INFORMATION TO USERS

This dissertation was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again - beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

University Microfilms
300 North Zeab Road
Ann Arbor, Michigan 48106
A Xerox Education Company
RIEDEL, Nelson Andrew, 1940-
DIGITALLY COMPANDED DELTA MODULATION.

The Ohio State University, Ph.D., 1972
Engineering, electrical

University Microfilms, A XEROX Company, Ann Arbor, Michigan
DIGITALLY COMPANDED
DELTA MODULATION

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of the Ohio State University

By


The Ohio State University
1972

Approved by

C. E. Warren
Adviser
Department of Electrical Engineering
PLEASE NOTE:

Some pages may have

indistinct print.

Filmed as received.

University Microfilms, A Xerox Education Company
ACKNOWLEDGMENT

I wish to express my sincere appreciation to my wife Marianne without whose continued encouragement these years of graduate study could not have been completed. To Professor C. E. Warren, my adviser, I offer sincere gratitude for his many helpful suggestions. The support granted by Bell Telephone Laboratories is gratefully acknowledged.

N. A. Riedel
VITA

October 14, 1940 — Born - Seville, Ohio

1960 - - - - - - - - Assoc. Elec. Tech., DeVry Technical Institute, Chicago, Illinois

1960-1963 - - - - - - Technician, Bell Telephone Laboratories, Columbus, Ohio

1963-1967 - - - - - - Engineer, Ohio Bell Telephone Company, Columbus, Ohio

1967 - - - - - - B.E.E., The Ohio State University, Columbus, Ohio

1967 - - - - - - M.Sc., The Ohio State University, Columbus, Ohio

1967 - - - - - - Member, Technical Staff, Bell Telephone Laboratories, Columbus, Ohio

FIELDS OF STUDY

Major Field: Electrical Engineering

Studies in Communication Theory - Professor C. E. Warren
Studies in Coding and Logic - Professor R. B. Lackey
Studies in Modern Control Theory - Professor H. Hemami
Studies in Mathematics - Professor C. Saltzer
Studies in Statistics - Professor R. C. Srivastava
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENT</td>
<td>11</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF SYMBOLS AND ABBREVIATIONS</td>
<td>x</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I.  INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II. DISCUSSION OF SPEECH SIGNAL</td>
<td>7</td>
</tr>
<tr>
<td>III. DISCUSSION OF LDM</td>
<td>14</td>
</tr>
<tr>
<td>IV. CHARACTERIZATION OF LDM WITH SINUSOIDAL INPUT SIGNALS</td>
<td>33</td>
</tr>
<tr>
<td>V. DISCUSSION OF COMPANDED DELTA MODULATION</td>
<td>79</td>
</tr>
<tr>
<td>VI. INVESTIGATION OF COMPANDED DELTA MODULATION</td>
<td>93</td>
</tr>
<tr>
<td>VII. SUMMARY</td>
<td>144</td>
</tr>
<tr>
<td>APPENDIX</td>
<td></td>
</tr>
<tr>
<td>A. Simulation of LDM</td>
<td>149</td>
</tr>
<tr>
<td>B. Computation of $\overline{E}(L)$ and Transient Response.</td>
<td>157</td>
</tr>
<tr>
<td>C. Calculation of the mms Error</td>
<td>165</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>171</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>2-1</td>
<td>Published Values of Speech Spectral Density Parameters.</td>
</tr>
<tr>
<td>4-1</td>
<td>Values of ( \frac{3}{A_1^2} R_e(1), (1 + \frac{3}{A_1^2} R_e(1)) ) ( \text{Computed By Simulation} )</td>
</tr>
<tr>
<td>6-1</td>
<td>Two-State Companding Rules.</td>
</tr>
<tr>
<td>6-2</td>
<td>Steady State Operating Point, Two-State Companding Rules - CDM_1.</td>
</tr>
<tr>
<td>6-3</td>
<td>Steady State Operating Point, Two-State Companding Rules - CDM_2.</td>
</tr>
<tr>
<td>6-4</td>
<td>Relation of Counter Length ( m ) to Parameter ( r ) and Attack Time Constant ( r_a ).</td>
</tr>
<tr>
<td>6-5</td>
<td>Computation of ( \langle T^2 \rangle ) for Two-State Companding Rule I.</td>
</tr>
<tr>
<td>6-6</td>
<td>Computation of ( \langle T^2 \rangle ) for Three-State Companding Rule.</td>
</tr>
<tr>
<td>B-1</td>
<td>Equations of ( b_1(k_{1_1}, k_{2_1}) ) for Two-State Companding Rules.</td>
</tr>
<tr>
<td>B-2</td>
<td>Function ( q_1(u_1) )</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>2-1</td>
<td>Speech Spectral Density</td>
</tr>
<tr>
<td>2-2</td>
<td>Speech Spectral Density Model</td>
</tr>
<tr>
<td>3-1</td>
<td>LDM System</td>
</tr>
<tr>
<td>3-2</td>
<td>Comparison, LDM System With Practical and Ideal Integrators</td>
</tr>
<tr>
<td>3-3</td>
<td>Waveforms of LDM System With Single Integration</td>
</tr>
<tr>
<td>3-4</td>
<td>LDM Quantizing Noise Power</td>
</tr>
<tr>
<td>3-5</td>
<td>LDM Feedback Network</td>
</tr>
<tr>
<td>3-6</td>
<td>Frequency Characteristic of Feedback Network</td>
</tr>
<tr>
<td>3-7</td>
<td>Comparison of Quantizing Noise Power of LDM₁ and LDM₂ Systems</td>
</tr>
<tr>
<td>3-8</td>
<td>Known Behavior of LDM₁ System for an 800 Hz Tone Input</td>
</tr>
<tr>
<td>3-9</td>
<td>Low Frequency Noise Caused by Unequal Step Sizes</td>
</tr>
<tr>
<td>3-10</td>
<td>Low Frequency Noise Caused by Comparator Offset Voltage</td>
</tr>
<tr>
<td>4-1</td>
<td>LDM₁ Encoder-Decoder</td>
</tr>
<tr>
<td>4-2</td>
<td>Illustration of $S_{20}$ and $\beta$</td>
</tr>
<tr>
<td>4-3</td>
<td>Error Spectral Density for Various Input Loading</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>5-1</td>
<td>Signal-to-Noise Ratio, Linear PCM and Log-PCM</td>
</tr>
<tr>
<td>5-2</td>
<td>Signal-to-Noise Ratio, 7 Bit Log-PCM, CDM₁, CDM₂</td>
</tr>
<tr>
<td>5-3</td>
<td>Analog Companding System</td>
</tr>
<tr>
<td>5-4</td>
<td>Digital Companding System</td>
</tr>
<tr>
<td>5-5</td>
<td>Various Digital Companding Methods</td>
</tr>
<tr>
<td>6-1</td>
<td>Basic CDM System, Two-State Companding Rule</td>
</tr>
<tr>
<td>6-2</td>
<td>Function $\bar{F}(L)$, LDM₁ System</td>
</tr>
<tr>
<td>6-3</td>
<td>Function $\bar{F}(L)$, LDM₂ System</td>
</tr>
<tr>
<td>6-4</td>
<td>Steady State Loading $\mathcal{L}_\gamma$ vs $\gamma$, Two-State</td>
</tr>
<tr>
<td></td>
<td>Companding Rules - CDM₁</td>
</tr>
<tr>
<td>6-5</td>
<td>Steady State Loading $\mathcal{L}_\gamma$ vs $\gamma$, Two-State</td>
</tr>
<tr>
<td></td>
<td>Companding Rules - CDM₂</td>
</tr>
<tr>
<td>6-6</td>
<td>CDM₁ System, $(S/N)$ and $q$</td>
</tr>
<tr>
<td>6-7</td>
<td>CDM₂ System, $(S/N)$ and $q$</td>
</tr>
<tr>
<td>6-8</td>
<td>Step Response of Linear Compandor</td>
</tr>
<tr>
<td>6-9</td>
<td>Companding Characteristic - Linear, Logarithmic and Log-Linear</td>
</tr>
<tr>
<td></td>
<td>Compandors</td>
</tr>
<tr>
<td>6-10</td>
<td>Block Diagram - Proposed Log-Linear Discrete Compandor</td>
</tr>
<tr>
<td>6-11</td>
<td>Three-State Compandor</td>
</tr>
<tr>
<td>6-12</td>
<td>Function $\bar{F}(L)$ for Three-State Companding Rule</td>
</tr>
<tr>
<td>6-13</td>
<td>Steady State Loading $\mathcal{L}_\gamma$ as a Function of $\gamma$</td>
</tr>
<tr>
<td></td>
<td>for Three-State Companding Rule</td>
</tr>
<tr>
<td>6-14</td>
<td>Transient Response CDM₁ - Three-State Companding Rule</td>
</tr>
</tbody>
</table>
6-15 Transient Response CDM₂ - Three-State Companding Rule ........................................................................ 130
6-16 Transient Response, CDM₁ .................................................................................................................. 132
6-17 Effects of an Isolated Channel Error on Encoder-Decoder Tracking .................................................. 138
6-18 Maximum (S/N) as a Function of Data Rate for CDM₁, CDM₂, and PCM ........................................... 141
A-1 Mathematical Model of LDM Encoder .................................................................................................. 150
B-1 Flowchart - Computation of $\overline{B}(L)$ ......................................................................................... 160
C-1 Waveform Ensemble .............................................................................................................................. 166
C-2 Function $X(t) - X(kt)$ ......................................................................................................................... 168
# List of Symbols and Abbreviations

<table>
<thead>
<tr>
<th>Symbol or Abbreviation</th>
<th>Meaning</th>
<th>Introduced In Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (subscript)</td>
<td>Single Integration Delta Modulation</td>
<td>1</td>
</tr>
<tr>
<td>2 (subscript)</td>
<td>Double Integration Delta Modulation</td>
<td>1</td>
</tr>
<tr>
<td>ai</td>
<td>Binary Codes Speech Signal</td>
<td>3.1</td>
</tr>
<tr>
<td>acf</td>
<td>Autocorrelation Function</td>
<td>App. A</td>
</tr>
<tr>
<td># (superscript)</td>
<td>Phase Corrected Variable</td>
<td>App. A</td>
</tr>
<tr>
<td>ba</td>
<td>Binary Companding Signal (Attack)</td>
<td>6.4</td>
</tr>
<tr>
<td>bi</td>
<td>Binary Coded Companding Signal</td>
<td>5</td>
</tr>
<tr>
<td>br</td>
<td>Binary Companding Signal (Recovery)</td>
<td>6.4</td>
</tr>
<tr>
<td>$\bar{B}$</td>
<td>Average Value of Binary Coded Companding Signal</td>
<td>6.2</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Equivalent Normalized Bandwidth</td>
<td>4.2</td>
</tr>
<tr>
<td>C</td>
<td>Amount of Companding</td>
<td>5</td>
</tr>
<tr>
<td>CDM</td>
<td>Companded Delta Modulation</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>Step Size</td>
<td>3.1</td>
</tr>
<tr>
<td>DM</td>
<td>Delta Modulation</td>
<td>1</td>
</tr>
<tr>
<td>DPCM</td>
<td>Differential Pulse Code Modulation</td>
<td>1</td>
</tr>
<tr>
<td>e(t)</td>
<td>Encoder Error</td>
<td>3.1</td>
</tr>
<tr>
<td>$\bar{e}^2$</td>
<td>Modified Mean Squared Error</td>
<td>3.4</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>E</td>
<td>Equivalent Voltage Source</td>
<td>4.1</td>
</tr>
<tr>
<td>f₁, f₂, f₃</td>
<td>Pole and Zero Values of Feedback Network</td>
<td>3.3</td>
</tr>
<tr>
<td>fₐ, fₐ, fₐ</td>
<td>Parameters of Speech Spectral Density Model</td>
<td>2</td>
</tr>
<tr>
<td>fₐ</td>
<td>Speech Bandwidth</td>
<td>2</td>
</tr>
<tr>
<td>f₁</td>
<td>Input Frequency</td>
<td>3.4</td>
</tr>
<tr>
<td>fₙ</td>
<td>Normalized Frequency</td>
<td>4.2</td>
</tr>
<tr>
<td>fₛ</td>
<td>Sampling Frequency</td>
<td>3.4</td>
</tr>
<tr>
<td>γ</td>
<td>Ratio of Recovery to Attack Time Constants</td>
<td>5</td>
</tr>
<tr>
<td>g</td>
<td>Increase in Δ Over One Sampling Period</td>
<td>6.2</td>
</tr>
<tr>
<td>H</td>
<td>Feedback Network Transfer Function</td>
<td>3.1</td>
</tr>
<tr>
<td>H₂</td>
<td>Second Integrator Transfer Function</td>
<td>4.6</td>
</tr>
<tr>
<td>Hₖ</td>
<td>Companding Filter Transfer Function</td>
<td>6.2</td>
</tr>
<tr>
<td>i</td>
<td>Current Source</td>
<td>4.1</td>
</tr>
<tr>
<td>i (subscript)</td>
<td>Value of Variable at ith Sample</td>
<td>App. A</td>
</tr>
<tr>
<td>k</td>
<td>Feedback Network Gain</td>
<td>3.3</td>
</tr>
<tr>
<td>kᵢ</td>
<td>Number of Samples Since the ith Past Change in the Binary Coded Speech Signal</td>
<td>App. A</td>
</tr>
<tr>
<td>L</td>
<td>Input Loading (LDM)</td>
<td>3.4</td>
</tr>
<tr>
<td>Lₒ</td>
<td>Optimum Loading Point</td>
<td>3.4</td>
</tr>
<tr>
<td>Lₖ</td>
<td>Steady-State Loading Point (CDM)</td>
<td>5</td>
</tr>
<tr>
<td>LDM</td>
<td>Linear Delta Modulation</td>
<td>1</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
<td>Value</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>( m )</td>
<td>Companding Counter Length</td>
<td>6.3</td>
</tr>
<tr>
<td>( \mu )</td>
<td>PCM Companding Constant</td>
<td>5</td>
</tr>
<tr>
<td>( N )</td>
<td>Audio Noise Power</td>
<td>3.4</td>
</tr>
<tr>
<td>( N_g )</td>
<td>Granular Noise</td>
<td>3.2</td>
</tr>
<tr>
<td>( N_o )</td>
<td>Overload Noise</td>
<td>3.2</td>
</tr>
<tr>
<td>( N_q )</td>
<td>Quantizing Noise</td>
<td>3.2</td>
</tr>
<tr>
<td>PCM</td>
<td>Pulse Code Modulation</td>
<td>1</td>
</tr>
<tr>
<td>( q )</td>
<td>Companding Coefficient</td>
<td>5</td>
</tr>
<tr>
<td>( q_{\text{max}} )</td>
<td>Maximum Value of Companding Control Signal</td>
<td>5</td>
</tr>
<tr>
<td>( Q )</td>
<td>Input Loading (CDM)</td>
<td>5</td>
</tr>
<tr>
<td>( r )</td>
<td>Ratio of Adjacent Step Sizes (CDM)</td>
<td>6.3</td>
</tr>
<tr>
<td>( (S/N) )</td>
<td>Signal-to-Noise Ratio</td>
<td>3.4</td>
</tr>
<tr>
<td>( (S/N)_0 )</td>
<td>Theoretical Maximum Signal-to-Noise Ratio</td>
<td>4.4</td>
</tr>
<tr>
<td>( (S/N)_{\text{max}} )</td>
<td>Maximum Signal to Noise Ratio</td>
<td>3.4</td>
</tr>
<tr>
<td>( Re(t) )</td>
<td>Error Autocorrelation Function</td>
<td>3.4</td>
</tr>
<tr>
<td>( Se(f) )</td>
<td>Error Spectral Density</td>
<td>3.4</td>
</tr>
<tr>
<td>( Sn(f) )</td>
<td>Normalized Error Spectral Density</td>
<td>4.2</td>
</tr>
<tr>
<td>( Sno )</td>
<td>Average Value of Normalized Error Spectral Density in Audio Band</td>
<td>4.2</td>
</tr>
<tr>
<td>( S_x(f) )</td>
<td>Input Spectral Density</td>
<td>3.4</td>
</tr>
<tr>
<td>( S_{\text{e}}^{*}(f) )</td>
<td>Signal Estimate Spectral Density</td>
<td>3.4</td>
</tr>
<tr>
<td>( T )</td>
<td>Tracking Constant</td>
<td>6.5</td>
</tr>
<tr>
<td>( \tau )</td>
<td>Sampling Period</td>
<td>4.1</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
<td>Section</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>$T_a$</td>
<td>Attack Time</td>
<td>6.1</td>
</tr>
<tr>
<td>$\tau_a$</td>
<td>Attack Time Constant</td>
<td>5</td>
</tr>
<tr>
<td>$T_r$</td>
<td>Recovery Time</td>
<td>6.1</td>
</tr>
<tr>
<td>$\tau_r$</td>
<td>Recovery Time Constant</td>
<td>5</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Phase Random Variable</td>
<td>App. A, C</td>
</tr>
<tr>
<td>$U(k)$</td>
<td>The Change in the Signal Estimate Due to a Constant Binary Signal of Duration $kT$</td>
<td>App. A</td>
</tr>
<tr>
<td>$u$</td>
<td>Number Stored in Companding Counter</td>
<td>6.3</td>
</tr>
<tr>
<td>$X$</td>
<td>RMS Value of Input</td>
<td>3.4</td>
</tr>
<tr>
<td>$X$</td>
<td>RMS Value of Derivative of Input Signal</td>
<td>5</td>
</tr>
<tr>
<td>$X_{max}$</td>
<td>RMS Design Maximum of Derivative of Input Signal</td>
<td>5</td>
</tr>
<tr>
<td>$X_m$</td>
<td>RMS Maximum Input Signal</td>
<td>3.4</td>
</tr>
<tr>
<td>$X_{max}$</td>
<td>RMS Design Maximum Input Signal (CDM)</td>
<td>5</td>
</tr>
<tr>
<td>$X(t)$</td>
<td>Input Signal</td>
<td>3.1</td>
</tr>
<tr>
<td>$X(t)$</td>
<td>Derivative of Input Signal</td>
<td>5</td>
</tr>
<tr>
<td>$\hat{X}(t)$</td>
<td>Signal Estimate</td>
<td>3.1</td>
</tr>
<tr>
<td>$y$</td>
<td>Time Random Variable</td>
<td>App. C</td>
</tr>
<tr>
<td>$Y(t)$</td>
<td>Decoder Output Signal</td>
<td>3.1</td>
</tr>
<tr>
<td>$z$</td>
<td>Error Random Variable</td>
<td>App. C</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

Since the invention of Pulse Code Modulation (PCM), numerous systems have been proposed for transmitting continuous messages using discrete or digital signals. In general, all of these systems consist of: (1) an encoder, transmitter or modulator for transforming the continuous message to a discrete signal; (2) a digital channel which conveys the discrete signal to the decoder; and (3) a decoder, receiver or demodulator which transforms the digital signal back to the continuous message. Historically, there have been two major advantages in using digital transmission systems. First, the digital pulse train can be transmitted over long distances and be regenerated without the accumulation of distortion and noise as occurs in analog transmission. Secondly, signals from radically different sources all look the same to the transmission system, permitting the transmission system to become a data line with no special consideration to whether the original signals are voice, video, data, etc., or a combination of these. Recently, a third and quite possibly the most important advantage has become apparent; it is the ease with which the telephone switching function can be combined with the transmission function. Currently,
in the telephone system, with the exception of the 101 ESS System\(^{(22)}\) which uses pulse amplitude modulation all switching is done using analog signals and a mechanical contact network. With recent advances in the computer-data processing field and the integrated circuit art, a switching network using digital solid-state devices seems promising from both a cost and reliability standpoint. Such a system, if implemented on a nationwide basis would permit encoder-decoders at the source and destination only, with transmission and switching systems being a merged digital data network. However, these advantages are not achieved without penalties. First, the transformation of the continuous message to a digital form and then back to the continuous form causes a degradation of the received message. This distortion is caused by forcing the continuous message to take on discrete values and is termed quantization noise. The quantization noise can be made arbitrarily small by using a finer quantization; however, this is at the expense of a greater channel bandwidth. For example, PCM coding of a speech signal of bandwidth \(f\), sampled at the Nyquist rate of \(2f\) and coded in seven\(^1\) bits requires a bit rate of \(14f\). Thus, the 3200 Hz speech band requires a bit rate of \(14 \times 3200\) or 44,800 bits per second. The bandwidth required to transmit a digital pulse train is at least one-half the bit rate. Therefore, the bandwidth required is seven times the bandwidth of the original signal. A second penalty is the cost of the coding and decoding hardware required for conversion.

\(^1\)PCM using levels that are equally spaced on a logarithmic scale requires at least 128 levels for good quality transmission.
between continuous and digital form. The challenge is the simultaneous minimization of the bandwidth and encoder-decoder cost while maintaining acceptable message quality.

Linear Delta Modulation (LDM) is the simplest form of a Differential Pulse Code Modulation (DPCM) System. These systems are based on the inventions of Debraine et al, deJagar, and Cutler. DPCM systems quantize the difference between successive sample values rather than the sample values themselves. One of the major characteristics of DPCM systems is that they are limited on the slope (derivative) of the input signal rather than the signal amplitude as is the case with PCM systems. This slope-limiting is well matched to speech which also tends to be slope-limited. When the quantizer is limited to only two levels, the system is called Delta Modulation, or DM. Both the modulator and demodulator make an estimate based on the present and past transmitted signals. In linear DM, (LDM) the value of the signal estimate at each sample time is a linear function of the quantized signal. The linear circuits used for transforming the quantized samples into the signal estimate fall into two types; a simple integrator, and a simple integrator followed by a second integrator which consists of a pole near the highest frequency in the signal band and a zero between the signal band and the sampling frequency. These systems are termed single and double integration delta modulation, respectively, and will be referred to as DM₁ and DM₂.

---

²The subscripts 1 and 2 are used on many variables and abbreviations and with the exception of the variables f₁ and f₂ refer only to single and double integration systems, respectively.
The literature abounds with analysis, modifications, and applications of LDM. (See References.) These references include data on sinusoidal input signals, gaussian and exponentially distributed input signals, broadband, television and speech signals, and feature analytical, simulation as well as measured results. Most of the results are lacking in that they fail to give simple closed form solutions for the nonideal integrator delta modulators together with the error spectral density for sinusoidal input signals. One of the objectives of this dissertation is to develop a useful closed form approximation of the noise performance of the LDM system for sinusoidal signals. By the use of simple expressions some insight can be gained into the operation of LDM with emphasis toward characterizing companded systems.

The noise characteristics of LDM are very sensitive to changes in the slope of the input signal. As a result, optimum signal-to-noise ratio operation is limited to a very narrow range of slope variation. Fortunately, speech tends to have constant rms slope over the audio frequency range for a signal with constant average power. However, the average power varies over a wide range causing a severe restriction on the usefulness of LDM. Companding is the process of adjusting the system parameters based on the signal parameters such that maximum signal-to-noise ratio is maintained over a wide range of input signals. It is well known that (see References 1, 2, 6, 19, 24, 36) companding can extend the useful range of the DM system sufficiently so that signal power variations are not a restriction.
The second purpose of this study is to investigate various means of determining the average loading (input signal level) of the system relative to the optimum point from the binary coded DM signal. This indication of loading, called the companding signal, is then used to adjust the quantizer step size of the DM system such that the system operates near the peak efficiency point. Such a system is abbreviated CIM. Some work has been done in this area including a system studied by Abate\(^2\) which changes the step size based on the two most recent quantized sample values in an ideal single integration delta modulator and a system proposed by Greifkes\(^19\) which adjusts the step size based on the last four quantized sample values in a double integration delta modulator. In addition to these digital companding methods, at least one analog companding system proposed by Brown and Brolin\(^6\) has appeared in the literature. Emphasis in this study is placed on answering the following questions: (1) what is the optimum number of samples that should be examined; (2) what is the best method to generate the companding signal; and (3) is this different for the single and double integration systems? A third and final purpose of this paper is to compare the performance of linear and companded delta modulation with that of PCM and with published data, where available, on the other companded delta modulation systems mentioned above. Some of the published results are for input signals with either gaussian or exponential probability distributions. Since this work is based entirely on sinusoidal input signals, a summary of the known statistical characteristics of speech
together with the relationship of the system performance for stochastic and sinusoidal input signals is included.
CHAPTER II

DISCUSSION OF SPEECH SIGNAL

The short-time speech signal corresponding to a single voiced sound is typically periodic,\(^3\) and has a line spectrum consisting of a fundamental frequency and harmonics which diminish in amplitude approximately as \(1/f^2\). Davenport\(^{10}\) has found from measurements that the duration of the basic voice pattern corresponding to a single voiced sound ranges from 4 msec to 9 msec with an average of about 7.6 msec.

The long-time speech process has been modeled as a stationary stochastic process. The amplitude probability distribution has been found to be approximately exponential by both Davenport and McDonald\(^{29}\). The spectral density from published data is shown in Figure 2-1. Curve a is the result published by Dunn and White\(^{13}\) for both men and women speakers that was modified by McDonald to reflect the attenuation of the local telephone plant and a telephone set. Curve b is the spectral density of a 5 second telephone signal which was also published by

\(^3\)Flanagan et al\(^{14}\) gives an excellent description of the short-time speech signal which includes graphs of the major spectral lines as a function of time for typical voiced sounds.
FIGURE 2 - 1

SPEECH SPECTRAL DENSITY:
(a) PUBLISHED DATA, MANY SPEAKERS;
(b) PUBLISHED DATA, SINGLE SPEAKER;
(c) MODEL
McDonald. Curve c is the spectral density model used by Brolin and Brown(6) in their study of CDM. The curves are different in detail; however, they do have the same shape and are sufficiently similar to establish curve c as a useful model.

The lower and upper cutoff frequencies for the audio band and the corner frequency for the 6 dB falloff of the spectral density has varied from author to author. Table 2-1 shows representative data from the literature for $f_a$, the lower cutoff frequency, $f_c$ the upper cutoff frequency, and $f_b$ the corner frequency. Throughout this paper the values of $f_a = 250$ Hz, $f_b = 800$ Hz, and $f_c = 3.1$ KHz will be used. This spectral density is shown on a logarithmic scale in Figure 2-2. The audio bandwidth, denoted by $f_d$ is defined as

$$f_d = f_c - f_a.$$  

(2-1)

In this paper a sinusoidal signal is used to characterize the DM system. This choice was made mainly to simplify measurement. However, it is not unreasonable considering that the power spectra of the short-time speech signal consists of discrete lines. Historically, a signal of approximately 800 Hz has been used to compare slope limited and amplitude limited systems. This originated with deJagars'\(^{(11)}\) observation that if, for an 800 Hz test signal, a slope limited and an amplitude limited system exhibit identical signal-to-noise ratios, and if neither system is overloading, then performance of the two systems to speech signals is approximately the same. For a first approximation, the LDM system can be viewed as a system which adds a constant power
<table>
<thead>
<tr>
<th>AUTHOR</th>
<th>LOW FREQUENCY CUTOFF $f_a$</th>
<th>CORNER FREQUENCY $f_b$</th>
<th>HIGH FREQUENCY CUTOFF $f_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abate(1)</td>
<td>0</td>
<td>0.23 $f_m$</td>
<td>Variable - $f_m$</td>
</tr>
<tr>
<td>Brolin &amp; Brown(6)</td>
<td>250 Hz</td>
<td>800 Hz</td>
<td>3.1 KHz</td>
</tr>
<tr>
<td>Greefkes &amp; Riemens(20)</td>
<td>200 Hz</td>
<td>500 Hz</td>
<td>3.6 KHz</td>
</tr>
<tr>
<td>Tomozana &amp; Kaneko(36)</td>
<td>300 Hz</td>
<td>-</td>
<td>3.4 KHz</td>
</tr>
</tbody>
</table>
FIGURE 2-2 SPEECH SPECTRAL DENSITY MODEL

- $f_a = 250\,\text{Hz}$
- $f_b = 800\,\text{Hz}$
- $f_c = 3.1\,\text{kHz}$
flat spectrum noise\textsuperscript{4} signal to the input, independent of the input waveform, provided the system is not overloaded. The fact that the noise is independent of the input is the basis for comparing the performance for an 800 Hz tone with that reported by Abate\textsuperscript{(2)} for simulated systems using stochastic input signals. This is not to be interpreted as meaning that the signal-to-noise ratios are comparable. This point is discussed in more detail in Chapter \textsuperscript{4}.

Although the model for speech assumes a stationary process, in actual telephone systems the average power varies widely. This is due to several causes, including different speakers, different losses in the individual subscriber telephone plant and the nature of the speech process. Previous studies of CDM have assumed that, in terms of the 800 Hz test signal, a power range of 40 to 50 dB is representative.\textsuperscript{(6,19,28,36)} For the purposes of this paper, the input signal range of the 800 Hz test tone is from 15 millivolts to 4.2 volts rms. This approximately corresponds to -35 dBm to +15 dBm for a 600 ohm termination, or a range of 50 dB. This choice of voltage range was dictated by on-hand integrated circuits and is about 6 dB high for telephone speech. This is not to be considered a constraint since all the results are normalized and, by using appropriate circuits, apply over any equivalent range.

\textsuperscript{4}For the purpose of this paper, noise is defined as that portion of the error spectral density which falls within the frequency range of the input signal, approximately the range $f_a$ to $f_c$. In this range, all noise is weighted equally, independent of frequency.
The 50 dB range is taken in the same sense as Reference 35. That is, the average level of the signal is taken to be 30 dB below the full load test signal with the weakest signal down 50 dB and the strongest signal down 10 dB. The 10 dB at the high end is to allow for variations between the average signal and the maximum instantaneous signal.
CHAPTER III

DISCUSSION OF LDM.

3.1 System Model

The basic LDM encoder consists of a two level quantizer, a sample-hold circuit and a linear feedback network, as illustrated in Figure 3-1. The quantizer produces an output of either +1 or -1, depending on whether the sign of the error \( e(t) \) at the sampling time is plus or minus. The error \( e(t) \) is given by

\[
e(t) = X(t) - \hat{X}(t),
\]

where \( X(t) \) is the input signal and \( \hat{X}(t) \) is the decoded signal estimate. The sample-hold circuit samples the output of the quantizer at frequency \( f_s \) and holds this value between sampling times. The output of the sample-hold circuit \( a_1 \) is applied to the linear network \( H(f) \) which produces the signal estimate \( \hat{X}(t) \). The decoder consists of the decoding network \( H(f) \) (identical to the feedback network of the encoder) and a low pass filter with the same bandwidth as the input signal. The linear network consists of an amplifier with gain \( k \) and one or more integrators. The signal estimate is a ramp function (the integration
FIGURE 3-1 LDM SYSTEM

ENCODER

INPUT SIGNAL $X(t)$ → $\pm e(t)$ → TWO LEVEL QUANTIZER → SAMPLE HOLD → $a_i$ DIGITAL CHANNEL

CLOCK

LINEAR FEEDBACK NETWORK $H(f)$

DECODER

DIGITAL CHANNEL $a_i$ → LINEAR FEEDBACK NETWORK $H(f)$ (IDENTICAL TO ENCODER) → LOW PASS FILTER → OUTPUT SIGNAL $Y(t)$
of the plus or minus constant amplitude output of the sample hold circuit). Many of the papers published on LDM use a model where the signal estimate is the sum of steps (ideal integration of an impulse) rather than the sum of ramps. Since the major statistic is the error at the sample times only, the two systems behave identically. The step size $\Delta$ is defined as the change in $X(t)$ over one sampling period. The value of $\Delta$ is illustrated for ideal and practical integrators in Figure 3-2.

3.2 Performance, Stochastic Input Signals

For stochastic input signals, the quantizing noise $N_q$ is defined as that portion of the error spectral density that falls within the audio band $f_a$ to $f_c$. In the LDM system the quantizing noise takes on two forms; granular noise $N_g$ resulting from forcing the signal to assume discrete values (i.e., multiples of $\Delta$) and overload noise $N_o$ which occurs when the slope of the signal exceeds the maximum slope that can be reproduced by the system (i.e., the product of the step size and sampling frequency). The quantizing noise is the sum of the granular noise and the slope-overload noise. Granular noise is similar to the quantizing noise in PCM systems, and as in PCM is a monotonic function of the step size, increasing as the step size increases. Slope-overload noise is also a monotonic function of the step size, but increases as the step size decreases, reaching the value of the signal as the step size approaches zero. Typical waveforms for the single integration delta modulation are illustrated in Figure 3-3. The quantizing noise as a function of the step size is
LDM SYSTEM WITH IDEAL INTEGRATOR

LDM SYSTEM WITH PRACTICAL INTEGRATOR

FIGURE 3-2
COMPARISON, LDM SYSTEM WITH PRACTICAL AND IDEAL INTEGRATORS.
FIGURE 3-3 WAVEFORMS OF LDM SYSTEM WITH SINGLE INTEGRATION
illustrated in Figure 3-4. The region where slope-overload noise predominates and the region where granular noise predominates are indicated. From Figure 3-4 it is seen that optimum performance is achieved over a very narrow range of step sizes.

Abate(2) has simulated the ideal integrator LDM system for both exponentially and gaussian distributed signals with spectral densities similar to the model illustrated in Figure 2-2. The results of these simulations show good agreement to those illustrated in Figure 3-4.

3.3 Double Integration LDM System

The feedback network $H(f)$ for the practical single integration system is a simple integrator with a pole located at frequency $f_1$. The value of $f_1$ is well below the lower signal frequency $f_a$. Typically $f_1$ is on the order of 120 Hz. In the double integration DM system, a second integrator consisting of a pole at frequency $f_2$ and a zero at frequency $f_3$ is added following the first integrator. It is not obvious at this point that the zero at $f_3$ is required. In Chapter 4 it is shown by measurement that presence of the zero improves the signal-to-noise ratio. This feedback network is illustrated in Figure 3-5. The purpose of the second integrator is to provide some adjustment in the step size $\Delta$ as a function of the input signal slope which results in an improved signal-to-noise ratio. The zero ($f_3$) is placed well below the sampling frequency, typically at $f_s/6$. The pole ($f_2$) is placed near the upper audio frequency $f_c$. For a signal of large slope, the binary pulse train contains many adjacent pulses
20

SIGNAL POWER

GRANULAR NOISE POWER, $N_G$

OVERLOAD NOISE, POWER, $N_O$

TOTAL $N_Q$

OPTIMUM PERFORMANCE RANGE.

FIGURE 3-4
LDM QUANTIZING NOISE POWER
FIGURE 3-5
LDM FEEDBACK NETWORK

FIGURE 3-6
FREQUENCY CHARACTERISTIC OF FEEDBACK NETWORK
of the same sign (i.e., an overload condition) resulting in the
fundamental frequencies of the triangular waveform signal estimate
being near frequency $f_2$. As a result, the value of $\Delta$, the step for
each sample is attenuated less, and the step is larger. For small
slope signals, the binary pulse train alternates every 2 or 3 samples
(refer to Figure 3-3) resulting in the fundamental frequencies of the
triangular waveform signal estimate being on the order of $f_2/6$ or
greater, and the step size $\Delta$ is at its smallest value. This is a
form of instantaneous companding. While the performance of the DM$_2$
system has not been analyzed in detail, the overall effect of this
second integrator on the noise curves of Figure 3-4 is known to move
the overload noise curve to the left resulting in slightly better
signal-to-noise ratio operation, as well as broadening and flattening
the region of optimum operation. This is illustrated in Figure 3-7.

3.4 Performance, Sinusoidal Signal

The sinusoidal input signal $X(t)$ is defined as

$$X(t) = \sqrt{2}X \sin(2\pi f_1 t)$$

where $X$ is defined as the rms value of the system input voltage and
$f_1$ is defined as the input frequency. The unfiltered output of the
decoder, termed the signal estimate is defined as $\hat{X}(t)$. The process
of encoding - decoding produces an error $e(t)$. This error is defined
in terms of $X(t)$ and $\hat{X}(t)$ as

$$e(t) = X(t) - \hat{X}(t).$$
FIGURE 3-7
COMPARISON OF QUANTIZATION NOISE POWER OF LDM₁ AND LDM₂ SYSTEMS
The one-sided spectral density of the input signal $S_X(f)$ is an impulse at the input frequency $f_i$. The error spectral density $S_e(f)$ is defined as the difference of the spectral density of the signal estimate $S_X(f)$ and the value of $S_X(f_i)$

$$S_e(f) = S_X(f) - S_X(f_i).$$

Conventionally, the mean squared error is defined as the area under a spectral density which is the Fourier transform of the autocorrelation function (the system is assumed to be ergodic)

$$R_e(v) = \frac{1}{2T} \lim_{T \to \infty} \int_{-T}^{T} e(t) e(t+v) \, dt. \quad (3-5)$$

If any phase shift is introduced by the system, the Fourier transform of $R_e(v)$ will contain an impulse at frequency $f_i$. In audio systems, a listener cannot detect a phase shift. However, a phase shift can radically alter the mean squared error. The error spectral density $S_e(f)$ as defined in equation (3-4) is zero at the input frequency $f_i$. Therefore, phase shifts do not alter the value of $S_e(f)$. Note that $S_e(f)$ is not, in general, the Fourier transform of $R_e(v)$.

In this paper a modified mean squared error criterion (abbreviated mms error) is used. The mms error is given by

$$\bar{e}^2 = \int_0^\infty S_e(f) \, df. \quad (3-6)$$
Since $S_e(f)$ is zero at $f = f_1$, the mms error is independent of any phase shift introduced by the system. The mean squared error and the modified mean squared error are identical only if the phase of the input $X(t)$ and the phase of the component of the signal estimate $\hat{X}(t)$ at the input frequency are equal. The mms error is also related to spectral density $S_X(f)$ as

$$\bar{e}^2 = \int_{0}^{f_1-} S_X(f) \, df + \int_{f_1+}^{\infty} S_X(f) \, df. \quad (3-7)$$

The audio noise power $N$ (unit resistances are assumed) is defined as that portion of the error spectral density which lies in the audio range $f_a$ to $f_c$, the lower and upper audio frequency cutoffs as defined in Section 2.

$$N = \int_{f_a}^{f_c} S_e(f) \, df = \int_{f_a}^{f_1-} S_X(f) \, df + \int_{f_1+}^{f_c} S_X(f) \, df. \quad (3-8)$$

The signal-to-noise ratio $(S/N)$ is defined as

$$(S/N) = \frac{X^2}{N}. \quad (3-9)$$

The rms maximum input voltage $X_m$ is defined as the rms value of that sinusoidal input signal for which the maximum slope is equal to the maximum slope which can be reproduced by the system. The maximum slope which can be reproduced by the system is the step size.
Δ multiplied by the sampling frequency \( f_s \). Thus,

$$\max \left[ \frac{\Delta}{\Delta t} (\sqrt{2} X_m \sin(2 f_1 t)) \right] = \Delta f_s,$$  \hspace{1cm} (3-10)

and

$$X_m = \frac{\sqrt{2} \Delta f_s}{4\pi f_1}.$$  \hspace{1cm} (3-11)

The system input loading \( L \) is defined as the ratio of the rms input voltage to \( X_m \)

$$L = \frac{X}{X_m}.$$  \hspace{1cm} (3-12)

The maximum value of the signal-to-noise ratio is defined as \( (S/N)_{\text{max}} \). That value of \( L \) for which \( (S/N)_{\text{max}} \) occurs is defined as the optimum loading \( L_0 \).

Frequently it is more convenient to plot the values of loading and signal-to-noise ratio on a logarithmic scale. When this is done the decibel is used and the notation \( \text{(dB)} \) is added to the variable.

That is,

$$L(\text{dB}) = 20 \log L,$$  \hspace{1cm} (3-13)

and

$$(S/N) (\text{dB}) = 10 \log (S/N).$$  \hspace{1cm} (3-14)

The general behavior of the LDM1 system for a sinusoidal input has been measured by deJager and others (see References 11, 36, 37,
This is summarized as follows. (Refer to Figure 3-8).

1. The value of $L_0$ is approximately -3 dB, and is independent of input frequency and sampling frequency.

2. For values of $L$ less than $L_0$, the audio noise power $N$ is approximately constant.

3. For values of $L$ in the range of $L_0$ to 0 dB, the audio noise power is a monotonically increasing function of $L_0$.

4. The behavior of the rms error $\overline{e^2}$ in relation to the loading is very similar to that of the audio noise power.

5. The signal-to-noise ratio for values of $L$ less than $L_0$ is an increasing function of $L$.

6. For values of $L$ greater than $L_0$, the $(S/N)$ is a decreasing function of $L_0$.

7. The audio noise power decreases approximately as the third power of the sampling frequency.

Tomozawa and Kaneko have developed equations for the audio noise power and the signal-to-noise ratio for both the LDM$_1$ and LDM$_2$ systems using the following assumptions.

1. The rms error voltage is approximately constant with respect to the loading $L$.

2. The rms error voltage is related to the step size by a constant $\alpha$. 
FIGURE 3-8
KNOWN BEHAVIOR OF LDM₁ SYSTEM FOR AN 800 Hz TONE INPUT.
fs = 40 kHz.
(3) The error spectral density is flat in the frequency range 0 to \( f_s/2 \), thus that portion of the error spectral density which lies in the audio band of width \( f_d \) is related to the mean squared error by the ratio \( 2f_d/f_s \).

The resulting audio noise power is thus

\[ N = \alpha^2 \Delta^2 \frac{2f_d}{f_s}. \]  

(3-15)

The constant \( \alpha \) was evaluated by measurement in reference 36. The agreement of equation (3-15) with measurements for the LDM₁ system was reported as good over a wide range of input frequencies, sampling frequencies, and loading. Comparison with measurements for the LDM₂ system was reported as good for small values of \( L \). However, for values of \( L \) near the optimum loading point the audio noise power was greater than predicted by equation (3-15) and the optimum loading point was found to be at a smaller value of \( L \) than in the LDM₁ system.

In Chapter IV a detailed examination is made of the error spectral density and the mms error. From this investigation it is found that the constant \( \alpha \) can be evaluated from the ensemble average of \( e^2 \). Further, it is found that the error spectral density is not flat, but contains a large peak with the location of the peak varying as the loading or input frequency is changed. However, it is found that the level of the error spectral density in the audio range is relatively constant, independent of input frequency and loading.
provided the system is operated well below the slope-overload point. An equation which is in general agreement with equation (3-15) is then developed for the system for loading well below the overload point. In addition, measurements are made on hardware models and the results of these measurements are used to characterize the system near the overload point.

3.5 Idle Circuit Noise

Another important parameter in DM systems is the idle circuit or zero input noise. Ideally, with no input, the output of the sample hold circuit is an alternating series of plus and minus pulses resulting in the idle circuit error spectrum having a series of harmonics beginning at one-half the sampling frequency, which is well above the audio band. Therefore, the idle circuit noise in the audio band is zero. In practical systems, however, this is not the case.

First, if the plus and minus step sizes are not identical, a low frequency occurrence of adjacent errors of the same sign results as is illustrated in Figure 3-9. A second problem results from the offset voltage of the quantizer. The offset voltage requires that \( \hat{X}(t) \) maintain a bias level which causes periodic adjacent errors of the same sign to replace the capacitor charge that is lost through the resistor in the nonideal integrator. This phenomenon is illustrated in Figure 3-10. The net effect of offset voltage and unequal step sizes is the possibility of a very distracting idle circuit tone on the system. Wang\(^{(39)}\) has analyzed this noise in detail. For the purpose of this study, the signal levels and step size are maintained sufficiently large to make the offset voltage problem negligible, and
FIGURE 3-9:
LOW FREQUENCY NOISE CAUSED BY UNEQUAL STEP SIZES

FIGURE 3-10:
LOW FREQUENCY NOISE CAUSED BY COMPARATOR OFFSET VOLTAGE
the plus and minus step sizes are adjusted to be identical during measurements.

3.6 Comparison LDM-PCM

Briefly, comparing delta modulation to PCM, the major differences are: (1) PCM overloads on amplitude whereas DM (and all DPCM systems) overloads on slope; and (2) the quantizer has only two levels in DM whereas in PCM typically 128 levels are used. For a slope limited input signal the PCM system contains a great deal of redundancy. That is, the amplitude range is determined by the peak input signal which, for speech, occurs at the lower frequencies whereas the sampling rate is determined by the maximum input frequency. Therefore, at the lower frequencies and greater power portion of the spectrum several times the number of necessary samples are taken and at the higher frequencies an amplitude range much greater than is necessary is provided. McDonald(29) has shown that for speech, the redundancy reduction achieved using DPCM results in sufficient performance improvement so that the 6-bit DPCM system is superior to the 7-bit PCM system. The improvement which results from matching the DPCM system to the input signal carries over to DM. This improved performance together with the simplicity of encoder-decoder design (2 level vs 128 level quantization) has lead to the current interest in DM as a viable alternative to PCM for speech transmission.
CHAPTER IV

CHARACTERIZATION OF LDM WITH SINUSOIDAL INPUT SIGNALS

The further characterization of the LDM systems in this chapter is necessary as background for the study of companded DM systems which is the major objective of this paper. An expression for the granular quantizing noise is developed. This expression is in agreement with those developed by others using different techniques. Curves of the error spectral density obtained by simulation are presented. These curves enable one to achieve an understanding which has been lacking in previous studies. Measured data from working systems are presented which show good agreement with the simulations over the granular noise region. Parameters derived from these measurements are used to characterize the system performance in the overload region. Finally, the system performance for sinusoidal input signals is compared with the known system performance for exponentially distributed random input signals and with recently published perceptual performance data for speech signals. This comparison strongly indicates that the system performance for an 800 Hz tone signal is a good measure of the perceptual performance of the system for speech signals.
4.1 LDM<sub>1</sub> Encoder-Decoder Circuits

The LDM<sub>1</sub> circuits used in the experimental study are illustrated in Figure 4-1. The decoder is identical to the feedback loop of the encoder and both systems produce the signal estimate \( \hat{x}(t) \). Therefore, if it is assumed that no errors occur in the digital channel, it is sufficient to study the encoder only. With reference to Figure 4-1, the operation of the system is as follows. If the error \( e(t) \) at sample \( i \) is positive, the output of the comparator is positive and the flip-flop is set to the "one" state. If the error is negative, the output of the comparator is negative and the flip-flop is set to the "zero" state. When the flip-flop is in the "one" state, transistor \( Q_1 \) is saturated and \( Q_0 \) is cut off. A constant current source provides a charging current \( I \) to capacitor \( C_{11} \). Conversely, if the flip-flop is in the "zero" state, capacitor \( C_{10} \) is charged with the constant current \( I \). The difference in voltage across the two capacitors is amplified by an amplifier with gain \( k \) to produce the signal estimate \( \hat{x}(t) \). Resistors \( R_{10} \) and \( R_{11} \) are provided to keep the capacitor dc voltages small. The values of \( R_{11} \) and \( R_{10} \) are equal and defined as \( R_1 \). The values of the capacitors are also equal and defined as \( C_1 \).

The sampling period \( \tau \) is defined in relation to the sampling frequency \( f_s \) as

\[
\tau = \frac{1}{f_s} \quad (4-1)
\]

The corner frequency \( f_1 \) of the integration network \( R_1 C_1 \) is

\[
f_1 = \frac{1}{2\pi R_1 C_1} \quad (4-2)
\]
ENCODER

CLOCK, FREQUENCY \( f_s \)

INPUT \( x(t) \) COMPARATOR

CHANNEL

CONSTANT CURRENT GENERATOR

DECODER

OUTPUT \( y(t) \) FILTER BANDPASS \( f_a \) TO \( f_c \)

STEP SIZE \( \Delta_i \)

VOLTAGE

- \( \Delta_i \)

+ \( \Delta_i \)

SAMPLING PERIOD \( t = 1/f_s \)

WAVEFORM AT \( \hat{x}_i(t) \) IN BOTH ENCODER AND DECODER WHEN \( x(t) = 0 \)

FIGURE 4-1

LDM \(_1\) ENCODER - DECODER
The step size $\Delta_1$ (subscript 1 for LDM$_1$ system) is defined as the peak to peak amplitude of the triangular wave seen at $\hat{X}(t)$ when the input is zero. (Refer to Figure 4-1.) In terms of the circuit parameters, $\Delta_1$ is

$$\Delta_1 = kR_1I \left(1 - \exp(-\tau/R_1C_1)\right). \tag{4-3}$$

For the LDM$_1$ system, $I$ is constant. For the companded systems discussed later, the value of $\Delta_1$ is made a variable based upon the average signal level. This is achieved by varying $I$.

An equivalent voltage source $E$ is defined as

$$E = kR_1I. \tag{4-4}$$

The time constant $R_1C_1$ is much greater than $\tau$. Thus, (4-3) can be simplified to

$$\Delta_1 \approx 2\pi E f_1/f_s. \tag{4-5}$$

With reference to the definition of the maximum input $X_m$, (equation (3-9)), $X_{m1}$ is defined in terms of the circuit parameters as

$$X_{m1} = \frac{\sqrt{2}}{2} E f_1/f_1. \tag{4-6}$$

In the following development the sampling frequency $f_s$ and the input frequency $f_1$ are treated as dependent variables. The input voltage is treated as the independent variable. Where convenient, the rms
value of the input voltage is expressed in terms of the loading factor \( L \). From equations (3-12) and (4-6),

\[
L = X \frac{\sqrt{2}}{E f_1}. \quad (4-7)
\]

For the simulations and measurements the sampling frequency is restricted to values of 28 KHz, 40 KHz, and 56 KHz with 40 KHz considered as a typical operational system sampling frequency.

4.2 Calculation of Audio Noise Power

Analysis of quantizing systems such as LDM and PCM is difficult. Bennett\(^{(4)}\) has shown that when these systems are exited with a sinusoidal signal the error spectral density in the audio band is very irregular, contains many impulses, and varies radically as the input voltage is altered. To overcome this problem, input signals which are random and similar to resistor noise in nature are often used. With this type of input, calculated error spectral densities are usually found to be smooth. However, when the error spectral density of an experimental system with a sinusoidal input is measured, it is found to be smooth and comparable to that calculated for a random input. This apparent contradiction is explained by noting that:

1) Measurements of spectra involve long time averages and cover finite widths of the spectrum. As a result, some irregularities are smoothed.

2) In a working system, the sampling frequency and the input frequency are not locked in phase. As a result, the measured spectra reflect some smoothing due to phase variations.
3) The decision level of the quantizer is corrupted by noise and thus, is not precise. This also tends to smooth the spectra.

4) The plus and minus step sizes are not precisely the same which also has a smoothing effect on the spectra.

The problem of determining the audio noise power for a sinusoidal input signal is solved in the following manner:

1) The system is assumed to be ergotic. An ensemble of records is taken by repeatedly turning the system on, allowing it to reach a steady state, then taking a record. Averages computed on this ensemble are assumed to be identical to long time averages taken on a single waveform. Variables which are computed from ensemble averages are denoted by the symbol $\langle \rangle$.

2) The rms error $\langle e^2 \rangle$ is determined by taking an ensemble average. This is found to be essentially a function of only the step size.

3) The error autocorrelation function at the sampling points is determined by taking the ensemble average of a large number of waveforms which are generated by computer simulation.

4) The error autocorrelation samples are Fourier transformed to obtain the error spectral density $\langle S_e(f) \rangle$. 
5) The normalized error spectral density is defined as

\[ \langle S_n(f_n) \rangle = \frac{\langle S_e(f_n) \rangle}{\langle e^2 \rangle}, \]  

(4-8)

where

\[ f_n = f/f_s. \]  

(4-9)

6) The amplitude of \( \langle S_n(f_n) \rangle \) in the audio band \( (f_a/f_s \leq f_n \leq f_c/f_s) \) is determined from the simulation to be constant independent of \( L, f_n, f_i, \) and \( f_s \) for values of \( L \) well below the slope overload region. This constant amplitude of \( \langle S_n(f_n) \rangle \) is defined as \( S_{n_0} \). A typical spectral density is illustrated in Figure 4-2.

7) The constant \( \beta \) is defined as that normalized frequency for which a flat spectrum of amplitude \( S_{n_0} \) and bandwidth \( \beta \) encloses unit area. Thus,

\[ \int_{0}^{\beta} S_{n_0} \, df_n = \beta \, S_{n_0} = 1. \]  

(4-10)

8) The audio noise power \( \langle N \rangle \) can now be computed using equations (3-8), (4-8), and (4-10).

\[ \langle N \rangle = \int_{f_a}^{f_c} \langle S_e(f) \rangle \, df = \langle e^2 \rangle \int_{f_a/f_s}^{f_c/f_s} \langle S_n(f_n) \rangle \, df_n = \]  

(cont.)


\[
\frac{\langle N \rangle}{\langle e^2 \rangle} = \frac{fd}{fs} S_{n_0}
\]

**FIGURE 4-2**

ILLUSTRATION OF \( S_{n_0} \) AND \( \beta \).
where \( f_d \) is the audio bandwidth \( f_c - f_a \) as defined in equation (2-1). When \( \beta \) is computed from the simulation results, the subscripts 1 and 2 are added to differentiate between the LDM\(_1\) and LDM\(_2\) systems.

### 4.3 Approximate Expression for \( \langle e^2 \rangle \)

The rms error \( \langle e^2 \rangle \) has been calculated by taking an ensemble average in Appendix C and found to be

\[
\langle e^2 \rangle = \frac{A^2}{3} [1 + \frac{3}{A^2} R_e(1)],
\]

(4-12)

where \( R_e(k) \) is the autocorrelation function of the error at the sample points.

\[
R_e(k) = \langle e(n), e(n+k) \rangle
\]

(4-13)

First, equation (4-12) is checked for zero input. With no input, the signal estimate \( \hat{X}(t) \) and the error \( e(t) \) are identical, a triangular wave with peak values of \( \pm \frac{A}{2} \), thus \( R_e(1) = -\frac{A^2}{4} \) and the value of \( \langle e^2 \rangle \) is

\[
\langle e^2 \rangle \bigg|_{L=0} = \frac{A^2}{12}.
\]

(4-14)
This agrees with the time average of the squared error $\bar{e}^2$ which is obtained by integrating $e(t)^2|_{L=0}$ over multiples of the time period $\tau$.

Next, consider what happens to $Re(1)$ as the loading is increased from zero. For small values of $L$, the step size $\Delta$ is large in relation to changes in the signal over one sampling period. The error tends to change sign every sample and $Re(1)$ is negative. Conversely, for heavy loading the changes of the input signal between the sampling time approaches the value of $\Delta$, and the error tends to be of the same sign for several sampling periods (for instance during the steep portion of the sinewave input). Thus, the value of $Re(1)$ tends to be positive. At some value of $L$, the term $\frac{3}{\Delta^2} Re(1)$ is zero.

$Re(1)$ was computed in the simulations described in Appendix A. The results of these simulations for the LDM$_1$ system are indicated in Table 4-1. First, considering runs 1 through 4, it is seen that the value of $\frac{3}{\Delta^1}$ $Re(1)$ varies from -.39 for a loading of -18 dB to a value of +.297 for a loading of -3 dB. The expression $(1 + \frac{3}{\Delta^2} Re(1))$ changes over the range of 0.61 to 1.30, a 3.3 dB change. The behavior shown on the other runs is similar. The value of $\Delta^2/3$ is approximately 20 dB below $X_{m1}$ which leads to the conclusion that the value of $\langle e^2 \rangle$ is relatively constant over the loading range of -18 dB to -3 dB. Thus, approximating the value of $3/\Delta^2 Re(1)$ as zero is reasonable, and $\langle e^2 \rangle$ is

$$\langle e^2 \rangle \approx \frac{\Delta^2}{3}. \quad (4-15)$$
TABLE 4-1

VALUES OF \( \frac{3}{\Delta_1^2} \text{Re}(l) \), \( 1 + \frac{3}{\Delta_1^2} \text{Re}(l) \)

COMPUTED BY SIMULATION

<table>
<thead>
<tr>
<th>RUN</th>
<th>( L )</th>
<th>( f_0 )</th>
<th>( f_1 )</th>
<th>( \frac{3}{\Delta_1^2} \text{Re}(l) )</th>
<th>( 1 + \frac{3}{\Delta_1^2} \text{Re}(l) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-18 dB</td>
<td>40 KHz</td>
<td>800 Hz</td>
<td>-0.392</td>
<td>0.61</td>
</tr>
<tr>
<td>2</td>
<td>-12 dB</td>
<td>40 KHz</td>
<td>800 Hz</td>
<td>-0.377</td>
<td>0.62</td>
</tr>
<tr>
<td>3</td>
<td>-6 dB</td>
<td>40 KHz</td>
<td>800 Hz</td>
<td>-0.093</td>
<td>0.91</td>
</tr>
<tr>
<td>4</td>
<td>-3 dB</td>
<td>40 KHz</td>
<td>800 Hz</td>
<td>+0.297</td>
<td>1.30</td>
</tr>
<tr>
<td>5</td>
<td>-6 dB</td>
<td>40 KHz</td>
<td>1600 Hz</td>
<td>-0.278</td>
<td>0.72</td>
</tr>
<tr>
<td>6</td>
<td>-18 dB</td>
<td>40 KHz</td>
<td>400 Hz</td>
<td>-0.516</td>
<td>0.48</td>
</tr>
<tr>
<td>7</td>
<td>-6 dB</td>
<td>28 KHz</td>
<td>800 Hz</td>
<td>-0.230</td>
<td>0.77</td>
</tr>
<tr>
<td>8</td>
<td>-6 dB</td>
<td>56 KHz</td>
<td>800 Hz</td>
<td>-0.241</td>
<td>0.76</td>
</tr>
</tbody>
</table>
4.4 Error Spectral Density - Simulation

The error spectral density has been determined using the simulation described in Appendix A. As mentioned earlier, this is an ensemble average, and the variables are indicated by using the symbol $\langle \rangle$ to differentiate from the measured values which are discussed later. The results of the simulation are shown in Figures 4-3, 4-4 and 4-5. Each curve has been normalized by dividing the amplitude of the spectral density by $\langle e^2 \rangle$ as given by equation (4-15) and by plotting on a frequency scale that is normalized with respect to the sampling frequency. This normalization was defined in equations (4-8) and (4-9).

The error spectral density for various input loadings at a sampling frequency of 40 KHz and an input frequency of 800 Hz are shown in Figure 4-3. Curves for loading of -12 and -18 dB are very similar and have peaks near one half the sampling frequency as is expected when the system is lightly loaded. The curve for a loading of -6 dB is somewhat flatter with the peak much closer to the audio band. For all three curves (-18, -12, and -6 dB loading) the density in the audio band is approximately the same. This indicates that in the loading range of -6 dB to -18 dB the audio noise power is constant. The mms error (area under each curve) is also approximately constant. The curve for -3 dB loading has a peak which is near the audio band, an increased mms error, and an increased density in the audio band. This loading corresponds to a point where the slope-overload noise is appreciable and thus causes an increase in the audio noise power.
FIGURE 4-3
ERROR SPECTRAL DENSITY FOR VARIOUS INPUT LOADING. (BY SIMULATION)

\( f_i = 800\text{Hz} \), \( f_s = 40\text{kHz} \)
Figure 4-4 illustrates the change in the error spectral density as the input frequency is changed. The input power is the same for each curve and corresponds to -12 dB loading at 800 Hz (i.e., -18 dB loading at 400 Hz and -6 dB at loading 1600 Hz). Although it is not obvious from the curves, the area under each curve was computed in the simulation and is essentially the same for each input frequency, leading to the conclusion that $\langle e^2 \rangle$ is independent of the input frequency provided that the input power at each frequency is well below the slope-overload point. The level of $\langle S_n(f_n) \rangle$ in the audio band is also the same for each curve leading to the conclusion that $\langle N \rangle$ is also independent of the input frequency. From Figures 4-3 and 4-4 it is seen that changing the input power at a fixed input frequency and changing the input frequency at a fixed input power results in similar changes to the error spectral density. That is, the system performance is based on the rms slope (derivative) of the input signal.

Figure 4-5 illustrates the variation in the error spectral density for sampling frequencies of 28, 40, and 56 KHz. The loading is fixed at -6 dB for each curve, well below the point where the overload noise makes an appreciable contribution to the quantizing noise. The curves for sampling frequencies of 40 KHz and 56 KHz are nearly identical whereas the curve for 28 KHz shows a substantial increase in the rms error; however, it closely follows the general shape of the other curves. Also, the level of the spectral density in the audio range is nearly equal for the three curves.
ERROR SPECTRAL DENSITY FOR VARIOUS INPUT FREQUENCIES. (BY SIMULATION)

\( f_s = 40 \text{kHz} \). INPUT CONSTANT
Figure 4-5

Error spectral density for various sampling frequencies. (By simulation)

\( f_i = 800 \text{Hz} \). \( L = -6 \text{dB} \).
From these curves, with the normalization that is used, two points are clear. First, for values of loading of -6 dB or less, the mms error \( <e^2> \) is independent of \( f_1, f_s \) (for \( f_s > 40 \text{ KHz} \)) and \( L \).

Second, for values of loading of -6 dB or less, the level of the error spectral density in the audio band is independent of \( f_n, f_1, f_s, \) and \( L \).

It was these two facts that were used to develop the expression for \( <N> \) in equation (4-11). The constant \( \beta_1 \) (subscript 1 for the LDM1 system) in equation (4-11) is now determined. Rearrangement of equation (4-11) gives

\[
\beta_1 = \frac{f_d \Delta_1^2}{3 f_s \int_{f_n}^{f_s} <s_e(f)> df}.
\]

The value of \( \beta_1 \) was computed using this equation and the simulation results for values of \( f_s = 40 \text{ KHz}, f_1 = 800 \text{ Hz}, \) and \( L = -6 \text{ dB} \) and was determined to be \( \beta_1 = 2.04 \). This value of \( \beta_1 \) is used throughout the remainder of this paper.

The mms error \( <e^2> \) (equation (4-15)) in terms of the maximum input \( X_{m1} \) (equation (3-11)) is

\[
<e^2> = \frac{2X_{m1}^2 (2\pi f_1)^2}{3 f_s^2}.
\]

Similarly, the audio noise power \( <N> \) (equation (4-11)) can be written as

\[
<N> = \frac{2X_{m1}^2 (2\pi f_1)^2 f_d}{3 \beta_1 f_s^2}.
\]
And finally, from equations (3-9) and (4-18), the signal-to-noise ratio is

\[ \langle S/N \rangle = \frac{3\beta_1 f_s^3 L^2}{2(2\pi f_1)^2 f_d} \]  

(4-19)

All the constants in (4-19) can be combined into a single constant \( \beta_1 = 2.04 \), and

\[ \langle S/N \rangle = 0.078 \frac{f_s^3}{f_1^2 f_d} \]  

(4-20)

Equation (4-20) is written in dB as

\[ \langle S/N \rangle (\text{dB}) = \langle S/N \rangle_0 (\text{dB}) + L(\text{dB}), \]  

(4-21)

where

\[ \langle S/N \rangle_0 = 0.078 \frac{f_s^3}{f_1^2 f_d} \]  

(4-22)

These equations are restricted to values of loading less than 

-6 dB (the granular noise region), and can be summarized as follows:

1) The rms error varies inversely as the square of the sampling frequency and is independent of loading.

2) The audio noise power varies as the inverse of the cube of the sampling frequency and is independent of loading.

3) The signal-to-noise ratio varies as the cube of the sampling frequency and as the square of the loading.
4.5 Measured Results, Comparison With Simulation

The experimental circuit used is identical to the circuit shown in Figure 4-1. The error \( e(t) \) is determined by passing the input \( X(t) \) and the signal estimate \( \hat{X}(t) \) through a differential amplifier, (equation (3-3)). A phase and gain adjustment is made on \( X(t) \) before it is applied to the differential amplifier so that average power at the input frequency \( f_1 \) is canceled. Thus, the measured mms error \( \bar{e}^2 \) is given by equation (3-6). Measurements of \( \bar{e}^2 \) were made using a Siemens 3D344 level meter. The audio noise power \( N \) was measured using a Western Electric 3A noise measuring set.\(^7\) A 3 KHz flat filter was used in the 3A set. This filter has very sharp cut-offs with 3 dB points at approximately 100 Hz and 3.0 KHz which closely matches the band \( f_c \) to \( f_a \) of 250 Hz to 3.1 KHz.

4.5.1 mms Error

The measured, calculated and simulated mms error as a function of input loading are shown in Figure 4-6. The input frequency is 800 Hz and the sampling frequency is 40 KHz. The simulation points are slightly greater than the measured points. The computed values of \( \langle e^2 \rangle \) using equation (4-12) and the simulation results for \( R_e(1) \) (Table 4-1) show good agreement with the measured values of \( \bar{e}^2 \). The value of \( \langle e^2 \rangle \) computed using equation (4-15) is higher than that measured over a wide range of loading. This is due to approximating the term \( 3/\Delta t^2 \) \( R_e(1) \) as zero rather than a negative value to reflect the mms error for a lighter loading. In general, it can be concluded that the simulation and measured data show good agreement.
A SIMULATION POINTS
• COMPUTED POINTS USING EQUATION (4-12)

\[ \frac{\langle \theta^2 \rangle}{Xm_1^2} \]

\[ \frac{X^2}{Xm_1^2} \]

\[ \frac{\langle \theta^2 \rangle}{Xm_1^2} \text{ COMPUTED EQUATION (4-15)} \]

\[ \frac{\langle \theta^2 \rangle}{Xm_1^2} \text{ MEASURED} \]

\[ \frac{\langle \theta^2 \rangle}{Xm_1^2}, \frac{X^2}{Xm_1^2} \]

FIGURE 4-6
MMS ERROR AS A FUNCTION OF LOADING;
COMPUTED, MEASURED AND SIMULATED.

\[ f_i = 800\text{Hz} \quad f_s = 40\text{kHz} \]
4.5.2 Error Spectral Density

The error spectral density was measured for a sampling frequency of 40 KHz and an input frequency of 800 Hz at a loading of -12 dB. The measurements were made by using the 50 Hz bandwidth filter supplied with the Siemens 3D3414 meter. The measurements were smoothed using equation (A-12). This measurement is compared with results obtained from the simulation in Figure 4-7. The agreement between the two is very good which supports the accuracy of the simulation.

4.5.3 Signal-To-Noise Ratio

A comparison of the measured signal-to-noise ratio and that computed (equation (4-21)) together with the simulation results in shown in Figure 4-8. The agreement between the simulation and the measured data is good which again supports the simulation. Equation (4-21) shows good agreement over the loading range of -3 dB to -18 dB. The maximum point on the curve occurs at the optimum loading point $L_0$. From the measured data $L_0$ occurs at -3.5 dB. For $L$ greater than $L_0$ the $S/N$ characteristic is approximated by the straight line on Figure 4-8. This corresponds to a 2 dB reduction in $(S/N)$ for each dB increase in loading. Using this approximation for $L > L_0$, and equation (4-21) for $L < L_0$, one can write the $S/N$ as the two line segments

\[
(S/N)(dB) = \begin{cases} 
(S/N)_0 (dB) + L(dB), & L < -3.5 \text{ dB}, \\
(S/N)_0 (dB) - 10.5 \text{ dB} - 2 L(dB), & -3.5 \text{ dB} < L < 0 \text{ dB}, 
\end{cases}
\]  

where $(S/N)_0$ is the value of equation (4-21) for $L = 0$. 

FIGURE 4-7
COMPARISON, SIMULATION AND MEASURED
ERROR SPECTRAL DENSITY $f_i = 800\text{Hz}$ $f_s = 40\text{kHz}$.
$L = -12\text{dB}$
FIGURE 4-8
COMPARISON, MEASURED AND SIMULATION SIGNAL-TO-NOISE RATIO.-LDM

\[ f_i = 800\text{HZ} \quad f_s = 40\text{KHz} \]
\[(S/N)_o \ (dB) = 10 \log \left(0.78 \frac{f_s^3}{f_1^2 f_d}\right). \] \hspace{1cm} (4-24)

The maximum signal-to-noise ratio \((S/N)_{\text{max}}\) occurs at \(L_o\). From equations \((4-23)\) and \((4-24)\),

\[ (S/N)_{\text{max}} \ (dB) = 10 \log \left(0.35 \frac{f_s^3}{f_1^2 f_d}\right). \] \hspace{1cm} (4-25)

Figures 4-9 and 4-10 show measured values of \((S/N)\) as a function of loading for variations in input frequency (Figure 4-9) and sampling frequency (Figure 4-10) along with the computed values using equation (4-23). The agreement is good over a wide range of loading, and it is concluded that equation (4-23) is a good approximation of the signal-to-noise ratio.

### 4.6 LDM2 Circuit

The double integration LDM encoder circuit is illustrated in Figure 4-11. It is identical to the LDM1 encoder shown in Figure 4-1 with the exception that a second integration consisting of \(R_2, R_3, \) and \(C_2\) has been added. This second integrator is also added to the decoder, and as in the LDM1 system, only the encoder circuit is considered. The transfer characteristic of the combined integration networks is illustrated in Figure 4-12. With reference to Figure 4-11, the pole of the second integrator is at frequency \(f_2\), which in relation to the circuit elements is

\[ f_2 = \frac{1}{2\pi(R_2+R_3)C_2}. \] \hspace{1cm} (4-26)
FIGURE 4-9
MEASURED (S/N) AS A FUNCTION OF LOADING FOR VARIOUS INPUT FREQUENCIES -LDM \( f_s = 40\text{kHz} \)
FIGURE 4-10
MEASURED (S/N) FOR VARIOUS SAMPLING FREQUENCIES - LDM₁

\[ f_s = 56 \text{kHz} \]
\[ f_s = 40 \text{kHz} \]
\[ f_s = 28 \text{kHz} \]

SIGNAL-TO-NOISE RATIO (S/N) (dB)

CALCULATED (EQN 4-23)
MEASURED

INPUT LOADING L(dB)
**Figure 4-11**
LDM2 Encoder

**Figure 4-12**
Frequency characteristic of network $H(f)$
The zero of the second integrator occurs at frequency \( f_3 \), where
\[
f_3 = \frac{1}{2\pi R_3 C_2}.
\] (4-27)

The transfer function of the second integrator is
\[
H_2(f) = \frac{1 + j \frac{f}{f_3}}{1 + j \frac{f}{f_2}}.
\] (4-28)

The step size \( \Delta_2 \) of the LDM\(_2\) system is defined in the same manner as \( \Delta_1 \), the step size for the LDM\(_1\) system; the peak to peak value of \( \hat{x}(t) \) when the input is zero (refer to Figure 4-1). The triangular waveform has components at frequencies \( f_s/2, 3f_s/2, 5f_s/2 \) \ldots etc. The components at frequency \( 3f_s/2 \) and above are small and are attenuated in the encoder (and decoder) amplifier. Using only the fundamental frequency of the triangular wave, one can express \( \Delta_2 \) in terms of \( \Delta_1 \) as
\[
\Delta_2 = \Delta_1 |H_2(f_s/2)|.
\] (4-29)

The value of \( \Delta_1 \) is given in equation (4-5), and \( \Delta_2 \) can be expressed as
\[
\Delta_2 = 2\pi E \frac{f_1}{f_s} \left( \frac{1 + \frac{f_s^2}{4f_3^2}}{1 + \frac{f_s^2}{4f_2^2}} \right)^{1/2}.
\] (4-30)
The maximum input $x_{m_2}$ is defined as that input which causes the input to the second integrator to be the maximum input for the LDM$_1$ system. Thus,

$$x_{m_1} = \frac{x_{m_2}}{|H_2(f_1)|} \quad \text{(4-31)}$$

The maximum input in terms of the circuit parameters is determined by substituting for $x_{m_1}$ (equation (4-6)) and $H_2(f_1)$ (equation (4-28)).

$$x_{m_2} = \frac{\sqrt{2} E f_1}{2 f_1} \left( \frac{1 + \left(\frac{f_1}{f_2}\right)^2}{1 + \left(\frac{f_1}{f_2}\right)^2} \right)^{1/2} \quad \text{(4-32)}$$

The loading for the LDM$_2$ system is defined in the same manner as the LDM$_1$ system (equation 4-7).

$$L = \frac{x}{x_{m_2}} = x \left( \frac{\sqrt{2} f_1}{E f_1} \right) \left( \frac{1 + \left(\frac{f_1}{f_2}\right)^2}{1 + \left(\frac{f_1}{f_3}\right)^2} \right)^{1/2} \quad \text{(4-33)}$$

4.6.1 Location of $f_2$ and $f_3$

As discussed in Chapter 3, the second integrator can be viewed as an instantaneous step size adapter. If several adjacent quantized errors are of the same sign, which indicates an overload condition, the fundamental frequency of the ramp signal approximation is lower than $f_3$ causing the ramps to be attenuated less by $H_2(f)$ and, in effect, causing
the step size to increase. The maximum amount of increase is essentially determined by the ratio $f_3/f_2$. Therefore, to achieve the maximum improvement, it is desirable to separate $f_2$ and $f_3$ as far as possible.

The lower limit on $f_2$ is dictated by overload considerations. The term $(f_1/f_1)$ in equation (4-32) gives a 6 dB/octave rolloff which matches the slope limited characteristic of the speech signal (refer to Figure 2-1). The term $(f_2/f_1)$ gives an additional 6 dB/octave rolloff at frequencies above $f_2$. From the model of the speech spectrum shown in Figure 2-2, the additional 6 dB rolloff above $f_2$ will cause overloading at the higher input frequencies unless $f_2$ is equal to or greater than $f_c$. However, in the spectrum model the power density is somewhat greater at the high frequency end of the spectrum than that noted from measurements (refer again to Figure 2-1) which permits $f_3$ to be less than $f_c$. Greefkes (19) has found that a value of $f_2 = 1800$ Hz does not cause any perceptive overloading at the higher frequencies. This value of $f_2$ will be used in the remainder of this paper.

The upper limit on the value of $f_3$ is set by the sampling frequency. Increasing $f_3$ beyond $f_s/2$ does not reduce the zero input step size as is indicated by equation (4-32) and is of no advantage. For a change in error sign every sample, the major frequency component of $\hat{X}(t)$ is at $f_s/2$. For a change every other sample, it is at $f_s/4$, etc. The value of $\Delta_2$ should start increasing only after several errors of the same sign occur. Therefore, $f_3$ should be less than $f_s/2$, but not less than $f_s/6$ or $f_s/8$ (corresponding to 3 or 4 adjacent errors of the same
Figure 4-13 shows the measured improvement in the maximum value of the signal-to-noise ratio as a function of $f_3$. The maximum improvement is $3.5$ dB obtained for $f_3 = f_b/6$.

For $f_2$ on the order of $1800$ Hz and $f_3$ equal to $f_b/6$, $f_1$ is much less than $f_3$, $f_2$ is much less than $f_b$, and equations (4-30), (4-32), and (4-33) can be simplified to:

$$
\Delta_2 \approx 4\sqrt{10} \pi E \frac{f_1 f_2}{f_b^2},
$$

(4-34)

$$
X_m_2 \approx \frac{\sqrt{2} E f_1 f_2}{2 f_1 \sqrt{f_1^2 + f_2^2}},
$$

(4-35)

$$
L \approx X \left( \frac{\sqrt{2} f_1 \sqrt{f_1^2 + f_2^2}}{E f_1 f_2} \right).
$$

(4-36)

4.6.2 Error and Noise Spectral Density - Simulation

The error spectral densities for a sampling frequency of $40$ KHz, input frequency of $800$ Hz and loads ranging from $-18$ dB to $-3$ dB are shown in Figure 4-14. These curves are normalized using equation (4-8) and were determined using the simulation described in Appendix A. A comparison of Figure 4-14 with the results for the LDM$_1$ system (Figure 4-3) reveals the following:

1) The error spectral density for the LDM$_2$ system is much flatter with the high frequency peaks noted in the LDM$_1$ system eliminated.
FIGURE 4-13

IMPROVEMENT IN MAXIMUM SIGNAL-TO-NOISE RATIO AS A FUNCTION OF THE RATIO OF SAMPLING FREQUENCY TO SECOND INTEGRATOR ZERO FREQUENCY. \((f_s/f_3)\).

\[ f_i = 800\text{Hz}, \ f_s = 40\text{kHz} \]
NORMALIZED SPECTRAL DENSITY \( S_n(f_n) \)

\[ L = -3dB \]
\[ L = -6dB \]
\[ L = -12dB \]
\[ L = -18dB \]

NORMALIZED FREQUENCY \( f_n = f/f_s \)

FIGURE 4-14
ERROR SPECTRAL DENSITY \(-LDM_2\)
(By Simulation) \( f_i = 800Hz \), \( f_s = 40kHz \)
2) The area under the \(-18\, \text{dB}, -12\, \text{dB}, \text{and } -6\, \text{dB} \) curves in both systems is essentially constant. Thus, \( \langle e^2 \rangle \) is approximately independent of input for \(-18\, \text{dB} < L < 6\, \text{dB} \).

3) For both systems, the audio noise power (that portion of the error spectral density in the audio band) is essentially flat and constant for loading in the range \(-18\, \text{dB} < L < -6\, \text{dB} \).

4) For both systems, the \( L = -3\, \text{dB} \) curve shows a marked increase in the audio noise power over that for smaller inputs; this is the effect of slope-overload noise making an appreciable contribution to quantizing noise.

For input loading of \(-6\, \text{dB} \) or less, the behavior of the audio noise power for the two systems is very similar. This is the region where granular noise predominates. In this region, the effect of the second integrator is minimal. Following the technique used for the LDM1 system, a constant \( \beta_2 \) (subscript 2 for LDM2 system) is defined such that a flat spectrum of width \( \beta_2 f_S \) has the same height as the average value of the normalized error spectral density \( \langle S_n(f_n) \rangle \) in the audio region and has total area equal to one. From equation (4-11),

\[
\beta_2 = \frac{\langle e^2 \rangle}{\langle N \rangle f_S} \quad \text{fd}. \quad (4-37)
\]

The value of \( \beta_2 \) was computed using equation (4-37) and the value of \( \langle N \rangle \) obtained from the simulation for \( L = -6\, \text{dB}, f_1 = 800\, \text{Hz} \) and
\( f_s = 40 \text{ KHz} \) and determined to be 0.785. From equation (4-37), the audio noise power is

\[
\langle N \rangle = \frac{f_d}{\sigma_s^2} \langle e^2 \rangle. \tag{4-38}
\]

Finally, substituting equations (4-15) and (4-34) into (4-38),

\[
\langle N \rangle = \frac{160 \pi^2}{38^2} \frac{E^2 f_d f_1^2 f_2^2}{f_s^5}. \tag{4-39}
\]

The signal-to-noise ratio from equations (3-9) (4-36) and (4-39) is

\[
\langle S/N \rangle = \frac{38^2}{320 \pi^2} \frac{f_s^5}{f_d f_1^2(f_1^2 + f_2^2)} L^2. \tag{4-40}
\]

The constants in equation (4-40) can be combined into a single number, and

\[
\langle S/N \rangle = 7.50 \times 10^{-4} \frac{f_s^5}{f_d f_1^2(f_1^2 + f_2^2)} L^2. \tag{4-41}
\]

These equations are restricted to values of \( L \) less than \(-6 \text{ dB}\).

4.7 Measured Results, LDM2

The experimental circuit is identical to the circuit illustrated in Figure 4-11. The measurement techniques are identical to those used for the LDM₁ system (Section 4.5).

4.7.1 Error Spectral Density

Figure 4-15 shows a comparison of the measured error spectral density and the one determined by simulation. Agreement is good which again substantiates the validity of the simulation.
FIGURE 4-15
COMPARISON OF MEASURED AND SIMULATOR ERROR SPECTRAL DENSITY –LDM₂

\[ f = 800\text{Hz, } f_s = 40\text{KHz } x/x_m = -12\text{dB} \]
4.7.2 **Signal-To-Noise Ratio**

A comparison of the measured signal-to-noise ratio and that computed (equation 4-41) together with the simulation results is illustrated in Figure 4-16. Equation (4-41) shows good agreement for \( L \) equal to or less than \(-6\) dB. In the loading range \(-6\) dB to \(-4\) dB the \((S/N)\) is essentially flat. This is the region where the second integrator causes the step size \( \Delta_2 \) to increase as the loading increases. For loading greater than \(-4\) dB, the system overloads, and the \((S/N)\) is reduced approximately 2 dB for each dB increase in the loading. Using equation (4-41) in the loading range \( L < -6 \) dB, and the two line segments for \( L > -6 \) dB indicated in Figure 4-16, one can approximate the signal-to-noise ratio as

\[
(S/N)(dB) = \begin{cases} 
(S/N)_0(dB) + L(dB), & L < -6 \text{ dB}, \\
(S/N)_0(dB) - 6 \text{ dB}, & -6 \text{ dB} < L < -4 \text{ dB}, \\
(S/N)_0(dB) - 14 \text{ dB} - 2L(dB), & -4 \text{ dB} < L < 0 \text{ dB}, 
\end{cases} 
\]

where \((S/N)_0\) is given by equation (4-41) for \( L = 1 \) (i.e., 0 dB),

\[
(S/N)_0 \text{ (dB)} = 10 \log \left( \frac{7.5 \times 10^{-4} f_s^5}{f_d f_1^2(f_1^2 + f_2^2)} \right) . 
\]

The maximum point on the measured curve is at \( L_0 = -4 \) dB. The value of the \((S/N)\) at this point is (equation (4-42))

\[
(S/N)_{\text{max}} \text{ (dB)} = 10 \log \left( \frac{1.88 \times 10^{-4} f_s^5}{f_d f_1^2(f_1^2 + f_2^2)} \right) . 
\]
Figure 4-16
Comparison, Measured and Simulator Signal-to-Noise Ratio - LDM2

$S/N$, (dB) vs. Input Loading $L$, (dB)

- Measured Points
- Simulation Points

Approximation

$-6 \text{dB} < L < -4 \text{dB}$

Eqn 4-42

$L_0 = -4 \text{dB}$

$f_i = 800 \text{Hz}$ $f_s = 40 \text{KHz}$
Equation (4-42) was checked with measurements at sampling frequencies of 28 KHz and 56 KHz and shows good agreement.

4.8 Comparison LDM₁ and LDM₂

A comparison of the (S/N) for the LDM₁ and LDM₂ systems for a sampling frequency of 40 KHz and an input frequency of 800 Hz is illustrated in Figure 4-17. From this Figure it is seen that the LDM₂ system exhibits a broader peak at the optimum operating point. From equations (4-25) and (4-44) it is seen that (S/N)_{max} for the LDM₁ system varies as the third power of the sampling frequency whereas (S/N)_{max} varies as the fifth power of the sampling frequency for the LDM₂ system. These equations are plotted in Figure 4-18 where it is noted that the advantage of the LDM₂ system is marginal at a sampling frequency of 28 KHz but at sampling frequencies near or above 40 KHz the advantage is substantial.

4.9 Comparison LDM₁ System Performance for Sinewave and Simulated Speech Input Signals

When the LDM₁ system is operated below the optimum loading point, the audio noise power is constant, independent of input power or frequency (Figures 4-3, 4-4). Also, the optimum loading point is a function of the slope of the input signal, not the input power. From these two points it would be expected that the noise degradation to a broad band signal with a spectral density which is also slope limited would be similar. Abate[2] has simulated the response
FIGURE 4-17
COMPARISON, SIGNAL-TO-NOISE RATIO OF LDM\textsubscript{1} AND LDM\textsubscript{2} $f_i = 800\text{Hz}$ $f_s = 40k\text{Hz}$
Figure 4-18

Comparison, \( (S/N)_{\text{MAX}} \) of LDM\(_1\) and LDM\(_2\) as a function of sampling frequency.

\( f_i = 800\text{Hz} \)
of the LIM$_{1}$ system for both gaussian and exponentially distributed signals with spectral densities of the form

$$S(f) = \frac{1}{1 + \left(\frac{f}{800 \ \text{Hz}}\right)^2}, \quad f \leq 3500 \ \text{Hz.} \quad (4-45)$$

This signal is referred to as the simulated speech signal. Nearly identical results were reported for both distributions. Here the results of Abate's simulations are compared with the results for an 800 Hz tone. Since the system overloads on slope, it would be expected that if a sinuclodal signal has the same mean squared slope as the simulated speech signal, the noise degradation to both signals would be identical. It is desired to make the comparison based on the signal-to-noise ratio. Since the simulated speech signal has exactly 1/2 the input power of an 800 Hz tone with the same mean squared slope, 3 dB must be added to the (S/N) for the simulated speech signal to make the two comparable on a (S/N) basis. A comparison of the (S/N) for the two signals is shown in Figure 4-19. The ordinate is in units of root mean square slope normalized to the maximum slope which can be reproduced by the system, $\Delta f_{S}$. The sampling frequency for the comparison is 40 KHz. From Figure 4-19, it is noted that the signal-to-noise ratios agree for very small slope inputs. However, as the maximum point on the curve is approached, the (S/N) for an 800 Hz tone is greater, and the maximum point is for a larger value of rms input slope. The measurement of the (S/N) for the simulated speech signal was based on the mean squared error. The measurement for the 800 Hz
FIGURE 4-19
COMPARISON SIGNAL-TO-NOISE RATIO FOR SIMULATED SPEECH SIGNAL AND 800Hz TONE

\[ f_s = 40kHz \]
tone was based on a modified mean squared error for which all power at 800 Hz is canceled. Also, for the 800 Hz tone it was noted that as the input power was increased the amount of phase shift between the input and output increased. To reconcile the difference between the (S/N) for the two signals, a second measurement for the 800 Hz tone in which the phase adjustment was made to null out all the 800 Hz tone in the error for a value of the relative input slope of -20 dB, and then left in this position for the measurement of (S/N). This second measurement is indicated as the modified measured (S/N) on Figure 4–19. The modified (S/N) for the 800 Hz tone shows very good agreement with the simulated speech signal both in the value of the maximum and the rms slope at which the maximum point occurs. It is thus concluded that, when the data is treated properly, the performance for the 800 Hz tone and for the simulated speech signal is very similar.

Recently, Jayant and Rosenberg(26) have reported a perceptual preference for DM systems which are operating beyond the peak in the (S/N) curve, which is computed on the basis of a mean squared error at a sampling frequency of 40 KHz. A perceptual (S/N) curve was presented along with the mean squared error (S/N) curve and the relationship of the two is very similar to the relationship of 800 Hz tone and simulated speech signal (S/N) curves in Figure 4–19. The peak in the perceptual performance was at a greater rms input slope and the curve decreased at a greater rate as the input slope is reduced below the value which corresponds to the maximum point on the (S/N) curve. The preference for operating in the overload region was attributed to the fact that granular
noise is perceivable by a listener as an additive background noise while slope-overload distortion exists only in relation to an original signal, which is not known to the listener. This is exactly the point noted between the methods used for measurement of the curves of Figure 4-19. In terms of an 800 Hz tone, it can be similarly stated that a listener cannot perceive phase shifts or slight gain variations. From this data it is concluded that the 800 Hz tone performance, based on the measurement techniques used, is a good description of the system performance relative to a listener, and the best performance is very near the peak on the 800 Hz tone (S/N) curve.

The delay between \( X(t) \) and \( X(t) \) decreases as the sampling frequency is increased. Thus, the difference between the two measurement techniques is less pronounced at higher sampling frequencies. Similarly, the listener preference cited previously nearly coincides with the mean-squared error signal-to-noise ratio at higher sampling frequencies.

1.10 Summary - LDM Performance

The performance of the LDM systems is summarized as follows:

1) The (S/N) is a monotonically increasing function of the input loading, increasing one dB for each dB increase in loading for values of input loading less than the optimum loading value (granular noise region).

2) The (S/N) is a monotonically decreasing function of the input loading, decreasing two dB for each dB increase in loading for values of input loading greater than the optimum loading value (slope-overload noise region).
3) The LDM$_2$ system exhibits a greater maximum (S/N) than the LDM$_1$ system for sampling frequencies equal to or greater than 30 KHz, with the advantage increasing as the sampling frequency is increased.

4) The simulation based on taking an ensemble average of waveforms yields a value of the (S/N) which very closely matches the measured performance.

5) The error spectral density determined from an ensemble average of waveforms shows very good agreement with the measured error spectral density.

6) The measured performance of the LDM$_1$ system for an 800 Hz tone signal is very similar to the published performance data for a simulated speech signal.

7) The optimum operating point for speech signals is very close to that value of input loading which corresponds to the peak of the (S/N) curve for an 800 Hz tone input signal.
CHAPTER V

DISCUSSION OF COMPANDED DELTA MODULATION

In Chapter 4 it was demonstrated that both the LDM\(_1\) and LDM\(_2\) systems exhibit a peak in the signal-to-noise ratio termed \((S/N)_{\text{max}}\) at the optimum loading point \(L_0\). The value of \(L_0\) is independent of the input frequency. From equations (3-11) and (3-12), the input loading is

\[
L = \frac{1}{f_s} \frac{\sqrt{2} X 2\pi f_1}{\Delta}. \tag{5-1}
\]

The rms value of the derivative of the sinusoidal input signal is

\[
\dot{X} = X 2\pi f_1, \tag{5-2}
\]

thus,

\[
L = \frac{\sqrt{2}}{f_s} \frac{\dot{X}}{\Delta}. \tag{5-3}
\]

Next, consider a Companded Delta Modulation or simply CDM system. The objective is to vary the step size \(\Delta\) at the same rate as \(X\) to maintain optimum loading (and thus maximum signal-to-noise ratio) over a wider range of \(X\). This is similar to the effect of the second integrator used in the LDM\(_2\) system. However, in the LDM\(_2\) system operating at
\( f_s = 40 \text{ KHz} \) the ratio of the largest to the smallest step size is on the order of 2, whereas in the CDM system a much larger variation in \( \Lambda \) is desired. The maximum and minimum values of \( \Lambda \) are \( \Lambda_{\text{max}} \) and \( \Lambda_{\text{min}} \). The companding coefficient \( q \) is defined as

\[
q = \frac{\Lambda}{\Lambda_{\text{min}}}, \quad 1 < q < \frac{\Lambda_{\text{max}}}{\Lambda_{\text{min}}} \tag{5-4}
\]

and is the signal used to adjust \( \Lambda \). The maximum value of \( q \) is defined as

\[
q_{\text{max}} = \frac{\Lambda_{\text{max}}}{\Lambda_{\text{min}}} \tag{5-5}
\]

The amount of companding \( C \) is expressed in decibels and is defined as

\[
C = 20 \log (q_{\text{max}}) \tag{5-6}
\]

Since \( \Lambda \) is no longer constant, the input loading as previously defined no longer relates \( X \) to \( X_m \). The system is slope limited and thus the overload point is a function of the derivative of the input signal, \( \dot{X} \). The maximum system design rms value of \( \dot{X} \) is \( \dot{X}_{\text{max}} \). A new parameter \( Q \) is defined as

\[
Q = \frac{X}{\dot{X}_{\text{max}}} \tag{5-7}
\]

Note that

\[
Q = \frac{\dot{X}}{\dot{X}_{\text{max}}} = \frac{2\pi f_1 X}{2\pi f_1 X_{\text{max}}} = \frac{X}{X_{\text{max}}} \tag{5-8}
\]
where \( x_{\text{max}} \) is the maximum design value of the rms input voltage at frequency \( f_1 \). The parameter \( Q \) in the CDM system is very similar to the loading \( L \) of the LDM system and is called the CDM system input loading or simply input. Frequently, \( Q \) is expressed in decibels as

\[
Q \text{ (dB)} = 20 \log Q, \quad 0 < Q < 1. \quad (5-9)
\]

The loading \( L \) of the encoder (and decoder) when both \( X \) and \( \Delta \) are at the maximum values is defined as

\[
L_c = \frac{\sqrt{2}}{f_s} \frac{x_{\text{max}}}{A_{\text{max}}}. \quad (5-10)
\]

To maintain the loading at \( L_c \) as \( Q \) is varied requires that

\[
q = Q q_{\text{max}}. \quad (5-11)
\]

This expression is determined by substituting equation (5-10) for \( L \) in (5-3) and combining with equations (5-4), (5-5), and (5-7).

At this point the PCM systems are considered in order to determine an approximate value of \( C \) so that the CDM system will exhibit an 800 Hz tone performance which is similar to that of 7 bit log PCM.

When the quantizing levels of PCM are equally spaced on a voltage scale, the system is called linear-PCM. It is well known that the quantizing noise of the linear-PCM system is constant independent of input, and the signal-to-noise ratio is a monotonically increasing function of the input loading. This is very similar to the LDM systems for values of input loading less than \( L_0 \). When the quantizing levels of the PCM system are arranged on a logarithmic scale the system is termed log-PCM.
For the log-PCM systems the signal-to-noise ratio is lower at the maximum point but is nearly flat over a wide range of input loading. The ratio of the largest difference in adjacent quantizing levels to the smallest difference in adjacent quantizing levels is defined as the companding constant \( \mu \). Good quality transmission is achieved for 7 bit log-PCM systems with \( \mu = 100 \) (\( \mu = 1 \) corresponds to linear-PCM). This is the quality standard that will be applied to the DM systems. The signal-to-noise ratio curves for the PCM system are illustrated in Figure 5-1. As discussed in Chapter II, systems suitable for the transmission of speech signals should be capable of transmitting an 800 Hz test tone over a 50 dB power range, with a weak, average and strong signal -50 dB, -30 dB, and -10 dB below the maximum, respectively. These points are indicated on Figure 5-1.

Next, \( L_c \) is set equal to \( L_0 \) and it is assumed that a system can be designed to obey equation (5-11) for \( Q \) in the range \( \frac{1}{q_{\text{max}}} \leq Q \leq 1 \), and \( q = 1 \) for \( Q \leq \frac{1}{q_{\text{max}}} \). The (S/N) curves of the CDM system are determined from the (S/N) equations of the LDM system (equations (4-23) and (4-42)) with \( \Delta = q_{\text{min}} \). Curves of the 800 Hz tone (S/N) performance for the CDM_1 and CDM_2 systems for a 40 KHz sampling frequency and the 7 bit log-PCM system are shown in Figure 5-2. From the Figure it is seen that if \( C = 39 \) dB and 35 dB for the CDM_1 and CDM_2 systems respectively, the performance of the CDM systems are comparable to 7 bit log-PCM. From this it is concluded that about 40 dB of companding is required if the CDM systems are to operate at \( f_s = 40 \text{ KHz} \).
Figure 5-1

Signal-to-noise ratio linear PCM and log-PCM ($\mu = 100$), 800Hz tone input.
SIGNAL-TO-NOISE RATIO $S/N$ (dB)

INPUT $Q$ (dB)

FIGURE 5-2

SIGNAL-TO-NOISE RATIO, 7 BIT LOG PCM, CDM$_1$, CDM$_2$,
800Hz TONE INPUT
The challenge is to design a system which obeys equation (5-11); that is, $Q$ is measured and the step size is adjusted such that $q = Q q_{\text{max}}$. This type of system is well known and has been used to adjust the gain of voice frequency repeaters such that the average signal level over a channel remains relatively constant. The controlling circuitry is called a syllabic compandor or simply compandor. (3) The syllabic compandor adjusts the system operating point based on the average signal power for each spoken syllable. The preceding discussion has not considered the transient response of the compandor. This response is described by two time constants, $\tau_a$, the attack time constant, and $\tau_r$, the recovery time constant. Normally the attack time constant is shorter than the recovery time constant because: 1) the deterioration in the $(S/N)$ is greater for the overload condition than the deterioration due to underloading; and 2) it is desired that the system maintain operation near the average operating level between syllables and return to the idle level only during pauses in the speech. The attack time must be short enough such that only a small part of the beginning of a spoken phrase is highly distorted. The recovery time constant must be short enough such that background noise is reduced rapidly during pauses. International standards (3) specify that the time required for the system to respond to a change from $Q = 0$ to $Q = Q_{\text{max}}$ (attack) be equal to or less than 5 milliseconds and the time required for the system to respond to a change from $Q = Q_{\text{max}}$ to $Q = 0$ (recovery) be equal to or less than 22.5 milliseconds.
In Chapter 4 it was noted that the LDM system performance deteriorates approximately twice as much for an overload condition as for an underload condition. Thus, if the system is operating near the optimum loading point, the same reasoning as 1) and 2) in the preceding paragraph apply, and the attack time constant should be shorter than the recovery time constant. The ratio of the recovery time constant to the attack time constant is defined as $\gamma$.

$$\gamma = \frac{r_r}{r_a}$$

The value of $\gamma$ should be greater than 2 if the system is operated at the maximum point on the (S/N) curve ($L_c = L_o$). However, if the system is operated below the peak on the (S/N) curve (underloaded, $L_c < L_o$), a smaller value of $\gamma$ is suitable. The parameters of the CDM system are, in addition to the parameters of the LDM system:

1) The amount of companding $C$,
2) The steady-state operating point $L_c$,
3) The attack time constant $r_a$,
4) The ratio of the attack time constant to the recovery time constant $\gamma$.

If the attack and recovery time constants are long with respect to the sampling period and short with respect to the syllable length (which is the case), the CDM system can be treated as an LDM system where the step size is adjusted properly for each syllable that the system transmits. This point is discussed further in Chapter 6.
Several methods of generating the signal $q$ have appeared in the literature. These can be divided into two types; the analog companding system, and the digital companding system. The analog companding system is illustrated in Figure 5-3. The system operates by taking the derivative of the input signal at the encoder, rectifying and low pass filtering the derivative to generate a voltage $q$ which is proportional to the average value of $|X(t)|$. This voltage is then used to control the step size $\Delta$ at the encoder. To achieve the same control at the decoder, the voltage $q$ is passed to a second LDM encoder at the transmitter which encodes the voltage $q$ and then to a second LDM decoder at the receiver where the voltage $q$ is reconstructed and used to control the step size at the receiver. This system has been constructed and tested by Brolin and Brown. (6) Their test system used a 96 KHz sampling rate for the speech DM system and a 6 KHz sampling rate for the LDM system used for the voltage $q$. The two digital signals are multiplexed on the same digital line. The system uses the double integration arrangement for both DM encoders-decoders. With 26 dB of companding, the 800 Hz tone performance of this system equals that of 7 bit log-PCM. Less than 40 dB of companding is required because of the high sampling frequency. This system has the disadvantage of requiring a framing technique to unsort the speech and companding samples at the receiver.

The digital companding system is illustrated in Figure 5-4. The system contains identical companding circuits at the encoder and decoder. The input to the compandor is the binary coded speech signal $a_i$. The output of the compandor is a voltage $q$ which is related to the input
TRANSMITTER

\[ f_s = 96\text{kHz} \]

INPUT \( x(t) \)

\[ \frac{d}{dt} \]

LOW PASS FILTER

LDM SPEECH ENCODER

STEP ADJUSTMENT

MULTIPLEX

RECEIVER

\[ f_s = 6\text{kHz} \]

OUTPUT \( y(t) \)

LDM SPEECH DECODER

MULTIPLEX

LDM COMПANDING DECODER

STEP ADJUSTMENT

LDM COMПANDING ENCODER

MULTIPLEX

ANALOG COMПANDING SYSTEM

FIGURE 5-3
**DIGITAL COMPANDING SYSTEM**

**FIGURE 5-4**

**VARIOUS DIGITAL COMPANDING METHODS**

**FIGURE 5-5**
signal, the specific relationship being a function of the particular system. The voltage \( q \) is then used to control the step size \( \Delta \). To date, three different types of digital compandors have appeared in the literature. These are illustrated in Figure 5-5 and are summarized below.

1) Tomozawa and Kaneko\(^{(37)}\) - The basic LDM system uses double integration. The companding control is achieved by passing the binary coded speech signal to an integrator which essentially reconstructs the input signal. The output of the integrator is rectified and low pass filtered to generate a voltage \( q \) which is proportional to the average level of the input voltage. Voltage \( q \) is then used to control the system step size. This system was constructed and tested using a sampling frequency of 56 KHz. The attack time constant for this system was 1 millisecond and the recovery time constant was 4 milliseconds \((\gamma = 4)\). Performance comparable to 7 bit log PCM was reported using 26 dB of companding. This system has the drawback that the adjustment of the step size is based on the input voltage. As a result, the system tends to be underloaded at input frequencies less than 800 Hz.

2) Abate\(^{(1)}\) - This system has been called Adaptive Delta Modulation (ADM). The arrangement uses the basic single integration LDM system. Companding control is achieved by generating a binary signal \( b_i \) at each sampling time.
The value of $b_i$ is plus one if the present and immediate past signal samples are of the same sign, otherwise $b_i$ is minus one. A running total of $E_b_i$ is maintained. The value of $E_b_i$ is either zero or positive. If the number of minus ones exceeds the number of plus ones, then $E_b_i$ is zero. The value of $q$ is then determined by the equation

$$q = r^{E_b_i}.$$  \hspace{1cm} (5-13)

A value of $r$ on the order of 1.33 is used in the system. This system has been simulated for speech signals and found to perform comparable to 7 bit log-PCM at sampling frequencies on the order of 40 KHz. The value of $\gamma$ for this system is one.

3) Greefkes\textsuperscript{(19)} - This system uses the double integration LDM system. Companding is achieved by generating a signal $b_i$ each sampling time. The value of $b_i$ is set at plus one until the three most recent speech samples are of the pattern either one-zero-one or zero-one-zero at which time $b_i$ is set and held at minus one until the sample pattern of four adjacent pulses of the same sign occurs. The pulse train $b_i$ is then filtered using an integrator with time constants on the order of several milliseconds to generate the control voltage $q$. Good performance has been reported for this system operating at a sampling frequency of 40 KHz using 40 dB of companding.
The starting point for this investigation is Systems 2) and 3) above. It is noted that the method of generating \( q \) is radically different both in terms of the number of speech samples which are employed and the translation of the digital signal \( b_i \) to the voltage \( q \). Specifically, the following questions are considered:

1) **What is the optimum number of voice samples to examine and how should this be done to generate \( b_i \)?**

2) **What is the relationship of the steady-state operating point \( L_c \) to the ratio of the attack to recovery time constants \( \gamma \)?**

3) **What is the best method to convert the signal \( b_i \) to the control signal \( q \)?**

These questions are the subject of the investigation of Chapter 6. In this investigation, the amount of companding \( C \) is fixed at 40 dB and the sampling frequency is fixed at 40 KHz. Both the CDM\(_1\) and CDM\(_2\) systems are considered.
CHAPTER VI

INVESTIGATION OF COMPANDED DELTA MODULATION

In Chapter 5 it was shown that approximately 40 dB of companding is required for companded delta modulators operating at a 40 KHz sampling frequency to equal the signal-to-noise ratio performance of 7 bit log-PCM. In the first part of this chapter the steady state loading Lc that is maintained using five different two-state companding rules is determined. A two-state companding rule is one in which the step size $\Delta$ is either increased or decreased after each sample. Only digital companders in which the step size adjustment is determined by examining the present and past n-1 binary coded speech samples are considered. Included in these five rules are those which have been studied by Abate\(^{(2)}\) for the CDM\(_1\) system and by Greerkes\(^{(19)}\) for the CDM\(_2\) system. It is found that several of the rules are usable for both the CDM\(_1\) and CDM\(_2\) systems. For both the CDM\(_1\) and CDM\(_2\) systems, a rule which adjusts the step size based on the three most recent coded speech samples appears to be the best. This analysis is based entirely on computer simulation. To verify the accuracy of the simulation, the performance of laboratory models using one of the companding rules for both the CDM\(_1\) and CDM\(_2\) systems are measured. With the exception of the value of the signal-to-noise ratio,
the measured performance is in agreement with that predicted by simulation. The measured (S/N) was found to be less than computed for the LDM system. This is the result of a ripple on the companding control voltage. Performance equal to that calculated is observed when the ripple is reduced by increasing the companding time constants.

In the second part of this chapter the transient response of the compandor is investigated. Compandors which change the step size a fixed amount each sampling period are called linear compandors. If the step size adjustment is such that the ratio of the change in step size over one sampling period to the companding coefficient $q$ is fixed, the compandor is called logarithmic. It is shown that a linear compandor is unsuitable for systems that require on the order of 40 dB of companding. Further, it is shown that a logarithmic compandor offers the ideal characteristic. A combined log-linear compandor is proposed which is nearly as economical as the linear compandor but approaches the operating performance of the logarithmic compandor. The compandors are divided into two types; the continuous type in which $\Delta$ takes on all values between $\Lambda_{\min}$ and $\Lambda_{\max}$, and the discrete type in which $\Delta$ takes on only discrete values between $\Lambda_{\min}$ and $\Lambda_{\max}$. The laboratory models discussed previously use the continuous type linear compandor. The proposed log-linear compandor is of the discrete type.

The third part of this chapter is devoted to the investigation of a new three-state companding rule. A three-state rule is one in which the step size is increased when an overload condition occurs, the step size is reduced when an underload condition occurs, and otherwise, the
step size is not changed. The steady state operating point for this rule is near the optimum loading point for both the CDM$_1$ and CDM$_2$ systems. The transient response of this system is determined for a log-linear discrete compandor by computer simulation. It is shown that the ripple in the companding control signal is nearly one-half that for a similar two-state rule and thus the system should have a larger steady state signal-to-noise ratio. Based on this, it is concluded that the three-state rule is better than a two-state rule.

Next, a brief investigation of the effects of channel errors indicates that both the attack and recovery time constants must be very long with respect to sampling period to avoid large tracking errors between the encoder and decoder due to channel errors. The tracking error appears to the user as a transmission gain or loss.

Finally, the steady state signal-to-noise ratio vs sampling frequency curves for the CDM$_1$, CDM$_2$, and PCM systems are presented. From this data it is concluded that performance equal or superior to 7 bit log-PCM can be achieved at a lower sampling frequency. Further, based on considerations of the ease of convertability from CDM to PCM, it is concluded that the single integration CDM system is the better choice of the two CDM systems. This system is more economical, requires less channel bandwidth than the PCM system, and thus is considered a good alternative to PCM.

The conclusions reached in this chapter are based on the system performance for an 800 Hz tone signal. Although, from the discussion in Section 4.9, it is felt that this is a good indication of the perceptual
performance, tests using voice signals similar to those discussed in Reference 26 are required to verify the conclusions reached here. Such testing is beyond the scope of this work.

6.1 Performance Criteria

To judge the various companding systems, the following performance objectives are used.

1) Maximum signal-to-noise ratio for an 800 Hz tone input signal.

2) The attack time $T_a$ is defined as the time required for step size to reach a value that corresponds to a steady state signal-to-noise ratio which is 6 dB less than the maximum value when the system is subjected to a 12 dB increase in the input. Similarly, the recovery time, $T_r$, is defined as the time required for the step size to reach a value which corresponds to a steady state signal-to-noise ratio which is 6 dB less than the maximum value when the system is subjected to a 12 dB input reduction. Both times $T_a$ and $T_r$ should be small with respect to the average syllable duration of 7.6 milliseconds.

3) A value of $T_r > T_a$ is suitable; however, $T_r$ should not greatly exceed $2T_a$ (refer to discussion of syllabic compandor in Chapter 5).

4) Both steady state and transient performance should be independent of the average input level $Q$.

5) The effects of channel errors should be minimized.
6) The system should permit economical implementation.

7) The system should lend itself to direct conversion to 7 bit log-PCM (refer to Section 6.7).

6.2 Steady-State Performance, Two State Companding Rules

The CDM system is illustrated in Figure 6-1. A n-bit shift register keeps a record of the present and past n-1 bits of the binary coded speech signal. The output of the shift register is passed to two masks, Ms and Mr, which are termed the set and reset masks. Based upon the particular pulse pattern in the shift register, three mutually exclusive events can occur. These are:

1) Mask Ms has output one,
2) Mask Mr has output one,
3) Both masks have output zero.

The output of the masks are used to control the flip-flop. A one output of mask Ms sets the flip-flop and a one output of mask Mr resets the flip-flop. The output of the flip-flop, b₁, is either plus one (set state) or minus one (reset state). Signal b₁ is termed the binary companding signal. Signal b₁ is low pass filtered by filter Hc and then amplified to produce voltage q-1. If the output of filter Hc is negative, q-1 is zero. The output of the filter is then added to voltage 1 to produce q, the step size control voltage. In terms of the LDM circuits (Figures 4-1, 4-11), voltage q connects to a resistor which carries current I and replaces the constant current generator used in the LDM circuits.
INPUT X(t) ENCODER

BINARY CODED SIGNAL $a_j$

FILTER $H_C$

AMP

(q-1)

CLOCK

INPUT

BIT SHIFT REGISTER

COMPANDOR

BASIC CDM SYSTEM - TWO-STATE COMPAUNDING RULE

FIGURE 6-1
The attack time constant $\tau_a$ is defined as the time required for $q$ to increase from 1 to $q_{\text{max}}/2.73$. During an overload condition the output of the flip-flop $b_1$ is plus one, and the output of filter $H_C$ is rising at a constant rate ($H_C$ is an integrator) as are voltages $q-1$ and $q$. The increase in the step size $\Delta$ over one sampling period is defined as $g$, where

$$g = \frac{\Delta}{\tau}.$$  \hfill (6-1)

In terms of the attack time constant,

$$g = \frac{q_{\text{max}}}{2.73} \frac{\tau}{\tau_a}. \hfill (6-2)$$

The attack time constant is chosen to be approximately 2 milliseconds. Thus, from equation (6-2), at a sampling frequency of 40 KHz and with 40 dB of companding ($q_{\text{max}} = 100$), $g$ is 0.46. The decrease in the step size over one sampling period is $g/\gamma$, where $\gamma \geq 1$ ($\gamma$ is the ratio $\tau_T/\tau_a$). With $q$ greater than 2 or 3, the change in the step size each sampling period is small and the companding circuit and the delta modulator can be treated independently when the system is operating in the steady state ($Q$ fixed). The compandor described here is termed linear because $g$ is independent of $q$ and continuous since $q$ can take on all values between 1 and $q_{\text{max}}$.

The binary coded speech samples of the LDM systems tend to have the pattern -01010---- when the system is lightly loaded. When the system is overloaded, the pattern tends to be ---1111--- or ---0000---. Thus, mask $M_S$ should sense when a number of identical pulses occur and cause $q$
to increase. Conversely, mask $M_r$ should sense when the pulse pattern contains alternating pulses and cause $q$ to decrease. Five combinations of set and reset masks were investigated. These are indicated in Table 6-1. Each combination is termed a companding rule or simply rule and is given one of the numbers I through V. The set mask lengths are from two to four bits. The reset mask lengths are two and three bits. Initial consideration of longer masks indicated a severe sensitivity to input frequency variations. This is not unreasonable. At an input frequency of 1600 Hz, only 25 samples are devoted to the entire input period ($f_s = 40$ KHz). For a five bit set mask and a three bit reset mask a minimum of eight sample periods are required to reset and then set the flip-flop immediately after it has been set. The time duration of eight sample periods is not negligibly small with respect to the 25 samples which are allocated to each period of a 1600 Hz tone. This effect is more severe for either longer mask lengths or higher input frequencies. The behavior of the rules indicated in Table 6-1 is nearly independent of the input frequency.

The first step in analyzing this system is to consider the relationship of the binary companding signal $b_i$ to the system loading $L$ when the compandor is disconnected (open loop). A measure of $b_i$ is the ratio of the amount of time $b_i$ is plus one to the amount of time $b_i$ is minus one. A new parameter $\bar{B}(L)$ is defined as this ratio,

$$\bar{B}(L) = \lim_{n \to \infty} \frac{\sum_{i=0}^{n} (b_i + 1)}{2n}, \quad b_i = +1 \text{ or } -1.$$  

(6-3)
### TABLE 6-1

**TWO-STATE COMPANDING RULES**

**SET AND RESET MASK BIT PATTERNS**

<table>
<thead>
<tr>
<th>RULE</th>
<th>SET BIT PATTERN (MR)</th>
<th>RESET BIT PATTERN (MG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>11 or 00</td>
<td>10 or 01</td>
</tr>
<tr>
<td>II</td>
<td>111 or 000</td>
<td>-10 or -01</td>
</tr>
<tr>
<td>III</td>
<td>111 or 000</td>
<td>101 or 010</td>
</tr>
<tr>
<td>IV</td>
<td>1111 or 0000</td>
<td>--10 or --01</td>
</tr>
<tr>
<td>V</td>
<td>1111 or 0000</td>
<td>-101 or -010</td>
</tr>
</tbody>
</table>
The range of $\overline{B}(L)$ is 0 to 1. The percentage of the time $b_i = -1$ is $\overline{B}(L)$ and the percentage of time $b_i = 1$ is $(1 - \overline{B}(L))$. When the system is in the steady state the average value of $q$ is constant.

This means that the increase in $q$ when $b_i$ is plus one must be balanced by the decrease in $q$ when $b_i$ is minus one. The increase when $b_i$ is plus one is inversely proportional to the attack time constant $\tau_a$. The decrease in $q$ when $b_i$ is minus one is inversely proportional to the recovery time constant. The input loading for the encoder in the steady state is denoted by $L_c$. Thus, in the steady state

$$\frac{\overline{B}(L_c)}{\tau_a} = \frac{(1 - \overline{B}(L_c))}{\tau_r}.$$  \hfill (6-4)

The ratio of $\tau_r/\tau_a$ has been defined as $\gamma$ (equation (5-12)). Thus, (6-4) can be simplified to

$$\overline{B}(L_c) = \frac{1}{1 + \gamma}.$$  \hfill (6-5)

The function $\overline{B}(L)$ (equation (6-3)) has been determined for the LDM1 and LDM2 systems for each of the five two-state rules using the simulation described in Appendix B and is illustrated in Figures 6-2 and 6-3. The values indicated in these Figures are for an input frequency of 800 Hz and a sampling frequency of 40 KHz. The function $\overline{B}(L)$ for $f_1 = 400$ Hz and 1600 Hz was also determined and is essentially identical to that at $f_1 = 800$ Hz. Using these values for $\overline{B}(L)$ and equation (6-5), one can determine the steady-state loading point $L_c$ as a function of $\gamma$. 


INPUT LOADING $L$(dB)

FIGURE 6-2: FUNCTION $B(L)$, LDM$_1$ SYSTEM

$f_i = 800$Hz  \hspace{1cm} f_s = 40$ kHz
FIGURE 6-3: FUNCTION $B(L)$, LDM$_2$ SYSTEM

$f_i = 800$ Hz  
$f_s = 40$ kHz
6.2.1 CDM1 Steady State Operating Point

The steady state operating point for each of the five two-state companding rules is plotted in Figure 6-4. The characteristics of each are summarized in Table 6-2. All the rules with the exception of IV offer satisfactory operation near the optimum loading point $L_0$. Based on the performance criteria established previously, the best rule appears to be number II in that operation at $L_0$ is possible with $\gamma = 2$. However, it is questionable whether any difference between I, II, III, or V could be readily detected by a listener when speech signals are used.

6.2.2 CDM2 Steady State Operating Point

The steady state loading point $L_0$ as a function of $\gamma$ for the two-state rules in a double integration system is plotted in Figure 6-5. The characteristics of the best operating point for each rule are listed in Table 6-3. As with the CDM1 system, rule IV does not permit operation near the optimum loading and is rejected. Recall that in the LDM system the $(S/N)$ is flat over a 2 dB range in $L$ at the maximum point. All the rules permit operation in this region except rule IV, and on a steady state basis are equivalent. Rule III permits operation in the center of this flat $(S/N)$ region with a value of $\gamma$ of 1.5 to 2, and is deemed the best of the four. However, this judgment is subject to the same constraints noted for the CDM1 system.

6.2.3 Measured Performance

To verify that the steady state analysis which is based entirely on computer simulation is accurate, hardware models using rule II were measured. These models used circuitry which is essentially identical to
FIGURE 6-4:

STEADY STATE LOADING $L_c$ VS $\delta$, TWO-STATE COMPANDING RULES-CDM$_1$  
$f_i = 800$Hz, $f_s = 40$kHz
<table>
<thead>
<tr>
<th>RULE</th>
<th>Y</th>
<th>$I_c$</th>
<th>$\frac{S}{N}$ (EQUATION 4-23)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>1.4</td>
<td>-5 dB</td>
<td>29.5 dB</td>
</tr>
<tr>
<td>II</td>
<td>2.0</td>
<td>-3.5 dB</td>
<td>31.0 dB</td>
</tr>
<tr>
<td>III</td>
<td>1.5</td>
<td>-4.4 dB</td>
<td>30.1 dB</td>
</tr>
<tr>
<td>IV</td>
<td>NO ACCEPTABLE OPERATING POINT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>2.4</td>
<td>-3.5 dB</td>
<td>31.0 dB</td>
</tr>
</tbody>
</table>

TABLE 6-2

STEADY STATE OPERATING POINT

TWO-STATE COMPAANDING RULES-CIM1

($f_1 = 800$ Hz, $f_s = 40$ KHz)
FIGURE 6-5
STEADY STATE LOADING $L_c$ VS $\varepsilon$, TWO-STATE COM-PANDING RULES-CMDM$_2$  
$f_i = 800\text{Hz}  \quad f_s = 40\text{kHz}$
TABLE 6-3

STEADY STATE OPERATING POINT

TWO-STATE COMPANDING RULES-CIM₂

\((f₁ = 800 \text{ Hz}, f₂ = 40 \text{ kHz})\)

<table>
<thead>
<tr>
<th>RULE</th>
<th>(γ)</th>
<th>(L_C)</th>
<th>((S/N)(\text{EQUATION 4.42}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>1.4</td>
<td>-5.5 dB</td>
<td>34.5 dB</td>
</tr>
<tr>
<td>II</td>
<td>1.7</td>
<td>-4.8 dB</td>
<td>34.5 dB</td>
</tr>
<tr>
<td>III</td>
<td>2.3</td>
<td>-4.0 dB</td>
<td>34.5 dB</td>
</tr>
<tr>
<td>IV</td>
<td></td>
<td></td>
<td>NO ACCEPTABLE OPERATING POINT</td>
</tr>
<tr>
<td>V</td>
<td>3.0</td>
<td>-4.0 dB</td>
<td>34.5 dB</td>
</tr>
</tbody>
</table>
that indicated in Figure 6-1. Measurement techniques were identical to those used for the LDM system (refer to Section 4.5). The measured values of S/N and q together with calculated values are illustrated in Figures 6-6 and 6-7. The measured value of q agrees with the calculated curve over most of the input range. The deviation for very small inputs is attributed to hysteresis in the comparator which causes the zero input samples to be of the pattern 00110011 rather than the expected pattern 10101010. This situation occurs only for small step sizes, and because of the smaller idle circuit step size is more severe for the CDM$_2$ system. For both systems, the value of q for larger step sizes is nearly exactly as calculated and it is concluded that the data obtained from the computer simulations accurately reflects the circuit operation.

The agreement between the measured and calculated signal-to-noise ratio is not as was expected. The measured values are lower than calculated by one to two decibels. Close examination of the companding control voltage revealed a 1600 Hz ripple ($f_1 = 800$ Hz). The phase of this ripple is such that the maximum step size occurs at the peak in the input waveform and the minimum step size occurs near the steepest portion of the input waveform. This is the worst possible phase relationship. To determine if this ripple is the cause of the reduced (S/N), the time constants were increased to approximately 10 milliseconds. With these longer time constants the ripple is nearly eliminated, and the measured (S/N) increased to essentially the calculated values. The reduction of this ripple is the motivation for considering the three-state compandor which is described in Section 6.4.
COMPANDING COEFFICIENT $q$

SIGNAL-TO-NOISE RATIO (S/N)

$\begin{array}{c}
50 \\
40 \\
30 \\
20 \\
10 \\
\hline
-60 & -50 & -40 & -30 & -20 & -10 & 0
\end{array}$

$\text{INPUT } Q (\text{dB})$

$\text{CDM}_1 \text{ SYSTEM, (S/N) AND } q$

$f_i = 800\text{Hz}, f_s = 40\text{kHz}$

FIGURE 6-6
COMPANDING COEFFICIENT $q$, SIGNAL-TO-NOISE RATIO (S/N)

CALCULATED S/N

MEASURED S/N

MEASURED $q$

CALCULATED $q$

INPUT Q (dB)

CDM$_2$ SYSTEM, (S/N) AND $q$, $f_i = 800$Hz, $f_s = 40$kHz

FIGURE 6-7
These systems were also tested with voice signals. The compandors were found to perform as expected and the transmission quality appeared to be good. This did not involve comparisons with uncoded signals and can be interpreted only as meaning that the systems did not exhibit any unforeseen objectionable characteristics.

6.3 Transient Response - Nonlinear Compandors

Another major characteristic of the compandor is the manner in which the system responds to changes in the input $Q$. Consider the case where the system is subjected to a step increase in $Q$ of 12 dB. The coded speech signal will consist of a string of identical pulses (overload condition) and the flip-flop will be set to the one state with $b_1 = +1$. (Refer to Figure 6-1.) The voltage $q$ will start increasing, an increase of $0.46$ units per each sampling period if an attack time constant of 2 milliseconds is used, (equation (6-2)). If it is assumed that the coded speech signal during overload is composed entirely of the binary digits one, or all zero, then the value of $q$ is

$$q = q_0 + g_i, \quad i = 0, 1, \ldots$$

(6-6)

where $q_0$ is the value of $q$ in the steady state just prior to the overload and $i$ is the number of samples since the overload. Equation (6-6) is valid until the new steady state operating point is reached. Since the system is subjected to a 12 dB power jump, the new steady state operating point of the system is $4q_0$. Strictly speaking, equation (6-6) is not valid since the sign of the binary speech signal changes after every peak in the sinewave input. This causes the increase in $q$ to be somewhat less
than indicated. However, this discrepancy does not alter the point of concern here. Equation (6-6) can be written as

\[ \frac{q}{q_0} = 1 + \frac{g}{q_0} i, \quad i = 0, 1, \ldots \]  

(6-7)

Clearly, the number of samples required for \( q \) to respond to the 12 dB change (\( q/q_0 \) to reach \( \frac{q_{\text{max}}}{4} \)) is a function of \( q_0 \). The value of \( q_0 \) can range from 1 to \( \frac{q_{\text{max}}}{4} = 25 \) (12 dB below the maximum input). Equation (6-7) is plotted on Figure 6-8 for various values of \( q_0 \). Several CDMS systems have used this technique (refer to discussion of Chapter 5). However, most of these systems operated at sampling frequencies much greater than 40 KHz, and typically 26 dB or less of companding is required. For 26 dB of companding, the curves \( q_0 = 1 \) and \( q_0 = 5 \) represent the extremes. The curves for \( q_0 = 1 \) and \( q_0 = 5 \) are different however, it is doubtful that the difference would be noticeable to a listener. On the other hand, the ratio of the number of samples to reach steady state is 25 times greater for \( q_0 = 25 \) than for \( q_0 = 1 \). This would certainly be perceptible and is undesirable.

The parameter \( g \) was defined (equation (6-1)) as the change in \( q \) over one sampling period. To achieve a transient response which is independent of the operating point, the factor \( g/q \) must be constant. This can be accomplished in two ways. The first way is to make the relation between \( \Delta \) and \( q \) nonlinear. That is,

\[ \Delta = q^2 \Delta_{\text{min}}, \]  

(6-8)

where

\[ q^2_{\text{max}} = \frac{\Delta_{\text{min}}}{\Delta_{\text{max}}}. \]  

(6-9)
STEP RESPONSE OF LINEAR COMPANDOR \( g = 0.46 \)

FIGURE 6-8
With such a relation, (equation (6-1)),

\[ g = \frac{dA}{dt} \tau = 2q \Delta_{\text{min}} \tau, \]  

(6-10)

and \( g/q \) is constant. One method to achieve a characteristic similar to equation (6-8) is to use a nonlinear resistor to generate the step control circuit (refer to Figure 6-1). A semiconductor diode could be used, however, the characteristics of these devices are both temperature sensitive and difficult to reproduce to the close tolerances which are required to insure good encoder-decoder tracking in a system that requires 40 dB of companding.

A second method is to make \( g \) a function \( q \) directly. This is the technique used by Abate\(^2\). That is,

\[ q = r^{E_{bi}} \]  

(6-11)

and

\[ \Delta = q \Delta_{\text{min}}, \]  

(6-12)

where \( r \) is a constant. For this arrangement,

\[ g = rq \Delta_{\text{min}} \]  

(6-13)

and

\[ g/q = r \Delta_{\text{min}} = \text{constant}. \]  

(6-14)

To generate a characteristic similar to that indicated in (6-11), digital techniques are used and \( \Delta \) is restricted to discrete values. The characteristic in (6-11) is referred to as a logarithmic or log characteristic as compared to the linear characteristic in equation (6-6). The
discrete compandor uses a binary counter to generate $E_{bi}$. The step size $\Delta$ is generated by decoding the binary code to a one-out-of-$p$ code where

$$ p = 2^m $$

and $m$ is the number of cells of the counter. Each of the $p$ elements of the one-out-of-$p$ code is used to control a switch which connects a voltage source through separate resistors which carry current

$$ I = I_0 r^i, \quad i = 0, 1, \ldots, p-1. \quad (6-16) $$

Current $I_0$ is that value for which the step size $\Delta$ is at its minimum $\Delta_{\text{min}}$. For 40 dB of companding, $p$ and $r$ are related by

$$ r^p = \frac{\Delta_{\text{max}}}{\Delta_{\text{min}}} = q_{\text{max}} = 100. \quad (6-17) $$

The attack time constant is related to $r$ as (equation (6-2) revised)

$$ r \tau_a / \tau = 100/2.73. \quad (6-18) $$

The value of $\tau_a$ for various values of $p$ at a sampling frequency of 40 KHz is listed in Table 6-4. A value of $\tau_a = 1.25$ ms requires a $m = 6$ stage counter. For $m = 6$, the value of $p$ is $6^4$. The use of $6^4$ switches and $6^4$ separate resistors is expensive and a compromise in order. If a linear relationship is used, one merely uses one switch and one resistor for each stage of the binary counter. The current produced for each cell which is set to one is

$$ I_j = I_0 2^j, \quad j = 1, 2, \ldots, m. \quad (6-19) $$
## TABLE 6-4

RELATION OF COUNTER LENGTH m TO PARAMETER r AND ATTACK TIME CONSTANT \( \tau_a \).

40 dB OF COMPANDING, \( f_g = 40 \) KHz

<table>
<thead>
<tr>
<th>COUNTER LENGTH m</th>
<th>( 2^m )</th>
<th>r</th>
<th>( \tau_a/r )</th>
<th>( \tau_a ) (MILLISECONDS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>8</td>
<td>1.78</td>
<td>6.25</td>
<td>0.15</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>1.334</td>
<td>12.5</td>
<td>0.30</td>
</tr>
<tr>
<td>5</td>
<td>32</td>
<td>1.155</td>
<td>25.6</td>
<td>0.62</td>
</tr>
<tr>
<td>6</td>
<td>64</td>
<td>1.0745</td>
<td>50.5</td>
<td>1.25</td>
</tr>
<tr>
<td>7</td>
<td>128</td>
<td>1.0366</td>
<td>100</td>
<td>2.5</td>
</tr>
<tr>
<td>8</td>
<td>256</td>
<td>1.0172</td>
<td>200</td>
<td>5</td>
</tr>
</tbody>
</table>
The currents from each cell are summed to produce the current $I$ which controls the step size. This arrangement requires $m$ switches and $m$ resistors as compared to $2^m$ resistors and switches for the logarithmic rule.

This problem is very similar to that faced by the early designers of PCM systems. Recall that the best performance for the PCM system is achieved when the levels are arranged on a logarithmic scale. Early models of PCM used semiconductor diodes to generate the desired nonlinearity. For the reasons mentioned previously, these proved to be undesirable. Later models of the PCM system used discrete techniques. The 7-bit sample is stored and then translated to a current which is proportional to the sample value. Rather than arrange every level to follow a logarithmic rule, groups of adjacent levels are related linearly and the relationship of adjacent groups is logarithmic. This relationship is called segmented logarithmic. A segmented compandor requires $j + 2^k$ switches where $2^j$ is the number of levels per linear segment and $k$ is the number of segments. Note that $2^j + 2^k = 128$ for 7 bit PCM.

It is proposed that this same technique be applied to the discrete compandor for the same reason - economy. The number stored in the counter after the $i$th sample is defined as $u_i$. The relationship of $q$ to $u_i$ for a 64 level discrete compandor is illustrated in
Figure 6-9 for the linear, logarithmic and a proposed 4 segment log-linear compandor. The equation for \( q \) for the 4 segment compandor is

\[
q = \begin{cases} 
1 + 0.413 u_i, & 0 \leq u_i \leq 15, \\
7.6 + 0.826 (u_i - 16), & 16 \leq u_i \leq 31, \\
20.8 + 1.65 (u_i - 32), & 32 \leq u_i \leq 47, \\
47.2 + 3.30 (u_i - 48), & 48 \leq u_i \leq 63.
\end{cases} 
\tag{6-20}
\]

A circuit for implementing equation (6-20) is illustrated in Figure 6-10.

From Figure 6-9 it is seen that the weighting of each linear region is on the ratio of two to one. Thus, 24 dB of companding is achieved logarithmically and 16 dB is achieved linearly. The total number of switch resistor combinations is 6 for the linear compandor, 64 the logarithmic compandor and 8 for the segmented compandor. The system transient response for the linear and log-linear compandors is compared in the next section for a new three-state companding rule. The steady state behavior for each of these compandors is identical and the computations of Section 6.1 are applicable to each.

6.4 Three-State Companding Rule

For the two-state rules, the step size is either increased or decreased on each sample. For the three-state rule, the third state is that of neither increasing or decreasing the step size. The reason for using a three-state rule is to reduce the ripple on \( q \) and thus increase the steady state signal-to-noise ratio. Implementation is based
COMPANDING CHARACTERISTIC - LINEAR, LOGARITHMIC AND LOG - LINEAR COMPA NDORS

FIGURE 6-9
CURRENT I USED TO CONTROL A

+1 INCREMENT

-1 DECREMENT

CLOCK f_s

UP-DOWN COUNTER

BLOCK DIAGRAM - PROPOSED LOG-LINEAR DISCRETE COMPANDOR

FIGURE 6-10
on the use of the 4 segment discrete compandor described in the previous section. A system block diagram is illustrated in Figure 6-11. The system consists of a n-bit shift register which stores the present and past (n-1) binary coded speech samples; two masks, mask Ma the attack mask and mask Mr the recovery mask, and the counter-switch-register circuitry to generate the step size control current. The output of mask Ma, ba, is a binary one if an overload condition exists. Otherwise, ba is a binary zero. When ba is one the counter is advanced. The output of mask Mr, br, is a binary one if an underload condition exists. Otherwise, br is a binary zero. Signal br is passed to a divide circuit where it is divided by the factor \( \gamma \) (ratio of \( r_f/r_a \)). For example, if \( \gamma \) is 2 the divider circuit consists of an additional counter cell. For \( \gamma = 2 \), one is subtracted from the counter every other time sample that ba is a one. Using this arrangement, values of \( \gamma = 1, 2, 3, ..., \) are possible. Other values of \( \gamma \) can be implemented by using a divider on ba also, for example \( \gamma = 1.5 \) is implemented by dividing br by 3 and ba by 2.

Mask Ma should sense when an overload exists. Therefore, ba should be one when a string of voice samples of the same sign occur. Conversely, mask Mr should sense when the system is underloaded. Therefore, br should be one when a string of voice samples of alternating sign are encountered. A preliminary investigation revealed that a mask length of three has the desired characteristic of indicating an overload/underload condition over a wide range of input frequencies. This is the only three-state rule which has been
INPUT $X(t)$ \to CHANNEL

$K_q$ \uparrow

INCREMENT

I($q$) \uparrow

DECREMENT

SWITCH-RESISTOR NETWORK

COUNTER

DECIMATE

DIVIDE BY $\gamma$

MASH $M_a$

MASH $M_f$

$\alpha_i$ \rightarrow TO CHANNEL

SHIFT REGISTER

THREE-STATE COMPANDOR

FIGURE 6-11
considered in detail. The variable $b_a$ is one only when the present and past two coded speech samples are of the pattern 111 or 000. The variable $b_r$ is one only when the most recent three coded speech samples are of the pattern 101 or 010.

To determine the system steady-state loading $L_c$, an approach similar to that for the two-state is used. The parameter $\overline{B}(L)$ is defined as

$$\overline{B}(L) = \lim_{n \to \infty} \frac{\sum_{i=0}^{n} b_{ai}}{\sum_{i=0}^{n} a_{bi} + \sum_{i=0}^{n} b_{ri}} \quad i = 0,1,2,... \quad (6-21)$$

The steady-state loading $L_c$ is that value of $L$ for which (refer to equation (6-5))

$$\overline{B}(L) = \frac{1}{1+\gamma} \quad (6-22)$$

The function $\overline{B}(L)$ was determined using the simulation described in Appendix B and is illustrated in Figure 6-12. These curves are for $f_1 = 800$ Hz and $f_s = 40$ KHz. The steady-state loading point $L_c$ as a function of $\gamma$ was then determined using equation (6-22) and the value of $\overline{B}(L)$ from Figure 6-12. This function is shown in Figure 6-13. With reference to Figure 6-13, the best operating point for both the CDM$_1$ and CDM$_2$ systems is for $1.5 < \gamma < 2$. For both systems, the steady state loading point $L_c$ is approximately one dB below $L_0$, very close to the perceptually best operating point as discussed in Chapter 4.
LOADING $L\,\text{dB}$

FUNCTION $\bar{B}(L)$ FOR THREE-STATE COMPAANDING RULE.

$f_i=800\,\text{Hz}$ $f_s=40\,\text{kHz}$

FIGURE 6-12
STEADY STATE LOADING $L_C$ VRS $\gamma$ - THREE-STATE

COMPANDING RULES $f_1=800\text{Hz}$ $f_s=40\text{kHz}$

FIGURE 6-13
The transient response of this system was calculated using the simulation discussed in Appendix B and is illustrated in Figures 6-14 and 6-15 for the CDM\textsubscript{1} and CDM\textsubscript{2} systems, respectively. The Figures show the value of the companding control signal \( q \) when the system is subjected to a 12 dB step of 800 Hz tone. These curves are the average of five responses where the input step for each response occurred at different input phases between 0 and \( \pi \). The discrete 4 segment log-linear compandor was used with the operating point from 6 dB below to 6 dB above the break point between two linear segments. The attack time constant for these curves is 1.25 milliseconds (i.e., Table 4-1, 40 dB of companding with 6 bit counter).

For the CDM\textsubscript{1} system (Figure 6-14), \( T_a \) is 0.5 milliseconds for both \( \gamma = 1.5 \) and \( \gamma = 2 \). \( T_r \) varies from 0.8 to 0.6 milliseconds as \( \gamma \) is varied from 1.5 to 2. Based on an average syllable duration of 7 milliseconds, the response is judged to be good and \( \gamma \) in the range 1.5 to 2 is appropriate.

For the CDM\textsubscript{2} system (Figure 6-15), \( T_a \) is 0.6 milliseconds for \( \gamma = 1.5 \) and \( \gamma = 2 \). \( T_r \), however, is radically different for the two values of \( \gamma \); for \( \gamma = 1.5 \), \( T_r \) is 0.7 milliseconds, and for \( \gamma = 2 \), \( T_r \) is 1.75 milliseconds. For both values of \( \gamma \) the recovery is much slower than observed for the CDM\textsubscript{1} system. From Figure 6-12 it is noted that \( B(L) \) for the CDM\textsubscript{2} system tends to level off at the value 0.2 for \( L = -7 \) dB. This means that for an underloaded condition, rather than \( b_a \) being zero essentially all the time and \( b_r \) being one most of the time, the ratio of \( E_b_a \) to \( E_b_r \) is on the order of \( 1/4 \) and recovery is slower than for the
TRANSIENT RESPONSE CDM₁ - THREE-STATE COMPANDING RULE, 4 SEGMENT LOG-LINEAR COMPANDOR. 12dB INPUT CHANGE.

fi=800Hz  fs=40kHz, τα=1.25 msec AVERAGE OF FIVE RESPONSES.

FIGURE 6-14
TRANSIENT RESPONSE CDM₂, THREE STATE COMPANDING RULE, 4 SEGMENT LOG-LINEAR COMPANDOR. 12dB INPUT CHANGE. \( f_1 = 800Hz, f_s = 40kHz, \tau_a = 1.25 \text{ msec} \) AVERAGE OF 5 RESPONSES

FIGURE 6-15
CDM₁ system which under the same conditions, the ratio of $E_{ba}$ to $E_{br}$ is much less than $1/4$. The best value of $\gamma$ for the CDM₂ system is clearly less than 2 and $\gamma = 1.5$ is near the optimum.

As a comparison, curves were also computed for the three-state rule using both a linear compandor (equation (6-7), $Q_o = -20$ dB) and a segmented compandor; and for two-state rule III using the segmented compandor. This comparison is illustrated in Figure 6-16. These curves are for a single response. First, comparing the three-state rule for the two different compandors, the performance is as expected from the previous discussion. Even though the attack time constants are the same, the response of the linear compandor is much slower because of the large value of $Q_o$ used.

Two-state rule III (Table 6-1) is essentially identical to the three-state rule with the exception of the third no-change state. From Figure 6-16 it is seen that the attack times for the two are essentially identical. However, in the steady state the ripple for the two-state rule is nearly double that of the three-state rule. As discussed previously, the phase relationship of the ripple and the input signal is such that a large ripple reduces the signal-to-noise ratio. Thus, the three-state rule offers the more desirable operation.

6.5 Channel Errors

Up to this point the effects of channel errors have not been considered. It is well known that for low error rates the effects on the LDM system are difficult to detect by a listener and cause little degradation. This is also true for CDM systems which use linear
TRANSIENT RESPONSE, $\text{CDM}_1 \cdot 12\text{dB}$ INPUT CHANGE. $\tau_a = 1.25\text{ msec}$, $f_i = 800\text{Hz}$, $f_s = 40\text{kHz}$, $\gamma = 2$

FIGURE 6-16
continuous compandors with time constants on the order of several milliseconds. A channel error causes a mistracking between the encoder and decoder companding circuits and results in a system gain or loss. For the continuous compandor, this mistracking is not noticed for two reasons; first, if the time constant is long, the tracking error is small, and secondly, the tracking error goes to zero after several time constants (integrator leak). With the discrete compandor there is no integrator leak and a tracking error remains until the system has returned to the silent state for a sufficient length of time for both counters to return to zero at which time system synchronization is reestablished. With no channel errors, the tracking between the encoder and decoder is exact.

Other than the obvious fact that a longer time constant reduces any tracking error it is not clear at this point what the magnitude of the tracking error is and what effect longer memories (shift register) used in some of the compandors has on the tracking error of a discrete compandor. As a start, consider a system which uses two-state rule I. For rule I, if two adjacent voice samples are identical, the step size is increased by a factor $r$. Otherwise, the step size is reduced by a factor $r$ (logarithmic compandor with $\gamma = 1$). Assume the encoder transmits message $---000---$ and an error occurs and the decoder receives $---010---$. Label the position of the error as position 1. The step size after pulse number 0 is $\Delta_0$. After pulse 2 the step size at the encoder $\Delta_{e2}$, is $r^2\Delta_0$. However, due to the error, the step size at the decoder after pulse 2, $\Delta_{d2}$,
is $r^{-2}A_0$. Since the system voltage gain is equal to the ratio $\Delta d/\Delta e$, the system has suffered a voltage reduction (loss) of $r^4$. Conversely, if the encoder transmits message $\cdots010\cdots$ and the encoder receives $\cdots000\cdots$ the system will suffer a voltage gain of $r^4$. If for example $r = \sqrt{2}$, the first case causes a 12 dB loss and the second case causes a 12 dB gain. The two are equally objectionable. A new parameter $T$, termed the tracking constant, is defined as

$$T = \frac{\max(\Delta e_2, \Delta d_2)}{\min(\Delta e_2, \Delta d_2)}, \quad (6-23)$$

The value $T = 1$ corresponds to perfect tracking and $T$ is greater than 1 if a tracking error occurs. At this point only two of the eight possible combinations of three binary digits have been considered. If it is assumed that each combination is equally likely, then it is possible to determine a mean squared value of $T$. A mean squared criterion is applied since power gain (or loss) is the appropriate measure. The eight binary combinations are enumerated in Table 6-5. The mean squared tracking value is

$$\langle T^2 \rangle = 0.5 (1 + r^8). \quad (6-24)$$

Now a similar calculation is made for the three-state rule discussed in the previous section. Again a logarithmic compandor is used. Also, all bit combinations are assumed equally likely. This problem is complicated both by the fact that the rule adjusts the step size based on three binary digits, thus five-bit signal patterns must be
TABLE 6-5
COMPUTATION OF $\langle T^2 \rangle$ FOR
TWO-STATE COMPANDING RULE I

<table>
<thead>
<tr>
<th>ENCODER MESSAGE</th>
<th>$\Delta e_2/\Delta_0$</th>
<th>DECODER MESSAGE</th>
<th>$\Delta q_2/\Delta_0$</th>
<th>$T^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>000</td>
<td>$r^2$</td>
<td>010</td>
<td>$r^{-2}$</td>
<td>$r^8$</td>
</tr>
<tr>
<td>001</td>
<td>1</td>
<td>011</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>010</td>
<td>$r^{-2}$</td>
<td>000</td>
<td>$r^2$</td>
<td>$r^8$</td>
</tr>
<tr>
<td>011</td>
<td>1</td>
<td>001</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>110</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>101</td>
<td>$r^{-2}$</td>
<td>111</td>
<td>$r^2$</td>
<td>$r^8$</td>
</tr>
<tr>
<td>110</td>
<td>1</td>
<td>100</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>111</td>
<td>$r^2$</td>
<td>101</td>
<td>$r^{-2}$</td>
<td>$r^8$</td>
</tr>
</tbody>
</table>
considered, and by the fact that $\gamma$ is not normally 1. For this calculation a value of $\gamma = 2$ is used. The counter is decremented on every other pulse on lead $b_r$ (refer to Figure 6-11). The system operation is unchanged if the ones and zeros are reversed on any combination. Thus, only 16 of the 32 five-bit binary combinations need be considered. The recovery problem is handled by assuming that if one pulse of the recovery signal $b_r$ occurs, it is counted with probability 1/2, and it is not counted with probability 1/2. If two pulses occur on $b_r$, the counter is reduced by 1 with probability 1. If three pulses occur on $b_r$, the counter is reduced by 1 with probability 1/2, and by 2 with probability 1/2. This computation is illustrated in Table 6-6, and

$$\langle m^2 \rangle = \frac{1}{32} (12 + 4r^2 + r^4 + 12r^6 + 3r^8). \quad (6-25)$$

Equations (6-24) and (6-25) are plotted in Figure 6-17 as a function of the number of cells in the counter. The number of cells $m$, the attack time constant $\tau_a$, and parameter $r$ are all related as indicated in Table 6-4. From Figure 6-17, as expected, the tracking error is reduced as the counter length is increased. The effect of a single error is more severe for the three-state rule than for the two-state rule. This is attributed to the fact that the three-state rule uses a shift register of length 3 and thus an error corrupts three companding pulses; whereas the two-state rule I uses a shift register of length 2, and only two companding pulses are corrupted by a single error. The Adaptive Delta Modulator (ADM) is a two-state
### Table 6-6

**Computation of $\langle T^2 \rangle$ for Three-State Companding Rule**

<table>
<thead>
<tr>
<th>Encoder Message</th>
<th>$\Delta e_3/\Delta_0$</th>
<th>Decoder Message</th>
<th>$\Delta d_3/\Delta_0$</th>
<th>$T^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>00000</td>
<td>$r^3$</td>
<td>00100</td>
<td>$r^{-1}$ or 1</td>
<td>$1/2(r^6+r^8)$</td>
</tr>
<tr>
<td>00001</td>
<td>$r^2$</td>
<td>00101</td>
<td>$r^{-1}$</td>
<td>$r^6$</td>
</tr>
<tr>
<td>00010</td>
<td>$r^{-1}$ or 1</td>
<td>00110</td>
<td>$1$</td>
<td>$1/2(1+r^2)$</td>
</tr>
<tr>
<td>00011</td>
<td>$r$</td>
<td>00111</td>
<td>$r$</td>
<td>1</td>
</tr>
<tr>
<td>00100</td>
<td>$r^{-1}$ or 1</td>
<td>00000</td>
<td>$r^3$</td>
<td>$1/2(r^6+r^8)$</td>
</tr>
<tr>
<td>00101</td>
<td>$r^{-1}$</td>
<td>00001</td>
<td>$r^2$</td>
<td>$r^6$</td>
</tr>
<tr>
<td>00110</td>
<td>$1$</td>
<td>00010</td>
<td>$1$ or $r$</td>
<td>$1/2(1+r^2)$</td>
</tr>
<tr>
<td>00111</td>
<td>$r$</td>
<td>00011</td>
<td>$r$</td>
<td>1</td>
</tr>
<tr>
<td>01000</td>
<td>$1$ or $r$</td>
<td>01100</td>
<td>$r^{-1}$</td>
<td>$1/2(1+r^2)$</td>
</tr>
<tr>
<td>01001</td>
<td>$r^{-1}$ or 1</td>
<td>01101</td>
<td>$r^{-1}$ or 1</td>
<td>1</td>
</tr>
<tr>
<td>01010</td>
<td>$r^{-2}$ or $r^{-1}$</td>
<td>01110</td>
<td>$r$</td>
<td>$1/2(r^6+r^8)$</td>
</tr>
<tr>
<td>01011</td>
<td>$r^{-1}$</td>
<td>01111</td>
<td>$r^2$</td>
<td>$r^6$</td>
</tr>
<tr>
<td>01100</td>
<td>$1$</td>
<td>01000</td>
<td>$1$ or $r$</td>
<td>$1/2(1+r^2)$</td>
</tr>
<tr>
<td>01101</td>
<td>$r^{-1}$ or 1</td>
<td>01001</td>
<td>$r^{-1}$ or 1</td>
<td>1</td>
</tr>
<tr>
<td>01110</td>
<td>$r$</td>
<td>01010</td>
<td>$r^{-2}$ or $r^{-3}$</td>
<td>$1/2(r^6+r^8)$</td>
</tr>
<tr>
<td>01111</td>
<td>$r^2$</td>
<td>01011</td>
<td>$r^{-1}$</td>
<td>$r^6$</td>
</tr>
</tbody>
</table>
FIGURE 6-17

EFFECTS OF AN ISOLATED CHANNEL ERROR ON ENCODER-DECODER TRACKING

MEASURE OF TRACKING ERROR

$10 \log <T^2>_{dB}$

NUMBER COUNTER CELLS $m$

ADM OPR POINT

THREE STATE RULE

TWO-STATE RULE

SUGGESTED OPERATING POINT
rule I system and operates with \( r \) on the order of 1.33 which corresponds to a 4 cell counter. This point is indicated on Figure 6-17 where it is noted that the average effect of a single isolated error is a 3 dB tracking error. The suggested operating point for the three-state system is for \( m = 6 \), for which a single error causes a 1 dB tracking error. The numerical values indicated in Figure 6-17 must be used with caution since only a single isolated error is considered, and it is not clear that the assumption that all bit patterns are equally likely is valid.

6.6 **Time Constant Tradeoff**

At this point it is clear that a tradeoff exists in the selection of the attack time constant \( \tau_a \). On the one hand, \( \tau_a \) should be very short to minimize performance deterioration during rapid changes in the input level. On the other hand, \( \tau_a \) should be long to maximize the steady state performance (reduce ripple on the companding control signal) and to minimize tracking errors. Further research using voice signals is necessary to establish the optimum value of \( \tau_a \). However, based on the international standard for syllabic compandors of an attack of five milliseconds or less, a value of \( \tau_a \) in the range of one to two milliseconds appears to be near the optimum value. This value is short enough to avoid objectionable transient performance and at the same time is long enough that tracking errors are relatively small, and by the use of a three-state companding rule one may realize a steady-state S/N near the maximum value achieved with the LDM system.
6.7 Comparison - CDM₁, CDM₂, and PCM

The signal-to-noise ratio characteristics of the log-PCM systems \( \mu = 100 \) for sinusoidal signals are well known.\(^{(35)}\) The \( \text{(S/N)} \) for a fixed number of levels increases at 3 dB per octave as the sampling frequency is increased. The \( \text{(S/N)} \) is also a function of the number of levels, increasing 6 dB for each level added. The sampling frequency must be slightly greater than twice the highest frequency in the input speech signal. In the following discussion two sampling frequencies are considered, 6.6 KHz and 8 KHz. A sampling frequency of 6.6 KHz is near the theoretical minimum for telephone speech. In practical systems a sampling frequency on the order of 8 KHz is used to permit the use of a less expensive filter in the decoder (i.e., when sampling at the Nyquist rate a very sharp cutoff filter is required to reject the alias signal). Companding for DM as it has been discussed here resulted in extending the peak \( \text{(S/N)} \) point over a broad range of inputs without lowering the maximum point appreciably. In log-PCM quite the contrary occurs; the curves have a broader range of a flat \( \text{(S/N)} \) but this is at the expense of a lower maximum \( \text{(S/N)} \) as was illustrated in Figure 5-1.

The average midinput range signal-to-noise ratios for the CDM₁, CDM₂, and PCM systems are compared as a function of the channel data rate in Figure 6-18. From the Figure it is clear that 7 bit log-PCM performance (requiring data rates of 46 to 56K bits) is equaled at data rates less than 40K bits for the CDM₂ system and at approximately 45K bits for the CDM₁ system.
MAXIMUM S/N AS A FUNCTION OF DATA RATE FOR CDM<sub>1</sub>, CDM<sub>2</sub> AND PCM

FIGURE 6-18
The suggestion that the companding used in CDM and the logarithmic companding used in PCM are equivalent is misleading. In PCM the levels are set and remain unchanged independent of the input. In the CDM system, the companding is done over the average signal level and cannot respond to rapid volume changes. The tradeoff of peak S/N for a broad lower level that is involved in PCM can be offset if a volume adjusting companding is used on PCM. From Figure 6-18, for sampling at a 6.6 KHz rate, 6 bit linear PCM has a maximum S/N approximately 5 dB greater than 7 bit log-PCM. Thus, it would appear that using 6 bit PCM with a combination of some level companding and some volume companding, the S/N of the CDM₂ system could be equaled or exceeded. Wilkinson (39) has reported good results with a system called adaptive PCM which uses these techniques.

6.8 Conversion of CDM to PCM

A communications network will quite probably use a variety of coding techniques. The capability of converting from one form to another directly, eliminating the filters required in the digital-to-analog-to-digital conversion is desirable. Today, the PCM system can be considered as standard. Thus, any new CDM system should lend itself to direct conversion to PCM. Some work has been done in this area including that described in References 17 and 23. In terms of the CDM₁ and CDM₂ systems described previously, discrete compandors have been suggested as a logical way to generate the desired logarithmic characteristic. This choice is consistent with digital conversion techniques. The irregular step size caused by the second integrator of the CDM₂
system does not however lend itself well to such an arrangement. With reference to Figure 6-18, at a sampling frequency of 40 KHz, the CDM\textsubscript{2} system exhibits a 3 dB greater (S/N) than the CDM\textsubscript{1} system. It is doubtful that this 3 dB improvement is sufficient to offset the added expense necessary to convert the signal to PCM. For higher sampling frequencies, the advantage of using the second integrator increases substantially and the difficulty introduced by the second integrator is a good tradeoff for the far superior performance.

6.9 **Optimum System**

Based on the preceding investigation, the system which most closely meets the performance objectives set forth in Section 6.1 is one which uses: 1) a single integration delta modulator, 2) a discrete log-linear compandor, and 3) a three-state companding rule. This system provides a steady state signal-to-noise ratio which is essentially the same as 7 bit log-PCM at a sampling frequency of approximately 40 KHz which is a 28 percent bandwidth reduction over the PCM system. In addition, this system is less expensive than the 7 bit log-PCM system.

\[\text{For the CDM}_1 \text{ system, the step size can be assumed to be a function only of the companding coefficient } q. \text{ The converter contains the companding circuitry and generates } q_i \text{ for every sample. The step size for the } i^{th} \text{ sample is } \Delta_i = q_i \Delta_{\text{min}}. \text{ The value of the signal estimate } \hat{x}_i \text{ is } \hat{x}_{i-1} + \Delta_i \text{sgn}(a_{i-1} - 0.5), \text{ where } a_i \text{ is the binary coded speech signal. The value } \hat{x}_i \text{ is stored in a register which has an output that is coded in the PCM code. This register is sampled at the PCM sampling rate. For the CDM system, } \Delta_i \text{ is not only a function of } q_i \text{ but also a function of the past pulse pattern; thus, to generate } \Delta_i \text{ one must keep a record of the past pulse configuration and devise a way to combine it with } q_i \text{ to generate } \Delta_i. \text{ It is this complication that is referred to above.} \]
CHAPTER 7

SUMMARY

The performance of LDM and CDM systems to sinusoidal signals together with the characteristics of various digital companding rules have been presented. The results have been compared, where appropriate, with similar systems reported in the literature and with PCM. The emphasis has been on systems suitable for the transmission of high quality telephone speech.

The results presented can be grouped into two categories. First, the simulation and measured data together with the calculations for the linear delta modulation system can be summarized as follows:

1) The value of the modified mean squared error computed by taking an ensemble average very closely matches measured values. The computed mms error in the granular noise region is essentially a function of the step size only.

2) The signal-to-noise ratio and error spectral density computed by taking an ensemble average of waveforms generated by computer simulation agree with measured values.

3) The error spectral density for the LDM system with a sinusoidal input signal has a large peak which, for
small input signals is outside the audio band. This
is the granular noise region and the level of the
spectral density in the audio band is essentially
independent of the input power and frequency. The
location of the peak in the spectral density is a
function of the input loading; for small values of
loading the peak is near one-half the sampling fre-
quency, and as the loading is increased the peak
moves toward the audio band, moving into the audio
band in the slope-overload noise region.

4) The signal-to-noise ratio of the single integration
delta modulator can be accurately approximated by
two line segments. In the granular noise region
the \((S/N)\) is an increasing function of the signal
level, increasing one dB for each dB increase in
the signal level. In the slope-overload noise
region the \((S/N)\) is a decreasing function of the
signal level, decreasing two dB for each dB increase
in the input.

5) The signal-to-noise ratio of the double integration
delta modulator can be accurately approximated by
three line segments. In the granular noise region
the \((S/N)\) increases one dB for each dB increase in
the input. In the slope-overload noise region, the
\((S/N)\) decreases two dB for dB increase in the input.
Between the granular and the slope-overload noise regions the (S/N) is constant, independent of the signal level.

6) The signal-to-noise ratio for an 800 Hz tone signal measured using a technique which discounts phase distortion accurately describes the performance of the system for voice signals.

Secondly, the results of the investigation of companded delta modulation can be summarized as follows:

1) Nearly 40 dB of companding is required for companded delta modulators using a sampling frequency of 40 KHz to equal the performance of 7 bit log-PCM.

2) A ratio of attack to recovery time constant on the order of 1.5 to 2 provides symmetric operational around the maximum signal-to-noise ratio point.

3) Linear compandors are unsuitable for systems which require on the order of 40 dB of companding.

4) The ideal companding characteristic is logarithmic.

5) A new combined linear-logarithmic compandor is suggested which combines the economy of the linear compandor with the operating characteristics of the logarithmic compandor.

6) Discrete compandors which use a counter to control the step size are very sensitive to channel errors which result in a tracking error between the encoder
and decoder. The severity of the tracking error increases substantially if the attack and recovery time constants are reduced to less than one millisecond.

7) Companding rules which control the step size based on more than four speech samples are unsuitable due to a severe sensitivity to variations of the input frequency.

8) At least four two-state companding rules exist which permit operation near the peak on the signal-to-noise characteristic for both the single and double integration companded delta modulators. Based on the 800 Hz tone performance, essentially identical performance can be expected for each of the four rules and no clear optimum exists.

9) A new three-state companding rule has been proposed which appears to perform superior to all the two-state rules.

10) Direct conversion between the delta modulation format and the PCM format is more difficult for the double integration than for the single integration system.

11) For operation at a 40K bit data rate, the single integration companded delta modulation system is judged to be the best of the alternatives considered.
The best companding arrangement consists of a segmented log-linear compandor which adjusts the step size based on a three-state companding rule. A discrete compandor with 64 step sizes and an attack time constant of 1.25 milliseconds and a recovery time constant of 1.9 milliseconds provides rapid adjustment to input level changes while at the same time minimizes the effects of channel errors on system tracking.

While this analysis which used sinusoidal test signals provides a convenient method of comparing various systems, it does not replace listening tests using speech signals. Thus, the conclusions reached here are subject to this constraint.
APPENDIX A

SIMULATION OF LDM

A.1 Basic System

The LDM encoder is illustrated in Figure A-1. The equation which describes this system is

\[ \hat{X}_i = \hat{X}_{i-1} + A_i \text{sgn}(e_{i-1}), \quad i = -2,-1,0,+1, \ldots, \quad (A-1) \]

where

\[ e_i = \hat{X}_i - \hat{X}_i, \quad (A-2) \]

and

\[ A_i = f(e_i, e_{i-1}, \ldots). \quad (A-3) \]

The subscript \( i \) refers to the \( i \)th sample. Equation (A-1) is nonlinear and a closed form solution is difficult to obtain analytically. This equation is implemented on a computer which is used to iterate (A-1) many times and from this data an average or steady state value of the error autocorrelation function (acf) is determined. This is then Fourier transformed to determine the error spectral density and the audio noise power.

The system has been simulated in addition to being constructed. This apparent duplication of effort is not without justification. A simulation permits one to obtain data which is difficult to measure.
MATHEMATICAL MODEL OF LDM ENCODER

Figure A-1
(For example, the acf at the sample points.) Also, a simulation can be slowed down and the system operation observed in a degree of detail difficult to achieve in a working system. Another advantage of the simulator is the ability to implement system modifications rapidly and to optimize prior to modifying the experimental circuit, thus minimizing laboratory effort. This is apparent when a number of companding rules are simulated as described in Appendix B.

The problem that is solved here can at most be termed simple in comparison to the large system statistical simulations which tax even the highest speed data processing systems. This simplicity enables the use of time share operation, specifically, CPS on an IBM 370 computer. The interaction capabilities of the time share operation presents the additional advantage of enabling one to continuously observe system performance while altering system parameters.

A.2 System Initial Conditions

The input signal is a sinusoidal wave of known amplitude and frequency, i.e.,

$$X_i = \sqrt{2} X \sin\left(2\pi \frac{f_i}{f_s} i + \theta\right),$$  \hspace{1cm} (A-4)

where $\theta$ is the phase of the input relative to the 0th sample. $\theta$ is a random variable which is assumed to be uniformly distributed between 0 and $2\pi$.

A second random variable is the error at the starting point, $e_0$. The distribution of $e_0$ is known only to be zero mean. The system is started with $e_0 = 0$.\)
Since the integrators used in the system are not ideal, the step size \( \Delta_i \) varies as a function of the past pulse pattern. To estimate the response of the system to the actual pulse pattern, the following procedure is used. The present sample is labeled \( i \), the most recent past sample \( i-1 \), that preceding it \( i-2 \), \(--\), etc. A record, termed \( k_j \), is kept of the position of the last five reversals of the binary signal \( a_i \). For the \( i \)th sample this record is:

\[
k_{j_i} = \begin{cases} 
k_{j_i-1} + 1, & \text{if } a_i = a_{i-1}, \\
1, & \text{if } a_i \neq a_{i-1},
\end{cases}
\]  
(A-5)

and

\[
k_{j_i} = \begin{cases} 
k_{j_i-1} + 1, & \text{if } a_i = a_{i-1}, \\
k_{j-1_{i-1}} + 1, & \text{if } a_i \neq a_{i-1},
\end{cases}
\]  
(A-6)

\[j = 2,3,4,5.\]

The response of the feedback network \( H(f) \) to a unit step of duration \( kr \) is defined as \( U(k) \), \( k = 1,2,3, \ldots \), etc., where \( r \) is the sampling period. Using linear sums of these step responses, one can approximate the change in the signal estimate between the \( i \)th and \((i+1)\)th sample as

\[
\Delta_i = 2U(k_{1_i}) - 2U(k_{2_i}) + 2U(k_{3_i}) - 2U(k_{4_i}) + U(k_{5_i}).
\]  
(A-7)

The variables \( e \) and \( \phi \) are known to be correlated because there is a delay between \( X(t) \) and \( \hat{X}(t) \). Another way of putting this is that the error spectral density contains a line at the input frequency. As was pointed out in Chapter III, this delay is not considered an error
when the noise is computed. To differentiate between the system
error and the error that is used to compute the noise a modified
error is used. This modified error is defined as

\[ e_1^* = e_1 - K \cos(2\pi \frac{f_1}{f_S} i + \phi). \]  (A-8)

The phase adjustment is achieved by changing K, i.e., $K \cos(2\pi \frac{f_1}{f_S} i + \phi)$
is in quadrature with the input $\sqrt{2} X \sin(2\pi \frac{f_1}{f_S} i + \phi)$.

The following procedure is used to determine the distribution
of $e_1^*$. The system is started with zero error and some fixed phase
$\phi_j$, and allowed to iterate equation (A-1) k times. This procedure is
repeated for r values of $\phi_j$; where

\[ \phi_j = (j/r)2\pi, \quad j = 0,1,2,\ldots, \quad r-1 \]  (A-9)

The value of $(e_k^*)^2$ is averaged for the r values of $\phi_j$ and this average
value is taken to be $\langle e^{*2} \rangle$, the mms error. The symbol $\langle \rangle$ refers to an
ensemble average. This procedure is repeated for various values of K
until a minimum for the mms error is found. The values of K which min-
imized $\langle e^{*2} \rangle$ ranged from zero for small input signals to a value cor-
responding a delay of approximately one sample period for large input
signals. Various values of k and r were tried, and the results
stabilized for k equal to greater than 25 and r equal to or greater
than 100. For the simulation results reported in Chapter IV, values
of k = 50 and r = 200 were used. The values of $e_k^*$ are taken to be
representative of the steady-state error and are used for the system
initial condition.
A.3 **Discrete Fourier Transform**

The discrete Fourier Transform procedure involves converting discrete samples of the autocorrelation function \( \Re e^* \) into discrete values of the spectrum \( S_e^* \). The sampling interval of the time function, defined as \( dt \), is determined by the bandwidth of the spectrum. That is, if the bandwidth of the spectrum is \( f \), one must sample at a rate equal to or greater than \( 2f \), the Nyquist rate. From measurements of the error spectral density of the LDM system it has been determined that the bandwidth of the spectrum is greater than \( f_b/2 \) and less than \( f_s \), where \( f_s \) is the LDM sampling frequency. Thus, \( dt \) must be greater than \( 1/f_s \), and a value of \( 2/f_s \) is larger than necessary. Since samples at intervals \( 1/f_s \) are necessary in the generation of \( X_i \), it was decided to use these samples and to generate one more midway between each of the samples. Thus, \( dt = \tau/2 \), where \( \tau \) is the DM sampling period.

The number of time samples \( M \) determines the resulting number of frequency samples and the resulting spacing of the frequency samples \( df \). This relationship is

\[
df = \frac{1}{M dt}.
\]

(A-10)

Since the spectrum is even, real and positive, only the cosine term of the transform is needed, and this need be computed only for

---

\(^6\)Hamilton\(^{21}\) gives an excellent description of the discrete Fourier Transform from which this summary was taken.
positive values of time and frequency. The value of the spectrum at each point is

\[ S_e^*(m \text{df}) = dt(Re(0) + 2 \sum_{n=1}^{M-1} Re^*(n(dt)) \cos \left( \frac{mn}{M} \right)), \]

\[ m = 0, 1, \ldots, M-1. \quad (A-11) \]

As is well known\(^5\), the result of this transform is not a good estimate of the spectrum. It is not consistent; that is, its variance does not go to zero as \( M \) increases. To get a good estimate of \( S_e^*(f) \) one must smooth the discrete sample values, that is, take a weighted sum of adjacent values. To accomplish this, a rectangular\(^8\) window has been used. The smoothing is expressed as

\[ S_m^*(m \text{df}) = \frac{1}{(2j+1)} \sum_{k=-j}^{k=j} S_e^*((m+k) \text{df}). \quad (A-12) \]

For this work a value of \( j \) equal 7 was used.

A.4 Generation of Autocorrelation Function Samples

As pointed out above, to obtain an accurate estimate of the error spectral density the values of the acf \( Re_1^* \) must be determined for \( i = 0, 1/2, 1, \ldots, \) etc., which corresponds to two samples for each sample of the LDM system. The scheme here is to generate a number of spectra representative of the ensemble of acfs which results from the random nature of the initial conditions for equations (A-1) and (A-2). These are then averaged to obtain the overall system average behavior. Since the transform is a linear operation, an equivalent procedure is to average the acf samples and then take the transform of this single average acf. This second technique has been used.
The technique of obtaining the acf is similar to that used to obtain the mms error. A phase \( \phi_j \) is picked and the system is allowed to start with zero error and then iterate for \( k \) sample times to reach a steady state condition. The acf samples are then computed using the following equations.

\[
R_{0*} = (e_k^*)^2, \quad (A-13)
\]
\[
R_{n*} = (e_k^*) (e_{k+n}^*), \quad n = 1, 2, \ldots, \quad (A-14)
\]
\[
R_{n-1/2}^* = (e_k^*) (e_{k+n-1/2}^*), \quad n = 1, 2, \ldots. \quad (A-15)
\]

This procedure is repeated for \( r \) values of \( \phi \). The individual samples are averaged using the expression

\[
R_{e1}^* = \frac{1}{r} \sum_{j=0}^{r-1} R_{e1}^* (\phi_j), \quad (A-16)
\]

and this average is transformed to obtain \( \langle E_0^* (f) \rangle \). The use of the simple arithmetic average in equation (A-16) is justified on the basis that \( \phi \) is assumed to be uniformly distributed between 0 and \( 2\pi \). The symbol \( \langle \rangle \) indicates that the spectrum was obtained from an ensemble average. Where these results are presented in Chapter IV, the asterisk is dropped and it is understood that the phase correction has been applied to the error.
APPENDIX B

COMPUTATION OF $\mathbf{B}(L)$ AND TRANSIENT RESPONSE

The system equations, repeated from Appendix A are:

$$\hat{X}_i = \hat{X}_{i-1} + \Delta_i \text{sgn}(X_{i-1} - \hat{X}_{i-1}), \quad i = \ldots, -2, -1, 0, 1, \ldots,$$

(B-1)

where

$$X_i = \sqrt{2} X \sin(2\pi f_i t / f_s),$$

(B-2)

$$\Delta_i = f(k_{1i}, k_{2i}, \ldots, k_{5i}),$$

(B-3)

$$k_{1i} = \begin{cases} k_{i-1} + 1, & \text{if } a_i = a_{i-1}, \\ 1, & \text{otherwise}, \end{cases}$$

(B-4)

and

$$k_{ji} = \begin{cases} k_{j-i-1} + 1, & \text{if } a_i = a_{i-1}, \\ k_{j-1-i-1} + 1, & \text{otherwise}, \end{cases}$$

(B-5)

$j = 2, 3, 4, 5.$

B.1 Computation of $\mathbf{B}(L)$ for Two State Companding Rules

The function $\mathbf{B}(L)$ for the two state companding rule was defined (equation (6-3)) as

$$\mathbf{B}(L) = \lim_{{n \to \infty}} \frac{1}{2n} \sum_{{i=0}}^{n} (b_i + 1), \quad b_i = +1 \text{ or } -1.$$

(B-6)
The variable $b_i$ is related to the variable $k_{1i}$ and, in some cases $k_{2i}$, depending on the specific rule. Consider for example Rule III (Table 6-1). The value of $b_i$ is set to $+1$ when the pattern $a_{i-1}, a_{i-2}$ is 111 or 000, and held at this value until the pattern 101 or 010 occurs at which time $b_i$ is set to $-1$, and held at $-1$ until the pattern to set $b_i$ back to $+1$ occurs. This can be written in terms of the variables $k_1$ and $k_2$ as

$$b_i = \begin{cases} 
-1, & \text{if } k_{1 \ell} + 1 = k_{2 \ell}, \text{ and } k_{1j} = 2, \text{ for all } \ell \leq j \leq i, \\
+1, & \text{otherwise.}
\end{cases}$$

Similar equations for the five two-state rules are tabulated in Table B-1.

The function $\tilde{B}(L)$ is computed for a fixed $I$ by setting $X$ (equation (B-2)) to the appropriate value corresponding to $L$ (equation (3-12)) and then iterating equations (B-1) through (B-5). This is illustrated in flowchart form in Figure B-1. The initial conditions are set, and the system iterates $m$ times ($i = -m + 1$ to $i = 0$) to allow $(X_i - \hat{X}_i)$ to reach a steady state condition. Then the value of

$$\sum_{i=0}^{n} (b_i - 1)$$

is accumulated for $n$ iterations after which $\tilde{B}(L)$ is computed using equation (B-6). To produce the smooth curves of $\tilde{B}(L)$ in Figures 6-2 and 6-3 values of $m = 50$, and $n = 1000$ were used. The function stabilizes for $n$ of approximately 200. For $n = 1000$ the averaging time corresponds to 20 cycles of the 800 Hz tone input ($f_0 = 40$ KHz).
TABLE B-1

EQUATIONS OF $b_1(k_{1_i}, k_{2_i})$ FOR TWO-STATE COMPANDING RULES

<table>
<thead>
<tr>
<th>RULE</th>
<th>$b_1(k_{1_i}, k_{2_i})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>$b_1 = \begin{cases} +1, &amp; k_{1_i} &gt; 1, \ -1, &amp; k_{1_i} = 1. \end{cases}$</td>
</tr>
<tr>
<td>II</td>
<td>$b_1 = \begin{cases} +1, &amp; k_{1_i} &gt; 2, \ -1, &amp; k_{1_i} \leq 2. \end{cases}$</td>
</tr>
<tr>
<td>III</td>
<td>$b_1 = \begin{cases} -1, &amp; \text{if } k_{1_i} + 1 = k_{2_i} \text{ and } k_{1_j} \leq 2, \text{ for all } 2 \leq j \leq i, \ +1, &amp; \text{otherwise}. \end{cases}$</td>
</tr>
<tr>
<td>IV</td>
<td>$b_1 = \begin{cases} +1, &amp; k_{1_i} &gt; 3, \ -1, &amp; k_{1_i} \leq 3. \end{cases}$</td>
</tr>
<tr>
<td>V</td>
<td>$b_1 = \begin{cases} -1, &amp; \text{if } k_{1_i} + 1 = k_{2_i} \text{ and } k_{1_j} \leq 3, \text{ for all } 2 \leq j \leq i, \ +1, &amp; \text{otherwise}. \end{cases}$</td>
</tr>
</tbody>
</table>
START

\( l = m + 1, \lambda_m - x_m = 0 \)

\( k_1 = 1, k_2 = 2, k_3 = 3, k_4 = 4, k_5 = 5 \)

\(?\)

\( l = n \)

NO

COMPUTE \( X_1 = 1 \)

EQN (B-2)

COMPUTE \( \Delta_1 \)

EQN (B-3)

COMPUTE \( \lambda_1 \)

EQN (B-4), (B-5)

? \( i \geq 0 \)

NO

TWO

STATE

RULES

COMPUTE \( b_l \)

THREE

STATE

RULES

COMPUTE \( b_{a_l}, b_{r_l} \)

? \( i \geq 0 \)

YES

COMPUTE \( \sum_{j=0}^{l-1} b_j \)

COMPUTE \( \sum_{j=0}^{l-1} b_{r_j} \)

COMPUTE \( \sum_{j=0}^{l-1} b_{a_j} \)

END

FLOW CHART- COMPUTATION OF \( \tilde{B}(L) \)

FIGURE B-1
B.2 Computation of $B(L)$ For Three State Rule

The function $B(L)$ for the three-state rule (equation (6-21)) is

$$B(L) = \lim_{n \to \infty} \frac{\sum_{i=0}^{n} b_{ai}}{n} = \frac{\sum_{i=0}^{n} b_{ai} + \sum_{i=0}^{n} b_{ri}}{n}.$$  \tag{B-8}

Variable $b_{ai}$ is 1 only if the three most recent binary coded samples ($a_i, a_{i-1}, a_{i-2}$) are identical, thus

$$b_{ai} = \begin{cases} 1, & \text{if } k_{1i} > 2, \\ 0, & \text{otherwise}. \end{cases} \tag{B-9}$$

Similarly, $b_{ri}$ is 1 only if the three most recent samples are of the pattern 101 or 010, thus

$$b_{ri} = \begin{cases} 1, & \text{if } k_{1i} = 1, \text{ and } k_{2i} = 2, \\ 0, & \text{otherwise}. \end{cases} \tag{B-10}$$

The function $B(L)$ is computed in exactly the same manner as for the two-state rules (Figure B-1) with the exception of the different expression for $B(L)$.

B.3 Transient Response

The system transient response is computed in a manner very similar to that used to compute $B(L)$. First, consider the response to an increase in the input signal. The system is assumed to be in steady state prior to the signal increase, with the value of $Q$ at $Q_0$ and the value of $q$ at $q_0$. The system is subjected to a 12 dB increase in input
at sample \( i = 0 \), thus \( Q = 4Q_0 \) for \( i > 0 \). The function of interest is \( q_i \). The new steady state value of \( q \) is \( 4q_0 \). Function \( q_i \) is dependent on the specific type of compandor used, that is, linear, logarithmic or log-linear. Only discrete compandors are considered with the attack time constant fixed at 1.25 milliseconds. The variable \( u_i \) is defined as the number stored in the binary counter of the discrete compandor on the \( i \)th sample. The relation of \( q \) to \( u \) is tabulated for each of the three types of compandors in Table B-2. The value of \( q_0 \) is obtained from equation (5-11), that is

\[
q_0 = q_0 q_{\text{max}} 
\]

(B-11)

The value of \( u_0 \) is computed from the equations in Table B-2 using \( q = q_0 \). The steady state loading \( L_c(\gamma) \) prior to the input increase is a function of the particular companding rule and is determined from Figure 6-4, 6-5, or 6-13. From these data, the rms value of the input prior to the change, \( X_0 \), is computed using equation (5-1),

\[
X_0 = \frac{L_c(\gamma) f_s q_0 A_{\text{min}}}{\sqrt{2} 2\pi f_1} 
\]

(B-12)

The value of the input after the increase is \( 4X_0 \) (12 dB increase).

The function \( q_i/q_0 \) is computed in essentially the same manner as indicated in the flowchart illustrated in Figure B-1 with the exceptions that \( m = 0 \), (the system starts with zero error \( \hat{X}(0) = X(0) \)), and the step size is

\[
A_i = q_i f(k_{1i}, k_{2i}, \ldots, k_{5i}). \quad \text{(B-13)}
\]
TABLE B-2

FUNCTIONS $q_i(u_i)$, $\tau_a = 1.25$ MILLISECONDS

**Linear Compandor**

$$q_i = 1 + 0.735 u_i, \quad u_i = 0,1,2,...$$

**Logarithmic Compandor**

$$q_i = (1.0745)^{u_i}, \quad u_i = 0,1,2,...$$

**4 Segment Log-Linear Compandor**

$$q_i = \begin{cases} 
1 + 0.413 u_i, & 0 \leq u_i \leq 15, \\
7.6 + 0.826(u_i - 16), & 16 \leq u_i \leq 31, \\
20.8 + 1.65(u_i - 32), & 32 \leq u_i \leq 47, \\
47.2 + 3.30(u_i - 48), & 48 \leq u_i \leq 63.
\end{cases}$$
The number stored in the counter, $u_i$, is

(two-state rule)

$$u_i = \begin{cases} u_{i-1} + b_i, & b_i = 1 \\ u_{i-1} + b_i/\gamma, & b_i = -1 \end{cases} \quad (B-14)$$

(three-state rule)

$$u_i = u_{i-1} + b_{a_i} - b_{r_i}/\gamma \quad (B-15)$$

When $q_i$ is computed (equations Table B-2) $u_i$ is truncated and only the integer value is used. This corresponds to the circuit action (Figure 6-11) where fractions are stored in the divide circuit.

The function $q_i$ contains a ripple component at twice the input frequency (refer to Figure 6-16). The smooth curves illustrated in Figures 6-14 and 6-15 are obtained by averaging $q_i$ for the five different input phases $\phi = 0, 0.2\pi, 0.4\pi, 0.6\pi, \text{and } 0.8\pi$.

The response to a 12 dB decrease in input (recovery) is computed in exactly the same manner as the attack with the exception that the steady state operating point prior to the input change is $4Q_0$, $4q_0$, and $4X_0$. At $i = 0$ the input is reduced to $X_0$, and the system responds by changing $q$ from $4q_0$ to $q_0$. 

APPENDIX C

CALCULATION OF THE mms ERROR

The objective here is to determine the mms error by taking an ensemble average. Consider the ensemble of input sinusoidal waveforms $X_j(t)$ placed one above the other as illustrated in Figure C-1. The vertical axis drawn through these waveforms is labeled the zero time axis. The average is taken along this axis. The process is assumed to have started a long time prior to the zero axis, and each waveform $X_j(t), X_{j+1}(t), \ldots, X_{j+n}(t)$, has the same frequency $f_i$, and the same rms amplitude $X$. On each waveform the sample points are marked with a small circle. The distance between the most recent sample (labeled the 0th sample) and the zero axis is defined as $y_j$ (seconds). (See magnified waveform Figure C-1.) The phase of the input relative to the 0th sample point is $\phi_j$ (radians). Both $\phi_1$ and $y_j$ are treated as random variables. Thus, the jth input is

$$X_j(t) = \sqrt{2} X \sin(2\pi f_i(t+y_j)+\phi_j). \quad (C-1)$$

The difference between the input $X_j(t)$ and the signal estimate $\hat{X}_j(t)$ at the kth sample point $z_{jk}$ is also a random variable and is defined as

$$z_{jk} = X_j(-y_j+k\tau) - \hat{X}_j(-y_j+k\tau), \quad k = 0,1,\ldots, \quad (C-2)$$

where $\tau$ is the sampling period.

165
WAVEFORM ENSEMBLE

FIGURE C-1
The signal estimate \( \hat{X}_j(t) \) can be written in terms of 
\( \hat{X}_j(-y_j+k\tau) \), the signal estimate at the \( k \)th sample, and \( z_{jk} \) as

\[
\hat{X}_j(t) = X_j(-y_j+k\tau) + \frac{\Delta(y_j+t-k\tau)}{\tau} \text{sgn}(z_{jk}), \quad -y_j+k\tau \leq t < -y_j+(k+1)\tau. \tag{C-3}
\]

The error is

\[
e_j(t) = X_j(t) - \hat{X}_j(t). \tag{C-4}
\]

From (C-2), (C-3) and (C-4),

\[
e_j(t) = X_j(t) - X_j(-y_j+k\tau) + z_{jk} - \frac{\Delta(y_j+t-k\tau)}{\tau} \text{sgn}(z_{jk}),
\]

\[-y_j+k\tau \leq t < -y_j+(k+1)\tau. \tag{C-5}
\]

The term \( X_j(t) - X_j(-y_j+k\tau) \) in equation (C-5) can be written as

\[
X_j(t) - X_j(-y_j+k\tau) = \sqrt{2} X [\sin(2\pi f_1(t+y_j) + \phi_j) - \sin(2\pi f_1 k\tau + \phi_j)],
\]

\[-y_j+k\tau \leq t < -y_j+(k+1)\tau. \tag{C-6}
\]

It is noted that the function contains a component at frequency \( f_1 \) and components at \( f_S \) and harmonics of \( f_S \). This function is illustrated in Figure C-2. The definition of the mms error excludes all average power at the input frequency \( f_1 \), and that component of \( X_j(t) - X_j(-y_j+k\tau) \) at \( f_1 \) must be subtracted out to compute the mms error. The remaining part of (C-6) after all average power at \( f_1 \) is subtracted out has components at frequencies equal to or greater than \( f_S \) which is above the frequency where the amplitude of the error spectral density is negligible (refer
\[ \sqrt{2X} \]

\[ X(t) \quad X(k\tau) \]

**TIME**

**FUNCTION** \( X(t) - X(k\tau) \)

**FIGURE C-2**
to Section A.3, Appendix A) and is assumed to be zero in the calculation of the mms error. The modified error $e_j^*(t)$ is defined as $e_j(t) - [X_j(t) - X_j(-y_j + k\tau)]$, and from (C-5)

\[ e_j^*(t) = z_{j_0} - \frac{\Delta(y_j + t)}{\tau} \text{sgn}(z_{j_0}). \]  

(C-7)

The subscript $j$ is now dropped. The squared error at the zero axis is

\[ e_{\#}^2 = z_0^2 - \frac{2\Delta y|z_0|}{\tau} + \left(\frac{\Delta y}{\tau}\right)^2. \]  

(C-8)

Note that the phase variable does not appear in the expression for $e_{\#}^2$.

The random variable $y$ is taken to be independent of $z_0$ and is uniformly distributed between 0 and $\tau$. The distribution of $z_0$ is unknown and expressed as $p(z_0)$. Thus, the joint distribution is

\[ p(y,z_0) = \frac{y}{\tau} p(z_0), \quad 0 \leq y < \tau, \quad -\infty < z_0 < \infty, \]  

(C-9)

and the expected value of the squared error is

\[ \langle e_{\#}^2 \rangle = \int_0^\tau \int_{-\infty}^\infty e_{\#}^2 p(y,z_0) \, dy \, dz_0. \]  

(C-10)

The result of this integration is

\[ \langle e_{\#}^2 \rangle = \langle z_0^2 \rangle - 2\Delta \langle |z_0| \rangle + \frac{\Delta^2}{3}. \]  

(C-11)

Thus, the mms error at the zero axis is a function of only the step size and the error at the sampling points.
Now, consider the expression for the acf at the sample points

\[ R_{e}^{*}(k) = \langle e^{*}(y), e^{*}(-y+\kappa) \rangle, \quad k = 0, 1, \ldots \]  
(C-12)

The value of \( R_{e}^{*}(k) \) for \( k = 1 \) is

\[ R_{e}^{*}(1) = \int_{0}^{T} \int_{-\infty}^{\infty} z_{0}[z_{0} - \Delta \text{sgn}(z_{0})] p(y, z_{0}) \, dy \, dz_{0}. \]  
(C-13)

The result of this integration is

\[ R_{e}^{*}(1) = \langle z_{0}^{2} \rangle - \Delta \langle |z_{0}| \rangle. \]  
(C-14)

This is exactly the term that appears in equation (C-11), thus

\[ \langle e^{*2} \rangle = \frac{\Delta^{2}}{3} \left( 1 + \frac{3}{\Delta^{2}} R_{e}^{*}(1) \right). \]  
(C-15)

In Section 4.3 it was argued that the term \( \frac{3}{\Delta^{2}} R_{e}^{*}(1) \) is small with respect to one and \( \langle e^{*2} \rangle \) can be expressed as

\[ \langle e^{*2} \rangle \approx \frac{\Delta^{2}}{3}. \]  
(C-16)

Where these results are used in Chapter IV the asterisk is dropped and it is understood that all average power at the input frequency has been subtracted out.
REFERENCES


