THE BELL SYSTEM
TECHNICAL JOURNAL

VOLUME XLI NOVEMBER 1962 NUMBER 6

Copyright 1962, American Telephone and Telegraph Company

An Analysis of Inherent Distortion in Asynchronous Frequency-Shift Modulators

By L. R. BOWYER and W. H. HIGHLEYMAN

(Manuscript received May 15, 1962)

In many commonly used frequency-shift modulators, a phase error occurs at the time of switching. If a demodulator is used which utilizes only the zero-crossing information, then this phase error will cause time jitter in the received data transitions.

The magnitude of the peak time jitter for various modulators is derived, assuming an ideal zero-crossing detector. The modulators considered include the reactance tube and variable reactance modulators, the basic switched reactance modulators, and the multivibrator modulator. It is found that the switched reactance modulators cause the most jitter, and that the multivibrator modulator may be designed to cause as small a jitter as desired. The theory agrees well with some experimental measurements made on existing data sets, which show that this jitter accounts for most of the back-to-back data distortion in many wideband data systems.

Finally, a set of sufficient conditions is derived for jitter-free frequency-shift modulation, and an implementation of a modulator satisfying these conditions is described.

I. SUMMARY

For the reader who may be more interested in the results of this paper than in their derivation, the following summary is presented.

Data communication systems using frequency-shift channels commonly suffer from a form of fortuitous distortion called jitter. This is

particularly significant in systems in which the bit rate is not small compared to the carrier frequency. Jitter is the variation about the correct position of the transition between marking and spacing signals at the receiver output. It is desirable to keep the magnitude of the jitter small compared to the bit length.

This paper studies the jitter which is inherent in various types of frequency-shift modulators. For purposes of this analysis, it is assumed that the data source is jitter free, that the transmission channel is distortionless, and that the receiver is an ideal zero-crossing detector. The modulators to be studied shift frequency instantaneously at the time the data source goes through a mark-space transition. In an ideal frequency-shift modulator, this shift in frequency takes place with phase continuity. However, many commonly used modulators do not maintain phase continuity at the switching instant.

In an ideal frequency demodulator (i.e., one whose output is proportional to the instantaneous rate of change of phase at its input), such a phase discontinuity would not cause a time error in the output data transition; it would simply cause an impulse to be added at the time of transition to the otherwise correct transition. However, most present-day demodulators utilize only the information contained in the zero crossings of the received wave, since the first operation in the receiver is to limit, or clip, the wave. In such a receiver, phase discontinuities at the switching instant in the received wave may indeed cause a time error in the mark-space transition at the receiver output.

Such a receiver is modeled by an ideal zero-crossing detector. (An ideal zero-crossing detector approaches an ideal FM detector as the bandwidth becomes small with respect to the carrier frequency.) The analysis proceeds by first relating the phase error in the received wave to the transition time error at the output of an ideal zero-crossing detector. Then the peak phase error that may occur for each type of modulator is determined, and this is related to the peak time error, or jitter, by the above model.

The frequency-shift modulators to be studied include the switched reactance modulator (in which a reactance is switched into and out of the tank of an oscillator to modify its frequency), the reactance tube modulator (in which the effective output reactance of an active circuit is changed by changing the gain of the active element and this reactance is used to control the frequency of a separate oscillator), the variable reactance oscillator (in which the functions of variable reactance and oscillation are combined into a single active circuit), and the multivibrator.

It is shown that of the LC oscillators, the reactance tube and variable reactance type of oscillators have the minimum jitter. For all cases of practical interest, this time error has a peak value roughly one-twelfth of the period of the frequency midway between the marking and spacing frequencies, and may take on both positive and negative values with respect to the true transition time.

The switched reactance type of oscillators which are analyzed include all four ways of switching a single reactive element into a simple LC tank, i.e., an inductor or a capacitor switched in parallel with, or in series with, the tank. In all cases, the peak jitter is described by the same equation, although there are two distinctly different phenomena giving rise to the jitter. For a transition in one direction, the peak jitter is exactly that obtained with the reactance tube and variable reactance type of modulators. For the opposite transition, however, undesired dc quantities increase the jitter. This increase in jitter is sensitive to bit rate if there is a decay mechanism for the undesired dc quantity. In the worst case, the increase in jitter is almost an order of magnitude over that of the low jitter transition. As the bit rate becomes lower and lower, the peak jitter associated with the worse transition approaches that of the opposite transition.

The peak time error τ_{em} for the various LC modulators is plotted in Fig. 7(a) as a function of the frequency-shift ratio $A = \omega_1/\omega_2 < 1$, where ω_1 and ω_2 are the two modulator frequencies. T_o is the period of the frequency midway between the marking and spacing frequencies. The parameter r applies to the switched reactance modulators. It is the ratio of the bit length to the time constant of the undesired dc quantity.

It is shown that the multivibrator may in principle be designed to be jitter-free. However, practical multivibrators do have some inherent jitter. The amount of jitter is dependent upon a linearity factor which is determined by the multivibrator circuit. The jitter associated with a multivibrator is usually less than the jitter associated with any of the above LC modulators.

In Fig. 13 is shown the peak jitter, τ_{em}, for the multivibrator modulator as a function of the frequency-shift ratio A. T_o is as defined in Fig. 7. The linearity factor, β, is the ratio of the supply voltage to the maximum control (voltage V_s and V_c respectively in Fig. 10a). The maximum control voltage corresponds to the highest frequency.

These theoretical results are supported by experimental jitter measurements made on various types of voice band data sets, as summarized in Table I. The agreement with the theory is good.

Finally, it is shown that a sufficient condition for jitter-free frequency-

shift modulation (when using a zero-crossing detector) is to switch the parameters of an oscillating tank circuit in such a way that

1. the tank current and voltage are maintained at the instant of switching, and
2. the characteristic impedance of the tank is held constant.

A means for implementing such an oscillator is described using a pair of variable reactance oscillators in a balanced arrangement. Some experimental waveforms for this circuit are shown in Fig. 16.

II. INTRODUCTION

A large part of present-day data communications systems utilize binary frequency-shift channels for transmission. In such systems, particularly in the wideband systems (i.e., systems in which the bit rate is not small compared to the carrier frequency), the problem of "jitter" is important. Jitter is the error in the reproduction of the exact time of transition between marking and spacing signals at the receiver output.

This problem is indicated in Fig. 1, in which it is assumed that the indicated data-source waveform is jitter-free and is transmitted repetitively. The received waveforms are shown overlapping one another. Note that, in addition to a usually harmless transmission delay, the receiver transitions do not all occur at the same relative time, but are instead distributed about the correct transition. The measure of jitter which will be used in this paper will be the maximum time error which may occur in the system, denoted by τ_{em} in Fig. 1. (In all systems con-

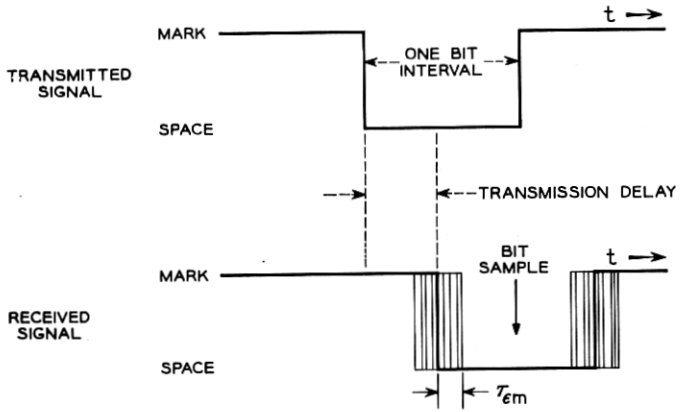

Fig. 1 — The problem of jitter.

sidered but one, the jitter is symmetric about the correct transition. However, the maximum jitter associated with a transition in one direction is not necessarily the same as that associated with the opposite transition. In these cases, a different $\tau_{\epsilon m}$ will be associated with each of the two transitions. The one exception to the above statement is the multivibrator modulator, in which the jitter occurs on only one side of the correct transition.)

As $\tau_{\epsilon m}$ increases, the time interval during which the received wave may be reliably sampled during each bit interval decreases. Thus it is important to keep this jitter to a minimum. Jitter may be induced from several sources. The data source itself may have inherent jitter. Jitter may be further induced in the modulation process. Distortion and noise present in the transmission channel will modify the modulated waveform, thus perhaps changing the apparent transition times. Finally, the receiver may contribute to the jitter.

This paper will analyze that jitter inherent in various asynchronous* frequency-shift modulation techniques. It will be assumed that the data source is jitter free, that the channel is ideal (i.e., noiseless and distortionless), and that the system utilizes an ideal zero-crossing type of detector. This detection process will be described more fully in the next section.

All the modulators to be studied shift frequency instantaneously at the time the data source goes through a mark-space transition. That is, the waveshape before the transition is given by

$$V_1 \cos (\omega_1 t + \theta_1)$$

and after the transition by

$$V_2 \cos (\omega_2 t + \theta_2).$$

The angular frequency before the transition is given by ω_1 and after the transition by ω_2 . The expressions $\omega_1 t + \theta_1$ and $\omega_2 t + \theta_2$ are defined as the *phase* of the wave, before and after the transition, respectively. Ideally, the shift in frequency should take place with phase continuity; i.e. $\omega_1 t + \theta_1$ should equal $\omega_2 t + \theta_2$ at the time of transition. However, it will be shown that many commonly used modulators have an inherent phase error that results in an inherent time jitter.

The modulators to be studied include the switched reactance modulator (in which a reactance is switched into and out of the tank of an oscillator to modify its frequency), the variable reactance oscillator,[1] the

* That is, the frequency may be suddenly shifted during any portion of the cycle.

reactance tube modulator,[2] and the multivibrator.[3] It will be shown that the switched-reactance type of modulator has the most jitter. For example, at a data rate corresponding to one cycle of carrier per bit, the switched reactance modulator has a minimum peak jitter equivalent to 7.96 per cent data distortion. The variable reactance oscillator and reactance tube type of modulators are equivalent and have less jitter. The multivibrator may in principle be designed to have negligible jitter, but produces a square rather than a sinusoidal wave. These theoretical results are then supported by experimental data.

III. THE IDEAL ZERO-CROSSING DETECTOR

Before pursuing the jitter analysis further, it is necessary to define in some detail the assumed detection process. An ideal FM detector is one which measures the instantaneous rate of change of phase of the received wave. For the types of modulators considered in this paper, there would be no jitter in the received wave using such a detector. Rather, the phase discontinuity at switching would simply cause an impulse to be added at the time of transition to the otherwise correct transition. However, most present-day detectors utilize only the zero-crossing information in the received wave. For example, consider the receiving portion of the 43A1 carrier telegraph system.[4] The incoming signals are first passed through a filter to reject out-of-band noise, and then amplified. The next step is to limit the wave. The purpose of this is to remove any amplitude modulation, but in addition it removes any information other than that carried by the zero-crossings. The wave is next passed through a discriminator to convert the frequency modulation to baseband information, followed by a dc amplifier. The output is then a square wave which duplicates closely the wave originally presented to the modulator. Although other data sets use different methods of demodulation, a limiter to remove amplitude modulation is part of many of these. Thus is it of practical interest to study the performance of frequency-shift modulators when working with a zero-crossing type of detector. A zero-crossing detector approaches an ideal FM detector as the bandwidth of the transmitted signal becomes small with respect to the carrier frequency. This relation is described more fully by Stumpers.[5]

The characteristics of the zero-crossing detector are shown graphically in Fig. 2. Let it be receiving a constant frequency sine wave with zero phase at the time origin. Each zero crossing is identified by the zero-crossing detector as an advancement of π radians in the received wave. It thus knows the rate of change of phase, and hence the frequency, of the incoming sinusoidal wave. This is represented by the solid line in

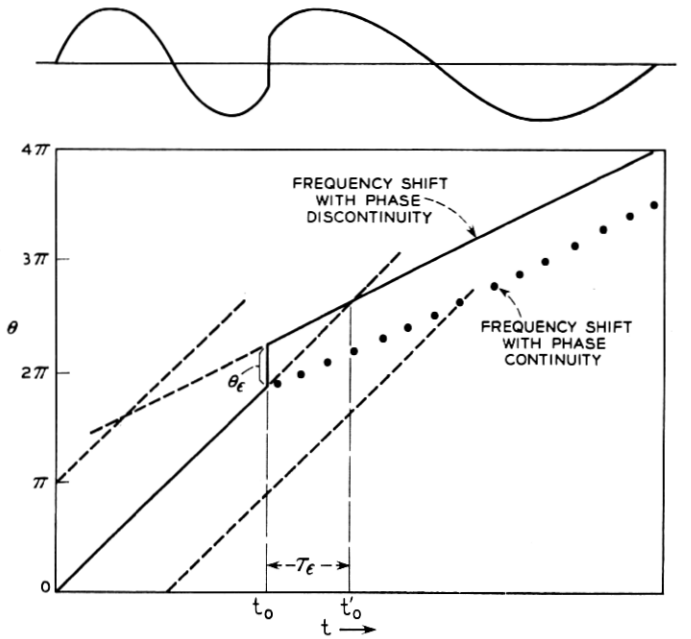

Fig. 2 — Graphical representation of the zero-crossing detector.

Fig. 2. However, the detector knows the phase of the wave only within an uncertainty factor of $n\pi$; the wave may equally well be represented by lines parallel to the solid line in Fig. 2, but of ordinate distance $n\pi$ away, where n is any integer. Two such lines are shown dashed.

If, at $t = t_o$, the frequency of the incoming wave is suddenly changed, the rate of zero crossings will be suitably changed. If the transition is made with phase continuity in the received wave, then the slope of the path in Fig. 2 simply changes suddenly, and the detector knows that the transition between frequencies occurred at the time of the break point. This case is illustrated by the dotted line in Fig. 2.

The problem of jitter is introduced when phase continuity is not preserved. In Fig. 2 is also shown a plot of the phase when the frequency is suddenly changed with an associated phase discontinuity of magnitude θ_ϵ. The detector has no knowledge of instantaneous phase except the zero-crossing information. It therefore assumes phase continuity, and determines the time t_o' at which a transition with phase continuity would have occurred to maintain the observed zero crossings. This time is determined graphically in Fig. 2 by the intersection of the two constant frequency lines determined by the detector. Since there is an indeter-

minacy of $n\pi$ in the absolute phase, there will be a multiplicity of such intersections, two of which are shown. It is assumed that the intersection yielding the minimum time error is chosen by the detector. Thus it is seen that a phase error of $n\pi$ radians causes no error in the detected transition time, as one would intuitively suspect. Furthermore, the phase error causing maximum time error is $\pi/2$ radians, or odd multiples thereof.

One problem has not been resolved. The zero-crossing detector is designed only for the two modulator frequencies at its input; it therefore is presumably incapable of recognizing zero crossings occurring at intervals closer than one-half a period of the upper frequency, since such zero crossings could not occur in a phase-continuous input wave. But with phase discontinuities, it is possible to get zero crossings closer than this interval. It is assumed that the zero-crossing detector ignores such zero crossings.

The relation between phase error and time error may be easily determined from Fig. 3 by noting that the slope of a line segment is the angular frequency, ω, of the input wave. Let the two modulator frequencies be denoted ω_1 and ω_2, where ω_1 is the frequency before switching, and ω_2 is the frequency after switching. Then from Fig. 3,

$$\omega_1 = \frac{\theta_1}{\tau_\epsilon}$$

$$\omega_2 = \frac{\theta_1 + \theta_\epsilon}{\tau_\epsilon}.$$

Thus

$$\tau_\epsilon = \frac{\theta_\epsilon}{\omega_2 - \omega_1}, \qquad \left|\theta_\epsilon\right| \leqq \frac{\pi}{2}. \tag{1}$$

For arbitrary θ_ϵ, this may be written

$$\tau_\epsilon = \min_n \frac{\theta_\epsilon - n\pi}{\omega_2 - \omega_1}, \qquad n \text{ an integer}, \tag{2}$$

where $\min_n f(n)$ means that $f(n)$ is calculated for that value of n which minimizes $f(n)$. The form (1) will usually be used with the understanding that (2) holds if

$$\left|\theta_\epsilon\right| > \frac{\pi}{2}.$$

Let $\theta_{\epsilon m}$ be the maximum possible phase error for a given system, and

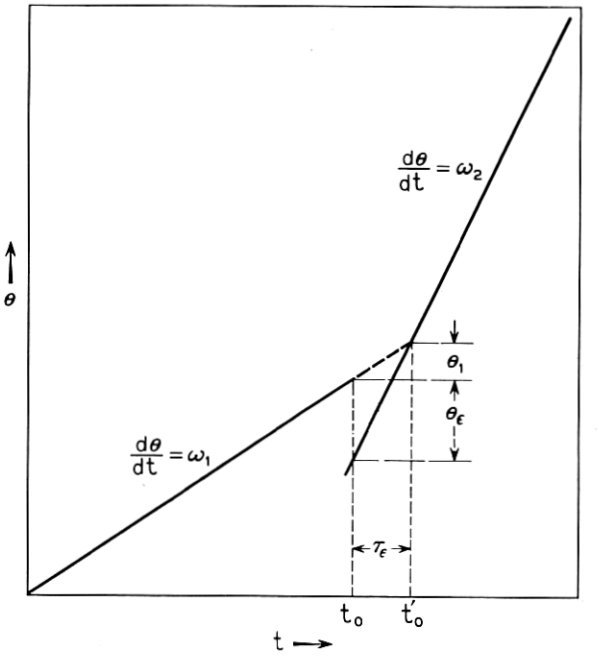

Fig. 3 — Relating phase error to time error.

let $\tau_{\epsilon m}$ be the resulting value of time error. Then (1) may be rewritten in a useful form as follows:

$$\tau_{\epsilon m} = \frac{\theta_{\epsilon m}}{(\omega_2 + \omega_1)\omega_2\left(1 - \frac{\omega_1}{\omega_2}\right)} \omega_2\left(1 + \frac{\omega_1}{\omega_2}\right),$$

or

$$\frac{\tau_{\epsilon m}}{T_0} = \frac{\theta_{\epsilon m}}{4\pi} \frac{1 + A}{1 - A}, \tag{3}$$

where $A = \omega_1/\omega_2$, and $T_o = 4\pi/(\omega_1 + \omega_2)$ is the period of the midfrequency (that frequency midway between ω_1 and ω_2). The value of (3) is thus a figure of merit for a system, giving the ratio of maximum jitter, $\tau_{\epsilon m}$, to the period of the midfrequency as a function of the frequency ratio A.

From reasoning such as that shown in Fig. 3, note that a positive phase error corresponds to a time advance for $A < 1$, but a time retarda-

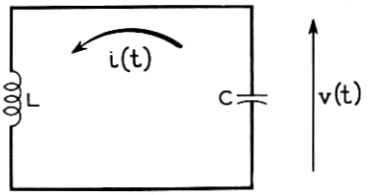

Fig. 4 — Initial tank conditions.

tion for $A > 1$. From (3), then, positive τ_{em} always corresponds to a time advance (an early transition) and negative τ_{em} corresponds to a time retardation (a late transition).

IV. PRELIMINARY REMARKS

The analysis of the inherent time jitter associated with a particular modulator will proceed by first determining the maximum phase error that would occur under the worst conditions, and then relating that phase error to the time error by the use of (3). Before going further, it will be useful to point out certain miscellaneous facts. Several of these remarks pertain to ideal (lossless) tank circuits, since most of the oscillators studied are LC oscillators which are representable by lossless tank circuits.

First note that only one frequency may exist in an undriven lossless tank, namely the resonant frequency. Therefore, when the frequency of oscillation of a tank circuit is changed by suddenly changing the value of one or more of the tank reactances, the frequency also changes suddenly to the new resonant frequency. There are no transients except for the instantaneous transition; i.e., the tank instantaneously reaches its steady-state condition of amplitude and frequency. In particular, if the tank voltage and current are of magnitude v_o and i_o at $t = 0$, and are in the directions indicated in Fig. 4, then it is well known that the tank oscillation voltage and current will be given by

$$v(t) = V \cos (\omega t + \theta) \tag{4}$$

$$i(t) = I \sin (\omega t + \theta) \tag{5}$$

where*

$$V = \sqrt{v_o^2 + \frac{L}{C} i_o^2} \tag{6}$$

* These are easily derivable, particularly from energy relations.

$$I = \sqrt{\frac{C}{L} v_o^2 + i_o^2} = V \sqrt{\frac{C}{L}} \tag{7}$$

$$\theta = \tan^{-1} \sqrt{\frac{L}{C}} \frac{i_o}{v_o}. \tag{8}$$

The energy stored in this tank is

$$E = \tfrac{1}{2} Cv^2(t) + \tfrac{1}{2} Li^2(t) = \tfrac{1}{2} CV^2 = \tfrac{1}{2} LI^2, \tag{9}$$

and is always constant as long as the tank is undisturbed.

It is of interest to physically relate the phase error to the frequency-shift waveform. Fig. 5 shows two independent ways in which the phase error may manifest itself in the time varying waveform when the frequency of a wave is suddenly changed. Fig. 5(a) shows the obvious type of phase discontinuity which is associated with a discontinuity in a waveform of constant amplitude. Fig. 5(b) shows the less obvious case in which the instantaneous voltage and current in the tank are continuous, but in which the amplitude of oscillation for one frequency is not the same as that for the other frequency. The phase error is determined from the condition of waveform continuity. Referring to Fig. 5(b) and considering switching at $t = 0$, one obtains

$$V_1 \cos \theta_1 = V_2 \cos \theta_2,$$

or

$$\frac{\cos \theta_1}{\cos \theta_2} = \frac{V_2}{V_1}.$$

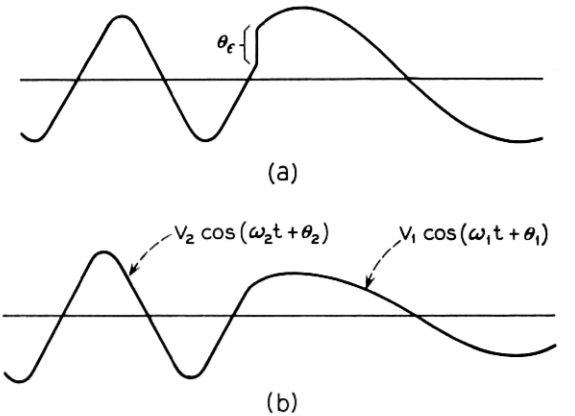

Fig. 5 — Relation of waveform to phase discontinuity: (a) phase discontinuity due to waveform discontinuity; (b) phase discontinuity due to amplitude change.

Thus the phase error will depend not only on the ratio of amplitudes, but also upon the phase angle at which the initial frequency is switched.

In the modulators to be described, either or both of these types of phase error may occur, depending upon the particular situation. In the case of amplitude variations, one might worry about the fact that a practical oscillator will try to maintain a constant amplitude. Thus, if the amplitude attempts to change as in Fig. 5(b), it will quickly recover to the initial value, either due to the drive of the active element in the case of an amplitude drop, or due to the limiting element in the case of an amplitude increase. However, it can be shown that this recovery is an exponential type of recovery; that is, it may be represented by

$$Ve^{\beta(t)t} \cos (\omega t + \theta).$$

Since $e^{\beta(t)t} > 0$ (β is finite in value), the zero crossings of the wave are unchanged. Thus the phase error is unaffected by oscillator recovery, and we will therefore not consider this problem further.

One final point should be made. Some of the discontinuities are rather complicated, in that the voltage discontinuity and current discontinuity may be seemingly unrelated, and the voltage and current amplitude changes may be different. However, the phase discontinuity associated with either the voltage or current waveform will be identical, since the voltage and current must always maintain a 90° phase relationship. Thus, in the following analysis, the calculation of the phase discontinuity for only one of these quantities is derived.

V. THE REACTANCE TUBE MODULATOR AND VARIABLE REACTANCE OSCILLATOR

The jitter associated with the reactance tube modulator and variable reactance oscillator will be derived first. The reactance tube type of modulator is well known in the art.[2] It comprises an oscillator which has one of its reactances determined in part by a virtual reactance, which in turn is determined by the gain of another active stage with appropriate reactance feedback. By varying the gain of the active element (or by varying its feedback parameters), the magnitude of the virtual reactance may be changed, thus changing the resonant frequency of the oscillator. The variable reactance oscillator[1] is essentially the combination of the functions of oscillation and variable reactance into a single active element.

If the frequency of oscillation is shifted by suddenly changing the gain of the reactance-determining active element, then it can be shown

that there is no discontinuity in the voltage and current waveforms at the time of switching. These modulators may therefore be modeled by a simple LC tank, as in Fig. 4, in which one or both of the reactances are suddenly changed to produce a frequency shift, but in such a way that the current and voltage are maintained at the switching instant. In this case, phase error is caused solely by an amplitude change.

Let the subscript 1 refer to the waveform before switching, and the subscript 2 refer to the waveform after switching. Then, using (4), (5), and (7), and assuming switching at $t = 0$, the conditions of voltage and current continuity require that

$$V_1 \cos \theta_1 = V_2 \cos \theta_2 \tag{10}$$

$$V_1 \sqrt{\frac{L_1}{C_1}} \sin \theta_1 = V_2 \sqrt{\frac{L_2}{C_2}} \sin \theta_2 . \tag{11}$$

Dividing (11) by (10) and rearranging yields

$$\tan \theta_2 = \sqrt{\frac{L_1}{C_1} \cdot \frac{C_2}{L_2}} \tan \theta_1 . \tag{12}$$

In most cases, only one of the reactances is changed. Let us assume that the value of the inductance is the shifted parameter, and further that $L_2 < L_1$ (that is, $\omega_2 > \omega_1$). All other cases (i.e., when $L_2 > L_1$, or when the capacitance is varied) are completely analogous and result in the same magnitude of jitter. Equation (12) may then be written as

$$\theta_2 = \tan^{-1} \frac{1}{A} \tan \theta_1 ,$$

where $A = \omega_1/\omega_2 < 1$ as previously defined in the discussion on zero-crossing detection.

The phase error as a function of the phase at shifting, θ_1, is then

$$\theta_\epsilon = \theta_2 - \theta_1 = \left(\tan^{-1} \frac{1}{A} \tan \theta_1 \right) - \theta_1 . \tag{13}$$

In order to determine $\theta_{\epsilon m}$, the maximum value of θ_ϵ, (13) is differentiated with respect to θ_1 and equated to zero (the value of θ_1 satisfying this condition will be denoted θ_{1m}):

$$\frac{1}{1 + \frac{1}{A^2} \tan^2 \theta_{1m}} \frac{1}{A} \sec^2 \theta_{1m} - 1 = 0.$$

Solving this for θ_{1m} yields

$$\theta_{1m} = \cos^{-1} \frac{\pm 1}{\sqrt{1 + A}} = \pm \tan^{-1} \sqrt{A} .$$

Thus

$$\theta_{\epsilon m} = \pm \left(\tan^{-1} \frac{1}{\sqrt{A}} - \tan^{-1} \sqrt{A} \right). \tag{14}$$

This is a monotonically decreasing function of A, and reaches a maximum for $A = 0$ (infinite frequency shift) of $\pi/2$ radians. For no frequency shift ($A = 1$), the phase error is zero as one would expect. This relation is shown plotted as the "∞" line in Fig. 6.

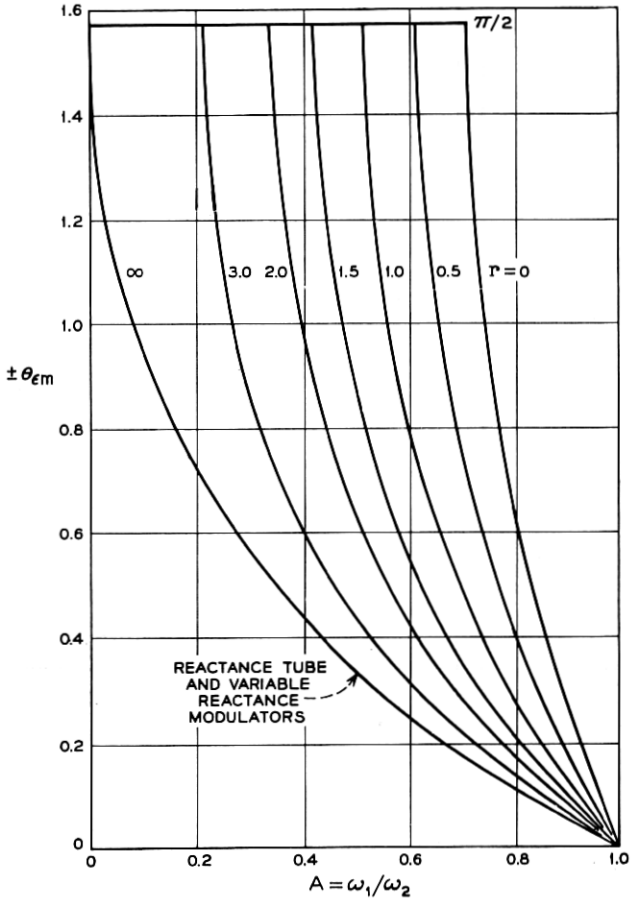

Fig. 6 — Peak phase error for the LC modulators; r = (bit length)/ (dc time constant).

Noting that the phase error never exceeds $\pi/2$ radians, the associated time error, or jitter, is determined from (3):

$$\frac{\tau_{\epsilon m}}{T_o} = \pm \frac{1}{4\pi} \frac{1 + A}{1 - A} \left(\tan^{-1} \frac{1}{\sqrt{A}} - \tan^{-1} \sqrt{A} \right). \tag{15}$$

For infinite frequency shift, this gives a jitter of one-eighth of a cycle of the midfrequency $(1/T_o)$. For zero frequency shift, the time jitter is, interestingly enough, not zero. Rather, (15) converges to a value of jitter equal to $1/4\pi$ of a cycle of the midfrequency. An insight into the reason for this may be obtained by referring to Fig. 2. Note that, as the two frequencies approach each other, the slopes of the phase loci also approach each other. Thus, for a fixed θ_ϵ, τ_ϵ increases as A approaches one. In the case just studied, θ_ϵ approaches zero as A approaches one, but evidently at about the same rate as τ_ϵ increases under the same condition.

Equation (15) is plotted as the lowest curve in Fig. 7(a). Note that the jitter measured in terms of the period of the midfrequency does not change appreciably with the magnitude of the frequency shift for these types of oscillators.

One final point is to be noted. From (15), it is seen that the time error has equal positive and negative excursions. Furthermore, (15) is symmetric in A; that is, replacing A by $1/A$ does not change the expression. Thus the opposite transition in which $\omega_2 < \omega_1$ has a time jitter function identical to (15).

VI. THE SWITCHED REACTANCE MODULATORS

A switched reactance modulator is one in which the resonant frequency of the tank is changed by physically switching an additional reactance in and out. Only those modulators consisting of a simple LC tank with a single additional switched reactance will be considered. There are four possible ways of doing this, leading to the four modulator models shown in Fig. 8. The modulators of Figs. 8 (a) and 8 (b) are clearly duals, and will be referred to as Type I modulators. The modulators of Figs. 8 (c) and 8 (d) are also duals, and will be referred to as Type II modulators. Because of the duality property, only one modulator of each type need be analyzed. It will be seen that all four types give rise to identical jitter expressions.

Let us first analyze the Type I modulator by considering the modulator of Fig. 8 (a). The case of switch closure is treated simply by noting that, since no current flows in the switched inductance prior to closure, the inductance value is changed (reduced) by the switch closure without

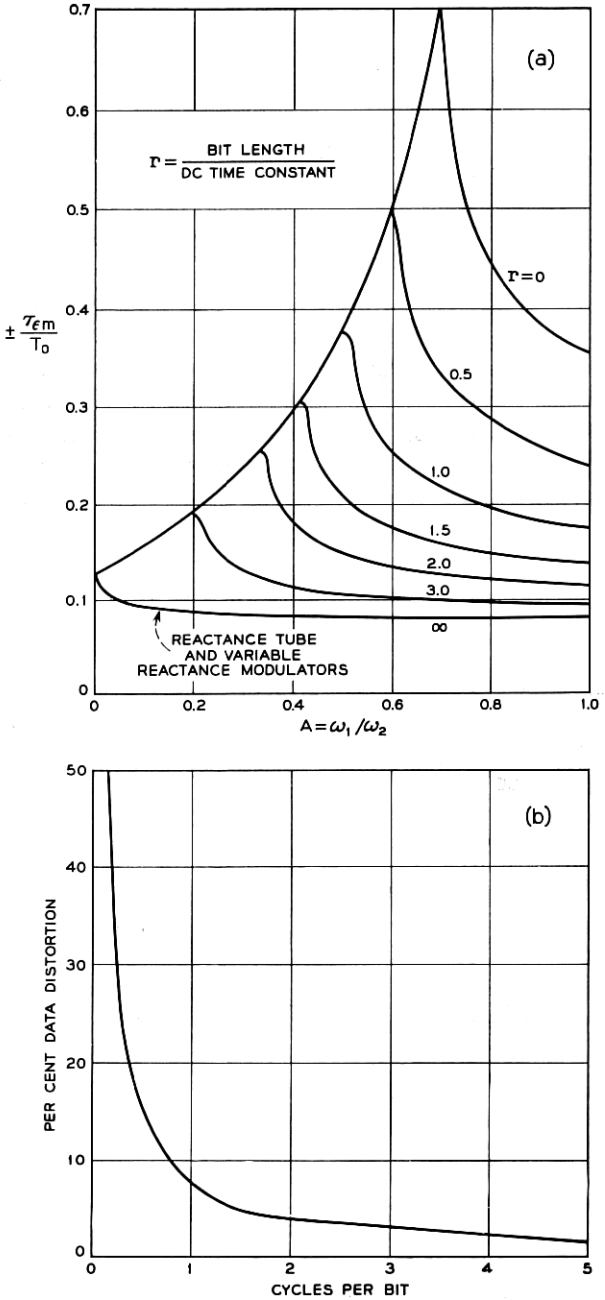

Fig. 7 — (a) Peak jitter for the LC modulators; (b) minimum peak jitter for the LC modulators.

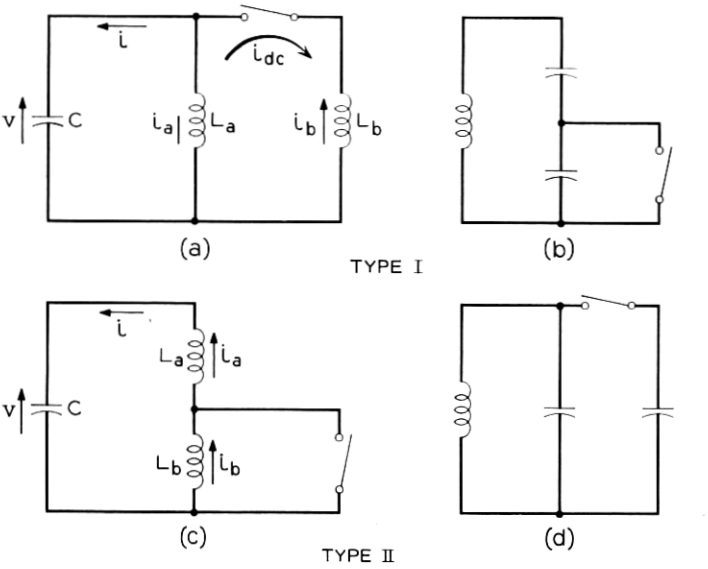

Fig. 8 — Switched reactance modulators.

disturbing the instantaneous voltage or current in the tank (the inductor looks like an open circuit at the time of switch closure). This is exactly the case considered in the previous section. Thus, (14) and (15) and the resulting curves of Figs. 6 and 7 describe the maximum phase and time errors, respectively, for this case.

In Fig. 8(b), the opening of the switch is the dual to the above case, since the switched capacitor initially looks like a short circuit. Thus the capacitance is changed without affecting the instantaneous values of the voltage and current.

In the case of switch opening in Fig. 8(a), current continuity in the capacitor is not preserved, since the current is suddenly reduced from the sum of the currents flowing in both inductors, $i_a + i_b$, to just that current flowing in the permanent tank inductor, i_a. The case of switch opening is furthermore complicated by the fact that a dc current may be flowing around the inductor loop just prior to switch opening. This may be seen more clearly from Fig. 9, in which is plotted the current, i_b, through the switched inductor, L_b. Consider the case shown of switch closure at the instant that the tank voltage is going through zero. The inductor current i_b is the integral of the applied voltage, and is as shown. Note that this current contains a dc component. The case

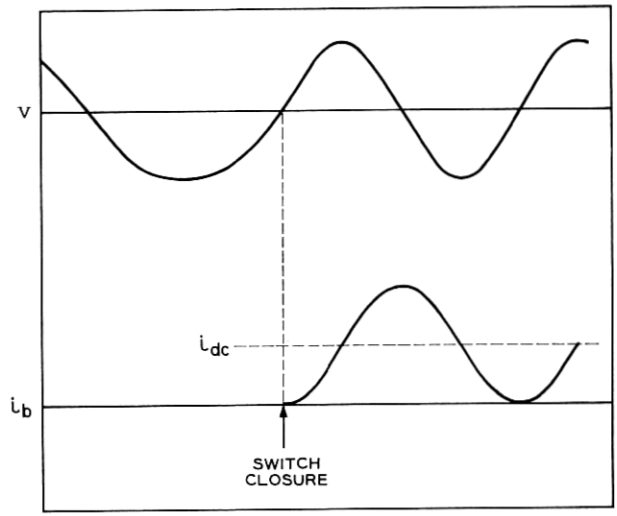

Fig. 9 — The dc component in a switched reactance modulator.

of Fig. 9 is the worst case; if the switch had been thrown at a voltage maximum, then the dc current component would have been zero.

This dc component will die exponentially with a time constant determined by the inductive loop. It does not affect the operation of the tank circuit as long as the switch is closed. However, whatever dc current remains is added to the current i [see Fig. 8(a)] at the time of switch opening. Thus the tank current i just before switch opening is given by $i_a + i_b$, and just after switch opening by $i_a + i_{dc}$. The dc current may be such as to actually reverse the tank current at the instant of switching. Thus large phase errors and hence large time errors may be expected.

Since the phase error will depend upon the magnitude of the dc component, which in turn involves the phase at which the switch was closed, then the phase error at switch opening depends on the phase at the time the switch is opened as well as the phase at the time the switch was previously closed. For purposes of the following analysis, let the following subscripts apply:

1 — just before switch closure.
2 — just after switch closure.
3 — just before switch opening.
4 — just after switch opening.

Then

$$L_1 = L_4 = L_a . \tag{16}$$

$$L_2 = L_3 = \frac{L_a L_b}{L_a + L_b} . \tag{17}$$

$$A^2 = \left(\frac{\omega_1}{\omega_2}\right)^2 = \frac{L_2}{L_1} = \frac{L_b}{L_a + L_b} < 1 . \tag{18}$$

At the time of switch closure, voltage and current are maintained. Thus

$$v_1 = V_1 \cos \theta_1 = v_2 . \tag{19}$$

$$i_1 = V_1 \sqrt{\frac{C}{L_1}} \sin \theta_1 = i_2 . \tag{20}$$

The dc current which is initiated upon switch closure may be determined by a transient analysis or more simply from the energy conditions:

$$E_1 = \tfrac{1}{2} L_1 i_1^2 + \tfrac{1}{2} C v_1^2 .$$

$$E_2 = \tfrac{1}{2} L_2 i_1^2 + \tfrac{1}{2} C v_1^2 .$$

The difference in energy must be due to the circulating dc current, since the initial current in the switched inductor is zero. Denote this energy difference by E_{dc}. Then

$$E_{dc} = \tfrac{1}{2} (L_a + L_b) i_{dc}^2 = E_1 - E_2 = \tfrac{1}{2} (L_1 - L_2) i_1^2 .$$

Making use of (16), (17), and (18), the resulting dc current may be written

$$i_{2dc} = (1 - A^2)\, i_1 = V_1 \sqrt{\frac{C}{L_1}} (1 - A^2) \sin \theta_1 . \tag{21}$$

The tank voltage is continuous during both switch closure and switch opening. Since the current is also continuous during switch closure (assuming zero initial current in L_b), the amplitude of oscillation after switch closure is, from (6), (19), and (20),

$$V_2 = V_3 = \left(\frac{L_2}{C} i_1^2 + v_1^2\right)^{\frac{1}{2}} = V_1 (A^2 \sin^2 \theta_1 + \cos^2 \theta_1)^{\frac{1}{2}} . \tag{22}$$

The tank voltage at switch opening may then be expressed as

$$v_3 = v_4 = V_3 \cos \theta_3 .$$

Then, from (22),

$$v_4 = V_1(A^2 \sin^2 \theta_1 + \cos^2 \theta_1)^{\frac{1}{2}} \cos \theta_3. \tag{23}$$

From (5), (7), and (22), the tank current immediately before switch opening is

$$i_3 = V_3 \sqrt{\frac{C}{L_2}} \sin \theta_3.$$

At this instant the share of the tank current through the permanent tank inductor is

$$i_{3a} = \frac{L_b}{L_a + L_b} i_3 = V_1 \sqrt{\frac{C}{L_2}} A^2 (A^2 \sin^2 \theta_1 + \cos^2 \theta_1)^{\frac{1}{2}} \sin \theta_3, \tag{24}$$

where (22) was used.

If the sum of the resistances of the two inductors is R, and the minimum time between switch closing and switch opening is τ_b (normally the bit length), then the dc current just before switch opening is at most

$$i_{3\text{dc}} = \exp\left(-\frac{R\tau_b}{L_a + L_b}\right) i_{2\text{dc}} = Di_{2\text{dc}},$$

where $D = \exp[-(R\tau_b)/(L_a + L_b)]$ will be called the dissipation parameter.

The actual current through the permanent inductor is the sum of its share of the tank current and the dc current. Immediately after switch opening, this current must be maintained, so that

$$i_4 = i_{3a} + Di_{2\text{dc}}.$$

From (21) and (24),

$$i_4 = V_1 \sqrt{\frac{C}{L_1}} [A(A^2 \sin^2 \theta_1 + \cos^2 \theta_1)^{\frac{1}{2}} \sin \theta_3 + D(1 - A^2) \sin \theta_1]. \tag{25}$$

The current and voltage immediately after switch opening have now been determined in terms of the phase angles at which the switch was initially closed (θ_1) and the phase angle at which it was then reopened (θ_3). These expressions are given by (23) and (25). Using (8), the phase angle immediately after switch opening may now be determined by the relation

$$\theta_4 = \tan^{-1} \sqrt{\frac{L_1}{C}} \frac{i_4}{v_4}$$

$$= \tan^{-1} \frac{A(A^2 \sin^2 \theta_1 + \cos^2 \theta_1)^{\frac{1}{2}} \sin \theta_3 + D(1 - A^2) \sin \theta_1}{(A^2 \sin^2 \theta_1 + \cos \theta_1)^{\frac{1}{2}} \cos \theta_3}.$$

The phase error at switch opening is then

$$\theta_\epsilon = \theta_3 - \theta_4$$

$$= \theta_3 - \tan^{-1}\left[A \tan \theta_3 + \frac{D(1 - A^2) \sin \theta_1}{(A^2 \sin^2 \theta_1 + \cos^2 \theta_1)^{\frac{1}{2}} \cos \theta_3}\right]. \tag{26}$$

We desire to find those values of θ_1 and θ_3 which maximize this function. These values will yield the peak phase error and thus the peak time jitter. Denote these values by θ_{1m} and θ_{3m}, respectively. It is shown in Appendix A that for $A_p \leqq A \leqq 1$,

$$\theta_{1m} = \pm \frac{\pi}{2}$$

and

$$\theta_{3m} = \pm \sin^{-1} \frac{1}{2A}\left[D(2A - 1) - \left(\frac{4(1 - D^2)A^2 + D^2(A + 5)}{A + 1}\right)^{\frac{1}{2}}\right]. \tag{27}$$

where A_p is the value of A for which $\theta_{3m} = \pm \pi/2$. It is further shown in Appendix A that, for $D = 0$, $A_p = 1/\sqrt{2}$; for $D = 1$, $A_p = 0$. For $0 \leqq A \leqq A_p$, the phase error may take on continuous values in excess of $\pi/2$ radians. Thus it can always take on the worst possible value, $\pi/2$. Hence

$$\theta_{\epsilon m} = \pm \theta_{3m} + \tan^{-1} \frac{A}{\cos \theta_{3m}}\left[\sin \theta_{3m} - D\left(\frac{1 - A^2}{A^2}\right)\right], \quad A_p \leqq A \leqq 1$$

$$\theta_{\epsilon m} = \pm \pi/2, \quad 0 \leqq A \leqq A_p,$$

where θ_{3m} is given by (27). Again, the peak phase error, and thus the peak jitter, is symmetric about the true value; that is, these functions may assume equal positive or negative value.

These phase error functions are plotted in Fig. 6, and the associated peak jitter functions as determined from (3) in Fig. 7(a), for various

values of the dissipation factor D. These curves were calculated with the aid of the IBM 7090 digital computer. The curve parameters in Figs. 6 and 7(a) are actually

$$r = -\ln D = \frac{R\tau_b}{L_a + L_b},$$

which is the ratio of the bit length, τ_b, to the time constant of the inductor loop. Note that the peak jitter is considerably greater for the case of small r (large dc currents), than for the case when the time constant is sufficiently short so that the dc current may die in one bit interval. For the case of zero dc current at switch opening ($D = 0$, or $r = \infty$) the quantity θ_{3m} becomes

$$\theta_{3m} = \pm \sin^{-1} \frac{1}{\sqrt{A+1}} = \pm \tan^{-1} \frac{1}{\sqrt{A}}.$$

Then the peak phase error is

$$\theta_{\epsilon m} = \pm\left(\tan^{-1} \frac{1}{\sqrt{A}} - \tan^{-1} \sqrt{A}\right).$$

But this is exactly the form of the peak phase error for the case of the variable reactance oscillator and reactance tube modulator. It may seem strange that the two classes of modulators should have the same peak jitter, since in one case a tank reactance is changed without affecting the instantaneous values of the currents and voltages, whereas in the other case a tank reactance is also changed, but a current discontinuity occurs (from i_3 to i_{3a} in the previous notation). However, a detailed examination of the equations will show that the phase error due to the current discontinuity is in such a direction as to partially compensate for the amplitude change, the result being a peak phase error equal to that resulting from the amplitude discontinuity alone.

As stated before, the modulator of Fig. 8(b) is the dual of the modulator just analyzed, and the same peak phase error and jitter equations therefore result. A switch opening in Fig. 8(b) is equivalent to a switch closure in Fig. 8(a). Analogous to the dc current in Fig. 8(a), a dc voltage may appear across the switched capacitor, and is balanced out by an equal and opposite voltage on the permanent tank capacitor. Then, upon reclosing the switch in Fig. 8(b), the tank voltage suddenly changes to the difference between the ac and dc voltages across the permanent tank capacitor, in complete analogy to the first modulator. Again, this troublesome dc voltage may be bled off by a resistor in parallel with each capacitor, leading again to the dissipation parameter.

It now remains to show that the modulators of Figs. 8(c) and 8(d) lead to the same phase error relationships as derived for the first two modulators. Consider the modulator of Fig. 8(c). The case of the switch closure is identical to the conditions for the variable reactance oscillator and reactance tube modulator; i.e., the value of the inductance is changed, but the instantaneous values of the tank current and voltage are unchanged. Hence the lowest curves for peak phase error and peak jitter in Figs. 6 and 7(a) hold for this case. However, the switched inductor, L_b, will have a current flowing through it just prior to switch closure, and will retain this current after the switch has closed. L_b may begin to discharge due to series resistance, but in general will have a nonzero current flowing through it at the time of switch opening. At switch opening, the flux in L_a and L_b will redistribute so that the currents through the two inductors are equalized, thus causing a current transient.

Using the same subscripts as before, define

$$L_1 = L_4 = L_a + L_b.$$

$$L_2 = L_3 = L_a.$$

$$A^2 = \left(\frac{\omega_1}{\omega_2}\right)^2 = \frac{L_2}{L_1} = \frac{L_a}{L_a + L_b} < 1.$$

Immediately after switch closure, the dc current flowing through L_b is

$$i_{2b} = i_{2dc} = i_1 = V_1 \sqrt{\frac{C}{L_1}} \sin \theta_1.$$

Just before switch opening, the dc current flowing through L_b is

$$i_{3dc} = Di_{2dc} = V_1 D \sqrt{\frac{C}{L_1}} \sin \theta_1,$$

where the dissipation factor, D, is now given by

$$D = \exp\left(-\frac{R\tau_b}{L_b}\right),$$

where R is the loop resistance containing L_b and the closed switch. The value of the current through L_a just before switch opening is

$$i_{3a} = V_3 \sqrt{\frac{C}{L_2}} \sin \theta_3 = V_1 \sqrt{\frac{C}{L_2}} (A^2 \sin^2 \theta_1 + \cos^2 \theta_1)^{\frac{1}{2}} \sin \theta_3,$$

where V_3 is determined exactly as in the previous case.

At the instant of switch opening, one is faced with the problem of two

inductors with different initial currents being connected in series. The series current must instantly adjust to some common value. To determine this common value, note that the tank voltage v must be continuous. This is true because there is no process by which an impulse in the tank current may be caused; thus the voltage across the capacitor and hence the tank voltage cannot change instantaneously.

Since the voltage across the two inductors is the same immediately before as after the switch opening, then the total flux stored in the two inductors must be preserved. The instantaneous redistribution of flux between the two inductors will cause a voltage impulse to appear across each inductor, but these impulses will be of equal magnitude and opposite sign, thus canceling. The total stored flux immediately before the switch opening is

$$\varphi_3 = L_a i_{3a} + L_b i_{3dc}.$$

Immediately after the switch opening, the stored flux is

$$\varphi_4 = (L_a + L_b) i_4 .$$

Equating φ_3 and φ_4, one obtains

$$\begin{aligned} i_4 &= A^2 i_{3a} + (1 - A^2) i_{3dc} \\ &= V_1 \sqrt{\frac{C}{L_1}} [A(A^2 \sin^2 \theta_1 + \cos^2 \theta_1)^{\frac{1}{2}} \sin \theta_3 + D(1 - A^2) \sin \theta_1]. \end{aligned} \tag{28}$$

The voltage immediately after switch opening is again

$$v_4 = V_3 \cos \theta_3 = V_1(A^2 \sin^2 \theta_1 + \cos^2 \theta_1) \cos \theta_3. \tag{29}$$

Equations (28) and (29) are identical to (25) and (23) respectively for the case of the Type I modulator. Thus the peak phase and jitter functions for the Type II modulator are identical to those for the Type I modulator [Figs. 6 and 7(a)], even though the waveform discontinuities are caused by an entirely different phenomenon.

The circuit of Fig. 8(d) is of course the dual to that of Fig. 8(c). Switch closure in one corresponds to switch opening in the other, and vice versa. In the circuit of Fig. 8(d), when the switch is opened, a voltage is left on the switched capacitor. It may decay through a leakage resistance, but upon switch closure the two different voltages on the two capacitors must instantly adjust to some common voltage according to the condition of charge conservation, in complete analogy to the circuits of Fig. 8(c).

Note that the jitter characteristic for all of the switched reactance mod-

ulators approaches that of the variable reactance-type modulators for low bit rates. This latter curve is almost a constant ($\tau_{em}/T_o = 1/4\pi$) over the range of A which might be commonly used. This then is a useful rule of thumb: In a modulator whose frequency is shifted by changing one reactance in the tank of an oscillator, the minimum inherent peak jitter is roughly $1/4\pi$ of the period of the midfrequency. This rule is shown plotted in Fig. 7(b) in a somewhat different form. The abscissa is the number of cycles of the midfrequency per bit:

$$\text{cycles/bit} = \frac{\tau_b}{T_o}.$$

The ordinate is the per cent data distortion, defined by

$$\text{distortion} = \frac{\tau_{em}}{\tau_b} = \frac{1}{4\pi}\frac{T_o}{\tau_b}.$$

Thus

$$(\text{distortion})(\text{cycles/bit}) = \frac{1}{4\pi}.$$

VII. THE MULTIVIBRATOR MODULATOR

A multivibrator type of modulator, using transistors as an example, is shown in Fig. 10(a). It consists of a standard symmetric multivibrator whose frequency may be controlled by varying the control voltage V_c. Often a low-pass filter follows the multivibrator. The function of this filter is to eliminate the harmonics of the multivibrator output, thus converting the square-wave output to a sine wave. This filter may add additional zero-crossing distortion, but this distortion will not be considered here. Only the zero-crossing distortion inherent in the multivibrator itself will be analyzed.

The analysis will proceed in the following manner. It will be assumed that the multivibrator is oscillating with some half-period η_1. At some time t_o, measured with respect to the last multivibrator transition [see Fig. 11(a)], the control voltage V_c is suddenly changed from its initial value V_{c1} to a new value V_{c2}. The multivibrator then oscillates with some new half-period η_2. The time between the two transitions on either side of t_o will be denoted η_o where η_o has a value between η_1 and η_2.

The resulting zero crossings will be compared to those of an ideal FM wave shown in Fig. 11(b) whose zero crossings correspond to those of the multivibrator before the switching instant, but are separated from those of the multivibrator by a time error of T_ϵ for times after the switch-

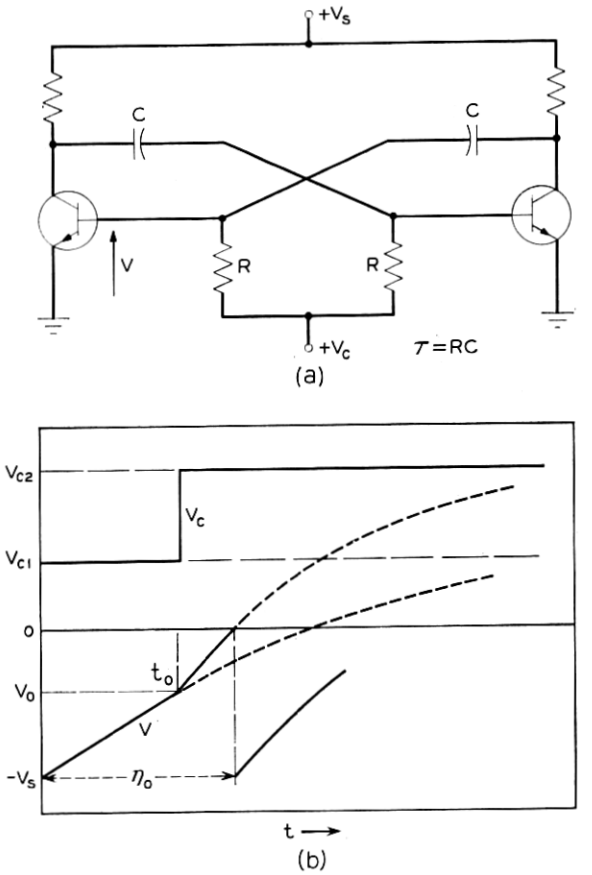

Fig. 10 — The multivibrator modulator: (a) multivibrator circuit; (b) base voltage waveform.

ing instant. This corresponds to a phase error of

$$\theta_\epsilon = 2\pi \frac{T_\epsilon}{2\eta_2}. \tag{30}$$

The worst value for the transition time will be determined, leading to a maximum value for θ_ϵ, denoted $\theta_{\epsilon m}$. From (3), the peak time error $\tau_{\epsilon m}$ in the data transition with respect to the period of the midfrequency T_o is

$$\frac{\tau_{\epsilon m}}{T_o} = \frac{1}{2}\frac{1+A}{1-A}\frac{T_{\epsilon m}}{2\eta_2}. \tag{31}$$

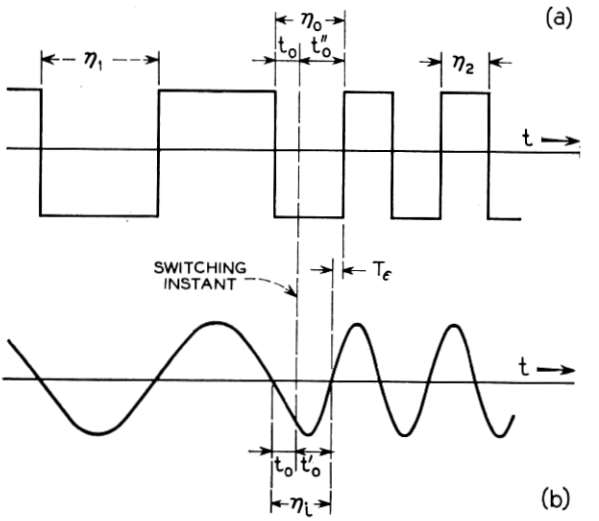

Fig. 11 — Comparing zero crossings of multivibrator modulator to those of ideal FM: (a) multivibrator outputs; (b) ideal FM.

Let us first derive the time between zero crossings on either side of the switching instant for an ideal FM wave. This time will be called η_i. From Fig. 11(b),

$$\eta_i = t_o + t_o'.$$

Because the amplitudes of the two sinusoids are the same, and the voltage is continuous at the transition, one may write

$$\sin \omega_1 t_o = \sin \omega_2(\eta_2 - t_o').$$

Thus

$$t_o' = \eta_2 \left(1 - \frac{t_o}{\eta_1}\right),$$

and

$$\eta_i = \eta_2 + t_o \left(1 - \frac{\eta_2}{\eta_1}\right). \tag{32}$$

We now derive the analogous time, η_0, for the multivibrator. In Fig. 10(b) is shown the operation of the multivibrator in terms of the voltage at a base. Immediately after a transition, the transistor is reverse biased by a voltage equal to the supply voltage, V_s. The voltage at this point

will start rising exponentially to the control voltage V_c, with a time constant τ, and is given by

$$v = V_c - (V_s + V_c)e^{-t/\tau},$$

where τ is defined in Fig. 10 as $\tau = RC$. For a constant control voltage, the half-period of oscillation, η, is determined by that time required for the voltage waveform to reach zero voltage (ignoring the transistor threshold voltage). Thus

$$\eta = \tau \ln\left(1 + \frac{V_s}{V_c}\right).$$

Let us define the following parameters:

$$A = \frac{\omega_1}{\omega_2} = \frac{T_2}{T_1} = \frac{\eta_2}{\eta_1}.$$

$$\alpha = \frac{V_s}{V_{c1}}.$$

$$\beta = \frac{V_s}{V_{c2}}.$$

Then

$$\eta_1 = \tau \ln(1 + \alpha),$$

$$\eta_2 = \tau \ln(1 + \beta).$$

At a time t_o after the last transition, there will be a base voltage v_o of

$$v_o = V_{c1} - (V_s + V_{c1}) \exp(-t_o/\tau).$$

If at this time the multivibrator control voltage is suddenly switched from V_{c1} to V_{c2}, the time from t_o to the next transition will be

$$t_o'' = \tau \ln\left(1 - \frac{v_o}{V_{c2}}\right) = \tau \ln\left[1 - \frac{V_{c1}}{V_{c2}} + \frac{V_{c1}}{V_{c2}}\left(1 + \frac{V_s}{V_{c1}}\right) \exp\left(-t_o/\tau\right)\right].$$

Then

$$\eta_o = t_o + t_o'' = t_o + \tau \ln\left[1 - \frac{\beta}{\alpha} + \frac{\beta}{\alpha}\left(1 + \alpha \exp\left(-t_o/\tau\right)\right)\right].$$

The difference between this and the like time, η_i for ideal FM is

$$T_\epsilon = \eta_o - \eta_i = \tau \ln\left[1 - \frac{\beta}{\alpha} + \frac{\beta}{\alpha}(1 + \alpha) \exp\left(-t_o/\tau\right)\right] + At_o - \eta_2. \tag{33}$$

From (30), θ_ϵ is maximized if T_ϵ is maximized. The value of t_o, denoted t_{om}, for which T_ϵ is a maximum is found by letting

$$\frac{\partial T_\epsilon}{\partial T_o} = \frac{\tau}{1 - \frac{\beta}{\alpha} + \frac{\beta}{\alpha}(1+\alpha)\exp(-t_{om}/\tau)} \frac{\beta}{\alpha}(1+\alpha)\left(-\frac{1}{\tau}\right)\exp(-t_{om}/\tau) + A = 0,$$

which results in

$$\exp(-t_{om}/\tau) = \frac{\left(1 - \frac{\beta}{\alpha}\right)}{\frac{\beta}{\alpha}(1+\alpha)\left(\frac{1}{A} - 1\right)}$$

or

$$t_{om} = \tau \ln \frac{(1+\alpha)(1-A)}{A\left(\frac{\alpha}{\beta} - 1\right)}. \tag{34}$$

Substituting (34) into (33), one obtains

$$T_{\epsilon m} = \tau \ln \frac{1 - \frac{\beta}{\alpha}}{1 - A} + A\tau \ln \frac{(1+\alpha)(1-A)}{A\left(\frac{\alpha}{\beta} - 1\right)} - \eta_2,$$

which can be rewritten as

$$T_{\epsilon m} = \tau\left[(1-A)\ln \frac{1 - \frac{\beta}{\alpha}}{1-A} + A \ln \frac{\beta}{A\alpha} + A \ln (1+\alpha) - \frac{\eta_2}{\tau}\right]. \tag{35}$$

But, since

$$\eta_2 = \tau \ln (1+\beta)$$

and

$$A = \frac{\ln (1+\beta)}{\ln (1+\alpha)}, \tag{36}$$

then

$$\begin{aligned} \alpha &= (1+\beta)^{1/A} - 1, \\ \tau &= \frac{\eta_2}{A \ln (1+\alpha)}, \end{aligned} \tag{37}$$

and

$$T_{em} = \frac{\eta_2}{A \ln (1 + \alpha)} \left[(1 - A) \ln \frac{1 - \frac{\beta}{\alpha}}{1 - A} + A \ln \frac{\beta}{A\alpha} \right],$$

or

$$\theta_{em} = 2\pi \frac{T_{em}}{2\eta_2} = 2\pi \frac{\left(\frac{1 - A}{A}\right) \ln \frac{1 - \frac{\beta}{\alpha}}{1 - A} + \ln \frac{\beta}{A\alpha}}{2 \ln (1 + \alpha)}. \tag{38}$$

The peak time jitter is then determined from (3):

$$\frac{\tau_{em}}{T_o} = \frac{1}{4} \frac{(1 + A)}{(1 - A)} \frac{\left(\frac{1 - A}{A}\right) \ln \frac{1 - \frac{\beta}{\alpha}}{1 - A} + \ln \frac{\beta}{A\alpha}}{\ln (1 + \alpha)}. \tag{39}$$

Equations (37) and (38) give the peak phase error for a transition in either direction (i.e., from the lower to the higher frequency, or vice versa) as a function of the two variables A and β. It is shown in Appendix B that the peak phase error for the transition in one direction is equal in magnitude but opposite in sign to the peak phase error for the opposite transition.

From the relation (3), it is then seen that the peak time error for the multivibrator is equal in magnitude and is always negative for either transition. Thus the multivibrator differs in its jitter characteristics from the LC modulators in that the jitter is one-sided, rather than symmetric, about the true transition time. The jitter is such that the actual transitions are delayed in time from the true transition ($-\tau_{em}$ corresponds to a time retardation).

The parameter β is essentially a "linearity" parameter; i.e., as $\beta \to 0$, less and less of the exponential of Fig. 10 is used. For α and β sufficiently small, the base voltage is essentially linear. In this case, from (37),

$$\beta \approx \alpha A.$$

Under this condition, (39) goes to zero. Thus the multivibrator type of modulator may be designed to have as little jitter as desired by making the control voltage, V_c, very much greater than the supply voltage, V_s.

It is also of interest to study the other limiting condition, i.e., $\beta \to \infty$. From (37), for $\beta \gg 1$,

$$\alpha \approx \beta^{\frac{1}{2}}.$$

Substituting this into (38) for the peak phase error, and taking the limit as $\beta \to \infty$, one finds

$$\lim_{\beta \to \infty} \theta_{\epsilon m} = -(1 - A)\pi,$$

for $\frac{1}{2} \leqq A \leqq 1$, $| \theta_{\epsilon m} | \leqq \pi/2$; therefore, from (39),

$$\lim_{\beta \to \infty} \frac{\tau_{\epsilon m}}{T_o} = -\frac{1 + A}{4}, \quad \frac{1}{2} \leqq A \leqq 1. \tag{40}$$

However, for $0 \leqq A \leqq \frac{1}{2}$, θ_ϵ may take on values in excess of $\pi/2$. Thus

$$\lim_{\beta \to \infty} \theta_\epsilon = -\frac{\pi}{2}$$

$$0 \leqq A \leqq \frac{1}{2}.$$

$$\lim_{\beta \to \infty} \frac{\tau_{\epsilon m}}{T_o} = -\frac{1}{8}\frac{1 + A}{1 - A}$$

Plots of $\theta_{\epsilon m}$ and $\tau_{\epsilon m}/T_o$ for the multivibrator are given in Figs. 12 and 13 respectively. Fig. 13(b) shows peak jitter on an expanded scale for small values. For comparison, the corresponding curve for the peak jitter inherent in the reactance tube type of modulator is shown dotted in Figs. 13(a) and 13(b).

Note that a given time error for the multivibrator corresponds to only half the distortion caused by the same time error associated with one of the LC modulators because of the one-sided property of the multivibrator jitter.

VIII. EXPERIMENTAL VERIFICATION

The results of the theory just presented were checked by measuring the peak jitter in some existing systems. In Table I are shown some experimental and theoretical results for the modulators of the Bell System DATA-PHONE Data Sets 101A, 103A, and 202A.[3] In all cases, a dotting (alternate marking and spacing) signal was used to avoid the effects of intersymbol interference. The receiver used for the 100 series data sets was the 101A receiver. It contains a limiter followed by a standard discriminator and slicing amplifier. Because of the limiter, the zero-crossing information is all that is retained. The 202A receiver was used with its modulator. This receiver is indeed a zero-crossing detector. It generates a pulse for each zero crossing, integrates the resulting pulse train, and delivers this resulting signal to a slicing circuit.

The 100 series data sets are capable of operating in either of two bands. These bands are denoted the f_1 band and the f_2 band. In the case of the

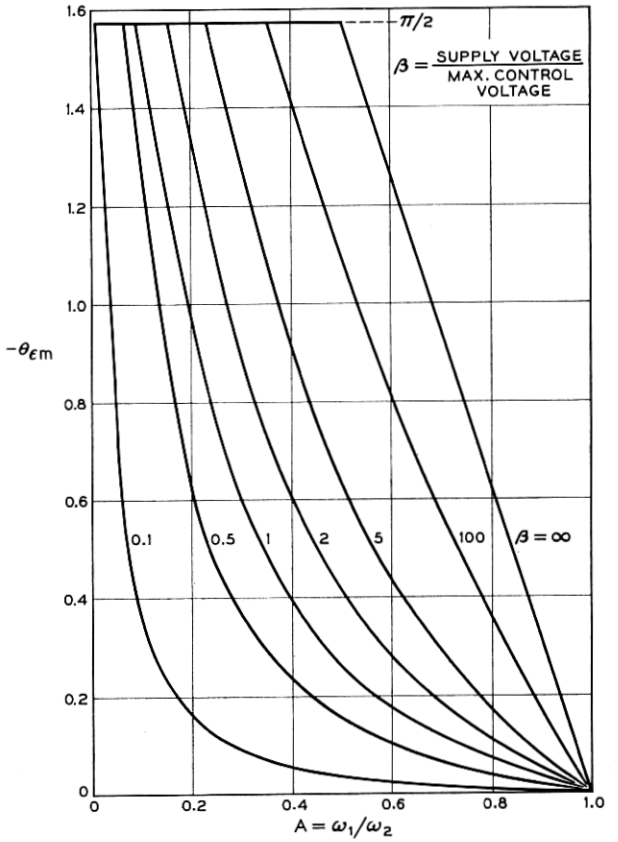

Fig. 12 — Peak phase error for the multivibrator modulator.

TABLE I — EXPERIMENTAL VERIFICATION

Modulator	Type	ω_1	ω_2	Bit Rate	Transition	Experimental Peak Jitter (μsec)	Theoretical Peak Jitter (μsec)
101	Switched Reactance	1070	1270	100	$\omega_1 \rightarrow \omega_2$	165	136
					$\omega_2 \rightarrow \omega_1$	280	280
		2025	2225	100	$\omega_1 \rightarrow \omega_2$	100	72
					$\omega_2 \rightarrow \omega_1$	130	112
103	Reactance Transistor	1070	1270	100	Either	165	136
				200	Either	165	136
202	Multivibrator	1200	2200	1200	Either	50	44.5

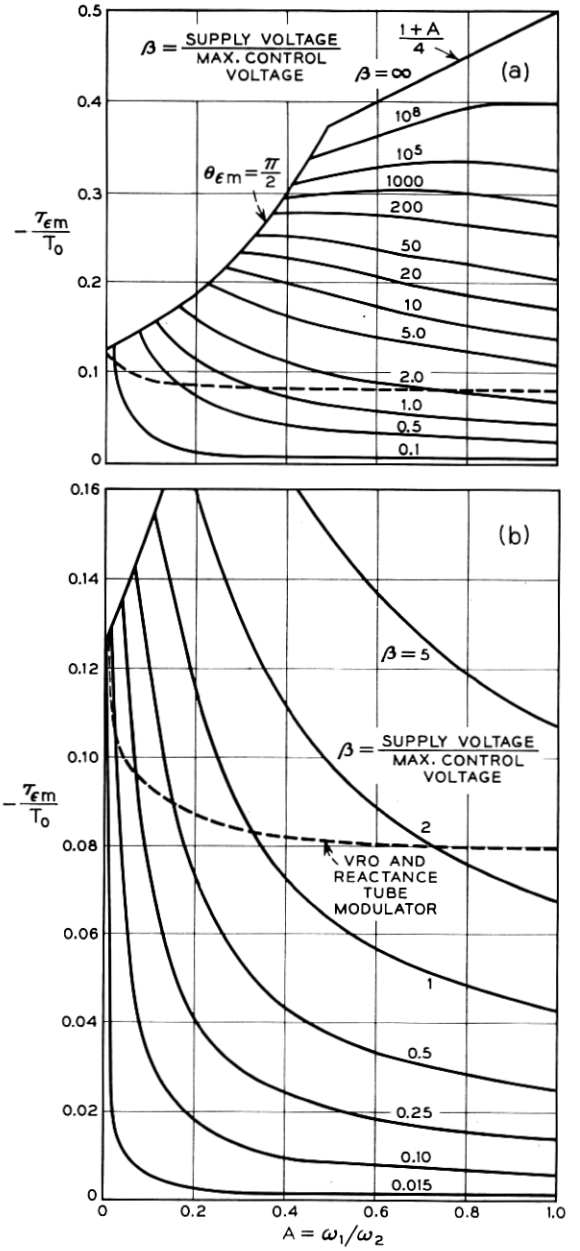

Fig. 13 — (a) Peak jitter for the multivibrator modulator; (b) peak jitter for the multivibrator modulator — expanded scale.

101A modulator, both the f_1 band and the f_2 band were used for the test. The marking and spacing frequencies in the f_1 band are 1070 and 1270 cycles, resulting in an A of 0.843. In the f_2 band, the corresponding frequencies are 2025 and 2225 cps, giving an A of 0.910. The modulator is a switched reactance modulator of the type shown in Fig. 8(a). For the f_1 band, the sum of the two inductances is 0.324 henry and the total series resistance is 44.5 ohms. For the f_2 band these parameters are 0.151 henry and 28.7 ohms. This modulator was tested at 100 bits per second, the corresponding values of r being 1.37 and 1.90 for the f_1 band and f_2 band respectively.

The 103A modulator was tested at 100 and 200 bits per second using the f_1 band and again the 101A demodulator. The 103A modulator is of the reactance transistor type (analogous to the reactance tube modulator). The results agree well with theory.

The 202A modulator was tested at 1200 bits per second, using 1200 cps and 2200 cps as the two frequencies. This corresponds to an A of 0.54. In this modulator, the value of β is 1.405 ($V_{c2} = 12.8$ v, $V_s = 18$ v). Again, the agreement with theory is quite good.

Because all of the above experiments yielded jitter comparable to that expected from the modulator alone, one is led to the conclusion that most of the back-to-back jitter in these data sets originates in the modulator, and that the demodulators approach fairly closely the ideal zero-crossing detector.

IX. DISTORTIONLESS, ASYNCHRONOUS, SINUSOIDAL FREQUENCY MODULATION

From (8), it is possible to obtain the conditions for distortionless modulation when the modulation is accomplished by changing the parameters of an LC tank. Let the subscript 1 refer to the instant before the transition, and the subscript 2 to the instant after. There will be no jitter if there is no phase error, that is, if

$$\theta_1 = \theta_2 .$$

From (8), this requires that

$$\sqrt{\frac{L_1}{C_1}}\frac{i_1}{v_1} = \sqrt{\frac{L_2}{C_2}}\frac{i_2}{v_2}.$$

One way of insuring that this condition holds is to maintain voltage and current continuity at the transition ($i_1 = i_2$, $v_1 = v_2$), and then to change both the inductance and capacitance by such an amount that

their ratio remains constant. That is, if the resonant frequency of a tank circuit is changed in such a way that the instantaneous tank voltage and current are continuous, and so that the characteristic tank impedance, $\sqrt{L/C}$, is maintained, then there will be no phase error.

Such a modulator may be implemented by using a balanced VRO[1] (Variable Reactance Oscillator) configuration as shown in Fig. 14, in which one active element displays a variable capacitance and negative resistance, and the other a variable inductance and negative resistance. The g_m's of each active element are switched in such a way that the quantity L/C remains constant. The modulator of Fig. 15 was constructed and tested. To give an idea of its performance, typical frequency transitions between 1950 cps and 2250 cps are shown in the oscillogram of Fig. 16.

X. CONCLUSION

It has been shown that phase error at a transition in a frequency shift signal will cause a time error in the output of a zero-crossing detector. The maximum value of this time error may be calculated, leading to the determination of the peak jitter.

It has been further shown that the switched reactance modulators have the largest inherent peak jitter, and that this jitter is bit-rate dependent.

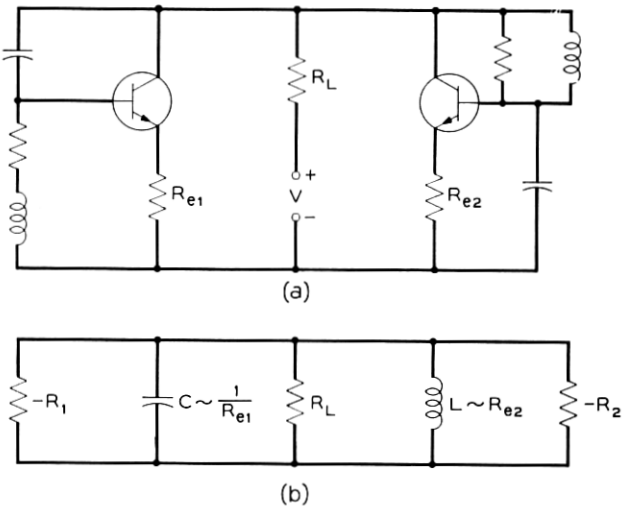

Fig. 14 — The balanced VRO configuration for distortionless frequency modulation: (a) basic balanced VRO; (b) equivalent circuit of basic balanced VRO.

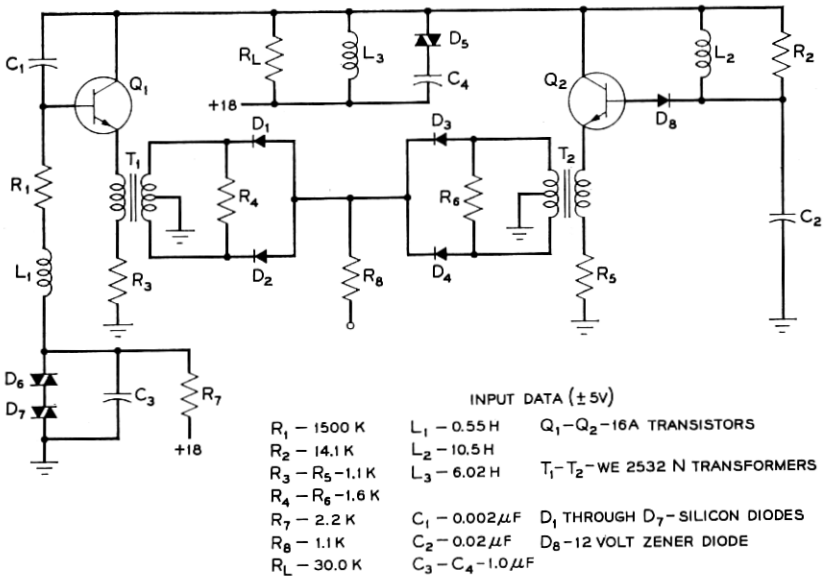

Fig. 15 — Balanced VRO modulator schematic diagram.

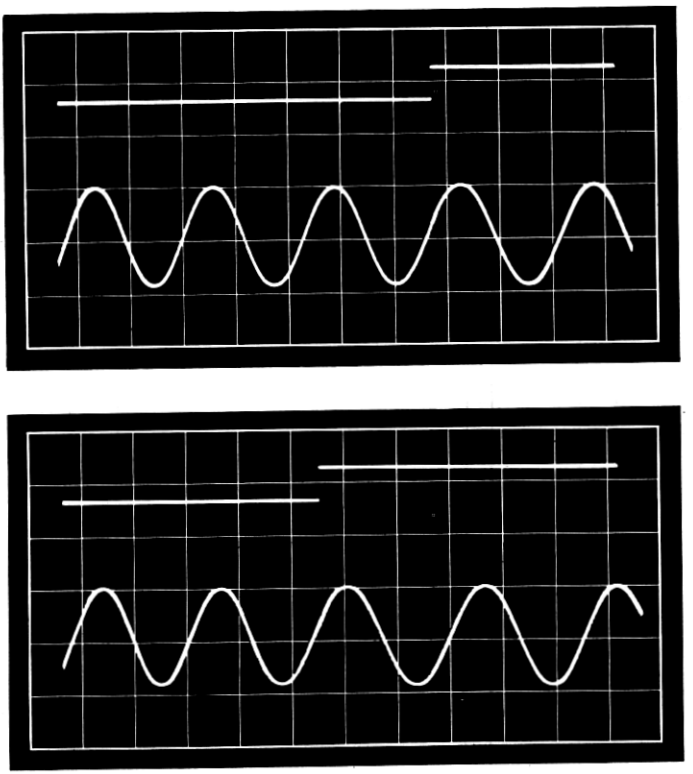

Fig. 16 — Examples of switching performance of balanced VRO: upper trace is data input; lower trace is modulator output; frequencies are 1950 and 2250 cps.

The reactance tube type of modulator and variable reactance oscillator have less peak jitter. The multivibrator can be designed to have as low a value of peak jitter as desired. By using an LC type of modulator in which the tank impedance is held constant at switching, it is possible in principle to design a frequency shift modulator having zero peak jitter.

All of the main points of the theory have been experimentally verified.

XI. ACKNOWLEDGMENT

The authors would like to thank Messrs. J. R. Livingston and D. C. Rife for their suggestion of the model for the zero-crossing detector which made this work possible. They would also like to thank these people and Messrs. J. R. Davey, I. Dorros, S. T. Meyers, and C. W. Farrow for their many helpful discussions.

APPENDIX A

Maximum Phase Error for the Switched Reactance Modulator

The phase error associated with the more severe transition in the switched reactance modulator (e.g., switch opening in the parallel inductor case) is given by (26):

$$\theta_\epsilon = \theta_3 - \tan^{-1}\left[A \tan \theta_3 + \frac{D(1 - A^2) \sin \theta_1}{(A \sin^2 \theta_1 + \cos^2 \theta_1)^{\frac{1}{2}} \cos \theta_3}\right]. \quad (41)$$

It is desired to find those values of θ_1 and θ_3 which maximize this expression. These values will be denoted θ_{1m} and θ_{3m} respectively.

Consider θ_1 first. Differentiating (41) with respect to θ_1 yields

$$\frac{\partial \theta_\epsilon}{\partial \theta_1} = \frac{\gamma \cos \theta_1}{(1 + \beta^2) \cos \theta_3}\left[1 + \frac{(1 - A^2) \sin^2 \theta_1}{A^2 \sin^2 \theta_1 + \cos^2 \theta_1}\right], \quad (42)$$

where

$$\beta = A \tan \theta_3 + \frac{\gamma \sin \theta_1}{\cos \theta_3},$$

and

$$\gamma = \frac{D(1 - A^2)}{(A^2 \sin^2 \theta_1 + \cos^2 \theta_1)^{\frac{1}{2}}}.$$

We are interested in those values of θ for which $\partial \theta_\epsilon / \partial \theta_1 = 0$. Since $A \leqq 1$, all quantities enclosed within the brackets in (42) are always nonnegative; therefore, this term will never become zero. Furthermore,

θ_1 can never cause β to become infinite, nor γ to become zero. Therefore

$$\theta_{1m} = \pm \frac{\pi}{2}$$

is the only value of θ_1 which causes (42) to go to zero. It can be shown that this value of θ_1 corresponds to a maximum point for θ_ϵ. θ_{1m} represents switching at the instant corresponding to the creation of the maximum value for the undesired dc quantity.

Investigation of (42) will reveal that $\theta_3 = \pm\pi/2$ will also cause $d\theta_\epsilon/d\theta_1$ to be zero; however, it can be shown that this corresponds to a minimum point for θ_ϵ. No other combinations of θ_1 and θ_3 will cause (42) to go to zero.

Equation (42) may now be simplified by allowing θ_1 to be $\pm\pi/2$:

$$\theta_\epsilon = \theta_3 - \tan^{-1}\left[A \tan \theta_3 \pm D \frac{(1 - A^2)}{A} \frac{1}{\cos \theta_3}\right]. \tag{43}$$

Then

$$\frac{\partial \theta_\epsilon}{\partial \theta_3} = 1 - \frac{1}{(1 + \beta^2) \cos^2 \theta_3} (A \pm \gamma \sin \theta_3), \tag{44}$$

where β and γ are now given by the simplified expressions

$$\beta = A \tan \theta_3 \pm \frac{\gamma}{\cos \theta_3} \tag{45}$$

$$\gamma = D \frac{(1 - A^2)}{A}. \tag{46}$$

We are interested now in determining those values of θ_3 which cause $\partial\theta_\epsilon/\partial\theta_3$ to be zero. Setting (44) to zero, one obtains

$$A \pm \gamma \sin \theta_{3m} = (1 + \beta^2) \cos^2\theta_{3m}.$$

Using the expression (45) for β, and making the substitution $\cos^2\theta_{3m} = 1 - \sin^2\theta_{3m}$, a quadratic in $\sin^2\theta_{3m}$ is obtained:

$$(A^2 - 1) \sin^2\theta_{3m} \pm \gamma(2A - 1) \sin \theta_{3m} + (\gamma^2 + 1 - A) = 0.$$

Replacing γ with (46) and rearranging, this quadratic becomes

$$\sin^2 \theta_{3m} \mp \frac{D(2A - 1)}{A} \sin \theta_{3m} - \left(\frac{D^2(1 - A^2)(1 + A) + A^2}{A^2(1 + A)}\right) = 0.$$

Thus

$$\sin \theta_{3m} = \pm \frac{D}{2A} (2A - 1) \pm \left[\frac{4A^2 \frac{(1 - D^2)}{D^2} + (A + 5)}{A + 1} \right]^{\frac{1}{2}}. \quad (47)$$

The range of interest for both A and D is from zero to one. Consider the contents of the square brackets in (47). This is minimum for $D = 1$, for which

$$\sin \theta_{3m} |_{D=1} = \pm \frac{1}{2A} \left[(2A - 1) \pm \left(\frac{A + 5}{A + 1} \right)^{\frac{1}{2}} \right].$$

If the plus sign in front of the term $[(A + 5)/A + 1]^{\frac{1}{2}}$ were to be used, then

$$\sin \theta_{3m} | _{D=1} > 1,$$

for all A. Since $\sin \theta_{3m}$ must be less than one, this is not possible. For $D < 1$, the magnitude of $\sin \theta_{3m}$ is even greater, since the term in the square brackets of (47) is now larger. Thus the minus sign must be used, and

$$\sin \theta_{3m} = \pm \frac{D}{2A} (2A - 1) - \left[\frac{4A^2 \frac{(1 - D^2)}{D^2} + (A + 5)}{A + 1} \right]^{\frac{1}{2}}. \quad (48)$$

Now, for any D, and $A = 1$, the magnitude of $(\sin \theta_{3m})$ is less than one. However, as A decreases, $| \sin \theta_{3m} |$ increases, and finally reaches unity for some $A = A_p$. For instance, for $D = 1$ (infinite dc time constant), $\sin \theta_{3m} = \pm 1$ for $A = A_p = 1/\sqrt{2}$; for $D = 0$ (zero dc time constant), $A_p = 0$. For other values of $0 < D < 1$, $0 < A_p < 1/\sqrt{2}$.

Thus for a given D and for $A_p < A \leqq 1$, the magnitude of θ_{3m} is less than $\pi/2$ and thus may be used to calculate the peak phase error by substitution into (43). For $A = A_p$, $\theta_{3m} = \pm \pi/2$, and it can be shown that at this point $\theta_{\epsilon m} = \pm \pi/2$. Furthermore, for $A < A_p$, a more detailed study of the expression for phase error (43) will reveal that the phase error may take on any value. Since it is assumed that $\pm \pi/2$ is the worst possible phase error, then we set $\theta_{\epsilon m} = \pm \pi/2$ for $A < A_p$

Summarizing these results, the peak phase error is given by

$$\theta_{\epsilon m} = \pm \theta_{3m} + \tan^{-1} \frac{A}{\cos \theta_{3m}} \left[\sin \theta_{3m} + D \left(\frac{1 - A^2}{A^2} \right) \right], \quad A_p \leqq A \leqq 1,$$

where

$$\sin \theta_{3m} = \frac{D}{2A}(2A - 1) - \left[\frac{4A^2\frac{(1 - D^2)}{D^2} + (A + 5)}{A + 1}\right]^{\frac{1}{2}},$$

and

$$\theta_{\epsilon m} = \pm\frac{\pi}{2} \quad \text{for} \quad 0 \leqq A \leqq A_p.$$

A_p is the value of A for which $|\sin \theta_{3m}| = 1$.

APPENDIX B

Jitter Symmetry in the Multivibrator Modulator

The expression for peak phase error in the multivibrator is given by (38):

$$\theta_{\epsilon m} = 2\pi \frac{\left(\frac{1 - A}{A}\right)\ln\frac{1 - \frac{\beta}{\alpha}}{1 - A} + \ln\frac{\beta}{A\alpha}}{2\ln(1 + \alpha)}. \tag{49}$$

Unlike the LC modulators, this phase error is of only one sign for any one transition. In fact, from Fig. 12, it is seen that this phase error is always negative for $A < 1$ (i.e., $\omega_1 < \omega_2$, or switching from the lower to the higher frequency).

It is to be shown here that the phase error for the opposite transition is of the same magnitude but of opposite sign (thus $\theta_{\epsilon m} > 0$ for $A > 1$). This is accomplished by replacing in (49) A by $1/A$, α by β, and β by α. This replacement yields an expression which describes the peak phase error, $\theta_{\epsilon m}'$, for the opposite transition:

$$\theta_{\epsilon m}' = 2\pi \frac{(A - 1)\ln\frac{1 - \frac{\alpha}{\beta}}{1 - \frac{1}{A}} + \ln\frac{A\alpha}{\beta}}{2\ln(1 + \beta)}.$$

β is related to α by (37):

$$\ln(1 + \beta) = A\ln(1 + \alpha). \tag{50}$$

Equation (49) may then be rewritten as

$$\theta_{em}' = \frac{\frac{(A-1)}{A}\ln\frac{A\left(\frac{\alpha}{\beta}-1\right)}{1-A} + \frac{1}{A}\ln\frac{A\alpha}{\beta}}{2\ln(1+\alpha)}$$

$$= \frac{-\left(\frac{1-A}{A}\right)\ln\frac{\left(1-\frac{\beta}{\alpha}\right)}{1-A}\cdot\frac{A\left(\frac{\alpha}{\beta}-1\right)}{\left(1-\frac{\beta}{\alpha}\right)} + \frac{1}{A}\ln\frac{A\alpha}{\beta}}{2\ln(1+\alpha)}$$

$$= \frac{-\left(\frac{1-A}{A}\right)\ln\left(\frac{1-\frac{\beta}{\alpha}}{1-A}\right) + \ln\frac{\beta}{A\alpha} + \left(\frac{1-A}{A}\right)\ln\frac{A\alpha}{\beta} - \left(\frac{1}{A}-1\right)\ln\frac{A\alpha}{\beta}}{2\ln(1+\alpha)}$$

$$= -\frac{\left(\frac{1-A}{A}\right)\ln\left(\frac{1-\frac{\beta}{\alpha}}{1-A}\right) + \ln\frac{\beta}{A\alpha}}{2\ln(1+\alpha)}.$$

Thus

$$\theta_{em}' = -\theta_{em}. \tag{51}$$

Now consider the relation (3) between peak phase error and peak time error:

$$\frac{\tau_{em}}{T_o} = \frac{\theta_{em}}{4\pi}\frac{1+A}{1-A}. \tag{52}$$

Let us again study the reverse transition by replacing θ_{em} with θ_{em}', A by $1/A$, and τ_{em} by τ_{em}':

$$\tau_{em}' = \frac{\theta_{em}'}{4\pi}\frac{\left(1+\frac{1}{A}\right)}{\left(1-\frac{1}{A}\right)} = \frac{\theta_{em}'}{4\pi}\frac{A+1}{A-1}.$$

Using (51),

$$\tau_{em}' = \frac{\theta_{em}}{4\pi}\frac{1+A}{1-A} = \tau_{em},$$

and the peak time jitter is of the same sign and magnitude for each transition. Reference to Fig. 13 shows that the sign of the peak jitter is always negative. Physically, this means that the apparent transition occurs at a time later than the true transition would occur.

REFERENCES

1. Caputo, S., and Highleyman, W. H., An Analysis of a Class of Variable-Frequency Sinusoidal Oscillators Using a Single Active Element, Proc. I.R.E., **50,** February, 1962, pp. 162–170.
2. Crosby, M. G., Reactance-Tube Frequency Modulators, RCA Review, **5,** July, 1940, pp. 89–96.
3. Meyers, S. T., An FM Data Set for Serial Transmission up to 1600 Bits per Second, Conference Paper CP61-1166, presented at Fall General Meeting of the A.I.E.E., Detroit, Michigan, October, 1961.
4. Hysko, J. L., Rea, W. T., and Roberts, L. C., A Carrier Telegraph System for Short-Haul Applications, B.S.T.J., **31,** July, 1952, pp. 666–687.
5. Stumpers, F. L. H. M., Theory of Frequency-Modulation Noise, Proc. I.R.E., **36,** September, 1948, pp. 1081–1092.