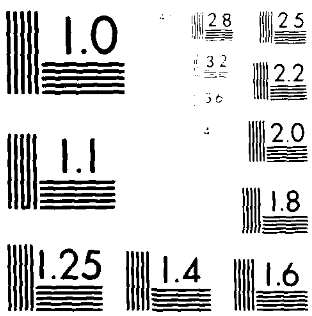
AD-A098 926 BROWN UNIV PROVIDENCE RI LEFSCHETZ CENTER FOR DYNAMI--ETC F/G 9/3
DIFFUSION APPROXIMATIONS FOR THE ANALYSIS OF DIGITAL PHASE LOCK--ETC(U)
NOV 80 H J KUSHNER, H HUANG, W T JU N00014-76-C-0279
UNCLASSIFIED LCDS-TR-80-2 NL

1 of 1
AD A
098926

END
DATE
FILMED
6-81
DTIC



MICROCOPY RESOLUTION TEST CHART

LCDS TECHNICAL REPORT 80-2

DIFFUSION APPROXIMATIONS FOR THE ANALYSIS OF DIGITAL PHASE LOCKED LOOPS

H. J. KUSHNER AND HAI HUANG

II. DIFFUSION APPROXIMATIONS FOR NONLINEAR PHASE LOOP-TYPE SYSTEMS WITH WIDE BAND INPUTS

H. J. KUSHNER AND WILFRED T. Y. JU

NOVEMBER 1980

Lefschetz Center for Dynamical Systems

Division of Applied Mathematics

DIFFUSION APPROXIMATIONS FOR THE ANALYSIS OF DIGITAL PHASE LOCKED LOOPS

Harold J. Kushner*

Hai Huang**

Abstract

Recent results for getting diffusion limits of a sequence of suitably scaled stochastic processes are applied to the synchronization problem for a digital phase locked loop (DPLL). The discrete time parameter error processes is suitably amplitude scaled and interpolated into a continuous time parameter process. For small filter gains and symbol intervals, a diffusion process approximation is rigorously obtained. This approximation is a Gauss Markov process and it yields approximate error variances, passage time distributions, correlation properties, (among other properties) for the DPLL. The tracking problem when the clock drifts is also treated. The technique is applicable to a wide variety of related problems, to get continuous time Markov systems which are easier to analyze the original (continuous or discrete time) systems which they approximate.

*Division of Applied Mathematics and Engineering, Brown University, Providence, R.I. This research was supported in part by the Air Force Office of Scientific Research under AFOSR 76-3063C, in part by the National Science Foundation under NSF Eng. 77-12946 and in part by the Office of Naval Research under N0014-76-C-0279-P0004.

**Division of Applied Mathematics, Brown University, Providence, R.I. This research was supported in part by the National Science Foundation under NSF-Eng. 77-12946 and in part by the Office of Naval Research under N0014-76-C-0279-P0004.

May, 1980

1. Introduction

In Chapter 9 of [1], Lindsey and Simon develop several interesting digital phase locked loops (DPLL) for the purpose of symbol synchronization. In theireffort to estimate the variances of the epoch estimation errors, it was assumed that the DPLL adjustment rate is slow and the errors small. Then an "equivalent" phase locked loop (PLL) was found, with an "equivalent" white noise input. The error variance of a linearized form of this PLL was then used as an approximation to the error variance of the DPLL. In the development of such a continuous time parameter approximation there are (or must be) either implicit or explicit amplitude scalings of the signal and noise and of the system gains. By speaking of an "equivalent PLL", and using it to estimate the error variances, there is at least the tacit recognition that for some suitable amplitude scaling of the error sequence, there is a continuous parameter interpolation of the error sequence which is close in some statistical sense to the output of the "equivalent" PLL. But the exact sense in which the PLL is "equivalent" or close is not clear, owing to the informality of the development and the use of a spectral analysis technique which fixed the state variable, and does not allow it to vary naturally. The general idea is useful, however, since owing to "central limit theorem" like effects, the complicated detailed structure of the DPLL would be replaced by a PLL with a white noise input, which is easier to analyze. When speaking of closeness of a DPLL and a PLL, we might mean that if the DPLL were parametrized (by, say, the symbol interval T or by a system gain), then as the parameter converged to (say) zero, the output of the DPLL converged in a suitable sense to the output of the PLL. Here, a systematic and rigorous way of doing this is developed. The technique has wide applicability. The specific end results are of the same type as obtained in [1], except that owing to the "weak convergence" nature of the approximation, much information on the DPLL beyond the error variances can be (approximately) obtained from the limit process.

Recently a very useful technique [2] has been developed for getting precise (in a sense to be described below) diffusion limits of a sequence of suitably scaled (and suitably interpolated into a continuous parameter processes) stochastic difference equations. Here, these methods are applied to the synchronization problem, and the correct approximating diffusion is obtained in a mathematically rigorous way. The limit could conceivably be interpreted as the output process of a particular PLL whose input noise is white Gaussian.+ But the important thing is that it is not necessary to make ad-hoc assumptions in the development. The method can be used to handle a wide variety of structurally similar problems in a systematic way. For specificity, we treat the scheme of Figure 9.34 of [1] under the noise assumptions there. See Figure 1 for the system. The same general scheme has been applied to other problems in [3]; namely, to get diffusion approximations to the "state" processes of a learning automata for adaptive telephone routing and an adaptive quantizer. The diffusion approximations are much easier to study than the original processes. Related "continuous time" methods have been applied in [4] to several "continuous time" problems.

+We do not emphasize this because the limit equation is quite simple-and an interpretation is not helpful.

The specific problem, scaling and interpolation will be developed in Section 2. The development is for the simple case connected with ([1], Figure 9.34). Extensions to more general noise, intersymbol interference and clock drift are discussed in Section 5. As will be clear, the technique gives more information than simply an approximation to the error variance. In Section 3, the general background theorem is given, together with some definitions from the theory of weak convergence of a sequence of stochastic processes. In Section 4, the theorems of Section 3 are applied to the problem of Section 2, and the main limit theorem obtained.

2. The DPLL; formulation, scaling and interpolation

The circuit is given in Figure 1, and Figure 2 gives the timing sequences. In this section and in Section 4, the signal sequence $\{s_n\}$ is a sequence of independent random variables, where $s_n = \pm A_0$ and $P\{s_n = A_0\} = \frac{1}{2}$, and $s(t)$ = input signal = s_n in the interval $[nT+\delta_0, nT+T+\delta_0)$, where δ_0 is the unknown epoch which is to be estimated. Since only the estimation errors are important, with no loss of generality we set $\delta_0 = 0$. The input noise $n_T(t)$ is white Gaussian, and its' power will be given below. Let $w_T(t) = \int_0^t n_T(s)ds$ = Wiener process with variance $\sigma_T^2 t$. We subscript $n_T(\cdot)$ and $w_T(\cdot)$ by T for reasons to be discussed below (2.3). More general signal and noise models will be discussed in Section 5. Let $\hat{\varepsilon}_n$ denote the n^{th} estimate of δ_0 and set $\lambda_n = (\hat{\varepsilon}_n - \delta_0)/T = \hat{\varepsilon}_n/T$.

The algorithm. Using the two parameters, λ_{n-1}, λ_n, define $e_n(\cdot,\cdot)$ (see Figure 1) by

$$
\begin{aligned}
(2.1) \qquad e_n(\lambda_{n-1},\lambda_n) = {} & |s_n(1-\Delta-\lambda_{n-1})T + s_{n+1}(\Delta+\lambda_n)T \\
& + w_T((n+1+\Delta+\lambda_n)T) - w_T((n+\Delta+\lambda_{n-1})T)| \\
& - |s_n(\Delta-\lambda_{n-1})T + s_{n+1}(1-\Delta+\lambda_n)T \\
& + w_T((n+2-\Delta+\lambda_n)T) - w_T((n+1-\Delta+\lambda_{n-1})T)|.
\end{aligned}
$$

Throughout, it is assumed (as in [1]) that $\Delta \le 1/4$. With use of a general finite memory linear filter, $\{\hat{\varepsilon}_n\}$, (λ_n), is defined by

$$
\begin{aligned}
\hat{\varepsilon}_{n+1} &= \hat{\varepsilon}_n + \gamma \sum_{i=0}^{K} \alpha_i e_{n-i}(\lambda_{n-i-1}, \lambda_{n-i}), \\
\lambda_{n+1} &= \lambda_n + \frac{\gamma}{T} \sum_{i=0}^{K} \alpha_i e_{n-i}(\lambda_{n-i-1}, \lambda_{n-i}),
\end{aligned} \tag{2.2}
$$

where $\gamma > 0$. The technique for (2.2) is very similar to that for (2.3). The limits all have the form (4.1) and differ only in the power of the input noise and in θ being replaced by $\theta(\sum \alpha_i)$. We will work with (2.3) for simplicity, where $g_n(\lambda, \lambda') = e_n(\lambda, \lambda')/T$.

$$
\lambda_{n+1} = \lambda_n + \frac{\gamma}{T} e_n(\lambda_{n-1}, \lambda_n) = \lambda_n + \gamma\, g_n(\lambda_{n-1}, \lambda_n). \tag{2.3}
$$

In any particular application, where T is fixed a-priori, σ_T^2 is determined from the problem data. But, consider a sequence of systems of the form (2.3), the sequence being parametrized by $T \to 0$. Assume, for purposes of this argument, that $\hat{\varepsilon}_n = 0$, and that $s_0 T + \int_0^T n_T(s)ds = y_0$ is used to estimate s_0 via a likelihood ratio. Thus, if $y_0 > 0$, then $s_0 = A_0 > 0$ is chosen. Note

$$
P\{\text{choosing } s_0 = A_0 > 0 \mid s_0 = -A_0\} = \int_{\frac{A_0 T}{\sigma_T \sqrt{T}}}^{\infty} dN(0,1),
$$

where $N(0,1)$ is the standard normal distribution. Thus, a natural parametrization is $\sigma_T^2 = \sigma^2 T$ for some constant σ. Note that the (noise power in a bandwidth of order $1/T$)/signal power) is constant, under the above scaling.

As $\gamma \to 0$, the continuous parameter interpolation (interpolation interval γ) of $\{\lambda_n\}$ converges to the solution of an ordinary differential equation which will not yield the detailed path information which is desired. A further normalization is required for this. It is convenient to write $g_n(\lambda_{n-1},\lambda_n)$ in the form

$$g_n(\lambda_{n-1},\lambda_n) = [(1-\Delta-\lambda_{n-1})s_n + (\Delta+\lambda_n)s_{n+1}$$
$$+ w(n+1+\Delta+\lambda_n) - w(n+\Delta+\lambda_{n-1})]$$
$$- [(\Delta-\lambda_{n-1})s_n + (1-\Delta+\lambda_n)s_{n+1} + w(n+2-\Delta+\lambda_n) - w(n+1-\Delta+\lambda_{n-1})]$$

where $w(\cdot)$ is a <u>Wiener process with variance</u>[+] $\sigma^2 t$. Next, define $U_n^\gamma = \lambda_n/\sqrt{\gamma}$ and define the process $U^\gamma(\cdot)$ by $U^\gamma(t) = U_n^\gamma$ on each interval $[n\gamma, n\gamma+\gamma)$. With parameter λ,λ' replacing λ_{n-1},λ_n, define $\bar{g}(\lambda,\lambda') \equiv Eg_n(\lambda,\lambda')$, and define $\xi_n^\gamma(\lambda_{n-1},\lambda_n) \equiv g_n(\lambda_{n-1},\lambda_n) - \bar{g}(\lambda_{n-1},\lambda_n)$. We can now write the normalized and centered iteration as

$$U_{n+1}^\gamma = U_n^\gamma + \gamma(\bar{g}(\lambda_{n-1},\lambda_n)/\sqrt{\gamma}) + \sqrt{\gamma}\,\xi_n^\gamma(\lambda_{n-1},\lambda_n). \tag{2.4}$$

Define the derivative $\frac{d}{d\lambda}\bar{g}(\lambda,\lambda)|_{\lambda=0} \equiv -\theta$. It can be shown that $\theta > 0$. For the analysis, it is convenient to expand (2.4) as

$$U_{n+1}^\gamma = U_n^\gamma - \gamma\theta U_n^\gamma + \gamma v_n + \sqrt{\gamma}\,\xi_n^\gamma(\lambda_{n-1},\lambda_n), \tag{2.5}$$

where v_n are $O(|\lambda_n|^2 + |\lambda_{n-1}|^2 + |U_n^\gamma - U_{n-1}^\gamma|)$ and $O(x)/|x|$ is bounded.

The limit theorem of Section 4 implies that $\{U^\gamma(\cdot)\}$ converges in distribution to a particular Gauss-Markov diffusion $U(\cdot)$ as $\gamma \to 0$.

[+]We can define $w(t) = w_T(tT)/T$.

This limit would be the output of the "equivalent" PLL, and only makes sense as a "near equivalence" if γ is small. In the cited section of [1], it is also supposed that the error estimates change slowly (small γ) and, in fact, that the λ_n are "constant" over a "long period" of time. The latter extraneous assumption is not needed here.

Properties of $\{\lambda_n\}$ are obtained from $\lambda_n = \sqrt{\gamma}\, U_n^{\gamma} \approx \sqrt{\gamma}\, U(n\gamma)$ or, equivalently, from $(n\gamma = t)$

$$\text{(2.6)} \qquad \sqrt{\gamma}\, U(t) \approx \lambda_{[t/\gamma]} \approx \sqrt{\gamma}\, U(nT \cdot \tfrac{\gamma}{T}).$$

Although the result does not depend on it, γ would normally depend on T, and the limit results suggest the appropriate form of the dependence. Since we are concerned with the behavior of $\{\lambda_n\}$ over real time intervals $\{n: nT \leq t\}$, by (2.6) we should have $\gamma \to 0$ as $T \to 0$. If $\gamma/T \to 0$ as $T \to 0$, then (2.6) implies that the system output becomes $\{U_n\}$, constant on any finite time interval as $T \to 0$. Let $nT = t$. If $\gamma/T \to \infty$ as $T \to \infty$, then $\lambda_{[t/T]}/\sqrt{\gamma} \approx U(nT\cdot\gamma/T) \approx U(\infty)$ ($U(\infty)$ has the limit distribution, as $t \to \infty$, of $U(t)$). In particular, let $\gamma = cT^{\alpha}$, $\alpha < 1$. Then the smaller is α, the larger are the errors. The best and most natural form in $\gamma = cT$. Then the change of λ_n per sample is proportional to the symbol interval width. The initial error λ_0 must be $O(\sqrt{\gamma})$, for otherwise the system (2.3) will not be able to improve the estimate for small γ.

Via the method of the next section, it can be shown that the v_n terms in (2.5) contribute nothing to the limit. For the sake of simplicity, we drop them now. Thus, henceforth we work with the partially linearized form

$$U^{\lambda}_{n+1} = U^{\lambda}_n - \lambda\beta U^{\lambda}_n + \sqrt{\lambda}\,\xi_n(\text{[illegible]}). \tag{2.7}$$

3. Mathematical Background

3a. Remarks on weak convergence theory. The theory of weak convergence of a sequence of probability measures is a powerful tool which has found applications in many areas of applied probability [5], [6], [9]. Only a few comments will be made here. For a full treatment, see [10]. Let $D[0,\infty)$ denote the space of real valued functions which are right continuous and have left hand limits. The piecewise constant process $U^{\gamma}(\cdot)$ can be treated as an abstract random variable with values in $D[0,\infty)$, and it induces a probability measure P_{γ} on it, (actually on the sets of $D[0,\infty)$ defined by a certain topology, called the Skorokhod topology, but this need not concern us here). The sequence $\{U^{\gamma}(\cdot)\}$ is said to be tight if for each $\delta > 0$, there is a compact set $K_{\delta} \in D[0,\infty)$ such that $P\{U^{\gamma}(\cdot) \in K_{\delta}\} \geq 1 - \delta$ for each δ. The sequence $\{U^{\gamma}(\cdot)\}$ converges weakly to a process $U(\cdot)$ if $U(\cdot)$ has paths in $D[0,\infty)$ and induces a measure P on it, and if for every bounded and continuous real valued function $F(\cdot)$ on $D[0,\infty)$,

$\int F(v)dP_{\gamma}(v) \to \int F(v)dP(v)$ as $\gamma \to 0$. If $\{U^{\gamma}(\cdot)\}$ is tight, then each subsequence contains a further subsequence which converges weakly to some process with paths in $D[0,\infty)$. In Section 4, it will be shown that for our problem all limits are actually the same Gauss-Markov process. The limit will give us the desired information about the errors and dynamics of $\{U_n^{\gamma}\}$ for small γ. Weak convergence is a substantial generalization of convergence in distribution.

Theorem 1 below gives criteria for tightness and weak convergence to a specific limit which are readily verifiable for our problem directly in terms of the problem data. Despite the abstract framework, the techniques are readily usable on problems such as the one of Section 2 and extensions, and the method of proof of Theorem 2 illustrates the relatively straightforward way in which the abstract Theorem 1 can often be applied.

3b. Remarks on the limit theorem. Let $B(\cdot)$ denote a standard Wiener process (covariance t) and $x(\cdot)$ the solution to the scalar stochastic differential equation

$$(3.1)\qquad dx = k(x,t)dt + v(x,t)dB,$$

where we suppose that $k(\cdot,\cdot)$ and $v(\cdot,\cdot)$ are continuous and that (3.1) has a unique solution (in the sense of distributions). Let R_n^γ denote the set $\{\xi_j^\gamma,\ j < n,\ U_j^\gamma,\ j \le n\}$, and let E_n^γ denote the conditional expectation given R_n^γ. Define the conditional "average difference" operator $\hat{A}^\gamma$ by $\hat{A}^\gamma f(n\gamma) = [E_n^\gamma f(n\gamma+\gamma) - f(n\gamma)]/\gamma$, where $f(\cdot)$ is a function which is constant on the intervals $[n\gamma, n\gamma+\gamma)$ and which depends on at most R_n^γ at time $n\gamma$. The operator A defined by

$$(3.2)\qquad Af(x,t) = k(x,t)\frac{\partial f}{\partial x} + \frac{1}{2}v^2(x,t)\frac{\partial^2 f}{\partial x^2}$$

is the differential generator of the process (3.1). If

$$(3.3)\qquad \hat{A}^\gamma f(U^\gamma(n\gamma),n\gamma) - (A+\partial/\partial t)f(U^\gamma(n\gamma),n\gamma) \to 0$$

as $\gamma \to 0$ for a suitably large class of functions f, then, under some subsidiary conditions one could conclude that $U^\gamma(\cdot) \to x(\cdot)$ weakly. Unfortunately, (3.3) is hard to get and does not hold in our case. Kurtz [7] showed that if (3.3) holds when the left hand f is "perturbed" to some f^γ which is close to f, then under some subsidiary conditions the processes will converge weakly. This point of view was developed and simplified in [8], [illegible] here, we use the form developed in [2], which is the most convenient for the purposes of this paper.

For purely technical reasons in the proof it is convenient to bound the process $U^\gamma(\cdot)$ in the manner given below. This bounding is used only in the theorem statement and as a technical device in the proofs. It does not affect the result. If, for each N, the [illegible] sequence of bounded processes converges weakly, then the original sequence converges as desired. Let $b_N(\cdot)$ denote a continuous function which is zero in $\{x: |x| \geq N+1\}$, equal to unity in $\{x: |x| \leq N\}$, and is infinitely differentiable. Define $\{U_n^{\gamma,N}\}$ by

$$(3.4) \qquad U_{n+1}^{\gamma,N} = U_n^{\gamma,N} - [\gamma\theta U_n^{\gamma,N} + \sqrt{\gamma}\,\xi_n^{\gamma,N}(\lambda_{n-1}^N, \lambda_n^N)]b_N(U_n^{\gamma,N}).$$

Here $\lambda_n^N/\sqrt{\gamma} = U_n^{\gamma,N}$ defines λ_n^N. The sequence $\{|U_n^{\gamma,N}|\}$ is stopped once it passes $N+1$. Let $U^{\gamma,N}(\cdot)$ be the piecewise constant process which equals $U_n^{\gamma,N}$ on the intervals $[n\gamma, n\gamma+\gamma)$. In Theorem 1, for each N, A^N stands for an operator of the form (3.2) whose coefficients are continuous and equal those of the operator A in the set $\{x: |x| \leq N\}$. The expressions $E_n^{\gamma,N}$ and

$\hat{A}^{\gamma,N}$ denote (resp.) expectation conditioned on $R_n^{\gamma,N} = \{U_j^{\gamma,N}, j \le n, \xi_j^{\gamma,N}(\lambda_{j-1}^N, \lambda_j^N), j < n\}$ and the "conditional average difference" operator

$$A^{\gamma,N} f(n\gamma) = [E_n^{\gamma,N} f(n\gamma+\gamma) - f(n\gamma)]/\gamma.$$

Theorem 1 is an adaptation of Theorems 2 and 3 of [2] to our problem. We use $\mathscr{D}$ = set of functions of (x,t) with compact support and whose mixed partial derivatives up to order 3 are continuous.

3c. The main background theorem.

Theorem 1. Assume the conditions on the coefficients of A^N and A given above, and on the uniqueness to the solution of (3.1). For each integer N and $f(\cdot,\cdot) \in \mathscr{D}$, suppose that there is a sequence of random functions $f^{\gamma,N}(\cdot)$ satisfying the following conditions: $f^{\gamma,N}(\cdot)$ is constant on each $[\gamma n, \gamma n+\gamma)$ interval and depends only on $U_j^{\gamma,N}$, $j \le n$, $\xi_j^{\gamma,N}$, $j < n$, there. For each N and $t_0 < \infty$ (recall $U_n^{\gamma,N} = U^{\gamma,N}(n\gamma)$)

(3.5) $\quad \sup\limits_{n,\gamma} E|f^{\gamma,N}(n\gamma)| < \infty, \quad \sup\limits_{n,\gamma} E|\hat{A}^{\gamma,N} f^{\gamma,N}(n\gamma)| < \infty, \quad n\gamma \le t_0$

(3.6) $\quad E|f^{\gamma,N}(n\gamma) - f(U_n^{\gamma,N}, n\gamma)| \to 0,$

$\quad E|\hat{A}^{\gamma,N} f^{\gamma,N}(n\gamma) - (\frac{\partial}{\partial t} + A^N) f(U_n^{\gamma,N}, n\gamma)| \to 0$ as $\gamma \to 0$, and $n\gamma \to t_0$.

Then if $\{U^{\gamma,N}(\cdot),\ \gamma > 0\}$ *is tight for each* N, *and* $U^{\gamma,N}(0)$ *converges to some* $U(0)$ *in distribution as* $\gamma \to 0$, *we have* $U^{\gamma}(\cdot) \to x(\cdot)$ *weakly, where* $x(\cdot)$ *solves* (3.1) *with* $x(0) = U(0)$.

The sequence $\{U^{\gamma,N}(\cdot),\ \gamma > 0\}$ *is tight for each* N *if for each* $t_0 < \infty$,

$$\lim_{\gamma\to 0} P\{\sup_{\gamma n\le t_0} |f^{\gamma,N}(n\gamma) - f(U_n^{\gamma,N}, n\gamma)| \ge \varepsilon\} = 0, \text{ each } \varepsilon > 0, \tag{3.7}$$

$$\lim_{K\to\infty} \overline{\lim_{\gamma>0}}\, P\{\sup_{\gamma n\le t_0} |\hat{A}^{\gamma,N} f^{\gamma,N}(n\gamma)| \ge K\} = 0. \tag{3.8}$$

4. The Limit Theorem

As noted in Section 2, the magnitude of the initial errors must be commensurate with the gain γ. Otherwise the system of Figure 1 will not function for small γ and T. For simplicity assume that there is a random variable U_0 such that $\lambda_0/\sqrt{\gamma} \to U_0$ as $\gamma \to 0$. In the work connected with [1, Figure 9.34], it was implicitly assumed that U_0 = stady state solution to (4.1) below.

Theorem 2. $\{U^\gamma(\cdot)\}$ converges weakly to the solution $U(\cdot)$ of the Gauss-Markov equation (4.1), as $\gamma \to 0$.

$$(4.1) \qquad dU = -\theta U dt + v dB, \quad u(0) = U_0.$$

In (4.1), $B(\cdot)$ is a standard Wiener process and (see above (2.4) for the definitions of $g_n, \bar{g}$) and v is given by

$$v^2 = 2E[g_{n+1}(0,0) - \bar{g}(0,0)][g_n(0,0) - g(0,0)]$$
$$+ E[g_n(0,0) - \bar{g}(0,0)]^2, \text{ any } n \geq 1,$$

Note. The form of v^2 is similar to that obtained formally by the method of [1, p. 445-447].

Proof. We need only verify the conditions of Theorem 1 for the process $\{U_n^{\gamma,N}\}$ of (3.4), for each fixed N. The proof is relatively straightforward. The systematic way in which the $f^{\gamma,N}(\cdot)$ are constructed is typical of the method in other problems. Henceforth, the test function $f(\cdot,\cdot) \in \mathcal{D}$ is fixed (recall the definition of $\mathcal{D}$ given above Theorem 1) and for notational convenience, we omit the superscript N on everything except $\hat{A}^{\gamma,N}$ and $E_n^{\gamma,N}$. We will get the perturbed test function $f^\gamma(\cdot)$ in the form $f^\gamma(n\gamma) = f(U_n^\gamma, n\gamma) + f_0^\gamma(n\gamma) + f_1^\gamma(n\gamma) + f_2^\gamma(n\gamma)$, where the f_i^γ are to be chosen sequentially such that the conditions of Theorem 1 hold.

Start by applying $\hat{A}^{\gamma,N}$ to $f(\cdot,\cdot)$:

$$
\begin{aligned}
(4.3)\quad \hat{A}^{\gamma,N} f(U_n^\gamma, n\gamma) = {} & \gamma f_t(U_n^\gamma, n\gamma) \quad + \quad o(\gamma) - \gamma f_u(U_n^\gamma, n\gamma)\, g(U_n^\gamma) b_N(\cdot) \\
& + \sqrt{\gamma}\, f_u(U_n^\gamma, n\gamma) b_N(U_n^\gamma) E_n^{\gamma,N} \xi_n^\gamma(\lambda_{n-1}, \lambda_n) \\
& + \gamma \frac{f_{uu}(U_n^\gamma, n\gamma)}{2} b_N^2(U_n^\gamma) E_n^{\gamma,N} (\xi_n^\gamma(\lambda_{n-1}, \lambda_n))^2 \\
& + o_{1n}^\gamma .
\end{aligned}
$$

The o_{1n}^γ is a remainder term in the truncated Taylor expansion and satisfies

$$
(4.4)\qquad o_{1n}^\gamma = E_n^{\gamma,N}\, O(\gamma^{3/2}(|\xi_n^\gamma(\lambda_{n-1}, \lambda_n)|^3 + 1)).
$$

For future use, note that (owing to the properties of the Wiener process) for each $t_0 < \infty$, and $N > 0$,

$$
(4.5)\qquad
\begin{aligned}
& \lim_{\gamma \to 0} \sup_{n\gamma \le t_0} |o_{1n}^\gamma / \gamma| = 0 \quad \text{w.p.1.}, \\
& \lim_{\gamma \to 0} \sup_{n\gamma \le t_0} E|o_{1n}^\gamma / \gamma| = 0.
\end{aligned}
$$

All ρ^{γ}_{kn} introduced below also satisfy (4.5).

Only the first and third terms on the right of (4.3) can be part of an operator such as $(\partial/\partial t + A^N)$. The 4th and 5th terms of (4.3) depend on the noise $\xi^{\gamma}_n(\lambda_{n-1},\lambda_n)$ as well as on $(U^{\gamma}_n, n\gamma)$ and need to be "averaged out". The perturbation f^{γ}_0 is chosen to "average out" the f_{uu} term. Define+

$$(4.6)\quad f^{\gamma}_0(n\gamma) = \frac{\gamma f_{uu}(U^{\gamma}_n,n\gamma)}{2} b_N(U^{\gamma}_n) \sum_{j=n}^{\infty} [E^{\gamma,N}_n(\xi^{\gamma}_j(\lambda_{n-1},\lambda_n))^2 - E(\xi^{\gamma}_j(\lambda_{n-1},\lambda_n))^2]$$

$$= \frac{\gamma f_{uu}(U^{\gamma}_n,n\gamma)}{2} b_N(U^{\gamma}_n)[E^{\gamma,N}_n(\xi^{\gamma}_n(\lambda_{n-1},\lambda_n))^2 - E(\xi^{\gamma}_n(\lambda_{n-1},\lambda_n))^2].$$

For our particular problem, due to the truncation effects of $b_N(\cdot)$, $|\lambda_n| \leq \delta \leq 1/4$ for small γ and the signal and Wiener process components of $\xi^{\gamma}_j(\lambda_{n-1},\lambda_n)$ (for $j > n$) are independent of $\{\xi^{\gamma}_j, j \leq n, U^{\gamma}_j, j \leq n\}$. Thus, the sum in (4.6) reduces to a single term. The general (and more complicated than needed) summation form for f^{γ}_0 is introduced here only because it is the appropriate form of $f^{\gamma}_0(n\gamma)$ for the generalizations of Section 5, and will facilitate the discussion there. For the same reason, $f^{\gamma}_1, f^{\gamma}_2$ are introduced in a summation form below, even though for the problem of this section, the sum reduces to a single term.

+When expectations of the form $E\,\xi^{\gamma}_j(\lambda_{n-1},\lambda_n)$, etc., are written, we mean $[E\,\xi^{\gamma}_j(\lambda,\lambda')]_{\lambda=\lambda_{n-1},\lambda'=\lambda_n}$; i.e., the λ_{n-1},λ_n are treated as parameters and considered to be fixed when computing the expectations. Also $\xi^{\gamma}_j(\lambda_{n-1},\lambda_n)$, $j \geq n$, is defined by $\xi^{\gamma}_j(\lambda,\lambda')$ with $\lambda = \lambda_{n-1}$, $\lambda' = \lambda_n$.

Now, applying $\hat{A}^{\gamma,N}$ to $f_0^{\gamma}(n\gamma)$ yields

$$\gamma \hat{A}^{\gamma,N} f_0^{\gamma}(n\gamma) = \gamma E_n^{\gamma,N} \frac{f_{uu}(U_{n+1}^{\gamma}, n\gamma+\gamma)}{2} b_N(U_{n+1}^{\gamma}) \cdot$$

$$\cdot [(\xi_{n+1}^{\gamma}(\lambda_n, \lambda_{n+1}))^2 - E(\xi_{n+1}^{\gamma}(\lambda_n, \lambda_{n+1}))^2]$$

$$(4.7) \qquad - \frac{\gamma f_{uu}(U_n^{\gamma}, n\gamma)}{2} b_N(U_n^{\gamma}) [E_n^{\gamma,N}(\xi_n^{\gamma}(\lambda_{n-1}, \lambda_n))^2$$

$$- E(\xi_n^{\gamma}(\lambda_{n-1}, \lambda_n))^2].$$

The part of the last term on the right containing the $E_n^{\gamma,N}$ is the <u>negative</u> of the next to the last term of (4.3). Also,

$$(4.8) \qquad \gamma E(\xi_n^{\gamma}(\lambda_{n-1}, \lambda_n))^2 = \gamma E(\xi_n^{\gamma}(\lambda_n, \lambda_n))^2 + o(\gamma).$$

For future use note that for each $t_0 > 0$ and $i = 0$,

$$(4.9) \qquad \lim_{\gamma\to 0} \sup_{n\gamma\le t_0} |f_i^{\gamma}(n\gamma)| = 0 \text{ w.p.1}, \quad \lim_{\gamma\to 0} \sup_{n\gamma\le t_0} E|f_i^{\gamma}(n\gamma)| = 0.$$

<u>Equation</u> (4.9) <u>will also hold for the</u> $f_1^{\gamma}, f_2^{\gamma}$ <u>introduced below</u>.

Replacing U_{n+1}^{γ} and λ_{n+1} by U_n^{γ} and λ_n, resp., in the <u>first term of</u> (4.7) alters that term only by a quantity o_{2n}^{γ} satisfying (4.5). In fact, o_{2n}^{γ} is bounded by (4.10). All the o_{kn}^{γ} introduced subsequently satisfy (4.5), but explicit bounds will not be given.

$$(4.10) \qquad O(\gamma^{3/2})E_n^{\gamma,N}[1+|\xi_n^\gamma(\lambda_{n-1},\lambda_n)|(1+|\xi_{n+1}^\gamma(\lambda_n,\lambda_n)|^2)] +$$
$$+ O(\gamma)E_n^{\gamma,N}|\xi_{n+1}^\gamma(\lambda_n,\lambda_{n+1})||w(n+2-\Delta+\lambda_{n+1})-w(n+2-\Delta+\lambda_n)| .$$

With $U_{n+1}^\gamma,\lambda_{n+1}$ replaced by U_n^γ,λ_n in the first term of (4.7) that term has the value zero, due to the independence of the increments of the Wiener process over non-overlapping time intervals.

Next, we turn to "averaging out" the $\sqrt{\gamma}$ term of (4.3). This will be done in two steps. Define

$$f_1^\gamma(n\gamma) = \sqrt{\gamma}\, f_u(U_n^\gamma,n\gamma)b_N(U_n^\gamma) \sum_{j=n}^{\infty} E_n^{\gamma,N}\xi_j^\gamma(\lambda_{n-1},\lambda_n)$$
$$= \sqrt{\gamma}\, f_u(U_n^\gamma,n\gamma)b_N(U_n^\gamma)E_n^{\gamma,N}\xi_n^\gamma(\lambda_{n-1},\lambda_n).$$

Applying $\hat{A}^{\gamma,N}$ to $f_1^\gamma(n\gamma)$ yields

$$(4.11) \quad \gamma\hat{A}^{\gamma,N}f_1^\gamma(n\gamma) = -f_1^\gamma(n\gamma) +$$
$$+ \sqrt{\gamma}\, E_n^{\gamma,N}b_N(U_{n+1}^\gamma)f_u(U_{n+1}^\gamma,n\gamma)\xi_{n+1}^\gamma(\lambda_n,\lambda_{n+1}) + \rho_{3n}^\gamma$$

where the ρ_{3n}^γ is due to the replacement of $n\gamma+\gamma$ by $n\gamma$. The function f_1^γ satisfies (4.9). The first term of (4.11) is the negative of the $\sqrt{\gamma}$ term of (4.3). The middle term of (4.11) will have to be averaged further.

The middle term on the right side of (4.11) can be expanded as

$$(4.12)\qquad \sqrt{\gamma}\, E_n^{\gamma,N}[f_u(U_n^{\gamma},n\gamma)b_N(U_n^{\gamma}) + (f_u(U_n^{\gamma},n\gamma)b_N(U_n^{\gamma}))_u(U_{n+1}^{\gamma}-U_n^{\gamma})]\xi_{n+1}^{\gamma}(\lambda_n,\lambda_{n+1}) + o_{4n}^{\gamma}.$$

The component of (4.12) involving $(f_u b_N)_u$ can be written as

$$(4.13)\qquad \gamma(f_u(U_n^{\gamma},n\gamma)b_N(U_n^{\gamma}))_u E_n^{\gamma,N}\xi_{n+1}^{\gamma}(\lambda_n,\lambda_{n+1})\xi_n^{\gamma}(\lambda_{n-1},\lambda_n) + o(\gamma)$$

$$= \gamma(f_u(U_n^{\gamma},n\gamma)b_N(U_n^{\gamma}))_u E_n^{\gamma,N}\xi_{n+1}^{\gamma}(\lambda_n,\lambda_n)\xi_n^{\gamma}(\lambda_{n-1},\lambda_n) + o_{5n}^{\gamma}.$$

The component of (4.12) involving $(f_u b_N)$ can be written as

$$\sqrt{\gamma}\, f_u(U_n^{\gamma},n\gamma)b_N(U_n^{\gamma})E_n^{\gamma,N}\xi_{n+1}^{\gamma}(\lambda_n,\lambda_n) + o_{6n}^{\gamma}(\gamma),$$

the first term of which equals zero by virtue of the independence of the increments of the Wiener process over non-overlapping time intervals.

Next, define f_2^{γ} by

$$f_2^{\gamma}(n\gamma) = \gamma(f_u(U_n^{\gamma},n\gamma)b_N(U_n^{\gamma}))_u \sum_{j=n}^{\infty}\sum_{k=j+1}^{\infty}[E_n^{\gamma,N}\xi_k^{\gamma}(\lambda_n,\lambda_n)\xi_j^{\gamma}(\lambda_{n-1},\lambda_n)$$

$$(4.14)\qquad - E\xi_k^{\gamma}(\lambda_n,\lambda_n)\xi_j^{\gamma}(\lambda_{n-1},\lambda_n)]$$

$$+ \sqrt{\gamma}\, f_u(U_n^{\gamma},n\gamma)b_N(U_n^{\gamma})\sum_{j=n}^{\infty}E_n^{\gamma,N}\xi_{j+1}^{\gamma}(\lambda_n,\lambda_n).$$

By the comment above (4.14), the second sum is zero (in the more general cases of Section 5, it will not necessarily be zero). Again, by the independence of the increments of $w(\cdot)$ over non-overlapping time intervals,

$$E_n^{\gamma,N}\zeta_k^\gamma(\lambda_n,\lambda_n)\zeta_j^\gamma(\lambda_{n-1},\lambda_n) = E\zeta_k^\gamma(\lambda_n,\lambda_n)\zeta_j^\gamma(\lambda_{n-1},\lambda_n), \ (k \geq j),$$

for $j \geq n+1$ and also for $j = n$ if $k > n+1$ (in which case both sides of the above equations equal zero). Thus (4.14) equals

$$(4.15) \quad \gamma(f_u(U_n^\gamma,n\gamma)b_N(U_n^\gamma))_u[E_n^{\gamma,N}\zeta_{n+1}^\gamma(\lambda_n,\lambda_n)\zeta_n^\gamma(\lambda_{n-1},\lambda_n) - E\zeta_{n+1}^\gamma(\lambda_n,\lambda_n)\zeta_n^\gamma(\lambda_{n-1},\lambda_n)].$$

It can be shown that

$$\gamma\hat{A}^{\gamma,N}f_2^\gamma(n\gamma) = -(4.15) + o_{7n}^\gamma + o(\gamma).$$

One component of (minus (4.15)) is the negative of the principal part of (4.13), the other component is the "averaged" centering term. Also, f_2^γ satisfies (4.9).

Summarizing the above calculations and recalling that $f^\gamma(n\gamma) = f(U_n^\gamma,n\gamma) + \sum_{i=0}^{2} f_i^\gamma(n\gamma)$, for each $t_0 < \infty$, $N < \infty$,

$$(4.16) \qquad \lim_{\gamma\to 0}\ \sup_{n\gamma\leq t_0} E|f^\gamma(n\gamma) - f(U_n^\gamma,n\gamma)| = 0,$$

$$\lim_{\gamma\to 0}\ \sup_{n\gamma\leq t_0} |f_1^\gamma(n\gamma) - f(U_n^\gamma,n\gamma)| = 0 \quad \text{w.p.1.}$$

Also, taking advantage of the cancellations in $\hat{A}^{\gamma,N}[f+f_0^\gamma+f_1^\gamma+f_2^\gamma]$, we have

$$\begin{aligned}\hat{A}^{\gamma,N}f^\gamma(n\gamma) = {} & f_t(U_n^\gamma,n\gamma) - (\upsilon U_n^\gamma)f_u(U_n^\gamma,n\gamma)b_N(U_\gamma^n) \\ (4.17) \qquad & + \frac{1}{2}\left(f_{uu}(U_n^\gamma,n\gamma)b_N(U_n^\gamma)E(\zeta_n^\gamma(\lambda_n,\lambda_n))^2\right. \\ & + (f_u(U_n^\gamma,n\gamma)b_N(U_n^\gamma))_u E\,\zeta_{n+1}^\gamma(\lambda_n,\lambda_n)\zeta_n^\gamma(\lambda_{n-1},\lambda_n) + o_{8n}^\gamma/\gamma.\end{aligned}$$

Changing ξ_n, ξ_{n+1} to zero in the right side of (4.17) alters that term by $O(\sqrt{\gamma})$. Define the operator A^N by

$$(\frac{\partial}{\partial t} + A^N) f(U_n^\gamma, n\gamma) = [\frac{\partial}{\partial t} + \frac{v_N^2(U_n^2, n\gamma)}{2} \frac{\partial^2}{\partial u^2} + k_N(U_n^\gamma, n\gamma)\frac{\partial}{\partial u}] f(U_n^\gamma, n\gamma)$$

= first four terms on right of (4.17), but with ξ_n, ξ_{n+1} replaced by 0.

By the properties of $b_N(\cdot)$, we have $v_N^2(u,t) = v^2$, $k_N(u,t) = \alpha u$ when $|u| \leq N$. Finally, since the solution of (4.1) is unique, all the conditions of Theorem 1 hold and the proof is completed. Q.E.D.

5. Extensions

5.1 General noise and intersymbol interference. Only an informal discussion will be given. With the appropriate scaling, the method is similar to that of the last section. Let $s_T(t) + \psi_T(t)$ denote the input, where $s_T(\cdot)$ = input signal, $\psi_T(\cdot)$ = stationary input noise with zero mean value.

Suppose that the channel memory is given by a function $h_T(\cdot)$. To keep the system from degenerating as $T \to 0$, we use $h_T(t) = h(t/T)$ for some transfer function $h(\cdot)$. For notational convenience assume that $h(\cdot) \neq 0$ only over a finite interval; in particular let $h(t) = 0$ for $t > Q$ for some integer Q. Let the waveform transmitted in the interval $[iT, iT+T)$ have the form $s_i q(t-iT)$, where $q(u) = 0$ out of $[0,T]$ and $\{s_i\}$ is a stationary sequence. Then

$$s_T(t) = \sum_{i=[\frac{t}{T}]-Q}^{[t/T]} s_i \int_0^t h(\frac{t-\tau}{T}) q(\tau - iT) d\tau .$$

Define

$$S_T(t) = \int_0^t s_T(u)du, \quad S(t) = S_T(t)/T = \int_0^{t/T} s_T(vT)dv.$$

The noise model is based on two considerations. First, for simplicity, we want the process $\psi_T(\cdot)$ to have only a finite memory (convenient, but not essential). Second, the considerations discussed below (2.3) still hold here; i.e., we want

$\operatorname{var} \int_0^T \psi_T(s)ds = \sigma_T^2 T = \sigma^2 T^2$ for some σ. To accomplish these aims we introduce a stationary random process $\psi(\cdot)$, define $\Psi(t) = \int_0^t \psi(s)ds$, assume that there is an integer R such that for each t_0, $\{\psi(s), s \leq t_0\}$ is independent of $\{\psi(s), s \geq t_0 + R\}$, and set $\psi_T(t) = \psi(t/T)$.

The finiteness assumptions connected with R, Q, guarantee that the sums defining $f_i^{\Delta,N}$ in Theorem 2 contain only a finite number of terms (with the signal and noise of this subsection used). The "tails" of these sums are all zero by the finiteness and independence assumptions.

Next, define $g_n(\lambda,\lambda')$, $\bar{g}(\lambda,\lambda')$, $U_n^\gamma,\xi_n^\gamma(\lambda,\lambda')$ as above (2.4), but with the noise and signal processes of this subsection; e.g., $g_n(\lambda,\lambda')$ has the representation

$$g_n(\lambda,\lambda') = [S(n+1+\Delta+\lambda') - S(n+\Delta+\lambda) + \Psi(n+1+\Delta+\lambda') - \Psi(n+\Delta+\lambda)]$$
$$- [S(n+2-\Delta+\lambda') - S(n+1-\Delta+\lambda) + \Psi(n+2-\Delta+\lambda') - \Psi(n+1-\Delta+\lambda)].$$

Again, set $-\mu = \frac{d}{d\lambda}\bar{g}(\lambda,\lambda)|_{\lambda=0}$ and suppose that $\mu > 0$, for otherwise the system (2.4) will be unstable for small $\gamma > 0$.

We will not go into the details, but the method of Theorem 2 works here also. Given $f \in \mathcal{D}$, the general (finite) summation forms of the $f_i^{\gamma,N}$ are used to get the perturbed test function $f^{\gamma,N}$ (recall that superscripts N were usually omitted in Theorem 2). We need to verify that (4.9) holds, and that (4.5) holds for the $o_{kn}^{\gamma,N}$ error terms. There can be verified under reasonable conditions on $\psi(\cdot)$. Assuming this, Theorem 2 holds but with the first term of v^2 replaced by (the sum contains at most $\max(Q,R)+1$ terms).

$$2\sum_{k=0}^{\infty} E\xi_{k+1}^\gamma(0,0)\xi_0^\gamma(0,0).$$

5.2. Random clock drift. For simplicity, return to the problem formulation of Sections 1 and 2, but suppose that the transmitter clock drifts. In particular, let the signal take the value $s(t) = s_i$ on $[\tau_i,\tau_{i+1})$ rather than as $[iT,iT+T)$, where $\tau_0 = \delta_0$, $\tau_{n+1} = \tau_n + T + \delta_{n+1}$, where δ_n is a zero mean random variable such that δ_n/T is small. Write $\delta_0^n = \sum_{i=0}^{n} \delta_i$. $\varepsilon_n = nT + \delta_0^n$. The system is given by Figure 1, and $\hat{\varepsilon}_n$ still denotes the estimate of the epoch δ_0^n. Set $\lambda_n = [\hat{\varepsilon}_n - \delta_0^n]/T$. We use the algorithm (2.3), (2.4) which we write in the form

$$\hat{\varepsilon}_{n+1} = \hat{\varepsilon}_n + \gamma e_n(\hat{\varepsilon}_{n-1}/T,\ \hat{\varepsilon}_n/T). \tag{5.1}$$

The integrator dumping timing is still given by Figure 2 but with the current definition of $\{\hat{\varepsilon}_n\}$. Figure 3 is merely a translation of Figure 2 into the "ε_n" notation. In particular, note that $(n+\Delta)T + \hat{\varepsilon}_{n-1} = \Delta T + \lambda_{n-1}T + \varepsilon_n - \delta_n$, and $(n+1-\Delta)T + \hat{\varepsilon}_{n-1} = (1-\Delta)T + \lambda_{n-1}T + \varepsilon_n - \delta_n$.

Define $g_n(\lambda_{n-1},\lambda_n) \equiv e_n(\hat{\varepsilon}_{n-1}/T, \hat{\varepsilon}_n/T)/T$ as in Section 2. Referring to Figure 3, note that

$$\begin{aligned} g_n(\lambda_{n-1},\lambda_n) = {} & [w(\Delta+\lambda_n + \tau_{n+1}/T - \delta_{n+1}/T) - w(\Delta+\lambda_{n-1} + \tau_n/T - \delta_n/T) \\ & + s_n[(1-\Delta-\lambda_{n-1} + (\delta_n+\delta_{n+1})/T)] + s_{n+1}(\lambda_n + \Delta - \delta_{n+1}/T) \\ & - [w(1-\Delta+\lambda_n + \tau_{n+1}/T - \delta_{n+1}/T) - [w(1-\Delta+\lambda_{n-1} + \tau_n/T - \delta_n/T) \\ & + s_n[(\Delta-\lambda_{n-1} + (\delta_n+\delta_{n+1})/T)] + s_{n+1}[1-\Delta+\lambda_n - \delta_{n+1}/T]]. \end{aligned} \tag{5.2}$$

$g_n(\lambda,\lambda')$ is defined as in (5.2) with parameters λ,λ' replacing λ_{n-1},λ_n, resp. Let $\bar{g}(\lambda,\lambda') = Eg_n(\lambda,\lambda')$. Next rewrite (5.1) as

$$\lambda_{n+1} = \lambda_n - \delta_{n+1}/T + \gamma g_n(\lambda_{n-1},\lambda_n). \tag{5.3}$$

Define ψ_n^γ by $\delta_{n+1}/T\sqrt{\gamma} \equiv \sqrt{\gamma}\,\psi_n^\gamma$. Define $g_n^\gamma(\lambda,\lambda')$, U_n^γ and $U^\gamma(\cdot)$ as above (2.4), but using the g_n of (5.2). Let $\alpha = -\frac{d}{d\lambda} g(\lambda,\bar{\lambda})|_{\lambda=\bar{\lambda}}$ and assume that $\alpha > 0$. As in Sections 2, 4, we work with the partially linearized system, which is

$$(5.4) \qquad U^{\gamma}_{n+1} = U^{\gamma}_n - \gamma\alpha U^{\gamma}_n + \sqrt{\gamma}\, \xi^{\gamma}_n(\lambda_{n-1},\lambda_n) - \sqrt{\gamma}\, \psi^{\gamma}_n .$$

In fact, the "linearization errors" go to zero on any finite interval $\{n: n\gamma \le T_0\}$ as $\gamma \to 0$.

For the sake of simplicity, let there be an N_0 not depending on T or γ such that $\{\psi^{\gamma}_i, i \le n\}$ is independent of $\{\psi^{\gamma}_i, i \ge n + N_0\}$ for each n. This is used only to assure that the sums $f^{\gamma,N}_i$ defined in Theorem 2 have a finite number of non-zero terms. Also, suppose that $\{\psi^{\gamma}_i\}$ is independent of $w(\cdot)$ (these assumptions are not necessary, but simplify the discussion). In order to effectively track changes in the timing, the drift terms $\{\delta_n\}$ must be "of the order of $\sqrt{\gamma}$" (loosely speaking). In particular, we assume that $\{\psi^{\gamma}_i\}$ is stationary, with a covariance not depending on γ.

The method of Theorem 2 can now be applied and the limit process is

$$(5.5) \qquad dU = -\alpha U dt + v_1 dB,$$

where (v^2 is defined in Theorem 2)

$$(5.6) \qquad v_1^2 = v^2 + [E(\psi^{\gamma}_n)^2 + 2 \sum_{i=1}^{\infty} E\psi^{\gamma}_i \psi^{\gamma}_0].$$

The added term in (5.6) is due to the $\{\sqrt{\gamma}\, \psi^{\gamma}_n\}$ clock drift process. If $\{\psi^{\gamma}_n\}$ were not independent of $w(\cdot)$, then there would be an additional "cross" term in (5.6).

REFERENCES

1. W.C. Lindsey, M.K. Simon, Telecommunication Systems Engineering, Prentice-Hall, Englewood Cliffs, 1973.

2. H.J. Kushner, "A martingale method for the convergence of a sequence of processes to a jump-diffusion process, to appear in Z. Wahrscheinlichkeitsteorie, Summer, 1980.

3. H.J. Kushner, Hai Huang, "Averaging methods for the asymptotic analysis of learning and adaptive systems, with small adjustment rate", submitted to SIAM J. on Control and Optimization.

4. H.J. Kushner, "Diffusion approximations to output processes of non-linear systems with wide band inputs, with applications", to appear IEEE Trans. on Information Theory.

5. D.E. Iglehart, "Diffusion approximations in applied probability", Math. of the decision sciences, Part II, Lectures in Appl. Math., Vol. 12, Amer. Math. Soc., Providence, 1968.

6. H.J. Kushner, Probability Methods for Approximations in Stochastic Control and for Elliptic Equations, Academic Press, New York, 1977.

7. T.G. Kurtz, "Semigroups of conditional shifts and approximation of Markov processes", Ann. Prob. 4, pp. 618-642, 1975.

8. H.J. Kushner, "Jump-diffusion approximations for ordinary differential equations with wide-band random right hand sides", SIAM J. on Control and Optimization, 17, pp.729-744, 1979.

9. G. Blankenship, G.C. Papanicolaou, "Stability and control of stochastic systems with wide-band noise disturbances, SIAM J. on Appl. Math. 34, pp. 437-476, 1978.

10. P. Billingsley, Convergence of Probability Measures, John Wiley and Sons, New York, 1968.

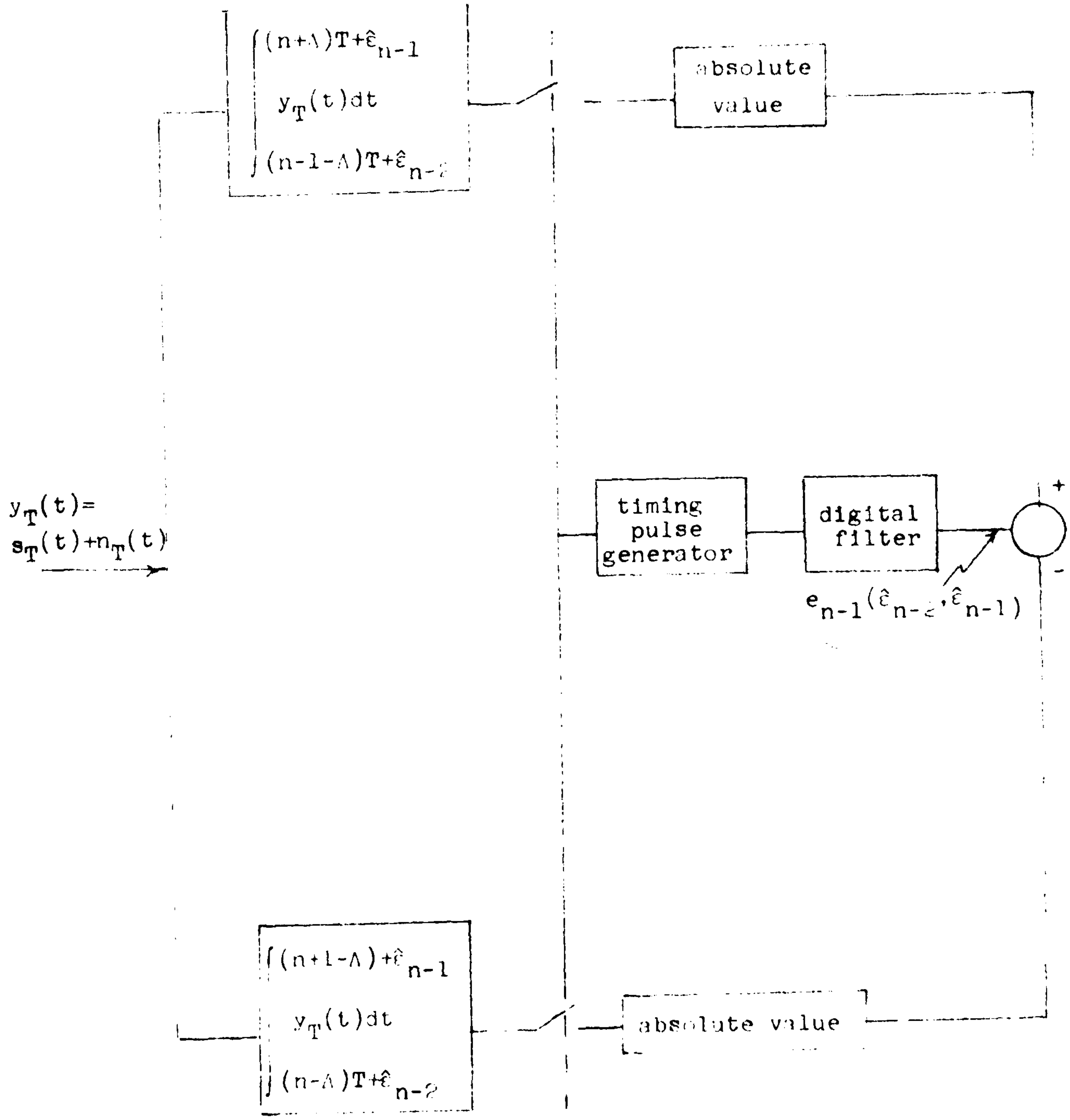
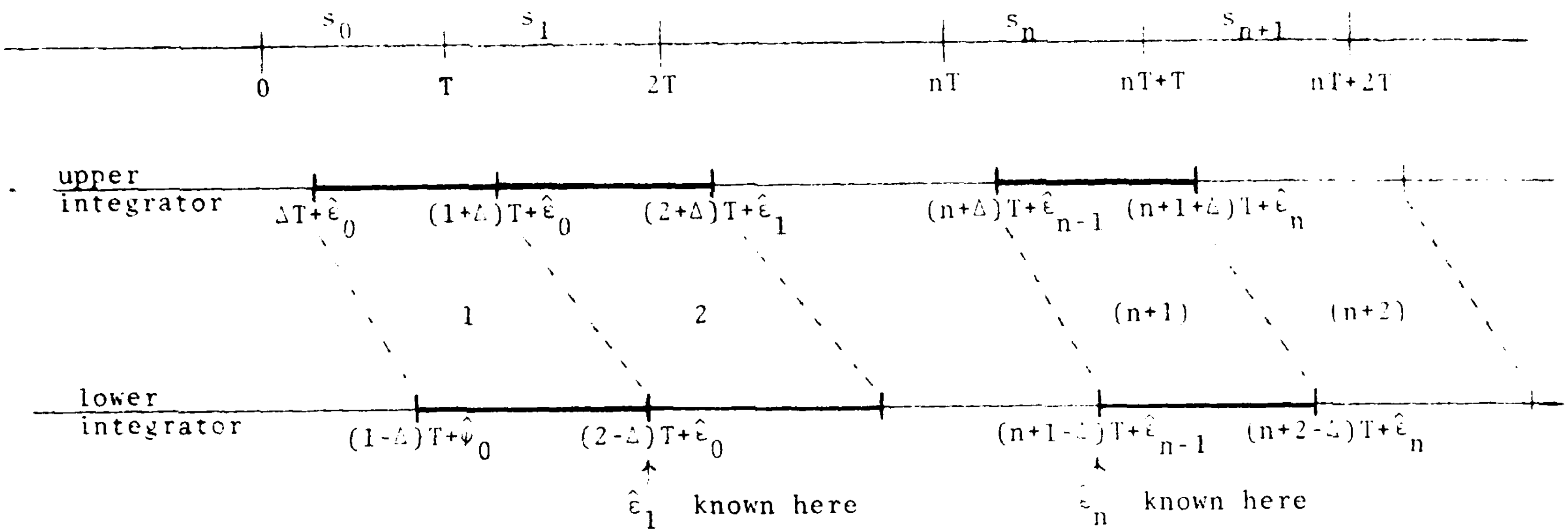


Fig. 1

The DPLL model, $\Delta \leq \frac{1}{4}$



The intervals up to and including the $(n+1)^{st}$ are used to get $\hat{\varepsilon}_{n+1}$ and $\xi_n^{\gamma}(\Delta_{n-1},\Delta_n)$, $\hat{\varepsilon}_n/T = \Delta_n$.

FIGURE 2. The Timing Sequence; Integrator Dump Times.

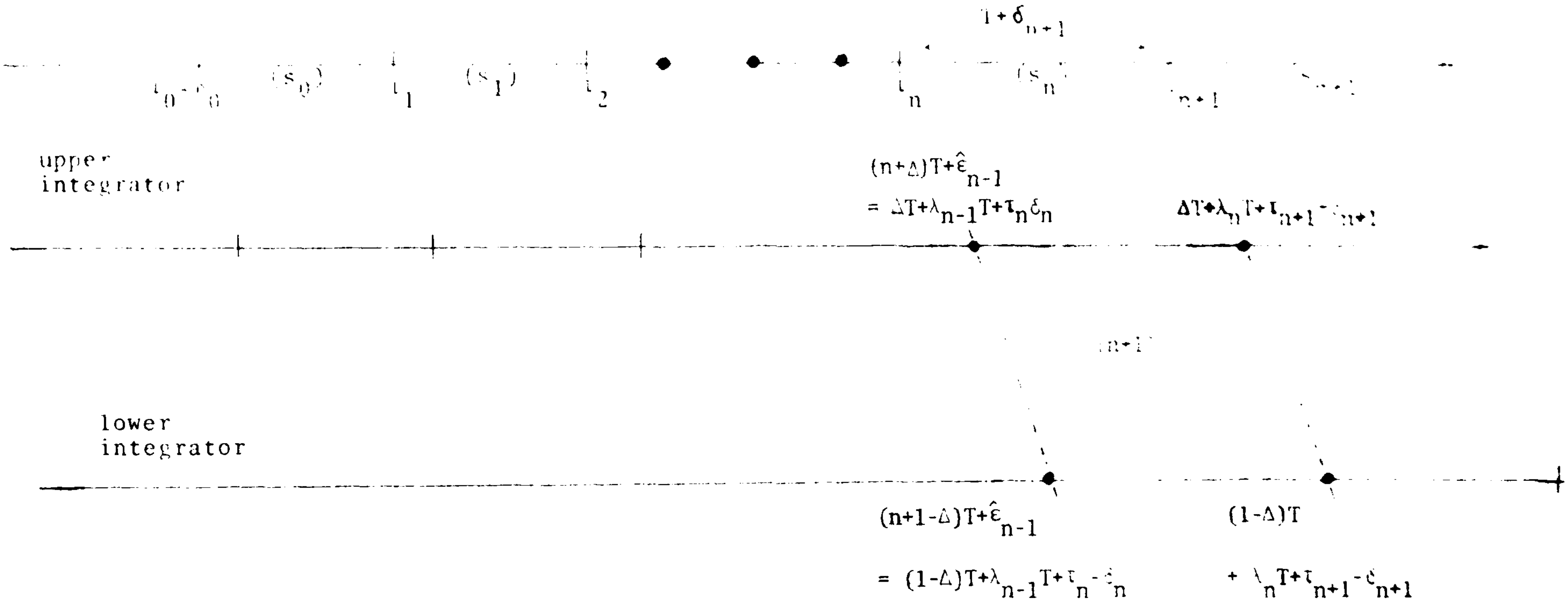


FIGURE 3. Timing with clock drift.

DIFFUSION APPROXIMATIONS FOR NONLINEAR PHASE LOCKED LOOP-TYPE SYSTEMS WITH WIDE BAND INPUTS.

by

Harold J. Kushner[†]
Division of Applied Mathematics & Engineering

Wilfred T.Y. Ju*
Division of Applied Mathematics
Brown University
Providence, R.I. 02912

November 1, 1980

[†] This research has been supported in part under contract #AFOSR-763063D of the Air Force of Scientific Research, in part by the National Science Foundation under grant NSF-Eng. 77 12946, and in part by the Office of Naval Research under grant N00014-76-C-0279-P0004.

*This research was supported in part by the National Science Foundation, under grant NSF-ENG. 77-12946.

ABSTRACT

Communication systems often involve differential equations models whose inputs are noises and signals with wide bandwidths. It is frequently of interest to approximate them by some Markov-diffusion process, since then many analytical and numerical methods can be used. Here, recent results on getting diffusion approximations to systems with such inputs are applied to three classes of detection systems which are very important in applications: (1). A phase locked loop with a limiter; (2) a quadricorrelator with and without a limiter (the function is to track changes in phase and frequency); (3) a 'squaring' loop, whose purpose is the tracking of the carrier frequency, despite the carrier modulation. In (3), a type of pulse phase modulation is used. The method is natural, systematic and relatively straightforward. Under natural scalings of the signals and noises, the appropriate diffusion approximations (for band-pass, but wide-band noise) are obtained. The approximation is in the sense of weak convergence. The first two problems have been hard to analyze owing to the nature of the non-linearity, and the results clearly indicate the advantage and disadvantages of the use of the limiter. The third problem has been difficult to analyze, partly due to the periodicities which occur naturally in such problems. All three classes represent widely used and important systems, and much information can be obtained from the limit process. For example, the results show that the use of a limiter can actually improve the tracking ability of the systems, when the noise is small. The

system signal and noise models to which the methods can be applied is much broader than those used here. But the results, together with the results in [4] for different classes of problems, illustrate the great potential of the approximation methods for problems in control and communication theory. In certain cases, the limit processes are of the type which have been obtained via more formal arguments.

I. Introduction

Diffusion approximations to the output and state variable processes for several types of phase locked loops (PLL), Costas loops, and related systems are obtained when the input noise is 'bandpass,' but with a wide bandwidth. The systems are commonly used to estimate and track the phase and frequency of received signals (with additive noise). There is a vast communication theory literature on the subject, and there are very many useful methods for the analysis of such systems [1]-[3]. Yet, it is only recently that rigorous methods for getting the diffusion approximations for more complicated and non-linear systems have become available. We will use one such method here.

Three important cases are of particular interest where, owing to the nature of the non-linearity or other system feature, the analysis has been difficult. In the first two cases, the system contains limiters (Figure 1b), a frequently used type of non-linearity.

Markov-diffusion approximations to the output and state processes of non-linear systems with wideband inputs are a major concern in communication (and control) theory because a large number of analytical and numerical techniques can be used on the approximation. The original system is often too complicated for much insight into its properties to be obtained otherwise. The fact that the bandwidth of the input process is often wide allows diffusion approximation or averaging methods to be fruitfully used to get the approximations.

Reference [4] illustrated the application of the general method of [5] to get diffusion approximations for several standard problems in communication theory. Using a related result, the investigation is continued here on the different (and perhaps harder) problems cited above. Reference [6] extends the result in [5] and provides a simpler proof under simpler conditions; but from the point of view of applications, the theorems of [5] and [6] are used in exactly the same way. Here we use the theorem in [6], because the conditions are simpler. In Section II, the main background theorem is stated. The basic idea is that the original system state, $x^{\varepsilon}(\cdot)$, is parameterized by ε, and as $\varepsilon \to 0$, the input noise bandwidth (BW) goes to ∞. Under reasonable conditions, the basic background theorem allows us to conclude that $x^{\varepsilon}(\cdot)$ converges weakly to a particular diffusion process $x(\cdot)$. Section III deals with the basic phase locked loop, with and without a limiter. Section IV treats a form of quadricorrelator with no limiter, and the limiter is added in Section V. (This system is a more sophisticated form of phase locked loop. It is used to track when the frequency errors are larger.) In Section VI, we treat a "squaring loop" whose purpose is to accurately track changes in the carrier frequency in presence of modulation, and we investigate the effects of the carrier modulation on the tracking errors when the noise intensity is small. Despite the mathematical nature of Section II, the basic results can often be used in a relatively straightforward way.

Owing to the differences in the problems treated here and in

[4], and in the types of noise used, many of the details are different. We concentrate on the differences, building on the results in [4] where possible, but often omitting details where they are similar to those in [4].

II. Mathematical Background

We suppose that the reader is familiar with the weak convergence terms and ideas as used, for example in [4], Section 2. Formally, suppose that the system is given by $\dot{x}^\varepsilon = H^\varepsilon(n^\varepsilon, x^\varepsilon)$, where n^ε is an input noise process whose BW $\to \infty$ as $\varepsilon \to 0$. We are interested in showing that $x^\varepsilon(\cdot)$ converges weakly to some diffusion $x(\cdot)$

$$dx = \alpha(x)dt + \sigma(x)dw, \tag{2.1}$$

with differential generator $A = \sum_i \alpha_i(x)\partial/\partial x_i + \frac{1}{2}\sum_{i,j} a_{ij}(x)\partial^2/\partial x_i\partial x_j$ where $a(x) = \{a_{ij}(x)\} = \sigma(x)\sigma'(x)$. Define the truncated process $x^{\varepsilon,N}(\cdot)$ by $\dot{x}^{\varepsilon,N} = H^\varepsilon(n^\varepsilon, x^{\varepsilon,N})b_N(x^{\varepsilon,N})$, where $b_N(x) = 1$ for $|x| \le N$ and equals zero for $|x| > N+1$. Let A^N be the differential generator of a diffusion process $x^N(\cdot)$ with coefficients $\alpha^N(\cdot)$, $\sigma^N(\cdot)$ equal to $\alpha(\cdot)$, and $\sigma(\cdot)$ in $\{x: |x| \le N\}$. If [6] $\{x^{\varepsilon,N}(\cdot)\} \to x^N(\cdot)$ weakly for each N, then $\{x^\varepsilon(\cdot)\} \to x(\cdot)$ weakly. The truncation is used because it is easier to work with bounded processes in the proof of the background theorem. It is a technical device, not an assumption on the original problem data. Next, we define some terms and then state the basic background theorem which is to be applied in the sequel. $b_N(\cdot)$ is assumed to be continuously differentiable.

Let $\hat{\mathcal{C}}_0$ denote the space of real valued continuous functions on R^r with compact support and $\hat{\mathcal{C}}_0^{\alpha,\beta}$ the subspace of functions in $\hat{\mathcal{C}}_0$ whose mixed α partial t-derivatives and β partial x-derivatives are continuous. Let $\{\mathcal{F}_t^\varepsilon\}$ be a non-decreasing sequence of σ-algebras with $\mathcal{F}_t^\varepsilon$ measuring $\{\eta^\varepsilon(s), s \le t\}$. Let $\mathcal{L}$ denote the class of real valued (progressively) measured (ω,t) functions such that if $g(\cdot) \in \mathcal{L}$, then $\sup_t E|g(t)| < \infty$, $E|g(t+\delta) - g(t)| \to 0$ as $\delta \to 0$, and $g(t)$ depends only on $\{\eta^\varepsilon(s), s \le t\}$. Let E_t^ε denote expectation conditioned on $\mathcal{F}_t^\varepsilon$. We say $\text{p-lim}_{\delta\to 0} f^\delta(\cdot) = 0$ if $\sup_{\substack{\delta>0\\ t>0}} E|f^\delta(t)| < \infty$ and for each t, $E|f^\delta(t)| \to 0$ as $\delta \to 0$. Define the operator $\hat{A}^\varepsilon$ with domain $\mathcal{D}(\hat{A}^\varepsilon)$ as follows: $g \in \mathcal{D}(\hat{A}^\varepsilon)$ and $\hat{A}^\varepsilon g(\cdot) = q(\cdot)$ iff $g(\cdot)$ and $q(\cdot)$ are in $\mathcal{L}$ and $\text{p-lim}_{\delta\to 0} \left[\frac{E_t g(\cdot + \delta) - g(\cdot)}{\delta} - q(\cdot)\right] = 0$. So, $\hat{A}^\varepsilon$ is a type of infinitesimal operator. The following theorem is a special case of that in [6]. A more complicated form was used in [4].

Theorem 1. Let (2.1) have a unique solution in the sense that any two solutions induce the same measure on the usual space of continuous functions. Fix N. For each $f(\cdot) \in \hat{C}^{2,3}$ let $\mathcal{L}$ be a sequence $\{f^{\varepsilon,N}(\cdot)\} \in \mathcal{L}$ such that

$$\text{p-lim}_{\varepsilon\to 0} |f^{\varepsilon,N}(\cdot) - f(x^{\varepsilon,N}(\cdot),\cdot)| = 0$$

$$\text{p-lim}_{\varepsilon\to 0} |\hat{A}^\varepsilon f^{\varepsilon,N}(\cdot) - (A^N + \frac{\partial}{\partial t}) f(x^{\varepsilon,N}(\cdot),\cdot)| = 0 \quad .$$

Then, if $\{x^{\varepsilon,N}(\cdot)\}$ is tight for each N, $\{x^\varepsilon(\cdot)\}$ converges weakly to $x(\cdot)$ as $\varepsilon \to 0$.

Note. Tightness is often not hard to prove. For our case the method of [7] as adapted in [6] can easily be used. The $\{f^{\varepsilon,N}(\cdot)\}$ are found by essentially the same method as used in [4], [5], [6], and [7]. We use the form $f^{\varepsilon,N}(t) = f(x^{\varepsilon,N}(t),t) + \sum_{i=0 \text{ or } 1}^{2} f_i^{\varepsilon,N}(t)$, where the $f^{\varepsilon,N}(\cdot)$ will be defined in the following sections. Henceforth, in order to minimize notation and detail, the N and b_N will be dropped, and where needed we simply assume that the processes $x^{\varepsilon}(\cdot)$ are bounded (as they are because we work with the truncation $x^{\varepsilon,N}(\cdot)$).

III. Phase Locked Loops with a Limiter

The system is described in Figure 1. First (Section III.1) we work with the smooth approximation $g_\alpha(\cdot)$ to the ideal limiter $g(\cdot)$. See Figure 1a. We get the diffusion process limit $x(\cdot)$ of the sequence $\{x^\varepsilon(\cdot)\} = \{v^\varepsilon(\cdot),\theta^\varepsilon(\cdot)\}$ as $\varepsilon \to 0$, then $\alpha \to 0$. The derivative of $g_\alpha(\cdot)$ is assumed bounded by some K/α and the filter in Figure 1 is simply the state variable representation of an arbitrary low pass filter. In Section III.2, we work with the hard limiter $g(\cdot)$ of Figure 1b directly. The limit diffusion is the same in both cases, and we develop the result for both cases in order to illustrate the robustness of the performance of the system of Figure 1 to mild changes in the non-linearity. This robustness is clearly necessary for a practical system.

In analyses of PLL's (even without limiters) it is usually supposed that the input noise is wide-band [9] and the limits sought (explicitly or implicitly) as the BW $\to \infty$. It is possible for both A_0 and θ^ε to depend on time, and the signal might then consist of the variations in θ^ε or A_0. But for our calculations in this section, A_0 is held fixed and $\theta^\varepsilon(\cdot) = \theta(\cdot)$, a differentiable function. If a more general $\theta(\cdot)$ were used (say a right-continous Markov process), then the infinitesimal operator of that process would play the role that the differential operator plays in the sequel. The result is the same. We are interested in the problem for large input noise and signal BW, say of the order $O(1/q_\varepsilon^2)$, where $q_\varepsilon \to 0$ as $\varepsilon \to 0$. Thus the center frequency ω_0^ε must tend to ∞ as $\varepsilon \to 0$.

We use $\omega_0^\varepsilon = \omega_0/\varepsilon^2$, $\varepsilon/q_\varepsilon \to 0$ as $\varepsilon \to 0$, so that the center frequency is large relative to the bandwidth, as in practical systems. This scaling is appropriate for the problem and consistent with heuristic methods for analyzing such systems. The gain $L_\varepsilon = L/q_\varepsilon$ is needed, either before or just after the filter, because otherwise the input to the VCO[+] will go to zero as $\varepsilon \to 0$ owing to the effects of the wide-band input noise [4].

We next describe the noise model. The noise model is a standard one for band limited noise and is suitably scaled for our method. We suppose that the noise is Gaussian, although this is not always necessary and the modification of the result for non-Gaussian noise will be stated when available. Let $z_i(\cdot)$, $i = 1,2$, denote independent real-valued stationary continuous Gaussian processes with unit variance and a correlation function $\rho(\cdot)$ which decreases to zero at an exponential rate. Let φ_i, $i = 1,2$ be random variables, uniformly distributed on $[0,2\pi]$ and such that $\{z_i(\cdot), \varphi_i, i = 1,2\}$ are mutually independent. Write $z_i^\varepsilon(t) = z_i(t/q_\varepsilon^2)$ and define the noise

$$n^\varepsilon(t) = [z_1^\varepsilon(t)\cos(\omega_0^\varepsilon t+\varphi_1) + z_2^\varepsilon(t)\sin(\omega_0^\varepsilon t+\varphi_2)]\sigma/q_\varepsilon. \tag{3.1}$$

If $S(\omega)$ is the spectral density of $z_i(\cdot)$, then the spectral density of $n^\varepsilon(\cdot)$ is $S(q_\varepsilon^2(\omega-\omega_0^\varepsilon)) + S(q_\varepsilon^2(\omega+\omega_0^\varepsilon))$. The choice of ω_0, ε, q_ε in any particular problem is determined by the problem

---

[+] The VCO (voltage controlled oscillator) is an oscilllator whose frequency deviation from a 'central' frequency is proportional to the input signal.

data and will be commented on below. For simplicity we set $L=k_0=1$. Their values can be incorporated into C (see Fig. 1).

We now make some simplications. First, we note that the center noise frequency ω_0^ε can be changed to any ω_1^ε such that $q_\varepsilon|\omega_1^\varepsilon-\omega_0^\varepsilon| \to 0$ as $\varepsilon \to 0$, without altering the results. Next, we drop the ξ_i from (3.1) for notational simplicity. This does not alter the results. Also, for notational simplicity we specialize the noise to the following Gauss-Markov case $z_i(t) = C_1 Z_i(t)$, $dZ_i = A_1 Z_i dt + B_1 dw$, where $w(\cdot)$ is a standard Wiener process and the roots of A_1 are in the open left hand plane. E_t^ε denotes conditioning on $\{Z^\varepsilon(s),\ s \le t\}$, where $Z(t) = \{Z_1(t), Z_2(t)\}$ and $Z^\varepsilon(t) = Z(t/q_\varepsilon^2)$.

Assuming (for the moment) that the multiplier device does nothing but multiply, its output is

$$\frac{\sigma}{2q_\varepsilon}[z_1^\varepsilon(t)(\cos\hat\theta^\varepsilon + \cos(2\omega_0^\varepsilon t + \hat\theta^\varepsilon)) + z_2^\varepsilon(t)(\sin(-\hat\theta^\varepsilon) + \sin(2\omega_0^\varepsilon t + \hat\theta^\varepsilon))] + \frac{A_0}{2}[\sin(\theta^\varepsilon-\hat\theta^\varepsilon) + \sin(\theta^\varepsilon+\hat\theta^\varepsilon+2\omega_0^\varepsilon t)]. \tag{3.2}$$

In analyzing systems with practical rather than with ideal multipliers, it is common practice to assume that the multiplier has a "low pass filter" incorporated within it, and to drop the terms in (3.2) containing $2\omega_0^\varepsilon t$. We make this assumption also.

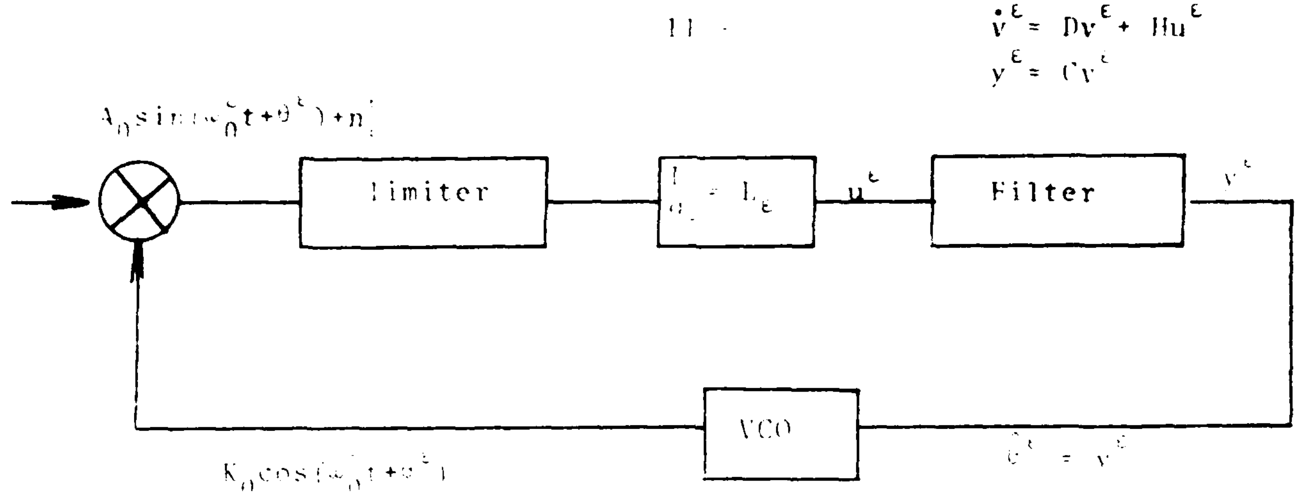
We want to retain a structure which allows the signal BW to be $O(1/q_\varepsilon^2)$. In fact, a filter would often be used before the multiplier to limit the input noise BW to that of the signal. Thus, for the moment, suppose that there is a filter in the multiplier with cutoff frequency $O(1/q_\varepsilon^2) << 2\omega_0^\varepsilon$. In the theoretical analysis (see Section II), the true $\dot{\theta}^\varepsilon$, $\dot{v}^\varepsilon$ are actually multiplied by $b_N(\cdot)$ and limits taken as $N \to \infty$, <u>then</u> $\varepsilon \to 0$. Thus, in the analysis, the derivatives are bounded uniformly in ε, for each N. This fact can be used to show that the terms in (3.2) containing $2\omega_0^\varepsilon t$ have no effect in the limit. But it is easier to simply make the assumption in the sentence below (3.2). All other filtering actions are incorporated explicitly into the filter box in Figure 1.

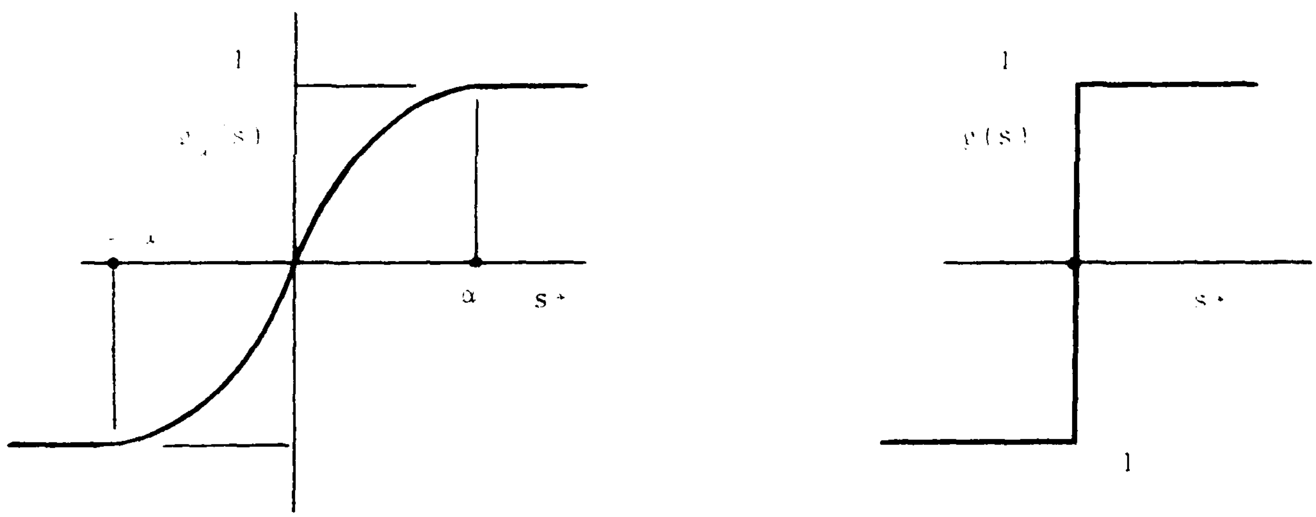
The input level A_0 can either be constant or time varying. We suppose for convenience that it is constant, and note the following for the time varying case. Let $A_1(\cdot)$ denote a bounded process with mean $EA_1(t) = \bar{A}_0$. (If the mean value $\bar{A}_0$ is periodic rather than a constant, use the arithmetic mean over the period instead of the mean value.) Suppose that the input modulation has the form $A_0(t) = A_1(t/q_\varepsilon^2)$ (bandwidth $O(1/q_\varepsilon^2)$). Then, loosely speaking, if $A_1(\cdot)$ is sufficiently strongly mixing*, the limit results are the same as for the constant A_0 case, but where $\bar{A}_0$ replaces A_0 in the limit formulas. The calculations required for the proof use a combination of the ideas of this section and of Section VI, where we consider the effect of variations in $A_0(\cdot)$ on the errors in carrier frequency tracking, when a "squaring loop" is used.

The main result is the following. As $\varepsilon \to 0$, $\{v^\varepsilon(\cdot),\theta^\varepsilon(\cdot)\}$ converges weakly to the diffusion process $v(\cdot),\theta(\cdot)$ given by

---

* I.e., the conditional distributions of $A_1(t)$, $A_1(t+s)$ given $\{A_1(u)u \le 0\}$ converge fast enough to the unconditional distributions as t, $t+s \to \infty$.





(a) Limiter Approximation (b) Limiter

FIGURE 1.

Phase Locked Loop

$$dv = [Dv + \frac{A_0}{\sigma} H\sqrt{\frac{2}{\pi}} \sin(\theta-\hat{\theta})]dt + H\,\sigma_0 dB \tag{3.3}$$

$$d\hat{\theta} = Cvdt, \quad \hat{\theta} \text{ given},$$

where $B(\cdot)$ is a standard Wiener process and

$$\sigma_0^2 = 4\int_0^\infty [P\{z_1(u) > 0,\ z_1(0) > 0\} - P\{z_1(u) < 0,\ z_1(0) > 0\}]\,du$$
$$= \frac{4}{\pi}\int_0^\infty \sin^{-1}\rho(u)du \tag{3.4}$$

where $\rho(\cdot)$ is the normalized (such that $\rho(0) = 1$) correlation function of $z_1(\cdot)$. If $\rho(t) = \exp -a|t|$, $a > 0$, then the integral can be evaluated and $\hat{\sigma}_0^2 = 2 \ln 2/a$ [4], Section 6.

For the system without a limiter (and $L/q_\varepsilon = L_\varepsilon$ replaced by a unity gain), the limiting process is defined by

$$dv = [Dv + H\frac{A_0}{2}\sin(\theta-\hat{\theta})]dt + \sigma_1 HdB,$$
$$d\hat{\theta} = Cvdt, \tag{3.5}$$

where $\sigma_1^2 = \frac{\sigma^2}{2}\int_0^\infty \rho(u)du.$

Note the "$1/\sigma$" effect in (3.3). For small σ, the system with the limiter is preferable to the system without the limiter. The result (3.5) remains true for non-Gaussian noise. The $1/\sigma$ effect has been demonstrated by simulations on systems similar to those of this section. These simulations suggest that the limit results are often 'worst case', in that, for small $\varepsilon > 0$, the actual system often performs better than indicated by the limit results. In particular, if the limit results indicate that the limiter improves the operation, then the performance might be even better with the actual system, if the effective value of ε is small. An equation of the form of (3.3) can also be obtained for non-Gaussian process, under suitable conditions on the $z_i(\cdot)$. Then σ_0^2 will be given by the top line of (3.4), but the $\sqrt{\frac{2}{\pi}}$ in (3.3) will be replaced by a different constant.

<u>Comment on the choice of $\omega_0, \varepsilon, q_\varepsilon, a, \sigma$ in a particular practical problem.</u> For the limit results (3.3) with $\rho(t) = \exp - a|t|$, both σ and a are needed. Even without knowledge of these values, (3.3) gives the primary qualitative properties. Also, since the gain $L_\varepsilon = 1/q_\varepsilon$ was used, it was implicitly assumed that q_ε was known. This is not necessary. We can estimate a/q_ε^2, σ^2/q_ε^2 and $\sigma^2/2a$ from the data (from the normalized correlation function, the variance, and the power density at center frequency). Let $L_\varepsilon = (\sqrt{\sigma}/q_\varepsilon)(\sqrt{a}/\sigma) = \sqrt{a}/q_\varepsilon$, a quantity which can be estimated. Then the $1/\sigma$ and ω_0 of (3.3) are replaced by $\sqrt{a}/\sigma$ and $\sqrt{2}\ \ln 2$ resp. Thus, prior knowledge of q_ε σ or a is not needed. This is the case for all the problems which we have examined. For the case where the $z_i(\cdot)$ are Gaussian but with correlation function $\rho(\cdot)$ going to zero exponentially, the $2\ln 2/a$ is replaced by $(4/\pi)\int_0^\infty \sin^{-1}\rho(t)dt$ [4, Section 6]. The spectral density of the noise (near the center frequency) after the limiter is approximately $q_\varepsilon^2(4/\pi)\int_0^\infty \sin^{-1}\rho(t)dt$, a quantity which can be estimated. If we let L_ε be proportional to the inverse of the square root of this quantity, the "$1/\sigma$ effect" noted in (3.3) is maintained.

III.1. The Smooth Limiter $g_\alpha(\cdot)$

Now, we restrict attention to use of the smooth limiter $g_\alpha(\cdot)$ and get the limits as $\varepsilon \to 0$ and $\alpha \to 0$. When $\varepsilon \to 0$, $\alpha \to 0$ is stated, we mean that both α and $\varepsilon \to 0$, but in such a way that $q_\varepsilon/\alpha \leq Kq_\varepsilon^{\delta}$ for some $\delta > 0$. This condition can be weakened. Dropping the $2\omega_0^\varepsilon t$ components of (3.2), the input to the filter is $u_\alpha^\varepsilon(t,\hat{\theta}^\varepsilon(t),\theta(t))$, where

$$u_\alpha^\varepsilon(t,\hat{\theta}^\varepsilon,\theta) = \frac{1}{q_\varepsilon} g_\alpha\left[\frac{\sigma}{2q_\varepsilon}\left(z_1^\varepsilon(t)\cos\hat{\theta}^\varepsilon - z_2^\varepsilon(t)\sin\hat{\theta}^\varepsilon\right) + \frac{A_0}{2}\sin(\theta - \hat{\theta}^\varepsilon)\right]$$

and

$$\begin{aligned} \dot{v}^\varepsilon &= Dv^\varepsilon + HEu_\alpha^\varepsilon + H(u_\alpha^\varepsilon - Eu_\alpha^\varepsilon) \\ \dot{\hat{\theta}}^\varepsilon &= Cv^\varepsilon, \end{aligned}$$

<u>where the expectation</u> E <u>is over the</u> $z_i^\varepsilon(t)$ <u>only and</u>

$$Eu_\alpha^\varepsilon(t,\hat{\theta}^\varepsilon,\theta) = \sqrt{\frac{2}{\pi}}\,\frac{A_0}{\sigma}\sin(\theta - \hat{\theta}^\varepsilon) + O(q_\varepsilon\alpha)/q_\varepsilon .$$

Now, Theorem 1 will be applied. Given the test function $f(\cdot) \in \hat{\mathcal{D}}$, we must find a sequence of perturbed test functions $\{f^\varepsilon(\cdot)\}$ satisfying the conditions of Theorem 1. (Recall, we drop N and b_N in our calculations, for notational convenience.) The averaging method and the technique of proof is the same as that used in [4] and [6] and very close to that in [8]. Where the details overlap those in [4] or [6] only a sketch will be given. For given $f(\cdot)$, we use f^ε of the form noted below Theorem 1.

For $f(\cdot) \in \mathcal{D}$ (write $x^\varepsilon = x^\varepsilon(t)$), we start by applying $\hat{A}^\varepsilon$ to $f(x^\varepsilon(\cdot),\cdot)$. In this case the $\hat{A}^\varepsilon$ operation is merely a right derivative

$$\hat{A}^\varepsilon f(x^\varepsilon,t) = f_t(x^\varepsilon,t) + f_\theta'(x^\varepsilon,t)Cv^\varepsilon \tag{3.6}$$
$$+ f_v'(x^\varepsilon,t)[Dv^\varepsilon + H(Eu_\alpha^\varepsilon(t,\hat{\theta}^\varepsilon,\theta(t)) + H(u_\alpha^\varepsilon(t,\hat{\theta}^\varepsilon,\theta(t))$$
$$- E(u_\alpha^\varepsilon(t,\hat{\theta}^\varepsilon,\theta(t)))].$$

Only the "noise term", $f_v'(x^\varepsilon,t)H(u_\alpha^\varepsilon - Eu_\alpha^\varepsilon)$, of (3.6) needs to be averaged out. The other terms are part of or close to components of $(\frac{\partial}{\partial t} + A)f(x^\varepsilon,t)$, where A is the operator of (3.5). Define the first perturbation $f_1^\varepsilon(t) = f_1^\varepsilon(x^\varepsilon(t),t)$, where

$$f_1^\varepsilon(x^\varepsilon,t) = \int_0^\infty ds E_t^\varepsilon f_v'(x^\varepsilon,t+s)H[u^\varepsilon(t+s,\hat{\theta}^\varepsilon(t),\theta(t)) - \tag{3.7}$$
$$- Eu_\alpha^\varepsilon(t+s,\hat{\theta}^\varepsilon(t),\theta(t))]].$$

<u>Note that the integrand at</u> $s = 0$ <u>is just the "noise" term of</u> (3.6). <u>In all expressions of the type</u> $Eu_\alpha^\varepsilon(t+s,\hat{\theta}^\varepsilon(t),\theta(t))$ (<u>or with</u> E_t^ε <u>replacing</u> E), <u>the expectation is over the</u> $\xi_1^\varepsilon(t+s)$ <u>only, not over</u> $\hat{\theta}^\varepsilon(t)$ <u>or</u> $x^\varepsilon(t)$. Via the change of variables $s/q_\varepsilon^2 \to s$, (3.7) can be written as

$$q_\varepsilon^2 \int_0^\infty ds E_t^\varepsilon f_v'(x^\varepsilon, t+q_\varepsilon^2 s) H[u_\alpha^\varepsilon(t+q_\varepsilon^2 s, \hat{\vartheta}^\varepsilon(t), \vartheta(t)) -$$

$$\bar{u}_\alpha^\varepsilon(t+q_\varepsilon^2 s, \hat{\vartheta}^\varepsilon(t), \vartheta(t))]$$

which is bounded in absolute value by

$$O(q_\varepsilon)[1 + |\xi(t/q_\varepsilon^2)|] \tag{3.8}$$

(See [10] for a related calculation; here we use a different noise model than in [10], and a multiplier rather than an adder. These require somewhat different details and yield results which are not directly deducible from the results of [10].) It can be checked that $f_1^\varepsilon(\cdot) \in \mathscr{D}(\hat{A}^\varepsilon)$ and that (write $x^\varepsilon(t) = x$; the term denoted simply by the bracket { } is the integrand in (3.7)), and the subscripts $\hat{\vartheta}$, ϑ and v still denote the partial derivatives or gradient)

$$\hat{A}^\varepsilon f_1^\varepsilon(t) = -f_v'(x^\varepsilon, t) H(u_\alpha^\varepsilon(t, \hat{\vartheta}^\varepsilon, \vartheta(t)) - \bar{u}_\alpha^\varepsilon(t, \hat{\vartheta}^\varepsilon, \vartheta(t)))$$

$$+ \int_0^\infty ds E_t^\varepsilon \{ \quad \}_{\hat{\vartheta}} \dot{\hat{\vartheta}}^\varepsilon(t) + \int_0^\infty ds E_t^\varepsilon \{ \quad \}_\vartheta \dot{\vartheta}(t) \tag{3.9}$$

$$+ \int_0^\infty ds E_t^\varepsilon \{ \quad \}_v' \dot{v}^\varepsilon(t).$$

The first term of (3.9) is the negative of the "noise term" of (3.6).

Now, examine the second term of (3.9), which we write in greater detail as (see below (3.10) for the definitions of the new terms

$$\frac{1}{q_\varepsilon}\int_0^\infty ds f'_{\hat{v}}(x^\varepsilon,t+s)H(E_t^\varepsilon g_\alpha[t+s] - Eg_\alpha[t+s])Cv^\varepsilon(t)$$

$$+\frac{1}{q_\varepsilon}\int_0^\infty ds f'_v(x^\varepsilon,t+s)H(E_t^\varepsilon \dot{g}_\alpha[t+s]\gamma^\varepsilon(t,t+s) - E\dot{g}_\alpha[t+s]\gamma^\varepsilon(t,t+s))Cv_t^\varepsilon \tag{3.10}$$

where $\dot{g}_\alpha(u) = \frac{\partial g_\alpha}{\partial u}(u)$ and $\gamma^\varepsilon(t,t+s) = \{-\frac{\sigma}{2q_\varepsilon}(Y_1^\varepsilon(\hat{\vartheta}^\varepsilon(t), z^\varepsilon(t+s))$ $-\frac{A_0}{2}\cos(\vartheta(t) - \hat{\vartheta}^\varepsilon(t))\}$. In (3.10) we used the definition

$$[t+s] \equiv \frac{\sigma}{2q_\varepsilon} Y_2^\varepsilon(\hat{\vartheta}^\varepsilon(t), z^\varepsilon(t+s)) + \frac{A_0}{2}\sin(\vartheta(t) - \hat{\vartheta}^\varepsilon(t))$$

for the arguement of $g_\alpha(\cdot)$, and the $Y_i^\varepsilon(\cdot)$ are defined by

$$Y_1^\varepsilon(\hat{\vartheta}^\varepsilon, z^\varepsilon(t+s)) = [z_1^\varepsilon(t+s)\sin\hat{\vartheta}^\varepsilon + z_2^\varepsilon(t+s)\cos\hat{\vartheta}^\varepsilon]$$

$$Y_2^\varepsilon(\hat{\vartheta}^\varepsilon, z^\varepsilon(t+s)) = [z_1^\varepsilon(t+s)\cos\hat{\vartheta}^\varepsilon - z_2^\varepsilon(t+s)\sin\hat{\vartheta}^\varepsilon].$$

The second term of (3.10) occurs since $g_\alpha[t+s]$ depends on $\hat{\vartheta}^\varepsilon(t)$. Now note the important fact that for each fixed t, the processes $Y_1^\varepsilon(\hat{\vartheta}(t), z^\varepsilon(t+\cdot))$ and $Y_2^\varepsilon(\hat{\vartheta}(t), z^\varepsilon(t+\cdot))$ are independent. This property, which will be used frequently, is due to the Gaussian assumption on the $z_i(\cdot)$. Without it the $\sigma_0 dB$ in (3.3) would be replaced by a more complicated expression.

The first term in (3.10) can be shown to satisfy (3.8), and so does the component of the second term which is linear in A_0 if (3.8) is divided by α. By the independence cited in the last paragraph, the expectation term, $E\dot{g}_\alpha[t+s]Y_1^\varepsilon(\hat{\vartheta}^\varepsilon(t), z^\varepsilon(t+s))$, in the second component of (3.10) is zero, and the conditional expectation there equals

$$-\frac{\varepsilon}{2q_\varepsilon} E_t^\varepsilon \dot{g}_\alpha[t+s] E_t^\varepsilon Y_1^\varepsilon(\hat{\theta}^\varepsilon(t), z^\varepsilon(t+s)) =$$

$$-\frac{\varepsilon}{2q_\varepsilon} E_t^\varepsilon \dot{g}_\alpha[t+s] \rho(s/q_\varepsilon^2) Y_1^\varepsilon(\hat{\theta}^\varepsilon(t), z^\varepsilon(t)).$$

From this, some manipulations using the change of variable $s/q_\varepsilon^2 \to s$ and the facts that $|\dot{g}_\alpha(u)| \le K/\alpha$ and that $\dot{g}_\alpha(u) = 0$ for $|u| > \alpha$ yield that the remaining component of (3.10) is bounded by (3.8). The third term of (3.9) can be treated in the same way as the second was treated and with the same result.

Thus $\underset{\substack{\varepsilon \to 0 \\ \alpha \to 0}}{\text{p-lim}} [2^{nd} + 3^{rd} \text{ term of } (3.9)] = 0.$

Only the last term of (3.9) remains. The part of the term which contains the $(Dv^\varepsilon + HEu_\alpha^\varepsilon)$ component of $\dot{v}^\varepsilon$ is also bounded by (3.8). Thus, it is negligible and only the remaining component of (3.9), which contains the $H(u_\alpha^\varepsilon - Eu_\alpha^\varepsilon)$ part of $\dot{v}^\varepsilon$ must be averaged further. We denote this component by k^ε and write it in the form (3.12).

$$\begin{aligned} & k^\varepsilon(t, \hat{\theta}^\varepsilon(t)) \\ & \quad = \frac{1}{q_\varepsilon^2} \int_0^\infty ds\, H' f_{vv}(x^\varepsilon(t), t+s) H \cdot \\ & \quad \{E_t^\varepsilon g_\alpha[t+s] - Eg_\alpha[t+s]\}\{g_\alpha[t] - Eg_\alpha[t]\}, \end{aligned} \tag{3.12}$$

where $[t]$, $[t+s]$ are defined below (3.10)(the expectation is only over the $g_\alpha[t+s]$ term). This term is not negligable and must be averaged further. Define the second perturbation $f_2^\epsilon(t) = f_2^\epsilon(x^\epsilon(t),t)$, where

$$f_2^\epsilon(x^\epsilon,t) = \int_0^\alpha d\tau\,[E_t^\epsilon k^\epsilon(t+\tau,\hat{\theta}^\epsilon) - Ek^\epsilon(t+\tau,\hat{\theta}^\epsilon)]. \tag{3.13}$$

It can be shown that $|f_2^\epsilon(t)| = O(q_\epsilon^2)[1+|x^\epsilon(t)|^2]$, that $f_2^\epsilon(\cdot) \in \mathcal{D}(\hat{A}^\epsilon)$ and that

$$\hat{A}^\epsilon f_2^\epsilon(t) = -(3.12) + Ek^\epsilon(t,\hat{\theta}^\epsilon(t)) + \tag{3.14}$$

$$+ \text{(terms whose p-lim is zero)}.$$

The last term on the right side is $[1 + |x^\epsilon(t)|^3]O(q_\epsilon/\alpha)$.

The middle term of (3.14) minus $H'f_{vv}(x^\epsilon(t),t)H\sigma_0^2/2$ tends to zero in the p-lim sense as $\epsilon \to 0$, $\alpha \to 0$. This last result follows by the same sort of argument as used in [4], Section 6.

Summarizing the above calculations and using the $f^\epsilon(\cdot) = f(x^\epsilon(\cdot),\cdot) + f_1^\epsilon(\cdot) + f_2^\epsilon(\cdot)$ yields that the two required limits of Theorem1 hold, where A is the operator of (3.3); i.e.,

$$\underset{\substack{\alpha\to 0\\ \epsilon\to 0}}{\text{p-lim}}\ [\hat{A}^\epsilon f^\epsilon(\cdot) - (\tfrac{\partial}{\partial t} + A)f(x^\epsilon(\cdot),\cdot)] = 0$$

$$\underset{\substack{\epsilon\to 0\\ \alpha\to 0}}{\text{p-lim}}\ [f_i^\epsilon(\cdot)] = 0,\quad i = 1,2.$$

By [6], Theorem 2, tightness of $\{x^\epsilon(\cdot)\}$ follows from the given bounds on $f_i^\epsilon(\cdot)$ and on $\hat{A}^\epsilon f^\epsilon(\cdot)$. Thus, the $\{x^\epsilon(\cdot)\}$ converge weakly to the solution of (3.3) as $\epsilon \to 0$, $\alpha \to 0$.

III.2. The System with the Limiter of Figure 1b

Now, the $g_a(\cdot)$ of III.1 is replaced by $\text{sign}(\cdot) = g(\cdot)$ and u_a^ε is replaced by u^ε where

$$u^\varepsilon(t,\hat{\theta}^\varepsilon,\theta) = \frac{1}{q_\varepsilon}\, \text{sign}\Big[\frac{\sigma}{2q_\varepsilon}\, (z_1^\varepsilon(t)\cos\hat{\theta}^\varepsilon - z_2^\varepsilon(t)\sin\hat{\theta}^\varepsilon) + \frac{A_0}{2}\sin(\theta-\hat{\theta}^\varepsilon)\Big]$$

and (3.6) holds (u^ε used). Also, $Eu^\varepsilon(t,\hat{\theta}^\varepsilon,\theta) = \sqrt{\frac{2}{\pi}}\, \frac{A_0}{\sigma}\sin(\theta-\hat{\theta}^\varepsilon) + O(q_\varepsilon)$. The function $f_1^\varepsilon(\cdot)$ is still defined by (3.7) (u used) and (3.8) still holds. Note that the distribution of (write $\hat{\theta}^\varepsilon = \hat{\theta}^\varepsilon(t), \theta = \theta(t)$) $[z_1^\varepsilon(t+s)\cos\hat{\theta}^\varepsilon - z_2^\varepsilon(t+s)\sin\hat{\theta}^\varepsilon + \frac{q_\varepsilon A_0}{\sigma}\sin(\theta-\hat{\theta}^\varepsilon)]$ <u>conditioned on data up to time</u> t is $N(\text{mean}^\varepsilon(s), \text{var}^\varepsilon(s))$, where $N(\alpha,\beta)$ is the normal distribution with mean α and variance β and

$$\text{mean}^\varepsilon(s) = \rho(s/q_\varepsilon^2)(z_1^\varepsilon(t)\cos\hat{\theta}^\varepsilon - z_2^\varepsilon(t)\sin\hat{\theta}^\varepsilon) + \frac{A_0 q_\varepsilon}{\sigma}\sin(\theta-\hat{\theta}^\varepsilon)$$
$$\text{var}^\varepsilon(s) = 1 - \rho^2(s/q_\varepsilon^2).$$

It is convenient to write $f_1^\varepsilon(\cdot)$ in the more explicit form (3.15) (with obvious notation, $f_1^\varepsilon(t) = f_1^\varepsilon(t,\theta^\varepsilon(t),\hat{\theta}^\varepsilon(t))$ where $f_1^\varepsilon(t,\hat{\theta}^\varepsilon,\theta)$ is defined by

$$f_1^\varepsilon(t,\hat{\theta}^\varepsilon,\theta) = \frac{1}{q_\varepsilon}\int_0^\infty f_v'(x^\varepsilon,t+s)H\Big\{\int_{-\infty}^\infty (\text{sign}\,\xi)\,dN(\text{mean}^\varepsilon(s), \text{var}^\varepsilon(s)) \tag{3.15}$$
$$- \int_{-\infty}^\infty (\text{sign}\,\xi)\,dN\Big(\frac{A_0 q_\varepsilon}{\sigma}\sin(\theta-\hat{\theta}^\varepsilon), 1\Big)\Big\}ds\ .$$

It can be shown that $f_1^\varepsilon(\cdot) \in \mathcal{D}(\hat{A}^\varepsilon)$ and that $\hat{A}^\varepsilon f_1^\varepsilon(\cdot)$ is given by (3.15).

$$\hat{A}^\varepsilon f_1^\varepsilon(t) = -f_v'(x^\varepsilon(t),t)H(u^\varepsilon(t,\vartheta^\varepsilon(t),\vartheta(t)) - Eu^\varepsilon(t,\vartheta^\varepsilon(t),\vartheta(t)))$$

$$+ \frac{1}{q_\varepsilon}\int_0^\infty (E_t^\varepsilon\{\quad\})_{\vartheta} ds\, \dot{\vartheta}^\varepsilon(t) + \frac{1}{q_\varepsilon}\int_0^\infty (E_t^\varepsilon\{\quad\})_{\vartheta} ds\, \dot{\vartheta}(t) \qquad (3.16)$$

$$+ \frac{1}{q_\varepsilon}\int_0^\infty (E_t^\varepsilon\{\quad\})_v' ds\, \dot{v}^\varepsilon(t),$$

where $\{\quad\}$ is the integrand in (3.7) with u_α^ε replaced by u^ε. When calculating the derivatives in (3.16), we use the explicit representation (3.15), so only derivatives with respect to parameters in the normal density function are taken. No derivatives of the sign function are taken. Proceeding as in the preceding subsection, the first two integrals on the right side of (3.16) are bounded by (3.8). Modulo a term which satisfies (3.8), the last term of (3.16) is

$$\frac{1}{q_\varepsilon^2}\int_0^\infty ds\; H'f_{vv}(x^\varepsilon(t),t+s)H(E_t^\varepsilon \text{sign}[t+s] -$$

$$- E\,\text{sign}[t+s])(\text{sign}[t] - E\,\text{sign}[t]),$$

where $[t]$, $[t+s]$ are defined below (3.10).

The rest of the development is exactly as in Section III.1, except that, when taking derivatives in the calculation of $\hat{A}^\varepsilon f_2^\varepsilon(\cdot)$, the explicit form of the expectation and conditional expectation are used, analogous to what was done with $f_1^\varepsilon(\cdot)$, and the bound below (3.14) is replaced by

$$O(q_\varepsilon)[1 + |z^\varepsilon(t)|^3) \tag{3.17}$$

The limit is still (3.3).

IV. The Quadricorrelator Without A Limiter

The system and some of the notation is given by Figure 2. As in Section III, the equations defining the linear filters are merely state variable representations of low pass filters. The noise model $n^\epsilon(\cdot)$ of Section III is used here and in the next section. Let the initial frequency error $\omega_0^\epsilon - \omega_1^\epsilon = \Delta\omega$ not depend on ϵ, and set $\omega_1^\epsilon = \omega_1/\epsilon^2$, $\omega_0^\epsilon = \omega_0/\epsilon^2$ and $K_0 = 1$ (absorb it into A_0, σ). For the upper filter, let $C_1 H_1 = 0$. This is necessary to guarantee that the limit of the input as $\epsilon \to 0$ to the differentiator is differentiable. This restriction is normally satisfied in practical systems. Now the system state is $x^\epsilon = (v_1^\epsilon, v_2^\epsilon, \hat\theta^\epsilon)$. As in Section II, the noise center frequency need not be ω_0^ϵ, provided that (noise center frequency minus $\omega_1^\epsilon)\cdot q_\epsilon^2 \to 0$ as $\epsilon \to 0$. The purpose of the system [3] is the estimation of and tracking of the (possibly time varying) input frequency. Unlike the phase locked loop, the input to the VCO depends on a frequency as well as a phase error. See the heuristic comments in [3], which are repeated below (4.1). We use the assumptions on A_0 and θ of the previous section.

The analysis is very similar to that of Section III, and only a few details will be given. The sequence $\{x^\epsilon(\cdot)\}$ converges weakly to $x(\cdot) = (v_1(\cdot), v_2(\cdot), \hat\theta(\cdot))$, where

$$
\begin{aligned}
dv_1 &= D_1 v_1 dt + \frac{A_0}{2} H_1 \sin(\Delta\omega t + \theta - \hat\theta)dt + \frac{\sigma_2}{2} H_1 dB_1 \\
dv_2 &= D_2 v_2 dt + \frac{A_0}{2} H_2 \cos(\Delta\omega t + \theta - \hat\theta)dt + \frac{\sigma_2}{2} H_2 dB_2 \qquad (4.1) \\
d\hat\theta &= v_1' D_1' C_1' C_2 v_2 dt, \qquad \sigma_2^2 = \sigma^2 \int^{\infty} \rho(s)ds,
\end{aligned}
$$

where the $B_i(\cdot)$ are independent Wiener processes. The independence of the $B_i(\cdot)$ is a consequence of the orthogonality of the $z_i(\cdot)$. It does not depend on the Gaussian assumption on the $z_i(\cdot)$ (as a similar independence assertion does for the case of the next section).

Remark. Assume, purely formally, that the filters are such that for some positive α_i their outputs are roughly $v_1^\epsilon(t) = \alpha_0 \sin(\Delta\omega t + \theta - \hat\theta^\epsilon) + m_1^\epsilon(t)$, $m_1^\epsilon(\cdot)$ being some differentiable noise, and $v_2^\epsilon(t) = \alpha_1 \cos(\Delta\omega t + \theta + \hat\theta^\epsilon) + m_2^\epsilon(t)$, $m_2^\epsilon(\cdot)$ being some noise process. Then the VCO input is $= \alpha_0\alpha_1(\Delta\omega + (\dot\theta - \dot{\hat\theta}^\epsilon))\cos^2(\Delta\omega t + \theta - \hat\theta^\epsilon)$ + noise terms. Thus the VCO action depends on the frequency offset $\Delta\omega$ as well as on the phase estimation error.

Continuing with the development, define

$$
\begin{aligned}
u_1^\epsilon(t,t+s,\hat\theta^\epsilon,\theta) &= \frac{A_0}{2}\sin(\Delta\omega t+\theta-\hat\theta^\epsilon) + \frac{\sigma}{2q_\epsilon}\, u_1^{\epsilon,0}(t,t+s,\hat\theta^\epsilon) \\
u_2^\epsilon(t,t+s,\hat\theta^\epsilon,\theta) &= \frac{A_0}{2}\cos(\Delta\omega t+\theta-\hat\theta^\epsilon) + \frac{\sigma}{2q_\epsilon}\, u_2^{\epsilon,0}(t,t+s,\hat\theta^\epsilon),
\end{aligned}
\tag{4.2}
$$

where we define

$$
\begin{aligned}
u_1^{\epsilon,0}(t,t+s,\hat\theta^\epsilon) &= z_1^\epsilon(t+s)\cos(\Delta\omega t-\hat\theta^\epsilon) + z_2^\epsilon(t+s)\sin(\Delta\omega t-\hat\theta^\epsilon) \\
u_2^{\epsilon,0}(t,t+s,\hat\theta^\epsilon) &= -z_1^\epsilon(t+s)\sin(\Delta\omega t-\hat\theta^\epsilon) + z_2^\epsilon(t+s)\cos(\Delta\omega t-\hat\theta^\epsilon).
\end{aligned}
$$

Proceeding as in Section III, following the argument below (3.2), and dropping the terms dropped there, the inputs to the linear filters at time t are $u_i^\epsilon(t,t,\hat\theta^\epsilon(t),\theta(t))$, $i = 1,2$. Then

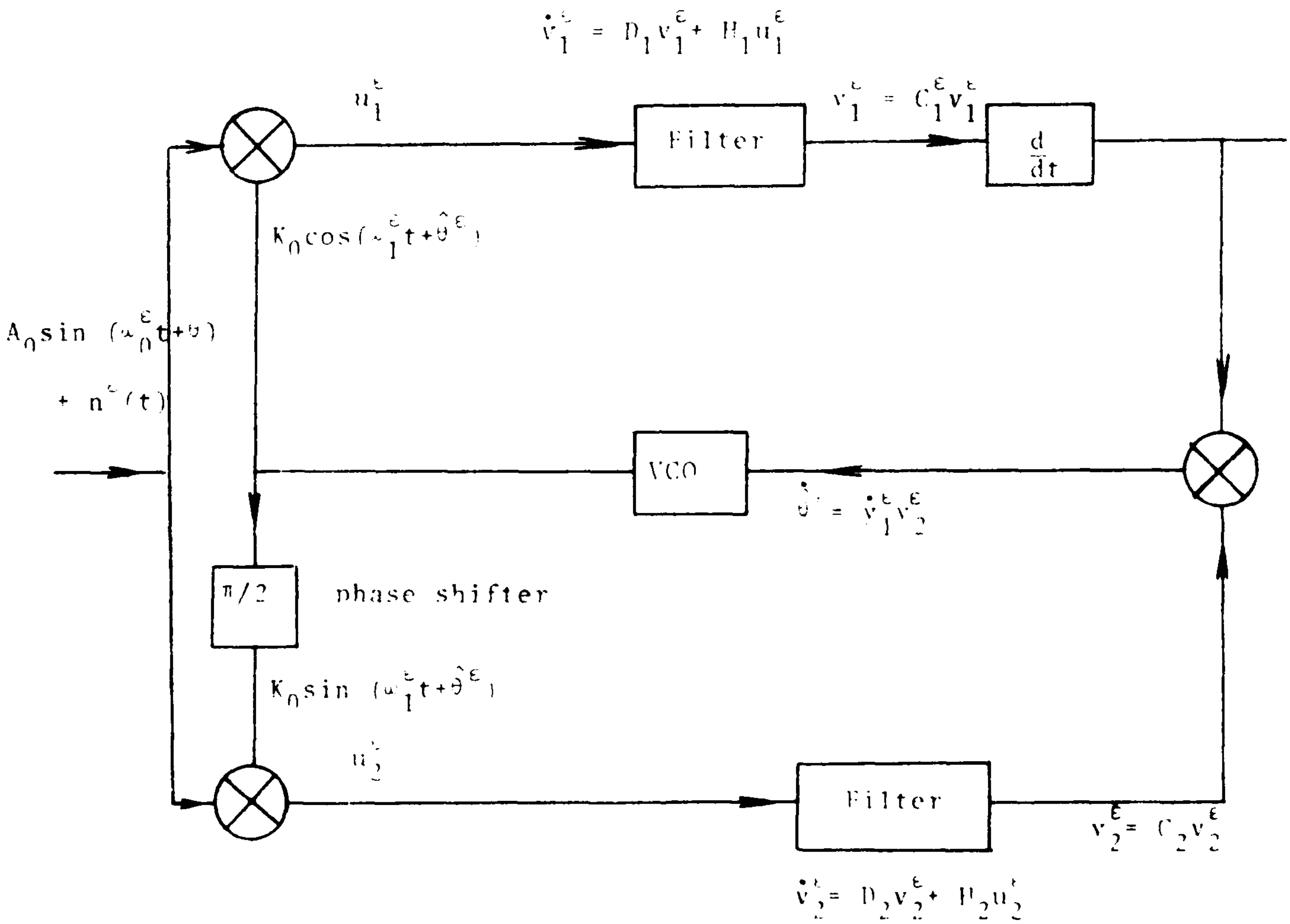


FIGURE 2

The Quadricorrelator Without Limiters

$$\dot{x}^\varepsilon = \begin{Bmatrix} D_1 v_1^\varepsilon + H_1 u_1^\varepsilon(t,t,\hat{\psi}^\varepsilon(t),\psi(t)) \\ D_2 v_2^\varepsilon + H_2 u_2^\varepsilon(t,t,\psi^\varepsilon(t),\psi(t)) \\ v_1^{\varepsilon\prime} D_1' C_1' C_2 v_2^\varepsilon \end{Bmatrix}.$$

For $f(\cdot) \in \mathcal{D}$, (write $x^\varepsilon = x^\varepsilon(t)$)

$$\hat{A}^\varepsilon f(x^\varepsilon,t) = f_t(x^\varepsilon,t) + f'_{v_1}(x^\varepsilon,t)\dot{v}_1^\varepsilon + f'_{v_2}(x^\varepsilon,t)\dot{v}_2^\varepsilon + \ldots$$

Only the components

$$\frac{\sigma}{2q_\varepsilon}[f'_{v_1}(x^\varepsilon,t)H_1 u_1^{\varepsilon,0}(t,t,\psi^\varepsilon) + f'_{v_2}(x^\varepsilon,t)H_2 u_2^{\varepsilon,0}(t,t,\ldots)]$$

of (4.3) need to be averaged out; the other components of (4.3) are part of $(\frac{\partial}{\partial t} + A)f(x^\varepsilon,t)$, where A is the operator of the process defined by (4.1).

The function $f_1^\varepsilon(t) = f_1^\varepsilon(x^\varepsilon(t),t)$ is defined in the usual way; namely

$$f_1^\varepsilon(x^\varepsilon,t) = \frac{\sigma}{2q_\varepsilon}\int_0^\infty ds E_t^\varepsilon\{f'_{v_1}(x^\varepsilon,t+s)H_1 u_1^{\varepsilon,0}(t,t+s,\psi^\varepsilon)$$

$$+ f'_{v_2}(x^\varepsilon,t+s)H_2 u_2^{\varepsilon,0}(t,t+s,\psi^\varepsilon)\}.$$

The bound (5.8) on $f_1^\varepsilon(\cdot)$ holds, $f_1^\varepsilon(\cdot) \in \mathcal{D}(\hat{A}^\varepsilon)$ and

$$\hat{A}^\varepsilon f_1^\varepsilon(x^\varepsilon,t) = -(4.3) + k^\varepsilon(x^\varepsilon,t),$$

where $k^\varepsilon(x^\varepsilon,t)$ is now defined to be $(f^\varepsilon_{1,x}(x^\varepsilon,\cdot))'\dot{x}^\varepsilon$ and equals

$$k^\varepsilon(x^\varepsilon,t) = \frac{\sigma^2}{4q_\varepsilon^2}\int_0^\infty \Big\{ H_1' f_{v_1v_1}(x^\varepsilon,t+s)H_1 E_t^\varepsilon u_1^{\varepsilon,0}(t,t+s,\ldots)u_1^{\varepsilon,0}(t,t,\ldots)$$

$$+ H_1' f_{v_1v_2}(x^\varepsilon,t+s)H_2 E_t^\varepsilon u_1^{\varepsilon,0}(t,t+s,\ldots)u_2^{\varepsilon,0}(t,t,\ldots)$$

$$+ H_2' f_{v_2v_1}(x^\varepsilon,t+s)H_1 E_t^\varepsilon u_2^{\varepsilon,0}(t,t+s,\ldots)u_1^{\varepsilon,0}(t,t,\ldots)$$

$$+ H_2' f_{v_2v_2}(x^\varepsilon,t+s)H_2 E_t^\varepsilon u_2^{\varepsilon,0}(t,t+s,\ldots)u_2^{\varepsilon,0}(t,t,\ldots)\Big\} ds$$

+ terms bounded by (3.8).

For each t and $x^\varepsilon, u_1^{\varepsilon,0}(t,t+\cdot,\ldots)$ is independent of $u_2^{\varepsilon,0}(t,t+\cdot,\ldots)$ and the expectation of the second and third components of the integral (4.7) are zero. Actually only orthogonality and not independence is required here. Defining $f_2^\varepsilon(\cdot)$ by (3.15), we can show that $f_2^\varepsilon(\cdot)$ is bounded by (3.15) and $f_2^\varepsilon(\cdot) \in \mathcal{D}(\hat{A}^\varepsilon)$ and that (writing $x^\varepsilon = x^\varepsilon(t)$)

$$\hat{A}^\varepsilon f_2^\varepsilon(t) = -(k^\varepsilon(x^\varepsilon,t) - Ek^\varepsilon(x^\varepsilon,t))$$

$$+ \text{ terms bounded by } (3.17).$$

Changing variables $s/q_\varepsilon^2 \to s$ in (4.7) and evaluating $Ek^\varepsilon(x^\varepsilon,\cdot)$

$$\text{p-}\lim_{\varepsilon\to 0}[Ek^\varepsilon(x^\varepsilon(\cdot),\cdot) - \frac{\sigma^2}{4}(H_1' f_{v_1v_1}(x^\varepsilon(\cdot),\cdot)H_1 + H_2' f_{v_2v_2}\ldots)\ldots] = 0 \ldots$$

Combining (4.8) with (4.5) and (4.6) yields

[illegible]

where A is the operator of the process [illegible] by (3.1) [illegible] bounds obtained on the [illegible] and on the [illegible] via [6, Theorem 2]. Thus, [illegible] converges [illegible] of (3.1).

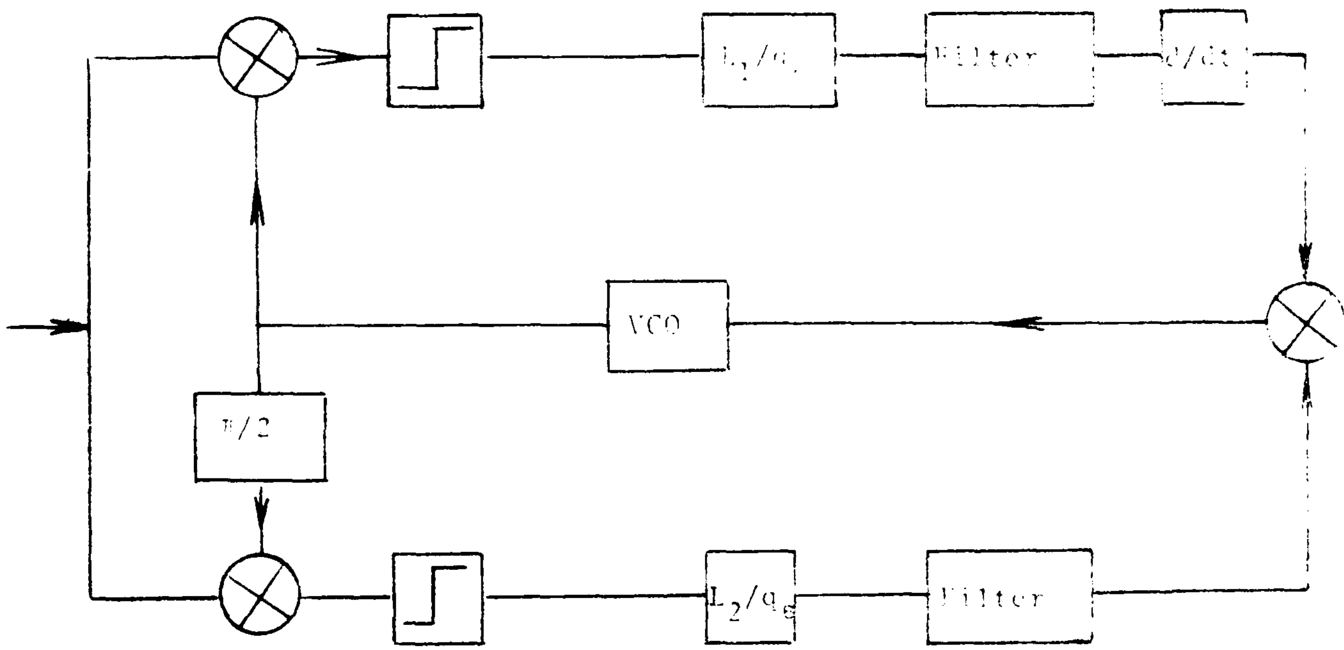


FIGURE 3

Quadricorrelator with a limiter in each arm.

V. Quadricorrelator with Limiter

The system, described by Figure 5, differs from the system of the previous section only in the inclusion of the limiter and normalizing gains L_i/q_ε. Those gains are required here, either before or after the low pass filters for the same reason that they were used in Section III. Following the method of the previous section, the inputs to the low pass filters are the signs of the u_i^ε of (4.2) times L_i/q_ε. Equation (4.5) holds, and its component

$\sum_{i=1}^{2} f'_{v_i}(x^\varepsilon,t)H_i \operatorname{sign}(u_i^\varepsilon(t,t,\theta^\varepsilon,\psi(t)))L_i/q_\varepsilon$ must be averaged out. Hence, as usual, we define the first perturbation $f_1^\varepsilon(t) = f_1^\varepsilon(x^\varepsilon(t),t)$, where

$$f_1^\varepsilon(x^\varepsilon,t) = \sum_{i=1}^{2} \frac{L_i}{q_\varepsilon}\int_0^\infty ds\, f'_{v_i}(x^\varepsilon,t+s)H_i[E_t^\varepsilon \operatorname{sign}(u_i^\varepsilon(t,t+s,\theta^\varepsilon,\psi(t))) - E \operatorname{sign}(u_i^\varepsilon(t,t+s,\theta^\varepsilon,\psi(t)))] . \tag{5.1}$$

Due to the non-differentiability of the sign function, an explicit integral representation of (5.1) of the form [illegible] used when calculating $\hat{A}^\varepsilon f_i^\varepsilon(\cdot)$. By a combination of the ideas of Sections III.2 and IV, we can show that $\{x^\varepsilon(\cdot)\}$ converges weakly to the diffusion $x(\cdot) = (v_1(\cdot), v_2(\cdot), \theta(\cdot))$:

$$dv_1 = (D_1v_1 + \frac{H_1\Lambda_0L_1}{\sigma}\sqrt{\frac{2}{\pi}}\sin(\Delta\omega t + \psi - \theta))dt + L_1H_1 [illegible]$$

$$dv_2 = (D_2v_2 + \frac{H_2\Lambda_0L_2}{\sigma}\sqrt{\frac{2}{\pi}}\cos(\Delta\omega t + \psi - \theta))dt + L_2H_2 [illegible]$$

$$d\theta = v_1'D_1'C_1'C_2v_2dt,$$

where $B_i(\cdot)$ are independent standard Brown motions and [illegible] is defined by (3.4). Note the "$1/\sigma$ effect" in (5.2). The $B_i(\cdot)$ are independent owing to the Gaussian assumption, because of which the $u_i^{\varepsilon,0}(t,t+\cdot,\hat{\theta}^{\varepsilon})$, $i = 1,2$, are independent for each $t,\hat{\theta}^{\varepsilon}$. This, in turn, is due to the $\pi/2$ phase shifter.

VI. The Squaring Loop

In this section, we do an asymptotic analysis of the system of Figure 4, whose purpose is to "track the carrier frequency" irrespective of the input modulation process $m_\varepsilon(\cdot)$. We will study the effects of the $m_\varepsilon(\cdot)$ on the estimation error process $(\theta^\varepsilon(\cdot) - \hat{\theta}^\varepsilon(\cdot))$ for small noise, and, eventually, do a linearization analysis. For notational convenience, scale such that $K_2 = 1$. In a sense, the procedure below is an attempt to rigorize the heuristic analysis of Gardner [3], Appendix B. The limit equations are given in (6.14)-(6.15).

For specificity, let the transmitted signal be pulse phase shift or amplitude modulated in the following way. We will scale the problem so that a meaningful result can be obtained for small pulse widths and small noise. Let $p(\cdot)$ denote a realizable transfer function, continuous for $t > 0$ and right continuous at 0, and set $p_\varepsilon(t) = p(t/q_\varepsilon^2)$, where $\varepsilon/q_\varepsilon \to 0$ as $\varepsilon \to 0$. Let $\{a_k\}$ be a sequence of bounded zero mean independent random variables with unit variance, and set $m_\varepsilon(t) = \sum_{k=-\infty}^{\infty} a_k p_\varepsilon(t-kT_\varepsilon)$, where $T_\varepsilon = Tq_\varepsilon^2$ denotes the width of the pulse interval, T being some given constant. Suppose that there is a bounded non-increasing function $\bar{p}(\cdot)$ such that $|p(t)| \le \bar{p}(t)$ and $\int_0^\infty dt \int_t^\infty du\, \bar{p}(u) < \infty$. The general technique of this section can be used with a greater variety of modulation types, and the independence of $\{a_k\}$ can be weakened. The input noise $\tilde{n}^\varepsilon(\cdot)$ has the form:

$$\tilde{n}^\varepsilon(t) = [z_1^\varepsilon(t)\cos \omega_0^\varepsilon t + z_2^\varepsilon(t)\sin \omega_0^\varepsilon t]\sigma.$$

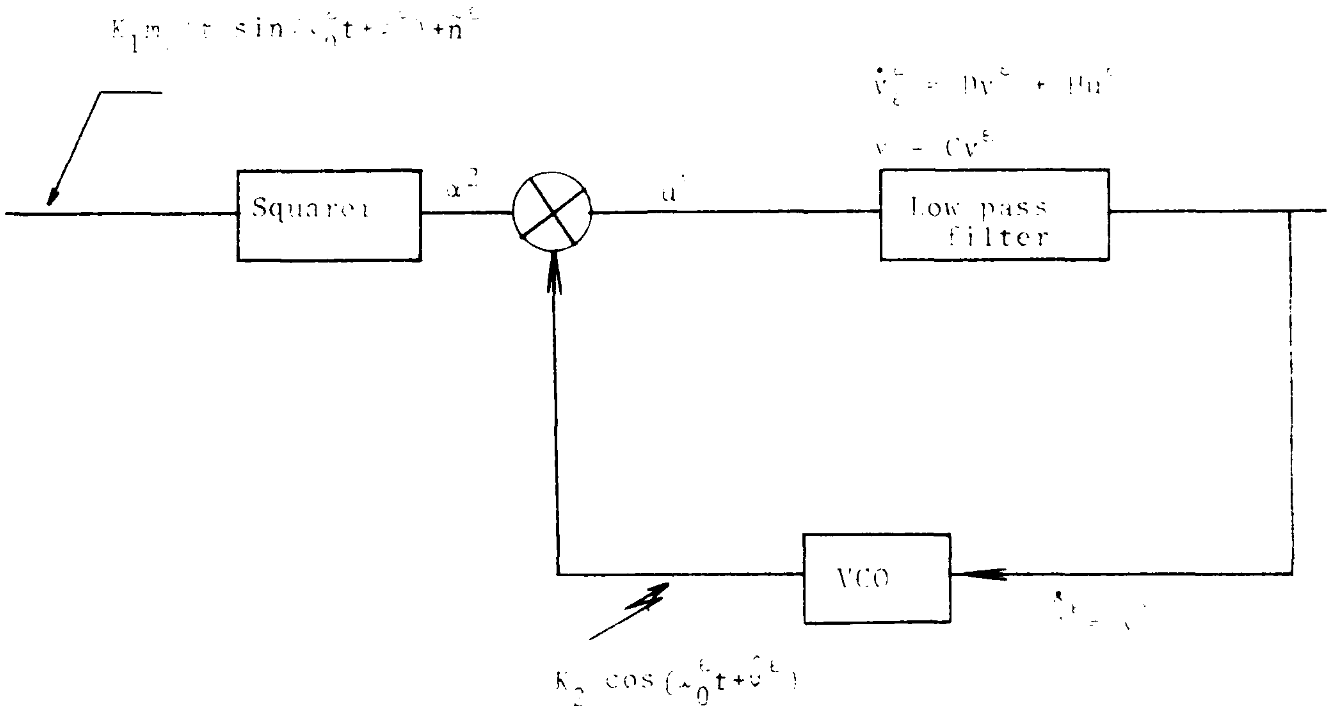


FIGURE 4

The Squaring Loop.

where $z_i^\varepsilon(\cdot)$ and $\tilde{z}_i^\varepsilon(\cdot)$ are defined in Section III.

In this section, the analysis is done for small noise, as essentially assumed in [3], Appendix B, so we do not divide $\tilde{n}^\varepsilon(\cdot)$ by q_ε. Suppose that $\dot{\theta}^\varepsilon = O(q_\varepsilon)$, and that the $z_i(\cdot)$ are independent of $\{a_k\}$.

The scaling is of the correct order for a meaningful problem. The signal and noise BW's are both $O(q_\varepsilon^{-2})$. The sequence $\{a_k\}$ is the transmitted signal sequence and the $p_\varepsilon(\cdot)$ yields the transmission channel **behavior** ("intersymbol interference"). The "memory" in the "symbol interference" system is roughly the same number of pulse intervals, irrespective of ε. Also, for small ε, the average noise energy per pulse interval is of the same order as the signal energy in that interval. With an inappropriate scaling, the problem would degenerate, as $\varepsilon \to 0$, to one in which detection was either perfect or purely random. In our case, $E[\int_0^{T_\varepsilon} \tilde{n}^\varepsilon(s)ds]^2 = O(T_\varepsilon^2)$ and $\int_0^{T_\varepsilon}[\text{signal }(s) + \tilde{n}(s)]ds \sim N[O(T_\varepsilon), O(T_\varepsilon^2)]$. So the detection problem does not degenerate as $\varepsilon \to 0$.

Since we wish to work with small errors and to eventually do a valid linearization, let $\theta^\varepsilon(0) - \hat{\theta}^\varepsilon(0) = O(q_\varepsilon)$. This will guarantee that $\delta\theta^\varepsilon(t) = \theta^\varepsilon(t) - \hat{\theta}^\varepsilon(t) = O(q_\varepsilon)$ for any finite t. If we wish to assume $\delta\theta^\varepsilon(0) = O(q_\varepsilon^\alpha)$ for some $\alpha \in (0,1)$, then a different scaling would be used in the sequel, but this point will not be pursued.

The circuit of Figure 4 was investigated by Gardner [3], Appendix B, and elsewhere, using good engineering intuition, but without the benefit of the asymptotic theory, and required some ad-hoc assumptions (e.g., holding the state variable $\hat{\theta}^\varepsilon(t)$ fixed throughout, and using "approximate" spectral methods on a non-linear problem. Any verification of that technique and result must apparantly deal with an "asymptotic situation" scaled essentially as above.

The analysis is started in a conventional way, by making certain expansions and dropping some of the terms. This procedure, commented on below, can be rigorized at the expense of additional detail and notation. First, square the input, expand the products of the obtained trigonometric functions in the usual way in terms of the sums of sines and cosines' of the sums and difference of the angles. Drop the terms whose sin or cos factor does not contain a $2\omega_0^\varepsilon$. We suppose that the squarer contains a linear high pass filter which does this. Since the BW of the dropped terms and the BW about $2\omega_0^\varepsilon$ of the retained terms is $O(q_\varepsilon^{-2})$, there is no problem in explicitly introducing such a filter, but it unnecessarily complicates the notation. Next, having dropped the cited terms, multiply the remainder by the VCO output $\sin 2(\omega_0^\varepsilon t+\hat{\theta}^\varepsilon(t))$, expand the products of the trigonometric functions as above, but now drop the terms whose sine or cosine factor contains an ω_0^ε in the argument. We suppose, as in previous sections, that the multiplier contains a low pass filter which does this. In fact, such an assumption is not necessary, and if the dropped terms were carried through the analysis, they would not affect the limit. But, for notational convenience, it is helpful to drop them at this point. Denote the resulting term, the multiplier output, by $u^\varepsilon(t) = u^\varepsilon(t, \hat{\theta}^\varepsilon(t),\theta^\varepsilon(t))$, where

$$u^\varepsilon(t,\hat{\theta}^\varepsilon,\theta^\varepsilon) = \frac{-K_1^2}{4} m_\varepsilon^2(t)\sin 2(\theta^\varepsilon-\hat{\theta}^\varepsilon) + \frac{K_1\sigma}{2} m_\varepsilon(t)[-z_1^\varepsilon(t)\sin(\theta^\varepsilon-2\hat{\theta}^\varepsilon)$$

$$+ z_2^\varepsilon(t)\cos(\theta^\varepsilon - 2\hat{\theta}^\varepsilon)] + \frac{\sigma^2}{4}[((z_1^\varepsilon(t))^2 - (z_2^\varepsilon(t))^2)\sin 2\hat{\theta}^\varepsilon$$

$$+ 2z_1^\varepsilon(t)z_2^\varepsilon(t)\cos 2\hat{\theta}^\varepsilon]. \tag{6.1}$$

Define $x_2^\varepsilon = (\theta^\varepsilon - \hat{\theta}^\varepsilon)/q_\varepsilon$, $x_1^\varepsilon = v^\varepsilon/q_\varepsilon$. Then

$$\dot{x}_1^\varepsilon = Dx_1^\varepsilon + H[\frac{-K_1^2 m_\varepsilon^2(t)}{4} \frac{\sin 2(\theta^\varepsilon - \hat{\theta}^\varepsilon)}{q_\varepsilon}]$$

$$+ \frac{HK_1 c}{2} \frac{m_\varepsilon(t)}{q_\varepsilon}[-z_1^\varepsilon(t)\sin(\theta^\varepsilon - 2\hat{\theta}^\varepsilon) + z_2^\varepsilon(t)\cos(\theta^\varepsilon - 2\hat{\theta}^\varepsilon)] \tag{6.2a}$$

$$+ \frac{H\sigma^2}{4}[\frac{((z_1^\varepsilon(t))^2 - (z_2^\varepsilon(t))^2)}{q_\varepsilon}\sin 2\hat{\theta}^\varepsilon + \frac{2z_1^\varepsilon(t)z_2^\varepsilon(t)}{q_\varepsilon}\cos 2\hat{\theta}^\varepsilon]$$

$$\equiv Dx_1^\varepsilon + H[\frac{-K_1^2 m_\varepsilon^2(t)\sin 2(\theta^\varepsilon - \hat{\theta}^\varepsilon)}{4q_\varepsilon}] + \frac{HK_1\sigma}{2q_\varepsilon} v_1^\varepsilon(t,\theta^\varepsilon,\hat{\theta}^\varepsilon)m_\varepsilon(t)$$

$$+ \frac{H\sigma^2}{4q_\varepsilon} v_2^\varepsilon(t,\theta^\varepsilon,\hat{\theta}^\varepsilon) \tag{6.2b}$$

$$\dot{x}_2^\varepsilon = \dot{\theta}^\varepsilon/q_\varepsilon - Cx_1^\varepsilon, \tag{6.2c}$$

where (6.2) defines the $v_i^\varepsilon(\cdot)$. Owing to the normalization, we require that $\dot{\theta}^\varepsilon(\cdot)/q_\varepsilon$ converge to a right continuous function $\tilde{\theta}(\cdot)$ (with left-hand limits) as $\varepsilon \to 0$, or else we deal with a subsequence which does converge. We assumed $\theta^\varepsilon = O(q_\varepsilon)$ in order to be able to do the linearization analysis below. For such a linearization to be valid, we must have $\theta^\varepsilon(\cdot) \to 0$ as $\varepsilon \to 0$. If the rate is slower than $O(q_\varepsilon)$, the effect of the $\theta^\varepsilon(\cdot)$ variations on the tracking errors increase to ∞ relative to the effects of the $m_\varepsilon(\cdot)$ and $\tilde{n}^\varepsilon(\cdot)$. Our scaling is the unique one for which the noise and $\theta^\varepsilon(\cdot)$ effects are commensurate, for small $\varepsilon > 0$.

Define

$$\sigma_m^2 = \frac{1}{T_\varepsilon}\int_0^{T_\varepsilon} Em_\varepsilon^2(s)ds = \frac{1}{T}\int_0^T \sum_{k=-\infty}^{\infty} p^2(s-kT)ds.$$

We now follow a procedure very similar to those of the previous sections, again omitting most details. The main differences being due to the periodicity of $Em_\varepsilon^2(s)$ and of (6.9), which forces us to use an additional averaging. We have $\sin 2(\theta^\varepsilon-\hat{\theta}^\varepsilon)/q_\varepsilon = 2x_2^\varepsilon + x_2^\varepsilon O(|\theta^\varepsilon-\hat{\theta}^\varepsilon|)$. Owing to the assumption on the initial conditions, if the $O(\cdot)$ above were carried through the analysis, it would contribute nothing to the limit. For convenience in the analysis here, it is dropped henceforth and we replace $\sin 2(\theta^\varepsilon-\hat{\theta}^\varepsilon)/q_\varepsilon$ in (6.2) by $2x_2^\varepsilon$. This must be kept in mind in the manipulations below. We use the form $f^\varepsilon(t) = f(x^\varepsilon,t) + \sum_{i=0}^{2} f_i^\varepsilon(t)$ for the perturbed test function.

As usual for $f(\cdot) \in \mathcal{D}$ (write $x^\varepsilon = x^\varepsilon(t)$),

$$\hat{A}^\varepsilon f(x^\varepsilon,t) = f_t(x^\varepsilon,t) + f'_{x_1}(x^\varepsilon,t)\dot{x}_1^\varepsilon + f'_{x_2}(x^\varepsilon,t)\dot{x}_2^\varepsilon. \tag{6.3}$$

Again, the components of (6.3) containing the processes $m_\varepsilon(\cdot)$ or $z^\varepsilon(\cdot)$ must be averaged out. The first term which we will average out is $f'_x(x^\varepsilon,t)H[-K_1^2 m_\varepsilon^2(t)x_2^\varepsilon/2]$. For this, we use the test function perturbation $f_0^\varepsilon(t) = f_0^\varepsilon(x^\varepsilon(t),t)$, where we define

$$\begin{aligned} f_0^\varepsilon(x^\varepsilon,t) &= \frac{-K_1^2}{2}\int_0^\infty f'_{x_1}(x^\varepsilon,t+s)HE_t^\varepsilon[m_\varepsilon^2(t+s) - \sigma_m^2]x_2^\varepsilon ds \\ &= -\frac{K_1^2 q_\varepsilon^2}{2}\int_0^\infty f'_{x_1}(x^\varepsilon,t+q_\varepsilon^2 s)HE_t^\varepsilon[m_\varepsilon^2(t+q_\varepsilon^2 s) - \sigma_m^2]x_2^\varepsilon ds. \end{aligned} \tag{6.4}$$

By the assumptions on $p(\cdot)$ and on $\{a_k\}$, $|E_t^\varepsilon m_\varepsilon^2(t+q_\varepsilon^2 s)$

$\sum_{-\infty}^{\infty} p^2(\frac{t}{q_\varepsilon^2} + s - kT)| \le \delta^\varepsilon(t,s) \to 0$ as $s \to \infty$ and is integrable in

s, uniformly in t,ε. Since the above sum is periodic with period T, we center $m_\varepsilon^2(\cdot)$ about its arithmetic mean σ_m^2 in (6.4). The integral in (6.4) is $O(q_\varepsilon^2)$.

We have $f_0^\varepsilon(\cdot) \in \mathcal{D}(\hat{A}^\varepsilon)$ and it can readily be shown that

$$\hat{A}^\varepsilon f_0^\varepsilon(x^\varepsilon,t) = -\frac{K_1^2 x_2^\varepsilon}{2} f'_{x_1}(x^\varepsilon,t)H(m_\varepsilon^2(t) - \sigma_m^2) + O(q_\varepsilon).$$

Thus, the $m_\varepsilon^2(\cdot)$ term in (6.2) contributes only its mean value to the limit equation. See (6.14) for the summarizing calculation.

Next, we average out the "remaining" noise terms in (6.3). This requires (as usual) $f_1^\varepsilon(x^\varepsilon(t),t) = f_1^\varepsilon(t)$, where

$$f_1^\varepsilon(x^\varepsilon,t) = \int_0^\infty \frac{\sigma K_1}{2} H' f_{x_1}(x^\varepsilon,t+s)E_t^\varepsilon \frac{v_1^\varepsilon(t+s,\theta^\varepsilon(t),\hat{\theta}^\varepsilon)m_\varepsilon(t+s)ds}{q_\varepsilon}$$
$$+ \int_0^\infty \frac{\sigma^2}{4} H' f_{x_1}(x^\varepsilon,t+s)E_t^\varepsilon \frac{v_2^\varepsilon(t+s,\theta^\varepsilon(t),\hat{\theta}^\varepsilon)ds}{q_\varepsilon} \quad . \tag{6.5}$$

As usual, the "remaining" noise term in (6.3) is just the sum of the integrands in (6.5), evaluated at $s = 0$. By using the change of variable $s/q_\varepsilon^2 \to s$, we can show that $f_1^\varepsilon(t)$ is bounded by $O(q_\varepsilon)[1+|z^\varepsilon(t)|^2]$.

Proceeding as usual, we next get that

$$\hat{A}^{\varepsilon} f_1^{\varepsilon}(x^{\varepsilon},t) = -(\text{integrand of (6.5) evaluated at } s = 0) \tag{6.6}$$
$$+ (\text{terms whose } \underset{\varepsilon\to 0}{\text{p-lim}} \text{ is zero}) + Q^{\varepsilon}(x^{\varepsilon},\theta^{\varepsilon}(t),t),$$

where

$$Q^{\varepsilon}(x^{\varepsilon},\theta^{\varepsilon}(t),t) = (f^{\varepsilon}_{1,x_1}(x^{\varepsilon},t))'\cdot(\text{last two terms on right side of (6.2a)}). \tag{6.7}$$

The second term on the right side of (6.6) is bounded by $O(q_\varepsilon)[1+|Z^{\varepsilon}(t)|^4]$. Actually, $Q^{\varepsilon}(x^{\varepsilon},\theta^{\varepsilon}(t),t)$ is just (6.5) with f_{x_1} replaced by $f_{x_1x_1}$ and multiplied on the right by the last two terms on the right side of (6.2a).

Using the mutual independence of the $\{z_i^{\varepsilon}(\cdot),\ i = 1,2,m_\varepsilon(\cdot)\}$ processes together with the fact that the $z_i^{\varepsilon}(\cdot)$ have mean zero and are Gaussian and that $\sin^2 u + \cos^2 u \equiv 1$, we get that

$$EQ^{\varepsilon}(x^{\varepsilon},\theta^{\varepsilon}(t),t) = \frac{K_1^2\sigma^2}{4q_\varepsilon^2}\int_0^\infty H' f_{x_1x_1}(x^{\varepsilon},t+s)H\rho(s/q_\varepsilon^2)Em_\varepsilon(t+s)m_\varepsilon(t)ds$$
$$+ \frac{4}{q_\varepsilon^2}\left(\frac{\sigma^2}{4}\right)^2\int_0^\infty H' f_{x_1x_1}(x^{\varepsilon},t+s)H\rho^2(s/q_\varepsilon^2)ds, \tag{6.8}$$

where

$$Em_\varepsilon(t+s)m_\varepsilon(t) = \sum_{k=-\infty}^{\infty} p_\varepsilon(t+s-kT_\varepsilon)p_\varepsilon(t-kT_\varepsilon). \tag{6.9}$$

The expression (6.9) is periodic in t (period T_ε). Because of this periodicity, an arithmetic average of (6.9) is used in defining the centering term for $f_2^{\varepsilon}(\cdot)$.

Let us write

$$\frac{1}{T_\varepsilon}\int_0^{T_\varepsilon} Em_\varepsilon(t+\tau+s)m_\varepsilon(t+\tau)d\tau \equiv V(s/q_\varepsilon^2) =$$

$$\frac{1}{T}\int_0^T \sum_{k=-\infty}^{\infty} p(\frac{t+s}{q_\varepsilon^2} + \tau - kT)p(\frac{t}{q_\varepsilon^2} + \tau - kT)d\tau \qquad (6.10)$$

$$= \frac{1}{T}\int_0^T \sum_{k=-\infty}^{\infty} p(\frac{s}{q_\varepsilon^2} + \tau - kT)p(\tau - kT)d\tau .$$

Now, define $f_2^{\varepsilon}(\cdot)$ by $f_2^{\varepsilon}(t) = f_2^{\varepsilon}(x^{\varepsilon}(t),t)$, where

$$f_2^{\varepsilon}(x^{\varepsilon},t) = \frac{K_1^2\sigma^2}{4q_\varepsilon^2}\int_0^\infty d\tau \int_0^\infty ds H' f_{x_1x_1}(x^{\varepsilon},t+\tau+s)H\{E_t^{\varepsilon}[v_1^{\varepsilon}(t+\tau+s)v_1^{\varepsilon}(t+\tau)m_\varepsilon(t+\tau+s)m_\varepsilon(t+\tau)]$$

$$- \rho(s/q_\varepsilon^2)V(s/q_\varepsilon^2)\}$$

$$+ \frac{K_1\sigma}{2q_\varepsilon^2}(\frac{\sigma^2}{4})^2\int_0^\infty d\tau \int_0^\infty ds\, H' f_{x_1x_1}(x^{\varepsilon},t+\tau+s)HE_t^{\varepsilon}[v_1^{\varepsilon}(t+\tau+s)m_\varepsilon(t+\tau+s)v_2^{\varepsilon}(t+\tau)$$

$$+ v_2^{\varepsilon}(t+\tau+s)v_1^{\varepsilon}(t+\tau)m_\varepsilon(t+\tau)] \qquad (6.11)$$

$$+ \frac{1}{q_\varepsilon^2}(\frac{\sigma^2}{4})^2\int_0^\infty d\tau \int_0^\infty ds H' f_{x_1x_1}(x^{\varepsilon},t+\tau+s)H\{E_t^{\varepsilon}v_2^{\varepsilon}(t+\tau+s)v_2^{\varepsilon}(t+\tau)$$

$$- Ev_2^{\varepsilon}(t+\tau+s)v_2^{\varepsilon}(t+\tau)\}.$$

The centering term in (6.11) is precisely (6.8) but with $Em_\varepsilon(t+s)m_\varepsilon(t)$ replaced by its arithmetic mean $V(s/q_\varepsilon^2)$. It can be shown that $f_2^{\varepsilon}(t)$ is bounded by

$$O(q_\varepsilon^2)[1+|z^\varepsilon(t)|^6] \tag{6.12}$$

and that

$$\hat{A}^\varepsilon f_2^\varepsilon(t) = (6.8) \text{ (but with } Em_\varepsilon(t+s)m_\varepsilon(t) \text{ replaced by } V(s/q_\varepsilon^2)) - Q^\varepsilon(x^\varepsilon(t),\rho^\varepsilon(t),t) + \text{terms satisfying (6.12).} \tag{6.13}$$

Summarizing the above calculations and writing x^ε for $x^\varepsilon(t)$ yields

$$\underset{\varepsilon\to 0}{\text{p-lim}} \left[\sum_{i=0}^{2} f_i^\varepsilon(x^\varepsilon,t)\right] = 0$$

$$\underset{\varepsilon\to 0}{\text{p-lim}} \left[\hat{A}^\varepsilon f^\varepsilon(x^\varepsilon,t) - \left(\frac{\partial}{\partial t} + A\right)f(x^\varepsilon,t)\right] = 0,$$

where

$$\begin{aligned} Af(x,t) = {} & f'_{x_1}(x,t)\left[Dx_1 - H\frac{K_1^2\sigma_m^2 x_2}{2}\right] + f_{x_2}(x,t)[\tilde{\theta}-Cx_1] \\ & + \left[\frac{\sigma^4}{4}\int_0^\infty \rho^2(s)ds + \frac{K_1^2\sigma^2}{4}\int_0^\infty \rho(s)V(s)ds\right]H'f_{x_1x_1}(x,t)H. \end{aligned} \tag{6.14}$$

Tightness can also be shown by applying [6, Theorem 1] with the given order estimates. By Theorem 1, $\{x^\varepsilon(\cdot)\}$ converges weakly to $x(\cdot) = (x_1(\cdot),x_2(\cdot))$, where

$$\begin{aligned} dx_1 &= [Dx_1 - HK_1^2\sigma_m^2 x_2/2]dt + \sqrt{2}\,\sigma_3 HdB \\ dx_2 &= [\tilde{\theta} - Cx_1]dt, \end{aligned} \tag{6.15}$$

$B(\cdot)$ = standard Brownian motion,

where σ_3^2 is the last bracketed term in (6.14). With non-Gaussian noise, the limit is the same except for the value of σ_3^2.

REFERENCES

[1] W.C. Lindsey, Synchronization Systems in Communications and Control, Prentice-Hall, Englewood Cliffs, N.J., 1972.

[2] D. Williams, "Improved phase-locked loop performance via non-linear loop filters," IEEE Trans. on Communications, COM-27, 1979, pp. 542-556.

[3] F.L. Gardner, Phaselock Techniques, 2nd Ed., Wiley, New York, 1979.

[4] H.J. Kushner, "Diffusion approximations to output processes of non-linear systems with wide inputs," IEEE Trans. on Information Theory, to appear late 1980.

[5] H.J. Kushner, "Jump-diffusion approximations for ordinary differential equations with wide-band random right hand sides," SIAM Journal on Control and Optimization. 17, pp. 729-744, 1979.

[6] H.J. Kushner, "A martingale method for the convergence of a sequence of processes to a jump-diffusion process," Z. Wahrscheinlichtkeitsteorie, 53, pp. 207-219, 1980.

[7] T.G. Kurtz, "Semigroups of conditional shifts and approximation of Markov processes," Ann. Prob. 4, pp. 618-642, 1975.

[8] G. Blankenship, G.C. Papanicolaou, "Stability and control of stochastic systems with wide-band noise disturbances," SIAM J. on Applied Math. 34, pp. 437-476, 1978.

[9] A.J. Viterbi, Principles of Coherent Communication, McGraw-Hill, New York, 1966.

[10] H.J. Kushner, Y. Bar-Ness, "Analysis of non-linear systems with wide-band inputs," IEEE Trans. on Automatic Control, to appear late 1980.