NONLINEAR ANALOG COMPUTATION AND MULTIFUNCTION GENERATION WITH NONLINEAR PULSE WIDTH MODULATION

JOANNES NICOLAAS MATTHEUS DE JONG
University of New Hampshire, Durham

Follow this and additional works at: https://scholars.unh.edu/dissertation

Recommended Citation
https://scholars.unh.edu/dissertation/2371

This Dissertation is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact nicole.hentz@unh.edu.
NONLINEAR ANALOG COMPUTATION AND MULTIFUNCTION GENERATION WITH NONLINEAR PULSE WIDTH MODULATION

Keywords
Engineering, Electronics and Electrical
DE JONG, Joannes Nicolaas Mattheus, 1948-
NONLINEAR ANALOG COMPUTATION AND
MULTIFUNCTION GENERATION WITH
NONLINEAR PULSE WIDTH MODULATION.

University of New Hampshire, Ph.D., 1976
Engineering, electronics and electrical

Xerox University Microfilms, Ann Arbor, Michigan 48106
NONLINEAR ANALOG COMPUTATION AND MULTIFUNCTION
GENERATION WITH NONLINEAR PULSE
WIDTH MODULATION

by

JOANNES NICOLAAS MATTHEUS DE JONG

Werktuigbouwkundig Ingenieur
Delft, University of Technology, The Netherlands
1972

A DISSERTATION
Submitted to the University of New Hampshire
In Partial Fulfillment of
The Requirements for the Degree of
Doctor of Philosophy
Graduate School
Engineering Ph.D. Program
System Design Area
May 1976
This dissertation has been examined and approved.

C. I. Taft
Dissertation Director, Charles K. Taft
Professor of Mechanical Engineering

Robert W. Corell
Professor of Mechanical Engineering

David E. Limbert
Associate Professor of Mechanical Engineering

John L. Pokoski
Associate Professor of Electrical Engineering

Linda G. Sprague
Associate Professor of Business Administration

May 7, 1976
Date
ACKNOWLEDGEMENTS

I would like to take this opportunity to thank Dr. Charles K. Taft who, both as an advisor and as a person, has helped to make my sojourn at the University of New Hampshire so successful. I would also like to express my appreciation of Dr. Charles K. Taft, Dr. David E. Limbert, Dr. Robert W. Corell and Dr. Linda G. Sprague whose encouragement and helpful discussions were indispensable.

Many thanks are also deserving Mrs. Donna Kavanagh, who has taken my pages of writing and magically typed them into a presentable dissertation.

My appreciation is also offered to Corning Glass Works for providing fluidic devices used in the research.

Finally, as this is the last instance that I may have to do so, I would like to thank the members of my research group, and others I have had contact with, who have helped to make my sojourn here successful.
PREFACE

This dissertation deals with a novel concept in nonlinear analog computation which was implemented with electrical as well as fluidic components. Moreover, consideration is given to the problems involving the "Technology Transfer" of this new concept. In order to improve the readability of this dissertation, it is divided into three parts, namely:

PART I  GENERAL THEORY AND IMPLEMENTATION WITH ELECTRICAL COMPONENTS

PART II  IMPLEMENTATION WITH FLUIDIC COMPONENTS

PART III  STUDENT TECHNOLOGY TRANSFER

The three parts are written in such a way that they can be read separately.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOMENCLATURE</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xvi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xvii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>xx</td>
</tr>
<tr>
<td>PART I GENERAL THEORY AND IMPLEMENTATION WITH ELECTRICAL COMPONENTS</td>
<td></td>
</tr>
<tr>
<td>Chapter 1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Chapter 2. Analog function generation: How it is done</td>
<td>4</td>
</tr>
<tr>
<td>Chapter 3. The basic operations used in function generation</td>
<td>6</td>
</tr>
<tr>
<td>A. Logarithm</td>
<td>6</td>
</tr>
<tr>
<td>B. Exponential</td>
<td>13</td>
</tr>
<tr>
<td>C. Multiplication</td>
<td>13</td>
</tr>
<tr>
<td>1) variable transconductance multiplier</td>
<td>14</td>
</tr>
<tr>
<td>2) pulse width/pulse height modulation</td>
<td>20</td>
</tr>
<tr>
<td>3) log-antilog multiplier</td>
<td>23</td>
</tr>
<tr>
<td>D. Division</td>
<td>26</td>
</tr>
<tr>
<td>1) inverted multiplier</td>
<td>26</td>
</tr>
<tr>
<td>2) direct variable transconductance divider</td>
<td>28</td>
</tr>
<tr>
<td>3) log-antilog divider</td>
<td>28</td>
</tr>
<tr>
<td>E. Squaring and square rooting</td>
<td>29</td>
</tr>
<tr>
<td>F. Multifunction generation</td>
<td>29</td>
</tr>
<tr>
<td>Chapter 4. Features of pulse width modulated systems</td>
<td>33</td>
</tr>
<tr>
<td>Chapter 5. General nonlinear modulation</td>
<td>37</td>
</tr>
<tr>
<td>Chapter 6. Nonlinear pulse width modulation</td>
<td>40</td>
</tr>
</tbody>
</table>
Chapter 7. Demodulation................................................................. 51
   A. Filtering of the pulse width modulated signal.... 51
   B. Integration of the pulse width modulated signal
      over one cycle............................................................... 56
   C. Pulse width modulation in the feedback loop
      of an integrator............................................................ 58

Chapter 8. Computations with nonlinear pulse width modulation... 64

Chapter 9. Exponentiation of signals................................. 77

Chapter 10. Response of the system to a general input......... 87

Chapter 11. Response of the system to a sinusoidal input signal.. 100

Chapter 12. Simulation on digital computer......................... 100

Chapter 13. Accuracy of the system................................. 110
   A. Carrier signals....................................................... 110
   B. Comparators......................................................... 112
   C. Integrator.............................................................. 112
   D. Accuracy............................................................... 113
   E. Sensitivity analysis............................................... 114
   F. Components......................................................... 116

Chapter 14. Experiments on the analog computer............... 120

Conclusion........................................................................ 128

PART II IMPLEMENTATION WITH FLUIDIC COMPONENTS

Chapter 15. Introduction to fluidic nonlinear computation..... 132

Chapter 16. Characteristics of fluidic devices.................. 136
   A. Resistors.............................................................. 136
   B. Volumes............................................................... 138
   C. Resistor-volume network ....................................... 138
   D. Amplifiers........................................................... 140
NOMENCLATURE

a - constant

$a_0$ - average value in fourier series approximation

$a_n$ - fourier cosine coefficient

$\hat{A}$ - dc component of sinusoidal input signal

$A$ - output level of comparator

$A_j$ - output level of comparator for signal $j(.)$, $j = x,y,w$

$Ave$ - average value

b - constant

$b_n$ - fourier sine coefficient

$\hat{B}$ - amplitude of sinusoidal input signal

$B$ - output level of comparator

$B_j$ - output level of comparator for signal $j(.)$, $j = x,y,w$

$B_k$ - bernoulli number of order $k$

$B_n$ - fourier coefficient (exponential notation)

C - denotes capacitor or capacitance

- amplitude of sawtooth carrier signal

$C_j$ - dc offset of the amplifier for signal $j(.)$, $j = x,w$

D - diameter of capillary tube

- denotes diode

D(.) - carrier signal

$D_j(.)$ - carrier signal for signal $j(.)$, $j = x,y,w$

$D_j^{-1}$ - inverse of the function $D_j(.)$

$e$ - 2.718.....

$e_l(.)$ - input signal of a network

- difference of input and carrier signal
$e_o(.)$ - output signal of a network

$E$ - energy in ripple

$E_c$ - levels of deadzone comparator

$E_i$ - energy of ripple into filter

$E_{in}$ - input voltage

$E_o$ - energy of the ripple out of filter

$E_{oj}$ - output voltage

$E_{i\text{max}}$ - maximum energy of the ripple into filter

$E_{o\text{max}}$ - maximum energy of the ripple out of filter

$f(.)$ - function of an argument

$g(.)$ - independent input signal

$G$ - parameter in sensitivity analysis

$h(.)$ - impulse response of a filter

$h_j$ - hysteresis of the comparator for $j(.)$, $j = x, w$

$H(jw)$ - transfer function of a filter with impulse response $h(.)$

$i$ - $1,2$: subscript for switching time instants denoting 1st and 2nd

$I$ - current through p-n junction

$I_B$ - base current

$I_c$ - collector current

$I_{cj}$ - collector current $j = 1,2...$

$I_e$ - emitter current

$I_{ej}$ - emitter current $j = 1,2...$

$I_{es}$ - emitter saturation current

$I_{esj}$ - emitter saturation current $j = 1,2...$

$I_{in}$ - input current

$I_0$ - extrapolated current for zero voltage across p-n junction
\( I_x \) - input current
\( I_y \) - input current
\( I_z \) - input current
\( I_{SE} \) - integral of the error squared
\( j \) - complex variable
\( x, y, w \): subscript to denote a variable associated with signal
\( k \) - constant
\( k_0 \) - constant
\( k_1 \) - constant
\( k_2 \) - constant
\( K \) - amplitude of carrier signal
\( K_j \) - amplitude of exponential carrier signal for signal \( j(.) \),
\( j = x, y, w \)
\( \ln \) - logarithm to the base \( e \)
\( L \) - period of sinusoidal input signal
\( m \) - exponent
\( n \) - 0, 1, 2, 3, 4, ....
\( N \) - frequency ratio of carrier signal and input signal
\( o(.) \) - output signal of filter
\( o_{dc}(.) \) - dc component of the output signal of filter
O(.) - response of filter for square wave

P - pressure
  - upper level of $S_q(.)$

P_1 - pressure upstream of a resistor

P_2 - pressure downstream of a resistor

P_{c_j} - input or control pressure of a fluidic device at port c_j,
  \quad j = 1, 2, 3, 4

P_i(.) - square wave pressure signal

P_{i_{\text{max}}} - maximum amplitude of $P_i(.)$

P_o - amplitude of exponential carrier signal pressure

P_{o_{j}} - output pressure of a fluidic element at port j, \ j = 1, 2

P_R - amplitude of pressure pulses out of modulator

P_S - supply pressure of fluidic elements

P_{S_2} - increased supply pressure for digital amplification

P_x - input pressure of the system

P_x' - input pressure of operational amplifier

P_w - output pressure of the system

P_w' - input pressure of operational amplifier

PWM - denotes pulse width modulator (modulation)

PWM(.) - pulse width modulated signal

q - electron charge

Q - volume flow
  - lower level of $S_q(.)$

Q_j - denotes transistor j = 1, 2, 3...

R - denotes resistance
  - external resistance for input and output of system

R_l - resistor in exponential wave shaping network

R_f - resistor in filter network for the output of the system
\( R_g \) - gas constant
\( R_L \) - load resistance, input resistance of operational amplifier
\( s \) - laplace operator
\( S \) - denotes a switch
\( S_j \) - denotes a switch \( j = 1, 2, \ldots \)
\( S_{q(.)} \) - square wave signal
\( t \) - time
\( t^* \) - \( t - nT \)
\( t' \) - \( t/T \)
\( t_0 \) - starting time
\( t_1 \) - length of a pulse
\( t_1 \) - time instant at which switching from level A to level B occurs
\( t_2 \) - time instant at which switching from level B to level A occurs
\( t_{e_n}^{'} \) - time instant at which the output signal equals the carrier signal for the second time
\( t_{ij} \) - switching time instant of the switch for \( j(.) \), \( j = x, y, w \)
\( t_{ij}' \) - \( t_{ij}/T \)
\( t_{jn} \) - switching time instant of the switch for signal \( j(.) \) in the interval \( (n-1)T < t < nT \)
\( t_{jn}' \) - \( t_{jn}/T \)
\( t_x \) - switching time instant of the comparator in the input modulator (fluidic)
\( T_x \) - length of pulse out of modulator for \( x(t) \) or \( P_x \)
\( t_w \) - switching time instant of the comparator in the feedback modulator (fluidic)
\( T \) - period of carrier signal
- absolute temperature

\( T_L \) - length of the pulse out of the digital amplifier

\( V \) - voltage across p-n junction

- volume

\( V_{be} \) - base emitter voltage

\( V_{bc} \) - base collector voltage

\( V_c \) - supply voltage

\( V_{in} \) - input voltage

\( V_o \) - volume in output filter network

\( V_{ol} \) - amplitude of square wave

\( V_{o2} \) - amplitude of pulses out of multivibrator

\( V_s \) - input signal voltage

\( V_w \) - volume in feedback modulator

\( V_x \) - input voltage

- volume in input modulator

\( V_y \) - input voltage

\( V_z \) - input voltage

\( w \) - running variable denoting output

\( w_0 \) - initial condition \( w(t_0) \)

\( w_n \) - value of the output at \( t = nT \) or \( t' = n \)

\( w(.) \) - output signal of the system

\( w'(.) \) - output signal of the modulator for \( w(.) \)

\( w'_{dc} \) - dc component of \( w'(.) \)

\( W \) - constant value of \( w(.) \)

\( x \) - running variable denoting input signal

\( x(.) \) - input signal of the system

\( x_j(.) \) - input signal \( j = 1,2,\ldots \)
$x'(.)$ - output signal of the modulator for $x(.)$

$x''(.)$ - $-x'(.)$

$x'_d$ - dc component of $x'(.)$

$X$ - constant value of $x(.)$

$X^*$ - $1 - X$

$y$ - running variable

$y(.)$ - additional input signal of the system

$y'(.)$ - output signal of the modulator for $y(.)$

$y'_d$ - dc component of $y'(.)$

$Y$ - constant value of $y(.)$

$z$ - running variable denoting input signal

$z(.)$ - input signal of the system

$Z$ - constant value of $z(.)$

$\alpha$ - ratio $I_c/I_e$

- denotes inverse time constant

$\alpha_1$ - inverse time constant

$\alpha_2$ - inverse time constant

$\alpha_j$ - inverse time constant for signal $j(.)$ or pressure $P_j$, $j = x,w$

$\alpha^*_j$ - $\alpha_j/100$

$\beta$ - ratio $P'_X/P_o$ or $P'_W/P_o$

- transistor gain

$\Delta$ - small increment

$\mu$ - absolute viscosity

$\pi$ - $3.14...$

$\rho$ - density

$\Sigma$ - summation sign
\( \tau \) - denotes time constant
\( \tau' \) - \( \tau / T \)
\( \tau_c \) - time constant of exponential carrier signal (fluidic)
\( \tau_i \) - time constant of integrator for demodulator
\( \tau'_i \) - \( \tau_i / T \)
\( \tau_r \) - time constant of output filter network
\( \tau_{rm} \) - measured time constant of the response
\( \phi \) - phase angle
\( \omega \) - radial frequency
\( \cdot \) - argument of a function
\( \neg x \) - denotes logical inverse of \( x \)
LIST OF TABLES

Table 8.1  dc components of pulse width modulated
signals ............................................................ 68
Table 8.2  Realizable nonlinear functions .................. 69
Table 9.1  Quadrants of operation .......................... 84
Table 10.1  Comparison of real and approximated
responses ...................................................... 93
Table 13.1  Minimum absolute values of the input signal
x(t) ............................................................... 111
Table 14.1  Values of the exponent in the experiments ... 122
Table 14.2  Measured and theoretical values of input and
output ......................................................... 127
Table 16.1  Truth table for flip-flop ........................ 144
Table 18.1  Different sizes of volumes yielding different
exponents ..................................................... 166
<table>
<thead>
<tr>
<th>Figure</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure</td>
<td>3.1</td>
<td>Voltage-current relation of a forward biased p-n junction</td>
</tr>
<tr>
<td>Figure</td>
<td>3.2</td>
<td>Generating a logarithmic function with a diode</td>
</tr>
<tr>
<td>Figure</td>
<td>3.3A</td>
<td>Transdiode</td>
</tr>
<tr>
<td>Figure</td>
<td>3.3B</td>
<td>Diode connected transistor</td>
</tr>
<tr>
<td>Figure</td>
<td>3.4</td>
<td>Logarithmic function generator</td>
</tr>
<tr>
<td>Figure</td>
<td>3.5</td>
<td>NPN transistor</td>
</tr>
<tr>
<td>Figure</td>
<td>3.6</td>
<td>Nonlinearity and temperature compensation</td>
</tr>
<tr>
<td>Figure</td>
<td>3.7</td>
<td>Linearized transconductance amplifier</td>
</tr>
<tr>
<td>Figure</td>
<td>3.8</td>
<td>Pulse width/pulse height modulation</td>
</tr>
<tr>
<td>Figure</td>
<td>3.9</td>
<td>PWM with external sawtooth signal</td>
</tr>
<tr>
<td>Figure</td>
<td>3.10</td>
<td>PWM without external signal</td>
</tr>
<tr>
<td>Figure</td>
<td>3.11</td>
<td>Pulse height modulation</td>
</tr>
<tr>
<td>Figure</td>
<td>3.12</td>
<td>Inverted multiplier</td>
</tr>
<tr>
<td>Figure</td>
<td>3.13</td>
<td>Square rooting</td>
</tr>
<tr>
<td>Figure</td>
<td>3.14</td>
<td>Multifunction module</td>
</tr>
<tr>
<td>Figure</td>
<td>5.1</td>
<td>General non-linear modulation</td>
</tr>
<tr>
<td>Figure</td>
<td>6.1</td>
<td>Nonlinear PWM</td>
</tr>
<tr>
<td>Figure</td>
<td>6.2</td>
<td>Sawtooth carrier signal</td>
</tr>
<tr>
<td>Figure</td>
<td>6.3</td>
<td>Exponential carrier signal</td>
</tr>
<tr>
<td>Figure</td>
<td>6.4</td>
<td>Generation of second order saturation function carrier signal</td>
</tr>
<tr>
<td>Figure</td>
<td>6.5</td>
<td>Second order saturation function</td>
</tr>
<tr>
<td>Figure</td>
<td>7.1</td>
<td>PWM signal</td>
</tr>
<tr>
<td>Figure</td>
<td>7.2</td>
<td>Demodulation by integration over one cycle</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Figure 7.3</td>
<td>Recovering of PWM signals</td>
<td>59</td>
</tr>
<tr>
<td>Figure 7.4</td>
<td>Improved feedback pulse width demodulator</td>
<td>61</td>
</tr>
<tr>
<td>Figure 7.5</td>
<td>Pulse width subtraction</td>
<td>62</td>
</tr>
<tr>
<td>Figure 8.1</td>
<td>System which performs nonlinear computation</td>
<td>65</td>
</tr>
<tr>
<td>Figure 8.2</td>
<td>Raising a signal to a power</td>
<td>72</td>
</tr>
<tr>
<td>Figure 8.3</td>
<td>Sinusoidal and triangular wave carrier signal</td>
<td>73</td>
</tr>
<tr>
<td>Figure 8.4</td>
<td>Ripple due to phase difference in carrier signals</td>
<td>75</td>
</tr>
<tr>
<td>Figure 9.1</td>
<td>Generation of an exponential carrier signal</td>
<td>78</td>
</tr>
<tr>
<td>Figure 9.2</td>
<td>Generation of an exponential carrier signal</td>
<td>79</td>
</tr>
<tr>
<td>Figure 9.3</td>
<td>Pulse height modulation</td>
<td>82</td>
</tr>
<tr>
<td>Figure 9.4</td>
<td>System that performs multifunction operation</td>
<td>83</td>
</tr>
<tr>
<td>Figure 9.5</td>
<td>Stability analysis</td>
<td>86</td>
</tr>
<tr>
<td>Figure 10.1</td>
<td>Response to Input $x(t')$, $t' &lt; t$</td>
<td>90</td>
</tr>
<tr>
<td>Figure 10.2</td>
<td>Three different types of responses to input $x(t')$, $t' &gt; t$</td>
<td>95</td>
</tr>
<tr>
<td>Figure 10.3</td>
<td>Step response</td>
<td>98</td>
</tr>
<tr>
<td>Figure 11.1</td>
<td>Response to a sinusoidal signal with $D_x(t') = D_w(t')$, so $w(t') = x(t')$</td>
<td>102</td>
</tr>
<tr>
<td>Figure 12.1</td>
<td>Amplitude and phase of first harmonic</td>
<td>108</td>
</tr>
<tr>
<td>Figure 12.2</td>
<td>Integral error squared</td>
<td>109</td>
</tr>
<tr>
<td>Figure 14.1</td>
<td>Elimination of unequal levels</td>
<td>121</td>
</tr>
<tr>
<td>Figure 14.2</td>
<td>Deadzone comparator and symbol</td>
<td>123</td>
</tr>
<tr>
<td>Figure 14.3</td>
<td>Network for generating carrier signals</td>
<td>124</td>
</tr>
<tr>
<td>Figure 14.4</td>
<td>Analog computer diagram</td>
<td>126</td>
</tr>
<tr>
<td>Figure 16.1</td>
<td>Diagram of resistor and volume components</td>
<td>137</td>
</tr>
<tr>
<td>Figure 16.2</td>
<td>Passive resistor volume network</td>
<td>139</td>
</tr>
<tr>
<td>Figure 16.3</td>
<td>Beam deflection amplifier and gain characteristic</td>
<td>141</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>16.4</td>
<td>Flip-flop characteristic</td>
<td>143</td>
</tr>
<tr>
<td>16.5</td>
<td>OR-NOR characteristics</td>
<td>145</td>
</tr>
<tr>
<td>16.6</td>
<td>Digital amplifier</td>
<td>147</td>
</tr>
<tr>
<td>17.1</td>
<td>Logarithmic PWM</td>
<td>149</td>
</tr>
<tr>
<td>17.2</td>
<td>Fluidic oscillators</td>
<td>150</td>
</tr>
<tr>
<td>17.3</td>
<td>Resistor-volume network</td>
<td>152</td>
</tr>
<tr>
<td>17.4</td>
<td>Comparator and logarithmic PWM</td>
<td>154</td>
</tr>
<tr>
<td>18.1</td>
<td>Nonlinear computation</td>
<td>157</td>
</tr>
<tr>
<td>18.2</td>
<td>Computation of powers and roots</td>
<td>159</td>
</tr>
<tr>
<td>18.3</td>
<td>Pulses out of NOR</td>
<td>161</td>
</tr>
<tr>
<td>18.4</td>
<td>Square wave</td>
<td>163</td>
</tr>
<tr>
<td>18.5</td>
<td>Exponential carrier signal</td>
<td>164</td>
</tr>
<tr>
<td>18.6</td>
<td>Measurements</td>
<td>167</td>
</tr>
<tr>
<td>18.7</td>
<td>Step response</td>
<td>170</td>
</tr>
<tr>
<td>19.1</td>
<td>Model of change</td>
<td>177</td>
</tr>
<tr>
<td>19.2</td>
<td>Transmitter-Channel-Receiver model</td>
<td>182</td>
</tr>
<tr>
<td>A111.1</td>
<td>Square wave signal</td>
<td>210</td>
</tr>
</tbody>
</table>
ABSTRACT

This dissertation describes a new method for function generation and nonlinear analog computation. The method is based on pulse width modulation with a periodic carrier signal that is not a saw-tooth or triangular wave signal. Instead, the carrier signal wave forms are sinusoidal, exponential, etc. If, for example, the carrier signal is a sine wave, then the average value of the output signal of the modulator is the arcsin function of the input modulating signal. Demodulation is done by feeding the pulse width modulated signal into an integrator which has a pulse width modulator in its feedback path. Using various carrier signal wave forms for input and feedback pulse width modulators, nonlinear computations like $X^aY^b$, $Ye^aX$, $aX+b\ln X$, $b(\arcsin X)$, etc. can be performed.

The generation of the function $W = \text{sign}(X).Z(X/Y)^m$ is described in more detail. Experiments on the analog computer showed that an accuracy of better than 0.2% of full scale was obtainable. The step frequency response and the signal to noise ratio were determined through simulation on a digital computer.

This new method for nonlinear computation also opened the way to fluidic nonlinear computation. A system that computed $W = X^a$ was implemented with fluidic components for several values of the exponent $a$. The accuracy was 1-2% of full scale, except for very large or very small signals in which case the accuracy decreased to 5-10% of full scale.
Finally, the progress of research leading to a dissertation is viewed as a process of technological change. In this process, the movement of technology from industry to student and vice versa is an indispensable component and methods for promoting it are investigated.
PART I

GENERAL THEORY AND IMPLEMENTATION WITH
ELECTRICAL COMPONENTS

Chapter I. Introduction

The rapid development of digital computation over the past twenty years has gradually diminished the need for solving problems by way of analog computation. The introduction of simulation languages (i.e. CSMP) for the digital computer caused many people to desert the analog computer, especially because these languages could handle nonlinear problems quite easily, which is a major obstacle in doing analog computation. However, the analog computer still has its advantages, mainly in the following areas:

- The simulation of modulated systems.

The modulation of signals occurs in communication systems and increasingly more so in control systems. In most of these systems the carrier signal has a much higher frequency than the modulating signal. If these systems are simulated on a digital computer, the choice of the basic solution increment must be based on the period of the carrier signal and consequently, is quite small. However, in most cases, the transient response of the system is of the same order as the modulating signal. Because of this time scale difference the digital computer would require many computations to solve for one transient response, resulting in a great amount of computer time. One author found that the simulation of a pulse
Width modulated control system on a digital computer was extremely time consuming [1]. Since the time scale is not quantized in analog simulation, the above problem does not exist, provided that the frequency response of the elements is satisfactory for the spectrum of the signals.

Determining solutions which depend on many parameters.

In this type of solution it is time consuming to use a digital simulation to determine the response of the system for different parameters. With the analog computer it is possible to quickly see the influence of a change especially when the analog computer is used in the repetitive operation mode.

Despite the advantages mentioned above, the use of the analog computer to solve nonlinear problems still presented enough difficulties to preclude its frequent use. However, in the past few years, commercially available inexpensive, versatile, nonlinear analog multifunction modules have revived the interest in analog computation. These modules can generate a variety of nonlinear functions and thus overcome one of the major disadvantages of analog computers. Besides the application in analog computers, these nonlinear analog function modules can perform signal processing to provide nonlinear analog computation [2]. The basic functions that most of these modules can realize are [3]: multiplication, division, squaring, square rooting, logarithms and exponentials. The functions that can be derived from these are: arbitrary exponents, root mean square, log ratio, sinh⁻¹, vector sum and trigonometric functions. This variety of functions finds applications in many areas. To name a few:
mass flow computation [4]
oxygen concentration measurement [4]
dynamometer control [5]
automatic gain control [6, 16]
phase detection between two signals [6]
power measurement [6, 16]
linearization of transducer output [7, 16, 18]
logarithmic converters [8]

From the above examples it is seen that nonlinear analog multifunction modules have applications in the areas of signal processing, instrumentation and control. They form building blocks which design engineers can easily interconnect to realize desired input-output relationships. Moreover, they reduce the effort required, so they save time and money [9, 14, 18].

This dissertation will treat a new nonlinear multifunction generator and deal with some of its applications. Particular attention will be paid to the electrical and fluidic implementation of the novel concept of nonlinear multifunction generation. However, an overview of the existing techniques for nonlinear multifunction generation will be given first.
Chapter 2. Analog Function Generation: How It Is Done

This chapter will give an overview of how nonlinear analog function generation is generally accomplished. No distinction will be made between nonlinear analog function generation and nonlinear analog computation and they will be used interchangeably throughout the text. In nonlinear function generation the output \( w(t) \) is a nonlinear function of one or more variables \( x_1(t), \ldots, x_n(t) \), i.e. \( w(t) = \prod_{i=1}^{n} x_i(t) \). The nonlinear function to be represented can either be derived from collected data or can follow from theoretical considerations.

There are two different types of functions that are generated; namely, rational functions and arbitrary functions [3]. Rational functions are functions that can be generated exactly by using the basic operations: multiplication, division, squaring, square rooting, logarithms, exponentials, addition and subtraction. An example is the generation of a polynomial. Also, the generation of inverse functions by using rational functions in the feedback path of a high gain amplifier yields a rational function [7], as well as the solution of implicit equations implemented with circuits containing the basic operations. In contrast, arbitrary functions are functions that cannot be generated by use of the basic operations. One way this problem has been solved is by using the well known diode function generator [3, 10, 18], which can generate an arbitrary function \( w = f(x) \) by approximating it using straight line segments. This diode function generator has great flexibility, but is not very accurate for small input signals, unless a large number of diodes is used. Moreover, choosing
the breakpoints of the straight line approximation might require some computation and can be rather cumbersome [11, 19]. One of the best known applications of the diode function generator is in the quarter square multiplier, which operates on the principle:

\[ xy = \frac{1}{4}(x+y)^2 - \frac{1}{4}(x-y)^2 \]

In this case, the diode function generator is used to provide the squaring operation. It is also possible to generate arbitrary functions of \( n \) variables with diode function generator [12]. However, the problems associated with the segment specification become more difficult. An interesting application would be to design a diode function generator suitable for two variables to perform the multiplication or division operation. A second way of generation arbitrary functions is by using a number of terms of a series expansion around a suitable operating point and implementing the series with a circuit using the basic rational functions operations [3]. An advantage of this approach is that the function is smooth instead of piecewise linear. Approximating \( \sin \theta \), \( \cos \theta \), and \( \arctan(y/x) \) has been reported [13, 15]. In another case, the summation of a number of exponential curves was used to approximate the time function \( 1/t \) [27].

From the previous discussion about nonlinear function generation it is seen that the basic rational function operations are important in order to be able to generate nonlinear functions. Therefore, the next chapter will give a brief treatment of the existing techniques to perform the basic operations.
Chapter 3. The Basic Operations Used In Function Generation

The basic operations used in function generation are: multiplication, division, squaring, square rooting, logarithms, exponentials, addition and subtraction. Since the implementation of the addition and subtraction operation is well known, it will not be dealt with here. The remainder of the basic operations will be treated in this section. Since the logarithm and exponential operations sometimes are used to perform the other operations, they will be treated first.

A. Logarithm

Generating the logarithm is usually done by making use of the exponential behavior of a semiconductor junction. The generalized voltage-current relationship for a p-n junction is shown (Fig. 3.1) [48]. If only diffusion current is considered, the following simple voltage-current relation for a forward biased p-n junction can be derived:

\[ I = I_0 (e^{qV/kT} - 1) = I_0 e^{qV/kT} \]  

\( I \) - current through the Junction [Amp.]
\( I_0 \) - extrapolated current for \( V = 0 \) [Amp.]
\( V \) - voltage across the junction [V]
\( q \) - electron charge \( 1.60219 \times 10^{-9} \) [C]
\( k \) - Boltzmann's constant \( 1.38062 \times 10^{-23} \) [J/°K]
\( T \) - absolute temperature [°K]

Taking the logarithm of both sides of equation 3.1 yields the following voltage-current relationship:

\[ V = \frac{kT}{q} \ln \frac{I}{I_0} \]
Fig. 3.1 Voltage-current relation of a forward biased p-n junction
A way to make use of this logarithmic relationship is to put a diode junction in the feedback path of an operational amplifier (Fig. 3.2) [3]. The output voltage $E_0$ is given by:

$$E_0 = \frac{kT}{q} \ln \frac{I_n}{I_0} = \frac{kT}{q} \ln \frac{V_{in}}{V_{in0}}$$

The above circuit has the disadvantage that the logarithmic relationship is not very accurate, so its useful range is only one or two decades. The inaccuracies are due to surface losses and generation-recombination effects in the p-n junction [21, 26]. Using the junction in a transistor to obtain the logarithmic relationship gives a much greater conformity to the desired law (Figs. 3.3A, 3.3B). In this case, the collector to base voltage is equal to zero and the above effects of surface losses and generation-recombination effects do not play a role [21]. The main difference between the two circuits is that the diode connected transistor can accept signals of either polarity, whereas the transdiode configuration only works for negative input signals. Since the junction characteristics are very much dependent on the temperature, a temperature compensation is needed. If this is done, both circuits will have a useful range of about six decades and an accuracy of about 1% full scale [2, 18].

The possibilities of performing logarithmic conversion by using the exponential characteristics of RC networks have not been generally recognized. However, one author writes [23]:

An interesting method for generating the logarithmic function is to form a wave train of high repetitive frequency consisting of functions of the form $Ke^{-a^t}$ and to measure the time $t_1$ for each exponential to decay to the value $x$. If $x$ is essentially constant, we have:

$$x = Ke^{-a^t}$$

$$t_1 = - \frac{1}{a} \ln x/K$$
Fig. 3.2 Generating a logarithmic function with a diode
Fig. 3.3A
Transdiode

Fig. 3.3B
Diode connected transistor
Using this principle, a logarithmic function generator has been built [24, 44]. The construction of this generator is shown [Fig. 3.4].

The generator (Fig. 4) includes an astable multivibrator feeding into an RC circuit. The resulting exponentially decaying voltage is compared with an input signal. The generator output is proportional to the time required for the exponential voltage to decay from the preset reference level to the level of input signal.

The time constant of the RC differentiator circuit is selected according to the desired dynamic range of the converter. Diode limiter D ensures a zero voltage across the differentiator at the beginning of each cycle.

Both \( V_B \) and a dc voltage \( V_S \) proportional to the input signal are fed to a differential comparator. The comparator has a differentiated output which triggers a bistable multivibrator. Positive pulses at times \( T = 0, T, \) etc. cause the multivibrator positive transition (leading edge). The negative pulses of \( V_C \) at times \( t = t_1, T + t_1, \) etc. are used to reset the multivibrator (negative edge). The resulting pulse width is proportional to \( t \) and is a measure of \( \log (V_0 / V_S) \).

This way of logarithmic function generation has improved accuracy over the semiconductor junction method. Another application of the same principle was found in logarithmic analog-to-digital conversion [25]. The exponential signal was used to generate the logarithm and the pulse output gated the output of a clock oscillator to a counter. By counting the number of pulses from the clock oscillator, logarithmic A/D conversion was accomplished. This resulted in logarithms with errors that were less than 1%. Limitations on the accuracy are caused by [26]:

- the characteristics of timing components \( R \) and \( C \) and their variations with temperature
- the stability of the voltage \( V_0 \)
- the offset voltage and the bias current variations of the comparator.
Fig. 3.4 Logarithmic function generator [34]
A disadvantage of logarithmic conversion with RC networks is that it is only suitable for positive signals. Later on, a logarithmic function generator will be developed which can accept signals of either polarity and which is comparable to the diode connected transistor logarithmic converter.

B. Exponential

Generating the exponential function using a semiconductor junction is quite similar to generating the logarithm. The difference is that the places of the input resistor and feedback transistor are interchanged (see Figs. 3A and 3B), so the transistor is in the input path and the resistor is in the feedback path. Since no basic changes in the circuitry have been made, the performance of the exponential converter is equivalent to that of the logarithmic converter.

It would also be possible to generate exponential functions by using RC logarithmic converters in the feedback path of an operational amplifier. Because the logarithmic function is generated very accurately, this method should give accurate exponential functions as well. However, no applications of this method were found.

C. Multiplication

There are various ways in which multiplication can be performed. A list of those include [10]:

- pulse width/pulse height modulation
- amplitude/phase modulation
- dual pulse width modulation (coincidence)
- dual amplitude modulation
- servo multipliers
- heat transfer
-Hall effect and magneto resistance
-pulse rate
-quarter square
-triangle averaging
-variable transconductance
-log-antilog

At present, the two most common means of performing electronic analog multiplication are variable transconductance and pulse width/pulse height modulation [17, 3]. A third method, log-antilog, is gaining in popularity, particularly for low speed higher accuracy calculations [3]. The variable transconductance multiplier has a large bandwidth, 1-10 MHz, but mediocre accuracy, 0.5-2.0% of full scale [2]. The pulse width/pulse height multiplier has a smaller bandwidth, 100 kHz, but is very accurate, 0.1% of full scale. However, it is more expensive. Finally, the log-antilog multiplier has a bandwidth of about 0.5-1.0 MHz and an accuracy of 0.25-0.5% of full scale. It can, however, only work in one quadrant, yet it is more versatile in that it can easily compute powers and roots. A brief description of the three most common multipliers will be given.

1) Variable transconductance multiplier.

The variable transconductance principle is that one input controls the gain (transconductance) of an amplifier, which amplifies the other input proportional to the control input [3]. For a transistor (Fig. 3.5), the multiplication property can be observed by differencing the simplified junction equation:
Fig. 3.5 NPN transistor
\[ l_c = l_{es} e^{\frac{qV_{be}}{kT}} \]

- \( l_{es} \) - emitter saturation current
- \( V_{be} \) - base emitter voltage
- \( l_c \) - collector current

Taking the difference yields:

\[ \Delta l_c = \frac{q}{kT} l_c \Delta V_{be} \]

Multiplication elements mechanized by making \( l_c \) proportional to one variable and \( \Delta V_{be} \) to the other, suffer from two problems:

- nonlinear logarithmic relationship between base emitter voltage and collector current, which nonlinearity would only be neglectable for very small \( \Delta V_{be} \).
- considerable temperature dependence

It is possible to compensate for the nonlinearity and temperature dependence by constructing the following circuit (Fig. 3.6) [28]. The logarithmic nonlinearity of transistor \( Q_2 \) is compensated for by the antilog connection of the diode connected transistor \( Q_1 \). The equation for the circuit is:

\[ \frac{kT_1}{q} l_{e1} \frac{l_{e1}}{l_{es1}} = \frac{kT_2}{q} l_{e2} \frac{l_{e2}}{l_{es2}} \]

- \( l_{es1} \) - saturation current of transistor \( Q_1 \)
- \( l_{es2} \) - saturation current of transistor \( Q_2 \)

The above yields:

\[ l_{e2} = l_{es2} \frac{l_{e1}}{l_{es1}} \]
Fig. 3.6 Nonlinearity and temperature compensation [28]
Now, the influence of the temperature and the nonlinearity has disappeared. The ratio $\frac{I_{s2}}{I_{s1}}$ is fixed for a given transistor pair, but it does not have to be unity, in which case there is a fixed proportional gain. To obtain the transconductance property, this compensated circuit is applied in a differential amplifier configuration (Fig. 3.7), which was invented by Gilbert [29]. For an input signal $\Delta x$, this circuit has a differential input $I_{x+\Delta x}, I_{x-\Delta x}$ and emitter currents $I_{e+\Delta e}, I_{e-\Delta e}$, where $I_{e}$ is kept constant. For the analysis the following assumptions are made:

1) $I_{es} = I_{es_j}$, $j = 1, 2, 3, 4$
2) perfect exponential relation between base emitter voltage and collector current
3) each transistor has a beta which is nearly infinite

Summing the emitter voltages around the $Q_1-Q_4$ loop yields:

$$\frac{kT}{q} \left[ \ln \frac{I_{x+\Delta x}}{I_{es}} - \ln \frac{I_{e+\Delta e}}{I_{es}} - \ln \frac{I_{x-\Delta x}}{I_{es}} + \ln \frac{I_{e-\Delta e}}{I_{es}} \right] = 0$$

From this, the following relation can be derived:

$$\frac{I_{x+\Delta x}}{I_{x-\Delta x}} = \frac{I_{e+\Delta e}}{I_{e-\Delta e}}$$ (3.2)

or

$$\Delta I_e = \frac{\Delta I_c}{I_c}$$ (3.3)

which demonstrates the linear transconductance property. The multiplication operation can readily be shown. Let the output be the difference $\Delta I_c$ of the collector currents $I_{c_2}$ and $I_{c_3}$:
Fig. 3.7 Linearized transconductance amplifier [28]
If the $y$ input is the sum of the two emitter currents, then

$$I_y = 2I_e$$

Substitution of the above equations in equation 3.3 yields:

$$\Delta I_c = \frac{\Delta I_x I_y}{I_x}$$  \hspace{1cm} (3.4)

The multiplication is thus accomplished.

The main sources of error are due to the afore-mentioned assumptions. With better processing and matching techniques the accuracy is expected to improve over the current 1% of full scale.

2) Pulse width/pulse height modulation

The pulse width/pulse height modulation multiplication method operates on the principle that one input signal ($y$) controls the duty cycle or pulse width of the pulse train (with period $T$), while the other input signal ($x$) controls the height of the pulse (Fig. 3.8). If this pulse train is filtered, the output of the filter will give the average value of the pulse train. This is expressed by:

$$\text{Ave} = \frac{(T/2+ky)x}{T} - \frac{(T/2-ky)x}{T} = 2kxy$$

It is seen that the output of the filter will be proportional to the product $xy$ of the input signals $x$ and $y$.

There are different ways in which duty cycle generation or pulse width modulation are accomplished:

a) Comparing the input signal with a sawtooth signal or a triangular wave signal (Fig. 3.9). For a sawtooth signal the width $t_1$ of the pulse is given by the equation:
Fig. 3.8 Pulse width/pulse height modulation
Fig. 3.9 PWM with external sawtooth signal
\[ y = -C + 2Ct/T \]
or: \[ t_1 = T/2 + Ty/2C \]

Observe that a zero input signal \( (y=0) \) yields a pulse width of \( T/2 \)
b) Constructing a system that is in a limit cycle with a comparator with hysteresis and an integrator (Fig. 3.10) [3, 10]. In the steady state, the integrator forces the sum of the dc components of its input to zero, causing the average value and thus the length of the pulses out of the comparator to be proportional to the input signal \( y \). This circuit tends to be more accurate, because the feedback path compensates for any timing errors or switching delays in the comparator. It can be shown that the period of the pulse train depends on the magnitude of the input signal \( y \). However, in many applications this is not a disadvantage of the circuit.

The pulse height modulation can be accomplished quite easily by driving a mechanical relay or semiconductor switch with the pulse width modulated signal, which alternately drives the switch to \( x \) and \( -x \) (Fig. 3.11). The output will be the pulse width/pulse height modulated signal (Fig. 3.8), which after filtering yields the product. The necessity of filtering decreases the bandwidth and makes it a factor 10 or 100 less than the bandwidth of the variable transconductance multiplier. The high accuracy of the pulse width/pulse height modulation multiplication is due to the fact that it does not rely on the exact relationship of a nonlinear element (like the semiconductor junction), but instead uses the properties of switches.

3) Log-antilog multiplier

The log-antilog multiplier is based on the identity:

\[ xy = \text{antilog}(\log x + \log y) \]
Fig. 3.10 PWM without external signal
Fig. 3.11 Pulse height modulation
The generation of the logarithmic and exponential relation was already treated, so the multiplication becomes obvious. The logarithms are generated by means of the transdiode or the diode connected transistor. The addition is performed with operational amplifiers, after which the antilog operation yields the product.

It is possible to generate powers and roots with logarithms, because the following holds:

\[ x^a y^b = \text{antilog}(\text{alog}x + \text{blog}y) \]

This is easily implemented. With temperature compensation it is possible to achieve accuracies of 0.2-0.5% of full scale and better with a bandwidth of about 1 MHz \[2\].

Performing log-antilog multiplication with RC logarithmic converters has not been mentioned in the literature. However, it is a method that is quite feasible and it will be dealt with in the following chapters.

D. Division

Since most of the principles of multiplication can also be applied to division, only a brief treatment of the division operation will be given here. The three most common ways in which division is performed are \[3\]:

1) inverted multiplier
2) direct variable transconductance divider
3) log-antilog divider

1) Inverted multiplier

This technique is based on putting a multiplier in the feedback path of a high gain amplifier (Fig. 3.12). For this circuit the following relationship holds:
Fig. 3.12  Inverted multiplier
\[ z = -xy \quad \text{or} \quad E_0 = y = - \frac{z}{x} \]

Of course, the performance of this divider depends on the performance of the multiplier. However, in most cases, the performance of the divider is much worse (it can have about 10 times the error of a multiplier), so an accurate multiplier is desirable in this configuration [30].

2) Direct variable transconductance divider

The variable transconductance divider can easily be explained by investigating equations 3.2, 3.3 and 3.4. At first, \( l_x + \Delta l_x \) was one signal and \( l_x - \Delta l_x \) the other. The current \( l_x \) was a constant and the input was \( \Delta l_x \). However, if \( l_x \) is also considered to be an input and the former input \( \Delta l_x \) is changed to the input \( l_z \), then the division operation follows from equation 3.4:

\[ l_c = \frac{l_z}{l_x} \cdot y \]

The signal \( l_y \) is used for scaling purposes. As general advantages of the variable transconductance divider, its good accuracy of about 0.5% of full scale and its large bandwidth of 500 kHz should be mentioned [40].

3) Log-antilog divider

The log-antilog divider is based on the identity:

\[ \frac{x}{y} = \text{antilog}(\log x - \log y) \]

The implementation is simple and was treated in the sections dealing with logarithm generation and multiplication. The log-antilog divider has the highest accuracy, 0.25% of full scale, but has less bandwidth [31, 40].
E. **Squaring and square rooting**

The squaring operation is mainly accomplished by having two identical inputs to a multiplier. Of course, all the remarks made in dealing with the multiplication operation are valid for the squaring operation as well.

The square root is usually obtained in one of the following two ways:

- **log-antilog**
  \[ x^{1/2} = \text{antilog}(1/2 \log x) \]

- feedback around a divider (Fig. 3.13)

The following equation holds:

\[ z = \frac{y}{z} \quad \text{or} \quad z = y^{1/2} \]

Both methods give fairly accurate results.

So far, this chapter gave an overview of the existing methods for performing the basic operations. The next section will deal with the multifunction generator which is a module that can perform most of the basic operations plus a few additional ones simply by changing some external connections.

F. **Multifunction generation**

The accuracy and ease with which the logarithmic function can be generated has facilitated the construction of multifunction modules. Such modules generate the function:

\[ E_o = V_y \left( \frac{V_z}{V_x} \right)^m \quad \text{generally} \ m = 0.2 \text{ to } 5.0 \]

This transfer function includes many of the basic operations, and thus, is very useful. The circuit is shown (Fig. 3.14) [3]; the dotted lines enclose the module. By connecting external resistors between A, B and
Fig. 3.13 Square rooting
Fig. 3.14 Multifunction module [3]
the exponent \( m \) can be adjusted. With some more external circuitry the sine, cosine and arctangent can also be computed. The accuracy of the multifunction modules is good, about 0.3\% of full scale [3] in the division operation, which is the least accurate mode of operation, is typical.

From the above discussion, it is clear that over the past few years a great expansion in the field of nonlinear analog computation has taken place. This culminated in the construction of a multifunction module. This dissertation describes a different way for performing multifunction operation, which greatly expands the number of functions that can be generated. The approach is based on pulse width modulation which is inherently accurate as was demonstrated in the pulse width/pulse height multiplier. However, before dealing with this method, the general features of pulse modulated systems will be outlined.
Chapter 4. Features of Pulse Width Modulated Systems

Linear pulse width modulation has been applied in instrumentation and control systems where it is desirable to convert an analog signal into the time duration of a pulse. The main advantages of applying pulse width modulation are provided by its on-off nature, which has high accuracy, high reliability and allows more efficient power amplification. In power amplification, the pulse width modulation notion permits power transistors to be controlled in the on-off mode, where the losses in their junctions are minimized [32,33,34]. In control systems, the pulse width modulation provides dither and, hence, greatly reduces the influence of coulomb friction and, at the same time, linearizes the nonlinear element. Moreover, it provides a limit on the maximum output which can be important in electric motors [45]. Further, the modulator gain can be adjusted. The amplitude of the carrier signal determines the gain characteristic of the pulse width modulator. This approach was used in one of the early flight control systems with a sinusoidal carrier signal to provide gain control in the system. In that system the amplitude of the sinusoidal signal was used to exercise adaptive control over the gain. In signal processing, analog multiplication can be accomplished by combined pulse width/pulse height modulation [10, 35, 2]. This was treated before and it is known that the method gives very accurate results. Furthermore, the information can be recovered in both analog and digital form. If the pulse train is filtered, then the signal is recovered in analog form. If the duration of the pulse is used to gate a clock into a counter, then the information may be recovered in digital form, which is one way of constructing analog-to-digital converters [10]. So, the
Information is easily converted by the pulse width modulation process into a form which can be recovered by analog or digital methods. Finally, in communication systems, the pulse width modulation is known to have superior signal to noise characteristics at the cost of greater transmission bandwidth [49].

This dissertation describes a kind of nonlinear pulse width modulation in which the carrier signal is not a linear function of time, like a sawtooth, but some nonlinear function of time, such as a sine wave or an exponential. It can be shown that if the input analog signal is mixed with a carrier signal, such as a sinusoidal or exponential carrier signal, the dc level of the resulting output signal will be proportional to the inverse of the function that represents the carrier signal over one cycle. Examples of nonlinear functions that can be realized are $\ln x$ and $\sin^{-1}x$. In the literature, this function generation feature is realized only to a limited extent. Usually, it appears in analyzing nonlinear quantized control systems in which the addition of a high frequency periodic signal (dither) to the error signal results in an altered characteristic. This altered characteristic, which is the average over one period of the pulse width modulated signal, is usually determined by the dual input describing function technique [36, 37]. The dithers that are commonly in use are triangular wave or sawtooth, sinusoidal and square wave. Then, the altered characteristics for a two level comparator without hysteresis become a linear function, an arcsin function and a three level quantizer respectively [38, 39].

One of the primary advantages of this approach in function generation is that it has great flexibility. It is possible to
completely change the character of the function generator by changing the carrier signal. For example, the same hardware can produce the arcsin $x$ of the analog signal $x(t)$ by using a sinusoidal carrier or it can produce the $\ln x$ by using an exponential carrier signal. Another feature is that the idea can be implemented in any process or system that has a comparator, relay or bistable element. Hence, its applications are in the mechanical, electrical and electronic field as well as in the fluidics field. The method would also be useful for signal conditioning in measurement systems where the information could be processed at the transducer by rather simple electronics to perform the simple linearization or data expansion and contraction and at the same time would provide a signal which would be very easily transmitted over long distances. Information in a pulse width modulated like other binary information can be transmitted very accurately.

It is also possible to use nonlinear pulse width modulation techniques in a simple feedback system to perform nonlinear computations. This is the basis for a new concept in multifunction module mechanization. Amongst the computations that can be performed are $x^a y^b$, $y e^{ax}$, $ax + b \ln x$, etc. A list of possibilities will be given in Chapter 8. The powers and roots can be taken and multiplication and division can be executed by using simple exponential carrier signals in such feedback systems. The feedback system acts as a filter to convert the information in the system to an analog form. With this type of system it would be very easy to perform the conversion of electrical pressure measurements across an orifice flow meter into flow signals, linearize hot wire anemometer characteristics or many other kinds of instrumentation conditioning tasks. In these situations, it is desirable to have not
only a nonintegral power or root of these signals generated, but it is also desirable to tailor the exponent to the particular instrumentation situation. Lastly, it appears possible to build an analog multifunction module which is capable of performing many different kinds of nonlinear computations. With this module it is only necessary to change the waveform or amplitude of the carrier signals to change the characteristics of the module.

In the next chapters, the notions of nonlinear pulse width modulation and its function generation feature will be examined. It will be shown how this can be applied in a feedback module to perform nonlinear analog computations.
Chapter 5. General Nonlinear Modulation

Before dealing with the nonlinear pulse width modulation topic, the more general case of nonlinear modulation will be discussed. General nonlinear modulation is done by feeding the difference between a high frequency carrier signal $D(t)$ (period $T$) and a low frequency modulating signal $x(t)$ into a nonlinear element (Fig. 5.1). It is seen that in the case that the nonlinear element is a comparator, nonlinear pulse width modulation is accomplished. This will be dealt with later. If the output signal $e_o(t)$ is filtered to suppress the carrier signal, the average value of the output is given by:

$$\text{Ave} = \frac{1}{T} \int_0^T e_o(t) \, dt$$

$$= \frac{1}{T} \int_0^T f[e_i(t)] \, dt$$

$$= \frac{1}{T} \int_0^T f[x(t) - D(t)] \, dt$$

It is assumed that $x(t)$ is essentially constant over the period $T$ of the carrier signal, so:

$$x(t) \approx X \quad \text{where } X \text{ is constant}$$

Then, the above equation becomes:

$$\text{Ave} = \frac{1}{T} \int_0^T f[X - D(t)] \, dt \quad (5.1)$$

Since in equation 5.1 both the characteristics of the carrier signal and the nonlinear element can be chosen, the number of possibilities are numerous. A few illustrative ones will be given.
Fig. 5.1 General nonlinear modulation
Example 5.1 Squarer and sawtooth carrier signal

nonlinear element \( e_o(t) = e_1^2(t) \)
carrier signal \( D(t) = -C+2Ct/T \)

Equation 5.1 becomes:

\[
Ave = \frac{1}{T} \int_0^T (X+C-2Ct/T)^2 dt
\]
\[
= X^2 + 1/3C^2
\]

This result is not very spectacular

Example 5.2 Square rooter and sawtooth signal

nonlinear element \( e_o(t) = e_1^{1/2}(t) \)
carrier signal \( D(t) = -Ct/T \)

Equation 5.1 becomes:

\[
Ave = \frac{1}{T} \int_0^T (X+Ct/T)^{1/2} dt
\]
\[
Ave = \frac{2}{3C} [(X+C)^{3/2} - X^{3/2}]
\]

By choosing different carrier signals and different nonlinear elements, many different functions can be realized. The main advantage of this approach is that only one nonlinear element is needed to obtain a rather complicated nonlinear function, which would otherwise require several linear and nonlinear elements. However, the functions obtained in this way are not of the type that are frequently in need of being generated.
Nonlinear pulse width modulation is accomplished by feeding the difference between a modulating signal $x(t)$ and a high frequency periodic carrier signal whose equation over one period $T$ is $D(t)$ into a two level output comparator with levels $A$ and $B$. For simplicity, it will be assumed that all comparators switch when $x(t) - D(t) = 0$. If $A$ and $B$ have a positive value, then the output of the modulator would appear as in Fig. 6.1C. Suppose that in the period $T$ of the carrier the comparator will only switch twice, then:

$$D(t) = x(t) \text{ for } t = t_1 \text{ and } t = t_2$$

$t_1$ - time instant at which switching from level $A$ to level $B$ occurs.

t_2$ - time instant at which switching from level $B$ to level $A$ occurs.

For the comparator given, the dc component of the output is:

$$x_{dc} = A + (B-A) \frac{t_2 - t_1}{T} \quad (6.1)$$

For many carrier waveforms $t_1$ and $t_2$ can be obtained in closed form by operating on $x(t)$ with a function which is the inverse of the operation represented by $D(t)$, where the inverse of the operator, $D^{-1}(.)$:

$$D^{-1}[D(t_i)] = t_i \quad i = 1, 2 \quad (6.2)$$

$D^{-1}$ inverse operation

For an at zero switching comparator, the switching time instants $t_1$ are given by:

$$x(t_i) = D(t_i) \text{ so } t_i = D^{-1}[x(t_i)]$$
Fig. 6.1 Nonlinear PWM
This makes it possible to generate an output $x_{dc}$ which is a nonlinear function of the input $x(t)$.

**Example 6.1** Constant signal with sawtooth carrier signal

$$x(t) = X \text{ where } X \text{ is constant}$$

$$D(t) = C - 2Ct/T \text{ (sawtooth carrier signal, Figure 6.2), } 0 < t < T$$

We need to find the switching time instants to calculate the dc level of $x(t)$. One switching occurs at $t=0, T, 2T, \ldots$ So, $t_1=0$. The other switching time instant must be found from the inverse of the carrier signal. This is done by solving the $D(t)=C-2Ct/T$ function for $t$.

That is:

$$t = \frac{T}{2} - \frac{T}{2C} D(t) \quad 0 < t < T$$

so

$$D^{-1}[D(t)] = \frac{T}{2} - \frac{T}{2C} D(T)$$

The switching time instant $t_2$ occurs when $D(t) = x(t) = X$, so:

$$t_2 = \frac{T}{2} - \frac{T}{2C} X$$

Using equation 6.1, the dc component of the output will be

$$x_{dc} = A + \frac{1}{2} (B-A)(1-X/C)$$

If $A = -B$, so that the comparator is symmetrical, the above equation becomes:

$$x_{dc} = AX/C$$

in this case, the dc component of the output is proportional to the magnitude of the input. This is known as linear pulse width modulation.

**Example 6.2** Constant signal with sinusoidal carrier signal

$$x(t) = X \text{ where } X \text{ is a constant}$$

$$D(t) = \sin(2\pi/T)t$$

The inverse of the carrier signal becomes:
Fig. 6.2 Sawtooth carrier signal
\[ t = \frac{T}{2\pi} \arcsin D(t) \]

so

\[ D^{-1}[D(t)] = \frac{T}{2\pi} \arcsin D(t) \]

The switching time instants \( t_i \) occur when \( D(t) = x(t) = X \), so

\[ t_1 = \frac{T}{2\pi} \arcsin X, \quad t_2 = \frac{T}{2} - \frac{T}{2\pi} \arcsin X \quad 0 < t < T \]

Using equation 6.1, the dc component of the output will be:

\[ x'_{dc} = A + \frac{1}{2} (B-A) - \frac{B-A}{\pi} \arcsin X \]

If \( A = -B \), so that the comparator is symmetrical, the above equation becomes:

\[ x'_{dc} = \frac{2A}{\pi} \arcsin X \]

Example 6.3 Constant signal with exponential carrier signal

\[ x(t) = X \quad \text{where } X \text{ is a constant} \]

\[ D(t) = Ke^{-at} \quad \text{(exponential carrier signal, Fig. 6.3)} \]

The switching time instants \( t_i \) are:

\[ t_1 = 0 \]
\[ t_2 = -\frac{1}{a} \ln \frac{X}{K} \]

Using equation 6.1, the dc component of the output will be:

\[ x'_{dc} = A + (B-A)(-\frac{1}{aT} \ln \frac{X}{K}) \]

By choosing \( A=0 \), the above equation becomes:

\[ x'_{dc} = \frac{-B}{aT} \ln \frac{X}{K} \]

As was mentioned before, this type of logarithmic converter gives very accurate results [24].

Example 6.4 Constant signal with second order saturation function

\[ x(t) = X \quad \text{where } X \text{ is a constant} \]

\[ D(t) = 8t/T - 16(t/T)^2 \quad 0 < t/T < 1/2 \]
Fig. 6.3 Exponential carrier signal
-8(t/T-1/2) + 16(t/T-1/2)^2 \quad \frac{1}{2} < t/T < 1

Saturation functions occur in bang-bang or relay control systems that are in a limit cycle [41]. If the output of the relay is fed into an integrator, the triangular wave output is a first order saturation function. Pulse width modulation accomplished with a comparator with hysteresis and an integrator in the feedback loop generates a first order saturation function (Fig. 3.10). To obtain a second order saturation function the feedback loop must have two integrators (Fig. 6.4). The output of the double integrator yields the second order saturation function (Fig. 6.5). It is possible to continue this process and to construct third, fourth, etc., order saturation functions. However, they do not appear to be interesting. Only the limit case of the \( n^{th} \) order saturation function, where \( n \to \infty \), is of importance, since it yields the sinusoidal function (Example 6.2).

For the second order saturation function, the switching time instants are given by:
\[
X = 8t/T - 16(t/T)^2 \quad \text{for } X > 0
\]
\[
(t/T)_{1,2} = 1/4 \pm \frac{1}{4}(1-X)^{1/2}
\]
Using equation 6.1, the dc component of the output will be:
\[
x_{dc}' = A + \frac{1}{2}(B-A)(1-X)^{1/2}
\]
If \( A=0 \), then
\[
x_{dc}' = \frac{1}{2}B(1-X)^{1/2}
\]
(6.3)

For \( X < 0 \), an equivalent expression for the dc component is given by:
\[
x_{dc}' = B - \frac{1}{2}B(1+X)^{1/2}
\]
These results are interesting since it opens the possibility for square root extraction. If the following substitution is made:
Fig. 6.4 Generation of second order saturation function carrier signal $o(t)$
Fig. 6.5 Second order saturation function
then equation 6.3 will yield:

\[ x_{dc} = \frac{1}{2}Bx^{1/2} \]

Thus, the square root operation is obtained. Another application of equation 6.3 can be found in the simulation of hydraulic valve flow equations. Mathematical models of these valves involve terms like \((1-x)^{1/2}\).

It would be useful to obtain a pulse width which is proportional to the inverse of the modulating signal. In this way, it would be possible to perform the division operation by the pulse width modulation method. For this, it is necessary to generate the time function \(1/t\). One author approximated this time function by [27]:

\[ \frac{1}{t} \approx k_0 + k_1 e^{-\alpha_1 t} + k_2 e^{-\alpha_2 t} \]

The negative exponentials can be obtained by discharging capacitors through resistors. The coefficients \(k_0, k_1, k_2\) and the parameters \(\alpha_1\) and \(\alpha_2\) can be found by a least squares approximation to \(1/t\). The method has an accuracy of about 2% and, of course, a limited range. Another author [43] reported the construction of a multivibrator which produced pulse widths that were inversely proportional to the input. It was used to perform the division operation and it had an accuracy of 2% of full scale.

Generally, it is possible to generate an output dc level that is equal to the inverse of the function representing the carrier signal over one cycle. The number of functions that can be generated this way is only limited by the carrier signal forms available. The results are exact for input signals \(x(t)\) that are constant. However, for time varying input signals the results will only be accurate if the frequency of the carrier signal is considerably higher than the
highest frequency component present in the input signal. This will be discussed in more detail in chapter II.
Chapter 7. Demodulation

The nonlinear pulse width modulation method yields a two level signal whose average value is a nonlinear function of the input signal. This average value is usually obtained by one of the following demodulation methods:

1) filtering of the pulse width modulated signal
2) integration of the pulse width modulated signal over one cycle
3) pulse width modulation in the feedback loop of an integrator

The three demodulation methods will be investigated analytically in this chapter. The evaluation will consider output ripple and step response. A frequency response study will not be done here; it can be done by using Bessel functions which result from the analysis [45].

A. Filtering of the pulse width modulated signal

The design of a filter for demodulating pulse width modulated signals is a compromise between the speed of response and the left-over ripple. In turn, the speed of response and the left-over ripple are determined by the bandwidth and the order of the filter. Increasing the bandwidth will increase the speed of response, but will also increase the left-over ripple, whereas increasing the order of the filter will reduce the left-over ripple, but increase the delay in the response.

In order to develop expressions for the ripple and the response, the fourier series of a pulse width modulated signal will be considered (Fig. 7.1). The fourier series representation of this signal is given by:

\[ \text{PWM}(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos \frac{2\pi n t}{T} + b_n \sin \frac{2\pi n t}{T} \]
Fig. 7.1 PWM signal
\[ a_0 = B + (A-B)t/T \quad (7.1A) \]
\[ a_n = \frac{A-B}{\pi n} \sin \frac{2\pi n}{T} t \quad (7.1B) \]
\[ b_n = \frac{A-B}{\pi n} (1-\cos \frac{2\pi n}{T} t) \quad (7.1C) \]

In the above equation the term \( a_0 \) is the wanted component and the ripple is represented by the sum of sine and cosine waves. If the pulse width modulated signal is fed into a filter, the response of the "dc" component of the output to a change in the dc component \( a_0 \) is given by the convolution integral:

\[ o_{dc}(t) = \int_0^T a_0(t) h(t-\tau) d\tau \]

\( o_{dc}(t) \) - dc component of the output of the filter
\( a_0(t) \) - dc component of the pulse width modulated signal
\( h(t) \) - impulse response of the filter

\( h(t) = 0 \quad t < 0 \) causal filter

With the above equation, the filter can be designed so that the response will be satisfactory.

The energy \( E \) in the ripple can be computed by using Parseval's theorem, which states

\[ E = \frac{1}{T} \int_0^T \text{PWM}^2(t) dt = a_0^2 + \frac{1}{2} \sum_{n=1}^{\infty} a_n^2 + b_n^2 \]

The last term on the right hand side of the above equation gives the energy in the unwanted components of the pulse width modulated signal that is fed into the filter, namely:
\[
E_I = \frac{1}{2} \sum_{n=1}^{\infty} a_n^2 + b_n^2 \quad E_I - \text{energy in the unwanted components of the pulse width modulated signal}
\]

Substituting the values for \(a_n\) and \(b_n\) from equations 7.1B and 7.1C yields:

\[
E_I = \sum_{n=1}^{\infty} \left( \frac{1}{\pi n} \right)^2 (A-B)^2 (1-\cos \frac{2\pi n}{T} t_1) 
\]

or equivalently:

\[
E_I = 2 \sum_{n=1}^{\infty} \left( \frac{1}{\pi n} \right)^2 (A-B)^2 \sin^2 \left( \frac{2\pi n}{T} t_1 \right)
\]

Suppose a filter of the order \(k\) is chosen and that the following approximation of the filter is valid for its frequency response above the cut-off frequency:

\[
H(jw) = \frac{1}{(j\omega T)^k}
\]

then for each harmonic, the value of this frequency response function is:

\[
H(\frac{2\pi n}{T}) = \frac{1}{(\frac{2\pi n}{T})^k} \quad n = 1,2,3,\ldots
\]

\(H(.)\) - fourier transform of the impulse response of the filter, \(h(t)\)

Then, the energy of the ripple of the output signal is given by:

\[
E_o = \sum_{n=1}^{\infty} \frac{1}{(\frac{2\pi n}{T})^{2k}} \left( \frac{1}{\pi n} \right)^2 (A-B)^2 (1-\cos \frac{2\pi n}{T} t_1) 
\]

\[
E_o = \sum_{n=1}^{\infty} \frac{(A-B)^2}{\pi(\frac{2\pi n}{T})^{2k+2}} \frac{1}{n^{2k+2}} (1-\cos \frac{2\pi n T}{T} t_1) 
\]

\[
E_o = \sum_{n=1}^{\infty} \frac{(A-B)^2}{\pi(\frac{2\pi n}{T})^{2k+2}} \frac{1}{n^{2k+2}} (1-\cos \frac{2\pi n T}{T} t_1) 
\]

(7.3)

The above series converges and the sum can be computed. This yields

(See Appendix I):
\[ E_o = \frac{2(A-B)^2}{(\frac{T}{2})^{2k+2}} \left[ |B_{2k+2}| - (-1)^k B_{2k+2}\left( \frac{1}{T} \right) \right] \]

\[ B_{2k+2} \] is the \( 2k+2 \) Bernoulli number

\[ B_{2k+2}\left( \frac{1}{T} \right) \] is the \( 2k+2 \) Bernoulli polynomial

With the above equation it is possible to compute the energy of the ripple of the output signal. However, it can be proven (see Appendix I) that the worst case, that is maximal energy of the ripple, occurs when \( t_j/T = 1/2 \). In this case the pulse width modulated signal consists of a dc component plus a square wave. This maximum energy of the ripple of the output signal of the filter is given by (see Appendix I):

\[ E_o^{\max} = \frac{4(A-B)^2}{(\frac{T}{2})^{2k+2}} |B_{2k+2}[1-2^{-(2k+2)}] | \]

For \( t_j/T = 1/2 \), the energy of the ripple of the input signal of the filter is given by:

\[ E_i^{\max} = \frac{2(A-B)^2}{\pi^2} \sum_{n=1,3,5} \frac{(-1)^2}{n^2} \]

\[ E_i^{\max} = 1/4(A-B)^2 \]

Normalization of the maximum output ripple with the maximum input ripple yields:

\[ \frac{E_o^{\max}}{E_i^{\max}} = \frac{16}{(\frac{T}{2})^{2k+2}} |B_{2k+2}[1-2^{-(2k+2)}] | \]
For \( k = 0, 1, 2 \) and 3, the Bernoulli numbers \( B_{2k+2} \) and the ratios
\[
\frac{E_0}{E_i}_{\text{max}} \quad \text{are:}
\]

<table>
<thead>
<tr>
<th>( k )</th>
<th>( B_k )</th>
<th>( \frac{E_0}{E_i}_{\text{max}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1/6</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>-1/30</td>
<td>2.08x10^{-2}/(\tau/T)^2</td>
</tr>
<tr>
<td>2</td>
<td>1/42</td>
<td>1.56x10^{-3}/(\tau/T)^4</td>
</tr>
<tr>
<td>3</td>
<td>-1/30</td>
<td>3.95x10^{-6}/(\tau/T)^6</td>
</tr>
</tbody>
</table>

From the above equation for the normalized energy of the ripple of the output signal it can be seen that the ratio of the time constant of the filter and the period of the carrier signal should be made large. Moreover, increasing the order of the filter \( k \) also decreases the energy of the ripple significantly, however, this also increases the time delay in the transient response.

The pulse width modulation of time varying modulating signals necessarily leads to some distortion. It was shown [50] that in the spectrum of pulse width modulated sinusoidal signals sidebands will appear even at very low frequency. This makes it impossible to recover a time varying modulating signal without distortion, even if an ideal filter with no phase lag and perfectly sharp cut-off would be used. The significance of the sidebands depends on the ratio of carrier frequency and modulating signal frequency.

B. Integration of the pulse width modulated signal over one cycle

The average value of the pulse width modulated signal can be obtained by integrating the pulse width modulated signal over one cycle [3, 46] (Fig. 7.2) After the integrator has computed the average value over one cycle, the output \( E_0 \) is fed into a sample and hold element.
Fig. 7.2 Demodulation by integration over one cycle
Then, the integrator is reset and an integration over the next period is performed. The method has the following advantages and disadvantages:

**Advantages:**
- The response of the demodulator to a change in input signal is almost instantaneous. There will only be a maximum delay of one period.
- The pulses for the reset and sample and hold do not have to be in phase with the carrier signal. However, they must have the same period.
- There is no steady state ripple.

**Disadvantages:**
- The reset and sample and hold mode require some time, which causes the circuit to be more complex if one wants to preserve the requirement that the integration has to be exactly over one period.
- Variations in the integration time constant $RC$ (temperature, etc.) will give an error in the output.

C. Pulse width modulation in the feedback loop of an integrator

A better method to perform demodulation is to feed the pulse width modulated signal into an integrator which has a pulse width modulator in the feedback loop (Fig. 7.3). The integrator and its feedback loop are the demodulator. If the comparator and carrier signal of the demodulator are different than those of the modulator, then the system performs nonlinear computation. This will be covered in the next chapter. If the comparator and the carrier signal of the modulator are the same as those of the demodulator, then the input signal $x(t)$ can be recovered at the output. With the input signal $x(t)$
Fig. 7.3 Recovering of PWM signals

note: all amplifiers invert sign

\[ C \rightarrow K \rightarrow \text{-CK} \]

\[ C \rightarrow \tau_i \rightarrow \text{-Ct/}\tau_i \]
a constant, the feedback loop will force the sum of the dc components of the input signals to the Integrator to zero. This can only happen if $x'(t)$ is equal to $w'(t)$ and this is only true when $x(t)$ is equal to $w(t)$. However, there will be a slight distortion for time varying input signals. The carrier signals can easily be made equal by obtaining both of them from the same source. However, the levels $A$ and $B$ of the two comparators will not exactly be equal, which results in an error. One way to minimize this error is to match the comparators. However, if the levels of the comparator are adapted for use with logic elements, a pulse width subtraction technique described in [47] provides more accurate demodulation. The demodulator and waveforms illustrating its operation are shown in Figures 7.4 and 7.5. If the signal $x$ is less than the signal $w$ the nor element $NOR_1$ will have a pulse output, causing the switch $S_1$ to close and the output $W$ to decrease. If signal $x$ is larger than signal $w$ the nor element $NOR_2$ will have a pulse output, causing switch $S_2$ to close and the output $w$ to increase. If $w = x$, neither nor element will have a pulse output and the output $w$ will remain constant. The main advantage of this circuit is that the accuracy does not depend on the levels of the comparators but uses the more favorable properties of nor elements and switches instead. The advantages and disadvantages of the feedback pulse width demodulator are:

**advantages:**

- No timing circuit for reset and sample and hold is required as was the case in the demodulation by integration over one cycle.
- The exact value of the integration time constant $RC$ is not important, so variations in $R$ and $C$ have no longer effect on the steady state $w$ versus $x$ relation.
Fig. 7.4 Improved feedback pulse width demodulator [47]
Fig. 7.5 Pulse width subtraction
There is no steady state ripple.

**Disadvantages:** The response of the demodulator is not instantaneous. It depends on the integration time constant $RC$. If the integration time constant is small, the response will be fast. Limitations on the value of the integration time constant are caused by operational amplifier noise, offset and the stability of the feedback system.

The above circuits (Figs. 7.3 and 7.4) will form the basis for a new multifunction generation concept. It was shown before that nonlinear functions could be generated by the nonlinear pulse width modulation method. In the following chapters, it will be shown that by choosing different carrier signals for the modulator and demodulator, nonlinear computation can be performed.
Chapter 8. Computation With Nonlinear Pulse Width Modulation

In the previous chapters it was shown that with different forms of the carrier signal different functions can be realized (Examples 6.1, 6.2, 6.3, 6.4). By choosing different carrier signals for the modulator and demodulator and different comparator output levels, various nonlinear computations can be performed. In the more general case there are two input signals $x(t)$ and $y(t)$ with carrier signals $D_x(t)$ and $D_y(t)$ and one output signal $w(t)$ with carrier signal $D_w(t)$. The comparator levels are $A_x, B_x, A_y, B_y, A_w, B_w$ respectively (Fig. 8.1). In the steady state, the integrator will force the dc component of its input to zero, so the following equation holds:

$$w_{dc} = x_{dc} + y_{dc}$$

where $w_{dc}$ is the dc component of the output of the pulse width modulator for $j(t)$ ($j = x, y, w$).

Computing the dc components of the signals $x'(t)$, $y'(t)$ and $w'(t)$, using the methods described in Chapter 6, Eq. 6.1 and substituting these in the above equation yields:

$$A_w + (B_w - A_w) \frac{t_{2w} - t_{1w}}{T} = A_x + (B_x - A_x) \frac{t_{2x} - t_{1x}}{T} +$$

$$A_y + (B_y - A_y) \frac{t_{2y} - t_{1y}}{T}$$

$$t_{ij} = D^{-1}_{j}[j(t)] \quad i = 1, 2 \quad j = x, y, w$$

It is shown in examples 6.1, 6.2, 6.3 and 6.4 that by choosing different forms for the carrier signal $D_j(t)$ different functions can be realized, so equation 8.1 describes different nonlinear computations. This is illustrated in the following examples.
Fig. 8.1 System which performs nonlinear computation
Example 8.1  Raising signals to arbitrary exponents by use of exponential carrier signals

Suppose: \( D_j(t) = K_j e^{-a_j t} \), \( j = x, y, w \) (Fig. 6.3)

\[ t_{1j} = 0 \quad j = x, y, w \]

\[ t_{2j} = -\frac{1}{a_j} \ln \frac{j(t)}{K_j} \quad j = x, y, w \]

Equation 8.1 becomes:

\[ A_w + (B_w - A_w) \frac{-1}{a_w T} \ln \frac{w(t)}{K_w} = A_x + (B_x - A_x) \frac{-1}{a_x T} \ln \frac{x(t)}{K_x} + A_y + (B_y - A_y) \frac{-1}{a_y T} \ln \frac{y(t)}{K_y} \]  (8.2)

By choosing suitable values for the parameters it is possible to eliminate unwanted components and simplify the above equation. If we choose:

1) \( A_j = 0 \quad j = x, y, w \)

2) \( K_j = 1 \quad j = x, y, w \)

then equation 8.2 becomes:

\[ \frac{-B_w}{a_w T} \ln w(t) = \frac{-B_x}{a_x T} \ln x(t) + \frac{-B_y}{a_y T} \ln y(t) \]

which when solved for \( w(t) \) gives:

\[ w(t) = x(t)^{\frac{B_x}{B_w}} y(t)^{\frac{B_y}{B_w}} \]  (8.3)
Raising signal to arbitrary exponents can be done this way by adjusting either the levels of the comparators or the inverse of the time constants.

Example 8.2 Exponential and sinusoidal carrier signal

Suppose: \( D_w(t) = Ke^{-at} \)
\( D_x(t) = K \sin \frac{2\pi}{T} t \)

no input signal \( y(t) \)

\( t_{1w} = 0 \)
\( t_{2w} = -\frac{1}{a} \ln \frac{w(t)}{K} \)
\( t_{1x} = \frac{T}{2\pi} \arcsin \frac{x(t)}{K} \)
\( t_{2x} = \frac{T}{2} - \frac{T}{2\pi} \arcsin \frac{x(t)}{K} \)  (See example 6.2)

Equation 8.1 becomes:

\[
A_w + (B_w - A_w) \left( \frac{-1}{at} \ln \frac{w(t)}{K} \right) = A_x + 1/2 (B_x - A_x) \frac{B_x - A_x}{\pi} \arcsin \frac{x(t)}{K}
\]

If we choose:

1) \( A_w = 0 \)
2) \( B_x = -A_x \)

then the above equation becomes:

\[
-\frac{B_w}{at} \ln \frac{w(t)}{K} = \frac{2A_x}{\pi} \arcsin \frac{x(t)}{K}
\]

or:

\[
-\frac{2aA_x}{\pi B_x} \arcsin \frac{x(t)}{K} = \frac{2aA_x}{\pi B_x} \arcsin \frac{x(t)}{K}
\]

\[w(t) = Ke^{-at}\]
It can be seen that different functions can easily be generated by using different carrier signals. As an example, the functions that can be generated with a sawtooth, sinusoidal and exponential carrier signal are explored. For these carrier signals the dc components of the pulse width modulated signals are shown (Table 8.1). It is assumed that the input signals are constant, which is denoted by a capital letter. Also, in this table the following comparator levels and carrier wave forms are assumed:

<table>
<thead>
<tr>
<th>carrier signal</th>
<th>equation $0 &lt; t &lt; T$</th>
<th>comparator levels</th>
<th>dc component of comparator output</th>
</tr>
</thead>
<tbody>
<tr>
<td>sawtooth (sawt)</td>
<td>$1 - 2t/T$</td>
<td>$B = -A$</td>
<td>$AX$</td>
</tr>
<tr>
<td>sinusoidal (sin)</td>
<td>$\sin \frac{2\pi t}{T}$</td>
<td>$B = -A$</td>
<td>$\frac{2A}{\pi} \arcsin X$</td>
</tr>
<tr>
<td>exponential (exp)</td>
<td>$e^{-\alpha t}$</td>
<td>$A = 0$</td>
<td>$\frac{-B}{\alpha T} \ln X$</td>
</tr>
</tbody>
</table>

Table 8.1 dc components of pulse width modulated signals

Combination of these carrier signals yield the nonlinear computations that are shown in Table 8.2.
Table 8.2. Realizable Nonlinear Functions

<table>
<thead>
<tr>
<th>$D_x(t)$</th>
<th>$D_y(t)$</th>
<th>$D_w(t)$</th>
<th>$W$ $(\alpha = \frac{A\alpha T}{B})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>sawt</td>
<td>sawt</td>
<td>sawt</td>
<td>$X+Y$</td>
</tr>
<tr>
<td>sawt</td>
<td>sawt</td>
<td>sin</td>
<td>$\sin \left(\frac{\pi}{2}(X+Y)\right)$</td>
</tr>
<tr>
<td>sawt</td>
<td>sawt</td>
<td>exp</td>
<td>$e^{-\alpha(X+Y)}$</td>
</tr>
<tr>
<td>sawt</td>
<td>sin</td>
<td>saw</td>
<td>$X + \frac{2}{\pi} \arcsin Y$</td>
</tr>
<tr>
<td>sawt</td>
<td>sin</td>
<td>sin</td>
<td>$\sin \left(\frac{\pi}{2}X + \arcsin Y\right)$</td>
</tr>
<tr>
<td>sawt</td>
<td>sin</td>
<td>exp</td>
<td>$e^{-\alpha(X + \frac{2}{\pi} \arcsin Y)}$</td>
</tr>
<tr>
<td>sawt</td>
<td>exp</td>
<td>sawt</td>
<td>$X - \frac{1}{a} \ln Y$</td>
</tr>
<tr>
<td>sawt</td>
<td>exp</td>
<td>sin</td>
<td>$\sin \left(\frac{\pi}{2}X - \frac{\pi}{2a} \ln Y\right)$</td>
</tr>
<tr>
<td>sawt</td>
<td>exp</td>
<td>exp</td>
<td>$e^{-\alpha_X \frac{\alpha}{\alpha}}$</td>
</tr>
<tr>
<td>sin</td>
<td>sin</td>
<td>saw</td>
<td>$\frac{2}{\pi} \arcsin X + \arcsin Y$</td>
</tr>
<tr>
<td>sin</td>
<td>sin</td>
<td>sin</td>
<td>$\sin(\arcsin X + \arcsin Y)$</td>
</tr>
<tr>
<td>sin</td>
<td>sin</td>
<td>exp</td>
<td>$e^{-2\alpha X \arcsin Y}$</td>
</tr>
<tr>
<td>sin</td>
<td>exp</td>
<td>saw</td>
<td>$\frac{2}{\pi} \arcsin X - \frac{1}{a} \ln Y$</td>
</tr>
<tr>
<td>sin</td>
<td>exp</td>
<td>sin</td>
<td>$\sin(\arcsin X - \frac{1}{2a} \ln Y)$</td>
</tr>
<tr>
<td>sin</td>
<td>exp</td>
<td>exp</td>
<td>$e^{-2\alpha X \arcsin Y \frac{\alpha}{\alpha_X}}$</td>
</tr>
<tr>
<td>exp</td>
<td>exp</td>
<td>saw</td>
<td>$-\frac{1}{a} (\ln X + \ln Y)$</td>
</tr>
<tr>
<td>exp</td>
<td>exp</td>
<td>sin</td>
<td>$\sin \left(-\frac{\pi}{2a}(\ln X + \ln Y)\right)$</td>
</tr>
<tr>
<td>exp</td>
<td>exp</td>
<td>exp</td>
<td>$\frac{\alpha}{\alpha_X} \frac{\alpha}{\alpha_Y}$</td>
</tr>
<tr>
<td>exp</td>
<td>exp</td>
<td>exp</td>
<td>$X_w \frac{\alpha}{\alpha_Y} Y_w$</td>
</tr>
<tr>
<td>exp</td>
<td>exp</td>
<td>exp</td>
<td>$X_{\frac{\alpha}{\alpha_X}} Y_{\frac{\alpha}{\alpha_Y}}$</td>
</tr>
</tbody>
</table>
The above analysis was based on computing the dc components of pulse width modulated signals. For this analysis, the form, phase and period of the carrier signals were not important. However, it will appear that for effective demodulation the form, phase and period are of considerable importance. An effective demodulation method is a demodulation method in which there is no steady state ripple in the output signal for constant input signals. For this it is necessary that the pulses out of the modulator for the output signal are identical to the pulses out of the modulator for the input signal. Consequently, this limits effective demodulation to only single Input, single output systems. Moreover, the pulses can only be identical if the levels of the two comparators are identical and the switching time instants \( t_{1x} \) are equal to \( t_{1w} \) (\( i = 1,2 \)) in the steady state. Furthermore, the period and the phase of the carrier signals have to be identical too. If the levels of the comparator are adapted for use with logic elements, then the NOR pulse width subtraction method can be used (Fig. 7.4). This has the advantage that the levels of the comparators do not exactly have to be the same. In the case that for both input signal and output signal the comparator levels and carrier signals are identical, effective demodulation occurs (Fig. 7.3) However, no nonlinear computation is performed in this case \( w(t) = \dot{x}(t) \). It is still possible to perform nonlinear computations by carefully selecting the carrier signals, which is illustrated in the following examples.

**Example 8.3** Raising a signal to an arbitrary power by use of an exponential carrier signal (See also example 8.1)
If exponential carrier signals are applied in a single input, single output system, the nonlinear computation that is performed is given by (See equation 8.3):

\[
\frac{B_x}{B_w} \frac{\alpha_w}{\alpha_x} \quad w(t) = x(t)
\]

Since the levels of the comparators have to be chosen the same, the adjustment of the exponent is done by selecting the ratio of the inverse of the time constants \( \alpha_w \) and \( \alpha_x \). This case is illustrated (Fig. 8.2). The input signal \( x(t) \) has a carrier signal with inverse time constant \( \alpha_x \) and the output signal \( w(t) \) has a carrier signal with inverse time constant \( \alpha_w \). Since \( t_{l_x} = t_{l_w} = 0 \), the only way that the pulses can be identical in the steady state is when \( t_{2x} = t_{2w} \) (dashed line). These time instants are given by:

\[
t_{2x} = -\frac{1}{\alpha_x} \ln x(t) \\
t_{2w} = -\frac{1}{\alpha_w} \ln w(t)
\]

so:

\[
w(t) = x(t) \frac{\alpha_w}{\alpha_x}
\]

In this case there is no ripple in the output signal in the steady state, since the input of the integrator is zero over the whole period.

**Example 8.4** Triangular wave and sinusoidal carrier signal (Fig. 8.3)

The input has a triangular wave and the output has a sinusoidal carrier signal. Again, if the levels of the comparators are identical, the integrator input can only be zero if the switching time instants \( t_{l_x} \) and \( t_{l_w} \) \((i = 1,2)\) are equal. If the levels \( B_x = B_w = -A \) and \( A_x = A_w = A \), then the following nonlinear computation is performed (See also examples 6.1 and 6.2):
Fig. 8.2 Raising a signal to a power.
Fig. 8.3 Sinusoidal and triangular wave carrier signal
\[ W = \sin \frac{\pi}{2} X \]

Effective demodulation cannot be performed in all cases. An example is the computation with an exponential and a sinusoidal carrier signal. Because \( t_1 = 0 \) for the exponential carrier, but \( \sin^{-1} \) for the sinusoidal carrier, the pulses cannot be identical in the steady state.

So far, it was assumed that the carrier signals had the same period and that there was no phase difference. The influence of a phase difference between two identical carrier signals will result in pulses that are shifted in time with respect to each other (Fig. 8.4). This causes the input of the integrator not to be zero over the whole period even when its dc level is zero. Consequently, the output signal \( w(t) \) will have a ripple at carrier frequency. In the case that the carrier signals have a different period, the leading and trailing edges of the pulses again will not coincide, so a similar kind of ripple at carrier frequency as in the previous case will occur.

The above analysis showed that it is important that both carrier signals have the same period, phase and both have exponential, sinusoidal, etc. forms. The period and phase can usually be made equal by obtaining them both from the same source. However, the form of the carrier signal depends on the function that has to be generated and so it cannot always be adapted as to not give any ripple in the steady state.

This chapter gave a general overview of the nonlinear computations that can be done with the nonlinear pulse width modulation method. It is impossible to treat all the possibilities that were mentioned in greater detail. However, since the nonlinear computation \( w = x^a y^b \) occurs most frequently, a detailed analysis of the system that performs this computation is given. For other nonlinear
Fig. 8.4 Ripple due to phase difference in carrier signals
computations the same method of analysis can be applied.
Chapter 9. Exponentiation of Signals

A more extensive study of the exponentiation of signals \( w = x^a \cdot y^b \) will be undertaken. For this computation, an exponential carrier signal is needed (See example 8.3). At first, it makes sense to generate an exponential carrier signal by feeding pulses into a first order filter. Ideally, the pulses should be of infinite height and zero width. This is not possible. The finite width of the pulses causes an error which is far too big, about 1%. A better method is to feed a square wave into an RC network (Fig. 9.1). The response of the circuit to a square wave signal is the exponential carrier signal. A slightly better method is to feed the square wave into a first order filter and to subtract the output from the input square wave (Fig. 9.2). The circuit has the transfer function:

\[
H(s) = \frac{-s}{s + a}
\]

This method has the advantage that it eliminates the influence of any dc component in the square wave. With this form of the carrier signal, it is possible to perform nonlinear computations with negative signals too. Using the above notions, the carrier signal over one cycle, \( 0 < t < T \) becomes:

\[
D_j(t) = \begin{cases} 
K_j e^{-at} & 0 < t < T/2 \\
-K_j e^{(-a)(t-T/2)} & T/2 < t < T 
\end{cases}
\]

\( j = x, w \)

If the difference between the signal \( j(t) \) and its carrier signal \( D_j(t) \) is fed into a comparator, then the switching time instants are:
Fig. 9.1 Generation of an exponential carrier signal
Fig. 9.2 Generation of an exponential carrier signal
\[ t_{1j} = 0 \quad j = x,w \quad x > 0, w > 0 \quad 0 < t < T/2 \]
\[ t_{2j} = -\frac{1}{\alpha_j} \ln \frac{j(t)}{K_j} \quad (9.1) \]

\[ t_{1j} = 0 \quad j = x,w \quad x < 0, w < 0 \quad T/2 < t < T \]
\[ t_{2j} = \frac{T}{2} - \frac{1}{\alpha_j} \ln \frac{j(t)}{K_j} \]

Applying these carrier signals in the system shown in Fig. 8.1, yields for \( x(t) > 0 \) and \( w(t) > 0 \) the following nonlinear computation (See equation 8.1)

\[ A_w + (B_w - A_w) \frac{-1}{\alpha_w} \ln \frac{w(t)}{K_w} = A_x + (B_x - A_x) \frac{-1}{\alpha_x} \ln \frac{x(t)}{K_x} \quad (9.1A) \]

If \( A_j = 0 \), then

\[ \frac{B_w}{\alpha_w} \ln \frac{w(t)}{K_w} = \frac{B_x}{\alpha_x} \ln \frac{x(t)}{K_x} \]

or

\[ w(t) = K_w \left[ \frac{x(t)}{K_x} \right]^{\frac{\alpha_w}{\alpha_x}} \quad (9.2) \]

A similar derivation for \( x(t) < 0 \) and \( w(t) < 0 \) yields the result:

\[ w(t) = -K_w \left[ \frac{x(t)}{K_x} \right]^{\frac{\alpha_w}{\alpha_x}} \quad (9.3) \]

It was mentioned before that if the amplitude of the square wave and, consequently, the amplitude of the exponential carrier signal was made unity (\( K_x = K_w = 1 \)), then the input signal \( x(t) \) could be raised to the power \( \frac{\alpha_w}{\alpha_x} \). However, it is not necessary to choose unity for the magnitude of the square wave. If the magnitude of the square wave for the carrier signal for \( x(t) \) is chosen to be an input \( Y \) and, equivalently, the magnitude \( K_w \) of the carrier signal for \( w(t) \) is chosen to be an input signal
Z, then equations 9.2 and 9.3 become:

\[ W = Z \left( \frac{X}{\sqrt{Y}} \right)^{\alpha_w/\alpha_x} \quad X > 0, W > 0 \quad (9.4A) \]

\[ W = -Z \left( \frac{X}{\sqrt{Y}} \right)^{\alpha_w/\alpha_x} \quad X < 0, W < 0 \quad (9.4B) \]

These equations describe the well known multifunction operation (See Fig. 3.11). The magnitude of the square waves, \( K_x, K_w \), can be made a function of the input signals \( Y, Z \) respectively by means of the pulse height modulation method (Fig. 9.3). If the switch \( S \) is driven by a square wave, the output will be a square wave with amplitude \( Y \), which when fed into the network for generating the exponential carrier gives the required carrier signal. Similarly, pulse height modulation of the signal \( Z \) is also done and this signal is used to generate the exponential carrier signal for the demodulator. Of course, it is necessary that \( |X| < |Y| \), since otherwise no switching will occur. If this condition is fulfilled, the necessary condition \( |W| < |Z| \) will always hold for any \( \alpha_w/\alpha_x > 0 \). In the case that the signals \( Y \) and \( Z \) are time varying signals, the results will still be accurate, provided that the carrier signal frequency is much higher than the highest frequency in either signal \( Y \) or \( Z \). A complete diagram of the circuit is shown (Fig. 9.4). The switches \( S_1 \) and \( S_2 \) are synchronously driven by a square wave, so the carrier signals will have the same period and there will be no phase difference. Since the levels of the comparators are identical, there will be no steady state ripple. Moreover, it would also be possible to use the NOR pulse width subtraction method. In order for the carrier signals to be in phase, the signals \( Y \) and \( Z \) should both be either
Fig. 9.3 Pulse height modulation
Fig. 9.4 System that performs multifunction operation
positive or negative. If they are of the opposite polarity, the carrier signals will have opposite polarity too, causing the demodulation to be with ripple. The results of an investigation into the quadrants of operation are shown in table 9.1.

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
<td>+</td>
<td>( \frac{a_y}{a_x} ), no ripple</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>-</td>
<td>as above, ripple</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>+</td>
<td>as above, ripple</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>-</td>
<td>as above, no ripple</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>+</td>
<td>(-\frac{a_y}{a_x}), no ripple</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>-</td>
<td>as above, ripple</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>+</td>
<td>as above, ripple</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>as above, no ripple</td>
</tr>
</tbody>
</table>

Table 9.1 Quadrants of operation

From this table it can be seen that the sign of the output signal \( W \) follows the sign of the input \( X \). Hence, for \( \frac{a_y}{a_x} = 1 \) two quadrant multiplication and division can be accomplished. The existing multifunction modules that are mentioned in the literature only work in one quadrant. Since the polarity of the signals \( Y \) and \( Z \) do not effect the function that is generated, we will confine the signals \( Y \) and \( Z \) to have positive values only.

The levels \( B_x \) and \( B_w \) are chosen to be -1, so the integrator will have a negative feedback and the system will be stable (Fig. 9.4). This will be shown in the following. As an example we choose \( x(t) \) to be
a positive constant signal \( X \) and \( D_x(t) \) to be equal to \( D_w(t) \) \( (\alpha_w = \alpha_x, K_x = K_w) \). Then the pulses \( x'(t) \) from the modulator for \( x(t) \) will be as shown in Figs. 9.5a and 9.5b. The pulse \( x''(t) \) will be the inverse of the pulse \( x'(t) \) (Fig. 9.5c). Now, suppose \( w(t) > X \), then the pulses \( w'(t) \) from the modulator for \( w(t) \) will be shown in Figs. 9.5a and 9.5d. The sum of the two pulses \( x''(t) \) and \( w'(t) \) will be a positive pulse (Fig. 9.5e). The integrator inverts this pulse and integrates it, so the output \( w(t) \) of the integrator will be decreasing (Fig. 9.5f) and this will make pulse \( w' \) wider and this will force \( w' + x' \) and the system will go to an equilibrium. The analysis for \( w(t) < X, \alpha_w \neq \alpha_x \) and \( K_x \neq K_w \) can be done in the same way and shows that the system is stable.

For this system (Fig. 9.4) the response to a general time varying input signal will be investigated in the next chapter.
Fig. 9.5 Stability analysis
Chapter 10. Response of the System to a General Input Signal

In this chapter the response of the system (Fig. 9.4) to a general time varying input signal will be investigated. In order to simplify the analysis some assumptions are made. A reasonable assumption is to suppose that the amplitudes of the exponential carrier signals $K_x$ and $K_w$ (signals $Y$ and $Z$) are constant over the carrier signal period and that changes in their value only occur instantaneously at the beginning of each period. For the input signal $x(t)$ this assumption is not made. No distinction will be made between the amplitudes $K_x$ and $K_w$ of the carrier signals and the input signals $Y$ and $Z$ and they will be used interchangeably.

Since the input signal is time varying, it is necessary to consider the response of the system in the time interval:

$$(n-1)T < t < nT \quad n = 1, 2, 3, 4, \ldots$$

$T$ period of the carrier signal

To reduce the number of parameters a dimensionless time $t'$ is introduced:

$$t' = t/T$$

With these definitions, the switching time instants for $x(t) > 0$ and $w(t) > 0$ become:

$$t'_{1j} = n-1 \quad j=x, w \quad x > 0, \ w > 0 \quad (10.1)$$

$$t'_{2j} = n-1 - \frac{1}{a_j T} \ln \frac{j_{j(t')}}{K_j} \quad n-1 < t' < n-1/2$$

From these equations it is seen that the switching time instant $t'_{1x}$ is equal to $t'_{1w}$. So, at this instant, both switches switch. However,
with a time varying input signal \( x(t') \), the switching time instants \( t_{2x} \) and \( t_{2w} \) will not be equal. Hence, in the time interval between \( t_{2x} \) and \( t_{2w} \), the integrator has a dc input signal and the output will change. If \( x(t') < w(t') \), then, \( t_{2x} > t_{2w} \); if \( x(t') > w(t') \), then \( t_{2x} < t_{2w} \).

To simplify the notation the following definitions are made:

- \( t_{x_n} \) – time instant in the interval \((n-1)T < t < nT\) at which the signal \( x(t) \) changes from -1 to 0 (See Figs. 9.4 and 9.5).
- \( t_{w_n} \) – time instant in the interval \((n-1)T < t < nT\) at which the signal \( w(t) \) changes from -1 to 0.
- \( t_{x_n}' = \frac{t_{x_n}}{T} \) and \( t_{w_n}' = \frac{t_{w_n}}{T} \)
- \( w_n \) – value of the output at \( t=nT \) or \( t=n \)

Because the derivations for \( x(t') < 0 \) and \( x(t') > 0 \) are quite similar, only the case \( x(t') > 0 \) will be derived. Expressions for \( x(t') < 0 \) will also be shown. Furthermore, it will appear that the response of the system is different for \( t_{x_n}' < t_{w_n}' \) and \( t_{x_n}' > t_{w_n}' \). So, the two cases will be treated separately.

The switching time instant in the interval \((n-1) < t' < n\) is given by the solution to the equation:

\[
x(t') = K_x e^{-a_x(T' - (n-1))}
\]

\[
t' = n-1 - \frac{1}{a_x T} \ln \frac{x(t')}{K_x}
\]
In general, this is a transcendental equation. The solution satisfies the equation:

\[ t' = n-1 - \frac{1}{a_x} \ln \frac{K_x}{x_n} \]  

\[ (10.2) \]

After the switch for signal \( x(t') \) has switched, the integrator starts integrating (Fig. 10.1). Suppose, the output of the integrator at \( t' = n-1 \) is \( w_{n-1} \). Then, the equation that describes the response \( w(t') \) is:

\[ w(t') = w_{n-1} + \frac{t'}{\tau_i} x_n \]

\[ (10.3) \]

\[ \tau_i = \tau_i / T \text{ normalized integration time constant} \]

The switching time instant \( t_w \) for the switch for \( w(t') \) is given by:

\[ w(t') = D_w(t') = K_w e^{-a T(t'-(n-1))} \]

With equation 10.3, this becomes:

\[ w_{n-1} + \frac{t'}{\tau_i} x_n = K_w e^{-a T(t'-(n-1))} \]

\[ (10.4) \]

Again, this is a transcendental equation. The solution \( t_w \) satisfies the equation:

\[ t'_w = n-1 - \frac{1}{a_w} \ln \frac{K_w}{w_n} \]

\[ (10.5) \]

Furthermore, the relation between \( w_n \) and \( w_{n-1} \) is:
Fig. 10.1  Response to input $x(t')$, $t_{x_n} < t_{w_n}$
By solving the equations 10.4, 10.5, and 10.6, the response of the system to any input can be computed.

In order to show the effects of various parameters on the performance of the system, an analysis of the approximate behavior will be executed. Therefore, in the following derivations, the system equations will be linearized around an operating point. With equation 10.6, equation 10.5 becomes:

\[
\dot{w}_n = n-1 - \frac{1}{\alpha_T} \ln \frac{w_n}{K_w}
\]

Substituting equations 10.2 and 10.7 in equation 10.6 yields:

\[
\ln \left( \frac{w_n}{K_w} \right) = n-1 - \frac{1}{\alpha_T} \ln \frac{w_n}{K_w}
\]

A Taylor series approximation of the logarithmic term around the operating point \( w(t) = K_w \) gives the following:

\[
\ln \left( \frac{w_n}{K_w} \right) = \ln \left( \frac{w_n}{K_w} \right) - \left( \frac{w_n}{K_w} - 1 \right)^{1/2} \left( \frac{w_n}{K_w} - 1 \right)^{1/2} \left( \frac{w_n}{K_w} - 1 \right)^{-1/2} \left( \frac{w_n}{K_w} - 1 \right)^{-1/2}
\]

\[
\ln \left( \frac{w_n}{K_w} \right) = \left( \frac{w_n}{K_w} - 1 \right)^{1/2} \left( \frac{w_n}{K_w} - 1 \right)^{1/2} \left( \frac{w_n}{K_w} - 1 \right)^{-1/2} \left( \frac{w_n}{K_w} - 1 \right)^{-1/2}
\]

If the system is close to the operating point, then the higher order terms can be neglected, so:
Now, only the first term in the series approximation has to be taken into account. This yields:

\[ w_n = \frac{\alpha_w T_i K [x(t_n')/K_x]^\alpha_w/\alpha_x}{1 + \alpha_w T_i K [x(t_n')/K_x]^\alpha_w/\alpha_x} \cdot w_{n-1} + \frac{-K [x(t_n')/K_x]^\alpha_w/\alpha_x}{1 + \alpha_w T_i K [x(t_n')/K_x]^\alpha_w/\alpha_x} \]  

(10.9A)

In general, this equation is a time varying linear first order difference equation. If the input signals \( x(t'), y(t'), \) and \( z(t') \) are constant, then the term \( K [x(t_n')/K_x]^\alpha_w/\alpha_x \) will also be a constant, in which case the above equations will be stationary and linear, and, therefore, can easily be solved. The equation shows that the response of the system is fast when the term \( \alpha_w T_i K [x(t_n')/K_x]^\alpha_w/\alpha_x \) is small. Hence, small input signals \( x(t) \) and \( y(t) \) and large input signal \( z(t) \) improve the speed of response. Moreover, for a fast response the integrator time constant \( T_i \) should be made as small as possible. To investigate the goodness of the linear approximation of the system a comparison is made between the real (eq. 10.8) and linearized response. As an example, the following parameters were chosen:

\[ K_x = K_w = 1 \]
\[ x(t_n') = 0.5 \]
\[ \alpha_w = \alpha_x \]
\[ \alpha_w T_i = 2 \]
In this case the system makes the output \( w(t) \) equal to the input \( x(t) \). The response of the system is given by the transcendental equation (equation 10.8):

\[
  w_n = w_{n-1} - 1/2 \ln 2w_n
\]  

(10.10)

The approximation of the response is given by (equation 10.9):

\[
  w_n = 1/2w_{n-1} + 1/4
\]  

(10.11)

The following table compares the real and approximated response at \( n = 1, 2, 3, \ldots \) for an initial condition of the system, \( w_0 = 0 \) and an input signal \( x(t) = 0.5 \) (Table 10.1).

<table>
<thead>
<tr>
<th>( n )</th>
<th>( w_n ) (eq. 10.10)</th>
<th>( w_n ) (eq. 10.11)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Real</td>
<td>Approximated</td>
</tr>
<tr>
<td>1</td>
<td>0.284</td>
<td>0.250</td>
</tr>
<tr>
<td>2</td>
<td>0.396</td>
<td>0.375</td>
</tr>
<tr>
<td>3</td>
<td>0.450</td>
<td>0.437</td>
</tr>
<tr>
<td>4</td>
<td>0.475</td>
<td>0.468</td>
</tr>
<tr>
<td>5</td>
<td>0.488</td>
<td>0.484</td>
</tr>
</tbody>
</table>

Table 10.1. Comparison of real and approximated response

It is seen that the linear approximation of the response is very good.

It slightly underestimates the performance of the system.

To provide more insight into the dynamics of the system, the continuous time equivalent of the difference equation that describes the linearized behavior of the system will be helpful (equation 10.9). It is given by (see Appendix II):
This equation is a time varying linear first order differential equation. Attempts to solve it in closed form were not successful. If the input signals are constant, then the time constant of the resulting constant linear first order differential equation is given by:

\[
\tau' = \frac{-1}{\ln \left( \frac{\alpha_w/\alpha_x}{1+\alpha_w/\alpha_x} \right)} \tag{10.13}
\]

Again, it is seen that decreasing the term \( \alpha_w/\alpha_x \) will increase the speed of response of the system.

The previous analysis of the response of the system to a time varying input signal was executed for the case that \( t_x < t_w \) (\( x(t) > w(t) \)). In the next section, the behavior of the system for \( t_x > t_w \) (\( x(t) < w(t) \)) will be investigated.

In this case, the switching time instant for the switch for \( x(t) \) is also given by equation 10.2. However, depending on the value of the integration time constant \( \tau_i \), three different possibilities for the switching time instant for the switch for \( w(t) \) can occur (Fig. 10.2).
Fig. 10.2 Three different types of responses to input $x(t')$, $t'_x > t'_w$.
A) \( \frac{1}{r} \) bigger than the absolute value of the slope of the carrier signal \( D_w(t') \) at \( t' = t' \). In this case the output of the system will follow the carrier signal (Fig. 10.2A) and the switch for \( w(t') \) will oscillate at a high frequency. The output at \( t' = n \) is given by:

\[
\begin{align*}
\frac{-\alpha T(t'_{n})}{-\alpha T(t'_{(n-1)})} = \frac{K_w}{K_w} e^{x_n} x_n \quad (10.14A)
\end{align*}
\]

If \( x(t') < 0 \), a similar derivation gives:

\[
\begin{align*}
\frac{-\alpha T(t'_{n})}{-\alpha T(t'_{(n-1)})} = \frac{K_w}{K_w} e^{x_n} x_n \quad (10.14B)
\end{align*}
\]

B) \( \frac{1}{r} \) less than the absolute value of the slope of the carrier signal \( D_w(t') \), but the output equals the carrier signal \( D_w(t') \) at \( t' = t' \), which is smaller than \( t' \). First the output \( w(t') \) will decrease as a ramp, then it will follow the carrier signal (Fig. 10.2B). The output at \( t' = n \) is given by equation 10.14.

C) \( \frac{1}{r} \) less than the absolute value of the slope of the carrier signal \( D_w(t') \) at \( t' = t' \), but the output \( w(t') \) does not equal the carrier signal \( D_w(t') \) before \( t' = t' \) (Fig. 10.2C). The output at \( t' = n \) is given by:

\[
\begin{align*}
w_n = w_{n-1} - \frac{v_{n-1}}{\tau} \quad (10.15A)
\end{align*}
\]

when \( x(t') < 0 \), then:

\[
\begin{align*}
w_n = w_{n-1} - \frac{v_{n-1}}{\tau} \quad (10.15B)
\end{align*}
\]
For certain forms of the input, it is also possible that the transcendental equation 10.2 has two solutions in the interval \( n-1 < t' < n-1/2 \)

However, the analysis of how the system will respond is not given here, but it will be quite similar. Since it is desirable that the integration time constant \( \tau_i \) be small, the first one of the three possibilities will frequently occur. A sufficient condition for the system to always show that behavior is:

\[
\frac{1}{\tau_i} > \max \left| \frac{d}{dt} \int K_w e^{-\alpha_w t'(n-1)} \right|
\]

or

\[
\alpha_w \tau_i K_w < 1
\]

This analysis shows that for \( \tau_i > 1/\alpha_w T_K \), a decrease in the integration time constant \( \tau_i \) will speed up the response (Fig. 10.2). If \( \tau_i < 1/\alpha_w T_K \), a further decrease of the integration time constant will not have any effect on the speed of the response. However, in this case the system will go to its required value in one period.

To illustrate the above analysis the step response of the system for \( D_x(t') = D_w(t') \) and \( K_x = K_w = 1 \) is shown (Fig. 10.3).

Whether or not the integrator has a nonzero input is determined by the sign of \( (x(t') - D_x(t')) \) and the sign of \( (w(t') - D_w(t')) \). If those signs are equal, then the integrator input is zero, since the output signals of the two switches cancel. If they are not equal, then the integrator has a dc input (See also Figs. 9.4 and 9.5). So, as soon as the step input signal \( x(t') \) is applied, the integrator starts integrating and the output signal \( w(t') \) will be a ramp. At \( t' = 1/2 \) the dc level of the input to the integrator is zero and the output \( w(t') \) will be a constant. However, when \( w(t') \) equals the carrier signal \( D_x(t') \), the integrator
Fig. 10.3 Step response
has a dc input and, consequently, the output will be a ramp. At $t' = 0.8$ the output signal $w(t')$ equals the carrier signal again. Since, at this time instant the absolute value of the slope of the ramp is bigger than the slope of the carrier signal, the output will follow the carrier signal (See also Fig. 10.2B). For $t' > 1$, the characteristic of the response will be the same for every period (See also Fig. 10.1). In the time interval $0 < t' < 1$, the value of the integration time constant $\tau_i'$ does not have much influence on the speed of the response. However, decreasing $\tau_i'$ will substantially increase the speed of the response when $t' > 1$.

In conclusion, to speed up the overall response of the system, one should make the integration time constant $\tau_i'$ as small as possible.
Chapter 11. **Response of the System to a Sinusoidal Input Signal**

The response of the system to a sinusoidal input signal can be computed by solving the transcendental equations 10.2, 10.5 and 10.6. Because of the transcendental nature of the equations and because there are three input signals to be considered, it makes sense to assume the input signal \( y(t) \) and \( z(t) \) to be constant and equal to unity and the input \( x(t) \) to vary sinusoidally, according to the following equations:

\[
\begin{align*}
y(t') &= z(t') = 1 \\
x(t') &= A + B\sin \left( \frac{2\pi t'}{N} \right) \quad (11.1)
\end{align*}
\]

\[\begin{align*}
N &= L/T \\
T &= \text{period of the carrier signal} \\
L &= \text{period of the sinusoidal input signal}
\end{align*}\]

The transcendental equations for finding the switching time instants of the comparator for \( x(t) \) are given by (See equation 10.2):

\[
\begin{align*}
t_{x}^{'} &= n - 1 - \frac{1}{\alpha_x} \ln(\hat{A} + B\sin \frac{2\pi t'}{N}) \quad x(t') > 0 \quad (11.2A) \\
t_{x}^{'} &= n - 1/2 - \frac{1}{\alpha_x} \ln |\hat{A} + B\sin \frac{2\pi t'}{N}| \quad x(t') < 0 \quad (11.2B)
\end{align*}
\]

If \( t_{x_{n}}' < t_{w_{n}}' \), then the switching time instant of the comparator for \( w(t) \) is given by (See equation 10.5):

\[
\begin{align*}
t_{w}^{'} &= n - 1 - \frac{1}{\alpha_w} \ln(w_{n-1}) \quad (11.3A) \\
t_{w}^{'} &= n - 1/2 - \frac{1}{\alpha_w} \ln |w_{n-1}| \quad (11.3B)
\end{align*}
\]
For \( t' > t_n \), it will be assumed that the integration time constant is sufficiently small, so that the output \( w(t) \) follows the carrier signal \( W(t) \) (See equation 10.14). In order to find the response \( w(t') \) the above equations were solved using the digital computer. A typical response is shown (Fig. 11.1). It shows 4 different characteristics in 4 different regions, namely:

1. \( 1 < t' < 5 \) equations II.2A and II.3A describe response
2. \( 6 < t' < 8 \frac{1}{2} \) equation 10.14A describes response
3. \( 9 < t' < 13 \) equations II.2B and II.3B describe response
4. \( 14 < t' < 16 \) equation 10.14B describes response

Before dealing with the influence of the parameters on the response of the system, an attempt will be made to provide insight into the behavior of the system, by deriving an approximate mathematical model. For this, the continuous time equivalent of the difference equation will be helpful. The continuous time equivalent of the difference equation that describes the linearized behavior when \( t_n < t' \) is given by (See equation 10.12):

\[
\frac{dw(t')}{dt'} = ln \left( \frac{\alpha x / \alpha x}{1 + \alpha W / \alpha x} \right) \left[ w(t') - x(t') \frac{\alpha w / \alpha x}{\alpha x} \right]
\]

This equation is a linear time varying differential equation.

For \( t_n < t' \), it is also possible to derive a linearized equation for the response of the system. With the assumption that \( \alpha \tau_i < 1 \) (\( K_w = 1 \)), the following holds (see equation 10.14A):
Fig. 11.1 Response to a sinusoidal signal with $D_x(t') = D_w(t')$ $\Rightarrow w(t') \approx x(t')$ (system Fig. 9.4)
The switching time instant \( t' \) is given by the solution to:

\[
-w_n = e^{w\cdot t' \cdot x_n}
\]

Linearization of this equation around the operating point \( t' \) yields:

\[
A + B\sin \frac{2\pi t}{N} + 2\pi B \cos \frac{2\pi t}{N} + 2\pi A \cos \frac{2\pi t}{N} \cdot x_n \cdot n-l - e^{-\alpha t \cdot x_n \cdot n-l} = 0
\]

Subtracting the operating point values and solving the \( t' \) gives:

\[
t' = \frac{1}{\text{something}}
\]

Multiplying both sides by \(-1\), adding \( n-1 \) and multiplying by \( \alpha_w \cdot T \) yields:

\[
-\alpha_w T \cdot x_n \cdot n-l = -\alpha_w T \cdot x_n \cdot n-l + \frac{2\pi B \cos \frac{2\pi t}{N} \cdot x_n \cdot n-l}{\alpha_w \cdot T \cdot x_n \cdot n-l}
\]

Conversion to the continuous time equivalent gives:

\[
-\alpha_w T \cdot x_n \cdot n-l = \frac{\alpha_w}{\alpha_x} \ln(A + B\sin \frac{2\pi t}{N} \cdot x_n)
\]

Taking the antilogarithm yields (See also Eq. 10.14A):
\[ w(t') = \left[ \hat{A} + B \sin \frac{2\pi}{N} (t' - \phi) \right]^{\alpha_w/\alpha_x} \]  

(11.5)

where \( \phi \) is some value between 0 and 1, depending on how the limit is taken in the conversion from the difference equation to the differential equation.

The equations 11.4 and 11.5 form the basis of an analysis of the approximate behavior of the system. Since the differential equation (Eq. 11.4) cannot be solved in closed form, a worst case approximation will be done. As a worst case approximation, the slowest response of the system will be investigated. This occurs when:

\[ |x(t')|^{\alpha_w/\alpha_x} = |x(t')|_{\text{max}}^{\alpha_w/\alpha_x} = |A+B|^{\alpha_w/\alpha_x} \]

So, the differential equation 11.4 is replaced by:

\[ \frac{dw(t')}{dt} = \ln \frac{\alpha_w T \tau |A+B|^{\alpha_w/\alpha_x}}{1 + \alpha_w T \tau |A+B|^{\alpha_w/\alpha_x}} \left[ w(t')^{\alpha_w/\alpha_x} \right] \]  

(11.6)

Now, if \( \alpha_w/\alpha_x \) is a non-negative integer, this differential equation can be solved. The solution for \( \alpha_w/\alpha_x = 1 \), is:

\[ w(t') = \hat{A} + \frac{2}{\sqrt{(2\pi \tau)^2 + 1}} \sin \left( 2\pi t' - \arctan \frac{2\pi}{\tau} \right) \]  

(11.7A)

\[ \tau = \frac{-1}{\ln \frac{1 + \alpha_w T \tau |A+B|^{\alpha_w/\alpha_x}}{\alpha_w T \tau |A+B|^{\alpha_w/\alpha_x}}} \]  

(11.7B)

The solutions for \( \alpha_w/\alpha_x = 2, 3, \ldots \) can also be computed. This will not be done here. However, if \( \alpha_w/\alpha_x \neq 1 \), it can be seen that some
nonlinear deformation will occur. This is due to the filtering characteristic of a first order filter. The magnitude of this distortion depends mainly on the time constant \( \tau \). When \( t_n' > t_w' \), equation 11.5 describes the behavior of the system. Except for a little phase lag, this behavior agrees quite well with the ideal behavior.

In conclusion, the equations 11.5 and 11.6 show that the system has greater bandwidth or less attenuation and phase lag with increasing frequency ratio \( N \) and decreasing time constant \( \tau \) (equation 11.7B). In turn, the time constant \( \tau \) decreases when \( \alpha_w \tau \mid A + B \mid \alpha_w / \alpha_x \) decreases. So, a small value of \( \tau_1 \), small signals and a large value of \( \alpha_w / \alpha_x \) will make the system respond fast. In the next section graphs will be presented which compare the results of this approximate analysis with the results of experiments that were performed.
Chapter 12. Simulation on Digital Computer

In order to verify the results of the analysis of the approximate behavior of the system (Fig. 9.4), the system was simulated on the digital computer. The amplitudes and the time constants of the carrier signals were chosen to be equal, so output signal $w(t')$ will follow the input signal $x(t')$. The switching time $t_{x_n}^t$ was computed by solving the transcendental equation 11.2. In the case that $t_{x_n}^t < t_{w_n}^t$, the value for $t_{w_n}^t$ was computed with equation 10.7 and depending on the value of the integration time constant $\tau_i$, equation 10.14 or 10.15 gave the value for $w_n$. The input $x(t')$ was a sinusoidal input signal with a dc component $\hat{A}$ equal to zero and an amplitude $\hat{B}$ (Eq. 11.1). In the experiments the following parameters were varied:

- amplitude sinusoidal input $\hat{B}$: $\hat{B} = 0.5, 0.8$
- frequency ratio carrier/input $N$: $N = 4, 8, 16, 32, 64, 128$
- integration time constant $\tau_i$: $\tau_i = 1.0, 0.1, 0.01$

whereas $\omega_w = \omega_x$, $T = 10$, and $K_x = K_w = 1$, so the system computed $w(t') = x(t')$.

The computer program computed the output $w(t')$ and a Fourier series approximation of the output, defined as:

$$w(t') = a_0 + \sum_{n=1}^{\infty} B_n e^{-j\frac{2\pi n}{N} t'}$$

In addition, the programs computed the integral of the error squared, which is defined as:

$$ISE = \frac{1}{N} \int_0^N \left[ \hat{A} + \hat{B} \sin \frac{2\pi}{N} t' - w(t') \right]^2 dt'$$

In Fig. 12.1 the amplitude and the phase shift of the fundamental of the output are plotted. The amplitude is normalized with respect to the
amplitude of the sinusoidal input signal $B$. As can be seen, the 
amplitude ratio $B_1/B$ improves substantially with decreasing integration
time constant $\tau_i$ and with increasing frequency ratio $N$. A typical 
error for the amplitude ratio with $\tau_i = 0.1$ and $N=100$ is about 1%. For 
small values of $\tau_i$ a further decrease in $\tau_i$ (N constant) does not seem 
to improve the response very much. This can be explained by realizing 
that the phase shift is mainly dependent on $N$ when $\tau_i$ is small (See 
Fig. 11.1). For $\tau_i = 0.1$, the estimates of the amplitude ratio and 
the phase shift from equations 11.5 and 11.7 are plotted. If $\tau_i$ is small, 
a reasonable value for the phase shift is $\frac{1}{2N}$. The estimates agree rather well. 
Especially when it is taken into account that the approximation was coarse. 

The results for the integral squared error $ISE$ show similar 
characteristics (Fig. 12.2). Now it makes sense to normalize the $ISE$ 
with respect to the power of the input signal $B_0^2/2$. Again, decreasing 
$\tau_i$ and increasing $N$ improves the response substantially, whereas the 
improvement is not very big when $\tau_i$ is small (N constant). A typical 
error for $\tau_i = 0.1$ and $N=100$ is 0.1%. 

In summary, the following variations give less attenuation and 
phase lag in the frequency response and a smaller integral squared error: 

1. Increasing the frequency ratio $N$
2. Decreasing the integration time constant $\tau_i$
3. Decreasing the amplitude of the input signals

The experiments confirm the deductions that were made from the approximate analysis in the previous chapter.
Fig. 12.1 Amplitude and phase of first harmonic

\[ \alpha_x T = \alpha_w T = 10 \]

\[ \tau_i : \circ \text{-} 1 \]
\[ \times \text{-} 0.1 \]
\[ \bullet \text{-} 0.01 \]

Approximation of \( x \) (eq. 11.7) for \( \tau_i = 0.1 \)
Fig. 12.2 Integral error squared

\[ x_0 T = \alpha_w T = 10 \]

\[ \hat{A} = 0 \quad \hat{B} = 0.5 \]

\[ \hat{A} = 0 \quad \hat{B} = 0.8 \]

\[ \tau'_i : \circ - 1 \quad x - 0.1 \quad \bullet - 0.01 \]
Chapter 13. Accuracy of the System

In the previous chapters the accuracy of the system (Fig. 9.4) with respect to its ability to perform linear and nonlinear operations on time varying signals was investigated. In this chapter the functional accuracy of the system as determined by nonideal component characteristics will be determined. The main sources of error are caused by variations in the carrier signals, comparators and integrator. These sources will be treated consecutively in the following section, after which their influence on the multifunction operation accuracy will be determined. Since the derivations and conclusions for \( x(t) > 0 \) and \( x(t) < 0 \) are similar, only the case \( x(t) > 0 \) will be considered.

A. Carrier signals

The carrier signals are generated by the pulse height modulation method (Fig. 9.4). The two switches are driven simultaneously by the same square wave signal, so it is reasonable to assume that the carrier signal have the same period and are in phase. Variations in the period of the carrier signal and in the time durations of the positive and negative part of the signal pulse are not important as long as the variations are the same in both carrier signals. Hence, the multivibrator that drives the switches does not have stringent frequency stability requirements and can be inexpensive.

From the equation of the ideal carrier signal

\[
D_j(t') = K_j e^{-\alpha_j t} \quad j = x, w
\]

(13.1)

It can be seen that inaccuracies can be due to variations in the amplitude \( K_j \) and the time constant \( \alpha_j \). In addition, there can also
be a dc offset. These notions lead to the equation of the real
carrier signal:
\[ D_j(t') = (K_j + \Delta K_j) e^{-(\alpha_j + \Delta \alpha_j)T/2} + C_j \quad j = x, w \quad (13.2) \]

\( \Delta K_j \) - variation in gain due to the finite forward resistance of
the switches and the variation in the gain of the shaping
network for the exponential carrier signal (variations
in \( R \) and \( C \)).

\( \Delta \alpha_j \) - variations in the inverse time constant due to variations
in the shaping network for the exponential carrier signal
(variations in \( R \) and \( C \)).

\( C_j \) - dc offset of amplifiers.

At \( t' = 1/2k \) (\( k = 0, 1, 2, \ldots \)) the carrier signal reaches its smallest
value. As a consequence, input signals \( x(t) \) that have an absolute value
that is smaller than the smallest absolute value of the carrier signal
cannot be processed, since in this case the switching time instants will
always be \( 1/2k \) (\( k = 0, 1, 2, \ldots \)). This smallest absolute value is given
by:
\[ |x(t)|_{\text{min}} = K_x e^{-\alpha_x T/2} \]

As an example some values of the normalized minimum absolute input signal
\( |x(t)|_{\text{min}}/K_x \) are shown for different values of \( \alpha_x T \) (Table 13.1).

| \( \alpha_x T \) | \( |x(t)|_{\text{min}}/K_x \) |
|-------|-------------|
| 6     | 0.0498      |
| 8     | 0.0183      |
| 10    | 0.0067      |
| 12    | 0.0025      |

Table 13.1 Minimum absolute value of the input signal \( x(t) \)
B. Comparators

The main sources of error that are caused by the comparators are due to hysteresis and variations in the output levels. However, if the NOR pulse width subtraction method is used, the influence of the variations in output levels will be eliminated (Fig. 7.4). In the following analysis it will be assumed that the NOR pulse width subtraction is not used so that we will be able to compare both methods later on.

We can write for the variations in the output levels of the comparator:

\[ A_j = A_j + \Delta A_j \quad j = x,w \]
\[ B_j = B_j + \Delta B_j \quad j = x,w \]  

(13.3)

For the hysteresis of the comparator it will be assumed that it is symmetrical around zero and it has a value of \( 2h_j \) \((j = x,w)\), so

\[ j(t) - D_j(t) - h_j > 0 \quad \text{switch to level } A_j \quad j = x,w \]
\[ j(t) - D_j(t) + h_j < 0 \quad \text{switch to level } B_j \quad j = x,w \]

There is no influence of the hysteresis of the comparator at \( t = t' = 1/2k \), since at those time instants an instantaneous change from 0 to \( \pm K \) in the carrier signal occurs. With this notion, the switching time instant \( t_{2j} \) becomes equal to zero and the switching time instant \( t_{1j} \) is given by:

\[ D_j(t') = j(t') - h_j \quad j = x,w \]  

(13.4)

This equation will be used together with equation 13.2 to determine the accuracy of the system.

C. Integrator

As was mentioned before, the errors due to the integrator are caused by dc offset and noise. It will be assumed that the integration time constant is sufficiently large so that the influence of the noise
can be neglected, while the dc offset is accounted for by lumping it
together with the dc offset \( C_w \) of the comparator for \( w(t) \).

D. Accuracy

An investigation into the accuracy of the system can be done by
considering the variations in the components and computing the actual
output \( w(t) \) of the system, which is compared with the ideal output. The
ideal output follows from equation 9.1A and is given by (equation 9.2):

\[
\frac{B_x}{B_w} \frac{\alpha_w}{\alpha_x} \int_{t}^{t+h} \frac{x(t)}{K} \, dt = w(t)
\]

Due to the variations the actual switching time instants of the switches
are given by (See equations 13.2 and 13.4):

\[
\begin{align*}
t_1^j &= 0 \\
t_2^j &= -\frac{1}{(\alpha_j + \Delta \alpha_j)T} \ln \frac{j(t) - h - C_j}{k_j + \Delta k_j} \\
\end{align*}
\]

These equations are used to compute the dc components of the pulse width
modulated signals (equation 6.1), which are made equal by the integrator.
Substituting equations 13.3 and 13.5 in equation 9.1A yields:

\[
A_w + \Delta A_w \cdot \frac{B_x + \Delta B - A_x - \Delta A_x}{(\alpha_x + \Delta \alpha_x)T} \ln \frac{x(t) - h - C_x}{k_x + \Delta k_x} = A_w + \Delta A_w
\]

If we choose:

1) \( A_j = 0 \) \( j = x, w \)

2) \( B_x = B_w = B \)

then the above equation solved for \( w(t) \) yields:
From this equation the following conclusions can be drawn:

- A hysteresis $h_j$ of the comparator and a dc component $C_j$ in the carrier signal have the same influence. They shift the desired operating point by the value $h_j + C_j$.
- A change in gain $K_w$ will multiply the output $w(t)$ by a constant.
- A change in gain $K_x$ will also multiply the output by a constant, however, this constant depends on the exponent of $x(t)$.
- Changes in the levels, $\Delta A_x$, $\Delta A_w$, $\Delta B_x$, $\Delta B_w$ and the inverse time constant $\Delta \alpha_x$ multiply the output by a constant and change the exponent of $x(t)$.
- A change in the inverse time constant $\Delta \alpha_x$ will only change the exponent of $x(t)$.

Making use of the NOR pulse width subtraction method eliminates the influence of the comparator levels and thus eliminates a significant source of error. In this case the above equation reduces to:

$$w(t) = h_w + C_w + (K_w + \Delta K_w) \left( \frac{x(t) - h_x - C_x}{K_x + \Delta K_x} \right) \frac{\alpha_w + \Delta \alpha_w}{\alpha_x + \Delta \alpha_x}$$

In the following, an investigation into the magnitude of the influence of inaccuracies will be performed by way of a sensitivity analysis.

E. Sensitivity analysis

The sensitivity analysis will be performed on the ideal equation, namely:

$$w = K \left[ \frac{x}{K_x} \right] \frac{\alpha_w}{\alpha_x}$$
The sensitivity of the output $w(t)$ with respect to a parameter $G$ is defined as [56]:

$$S_G = \frac{G}{w} \frac{dw}{dG} \quad G = \text{parameter}$$

Applying sensitivity analysis to equation 13.7 yields:

$S_w = 1$  
Errors in the comparator (dc offset and hysteresis) and in the carrier signal (dc component) for $w(t)$ directly influences the output $w(t)$.

$S_x = \alpha_w/\alpha_x$  
See above comment. However, the errors will be worse for $\alpha_w/\alpha_x > 1$ than for $\alpha_w/\alpha_x < 1$.

The influence of hysteresis in the comparators for $x(t)$ and $w(t)$ tend to cancel each other somewhat. For example, if (equation 13.6):

$$C_j = 0 \quad j = x, w$$
$$\Delta K_j = 0 \quad j = x, w$$
$$\Delta \alpha_j = 0 \quad j = x, w$$
$$\alpha_w/\alpha_x = 1$$

then $w = h_w + x - h_x$

Since the hysteresis is always positive and probably the two comparators will have about the same magnitude of hysteresis, its influence will be reduced.

$S_K = 1$  
Error in gain directly influences the output.

$S_K = -\alpha_w/\alpha_x$  
See above comment. However, errors will be worse for $\alpha_w/\alpha_x > 1$ than for $\alpha_w/\alpha_x < 1$.

Gain errors also tend to cancel each other somewhat, because an increase in both gains (due to the same cause, like temperature) increase both numerator and denominator of equation 13.6, so the
in the quotient will be reduced.

\[ S_{\alpha_w/\alpha_x} = (\alpha_w/\alpha_x) \ln K_x/K_w \]

This sensitivity is large for small input signals \( x(t) \) and \( z(t) \) \((K_w)\). It becomes better if signal \( y(t) \) \((K_x)\) decreases.

In general, this analysis shows that the inaccuracies are larger for \( \alpha_w/\alpha_x > 1 \) than for \( \alpha_w/\alpha_x < 1 \). In the next section the accuracies of the components will be treated in greater detail. Together with the above sensitivity analysis, this will give a good insight into the accuracy of the overall system.

F. Components

1) Resistors: The resistors are used in the exponential carrier signal generator and various operational amplifiers. For these resistors accurate values can be obtained with thin film processing techniques. Commercially available are resistors with tolerances to \( 0.01\% \) and resistor ratio networks with tolerances to \( 0.005\% \) since the resistors can be matched. This accuracy is more than adequate. The variations with temperature are also important, since they could cause large RC time constants to vary significantly. Temperature coefficients of 25 ppm/°C of individual resistors and 1 ppm/°C of resistor ratio networks are available. This means that, due to temperature dependence of resistors, the time constant RC or \( 1/\alpha \) may change as little as \( 10^{-6} \) per degree C.

2) Capacitors: Capacitors are used in the exponential carrier signal generator and in the integrator. The tolerances are usually 5-10%, which is too inaccurate to be able to use them in an
off the shelf manner except by trimming or matching. By matching capacitors accurate RC or $1/\alpha$ ratios can be obtained. The temperature coefficient of individual capacitors is about 30 ppm/°C, which means that RC time constant changes with temperature are due more to capacitors than to resistors.

3) Operational Amplifiers: The main source of error from the operational amplifier results from the input offset voltage. This offset voltage is usually about 0.5 - 2mV, but can be as low as 25 - 125μV with an offset voltage drift of 0.2μV/°C, which is negligibly low. Of course, at a greater expense, it is possible to apply offset nulling. As was shown in the sensitivity analysis, offset voltage directly influences the accuracy of the output. For instance, a 1mV offset in an operational amplifier for the summation of $w(t)$ and $D_w(t)$ gives a 1mV error in the output $w(t)$. However, referred to a full scale output of 10V, a 1mV error gives only an inaccuracy of 0.01%.

4) Switches: For the switches either FET or transistor switches can be chosen. The advantage of the FET switch over the transistor switch is that there is no inherent offset voltage. However, the FET switch has a nonzero ON resistance, which is the main source of error [54, 55]. However, with compensation techniques this value can be brought down to 5 - 50Ω, which with proper choice of the circuit impedance levels has almost negligible effect (10k load, inaccuracy 0.05-0.5%). The response time of FET switches is usually about 100 - 300 ns.

5) Comparators: As with the operational amplifiers, the main source of error is due to input offset voltage. Typical values for
this offset voltage is 1 - 2mV, although 0.3 - 0.5mV with easy offset mulling capability and an offset drift of about 1μV/°C, is available. The response time of the comparator is usually about 100 ns.

6) Response of the exponential wave shaping network to a square wave: In order to obtain the exponential carrier signal a square wave of amplitude Y and period T is fed into a network with transfer function \(-s/(s+\alpha)\). The response of this network is given by (See Appendix III):

\[
D(t) = \begin{cases} 
(1-e^{-1/2\alpha T})Ye^{-\alpha t} & 0 < t < T/2 \\
(1-e^{-1/2\alpha T})Ye^{-\alpha t} & T/2 < t < T 
\end{cases}
\]

As can be seen, some inaccuracy is caused by an unwanted factor, \(1-e^{-1/2\alpha T}\), by which the carrier signal is multiplied. This is due to the fact that at the end of half the period, T/2, the output D(t) has not reached its asymptotic value yet. The factor is small when \(\alpha T\) is large. For \(\alpha T = 10, 12, 14, 16\) the error is 0.67, 0.25, 0.09, 0.03% respectively.

This discussion shows that the accuracy of individual components is good. The total error will be the net effect of the contributions of the inaccuracies of the individual components, neither of which are large and some of them will cancel. It appears that one of the main sources of error is due to the dc offset voltage of comparators and operational amplifiers. A reasonable figure for the offset voltage is about 1 mV, which on a 10 V maximum output scale would result in an error of 0.01% of full scale. Of course, offset compensation can be applied at the cost of more complicated circuitry. The effects of the finite ON resistance of the switches as well as the exponential wave shaping network can be reduced significantly by adjusting the gain of the amplifiers slightly greater than unity.
For the maximum frequency response of the system it is more difficult to obtain an estimate. In order to generate an accurate carrier signal, the period of the carrier signal should be about 10-50 times longer than the response time of the elements, which gives a maximum carrier frequency of about 100-300kHz. In addition, the frequency of the input signal should be considerably lower than the carrier signal, which would give a maximum frequency of the input signal of about 10kHz.

Analog computer simulations indicated that for maximum signal levels of ±10V, an accuracy of 0.05% could be obtained. The inherent accuracy of nonlinear junctions constructed using this pulse width modulation principle will depend on the accuracy of bias levels, resistor, capacitors, etc. If the circuit were integrated on a single chip, one might expect that a very accurate device would result.
Chapter 14. Experiments on the Analog Computer

The multifunction operation concept was implemented on the analog computer. However, there were some differences between the actual implementation and the multifunction circuit as shown in Figure 9.4. At the time that the experiments were performed, the advantages of the pulse height modulation method and the NOR pulse width subtraction method were not conceived yet. Hence, the amplitude of the carrier signals were chosen to be unity and the circuit performed the following operation (See equation 9.2):

\[
\frac{a_B}{w_x} = \frac{x}{w}
\]

As was pointed out before, making the levels of the two comparators equal is advantageous. Hence, two comparators that had similar characteristics were chosen. However, the inaccuracies due to unequal levels were still unacceptable. The characteristics of the comparators were:

- comparator for \( x(t) \): \( x(t) + D_x(t) > 1.0 \text{mV} \) output = 33.2mV
- \( x(t) + D_x(t) < 1.0 \text{mV} \) output = 3.80V
- comparator for \( w(t) \): \( w(t) + D_w(t) > 1.0 \text{mV} \) output = 44.5mV
- \( w(t) + D_w(t) < 1.0 \text{mV} \) output = 3.84V

It can be seen that the comparators had an offset of 1mV and that the output levels were not equal at all. In order to eliminate this, the difference of the pulses out of the comparator was fed into a deadzone comparator, which had a deadzone that was bigger than difference in output levels of the comparators (Fig. 14.1). The pulses from the comparator for \( x(t) \) and the comparator for \( w(t) \) (Fig. 14.1A) are sub-
Fig. 14.1 Elimination of unequal levels
tracted and fed into a deadzone comparator with a symmetrical deadzone of $2E_c$ (Fig. 14.1B, Fig. 14.2), eliminating any influence of the level inaccuracy. The network for the generation of the carrier signals is shown (Fig. 14.3, see also Fig. 9.2). The output of the signal generator (S.G.) was a square wave with a period of 40 ms, which was fed into two networks with transfer functions:

$$H_j(s) = \frac{s}{s + 100\alpha_j}, \quad j = x, w$$

The magnitude of the square wave was adjusted as to make the amplitudes of the exponential carrier signals exactly equal to unity. This resulted in the following carrier signal:

$$D_j(t) = -100\alpha_j^* \begin{cases} e^{j = x, w} & 0 < t < T/2 \\ e^{-100\alpha_j^*(t-T/2)} & T/2 < t < T \end{cases}$$

The ratio of $\alpha_w^*/\alpha_x^*$ determines the nonlinear computation that was performed. For this ratio the following values were chosen (Table 14.1).

<table>
<thead>
<tr>
<th>$\alpha_w^*$</th>
<th>$\alpha_x^*$</th>
<th>$\alpha_w^<em>/\alpha_x^</em>$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>10/3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>1/2</td>
</tr>
<tr>
<td>10/3</td>
<td>10</td>
<td>1/3</td>
</tr>
</tbody>
</table>

Table 14.1 Values of the exponent in the experiments
Fig. 14.2 Deadzone comparator and symbol
Fig. 14.3 Network for generating carrier signals
The above notions led to the following analog computer set up for the implementation of the nonlinear computation (Fig. 14.4). The input signal $x$ was obtained from a 10V supply on the analog computer. This voltage was multiplied by 0.1 and the value of the input signal was set with a potentiometer between plus and minus 1V. Both input signal $x$ and output signal $y$ were measured with a 3 digit digital voltmeter. The results of the measurements are shown in Table 4.2. These results show very good agreement between the theoretical and experimental values. In general, it can be stated that the accuracy in the middle range is about 0.2% of the full scale of 1V. The results tend to be a little less accurate for very small absolute values of the input signal in the case that $\frac{w_x}{x} > 1$. This agrees with the results of the sensitivity analysis which showed increased sensitivity to nonideal components in this case. For input signals that were very close to $\pm 1V$, the errors were also larger. Because the pulses out of the comparators are of very short duration and because the absolute value of the slope of the carrier signal is large, the influence of the propagation delays and the influence of nonideal pulses are increased. In the extreme case of $x = 0.99$ and the computation of the square or cube root, the comparator for $w(t)$ was not able to generate a feedback pulse anymore, causing the output $w(t)$ to increase and the amplifier in the integrator circuit to saturate.

The influence of inaccuracies like hysteresis, drift and offset voltage could be reduced by using the total range of the analog computer, which is $\pm 10V$. For example, the influence of the 1mV offset of the comparator is of the order of 0.1% on a 1V full scale, but becomes of the order of 0.01% on a 10V full scale. Hence, it would be advantageous to adjust the amplitudes of the carrier signals as to make use of the full
Fig. 14.4 Analog computer diagram
scale of the analog computer. However, then the output would be multiplied by the constant \( \alpha_w/\alpha_x \) \( (1/K_x)^w \), which deviates from unity for \( K_w = K_x \neq 1 \). Hence for every \( \alpha_w/\alpha_x \) the constant would be different, requiring an additional amplification.

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( x )</td>
<td>( w )</td>
<td>( x^2 )</td>
<td>( w )</td>
<td>( x^3 )</td>
<td>( w )</td>
<td>( x^{1/2} )</td>
<td>( w )</td>
</tr>
<tr>
<td>-0.990</td>
<td>-0.995</td>
<td>0.981</td>
<td>-0.971</td>
<td>-0.970</td>
<td>-0.959</td>
<td>0.995</td>
<td>-</td>
</tr>
<tr>
<td>-0.900</td>
<td>-0.902</td>
<td>0.810</td>
<td>-0.809</td>
<td>-0.729</td>
<td>-0.729</td>
<td>0.949</td>
<td>-0.953</td>
</tr>
<tr>
<td>-0.800</td>
<td>-0.802</td>
<td>0.640</td>
<td>-0.640</td>
<td>-0.512</td>
<td>-0.512</td>
<td>0.894</td>
<td>-0.899</td>
</tr>
<tr>
<td>-0.700</td>
<td>-0.702</td>
<td>0.490</td>
<td>-0.491</td>
<td>-0.343</td>
<td>-0.345</td>
<td>0.837</td>
<td>-0.840</td>
</tr>
<tr>
<td>-0.600</td>
<td>-0.601</td>
<td>0.360</td>
<td>-0.361</td>
<td>-0.216</td>
<td>-0.218</td>
<td>0.776</td>
<td>-0.777</td>
</tr>
<tr>
<td>-0.500</td>
<td>-0.501</td>
<td>0.250</td>
<td>-0.251</td>
<td>-0.125</td>
<td>-0.127</td>
<td>0.707</td>
<td>-0.709</td>
</tr>
<tr>
<td>-0.400</td>
<td>-0.400</td>
<td>0.160</td>
<td>-0.161</td>
<td>-0.064</td>
<td>-0.067</td>
<td>0.632</td>
<td>-0.634</td>
</tr>
<tr>
<td>-0.300</td>
<td>-0.300</td>
<td>0.090</td>
<td>-0.092</td>
<td>-0.037</td>
<td>-0.032</td>
<td>0.548</td>
<td>-0.548</td>
</tr>
<tr>
<td>-0.200</td>
<td>-0.201</td>
<td>0.040</td>
<td>-0.045</td>
<td>-0.008</td>
<td>-0.016</td>
<td>0.447</td>
<td>-0.447</td>
</tr>
<tr>
<td>-0.100</td>
<td>-0.100</td>
<td>0.010</td>
<td>-0.017</td>
<td>-0.001</td>
<td>-0.015</td>
<td>0.316</td>
<td>-0.314</td>
</tr>
<tr>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>0.100</td>
<td>0.100</td>
<td>0.010</td>
<td>0.011</td>
<td>0.001</td>
<td>0.004</td>
<td>0.316</td>
<td>0.315</td>
</tr>
<tr>
<td>0.200</td>
<td>0.200</td>
<td>0.040</td>
<td>0.040</td>
<td>0.008</td>
<td>0.010</td>
<td>0.447</td>
<td>0.446</td>
</tr>
<tr>
<td>0.300</td>
<td>0.300</td>
<td>0.090</td>
<td>0.091</td>
<td>0.027</td>
<td>0.027</td>
<td>0.548</td>
<td>0.548</td>
</tr>
<tr>
<td>0.400</td>
<td>0.399</td>
<td>0.160</td>
<td>0.160</td>
<td>0.064</td>
<td>0.064</td>
<td>0.632</td>
<td>0.633</td>
</tr>
<tr>
<td>0.500</td>
<td>0.500</td>
<td>0.250</td>
<td>0.250</td>
<td>0.125</td>
<td>0.126</td>
<td>0.707</td>
<td>0.708</td>
</tr>
<tr>
<td>0.600</td>
<td>0.600</td>
<td>0.360</td>
<td>0.361</td>
<td>0.216</td>
<td>0.218</td>
<td>0.776</td>
<td>0.777</td>
</tr>
<tr>
<td>0.700</td>
<td>0.700</td>
<td>0.490</td>
<td>0.489</td>
<td>0.343</td>
<td>0.343</td>
<td>0.837</td>
<td>0.838</td>
</tr>
<tr>
<td>0.800</td>
<td>0.801</td>
<td>0.640</td>
<td>0.640</td>
<td>0.512</td>
<td>0.514</td>
<td>0.894</td>
<td>0.897</td>
</tr>
<tr>
<td>0.900</td>
<td>0.900</td>
<td>0.810</td>
<td>0.808</td>
<td>0.729</td>
<td>0.730</td>
<td>0.949</td>
<td>0.952</td>
</tr>
<tr>
<td>0.990</td>
<td>0.992</td>
<td>0.981</td>
<td>0.969</td>
<td>0.970</td>
<td>0.958</td>
<td>0.995</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 14.2 Measured and theoretical values of input and output
Conclusions

It was demonstrated that the nonlinear pulse width modulation method has important function generation features. Nonlinear pulse width modulation is pulse width modulation in which the carrier signal is a nonlinear function of time over one period. It can be shown that the dc component of the output of the modulator is the inverse of the function that represents the carrier signal over one cycle. Hence, functions like \( \arcsin x \) and \( \ln x \) can easily be generated by using sinusoidal and exponential carrier signals. Applying different carrier signals yield different nonlinear functions. The demodulation was performed by putting a nonlinear pulse width modulator in the feedback path of an integrator. If the carrier signals and the modulators in the input and the feedback path are identical, then demodulation without any steady state ripple in the output signal is performed, while at the same time the response of the demodulator is excellent. If the carrier signals and the modulators are different, then nonlinear computation is performed. A list of possible computations by using sinusoidal, exponential and sawtooth carrier signals was given (See Table 8.2).

Particular attention was paid to the nonlinear computation \( w = z(x/y)^m \) (Fig. 9.4). A model for the response of this system to a step and a sine wave was developed. This showed that if the ratio \( N \) of the frequency of the carrier signal and the frequency of the sine wave input is large, then the system will respond quickly and accurately. Moreover, the magnitude of the quotient \( T_j \) of the time constant of the integrator and the period of the carrier signal is important too. A small value of this normalized integration time constant will give a fast response. In the case that the input signals \( y \) and \( z \) were chosen
to be unity, the response of the system to a sinusoidal input signal of amplitude B was investigated. This was done by developing a Fourier series of the output w(t) of the system. For a frequency of the input signal which is 1/100 (N=100) of the frequency of the carrier signal and a time constant of the integrator that is 1/10 of the period of the carrier signal, the fundamental of the output signal had a typical error of 1% and the phase shift was 1.6°. The error between the output signal and input signal was squared, integrated over the period of the sinusoidal input signal and normalized with respect to the power of the input signal E^2/2. A typical value of this quotient is 0.001, which gives a signal to noise ratio of 1000. By increasing the frequency ratio N and decreasing the normalized integration time constant \( \tau_1 \), the system can be made even more accurate. The approximate analysis gives estimates of the magnitude of the amplitude and phase of the fundamental and clearly indicates the influence of the parameters on the response.

A estimation of the accuracy of the system was performed. Inaccuracies are due to nonideal components like the modulators, carrier signal generators and dc offset of amplifiers. A sensitivity analysis showed that an important source of inaccuracy is caused by the dc offset of operational amplifiers used in comparator, integrator circuit, etc.. It also showed that the sensitivity to nonideal components is greater for \( m > 1 \) than for \( m < 1 \).

Finally, the system was implemented on the analog computer. The signals y and z were chosen to be one, while for the exponent m the values 1/3, 1/2, 1, 2 and 3 were chosen. The output had an accuracy of about 0.2% of full scale, which was 1 Volt. This accuracy could be improved significantly by choosing a 10 V full scale, which would reduce the influence of hysteresis, dc offset, etc., by a great amount.
In conclusion, it was shown that this kind of analog multifunction is a versatile accurate method for performing nonlinear computations.

In the next part of this dissertation (Part II) the nonlinear function generation feature is implemented with fluidic devices to perform "FLUIDIC NONLINEAR COMPUTATION".
PART II

IMPLEMENTATION WITH FLUIDIC COMPONENTS

Abstract

Fluidic nonlinear computation has received very little attention in the literature. Only the multiplication operation has been dealt with to some extent. This part describes how the nonlinear computation \( w = x^ay^b \) can be performed and implemented with fluidic elements. The principle is based on the log-antilog method which yields the relation \( w = \text{antilog}(\text{alog} \ x + \text{blog} \ y) \). Taking the logarithm of a signal can be done with the logarithmic pulse width modulation method, which is pulse width modulation using an exponential carrier signal. This results in an average value of the output of the modulator that is proportional to the logarithm of the input signal. The antilog operation can be performed by constructing a feedback loop around high gain amplification and putting the logarithmic pulse width modulator in the feedback path. The concept was demonstrated in the implementation of a system that computed \( w = x^a \) (\( a = 0.33, 0.5, 0.66, 1, 1.5, 2, 3 \)). The accuracy was \( 1-2\% \) of full scale, except for very large or very small signals in which case the accuracy decreased to \( 5-10\% \) of full scale.
Chapter 15. Introduction to Fluidic Nonlinear Computation

Linear analog computation has been performed with fluidic elements. Equivalent to analog computation with electrical signals, the elementary building blocks are resistors, capacitors, amplifiers and high gain operational amplifiers. With these elements, linear operations like addition, subtraction, differentiation and integration have been executed more or less successfully [57, 59]. Moreover, other linear networks like lead, lag, lead-lag and networks with complex poles, have been synthesized and implemented [58, 59, 70]. In contrast with linear analog computation, nonlinear analog computation with fluidic elements has received little attention, although being able to perform the elementary nonlinear computations like multiplication, division, raising to the \( n^{th} \) power, taking the \( n^{th} \) root would be of significant value. An example in process control would be the linearization of the square root relationship between the pressure drop across and the flow through a sharp edged orifice using a fluidic squarer. Moreover, the multiplication operation has a useful application in the measurement of power. For instance, after linearization of the square root relationship between pressure and flow, the power which is the product of pressure and flow can be computed. It will be shown that this kind of nonlinear computation, which is characterized by the expression \( x^2y \), can be implemented in a simple module.

In the literature, Parker and Rexford [60, 61] describe the multiplication operation using fluidic devices. The method was based on the pulse width/pulse height modulation method. However, the operation of the device was not satisfactory, mainly because of the difficulties in performing the pulse height modulation, which required
a bistable amplifier to be used as a passive device. A different method to perform the multiplication operation was based on the stochastic computing principle [68]. In this method, an analog signal is mapped into the probability space. This means that the analog signal is represented by a random binary pulse train, whose average value is proportional to the analog signal. It can be shown that if two statistically independent random binary pulse trains (representing two signals to be multiplied) are fed into an AND gate, the average value of the output of the AND gate will be the product. The author expected very low bandwidth (0.5 Hz), moderate accuracy and tough problems in realizing the analog to binary pulse train converter. Although approximate division is also possible with this method, the problems to be solved are more difficult.

The reason that fluidic nonlinear computation encounters many difficulties is that no equivalent of the logarithmic relationship of the transistor junction is known in the fluidic field. Therefore, the powerful method of performing nonlinear computation by this particular log-antilog method cannot be used. However, with the logarithmic pulse width modulation method which will be explained in the next chapter, it is also possible to perform this log-antilog operation.

Linear pulse width modulation methods, in which the carrier signal is a linear function of time (sawtooth, triangular wave), have extensively been applied in fluidic measurement and control systems. The feasibility of this was, for example, demonstrated in the following applications:
- fluidic multiplier based on pulse width/pulse height modulation [60, 61].
- thrust vector control in missile systems [59].
- fluidic PID controller with PWM signals [66].
- fluidic servoamplifier operated by the PWM mode [64].
- pulse width modulation for phase measurement [65].
- fluidic proportional amplifier using pulse width modulation techniques [63].
- fluidic attitude control [62, 69].

In these cases, the average value of the pulse width modulated signal was used as a control or measurement signal. This average value is a linear function of the input modulating signal. By applying a carrier signal that is a nonlinear function of time over one period, the average value of the output of the modulator will be a nonlinear function of the input signal, namely the inverse of the function that represents the carrier signal over one cycle. Hence, for an exponential carrier signal, this average value will be proportional to the logarithm of the input signal. Fluidically, it is possible to generate this exponential carrier signal quite easily with resistor-volume networks, as will be demonstrated later. Since the basic elements for performing pulse width modulation and other linear computations are also available, it becomes possible to perform multifunction operation \(x^a y^b\), based on the log-antilog method.

Before the module that implements this multifunction operation will be dealt with in greater detail, a treatment of the different components that are used in the implementation will be given. After that, the logarithmic pulse width modulator will be dealt with and, finally,
the construction of the multifunction module will be treated.
Chapter 16. Characteristics of Fluidic Devices

This chapter will give an overview of the fluidic devices that are used for the implementation of the multifunction operation. Also simple lumped parameter models of resistors, volumes, amplifiers, operational amplifiers, flip-flop and OR-NOR elements will be described. The fluidic devices used were supplied by Corning Glass Works. A list of devices used in the implementation of the system is given in Appendix IV.

A. Resistors

The relation between the pressure drop across and the flow through a resistor can be modeled by (See Fig. 16.1A):

\[ P_1 - P_2 = \rho \rho Q \]

- \( P_1 - P_2 \) - pressure drop across resistor [lbf/in^2]
- \( Q \) - volume flow through the resistor [in^3/sec]
- \( \rho \) - density [lbfsec^2/in^4]
- \( R \) - resistance [/Insec]

In order for the resistance to be independent of flow and pressure and thus to yield a linear relation, the flow has to be incompressible and laminar. For an incompressible and laminar flow through a pipe, the resistance is linear and is given by the Hagen-Poiseulle relation. Since the pressures in fluidic elements are quite low and the pressure drops are much less than the bulk modulus of air the flow is nearly incompressible and the resistance can be represented with good accuracy by this relation:

\[ R = \frac{128\mu L}{\pi D^4 \rho} \]

- \( \mu \) - absolute viscosity [lbfsec/in^2]
- \( L \) - length of the pipe [in]
- \( D \) - diameter of the pipe [in]
Fig. 16.1 Diagram of resistor and volume components
Linear resistances based on this principle are commercially available.

B. Volumes

A volume can store energy and hence is equivalent to a capacitor. The capacitance of a volume is given by the equation (See Fig. 16.1b):

\[ \rho Q = C \frac{dP}{dt} \]

- \( C \) - capacitance of the volume \( \text{[in} \text{sec}^2 \text{]} \)
- \( P \) - pressure in the volume \( \text{[lbf/in}^2 \text{]} \)
- \( \rho Q \) - mass flow into the volume \( \text{[lbf sec/in]} \)

The capacitance \( C \) of the volume \( V \) can be computed by using the ideal gas law. This results in the following values of the capacitance for a polytropic process:

\[ C = \frac{V}{nR_g T} \]

- \( V \) - volume \( \text{[in}^3 \text{]} \)
- \( R_g \) - gas constant \( \text{[in}^2 \text{sec}^{-2} \text{°K}] \)
- \( T \) - absolute temperature \( \text{[°K]} \)
- \( n \) - polytropic constant

\( n=1 \) for isothermal, 1.4 for adiabatic (air)

It can be seen that the capacitance of a volume varies significantly depending on whether the expansion is isothermal or adiabatic. Usually, it is not obvious what the value for \( n \) is, hence, it is difficult to determine the theoretical capacitance exactly. However, for fluidic systems a value of \( n=1.4 \) (adiabatic) seems to be most satisfactory, since the operating frequencies are high so that little heat transfer occurs.

C. Resistor-volume network

A passive filter can be obtained by connecting a resistor and a volume together (Fig. 16.2). The transfer function of this network is given by [58]:
PLEASE NOTE:

This page not included in material received from the Graduate School. Filmed as received.

UNIVERSITY MICROFILMS
Fig. 16.2 Passive resistor-volume network
\[
\frac{P_O}{P_i} = \frac{R_L}{R_1 + R_L} \frac{R_1 R_L}{R_1 + R_L} \frac{V}{\eta R T} s + 1 = \frac{K}{\tau s + 1}
\]

pressure gain \( K = \frac{R_L}{R_1 + R_L} \)

time constant \( \tau = \frac{R_1 R_L}{R_1 + R_L} \frac{V}{\eta R T} \)

\( P_O \) - output pressure
\( P_i \) - input pressure
\( R_1 \) - resistance
\( R_L \) - load resistance
\( s \) - laplace operator

Since this is a first order transfer function, it can be seen that the response of this network to a square wave input consists of exponential signals, which will be used later to construct the exponential carrier signal.

D. Amplifiers

The amplifier that is most used in analog fluidic systems is the beam deflection amplifier (Fig. 16.3). The output pressure difference \( P_{\text{out}} = P_{\text{out}} \) is a nonlinear function of the input pressure difference \( P_{\text{in}} = P_{\text{in}} \), so the device has a nonlinear gain characteristic [67].

Especially when the input pressure becomes large, the nonlinearity is quite undesirable. However, around the origin the relation is nearly linear and satisfactory for many amplification purposes. Another disadvantage of the beam deflection amplifier is its relatively low input resistance, so it consumes a great amount of flow for a given input pressure. For the above reasons, single beam deflection amplifiers were not used, but instead the high gain operational
Fig. 16.3 Beam deflection amplifier and gain characteristic
amplifiers which consisted of many cascaded stages were chosen. The first stage of such an operational amplifier element has much smaller input ports which gives a much higher input resistance.

E. Operational amplifiers

High gain operational amplifiers are obtained by cascading 2, 3, 4 or 5 amplifiers. The flow gain is about 5-6 per stage and when saturated, the output remains constant. Moreover, the first stage is designed to have a much higher input impedance, so the operational amplifiers do not have the undesirable properties of single beam deflection amplifiers. By constructing feedback paths around the operational amplifier it is possible to perform linear computations like addition, subtraction, multiplication by a constant, lag-lead, lead-lag, differentiation and integration. However, in order for the system to be stable, stabilizing volumes and resistors are needed, which make the whole construction cumbersome. Moreover, since there is no equivalent of the blocking capacitor in fluidic systems, a pure integrator cannot be built. Therefore, a bootstrap configuration must be chosen which requires careful matching of the resistors. Since most of the characteristics of fluidic elements are dependent on the supply pressure and the temperature, the adjustment of the integrator is a difficult task [70]. The aforementioned disadvantages limit the applicability of the operational amplifier. However, because of its low hysteresis, the open loop operational amplifier is useful as a comparator which will be described in the next section.

F. Flip-flops

A fluidic flip-flop is a bistable device with memory characteristic. Its characteristics and truth table are shown (Fig. 16.4,
Fig. 16.4 Flip-flop characteristics
Table 16.1 Truth table for flip-flop.

<table>
<thead>
<tr>
<th>$P_{c_1}$</th>
<th>$P_{c_2}$</th>
<th>$P_{o_1}$</th>
<th>$P_{o_2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>unallowable</td>
<td></td>
</tr>
</tbody>
</table>

Table 16.1 Truth table for flip-flop.

If there is a control input $P_{c_1}$ and no control input $P_{c_2}$, then there is an output $P_{o_1}$ and no output $P_{o_2}$. Similarly, if there is a control input $P_{c_2}$ and no control input $P_{c_1}$, then there is an output $P_{o_2}$ and no output $P_{o_1}$. In the absence of both control inputs, the flip-flop remains in the same state; this state was determined by the last applied control input pair $P_{c_1}, P_{c_2}$. This causes the device to possess memory characteristics. In order for the flip-flop to switch state, the control pressure should be about 10% of the supply pressure. Note that the situation in which both control ports have an input signal is not allowed.

G. OR-NOR

The symbol for the OR-NOR element is shown (Fig. 16.5). If either of the control pressures $P_{c_1}, P_{c_2}, P_{c_3}$, and $P_{c_4}$ is present, then there will be an output pressure $P_{o_1}$ and no output pressure $P_{o_2}$. If neither of the control pressures are present, then the output pressure $P_{o_2}$ will be present and there will be no output pressure $P_{o_1}$. This device generates the OR-NOR function with the output $P_{o_1}$ being the OR function and the output $P_{o_2}$ being the NOR function. In order for this
Fig. 16.5 OR-NOR characteristics
device to switch to $P_{o1}$, the control pressure should be about 10\% of the supply pressure, while the return pressure is about 2\% of the supply pressure. Generally, the operation in the NOR state is more favourable since the output is directly connected to the supply, which causes the loading characteristic to be better than that of other fluidic devices (e.g. flip-flop).

H. Digital amplifier

The symbol for the digital amplifier is shown (Fig. 16.6). The difference between a digital amplifier and a flip-flop is that a digital amplifier has a ratio of output pressure to supply pressure that is much higher and an hysteresis that is much lower than that of a flip-flop. Because of the increased output pressure a digital amplifier is better able to drive a load. Either pressure $P_{c1}$ or $P_{c2}$ should be present at all times, but at no time should be present simultaneously.

The fluidic devices that were treated above form the basic building blocks for the nonlinear pulse width modulator that will be treated in the next chapter.
Fig. 16.6 Digital amplifier
Chapter 17. Fluidic Logarithmic Pulse Width Modulation

Logarithmic pulse width modulation is performed by comparing a high frequency exponential carrier signal (period $T$) with an input modulating signal (Fig. 17.1) [74]. Switching of the comparator occurs when the exponential carrier signal equals the input signal $P_x$. One switching occurs at $t = 0, T, 2T ...$. For the first cycle the other switching time instant is given by:

$$P_x' = P_o e^{-\alpha_x t}$$

or:

$$t_x = -\frac{1}{\alpha_x} \ln \frac{P_x'}{P_o}$$

(17.1)

$\alpha_x$ = inverse of the time constant of the exponential carrier signal
$P_x'$ = Input modulating signal
$P_o$ = amplitude of the exponential carrier signal

Hence, it can be seen that the average value of the output of the modulator is proportional to the logarithm of the input signal $P_x'$. This type of modulation will be called logarithmic pulse width modulation. The exponential carrier signal can be generated by feeding a square wave into a resistor-volume network and the comparator function can be performed by a Schmitt trigger element. In the following section the fluidic oscillator that generates the square wave, the resistor-volume network and the comparator (Schmitt trigger) will be treated.

A. Fluidic oscillator

Fluidic oscillators can be built in a variety of ways [71, 72]. The most common method is based on feedback around a bistable flip-flop element (Fig. 17.2). The period of oscillation is determined by the
Fig. 17.1 Logarithmic PWM
Fig. 17.2 Fluidic oscillators
length of the delay line between the output and the control input. This
method is very simple, however, the period of the oscillation and the
length of the pulses vary considerably. This is probably due to the
variations in the switching levels of the flip-flop element. A better
method is to construct a feedback loop around an operational amplifier
(Fig. 17.2B). This gives a very stable and accurate square wave [72].
The outputs of the operational amplifier are connected to a flip-flop
element, which output is fed into an OR-NOR element to act as a buffer.
This method (Fig. 17.2B) was used to obtain the square wave that was
used in the generation of the exponential carrier signal.

B. **Resistor-volume network**

The steady state response of a resistor-volume network to a
square wave with period T consists of periodic exponential waves
(Fig. 17.3). It can be seen that in the time interval T/2 < t < T,
where t = 0 when \( P_{i} \) switches from 0 to \( P_{i} \), the desired exponential
carrier signal is generated. If it were possible to eliminate the
influence of the first part (0 < t < T/2) of the carrier signal, then the
second part (T/2 < t < T) will determine the switching time instants of
the comparator and, hence, generate the logarithm. The first part of
the carrier signal can be eliminated as shown in the next section. If
this is done, then the carrier signal for the second part can be described
by:

\[
D(t) = P_{0} e^{-\alpha t} = P_{0} e^{-t/\tau}
\]

where

\[
P_{0} = \frac{R_{L}}{R_{i} + R_{L}} P_{i} \max \\
(\text{approximation for large } T/\tau)
\]

\[
\tau = \frac{1}{\alpha} = \frac{R_{i} R_{L}}{R_{i} + R_{L}} \frac{V}{n R \xi}
\]

(17.2)
Fig. 17.3 Resistor-volume network
D(t) - carrier signal

$P_o$ - amplitude of carrier signal \([\text{lbf/in}^2]\)

$\alpha$ - inverse time constant \([\text{sec}^{-1}]\)

$\tau$ - time constant \([\text{sec}]\)

$R_i$ - input resistance \([\text{in/sec}]\)

$R_L$ - load resistance \([\text{in/sec}]\)

$P_{max}$ - amplitude of square wave \([\text{lbf/in}^2]\)

$V$ - volume \([\text{in}^3]\)

$n$ - polytropic constant

$R_g$ - gas constant \([\text{in}^2/\text{sec}^2\cdot\text{K}]\)

$T$ - absolute temperature \([\text{K}]\)

The amplitude $P_o$ and the time constant $\tau$ are a function of the load resistance $R_L$. This load resistance is the input resistance of the fluidic element (comparator in this case) that is connected to the output. Since the input resistance $R_i$ of a fluidic element is not constant, but is load dependent, the amplitude and time constant of the exponential carrier signal will vary. The variation will be negligible if the ratio $R_i/R_L$ is small, which is accomplished by choosing $R_i$ small. Because of the output flow-pressure characteristic of the NOR element, it does not appear as a pressure source and the first part of the square wave when the volume is charging, will be deformed \((0 < t < T/2)\). However, since only the second part of the carrier signal is used, this deformation does not influence the quality of the exponential carrier signal.

C. Comparator

The comparator function can be accomplished by connecting a 3 stage operational amplifier, a flip-flop and a NOR element In series (Fig. 17.4). This element is known as a Schmitt trigger [73]. The
Fig. 17.4 Comparator and logarithmic PWM
operational amplifier changes state whenever the input signal $P'_x(t)$ equals the carrier signal $D(t)$. The output of the operational amplifier is fed into a flip-flop. The flip-flop output is fed into a NOR element. For the exponential carrier signal and the input signal shown in the figure (Fig. 17.4), the pulse output of the NOR element is depicted. The unwanted part of the pulse ($0 < t < T/2$) is eliminated by feeding the input square wave into the NOR element (dashed line in the figure), which causes the output of the NOR element to be zero in the interval $0 < t < T/2$. Hence, the pulse length (dashed line) out of the NOR element is proportional to the logarithm. The length of the pulse is given by (See equation 17.1)

$$t_x = -\frac{1}{\alpha_x} \ln \frac{P'_x}{P_0}$$

The relation between the input signal $P_x$ and the signal $P'_x$ is given by:

$$P'_x = \frac{R}{R+R_L} P_x = \beta P_x$$

thus:

$$t_x = -\frac{1}{\alpha_x} \ln \frac{\beta P_x}{P_0}$$  \hspace{1cm} (17.3)

Again, the load resistance $R_L$ will not be constant and it would be advantageous to choose $R$ much smaller than $R_L$. However, in general, the signal to be measured will be loaded by the resistance $R$ and, therefore, $R$ should be made as big as possible. As a compromise, the resistance $R$ was chosen to be about equal to the input resistance $R_L$ of the operational amplifier. Since the variations in $R_L$ are not very large and since their influence is reduced by the choice $R = R_L$, it will be assumed that the parameter $\beta$ is a constant.
Chapter 18. Fluidic Nonlinear Computation

A. Description of the System

In this section the nonlinear computation feature of the logarithmic pulse width modulation method will be demonstrated for the nonlinear computation $P_w = P_x^a$. The implementation of this nonlinear computation is based on the relation:

$$\log P_w = a \log P_x \quad (18.1)$$

or:

$$P_w = \text{antilog}(a \log P_x)$$

In the previous section it was shown how the logarithmic function could be obtained with the logarithmic pulse width modulation method. It is well known that the inverse of a function can be obtained by putting this function in the feedback path of an integrator [74] (Fig. 18.1). Hence, by putting a similar logarithmic pulse width modulator in the feedback path, the antilog function is implemented. In the steady state, the average value of the sum of the input signals to the integrator has to be equal to zero, so equation 18.1 is implemented and the system computes $P_w = P_x^a$. The carrier signals for the input and feedback modulator are chosen to have the same period, amplitude and no phase difference, but to have a different inverse time constant $\alpha_x, \alpha_w$, hence:

carrier signal for input $D_x(t) = P_o e^{-\alpha_x t}$

carrier signal for feedback $D_w(t) = P_o e^{-\alpha_w t}$

As was done in equation 17.3, the switching time instants $t_x$ and $t_w$ of the comparators for the input and output signal can be computed from:

$$t_x = -\frac{1}{\alpha_x} \ln \frac{BP_x}{P_o} \quad \quad t_w = -\frac{1}{\alpha_w} \ln \frac{BP_w}{P_o} \quad (18.2)$$
Fig. 18.1 Nonlinear computation
If the amplitudes of the pulses out of the input and feedback modulator are chosen to be the same, then the only way in which the average value of the sum of the input signals to the integrator can be zero in the steady state is for \( t_x \) to equal \( t_w \), hence:

\[
- \frac{1}{\alpha_w} \ln \frac{\beta P_w}{P_o} = - \frac{1}{\alpha_x} \ln \frac{\beta P_x}{P_o}
\]

or:

\[
\frac{\beta P_w}{P_o} = \left[ \frac{\beta P_x}{P_o} \right] \frac{\alpha_w}{\alpha_x}
\]

(18.2A)

and the nonlinear computation is performed. The factor \( \beta/P_o \) can be seen as a normalizing or scaling factor.

The complete circuit is shown (Fig. 18.2). The output of the oscillator which generates the square wave is fed into the resistor-volume networks which perform the exponential wave shaping. These networks have the same resistance \( R_1 \), but a different size of the volume; \( V_x \) for the input and \( V_w \) for the feedback modulator. Hence, the ratio of the inverse of the time constants \( \alpha_w/\alpha_x \) of the exponential carrier signals is equal to \( V_x/V_w \) (See equation 17.2), which causes the following computation to be performed:

\[
\frac{\beta P_w}{P_o} = \left[ \frac{\beta P_x}{P_o} \right] \frac{V_x}{V_w}
\]

By simply changing the sizes of the volumes it is possible to adjust the exponent in this relation.

In the feedback configuration (Fig. 18.1) the integrator would force the length of the pulses to be equal. However, fluidic integrators are troublesome and difficult to construct [70]. For this reason a different feedback scheme was devised (Fig. 18.2). Let us first consider
APPENDIX IV list of devices and values

Fig. 18.2
Computation of powers and roots
the last three stages of amplification which consist of a flip-flop FF\(_2\) and two digital amplifiers. The last digital amplifier has an increased supply pressure of 15 psi (1034 mbar) to provide sufficient amplification. If the output of the last digital amplifier is at port L, then the pressure in the volume \(V_o\) will increase and, consequently, if the output is at port R, then the pressure \(P_w\) in the volume will decrease since it is being depleted through the resistors \(R_f\) and \(R\). Because a digital amplifier is a bistable device, the pressure \(P_w\) will oscillate. However, for a sufficient large volume \(V_o\) the amplitude of the oscillations will be neglectable. The state of the last digital amplifier is determined by the state of flip-flop FF\(_2\) which, in turn, follows from its inputs. These inputs are determined by the output of the flip-flop FF\(_1\) and NOR element \(N_1\). It will appear that the output of NOR \(N_1\) depends on whether the pulse length \(t_x\) out of flip-flop FF\(_x\) is longer or shorter than the pulse length \(t_w\) out of flip-flop FF\(_w\), so two cases can be distinguished:

\[
\frac{t_x}{t_w} > 1
\]

For this case (see equation 18.2):

\[
\frac{\beta P_w}{P_o} > \left[ \frac{\beta P_x}{P_o} \right] \frac{V_x}{V_w}
\]

(18.3)

The pulses out of the flip-flops FF\(_1\), FF\(_x\) and FF\(_w\) which are the input of NOR \(N_1\) are shown (Fig. 18.3A). From this, it appears that the output of NOR \(N_1\) is zero over the entire period. The output of FF\(_1\) causes the flip-flop FF\(_2\) and, consequently, the last digital amplifier to be set to port R. Hence, the pressure \(P_w\) in the volume \(V_o\) will decrease, which causes the system to go to its equilibrium given by equation 18.2A.
Fig. 18.3 Pulses out of NOR $N_1$
For this case

\[
\frac{\beta P^x_{\text{L}}}{P_o} < \left[ \frac{P_x}{P_o} \right] \frac{V_x}{V_w}
\]

(18.4)

From Fig. 18.3B it can be seen that in this case the NOR \( N_1 \) will have a pulse output. The length of this pulse is equal to \( t_w - t_x \), so this NOR element essentially performs pulse width subtraction. The pulse causes the flip-flop \( FF_2 \) and, consequently, the last digital amplifier to be set to port \( L \) and the pressure in the volume will increase. At the beginning of each period the flip-flop is reset by the output of flip-flop \( FF_1 \), so the total length \( T_L \) of the pulse out of the \( L \) port of the last digital amplifier is given by:

\[
T_L = T - \left( \frac{T}{2} - \frac{1}{\alpha_x} \ln \frac{\beta P^x_{\text{L}}}{P_o} \right) = \frac{T}{2} + \frac{1}{\alpha_x} \ln \frac{\beta P^x_{\text{L}}}{P_o}
\]

Hence, the bigger the input signal \( P_x \), the longer the pulse. This pulse is being amplified and filtered again to cause the pressure \( P_w \) in the volume to increase. The advantage of the variable pulse length can be seen by realizing that the bigger \( P_x \), the bigger \( P_w \) and the faster the depletion of air out of the volume, which makes a greater pulse length desirable, since it compensates for this depletion.

B. Measurements

1) Square wave and exponential carrier signal

The square wave from the oscillator circuit (Fig. 18.2) and the exponential carrier signal, which is the response of the resistor volume network are shown (Figs. 18.4 and 18.5). The picture of the square wave was taken with resistance \( R_1 \) disconnected and the pressure transducer
Fig. 18.4 Square wave
Fig. 18.5 Exponential signal
(pressure to electrical signal) connected to the output of NOR element N_x. Hence, the output was blocked and there was no loading effect. The period of the square wave was 0.164 sec. and the amplitude was 1.1 psi (76 mbar). The switching from "1" to "0" does not exactly occur at T/2 sec., but this does not affect operation. The exponential carrier signal had an amplitude of 0.84 psi (58 mbar) and a measured time constant of 16 msec. The reduced amplitude of the exponential carrier signal is due to loading effects which, in effect, deform and reduce the amplitude of the square wave. The time constant of the exponential carrier signal can be computed from equation 17.2. With R_L << R_L, the computed time constant τ_c is given by:

\[ \tau_c = \frac{R_v L_x}{nR_g T} \]

With

\[ R_v = 0.26 \times 10^7 \text{in/sec} \quad (10.2 \times 10^7 \text{m/sec}) \]

\[ V = 17.4 \text{ml} \quad (1.06 \text{in}^3) \]

\[ n = 1.4 \quad \text{(adiabatic process)} \]

\[ R_g = 2.48 \times 10^5 \text{in}^2/\text{sec}^2 \cdot ^\circ R \quad (4.46 \times 10^5 \text{in}^2/\text{sec}^2 \cdot ^\circ K) \]

\[ T = 295 \ ^\circ K \]

For these values: \( \tau_c = 0.015 \text{sec} \)

which agrees with a measured value of 16 msec.

2) **Powers and Roots**

Measurements of the system computing powers and roots were made. By varying the sizes of the volumes \( V_x \) and \( V_w \), the values for the exponent \( V_x/V_w \) as shown in table 18.1 were chosen.
Table 18.1 Different sizes of volumes yielding different exponents.

The input pressure $P_X$ and the output pressure $P_W$ were normalized with respect to $\beta/P_o$, which had a measured value of 1/2 psi (34.5 mbar). Since the amplitude of the carrier signal was 0.84 psi (58 mbar), the measured value for $\beta$ is 0.42, which agrees with the design consideration indicating that a $\beta$ of about 0.5 was a good compromise. In order to be able to plot the results on an X-Y recorder a pressure transducer which converted a pressure to a voltage was used. The gains of the pressure transducer for the input $P_X$ and the output $P_W$ were chosen to be equal and was 1 V/psi (0.0145 V/mbar). The gains of the X input and Y input of the X-Y recorder were both 2 in/V (5 cm/V), giving an overall gain of 2 in/psi (0.0725 cm/mbar). Plots of the normalized output pressure $P_W/P_o$ versus the normalized input pressure $\beta P_X/P_o$ are shown for different values of the exponent $V_X/V_W$ (Fig. 18.6). For the exponent the values 0.33, 0.5, 0.66, 1, 1.5, 2 and 3 were chosen. The theoretical and measured values agree quite well, except for very large or very small
Fig. 18.6 Measurements
values of the input and output pressures. This can be caused by a
variety of phenomena, some of which will be dealt with in the following
paragraph where the sources of inaccuracies will be discussed.

First of all, the input resistances $R$ for the input pressure
$P_x$ and the output pressure $P_w$ are not exactly equal. They were taken from
a commercially available set, which had a tolerance of $\pm 10\%$. However,
since only two resistors of value $R$ were available, matching was not
possible. For the resistances $R_i$ which were used in the exponential wave
shaping networks, the situation was the same. It is clear that matching
would improve accuracy. Secondly, different sizes of volumes $V_x$ and $V_w$
will load the NOR's $N_x$ and $N_w$ differently and thus the amplitudes $P_o$
of the exponential carrier signals for the input and feedback modulator are
not exactly the same. An attempt to eliminate this influence by connect­
ing the resistor-volume network for the input and for the feedback modulator
to one single NOR element failed because there appeared to be too much
interaction between the pressures in the volumes $V_x$ and $V_w$. Thirdly, the
NOR elements $N_x$ and $N_w$ are not identical and thus may have different output
pressures and different loading characteristics. An investigation into
this showed output pressure and loading were not significant sources of
inaccuracy. Fourthly, the fluidic elements have a limited speed of
response. As a ballpark figure the delay of a signal through a IX Corning
fluidic element is of the order of 1 ms. Hence, it is impossible for the
NOR element $N_i$ to perform the pulse width subtraction for pulses that have
a pulse length difference that is shorter than about 1 ms. For large
input signals $P_x$ and $P_w$ which yield short pulses, this influence can be
significant. Moreover, the delay through the fluidic elements also makes
it necessary to adjust the length of the tube between the output of the
flip-flop FF₁ and the input of flip-flop FF₂ and the input of NOR N₁. 

Fifthly, the 3 stage operational amplifiers have an offset, hysteresis and a changing input resistance $R_L$ due to loading and control signals on the opposite ports. These quantities are not only a function of the difference of the two input signals to the operational amplifier, but also are a function of the amplitude of the individual signals. The offset and hysteresis play an important role when the input signals are small, while large signals were noticed to influence the input resistance $R_L$.

Despite the influence of the sources of the inaccuracy mentioned above, the accuracy of the system was still good. In general, this accuracy is about 1-2% of full scale. This accuracy decreases to about 5-10% of full scale for very small or very large input signals. Further investigation is needed to better identify and eliminate the sources of inaccuracy, especially for small and large signals. This might improve the overall accuracy of the system significantly.

3) Step response of the system

The overall response of the system to a step change in input signal $P_\times$ is shown (Fig. 18.7). The quantity $\beta P_\times/P_0$ was changed from 0.16 to 0.5 and back. For this test $V_\times$ was chosen to be equal to $V_\ast$, so the output $\beta P_\ast/P_0$ had to be equal to $\beta P_\times/P_0$ in the steady state. The speed of response depends on the resistor-volume-resistor network that is connected to the last digital amplifier. It is clear that for a step decrease in $\beta P_\times/P_0$ the time constant of the response is determined by the time constant $\tau_r$ of the network, which is given by (See equation 17.2):

$$\tau_r = \frac{R_f (R+R_L)}{R_f + R + R_L} \frac{V_0}{nRg}$$
Fig. 18.7 Stepresponse

\[ V_x = V_w = 17.4 \text{ ml} \]

\[ P_x = P_w \]

\[ \tau_{rm} = 6.7 \text{ s} \]
Substituting the values for \( R, R^*, V_0 \) \( n=1.4, R_g \) \( T \) and \( R_L = 0.72 R \) (computed from \( \beta = 0.42 \)) yields:

\[
\tau_r = 7.7 \text{ sec.}
\]

This approximately agrees with the measured value \( \tau_{rm} = 6.7 \) sec. For a step increase in \( \frac{BP_x}{P_o} \) the response of the resistor-volume-resistor network to a pulse train has to be computed. It can be proven that the time constant of the response is the same as for a step decrease and thus is given by equation 18.5. It should be noticed that the steady state value of \( \frac{BP_w}{P_o} \) is not the asymptote of the response. The asymptote for a step decrease is given by \( P_w = 0 \) and the asymptote for a step decrease is determined by the average value of the pulse train out of the last digital amplifier, which depends on its supply pressure and the input signal \( P_x \). This causes the actual speed of response to be faster than the time constant would indicate. As can be seen, the small signal response is much better. For instance, a step change in \( \frac{BP_x}{P_o} \) from 0.5 to 0.4 would only take the response 1.6 sec. to reach its steady state value. In general, it takes about 1-2 sec. for the response to a step change of 10% of full scale to reach its new steady value. From equation 18.5, it also appears that the volume \( V_o \) is the most important factor in the speed of the response. Decreasing the volume would increase the speed of response, but at the cost of increasing the amplitude of the steady state oscillations, which are about 1-2% of full scale for \( \frac{BP_x}{P_o} = 0.5 \) (Fig. 18.7).
Conclusion

This part described how the nonlinear computation \( w = x^a y^b \) can be performed and implemented with fluidic elements. The principle is based on the log-antilog method which yields the relation \( w = \text{antilog} (a \log x + b \log y) \). The logarithm of a signal was obtained using the logarithmic pulse width modulation method, which is pulse width modulation with an exponential carrier signal. This resulted in an average value of the output of the modulator that is proportional to the logarithm of the input signal. The antilog operation was performed by constructing a feedback loop around high gain amplification and putting the logarithmic pulse width modulator in the feedback path. In this way the above multifunction operation can be performed. The concept was demonstrated in the implementation of a system that computed \( w = x^a \). Plots of the output versus the input were shown for \( a = 0.33, 0.5, 0.66, 1, 1.5, 2 \) and 3 were shown and the results were compared with the theoretical values (Fig. 18.6). The results indicated an accuracy of 1-2\% of full scale, except for very large or very small signals in which case the accuracy decreased to about 5-10\% of full scale. Some of the main sources of inaccuracy were due to the nonideal characteristics of the operational amplifiers in the comparator for the logarithmic pulse width modulator. These operational amplifiers had an offset, hysteresis and a variable input resistance. These characteristics especially influenced the accuracy at small signals. The response of the system to a step was shown and it was concluded that for a step change of 10\% of full scale, the output would reach its steady state value in about 1-2 sec.
PART III

STUDENT TECHNOLOGY TRANSFER

Abstract

This part of the dissertation discusses the problems of movement of technology as encountered by the student. The process which leads to a dissertation is viewed as a process of technological change. In this process of technological change, the movement of knowledge and technology from industry to student, and vice versa, can be very important. This process of movement and methods promoting it are investigated.
Chapter 19. Introduction

Most researchers who have undertaken an investigation of the process of change agree that change is very important in the capitalistic society [75, 76]. They think that studying the process of change will enhance our knowledge of how change occurs and that, as a result, this study might give us some ways to influence the process of change. Among the types of change that can occur, social and technological change are of major importance. The definition of social change as given by Rodgers is:

Social change is the process by which alteration occurs in the structure of a social system. National revolution, invention of a new manufacturing technique, founding of a village improvement council, adoption of birth control methods - are all examples of social change [77].

In order to define technological change, we will first have to define technology. A broad definition given by Schon is:

Technology is any tool or technique, any product or process, any physical equipment or method of doing or making by which human capability is extended [78].

In this sense, technology is defined as the "state of the art" or knowledge that is available to mankind; any addition to this "state of the art" is referred to as technological change. In the past, technological change has been the prime mover behind social change; for instance, the Industrial revolution, the boom of technology in the past few decades, have caused a great deal of social change.
However, the process can also operate in the other direction: social change can induce technological change [79]. For example, the change in attitude toward pollution has forced research into pollution control.

There are a variety of driving forces for technological change — among them are competition, the fulfillment of personal needs, all the way from survival to the sense of personal accomplishment, and fulfillment of the needs of society. Those driving forces will not be dealt with in this dissertation; instead, a model will be devised that specifically describes the processes of technological change, but which can be applied to change in the more general sense, also. This model of technological change will be applied to the student working on a project for his dissertation. It will appear that many of the specific problems encountered by the student are similar to the general problems of technological change.

**Models of change**

Social and/or technological change occur in a number of phases. Although the fields of social sciences and technology exhibit quite different characteristics, in the emphasis on change, the processes of social and technological change are quite similar. These similarities are illustrated by the definitions of the different phases in the processes of social and technological change as distinguished by different authors. According to Rodgers, social change consists of:

- **Invention:** the process by which new ideas are created or developed.
- **Diffusion:** the process by which these new ideas are communicated to members of a social system.
According to Hollomon, technological change has the following steps:

- **invention**: the concept or notion that will improve goods, services or processes.
- **innovation**: the introduction of the new concept into the economy.
- **diffusion**: spread of new technology.

From the above descriptions of social and technological change it is clear that the two processes are quite alike. Most differences are of word choice, of definitions rather than conceptual differences. Hence, the characteristics of social and technological change can be merged into a more general model of change, which then can be applied to both fields.

An example of such a model is the block diagram of figure 19.1 which shows the different phases in the process of change - invention, innovation and consequences. The progression of the invention to the innovation stage as well as the movement (which can sometimes be diffusion) of the innovation to the audience of receivers resulting in consequences is indicated by arrows. During the invention/innovation process the movement of knowledge to and from the knowledge pool plays an important role in the progress of this process. However, before a more detailed discussion of this process can be pursued, it is necessary to define the terms invention, innovation, consequences, and knowledge pool as they will be adhered to throughout the following discourse in which we will confine ourselves in the process of technological change.
process of change

invention

innovation

consequences

knowledge pool

Fig. 19.1 Model of change
Invention: first creation of a piece of hardware or software, usually in prototype or in demonstration form, which proves the feasibility of an idea or concept [81].

Innovation: the technical, industrial and commercial steps which lead to the marketing of new manufactured products and to commercial use of new technical processes and equipment [81].

Consequences: changes that occur in industry or society as a result of the rejection or adoption of the innovation [77].

Knowledge pool: all information and resources available to mankind. The distinction between invention and innovation can be distinct or unclear. If inventor and innovator are different people or institutions, the distinction is sharp. However, if the inventor and the innovator are the same person or institution, invention may be a first step and innovation a subsequent step, continuous development of an idea or concept. As was indicated before, in the processes of both invention and innovation, the movement of knowledge to and from the knowledge pool is of great importance. This pool contains information about the "state of the art", which can be existing technology or the knowledge of how things can be done. In turn, this knowledge is available in, for instance, the literature or even one's own experience. One of the benefits of the invention/innovation process is the contribution which the process itself may make to the knowledge pool.

From the above description it is clear that the movement of knowledge or technology plays a vital role in the process of technological change. This movement can occur in two different directions - vertical and horizontal:

Vertical movement of technology refers to the process by which new scientific knowledge is incorporated into technology and by
which a "state of the art" becomes embodied in a system and by which the confluence of several different and apparently unrelated technologies lead to a new technology.

Horizontal movement of technology occurs through the adaptation of a technology from one application to another - possibly wholly unrelated to the first, e.g., adaptation of military aircraft to civilian air transport.

These definitions are adapted from Brooks [84] who first introduced the notion of vertical and horizontal transfer of technology. They imply that horizontal movement of technology occurs across institutional or organizational boundaries, and vertical movement of technology occurs within institutional boundaries [83]. As an example illustrating the occurrence of both horizontal and vertical movement of technology, consider the introduction of a new packaging method. This innovation could be the result of a confluence of different technologies which provide the necessary conditions for the innovation to take place (vertical movement), whereas the adaptation of this new packaging method to a great variety of different products from different industries would be horizontal movement.

Kottenstette and Rusnak [83] distinguish two different and distinct mechanisms for the movement of technology - diffusion and transfer. The process of transfer differs from the process of diffusion in that, in the process of diffusion a professional infrastructure exists to facilitate the movement of technology. An infrastructure can be seen as a channel through which knowledge or technology moves from one place to another - the classroom, meetings and professional
societies, for example. No implication should be drawn about the speed or effectiveness of the movement of technology which occurs through diffusion; diffusion can be fast and effective.

The process of transfer is characterized by the absence of a professional infrastructure. Hence, if knowledge or technology is to move from one place to another, overt action must be taken. The creation of, and/or improvement on, an infrastructure are examples of overt action and technology transfer. It is important to realize that some diffusion always has to take place if the movement of technology is to be complete. Because the acceptance and utilization of knowledge and technology is the important issue, the receivers have to be convinced of its usefulness. This is clearly a process of diffusion in which the establishment of favourable external conditions (coercion, incentives, explaining patiently, providing technical information, etc.) are acts of transfer. In this light, transfer can be seen as any overt action that facilitates the movement of technology which ultimately takes place through diffusion.

In the next section, a model for the movement of technology will be devised which will provide an insight into important factors that influence the movement of technology.

Model of the movement of technology

Rodgers [77] models the characteristics of the cross-cultural communication of innovations as a Source-Message-Channel-Receiver system. Gruber and Marquis [93] relate problems in the transfer of technology to problems of interpersonal communication and communication
between groups. They also note that communication within a firm and to the external environment have explanatory power when differences in performances are identified. Hence, effectiveness of communication appears to be indispensable for effective movement of technology. Based on these ideas, it appears useful to model the process of movement of technology as a Transmitter-Channel-Receiver system used for the transmission of signals in a communication system (Fig. 19.2). The quality of transmission in these systems depends mainly on the properties of the transmitter, channel and receiver, and on the interference level from the environment. The similarities between the transmission of signals and the movement of technology will be shown by relating the concepts and definitions from communication systems [50, 98] to those involved in the movement of technology. This will pinpoint factors that influence, promote, or inhibit the movement of technology.

Transmitter-Channel-Receiver Movement of Technology Model and Knowledge

1. Characteristics of transmitter and receiver

The quality of transmitter and receiver influences the overall quality of the transmission of signals as well as the movement of technology. The quality can be evaluated with respect to:

Power of transmission and sensitivity of receiver: more powerful transmitter yields a more favourable signal-to-noise ratio and a sensitive receiver will be better able to detect a signal.

The authority and status of the person trying to accomplish movement can, to some extent, determine the receiver's motivation, willingness to listen and degree of trust in the importance of the message. More-
Fig. 19.2 Transmitter-Channel-Receiver Model
The capability of transmitter and receiver to use different codes (AM, FM, etc.): the transmitter and receiver must be capable of using the same code and must be in the same code mode during any specific transmission.

According to Allen [94, 95], a mismatch in information coding schemes appears to be responsible for the ineffectiveness of communication across organizational boundaries. Examples are the use of high level mathematics which can provide precise formulation, but which cannot be understood by the receiver, or the inability to understand a foreign language.

The application of a filter to reduce the influence of noise can improve the quality of reception; however, it may also block part of the signal.

An overwhelming quantity of information may make it necessary to classify information as to its expected usefulness. However, this may result in discarding useful information.

2. Characteristics of the channel or medium

A short channel improves the quality of transmission since distortion, attenuation and signal-to-noise ratio are thus minimized. Movement of technology is facilitated by 1) minimizing transfer points [87] and 2) minimizing the distance between transfer points (minimizing diffusion...
Channel capacity determines how many messages can be sent or received at the same time. If too many messages are sent, the receiver will not be able to distinguish among messages.

3. Frequency of communication and feedback

The more frequent the communication the better the chance that the message will get across: this can be accomplished through repeated broadcasting with one or more transmitters through a number of channels (redundancy).

Two-way communication (transmitter becomes receiver and vice versa) makes a dialogue possible which improves the chances of effective communication. The quality of transmission does not have to length.) Examples are manufacturing participation in R&D and vice versa [97], closeness of vendors and users [88] and graduate internship in industry [82].

One can only receive and digest a limited amount of information over a given time period. Hence, the person promoting the movement of technology should be aware of the capabilities of the receiver.

The duration and the frequency of transfer are important variables in the transfer of technology [96]. The success of transfer depends on the frequency with which the idea is presented to receivers through publications, conferences, oral presentations, etc. (comparison with advertising).

Communication has to occur in both directions in order for the transfer of technology to be successful [96]. The two-way communication opens the possibility for feedback which results in the establishment, clarification,
be good; the feedback causes and adjustments of goals and objec-
tive s of both parties.

The above comparison shows that the transmitter-channel-receiver model is helpful in explaining and detecting problems encountered in the transmission of technology. However, it does not take into account explicitly other such important variables in the process of transmission as personal objectives and goals, motivation and behavioral aspects of the individuals involved in the transmission, etc. Important questions like "Why did he not turn on or tune his receiver" do not fit this model. However, the concepts originating from this model still prove useful, although they can be used only to a limited extent.

In the next chapter, the above definitions, ideas and models will be applied to the case of a doctoral student doing his dissertation. It will appear that the process that leads to a dissertation is similar to the process of technological change and, thus, has many of its characteristics. In this analysis, special attention will be paid to problems in the movement of technology encountered by the student.
Chapter 20. Technological Change and the Student

In the previous chapter it was shown that the process of technological change proceeds in three stages - invention, innovation and consequences. In the majority of these cases, the student doing dissertation research does not finish the complete sequence in the process of technological change, but completes stages one or two, paying little or no attention to the last. However, viewing the efforts of the student on the road to his dissertation as a process of technological change can provide insight into the problems encountered by the student and may give indications as to where some of the solutions to the problems encountered in this partial completion of the process of technological change are located.

The first stage - the process of invention - is always completed, for almost all universities require that their doctoral students carry out some original research [85]. This leads to the creation or development of a new idea or concept and, hence, the result can be classified as an invention according to the previous definitions.

The process which leads to a dissertation topic is mostly a process of vertical diffusion. This is because in a majority of the cases the student is working within the professional infrastructure of the university, and the university facilitates the confluence of all information obtained through courses, literature, meetings, etc., leading to a project for a dissertation. In this process of vertical diffusion, the movement of knowledge, technology, and ideas from the knowledge pool is promoted by the short channel (same office, same building) between the knowledge pool and the receivers (faculty-student,
library-student, student-student) which especially improves frequency of communication and feedback.

In the case of a student/inventor, the distinction between invention and innovation is not very sharp because the inventor and the innovator is the same person. Usually, the process of technological change halts somewhere after the invention stage, leaving subsequent stages incomplete. In the process of innovation, one would need to take the technical, industrial and commercial steps that would lead to, for instance, a new product. Frequently, universities are not well-equipped to take these necessary steps. In contrast, industries have the know-how, experience, capability of products, and the motivation (competition and survival) necessary to successfully complete the innovation stage. It could be advantageous for the student to be able to make use of these resources, which can become available to him through the process of transfer and diffusion.

One of the main reasons that the problems arise in the movement of technology and knowledge from industry to student, and vice versa, is caused by differences in goals and objectives of both parties. The goals and objectives of the student are primarily education, obtaining a degree, adding to the pool of knowledge and doing research, and preparing for a future job; industry, however, is primarily concerned with profitability and, hence, production and sales. Both parties tend to focus on their own goals and objectives and feel that they do not necessarily need each other to satisfy these. It is important to note that the university generally does not require the student to promote or complete the movement of the dissertation
research to industry: for one reason, the process would involve high risks (flunking) and would be extremely time-consuming to the student. The only contribution to the movement required by the university is within its professional infrastructure - the writing and defense of a dissertation. Hence, different motives must play a role if the student is promoting movement of knowledge and technology to industry. This decision as to whether or not to promote movement to industry is usually determined by a trade-off between benefits and costs. The benefits could be educational, the desire to make use of industrial experience, possible financial gain, merely ideological, and the satisfaction of success; the costs would be time, effort, money, and the possibility of eventual failure.

Should the student try to promote movement to industry, he would be seeking an overlap of his goals and objectives with those of industry. Thus, he would 1) have to be interested in the problems which industry wants to solve or 2) have to find problems in industry for which his solution method is suitable. An example of the former case is industry's funding of universities in the form of grants to faculty which subsequently are offered to students in the form of research assistantships which may lead to a dissertation project. In the latter case, however, the student is required to spend time and effort in order to locate suitable problems in industry for which his solution method is appropriate. It is clear that this is a difficult and time-consuming task for the student, and in this instance, the help of the faculty is indispensable since the faculty has the experience, the status, the connections, etc., to best accomplish
the task.

Once an overlap in goals and objectives has been established, one would like to promote the movement of technology and knowledge from university to industry, and vice versa: this movement can take place through transfer and diffusion. In the next section, the different methods for promoting such movement, and the advantages and disadvantages of each will be examined.

**Student-Industry: Transfer and Diffusion**

The characteristics of five different methods for promoting the transfer and diffusion of knowledge and technology between student and industry will be discussed in the following.

1. **Obtaining funding from industry**

At the beginning of a project, the motivation to start the process of movement to industry is emphasized by possible industrial funding, which provides necessary development expenses as well as bread and butter for the student. This experience of seeking and obtaining funds is important to the student, but chances of success are slim. The transient status of the student as well as the indeterminate and not yet well-defined state of the project might be of concern to the industry. Moreover, in many cases the process is slow and the student cannot wait for the funding to take place [90]. If industry's funding is obtained, the industry has displayed an active interest in the project by providing the funds for it. Hence, a professional infrastructure has been established in which both the funder and the researcher would like to promote the process of diffusion so that their understanding will be improved and their benefits maximized.
However, the established overlap of goals and objectives of the student and Industry is only a small subset of the respective goals and objectives of both parties. Since both student and Industry are still working within their professional infrastructures, the goals and objectives of their organizations which form a much larger set tend to be emphasized. Also, frequently, both the funder and the researcher become lax in the transfer and diffusion process. The researcher left alone "feels" that he can get more done. The funder may not need the results to solve his immediate problems and does not press for progress reports. In this case, expectations of both parties can diverge to the point where the results of the research are unsatisfactory to the funder and funding ends. The commonality of goals can only be maintained, updated, and reinforced through frequent communication. However, geographical separation (long channel) makes frequent communication difficult and, thus, inhibits the diffusion of knowledge and technology.

2. Internship in industry

The movement of knowledge and technology through internship in industry can be seen as a mixture of transfer and diffusion. The actual move of the student into the industrial environment is clearly a process of transfer which significantly shortens the student-industry channel. The process of diffusion begins as the student works within the professional infrastructure of Industry. The rate of the diffusion is enhanced in both directions, i.e. student-to-Industry and Industry-to-student, as there is now commonality of goals since both the industry and the student are concerned with the student's
progress in a specific area. In the limited domain around the student, the academic and the industrial approach to problem solving are incorporated into a single solution. Thus, internship is perhaps the most efficient means of diffusion of ideas and methods. Unfortunately, internships are not easily obtained unless a program with industry exists already within the university. Moreover, it is infrequent that the participating industry has problems which can be approached suitably by the student's solution method. The reasons behind this are two-fold: first, the student's solution method often is not yet developed at the time the question of internship becomes relevant and, thus, it is difficult to match this with specific industrial problems and, secondly, the areas of concern of an industry associated with a particular university do not necessarily encompass the specific fields of interest of all students doing research in the discipline.

The characteristics of internship are quite similar to those of funding which were previously discussed. Here, also, are problems of obtainability, as well as the trade-off between benefits and costs. Moreover, the process of the movement of technology and the problems associated with this are quite similar to those of funding. The main difference is that the industrially funded research is focused towards satisfying academic goals, whereas the work of an intern is directed more towards satisfying industrial goals. In either case, frequent communication to match different goals and objectives remains indispensable.

3. Publications and conferences

The standard way in which universities contribute to the
movement of knowledge and technology is by publishing in professional journals and/or presenting papers at conferences. It is important to note that, in many cases, this contribution to the movement is sufficient for recognition. Although publications and presentations of research results can establish an overlap of industry and student work, they do not necessarily make both parties aware of the existence of such an overlap. In other words, new technology and knowledge is deposited in the knowledge pool and left alone by the researcher; independently, prospective users search through this pool for information suitable for their needs. Hence, the existing structure of publishing in journals and presentation at conferences is not a very effective channel for the movement of technology to the specific user, since no direct transfer and diffusion to this user takes place as does occur in the cases of funding and internship.

4. Finding a job with a particular industry

Obtaining a job after graduation with the industry of interest is an effective way to promote the movement of knowledge and technology through transfer and diffusion. Similar to the case of internship, transfer consists of the move to industry, and diffusion takes place through working in the professional infrastructure of the industry. However, in contrast to the case of internship, no benefits are obtained from the industry during the process of innovation at the university. Moreover, the different goals and objectives of the student are clearly time-separated - first, university, then industry goals. Hence, no conflicts of goals of university and industry arise.
5. **Selling Innovation to Industry**

The successful termination of a project can yield a product or device that might be of interest to Industry. Supposedly, the dissertation has shown the value and the feasibility of the product or device by way of successful analytical and experimental work on an evolved model or prototype. Hence, one is better equipped to convince Industry of its use and, in turn, Industry is better able to evaluate its potential.

The nonlinear multifunction module based on the nonlinear pulse width modulation method described in Parts I and II of this dissertation is at this stage. For successful transfer and diffusion, however, there are still many barriers to overcome. For instance, it is difficult to evaluate market potential of the product, and problems in manufacturing are difficult to foresee. Moreover, it appears that patenting is necessary for successful negotiations, because Industry is most interested in ideas with patent protection [89] in order to avoid problems of ownership and duplication. However, with the new pending patent laws, it will become more difficult to obtain a patent [92]: it will mean additional costs to the inventor that can only be outweighed by expected benefits.
Conclusion

This part of the dissertation discussed the problems of movement of technology as encountered by the student. It was shown that the process which leads to a dissertation can be seen as a process of technological change. In this process of technological change, the movement of knowledge and technology from industry to student, and vice versa, is very important. This movement takes place through transfer and diffusion. The characteristics of the transfer and diffusion process were investigated for the following cases: 1) Obtaining funding from industry; 2) Internship in industry; 3) Publications and conferences; 4) Finding a job with a particular industry; 5) Selling the innovation to industry. It was concluded that the two major barriers to successful transfer and diffusion were lack of frequent communication and different goals and objectives of industry and university.


APPENDIX I. COMPUTATION OF THE ENERGY OF THE RIPPLE

For the energy of the ripple the following equation can be derived (See equation 7.3):

\[ E_o = \sum_{n=1}^{\infty} \frac{(A-B)^2}{\pi^2 \frac{(2\pi t)^{2k}}{T}} \frac{1}{2k+2} \left(1-\cos \frac{2\pi n x}{T} \right) \]

The sum of this series can be computed. It is given by [51]:

\[ E_o = \frac{(A-B)^2}{\pi^2 \frac{(2\pi t)^{2k}}{T}} \left[ \frac{2}{(2k+2)!} B_{2k+2} \right] - \frac{(-1)^k (2\pi)^{2k+2}}{2(2k+2)!} B_{2k+2} \left(\frac{t}{T}\right) \]

\[ B_{2k+2} \text{ is the 2k+2 Bernoulli number} \]

\[ B_{2k+2} \left(\frac{t}{T}\right) \text{ is the 2k+2 Bernoulli polynomial} \]

The Bernoulli numbers and polynomials for k = 0, 1 and 2 are:

\[ k = 0 \quad B_2 = 1/6 \quad B_2(x) = x^2 - x + 1/6 \]
\[ k = 1 \quad B_4 = -1/30 \quad B_4(x) = x^4 - 2x^3 + x^2 - 1/30 \]
\[ k = 2 \quad B_6 = 1/42 \quad B_6(x) = x^6 - 3x^5 + \frac{5}{2}x^4 - 1/42 \]

Simplifying equation A1.1 yields:

\[ E_o = \frac{2(A-B)^2}{(\pi^2 \frac{(2\pi t)^{2k}}{T})^{2k+2}} \left[ B_{2k+2} \right] - \frac{(-1)^k (2\pi)^{2k+2}}{2(2k+2)!} \frac{t}{T} \]

From the above equation it can be seen that the energy in the ripple is a function of the pulse width \( t_1/T \). The maximum energy of the ripple can be obtained by differentiating the above equation with respect to \( t_1/T \). In order to compute this derivative, the following properties of Bernoulli numbers are used [51]:

\[ \frac{d}{dx} B_k(x) = kB_{k-1}(x) \quad (A1.3a) \]
\[ B_k(1-x) = (-1)^k B_k(x) \quad (A1.3b) \]
Substituting $x = 1/2$ yields:

$$B_k(1/2) = (-1)^k B_k(1/2) \quad \text{(A1.3c)}$$

or:

$$B_k(1/2) = 0 \text{ for } k \text{ odd} \quad \text{(A1.3d)}$$

The derivative of equation A1.2 is given by:

$$\frac{dE}{dt} = C(2k+2)B_{2k+1}(\frac{t}{T}) \quad C - \text{constant}$$

Using equation A1.3d it is seen that the first derivative is zero for $t/T = 1/2$. This proves that in the case that the pulse width modulated signal is a dc component plus a square wave, the energy in the ripple of the output signal will be maximal. This maximal energy is given by:

$$E_{max} = \frac{2(A-B)^2}{(\frac{t}{T})^{2k+1}(2k+2)!} \left[ |B_{2k+2}| (\frac{t}{T})^{2k+1} \right] \quad \text{(A1.4)}$$

In simplifying this equation the following properties of Bernoulli numbers are used [51,52]:

$$B_k(1/2) = -(1-2^{1-k})B_k$$

or:

$$B_{2k+2} = -(1-2^{1-k})B_{2k+2}$$

and

$$B_{2k} = (-1)^{k-1}B_{2k}$$

or:

$$|B_{2k+2}| = (-1)^{k}B_{2k+2}$$

With these properties, equation A1.4 becomes

$$E_{max} = \frac{4(A-B)^2}{(\frac{t}{T})^{2k}(2k+2)!} |B_{2k+2}| (1-2^{1-(2k+2)})$$
Using this equation it is possible to compute the energy of the left over ripple for a pulse width modulated signal that consists of a dc component plus a square wave. It is seen that the ratio of the time constant of the filter and the period of the carrier signal, \( \tau / T \) should be made small. Moreover, increasing the order \( k \) of the filter reduces the ripple significantly.
APPENDIX II. CONVERSION FROM DIFFERENCE TO DIFFERENTIAL EQUATION

In this appendix the continuous time equivalent of the difference equation that describes the response of the system will be derived (equation 10.9a). This equation was:

\[ w_n = \frac{a_w}{a_x} \left( \frac{x(t_n)}{K_x} \right) + \frac{K_w}{1+a_w T K_w} \left( \frac{x(t_n)}{K_x} \right) \]  

(A11.1)

In order to find the equivalent differential equation, we will consider the following differential equation:

\[ \frac{dw(t)}{dt} = f(t)w(t) + g(t) \]  

(A11.2)

We will derive an equivalent difference equation for this differential equation, which will turn out to look like equation A11.1. Comparing the coefficients of the two difference equations will give the values for \( f(t) \) and \( g(t) \).

The general solution of the differential equation A11.2 is [53]:

\[ w(t) = e^{\int_{t_0}^{t} f(\tau) d\tau} \left( w(t_0) + e^{\int_{t_0}^{t} g(\tau) d\tau} \right) \]

Computing the values of the output signal \( w(t) \) at \( t = (k+1)T \) and \( t = kT \) yields:

\[ w((k+1)T) = e^{\int_{t_0}^{(k+1)T} f(\tau) d\tau} \left( w(t_0^+) + e^{\int_{t_0^+}^{(k+1)T} g(\tau) d\tau} \right) \]

(A11.3)
\[
\begin{align*}
\int_{0}^{t} f(\tau) d\tau & \quad \int_{0}^{t} f(\tau) d\tau - \int_{0}^{t} f(\tau^*) d\tau^* \\
\text{Multiply equation A11.4 with:} \quad \int_{kT}^{(k+1)T} f(\tau) d\tau \\
\text{and subtract equation A11.4 from A11.3. This yields:} \\
\int_{kT}^{(k+1)T} f(\tau) d\tau & \quad \int_{0}^{t} f(\tau) d\tau - \int_{0}^{t} f(\tau^*) d\tau^* \\
\text{Define } \phi[(k+1)T,kT] = e^{+1} \\
\text{Then the above equation becomes:} \\
\int_{kT}^{(k+1)T} f(\tau) d\tau & \quad \int_{0}^{t} f(\tau) d\tau - \int_{0}^{t} f(\tau^*) d\tau^* \\
\text{Assume that } f(\tau) \text{ is essentially constant over the interval } kT < \tau < (k+1)T, \\
\text{so:} \\
f(\tau) & \approx f(kT) \\
\text{then equation A11.5 becomes:} \\
\int_{kT}^{(k+1)T} f(\tau) d\tau & \quad \int_{0}^{t} f(\tau) d\tau - \int_{0}^{t} f(\tau^*) d\tau^* \\
\end{align*}
\]
\[ w((k+1)T) = e^{f(kT)T}w(kT) + \int_{kT}^{(k+1)T} e^{f(kT)[(k+1)T-\tau]} g(\tau) d\tau \quad kT \leq \tau \leq (k+1)T \]

If it is assumed that \( g(\tau) \) is essentially constant over the interval \( kT \leq \tau \leq (k+1)T \), then the above equation becomes:

\[ w((k+1)T) = e^{f(kT)T}w(kT) + e^{f(kT)} (k+1)T g(kT) \quad kT \leq \tau \leq (k+1)T \]

Comparing the coefficients of equations All.1 and All.6 gives:

\[ e^{f(kT)T} = \frac{\alpha_w T^\tau_i K_w [x(t_{x_k})/K_{x_w}]^x}{1 + \alpha_w T^\tau_i K_w [x(t_{x_k})/K_{x_w}]^x} \]

or \( f(t) = \frac{\alpha_w T^\tau_i K_w [x(t)/K_{x_w}]^x}{1 + \alpha_w T^\tau_i K_w [x(t)/K_{x_w}]^x} \)

and \( g(t) = \frac{\alpha_w T^\tau_i K_w [x(t)/K_{x_w}]^x}{1 + \alpha_w T^\tau_i K_w [x(t)/K_{x_w}]^x} \)

Substitution of the equations for \( f(t) \) and \( g(t) \) results in the differential equation:
\[
\frac{dw(t)}{dt} = \ln \frac{\alpha \tau \alpha x / \alpha w}{1 + \alpha \tau \alpha x / \alpha w} \left[ w(t) - K_w \left( \frac{x(t)}{K_x} \right)^{\alpha_w / \alpha_x} \right]
\]

With \( t' = t/T \) this becomes:

\[
\frac{dw(t')}{dt'} = \ln \frac{\alpha \tau \alpha x / \alpha w}{1 + \alpha \tau \alpha x / \alpha w} \left[ w(t') - K_w \left( \frac{x(t')}{K_x} \right)^{\alpha_w / \alpha_x} \right]
\]
APPENDIX III. RESPONSE OF A FIRST ORDER SYSTEM TO A SQUARE WAVE

In this chapter the response of a first order system with a time constant $1/\alpha$ will be investigated. The results that will be obtained will be used to determine the accuracy of the carrier signal and, hence, the accuracy of the system. The square wave that is fed into the first order filter is defined as (See also Fig. AIII.1):

$$\text{Sq}(t) = \begin{cases} 
P & nT < t < nT+L \\
Q & nT+L < t < (n+1)T \end{cases}$$

The response of the filter can be calculated as follows:

$$0 < t < T \quad 0(t) = Q + (P-Q)(1-e^{-\alpha t}) \quad 0(0) = 0$$

$$L < t < T \quad 0(t) = Q + (P-Q)(1-e^{-\alpha t}) - (P-Q)(1-e^{-\alpha(t-L)})$$

$$T < t < T+L \quad 0(t) = Q + (P-Q)(1-e^{-\alpha t}) - (P-Q)(1-e^{-\alpha(t-L)}) + (P-Q)(1-e^{-\alpha(t-T)})$$

etc.

In the following derivation we will compute the response of the system in the time intervals $nT < t < nT+L$ and $nT+L < t < (n+1)T$, after which we investigate the limit case $n \to \infty$ to obtain the steady state response. Because the derivations are similar for both intervals, only the derivation for the interval $nT < t < nT+L$ will be given. In this interval the output signal is given by:

$$0(t) = Q + (P-Q)(1-e^{-\alpha t}) + (P-Q)(1-e^{-\alpha(t-L)}) + \cdots + (P-Q)(1-e^{-\alpha(t-nT)})$$

$$- (P-Q)(1-e^{-\alpha(t-L)}) - \cdots - (P-Q)(1-e^{-\alpha(t-(n-1)T-L)})$$

$$0(t) = P -(P-Q)e^{-\alpha t} (1 + e^{-\alpha T} + e^{2\alpha T} + \cdots + e^{n\alpha T})$$

$$+ (P-Q)e^{-\alpha(t-L)} (1 + e^{\alpha T} + e^{2\alpha T} + \cdots + e^{(n-1)\alpha T})$$
Fig. AIII.1 Square wave signal
Computing the sum of the series yields:

\[ O(t) = P - (P-Q)e^{-\alpha t} \frac{1-e^{\alpha (n+1)T}}{1-e^{\alpha T}} + (P-Q)e^{-\alpha (t-L)} \frac{1-e^{\alpha T}}{1-e^{\alpha T}} \]

\[ O(t) = P - (P-Q) \frac{e^{-\alpha t}}{1-e^{\alpha T}} [e^{\alpha L} - (1-e^{\alpha (n+1)T})] \]

Let \( t = nT+t^\star \), then for \( 0 < t^\star < L \) the following holds

\[ O(t^\star) = P - (P-Q) \frac{e^{-\alpha t^\star}}{1-e^{\alpha T}} [e^{\alpha (L-nT)} - e^{\alpha L} - nT+e^{\alpha T}] \]

Now, let \( n \to \infty \), to obtain the response after the transient has died out and subtract the response of the system from the input signal to obtain the carrier signal:

\[ D(t^\star) = S\tau(t^\star) - O(t^\star) \]

or

\[ D(t^\star) = -(P-Q) \frac{e^{-\alpha t^\star}}{e^{\alpha T}-1} e^{-\alpha t^\star} \quad 0 < t^\star < L \quad (AIII.1) \]

A similar derivation for \( L < t^\star < T \) gives:

\[ D(t^\star) = (P-Q) \frac{e^{\alpha T} - e^{\alpha (T-L)}}{e^{\alpha T}-1} e^{-\alpha (t^\star-L)} \quad L < t^\star < T \quad (AIII.2) \]

For \( L = 1/2T \), the above two equations describe the ideal carrier signal:

\[ D(t^\star) = -Ke^{-\alpha t^\star} \quad 0 < t^\star < 1/2T \]
\[ D(t^\star) = Ke^{-\alpha (t^\star-T/2)} \quad 1/2T < t^\star < T \]

However, it can be seen that some inaccuracy is caused by an unwanted factor with which the carrier signal is multiplied by, namely:

\[ K = \frac{e^{\alpha T} - e^{-1/2\alpha T}}{e^{\alpha T}-1} \approx 1-e^{-1/2\alpha T} \quad \text{for } \alpha T \text{ large} \]

The factor can be neglected if \( \alpha T \) is large.
APPENDIX IV.

The fluidic devices that were used in the construction of the system (Fig. 18.2) were Corning FICM (Fluidic Industrial Control Modules) fluidic devices which were mounted on a manifold. The supply pressure of all fluidic devices was 6.5 psl (448 mbar). Only the supply pressure of the last digital amplifier was increased to 15 psl (1034 mbar). A list of the devices is given below.

- **DA_1** - digital amplifier, type 191460
- **DA_2** - digital amplifier, type 191452
- **FF_1**
- **FF_2** - flip-flop, type 191454
- **FF_x**
- **FF_w**
- **OP_o** - 2 stage operational amplifier, no number available
- **OP_x** - 3 stage operational amplifier, no number available
- **OP_w**
- **R** - resistor: 3.62x10^7 ./insec. (143x10^7 ./msec.)
- **R_t** - resistor: 0.26x10^7 ./insec. (10.2x10^7 ./msec.)
- **R_t** - resistor: 4.05x10^7 ./insec. (159x10^7 ./msec.)
- **V_o** - volume: 57.7 in^3 (946 cm^3)
- **V_x** - variable volumes
- **V_w**