A STUDY OF THE EFFECTS OF LINEAR NETWORKS ON FM WAVES

by

Preston Benton Johnson

Thesis Submitted to the Graduate Faculty

of the Virginia Polytechnic Institute

in Candidacy for the Degree of

Doctor of Philosophy

in

Electrical Engineering

APPROVED:

Dr. Harry K. Ebert, Jr.
Committee Chairman

Dr. W. W. Cannon

Dr. R. H. Miller

Dr. M. Y. Rhee

Prof. Ralph R. Wright

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

When a linear network whose transmission characteristics are nonlinear functions of frequency is excited by a frequency-modulated wave, the response will generally be a distorted "hybrid" wave. A hybrid wave is one which exhibits both amplitude modulation and frequency modulation. The amplitude modulation arises from the fact that the sideband components of the FM wave are subjected to different attenuation due to the frequency-dependent attenuation characteristic. If, in addition, the attenuation characteristic is nonlinear, the amplitude modulation will also be distorted. Generally speaking, the presence of amplitude modulation, distorted or not, is not considered to be of great concern since most FM systems utilize amplitude limiters to suppress this characteristic.

Of primary importance is the phase distortion generated in the FM wave due to a nonlinear phase characteristic. The distortion arises because each sideband undergoes a phase shift which is not linearly related to the phase shifts of the other sideband components. If the network phase shift is a linear function of frequency, the output wave will be shifted with respect to the input wave but no distortion will result. This will be illustrated during the course of this investigation. Since the output of the FM discriminator is proportional to the derivative of the phase of the applied wave, any distortion in the phase will result in distortion of the detected signal. Furthermore, one cannot easily compensate for this distortion, so it is highly desirable that it be avoided.

The analysis of distortion generated in FM waves by linear networks has been the subject of considerable research dating back to the 1930's when frequency modulation was shown to be a practical and sometimes desirable means of transmitting intelligence. There are essentially two basic approaches to the analytical analysis of the distortion. The first of these is known as the Fourier method in which the frequency-modulated wave is broken up into its Fourier spectrum. Each of these "sidebands" is then operated on by the network, and the output wave is then obtained by taking the vector sum of the "weighted" sidebands. This method is simple in concept and straightforward in application. It is an example of the powerful superposition principle used so often in engineering to solve very complex problems. The chief disadvantage of the Fourier method is the extremely large number of computations involved whenever the modulation index is large (resulting in a large number of sidebands). In this case, the task is simply too formidable to be done by hand. In the past, the Fourier method has usually been abandoned at this point in favor of other schemes.

The other approach which has been employed to analyze FM distortion is the so called "Quasi-steady-state" or "Quasi-stationary" method. In this method the frequency-modulated wave is considered to be varying slowly enough in frequency that the transients caused by the fact that the frequency is varying are negligible. This method is particularly applicable at very low modulating frequencies with respect to the frequency deviation (large values of modulation index). As the frequency of the modulating wave is allowed to become higher, the

assumption that the transient components are negligible becomes less acceptable and correction factors must be added to yield an accurate response. These correction factors usually appear as an infinite series of very complex terms, but supposedly one can obtain the desired degree of accuracy by including more and more of these terms.

In the following chapters the two methods given above will be explored with the main emphasis on developing digital computer techniques for application to the Fourier method. It will be shown that the use of the digital computer now makes it practical to use the Fourier method even for relatively large values of modulation index. This method will be extended to apply to modulation by a pair of sinusoidal waves.

A fairly simple form of the Quasi-steady-state method is also programmed on the digital computer. Although an extended rigorous investigation of this method is not given, certain characteristics inherent in the analysis are pointed out for the sake of comparison with the Fourier method. The comparison of the two methods features an application to a reasonably complex multi-section filter.

The experimental verification of the theoretical predictions is obtained by analog simulation techniques. This rather novel approach is found to have some distinct advantages over more conventional techniques.

Digital computer programs written in the FORTRAN IV language for processing on the IBM 7040/1401 system are given for each method of analysis along with various supporting programs. Some of these

supporting programs are capable of being applied to other problems since they are complete and independent on the main program. Among these are programs for computing Bessel functions of the first kind and for computing the phase characteristic of a network when the attenuation characteristic is known.

II. THE REVIEW OF LITERATURE

The first publication of significance to treat in detail the problem of distortion in frequency-modulated systems was the classical work of Carson and Fry [1][1] published in 1937. This paper includes a rigorous mathematical treatment of the theory involved when waves with varying frequency are applied to electrical networks. It furnishes the mathematical foundation for the "Quasi-steady-state" method and expresses the correction terms required by this method in terms of an infinite series. Further investigation of this method was done by Van der Pol [2] and Stumpers [3]. Stumpers limited his analysis to periodic frequency modulation and succeeded in demonstrating the convergence properties of his correction-term series as well as for the series of Carson and Fry. Baghdady [4,5] summarized the series of Carson and Fry, Van der Pol, and Stumpers and then presented his own development. Rowe [6] pointed out certain errors in the analysis of Baghdady that appeared in both [4] and [5]. Other investigators [7-10] made significant contributions to the further development of the Quasi-steady-state analysis.

The Fourier method of analysis has received much less attention in the literature than has the Quasi-steady-state method. One of the earlier publications was due to Frantz [11] who developed an elaborate scheme for manually calculating the output. Following closely the work

---

[1] Numbers appearing in brackets refer to the Bibliography.

of Cherry and Rivlin [12], Frantz's method required that the network response be expressible in terms of a finite power series or trigonometric series. The following procedure is also outlined by Frantz as a rigorous and general method of obtaining the output-voltage envelope for a network excited by a frequency-modulated wave:

1. Measure or calculate the steady-state amplification versus frequency and phase-shift versus frequency characteristics of the network in question.
2. Express the frequency-modulated input wave in terms of its steady-state spectrum of sinusoidal side frequencies.
3. Pass the individual side frequencies through the network, altering the amplitude and phase of each according to the steady-state amplification and phase characteristics of the network at the particular frequency of the side frequency.
4. Plot the sum of the altered side frequencies point by point to obtain the output wave as a function of time.
5. Draw a smooth curve through the carrier voltage peaks to obtain the response envelope of the output voltage.

Frantz goes on to state, "The method is rigorous and general, but not practical." This conclusion was based on the assumption that the calculations involved in applying this method were to be performed by hand. The development and availability of modern high speed digital computers renders this judgement rather harsh. At this point Frantz leaves the "rigorous and general" approach and concentrates on developing approximations to be applied in more of a "pseudo-Fourier"

method.

Certain other investigators [13,14] have applied this pseudo-Fourier method to networks for special cases where the departure from linearity of the network characteristics was small. However, no references were found which indicate that the exact Fourier method has been adapted for solution on a digital computer.

III. THE FOURIER METHOD OF ANALYSIS

Mathematical Development

Consider the equation of an angle modulated wave

$$e(t) = E \sin \Phi(t) \qquad \text{(III-1)}$$

where E is a constant independent of time and $\Phi(t)$ is the phase angle which is assumed to vary with time in accordance with the modulating signal. In the case of frequency modulation, which is classified as a special case of angle modulation, the frequency of the sine function in equation (III-1) needs some clarification. Since the argument of the sine function is continuously changing, to speak of frequency in the conventional sense has little meaning. To avoid this ambiguity, it has been the practice to introduce a quantity known as the "instantaneous frequency" defined as the time derivative of the phase angle, just as angular velocity is defined as the time derivative of angular displacement. Mathematically, the instantaneous frequency is

$$\omega_i = \frac{d\,\Phi(t)}{dt} . \qquad \text{(III-2)}$$

In frequency-modulated systems the frequency of the carrier wave is varied by the modulating signal and can be represented as

$$\omega_i = \omega_c + s(t) \qquad \text{(III-3)}$$

where ω_c is the frequency of the unmodulated carrier and s(t) is proportional to the modulating signal. Substituting (III-3) into

(III-2) and solving for Φ (t) yields

$$\Phi(t) = \int_o^t \omega_i dt + \varphi_o$$

$$= \int_o^t [\omega_c + s(t)]\, dt + \varphi_o$$

$$= \omega_c t + \int_o^t s(t)dt + \varphi_o . \qquad \text{(III-4)}$$

In this equation φ_o is the phase angle of the carrier at t = 0 as required in the case where s(t) is always zero and $\Phi(t) = \omega_c t + \varphi_o$. The equation for the frequency-modulated wave can now be written as

$$e(t) = E \sin \left[\omega_c t + \int_o^t s(t)dt + \varphi_o \right] . \qquad \text{(III-5)}$$

This equation, which is applicable for any modulating signal, is general and will be referred to several times later in this paper.

Modulation by a Single-Frequency Sinusoid

As the first study in this analysis, the very important case of single-frequency sinusoidal modulation will be considered. This development is especially important to some of the later distortion analyses. Let the modulating wave be

$$e_m(t) = K \cos \omega_m t \qquad \text{(III-6)}$$

where K is the peak value of the modulating wave and ω_m is the angular frequency of the modulating wave. Since s(t) in equation (III-3) is proportional to the modulating signal, then

$$s(t) = \Delta\omega \cos \omega_m t \qquad \text{(III-7)}$$

where $\Delta\omega$ is known as the "maximum frequency deviation" and is linearly related to K by the modulation circuit constants. Substituting this value of s(t) into Equation (III-5) yields

$$e(t) = E \sin \left[\omega_c t + \frac{\Delta\omega}{\omega_m} \sin \omega_m t + \varphi_o \right]$$

$$= E \sin \left[\omega_c t + m_f \sin \omega_m t + \varphi_o \right] . \qquad \text{(III-8)}$$

The quantity $m_f = \Delta\omega/\omega_m$ is termed the "modulation index" and is large for low modulation frequencies and small for high modulation frequencies. In the following analysis φ_o will be assumed to be zero. No loss of generality occurs because of this simplification. Thus,

$$e(t) = E \sin \left[\omega_c t + m_f \sin \omega_m t \right] . \qquad \text{(III-9)}$$

This expression can be expanded by trigonometric identities to yield

$$e(t) = E \left[\sin \omega_c t \cos (m_f \sin \omega_m t) + \cos \omega_c t \sin (m_f \sin \omega_m t) \right] . \qquad \text{(III-10)}$$

This equation can be further expanded in terms of Bessel functions by making use of the following identities:

$$\cos (x \sin y) = J_o(x) + 2 \sum_{K=1}^{\infty} J_{2k}(x) \cos 2ky; \qquad \text{(III-11)}$$

$$\sin (x \sin y) = 2 \sum_{k=1}^{\infty} J_{2k-1}(x) \sin (2k-1)y. \qquad \text{(III-12)}$$

Making the substitutions suggested above and introducing the normalized variable $F(t) = e(t)/E$, Equation (III-10) can now be written

$$F(t) = \sin \omega_c t \left[J_o(m_f) + 2 \sum_{k=1}^{\infty} J_{2k}(m_f) \cos 2k \omega_m t \right]$$
$$+ \cos \omega_c t \left[2 \sum_{k=1}^{\infty} J_{2k-1}(m_f) \sin (2k-1) \omega_m t \right]$$
$$= \sum_{k=1}^{\infty} \left[J_o(m_f) \sin \omega_c t + 2 J_{2k}(m_f) \sin \omega_c t \cos 2k\omega_m t \right.$$
$$\left. + 2 J_{2k-1}(m_f) \cos \omega_c t \sin (2k-1) \omega_m t \right] . \qquad \text{(III-13)}$$

Recombining the trigonometric terms yields

$$F(t) = \sum_{k=1}^{\infty} \left\{ J_o(m_f) \sin \omega_c t + J_{2k}(m_f) \left[\sin (\omega_c + 2k\omega_m) t \right.\right.$$
$$\left. + \sin(\omega_c - 2k\omega_m) t \right] + J_{2k-1}(m_f) \left[\sin \left\{ \omega_c + (2k-1)\omega_m \right\} t \right.$$
$$\left.\left. - \sin \left\{ \omega_c - (2k-1) \omega_m \right\} t \right] \right\} . \qquad \text{(III-14)}$$

Expansion of this equation yields the Fourier spectrum of the FM wave. In order to more fully appreciate the properties of this spectrum, the expansion will be carried out for a few terms.

$$F(t) = J_o(m_f) \sin \omega_c t + J_1(m_f) \left[\sin (\omega_c + \omega_m) t - \sin (\omega_c - \omega_m) t \right]$$
$$+ J_2(m_f) \left[\sin(\omega_c + 2\omega_m) t + \sin (\omega_o - 2\omega_m) t \right]$$
$$+ J_3(m_f) \left[\sin(\omega_c + 3\omega_m) t - \sin (\omega_c - 3\omega_m) t \right]$$
$$+ J_4(m_f) \left[\sin(\omega_c + 4\omega_m) t + \sin(\omega_c - 4\omega_m) t \right]$$
$$+ \ldots . \qquad \text{(III-15)}$$

The spectrum of this FM wave is found in Figure 1 for a modulating frequency of 1000 hertz and a peak frequency deviation of 2000 hertz.

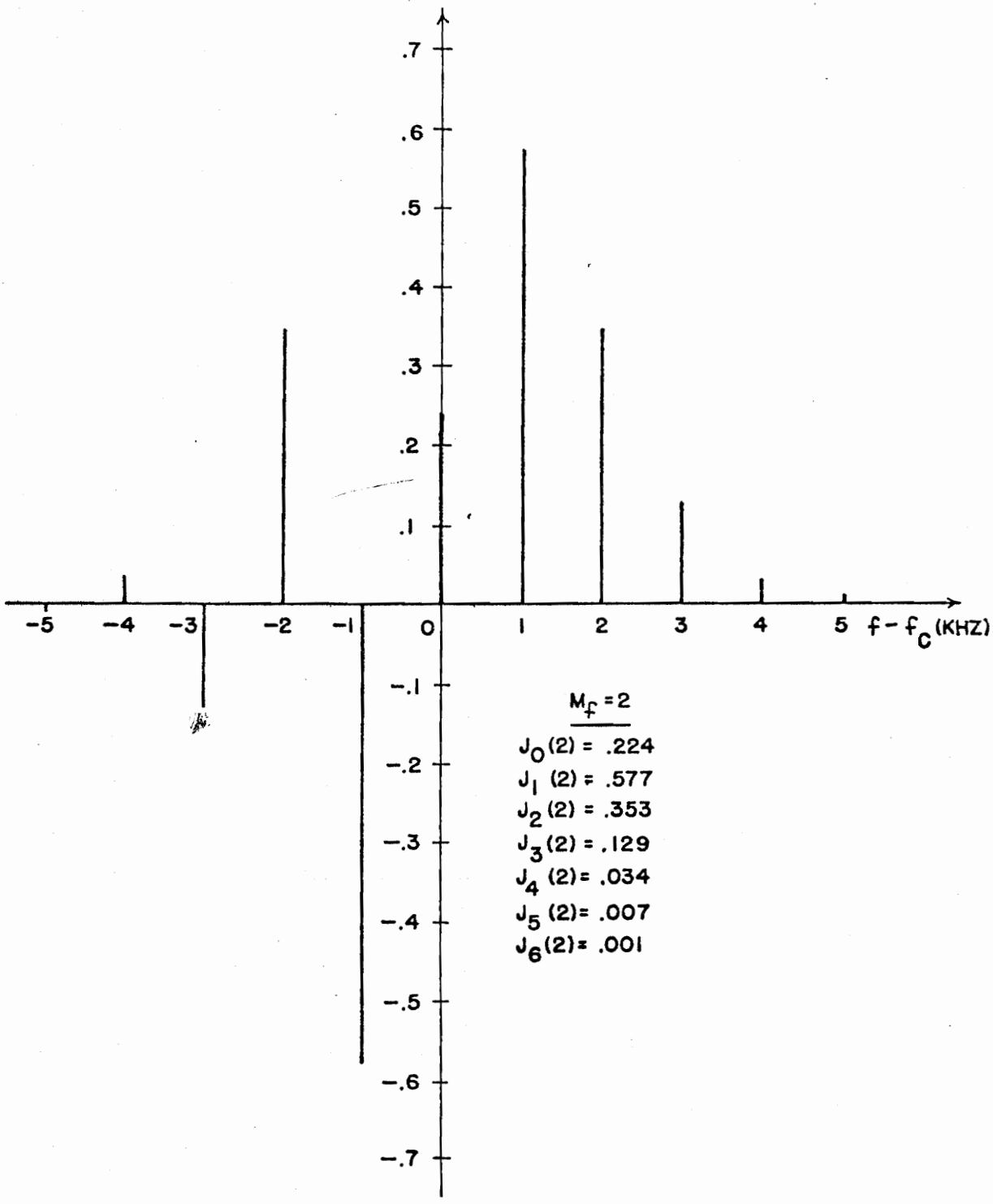


FIGURE I — SPECTRUM OF FM WAVE GIVEN BY EQUATION (III-15) WITH $\omega_M = 2\pi \times 1000$ AND $\Delta\omega = 2\pi \times 2000$

Equation (III-15) can be expressed in a more compact form by noting that $J_n(X) = (-1)^n J_{-n}(X)$. Thus,

$$F(t) = \sum_{n=-\infty}^{\infty} J_n(m_f) \sin (\omega_c + n\,\omega_m)t \quad . \qquad \text{(III-16)}$$

Distortion Analysis

Let F(t) be applied to the input of a linear network whose transmission characteristics are functions of frequency as depicted in Figure 2. For simplicity, let $A_n = A(\omega_n)$ and $\theta_n = \theta(\omega_n)$ where n

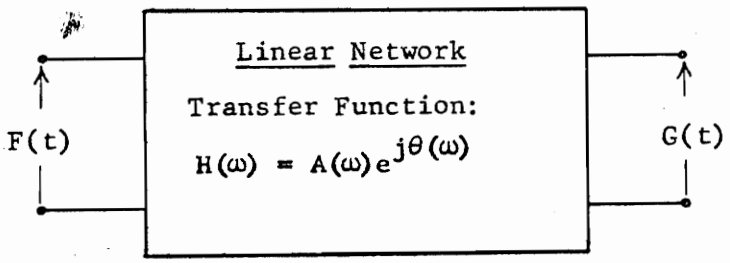


Figure 2. Transmission of FM wave through linear network.

is the order of the sideband component. For example, A_{-3} indicates the network gain at the 3rd lower sideband frequency. Since each sideband will be operated on by the linear network, the output wave will have the form

$$G(t) = \sum_{n=-\infty}^{\infty} A_n J_n(m_f) \sin (\omega_c t + n\omega_m t + \theta_n) \qquad \text{(III-17)}$$

$$= \sum_{n=-\infty}^{\infty} A_n J_n(m_f) \left[\sin \omega_c t \cos (n\omega_m t + \theta_n) \right.$$

$$\left. + \cos \omega_c t \sin (n\omega_m t + \theta_n) \right] . \qquad \text{(III-18)}$$

At this time it will be convenient to write G(t) also as

$$G(t) = R \sin \omega_c t + S \cos \omega_c t \qquad \text{(III-19)}$$

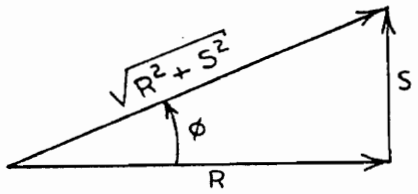
where R and S are given by

$$R = \sum_{n=-\infty}^{\infty} A_n J_n(m_f) \cos (n\omega_m t + \theta_n) \qquad \text{(III-20)}$$

and

$$S = \sum_{n=-\infty}^{\infty} A_n J_n(m_f) \sin (n\omega_m t + \theta_n) . \qquad \text{(III-21)}$$

Equation (III-19) shows that R and S can also be interpreted as the horizontal and vertical components of the angular displacement of the output wave, respectively. This is illustrated in Figure 3. From the figure it is seen that



Magnitude = $\sqrt{R^2 + S^2}$

Phase Angle = ϕ = arctan (S/R)

(Reference $\omega_c t = 90^\circ$)

Figure 3. Phasor diagram for G(t).

$$R = \sqrt{R^2+S^2} \cos \phi \text{ and } S = \sqrt{R^2+S^2} \sin \phi . \qquad \text{(III-22)}$$

Making these substitutions into Equation (III-19) yields

$$G(t) = \sqrt{R^2+S^2}\ \left[\sin \omega_c t \cos \phi + \cos \omega_c t \sin \phi\right]$$

$$= \sqrt{R^2+S^2}\ \sin (\omega_c t + \phi) . \qquad \text{(III-23)}$$

The instantaneous frequency of the output wave can now be found by recalling the definition given in Equation (III-2). Thus,

$$\omega_i = \frac{d\tilde{\phi}}{dt} = \frac{d}{dt} \quad (\omega_c t + \phi)$$

$$= \omega_c + \frac{d\phi}{dt} . \qquad \text{(III-24)}$$

Making use of the fact that $\phi = \arctan (S/R)$, ω_i can be written

$$\omega_i = \omega_c + \frac{R(dS/dt) - S(dR/dt)}{R^2 + S^2} \qquad \text{(III-25)}$$

where dR/dt and dS/dt are obtained by differentiating Equations (III-20) and (III-21) with respect to time. Thus,

$$\frac{dR}{dt} = - \sum_{n=-\infty}^{\infty} n\, \omega_m A_n J_n(m_f) \sin (n\omega_m t + \theta_n), \qquad \text{(III-26)}$$

and

$$\frac{dS}{dt} = \sum_{n=-\infty}^{\infty} n\, \omega_m A_n J_n(m_f) \cos (n\omega_m t + \theta_n). \qquad \text{(III-27)}$$

The output of the discriminator will be proportional to $d\phi/dt$ or

$$e_o = K \left[\frac{R(dS/dt) - S(dR/dt)}{R^2 + S^2} \right] \qquad \text{(III-28)}$$

where K is the constant of proportionality. It should be noted that Equation (III-26) assumes that the amplitude variations given by

$\sqrt{R^2 + S^2}$ have been eliminated by limiting action before reaching the discriminator.

It is of interest to observe what effects the network has on the amplitude of the output wave before limiting. In the case of constant attenuation at all frequencies, the amplitude of the output will be constant. When the network attenuation varies with frequency, the output wave will exhibit amplitude modulation in addition to the frequency modulation. If the variation of the attenuation characteristic with frequency is linear, the amplitude modulation will be undistorted, but a nonlinear attenuation characteristic will result in distorted amplitude modulation. This action is illustrated in Figure 4. A few comments can also be made concerning the phase characteristic of the network. If the phase characteristic of the network is constant, the output wave will be identical in phase to the input. In the case of a phase characteristic which varies linearly with frequency, the output will not exhibit phase distortion but will suffer a time delay with respect to the input. Finally, a nonlinear phase characteristic introduces phase distortion in the output wave. This behavior can be understood by recalling that the discriminator output is proportional to the derivative of the instantaneous phase of the applied wave.

Equation (III-28) specifies the demodulated signal which results when the FM wave is passed through any network whose transmission characteristic can be determined at the various sideband frequencies. The distortion in the output signal can now be determined by using a Fourier series to represent the waveform and then using the following

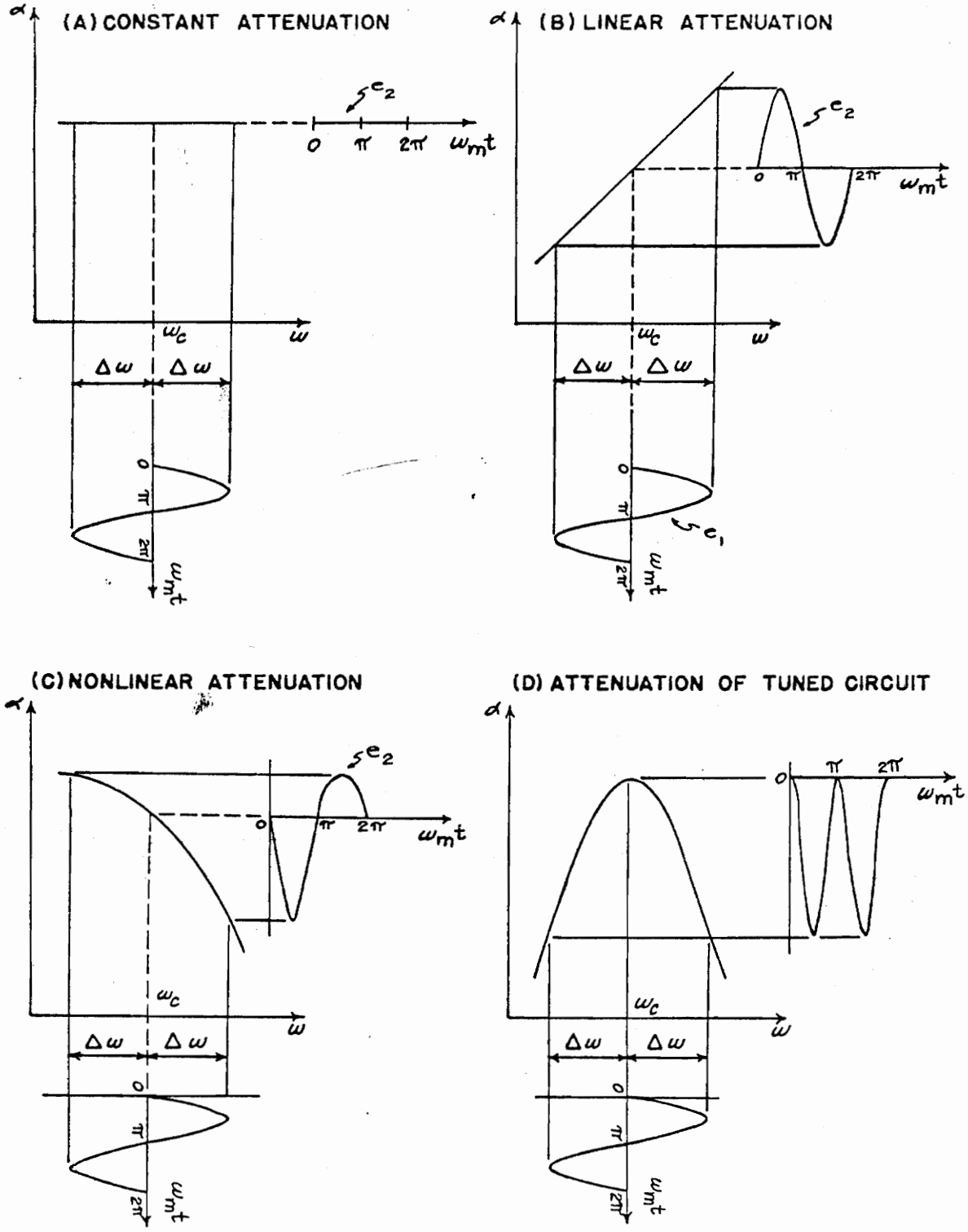


FIGURE 4 – EFFECTS OF ATTENUATION CHARACTERISTIC ON A FM WAVE

formula for total harmonic distortion:

$$\%_{HD} = \sqrt{\frac{E_2^2 + E_3^2 + \cdots}{E_1^2 + E_2^2 + E_3^2 + \cdots}} \times 100 \qquad \text{(III-29)}$$

where E_n represents the r.m.s. value (or peak value if more convenient) of the $n^{\underline{th}}$ harmonic.

It should be pointed out that Equation (III-28) suffers no restrictions with respect to size of modulation index or small departure from linearity of the transmission characteristics as do some of the other popular methods of distortion analysis for FM waves. The most serious drawback to this method in the past has been the enormous amount of labor involved in calculations whenever the modulation index was even moderately large. The development of ultra-fast digital computers now essentially eliminates this disadvantage.

Digital Computer Analysis

Pages 27 through 29 show a digital computer program which is used to evaluate Equation (III-28). Sample output data is given on pages 30 through 32. In addition to the main program, there are also some supporting subprograms which perform various important tasks. Although it has been stated that Equation (III-28) suffers no restrictions, the limits of practicability require placing restraints on the computer analysis. Since the digital computer is inherently an approximation device, the question arises, "Just how much accuracy is required in the

```
C
C       FM DISTORTION ANALYSIS   MAIN PROGRAM   6TH REVISION  3/20/66
C       MOD. INDEX, DWTD IS INCREMENT OF ANG. DISPL. OF MOD. WAVE,
C       AND DEV IS THE MAXIMUM DEVIATION IN RADIANS/SECOND.
C
C
      DIMENSION AH(60),AL(60),ANGH(60),ANGL(60),FT(360),A(10),B(10)
      COMMON AH,AL,ANGH,ANGL,A0,ANG0
      EXTERNAL NOLMUN
   33 READ(5,1) WC,W,DEV,DWTD
    1 FORMAT(4F10.0)
      Z=DEV/W
C
C       WC IS ANG. VEL. OF CARRIER, W IS ANG. VEL. OF MOD. WAVE, Z IS
      WRITE(6,70) WC,W,DEV,Z
   70 FORMAT(4H1WC=F10.0,5H   W=F10.0,8H    DEV=F10.0,5H   Z=F6.3//)
      IF(Z-.4) 2,2,3
    2 IMAX=1
      GO TO 37
    3 IF(Z-1.) 4,4,7
    4 IMAX=3
      GO TO 37
    7 IMAX=Z+5.
   37 PI=3.1415927
      DWTR=DWTD*PI/180.
   14 CALL TFCN2(WC,W,IMAX)
C       THE VALUES FROM THE TRANSFER FUNCTION ARE NOW AVAILABLE TO THE
C       MAIN PROGRAM BY VIRTUE OF THE COMMON STATEMENT.
      WRITE(6,49) A0,ANG0
   49 FORMAT(16H                A0=F11.8,18X,5HANG0=F11.8)
      DO 81 J=1,IMAX
   81 WRITE(6,82) J,AL(J),J,AH(J),J,ANGL(J),J,ANGH(J)
   82 FORMAT(4H AL(I2,2H)=F11.8,4H AH(I2,2H)=F11.8,6H ANGL(I2,2H)=F11.8,
     1       6H ANGH(I2,2H)=F11.8)
      WRITE(6,92)
```

```
   92 FORMAT(1H1)
      K=0
      MX=360.05/DWTD
      WTR=0
      DO 62 J=1,MX
   50 R=BES(Z,0)*A0
C     (NOTE. BES(Z,N) IS A BESSEL FUNCTION WITH ARGUMENT Z AND ORDER N.)
      S=0.
      C=1.
      DO 5 N=1,IMAX
      R=R+BES(Z,N)*(AH(N)*COS(C*WTR+ANGH(N))+(-1.)**N*AL(N)
     1  *COS(C*WTR-ANGL(N)))
      S=S+BES(Z,N)*(AH(N)*SIN(C*WTR+ANGH(N))-(-1.)**N*AL(N)
     1  *SIN(C*WTR-ANGL(N)))
    5 C=C+1.
      DR=0.
      DS=0.
      C=1.
      DO 10 N=1,IMAX
      DR=DR-BES(Z,N)*C*W*(AH(N)*SIN(C*WTR+ANGH(N))+(-1.)**N
     1  *AL(N)*SIN(C*WTR-ANGL(N)))
      DS=DS+BES(Z,N)*C*W*(AH(N)*COS(C*WTR+ANGH(N))-(-1.)**N
     1  *AL(N)*COS(C*WTR-ANGL(N)))
   10 C=C+1.
      PHI=ATAN2(S,R)
      Q=R*R+S*S
      RMAG=SQRT(Q)
      DWI=(R*DS-S*DR)/Q
      WTD=WTR*180./PI
      K=K+1
      IF(K-1)20,20,30
   20 WRITE(6,21)
   21 FORMAT(/12H WT(DEGREES),6X,9HMAGNITUDE,8X,12HPHI(RADIANS),5X,
     1  16HDWI(RADIANS/SEC)/)
   30 WRITE(6,31)WTD,RMAG,PHI,DWI
```

```
   31 FORMAT(1X,4(1PE15.8,2X))
      WTR=WTR+DWTR
C
CC    EVALUATION OF HARMONIC DISTORTION.
C
   62 FT(J)=DWI
      WRITE(6,92)
      DO26 I=1,10
      S1=0
      S2=0
      S3=0
      DO25 M=1,MX
      Y=I
      V=M
      S1=S1+(FT(M)*DWTR)/(2.*PI)
      S2=S2+(FT(M)*COS(Y*DWTR*(V-.5))*DWTR)/PI
   25 S3=S3+(FT(M)*SIN(Y*DWTR*(V-.5))*DWTR)/PI
      DC=S1
      A(I)=S2
      B(I)=S3
   26 WRITE(6,101) I,A(I),I,B(I)
  101 FORMAT(3H A(I2,2H)=F12.7,8X,2HB(I2,2H)=F12.7)
      WRITE(7,102) DC
  102 FORMAT(15X,3HA0=F12.7)
      C1=A(1)**2+B(1)**2
      TOP=0
      DO 121 J=2,10
  121 TOP=TOP+A(J)**2 + B(J)**2
      DIST=SQRT(TOP/(TOP+C1))*100.
      WRITE(6,130) DIST
  130 FORMAT(///21H PERCENT DISTORTION = F7.2)
      GO TO 33
   60 STOP
      END
```

WC= 10000. W= 10. DEV= 200. Z=20.000

A0= 0.50000000 ANG0=-0.00000000

AL(1)=	0.49028089	AH(1)=	0.49029975	ANGL(1)=	0.19749181	ANGH(1)=	-0.19729950
AL(2)=	0.46417416	AH(2)=	0.46430223	ANGL(2)=	0.38085185	ANGH(2)=	-0.38016219
AL(3)=	0.42857568	AH(3)=	0.42891616	ANGL(3)=	0.54108299	ANGH(3)=	-0.53975945
AL(4)=	0.39012850	AH(4)=	0.39073797	ANGL(4)=	0.67571970	ANGH(4)=	-0.67376846
AL(5)=	0.35310951	AH(5)=	0.35399340	ANGL(5)=	0.78665287	ANGH(5)=	-0.78415284
AL(6)=	0.31952272	AH(6)=	0.32065618	ANGL(6)=	0.87753973	ANGH(6)=	-0.87458885
AL(7)=	0.28994200	AH(7)=	0.29128909	ANGL(7)=	0.95221003	ANGH(7)=	-0.94889915
AL(8)=	0.26423287	AH(8)=	0.26575740	ANGL(8)=	1.01400402	ANGH(8)=	-1.01040842
AL(9)=	0.24198136	AH(9)=	0.24365139	ANGL(9)=	1.06561887	ANGH(9)=	-1.06179801
AL(10)=	0.22270652	AH(10)=	0.22449545	ANGL(10)=	1.10916078	ANGH(10)=	-1.10516064
AL(11)=	0.20595180	AH(11)=	0.20783810	ANGL(11)=	1.14625418	ANGH(11)=	-1.14211018
AL(12)=	0.19131728	AH(12)=	0.19328369	ANGL(12)=	1.17815015	ANGH(12)=	-1.17388961
AL(13)=	0.17846526	AH(13)=	0.18049805	ANGL(13)=	1.20581642	ANGH(13)=	-1.20146053
AL(14)=	0.16711557	AH(14)=	0.16920373	ANGL(14)=	1.23000699	ANGH(14)=	-1.22557233
AL(15)=	0.15703733	AH(15)=	0.15917202	ANGL(15)=	1.25131448	ANGH(15)=	-1.24681418
AL(16)=	0.14804060	AH(16)=	0.15021468	ANGL(16)=	1.27020904	ANGH(16)=	-1.26565355
AL(17)=	0.13996883	AH(17)=	0.14217648	ANGL(17)=	1.28706715	ANGH(17)=	-1.28246485
AL(18)=	0.13269255	AH(18)=	0.13492901	ANGL(18)=	1.30219299	ANGH(18)=	-1.29755074
AL(19)=	0.12610419	AH(19)=	0.12836553	ANGL(19)=	1.31583458	ANGH(19)=	-1.31115794
AL(20)=	0.12011391	AH(20)=	0.12239687	ANGL(20)=	1.32819578	ANGH(20)=	-1.32348938
AL(21)=	0.11464629	AH(21)=	0.11694812	ANGL(21)=	1.33944549	ANGH(21)=	-1.33471316
AL(22)=	0.10963765	AH(22)=	0.11195607	ANGL(22)=	1.34972468	ANGH(22)=	-1.34496965
AL(23)=	0.10503398	AH(23)=	0.10736702	ANGL(23)=	1.35915186	ANGH(23)=	-1.35437681
AL(24)=	0.10078916	AH(24)=	0.10313518	ANGL(24)=	1.36782727	ANGH(24)=	-1.36303452
AL(25)=	0.09686370	AH(25)=	0.09922127	ANGL(25)=	1.37583622	ANGH(25)=	-1.37102772

WT(DEGREES)	MAGNITUDE	PHI(RADIANS)	DWI(RADIANS/SEC)
0.	1.2172946 E-01	-1.3212519 E 00	1.9878537 E 02
1.5000000 E 01	1.2567063 E-01	-2.4173585 E 00	1.9347453 E 02
3.0000000 E 01	1.3783579 E-01	2.4556480 E 00	1.7525118 E 02
4.4999999 E 01	1.6019881 E-01	3.8441239 E-01	1.4554938 E 02
5.9999998 E 01	2.0473462 E-01	-2.5830240 E 00	1.0700558 E 02
7.4999997 E 01	2.7777351 E-01	-3.3142716 E-01	6.5212239 E 01
8.9999997 E 01	3.8458743 E-01	8.0345548 E-01	2.1931222 E 01
1.0500000 E 02	4.6060142 E-01	8.3756386 E-01	-1.9689184 E 01
1.1999999 E 02	3.4230031 E-01	-2.7390651 E-01	-6.9005179 E 01
1.3499999 E 02	1.4259711 E-01	2.8664112 E 00	-1.6699820 E 02
1.4999999 E 02	1.4383285 E-01	-1.1965472 E 00	-1.6968253 E 02
1.6499999 E 02	1.2672659 E-01	2.1588349 E-01	-1.8967768 E 02
1.7999999 E 02	1.1945174 E-01	1.3259280 E 00	-1.9877892 E 02
1.9499998 E 02	1.2340428 E-01	2.4217796 E 00	-1.9344986 E 02
2.0999998 E 02	1.3561376 E-01	-2.4515835 E 00	-1.7524663 E 02
2.2499998 E 02	1.5810766 E-01	-3.8082312 E-01	-1.4560340 E 02
2.3999998 E 02	2.0287943 E-01	2.5860641 E 00	-1.0702140 E 02
2.5499997 E 02	2.7644276 E-01	3.3343653 E-01	-6.5256109 E 01
2.6999997 E 02	3.8429183 E-01	-8.0238427 E-01	-2.1955908 E 01
2.8499997 E 02	4.6154709 E-01	-8.3673433 E-01	1.9717611 E 01
2.9999997 E 02	3.4345209 E-01	2.7743019 E-01	6.9259310 E 01
3.1499996 E 02	1.4487763 E-01	-2.8579608 E 00	1.6626415 E 02
3.2999996 E 02	1.4591841 E-01	1.2021045 E 00	1.6990759 E 02
3.4499996 E 02	1.2901704 E-01	-2.1132642 E-01	1.8973760 E 02

A(1)=	196.1421833	B(1)=	40.5384636
A(2)=	0.0047915	B(2)=	0.0186620
A(3)=	5.8493007	B(3)=	-8.5666604
A(4)=	0.0380563	B(4)=	0.0291722
A(5)=	-9.0597389	B(5)=	-2.2957667
A(6)=	0.0487638	B(6)=	-0.0440578
A(7)=	-2.2830116	B(7)=	7.5473094
A(8)=	-0.0440062	B(8)=	-0.0780151
A(9)=	6.3760155	B(9)=	3.2252536
A(10)=	-0.0858870	B(10)=	0.0178518

A0= -0.0112197

PERCENT DISTORTION = 8.73

answer?" Thus, economics is brought into the picture and plays an important role in the overall analysis. One must weigh improved accuracy against increased cost and strive to obtain the optimum efficiency. It is with this philosophy in mind that the computer program will now be analyzed.

In order to calculate the output waveform as a function of time using Equation (III-28), the values of ω_c, ω_m and $\Delta\omega$ are read into the program. Also supplied initially is the increment in the frequency of the modulating wave which determines the value of ω that will be used in the calculations. After these quantities have been entered, the modulation index m_f is calculated. At this point in the program, a very important decision must be made concerning the number of sidebands to take in order to achieve the desired accuracy. Equation (III-16) shows that each sideband is multiplied by a Bessel function of the same order as the sideband. Figure 5 shows some properties of Bessel functions of the first kind which will help in making the decision. As seen from the curves the magnitude of the Bessel function is very small whenever the order is approximately five greater than the argument. This means that little error would be introduced by neglecting sidebands of order greater than $m_f + 5$. The curve in Figure 6 gives some indication of the error introduced by neglecting these higher sidebands. This curve is actually a plot of the magnitude of the sideband of order $m_f + 6$ and shows that the coefficient of the $m_f + 6$ order harmonic is equal to or less than 0.01 for m_f equal to or less than 44. Thus, the error introduced by neglecting these sideband components (and all higher

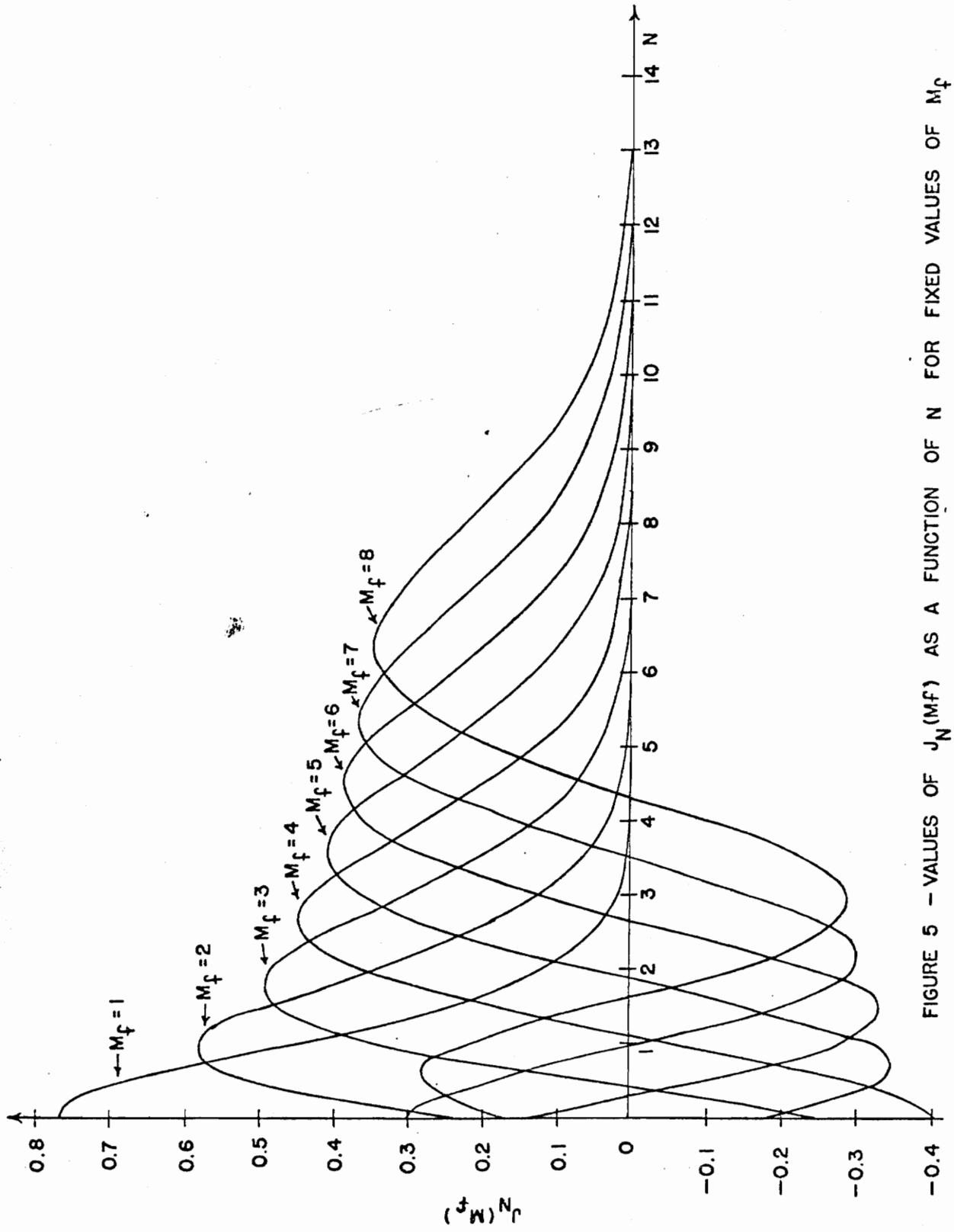


FIGURE 5 – VALUES OF $J_N(M_f)$ AS A FUNCTION OF N FOR FIXED VALUES OF M_f

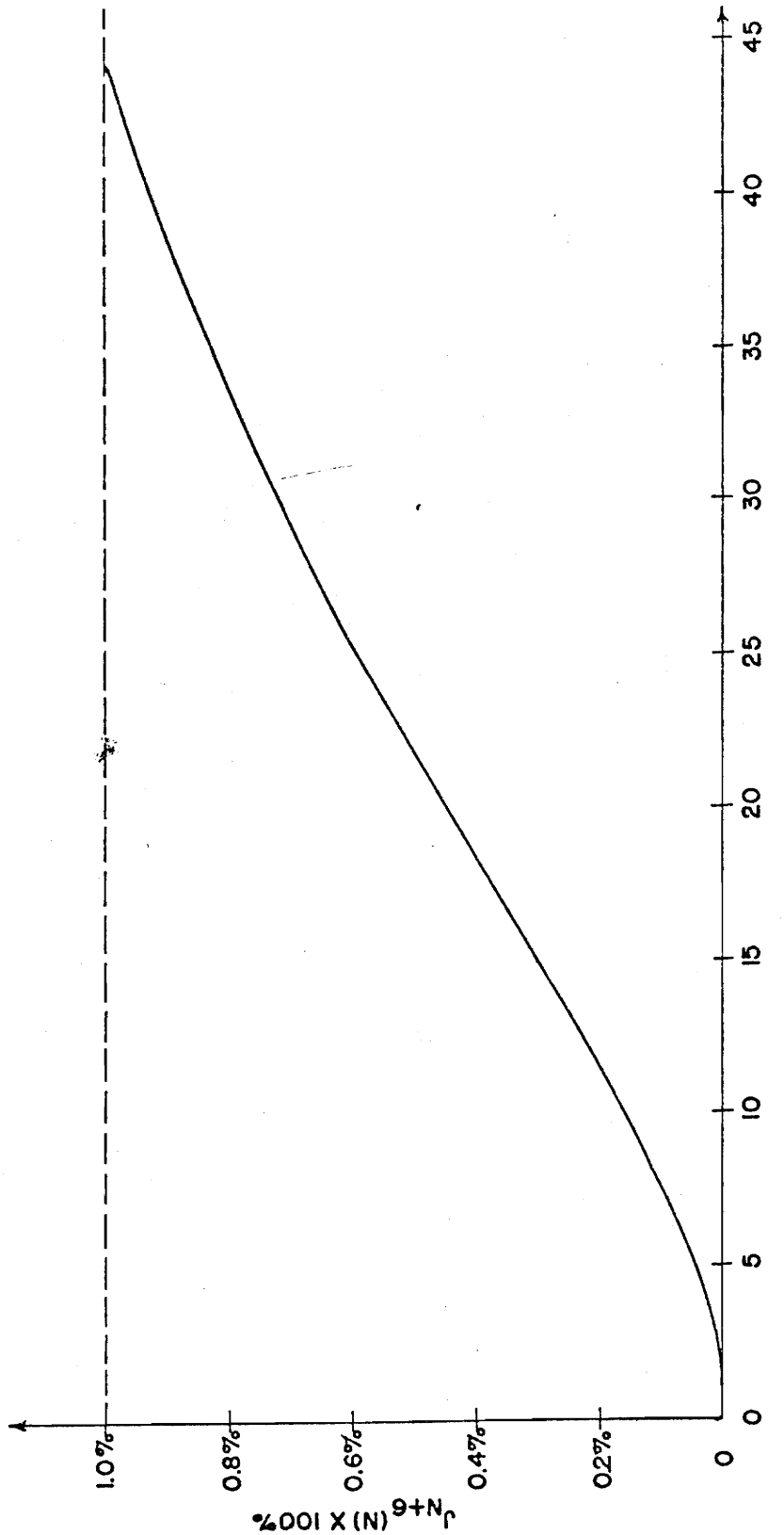


FIGURE 6 – MAGNITUDE OF $J_{N+6}(N)$ AS A FUNCTION OF N AS A PERCENTAGE OF UNITY

order components) is indeed small. Figure 6 also shows that the error is very small at low values of modulation index which indicates that sideband components of lower order than the $6^{\underline{th}}$ may possibly be neglected. In fact, Hund [18] shows that a single pair will suffice up to a modulation index of 0.4 and only the first 3 orders are necessary for values of modulation index up to and including unity. Hence, the number of sidebands to be used in the calculations will be either 1,3, or $m_f + 5$ depending on whether $0 < m_f \leq .4$, $.4 < m_f \leq 1$, or $m_f > 1$, respectively. The computer variable which determines the number of sideband pairs to be used is IMAX. The group of statements starting after statement number 70 down to and including statement 7 is where the computer decides how many sideband pairs to use.

Statement number 14 in the program is a call to a subroutine named TFCN2 which is the network transfer function. FORTRAN programs of a few typical transfer functions are contained in Appendix E. At this time the TFCN2 subroutine computes the magnitude and phase of the network transfer function for all sideband frequencies to be used later in the program and stores the values in storage common to the main program. The reason for this is to save time by not having to call the subroutine each time a magnitude or phase angle is needed. The cost of this time saving is the loss of storage space. These values are also printed out for convenience and checking. At this point a simple example will be worked out by hand to show the sequence of calculations to be made so that the rest of the computer program will be easier to understand. For convenience the phase shift and attenuation

characteristics will be assumed to be linear functions of frequency so that the demodulated output will be undistorted but shifted by an angle equal to the amount that the first sideband is shifted by the network. In addition, the FM wave should exhibit undistorted amplitude modulation due to the linear amplitude variations with frequency.

Example: Let the frequency-modulated wave be represented by the equation

$$e(t) = E \sin (\omega_c t + 1 \sin \omega_m t) \qquad \text{(III-30)}$$

so that $m_f = 1$. The normalized function can then be represented by Equation (III-15) as

$$\begin{aligned} F(t) = {} & J_0(1) \sin \omega_c t \\ & + J_1(1) \left[\sin (\omega_c+\omega_m)t - \sin(\omega_c-\omega_m)t \right] \\ & + J_2(1) \left[\sin (\omega_c+2\omega_m)t + \sin(\omega_c-2\omega_m)t \right] \\ & + J_3(1) \left[\sin(\omega_c+3\omega_m)t - \sin(\omega_c-3\omega_m)t \right] . \end{aligned} \qquad \text{(III-31)}$$

This series has been terminated after only 3 sidebands have been taken into account. This is in accordance with the previous discussion. The normalized output wave is now obtained by multiplying Equation (III-31) by the network transfer function. Hence,

$$G(t) = A_o J_o(1) \sin(\omega_c t + \theta_o)$$
$$+ J_1(1)\left[A_1 \sin(\omega_c t + \omega_m t + \theta_1) - A_{-1} \sin(\omega_c t - \omega_m t + \theta_{-1})\right]$$
$$+ J_2(1)\left[A_2 \sin(\omega_c t + 2\omega_m t + \theta_2) + A_{-2} \sin(\omega_c t - 2\omega_m t + \theta_{-2})\right]$$
$$+ J_3(1)\left[A_3 \sin(\omega_c t + 3\omega_m t + \theta_3) - A_{-3} \sin(\omega_c t - 3\omega_m t + \theta_{-3})\right]. \quad \text{(III-32)}$$

It will now be assumed that the gain and phase characteristics of the network are linear and can be represented as follows:

$A_o = 1.0$ $\qquad$ $\theta_o = 0^o$

$A_1 = 0.9$	$A_{-1} = 1.1$	$\theta_1 = -10^o$	$\theta_{-1} = 10^o$
$A_2 = 0.8$	$A_{-2} = 1.2$	$\theta_2 = -20^o$	$\theta_{-2} = 20^o$
$A_3 = 0.7$	$A_{-3} = 1.3$	$\theta_3 = -30^o$	$\theta_{-3} = 30^o$

The values of the Bessel functions with orders 0, 1, 2, 3, for argument 1.0 are

$J_o(1) = 0.7652$	$J_1(1) = 0.4401$
$J_2(1) = 0.1149$	$J_3(1) = 0.01956.$

In order to simplify the calculations, Equation (III-32) will be expressed in phasor form taking $\sin \omega_c t$ as reference. Thus,

$$\tilde{G}(t) = A_o J_o(1) \underline{/\theta_o} + A_1 J_1(1) \underline{/\omega_m t+\theta_1} - A_{-1} J_1(1) \underline{/-\omega_m t+\theta_{-1}}$$

$$+ A_2 J_2(1) \underline{/2\omega_m t+\theta_2} + A_{-2} J_2(1) \underline{/-2\omega_m t+\theta_{-2}}$$

$$+ A_3 J_3(1) \underline{/3\omega_m t+\theta_3} - A_{-3} J_3(1) \underline{/-3\omega_m t+\theta_{-3}} \ .$$

Substituting the values from the transfer function and the Bessel function values yields

$$\tilde{G}(t) = 0.7652(1.0) \underline{/0^o} + 0.4401 \ (.9 \underline{/\omega_m t-10^o} - 1.1 \underline{/-\omega_m t+10^o})$$

$$+ 0.1149 \ (.8 \underline{/2\omega_m t-20^o} + 1.2 \underline{/-2\omega_m t + 20^o})$$

$$+ 0.01956 \ (.7 \underline{/3\omega_m t-30^o} - 1.3 \underline{/-3\omega_m t+30^o}) \ .$$

In order to complete the calculations, certain values of the modulating wave will be substituted in and the corresponding value of the output wave will be obtained. The plot of the resulting data will yield the output waveform.

$\underline{\omega_m t = 0^o}$

$$\tilde{G} = 0.7652 + 0.3961 \underline{/-10^o} - 0.4841 \underline{/10^o}$$

$$+ 0.0919 \underline{/-20^o} + 0.1379 \underline{/20^o}$$

$$+ 0.0137 \underline{/-30^o} - 0.0255 \underline{/30^o}$$

$$= 0.898 \underline{/-10.8^o}$$

$\omega_m t = 30^\circ$

$$\tilde{G} = 0.7652 + 0.3961 \angle 20^\circ - 0.4841 \angle -20^\circ$$

$$+ 0.0919 \angle 40^\circ + 0.1379 \angle -40^\circ$$

$$+ 0.0137 \angle 60^\circ - 0.0255 \angle -60^\circ$$

$$= 0.905 \angle 19.7^\circ .$$

These two calculations illustrate the method and Table (1) summarizes the results of other calculations. These values are plotted in Figures 7 and 8 in order to analyze the results. Referring first to Figure 7 one observes that the phase of the output wave as a function of time is a sine wave shifted by 10° from the origin. Since the demodulated signal is proportional to the derivative of this phase angle, the output signal will be a cosine wave delayed by 10° with respect to the modulating wave. This is in agreement with what was predicted. The peak value of the phase variation is seen to be 57.5° or approximately one radian which agrees with the value given in Equation (III-30). Figure 8 shows the sinusoidal amplitude variation of the FM wave that was caused by the linear attenuation characteristic of the network. It should be pointed out that some inaccuracies are encountered in this example, especially in the amplitude calculations.

TABLE 1

COMPUTED VALUES FOR THE EXAMPLE PROBLEM

$\omega_m t$ (degrees)	$\tilde{G}(t)$	$\omega_m t$ (degrees)	$\tilde{G}(t)$
0	0.898 /-10.8°	180	1.094 /9.92°
10	0.895 / 0.0°	190	1.095 /0.0°
20	0.898 / 10.8°	200	1.094 /-9.9°
30	0.905 / 19.7°	210	1.092 /-19.5°
40	0.915 / 28.7°	220	1.083 /-28.7°
50	0.928 / 36.7°	230	1.080 /-36.7°
60	0.939 / 43.7°	240	1.069 /-43.8°
70	0.950 / 49.4°	250	1.053 /-49.6°
80	0.964 / 53.8°	260	1.042 /-54.2°
90	0.981 / 56.5°	270	1.016 /-56.6°
100	0.996 / 57.5°	280	0.996 /-57.5°
110	1.016 / 56.6°	290	0.981 /-56.5°
120	1.042 / 54.2°	300	0.964 /-53.8°
130	1.053 / 49.6°	310	0.950 /-49.4°
140	1.069 / 43.8°	320	0.939 /-43.7°
150	1.080 / 36.7°	330	0.928 /-36.7°
160	1.083 / 28.7°	340	0.915 /-28.7°
170	1.092 / 19.5°	350	0.905 /-19.7°

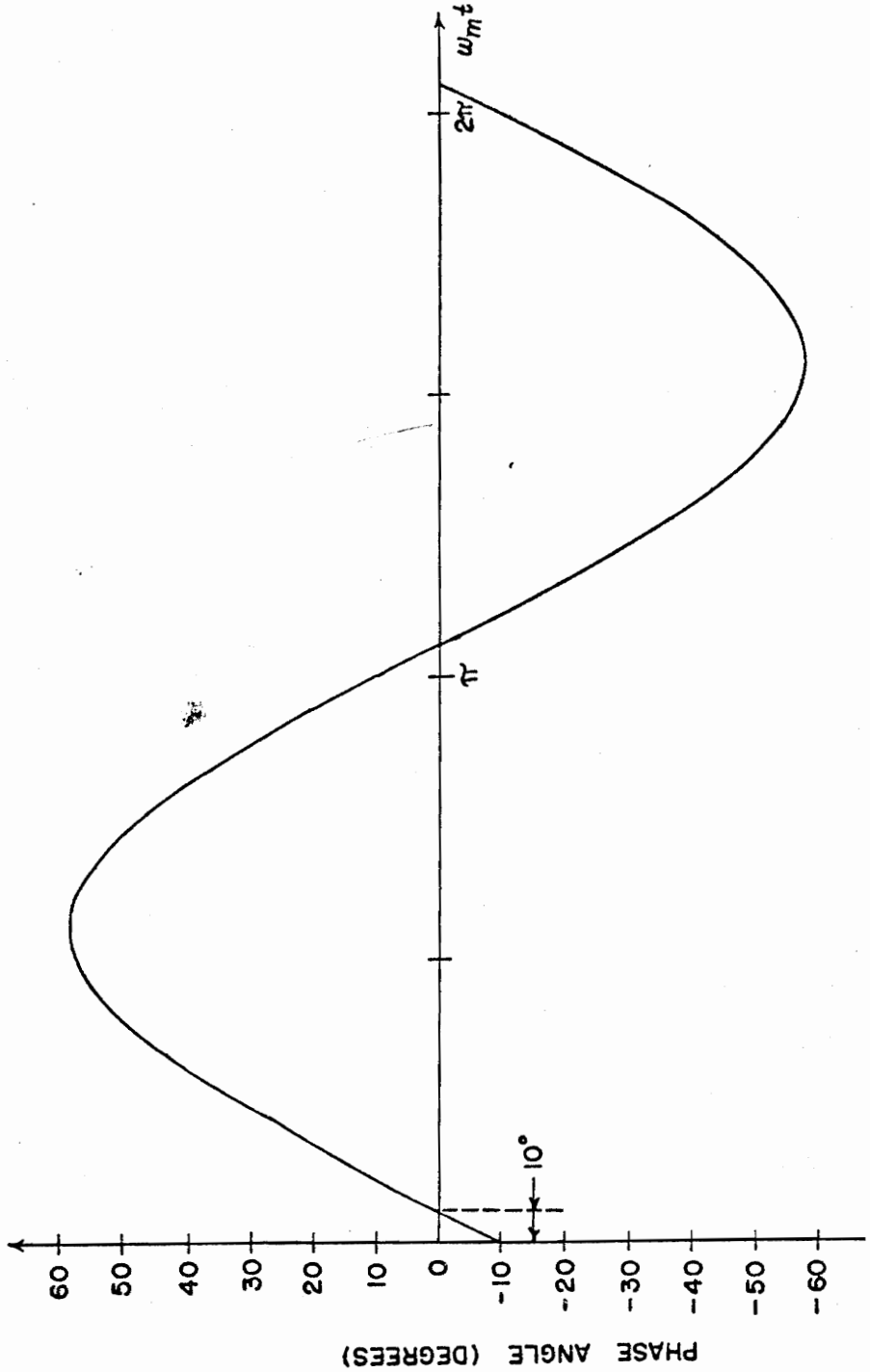


FIGURE 7 – PLOT OF THE PHASE OF THE FM WAVE AFTER PASSING THROUGH LINEAR NETWORK

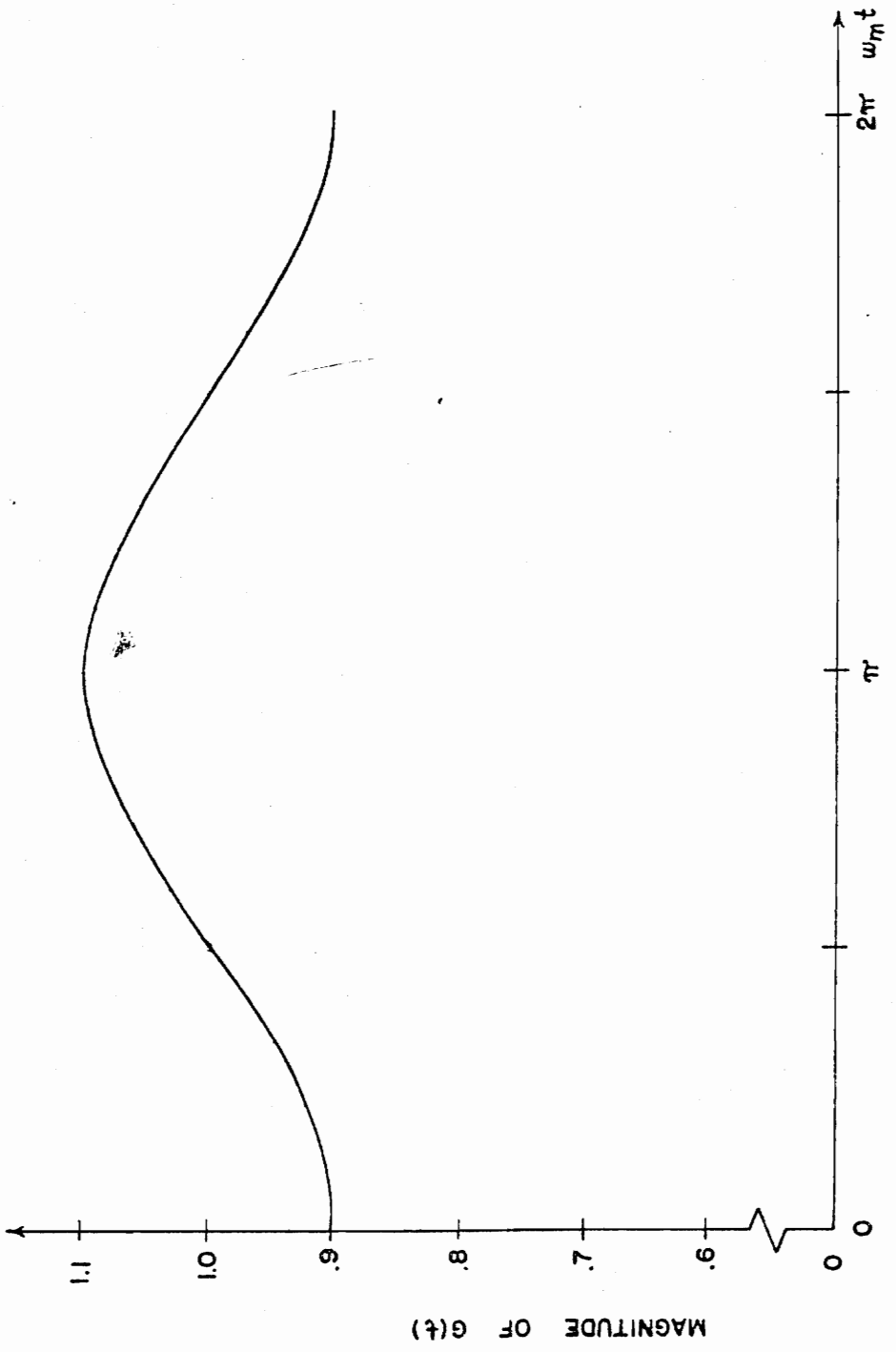


FIGURE 8 – PLOT OF THE AMPLITUDE OF THE FM WAVE AFTER PASSING THROUGH LINEAR NETWORK

These inaccuracies result partly because of being restricted to slide rule accuracy in the calculations and partly because of neglecting some of the higher order sidebands. It should also be observed that the demodulated wave was never actually calculated by the method given in the section on distortion analysis (see Equation (III-28)). The reason for this was that the work inyolved was too tedious to be readily performed by hand even for this simple example. In addition, since no distortion was present, it was easy enough to deduce the output from the curve of phase as a function of time. The chief purpose of this example has been to understand the basic mechanics involved in making certain calculations.

With the above example in mind, let us now return to the analysis of the computer program. Assume that the computer has determined how many sidebands to use and has called the transfer function subroutine which has already calculated the values for the gain and phase angle at each one of the sideband frequencies and has these values stored. The next step is to determine at what values of time it would be best to compute the output wave. The obvious choice is to obtain values of the output signal over a full cycle of the modulating wave. This will assure that all pertinent data is available. The value of the

angular increment in the modulating wave that is initially read into the program (DWTD) allows for the selection of the times (or angles) into which the wave will be divided. The statement MX = 360.05/DWTD is the way the computer decides how many sets of calculations it is to make. For example, if one wished to obtain the output signal at each 10^{o} of the modulating wave throughout a full cycle of the modulating wave, it would be necessary to make a total of 36 sets of calculations--one for every 10^{o} of the 360^{o} in a full cycle. Note that since MX is a fixed point constant, the calculation made by the computer will be made in floating-point arithmetic and then everything to the right of the decimal will be truncated. The reason for the decimal fraction added to 360 to give 360.05 is to assure that no relevant integers are lost during the floating-point to fixed-point arithmetic conversion.

The program now enters that portion where the calculations are made to determine the magnitude and phase of the frequency-modulated output wave and the values of the demodulated signal as a function of time. The first set of calculations is made at $\omega_m t = 0^{o}$ and each subsequent set is made at $\omega_m t$ plus the increment as stated above. During the course of the calculations, it is necessary to use certain Bessel functions as has already been seen in the sample problem. This requires that a scheme be made available to the computer to compute Bessel functions. This is accomplished by using a function subprogram with the main program so that the value of a certain Bessel function is computed at the time it is needed. The details of the Bessel function

subprogram BES (Z,N) are found in Appendix A. It will be sufficient here just to state that BES (Z,N) is a function subprogram which computes Bessel functions of the first kind of order N and argument Z.

The portion of the program included in the DO-loop designated by "DO26" is the heart of the program since it is here that the main calculations are made. In effect, Equation (III-28) is evaluated with the constant K assumed to be unity. In the process of calculating the demodulated signal the amplitude and phase of the FM wave are also obtained. These quantities along with the values of $\omega_m t$ at which the calculations are made are all printed out for evaluation purposes. After the output wave is obtained it is analyzed for distortion by first computing its Fourier series and applying Equation (III-29). In determining the Fourier series by numerical integration it is necessary to choose intervals of the output wave that are compatible with the main program. In other words, the values of ωt chosen for the evaluation of the Fourier series must be equal to or at least integral multiples of $\omega_m t$. In this program the same number of intervals is chosen for the Fourier series computation as for the evaluation of the output wave, namely, 360/DWTD. Provisions are made in the DIMENSION statement to accommodate 360 intervals. Under this restriction DWTD is limited to values of 1^o or greater. This is regarded as sufficient for most purposes. If smaller increments are desired, it is necessary only to change the DIMENSION statement to accommodate the increased number of increments. It should be noted that the first 10 harmonics are calculated. This again should be

sufficient but, if not, the only change required to increase (or decrease) the number of harmonics to be calculated is to change the upper limit of the DO statement, DO26. Finally, the distortion is calculated from Fourier series coefficients and printed out as the end result.

The computer time required to make a complete analysis varies from a couple of minutes to about 15 or 20 minutes on the IBM 7040/1401 system depending upon several factors. For example, if the modulation index is low so that only a few sidebands are needed and the incremental change in $\omega_m t$ is large so that only a few sets of calculations are required, then the time will be short. On the other hand, an increase in either of the above quantities will result in a corresponding increase in computing time. An actual computer solution for the case of a modulation index of unity evaluated for every 10° of the modulating wave took a total of four minutes including two minutes of output time. A second run made for the case when the modulation index was 40 and evaluated for each 3° of the modulating wave consumed 16 minutes. Thus, it can be seen that indiscriminate choices of input data can result in long and expensive computer runs.

Modulation by Two Sinusoidal Waves

There are many cases where it would be convenient to be able to determine the FM output from a network when the input FM wave has been modulated by signals other than a single sinusoid. One such case concerns

modulation produced by a pair of sinusoidal waves. Such modulation can be used to determine the intermodulation distortion produced in an FM wave due to the nonlinear transmission characteristics of the network. Let the instantaneous frequency be represented by

$$\omega_i = \omega_c + \Delta\omega_1 \cos \omega_1 t + \Delta\omega_2 \cos \omega_2 t \quad . \qquad \text{(III-33)}$$

The equation of the FM wave is obtained as before by integrating ω_i with respect to time and substituting into Equation (III-5). Again taking φ_o to be zero yields

$$e(t) = E \sin \left[\omega_c t + m_1 \sin \omega_1 t + m_2 \sin \omega_2 t \right] \qquad \text{(III-34)}$$

where $m_1 = \Delta\omega_1/\omega_1$ and $m_2 = \Delta\omega_2/\omega_1$. Equation (III-34) can now be expanded as before except now there will be many more terms to consider. Due to the additional complexity of the resulting expansions, it will be advantageous to use an alternate derivation. In order to do this it is necessary to express the wave in terms of exponentials. Hence,

$$e(t) = E \cdot \mathrm{Im}\left\{ e^{\,j(\omega_c t + m_1 \sin \omega_1 t + m_2 \sin \omega_2 t)} \right\} . \qquad \text{(III-35)}$$

For ease of manipulation, the complex quantity C(t) will be introduced and defined as follows:

$$C(t) = e^{\,j(\omega_c t + m_1 \sin \omega_1 t + m_2 \sin \omega_2 t)} \qquad \text{(III-36)}$$

so that the FM wave normalized in amplitude is

$$F(t) = \frac{e(t)}{E} = \mathrm{Im}\left\{ C(t) \right\} \quad . \qquad \text{(III-37)}$$

Now Equation (III-36) can be written

$$C(t) = e^{j\omega_c t} \cdot e^{jm_1 \sin \omega_1 t} \cdot e^{jm_2 \sin \omega_2 t} \qquad \text{(III-38)}$$

By making use of the following identity, this expression can be expanded into an infinite series with Bessel function coefficients:

$$e^{jx \sin y} = \sum_{n=-\infty}^{\infty} J_n(x)\, e^{jny} . \qquad \text{(III-39)}$$

Equation (III-38) now becomes

$$C(t) = \left[\sum_{N=-\infty}^{\infty} J_N(m_1)\, e^{jN\omega_1 t}\right] \cdot \left[\sum_{K=-\infty}^{\infty} J_K(m_2)\, e^{jK\omega_2 t}\right] \cdot e^{j\omega_c t} . \qquad \text{(III-40)}$$

Since these two summations are independent of each other and neither influences the $e^{j\omega_c t}$ term, both summation symbols can be grouped together in front of the other terms. The result is

$$C(t) = \sum_{N=-\infty}^{\infty} \sum_{K=-\infty}^{\infty} J_N(m_1)\, e^{jN\omega_1 t} \cdot J_K(m_2)\, e^{jK\omega_2 t} \cdot e^{j\omega_c t}$$

$$= \sum_{N=-\infty}^{\infty} \sum_{K=-\infty}^{\infty} J_N(m_1)\, J_K(m_2)\, e^{j(\omega_c + N\omega_1 + K\omega_2)t} . \qquad \text{(III-41)}$$

The actual FM wave is obtained by taking the imaginary part of this expression. Hence,

$$F(t) = \sum_{N=-\infty}^{\infty} \sum_{K=-\infty}^{\infty} J_N(m_1) J_K(m_2) \sin (\omega_c + N\omega_1 + K\omega_2)t. \qquad \text{(III-42)}$$

Equation (III-42) results in spectral components of the following frequency and amplitude:

1. Carrier component: $J_o(m_1)J_o(m_2)\sin \omega_c t$
2. Sidebands due to ω_1: $J_n(m_1)J_o(m_2)\sin(\omega_c \pm N\omega_1)t$
3. Sidebands due to ω_2: $J_o(m_1)J_k(m_2)\sin(\omega_c \pm K\omega_2)t$
4. Sidebands due to ω_1 and ω_2 : $J_n(m_1)J_k(m_2)\sin(\omega_c \pm N\omega_1 \pm k\omega_2)t$

Figures 9 (a), (b), (c), and (d) show the plots of the magnitude of these sideband components for an example case. Figure 9(e) gives the total spectrum for the wave taking the signs into account. In this example the carrier is normalized to 100 for convenience while the modulating frequencies are $\omega_1 = 1$ and $\omega_2 = 2$. The values for the modulation index were taken as $m_1 = 2$ and $m_2 = 1$. While one must be very careful in drawing any concrete conclusions from one specific example, there are a few interesting generalizations which can be made. First and most obvious is the large number of sidebands that arise because of ω_1, ω_2, and particularly due to the interactions of ω_1 and ω_2. It should be pointed out that the spectrum shown in Figure 9(e) is misleading in this regard. The number of sidebands shown in the total spectrum do not greatly exceed those of the individual plots of Figures 9 (a), (b), (c), and (d). The reason for this is that the value of ω_2 turns out to be the first harmonic of ω_1. In the general case where the two modulating frequencies are not harmonically related, all of the sideband components would show up in the total spectrum. Nevertheless, it is interesting to observe how the sidebands behave in the above case.

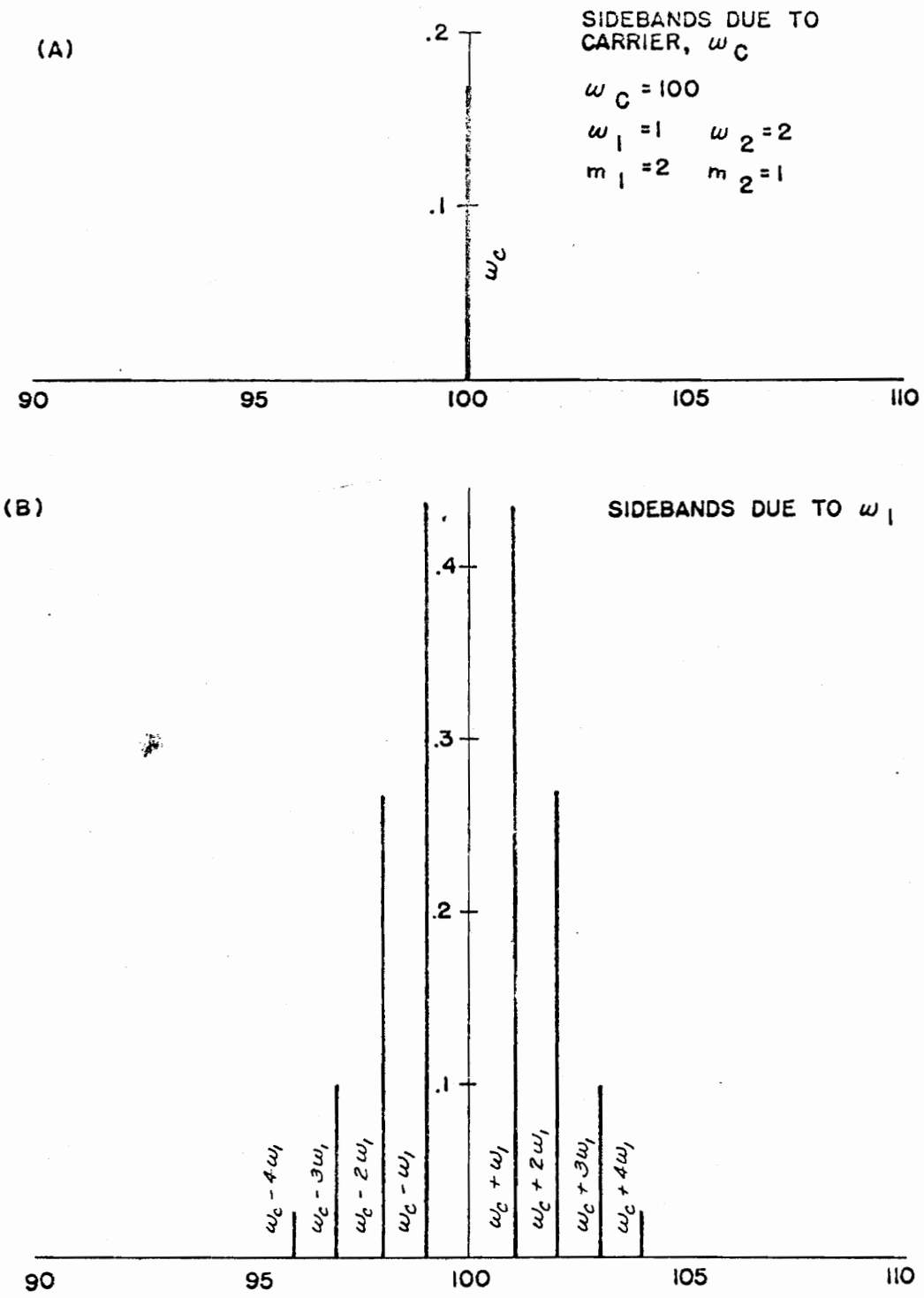
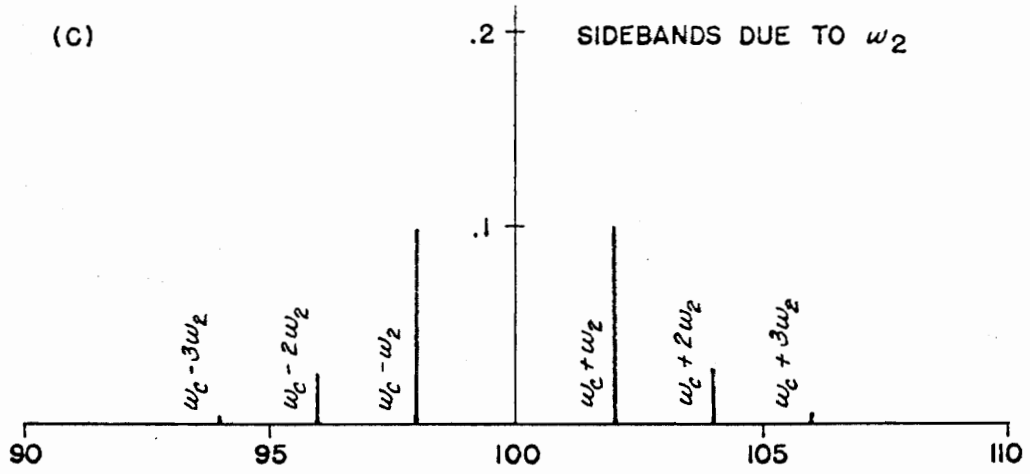


FIGURE 9 (A & B) – MAGNITUDE OF SIDEBAND COMPONENTS FOR MODULATION BY TWO SINUSOIDS



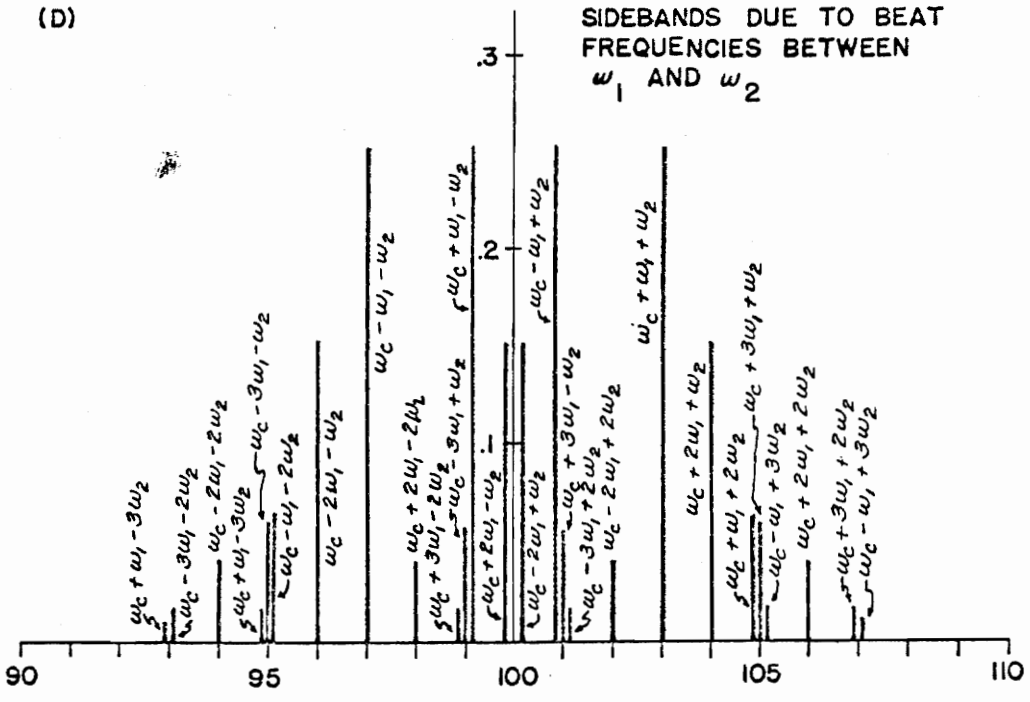


FIGURE 9 (C & D) – MAGNITUDE OF SIDEBAND COMPONENTS FOR MODULATION BY TWO SINUSOIDS

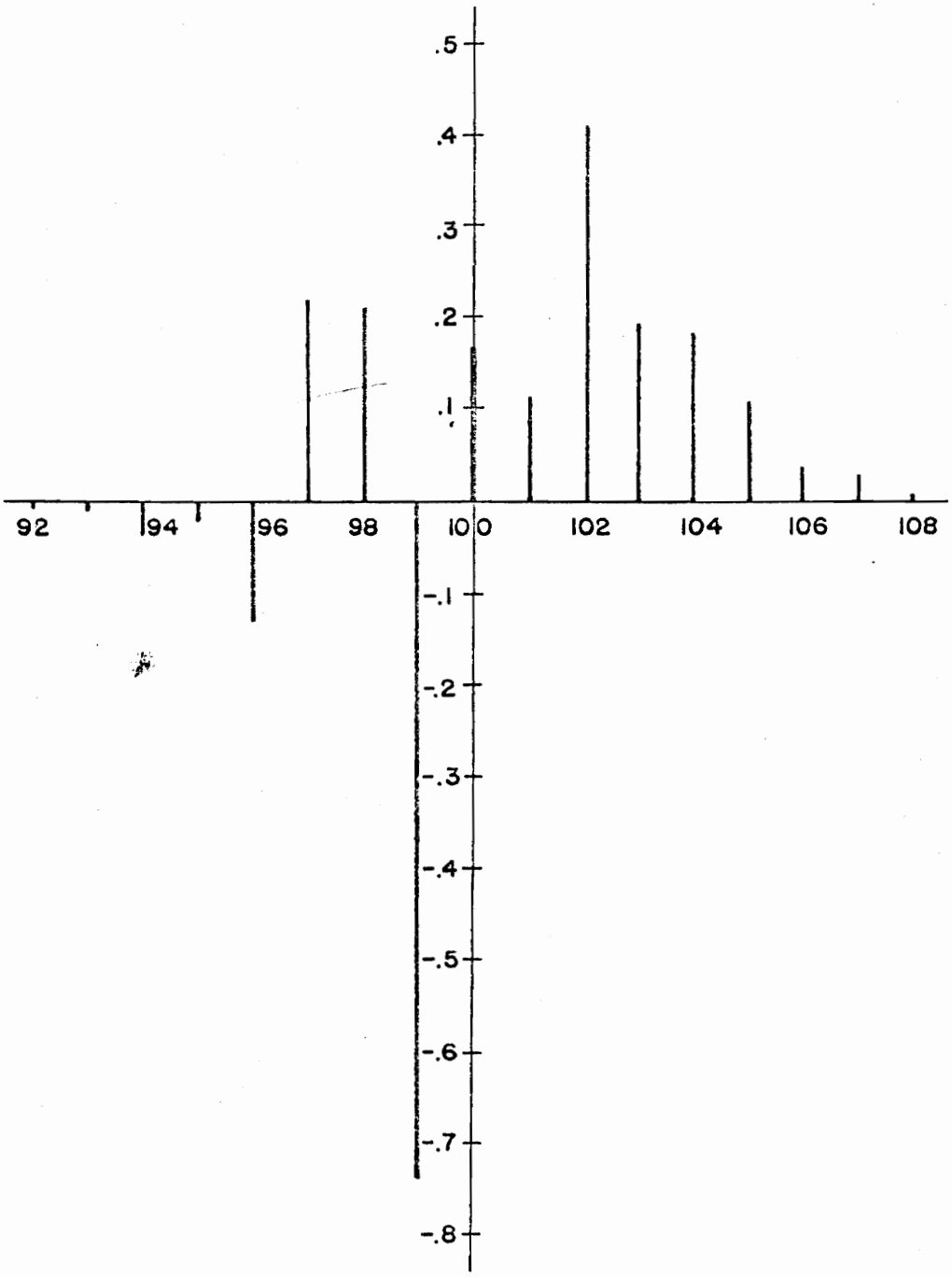


FIGURE 9(E)—TOTAL SPECTRUM OF FM WAVE INCLUDING POLARITY

The spectrum takes on a very unsymmetrical shape as far as magnitude and sign is concerned. The increased bandwidth of the spectrum is also observed. Although the increase is not very large, it does indicate a necessity for further investigation.

In making a study of the required bandwidth for a network to pass an FM wave modulated with a pair of sinusoids, one first observes that each sideband has as its coefficient the produce of two Bessel functions. Certain properties of Bessel functions have already been discussed in connection with the single frequency modulation. The study of these properties indicated that an error of less than 1% was obtained for the single frequency modulation if sidebands higher than the 5th were neglected. This same criterion seems to be a convenient starting place in determining the maximum number of sidebands to include in the case of double-frequency modulation. This case, however, is complicated by the fact that there are two sets of Bessel functions with different arguments to be considered. It is not an easy task to deduce for certain from looking at the expressions for the sidebands when the product of the two Bessel functions becomes negligible. Computer results show that if one uses $m_1 \neq m_2 \neq 5$ as the number of significant sidebands, the results are accurate to 6 significant figures. If this is more accuracy than is really needed, one can use the number of sidebands as $m_1 \neq m_2$. In this case, the answers should be good to roughly three significant figures or slide rule accuracy which may be acceptable for many purposes. The decreased accuracy would be exchanged for less computer time.

Digital Computer Analysis Using Double-Frequency Modulation

The complete computer program which evaluates the output wave for the case of double-frequency modulation is found in Appendix B along with the output results. The basic philosophy of this program is the same as that of the program using single-frequency modulation. However, the increased complexity of the calculations is immediately apparent. As mentioned earlier, the result of nonlinear network characteristics will be intermodulation of the two modulating waves. The program does not attempt to calculate a value for the intermodulation distortion because no simple method could be obtained for determining the necessary values. The reason for this is due to the dependence of the shape of the intermodulated wave on the network characteristic. The best solution for this problem seems to be to plot the output wave and calculate the intermodulation distortion from this plot. There are two methods of defining the intermodulation distortion generally used.[1] The most common of these methods, referred to as the SMPTE (Society of Motion Picture and Television Engineers) method, is characterized by applying two sine waves of widely differing frequencies to a nonlinear network and observing the sum, difference, and other frequency components generated by the nonlinearities. Typical frequencies used in this method are 60 and 6000 hertz mixed with an amplitude ratio of 4:1, respectively.

---

[1] See pages 335-341, Terman and Pettit [19].

The other method employed in measuring intermodulation distortion is the so-called CCIF method recommended by the International Telephonic Consultative Committee. In this method the modulating frequencies are both of the same amplitude while the frequencies are high but slightly different. Both of these methods have advantages and disadvantages with regard to solution on the digital computer. One of the major drawbacks of the SMPTE method is that the frequency and amplitude requirements for the modulating waves result in a modulation index ratio of 400:1 - a value completely out of the range of the program given in Appendix B. It is observed however, that a frequency ratio of 10:1 (say 60 hertz and 600 hertz) instead of the 100:1 normally used results in a modulation index ratio of 40:1 which turns out to be just in the range of the computer program.

The values of modulation index cause no problems in the CCIF method since the amplitudes of the modulating waves are equal and the frequencies are very close to being equal. The main problem encountered in this method is that the difference frequency between the two waves carries considerable significance and to obtain this waveform accurately the two modulating waves must be evaluated at a large number of points in each cycle. In addition, the calculations must be made over a complete cycle of the beat frequency wave. For example, if the frequencies of the modulating waves are 1000 and 1010 hertz, calculations would be necessary every few microseconds (since the period of each modulating wave is about 1 millisecond) in order to obtain an adequate beat frequency component of the output wave. In addition,

one must make sure that the calculations are made over the complete period of the 10 hertz beat frequency wave. Thus if the high frequency waves are sampled about every 10^{o} and over a full cycle of the 10 hertz wave, calculations would be necessary at approximately every 25 micro-seconds over a period of 10 milliseconds or about 400 sets of calculations. Since each set of calculations involves evaluating several equations each of which is composed of a large number of sidebands, it is easy to see that the work soon gets prohibitive even for a digital computer.

Based on the above discussion and the relative ease with which the answer can be obtained from a plot of the data, the SMPTE method seems to be the most promising. Figure 10 illustrates this method in a little more detail.

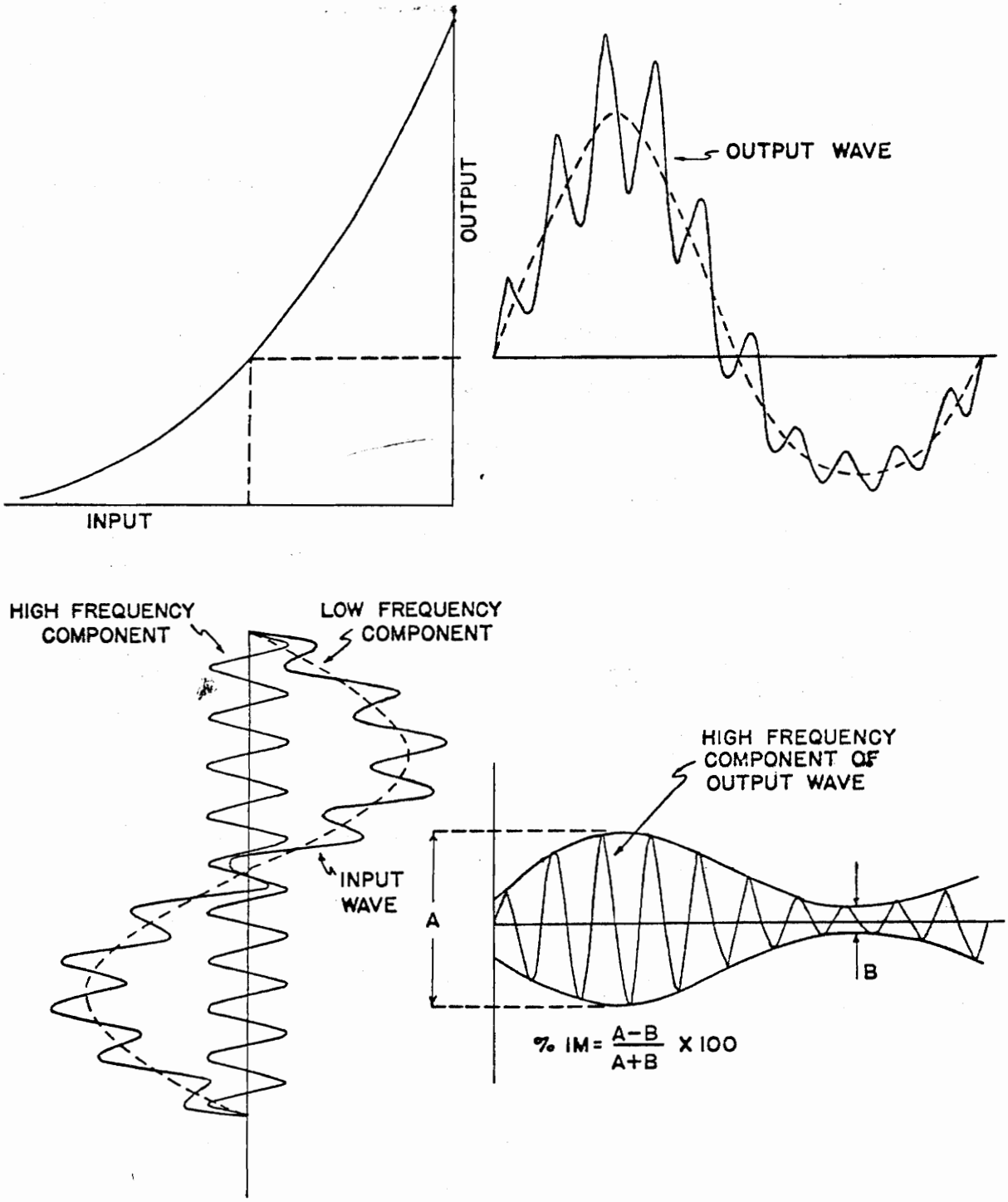


FIGURE 10 – EVALUATION OF INTERMODULATION DISTORTION BY THE SMPTE METHOD

IV. THE QUASI-STEADY-STATE ANALYSIS

The analysis of distortion in FM systems by approximations based on "quasi-steady-state" conditions was first introduced by Carson and Fry [1] in a paper entitled Variable-frequency Electric Circuit Theory. Basically, this method amounts to assuming that the input frequency is varying slowly enough so that the output frequency is at every instant equal to the input frequency plus a small (hopefully) correction term. The correction term is necessary since a varying frequency will give rise to transients no matter how small the variation. It is very important to recognize that error is inherent in this method of analysis and most of the effort in this area has been directed toward reducing the error to acceptable proportions.

Mathematical Treatment

Again, consider the FM wave at the input to the filter to be

$$e_i = E \sin(\omega_c t + \frac{\Delta \omega}{\omega m} \sin \omega m t) \ . \qquad \text{(IV-1)}$$

The phase angle of this wave is the argument of the sine function and the instantaneous frequency is the time derivative of the phase angle. Likewise, the instantaneous frequency of the output is the time derivative of the output phase angle. To a first approximation the output phase will be equal to the phase angle of the input plus whatever phase shift is introduced by the network. Hence,

$$\phi_{out} = \phi_{in} + \theta \qquad \text{(IV-2)}$$

where ϕ_{out} is the phase angle of the output wave, ϕ_{in} is the phase angle of the input wave, and θ is the phase shift introduced by the network. The instantaneous frequency of the output is obtained by differentiating Equation (IV-2):

$$(\omega_i)_{out} = \frac{d\phi_{out}}{dt} = \frac{d\phi_{in}}{dt} + \frac{d\theta}{dt} \qquad \text{(IV-3)}$$

where $d\phi_{in}/dt$ is recognized as the instantaneous frequency of the input wave and $d\theta/dt$ is a characteristic of the network. It is the $d\theta/dt$ term that gives rise to a correction term to compensate for the inability of the network to respond instantaneously to a change in the input frequency. From Equation (IV-1) the phase angle of the input wave is

$$\phi_{in} = \omega_c t + \frac{\Delta\omega}{\omega_m} \sin \omega_m t \qquad \text{(IV-4)}$$

and

$$\frac{d\phi_{in}}{dt} = \omega_c + \Delta\omega \cos \omega_m t. \qquad \text{(IV-5)}$$

The instantaneous frequency of the output can now be written as

$$(\omega_i)_{out} = \omega_c + \Delta\omega \cos \omega_m t + \frac{d\theta}{dt} \quad . \qquad \text{(IV-6)}$$

Since the phase response of the network is usually easier to obtain in the frequency domain than as a function of time, it will be necessary to express $d\theta/dt$ in terms of the frequency domain. This is easily accomplished by observing that

$$\frac{d\theta}{dt} = \frac{d\theta}{d\omega_i} \cdot \frac{d\omega_i}{dt} \qquad \text{(IV-7)}$$

where $d\theta/d\omega_i$ is the desired network characteristic and $d\omega_i/dt$ is easy to obtain from Equation (IV-5) as

$$\frac{d\omega_i}{dt} = - \omega_m \Delta\omega \sin \omega_m t. \qquad \text{(IV-8)}$$

Making these substitutions into Equation (IV-6) yields

$$(\omega_i)_{out} = \omega_c + \Delta\omega \left[\cos \omega_m t - \omega_m \sin \omega_m t \cdot \frac{d\theta}{d\omega_i}\right]. \qquad \text{(IV-9)}$$

The demodulated output signal is now obtained by separating the carrier frequency from the total instantaneous frequency:

$$e_{sig} \alpha \Delta\omega \cos \omega_m t - \frac{d\theta}{d\omega_i} \cdot \omega_m \Delta\omega \sin \omega_m t. \qquad \text{(IV-10)}$$

Equation (IV-10) clearly indicates the dependence of the distortion term on both the network phase characteristic and the modulating frequency. Criteria for low distortion can be inferred from a study of this equation. Low distortion will result in the case of networks which have a small change in phase angle as the frequency is changed and then only if the modulating frequency is relatively low. It is important to note that both a small $d\theta/d\omega_i$ and a low ω_m are necessary in order to insure low distortion. The results obtained for a specific application and reported in Chapter V graphically illustrate this point.

Digital Computer Program

In order to compare the Quasi-steady-state method of analysis with the Fourier method, Equation (IV-10) was programmed on the digital computer. Appendix C contains the resulting programs which consist of

a main program and two subprograms. Again, the main program establishes the program philosophy and directs the solution procedure while the subprograms assist the main program by making certain frequently occuring routine calculations.

Network data are first entered into the program in numerical form as a table giving the frequency and the corresponding network phase shift. These values allow the subprogram named "FUNCTION DER(W)" to calculate $d\theta/d\omega_i$ with the aid of the subprogram "FUNCTION SLOPE(L)". After the network data have been entered, the program reads in values for the carrier frequency, WC, the modulating wave frequency, WM, the maximum frequency deviation, DEV, and the incremental steps in frequency, DWTD at which successive calculations are made. This portion of the program is consistent with the program for the Fourier method as given in Chapter III.

After a few mechanical details such as determining how many calculations to make and at what frequencies to make them, the computer evaluates Equation (IV-10) for angles of the modulating wave from 0^{o} to 360^{o} spaced DWTD degrees apart. The values of the output wave are then printed out along with the angle of the modulating wave at which they were calculated and the corresponding values of the input wave. All input data are also printed out for ready reference. For more details concerning the digital computer program refer to Appendix C.

Higher-Order Corrections to the Quasi-Steady-State Method

As previously stated, Equation (IV-3) is only a first approximation to the output wave. The correction term $d\theta/dt$ is not adequate to compensate for the error when the network phase shift is large and/or the modulating frequency is high. The problem of developing an equation for the output that is less restrictive has been the subject of intensive study since the early days of FM. One of the better known expressions for the frequency-modulated wave at the output of a network is that due to Van der Pol and Stumpers [2]. A modified form of this expression is

$$e_o = E\, e^{j\left[\omega_c t + \int_o^t s(t)dt\right]} \cdot \left\{ H(\omega) + j\frac{\omega_m}{2}\left[\frac{ds(t)}{dt}\right]\frac{d^2H(\omega)}{d\omega^2} + j\frac{\omega_m^2}{6}\left[\frac{d^2s(t)}{dt^2}\right]\frac{d^3H(\omega)}{d\omega^3} - \frac{\omega_m^2}{8}\left[\frac{ds(t)}{dt}\right]\frac{d^4H(\omega)}{d\omega^4} + \cdots \right\} \qquad \text{(IV-11)}$$

where e_o is the FM output wave, E is the amplitude of the input wave and assumed to be constant, s(t) is the modulating wave, and $H(\omega)$ is the network transfer function. Equation (IV-11) is a rather complicated expression involving both time - and frequency-domain terms. In fact, it is the difficulty encountered in attempting to apply this expression to a practical problem that makes the Fourier approach look promising. In order to more fully appreciate this last statement, consider that the first two correction terms in the braces are sufficient to yield the desired accuracy. Also consider sinusoidal modulation of the form

$$s(t) = \Delta\omega \cos \omega_m t \tag{IV-12}$$

so that

$$\int_0^t s(t)dt = \frac{\Delta\omega}{\omega_m} \sin \omega_m t \tag{IV-13}$$

and

$$\frac{ds(t)}{dt} = -\omega_m \Delta\omega \sin \omega_m t. \tag{IV-14}$$

Equation (IV-11) can now be written

$$e_o = E\, e^{j(\omega_c t + \frac{\Delta\omega}{\omega_m} \sin \omega_m t)} \cdot \left\{ H(\omega) - j \frac{\omega_m^2}{2} \Delta\omega \cdot \sin \omega_m t \cdot \frac{d^2H(\omega)}{d(\omega)^2} \right\} . \tag{IV-15}$$

Now assume that the network transfer function can be expressed as

$$H(\omega) = A(\omega)e^{j\theta(\omega)} \tag{IV-16}$$

Differentiating this expression twice with respect to ω yields

$$\frac{d^2H}{d\omega^2} = e^{j\theta} \left\{ \left[\frac{d^2A}{d\omega^2} - A \left(\frac{d\theta}{d\omega}\right)^2 \right] + j \left[A\left(\frac{d^2\theta}{d\omega^2}\right) + 2 \left(\frac{d\theta}{d\omega}\right)\left(\frac{dA}{d\omega}\right) \right] \right\} \tag{IV-17}$$

where $H = H(\omega)$, $\theta = \theta(\omega)$, and $A = A(\omega)$. Making this substitution into Equation (IV-15) gives for the output wave

$$\begin{aligned} e_o = E\, e^{j(\omega_c t + \frac{\Delta\omega}{\omega_m} \sin \omega_m t)} \cdot e^{j\theta} \Bigg\{ & A \\ & + \frac{\omega_m^2}{2} \Delta\omega \sin \omega_m t \left[A \left(\frac{d^2\theta}{d\omega^2}\right) + 2 \left(\frac{d\theta}{d\omega}\right)\left(\frac{dA}{d\omega}\right) \right] \\ & - j \frac{\omega_m^2}{2} \Delta\omega \sin \omega_m t \left[\frac{d^2A}{d\omega^2} - A \left(\frac{d\theta}{d\omega}\right)^2 \right] \Bigg\} . \end{aligned} \tag{IV-18}$$

It is observed that the quantities within the braces of the above expression represent the real and imaginary parts of a complex function. Let this complex function be designated by

$$Q(\omega)e^{j\gamma(\omega)} = Qe^{j\gamma}$$

where Q is the square root of the sum of the squares of the real and imaginary parts and γ is the angle whose tangent is the imaginary part divided by the real part. Equation (IV-18) now simplifies to

$$e_o = Ee^{j(\omega_c t + \frac{\Delta\omega}{\omega_m}\sin\omega_m t)} \cdot e^{j\theta} \cdot Qe^{j\gamma} \qquad \text{(IV-19)}$$

Combining the exponential terms yields

$$e_o = EQe^{j(\omega_c t + \frac{\Delta\omega}{\omega_m}\sin\omega_m t + \theta + \gamma)} \qquad \text{(IV-20)}$$

Finally, the instantaneous frequency is obtained by taking the time derivative of the phase angle. Thus,

$$(\omega_i)_{out} = \omega_c + \Delta\omega\cos\omega_m t + \frac{d\theta}{dt} + \frac{d\gamma}{dt} \qquad \text{(IV-21)}$$

where

$$\gamma = \tan^{-1}\left\{\frac{-\frac{\omega_m^2}{2}\Delta\omega\sin\omega_m t\left[\frac{d^2A}{d\omega^2} - A\left(\frac{d\theta}{d\omega}\right)^2\right]}{A + \frac{\omega_m^2}{2}\Delta\omega\sin\omega_m t\left[A\frac{d^2\theta}{d\omega^2} + 2\left(\frac{d\theta}{d\omega}\right)\left(\frac{dA}{d\omega}\right)\right]}\right\} \qquad \text{(IV-22)}$$

The first three terms of this equation are recognized to be the same as were obtained by the first approximation method in Equation (IV-6). The $d\gamma/dt$ term gives a second-order correction to the output wave and, in doing so, the accuracy of the Quasi-steady-state method is extended. The problem encountered in trying to apply Equation (IV-21) is the

evaluation of $d\gamma/dt$. Equation (IV-22) shows that γ depends on the first and second derivatives of the amplitude and phase characteristics in addition to other quantities. Needless to say, the algebra involved in obtaining an analytic expression for γ would be exceedingly complicated for any networks except very simple ones. Even assuming that one did obtain the analytic expression for γ, it must then be differentiated before it could be applied to Equation (IV-21).

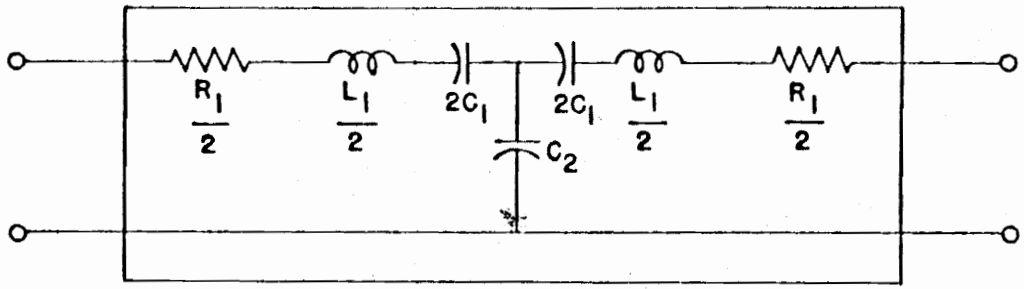
One might argue that $d\gamma/dt$ could be evaluated numerically instead of analytically as was done in the case of the first-order approximation. Indeed, this could be done, but by the time one has differentiated γ, the first, second, and third derivatives of the amplitude and phase have been evaluated. The inherent error in numerical differentiation would either yield meaningless results or else require such small increments in the phase and amplitude versus frequency input data to make this approach impractical.

As can be observed from the above discussion, the addition of even one extra correction term results in very complex expressions. No doubt higher-ordered terms would become so complex as to be completely impractical to apply to more complicated networks. The Van der Pol-Stumpers expansion used here is typical of the approach and is considered to be a refinement of previous efforts. In no case was this method found to be practically applied to a circuit more complicated than the simple three-branch, parallel-tuned circuit.

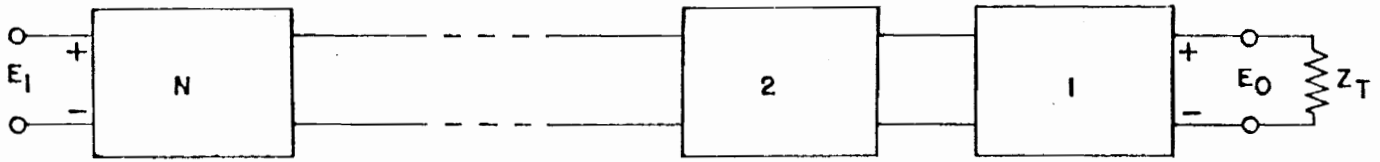
V. APPLICATION TO A MULTI-SECTION FILTER NETWORK

The Filter Network

In order to test and compare the methods of analysis discussed in Chapters III and IV, a practical bandpass filter network was selected. The network used was composed of a number of lossy constant-K filter sections connected in tandem. Figure 11 shows the details of the filter including the circuit diagram of a single section and the interconnection of the various sections. For this analysis, data giving the magnitude and phase as a function of frequency were obtained for filters composed of 10, 6, 3, and 1 sections. Appendix E contains the digital computer program used to evaluate the network. The filter was designed so that each section would have a bandpass of about 5 kilohertz. This design was made assuming lossless components and when the dissipation in the series arm (representing a coil with a Q of 100 at 100 kilohertz) is taken into account, the actual stage bandpass is found to be closer to 6 kilohertz as can be seen in Figure 12 for $N = 1$. Figure 12 also shows the magnitude of the filter transfer function as a function of frequency for the 3-section, 6-section, and the 10-section filters when the terminating impedance is 1000 ohms resistance. The effect of the series arm dissipation is apparent from the curves inasmuch as the attenuation in the passband gets progressively larger as more sections are added. Without attenuation the magnitude of the transfer function would always be unity within the passband. Figures 13, 14, and 15 show the phase characteristics



(A) CIRCUIT CONFIGURATION OF SINGLE SECTION.



(B) TANDEM CONNECTION OF N SECTIONS SHOWING TERMINATING IMPEDANCE, Z_T.

FIGURE II – DETAILS OF MULTI-SECTION FILTER

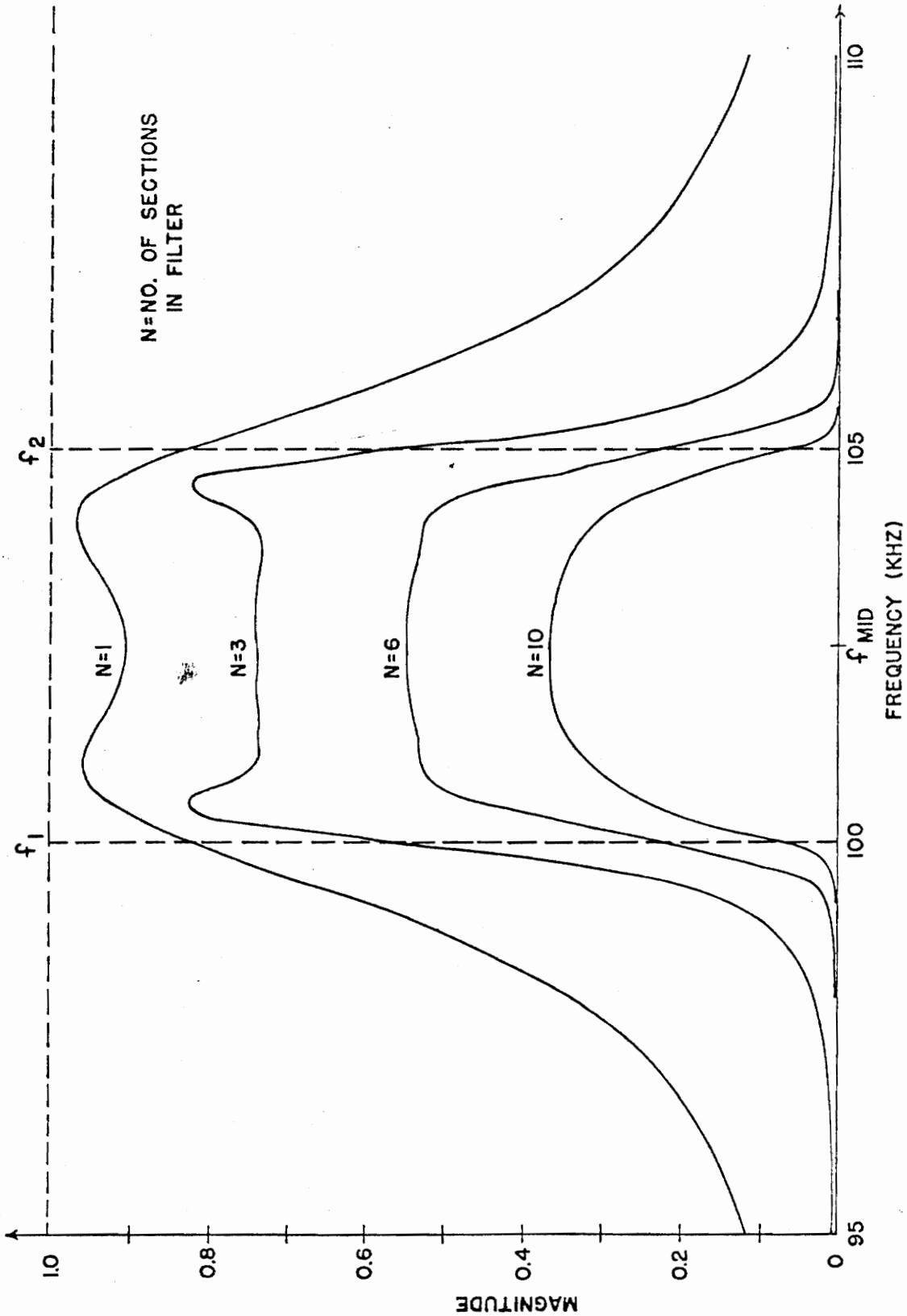


FIGURE 12 — TRANSFER FUNCTION MAGNITUDE AS A FUNCTION OF FREQUENCY FOR FILTER

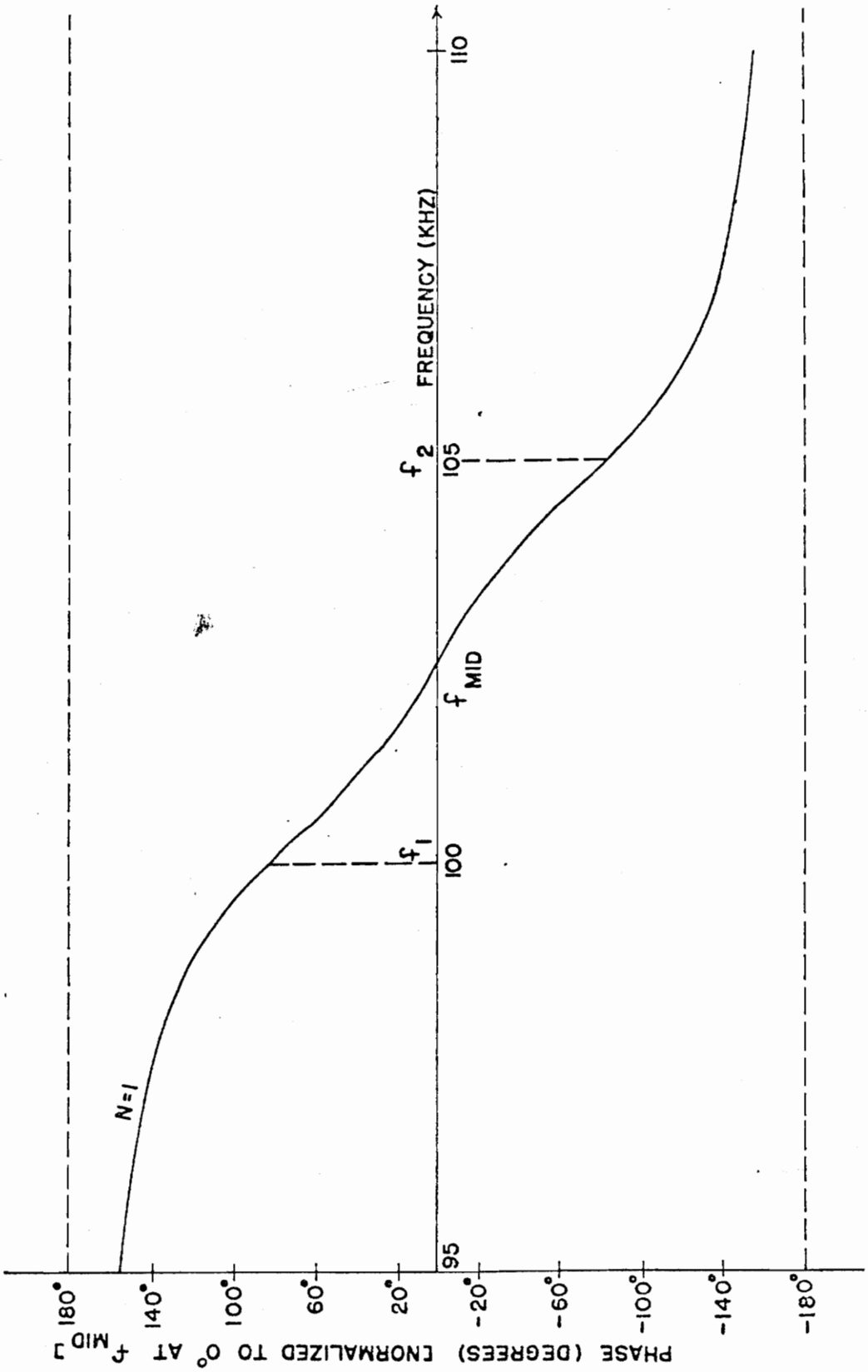


FIGURE 13 — PHASE CHARACTERISTIC OF SINGLE-SECTION FILTER

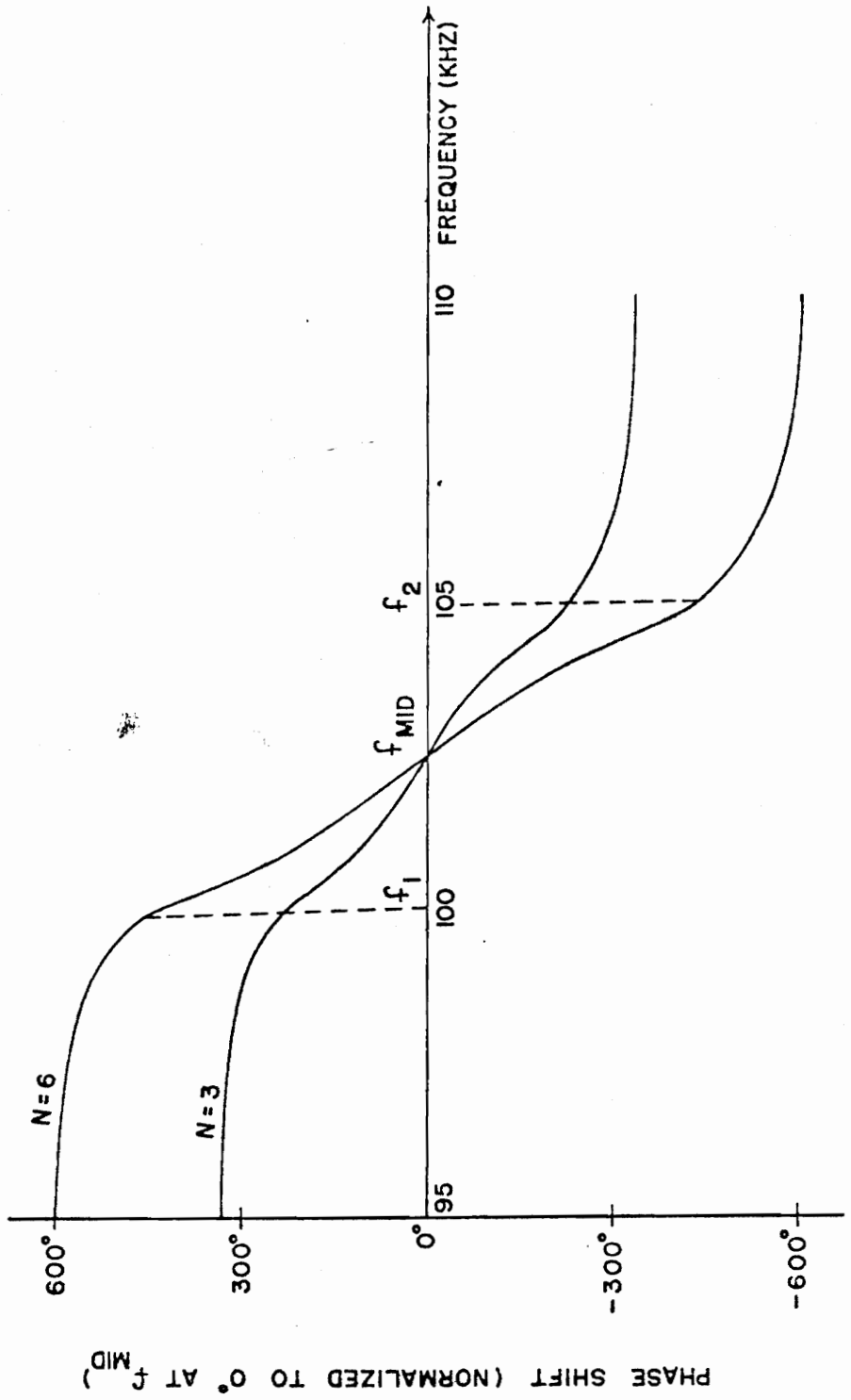


FIGURE 14 – PHASE CHARACTERISTICS OF 3-SECTION AND 6-SECTION FILTER

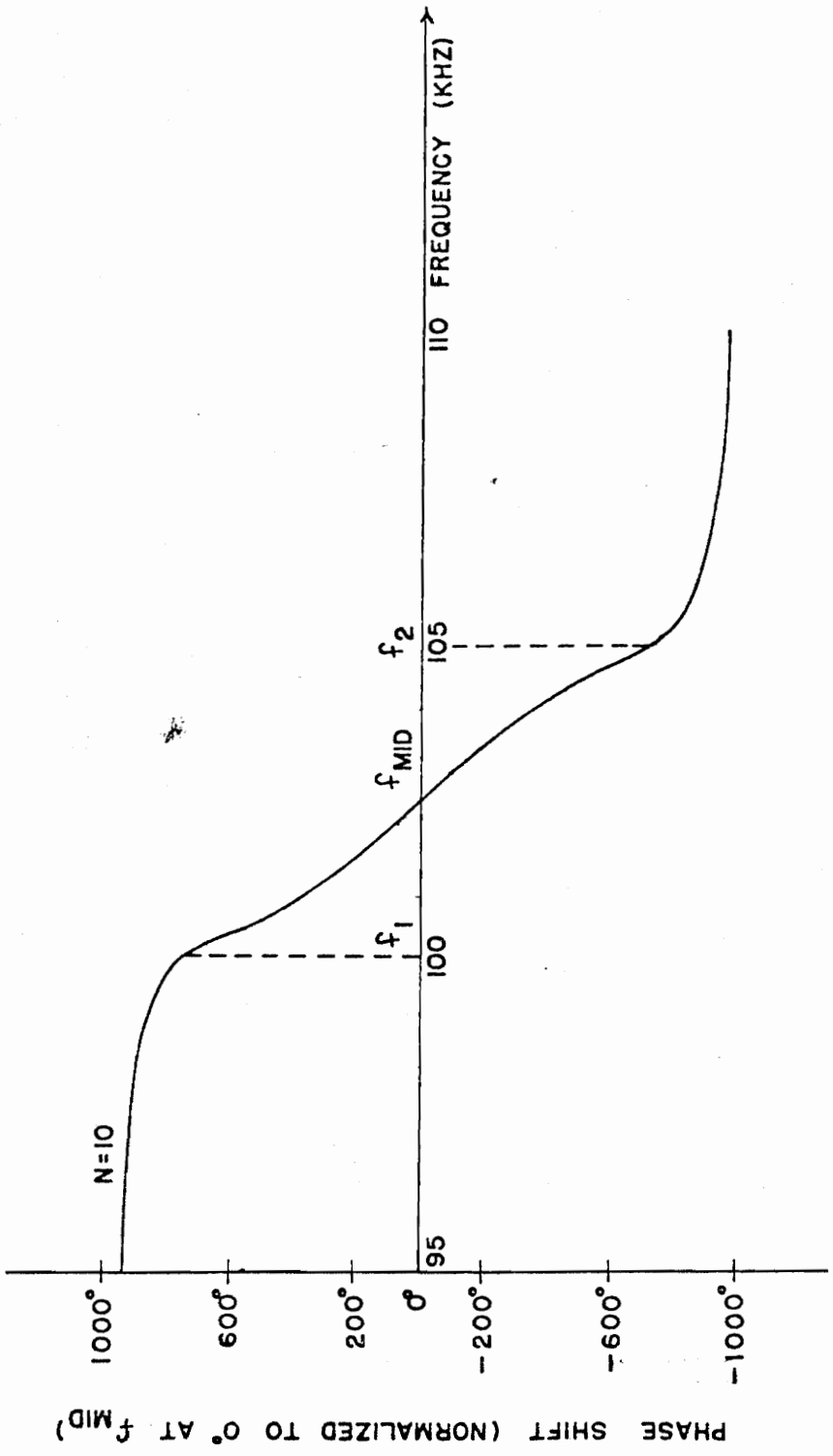


FIGURE 15 — PHASE CHARACTERISTIC OF 10-SECTION FILTER

of the network as a function of frequency. The reason for plotting these curves on three separate sheets was to display more clearly any nonlinear properties that may be present. From the curves one observes how the phase shift varies with frequency for each of the filters being considered and also the way in which the magnitude of the total phase shift at any given frequency increases as more sections are added. Figure 13 which is plotted for the single-section filter clearly shows some small nonlinearities present in the curve while the curve for the 10-section filter shown in Figure 15 seems to be very linear over most of the pass band. This is contrary to what one would expect since it would be assumed that nonlinearities appearing in the single-section filter would be magnified when 10 sections are taken into account. The fact is that there are indeed relatively large nonlinearities present in the curve for the 10-section filter, but they are masked by the extremely large total phase shift. In order to better observe the characteristics of these nonlinearities, the data has been altered and re-plotted in Figure 16 for the 10-section filter. The single-section filter phase characteristic is also replotted for comparison. The extent of the alteration in the data was merely to subtract a linear ramp function from the curve. According to the theory of Chapter III, this should result in a fixed phase shift of the output signal but should not affect the distortion. Hence, Figure 16 more realistically compares the nonlinearity of the 10-section filter with that of the single section.

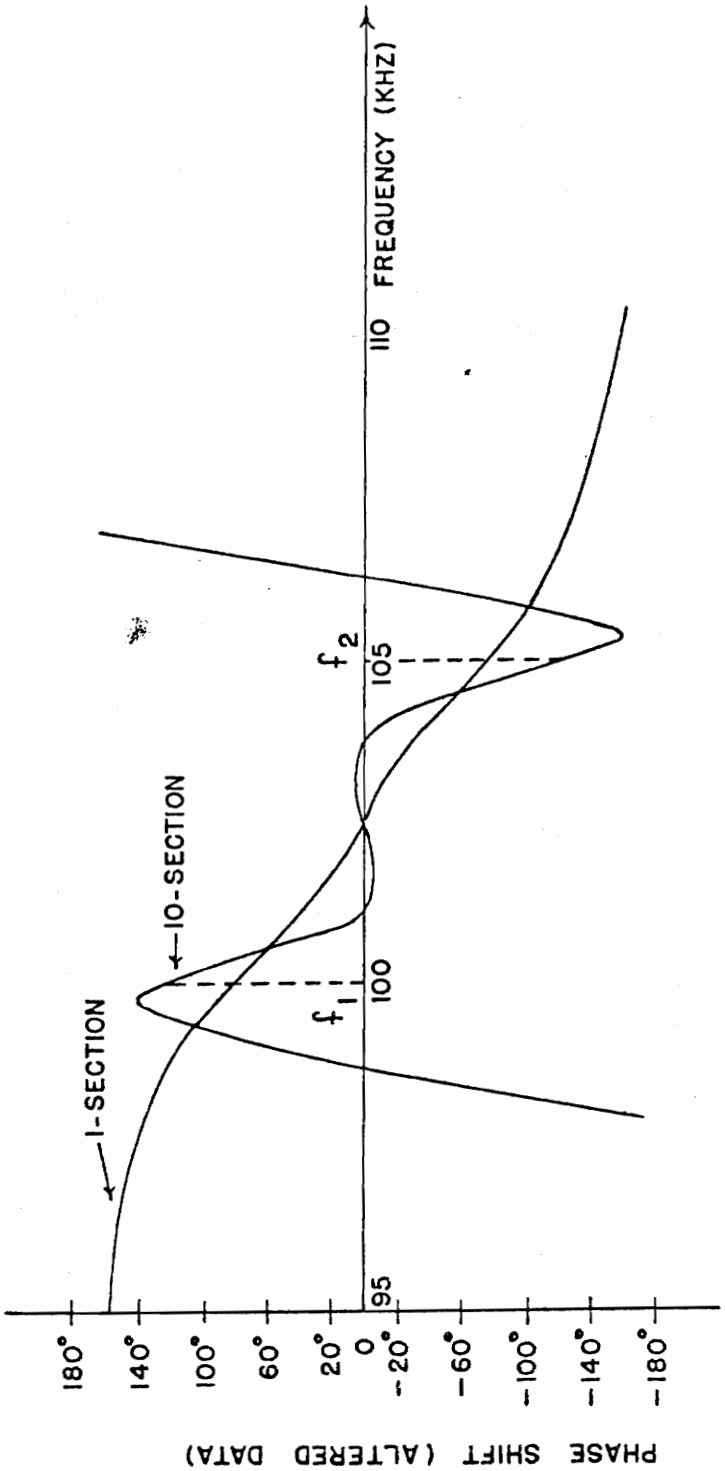


FIGURE 16 - PHASE CHARACTERISTICS FOR 10-SECTION AND 1-SECTION FILTERS SHOWING NONLINEARITIES

Exact Analysis by Fourier Method

A frequency-modulated wave is now to be applied to the input terminals of this filter and the wave at the output of the filter is to be determined. Since the Fourier analysis gives the exact response (within about 1% as indicated in Chapter III), these data will be used as reference and the values obtained from the Quasi-steady-state will be compared to them. Figure 17 illustrates how the output wave is expected to vary as a function of the number of filter sections when the frequency deviation is 2000 radians/second and the modulating frequency is 100 radians/second. One observes that the curves are not greatly distorted even for a large number of sections. This is because of the relatively large modulation index which results in a large number of closely spaced sidebands of small amplitude which are essentially confined to the passband. In addition, the larger of the sidebands is located close to the center of the passband where the phase characteristic is fairly linear. On the other hand, one would expect the distortion to increase greatly as the modulating frequency increases since the sidebands become more spread out and would tend to have significant components in the region of extreme nonlinearity. This is indeed the case as can be seen from Figure 18, which shows the output wave for various modulating frequencies when the frequency deviation is held constant.

Returning once again to Figure 17, it is interesting to make some observations concerning the phase shift (or delay). The discussion

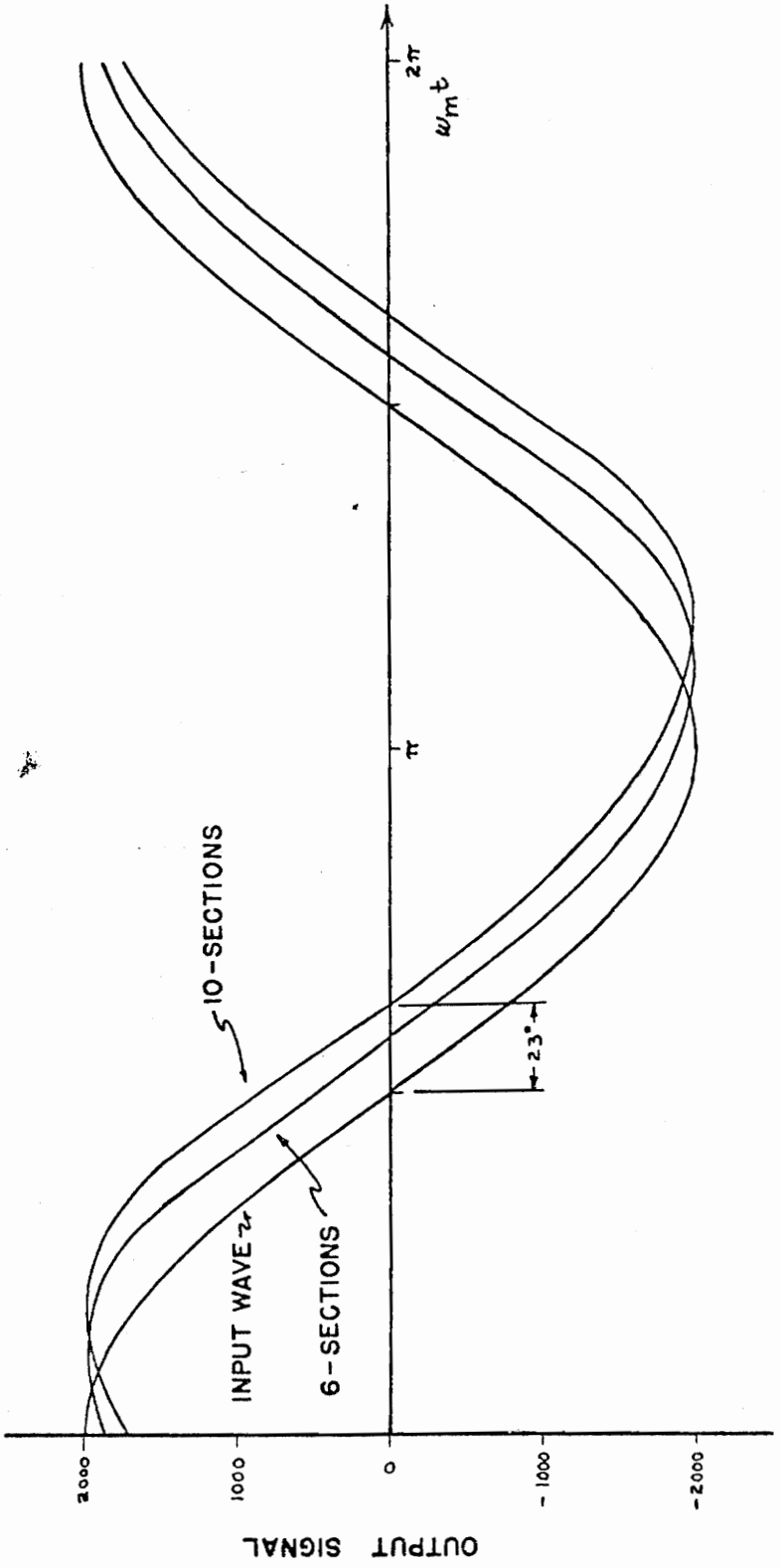


FIGURE 17–VARIATION OF OUTPUT WAVE AS A FUNCTION OF THE TOTAL PHASE SHIFT ($\Delta\omega = 2000, \omega_m = 100$)

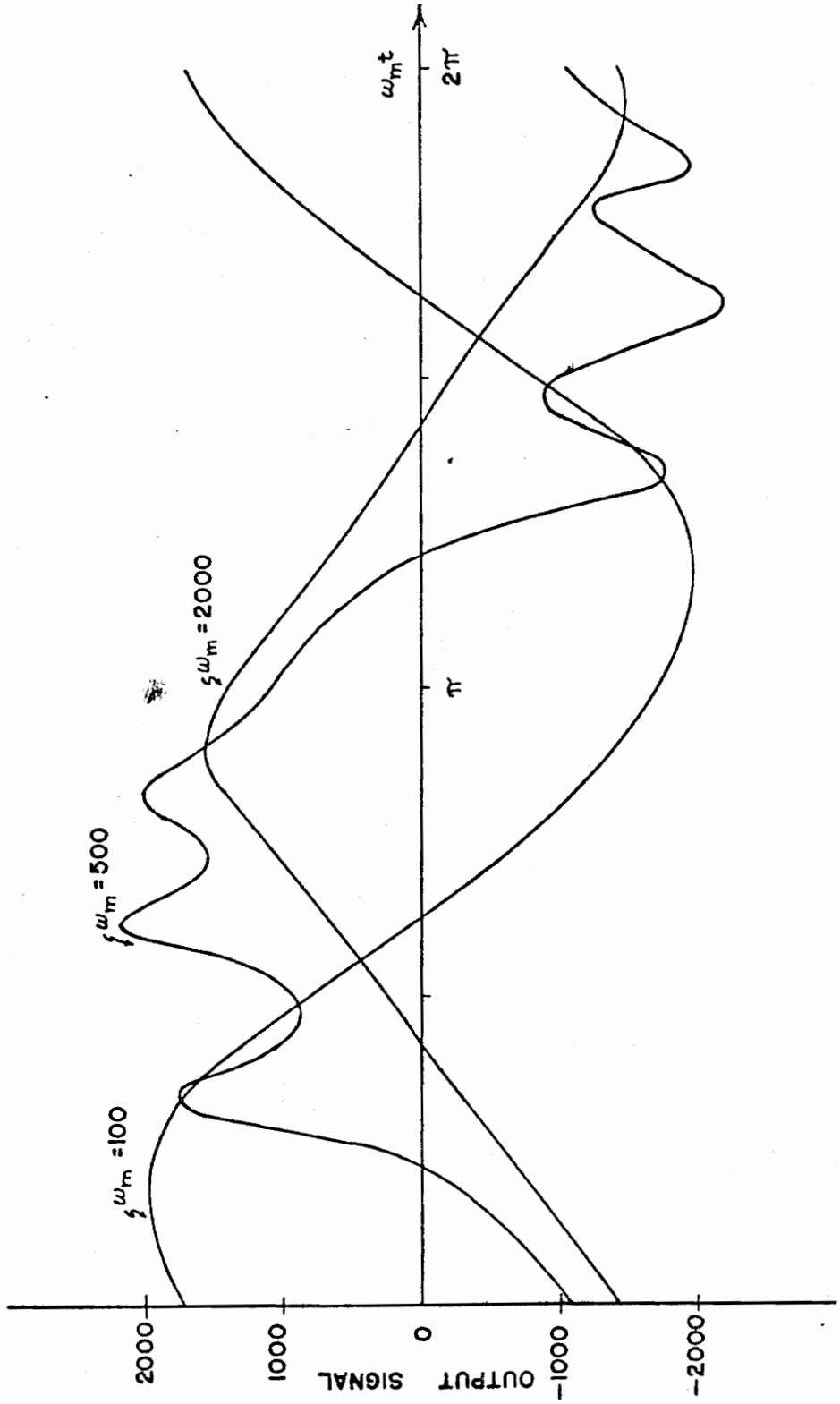


FIGURE 18—OUTPUT SIGNAL FOR 10-SECTION FILTER AS MODULATING FREQUENCY IS VARIED ($\Delta\omega = 2000$ RADIANS/SECOND)

in Chapter III indicated that the phase shift of an undistorted output wave arising from passage of the wave through a network with a linear phase characteristic should be equal in value to the phase shift occurring in the first sideband. From the plotted data one should be able to verify this theory even though the waves are somewhat distorted. Consider the case of the 10-section filter in Figure 17. The phase delay of this wave with respect to the input is approximately 23^{o}. The phase shift of the 10-section filter at the first sideband is obtained from Figure 15 as approximately 20^{o} (for a frequency of 100 hertz). The accuracy of this reading is somewhat poor due to the scale factors; therefore, the actual data were consulted which gave a value of 22.9^{o} for the phase shift at the first sideband. This value is very close to that read from Figure 17.

Analysis by Quasi-Steady-State Method

It is now desirable to investigate the Quasi-steady-state method of analysis to verify the theory presented in Chapter IV. The approach using only the first approximation was explored because of the difficulty in applying the higher-order correction terms to a complicated network. Figures 19, 20, and 21 compare the output predicted by the Quasi-steady-state method with the exact output as given by the Fourier method. The curves of Figure 19 are plotted for a 3-section filter, while the curves for the 6-section and 10-section filters are shown in Figures 20 and 21, respectively. The reason the single-section results are not shown is that the curves practically lie

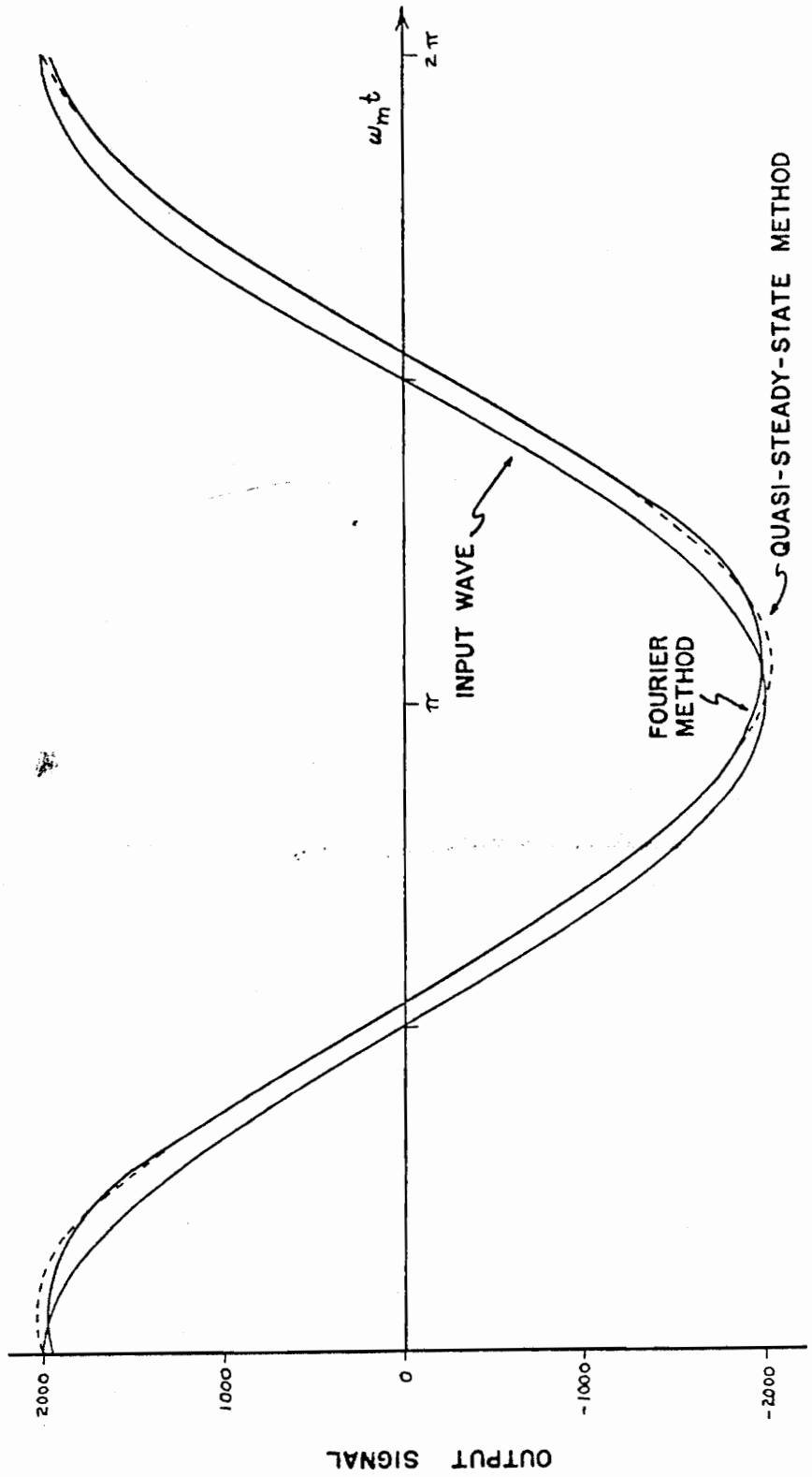


FIGURE 19 – COMPARISON OF METHODS OF ANALYSIS FOR 3-SECTION CONSTANT-k FILTER ($\Delta\omega = 2000$, $\omega_m = 100$)

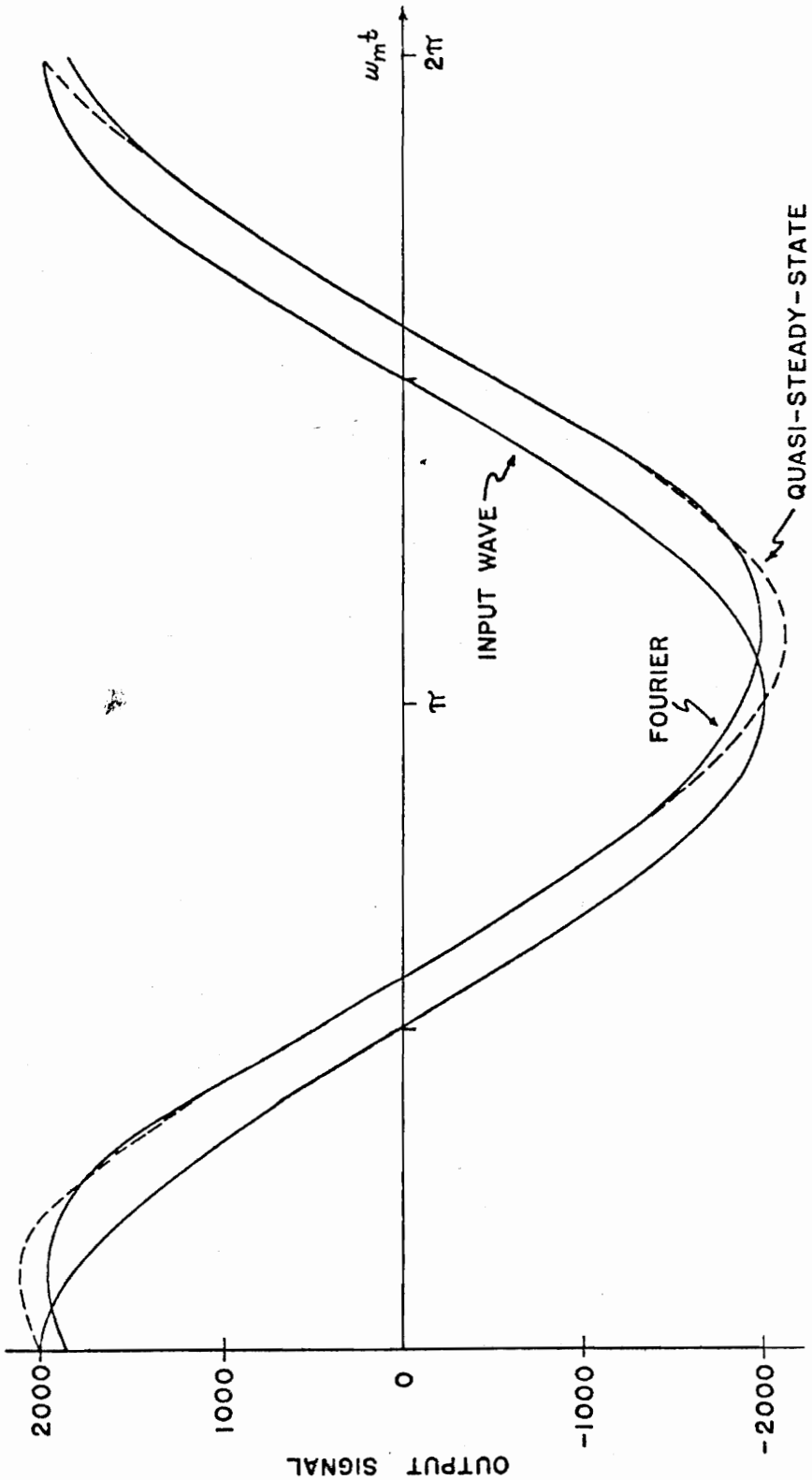


FIGURE 20-COMPARISON OF METHODS OF ANALYSIS FOR 6-SECTION CONSTANT-K FILTER ($\Delta\omega = 2000, \omega_m = 100$)

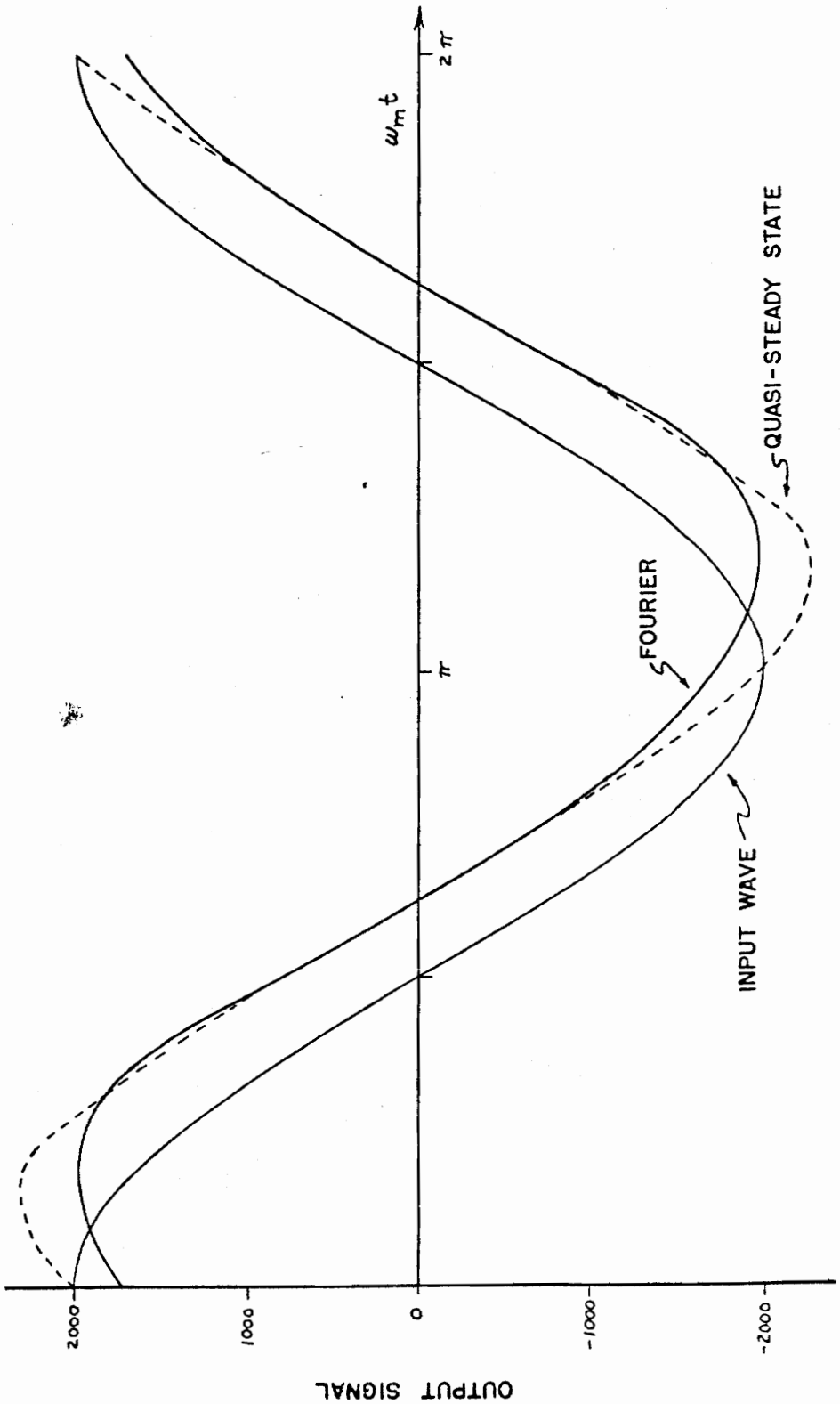


FIGURE 21-COMPARISON OF METHODS OF ANALYSIS FOR 10-SECTION FILTER ($\Delta\omega = 2000$, $\omega_m = 100$)

on top of one another, and the output wave as predicted by the Quasi-steady-state method begins to depart significantly from the exact output wave. This is in agreement with the behavior expected from the analysis of Equation (IV-10). It is significant to note that the greatest error in the Quasi-steady-state method occurs at the peak of the waves while in between the peaks the error is quite small. In fact, the predicted wave is greater in amplitude than the input wave indicating that the deviation has been increased by the filter; that is, the results of the Quasi-steady-state analysis would have one believe that as the phase shift of a network is increased, the peak frequency deviation is also increased. This comes about as a result of assuming that the output follows the input almost instantaneously which obviously cannot be so. Figure 22 shows the error in the Quasi-steady-state method as compared to the Fourier method for the cases illustrated in Figures 19, 20, and 21. The error for the case of the single-section filter is equal to or less than 1.5% at all times, so it does not appear on Figure 22.

As was stated in Chapter IV, the Quasi-steady-state method also becomes inaccurate at low values of modulation index (i.e., higher modulating frequencies). Results which verify this behavior for the case of the 10-section filter are found by comparing Figures 21 and 23. Higher modulating frequencies result in even more error than Figure 23 shows. Of course, the explanation of this phenomenon is simply that as the frequency increases, the network can no longer respond in the fashion assumed by the Quasi-steady-state analysis.

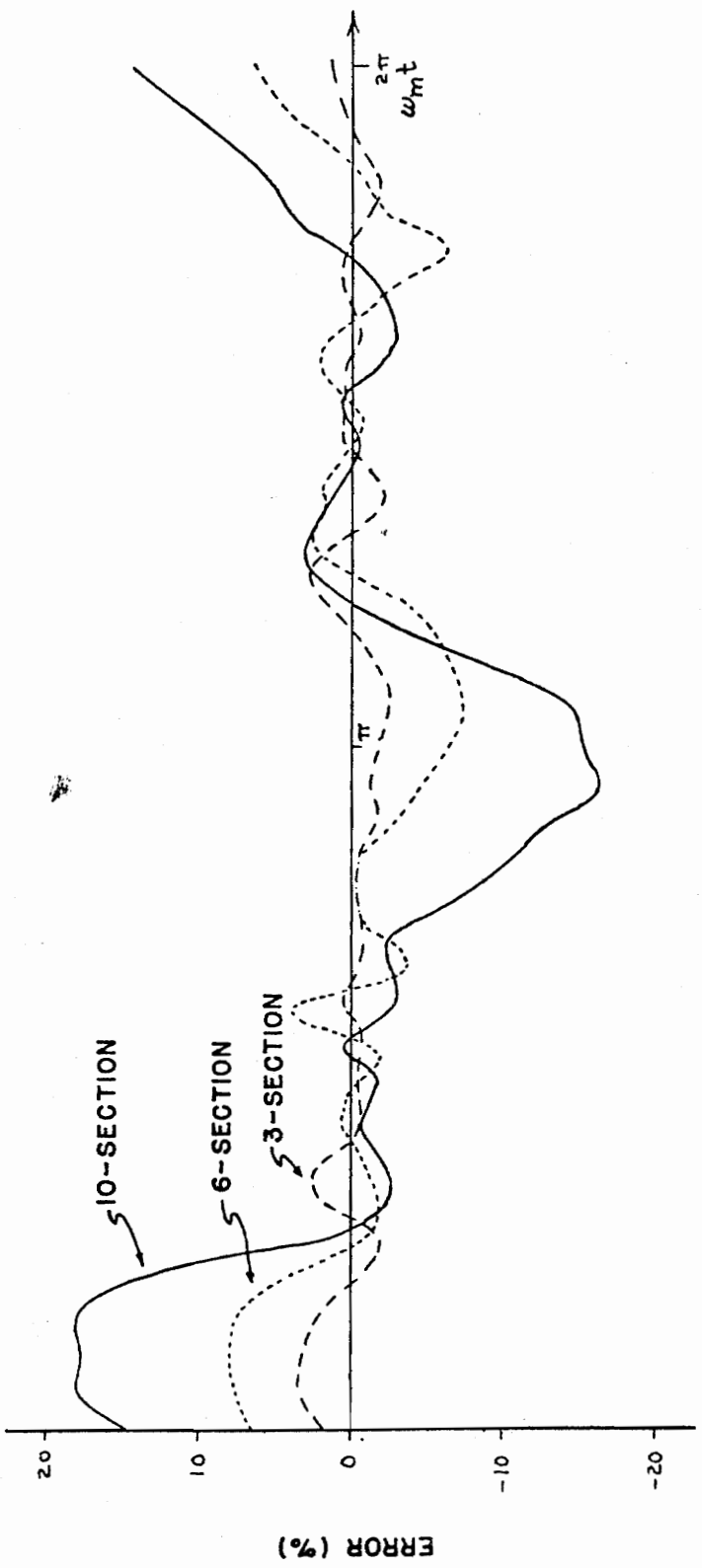


FIGURE 22-ERROR IN QUASI-STEADY-STATE METHOD AS BASED ON THE FOURIER METHOD ($\Delta\omega = 2000, \omega_m = 100$)

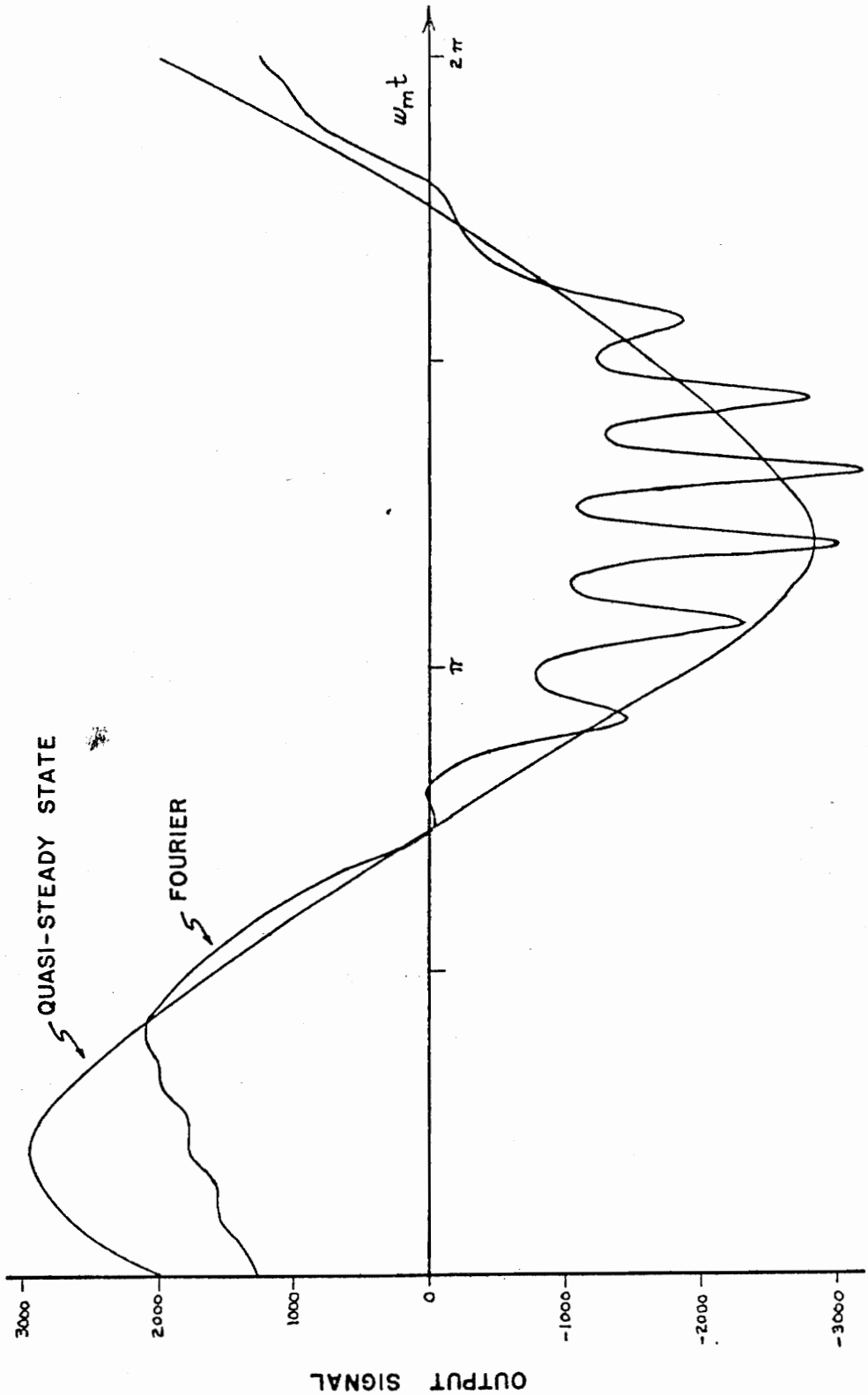


FIGURE 23–COMPARISON OF METHODS OF ANALYSIS FOR 10-SECTION FILTER ($\Delta\omega$=2000, ω_m=200)

VI. ANALOG COMPUTER SIMULATION

As a means of experimentally verifying the theory set forth in the foregoing chapters, an FM system was simulated on the analog computer. Until recently, the simulation of electronic systems in which there occur widely separated frequencies has been considered to be impractical due to difficulties encountered in time scaling such systems. For example, consider the case of a conventional FM broadcast receiver which receives radio frequency waves around 100 megahertz. These RF waves are then converted into an intermediate frequency of 10.7 megahertz which in turn are passed through a detector to yield the intelligence-bearing signal which lies in the audio range of 50-15,000 hertz. An attempt to simulate this system on the analog computer as it stands would probably be unsuccessful, since time scaling, which would result in translating the carrier frequency down to the usable range of the computer (normally from DC to a few thousand hertz), would result in audio frequencies so low that they would become indistinguishable from DC drift. More successful simulation schemes are based on the fact that the higher frequencies do not explicitly make any contribution to the intelligence being transmitted [20,21]; that is, as long as the frequency of the carrier is high enough to pass all of the sidebands involved, the same intelligence could be transmitted at a much lower carrier frequency than is conventionally employed. In fact, since the carrier wave is merely the vehicle of propogation, one is free to choose almost any convenient carrier frequency for the purpose of analog

simulation. This choice would be subject to some rather loose restrictions. First, one must make sure that the frequency is high enough so that all sidebands of the frequency-modulated wave are transmitted without interference. Next, the frequency should be high enough so that wave forms throughout the simulated system have the same characteristics as those of the actual system under study. This is especially true whenever waveforms such as the envelope of the carrier wave are to be studied. Finally and most importantly, the frequency must be low enough so that the relatively limited bandpass of the computer amplifiers does not degrade the performance of the overall system.

There are many advantages to using analog computer simulation over more conventional experimental techniques in the case of complex systems such as the one studied here. In general, the simulation technique allows one to predict, study, and optimize a system without actually constructing it. It was especially advantageous when applied to the present system inasmuch as it avoids certain problems that would be encountered at the higher frequencies usually employed. Thus, the use of high-frequency techniques in construction and measurement, which are more of an art than a science, are unnecessary. In addition, the components of the system are completely accessible for purposes of alteration or making detailed studies. In the case of the present problem, there exists the additional advantage of being able to simulate any linear network for which the transfer function can be obtained (of course, this implies that one has at his disposal a

relatively large analog computer).

Configuration of Simulated System

The block diagram for the complete simulated system is shown in Figure 24. The blocks inside of the dashed lines represent the portion of system which is replaced by the analog computer while the other blocks represent necessary peripheral equipment. The analog computer actually used in this set-up was a combination of two Heath Model ES-400 units slaved together. While having serious limitations, the availability and ease of operation of these units more than justified their use. Some of these limitations will be discussed later in the chapter. A Hewlett-Packard Model 3300A function generator was used as the FM modulator. This piece of equipment was operated as a voltage-controlled oscillator with the modulating voltage being supplied from a Hewlett-Packard Model 202A low-frequency function generator. A frequency of 100 hertz was selected as the frequency of the unmodulated carrier based on the factors already discussed above. Maximum frequency deviations up to approximately five hertz were employed during the course of this investigation although this figure should not be construed to be any sort of limiting value as far as the system is concerned. By using values of frequency deviation in the range $0 \leq \Delta f \leq 5$ and modulating frequencies in the range $0.1 \leq f_m \leq 1.0$ one is able to obtain values of the modulation index from zero to 50. Thus, the modulation index has roughly the same range as was found to be reasonably economical in the case of the digital computer analysis.

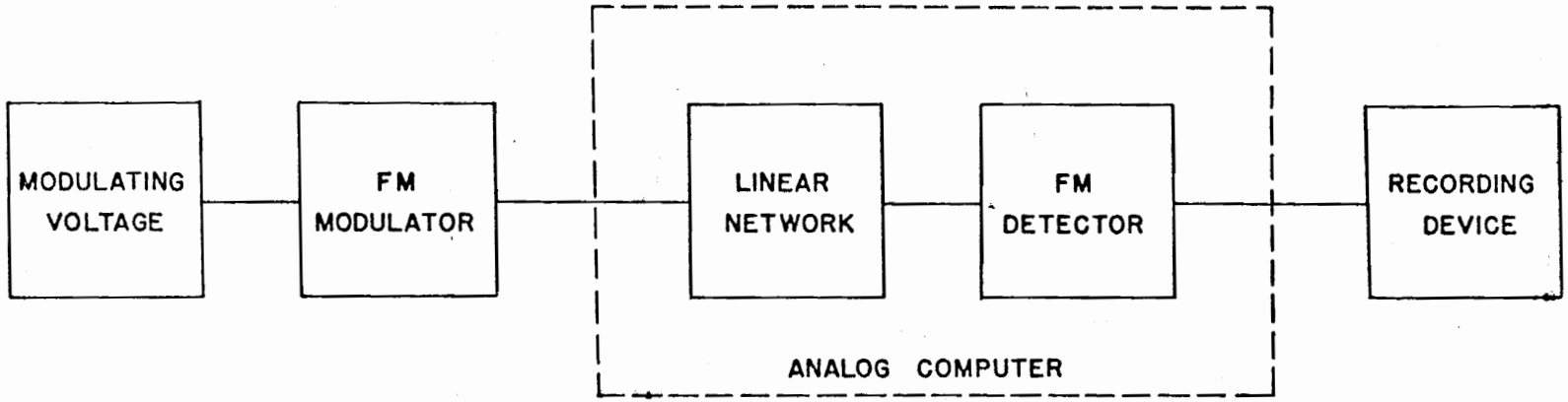


FIGURE 24 - SIMULATION OF FM DISTORTION PROBLEM

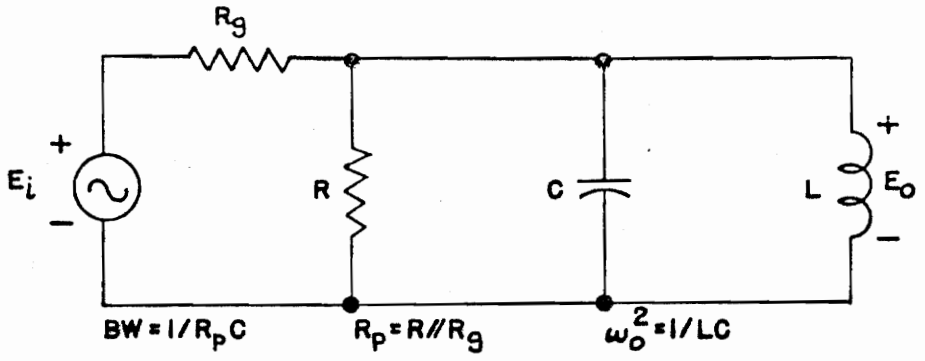
The recording device used was a Tektronix Model 564 storage oscilloscope. This piece of equipment proved to be almost indispensible for viewing the details of the very low frequency waveforms.

The heart of the simulator is the FM detector which can be further broken down into three basic components as illustrated in Figure 25. The actual analog computer circuits represented by each of these blocks will be postponed briefly in favor of investigating the simulation of a single-tuned parallel RLC network. This circuit will form the basis for several of the analog computer programs which follow. Referring to Figure 26(a) and writing the node equation at the output terminal, one obtains after some simplification

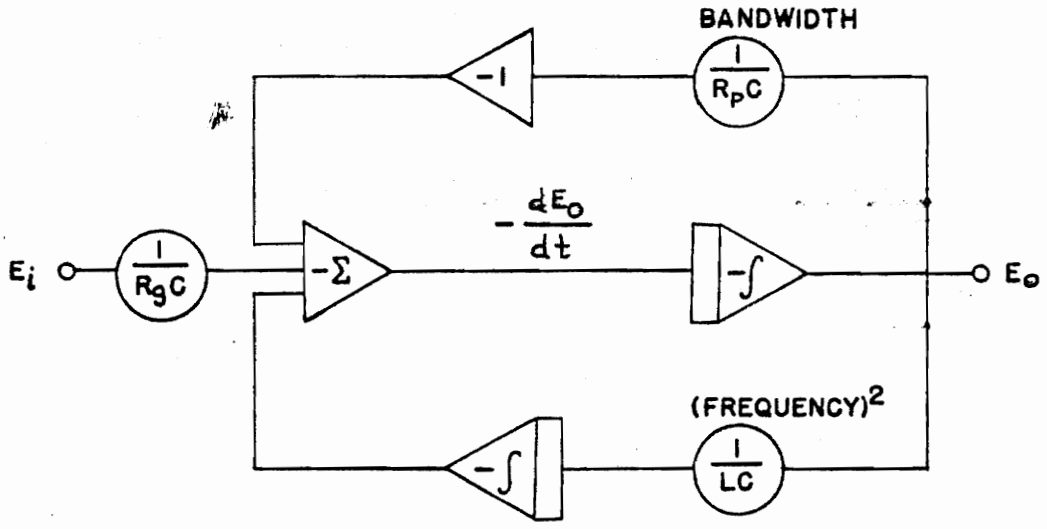
$$\frac{dEo}{dt} = \frac{Ei}{RgC} - \frac{Eo}{RpC} - \frac{1}{LC}\int_{o}^{t} E_o dt \ . \qquad \text{(VI-1)}$$

The unscaled computer program for this equation is shown in Figure 26(b). Note that the bandwidth and the frequency are controlled by independent potentiometers resulting in a very versatile circuit. One use of the single-tuned circuit will be in the construction of a frequency discriminator. When a pair of single-tuned circuits are connected as shown in Figure 27 so that their individual amplitude-versus-frequency characteristics subtract, the result is a frequency discriminator with an overall characteristic as shown in Figure 28. In this case, the two frequency controls determine the bandwidth of the discriminator by determining f_1 and f_2 while the individual bandwidth control potentiometers determine the steepness of the linear portion of the curve. Figure 28, which was obtained during

FIGURE 25 – BLOCK DIAGRAM OF FM DETECTOR



(A) SINGLE-TUNED CIRCUIT



(B) UNSCALED COMPUTER PROGRAM FOR (A)

FIGURE 26 —ANALOG SIMULATION OF SINGLE-TUNED CIRCUIT

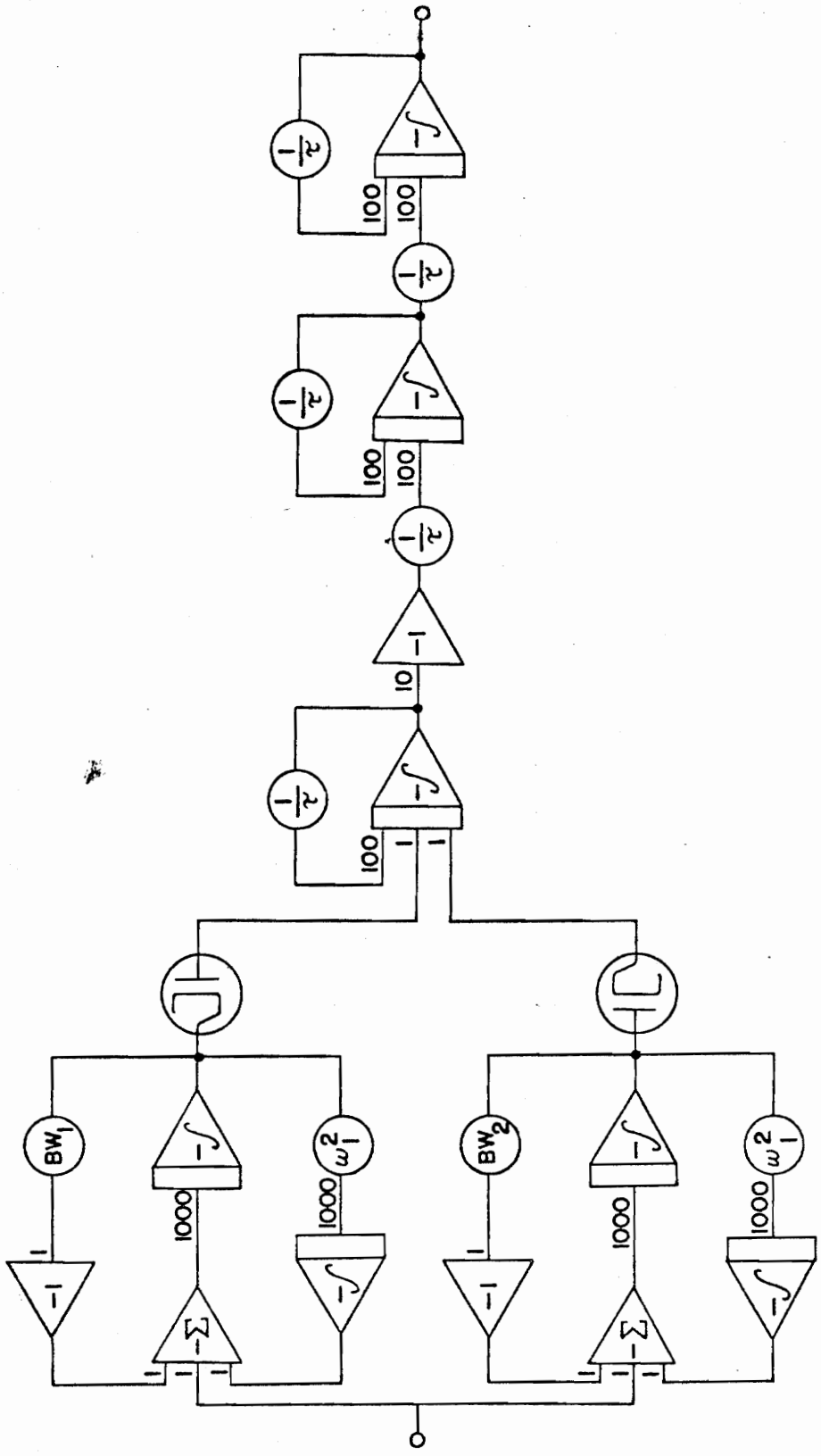


FIGURE 27 – DISCRIMINATOR AND LOW-PASS FILTER CIRCUITS

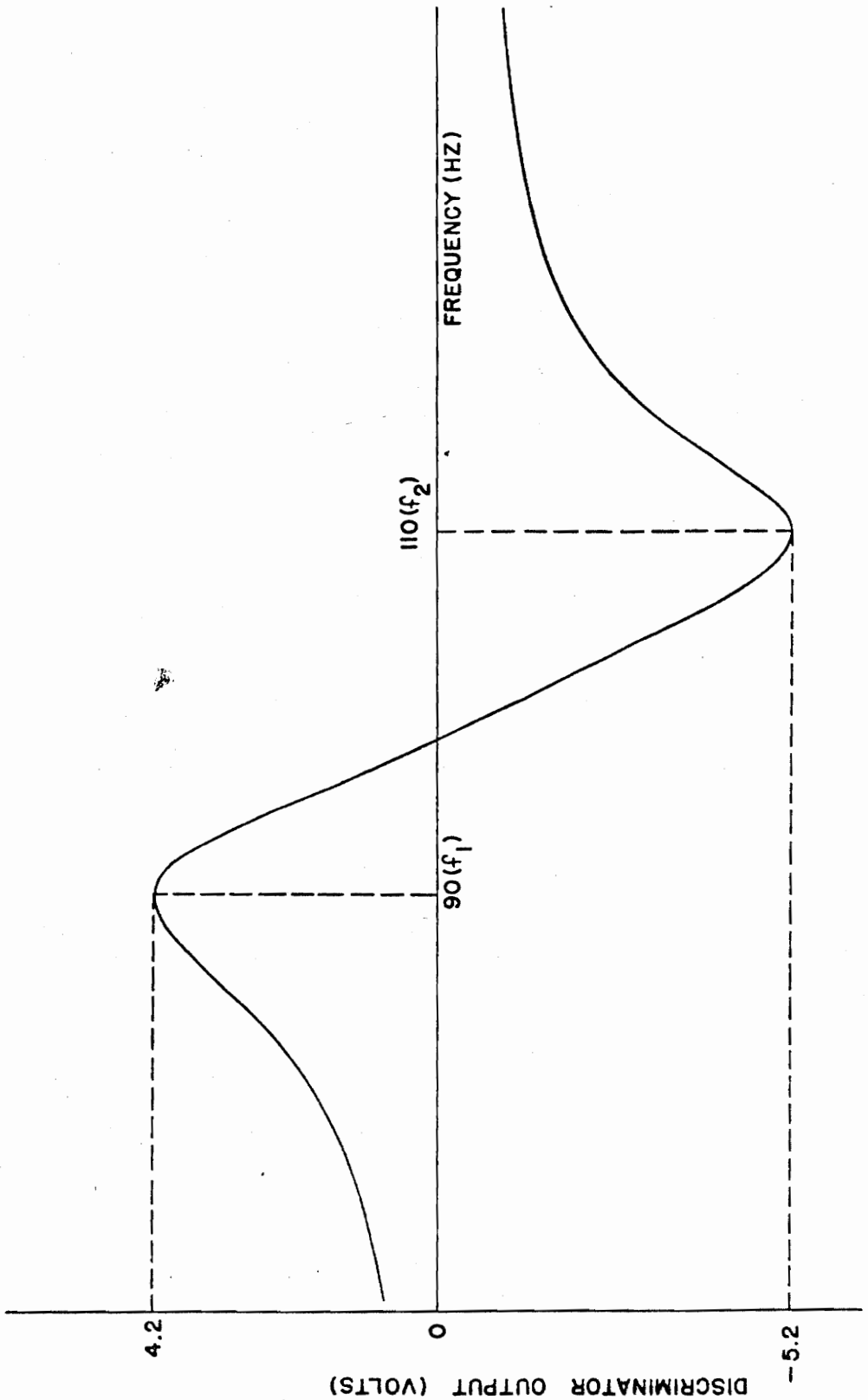


FIGURE 28 – CHARACTERISTIC OF SIMULATED FM DISCRIMINATOR

actual operation, shows a bandwidth in the linear region of about 15 hertz which is more than adequate to accomodate the wave discussed earlier in this chapter. Referring again to Figure 27, the four amplifiers to the right of the diodes comprise a three-stage low-pass filter with an amplifier to boost the signal. The output of the low-pass filter is essentially free from higher frequencies. Although this is the desired behavior, there are drawbacks. These will be pointed out later in conjunction with some of the experimental results.

In Chapter III it was pointed out that a frequency-modulated wave after passing through a frequency selective network usually exhibits amplitude modulation. Since the discriminator is amplitude sensitive as well as being frequency sensitive, this causes the problem of simultaneously detecting both AM and FM. This problem is practically eliminated by the use of a limiter ahead of the discriminator as shown in Figure 27. The analog computer simulation which accomplishes this operation is found in Figure 29. The limiter is composed of a single-tuned resonant circuit connected to the output of an amplifier which is driven into saturation on the positive and negative peaks of the input wave. The degree of saturation of the amplifier is easily controllable by the two potentiometers which are connected to the reference voltages. For the particular connection shown in Figure 29, the point of saturation of the input wave is variable from zero to 50 volts with the negative and positive saturation points being independently controllable. The output wave which has an appearance approximating a square wave is then applied to the input of the single-tuned resonant

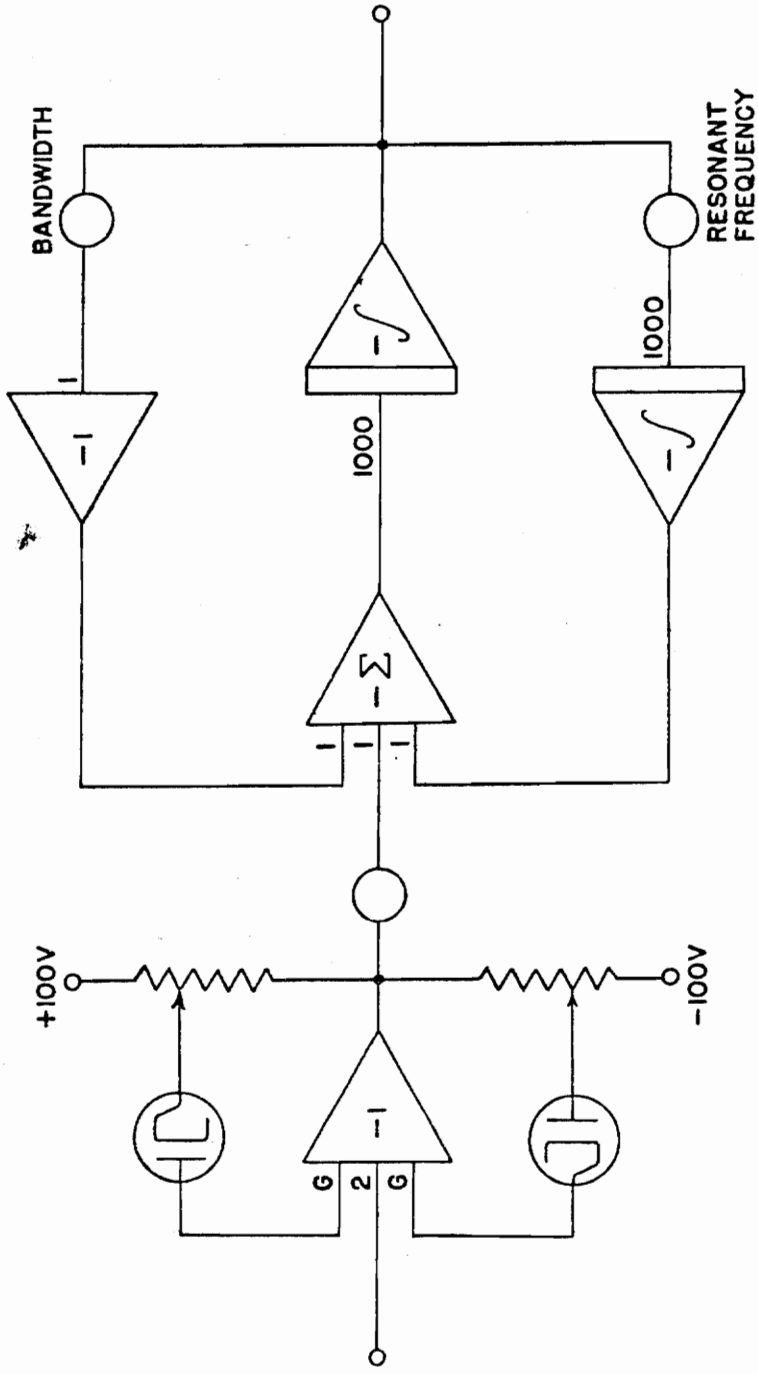


FIGURE 29 —ANALOG COMPUTER PROGRAM FOR LIMITER

circuit. The tuning of this circuit is rather broad so as to pass the carrier frequency and its sideband components while filtering out the undesirable frequency components such as contained in the "corners" of the square wave. Also, broad tuning is necessary in order to assure that the tuned circuit itself does not introduce amplitude modulation. The net result is once more an FM wave with negligible amplitude modulation. This wave can now be applied to the discriminator to obtain the true demodulated output without interference due to amplitude modulation.

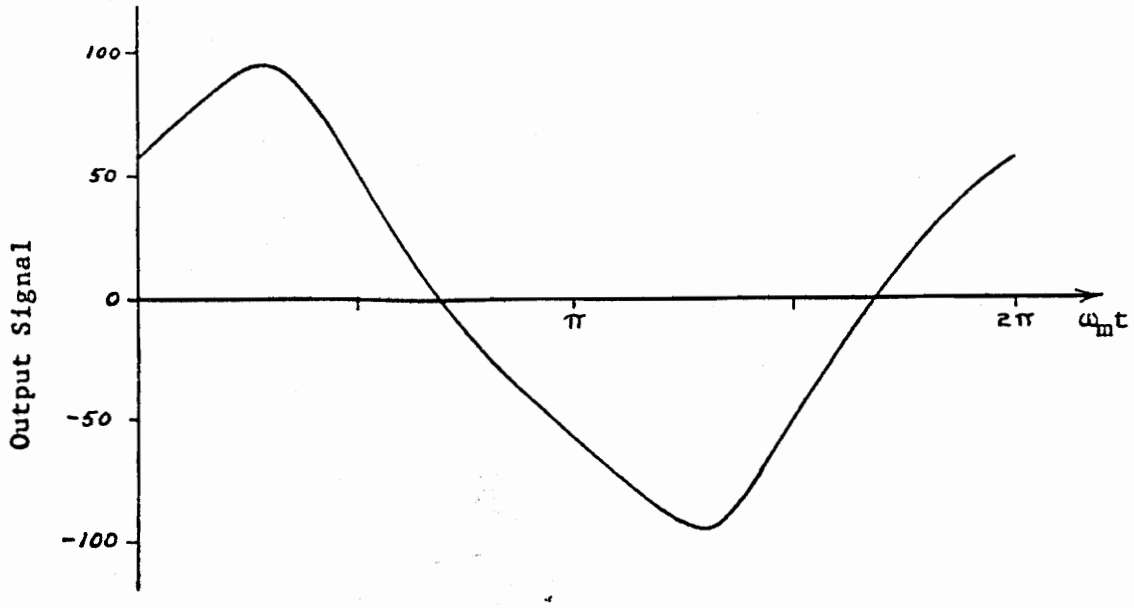
Comparison of Experimental Results with Theory

For the experimental investigation of the theory presented in Chapter III, a single-tuned resonant circuit was chosen as the linear network represented by the block labeled "linear network" in Figure 24. This network has the same configuration as shown in Figure 26 and used several times previously, except in the present case it is sharply tuned to purposely introduce distortion. A resonant frequency of 10,000 radians/second, a bandwidth of 100 radians/second, and a mid-band gain of 0.5 were selected as characteristics for the tuned-circuit. The FM wave applied to the input of this network was assumed to have a carrier frequency of 10,000 radians/second while the peak frequency deviation and the modulating frequency were allowed to take on different values.

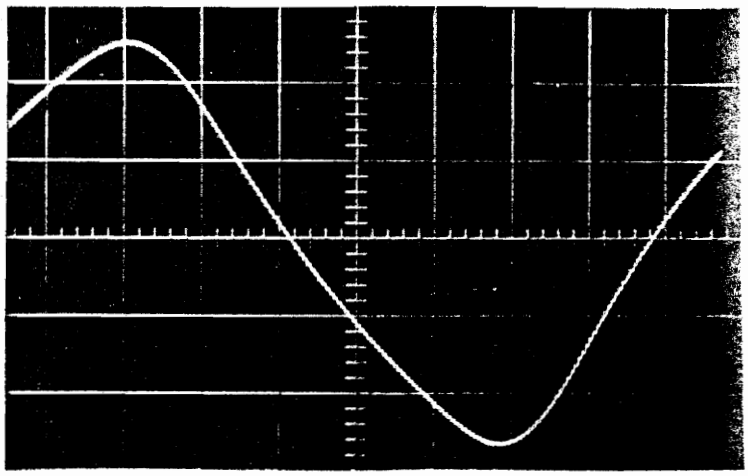
For the analog simulation, the tuned circuit was operated at a

resonant frequency of 100 hertz with a bandwidth of one hertz. Since the carrier frequency must be the same as the resonant frequency of the tuned circuit, its value must also be 100 hertz. The frequency of the analog simulation has been lowered by a factor of $100/2\pi$ which is observed from the bandwidth requirements in the two cases above. It should be re-emphasized that it is not necessary for the two carriers to be related by the time-scale factor. The present case is merely a coincidence, and the same results would be obtained if the actual carrier frequency were allowed to be 100 kilohertz as long as the bandwidth remained at 100 radians/second. It will be noted that the mention of circuit Q has been avoided since if this quantity is specified, the ratio of carrier frequency to bandwidth is fixed. Thus one would have to scale the carrier frequency by the same factor that is used for the bandwidth. As mentioned previously, this requirement would place a severe limitation on the practical aspects of the simulation. For this problem, the bandwidths must be scaled by the same factor as are the peak frequency deviations and the modulating frequencies. The only requirement on the carrier frequency is that it be equal to the resonant frequency of the tuned circuit.

Based on the above information, the digital computer was used to analyze the distortion for several different combinations of frequency deviation and modulating frequency. The results of these predictions are shown in Figures 30-37 along with photographs obtained from the experimental apparatus. The agreement between the predicted and experimental waveforms is extremely good in all cases. The only

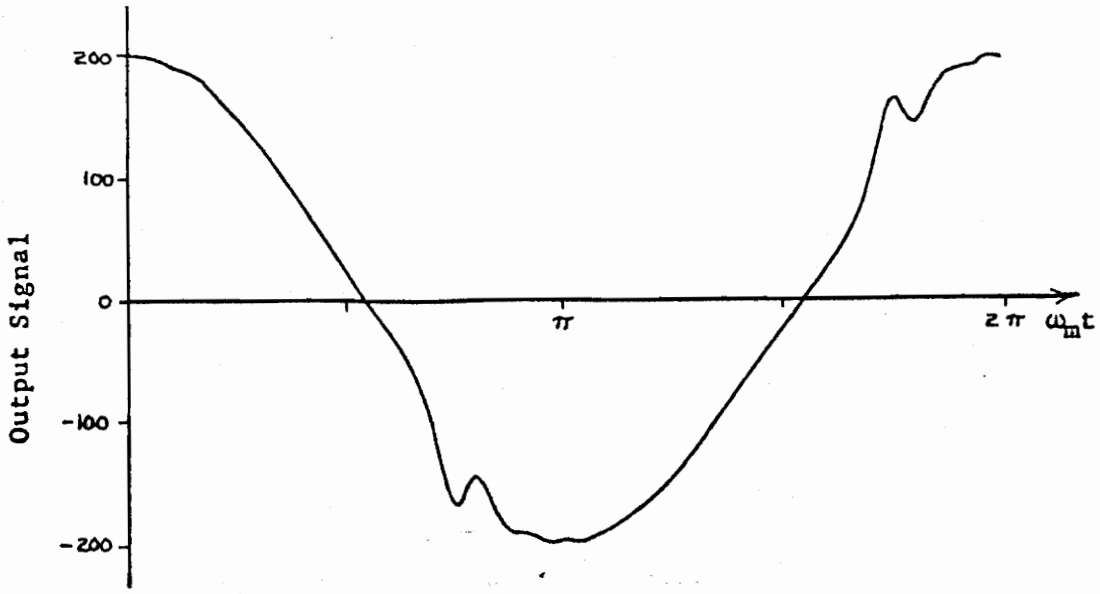


(a) Output signal predicted by digital computer.

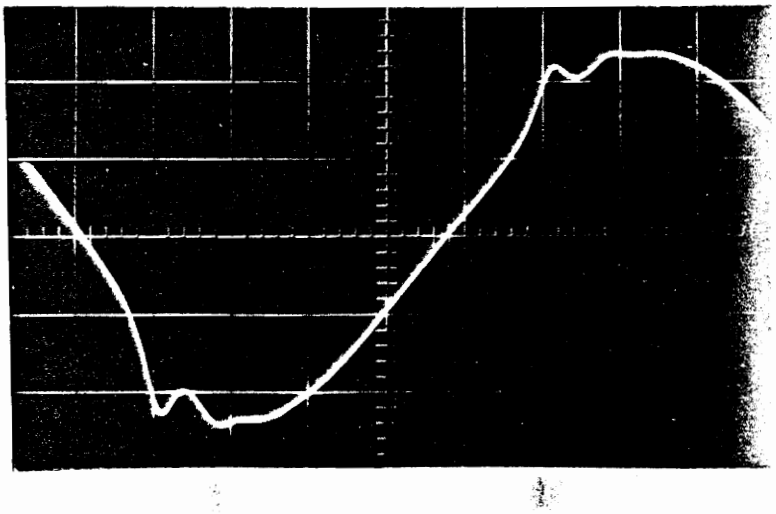


(b) Experimental waveform from analog simulation.

Figure 30. Output waveforms for $\Delta\omega = 100$ and $\omega_m = 50$ radians/second.

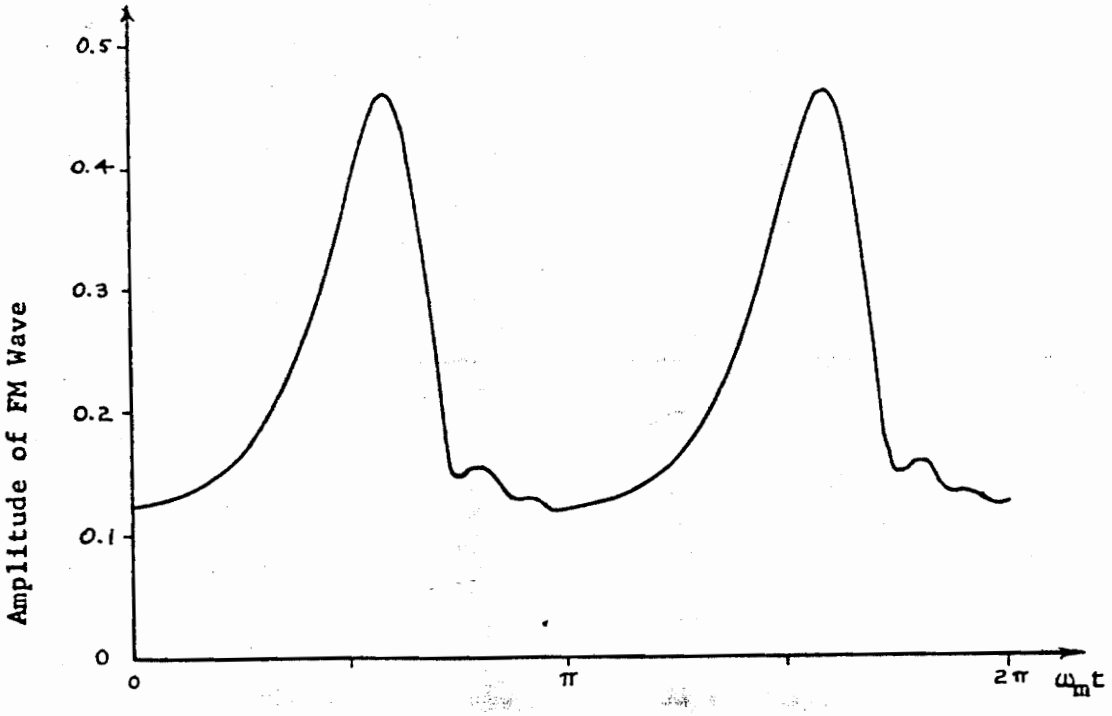


(a) Output signal predicted by digital computer.

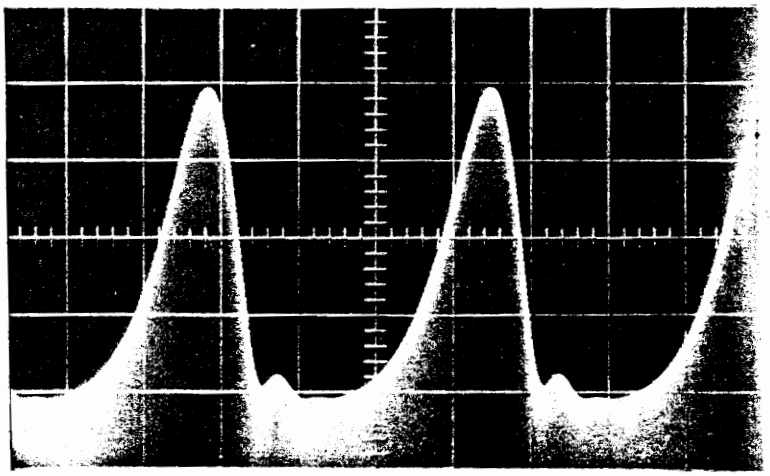


(b) Experimental waveform from analog simulation.

Figure 31. Output waveforms for $\Delta\omega = 200$ and $\omega_m = 10$ radians/second.

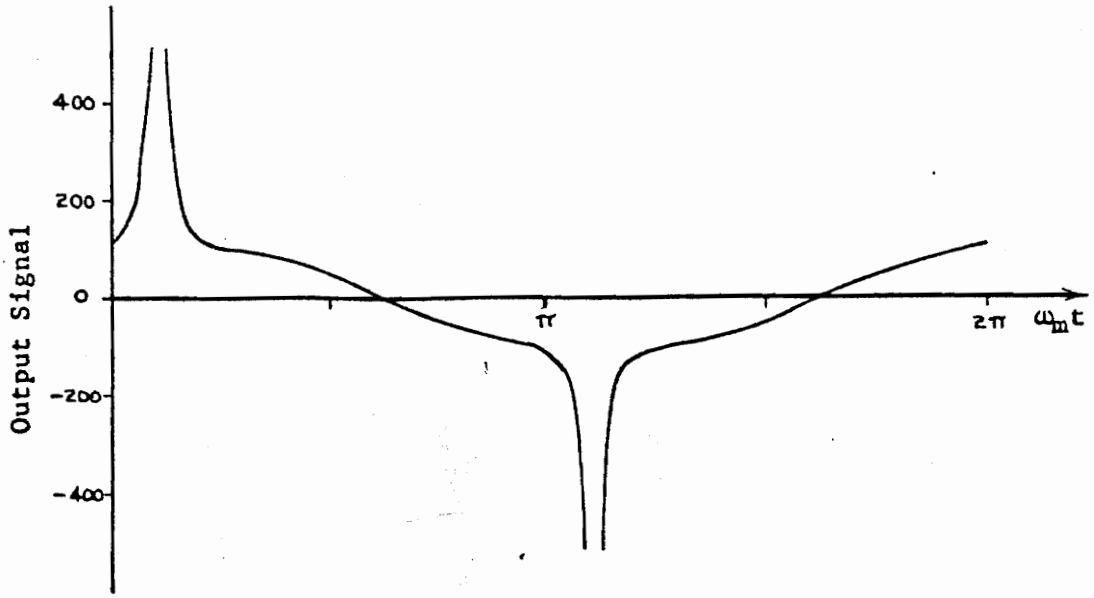


(a) Envelope of FM wave as predicted by digital computer.

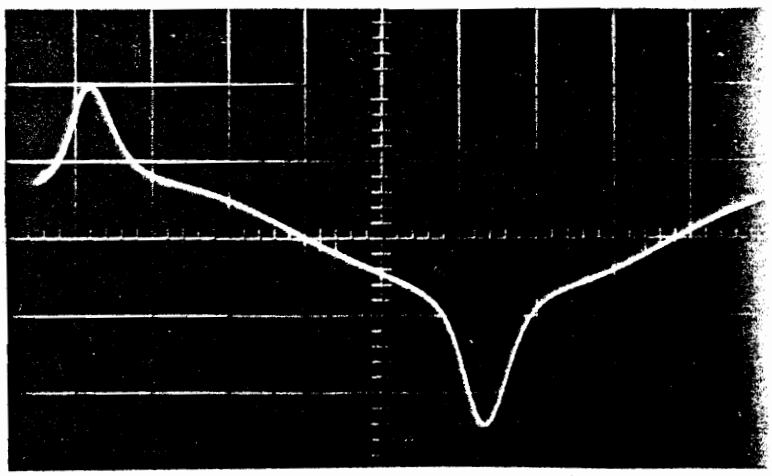


(b) Experimental waveform from analog simulation.

Figure 32. Envelope waveforms showing amplitude modulation ($\Delta\omega = 200$ $\omega_m = 10$ radians/second).

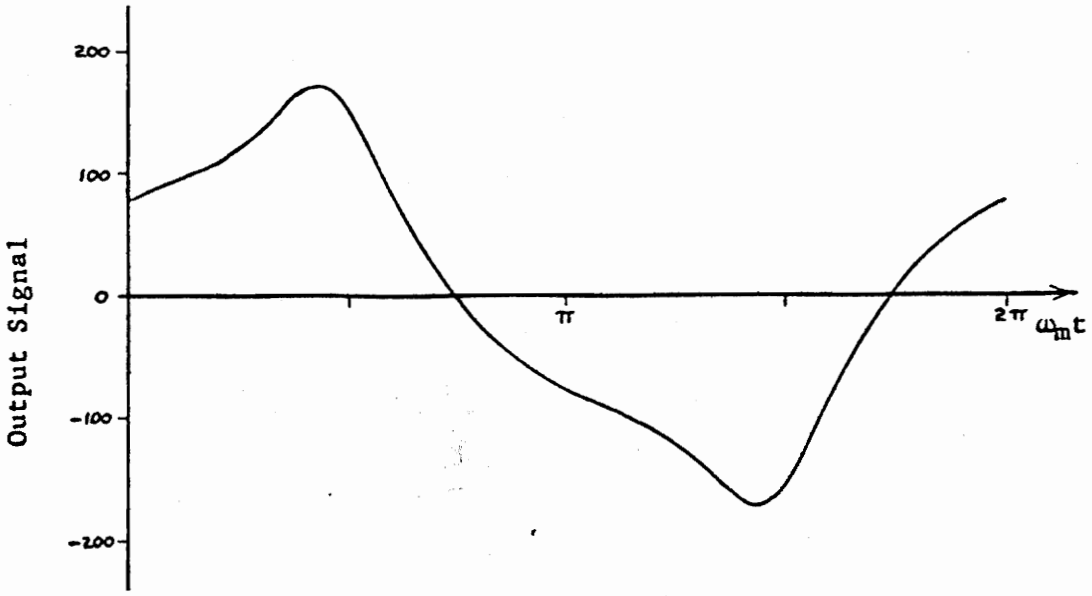


(a) Output signal predicted by digital computer.

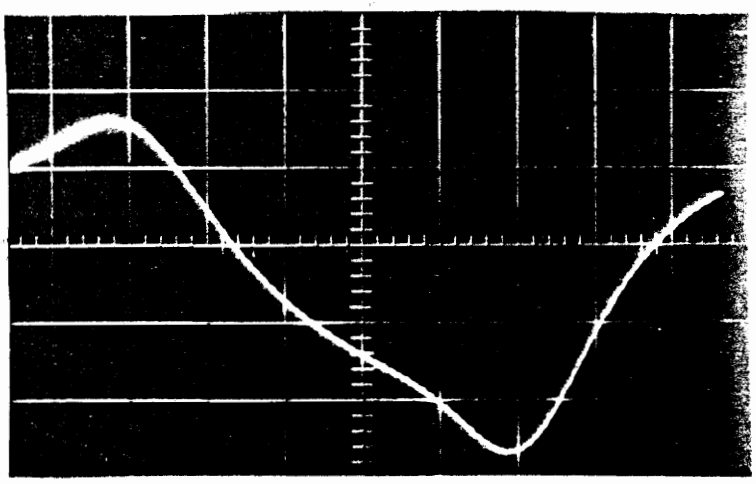


(b) Experimental waveform from analog simulation.

Figure 33. Output waveforms for $\Delta\omega = 200$ and $\omega_m = 50$ radians/second.

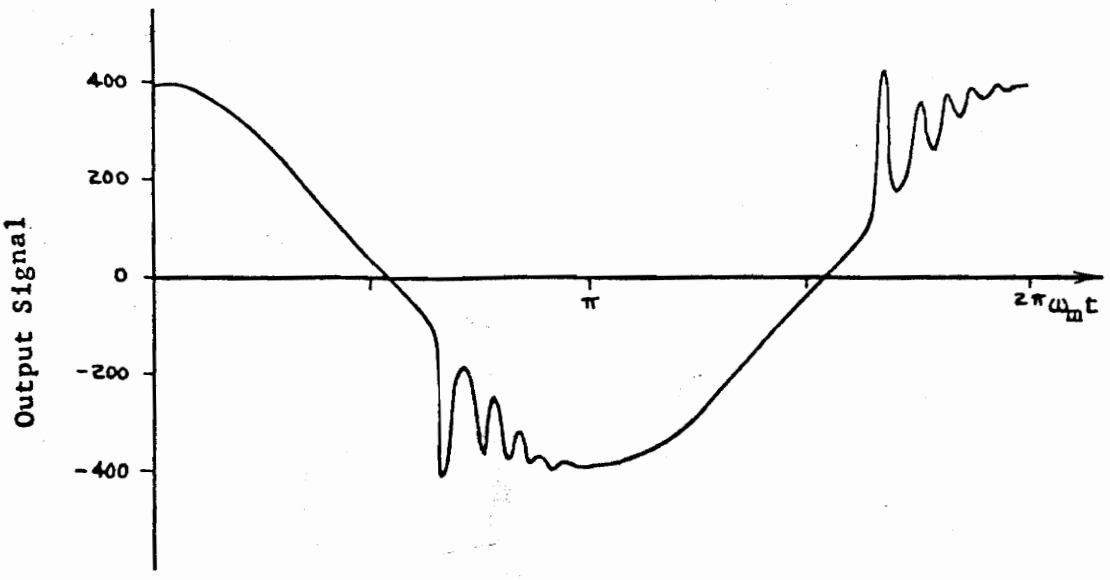


(a) Output signal predicted by digital computer.

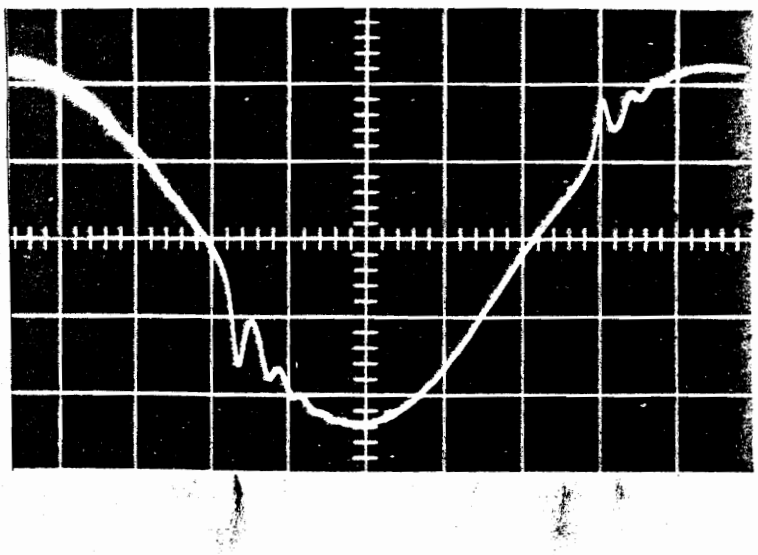


(b) Experimental waveform from analog simulation.

Figure 34. Output waveforms for $\Delta\omega = 200$ and $\omega_m = 100$ radians/second.

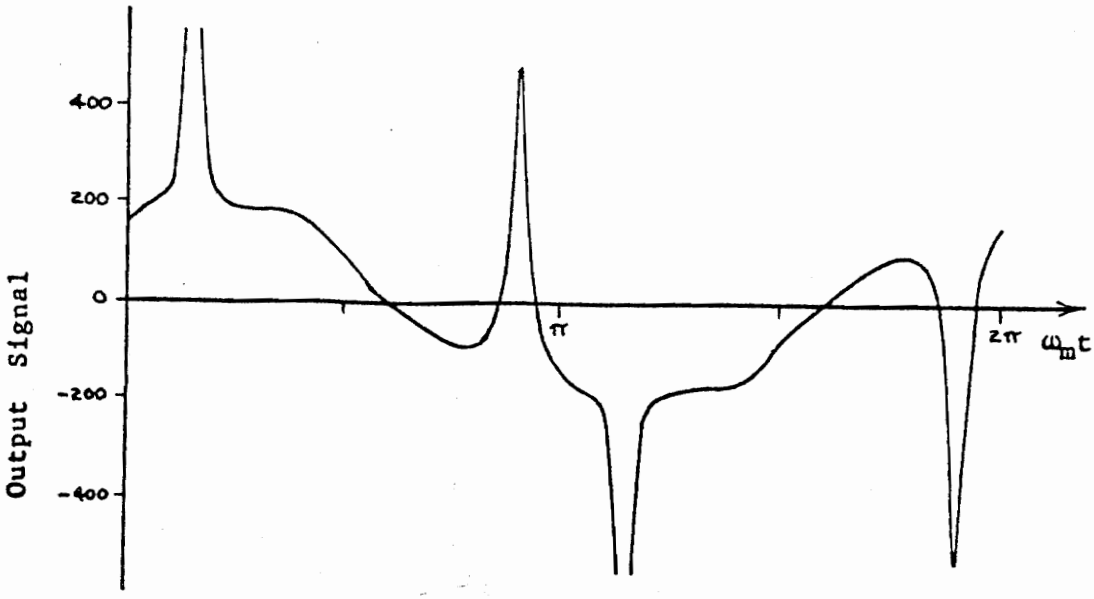


(a) Output signal predicted by digital computer.

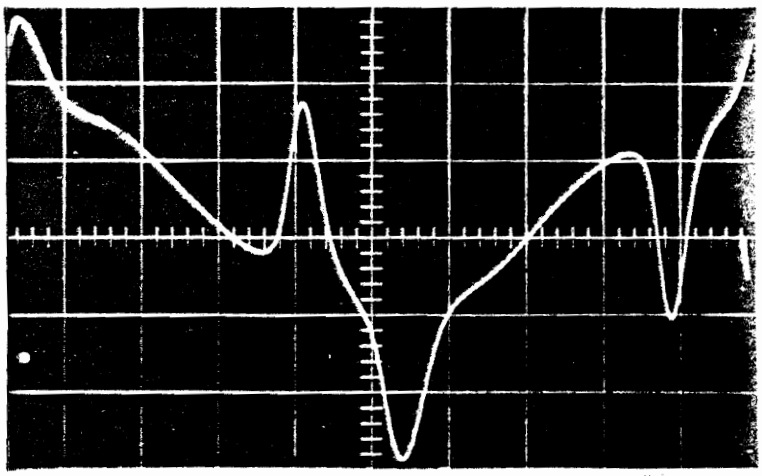


(b) Experimental waveform from analog simulation.

Figure 35. Output waveforms for $\Delta\omega = 400$ and $\omega_m = 10$ radians/second.

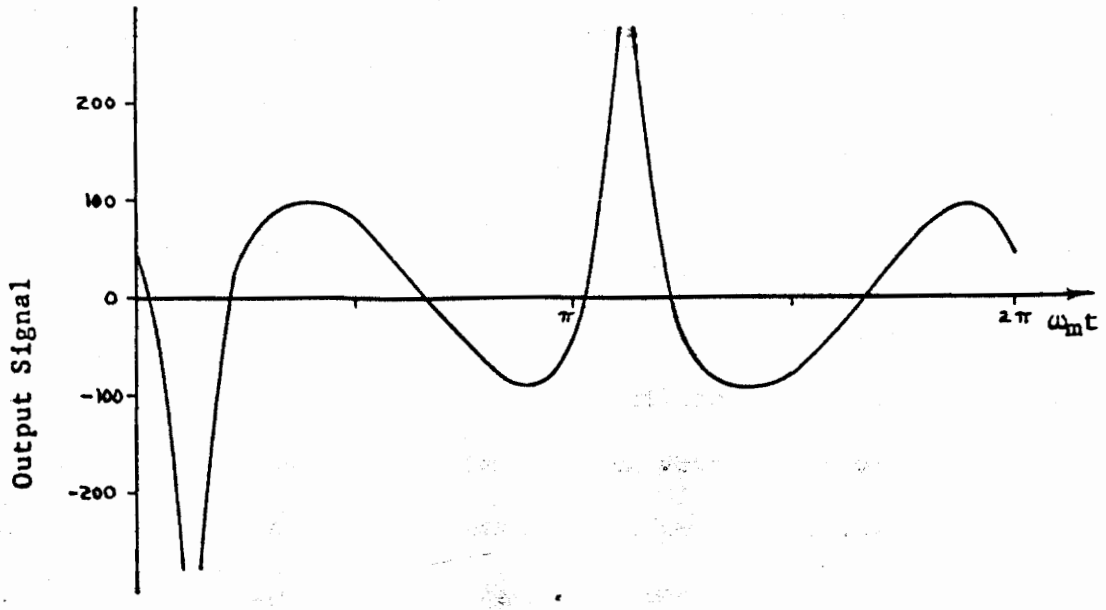


(a) Output signal predicted by digital computer.

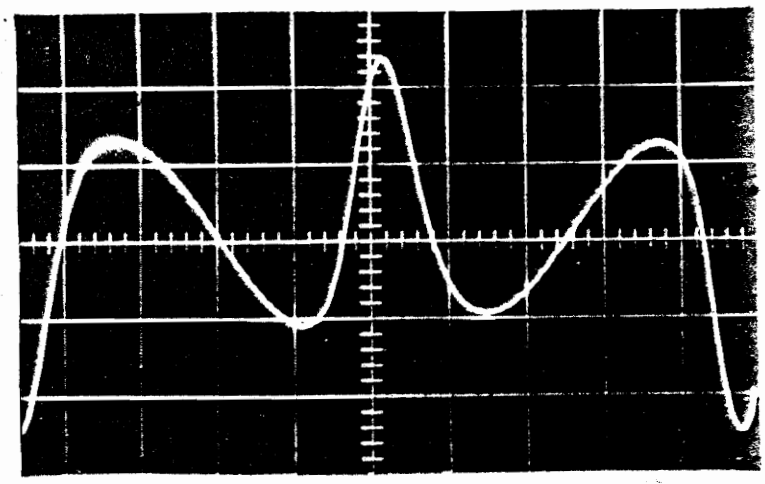


(b) Experimental waveform from analog simulation.

Figure 36. Output waveforms for $\Delta\omega = 400$ and $\omega_m = 50$ radians/second.



(a) Output signal predicted by digital computer.



(b) Experimental waveform from analog simulation.

Figure 37. Output waveforms for $\Delta\omega = 400$ and $\omega_m = 100$ radians/second.

discrepancies of notable significance occur when the predicted waveform exhibits very large amplitude spikes. In these cases, the experimental results show the proper spikes but at much lower amplitudes. This behavior is illustrated in Figures 33, 36, and 37 and to a lesser extent in Figure 35. The reason for the apparent conflict between theory and experiment is due to poor high frequency response in the simulated system caused by the low-pass filter network following the discriminator. A solution to the problem would be to use a higher carrier frequency so that the break frequency of the low-pass filter could be high enough to essentially pass the spikes while still effectively eliminating the carrier. This could easily be done on some of the more modern and elaborate analog computers but was not considered to be practical using this experimental arrangement. In fact, the results were good enough to completely verify the theory from the data obtained. In all cases the lower frequency characteristics of the waveforms were in very close agreement.

The waveforms shown in Figure 32 offer further proof of the validity of the digital computer analysis. These waveforms differ from the other waves shown inasmuch as they show the amplitude modulation existing on the waveform after passing through the linear network and before limiting. The comparison between the predicted wave and the experimental wave is remarkable. The results presented in this chapter leave no doubt as to the validity of the digital computer analysis.

VII. CONCLUSIONS

Two methods of predicting the distortion which is generated in frequency-modulated waves during passage through linear networks have been investigated in this paper. The major emphasis has been placed on the Fourier method while the Quasi-steady-state analysis received attention only to the extent of discovering some of its inherent limitations and comparing its practical application with that of the Fourier method. In Chapter V both analyses were applied to the case of a practical multi-stage filter circuit. Results of that study revealed that for the Quasi-steady-state approach to be applicable, it is necessary for the modulating frequency to be low and for the network phase shift to be relatively linear. Although inclusion of additional correction terms improves the accuracy, the above restrictions still apply, relatively speaking. In addition, the mathematics involved for the higher-order correction terms soon gets prohibitive. Imagine obtaining the overall transfer function for the 10-section filter of Chapter V, finding the magnitude and phase of this function, and then obtaining first and second derivatives of the amplitude and phase as a function of frequency. Nevertheless, these are just some of the steps involved in applying the second-order correction term of the Quasi-steady-state method to the filter problem. The Fourier method, on the other hand, suffers from none of the above limitations, but instead, from only the enormous number of computations required during the course of the analysis. But this is precisely the

type of work for which the digital computer was designed and the kind it performs best. Both of the methods mentioned above are actually complimentary rather than competetive. To be more explicit, the Fourier method is applicable at high modulating frequencies (low modulation index) where the Quasi-steady-state method is invalid, and conversely, the Quasi-steady-state analysis is applicable at very low modulating frequencies where the Fourier method is impractical. It is the middle range of modulating frequencies that causes most of the problems. Most of the existing literature on this subject is concerned with extending the range of the Quasi-steady-state analysis to take in more of the middle range of modulating frequencies. The Fourier method given in this paper attempts to close the gap from the other end of the frequency range. It would be desirable if one could devise some rule which could be applied in general to an FM wave to determine which of the two methods of analysis would be more economical and practical to perform. The formulation of such a rule does not seem likely however, since, as has been already shown in Chapter V, the network itself exerts a strong influence on the accuracy of the Quasi-steady-state method. It seems that experience would offer the best guide in making a suitable choice between the two methods.

A method was also devised for obtaining the output wave of a network when the FM wave was modulated by a pair of sine waves. The Fourier approach was also the basis of this analysis and the remarks made concerning single-frequency modulation generally apply to the double-frequency case also. However, due to a much greater number of

computations arising from the additional sidebands, it is recommended that this method be avoided except in special cases. Since harmonic distortion and intermodulation distortion are in actuality measurements of the same phenomena (network nonlinearity), it would be much more economical to use only harmonic distortion. In fact, the results obtained in this paper show that "percent distortion" has little significance except in cases of relatively small distortion. The main accomplishment of this study is the ability to predict the output waveform for any distortion.

In Chapter VI an experimental technique employing analog computer simulation was developed in order to verify the predictions made by the digital computer. Advantages of this technique over more conventional arrangements are discussed. Although not a topic of direct concern to this paper, it was observed that circuit parameters such as inductors, capacitors, and resistors in addition to other circuit properties were easily altered for optimum performance, usually by the adjustment of a potentiometer. This is an important feature of the analog simulation technique and proves to be a valuable tool to the circuit or systems designer. The results obtained from the simulated system were so close to the predicted values as to leave no doubt to the validity of the digital computer analysis and the analog simulation technique. Indeed, the only discrepancies observed in the experimental data were easily explained.

In evaluating the calculations for the Fourier method it was necessary to have at one's disposal a means of evaluating Bessel

functions. A FORTRAN program is located in Appendix A which performs this task. The program is completely self-contained and can be used to compute Bessel functions for other types of problems. The details of the program are discussed in the appendix.

Appendix D contains the development of a method by which the minimum-phase characteristic of a network can be obtained when only the attenuation characteristic is known. Although the program appears here as one of the accessories which makes the analysis a little more versatile, it can also be used independently.

FORTRAN programs for a few typical network transfer functions are included in Appendix E. Of particular interest are the program for the filter network used in Chapter V and the program for the single-tuned circuit used in Chapter VI to experimentally verify theoretical predictions. Although all of the programs in this paper were compiled in the FORTRAN IV language on the IBM 7040/1401 system, minor changes would allow for processing using other versions of the FORTRAN language on different computers. However, it is not recommended that the longer running programs be solved on computers that are much slower than the one used in this study unless absolutely necessary.

It has been the chief goal of this investigation to develop and explore methods whereby the distortion in FM systems could be practically and conveniently predicted. The success or failure of this venture can be measured by the ability to obtain answers of a desired accuracy as quickly and efficiently as possible. The degree of flexibility of the analysis in making changes in the FM wave and/or the networks to which

the FM wave is applied is another factor to be considered. The results obtained from the detailed studies and experimental investigations indicate that this goal has been attained. It is hoped that the methods and results presented in this paper will prove to be valuable to those investigators presently engaged in research in this field.

ACKNOWLEDGEMENTS

The author would like to express his sincere appreciation to Dr. Harry K. Ebert, Jr. for his patience and indispensable guidance during the course of this investigation. Special thanks go to the members of the author's graduate committee and the many others who have made valuable contributions. Recognition is due to Mrs. Wanda C. Hall for undertaking the task of transforming the manuscript into its final form. The author finds it difficult to express the deep feeling of gratitude and pride that he has for his wife and family who, in very trying times, responded only with encouragement and understanding.

BIBLIOGRAPHY

1. Carson, J. R., and T. C. Fry, "Variable-frequency electric circuit theory," Bell Sys. Tech. J., vol. 16, Oct. 1937, pp. 513-540.

2. Van der Pol, B., "The fundamental principles of frequency modulation," J.IEE (London), pt. 3, vol. 93, May 1946, pp. 153-158.

3. Stumpers, F. L. H. M., "Distortion of frequency-modulated signals in electrical networks," Commun. News, vol. 9, April 1948, pp. 82-92.

4. Baghdady, E. J., "Theory of low-distortion reproduction of FM signals in linear systems," IRE Trans. on Circuit Theory, vol. CT-5, Sept. 1958, pp. 202-214.

5. Baghdady, E. J., "Lectures on Communication System Theory," McGraw-Hill, New York, 1961, ch. 19, pp. 439-455.

6. Rowe, H. E., "Distortion of angle-modulated waves by linear networks," IRE Trans. on Circuit Theory (Correspondence), vol. CT-9. Sept. 1962, pp. 286-290.

7. Hupert, J. J., "Transient response of narrow-band networks to angle-modulated signals," Proc. NEC, vol. 18, Oct. 1962, pp. 458-468.

8. Hess. D. T., "Transmission of FM signals through linear systems," Proc. NEC, vol. 18, Oct. 1962, pp. 469-476.

9. Weiner, D. D., and B. J. Leon, "The quasi-stationary response of linear systems to modulated waveforms," Proc. IEEE, vol. 53, June 1965, pp. 564-575.

10. Weiner, D. D. and B. J. Leon, "On the quasi-stationary response of linear time-invariant filters to arbitrary FM signals," IEEE Trans. on Circuit Theory (Correspondence), vol. CT-11, June 1964, pp. 308-309.

11. Frantz, W. J., "The transmission of a frequency-modulated wave through a network," Proc. IRE, vol. 34, March 1946, pp. 114p-125p.

12. Cherry, E. C. and R. S. Rivlin, "Nonlinear distortion, with particular reference to the theory of frequency-modulated waves," pt. II, Phil. Mag., vol. 33, April 1942, pp. 272-293.

13. Medhurst, R. G., "Harmonic distortion of FM waves by linear networks," Proc. IEE (London), paper 1650, vol. 101, pt. 3, May 1954, p. 171.

14. Assadourian, F., "Distortion of a frequency-modulated signal by small loss and phase variations," Proc. IEE, vol. 40, 1952, p. 172.

15. Fagot, J., and P. Magne, "Frequency Modulation Theory, " Pergamon Press, New York, 1961.

16. Black, H. S., "Modulation Theory," D. Van Nostrand, Princeton, New Jersey, 1953.

17. Panter, P. F., "Modulation, Noise, and Spectral Analysis," McGraw-Hill, New York, 1965.

18. Hund, A., "Frequency Modulation," McGraw-Hill, New York, 1942.

19. Terman, F. E., and J. M. Pettit, "Electronic Measurements," McGraw-Hill, 1952.

20. Berger, E. L., and R. M. Taylor, "Optimization of a radar in its environment by GEESE techniques," paper presented at the Joint IRE-AIEE-ACM Computer Conference, Los Angeles, May 1961.

21. Johnson, C. L., "Analog Computer Techniques," 2nd Ed., McGraw-Hill, New York, 1963, Ch. 9.

22. McCormick, J. M., and M. G. Salvadori, "Numerical Methods in FORTRAN," Prentice-Hall, Englewood Cliffs, New Jersey, 1964.

23. Jahnke, E., and F. Emde, "Tables of Functions," Dover Publications, New York, 1945, pp. 128, 138.

24. Bode, H. W., "Network Analysis and Feedback Amplifier Design," D. Van Nostrand, Princeton, New Jersey, 1945, ch. 14.

25. Tuttle, D. F., "Electric Networks: analysis and synthesis," McGraw-Hill, New York, 1965, ch. 11.

26."Tables of Bessel Functions," Harvard University Press, Cambridge, Massachusetts, 1947.

VITA

Preston Benton Johnson was born on March 7, 1932 in Benson, North Carolina. He was graduated from Benson High School, Benson, North Carolina in June, 1950. From March, 1951 until January, 1955, he served in the United States Navy where he attended various electronics schools. In February, 1955 he entered North Carolina State University where he was granted a Bachelor of Electrical Engineering degree in June, 1958. He was employed by the Martin Company in Orlando, Florida from June, 1958 until September, 1959 at which time he returned to North Carolina State University to accept a part-time teaching position and to continue his education. He received the Master of Science degree from North Carolina State University in January, 1962 and remained at that institution in the capacity of a full-time instructor until August, 1962. In September, 1962 he accepted an appointment as Assistant Professor of Electrical Engineering at Virginia Polytechnic Institute. In September, 1963 his teaching duties were reduced to one-half time in order to allow him time to pursue the Doctor of Philosophy degree. He participated in the Ford Foundation Forgivable Loan Program from September, 1963 until June, 1966.

In June, 1954 he married the former Miss June Dale McLamb of Benson, North Carolina. They presently reside in Blacksburg, Virginia with their three sons, Ronald, Mark, and Gregg.

Preston B. Johnson

APPENDIX A

Program for Computing Bessel Functions of the First Kind

The program appearing in this appendix is a modification of one published by McCormick and Salvadori [22] . The original program also computed ordinary Bessel functions of the second kind and modified Bessel functions of the first and second kinds. This feature was deleted in the present version since there was no need for these functions. However, the present program is able to compute Bessel functions for negative orders. The original program was found to be extremely inaccurate at some very critical points, especially for the higher orders. Correction of this problem was achieved by selective use of double-precision arithmetic. This point will be discussed in more detail below.

In the computer program the function BES(X,N) represents a Bessel function of the first kind of order N and with an argument X. The program makes use of both a power series expansion and an asymptotic expansion depending on the relationship between the order and the argument. More specifically, the power series is used for arguments less than the order plus six and the asymptotic expansion is used for larger arguments.

The actual expansions used are [23]

$\underline{x < n+6}$

$$J_n(x) = \frac{\left(\frac{x}{2}\right)^n}{0!\,n!} - \frac{\left(\frac{x}{2}\right)^{n+2}}{1!\,(n+1)!} + \frac{\left(\frac{x}{2}\right)^{n+4}}{2!\,(n+2)!} - \cdots \qquad \text{(A-1)}$$

$\underline{x \geq n+6}$

$$J_n(x) = \sqrt{\frac{2}{\pi x}}\left[P_n(x)\cos\phi - Q_n(x)\sin\phi\right] \qquad \text{(A-2)}$$

where $\phi = x - \left(\frac{n\pi}{2} + \frac{\pi}{4}\right)$,

$$P_n(x) = 1 - \frac{(4n^2-1)(4n^2-9)}{2!\,(8x)^2} + \frac{(4n2-1)(4n^2-9)(4n^2-25)(4n^2-49)}{4!\,(8x)^4} - \cdots,$$

and

$$Q_n(x) = \frac{4n^2-1}{8x} - \frac{(4n^2-1)(4n^2-9)(4n^2-25)}{3!\,(8x)^3} + \cdots .$$

The reason that the power series is not used for large arguments is because of gross inaccuracies which arise due to the limited number of significant figures carried by the computer. To illustrate this, consider first the case of the power series expansion for X = N = 1. Thus,

$$J_1(1) = \frac{(1/2)}{1} - \frac{(1/2)^3}{2!} + \frac{(1/2)^5}{2!\,3!} - \cdots .$$

The computer of course handles these computations using decimal fractions with only 8 significant figures. Therefore, in the computer the above expansion would be

$$J_1(1) = .50000000 - .06250000 + .00260417 - \cdots$$

and the results would be accurate to 8 figures after only a few terms.

On the other hand, if the computation is made for X = 20 and n = 1, the series used by the computer would be

$$J_1(10) = 10.000000 - 500.00000 + 8666.6667$$
$$- 69444.444 + 347222.22 - 11574074.$$
$$+ 275573190. - \ldots .$$

The desired answer for J_1 (10) is approximately 0.04347. From the series one observes that the terms are still growing. However, the series would eventually reach a point where the terms would begin decreasing since the series is convergent. The important point is the loss of accuracy that occurs due to the limited number of significant figures. Since the computer carries only 8 figures, any number with more than 8 significant figures will be rounded off, and any significance that the rounded off numbers might have had is lost. This occurs in the above series after only a few terms and the significance of the second decimal place, which is to contain the most significant figure in the answer, becomes meaningless.

The asymptotic expansion is the semi-convergent series of Hankel and is good when the magnitude of the argument is much less than the magnitude of the order. The switch from the power series to the asymptotic expansion occurs when X = n $\neq$ 6 as indicated above. The main featuresof the computer program will now be discussed.

When the function subprogram is initially called, the order is tested to determine if it is negative. The following relationship is used for evaluating the function when the order is negative:

$$J_{-N}(x) = (-1)^{N} J_N(x) . \qquad \text{(A-3)}$$

In the computer program, this is accomplished by setting the variable IND equal to zero for positive orders or equal to one for negative orders. The actual computations are always performed with positive values of N and the final answer is decided by statements 600 and 601. Next, the decision must be made to determine whether the computations should be performed in single-precision or double-precision arithmetic. This is accomplished by the five IF statements appearing at the beginning of the program. A study of the results obtained from a program which had been run completely in single precision revealed that for orders greater than 13, a significant error began to show up for values of arguments of $n \neq 5$. For example, when the order is 20, an argument of 25 gives an answer that is off more than one part in the third decimal place. For higher orders the error is much worse until the point is reached where the results have no meaning at all. It turns out that the maximum error occurs on the last power series expansion before the switch is made to the asymptotic expansion. In order to compensate for this error, double-precision arithmetic is used, but only to a limited extent. Table 2 summarizes the results of the series of IF statements.

In the power series expansion the calculations continue term by term until a point is reached where the last term added makes no difference in the previous answer. The result is then returned to the

main program. The number of terms used in the asymptotic expansion is

$$m = x + 1 + (x^2 + n^2)^{\frac{1}{2}}$$

provided that a test similar to that used in the power series does not terminate the series first.

Double Precision		Single Precision	
Order	Argument	Order	Argument
0 - 13	None	0 - 13	All
14 - 33	$N - 2 \leq X \leq N + 5$	14 - 33	$N - 2 > X > N + 5$
34 - 50	$X \geq N - 8$	34 - 50	$X < N - 8$

Table 2. Summary of IF statements showing the type of arithmetic used in the computer depending on argument and order of the Bessel function.

The results from use of this program have been checked for integer orders from 0-50 and for integer arguments from 0-50 using published tables of Bessel functions [26] and found to be accurate to at least 4 decimal places.

```
C
C       FUNCTION SUBPROGRAM FOR COMPUTING J-N(X).  POWER SERIES EXPANSION
C       USED FOR X LESS THAN N+6, LARGER ARGUMENTS USE ASYMPTOTIC EXPANSION.
C
        FUNCTION BES(X,N)
C
        DIMENSION AH(60),AL(60),ANGH(60),ANGL(60),FT(360),A(10),B(10)
        COMMON AH,AL,ANGH,ANGL,A0,ANG0
        EXTERNAL NOLMUN
        DOUBLE PRECISION DX,DXA,DXB,DAN,DT,DS,DBES,DDEN,DC,DD,DCONST,
       *1                 DPHI,DU,DPN,DQN,DAI,DPI
C
        IND=0
        IF(N) 105,106,106
  105   IND=1
        N=-N
  106   CONTINUE
C
        PI=3.1415927
        DPI=3.141592653589793
C
        FN=N
        IF(FN-13.5) 60,9999,93
   93   IF(FN-33.5) 95,9999,97
   95   IF(X-FN+2.) 60,78,88
   88   IF(X-FN-5.) 78,78,60
   97   IF(X-FN+8.) 60,78,78
C
C       SINGLE PRECISION COMPUTATION FOR J-N(X).
C
   60   IF(X-FN-6.) 1,200,200
C
C       POWER SERIES EXPANSION FOR J-N(X).  SEE P. 128 IN JAHNKE AND EMDE.
C
    1   XA=X/2.
```

```
      XB=XA*XA
    3 AN=N
      T=1.
      S=-1.
   10 IF(AN) 9999,20,12
   12 T=T*XA/AN
      AN=AN-1.
      GO TO 10
   20 BES=T
      DO 30 K=1,100000
      DEN=K*(K+N)
      T=T*XB/DEN
      IF((BES+T)-BES) 25,600,25
   25 BES=BES+T*S
   30 S=-S
C
C     COMPUTATION OF J-N(X) BY ASYMPTOTIC EXPANSION.  SEE P. 138 IN
C     JAHNKE AND EMDE.
C
  200 C=4*N*N
      D=8.*X
      CONST=SQRT(2./(PI*X))
      AN=N
      PHI=X-(2.*AN+1.)/4.*PI
      M=X+1.+SQRT(X*X+AN*AN)
      T=(C-1.)/D
      S=1.
      U=1.
      PN=1.
      QN=T
C
      DO 240 I=2,M
      AI=I
      T=(C-(2.*AI-1.)**2)/D*T/AI
      IF(S) 220,9999,210
```

```
  210 PN=PN-T*S*U
      U=-U
      GO TO 230
  220 QN=QN-T*S*U
      IF((QN+T)-QN) 230,241,230
  230 S=-S
  240 CONTINUE
  241 BES=CONST*(PN*COS(PHI)-QN*SIN(PHI))
      GO TO 600
C
C
C     DOUBLE PRECISION COMPUTATION FOR J-N(X).
C
   78 DX=X
      IF(DX-FN-6.) 301,400,400
  301 DXA=DX/2.
      DXB=DXA*DXA
  303 DAN=N
      DT=1.
      DS=-1.
  310 IF(DAN) 9999,320,312
  312 DT=DT*DXA/DAN
      DAN=DAN-1.
      GO TO 310
  320 DBES=DT
      DO 330 K=1,100000
      DDEN=K*(K+N)
      DT=DT*DXB/DDEN
      IF((DBES+DT)-DBES) 325,326,325
  325 DBES=DBES +DT*DS
  330 DS=-DS
  326 BES=DBES
      GO TO 600
C
C     ASYMPTOTIC EXPANSION
```

```
C
  400 DC=4*N*N
      DD=8.*DX
      DCONST=DSQRT(2./(DPI*DX))
      DAN=N
      DPHI=DX-(2.*DAN+1.)/4.*DPI
      M=DX+1.+DSQRT(DX*DX+DAN*DAN)
      DT=(DC-1.)/DD
      DS=1.
      DU=1.
      DPN=1.
      DQN=DT
C
      DO 440 I=2,M
      DAI=I
      DT=(DC-(2.*DAI-1.)**2)/DD*DT/DAI
      IF(DS) 420,9999,410
  410 DPN=DPN-DT*DS*DU
      DU=-DU
      GO TO 430
  420 DQN=DQN-DT*DS*DU
      IF((DQN+DT)-DQN) 430,441,430
  430 DS=-DS
  440 CONTINUE
  441 BES=DCONST*(DPN*DCOS(DPHI)-DQN*DSIN(DPHI))
  600 IF(IND) 9999,1111,601
  601 BES=(-1.)**N*BES
      N=-N
 1111 RETURN
C
 9999 STOP
C
      END
```

APPENDIX B

Distortion Analysis for Modulation By Two Sinusoids

A FORTRAN program for determing the response of a network when the exciting function is an FM wave modulated by two sinusoids is given on pages 134 and 135. The philosophy of this program is the same as that of the program used to determine the response when the FM wave is modulated by a single frequency sinusoid. Before studying the details of the program, a discussion of the theory presented in Chapter III is in order. The normalized FM wave at the input of the network is given by Equation (III-42) and repeated here for convenience:

$$F(t) \triangleq \sum_{N=-\infty}^{\infty} \sum_{K=-\infty}^{\infty} J_N(m_1) J_K(m_2) \sin (\omega_c + N\omega_1 + K\omega_2)t. \qquad (B\text{-}1)$$

The spectral components represented by this equation are shown in Chapter III for a specific case. The distortion analysis is concerned with evaluating the spectrum after the wave is passed through the linear network. Again the network transfer function can be represented by

$$H(\omega) = A(\omega) e^{j\theta(\omega)} . \qquad (B\text{-}2)$$

It is necessary to evaluate this expression at the sideband frequencies which are functions of the carrier and the two modulating frequencies. As a matter of convenience, the following definitions will be made:

$$A(\omega_c + N\omega_1 + K\omega_2) = A_{N,K}$$

and $\theta(\omega_c+N\omega_1+K\omega_2) = \theta_{N,K}$.

The response of the network to the input wave is

$$G(t) = \sum_{N=-\infty}^{\infty} \sum_{K=-\infty}^{\infty} A_{N,K}\, J_N(m_1)\, J_K(m_2) \sin (\omega_c t + N\omega_1 t + K\omega_2 t + \theta_{N,K}) . \quad \text{(B-3)}$$

This can be expanded by trigonometric identities to yield

$$G(t) = \sum_{N=-\infty}^{\infty} \sum_{K=-\infty}^{\infty} A_{N,K}\, J_N(m_1) J_K(m_2) \left[\sin \omega_c t \cos (N\omega_1 t+K\omega_2 t+\theta_{N,K}) + \cos \omega_c t \sin (N\omega_1 t + K\omega_2 t + \theta_{N,K})\right] . \quad \text{(B-4)}$$

This equation is of the form

$$G(t) = R \sin \omega_c t + S \cos \omega_c t \quad \text{(B-5)}$$

which is the same as Equation (III-19). R and S are given by

$$R = \sum_{N=-\infty}^{\infty} \sum_{K=-\infty}^{\infty} A_{N,K}\, J_N(m_1) J_K(m_2) \cos (N\omega_1 t+K\omega_c t+\theta_{N,K}) \quad \text{(B-6)}$$

and

$$S = \sum_{N=-\infty}^{\infty} \sum_{K=-\infty}^{\infty} A_{N,K}\, J_N(m_1) J_K(m_2) \sin (N\omega_1 t+K\omega_2 t+\theta_{N,K}) . \quad \text{(B-7)}$$

As shown in Chapter III the instantaneous frequency of the FM output wave is

$$\omega_i = \omega_c + \frac{R(dS/dt) - S(dR/dt)}{R^2 + S^2} \quad \text{(B-8)}$$

where

$$\frac{dR}{dt} = -\sum_{N=-\infty}^{\infty} \sum_{K=-\infty}^{\infty} A_{N,K}(N\omega_1 + K\omega_2) J_N(m_1) J_K(m_2)$$

$$\cdot \sin (N\omega_1 t + K\omega_2 t + \theta_{N,K}) \qquad \text{(B-9)}$$

and

$$\frac{dS}{dt} = \sum_{N=-\infty}^{\infty} \sum_{K=-\infty}^{\infty} A_{N,K}(N\omega_1 + K\omega_2) J_N(m_1) J_K(m_2)$$

$$\cdot \cos(N\omega_1 t + K\omega_2 t + \theta_{N,K}). \qquad \text{(B-10)}$$

The amplitude of the FM output wave is obtained by

$$|G(t)| = \sqrt{R^2 + S^2} \quad . \qquad \text{(B-11)}$$

The demodulated output wave is now obtained by removing the carrier frequency from the FM wave. Thus,

$$e_o = K \left[\frac{R(dS/dt) - S(dR/dt)}{R^2 + S^2} \right] \qquad \text{(B-12)}$$

where K represents demodulation circuit parameters and will be considered to be unity for convenience. The FORTRAN program for evaluating Equations (B-11) and (B-12) will now be discussed.

The input data are fed into the main program which with the aid of the two subprograms makes the calculations necessary to solve the problem. The transfer function subprogram for the double-frequency modulation case differs from that of the single-frequency case in that it is actually composed of two subprograms - one for calculating the attenuation and the other for calculating the phase. Examples of these subprograms are found on pages 136 and 137 following the main program.

The FORTRAN program for the Bessel function subprogram is described in Appendix A.

The execution of the main program is quite similar to that of the single-frequency modulation program. It begins by calling in data which describe the FM input wave. The computer variables WC, W1, W2, DEV1, DEV2, and DWTD represent, respectively: the carrier frequency, the frequency of the first modulating wave, the frequency of the second modulating wave, the frequency deviation of the first modulating wave, the frequency deviation of the second modulating wave, and the incremental change in the angle of the first modulating wave at which successive calculations are to be made. All of the above values are in radians/second except DWTD which is in degrees. It is important that the lower modulating frequency be identified as W1 since it is for this wave that the calculations are made over a full cycle. Failure to comply with this requirement will result in incomplete output data. The program next computes the values of modulation index, Z1 and Z2 for the two waves. These values are then used to determine the maximum number of sidebands to be used during the course of the calculations as discussed in Chapter III. The remainder of the program is essentially the same as that for the single-frequency modulation case with the exception that the arithmetic statements used to evaluate R, S, DR, and DS appear in a different form. In the present case, Equations (B-6), (B-7), (B-9), and (B-10) were programmed directly on the computer. Each of these equations requires a double summation over intervals that include negative indices. The summations are easily implemented in the program by the two "DO 5" statements. However, FORTRAN IV does not

accept negative values for the index of a "DO" statement. This restriction was circumvented by the introduction of the variables NS, NP, and KP. The net effect is that the summations are made over the intervals from - NMAX to NMAX. These equations also make use of Bessel functions of negative orders as discussed in Appendix A.

Figure 38 shows the result of simultaneously applying two sinusoids with frequencies of 50 and 500 hertz to a sharply tuned resonant circuit. This is a case of extreme distortion and indicates the futility of trying to compute the intermodulation distortion by the computer program. In fact, the intermodulation effect cannot even be detected in the figure. Thus, one concludes that intermodulation measurements only have significance when the distortion is reasonably small.

As previously indicated in Chapter III, the execution of the program given in this appendix is generally uneconomical. The time required for the computations involved in obtaining data for the plot in Figure 38 was 22 minutes on the IBM 7040/1401 system. Several other times the program ran for more than 30 minutes. The long running times are due to several factors. First, the number of sidebands which is used in each calculation is greatly increased over that for the case of the single modulating frequency. Secondly, the output data must be obtained for smaller angular increments of the low frequency wave to assure that adequate information on the high frequency wave is obtained. For the curve shown in Figure 38, a calculation was made at every 3^{o} of angle of the low frequency input

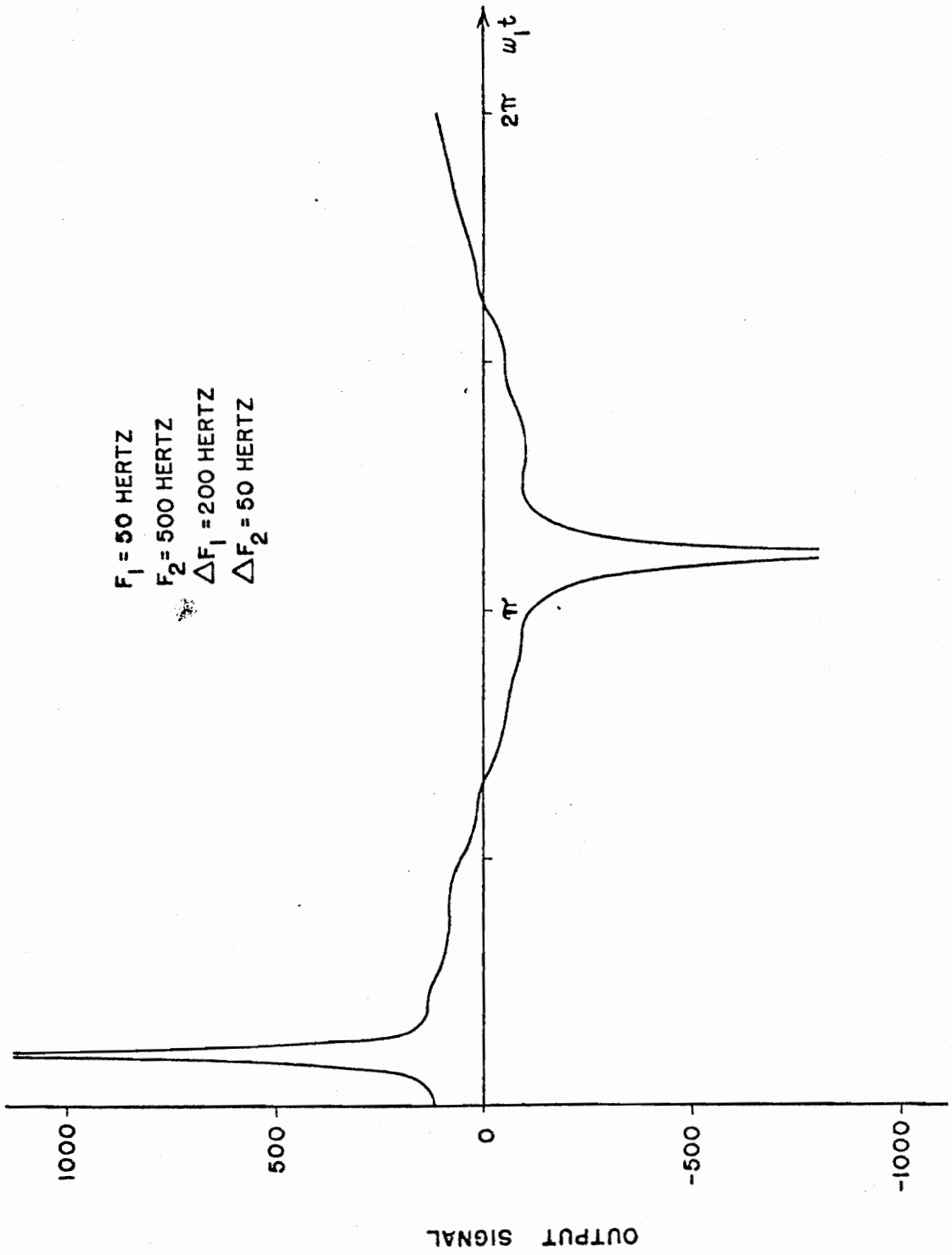


FIGURE 38-OUTPUT WAVE FOR MODULATION BY TWO SINUSOIDS

modulating wave. Since the ratio of the two modulating frequencies was 10:1, the calculations were only made at every 30^{o} of angle of the high frequency wave. Finally, another probable cause for the long running time was due to underflows occurring in the program. Normally, one is allowed 10 underflows (or overflows) in the execution of a program before the program is terminated. Each time an underflow is encountered, the computer generates an error message and then returns to the execution of the program. The statement "NOLMUN" protects the program from being terminated but does not prevent the computer from generating 30 or 40 extra instructions.

Based on the above results it is recommended that distortion analyses be made using a single modulating frequency when possible. However, it should be pointed out that the method presented in this appendix can theoretically be extended to accommodate any number of modulating frequencies. This would allow one "construct" waveforms of any type for purposes of modulation. At this time, however, the practical and economic aspects of such a venture are prohibitive.

```
C
CC      FM DISTORTION ANALYSIS USING DOUBLE-FREQUENCY MODULATION.
C
   33 READ(5,1) WC,W1,W2,DEV1,DEV2,DWTD
    1 FORMAT(6F10.0)
      Z1=DEV1/W1
      Z2=DEV2/W2
      WRITE(6,70) WC,W1,W2,DEV1
   70 FORMAT(4H1WC=F12.4,6H   W1=F12.4,6H   W2=F12.4,8H   DEV1=F12.4)
       WRITE(6,7000) DEV2,Z1,Z2
 7000 FORMAT(8H   DEV2=F12.4,6H   Z1=F12.4,6H   Z2=F12.4//)
      NMAX=Z1+Z2+5.
      WRITE(6,13) NMAX
   13 FORMAT(6H1NMAX=I3//)
      PI=3.1415927
      DWTR=DWTD*PI/180.
      I=0
      MX=360.05/DWTD
      WT1=0.
      DO 62 J=1,MX
      WT2=WT1*W2/W1
      R=0.
      S=0
      DR=0
      DS=0
      NS=2*NMAX+1
      DO 5 N=1,NS
      NP=N-NMAX-1
      AN=NP
      B1=BES(Z1,NP)
      DO 5 K=1,NS
      KP=K-NMAX-1
      AK=KP
      B2=BES(Z2,KP)
      R=R+A(WC+AN*W1+AK*W2)*B1*B2*COS(AN*WT1+AK*WT2+ANG(WC+AN*W1+AK*W2))
```

```
      S=S+A(WC+AN*W1+AK*W2)*B1*B2*SIN(AN*WT1+AK*WT2+ANG(WC+AN*W1+AK*W2))
      DR=DR-A(WC+AN*W1+AK*W2)*(AN*W1+AK*W2)*B1*B2*SIN(AN*WT1+AK*WT2
     1+ANG(WC+AN*W1+AK*W2))
      DS=DS+A(WC+AN*W1+AK*W2)*(AN*W1+AK*W2)*B1*B2*COS(AN*WT1+AK*WT2
     1+ANG(WC+AN*W1+AK*W2))
    5 CONTINUE
      PHI=ATAN2(S,R)
      Q=R*R+S*S
      RMAG=SQRT(Q)
      DWI=(R*DS-S*DR)/Q
      WTD=WT1*180./PI
      I=I+1
      IF(I-1) 20,20,30
   20 WRITE(6,21)
   21 FORMAT(14H  WT1(DEGREES),5X,9HMAGNITUDE,8X,12HPHI(RADIANS),5X,
     1       16HDWI(RADIANS/SEC)/)
   30 WRITE(6,31)WTD,RMAG,PHI,DWI
   31 FORMAT(1X,4(1PE15.8,2X))
   62 WT1=WT1+DWTR
      GO TO 33
   60 STOP
      END
```

```
C
C       MAGNITUDE OF TRANSFER FUNCTION FOR SINGLE-TUNED CIRCUIT.
C       RESONANT FREQUENCY IS 10000 RADIANS/SECOND AND Q IS 100.
C
        FUNCTION A(W)
        A=50.*W/SQRT((1.E8-W*W)**2 + W*W*1.E4)
        RETURN
        END
```

```
C
C       PHASE OF TRANSFER FUNCTION FOR SINGLE-TUNED CIRCUIT.
C       RESONANT FREQUENCY IS 10000 RADIANS/SECOND AND Q IS 100.
C
        FUNCTION ANG(W)
        ANG=ATAN2((1.E8-W*W),W*100.)
        RETURN
        END
```

APPENDIX C

FORTRAN Programs for Quasi-steady-state Analysis

Three FORTRAN programs, consisting of a main program and two subprograms, are used in the Quasi-steady-state analysis. These programs are located on pages 142 through 145. As stated in Chapter VI, the problem is simply to evaluate the equation.

$$e_{sig} \alpha \Delta\omega \cos \omega_m t - \left(\frac{d\theta}{d\omega_i}\right) \cdot \omega_m \Delta\omega \sin \omega_m t. \qquad (C\text{-}1)$$

A study of this equation reveals that every quantity on the right-hand side is known (except $d\theta/d\omega_i$). Thus, the problem actually boils down to finding the slope of the phase-versus-frequency curve. The subprogram DER(W) is used to make this computation with the aid of the subprogram named SLOPE(L). The method of numerical differentiation is used in evaluating the derivative. Figure 39 will be used to illustrate the method used in calculating $d\theta/d\omega_i$. Suppose that it is desired to find the value of $d\theta/d\omega_i$ at $\omega_i = \omega_p$ given that the other points on the curve are known. The method employed to approximate this value was to use

$$\frac{d\theta}{d\omega_i} \approx \frac{\Delta\theta}{\Delta\omega_i} \Bigg|_{\omega_i = \omega_p} . \qquad (C\text{-}2)$$

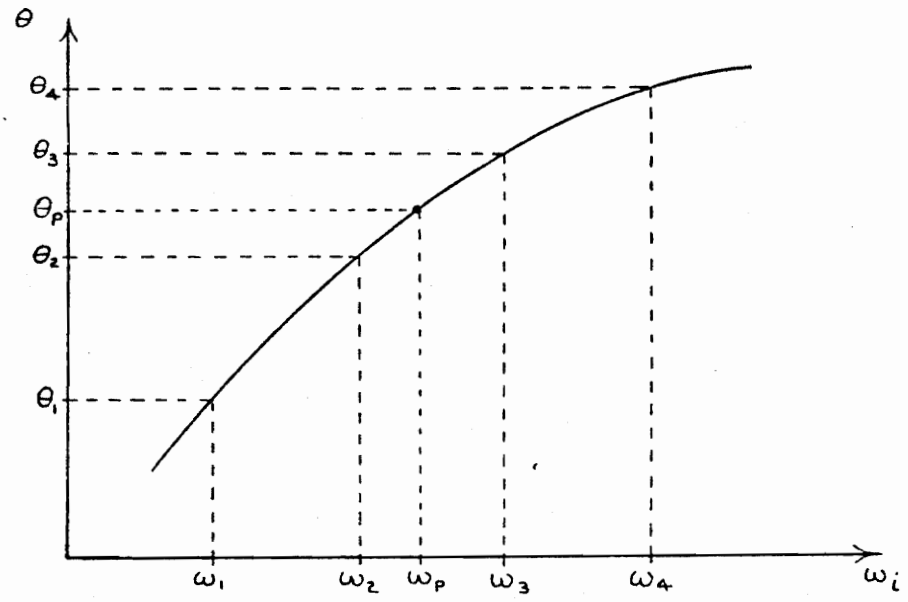


Figure 39. Typical curve of phase shift as a function of frequency for a linear network.

The simplest approach to solving Equation (C-2) is to assume that $\Delta\theta = \theta_3 - \theta_2$ and $\Delta\omega = \omega_3 - \omega_2$. This means that the slope at any frequency between ω_2 and ω_3 is considered to be the same. Depending on the shape of the curve in this range and the closeness of the points of known θ, the accuracy can vary considerably. There must be at least one frequency at which the approximation is exact but it may be in considerable error at other points even for a smooth curve. Frequencies close to the end points (ω_2 and ω_3) would usually show more error than those close to the center of the region if one assumes a monotonically increasing or decreasing function. This leads to a method of improving the accuracy of the approximation especially if equally spaced frequencies are used for the points of known θ. The

slope of the curve at ω_2 can be obtained to a fairly good accuracy since ω_2 is equally spaced between ω_1 and ω_3. The same is true for obtaining the slope at ω_3 since ω_3 lies exactly between ω_2 and ω_4. Thus, it can be assumed that the slopes of the curve are known at ω_2 and ω_3. It is now a simple matter to use linear interpolation to find the slope at any point in between ω_2 and ω_3. This method of improving the accuracy actually is equivalent to assuming a constant second derivative as opposed to the constant first derivative assumption of the previous method. Of course, further improvement could be obtained by working with higher derivatives but the second derivative method was sufficient for the present application.

In analyzing the main program, it is observed that the input data representing the variation of phase with frequency are entered in the form of two one-dimensional arrays called WL and THETA. The maximum size of each of these arrays is limited to 99 values by the DIMENSION statement and could be increased if desired. The arrays are actually used in the subprograms and are made available to them by means of the COMMON statement. The quantity N determines exactly how many values of WL and THETA are to be entered. The fixed point variable M determines the number of times the input arrays will be used. After the input data are all entered, they are printed out again for reference purposes. The program then calls for the values of the carrier frequency WC, the modulating frequency deviation DEV, and the angular increment DWTD.

After the input wave characteristics are entered, the program

proceeds to evaluate Equation (C-1). The values of the output wave, the input wave, and the angle at which they were calculated are then printed out. At this point the control returns to one of two places depending on the value of the fixed-point variable M which was entered initially with N and the arrays. The purpose of specifying M is to have the increased flexibility of being able to change networks as well as input waveforms. It is not necessary to specify M in which case the program will never call for other network data. If M is specified, read statement 33 will be executed M times and then the control will be automatically returned to statement 41 which in turn will call for new network data. Sample output data are found on page 146.

The execution time of this program is reasonably short, being in the order of three minutes for a single run taken at every three degrees of the input modulating wave. However, there are other disadvantages to using this method as discussed in Chapters IV and V.

```
C
C       FM DISTORTION ANALYSIS BY QUASI-STEADY-STATE METHOD   4/5/66  PBJ
C
        DIMENSION WL(99),THETA(99)
        COMMON WL,THETA,N
C
   41 READ(5,1) N,M,(WL(I),I=1,N),(THETA(I),I=1,N)
    1 FORMAT(2I2/8(F10.0))
C
C       N IS THE MAXIMUM NUMBER OF POINTS ON THE PHASE VERSUS FREQUENCY
C       CURVE. M IS THE NUMBER OF TIMES NEW DATA IS READ IN BY READ
C       STATEMENT NO. 33.  AFTER TABLE IS USED FOR M CALCULATIONS, A NEW
C       SET OF TABLE VALUES IS READ IN.  IF M IS NOT SPECIFIED OR IS SET
C       EQUAL TO ZERO, CONTROL WILL BE CONTINUOUSLY RETURNED TO STATEMENT
C       NO. 33 AND NO ADDITIONAL TABLE ENTRIES ARE POSSIBLE.
C
        WRITE(6,2) N,M
    2 FORMAT(3H1N=I2,5X,2HM=I2//)
C
        DO 3 I=1,N
    3 WRITE(6,4) I,WL(I),I,THETA(I)
    4 FORMAT(4H WL(I2,2H)=F10.0   ,5X,6HTHETA(I2,2H)=F11.7)
C
   33 READ(5,5) WC,WM,DEV,DWTD
    5 FORMAT(4F10.0)
C
        Z=DEV/WM
C
C       WC IS ANG. VEL. OF CARRIER, WM IS ANG. VEL. OF MOD. WAVE, Z IS
C       MOD. INDEX, DWTD IS INCREMENT OF ANG. DISPL. OF MOD. WAVE,
C       AND DEV IS THE MAXIMUM DEVIATION IN RADIANS/SECOND.
C
        WRITE(6,70) WC,WM,DEV,Z
   70 FORMAT(4H1WC=F10.0,5H  WM=F10.0,8H    DEV=F10.0,5H   Z=F6.3//)
C
```

```
      PI=3.1415927
      DWTR=DWTD*PI/180.
C
      WRITE(6,92)
   92 FORMAT(1H1)
C
      K=0
      MX=360.05/DWTD
      WTR=0
C
      DO 62 J=1,MX
      W=WC+DEV*COS(WTR)
      OUT=DEV*(COS(WTR)-WM*DER(W)*SIN(WTR))
      CK=W-WC
      WMT=WTR*180./PI
C
      K=K+1
      IF(K-1) 20,20,30
C
   20 WRITE(6,21)
   21 FORMAT(13H WMT(DEGREES),6X,13HOUTPUT SIGNAL,6X,12HINPUT SIGNAL/)
C
   30 WRITE(6,31) WMT,OUT,CK
   31 FORMAT(1X,3(1PE15.8,2X))
C
   62 WTR=WTR+DWTR
      M=M-1
C
      IF(M) 33,41,33
C
   60 STOP
      END
```

```
C
      FUNCTION DER(W)
C
      DIMENSION WL(99),THETA(99)
      COMMON WL,THETA,N,J
C
      J=0
C
    1 J=J+1
C
      IF(J-N) 4,4,6
C
    6 WRITE(6,16)
   16 FORMAT(38H FREQUENCY OUT OF RANGE OF INPUT DATA.)
      DER=99999999.
      RETURN
C
    4 IF(W-WL(J)) 3,2,1
C
    2 DER=SLOPE(J)
      RETURN
C
    3 IF(J-1) 60,6,10
   10 DER=SLOPE(J-1)+(SLOPE(J)-SLOPE(J-1))*(W-WL(J-1))/(WL(J)-WL(J-1))
      RETURN
   60 STOP
      END
```

```
C
      FUNCTION SLOPE(L)
C
      DIMENSION WL(99),THETA(99)
      COMMON WL,THETA,N,J
C
      IF(J-1) 60,5,6
    6 IF(J-N) 4,7,60
C
    4 SLOPE=(THETA(J+1)-THETA(J-1))/(WL(J+1)-WL(J-1))
      RETURN
C
C     LOWER END POINT
    5 SLOPE=(THETA(J+1)-THETA(J))/(WL(J+1)-WL(J))
      RETURN
C
C     UPPER END POINT
    7 SLOPE=(THETA(J)-THETA(J-1))/(WL(J)-WL(J-1))
      RETURN
C
   60 STOP
      END
```

N=30 M= 1

WL(1)=	99700.	THETA(1)=	0.1095444
WL(2)=	99900.	THETA(2)=	-0.0205399
WL(3)=	100000.	THETA(3)=	-0.0892085
WL(4)=	100200.	THETA(4)=	-0.2318195
WL(5)=	100400.	THETA(5)=	-0.3781413
WL(6)=	100600.	THETA(6)=	-0.5237910
WL(7)=	100800.	THETA(7)=	-0.6647824
WL(8)=	101000.	THETA(8)=	-0.7982374
WL(9)=	101200.	THETA(9)=	-0.9226716
WL(10)=	101400.	THETA(10)=	-1.0378648
WL(11)=	101600.	THETA(11)=	-1.1445171
WL(12)=	101800.	THETA(12)=	-1.2438938
WL(13)=	102000.	THETA(13)=	-1.3375430
WL(14)=	102200.	THETA(14)=	-1.4271316
WL(15)=	102400.	THETA(15)=	-1.5143550
WL(16)=	102600.	THETA(16)=	-1.6009104
WL(17)=	102800.	THETA(17)=	-1.6884979
WL(18)=	103000.	THETA(18)=	-1.7788287
WL(19)=	103200.	THETA(19)=	-1.8736258
WL(20)=	103400.	THETA(20)=	-1.9745836
WL(21)=	103600.	THETA(21)=	-2.0832724
WL(22)=	103800.	THETA(22)=	-2.2009405
WL(23)=	104000.	THETA(23)=	-2.3282076
WL(24)=	104200.	THETA(24)=	-2.4646715
WL(25)=	104400.	THETA(25)=	-2.6085587
WL(26)=	104600.	THETA(26)=	-2.7566414
WL(27)=	104800.	THETA(27)=	-2.9046184
WL(28)=	105000.	THETA(28)=	-3.0479459
WL(29)=	105100.	THETA(29)=	-3.1166329
WL(30)=	105300.	THETA(30)=	3.0370520

APPENDIX D

Calculation of the Phase Shift of a Network from the Attenuation.[1]

Frequently in practical situations, the attenuation (or gain) characteristic of a network will be known while the phase characteristic is not. This may result due to difficulty in measuring phase although the attenuation is easy to obtain; or design criteria may specify the attenuation while ignoring the phase specifications. Sometimes the attenuation may be specified by a set of numerical data or by a curve. In any event, there are times when it would be convenient to have a method by which the phase could be obtained from the attenuation in a relatively easy manner. The digital program included in this appendix enables one to perform this operation. The input data to the program is in the form of specific values of attenuation and the frequency corresponding to each value and the output consists of the frequency, the magnitude of the gain (corresponding to the attenuation), and the phase of the network transfer function.

In order to understand the philosophy of the calculations, certain integral relationships first developed by Bode will be derived. These relationships will then be linearized for numerical solution on

---

[1]This method of calculating the phase of a network from its attenuation characteristic was conceived by Bode [24] but the approach used here is basically that used by Tuttle [25].

the digital computer. Consider the function H(s) that is analytic (no singularities) within and on a closed contour, C. Then, according to a fundamental theorem due to Cauchy

$$\oint_C H(s)\, ds = 0 \qquad \text{(D-1)}$$

where $\oint_C$ indicates the integral around the closed contour, C. Now let the closed contour be the border of the right half plane as shown in Figure 40. Notice that the path of integration has avoided the singularity on the jw axis by taking a small semicircular path around it. Since the function is observed to have a pole on the jw axis, it can be written as

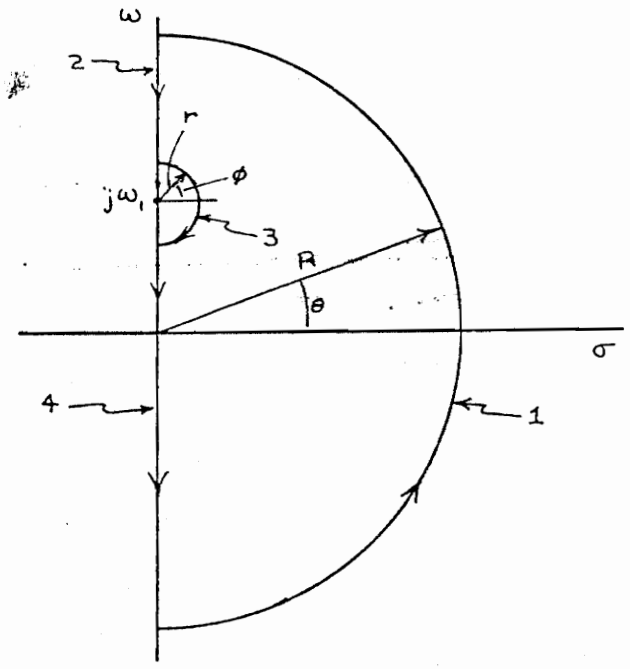


Figure 40. Path of Integration for evaluating the line integral around the closed contour, C.

$$H(s) = \frac{F(s)}{S - j\omega_1} \quad . \tag{D-2}$$

Furthermore, this function can be integrated around C to yield zero since the closed path has carefully avoided the point $S = jw_1$. Thus,

$$\oint_c \frac{F(s)}{S - j\omega_1} \, ds = 0 \quad . \tag{D-3}$$

The complete integral can now be evaluated by breaking the path up into the four sections shown in Figure 40:

$$\oint_c = \int_1 + \int_2 + \int_3 + \int_4 = 0 \quad . \tag{D-4}$$

These integrals will now be evaluated separately. The first integral on the right is to be taken over the semicircle specified by $s = Re^{j\theta}$ from the limits $-\pi/2$ to $+\pi/2$. Since it is desired to include the complete right half plane, the integral will be evaluated as $R \to \infty$. By use of the Cauchy Integral Formula the value of this integral is

$$\lim_{R \to \infty} \int_{-\pi/2}^{\pi/2} \frac{F(Re^{j\theta})}{Re^{j\theta} - j\omega_1} \, Re^{j\theta} \, jd\theta = j\pi F(\infty). \tag{D-5}$$

Integral 3 is evaluated in the same manner except that r is allowed to shrink to zero and the equation of the small semicircle is $s = jw_1 + re^{j\phi}$. Therefore, the value of this integral is

$$\lim_{r \to 0} \int_{\pi/2}^{-\pi/2} \frac{F(j\omega_1 + re^{j\phi})}{re^{j\phi}} \, re^{j\phi} \, jd\phi = -j\pi F(j\omega_1) \, . \tag{D-6}$$

The combination of integrals 2 and 4 represent integration along the entire jw axis except for the singular point at $s = jw_1$. These integrals will be written

$$\lim_{r \to 0} \left\{ \int_{\infty}^{\omega_1 + r} \frac{F(j\omega)}{\omega - \omega_1} \, d\omega + \int_{\omega_1 - r}^{-\infty} \frac{F(j\omega)}{\omega - \omega_1} \, d\omega \right\} = \int_{\infty}^{-\infty} \frac{F(j\omega)}{\omega - \omega_1} \, d\omega \qquad \text{(D-7)}$$

where the asterisk indicates that the integral must be evaluated by the limiting process indicated. Equation (D-4) can now be written

$$j\pi F(\infty) - j\pi F(j\omega_1) + \int_{\infty}^{-\infty} \frac{F(j\omega)}{\omega - \omega_1} \, d\omega = 0 \qquad \text{(D-8)}$$

or

$$F(j\omega_1) = F(\infty) + \frac{j}{\pi} \int_{\infty}^{-\infty} \frac{F(j\omega)}{\omega - \omega_1} \, d\omega \; . \qquad \text{(D-9)}$$

Now let $F(jw) = x(w) + jy(w)$ so that

$$x(\omega_1) + jy(\omega_1) = x(\infty) + jy(\infty) + \frac{j}{\pi} \int_{\infty}^{-\infty} \frac{x(\omega)}{\omega - \omega_1} \, d\omega - \frac{1}{\pi} \int_{\infty}^{-\infty} \frac{y(\omega)}{\omega - \omega_1} \, d\omega. \qquad \text{(D-10)}$$

Equating real and imaginary parts yields

$$x(\omega_1) = x(\infty) - \frac{1}{\pi} \int_{\infty}^{-\infty} \frac{y(\omega)}{\omega - \omega_1} \, d\omega \qquad \text{(D-11)}$$

and

$$y(\omega_1) = y(\infty) + \frac{1}{\pi} \int_{\infty}^{-\infty} \frac{x(\omega)}{\omega - \omega_1} d\omega \quad . \tag{D-12}$$

Various manipulations are used to reduce these equations to a more usable form[1]:

$$x(\omega_1) = x(\infty) + \frac{1}{\pi\omega_1} \int_0^{\infty} \frac{d(\omega y)}{d\omega} \ln \left| \frac{\omega - \omega_1}{\omega + \omega_1} \right| d\omega \; ; \tag{D-13}$$

$$y(\omega_1) = - \frac{1}{\pi} \int_0^{\infty} \frac{dx}{d\omega} \ln \left| \frac{\omega - \omega_1}{\omega + \omega_1} \right| d\omega \; . \tag{D-14}$$

In Equation (D-14), $y(\infty)$ has been taken to be zero. This will yield a so-called "minimum function".

Until now no restriction has been placed on F(s) as to the type of network function. The formulas are equally applicable whether F(s) is taken to be an impedance function, an admittance function, or a transfer function. However, the purpose of this study is to obtain a method of obtaining the phase of a network transfer function when the attenuation is given. In this case, it will be convenient to let F(s) be

$$F(s) = \alpha + j\beta \tag{D-15}$$

where α is the attenuation constant and β is the phase constant. Also

---

[1]For details see Tuttle, loc.cit.

since α is known, it is only necessary to write Equation (D-14) in terms of these quantities. Thus,

$$\beta(\omega_1) = -\frac{1}{\pi}\int_0^\infty \frac{d\alpha}{d\omega} \ln\left|\frac{\omega - \omega_1}{\omega + \omega_1}\right| d\omega . \qquad \text{(D-16)}$$

This equation is still rather difficult to use in its present algebraic form. However, a method using linear approximations allows one to get an approximate value for β . Note that if $d\alpha/d\omega$ is a constant, the integral simplifies considerably. In the case of the attenuation function, the behavior is such that the function becomes asymptotic to a straight line at high and low frequencies if one uses a logarithmic frequency scale. Furthermore, the function will be well behaved as $w \rightarrow 0$ and $w \rightarrow \infty$. Figure 41 shows how a typical attenuation function can be broken up into a series of linear segments.

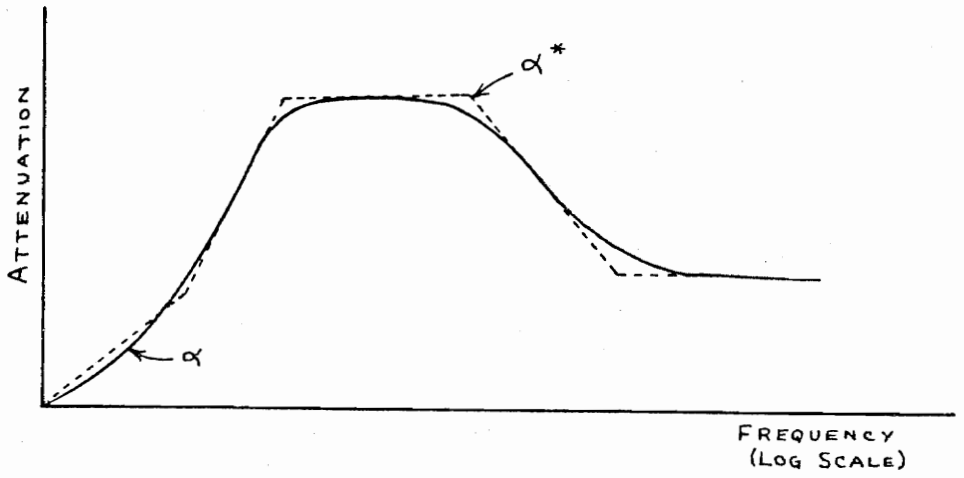


Figure 41. Typical curve of α and its linear approximation, α^* as a function of frequency.

The phase shift, β can now be approximated by Equation (D-16).

Rather than try to apply Equation (D-16) directly to the curve α^*, it will be much more convenient to use the method of representing the curve as a sum of "semi-infinite segments". Figure 42 indicates how this method works. Figure 42(a) shows the linear approximation of the attenuation curve of Figure 41 while Figure 42(b) shows how this approximate curve can be broken up into the sum of a series of ramp functions (called "semi-infinite segments"). Note how the slopes of the ramp functions behave in going from one linear segment to the next. First the slope of the previous ramp must be compensated for and then the actual slope of the segment as shown in Figure 42(a) is then added. The phase can now be calculated for the individual ramps and added to obtain the total phase at a particular frequency. The fact that the slopes are interrelated illustrates the network property whereby the attenuation characteristic over the whole frequency band influences the phase shift at any particular frequency. For example, the shape of the attenuation curve at 10 kilohertz will have some influence on the phase shift at 100 hertz however small it might be. One of the chief advantages of the ramp function approach is that the segmented curve is now represented by a series of functions of one basic type. The only difference between individual functions is the slope which is assumed to be constant. Hence, it is sufficient to calculate a representative phase shift for a ramp of a given slope. This will be done for a slope of unity and then the other phase shifts (due to other ramps) can be obtained by simply multiplying the representative phase shift by the

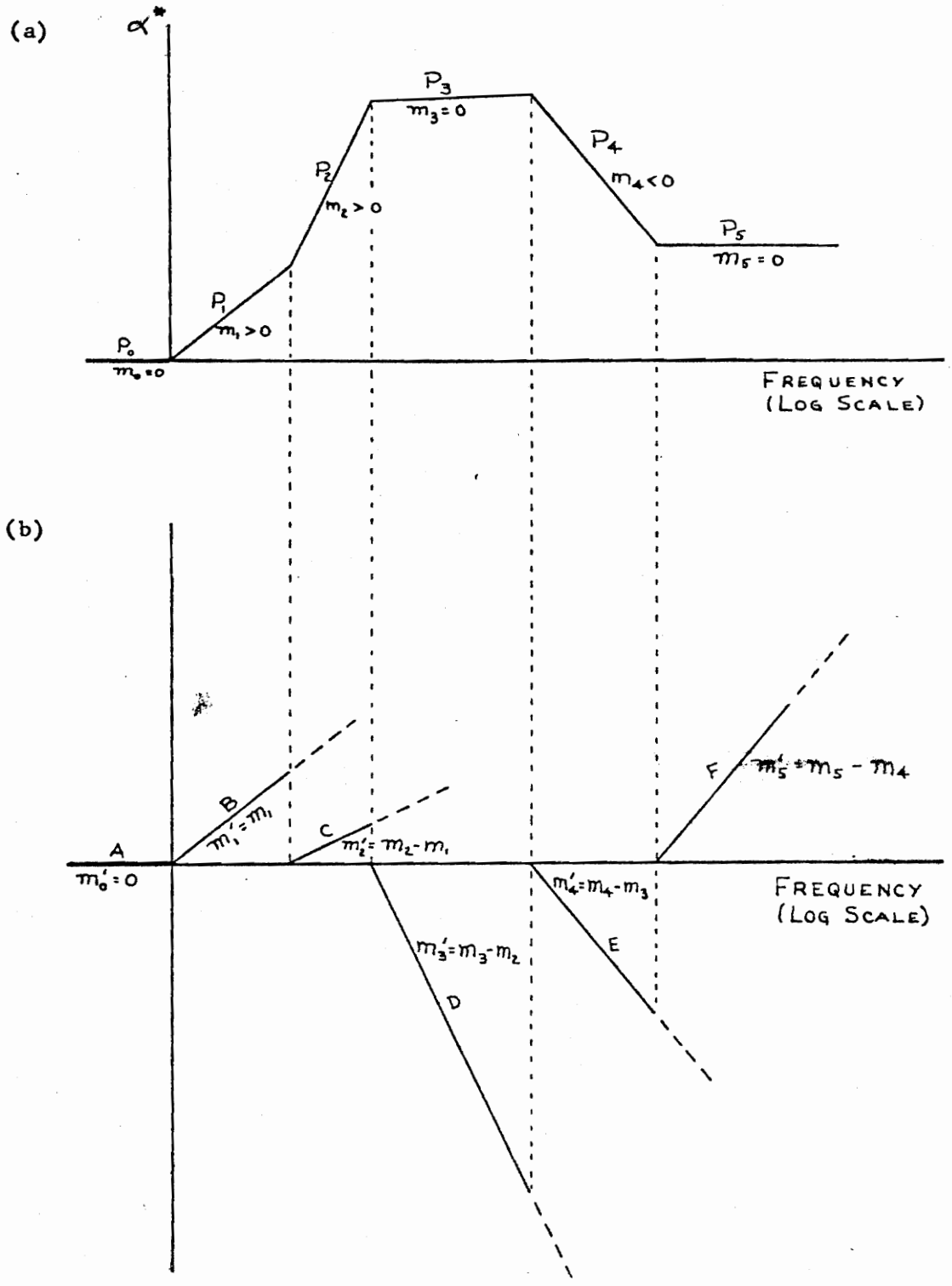


Figure 42. (a) Linear approximation of attenuation curve.
(b) Representation of (a) as a series of ramp functions.

proper slope.

Before proceeding it will be convenient to simplify Equation (D-16) to put it into a slightly different form[1]. The resulting equation is

$$\beta_o\left(\frac{\omega}{\omega_o}\right) = \frac{1}{\pi}\int_0^{x_o} \ln\left|\frac{1+x}{1-x}\right| \frac{dx}{x} \qquad \text{(D-17)}$$

where β_o is the phase of the representative ramp function whose slope is unity, $x = \omega_1/\omega$, and $x_o = \omega/\omega_o$, ω_o being the "break point" frequency of the ramp. As an illustration to show how Equation (D-17) is used to calculate the phase at a frequency ω, consider the case shown in Figure 43.

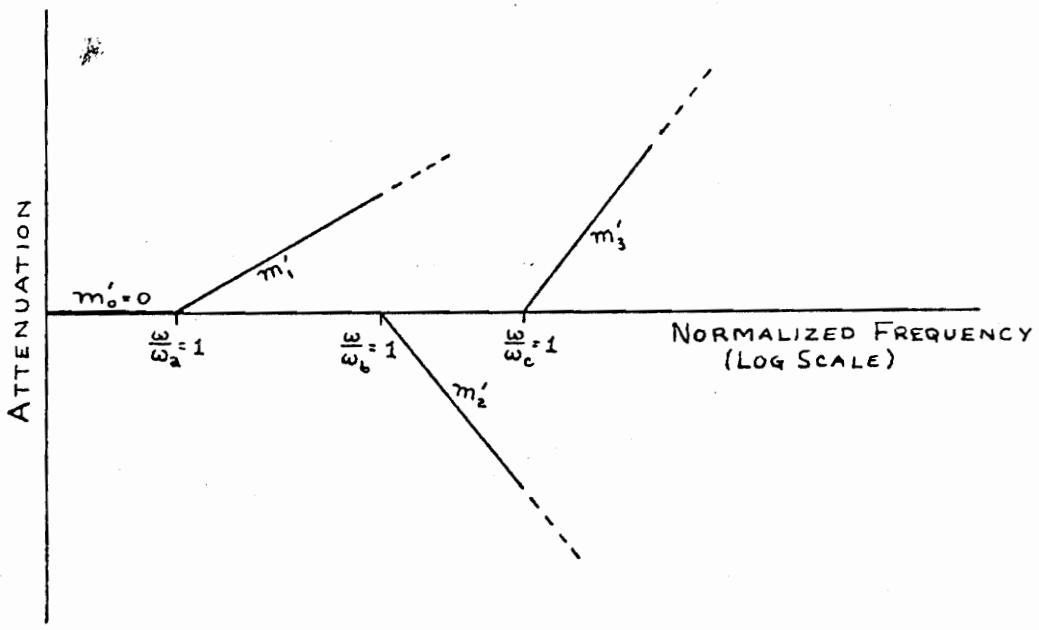


Figure 43. Network attenuation for illustration.

---

[1]For details, see chapter 14 in Bode, loc.cit.

The total phase shift will be equal to the sum of the component phase shifts. For the first segment where the slope of the attenuation characteristic is zero, the phase shift contributed is likewise zero. However, the phase shift that arises from the next three segments will not be zero. It can be calculated from Equation (D-17) by replacing ω_o by the break frequencies of the different ramps, multiplying by the proper slope, and finally taking the sum. The result is

$$\beta(\omega) = m_1' \beta_o\left(\frac{\omega}{\omega_a}\right) + m_2' \beta_o\left(\frac{\omega}{\omega_b}\right) + m_3' \beta_o\left(\frac{\omega}{\omega_c}\right) + \ldots \qquad \text{(D-18)}$$

Now it is necessary to find a method for evaluating β_o which will be relatively easy to implement on the digital computer. This will be accomplished by first expanding the integrand of Equation (D-17) by a Taylor series and then integrating.

$$\beta_o = \frac{1}{\pi} \int_o^{x_o} \ln \left|\frac{1 + x}{1 - x}\right| \frac{dx}{x}$$

$$= \frac{1}{\pi} \int_o^{x_o} \left[\left(x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots\right) - \left(-x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \ldots\right)\right] \frac{dx}{x}, \quad 0 < x_o < 1$$

$$= \frac{2}{\pi} \int_o^{x_o} \left(1 + \frac{x^2}{3} \quad \frac{x^4}{5} + \ldots\right) dx, \quad 0 < x_o < 1$$

$$\beta_o = \frac{2}{\pi}\left(x_o + \frac{x_o^3}{9} \quad \frac{x_o^5}{25} \quad \frac{x_o^7}{49} + \ldots\right), \quad 0 < x_o < 1. \qquad \text{(D-19)}$$

Equation (D-19) is only good for $\omega < \omega_o$ since the series only converges for these values. Since β_o is a symmetrical function, it can be shown

that

$$\beta_o\left(\frac{\omega}{\omega_o}\right) = \frac{\pi}{2} - \beta_o\left(\frac{\omega_o}{\omega}\right) \quad , \omega > \omega_o \quad . \tag{D-20}$$

This allows for Equation (D-19) to be used for all frequencies if interpreted properly with the aid of Equation (D-20).

Finally, before discussing the actual computer program, it would perhaps be wise to discuss what would be the most desirable forms for the input and output data. The slopes of the ramp functions are inherently in terms of nepers per unit of $\ln(\omega/\omega_o)$ because of the definitions made in Equation (D-15). That is, the attenuation constant as defined by Equation (D-15) is in nepers. Stated mathematically,

$$m = \frac{\Delta\alpha}{\Delta u} \tag{D-21}$$

where $u = \ln(\omega/\omega_o)$. If values of attenuation at frequencies ω_a and ω_b are used, then

$$m = \frac{\alpha_b - \alpha_a}{\ln(\omega_b/\omega_a)} \quad . \tag{D-22}$$

It is even more convenient to express this equation in terms of decibels in which case Equation (D-22) becomes

$$m = \frac{x_b - x_a}{20 \log_{10}(\omega_b/\omega_a)} \quad . \tag{D-23}$$

Where x_a and x_b are expressed in decibels and related to α_a and α_b by the formula $x = \frac{20\alpha}{\ln 10}$. Another form which the input data might take is that of a gain (e.g., output voltage divided by input voltage) instead of an attenuation. Also, it would be convenient to express this gain

as a simple ratio instead of a certain number of decibels. Since $x = 20 \log_{10} A$ where A is the gain, Equation (D-23) can be written as

$$m = \frac{\log_{10} (A_b/A_a)}{\log_{10} (\omega_b/\omega_a)} \quad . \qquad \text{(D-24)}$$

It is this equation that will be used in the digital computer solution.

The computer programs on pages 160 and 161 are the result of implementing the above method for numerical solution on the computer. Essentially, the program evaluates Equation (D-18) by use of Equations (D-19), (D-20), and (D-24). The program on page 160 is the main program which controls the overall solution, while the program on page 161 is just used to evaluate Equations (D-19) and (D-20).

The highlights of the FORTRAN programs will now be discussed. The execution of the main program begins by reading in NMAX, W(N), and A(N). The quantity NMAX is used to tell the computer how many sets of values of input data to enter. Each set of input data consists of a value of frequency represented by W(N) and the corresponding value of transfer function gain represented by A(N). The number of segments into which the input data is to be divided is then determined. This number is represented by the quantity MX and is, of course, equal to NMAX-1. The values of slope for each segment (corresponding to the values of M shown in Figure 42(a) are calculated next according to Equation (D-24) and stored as a one-dimensional array named P(I). From these values the slopes of the semi-infinite segments (corresponding to m' of Figure 42(b)) are evaluated and also stored as a one-dimensional array as S(I). Next Equation (D-18) is built up term-by-term until all segments have been

included. During the course of this calculation, the function subprogram FUNCTION BETA (Z) representing $\beta_o(\omega/\omega_o)$ is called upon many times to evaluate Equations (D-19) and (D-20). This subprogram simply makes a test to determine whether $\omega<\omega_o$ or not and makes the proper substitution into the Taylor series expansion. The series continues until the last term is less than 10^{-5} at which time adjustments are made according to Equation (D-20) if necessary. The value of β_o is then returned to the main program.

Finally, as the phase is computed at each frequency supplied by the input data, a set of output values are printed. A set of the output data consists of (1) the frequency at which the calculation was made, in radians per second, (2) the value of the network transfer function gain at this frequency, (3) the value of phase angle in degrees, and (4) this same value of phase angle except converted to radians to make it more versatile. Sample output data for a low-pass filter are shown on page 162. For the sample data, the program was altered slightly by substituting values calculated from the transfer function for the angle in radians.

In conclusion, it should be pointed out that the phase angle evaluated by the method outlined here is limited to the minimum value. This restriction came about as a result of the assumption made in Equation (D-12) where $y(\infty)$ was assumed to be zero. This is a necessary restriction, but fortunately, not normally a serious one.

```
C      BODE METHOD OF OBTAINING MINIMUM PHASE ANGLE FROM THE ATTENUATION.
C
       DIMENSION W(91),A(91),P(90),S(90)
   99  READ(5,1) NMAX,(W(N),N=1,NMAX),(A(N),N=1,NMAX)
    1  FORMAT(I3/(7F10.0))
       MX=NMAX-1
       WRITE(6,4)
    4  FORMAT(15H1W(RADIANS/SEC),6X,  9HMAGNITUDE,9X,14HANGLE(DEGREES),
      1       5X,14HANGLE(RADIANS)/)
       DO 6 I=1,MX
    6  P(I)=ALOG10(A(I+1)/A(I))/ALOG10(W(I+1)/W(I))
       S(1)=P(1)
C      P(J) IS THE INITIAL SLOPE AND S(J) IS THE ADJUSTED SLOPE.
       DO 10 I=2,MX
   10  S(I)=P(I)-P(I-1)
       DO 28 J=1,NMAX
       PART=0.
       DO 18 K=1,MX
   18  PART=PART+S(K)*BETA(W(J)/W(K))
       B=PART
       BR=3.1415927*B/180.
   28  WRITE(6,50) W(J),A(J),B,BR
   50  FORMAT(1X,4(1PE15.8,4X))
       GO TO 99
   60  STOP
       END
```

```
C
      FUNCTION BETA(Z)
      L=1
      C=360./3.14159265**2
      IF(Z.LT..999999) GO TO 1
      IF(Z.GT.1.000001) GO TO 5
      GO TO 4
    5 X=1./Z
      L=2
      GO TO 3
    1 X=Z
    3 BETA=0.
      DO 7 I=1,99,2
      T=I
      P=C*X**I/(T*T)
      IF(P-.00001)6,6,8
    8 BETA=BETA+P
    7 CONTINUE
    6 GO TO (10,9),L
    9 BETA=90. - BETA
      GO TO 10
    4 BETA=45.
   10 RETURN
      END
```

W(RADIANS/SEC)	MAGNITUDE	ANGLE(DEGREES)	CHECK VALUES
1.0000000 E 03	1.0049870 E 00	-5.6623568 E 00	-5.7675232 E 00
3.0000000 E 03	1.0436503 E 00	-1.8620564 E 01	-1.8245846 E 01
4.0000000 E 03	1.0748340 E 00	-2.5731437 E 01	-2.5463334 E 01
5.0000000 E 03	1.1094004 E 00	-3.3959103 E 01	-3.3690019 E 01
6.0000000 E 03	1.1399019 E 00	-4.3428602 E 01	-4.3152305 E 01
7.0000000 E 03	1.1546236 E 00	-5.4185954 E 01	-5.3923944 E 01
7.9999999 E 03	1.1399019 E 00	-6.5984971 E 01	-6.5772340 E 01
8.9999999 E 03	1.0871492 E 00	-7.8227701 E 01	-7.8078911 E 01
1.0000000 E 04	1.0000000 E 00	-9.0132137 E 01	-8.9999999 E 01
1.1000000 E 04	8.9296390 E-01	-1.0121192 E 02	-1.0080831 E 02
1.4000000 E 04	5.8909200 E-01	-1.2371293 E 02	-1.2443904 E 02
1.7000000 E 04	3.9338080 E-01	-1.3739765 E 02	-1.3802962 E 02
2.0000000 E 04	2.7735011 E-01	-1.4543761 E 02	-1.4631001 E 02
1.0000000 E 05	1.0048000 E-02	-1.7771171 E 02	-1.7999999 E 02

APPENDIX E

Other Related FORTRAN Programs

This appendix contains certain other FORTRAN programs that were used extensively during the course of the investigation. Some of these have been mentioned previously and some have not. All are connected with the network transfer function either as a means of making direct calculations or as a vehicle by which predetermined data can be entered into the program. The main features of each program will be discussed so that one can obtain an understanding of the philosophy involved in using the transfer function subprogram in conjunction with the main program. These programs are meant merely as examples and different programs would be written to fit particular applications.

Check Program for Testing Main Program

A FORTRAN program specifically designed for checking the execution of the main program is found on page 169. The function of this subroutine when called by the main program is simply to store ones in all locations designated by AH, AL, or AO and zeros in all locations designated by ANGH, ANGL, or ANGO. The quantities AH, AL, and AO are used to signify the gain at sideband frequencies higher than the carrier, the gain at sideband frequencies lower than the carrier, and the gain at the carrier frequency, respectively. The subscript shown by the number in the parenthesis in the program is the number of the

sideband. For example, the value AL(2) is the second sideband below the carrier. The same comments hold true for ANGH, ANGL, and ANGO except ANG is interpreted to be angle. These definitions are consistent for the single-frequency modulation programs throughout this paper. When the check program is used as the network transfer function, the output wave should be exactly equal to the input wave at all times. This allows one to evaluate the main program for errors and accuracy.

Program for Single-Tuned Network

The transfer function subroutine for a single-tuned circuit is found on page 170. This program directly evaluates the transfer function of the circuit shown in Figure 44 at each sideband frequency. The values of the gain and phase of the transfer function are stored in memory locations which are common to the main program so that they are available for computations. Although the values of the components in this circuit are fixed, it would be a simple matter to write the equations in terms of L, C, and R's. This is the circuit that was used to obtain the digital computer results which were verified in Chapter VI.

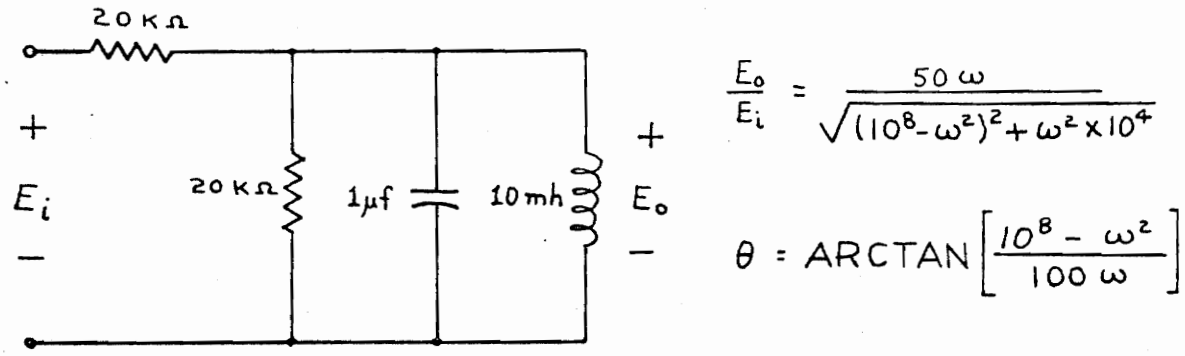


Figure 44. Single-tuned network and transfer characteristics.

Program for Entering Predetermined Data

A subroutine subprogram is shown on page 171 which allows one to enter network data in numerical form. The READ statement calls for IMAX sets of data to be entered in the form of four one-dimensional arrays called AH, AL, ANGH, and ANGL. These arrays have already been discussed in connection with the check program. The values of AO and ANGO are also entered into storage. It should be pointed out that it is imperative to this Fourier method that the phase angle at the carrier frequency be zero. However, if ANGO in predetermined data is not zero, the program will automatically normalize all of the angles so that this requirement is met.

Program for Computing the Transfer Function and Input Impedance of Multistage Filters.

This program was used to determine the transfer characteristics of the filter network described in Chapter V. Figure 11 in Chapter V shows the details of a single section and how sections are connected together to form the multistage filter. Actually, the program (see p. 172) is a great deal more general than is implied by Figure 11. For example, consider the general tee-configuration of the classical constant-k filter terminated by the impedance Z_T as shown in Figure 45. It is a simple matter to solve this circuit for the input impedance and the transfer function.

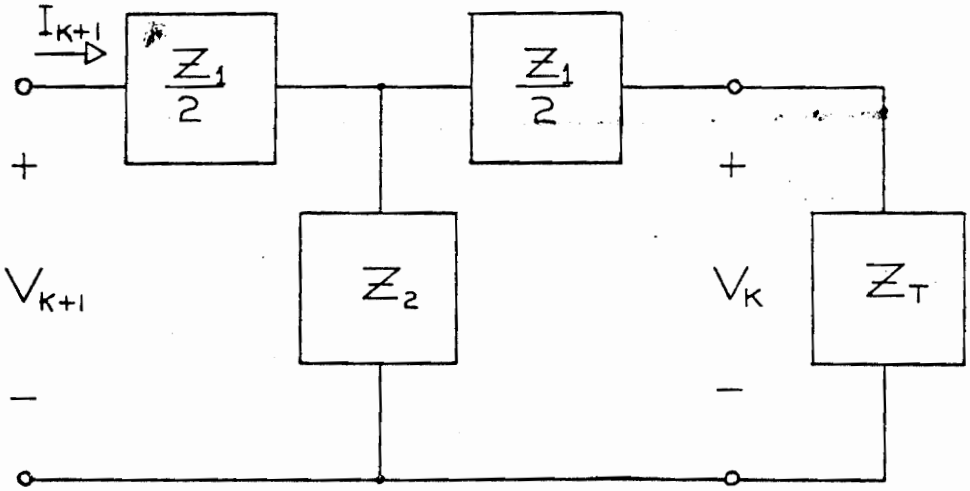


Figure 45. Constant-k filter, T-configuration.

The resulting equations are

$$Z_{k+1} = \frac{Z_1}{2} + \frac{Z_2\left(\frac{Z_1}{2} + Z_T\right)}{Z_2 + \frac{Z_1}{2} + Z_T} \qquad \text{(E-1)}$$

and

$$\frac{V_k}{V_{k+1}} = \frac{Z_2 Z_T}{Z_1 Z_2 + Z_2 Z_T + \frac{Z_1^2}{4} + \frac{Z_1 Z_T}{2}} \qquad \text{(E-2)}$$

Thus the input impedance and the transfer function have been obtained for the single section. Now if an identical section is placed in front of this one and if the input impedance just calculated is considered to be the terminating impedance of the section just added, then the input impedance of the added section can be obtained by again applying Equation (E-1). The transfer function of the new section can be obtained by making the same substitutions into Equation (E-2). The overall transfer function of both sections is obtained by taking the product of the individual transfer functions. This process can be repeated for as many sections as desired.

The FORTRAN program accomplishes this for the special case of the circuit of Figure 11. READ statement number 13 enters into the program the following values: NUM, R1, RT, F1, F2, DF, C1, C2, L1, FA, and FB. NUM determines how many sections are going to be used; R1, C1, and L1 are the components of the series arm; C2 is the shunt capacitor; F1 and F2 are the upper and lower cutoff frequencies for a single section; DF is the increment in frequency which determines the spacing between successive calculations; and FA and FB determine the range of frequencies over which the calculations are to be made. Following the

READ statement the program determines values for Z_T, Z_1, and Z_2. The input impedance and the transfer function are then computed and printed out. After this, the whole process is repeated as described above. It is to be noted that many of the computations are made in complex arithmetic and consequently the program is not directly applicable to earlier versions of the FORTRAN language. A set of sample output data is given on page 174.

This program is not directly compatible with the main program as a subroutine, so some other arrangements must be made for using the calculated data. This is done conveniently by use of the preceding program for entering predetermined data.

```
C
C       CHECK PROGRAM
C
        SUBROUTINE TFCN2(WC,W,IMAX)
        DIMENSION AH(60),AL(60),ANGH(60),ANGL(60),FT(360),A(10),B(10)
        COMMON AH,AL,ANGH,ANGL,A0,ANG0
        A0=1.
        ANG0=0.
        DO 3 I=1,IMAX
        AH(I)=1.
        AL(I)=1.
        ANGH(I)=0.
    3   ANGL(I)=0.
        RETURN
        END
```

```
C
C     TRANSFER FUNCTION OF SINGLE-TUNED CIRCUIT.
C
      SUBROUTINE TFCN2(WC,W,IMAX)
      DIMENSION AH(60),AL(60),ANGH(60),ANGL(60),FT(360),A(10),B(10)
      FMAG(A)=50.*A/SQRT((1.E8-A*A)**2 + A*A*1.E4)
      FANG(A)=ATAN2((1.E8-A*A),A*100.)
      COMMON AH,AL,ANGH,ANGL,A0,ANG0
      A0=FMAG(WC)
      ANG0=FANG(WC)
      DO 3 I=1,IMAX
      Q=I
      E=WC-Q*W
      D=WC+Q*W
      AL(I)=FMAG(E)
      AH(I)=FMAG(D)
      ANGL(I)=FANG(E)
    3 ANGH(I)=FANG(D)
      RETURN
      END
```

```
C
C      PROGRAM FOR ENTERING PREDETERMINED DATA.
C
       SUBROUTINE TFCN2(WC,W,IMAX)
       DIMENSION AH(60),AL(60),ANGH(60),ANGL(60),FT(360),A(10),B(10)
       COMMON AH,AL,ANGH,ANGL,A0,ANG0
       READ(5,500)A0,ANG0,(AH(J),J=1,IMAX),(AL(J),J=1,IMAX),
      1    (ANGH(J),J=1,IMAX),(ANGL(J),J=1,IMAX)
  500  FORMAT(8F10.0)
       DO 12 N=1,IMAX
       ANGH(N)=ANGH(N)-ANG0
   12  ANGL(N)=ANGL(N)-ANG0
       ANG0=0.
       RETURN
       END
```

```
C
C     PROGRAM FOR COMPUTING THE TRANSFER FUNCTION AND INPUT IMPEDANCE OF
C     MULTISTAGE FILTERS.                                P.B.JOHNSON   3/10/66
C
C
      EXTERNAL NOLMUN
      COMPLEX TFCN,Z1,Z2,ZI,ZT
      REAL L1
   13 READ(5,50) NUM,R1,RT,F1,F2,DF,C1,C2,L1,FA,FB
   50 FORMAT(I2/5F10.0/3E15.8/2F10.0)
      PI=3.1415927
      WA=2.*PI*FA
      WB=2.*PI*FB
      DW=2.*PI*DF
      W=WA
   30 ZT= CMPLX(RT,0.)
   31 F=W/(2.*PI)
      WRITE(6,60) F
   60 FORMAT(12H1FREQUENCY =1PE15.8//)
      WRITE(6,61)
   61 FORMAT(8H SECTION,9X,15HINPUT IMPEDANCE,16X,4HGAIN,11X,
     1         14HPHASE(DEGREES))
      WRITE(6,63)
   63 FORMAT(13X,1HR,18X,1HX/)
      Z1= CMPLX(R1,(W*L1-1./(W*C1)))
      Z2= CMPLX(0.,-1./(W*C2))
      TFCN=(1.,0.)
      DO 10 I=1,NUM
      ZI=Z1/(2.,0.)+Z2*(Z1/(2.,0.)+ZT)/(Z2+Z1/(2.,0.)+ZT)
      TFCN=TFCN*Z2*ZT/(Z2*(Z1+ZT)+Z1*(Z1/(4.,0.)+ZT/(2.,0.)))
      X=REAL(TFCN)
      Y=AIMAG(TFCN)
      GAIN=SQRT(X*X+Y*Y)
      PHI=ATAN2(Y,X)
      PHASE=180.*PHI/PI
```

```
      WRITE(6,62) I,ZI,GAIN,PHASE
   62 FORMAT(3X,I2,3X,4(1PE15.8,3X))
   10 ZT=ZI
      W=W+DW
      IF(WB-W) 13,30,30
  100 STOP
      END
```

FREQUENCY = 1.0520000 E 05

SECTION	INPUT IMPEDANCE R	INPUT IMPEDANCE X	GAIN	PHASE(DEGREES)
1	5.4982577 E 02	6.2314278 E 02	7.6949843 E-01	1.7763951 E 02
2	2.9063730 E 02	5.6217047 E 02	5.4730197 E-01	4.7310311 E 00
3	2.3409777 E 02	4.9937699 E 02	3.6795150 E-01	-1.6390538 E 02
4	2.2558984 E 02	4.7024048 E 02	2.3856012 E-01	2.8743109 E 01
5	2.2719131 E 02	4.5871968 E 02	1.5191673 E-01	-1.3848192 E 02
6	2.2962782 E 02	4.5484634 E 02	9.6066774 E-02	5.4159521 E 01
7	2.3114620 E 02	4.5384879 E 02	6.0619483 E-02	-1.1331611 E 02
8	2.3186120 E 02	4.5374066 E 02	3.8235430 E-02	7.9145296 E 01
9	2.3213770 E 02	4.5382100 E 02	2.4117535 E-02	-8.8420105 E 01
10	2.3222418 E 02	4.5389624 E 02	1.5213835 E-02	1.0400510 E 02

1FREQUENCY = 1.0540000 E 05

DISSERTATION ABSTRACT

A STUDY OF THE EFFECTS OF LINEAR NETWORKS ON FM WAVES

Preston Benton Johnson

Doctor of Philosophy, June 12, 1966
(M.S., North Carolina State University, 1962)
(B.E.E., North Carolina State University, 1958)

Directed by Dr. Harry K. Ebert, Jr.

The analysis of the distortion which results when frequency-modulated waves are passed through linear networks is investigated by the Fourier method and the Quasi-steady-state method. The major emphasis is placed on the Fourier method, and extensive digital computer programs are developed to allow this method to be implemented on the modern, high-speed digital computer. In the Fourier method, the frequency-modulated wave which is applied to the input of a linear network is broken up into its Fourier spectrum. Each of the resulting "sideband" frequencies is then passed through the network and is subjected to alterations in amplitude and phase. The output wave is then synthesized by taking the vector sum of the "weighted" sideband components. In contrast to the single pair of sideband frequencies generated by amplitude modulation, the spectrum of a frequency-modulated wave contains an infinite number of sideband components. Fortunately, only a

relatively small number of these sidebands have significant influence on the total makeup of the waveform. The number of significant sidebands is proportional to the value of modulation index. When the modulation index is high, the number of significant sidebands is very large and the number of computations required by the Fourier method becomes enormous. Previously considered to be completely impractical, the Fourier method was usually abandoned in favor of the Quasi-steady-state approach. However, the digital computer techniques developed in the course of this investigation allow for a fast, economical, and convenient analysis based on the Fourier method even when the modulation index is relatively high. Analyses were performed for values of modulation index up to 45 and techniques are discussed for increasing this range.

The Quasi-steady-state method is based on the assumption that the frequency of the input wave is changing slowly enough that the frequency of the output wave at any instant is equal to the "instantaneous frequency" of the input wave. This method is inherently in error since it neglects the transient terms generated by the changing frequency. To compensate for this error, it is the general practice to incorporate correction terms, usually in the form of an infinite series. The Quasi-steady-state method is more effective at low modulating frequencies (high modulation index). While the analysis contained in this paper considers in detail only a first-order correction, the application of higher-order correction terms is discussed. The results obtained from applying both analyses to a complex, multi-section filter indicate that

the computer solution of the Fourier method is preferable for intermediate values of modulation index.

Experimental verification of the Fourier method is obtained by simulating the system on an analog computer. The advantages of this rather novel approach are discussed in some detail. The agreement between the results predicted by the digital computer and those obtained experimentally leaves no doubt to the validity and accuracy of the analysis.

Digital computer programs for analyzing the distortion using each of the above methods are given. Subprograms are also included, some of which can be used independently. Among these are a program that computes Bessel functions of the first kind for positive and negative orders and a program that computes the minimum phase shift of a network from its attenuation. All programs are written in the FORTRAN IV computer language and were executed on the IBM 7040/1401 system.