ON ZERO-CROSSING ANALYSIS AND DESIGN OF SIGNAL PROCESSING HARDWARE

A Thesis
submitted in partial fulfilment of the
requirements for the degree of
DOCTOR OF PHILOSOPHY

By

Suresh V. Kibe

EEE GROUP
BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE
PILANI (Rajasthan)
JUNE, 1979

TO

MY PARENTS,

SARALA AND RATNAPRABHA

CERTIFICATE

This thesis is being submitted under Regulation 16.4 of the Academic Regulations for doctoral programmes which allows a faculty member of the Institute to do Ph.D research without the benefit of a supervisor.

I hereby certify that the thesis entitled 'On Zero-crossing Analysis and Design of Signal Processing Hardware' which I have submitted for award of Ph.D degree of the Institute embodies my original work.

Signature in full: _Suresh V. Kibe_

Name in Block Letters: SURESH V. KIBE

ID. No. 72E88006

Designation: Formerly, Lecturer in EEE Group.

Date: 19 June, 1979.

ACKNOWLEDGEMENTS

without the benefit of a Supervisor - a stupendous yet very rewarding experience.

Mr.Sadasivan and Mr.Chain Singh of ISRO HQ did a splendid job of typing and artwork for this thesis.

Suresh V. Kibe

ABSTRACT

In this thesis Stumper's basic of 'Zero-crossing FM Detection' has been extended so that noise in FM has been looked upon as the change in the position of the zero-crossings of an FM wave in the presence of Narrow-Band Gaussian noise or any interfering signal. The physical insight gained using this approach is exploited to derive the known formulae for calculation of Noise power due to spikes in the presence of additive Narrow-Band Gaussian noise using simple mathematics. Thus, this is an alternative approach to Rice's 'clicks' analysis and offers better physical insight into FM Threshold and Capture. It helps one to look at an FM demodulator as a system whose output is proportional to the axis-crossing rate of the input signal. To prove this, the important blocks of an IC PLL have been analysed.

In short, any application where the number of axis-crossings per second of a signal is used to determine the instantaneous frequency the FM demodulator could be used with advantage. This idea has been used to design a Formant Vocoder

...

which uses IC PLLs for real time Formant extraction from prefiltered speech signal. The hardware is simple.

An important spin-off of this study is that it helps one to visualize the use of digital building blocks not only in FM but other CW communication systems since the zero-crossing FM demodulator is essentially a simple digital system.

LIST OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES

CHAPTER - 1 : INTRODUCTION

Zero-crossing analysis as envisaged in this thesis deals with the estimation of the instantaneous frequency of a band limited signal by finding the number of axis crossings per second of that signal, within a certain time period. There have been independent attempts to use the zero-crossing analysis in various problems. Examples:

(1) Stumpers [1] as early as in 1948 demonstrated that all the information in a Frequency Modulated (FM) signal is contained in the points at which the FM signal crosses the origin, the so-called 'Zero-crossings'.

(2) One of the earliest attempts at automatic tracking of Formant frequencies in human speech signal was an average zero-crossing count [2].

(3) Dubnowski et.al. [3] have used an algorithm which uses centre clipping [level-crossing detection] and infinite peak clipping for a real-time Digital hardware pitch detector. Pollock and Licklider [4] have shown that Infinite peak clipping has no appreciable effect on the Intelligibility of speech.

(4) Recently, Zero-crossing analysis Techniques have been applied to Automatic Speech Recognition [5].

(5) Zero-crossing analysis studies have been used in many other areas e.g. fatigue studies in structures and loads, environmental and reliability testing, and analysis of flight test data and guidance systems [6].

This thesis attempts to analyze a class of systems in Communications and Speech Processing in an integrated manner. It is seen that with this approach it is possible to see similarities in these various attempts and exploit them to understand these systems better.

In general, the calculation of the number of zeros per second and the variation of the distance between consecutive zeros is a difficult mathematical problem when random signals are involved [7]. However, in narrow-band processes and signals both the instantaneous frequency and amplitude variations are small within an interval which is governed by the bandwidth and hence definite guidelines in the design of systems used to demodulate such signals are possible. The zero-crossing analysis is particularly helpful in giving better physical insight than is possible with the vectorial representation in explaining Noise in FM systems.

The choice of the 'zero' level for analysis is natural since most waveforms after removal of the d.c. component have zero average value. All the same, it has been shown that under certain conditions a level crossing detector might prove more useful. Incidently, in the detection of binary signals in additive Gaussian Noise, where the '1's and '0's are equally likely, the optimum decision threshold is taken to be the zero level.

In this thesis, systems used to process a class of signals whose instantaneous frequency varies over a definite range as a function of time have been analysed using the zero-crossing analysis approach. Chapter 2 deals essentially with explaining the phenomenon of FM Threshold. It is seen that this approach gives a much better insight into the phenomenon of FM Threshold than is possible with Rice's 'clicks' analysis and hence can be used to design Threshold Extension Systems better. In chapter 3 a class of hardware used in Speech Analysis and Synthesis systems has been analysed using the zero-crossing approach and it has been shown that Formant Tracking is possible using a PLL. To prove this, it has been shown that the IC PLL essentially is a device whose output is proportional to the axis crossing rate of the input signal. In chapter 4, a Formant Vocoder has been designed based on the analysis carried out in chapter 2 and 3.

An important spin-off of this study is that it helps one to visualize the use of zero-crossing detector a simple digital IC, not only in FM demodulation but other analog systems. With the cost of a digital IC going down simple and economic systems can be conceived using this approach.

It is believed that the zero-crossing analysis of FM Threshold and the consequent design of a Formant Vocoder using a PLL presented here is a new approach.

This approach needs to be further substantiated by both theoretical as well as practical work to cover all aspects of noise in FM systems.

CHAPTER 2: ZERO-CROSSING ANALYSIS OF FM THRESHOLD

In this chapter the phenomenon of FM Threshold has been analysed using the zero-crossing approach. Stumpers' basic idea of 'Zero-crossing FM Detection' has been extended so that noise in FM has been looked upon as the change in the position of the zero-crossings of an FM wave in the presence of Narrow-Band Guassian noise or any interfering signal. The physical insight gained using this approach is exploited to derive know formulae for calculation of Noise power due to spikes in the presence of additive Narrow-Band Gaussian noise using simple mathematics. Thus, this is an alternative approach to Rice's 'clicks' analysis [8] and offers better physical insight into FM Threshold and Capture. Simple explanations have been offered to the observed experimental results in FM receivers under threshold such as:

1. Occurence of positive and negative spikes at the output of the FM Receiver.

2. Increase in the number of negative spikes at the positive extreme of modulation (highest instantaneous frequency) and vice-versa.

3. Occurence of threshold at higher values of input SNR for higher deviation systems.

4. Enhanced increase in the number of spikes near and below threshold when the carrier is sinusoidally modulated than when it is unmodulated.

The positive spike has been likened to 'Extra' and the negative spike to 'Missing' pair of zeros in FM signal in the presence of Narrow Band Gaussian noise. This approach offers a clue to the time duration, frequency and the nature of the spike which are helpful in the design of Threshold Extension Devices (TEDS).

In the analysis presented only instantaneous values have been considered. No effort is made to justify the input SNR at which threshold sets-in.

2.1. REVIEW OF PREVIOUS WORK

Stumpers has shown that there is a direct relationship between the instantaneous frequency and the number of zero-crossings of an FM signal [see appendix A]. Leentvaar and Flint [9] have done an extensive analysis of FM capture. Some relevant part of their analysis is reproduced here.

Consider two signals at the input of the limiter of an FM receiver

$$v_1(t) = A_1 \cos \varphi_1(t) = \operatorname{Re} \left[A_1 e^{j\varphi_1(t)} \right] \qquad .. \ (2.1.1)$$

$$v_2(t) = A_2 \cos \varphi_2(t) = \operatorname{Re} \left[A_2 e^{j\varphi_2(t)} \right] \qquad .. \ (2.1.2)$$

Where $\quad \varphi_2(t) - \varphi_1(t) = \varphi(t)$

$$A_2/A_1 = A \qquad\qquad .. \ (2.1.3)$$

$$\omega_1 = \frac{d\varphi_1}{dt}$$

$$\omega_2 = \frac{d\varphi_2}{dt}$$

$$\omega = \omega_2 - \omega_1 \qquad\qquad \text{.. (2.1.4)}$$

The amplitude of the resulting signal R from Fig.1
is given by

$$R = \sqrt{A_1^2 + A_2^2 + 2A_1A_2 \cos \varphi} \qquad \text{.. (2.1.5)}$$

The phase angle of the resultant is given by

$$\psi = \varphi_1 + \arctan \frac{A \sin \varphi}{1+A \cos \varphi} \qquad \text{.. (2.1.6)}$$

The instantaneous frequency of the limiter output
is given by

$$\omega_r = \frac{d\psi}{dt} = \frac{d\varphi_1}{dt} + \frac{A^2 + A \cos \varphi}{1 + A^2 + 2A \cos \varphi} \cdot \frac{d\varphi}{dt}$$

$$= \omega_1 + \frac{A^2 + A \cos \varphi}{1 + A^2 + 2A \cos \varphi} (\omega_2 - \omega_1)$$

$$= \omega_1 + x (\omega_2 - \omega_1) \qquad\qquad \text{.. (2.1.7)}$$

Where
$$x = \frac{A^2 + A \cos \varphi}{1 + A^2 + 2A \cos \varphi}$$
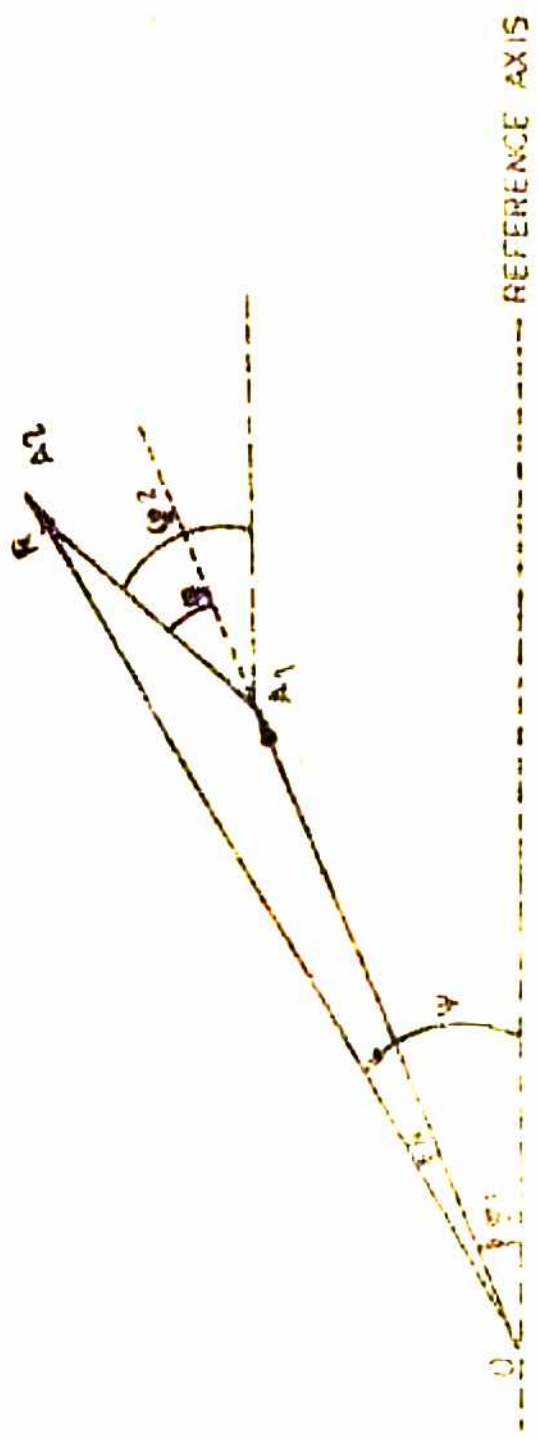
FIG. 1. PHASOR DIAGRAM OF TWO SUPERIMPOSED SIGNALS U1 AND U2

Fig.2(a) shows the variation of the function x with the instantaneous phase difference φ and Fig.2(b) shows the variation of the instantaneous frequency of the resultant. It has been shown by Leentvaar and Flint that for

$$A > 1 \quad W_{r\ mean} = W_2 \quad \text{and for} \quad A < 1 \quad W_{r\ mean} = W_1$$
$$.. (2.1.8)$$

Figures 2(a) and (b) convey the following important information:

(1) The mean frequency of the output of the limiter is equal to the frequency of the stronger signal. A zero-crossing count of the output signal will yield the instantaneous frequency of the stronger signal (Capture).

(2) There is a sudden change in the function x and hence in the instantaneous frequency of the resultant when the instantaneous phase difference between the two signals is π. These sudden changes (peaks) occur with a frequency $|W_2 - W_1|$ the beat frequency. The function x changes sign at $A = 1$.

(3) The peaks always 'point away' from the frequency of the weaker signal.

(4) The peaks are smaller when the frequency difference between W_1 and W_2 is small.

However, it is difficult to visualize an infinite frequency jump when A=1. What does it correspond to physically?. This question is answered in the next section.
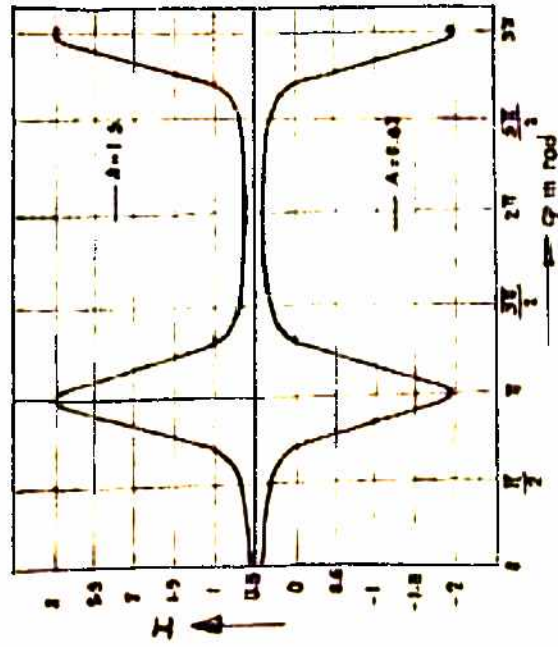
(a) The function $I = (A^2 + A\cos\phi)/(1 + A^2 + 2A\cos\phi)$ as a function of phase difference ϕ of the two input signals of the limiter

$A = A_2/A_1$ is the amplitude ratio of these two signals
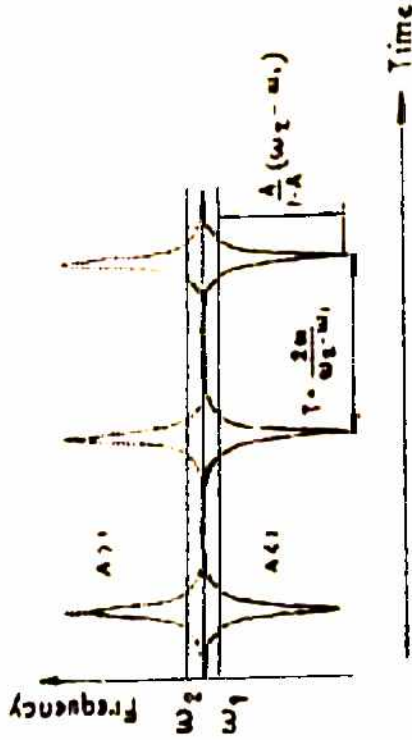
(b) Resultant frequency of the limiter output signal as a function of time when two signals with frequencies ω_1 and ω_2 are supplied to the limiter input. Amplitude ratio A of signals may be <1 (drawn graph) or >1 (dotted graph) ω_1 and ω_2 instantaneously constant

FIG. 2. VARIATION OF INSTANTANEOUS FREQUENCY OF THE RESULTANT IN FM CAPTURE

S.O.Rice in his monumental work on noise in
FM Receivers used the vectorial approach and explained
the phenomenon of FM Threshold using the famous heuristic
'clicks' analysis. Rice's conjecture of FM Threshold
is shown in Fig.3. In Fig.3 R(t) is the resultant
vector. For simplicity, the carrier is assumed to be
unmodulated and is given by

$$v(t) = A_s \cos \omega_c t \qquad ..(2.1.9)$$

Narrow-Band noise n(t) is represented by

$$n(t) = x(t) \cos \omega_c t - y(t) \sin \omega_c t \qquad ..(2.1.10)$$

$$= r(t) \cos (\omega_c t + \varphi_n) \qquad ..(2.1.11)$$

Where,

$$r(t) = \sqrt{x^2(t) + y^2(t)}$$

$$\varphi_n = \arctan \frac{y(t)}{x(t)} \qquad ..(2.1.12)$$

x(t) and y(t) are gaussian random variable with zero
mean and deviation ηB where η is spectual density of
white noise and B is the bandwidth of the IF filter.
Schilling [10] after an involved calculation [see
appendix B] has shown that the total number of spikes
occuring per second in the presence of carrier alone
(no modulation) is given by

$$N_c = N_- + N_+ = 2N_- = 2N_+ = \frac{B}{2\sqrt{3}} \, \text{erf}_c \sqrt{\frac{f_M}{B} \cdot \frac{S_i}{N_M}}$$

$$..(2.1.13)$$

Path of R(t)

$t > t_2$ R(t)   r(t) $t_2 + \Delta t/2$

$t = t_2$ $\theta(t)$ ΔS
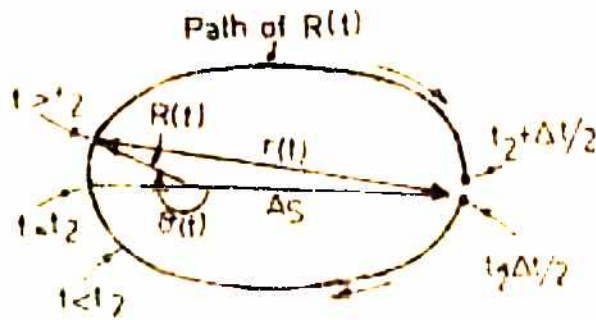
$t < t_2$ $t_2 \Delta t/2$

(A) LOCUS OF R (t) AND Θ(t) TO CAUSE A
NEGATIVE SPIKE BETWEEN t2 AND t2+Δt

$$\text{AREA} \quad \int_{t_2}^{t_2+\Delta t} \frac{d\theta}{dt} \cdot dt = \theta \Big|_{t_2}^{t_2+\Delta t} = 2\pi$$

θ(t)

2π

0   t2   t2 + Δt   t

(B)

$\frac{d\theta}{dt}$

AREA-2π   (C)

0   t

FIG.3. RICE'S CONJECTURE OF FM
THRESHOLD
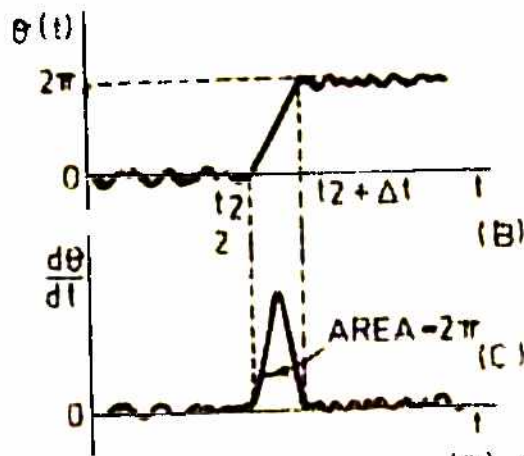
(B) A PLOT OF Θ (t) FOR A CASE
IN WHICH THE END POINT OF R(t)
IN FIG (A) EXECUTES A ROTATION
AROUND THE ORIGIN (C) A PLOT
OF dθ/dt AS A FUNCTION OF
TIME

WHITE-
NOISE

n(t)

Filter
H(f)   n(t)

2/B

|H(f)|

B   B

1/fc

fc   fc   f
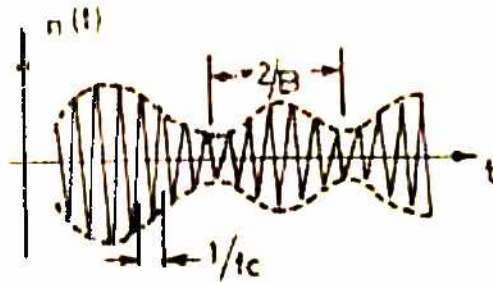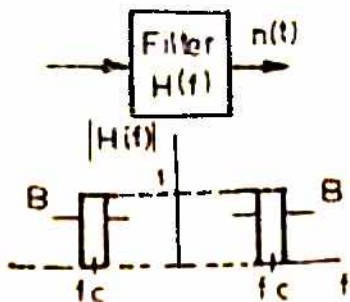
NARROWBAND NOISE

FIG.4. RESPONSE OF A NARROW BAND FILTER TO NOISE

where $N_M = \eta f_M$, $\quad S_i = \dfrac{A_s^2}{2}$ – input signal power

f_M – cut-off frequency of the base band filter at the output.

Schilling has also shown that in the presence of modulation, the average number of spikes per second increases by a factor given by

$$\overline{\delta N} = \frac{2\Delta f}{\pi}\, e^{-\left(\frac{f_M}{R}\right)} \times \left(\frac{S_i}{N_M}\right) \qquad \text{.. (2.1.14)}$$

where Δf = maximum frequency deviation of the carrier.

Near and below threshold $\overline{\delta N} \gg N_c$ therefore the total average number of spikes is $N \cong \overline{\delta N}$. Equation(2.1.14) is used to account for the noise due to spike in FM Receivers.

In this analysis some questions remains unanswered. What is the bound on the time interval Δt in which a spike occurs? What is the physical equivalent of a positive and negative spike? It is assumed that to ensure a spike output it is not actually necessary to observe a complete rotation of $R(t)$, but it is adequate that there be guaranteed a rotation of at least π radians. Why? All of these questions are answered in the next section.

2.2. ZERO-CROSSING ANALYSIS OF FM THRESHOLD

It is instructive to look at the phenomenon of Narrow-Band noise from the zero-crossings point of view. Consider the response of a Narrow-Band filter (ideally a notch-filter) to noise. It is found that the output has the appearance shown in Fig.4. If the filter were an ideal notch-filter the output would have been a sinusoid. Then, the distance between the zero-crossings would be fixed. As the band-width of the filter is increased slightly, one notices a fluctuation both in the amplitude of the envelope and also the distance between the zero-crossings. The output $n(t)$ is represented by equation (2.1.11) where $r(t)$ is Rayleigh distributed with the probability density function given by

$$f_r(r) = \begin{cases} re^{-r^2/2\eta B} & r \geqslant 0 \\ 0 & r < 0 \end{cases}$$

$$\cdots (2.2.1)$$

and $\varphi_n(t)$ is uniformly distributed with the probability density function

$$f_\Theta(\Theta) = \frac{1}{2\pi} \qquad -\pi \leqslant \Theta \leqslant \pi \qquad \cdots (2.2.2)$$

If the centre frequency of the filter is f_c the distance between two consecutive zero-crossings is

$$\frac{1}{2f_c} \qquad \text{when } B = 0$$

and within

$$\frac{1}{} \qquad \text{when B is fi...}$$

Any particular value of the envelope holds only
for 2/B seconds. In other words, the envelope can be
assumed to have a particular value for 2/B seconds
[Ref.11 and Art.10.7 of Ref.7]. As the bandwidth is
further relaxed the distance between zero-crossings
becomes more varied and the envelope fluctuates more
till when there is no filter the separation between
zero-crossings ranges between 0 and ∞ and the amplitude
fluctuates randomly. Thus, Narrow-Band noise if viewed
within an interval of 2/B seconds looks more or less
like a sinusoid. This representation of Narrow-Band
noise has been used in Fig.5 which gives a heuristic
model for FM Threshold based on the zero-crossing
approach. Fig.5 gives the following information.

(1) When $r < A_s$ and the noise frequency is greater
    or less than the signal frequency there is no
    change in the number of zero-crossings of the
    resultant signal.

(2) When $r > A_s$ for $W_n > W_s$ the resultant signal
    has an extra pair of zeros and for $W_n < W_s$
    a pair of zeros is missing.

(3) When $r = A_s$ the resultant signal in both
    the cases just touches the zero amplitude
    axis thus giving rise to a pair of zeros
    infinitesimally close to one another which
    corresponds to the infinite frequency case
    when A=1 in Fig.2.

AMPLITUDE

TIME

$\pi \; < \; A_s$

$\omega_n \; > \; \omega_s$

AMPLITUDE

$\pi \; > \; A_s$

$\omega_n \; > \; \omega_s$

AMPLITUDE

TIME

$\pi \; < \; A_s$

$\omega_n \; < \; \omega_s$

AMPLITUDE

$\pi \; > \; A_s$

$\omega_n \; < \; \omega_s$
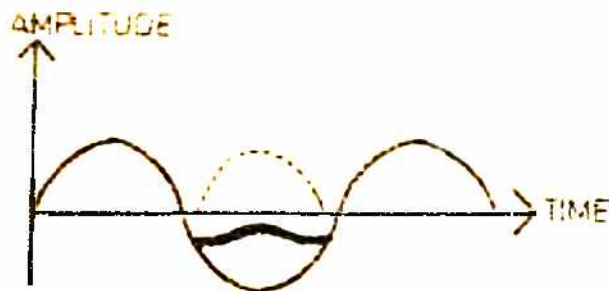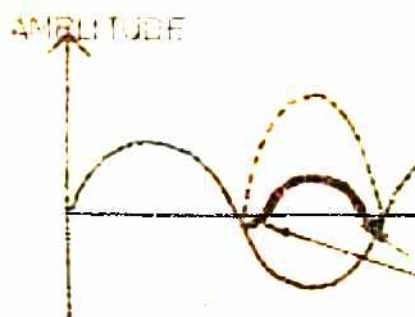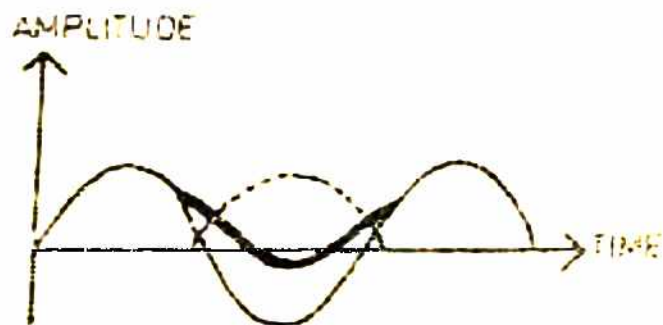
NO SPIKE
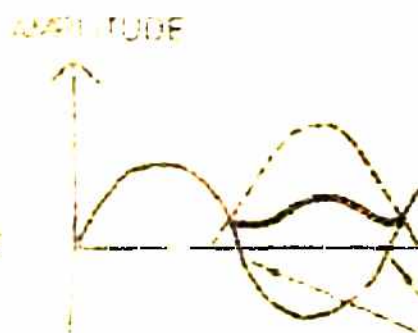
SPIKE

FIG.5. ZERO CROSSING MODEL OF

EXTRA ZEROS

- - - - NOISE
───── SIGNAL
━━━━━ RESULTANT

$n = As$

$\omega_n > \omega_s$

ZERO MISSED

AMPLITUDE

$n = As$

$\omega_n < \omega_s$
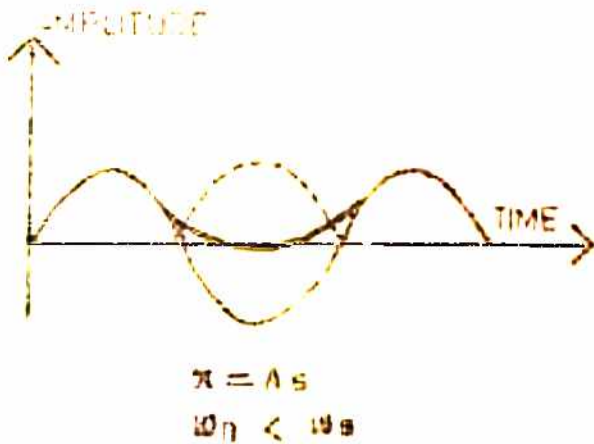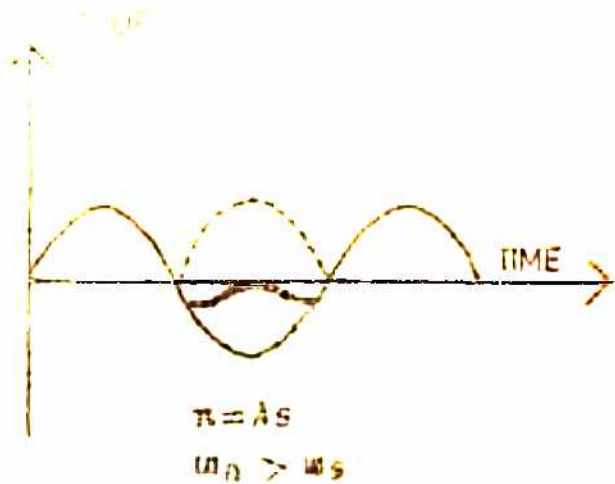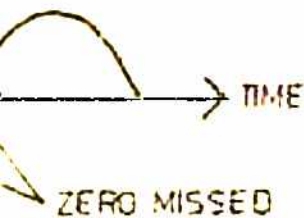
∞ SPIKE

F FM THRESHOLD

It may be noted that only the case when the signal
and noise have a phase difference of π has been considered
as frequency jumps are encountered only then. It is
instructive to study the position of the zero-crossings
for phase differences other than π. Fig.6(a) depicts
one such case when the phase difference is $\pi/2$. For
simplicity, the two frequencies are taken equal. It can
be seen in Fig.6(a) that the position of the zero-
crossings of the resultant is more close to the position
of the zero-crossings of the stronger signal but there
is no change in the number of zero-crossings. When the
amplitudes of the two signals are equal the position
of the zero-crossings of the resultant is mid-way between
those of the noise and signal. It is only when the
relative amplitude of the strong signal is very high
that the zero-crossings of the resultant move close to
its zero-crossings meaning that there is a very gradual
shift in the position of the zero-crossings when the
phase differences are not π. This fact is depicted
in the very slow variation of the function x with A
for phase differences other than π in Fig.2(a).

Thus, when $r \geqslant A_s$ (this condition corresponds to
an input SNR of 10 to 11 dB normally) and when the
instantaneous phase difference between signal and
noise is π a pair of zeros is either missed (-ve spike)
or an extra pair of zeros (+ve spike) is generated.
This is the onset of threshold. It is easy to visualise

(a) ZERO CROSSING CLOSE TO THOSE OF THE STRONGER SIGNAL
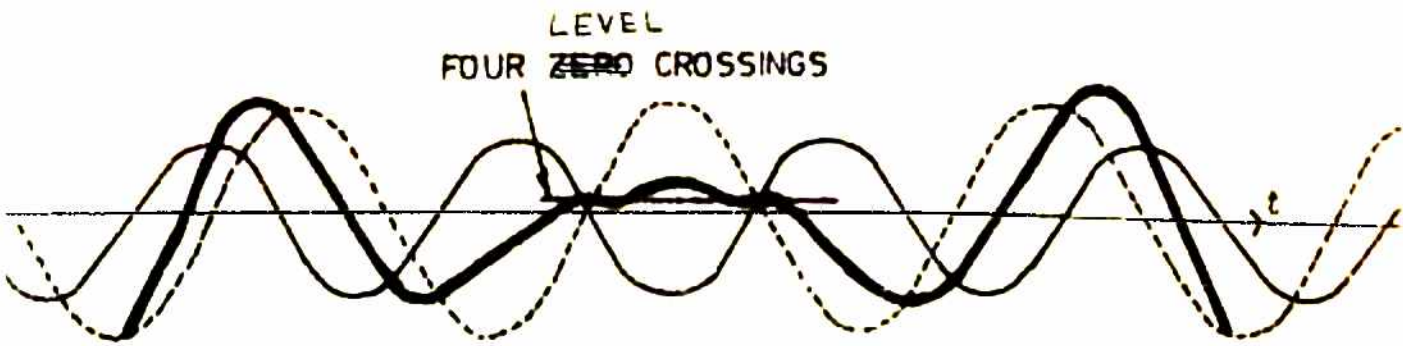
LEVEL
FOUR ZERO CROSSINGS

(b)

------- NOISE
——— SIGNAL
━━━━ RESULTANT
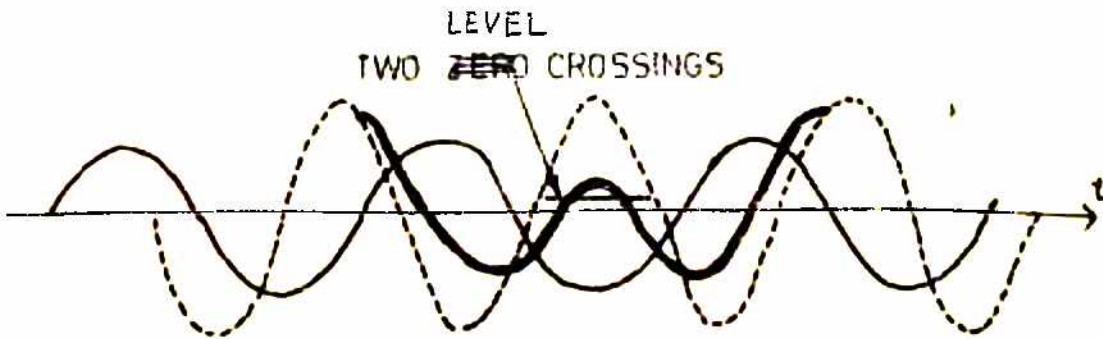
LEVEL
TWO ZERO CROSSINGS

(c)

FIG.6 . DETAILS OF ZERO CROSSING MODEL FOR FM THRESHOLD

that there are FOUR ways in which a spike (+ve or -ve) can occur. TWO are shown in Fig.5 (corresponding to $r > A_s$ case) and two others are the duals of these.

A missing pair of zeros corresponds to a phase change of -2π. An extra pair of zeros corresponds to a $+2\pi$ phase change. The area under the spike is therefore always 2π.

Having noted the condition under which a spike occurs it is easy to find the number of spikes per second. Since the noise envelope $r(t)$ is Rayleigh distributed probability that it is greater than A_s ($r > A_s$)

$$= e^{-\frac{A_s^2}{2\eta B}}$$

$$= e^{-\frac{f_M}{B} \times \frac{S_i}{\eta f_M}}$$

$$= e^{-\left(\frac{f_M}{B}\right) \times \left(\frac{S_i}{N_M}\right)} \qquad \qquad \text{.. (2.2.4)}$$

Where $S_i = \frac{A_s^2}{2}$ is the input Signal Power.

Since the phase $\varphi_n(t)$ of noise is uniformly distributed the probability that the phase of the noise waveform is just right to cause a spike is $\frac{1}{2\pi}$..

Therefore, the joint probability P that the condition for a spike is satisfied is

$$P = \frac{1}{2\pi} \times e^{-\left(\frac{f_M}{B}\right) \times \left(\frac{S_i}{N_M}\right)} \qquad \qquad \text{.. (2.2.5)}$$

Now, there are four different ways in which this can occur hence, the probability P_S that either a +ve or -ve spike is encountered is

$$P_s = \frac{4}{1} \times \frac{1}{2\pi} \, e^{-\left(\frac{f_M}{B}\right) \times \left(\frac{S_i}{N_M}\right)} \qquad \text{.. (2.2.6)}$$

Equation (2.2.6) is the probability of a spike in any 2/B seconds, the time for which the amplitude of the noise envelope remains quasi constant.

∴   Probability of occurence of a spike per unit time is

$$\frac{4 \times \dfrac{1}{2\pi} \times e^{-\left(\frac{f_M}{B}\right) \times \left(\frac{S_i}{N_M}\right)}}{2/B}$$

$$= \frac{B}{\pi} \times e^{-\left(\frac{f_M}{B}\right) \times \left(\frac{S_i}{N_M}\right)} \qquad \text{.. (2.2.7)}$$

Now B, the r.f. bandwidth of a sinusoidally modulated FM signal is given by

$$B = 2(\Delta f + fm) \simeq 2\Delta f \qquad \text{for large } \Delta f \quad (2.2.8)$$

Δf - maximum freq deviation

fm - frequency of the modulation signal.   -

From (2.2.7) and (2.2.8) we have probability of occurence of a spike per unit time $\overline{\delta N}$ as

$$\overline{\delta N} = \frac{2 \triangle f}{\pi} \, e^{-\left(\frac{f_M}{B}\right) \times \left(\frac{S_i}{N_M}\right)} \qquad \qquad .. \,(2.2.9)$$

The average time between spikes is then

$$T_s = \frac{1}{N} = \frac{1}{\overline{\delta N}} \qquad \qquad .. \,(2.2.10)$$

It may be noted that in Fig.5 the noise waveform has been shown/only for half-wavelengths and where the peaks of the noise and signal are opposite in phase. It can be easily shown that at other instants there is only a change in the position of the zero-crossings but a pair of zeros is neither missed nor an extra pair of zeros added. [see Fig.6]. It is seen in Fig.6(b) that the zero-crossings of the resultant are close to those of the noise waveform. Since the frequency of the noise waveform is less than that of the signal the number of zero-crossings would also be less. Hence, the number of zero-crossings of the resultant are also less than those of the signal and this transition takes place when the phase difference between the signal and noise is π. Thus, at that point a pair of zeros is missed.

Equation (2.2.9) for the number of spikes per second is the same as that obtained by Schilling.

The zero-crossing analysis offers the following simple explanations for other observed facts as well.

(i) <u>Earlier onset of Threshold for large deviation systems</u>

As the deviation of the system is increased, the r.f. bandwidth following Carson's rule increases correspondingly. Thus, the demodulation r.f. bandwidth increases accordingly. As a result of the increased bandwidth of the IF filter the noise envelope has larger fluctuations and hence the condition of spike $(r \geqslant A_s)$ is reached at higher values of input signal.

(ii) <u>-ve spikes at +ve extreme of modulation and vice versa</u>

It is clear from Fig.5(a) and (b) that a -ve spike (missing pair of zeros) occurs when $\mathrm{W}_n < \mathrm{W}_s$ and a +ve spike occurs (extra pair of zeros) when $\mathrm{W}_n > \mathrm{W}_s$. Thus at the +ve extreme of modulation, the frequency of noise W_n will always be less than the instantaneous frequency of the signal hence only -ve spikes will occur. The opposite will be the case at the -ve extreme of modulation.

(iii) <u>Enhanced increase in the number of spikes under modulation</u>

It has been shown in Fig.2(a) and (b) that the instantaneous frequency of the resultant experiences a sudden jump when $\varphi = \pi$ and that the frequency of occurence of such spikes is the absolute difference

in the instantaneous frequencies of signal and noise.
When the signal frequency occupies the centre of the
band (unmodulated carrier) the absolute difference in
the frequencies of signal and noise is the least.
Under modulation, the instantaneous frequency of the
signal shifts to +ve and -ve extremes increasing the
absolute difference in the frequencies of signal and
noise. Thus, under any modulation, the frequency of
spikes will increase.

2.3. APPLICATIONS

It is clear from the analysis presented in the
earlier section that the zero-crossing analysis gives
a definite clue to the time duration of the spike. It
is clear from figures 5 and 6 that the duration of the
spike is roughly $\frac{1}{2f_n}$ where f_n is the instantaneous
frequency of the noise or interfering signal. A closer
look at Fig.6 reveals that when the instantaneous phase
difference is exactly π, a pair of zeros is missed and
hence there is no zero-crossing in an interval $\frac{1}{2f_n}$.
The longest such interval clearly is the one which
corresponds to the lowest frequency of the noise waveform
or

$$\frac{1}{2\left[f_c - \frac{B}{2}\right]} \qquad\qquad ..\,(2.3.1)$$

Similarly, it can be shown that when the condition
for positive spike is satisfied the shortest interval
in which an extra pair of zeros is encountered is given by

$$\frac{1}{2[f_c + \frac{B}{2}]} \qquad \qquad .. (2.3.2)$$

The amount of frequency jump depends upon the instantaneous
frequency difference and the amplitude ratio of signal
and noise.   This information is of vital importance in
the design of Threshold Extension Devices (TEDs)
(cf, Ref [14] pp 348-351 for detailed discussions).

A zero-crossing FM demodulator with threshold
extension capability would therefore be an accurate
zero-crossing counter which is capable of ignoring
sudden changes in the zero-crossing count in the
interval given by eqn. (2.3.1) [see Fig.7]. A closer
look at Figures 6(b) and 6(c) reveals that the
detection of Missing or Extra pair of zeros is possible
since during spike formation definite patterns of the
resultant waveform are observed in the two conditions.
An appropriate level detector (where the decision level
is dependent on the Amplitude ratios) in conjunction
with the zero-crossing detector can be used to identify
the patterns. It is seen that FOUR positive-level
crossings are seen in Fig.6(b) while only TWO are present
at that level in Fig.6(c). It can be shown that such
patterns would be observed on the negative side when

FM INPUT → **ACCURATE ZERO CROSSING DETECTOR** → PULSES AT ZERO-CROSSINGS → **COUNTER** → **DEMODULATED OUTPUT**

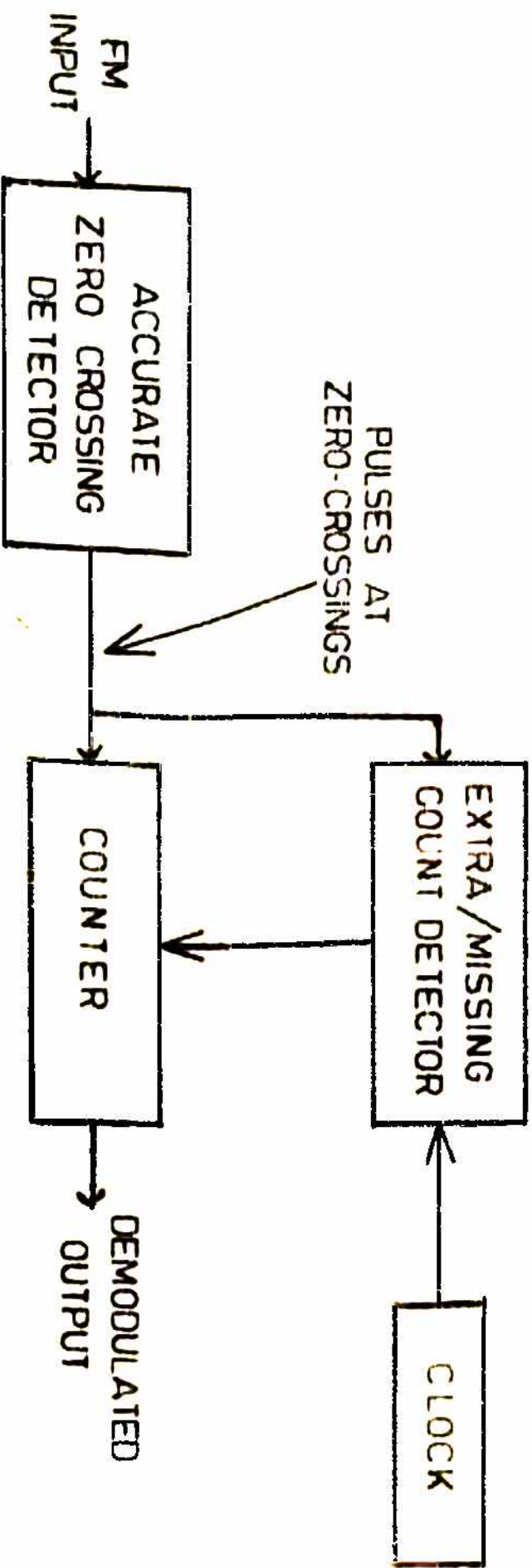**EXTRA/MISSING COUNT DETECTOR** ← **CLOCK**

FIG.7. MODEL OF FM DETECTOR WITH THRESHOLD EXTENSION

the condition for a spike is fulfilled on the negative peaks of the noise waveform. Any FM demodulator may be looked upon as a device whose output is proportional to the variation of the zero-crossing distance. The zero-crossing detector is essentially a device which can be made into a digital IC.

The 'Zero-crossing' approach is particularly useful for same or adjacent channel interference since the concepts of instantaneous frequency difference and amplitude ratios of the signal and the interfering signal (noise) are involved in the analysis. Once the condition for the onset of threshold is met, interference from a signal whose instantaneous frequency is farthest (and yet within the I.F. bandwidth) from the instantaneous frequency of the signal could be more harmful as the frequency of spikes in that case will be higher.

2.4 CONCLUSIONS:
It is seen that the Zero-crossing Analysis of FM threshold has the following advantages.

(1) The analysis is simple.
(2) It offers a physical insight into the phenomenon of spike formation.
(3) It offers definite clue to the duration of a spike which is essential to the design of threshold extension devices.
(4) This analysis may be extended to same of adjacent channel interference calculations.

CHAPTER 3: ZERO-CROSSING ANALYSIS OF SOME SIGNAL PROCESSING HARDWARE

It has been shown in Chapter 2 that there is a direct relationship between the average zero-crossing count and the instantaneous frequency of an FM wave and that a convincing explanation based on the zero-crossing approach is possible for various other observed facts in FM systems. In this chapter the phase detector used in some commercial IC PLLs has been analysed. It has been shown that the IC PLL can be used as a device which essentially uses the zero-crossing information of the input signal. A class of hardware used for measurement of short-time spectrum of human speech has been analysed using the zero-crossing approach. It has then been shown that real-time estimation of Formant frequencies of human speech is possible using the IC PLLs, as zero-crossing information is of vital importance in Formant extraction. Definite guidelines for the design of a Formant Vocoder using an IC PLL have been evolved.

3.1. ZERO-CROSSING ANALYSIS OF LINEAR PHASE DETECTORS

Consider the Amplitude Modulated (AM) signal demodulation using a phase lock loop (PLL) [12] shown in Fig.8. The PLL is used for carrier recovery. It may be noted that in an AM signal with modulation index $m < 1$ the zero-crossings correspond to those of the carrier. It is clear from Fig.8 that the PLL uses the
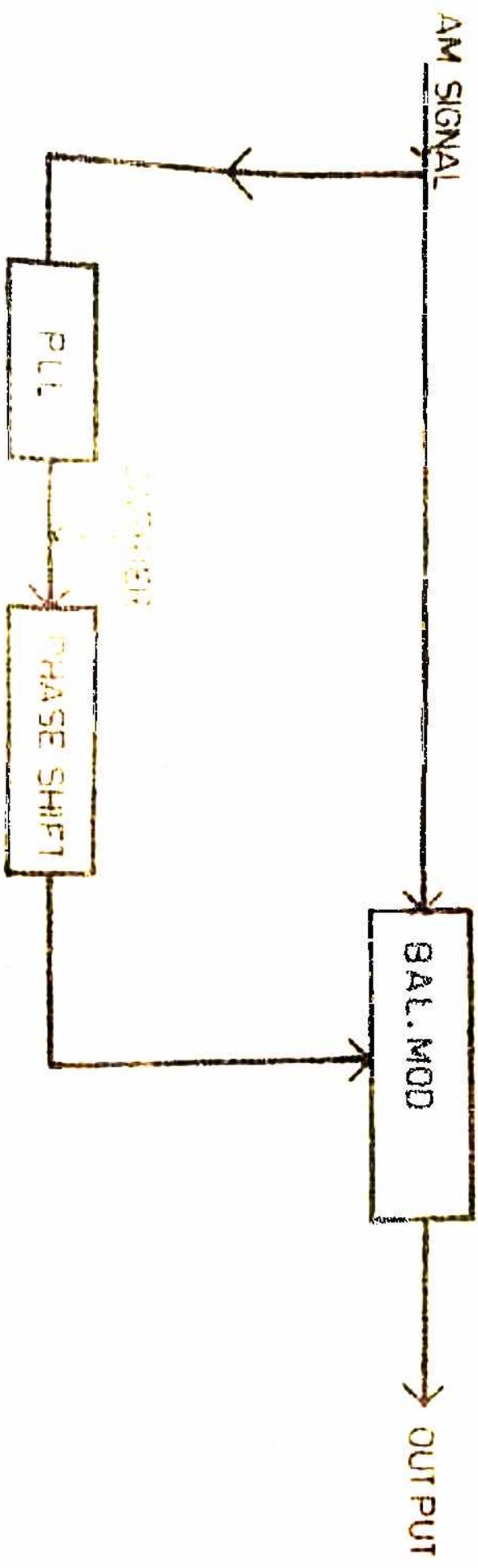
FIG.3. AM DEMODULATION USING PLL

zero-crossing information for carrier extraction. In fact, all linear phase detectors use the zero-crossing information. Fig.9(a) and (b) show the block schematic of a linear phase detector, with linear characteristics. Fig.9(c) is the timing diagram for Fig.9(b). The phase detector shown in Fig.9(a) is the familiar phase detector which is linear only for small values of Θ (since sin $\Theta=\Theta$ for Θ small). Fig.9(b) shows a digital linear phase detector [13] which is equivalent to the one shown in Fig.9(a). It is easy to prove from the timing diagram shown in Fig.9(c) that the transfer characteristics of this phase detector will be linear. In fact, most analogue phase detectors which are used in linear ICs use the same basic principle as that of the digital phase detector of Fig.9(b). It is clear from the timing diagram (Fig.9(c)) that these phase detectors use the zero-crossing information of the two input waveforms. Most PLL ICs use the phase detector shown in Fig.10(b) [14] which is an improved version of the basic circuit shown in Fig.10(a). It is clear from Fig.10(a) that I_e flows only when V_1 is positive and I_4 flows only when V_1 is positive and V_2 is negative - an enclusive OR operation. Similarly, in the Full-Wave coincidence detector I_{e1} flows only when V_1 is positive and I_{e2} when V_1 is negative. I_6 flows only when V_2 is positive and V_1 is negative. I_6 thus flows only after V_2 crosses zero with a positive slope and goes to zero when V_1 crosses zero with a positive slope. I_5 contains information about the

$$\frac{AB}{2} \sin(2w_c t + \theta_1 + \theta_2) + \frac{AB}{2} \sin(\theta_1 - \theta_2)$$

A sin $(w_c t + \theta_1)$

B cos $(w_c t + \theta_2)$

MULTIPLIER

LPF

C sin θ

(a)

v_A

v_B

SCHMITT TRIGGER

A

SCHMITT TRIGGER

B

XOR

C

INTEGRA-TOR

V_0

(b)

v_A

A

v_B

B

V_0

C

V_0

$-\pi/2$

$\pi/2$

π

φ

(c)
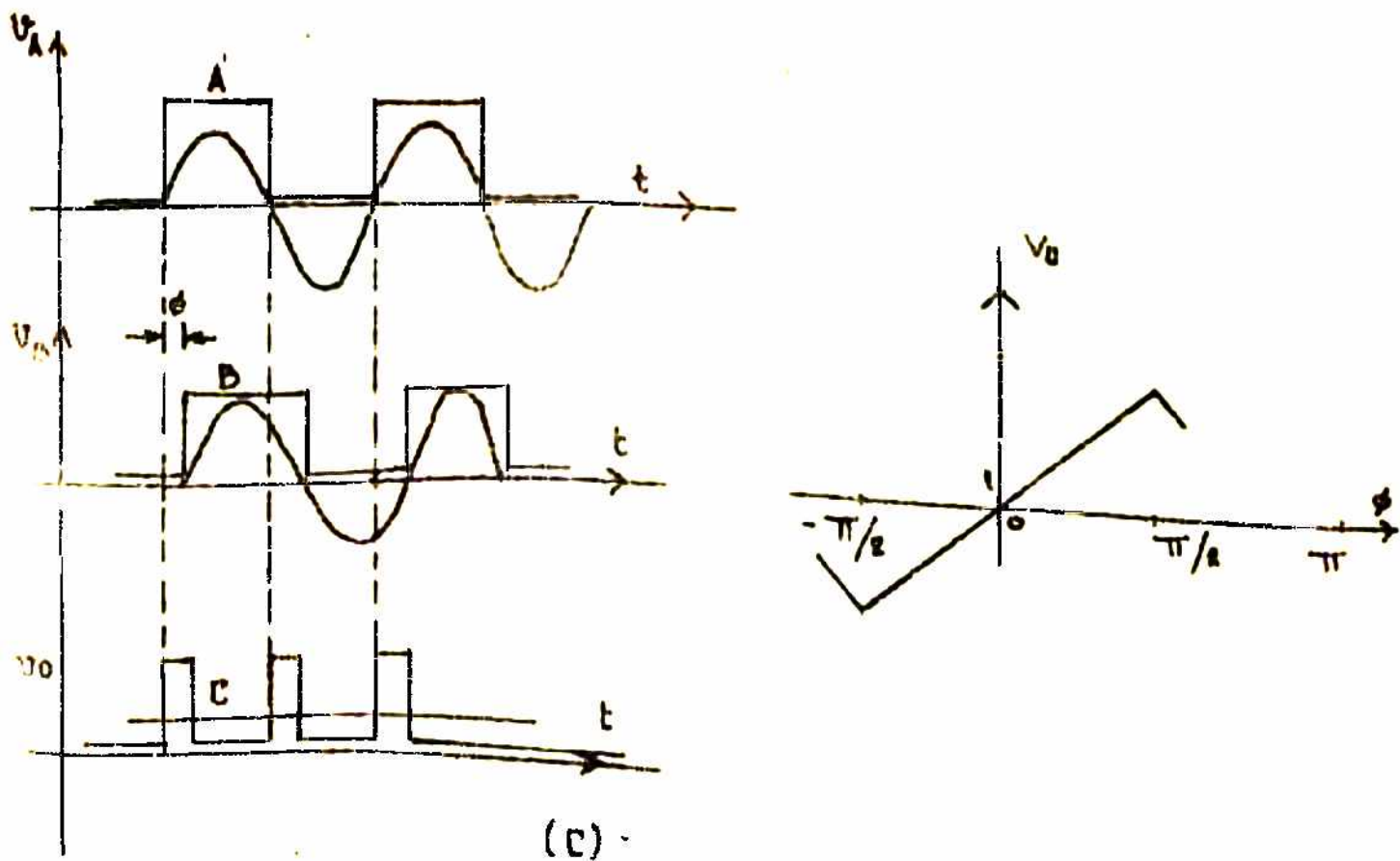
FIG.9. BLOCK SCHEMATIC OF LINEAR PHASE DETECTORS

(a) SINGLE BALANCED COINCIDENCE DETECTOR
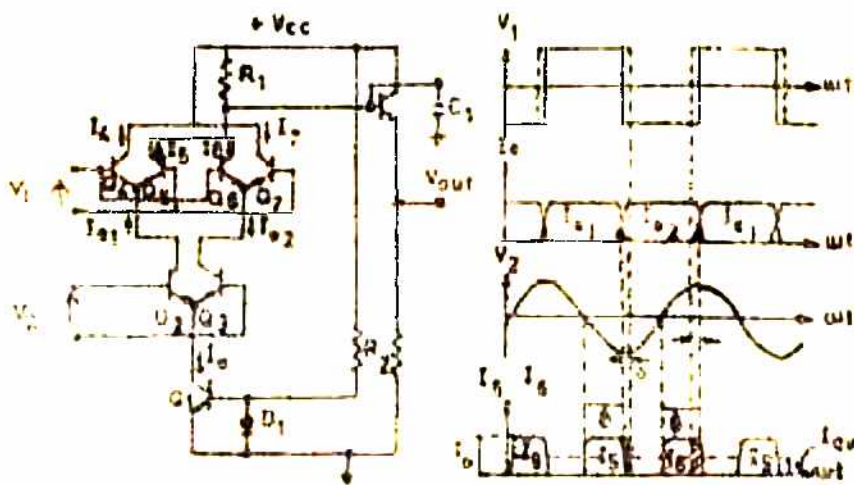
(b) FULL WAVE COINCIDENCE DETECTOR

FIG. 10. IC PHASE DETECTORS

negative zero-crossings of V_1 and V_2. Under the condition of no phase jitter in either V_1 or V_2 the width of the current pulses I_5 and I_6 corresponds to the distance between the negative and positive zero-crossings of V_1 and V_2 respectively. Also,

$$V_C \cdot I_{av} = \frac{I_0}{2\pi} \cdot \varphi \qquad \qquad \text{.. (3.1.1)}$$

where, I_0 is the pulse peak current and φ is the phase difference in radians. This therefore is a linear phase detector which involves the zero-crossing information of the inputs V_1 a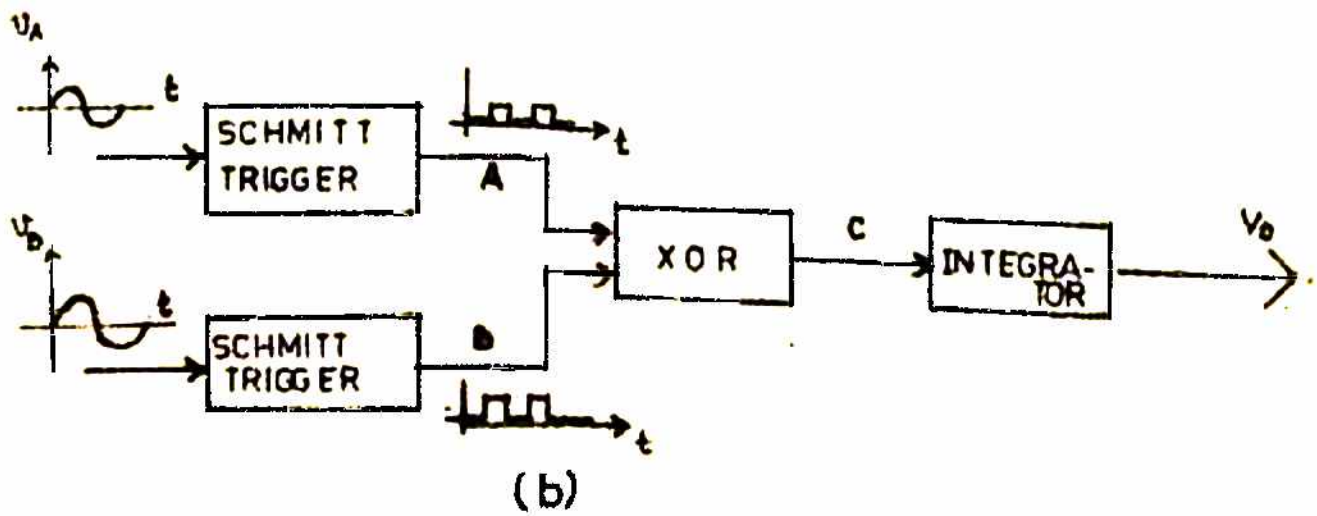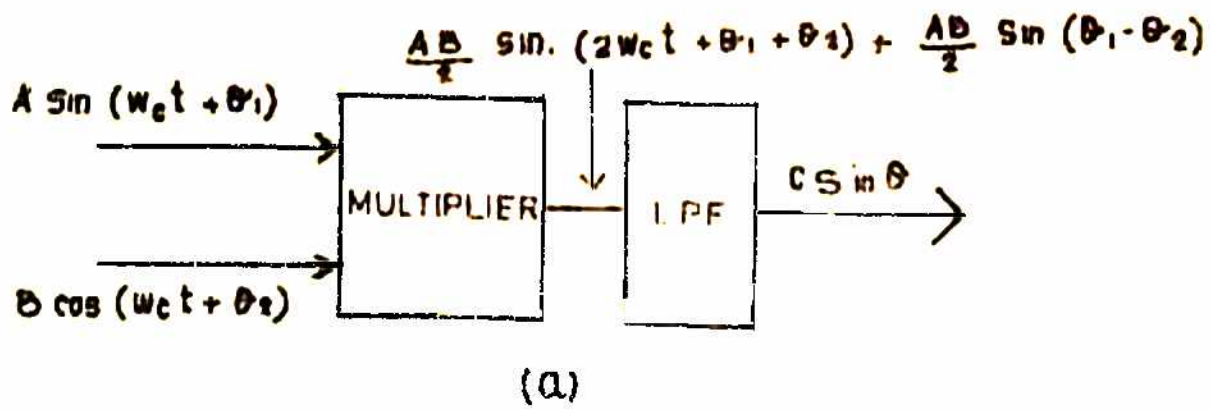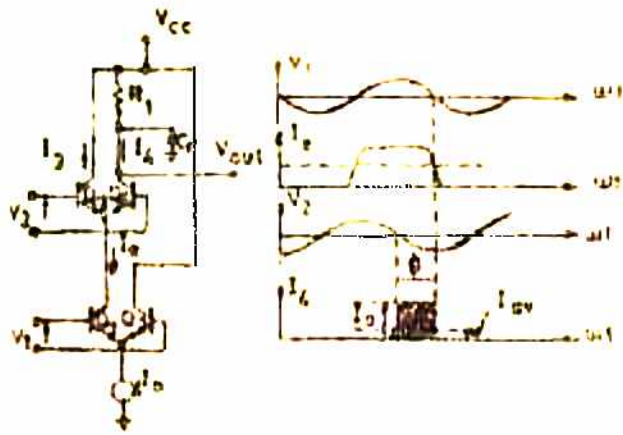nd V_2. In practice, however, it is not possible to get an accurate zero-crossing detector i.e., the transistors Q_4-Q_7 and Q_2-Q_3 show a definite threshold above which they show switching behavior. In Signetics IC PLL 565 the phase detector constant K_d (see Appendix D for the definition of K_d) reaches a constant value of 0.6 when the input is above 50 mVrms. This means that the phase detector of IC 565 behaves like a zero-crossing linear phase detector when the input voltage is above 50 mV rms.

3.2 ZERO-CROSSING INTERPRETATION OF THE SHORT-TIME AMPLITUDE SPECTRUM ANALYSIS HARDWARE

Fig.11 shows a block schematic of the system used fc estimating the short-time Amplitude Spectrum in the well-known sound spectrograph and in most filter-bank spectrum analyzers. Consider that the bandwidth of the input filter of Fig.11 were very narrow (ideally, a notc filter)

f(t)○

```
┌─────────┐        ┌──────────┐        ┌─────────┐
│   BP    │        │          │        │   LP    │
│ Filter  │───────▶│ Rectifier│───────▶│ Filter  │──▶ |F(w,t)|
│  p(t)   │        │          │        │  q(t)   │
└─────────┘        └──────────┘        └─────────┘
```

a'(w,t) or
b'(w,t)

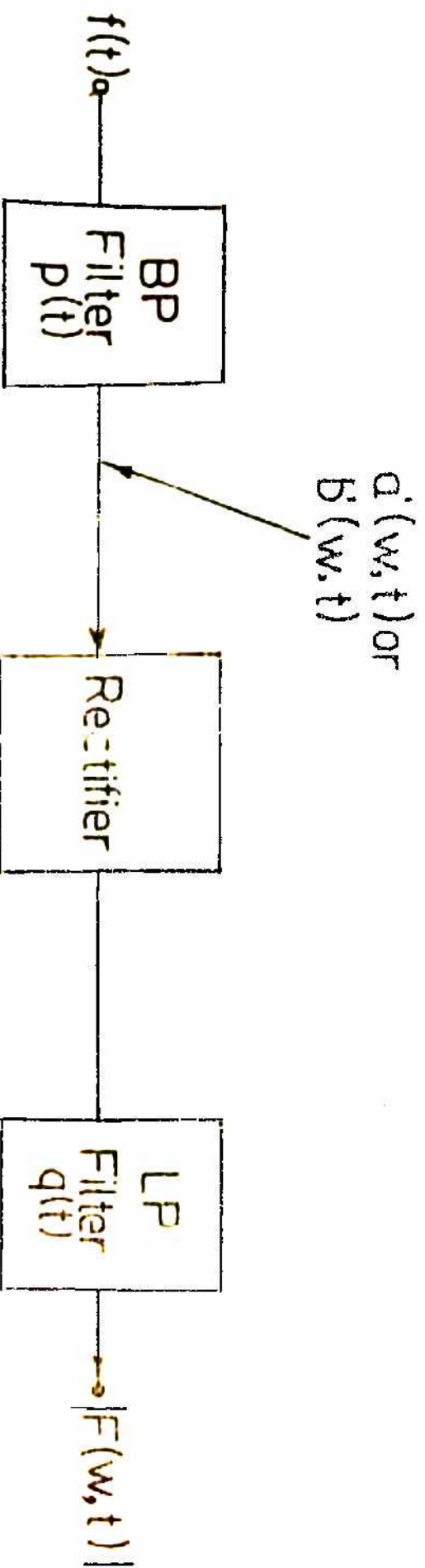FIG. 11 . PRACTICAL MEASUREMENT OF THE SHORT TIME
SPECTRUM F (w,t).

frequency at which the Amplitude Spectrum is to be estimated. The output of such a filter would be a near sinusoid and the peak rectifier would give the amplitude of the sinusoid at any given instant of time. Thus an ideal notch filter would give the desired $|F(\omega,t)|$ for all time. As the bandwidth of the filter is increased the amplitude as obtained at the output of the Rectifier Filter corresponds to the frequency obtained by finding the average zero-crossing count in that 2/B second interval where B is the bandwidth of the input filter. Thus it is necessary that the bandwidth of the input filter be as narrow as possible so that the interval over which the measurements are valid is long and accurate measurement is possible.

3.3 FORMANT TRACKING USING PHASE-LOCK LOOP

Fig.12 shows a functional block diagram of the human speech producing mechanism. This model consists of a vocal source function followed by a Vocal Tract Transfer Function, the output of this system being SPEECH. In the production of speech, the parameters of both the vocal source and vocal tract are modified as functions of time.

For the production of voiced speech the Vocal Source function is a quasiperiodic function of time with a pressure waveform which is roughly triangular in shape and with an average frequency of 130Hz for a male and 260Hz for a female. This function is shown as $S_1(t)$. For the productio of frictionally excited speech the vocal source function

is modelled as a noise generator $S_2(t)$ band limited from 50Hz to about 10000Hz. The spectral properties of this noise source are flat in the mid audio range [15].

The vocal tract transfer function indicated as $H(f,t)$ in Fig.12 modifies the output from the vocal source function to result in speech. This transfer function depends on the position of the articulators (tongue, lips, lower jaw and velum) as well as on overall physical dimensions of the vocal apparatus, and so for normal speech is a function of both frequency and time.

For voiced speech sounds the frequency response of the transfer function will have an approximate characteristic shape shown in Fig.13. The frequency of the first peak F_1, is called the first vocal resonance or first formant frequency. The frequency at the second peak labeled F_2, is called the second formant frequency and the frequency of the third peak labeled F_3 is called the third formant frequency. Although there are higher formant frequencies French and Stienberg [16] have shown that voiced speech can be reliably synthesized from a knowledge of only the first three formant frequencies.

A basic problem in speech analysis is the accurate determination of the formant frequencies, a process commonly called formant tracking. Formant tracking is useful in speech analysis, automatic recognition of speech by machine and bandwidth compression of speech with application of vocoders.
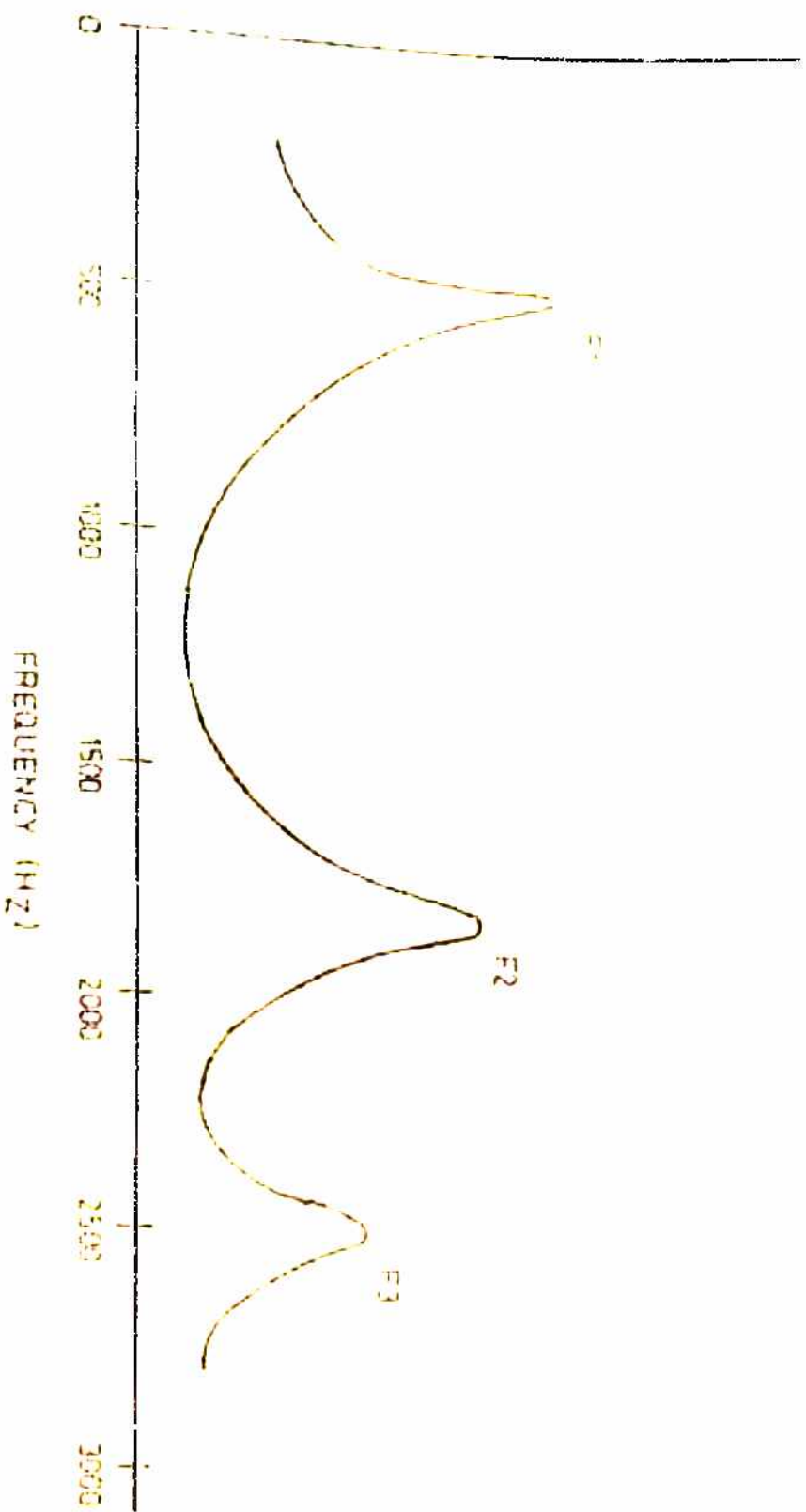
GAIN

FREQUENCY (Hz)

FIG. 13. APPROXIMATE TRANSFER FUNCTION OF THE VOCAL TRACT FOR /ε/ AVERAGE MALE SPEAKER.

Characteristics of Formants

The following are some of the characteristics of the first three formants for non-nasal voiced sounds uttered by an average male speaker.

1. Formant Frequency Ranges:

F_1: 300 Hz - 1000 Hz

F_2: 500 Hz - 2500 Hz

F_3:1700 Hz - 3000 Hz

2. Formant Bandwidth:

The first two formants have an average bandwidth of about 80 Hz. The higher the formant number the higher is the formant bandwidth [17]. The third formant has an average bandwidth of about 120 Hz.

3. The second formant frequency changes with an average slope of 40 Hz/ms. [18]. The highest slopes are encountered only for the second formant. The curves depicting the variation of the formant frequencies are band limited to 17 Hz [see Fig.14].

Many methods have been suggested for formant tracking. Approximation to the first and second formant frequencies of speech have been obtained by first filtering the speech into pass-bands corresponding roughly to the first two formant frequency ranges and taking average axis crossing density measurement - an approach based on the reasoning that Formants are the most prominent spectral components and consequently are
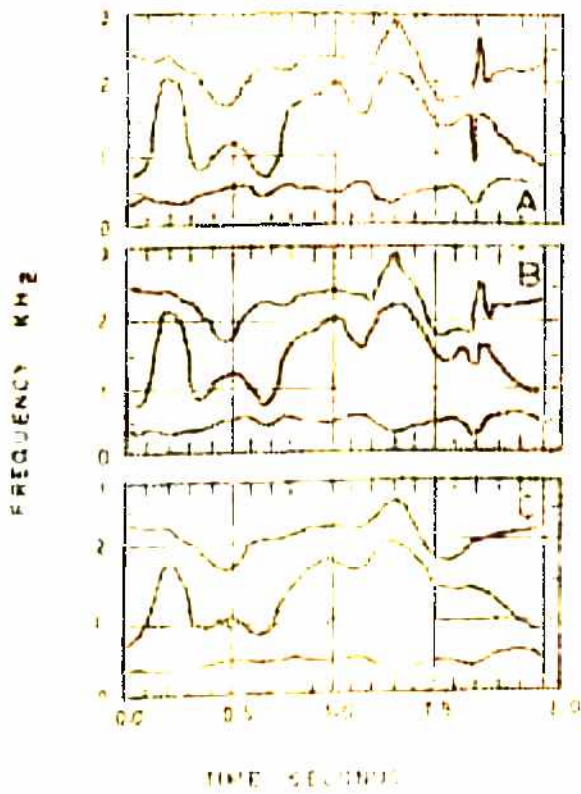
FIG. 14. SMOOTHING OF FORMANT SIGNALS. A – 50 Hz
BANDWIDTH (NO SMOOTHING). B – 16 Hz BANDWIDTH.
C 4 - Hz BANDWIDTH.

expected to have the strongest influence upon the axis-crossing rate.

Recent methods have been based on the method of moment calculations [19], on cepstrum techniques [20] zero-crossing wave analysis [21] and spectrum scanning and peak picking methods [22]. These recent methods are quite accurate but require elaborate digital processing of the speech signal. It should be noted that some special purpose extremely fast processors have been built which are capable of estimating speech parameters in real-time but generally these processors are either expensive or are not commercially available. Neiderjohn [23] has determined the distribution of axis crossing intervals of a prefiltered version of speech signal within each 20 msec speech segment using a computer and uses this information to determine the formant frequency. The method used to determine the formant frequencies and amplitudes in a typical formant vocoder is shown in Fig.15. [24].

It is seen that bandwidth of the input Filters is chosen to accomodate the full Formant frequency range. By the analysis presented earlier it is clear that with such large bandwidths of the input filters the corresponding time window is very narrow. Consequently, the Frequency and Amplitude information obtained from the output of such filters would have large fluctuations and would need larger bandwidth filters for further smoothing. Also, there is a possibility of two Formants
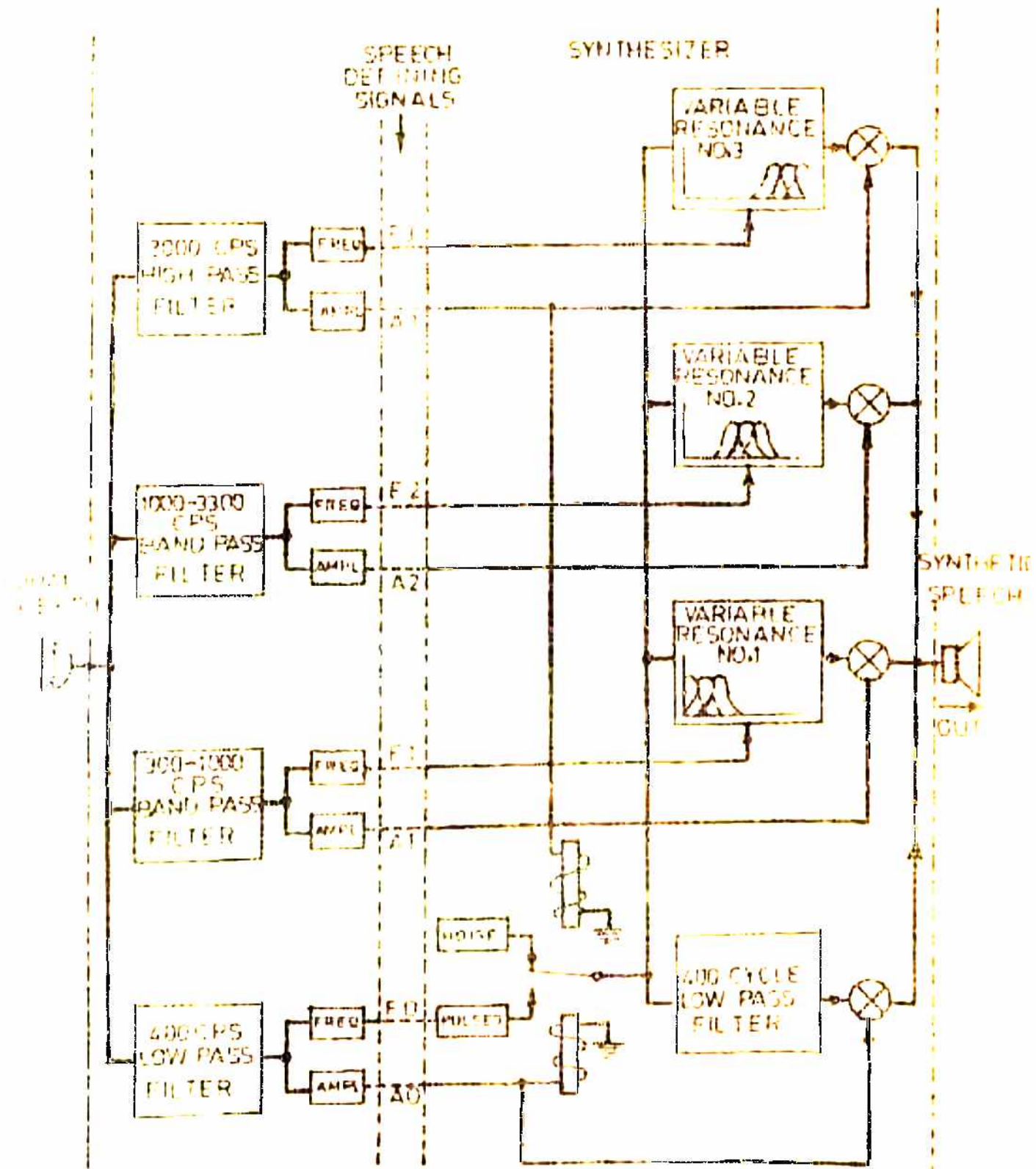
FIG.15.PARALLEL CONNECTED FORMANT VOCODER

falling within the bandwidth of a particular filter, thus fouling the entire measurement. It is therefore necessary to have narrow input filters whose centre frequency could be varied over the entire Formant frequency range. Fig.16 shows the block schematic of the Formant Vocoder based on the approach followed in this thesis. Imagine that the bandwidth of the input filters 1,2 and 3 is narrow and that the centre frequencies at any instant are close to the Formant frequencies at that instant respectively. The Formant frequency would have the greatest influence on the zero-crossings and the PLL would track any changes in the instantaneous frequency (the number of zero-crossings per second) and hence changes in the Formant frequency. The control signal to the VCO of the PLL could also be used to vary the centre frequency of the input BPF. Thus the centre frequency of the input BPF would vary in accordance with any change in the Formant frequency and the PLL would give an output which would indicate the variation of the Formant frequency with time. The amplitude of the PLL output at any instant would correspond to the amount of frequency difference of the Formant frequency from the free running frequency of the corresponding PLL as also the quiescent centre frequency of the input BPF.

It is seen that the block schematic of the synthesizer resembles the actual human speech producing mechanism more closely than the synthesizer given in Fig.15.

ANALYSER

FIGURE 3.0 SCHEMATIC OF INSTRUMENT

SYNTHESIZER
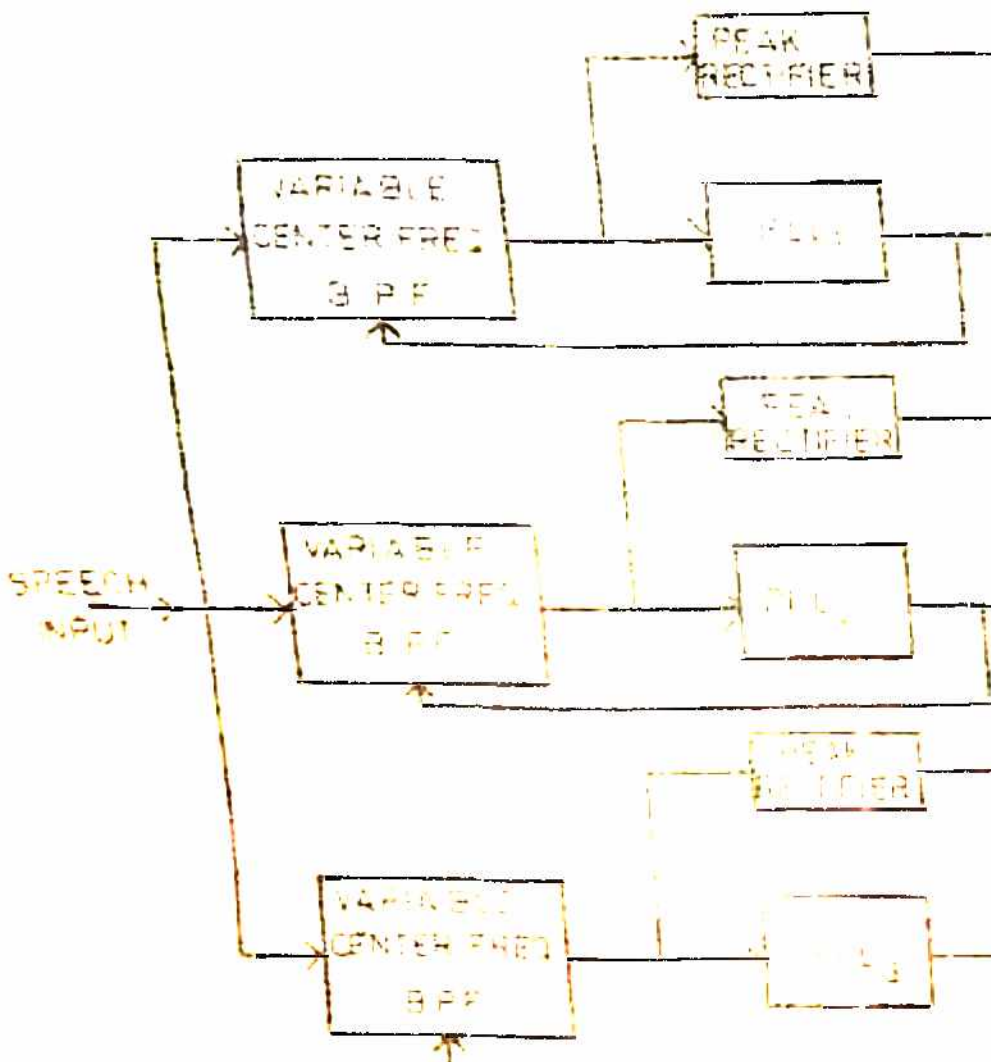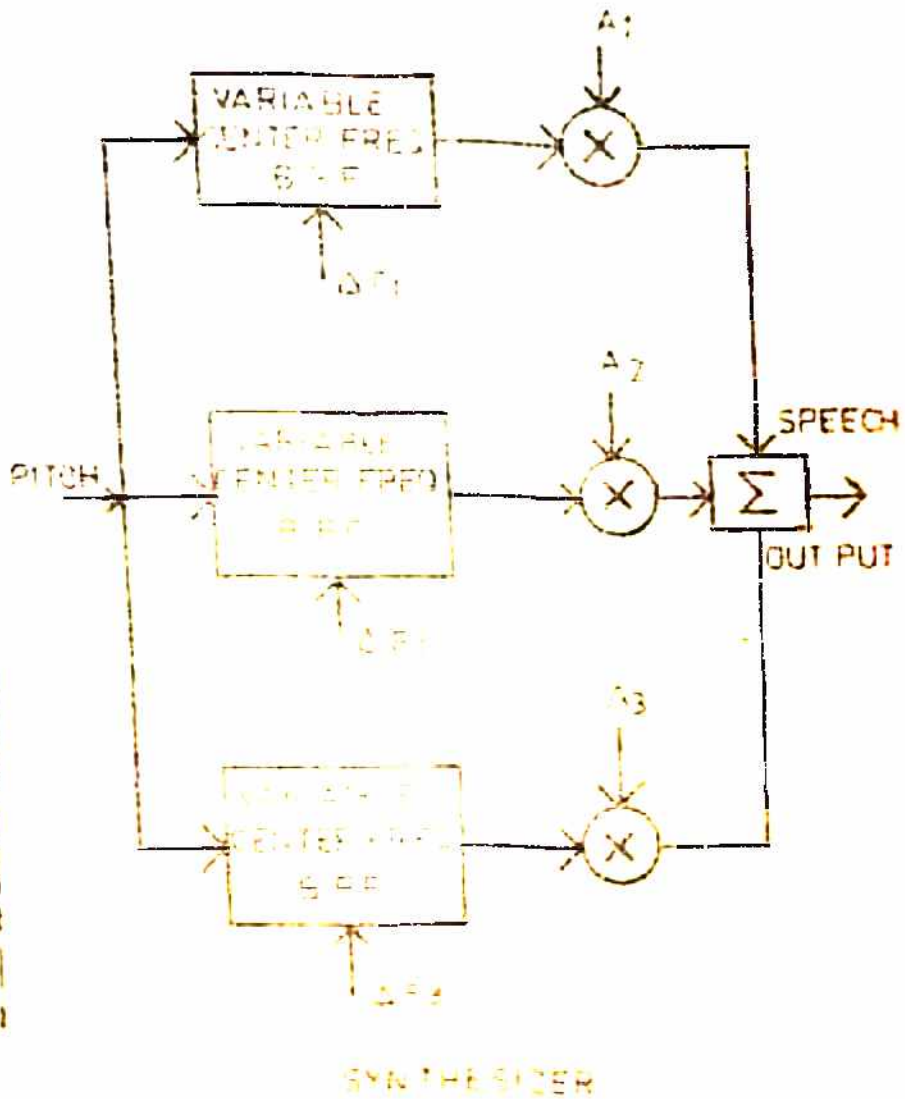
CHANNELS OF A VOCODER

The following are some guidelines for the design of a Formant Vocoder which uses the philosophy shown in Fig.16.

1. It should be possible to vary the centre frequency of the input BPFs over the entire Formant range.

2. The bandwidth of the filters should equal the bandwidth of the particular Formant and should remain constant as the centre frequency is varied. This is guided by the fact that a very narrow bandwidth of the input filter would be impractical since the initial acquisition on the Formant frequency would be difficult. Also, when the bandwidth of the input BPF is equal to Formant bandwidth, the synthesizer resembles the actual human speech producing mechanism and output from the BPF would be sufficiently large.

3. The lock range of the PLL should be larger than the Formant range in question.

4. The PLL should be able to accommodate the highest sweep rates. Maximum sweep rates are encountered in the second Formant normally.

5. The centre frequency of the BPF must follow the frequency of the VCO closely. For that it is essential that the gains of the VCO and that of the BPF centre frequency be the same.

6. The input to the PLL should be above that level where the phase detector gain is a constant and the PLL essentially uses the zero-crossing information of the input signal.

3.4 CONCLUSIONS:

The above analysis shows that

(1) Every linear phase detector uses the zero-crossing information of the inputs.

(2) Since Formant Tracking is possible by utilising the axis-crossing information of prefiltered speech a PLL with a linear phase detector can be used for Formant Extraction.

(3) Use of narrower input filters in a Formant Vocoder avoids Formant overlap, facilities frequency measurement and the Synthesizer resembles the human speech producing mechanism more closely.

In the next chapter, these guidelines have been used to design the three channels of a Formant Vocoder. Advantages and shortcomings of this approach have also been pointed out.

CHAPTER 4: DESIGN OF THE FORMANT CHANNELS OF A FORMANT VOCODER

Following the approach described in chapter 3 a design of the three Formant channels of a Formant Vocoder has been worked out and implemented. Signetics IC PLL 565 has been used as a device whose output is proportional to the axis crossing rate and hence the instantaneous frequency of prefiltered speech. It must be pointed out that the implementation of the design described in this chapter is only one typical scheme and many other schemes for satisfying the basic design goals set forth in the previous chapter are possible. Advantages and shortcomings of the hardware used have been mentioned.

The phase locked loops used to estimate the frequency must have the following characteristics.

1. Centre frequency 650 Hz - First channel, 1500 Hz Second channel 2500 Hz - Third channel. These frequencies are approximately in the centre of the Formant ranges.

2. Lock range: $\pm$ 350 Hz nominal - First channel +1000 Hz - Second channel +1000 Hz - Third channel. The capture range should also be about the same as the lock range. This point has been explained later.

3. Lock up time: Less than 2.5 msec. This time is 10 percent of $\frac{2}{B}$ seconds (25 msec) since B = 80 Hz.

4. The PLL must be able to track changes in the frequency of the order of 40 Hz/msec or the maximum sweep rate that the PLL can accomodate must be 40 KHz/sec.

The input band pass filters must have the following characteristics.

1. The centre frequencies of the filters corresponding to channel 1: 650 Hz, channel 2: 1500 Hz, channel 3: 2500 Hz.

2. The 3-dB bandwidth of these filters is 80 Hz since the average bandwidth of the 1st and 2nd Formant is 80 Hz for a male speaker and about 120 Hz for the third formant. This bandwidth should not change with change in the centre frequency. For simplicity all bandwidths were kept 80 Hz.

3. The variation of the centre frequencies with the O/P voltage must be similar to the variations of the VCO frequency.

4. The gain of the input BPFs at the centre frequencies should be about 10. This is necessary so that the input to the PLL is above 50 mVrms when the output from the tape-recorder is about 20 mVrms.

4.1. DESIGN OF THE INPUT BAND PASS FILTERS

Fig.17(a) shows the circuit diagram of a BPF. The transfer function of this filter is [25],[31]

$$\frac{V_O(s)}{V_i(s)} = \frac{S/R_1 C_1}{s^2 + \frac{C_1 + C_2}{R_3 C_1 C_2} s + \frac{1}{R' R_3 C_1 C_2}}$$

where $R' = R_1 || R_2$ .. (4.1.1)
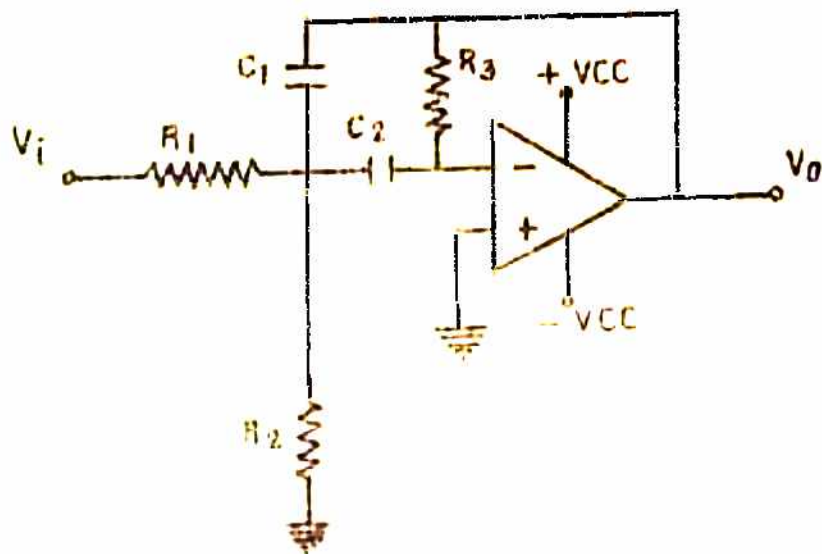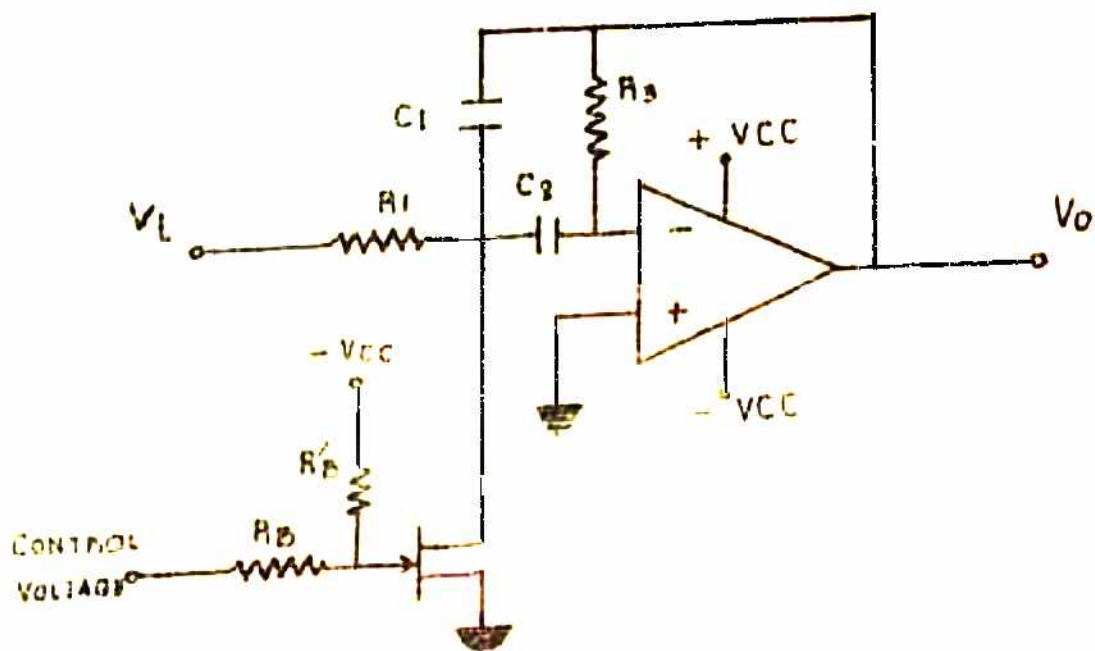
(A) BAND PASS FILTER

FIG.17. VARIABLE CENTRE FREQUENCY BAND PASS FILTER (VCF-BPF)

comparing with the standard expression for an active BPF

$$A_v(s) = -\frac{(W_o/Q) A_o s}{s^2 + (UU_o/Q) s + UU_o{}^2} \qquad \cdot\cdot \ (4.1.2)$$

where A_o = mid band gain

$$\frac{UU_o}{2\pi} = f_o = \text{centre frequency}$$

$$Q = \frac{f_o}{B} = \text{Quality factor} \qquad \cdot\cdot \ (4.1.3)$$

we get

$$R_1 C_1 = \frac{Q}{UU_o A_o} \qquad \cdot\cdot \ (4.1.4)$$

$$R_3 \cdot \frac{C_1 C_2}{C_1 + C_2} = \frac{Q}{UU_o} \qquad \cdot\cdot \ (4.1.5)$$

$$R' R_3 C_1 C_2 = \frac{1}{UU_o{}^2} \qquad \cdot\cdot \ (4.1.6)$$

It is clear from eqns (4.1.4) to (4.1.6) that

if R_2 alone is varied the centre frequency of the BPF will change while the bandwidth and the mid band gain will remain same. This can be achieved in practice by the arrangement shown in Fig.17(b) where the FET Drain to source resistance is varied by a control voltage which varies the bias on the FET Gate. The quiescent value of the FET drain to source voltage and hence the centre frequency of the filter can be adjusted by adjusting the values of resistors R_B and R'_B. However, this filter needs a compensated control voltage if linear variation

of the centre frequency is desired. This is so because
of the non-linearity of the FET characteristic when it is
used as a variable resistance. The compensated BPF is
shown in Fig.18.

The requirements are

$$A_o = 10, \quad B = \frac{f_o}{c} = 80 \text{ Hz}$$

Let $R_1 = 10$ K, then from eqn $(4.1.4)$

$$C_1 = 0.0198 \; \mu \; F$$

Let $C_1 = 0.015 \; \mu \; F$

Let $C_2 \quad 0.015 \; \mu \; F$

Then from eqn. $(4.1.5)$ $R_3 = 274$ K

Let $R_3 = 270$ K

The values of R and R' are adjusted to get a
suitable value of R_2 for obtaining the desired centre
freq. It was practically found that $A_o = 10.8$ and
B = 80 Hz were obtained with the following component
values.

$C_1 = 0.015 \; \mu F$, $C_2 = 0.01 \; \mu F$, $R_1 = 10$ K, $R_3 = 270$ K

The final circuit and the transfer characteristics
are given in Fig.19. Similar BPFs were fabricated for
the synthesizer.

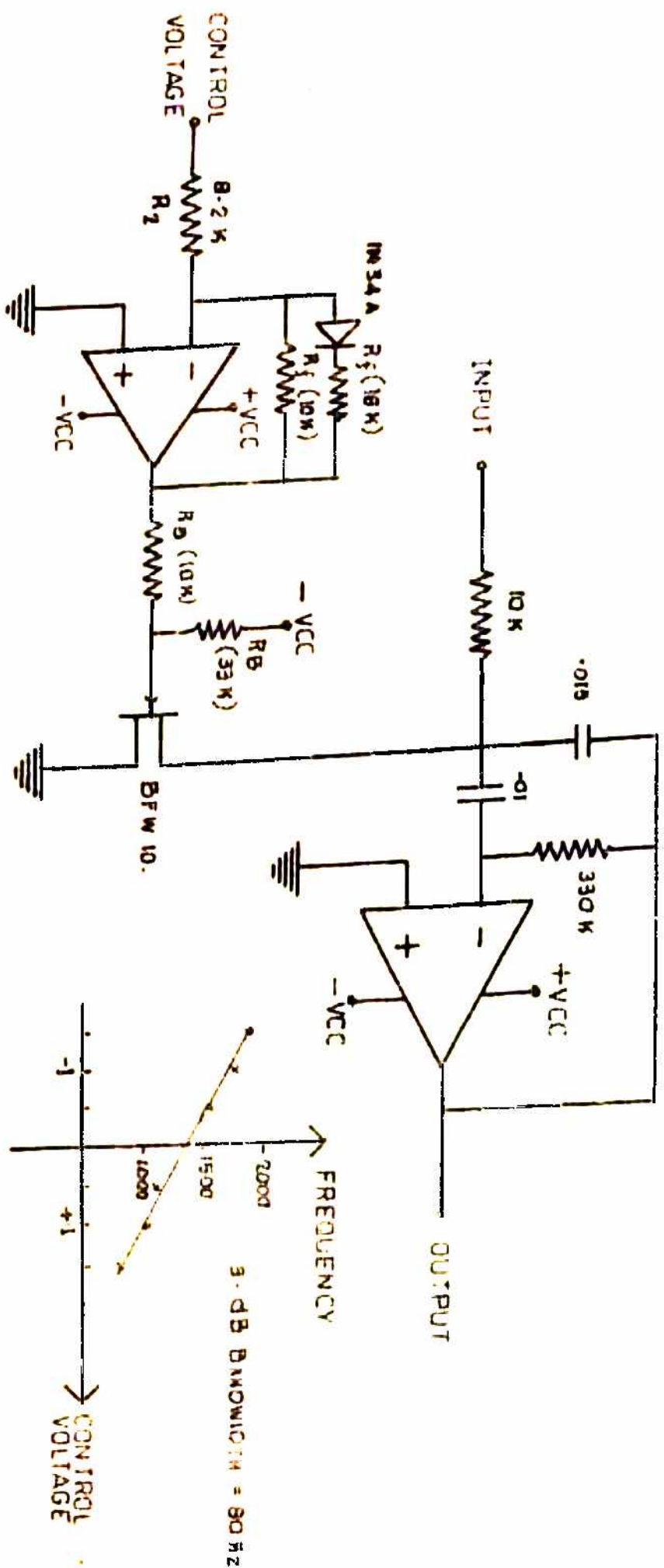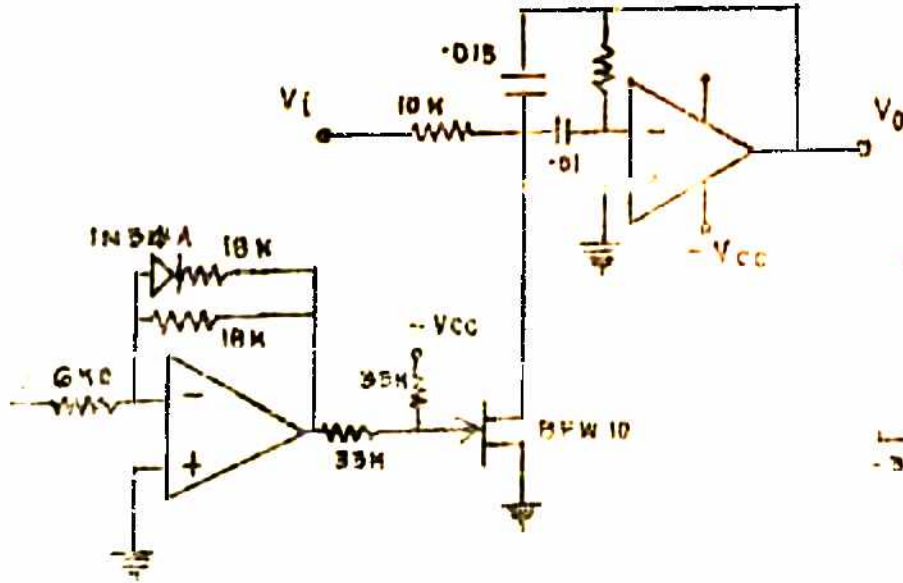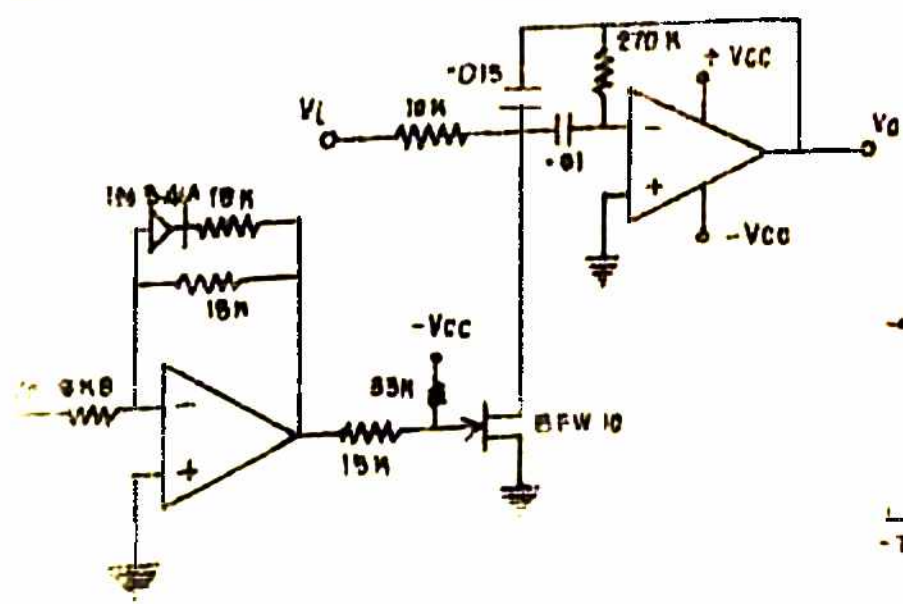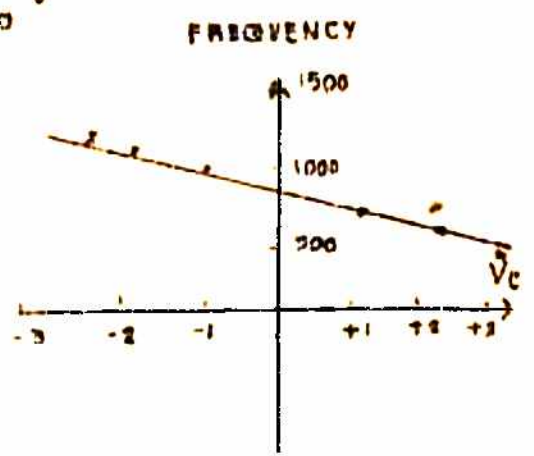FIG. 18. COMPENSATED VARIABLE CENTRE FREQUENCY BPF AND CHARACTERISTICS.

1 ST. CHANNEL

3 dB BANDWIDTH = 80 Hz
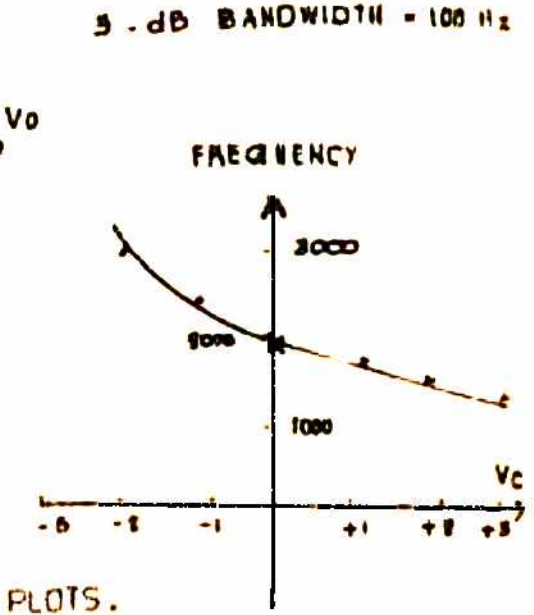
2 ND. CHANNEL

3 dB BANDWIDTH = 100 Hz
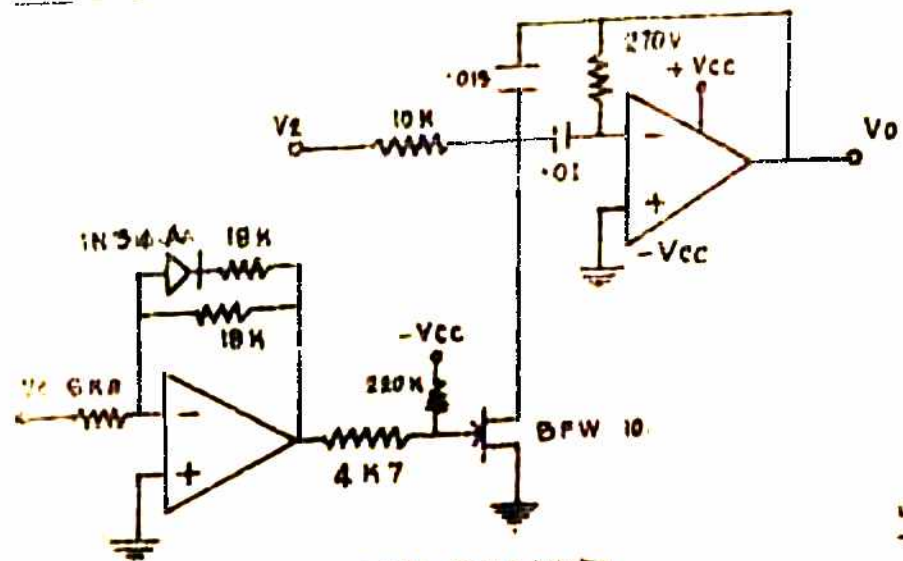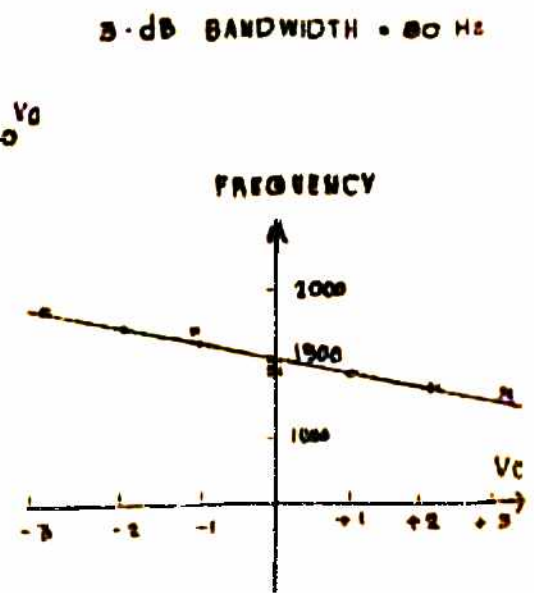
3 RD. CHANNEL

FIG. 19. B P F 5 AND FREQ VS. VOLTAGE PLOTS.

4.2 SMOOTHING FILTERS

The smoothing filters for the Amplitude and Formant Frequency variation signals should have a cut off frequency of less than 40 Hz since the bandwidth of the input filter is 80 Hz. From Fig.14 it is clear that a 16 Hz smoothing filter has negligible effect on the parameter variations.

Thus, the Smoothing Filters are simple R.C low-pass Filters with

R = 4.7K   and C=2 µF

4.3 DESIGN OF THE FORMANT TRACKING PLL

The signetics IC PLL 565 was used to track the Formant frequency changes. The design equations for 565 PLL [26,27] are given below [see Fig.20].

$$\text{Centre frequency } f_0 = \frac{1.2}{4R_1 C_1} \qquad \text{.. } (4.3.1)$$

$$\text{Lock range } f_L = \pm \frac{8f_0}{V_{cc}} \qquad \text{.. } (4.3.2)$$

$$\text{Capture range } f_C \quad \pm \frac{1}{2\pi} \sqrt{\frac{2\pi f_L}{\tau}}$$

$$\text{where } \tau = (3.6 \times 10^3) \times C_2 \qquad \text{.. } (4.3.3)$$

$$\text{Natural frequency } W_n = \sqrt{\frac{K_o K_d}{\tau}} \qquad \text{.. } (4.3.4)$$

$$\text{Damping factor } \zeta = \frac{W_n}{2K_o K_d} \qquad \text{.. } (4.3.5)$$

$$\text{Lock up time } T_p = \frac{(\Delta W)^2}{2 \zeta W_n^3} \qquad \text{.. } (4.3.6)$$
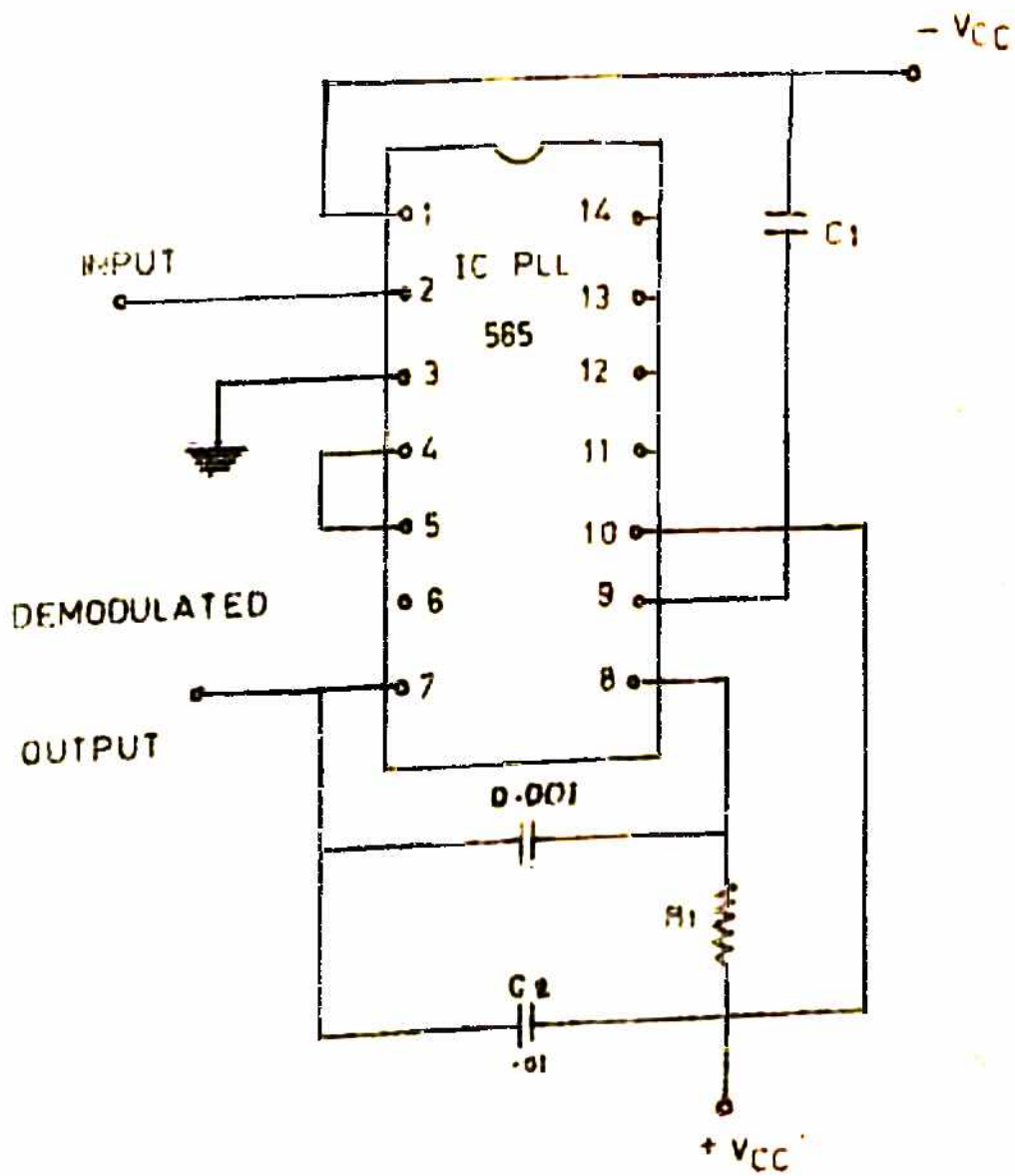
FIG. 20. BLOCK SCHEMATIC OF PLL AS FM DEMODULATOR

Maximum sweep rate that the loop can accept $= \Delta \dot{UU}$

and , $\dfrac{\Delta \dot{UU}}{UU_n{}^2} < 0.5$

where

$$K_o = \dfrac{8.0 \; UU_o}{\text{Total supply voltage}}$$

$K_d = 0.6$ for $V_i > 50$ mV rms

K_o and K_d are the gains of the VCO and the phase detector respectively.

The requirements for a Formant Vocoder are given on page 46.

Let $R_1 = 8.2$ K

for $f_o = 650$ Hz	$C_1 = 0.056 \; \mu F$
$f_o = 1500$ Hz	$C_1 = 0.024 \; \mu F$
$f_o = 2500$ Hz	$C_1 = 0.0146 \; \mu F$

1st channel

For

$V_{cc} = 8V$ and $C_2 = 0.01 \; \mu F$

$K_o = \dfrac{8 \times 2\pi \times 650}{16} = 2042$

$K_d = 0.6$

$UU_n = 5.83 \times 10^3$ rad

$\zeta = 2.38$

$T_p = 0.04$ msec $\ll 2.5$ msec

$f_L = \pm 650$ Hz

$f_c = \pm 1.696$ KHz

Notice that in the calculation of τ_p, ΔUU is taken to be $2\pi \times 1$ KHz since the PLL may be required to acquire lock over a range of ± 1 KHz.

$$\text{Maximum sweep rate } \Delta \dot{UU} < 0.5 \, W_n{}^2$$
$$< 16.99 \text{ MHz/sec}$$

Thus this loop can safely accomodate sweep rates of 40 kHz/sec. It can be easily shown that all the requirements for higher number channels can also be satisfied with $C_2 = 0.01 \, \mu F$. Thus, the circuit shown in Fig.22 may be used with appropriate values of C_1 for the second and third channel. Similarly, the amplitude signal, which is obtained by passing the output of the BPF through a rectifier filter which uses a 16 Hz filter. The complete circuit diagram of the Formant Analyser is shown in Fig.21 and that of the synthesizer in Fig.22.

4.4 RESULTS AND DISCUSSIONS

All the subsystems were tested separately and found to work satisfactorily. However, when the complete system was tested for synthetic speech using the laboratory set up shown in Fig.23 but it did not function as desired. Echoes of Standard male utterances with known spectrographs were played from the tape-recorder. The output synthetic speech was not intelligible. The possible reasons for this are

1. It is difficult to match the voltage to frequency transfer characteristics of the BPFs and the
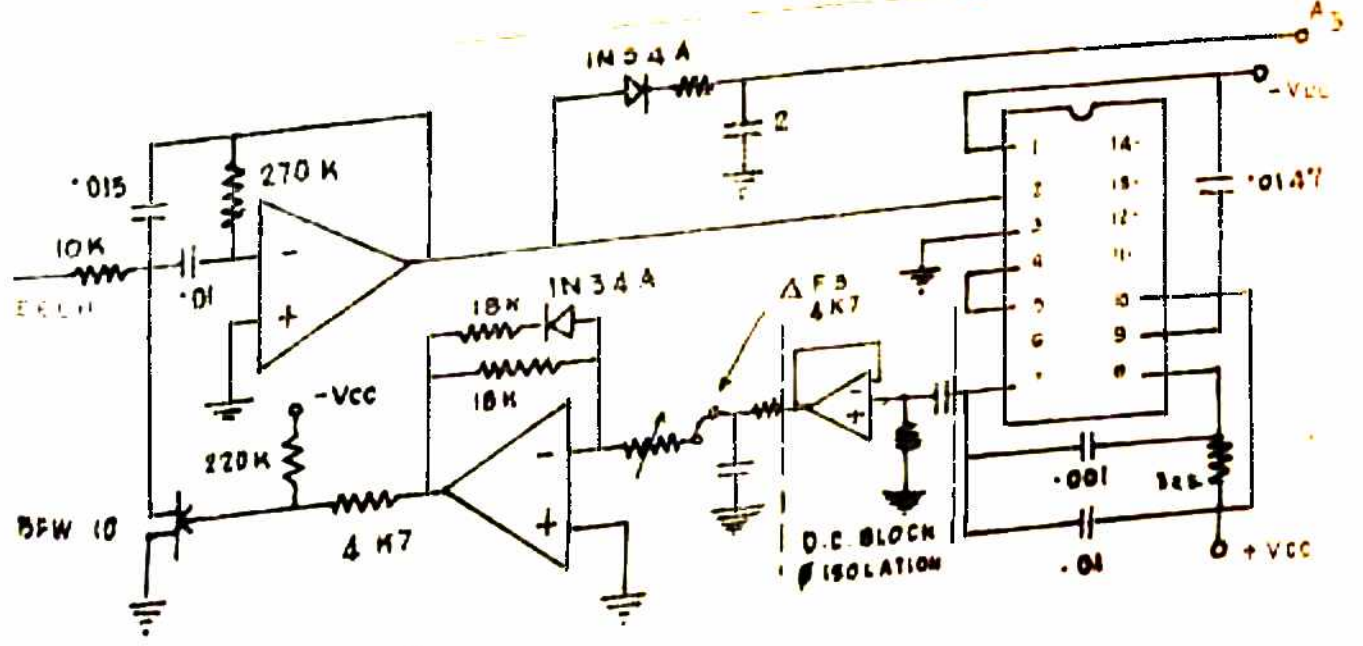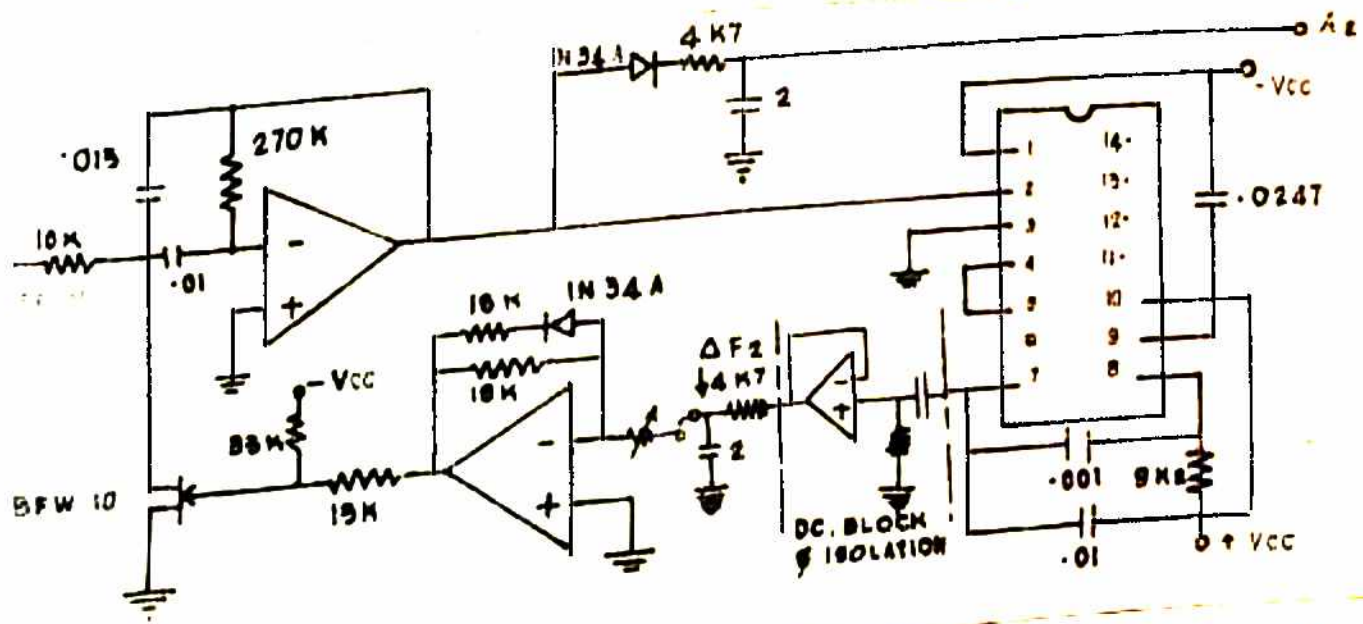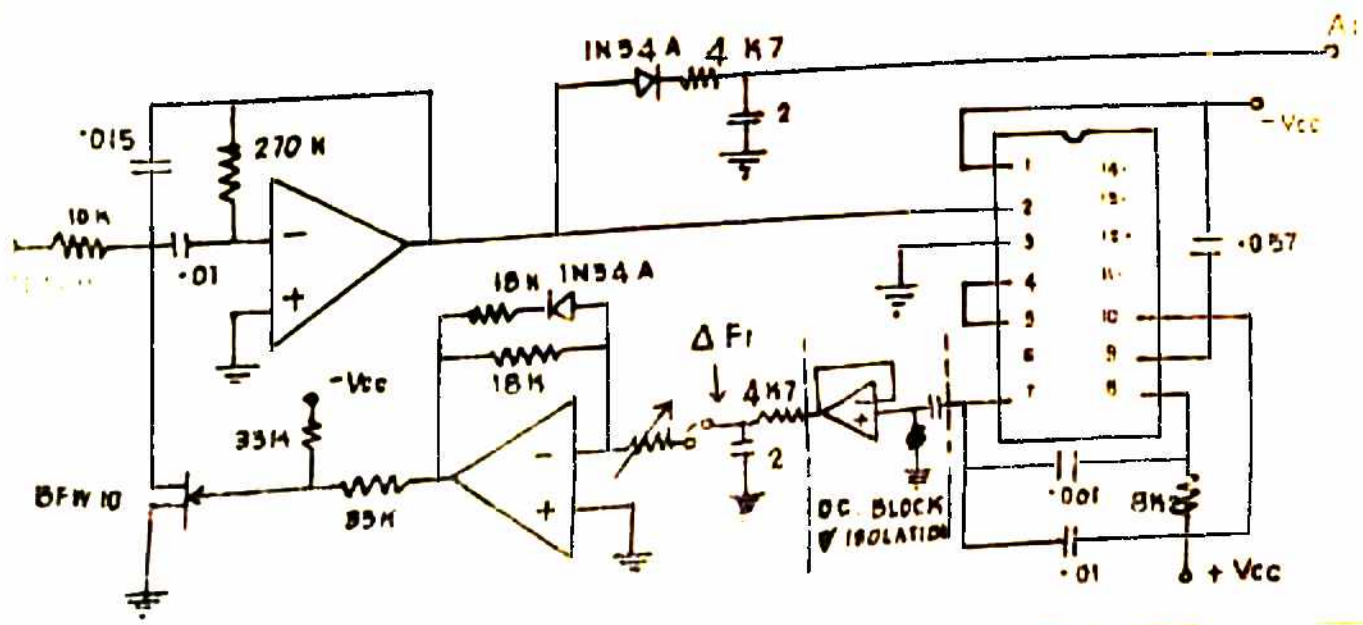
ALL CAPACITORS IN ·μF.

FIG. 21. CIRCUIT DIAGRAM OF FORMANT VOCODER ANALYSER

Fig. 22. CIRCUIT DIAGRAM OF FORMANT VOCODER SYNTHESIZER
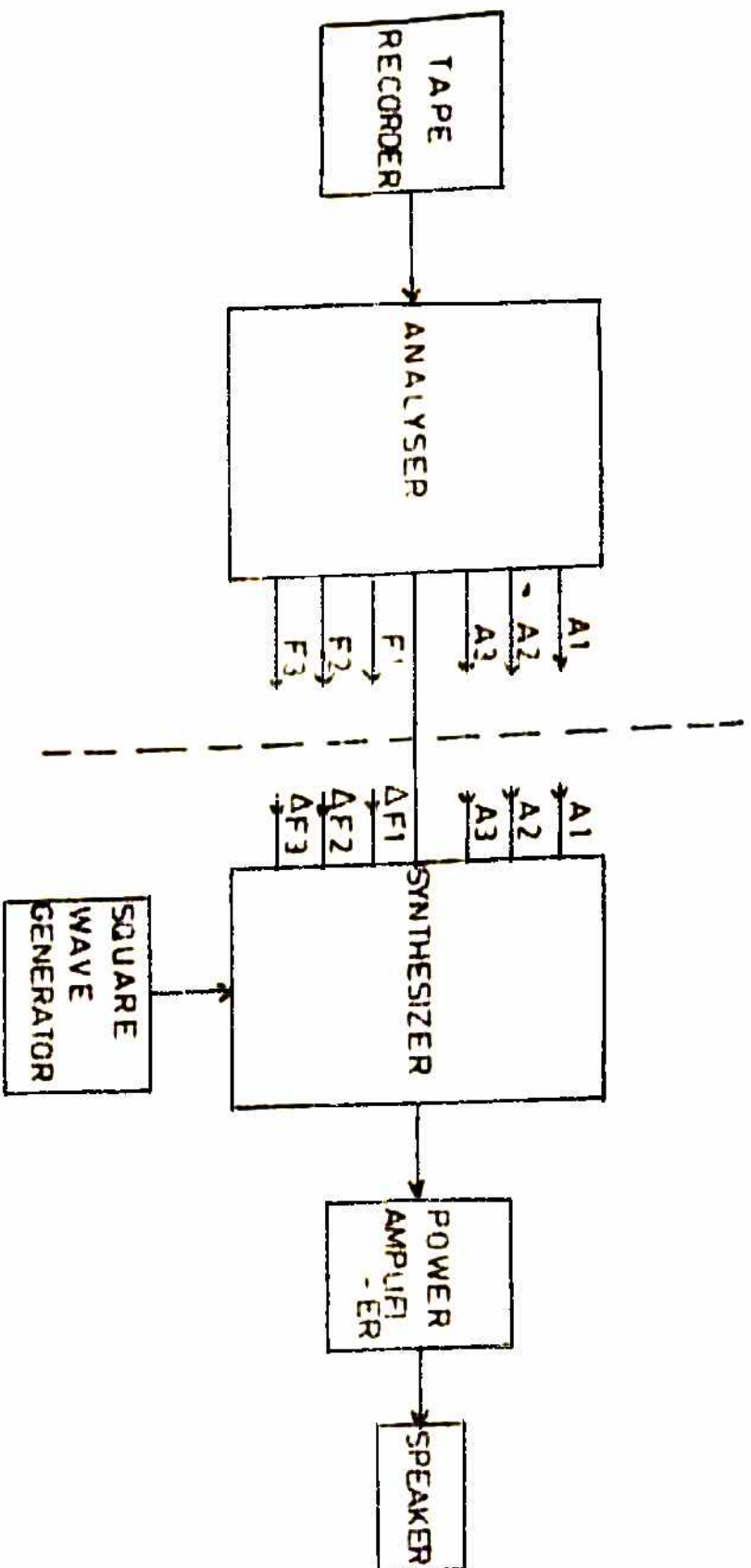
Fig. 23. LABORATORY SET UP FOR TESTING FORMANT VOCODER

internal VCOs of 565 for the three channels over
the entire operational frequency range. A closer
look at the voltage vs frequency characteristics
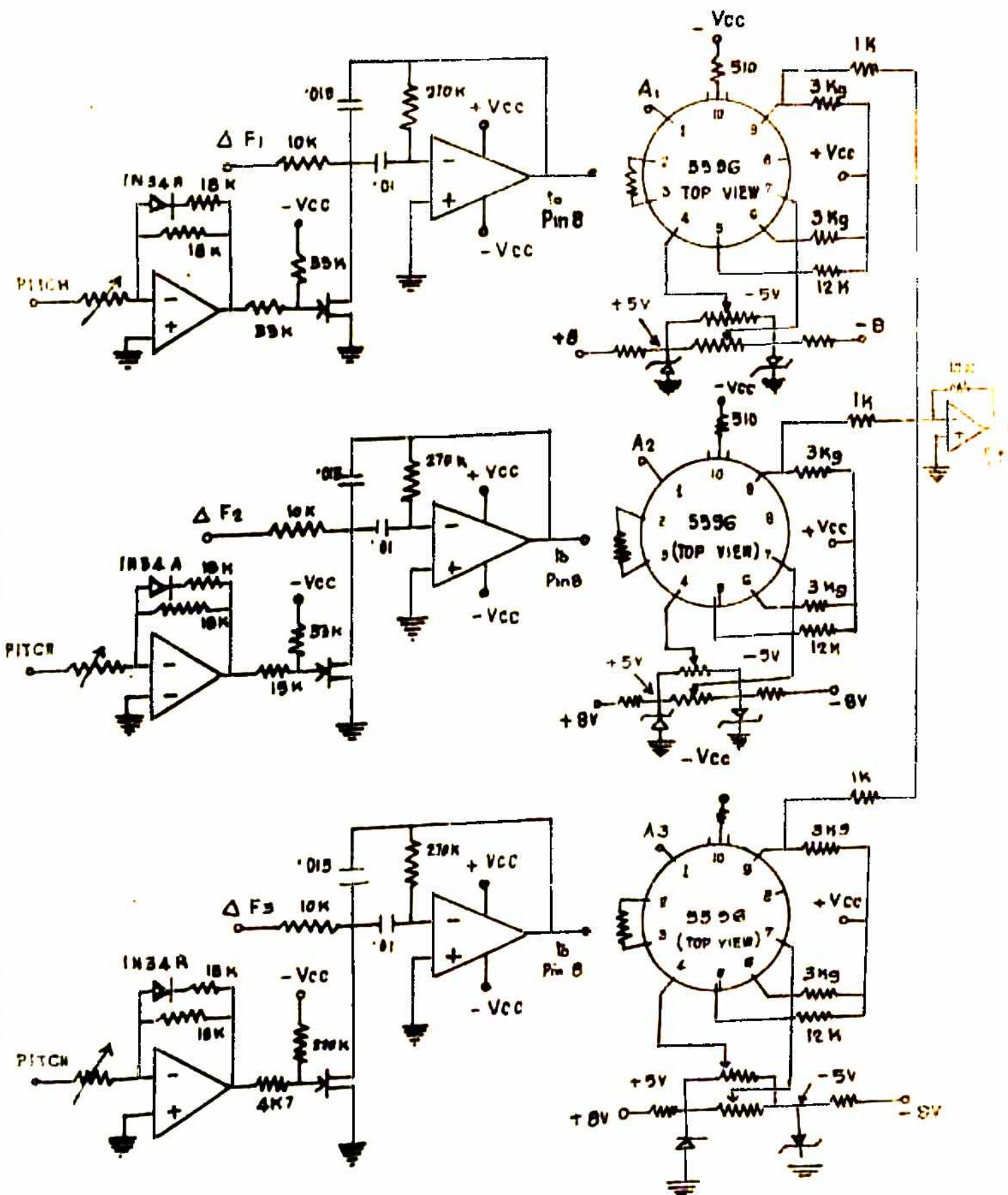of the filters reveals that though the overall
characteristics appear linear the change in frequency
with voltage may not be linear for small voltage
changes. The best linearity is observed between
the frequency range 1000 Hz to 2000 Hz since a
better control on the value of Drain to source
resistance is possible. The linearity is poorer
for the first and third filter. Fig.19(c) shows
that for -ve voltages the centre frequency increases
much faster since as the Drain to source resistance
becomes smaller the voltage to frequency charac-
teristics become much softer and a tight control
on the value of the resistance is difficult. This
seems to be the chief reason for the failure. Also,
it was not possible to push the centre frequency
of the third filter beyond 2100Hz.

It may be noted that the gain of the VCO of IC PLL 565
is 2042Hz/volt even for the first channel. Since the
bandwidth of the input filter is only 80Hz, linearity
down to a few millivolts is essential. This also brings
into focus the need for the gains of the VCO and the
filter to be exactly matched - a difficult task using
discrete components. In Fig.21, the feedback from the
output of the PLL to the BPF is given after smoothing
to maintain similarity between the Analyser and Synthe-
sizer. Feedback prior to smoothing was also tried.

The slope of the voltage vs frequency plot can be easily varied by varying either R_i, R_f and R'_f or R_B and R'_B (see Fig.18). But R_B and R'_B also decide the quiescent frequency and hence should not be disturbed. The compensating network, R_f and R'_f should also not be varried indiscretely otherwise the linearity suffers. Hence, R_i can be varied to adjust the slope and make it equal to that of the particular VCO. In Fig.19, R_i is taken to be 6 K 8 so that the voltage vs frequency plot can be obtained easily in the laboratory. With smaller values of R_i an accurate power supply with millivolts range should be available. R_i has been shown 'Variable' in Fig.21 as it can be adjusted in the laboratory.

2. The circuit suffers from initialisation i.e. it is likely that the Formant frequency lies outside the pass-band of the input filters initially. This would result in the loss of data only till the Formant frequency is within the pass-band of the filter because once under lock the loop has the capability to track any changes of the Formant Frequency over the entire range. This can be alienated only by making a quick guess of the formant frequencies initially (or after every major break) by some technique e.g. the peak-peaking technique and a suitable control voltage generated to change the centre frequency of the BPF to the desired values. The PLL can be designed have a capability to lock on to any frequency

within which the Formant is likely to lie. It is for this reason that the capture range has been taken to be $\pm$ 1 KHz.

The values of the biasing resistances of IC 5596 were adjusted to make it work as an ideal multiplier. Similarly, the values of the resistances R_1 (shown variable in Figs. 21 and 22) were adjusted in the laboratory.

In this design it has been assumed tacitly that the SNR at the output of a tape recorder when the recording is done professionally is very high. Then, the loop SNR for the phase lock-loop can be shown to be high enough for the linear analysis to remain valid $\angle$ 27,30 $\angle$.

CHAPTER 5: CONCLUDING REMARKS

It is believed that the analysis of FM Threshold presented in this thesis is the FIRST satisfactory explanation of the phenomenon of FM Threshold using the zero-crossing approach. The calculation of Noise effects in FM detection using this technique was considered tedious and the physical insight into the phenomena taking place particularly these leading to FM Threshold tended become obscure in the mathematical manipulations. It has been shown that a simple explanation to the phenomena of FM Threshold is possible and the mathematics is not complicated. Simple explanations have been offered to other observed facts in FM detection as well. Thus, this thesis places the zero-crossing approach to calculation of noise in FM systems on a very firm ground.

The physical explanation of the phenomenon of FM threshold comes as an extension of FM capture. In fact, very close analogies can be drawn between the phenomenon of FM capture and Threshold and FM Threshold can be termed as capture due to Noise if the condition for spike formation is satisfied.

It is seen that the working of the short-time Amplitude spectrum hardware for speech Analysis can be easily understood and definite design guidelines can be generated using this approach. The hardware described in this thesis helps one to use the phase lock loop as a device which uses the zero-crossing information for estimating the instantaneous frequency.

5.1 SCOPE FOR FURTHER WORK

1. FDM-FM is a recognised modulation scheme for Satellite Communication [28]. Methods of calculating the Wanted to Interfering Carrier Ratio in Geostationary Satellite Networks have been given in CCIR study group report 455-1 [29]. The specific effects on system service will depend on many additional factors such as (1) type of service, e.g., telephony, television, data, etc., (2) type of Modulation (3) Modulation parameters and (4) desired carrier-to-system thermal noise ratio. This is a subject of continuing investigation. Since the analysis presented in this thesis is in terms of differences in instantaneous frequencies and amplitude ratios, this study can be very well extended for calculation of Inter - ference effects on board the satellite as well as on Ground Stations in any specific Network when FDM-FM is used.

2. The zero-crossing model of FM Threshold gives a deeper insight into the formation and nature of spikes due to additive Gaussian noise. The duration of spike and the amount of Frequency jump encountered for a given Signal to Noise ratio and difference in instantaneous frequencies of signal and Noise can be estimated using this model. This information is of vital importance

in the Design of Threshold Extension Devices
(TEDs). [c.f. Ref.10. pp.348-351 for detailed
discussion].

3. Using the approach and design guidelines set
   forth in this thesis it is possible to think of
   a dedicated digital system based around a micro-
   processor for real time tracking of Formants in
   human speech. An important spin-off of this
   study is that it helps one to visualize the use
   of all digital systems in FM and other analog
   communication systems.

APPENDIX-A

Relationship between the instantaneous frequency and the average zero-crossing count of an FM wave.

FM detection is possible by counting the number of times the FM waveform crosses origin in a stipulated time interval.

Consider an FM signal given by

$$v(t) = A \cos \left[W_c t + \int_{-\infty}^{t} f(t) \, dt \right] \qquad .. \ (A-1)$$

$$= A \cos \left[\Theta(t) \right]$$

where

$$f_c = \frac{W_c}{2\pi} \qquad - \text{ is the frequency of the carrier and}$$

$f(t)$ is the modulating signal.

The modulating signal usually changes much more slowly than the carrier and hence may be assumed to be constant between two consecutive zero-crossings of the FM signal. Let t_1 be zero-crossing as shown in Fig.A-1 with $t_2 = t_1 + \Delta t$ the next zero-crossing. Then,

$$\Theta(t_2) - \Theta(t_1) = \pi = W_c (t_2 - t_1) + K \int_{t_1}^{t_2} f(t) \, dt.$$

$$= W_c (t_2 - t_1) + K \, f(t_1) (t_2 - t_1)$$

$$= [W_c + Kf(t_1)] (t_2 - t_1) \qquad .. \ (A-2)$$

From eqn. (A-2) we have

$$W_i = W_c + k \, f(t_1) = \frac{\pi}{t_2 - t_1}$$
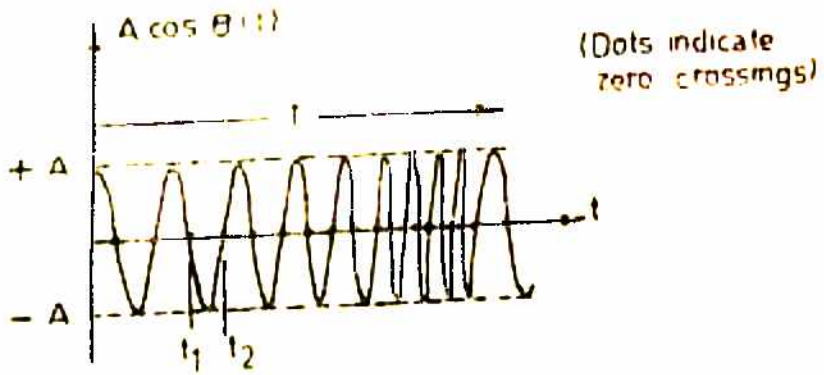
FIG. A-1. ZERO-CROSSING DETERMINATION
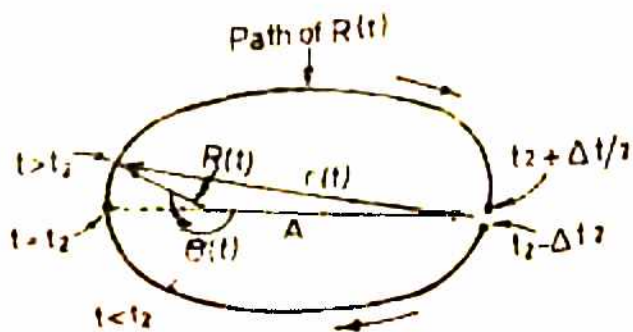

FIG. B-1. LOCUS OF R(t) AND θ(t) TO CAUSE A NEGATIVE
SPIKE BETWEEN t_2 AND $t_2 + \Delta t$

or $\quad f_1 = f_c + \frac{K}{2\pi} f(t_1) = \frac{1}{2(t_2 - t_1)}$ $\qquad$ .. (A-3)

where
$$f_i = \frac{\omega_i}{2\pi} = \text{instantaneous frequency} = \frac{1}{2\pi} \frac{d}{dt} [\theta(t)]$$
$\qquad$ .. (A-4)

The desired output may thus be found by measuring the spacing between zero-crossings. If positive going zero-crossings only are considered, we get

$$f_i = \frac{1}{t_2 - t_1}$$
$\qquad$ .. (A-5)

A simple way of measuring the spacing between zero-crossings is to actually count the number of zero-crossings in a given time interval.

Thus, considering a counting interval, T_c, long enough so that it counts a significant number of zero-crossings, yet short enough compared with $\frac{1}{B}$ so that $f(t)$ still does not change too-much in this time interval

$$\frac{1}{T_c} < T_c << \frac{1}{B}$$
$\qquad$ .. (A-6)

Let n_c, be the number of positive zero-crossings in T_c seconds. Then,

$$n_c = \frac{T_c}{t_2 - t_1}$$
$\qquad$ .. (A-7)

$\qquad$ .. (A-8)
$$f_i = \frac{n_c}{T_c}$$

The instantaneous frequency f_i, and hence the modulating signal $f(t)$ can thus be directly found in terms of the measured count n_c.

APPENDIX-B

Calculation of Mean Time between Spikes by Rice's Analysis

B.1 Properties of Narrow-Band Gaussian noise.

Narrow-Band noise, $n(t)$ is represented by

$$n(t) = x(t) \cos W_c t - y(t) \sin W_c t \quad \text{.. (B-1)}$$

Probability densities of $x(t)$ and $y(t)$ are given by

$$f(x) = \frac{1}{\sqrt{2\pi \eta B}} e^{-x^2/2\eta B}$$

$$f(y) = \frac{1}{\sqrt{2\pi \eta B}} e^{-y^2/2\eta B} \quad \text{.. (B-2)}$$

The probability densities of the time derivitives of $x(t)$ and $y(t)$ are given by

$$f(\dot{x}) = \frac{\exp \dfrac{-\dot{x}^2}{(2\pi^2/3)\eta B^3}}{\sqrt{(2\pi^3/3)\eta B^3}}$$

$$f(\dot{y}) = \frac{\exp \dfrac{-\dot{y}^2}{(2\pi^2/3)\eta B^3}}{\sqrt{(2\pi^3/3)\eta B^3}} \quad \text{.. (B-3)}$$

Rice's Conjecture of Spike Formation

Fig.B-1 shows Rice's conjecture of a negative spike. A spike results at the discriminator output when $\Theta(t)$ rotates 2π rad. If the rotation was counter clock-wise so that $\Theta(t)$ increased by 2π the spike was positive and if the rotation was clockwise so that $\Theta(t)$ decreased by 2π the spike was negative. It is assumed that if the end point of $R(t)$ has crossed the horizontal axis so that $\Theta(t)$ passes through π rad, then $\Theta(t)$ will continue to increase (or decrease) giving rise to a positive (or negative) spike. Thus, it is assumed that to ensure a spike output, it is not actually necessary to observe a complete rotation of $R(t)$ but it is adequate that there be guaranteed a rotation of at least π rad. If then at an instant of time $t = t_2$.

$$x(t_2) < -A$$

$$y(t_2) = 0$$

$$\text{and} \quad \dot{y}(t_2) > 0 \qquad\qquad ..(B-4)$$

then $\Theta(t)$ has decreased through π rad, and a negative spike will result.

The probability P_- of a negative spike occuring in the time interval t is the probability that the condition of eqn (B-4) are satisfied at some instant t_2 within the time interval t.

Then,

$$P_- = P\left[x(t_2) < -A, \; y(t_2) = 0. \; \frac{dy}{dt}\bigg|_{t_2} > 0\right] \quad (B-5)$$

The joint probability density of x, y and $\dot{y}$ is given by

$$f(x, y, \dot{y}) = \frac{e^{-x^2/2\eta B}}{\sqrt{2\pi}\,\eta B} \quad \times \quad \frac{e^{-y^2/2\eta B}}{\sqrt{2\pi}\,\eta B} \quad \times \quad \frac{e^{-\dot{y}^2/(2\pi^2/3)\,\eta B^3}}{\sqrt{(2\pi^3/3)\,\eta B^3}} \qquad \ldots (B\text{-}6)$$

Therefore,

$$P_{-} = \int_{-\infty}^{-A} dx \int_{0}^{\sim} d\dot{y} \int_{-\dot{y}/2}^{+\dot{y}/2} dy \quad f(x, y, \dot{y}) \qquad \ldots (B\text{-}7)$$

After an involved calculation it has been shown
(see Ref.11) that

$$P_{-} = \left[\frac{B}{4\sqrt{3}} \quad \operatorname{erf}_{c} \sqrt{\left(\frac{f_M}{B}\right) \times \left(\frac{S_i}{N_M}\right)} \right] \Delta t$$

In one second there are $\frac{1}{\Delta t}$ intervals in which a spike
might occur. Since P_{-} is the probability of occurence
of a negative spike within any such individual interval
the expected number of spikes in a second, N_{-} is

$$N_{\rightarrow} = \frac{P_{-}}{\Delta t}$$

$$= \frac{B}{4\sqrt{3}} \operatorname{erf}_{c} \sqrt{\frac{f_M}{B} \times \frac{S_i}{N_M}} \qquad \ldots (B\text{-}8)$$

It may be shown that if the frequency of the carrier is
offset by $\overline{\delta f}$ the number of spikes increases by an
amount $\overline{\delta N}$ given by

$$\overline{\delta N} = \overline{\delta f} \quad e^{-\left(\frac{f_M}{B}\right)\left(\frac{S_i}{N_M}\right)} \qquad \ldots (B\text{-}9)$$

if $\overline{\delta f} = \Delta f \cos 2\pi f_m t$ .. (B-10)

where Δf is the frequency deviation and $f_m =$ frequency of the modulation signal.

$$\overline{\delta f} = \frac{2 \Delta f}{\pi}$$ .. (B-11)

and

$$\overline{\delta N} = \frac{2 \Delta f}{\pi} \, e^{-\left(\frac{f_M}{B}\right) \times \left(\frac{S_i}{N_H}\right)}$$ .. (B-12)

APPENDIX-C

Short-time Frequency Analysis of Speech

The short-time transform· of the speech signal f(t)
is defined by

$$F(\omega, t) = \int_{-\infty}^{t} f(\lambda) \, h(t-\lambda) \, e^{-j\omega\lambda} \, d\lambda$$

or

$$F(\omega, t) = e^{-j\omega t} \int_{0}^{\infty} f(t-\lambda) \, h(\lambda) \, e^{j\omega\lambda} \, d\lambda \qquad ..(C-1)$$

where the weighting function h(t) is the impulse
response of physically realizable linear system and is
shown in Fig.C-1.
Eqn. (C-1) may be rewritten as

$$F(\omega, t) = e^{-j\omega t} \left[\int_{0}^{\infty} f(t-\lambda) h(\lambda) \cos\omega\lambda \, d\lambda + j \int_{0}^{\infty} f(t-\lambda) h(\lambda) \sin\omega\lambda \, d\lambda \right]$$

$$= [a'(\omega, t) + jb'(\omega, t)] \, e^{-j\omega t} \qquad ..(C-2)$$

F(ω, t) is the short-time power spectrum.
The measurement of F(ω, t) can therefore be implemented
by the operation shown in Fig.C-2.

It may be shown that the quantity F(ω, t) is
essentially the time envelope of either a'(ω, t) or
b'(ω, t) [or either a(ω, t) or b(ω, t)] provided the
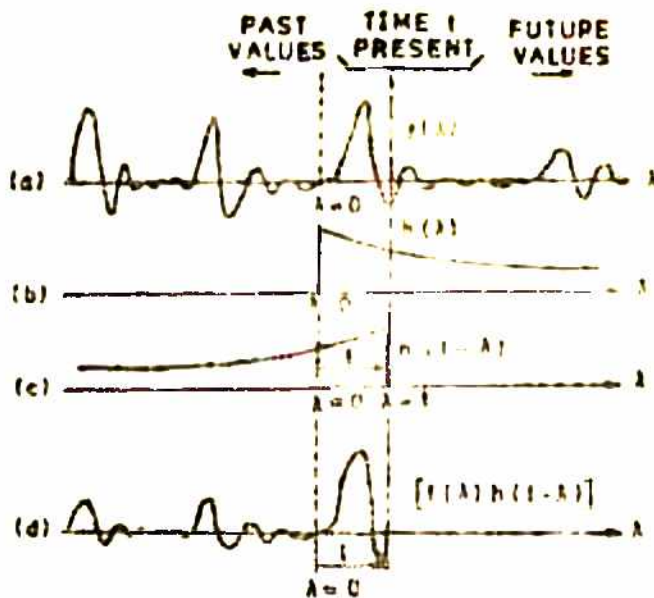spectrum of h(t), does not overlap ω. The envelope

FIG C 1.WEIGHTING OF AN ON GOING SIGNAL f(t)
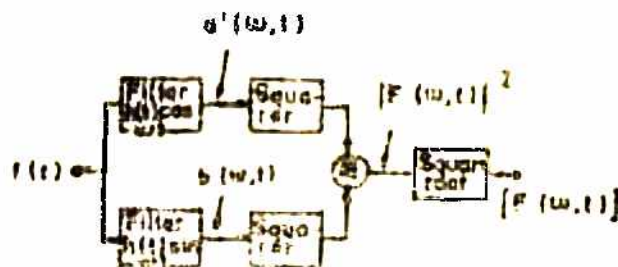BY A PHYSICALLY REALIZABLE TIME
WINDOW h(t)

FIG.C2.A METHOD OF MEASURING SHORT TIME
AMPLITUDE SPECTRUM |F(w,t)|
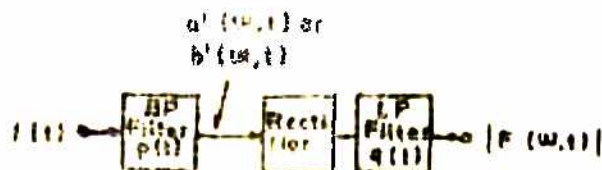

FIG.C 3.PRACTICAL MEASUREMENT OF THE
SHORT TIME SPECTRUM F(w,t).

can be approximated electrically by developing the envelope of either filter branch of Fig.C-2. This is conventionally done by linear rectification and low pass filtering as shown in Fig.C-3.

It has been shown that the impulse response of the filter of Fig.C-3 whose frequency domain response be

$$P(W) = 1 \quad (W_0 - W_1) \leqq W \leqq (W_0 + W_1)$$

$$P(W) = 1 \quad -(W_0 + W_1) \leqq W \leqq -(W_0 - W_1)$$

$$= 0 \quad \text{elsewhere}$$

is given by

$$P(t) = \left(\frac{2W_1}{\pi}\right)\left(\frac{\sin W_1 t}{W_1 t}\right) \cos W_0 t$$

$$= h(t) \cos W_0 t$$

Thus, the time window of this filter is the $\frac{\sin X}{X}$ envelope of the impulse response. The time between initial zeros of the envelope is taken as the effective duration, D, of the time window. If the bandwidth of the filter is 50 Hz, $D = 40$ msec $= \frac{2}{B}$.

APPENDIX-D

Second Order Phase Lock Loop Analysis

The block schematic of a second order PLL is shown in Fig.D-1.

From Fig.D-1 we have,

$$\Psi(s) = \Phi(s) - \frac{K_d K_O\ H(s)\Psi(s)}{s}$$

or $\quad \Psi(s) \cdot \left[1 + \frac{K_d K_O H(s)}{s} \right] = \Phi(s) \qquad \text{.. (D-1)}$

$$\frac{\Psi(s)}{\Phi(s)} = \frac{s}{s + K_d K_O\ H(s)} \qquad \text{.. (D-2)}$$

when,

$$H(s) = \frac{1}{1 + RC s}$$

$$= \frac{1}{1 + s\tau} \qquad \text{.. (D-3)}$$

where $\tau = RC$

Equation (D-2) becomes

$$\frac{\Psi(s)}{\Phi(s)} = \frac{s^2 + s/\tau}{s^2 + s/\tau + K_d K_O/\tau} \qquad \text{.. (D-4)}$$

Comparing the Denominator of eqn (D-4) with $s^2 + 2\zeta W_n s + W_n^2$ we have,

$$W_n = \sqrt{\frac{K_d K_O}{\tau}} \qquad \text{.. (D-5)}$$

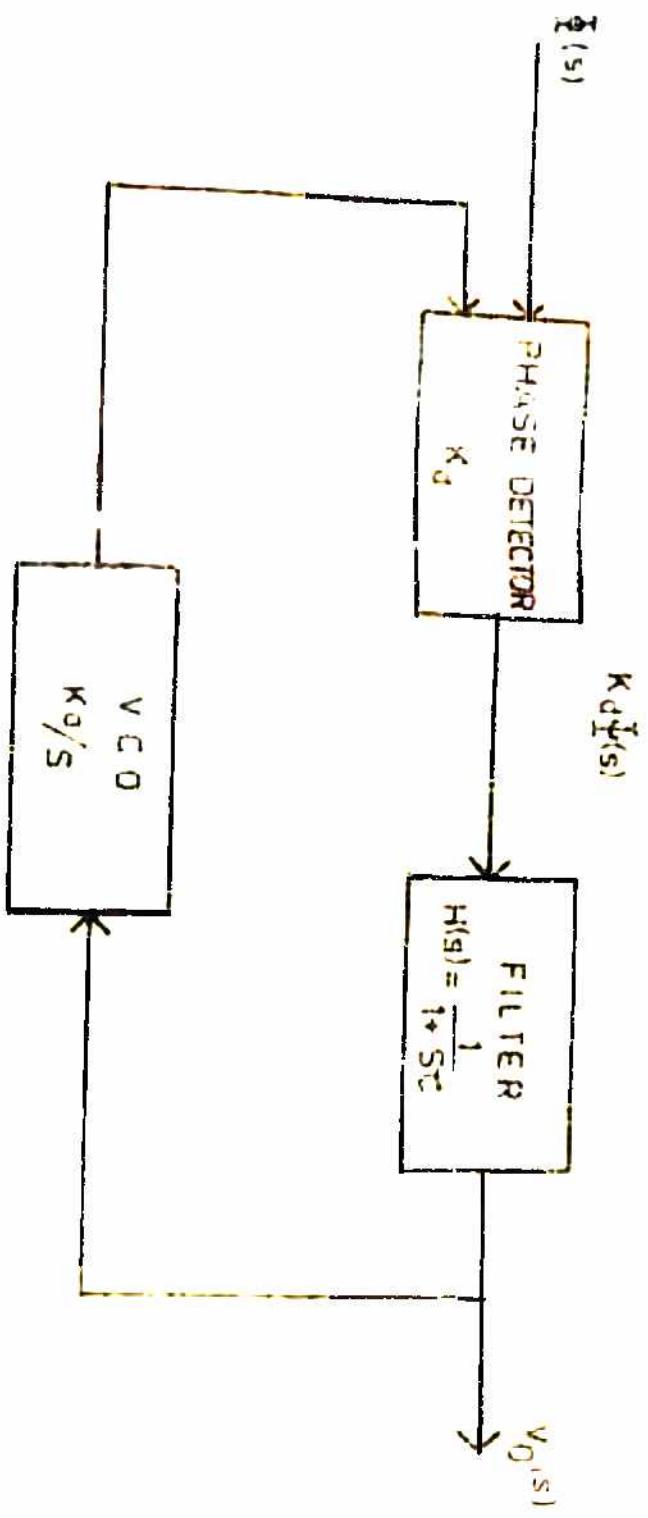FIG D: BLOCK SCHEMATIC OF SECOND ORDER PLL

$$\zeta = \frac{1}{2\tau_d K_0 \tau} \qquad \qquad .. \ (D\text{-}6)$$

It has been shown [30] that the lock-up time T_p for a second order PLL is given by,

$$T_p = \frac{(\Delta \omega)^2}{2\zeta \omega_n^3} \qquad \qquad .. \ (D\text{-}7)$$

where $\Delta \omega$ is the initial frequency offset. An absolute maximum allowable sweep rate is,

$$\Delta \omega < \omega_n^2 \qquad \qquad .. \ (D\text{-}8)$$

For $\zeta > 1$, the allowable sweep rate for probability of lock $P = 1$ is given by

$$\Delta \omega < \frac{1}{2} \omega_n^2 \qquad \qquad .. \ (D\text{-}9)$$

References

1. F.L.H.M.Stumpers, 'Theory of Frequency Modulation Noise' Proc.I.R.E. vol.36, pp1081-1092, September, 1948.

2. E.Peterson, 'Frequency detection and Speech Formants', J.Acoust. Soc. Am. vol.23, pp668-674, 1951.

3. J.J.Dubnowski, 'Real-Time Digital Hardware Pitch Detector', IEEE Trans, Acoust, Speech, Signal Processing vol.ASSP-24, No.1, pp2-8, February 1976.

4. Pollack and J.C.R.Licklider, 'Effects of Differentiation Integration and Infinite peak clipping upon the Intelligibility of Speech', J.Acoust. Soc. Am. vol.20, pp42-51, 1948.

5. R.J.Niederjohn, 'A Mathematical Formulation and comparison of Zero-crossing Analysis Techniques which have been applied to Automatic Speech Recognition' IEEE Trans. ASSP, vol.23, No.4, pp373-380, August, 1975.

6. J.S.Bendat, 'Principles and Applications of Random Noise Theory', New York Wiley, 1958.

7. A.Papoulis, 'Probability, Random Variables and Stochastic Processes', New York, Mc-Graw Hill, 1965.

8. S.O.Rice, 'Noise in FM Receivers', Chap.25, pp375-424, in 'Proceedings, Symposium on Time Series Analysis', M.Rosenblatt, (Ed.) John Wiley and Sons Inc., New York, 1963.

9. K.Leentvaar and J.H.Flint, 'The Capture Effect in FM Receivers' IEEE Tran. COMM. vol.COM-24, No.5, May 1976.

10. D.L.Schilling and H.Taub, 'Principles of Communication Systems' Mc Graw-Hill, 1971.

11. M.Schwartz 'Information Transmission, Modulation and Noise' Mc Graw-Hill, 1970.

12. C.S.Weaver, 'A New Approach to the Linear Design and Analysis of phase locked loops' IRE Trans. Space Electron. Telemetry, vol.SET-5 pp166-178, Dec, 1959.

13. A.W.Moore, 'Phase-locked loop for motor speed control' IEEE. Spectrum, pp.61-67, April, 1973.

14. A.Bilotti and R.S.Pepper, 'A Monolithic Limiter and Balanced Disciminator for FM and TV Receivers' IEEE Trans BTR-13, pp60-65, Nov. 1967.

15. J.L.Flanagan, 'Speech Analysis, synthesis and perception Springer-Verlag, New York, 1965.

16. N.R.French and J.C.Steinberg, 'Factors Governing the Intelligibility of speech sounds', J.Acoust. Soc. Am, vol.19, pp90-119, 1947.

17. C.G.M.Fant, 'On the Predictability of Formant Levels and Spectrum Envelopes from Formant Frequencies'. For Roman Jakobson, Mouton and Co., the Hague, pp109-120, 1956. Reprinted in 'Speech Synthesis' Benchmark Papers is Acoustics, Dowden, Hutchinson and Ross Inc, 1973.

18. J.D.Markel, 'Digital Inverse Filtering - A New Tool for Formant Trajectory Estimation', IEEE Trans. Audio Electroacoust, vol.AV-20, pp129-137, 1972.

19. J.Suzuki, Y.Kadokawa and K.Nakata, 'Formant Frequenc., Extraction by the Method of Moment Calculations', J.Acoust. Soc. Am, vol.35, pp1345-1353, 1963.

20. R.W.Schafer and L.R.Rabiner, 'System for Automatic Formant Analysis for Vocied Speech', J.Acoust. Soc.Am. vol.47, pp634-648, 1970.

21. T.Sakai and S.Doshita, 'Automatic Speech Recognition System for Conversational Sound', IEEE Trans. Electro Comp. vol.12, pp835-846, 1963.

22. J.L.Flanagan, 'Automatic Extraction of Formant Frequencies from Continuous Speech', J.Acoust. Soc.Am, Vol.28, pp110-118, 1956.

23. R.J.Niederjohn, 'Formant Tracking, Categorization, Segmentation and Phoneme Recognition of continuous speech', Ph.D. dissertation, Amberst, Mass. Grant No.N5 08306 Report 7, May 1971.

24. W.A.Munson and H.C.Montgomery, 'A Speech Analyser and synthesizer', J.Acoust. Soc. Am, vol.22, pp678(A) 1950.

25. Millman and Halkias, 'Integrated Electronics', Mc Graw Hill, 1972.

26. Signetics, 'Linear IC_s Application Manual'.

27. F.M.Gardner, 'Phase lock Techniques' Wiley Inc. 1966

28. F.L.H.M.Stumpers, 'Comparison of Modulation systems pp.245-270 in 'New Directions in Signal Processing in Communication and Control' J.K.Skwirzynski (Ed) Noordhoff International Publishing, Leydon, The Netherlands 1975.

29. CCIR Report 455-1, in CCIR Plenary Assembly, Geneva 1974, vol.IV, ITU, Geneva, 1974.

30. A Blanchard, 'Phase locked loops', John Wiley and Sons, Inc 1976.

31. V J Georgiou, 'Voltage-tuned Filter varies Centre Frequency Linearly', Electronics, pp 104,Nov 6,1972.