Lawrence Berkeley National Laboratory
Recent Work

Title
TIME QUANTIZATION IN A FEEDBACK SYSTEM

Permalink
https://escholarship.org/uc/item/59r8h3c1

Authors
Waddell, John F.
Morris, Harold D.

Publication Date
1953-05-01

eScholarship.org Powered by the California Digital Library
University of California

UCRL 2208
2nd Rev.

UNIVERSITY OF CALIFORNIA

Radiation Laboratory

TWO-WEEK LOAN COPY

This is a Library Circulating Copy which may be borrowed for two weeks. For a personal retention copy, call Tech. Info. Division, Ext. 5545

BERKELEY, CALIFORNIA

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

UCRL-2208(2nd Rev.)

UNIVERSITY OF CALIFORNIA
Radiation Laboratory
Berkeley, California

Contract No. W-7405-eng-48

A SPECTRAL APPROACH TO SAMPLED-DATA THEORY

John F. Waddell and Harold D. Morris

June 13, 1958

Printed for the U.S. Atomic Energy Commission

A SPECTRAL APPROACH TO SAMPLED-DATA THEORY

John F. Waddell and Harold D. Morris

Radiation Laboratory
University of California
Berkeley, California

June 13, 1958

ABSTRACT

This report discusses effects observed in a feedback system, otherwise linear, in which time is quantized. (These systems are often termed "sampled-data" systems). The time quantizer ("sampler"), the most important element in the system, is treated in detail as applying to any sort of sampling that can be approximated by a purely amplitude-sampling process.

Following a general physical description of amplitude-sampling, the output spectrum of an amplitude sampler of arbitrary pulse shape is derived and discussed. On the basis of the spectral analysis, a linear transfer function is developed which is applicable, with certain important limitations, to signals of all frequencies. The behavior of this transfer function is discussed, and several cases are distinguished.

The problem of stability in a feedback loop is treated, with particular attention to the important type of oscillations that are here dubbed "sampling oscillations." A general criterion for stability is developed, which includes all possible oscillations. Mathematical and physical interpretation of the stability criterion is provided, and a procedure is devised for the solution of practical problems.

Experimental evidence is offered in support of the theoretical results, with good agreement. The theory is also checked by calculation of the behavior of a special case whose sampling oscillations are amenable to transient analysis. Complete agreement between these calculations and the prediction of the theory is found.

A SPECTRAL APPROACH TO SAMPLED-DATA THEORY

John F. Waddell* and Harold D. Morris†

Radiation Laboratory
University of California
Berkeley, California

June 13, 1958

1. INTRODUCTION

When in an otherwise continuous-control feedback loop, some nonlinear element is inserted which imposes upon the system a quantization of time (a "sampling" process), the result is termed a "time-quantized," or "sampled-data" system. No matter how the time-quantizing element, or "sampler," operates, the result is the same: time is divided into equal intervals. At the beginning of each interval the loop is closed for a small fraction of a time quantum. The output of the sampler is a train of pulses, one for each quantum of time, each carrying the information derived from a momentary loop closure. In the usual form of sampler, the envelope of the output pulse train approximates the input signal. The analytical methods and concepts used here apply to any pulse shape, and with adaptation, to any type of sampling.

An example of a time-quantized feedback system is given in Fig. 1, where it is the error signal that is sampled. Other possible variations of such a system involve rearrangement of elements in the loop; to these apply the same concepts and methods, even as to systems containing only passive elements.

The earliest references in this field are those of MacColl[1] and Hurewicz,[2] each of whom uses an approach based upon the "Z-transformation,"

---

*Now at St. Mary's College, California.

†Now at Donner Scientific Company, Concord, California.

[1]MacColl, L. A. Fundamental Theory of Servomechanisms, (Van Nostrand, Princeton, N. J. 1944), Ch. 10.

[2]Hurewicz, W. in Theory of Servomechanisms, James, Nichols, and Phillips, Eds. (McGraw, New York, N. Y., 1947).

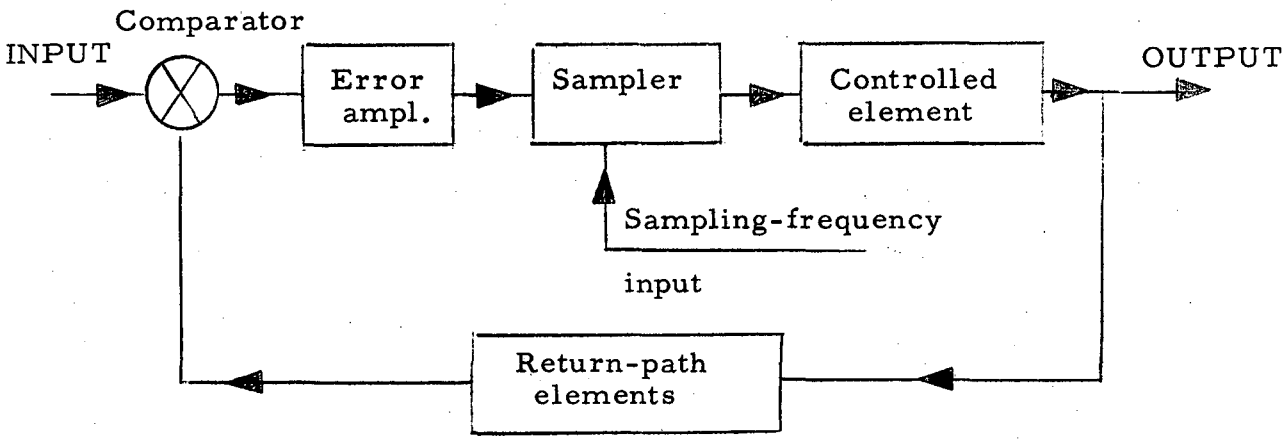

Fig. 1. A typical time-quantized feedback system.

a disguised Laplace transformation. The basis of the method is the use of sequences representing the values of input, output, and signals at other points in the feedback loop, at each instant of sampling. This involves calculations with knowledge of the pertinent signals only at the sampling instants.

Linvill takes a physical view of the problem, treating the particular case of "impulse" sampling in detail.[3] He regards the use of any sampling-pulse shape other than the impulse, or delta function, as being amenable to treatment as impulse sampling plus linear filtering. This is approximately true in most common cases, but is not an applicable point of view when the sampler pulse is shaped by nonlinear or by active networks, or when the sampling is done by means other than purely amplitude sampling. He develops expressions for the sampler input and output spectra and time functions, and uses these and the fact that the spectrum of the impulse-sampler output is periodic in the sampling frequency to derive an expression for the output time function of a feedback loop containing a sampler. An extension of Nyquist's stability criterion is implicit in his output equations, and he treats its application in the remainder of the paper. The spectral (Laplace-transform) point of view was applied to feedback systems containing a digital computer by Salzer.[4]

A very clear review of the field as of 1951 was contributed by Ragazzini

[3] W. K. Linvill, Sampled-data Control Systems Studies through Comparison of Sampling with Amplitude Modulation, Trans. Am. Inst. Elec. Eng'rs. 70, 1779 (1951).

[4] J. M. Salzer, Sc.D. Thesis, Mass. Inst. Tech. (1951). "A Treatment of Digital Control Systems and Numerical Processes in the Frequency Domain."

and Zadeh,[5] exponents of the Z-transform method, who most successfully relate their method to the Laplace-transform approach of Linvill.

Lago and Truxal treat the design of sampling systems, using the Z-transform method to treat the stability problem and using open-loop impulse-response calculations to derive the response between sampling instants.[6] However, this approach applies only to impulse sampling.

In another paper, Lago reveals a discrepancy between the results of the Laplace and Z-transform methods and proceeds to resolve the question.[7] In yet another contribution, Lago outlines a method by which sampled-data systems can be synthesized, using the Z-transform method.[8]

This work was undertaken to develop a theory of stability based upon the spectral approach, and to provide physical interpretation of the phenomena involved. In general we share the opinion of Linvill that the viewpoint which stresses frequency analysis lends itself most readily to physical interpretation. Consistent with the spectral point of view, we carry out the analysis in terms of a generalized pulse shape, subject only to the requirement that its Fourier integral converge. Generality is gained; the theory herein developed is not limited to cases where pulse shaping is performed by linear networks. The sampling process need not be a pure amplitude-sampling process at all, but may be accomplished by any means that yields a sampled output proportional to the amplitude of the input signal at the sampling instants. The theory developed here may be applied as long as the networks containing the sampling process are susceptible to transfer-function measurement and to determination of output spectrum.

---

[5] J.R. Ragazzini, and L.A. Zadeh, The Analysis of Sampled-Data Systems, Trans. Am. Inst. Elec. Eng'rs. 71, 225 (1952).

[6] G.V. Lago and J.G. Truxal The Design of Sampled-Data Feedback Systems, Trans. Am. Inst. Elec. Eng'rs. 73, 247 (1954).

[7] G.V. Lago Additions to Sampled-Data Theory, Proc. Nat'l. Electronics Conf. 10, 758 (1954).

[8] G.V. Lago A Synthesis Procedure for Sampled-Data System, Proc. Nat'l. Electronics Conf. 11, 351 (1955).

2. MECHANISM OF THE SAMPLING PROCESS

The sampling process is illustrated in Fig. 2.

The discussion to follow will concern itself principally with the transmission of a monochromatic input signal of arbitrary frequency having unit amplitude, as is usual in discussions of linear processes. The sampling process is a "linear" modulation process in which nonlinearity results in interaction between the sampling-frequency signal and the input-signal components, but not between the input components themselves. Thus to an input-signal component having some frequency ν, there corresponds an output component of the same frequency whose amplitude is proportional to that of the corresponding input component.

For an input signal whose frequency is very small compared with the sampling frequency, time quantization has negligible effect, and ordinary linear analysis is reasonably satisfactory. As the input frequency increases, so as to no longer be much smaller than the sampling frequency, effects of sampling become noticeable, and the degree of approximation to which the linear analysis applies becomes less. As the input frequency approaches sampling frequency, sampling effects become predominant, and ordinary feedback theory fails in several important respects. How this comes about is perhaps best explained from a purely physical point of view. For convenience, the discussion will be particularized to the case of "clamped" sampling.

Consider the case of an input signal whose frequency is exactly the sampling frequency (let the input signal be sampled at its own frequency); the output can then contain only a constant-amplitude (zero-frequency) term. The amplitude of the output then depends only upon the phase at which sampling occurs, and accordingly can take on any value of amplitude that might be repetitively reached by the input signal during successive sampling instants. Now suppose that the input-signal frequency is somewhat below sampling frequency (which we will call ν_o). Sampling will take place at intervals of somewhat less than a period of the input signal. The output will accordingly be a train of square pulses of length nearly equal to a period. This situation is shown in Fig. 3a. Note that the largest spectral component of the output has a frequency much lower than that of the input. Actually (as will be shown later), the low frequency is the difference between input and sampling frequencies. There is also a small component whose frequency is that of the input signal, and there are many smaller components whose frequencies are greater than the sampling frequency.

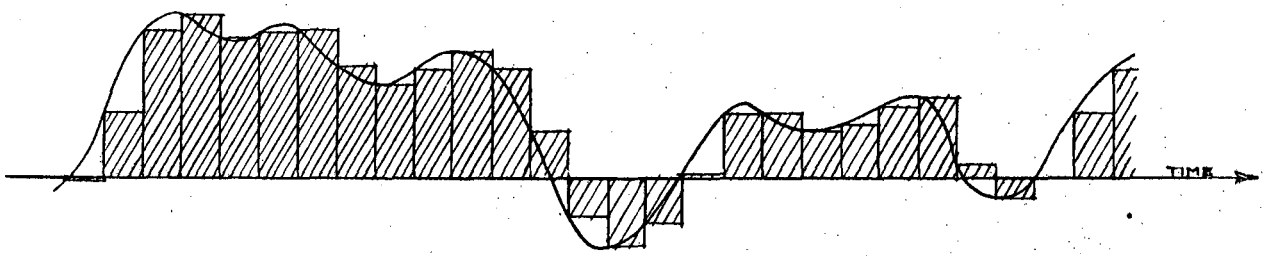

Fig. 2. Illustrating sampling for an input signal of arbitrary waveform.

One might now ask what the situation would be were the frequency of the input signal to have the lower value (i.e., lie close to zero). Figure 3b shows outputs due to input signals of unit amplitude and frequencies of $\nu_o/8$ and $7\ \nu_o/8$ to be identical in appearance. Also shown in Fig. 3b are the output spectral components of these frequencies, plotted from results shown subsequently; their sum is plotted in Fig. 3c as a matter of interest.

One can now form a picture of the output spectrum for frequencies less than ν_o. Given an input signal of unit amplitude and variable frequency, for input frequency near zero the output will contain a component of input frequency whose amplitude is close to unity. A difference-frequency component having small amplitude exists. As ν increases, amplitude of the component of input frequency decreases, and amplitude of the difference-frequency component increases.

At this point it is well to inquire into the happenings at half-sampling frequency, where the input and difference frequencies are equal. As for the case of equal input and sampling frequencies, magnitude of the output depends upon phase of the input signal at sampling instants, a situation illustrated graphically in Fig. 4. If input frequency were perturbed from $\nu_o/2$, the square-wave train would grow to maximum value, shrink to zero, then grow with opposite phase, pass maximum value, shrink again, etc. This is suggestive of interference between the input- and difference-frequency components (of equal magnitude at half-sampling frequency, their relative phases depending upon sampling phase).

Next consider the case in which $\nu = n\nu_o/2$ for $n > 2$. In the even case ($\nu = m\ \nu_o$), the result is obviously identical with the case $m = 1$, where the output has zero frequency with the amplitude dependent upon the phase at which sampling occurs. The only difference now is that sampling occurs every m periods of the signal. A similar relation holds for the half-multiples, $(2m-1)\ \nu_o/2 = (m - 1/2)\ \nu_o$. Here a train of square waves occurs, the period being $2/\nu_o$, sampling occurring every $(m - 1/2)$ periods of the signal. As will be shown later, the sum- and difference-frequency pattern repeats in every band of width ν_o sec^{-1}; with every multiple of sampling frequency including zero, there is associated a pair of spectrum lines of frequencies $m\ \nu_o + \nu$ and $(m + 1)\ \nu_o - \nu_o$. Physical behavior of the output is similar in the higher orders to that noted for the zeroth order,

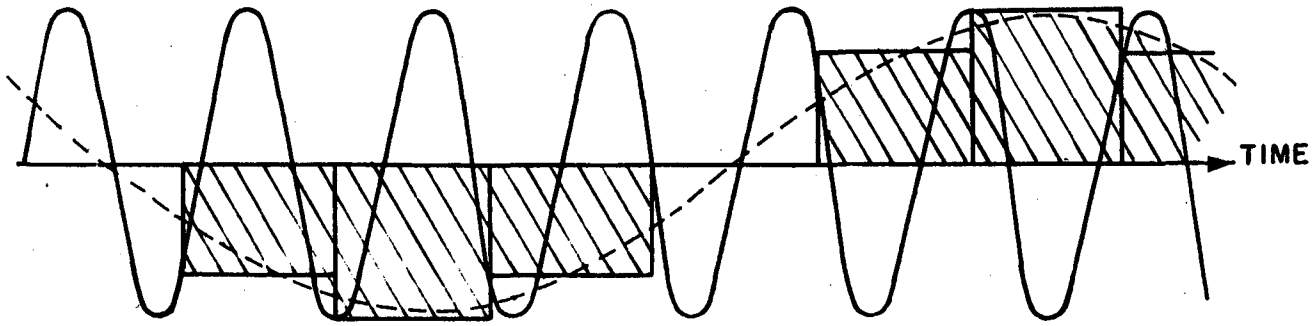

Fig. 3a. Clamped sampler input and output for input frequency of 7/8 sampling frequency. Note the generation of a low frequency (difference frequency) component nearly as large in amplitude as the input signal itself.

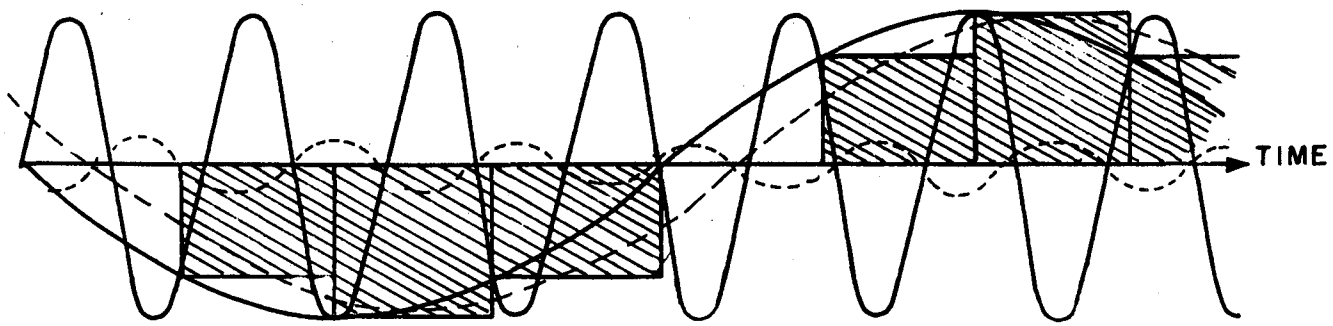

Fig. 3b. Clamped sampler output for input signals (solid lines) of 1/8 and 7/8 sampling frequency, and spectrum of output (i.e., output components) in the region below sampling frequency.

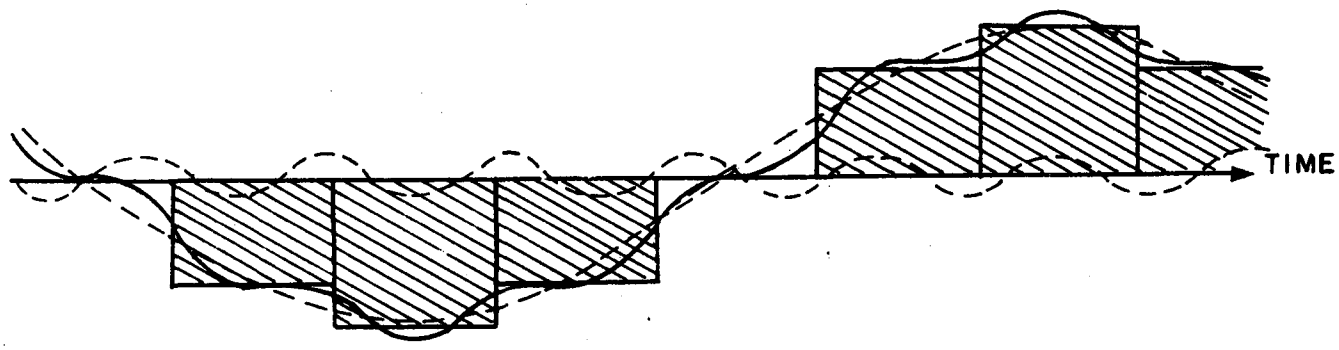

Fig. 3c. Clamped sampler output for input frequency of 1/8 or 7/8 sampling frequency. Components of the output having those frequencies are shown, together with their sum.

MUB-153

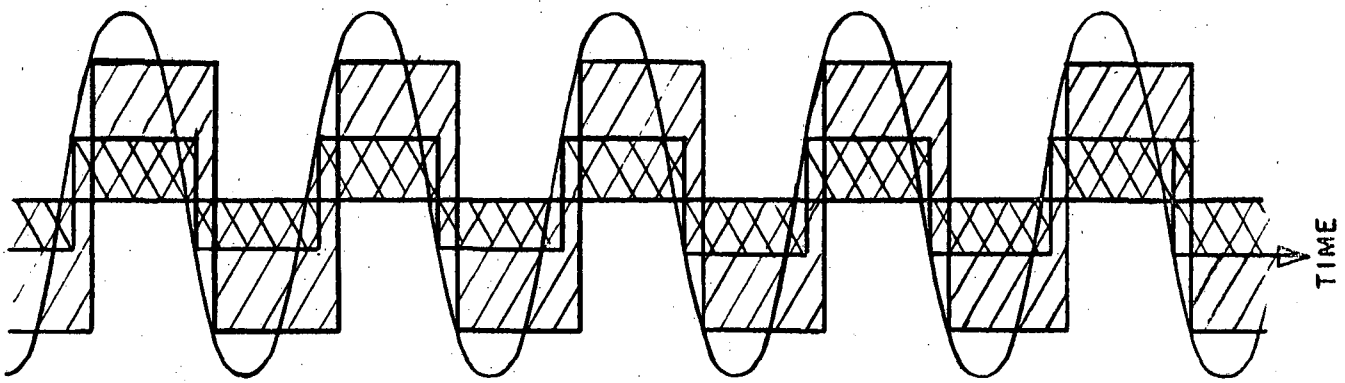

Fig. 4. Illustration of the effect of sampling phase upon clamped sampler output when input frequency is half the sampling frequency.

with the same interference behavior at $(m - 1/2)\ \nu_o$ and the zero-beat phenomena at $m\ \nu_o$. These effects were observed experimentally on a physical sampler up to approximately the fiftieth order, at which time imperfections in the sampler caused difficulty in the observations.

In discussing the transfer properties of a sampler, our interest lies almost exclusively in the region between zero and the sampling frequency. By "transfer properties" we mean the relationship existing between the input signal, a sinusoid of frequency ν, and that component of the output spectrum having the same frequency, ν. This relationship may conveniently be expressed in terms of a "transfer function," which is the complex ratio of output to input (considering only the output component having input frequency).

The term "transfer function" is used by most writers in feedback theory to describe something quite different from what is dealt with here.[9] Ordinarily, the transfer function of a network is the ratio of some operational transform of the output time function to that of the input for an arbitrary input function. For a linear passive network, the definition of "transfer function" used in this work agrees with the customary definition with the Fourier transform used, and is called by some writers the "frequency-transfer function". In this sense, a transfer function is a "spectral-modifying function" wherein the spectrum of the input is modified only by changes in amplitude and phase of the spectral components, these changes being a function of frequency of the component operated upon. This is an embodiment of the principle of superposition, and the output spectrum contains no component having any frequency not present in the input.

In the time-quantized case, however, the principle of superposition holds in another sense, namely that the output spectrum corresponding to an input spectrum made up of a series, or superposition of spectra, is the superposition of the output spectra due to each of the individual inputs. The sampler is a device having two inputs and one output, interaction occurring between the signals entering the two inputs. There is, however, no interaction between two signals simultaneously applied to the "signal input.", or the input. Therefore, no unique transfer function exists in the transform sense. In general, no function can be written down that expresses the ratio

---

[9] Cf. Greenwood, Holdam, and MacRae, Electronic Instruments, (McGraw-Hill, New York, N. Y., 1948) p. 230 et. seq; or Theory of Servomechanisms, James, Nichols, and Phillips, Eds., (McGraw-Hill, New York, N. Y., 1947). p. 58 , et. seq.

of the output spectrum to that of the input, excepting in a purely symbolic way. The symbolic "transfer function" one would thus obtain is dependent upon the input waveform and is not a function of the sampler alone.

The general nature of the amplitude of the previously defined sampler transfer function has been established as a monotone decreasing function having unit amplitude at zero frequency and zero amplitude at ν_o. A clue to the nature of the phase function may be found in Fig. 2. It seems reasonable to conclude that a constant delay of approximately half a sampling period occurs. Again, reference to Fig. 3b suggests the correctness of such a conclusion at the lower frequencies, although little can be said regarding frequencies in the region near ν_o. Thus, in the lower range there is reason to conclude that the sampler is, or is nearly, a constant-delay (linear-phase) device. Proof of this conclusion results from mathematical analysis of the output spectrum.

The transfer function of a sampler is all-important in considering such a device as part of a feedback loop; this paper is devoted principally to a discussion of transfer properties. Based upon physical reasoning, a picture has been drawn of this function on which only one blot appears aside from waveform distortion, namely, that at half-sampling frequency there is a discontinuity in behavior. This is of course due to the fact that $\nu_o/2$ is the only frequency for which $\nu_o - \nu = \nu$; therefore at $\nu_o/2$ one might expect the sampler to be a poor risk in a feedback loop.

3. THE OUTPUT SPECTRUM

An excellent discussion of factors affecting spectra composition of a sampler output is given by Bennett, who deals with the "flat" spectrum resulting from "instantaneous" sampling (sampling pulses are modulated "delta-functions").[10] For pulses of finite length, the spectrum is modified by an "aperture effect" which is a function of pulse shape and length. The most concise expression for the output spectrum results from the work of Kleene,[11] whose paper is a mathematical analysis yielding the output of a sampler having generalized pulse shape, assuming a generalized input waveform. His results will be used, and enlarged upon, in this work.

Let there be given, then, a sampler whose pulse shape g(t) is arbitrary, save that its Fourier integral converges. An example of such a waveform is provided in Fig. 5a, which depicts the unmodulated output of some sampler as a function of time. The sequence of instants marking the beginning of each sampling period we call $\left\{t_k\right\}$, i.e., $\left\{t_k\right\} = \left\{k/\nu_o\right\}$.

Assume that the input signal f(t) possesses an enumerable set of spectrum lines, implying that the input can be expressed as the sum of a finite or enumerably infinite set of periodic functions. Thus, we can express f(t) as a sum of Fourier series,

$$f(t) = \sum_{s=1}^{\infty} \sum_{m=-\infty}^{\infty} c_{ms}\, e^{2\pi i m \nu_s t} \tag{3.1}$$

where $1/\nu_s$ is the period of the sth periodic component of f(t).

Now by definition of the sampling process, an example of which is provided in Fig. 5b, the output of the sampler is

$$F(t) = f(\left\{t_k\right\})\, g(t - \left\{t_k\right\}), \tag{3.2}$$

consisting of the pulse train $g(t - \left\{t_k\right\})$ modulated by the sequence of input values at the sampling instants, $f(\left\{t_k\right\})$.

---

[10] W. R. Bennett, Spectra of Quantized Signals, Bell System Tech. J. 27, 446 (1948).

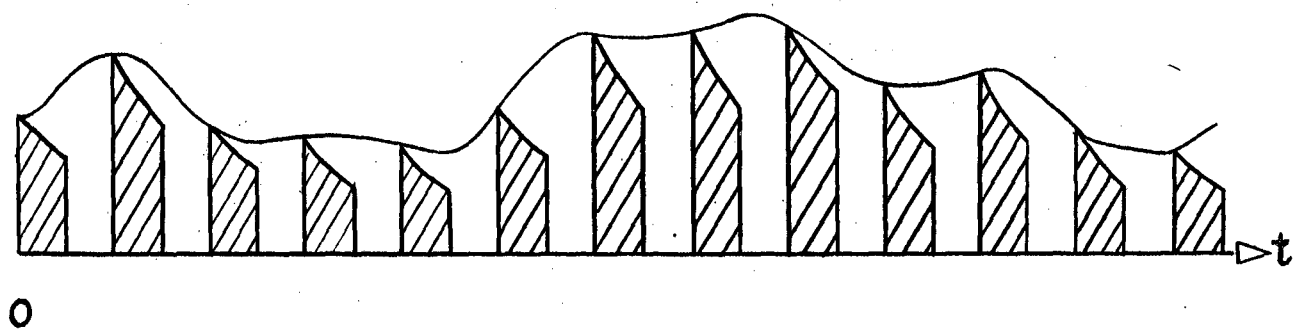

Fig. 5. Input and output of a sampler having some (arbitrary) pulse shape g(t). (a) Unmodulated pulse train. (b) Modulated pulse train F(t) (cross-hatched), with modulating input function f(t) (solid-line curve).

When Eq. (3.1) is written for the sequence $\{t_k\}$, and substituted into Eq. (3.2), with suitable transformations,[11] the result is

$$F(t) = \sum_{s=1}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} c_{ms}\, W(n\nu_o + m\nu_s) e^{2\pi i(n\nu_o + m\nu_s)t}, \quad (3.3)$$

where m and n are integers, and

$$W(\nu) = \nu_o \int_0^{1/\nu_o} g(t)\, e^{-2\pi i \nu t}\, dt . \quad (3.4)$$

The nature of the output spectrum is apparent: there are spectrum lines for all values of m and n at $n\nu_o + m\nu_s$. The complete spectrum of the input signal is transposed in frequency to cluster about each integral multiple ("harmonic") of the sampling frequency. The entire spectrum is modified by that of the single sampling puse—i.e., each spectral component is modified by the Fourier transform of the sampling pulse, evaluated at the frequency of the line in question. Illustration of the above is given in Figs. 6 and 7.

In the case where the input is a simple sinusoid, the relations take the form

$$f(t) = a \cos(2\pi\nu t) \quad (3.5)$$

or

$$f(t) = \frac{a}{2} e^{2\pi i \nu t} + \frac{a}{2} e^{-2\pi i \nu t} . \quad (3.6)$$

Thus we have

$$F(t) = \sum_{n=-\infty}^{\infty} \frac{a}{2} \left[W(n\nu_o - \nu) e^{2\pi i (n\nu_o - \nu)t} + W(n\nu_o + \nu) e^{2\pi i (n\nu_o + \nu)t} \right] . \quad (3.7)$$

Each integral multiple of ν_o has "sidebands" differing from $n\nu_o$ by $\pm\nu$.

It is of interest to consider the spectrum as input frequency varies from zero to many times the sampling frequency. As the input frequency term ν starts up from zero, the difference-frequency term $\nu_o - \nu$ starts down from ν_o, and a set of sum-frequency terms $n\nu_o + \nu$ and of difference-frequency

---

[11] S. C. Kleene, Analysis of Lengthening of Modulated Repetitive Pulses, Proc. IRE 35, 1049 (1947).

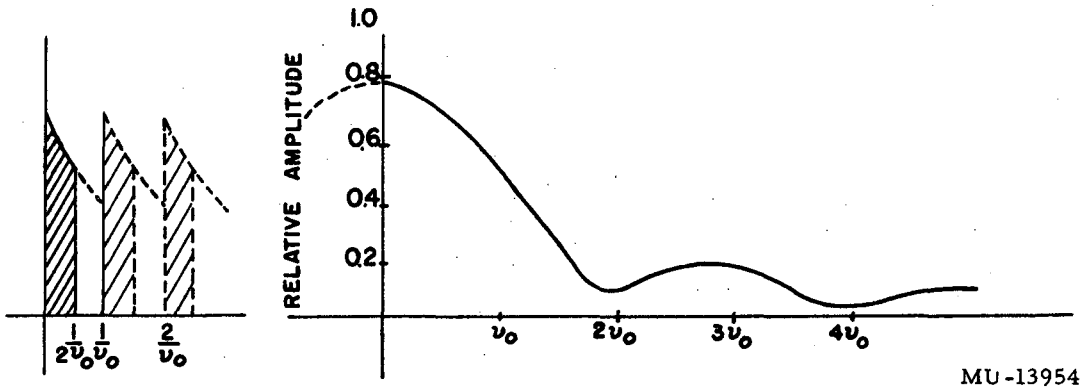

Fig. 6. Pulse train of Fig. 5a. and magnitude of the spectral density (Fourier transform) of a single pulse (positive-frequency portion).

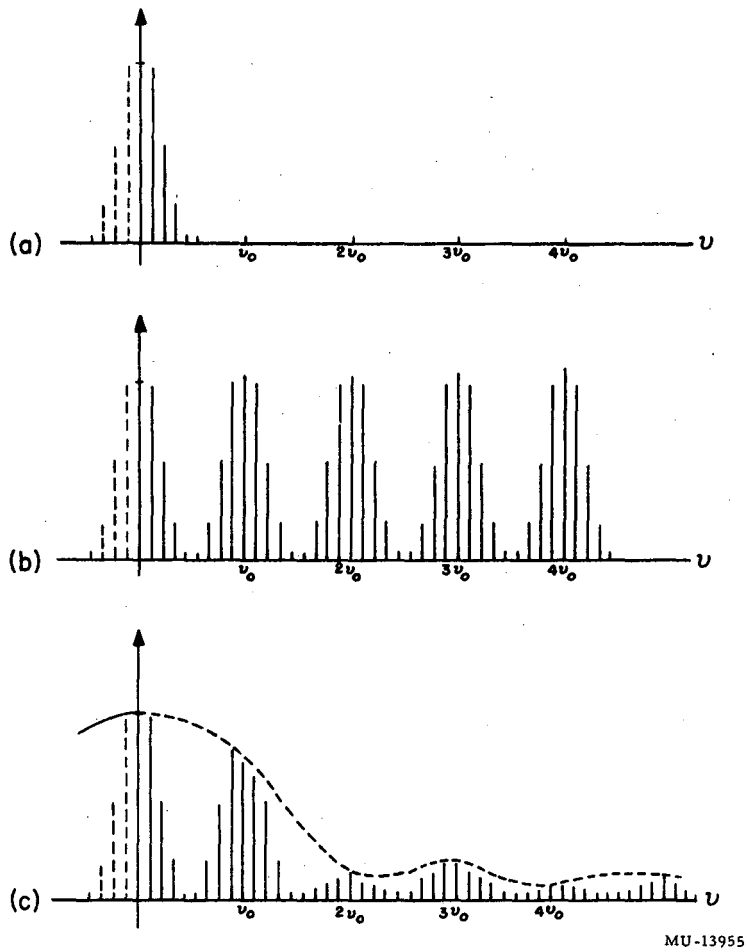

Fig. 7. Structure of sampler output spectrum, showing (a) amplitude spectrum of some arbitrary input signal; (b) input-signal spectrum transposed to multiples of the sampling frequency (This is not the output spectrum.); and (c) input signal transposed and modified by spectrum of sampling pulse. Only amplitude is shown, of course (This is the output spectrum).

terms $(n + 1)\nu_o - \nu$ approach each other in every interval of width ν_o. Each term is of course modified by $W(\nu)$.

As the input-frequency term comes to the value $\nu_o/2$, all of the sum and difference terms in each interval of width ν_o coalesce. When ν becomes larger than $\nu_o/2$, the spectrum lines change roles (as though they had "passed" each other at the moment of coalescence); behavior of the spectrum is symmetrical about $\nu_o/2$. When ν is any integral multiple of ν_o, the spectrum is identical to that for $\nu = 0$. Indeed, inspection of Eq. (3.7) shows that the same spectrum results for a unit amplitude input applied, having the frequency of any spectrum line. Figure 3 illustrates the effects mentioned above.

One further effect is left to be explored, the possibility of cross-modulation. In Eq. (3.3) we let the index s take on the values 1 and 2, and m the values ± 1, corresponding to an input signal consisting of two equal sinusoids, with frequencies ν_1 and ν_2. Thus we obtain

$$F(t)\, C \sum_{n=-\infty}^{\infty} \Big[W(n\nu_o - \nu_1)\, e^{2\pi i(n\nu_o - \nu_1)t} + W(n\nu_o + \nu_1) e^{2\pi i(n\nu_o + \nu_1)t} + W(n\nu_o - \nu_2) e^{2\pi i(n\nu_o - \nu_2)t} + W(n\nu_o + \nu_2) e^{2\pi i(n\nu_o + \nu_2)t} \Big], \quad (3.8)$$

since we have $C_{-1s} = C_{1s} = C_s = C_1 = C_2 = C$. The result is a linear combination of terms in frequencies ν_1 and ν_2 separately; no terms of frequency $\nu_1 \pm \nu_2$ or of any other combination of ν_1 and ν_2 occur, and we are led to the obvious conclusion that there can be no cross modulation. Hence we have the term "linear modulation," used previously. Therefore the principle of superposition can be correctly applied to the sampling process. Experimental verification of the above result was obtained.

4. THE TRANSFER FUNCTION

Here is examined the behavior of the particular line of the output spectrum having input frequency, as the input frequency is varied over the entire domain of real values. The input signal will be throught of as a real function with a positive real frequency, ν, expressible by Eq. (3.5). Because we are using complex Fourier series representation, there are two spectrum lines, at $\pm \nu$ (i.e., $m = \pm 1$), as expressed in Eq. (3.6). The line of frequency ν is represented in Eq. (3.7) by the term in $+\nu$ for $n = 0$, namely $W(\nu) \exp(2\pi i \nu t)$. The coefficient of this term, $W(\nu)$, is the transfer function of the sampler everywhere except on the discrete set of frequencies $\nu = (2n - 1)\nu_o/2$ (a case to be discussed in detail subsequently). $W(\nu)$ is defined by Eq. (3.4) (which is assumed to be convergent); hence $W(\nu)$ is an analytic function of ν.

In the discussion to follow, it will be useful to have as an illustrative example a specific type of sampler. The clamped sampler provides a suitable example. The discussion of the transfer function will, however, be kept in general terms, applicable to any physical pulse shape. Application of the principles will subsequently be made to the delta-function, or impulse-type sampler, which has some interesting properties.

In the case of the clamped sampler, the pulse is one of simple square shape and of duration equal to the repetition period. Thus. Eq. (3.4) integrates to

$$W(\nu) = e^{-i\pi \frac{\nu}{\nu_o}} \frac{\sin \pi \frac{\nu}{\nu_o}}{\pi \frac{\nu}{\nu_o}} , \tag{4.1}$$

or

$$\left| W(\nu) \right| = \left| \frac{\sin \pi \frac{\nu}{\nu_o}}{\pi \frac{\nu}{\nu_o}} \right| \tag{4.2}$$

and

$$\arg W(\nu) = - \pi \frac{\nu}{\nu_o} + \arg \left(\frac{\sin \pi \frac{\nu}{\nu_o}}{\pi \frac{\nu}{\nu_o}} \right) . \tag{4.3}$$

Equation (4.3) embodies the linear-phase property discussed previously, the function arg (sin u/u) being $-k\pi$, where k is the number of zeros of sin u/u which lie between 0 and u. The phase properties are an outgrowth of the pulse shape alone.

Figure 8 shows the locus of this transfer function in the complex plane for $\arg W \leq 8\pi$, i.e., for the first four orders of the spectrum.

The function $W(\nu)$ as described in Eq. (4.1) is everywhere single-valued and continuous in the domain of real frequencies. It has a zero at each point of the set $\nu = n\nu_o$ (n is any integer > 0), implying that there is no output at sampling frequency or any integral multiple thereof greater than zero. Its curvature is continuous at the zeros, as examination of Eq. (4.3) will show.

As previously mentioned, occurrence of the input frequency at any odd integral multiple of half-sampling frequency gives rise to discontinuous behavior of the transfer function. Physically, this result may be ascribed to an interference between the sum and difference frequencies in each spectral order, as is shown in the analysis to follow.

The method is one of perturbation of the input frequency from any member of the set $(2n - 1)\ \nu_o/2$; recalling the property of the spectrum that identical spectra are obtained in response to unit inputs having the frequency of any line of the spectrum we use the value $\nu_o/2$ for input frequency. Thus we write Eq. (3.7), and we change the index of summation from the multiple of sampling frequency to the spectral order; i.e., the interval of any order k lies between the zeros of the (k - 1) and the k orders. Then let

$$\nu = \frac{\nu_o}{2} - \delta\nu \tag{4.4}$$

with

$$\frac{\delta\nu}{\nu_o} << 1 .$$

If we substitute Eq. (4.4) into the modified Eq. (3.7), and simplify the expressions for the frequencies, then $W(\nu)$, being analytic, may be replaced by its Taylor expansion in the neighborhood of $(k - 1/2)\ \nu_o$, k being any integer. In view of the restriction of Eq. (4.4), we may replace $W\left[(k - 1/2)\ \nu_o \mp \delta\nu\right]$ by the leading term. The result is

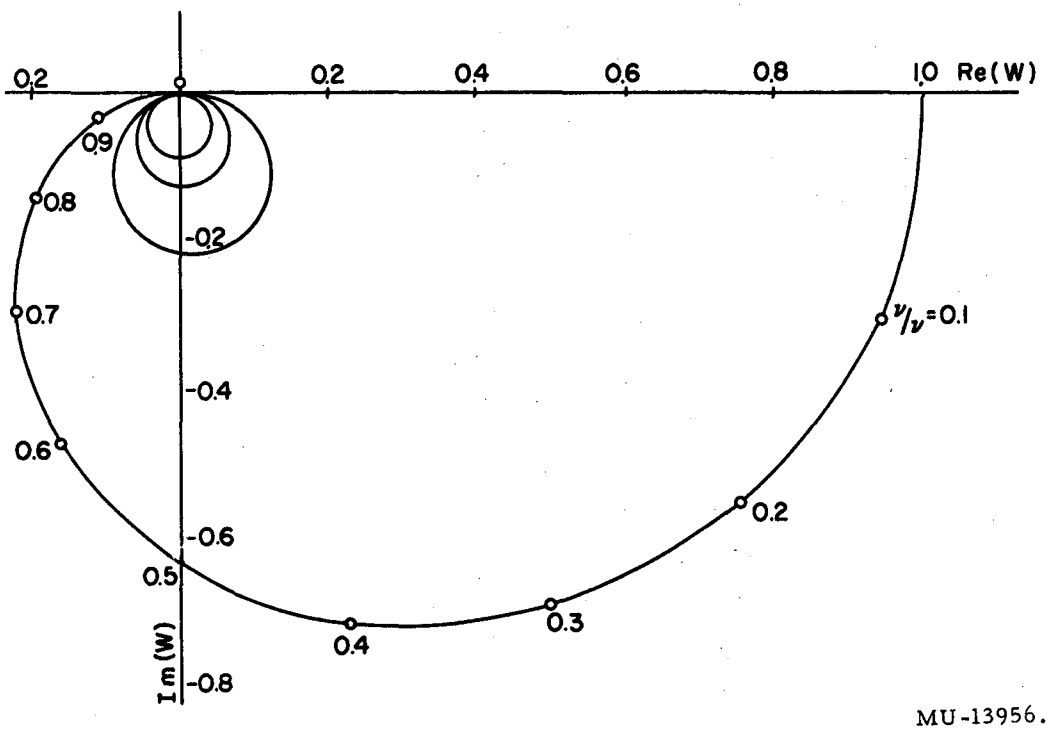

Fig. 8. Partial transfer function of a perfect amplitude sampler of the clamped type. Numbers shown are the frequency normalized to sampling frequency. Only the first four orders of the function are shown.

$$F(t) = \sum_{k=-\infty}^{\infty} \frac{a}{2} W\left[(k-1/2)\,\nu_o\right] \left\{ e^{2\pi i \left[(k-1/2)\nu_o - \delta\nu\right] t} + e^{2\pi i \left[(k-1/2)\nu_o + \delta\nu\right] t} \right\} \tag{4.5}$$

Now we consider the behavior of each term (spectral order) independently of the others, phase being referred to the lower-frequency (difference-frequency) line, with the result

$$F_k(t) \exp\left\{-2\pi i \left[(k-1/2)\nu_o - \delta\nu\right] t\right\} = \frac{a}{2} W\left[(k-1/2)\nu_o\right]\left[1 + e^{4\pi i \delta\nu t}\right]. \tag{4.6}$$

In each order the output consists of the sum of two components which interfere in the limit, when their frequencies are $k\nu_o/2$, to produce an output magnitude that is in accordance with the phase of sampling. If we allow the perturbation to remain, we see the behavior very clearly. For <u>every order</u>, the rate of precession in phase of the upper component with respect to the lower has the same value, namely $2(2\pi\,\delta\nu)$. The time required for a precession through one period of the output is the same for all orders, irrespective of their frequency. (Note that their frequencies are perturbed integral multiples of the signal frequency.) Thus the output waveform grows from nothing to its maximum value and then shrinks again to nothing, the cycle being repeated as long as the perturbation remains. Throughout all this process the waveform remains unchanged, because the interference is in each order of the same degree as in all the others. All the orders interfere to the same degree, simultaneously. If at any time in the cycle that is decribed here the perturbation is removed, the output remains frozen at the condition then existing. The exponent $2(2\pi\delta\nu t)$ is seen to be just twice the phase of sampling, the transfer function thus having period π with respect to the variable ϕ. Equation (4.6) may therefore be written in the limit

$$F_k(t) \exp\left[-2\pi i (k-1/2)\,\nu_o t\right] = \frac{a}{2} W\left[(k-1/2)\nu_o\right]\left[1 + e^{2i\phi}\right], \tag{4.7}$$

where ϕ is the phase at which sampling occurs.

Equation (4.7) is the kth term of the series in Eq. (4.5), in the limit $\delta\nu = 0$. The output is therefore

$$F(t) = \frac{a}{2}\left(1 + e^{2i\phi}\right) \sum_{k=-\infty}^{\infty} W\left[(k-1/2)\nu_o\right] e^{2\pi i(k-1/2)\,\nu_o t} . \qquad (4.8)$$

This is the response to a monochromatic input having for its frequency any odd integral multiple of half the sampling frequency. If the function $W(\nu)$ is represented by a contour on the complex plane (e.g., Fig. 8), each of the interference terms in the output (Eqs. 4.6, 4.7, and 4.8) adds a circle centered on $W\left[(k-1/2)\nu_o\right]$ and passing through the origin.

The behavior described above also occurs at any integral multiple of sampling frequency. To show this, we take Eq. (3.7), and without rearranging, we let $\nu = \delta\nu$ for $\delta\nu/\nu_o << 1$ as before, realizing that, as previously, in the limit, $W(n\nu_o \pm \delta\nu)$ approaches $W(n\nu_o)$ as $\delta\nu$ is allowed to vanish. Thus the factors $W(n\nu_o) \exp(2\pi i n\nu_o t)$ are removed. Again, as before, phase is referred in each term to the lower-frequency spectrum line. Thus we have the set of equations

$$F_n(t) \exp\left[-2\pi i\,(n\nu_o - \delta\nu)t\right] = \frac{a}{2}\, W(n\nu_o)\left[1 + e^{4\pi i\delta\nu t}\right]. \qquad (4.9)$$

All the remarks made with reference to the $\nu_o/2$ case apply here. In the clamped sampler, $W(n\nu_o) = 0$ for all values of $n > 0$. For samplers having pulse shapes other than square, $W(n\nu_o)$ does not in general vanish, and one thus finds interference at all integral multiples of half the sampling frequency.

A case in point is the clamped sampler, the pulse shape being square and of length $1/\nu_o$, whence we write

$$W\left[(k-1/2)\nu_o\right] = -i\,\frac{2}{(2k-1)\,\pi} . \qquad (4.10)$$

Thus we obtain

$$F(t) = -i\,\frac{a}{2}\,\frac{2}{\pi}\left(1 + e^{2i\phi}\right) \sum_{k=-\infty}^{\infty} \frac{e^{2\pi i(k-1/2)\,\nu_o t}}{2k-1} . \qquad (4.11)$$

This case is illustrated in Fig. 9. The first order shows the largest and most important effect, and hence is given special attention in Fig. 10. If we use the clamped sampler as an illustration of the general case and take

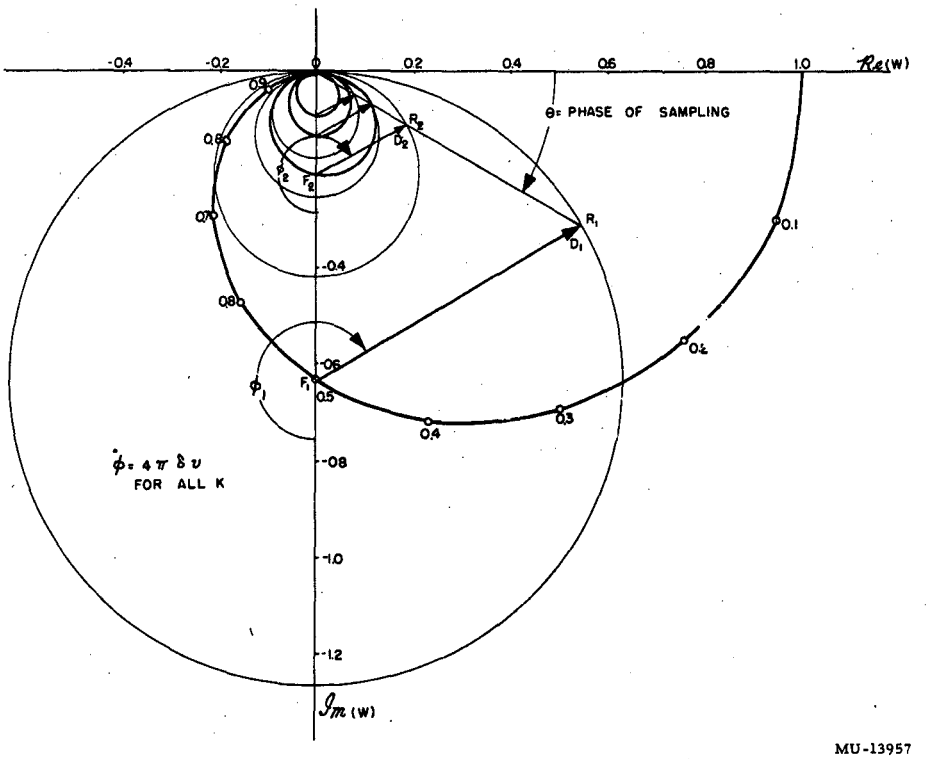

Fig. 9. Complete transfer function for the clamped sampler in the first four orders, showing the interference structure occurring at $n\,\nu_o/2$. Here F_k are sum-frequency components of kth order, D_k are difference-frequency components, and R_k are the resultants.

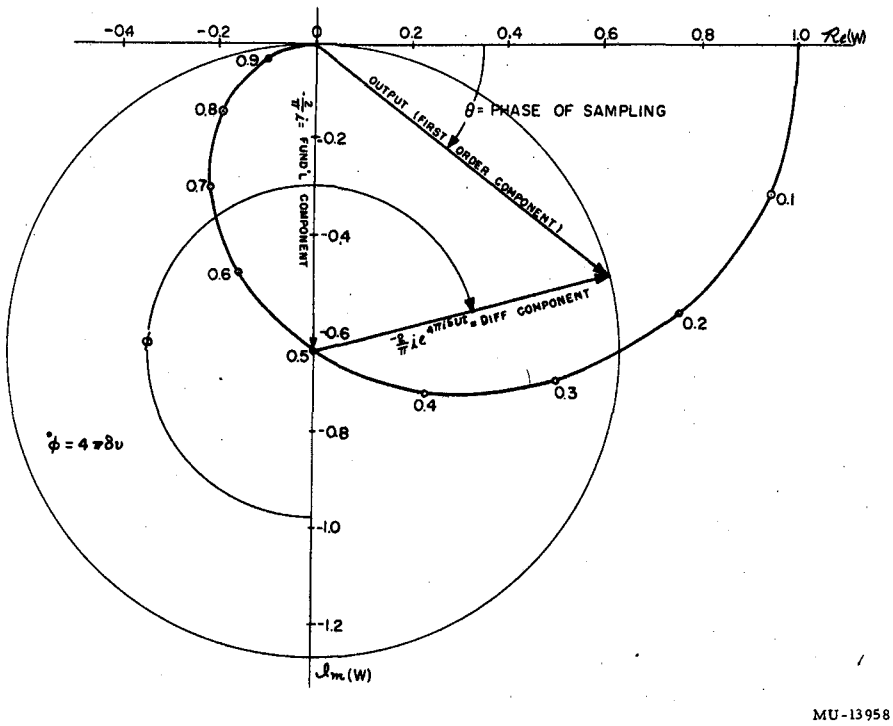

Fig. 10. Region below the first zero of the sampler transfer function of Fig. 9, showing mechanism of interference between spectral components of sampler output at half-sampling frequency.

the perturbation once more, the input-frequency component lies arbitrarily close to $\nu_o/2$, and the difference-frequency component is symmetrically disposed beyond $\nu_o/2$. The latter case precesses in phase about the "fundamental," or input-frequency component at a rate equal, in precession periods per second, to the separation $(2\delta\nu)$ between the two spectrum lines. If at any time we remove the perturbation, we have a possible combination of circumstances in the limit at $\nu_o/2$. Thus, solely because of chance, the output phase at $\nu_o/2$ could lie anywhere on the circular locus shown. The accidental nature of this interference phenomenon makes the transfer function discontinuous at $(2n-1)\,\nu_o/2$, with the transfer function consisting of $W(\nu)$ plus the set of circles discussed above.

The case where the sampling pulses are unit impulses is of interest, as it has been used in the work of both Linvill and Bennett. Let the pulse shape now be

$$g(t) = \delta(t), \tag{4.12}$$

where $\delta(t)$ is the Dirac δ function. Thus we write

$$\delta(t) = \lim_{a \to 0} S(t, a) \quad , \tag{4.13}$$

where

$$S(t, a) = \frac{1}{\sqrt{2\pi}} \frac{1}{2} e^{-\frac{t^2}{2a^2}} \quad . \tag{4.14}$$

The function $S(t, a)$ is a Gaussian of width a and amplitude $1/a$, hence of unit area whatever the value of a.

Substituting the above into Eq. (3.4) and performing the integration, we have

$$W(\nu) = \lim_{a \to 0} W(\nu, a) \tag{4.15}$$

and

$$W(\nu, a) = e^{-2\pi^2 a^2 \nu^2} \quad . \tag{4.16}$$

Combined, Eqs. (4.6) and (4.9) are

$$F_n(t) \exp\left[-2\pi i \left(m \frac{\nu_o}{2} - \delta\nu\right)t\right] = W\left(m\frac{\nu_o}{2}\right)\left[1 + e^{4\pi i \delta\nu t}\right] \quad .$$

For $W(m\nu_o/2)$ we substitute the expression of Eq. (4.16), with $\nu = \nu_o/2$, and pass to the limit $a = 0$. Thus we obtain

$$F_n(t) \quad \exp\left[-2\pi i\left(m\,\frac{\nu_o}{2} - \delta\nu\right)t\right] = 1 + e^{2\pi i(2\delta\nu)t} \quad . \quad (4.17)$$

All orders have their interference circles superposed in a single one centered on the point $1 + 0i$.

To consider that the sampling function will occur on an impulse basis in any physical system is of course a mathematical idealization. We can profitably make this idealization if the actual pulse shape can be attributed to linear passive networks, because the function of these can be readily described in terms of a transfer function. If, however, the pulse shaping is due to the action of non-linear or active networks, things are not so simple, and therefore describing the transfer properties of a physical sampler in terms of its pulse shape has the advantage of simplicity and realism. In any event, the δ-function idealization does not in any respect free us from the interference phenomenon. It is an essential result of the quantization of time, and of nothing else.

We might also note that the impulse sampler is a reasonable idealization of a physical sampler using short pulses, and illustrates one significant feature of the latter: the shorter the sampling pulse, the smaller the output amplitude for a given input. In fact, the output-to-input amplitude ratio is proportional to the pulse width, a, for the Gaussian shape, which comes fairly close to being representative of very short electronically-generated pulses.

5. ARBITRARY INPUT, PERIOD NEAR $2/\nu_o$

To facilitate the discussion in the next section, we will here consider the output resulting from an arbitrary periodic input whose period is close to twice the sampling period. To start with, the input and output are given by Eqs. (3.1) and (3.3), respectively, with S restricted to unity. Then Eq. (3.3) becomes

$$F(t) = \sum_{n=-\infty}^{\infty} c_o W(n\omega_o)\, e^{in\omega_o t} + \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} \left[c_m W(n\omega_o + m\omega)\, e^{i(n\omega_o + m\omega)t} + c_{-m} W(n\omega_o - m\omega)\, e^{i(n\omega_o - m\omega)t} \right]. \tag{5.1}$$

For convenience, we call the double-sum term by the name H(t). Applying Eq. (4.4), we obtain

$$H(t) = \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} \left\{ c_m W\left(n\omega_o + m\frac{\omega_o}{2} - m\delta\omega\right) e^{i\left(n\omega_o + m\frac{\omega_o}{2} - m\delta\omega\right)t} + c_{-m} W\left(n\omega_o - m\frac{\omega_o}{2} + m\delta\omega\right) e^{i\left(n\omega_o - m\frac{\omega_o}{2} + m\delta\omega\right)t} \right\}.$$

We next shift the order of summation so as to pair off terms whose frequency arguments are separated only $\delta\omega$, making the explicit provision that the analysis will apply only to pulse shapes such that the function $W(\omega)$ is analytic (see Eq. 3.4). Imagine the second n sum (that involving c_{-m}) in Eq. (5.2) to have an index k instead of n. Then impose the condition

$$n\omega_o + m\frac{\omega_o}{2} = k\omega_o - m\frac{\omega_o}{2}$$

or

$$n + m = k \quad .$$

This is substituted into the first n sum of Eq. (5.2); the expression is now summed on the index k, resulting in

$$H(t) = \sum_{m=1}^{\infty} \sum_{k=-\infty}^{\infty} \left\{ c_m W\left[(2k-m)\,\frac{\omega_o}{2} - m\delta\omega\right] e^{i\left[(2k-m)\frac{\omega_o}{2} - m\delta\omega\right]t} + c_{-m}\, W\left[(2k-m)\,\frac{\omega_o}{2} + m\delta\omega\right] e^{i\left[(2k-m)\,\frac{\omega_o}{2} + m\delta\omega\right]t} \right\} .$$

For sufficiently small values of $\delta\omega/\omega$, we have

$$H(t) = \sum_{m=1}^{\infty} \sum_{k=-\infty}^{\infty} c_m W\left[(2k-m)\,\frac{\omega_o}{2}\right] e^{i\left[(2k-m)\,\frac{\omega_o}{2} - m\delta\omega\right]t} \times \left[1 + \frac{c_{-m}}{c_m} e^{2im\delta\omega t}\right]. \qquad (5.3)$$

The quantity $2\delta\omega t$ is the precession rate of the sampling instant. If the perturbation is removed, the process is stopped at some particular value of the sampling phase, reckoned on the scale $2\pi = 2/\nu_o$ (the phase is referred to the fundamental frequency component of the spectrum). We call

$$t\delta\omega = \phi(t) \qquad (5.4)$$

the sampling phase, whose zero is reckoned from the maximum value of the interference factor in Eq. (5.3), (i.e., from the time of reinforcement).

Referring to the expression for f(t), we see that in order for f(t) to be real, we have

$$c_m = c^*_{-m} \text{ (complex conjugate) and thus}$$

$$c_{\pm m} = \mathscr{C}_m\, e^{\pm i\theta_m} .$$

Accordingly we find

$$\frac{c_{-m}}{c_m} = e^{-2i\theta_m} ,$$

and the last (bracketed) factor in Eq. (5.3) is

$$1 + e^{2i(m\phi - \theta_m)} .$$

Particular interest lies in the case where input frequency is precisely $\nu_o/2$. Equation (5.3) is, with the perturbation removed (and with slight rearrangement),

$$H(t) = \sum_{m=1}^{\infty} c_m \left[1 + e^{2i(m\phi - \theta_m)}\right] \sum_{k=-\infty}^{\infty} W\left[\left(k - \frac{m}{2}\right)\omega_o\right] e^{i\left(k - \frac{m}{2}\right)\omega_o t} . \qquad (5.5)$$

The entire output is

$$F(t) = \sum_{k=-\infty}^{\infty} c_o W(k\omega_o)\, e^{ik\omega_o t} + H(t) . \qquad (5.6)$$

As noted above, θ_m is the phase of c_m, and ϕ is the phase of sampling. With regard to the latter, some remarks may be pertinent, as considerable potential for ambiguity exists.

In the case where $m = 1$ (the so-called "fundamental" component of the input spectrum), the argument of the interference term is $2i(\theta_1 - \phi)$, giving maximum value to the term at $\phi = \theta_1$. Likewise, for each order, the term reaches its maximum value when $m\phi = \theta_m$. Now ϕ is the phase of the "fundamental," or first-order term at which sampling occurs, give or take some constant that depends upon input waveform. The zero of ϕ is determined by the fact that maximum first-order output obtains if sampling occurs at $\phi = \theta_1$, the phase of any component being referred back to the interval between successive occurrences of any particular value of the input time function.

In Eq. (5.6), the first series represents the output corresponding to the average value, or zero-order term, of the input, and consists of a track of identical pulses. Superposed upon this is the dynamic output signal, H(t), of Eq. (5.5).

The function H(t) (Eq. 5.5) is to be compared with the function F(t) of Eq. (4.8). In Eq. (5.5), each term of the sum on m consists of, firstly, the Fourier coefficient of the input component of mth order; and secondly, the mth-order interference factor $1 + \exp(2i\psi_m)$, where $\psi_m = m\phi - \theta_m$ is the phase

at which the mth-order component is sampled. Thirdly, there is the output pulse train embodied in the k sum. The (partial) argument $(m/2)\omega_o$ is the frequency of the mth-order spectrum line of the input and output. The complete argument $(k-m/2)\omega_o$ is the set of frequencies of the set of spectrum lines resulting from the mth-order input component.

It is profitable to write H(t) in another way. Because the argument of the k series in Eq. (5.5), may be written $(2k-m)\omega_o/2$, and since the factor (2k-m) is a dummy variable, it may be used as the index of summation. This factor is dependent upon m only in the matter of oddness or evenness; i.e., if m is odd, (2k-m) is odd; if m is even, (2k-m) is even. Thus it is possible to break up H(t) into two double sums, one involving only even values both of indices m and k, the other involving only odd values of both indices. At this juncture it becomes apparent that each sum, odd and even, is the product of two series, which can be rearranged to render the output

$$F(t) \sum_{k=-\infty}^{\infty} \left\{ \left[c_o + \sum_{m=1}^{\infty} c_{2m} \left(1 + e^{2i(2m\phi - \theta_{2m})}\right)\right] W\left(2k\frac{\omega_o}{2}\right) e^{2ik\frac{\omega_o}{2}t} + \sum_{m=1}^{\infty} c_{2m-1}\left[1 + e^{2i\left[(2m-1)\phi - \theta_{2m-1}\right]}\right] W\left[(2k-1)\frac{\omega_o}{2}\right] e^{i(2k-1)\frac{\omega_o}{2}t} \right\}. \tag{5.7}$$

It is instructive to compare Eqs. (5.7) and (4.8). In the case of a monochromatic input of period $2/\nu_o$, a symmetrical output is generated. However, when the input is not symmetrical, albeit of period $2/\nu_o$, the output cannot be symmetrical unless (and only unless) the sampling instants happen to fall at equal positive and negative values of the input. Such an event would be fortuitous indeed. In general, a nonsymmetrical periodic input generates an output in which the even-order spectrum lines are quite as important as those of odd order.

As an informative application we consider the input to be

$$f(t) \sum_{m=-\infty}^{\infty} c_m e^{im\omega t} = a\left(1 + \cos\frac{\omega_o}{2}t\right) .$$

Here we have defined $c_o = a$, $c_{-1} = c_1 = a/2$, $c_m = 0$ for all values of $m \neq -1, 0$ or $+1$ and $c_2\theta_m = 0$. In Eq. (3.8), let $\nu_1 = 0$, $\nu_2 = \frac{\nu_o}{2}$. There results an equation for F(t) which is the sum of sampler outputs corresponding to (a) the dc

component of the input (above), and (b) its sinusoidal part. We have the latter output expressed precisely in Eq. (4.8). If we substitute the above values for the c_m and the θ_m in Eq. (5.7), the same result is found.

6. STABILITY IN A FEEDBACK LOOP

In an ordinary (continuous-control) feedback system, the criterion of Nyquist applies to the frequency transfer function of the entire loop, i.e., to the transfer function that operates upon signals passing around the loop.

One must hold certain reservations when he uses the term "transfer function" in connection with a sampler, as the sampler is essentially a nonlinear device. What we have called the "transfer function" of a sampler certainly cannot be combined with those of a sequence of linear elements in the usual manner and directly subjected to Nyquist's criterion. This "transfer function" for the whole loop, so derived, however, can be used to generate a function having properties similar to those of the true transfer function for a series of cascaded linear networks. This function is characteristic of the particular time-quantized feedback loop under consideration, and to it the stability criterion of Nyquist can be directly applied as for a completely linear system.

This extended Nyquist criterion is developed by Linvill[3] and by Ragazzini and Zadeh.[5] These authors do not, however, explicitly recognize the "sampling oscillations" which occur with a period of $2/\nu_o$. In fact, Linvill seems to attribute such oscillations to nonlinearities in the output member of the system ("saturation effects"). The trouble, rather, arises in the sampler itself, and solely from the sampling process. Lago and Truxal recognize sampling oscillations, and they insist that the extended Nyquist criterion applies.[6]

The authors have felt certain doubts as to whether the extended criterion as so far put forth fully accounts for all of the troubles at half-sampling frequency. Therefore, in the section to follow there is a somewhat lengthy, but general, demonstration of the fact that it does. In a later section there is a derivation of the extended criterion from the point of view adopted with reference to sampling oscillations, and the criterion is stated in a manner readily interpretable from a physical (geometrical) standpoint, lending itself to a geometrical method for estimation of critical gain in practical applications. These interpretations are subsequently developed.

Oscillations of a Period $2/\nu_o$.

One might think that Nyquist's criterion could be extended to the time-quantized situation by mere inclusion of the interference circles as parts of the transfer function, which certainly they are. But it is not enough to state that self-sustaining oscillations will occur if a circle encloses the critical point.

In some cases a system with sampling will oscillate at values of gain materially lower than that required to cause enclosure of the critical point. In other cases, enclosure can occur without ill effects. Such oscillations would occur only at half the sampling frequency, because an oscillation at any odd multiple of $\nu_o/2$ will be reduced to frequency $\nu_o/2$ by the sampling process, and any signal whose frequency is an integral multiple of sampling frequency will be reduced by the sampling process to a train of identical (sampling) pulses.

The term "oscillation 'at' half-sampling frequency" suggests a looseness of usage which will here be convenient. It means simply that the output waveform is that associated with a monochromatic input of frequency $\nu_o/2$, and of course, the input is a periodic function having period $2/\nu_o$, with a waveform which depends upon the sampler pulse shape and upon the transfer properties of the remainder of the loop.

In general, existence of oscillations at a gain level other than that required to cause enclosure of the critical point by an interference circle is due to the waveform effect, or to put it another way, to the cumulative effects of interference in all of the various orders. Thus, the gain required to produce these oscillations is dependent upon the transfer function $W(\nu)$ and upon the loop transfer properties, because these literally determine how many orders are important contributors to the oscillation.

Consider a feedback loop from the point of view of Fig. 11. At A, the sampler input is f(t), having a line spectrum $f(\omega)$, whose line of lowest frequency is at $\nu_o/2$. At B, the sampler output is F(t), with spectrum $F(\omega)$, also a line spectrum. It will be assumed that the loop is in a barely oscillating condition, therefore the signals at A and B are related by the oscillation condition

$$f(\omega) = - KG(\omega)\, F(\omega), \tag{6.1}$$

where f and F are corresponding sampler input and output, in the notation of Section 3. Equation (6.1) represents the operation of $-KG(\omega)$ on each spectrum line (Fourier coefficient) of F(t), and implies that f(t) is the recombination of the resulting spectrum in a time function expressed by a Fourier series. Thus, if we have

$$F(t) = \sum_{m=-\infty}^{\infty} \lambda_m e^{\,im \frac{\omega_o}{2} t} \tag{6.2}$$

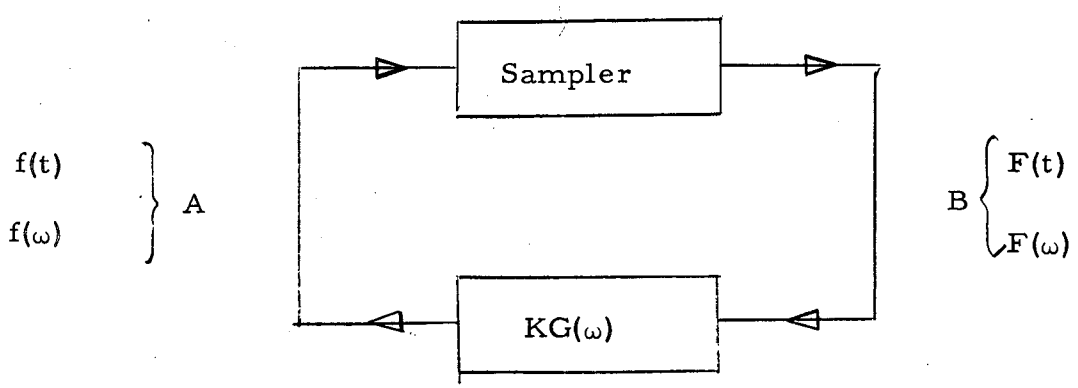

Fig. 11. Feedback loop containing a sampler and elements whose transfer properties can be described by the function $KG(\omega)$. K is the return-path gain; $G(\omega)$ is its frequency characteristic.

and

$$f(t) = \sum_{m=-\infty}^{\infty} c_m e^{im \frac{\omega_o}{2} t} , \tag{6.3}$$

then we obtain

$$c_m = - KG(m \frac{\omega_o}{2}) \lambda_m . \tag{6.4}$$

The λ_m can be determined from Eq. (5.7), which describes the sampler output corresponding to an arbitrarily shaped periodic input having twice the sampling period. Certain observations, however, can be made regarding the possible nature of the oscillations and can save much labor in the calculations. To start with, a zero-frequency component will not circulate around the loop, because feedback is by definition "inverse" to frequencies near zero. Furthermore, were the waveform unsymmetrical, it would possess components having frequencies that were even multiples of half the sampling frequency, and vice versa. Insofar as the results of the sampling process are concerned, however, any input having sampling frequency or any integral multiple thereof is tantamount to a "dc" input, which owing to the action of the feedback process, is suppressed. Thus only a symmetrical oscillation can exist in the steady state and therefore can contain only components whose frequencies are odd multiples of $\nu_o/2$. Thus, the most general form of the sampler output F(t) under the oscillation condition is given by the sum of the odd-order terms in Eq. (5.7), wherein the notation with respect to the summation indices k and m will be interchanged for the sake of convenience. The resulting expression is

$$F(t) = \sum_{m=-\infty}^{\infty} \sum_{k=1}^{\infty} c_{2k-1} \left\{ 1 + e^{2i\left[(2k-1)\phi - \theta_{2k-1}\right]} \right\} W\left[(2m-1)\frac{\omega_o}{2}\right] e^{i(2m-1)\frac{\omega_o}{2} t} . \tag{6.5}$$

Thus by comparing Eqs. (6.2) and (6.5), we have the expressions for the λ_m:

$$\lambda_{2m} = c_{2m} = 0 \tag{6.6a}$$

and

$$\lambda_{2m-1} = \sum_{k=1}^{\infty} c_{2k-1} \left\{ 1 + e^{2i\left[(2k-1)\phi - \theta_{2k-1}\right]} \right\} W\left[(2m-1)\frac{\omega_o}{2}\right] \tag{6.6b}$$

Then, from Eq. (6.4), we have

$$c_{2m-1} = - K \sum_{n=1}^{\infty} c_{2n-1} \left\{ 1 + e^{2i\left[(2n-1)\phi - \theta_{2n-1}\right]} \right\} W\left[(2m-1)\frac{\omega_o}{2}\right] G\left[(2m-1)\frac{\omega_o}{2}\right] \tag{6.7}$$

Now Eq. (6.5) gives F(t) in terms of input waveform parameters and sampling phase. We can also write F(t) in terms of sampler pulse shape and an arbitrary amplitude coefficient, and then apply Eq. (6.4) to get a relation between the c_m and c_{-m}, because the output waveform corresponding to an arbitrary, symmetrical, input of period $2/\nu_o$ is the same as that corresponding to some sinusoidal input whose amplitude depends somehow upon sampling phase and waveform. Thus, we use Eq. (3.7) with $\omega = \frac{\omega_o}{2}$, which when simplified becomes

$$F(t) = a \sum_{m=-\infty}^{\infty} W\left[(2m-1)\frac{\omega_o}{2}\right] e^{i(2m-1)\frac{\omega_o}{2}t} . \tag{6.8}$$

The coefficient, a, now is a function of ϕ and the θ_{2m-1}.

Now a value of a may be found that will make Eq. (6.8) identical with Eq. (6.5) for any waveform of the type under discussion. Comparing Eq. (6.8) with Eq. (6.2) we obtain

$$\lambda_{2m-1} = a\, W\left[(2m-1)\frac{\omega_o}{2}\right] , \tag{6.9}$$

and from (6.4),

$$c_{2m-1} = - a\, KG\left[(2m-1)\frac{\omega_o}{2}\right] W\left[(2m-1)\frac{\omega_o}{2}\right]. \tag{6.10}$$

Equating either Eq. (6.9) to Eq. (6.6b), or Eq. (6.10) to Eq. (6.7) gives

$$a = \sum_{k=1}^{\infty} c_{2k-1} \left\{ 1 + e^{2i\left[(2k-1)\phi - \theta_{2k-1}\right]} \right\}. \tag{6.11}$$

Now multiply Eq. (6.10) by $1 + \exp\left\{2i\left[(2m-1)\phi - \theta_{2m-1}\right]\right\}$, and sum on m. The member on the left is precisely a, yielding

$$-\frac{1}{K} = \sum_{m=1}^{\infty} \left\{ 1 + e^{2i\left[(2m-1)\phi - \theta_{2m-1}\right]} \right\} G\left[(2m-1)\frac{\omega_o}{2}\right] W\left[(2m-1)\frac{\omega_o}{2}\right]. \tag{6.12}$$

The quantities θ_{2m-1} are the phases of the Fourier coefficients of the sampler input. Using Eq. (6.10) once more, we have

$$e^{-2i\theta_{2m-1}} = \frac{c_{-(2m-1)}}{c_{2m-1}} = \frac{G\left[-(2m-1)\frac{\omega_o}{2}\right] W\left[-(2m-1)\frac{\omega_o}{2}\right]}{G\left[(2m-1)\frac{\omega_o}{2}\right] W\left[(2m-1)\frac{\omega_o}{2}\right]}. \tag{6.13}$$

In order for $G(\omega)$ to be the transfer function of a physically realizable network,[12] $G(-\omega)$ must equal $G^*(\omega)$. Moreover, from Eq. (3.4), $W(-\omega)$ equals $W^*(\omega)$ (the asterisk denotes the complex conjugate). Then, we obtain

$$\theta_{2m-1} = \arg G\left[(2m-1)\frac{\omega_o}{2}\right] + \arg W\left[(2m-1)\frac{\omega_o}{2}\right]. \tag{6.14}$$

Equation (6.12) is the condition for oscillation; to separate it into real and imaginary parts, Eq. (6.13) can be helpful:

$$-\frac{1}{K} = \sum_{m=1}^{\infty} \left\{ G\left[(2m-1)\frac{\omega_o}{2}\right] W\left[(2m-1)\frac{\omega_o}{2}\right] + e^{2i(2m-1)\phi} G\left[-(2m-1)\frac{\omega_o}{2}\right] W\left[-(2m-1)\frac{\omega_o}{2}\right] \right\} \tag{6.15}$$

---

[12] H. W. Bode, Network Analysis and Feedback Amplifier Design, (Van Nostrand, New York, N.Y. 1945), Sec. 7.3, pp. 106 ff.

Here G and W are each complex functions of the real variable ω. Letting X, Y, U, and V be real, even functions of ω, we write

$$G(\pm\omega) = X(\omega) \pm iY(\omega) \tag{6.16a}$$

and

$$W(\pm\omega) = U(\omega) \pm iV(\omega) \,. \tag{6.16b}$$

Therefore the real part of Eq. (6.15) is

$$\sum_{m=1}^{\infty} \left\{ \left[X_m U_m - Y_m V_m \right] \left[1 + \cos 2(2m-1)\phi \right] + \sin 2(2m-1)\phi \left[X_m V_m + Y_m U_m \right] \right\} = - \frac{1}{K} \,, \tag{6.17}$$

and the imaginary part is

$$\sum_{m=1}^{\infty} \left[X_m V_m + Y_m U_m \right] \left[1 - \cos 2(2m-1)\phi \right] + \sin 2(2m-1)\phi \left[X_m U_m - Y_m V_m \right] = 0 \,, \tag{6.18}$$

where the subscript m indicates evaluation of the function at $\omega = (2m-1)\omega_o/2$.

Equation (6.18) is satisfied by the condition $\phi = n\pi$, n being zero or a positive integer. If one looks momentarily at the oscillation process as a repetitive transient phenomenon, it is readily apparent that the phase of sampling, in the oscillation waveform, is an invariant of the system, as is the waveform at the sampler input; it cannot differ. For suppose that somehow sampling were to occur at a "different" phase. Then the transient response to the sampling pulse would occur as before (i.e., the sampler output pulse would be modified in the same way as before), the only possibly excepted aspect being the amplitude of the waveform. Sampling would occur after the usual interval $(1/\nu_o)$ and the same transient (modified sampling-pulse shape) would have occured in the interim. Therefore sampling phase was not "different"

after all. We are at liberty to reckon the zero of phase from this point of the waveform. Indeed, Eq. (6.18) dictates this, and the phase of sampling is consequently taken as zero. Eq. (6.17) reduces immediately to

$$\sum_{m=1}^{\infty} 2(X_m U_m - Y_m V_m) = -\frac{1}{K} \quad . \tag{6.19}$$

This is the criterion for stability that applies to oscillation at half the sampling frequency and that extends the criterion of Nyquist to the time-quantized feedback system.

The criterion of Eq. (6.19) has a simple and direct physical interpretation. The (normalized) loop transfer function is $G(\omega)\, W(\omega)$. If we refer to Eqs. (6.16a) and (6.16b), Eq. (6.19) is precisely

$$\sum_{m=1}^{\infty} 2\mathrm{Re}\left\{G(\omega_m)\, W(\omega_m)\right\} = -\frac{1}{K} \quad . \tag{6.20}$$

Thus, the cumulative algebraic sum of the real-axis intercepts of the interference circles must not reach the point $(1/K) - 0i$, or sampling oscillations will occur. The geometrical situation is shown in Fig. 12. It is to be noted that in the case chosen for Fig. 12, the sampling oscillations could not occur, because the loop gain must be a positive quantity, and because the sum of Eq. (6.20) is itself positive. Hence the oscillation condition can never be reached in this example. Therefore the question as to whether oscillation must occur if some interference circle should loop about the critical point is answered negatively because the waveform at sampler input is made up of the sum of contributions from all the orders. If one of these by itself were large enough to encircle the critical point, its effects would be modified by the cumulative sum of the contributions from all of the other orders. It is the resultant effect alone that can cause oscillations.

In the case of clamped sampling the application of Eq. (6.19) is made by recourse to Eq. (4.10), which in Eq. (6.16b) gives

$$U_m \equiv 0$$

and

$$V_m = -\frac{2}{\pi}\,\frac{1}{2m-1} \quad .$$

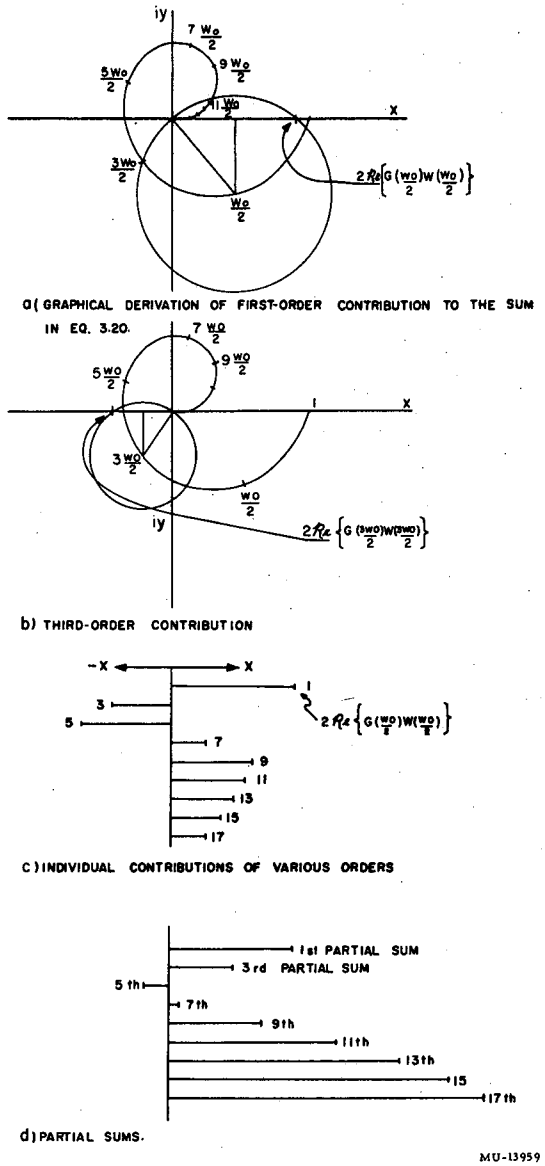

Fig. 12. Geometrical interpretation of the stability criterion of Eq. (6.20) for a hypothetical feedback loop having a frequency characteristic $G(\omega)\ W(\omega)$. Note (from d) that this hypothetical system could not possibly undergo sampling oscillations.

Consequently Eq. (6.19) becomes (by the use of the more customary notation in place of the subscripts)

$$\frac{4}{\pi} \sum_{m=1}^{\infty} \frac{1}{2m-1} \; Y\left[(2m-1)\,\frac{\omega_o}{2}\right] = -\frac{1}{K}\,, \qquad (6.21)$$

with Y being defined in Eq. (6.16a) as the imaginary part of the (normalized) loop frequency characteristic, and K being the "loop gain" (normalizing constant).

The impulse sampler is quite different. Here we have

$$W(\nu) = 1 \quad .$$

Hence it follows that

$$U_m = 1$$

and

$$V_m \equiv 0 \quad .$$

Equation (6.19) becomes

$$\sum_{m=1}^{\infty} 2X\left[(2m-1)\,\frac{\omega_o}{2}\right] = -\frac{1}{K} \quad . \qquad (6.22)$$

This is in interesting contrast to the clamped case.

General Case of Feedback Oscillations

The oscillation condition for the general case proves easier to derive than that for oscillations of period $2/\nu_o$. Once again, we refer to Fig. 11 and to Eq. (6.1), where ω is no longer considered to be necessarily related to $\omega_o/2$. The loop is considered to be closed at A, where in the oscillating condition the input equals the output, with reversed sign (feedback "inverse" for low frequencies). Because the oscillation (in order to be an oscillation) is periodic, Eqs. (3.1) and (3.3) hold with all the quantities $c_{ms} = 0$ excepting the set $c_{m1} = c_m$. Thus if we assume an oscillation having a period $2\pi/\omega$, these equations are written,

$$f(t) = \sum_{m=-\infty}^{\infty} c_m e^{im\omega t} \qquad (6.23)$$

and $$F(t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_m W(n\omega_o + m\omega)\, e^{i(n\omega_o + m\omega)t} . \qquad (6.24)$$

Because F(t) and f(t) are both real, c_{-m} equals c_m, and $W(-\omega)$ equals $W^*(\omega)$.

Now there are precisely as many lines in the spectrum of f(t) as there are in that of F(t). And to every line of F applies the condition of Eq. (6.1) to make up the corresponding line in f. Thus, we have

$$f(t) = -K \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_m GW(n\omega_o + m\omega)\, e^{i(n\omega_o + m\omega)t} , \qquad (6.25)$$

where $GW(\omega) \equiv G(\omega)\, W(\omega)$.

Following Linvill, we trace the line of frequency ω. In the sampling process f(t) becomes F(t), so

$$c_o\, e^{i\omega t} \text{ becomes } c_o\, W(\omega)\, e^{i\omega t} \text{ plus other terms.} \qquad (6.26)$$

Likewise, every component of f(t) having a frequency in the set $(n\omega_o + \omega)$ contributes a line of frequency ω in F as it passes through the sampler. Because these all arise from the $m = 0$ term (frequency ω), they have the coefficient c_o. Thus

$$-K\, c_o\, GW(n\omega_o + \omega)\, e^{i(n\omega_o + \omega)t} \qquad (6.27)$$

becomes

$$-K\, c_o\, GW(n\omega_o + \omega)\, W(\omega)\, e^{i\omega t}$$

plus other terms. The total amplitude of the term of frequency ω will be the sum of all of the contributions. Therefore we have

$$-K\, c_o\, W(\omega)\, e^{i\omega t} \sum_{n=-\infty}^{\infty} GW(n\omega_o + \omega) . \qquad (6.28)$$

This is the term of frequency ω in F(t). Equating this to the term value in expression (6.26) yields the oscillation condition for the (arbitrary) frequency ω:

$$1 + K \sum_{n=-\infty}^{\infty} GW(n\omega_o + \omega) = 0. \qquad (6.29)$$

The left-hand member of Eq. (6.29) is the denominator of Linvill's Eq. (15)[3] and is the stability-determining factor in all of Ragazzini and Zadeh's single-sampler systems listed in their Table II.[5]

Equation (6.29) is the general oscillation condition. To the function

$$K \sum_{n=-\infty}^{\infty} GW(n\omega_o + \omega) \qquad (6.30)$$

the criterion of Nyquist is applied in the customary manner. As will be presently seen, Eq. (6.29) includes Eq. (6.20) as a special case. Out of that demonstration will, moreover, come some interesting results regarding properties of Expression (6.30) .

The sum in Expression (6.30) is

$$S(\omega) = GW(\omega) + \sum_{n=1}^{\infty} GW(n\omega_o + \omega) + GW(-n\omega_o + \omega) \ .$$

Since $(-n\omega_o + \omega) = -(n\omega_o - \omega)$ and considering Eq. (6.16), we have

$$GW(-n\omega_o + \omega) = G^*W^*\left[-(n\omega_o - \omega)\right] = \overline{GW}^*\left[-(n\omega_o - \omega)\right] .$$

The result is

$$S(\omega) = GW(\omega) + \sum_{n=1}^{\infty} \left[GW(n\omega_o + \omega) + \overline{GW}^*(n\omega_o - \omega)\right], \qquad (6.31)$$

which is a complex function of the real variable, ω. Let $\omega = \omega_o/2$. Then in Eq. (6.31) the argument of the first term is $\omega_o/2$, that of the first term in the bracket is $(2n+1)\,\omega_o/2$, and that of the second bracketed term is $(2n-1)\omega_o/2$. If we consider the bracket as two separate sums, then the leading term of the second sum has the argument $\omega_o/2$, and the second term (n=2) has the argument $3\,\omega_o/2$, which is the same as that of the leading term of the first sum. Because n is a dummy variable, the argument of the general term in the first sum may be made $(2n-1)\,\omega_o/2$ and the summation range changed from n = 2 to ∞. Thus, all the terms in Eq. (6.31) combine to give

$$S\left(\frac{\omega_o}{2}\right) = \sum_{n=1}^{\infty} \left\{ GW\left[(2n-1)\frac{\omega_o}{2}\right] + \overline{GW}^*\left[(2n-1)\frac{\omega_o}{2}\right] \right\},$$

which is nothing more than

$$S = \sum_{n=1}^{\infty} 2\,\mathrm{Re}\; GW\left[(2n-1)\frac{\omega_o}{2}\right]. \qquad (6.32)$$

Thus $S\left(\frac{\omega_o}{2}\right)$ is identical with the sum function of Eq. (6.20).

The first interesting result alluded to above is that $S(\omega_o/2)$ is always real. Thus, if the imaginary part of $S(\omega)$ does not vanish at a frequency smaller than $\nu_o/2$, sampling oscillations are the only possible oscillations of the system. This is because oscillation is not possible at frequencies greater than $\nu_o/2$, owing to the translation property of the sampler output spectrum. The sampler output corresponding to an input having a period less than $2/\nu_o$ will have a spectrum line in the interval between frequency zero and $\nu_o/2$ because of the first difference. But this corresponds to an input of period $1/(\nu_o-\nu)$, greater than $2/\nu_o$. This signal will pass around the loop with its period unchanged (the loop is linear except for the sampler). Hence the oscillation can not have a period less than $2/\nu_o$. Thus, the sampling system can have no oscillation whose "fundamental" frequency (reciprocal period) is greater than half the sampling frequency. This considerably simplifies the task of evaluating $S(\omega)$, for we need to investigate only the region $o < \omega < \omega_o/2$, although this means evaluating the terms of S for all spectrum lines at frequencies for which G and W can be measured or computed.

7. EXPERIMENTAL VERIFICATION OF THE THEORY

The experiments were performed using a model sampler that consisted of a bilateral gate through which the signal from the input amplifier was impressed upon the storage capacitor during the gate time, approximately 4 μsec out of a sampling period of 417 μsec. Gating pulses were obtained from a multivibrator-pulse amplifier chain. Figure 13 is a diagram of the sampler.

For measurement of the transfer function the model sampler derived its input signal from a Hewlett-Packard model-650A signal generator operated at an output level of 0.10 rms. Sampler output was monitored on a Tektronix dc oscilloscope and amplitude measurements were made with a Hewlett-Packard model-300A spectrum analyzer.

In adjusting the data it was determined that by far the largest error was due to finite sampler impedance and that, at least in functional dependence upon frequency, any and all other effects contributed little to the errors. The impedance-correction function was determined by a least-squares reduction of supplementary data near half-sampling frequency where the effect was largest. In Fig. 14 are shown the data, as adjusted, compared with the function $\left| W(\nu/\nu_o \right|$ of Eq. (4.2). It is evident that the model sampler has a transfer characteristic closely approximating the ideal.

Early in the course of this work several spectral analysis experiments were performed. These were done before the theory as it now stands was available and were directed toward acquisition of knowledge regarding the nature of the sampling process. It was through these experiments that the sum-and-difference pattern of spectrum line frequencies was established experimentally.

There remained but one prediction of theory to be verified, namely the proposition that the same spectrum resulted from an input signal of unit amplitude having the frequency of any spectrum line. To prove this, data were taken with the same setup as was used for the measurement of the transfer function. The frequency $\nu/\nu_o = 0.10$ was chosen. Sampling frequency was 2400 cps, making the lowest frequency (largest amplitude) line occur at 240 cps.

The spectrum analyzer was adjusted on the lowest line, with sampler input signal amplitude of 0.15 rms. The spectrum analyzer was then tuned to the frequency of the line to be observed, and peaked. The spectrum line

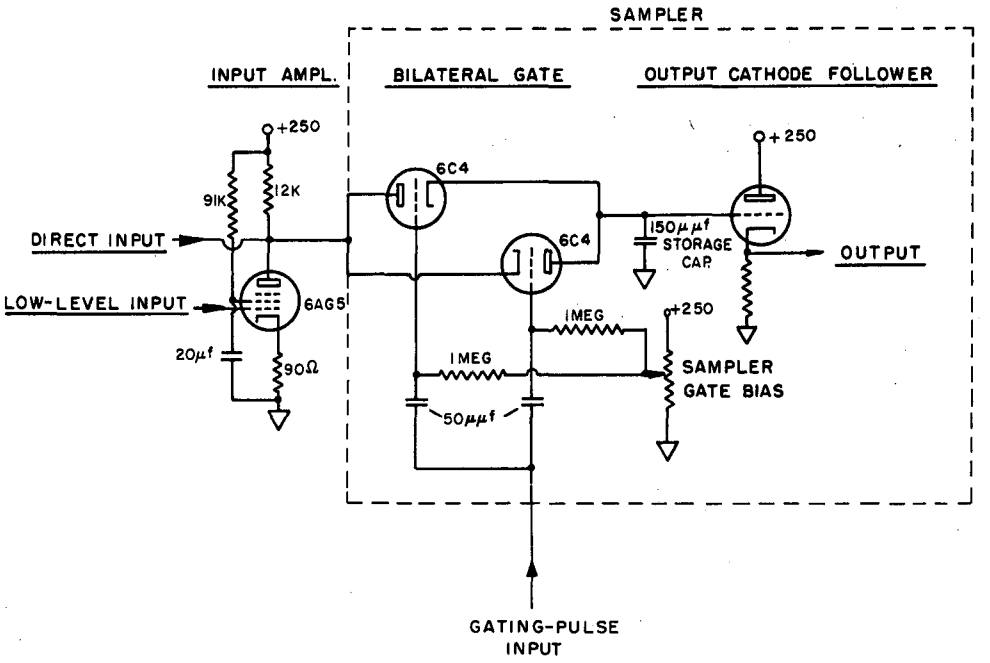

Fig. 13. Circuit diagram of model sampler used in experiments described in text.

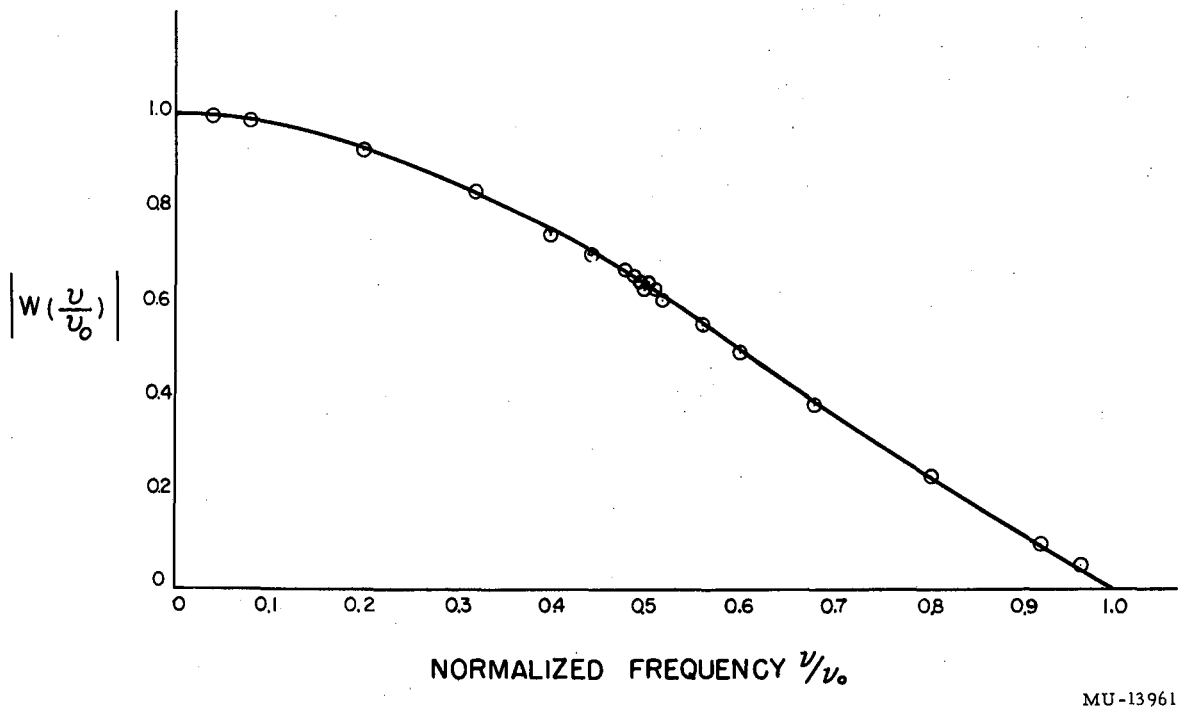

Fig. 14. Magnitude of transfer function of clamped sampler (Eq. 4.2), with experimental points measured from model sampler.

frequency was considered as a parameter, with input frequency the independent variable. The signal generator was tuned successively in increasing frequency to the frequency of each line in the spectrum, and the amplitude of the line being observed was read from the spectrum analyzer. At the start of each run (i.e., whenever the spectrum analyzer was tuned to a new line frequency) the signal generator and spectrum analyzer were renormalized on the lowest-frequency line.

This experiment resulted in verification of the prediction of Section 3. Interelectrode capacitance in the tubes of the bilateral gate contributed small deviations from predicted behavior in certain of the spectrum lines. Further experimental procedures sufficed to define this effect quantitatively.

Closed-Loop Stability Experiment

In order to verify the criterion for stability, an experiment was designed which would meet two needs: (a) to involve sufficiently simple networks in the loop as to make possible a calculation of the oscillation condition by transient analysis, thus providing an independent calculation with which to compare results both of theory and of experiment; (b) to reduce the effect of unaccountable variables to a minimum. Thus, there are derived three sets of gain values: (a) the results of application of Eq. (6.19), (b) the results of the transient-analysis calculation, and (c) the results of experiment. The simplest type of network suitable for attaining these objectives is one that involves a single energy-storage element. For convenience the capacitive phase-lag configuration was chosen. A schematic diagram of the apparatus is shown in Fig. 15. At the top of the figure are shown the sampler and the delay network, which was so designed that with convenient values of C and ν_o, three cases of RC could be measured; the phase shift at $\nu_o/2$ was to be approximately (a) $\pi/6$, (b) $\pi/3$, or (c) $5\pi/12$ (75^o). The values of phase shift used therefore corresponded roughly to normalized output amplitudes of (a) 1/2, (b) 1/3, and (c) 1/6. These values encompass a six-to-one range in bandwidth, and the results obtained illustrate the simplicity gained in the stability criterion with decreasing bandwidth. (Saying it another way, they illustrate the complication introduced into the stability criterion when the bandwidth is made near to or greater than half the sampling frequency.) The return-path amplifier was so designed as to have a much larger bandwidth than that of the delay network at the lowest value of C used. A switch

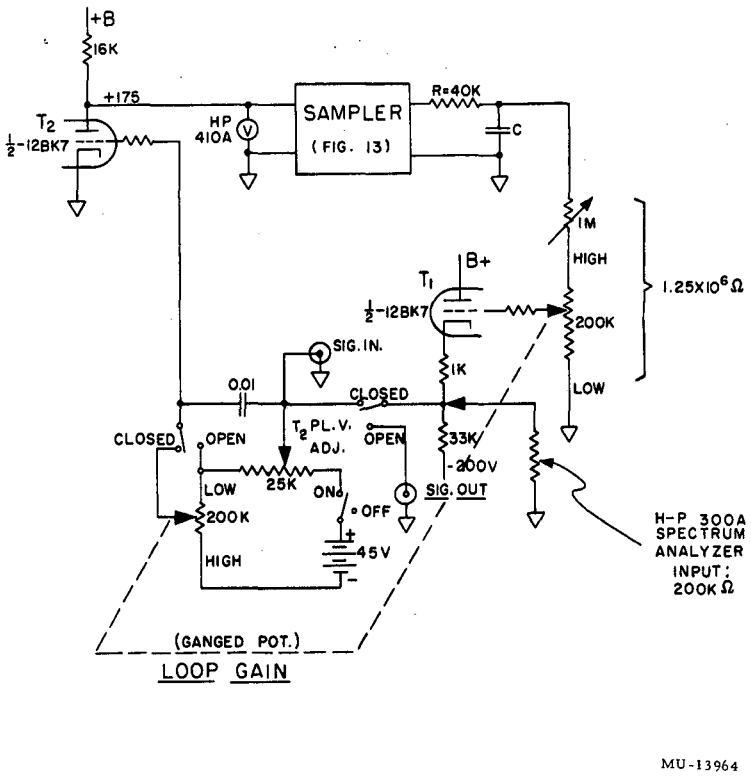

Fig. 15. Sampler - RC feedback apparatus, schematic diagram. Input of spectrum analyzer connected to "sig. in" when loop is closed, and to "sig. out" when loop is open. Effect is to have it shunted across the cathode-follower output at all times, as shown in diagram. Total (load) resistance shunting C is 1.25×10^6 ohms.

was provided in order that the loop might be opened or closed conveniently, with appropriate circuits arranged so as to maintain the necessary voltages at all points regardless of whether the loop was closed or open. Note that the Hewlett-Packard model 300A spectrum analyzer used in the gain measurements was kept connected across the cathode-follower output in the closed-loop condition. The gain measurements were made in the open-loop condition.

With each of the three values of the capacitor C, several gain measurements were made. Firstly, the loop was closed and voltages were adjusted to the proper values; loop gain was raised very slowly until oscillations were imminent. The loop was then opened, voltages were again read, and the gain was measured. Preceding each loop closure, the instrumental adjustments were checked carefully. The measurements of gain were made at $\nu_o/10$, because of spectrum-analyzer bandwidth limitation. Measurements were made also at $\nu_o/2$, for checking purposes. It was assumed that in correction to zero frequency of the $\nu_o/10$ values the transfer function of the idealized delay network could be used without significant error.

The transfer function of the apparatus in the frequency range of zero to ν_o was measured experimentally. This was done in order that the actual apparatus could be replaced in the calculations by an "ideal" model, one in which the parasitic effects were accounted for through the use of experimentally measured parameters.

Measurements of the values of R and C were made on a General Radio model-650A impedance bridge, the accuracy of which is stated by the manufacturer to be ±1% at the values in question.

An inspection of Fig. 15 will assure the reader that although to a first order the loop frequency characteristic is that of a simple RC delay network, a tangible second-order correction can be expected, particularly in the wider-band conditions. The most obvious effect is that of the output impedance, which was made purposely as large as feasible with minimum capacitance. Because the accountable shunting capacitance can be placed in the order of 10^{-12} farad, it will not be explicitly considered. The equivalent network is shown in Fig. 16. The transfer function may be written down immediately, and simplifies to

$$v(\omega) = \frac{1}{(1 + r) + iCR\omega} , \qquad (7.1)$$

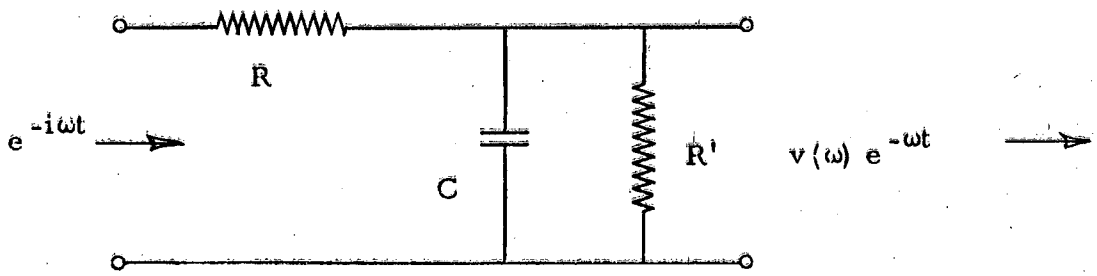

Fig. 16. RC delay network with output shunting resistance R'.

where $r = R/R'$. This can be compared with the case where R' is infinite ($r = 0$), where

$$v(\omega) = \frac{1}{1 + iCR\omega} \quad . \tag{7.2}$$

It is desirable to normalize the transfer functions to their zero-frequency values, making each the product of a gain constant and a frequency characteristic (normalized transfer function). This is done for Eq. (7.1) by multiplying by $(1 + r)$. Thus, we have

$$v_o(\omega) = \frac{1}{1 + i\,\sigma'} \quad , \tag{7.3}$$

in which

$$\sigma' = \frac{CR}{1 + r}\,\omega = \tau'\,\omega \quad , \tag{7.4}$$

whereas in the limiting case $r = 0$,

$$v_o(\omega) = v(\omega) = \frac{1}{1 + i\,\sigma} \quad , \tag{7.5}$$

where

$$\sigma = CR\omega = \tau\,\omega. \tag{7.6}$$

Therefore, it appears, the normalized transfer function of the shunted-output network is the same as that of the unshunted one, provided one uses a "corrected" resistance value $R/(1+r)$.

Another effect to consider is the cumulative effect of all of the unaccountable parameters in the system, which become particularly troublesome when the wider bandwidths are used. As is subsequently shown, even the use of the "corrected" resistance in the network calculation can lead to a demonstrably incorrect result. One way in which a correction may be introduced is to measure the transfer function experimentally and fit a set of parameters to a suitable model. In this case, the authors were unable to do more than to measure the amplitude of the transfer function for each value of C at nine values of frequency between zero and the sampling frequency. These values, when plotted, gave curves essentially similar to the amplitude of Eq. (7.3) or Eq. (7.5), lying somewhat below the calculated "model" curves. The experimental points were distributed about the average curve in what seemed to be a fairly random manner. It was assumed that the equivalent

network could be taken to be that of Fig. 16 without serious error, and that the time constants τ(or τ') should be calculated by taking the arithmetical average of the values of $\tau = \alpha/\omega$ obtained from the various amplitude measurements for each case. The values so obtained differed by very tangible amounts from the time constants of the idealized networks; the smaller τ (greater bandwidth) the greater the difference. This is as expected, the sizes of the discrepancies, however, are not comforting, and the assumption that the physical network differs from the "equivalent" one by only a "small" degree really generates only a first-order correction. One interesting aspect of the problem is that the differences in the time constants can be explained by increasing each of the values of C by approximately 4.5×10^{-10} farad, which seems to indicate that at frequencies less than the sampling frequency the "stray" effects appear as approximately 450 $\mu\mu$f of shunting capacitance across the output of the "equivalent" network. Certainly at higher frequencies, however, stray inductances come into the picture, and phase shifts certainly exceed $\pi/2$. Amplitudes also will be finite but small, and the slowly converging series that embodies the stability condition will not converge slowly enough to match the physical situation.

Critical Gain-Prediction by Transient Analysis

In view of the experimental limitations it is fortunate indeed that a network could be chosen whose properties make the feedback oscillations susceptible to simple analysis, completely independent of the theory of Section 6. Indeed, agreement between two such independent calculations makes a resort to experiment unnecessary for the sole purpose of confirming a theory. The experimental results in this instance do serve a highly useful purpose, that of illustrating the importance of stray effects in the time-quantized feedback system. This importance is quite out of proportion to that accorded strays in ordinary feedback practice.

The transient method is just that. One knows what the sampler-output waveform is, and, were one able to more readily solve a transient problem than to measure a transfer function, the need for a frequency-domain stability criterion would not exist. It is seldom, however, that one can make time-domain measurements and calculations with the ease and certainty of those in the frequency domain.

Figure 17 shows two waveforms in superposition. They are the sampler output (square wave) and the sampler input (nothing more than the square wave modified by the low-pass network and amplified). Consider now the time sequence of events. At some sampling instant, the network output has some value, $-y_o$. Sampling occurs, and the sampler output is clamped at its positive value, which we call unity. This is the input applied to the network, whose output momentarily has the value $-y_o$. The network output takes an exponential form, the asymptote being the network input (sampler output) value, unity. After the elapse of a sampling period, the sampler again samples and clamps, this time at some other value of voltage. The network output then begins to approach the new value exponentially.

It is, of course, understood that the loop is oscillating at half the sampling frequency, and therefore, both waveforms are symmetrical. Hence if the initial value of sampler input was $-y_o$, the alternate value is $+y_o$. Thus in any sampling period we have

$$2nT < t < (2n + 1)\,T, \quad y(t) = y_1(t)$$

the output voltage of the network. In the next period, we have

$$(2n + 1)\,T < t < 2(n + 1)T, \quad y(t) = y_2(t)\,.$$

The sampler output oscillates between +1 and -1. Then (see Fig. 17) we write

$$y_1(t) = 1 - (1 + y_o)\, e^{-\frac{t}{\tau}} \tag{7.7}$$

and likewise

$$y_2(t) = (1 + y_o)\, e^{-\frac{t}{\tau}} - 1\,. \tag{7.8}$$

Now we can equate

$$y_1(T) = y_2(0) = y_o \tag{7.9}$$

and

$$y_2(T) = y_1(0) = -y_o\,. \tag{7.10}$$

Combining Eqs. (7.7) and (7.9), or Eqs. (7.8) and (7.10), we obtain

$$y_o\left(1 + e^{-\frac{T}{\tau}}\right) = 1 - e^{-\frac{T}{\tau}}\,.$$

The gain required to amplify the function $y(t)$ to unit amplitude is just $1/y_o$. Thus, we write

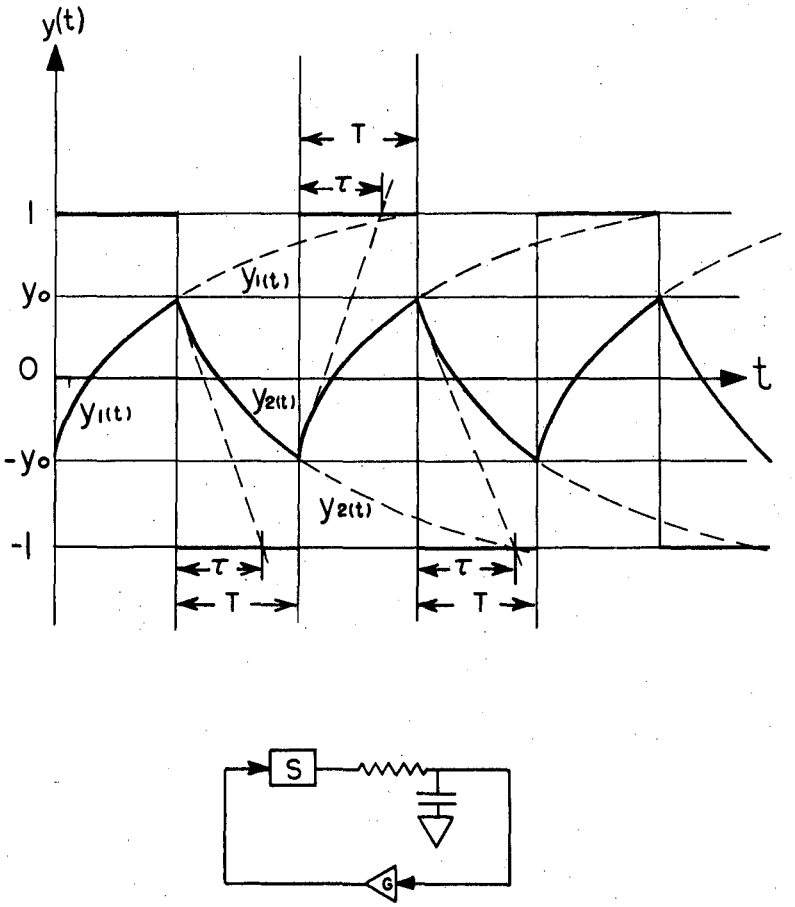

Fig. 17. Oscillations at half the sampling frequency in a feedback loop whose frequency properties are determined by a simple RC low-pass filter section.

$$G(T, \tau) = \frac{1 + e^{-\frac{T}{\tau}}}{1 + e^{-\frac{T}{\tau}}} \quad . \tag{7.11}$$

Application of the Stability Criterion

Application of the theory of Section 6 is very simply made. Using the notation of that section and Eqs. (7.3) and (7.4), or Eqs. (7.5) and (7.6), we have

$$G(\omega) = \frac{1}{1 + i\tau\omega} \quad , \tag{7.12}$$

$$X(\omega) = \frac{1}{1 + (\tau\omega)^2} \quad , \tag{7.13}$$

and

$$Y(\omega) = \frac{-\tau\omega}{1 + (\tau\omega)^2} \quad . \tag{7.14}$$

From this, we obtain

$$Y\left(\left[2m-1\right]\frac{\omega_o}{2}\right) = \frac{-(2m-1)\frac{\tau\omega_o}{2}}{1 + \left[(2m-1)\frac{\tau\omega_o}{2}\right]^2} \quad . \tag{7.15}$$

Hence Eq. (6.21) becomes

$$\frac{1}{K} = \frac{4}{\pi}\left(\frac{2}{\tau\omega_o}\right)\sum_{m=1}^{\infty}\frac{1}{\left(\frac{2}{\tau\omega_o}\right)^2 + (2m-1)^2} \quad , \tag{7.16}$$

a very slowly converging series indeed.

By direct application of a method due to Gumowski[13] the series of Eq. (7.16) can be accurately summed with the expenditure of a very modest effort, whereas, were the summation to be made term by term, calculations

[13] I. Gumowski, Summation of Slowly Converging Series, Letter in J. Appl. Phys. 24, 1068 (1953).

show that several hundred terms would be necessary. The results shown in the subsequent subsection as calculated from Eq. (7.16) were obtained in this way.

Tabulation and Comparison of Results

In Fig. 18 are represented the three experimental cases, each being characterized by a particular value of C. All three transfer functions are superposed, through the use of the dimensionless frequency variable, δ. The effect of strays is immediately apparent, and one would find it difficult to avoid doubting the realism of the idealized transfer function at values of σ several times the sampling value. The problem is particularly acute when one considers the cumulative importance of the higher orders, as is indicated by the slowness of convergence of the series of Eq. (7.16).

In Table I comparison is made between time-constant values, from which a low frequency " equivalent" shunting capacitance was calculated. The value of this capacitance seems remarkably consistent through the three cases, though the time constants cover an order of magnitude in range. In Table II are exhibited for comparison the variously derived values of loop gain required to produce oscillations. As a standard of comparison we took the values calculated from transient analysis. The theory, as exemplified in Eq. (7.16), agrees with these perfectly. Because the networks used in the two calculations were the same (ideal networks), and because the transient-analysis approach must produce the correct value, one is justified in stating that the theory has met its test.

The experimental results are listed in the fourth and fifth columns of Table II, the former column being the gain values and the latter column being a comparison with the calculated values. It is instructive to compare these numbers with reference to the approximate bandwidths listed in column 6. Note that in case 3, where the bandwidth is approximately $\nu_o/8$, the strays are quite unimportant, whereas when the bandwidth is little more than doubled, the importance of the strays (through the higher orders of the spectrum) has increased more than five-fold. Some light into the specific nature of the effect that the strays had upon the experimental results is shed by Fig. 19, which shows the transfer function of the entire feedback loop for Case 1 (neglecting strays, excepting that the measured time constants are used). The further

Table I. Parameters of network "equivalent" of feedback loop in stability experiment

Case	C measured (farad)	τ calculated (sec)	τ measured (sec)	Difference (%)	Shunt capacity "equivalent" of difference (farad)
1	1.740×10^{-9}	6.744×10^{-5}	8.428×10^{-5}	+25	4.4×10^{-10}
2	5.432×10^{-9}	2.105×10^{-4}	2.266×10^{-4}	+ 7.6	4.1×10^{-10}
3	1.45×10^{-8}	5.620×10^{-4}	5.794×10^{-4}	+ 3.1	4.5×10^{-10}

Table II. Comparison of variously derived values of gain required to produce sampling oscillations

Case	Gain required to produce oscillations				
	transient (Eq. 7.11)	theory (Eq. 7.16)	experimental	$\frac{\text{meas.}}{\text{calc.}}-1$	Bandwidth (approximate) (sec^{-1})
1	1.058	1.058	1.24	+17.2%	$\frac{3}{2}$ $(\nu_o/2)$
2	1.438	1.438	1.56	+ 8.5%	$\frac{3}{5}$ $(\nu_o/2)$
3	3.028	3.028	3.07	+ 1.4%	$\frac{1}{4}$ $(\nu_o/2)$

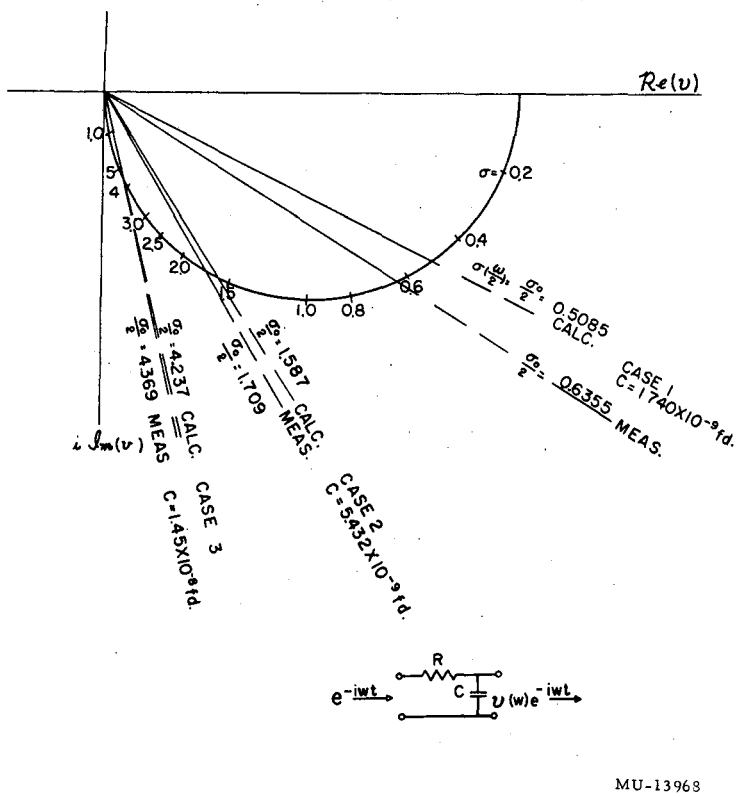

Fig. 18. Transfer function of RC delay network, with frequency shown in terms of the dimensionless variable $\sigma=\tau\omega$. Shown here are the three cases for which critical closed-loop gain measurements were made at half the sampling frequency.

Note: Calculated values include only the RC network plus its load resistance; measured values include entire apparatus. For technique and remarks re. limitations of the method, see text.

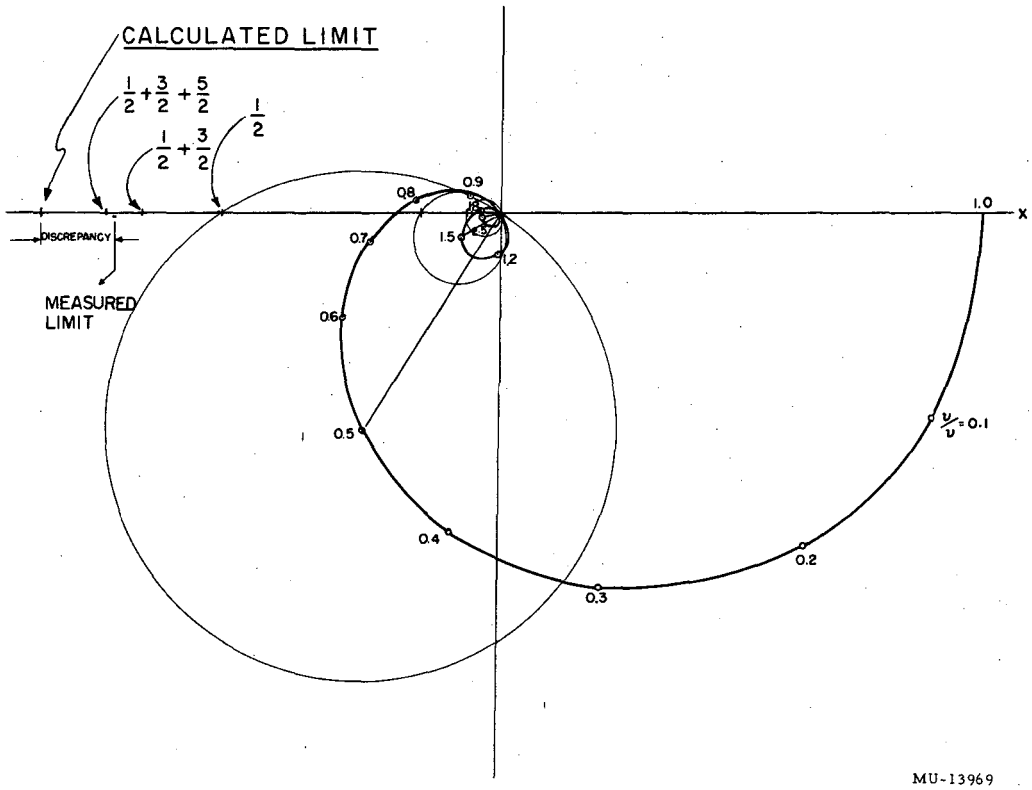

Fig. 19. Transfer function of feedback loop, case 1 (ideal network, measured value of τ) showing partial sums to order 5/2, and limits.

effects of strays can be expected to arise primarily in the influence of lead inductances, which will bring about additional phase shifts in the higher orders. In Fig. 19 the interference circles for the first three orders (1, 3, and 5) are shown along with the partial sums of their intercepts. Also shown are the limits. Note that the measured value of critical gain is less than the calculated value. Note further that the difference is roughly equal to the sum of all the orders above the fifth, and also that the fifth order could have considerable additional phase shift without affecting the results materially. (The results are insensitive to a relatively small added phase shift in the fifth order, but any phase shift will be greatly magnified in the higher orders, and the contribution of any given order might well be zero, or positive.) Indeed one might conclude from the gain figures listed, and from the above, that the net effect of all of the higher orders is close to zero, indicating that about as many of the contributions were positive as were negative, or that the high-order parts of the transfer function were "wrapped" around the origin by the effect of strays, instead of all lying on the same (negative) side.

8. CONCLUSIONS

In the discussion concerning the properties of the output spectrum of a sampler, it was shown that a sampler can be regarded as having a transfer function, subject to certain definite limitations. This is due to the fact that the input spectrum appears in the output, with its components modified in amplitude and phase. The sampling process is an essentially nonlinear one in that new spectrum lines are generated, but it was shown that no interaction effects occur between input components.

The transfer function is continuous excepting at multiples of half the sampling frequency. The existence of such a transfer function makes possible the use of Nyquist's criterion for stability, so extended as to account for the complexity of the waveform of feedback oscillations, when they exist. The peculiar behavior of the transfer function at multiples of $\nu_o/2$ is due to interference between the spectral components of those frequencies.

Because previous publications in this field have included no more than an occasional mention of sampling oscillations, the subject was treated in detail here. The extension of Nyquist's stability criterion implied by Linvill was shown to include the case of sampling oscillations.

This extension of Nyquist's criterion is embodied in the application of the criterion in the customary manner, not to the loop-transfer function, but to a sum function (see Eq. 6.29), which embodies the contributions to the sampler input waveform of all the spectral orders in the sampler output, as modified by the loop transfer function. It was seen that the sum function evaluated at half-sampling frequency is always real, with the implication that if this sum function does not become real (and negative) at any frequency below half sampling, the only oscillations that the system can undergo are the sampling oscillations. No oscillations can ever occur with a period shorter than twice the sampling period, because of the transposition of spectra in the sampling process. Thus, in stability calculations, no base frequencies above half sampling need be used. However, terms for the sum function must be calculated (or scaled graphically) from the transfer function of the loop (including sampler) for spectral orders as high as can be calculated (or scaled) considering the number of significant figures in the calculation. Indeed, the sum function may be a very slowly converging series, in which case any graphical estimate of critical gain is a rough one and must include some sensible accounting of high-frequency "stray" effects.

When the loop is subject to sampling oscillations, the critical gain can be roughly estimated by simple graphical means. The procedure is as follows (Fig. 19): (a) plot the transfer function of the entire loop, including sampler; (b) place the compass point on the point of the transfer function corresponding to the frequency $m\,\omega_o/2$ (m an odd integer); (c) set the compass so as to pass a circle through the origin; (d) strike an arc across the real axis; (e) repeat for every order for which amplitude of the transfer function is sufficient to allow the construction; (f) add these algebraically (easily done with a scale); (g) mark off the sum on the real axis. That is the critical point for sampling oscillations. If it is on the positive real axis, or if it lies on the negative real axis inside the continuous portion of the transfer function, sampling oscillations cannot occur, and the conventional Nyquist criterion provides the limitation. Note that the estimate is only that, because the cumulative effect of the neglected orders may be quite noticeable.

9. ACKNOWLEDGMENTS

The authors wish to acknowledge with gratitude the contribution of their friend and colleague, Dr. Bob H. Smith, who kindly consented to read the manuscript, and the encouragement given by George M. Farly during the course of this work.

This work was performed under the auspices of the U.S. Atomic Energy Commission.