STANFORD ARTIFICIAL INTELLIGENCE LABORATORY
MEMO AIM- 208

STAN-CS-73-379

THE HETRODYNE FILTER AS A TOOL FOR ANALYSIS OF TRANSIENT WAVEFORMS

BY

JAMES ANDERSON MOORER

SUPPORTED BY

ADVANCED RESEARCH PROJECTS AGENCY

ARPA ORDER NO. 457

JULY 1973

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

STANFORD ARTIFICIAL INTELLIGENCE LABORATORY JULY 14, 1973
MEMO A I M-208

COMPUTER SCIENCE DEPARTMENT
REPORT CS-379

THE HETROOYNE FILTER AS A TOOL FOR ANALYSIS OF TRANSIENT WAVEFORMS

by

James Anderson Moorer

ABSTRACT: A method of analysis of transient waveforms is discussed. Its properties and limitations are presented in the context of musical tones. The method is shown to be useful when the risetimes of the partials of the tone are not too short. An extention to inharmonic partials and polyphonic musical sound is discussed.

This research was supported in part by the Advanced Research Projects Agency of the Office of the Secretary of Defense under Contract No. SD-183.

List of Figures

INTRODUCTION

The analysis of the attack transients of vocal or musical tones goes back as far as 1932 with Backhaus' [1,2] tunable resonator and drum recorder. Luce and Clark [3,4,5] used fi l tering methods to select partial tones for analysis and recording. More recently, with the advent of computer music, analysis of musical instruments for the purpose of simulation of timbre has been done by what will be called a "hetrodyne f i l ter" for want of a better name, Beauchamp [6] analysed each partial of a complex waveform by first multiplying the waveform by a sin and cosine at the frequency of the partial in question. The result was then low-pass filtered, then squared and summed. Freedman [7,8,9], and later Keeler [10,11] used a discrete finite summation over one period of the fundamental frequency in place of Beauchamp's low-pass filter, an effect which as we shall show later conveniently places a zero of transmission at all harmonic partials other than the one in question.

It is the purpose of this article to explore this method,' report its characteristics, its limitations, its uses and some simple extensions.

THE METHOD

Let us define the hetrodyne filter as follows, We begin with a discrete function Fi which represents a continuous function F(t) at discrete intervals t=ih, where h is the time between samples. h is called the "sampling interval." The reciprocal of h is called the "sampling frequency" or the "sampling rate." Let us define a and b as follows:

$$a = \sum_{i=\alpha}^{\alpha+N-1} F_i \cos(\omega_o ih + \theta_o)$$

$$b = \sum_{i=\alpha}^{\alpha+N-1} F_i \sin(\omega_o ih + \theta_o) \qquad (1)$$

ω_c will be called the "center frequency".

Without loss of generality, we may define

$$\varphi_o = \theta_o + \alpha h \qquad (2)$$

and thus rewrite the sum as going from 0 to N-1.

$$a = \sum_{i=0}^{N-1} F_i \cos(\omega_o ih + \varphi_o)$$

$$b = \sum_{i=0}^{N-1} F_i \sin(\omega_o ih + \varphi_o) \qquad (3)$$

This change of variables hides the time-position of the filter in the phase. We must remember that as the filter is advanced through time, the phase angle will increase, and that any results which depend on this phase angle will be functions of time.

Since the summation operation is linear, we may represent the input waveform F_i as a sum of sinusiods and may thus examine the response of the filter to a sinusiodal excitation, as is commonly done with linear filters.

$$a = \sum_{i=0}^{N-1} A \cos(\omega ih + \varphi_o)$$

$$b = \sum_{i=0}^{N-1} A \cos(\omega ih + \varphi) \sin(\omega_o ih + \varphi_o) \qquad (4)$$

With the help of the summation calculus [12] and some trigonometric identities, we may compute the summation in closed form without error as follows:

$$a = \frac{A}{2N}\left\{\frac{\sin[(\omega+\omega_o)Nh/2]\ \cos[(\omega+\omega_o)(N-1)h/2+\varphi+\varphi_o]}{\sin[(\omega-\omega_o)h/2]}\right.$$

$$\left.+\ \frac{\sin[(\omega-\omega_o)Nh/2]\ \cos[(\omega-\omega_o)(N-1)h/2+\varphi-\varphi_o]}{\sin[(\omega-\omega_o)h/2]}\right\} \qquad (5)$$

$$b = \frac{A}{2N}\left\{\frac{\sin[(\omega+\omega_o)Nh/2]\ \sin[(\omega+\omega_o)(N-1)n/2+\varphi+\varphi_o]}{\sin[(\omega+\omega_o)h/2]}\right.$$

$$\left.+\ \frac{\sin[(\omega-\omega_o)Nh/2]\sin[(\omega-\omega_o)(N-1)h/2+\varphi-\varphi_o]}{\sin[(\omega-\omega_o)h/2]}\right\}$$

This is not a very useful expression as it stands, but it may be simplified somewhat by computing the sum of the squares of a and b.

$$a^2+b^2 = \frac{A^2}{4N^2}\left\{\frac{\sin^2[(\omega+\omega_o)Nh/2]}{\sin^2[(\omega+\omega_o)h/2]}+\frac{\sin^2[(\omega-\omega_o)Nh/2]}{\sin^2[(\omega-\omega_o)h/2]}\right.$$

$$\left.+\ \frac{\sin[(\omega+\omega_o)Nh/2]\sin[(\omega-\omega_o)Nh/2]}{\sin[(\omega+\omega_o)h/2]\ \sin[(\omega-\omega_o)h/2]}\cos[\omega_o(N-1)h+2\varphi_o]\right\}$$

(6)

Now if one chooses N to be such that

$$Nh\omega_o = 2\pi k,\ k = \text{any integer} \neq 0$$

Then some terms in equation 6 collapse to produce

$$a^2+b^2 = \frac{A^2}{4N^2}\sin^2(\omega Nh/2)\left\{\frac{1}{\sin^2[(\omega+\omega_o)h/2]} + \frac{1}{\sin^2[(\omega-\omega_o)h/2]} + \frac{2\cos[\omega_o h - 2\varphi_o]}{\sin[(\omega+\omega_o)h/2]\,\sin[(\omega-\omega_o)h/2]}\right\} \quad (7)$$

The square root of the above expression will be termed the "magnitude" of the output of the hetrodyne filter. The arctangent of the ratio of a to b will be called the "phase" of the output of the hetrodyne filter.

This process is similar to the discrete Fourier transform, except that only one frequency is processed instead of many. The results of this analysis can easily be generalized to represent the output of the DFT by setting the period of the center frequency to a multiple of the sampling interval.

If one further assumes that the frequency of the input sinusoid is close to the center frequency, then we see that

$$\lim_{\omega \to \omega o} a^2+b^2 = \frac{A^2}{4N^2}\left\{ 0 + N^2 + 0 \right\} = \tfrac{1}{4}A^2$$

define $\Delta\omega = \omega - \omega o$

$$\lim_{\omega \to \omega o} \frac{b}{a} = \frac{\sin\{2\omega oh[(N-1)/2+\alpha]\}+N\sin\{\Delta\omega h[(N-1)/2+\alpha]\}}{\cos\{2\omega oh[(N-1)/2+\alpha]\}+N\cos\{\Delta\omega h[(N-1)/2+\alpha]\}}$$

if $N \gg 1$ then

$$\lim \frac{b}{a} \approx \tan\{\Delta\omega h[(N-1)/2 + \alpha]\}$$

Thus we see that in the limit, the magnitude of the output of the filter becomes independant of time and becomes a measure of the amplitude of the input sinusoid, The phase of the output of the filter remains always a function of time, but is also a linear function of time, its slope being determined by the difference of the center frequency and the input frequency.

This reveals a method of determining the amplitude of the input sinusiod and getting a better estimate of its frequency. If the center frequency of the summation is near the actual frequency of the input sinusoid, the phase will be very nearly a linear function of time, thus we may find the frequency deviation by fitting the phase with a straight line and observing its slope.

The consequences of choosing N as above are significant. If the center frequency is a multiple of some fundamental frequency, then we may choose N to coincide with the period of the fundamental and thus cancel out all the harmonic partials except the center frequency. Figure 1 shows the log of the magnitude of the output of the hetrodyne filter for a range of sinusoidal inputs. The center frequency in this plot is 400 Hz and the summation period, Nh, is 10 milliseconds. Figure 2 shows the log of the magnitude versus frequency for a center frequency of 100 Hz and the same summation period. Notice the zeros of transmission at all multiples of the summation period except the center frequency.

ERROR ANALYSIS

Freedman [9] and Keeler [11] both show that this method is sufficiently accurate for their purposes even when the input signal is not a perfect sum of sinusoids. Keeler does not even bother computing the summation at each point, but at regular intervals only, and presents us with an elegant proof that the error in doing so is negligable. The above work may well seduce one as it did the author into believing that this is a perfectly accurate method, universally applicable. This is not so. To persue the matter further, let us reformulate the equations somewhat. We shall compute not two summations but one:

$$G_\alpha = \sum_{i=\alpha}^{\alpha+N-1} F_i e^{i\omega_o h+\theta_o} \tag{9}$$

The result will be a complex quantity whose real and imaginary parts correspond to the a and b discussed earlier. We will be interested in the magnitude and the phase of G. For the input waveform, Fi, we shall take a complex sinusoid with exponential decay.

$$G_\alpha = \sum_{i=o}^{\alpha+N-1} e^{(\sigma+j\omega)ih+\theta} e^{j\omega_o ih+\theta_o}$$

$$= e^{\alpha h[\sigma+j(\omega+\omega_o)]+j(\theta+\theta_o)} \frac{e^{Nh[\sigma+j(\omega+\omega_o)]}-1}{e^{h[\sigma+j(\omega+\omega_o)]}-1} \tag{10}$$

Again, we see the magnitude is related to the amplitude of the input sinusoid and the phase drift with time is related to the frequency differece. The exponential decay of the input signal causes imperfect cancellation of other harmonic partials, and depending on the speed of the attack, the deviation can be important.

A more reveal ing case would be to assume the signal begins at zero amplitude, and rises exponentially to its steady-state value and that the signal begins somewhere during the summation, say at $i=\beta$.

$$G_{\alpha,\beta} = \sum_{i=\alpha+\beta}^{\alpha+N-1} (1-e^{\partial ih})\, e^{j\omega ih+\theta}\, e^{j\omega_o ih+\theta_o}$$

$$= e^{(\alpha+\beta)jh(\omega+\omega_o)+j(\theta+\theta_o)} \frac{e^{(N-\beta)jh(\omega+\omega_o)}-1}{e^{jh(\omega+\omega_o)}-1} \qquad (11)$$

$$- e^{\partial h(\alpha+\beta)} \frac{e^{(N-\beta)h[\partial+j(\omega+\omega_o)]}-1}{e^{h[\partial+j(\omega+\omega_o)]}-1}$$

This is equivalent to setting N to some smaller value. As the filter progresses through the attack, the effective width of the window wi l l approach N, and the response of the filter w i l l become more representative.

Here we see that even if N and the center frequency are carefully chosen, the frequency response is not the same, Figure 3 shows the response for a center frequency of 300 Hz, a summation interval of 10 milliseconds, and $\beta=N/2$. We see that the neighboring harmonics are not canceled out. This shows that if the signal begins anywhere within the sindow, the output should not be taken to be an accurate indication of the amplitude of the partial. It is exactly analagous to taking N to be a non-multiple of the period of the fundamental frequency. Figure 4 shows the output for an averaging window of 10 milliseconds and center frequencies of 100, 200, and 300 Hz. The input was a sum of sinusoids of unit amplitude and frequencies 100, 200, and 300 Hz with exponential attacks of time constants 30, 20, and 10 milliseconds respectively. The leakage among the harmonics is apparent here.

If the attack is not exponential, another form of distortion can occur. Figure 5 shows the magnitude of the filter output for the same input and center frequencies as figure 4, but with linear attacks rather than exponential.

The presence of inharmonic partials can cause an effect similar to amplitude modulation. Figure 6 shows the magnitude and phase of the fundamental of a guitar note at 132 Hz. The apparent modulation is caused by an inharmonic partial at 186 Hz, the frequency of a known box resonance.

Another source of error is that of frequency quantization. N can not in general be chosen such that Nh is exactly the period of the fundamental frequency and still have N be an integer. This can be tolerated, but it also implies that the center frequency must be a multiple of $2\pi/Nh$ rather than a multiple of the fundamental frequency. If we do not set the center frequency to exactly a multiple of $2\pi/Nh$, we get imperfect cancellation of a pole and a zero. Figure 7 shows the magnitude versus frequency of such a case. Note the doublet around 400 Hz, the center frequency.

Since Nh is not exactly the period of the fundamental frequency, the harmonic partials are not exactly cancelled out. Fur thermore, most string instruments show a deviation from perfect integral multiples of the fundamental frequency. This deviation also contributes to leakage among the harmonics. Equation 10 may be used to determine exactly how much leakage is present.

A SIMPLE EXTENSION

Although there would seem to be no good solution to the problem of the note beginning within the summation period, there is a technique for dealing with the inharmonic partials. if the attack time of the partial in question is not too swift, it can be filtered out before analysis of the harmonic partials is done. The harmonic partials may then be filtered out to allow analysis of the inharmonic partials.

The filter advocated here is a comb filter. This is described simply by the recurrence relation:

$$Y_n = X_n - X_{n\ m} \qquad (12)$$

With frequency response

$$|Hm(\omega)| = \sqrt{\sin^2(\omega mh) + [\cos(\omega mh)-1]^2} \qquad (13)$$

. We see that the comb filter has a zero of transmission at all multiples of the base frequency 1/mh. The only hazards with the comb filter are those of transient response and preturbing the harmonic partials. The transient response of a comb filter is explicit. It is identical ly zero beyond mh seconds. If this can be tolerated, then the f i lter may be useful,

If the frequencyy of some harmonic partial falls near one of the zeros of transmission of the comb, it will be attenuated. We may prevent this by a method which appeared in Gold and Rader [13,14]. At those zeros of the comb that we wish to eliminate, a digital resonater is used to cancel out the zero. The details of the configuration are described in ful l detail in the references and will not be repeated here.

For analysis of an inharmonic partial, the harmonic partials may all be eliminated with a single comb, subject to the limitation that the partials may not be exact multiples of the fundamental, and that the frequency of the fundamental becomes quantized when the comb length, m, is chosen.

The comb wi l l, of course, attenuate the signal under analysis by some amount. We may predict the attenuation from equation 13 and then multiply the results of the hetrodyne analysis by the reciprocal of the attenuation factor to obtain a better estimate of the amplitudes of the partials.

CURRENT USES

The technique of combing out unwanted signals and then analysing the remaining waveform is the basis for an automated system for the analysis of polyphonic music currently being developed by the author. When two or more instruments are playing simultaneously, a Fourier analysis is used to get an estimate of the pitch and duration of each note. All notes but one are then eliminated by combing, and the hetrodyne filter is applied to determine the attack time of the note and to correct the estimated pitch of the note. This is iterated for each of the notes in the piece. The result is subjected to a heuristic analysis and is eventually displayed as a musical score of the piece under analysis.

Also under study by a colleague is how the physical parameters of musical tones such as risetimes of the partials, trenielo, and steady-state value contribute to the perceived timbre of the instrument. The hetrodyne method as outlined above is used to determine these parameters from digitized recordings of tones of actual musical instruments.

CONCLUSION

The hetrodyne filter has been shown to be a useful method for the analysis of the partials of musical tones as long as its limitations are observed. It can fail if the attack times are too quick, if the frequencies of the partials deviate too far from perfect integral multiples of the fundamental frequency, or if the sampling rate is so low that frequency quantization effects become significant, It can also fail if substantial frequency modulation (vibrato) is present.

An extention of the method to tones with inharmonic partials and even to multiple simultaneous notes was shown to be possible by use of the comb filter as long as the effects of the transient response of the comb was judged to be tolerable.

These methods are currently in use by the author and his colleagues in the analysis of digitized musical sound,

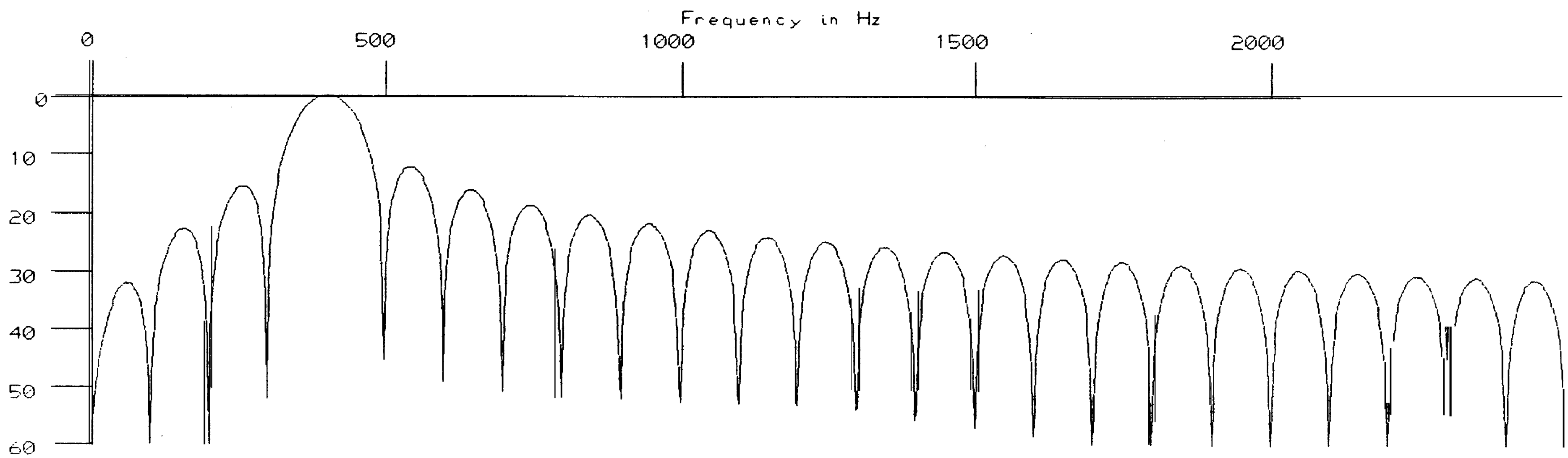

Figure 1

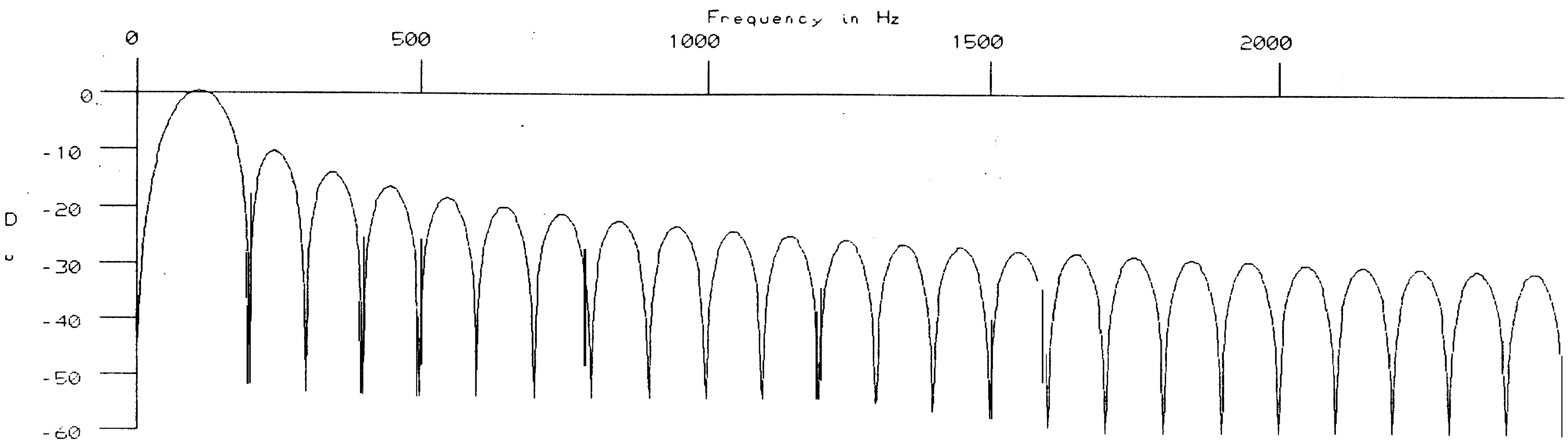

Figure 2

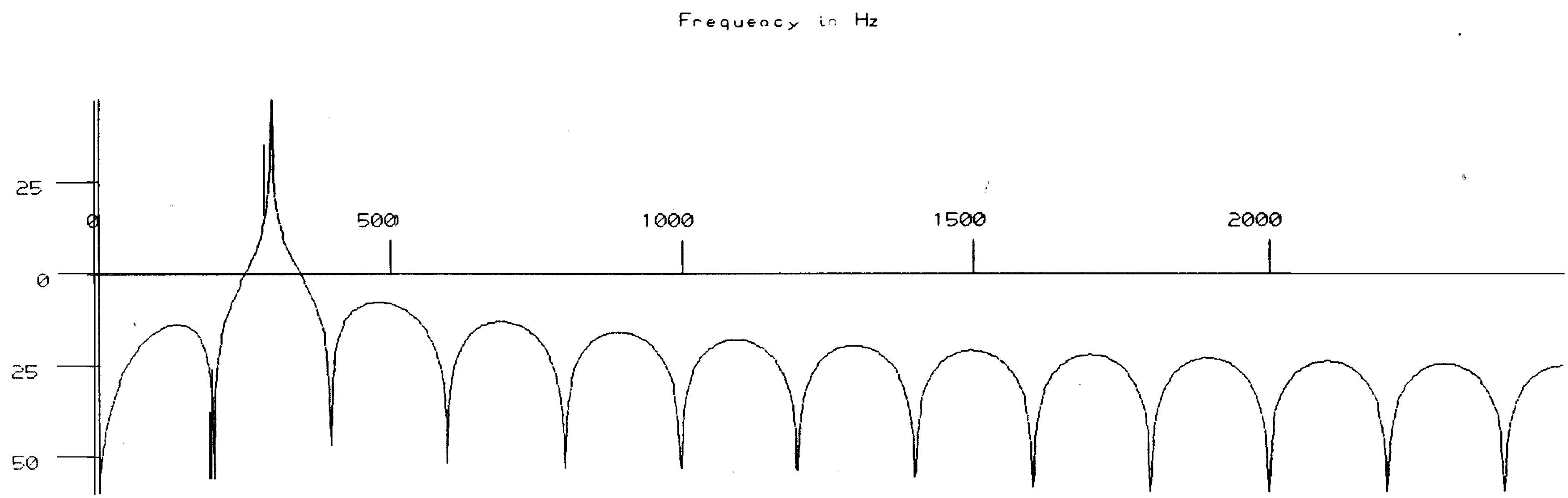

Figure 3

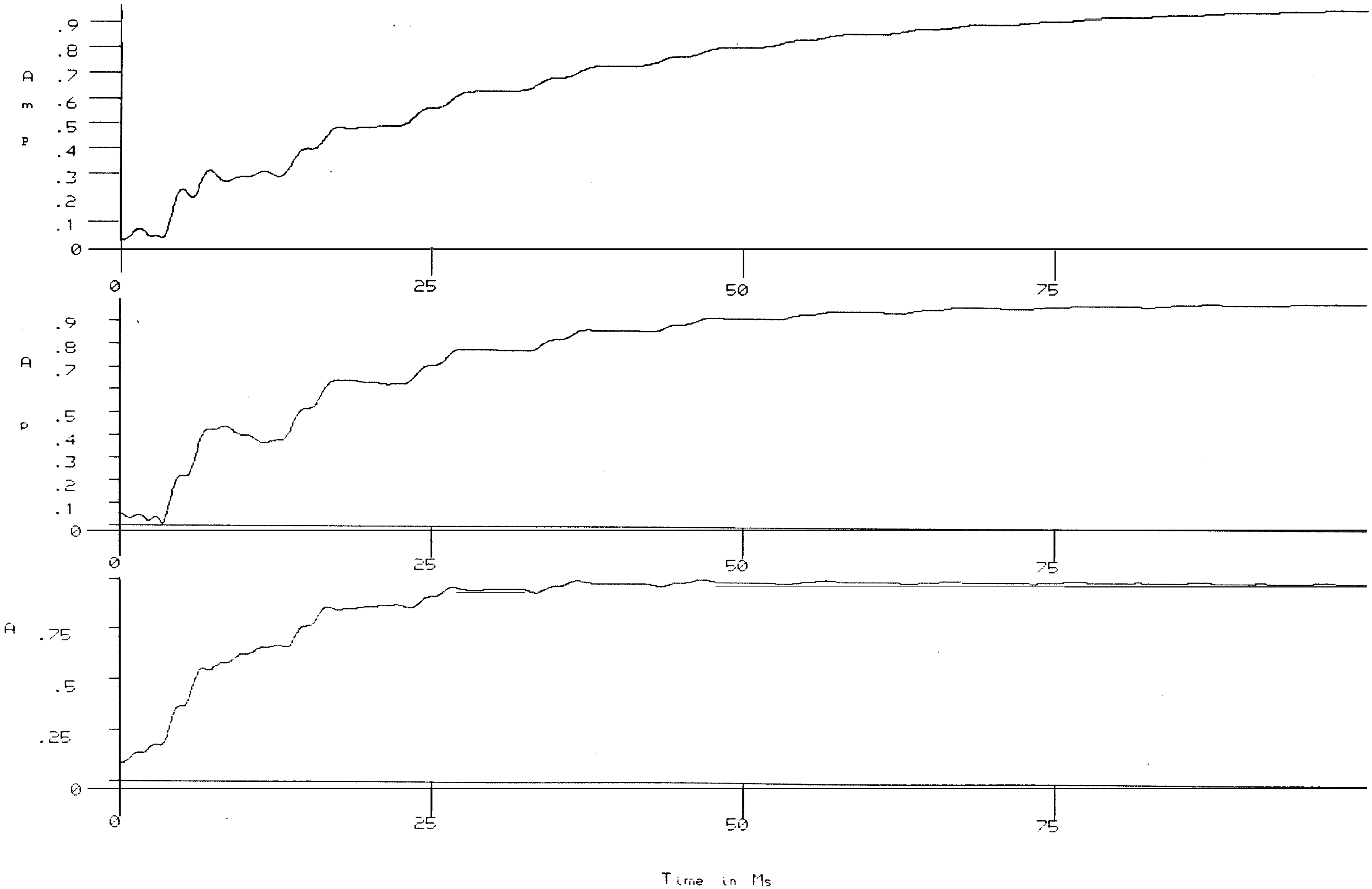
A
m
p
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
0
25
50
75
A
p
.9
.8
.7
.5
.4
.3
.2
.1
0
0
25
50
75
A
.75
.5
.25
0
0
25
50
75
Time in Ms

Figure 4

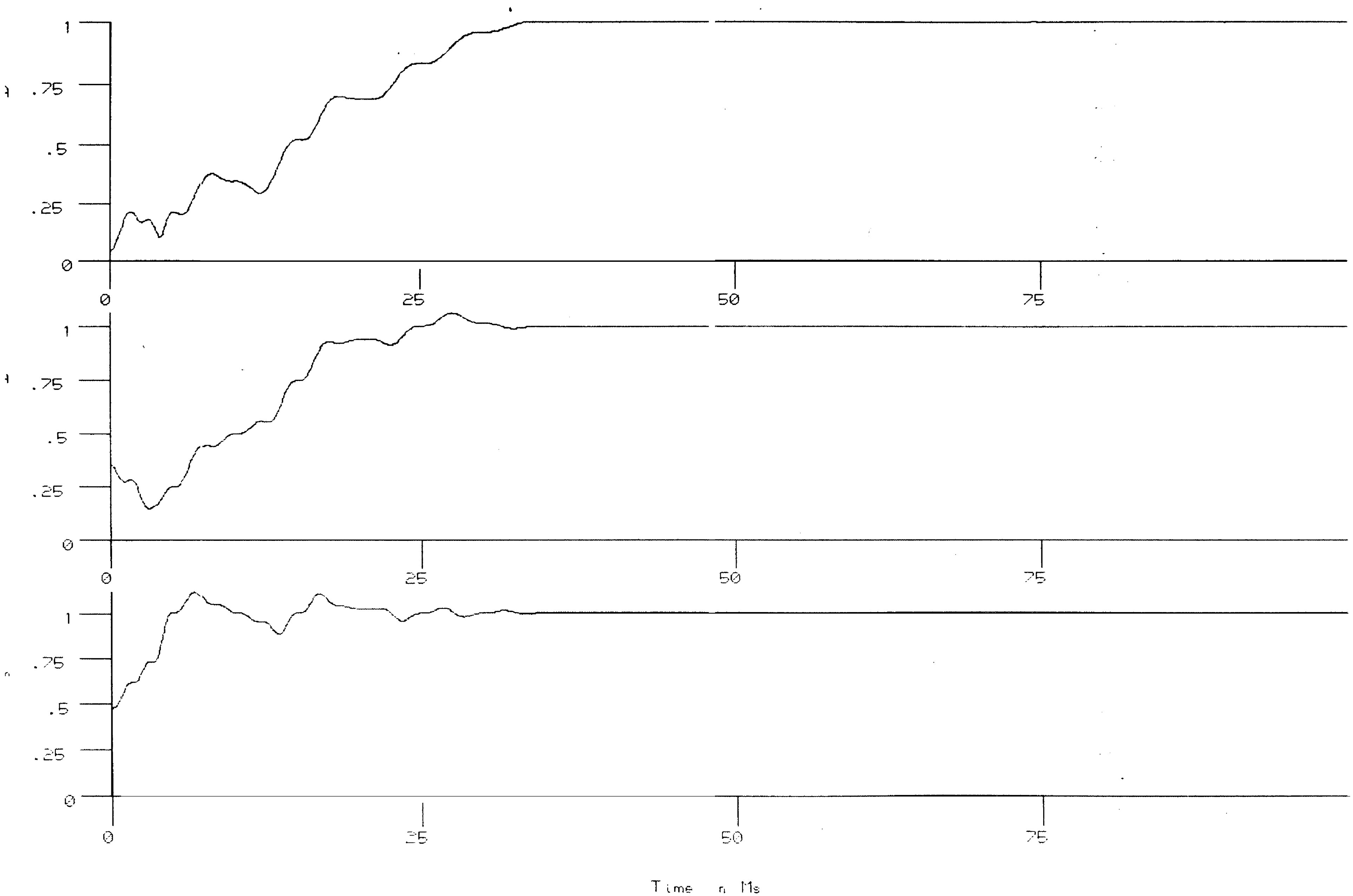

Figure 5

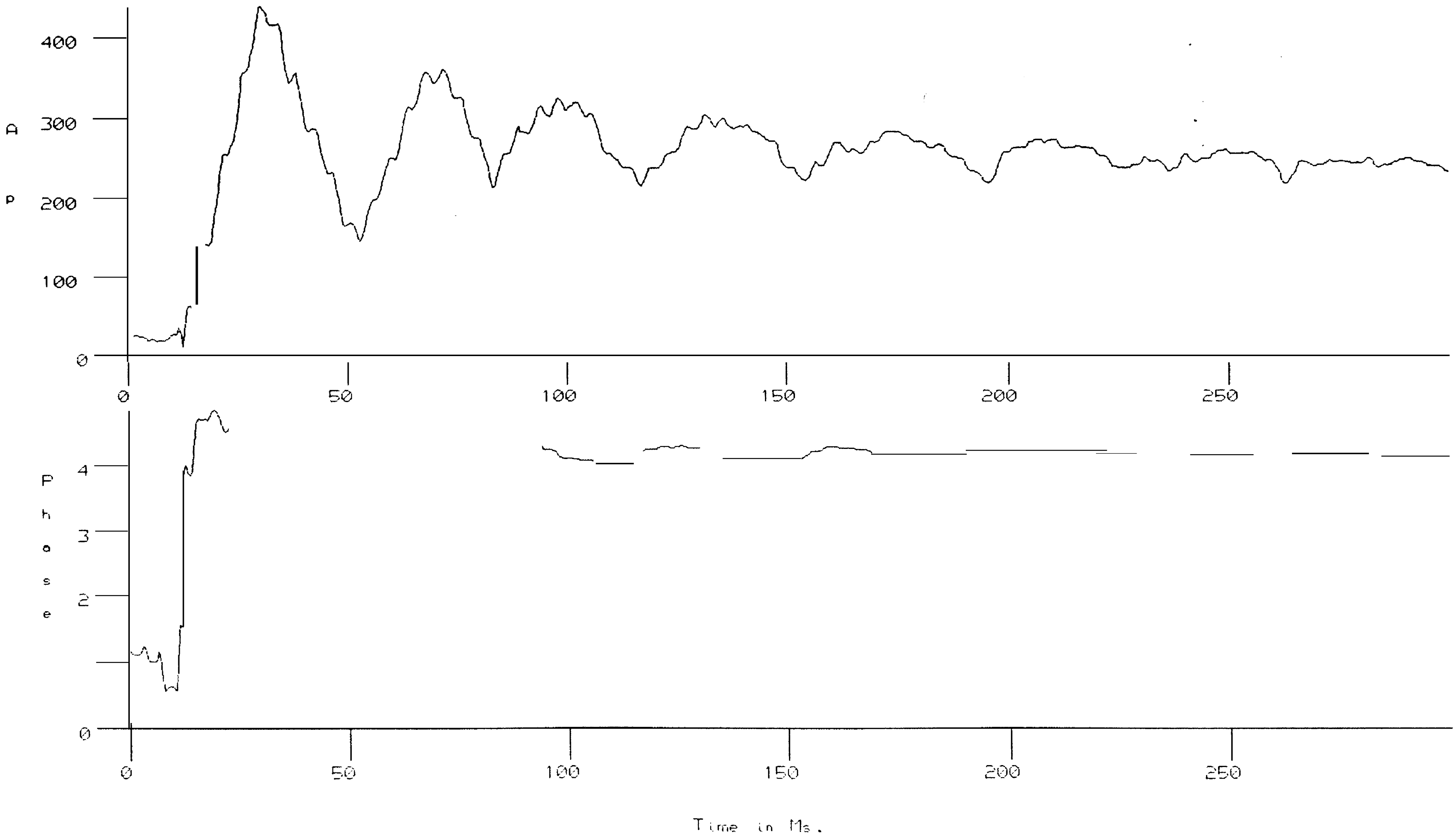
A
p
400
300
200
100
0
0
50
100
150
200
250
P
h
a
s
e
4
3
2
0
0
50
100
150
200
250
Time in Ms.

Figure 6

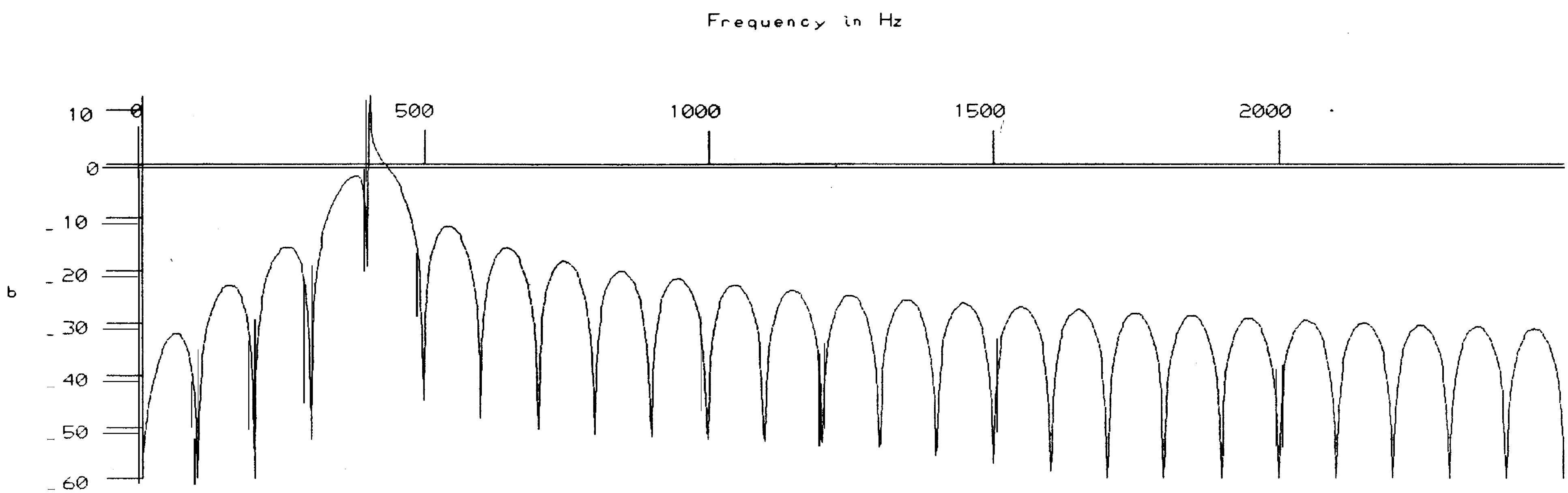

Figure 7

REFERENCES

1) H. Backhaus, "Uber Geigenklange," Zeitschrift fur Technische Physik, Vol 8, #11, 1927, pp509-515

2) H. Backhaus, "Uber die Bedeutung der Ausgleichsvorgange in der Austrik," Zeitschrift fur Technische Physik, vol 13, #1, pp31-46, 1932

3) David A. Luce "Physical Correlates of Nonpercussive Musical Instrument Tones," PhD thesis, MIT, 1963

4) David Luce, Melville Clark, 'Physical Correlates of Brass Instrument Tones," J. Acoust, Soc. Am, vol 42, p1232, 1967.

5) David Luce, Melville Clark, 'Duration of Attack Transients of Nonpercussive Orchestral Instruments," J. Audio Eng. Soc., vol 13, #3, p194, 1965.

6) James W. Beauchamp, "A Computer System for Time-Variant Harmonic Analysis and Synthesis of Musical Tones," in "Music by Computers," J. W. Beauchamp and Heinz Von Foerster eds, John Wiley and Sons, New York,

7) M.D. Freedman, "Analysis of Musical Instrument Tones,' J. Acoust. Soc. Am vol 41, p793, 1967

8) M.D. Freedman, "A Method For Analysing Musical Tones," J Audio Eng. Soc., Vol 16, #4, p419, Oct. 1968

9) M.D. Freedman, "A Technique for Analysis of Musical Instrument Tones," PhD dissertation, University of Illinois, Urbana. 1965

10) J.S. Keeler, 'The Attack Transients of Some Organ Pipes," IEEE Tran. on Audio and Electroacoustics, Vol AU-20, #5, pp378-391, December 1972

11) J.S. Keeler, 'Piecewise-Periodic Analysis of Almost-Periodic Sounds and Musical Transients," IEEE Tran. on Audio and Electroacoustics, Vol AU-20, #5, pp338-344, December 1972

12) R.W. Hamming, 'Numerical Methods for Scientists and Engineers," New 'fork: McGraw-Hill, 1962

13) Charles M. Rader, Bernard Gold, "Digitat Filter Design Techniques in the Frequency Domain," Proc. IEEE, vol 55, #2, Feb. 1967, pp149-171

14) Bernard Gold, Charles M. Rader, "Digital Processing of Signals," McGraw-Hill Book Company, New York, 1969