Fourier Series and Estimation on the Circle with Applications to Synchronous Communication—Part II: Implementation

ALAN S. WILLSKY, MEMBER, IEEE

***Abstract*—The practical implementation of the infinite-dimensional optimal estimation results presented in Part I of this series is considered. Several techniques are described in detail. Included among these is the so-called "assumed density" approximation technique. Finite-dimensional suboptimal filtering equations based on this method are derived for several of the phase-tracking/demodulation problems studied in Part I. Finally, these techniques are applied to a phase tracking problem of importance in navigation systems such as Omega, and simulation results are reported that favorably compare a system designed using these techniques to an optimal phase-lock loop and an optimal linear system.**

I. Introduction

In PART I of this series [1] we studied a wide variety of discrete- and continuous-time phase-tracking and demodulation problems in the presence of a number of different noise sources. We derived infinite-dimensional optimal estimation equations by considering the stochastic differential equations satisfied by the conditional expectations of certain functions of signal phase, frequency, and amplitude. These equations display the rich structure present in rather large classes of estimation problems on the circle S^1. (See [2]–[12] for other results on S^1 estimation.)

However, for any practical application of these results we must approximate the optimal estimation equations. In Section III we will discuss several methods for truncating the infinite-dimensional estimation equations derived in Part I. As in [1], we develop these general techniques by examining several specific examples. Also, we concentrate on the continuous-time problem, but the extension to discrete-time problems is clear. Section IV contains the results of simulations that compare the performance of a system designed using these techniques to that of two other systems for a phase-tracking problem important in navigation systems such as Omega [5], [15]. In the next section we briefly review the design of phase-lock loop (PLL) systems, since PLL systems provide several interesting comparisons with the Fourier series techniques developed here and in [1].

Manuscript received July 26, 1973; revised April 18, 1974. This work was done in part while the author was a Fannie and John Hertz Foundation Fellow in the Department of Aeronautics and Astronautics, M.I.T., and in part at the Decision and Control Sciences Group of the M.I.T. Electronic Systems Laboratory with partial support provided by AFOSR under Grant 72-2273.
The author is with the Department of Electrical Engineering, Massachusetts Institute of Technology, Cambridge, Mass. 02139.

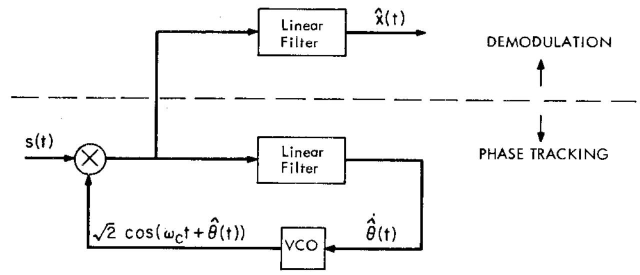

Fig. 1. Basic PLL model.

II. Phase-Lock Loop

In this section we consider a very important class of phase-tracking and demodulation systems. We will later use the PLL to understand the physical significance of the estimation techniques we develop and will also compare the performance of the two for an important example.

The basic phase-lock loop model [13] is illustrated in Fig. 1. The received signal is of the form

$$s(t) = \sqrt{2P} \sin(\omega_c t + \theta(t)) + \dot{w}(t) \tag{1}$$

where $\theta(t)$ is usually taken to be some type of linear diffusion process and $\dot{w}$ is a white noise. The part of the loop below the dashed line in Fig. 1 essentially performs the function of tracking the time-varying phase (which is all that is needed for some applications, such as Omega [15]), while the filtering above the dashed line performs the desired demodulation, e.g., if

$$\theta(t) = cx(t) \qquad \text{or} \qquad \theta(t) = c \int_0^t x(s)\, ds.$$

Following Van Trees [13], we can write the product of $s(t)$ and the output of the voltage controlled oscillator (VCO) as

$$\begin{aligned} \sqrt{2}\, s(t) \cos(\omega_c t + \hat{\theta}(t)) &= \sqrt{2}\, \dot{w}(t) \cos(\omega_c t + \hat{\theta}(t)) + \sqrt{P} \sin(\theta(t) - \hat{\theta}(t)) \\ &\quad + \sqrt{P} \sin(2\omega_c t + \theta(t) + \hat{\theta}(t)). \end{aligned} \tag{2}$$

The reader is referred to [13], in which it is argued that $n(t) = \sqrt{2}\, \dot{w}(t) \cos(\omega_c t + \hat{\theta}(t))$ is essentially a white noise process of strength equal to that of $\dot{w}$. Then, if we assume that the phase-tracking linear filter has a bandwidth much smaller than $2\omega_c$, we can ignore the double-frequency

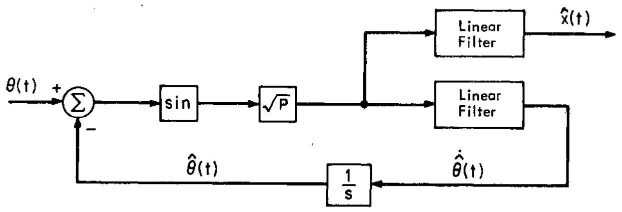

Fig. 2. Baseband PLL model.

term in (2). In this case, we obtain the baseband model of the PLL depicted in Fig. 2.

The simplest PLL system is the first-order loop, in which the phase-tracking linear filter is taken to be a constant gain. More complicated loops can be obtained by using standard Kalman filtering or Wiener–Hopf techniques. In this case, the assumption that the loop is "above threshold" [13], i.e., that the approximation

$$\sin(\theta(t) - \hat{\theta}(t)) \simeq \theta(t) - \hat{\theta}(t) \tag{3}$$

is valid, is used to linearize the PLL model in Fig. 2. Then standard techniques can be used to determine the optimum linear filter (given the statistical properties of $\theta(t)$).

III. Moment Truncation Methods for Phase-Tracking and Demodulation Problems

In this section we discuss a few methods for truncating the infinite sets of moment equations encountered in [1]. As mentioned earlier, we will specifically treat only continuous-time problems; however, these techniques are also applicable to discrete-time problems. In addition, as discussed in [14], some of these techniques are quite general and can be applied to large classes of nonlinear estimation problems.

In the present discussion we will examine only a few of the types of approximations that can be used. The reader is referred to [5], [7], [14], and [19] for other discussions of numerical methods in filtering theory. As in [1], we will treat the problem at hand by considering examples—a phase estimation problem and a phase demodulation example, both considered in [1]. (The techniques discussed here can be readily extended to the other problems studied in [1].)

Example 1

We consider the problem analyzed in Example 4 of [1]. We receive the signal

$$dz(t) = \sin\theta(t)\, dt + r^{1/2}(t)\, dw(t) \tag{4}$$

where

$$\theta(t) = \omega_c t + \int_0^t q^{1/2}(s)\, dv(s) + \theta_0 \tag{5}$$

and v and w are independent Brownian motions, $\mathscr{E}(dv^2(t)) = \mathscr{E}(dw^2(t)) = dt$, $q(t) \geq 0$, $r(t) > 0$, and $\omega_c > 0$. Also θ_0 is a random initial condition independent of v and w. The problem is to choose $\hat{\theta}(t \mid t)$ to minimize $\mathscr{E}[1 - \cos(\theta(t) - \hat{\theta}(t \mid t)) \mid z(s), 0 \leq s \leq t]$.

The optimal filter can be described [1] as follows: let

$$\begin{aligned} c_n(t) &= \frac{1}{2\pi}\, \mathscr{E}[e^{-in\theta(t)} \mid z(s), 0 \leq s \leq t] \\ &= b_n(t) - ia_n(t). \end{aligned} \tag{6}$$

Then

$$\begin{aligned} dc_n(t) = &- \left[in\omega_c + \frac{n^2}{2}\, q(t)\right] c_n(t)\, dt \\ &+ \left[\frac{(c_{n-1}(t) - c_{n+1}(t))}{2i} + 2\pi c_n(t)\, \mathrm{Im}\,(c_1(t))\right] \\ &\cdot \left[\frac{dz(t) + 2\pi\, \mathrm{Im}\,(c_1(t))\, dt}{r(t)}\right] \end{aligned} \tag{7}$$

$$\hat{\theta}(t \mid t) = \tan^{-1}\left(\frac{a_1(t)}{b_1(t)}\right). \tag{8}$$

As discussed in [1], this filter consists of an infinite bank of second-order filters (see Figs. 1 and 2 in [1]). Referring to Fig. 2 in [1] or to (7) we see that the c_n filter looks like a damped oscillator (bandpass filter) at the frequency $n\omega_c$, with nonlinear (product) coupling terms to the other filters and to the measurement dz. In particular, one of these coupling terms involves the multiplication of c_n by dz.

Now consider the PLL described in Section II. The VCO, operating at the frequency ω_c, produces an output that looks something like $2\pi\sqrt{2}\, b_1(t)$, and this output multiplies the received signal $s(t)$. This multiplication feature strongly resembles the product terms in the optimal system. Thus the optimal system can be (loosely) interpreted as an infinite bank of PLL's, with resonant frequencies being various multiples of ω_c.

The problem we wish to address here is a practical one. Can we find finite dimensional approximations to the infinite dimensional optimal equations (7) and (8); i.e., can we successfully truncate the infinite back of filters? In some sense, what we wish to do is to approximate the S^1 density

$$p_\theta(\xi,t) = \sum_{n=-\infty}^{+\infty} c_n(t) e^{in\xi} \tag{9}$$

by a density determined by a finite set of parameters.

A natural approximation is

$$\tilde{p}_\theta(\xi,t) = \sum_{n=-N}^{+N} c_n(t) e^{in\xi} \tag{10}$$

i.e., assume $c_m(t) = 0$, for all $m > N$. As discussed in [19], the Fourier coefficients usually fall off at least as $1/n^2$, and thus, for large N, this straightforward truncation method may work quite well. Of course one needs some numerical results to determine how many terms are enough. The reader is referred to Section IV for a discussion that indicates that this straightforward method is not as good as it first appears to be.

A second truncation method is suggested by the PLL. Recall from Section II that a crucial assumption in the design of PLL's is that the linear filter in the phase-tracking loop is low-pass and cuts off terms at carrier frequency $2\omega_c$. Thus, in some sense, the PLL can be regarded [5] as a truncation of the infinite bank of filters in which we keep only the lowest mode and remove the coupling to the higher modes by filtering out of the $2\omega_c$ term. This suggests an approximation method for (7). The Fourier coefficients $c_n(t)$ can be written

$$c_n(t) = \frac{1}{2\pi} e^{-i(n\omega_c t + \alpha_n(t))}. \tag{11}$$

Thus, if we use a low-pass filter that allows terms with frequencies $\leq N\omega_c$ to pass unamplified and not phase-shifted, but chops off all frequencies $\geq (N+1)\omega_c$, we can effectively truncate the infinite bank of filters. Consider the term in the differential equation for c_N that causes the difficulty, i.e., that contains the coupling to the higher modes. Specifically consider the term

$$\eta_N(t) = \left[\frac{c_{N-1}(t) - c_{N+1}(t)}{2i} + 2\pi c_N(t) \operatorname{Im}(c_1(t))\right]. \tag{12}$$

If we passed η_N through an ideal low-pass filter of the desired type the output would be

$$\tilde{\eta}_N(t) = \frac{c_{N-1}(t)}{2i} + \frac{\pi}{i} e^{-i((N-1)\omega_c t + \alpha_N(t) - \alpha_1(t))}. \tag{13}$$

The effect of the low-pass filter is precisely the same if we assume

$$c_{N+1}(t) = 2\pi c_N(t) c_1(t) \tag{14}$$

which is, in fact, true if we are tracking perfectly. Thus, if we are tracking well, (14) may be a reasonable approximation to use to truncate the bank of filters.

We now discuss several examples of what has been called the "assumed density" type of approximation (see [14], [20]). The basic idea of the approximation is the following: we assume that the conditional density has some known form that is specified by a finite set of parameters; then, having $\{c_n\}_{n=1}^N$, we compute the assumed density parameters and the associated value of c_{N+1}. For example, for the assumed density form (10), we have $c_{N+1} = 0$. A slightly more complicated example involves the assumption that $p_\theta(\xi,t)$ is a folded normal density (see [1])

$$p_\theta(\xi,t) = \frac{1}{2\pi} \sum_{n=-\infty}^{+\infty} e^{-n^2\gamma/2} e^{in(\xi-\eta)} = F(\xi;\eta,\gamma). \tag{15}$$

In this case, if we have computed c_1 and if we assume p_θ is given by (15), we can compute c_{N+1} from the equation

$$c_{N+1} = (2\pi)^{(N+1)^2-1} |c_1|^{N(N+1)} c_1^{(N+1)}. \tag{16}$$

In fact, for the folded normal, we can compute all of the c_k if we know any one of them other than c_0. Using these relationships among the c_i for the folded normal, we can obtain equations, such as (16), that can be used to truncate our bank of filters. In the next section we present results for this approximation method.

One could also consider approximating p_θ by a uniform density

$$p_\theta(\xi) = \begin{cases} \dfrac{1}{\theta_b - \theta_a}, & -\pi \leq \theta_a \leq \xi \leq \theta_b < \pi \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

We will not discuss this method here but will discuss an R^1 analog later in this section. In addition, we could consider approximations of the form

$$p_\theta(\xi) = \sum_{n=1}^{M} d_n F(\theta; \eta_n, \gamma_n) \tag{18}$$

(the set of such densities is dense in $L^1(-\pi,\pi)$, [3], [5]). The details of several approximation schemes based on (18) are given in [5]. We now consider some R^1 approximation techniques. Again, to be specific, suppose we consider the phase demodulation problem of Example 5 in [1].

Example 2

Consider the R^1 signal process $x(t)$ satisfying

$$dx(t) = a(t)x(t)\,dt + q^{1/2}(t)\,dv(t) \tag{19}$$

where v is as before and is independent of $x(0)$. We wish to compute $\hat{x}(t \mid t) = \mathscr{E}[x(t) \mid z(s), 0 \leq s \leq t]$, where z satisfies

$$dz(t) = \sin(\omega_c t + x(t))\,dt + r^{1/2}(t)\,dw(t). \tag{20}$$

Defining

$$\begin{aligned} c_{nm}(t) &= \frac{1}{2\pi} \mathscr{E}[x^n(t) e^{-im\theta(t)} \mid z(s), 0 \leq s \leq t] \\ &= b_{nm}(t) - i a_{nm}(t) \end{aligned} \tag{21}$$

where

$$\theta(t) = (\omega_c t + x(t)) \bmod 2\pi \tag{22}$$

we have

$$\begin{aligned} dc_{nm}(t) = &\left\{\left[na(t) - \frac{m^2 q(t)}{2} - im\omega_c\right] c_{nm}(t)\right. \\ &- ima(t)c_{n+1,m}(t) \\ &\left. + \frac{n(n-1)q(t)}{2} c_{n-2,m}(t) - iq(t)nmc_{n-1,m}(t)\right\} dt \\ &+ \left[\frac{c_{n,m-1}(t) - c_{n,m+1}(t)}{2i} + 2\pi c_{nm}(t) \operatorname{Im}(c_{01}(t))\right] \\ &\cdot \left[\frac{dz(t) + 2\pi \operatorname{Im}(c_{01}(t))\,dt}{r(t)}\right] \end{aligned} \tag{23}$$

$$\hat{x}(t \mid t) = 2\pi b_{10}(t). \tag{24}$$

Suppose we compute only c_{nm}, for $n = 1,\cdots,N$ and $m = 1,\cdots,M$. From (23) we see that to truncate the filter equations effectively, we must compute approximations for $\{c_{N+1,m}\}_{m=1}^M$ and $\{c_{n,M+1}\}_{n=1}^N$. In some sense this truncation problem is more difficult than the S^1 problem discussed previously, since, for example, the moments

$$c_{n0}(t) = \frac{1}{2\pi}\,\mathscr{E}[x^n(t) \mid z(s), 0 \le s \le t] \tag{25}$$

do not necessarily go to zero for large n.

We first note that we can come up with truncation techniques analogous to those for the S^1 problem by replacing folded normal densities with normal densities. (Note that finite linear combinations of normal densities are dense in $L^1(-\infty,\infty)$, [23].) For such techniques we need to know the form of $\mathscr{E}(x^n e^{-imx})$, if x has the density $N(x;\eta,\gamma)$ (normal density with mean η and variance γ). For such a random variable x, let

$$p_{nm} = \mathscr{E}(x^n e^{-imx}). \tag{26}$$

We can then show that for all $m \ge 0$, we have the following equations:

$$p_{0m} = e^{-m^2\gamma/2}e^{-im\eta} \tag{27}$$

$$p_{nm} = \eta p_{n-1,m} + \gamma[(n-1)p_{n-2,m} - imp_{n-1,m}], \qquad n \ge 1. \tag{28}$$

An obvious analog of one of the S^1 techniques is to approximate the density $p(x,t)$, for $x(t)$ conditioned on $z(s)$, $s \le t$, by

$$\tilde{p}(x,t) = N(x; 2\pi c_{10}(t), 2\pi c_{20}(t) - 4\pi^2 c_{10}^2(t)). \tag{29}$$

Then, if we compute c_{nm}, for $n = 1,\cdots,N$ and $m = 1,\cdots,M$, we can approximate $\{c_{N+1,m}\}_{m=1}^{M}$ and $\{c_{n,M+1}\}_{n=1}^{N}$ by using (26)–(28) with the approximation (29) and the relation

$$c_{nm}(t) = \frac{1}{2\pi}\,e^{-im\omega_c t}\mathscr{E}[x^n(t)e^{-imx(t)} \mid z(s), 0 \le s \le t]. \tag{30}$$

Using this approximation in (23), we can truncate the filter equations. We also note that analogous to the S^1 case, we can approximate $p(x,t)$ by a finite linear combination of normal densities.

We now consider another assumed form density technique. In this method we again use $c_{10}(t)$ and $c_{20}(t)$ to compute approximations for

$$\{c_{N+1,m}(t)\}_{m=1}^{M} \qquad \text{and} \qquad \{c_{n,M+1}(t)\}_{n=1}^{N}.$$

As discussed in [14], we assume that $\tilde{p}(x,t)$, our approximation to the conditional density $p(x,t)$, is the uniform density

$$\tilde{p}(x,t) = U(x;\eta(t),\sigma(t)) = \begin{cases} \dfrac{1}{2\sigma(t)}, & \eta(t) - \sigma(t) \le x \le \eta(t) + \sigma(t) \\ 0, & \text{otherwise} \end{cases} \tag{31}$$

where

$$\eta(t) = 2\pi c_{10}(t) \tag{32}$$

$$\sigma(t) = [2\pi c_{20}(t) - \eta^2(t)]^{1/2}. \tag{33}$$

As with the normal density, we can compute $d_{nm} = \mathscr{E}(x^n e^{-imx})$, if x has the density $U(x;\eta,\sigma)$

$$d_{n0} = \frac{(\eta+\sigma)^{n+1} - (\eta-\sigma)^{n+1}}{2\sigma(n+1)} \tag{34}$$

and for $n \ge 0$ and $m \ge 1$

$$d_{nm} = \frac{i}{2m\sigma}[(\eta+\sigma)^n e^{-im(\eta+\sigma)} - (\eta-\sigma)^n e^{-im(\eta-\sigma)}] - \frac{in}{m}\,d_{n-1,m}. \tag{35}$$

We can use these equations to truncate the phase demodulation equation (23).

Another technique involves the notion of *cumulants* of a probability density [24], [26]. Let x be a real-valued random variable with density $p(x)$, and let $\theta_x(u)$ be the characteristic function of x

$$\theta_x(u) = \mathscr{E}(e^{iux}) = \int_{-\infty}^{+\infty} e^{iux}p(x)\,dx. \tag{36}$$

If we write

$$\theta_x(u) = \exp\left\{\sum_{n=1}^{\infty}\frac{(iu)^n}{n!}k_n\right\} \tag{37}$$

the k_n are called the *cumulants* of x. The cumulants are related to the *moments*

$$m_n = \mathscr{E}(x^n) \tag{38}$$

by the formulas

$$\begin{aligned} k_1 &= m_1 \\ k_2 &= m_2 - m_1^2 \\ k_3 &= m_3 - 3m_1m_2 + 2m_1^3 \\ k_4 &= m_4 - 3m_2^2 - 4m_1m_3 + 12m_1^2m_2 - 6m_1^4. \\ &\vdots \end{aligned} \tag{39}$$

As mentioned earlier, we cannot assume that the $\mathscr{E}(x^n)$ tend to 0 as n approaches ∞; however, as discussed in [24], it is reasonable to assume that the cumulants tend to zero. Thus suppose we compute $\{c_{n0}(t)\}_{n=1}^{N}$ and assume $k_n(t) = 0$, for all $n > N$. We can then use (39) to compute $k_n(t)$, $n \le N$. Using the approximation

$$\tilde{\theta}_x(u,t) = \exp\left\{\sum_{n=1}^{N}\frac{(iu)^n}{n!}k_n(t)\right\} \tag{40}$$

we can compute

$$\tilde{p}(x,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iux}\tilde{\theta}_x(u,t)\,du \tag{41}$$

and the appropriate expectations.

The reader is referred to [7], [14], [24], and [25] for a variety of other techniques, including several using Hermite polynomials and quasi-moments. Finally, for the general multidimensional problem discussed in [1], we note that the preceding discussions suggest a general assumed density approach; i.e., we can approximate $p(y,t)$ by a multidimensional normal density, a sum of such terms, or a uniform density over some region of R^n. We will not discuss such techniques here, since the details are quite similar to those for the scalar problems previously considered.

IV. Phase-Tracking Problem in the Presence of Additive Channel Noise

In this section we will discuss the results of a series of simulations of several different types of phase-tracking systems. The tracking problem used is the one discussed in Example 1. We wish to track the phase $\theta(t)$ mod 2π, where

$$\theta(t) = \omega_c t + q^{1/2} v(t) \tag{42}$$

($v(t)$ is a standard Brownian motion process), and we observe

$$\dot{z}(t) = \sin \theta(t) + r^{1/2} \dot{w}(t) \tag{43}$$

(w is a standard Brownian motion independent of v). We note that $1/q$ is called the *oscillator coherence time* [13].

The first tracking method we discuss is a PLL system. The reader is referred to [13, pp. 37–41] for the development of the optimal PLL phase tracker. Referring to Fig. 1, the optimal (steady state) linear filter in the phase-tracking loop is a constant

$$k = \left(\frac{q}{r}\right)^{1/2}. \tag{44}$$

Also, the analysis in [13] yields the result that, if the linear assumption used to aid in the PLL analysis is valid (i.e., if the system is "above threshold"), the phase error variance (in radians) is

$$P_{\theta l} = \sqrt{2rq}. \tag{45}$$

A second phase-tracking system has been proposed by Gustafson and Speyer [9]. Essentially, their system is the optimal *linear* filter (in the sense of minimizing error variance). The reader is referred to [9] for the development of the filter equations.

The other two systems were motivated by the Fourier series results discussed in Example 1. As discussed there and in Section III, we must consider suboptimal filtering techniques that involve a truncation of the infinite Fourier series. The first method we have considered is the straightforward truncation procedure; i.e., assume all coefficients a_n and b_n are 0, for $n > N$. We will not present any simulation results for this method, since several runs were made with $N = 3$, and extremely poor results were obtained. An intuitive explanation for this is that the higher coefficients need *not* be negligible. For instance, suppose we know the phase perfectly; then the probability "density" is an impulse at the known value η, and the formal Fourier series expansion for this is

$$p(\theta) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{n=1}^{\infty} \cos n(\theta - \eta). \tag{46}$$

In this case, the various coefficients are of the same order. Thus, if we are tracking well (i.e., the density is nearly an impulse), the assumption that the higher coefficients are negligible is a poor one.

Thus it was necessary to devise an alternative truncation procedure. The one adopted was the folded normal assumed density approximation. The system that has been simulated is the simplest of this type; i.e., we only compute a_1 and b_1 and approximate a_2 and b_2 using (16). Referring to the Fourier series equation (7), our suboptimal filter equations are

$$\dot{a}_1 = \left(\omega_c b_1 - \frac{q}{2} a_1\right) + \frac{(\dot{z} - 2\pi a_1)}{r} \cdot \left[\frac{1}{2}\left(\frac{1}{2\pi} - 8\pi^3(b_1^4 - a_1^4) - 2\pi a_1^2\right]\right. \tag{47}$$

$$\dot{b}_1 = -\left(\omega_c a_1 + \frac{q}{2} b_1\right) + \frac{(\dot{z} - 2\pi a_1)}{r} \cdot [8\pi^3 a_1 b_1 (a_1^2 + b_1^2) - 2\pi a_1 b_1] \tag{48}$$

$$\hat{\theta} = \tan^{-1}\left(\frac{a_1}{b_1}\right). \tag{49}$$

Note that the right sides of (47) and (48) are polynomials in a_1 and b_1 and can-be computed easily.

We note that if we are tracking the phase perfectly, (16) gives *precisely* the correct values for a_2 and b_2. Then, since the differential equations for a_1 and b_1 do not explicitly depend on a_n and b_n, for $n > 2$, the finite dimensional filter (47)–(49) performs optimally; i.e., the a_1 and b_1 values are exactly the same as the values obtained from the optimal infinite dimensional filter, and thus our estimate $\hat{\theta} = \tan^{-1}(a_1/b_1)$ is the optimal one. Thus, for small noise variances, one would expect the highly nonlinear filter (47)–(49) to operate nearly optimally, where the "optimal" performance is that attained by the linearized PLL, i.e., $P_{\theta l}$. (Note that below threshold $P_{\theta l}$ is not actually achieved by the PLL.)

Finally, we note that the right sides of (47) and (48) are highly nonlinear, and this leads to two complications. The first of these concerns the existence of solutions to the equations, since the right sides do not satisfy the Lipschitz conditions that are used in the standard proof of the existence of solutions to Itô differential equations [25]. We do not prove the existence of a solution, but rather point out that the actual Fourier coefficients are bounded in magnitude by $\frac{1}{2}\pi$, so that we can replace the various terms on the right sides of (47) and (48) by "saturated" versions. For example, a_1^4 can be replaced by the function

$$f(a_1) = \begin{cases} a_1^4, & |a_1| \le \dfrac{1}{2\pi} \\ \dfrac{1}{16\pi^4}, & |a_1| > \dfrac{1}{2\pi}. \end{cases} \tag{50}$$

If we do this, we obtain equations that do satisfy the necessary Lipschitz conditions. We remark that this discussion is academic, since our simulations indicate that the performance of this filter is quite good, and the values of a_1 and b_1 in the simulations never exceeded $\frac{1}{2}\pi$.

TABLE I
PHASE TRACKING SIMULATION PERFORMANCE SUMMARY—RMS PHASE ERROR (DEGREES)

$P_{\theta\ell}$ (rad.^{2})	Linear Predicted RMSPE	Predicted PLL RMSPE	Actual PLL RMSPE	SDNF RMSPE	FCF RMSPE
0.041	11.6	11.6	12.0	11.9	11.6
0.130	20.6	20.6	20.3	20.1	19.8
0.225	27.1	29.1	28.8	28.8	28.1
0.4	36.6	44.0	43.4	42.0	41.2
0.7	48.2	62.5	61.2	57.3	57.0
1.0	57.3	72.5	71.9	69.1	68.9
1.3	65.1	78.2	76.2	73.1	73.0

A second complication caused by the nonlinearities in the filter equations arises in considering the numerical integration of the filter equations. Wong and Zakai [16], [17] have shown that in numerically integrating stochastic differential equations driven by white noise, one must include correction terms (which are nonzero only if the equations are nonlinear) in the equations to be integrated in order to obtain a numerical solution that faithfully approximates the solution to the original stochastic equations. Following [16], [17], we obtain the following equations that have been used in the numerical simulation of (47) and (48):

$$\dot{\alpha}_1 = \left(\omega_c\beta_1 - \frac{q}{2}\alpha_1\right) - \frac{1}{2}\left(\sigma_1\frac{\partial\sigma_1}{\partial\alpha_1} + \sigma_2\frac{\partial\sigma_1}{\partial\beta_1}\right) + \frac{(\dot{z} - 2\pi\alpha_1)}{r^{1/2}}\sigma_1 \tag{51}$$

$$\dot{\beta}_1 = -\left(\omega_c\alpha_1 - \frac{q}{2}\beta_1\right) - \frac{1}{2}\left(\sigma_1\frac{\partial\sigma_2}{\partial\alpha_1} + \sigma_2\frac{\partial\sigma_2}{\partial\beta_1}\right) + \frac{(\dot{z} - 2\pi\alpha_1)}{r^{1/2}}\sigma_2 \tag{52}$$

$$\sigma_1 = \frac{1}{r^{1/2}}\left[\frac{1}{2}\left(\frac{1}{2\pi} - 8\pi^3(\beta_1{}^4 - \alpha_1{}^4) - 2\pi\alpha_1{}^2\right)\right] \tag{53}$$

$$\sigma_2 = \frac{1}{r^{1/2}}\left[8\pi^3\alpha_1\beta_1(\alpha_1{}^2 + \beta_1{}^2) - 2\pi\alpha_1\beta_1\right] \tag{54}$$

$$\hat{\theta} = \tan^{-1}\left(\frac{\alpha_1}{\beta_1}\right). \tag{55}$$

We note that Wong and Zakai's results also require Lipschitz conditions on the right side, but the same type of arguments as shown previously can be used here. Also, the integration scheme used was the Runge–Kutta method, for which the "full" Wong–Zakai correction term is needed [27]. For other discussion along these lines, we refer the reader to [27]–[29].

As in [13] and [9], we use $P_{\theta l}$ as the variable to be varied in the simulations. We take the phase error variance and its square root, the RMS phase error, as our performance criteria. A carrier frequency $f_c = \omega_c/2\pi = 10\,000$ Hz was used in the simulation, and the value of r was varied in order to achieve the desired values of $P_{\theta l}$ (see (45)).

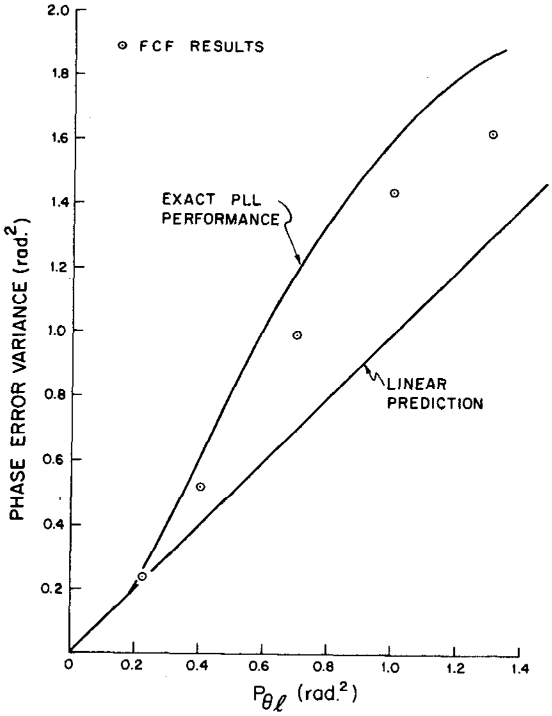

Fig. 3. Phase error variance results.

TABLE II
PHASE TRACKING SIMULATION PERFORMANCE SUMMARY—$\mathscr{E}[1 - \cos(\theta - \hat{\theta})]$

$P_{\theta\ell}$	PLL	SDNF	FCF
0.041	0.0217	0.0212	0.0202
0.130	0.0608	0.0601	0.0581
0.225	0.1171	0.1178	0.1120
0.4	0.2447	0.2335	0.2253
0.7	0.4098	0.3728	0.3687
1.0	0.5312	0.5040	0.5002
1.3	0.6158	0.5719	0.5692

The PLL, the Gustafson–Speyer "state-dependent noise filter" (SDNF), and the Fourier coefficient filter (FCF) (47)–(49) were all simulated using identical noise sequences to allow direct comparison. Table I contains the performances (as measured by RMS Phase Error (RMSPE)) of the various filters, Fig. 3 graphically displays the phase error variance for the PLL and FCF, and filter performance as measured by the criterion $\mathscr{E}(1 - \cos(\theta - \hat{\theta}))$ is reported in Table II. We note that the phase error density for the PLL tracking system has been analytically determined (using the baseband PLL model) [13], and thus our simulation of the PLL system provides a check on the validity of the overall simulation.

We make several comments on the simulation results. First of all, we note that the FCF performed consistently better than the other systems, although the SDNF per-

TABLE III
SUMMARY OF RESULTS FOR THE FCF WITHOUT WONG–ZAKAI CORRECTION TERMS

$P_{\theta\ell}$	RMSPE	$\mathscr{E}[1-\cos(\theta-\hat{\theta})]$
0.041	16.4	0.0396
0.4	44.9	0.2534
1.3	73.6	0.5774

formance is quite close. One interesting point involves the above-threshold performance. As mentioned earlier, the PLL is optimal, with respect to the minimum variance criterion, above threshold (see [13]), while, as discussed in [9], the SDNF is *not* optimal above threshold. The simulation results obtained indicate that the FCF may be optimal above threshold. The proof of this is an open question; however, an intuitive argument can be made that above threshold, the phase density looks like a folded normal, and thus the approximation used in the FCF is, in fact, very nearly correct. If this is so, the FCF should perform optimally or very nearly optimally. (We note that the FCF is designed using the criterion $\mathscr{E}[1 - \cos(\theta - \hat{\theta})]$, but for small phase errors and θ, a folded normal random variable $\hat{\theta}$ for the FCF will also minimize $\mathscr{E}[(\theta - \hat{\theta})^2]$). A related idea is that as $P_{\theta l}$ increases, the phase density looks less and less like a normal density, and thus the FCF approximation is not quite as good. Therefore, filter performance below threshold may be improved by including more Fourier coefficients in the FCF. Finally, we refer the reader to Table III in which we present the results of several simulation runs of the FCF *without* the Wong–Zakai correction terms. Note that the performance is somewhat worse, more so above threshold where the filter works best.

V. CONCLUSION

In this series of two papers we have studied a large class of discrete- and continuous-time phase-tracking and demodulation problems. By using Fourier series techniques, we have been able to uncover the inherent structure in these problems and have displayed infinite-dimensional optimal estimation equations.

The question of determining suitable finite-dimensional approximations to these equations has been considered at some length. One of these techniques was applied to a phase-tracking problem of importance in some navigation systems, and the performance of a rather simple approximate filter, one that keeps only the first Fourier coefficients, compares quite favorably to that of the optimal phase-lock loop and the optimal linear tracking filter.

REFERENCES

[1] A. S. Willsky, "Fourier series and estimation on the circle with applications to synchronous communication—Part I: Analysis," this issue, pp. 577–583.
[2] J. T. Lo and A. S. Willsky, "Estimation for rotational processes with one degree of freedom—I: Introduction and continuous time processes," to appear.
[3] A. S. Willsky and J. T. Lo, "Estimation for rotational processes with one degree of freedom—II: Discrete time processes," to appear.
[4] ——, "Estimation for rotational processes with one degree of freedom—III: Applications and implementation," to appear.
[5] A. S. Willsky, "Dynamical systems defined on groups: Structural properties and estimation," Ph.D. dissertation, Dep. Aeronautics and Astronautics, M.I.T., Cambridge, Mass., June 1973.
[6] J. T. Lo and A. S. Willsky, "Stochastic control of rotational processes with one degree of freedom," to appear.
[7] R. S. Bucy, C. Hecht, and K. D. Senne, "New methods for non-linear filtering," *Revue Francaise d'Automatique, Informatique et de Recherche Operationelle*, pp. 3–54, Feb. 1973.
[8] A. J. Mallinckrodt, R. S. Bucy, and S. Y. Chang, "Final project report for a design study for an optimal non-linear receiver/demodulator," Goddard Space Flight Center, Greenbelt, Md., NASA Contract NAS5-10789, 1970.
[9] D. E. Gustafson and J. L. Speyer, "Linear and asymptotic minimum variance filters applied to phase-lock loops," to appear.
[10] R. W. Brockett, "Lie theory and control systems on spheres," *SIAM J. Appl. Math.*, vol. 25, pp. 213–225, Sept. 1973.
[11] F. Perrin, "Étude mathématique du mouvement Brownien de rotation," *Ann. Ecole Norm. Sup.*, vol. 45, pp. 1–51, 1928.
[12] P. Lévy, "L'addition des variables aléatoires définies sur une circonférence," *Bull. Soc. Math. Fr.*, vol. 67, pp. 1–41, 1939.
[13] H. L. Van Trees, *Detection, Estimation, and Modulation Theory—Part II: Nonlinear Modulation Theory.* New York: Wiley, 1971.
[14] B. W. Licht, "Approximations in optimal nonlinear filtering," Systems Research Center, Case Western Reserve Univ., Cleveland, Ohio, Rep. no. SRC 70-1, Apr. 1970.
[15] J. A. Pierce, W. Palmer, A. D. Watt, and R. H. Woodward, "Omega—A world-wide navigational system specification and implementation—Second revision," Pickard and Burns Electronics, Waltham, Mass., May 1966.
[16] E. Wong and M. Zakai, "On the relation between ordinary and stochastic differential equations," *Int. J. Eng. Sci.*, vol. 3, pp. 213–229, 1965.
[17] ——, "On the convergence of ordinary integrals to stochastic integrals," *Ann. Math. Statist.*, vol. 36, pp. 1560–1564, 1965.
[18] H. L. Van Trees, *Detection, Estimation, and Modulation Theory—Part I: Detection, Estimation, and Linear Modulation Theory.* New York: Wiley, 1968.
[19] A. Kufner and J. Kadlec, *Fourier Series*, G. A. Toombs, transl. London: Iliffe Books, 1971.
[20] H. J. Kushner, "Approximations to optimal nonlinear filters," *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 546–556, Oct. 1967.
[21] R. W. Hamming, *Numerical Methods for Scientists and Engineers.* New York: McGraw-Hill, 1962.
[22] D. J. Wilde, *Optimum Seeking Methods.* Englewood Cliffs, N.J.: Prentice-Hall, 1964.
[23] J. T. Lo, "Finite-dimensional sensor orbits and optimal nonlinear filtering," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 583–588, Sept. 1972.
[24] R. L. Stratonovich, *Topics in the Theory of Random Noise*, vol. I, R. A. Silverman, transl. New York: Gordon and Breach, 1963.
[25] A. H. Jazwinski, *Stochastic Processes and Filtering Theory.* New York: Academic, 1970.
[26] T. Nakamizo, "On the state estimation for non-linear dynamical systems," *Int. J. Contr.*, vol. 11, no. 4, pp. 683–695, 1970.
[27] D. J. Wright, "The digital simulation of stochastic differential equations," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 75–76, Feb. 1974.
[28] E. J. McShane, "Stochastic differential equations and models of random processes," in *Proc. 6th Berkeley Symp.*, 1971.
[29] R. E. Mortenson, "Mathematical problems of modeling stochastic nonlinear dynamic systems," *J. Statistical Physics*, vol. 1, no. 2, pp. 271–296, 1969.