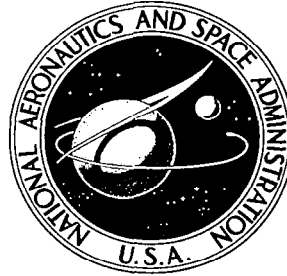


NASA CONTRACTOR
REPORT

NASA CR-682



NASA CR-682



TECH LIBRARY KAFB, NM

LOAN COPY: RETURN TO
AFWL (WLIL-2)
KIRTLAND AFB, N MEX

ADVANCED THRESHOLD
REDUCTION TECHNIQUES STUDY

by Cesar A. Filippi

Prepared by

ADCOM, INC.

Cambridge, Mass.

for Goddard Space Flight Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • JANUARY 1967



ADVANCED THRESHOLD REDUCTION TECHNIQUES STUDY

By Cesar A. Filippi

Distribution of this report is provided in the interest of information exchange. Responsibility for the contents resides in the author or organization that prepared it.

Prepared under Contract No. NAS 5-9011 by
ADCOM, INC.
Cambridge, Mass.

for Goddard Space Flight Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 - Price \$3.00



AUTHORSHIP

This report has been prepared by

Cesar A. Filippi

with the support of

Dr. Elie J. Baghdady

Dr. Phillip A. Bello

Charles J. Boardman

Robert L. Chufo

Richard A. DesRosiers

Harold G. Fritz

Alan C. Marshall

ABSTRACT

The threshold improvement capabilities of feedback and band-dividing demodulation techniques are investigated for PM or FM signals having a large index and/or deviation. The threshold performance of phase-locked and frequency-compressive feedback demodulators is analyzed and an optimum second-order phase-locked demodulator having a sinusoidal phase detector is developed and tested along with a conventional FM demodulator. Optimum higher-order loops are also derived as approximations to an optimum demodulator. Their threshold improvement capabilities are established and a third-order loop is developed and tested. The cycle slippage with a sawtooth phase detector is also investigated. Finally, several band-dividing systems based on either direct demodulation or signal estimation techniques are analyzed and compared to the phase-locked demodulator insofar as possible.

TABLE OF CONTENTS

<u>Section</u>	<u>Page</u>
I GENERAL INTRODUCTION	1
II THE PLL DEMODULATOR	3
2.1 Introduction.	3
2.2 The PLL Demodulator	4
2.3 Noise and Threshold Performance of a PLL	5
2.4 Discussion of Certain Published Analyses.	9
2.5 Conclusion	19
III THE SECOND ORDER PLL DEMODULATOR.	21
3.1 Introduction.	21
3.2 The Second Order PLL Demodulator	22
3.3 Case of a Rectangular PM Message Spectrum	24
3.4 Case of a Sinusoidal Message.	26
3.5 Second Order Loop Design.	29
3.6 Conclusion	31
IV HIGHER ORDER PLL DEMODULATORS	33
4.1 Introduction.	33
4.2 The Optimum Realizable PLL Demodulator	33
4.3 Approximation to the Optimum PLL Demodulator	34
4.4 The Third Order Loop	39
4.5 Third Order Loop Design	41
4.6 Applicability of Bode Filters to PLL Demodulators	47
4.7 Conclusion	50
V EXPERIMENTAL RESULTS	53
5.1 Introduction.	53
5.2 The Conventional FM Demodulator Results	61
5.3 Description of The PLL Demodulators	64
5.4 The Experimental PLL Thresholds	70
5.5 The Sawtooth Phase Detector	77
5.6 Sawtooth Detector Experiments	81
5.7 Cycle Slippage in the Sawtooth Phase Detector.	84
5.8 Conclusion	91

TABLE OF CONTENTS (Cont.)

Section	Page
VI THE FCF DEMODULATOR	95
6.1 Introduction	95
6.2 The FCF Demodulator	95
6.3 Threshold Performance and Design of FCF Demodulators	96
6.4 Comparison of FCF and PLL Demodulators	99
6.5 Conclusion	101
VII BAND-DIVIDING DEMODULATION	103
7.1 Introduction.	103
7.2 Comparison with the Second Order PLL Demodulator.	103
7.3 Effect of Mis-Selection Noise in Systems with Filter Overlap.	104
7.4 PLL Band-Dividing	108
7.5 Band-Dividing Signal Estimation Techniques	114
7.6 Conclusion	120
APPENDIX A	121
REFERENCES.	131

LIST OF ILLUSTRATIONS

Figure		Page
1	System representation of Eq. (2.1)	4
2	General phase-transfer model of a PLL (sinusoidal phase detector).	6
3	Linear phase-transfer model of a PLL.	6
4	Slip of a cycle	9
5	Special model of a PLL demodulator	10
6	Replot of data in Sanneman and Rowbotham's Fig.12	16
7	Compression vs (a, γ)	28
8	Second order loop filter	30
9	Characteristics of optimum demodulator	35
10	Threshold curves for the approximations to the optimum demodulator	38
11	Configuration of active filter network	44
12	Circuits for synthesizing impedances	45
13	Gain and phase shift of a Bode filter with 30° phase margin	49
14a	Frequency response of the 20 kc i-f filter.	54
14b	Frequency response of the 50 kc i-f filter.	55
14c	Frequency response of the 100 kc i-f filter	56
15	Frequency response of output lowpass filters	57
16	Effect of notches on the 6 kc lowpass filter	58
17	Manual demodulator threshold test system	60
18	Automatic demodulator threshold test system	62
19	Block diagram of second order PLL	65
20	Sinusoidal phase detector circuit	65
21	Second order loop filter and DC amplifier circuits.	66
22	VCO circuit	66
23	VCO linearity	68
24	VCO amplitude vs. frequency.	69
25	Third order loop filter circuit	70

LIST OF ILLUSTRATIONS (Continued)

Figure		Page
26a	PLL ₂ tests ($\delta = 10, f_m = 1 \text{ kc}, B_{if} = 23 \text{ kc}$)	72
26b	PLL ₂ tests ($\delta = 20, f_m = 1 \text{ kc}, B_{if} = 54 \text{ kc}$).	73
26c	PLL ₂ tests ($\delta = 30, f_m = 1 \text{ kc}, B_{if} = 110 \text{ kc}$)	74
27a	PLL ₃ tests ($\delta = 10, f_m = 1 \text{ kc}, B_{if} = 23 \text{ kc}$).	75
27b	PLL ₃ tests ($\delta = 30, f_m = 1 \text{ kc}, B_{if} = 110 \text{ kc}$)	76
28	Higher order PLL threshold curves for $\sigma^2 = 1/10 \text{ rad}^2$	78
29	Block diagram of sawtooth phase detector.	80
30	Sawtooth phase detector circuitry	82
31	Sawtooth phase detector 20 ms/in characteristic	83
32	Sawtooth detector results (PLL ₂ , $\delta = 30$)	85
33	Sawtooth detector results (PLL ₃ , $\delta = 30$)	86
34	PLL ₂ oscillograms	87
35	PLL ₃ oscillograms	88
36	Sawtooth detector model during cycle slippage.	89
37	Second order PLL model	90
38	Waveforms during cycle slippage	92
39	Block diagram of FCF demodulator	95
40	General phase transfer model of an FCF demodulator.	97
41	Linear phase transfer model of the FCF demodulator.	98
42	Contours of $P'_{n2}/P_{n2} = 1$, for various N	107
43	$\omega_2 - \omega_1 \ll \Omega_{\text{pull-in}}$	109
44	$\omega_2 - \omega_1 \sim \Omega_{\text{pull-in}}$	109
45	$\omega_2 - \omega_1 \sim \Omega_{\text{pull-in}}, \omega_2 - \omega_0 > \Omega_{\text{pull-in}}$	110
46	Frequency pull-in time of a second order phase-locked loop	112
47	Block diagram of maximum likelihood estimator (case of no-memory processing).	115
48	Darlington system configuration	119
49	System operation	119

Chapter I

GENERAL INTRODUCTION

This report summarizes the work performed at ADCOM, Inc. under Contract NAS 5-9011 during the period of July 1, 1964 - June 30, 1965. The purpose of the program is to study both theoretically and experimentally some promising FM threshold reduction techniques, in particular the feasibility of reducing threshold 3 db below that of a second order phase-locked loop (PLL) demodulator.

The theoretical effort on the program was concerned with analytical studies of feedback and band-dividing demodulation techniques, while the experimental effort involved the investigation of different promising PLL demodulation schemes. A conventional FM demodulator was also included for comparison purposes. The modulation waveforms were assumed to be large index PM or FM signals.

The contents of the report can be summarized briefly as follows:

Chapter II motivates feedback demodulators and characterizes the threshold performance of PLL demodulators while illustrating some of the analytical difficulties involved in their study.

Chapter III presents the logical design of the second order PLL demodulator. The optimum noise vs distortion compromise is established and lower bounds on the loop noise bandwidth and threshold are determined.

Chapter IV establishes the theoretical improvement capabilities of higher order PLL demodulators as approximations to an optimum demodulator. The logical design of a third order loop is discussed in detail.

Chapter V presents the experimental results obtained with the second and third order loops designed in the previous chapters. Also, this chapter discusses

the experimental results obtained when the sinusoidal phase detector is replaced by one having a sawtooth characteristic.

Chapter VI presents the frequency-compressive feedback (FCF) demodulator and discusses its threshold performance and logical design, while illustrating some of its uncertainties.

Chapter II

THE PLL DEMODULATOR

2.1 Introduction

The interest in feedback demodulator structures can be derived from general considerations. The Youla approach is to assume a finite-duration message (having gaussian statistics) and additive gaussian noise and then derive the optimum waveform that maximizes the a posteriori probability of detecting the message.¹

The solution to the problem results in a set of implicit integral equations which characterize the feedback structure. The formulation of the problem requires the entire message to be present and the structures obtained are usually unrealizable in real time. Still, the approximation with realizable systems can yield a real-time demodulator. The problem of the realizable maximum-likelihood demodulator has not yet been solved analytically.

In the case of FM signals, the integral equation in question may be found to be

$$\hat{m}(t) = \frac{1}{N_o} \int_0^T R_{mm}(\xi - t) \sin \left[\omega_c \xi + \int_{-\infty}^{\xi} \hat{m}(x) dx \right] \cdot m_i(\xi) d\xi, \quad (2.1)$$

$$0 \leq t \leq T$$

where $m(t)$, $m_i(t)$ and $\hat{m}(t)$ respectively represent the message, input (message plus noise) and output signals, N_o is the noise density assumed white, $R_{mm}(\tau)$ is the message autocorrelation function, ω_c is the carrier and $\int_{-\infty}^{\xi} \hat{m}(x) dx$ is the output FM signal. A structure that performs this operation is shown in Fig. 1. The $\sin []$ term in Eq. (2.1) represents the VCO output and is mixed with the input signal, and this product is filtered by $R_{mm}(\tau)$. This filter is unrealizable and the practical extension is to select realizable filters.

The PLL demodulator follows directly from Fig. 1. The FCF demodulator is more complicated since part of the filtering takes place

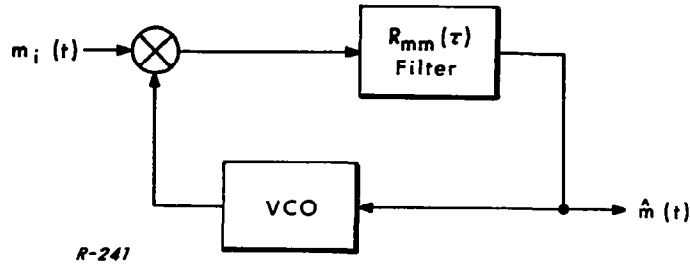


Fig. 1 System representation of Eq. (2.1).

at bandpass. Still, the overall closed-loop action is very similar to the PLL and analogous system equations result.

2.2 The PLL Demodulator

The analytical support of the PLL as an FM demodulator has been presented. We will now take a closer look at the PLL system and establish its signal-tracking and noise-rejection capabilities that eventually limit its threshold performance. The PLL system is best analyzed by a phase transfer model, a consequence of the fact that its behavior is tightly governed by the locking (phase error) conditions existing in the loop. Its threshold phenomena can be estimated from the statistics of the phase error.

The i-f input is assumed to be the sum of a CW signal plus a narrow-band gaussian random process, and the noise is decomposed into quadrature and cophasal components about the signal phase. This composite input may be written as

$$e_{if}(t) = E_c \left\{ \left[1 + \frac{x_c(t)}{E_c} \right] \sin \left[\omega_c t + \phi_s(t) \right] + \frac{x_q(t)}{E_c} \cos \left[\omega_c t + \phi_s(t) \right] \right\} \quad (2.2)$$

where E_c is the carrier amplitude, ω_c is the carrier frequency, $\phi_s(t)$ is the signal phase, and $x_q(t)$ and $x_c(t)$ are the quadrature and cophasal envelopes of the noise. Notice that the noise is decomposed about the modulated carrier phase and that the mean-square values of $x_q(t)$ and $x_c(t)$ will be similar to that of the original noise as long as the signal and noise are uncorrelated.

The most common type of PLL demodulator uses a sinusoidal phase detector whose operation may be characterized as multiplying its i-f and VCO inputs. If the VCO output signal is represented by $E_o \cos[\omega_c t + \phi_o(t)]$, where $\phi_o(t)$ represents the PLL estimation of the i-f input phase, then the phase detector output is proportional to

$$\left[1 + \frac{x_c(t)}{E_c}\right] \sin \left[\phi_s(t) - \phi_o(t)\right] + \frac{x_q(t)}{E_c} \cos \left[\phi_s(t) - \phi_o(t)\right] \quad (2.3)$$

where high-frequency terms have been omitted since they are filtered within the loop. The phase transfer model of the PLL is then that shown in Fig. 2 where $F(p)$ represents the (unity d-c gain) loop filter, K is the loop gain and the VCO is assumed to have a linear characteristic. In general, the PLL behaves as a randomly time-variant, nonlinear system. Notice that even in the absence of noise the system is still nonlinear (but time invariant) and can only be linearized if small phase error conditions exist.

2.3 Noise and Threshold Performance of a PLL

In the absence of noise, the loop design criterion is to provide the best possible reproduction of the signal phase at the VCO output. The model is then that of Fig. 3 with $x_q(t) = 0$ as long as the signal tracking error, denoted as linear distortion error, satisfies $\sin \phi(t) \approx \phi(t)$. If this relation does not hold then a sinusoidal operator on the phase error is required and the error becomes a nonlinear distortion error.

Assume now that a small amount of noise is added under linear distortion error conditions. Notice that the validity of the model of Fig. 3 not only

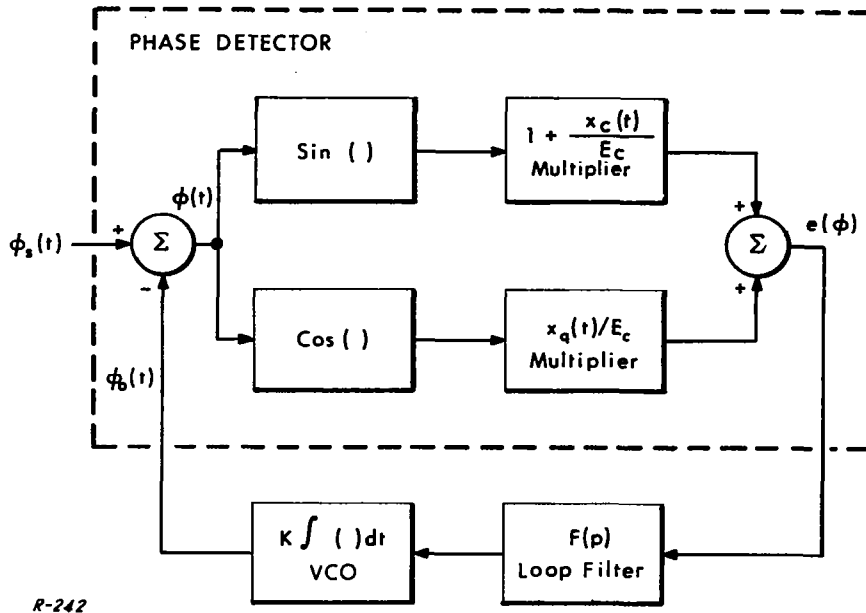


Fig. 2 General phase-transfer model of a PLL (sinusoidal phase detector).

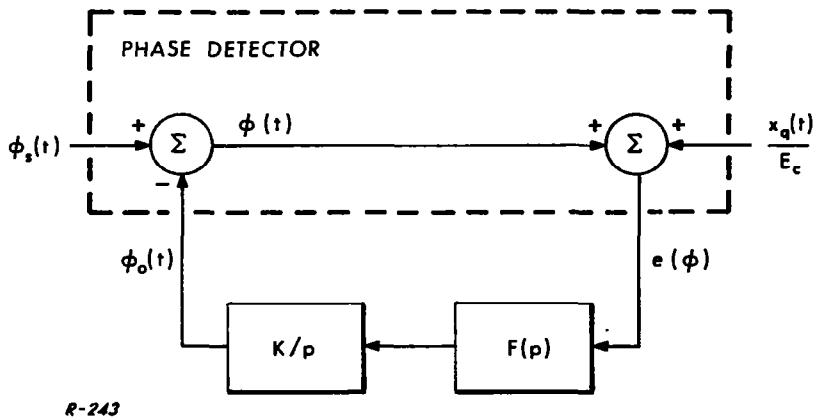


Fig. 3 Linear phase-transfer model of a PLL.

requires $\sin \phi \approx \phi$ and $x_c(t) \ll E_c$ but also $\cos \phi(t) \approx 1$, the latter requirement being stronger than the linear approximation of the sine. If all these conditions are met, then the phase error is now formed by the linear distortion error plus a noise error caused by the contribution of $x_q(t)$ to the VCO output phase $\phi_o(t)$.

The maximum static error the loop can tolerate without slipping a cycle is π rad. However, once the error exceeds $\pi/2$ then the system operates in the unstable region of the sinusoidal detector characteristic and a cycle is slipped in returning to a stable point. The noise error is random in nature and will have a finite probability of exceeding $\pi/2$. For large SNR conditions at the VCO output (small noise errors) this probability is small and the loop remains essentially in a locked state. Moreover, if the SNR is large enough, the phase error excursions in the nonlinear region of the sinusoidal detector are also rare and the linear model of Fig. 3 is valid almost all the time.

As the noise is increased, the small phase error approximation breaks down and the model of Fig. 2 becomes necessary. The contribution from random terms to the VCO output gets larger and the relative frequency of phase peaks greater than $\pi/2$ is larger. As the noise is increased further, this frequency increases to the point where the loop slips cycles often enough to remain essentially out of lock.

On this basis, it seems that two thresholds are pertinent in this system. The first one, defined as linear threshold, represents the departure from the linear model of Fig. 3 and will be characterized in a db-db plot of $(\text{SNR})_{\text{out}}$ vs $(\text{SNR})_{\text{in}}$ by the departure from the linear relation existing above threshold. The second one, defined as nonlinear threshold, represents the transition region where the output SNR becomes unacceptable according to the user fidelity criterion. It is evident that this last threshold depends on the individual application in question and does not lend itself to a general systems comparison criterion.

We have included it in this discussion to warn the reader that the system performance below the linear threshold is not necessarily unreliable, as will be illustrated in some experimental results.

The linear threshold also suffers from some arbitrariness when defining the phase error conditions at which the linear model ceases to be valid almost all the time. A knowledge of the phase error statistics will be very helpful in establishing a criterion since the probability of phase peaks exceeding $\pi/2$ could be formulated. However, the random noise error can be assumed to have gaussian statistics only as long as the linear model is valid and the evaluation of the large error statistics are a complex analytical problem. The use of a given mean-square value (or other measure) as a threshold phase error will be really informative when we can assign to it its correspondent relative frequency of peaks above a certain value, and to do this we must know the statistics.

In any case, once the critical error ($\pi/2$ for a sinusoidal phase detector) is exceeded, the slipping of a cycle appears as a new disturbance (besides the already existing noise) in the FM output signal. This signal is given by

$$\dot{\phi}_o(t) = K F(p) e[\phi(t)] \quad (2.4)$$

within a proportionality constant, and the occurrence of a peak above the critical error introduces a component in $e[\phi(t)]$ during the transition to the new stable point. For instance, in the sinusoidal detector this component appears when the operating point moves along the negative slope of the sinusoid and back up to the new stable point of Fig. 4, while in a sawtooth detector it appears as a 2π step plus the term due to the return to the stable point as shown in Fig. 4.

We know that the transition $\Delta\phi(t)$ implies a net step of 2π whose shape is governed by the loop dynamics. This shape in turn governs the detector output $e[\phi(t)]$ and will eventually represent the FM disturbance. We cannot note that the sawtooth detector may be expected to show a larger amplitude disturbance

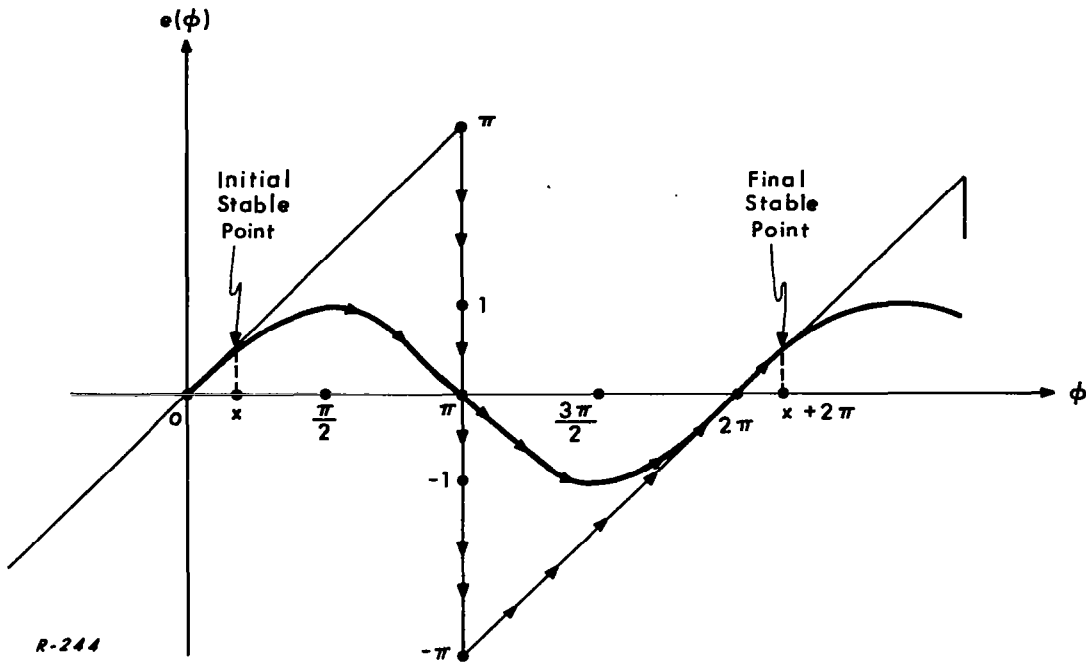


Fig. 4 Slip of a cycle.

according to its characteristic but also a narrower one since its slope is always larger than that of the sinusoidal detector and on this basis it is not clear which of the two cases will contribute more disturbance power per cycle slipped to the FM output. The sawtooth detector seems to have an advantage in its probability of occurrence of such a disturbance as a consequence of its largest critical error. In any case, the demodulated disturbance will cause a reduction in the output SNR and a departure from linearity in the threshold curves which can be used to establish a threshold measure.

2.4 Discussion of Certain Published Analyses

During the development of the program, ADCOM performed a theoretical study of certain approaches used by different authors in the analysis of PLL demodulators. We now summarize these investigations.

(a) The "independent" noise model

Perhaps the more relevant issue is the validity of the model shown in Fig. 5, where $n(t)$ is an independent gaussian process having the same statistics as the quadrature and cophasal components of the i-f input noise. It is very important to be aware of the assumptions inherent in this model due to its

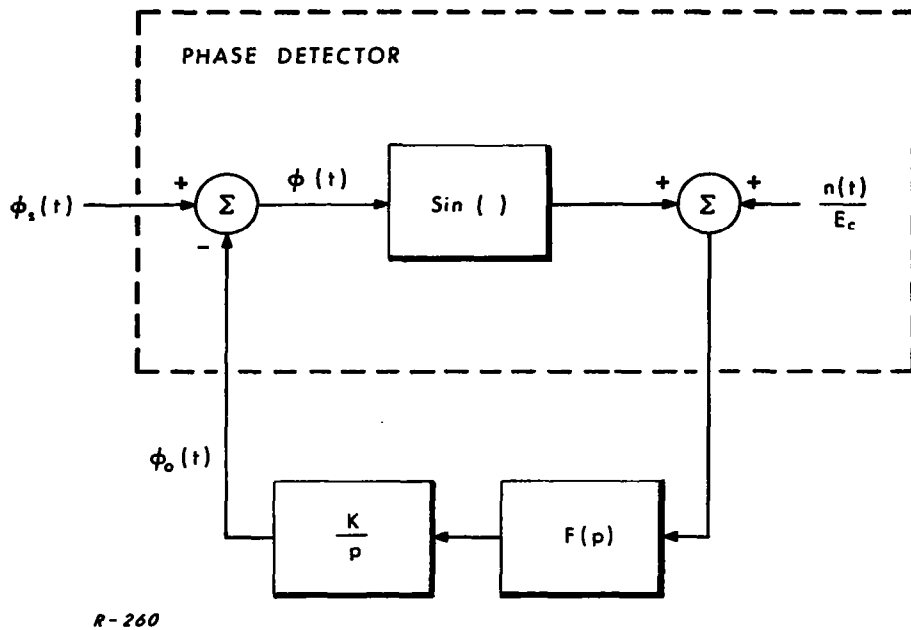


Fig. 5 Special model of a PLL demodulator.

popular use by various authors in different problems. In the absence of noise, there is clearly no distinction between this model and that of Fig. 2 with $n(t) = x_q(t) = x_c(t) = 0$.

In order to discuss the noisy case, we will first derive the exact expression of the random variable $n(t)$. Consider the i-f additive noise decomposed into quadrature and cophasal components about the unmodulated carrier phase so that Eq. (2.2) would read

$$e_{if}(t) = E_c \left\{ \sin \left[\omega_c t + \phi_s(t) \right] + \frac{\hat{x}_c(t)}{E_c} \sin \omega_c t + \frac{\hat{x}_q(t)}{E_c} \cos \omega_c t \right\} \quad (2.5)$$

where

$$\hat{x}_q(t) = x_q(t) \cos \phi_s(t) + x_c(t) \sin \phi_s(t) \quad (2.6)$$

$$\hat{x}_c(t) = -x_q(t) \sin \phi_s(t) + x_c(t) \cos \phi_s(t) \quad (2.7)$$

Thus, multiplication with the VCO signal $E_o \cos \left[\omega_c t + \phi_o(t) \right]$ yields a sinusoidal phase detector baseband output proportional to

$$\sin \left[\phi_s(t) - \phi_o(t) \right] - \frac{\hat{x}_c(t)}{E_c} \sin \phi_o(t) + \frac{\hat{x}_q(t)}{E_c} \cos \phi_o(t) \quad (2.8)$$

so that $n(t)$ in Fig. 5 is exactly given by

$$n(t) = -\hat{x}_c(t) \sin \phi_o(t) + \hat{x}_q(t) \cos \phi_o(t) \quad (2.9)$$

At this stage, the procedure is to claim that $n(t)$ is gaussian and has the same statistics as $\hat{x}_c(t)$ and $\hat{x}_q(t)$, and hence those of $x_c(t)$ and $x_q(t)$. The net effect is thus to move the baseband noise through the phase detector without any change in its statistics. After this, we can forget about Eq. (2.9) and use $n(t)$ as an independent input with the aforesaid statistics.

The claim essentially states that the input and VCO noise components are uncorrelated. The fact that this is not the case, as shown by Eq. (2.9), was properly acknowledged by Viterbi² in his discussion of the model in question.* The unknown issue is the actual effects of the approximation. We will try to illustrate some of the uncertainties with examples.

The output FM signal of the PLL demodulator satisfies Eq.(2.4). For the two models of Figs. 2 and 5, the phase detector outputs are respectively

$$e = \left(1 + \frac{x_c}{E_c}\right) \sin \phi + \frac{x_q}{E_c} \cos \phi \quad \text{for Fig. 2} \quad (2.10)$$

and

$$e = \sin \phi + \frac{n}{E_c} \quad \text{for Fig. 4} \quad (2.11)$$

where the time dependence is implicit. These two expressions can be equalized through Eq.(2.9) and the claim is that we could also use Eq.(2.11) while letting $n(t)$ be an independent process having the same statistics as $x_c(t)$ and $x_q(t)$.

As a first example of the assumptions involved in this claim, we now evaluate the mean-square phase detector output. The model of Fig. 5 yields

$$\overline{e^2} = \overline{\sin^2 \phi} + \frac{\overline{x^2}}{E_c^2} \quad (2.12)$$

where $\overline{x^2} = \overline{x_c^2} = \overline{x_q^2}$ while that of Fig. 2 yields

* Develet³ uses the model of Fig. 5 but does not offer any new comment about its validity while Van Trees⁴ "proves" gaussianity twice (in case "one does not want to use this type of argument") and both proofs depend on the a priori assumption of statistical independence between $n(t)$ and $\phi(t)$, which is not really the case.

$$\begin{aligned}
\overline{e^2} &= \overline{\sin^2 \phi} + \frac{x_c}{E_c} \left(2 + \frac{x_c}{E_c} \right) \overline{\sin^2 \phi} \\
&\quad + \left(\frac{x_q}{E_c} \right)^2 \overline{\cos^2 \phi} + 2 \frac{x_q}{E_c} \left(1 + \frac{x_q}{E_c} \right) \overline{\sin \phi \cos \phi}
\end{aligned} \tag{2.13}$$

and the last three terms of Eq. (2.13) will only yield the last term of Eq. (2.12) if $\phi(t)$, i. e., $\phi_o(t)$, is statistically independent of $x_c(t)$ and $x_q(t)$. The actual difference when this is not the case has not been discussed by any author. Notice that Eq. (2.10) can be rewritten as

$$e = \sin \phi + \frac{x_q}{E_c} + \frac{x_c}{E_c} \sin \phi - \frac{x_q}{E_c} \left(\frac{\phi^2}{2!} - \frac{\phi^4}{4!} \dots \right) \tag{2.14}$$

and the first two terms have contributions equivalent to those of Eq. (2.11) since $n(t)$ is claimed to have the same statistics as $x_q(t)$. The effect of neglecting the last two terms in Eq. (2.14) has not been discussed either.

Moreover, the logic behind the qualitative argument used to declare $n(t)$ independent of $\phi(t)$ (see Ref. 2) is not limited to sinusoidal phase detectors but would also apply to any narrowband loop using other detector characteristic by simply substituting $\sin \phi$ by the detector functional. For example, the output of a tanlock phase detector⁷ would appear as

$$e = \frac{(1+k) \sin \phi}{1+k \cos \phi} + \frac{n}{E_c} \tag{2.15}$$

whereas the actual output can be found to be

$$e = \frac{(1+k) \left[(1+x_c/E_c) \sin \phi + x_q/E_c \cos \phi \right]}{1+k \left[(1+x_c/E_c) \cos \phi - x_q/E_c \sin \phi \right]} \tag{2.16}$$

Notice that in this case the actual detector output cannot be even decomposed into a first term equal to the detector characteristic plus a second term that involves the noise explicitly, as in Eqs. (2.10) and (2.11). Again, the effects of using Eq. (2.16) instead of Eq. (2.15) when statistical independence between $\phi(t)$ and $n(t)$ is not allowed have yet to be evaluated.

On the basis of these examples, it seems that a more profound understanding of the VCO phase noise statistics and its dependence on $\phi(t)$ is needed to reject or permit the model of Fig. 5 in a given application. For instance, the use of Eq. (2.11), or any other quasi-linear approximation, in threshold evaluation is as good (or bad) as that obtained using a linear dependence on ϕ if the terms neglected in Eq. (2.14), such as $x_c \phi/E_c$ and $x_q \phi^2/2E_c$, are found to be as significant as terms such as $\phi^3/6$ and $\phi^5/20$ involved in the nonlinear relation. The answer to this question is not immediately obvious at all.

(b) The work of Sanneman and Rowbotham

The digital simulation of a second-order PLL by Sanneman and Rowbotham⁵ has produced an interesting result which was not fully developed in their paper. They used a sampled-data system to simulate the PLL; the sampling rate being chosen large enough to ensure correspondence with the continuous system. The computerized loop was supplied with a simulated carrier plus noise input and allowed to run (starting with prescribed initial conditions) until a break of lock occurred. A large number of runs were taken for each set of conditions so that an average time to unlock could be computed for various carrier-to-noise ratios and initial conditions.

The authors make an issue of the fact that the probability of unlocking in τ seconds can be described by the empirical law.

$$P(\omega_n \tau) = 1 - e^{-\omega_n \tau / \bar{\omega}_n \bar{\tau}} \quad (2.17)$$

where $\bar{\tau}$ is the average time to unlock and $\bar{\omega}_n$ is the resonance frequency. This function is simply the probability of one or more Poisson events in an interval $\omega_n \tau$ and would be obtained for a model where the probability of unlock is constant in any given time interval. In other words, their data supports the hypothesis that unlock events occur individually and collectively at random.

Of greater interest is the plot of average time to unlock versus phase error given in their Fig. 12. The abscissa of this plot is directly the rms phase error and the ordinate is the log of the normalized time to unlock. Viterbi² has obtained an expression for the mean time to unlock for the filterless PLL which takes the form

$$B_{\ell} \bar{\tau} \approx \frac{\pi}{4} \exp \left[\frac{2.0}{\sigma_{\phi}^2} \right], \quad (2.18)$$

where $2B_{\ell}$ is the noise bandwidth of the loop. A plot of $\log B_{\ell} \bar{\tau}$ versus $1/\sigma_{\phi}^2$ yields a straight line. This result immediately suggests a replotting of Sand R's Fig. 12 with an abscissa of $1/\sigma_{\phi}^2$. Such a plot is shown in our Fig. 6. The data now yields a straight line. For both initial conditions equal to zero, S and R's data is described by

$$\omega_n \bar{\tau} = 2 \exp \left[\frac{1.60}{\sigma_{\phi}^2} \right] \quad (2.19)$$

Equation(2.19) therefore appears to be an adequate empirical expression for the mean time to unlock in the second-order PLL, a result which Viterbi was unable to obtain theoretically. The use of Eq.(2.19) is further justified by the fact that it is very similar to the theoretical result in Eq.(2.18) for the filterless loop.

It is now possible to define a threshold value for σ_{ϕ}^2 in terms of a prescribed number of unlocks per second. For example, a second-order loop having a 16 kc natural frequency ($\omega_n = 105$) will have an average of 1 unlock per second at a mean-squared phase error of $\sigma_{\phi}^2 = 0.15 \text{ rad}^2$. For $\sigma_{\phi}^2 = 1$ as proposed by Develet^{3, 6} and others the corresponding unlock rate is 10,000 per second for $\omega_n / 2\pi = 16 \text{ kc}$.

These results are for an unmodulated carrier plus noise and zero steady-state phase error. If the mean -square phase error includes a random modulation - tracking error the unlock rate will certainly be no less than that given by Eq. (2.19).

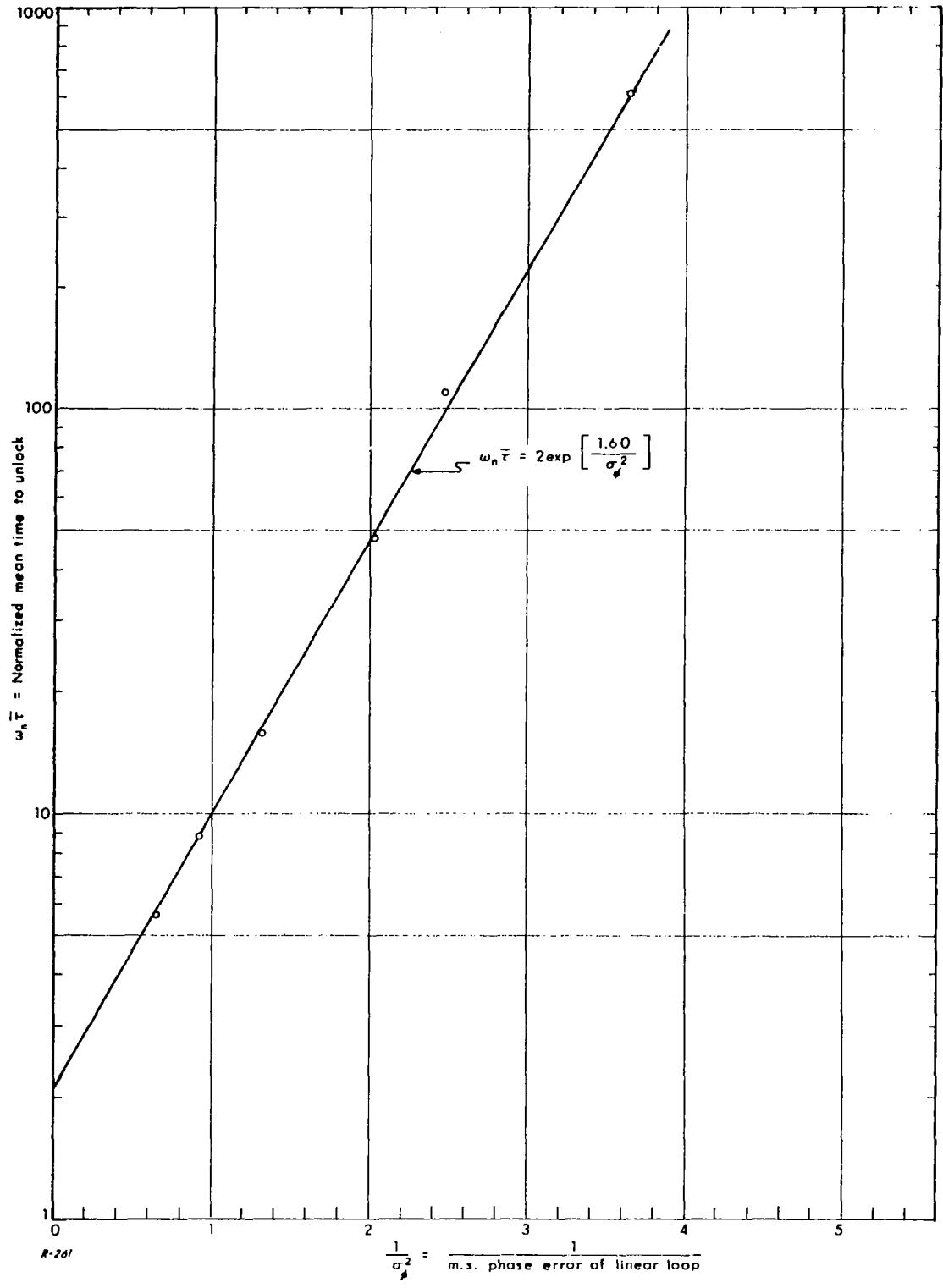


Fig. 6 Replot of data in Sanneman and Rowbotham's Fig. 12.

Therefore, for audio-type loop bandwidths and a threshold defined by an unlock rate of about 1 second the threshold occurs at a mean-square phase error of 0.15 rad^2 . The unlock rate is much more sensitive to the phase error than it is to the loop bandwidth; therefore, one need not be extremely precise in specifying the loop bandwidth or the exact unlock rate. For example, a 10 times decrease in unlock rate corresponds to $\sigma_\phi^2 = 0.123$ instead of 0.15 or about 0.9 db difference in SNR measured in the loop bandwidth.

On the basis of Sanneman and Rowbotham's results, it therefore appears that an appropriate threshold value of total mean-squared phase error in the second-order PLL is 0.15 rad^2 . At this level of error, the linearized model of the PLL is reasonably accurate. The average gain of the multiplier, as computed by Develet³, is 0.93 instead of 1. The threshold criteria of Develet^{3, 6} and Van Trees⁴ which are based upon a point of divergence of mean-square phase error in an approximate model yield unreasonably high rates of unlock. The divergence of phase error (or gross unlock in some sense) is a phenomenon associated with the particular model being used. This phenomenon of the model is not an adequate description of the random breaks of lock which are observed to occur near the threshold.

(c) The work of Balodis

We will first discuss the work of Balodis⁷ regarding the experimental threshold improvement capabilities of a tanlock phase detector over a sinusoidal one in a second order loop. In particular we will refer to Balodis' Fig. 9 which applies to a 10 cps sinusoidal modulation. The curves represent threshold curves for the different systems under comparison: the minimum SNR (defined in some bandwidth) that allows above-threshold operation can be determined in terms of the modulation index and the total (distortion plus noise) loop phase error that characterizes the threshold, and a choice of a value for this threshold error then yields the curves. These curves are analogous to those presented in Chapter IV except for the bandwidth in which the SNR is defined.

A problem present in Balodis' work is that he is using the same loop parameters for different modulation indices. While this is admissible for the loop gain and the damping factor, the noise bandwidth (or resonance frequency) should vary with the index in order to achieve an optimum performance (see the results of Chapter III). Thus, for any proper choice of threshold error, the curves of Balodis' Fig. 9 are only optimum at one index and their threshold performance can be improved at other indices by varying the noise bandwidth. For instance, on the basis of a linear model, Chapter III derives the optimum noise bandwidth for sinusoidal modulation in terms of the index δ and the rms threshold error σ as

$$B_n = 1.26 \left(\frac{\delta}{\sigma}\right)^{1/2} \omega_m \text{ cps} \quad (2.20)$$

where ω_m is the modulation frequency in radians. In turn, Balodis' noise bandwidth satisfies

$$B_n = 2.68 \omega_m \quad (2.21)$$

so that if we assume his design as optimum, then we can equate Eqs. (2.20) and (2.21) to yield

$$\sigma = 0.225 \delta \quad (2.22)$$

Thus, for indices of 2, 5 and 10 such "optimum" design would result in rms threshold errors of 0.45, 1.125 and 2 rads (based on a linear model). On this basis, it is no wonder that the sinusoidal PLL requires an infinite SNR for indices larger than 10 according to Balodis' results since it can only handle 1.57 rad. The effect of properly varying the bandwidth when the index increases will be to remove the progressive increase of the slope in Balodis' curves. Hence, we must claim that the threshold improvement capabilities of the tanlock over the sinusoidal phase detector remain unknown except at one index in the curves. Even at this index, the use of the same loop design for the sinusoidal

and tanlock detectors cannot be optimum for both since they must have different threshold errors, i. e. , the tanlock accepts more error before going out of lock, and an equation such as Eq. (2.20) will yield different noise bandwidths. All we can say, again based on a linear model, is that the systems in Fig. 9 of Balodis will only be optimum at the index

$$\delta = 4.44 \sigma \quad (2.23)$$

as given by Eq. (2.22), where σ is the threshold error which should be assumed different for the sinusoidal and tanlock detectors (perhaps even for the $k = 0.7$ and $k = 0.9$ tanlocks). For example, the choice of $\sigma = 1/2$ rad as the sinusoidal detector threshold will determine the optimum index as $\delta = 2.22$ and at this point the tanlock (whether optimum or not) is giving about 1 db of improvement. Even if we use $\sigma = \pi/2$ for the sinusoidal detector we cannot get further than $\delta = 7$ and any higher value of σ is clearly absurd for this detector. Any comparison with the tanlock at other indices is unfair.

2.5 Conclusion

In this chapter we have presented the fundamental concepts in the theory of PLL demodulators with particular attention to its threshold behavior. The PLL demodulator is motivated as a realizable approximation to the maximum-likelihood FM demodulator. Its analysis is best carried out in terms of phase transfer models since its behavior is tightly governed by its locking error conditions. In general, the loop acts as a nonlinear time-variant model that can only be reduced to a linear time-invariant model under small error, high SNR conditions, a requirement that is statistical in nature.

The PLL threshold occurs when the cycle slippage due to phase error peaks larger than those allowed by the phase detector become often enough. This criterion is somewhat arbitrary and its formulation can only be attempted in terms of departure from the linear time-invariant model due to analytical complexity considerations. The cycle slippage appears as an additional disturbance

(besides random noise) in the demodulated FM signal thus reducing the output SNR and causing departure from linearity in the standard threshold plots. On this basis, a promising experimental approach for the determination of the threshold is to identify it to a preset departure from linearity.

An evaluation of some pertinent published analyses is also included. The main issues discussed are:

- a) The validity of using a nonlinear time-invariant model where independent gaussian baseband noise is added at the output of the phase detector. The uncertainties and limitations introduced by this model are illustrated.
- b) The work of Sanneman and Rowbotham regarding the statistics of the unlock time of a PLL. Their analysis is extended to show the dependence of the average unlock time as a function of the mean-square error existing in the loop on the basis of a linear model.
- c) The work of Balodis regarding the experimental comparison of the threshold improvement capabilities of second order loops using sinusoidal and tanlock phase detectors. The fact that the experiments maintained a constant noise bandwidth as the modulation index varied is shown to introduce a serious limitation in the results.

Chapter III

THE SECOND ORDER PLL DEMODULATOR

3.1 Introduction

In the present chapter, we will analyze the second order PLL demodulator and optimize its threshold performance. The linear, time invariant phase model of Fig. 3 is assumed as a starting point. This inherently assumes an approximation to the problem, since the loop phase noise and the additive input noise will show some statistical peaks that exceed the linearity and time invariance conditions with a finite probability such that the model in question is not valid always. However, such behavior is assumed to have a small frequency of occurrence so that the model may be assumed as valid "almost all the time." The threshold phenomenon essentially represents the breakdown of this model and is characterized by the occurrence of the statistical peaks with a frequency larger than allowable.

If the closed-loop phase transfer function referred to the VCO output is denoted by $H(s)$, i. e. ,

$$H(s) = \frac{KF(s)}{s+KF(s)} \quad (3.1)$$

then the loop phase error in tracking the message term ϕ_s is given by

$$\begin{aligned} \phi(s) &= \phi_s(s) - \phi_o(s) = \phi_s(s) - H(s) [\phi_s(s) + \phi_n(s)] \\ &= [1 - H(s)] \phi_s(s) - H(s) \phi_n(s) \end{aligned} \quad (3.2)$$

where $\phi_n(s) = x_q(s)/E_c$ represents the equivalent phase noise input. The first term is the message distortion error $\phi_{e,d}$ while the second term is the noise error $\phi_{e,n}$. These two error contributions must be optimally weighted under some convenient criterion in order to minimize the total error. Thus, the loop design depends on the input message and noise characteristics.

Two different modulating signals will be considered: a PM message having a rectangular spectrum and a sinusoidal message. The first case will have more generality in the sense that it will yield a near-optimum design for a wide class of modulating signals. The second case will in turn be more useful in the experimental work since it will simplify keeping a close tag on the analysis assumptions as limitations after a system is built. The interesting note is that the optimum second order loop design for these two messages is extremely close to being almost equal, as will be made evident in the following sections.

The input noise will be assumed to be narrow-band gaussian with zero mean and a flat power density spectrum of Φ watts/cps (one-sided). The process $x_q(t)/E_c$ characterizing the input phase noise will have the same properties except for a $1/E_c^2$ reduction in its power density spectrum. On this basis, the mean-square noise error is given by

$$\overline{\phi_{e,n}^2} = \frac{\Phi}{2S_1} \int_{-\infty}^{\infty} |H(j\omega)|^2 \frac{d\omega}{2\pi} = \frac{\Phi B_n}{2S_1} \text{ rad}^2 \quad (3.3)$$

where $S_1 = E_c^2/2$ is the carrier power and $B_n = \int_{-\infty}^{\infty} |H(j\omega)|^2 df$ is the equivalent phase noise bandwidth of the system in cps.

3.2 The Second Order PLL Demodulator

The second order loop is characterized by a loop filter of the form

$$F(s) = \frac{1 + \tau_1 s}{1 + \tau_2 s} \quad (3.4)$$

so that the set of loop design parameters is (K, τ_1, τ_2) . The equivalent noise bandwidth for this loop can be found to be

$$\begin{aligned}
B_n &= \frac{K}{2} \frac{1 + \frac{K\tau_1^2}{\tau_2}}{1 + K\tau_1} \quad \text{cps} \\
&\approx \frac{1}{2} \left(\frac{1}{\tau_1} + \frac{K\tau_1}{\tau_2} \right) \quad \text{cps}
\end{aligned} \tag{3.5}$$

At this point it is useful to notice that this bandwidth uniquely specifies the noise error for a given normalized input noise density Φ/S_1 and phase noise statistics. In particular, several second order loops having different design parameters (K, τ_1, τ_2) but the same noise bandwidth B_n will have the same mean-square noise error and their relative merits will only be determined by their message tracking (distortion error) capabilities. Notice the same statement is true independent of the order of the loop. On this basis, it is very useful to change the set of design parameters (K, τ_1, τ_2) to a new set that includes B_n as one of its parameters. A look at the approximate relation in Eq. (3.5) suggests the transformation

$$\frac{1}{\tau_1} = aB_n, \quad \frac{K\tau_1}{\tau_2} = (2-a)B_n, \quad 0 < a < 2 \tag{3.6}$$

which defines the new set of design parameters (K, B_n, a) by

$$\tau_1 = \frac{1}{aB_n}, \quad \tau_2 = \frac{K}{a(2-a)B_n^2}, \quad \frac{K}{B_n} \gg a \tag{3.7}$$

To be precise, the exact noise bandwidth is actually given by $B_n / (1 + (aB_n/K))$ due to the approximate sign in Eq. (3.5). This expression approximates B_n for $K/B_n \gg a$, a reasonable relation since $0 < a < 2$.

3.3 Case of a Rectangular PM Message Spectrum

We will first consider the case where the input phase spectrum has the rectangular form.

$$S_m(\omega) = \begin{cases} \frac{\pi}{\alpha} & \text{for } |\omega| < \alpha \\ 0 & \text{otherwise} \end{cases} \quad (3.8)$$

and the loop will be optimized using a mean-square criterion. The mean-square distortion error, expressed in terms of the new design parameter, is given by

$$\begin{aligned} \overline{\phi_{e;d}^2} &= \frac{\pi}{\alpha} \delta^2 \int_{-\infty}^{\infty} |1-H(j\omega)|^2 \frac{d\omega}{2\pi} \\ &\approx \frac{\delta^2 \alpha^4}{5a^2(2-a)^2 B_n^4} \text{ rad}^2 \end{aligned} \quad (3.9)$$

for $1/\tau_1 \gg |\omega| \gg 1/\tau_2$ and $K/\tau_2 \gg \omega^2$, and where δ represents the modulation index since

$$\int_{-\infty}^{\infty} S_m(\omega) \frac{d\omega}{2\pi} = 1 \quad (3.10)$$

This error will be minimized by selecting $a=1$ without altering the noise error as a consequence of our choice of design parameters. Thus, the first round of optimization results in K large and $a=1$ with a total mean-square error given by

$$\overline{\phi^2} = \frac{\delta^2 \alpha^4}{5B_n^4} + \frac{\Phi B_n}{2S_1} \text{ rad}^2 \quad (3.11)$$

The final step is to minimize with respect to B_n which yields

$$B_n = \left(\frac{8}{5} \delta^2 \alpha^4 \frac{S_1}{\Phi}\right)^{1/5} \text{ cps} \quad (3.12)$$

and a minimum total mean-square error equal to

$$\overline{\phi^2} = 2^{-12/5} \cdot 5^{4/5} \cdot \delta^{2/5} \left(\frac{\Phi \alpha}{S_1}\right)^{4/5} \text{ rad}^2 \quad (3.13)$$

Thus, if the threshold characterizing the breakdown of the linear model is chosen as σ^2 rad, then above-threshold operation implies

$$\frac{S_1}{\Phi} \geq \frac{5\delta^{1/2} \alpha}{8\sigma^{5/2}} \frac{\text{watts}}{\text{watts/cps}} \quad (3.14)$$

and the input SNR at threshold can be directly found by dividing the right-hand side of Eq. (3.14) by the i-f bandwidth in cps.

It is of interest to note that the optimization results in the noise error term of Eq. (3.11) being exactly 4 times larger than the distortion error term, and this statement has a one-to-one correspondence with Eq. (3.12), i. e., it completely defines the optimum design. This fact implies that the minimized total error can be conceived as either 5 times the distortion or as 5/4 times the noise error. The first characterization can be used to establish the minimum optimum noise bandwidth by demanding the total error to be lower than the critical σ^2 value, and results in

$$B_n \geq \left(\frac{\delta}{\sigma}\right)^{1/2} \alpha \text{ cps} \quad (3.15)$$

The second characterization can in turn be used to establish the SNR requirements and results in

$$(\text{SNR})_{B_n} = \frac{S_1}{\Phi B_n} \geq \frac{5}{8\sigma^2} \quad (3.16)$$

from which Eq. (3.14) can be derived with the aid of Eq. (3.15).

3.4 Case of a Sinusoidal Message

Consider now the case where the modulating signal has the form

$$\phi_s(t) = \delta \sin \omega_m t \quad (3.17)$$

In this case the distortion error will be sinusoidal with an amplitude given by

$$|\phi_{e,d}| = \delta |1 - H(j\omega_m)| \text{ rad}$$

$$\approx \delta \left[\frac{\left(\frac{\omega_m}{K}\right)^2 + \frac{1}{a^2(2-a)^2} \left(\frac{\omega_m}{B_n}\right)^4}{1 + \frac{2-3a}{a^2(2-a)} \left(\frac{\omega_m}{B_n}\right)^2 + \frac{1}{a^2(2-a)^2} \left(\frac{\omega_m}{B_n}\right)^4} \right]^{1/2} \text{ rad} \quad (3.18)$$

for $K/B_n \gg 2a$, and by noting that the maximum value of $a(2-a)$ is 1 and that K can be made very large, the first term in the root numerator may be neglected. Also, it seems reasonable to expect that either the first or the last term in the root denominator will predominate, except perhaps for a small range of ω_m/B_n values, which results in a compression factor of the form $(x/1+x)^{1/2}$ and $x \ll 1$ is required if any noticeable compression is to occur. Hence,

$$|\phi_{e,d}| \approx \frac{\delta \omega_m^2}{a(2-a)B_n^2} \text{ rad} \quad (3.19)$$

and optimum behavior again occurs for $a=1$ (and K large).

A more exact analysis can be done by plotting the compression factor with only the first term in the root numerator being neglected, i. e.,

$$|\phi_{e,d}| \approx \delta \left[1 + \frac{(2-a)(2-3a)}{\gamma^2} + \frac{a^2(2-a)^2}{\gamma^4} \right]^{-1/2} \text{ rad} \quad (3.20)$$

where $\gamma = \omega_m / B_n$ and the results are shown in Fig. 7 for a constant γ , i. e., a given noise error. The interest is to operate at the largest possible abscissa in order to compress the modulation index as much as possible. Operation at $\gamma > 1$ is inadmissible since no essential compression occurs, while operation at very small γ 's is excellent though the price will be an extremely large noise bandwidth and noise error. In general, it is noted that $a=1$ is optimum for $\gamma < 1$ where a compression occurs. Moreover, for $\gamma < 0.7$ the approximation of the abscissa by γ^{-2} (at $a=1$) can be found to yield a negligible error. On this basis, our choice of Eq. (3.19) with $a=1$ is truly an optimum behavior. Notice finally that the neglected factor containing the ω_m / K term will not alter the a choice but will only introduce a reduction in the assumed compression.

Thus, the total mean-square error to be minimized with respect to the noise bandwidth is given by

$$\overline{\phi^2} = \frac{\delta^2 \omega_m^4}{2B_n^4} + \frac{\Phi B_n}{2S_1} \quad \text{rad}^2 \quad (3.21)$$

and a comparison with Eq. (3.11) shows that we can use the previous results replacing α by ω_m and δ by $\sqrt{2.5} \delta$. Hence, the optimum noise bandwidth is then

$$B_n = (4 \delta^2 \omega_m^4 \frac{S_1}{\Phi})^{1/5} \quad \text{cps} \quad (3.22)$$

while the minimum total mean-square error is given by

$$\overline{\phi^2} = 2^{-13/5} \cdot 5 \cdot \delta^{2/5} \left(\frac{\Phi \omega_m}{S_1} \right)^{4/5} \quad \text{rad}^2 \quad (3.23)$$

and the above-threshold SNR requirement reads

$$\frac{S_1}{\Phi} \geq \frac{5^{5/4} \delta^{1/2} \omega_m}{2^{13/4} \sigma^{5/2}} \quad \frac{\text{watts}}{\text{watts/cps}} \quad (3.24)$$

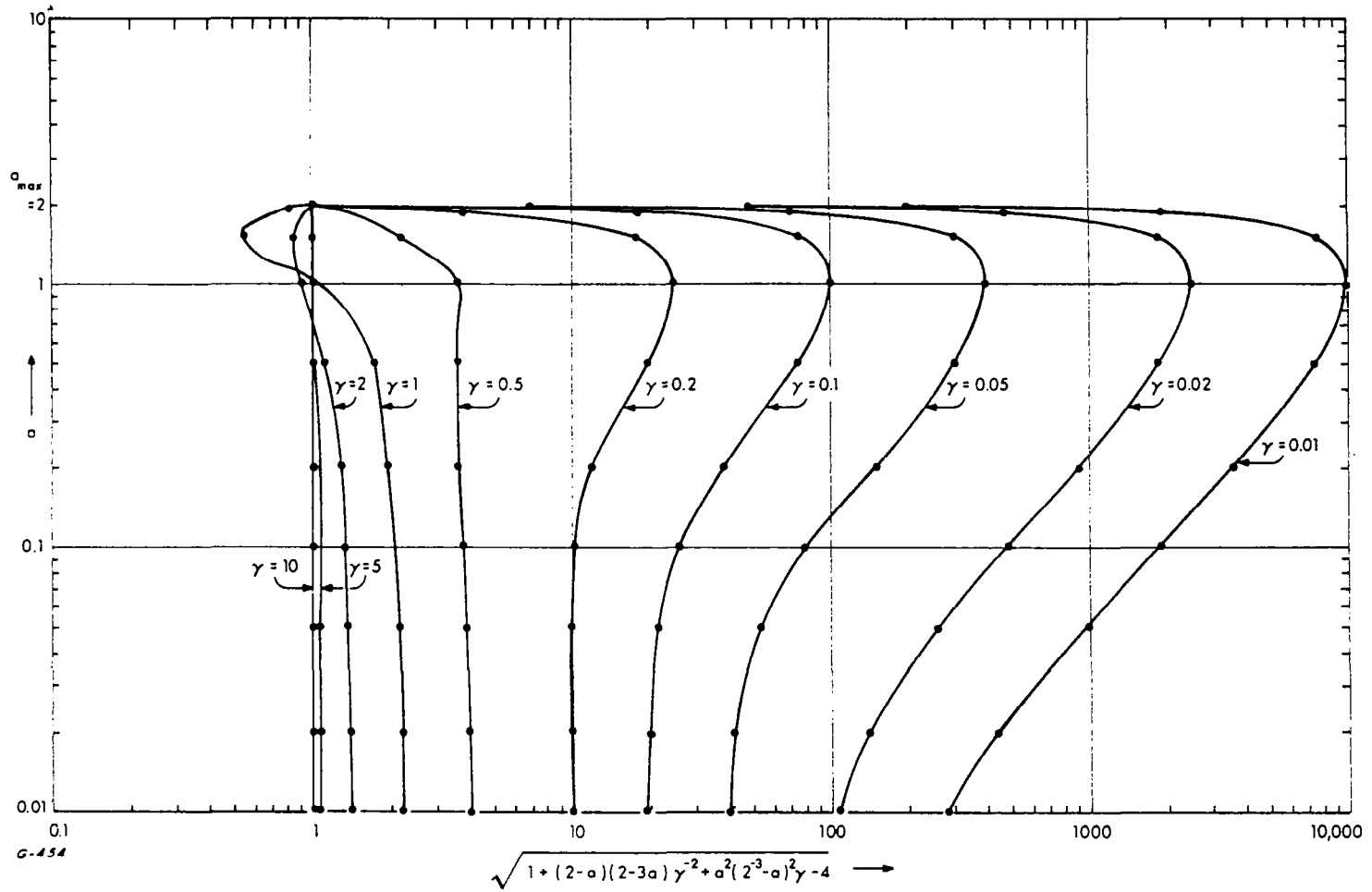


Fig. 7. Compression vs (a, γ) .

It can be checked that the optimization again results in the noise error being 4 times larger than the distortion error. The first characterization of the total error as 5 times the distortion error yields the minimum optimum noise bandwidth as

$$B_n \geq \left(\frac{5}{2}\right)^{1/4} \left(\frac{\delta}{\sigma}\right)^{1/2} \omega_m \quad \text{cps} \quad (3.25)$$

while the second characterization of the total error yields the SNR requirement as

$$(\text{SNR})_{B_n} = \frac{S_1}{\Phi B_n} \geq \frac{5}{8\sigma^2} \quad (3.26)$$

Thus a comparison of the optimum second order PLL demodulators for the two modulating signals under consideration shows that:

- a) An optimum design occurs when K is large and $a=1$ in both cases.
- b) The minimum noise bandwidth is larger in the sinusoidal modulation case by a factor of $(5/2)^{1/4} = 1.26$ when the two baseband message bandwidths are set similar ($\alpha = \omega_m$), as indicated by Eqs. (3.15) and (3.25).
- c) The input SNR at threshold is larger in the sinusoidal modulation case by the same $(5/2)^{1/4} = 1.26 = 1$ db factor, as indicated by Eqs. (3.14) and (3.24).
- d) The loop SNR (i. e., the SNR in the loop phase noise bandwidth) at threshold is the same in both cases, as indicated by Eqs. (3.16) and (3.26). This is a direct consequence of the fact that the same optimum relative weighting (1:4) exists between the distortion and noise error terms in both cases.

3.5 Second Order Loop Design

The experimental tests were conducted with a 1 kc tone as the modulating signal and on this basis the results of Sec. 3.4 were used to design the loop. The modulation indices were chosen as $\delta=10, 20$ and 30 and a mean-square phase error $\sigma^2=1/4 \text{ rad}^2$ was assumed to characterize the threshold

a priori. The reliability of this last choice was checked experimentally and will be discussed later.

The design logic is as follows: the value $a = 1$ has already been motivated as an optimum choice. The optimum noise bandwidth is then found by using Eq. (3.25) for the three indices under consideration. Finally, the loop gain K is chosen large enough to satisfy the approximations involved in Secs. 3.2 and 3.4, and this completes the loop design in terms of the set of parameters (K, B_n, a) . The original set (K, τ_1, τ_2) was then computed and the loop filter synthesized using the simple lag realization shown in Fig. 8. The design is now summarized with the following table:

δ	B_n (kc)	B_n / ω_m (cycles/rad)	$K(\text{sec}^{-1})$	τ_1 (μsec)	τ_2
10	35.4	5.63	$> 7 \times 10^5$	28	$K\tau_1^2$
20	50.1	7.97	$> 12 \times 10^5$	20	$K\tau_1^2$
30	61.3	9.76	$> 18 \times 10^5$	16.3	$K\tau_1^2$

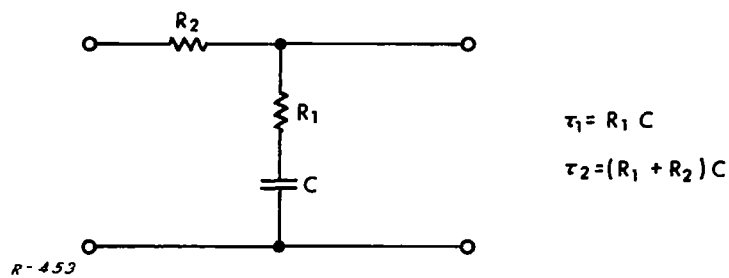


Fig. 8 Second order loop filter.

3.6 Conclusion

This chapter has presented the analytical optimization of second order PLL demodulators for the cases where the PM message has a rectangular spectrum and for sinusoidal modulation. A novel advantageous technique consists of first replacing the conventional set of design parameters (K, τ_1, τ_2) by another one that includes the equivalent noise bandwidth as an independent parameter. The design logic is then to minimize the distortion error with respect to all parameters but the noise bandwidth, and finally to minimize the total error (distortion plus noise) with respect to this bandwidth.*

The second order loops have been optimized (at threshold) for the two modulation messages in question under a mean-square error criterion and the two cases result in almost identical designs. In both cases we have the optimum $a = 1$ (same damping - see footnote). In both cases the minimum total error shows an optimum relative weighting equal to 4:1 between noise and distortion errors, which results in the same threshold SNR in the loop bandwidth. Also, the optimum noise bandwidths and input SNR's at threshold differ by only 1 db when the two baseband message bandwidths are set equal. The experimental design specifications are finally determined by assuming a mean-square error of $1/4 \text{ rad}^2$ at threshold.

* Another popular set of design parameters is (K, ω_n, ζ) where ω_n is the resonance frequency and ζ the damping factor of the linear transfer function $H(s)$. While this last set can be closely identified in servo notation and linear transient analysis, it does not characterize the loop phase noise directly nor independently of the distortion. Moreover, the applicability of such set is questionable for higher-order loops while the noise bandwidth can always be taken as an independent and extremely informative parameter in any loop (e. g., different order loops having the same noise bandwidth need only be compared in their distortion behavior to establish the best one under a mean-square error criterion). In any case, the relation between the two sets is given by

$$\left. \begin{aligned} \omega_n &\approx \sqrt{a(2-a)} B_n \\ \zeta &\approx 1/2 \sqrt{\frac{2-a}{a}} \end{aligned} \right\} \text{ for } \frac{K}{B_n} \gg a$$



Chapter IV

HIGHER ORDER PLL DEMODULATORS

4.1 Introduction

In this chapter we will derive the optimum realizable (linear) PLL demodulator for the case where the PM message has a rectangular spectrum using a mean-square error criterion. The resultant optimum loop filter is then approximated by a set of functions that essentially characterize the loop order, and the performance of these higher-order loops is evaluated. A similar analysis for sinusoidal modulation becomes too complex because of the impulsive nature of the message spectrum and a study of the loop optimality when this input is used is limited to experimental results.

On this basis, the baseband PM message is assumed to have a spectrum $S_m(\omega)$ satisfying Eq. (3.10) and the problem is to find the optimum realizable transfer function $H(s)$ that minimizes the total mean-square phase error. There exists a one-to-one correspondence between the transfer function $H(s)$ and the loop filter $F(s)$, as given by Eq. (3.1). It is of interest to note that in an ordinary open-loop filtering case we could omit the realizability condition and approximate any unrealizable filter that results by introducing a delay between its input and output. However, in the PLL the loop filter output is used to form the instantaneous message reproduction so that we cannot use these techniques.

4.2 The Optimum Realizable PLL Demodulator

The total mean-square error to be minimized is formed by two additive terms: a distortion error given by

$$\overline{\phi_{e,d}^2} = \int_{-\infty}^{\infty} \delta^2 S_m(\omega) |1 - H(j\omega)|^2 \frac{d\omega}{2\pi} \text{ rad}^2 \quad (4.1)$$

where δ is the modulation index, and a noise error given by Eq. (3.3). The optimum realizable linear filter has been derived by Yovits and Jackson⁸ and results in

$$|1 - H(j\omega)|^2 = \frac{1}{1 + \frac{2S_1}{\Phi} \delta^2 S_m(\omega)} \quad (4.2)$$

while the minimum total error is given by

$$\overline{\phi^2} = \frac{\Phi}{2S_1} \int_{-\infty}^{\infty} \log \left[1 + \frac{2S_1}{\Phi} \delta^2 S_m(\omega) \right] \frac{d\omega}{2\pi} \text{ rad}^2 \quad (4.3)$$

Consider now the case where $S_m(\omega)$ is given by Eq. (3.8) and for convenience define the SNR in the message bandwidth α as

$$\Lambda = \frac{2\pi S_1}{\Phi \alpha} \quad (4.4)$$

so that Eqs. (4.2) and (4.3) respectively become

$$|1 - H(j\omega)|^2 = \begin{cases} \frac{1}{1 + \Lambda \delta^2} & \text{for } |\omega| < \alpha \\ 1 & \text{otherwise} \end{cases} \quad (4.5)$$

and

$$\overline{\phi^2} = \frac{1}{\Lambda} \log(1 + \Lambda \delta^2) \quad (4.6)$$

A plot of Eq. (4.6) as a function of the total error $\overline{\phi^2}$ is shown in Fig. 9. The choice of a threshold value will select one of the many possible curves which will then play the same role as Eqs. (3.14) and (3.24); it describes the threshold SNR as a function of the message parameters.

4.3 Approximation to the Optimum PLL Demodulator

It is clear that to build a system having exactly the transfer function of Eq. (4.5) would require an infinite number of poles and zeros. However, we can certainly approximate it as closely as desired by using a system of high enough order. Because of the increase of complexity associated with higher

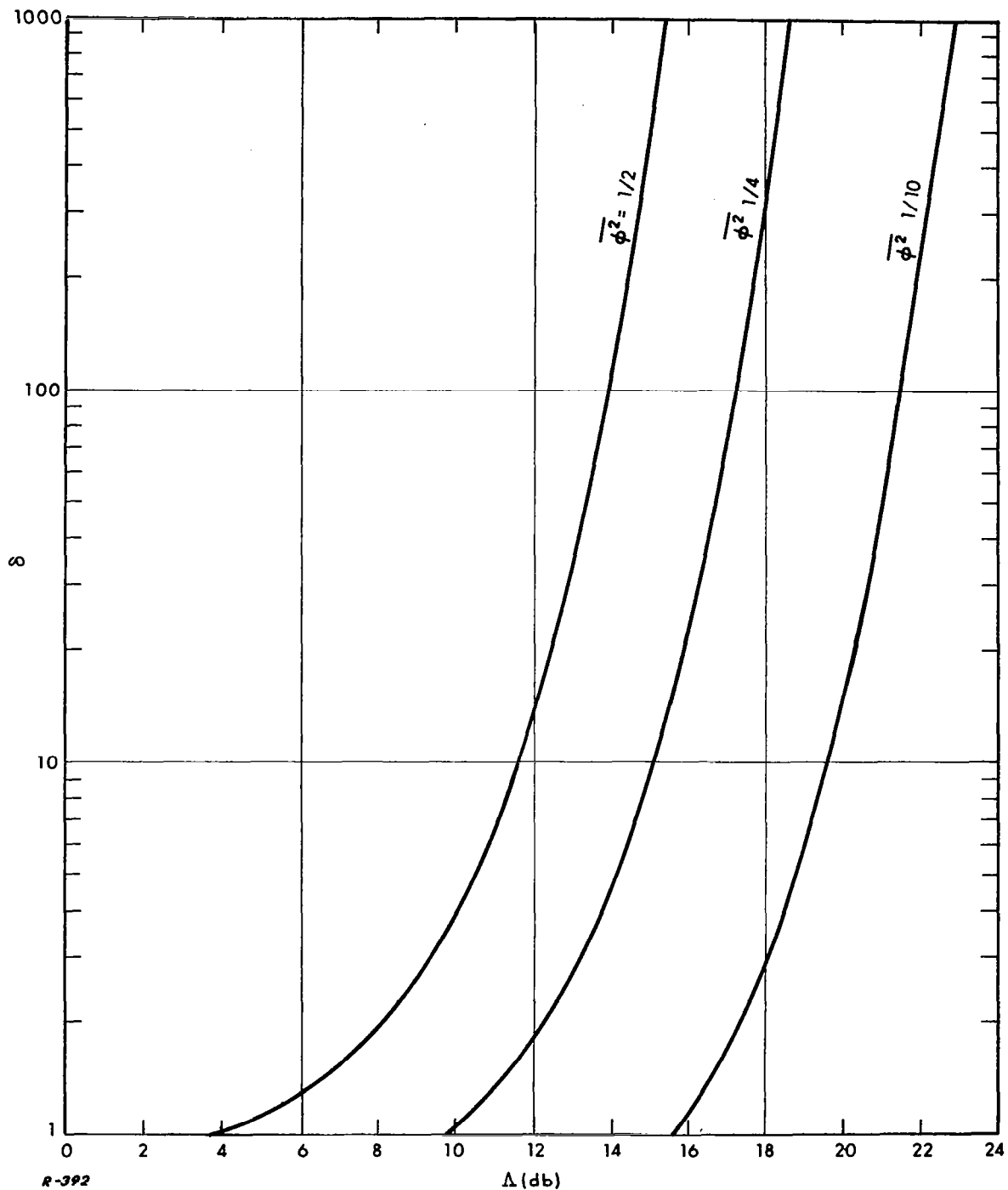


Fig. 9. Characteristics of optimum demodulator.

order filters we would like to keep the order, n , as small as possible while meeting a desired performance criterion. One way of approaching this problem might be to find the "best" n^{th} order filter by the following procedure. First, we would write the unknown filter $F(s)$ in terms of $2(n-1)$ parameters as

$$F(s) = \prod_{i=1}^{n-1} \frac{s + \alpha_i}{s + \beta_i} \quad (4.7)$$

It is easily verified by Eq. (3.1) that this gives an $H(s)$ with a denominator of order n . Next we could vary the $2(n-1)$ parameters $\alpha_1, \alpha_2, \dots, \alpha_{n-1}, \beta_1, \beta_2, \dots, \beta_{n-1}$ so as to minimize the resulting mean-square loop error. This procedure would be obviously impractical for n reasonably large. An alternative approach is to assume more of the structure initially, leaving less parameters to be varied.

Adopting the latter procedure, we notice that $|1 - H(j\omega)|^2$ is equal to the difference between a constant and a rectangle:

$$|1 - H(j\omega)|^2 = 1 - \left(1 - \frac{1}{1 + \Lambda \delta^2}\right) f(\omega) \quad (4.8)$$

where

$$f(\omega) = \begin{cases} 1 & |\omega| < \alpha \\ 0 & \text{otherwise} \end{cases} \quad (4.9)$$

A familiar way of approximating the rectangular function is with Butterworth functions of the form

$$f(\omega) = \frac{1}{1 + \left(\frac{\omega}{\alpha}\right)^{2n}} \quad n = 1, 2, \dots \quad (4.10)$$

It turns out that we will achieve much better performance by replacing α by a parameter β which we will vary to optimize the loop. With this change,

the substitution of Eq. (4.10) into Eq. (4.8) yields

$$|1 - H(j\omega)|^2 = \frac{\left(\frac{\omega}{\beta}\right)^{2n} + \frac{1}{1 + \Lambda \delta^2}}{1 + \left(\frac{\omega}{\beta}\right)^{2n}} \quad (4.11)$$

which approaches the optimum system of Eq. (4.5) both for very large and very small ω , and where the speed of convergence is controlled by n . A particular advantage of our choice of Butterworth functions is that in calculating $H(j\omega)$ (and $F(j\omega)$) we can take advantage of the fact that the factors of terms of the form $1 + x^{2n}$ have been previously studied. (See Ref. 9, p. 252).

We will now discuss the efficiency of these higher order loops in approximating the optimum demodulator performance by assuming a threshold phase error $\sigma^2 = 1/4 \text{ rad}^2$. The results are summarized in Fig. 10 where the performance of the $n = 2-4$ approximations is indicated. The dotted line corresponds to the optimum second order loop derived in Chapter III. Each curve characterizes the system threshold performance after being properly optimized to have a minimum total mean-square error, i. e., they represent plots analogous to Eqs. (3.14) and (3.24) for $\sigma^2 = 1/4 \text{ rad}^2$. The actual derivation of each curve is illustrated in the next section for the $n = 3$ case, the others following a similar analysis. It is noted that the $n = 2$ system is very close to the conventional second order loop of Chapter III. The threshold improvements of each system over this last loop are directly read as their horizontal separation for a given index. The fourth order loop can give a 3 db improvement for indices larger than 30 while a third order loop would require an index of the order of 200. As the order is increased, the loops yield diminishing returns in their improvement capability so that an extremely high order loop would be actually needed to approach the optimum performance.

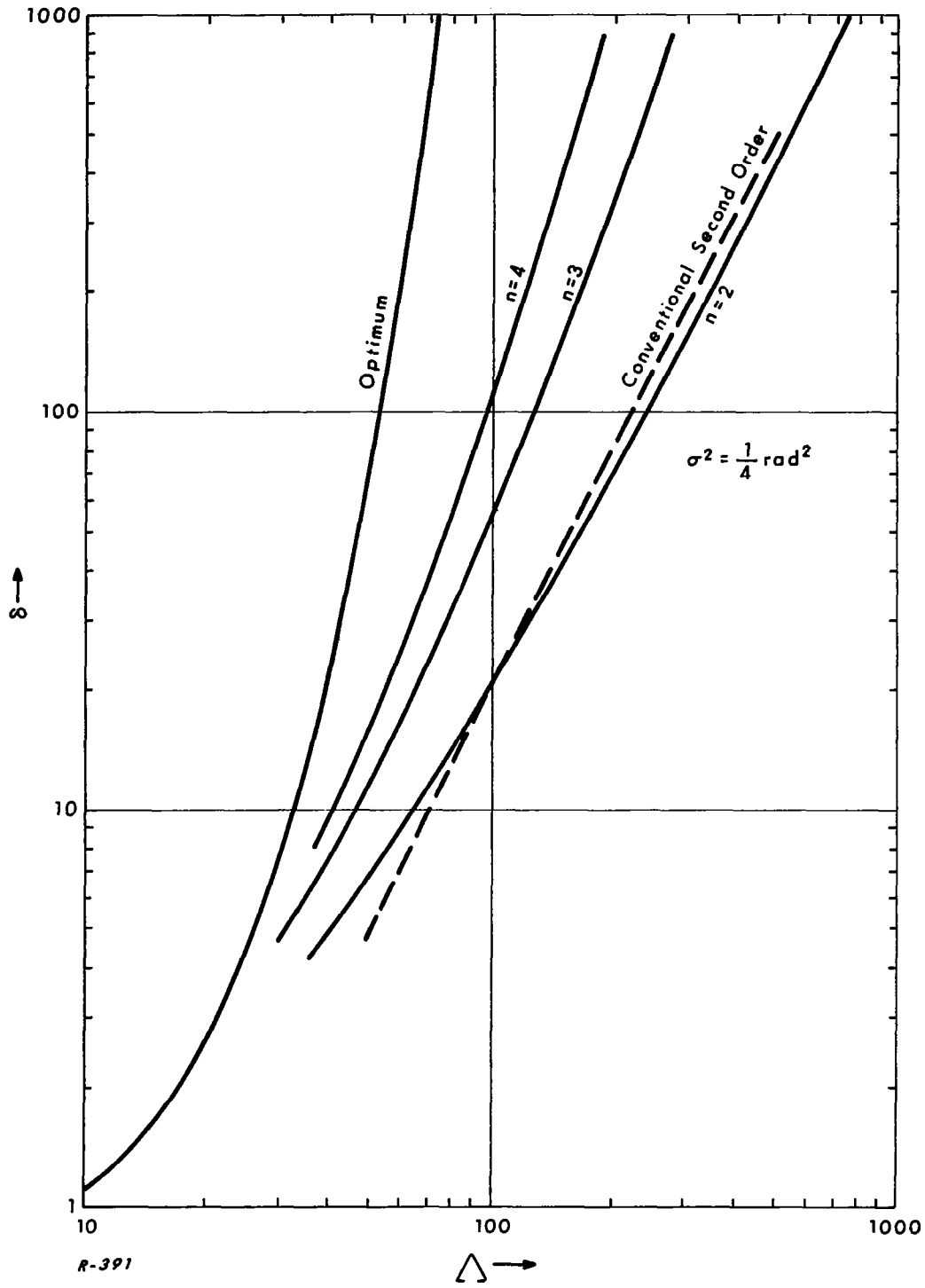


Fig. 10 Threshold curves for the approximations to the optimum demodulator.

4.4 The Third Order Loop

In order to check the experimental feasibility of these results a third order PLL demodulator corresponding to the use of $n = 3$ in Eq. (4.11) was synthesized. We will now derive the resultant loop filter and discuss its synthesis. To do this, we first rewrite Eq. (4.11) in the form

$$|1 - H(j\omega)|^2 = [1 - H(j\omega)][1 - H(-j\omega)] = \frac{1 + (\frac{\omega}{\beta'})^6}{(1 + \Lambda\delta^2)[1 + (\frac{\omega}{\beta})^6]} \quad (4.12)$$

where

$$\beta' = \frac{\beta}{(1 + \Lambda\delta^2)^{1/6}} \quad (4.13)$$

Next we note that the factors of $1 + x^6$ are given by

$$(1 + jx)(1 - jx)(1 + jx - x^2)(1 - jx - x^2) \quad (4.14)$$

so that assigning all singularities in the RHS of the complex plane to $1 - H(-j\omega)$ and the remaining ones to $1 - H(j\omega)$, we get

$$1 - H(j\omega) = \frac{1}{\sqrt{1 + \Lambda\delta^2}} \cdot \frac{[1 + j(\frac{\omega}{\beta'})][1 + j(\frac{\omega}{\beta'}) - (\frac{\omega}{\beta'})^2]}{[1 + j(\frac{\omega}{\beta})][1 + j(\frac{\omega}{\beta}) - (\frac{\omega}{\beta})^2]} \quad (4.15)$$

and transposing with the aid of Eq. (4.13) we finally have

$$H(j\omega) = \frac{\sqrt{1 + \Lambda\delta^2} - 1 + 2j[\sqrt{1 + \Lambda\delta^2} - 6\sqrt{1 + \Lambda\delta^2}](\frac{\omega}{\beta}) - 2[\sqrt{1 + \Lambda\delta^2} - 3\sqrt{1 + \Lambda\delta^2}](\frac{\omega}{\beta})^2}{\sqrt{1 + \Lambda\delta^2} [1 + j(\frac{\omega}{\beta})][1 + j(\frac{\omega}{\beta}) - (\frac{\omega}{\beta})^2]} \quad (4.16)$$

We have thus determined the closed loop transfer function except for the parameter β which should be chosen so as to minimize the total mean-square error. The distortion error of Eq. (4.1) can be evaluated with the aid of Eqs. (3.8)

and (4.11) to yield

$$\overline{\phi_{e, d}^2} = \frac{\delta^2 \alpha^6}{7\beta^6} \text{ rad}^2 \quad (4.17)$$

while the noise error of Eq. (3.3) yields

$$\overline{\phi_{e, n}^2} = \frac{\pi\beta}{3\alpha\Lambda(1+\Lambda\delta^2)} \left[5(1+\Lambda\delta^2) - 6(1+\Lambda\delta^2)^{5/6} + 1 \right] = \frac{\lambda\beta}{\Lambda\alpha} \text{ rad}^2 \quad (4.18)$$

where λ is introduced to simplify the notation.

The total mean-square error then becomes

$$\overline{\phi^2} = \frac{\delta^2 \alpha^6}{7\beta^6} + \frac{\lambda\beta}{\Lambda\alpha} \text{ rad}^2 \quad (4.19)$$

and minimizing with respect to β we obtain

$$\beta = \left(\frac{6\Lambda\delta^2}{7\lambda} \right)^{1/7} \alpha \quad (4.20)$$

while the corresponding minimum total error results in

$$\overline{\phi^2} = \left(\frac{7\lambda}{6\Lambda} \right)^{6/7} \delta^{2/7} \text{ rad}^2 \quad (4.21)$$

It is of interest to note that the optimum relative weighting between the noise and distortion error terms of Eq. (4.19) is 6:1, this statement having a one-to-one correspondence with Eq. (4.20).

The $n = 3$ curve of Fig. 10 is then plotted by assuming a threshold error $\sigma^2 = 1/4 \text{ rad}^2$ so that Eq. (4.21) becomes

$$\delta = 0.00495 \left(\frac{\Lambda}{\lambda} \right)^3 \quad (4.22)$$

which is an implicit equation since λ is given by

$$\lambda = \frac{\pi}{3(1+\Lambda\delta^2)} \left[5(1+\Lambda\delta^2) - 6(1+\Lambda\delta^2)^{5/6} + 1 \right] \quad (4.23)$$

The following procedure was used to plot the curve

- a) Assume a value of $\Lambda\lambda^2$, say ν , and draw the straight line (plotting δ vs Λ in logarithmic coordinates) corresponding to $\Lambda\delta^2 = \nu$.
- b) Use Eq. (4.23) to calculate λ for $\Lambda\delta^2 = \nu$ and draw the straight line corresponding to Eq. (4.22) with the resultant value of λ .
- c) The intersection of the two straight lines drawn in a) and b) is a point on the desired threshold curve. The procedure is then repeated with different values of ν to give the complete curve.

4.5 Third Order Loop Design

In the last section we have determined the transfer function $H(j\omega)$ in terms of Λ , δ and β/α for the third order loop. We now wish to synthesize the corresponding loop filter given by

$$F(s) = \frac{sH(s)}{K[1 - H(s)]} \quad (4.24)$$

so that for the $H(s)$ given by Eq. (4.16)

$$F(s) = \frac{s}{K} \frac{\sqrt{1+\Lambda\delta^2} - 1 + 2(\sqrt{1+\Lambda\delta^2} - \sqrt[6]{1+\Lambda\delta^2}) \left(\frac{s}{\beta}\right) + 2(\sqrt{1+\Lambda\delta^2} - \sqrt[3]{1+\Lambda\delta^2}) \left(\frac{s}{\beta}\right)^2}{\left[1 + \sqrt[6]{1+\Lambda\delta^2} \left(\frac{s}{\beta}\right)\right] \left[1 + \sqrt[6]{1+\Lambda\delta^2} \left(\frac{s}{\beta}\right) + \sqrt[3]{1+\Lambda\delta^2} \left(\frac{s}{\beta}\right)^2\right]} \quad (4.25)$$

We can write this in a more compact form by defining some new time constants as

$$F(s) = A_o \frac{A'}{A} \frac{sT_3'}{1 + sT_1' + s^2T_1'T_2'} \cdot \frac{1 + sT_1 + s^2T_1T_2}{1 + sT_3} \quad (4.26)$$

As it stands, $F(s)$ does not provide a d-c path. This will be necessary to keep the center frequency of the VCO from drifting away from the carrier frequency. For this reason we modify $F(s)$ to include a d-c path: as

$$F(s) = A_o \frac{A'}{A} \frac{1 + sT'_3}{1 + sT'_1 + s^2 T'_1 T'_2} \cdot \frac{1 + sT_1 + s^2 T_1 T_2}{1 + sT_3} \quad (4.27)$$

With this definition of the time constants we have the following equivalences:

$$KA_o T'_3 \frac{A'}{A} = \sqrt{1 + \Lambda \delta^2} - 1 = \text{Total loop gain} \quad (4.28)$$

$$T_1 = \frac{2}{\beta} \frac{\sqrt{1 + \Lambda \delta^2} - \sqrt[6]{1 + \Lambda \delta^2}}{\sqrt{1 + \Lambda \delta^2} - 1} \quad (4.29)$$

$$T_2 = \frac{1}{\beta} \frac{\sqrt{1 + \Lambda \delta^2} - \sqrt[3]{1 + \Lambda \delta^2}}{\sqrt{1 + \Lambda \delta^2} - \sqrt[6]{1 + \Lambda \delta^2}} \quad (4.30)$$

$$T_3 = T'_1 = T'_2 = \frac{1}{\beta} \sqrt[6]{1 + \Lambda \delta^2} \quad (4.31)$$

$T'_3 =$ (we must choose this so that $sT'_3 \gg 1$ except at frequencies near d-c)

From Fig. 10 we can read the values of the SNR Λ at threshold for loops operating at any given modulation index. The values of Λ and δ are then inserted into Eq. (4.23) to find the corresponding λ which in turn yields β/α from Eq. (4.20). The final use of $\alpha = 2\pi \times 10^3$ rad/sec permits us to evaluate the filter parameters. The design for $\delta = 10, 20$ and 30 is summarized in the following table:

δ	10	20	30
Δ (db)	16.7	18.1	18.9
β/α	2.71	3.42	3.91
$KA_o T_3' \frac{A'}{A} (\text{sec}^{-1})$	67.5	160.1	262
T_1 (sec)	1.121×10^{-4}	9.03×10^{-5}	8.00×10^{-5}
T_2 (sec)	4.72×10^{-5}	3.93×10^{-5}	3.53×10^{-5}
T_3 (sec)	2.41×10^{-4}	2.53×10^{-4}	2.62×10^{-4}
T_1' (sec)	2.41×10^{-4}	2.53×10^{-4}	2.62×10^{-4}
T_2' (sec)	2.41×10^{-4}	2.53×10^{-4}	2.62×10^{-4}
T_3' (sec)	1.00×10^{-2}	1.00×10^{-2}	1.00×10^{-2}

It is noted that the only loop parameter which is strongly dependent on the modulation index δ is the loop gain. This suggests the possibility of designing a loop which will operate well at different modulation indices with adjustment of the loop gain only.

The problem of synthesizing a network having the transfer function of Eq. (4.27) may be approached in many different ways. The approach chosen here is to design an active network including an operational amplifier. With the arrangement shown in Fig. 11, the overall transfer function is given by

$$F(s) = A_o \frac{Z'(s)}{Z(s)} \quad A_o \gg 1 \quad (4.32)$$

This will be identical to Eq. (4.27) if we let

$$Z'(s) = A' \frac{1 + sT'_3}{1 + sT'_1 + s^2 T'_1 T'_2} \quad (4.33)$$

and

$$Z(s) = A \frac{1 + sT_3}{1 + sT_1 + s^2 T_1 T_2} \quad (4.34)$$

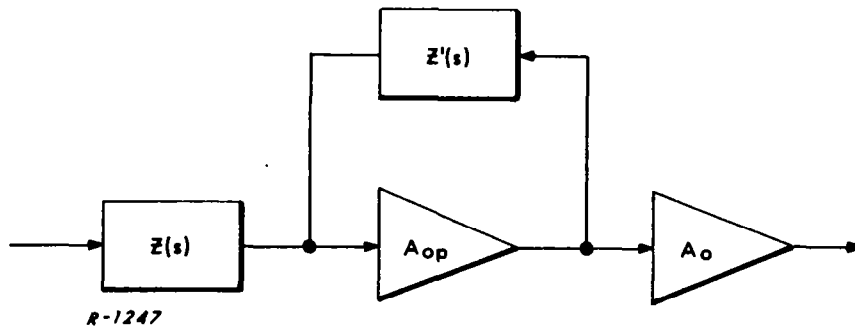
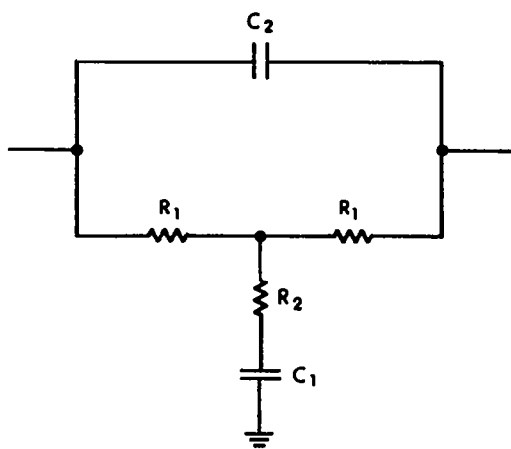
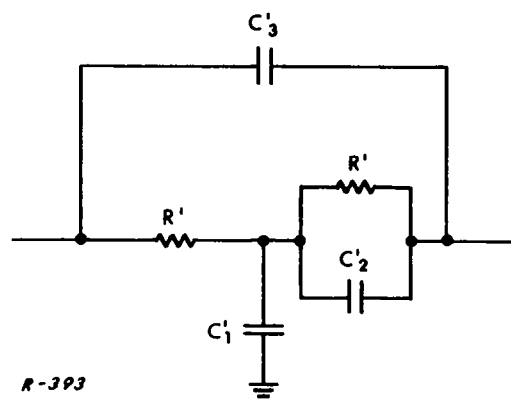


Fig. 11 Configuration of active filter network.

To synthesize the impedance $Z(s)$ we choose the circuit configuration shown in Fig. 12a. (We must use a configuration which allows complex poles.) For this circuit, the following relations hold (see Ref.10)



a Circuit of $Z(s)$



R-393

b Circuit of $Z'(s)$

Fig. 12 Circuits for synthesizing impedances.

$$R_1 = \frac{A}{2} \quad (4.35)$$

$$R_2 = \frac{AT_1(T_3 - T_2)}{4[T_3^2 - T_1(T_3 - T_2)]} \quad (4.36)$$

$$C_1 = \frac{4[T_3^2 - T_1(T_3 - T_2)]}{AT_3} \quad (4.37)$$

$$C_2 = \frac{T_1 T_2}{AT_3} \quad (4.38)$$

Using these equations and the tabulated values of the time constants T_1 , T_2 , T_3 given we can calculate the component values as follows

COMPONENT VALUES FOR Z(s)

	$\delta = 10$	$\delta = 20$	$\delta = 30$
R_1	0.5A	0.5A	0.5A
R_2	0.150A	0.1081A	.0899A
C_1	$\frac{6.03 \times 10^{-4}}{A}$	$\frac{7.07 \times 10^{-4}}{A}$	$\frac{7.71 \times 10^{-4}}{A}$
C_2	$\frac{2.20 \times 10^{-5}}{A}$	$\frac{1.402 \times 10^{-5}}{A}$	$\frac{1.078 \times 10^{-5}}{A}$

The parameter A may be chosen so that these component values are of reasonable magnitude.

To synthesize the impedance $Z'(s)$ we choose a different circuit configuration (Fig. 12b) which gives more reasonable component values in spite of the large value of T_3' . With this circuit we have the following relationships:

$$R' = \frac{A'}{2} \quad (4.39)$$

$$C'_1 = \frac{2[2T'_3{}^2 - T'_1(T'_3 - T'_2)]}{A'T'_3} \quad (4.40)$$

$$C'_2 = \frac{2T'_1(T'_3 - T'_2)}{A'T'_3} \quad (4.41)$$

$$C'_3 = \frac{T'_1 T'_2}{A'T'_3} \quad (4.42)$$

Substitution of the tabulated time constants into these equations gives the component values listed below. The parameter A' may be chosen so that these component values are of reasonable magnitude.

COMPONENT VALUES FOR $Z'(s)$

	$\delta = 10$	$\delta = 20$	$\delta = 30$
R'_1	$0.5A'$	$0.5A'$	$0.5A'$
C'_1	$\frac{3.95 \times 10^{-2}}{A'}$	$\frac{3.95 \times 10^{-2}}{A'}$	$\frac{3.95 \times 10^{-2}}{A'}$
C'_2	$\frac{4.70 \times 10^{-4}}{A'}$	$\frac{4.94 \times 10^{-4}}{A'}$	$\frac{5.10 \times 10^{-4}}{A'}$
C'_3	$\frac{5.81 \times 10^{-6}}{A'}$	$\frac{6.40 \times 10^{-6}}{A'}$	$\frac{6.86 \times 10^{-6}}{A'}$

4.6 Applicability of Bode Filters to PLL Demodulators

During the development of the program, the possibility of using Bode filters as the loop filters was mentioned. The term "Bode filter" describes an idealized, realizable lowpass transfer function with certain specific properties. These specific properties are:

- a) constant gain (A_0) in the passband;
- b) constant phase margin ($y\pi$ radians) outside the passband, i. e., the phase shift outside the passband is equal to $(1-y)\pi$ radians.

If we let $H(j\omega)$ be the Bode transfer function and define

$$\log H(j\omega) = A(j\omega) + jB(j\omega) \quad (4.43)$$

then we may restate the design properties of the Bode filter as

$$(a) \quad A(j\omega) = A_0 \quad \omega < \omega_0 \quad (4.44)$$

$$(b) \quad B(j\omega) = -(1-y)\pi \quad \omega > \omega_0 \quad (4.45)$$

where ω_0 defines the passband. The realizable transfer function corresponding to these two conditions is given by

$$\log H(j\omega) = A_0 - 2(1-y) \log \left[\sqrt{1 - \frac{\omega^2}{\omega_0^2}} + j \frac{\omega}{\omega_0} \right] \quad (4.46)$$

It is easy to show that in the separate regions of interest this expression can be reduced to

$$\left. \begin{aligned} A(j\omega) &= A_0 \\ B(j\omega) &= -2(1-y) \tan^{-1} \frac{\omega}{\sqrt{\omega_0^2 - \omega^2}} \end{aligned} \right\} \text{for } \omega < \omega_0 \quad (4.47)$$

$$\left. \begin{aligned} A(j\omega) &= A_0 - 2(1-y) \log \left[\sqrt{\frac{\omega^2}{\omega_0^2} - 1} + \frac{\omega}{\omega_0} \right] \\ B(j\omega) &= -(1-y)\pi \end{aligned} \right\} \text{for } \omega > \omega_0 \quad (4.48)$$

The amplitude and phase functions for a typical Bode filter (phase margin = 30°) are plotted in Fig. 13.

It should be emphasized that the Bode filter is "ideal" only in the sense that it satisfies the two "ideal" properties (a) and (b) above. If different requirements were specified, a different transfer function would be ideal. Furthermore, there is no direct way to synthesize a network having the transfer function given by Eq. (4.46)

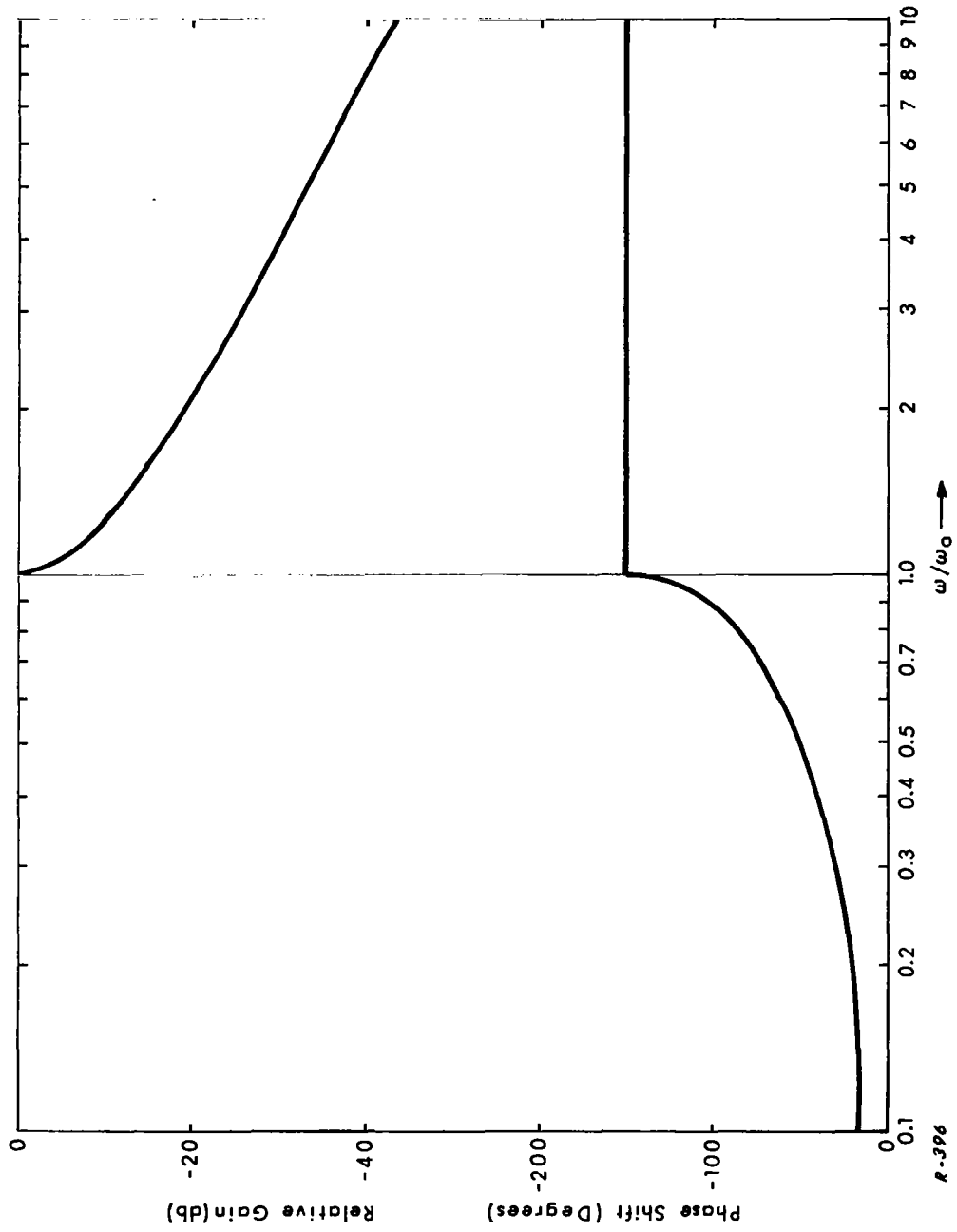


Fig. 13 Gain and phase shift of a Bode filter with 30° phase margin.

In considering the applicability of a Bode filter to the PLL demodulator, we must recognize that the properties which are desirable in the PLL filter do not necessarily coincide with those used to design the Bode filter. Specifically, the main objective in the PLL is to prevent the loss of lock which causes the threshold phenomenon and for this purpose it is necessary to minimize the total loop error. The design of PLL must thus be directed towards the minimization of the loop error and conditions (a) and (b) above do not have this property. Suppose we are given the task to select a loop filter for PLL demodulation. There are two possible approaches that we can take: we either select some truly optimum filter form derived from some adequate criterion and which may be difficult to synthesize (like the filters of the previous section of this Chapter IV) or we use a simpler form from synthesis considerations which will not be optimum (like the filter in Chapter III). Unfortunately, the Bode filter does not meet any of these properties, i. e., it is not optimum and it is not easy to synthesize. On this basis, we must reject it as an a priori logical choice in the design of PLL demodulators.

4.7 Conclusion

In this chapter we have derived the optimum (linear) PLL demodulator for a rectangular PM spectrum and established its threshold performance. The resultant closed-loop transfer function can be approximated with the aid of modified Butterworth functions whose order characterizes the loop order. The threshold performance of these higher order loops is presented with particular emphasis on:

- a) their threshold improvement capabilities relative to the second order loop discussed in Chapter II;
- b) their ability to approach the optimum demodulator performance.

A third order PLL demodulator is optimized resulting in a 6:1 relative weighting between noise and distortion error terms. The corresponding loop

filter is synthesized with RC active networks due to the presence of complex poles. The filter in question does not include a necessary d-c path and this is provided in such a way that the system performance is not essentially altered.



Chapter V

EXPERIMENTAL RESULTS

5.1 Introduction

In this chapter we will present and discuss the experimental results obtained. The threshold curves are plotted in the usual way: the x-axis represents the input SNR defined in the i-f bandwidth and the y-axis represents the output SNR defined in the output lowpass filter bandwidth. The i-f bandwidths used are shown in Figs. 14 a-c and their corresponding noise bandwidths are approximately 23, 54 and 110 kc (as obtained by graphical interpolation and integration).

The output lowpass filter bandwidth was selected to maintain the output SNR within the dynamic range of the system. A slight variation of the threshold took place as this bandwidth varied and it can be attributed to our threshold criterion of so many db below linearity together with the nonuniform variation of the output noise spectrum for different SNR's (for example, see Ref. 11). However, this variation was very small as long as reasonable i-f and output bandwidths were used according to the modulation index (10, 20, 30) and frequency (1 kc) of the sinusoidal FM signal being used in the experiments. The direction and magnitude of these variations were such that the comparison of different demodulators at different indices but with the same output lowpass filter was adequate for establishing relative threshold improvements. We have chosen a 6 kc (3 db point) lowpass filter for the presentation of the results because we can accommodate all demodulators and indices within the system dynamic range while still maintaining an accurate absolute threshold reading. Also, in order to evaluate the SNR, some notches were used to remove the signal and its harmonics and their effect should be accounted for if a rigorous evaluation of the output noise is desired. These effects are illustrated in Figs. 15 and 16; the first one shows the original 6 kc filter and the second one shows its modification introduced by the notches.

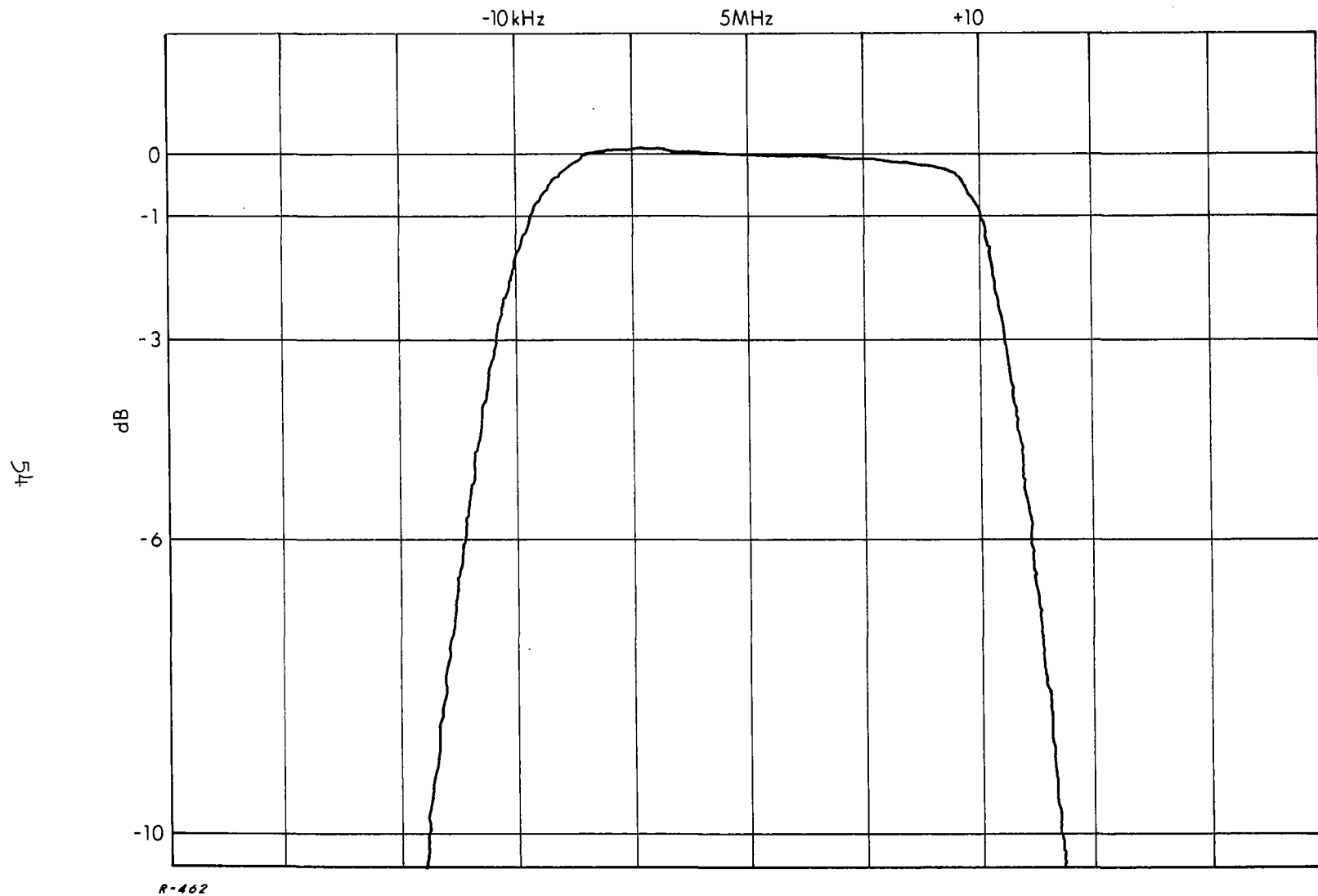
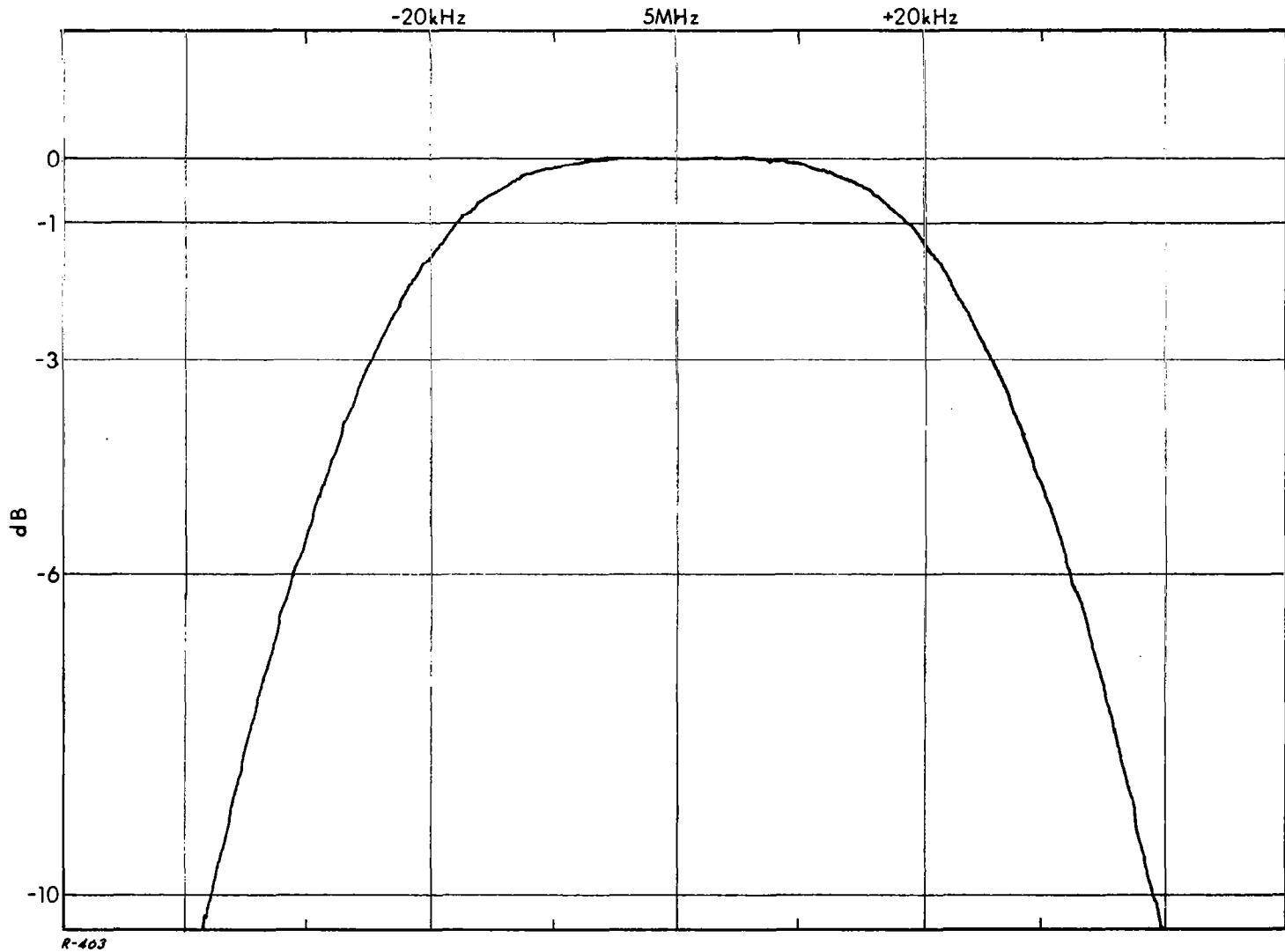


Fig. 14 a Frequency response of the 20 kc i-f filter.

55



R-463

Fig. 14b Frequency response of the 50 kc i-f filter.

56

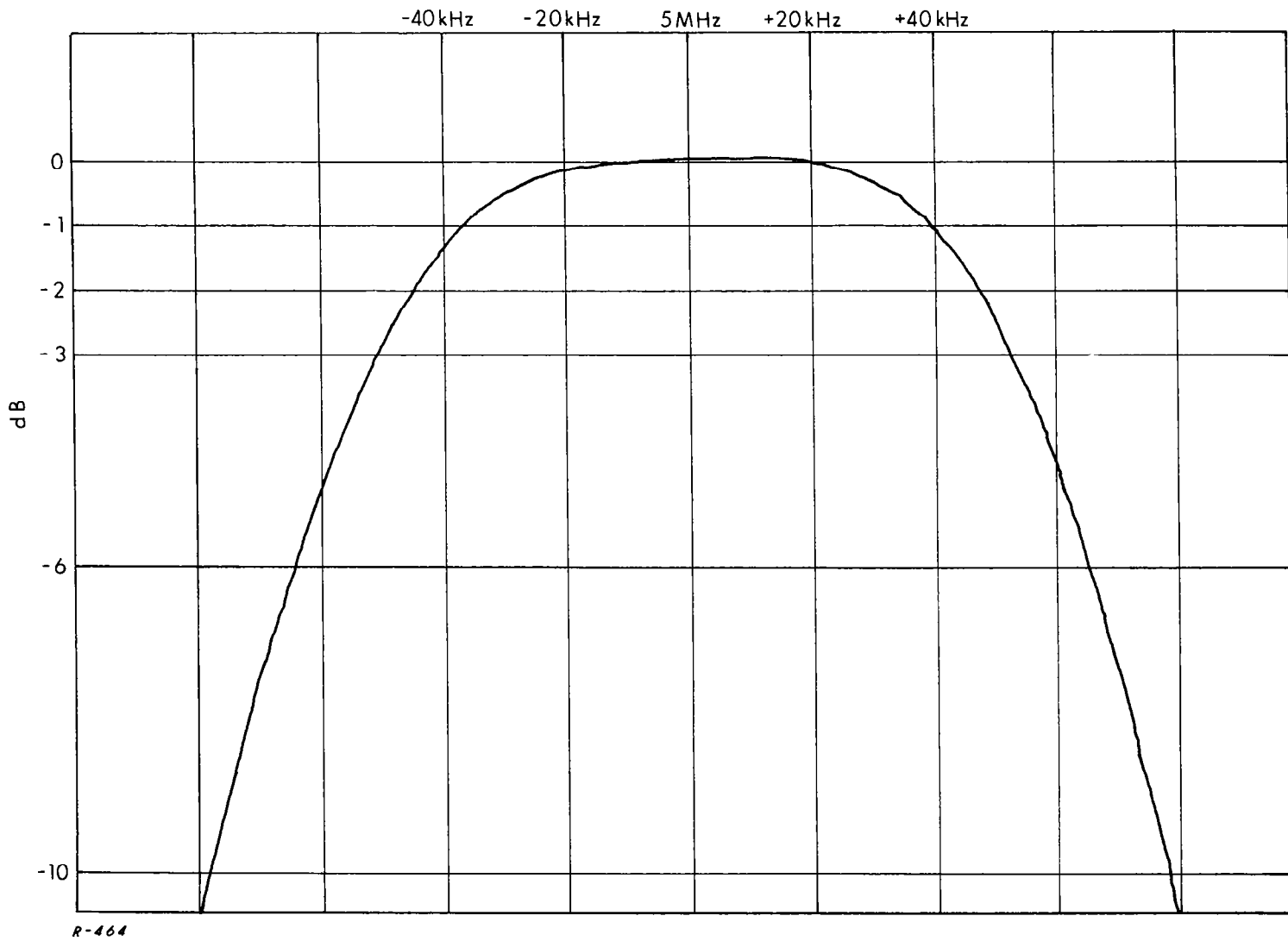


Fig. 14c Frequency response of the 100 kc i-f filter.

57

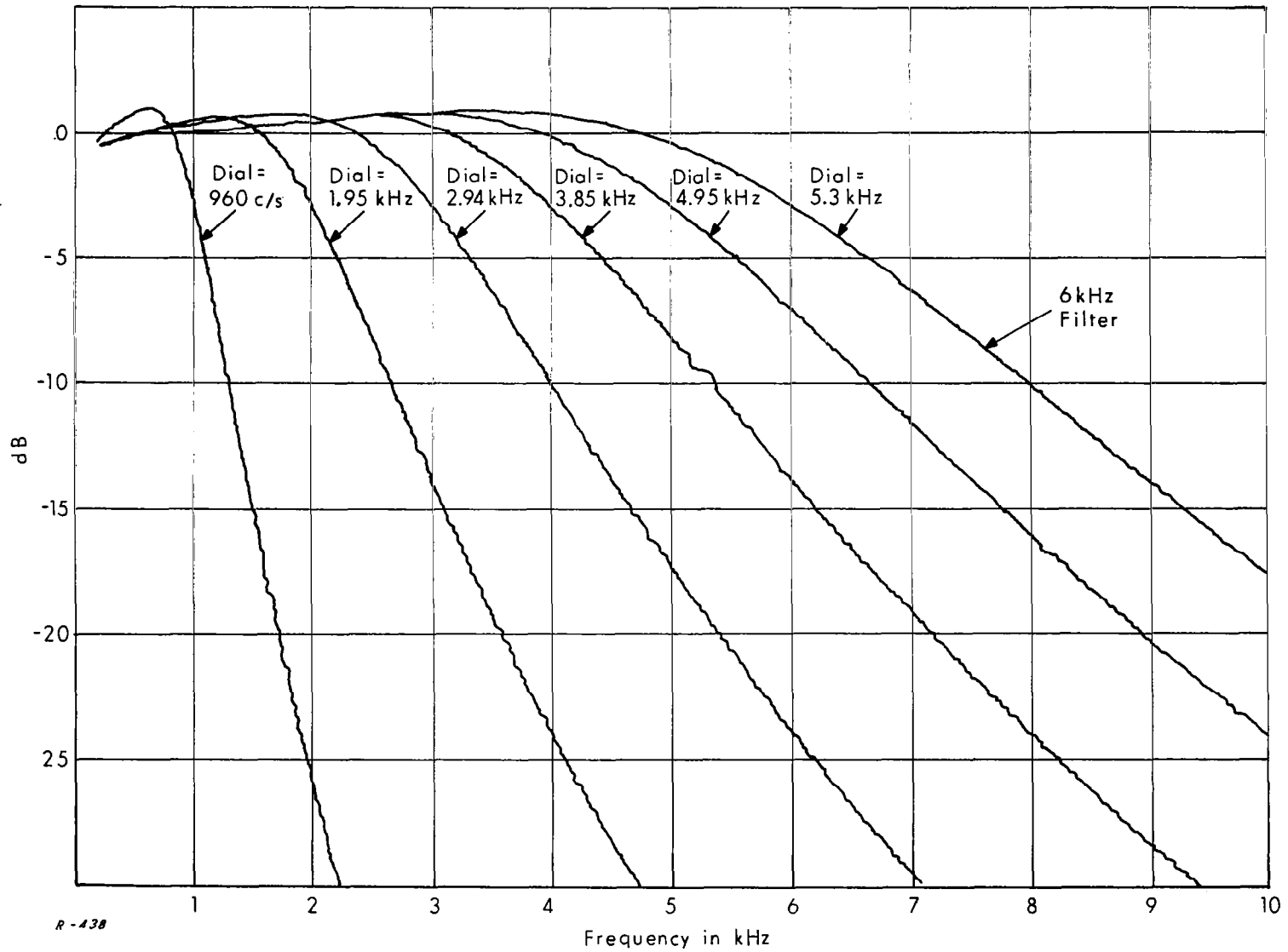


Fig. 15 Frequency response of output lowpass filters.

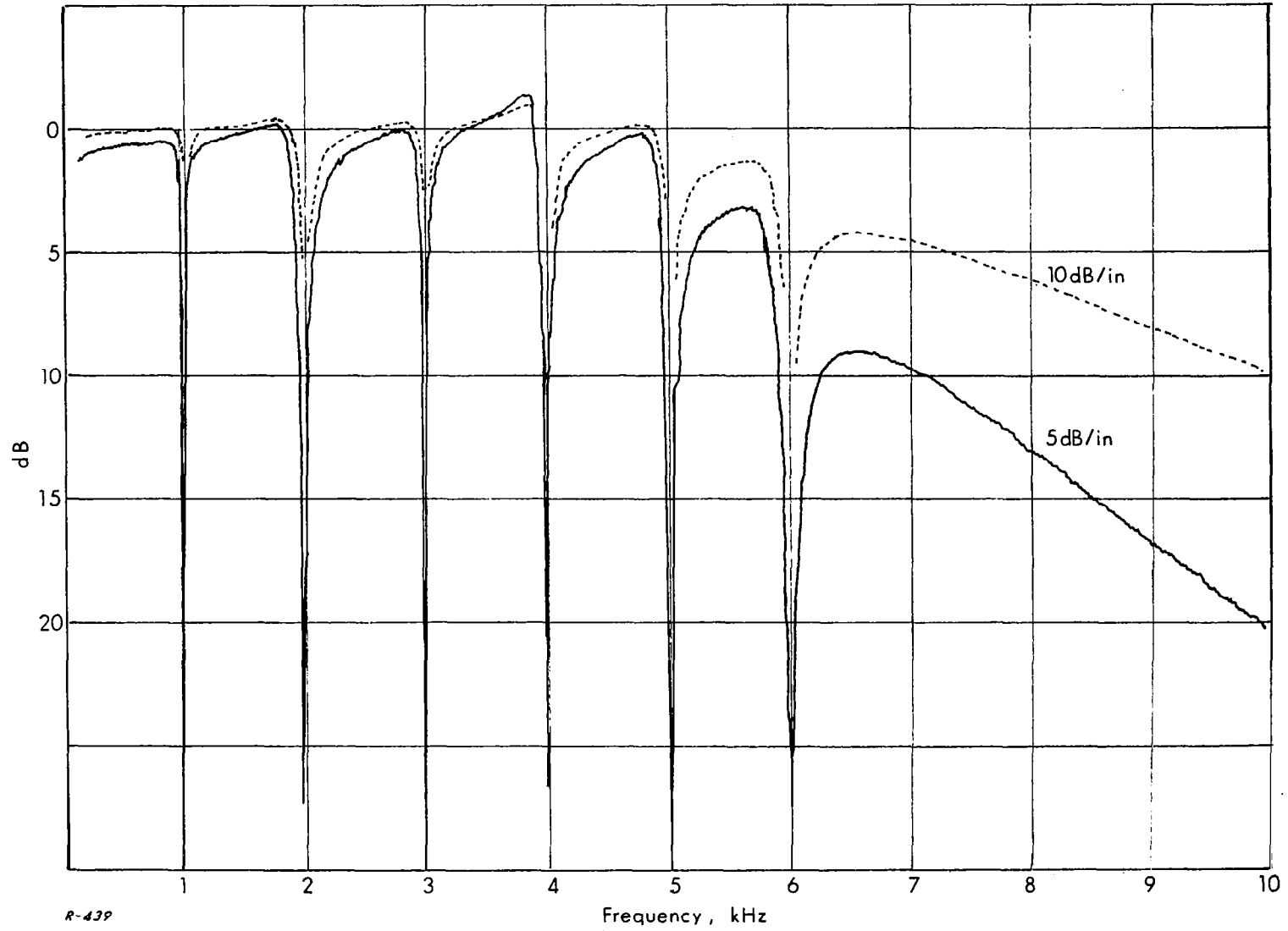
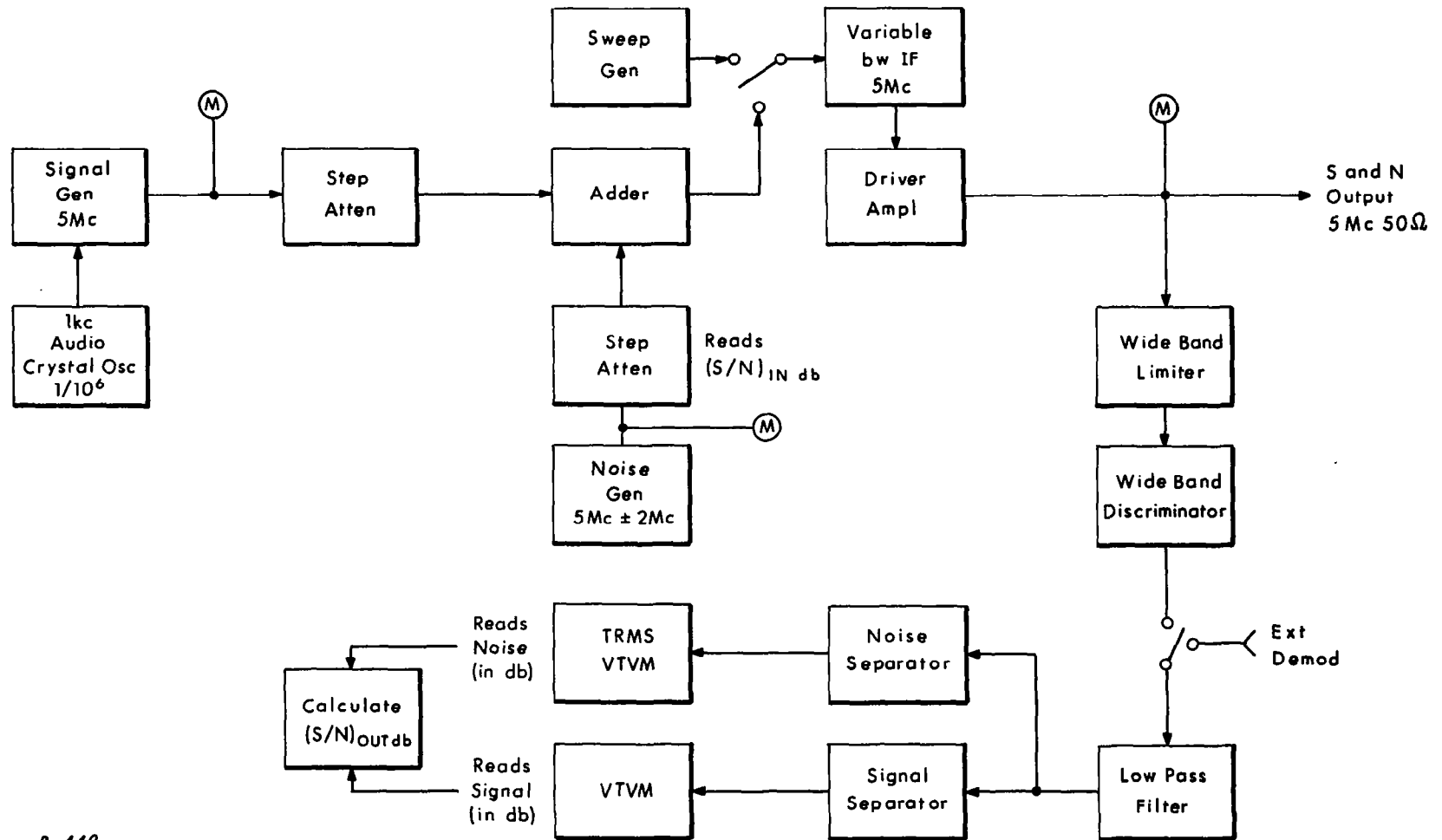


Fig. 16 Effect of notches on the 6 kc lowpass filter.

The experimental threshold is determined by reading the input SNR at which the departure from the linear dependence existing at high SNR reaches 0.5 db. This departure was measured by interpolating a straight line at high SNR and a smooth curve at low SNR in the continuous non-smooth curves obtained with the automatic plotter described in the next section.

The demodulator threshold test system is required to measure characteristics of the various FM demodulation schemes. The desire to obtain accurate results requires that the system allow for calibration repeatability to within one-half of a db ($\pm 5\%$) overall. To achieve this accuracy requires that the short-term (15-30 minute) stability of many of the individual components be ± 0.1 db or better. The long-term stability need not be of high accuracy because provision is made to calibrate the system at intervals. However, the requirements for calibration should be held to a minimum by providing monitoring circuitry.

One method of providing calibrated signal and noise levels is shown in the diagram of Fig. 17. The FM signal is generated, passed through a precision calibrated attenuator, and added to a noise source that has also been passed through an attenuator. The composite signal plus noise is filtered in stable measurable filters to provide precise noise bandwidths. This composite i-f signal feeds a linear driver amplifier having a peak capability of +20 db whose output is fed simultaneously to a wide bandwidth-limiter-discriminator (to be used as a standard of comparison for threshold improvements) and to the demodulator under test. The resultant demodulated signal is filtered with a calibrated lowpass filter. This provides a known baseband noise bandwidth for the precision measurements. The output signal level is measured with a narrowband spectrum analyzer and the noise level is measured with a true rms meter after the signal and its harmonics have been removed with a notch filter. Measurements on this system are taken by adjusting the input signal-to-noise ratio with the step attenuators in steps of one db. The



R-440

Fig. 17 Manual demodulator threshold test system.

output signal and noise levels are read and recorded. The threshold curve is plotted by calculating the signal-to-noise ratio at the output from the meter readings for each point on the curve.

The addition of certain laboratory equipment allows the time consuming point-by-point method to be replaced by a rapid and more accurate automatic plotting system. This system is shown in Fig. 18. The basic system remains unchanged with two exceptions. The step attenuator in the noise channel is replaced by a motor driven attenuator which is swept at 0.2 db per sec. The signal portion of the demodulator output is held constant by an AGC circuit to within ± 0.1 db. Provision is made for applying a d-c voltage proportional to the input signal-to-noise ratio (in db) to the x input of an x-y plotter. The y input of the plotter is derived from an audio logarithmic amplifier so that its output is proportional to the output signal-to-noise ratio in db.

5.2 The Conventional FM Demodulator Results

During the development of the program the question of whether or not the well-known FM improvement notion was directly measurable in the curves was asked. In order to answer this question it is necessary to be well aware of the assumptions inherent in any FM improvement expression. Under high SNR conditions, the output SNR of a conventional FM demodulator can be evaluated by dividing the separate contributions of signal and noise. The output signal power is given by

$$S_o = (\Delta F)^2 \overline{m^2(t)} \quad (5.1)$$

where ΔF is the frequency deviation, $m(t)$ is the FM message of unity amplitude and the discriminator constant is assumed to be unity since it affects signal and noise by the same amount. In the case of sinusoidal modulation this expression becomes

$$S_o = \frac{1}{2} (\Delta F)^2 \quad (5.2)$$

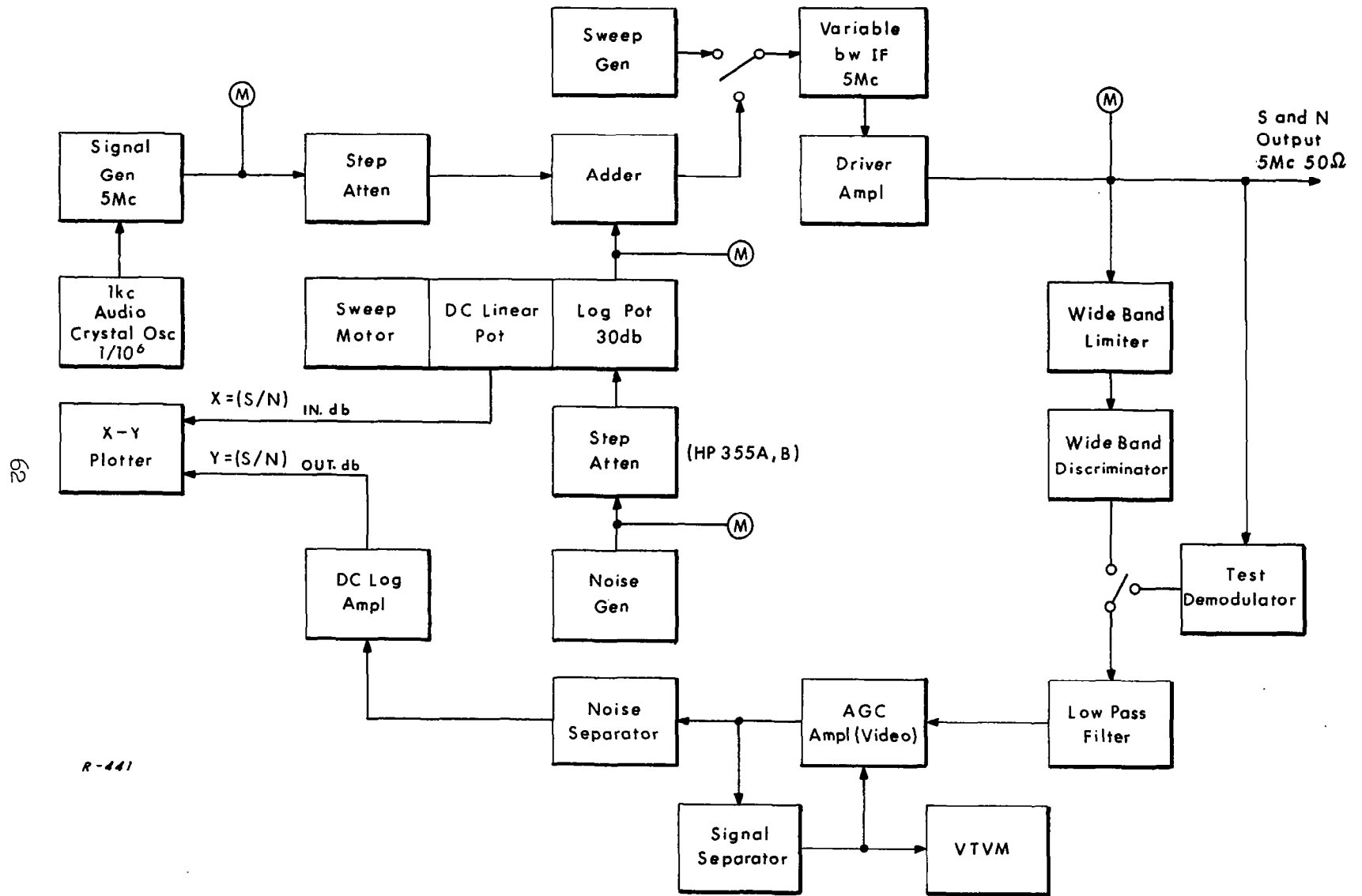


Fig. 18 Automatic demodulator threshold test system.

In turn, the output noise power is given by

$$N_o = \frac{\Phi}{S_1} \int_0^\infty f^2 |H_{\ell p}(j\omega)|^2 df \quad (5.3)$$

where $H_{\ell p}(s)$ represents the output lowpass filter. If a rectangular filter of width $B_{\ell p}$ is assumed as a first approximation, this expression reduces to

$$N_o = \frac{\Phi B_{\ell p}^3}{3 S_1} \quad (5.4)$$

and the resultant output SNR is thus

$$\left(\frac{S_o}{N_o}\right)_{\text{FM}} = \frac{3 S_1 (\Delta F)^2}{2 \Phi B_{\ell p}^3} \quad (5.5)$$

If we now denote the i-f noise bandwidth by B_{if} , this expression can be re-written as

$$\left(\frac{S_o}{N_o}\right)_{\text{FM}} = \frac{3}{2} \left(\frac{S_1}{N_1}\right)_{\text{FM}} \frac{(\Delta F)^2 B_{\text{if}}}{B_{\ell p}^3} \quad (5.6)$$

where $N_1 = \Phi B_{\text{if}}$ is the input noise power. If we now set

$$B_{\text{if}} = 2 B_{\ell p} \quad (5.7)$$

and

$$B_{\ell p} = f_m \quad (5.8)$$

then Eq. (5.6) becomes

$$\left(\frac{S_o}{N_o}\right)_{\text{FM}} = 3 \delta^2 \left(\frac{S_1}{N_1}\right)_{\text{FM}} \quad (5.9)$$

On the other hand, in AM systems the output and input SNR are related by

$$\left(\frac{S_o}{N_o}\right)_{\text{AM}} = \left(\frac{S_1}{N_1}\right)_{\text{AM}} \quad (5.10)$$

so that if we set $(S_1/N_1)_{\text{FM}} = (S_1/N_1)_{\text{AM}}$, i. e., we use not only the same carrier power but also the same i-f bandwidth (which could be absurd), then the FM improvement in output SNR over AM is given by

$$\left(\frac{S_o}{N_o}\right)_{\text{FM}} = 3 \delta^2 \left(\frac{S_1}{N_1}\right)_{\text{FM}} \quad (5.11)$$

and a db plot of $(S_o/N_o)_{\text{AM}}$ and $(S_o/N_o)_{\text{FM}}$ vs S_1/N_1 will show a 45° line for the AM case and a parallel line displaced by $10 \log 3 \delta^2$ db for the FM case.

However, if the conditions of Eqs. (5.7) and (5.8) are not met, as is the case in our tests, then Eq. (5.6) must be used (assuming the rectangular output filter approximation is adequate), and the FM line displacement (or y-axis intercept) will now be $10 \log [3(\Delta F)^2 B_{\text{if}}/2 B_{\text{lp}}^3]$ db. This relation can be checked for the different conventional FM curves presented, which are labeled STD (for standard) in the diagrams.

5.3 Description of The PLL Demodulators

A block diagram of the second order loop is shown in Fig. 19 and the corresponding circuits are given in Figs. 20 - 22. The loop filter components are chosen in accordance to the results of Chapter III. It is of interest to note that if we want to change the loop noise bandwidth while keeping the same values for the (K, a) parameters, this can be easily accomplished through

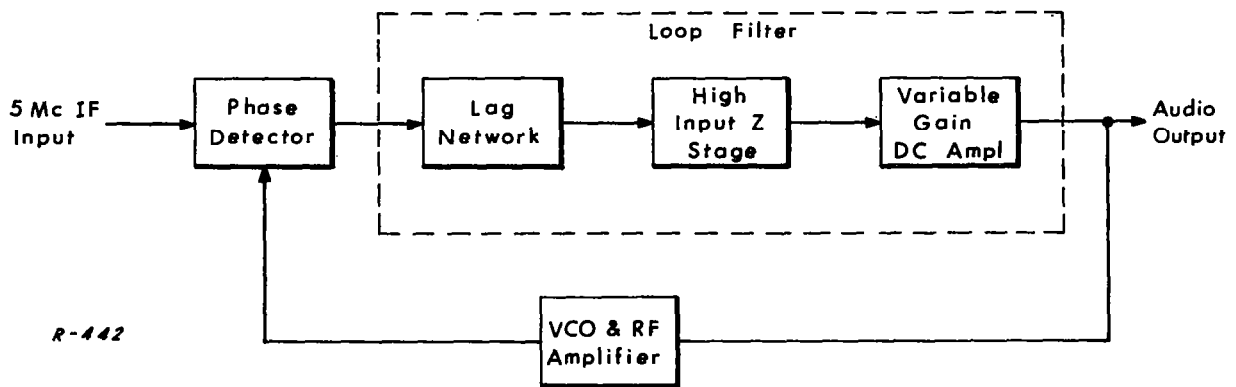


Fig. 19 Block diagram of second order PLL.

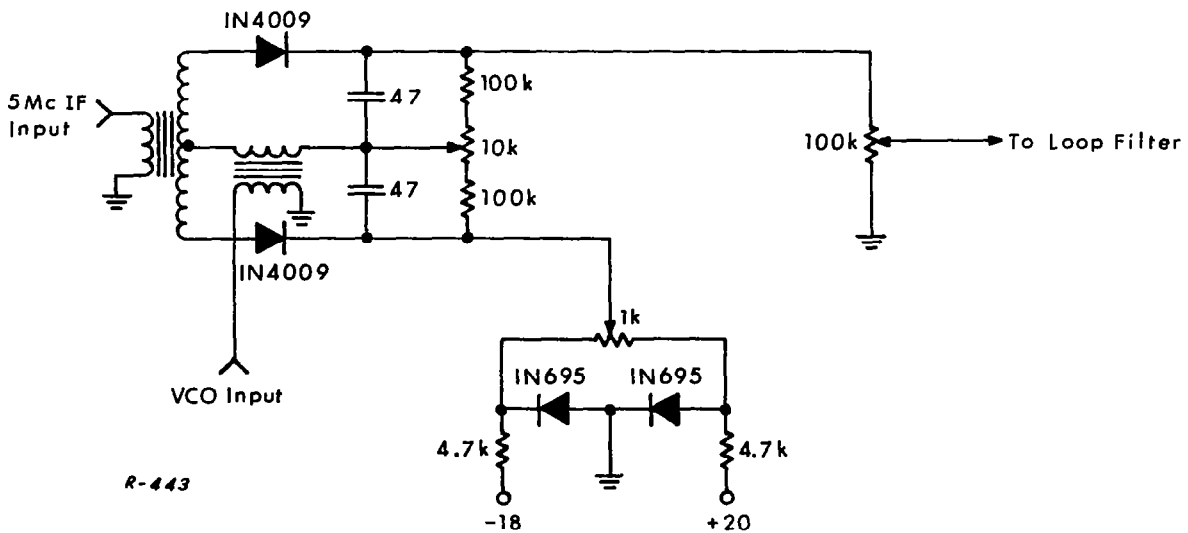


Fig. 20 Sinusoidal phase detector circuit.

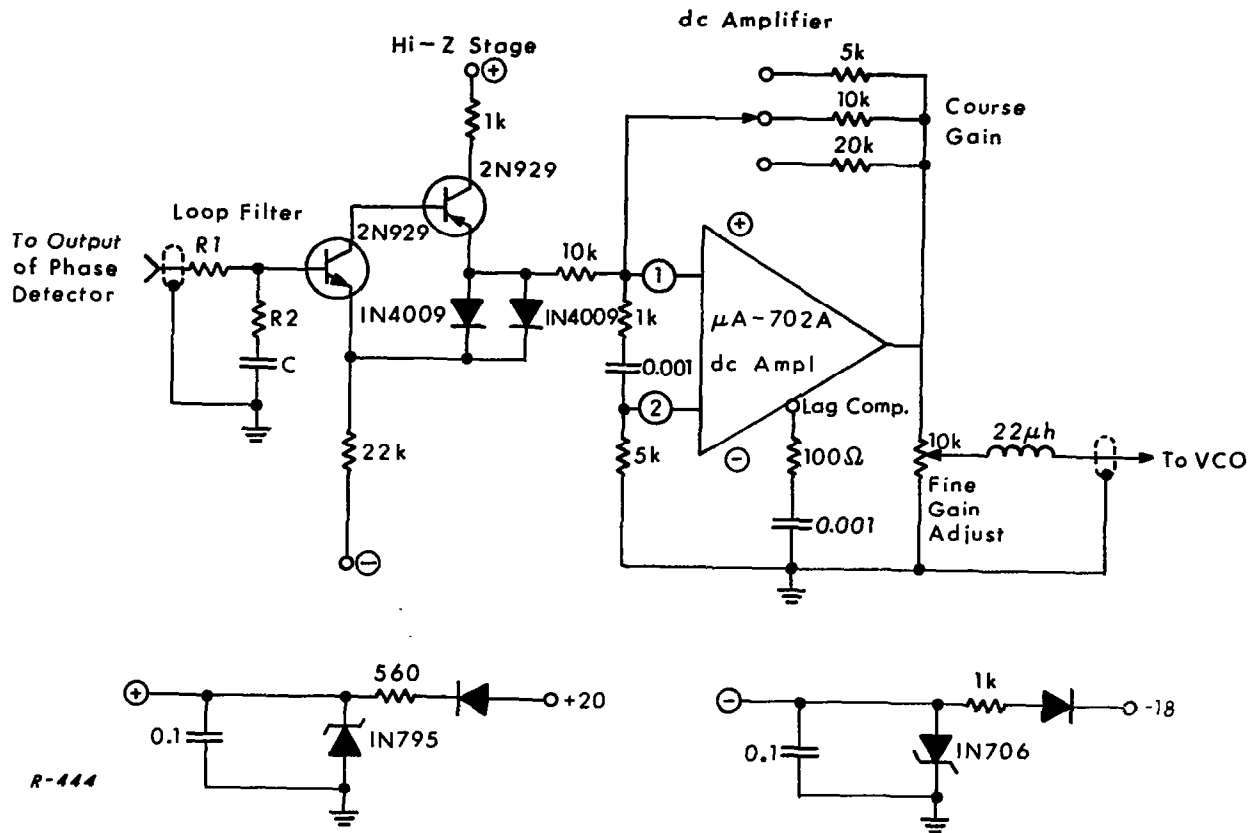


Fig. 21 Second order loop filter and DC amplifier circuits.

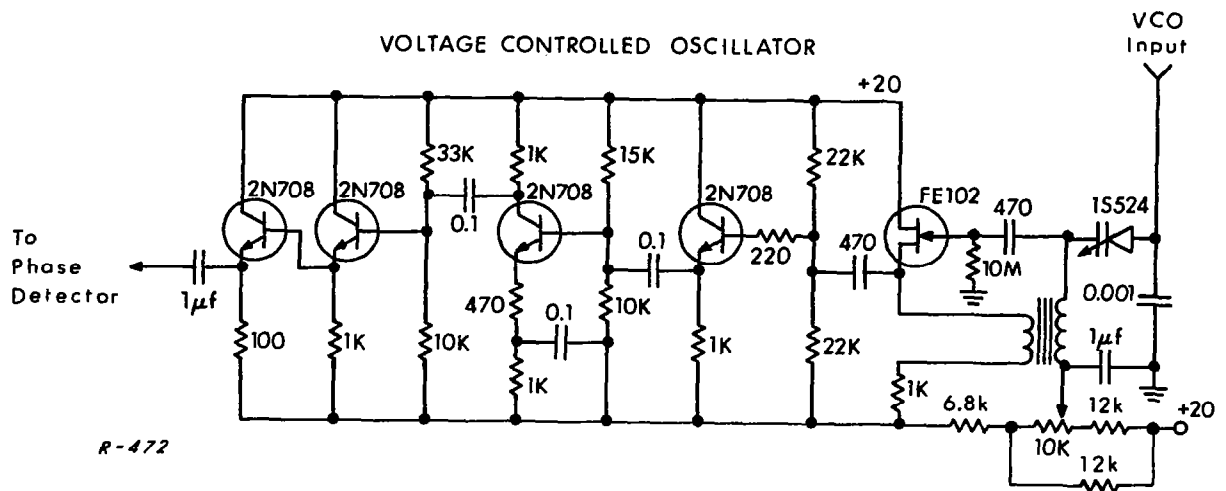


Fig. 22 VCO circuit.

the following transformations:

$$B_n \rightarrow rB_n \quad , \quad \tau_1 \rightarrow \frac{\tau_1}{r} \quad , \quad \tau_2 \rightarrow \frac{\tau_2}{r^2} \quad (5.12)$$

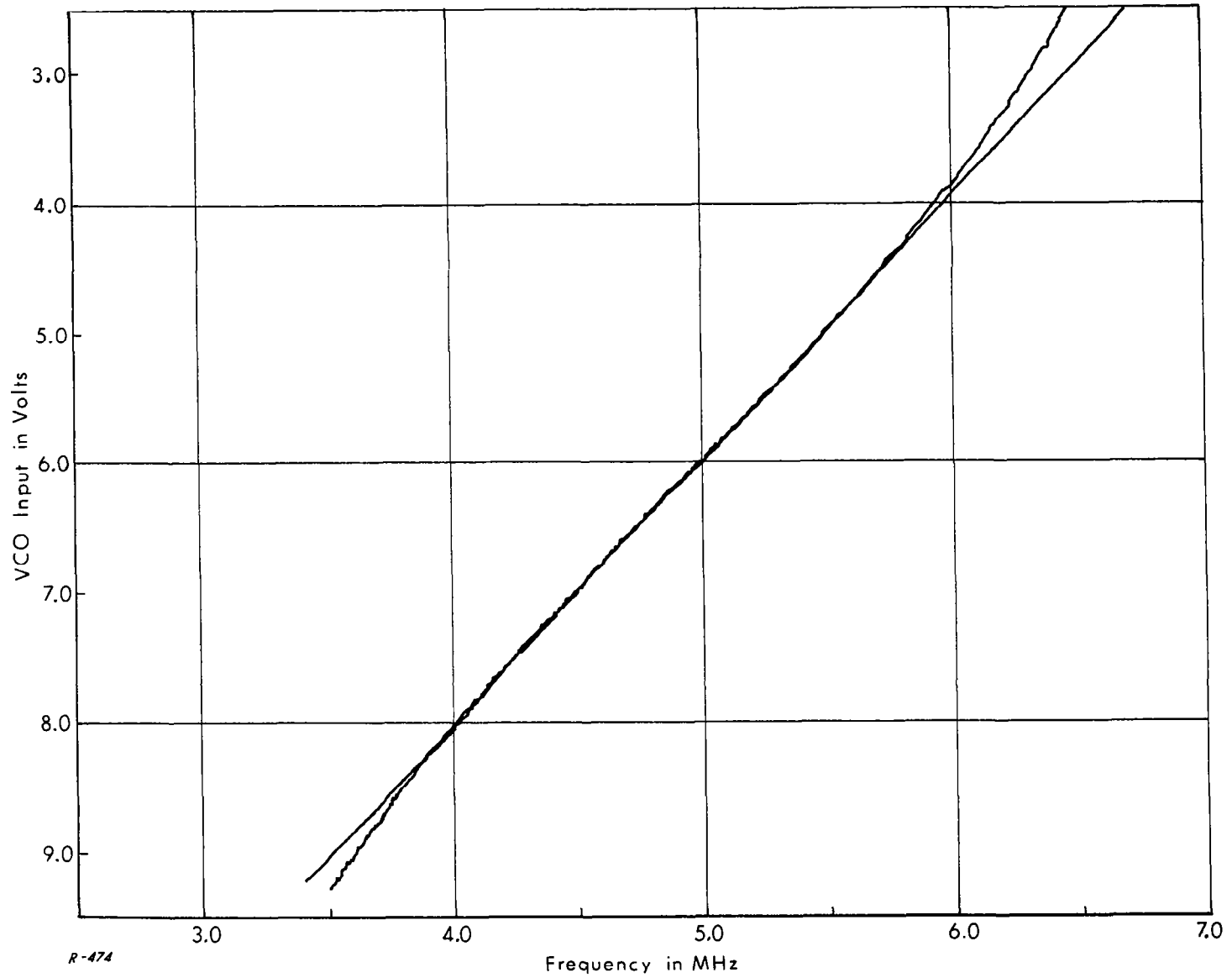
In the sinusoidal phase detector, the transformer turns ratio is chosen for maximum sensitivity while still maintaining the number of secondary turns below self-resonance. The maximum sensitivity is 1.5 volts per radian and is adjustable. Provision is also made to introduce a small d-c bias in series with the phase detector in order to compensate for any d-c offset component which may be introduced by additional circuitry in the loop.

Figure 21 shows the high input impedance stage used to match the output of the loop filter. This stage is designed to have zero d-c offset enabling it to be coupled directly to the following d-c operational amplifier. The gain of the operational amplifier can be varied coarsely in known steps by switching different feedback resistors into the circuit. A continuously variable gain control is also provided in the output.

The VCO, uses a variable-capacitance diode in a shunt tuned oscillator circuit. The diode was chosen for its linear characteristic and reversed biased to the center of its linear region. A field effect transistor is used as the oscillator to provide a high input impedance to the tuned circuit. The stages that follow provide isolation and gain to produce 1 v rms into the 50 Ω phase detector input impedance.

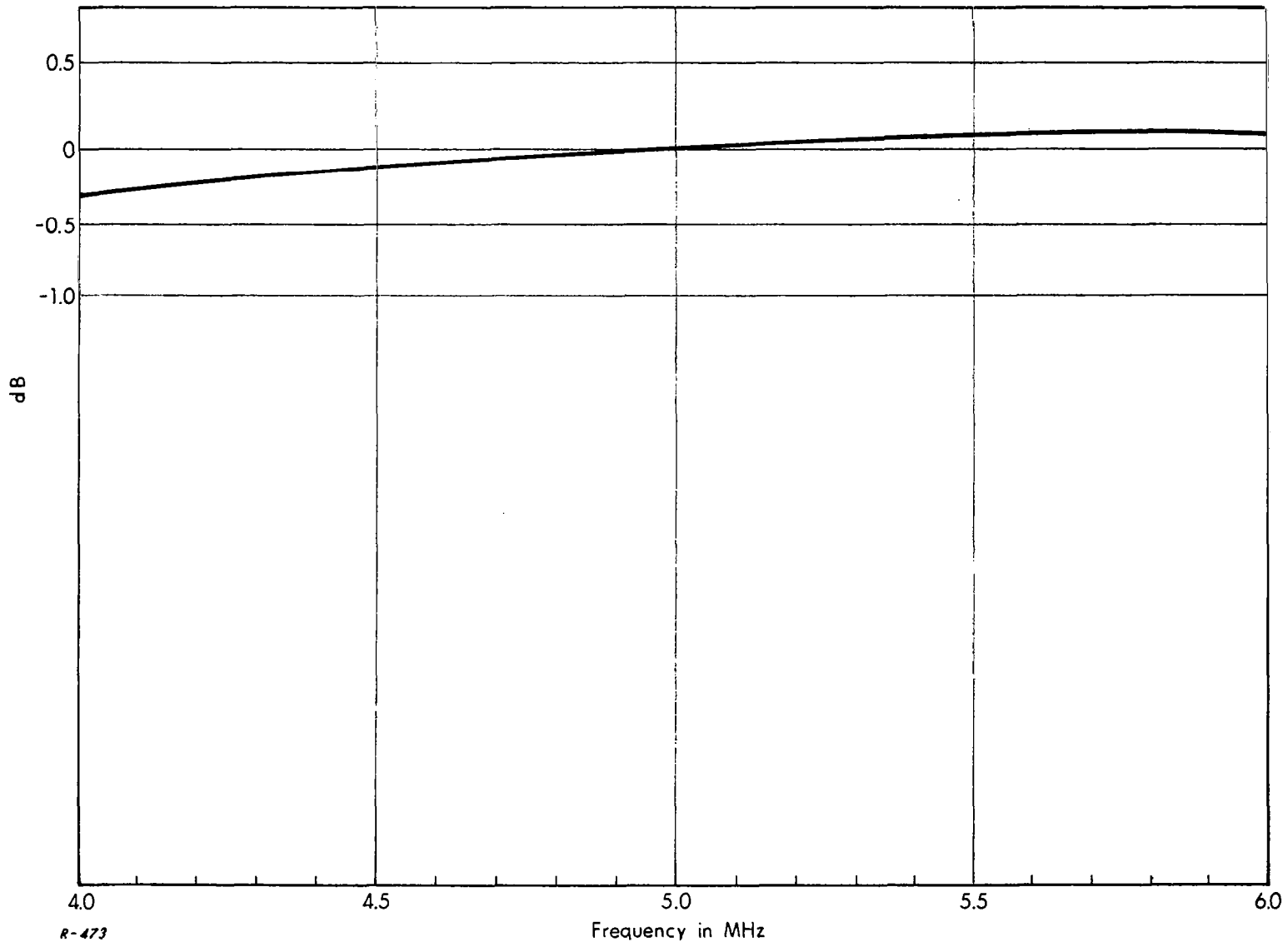
The overall VCO linearity is shown in Fig. 23 while its amplitude vs frequency characteristic is shown in Fig. 24. The VCO sensitivity is 500 kc/volt.

The third order loop differs from the above circuitry only in its loop filter, which is shown in Fig. 25 and designed in accordance to the results



R-474

Fig. .23 VCO linearity.



R-473

Fig. 24 VCO amplitude vs. frequency.

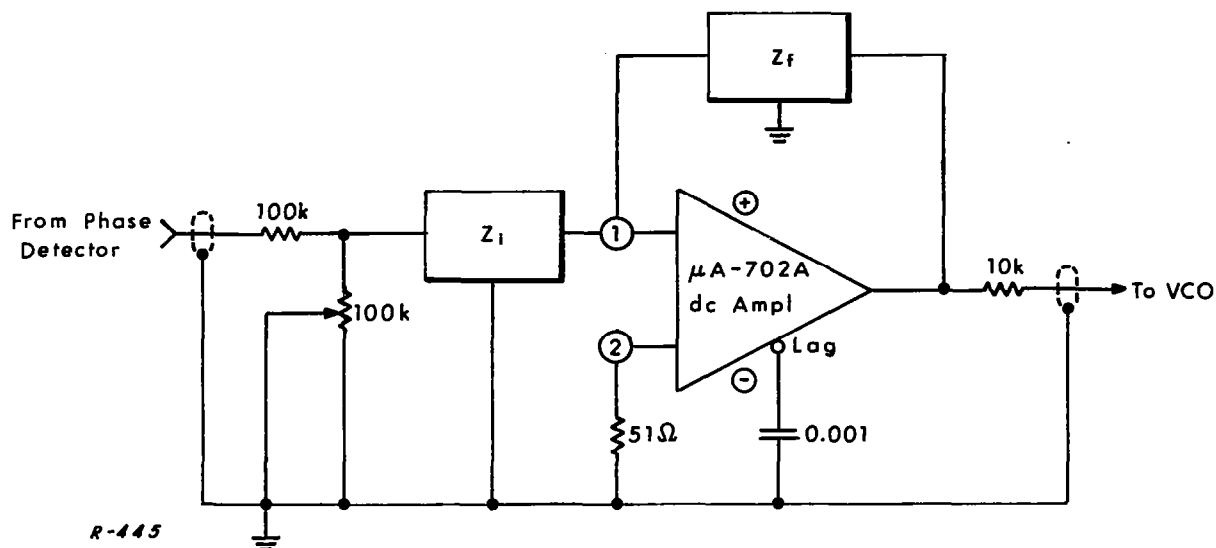


Fig. 25 Third order loop filter circuit.

of Chapter IV. A voltage divider in the input circuit provides for the adjustment of the loop gain. It is of interest to note that a change in the loop noise bandwidth can be accomplished through the following transformations, where the notation used is that of Chapter IV:

$$\begin{aligned}
 B_n &\rightarrow rB_n \quad , \quad C_1 \rightarrow \frac{C_1}{r} \quad , \quad C_2 \rightarrow \frac{C_2}{r} \quad , \\
 C_2' &\rightarrow \frac{C_2'}{r} \quad , \quad C_3' \rightarrow \frac{C_3'}{r^2}
 \end{aligned}
 \tag{5.13}$$

5.4 The Experimental PLL Thresholds

The PLL₂ and PLL₃ were designed on the basis of a $\sigma^2 = 1/4 \text{ rad}^2$ threshold error and the experimental curves are respectively shown in

Figs. 26 a-c for the PLL₂ (corresponding to $\delta = 10, 20, 30$) and in Figs. 27 a-b for the PLL₃ (corresponding to $\delta = 10, 20$). The thresholds were determined according to the 0.5 db linearity departure criterion previously explained and are indicated by vertical arrows. The experiment results are summarized in the first seven rows of the table that follows, where the demodulator entries represent the input SNR at threshold in db.

In the PLL₂ case, the theoretical thresholds for $\sigma^2 = 1/4 \text{ rad}^2$ and $\sigma^2 = 1/10 \text{ rad}^2$ can be derived from Eq. (3.24) normalized to the tabulated noise bandwidths and they are indicated in the 8th and 9th rows in the table. The experimental results hints that $1/10 < \sigma^2 < 1/4 \text{ rad}^2$ represents a proper range for the theoretical threshold error. A check of this statement was obtained by redesigning for $\delta_{\sigma}^2 = 1/10 \text{ rad}^2$ and noting that the experimental threshold was still in the theoretical range predicted by $1/10 < \sigma^2 < 1/4 \text{ rad}^2$. A similar result occurred when the PLL₂ was redesigned for a rectangular PM spectrum according to Eqs. (3.12) and (3.14) and for the two threshold errors in question. As a matter of fact, the optimum theoretical PLL₂

(1) δ	10	20	30	} experimental, sinusoidal modulation
(2) $B_{if}(kc)$	23	54	110	
(3) STD	9.0	9.0	9.0	
(4) PLL ₂	7.5	5.75	5.0	
(5) PLL ₃	6.5	-	3.5	
(6) STD/PLL ₂	1.5	3.25	4.0	
(7) PLL ₂ /PLL ₃	1.0	-	1.5	
(8) PLL ₂ ($\sigma^2=1/4$)	5.9	3.6	1.5	} theoretical, sinusoidal modulation (the rectangular spectrum entries are 1 db lower)
(9) PLL ₂ ($\sigma^2=1/10$)	10.9	8.6	6.5	
(10) PLL ₃ ($\sigma^2=1/4$)	3.0	0.7	-1.5	} theoretical, rectangular spectrum
(11) PLL ₃ ($\sigma^2=1/10$)	8.0	5.7	3.5	
(12) PLL ₂ /PLL ₃ ($\sigma^2=1/4$)	1.8	1.9	2.0	
(13) PLL ₂ /PLL ₃ ($\sigma^2=1/10$)	1.5	1.8	2.0	

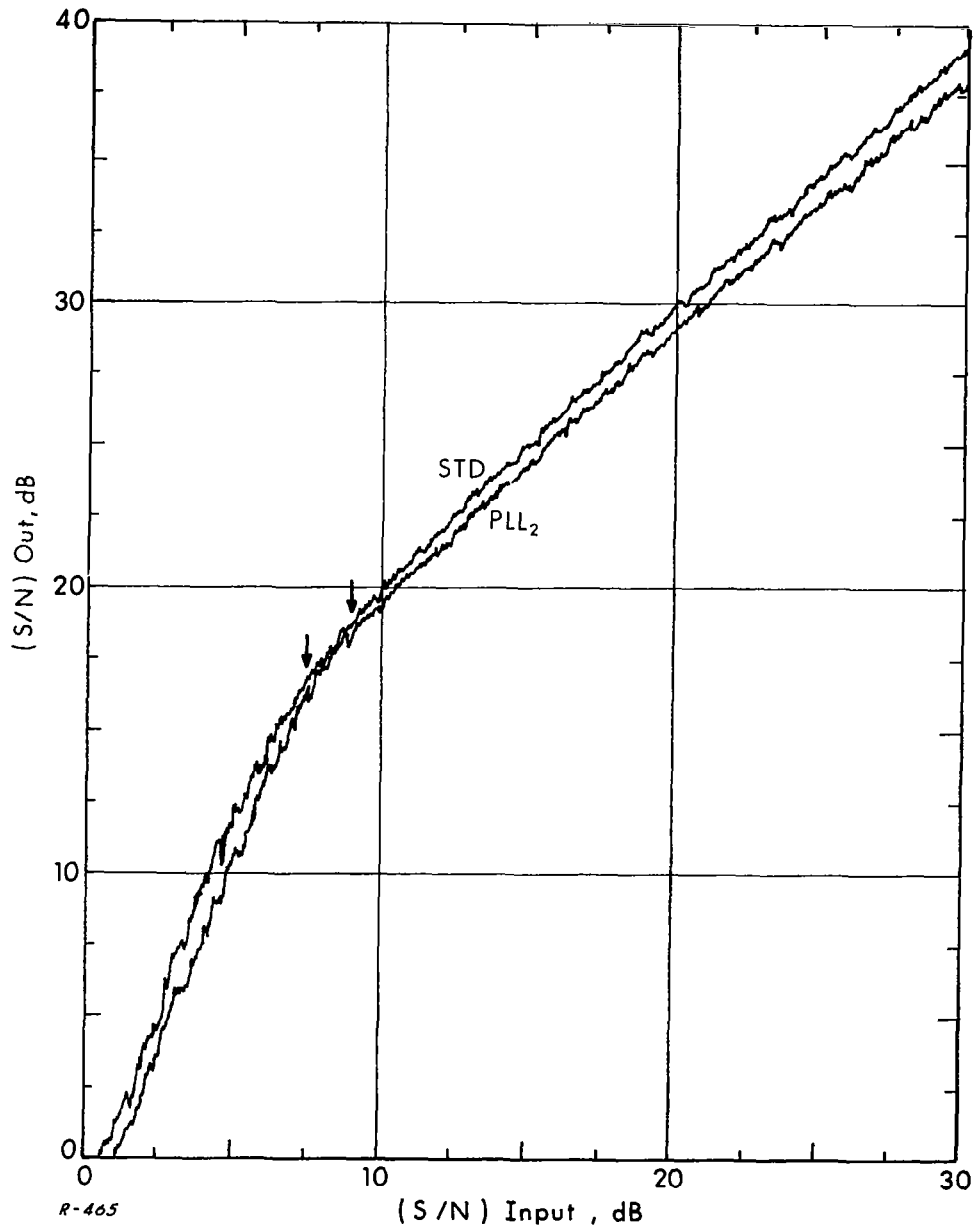


Fig. 26a PLL₂ tests ($\delta = 10$, $f_m = 1$ kc, $B_{if} = 23$ kc).

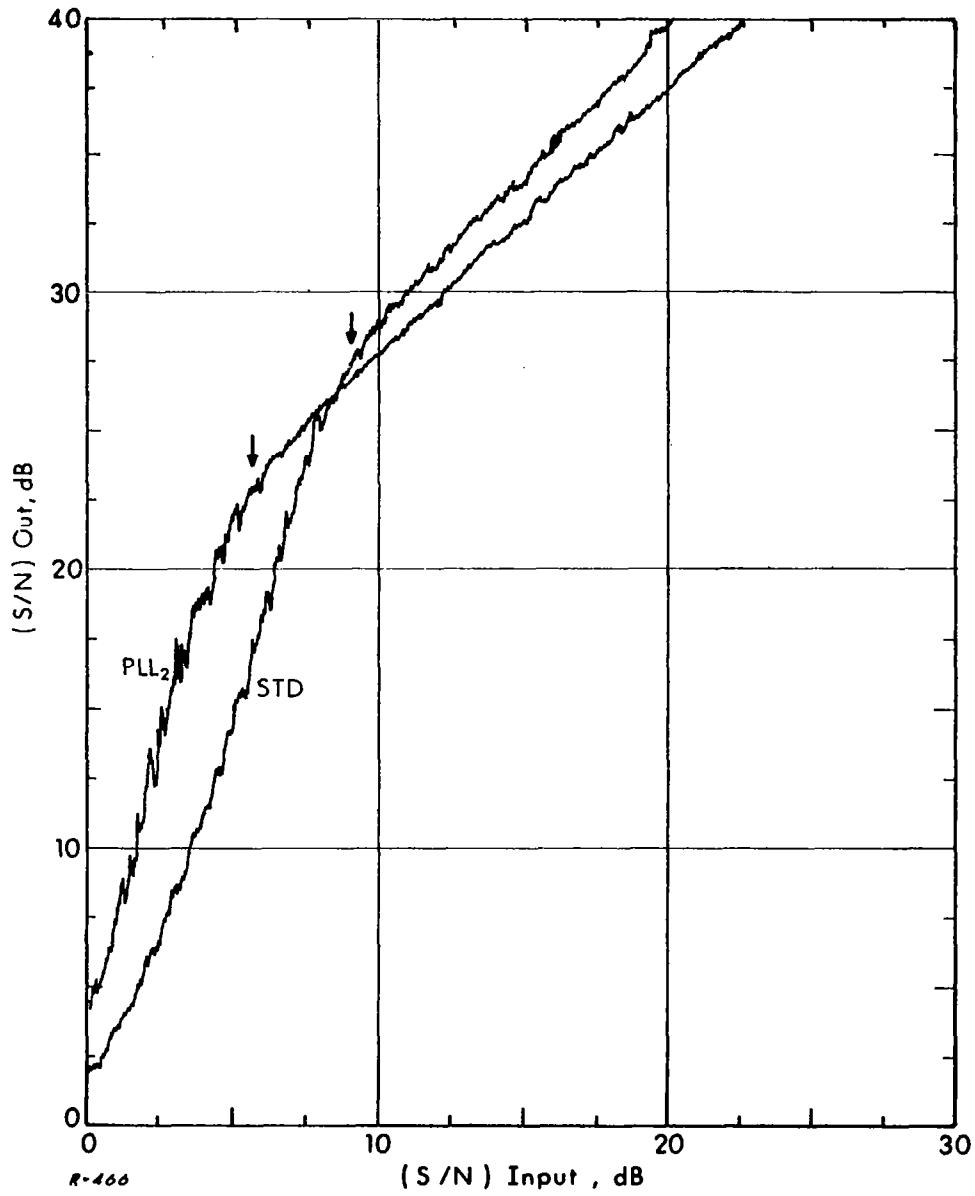


Fig. 26b PLL_2 tests ($\delta = 20, f_m = 1 \text{ kc}, B_{if} = 54 \text{ kc}$).

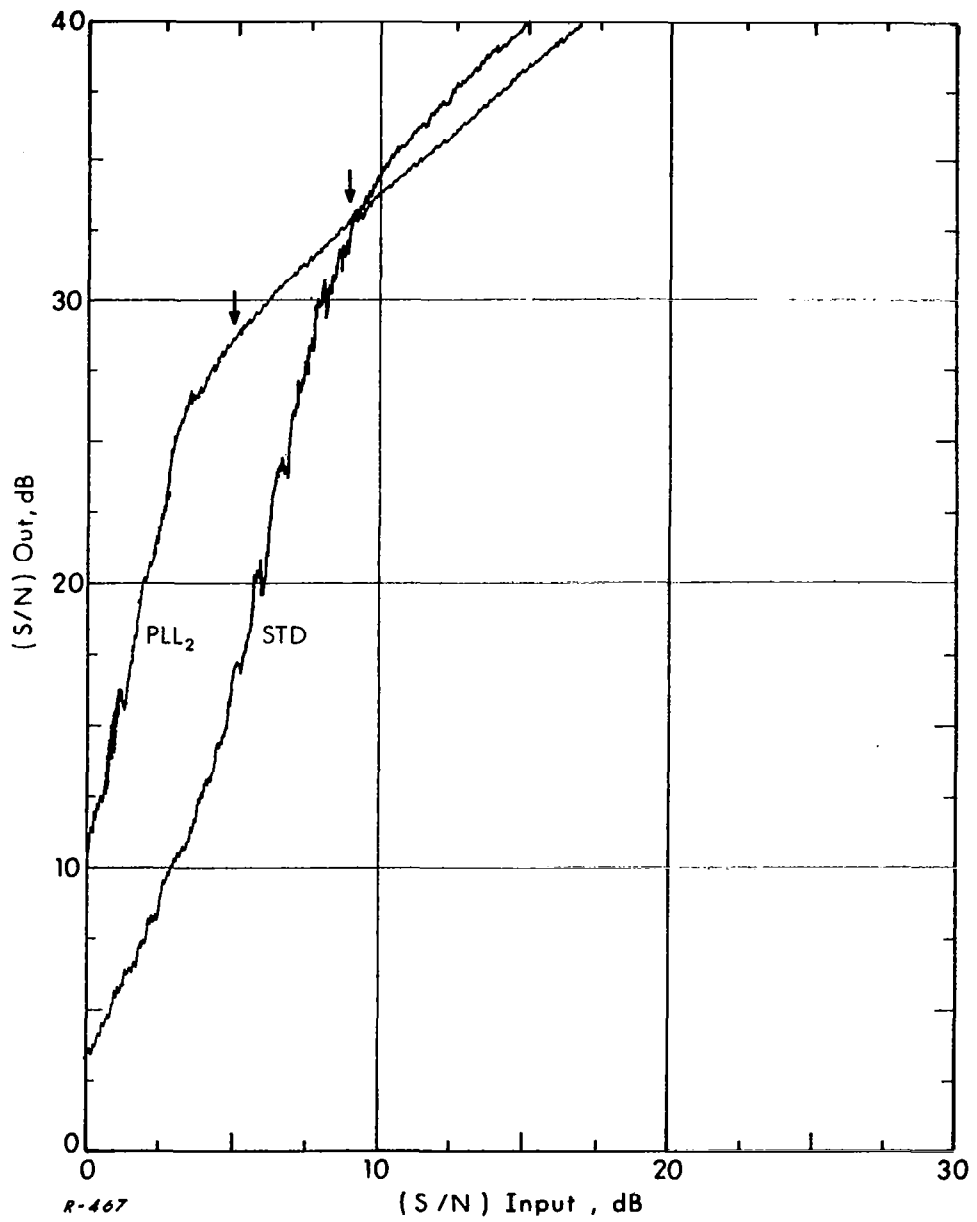


Fig. 26c PLL_2 tests ($\delta = 30$, $f_m = 1$ kc, $B_{if} = 110$ kc).

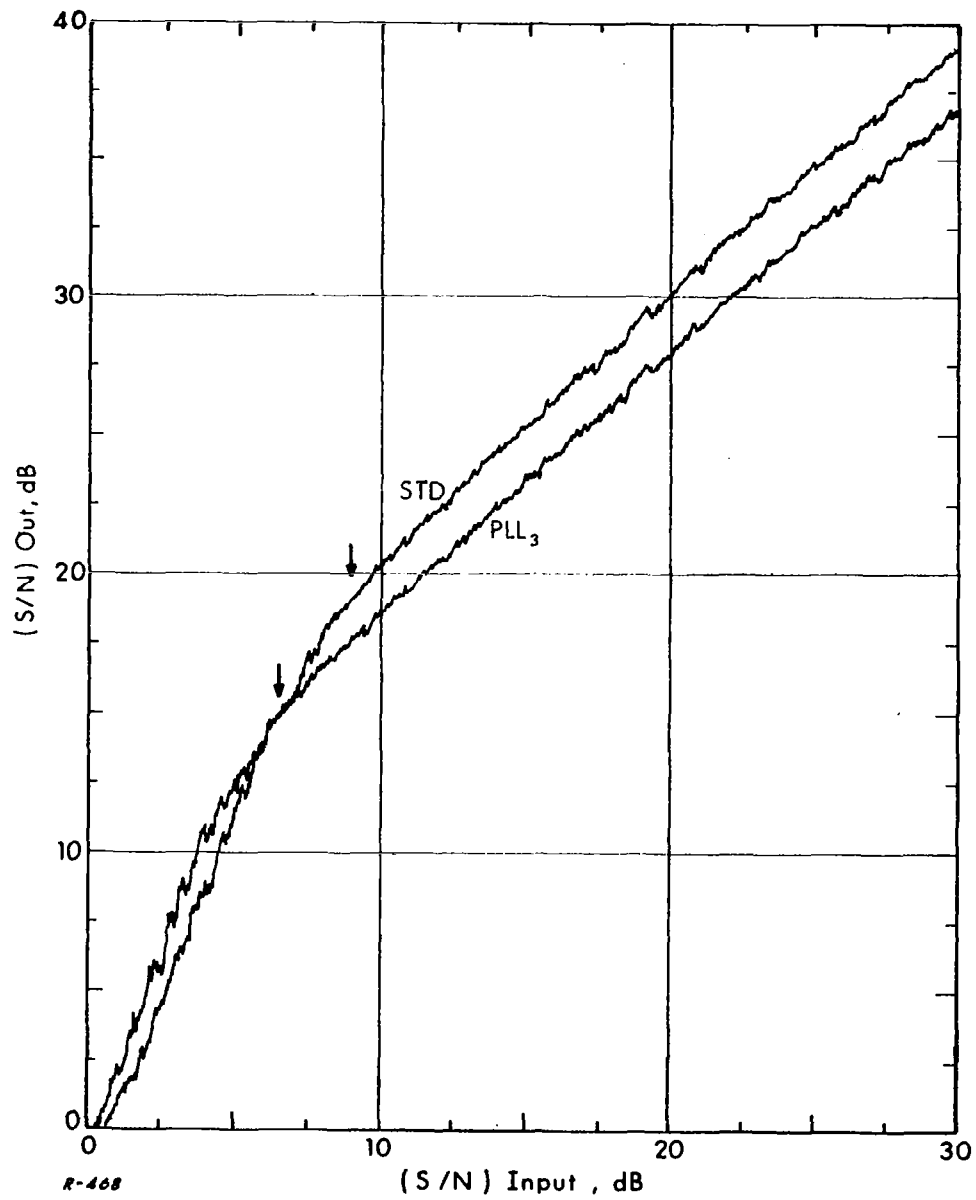


Fig. 27a PLL₃ tests ($\delta = 10$, $f_m = 1$ kc, $B_{if} = 23$ kc).

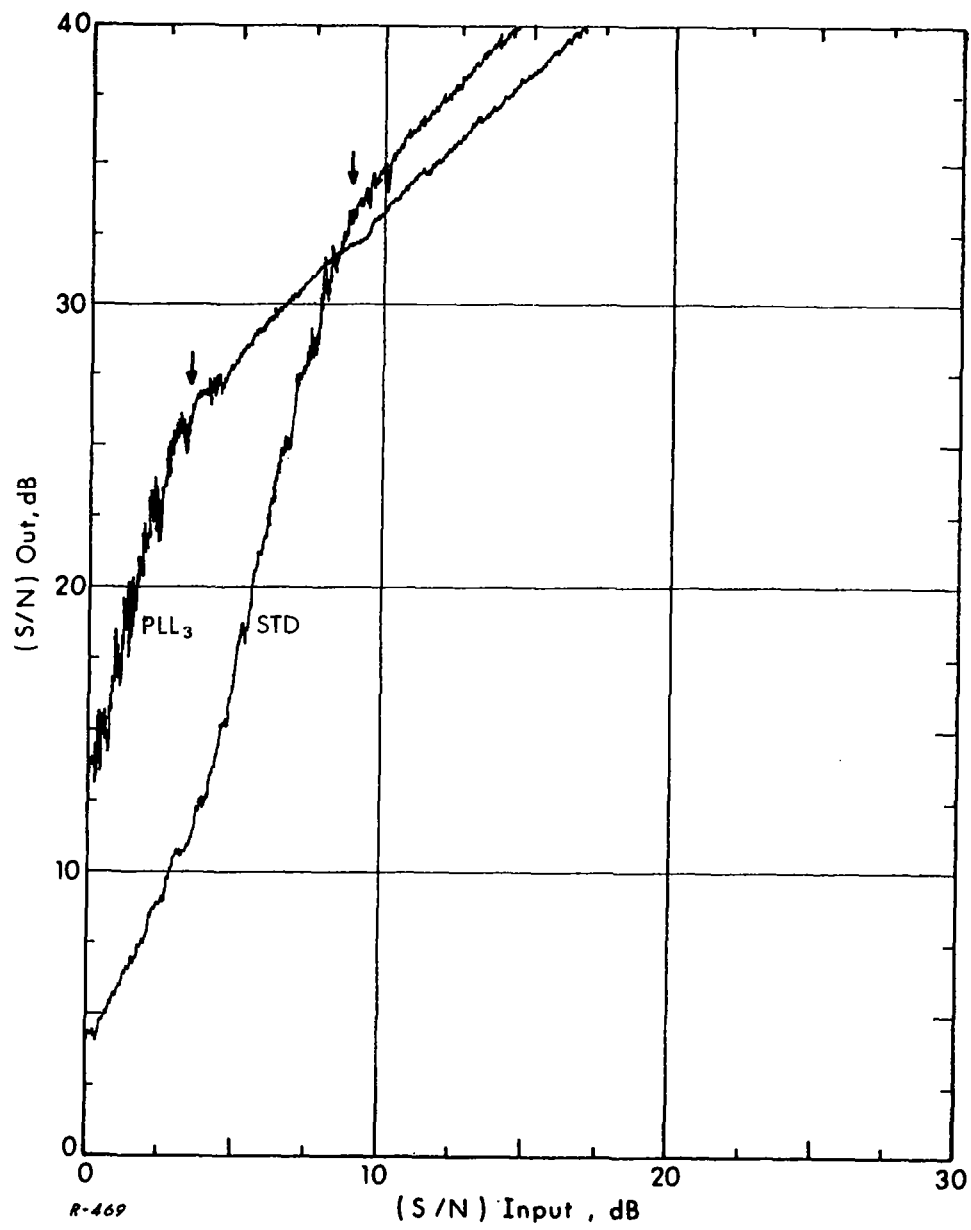


Fig. 27b PLL_3 tests ($\delta = 30$, $f_m = 1$ kc, $B_{if} = 110$ kc).

for sinusoidal modulation and $\sigma^2 = 1/4 \text{ rad}^2$ is the same as that for the rectangular PM spectrum and $\sigma^2 = 1/10 \text{ rad}^2$.

On the basis of these results, it seems useful to include the PLL₃ theoretical results for $\sigma^2 = 1/10 \text{ rad}^2$. The extension is not as simple as in the PLL₂ case where an explicit dependence on σ was available, and the elaborate procedure described in Chapter IV and used to derive Fig. 10 must be repeated. The resultant curves analogous to Fig. 10 are shown in Fig. 28 and the theoretical PLL₃ thresholds are indicated in the 10th and 11th rows in the table. The experimental thresholds were again found to be compatible for threshold errors in the range $1/10 < \sigma^2 < 1/4 \text{ rad}^2$, perhaps closer to the former. The theoretical threshold improvement of the PLL₃ over the PLL₂ is shown in the 12th and 13th rows in the table. There exists sufficient agreement with the experimental results so as to use the curves of Figs. 10 and 28 in order to evaluate the threshold improvement capabilities of higher order loops and the degree of loop filter complexity required for a given improvement.

Finally, it is of interest to comment on the fluctuations of the PLL output SNR, relative to the STD, at high input SNR conditions. This effect occurs because the loop filter design criterion has been to optimize the loop threshold performance and not to increase the output SNR under small noise conditions. This alternate design could be accomplished but at the expense of the threshold improvement. Still, once the loop is designed for optimum threshold performance, its output SNR at large input SNR conditions can be improved by post detection filtering without affecting the threshold.

5.5 The Sawtooth Phase Detector

In accordance to the discussion of Chapter II, the nonlinear reduction in output SNR that occurs at threshold is mainly due to the presence of new disturbances in the output signal caused by cycle slippages in the phase detector. The detector characteristic and the closed loop response determine

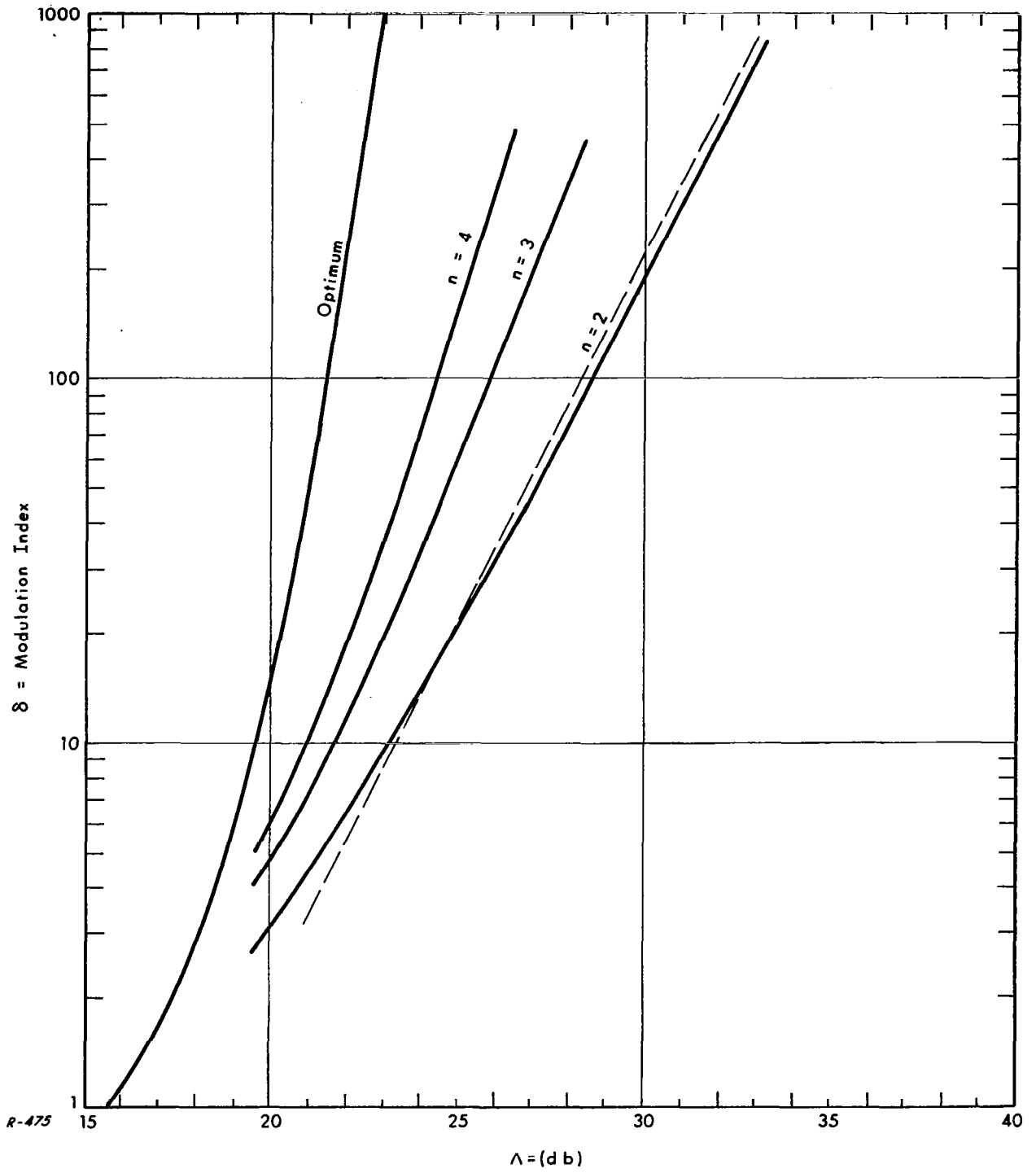


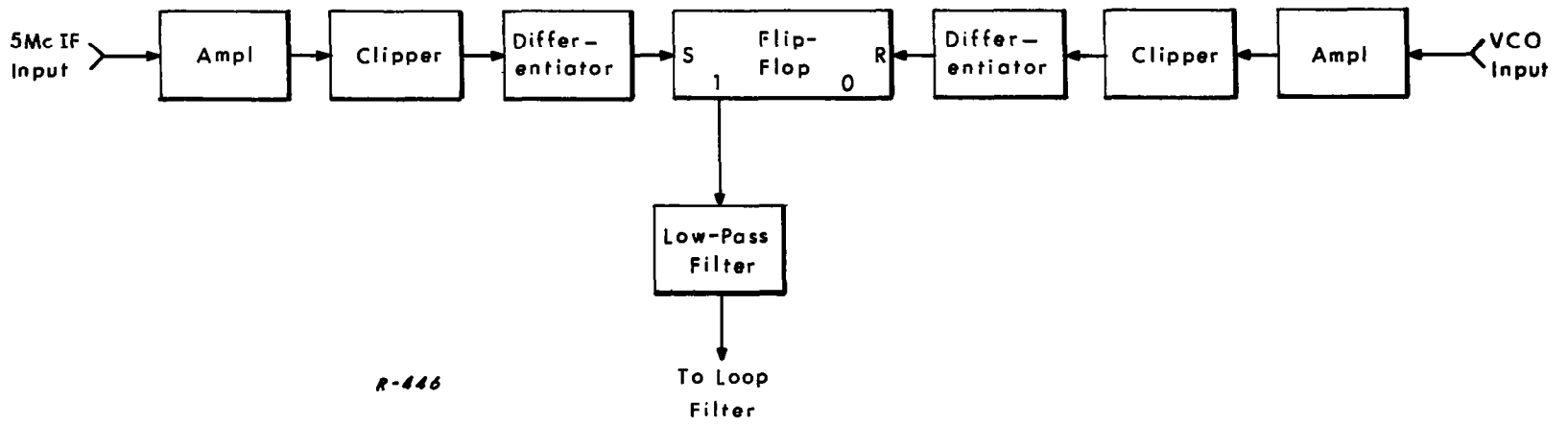
Fig. 28 Higher order PLL threshold curves for $\sigma^2 = 1/10 \text{ rad}^2$.

the shape and power per disturbance and they also control the frequency of occurrence of these disturbances, i. e., the detector characteristic defines the critical error beyond which a cycle slippage occurs while the closed loop response defines the noise error statistics.

Thus, it seems that some threshold improvement can be achieved by using a phase detector characteristic with a larger critical error than the $\pi/2$ rad existing in a conventional sinusoidal phase detector, provided that the different output power per cycle slippage in the two cases is a secondary effect when compared to the frequency of occurrence. On this basis, we decided to build a sawtooth phase detector having the characteristic of Fig. 4. The linear model is thus valid always as long as the critical error of π rad is not exceeded and threshold appears when phase error peaks above this value become regular. The phase noise statistics are gaussian for phase errors in the absence of a cycle slippage so that a probability measure of phase error peaks can be easily obtained. The mean-square threshold error can be selected according to this criterion or by trial and error using empirical results as in the previous section. In any case, this threshold error is expected to be larger than that of the sinusoidal detector so that theoretical thresholds can be computed using Eqs. (3.14) and (3.24) for the PLL_2 and developing plots analogous to Figs. 10 and 28 for higher-order loops.

The synthesis of the sawtooth phase detector is summarized by the block diagram of Fig. 29. The i-f and VCO signals are amplified, clipped and differentiated in two identical channels to produce short pulses characterizing to their zero crossings and which are used to control a high speed flip-flop. The i-f pulses are sent into the set terminal of the flip-flop and the VCO pulses into the reset terminal. Therefore, the time spent in the set state will be the time existing between the i-f pulse and the VCO pulse.

If the flip-flop puts out a positive voltage in the set state and an equal negative voltage in the reset state, the average output voltage will be a linear function of the phase error. The average output will be zero when the pulses



R-446

Fig. 29 Block diagram of sawtooth phase detector.

are 180° out of phase. If the phase error exceeds $\pm \pi$ the pulses will pass each other. There will be a sudden discontinuity, and the voltage will change quickly from one extreme to the other. If the i-f signal is turned off, the flip-flop acts like a binary counter, driven by the VCO signal, and the average output will be zero.

The average voltage is extracted from the flip-flop output by the lowpass filter. The filter cutoff frequency is low enough to remove signal harmonics.

The corresponding circuit diagram is shown in Fig. 30. Identical circuits are used to obtain pulses from the VCO and i-f signals. Transistors Q_1 amplify the input signals which are then clipped in the emitter coupled clipper circuit formed by Q_2 , Q_3 , Q_4 and Q_5 . The resulting clipped signal is differentiated by the R-C network. The resulting i-f spikes are sent into trigger transistor Q_9 , the set terminal of the flip-flop and the VCO pulses are sent into trigger transistor Q_6 , the reset terminal of the flip-flop. Transistors Q_7 and Q_8 are part of the actual flip-flop circuit. The rise time of the pulses is 8 nanoseconds which would imply a maximum phase uncertainty of $\pm 15^\circ$ at 5 mc.

The output voltage of the flip-flop is taken at the collector of Q_8 and sent to Q_{10} where its d-c level is changed to be compatible with the loop requirements. The lowpass filter is included in the collector circuit of Q_{10} .

The overall detector characteristic is illustrated in Fig. 31 which was obtained by recording the detector output when two different frequencies are applied to its inputs. The curve shows a sensitivity of 1 rad/volt but it was adjusted to operate at 0.7 rad/volt, which is the same operational sensitivity as the sinusoidal detector.

5.6 Sawtooth Detector Experiments

The threshold error of the sawtooth phase detector, which controls the optimum loop design, was established experimentally by comparing the

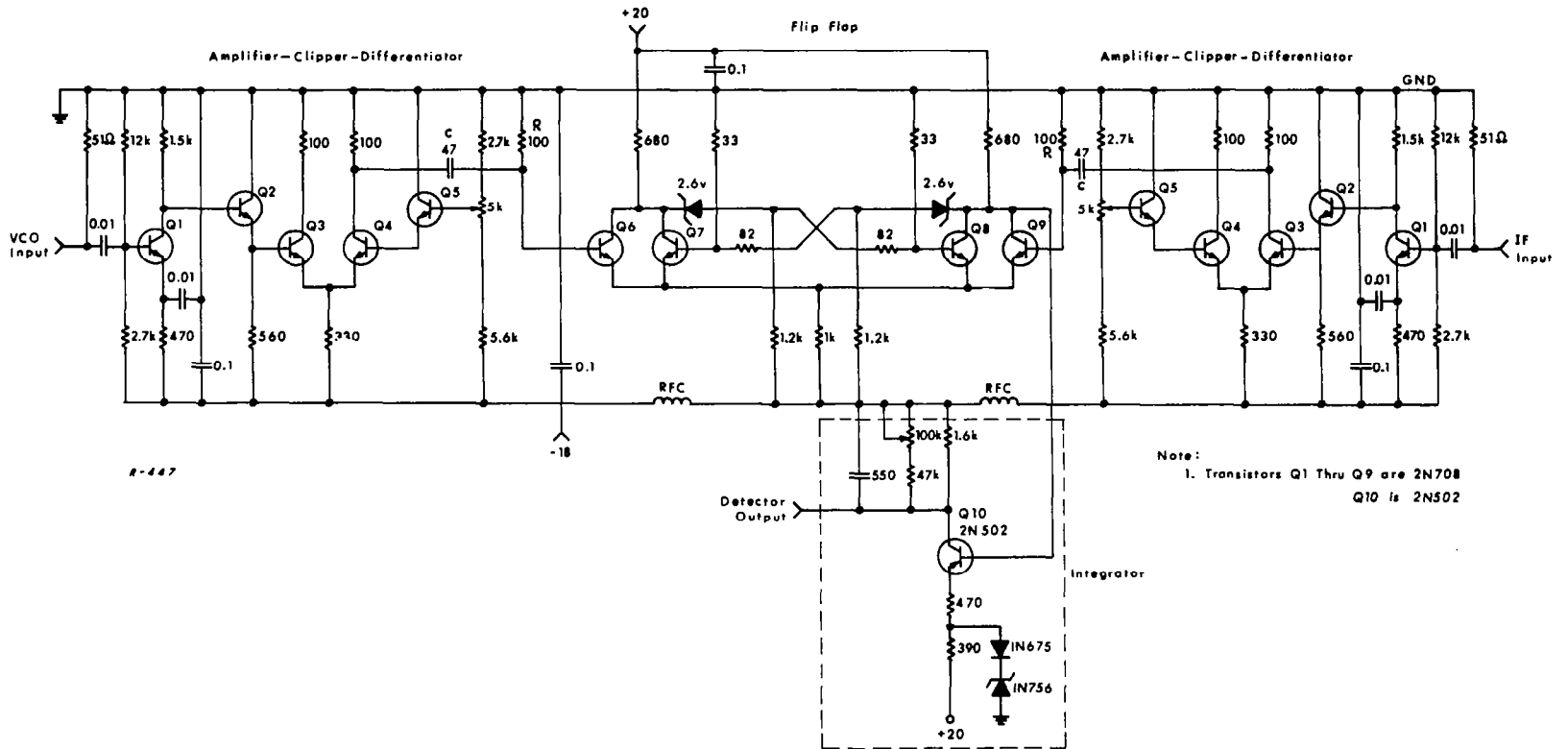
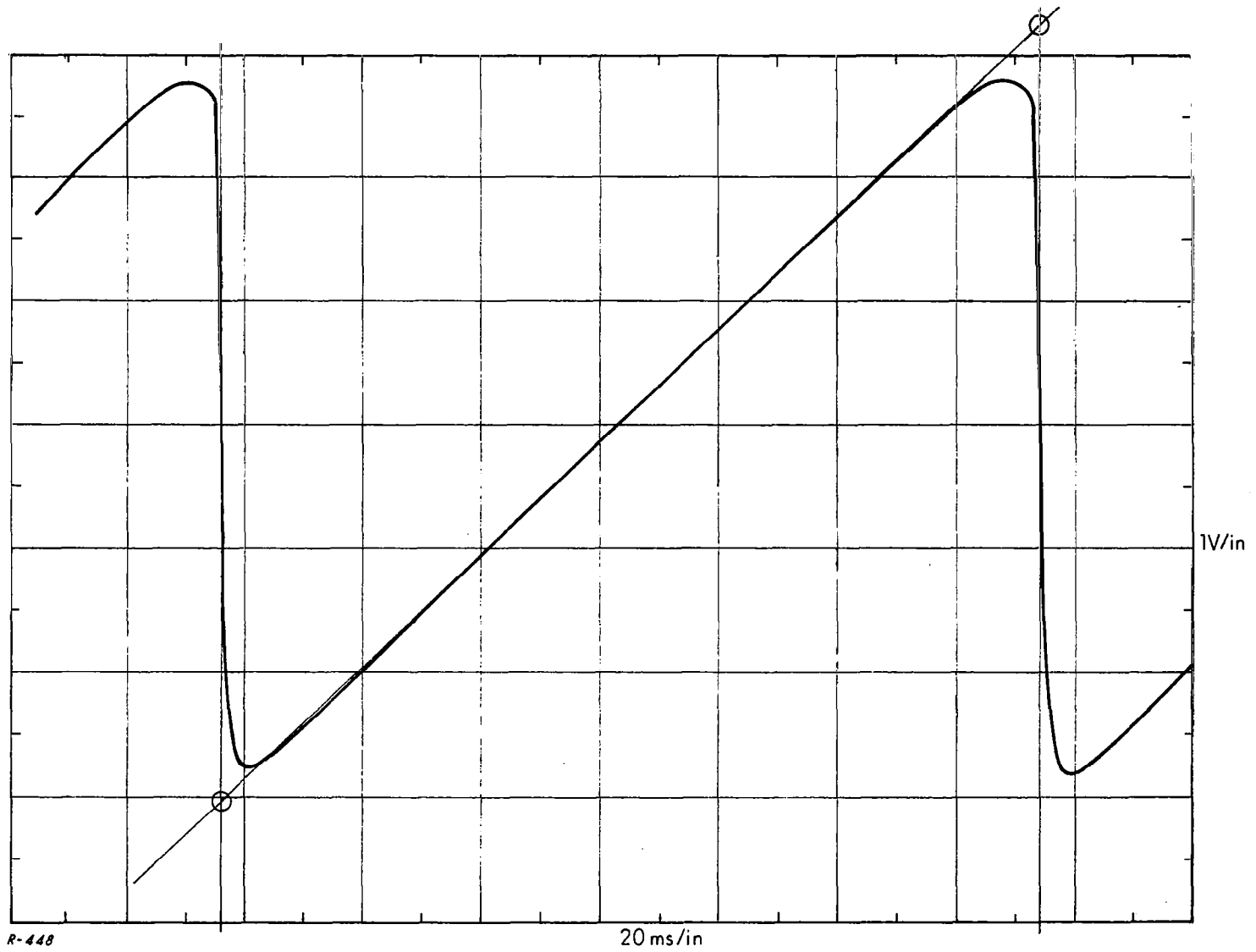


Fig. 30 Sawtooth phase detector circuitry.



R-448

Fig. 31 Sawtooth phase detector 20 ms/in characteristic.

threshold curves corresponding to different noise bandwidths and selecting the one giving the lower threshold. In the PLL₂, the result was a noise bandwidth equal to $1/\sqrt{2}$ times the optimum one found for the sinusoidal detector test which according to Eq. (3.25) implies a threshold error equal to two times that of the sinusoidal case. Even though this value seems reasonable, as well as interesting since the maximum tolerable error before a cycle slippage occurs is also a factor of $(\frac{\pi}{\pi/2}) = 2$, we cannot claim it as definite due to the fact that no threshold improvement (or degradation) over the sinusoidal detector results was present. This is illustrated in Figs. 32 and 33 which correspond to the (re-optimized) PLL₂ and PLL₃ results analogous to Figs. 26c and 27b for the sinusoidal detector. For instance, Eq. (3.24) predicts a theoretical improvement of $2^{5/2} = 7.5$ db in the PLL₂ case.

In an effort to discover the cause of the lack of threshold improvement, the sinusoidal and sawtooth phase detector outputs around the threshold region were studied (for the properly optimized $\delta = 30$ loops) by passing them first through a 200 kc lowpass filter and then observing their time fluctuations in a fraction of the sinusoidal (1 kc) distortion error period. The results are summarized in Figs. 34 and 35 where the amplitude scale is 0.5 v/cm for the sinusoidal detector pictures and 1 v/cm for the sawtooth detector pictures. The time scale is 25 μ sec/cm so that a whole picture represents 1/4 of the distortion error period. The sawtooth detector shows some vertical spikes which are absent in the sinusoidal detector case. The nature of these spikes and their effect on the threshold performance of the sawtooth phase detector loops are analyzed in the section that follows.

5.7 Cycle Slippage in the Sawtooth Phase Detector

At first sight, the behavior of the sawtooth phase detector during a PLL cycle slippage would seem to be a nonlinear problem. However, a linear formulation of the problem can be done by using a proper model of the detector operation and its output as a function of time can be determined analytically.

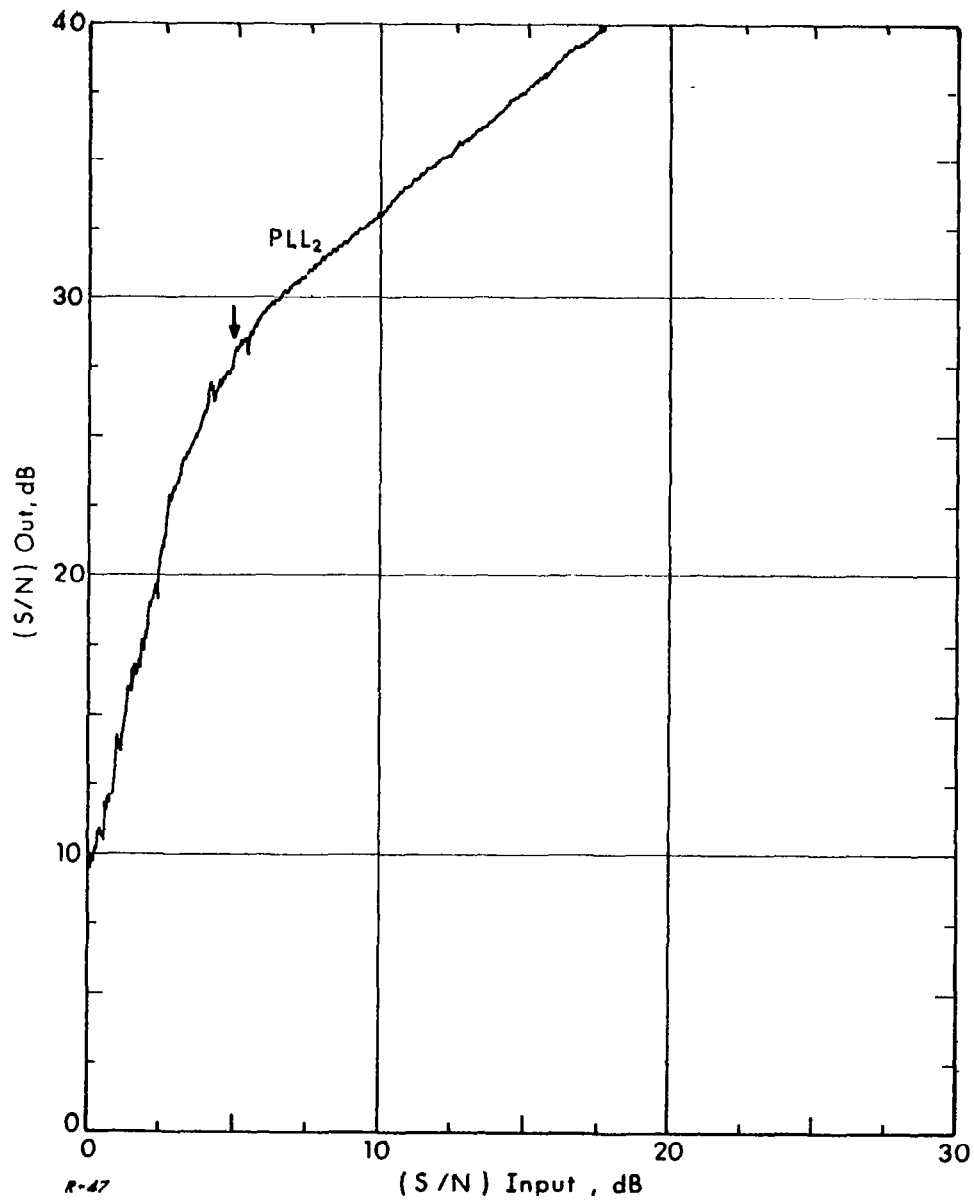


Fig. 32 Sawtooth detector results (PLL₂, $\delta = 30$).

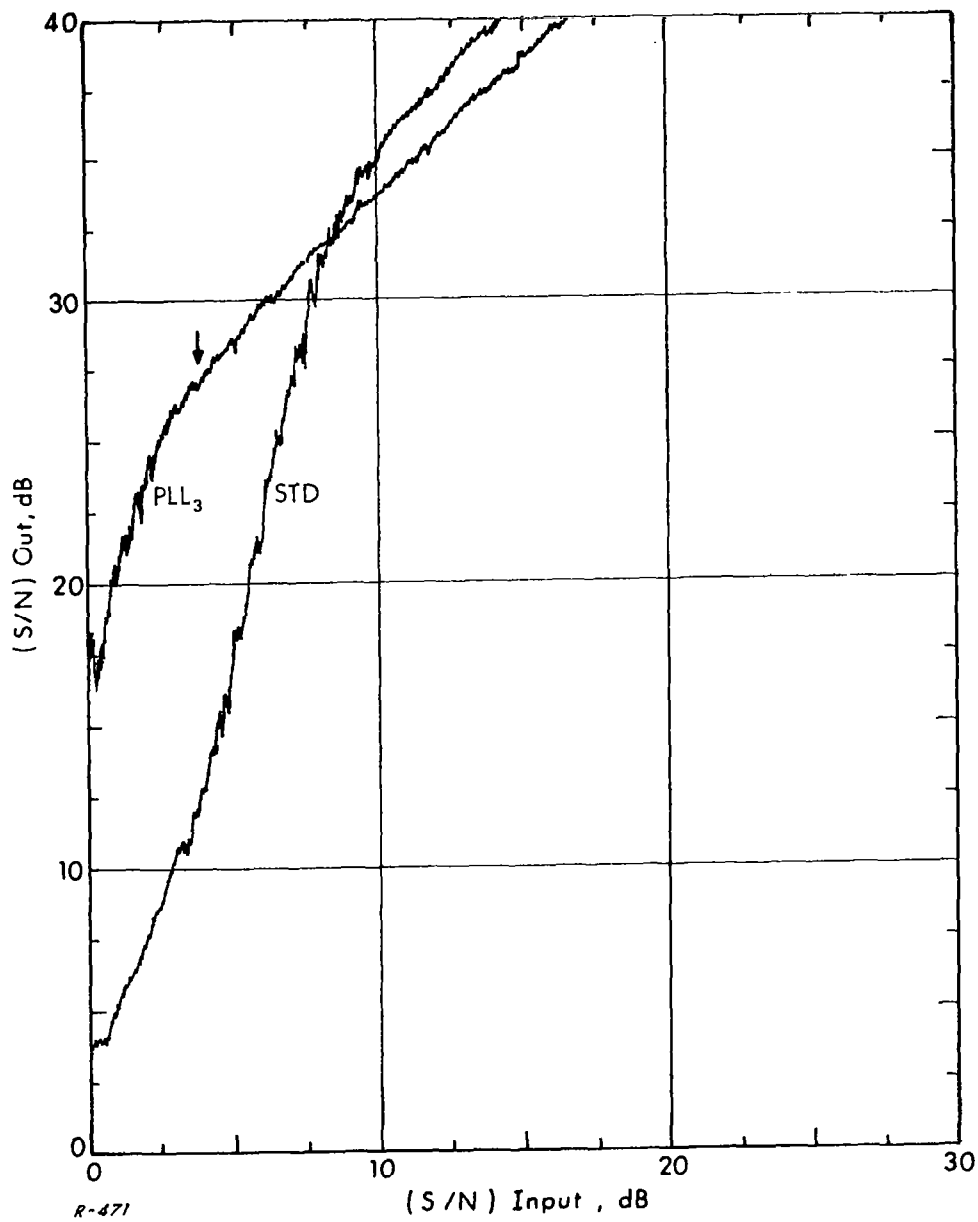
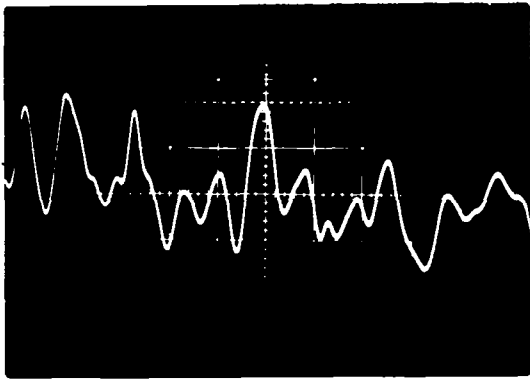
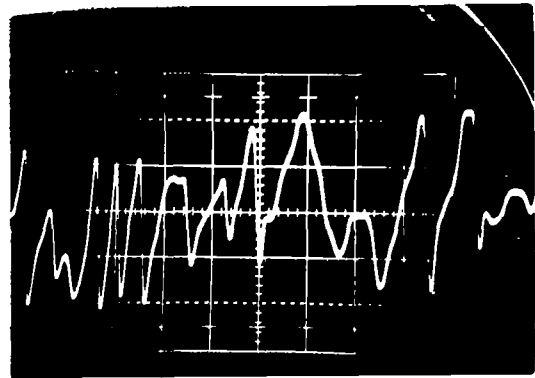


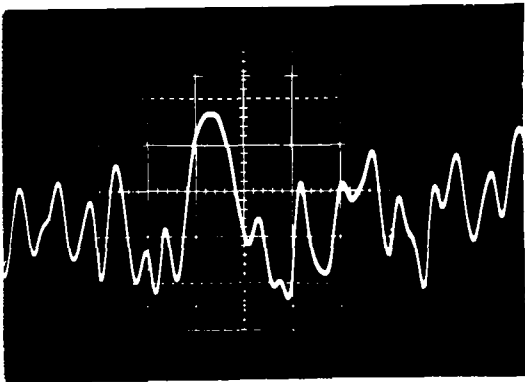
Fig. 33 Sawtooth detector results (PLL₃, $\delta = 30$).



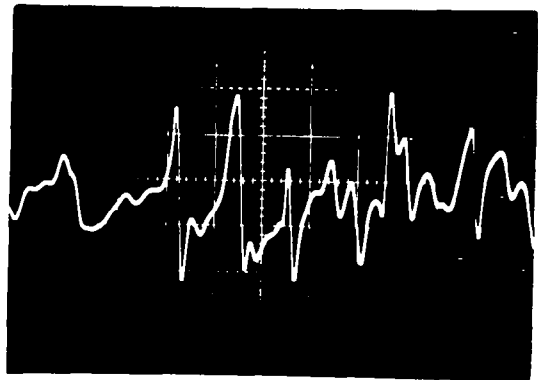
Sinusoidal detector: $(S/N)_{in} = 0$ db



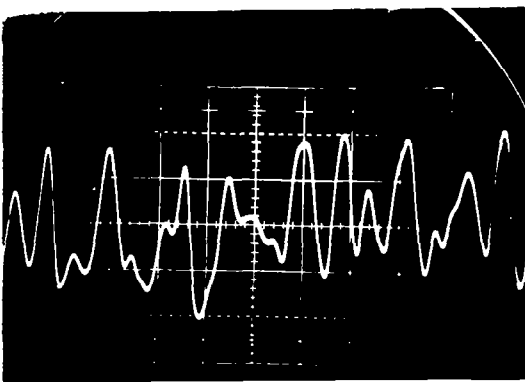
Sawtooth detector: $(S/N)_{in} = 0$ db



Sinusoidal detector: $(S/N)_{in} = -2$ db

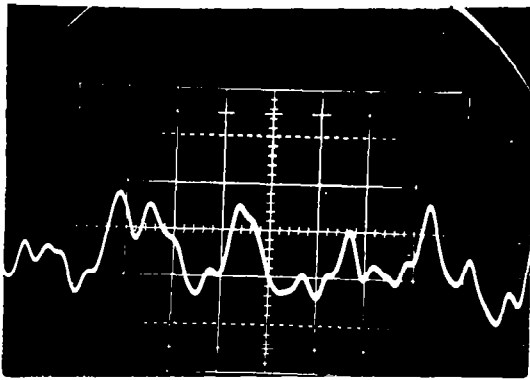


Sawtooth detector: $(S/N)_{in} = -10$ db

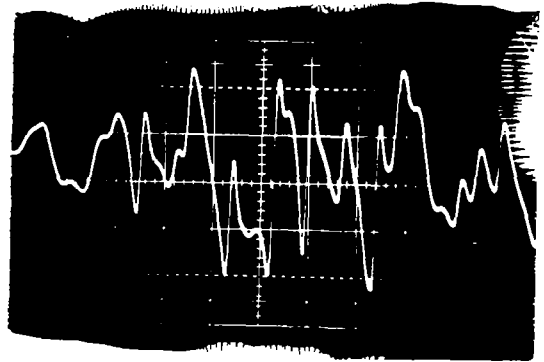


Sinusoidal detector: $(S/N)_{in} = -5$ db

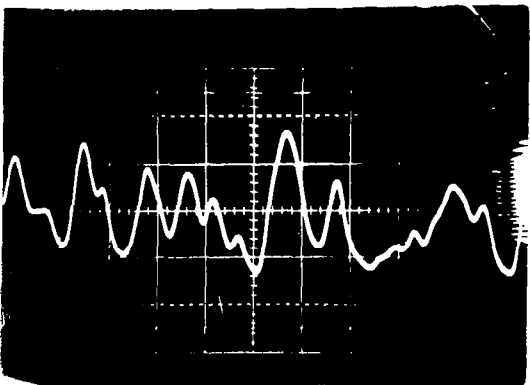
Fig. 34 PLL₂ oscillograms.



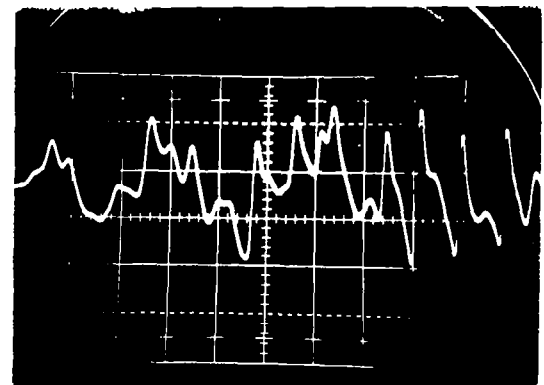
Sinusoidal detector: $(S/N)_{in} = 0$ db



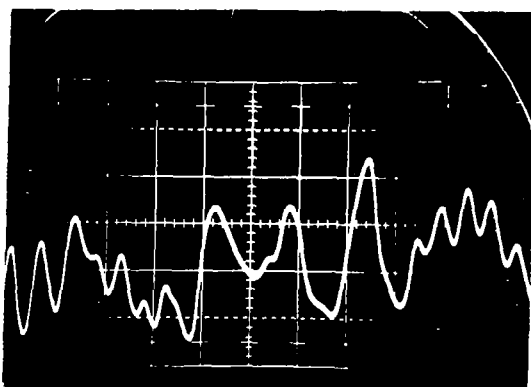
Sawtooth detector: $(S/N)_{in} = 0$ db



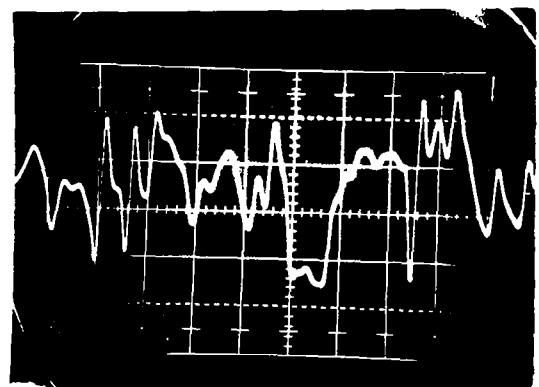
Sinusoidal detector: $(S/N)_{in} = -2$ db



Sawtooth detector: $(S/N)_{in} = -2$ db



Sinusoidal detector: $(S/N)_{in} = -5$ db



Sawtooth detector: $(S/N)_{in} = -5$ db

Fig. 35 PLL₃ oscillograms.

This approach is not applicable to the sinusoidal detector where a nonlinear analysis is truly required.

Consider the sawtooth detector characteristic shown in Fig. 4. The detector operation is described by a linear dependence between $e(\phi)$ and ϕ only as long as the loop operation is limited to $|\phi| < \pi$ rad. If a cycle slip-page occurs, the detector operation must now be characterized over the range $-\pi < \phi < 3\pi$ rad. This is accomplished by the model of Fig. 36 where a negative step is introduced at the time t^* when the (increasing) phase error reaches π rad. Note that aside from this step input function, the rest of the

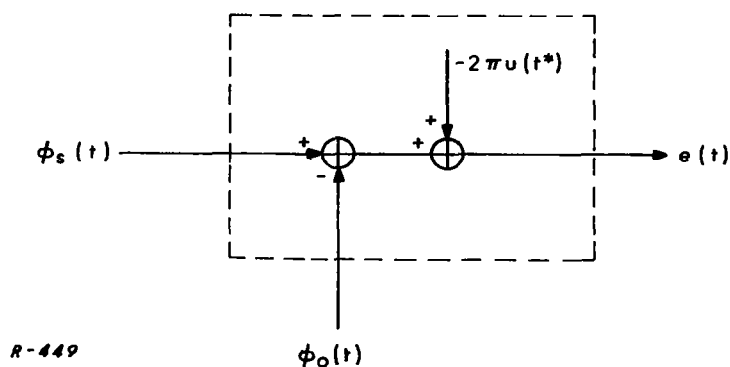


Fig. 36 Sawtooth detector model during cycle slippage.

model is identical to that existing when the loop operates in its locked condition. The fact that the system remains unchanged is extremely important; it allows us to compute the transient due to the step input and add it directly to the "steady-state" waveform which existed in the absence of cycle slippage.

For simplicity, we will consider the second order loop of Fig. 37 so that the system equations are summarized by

$$\begin{aligned} \phi_o(s) &= \frac{K}{s} v(s) = \frac{K}{s} \cdot \frac{1 + \tau_1 s}{1 + \tau_2 s} e(s) \\ &= \frac{K}{s} \cdot \frac{1 + \tau_1 s}{1 + \tau_2 s} \left[\phi_s(s) - \phi_o(s) - \frac{2\pi}{s} \right] \end{aligned} \quad (5.14)$$

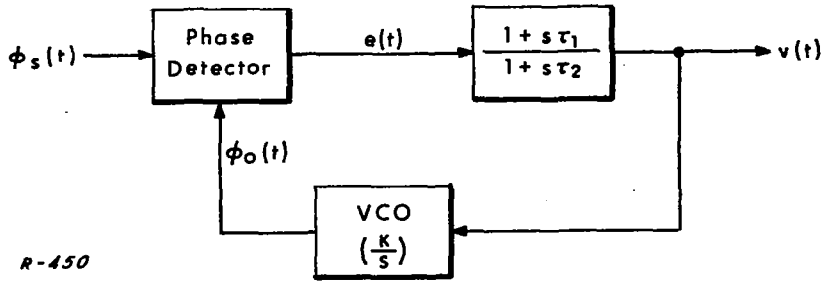


Fig. 37 Second order PLL model.

Here we have assumed for convenience that $t^* = 0$. This implies that our "steady-state" solution (possibly including noise) is such that $e(t)$ has increased to π at $t = 0$. Solving these equations for the transient term due to the 2π step we find

$$v(s) = -\frac{2\pi}{K} \frac{1 + \tau_1 s}{1 + (\tau_1 + \frac{1}{K})s + \frac{\tau_2}{K} s^2} \quad (5.15)$$

so that taking the inverse transform and using the optimum relations derived in Chapter III, ($\tau_2 = K\tau_1^2$, $K\tau_1 \gg 1$, $\tau_1 = 1/B_n$), we get

$$v(t) = -\frac{4\pi B_n}{\sqrt{3} K} e^{-\frac{B_n t}{2}} \cos\left(\frac{\sqrt{3} B_n t}{2} - \frac{\pi}{6}\right) \quad (5.16)$$

so that

$$\phi_o(t) = -\frac{4\pi}{\sqrt{3}} e^{-\frac{B_n t}{2}} \sin\left(\frac{\sqrt{3} B_n t}{2} - \frac{\pi}{3}\right) \quad (5.17)$$

where we have chosen the constant of integration such that $\phi_o(0^+) = 0$ and $\lim_{t \rightarrow \infty} \phi_o(t) = -2\pi$. Thus, we also have

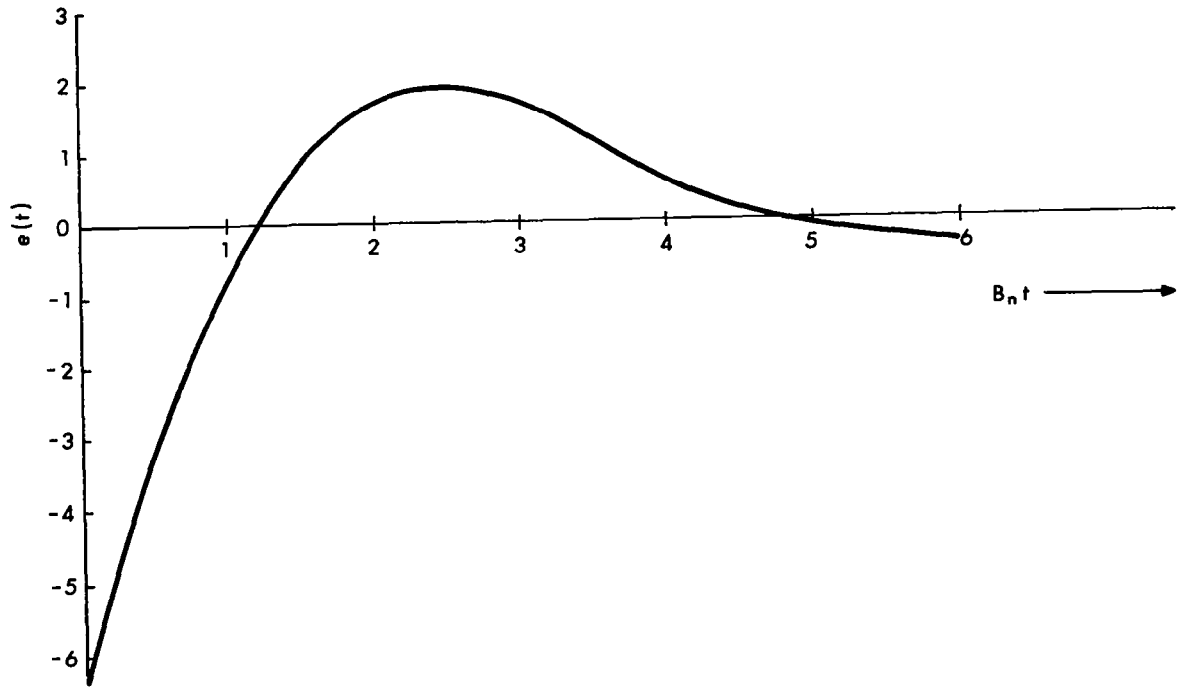
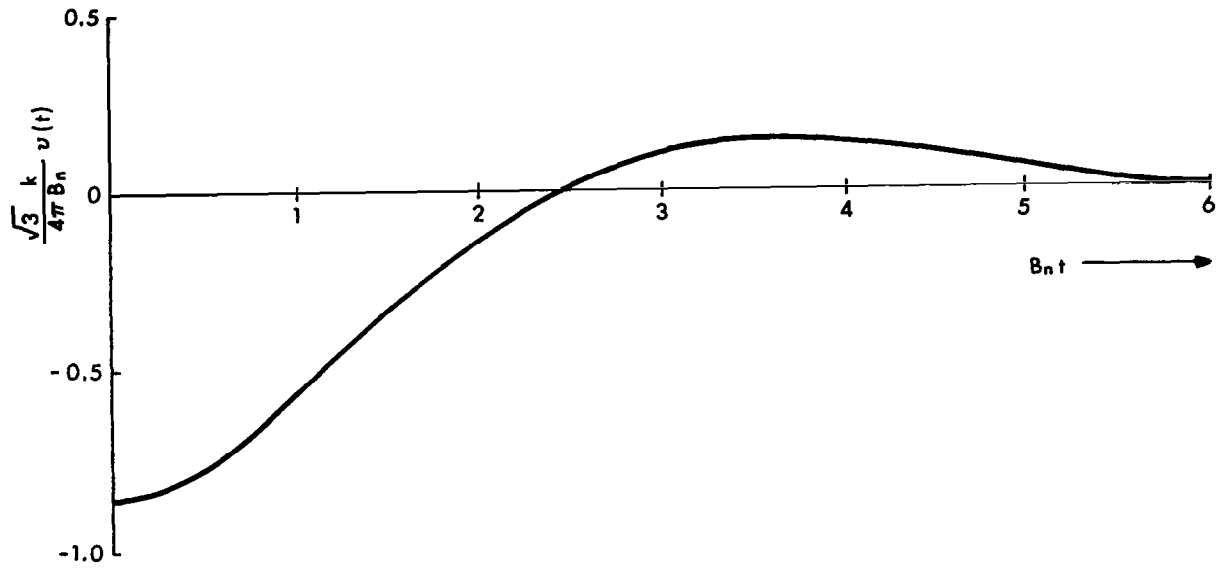
$$e(t) = \frac{4\pi}{\sqrt{3}} e^{-\frac{B_n t}{2}} \sin\left(\frac{\sqrt{3} B_n t}{2} - \frac{\pi}{3}\right) \quad (5.18)$$

The transient waveforms $v(t)$ and $e(t)$ are shown in Fig. 38. The PLL₂ sawtooth detector output during the cycle slippage will thus be characterized by an instantaneous 2π discontinuity (assuming a 1v/rad sensitivity) followed by a transient whose duration depends on the loop noise bandwidth. The experimental sensitivity is 0.7 v/rad so that the instantaneous step will correspond to about 4v according to Fig. 31. Also, the transient duration is small compared to the distortion error period according to the loop design for $\delta = 30$. Finally, we can claim the absence of the step discontinuity in the sinusoidal phase detector output due to its smooth characteristic.

This last statement can be verified experimentally by identifying the spikes in the sawtooth detector oscillograms of Figs. 34 and 35 as the step discontinuity previously derived, and noting their absence in the sinusoidal detector pictures. However, in the sawtooth pictures we not only discover the 4v steps but also some smaller ones which indicate that the cycle slippage somehow occurs before the phase error reaches π rad. It is evident that when this occurs the advantage of the sawtooth detector over the sinusoidal detector is no longer present and the threshold improvement disappears. The actual cause of the cycle slippage before the error reaches π rad is still being investigated.

5.8 Conclusion

In this chapter we have discussed the experimental results. The PLL₂ and PLL₃ experimental thresholds defined by a 0.5 db linearity departure were found to be compatible with theoretical threshold errors in the



R-451

Fig. 38 Waveforms during cycle slippage.

range $1/10 < \sigma^2 < 1/4 \text{ rad}^2$ (for the sinusoidal phase detector). The threshold improvement of the PLL₃ over the PLL₂ showed sufficient agreement with the theoretical curves so as to rely on the latter for establishing the loop filter complexity required for a given improvement.

The use of a sawtooth phase detector whose operation is based on digital techniques did not introduce any threshold improvement. An analytical study of the sawtooth detector output during a PLL cycle slippage allowed us to interpret the experimental results and to discover that the lack of improvement was due to the fact that the PLL cycle slippage took place before the phase error reached π rad thus eliminating the inherent advantages of the sawtooth detector relative to the sinusoidal one.

Chapter VI

THE FCF DEMODULATOR

6.1 Introduction

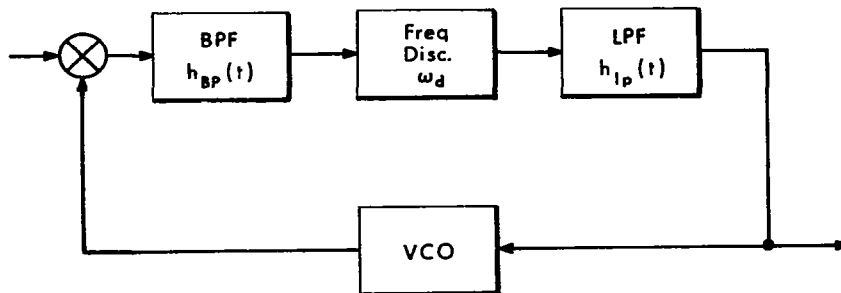
The maximum-likelihood demodulator presented in Chapter II immediately suggests the use of a PLL as its realizable approximation due to its similarity with the linear phase transfer model of the latter. The FCF demodulator can be similarly motivated from its lowpass analog model, though the existence of an open-loop i-f threshold implies a more complex analysis. In this chapter we will discuss the FCF demodulator threshold and compare its performance to the PLL demodulator.

6.2 The FCF Demodulator

The block diagram of an FCF demodulator is shown in Fig. 39 where the input and VCO output are similar to those existing in the PLL except for a difference between the carrier and free-running VCO frequencies. The mixer output is then proportional to

$$\left[1 + \frac{x_c(t)}{E_c} \right] \cos \left[\omega_d t + \phi(t) \right] - \frac{x_q(t)}{E_c} \sin \left[\omega_d t + \phi(t) \right] \quad (6.1)$$

where ω_d is the i-f frequency representing the quiescent operation of the discriminator. If the latter is assumed linear and amplitude - insensitive, the VCO output phase is given by



R-261

Fig. 39 Block diagram of FCF demodulator.

$$\phi_o(t) = K h_{lp}(t) \otimes \tan^{-1} \frac{A_q(t) \otimes h_{LP}(t)}{A_c(t) \otimes h_{LP}(t)} \quad (6.2)$$

where

$$A_c = \left(1 + \frac{x_c}{E_s}\right) \cos \phi - \frac{x_q}{E_s} \sin \phi \quad (6.3)$$

$$A_q = \left(1 + \frac{x_c}{E_s}\right) \sin \phi + \frac{x_q}{E_s} \cos \phi \quad (6.4)$$

and $h_{LP}(t)$ represents the impulse response of the lowpass analog of $h_{BP}(t)$, while K is again the open loop gain. The general nonlinear phase transfer model of the FCF is thus that shown in Fig. 40 and is noted to be much more complicated than the equivalent one for the PLL given by Fig. 2.

6.3 Threshold Performance and Design of FCF Demodulators

In the absence of noise Eq. (6.2) simplifies to

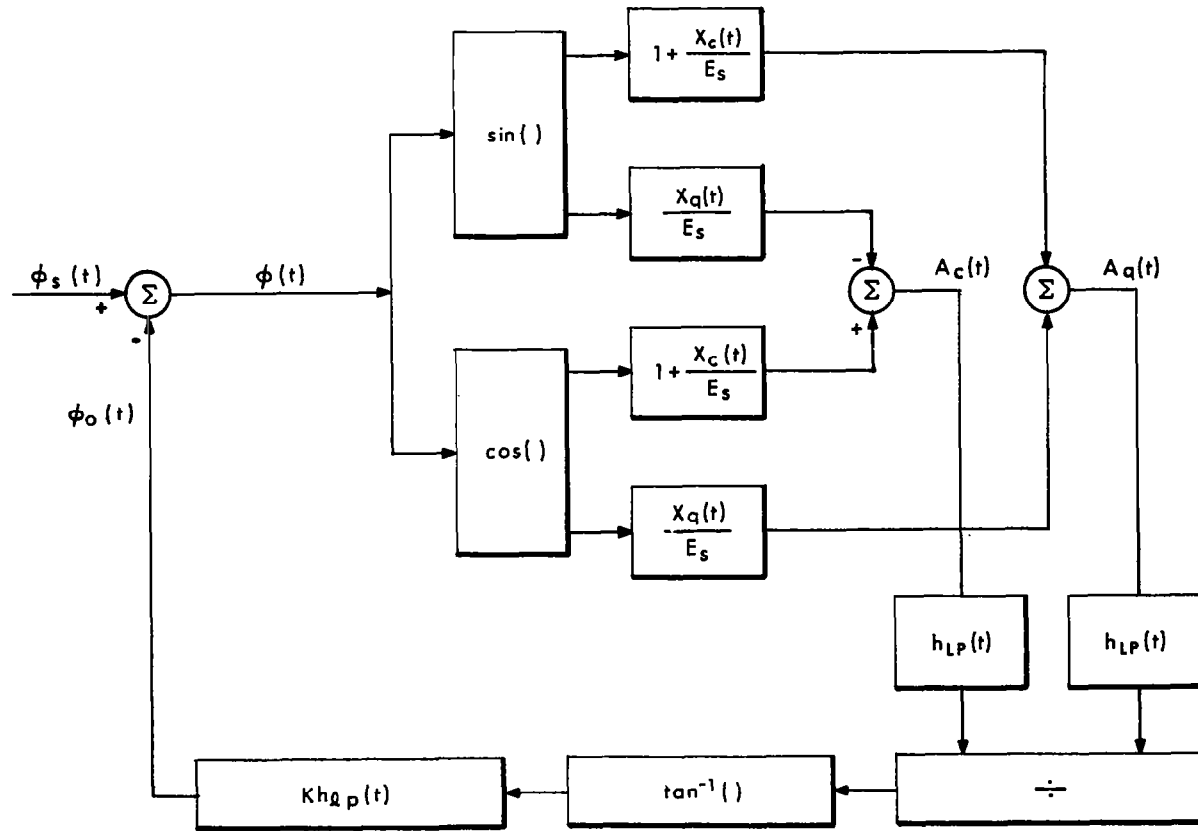
$$\phi_o(t) = K h_{lp}(t) \otimes \tan^{-1} \frac{h_{LP}(t) \otimes \sin \phi(t)}{h_{LP}(t) \otimes \cos \phi(t)} \quad (6.5)$$

and if we further assume small phase error conditions this expression may be approximated by

$$\phi_o(t) = K h_{lp}(t) \otimes \left[h_{LP}(t) \otimes \phi(t) \right] \quad (6.6)$$

which defines a linear (noiseless) model. The linearity is maintained in the presence of noise if $|\phi(t) + \theta(t)| \ll 1$, where $\theta(t)$ is the i-f phase noise which approximates $x_q(t)/E_s$ under high SNR conditions. Thus, the small error, high SNR linear model of the FCF demodulator is that shown in Fig. 41.

One of the difficulties in establishing an analytical threshold criterion for the FCF demodulator is the absence of a bound analogous to the critical



P-262

Fig. 40 General phase transfer model of an FCF demodulator.

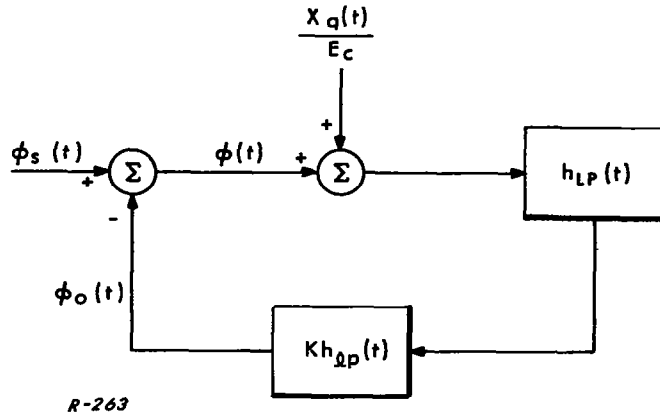


Fig. 41 Linear phase transfer model of the FCF demodulator.

peak error in the PLL phase detector and its corresponding cycle slippage whose frequency of occurrence can be identified to the reduction in output SNR. In the FCF, the analogous cycle slippage is obtained through the i-f encirclements of the origin in a phasor diagram which are detected by the discriminator. The FCF threshold is thus closely related to the i-f discriminator threshold and the common approach is to equate them by selecting an rms mean-square phase noise error and an i-f input SNR as their respective characteristic phenomena. While the i-f threshold SNR choice is more or less conventional, the critical noise error is not easily motivated analytically and its choice is usually based on empirical evidence.

Thus, if we assume a discriminator i-f threshold SNR of $\frac{S_1}{\Phi B_{if}} = \rho$ and a FCF threshold noise error of $\frac{\Phi B_n}{2S_1} = \sigma_n^2$, then the simultaneous occurrence of these two phenomena implies

$$B_n = 2\rho \sigma_n^2 B_{if} \text{ cps} \quad (6.5)$$

If we denote the closed loop phase transfer function of Fig. 41 by

$$H(s) = \frac{KH_{LP}(s) H_{lp}(s)}{s + KH_{LP}(s) H_{lp}(s)} \quad (6.6)$$

the phase noise bandwidth B_n is given by

$$B_n = \int_{-\infty}^{\infty} |H(j\omega)|^2 \frac{d\omega}{2\pi} \text{ cps} \quad (6.7)$$

where $H_{LP}(s)$ and $H_{lp}(s)$ are assumed to have unity d-c gain. In turn, the i-f bandwidth is selected according to a compressed message distortion criterion. For example, in the case of sinusoidal modulation of δ and frequency ω_m rad/sec, the i-f bandwidth is selected as that required to pass a sinusoidal FM signal of index $\delta|1-H(j\omega_m)|$ and frequency ω_m rad/sec in accordance to some distortion criterion.

Finally, the time constants of $H_{LP}(s)$ and $H_{lp}(s)$ are chosen so as to minimize the threshold

$$\frac{S}{\Phi} = \rho B_{if} = \frac{B_n}{2\sigma_n^2} \frac{\text{watts}}{\text{watts/cps}} \quad (6.8)$$

i. e., to minimize B_n (or B_{if}) while satisfying Eq. (6.5). An analytical optimization of $H(s)$ analogous to that effected for the PLL in Chapter IV is not straightforward due to the different formulation of the problem. The common approach is to select some simple transfer functions for $H_{LP}(s)$ and $H_{lp}(s)$ and then select the time constants in a logical way.

6.4 Comparison of FCF and PLL Demodulators

A good example of the aforesaid design principles is the work of Heitzman¹² where he considers several transfer functions for the case of sinusoidal modulation. We will in particular refer to his choice of

$$H_{LP}(s) H_{lp}(s) = \frac{1 + \tau_2 s}{(1 + \tau_1 s)(1 + \tau_3 s)} \quad (6.9)$$

which gives the best performance. Heitzman uses $\rho = 10$ and $\sigma_n^2 = (1/3.11)^2$ as his threshold criteria and develops an expression for the FCF threshold. However, his Eq. (25) is only valid for modulation indices (M in his notation) satisfying $\delta^{1/3} \gg 1$ which is somewhat limited. Be that as it may, the resultant FCF threshold is then given by

$$\frac{S_1}{\Phi} = 5 \delta^{1/3} \omega_m \quad (6.10)$$

Heitzman also includes a comparison with a second order PLL whose bandwidth is optimized for minimum error after setting $a = 2/3$ ($\xi = 1/\sqrt{2}$) arbitrarily and assuming $\sigma^2 = 1/4 \text{ rad}^2$ as the PLL threshold error. The optimum choice of $a = 1$ and the corresponding threshold of Eq. (3.24) indicated in Chapter III can be shown to yield a 1 db threshold improvement over Heitzman's design independent of the threshold error. Thus, the FCF threshold improvement over the second order PLL can be determined from Eqs. (3.24) and (6.10) to be

$$\frac{\text{PLL}_2}{\text{FCF}} = -8 - 25 \log \sigma + \frac{5}{3} \log \delta \text{ db} \quad (6.11)$$

This improvement is not only strongly dependent on the FCF threshold values assumed for ρ and σ_n^2 , but also on that assumed for the PLL threshold error σ^2 . For instance, Eq. (6.11) becomes

$$\frac{\text{PLL}_2}{\text{FCF}} = -0.5 + \frac{5}{3} \log \delta \quad \text{for } \sigma^2 = \frac{1}{4} \text{ rad}^2 \quad (6.12)$$

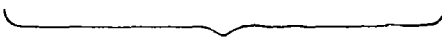

and

$$\frac{\text{PLL}_2}{\text{FCF}} = 4.5 + \frac{5}{3} \log \delta \quad \text{for } \sigma^2 = \frac{1}{10} \text{ rad}^2 \quad (6.13)$$

A comparison between the threshold improvement capabilities of Heitzman's FCF and the higher order PLL demodulators is summarized in the table that

follows where the entries represent the system improvement over the second order PLL. It is recalled that the experimental PLL threshold was found to be somewhere in this range of σ^2 values. On this basis, it seems that a more refined study becomes necessary in order to compare these systems. In particular, the more obscure point is perhaps the choice of the noise error threshold σ_n^2 in the FCF demodulator since it is not directly motivated and it does alter the FCF threshold entries by a considerable amount, a phenomenon analogous to the effect of varying the PLL threshold error σ^2 .

δ	<u>PLL</u> ₃	<u>PLL</u> ₄	<u>FCF</u>	<u>PLL</u> ₃	<u>PLL</u> ₄	<u>FCF</u>
10	1.8	2.5	1.2	1.5	2.2	6.2
20	1.9	2.7	1.7	1.8	3.0	6.7
30	2.0	3.0	2.0	2.0	4.1	7.0

 <p>PLL threshold at $\sigma^2 = 1/4 \text{ rad}^2$</p>	 <p>PLL threshold at $\sigma^2 = 1/10 \text{ rad}^2$</p>
---	---

6.5 Conclusion

In this chapter we have studied the FCF demodulator through phase transfer models analogous to those existing for the PLL but rather more complex. The FCF threshold is also difficult to characterize and the common approach is to equate the i-f discriminator threshold to the occurrence of a closed-loop mean-square phase noise error, which is somewhat arbitrary, and to minimize the (linear) loop noise bandwidth. The actual choice of

threshold errors strongly influences the theoretical threshold improvement capabilities of PLL and FCF demodulators and some empirical tests for the FCF similar to those discussed in Chapter V for the PLL seem desirable in order to establish the threshold noise error of this system.

Chapter VII

BAND-DIVIDING DEMODULATION

7.1 Introduction

The band-dividing system presented by Akima¹³ shows an analytical SNR improvement over conventional FM demodulators based on the use of an effective predetection bandwidth smaller than a conventional i-f bandwidth. The procedure is to divide the i-f by means of a filter bank, identify the presence of the signal in one filter and demodulate the individual filter output in a conventional way. If the correct filter is always selected, then the SNR (and threshold) improvement is given by B_{if}/B_f , the ratio of the i-f to the filter bandwidth. The presence of noise causes incorrect filter selection thus reducing the improvement capabilities this effect defined as mis-selection noise.

In the following pages, we first compare the Akima system to a PLL demodulator and then present some alternate band-dividing techniques that attempt to eliminate some of the system limitations. The most desirable ones were sufficiently complex to preclude their experimental investigation on this program.

7.2 Comparison with the Second Order PLL Demodulator

The ultimate SNR improvement and mis-selection immunity of the Akima system is limited by the minimum bandwidth required for the individual filters in order to reproduce the signal in a quasi-stationary way. The requirement in question is

$$\text{max. rate of frequency change} \times \text{filter response time} \ll \text{filter bandwidth} \quad (7.1)$$

For the case of sinusoidal modulation of index δ and frequency ω_m , the maximum rate of change of frequency is $\delta \omega_m^2$ rad/sec². If the filter response time is assumed to be of the order of the inverse of its bandwidth, then Eq. (7.1) reads

$$\frac{\delta \omega_m^2}{2\pi} \times \frac{1}{B_f} \ll B_f \quad (7.2)$$

or

$$B_f \gg \frac{1}{\sqrt{2\pi}} \delta^{1/2} \omega_m \quad (7.3)$$

mis-selection which imposes a lower bound on performance since noise has yet to be considered.

The result of Eq. (7.3) shows that the individual filter bandwidth is of the same order of magnitude of the second-order PLL bandwidth when the latter system is optimized for this type of modulation. The presence of mis-selection noise will strongly degrade the Akima system, particularly at threshold, and this seems to be sufficient evidence to direct our attention to other solutions at this stage.

It should be noted that these results are applicable to any band-dividing scheme based on a filter bank followed by a conventional FM demodulator, i. e., where the SNR improvement is given by B_{if}/B_f . The actual amount of filter overlap does not change this result but only alters the mis-selection noise performance, i. e., the extent to which the ideal improvement of B_{if}/B_f is approximated.

7.3 Effect of Mis-Selection Noise in Systems with Filter Overlap

In any band-dividing technique which selects the filter with maximum output, and regardless of whether or not a more refined estimate is subsequently made, noise power is produced due to filter mis-selection. Akima¹³ calculates this noise power for a system in which the band-dividing channels are orthogonal. He assumes that the correct channel is selected with probability q and each of the remaining $N-1$ channels have probability p of being mistakenly selected. The result of his calculation is that the mis-selection noise power is given by

$$P_{n2} = \frac{Np(N^2-1)}{12} \quad (7.4)$$

As a first attempt, we will modify the problem by assuming that the two channels adjacent to the correct channel have higher probability of being mis-selected than more remote channels. This is a good representation of the situation where the frequency responses of adjacent filters are overlapping to some extent. Thus, we let

q_c = probability of selecting the correct channel
 q_a = probability of selecting either adjacent channel
 p' = probability of selecting either one of the other channels.

With these assumptions we will calculate a new value of noise power P_{n2}' and compare this with Akima's result. First we will compute the expected value of the channel number actually chosen (\bar{i}), assuming that the signal actually occurs in the j^{th} channel.

$$\bar{i} = [1 + 2 + \dots + N - (j-1) - j - (j+1)] p' + (j-1) q_a + j q_c + (j+1) q_a \quad (7.5)$$

or

$$\bar{i} = \frac{N(N+1)}{2} p' - 3 j p' + j (q_c + 2q_a) \quad (7.6)$$

Now we use the fact that

$$(N-3) p' + q_c + 2q_a = 1 \quad (7.7)$$

to eliminate the quantity $q_c + 2q_a$ with the result

$$\bar{i} - \frac{N+1}{2} = (j - \frac{N+1}{2}) (1 - Np'). \quad (7.8)$$

This is the same result as Akima's, except that we have replaced p with p' , and some consideration can show that this could have been predicted from the symmetry of our probability assignment.

Next, we note that, as in Akima's work,

$$\begin{aligned}
 P_{n2}' &= \overline{(i-\bar{i})^2} = \overline{i^2} - \bar{i}^2 \\
 &= [1^2 + 2^2 + \dots + N^2 - (j-1)^2 - j^2 - (j+1)^2] p' + (j-1)^2 q_a \\
 &\quad + j^2 q_c + (j+1)^2 q_a - \bar{i}^2. \quad (7.9)
 \end{aligned}$$

Expanding and collecting with the use of Eq. (7.7) and the relationship

$$1^2 + 2^2 + \dots + N^2 = \frac{N}{6} (N+1) (2N+1), \quad (7.10)$$

we have

$$P_{n2}' = \frac{N}{6} (N+1) (2N+1) p' + j^2 (1 - Np') + \bar{i}^2 + 2 (q_a - p). \quad (7.11)$$

Now we restrict ourselves, as Akima did, to the noise power in the absence of modulation. Hence,

$$j = \frac{N+1}{2} \quad (7.12)$$

This in turn implies, from Eq. (7.8) that

$$\bar{i} = \frac{N+1}{2} \quad (7.13)$$

Substituting these two relations in Eq. (7.11) we find

$$P_{n2}' = \frac{Np'(N^2-1)}{12} + 2 (q_a - p') \quad (7.14)$$

This is the desired result and for purposes of comparison we form the ratio of P_{n2}' and P_{n2} from Eq. (7.19):

$$\frac{P_{n2}'}{P_{n2}} = \frac{p'}{p} \left[1 - \frac{24}{N(N^2-1)} \right] + \frac{24}{N(N^2-1)} \frac{q_a}{p} \quad (7.15)$$

When this ratio is greater than one, performance of the system with overlap is inferior to performance of the system without, when the ratio is less than one the reverse is true. We note that with Eqs. (7.7) and (7.15) define four variables, p'/p , q_a/p , q_c/p , and P_{n2}'/P_{n2} , any two of which may be considered independent. In Fig. 42 we treat p'/p and q_a/p as the independent variables and only the contour $P_{n2}'/P_{n2} = 1$ is drawn so that curves for various values of N may be included. Above this contour (for a given N) performance of the system with overlap is inferior to the system without; below the contour it is superior.

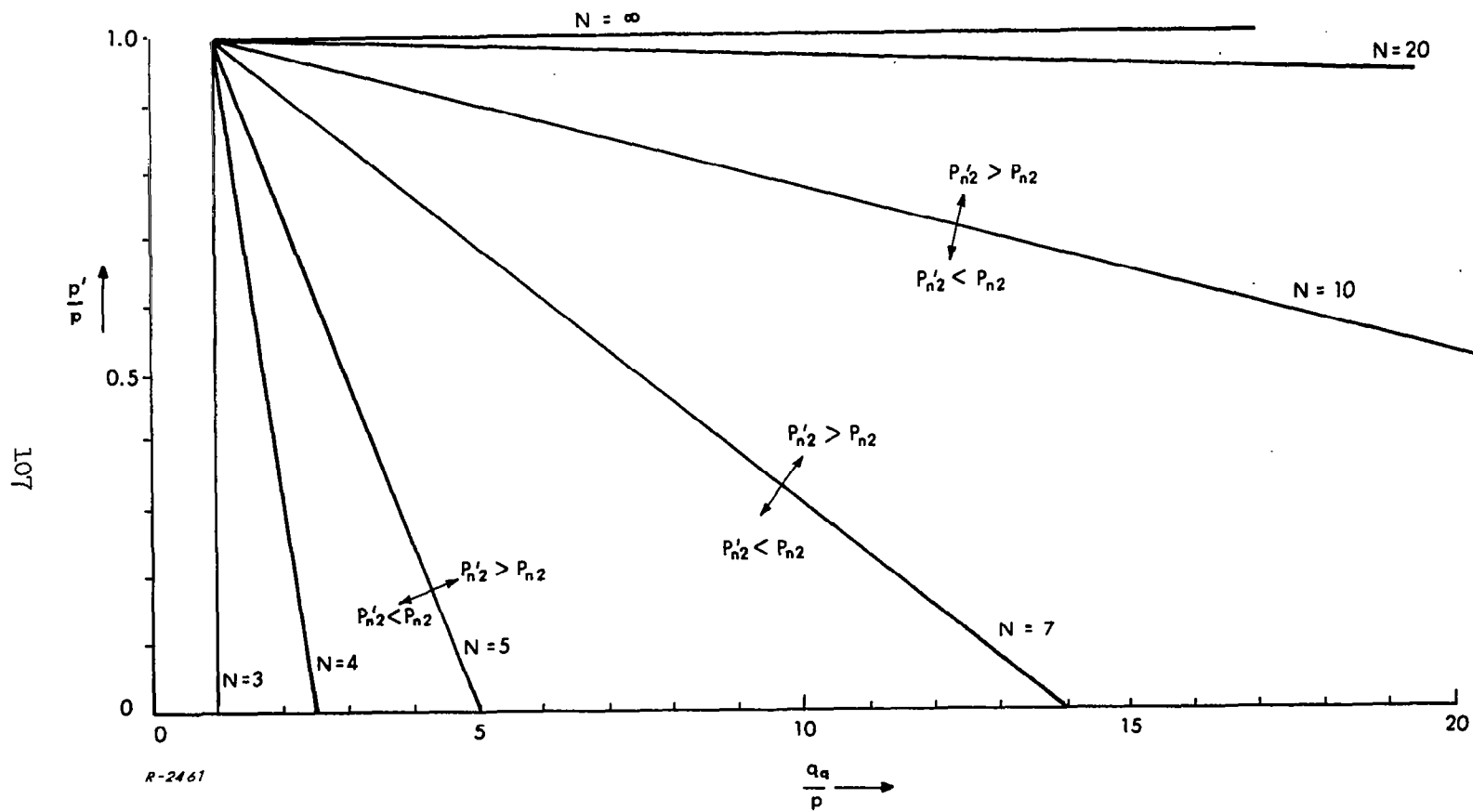


Fig. 42 Contours of $\frac{P'_{n2}}{P_{n2}} = 1$, for various N.

7.4 PLL Band-Dividing

Another technique based on band-dividing consists of having each i-f band division to represent the double-sided pull-in range* of a PLL whose free running VCO is at the center of the division in question. If the i-f divisions (pull-in ranges) are orthogonal and span the whole i-f band, then a signal frequency that remains long enough within the pull-in range of a loop will be captured only by that loop. On this basis, the i-f bandwidth equals the product of the pull-in range of a loop times the number of loops required to cover the whole i-f. The overall system is completed by the inclusion of a decision circuit to select the proper loop and a signal extractor to reproduce the modulation. The interest in this system lies in the possibility of providing an i-f division width (pull-in range) significantly larger than the noise bandwidth of the filter (PLL). In order to compare this technique with other methods, the minimum noise bandwidth required for each loop must be evaluated.

In analogy with the passive filter bank, a necessary condition for the signal reproduction is that the signal frequency variation during the acquisition time remains within the pull-in range of the loop in question. In general, one can distinguish between two acquisition modes for a signal with arbitrary phase dynamics. The first case consists of the semi-instantaneous acquisition of the signal so that the difference between the signal frequency at $t = 0$ and at $t = t_{acq}$ is very small relative to the pull-in range, as shown in Fig. 43. This case may be treated analytically as the acquisition of a constant frequency step

$$\Omega \approx \omega_0 - \omega_1 \approx \omega_0 - \omega_2. \quad (7.16)$$

Notice that even in the case where

$$|\omega_1 - \omega_0| < \Omega_{\text{pull-in}}$$

and

$$|\omega_2 - \omega_0| > \Omega_{\text{pull-in}}$$

*The pull-in range of a PLL is defined as the maximum constant frequency step, referred to the free running VCO, that can be required.

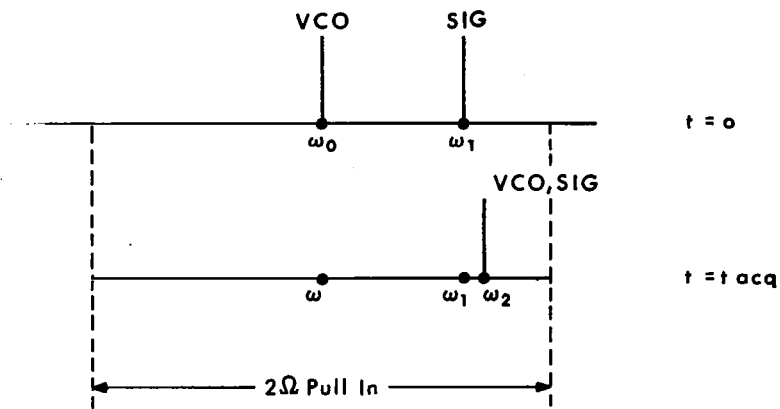
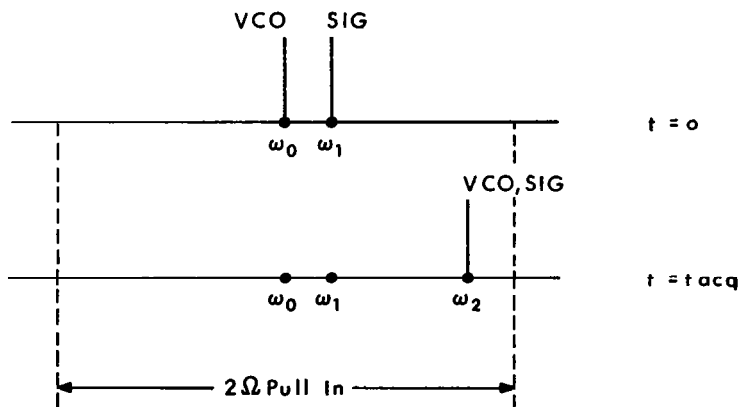


Fig. 43 $\omega_2 - \omega_1 \ll \Omega_{\text{pull-in}}$.



R-2462

Fig. 44 $\omega_2 - \omega_1 \sim \Omega_{\text{pull-in}}$.

the analysis makes sense since it is just a question of whether a given loop or its adjacent one acquires the signal. The second case consists of the alternate possibility where the difference between the signal frequency at $t = 0$ and t_{acq} is comparable to the pull-in range, as shown in Fig. 44. The analysis in this case must include formulation of the actual phase fluctuations and it becomes very complicated due to the nonlinearities involved in the acquisition process. Still, one can note an undesirable feature in this case. Consider the possibility of

$$|\omega_1 - \omega_0| < \Omega_{\text{pull-in}}$$

and

$$|\omega_2 - \omega_0| > \Omega_{\text{pull-in}}$$

which should happen rather frequently for the case in question and is shown in Fig. 45. It is noted that even if the adjacent loop were to acquire, which is not certain, the signal is lost during the transition time to the adjacent loop.

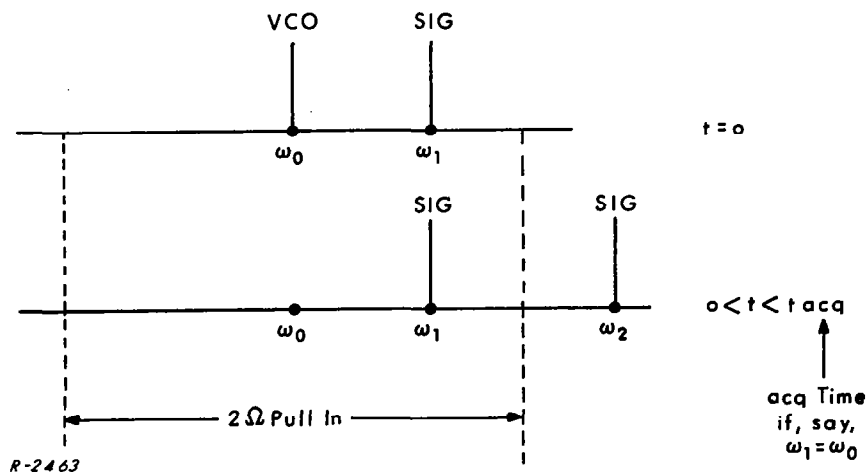


Fig. 45 $\omega_2 - \omega_1 \sim \Omega_{\text{pull-in}}$, $|\omega_2 - \omega_0| > \Omega_{\text{pull-in}}$.

This phenomenon was not relevant in the first case because these transitions do not occur frequently since acquisition takes place before the signal has moved away, and whenever they occur, the transition time is of the order of the acquisition time. It is evident that the second acquisition mode seems undesirable from reliability considerations and that the first mode is thus desired. The acquisition requirement for sinusoidal modulation is then

$$\begin{aligned} \text{max rate of} & \times \text{acquisition} &<< \text{pull-in} \\ \text{freq change} & \text{time} & \text{range} \\ (\delta\omega_m^2) & (t_{\text{acq}}) & (\Omega_{\text{pull-in}}) \end{aligned} \quad (7.17)$$

The analysis that follows assumes a conventional second order loop having the set (K, B_n, a) introduced in Chapter III as design parameters. The acquisition time of a frequency step by such a system may be approximated by

$$\Omega^2 / a(2-a)^2 B_n^3, \quad (7.18)$$

where Ω is the initial frequency shift in rad/sec, as long as Ω is sufficiently smaller than the pull-in range. A more detailed plot of the acquisition time is shown in Fig. 46 and is based on the analytical work of Richman.¹⁴ As Ω gets closer to the pull-in range, the actual acquisition time becomes increasingly larger than that given by the previous expression and finally $t_{\text{acq}} = \infty$ for $\Omega = \Omega_{\text{pull-in}}$. Therefore, Eq. (7.17) becomes

$$\delta\omega_m^2 \times \frac{\Omega^2}{a(2-a)^2 B_n^3} \ll \Omega_{\text{pull-in}}^{\text{suff}} \text{ for } \Omega < \Omega_{\text{pull-in}}. \quad (7.19)$$

The worst case for acquisition occurs at $\Omega = \Omega_{\text{pull-in}}$. Even though the previous formulation is not valid at this value, it is still of interest to use $\Omega = \Omega_{\text{pull-in}}$ because if an undesirable performance results, as will be the case, one can then claim that the actual performance is even worse and conclude that the system is inferior to other techniques.

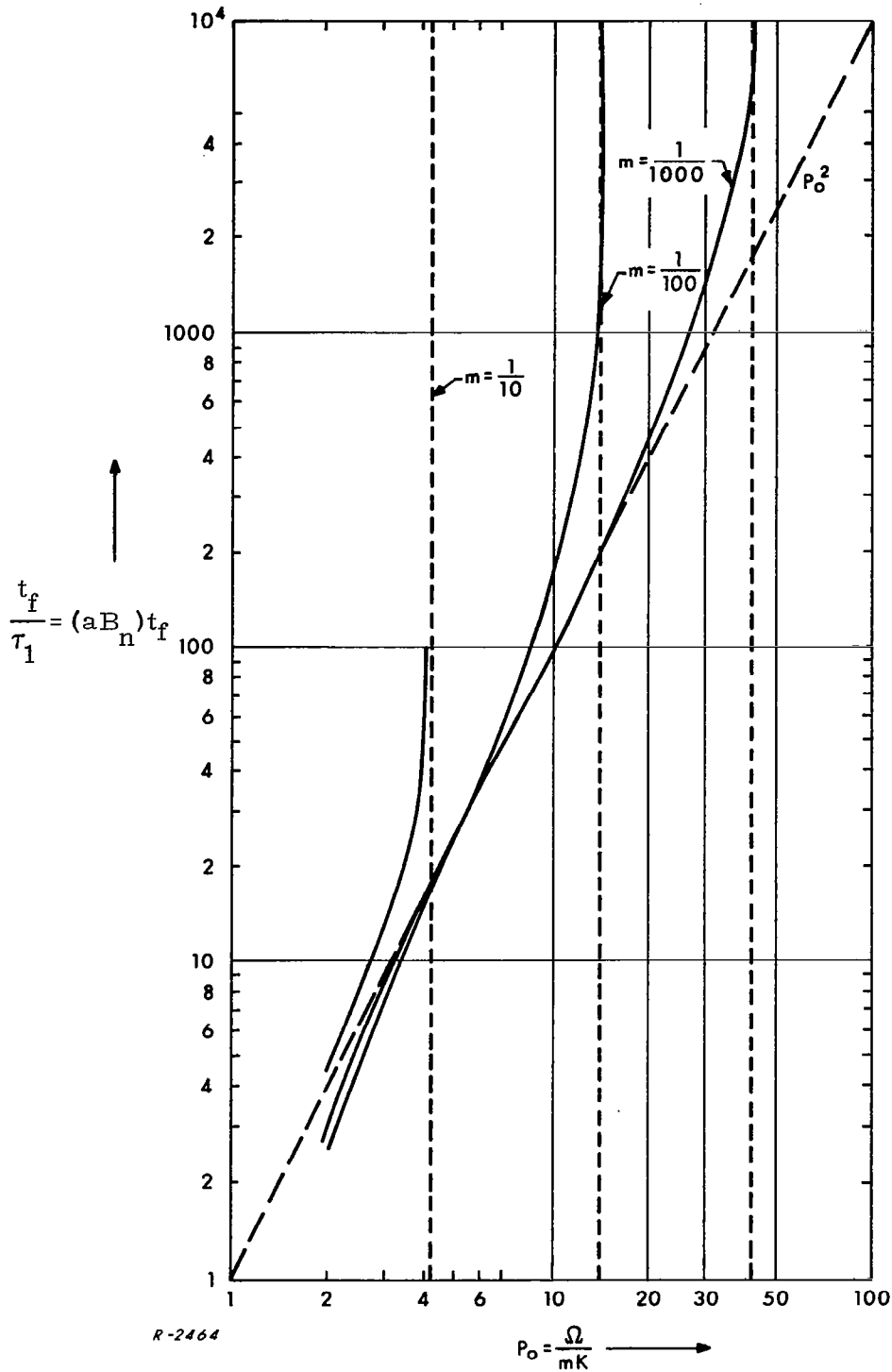


Fig. 46 Frequency pull-in time of a second order phase-locked loop. The dotted lines characterize the pull-in ranges for the different

$$m = \tau_1 / \tau_2 = \frac{(2-a)B_n}{K}$$

The pull-in range of the second order loop in question is given by *

$$\Omega_{\text{pull-in}} \approx \sqrt{2(2-a) B_n K} \quad (7.20)$$

so that the acquisition requirement reads

$$B_n^{5/2} \gg \frac{\sqrt{2K}}{a(2-a)^{3/2}} \delta \omega_m^2 \quad (7.21)$$

Using a factor of 10 to account for the much greater sign and setting $a = 1$ which gives a near-optimum behavior, then Eq. (7.21) reduces to

$$B_n > 1.2 \left(\frac{K}{B_n} \right)^{1/4} \sqrt{\delta} \omega_m \quad \text{cps.} \quad (7.22)$$

A comparison with the optimized second order PLL demodulator shows that each PLL in the bank requires a noise bandwidth equal to or larger than that of a single PLL processing the whole i-f signal since $K \gg B_n$. Notice that the K parameter for the PLL bank is established from the steady-state error $\phi = \Omega/K$ so that using $\phi < \frac{1}{2}$ radian (certainly an upper bound), for instance,

$$K > \frac{\Omega}{\phi} = \frac{1}{\phi} \sqrt{2B_n K} \quad \text{for } \Omega = \Omega_{\text{pull-in}}, \quad a = 1 \quad (7.23)$$

or

$$\left(\frac{K}{B_n} \right)^{1/4} > \frac{1.2}{\sqrt{\phi}} > 1.7 \quad (7.24)$$

so that

$$B_n > 2 \sqrt{\delta} \omega_m \quad \text{cps} \quad (7.25)$$

which is comparable to the noise bandwidth of the single PLL demodulator.

This result seems to be sufficient evidence to discard the conventional PLL band-dividing scheme, particularly if the aforesaid correction for the acquisition time is accounted for. This statement must be revised if faster acquisition capabilities are included, such as VCO sweeps (which is probably undesired

*The derivation of this result is rather complex. On this basis, it is better to give the reference than to try to summarize the analysis. The detailed derivation appears in Ref. 15.

because of the on-off synchronization complexity and noise performance) or non-sinusoidal phase detectors (which could decrease the acquisition time by a considerable amount and thus be desirable).

7.5 Band-Dividing Signal Estimation Techniques

The results of the previous sections show that the individual filter bandwidth constraints derived from signal acquisition considerations limit the performance capabilities of systems whose SNR improvement is based on the ratio of the i-f and filter bandwidths. These conclusions suggest that we direct our approach towards techniques whose SNR improvement is not strongly dependent on B_{if}/B_f at a first glance. As a simple example, consider the case where the output signal is a quantized waveform obtained by selecting the filter with the signal from the i-f bank and developing a quantum level in accordance to the filter selected. It is evident that such a technique would require a large amount of filter overlap from resolution considerations and a more promising method would consist of estimating the signal frequency from the output of more than just a single filter.

The problem at hand is to estimate the signal frequency when the outputs of the set of band-dividing filters are available. A maximum likelihood estimator is derived in Appendix A for the case where the filter outputs are processed at the same time instant. The block diagram of the resultant system is shown in Fig. 47 where $2N + 1$ filters are assumed. A generalized filter output $\omega(t, \nu)$ is both a function of time and frequency, the latter distinguishing the filter from its neighbors. For the no-memory processing system under consideration, the time dependency is deleted in Fig. 47 and ν is taken as kF , $k = 0, \pm 1, \dots, \pm N$, for equally spaced filters. The linear transformation T yields statistical independent noise components in different filters (unnecessary if the filter impulse responses are orthogonal) and the maximum likelihood estimator can then be formulated using conventional techniques. The estimator operation is defined by Eq. (31) in Appendix A and consists of a reference signal generator $S_k^*(\zeta)$ (where ζ is the

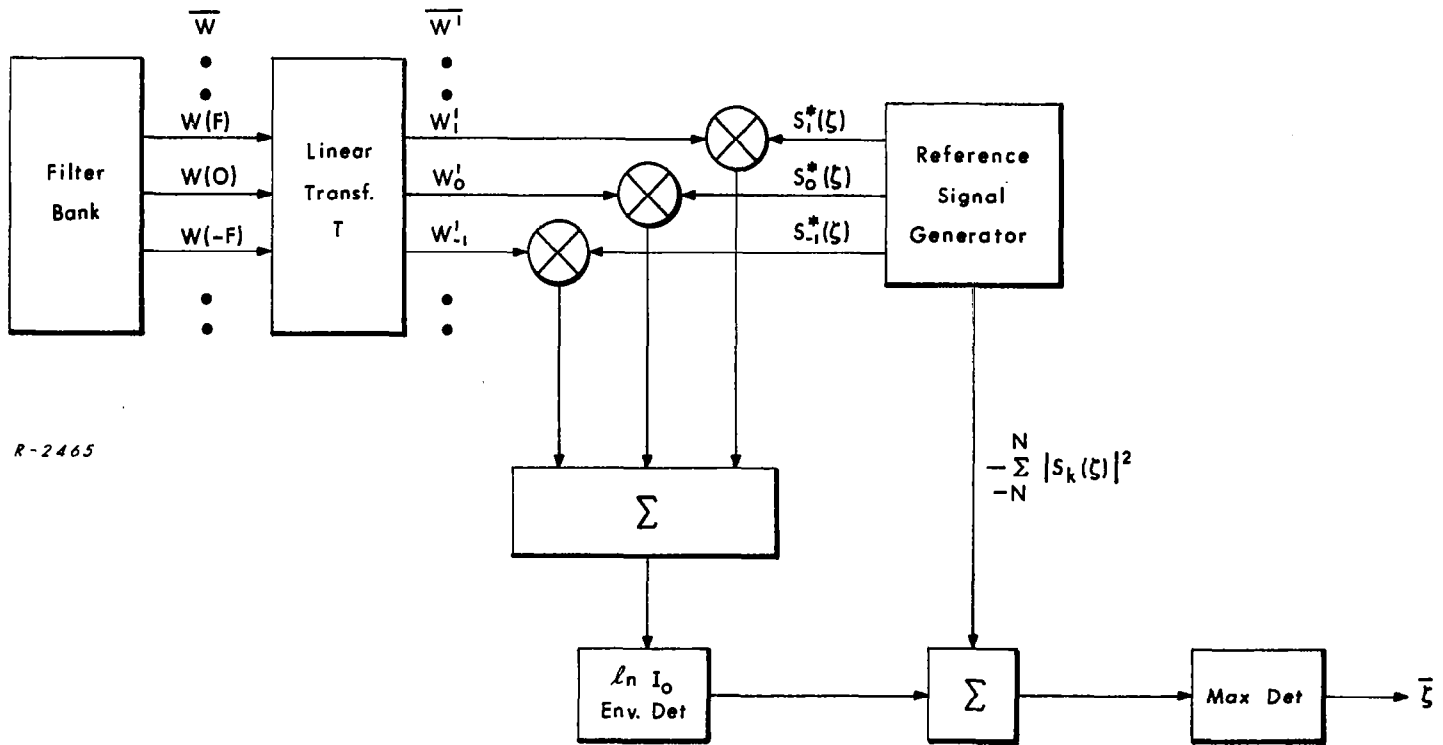


Fig. 47 Block diagram of maximum likelihood estimator (case of no-memory processing).

signal frequency), a product bank followed by an adder and a nonlinear envelope detector, and a comparator that yields the maximum likelihood estimate $\bar{\zeta}$.

A simpler interpolation scheme from synthesis considerations consists of examining the set $|\omega(t, kF)|$ of envelope detected outputs and forming a smooth curve that approximates $|\omega(t, \nu)|$. The value of ν at which this curve has its maximum would be the estimate of the signal frequency ζ . For such a procedure to be effective, the frequency separation F between the center frequencies of adjacent filters should be small enough for the filter transfer function to be essentially constant for changes in ν of this order of magnitude. In addition, of course, the SNR must be sufficiently high.

As an example, consider a quadratic interpolation scheme which involves passing a third degree polynomial through three adjacent points. There is no need to do this for all triplets of points but only for the triplet which surrounds the desired maximum. Thus one may first determine the filter whose envelope is maximum and then select as the desired triplet the outputs of this filter and the adjacent filters. In the event that the maximum envelope output occurs at one of the two end filters, the two adjacent interior filters should be used. Our discussion ignores this end point situation which is readily handled by slight modifications of the procedure discussed below.

Since the interpolation operation is a no-memory operation, we need not introduce time in our discussion. Let $E(\nu) = |\omega(\nu)|$ represent the continuous function which is sampled by the filter output envelopes. For simplicity, suppose that the desired triplet of filters is centered on $\nu = 0$ and let

$$\begin{aligned} E(-F) &\equiv E_- & (7.26) \\ E(0) &\equiv E_0 \\ E(F) &\equiv E_+ \end{aligned}$$

Then, if $f(\nu) = a\nu^2 + b\nu + c$ is the quadratic passing through the points $(E_-, -F)$, $(E_0, 0)$ and (E_+, F) , it satisfies the equations

$$\begin{aligned} E_- &= aF^2 - bF + c \\ E_0 &= c \\ E_+ &= aF^2 + bF + c. \end{aligned} \tag{7.27}$$

where the coefficient c is given directly and the coefficients a and b are readily computed as

$$\begin{aligned} b &= \frac{E_+ - E_-}{2F} \approx \left[\frac{dE(\nu)}{d\nu} \right]_{\nu=0} \equiv \dot{E}(0) \\ a &= \frac{(E_+ - E_0) - (E_0 - E_-)}{2F^2} \approx \frac{1}{2} \left[\frac{d^2E(\nu)}{d\nu^2} \right]_{\nu=0} \equiv \frac{1}{2} \ddot{E}(0) \end{aligned} \tag{7.28}$$

and where the approximations indicated are valid for sufficiently small F .

One may readily determine that the maximum value of $f(\nu)$ occurs at

$$\nu = -\frac{b}{2a} = \frac{F}{2} \left[\frac{E_- - E_+}{E_+ + E_- - 2E_0} \right] \approx -\frac{\dot{E}(0)}{\ddot{E}(0)}. \tag{7.29}$$

It follows from the above that if the filter which has the maximum output envelope has a center frequency located at $\nu = kF$, the location of the maximum $\hat{\nu}$ of the interpolated function is given by

$$\hat{\nu} = F \left[K + \frac{1}{2} \frac{E_- - E_+}{E_+ + E_- - 2E_0} \right] \tag{7.30}$$

A signal estimation technique that converts frequency resolution into time resolution prior to the variable estimation has been presented by Darlington.¹⁶ He analyzes demodulators that estimate the frequency of a transmitted FM signal derived from a sample - and - held waveform and denotes this method of demodulation as spectrum analysis. One example would be a bank of narrow -band filters as suggested by Akima and Darlington points out that this system suffers from the practical disadvantage of estimating an unquantized variable (frequency) with a finite number of filters. In order to obtain a very precise estimate, an awkwardly large number of filters is required.

To avoid these difficulties, Darlington suggests an alternate receiver design as shown in Fig. 48. The input signal (constant frequency in any one sample interval) is mixed with a linearly swept oscillator and the output is then applied to a dispersive line which has a delay proportional to frequency. The output of the delay line is a pulse delayed by an amount proportional to the frequency of the input (see Fig. 49). The only remaining operation is to estimate the time of occurrence of this pulse.

Darlington claims that this is an easier operation than the original operation of using a finite number of filters to estimate the input frequency. However, the advantages are more apparent than real. A simple threshold type level detector will give nowhere near the optimum mean-square error which Darlington calculates. A more sophisticated device which detects both the upward and downward passage through a threshold and averages the time of these two events will similarly give substantially worse performance than the optimum. To approach the optimum one might consider correlating the pulse with a set of delayed versions of the expected pulse shape and whichever member of the set had a delay nearest to the actual delay would produce the largest output. However, this is exactly the situation we set out to avoid in the first place, so that converting the frequency axis to a time axis presents no real advantage. Therefore, a system of this type offers no advantage over a bank of filters as described by Akima.

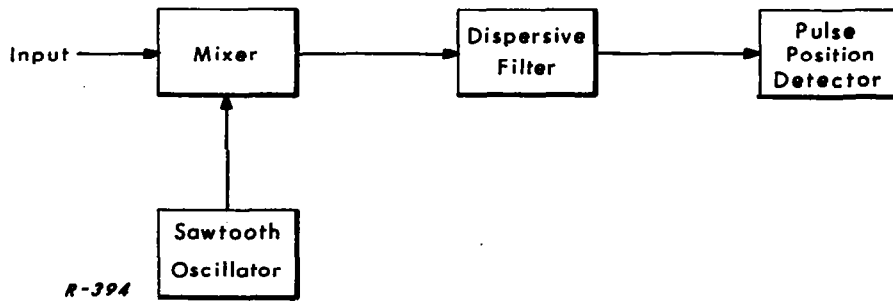


Fig. 48 Darlington system configuration.

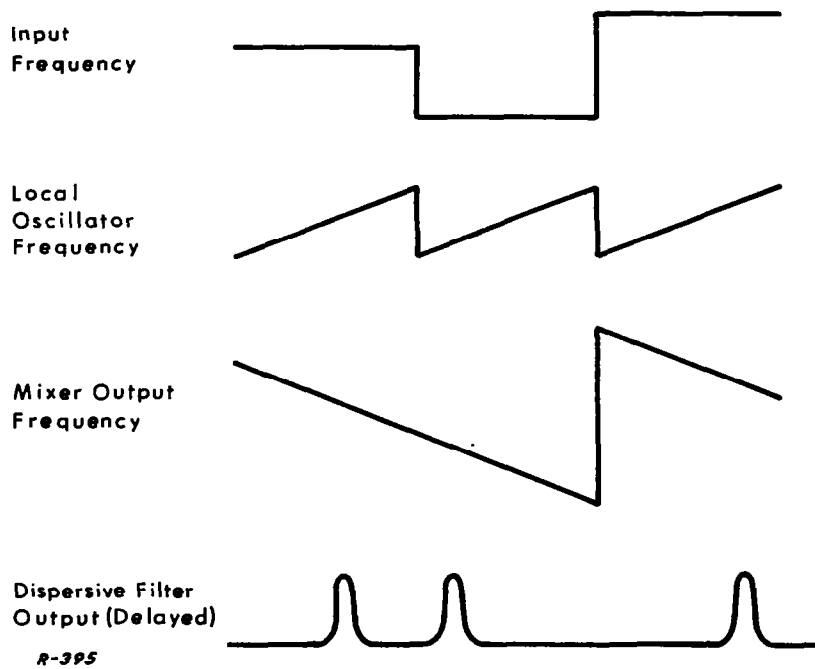


Fig. 49 System operation.

7.6 Conclusion

In this chapter we have studied different band-dividing demodulation techniques. The original system of Akima was found to be at most comparable to a second order PLL demodulator when the individual filter bandwidth constraints arising from quasi-stationary response considerations were accounted for. Also, the effect of mis-selection noise due to the filter overlap is analyzed.

The limitations of the Akima system directly suggest the use of i-f divisions whose widths are larger than their noise bandwidths, and the use of a PLL bank seems a possible solution if we identify the pull-in range with the i-f division. The acquisition implications inherent in this operation are found to result in a system inferior to the second order PLL demodulator due to the long acquisition time involved when sinusoidal phase detectors are used in the PLL bank.

The failure of band-dividing systems whose improvement is based on the ratio of i-f and the individual filter bandwidth to show any promising improvement relative to a PLL demodulator suggests the study of band-dividing signal estimation techniques. A maximum-likelihood band-dividing estimator is derived and some practical approximations are discussed. The SNR improvement analysis of these systems needs further investigation to motivate any development at this stage.

REFERENCES

1. Youla, D. C. , "The Use of the Method of Maximum Likelihood in Estimating Continuously Modulated Intelligence Which Has Been Corrupted by Noise", IRE Trans. on Information Theory, March 1954.
2. Viterbi, A. J. , "Phase-Locked Loop Dynamics in the Presence of Noise by Fokker-Planck Techniques", Proc. IEEE, December 1963.
3. Develet, J. A. , "An Analytic Approximation of Phase-Lock Receiver Threshold", IEEE Trans. On Space Electronics and Telemetry, March 1963.
4. Van Trees, H. L. , "An Introduction to Feedback Demodulation", AD-416586.
5. Sanneman, R. W. , and Rowbotham, J. R. , "Unlock Characteristics of the Optimum Type II Phase-Locked Loop", IEEE Trans. on Aerospace and Navigational Electronics, March 1964.
6. Develet, J. A. , "A Threshold Criterion for Phase-Lock Demodulation", Proc. IEEE, February 1963.
7. Balodis, M. , "Laboratory Comparison of Tanlock and Phase-Lock Receivers", Natl. Telem. Conf. , 1964.
8. Yovits, M. C. and Jackson, J. L. , "Linear Filter Optimization with Game Theory Considerations", 1955 IRE National Convention Record, Part 4, pp.193-199.
9. Seshu, S. and Balabanian, Norman, Linear Network Analysis, Wiley, New York, 1959.
10. Landee, Davis and Albrecht, Electronic Designers' Handbook, McGraw-Hill, New York, 1957.
11. Downing, J. J. , Modulation Systems and Noise, Chapter V, Prentice-Hall, Inc. , New Jersey, 1964.
12. Heitzman, R. E. , "A Study of the Threshold Power Requirements of FMFB Receivers", IRE Trans. on Space Electronics and Telemetry, December 1962.
13. Akima, H. , "Theoretical Studies of SNR Characteristics of an FM System", IEEE Trans. on Space Electronics and Telemetry, December, 1963.

REFERENCES (CONT'D)

14. Richman, D. , "Color-Carrier Prereference Phase Synchronization in NTSC Color TV " , Proc. IRE, January, 1954.
15. ADCOM, Inc. , "High Accuracy Satellite Tracking Systems" , 9th Quarterly Progress Report, Contract No. NAS 5-1187, June, 1963.
16. Darlington, S. , "Demodulation of Wideband, Low-Power FM Signals " , BSTJ, January, 1964.

Appendix A

A MAXIMUM LIKELIHOOD BAND-DIVIDING DEMODULATOR

This Appendix studies the operation of an FM demodulator that uses the outputs of a bank of filters for demodulating an FM signal. After discussing the characteristics of the noise and signal terms at the outputs of these filters, we describe a maximum likelihood estimator which processes the filter outputs to obtain an estimate of the desired FM.

Signal and Noise Characterization

Let $h(t)$ represent the complex envelope of the impulse response of the filter located at the "center" of the bank of filters. Then $h(t) e^{j\pi\nu t}$ represents the complex envelope of a filter removed ν cps from the center filter (assuming that our reference or "carrier" frequency for the definition of complex envelopes is chosen equal to the center frequency of the center filter). In the absence of noise the output (complex envelope) of the filter located at ν cps is given (apart from an irrelevant factor of $1/2$) by

$$Y(t, \nu) = \int S(t - \zeta) h(\zeta) e^{j\pi\nu\zeta} d\zeta \quad (\text{A. 1})$$

where $S(t)$ is the input signal (complex envelope). Similarly the output noise is given by

$$Z(t, \nu) = \int N(t - \zeta) h(\zeta) e^{j\pi\nu\zeta} d\zeta \quad (\text{A. 2})$$

where $N(t)$ is the input noise. If we momentarily consider the possibility of a continuum of filters, then $Y(t, \nu)$, $Z(t, \nu)$ are random processes in two independent variables (t, ν) which describe the signal and noise outputs of these filters.

Since $Z(t, \nu)$ is derived by a linear operation on $N(t)$ one may assume $Z(\cdot)$ to be a gaussian process if $N(\cdot)$ is. The autocorrelation function of $Z(t, \nu)$ is given by

$$\begin{aligned}
& \overline{Z^*(t, \nu) Z(t + \tau, \nu + \Omega)} \\
&= \iint N^*(t - \zeta) N(t + \tau - \eta) h^*(\zeta) h(\eta) e^{-j2\pi [\nu\zeta - (\nu + \Omega)\eta]} d\zeta d\eta \\
&= \int R(\tau - \sigma) H(\sigma, \Omega) e^{j2\pi\nu\sigma} d\sigma \tag{A. 3}
\end{aligned}$$

where $R(\tau) = \overline{N^*(t) N(t + \tau)}$ is the autocorrelation function of the input noise and

$$H(\sigma, \Omega) = \int h^*(\tau - \sigma) h(\eta) e^{j2\pi\Omega\eta} d\eta \tag{A. 4}$$

may be taken as the ambiguity function of the filter impulse response.

Assuming that the filter outputs are examined at the same time instant and that the noise is white with unit power density, the noise correlation function of interest is

$$\begin{aligned}
\overline{Z^*(t, \nu) Z(t, \nu + \Omega)} &= H(0, \Omega) \\
&= \int |h(\zeta)|^2 e^{j2\pi\Omega\zeta} d\zeta \\
&= \int H^*(f) H(f - \Omega) df \equiv R_H(\Omega) \tag{A. 5}
\end{aligned}$$

where $H(f)$ is the center filter transfer function (lowpass equivalent). In view of Eq.(A. 5) the process $Z(t, \nu)$ with t fixed is statistically stationary in the variable ν when the input noise is white.

We turn our attention now to the signal term $Y(t, \nu)$. We shall briefly study the conditions on the filter characteristic for an input frequency-modulated signal so that the output differs little from the quasi-stationary output. It is desirable to explicitly indicate the group delay provided by the filter (at zero frequency). Thus let

$$h(t) = k(t - \Delta) \tag{A. 6}$$

where Δ is the filter phase slope at zero frequency and $k(t)$ is the impulse response of a hypothetical filter with zero phase slope at zero frequency.

In the case of FM or PM the input signal can be represented by $S(t) = e^{j\phi(t)}$ so that using this representation and Eq. (A. 6) in Eq. (A. 1) we find

$$\begin{aligned} Y(t, \nu) &= \int e^{j[\phi(t-\zeta) + 2\pi\nu\zeta]} k(\zeta - \Delta) d\zeta \\ &= e^{j[\phi(t-\Delta) + 2\pi\nu\Delta]} \int e^{j[\phi(t-\zeta-\Delta) - \phi(t-\Delta)] + j2\pi\nu\zeta} k(\zeta) d\zeta \end{aligned} \quad (\text{A. 7})$$

and upon using the definition

$$R(t, \zeta) = \phi(t-\zeta) - \phi(t) + \zeta \dot{\phi}(t) \quad (\text{A. 8})$$

we find that

$$Y(t, \nu) = e^{j[\phi(t-\Delta) + 2\pi\nu\Delta]} K\left(\nu - \frac{\dot{\phi}(t-\Delta)}{2\pi}\right) + E(t, \nu) \quad (\text{A. 9})$$

where

$K(\cdot)$ is the transform of $k(t)$ and

$$E(t, \nu) = e^{j[\phi(t-\Delta) + 2\pi\nu\Delta]} \int e^{j\zeta[2\pi\nu - \dot{\phi}(t-\Delta)]} \left[e^{jR(t-\Delta, \zeta)} - 1 \right] R(\zeta) d\zeta \quad (\text{A. 10})$$

It may be recognized that the first term in Eq. (A. 9) is the quasi-stationary portion of the output signal. The magnitude of the error term is bounded as follows

$$\begin{aligned} |E(t, \nu)| &\leq \int \left| e^{jR(t-\Delta, \zeta)} - 1 \right| |k(\zeta)| d\zeta \\ &= \int 2 \left| \sin \frac{R(t-\Delta, \zeta)}{2} \right| |k(\zeta)| d\zeta. \end{aligned} \quad (\text{A. 11})$$

so that a sufficient condition for small distortion assuming $K(0) = 1$ is

$$\left| R(t-\Delta, \zeta) \right| \ll 1 \text{ for } -\frac{T}{2} < \zeta < \frac{T}{2} \quad (\text{A. 12})$$

where $|k(t)|$ is assumed to be negligible for t outside an interval of duration T centered on $t = 0$. Assuming that the indicated derivatives exist, then $R(t, \zeta)$ has the expansion

$$R(t, \zeta) = \sum_{n=2}^N \frac{(-\zeta)^n}{n!} \frac{d^n \phi(t)}{dt^n} + \frac{(-\zeta)^N}{N!} \frac{d^N \phi(\theta)}{d\theta^N}; t - \zeta < \theta < t \quad (\text{A. 13})$$

which leads to the sufficient condition

$$\frac{\zeta^2}{2} \left| \frac{d^2 \phi(\theta)}{d\theta^2} \right| \ll 1 \text{ for } -\frac{T}{2} < \zeta < \frac{T}{2}, t - \zeta < \theta < t. \quad (\text{A. 14})$$

and the least tight condition is clearly*

$$\text{Max} \left| \frac{d^2 \phi(t)}{dt^2} \right| \ll \frac{2}{T^2} \quad (\text{A. 15})$$

For the case of sinusoidal modulation given by $\phi(t) = \delta \sin \omega_m t$ then Eq. (A. 15) becomes after a slight rearrangement

$$f_m^2 T^2 \ll \frac{1}{2\delta\pi^2} \quad (\text{A. 16})$$

so that for a large modulation index the product $f_m T$ must be very much less than one if the sufficient quasi-stationary condition Eq. (A. 15) is to hold for sinusoidal modulation. Even for values of δ of the order of unity Eq. (A. 15) implies that $f_m T \ll 1$.

* When $\phi(t)$ is a random process Eq. (A. 14) and Eq. (A. 15) are not so convenient. In such a case one should replace Eq. (A. 16) by

$$\sqrt{|R(t-\Delta, \zeta)|^2} \ll 1, \quad -\frac{T}{2} < \zeta < \frac{T}{2}.$$

Signal Estimation

In terms of the hypothetical filter with zero group delay we may represent the output of a filter with center frequency ν as

$$w(t, \nu) = \left[Z(t, \nu) + e^{j\dot{\phi}(t-\Delta)} K\left(\nu - \frac{\dot{\phi}(t-\Delta)}{2\pi}\right) \right] \exp [j2\pi\nu\Delta] \quad (\text{A. 17})$$

where the noise

$$Z(t, \nu) = \int N(t - \zeta - \Delta) k(\zeta) e^{j2\pi\nu\zeta} d\zeta \quad (\text{A. 18})$$

has correlation function

$$\overline{Z^*(t, \nu) Z(t, \nu + \Omega)} = \int K(f) K^*(f + \Omega) df \quad (\text{A. 19})$$

assuming $N(t)$ to be white and of unit spectral density.

If $2N + 1$ equally spaced filters are available we may state our estimation problem as follows: given the $2N + 1$ random processes $w(t, kF)$; $k = 0, \pm 1, \dots, \pm N$, then determine the frequency modulation $\dot{\phi}(t)/2\pi$.

An obvious heuristic solution to this problem arises from the observation that the magnitude of the signal component in Eq. (A. 17) has a maximum at

$$\nu(t) = \nu_{\max} + \frac{\dot{\phi}(t-\Delta)}{2\pi} \quad (\text{A. 20})$$

where

$$\text{Max}_{\nu} |K(\nu)| = |K(\nu_{\max})|. \quad (\text{A. 21})$$

Assuming a symmetrical unimodal filter

$$\nu_{\max} = 0 \quad (\text{A. 22})$$

and the signal component has its maximum magnitude at the filter whose center frequency equals the "instantaneous" frequency of the input. Thus, neglecting

noise, one possible estimation procedure is to determine $|w(t, kF)|$ and by an interpolation procedure estimate the value of ν for which $|w(t, \nu)|$ is maximum. This value of ν , say $\tilde{\nu}(t)$, will be an estimate of $\dot{\phi}(t-\Delta) 2\pi$.

Unfortunately there is no proof that this procedure will result in an optimum estimate of the input frequency modulation. We shall determine an optimum procedure for the particular case in which the outputs of the $2N + 1$ filter are examined at the same time instant, i. e., a no-memory processing of the filter outputs. For convenience in notation let

$$\begin{aligned} Z(t, \nu) &= Z(\nu) \\ \phi(t - \Delta) &= \phi \\ \frac{\dot{\phi}(t - \Delta)}{2\pi} &= \zeta \\ w(t, \nu) &= w(\nu) \end{aligned} \tag{A. 23}$$

so that

$$W(\nu) = w(\nu) e^{-j2\pi\nu\Delta} = z(\nu) + e^{j\phi} K(\nu - \zeta) \tag{A. 24}$$

Since Δ is known we may assume that $w(\nu) e^{-j2\pi\nu\Delta} \equiv W(\nu)$ is observed rather than $w(\nu)$ alone. A study of Eq.(A. 24) reveals that if $w(\nu) e^{-j2\pi\nu\Delta}$ were known for all values of ν , the problem of estimating ζ would be formally identical to the problem of determining the position of a pulse of known shape but unknown phase (assuming ϕ unknown) and position in stationary colored gaussian noise of autocorrelation function given by Eq. (A. 19).

Since $w(\nu)$ is known only for a finite set of values of ν our problem is a sampled version of the above problem, namely: determine the position of a pulse of unknown phase and position in stationary colored gaussian noise by observation of a finite number of periodically spaced samples.

Using vector and matrix notation, we define the observed filter outputs and the signal and noise terms as

$$\bar{W} = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ W(F) \\ W(o) \\ W(-F) \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \quad (\text{A. 25})$$

$$\bar{Z} = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ Z(F) \\ Z(o) \\ Z(-F) \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \quad (\text{A. 26})$$

$$\bar{K}(\xi_o) = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ K(F-\xi_o) \\ K(-\xi_o) \\ K(-F-\xi_o) \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \quad (\text{A. 27})$$

so that the observed vector \bar{W} is given by

$$\bar{W} = \bar{Z} + e^{j\phi} \overline{K(\zeta_0)} \quad (\text{A. 28})$$

when the parameter value $\zeta = \zeta_0$.

This problem of optimally estimating ζ may be solved in two steps. First one may determine a linear transformation of the set of samples (i. e., \bar{W}) which will yield independent noises and then one may solve the simpler independent noise case. Since the linear transformation is reversible no information is lost and the overall optimization problem will remain unchanged if we deal with the transformed variables. Let the noise whitening transformation be denoted by T , the transformed samples by \bar{W}' , and the transformed noises by \bar{Z}' . Then

$$\bar{W}' = \bar{Z}' + e^{j\phi} T \overline{K(\zeta_0)} \quad (\text{A. 29})$$

where

$$\begin{aligned} \bar{W}' &= T \bar{W} \\ \bar{Z}' &= T \bar{Z}. \end{aligned} \quad (\text{A. 30})$$

In the transformed problem the noises are independent and the optimization problem is considerably simplified. Standard statistical techniques may be used to obtain minimum mean squared error or maximum likelihood estimates. Thus it may be shown that the maximum likelihood estimate of ζ is given by $\tilde{\zeta}$ below,

$$\tilde{\zeta} = \text{Max}_{\zeta} \left\{ \ln I_0 \left[\left| \sum_{-N}^N w_k' s_k^*(\zeta) \right| \right] - \sum_{-N}^N |s_k(\zeta)|^2 \right\} \quad (\text{A. 31})$$

where $\{s_k(\zeta)\}$ are the elements of $T \overline{K(\zeta)}$

$$TK(\zeta) = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ s_1(\zeta) \\ s_0(\zeta) \\ s_{-1}(\zeta) \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \quad (\text{A. 32})$$

and $\{w_k\}$ are the elements of \overline{W}^T ,

$$\overline{W}^T = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ w_1^T \\ w_0^T \\ w_{-1}^T \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \quad (\text{A. 33})$$

The first sum in Eq. (A. 31) represents a determination of the correlation between the transformed input and expected signal while the second term is a bias term depending upon the energy of the transformed signal vector. The function $I_0 []$ is the modified Bessel function of the first kind and order zero. One may interpret the operations $\ln I_0 (|)$ as a nonlinear envelope detection operation.

A block diagram of the maximum likelihood estimator is shown in Fig. 45. Since this estimator processes the filter bank outputs at the same time instant, i. e., since it is a no-memory processing, further processing which involves filtering operations can be expected to reduce the noise still further. The most obvious processing of $\tilde{\zeta}(t)$ would be a straightforward filtering to the known bandwidth of $\phi(t)$.

One may formulate a more involved optimal estimation problem that involves general processing of the filter bank outputs with memory. It appears that such processing is too complicated to be of practical importance. Even the scheme of Fig. 46 is rather involved. There is one special situation in which the operation of Fig. 46 simplify, namely, when the impulse responses (and thus transfer functions) of the filters are orthogonal. In such a case the noises at the filters outputs become independent and the transformation T is unnecessary.

If in addition N is chosen large enough so that the set of values $\{K(k-\zeta); k = 0, \pm 1, \dots, \pm N\}$ represents a closely sampled version of $K(\nu-\zeta)$ for any expected range of values of ζ then the bias term becomes independent of ζ and may be removed. Also since $\ln I_0[x]$ is monotonically related to x a simple envelope detector will be satisfactory.