UNITED STATES DEPARTMENT OF COMMERCE
Alexander B. Trowbridge, Secretary
NATIONAL BUREAU OF STANDARDS • A. V. Astin, Director

NBS TECHNICAL NOTE 357

ISSUED NOVEMBER 2, 1967

SIGNAL DESIGN FOR TIME DISSEMINATION: SOME ASPECTS

J. L. JESPERSEN

Radio Standards Laboratory
Institute for Basic Standards
National Bureau of Standards
Boulder, Colorado 80302

NBS Technical Notes are designed to supplement the Bureau's regular publications program. They provide a means for making available scientific data that are of transient or limited interest. Technical Notes may be listed or referred to in the open literature.

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 20402
Price: 30 cents.

TABLE OF CONTENTS

SIGNAL DESIGN FOR TIME DISSEMINATION: SOME ASPECTS

J. L. Jespersen

The purpose of this paper is to discuss, in a general way, some problems of time signal dissemination in a noisy environment. Most of the paper applies to any timing system, but particular emphasis is given to a CW two-frequency system. This is done for two reasons: first, as will be shown, a two-frequency CW system evolves naturally from fundamental considerations to meet certain user requirements; and second, a two-frequency VLF system is being investigated experimentally at the present time.

Key Words: VHF, HF, VLF, time and frequency dissemination, synchronization, satellite, noise.

1. Introduction

A timing signal usually consists of some waveform which is repeated at uniform intervals of time, ΔT, as illustrated in figure 1. To know the time means knowing how many units and fractional units of ΔT have elapsed from some starting point T_E (see fig. 1). In most timing systems, as illustrated in figure 1, there is some finite path delay, δt, between the time the signal is broadcast and the time it is received. Although determining the path delay is not a trivial problem, it will be assumed known in this paper.

In the design of signals for time dissemination, it is important to know whether the user is primarily interested in obtaining T_E or ΔT or both from the timing signal. In addition, these determinations must be made in the presence of noise, which will, for example, cause the timing signal arrival time to jitter as illustrated by the double arrow in

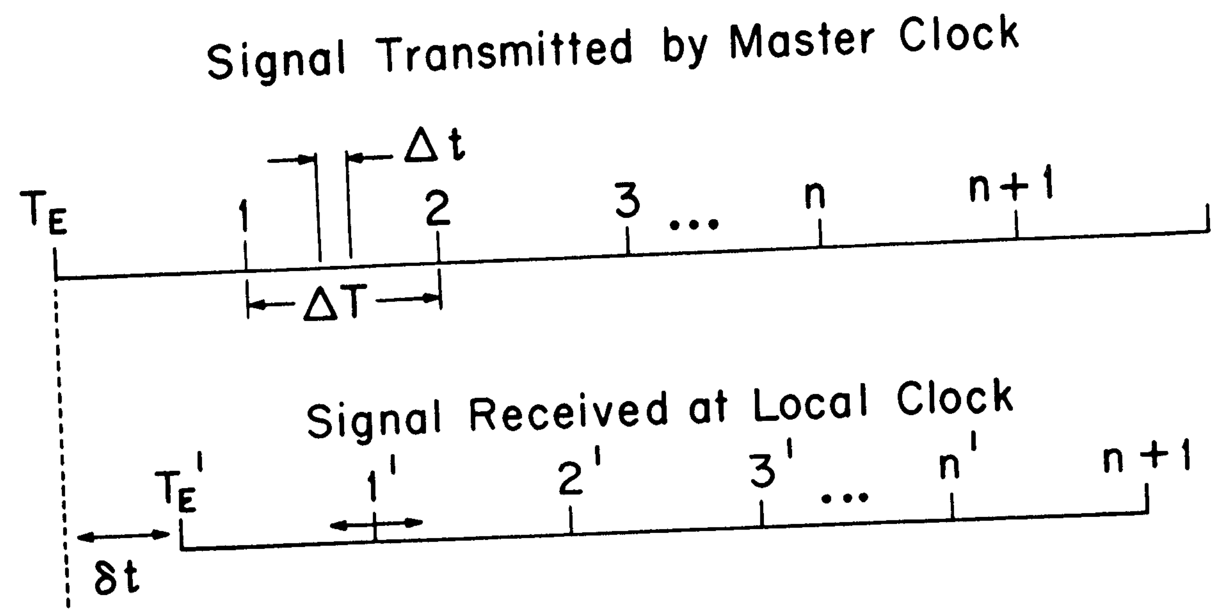

Figure 1. Elements of a timing system.

figure 1. Although there is a wide variety of noise which may be encountered in actual practice, we will, for the most part, assume noise with a uniform spectral density N_0 (white noise).

Although we will be considering various aspects of the timing signal design in some detail in later sections, it will be useful at this point to consider a simple example which will serve to illustrate the kind of problems which will be encountered.

Consider a timing system which consists of a single pulse of CW signal at frequency f and duration q, which we know ahead of time is to designate a certain moment in time with respect to T_E. The standard deviation in the time of arrival of the leading edge of this signal is [Skolnik, 1962]

$$\sigma(t) = \frac{t_r}{(2S/N)^{1/2}}$$

where t_r is the pulse rise time and S/N is the pulse signal-to-noise power ratio. The bandwidth occupied by the signal is given approximately by $1/t_r$. If we decreased the pulse length q while maintaining the same energy in the pulse, we could better our measurement of the time of arrival, since we have increased the pulse S/N. On the other hand, since the signal consists of a single pulse, we have no way of obtaining interval information, ΔT, unless we measure the carrier frequency, f, of the pulse. The standard deviation of the carrier frequency measurement σ (f) is a function of the reciprocal of the duration, q, of the pulse. (See Section 3.3) Therefore, we see that for a single pulse the time of arrival measurement and the frequency measurement are related, and we can only improve one measurement at the expense of the other.

Suppose now that we have a single pulse of energy which lasts for q seconds, and we find that we can measure the frequency of the carrier to sufficient precision for our needs. Is there any way in which we could improve our time of arrival measurement without sacrificing the frequency measurement? The time of arrival measurement depends upon the S/N at the leading edge of the pulse and does not depend upon how long the pulse lasts. This suggests that we could improve our measurement of arrival time if we could have more leading edges. Thus, suppose we took our pulse of duration q and signal level S and divided it into m pulses each of duration q/m with the same signal level S. Our total signal now consists of m pulses but the duration of these m pulses is still q seconds so that we haven't changed the precision of our frequency measurement. On the other hand, we now have m leading edges to measure so that we have decreased the standard deviation of our arrival time measurement by a factor $(m)^{-1/2}$, if each of the m measurements are independent. This implies that the time between successive pulses is at least 1/B seconds where B is bandwidth; otherwise, the noise fluctuations on successive pulses will not be independent. We have now improved our time of arrival measurement without damaging the frequency measurement. But in doing so, we have introduced another difficulty. Unless we have some additional information, we no longer know which one of the m pulses represents the time. Thus, for the privilege of improving our time of arrival measurement, while maintaining the same frequency error, we have introduced ambiguity.

Although we have introduced the idea of several pulses from the point of view of obtaining precise frequency and time of arrival measurements simultaneously, there is another important practical reason for having timing signals available frequently and that is simply that the user does not wish to wait long periods of time for timing information. Again, however, the penalty for frequent timing signals is ambiguity, unless means are taken to keep the signals distinguishable.

However, to keep the signals distinguishable, we must also pay a price. To illustrate how this comes about, consider a timing system which consists of one per second pulses. In this system no user has to wait more than one second for a second marker; however, if the user does not already know the time to the nearest second, the system is ambiguous for him. Presumably we could make the system unambiguous for our user by going to a lower pulse rate, but in doing so we would make the users who already knew the time to the nearest second unhappy; because,now they have to wait longer for a timing mark. If we wanted to maintain the time markers at the one per second rate and still satisfy the user who didn't know the time to the nearest second, we would have to provide additional information to remove the ambiguity. (WWV uses such a system where the basic pulse train is supplemented by voice announcements and omissions of certain pulses.)

If only a certain amount of energy is available, the energy used for identification (to reduce the signal ambiguity) will result in reduced knowledge of the time of arrival of the pulses. This is attributable to the lowered signal-to-noise power ratio, S/N. That is, the timing signal design has traded measurement precision for reduced ambiguity. This trade-off can be carried too far. The user of a timing pulse occurring each second may wish certain pulses to be omitted (more identification). A second user may be satisfied more quickly by leaving in all the pulses to improve his measurement precision. Thus, a signal can be too precise (ambiguous) for one user and overidentified for another.

A timing signal may be used simultaneously for both "time of arrival" and identification. We could transmit a signal consisting of a group of coded pulses. One of the pulses in the group could be designated as the synchronization pulse. It would then appear that both synchronization and identification have been satisfied. It will be shown later, however, that optimum determination of the time of arrival of the signal requires a priori knowledge of the signal shape. But if we already know the shape, then the signal does not contain any new identification information; thus, we cannot optimumly determine the signal arrival time and simultaneously gain identification information. If, however, the incoming signal were one of a number of possible signals, then we could gain information by using a number of parallel decoders optimized for each of the possible signals.

There does not appear to be any general method for designing signals which will simultaneously satisfy the requirements of ambiguity, time of arrival, and interval (or frequency) measurement precision. It is intuitively obvious that the reduction of ambiguity will require a signal whose frequency components are spread throughout the available bandwidth. However, for the case of increased precision in time of arrival measurement, we will see that the frequency components should be concentrated near the edges of the available bandwidth, whereas for increased precision in the frequency (or interval) measurements, the signal should last as long as possible.

2. Information Theory and Time Dissemination

Before considering in detail the problems of ambiguity, time of arrival and frequency measurement precision, it will be useful, using information theory, to determine how well the optimum system should operate for a given bandwidth, B, signal-to-noise power ratio, S/N, and observing time, t.

Ordinarily, the user does not need a signal which completely specifies the time, but rather he already knows the time down to some interval, ΔT, and wishes to know the time down to some smaller interval, Δt. For example, he may know the time to the nearest hour ($\Delta T = 1$ hour) and wish to know the time to the nearest minute ($\Delta t = 1$ minute). If we assume that the correct time is equally likely at all points within the interval ΔT, then the time has equal probability of being in any one of k intervals where

$$k = \Delta T/\Delta t.$$

The amount of information in bits required to designate one of the k intervals is

$$I = \log_2 k = \log_2 \Delta T/\Delta t. \tag{1}$$

A system which is optimum with a bandwidth B, and a signal-to-noise power ratio S/N can transmit information at a maximum rate [Schwartz, 1959] given by

$$R = B \log_2(1+S/N).$$

The total amount of information transmitted during a time t will be

$$I = t B \log_2 (1 + S/N). \tag{2}$$

Equating (1) and (2) we see that the improvement factor $\frac{\Delta T}{\Delta t}$, which may be obtained with the optimum system with a given B, S/N and observing time t is

$$\Delta T/\Delta t = (1 + S/N)^{+tB}. \tag{3}$$

To illustrate this result, consider a timing system which consists of a sine wave at frequency f. The standard deviation, $\sigma(t)$, in the time of arrival of the zero-crossings of the sine wave is

$$\sigma(t) = \frac{1}{2\pi f (2\ S/N)^{1/2}} \qquad \text{(see Section 3.2, Equation 7)} \tag{4}$$

if $S/N >> 1$. [Skolnik, 1962; also see Section 2 for further discussion of the sine wave system.] In this system it is assumed that the user knows the time to the period of the sine wave (or less) so that

$$\Delta T = 1/f.$$

In addition, the smallest interval which may be distinguished by the user is $\sigma(t)$; so that, $\Delta t = \sigma(t)$. Therefore

$$\left(\frac{\Delta T}{\Delta t}\right)_{\text{sine wave}} = 2\pi\,(S/N)^{1/2}.$$

In order to compare this result with the optimum system, we will assume that $tB = 1$. This condition implies that the sine wave system bandwidth is just wide enough to accommodate the spectral width of a sine wave which has a duration of t seconds. With this assumption, and with the assumption already made that $S/N >> 1$, (3) becomes approximately

$$\frac{\Delta T}{\Delta t} = S/N.$$

Thus the optimum system is a factor $(S/N)^{1/2}/2\pi$ better than the sine wave system.

If we increased B for the sine wave case and kept the same observing time, we would not improve our result since we would be obtaining more independent samples per unit time on the one hand and letting in more noise on the other. This is easily shown from (4) since independent samples may only be obtained every $\frac{1}{B}$ seconds. We can, however, improve our result for the optimum case by increasing B. To illustrate this, assume that the noise is white with spectral density N_o: then $N = N_o B$. As we increase B, the term $S/N_o B$ becomes small compared to 1 and we can approximate (2) by

$$tB \log(1 + S/N_o B) \cong (t)\,1.45\,S/N_o$$

which allows us to rewrite (3) as

$$\frac{\Delta T}{\Delta t} = 2^{t(1.45\,S/N_0)}.$$

A signal which approached this limit might consist of a long sequence of pulses almost completely masked by noise. The information could be recovered by multiplying the incoming signal, pulse by pulse, by a locally-stored replica of the incoming signal. However, for timing we are generally interested in signals with $S/N \gg 1$ which are readily received by simple equipment. Nevertheless, low signal-to-noise signals with large bandwidths are of interest in specialized cases, as might be encountered in transmitting time to a distant space craft.

Information theory does not tell us how to design an optimum system. In most systems there is a high degree of redundancy. But in systems approaching the optimum, there can be little redundancy. In fact, systems which approach the optimum transmit only an error signal [Pierce and Cutler, 1959]. From a timing point of view, this means that the user estimates the time as well as he can from the preceding signal and asks that the future signal contain only such information as will reduce his error. Since no two users of the same timing signal are likely to have the same errors as a function of time, the optimum signal for the two users quickly diverges. Such a system would be difficult and expensive to provide for a large number of users. Nevertheless, there are some instances in time dissemination where this kind of procedure might be approached. These systems might be classified as "active" systems where the user has some way of communicating with the signal source. The meteor trail [Gatterer, 1967] and satellite-transponder [Markowitz, et al., 1966] timing systems are possible examples. In these systems, a signal may be transmitted from the master to slave clock and back again, allowing the propagation delay and the slave clock error to be determined. At this point, the signals from the master station could be advanced in time so as to arrive on time at the slave station. Thus, in a sense, a signal is being transmitted which is tailored to the users' specific needs.

3. Bandwidth and the Time of Arrival Problem

In a recent CCIR document [1966] it is stated:

> "There is an increasing number of applications requiring the use of very precise time reference. In an effort to achieve greater precision, it is expected that the system of time distribution will tend to make use of an increased bandwidth up to the limits imposed by the allocated band, the instabilities of the propagation path, or by consideration of noise and interference."

The truth of this statement is well demonstrated by (3). Nevertheless, as stated previously in instances where the problem is essentially "time of arrival" rather than identification, what is required for an optimum time of arrival determination is that the signal be concentrated in a narrow frequency range which implies reduced rather than increased bandwidth B.

To further amplify this point, consider the similarity between the problem of using a timing signal for synchronization and the problem of detecting and using a radar echo. To demonstrate this similarity, consider the steps which are required in the radar operation:

1. Transmit a signal at some known time.
2. Listen for the return echo.
3. If one is present, measure its arrival time so that the range to the reflecting object may be determined.
4. Determine the carrier frequency of the echo.

Steps 2 and 3 must also be taken by the user of the timing signal. In addition, if the user needs the time interval as well as the epoch, he may want to take step 4.

People interested in the radar problem have devoted considerable effort to the problem of radar signal design so the required steps can be taken efficiently. Asked simply, for a given energy and noise environment, what signal shape allows the four operations listed above to be accomplished with the least uncertainty? Approaches have been developed to answer the signal design problem for radar and have obvious application to the design of timing signals. We will consider these steps in detail now.

3.1 Recognition of the Presence of the Signal

Increasing the amount of energy in the signal for a given noise environment will certainly enhance the chances of recognizing the presence of the signal. However, for a given amount of energy, might it be easier to recognize the signal if it had some distinctive shape?

To make meaningful comparison of various shapes requires that the means of detection be specified. A device known to maximize the output S/N is the so-called "matched filter". We will assume in the signal comparison that a matched filter is always used.

Suppose that the input signal s(t) has a voltage spectrum S(f), that is, S(f) is the Fourier transform of s(t). Further, let H(f) be the frequency-response function of the filter (not necessarily a matched filter), and let N_o be the input noise power per unit bandwidth. Then it may be shown [Skolnik, 1962] that the signal-to-noise power ratio R out of the filter is given by

$$R \le \frac{\left|\int_{-\infty}^{\infty} S(f)\, H(f)\, \exp\,(j\, 2\pi\, f\, t)\, df\right|^2}{\frac{N_o}{2} \int_{-\infty}^{\infty} |H(f)|^2\, df} . \tag{5}$$

If $H(f) = S^*(f) \exp(-j 2 f t_1)$ (* means complex conjugate), then the equality sign in the above equation holds. That is, H(f) is now a matched filter for s(t). Since

$$\int_{-\infty}^{\infty} |S(f)|^2 \, df = E, \text{ the total energy,}$$

we obtain for the optimum case

$$R = \frac{2E}{N_o} .$$

The important thing to note about this result is that it is independent of the signal shape S(f). That is, as long as we use a matched filter, only the energy in the signal and the noise determine our ability to recognize the presence of the signal. Note, however, that (5), which gives the expression for the matched filter, involves the signal spectrum. This is not surprising since we expect to have a better chance of recognizing the signal if we know in advance what it looks like.

It is relatively difficult to build matched filters for signals with complicated spectra. On the other hand, the detection of CW signals by phase-lock receivers closely approaches a matched filter under certain conditions. Phase-lock receivers are not usually described in terms of matched filters, but rather are considered in terms of cross-correlation where the incoming signal is multiplied by a replica of itself. It may be shown [Skolnik, 1962], however, that the output of a matched filter is proportional to the input signal multiplied by a replica, thus indicating the equivalence between matched filters and cross-correlation or phase-lock detectors.

3.2 Measurement of Signal Arrival Time

We have seen in the previous discussion that the signal shape is of no importance in determining its presence when a matched filter is used. Let us assume now that the signal-to-noise ratio is large and we are using a matched filter. Then, does our ability to measure the arrival time depend upon the signal shape? There are several approaches to this problem. Using a statistical approach, Berkowitz [1965] shows that the standard deviation in the time of arrival of the signal is given by

$$\sigma(t) = \frac{1}{\beta(2E\,|N_o)^{1/2}}\,. \tag{6}$$

Notice that this expression depends upon the total signal energy E, the noise background density N_o, and also involves an additional factor β, defined by

$$\beta^2 = \frac{1}{E}\int_{-\infty}^{\infty} (2\pi f)^2\,|S(f)|^2\,df;$$

an expression which depends very much upon the signal shape. Whereas recognizing the presence of the signal was independent of signal shape, measuring the arrival time is not. β is called the effective bandwidth and differs from the usual definition of bandwidth.

Equation (6) shows that the time-of-arrival error for a given signal-to-noise ratio may be minimized by maximizing β. The form of β is similar to the moment of inertia of a body about its center of mass: the moment of inertia of a body confined to some radius r will be maximized by putting all of the material in the body at the radius r. So, if external factors limit our bandwidth to B, then the spectrum which produces the largest effective spectrum β is one with all of the energy crowded at the two ends of the band, i.e., two sine waves at frequencies $f_o - B/2$ and $f_o + B/2$, where f_o is the carrier frequency and B is the bandwidth.

The spectrum of such a signal is given by

$$\delta(f) = \delta(f - f_o - B/2) + \delta(f - f_o + B/2)$$

where δ is the delta function. Substituting $\delta(f)$ into (5), we obtain

$$\sigma(t)_c = \frac{1}{2\pi(2f_o^2 + 2f_o B + B^2)^{1/2}(2E/N_o)^{1/2}} .$$

If only the envelope of the signal is considered, the standard deviation in time of arrival becomes

$$\sigma(t)_e = \frac{1}{2\pi B(2E/N_o)^{1/2}} \tag{7}$$

which is the same as (4) in Section 2. At the present time [Morgan, 1959; Watt, et al., 1961; Fey and Looney, 1966], a two-frequency timing system is being tested in the VLF radio spectrum which does make use of both the carrier and the envelope. This system will be discussed in more detail later.

3.3 Measurement of Time Interval

If a relatively low frequency timing signal is transmitted as modulation on a higher frequency carrier, time interval may be measured by demodulating the signal or by direct observation of the carrier itself. The statistical approach mentioned in the preceding section may also be used to estimate the precision of the interval measurement: either carrier, modulation, or both. The result is similar to that obtained for the time of arrival problem [Berkowitz, 1965]; namely,

$$\sigma(f) = \frac{1}{\alpha(2E/N_o)^{1/2}}$$

where

$$\alpha = \left(\frac{\int_{-\infty}^{\infty} (2\pi t)^2 \, {}^2(t) \, dt}{\int_{-\infty}^{\infty} {}^2(t) \, dt} \right)^{1/2} .$$

α is called the effective time duration of the signal. The primary implication of this result is that the longer the signal lasts, the more precisely its frequency and therefore its interval may be measured. For an application of this result, suppose a signal consists of a sine wave one cycle long with frequency f; i.e.,

$$S(t) = A \sin 2\pi ft, \quad 0 \le t \le 1/f$$

$$= 0 \text{ otherwise.}$$

From the two equations above

$$\sigma(f) = \frac{\sqrt{6}\, f \sqrt{N_o}}{\sqrt{(8\pi^2 - 3)}\sqrt{2E}}$$

$$\sim \frac{\sqrt{3N_o\ f}}{2\pi\sqrt{2E}} .$$

As long as $E/N_o >> 1$ [Skolnik, 1962], E, in the expression above, can be considered the total signal energy resulting from several measurements. Thus, if the sine wave power = P, E = Pt where t is the duration of the signal. With this result, we can rewrite the expression for σ (f) to obtain

$$\sigma(f) = \frac{\sqrt{3N_o}\, f}{2\pi\sqrt{Pt}}$$

This result shows the $1/\sqrt{t}$ dependence in the standard deviation of the frequency measurement which was pointed out in the introduction.

3.4 Simultaneous Measurement of Time of Arrival and Frequency

The product of the rms time of arrival and frequency errors is

$$\sigma(f)\,\sigma(t) = \frac{1}{\alpha\beta\ (2E/N_o)}\ .$$

This shows that the time of arrival and frequency (or interval) may be measured simultaneously with as small an error as desired by increasing the S/N or the $\alpha\beta$ product, or both. As a matter of interest, it may be shown that the product $\alpha\beta$ always satisfies the relation

$$\alpha\beta \geq \pi$$

[Skolnik, 1962]. This result is referred to as the radar uncertainty relation. Thus,

$$\sigma(f)\,\sigma(t) \leq \frac{1}{\pi\,(2E/N_o)}\ . \qquad (8)$$

For a Gaussian-shaped pulse

$$\beta\alpha = \pi$$

so that from (8), this pulse is the poorest possible waveform for the simultaneous measurement of time of arrival and frequency where time of arrival refers to the pulse itself and where f means the carrier frequency for the pulse.

4. The Multiple CW System

It was shown in the previous sections that if signal is restricted to some bandwidth B, a sine wave is optimum for time of arrival measurements. But this system will be ambiguous to all who do not know the time unambiguously to at least the period of the sine wave. The system could serve a larger class of users by transmitting a code, as discussed previously, or by increasing the period of the sine wave with the cost, however, of making some users who could use the higher frequency wait longer for a timing mark. To satisfy more users, we can transmit two sine waves. As an illustration, suppose we have some users who know the time to the nearest second and another group of users who know the time to the nearest 10 seconds. We could satisfy both of these groups by transmitting two sine waves simultaneously with periods of 1 and 10 seconds. As more requirements were added, more CW components could be transmitted at both higher and lower frequencies. In the limit, the spectrum of the composite signal would approach that of a pulse, which, of course, is unambiguous but occupies more bandwidth B.

Consider this multiple CW system in somewhat more detail. Suppose two sine waves are transmitted one after the other with the same power P at frequencies f_1 and f_2 where $f_2 > f_1$. Assume that f_1 is the lowest frequency which is not ambiguous for users. In addition, assume that the user requires a measurement precision of time of arrival to σ seconds. To use the system requires that the position of the zero-crossing of the lowest frequency f_1 be known well enough to coarsely identify the correct cycle of the higher frequency f_2. That is, from (7), the phase jitter on the signal at frequency f_1,

$$\left(\frac{N_o}{8\pi^2 f_1^2 E}\right)^{1/2} \leq \frac{1}{f_2}$$

to correctly identify a cycle on the signal with frequency f_2. Since we are transmitting a sine wave of constant amplitude, we may replace E by PT where P is the power in the signal and T is the observing time. Solving this equation for T, the minimum time required to identify a cycle at f_1 is

$$T_1 = \frac{N_o}{8\pi^2 P} \left[\frac{f_2}{f_1}\right]^2 .$$

Similarly, the time required to obtain the desired precision σ by observing the signal at f_2 is

$$T_2 = \frac{N_o}{8\pi^2 P} \left[\frac{1}{f_2 \sigma}\right]^2 .$$

The total time required is

$$T = T_1 + T_2 = \frac{N_o}{8\pi^2 P}\left[\left(\frac{f_2}{f_1}\right)^2 + \left(\frac{1}{f_2 \sigma}\right)^2\right]$$

In a similar fashion, if we have n frequencies, $f_1 \ldots f_n$, the time required to obtain a precision σ starting at the lowest frequency f_1 is

$$T = \frac{N_o}{8\pi^2 P} \sum_{i=1}^{n} \left(\frac{f_{i+1}}{f_n}\right)^2 + \left(\frac{1}{\sigma f_{n+1}}\right)^2 . \tag{9}$$

To determine the optimum spacing of the frequencies, differentiate this expression with respect to frequencies f_2 through f_{n+1}, and obtain a set of simultaneous equations. It may be shown that T is a minimum if

$$f_i = \left[\left(\frac{1}{\sigma}\right)^{i-1} (f_1)^{n-i+2}\right] \; ; \; i = 2, 3, \ldots n+1. \tag{10}$$

Substituting (10) into (9), we obtain

$$T_{min} = \frac{n+1}{(\sigma f_1)\, 2/(n+1)} \; \frac{N_o}{8\pi^2 P} .$$

Now, to optimize the number of frequencies used, differentiate this expression with respect to n. We find that if

$$n = -(1 + 2 \ln \sigma f_1) = -(1 + 4.65 \log_{10} \sigma f_1)$$
$$= n_{op} ,$$

then T is a minimum with respect to both n and the f_i, and is given by

$$T_{min} = (-4.65 \log_{10} \sigma f_1) \frac{N_o}{8\pi^2 P} = (n_{op} + 1) \frac{N_o}{8\pi^2 P} .$$

Figure 2 shows how T varies with n for the case $f_1 =$ 100 Hz,

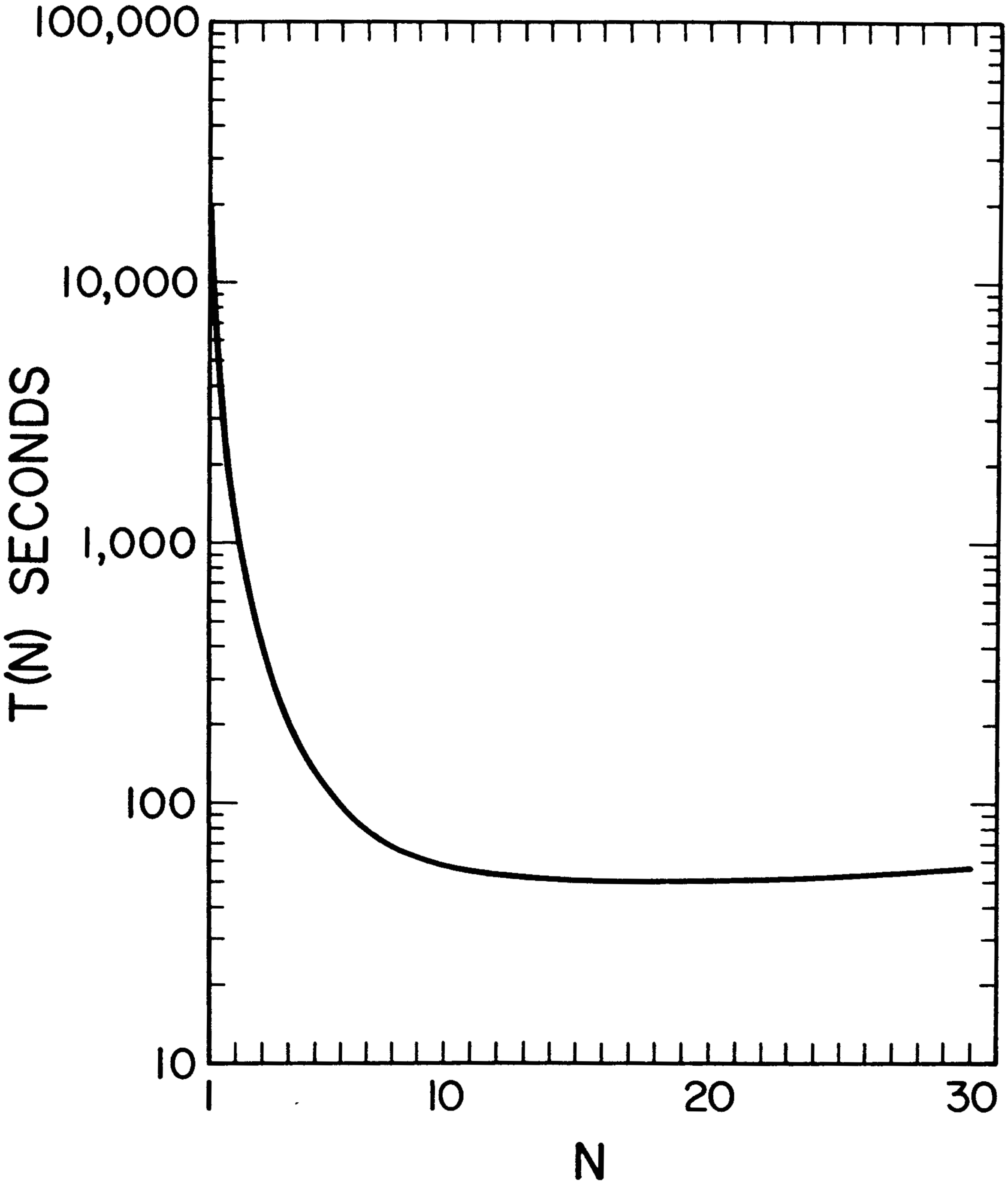

Figure 2. Variation in the required transmission time, T, in seconds, as a function of the number of signals transmitted, N.

$8\pi^2 P/N_o = 1000$, $\sigma = 1\mu s$, and where the f_i have the optimum spacings given by (10). The figure illustrates that a significant gain in precision may be made by using more than one frequency, but beyond $n = 2$ or 3, the gain diminishes rather rapidly as the value of n, which minimizes T, is approached. In some situations the value of n, which minimizes T, may not be realistic in the following sense. As n approaches n_{op}, the transmission time per frequency becomes small. If this transmission time becomes short compared to the period (i.e., 1/f) of the transmitted signal, then the foregoing discussion becomes invalid since $\sigma(t)$ is no longer accurately given by (6). Although we have transmitted the signals sequentially in this discussion, the same result is obtained, if all signals are transmitted simultaneously, since in this case each signal-to-noise ratio must be reduced by a factor 1/n to maintain constant energy.

5. The VLF Two-Frequency Timing System

In the VLF two-frequency timing system presently being used, the low frequency, Δf, (called f_1 in the previous section), is obtained by observing the beat note between two CW VLF signals at frequencies $f_1 = 20$ kHz and $f_2 = 19.9$ kHz. The next higher frequency can be either f_1, f_2, or the average carrier frequency

$$f_c = \frac{f_1 + f_2}{2}.$$

As mentioned, the principle of the system is to observe the low frequency long enough to resolve the ambiguity at the higher frequency and then observe this unambiguous higher frequency until the desired timing precision is obtained.

In this system, used over an ionospheric path, the magnitude of the phase fluctuations on the two carriers are not necessarily equal and may have some correlation ρ. For these conditions, it may be shown from

(9) and (16) (see Appendix) that the optimum value of carrier frequency is given by

$$f_c = \sqrt{\frac{\Delta f}{\sigma(1 - 2\gamma\rho + \gamma^2)^{1/2}}} \;. \qquad (11)$$

Here γ is the ratio of the rms phase fluctuations on the signals at frequencies f_2 and f_1, respectively.

As an illustration, assume $\rho = 0$, $\gamma = 1$, $\Delta f = 100$ Hz, and the user would like a measurement precision of $\sigma = 10\ \mu s$. From (11) the optimum carrier frequency is

$$f_c \cong 8.4 \text{ kHz}.$$

If $\rho = 0.5$, then

$$f_c \cong 10.0 \text{ kHz}$$

and for $\rho = 0.8$

$$f_c \cong 22.0 \text{ kHz}$$

which is near the presently used carrier frequencies. As ρ approaches 1.0, f_c approaches infinity if γ is held at 1.0. In actual practice γ would probably not approach 1.0 as ρ approached 1.0. The reason for this is that a highly correlated motion of the carriers at frequencies f_1 and f_2 would probably be due to a general movement of the entire reflection region in the propagation medium (path length fluctuations). This would result in the delay of the two carriers changing the same amount in time rather than in phase angle. Since γ refers to standard deviation of the phase, an equal movement of the two carriers in time requires that $\gamma = f_2/f_1$. For this particular case, the standard deviation of the motion of the envelope is just the standard deviation of the motion of either one of the carriers. Equation (11) reduces to

$$f_c = \left(\sqrt{\frac{f_1}{\sigma}}\right)_{\rho = 1,\ \gamma = f_2/f_1}$$

in this case.

It is not always possible, because of other considerations, to use the optimum f_c given in (11). Figure 3 shows the variation in observation time T as a function f_c/f_{co} (where f_{co} is the optimum carrier frequency) for several values of ρ where $\gamma = 1.0$, $\frac{8\pi^2 P}{N_o} = 1000.0$, and $\Delta f = 100$ Hz.

6. Characterization of the Noise

Noise affecting a timing system may be placed into two categories. The first is usually called additive noise, and refers to noise energy added to the spectral region occupied by the signal. Examples are: ignition noise, noise generated by the receiver, atmospheric noise, and cosmic noise. The effects of additive noise can be made as small as desired b increasing the strength of the signal. There is also multiplicative noise, which cannot be diminished by increasing the signal strength. This noise is produced, for example, when the propagation medium distorts or modulates the signal. This distortion is always present in the signal no matter what its strength. Diversity techniques have been developed to offset multiplicative noise for communication channels, but appear to have had little or no application with respect to time dissemination. The concept of diversity is to obtain a greater number of independent measurements of the signal for a given amount of signal energy. Thus, the variance of several independent measurements

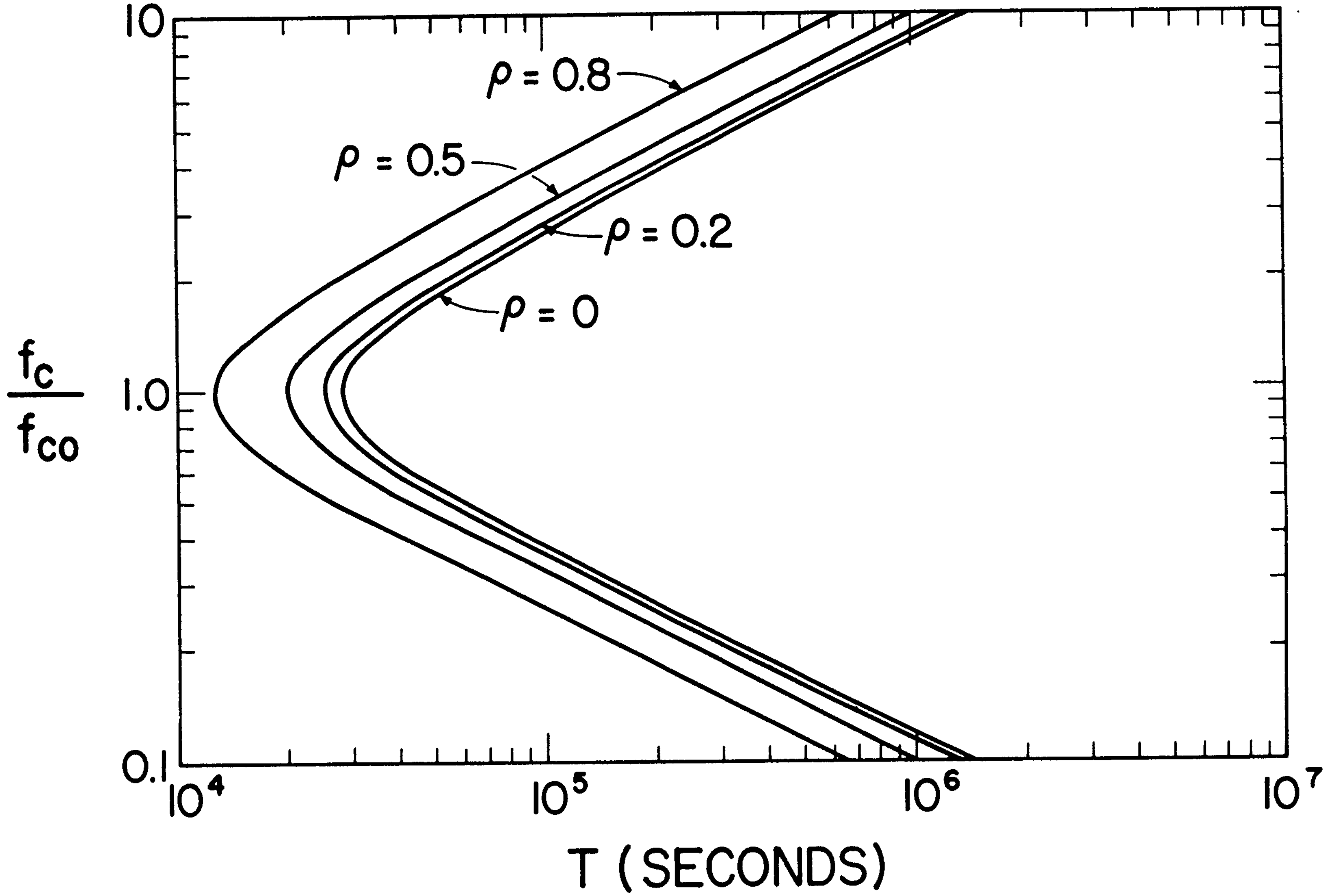

Figure 3. Variation in the required transmission time, T, as a function of the ratio of the frequency actually broadcast, f_c, to the optimum frequency, f_{co}.

is given by

$$\sigma^2 = \frac{\sigma_m^2}{n}$$

where σ_m^2 is the error for each independent measurement, and n is the number of independent samples: independence meaning the noise is uncorrelated from one sample to the next. As an example of multiplicative noise, consider a signal from an artificial earth satellite which passes through a set of ionospheric irregularities whose average size is L (see fig. 4). If L is much larger than the signal radio frequency wavelength λ, and if the satellite is far from the irregularities, then each irregularity will scatter the signal over an angle $\theta \cong \lambda/L$. To an observer who is at a distance D from the irregularities, it will appear to him as though the signal is coming from an area subtended by the angle θ at the distance D. Suppose now the satellite transmits a pulse. Part of the pulse energy will travel directly to the observer along the path d_1. Other portions of the pulse will travel over more distant paths out to the greatest path d_2 given by

$$d_2 = \sqrt{D^2 + D^2\,\theta^2/4}$$

$$= D\sqrt{1 + \lambda^2/2L^2} \quad .$$

Thus, a pulse observed at the ground will be spread over a time interval

$$t = \frac{d_2 - d_1}{V} = \frac{D(1 - \sqrt{1 + \lambda^2/2L^2})}{V}$$

where V = group velocity of the pulse.

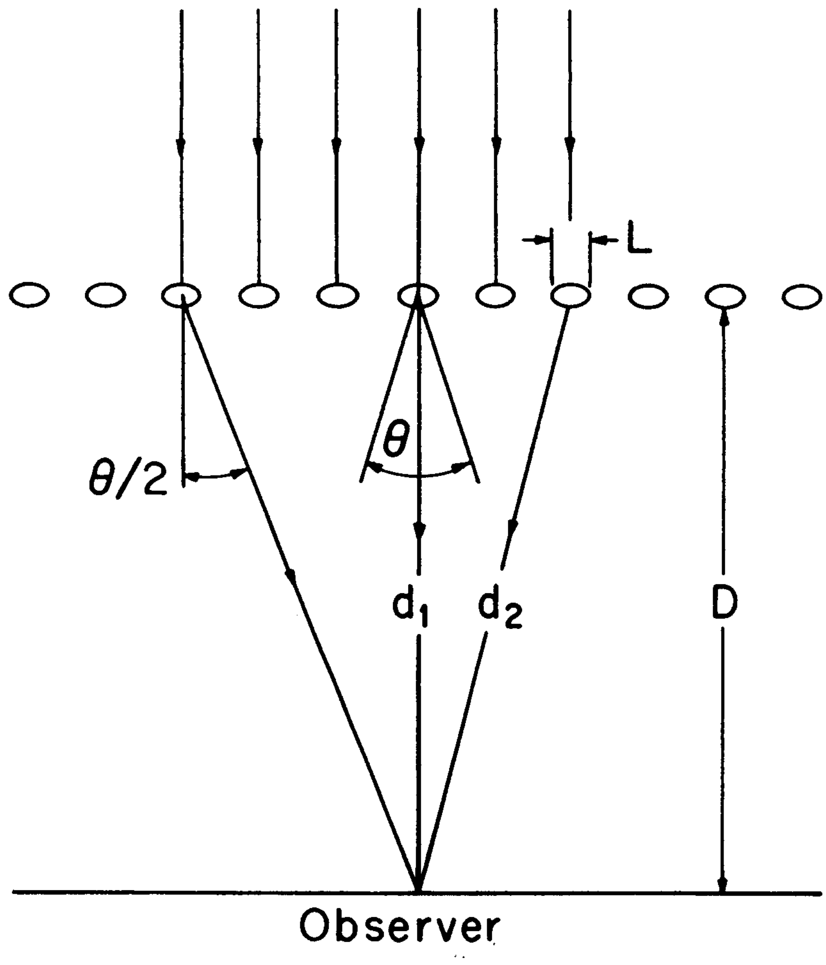

Figure 4. Schematic representation of scattering of satellite signal.

In addition, the irregularities may have motion. This would produce a random phase modulation of the received signal so that its spectrum would appear to be spread. If the satellite is transmitting a CW signal of frequency f_o, the received signal will appear to have spectrum spread about f_o. This is true only if there is no average motion of the irregularities either away from or toward the observer. Otherwise, the spectrum will also have its average value Doppler shifted by some amount f_d. If the spectrum is spread about the average frequency over a range BHz, the received signal will fade with a period of approximately 1/B seconds.

We thus arrive at a fairly simple picture of the fading produced by this channel. First, the signal will appear to fade at a rate B c/s; and second, portions of the signal spectrum within 1/t c/s of each other will appear to fade coherently. The fading rate is governed by the motion of the irregularities, and the correlation is governed by the spread in the multi-path.

If we have a signal which is spread over T seconds in the time domain and over F c/s in the frequency domain, then we could think of the signal as containing N independent diversity elements where

$$N = \frac{TF}{(1/B)(1/t)} = (Bt)(TF).$$

The precision σ^2_{TF} obtained over a time T with a signal spread over a spectrum F is

$$\sigma^2_{TF} = \frac{\sigma^2_m}{N} = \frac{\sigma^2_m}{(Bt)(TF)} .$$

In principle at least, if the noise in our measurement always adds incoherently and the signals add coherently, the lack of measurement precision due to the multiplicative propagation noise can be made arbitrarily small if the available energy is divided into enough diversity elements. Eventually the additive noise will dominate over the multiplicative noise and will represent the fundamental limitation.

In most practical cases there is another type of limitation which is related to the particular detector being used. It is a characteristic of most detectors that the output signal to noise deteriorates rapidly when the input signal to noise reaches some minimum value (threshold effect). If the available energy is divided into too many elements, the detector input signal-to-noise ratio will be too small for the detector to operate efficiently. For example, in a square law detector where S is the input signal and N is the noise, the output of the detector is

$$(S + N)^2 = S^2 + 2NS + N^2 .$$

The output signal-to-noise ratio is

$$r_o = \frac{S^2}{2NS + N^2} \cong \frac{1}{2}\frac{S}{N} \text{ if } \frac{S}{N} >> 1.$$

If $S/N << 1$

$$r_o \cong \left(\frac{S}{N}\right)^2$$

which deteriorates rapidly for small S/N.

With some kinds of diversity it is not necessary to divide the available energy into packets thus reducing the signal energy per packet. An example is space diversity. Here the signal is observed simultaneously at points on the ground sufficiently far apart for the noise fading to be uncorrelated. Other examples are polarization and angle of arrival diversity.

The primary purpose of the preceding discussion was to point out the possible utility of using diversity techniques in time dissemination and not to discuss any particular application in detail. However, there is evidence in the literature that multiplicative noise rather than additive noise is a limiting factor in some existing time and frequency systems. In a study made by Blair and Morgan [1965] of the short term frequency stability of WWVL as recorded at Greenbelt, Maryland, they found the standard deviation of the frequency fluctuations decreased by a factor of two during the daytime when the transmitted power was increased by a factor of about 50. If the frequency fluctuations had been produced entirely by additive noise, the expected decrease would have been by a factor of about seven (see (6)). In this particular instance the multiplicative noise was probably due both to propagation effects and to frequency deviations in the signal as transmitted. During the night there was no appreciable change in the standard deviation of the phase fluctuations from low to high transmitted power. If we assume that the part of the multiplicative noise due to fluctuations in the transmitted signal was about the same both day and night, then it appears that for the night time observations the propagation effects dominate over both additive noise and transmitted signal fluctuations. In any event, these observations demonstrate that one must be aware of the relative importance of multiplicative and additive noises as they affect a specific system, before any meaningful steps may be taken to improve the situation.

7. Appendix

7.1 Effects of Noise on Two-Frequency Timing Systems

If the phase fluctuations on the two carriers are not perfectly correlated, the position of the envelope will move considerably in excess of the noise on either one of the carriers. In the event that the noise on the two carriers becomes completely uncorrelated, the fluctuation in the position of the beat note becomes inversely proportional to the frequency separation [Watt, et al., 1961]. Consider the general case where the correlation between the two carriers phase fluctuations is arbitrary and where the standard deviation of the fluctuations on each of the two carriers are not necessarily equal.

The two carriers may be written in the form

$$x(t) = A \sin 2\pi f_1 (t - \tau_x)$$

and

$$y(t) = A \sin 2\pi f_2 (t - \alpha k \tau_x + \alpha \tau_x^1)$$

where τ_x is the total displacement of the zero crossing of x (t) about its mean position, $k\tau_x$ is the portion $(0 \le k \le 1)$ of the displacement of the zero crossing of y (t) which is correlated with x (t), τ_x^1 is the uncorrelated part, and α is some constant to take care of the possibility that the variance of the zero crossing times of x (t) is not the same as that for y (t). Then*

$$x(t)\,y(t) = \tfrac{1}{2}\cos\,[2\pi f_1(t - \tau_x) - 2\pi f_2(t - \alpha k\tau_x + \alpha\tau_x^1)]$$
$$- \tfrac{1}{2}\cos\,[2\pi f_1(t - \tau_x) + 2\pi f_2(t - \alpha k\tau_x + \alpha\tau_x^1)].$$

The first term represents the envelope component and the second is the carrier frequency.

We want to evaluate the jitter of the envelope, i.e., we are interested in the times t_o when the envelope is zero. Thus,

$$\text{Arg}\,(E) = 2\pi f_1 (t_o - \tau_x) - 2\pi f_2 (t_o - \alpha k\tau_x + \alpha\tau_x^1) = \pm\, n\pi,\ n = 0, 1, 2$$

*The beat note may also be produced by adding the signals, although it does not change the result obtained here.

or rewriting

$$\mathrm{Arg}(E) = 2\pi t_o (f_1 - f_2) - 2\pi\tau_x (f_1 - \alpha k f_2) - 2\pi\alpha\tau_x^1 f_2 = \pm n\pi .$$

Solving for t_o,

$$t_o = \frac{\pm n\pi + 2\pi\tau_x (f_1 - \alpha k f_2) + 2\pi\alpha\tau_x^1\, f_2}{2\pi (f_1 - f_2)} .$$

Then the variance of t_o is

$$\mathrm{Var}\ t_o = \sigma^2_{\tau_x} \left[\frac{f_1 - \alpha k f_2}{f_1 - f_2}\right]^2 + \alpha^2 \sigma^2_{\tau_x 1} \left[\frac{f_2}{f_1 - f_2}\right]^2 \qquad (12)$$

where $\sigma^2_{\tau_x}$ is the variance of τ_x and $\sigma^2_{\tau_x 1}$ is the variance of τ_x^1.

Now

$$\tau_x = k\tau_x + \tau_x^1$$

$$\therefore \quad \sigma^2_{\tau_x} = k^2 \sigma^2_{\tau} + \sigma^2 \tau_x^1 \qquad \text{or}$$

$$\sigma^2_{\tau_x} = \sigma^2_{\tau_x} (1 - k^2) . \qquad (13)$$

Substituting (12) → (13)

$$\text{Var } t_o = \sigma^2_{\tau_x}\left(\frac{f_1 - \alpha k f_2}{f_1 - f_2}\right)^2 + \alpha^2 \sigma^2_{\tau_x}(1 - k^2)\left(\frac{f_2}{f_1 - f_2}\right)^2$$

$$= \frac{\sigma^2_{\tau_x}}{(f_1 - f_2)^2}\left[(f_1 - \alpha k f_2)^2 + \alpha^2(1 - k^2) f_2^2\right]$$

$$= \frac{\sigma^2_{\tau_x}}{(f_1 - f_2)^2}\left[f_1^2 - 2\alpha k f_1 f_2 + \alpha^2 f_2^2\right] . \tag{14}$$

The total displacement of the phase of y(t) is $\alpha\tau_x$ or

$$\text{Var } (\alpha\tau_x) = \sigma^2_{\tau_y} = \alpha^2 \sigma^2_{\tau_x} .$$

Therefore, α^2 represents the ratio of the variance of time variation for y(t) and x(t). Also the phase correlation, ρ, between x(t) and y(t) is

$$\rho = \frac{\langle \tau_x \tau_y \rangle}{\sigma_{\tau_x} \sigma_{\tau_y}}$$

$$= \frac{\langle (\tau_x)\alpha(k\tau_x + \tau_x^1)\rangle}{\sigma_{\tau_x} \sigma_{\tau_y}}$$

$$= \frac{\langle \alpha k \tau_x^2 \rangle}{\sigma_{\tau_x} \sigma_{\tau_y}} + \frac{\langle \alpha \tau_x \tau_x^1 \rangle}{\sigma_{\tau_x} \sigma_{\tau_y}} .$$

But the last term equals zero since τ_x and τ_x^1 are not correlated.

$$\therefore \quad \rho = \frac{\alpha k \sigma^2 \tau_x}{\alpha \sigma^2 \tau_x} = k. \tag{15}$$

Substituting (15) → (14), we obtain

$$\mathrm{Var}\ t_o = \frac{\sigma^2 \tau_x}{(f_1 - f_2)^2}\left[f_1^2 - 2\alpha\rho f_1 f_2 + \alpha^2 f_2^2\right]$$

$$= \frac{\varphi_x^2}{4\pi^2 (f_1 - f_2)^2}\left[1 - 2\gamma\rho + \gamma^2\right] \tag{16}$$

where

$$\gamma^2 = \frac{\varphi_y^2}{\varphi_x^2}$$

and φ_y^2 and φ_x^2 are the variances in phase rather than in time on the two carriers. Some special cases:

A. $\rho = 1, \ \alpha = 1$

$$\mathrm{Var}\ t_o = \sigma^2 \tau_x$$

That is, the jitter on the beat note is just the jitter on either one of the carriers.

B. $\rho = 0, \ \alpha^2 = 1$

$$\mathrm{Var}\ t_o = \sigma^2 \tau_x \left(\frac{f_1^2 + f_2^2}{f_1^2 - f_2^2}\right)$$

C. $\rho = 0, \alpha^2 = \frac{(2\pi f_1)^2}{(2\pi f_2)^2}$, i.e., $\langle\varphi^2\rangle x = \langle\varphi^2\rangle y$

which is the more usual case since $\langle\varphi^2\rangle \alpha \frac{1}{2\,S/N}$. Then

$$\text{Var } t_o = \frac{\sigma^2_{\tau_x}}{(f_1 - f_2)^2}\left[f_1^2 + \frac{\omega_1^2}{\omega_2^2} f_2^2\right]$$

$$= \frac{\sigma^2_{\tau_x}}{(f_1 - f_2)^2}\left[f_1^2 + \frac{f_1^2 f_2^2}{f_2^2}\right]$$

$$= \frac{\sigma^2_{\tau_x}}{(f_1 - f_2)^2}\left[2f_1^2\right].$$

Now

$$\sigma^2_{\tau_x} = \frac{\langle\varphi_x^2\rangle}{4\pi^2 f_1^2}$$

$$\therefore \text{Var } t_o = \frac{2\langle\varphi_x^2\rangle}{4\pi^2 (f_1 - f_2)^2}$$

or

$$\sigma_{t_o} = \frac{\sqrt{2}\,\varphi_x}{\omega_1 - \omega_2}$$

which is the result given by Watt, et al. [1961].

Although the results obtained here apply generally, the system is presently being investigated at VLF. However, very little information is available about the correlation bandwidth at VLF. Presumably, it is considerably greater than 100 c/s at VHF and up; that is, two signals separated by less than 100 c/s would have their fading highly correlated while signals separated by frequencies appreciably greater than 100 c/s would undergo uncorrelated phase fading. For purposes of discussion, assume that the correlation has a gaussian dependence. Therefore,

$$\mathrm{Var}\ t_o = \frac{\varphi^2}{f_s^2}\left(2 - \exp - \frac{f_s^2}{2 f_o^2}\right)$$

where $f_s = f_1 - f_2$. Figure 5 shows Var t_o assuming that the correlation bandwidth is 100 c/s and with the frequency separation varying from 0 to 1000 c/s. It is obvious from this figure that the amount of correlation between the phase jitter of the two carriers is not nearly as important a factor as is the frequency separation. This may be easily seen by looking at (16), which shows that for a given frequency separation, f_s, that Var t_o will only be a factor of two larger for the case of $\rho = 0$ as compared to the $\rho = 1$ case. On the other hand, for a given value of ρ, the variance may be reduced by a factor of two by increasing the frequency separation by a factor of $\sqrt{2}$ and by a greater amount for greater frequency separations. However, this cannot go on indefinitely because the problem of ambiguity will crop up.

8. Acknowledgments

I am indebted to E. L. Crow for suggestions and comments on the material contained in the Appendix. In addition, I have had many stimulating discussions with L. E. Gatterer, G. Kamas, and A. H. Morgan

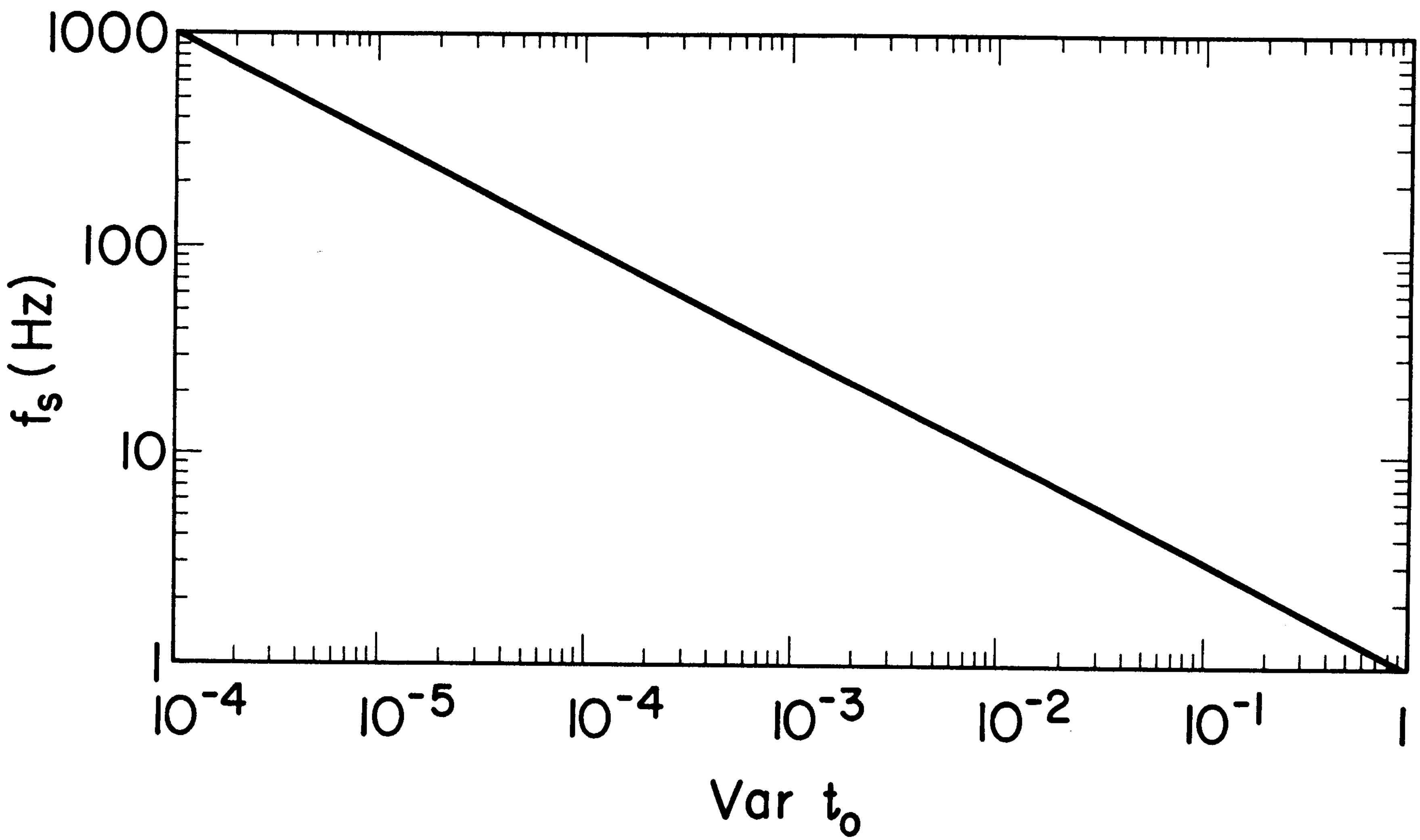

Figure 5. Variance of the signal arrival time, Var t_o, as a function of frequency separation, f_s.

on most of the material covered in this paper. I also thank R. Ruppe for checking the mathematical results and for developing computer programs to calculate the material presented in figs. 2, 3, and 5.