac{T}{2f_1}} \left| \int_{\mu_2}^{\mu_1} \exp\left[j\pi\, y^2/2 \right] dy \right|$$

where $\mu_1 = \sqrt{2f_1T} + (f_o - f)\sqrt{2T/f_1}$

$$\mu_2 = (f_o - f)\sqrt{2T/f_1}$$

The above integral can be evaluated as the difference of two Fresnel integrals, $Z(\mu_1) - Z(\mu_2)$, where

$$Z(\mu) = \int_o^{\mu} \exp\left[j\pi\, y^2/2\right] dy.$$

The normalized amplitude spectrum $\left|P(f)\right| / \sqrt{T/f_1} = \frac{1}{\sqrt{2}}\left|Z(\mu_1) - Z(\mu_2)\right|$

is plotted in Figure 3-14 for several values of the compression ratio f_1T. The absicissa is the normalized one-sided frequency off-set relative to the RF center frequency, $(f - f_c)/f_1$. The spectrum is symmetrical relative to $(f - f_c)/f_1 = 0$.

The amplitude spectrum for the ideal rectangular pulse is also plotted for comparison.

Examination of Figure 3-14 indicates that there is considerable high frequency content in the chirp pulses, significantly more than in the other sensor MTF's considered earlier. This distinction is important in determining A-to-D conversion characteristics.

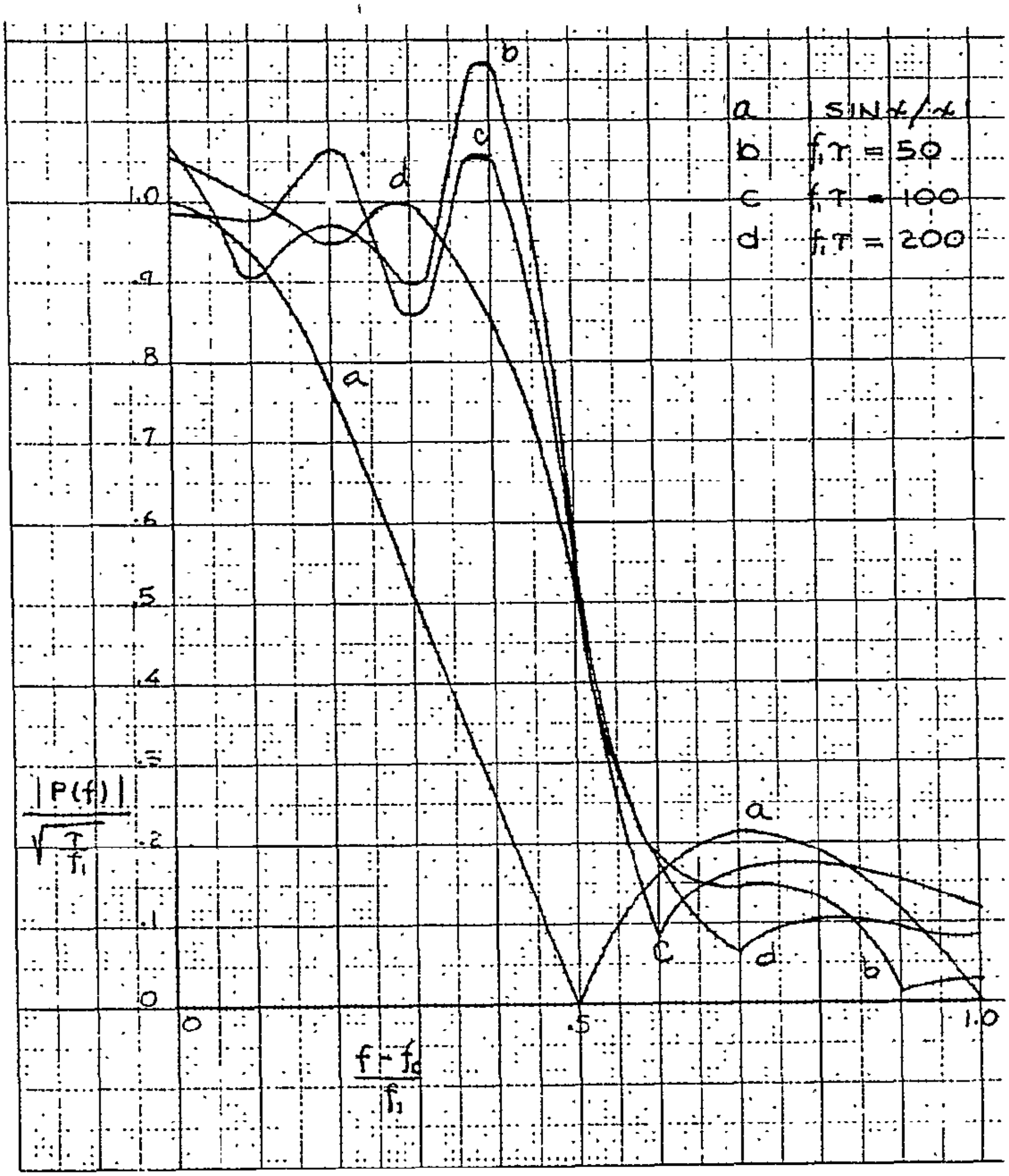


Figure 3-14 Normalized Amplitude Spectrum for Coherent Side Looking Radar

Section 4

HIGH SPEED A/D CONVERTER TECHNOLOGY

4.1 HIGH SPEED A/D CONVERTERS

A/D converts operating in the 40 to 400 Mbps rate range, ust use high speed components and a minimum number of successive logic levels. The switching technology which related to these speeds has been discussed in Section 1 of this volume; this section will discuss the various converters which can be used at these speeds, and related to Section 4 of Volume 1.

4.2 TYPES OF HIGH SPEED A/D CONVERTERS

At the 40-400 Mbps bit-stream speeds, few A/D converter circuits can meet the requirement of speed, low noise and with accuracy. The following paragraphs will discuss the basic aspects of several candidate ultra high speed A/D converter techniques which can operate at these speeds, i.e.

- Cyclic conversion encoder
- Phase-plane conversion encoding
- Threshold encoders
- Scanning tube-encoder
- Delta modulation method of feedback encoding (redundancy removal).

Delta modulation technology is particularly suited to the transmission of high speed sensor data, and a computer analysis will be provided at the end of this section.

4.2.1 Folding Converter A/D Conversion

In 1956, B. D. Smith published a novel analog-to-digital conversion method which has been adopted for use in many modern systems including the Bell Telephone 224 Megabit PCM system.

This system employs the "folding" or cyclic technique employing a series of identical stages as shown in Fig. 4-1a. Each stage provides one digit of the binary code corresponding to the sample. Each stage has the same inverted "V" input-output characteristic such that it has a slope of +2 for negative input signals and a slope of -2 for positive input signals. The coder operates directly in the Gray code. The digits are obtained from each stage; if the input to a stage is less than half the maximum amplitude, the digital output is zero; if it is greater than half the maximum, then the digit output is one. As each stage develops its output, it yields a "residue" which is used to drive the next stage. The digit outputs of each of the stages are passed through delay lines to provide the serial bit stream. (Fig.4-1b)

If, in the circuit of Fig. 4-1a, the operation of Stage 1 is such that

$$E_{D1} = \begin{cases} 1 & \text{if} \quad E_1 > 0 \\ 0 & \text{if} \quad E_1 < 0 \end{cases}$$

then the residue output, E_2, to Stage 2 is related to input voltage, E_1 by

$$E_2 = E_{ref} - 2|E_1|$$

where E_{ref} is a dc reference voltage which is related to the bit being determined. In a pure binary system, for example, the reference voltage would refer to the 2^o, 2^1, 2^2, . . . , 2^n levels, respectively.

4.2.1.1 Fisher High Speed Folding Encoder. R. E. Fisher of Bell Telephone has developed a 1200-megabit-per-second analog-to-digital converter which provides 5-bit coding and which utilizes microwave equivalents of low-frequency logic circuits, using the folding techniques described in connection with Fig. 4-1a using phase domain technology which is discussed in the following paragraphs.

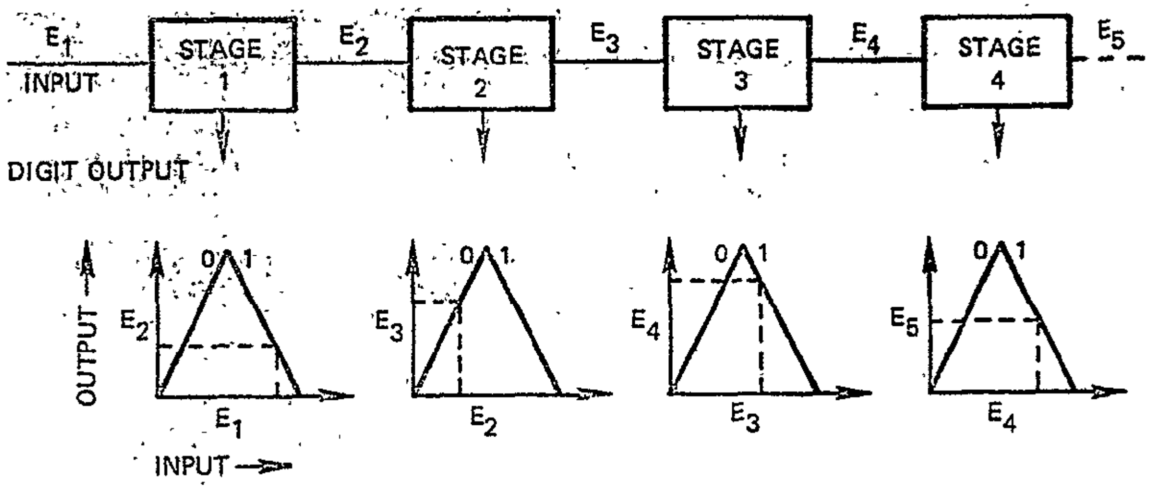

E1
INPUT
STAGE 1
E2
STAGE 2
E3
STAGE 3
E4
STAGE 4
E5
DIGIT OUTPUT
OUTPUT
INPUT
0 1
E2
E1
E3
E2
E4
E3
E5
E4

Fig. 4-1.a Output versus Input Wave Forms of the Stages of the Folding Coder

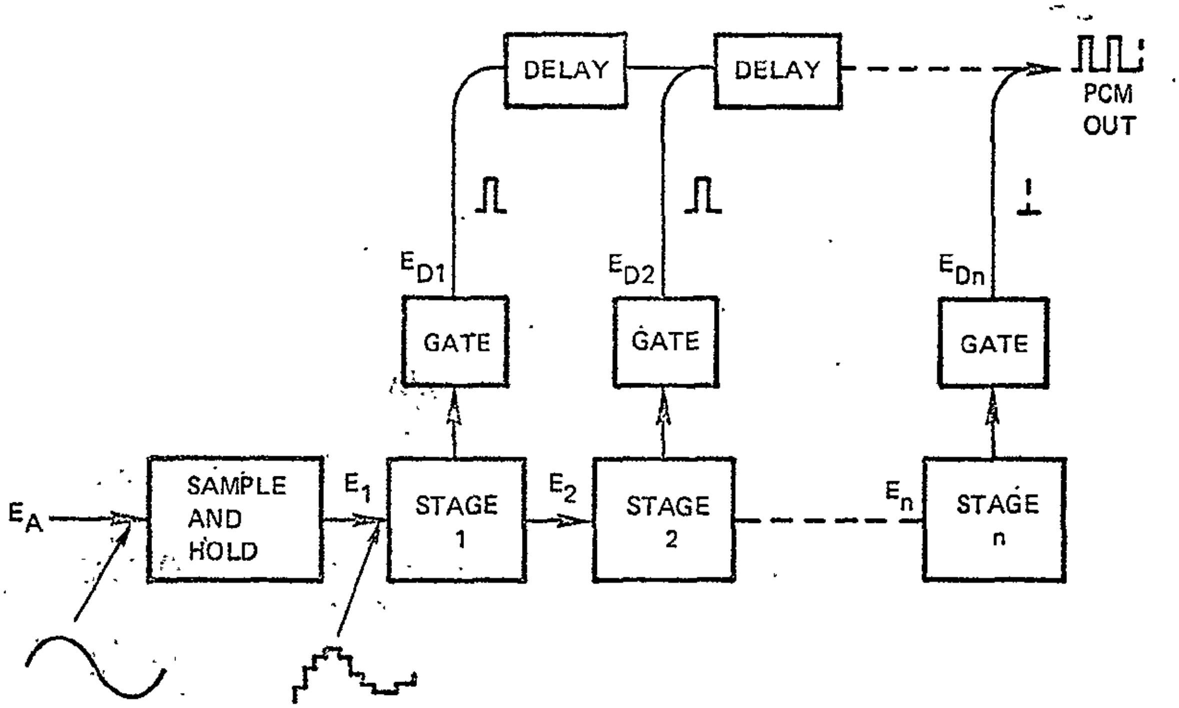

DELAY
DELAY
PCM OUT
ED1
ED2
EDn
GATE
GATE
GATE
EA
SAMPLE AND HOLD
E1
STAGE 1
E2
STAGE 2
En
STAGE n

Fig. 4-1.b N-Stage Serial Type of Folding Coder for Analog-to-Digital Conversion

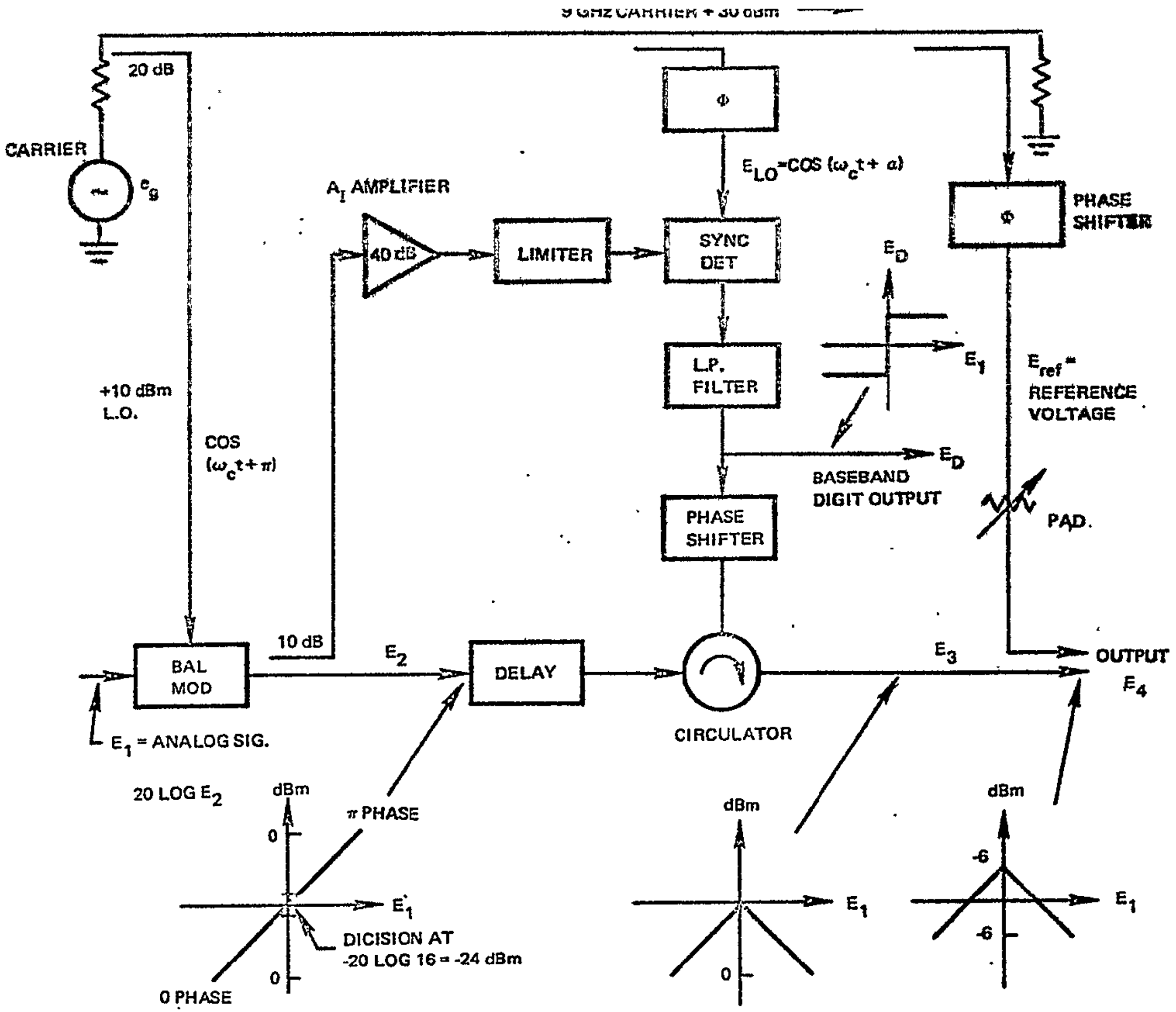
Stage 1 of this circuit is shown in Fig.4-2; it was designed to sample and hold at a 240-MHz rate and encode the sampled analog signal into 5-bit words (32 quantization levels) in the Gray code. This sample and hold rate, combined with 5-bit accuracy makes such a system attractive for the transmission of high speed image data.

In this circuit, the baseband analog signal amplitude is replaced by the amplitude of a 9-GHz microwave carrier while the baseband signal polarity is described by microwave carrier phase.

In the circuit of Fig. 4-2, the balanced modulator output, E_2, has either Π phase or zero phase. If E_2 has Π phase, then the stage must emit E_o in the form of the digit 1 and shift E_3 to zero phase. On the other hand, if E_2 has zero phase, the system must emit E_o in the form of a digit zero and maintain E_3 at zero phase. A CW carrier of correct magnitude and phase (E_{ref}) is added to E_3 to provide the output signal to drive the next stage. The implementation of the phase domain will be discussed in the next paragraphs.

4.2.2 Phase Domain Encoders

Phase-domain encoding is capable of highest speeds, using balanced-modulator techniques introduced in the Fisher A/D converter of Fig. 4-2. Phase domain encoding, using the phase-locked oscillator was the first approach considered for high-speed computers. The phase-locked oscillator, consists of a tank circuit which is forced to provide parametric oscillations by a pump operating at twice pump frequency. Sustained oscillations can occur in either of two possible phases 180^o apart. The circuit can be caused to change phase states by "Locking" the oscillations to the phase state of a locking signal. Computer logic systems using phase-locked oscillators were built in 1958-1959 using clock rates above one gigahertz, performing logic at rates as high as 1.85 GHz. However, the advent of the tunnel diode and the inherent complexity of a PLO computer caused interest in this technology to wane until the present era of high-speed bit stream requirements.


9 GHz CARRIER + 30 dBm
20 dB
CARRIER
e_g
A_I AMPLIFIER
40 dB
LIMITER
SYNC DET
Φ
E_{LO}=COS ($\omega_c t+\alpha$)
Φ
PHASE SHIFTER
E_D
E_1
E_{ref} = REFERENCE VOLTAGE
L.P. FILTER
+10 dBm L.O.
COS ($\omega_c t+\pi$)
E_D
BASEBAND DIGIT OUTPUT
PHASE SHIFTER
PAD.
10 dB
E_2
BAL MOD
DELAY
CIRCULATOR
E_3
OUTPUT
E_4
E_1 = ANALOG SIG.
20 LOG E_2
dBm
π PHASE
0
E_1
DICISION AT -20 LOG 16 = -24 dBm
0
0 PHASE
dBm
E_1
0
dBm
-6
E_1
-6

Five-bit microwave coder, schematic diagram of stage 1.

Fig. 4-2 Details of One Stage of the Fisher Microwave Analog/Digital Converter to Gray Code Including Microwave Circuit Equivalent of One Stage for 5-bit 1200 Megabit Transmission

Many basic limitations to speed can be removed by at least an order of magnitude by transformation of the voltage variable into the phase domain.

By operating the phase domain, it is no longer necessary to amplify, rectify, compare, or otherwise operate upon the amplitude of a voltage analog signal. Instead, these operations are performed upon the phase of a carrier, where the amplitude of the carrier is not of primary importance but need only be sufficiently greater than noise.

All analog-to-digital converters generate a binary code or some other code made up of a sequence of only two digit value. By way of definition, the term "binary" will refer to the natural binary number system specified by the series:

$$\sum_{n=0}^{N} B_n 2^{N-n} \qquad n = 0, 1, 2, \ldots, N$$

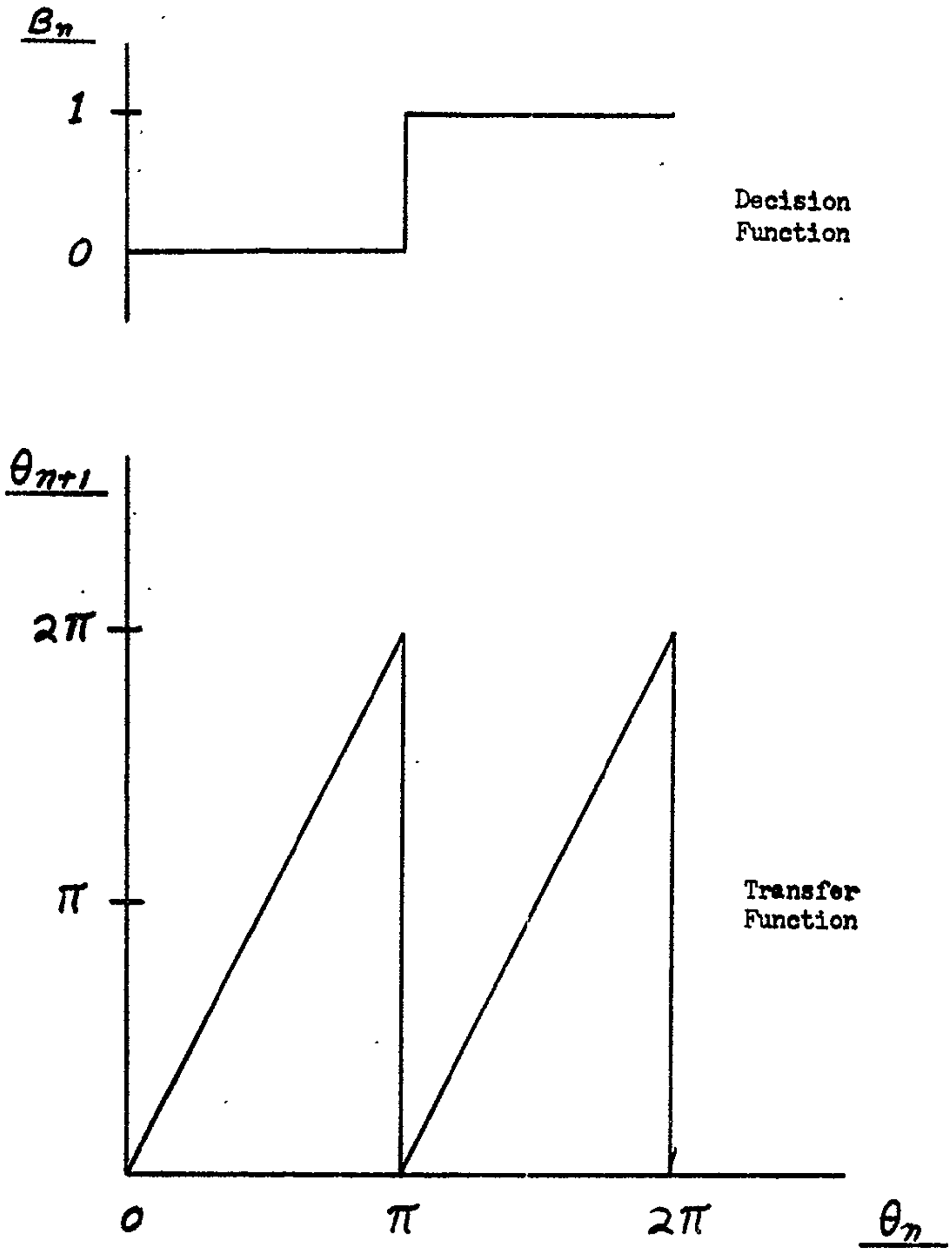
Let the analog variable, f(t), be represented by a proportional phase variable, $\theta_o(t)$, having the range $0 \leq \theta_o < 2\pi$. Let $\theta_n = 2^n \theta_o \pmod{2\pi}$ where $n = 0, 1, 2 \cdots$. Then the $(n+1)^{th}$ most significant bit, B_n, is determined by the formula:

$$B_n = 0 \quad \text{for} \quad 0 \leq \theta_n < \pi$$

$$B_n = 1 \quad \text{for} \quad \pi \leq \theta_n < 2\pi$$

Given any θ_o, one can find the corresponding B_n from the transfer and decision functions illustrated in Fig. 4-3. The transfer function has a slope of two and is modulu - 2π. The decision function has the value zero or one, depending upon θ_n. The relationship between θ_o, the input phase variable, and the resulting sequence of binary bits is given by the formula:

$$\theta_o \simeq \frac{2\pi}{2^{N+1}} \sum_{n=0}^{N} B_n 2^{N-n} \quad \text{radians}$$



$$\theta_0 \simeq \frac{2\pi}{2^{N+1}} \sum_{n=0}^{N} B_n 2^{N-n} \text{ radians,} \quad n = 0, 1, 2, \cdots, N$$

Figure 4-3 Encoder Logic

twice the phase variation. This is shown in the following derivation of the output of the n^{th} encoder stage.

Let the input to the first encoding stage be $e_o(t) = \sin(\omega t + \theta_o)$. If we take the second harmonic of $e_o(t)$ and mix it with the third harmonic, $\cos 3\omega t$, of the carrier by means of a balanced modulator, we have:

$$\sin(2\omega t + 2\theta_o) \cdot \cos 3\omega t = -\tfrac{1}{2} \sin(\omega t - 2\theta_o) + \tfrac{1}{2} \sin(5\omega t + 2\theta_o)$$

The second of the modulation products is filtered out. The first is amplified by minus two and becomes the input to the second encoding stage:

$$e_1(t) = \sin(\omega t - 2\theta_o)$$

Similary, we have as an input to the third encoding stage:

$$e_2(t) = \sin(\omega t + 4\theta_o)$$

Therefore the output of the n^{th} encoding stage is

$$e_n(t) = \sin[\omega t + (-2)^n \theta_\theta]$$

or, since we have defined $\theta_n = 2^n \theta_o$,

$$e_n(t) = \sin[\omega t + (-1)^n \theta_n]$$

It will be noticed that the sign of the phase reverses in each succeeding stage. In the stages of odd n this results in merely generating the compliment of B_n.

The first encoder stage produces the most significant bit, B_o, by comparing the phase of the phase modulated signal, $\sin(\omega t + \theta_o)$,

where $n = 0, 1, 2, \cdots, N$
and $N+1$ = number of bits/word.

The encoding scheme described in the preceding section can be implemented in the following manner. The binary number is derived by feeding the analog signal, f(t), to a voltage-to-phase converter, or phase modulator, the output of which is the phase modulated signal, $\sin[\omega t + \theta_o(t)]$, whose phase, $\theta_o(t)$, compared with a reference phase, δ, of the carrier, $\sin(\omega t + \delta)$, varies in proportion to the applied input voltage, f(t). The phase modulated output is sensed by a phase comparator which compares the phase, $\theta_o(t)$, of the modulated signal with a reference phase, δ, of the carrier. In all the following definitions of phase relationships we shall assume $\delta = 0$. The output of the phase comparator is one of two voltage levels representing a one or zero, indicating whether the phase of the modulated signal was greater or less than the reference phase. By this means, the first most significant bit, B_o, in the natural binary system is obtained.

Successively lesser significant bits are obtained from cascaded encoding stages having the transfer and decision functions illustrated in Fig. 4-3/ The transfer function is implemented as follows: The phase modulated signal is applied to a doubler circuit which multiplies the frequency and phase by a factor of two. The multiplied signal is then mixed with the third harmonic of the carrier and the difference frequency obtained by filtering. The output of the filter is amplified by minus two. The output of the amplifier, then, is a phase modulated signal having the same amplitude as the input to the encoding stage but having with the phase of the carrier having a fixed phase shift, φ. This comparison is easily accomplished by first (a), generating a narrow pulse, $u'(\sin \omega t + \theta_o) = u(\theta_o)$ at the time of the positive going zero crossing of the phase modulated signal, and (b), generating the function, $F'(\sin \omega t + \varphi) = F(\varphi)$. $F(\varphi)$ is a simple limiting operation defined as:

$$F(\varphi) = 0 \quad \text{for } \sin(\omega t + \varphi) < 0$$

$$F(\varphi) = 1 \quad \text{for } \sin(\omega t + \varphi) \geq 0$$

Now, if we let $\varphi = \pi$, then B_o will be obtained at the output of an "and" gate having the inputs $u(\theta_n)$ and $F(\pi)$. In general, B_n will be obtained at the output of an "and" gate having the inputs $u(\theta_n)$ and $F(\pi)$. Fig. 4-4 is a waveform diagram showing the relationship between $u(\theta_n)$ and $F(\pi)$.

A block diagram of three cascaded encoding stages corresponding to the implementation described in this section is shown in Fig. 4-5.

4.2.3 Fast Shift Register Encoders

Fast shift registers can be used for storage or encoding using the shift register whose basic circuit at microwave frequencies is a series of flip-flops triggered by a clock or timing circuit. This unit encodes the output of the quantizer..

A series of bits is applied to the input in synchronization with a clock which actuates all flip-flops simultaneously so that a given series of bits moves or progresses through the shift register and occupies the shift register - one bit per flip-flop at a rate determined by the clock. Once the shift register is occupied, the serial input is removed. Application of the clock again to the flip-flop will read out the stored bits as a serial bit stream output. If during the readout the clock is very fast, the bit stream will comprise a series of bit bursts. For gigahertz data rates, the shift register should be capable of moving bits from flip-flop to flip-flop in time intervals equal to or less than a nanosecond; the serial bit stream read-in can be at a much lower rate.

Present shift registers can now shift state data from stage to stage at rates up to one gigahertz using, for example, tunnel diode shift registers shown in Fig. 4-6 (due to B. Sear)* or using emitter coupled logic with clock routing, interconnect and bit stream distribution circuits designed to accommodate these high speeds.

---

*B. Sear, "Research for High-Speed Analog-to-Digital Conversion Techniques," WPAFB Technical Documentary Report, No. AL TDR-64-154.

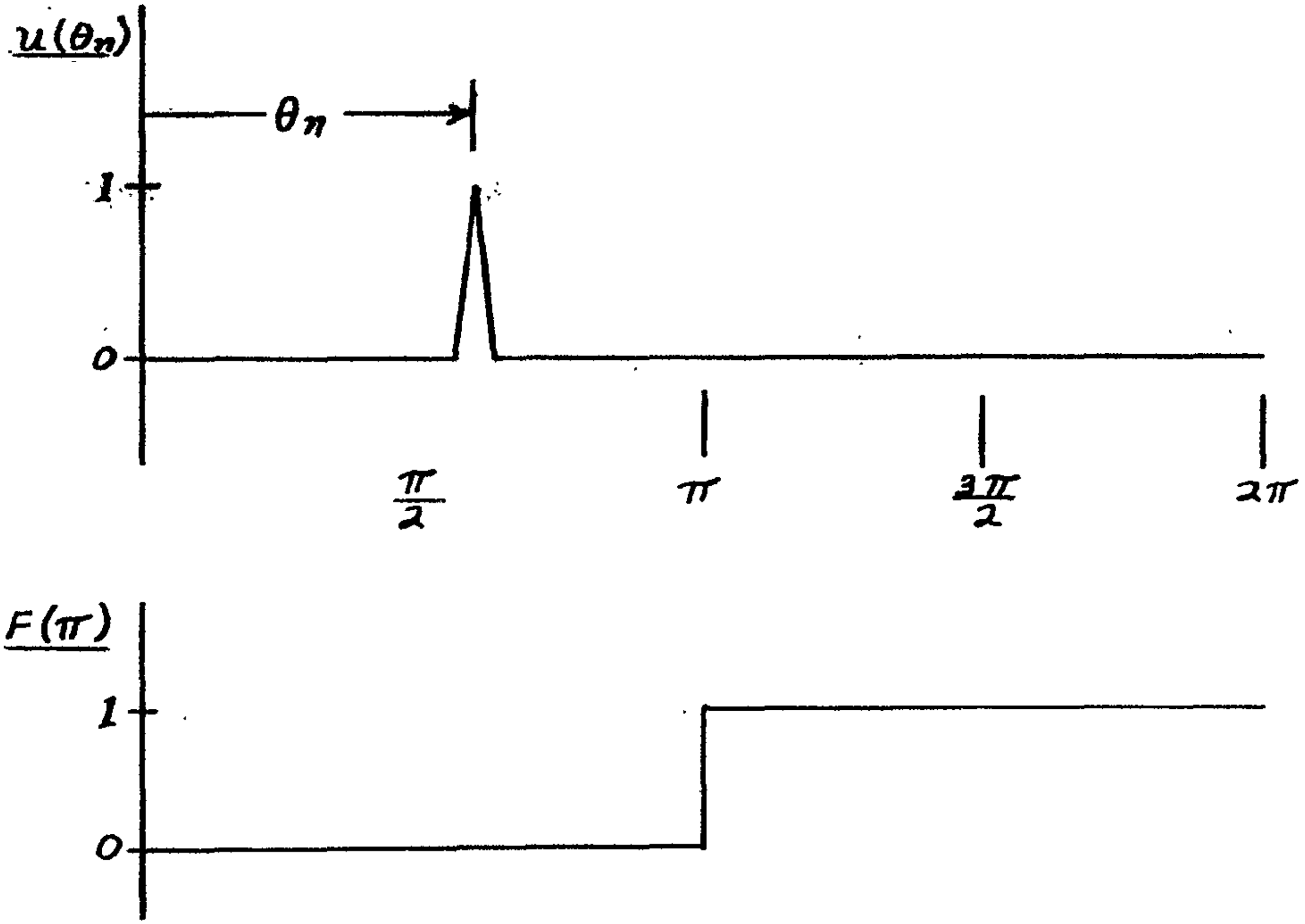


Fig.4-4 Waveform Diagram of u(θ_n) and F(π)

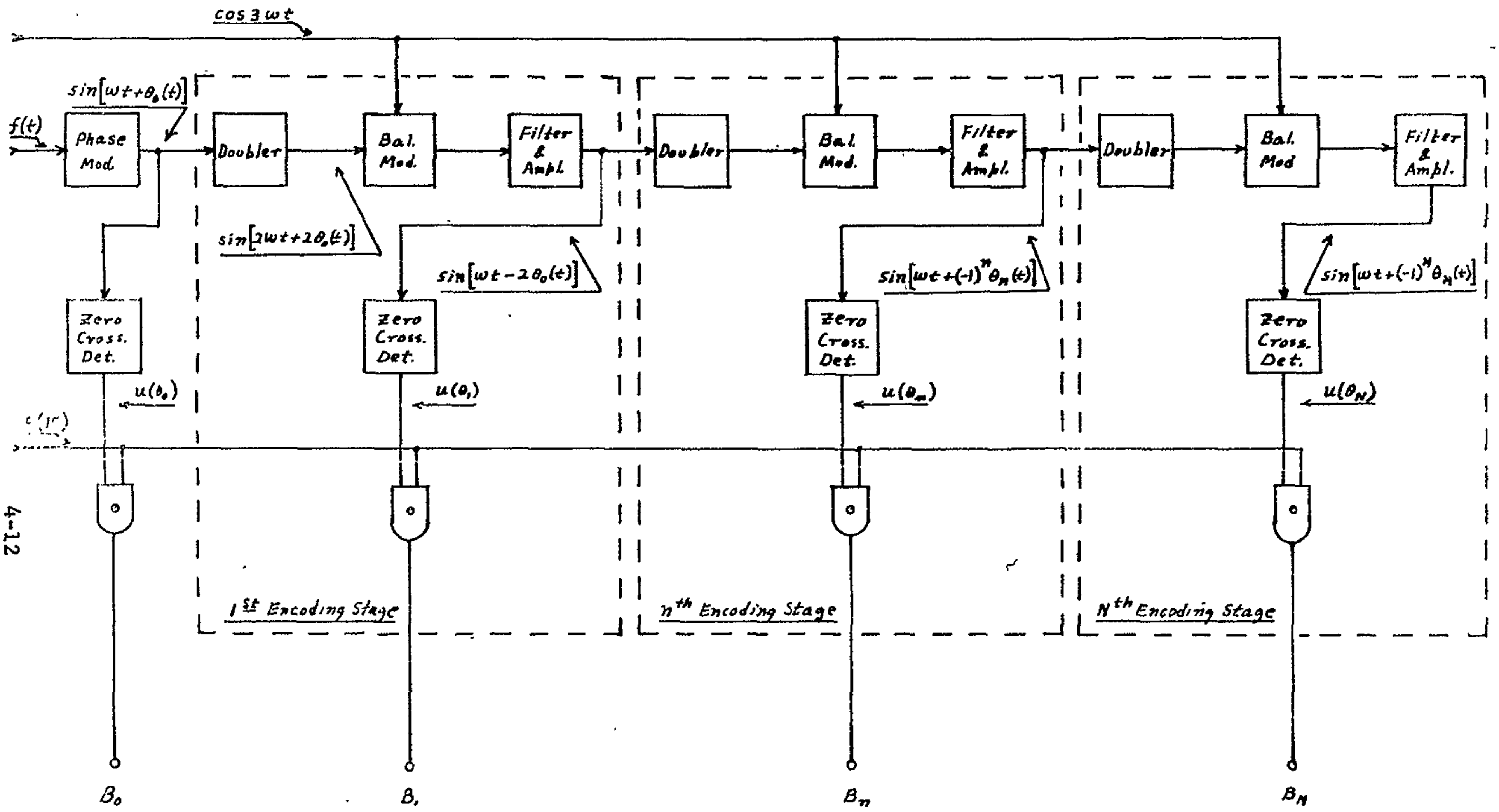

cos 3ωt
f(t)
sin[ωt+θ₀(t)]
Phase Mod.
Doubler
Bal. Mod.
Filter & Ampl.
sin[2ωt+2θ₀(t)]
sin[ωt−2θ₀(t)]
sin[ωt+(-1)ⁿθₙ(t)]
sin[ωt+(-1)ᴺθ_N(t)]
Zero Cross. Det.
u(θ₀)
u(θ₁)
u(θₙ)
u(θ_N)
1st Encoding Stage
nth Encoding Stage
Nth Encoding Stage
B₀
B₁
Bₙ
B_N

Fig. 4-5 Block Diagram of Binary Encoder

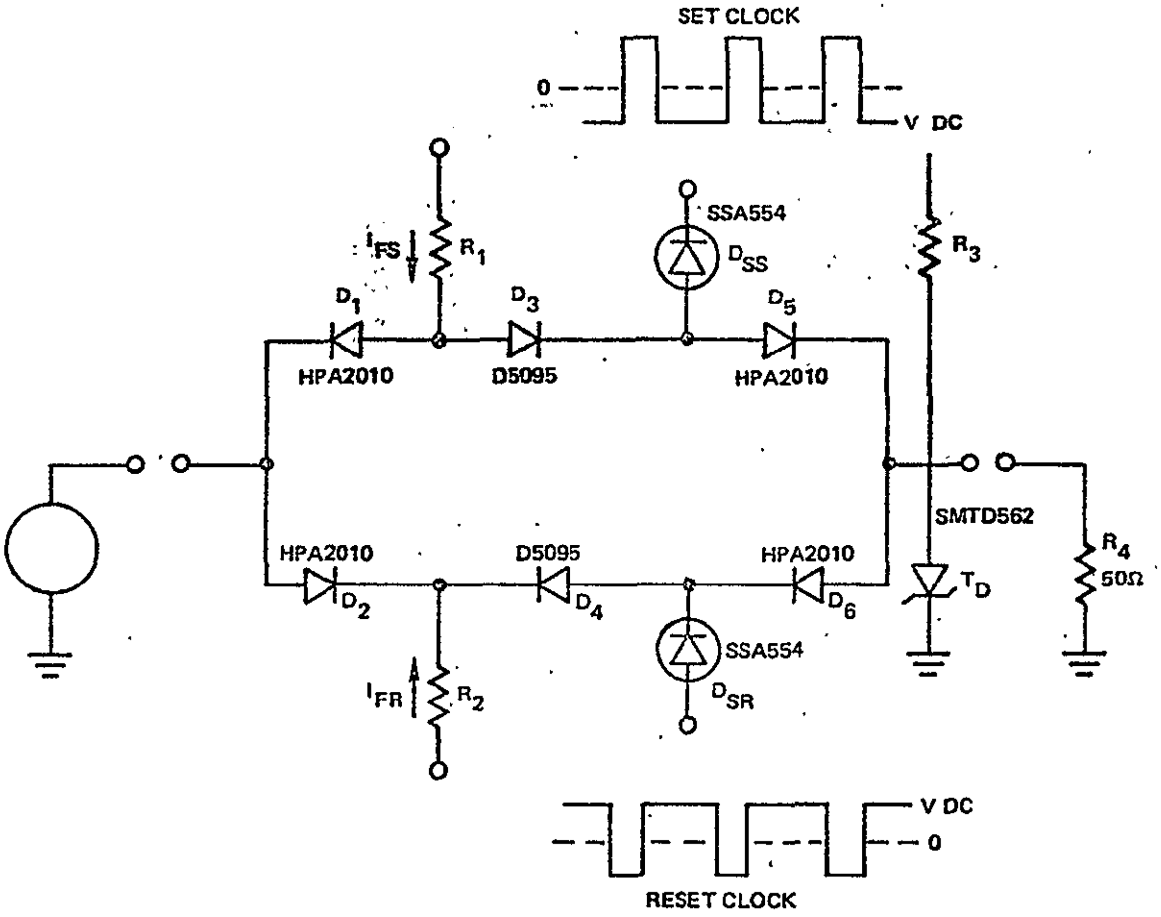

SET CLOCK
0
V DC
SSA554
D_{SS}
I_{FS}
R_1
R_3
D_1
D_3
D_5
HPA2010
D5095
HPA2010
SMTD562
R_4
50Ω
T_D
HPA2010
D5095
HPA2010
D_2
D_4
D_6
SSA554
D_{SR}
I_{FR}
R_2
V DC
0
RESET CLOCK

Fig. 4-6 Circuit Diagram of a One-gigahertz Shift Register Stage Due to B. Sear. Tunnel Diodes are used.

Diffuse threshold encoders and decoders can use ultra fast shift registers to perform at very high data rates.

The threshold decoder Fig. 4-7 has been built which can correct up to two random errors or an error burst of up to 8 bits received in a block of 11 bits. Rate one-half convolutional encoding is employed; that is, a parity bit is generated by the convolutional encoder and interleaved between each data bit.

The convolutional encoder-sequential decoder has been tested at the U.S. Army SATCOM Agency with a QPSK modem and has provided gains in error performance of from 4.3 to 5.2 dB over straight QPSK at a P_e of 10^{-5}. Coding gain is primarily a function of decoder memory length and search depth.

4.2.4 Cathode-ray-tube A/D Converters

The cathode-ray-tube (CRT) converter was yesterday's answer to high-speed encoding. Today, with high-speed integrated circuits and other high-speed semiconductor components, much higher conversion speeds are obtainable at smaller size, weight and cost.

Essentially, the CRT A/D converter is a special type of cathode ray tube, in which an aperture and a segmented output plate have been added to the conventional electron gun and deflection plates (Fig. 4-8). The openings in the aperture plate are arranged in several columns, with each column representing one binary digit. The length of the apertures is one increment high for the least-significant bit, 2^0, two increments high for 2^1, four for 2^2, etc. The openings in the aperture plate thus conform to a standard binary pattern.

With new advances in high speed scanning-beam technology, as for example, used in the fastest vidicons, the potentialities of the fast-scanned CRT A/D converter bear careful consideration as a candidate for space applications;

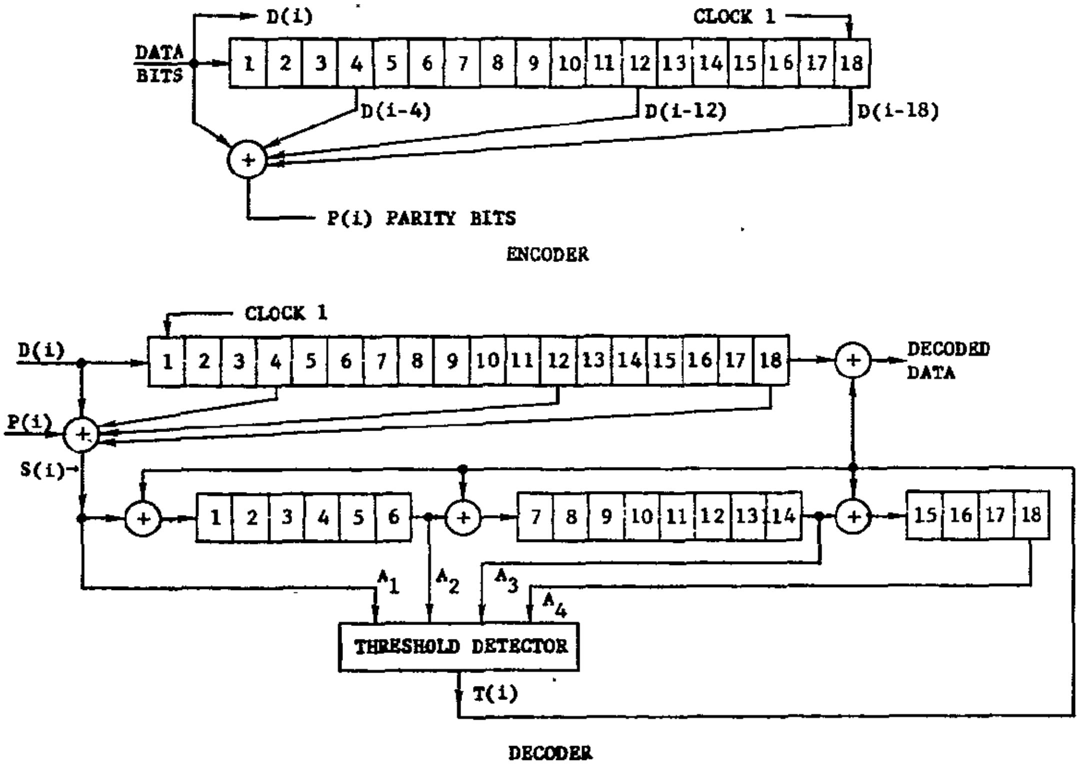

D(i)
CLOCK 1
DATA BITS
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
D(i-4)
D(i-12)
D(i-18)
+
P(i) PARITY BITS
ENCODER
CLOCK 1
D(i)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
DECODED DATA
P(i)
S(i)
1 2 3 4 5 6
7 8 9 10 11 12 13 14
15 16 17 18
A1
A2
A3
A4
THRESHOLD DETECTOR
T(i)
DECODER

Figure 4-7 DIFFUSE THRESHOLD DECODER/ENCODER (N = 2) INTERLEAVED TO ORDER 2

to take advantage of the inherent circuit simplicity involved, and to take advantage of the excellent reliability and long life being demonstrated in space by the vidicons of Tiros, Nimbus, and other sensor satellites.

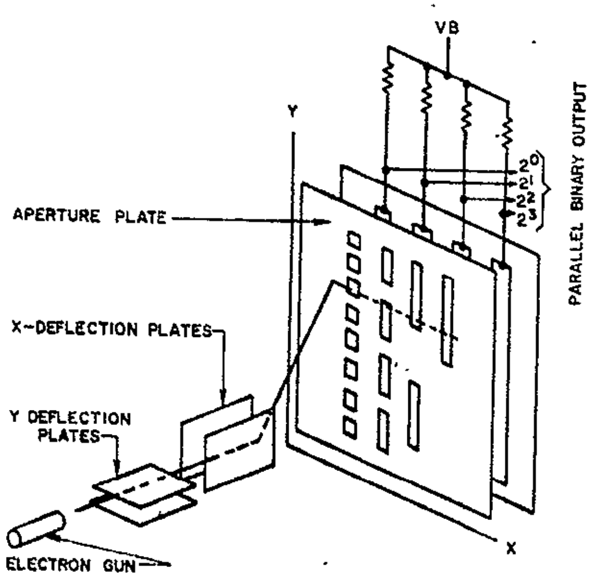

VB
Y
APERTURE PLATE
X-DEFLECTION PLATES
Y DEFLECTION PLATES
ELECTRON GUN
X
2^0
2^1
2^2
2^3
PARALLEL BINARY OUTPUT

Figure 4-8. Cathode-Ray Tube A/D Converter produces a digital output as the electron beam is moved across the aperture plate by the analog voltage

4.3 Redundancy-Removal Encoders

It is the nature of all practical sensors that their output voltage does not change from minimum to maximum response within one sample interval, provided, of course, that the sampling rate is equal to or faster than the Nyquist rate. Another way of saying this is that adjacent sample values are dependent. It, therefore, seems appropriate to take advantage of this dependence in devising an A-to-D conversion technique. Note that conventional PCM treats each sample value independently.

There are two basic methods for taking advantage of sample-to-sample dependence and thus removing redundancy from the coded data. The first method is commonly referred to as data compression. It involves digital processing of a PCM data stream according to some algorithm that uses past data samples to predict future sample values and transmits only that data necessary to update the prediction. There are many types of data compression ranging from relatively simple run length encoding to more complex polynomial prediction algorithms. Data compression techniques will not be explored further here because of the hardware complexity normally required and its doubtful value in adequately handling complex scene imagery.

The second basic method for redundancy removal is known as feedback encoding. Two well known types of feedback encoding are differential PCM* and delta modulation. Feedback encoding methods compare each sample value against some form of reference signal derived from previous samples. The resultant error voltage is quantized, digitally encoded and transmitted. Thus, only changes from sample to sample are transmitted.

Since only changes are transmitted the precision of quantization can be reduced. Thus, for the same sampling rate, two-bit precision in the feedback encoder will permit a 3-to-1 reduction in required data rate relative to 6-bit PCM for typical imagery-type sensors. Another advantage of feedback encoders is that they are not as dependent upon sensor dynamic range variations as are PCM encoders. When PCM is used, the quantization levels must be chosen to accommodate the maximum expected signal values, and accuracy and/or data rate are often wasted on seldom reached peak signal levels. Feedback encoding, on the other hand, transmits only changes* in signal level and is not constrained by maximum signal level.

---

*J. B. O'Neal, "Predictive Quantizing Systems for the Transmission of Television Signals," BSTJ, May-June, 1966.

Considerable work has been performed at Philco-Ford with the transmission of imagery using delta modulation. Laser scanner and vidicon signals have been handled using two-bit delta modulation. The parameters of the delta modulator were optimized for the transmission of high resolution aerial reconnaissance pictures with excellent results. Other sensors and other data applications may require a different feedback encoder configuration and optimization of parameters. For example, weather photographs require accurate edge response and fine gray scale resolution for the analysis of cloud motion and density.

A major advantage of delta modulation when compared to low precision PCM such as 3- or 4-bit PCM is that contouring** is not dominant as is the case for both 3- and 4-bit PCM. Pseudo-noise coding can be used with 3- or 4-bit PCM to dither the analog level and eliminate contouring. However, dither does not decrease the variance of quantizing noise.*** Four bit PCM with

---

*Typical rms prediction error σ_e in signal level when using a 2-sample linear predictor are from 10 log σ/σ_e = 7 to 12 dB below that of the input signal for typical scenes, and they have a probability density which is approximately Laplacian. This differential quantization provides an output signal-to-noise ratio 7 to 12 dB higher than a PCM quantizer using no feedback. The output signal-to-noise ratio of the n bit predictive quantizer is

$$\mathrm{SNR}_{out} = 2(2^n/3)^2 \, (\sigma/\sigma_e)^2$$

**Delta modulation noise is exhibited in 4 forms: grainy noise, slope overload, contouring, and edge busyness. The fact that quantizing error is split into 4 visually distinguishable categories suggests that a given rms quantizing noise is subjectively less objectionable than the same rms noise in PCM. See J. B. O'Neal, BSTJ, January, 1966.

***L. Schuchman, "Dither Signals and Their Effect on Quantizing Noise," IEEE Trans. Comm. Technology, December, 1964.

pseudo-noise at twice the bit rate is thus inferior or equal to 2-bit delta modulation when optimally matched to the input data.*

4.3.1 Introduction to Delta Modulation. Delta modulation is a data encoding technique for the imagery system which has the required high precision of five or six bit PCM but which makes use of the inherent redundancy or limited slew rate of the information. Delta modulation has been used by Philco-Ford in the "Compass Link" program and the "Quick-Look" demonstrations and has resulted in encoding efficiency three times better than six bit PCM producing output images which cannot be discriminated from images where sensor and recorder are connected back-to-back in the analog mode. In addition, delta modulation can be made to operate in noisy environments producing error rates as poor as 10^{-3} and can produce image information of useful content although of degraded quality.

Delta modulation is essentially a type of analog-to-digital conversion technique in which the changes in the analog signal input in succeeding sample intervals are transmitted. In the delta modulator, the instantaneous analog video input is compared at each digitizing sample time to determine whether a potential on an analog integrator (which holds a value representing the previous sample) should be incremented or decremented in order to reduce the difference. The increments or decrements applied are discrete and are determined by digitally quantizing the error difference with sometimes one-bit, or more usually, with two-bit precision.

These bits are then transmitted to the delta demodulator where they are utilized to produce a series of like charge increments (or decrements) which are applied to a like integrator. Since the same pattern of charge increments is applied to similar integrators in both the delta modulator and demodulator, the integrator signals track identically. By a suitable choice of parameters, the error between the integrator voltages and the current video input level can be kept small; thus, the video input is reproduced accurately in the delta demodulator.

---

*Ref. O'Neal, op cit, Fig. 7.

4.3.2 One-Bit Delta Modulation. The simplest delta modulator shown in Fig. 4-9, employs only one-bit quantizing of the error between the integrator and the analog input levels. Thus, a fixed quantity of potential is added or subtracted during each sample time. This quantity then represents a peak-to-peak hunting oscillation or noise. To keep this noise low,* the potential step change may be reduced to a small value, e.g., 1/100 of the maximum peak-to-peak analog level. However, the integrator would then require 100 and more samples to slew the entire dynamic range in response to a video step, which is impractially slow for most video systems. Thus, the one-bit delta modulator is not most useful for video systems where a practical limit on sampling rate exists; multi-bit operation is shown in Fig.4-10.

A very simple type of delta modulator is shown in Fig. 4-11. This circuit was producing video information at a 100 megabit rate in 1961.

---

*De Jager ("Delta Modulation," in Communication Theory, Butterworth Scientific Publication, 1953) and Van de Weg (Quantizing Noise in a Single Integration Delta Modulation System, Philips Research Report, 1953) have shown that the output SNR for a n bit delta modulator with a sinusoidal input of f_m sampling rate f_s, and output bandwidth f_o is

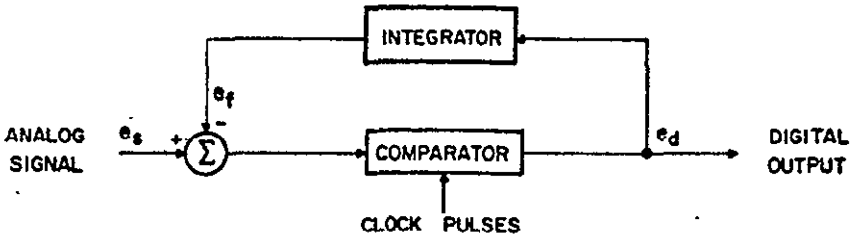
$$SNR = (2^n-1)^2 \ (0.04) \ \frac{f_s^{\ 3}}{f_m^{\ 2} f_o}$$

Thus increasing sampling rate by 2 increases output SNR by 9 dB, whereas increasing from 1 to 2 bit ΔM increases output SNR by 10 dB.

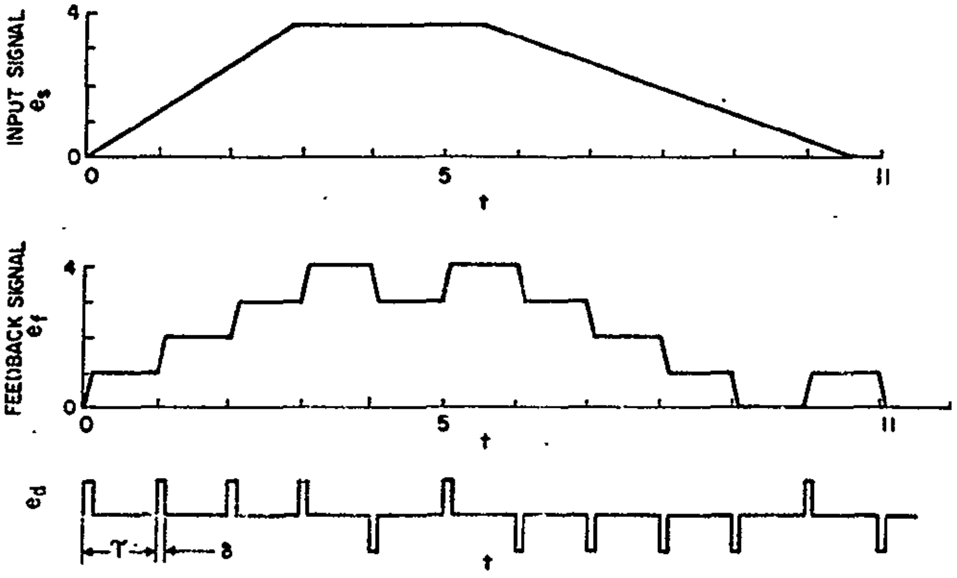
Band limited Gaussian noise of bandwidth f_m the corresponding ratio is

$$SNR = \frac{9(2^n-1)^2}{128\pi^2} \left(\frac{f_s}{f_m}\right)^3$$

behavior in a similar manner with respect to n bits/sample and sampling frequency f_s.

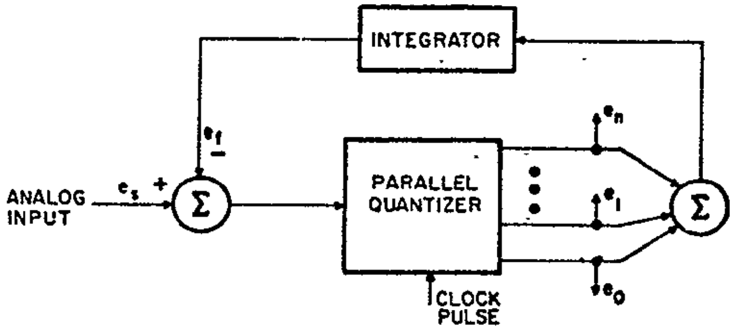

INTEGRATOR
e_f
−
ANALOG SIGNAL
e_s
+
Σ
COMPARATOR
e_d
DIGITAL OUTPUT
CLOCK PULSES

a. BLOCK DIAGRAM

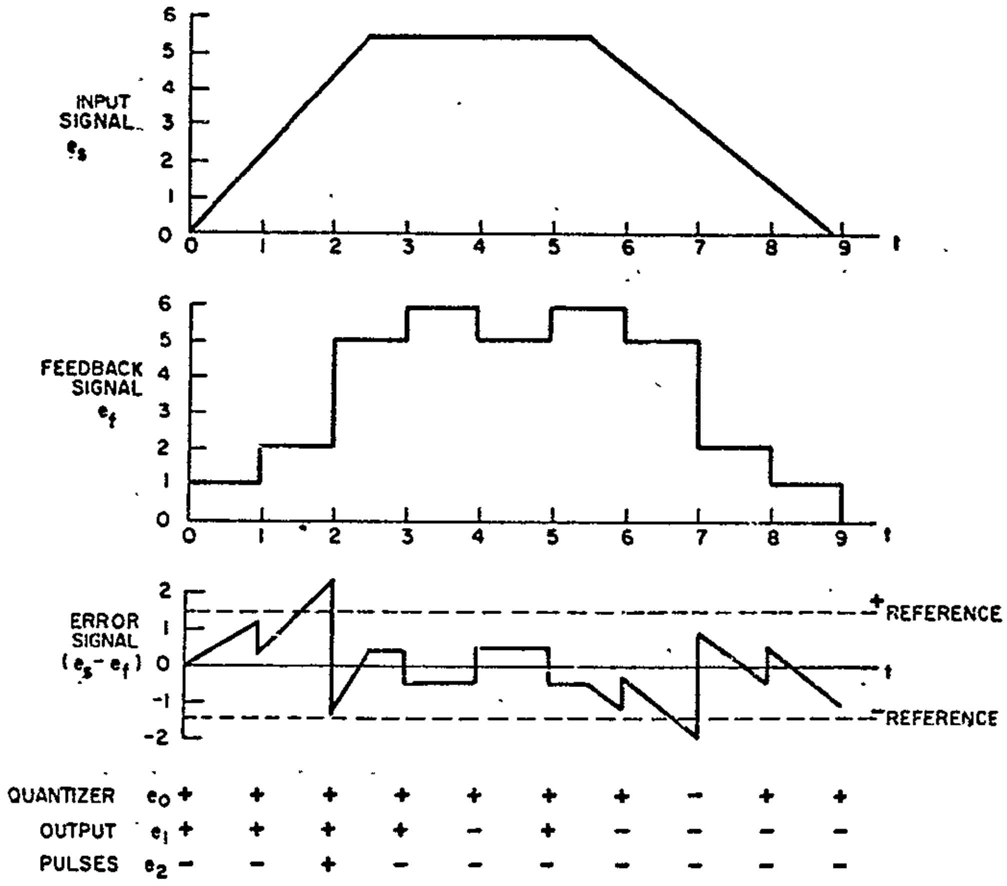


b. WAVEFORMS

FIGURE 4-9 ONE-BIT DELTA MODULATION SYSTEM


INTEGRATOR
e_f
−
ANALOG INPUT
e_s
+
Σ
PARALLEL QUANTIZER
CLOCK PULSE
e_n
e_1
e_0
Σ

a. BLOCK DIAGRAM



b. WAVEFORMS

Figure 4-10 MULTIBIT DELTA MODULATION SYSTEM

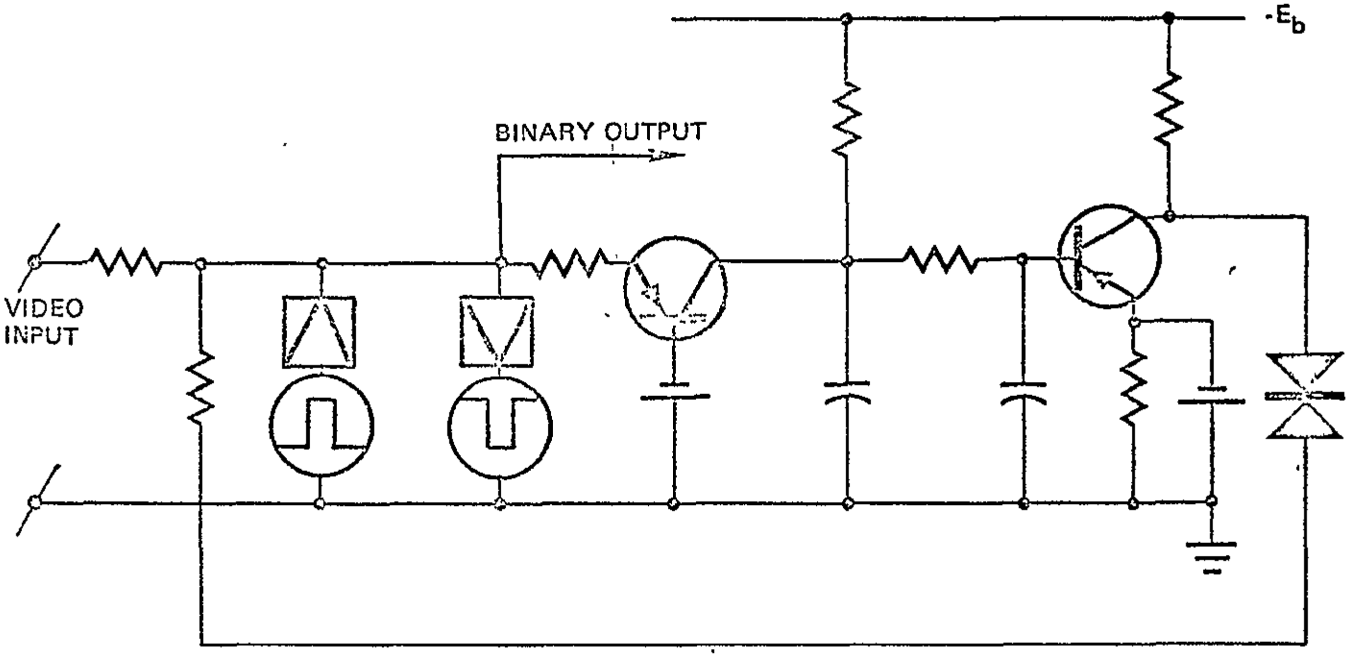

-E$_b$
BINARY OUTPUT
VIDEO
INPUT

Fig.4-11 Tunnel Diode Delta-Modulator for Video Information - after Balder and Kramer

4.4 THE ACCURACY OF DELTA MODULATION

It was pointed out that a single-bit delta modulator cannot meet the accuracy and transient requirements simultaneously. Therefore, two-bit, three-bit, and a four-bit delta modulator technique with exponentially increasing quantization levels were investigated by simulation with a digital computer.

The input signal used in the simulation is a ramp function which represents the output of the radiometer scanning a sharp black-to-white transition. The sampling rate used in the simulation is representative of actual link design. The sampling rate in the 2-bit delta modulator is twice as high as the 4-bit delta modulator to make the output bit rate constant.

The advantages of a 4-bit delta modulator over a 2-bit delta modulator will be clearly demonstrated as far as accuracy is concerned, with 3-bit delta modulator providing a good compromise between performance and complexity.

4.4.1 N-Bit Delta Modulator

In order to evaluate the response of an n-bit delta modulator to specific input functions, a computer model of the encoder was developed. The model shown in Figure 4-12 simulates 1, 2, 3 or 4-bit delta modulators with the capability of adjusting the several significant parameters in the loop.

The integrator has a time constant of N_1 samples, while G and A are the gains on the quantizer output R and the integrator output Z, respectively. It can be shown that only the product AG need be varied to obtain the complete set of responses to the system. Thus, we may normalize by setting A = 1 and vary just G.

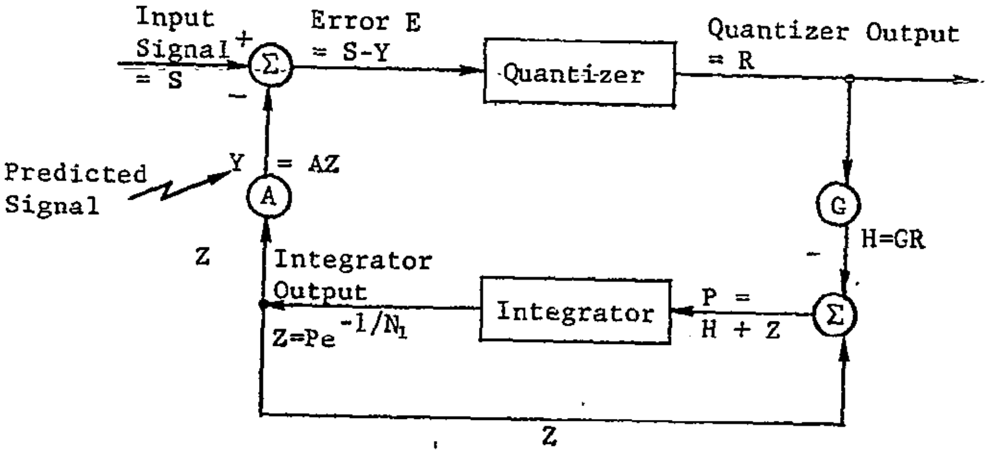

Input
Signal
= S
+
−
Σ
Error E
= S-Y
Quantizer
Quantizer Output
= R
Predicted
Signal
Y
= AZ
A
G
H=GR
−
Z
Integrator
Output
Z=Pe$^{-1/N_1}$
Integrator
P =
H + Z
Σ
Z

Figure 4-12 Delta Modulator Model

The quantizer is an M-bit exponential device, so that it has 2^M quantization levels. In this case, we shall consider M = 2 and 4. The quantity q is the minimum step size, which can be arbitrarily chosen.

The quantizer's exponential character here is in contrast to the quantizer with uniform step size, where all the steps are of the same size, for the ratio of each step to its predecessor is the quantity kept constant and is larger than unity.

The choice of quantization levels is constrained by several factors, the most critical being the choice of "break values"; i.e., the values of E for which R changes levels, and the stability requirements. In this simulation we have chosen to fix quantization levels and then pick those values which yield optimum response.

4.4.1.1 The 2-Bit Delta Modulator. Consider a 2-bit delta modulator having the characteristic shown in Figure 4-13. Assume an integrator time constant $N_1 = 10$, and that the higher quantization level (Q_2) is 8 times the lower level, with the break values at $E = \pm L(1)$. For some fixed gain G, the quantity L (1) can be varied over a suitable range of values and both RMS error and mean error can be calculated over an appropriate set of input step functions. From this calculation, we may choose that L(1) that minimizes the RMS or mean error and repeat the same process but with a variable G, and so forth, hopefully converging to some absolute minimum error.

The computer program which accomplished this calculation used two sets of computations, one using the minimum RMS error criterion, the other the minimum error criterion. The inputs are all step functions which equal 0 at t = 0 and thereafter equal S_o where S_o = 0, .25, .50, ..., 8.0. The results are shown in Table 4-2.

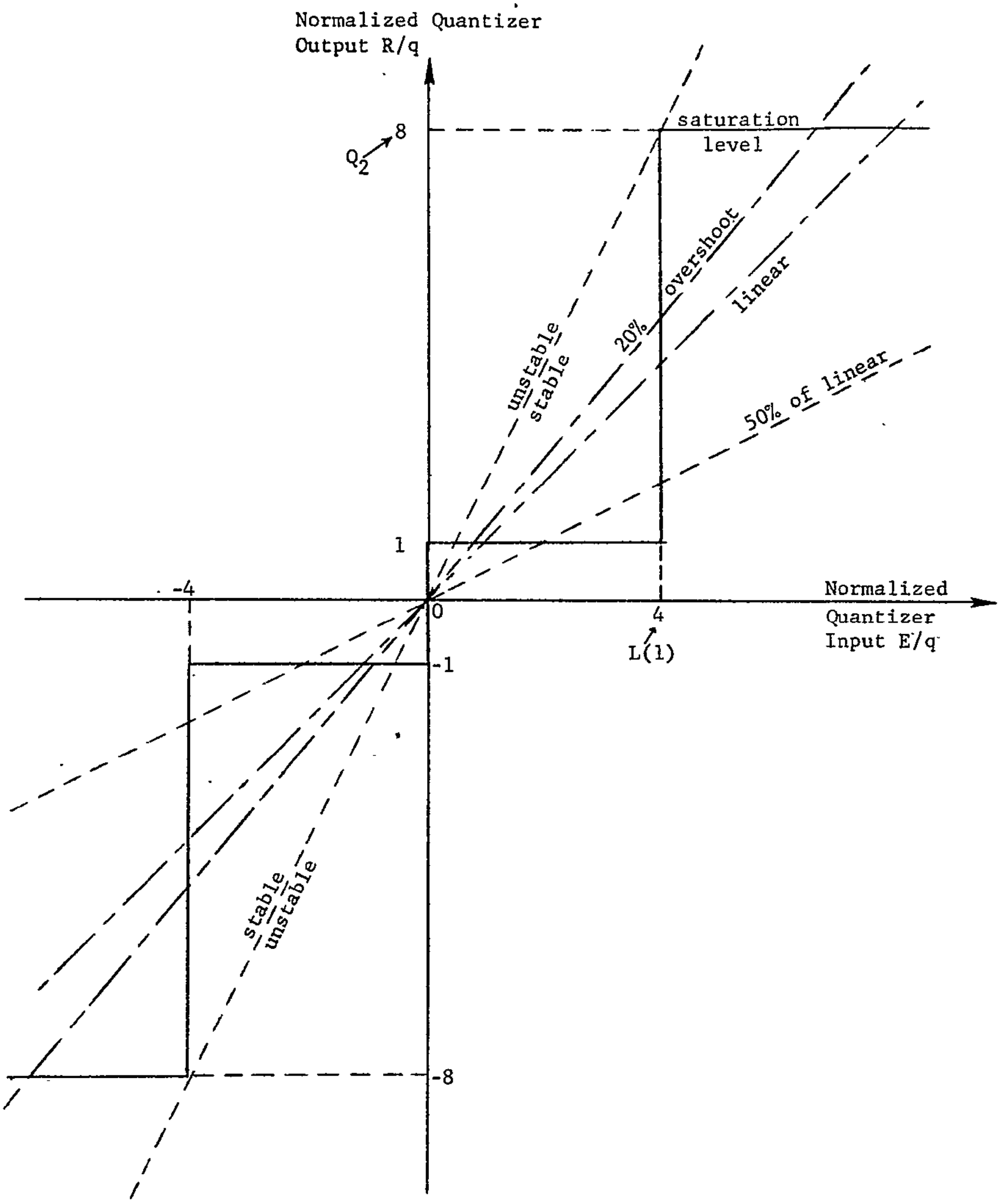


Figure 4-13 Typical 2-Bit Exponential Quantizer Characteristic (4 levels). An infinite time constant for the integrator is assumed.

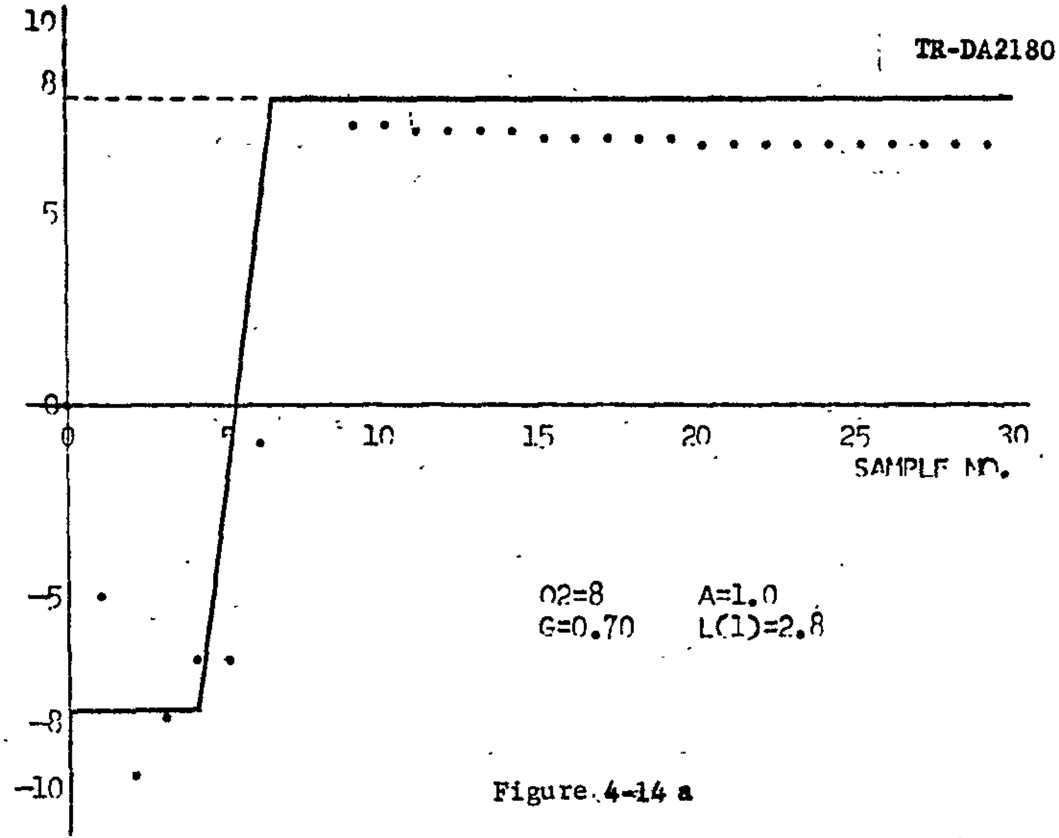


Figure 4-14 a

RESPONSE OF 2-BIT DELTA MODULATOR TO A RAMP FUNCTION

WIDTH OF RAMP PORTION = 2.5 X SAMPLING PERIOD
ABCISSA OF LOWER CORNER = 4.00

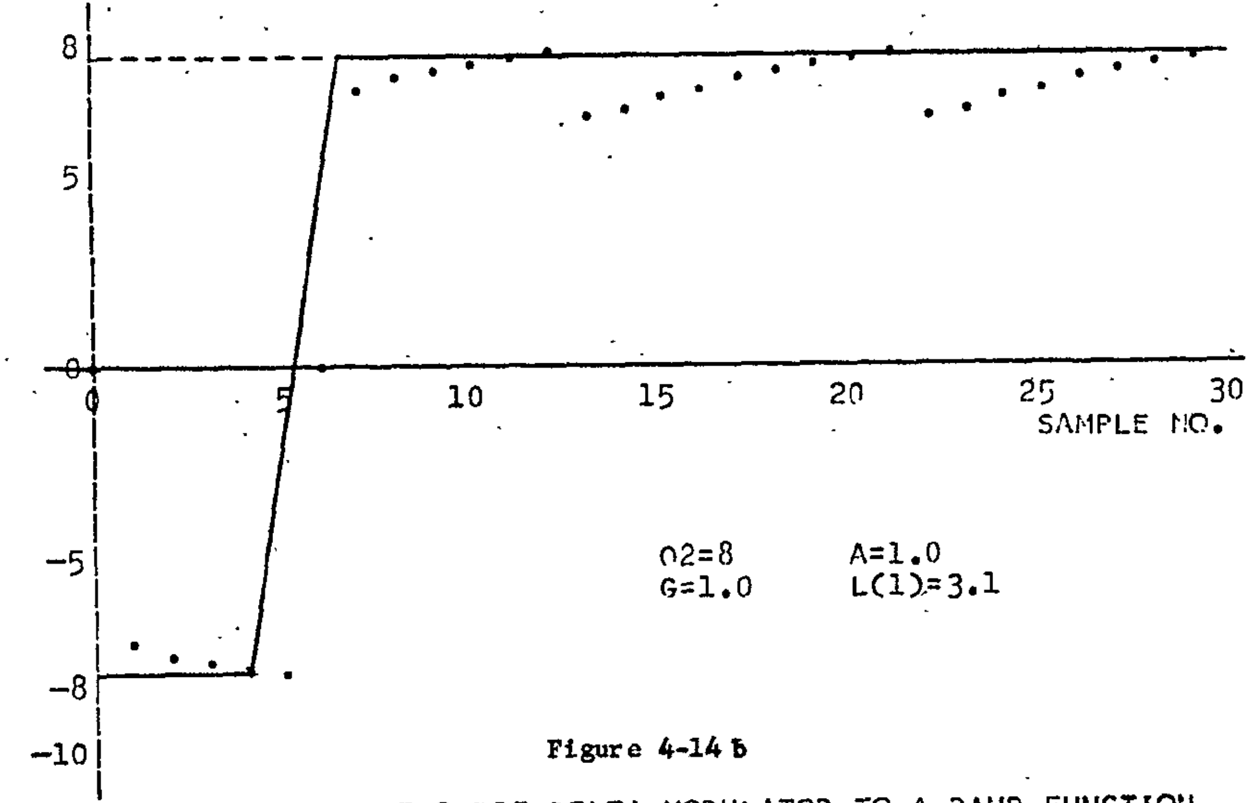


Figure 4-14 b

RESPONSE OF 2-BIT DELTA MODULATOR TO A RAMP FUNCTION

WIDTH OF RAMP PORTION = 2.5 X SAMPLING PERIOD
ABCISSA OF LOWER CORNER = 4.00

	L(1)	G	Minimum Error
Case 1 (minimum RMS error):	2.8	0.7	1.51 (RMS)
Case 2 (minimum mean error):	3.1	1.0	0.62 (mean)

TABLE 4-1 Optimized Values for L(1), G for 2-Bit Delta Modulator

It is observed that at least for the values computed, the two types of errors do not have their minima at the same parameter values.

The response of 2-bit delta modulators with these parameters to a ramp function are shown in Figure 4-14. It is clear that the response of Case 2 is better - the mean error being about half those in Case 1.

4.4.1.2 The 4-Bit Delta Modulator. As has been observed from other sample runs, when the integrator time constant $N \le 20$ the response to a typical step function is unacceptable, while $N_1 \ge 200$ yields an extremely good response. Such a large figure is unacceptable for purposes of smoothing out bits that are erroneously transmitted. In the 4-bit case, therefore, a compromise figure $N_1 = 50$ samples is used. Also assume a minimum step size $q = 1$.

With a quantizer characteristic as given in Figure 4-15, a program was constructed to count the number of times the quantized error R was = ± 1 for an entire set of step functions. These functions equal 0 initially and equal S_o thereafter, where $S_o = 128, -126, \ldots, 0.2, \ldots, 126, 128$, from the 2nd through the 41st samples. The maximum count is therefore 40 x 129 or 5,160. The result is shown in Table 4-2.

Gain G	No. Times R = 1
0.5	1,398
0.7	2,087
0.9	2,942
1.1	3,336
1.3	3,682
1.5	3,802
1.7	3,467
1.9	2,144

TABLE 4-2 4-Bit Delta Modulator Response for A=1, $N_1 = 50$

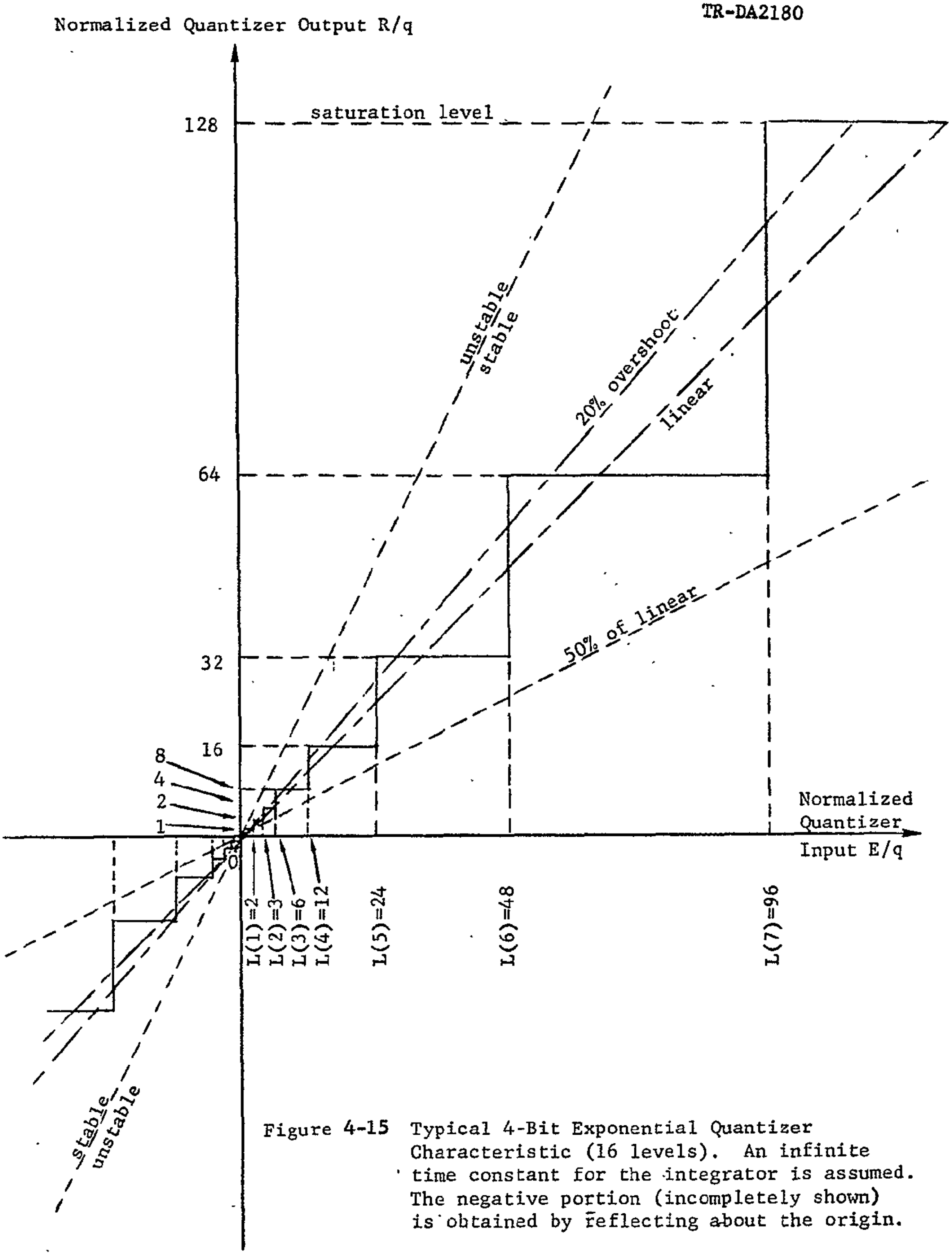


Figure 4-15 Typical 4-Bit Exponential Quantizer Characteristic (16 levels). An infinite time constant for the integrator is assumed. The negative portion (incompletely shown) is obtained by reflecting about the origin.

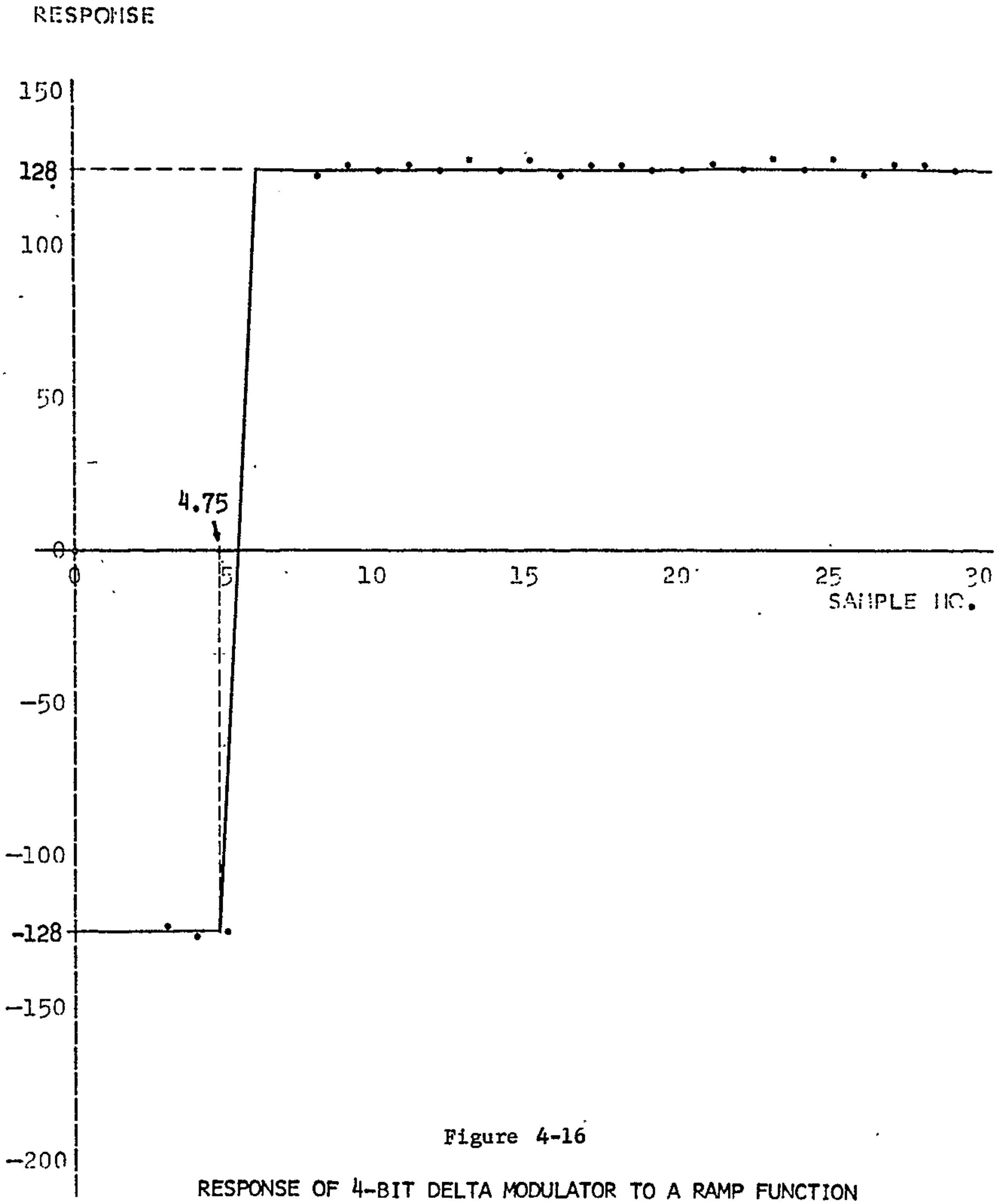


Figure 4-16

RESPONSE OF 4-BIT DELTA MODULATOR TO A RAMP FUNCTION
WIDTH OF RAMP PORTION = 1.25 X SAMPLING PERIOD

The best performance occurs at G = 1.5, which figure is used in an optimization of the quantizer characteristic. The result of this effort yielded the following values as an optimum set of break values with the levels.

2. 2.7, 9, 21, 40, 77, 126.

As a check, the number of times $|R| = 1$ was computed using this configuration, while G was varied from 1 through 2, inclusive. Again G seemed to be optimum at 1.5.

One response of the optimized system to a ramp of height 256 and a ramp portion with width 1.25 times the sampling period, is shown in Figure 4-16.

4.4.1.3 Optimization of 3-Bit Delta Modulator. 3-bit delta modulation is the best compromise between accuracy and circuit complexity, and is given the following detailed analysis:

A computer program ("DM3") was written to calculate the following quantities:

$$\text{AMNER} = 100 \sum_{S_o=1}^{32} \frac{1}{S_o/2} \sum_{j=6}^{40} E(j,S_o,B_1,B_2,B_3,G) \Big/ (35\text{x}32)$$

$$\text{RMS} = 100 \left[\sum_{S_o=1}^{32} \frac{1}{(S_o/2)^2} \sum_{j=1}^{5} E^2(j,S_o,B_1,B_2,B_3,G) \Big/ (5\text{x}32) \right]^{\frac{1}{2}}$$

where

$S_o/2$ = height of step function (=.5,1.0,1.5,...,16)

$E(j,S_o,B_1,B^2,B^3,G)$ = error of delta modulator in jth sample when input is step function of height $S_o/2$, with break values B_1,B_2,B_3, gain G.

RMS is a measure of how well the initial rise in the step functions is approximated.

AMNER is a measure of how well the "steady state" portion is approximated.

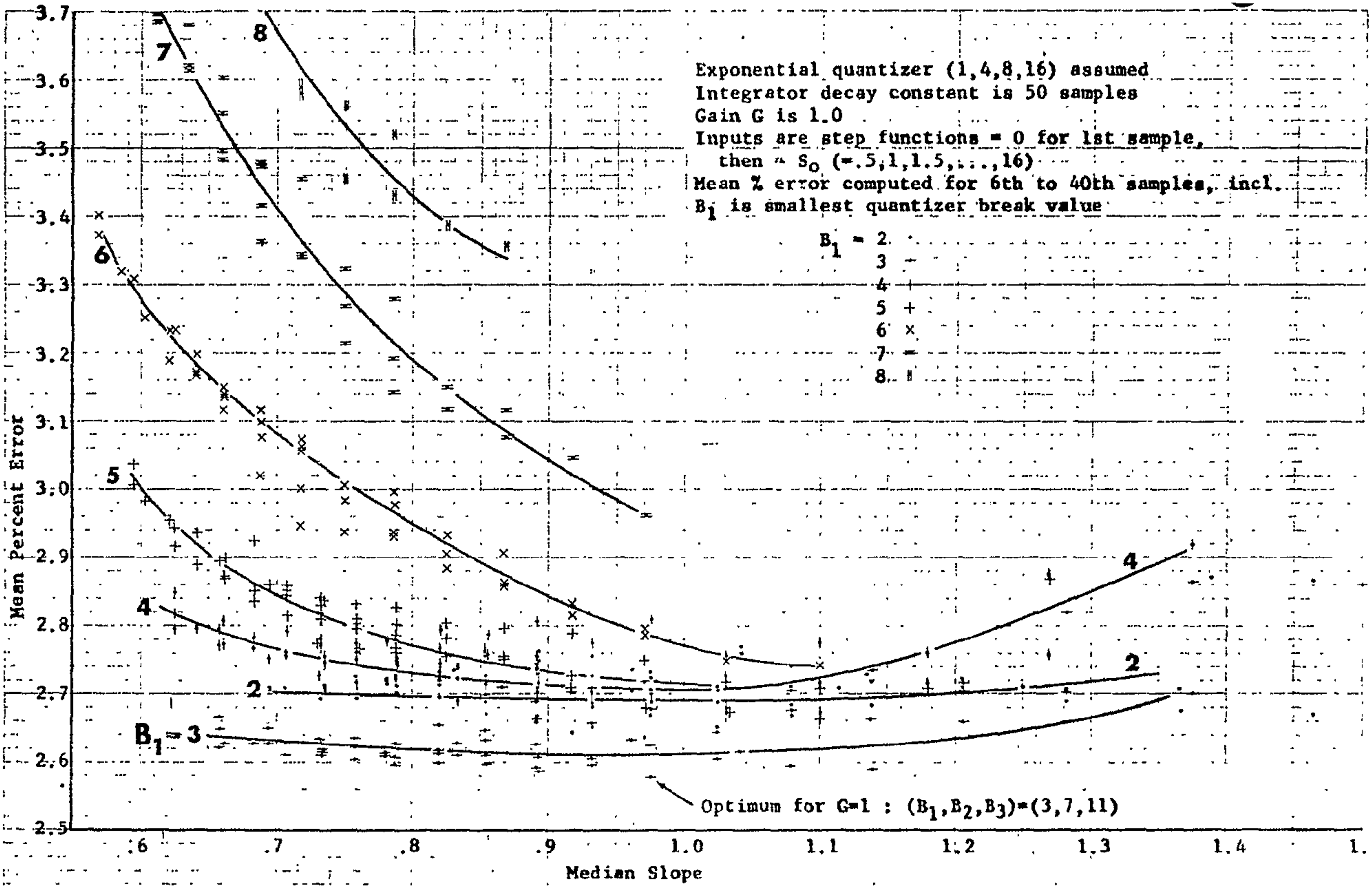
A quantizer characteristic with levels 1, 4, 8, 16 is assumed. The division by $S_o/2$ in the above sums is a normalization procedure, so that multiplication by 100 yields a percentage error type figure. Figures 4-17 and 4-18 display two interesting runs of DM3.

To assist visualization and categorization, the concept of the "median slope" of a quantizer characteristic is introduced. This quantity is roughly a measure of the steepness of the characteristic (the median slope of the uniform step quantizer characteristic is 1). For a typical characteristic, let ℓ be a straight line drawn through the origin, cutting the levels of the characteristic into segments. Then the median slope of the characteristic is defined as the slope of that ℓ for which

$$F_1+F_2+F_3= G_2+G_3+G_4.$$

Figure 4-17 is a plot of the most interesting part of the run for the case G=1.0 (no gain). In this run, each case for which the quantizer error in the 21st sample (for any S_o) is not ± 1 is defined as unstable and is discarded for optimization purposes. The quantities plotted are AMNER, the mean % error for the 32 step functions from the 6th to 40th samples, inclusive, and the median slope. In general the mean % error tends to increase with decreasing median slope and with an increasing value of B_1, the smallest break value, for $B_1 \gtrsim 3$. The minimum mean % error, 2.6%, was found for the case $B_1=3$, $B_2=7$, $B_3=11$.

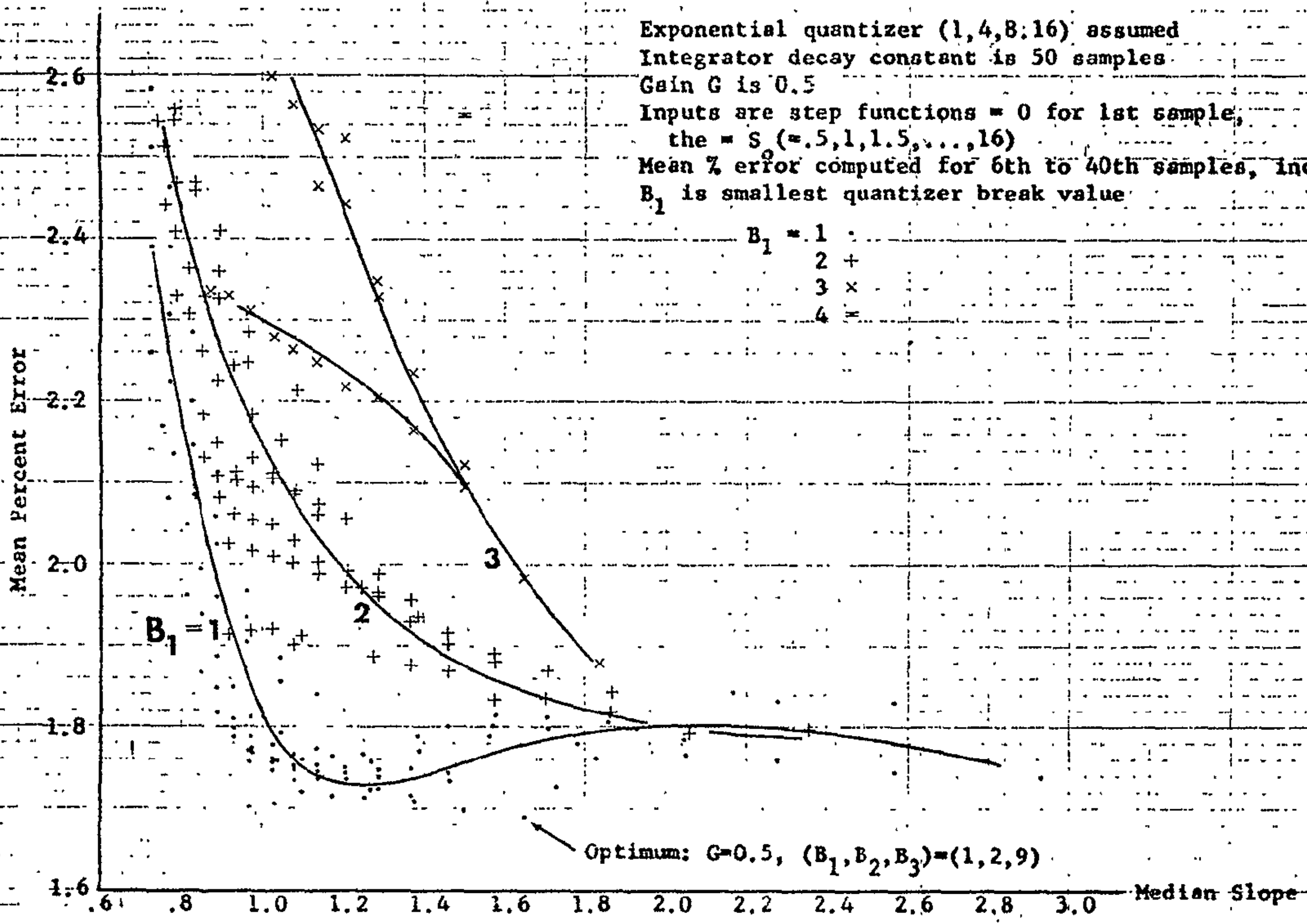
Figure 4-18 results from a much more extensive run, for G=0.1, 0.2, ..., 2.0 and the same break values as above. To concentrate on only the most interesting figures, the output was restricted to those cases for which AMNER $\leq$ 2.6%. Table 4-3 is a count of the total number of such cases for each value of G.



Exponential quantizer (1,4,8,16) assumed
Integrator decay constant is 50 samples
Gain G is 1.0
Inputs are step functions = 0 for 1st sample, then = S_0 (=.5,1,1.5,...,16)
Mean % error computed for 6th to 40th samples, incl.
B_1 is smallest quantizer break value

Optimum for G=1 : $(B_1,B_2,B_3)=(3,7,11)$

Fig. 4-17 Mean Percent Error of Output of 3-Bit Delta Modulator vs. Median Slope of Quantizer Characteristic.


Exponential quantizer (1,4,8,16) assumed
Integrator decay constant is 50 samples
Gain G is 0.5
Inputs are step functions = 0 for 1st sample,
the = S_o (=.5,1,1.5,...,16)
Mean % error computed for 6th to 40th samples, incl.
B_1 is smallest quantizer break value

Optimum: G=0.5, (B_1,B_2,B_3)=(1,2,9)

Fig. 4-18 Mean Percent Error of Output of 3-Bit Delta Modulator vs. Median Slope of Quantizer Characteristic.

G	Total No. Cases for which AMNER 2.6%
0.1	0
0.2	0
0.3	21
0.4	67
0.5	207
0.6	91
0.7	9
0.8	0
0.9	0
1.0	9
1.1	0
.	.
.	.
.	.
2.0	0

TABLE 4-3 Number of cases for which AMNER 2.6% vs. G.

4.5 QUANTIZING NOISE IN A/D CONVERSION

Quantizing is a nonlinear operation which takes an input signal $x(t)$ of arbitrary amplitude distribution and produces an output $y(t)$ whose possible amplitude is, at most, a countably infinite number of values. The quantizer is used in any system where an analog-to-digital conversion is required. If the input into the quantizer is considered to be an ergodic stochastic process, a knowledge of the statistics of the quantizer output can be used to minimize the loss of information inherent in the quantizing operation. B. Widrow[1] has shown that, if interest is in just first- and second-order statistics, there are some random functions which, when passed through a quantizer, do not lose first-and second-order statistical information. In such cases, the quantizer operation can be considered as a summing operation of the signal and an independent noise source.

Ergoticity is a property of a stochastic process which implies that the ensemble statistics of the process will become equal to the time averaged statistics of a single sample function of the stochastic process as the time-averaging period approaches infinity. This property allows the statistical analysis of ergodic stochastic processes to be carried out with respect to ensemble statistics so that the statistical behavior of a time series can be predicted. However, when the time-averaging period is finite, errors can occur since the equality between ensemble statistics and time-averaged statistics is only true as a limiting relationship. It, therefore, is of interest to see how much information is lost due to quantization when the time-averaging period is finite.

---

[1] B. Widrow, "Statistical Analysis of Amplitude-Quantized Sample Data Systems", Applications and Industry, AIEE; January, 1961.

It is possible to minimize the loss of statistical information due to quantization for an arbitrary information signal when the quantized signal is observed over a finite interval of time by adding a second signal (known as a dither signal) before the quantizer and subtracting the dither signal after the quantizer. The function of the dither signal is to change the statistical relationship between the information signal and the quantizer output so that the quantizer operation can be considered a summing operation of the input signal and a statistically known and independent noise source. However, the case of the use of a dither signal is not considered further here.

4.5.1 Quantizer Representation

In general, random signals, when passed through a quantizer, lose first- and second-order statistical information. This loss is a function of the mean, variance, and amplitude range for a number of random inputs.

Symbolically, a quantizer may be represented as in Figure 4-19. In Figure 4-20 a typical input-output quantizer characteristic is drawn for a uniform quantizer. There are k negative quantizing levels and a total of M quantizing levels (this includes the zero level). It is assumed that saturation cannot occur.

The quantizing noise, as a function of the signal x, has simply a saw-tooth characteristic as shown in Figure 4-21. (This assumes the input-output characteristic of Figure 4-20.)

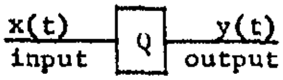


Fig.4-19 Block diagram of quantizer.

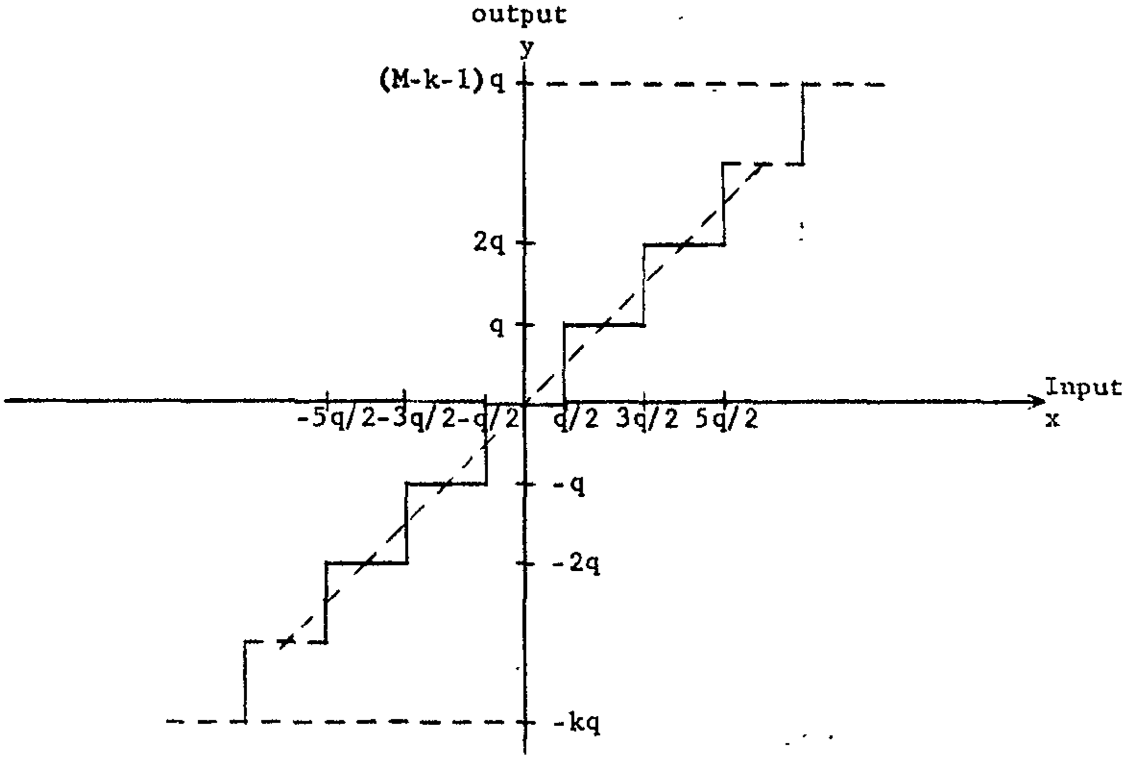


Fig.4-20 Quantizer input-output characteristic.

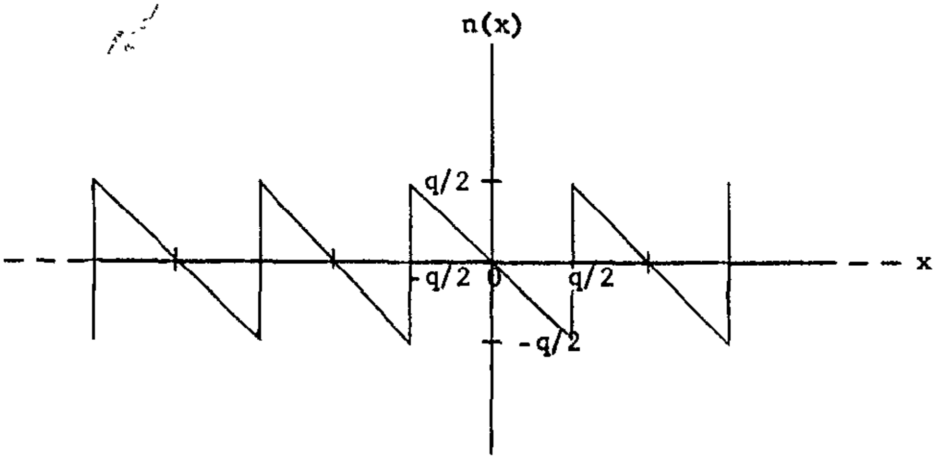


Fig.4-21 Quantizer noise as a function of the input.

Using Figure 4-20 as the quantizer model, the second-order statistics of the output of the quantizer can be described in the following manner.

$$E\left[y(t_1)y(t_2)\right] = E\left\{\left[n(t_1) + x(t_1)\right]\left[n(t_2) + x(t_2)\right]\right\}$$

Widrow has shown that, if x(t) has a characteristic density function $W_x(u)$ which satisfies the following criteria,

$$W_x(u) = 0 \qquad \text{for all } u > \frac{2\pi}{q}, \tag{1}$$

then

$$E\left[x(t_1)n(t_2)\right] = 0.$$

There are not many random signals that can satisfy this criteria, but Widrow has shown that this is a reasonable approximation for many important random signals.

To illustrate the validity of the approximation, an x_1 with a gaussian probability density function with zero mean yields

$$E\left[n(t_i)x(t_j)\right] = 0$$

if the variance of the gaussian distribution is greater than q.

Furthermore, for gaussian distributed input signals, the autocorrelation function of the noise is

$$E\left[n(t_i)n(t_j)\right] = \begin{cases} q^2/12 & t_i = t_j \\ 0 & t_i \neq t_j \end{cases}$$

even when the correlation coefficient of the input signal is moderately close to one.

There is a temptation at this point to consider the quantizer as a summation of an independent white noise signal and an information signal when interest is only in the first- and second-order statistics of the output.

In all physical problems, one deals with the observation of a random time function observed over a finite time interval T. The choice of statistics that adequately describe the physical situation is not always simple.

The equivalence of ensemble statistics and the time-averaged statistics for an ergodic process is strictly valid when the averaging interval T approaches infinity. If the time interval is finite and the process is ergodic, the use of such statistics may lead to erroneous conclusions.

We shall, therefore, treat the finite time average of $n(t_k)x(t_k)$ as a random variable and define it as

$$M(t_k) = \frac{1}{k} \sum_{i=1}^{k} n(t_i)x(t_i).$$

If the quantizer is to be accurately modeled by a linear operation, one requires at least the following:

$$E\, M(t_k) = 0 \tag{2}$$

$$E\left[M(t_k)^2\right] = \frac{E(x^2)\,\frac{q^2}{12}}{k} \tag{3}$$

where $\left\{E(x^2)q^2/12\right\}/k$ is the variance of the sample mean when $n(t_k)$ and $x(t_k)$ are truly independent and $n(t)$ has a white spectral density.

For purposes of comparison, the following random variable is defined. Let

$$G(t_k) = \frac{1}{k} \sum_{i=1}^{k} n_g(t_i)x(t_i)$$

where $n_g\,(t_k)$ and $x(t_k)$ are independent random variables. $n_g(t_k)$ is uniformly distributed with

$$|n_g| \leq \frac{q}{2}$$

and the samples of $n_g(t_k)$ are independent of one another.

We now wish to compare the statistics of $M(t_k)$ with those of $G(t_k)$. Table 4-4 summarizes results derived for an input signal which either satisfies (1) or has a gaussian probability density distribution. It can be seen that, if the input samples are independent, the linear model is valid. However, if $x(t)$ is maintained essentially constant over the averaging period, treating the quantizer as a summer leads to errors.

TABLE 4-4

Comparison of the first- and second-order statistics of $M(t_k)$ and $G(t_k)$ when the characteristic function of $x(t)$ satisfies (1) or has a gaussian shape.*

	$E[M(t_k)]$	$E[G(t_k)]$	$E[M^2(t_k)]$	$E[G^2(t_k)]$
1. Independent Samples	0	0	$\frac{\frac{q^2}{12} E(x^2)}{k}$	$\frac{\frac{q^2}{12} E(x^2)}{k}$
2. $x(t_k) = x$ for all k samples.	0	0	$\frac{q^2}{12} E(x^2)$	$\frac{\frac{q^2}{12} E(x^2)}{k}$

*If $x(t)$ is a gaussian random variable with mean m and variance σ, then the entries in the table are valid if,

$$\sigma > q$$

$$2\sigma^2 + m \ll e^{2\pi^2}$$

TABLE 4-5

Evaluation of $E\ M(t_k)$ when x is a uniformly distributed random variable with mean m and variance $\frac{\gamma^2}{3}$.

M				$E\ M(t_k)$
0	$\frac{\gamma}{2} q$	odd integer		$+ \frac{q^2}{12}$
0	$\frac{\gamma}{2} q$	even integer		$+ \frac{q^2}{24}$
m	φ	γ	δ	
$\varphi + \frac{1}{4}$	even positive integer	$\frac{\delta}{4} q$	positive int.	$-\left(\frac{\varphi + \frac{1}{2}}{4\delta - 1}\right) \frac{q^2}{4} + \frac{q^2}{48}$

TABLE 4-6

Evaluation of $E\left[M(t_k)\right]$ when $f(x) = \frac{\alpha}{2} e^{-\alpha(x-m)}$

m	α			$E\left[M(t_k)\right]$	
				less than	greater than
0	$\frac{2\pi}{q}$			$+\frac{0.22\ q^2}{\pi^2}$	$+\frac{0.21\ q^2}{\pi^3}$
0	$\frac{1}{q}$			$+0.164\ q^2$	$+0.149\ q^2$
m	φ	α	Ω	$E\left[M(t_k)\right]$	
$\varphi + \frac{1}{4}$	even integer	$\frac{1}{\Omega q}$	1	$\frac{q^2}{\pi}\left(\varphi + \frac{1}{4}\right)(0.025)$	
$\varphi + \frac{1}{4}$	even integer	$\frac{1}{\Omega q}$	>>1	$\frac{q^2}{\pi^3}\left(\frac{\varphi + \frac{1}{4}}{\Omega^3}\right)$	

Many signals of importance cannot be approximated as gaussian random variables. For such signals, a linear analysis of the quantizer, even when the samples are independent, may not be valid. Tables 4-5 and 4-6 summarize 4 results which illustrate the nonlinearity of the quantization operation for two density functions. When the probability density of the data is

$$f(x) = \begin{cases} \frac{1}{2}\gamma & m - \gamma < x < m + \gamma \\ 0 & \text{Otherwise} \end{cases}$$

Then $E\left[M(t_k)\right]$ varies as a function of the mean m and the base 2 , as listed in Table 4-5. There exist large values of γ which can make the characteristic function of f(x) come arbitrarily close to satisfying (1) and yet the value of $E\left[M(t_k)\right]$ does not approach zero.

In Table 4-6 the density function studied is taken as

$$f(x) = \frac{\alpha}{2} e^{-\alpha(x-m)}$$

Here, again, we have a random input into the quantizer whose mean and variance affect the value of E $M(t_k)$.

4.5.2 Linearization of the Quantizer and the Recovery of the Unquantized Signal Statistics

If the time-averaging period to observe a quantized random signal can be considered infinite, then conditions which allow the quantizer to be modeled by a summer fed by additive noise also allow the complete recovery of the signal statistical information. More specifically, an infinite level quantizer can be considered linear if the information signal is either a gaussian distributed random variable or has a characteristic function which has the bandlimited property.

If the time-averaging period to observe a quantized random signal cannot be considered infinite, then conditions which allow the quantizer to be considered linear minimize the loss of information due to quantization.

4.5.3 Linearization of the Quantizer and the Reconstruction of the Unquantized Information Signal. - A Heuristic Discussion

Quantization without the use of a dither signal is an operation that always produces the same noise value whenever the input signal is a given amplitude. Hence, whenever the input signal has some waveform structure, there is a corresponding waveform structure to the noise.

If the input is considered as a random signal represented by random samples, then strongly correlated samples observed in an interval T imply that the filtered and continuous waveform represented by these samples has a simple waveform structure. One would characterize such a waveform with words such as; smooth, slowly varying, and predictable. On the other hand, if the samples are independent, then the continuous waveform represented by these samples has a complex waveform and the resulting quantized output will have a complex waveform.

Section 5

THEORY OF QUADRIPHASE MODULATION

5.1 PHASE MODULATION TECHNIQUES FOR HIGH SPEED DIGITAL DATA

For digital bit stream data rates up to 5 Mbps, many modulation techniques are available, to provide both the modulated RF carrier which represents the bit stream and also demodulate the bit stream from the modulated carrier.

At the data rates of 40 to 400 Mbps which are being addressed in this study phase-shift-keying techniques conserve bandwidth and provide carrier reconstruction at the receiver which eliminates the need of coherence or the transmission of modulation-synchronization information.

This section will deal with the theory of quadriphase modulation and demodulation and include effects of phase unbalance, differential encoding and decoding and the problems associated with the reception of quadriphase modulated carriers by a coherent (phase-lock-oscillator) receiver in the presence of noise and phase ambiguity. In effect, this section provides a theoretical basis of Section 5 of Volume 1.

Special mention is made of contributions by Dr. D. T. Magill who developed this theory prior to this study in a Philco Ford Technical Memo #135.

A quadriphase modulator is essentially two biphase modulators in phase quadrature.

5.2 BASIC QUADRIPHASE COMMUNICATION SYSTEM THEORY

A quadriphase signal is a phase modulated signal such that it may have one of four possible phase positions (illustrated in Fig.5-1) for the duration of two input binary digits. Thus, $\log_2 4 = 2$ bits are carried by the carrier phase angle. A quadriphase signal may be written as

$$s(t) = \cos \{\omega_o t + e(t) \left(\frac{\pi}{2} \right) + \theta\}$$

where e(t) is either 0, 1, 2, or 3 (i.e., modulo 4) for the duration of two input binary symbols or one input quarternary symbol. Figure 5-2 presents a possible truth table relating e to two binary input streams, x and y, respectively.

5.2.1 Advantages of Quadriphase Signalling

Quadriphase signalling when detected by a coherent matched-filter (or correlation) detector in a gaussian noise environment yields the minimum theoretical binary error rate.* Thus, quadriphase offers excellent energy efficiency. Furthermore, quadriphase signalling has twice the bandwidth efficiency of ordinary biphase signalling, i.e., phase-reversal keying. This well-known fact may readily be established by considering a quadriphase signal as the resultant of two orthogonal biphase signals. Orthogonality is achieved since the sine and cosine functions are orthogonal to each other. Thus, in some sense quadriphase represents phase-division multiplexing (PDM) of two binary digit streams rather than the more conventional approaches of frequency-division multiplex (FDM) or time-division multiplex (TDM).

---

*See Section III.

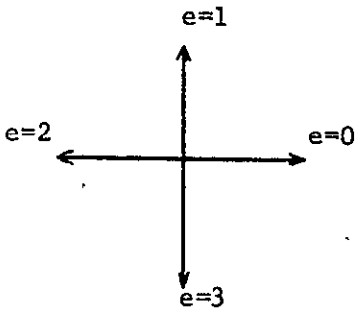


Fig.5-1 Quadri-phase phasor diagram

x \ y	0	1
0	0	3
1	1	2

Fig.5-2 Quarternary truth table for two inputs (binary)

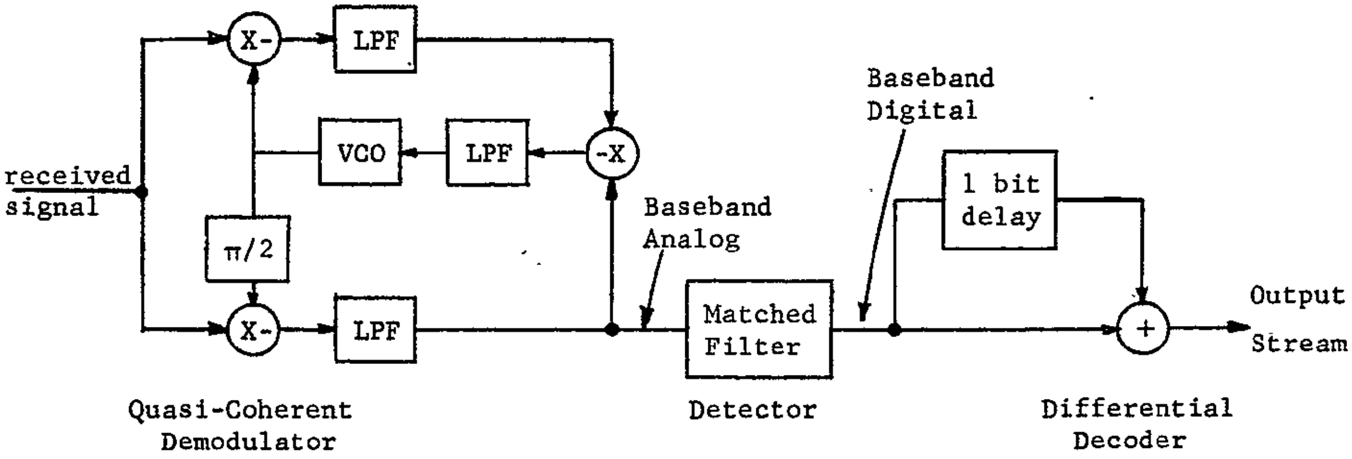

received signal
X-
LPF
VCO
LPF
-X
π/2
X-
LPF
Baseband Analog
Matched Filter
Baseband Digital
1 bit delay
+
Output Stream
Quasi-Coherent Demodulator
Detector
Differential Decoder

Fig.5-3 Quasi-coherent demodulator and differential decoder approach for bi-phase signals.

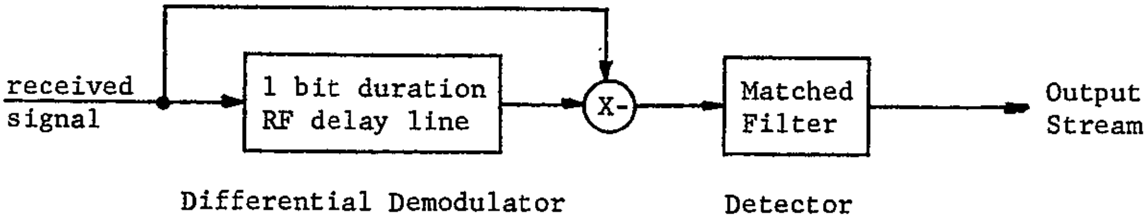

received signal
1 bit duration RF delay line
X-
Matched Filter
Output Stream
Differential Demodulator
Detector

Fig 5-4 Differential demodulation and detection approach for bi-phase signals.

5.2.2 Problems with Quadriphase

The major difficulty with quadriphase signalling is the resolution of phase ambiguities of $\pi/2$ radians (note that $4 \times \pi/2 = 2\pi$ radians). This situation is analogous to that of ordinary anti-podal biphase which has phase ambiguities of π radians (note that $2 \times \pi = 2\pi$). An alternate statement (equally valid) of quadriphase ambiguities is that there exists the ordinary sign ambiguity on both the inphase (I) and quadrature (Q) channels plus an ambiguity as to which is the I channel and which is the Q channel.

Conventional anti-podal biphase modulation avoids the sign ambiguity by employing binary differential encoding. Thus, it is desired to find the quarternary differential encoding analog.

5.2.3 Two Fundamental Approaches to Multi-Phase Detection

There are two fundamentally different approaches to multi-phase detection. The first, and superior approach (E/N_o) derives a quasi-coherent reference from the received signal, uses the quasi-coherent reference to demodulate (to baseband) the received signal, optimally detects the baseband signal(s), e.g., in matched filters, and finally logically decodes the resultant bit stream(s) to resolve ambiguities.

The second approach stores the past received analog signal (in a delay line or tuned filter) and uses it to differentially demodulate (to baseband) the present received signal and optimally detects the baseband signal(s). The ambiguities are resolved in the differential demodulation operation. Consequently, no logical decoding is required.

It is well known that for conventional channels and normal error rate requirements the first approach (quasi-coherent demodulation and differential decoding) is several dB superior to the second approach (differential demodulation and detection).

Figures 5-3 and 5-4 illustrate these two approaches for biphase modulation.

5.3 QUADRIPHASE MODULATORS - FUNCTIONAL DESCRIPTION

The quadriphase modulator represents an efficient method of transforming two serial bit streams into a phase-shift keyed millimeter wave carrier.

Before discussing the various circuit aspects of biphase and quadriphase modulators, consider first a detailed functional description of the quadriphase signalling system which will lead to a description of system degradation due to phase imbalance.

5.3.1 Functional Description of Quadriphase Modulator

The function of the transmitter is to generate the following signal

$$s(t) = \cos\left\{\omega_o t + d(t)\left(\frac{\pi}{2}\right) + e(t-1)\left(\frac{\pi}{2}\right) + \theta\right\}$$

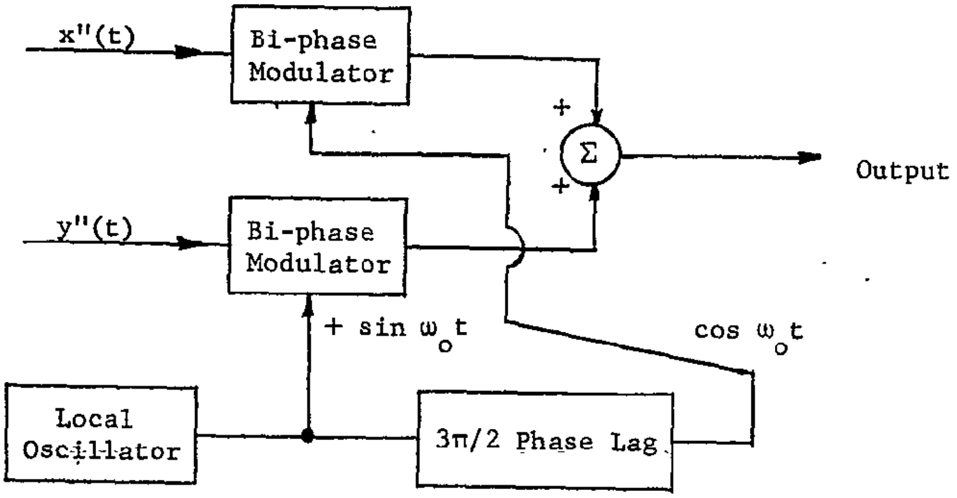
where d(t) is either 0, 1, 2 or 3 (i.e., modulo 4) for the duration of two binary input symbols, e(t-1) is the previous quarternary information symbol, ω_o is the carrier angular frequency, and θ is an arbitrary phase.

A quadriphase signal is obtained with the realization that

$$e(t) = d(t) + e(t-1).$$

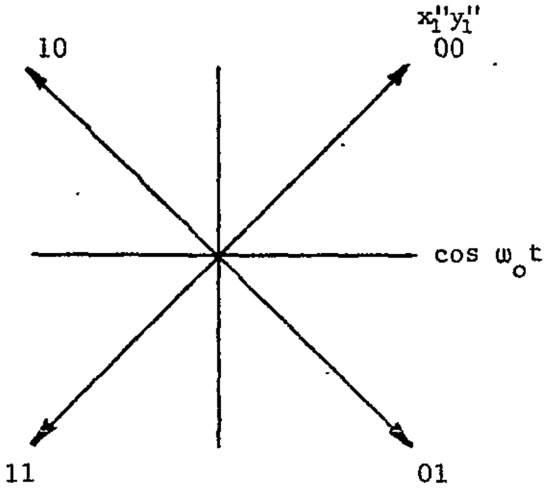
Thus, differential quarternary encoding is realized where d(t) represents the quarternary difference (two information bits). Figure 5-2 represents the desired truth table in terms of two input binary digit streams [x(t) and y(t)] and d(t) and the phase difference, $\varphi_d(t) = d(t)(\pi/2)$. A Gray code has been employed since this code minimizes the average bit error rate.

It is desired to realize the quadriphase modulator as shown in Fig. 5-5 since this approach is superior in the sense that it may be readily implemented at RF frequencies as well as at IF frequencies. Some alternate techniques, e.g., those employing counters, must be accomplished at relatively low, i.e., IF, frequencies. The variables x''(t) and y''(t) represent



x''(t)
Bi-phase Modulator
y''(t)
Bi-phase Modulator
+
Σ
+
Output
+ sin $\omega_o t$
cos $\omega_o t$
Local Oscillator
3π/2 Phase Lag

Fig.5-5 Block diagram of quadri-phase modulator.



$x_1''y_1''$
00
10
cos $\omega_o t$
11
01

Fig. 5-6 Phasor diagram for $\emptyset_e(t-1)$ as a function of variables x_1'' and y_1''.

appropriately encoded versions of x(t) and y(t), i.e., the differential encoder outputs.

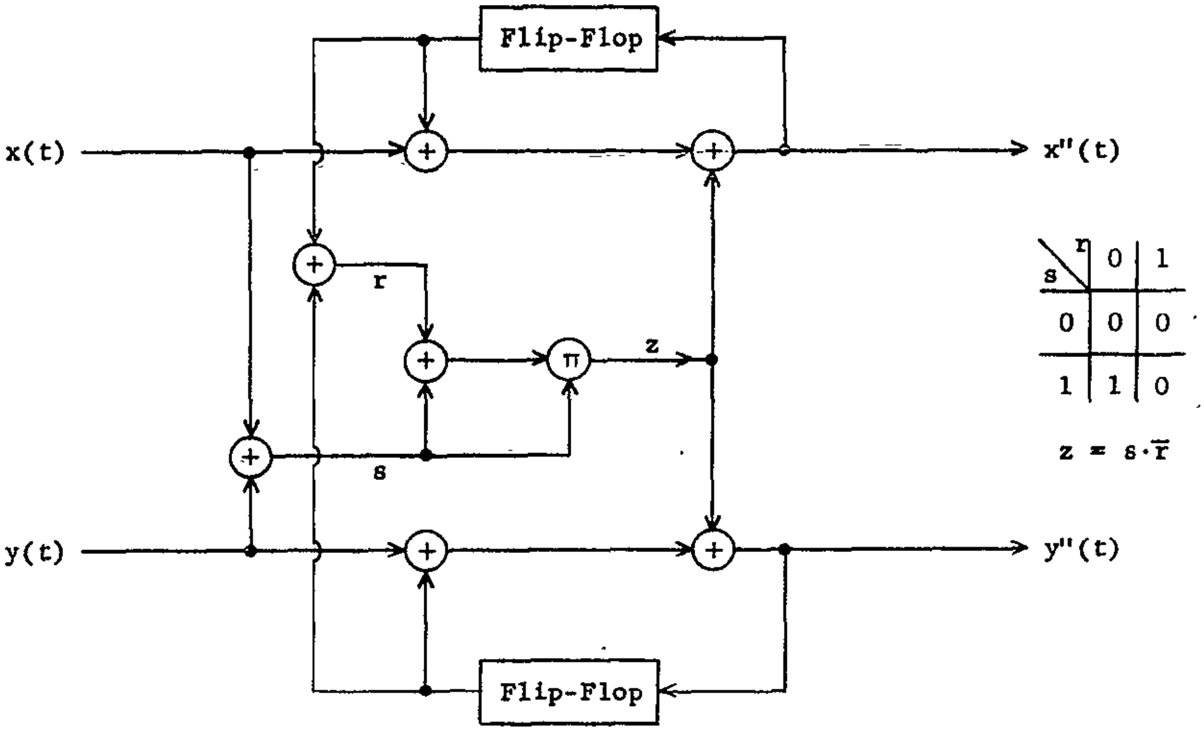
The purpose of the differential encoder is generate the appropriate binary variables $x''(t)$ and $y''(t)$ such that the proper phase difference $\phi_d(t)$ (given by truth table of Table 5-1) is created by the data streams x(t) and y(t).

The validity of this truth table may be established by considering the phasor diagram of Fig. 5-6 which illustrates the four possible phase positions for the preceding phase, $\varphi_e(t-1) = \varphi_{e1} = e(t-1)(\pi/2)$.

A method for implementing the truth table of Table 5-2 is now presented. An obvious approach is to implement the rule $x_2'' = x_1'' \oplus x_2$, $y_2'' = y_1'' \oplus y_2$ for the majority of the cases and the rule $x_2'' = \overline{x_1'' \oplus x_2}$, $y_2'' = \overline{y_1'' \oplus y_2}$ (where the superbar denotes complement) for the cases (four) marked by an asterisk. Figure 5-6 is a block diagram of the differential encoder that realizes the truth table. **The encoder is shown in Figure 5-7.**

TABLE 5-1 TRANSMITTER TRUTH TABLE

x	y	d	ϕ_d(rads)
0	0	0	0
0	1	1	$\pi/2$
1	1	2	π
1	0	3	$(3\pi)/2$


Flip-Flop
x(t)
x''(t)
r
z
s
y(t)
y''(t)
Flip-Flop

s \ r	0	1
0	0	0
1	1	0

$z = s \cdot \bar{r}$

Fig. 5-7 Block diagram of quarternary differential encoder. The symbol $\oplus$ denotes mod-2 addition and $\textcircled{\pi}$ denotes analog multiplication or gating (digital).

TABLE 5-2 TRUTH TABLE FOR DIFFERENTIAL ENCODER OUTPUT VARIABLES x_2'' and y_2''.
Note that those cases which deviate from the rule $x_2'' = x_1'' \oplus x_2$, $y_2'' = y_1'' \oplus y_2$ have been identified by an asterisk.

		$\phi_e(t-1)$ x_1'' y_1''	$\pi/4$	$-\pi/4$	$-(3\pi/4)$	$+(3\pi/4)$
$\phi_d(t)$	x_2	y_2	0 0	0 1	1 1	1 0
0	0	0	0 0	0 1	1 1	1 0
$\pi/2$	0	1	1 0*	0 0	0 1*	1 1
π	1	1	1 1	1 0	0 0	0 1
$3\pi/2$	1	0	0 1*	1 1	1 0*	0 0

5.3.2 Sensitivity Analysis

Four sources of steady-state quadriphase modulator loss are: 1) (I,Q) channel amplitude imbalance, 2) (1,0) amplitude imbalance in each channel, 3) (I,Q) channel phase imbalance, and 4) (1,0) phase imbalance. The losses stemming from each of these imbalances are evaluated below.

5.3.2.1 (I,Q) Channel Amplitude Imbalance. - Imbalanced amplitude between the two biphase channels can potentially create an unbalanced power loss and an AM-PM distortion loss.

5.3.2.1.1 Unbalanced Power Loss. - The average probability of bit error for the quadriphase signal is given by

$$P_e = \frac{1}{2} P_{ex} + \frac{1}{2} P_{ey}$$

where P_{ex} and P_{ey} are the error probabilities in the in-phase (I) and quadrature (Q) channels, respectively. For bit error rates on the order of $5 \cdot 10^{-7}$, a change in error rate by a factor of two corresponds to a 0.25 dB energy change.

The channel error probabilities are given by

$$P_{ex} = P_e\{(1 + \alpha)\ \rho_o\} = P_e(\rho)$$

where $\rho \overset{\Delta}{=} (1 + \alpha)\ \rho_o$

and $P_{ey} = P_e\{(1 - \alpha)\ \rho_o\} = P_e(\gamma\rho)$

where $\gamma = (1 + \alpha)/(1 - \alpha)$ is the channel power imbalance factor. For a 1-dB imbalance

$$P_e \approx \frac{1}{2} P_{ey} = \frac{1}{2} P_e\{(1 - \alpha)\ \rho_o\} = P_e\{\rho_o - 0.5 \text{ dB} + 0.25 \text{ dB}\}$$

$$= P_e\{\rho_o - 0.25 \text{ dB}\}$$

Thus, a 1-dB imbalance produces only a 0.25-dB degradation. Figure 5-8 plots the degradation as a function of the channel power imbalance factor.

5.3.2.1.2 (1,0) Pulse Amplitude Imbalance (Each Channel). - A (1,0) pulse amplitude imbalance causes the following losses: 1) average power loss in the lower level signal and 2) AM-PM conversion losses. The magnitudes of these degradations are estimated below.

1) Average Power Loss. - The average power loss due to (1,0) amplitude imbalance can be assessed from Fig. 5-9. Consider the probability of bit error in the in-phase channel

$$P_{ex} = \frac{1}{4} P_{e00} + \frac{1}{4} P_{e10} + \frac{1}{4} P_{e11} + \frac{1}{4} P_{e01}$$

where

$$P_{e00} = P_e\{1 \cdot \rho_o\}$$

$$P_{e10} = P_e\left\{\frac{2}{1 + \beta}\ \rho_o\right\}$$

$$P_{e11} = P_e\{1 \cdot \rho_o\}$$

$$P_{e01} = P_e\left\{\frac{2\beta}{1 + \beta}\ \rho_o\right\}$$

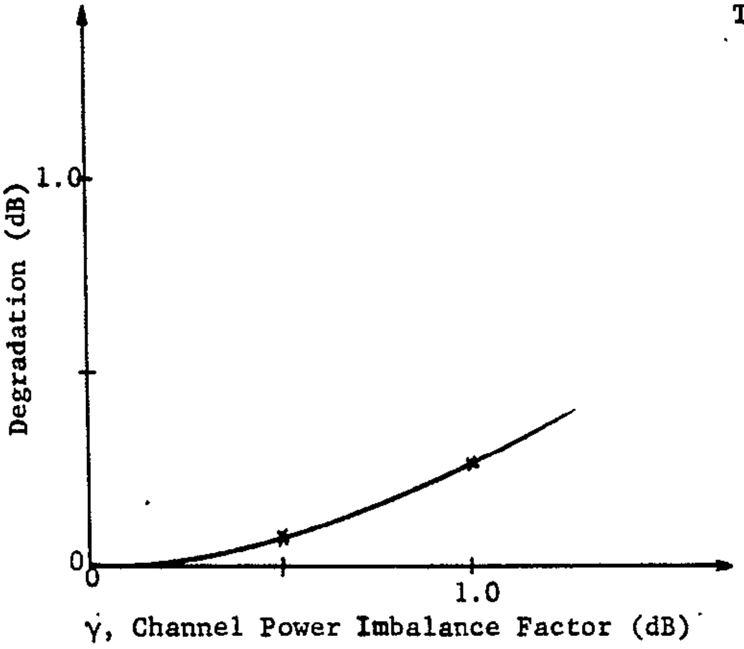


Fig. 5-8 Degradation as a function of channel power imbalance factor.

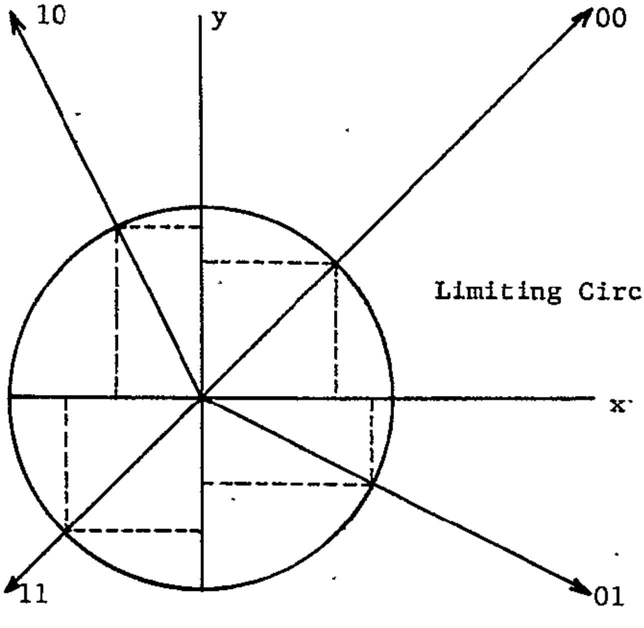


Fig. 5-9 Phasor diagram illustrating (exaggerated for convenience) effect of (1,0) pulse amplitude imbalance. Identical imbalances are assumed for both channels.

For the case of 1-dB imbalance, i.e., B = 1.26,

$$P_{ex} = P_e\left\{\frac{2}{1+\beta}\,\rho_o + 0.5\ \text{dB}\right\}$$

$$= P_e\{\rho_o - 0.53\ \text{dB} + 0.5\ \text{dB}\} = P_e\{\rho_o - 0.03\ \text{dB}\}$$

Thus, this portion of the loss due to (1,0) imbalance is negligible. Figure **5-10** plots this loss as a function of the (1,0) amplitude imbalance factor, β.

2) AM-PM Conversion Loss. - The effect of AM-PM conversion on system performance can be assessed best from the phasor diagram of Fig. **5-11** which also includes the (1,0) pulse amplitude loss effect. Identical imbalances are assumed for each channel.

00 Phase Position - In this position there is no phase distortion and, thus, no degradation.

$$P_{e00} = P_e\{1\cdot\rho_o\}$$

10 Phase Position - The degradation in this position is given by $[\sqrt{2}\,\sin\,(\theta_{10} + \varphi_{10})]^2$ for the x channel and $[\sqrt{2}\,\cos\,(\theta_{10} + \varphi_{10})]^2$ for the y channel where θ_{10} is the AM-PM phase shift (a function of drive level) and φ_{10} is the pulse amplitude imbalance phase shift.

$$\varphi_{10} = \tan^{-1}\left\{\frac{1}{\sqrt{\beta}}\right\}$$

Assuming $\beta = 1.26$ and an AM-PM conversion characteristic of 15^o/dB, one finds $\varphi_{10} = 41.7^o$ and $\theta_{10} = 7.1^o$. Thus, $D_{10x} = 2\cos^2 48.8 = 0.87$ and a degradation of 0.61 dB exists for the y channel.

11 Phase Position - The degradation in this position is given by $[\sqrt{2}\,\cos\,(\theta_{11} + \varphi_{11})]^2$ for the x channel and $[\sqrt{2}\,\sin\,(\theta_{11} + \varphi_{11})]^2$ for the y channel where θ_{11} is defined as before and $\varphi_{11} = 45^o$ for β - 1.26, $\theta = 15^o$. Thus, $D_{11x} = 2\cos^2 60^o = 0.5$ and a 3-dB degradation occurs. $D_{11y} = 2\sin^2 60^o = 1.5$ and an enhancement of 1.76 dB.

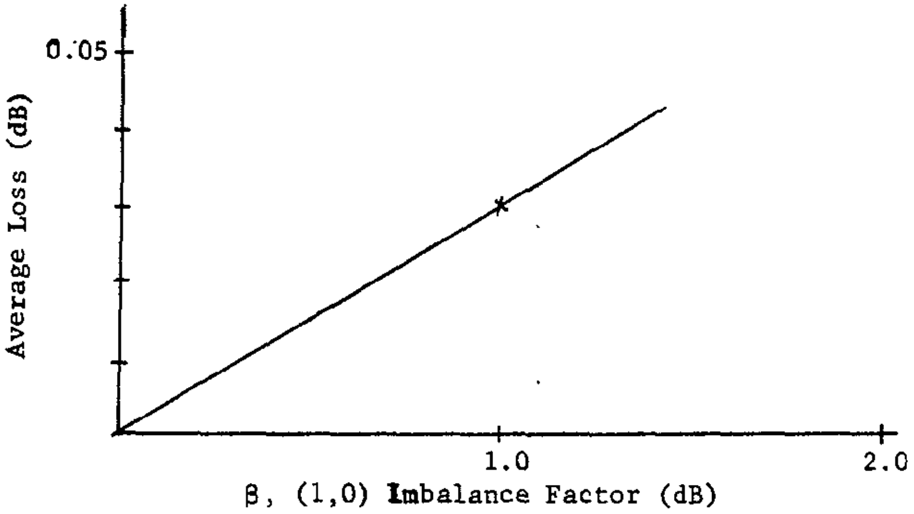


Fig. 5-10 Average power loss due to (1,0) pulse amplitude imbalance.

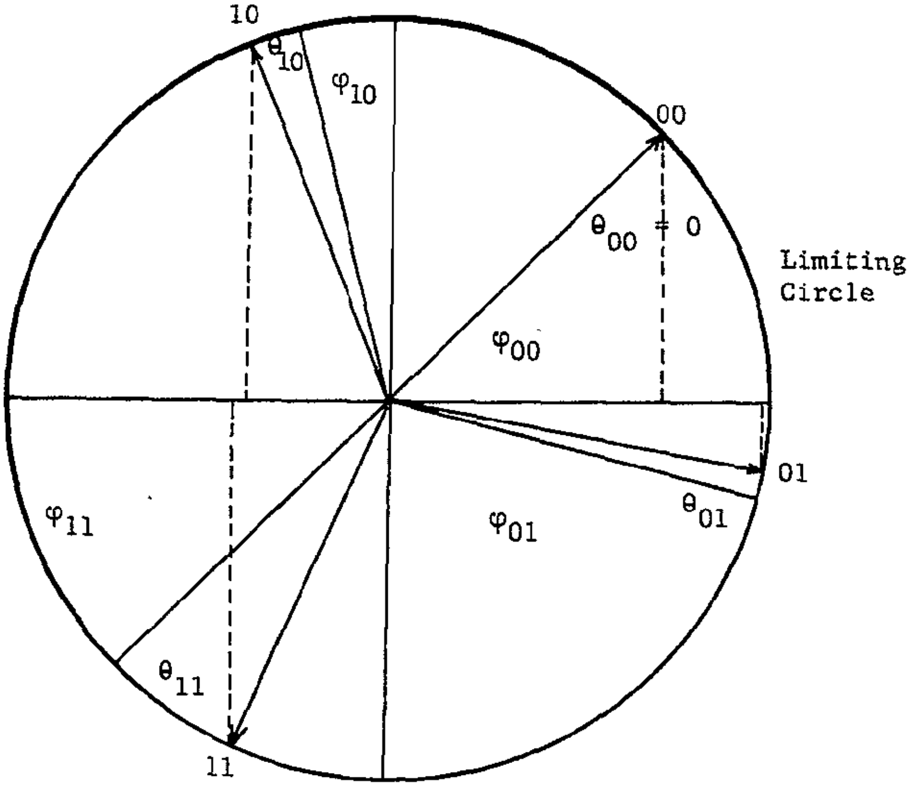

10
θ10
φ10
00
θ00 ≠ 0
Limiting Circle
φ00
01
θ01
φ01
φ11
θ11
11

Fig. 5-11 Phasor diagram illustrating (exaggerated for convenience) (1,0) imbalance effect with AM-PM conversion distortion shown.

01 Phase Position - The degradation in this position is given by $[\sqrt{2}\ \sin\ (\theta_{01} + \varphi_{01})]^2$ and $[\sqrt{2}\ \cos\ (\theta_{01} + \varphi_{01})]^2$ for x and y channels, respectively, where θ_{01} is defined as before and

$$\varphi_{01} = \tan^{-1} \{\sqrt{\beta}\}$$

Assuming $\beta = 1.26$ and 15^{o}/dB AM-PM characteristic, one finds $\varphi = 50.7$ and $\theta_{01} = 7.1$. Thus, $D_{01x} = 2\ \sin^2 57.8 = 1.43$ and an enhancement of 1.56 dB occurs. $D_{01y} = 2\ \cos^2 57.8 = 0.567$ and a degradation of 2.47 dB occurs. The resultant average probability of error is given by

$$\begin{aligned} P_e = \frac{2}{8} P_e\{\rho_o\} + \frac{1}{8} P_e\{\rho_o + 0.54\} + \frac{1}{8} P_e\{\rho_o - 0.61\} \\ + \frac{1}{8} P_e\{\rho_o - 3\} + \frac{1}{8} P_e\{\rho^o + 1.76\} \\ + \frac{1}{8} P_e\{\rho_o + 1.56\} + \frac{1}{8} P_e\{\rho_o - 2.47\} \end{aligned}$$

which can be approximated as

$$\begin{aligned} P_e &\approx P_e\{\rho_o - 3 + 0.75\} + P_e\{\rho_o - 2.47 + 0.75\} \\ &\approx P_e\{\rho_o - 2.3\} \end{aligned}$$

Figure 5-12 plots the degradation as a function of (1,0) pulse amplitude imbalance for a 15^{o}/dB AM-PM conversion coefficient. In practice, the phase-lock loop PCM demodulator adopts an average (compromise) phase reference position such that the dominant degradation terms, D_{11} and D_{01}, are significantly reduced. Curve (b) of Fig. 5-12 illustrates the performance when phase-lock loop compensation is properly accounted.

5.3.3 Pulse Width Asymmetry

The effect of pulse width asymmetry, i.e., different durations for ones and zeros, is to degrade performance as given by

$$P_e = \frac{1}{2} P_e\{(1 + \delta)\rho_o\} + \frac{1}{2} P_e\{(1 - \delta)\rho_o\}$$

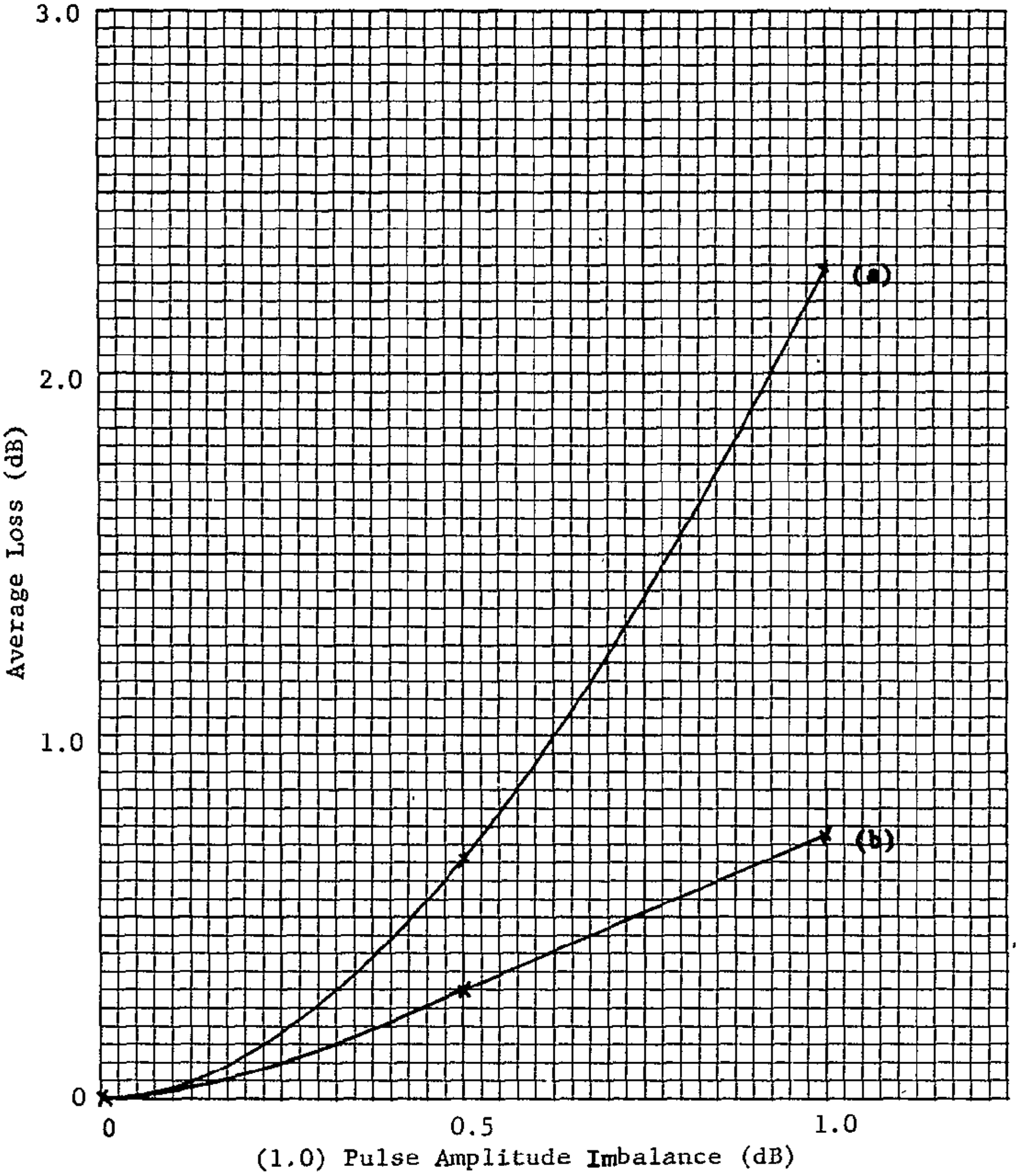


Fig. 5-12 Average total steady-state loss due to (1,0) pulse amplitude imbalance - includes average power loss, **AM-PM distortion loss, compensation. Curve (a) ignores the compensation provided by the receiver phase-lock loop. Curve (b) shows the performance when phase-lock loop compensation is properly accounted.**

For the case of $(1 + \delta)/(1 - \delta) = 1.5$

$$P_e \approx P_e\{\rho_o + 0.25 - 0.97\} = P_e\{\rho_o - 0.72\}$$

Figure **5-13** plots the average degradation as a function of asymmetry factor. It is to be noted that these results, while derived for rectangular pulses, also apply approximately to trapezoidal pulses (assuming symmetry about the zero-crossing axis).

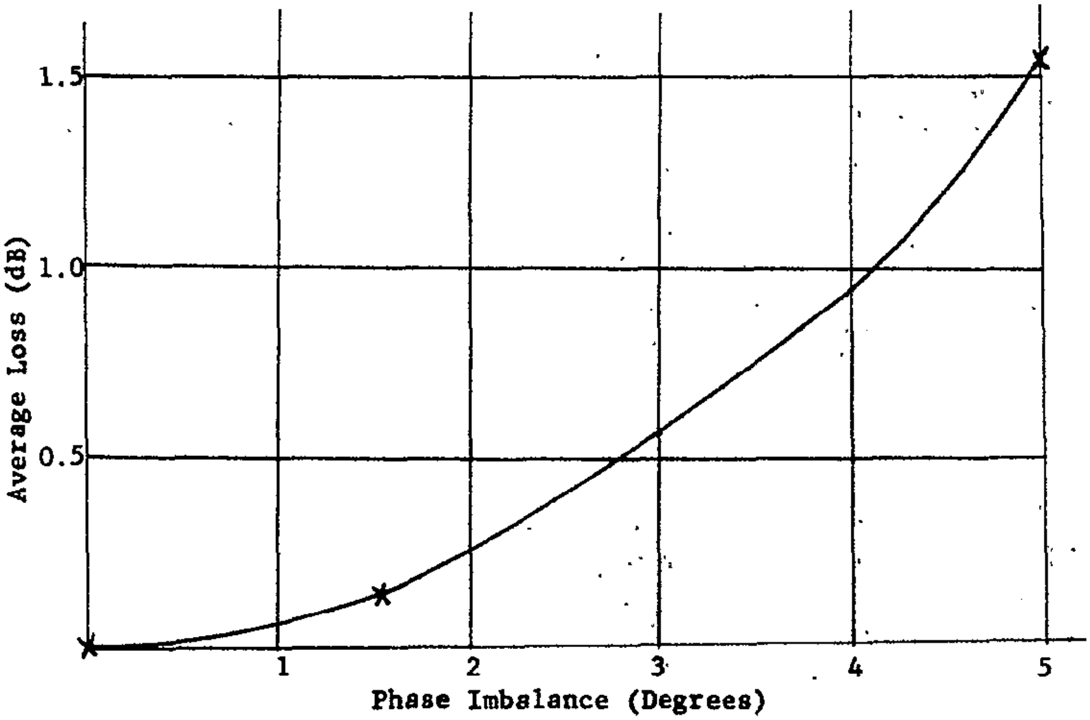


Fig. 5-13 Average loss due to (1,0) phase imbalance under worst-case assumptions.

5.4 COHERENT BPSK AND QPSK DEMODULATION AND DETECTION

The coherent BPSK and QPSK demodulator must perform many functions in order to recover a bit stream - or two bit streams - from a modulated carrier:

- generate or reconstruct a quasi-coherent "local oscillator" signal from the received modulated carrier.
- matched-filter detect.
- bit -synchronize
- logically decode.

These functions are discussed in the following paragraphs.

5.4.1 Functional Description of BPSK and QPSK Receivers

The function of the receiver is to generate a quasi-coherent reference, demodulate the carrier to baseband. The received signal is given by

$$r(t) = \cos\{\omega_1 t + e(t-\tau) \cdot (\pi/2) + \theta_r\}$$

where τ represents the propagation delay between transmitter and receiver, $\theta_r \not\equiv \theta$ is an unknown random phase angle which includes path delay and ω_1 is a potentially doppler shifted version of ω_o, the transmitted angular frequency. A diagram of the complete receiver is shown in Fig. 5-14.

5.4.1.1 Quadri-Phase Demodulator. - Several theoretically equivalent quadriphase demodulators can be used using different approaches to carrier reconstruction. The one shown in Fig.5-15 has been selected for its simplicity and is referred to as the x4 or fourth harmonic approach. A carrier is restored for tracking purposes since

$$\cos\{4\omega_1 t + 4\theta_r + 4e(t-\tau)(\pi/2)\} = \cos\{4\omega_1 t + 4\theta_r\}$$

This and other types of demodulators will be discussed later in this section.

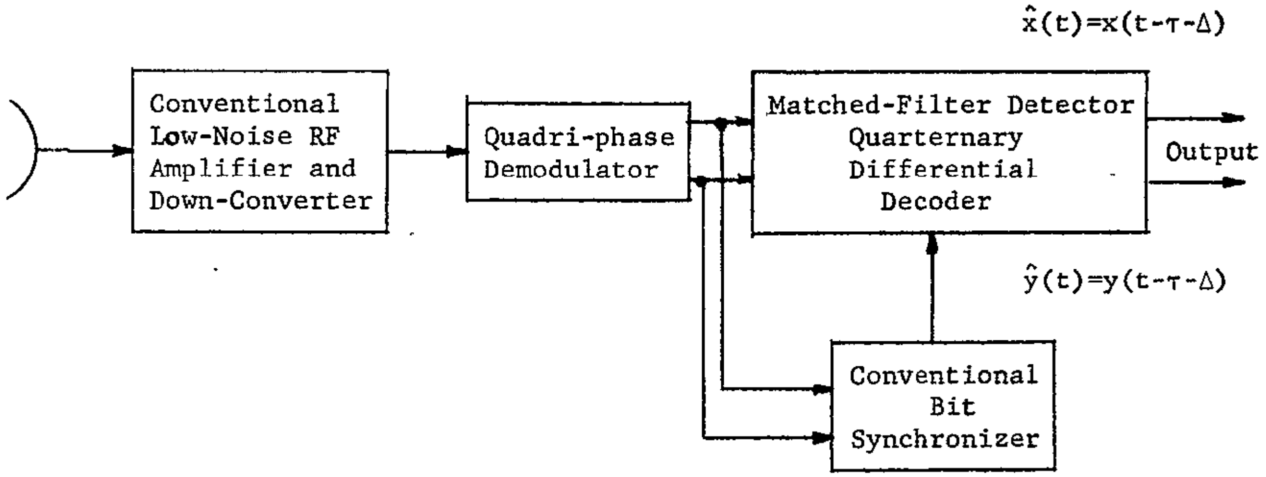

$\hat{x}(t)=x(t-\tau-\Delta)$
Conventional Low-Noise RF Amplifier and Down-Converter
Quadri-phase Demodulator
Matched-Filter Detector Quarternary Differential Decoder
Output
$\hat{y}(t)=y(t-\tau-\Delta)$
Conventional Bit Synchronizer

Fig. 5-14 Block diagram of receiver.

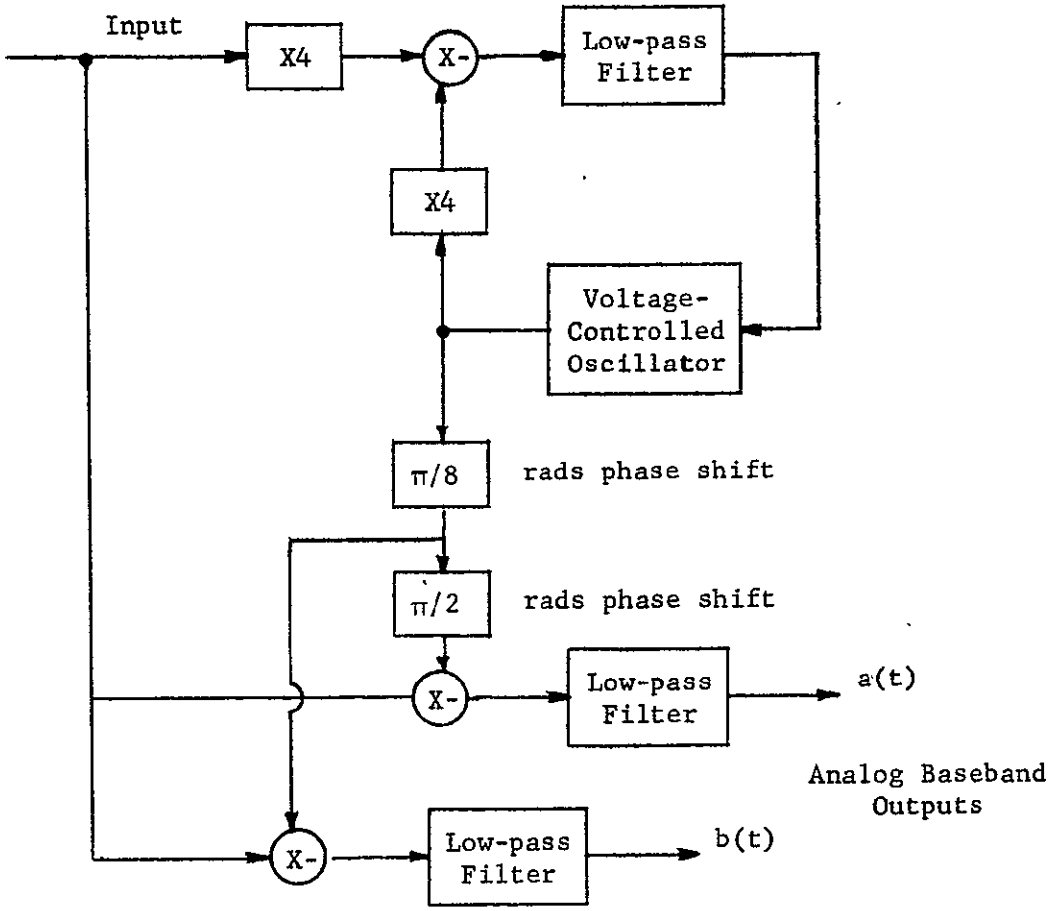

Input
X4
X-
Low-pass Filter
X4
Voltage-Controlled Oscillator
π/8
rads phase shift
π/2
rads phase shift
X-
Low-pass Filter
a(t)
Analog Baseband Outputs
X-
Low-pass Filter
b(t)

Fig. 5-15 Block diagram of one possible quadri-phase demodulator.

A vector decomposition of quadri-phase signals can be made into in-phase (I) and quadrature (Q) bi-phase signal channels. Each channel conveys one bit of information per quadri-phase symbol duration and possesses one-half the signal power. Since each channel is transmitting at one-half the total rate, identical performance to bi-phase signalling is achieved so long as the two channels do not produce mutual interference. They do not since the sine and cosine functions, corresponding to the Q and I channels, respectively, are mutually orthogonal. Thus, coherent detection of quadri-phase signals yields indentical average bit error rate performance to coherent detection of bi-phase signals.

Some intuitive arguments might mislead one into believing quadri-phase signalling inferior to bi-phase signalling. For example, quadri-phase possible vector positions are "closer" together, i.e., not anti-podal, than bi-phase possible vector positions. Thus, it might appear that quadri-phase decisions are more difficult for the same signal power and noise density. However, quadri-phase signalling has twice the symbol duration permitting a decision enhancement. It is an interesting fact that these two effects exactly cancel yielding indentical performance for bi-phase and quadri-phase systems in terms of average bit error rate.

5.4.1.2 Matched Filter Detector and Differential Decoder. - The function of the matched filters, i.e., finite-time integrators, samplers, and decision threshold (hard limiter or signum function) is to optimally detect the analog signals a(t) and b(t) and convert them to binary digital streams x'(t) and y'(t). Figure 5-16 is a block diagram of the matched-filter detectors. Note that this portion of the receiver represents conventional design - however, two channels are required in this application.

The function of the differential decoder is to unambiguously transform the binary digit streams x'(t) and y'(t) so that the decoder outputs $\hat{x}(t)$ and $\hat{y}(t)$ satisfy the relation

$$\hat{x}(t) = x(t - \tau - \Delta) \qquad \hat{y}(t) = y(t - \tau - \Delta)$$

where τ is the propagation delay and Δ is the individual bit duration.

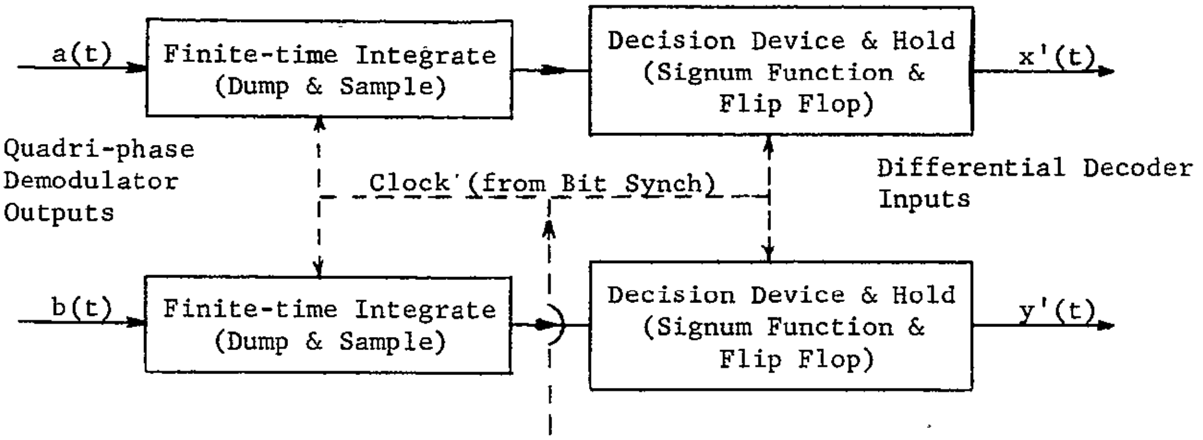

a(t)
Finite-time Integrate (Dump & Sample)
Decision Device & Hold (Signum Function & Flip Flop)
x'(t)
Quadri-phase Demodulator Outputs
Clock (from Bit Synch)
Differential Decoder Inputs
b(t)
Finite-time Integrate (Dump & Sample)
Decision Device & Hold (Signum Function & Flip Flop)
y'(t)

Fig. 5-16 Block diagram of matched-filter detectors.

The desired decoder truth table is now derived. The received signal is given by r(t) and the quasi-coherent reference signals by cos $\{\omega_1 + \theta_r + \alpha - (\pi/4)\}$ and sin $\{\omega_1 t + \theta_r + \alpha - (\pi/4)\}$ where $\alpha = n \cdot (\pi/2)$ for $n = 0, 1, 2, 3, \ldots$, i.e., α denotes the potential $\pi/2$ radian ambiguity. Thus, the analog outputs of the quadri-phase demodulator are given by

$$a(t) = \cos \{e(t-\tau)(\pi/2) - \alpha + (\pi/4)\}$$

$$b(t) = \sin \{e(t-\tau)(\pi/2) - \alpha + (\pi/4)\}.$$

The sampled analog outputs of the finite-time integrators are also given by a(t) and b(t) since $e(t - \tau)$ remains fixed over the period of the finite-time integration. Note the operation of the finite-time integration is defined by

$$c_o(\Delta) = \frac{1}{\Delta} \int_0^{\Delta} c_i(t) \quad dt$$

where c_o and c_i are the output and input variables, respectively. Thus, the hard-bit decisions are given by

$$x'(t) = \mathrm{sgn}\,\{\cos [e(t-\tau)(\pi/2) - \alpha + (\pi/4)]\}$$

$$y'(t) = \mathrm{sgn}\,\{\sin [e(t-\tau)(\pi/2) - \alpha + (\pi/4)]\}$$

where sgn $\{\cdot\}$ denotes the signum or hard-limiter function. Similarly,

$$x'(t-\Delta) = \mathrm{sgn}\ \{\cos\ [e(t-\tau-\Delta)(\pi/2) - \alpha + (\pi/4)]\}$$

$$y'(t-\Delta) = -\mathrm{sgn}\ \{\sin\ [e(t-\tau-\Delta)(\pi/2) - \alpha + (\pi/4)]\}$$

Defining

$$\psi(t-\Delta) = e\ (t-T-\Delta)\ (\pi/2) - \alpha + (\pi/4)$$

one finds that $x'(t)$ and $y'(t)$ may be expanded using trigonometric identities to yield

$$x'(t) = \mathrm{sgn}\ \{\cos\ [d(t-\tau-\Delta)(\pi/2)]\ \cos\ [\Psi(t-\Delta)] - \sin\ [d(t-\tau-\Delta)(\pi/2)]\ \sin\ [\Psi(t-\Delta)]\}$$

$$y'(t) = \mathrm{sgn}\ \{\sin\ [d(t-\tau-\Delta)(\pi/2)]\ \cos\ [\Psi(t-\Delta)] + \cos\ [d(t-\tau-\Delta)(\pi/2)]\ \sin\ [\Psi(t-\Delta)]\}$$

The desired truth table is now found by evaluating the above for the four cases $d(t - \tau - \Delta) = 0, 1, 2,$ and 3.

i) Case #1 $d(t-\tau-\Delta) = 0$

$$x'(t) = \mathrm{sgn}\{\cos[\Psi(t-\Delta)]\} = x'(t-\Delta)$$
$$y'(t) = -\mathrm{sgn}\{\sin[\Psi(t-\Delta)]\} = y'(t-\Delta)$$

ii) Case #2 $d(t-\tau-\Delta) = 1$

$$x'(t) = \mathrm{sgn}\{-\sin[\Psi(t-\Delta)]\} = -\mathrm{sgn}\{\sin[\Psi(t-\Delta)]\} = y'(t-\Delta)$$
$$y'(t) = \mathrm{sgn}\{\cos\ [\Psi(t-\Delta)]\} = -x'(t-\Delta)$$

iii) Case #3 $d = 2$

$$x'(t) = \mathrm{sgn}\{-\cos[\Psi(t-\Delta)]\} = -\mathrm{sgn}\{\cos[\Psi(t-\Delta)]\} = -x'(t-\Delta)$$
$$y'(t) = -\mathrm{sgn}\{-\sin[\Psi(t-\Delta)]\} = +\mathrm{sgn}\{\sin[\Psi(t-\Delta)]\} = -y'(t-\Delta)$$

iv) Case #4 $d = 3$

$$x'(t) = \mathrm{sgn}\{\sin[\Psi(t-\Delta)]\} = -y'(t-\Delta)$$
$$y'(t) = -\mathrm{sgn}\{-\cos[\Psi(t-\Delta)]\} = +x'(t-\Delta)$$

The resulting truth table is presented in Table 5-3, the entries of which enable $\hat{x}_2 = x_2$ and $\hat{y}_2 = y_2$ to be realized.

The truth table of Table 5-3 may be implemented in several manners. The simplest approach utilizes the quaternary decoder illustrated in Fig. 5-17.

Table 5-3 Differential decoder truth table. Note that $x_2' = x'(t - \tau - \Delta)$ and $x_1' = x'(t - \tau - 2\Delta)$ and corresponding for other variables.

x_2	y_2	d	x_2'	y_2'
0	0	0	$+x_1'$	$+y_1'$
0	1	1	$+y_1'$	$-x_1'$
1	1	2	$-x_1'$	$-y_1'$
1	0	3	$-y_1'$	$+x_1'$

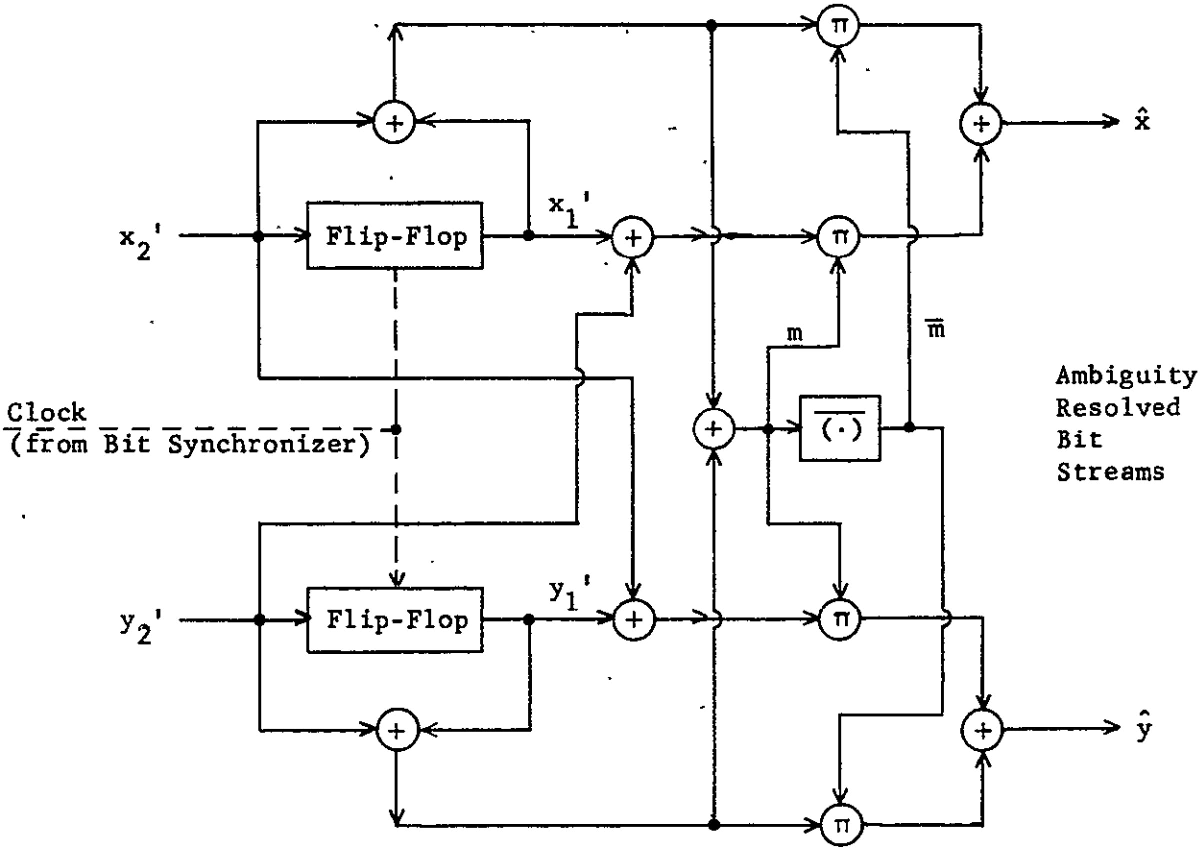

x_2'
Flip-Flop
x_1'
Clock (from Bit Synchronizer)
y_2'
Flip-Flop
y_1'
m
$\overline{m}$
$\overline{(\cdot)}$
π
$\hat{x}$
$\hat{y}$
Ambiguity Resolved Bit Streams

Fig. 5-17 Block diagram of quaternary differential decoder.

The use of the quaternary differential decoder is based on the fact that true coherent matched-filter detection of phase-modulated signals is not possible, in general, due to the unknowns of transmitter phase, path length, etc. At best, a quasi-coherent (regeneration of a noisy replica of the carrier) detection is possible. The result is that a carrier phase ambiguity on $n \cdot (\pi/2)$ radians exists for quadri-phase signals. In order to resolve the ambiguity, it is necessary to relate the present symbol to previous symbols - thus, propagating the effect of errors. The ambiguity resolution approach can double the average bit error rate. For an average bit error rate of 10^{-6}, this corresponds to an increase in the required energy-per-bit-to-noise-power spectral-density ratio of approximately 0.25 dB.

By employing the Gray code, a single error in one of the four binary digits (two present and two previous) used in the decoder can produce a single error in one of the two decoded output binary digits. Since there are four possibilities for single errors (by far, the most probable), it might appear that the differential encoding and decoding increases the error rate by four. However, each symbol contains two information bits.

The effect of quaternary differential coding is identical to binary differential coding; the error rate is doubled.

SECTION 6

SEARCH AND TRACKING CALCULATIONS

6.1 INTRODUCTION

Section 6 of Volume 1 discussed the basic system parameters and error budgets involved in the determination of the acquisition, search, and tracking action of a data relay satellite system in space. This section will discuss the detailed calculations involved in the development of these error budgets.

6.2 SEARCH CALCULATIONS

Assume an uncertainty region with 20 mr on a side as illustrated in Figure 6-1.

If acquisition is to occur between a relay and sensor vehicle, Vehicle B must direct its antenna in a given direction at the time the sequence is initiated and Vehicle A must search over the uncertainty region. At the end of the first search period (time for A to cover the search sector), Vehicle B must increment its antenna in angle and A will repeat its sweep over the sector.

The search sector is shown in Figure 6-2. Vehicle A antenna beamwidth is taken as 3.5 mr and is illustrated in the figure. Search consists of sweeping the antenna across the sector and then back after raising the angle. Overlap is shown to provide full coverage.

A scan rate of 10 milliradians per second is now assumed. This rate is considered reasonable from the mechanical standpoint. For seven scans, a minimum of 14 seconds is required. When allowance is made

9

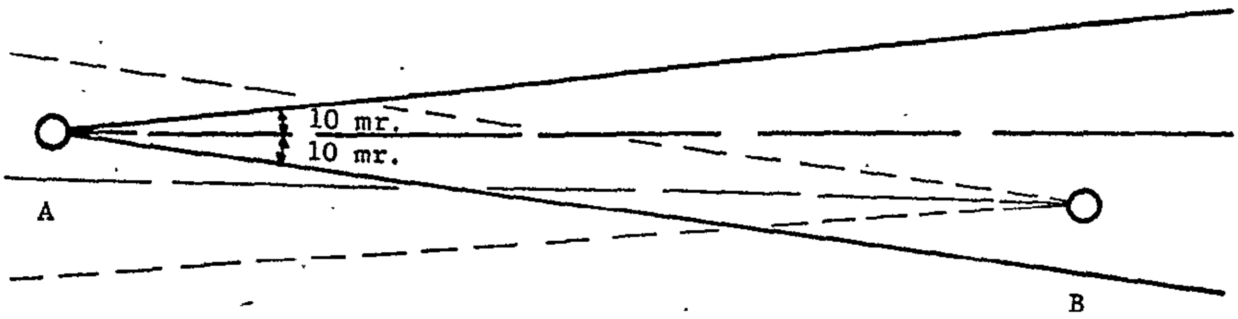

10 mr.
10 mr.
A
B

Figure 6-1 Antenna Pointing Uncertainty Geometry

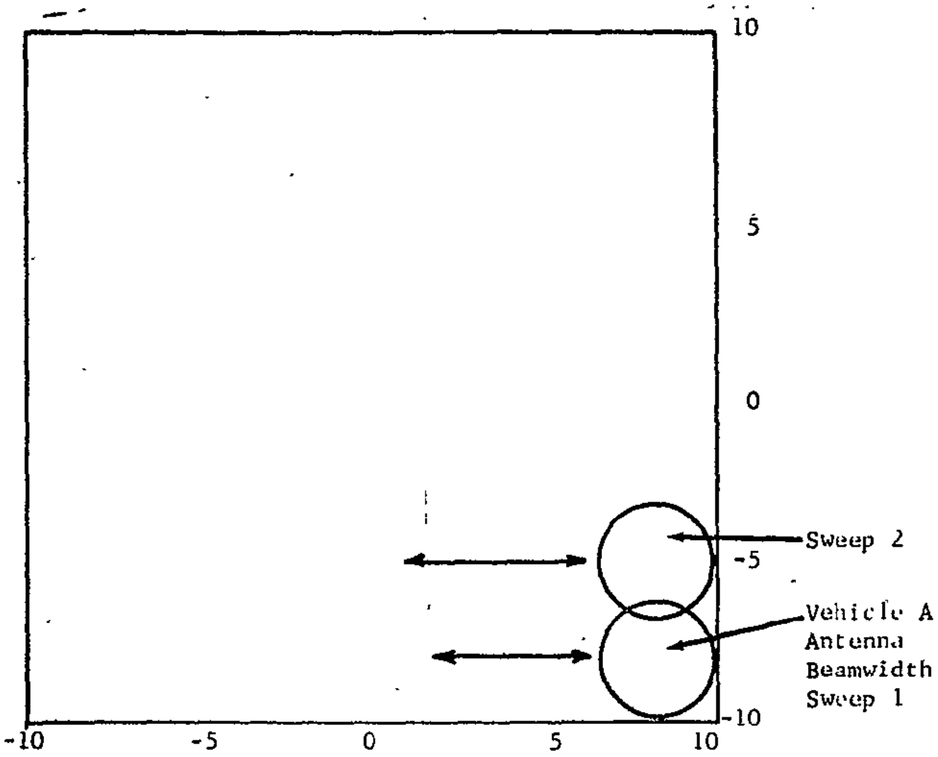

10
5
0
-5
-10
-10
-5
0
5
10
Sweep 2
Vehicle A
Antenna
Beamwidth
Sweep 1

Figure 6-2 Search Sector

for reversing the antenna at the end of the sweep, the time will be more on the order of 45 seconds. The number of times the entire sector must be searched, and hence total acquisition time, depends on the signal strength, i.e., usable beamwidth of Antenna B.

Assume a signal strength margin such that the usable beamwidth of Vehicle B is 5 mr (approximately 10-dB down). Vehicle B is contained in the search sector of Vehicle A. The antenna on B, however, need not be pointing at A since B has the same uncertainty region as A. The uncertainty region of B is divided into 5 mr cells, of which there are 16. At the rate of 45 seconds per cell, 12 minutes will be required to cover the complete sector. Thus, the maximum time to acquire is 12 minutes, although it is highly unlikely that it would ever take this long. Nevertheless, the system must be capable of accepting maximum acquisition time even though it occurs rarely.

During the time the search is in progress, the relative position of the two vehicles will change due to the relative motion. Between two relay satellites the change will be negligible.

In the worst-case condition, i.e., when the velocity vector of the sensor satellite is perpendicular to the line connecting the two vehicles, the angle will be as shown in Figure 6-3.

$$\tan \theta = \frac{v\ t}{19.027}$$

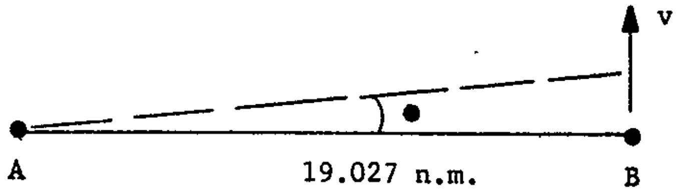
where v = velocity = 4.1 nm/sec

t = time = 12 minutes = 720 sec

θ = 0.15 radians

Although acquisition will normally take place in less than 12 minutes and at a point in space where the velocity vector is nearly parallel to the line connecting the vehicles, the problem of relative motion

Fig. 6-3


v
A
19.027 n.m.
B

Basic Acquisition Geometry

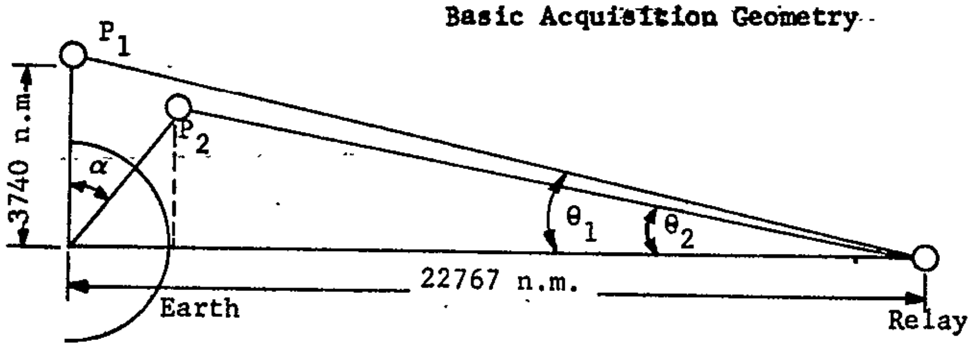

P_1
P_2
α
3740 n.m.
θ_1
θ_2
22767 n.m.
Earth
Relay

Fig. 6-4 Acquisition Geometry, Sensor Satellite in view

between the two does not vanish. The simplest solution to the problem is to restrict the region where acquisition can occur and then program in a slew rate, superimposed on the search motion, that will cancel the relative motion error.

A similar calculation will now be made for an acquisition which occurs when the sensor satellite comes into view of the relay satellite. The geometry is shown in Figure 6-4.

Let P_1 be the initial position of the sensor and P_2 the position 12 minutes later. For a 300 n.m. orbit altitude, the earth revolution period is 100 minutes and $\alpha - 43^{\circ}$,

$$\tan \theta_1 = 0.165 \approx \theta_1$$

$$\tan \theta_2 = 0.135 \approx \theta_2$$

$$\Delta\theta = 30 \text{ milliradians}$$

which exceeds the search width.

A phase lock tracking receiver is assumed for this configuration. A frequency search by the loop VCO is required simultaneously with spatial search to detect the signal if it is present in the antenna beam. The width of the search depends on the combined effect of oscillator drift and doppler. At the frequencies of interest, the doppler shift is much more significant than oscillator drift, hence the search range is essentially defined by doppler considerations.

Assume a 300 n.m. orbit for the sensor satellite and link frequency of 50 GHz. Maximum doppler shift will be 1.25 MHz. The sense of the shift ($\pm$) depends on the relative motion of the vehicle - approaching or receding.

6.3 PHASE-LOCK LOOP BANDWIDTH

Phase-lock loop bandwidth can be determined from the preceding paragraph that will permit the above search rate and provide a high probability of detection.

Assume a signal strength that will provide a signal-to-noise ratio in the loop acquisition bandwidth of 15 dB.

$$\frac{\Delta\dot{\omega}\ \mathrm{max}}{\omega n^2} = 1.05$$

$$\omega n^2 = \frac{\Delta\dot{\omega}\ \mathrm{max}}{1.05} = \frac{2\ \pi \times 1.67 \times 10^6}{1.05} = 10 \times 10^6$$

$$\omega n = 3.16 \times 10^3 = \text{loop matural frequency}$$

Assume a second-order loop with damping factor of 0.7

$$\frac{B_L}{W_m} = 0.5$$

$$B_L = 1.58 \times 10^3$$

$$2B_L = 3.15 \times 10^3 \text{ Hz} = \text{two-sided loop noise bandwidth}$$

This represents a minimum value. To insure a high probability of detection, the bandwidth should be increased to at least 5 kHz.

6.4 TIME FOR FREQUENCY SEARCH

Antenna scan rate has been assumed to be 10 milliradians per second. Time allowed for a frequency search will then be approximately

$$t = \frac{3\ \mathrm{mr}}{10\ \mathrm{mr/sec}} = \frac{\text{beamwidth}}{\text{velocity}}$$

$$t = 0.3 \text{ sec}$$

Beamwidth is taken as 3 mr to allow the vehicle to pass through off-beam center and be detected. Thus, full beamwidth of 3.5 mr only occurs on the beam center as the antenna is swept across the search sector.

6.5 FREQUENCY SEARCH RATE

Assume a frequency search sector of 0.5 MHz. Frequency search rate is therefore determined by

$$\dot{f}_s = \frac{0.5\ \text{MHz}}{0.3\ \text{sec}} = 1.67\ \text{MHz/sec}$$

Increasing the frequency search sector to 2.5 MHz requires an increase in search rate by a factor of 5. The loop bandwidth will increase by a factor of 2.2, i.e., minimum loop bandwidth becomes 7 kHz.

6.6 TRACKING ERRORS

The Tracking Error Budget of Table 6-2 of Section 6, Volume 1, involves many parameters which contribute to the budget. The following paragraphs discuss in detail the various tracking errors involved.

6.6.1 Loop Noise Bandwidth

Loop noise bandwidth is determined from the dynamic characteristics of the signal that must be tracked, assuming that the loop is high gain and of order two.

Dynamic tracking error is defined

$$\theta a = \frac{\Delta \dot{\omega}}{n^2} \text{ radians}$$

For a condition where $\theta a \leq 0.1$ radians, then

$$\dot{\omega} n^2 = \frac{\Delta \dot{\omega}}{0.1} \text{ minimum}$$

$$\omega n = \sqrt{\frac{\Delta \dot{\omega}}{0.1}}$$

In a high gain second order loop with damping factor of 0.7, the two sided noise bandwidth is

$$2BL = \omega n$$

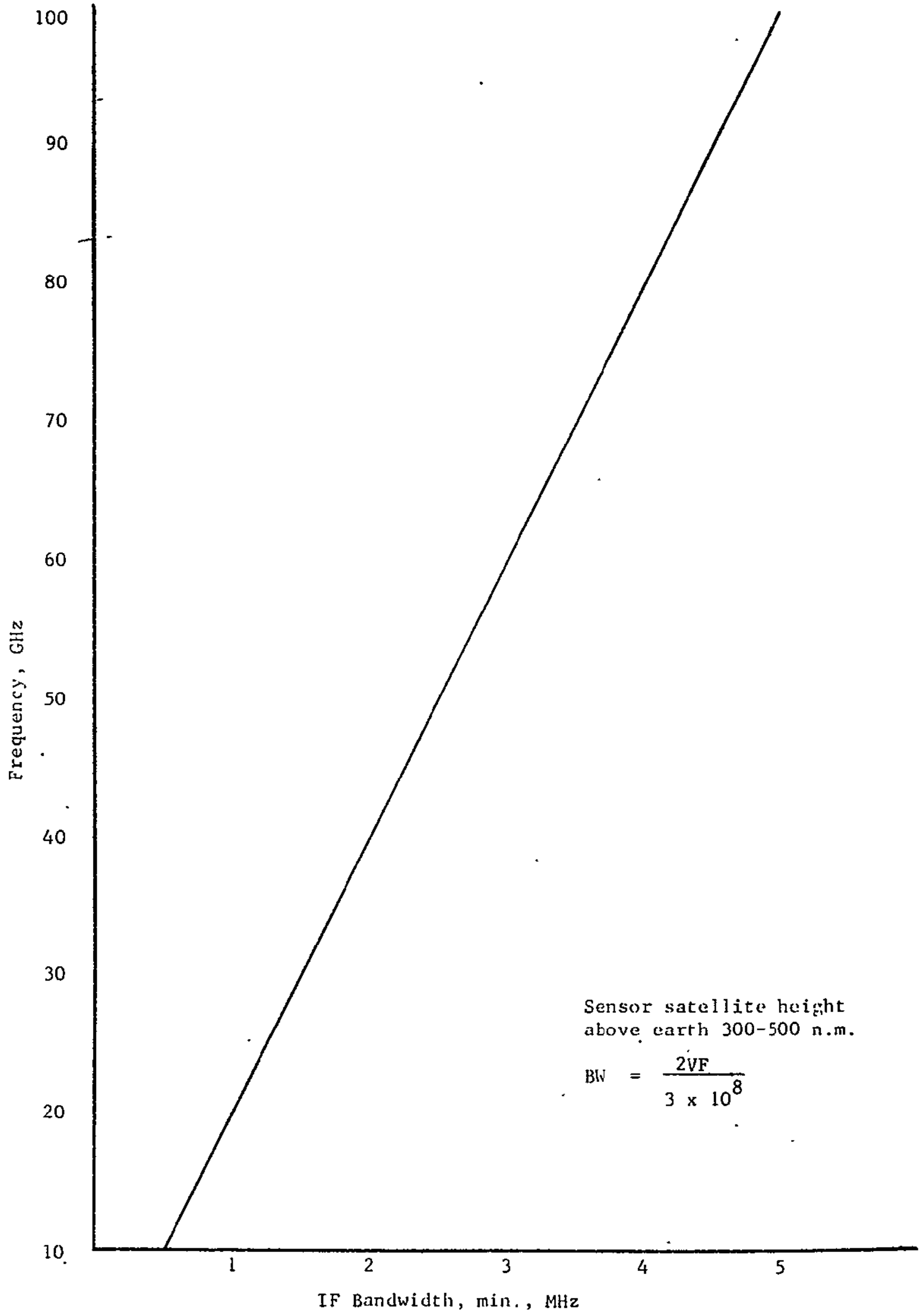
6.6.2 IF and Loop Bandwidth vs Doppler Shift

IF bandwidth vs frequency is shown in Figure 6-5 to show the effect of doppler shift based on the relative motion of the sensor satellite with respect to the relay satellite.

A plot of maximum doppler rate versus frequency is given in Figure 6-6 for 300 nm and 500 nm altitude orbits. Loop bandwidth versus maximum doppler rate is given in Figure 6-7. Loop bandwidth in this figure should be interpreted as minimum value for a given doppler rate.

6.6.3 Tracking Accuracy

When in the autotrack mode, it is desirable to maintain the peak of the antenna beam directed at the target at all times to take full advantage of the peak antenna gain. When the tracking antenna is highly directive, small angular errors can cause a significant reduction in antenna gain which, in turn degrades link performance.



Sensor satellite height above earth 300-500 n.m.

$$BW = \frac{2VF}{3 \times 10^8}$$

Figure 6-5 IF Bandwidth required due to Doppler Shift versus Carrier Frequency

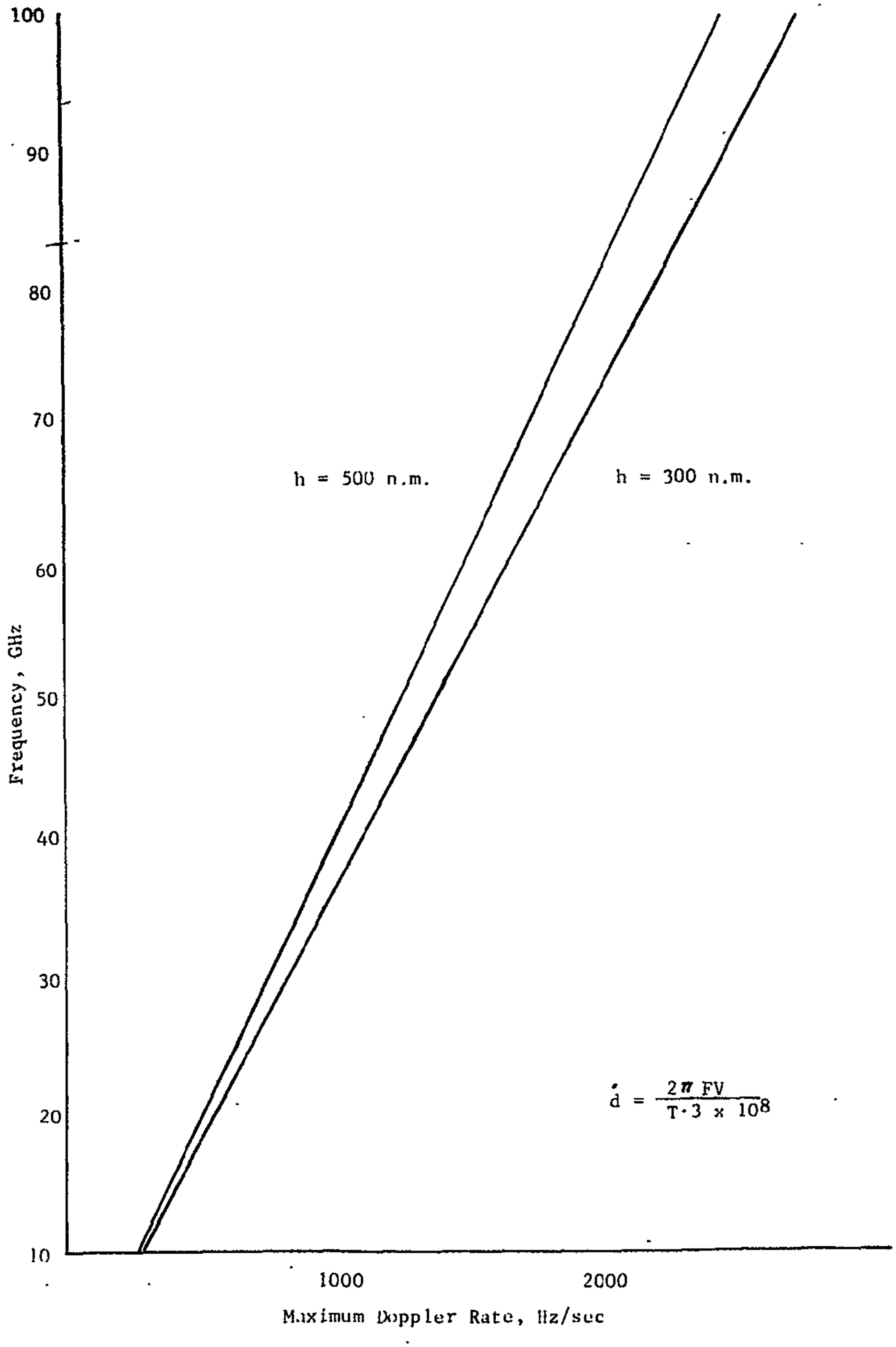


Figure 6-6 Maximum Doppler Rate versus Frequency for 300 n.m. and 500 n m. Sensor Satellite Earth Orbit

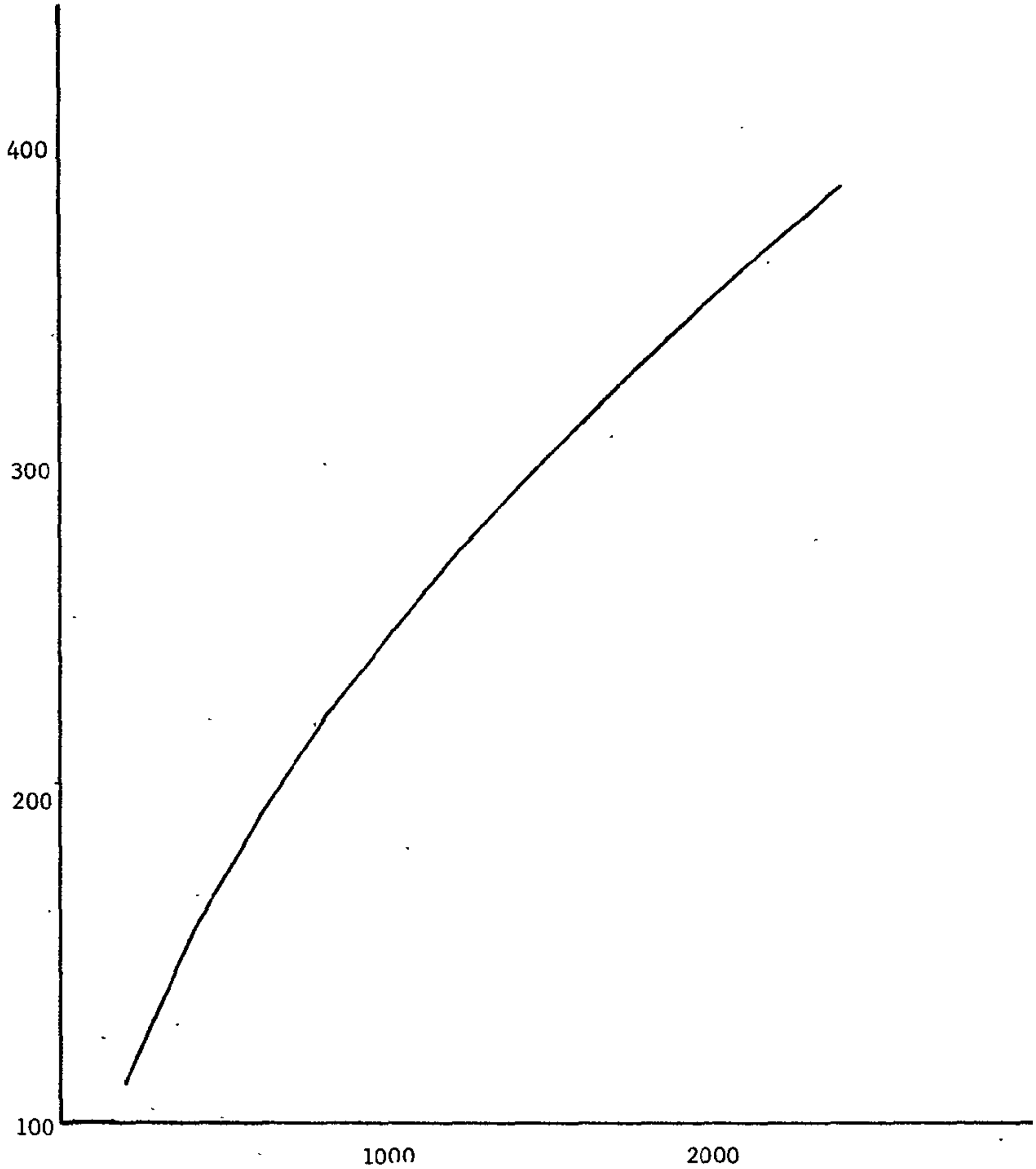


Figure 6-7. Loop Bandwidth versus Maximum Doppler Rate

A criteria of acceptability for tracking error can, therefore, be established based on the antenna gain degradation that can be tolerated by the systems. For purposes here, a degradation of 0.5 dB maximum will be assumed acceptable.

In the following paragraphs, tracking errors are identified and contributions assessed. Total error is then determined to ascertain if a specified allowable error is attainable.

6.6.4 Tracking Error Due to Thermal Noise

Angle tracking error due to the thermal noise in the tracking receiver can be calculated for the phase lock receiver configuration.

A plot of the RMS error, θ, versus signal-to-noise ratio in the receiver phase-lock-loop bandwidth is shown in Figure 6-8.

6.6.5 Tracking Error Due to Pre-Comparator and Post-Comparator Differential Phase Shift

Differential phase shifts before and after the comparator, shift the apparent null in antenna angle causing the antenna to track the signal at a point off the peak of the sum channel pattern peak.

Tracking error is the antenna angle error from the sum channel peak; in the ideal case, to the null.

A simplified block diagram of the tracking antenna and receiver is shown in Figure 6-9. For simplicity, two feeds are shown on the antenna instead of four; i.e., tracking in one axis instead of two. The outputs of the two feeds are combined in the comparator (amplitude comparison simultaneous lobing technique) to produce the sum (Σ) and difference (Δ) channel.

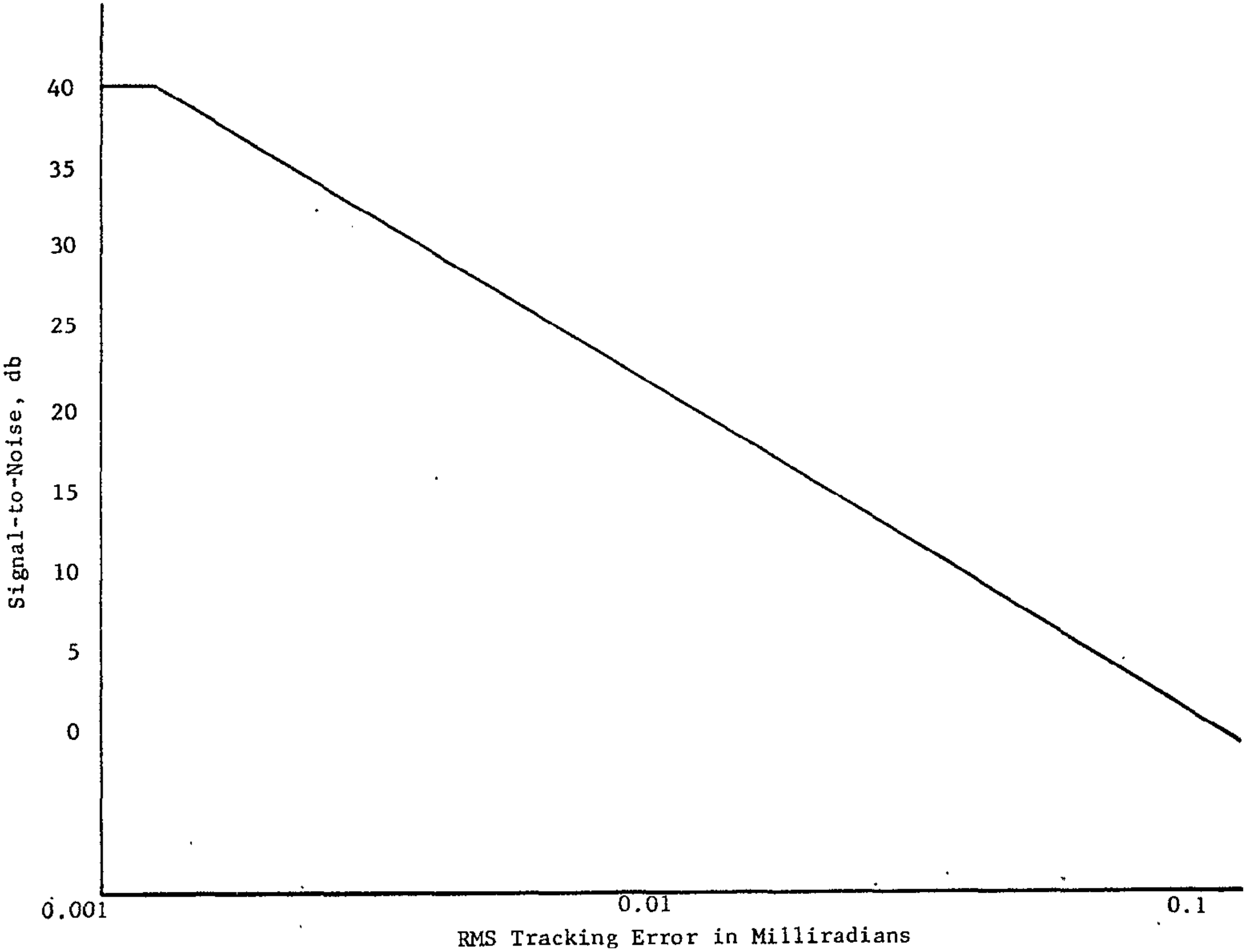


Figure 6-8 Tracking Error versus Signal-to-Noise Ratio

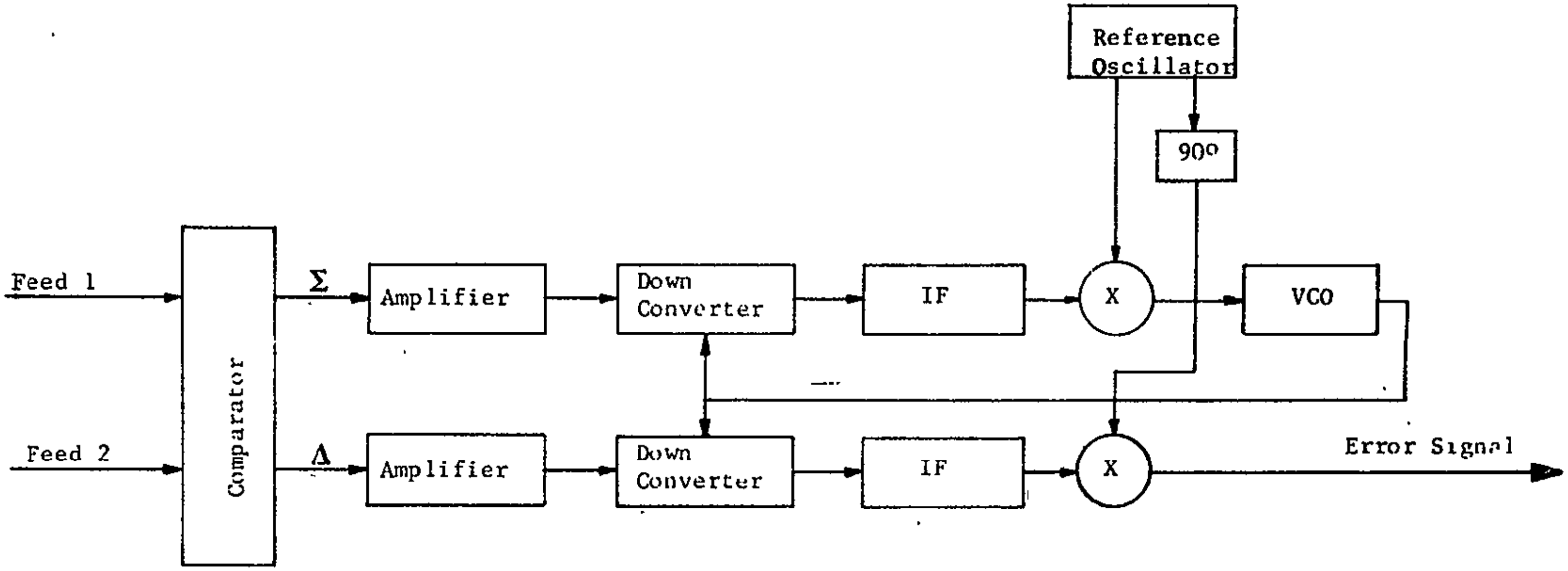

Reference Oscillator
90°
Feed 1
Feed 2
Comparator
Σ
Δ
Amplifier
Amplifier
Down Converter
Down Converter
IF
IF
X
X
VCO
Error Signal

Figure 6-9. Simplified Tracking Antenna/Receiver Configuration

The sum channel signal is amplified and down-converted in the receiver. The local oscillator signal is provided by the receiver VCO, which is phase locked to the carrier component of the signal. Receiver AGC is derived from the sum channel signal and controls the gain in both the sun and difference channels.

The difference channel signal is amplified and translated in the same manner as in the sum channel. Note that any phase variation in the sum channel is introduced into the difference channel by the VCO since the VCO is phase locked to the sum channel. The DC error voltage is obtained from the product detector shown in the figure.

The signal at the antenna can be written as follows:

$f(t) = A \sin[\omega_c t + a + \phi(t)]$

ωc = Radian carrier frequency

a = Arbitrary reference phase

$\phi(t)$ = Angle modulation on the carrier

This signal will be present in each of the feeds. The amplitude in each feed will depend upon the position of the signal source relative to antenna boresight. When autotracking, the two amplitudes will be nearly equal.

Feed #1 $f_1(t) = A_1 \sin[\omega_c t + a + \beta 12 + \phi(t)]$

Feed #2 $f_2(t) = A_2 \sin[\omega_c t + a - \beta 12 + \phi(t)]$

where β = relative phase shift between $f_1(t)$ and $f_2(t)$, i.e., pre-comparator differential phase shift.

Phase shift β can result from differences in feed line lengths to the comparator or from other dissimilar electrical characteristics of the feeds. It can also result from a received wavefront which is not perpendicular to the boresight axis. When autotracking this condition is essentially met so that little, if any, relative phase shift results from a non-perpendicular wavefront.

This analysis will be concerned with the relative phase shift between the feeds and between the sum and difference channels. Thus, the arbitrary phase reference, a, can be set equal to zero.

At the comparator output the sum channel signal can be expressed as the sum of the two feed signals:

$$f\Sigma\ (t) = A_1 \sin\left[\omega_c t + \beta/2 + \phi(t)\right] + A_2 \sin\left[\omega_c t - \beta/2 + \theta(t)\right]$$

and the difference channel:

$$f\Delta\ (t) = A_1 \sin\left[\omega_c t + \beta/2 + \phi(t)\right] - A_2 \sin\left[\omega_c t - \beta/2 + \theta(t)\right]$$

It should be noted, at this point, that a phase shift to the carrier has been introduced instead of a time delay to the overall modulated signal. The reason for this is that the receiver is assumed to track only the carrier component of the signal. That is, the phase-lock-loop is a carrier tracking loop and the modulation is assumed to be filtered out prior to the difference channel detector.

Now, the post detection differential phase shift can be introduced and the error signal at the output of the difference channel detector derived.

$$g\Sigma\ (t) = G\left\{ A_1 \sin\left[\omega_{ct} + \beta/2 + \gamma/2\right] + A_2 \sin\left[\omega_{ct} - \beta/2 + \gamma/2\right]\right\}$$

$$g\Delta\ (t) = G\left\{ A_1 \sin\left[\omega_{ct} + \beta/2 - \gamma/2\right] - A_2 \sin\left[\omega_{ct} - \beta/2 - \gamma/2\right]\right\}$$

where, G = channel gain

γ = post comparator differential phase shift

Note the modulation has been dropped for reasons presented above.

The error signal is:

$$E(t) = k\, g_{\Sigma}(t) \times g_{\Delta}(t)$$

k = gain constant of the detector.

$$E(t) = k\, G^2\, 1/2\, (A_1^{\,2} - A_2^{\,2}) \operatorname{Cos} \gamma + \frac{A_1 A_2}{2}\left[\operatorname{Cos}(\gamma - \beta) - \operatorname{Cos}(\gamma + \beta)\right]$$
$$- \frac{A_1^{\,2}}{2} \operatorname{Cos}(2\, \omega_{ct} + \beta) + \frac{A_2^{\,2}}{2} \operatorname{Cos}(2\, \omega_{ct} - \beta)$$

The double frequency terms can be ignored since they will be filtered out. The DC component is of concern here and is given by the first two terms of the expression.

The final step in the analysis requires the amplitudes A_1 and A_2 be expressed as a function of antenna angle. This is done by assuming the individual feed patterns to be Gaussian in form and, as such, can be expressed:

$$A_1(\theta) = \exp\left[\frac{-2.78}{2}\left(\frac{\theta + 0.045}{0.13}\right)^2\right]$$

$$A_2(\theta) = \exp\left[\frac{-2.78}{2}\left(\frac{\theta - 0.045}{0.13}\right)^2\right]$$

where θ is the spatial angle in degrees. Sum channel antenna beamwidth is 3.5 mr.

Arbitrarily select $kG^2 = 2$, since kG^2 is a constant affecting only the gradient of the output error signal. E (t) now becomes:

$$E(t) = (A_1^{\,2} - A_2^{\,2}) \operatorname{Cos} \gamma + A_1 A_2 \left[\operatorname{Cos}(\gamma - \beta) - \operatorname{Cos}(\gamma + \beta)\right]$$

Tracking error is obtained from the above expressions when $\Sigma(t) = 0$.

A plot of track error versus β and γ is given in Figure 6-10. Figure 6-11 is a plot of the error voltage versus angle θ to show the effects of γ and β.

It can be seen that when $\gamma = 0$ no error results even though $\beta = 0$. It should be noted that if either γ or β is zero the tracking error will be zero. Curves are also given for $\gamma = 10^{\circ}$ and $\beta = \pm 4^{\circ}$. Note that the phase reversal in β shifts the crossover from positive to negative error but that the absolute value of the error remains the same. The same result is obtained if the phase of γ is reversed instead of β. Finally, the curves are linear over the range(s) of investigation. Increasing the range introduces non-linearity at the end points; however, this is of little concern since tracking is usually constrained to small angles off boresight.

6.6.6 Tracking and Pointing Error Due to Amplitude Unbalance

In a two-axis, simultaneous-lobing, amplitude-comparison, tracking antenna, four individual feeds are combined in a comparator to produce the sum and difference of the outputs of the individual feeds. Ideally, when the individual patterns are identical, the difference channel null and electrical axis (each axis) coincides, in angle, with the sum channel; and since the servos track the null, the signal is received on the sum channel pattern peak.

An amplitude unbalance between the individual feed patterns will shift angle of the difference channel null and cause the antenna to track the signal at a point off the sum channel patterns peak.

The individual beam patterns are assumed Gaussian in form and, as such, can be represented as follows:

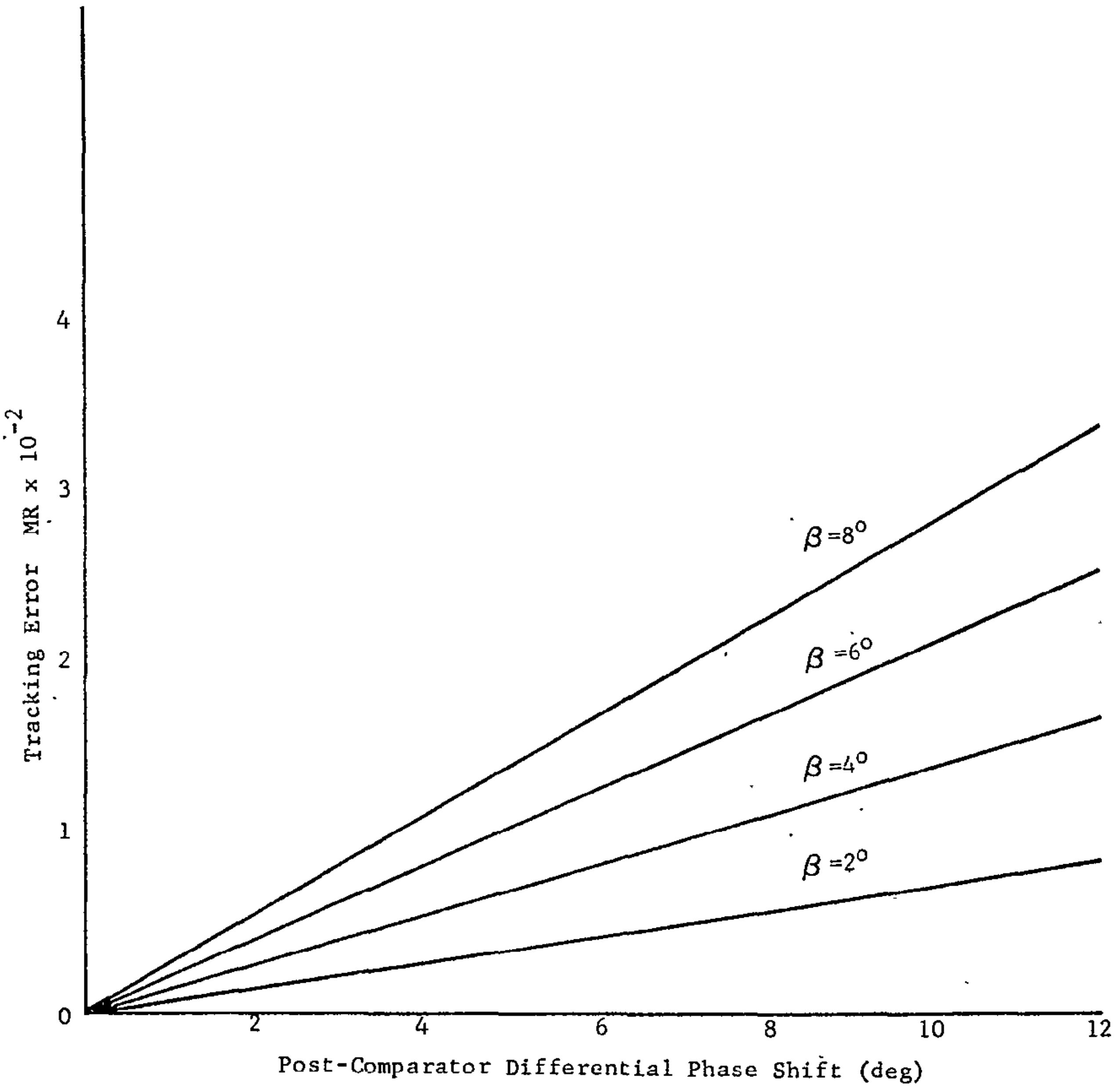


Figure **6-10** Tracking Error versus Post-Comparator Differential Phase Shift and Pre-Comparator Differential Phase Shift

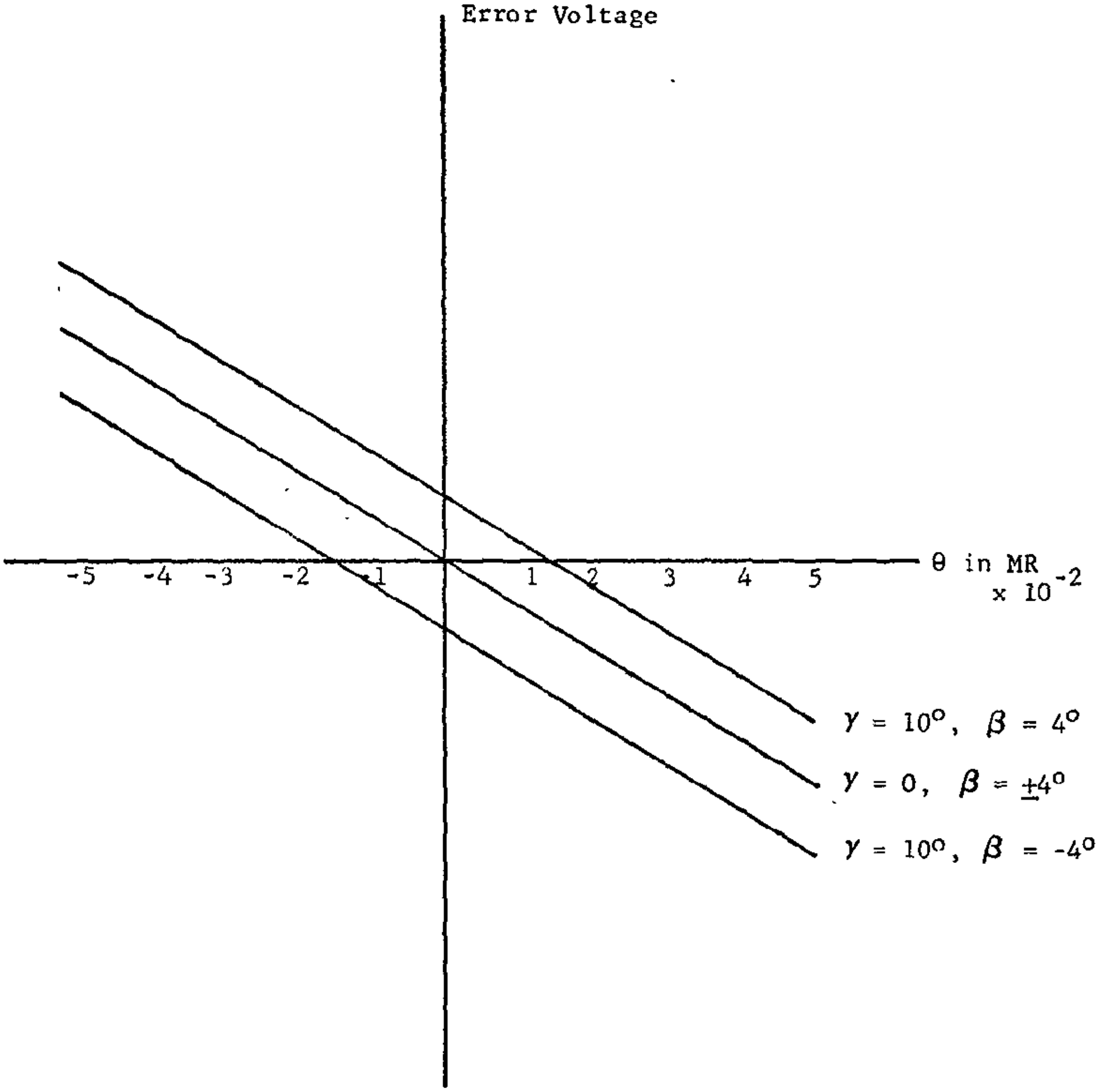


Figure 6-11 Error Voltage versus Angle γ

$$P_1(\theta) = G_1 \exp\left[-2.78\left(\frac{\theta + 0.045}{0.13}\right)^2\right]$$

$$P_2(\theta) = G_2 \exp\left[-2.78\left(\frac{\theta - 0.045}{0.13}\right)^2\right]$$

where θ is in degrees.

The sum channel is given by:

$$P\Sigma\cdot(\theta) = P_1 \;\&\; P_2$$

At the crossover (θc), $P_1 = P_2$. Thus an expression can be written which relates G_1, G_2 and θc where G_1 and G_2 are the amplitudes of the individual patterns.

$$G_1 \exp\left[-2.78\left(\frac{\theta c + 0.045}{0.13}\right)^2\right] = G_2 \exp\left[-2.78\left(\frac{\theta c - 0.045}{0.13}\right)^2\right]$$

$$\frac{G_1}{G_2} = \frac{\exp\left[-2.78\left(\frac{\theta c - 0.045}{0.13}\right)^2\right]}{\exp\left[-2.78\left(\frac{\theta c + 0.045}{0.13}\right)^2\right]}$$

$$\frac{G_1}{G_2} = \exp\left[29.6\ \theta c\right]$$

Amplitude unbalance (G_1/G_2) versus boresight shift θc is plotted in Figure 6-12.

6.6.7 Tracking and Pointing Error Due to Unequal Individual Beamwidth

Tracking error due to amplitude unbalance, treated previously, considered the effect of varying the individual beam amplitudes while holding the beamwidth constant. In this section the amplitude of the two individual patterns is held constant and the beamwidth of one of the individual patterns is allowed to vary with respect to the other.

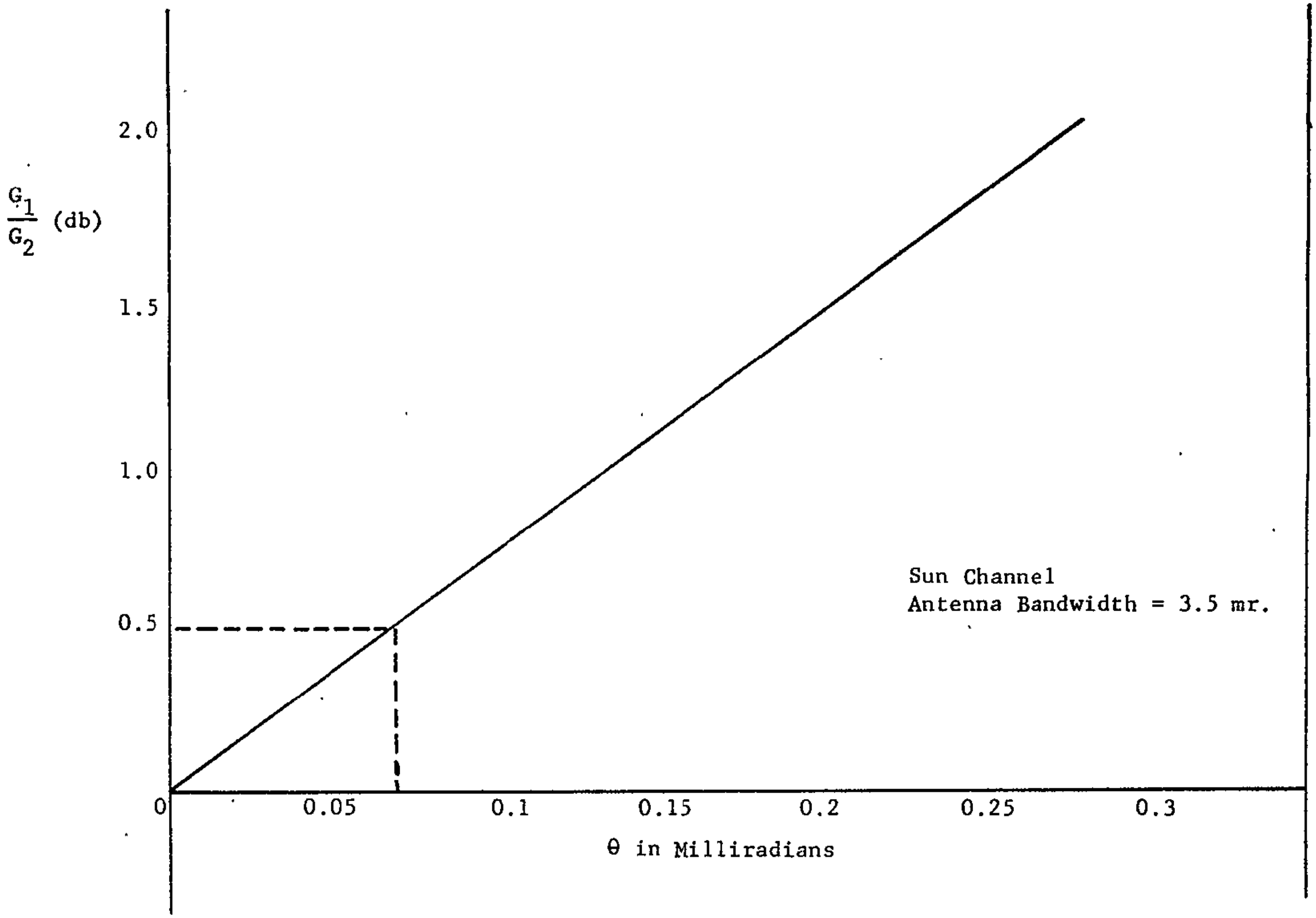


Figure 6-12 Amplitude Unbalance versus Boresight Shift

As before, the individual beam patterns are assumed Gaussian in form.

$$P_1(\theta) = \exp\left[-2.78\left(\frac{\theta + 0.045}{\theta_{B_1}}\right)^2\right]$$

$$P_2(\theta) = \exp\left[-2.78\left(\frac{\theta - 0.045}{\theta_{B_2}}\right)^2\right]$$

where θ and θ_B (3 dB beamwidth) are in degrees.

As before, the null occurs when

$$P_1 = P_2$$

$$\exp\left[-2.78\left(\frac{\theta + 0.045}{\theta_{B_1}}\right)^2\right] = \exp\left[-2.78\left(\frac{\theta - 0.045}{\theta_{B_2}}\right)^2\right]$$

For this case the solution is found by letting the computer search for the null. The results are plotted in Figure 6-13.

6.6.8 Tracking Error Due to Antenna Temperature

Error due to temperature refers to the affect of non-uniform heating of the antenna, and components or cables in the transmission path. Error may be introduced (existing error allocations increased) due to one side of the spacecraft being exposed to the sun for prolonged periods of time. For the purpose of this analysis an error of 0.1 mr is allocated for temperature variations. It is expected that the contribution from the antenna will be the only significant error and that it will be within the allocated value.

The mean-square tracking error is given by

$$\theta^2 = \frac{5\,\psi^2}{2\,\mathrm{DBc}\,\beta^2 \ln 2 \ln 10 \left[\prod_{k=0}^{N} \mathrm{Jo}\,(\Phi_k)\right]^2} \qquad \frac{\phi\,B_n}{S}$$

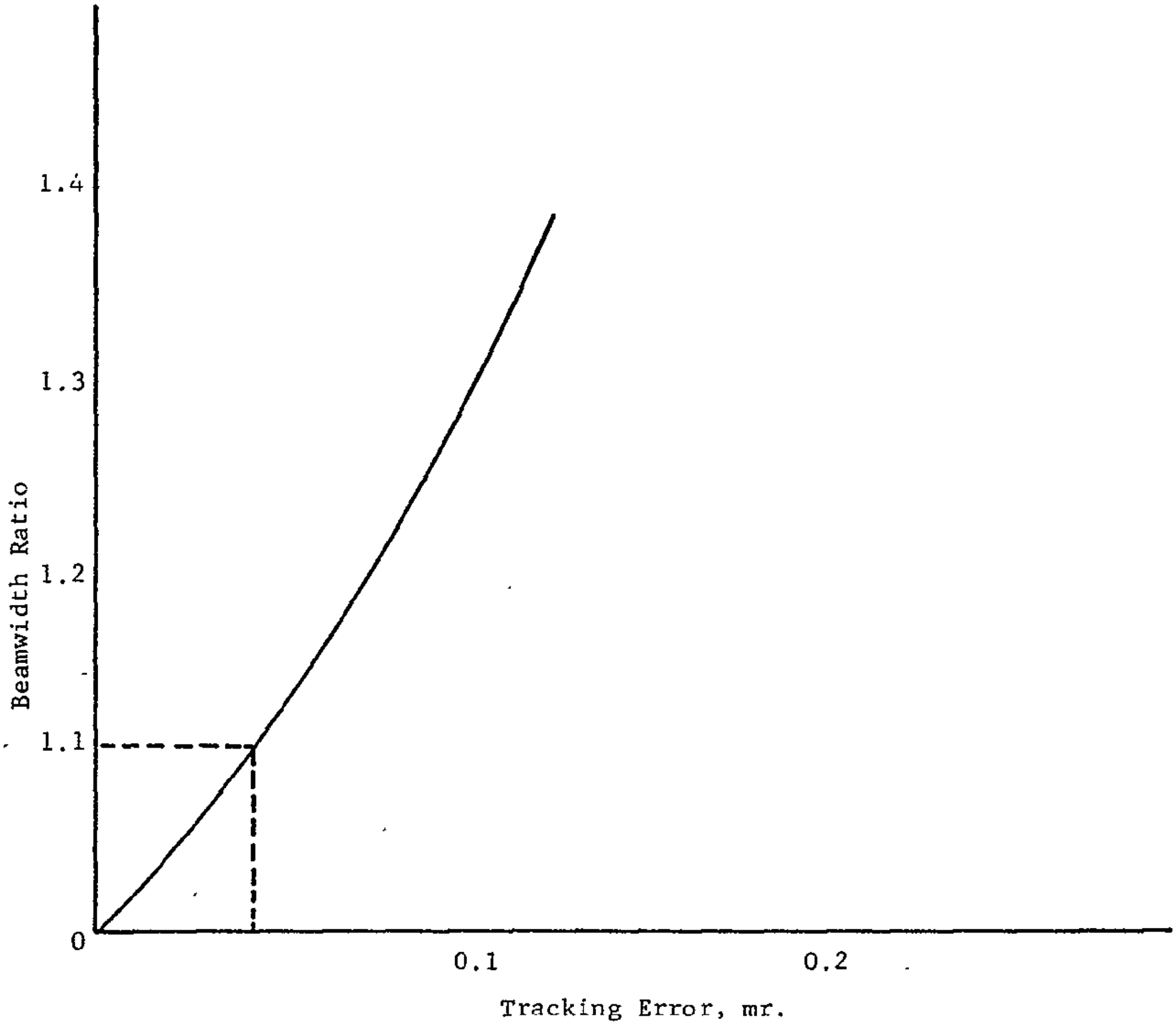


Figure 6-13 Tracking Error due to Unequal Individual Beamwidths Sum Channel Antenna Beamwidth = 3.4 mr.

ψ = Sum pattern 3 dB beamwidth (radians)

DBc = Crossover point in dB down from the individual beam pattern peak

β = Ratio of sum pattern 3 dB beamwidth to individual pattern of 3 dB beamwidth

ln ( ) = Natural logarithm

Φ = One-sided noise density in difference channel

B_n = Servo noise bandwidth (one-sided)

S = Received signal power on boresight

Jo ( ) = Zero order Bessel function of first kind

The factor in the denominator, Π Jo (Φ_k), accounts for the carrier loss due to modulation. If the signal power is taken as carrier power this term can be ignored. This is done here and S_c will denote carrier power.

It can be seen that tracking error is a function of the noise power in the servo bandwidth. It is more convenient to plot error as a function of signal-to-noise ratio in the tracking receiver phase lock loop bandwidth rather than servo bandwidth. The equation can be modified to do this as follows:

$$\frac{S_c}{\Phi B_n} = \frac{\Phi s \; B_T}{\Phi B_n} \cdot \frac{S_c}{\Phi_s \; B_T}$$

where Φ_s = noise density in the sum channel

B_T = phase lock loop double sided noise bandwidth

For this analysis, Φ_s = Φ and the mean square error can be written.

$$\theta^2 = \frac{5\,\psi^2}{2\ \mathrm{DBc}\ \beta^2\ \ln 2\ \ln 10} \cdot \frac{B_n}{B_T} \cdot \frac{N}{S_c}\ \text{radians}^2$$

where

$$N = \Phi B_T$$

6.6.9 Servo Lag Errors

Servo lag is the term applied to the angle tracking error incurred due to the characteristics of the servo system when tracking a moving vehicle.

Steady state tracking error is determined as follows:

$$E = \frac{\text{angular acceleration}}{K_a} = \frac{\dot{w}}{K_a}$$

K_a = Servo loop acceleration error constant

The maximum tracking acceleration requirement for the synchronous relay satellite is on the order of 0.023 milliradians/sec^2. To maintain tracking error of 0.1 milliradians or less will require a servo error constant,

$$K_a = \frac{0.023}{0.1} = 0.23$$

which should be readily obtained for the antenna under consideration.

6.6.10 Servo Bias Error

Servo bias error refers to an offset error that commonly occurs with dc amplifiers. For example, the error voltage at the receiver is a dc voltage with positive value for errors of one sense and negative for errors of the opposite sense. Zero voltage corresponds to zero error. A dc offset would introduce a bias error called servo bias error. Errors of this type are minimized by careful design.

An allocation of 0.1 mr is usually given to servo bias error.

SECTION 7

MILLIMETER WAVE ANTENNAS

7.1 THE HIGH-GAIN MILLIMETER WAVE ANTENNA

Antenna gains must be high at millimeter waves - of the order of 60 dB, to maximize ERP with existing transmitting tubes and to maximize G/T using existing relatively high noise figure techniques.

Therefore, the millimeter wave high gain antenna must be able to (1) provide very high gain while also operating in the changing solar environment without deformation or change in surface accuracy which would greatly reduce gain and distort the antenna pattern, and (2) be capable of being constructed with minimum weight and inertial mass.

These requirements which are unique and critical to millimeter wave high gain operation are met by using a cassegrain antenna with high efficiency surfaces, and by using special materials which maintain contour accuracy and minimize structure weight.

The following paragraphs address those specific considerations.

7.1.1 Antenna Sizes

The relationships between antenna diameter, frequency, gain and beamwidth for the frequency and beamwidth range of interest are illustrated in Figure 7-1. For example, at 60 GHz an antenna aperture must be about five feet in diameter to produce a 0.2° beamwidth.

Note from Figure 7-1 that antennas between six to eight feet in diameter are required for the maximum gain to maximize ERP and G/T.

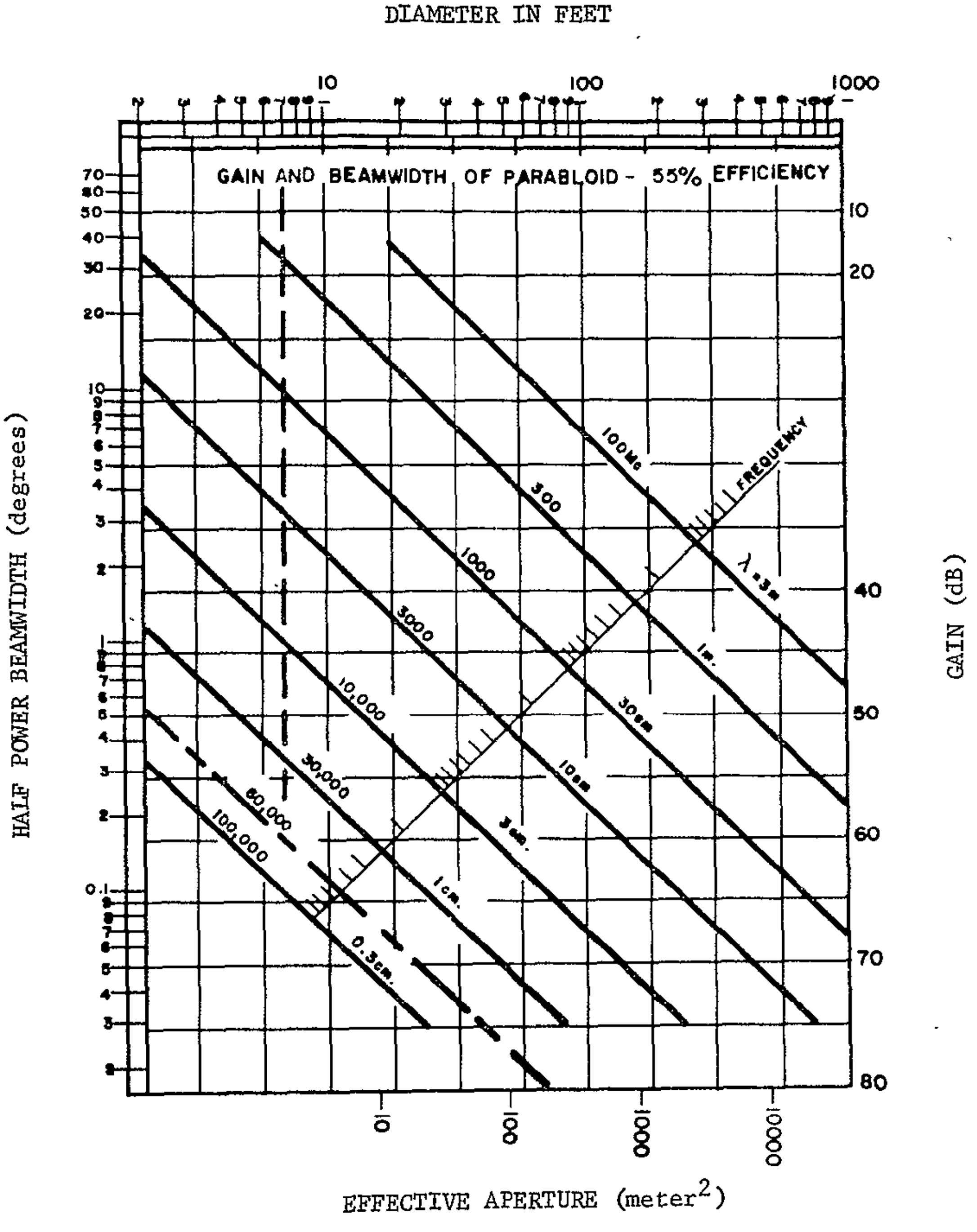


FIGURE 7-1 GAIN AND BEAMWIDTH OF A PARABOLIC ANTENNA (Efficiency 55%)

7.2 HIGH-EFFICIENCY REFLECTOR DESIGN

The gain achieved by conventional Cassegrain systems can be substantially improved by using special contoured reflectors (High Efficiency). As discussed by Galindo[1] and Williams[2], the energy incident on the subreflector from the feed can be redistributed on the main reflector to yield uniform illumination and therefore maximum gain from an antenna of specified size. The perturbation in the path length introduced by changing the subreflector contour is corrected by a slight modification in the main reflector surface. This correction is frequency independent. A gain increase of 0.5 dB is typically achieved using this technique.

In conventional systems, a 10 dB edge illumination for the subreflector is commonly chosen to optimize the trade-off between illumination and spillover efficiencies. The combination efficiency is typically 80 percent. In the high efficiency antenna, since uniform illumination can be obtained independent of the subreflector edge illumination, typically 94 percent can be obtained for the combined efficiencies. This is achieved by lowering the subreflector edge illumination to approximately 17 to 2 dB.

High efficiency surfaces are largely standard with large (90-100 ft diamter) Comsat-type antennas all over the world, where the high efficiency technique has provided 60 dB gain from an antenna diameter of 97 feet at 4 GHz where in excess of 120 ft diameter could be required to achieve this gain increase using the standard 55% efficient Cassegrain antenna.

The use of the high-efficiency dual reflector system (Figure 7-2) can produce efficiencies approaching 90% for the directive gain. This efficiency includes all loss factors down to the antenna terminals, i.e., aperture illumination, main dish spillover, sub-dish spillover, cross polarization,

---

[1] Galindo, Victor, "Design of Dual-Reflector Antennas With Arbitrary Phase and Amplitude Distribution," IEEE, Transaction on Antenna and Prop. July, 1964.

[2] Williams, William F., "High Efficiency Antenna Reflector," Microwave Journal, June 1964.

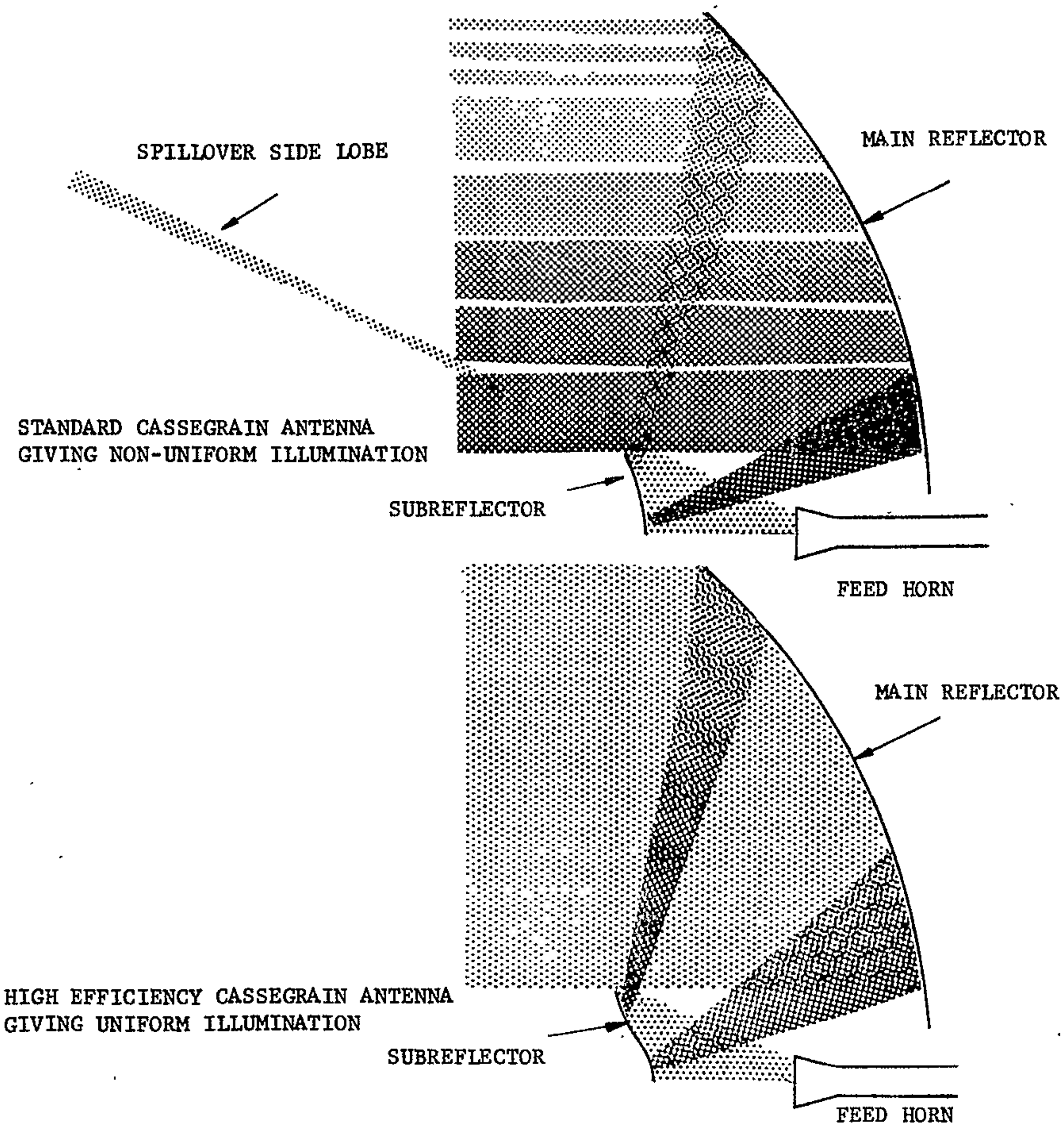

SPILLOVER SIDE LOBE
MAIN REFLECTOR
STANDARD CASSEGRAIN ANTENNA
GIVING NON-UNIFORM ILLUMINATION
SUBREFLECTOR
FEED HORN
MAIN REFLECTOR
HIGH EFFICIENCY CASSEGRAIN ANTENNA
GIVING UNIFORM ILLUMINATION
SUBREFLECTOR
FEED HORN

Figure 7-2 Comparison of the Illumination Characteristics of a Standard and High Efficiency Cassegrain Antenna System

TABLE 7-1

EFFICIENCY COMPARISON

Parameter	High-Efficiency	Conventional
Aperture	1.00	0.875
Spillover	0.973	0.903
Curvature (estimated)	0.98	0.99
Cross Polarization	0.9975	0.9975
Phase	1.00	1.00
RMS Surface	0.975	0.975
Blockage	0.980	0.993
Diffraction	0.994	0.936
Calculated Efficiency	0.903	0.707

	High-Efficiency		Conventional	
	Theory	Measured	Theory	Measured
Gain (73.57 Gc)	53.09	52.94	51.94	51.74 dB
3 dB Beamwidth	0.395°	0.39°	0.44°	0.45°
Null Width of Main Beam	0.917°	0.88°	1.205°	1.12°
E plane Sidelobes	16.9	17.3	24	20 dB
H plane Sidelobes	16.9	16.5	24	19.5 dB

phase, and blockage. This technique achieves the high efficiency by providing nearly uniform aperture illumination with very low spillover loss and has been demonstrated at 73.57 GHz by John Bowes of Philco-Ford as detailed in the actual comparative performance data of Table 7-1 where 1.2 dB increase in gain was achieved.

It is evident that this technique is a highly useful one for minimizing antenna size and weight for a given gain.

7.3 GAIN VERSUS ANTENNA SURFACE DISTORTION

Antenna dish distortion can cause significant gain loss and beam squint, depending on the size and construction of the dish. Figure 7-3 illustrates the effect on gain of uniform distortion in terms of edge deflection and paraboloid F/D ratio. It is significant that a total edge deflection of only 0.020" can cause significant gain loss at 60 GHz.

For Cassegrain antenna systems, feed and sub-reflector movements resulting from thermal or other effects are of significance in the application. Sub-reflector movement along the focal axis has approximately the same effect as moving the feed of a focal point paraboloid along the focal axis. This results in a defocusing effect which produces gain loss. Sub-reflector notations and translations off axis also result in gain loss as well as beam squint. Thermal effects in its support structure must therefore be taken into account in the design of the antenna.

Surface tolerance effects has been thoroughly investigated by J. Ruze who has shown that the relative loss of gain is -

$$\frac{G}{G_o} = \exp\left[\frac{4 \pi \delta}{\lambda}\right]$$

where G = realized gain

G_o = ideal paraboloid gain

δ = rms error

λ = wavelength

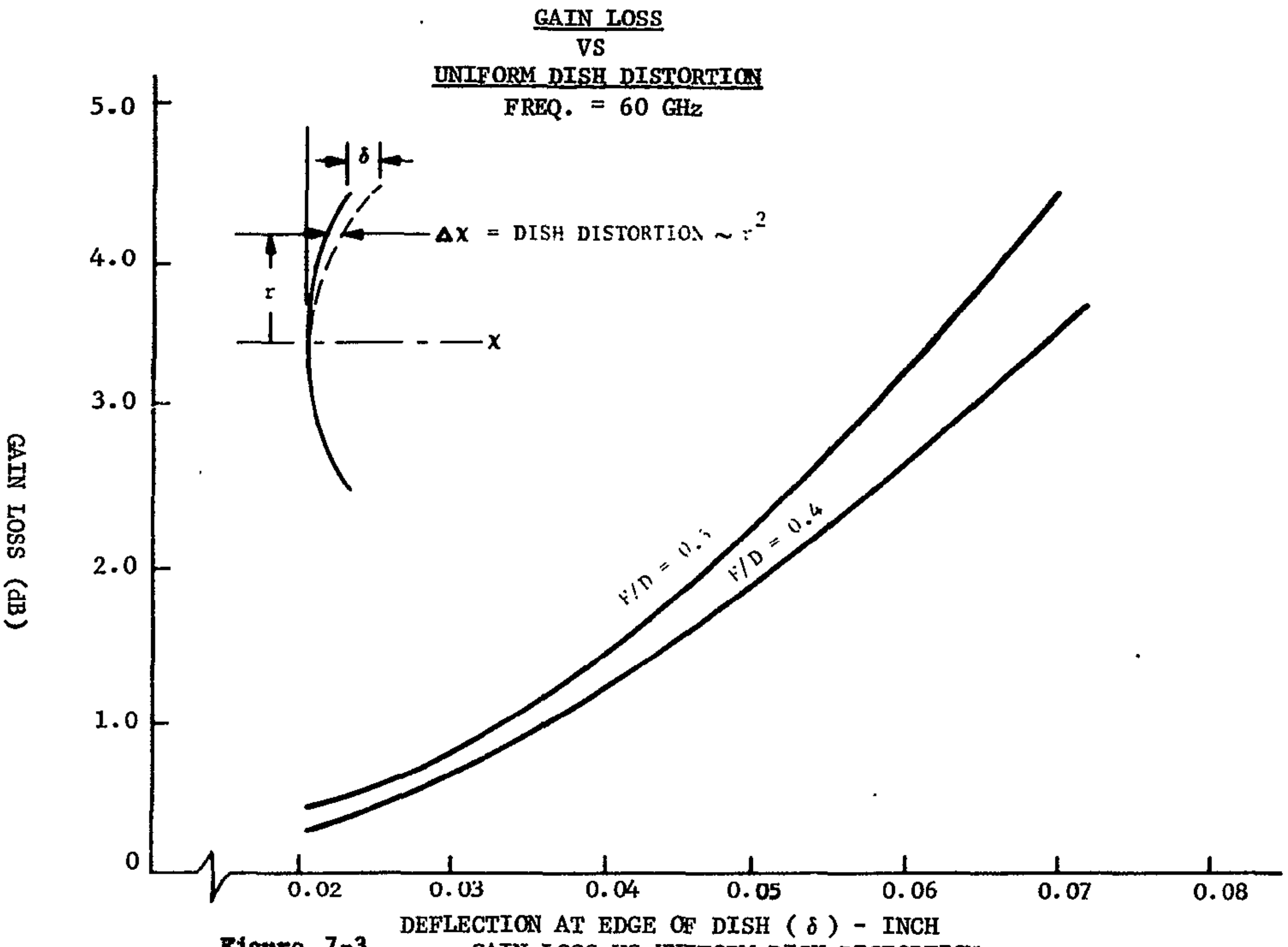


Figure 7-3 GAIN LOSS VS UNIFORM DISH DISTORTION

This relationship is shown in the graph of Figure 7-4. For a Cassegrain system -

$$\delta = \sqrt{M_{rms}^2 + S_{rms}^2}$$

where M_{rms} = main-reflector rms error

S_{rms} = sub-reflector rms error

The graph shows that for a 0.2 dB gain loss at 60 GHz, $\delta = 0.003$ inch rms.

7.4 BASELINE MATERIAL DESIGN

The basic millimeter wave antenna design depends critically on the ability of the structure to maintain accurate contours in the changing solar environment. Consequently, extensive structure and materials analyses of the main reflector and subsequent preliminary design of antenna structures were made with several candidate materials.

Analyses were conducted in Philco-Ford based on a thin shell consisting of two face sheets of the material under consideration separated by an aluminum honeycomb core. A number of materials were initially evaluated and discarded because of one or more disallowable properties such as high thermal coefficient of expansion, excessive weight, or poor fabricability.

Graphite/epoxy, invar, and beryllium were the candidates remaining after this process and they have been studied intensively during the reporting period. Since initial results indicated unacceptably large thermal deflections for the Be face sheet case, a thin solid shell of Be was investigated as an alternate design.

Results for the three primary candidate designs are summarized in Figure 7-5. Graphite/epoxy is the clear choice for both RF performance and weight considerations.

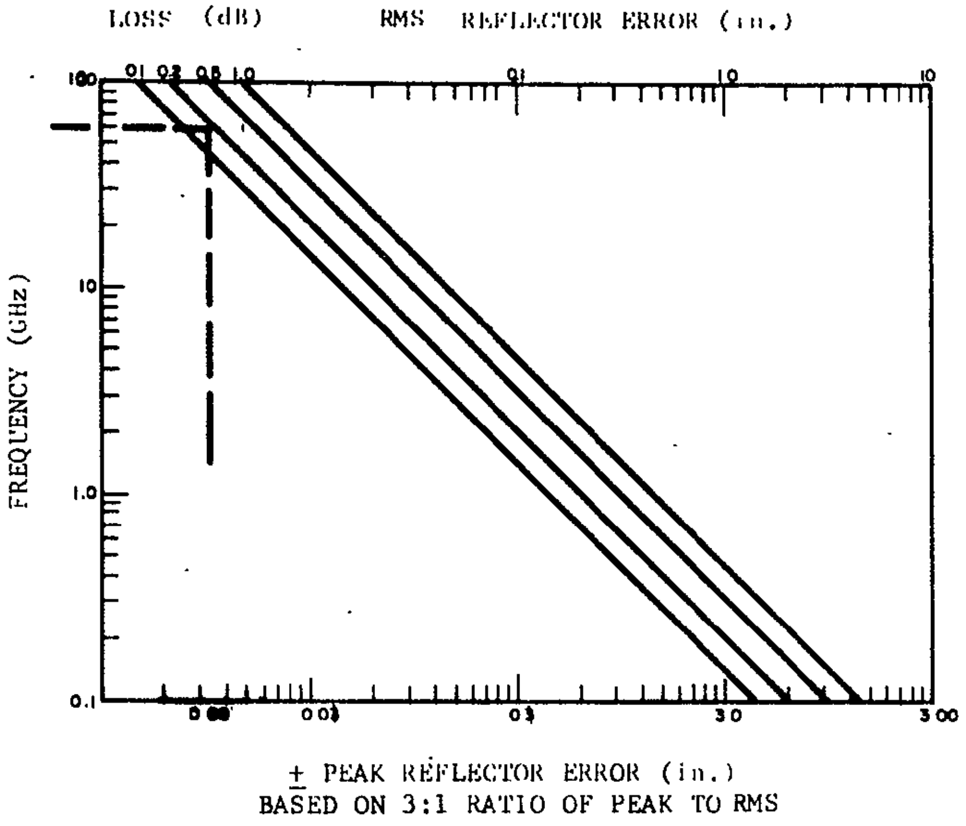


Figure 7-4 Maximum Operating Frequency of Parabolic Antenna vs Reflector Error

Material & Fabrication Technique	Temp. Coef. of Expansion Used in Analysis (in/in/°F	Weight (lb.) Dish, Sub-Reflector & Strusts Including Paint	Solar-Induced Effects			
			Maximum Edge Defl. (inches)	Beam Squint (degrees)	Gain Loss (dB)	Diff. Boresight* (degrees)
0.010" Graphite-epoxy skins on 3/4" aluminum honeycomb	0.92×10^{-6}	27.24	0.008	0.0075	0.14	< 0.01
0.010" Invar skins on 3/4" aluminum honeycomb		62.99	0.015	0.013	0.40	< 0.01
1/16" Solid Beryllium Shell	6.5×10^{-6}	35.87	0.043	0.088	3.96	< 0.01

*Divergence between main beam peak and tracking pattern null.

Figure 7-5 Comparison of Candidate Antenna Structural Approaches

7.4.1 Analytical Approach

The antenna was subdivided into three major thermal regions: 1) the main reflector, 2) the subreflector, and 3) the struts. For the main reflector, it was assumed that the thermal effects of the subreflector and struts were minor and therefore could be neglected. For the subreflector, it was assumed that the thermal effects of the struts were minor and therefore, could be neglected. However, the radiant heat transfer from the main reflector to the subreflector was incorporated into the thermal model. For the struts, it was assumed that at their ends where they are jointed to the main reflector and subreflector the temperature was undisturbed; i.e., the strut did not significantly affect the local temperature distribution on the reflector. This localized thermal effect appears to be minor but requires further evaluation.

For the main reflector and subreflector, heat is conducted through the face sheets as well as the honeycomb core in three directions: normal, circumferential and radial. For a composite structure, the thermal conductance through each component was added to form a single thermal resistance element in the thermal network. For the antenna designs studied, the temperature difference across the reflector wall was calculated to be of the order of 1 to 4°F. Therefore, thermal resistance in the normal direction was neglected. The struts behave as thermal fins inadvertently and transfer heat either to or from the main reflector or subreflector. There is no significant exchange of heat between the main reflector and subreflector through the struts since the effective thermal conductivity of the strut is small, 0.60 BTU/hr.ft.$^{\circ}$F, compared to the thermal conductivity of reflector materials. Therefore, relative thermal isolation is afforded at the strut-reflector interface.

For all design studies the antenna was assumed to have a surface coating with an emissivity value of 0.90 and to radiate thermal energy to deep space. For those antenna surfaces that exchange radiant energy with other surfaces as well as deep space, the appropriate view factors were determined.

The concave surface of the reflector exchanges thermal radiation with itself. Radiant energy from the main reflector to the struts and subreflector was approximated based upon mean temperatures and overall view factors. Radiant solar energy was absorbed by that portion of the antenna exposed to solar illumination.

7.4.2 Thermal Model

Due to thermal symmetry, half the antenna was modelled. The main reflector was subdivided into 85 sections, each section being represented by a node. Nodes in the circumferential and axial (Z) directions are thermally connected by thermal resistances. Each node is thermally connected to deep space by means of a radiant interchange factor which accounts for the radiant transfer of energy from both sides of the reflector. Since the temperature difference across the reflector wall is small, only one node was used to represent the temperature of a given section. The RF package, located on the convex side of the main reflector, was assumed to be thermally isolated from the antenna. Solar energy was absorbed by those nodes that are illuminated. The thermal model also predicted the radiant energy incident upon the subreflector. The subreflector was modelled in a similar manner to that of the main reflector with the two exceptions: 1) the subreflector was subdivided into 19 sections and 2) the heat input from the main reflector was uniformly distributed over the convex surface of the subreflector.

The maximum (minimum) temperature for all antenna designs occurs at the radial edge at an angle of 0° (180°). The difference between these two temperatures is the maximum temperature difference for the reflector. Results for the three cases considered are shown below:

	Maximum Temperature Difference (°F)	
Case	Main Reflector	Subreflector
Graphite/Epoxy	354	178
Invar	353	158
Beryllium	306	18

The variation of the maximum temperature difference with the thermal conductivity of the face sheet material, at a given face sheet thickness is shown in Figure 7-6 for both the main reflector and the subreflector. The main reflector temperature difference is primarily established by the exchange of radiant energy on the concave surface of the reflector. Therefore, thermal conductivity and thickness are weak thermal control parameters for this reflector. The results shown in Figure 7-6 are representative of antenna designs which utilize either invar or graphite/epoxy face sheets.

The temperature profile around the circumference of a tubular strut has been predicted for the condition of maximum solar illumination when the solar vector and tube center-line are normal. For this condition, the largest temperature difference diametrically across the tube was 36°F and represents the maximum expected temperature difference. This result has been taken into account in the structural analyses discussed in the next section.

7.4.3 Antenna Pattern Analysis

A computer program used to transform the antenna geometry and feed into far-field patterns and gain has been developed by Philco-Ford and is known as SECPAT (Secondary Patterns). It transforms a subreflector scatter pattern into an aperture amplitude and phase distribution from which it computes far-field patterns and gain. This computation consists of fitting Fourier series as functions of circumference for annuli in the aperture plane, such that the specific phase errors calculated will be satisfied, the Fourier series then representing phase functions. Integration over the entire surface in polar coordinates results in a radiation pattern and gain. The amplitude function assumed is, of course, included in the integration. Blockage effects, reflector errors, including RMS surface and specified distortions, and spillover, are included. A high degree of accuracy (≈ 0.01 dB) is achieved out to the third sidelobe; this has been determined by analyzing cases for which closed-form solutions exist and comparing the SECPAT and closed-form results.

- Steady State
- Aluminum Honeycomb

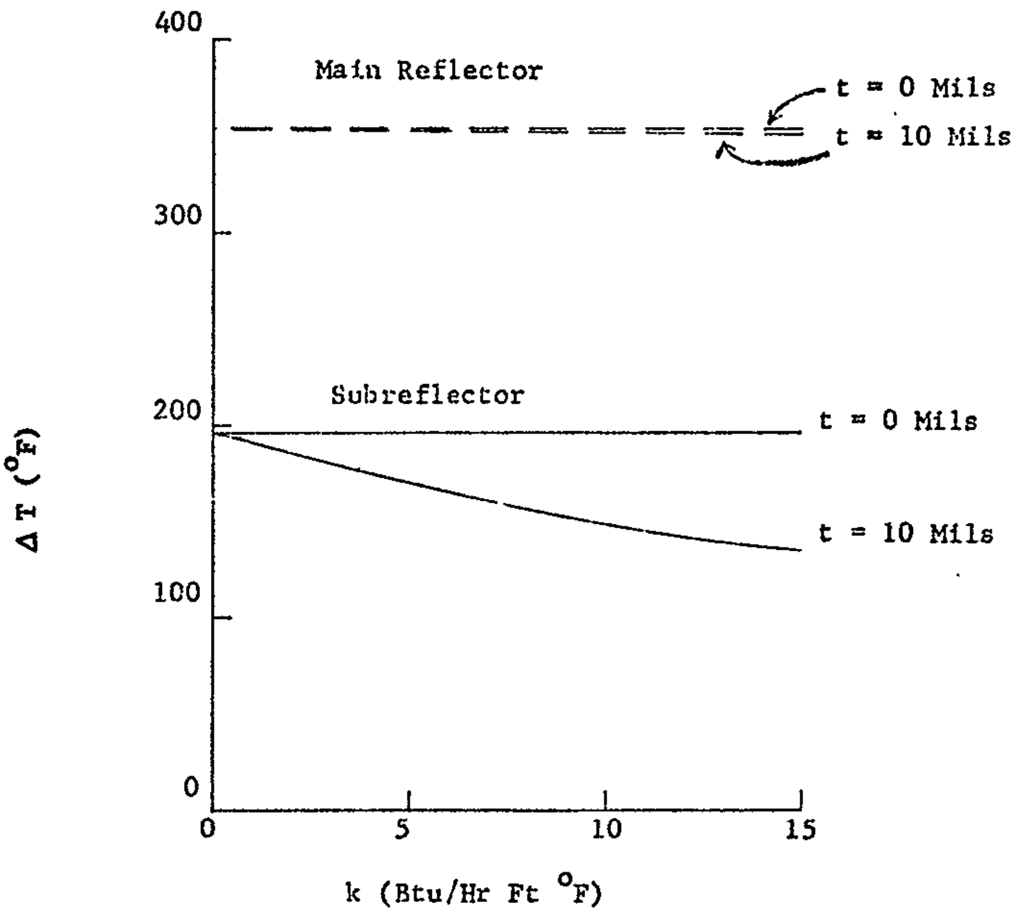


Figure 7-6 MAXIMUM TEMPERATURE DIFFERENCE (Δ T) VERSUS THERMAL CONDUCTIVITY (k) OF FACE SHEET MATERIAL FOR CONSTANT FACE SHEET THICKNESS (t)

The following assumptions were made for the analysis of the distorted dishes described above:

1. Parabolic main reflector
2. Uniform illumination of the main reflector aperture (very closely approximated with high-efficiency surfaces)
3. No subreflector or feed phase errors (automatically compensated for in the high efficiency design)
4. 0.002" RMS surface error with large correlation internal.

Computed radiation patterns in the plane containing the sun for the three cases of interest (carbon, invar and beryllium) plus a "no sun" case may be found in Figure 7-7 a-d. It is apparent that the graphite/epoxy antenna experiences the least gain loss and beam squint of the three cases analyzed.

7.4.4 Worst Case Thermal Environment Definition

Detailed studies of the thermal distortion of parabolic antennas performed by Philco-Ford under contract to NASA/Lewis Research Center (NAS 3-11525) have shown that maximum distortion of the antenna reflector occurs at the solar aspect that produces the maximum temperature difference across the surface of the antenna. The temperature difference for a typical parabolic antenna is shown in Figure 7-8 as a function of sun angle, θ. θ is defined as the angle between the solar vector and the aperture plane of the reflector and has a 24-hour period of revolution at synchronous altitude. The maximum temperature difference is shown to occur near $\theta = 0^o$ and this orientation has been assumed for all antenna thermal analyses performed for this study. The synchronous altitude case was selected because of the obvious utility of this orbit for such an application.

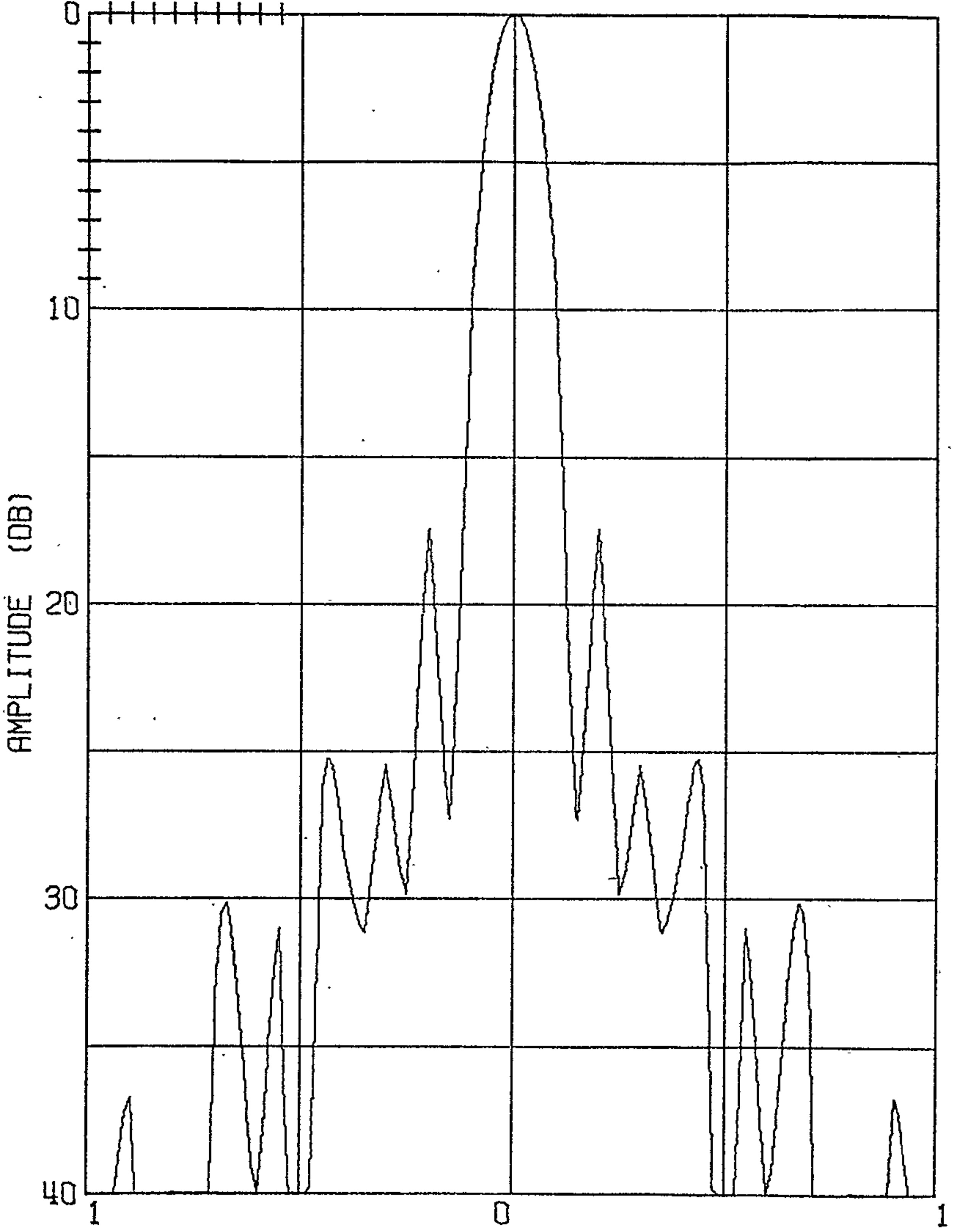


Figure 7-7 a

Calculated Radiation Pattern for 8' Dish at 60 GHz - No Sun

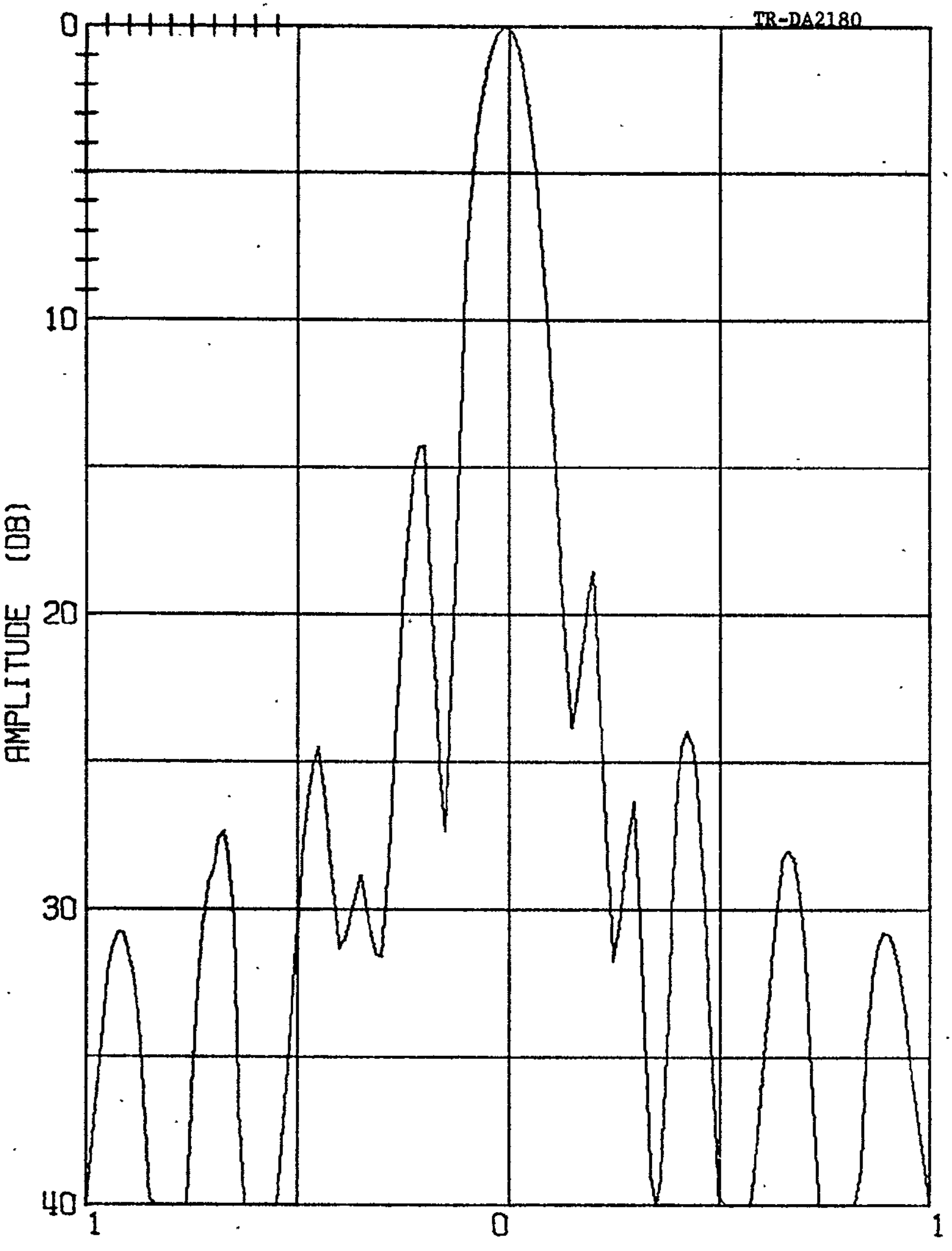


Figure 7-7 b

Calculated Radiation Pattern for 8' Dish at 60 GHz -
Worst Case Sun, Graphite/Epoxy

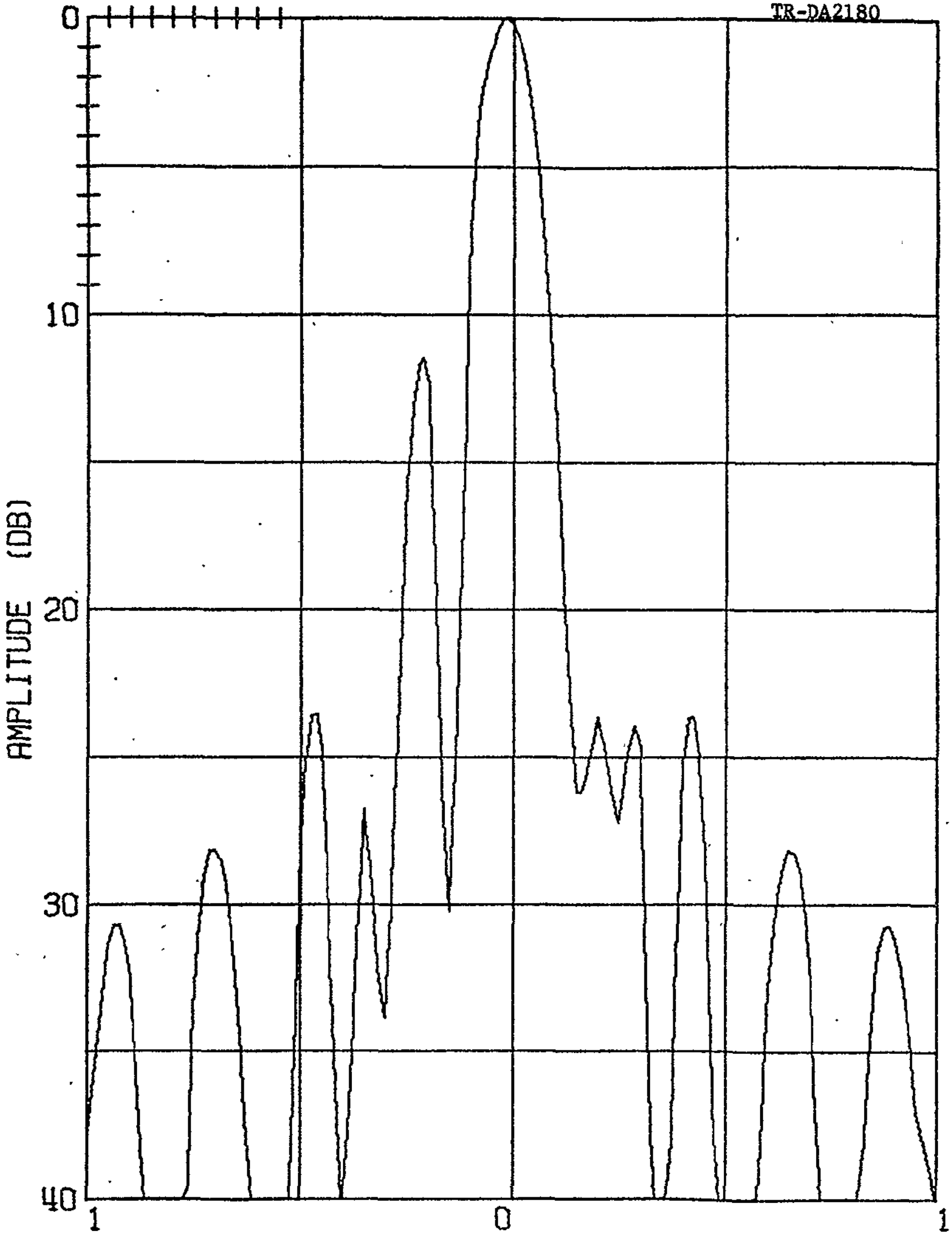


Figure 7-7 c
Calculated Radiation Pattern for 8' Dish at 60 GHz -
Worst Case Sun, Invar

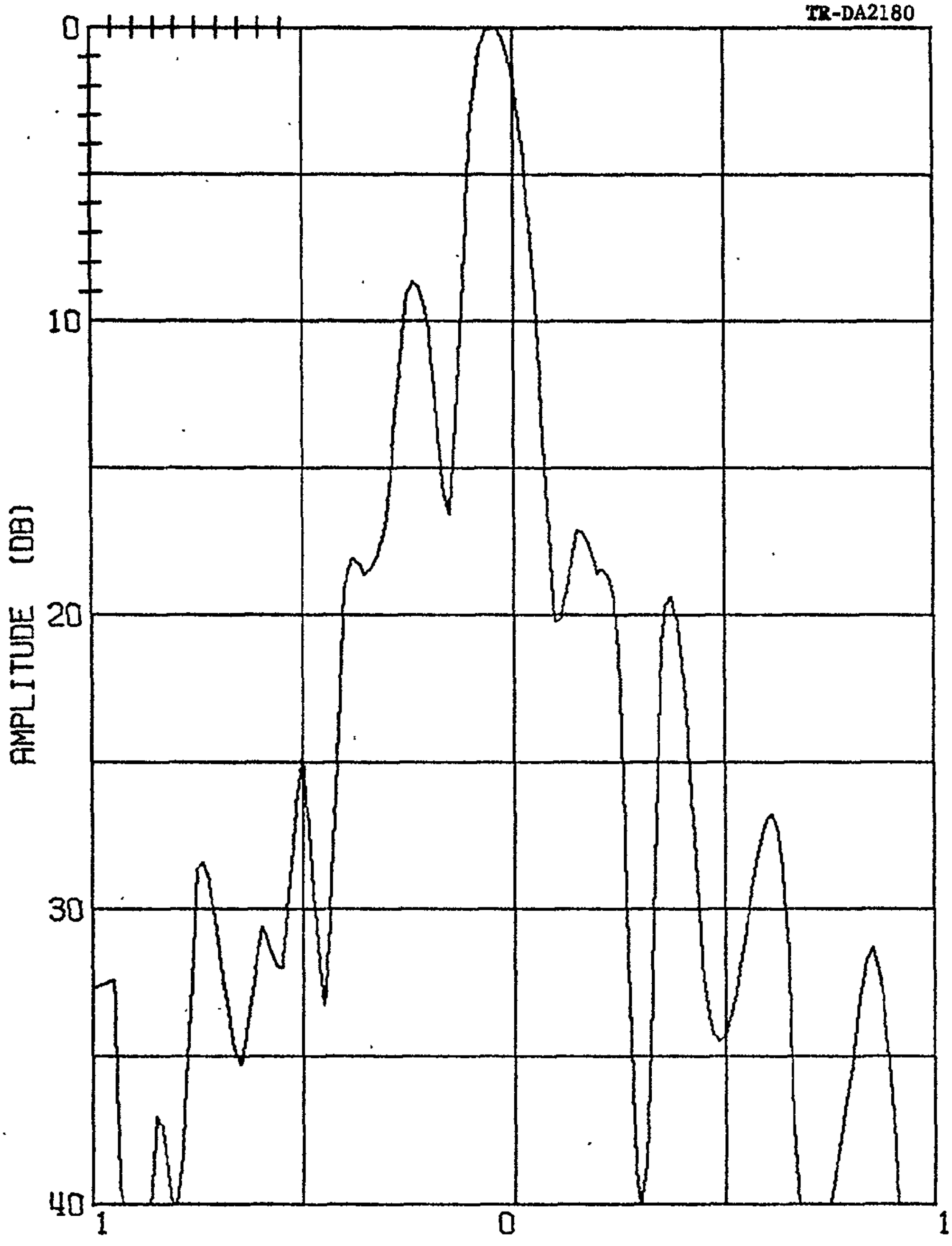


Figure 7-7 d

Calculated Radiation Pattern for 8' Dish at 60 GHz -
Worst Case Sun, Beryllium

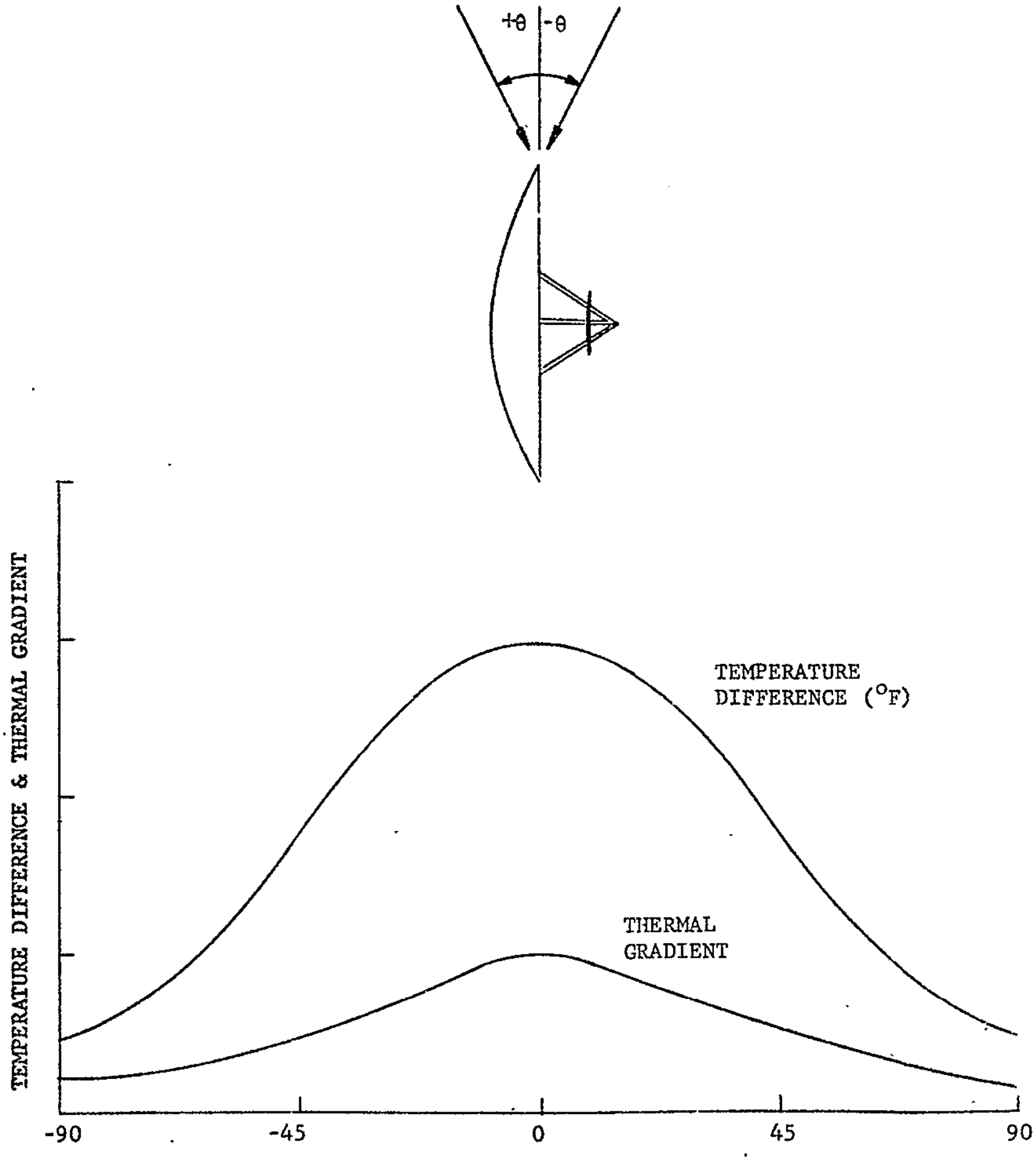


FIGURE 7-8 TEMPERATURE DIFFERENCE AND GRADIENT VERSUS SUN ANGLE

7.5 TRACKING FEEDS IN HIGH GAIN MILLIMETER WAVE ANTENNAS

A number of feed configurations have been developed over the years to provide analog error signals simultaneously with main beam patterns from a common reflector. Figure 7-9 illustrates the feed system used with a multi-mode feed.

A summary of several feed types as regards main beam and side lobe efficiency are listed in Tables 7-2 and 7-3.

Multiple horn and multimode feeds are both rather complex and difficult to fabricate at millimeter wavelengths because of the tolerances involved in both path lengths and angular alignment. The design of a truly optimized millimeter tracking feed for spacecraft use would incorporate considerations such as these, and would hopefully result in a relatively simple, multi-mode structure.

The classic four-horn feed uses four identical horns in a cluster about the antenna focal point; the outputs of these horns are combined with four-port hybrid junctions in such a way that both the sum of all four horns and the difference between elevation and azimuth-pairs are obtained.

Four-horn feeds generally suffer from the fact that the arrangement which yields good tracking information seldom results in the maximum on-axis gain for the sum pattern. For this reason, the newer designs have begun employing separate feed systems for tracking and main-beam formation. Thus, for example, a five-horn cluster may be used with the center horn optimized for main beam gain by the use of higher order modes and the four outer horns designed for the best possible tracking performance. The five-horn feed has an advantage at millimeter wave frequencies of requiring a comparator using minimum waveguide lengths and therefore of minimum loss.

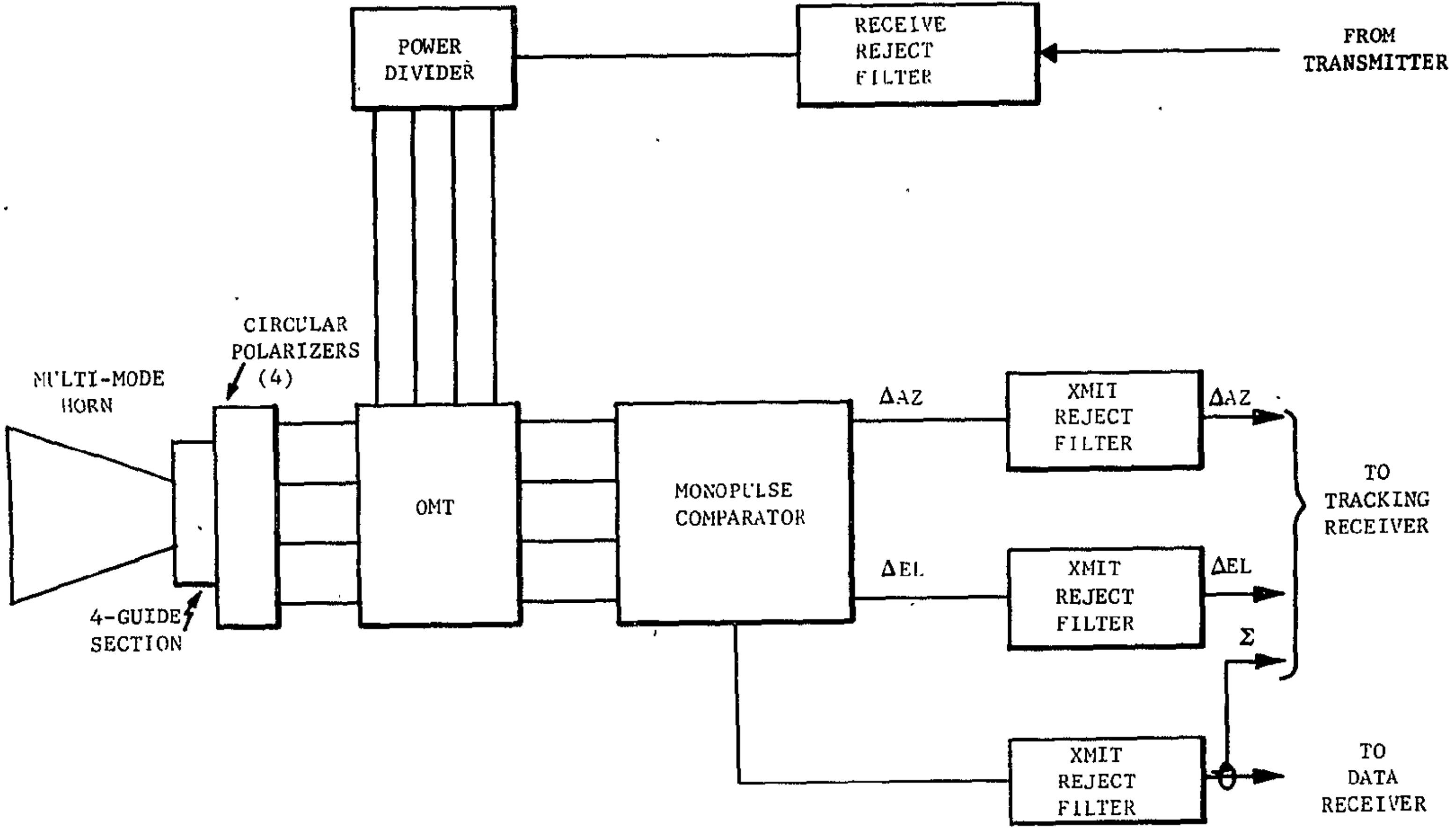

POWER DIVIDER
RECEIVE REJECT FILTER
FROM TRANSMITTER
CIRCULAR POLARIZERS (4)
MULTI-MODE HORN
4-GUIDE SECTION
OMT
MONOPULSE COMPARATOR
ΔAZ
ΔEL
XMIT REJECT FILTER
XMIT REJECT FILTER
XMIT REJECT FILTER
ΔAZ
ΔEL
Σ
TO TRACKING RECEIVER
TO DATA RECEIVER

Figure 7-9 Antenna Feed System Block Diagram

TABLE 7-2

AVERAGE BEAM EFFICIENCY OF FEED HORNS

Feed Configuration	10 dB Beam Efficiency	Main Beam Efficiency	Sidelobe Efficiency
4-Horn	58%	65%	37%
5-Horn (Sum outer 4 horns only)	---	27%	73%
5-Horn (Sum 5 horns)	41%	54%	46%
Multi-mode (Suppressed sidelobes)	91%	99%	1%
Multi-mode (No sidelobe suppression)	80%	97%	3%

TABLE 7-3

TRACKING FEED CHARACTERISTICS

Feed Configuration	Aperture Efficiency*	Tracking Gain Relative to Main Beam (Typical)
4-Horn	50%	-5 dB
5-Horn optimized center horn for main beam, outer 4 horns for tracking	78%	-15 dB
Multi-mode Tracking Horn	80%	-2 dB

* Includes Cassegrain efficiency and spillover factors.

Tracking feeds employing a single aperture (multi-mode) which would simultaneously afford maximum beam gain and maximum tracking gain have been constructed and have found wide applications, particularly at 2.2 and 4 and 6 GHz. Multi-mode feeds have been constructed by Philco-Ford at millimeter wavelengths of both 24 GHz and 60 GHz, and tracking characteristics comparable to presently used multi-mode tracking feeds in the microwave frequency regions have been achieved. While the efficiency characteristics of the multi-mode feed exceed those of multiple-horn feeds, the comparator circuit required can require a large number of waveguide components and thereby be significantly lossy.