independent, the second-stage information rate for an exponential distribution of pulse intervals becomes

\[
\frac{H' - N'}{T_{\text{min}} + T_c} = \log \left( \frac{2T_c}{\delta} \right) + \log \frac{K(\delta/8w)^{1/2}}{T + 0.37T_c}.
\]

Since this rate was obtained under the peak-limited (and hence band-limited) conditions stated above, it should be less than

\[
\frac{\log (2T_c/\delta)}{T}.
\]

On the Problem of Time Jitter in Sampling*

A. V. RAIKRISHNAN†, SENIOR MEMBER, IRE

Summary—There are many communication as well as control systems—in fact, an increasing number of them—in which at some stage a continuous data source is sampled “periodically,” at the nominal rate of 2W samples a second, W being the highest frequency component of the data. There are, generally speaking, two kinds of errors introduced by the sampling mechanism: errors in amplitude and errors in timing, or “time jitter.” This paper is concerned with the latter.

We assume a random model for the jitter. We begin with a study of the properties of the jittered samples for both deterministic and stochastic signals. Depending on the stochastic properties of the jitter, the presence of a discrete component in the signal may give rise to new discrete components as a result of jitter. Generally speaking, however, the effect of jitter is to produce a (frequency) selective attenuation as well as a uniform spectral density component. The more correlation in the jitter, the less the spectral distribution is affected.

Various measures of the “error” due to jitter are estimated. Thus the error may be the mean square in the fitted samples or some linear or nonlinear operation thereof. Also, weighted mean-square errors are considered, and general methods of estimating these errors are developed.

The problem of optimal use of the jittered samples is next considered. Interpreting the optimality to be in the mean-square sense, an explicit solution for optimal linear operation is obtained. Also for a wide class of signals it is shown that jitter does not affect the nature of the optimal operations; linear operations, for instance, remain linear, although with different weights.

To illustrate the methods an example drawn from telemetry is given, where the timing is derived from the zero crossings of a sine wave and the time jitter is taken as due to noise. The jitter is highly correlated and the results involve some lengthy calculations.


Various measures of the "error" due to jitter are estimated in Section IV. Thus the error may be the mean square in the fitted samples or some linear or nonlinear operation thereof. Also, weighted mean-square errors may be considered. General methods of evaluating these errors are developed.

In Section V we study the problem of optimal use of the jittered samples. Interpreting the optimality to be in the mean-square sense, it is shown how an explicit solution for optimal linear operation may be obtained. Also for a wide class of signals it is shown that jitter does not affect the nature of the optimal operations; linear operations, for instance, remain linear, although with different weights.

Finally, in Section VI an example drawn from telemetry is given. Here the timing is derived from the zero crossings of a sine wave and the time jitter is taken as due to noise. The jitter is highly correlated and the results involve some lengthy calculations.

II. LET x(t) BE CONTINUOUS PARAMETER "SIGNAL" WHICH IS BEING SAMPLED AT DISCRETE MULTIPLES OF THE PERIOD T.

The samples
\[ x_n = x(nT) \] (1)
provide a discrete-parameter process. There are two kinds of errors introduced by the sampling mechanisms: errors in amplitude and errors in timing. In the latter, the sampling epochs are subject to "jitter" so that samples are now taken not at nT but at \((nT + \xi)\) where \(\xi\) represents the error of jitter, and the corresponding samples now become
\[ y_n = x(nT + \xi). \] (2)

Any quantitative evaluation, has, however, to depend on the kind of assumption made regarding the jitter as well as the class of signals considered. Thus we can consider each as either random or deterministic. However there would appear to be little ground for assuming deterministic jitter as far as sampling errors are concerned and in any case, given the particular deterministic jitter the pertinent theory here must be considered well known, going back to the basic work of Wiener and Levinson.\(^3\)

We shall here consider only random model for the jitter. Specifically, we shall assume that \(\Xi\) is a strictly stationary discrete-parameter process, with finite variance. We shall show later by an example that this is a natural assumption. As far as the signal is concerned, we shall make no assumptions and consider both the deterministic and stochastic case.


III. A STUDY OF THE SAMPLES WITH TIME JITTER

**Stochastic Signals**

Let us first consider the signal \(x(t)\) to be stationary, random with zero mean and with correlation \(R(t)\) so that
\[ R(t) = \int_{-\infty}^{\infty} e^{i\omega t} dP(\omega) \] (3)
where \(dP(\omega)\) is the differential spectrum. Then it is readily verified that
\[ E[y_n] = 0 \]
\[ E[y_n y_m] = \int_{-\infty}^{\infty} e^{2\pi i(f-n-m)T} C_{x,n}(f) dP(f). \]

Where we have used the independence of the \(x(t)\) and \(\xi\), and \(C_{x,n}(f - f)\) is the characteristic function
\[ C_{x,n}(f, -f) = E[exp 2\pi i(\xi_n - \xi_m)] \]
then, of course,
\[ C_{x,n}(f, -f) = 1. \]

Since the \(\xi\) process is assumed to be strictly stationary, we have that \(C_{x,n}(f - f)\) depends only on \((n - m)\), so that we obtain that the \(y_n\) process has a stationary covariance. This covariance, \(R_x(n)\), is given by
\[ R_x(n) = \int_{-\infty}^{\infty} e^{2\pi i f n} C_x(f) dP(f) \] (4)
where
\[ C_x(f) = C_{x,2x}(f, f). \]

Note that, if there were no jitter,
\[ R_x(n) = R_x(n) = \int_{-\infty}^{\infty} e^{2\pi i f n} dP(f) \]
where \(R_x(n)\) is the covariance of \(\{x(nT)\}\). Also, \(C_x(f)\) being a characteristic function, we have that
\[ |C_x(f)| \leq 1, \quad C_x(f) = 1 \]
\[ R_x(0) = R_x(0). \]

In particular then, time jitter does not affect the average signal power. Proceeding further, let us examine the spectrum of the \(y_n\) process. Now, \(R_x(n)\) has the representation which can be obtained from (3):
\[ R_x(n) = \int_{-1/2}^{1/2} e^{2\pi i n \lambda} d\Phi(\lambda) \] (5)
where, taking \(T\) as \(1/2W\),
\[ d\Phi(\lambda) = \sum_{k=-n}^{n} dP(2W\lambda + 2kW). \]

By a similar change of variable in (4), we have
\[ R_x(n) = \int_{-1/2}^{1/2} e^{2\pi i n \lambda} d\Phi(\lambda) \] (6)
where

\[ d\psi_\lambda = \sum_{k=-\infty}^{\infty} C_\lambda(2W\lambda + 2kW) \, dP(2W\lambda + 2kW). \]

The spectrum of the \( y_n \) process being capable of more direct interpretation, the next step is to evaluate it from (6). Now, the spectrum is not given by the factor of \( \exp 2\pi\in\lambda \) in (6). However, some simplification arises if the jitter \( \xi \) is such that \( \xi \) is independent of \( \xi_n \). We shall study this case first.

"Pure White" Jitter

If the sequence is "pure white" so that \( \xi \) is independent of \( \xi_n \), we have, for \( n \neq 0 \),

\[ C_{\xi}(f) = E[e^{2\pi i \xi_n}] = [C(f)]^2 \]

where \( C(f) \) is the characteristic function of the first order distribution of \( \xi \) and is independent of \( n \). Hence,

\[ R_\xi(n) = \int_{-1/2}^{1/2} e^{2\pi i \lambda n} \left| C(2W\lambda + 2Wk) \right|^2 dP(2W\lambda + 2Wk), \quad n \neq 0. \]

This still does not mean (!) that the spectrum of \( y_n \) process is the factor of \( \exp 2\pi\in \) in (7) since

\[ R_\xi(0) = \int_{-1/2}^{1/2} d\psi(\lambda). \]

To obtain the spectrum let

\[ \int_{-1/2}^{1/2} \left| C(2W\lambda + 2Wk) \right|^2 dP(2W\lambda + 2Wk) = \sigma^2. \]

Then, writing

\[ R_\xi(n) = R_\xi(0) = (R_\xi(0) - \sigma^2) + \int_{-1/2}^{1/2} \sum_{k=-\infty}^{\infty} \left| C(2W\lambda + 2Wk) \right|^2 dP(2W\lambda + 2Wk), \]

we obtain for the differential spectrum of the \( y_n \) process

\[ dP_\xi(\lambda) = \sum_{k=-\infty}^{\infty} \left| C(2W\lambda + 2Wk) \right|^2 dP(2W\lambda + 2Wk) + (R_\xi(0) - \sigma^2) \, d\lambda. \]

Let us assume next that the signal is band-limited to \((-W, W)\), so as to isolate the effect due to jitter. Then, of course,

\[ dP_\xi(\lambda) = \left| C(2W\lambda) \right|^2 dP(2W\lambda) + (R_\xi(0) - \sigma^2) \, d\lambda \]

where the second term is also positive, since

\[ \left| C(2W\lambda) \right|^2 \leq 1. \]

Now (8) yields a simple interpretation. It is that the jitter selectively attenuates all frequencies and adds a uniform spectral density given by

\[ (R_\xi(0) - \sigma^2). \]

It does not alter the composition of the spectrum. Thus if the spectrum contains discrete components, these are retained (unless \( C(f) \) vanishes at these frequencies) although attenuated and we have in addition a uniform spectral density. If we use the "ideal" method of fit to obtain \( y(t) \), viz: if

\[ y(t) = \sum_{n=-\infty}^{\infty} y_n \sin \frac{\pi(2Wt - n)}{\pi(2Wt - n)} \]

then \( y(t) \) is a stationary process (assuming \( x(t) \) has no discrete components at \( \pm W \)). Its differential spectrum is given by, if \( x(t) \) is band-limited to \((-W, W)\):

\[ \left| C(f) \right|^2 dP(f) + (R_\xi(0) - \sigma^2) \, df. \]

(9)

A measure of the frequency distortion due to jitter is then the attenuation in

\[ 10 \log_a \left| C(f) \right|^2. \]

For Gaussian jitter of variance \( \sigma^2 \) this is

\[ -10 \log_{10} \exp 4\pi^2 \sigma^2 = (40\pi^2 \sigma)^2 (\log_{10} e) \]

and increases monotonically with frequency. If the jitter is taken to be uniformly distributed between \(-\gamma T\) and \(+\gamma T\), this is

\[ -10 \log_{10} \left| \frac{\sin 2\pi\gamma T}{2\pi\gamma T} \right|^2 \]

(10)

and the attenuation again increases with frequency at least for band-limited signals, limited to \([-1/T, 1/T]\), as would be the case normally. In general, for jitter distributions absolutely continuous with respect to Lebesque measure we should expect this to be the case. On the other hand, if we take an extreme case and assume that the jitter consists of just two amplitudes \(-\gamma T\) and \(+\gamma T\) with probabilities \( P \) and \( 1-P \), we have

\[ C(f) = (1 - P)^2 + P^2 + 2P(1 - P) \cos 4\pi\gamma T \]

and we cannot clearly make any statement about increasing attenuation with frequency.

Correlated Jitter

We next consider the general case when the jitter is correlated. Here it is convenient to isolate the effect due to correlation. Thus we begin by rewriting (6) as

\[ R_\xi(n) = R_\xi(0) + \int_{-1/2}^{1/2} \left[ C_\lambda(2W\lambda) - \left| C(2W\lambda) \right|^2 \right] \, d\psi(\lambda). \]

(11)

where \( R_\xi(n) \) is the covariance assuming \( \xi \) to be pure white and \( D(n) \) is the perturbation due to correlation in the jitter. In particular,

\[ D(0) = 0 \]

and the average power is again unchanged.

The spectrum of the $y_n$ process can then similarly be decomposed with a perturbation term due to correlation in the jitter. This perturbation term, for example, $P_b(\lambda)$, has the Fourier Coefficients $D(n)$, and can, of course, be determined from the formula

$$
P_b(\lambda_0) + P_b(\lambda_0) = \frac{1}{2} \lim_{n \to -\infty} \sum_{n} D(n) \sin \frac{2\pi n \lambda_0 - 2\pi n \lambda}{n}
$$

where we have used the fact the $D(n)$ is real.

To begin with, let us consider the jitter to be Gaussian with variance $\sigma^2$ and covariance

$$
E[\zeta_{n+1}] = r(k).
$$

Then in (11),

$$
C_s(f) - |C(f)|^2 = \left(\exp - 4\pi^2 f^2 \sigma^2 \right) \left(\exp 4\pi^2 f r(n) - 1\right)
$$

and thus

$$
D_s = \sum_{k=1}^{\infty} \int_{-1/2}^{1/2} \frac{[4\pi^2 (2Wn)^3 r(n)]^k}{k!} \cdot \sin \frac{2\pi n \lambda_0}{n} \cdot \exp - 4\pi^2 (2Wn)^2 \sigma_n^2 d\lambda(k).
$$

If we assume that the correlation goes to zero with $k$, then

$$
|D_s| \leq (\text{constant}) |r(n)|
$$

and, for example,

$$
\sum_{n=1}^{\infty} |r(n)| < \infty,
$$

then it follows that $P_b(\lambda)$ is absolutely continuous with derivative given by

$$
P_b(\lambda) = 2 \sum_{n=1}^{\infty} D_s \cos 2\pi n \lambda.
$$

Substituting in (13) and changing the order of summation we obtain finally

$$
P_b(\lambda_0) = 2 \sum_{n=1}^{\infty} \int_{-1/2}^{1/2} \frac{[4\pi^2 (2Wn)^3 r(n)]^k}{k!} \cdot \int_{-1/2}^{1/2} \left(e^{r^4 (2Wn)^3} \cdot d\lambda(k)\right)
$$

For example, if a Markov model is used for the jitter, so that

$$
r(n) = r^{(n)} \sigma^2, \quad 0 < r < 1,
$$

the first series can be summed and we have

$$
P_b(\lambda_0) = 2 \sum_{n=1}^{\infty} \frac{[4\pi^2 (2Wn)^3 \sigma_n^2]^4}{k!} \cdot \int_{-1/2}^{1/2} \frac{r^4 \cos 2\pi n (\lambda_0 - \lambda)}{1 - 2r^4 \cos 2\pi n (\lambda_0 - \lambda) + r^{2k}} \cdot \exp - 4\pi^2 (2Wn)^2 \sigma_n^2 d\lambda(k).
$$

It should be noted that $p_b(\lambda)$ is not a positive spectral density. Indeed, we have

$$
\int_{-1/2}^{1/2} p_b(\lambda) d\lambda = 0.
$$

We see then that the effect of correlation is to introduce a further spectral density term which is no longer a constant. If, for instance, we consider the effect on a discrete component at $\lambda = \pm \lambda_0$ in the spectrum of $\{x_n\}$, we have

$$
p_b(\lambda) = 2 \sum_{n=1}^{\infty} \frac{(4\pi^2 (2Wn)^3 \sigma_n^2)^k}{k!} \cdot \frac{r^4 \cos 2\pi n (\lambda_0 - \lambda) - r^2}{1 - 2r^4 \cos 2\pi n (\lambda_0 - \lambda) + r^{2k}}
$$

$$
\cdot \exp - 4\pi^2 (2Wn)^2 \sigma_n^2.
$$

A typical term in this,

$$
\frac{a \cos 2\pi n (\lambda_0 - \lambda) - a^2}{1 - 2a \cos 2\pi n (\lambda_0 - \lambda) + a^2},
$$

is plotted in Fig. 1 as a function of $\lambda_0 - \lambda$ for two values of $\alpha$. We note that there is now a positive peak at $\lambda = \lambda_0$ with increasing steepness as $\alpha$ increases.

More generally, we can now describe a method which yields a representation of the spectrum for a large class of jitter statistics. For this let $p(x, y; n)$ be the joint density of $\xi_n, \xi_{n+k}$ and let us assume that it can be expanded in terms of the first order densities $p(\cdot)$ as

$$
p(x, y; n) = \sum_{m=0}^{\infty} a^m p(x) p(y) P(x) p(y)
$$

where $P(x)$ are orthogonal with respect to $P(x)$ and

$$
a^n = 1.
$$

We have then that

$$
C_s(f) - |C(f)|^2 = \int e^{2\pi i f x} [P(x, y; n) - P(x) P(y)] dx dy
$$

$$
= \sum_{m=0}^{\infty} a^m |g_s(f)|^2
$$

where

$$
g_s(f) = \int e^{2\pi i f x} p(x) P(x) dx
$$

and substituting in (12), gives

$$
D(n) = \sum_{k=1}^{\infty} a^k \int_{-1/2}^{1/2} |g_s(2Wn)|^2 \cdot \exp - 4\pi^2 (2Wn)^2 \sigma_n^2 d\lambda(k).
$$

Assuming that $a^k$ goes to zero sufficiently rapidly with $k$ and $n$, we can conclude from this that the effect of
Fig. 1.

correlation in the jitter is to add another and nonuniform density term \( p_D(X) \) which we can explicitly express as

\[
p_D(X) = \frac{2}{\pi W^2} I_0(2\pi W X) e^{-\pi W^2 X^2}
\]

yielding a general solution to the spectrum problem. In the Gaussian case previously considered, for instance, \( a^2 = \pi^2 \) and (19) specializes to the prior result.

If the correlation \( r(n) \) does not vanish as \( n \) goes to infinity, this general result has to be modified. In particular, new discrete components may occur in the spectrum of \( y_n \) which did not exist in the \( x(t) \) signal process. The simplest example of this type is to take a Gaussian jitter with

\[
r(n) = \sigma^2 \cos 2\pi \lambda n
\]

where \( \lambda \) is the linear frequency of the discrete harmonic component. In this case it is more convenient to go back to \( R_x(n) \) directly:

\[
\begin{align*}
R_x(n) &= 2 \int_{-1/2}^{1/2} e^{2\pi \lambda n} (\exp - 4\pi^2 W^2 \lambda^2 \sigma^2) \\
&\quad \cdot (\exp + 4\pi^2 W^2 \lambda^2 \sigma^2 \cos 2\pi \lambda n) d\lambda \\
&\quad - \int_{-1/2}^{1/2} e^{2\pi \lambda n} (\exp + 4\pi^2 W^2 \lambda^2 \sigma^2) I_0(16\pi^2 W^2 \lambda^2 \sigma^2) e^{-4\pi^2 W^2 \lambda^2 \sigma^2} d\lambda \\
&= \sum_k \int_{-1/2}^{1/2} \left( e^{2\pi \lambda (n-k)} I_0(16\pi^2 W^2 \lambda^2 \sigma^2) e^{-4\pi^2 W^2 \lambda^2 \sigma^2} d\lambda \right)
\end{align*}
\]

where \( I_0(\cdot) \) is the 0th order modified Bessel function. A typical term in this summation is

\[
\int_{-1/2}^{1/2} e^{2\pi \lambda (n-k)} I_0(a^2 \lambda^2) e^{-a^2 \lambda^2} d\lambda
\]

where

\[
a^2 = 16\pi^2 W^2 \sigma^2.
\]

For each \( n \), let \( \lambda_n \) be a stationary Markov process with the transition matrix

\[
\begin{pmatrix}
1 - P & P \\
P & 1 - P
\end{pmatrix}
\]

so that for \( n \neq 0 \)

\[
D_n = \int_{-1/2}^{1/2} \left( 1 - 2P \right)^n (1 - \cos 4\pi (2W) \lambda) e^{-\pi (\lambda - \lambda_n)^2} d\lambda.
\]

If \( 0 < P < 1 \), it is clear that

\[
\sum |D_n| < \infty,
\]
so that the effect of correlation is to add a density term
\[ p_D(\lambda) = 2 \sum_{i} D_n \cos 2\pi n \lambda \]
\[ = 2 \int_{-1/2}^{1/2} \frac{a \cos 2\pi(\lambda - \sigma) - \sigma^2}{1 - 2a \cos 2\pi(\lambda - \sigma) + \sigma^2} \cdot (1 - \cos 4\pi W \sigma b) \, d\sigma \]
where
\[ a = 1 - 2P. \]

It follows from this that no new discrete components are introduced. In the extreme case where \( p = 0 \), we have complete correlation. In this case, in fact, since
\[ C_1(f) = 1 \]
the total differential spectrum of the \( y_n \) process is given by
\[ dP_y(\lambda) = d\psi(\lambda) \]
and the spectrum is unaffected.

The method can be extended to more than two states as well as higher-order Markov processes.

IV. DETERMINISTIC SIGNALS

We shall next consider deterministic signals and random jitter. We shall consider the signals to be represented by a trigonometric sum
\[ x(t) = \sum_{i} a_i e^{i2\pi f_i t} \quad (23) \]
where \( f_i \) are real and \( a_i \) are complex and we may take \( x(t) \) to be real. The spectrum of \( x(t) \) consists of discrete components at \( = tf_i \). If \( |f_1| < W \), then of course \( \{ z_i \} \) will completely describe \( x(t) \). We shall now examine the effect of random jitter on the samples.

Here
\[ y_n = x(nT + J_n) \]
is now a random process, although, of course, no longer stationary. Thus
\[ E[y_n] = \sum_{i} a_i e^{i2\pi f_i n} C(f_i) \quad (24) \]
where
\[ C(f) = E[e^{i2\pi f_n}] \]
and we note the appearance of the attenuation factor \( C(f_i) \) in contrast to (23). The correlation is
\[ E[y_n g_n] = \sum_{i} \sum_{m} a_i a_m e^{i2\pi (f_i - f_j) n} \tau C_{n-m}(f_i, -f_j) \]
where \( C_{n-m}(f_i, -f_j) \) is defined as before in (4). The first order distribution of \( y_n \) is conveniently obtained from the characteristic function
\[ E[e^{iux}] = \sum_{n} \cdots \sum_{n} J_{s_1}(uB_1) \cdots J_{s_v}(uB_v) \cdot C(n_{s_1} + n_{s_2} + \cdots + n_{s_v}) \cdot \exp i \sum_{i} C(2\pi f_i nT + \theta_i) n_i \quad (25) \]
where \( x(t) \) is written in the form
\[ x(t) = \sum_{i} B_i \sin(2\pi f_i t + \theta_i). \]
The appearance of the Bessel functions is characteristic, since the jitter can be thought as a sort of discrete frequency modulation. It is clear that the first-order distribution is not easy of explicit expression. The characteristic function corresponding to the second-order distribution of \( z_n, z_{n+m} \) can be computed similarly. Now if \( z_n, z_{n+m} \) are independent, a simplification arises, for in that case \( y_n \) and \( y_{n+m} \) are independent also. In that case (25) yields all the desired statistics.

Since the process is nonstationary, we shall define an average (intensity) spectrum as done by Lawson and Uhlenbeck.\(^6\) We know that the spectrum will contain a discrete part and/or continuous part. The discrete part is given by
\[ \lim_{M \to \infty} \frac{1}{M^2} \left[ \sum_{i} y_n e^{i2\pi f_i n} \right]^2. \]
We then have, in the present case
\[ E\left[ \frac{1}{M^2} \left| \sum_{i} y_n e^{i2\pi f_i n} \right|^2 \right] = \frac{1}{M^2} \sum_{i} \sum_{m} E[y_n g_m] e^{i2\pi (n-m)\lambda}. \quad (26) \]
Now for a typical term in this
\[ E[y_n g_m] e^{i2\pi (n-m)\lambda} = \sum_{i} \sum_{m} a_i a_m e^{i2\pi (f_i - f_j) n} \cdot C_{n-m}(f_i, -f_j). \]
Hence we can, by a change of order of summation, consider limits of the form
\[ \lim_{M \to \infty} \frac{1}{M^2} \sum_{i} \sum_{m} C_{n-m}(f_i, -f_j) \cdot \exp i 2\pi (mf_i T) T + n - m\lambda). \quad (27) \]
If the jitter is pure white, so that,
\[ C_{n-m}(f_i, -f_j) = C(f_i)C(f_j) \]
the limit in (27) simplifies to yield zero for \( k \neq j \), while for \( k = j \) we obtain for \( \lambda = -f_i T \) and \( 1 - f_i T, \)
\[ \frac{1}{2} \left| a_i \right|^2 \left| C(f_i) \right|^2 \quad (28) \]
and zero for other values of \( \lambda \). We note again the attenua-

tion factor \( |C(f)|^2 \). Now the average intensity in this case
\[
\lim_{M \to \infty} \frac{1}{M} \sum_{k=1}^{M} E(y^2_k) \]
is easily calculated to be
\[
\frac{1}{2} \sum_{k=1}^{M} \left| a_k \right|^2 \tag{29}
\]
and hence the difference
\[
\frac{1}{2} \sum_{k=1}^{M} \left| a_k \right|^2 \left[ 1 - |C(f)|^2 \right]
\]
appears as in the form of a continuous spectrum. When the jitter is correlated, \( \delta \)-functions appear again at
\[
\lambda = -f_sT \quad \text{and} \quad 1 - f_sT
\]
with magnitude
\[
\lim_{\infty} a^2 \sum_{k=1}^{M} \frac{(M - P)}{M^2} C_P(f_s - f_h) \tag{30}
\]
and the difference between this and (29) appears as the continuous part of the spectrum unless the jitter contains discrete harmonic components. As expected, the results are similar to those for random signals.

V. Errors Due to Jitter

Stochastic Signals

Let us now consider the error due to time jitter. At the first instance one may be interested in the error in the samples themselves:
\[
y_n - x_n
\]
or the error in the “fitted” function \( y(t) \) determined from the samples:
\[
y(t) - x(t)
\]
More generally, if the output of the system is some operation, linear or nonlinear, on the samples, we may consider the error:
\[
0[y_n] - 0[x_n]
\]
The point here is that this error can be small, even though the “fit” error above is large. As for the quantitative measure of the error itself, the (normalized) mean-square error appears to be convenient. Since we are talking about the errors in a quantity which can itself be zero, we can also consider a weighted mean-square error such as
\[
E\left[ \frac{(y(t) - x(t))^2}{\Delta^2 + x(t)^2} \right].
\]
We shall now develop some general methods for calculating these errors.

First we shall consider signals band-limiting to \([-W, W]\) so as to concentrate on the error due to jitter alone. Then
\[
E[(y_n - x_n)^2] = R_s(0) + R_s(0) - 2E[y_n x_n]
\]
and
\[
= 2R_s(0) - 2 \int_{-W}^{W} C(f) \, dP(f)
\]
\[
= 2 \int_{-W}^{W} (1 - C(f)) \, dP(f). \tag{31}
\]
If \( C(f) \) has the power series expansion
\[
C(f) = 1 + \sum_{k=1}^{\infty} a_k(2\pi f)^n,
\]
Then (31) can be expressed as
\[
2 \int_{-W}^{W} \sum_{k=1}^{\infty} a_k(2\pi f)^n \, dP(f) = 2 \sum_{k=1}^{\infty} a_k R_s(0) \tag{32}
\]
provided the series in (32) converges uniformly or boundedly in \([-W, W]\), as is true in all the examples studied. A good approximation would be to take (provided the jitter mean is zero) the first term
\[
2a^2 R_s(0) \tag{33}
\]
\( \sigma^2 \) being the jitter variance. For signals with a flat spectrum the normalized error (as a fraction of signal power) is
\[
16\pi^2 \sigma^4 W^2 = 8\pi^2 \sigma^4 W^2 \frac{1}{3} \tag{34}
\]
and is, in particular, proportional to the square of the bandwidth. This is also the mean-square error in the “fitted” \( y(t) \) using
\[
y(t) = \sum_{n=1}^{M} y_n \frac{\sin \pi(2Wt - n)}{\pi(2Wt - n)}
\]
since \( y(t) \) is also stationary, as already noted. The mean-square error (34) is meaningful in itself, since \( x_n \text{ and } y_n \) have the same power so that no scale factor is involved.

Next let us consider a weighted mean-square error
\[
E\left[ \frac{(y(t) - x(t))^2}{\Delta^2 + x(t)^2} \right]. \tag{35}
\]
First of all it is convenient to express this as
\[
E^2(\Delta) = \left[ \frac{(y_n - x_n)^2}{\Delta^2 + x_n^2} \right] \tag{36}
\]
where
\[
\sigma^2 = R_s(0).
\]
Note that the normalized mean-square error \( \varepsilon^2 \) is then given by
\[
\varepsilon^2 = \lim_{\Delta \to 0} \Delta^2 E^2(\Delta).
\]
For a general class of signals for which 2nd order joint densities can be expanded as (17), we can simplify (36) further. Thus let \( p(t; x, y) \) the joint density of \( x(t), \ x(t + \tau) \) be expressible as
\[
p(t; x, y) = \sum_{k=0}^{\infty} a_k(\tau)Q_k(y)Q_k(x)P(x)P(y).
\]
Then, after some simplification, we have
\[ e^2(\Delta) = b_0(\Delta) + (1 - \Delta^2 b_0(\Delta))(1 - 2a_{\Delta}) + a_{\Delta} b_0(\Delta) \]  
(37)

where
\[ b_0(\Delta) = \left( \int y^2 Q_0(y) P(y) \, dy \right) \left( \int \frac{Q_1(x)}{a_0^2 \sigma^2} \, \sigma^2 P(x) \, dx \right) \]
\[ a_{\Delta} = E[a_0(\Delta)]. \]

If the third moment is zero, \( b_1(\Delta) \) vanishes; for Gaussian signals further simplification is possible, and \( b_0(\Delta), b_1(\Delta) \) can be explicitly evaluated. We omit the details.

If the desired output is the result of a linear operation, let us say
\[ Z_n = \sum_{m=-\infty}^{\infty} W_m x_{m-k}, \]
then
\[ Z_n = \sum_{m=-\infty}^{\infty} W_m y_{m-k} \]
and the mean-square error is now
\[ E[(Z_n - Z_n)^2] = \left[ \sum_{m=-\infty}^{\infty} W_m^2 (x_{m-k} - y_{m-k}) \right]^2 \]
and for band-limited signals, in terms of spectra this is
\[ \int_{-1/2}^{1/2} \left| L(\lambda) \right|^2 \left[ d\Psi(\lambda) + dP_\psi(\lambda) - 2C(2W\lambda) \right] d\Psi(\lambda) \]
where \( L(\lambda) \) is the transform of the linear weights
\[ L(\lambda) = \sum_{m=-\infty}^{\infty} W_m e^{2\pi i m \lambda}. \]

In contrast to (31) we have now a frequency weighting, and, of course, dependence on jitter correlation. Next let us consider nonlinear functionals. Suppose the desired is \( f(z_n) \). Then the mean-square error is
\[ E[(f(z_n) - f(y_n))^2]. \]
(39)

To obtain a general result, we may expand \( f(x) \) in orthogonal polynomials. Let us, to be specific, assume \( x(t) \) Gaussian of unit variance. Then let
\[ f(x) = \sum_{n=-\infty}^{\infty} a_n H_n(x) \]
then
\[ E[f(x_n) f(y_n)] = \sum_{n=0}^{\infty} a_n^2 \gamma^n \]
where \( H_n(\cdot) \) are Hermite polynomials
\[ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H_n(x) H_n(x) \exp - x^2 / 2 \, dx = \delta_n^0, \]
\[ \gamma = \int_{-W}^{W} C(f) \, dP(f) \]  
and by assumption \( \int_{-W}^{W} dP(f) = 1 \).

Hence mean-square error is
\[ E[(f(x_n) - f(y_n))^2] = 2 \sum_{n=0}^{\infty} a_n^2 [1 - \gamma^n]. \]
(40)

This result is, of course, independent of whether the jitter is correlated or not. A similar result holds if the desired signal is of the form:
\[ f(\sum_{m=-\infty}^{\infty} W_m x_{m-k}). \]

VI. OPTIMIZATION OF JITTERED SAMPLES

In this section we shall consider methods of optimum processing of jittered data. The first case we shall examine is the one where \( x(t) \) is band-limited and it is desired to obtain a best “fit” for \( x(t) \) from the samples. Here and throughout this section we shall use a method introduced by the author in (4). By best “fit” we shall mean the optimum mean-square estimate of \( x(t) \) from the jittered samples \( y_n \). We use the entire sequence \( y_n \) since we are allowing two-sided interpolation. First we note that, regardless of statistics, whether Gaussian or not, the best mean-square estimate of \( x(t) \) from \( x_n \) is linear, and
\[ x(t) = \sum_{n=-\infty}^{\infty} x_n \sin \pi(2Wt - n) \]
(41)
where the limit is taken in the mean of order two, as proved in (4), the error being zero.

Band-Limited Signals—Pure White Jitter

We shall begin by considering pure white jitter since this should yield some simplification. We shall first consider the optimum linear estimate. Let
\[ \sum_{n=-\infty}^{\infty} a_n(t) y_n = y(t) \]
be the optimal estimate. Then we must have
\[ E(\sum_{n=-\infty}^{\infty} a_n(t) y_n) = E[x(t) y_n] \]
yielding the Wiener Hopf equation
\[ \sum_{n=-\infty}^{\infty} a_n(t) R_s(u - k) = E[x(t) y_n]. \]
(42)

To solve this we proceed as follows:
Since \( y_n \) is also band-limited, we have
\[ R_s(u) = \int_{-1/2}^{1/2} e^{2\pi i u \lambda} dP_s(\lambda) \]
\[ E[x(t) y_n] = \int_{-1/2}^{1/2} e^{-2\pi i u \lambda} \psi_s(\lambda) \\
\]  
(43)

Now let us first for simplicity, assume that
\[ d\Psi(\lambda) = \psi_s(\lambda) d\lambda \]
so that in this case, also:
\[ dP_s(\lambda) = \psi_s(\lambda) d\lambda. \]
We note that the ratio
\[ \frac{\psi_z(x)}{\psi_y(x)} = g(x) = \frac{\psi_z(x)}{C(2W\lambda) - \psi_y(x)} + a^2 \]  
(45)

where
\[ a^2 = \int_{-1/2}^{1/2} \left(1 - \left| C(2W\lambda) - \psi_y(x) \right| \right) \psi_y(x) \, d\lambda \]
is well defined. We then rewrite (44) as
\[ E[x(t)y_i] = \int_{-1/2}^{1/2} e^{-2\pi ikx} C(2W\lambda) g(\lambda) \psi_y(x) \, d\lambda. \]  
(46)

Substituting these into (42) we have
\[ \int_{-1/2}^{1/2} e^{-2\pi ikx} \sum a_i e^{2\pi ikx} g(\lambda) \psi_y(x) \, d\lambda = 0. \]
Hence the solution is given by
\[ \sum a_i e^{2\pi ikx} = C(2W\lambda) g(\lambda) e^{2\pi ikx}. \]
Hence
\[ a_i = \int_{-1/2}^{1/2} C(2W\lambda) g(\lambda) e^{2\pi ikx} \, d\lambda \]  
(47)

which thus generalizes the result (41).

In the general case, \( g(\lambda) \) is simply the Radon-Nikodym derivative of \( d\Psi(\lambda) \) with respect to \( dP_\lambda(\lambda) \), since the relation
\[ dP_\lambda(\lambda) = C(2W\lambda) g(\lambda) d\Psi(\lambda) + a^2 \]
implies that \( d\Psi(\lambda) \) is absolutely continuous with respect to \( dP_\lambda(\lambda) \), provided the characteristic function does not vanish.

The next step is to evaluate the error in using (47). Being an optimal estimate, the mean-square error is
\[ E[(x(t) - y(t))^2] = E[x(t)^2] - E[x(t)g(t)]. \]
Now
\[ E[x(t)g(t)] = \int_{-1/2}^{1/2} C(2W\lambda) g(\lambda) \, d\lambda \]
Hence
\[ E[(x(t) - y(t))^2] = \int_{-1/2}^{1/2} [1 - C(2W\lambda)] g(\lambda) \, d\lambda. \]  
(48)

As is to be expected, when we have no jitter, so that
\[ C(2W\lambda) g(\lambda) = 1, \]
we have zero error.

**Band-Limited Signals in Correlated Jitter**

We can proceed in a manner already outlined, for the general case where the jitter is also correlated. The basic (42) is again the same. The only difference is in the expression for \( g(\lambda) \), owing to the correlation. (The required absolute continuity of \( d\Psi(\lambda) \) with respect to \( dP_\lambda(\lambda) \) continues to hold.) The error is again given by (48).

Instead of "fitting" to obtain \( y(t) \), we may consider optimal operations on the samples to obtain a best (mean-square) estimate for some random variable \( Z \). In the case of no jitter, the best linear estimate is then given by
\[ \sum a_i x_i \]
where the coefficients \( a_i \) are of course determined from
\[ \sum a_i R_n(j - k) = E[Zx_i]. \]
To find how jitter modifies things, let us assume that for every \( t \),
\[ E[Zx(t)] = \int_{-W}^{W} e^{2\pi ikf} Z(f) \, df. \]
Then let the optimal weights be \( b_i \) so that
\[ \sum b_i y_i = Z^* \]  
and \( b_i \) must satisfy
\[ \sum b_i E[y_i y_i] = E[Zy_i] \]
which transforms to
\[ \sum b_i R_n(j - k) = \int_{-W}^{W} e^{2\pi ikf} C(2W\lambda) Z(2W\lambda) \, df. \]
Transforming the left side, we obtain
\[ \int_{-1/2}^{1/2} e^{2\pi ilk} \sum b_i e^{2\pi ikx} dP_\lambda(\lambda) \]
so that the optimal weights are now given by
\[ b_i = \int_{-1/2}^{1/2} e^{-2\pi ilk} C(2W\lambda) Z(2W\lambda) g(\lambda) \, d\lambda. \]  
(50)

In the jitter-free case, of course,
\[ a_i = \int_{-1/2}^{1/2} e^{-2\pi ilk} Z(2W\lambda) \, d\lambda. \]
The relation (50) thus represents a very general result.

**Optimal Nonlinear Operations**

So far we have only considered optimal linear operations. If the primary signal \( x(t) \) is not Gaussian, then the question arises as to how much better one can do using nonlinear operations on the jittered signal samples. In estimating \( x(t) \) from samples in the nonjittered case, the optimal operation is linear, with zero error, regardless of statistics. Hence one would expect that perhaps not much can be gained by additional nonlinear operations.
Without going into the most general nonlinear operation, let us consider first an estimate for $x(t)$ in terms of the $y_n$ as

$$x^*(t) = \sum_{n} P_n(t; y_n). \quad (51)$$

Let $P(t; n; x/y)$ be the conditional probability of $x(t)$ given $y_n$ and $P(n - m; x/y)$ that of $y_m$ given $y_n$. In this case the optimal mean-square estimate can be shown to satisfy the integral equation

$$\int xP(t, n; x/y_n) dx = \int \sum P_n(t; y_n)P(y_n/x_n) dy_n. \quad (52)$$

The conditional probabilities required here can be determined from the joint densities as follows. Thus let the joint density of $x(t_1), x(t_2)$ be

$$P(t_1, t_2; x_1, x_2).$$

Then the joint density of $y_n$ and $y_m$ is given by

$$\int P(n - mT + \xi_n - \xi_m; y_n, y_m)P(\xi_n, \xi_m) d\xi_n d\xi_m$$

and of course depends only $(n - m)$. Similarly the joint density of $x(t)$ and $y_n$ is given by

$$P(t; x, y) = \int P(t - nT - \xi_n; x, y)P(\xi_n) d\xi_n. \quad (53)$$

Let us assume the $P(t_1 - t_2; x_1, x_2)$ has an expansion of the form:

$$P(t_1 - t_2; x_1, x_2) = \sum_{n} a_n(t_1 - t_2)Q_n(x_1)Q_n(x_2)P(x_1)P(x_2).$$

Then a similar expansion holds for the joint density of $y_n$ and $y_m$:

$$P(y_n, y_m) = \sum_{n} b_n(n - m)Q_n(y_n)Q_n(y_m)P(y_n)P(y_m), \quad (54)$$

where

$$b_n(n - m) = E[a_n(n - mT + \xi_n - \xi_m)]. \quad (55)$$

Similarly in (53)

$$P(t; x, y) = \sum c(t - nT)Q(x)Q(y)P(x)P(y)\quad (56)$$

where

$$c(t - nT) = E[a_n(t - nT - \xi_n)]. \quad (57)$$

Substituting (57) into (52) we note that

$$\int xP(t, n; x/y_n) dx = C_n(t - nT)Q_n(y_n)$$

$$\cdot \int xQ_n(x)P(x) dx = C_n(t - nT)Q_n(y_n) \sqrt{R_n(0)},$$

since we can expand $P_n(t, y)$ as

$$P_n(t; y) = \sum_{k=0}^{\infty} a_n^*(t)Q_k(y).$$

Upon substitution into the right side of (52) it follows that

$$a_n^*(t) = 0 \quad \text{or} \quad k \neq 1$$

and

$$\sum_{n} b_n(n - m)a_n^*(t) = C_n(t - nT)$$

since

$$b_n(n - m) = E[R_n(n - mT + \xi_n - \xi_m)]$$

$$C_n(t - nT) = E[R_n(t - nT + \xi_n - \xi_m)]$$

$$Q_n(y) = \frac{y}{\sqrt{R_n(0)}},$$

we note that (58) reduces to (42). In other words, the optimum nonlinear estimate of the form (51) reduces to the linear, as we expected. In particular then the error (48) cannot be improved upon. Whether a more general nonlinear form than (51), including cross-producing, can lead to improvement is not clear, although doubtful.

Let us assume the $P(t_1 - t_2; x_1, x_2)$ has an expansion of the form:

$$P(t_1 - t_2; x_1, x_2) = \sum_{n} a_n(t_1 - t_2)Q_n(x_1)Q_n(x_2)P(x_1)P(x_2).$$

Then a similar expansion holds for the joint density of $y_n$ and $y_m$:

$$P(y_n, y_m) = \sum_{n} b_n(n - m)Q_n(y_n)Q_n(y_m)P(y_n)P(y_m), \quad (54)$$

where

$$b_n(n - m) = E[a_n(n - mT + \xi_n - \xi_m)]. \quad (55)$$

Similarly in (53)

$$P(t; x, y) = \sum c(t - nT)Q(x)Q(y)P(x)P(y)\quad (56)$$

where

$$c(t - nT) = E[a_n(t - nT - \xi_n)]. \quad (57)$$

Substituting (57) into (52) we note that

$$\int xP(t, n; x/y_n) dx = C_n(t - nT)Q_n(y_n)$$

$$\cdot \int xQ_n(x)P(x) dx = C_n(t - nT)Q_n(y_n) \sqrt{R_n(0)},$$

since we can expand $P_n(t, y)$ as

$$P_n(t; y) = \sum_{k=0}^{\infty} a_n^*(t)Q_k(y).$$

Upon substitution into the right side of (52) it follows that

$$a_n^*(t) = 0 \quad \text{or} \quad k \neq 1$$

and

$$\sum_{n} b_n(n - m)a_n^*(t) = C_n(t - nT)$$

since

$$b_n(n - m) = E[R_n(n - mT + \xi_n - \xi_m)]$$

$$C_n(t - nT) = E[R_n(t - nT + \xi_n - \xi_m)]$$

$$Q_n(y) = \frac{y}{\sqrt{R_n(0)}},$$

we note that (58) reduces to (42). In other words, the optimum nonlinear estimate of the form (51) reduces to the linear, as we expected. In particular then the error (48) cannot be improved upon. Whether a more general nonlinear form than (51), including cross-producing, can lead to improvement is not clear, although doubtful.

We have so far considered estimating $x(t)$. More generally we can consider estimates of a random variable $Z(t)$. It would appear that again the form of the optimal estimator is not changed by the jitter, although the actual coefficients change.

VII. AN APPLICATION TO TELEMETRY SYSTEMS

We shall now consider an application of some of the foregoing results to sampling problems in a class of telemetry systems where the sampling times are determined by a clock derived from the axis crossing of a sine wave and the jitter is that associated with noise in the system. We shall consider a model with additive noise. Let the sine wave be

$$A \sin \pi Wt$$

where the axis crossings occur at

$$t = nT$$

in the absence of noise. We will take the noise to be Gaussian, narrow-banded symmetric about the center frequency ($W/2$) and in the usual way, the noisy sine wave can then be expressed

$$A(t) \sin (\pi Wt + \phi(t))$$

where

$$A(t)^2 = x(t)^2 + (y(t) + A)^2$$

$$A(t) \sin \phi(t) = A + y(t)$$

$$A(t) \cos \phi(t) = x(t)$$

and $x(t)$ and $y(t)$ are independent Gaussians with band widths much less than $W/2$. Assuming $A^2$ is large in comparison with noise variance, the timing jitter can be taken to be solely due to the phase jitter. Also $\phi(t)$ is slow varying, so that the jitter $\xi_n$ as we have been using
it, can be taken to be defined by
\[ W_1 + \phi(nT + z) - 0 \]
which can further be approximated as
\[ W_1 - \phi(nT) + z, \phi'(nT) = 0 \]
and since \( \phi'(nT) \) is small compared to \( W \), we shall finally take
\[ z_n = -\frac{\phi(nT)}{\pi W}. \]  
(61)

Here we can take \( -\pi \leq \phi(nT) \leq \pi \), so that the jitter in this model is such that
\[ \frac{-T}{2} \leq z_n \leq \frac{T}{2}. \]

Now the statistics of the sequence \( \phi(nT) \) can be determined from (60), and have been given by various authors.\(^2,5,7\)

For our purposes, we need to obtain the characteristic functions, not readily available. Here by the usual change of variable to polar coordinates, we obtain
\[
C(u) = E[\exp iut] = \int_0^\pi \frac{1}{2\pi\sigma} e^{-Ax/2\sigma} e^{-Bx/2\sigma} R dR
\]
\[
= \sum_n C(n) \frac{\sin \pi u + n}{\pi(u + n)} \sin \pi u
\]  
(62)

where
\[
C(n) = e^{-A^2/2\sigma^2} \int_x^0 e^{-xI_1(A\sqrt{2x}/\sigma)} dx
\]
\[
= \frac{\Gamma(n + 1)}{\Gamma(n + 1)} \rho \cdot \mathbf{F}_1(-; -; -) \]

and \( \mathbf{F}_1(-; -; -) \) is the confluent hypergeometric function.

In particular, the mean is zero and the variance is
(by differentiating in (62)
\[ \sigma^2 = \frac{\pi^2}{3} + 4 \sum_n C(n) \frac{\cos \pi n}{n^2}. \]  
(63)

The characteristic function (two-variate)
\[ C(u_1, u_2) = E[\exp iu_1\phi(t) + iu_2\phi(t + \tau)] \]
can be similarly evaluated as
\[ C(u_1, u_2) = (1 - \lambda) \exp \left[ -\frac{A^2}{2\sigma^2} \frac{\lambda}{1 - \lambda} \sum_n \sum_r A_{m, n, r} \frac{\sin \pi (r + n + u_1) \sin \pi (m - r + u_2)}{\pi (r + n + u_1) \pi (m - r + u_2)} \right] \]  
(64)

where
\[ A_{m, n, r} = \int_0^\pi \int_0^\pi e^{-r_1^2 + r_2^2/2} I_n(m) I_r(n) I_r(n) \frac{1}{\sigma^2} r_1 \sqrt{1 + \lambda} \]  
\[ \lambda = E[x(t)x(t + \tau)]/E[x(t)^2]. \]  
(65)

In spite of the voluminous literature on the subject,\(^2,5,7\) simpler expressions than these for the characteristic functions appear to be unavailable.

Proceeding with the example, we can first calculate the (normalized) mean-square error of fit using (34) and (64) as approximately assuming a flat signal spectrum
\[ \varepsilon^2 = \frac{8}{3} \left( \frac{\pi^2}{3} + 4 \sum_n C(n) \frac{\cos \pi n}{n^2} \right) = \frac{8}{3} \left( \frac{\pi^2}{3} - \frac{4}{\sqrt{2}} \right) \]

retaining only the first term in the series, for large enough ratio \( A/\sigma \). The error of course goes to zero as \( A/\sigma \to \infty \), since \( \sigma^2 \) goes to zero.

To consider the effect on the spectrum we need to evaluate (65). The correlation coefficient \( \lambda \) therein can be estimated using again a flat spectrum for the noise, with \( B \) as the highest frequency, so that
\[ \lambda = \frac{\sin 2\pi B^2}{2\pi B^2}. \]

Hence for
\[ \tau = KT, \]
we have for the corresponding correlation
\[ \lambda = \frac{\sin 2\pi K^2}{2\pi K^2}, \]
where
\[ r = BT = \frac{B}{2W} \ll 1. \]

Hence the jitter is highly correlated. For \( \lambda \) close enough to unity, we can approximate (65) as
\[
(1 - \lambda)^2 e^{-A^2/2\sigma^2(1 - \lambda^2)} \sum_{-\infty}^{\infty} A_{m, n, r} \frac{\sin \pi (r + n + u_1) \sin \pi (m - r + u_2)}{\pi (r + n + u_1) \pi (m - r + u_2)}.
\]

Rather than go into lengthy calculations, we note at this point that for large signal-noise-ratio in the clock pulses, we may consider the jitter as approximately Gaussian with correlation close to unity, variance being given by (64). Thus the effect on the spectrum would appear to be small and can be evaluated as in Section III.