Frequency Line Tracking, Extended Kalman Filters and Some HMM Problems[1]

Brian D O Anderson[2]

Ben James[2]

Robert C Williamson[2]

December 1991

[1] Contribution to Workshop on Hidden Markov Models for Tracking, Wirrina Cove, South Australia, February 1992.

[2] Department of Systems Engineering and Centre for Robust and Adaptive Systems, Research School of Physical Sciences, Australian National University, GPO Box 4, Canberra, ACT 2601, Australia.

Abstract

We review the use of Kalman filter and HMM techniques for frequency tracking (single-tone and multiharmonic cases). A possible new HMM structure is suggested and issues related to thresholding phenomena are discussed.

1 Introduction

The purpose of this paper is to raise some speculative issues concerning Hidden Markov Model filters. The vehicle for doing this involves consideration of some frequency tracking problems. In such problems, the received signal comprises an underlying sinusoid, sum of sinusoids or sum of harmonically related sinusoids, together with additive measurement noise. The frequency tracking problem (single tone, multi-tone and multiharmonic) is the problem of estimating the underlying frequency or frequencies, which may be slowly varying. In the single-tone case, the frequency tracking problem can more or less be viewed as FM demodulation, about which much is known.

As described in Section 2, single-tone and multiharmonic frequency tracking problems can both be addressed by Kalman filtering theory, or in the constant parameter case, efficacious maximum likelihood estimation algorithms. Threshold phenomena are observed.

In Section 3, the HMM approach to these problems is described. Two methods are available in the literature, and a third is proposed which appears more closely tied to the Kalman filtering theory.

In Section 4, there is a brief discussion concerning threshold phenomena in HMM filtering.

2 Frequency Tracking — Continuous State-Space Filters

FM detection [VA1] provides a simple example of a frequency tracking problem with a single tone:

$$z(t) = A\cos(\omega_0 t + \beta \int_{t_0}^{t} a(\tau)d\tau) + n(t) \tag{2.1}$$

Here, A is typically constant and known (though not always so), $n(t)$ is white measurement noise of covariance $N_0\delta(t-s)$, and the instantaneous frequency $\omega_0 + \beta a(t)$ often is subject to certain assumptions, for example

1) $\beta a(t) = \omega_1$, where ω_1 is unknown while ω_o is known. Thus in (2.1), there is a constant unknown frequency and phase. This is the single-tone problem with constant frequency.

2) $\beta a(t) = \omega_1 + \omega_2 t$, with at least one of ω_1 and ω_2 unknown while ω_0 is known; thus the frequency has a ramp variation

3) $a(t)$ is the output of a low bandwidth linear system excited by white noise, e.g. $\dot{a}+\gamma a = \lambda\xi(t)$, for some unit variance white noise $\xi(t)$ with $\gamma > 0$, $\lambda \neq 0$; this is the single-tone problem with slowly varying frequency.

Classical [VA1] or modern (i.e. Kalman filter based) [AN1,SN1] approaches lead to the phase locked loop as an estimator of $a(t)$, or the instantaneous frequency, see Figure 2.1 for example. The figure depicts the case when A is known and a variant must be used when A is also to be estimated; the values for ρ and τ in Figure 2.1c are not here particularly important.

An alternative approach to single-tone tracking is to postulate that over intervals of a certain length, T say, the frequency of the received signal can be assumed to be constant. If a discrete-time model is used, this means one postulates that

$$z(t) = A\sin(\omega_0 t + \theta_0) + n(t) \quad t = 0, 1, 2, \cdots, N-1$$

with A, ω_0 and θ_0 parametrizing the signal, and $n(t)$ a white noise sequence of variance N_0. Usually, ω_0 and θ_0 are unknown, and A may or may not be known, [RI1].

A maximum likelihood estimator is available for the triple ω_0, θ_0 and A (or, if A is known, ω_0 and θ_0). The implementation basically involves Fourier transforming the data. For example, if instead of $z(t)$ one works with the analytic signal

$$Z(t) \triangleq z(t) + \check{z}(t) = A\exp j(\omega_0 t + \theta_0) + (n(t) + j\check{n}(t))$$

as per [RI1], and if one defines

$$\bar{Z}(\omega) = \frac{1}{N}\sum_{t=0}^{N-1} Z(t)\exp(-jt\omega T)$$

where $T = \frac{1}{N}$ then $\hat{\omega} = \arg\max|\bar{Z}(\omega)|$ is the maximum likelihood estimate over the interval — see Figure 2.2. In the next interval, a new maximum likelihood estimate is obtained, and so on. The estimated frequency track is thus piecewise constant.

Threshold phenomena are observed with both these types of estimator. Above the threshold signal to noise ratio:

- the performance of the phase-locked loop is characterizable by analysis using linearization, and no "cycle-skipping" occurs
- the performance of the MLE is characterized by (essentially) non occurrence of outliers (i.e. noise so contributing to $|\bar{Z}(\omega)|$ in Figure 2.2 that the peak occurs well away from ω_0) and by the Cramer-Rao bounds holding with equality.

Below the threshold SNR, in both cases the quality of estimate deteriorates sharply, i.e. a small reduction in SNR can result in a big increase in estimation error variance.

Characterization of the onset of threshold has been available for some time in the PLL approach: it is associated with the phase error variance reaching a value of about 0.25 rad^2 for a particular type of frequency variation [VA1, see p.60]. A characterization has recently became available for the MLE single tone problem, which again involves a critical value of the phase error variance, [JA1].

Since in both cases, the above threshold behaviour, including phase error variance, is computable using the problem data, a priori prediction of whether performance will be above or below threshold is possible.

Now let us turn to the multiharmonic frequency tracking problem, but still using a continuous state-space viewpoint. There is an analog for each of the previous tracking schemes. The first of these is described in [PA1]. The received signal is

$$z(t) = \sum_{k=1}^{m} A_k \sin(k\omega t + \phi_k) + n(t)$$

One can model ω in any of the ways already described, and even postulate a separate (slow) variation in ϕ_k. Also, one can treat the case of known or unknown A_k. Kalman filtering theory leads to a structure involving coupled phase-locked loops, see Figure 2.3 for the case of two harmonics only for an illustration of that part of the receiver structure which estimates ω. The signals $\widehat{\theta}_1, \widehat{\theta}_2$ are estimates of $\omega t + \phi_1$ and $2\omega t + \phi_2$ respectively. If the low pass filters are neglected and the coupling between the loops neglected, the upper loop is like that of Figure 2a with

$$\frac{\beta A}{N_0}[\frac{\rho}{s+\sigma} + \tau] = \frac{k_{12}}{s} + k_{22}$$

For the MLE problem, we assume the received signal is available in discrete time as

$$Z(t) = \sum_{k=1}^{m} A_k \exp[k\omega_0 t + \theta_k^0] + [n(t) + j\check{n}(t)] \quad t = 0, 1, \cdots, N-1$$

The unknowns are the fundamental frequency ω_0, the phases $\theta_k^0, k = 1, \cdots, m$ and possibly the amplitudes A_k, and all these unknowns are assumed constant over the measurement interval. The noise is white.

Once again, the discrete Fourier transform of $Z(t)$ proves highly relevant in generating the desired ML estimates. Because the separate tones are harmonically related, the algorithm is not the same as that of [RI2]. The theory is set out in [BA1], and $\widehat{\omega}$ is defined by

$$\widehat{\omega} = \arg\max \sum_{i=1}^{m} |\bar{Z}(i\omega)|^2$$

(at least for the case when amplitudes and phases are initially unknown, which is typical in the multitone problem).

For both problems, simulation establishes the existence of a threshold phenomenon. Above threshold

- Coupled PLL performance is predictable by Kalman filtering theory, using linearization.

- MLE performance is predictable by Cramer-Rao calculations, the bounds defining the actual performance.
- The MLE and coupled PLL approaches can be harmonised, through making certain specialisations in the signal model used for the PLL, [JA3].
- The quantity $\Sigma k^2 A_k^2 / N_0$ functions as a SNR, [JA2], [JA3].

As for characterizing the threshold itself, rather less is known than for the single tone case; however, for the MLE problem:

- Below threshold performance can be computed, as well as being establishable by simulations, [WI1], the computations are substantial, and not very insightful. Nonetheless, the fact that they can be performed is itself important.
- The threshold SNR level can be characterized approximately by a readily computable quantity; in the single tone case, this is equivalent to the phase-error variance, [JA1]. The quantity computed remains unaltered so long as the ratio of the SNR to the number of measurements, N, is kept constant. This quantity remains roughly invariant at threshold for different values of N as may straightforwardly be checked by successively doubling N while observing a 3dB decrease in (or in absolute terms, a halving of) the SNR value at the point of threshold, (see [JA4]). In the multiharmonic case, the quantity is closely related to each of the harmonic phase error variances.

3 Frequency Tracking — HMM Filters

We begin by discussing the scheme of [ST1] for the single-tone problem. The range of possible frequencies for a single tone is termed a gate; the gate is divided into n adjacent cells, with centre frequencies $\tilde{\omega}_i, i = 1, \cdots, n$. A model is postulated in which over intervals of length T, say, the frequency track is assumed to be in the i-th cell or to be absent, and the location of the track over the next time interval is governed by a transition probability from the i-th to the j-th cell (which in turn can be obtained by postulating that frequencies vary as a result of process noise).

The data over one interval length is collected, and subjected to Fourier transformation. The DFT is evaluated at the centre frequency of each cell, and the maximum magnitude, if above a threshold[1], is held to define a "measurement" of the active cell. If the threshold magnitude is not reached, the "measurement" is that no signal is present. Knowing the actual signal to noise ratio allows calculation of the probabilities of making correct and incorrect measurements. When these probabilities, the cell-to-cell transition probabilities and the "measurements" are put together, track estimates over a number of intervals (which may involve initiation and termination) can be found. [A track estimate over a number of intervals is simply an estimate of the series of cells through which the frequency passes.]

[1]This threshold is quite distinct from the "threshold phenomena" discussed in the previous section.

Some key advantages of this arrangement are

- The easy way in which track initiation and termination are included.
- The fact that smoothed as opposed to filtered estimates (fixed-interval, or fixed-lag if desired) can be obtained.
- The attractive computational burden.
- The possibility of training using standard HMM techniques, to improve estimates of the underlying probabilities.

However, on the other side, let us note that

- The transition and measurement probabilities are derived effectively on the assumption that the actual tracks are piecewise constant, which is not the case at all.
- All phase information is being thrown away.

For the resolution of the first point, one could examine the MLE calculations of Rife and Boorstyn [RI1] with a view to assessing the damage caused by a varying frequency in an interval. Also, as an additional or alternative approach, one could use the training facility of HMM to improve the probability estimates.

A step towards resolution of the second point is described below.

What of the multitone harmonic tracking problem?

The same approach as above could be contemplated. The key variations would arise firstly in the algorithm which took measurements over one time interval and determined a frequency estimate (the "measurement"), and secondly in calculating the probabilities of making correct and incorrect measurements in one interval. Here, a number of the ideas of [WI1] may be relevant — these calculations are far from easy.

We now discuss an improvement [BA2] of the above arrangement which is linked to a modified concept of "measurement". For each interval of time-series data, and each cell, the amplitude and phase of the DFT evaluated at the centre frequency of the cell are obtained; complete specification of these quantities in all gate cells over *two adjacent* intervals is now deemed a "measurement". If one assumes that the signal frequency is actually located in the same cell in two successive time intervals, the phase information allows comparatively accurate estimation of the frequency, i.e. one estimates the frequency not via the centre frequency of the cell, but by this centre frequency corrected with the aid of the phase data. This is based on work in [McM1, McM2].

Apart from improved frequency estimation within a cell, there is a change in the measurement probability matrix. Previously, the measurement probability matrix was defined by the probabilities of incorrect or correct measurements; i.e. for each i, the collection $Pr\{$"measurements" $=$

cell i | true frequency in cell j}. Under this new scheme, one needs as a measurement likelihood function {Probability of signal DFT amplitudes R_k, R_{k+1} in interval I_k and I_{k+1} and phase difference $\phi_{k+1} - \phi_k$ given true amplitude A and frequency $\tilde{\omega}$}.

There is a third possible HMM scheme which could be used in the single or multi-tone case, which so far seems not to have been used. In [ST2], a description of a conventional Kalman filter has been given as an *infinite state* hidden Markov model. The underlying idea of [ST2] could be carried over to the problem of single-tone (or multitone) estimation in the following way.

Consider the discrete-time signal model

$$\begin{aligned} z(kT_s) &= A\cos(\omega_o kT_s + \beta \sum_0^{kT_s} a(jT_s)) + n(kT_s) \\ a(\overline{j+1}T_s) &= \alpha a(jT_s) + \xi(jT_s) \end{aligned}$$

which can be rewritten, introducing a two dimensional state vector $[a(kT_s), \theta(kT_s)]$ where $\theta(kT_s)$ is the sum of the previous $a(jT_s)$:

$$\begin{aligned} z(kT_s) &= A\cos[\theta(kT_s)] + n(kT_s) \\ a(\overline{k+1}T_s) &= \alpha a(kT_s) + \xi(kT_s) \\ \theta(kT_s) &= \theta(\overline{k-1}T_s) + \omega_0 T_s + a(kT_s) \end{aligned}$$

This is a discretized version of the scheme considered in Section 2. For convenience, assume A is known.

The (a, θ) space is gridded, and in this way the $[a(kT_s), \theta(kT_s)]$ process is approximated by a stationary finite-state Markov process, with transition probabilities computable from the above equations, the grid points and the statistics of $\xi(\cdot)$. The measurements are also discretized and a measurement probability matrix $Pr\{z(kT_s) = i \mid \theta(kT_s) = j\}$ can be determined.

Now the tracking problem is an HMM problem, albeit one in what may be a high dimensional space — 200 or 300 values of θ may be needed, as well as a number of values for a, the estimation of which yields the desired frequency estimate.

As with any HMM scheme, probability vectors can be obtained, smoothing is relatively easy, and training is possible.

Why might this procedure be preferable to a phase-locked loop? The phased-locked loop is, roughly speaking, an approximate conditional probability computer, and so is an HMM receiver. The former approximates via a gaussian density, the latter by a discrete collection of delta functions. Since these approximations are different, different performance is to be expected, particularly below threshold.

The multitone versions could be of much higher dimension again, though the transition probability matrix will be sparse. The computational burden has not been investigated. In general, one would hope that the structure of the HMM estimator *would parallel that of the associated extended Kalman filter*, and this is the coupled phase-locked loop structure. Thus one might discern the equivalent of several HMM single tone estimators with some cross-coupling.

This loosely sketched scheme, related perhaps more closely to a signal model with truly continuously varying frequency than either of the first two schemes (which presume piecewise constant frequency) has yet to be evaluated, but would seem promising.

4 A General HMM Issue

Some estimation problems naturally involve a discrete state-space, and discrete values of measurements. Others are only presented this way after some "restructuring". Frequency tracking is certainly an example of the latter, since frequency is naturally a continuous variable, and noisy sine-wave measurements equally so.

Problems where the state and measurement spaces are continuous can often be tackled by extended Kalman filter ideas, [AN1]. The EKF is derived using linearization assumptions, the violation of which is typically associated with a threshold-type phenomenon. Since the HMM filter may well be a discrete state, discrete measurement approximation of the EKF, then questions like the following arise: are threshold phenomena well recognized in HMM applications? Is anything systematic known about them? Since an HMM does not rely on linearization in the way an EKF does, might threshold occur for lower SNRs than for an associated EKF?

References

[AN1] B D O Anderson and J B Moore, *Optimal Filtering*, Prentice Hall, Inc., 1981.

[BA1] R F Barrett and D R A McMahon, "ML estimation of the fundamental frequency of a harmonic series", *Proc ISSPA 87*, Brisbane, 1987, pp. 333-336.

[BA2] R F Barrett and R L Streit, "Frequency line tracking using hidden Markov models with phase information", *Proc ISSPA 90*, Gold Coast, Australia, 1990, pp. 243-236.

[JA1] B James, B D O Anderson and R C Williamson, "Characterization of threshold for single tone maximum likelihood frequency estimation", submitted to *IEEE Trans Sig Proc*, 1991.

[JA2] B James and B D O Anderson, "The amplitude, phase and frequency estimation of multi-harmonic signals in noise — an investigation of the general phase-frequency estimator", *Proc ISSPA 90*, Gold coast, Australia, 1990, pp. 141-146.

[JA3] B James, B D O Anderson and R C Williamson, "Conditional mean and maximum likelihood approaches to multiharmonic frequency estimation", submitted to *IEEE Trans Sig Proc*, 1991.

[JA4] B James, B D O Anderson and R C Williamson, "Characterization of threshold for multiharmonic maximum likelihood frequency estimation", in preparation.

[McM1] D R A McMahon and R F Barrett, "An efficient method for the estimation of the frequency of a single tone in noise from the phases of discrete Fourier transform", *Signal Processing*, Vol. 11, 1986, pp. 169-177.

[McM2] D R A McMahon and R F Barrett, "Generalization of the method for the estimation of the frequencies of tones in noise from the phases of discrete Fourier transforms", *Signal Processing*, Vol. 12, 1987, pp. 371-383.

[PA1] P J Parker and B D O Anderson, "Frequency tracking of nonsinusoidal periodic signals in noise", *Signal Processing*, Vol. 20, 1990, pp. 127-152.

[RI1] D C Rife and R R Boorstyn, "Single-tone parameter estimation from discrete-time observations", *IEEE Trans. Info. Theory*, Vol. It-20, 1974, pp. 591-598.

[RI2] D C Rife and R R Boorstyn, "Multiple-tone parameter estimation from discrete-time observations", *Bell Syst Tech J*, Vol. 55, 1976, pp. 1389-1410.

[SN1] D L Snyder, *The State Variable Approach to Continuous Estimation*, MIT Press, Cambridge, Mass, 1969.

[ST1] R L Streit and R F Barrett, "Frequency line tracking using hidden Markov models", *IEEE Trans. ASSP*, Vol. 38, 1990, pp. 586-598.

[ST2] R L Streit, "The Kalman filter is an infinite-state hidden Markov model", unpublished manuscript.

[VA1] H L Van Trees, *Detection, Estimation and Modulation Theory*, John Wiley, New York, 1971.

[WI1] R C Williamson, B D O Anderson and B James, "Threshold effects in maximum likelihood multiharmonic frequency estimation", in preparation.

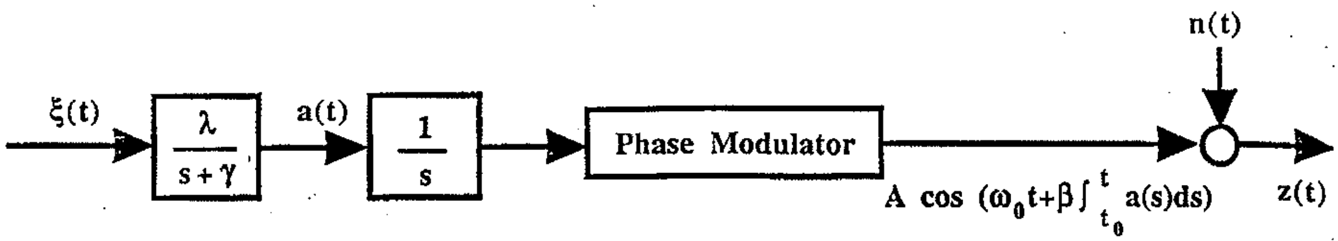

Figure 2.1a Signal model typical for FM

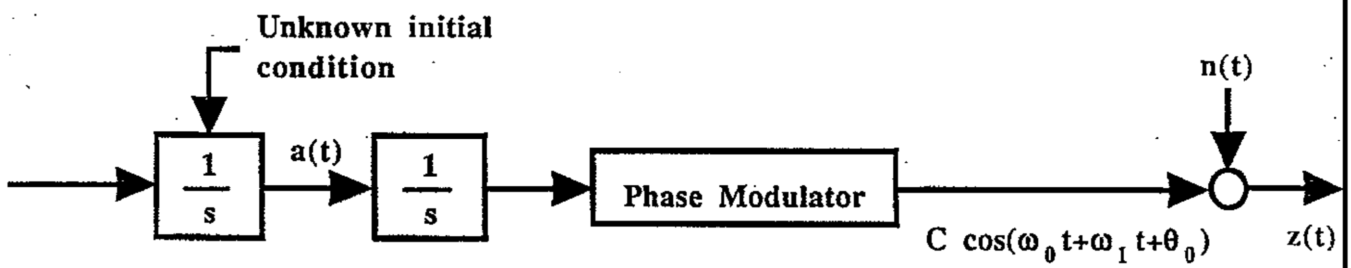

Figure 2.1b Signal model for constant frequency tone estimation

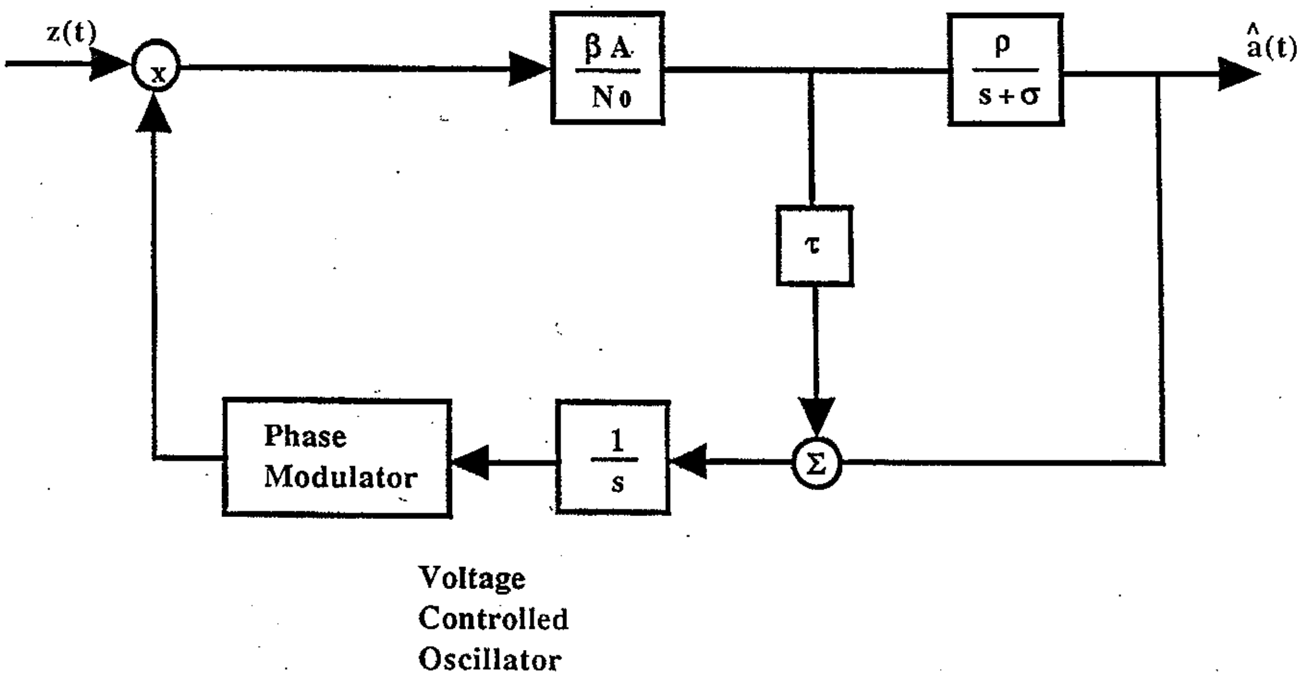

Figure 2.1c Phase locked loop receiver

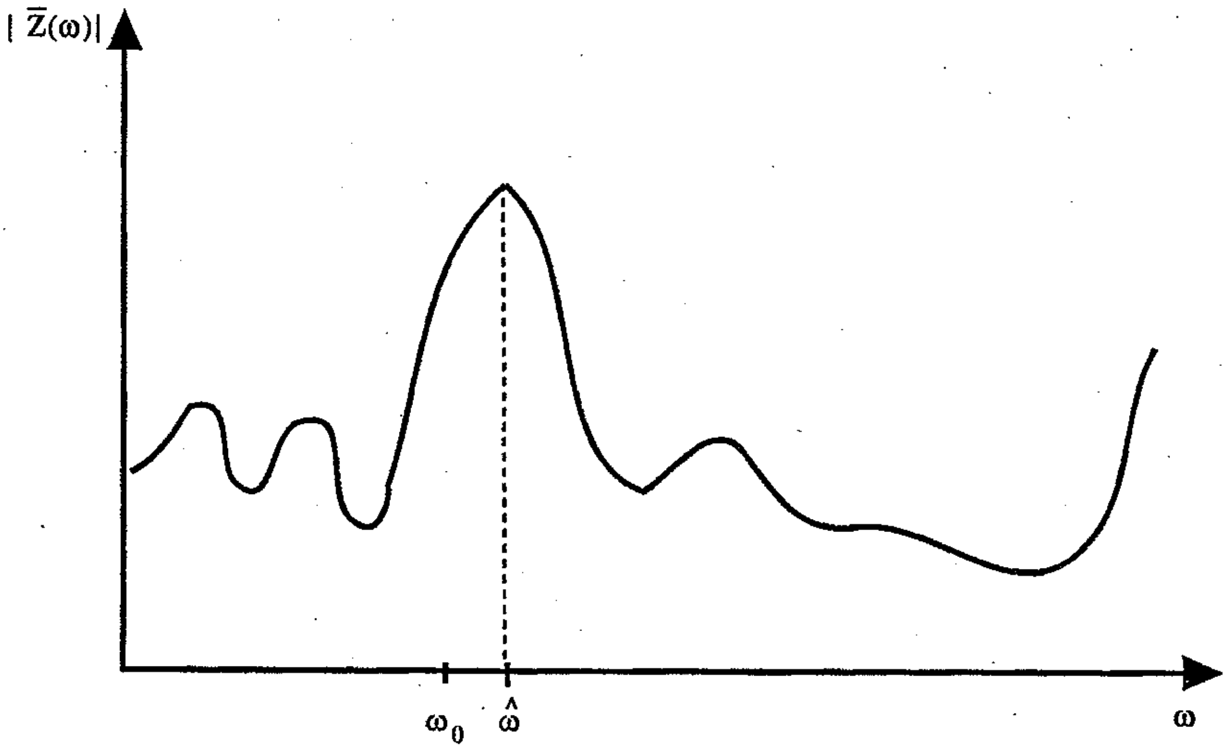

Figure 2.2 Typical plot of $|\bar{Z}(\omega)|$ without outlier interference

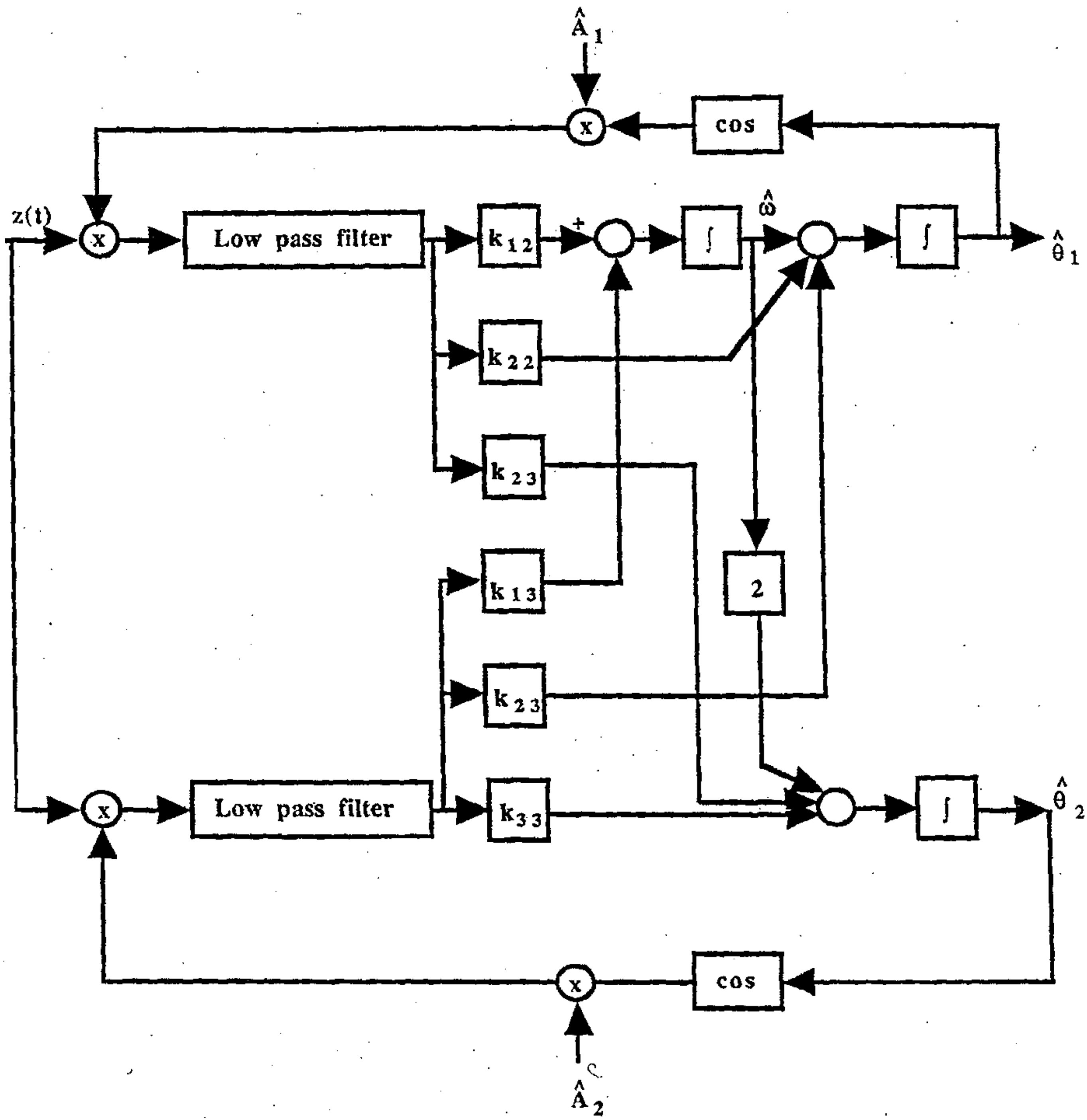

Figure 2.3 Estimator of frequency for periodic signal with fundamental plus one harmonic