ESTIMATION OF THE DYNAMICS OF A MOVING SOURCE BY AN INTEGRATED SPATIAL AND TEMPORAL PROCESSING

by

JOSÉ MANUEL FONSECA DE MOURA

Eng. Elect., Instituto Superior Técnico,
Universidade de Lisboa
(1970)

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

and

ELECTRICAL ENGINEER

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February, 1973

Signature of Author
Department of Electrical Engineering, September 15, 1972

Certified by
(Harry L. Van Trees, Thesis Supervisor)

Accepted by
Chairman, Departmental Committee on Graduate Students

ESTIMATION OF THE DYNAMICS OF A MOVING SOURCE BY AN INTEGRATED SPATIAL AND TEMPORAL PROCESSING

by

José Manuel Fonseca de Moura

Submitted to the Department of Electrical Engineering on September 15, 1972 in partial fulfillment of the requirements for the Degree of Master of Science and Electrical Engineer.

ABSTRACT

This thesis considers from a new standpoint the problem of tracking a moving source by a passive observer.

The changes induced by the source dynamics in the spatial and temporal structure of the emitted narrow-band signal are processed by a spatially and temporally coupled receiver which simultaneously estimates the range, the bearing and their time derivatives.

Unlike the classic "bearings only" problem, we model the bearing and range waveforms as finite-state-dimension stochastic processes.

Depending on the particular choice of the coordinate frame, we are led to different system representations. The two frameworks here considered, polar and cartesian, both led to a nonlinear estimation problem. The Nonlinear Filtering Theory for lumped-state systems is then applied to process the received signal.

Two linear approximations of the (∞-dimensional) optimal processor are considered: the Extended Kalman filter and the Maximum A Posteriori filter.

A mathematical analysis shows that the (sub)optimal receivers essentially perform two beams which are the inputs to two phase lock loops tracking respectively the bearing and range waveforms. We derive bounds on the expected performance of these channels.

After formulating the discrete version of the continuous nonlinear estimation problem extensive simulation studies are carried out in a digital computer in order to assess how the filter performance is affected by the geometric configurations and the several parameters (signal to noise ratio, driving noise level, array length). The simulation results confirm the main analytical conclusions of Chapter V and show the existence of a threshold for the signal to noise ratio.

When using the polar frame, to overcome numerical innaccuracies which make the propagated error covariance to lose its positive semidefiniteness character, we implemented the square root algorithm which improved the overall filters performance.

Finally to initialize the (Kalman) filters some a priori information is needed. A bearing maximum likelihood estimation followed by a triangularization procedure solves for the starting point of the filters.

THESIS SUPERVISOR: Harry L. Van Trees

TITLE: Professor of Electrical Engineering

ACKNOWLEDGEMENTS

I want to thank my advisor Professor Harry Van Trees for his guidance and valuable advice during the course of this thesis.

Part of this work used the facilities of the Information Processing Center at the Massachusetts Institute of Technology.

I was supported by a fellowship of the Instituto de Alta Cultura, Portugal. I sincerely acknowledge this financial support.

The discussions held with my officemate José Paulo Albuquerque were very useful.

I thank Miss Diane Horner for typing the whole manuscript.

Tereza who with her own burden of study always found time to devote to me and to Bárbara, and Bárbara with her cheerful smile gave me the necessary moral strength to accomplish this work.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

Symbols

B1, B2, B3	Covariance Bounds
c	Medium Wave Propagation Velocity
$E(.\|.)$	Expectation Operator
$\underline{f}(t, \underline{x}(t))$	System Matrix Function
$\underline{g}(t, \underline{x}(t))$	Driving Matrix Coefficient
$\underline{h}(t, \underline{x}(t))$	Signal Process
$\underline{M}(t)$	Time Update Error Covariance Matrix
$N_o/2$	Measurement Noise Power Level
N	Number of Array Elements
ϕ_N	Ambiguity Function
$\underline{p} = \begin{bmatrix} p_1 \\ \cdot \\ \cdot \\ p_N \end{bmatrix}$	Array Position Vector
P	Signal Power
$\underline{P}$	Error Covariance Matrix
P_r	Probability of (Large) Errors
$2PT/N_o$	Equivalent Signal to Noise Ratio
Q	Driving Noise Power Level
$\underline{r}(t)$	Received Vector
R(t)	Range
$\underline{R}(t)$	Measurement Noise Covariance Matrix

Symbols

$\sigma = \Sigma p_i^2$	Array Second Geometric Moment
$\underline{s}(t, \underline{x}(t))$	Signal Vector
$\underline{S}$	Square Root Form of $\underline{M}$
S1, S3	Mathematical Input Signal Components
$\Theta(t)$	Bearing Angle
T	Sampling Rate
$\underline{u}(t) = \begin{bmatrix} u_x \\ u_y \end{bmatrix}$	Driving White Process
$\underline{w}(t) = \begin{bmatrix} w_1 \\ \cdot \\ \cdot \\ w_N \end{bmatrix}$	Measurement White Noise
w_c	Carrier Angular Frequency
$\underline{W}$	Square Root Form of $\underline{P}$
W1, W3	Mathematical Input Noise Components
$\underline{x}(t)$	System State Vector
$\bar{\underline{x}}(t)$	Time Update Estimate
$\hat{\underline{x}}(t)$	Measurement Update Estimate
X_T, Y_T	Source Cartesian Coordinates
Z1, Z3	Mathematical Input Beams

Abbreviations

EKF	Extended Kalman Filter
ILP	Ideal Low Pass Filter
LRT	Likelihood Ratio Test
MAP	Maximum A Posteriori Filter
ML	Maximum Likelihood
SNR	Signal to Noise Ratio
SQRT	Square Root Algorithm
STD	Standard Deviation
VAR	Variance

Chapter I

Introduction

The problem of concern in this work is the study of performance of several processors for the estimation of the position and velocity of a moving source generating a narrow band signal.

A new model is considered, developed by Van Trees [B-1], where a spatial-temporally coupled tracking allows the simultaneous estimation of range, range rate, bearing and bearing rate.

The moving target induces a change on the temporal structure of the narrow band signal and causes a geometric change reflected in its spatial structure. By considering these changes, Van Trees successfully developed a model which utilizes the temporal and spatial effects in an optimally-coupled processor.

In this work we consider the case of a stationary observer. Although this may not be the best strategy, the main theoretical questions arise and more easily can be dealt with in this simpler context. The main issues we consider relate to the choice of the coordinate systems (polar versus rectangular), of the nonlinear processor and finding regions of convergence as well as range of the important parameters (e.g. signal to noise ratio, driving noise power level, etc.) for which the filters converge and exhibit an acceptable performance.

As for the organization of this report we present in Chapter 2 the development of the model leading to the dynamical system (state variable model) representing the target's motion and establish the observed waveform structure in both polar and rectangular frameworks.

In Chapter 3 we set up the general nonlinear estimation problem within the framework of the lumpen - state variable theory. After a brief discussion we are led for reasons indicated in the chapter, to the choice of two first order approximations of the optimal processor, the Extended Kalman filter (EKF) and the Maximum a Posteriori (MAP) filter.

In Chapter 4 we reformulate the problem in its discrete version and derive the actual implementation of the several receivers. For reasons that will become apparent later, we develop in section 4.5 the square root algorithm implementing the EKF and the MAP filter.

In Chapter 5 we present a mathematical model of the filter and analyse its behavior. We conclude that essentially the filter's action is equivalent to a two beam forming operations followed by a 4-dimensional system, the receiver's linearized "copy" of the dynamical system, where we are able to distinguish two phase-lock loops one essentially tracking the range and the second locking the bearing angle.

In Chapter 6 we present the simulation results and compare them with the mathematical analysis of Chapter 5.

In Chapter 7 we resort to a maximum likelihood estimation procedure to initialize the Kalman filter.

Finally, Chapter 8 recaptures the main conclusions drawn from this work and presents the pertinent bibliography.

Chapter II

Model

In this section we intend to suitably characterize by a stochastic dynamical system the motion of the moving source and then develop the assumed temporal and spatial structure of the received signal. For simplicity we will consider a discrete linear array and a planar geometry, i.e. that the target moves on the plane defined by the linear observer and the target's position at time t = o. We further assume the far-field hypothesis, i.e. that the observer is sufficiently far from the source so that the receiving waveforms can be considered as planar.

As pointed out earlier this model although not considering all the intricacies of the real physical problem is sufficiently complex to give a first account on the main theoretical issues arising in the real situation and allows us with enough information to draw important conclusions about the expected performance of the designed processors.

We first establish the quantitative aspects of the target motion and observer geometry. Then we proceed by choosing a temporal model for the emitted signal and discuss its integrated spatial and temporal structure at the receiver.

2.1 Geometry

Figure 2-1 gives a sketch of the plane where the target motion evolves and the observer is fixed. We choose the coordinate system, as indicated in the figure, with the x-axis pointing east, along the linear array, and the y-axis pointing north.

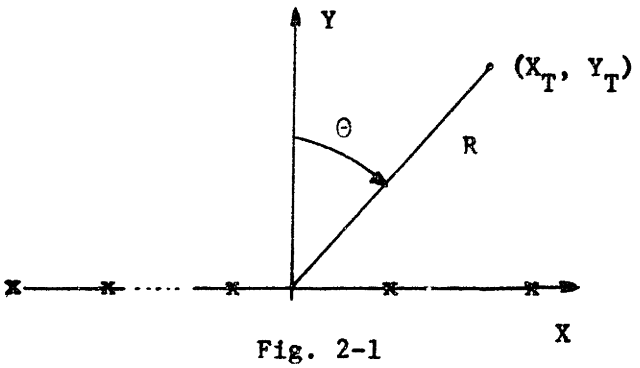


Fig. 2-1

The observer has N sensors distributed along its main axis (X-axis) with positions given by the position vector

$$\underline{p} = \begin{bmatrix} p_1 \\ \vdots \\ p_N \end{bmatrix} \tag{2-1}$$

The target location is given in rectangular coordinates by (X_T, Y_T) and in polar coordinates by (R, Θ) with the usual relations:

$$R = \sqrt{X_T^2 + Y_T^2} \tag{2-2}$$

$$\Theta = \tan^{-1}(X_T/Y_T) \tag{2-3}$$

2.2 Dynamical System

We consider that the source moves with a nominal velocity which is perturbed by random acceleration components along the X and Y directions.

In the classical analysis the target's velocity is considered constant, so that this model allows much more freedom and is more realistic. In the rectangular frame we define the 4-state vector

$$\begin{cases} x_1 = X_T \\ x_2 = \dot{X}_T \\ x_3 = Y_T \\ x_4 = \dot{Y}_T \end{cases} \tag{2-4}$$

and the source motion is governed by the linear dynamical system

$$\begin{cases} \dot{\underline{x}}(t) = \underline{F}\,\underline{x}(t) + \underline{G}\,\underline{u}(t) \\ \underline{x}(o) = \underline{x}_o \end{cases} \tag{2-5}$$

where

$$\underline{F} = \begin{bmatrix} o & 1 & o & o \\ o & o & o & o \\ o & o & o & 1 \\ o & o & o & o \end{bmatrix}$$

$$\underline{G} = \begin{bmatrix} o & o \\ 1 & o \\ o & o \\ o & 1 \end{bmatrix} \tag{2-6}$$

$$\underline{u}(t) = \begin{bmatrix} u_x(t) \\ u_y(t) \end{bmatrix}$$

We model the driving term $\underline{u}(t)$ as a zero mean white Gaussian noise with covariance matrix

$$\text{Cov}\,\{\underline{u}(t),\, \underline{u}(\tau)\} = E\,\{\underline{u}(t)\,\underline{u}^T(\tau)\} = \underline{Q}\,\delta(t - \tau) \tag{2-7}$$

We assume known the statistics of $\underline{x}_o$ with:

$$\begin{cases} E\,\underline{x}_o = \hat{\underline{x}}_o \\ \text{Cov}\,\{\underline{x}_o,\, \underline{x}_o\} = \underline{P}_o \end{cases} \tag{2-8}$$

and furthermore

$$E\, x_i(o)\, u_\ell(t) = o \qquad \begin{matrix} i = 1, \ldots, 4 \\ \ell = x,y \\ t \geq o \end{matrix} \tag{2-9}$$

We see that we have modeled the target's acceleration components

$$\begin{cases} \dot{x}_2 = \ddot{X}_T = u_x \\ \dot{x}_4 = \ddot{Y}_T = u_y \end{cases} \tag{2-10}$$

by a sample function of a (mathematical) white noise vector process.

More complex situations could be handled by augmenting the state-vector with for example $\dddot{X}_T$ and $\dddot{Y}_T$. Also we could, for example, have modeled the accelerations as a one pole stochastic vector process. The range and bearing can easily be computed through the nonlinear relation (2-2) and (2-3).

In the polar coordinates geometry, as fixed in Fig. 2-1, we define the state vector as

$$\begin{cases} x_1 = R \\ x_2 = \dot{R} \\ x_3 = \Theta \\ x_4 = \dot{\Theta} \end{cases} \tag{2-11}$$

and the dynamical system is obtained by translating into polar coordinates the velocity and acceleration of the source. We obtain

$$\begin{cases} \dot{\underline{x}} = \underline{f}(\underline{x}(t)) + \underline{g}(\underline{x}(t))\, \underline{u}(t) \\ \underline{x}(o) = \underline{x}_o \end{cases} \tag{2-12}$$

where

$$\underline{f}(\underline{x}(t)) = \begin{bmatrix} x_2(t) \\ x_1(t)\ x_4^2(t) \\ x_4(t) \\ -2x_2(t)\ x_4(t)/x_1(t) \end{bmatrix} \tag{2-13}$$

and

$$\underline{g}(\underline{x}(t)) = \begin{bmatrix} 0 & 0 \\ \sin x_3(t) & \cos x_3(t) \\ 0 & 0 \\ \cos x_3(t)/x_1(t) & -\sin x_3(t)/x_1(t) \end{bmatrix} \tag{2-14}$$

and $\underline{u}(t)$ and $\underline{x}_o$ have predefined statistics.

We observe that now the range, range rate, bearing and bearing rate are directly available in the state-vector, and so no nonlinear transformation is needed to get back these quantities. But we ended up with a nonlinear dynamic system, where, the driving term matrix coefficient is a nonlinear function of the state vector and so a random matrix.

Warning

Since the dynamical systems (2-5) and (2-12) are stochastic differential equations (2-5) and (2-12) are merely formal descriptions of the model.

More properly we should have written the Langevin equations

$$\begin{cases} d\underline{x}(t) = \underline{F}\underline{x}(t)dt + \underline{G}d\underline{\beta}(t) \\ \underline{x}(o) = \underline{x}_o \end{cases} \tag{2-5'}$$

and

$$\begin{cases} d\underline{x}(t) = \underline{f}(\underline{x}(t),\ t)dt + \underline{g}(\underline{x}(t))\ d\underline{\beta}(t) \\ \underline{x}(o) = \underline{x}_o \end{cases} \tag{2-12'}$$

where $\{\underline{\beta}(t),\ t \geq o\}$ is a random vector Brownian motion (Wiener process) with covariance parameter $\underline{Q}$.

The differentials (2-5') and (2-12') are defined in terms of their integral counterparts:

$$\underline{x}(t) - \underline{x}_o = \int_o^t \underline{F}\underline{x}(t)dt + \int_o^t \underline{G}\, d\underline{\beta}(t) \tag{2-5"}$$

and

$$\underline{x}(t) - \underline{x}_o = \int_o^t \underline{f}(\underline{x}(t))dt + \int_o^t \underline{g}(\underline{x}(t))\, d\underline{\beta}(t) \tag{2-12"}$$

Because of the particular form of the matrices $\underline{F}$, $\underline{G}$, $\underline{f}(.)$ and $\underline{g}(.)$ (equations (2-6), (2-13), (2-14)), and the statistical hypothesis assumed the stochastic process $\underline{x}(t)$ is mean square integrable so that

$$\int_o^t \underline{F}\underline{x}(t)dt \tag{2-15}$$

and

$$\int_o^t \underline{f}(\underline{x}(t))dt \tag{2-16}$$

are interpreted as well defined m.s. Riemman integrals for the sample functions.

It remains to interpret the remaining integrals in (2-5") and (2-12"). However, no ambiguity arises in either case as we shall see in a moment.

In (2-5") because $\underline{G}$ is independent of $\underline{\beta}(t)$ the Itô and Stratonovitch formulations give the same answer ((2-5') is a linear system).

As for

$$\int_o^t \underline{g}(\underline{x}(t))\, d\underline{\beta}(t) \tag{2-17}$$

we interpret it in the Itô(I) sense. But if we consider $\underline{g}(\underline{x}(t))$ as an explicit function of $\underline{\beta}(t)$ through

$$\underline{g}(\underline{x}(\underline{\beta}(t))) \tag{2-18}$$

we also could interpret (2-17) as a Stratonovitch (S) integral. Anyway, in this example they both also coincide. In fact for the ith component we would have the transformation, see Jazwinski [A-4],

$$(I) \leftrightarrow (S) \qquad : \qquad \frac{1}{2} \sum_{j=1}^{2} \sum_{k=1}^{4} g_{kj} \, \partial g_{ij} / \partial x_k \tag{2-19}$$

but given the particular form of (2-14)

$$g_{kj} \, \partial g_{ij} / \partial x_k = o \qquad \text{for} \quad \begin{matrix} i = 1, \ldots, 4 \\ k = 1, \ldots, 4 \\ j = 1, 2 \end{matrix} \tag{2-20}$$

so that both interpretations give the same value for (2-17) and we can speak of (2-5) and (2-12) as formal descriptions of (2-5") and (2-12").

Finally we observe that the vector random process $\underline{x}(t)$ as conceptually generated by (2-5") is a Gauss-Markov stochastic process (for Gaussian initial condition $\underline{x}_o$) and that $\underline{x}$ as conceptually generated by (2-12") is in any case a Markov process.

2.3 Structure of the Received Signal

We proceed by first choosing a temporal structure for the signal and then by discussing the temporal and spatial structures of the received signal induced by the target's dynamics.

We assume that the source generates a narrow band signal that we simply model as a sinewave with carrier frequency w_c, i.e.

$$S(t) = \sqrt{2P} \sin w_c t \tag{2-21}$$

where P is the transmitted power assumed constant throughout.

At the reference sensor a delayed version of S(t) is received plus an additive white noise component

$$r_o(t) = s_o(t) + w_o(t) \tag{2-22}$$

where

$$s_o(t) = s(t - \tau_o) \tag{2-23}$$

and τ_o is the delay corresponding to the travel time of the wavefront from the source to the reference element, given by

$$\tau_o(t) = \frac{R(t - \tau_o(t))}{c} \tag{2-24}$$

where c is the medium propagation velocity.

At the ith sensor the received signal is

$$r_i(t) = s_i(t) + w_i(t) \tag{2-25}$$

with

$$\begin{aligned} s_i(t) &= s_o(t + \tau_i(t)) \\ &= \sqrt{2P}\ \sin w_c\left(t + \tau_i - \frac{R(t + \tau_i - \tau_o)}{c}\right) \end{aligned} \tag{2-26}$$

and the relative delay τ_i to the reference element is given by

$$\tau_i = \frac{\sin\Theta(t - \tau_o(t + \tau_i) + \tau_i(t))}{c} p_i \quad * \tag{2-27}$$

We observe that (2-24) and (2-27) are memory functions of the dynamics. In order to apply the lumpen state variable theory we have to approximate (2-24) and (2-27) by no memory functions. The difficulty lies in the interdependence of the range R(t) and bearing angle $\Theta(t)$ on the several delays τ_o, τ_i, i = 1, ... N.

---

* Obviously τ_i can be $\lesseqgtr$ o. For (2-26) we considered + τ_i so that if we choose the reference element to be the geometric center of the array and also the center of the coordinate system the sign convention for the x-axis for the p_i and τ_i, as well as Θ, will all be consistent.

We proceed by considering Taylor's series expansions of $R(t - \tau_o)$, $R(t - \tau_o + \tau_i)$ and $\Theta(t - \tau_o + \tau_i)$ and truncating these expansions to the linear terms

$$R(t - \tau_o) \simeq R(t) - \dot{R}(t)\tau_o \tag{2-28}$$

Substitution of (2-28) in (2-24) leads to

$$\tau_o = \frac{R}{c} \frac{1}{1 + \frac{\dot{R}}{c}} \tag{2-29}$$

for the dynamics we are interested in

$$\frac{\dot{R}}{c} << 1 \tag{2-30}$$

and so (2-29) can in fact be approximated by

$$\tau_o \doteq \frac{R(t)}{c} \tag{2-31}$$

Since we have for the usual arrays and usual geometries

$$\tau_o >> \tau_i \qquad (p_i << R) \tag{2-32}$$

we can approximate (2-27) by

$$\tau_i \simeq \frac{\sin\Theta(t - \tau_o)}{c} p_i \tag{2-33}$$

Expanding in Taylor's series the argument of the sin(.) function and retaining the linear terms

$$\Theta(t - \frac{R}{c}) \simeq \Theta(t) - \frac{R}{c}\dot{\Theta}(t) \tag{2-34}$$

substitution of (2-34) in (2-33) and expanding the sin(.) leads directly to

$$\tau_i \simeq \frac{p_i}{c} \{\sin \Theta(t) \cos(\dot{\Theta}\frac{R}{c}) - \sin(\frac{R}{c}\dot{\Theta}) \cos\Theta(t)\} \tag{2-35}$$

Unless $\Theta = o$, for the usual dynamics we consider, we can further simplify (2-35) to

$$\tau_i \simeq \frac{p_i}{c} \sin\Theta(t) \tag{2-36}$$

For a better feeling of how these approximations might relate to the actual world we present a numerical example.

Let

$R \sim 10^4$ feet

$c \sim 5000$ feet

$\dot{R} \sim 30$ feet

$p_i \sim 250$ feet (array of 500 feet)

$\Theta \sim 10^{-1}$ rad

$\dot{\Theta} \sim 10^{-3}$ rad/sec

then we get

$\tau_o \sim 2$ seconds

$\tau_i \simeq \frac{p_i}{c} \sin\Theta(t - \tau_i) < \frac{1}{20}$ sec

$\frac{R\dot{\Theta}}{c} \sim 2 \times 10^{-3}$ rad $\ll \Theta \sim 10^{-1}$ rad

$\frac{\dot{R}}{c} \sim \frac{3}{500} \ll 1$

Returning to (2-26) we can write

$$s_i(t) = \sqrt{2P} \sin\left[w_c\left(t + \tau_i - \frac{R(t - \tau_o + \tau_i)}{c}\right) \right]$$

$$\doteq \sqrt{2P} \sin\left[w_c\left(t + \tau_i - \frac{1}{c}(R(t) + \dot{R}\tau_i)\right)\right]$$

$$= \sqrt{2P} \sin\left[w_c\left(t - \frac{R}{c} + \tau_i\left(1 - \frac{\dot{R}}{c}\right)\right)\right] \tag{2-37}$$

Once again we observe that

$$\frac{\dot{R}}{c} \ll 1 \tag{2-38}$$

and so approximate

$$s_i(t) \simeq \sqrt{2P} \sin w_c(t - \frac{R}{c} + \tau_i(t)) \tag{2-39}$$

with

$$\tau_i(t) = \frac{p_i}{c} \sin\Theta(t) \tag{2-40}$$

In summary we conclude that at the ith sensor the received signal is

$$\begin{aligned} r_i(t) &= s_i(t) + w_i(t) \\ &= \sqrt{2P} \sin w_c(t - \frac{R(t)}{c} + \frac{p_i}{c} \sin\Theta(t)) + w_i(t) \end{aligned} \tag{2-41}$$

(2-41) is the modeled received waveform. We see that because of the source motion the available signal at the receiver is phase modulated. It is this phase modulation, together with all the other a priori information that allows us to estimate the target dynamics, in particular the range, range rate, bearing and bearing rate.

For completely specifying the structure of the received signal we have to indicate the assumed statistics for the additive noise in (2-41).

In this work we will only assume a white Gaussian disturbance. So we have that

$$\underline{w}(t) = \begin{bmatrix} w_1(t) \\ \vdots \\ w_N(t) \end{bmatrix} \tag{2-42}$$

is a sample function from a zero mean, white Gaussian real random vector, with covariance matrix

$$\text{cov}\{\underline{w}(t), \underline{w}(\tau)\} = E\{\underline{w}(t)\ \underline{w}^T(\tau)\} = \underline{R}(t)\ \delta(t - \tau) \tag{2-43}$$

We further assume that

$$\underline{R}(t) = \frac{N_o}{2} \underline{I} \tag{2-44}$$

i.e. the noises are spatially and temporally uncorrelated. Furthermore, we take

$$E\{w_i(t)\, u_j(\tau)\} = o \qquad \begin{array}{l} \text{all } i = 1, \ldots, N \\ j = x, y, t, \tau \end{array}$$

and (2-45)

$$E\{w_i(t)\, x_\ell(o)\} = o \qquad \begin{array}{l} \text{all } i = 1, \ldots, N \\ \text{all } \ell = 1, \ldots, 4, \text{ all } t \end{array}$$

These restrictive assumptions on the statistics and joint statistics of the several noises and initial conditions are assumed for simplicity. We could relax to the case of coloured noise or spatially correlated noise since there are available several techniques reducing the problem to the one sketched above. We proceed in the simpler context of (2-43) to (2-45).

2.4 Summary

In brief in this section we established the geometry of the problem in paragraph 2-1, the dynamical system modeling the source motion in 2-2 and the temporal and spatial structure of the signal at the receiver.

As for now we have two different characterizations of the dynamical system. In the rectangular frame the system is linear

$$\begin{cases} \dot{\underline{x}}(t) = \underline{F}\, \underline{x}(t) + \underline{G}\, \underline{x}(t) \\ \underline{x}(o) = \underline{x}_o \end{cases} \tag{2-5}$$

and in the polar case we have the nonlinear stochastic dynamical system

$$\begin{cases} \dot{\underline{x}} = \underline{f}(\underline{x}(t)) + \underline{g}(\underline{x}(t))\, \underline{u}(t) \\ \underline{x}(o) = \underline{x}_o \end{cases} \tag{2-12}$$

We modeled the emitted signal as a narrow-band process, for simplicity the sine wave

$$S(t) = \sqrt{2P} \sin w_c t \tag{2-46}$$

and at the ith sensor we derived the structure of the received waveform to be

$$r_i(t) = \sqrt{2P} \sin w_c\left(t - \frac{R(t)}{c} + \frac{p_i}{c}\sin(\Theta(t))\right) + w_i(t) \quad i = 1, .., N \tag{2-47}$$

We define

$$a_i(t) = \frac{w_c}{c}[p_i \sin\Theta(t) - R(t)] \tag{2-48}$$

In the rectangular coordinates we can express (2-48) as

$$a_i^r(t) = \frac{w_c}{c}\left(p_i \frac{x_1}{\sqrt{x_1^2 + x_3^2}} - \sqrt{x_1^2 + x_3^2}\right) \tag{2-49}$$

and in polar coordinates as

$$a_i^p(t) = \frac{w_c}{c}[p_i \sin x_3 - x_1] \tag{2-50}$$

Collecting (2-47) in a N-vector

$$\underline{r}(t) = \underline{s}(\underline{x}(t), t) + \underline{w}(t)$$

with

$$\underline{s}(x(t), t) = \sqrt{2P} \begin{bmatrix} \sin(w_c t + a_1(t)) \\ \cdots \\ \sin(w_c t + a_N(t)) \end{bmatrix} \tag{2-51}$$

with $a_i(t)$ given in the general setting of (2-48).

The statistics of the driving vector $\underline{u}(t)$, the additive disturbance $\underline{w}(t)$ and the initial state were previously defined.

We observe that we did not consider any amplitude fading or phase drift in the received signal, the only disturbance being additive.

We also assume known all the other parameters as the signal power P, the carrier frequency w_c, the noise levels $\frac{N_o}{2}$ and Q. This is a simpler approach taken at this stage in order to explore the essential features of the problem. Van Trees [B-2] considers several generalizations of this model.

What we want is from the available data, i.e. (2-51) and all the other knowledge we have to do an estimation of the range, range rate, bearing and bearing rate which is good in some sense (e.g. MMSE).

In the next chapter we discuss the nonlinear estimation problem in its general context and justify the choice of the optimal (in a given sense) processors we are going to consider in this work.

Chapter III

General Nonlinear Estimation Processor

In the previous chapter we saw how the assumed time structure of the emitted signal and the phase modulation of the signal by the target dynamics led us to an inherently nonlinear observation model. We also saw that in the case of the polar frame the stochastic dynamical system was nonlinear.

In this chapter we want to discuss the solution of the essentially nonlinear estimation problem where strong nonlinearities of the system function and/or of the observation function arise.

In the framework of a dynamical differential equation description the nonlinear estimation problem has essentially been solved by Stratonovitch [C-3], [A-7]. Later Kushner [C-1], [C-2] also derived the optimal solution.

This optimal solution involves the knowledge of the a posteriori probability which requires in general the solution of partial integro-differential equations. As observed by Kushner the general optimal nonlinear processor is an ∞ - dimensional system and as such unfeasible in practice. In order to get computationally effective algorithms, one has to resort to some kind of approximations. In most of them one of the steps in the derivation of the suboptimal processors is a truncated Taylor's series expansion of the system and observation function about a

nominal trajectory or the previous estimate. Depending on the various optimality criteria (e.g. MMSE, MAP, etc.) and on the nature of the involved approximations (e.g. finite parametrization of the a posterior probability, assumptions on the shape and moments of the conditional probability, order of the retained terms in the Taylor's series expansions, etc.), one is led to several approximate filters which in turn are computationally practicable.

The literature is indeed very vast in this subject and an appreciable number of suboptimal algorithms have been developed. A good account on the "state of the art" is given in Jazwinski [A-4]. Other suitable references include Snyder[A-6], Sage and Melsa [A-5]. Papers dealing with the subject include references [D].

Unfortunately the derivations of these filters do not give an intuitive measure of the relative importance of the several approximations made.

In other words the particular choice of a specific solution for a given problem has to be determined by some kind of an arbitrary criterion if not based first on a simulation study of how well the several filters do perform in the situation under consideration.

Filters with low order approximations are particularly relevant since they led to simpler solutions to implement.

In this work we will only consider filters with a first order approximation (only linear terms are retained in the Taylor series' expansions). In a subsequent work we intend to carry out the simulation of several second order filters.

We are going to simulate the so-called Extended Kalman filter (EKF) and the first order Maximum a Posterior filter (MAP).

The filter equations consist of the equations of propagation of the first and second moment of the conditional probability density function. The EKF is derived assuming the conditional mean estimate $\underline{\hat{x}}$ (MMSE estimate) is known and is used to expand the message and observation model in a Taylor series about

$$\underline{x}\ (t) = \underline{\hat{x}}(t) \tag{3-1}$$

and retaining the linear terms. With the linearized model we can then apply the results from linear filtering, i.e. the Kalman-Bucy theory.

The MAP* filter is derived, e.g. Sage and Melsa [A-5], by use of the maximum principle and a quadratric cost functional where the weighting matrices are the inverses of the system driving noise and of the observation noise covariance matrices.

We now summarize the general nonlinear estimation problem and the two approximate solutions we implement.

Dynamical system

$$\begin{cases} \dot{\underline{x}}(t) = \underline{f}(\underline{x}(t),\ t) + \underline{g}(\underline{x}(t),\ t)\ \underline{u}(t) \\ \underline{x}(t) = \underline{x}_o \end{cases} \tag{3-2}$$

Received signal

$$\underline{r}(t) = \underline{h}(t,\ \underline{x}(t)) + \underline{w}(t) \tag{3-3}$$

Noise statistics

$\underline{u}(t)$, $\underline{w}(t)$ are sample functions of zero mean, statistically

---

* Because of the wide number of available filters in the literature and because of the fact that alternate derivations using approximations of the same order lead in some cases to essentially equivalent algorithms the nomenclature is not uniform. The continuous version of the algorithm we call therein MAP filter is Snyder's filter derived using a MMSE criterion. Sage and Melsa rederive the discrete version result using a MAP criterion.

independent Gaussian white noise real vector processes.

$\underline{x}_o$ has (a priori) known statistics and is statistically independent of $\underline{u}(\tau)$ and $\underline{w}(t)$ for $\tau > o$ and every t.

Filter equations

Case I Extended Kalman filter

a) Estimator equation

$$\dot{\hat{\underline{x}}} = \underline{f}(\hat{\underline{x}}, t) + \underline{P}(t) \frac{\partial \underline{h}^T(\hat{\underline{x}}(t))}{\partial \hat{\underline{x}}} \underline{R}^{-1}[\underline{r}(t) - \underline{h}(\hat{\underline{x}}(t))] \tag{3-4}$$

b) Error covariance equation

$$\dot{\underline{P}}(t) = \frac{\partial \underline{f}}{\partial \hat{\underline{x}}} \underline{P}(t) + \underline{P}(t) \frac{\partial \underline{f}^T}{\partial \hat{\underline{x}}} + \underline{g}(\hat{\underline{x}}(t)) \underline{Q} \underline{g}^T(\hat{\underline{x}}(t)) - \underline{P}(t) \frac{\partial \underline{h}^T}{\partial \hat{\underline{x}}} \underline{R}^{-1} \frac{\partial \underline{h}}{\partial \hat{\underline{x}}} \underline{P}(t) \tag{3-5}$$

c) Initial conditions

$$\begin{cases} \hat{\underline{x}}(o) = E\, \underline{x}(o) \\ \underline{P}(o) = \underline{P}_o = \mathrm{cov}(\underline{x}(o), \underline{x}(o)) \end{cases} \tag{3-6}$$

Case II MAP filter

a) Estimator equation

$$\dot{\hat{\underline{x}}}(t) = \underline{f}(\hat{\underline{x}}(t)) + \underline{P}(t) \frac{\partial \underline{h}^T}{\partial \hat{\underline{x}}} \underline{R}^{-1}[\underline{r}(t) - \underline{h}(\hat{\underline{x}}(t))] \tag{3-7}$$

b) Error covariance equation

$$\dot{\underline{P}}(t) = \frac{\partial \underline{f}}{\partial \hat{\underline{x}}} \underline{P}(t) + \underline{P}(t) \frac{\partial \underline{f}^T}{\partial \hat{\underline{x}}} + \underline{g}(\hat{\underline{x}}(t)) \underline{Q} \underline{g}^T(\hat{\underline{x}}(t)) + \underline{P}(t) \frac{\partial}{\partial \hat{\underline{x}}} \left\{ \frac{\partial \underline{h}^T(\hat{\underline{x}}(t))}{\partial \hat{\underline{x}}} \underline{R}^{-1}[\underline{r}(t) - \underline{h}(\hat{\underline{x}}(t))] \right\} \underline{P}(t) \tag{3-8}$$

c) Initial conditions

$$\left|\begin{array}{l}\hat{\underline{x}}(o) = E\,\underline{x}(o) \\ \underline{P}_o = \mathrm{cov}(\underline{x}(o), \underline{x}(o))\end{array}\right. \tag{3-9}$$

Notation:

By $\frac{\partial \underline{f}}{\partial \underline{x}}$, $\frac{\partial \underline{h}}{\partial \underline{x}}$ we mean the Jacobians $\left.\frac{\partial \underline{f}(\underline{x}(t))}{\partial \underline{x}(t)}\right|_{\underline{x}(t) = \hat{\underline{x}}(t)}$ and $\left.\frac{\partial \underline{h}(\underline{x}(t))}{\partial \underline{x}(t)}\right|_{\underline{x}(t) = \hat{\underline{x}}(t)}$ and by $\frac{\partial \underline{f}^T}{\partial \underline{x}}$, $\frac{\partial \underline{h}^T}{\partial \underline{x}}$ their transposes. Hat stands for the expectation operator.

The particular form of equations(3-2) to (3-9) depends on the particular choice of the coordinate systems as was discussed in Chapter 2.

Comparison of equations (3-4) to (3-6) with (3-7) to (3-9) shows that the EKF and the MAP are very similar. In particular the estimator equations are identical and the only difference is on the quadratic term in the error covariance propagation equation. But while in the EKF the covariance does not depend explicitly on the received signal, in the MAP covariance propagation equation the received waveform enters explicitly. A priori one can expect an improvement of performance (at least when the underlying linearized assumption holds) with the MAP filter since the actual difference between the received signal and our best estimate influences directly the covariance determination.

On the other hand the implicit presence of the residual (or innovations process)

$$\underline{r}(t) - \underline{h}(\hat{\underline{x}}(t))$$

in the MAP covariance equation leads one to suspect that this filter will be more sensitive to, for example, the value of the signal to noise ratio and the actual geometry with the danger of eventual instabilities due to the theoretical impossibility but actual possibility of the propagated error covariance as computed by (3-5), loosing its positive semidefiniteness and so its practical significance leading to an unstable filter. These and other questions are explored in Chapter 6 where we present the simulation results.

As argued before (end of section 2-2) equations (3-2) are formal descriptions of the actual mathematical model. A similar discussion should be made with respect to equations (3-3) to (3-9) which are once again mere formal descriptions of the mathematical model. We should rewrite them in terms of differentials, to be interpreted in the sense of their corresponding Itô stochastic integral equations, given that the Stratanovitch integral seems to have problems of interpretation in nonlinear estimation theory (with this respect see Jazwinski [A-4] or Kushner [D-2], but also Stratonovitch [A-7], pp 43).

In the next chapter we reformulate the just described nonlinear estimation problem in the discrete framework of a digital computer.

Chapter IV

Implementation of the Nonlinear Receiver

Our comparative studies of the several processors, for different sets of parameters, different initial conditions and different tracks are to be made in a digital computer.

Before we actually simulate the filters we reformulate the continuous time, narrow band model as a discrete, low pass problem.

In section 4-1 we consider the quadrature component model for the narrow band signal problem of Chapter 2 and develop its low pass version. In sections 4-2 to 4-4, we approximate the continuous time by a discrete time stochastic model and write down the actual form of the several filters. Finally in section 4-5 we present an alternate algorithm, the square root filter (SQRT) implementation, to which we have to resort in order to circumvent practical problems due to computer roundoff errors.

4.1 Quadrature Components and Low Pass Model

We recall from Chapter 2 that the received signal at the ith sensor can be written as

$$r_i(t) = s_i(t, \underline{x}(t)) + w_i(t)$$

$$= \sqrt{2P}\, \sin w_c\left(t + \frac{p_i}{c} \sin\Theta(t) - \frac{R(t)}{c}\right) + w_i(t)$$

$$i = 1, \ldots, N \qquad (4\text{-}1)$$

We modeled $w_i(t)$ as a sample function from a white noise Gaussian process with spectral height of $\frac{N_o}{2}$. This model is just a mathematical realization of a true physical wideband pass process (noise at the sensors) which is present and corrupts the signal in the measured waveforms.

We assume that $\Theta(t)$ and $R(t)$ are low pass stochastic processes, so that $r_i(t)$, $i = 1, \ldots, N$, are in fact band-pass stochastic processes.

We recall, Van Trees [A-8] problem 3.3.8, or Viterbi [A-12], Appendix A, that for a bandpass stationary Gaussian process $n(t)$ with mean $\hat{n}$ and spectral density

$$\begin{aligned} &S_n(w) = S_L(w - w_o) + S_L(-w - w_o) \\ &\text{where: } S_L(w) \text{ is negligible for } |w| > w_o \end{aligned} \tag{4-2}$$

a sample function decomposition can be made

$$n(t) = \sqrt{2}\ n_1(t) \sin w_o t + \sqrt{2}\ n_2(t) \cos w_o t + \hat{n} \tag{4-3}$$

where $n_1(t)$ and $n_2(t)$ are sample functions of stationary, zero mean Gaussian processes, whose spectral densities are negligible for $|w| > w_o$, i.e. are essentially low pass stochastic processes.

If we let $R(\tau)$, $R_1(\tau)$, $R_2(\tau)$ be the covariance functions for $n(t)$, $n_1(t)$ and $n_2(t)$ respectively and $R_{12}(\tau)$, $R_{21}(\tau)$ the crosscovariances of $n_1(t)$ and $n_2(t)$ then the following relations hold:

$$\begin{cases} R_1(\tau) = R_2(\tau) \\ R_{12}(\tau) = -R_{21}(\tau) \\ R(\tau) = 2[R_1(\tau) \cos w_o\tau - R_{12}(\tau) \sin w_o\tau] \end{cases} \tag{4-4}$$

Furthermore if $S_L(w)$ is an even function then

$$R_{12}(\tau) = -R_{21}(\tau) = o \tag{4-5}$$

and the two processes $n_1(t)$ and $n_2(t)$ are uncorrelated and therefore, being Gaussian, are statistically independent.

Applying these results to the received waveform $r_i(t)$ we have:

$$r_i(t) = \sqrt{2}\ r_{c_i}(t) \cos w_o t + \sqrt{2}\ r_{s_i}(t) \sin w_o t$$

and we recover $r_{c_i}(t)$ and $r_{s_i}(t)$ as shown in Figure 4-1:

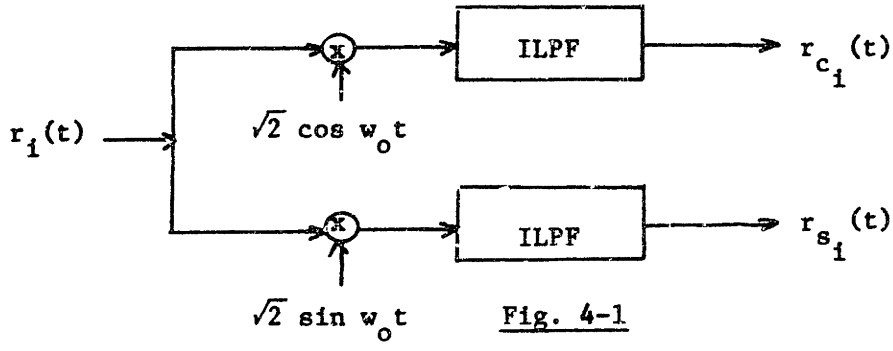


Fig. 4-1

where: ILPF stands for Ideal Low Pass Filter

$$\begin{aligned} r_{c_i}(t) &= \sqrt{P}\ \sin \tfrac{w_c}{c}[p_i \sin\Theta(t) - R(t)] + w_{c_i}(t) \\ r_{s_i}(t) &= +\sqrt{P}\ \cos \tfrac{w_c}{c}[p_i \sin\Theta(t) - R(t)] + w_{s_i}(t) \end{aligned} \tag{4-6}$$

$w_{c_i}(t)$, $w_{s_i}(t)$, $i = 1, \ldots, N$ are sample functions of uncorrelated zero mean, white Gaussian noise processes with spectral height of $\frac{N_o}{2}$.

In the simulation we work with the low pass components (4-6) instead of (4-1). We define

$$\begin{aligned} h_{c_i}(t) &= \sqrt{P}\ \sin \tfrac{w_c}{c}[p_i \sin\Theta(t) - R(t)] \\ h_{s_i}(t) &= +\sqrt{P}\ \cos \tfrac{w_c}{c}[p_i \sin\Theta(t) - R(t)] \end{aligned} \tag{4-7}$$

$$i = 1, \ldots, N$$

In vector notation

$$\underline{r}(t) = \begin{bmatrix} \underline{r_c}(t) \\ \cdots \\ \underline{r_s}(t) \end{bmatrix} = \underline{h}(t, \underline{x}(t)) + \underline{w}(t) \tag{4-8}$$

with:

$$\underline{r}(t) \in \mathbb{R}^{2N},\ \underline{h}(\underline{x}(t),\ t) \in \mathbb{R}^{2N},\ \underline{w}(t) \in \mathbb{R}^{2N}$$

$$\underline{r}_c(t) = \underline{h}_c(\underline{x}(t),\ t) + \underline{w}_c(t) \tag{4-9}$$

$$\underline{r}_s(t) = \underline{h}_s(\underline{x}(t),\ t) + \underline{w}_s(t)$$

and

$$\underline{h}(\underline{x}(t),\ t) = \begin{bmatrix} \underline{h}_c(\underline{x}(t),\ t) \\ \cdots\cdots \\ \underline{h}_s(\underline{x}(t),\ t) \end{bmatrix} = \begin{bmatrix} h_{c_1}(\underline{x}(t),\ t) \\ \cdots \\ h_{c_N}(\underline{x}(t),\ t) \\ h_{s_1}(\underline{x}(t),\ t) \\ \cdots \\ h_{s_N}(\underline{x}(t),\ t) \end{bmatrix}$$

$$\underline{w}(t) = \begin{bmatrix} \underline{w}_c(t) \\ \cdots \\ \underline{w}_s(t) \end{bmatrix} = \begin{bmatrix} w_{c_1}(t) \\ \cdots \\ w_{c_N}(t) \\ w_{s_1}(t) \\ \cdots \\ w_{s_N}(t) \end{bmatrix} \tag{4-10}$$

We conclude that at the receiver we have available a 2N-dimensional stochastic vector waveform phase modulated by the system dynamics and which is to be processed in some optimal way in order to estimate the range, range rate, bearing and bearing rate.

4.2 Sampled Data Version of the Continuous Time Stochastic System

In Chapter 2 we presented two (formal) descriptions of the dynamical system, summarized by equations (2-5') and (2-12')

Rectangular:

$$\begin{cases} d\underline{x}_r(t) = \underline{F}\,\underline{x}_r(t)\,dt + \underline{G}\,d\underline{\beta}(t) \\ \underline{x}_r(o) = \underline{x}(o) \end{cases} \tag{4-11}$$

Polar:

$$\begin{cases} d\underline{x}_p(t) = \underline{f}(\underline{x}_p(t),\ t)\ dt + \underline{g}(\underline{x}_p(t),\ t)\ d\underline{\beta}(t) \\ \underline{x}_p(o) = \underline{x}_o \end{cases} \tag{4-12}$$

depending on the particular choice of the coordinate system.

We observe that the state vectors in (4-11) and (4-12) are related by the nonlinear transformations:

$$\begin{aligned} x_{p_1} &= R = \sqrt{x_{r_1}^2 + x_{r_3}^2} = \sqrt{X_T^2 + Y_T^2} \\ x_{p_3} &= \Theta = \tan^{-1}\frac{x_{r_1}}{x_{r_3}} = \tan^{-1}\frac{X_T}{Y_T} \end{aligned} \tag{4-13}$$

The strategy we adopt in our simulation studies is to generate the target's motion in rectangular coordinates and obtain the state vector $\underline{x}_p$ through the nonlinear transformation (4-13).

When simulating the dynamical system (2-5) care must be taken because of its stochastic nature. Essentially at stake is the best way of approximating a stochastic continuous process by a stochastic sequence.

Wong and Zakai [C-4], [C-5], see Jazwinski [A-4], considered a polygonal approximation to the Wiener process β_t

$$\beta_t^\rho = \beta_{t_i}^\rho + \frac{\beta_{t_{i+1}}^\rho - \beta_{t_i}^\rho}{t_{i+1} - t_i}(t - t_i), \qquad t_i \le t \le t_{i+1} \tag{4-14}$$

$$\rho = \max_i (t_{i+1} - t_i)$$

and showed that for broad conditions on the system functions f(.) and g(.) (e.g. Lipschitz type on both arguments) and on the initial state the Markov processes generated by the sequence of differential equations

$$\begin{cases} dx_t^\rho = f(x_t^\rho, t)dt + g(x_t^\rho, t)d\beta_t^\rho \\ x^\rho(o) = x(o) \end{cases} \tag{4-15}$$

would converge in the mean square sense to x(t) as generated by

$$dx_t = f(x_t, t)dt + g(x_t, t)d\beta_t \tag{4-16}$$

i.e.

$$\text{l.i.m.}_{\rho \to o}\ x_t^\rho = x_t \tag{4-17}$$

This important result suggests that we approximate the white noise process $\underline{w}(t)$ by its sampled version, in order to get a discrete white noise sequence, after conveniently adjusting the white noise level.

Example (Kalman) (see [A-4])

Let $\{\underline{x}_n,\ n = 1, 2, \ldots\}$ be a white Gaussian sequence with statistics

$$\begin{cases} E\,\underline{x}_n = \underline{o}\ , \text{ every } n \\ \text{cov}(\underline{x}_n, \underline{x}_m) = \underline{Q}\,\delta_{mn} \end{cases} \tag{4-18}$$

where δ_{mn} stands for the Kronecker symbol and $\underline{Q}$ is a constant matrix.

Define the stochastic process $\underline{y}(t)$ by

$$\begin{cases} \dot{\underline{y}}(t) = \underline{x}_T(t) \\ \underline{y}(o) = \underline{o} \end{cases} \tag{4-19}$$

where

$$\underline{x}_T(t) = \underline{x}_n,\ \frac{1}{T} \text{ is an integer},\ nT \leq t < (n+1)T \tag{4-20}$$

Consider the random vector

$$\underline{y}(1) = \int_o^1 \underline{x}_T(t)dt \tag{4-21}$$

and compute the variance of $|\underline{y}(1)|$:

$$\text{var}(|\underline{y}(1)|) = \text{tr} \sum_{i=1}^{1/T} \underline{Q}_T^2 = \text{tr}\ \underline{Q}_T \tag{4-22}$$

which goes to zero as $T \to o$. In order to circumvent this result and to keep the variance of $\underline{y}(1)$ constant as $T \to o$ we have to adjust the noise level by a factor of $\frac{1}{T}$ when we approximate the continuous time stochastic process by its sampled version.

Returning to the actual implementation of the stochastic dynamical system (2-5) we construct the following discrete version:

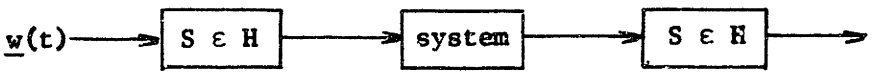
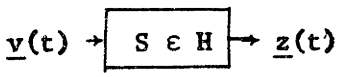


Figure 4-2

where



is a sample and hold device with sampling period T, i.e.:

$$\underline{z}(t) = \underline{v}(kT) \ , \quad kT \leq t < (k+1)T \tag{4-23}$$

Defining

$$\begin{aligned} \underline{x}_k &= \underline{x}(kT) \\ \underline{u}_k &= \underline{u}(kT) \end{aligned} \tag{4-24}$$

the variation of constants formula, e.g. R.W. Brocket [A-2], gives for the system (2-5):

$$\underline{x}(t) = e^{\underline{F}(t - t_o)}\, \underline{x}(t_o) + \int_{t_o}^{t} e^{\underline{F}(t - \sigma)}\, \underline{G}\, \underline{u}(\sigma) d\sigma \tag{4-25}$$

Choosing:

$$\begin{aligned} t_o &= kT \\ t &\in [kT,\ (k+1)T] \end{aligned} \tag{4-26}$$

we get for the sampled data version of (2-5)

$$\underline{x}_{k+1} = e^{\underline{F}T}\,\underline{x}_k + \left(\int_o^T e^{\underline{F}(T-\sigma)}d\sigma\right)\,\underline{G}\,\underline{u}(kT) \tag{4-27}$$

and at intermediate points:

$$\underline{x}(t) = e^{\underline{F}(t-kT)}\underline{x}_k + \left(\int_o^{t-kT} e^{\underline{F}(t-kT-\sigma)}d\sigma\right)\,\underline{G}\,\underline{u}(kT), \qquad kT \leq t < (k+1)T \tag{4-28}$$

In (4-27) and (4-28) we approximated the white noise process $\underline{u}(t)$ by a "flat white noise."

$$\underline{u}(t) \simeq \tilde{\underline{u}}(t) = \underline{u}(kT) \qquad kT \leq t < (k+1)T \tag{4-29}$$

Hence the sampled data mathematical description of the dynamical system (2-5):

$$\begin{cases} \dot{\underline{x}}(t) = \underline{F}\,\underline{x}(t) + \underline{G}\,\underline{u}(t) \\ \underline{u}(o) = \underline{x}_o \end{cases} \tag{2-5}$$

is:

$$\begin{cases} \underline{x}_{k+1} = \tilde{\underline{F}}\,\underline{x}_k + \tilde{\underline{G}}\,\underline{u}_k \\ \underline{x}_o = \underline{x}_o \end{cases} \tag{4-30}$$

where

$$\begin{cases} E\,\underline{u}_k = \underline{o} \\ E\,\underline{u}_k\,\underline{u}_j^T = \frac{1}{T}\,\underline{Q}\,\delta_{kj} \end{cases} \tag{4-31}$$

$$\begin{aligned} \underset{\sim}{F} &= e^{\underline{F}T} \\ \underset{\sim}{G} &= \int_o^T e^{\underline{F}(T-\sigma)} d\sigma \, \underline{G} \end{aligned} \tag{4-32}$$

For the particular matrices $\underline{F}$ and $\underline{G}$ of Chapter 2, we obtain:

$$\underset{\sim}{F} = \begin{bmatrix} 1 & T & & \\ o & 1 & \underline{o} & \\ & & 1 & T \\ \underline{o} & & o & 1 \end{bmatrix}$$

(4-33)

$$\underset{\sim}{G} = \begin{bmatrix} \frac{T^2}{2} & o \\ T & o \\ o & \frac{T^2}{2} \\ o & T \end{bmatrix}$$

For the nonlinear system (2-12) a simple discrete approximation can be derived (Euler's approximation), in a nonrigorous way, just by considering the continuous stochastic process $\underline{x}(t)$ to be the limit, as the samples become dense ($T \to o$) of the stochastic sequence generated by the discrete system:

$$\begin{cases} \underline{x}_{k+1} = \underset{\sim}{f}(\underline{x}_k) + \underset{\sim}{g}(\underline{x}_k)\, \underline{u}_k \\ \underline{x}_o = \underline{x}_o \end{cases} \tag{4-34}$$

and where

$$\begin{aligned} \underset{\sim}{f}(\underline{x}_k) &= [T\, \underline{f}(\underline{x}(t)) + \underline{x}(t)] \Big|_{t = kT} \\ \underset{\sim}{g}(\underline{x}_k) &= T\, \underline{G}(x(t)) \Big|_{t = kT} \end{aligned} \tag{4-35}$$

and $\{\underline{u}_k,\ k = 0, 1, 2, \ldots\}$ is a white Gaussian sequence

$$\underline{u}_k = \underline{u}(t)\Big|_{t = kT} \tag{4-36}$$

with adjusted spectral level of $\frac{1}{T}\underline{Q}$.

Note that in the linear case the sampled version is valid for all T whereas in the nonlinear case it is a small T approximation.

4.3 Received Signal

In the discretized version of the continuous physical model we assume that the measurements (4-1) are available at discrete points of time.

The received signal in the discrete model is

$$\begin{aligned} \underline{r}_k &= \underline{r}(t)\Big|_{t = kT} \\ &= \underline{h}(\underline{x}_k,\ kT) + \underline{w}_k \end{aligned} \tag{4-37}$$

with

$$\underline{w}_k = \underline{w}(kT) \tag{4-38}$$

and where once again we normalize the spectral height of the white noise sequence $\{\underline{w}_k,\ k = 1, 2, \ldots\}$ to $\frac{1}{T}\frac{N_o}{2}$.

4.4 Discrete Extended Kalman Filter - Discrete MAP Filter

We saw in the previous paragraphs that for simplicity the solution to the nonlinear estimation problem that we simulate in this work is the discrete version of the nonlinear algorithms of Chapter 3.

In this section we present the several suboptimal solutions of the nonlinear discrete estimation problem.

Before we proceed we will make preliminary definitions. Because the filters we are going to implement, the Extended Kalman filter (EKF) and the Maximum A Posteriors (MAP) filter, correspond to first order approximations, we need the following Jacobian matrices:

$$\frac{\partial \underline{\tilde{f}}}{\partial \underline{x}_k} \triangleq \frac{\partial \underline{\tilde{f}}(\underline{x}_k)}{\partial \underline{x}_k} = \begin{bmatrix} \frac{\partial f_1(\underline{x}_k)}{\partial x_{1_k}} & \cdots & \frac{\partial f_1(\underline{x}_k)}{\partial x_{4_k}} \\ \vdots & & \\ \frac{\partial f_4(\underline{x}_k)}{\partial x_{1_k}} & \cdots & \frac{\partial f_4(\underline{x}_k)}{\partial x_{4_k}} \end{bmatrix} \tag{4-39}$$

$$\underline{\underline{H}}_k \triangleq \frac{\partial \underline{h}(\underline{x}_k)}{\partial \underline{x}_k} = \begin{bmatrix} \frac{\partial \underline{h}_c(\underline{x}_k)}{\partial \underline{x}_k} \\ \cdots \\ \frac{\partial \underline{h}_s(\underline{x}_k)}{\partial \underline{x}_k} \end{bmatrix} = \begin{bmatrix} \frac{\partial h_{c_1}(\underline{x}_k)}{\partial x_{1_k}} & \cdots & \frac{\partial h_{c_1}(\underline{x}_k)}{\partial x_{4_k}} \\ \vdots & & \\ \frac{\partial h_{c_N}(\underline{x}_k)}{\partial x_{1_k}} & \cdots & \frac{\partial h_{c_N}(\underline{x}_k)}{\partial x_{4_k}} \\ \frac{\partial h_{s_1}(\underline{x}_k)}{\partial x_{1_k}} & \cdots & \frac{\partial h_{s_1}(\underline{x}_k)}{\partial x_{4_k}} \\ \vdots & & \\ \frac{\partial h_{s_N}(\underline{x}_k)}{\partial x_{1_k}} & \cdots & \frac{\partial h_{s_N}(\underline{x}_k)}{\partial x_{4_k}} \end{bmatrix} \tag{4-40}$$

Applying these definitions to both cases, rectangular and polar coordinates, we have:

Rectangular coordinates:

$$\frac{\partial \underline{\underline{f}}}{\partial \underline{x}_k} = \underline{\underline{\Phi}} = e^{\underline{\underline{F}}\,T}$$

$$\underline{H}_k = \frac{\partial \underline{h}_r(\underline{x}_k)}{\partial \underline{x}_k} \tag{4-41}$$

Polar coordinates:

$$\frac{\partial \underline{\underline{f}}}{\partial \underline{x}_k} = T\,\frac{\partial \underline{\underline{f}}}{\partial \underline{x}_k} + \underline{\underline{I}}$$

$$\underline{H}_k = \frac{\partial \underline{h}_p(\underline{x}_k)}{\partial \underline{x}_k} \tag{4-42}$$

We next define a composite observation as

$$\P_k \triangleq \{\underline{r}_j,\ j = 1, \ldots, k\} \tag{4-43}$$

i.e., $\P_k$ is the set of received vectors up to and including time k.

Let:

$$\hat{\underline{x}}_k \triangleq E\{\underline{x}_k \mid \P_k\}$$

$$\bar{\underline{x}}_k \triangleq E\{\underline{x}_k \mid \P_{k-1}\} \tag{4-44}$$

$$\underline{\underline{P}}_k \triangleq E\{(\underline{x}_k - \hat{\underline{x}}_k)\,(\underline{x}_k - \hat{\underline{x}}_k)^T \mid \P_k\}$$

$$\underline{\underline{M}}_k \triangleq E\{(\underline{x}_k - \bar{\underline{x}}_k)(\underline{x}_k - \bar{\underline{x}}_k)^T \mid \P_{k-1}\} \tag{4-45}$$

i.e.

$\hat{\underline{x}}_k$ is the corrected estimate of $\underline{x}_k$ after we incorporate the measured $\underline{r}_k$;

$\bar{\underline{x}}_k$ is the predicted value of the estimate of $\underline{x}_k$, before we do the measurement at time k;

$\underline{P}_k$ is the covariance of the corrected error (after we measured $\underline{r}_k$); and

$\underline{M}_k$ is the covariance of the predicted error at time k.

Within the above framework we present the Discrete Extended Kalman Filter. The most common format is as follows:

I. Measurement Update Equations (correction)

A. Estimator Equation

$$\begin{cases} \hat{\underline{x}}_k = \bar{\underline{x}}_k + \underline{K}_k(\underline{r}_k - \underline{h}(\bar{\underline{x}}_k, k)) \\ \hat{\underline{x}}_o = E\,\underline{x}_o \end{cases} \tag{4-46}$$

B. Kalman Gain

$$\underline{K}_k = \underline{P}_k \underline{H}_k^T \tilde{\underline{R}}_k^{-1} \tag{4-47}$$

C. Covariance Propagation Equation

$$\begin{cases} \underline{P}_k = \underline{M}_k - \underline{M}_k \underline{H}_k^T [\underline{H}_k \underline{M}_k \underline{H}_k^T + \tilde{\underline{R}}_k]^{-1} \underline{H}_k \underline{M}_k \\ \underline{P}_o = \mathrm{cov}(\underline{x}_o, \underline{x}_o) \end{cases} \tag{4-48}$$

II. Time Update Equations (prediction)

A. Estimator Equation

$$\bar{\underline{x}}_{k+1} = \tilde{\underline{f}}(\hat{\underline{x}}_k) + \tilde{\underline{g}}(\hat{\underline{x}}_k)\,\underline{u}_k \tag{4-49}$$

B. Covariance Propagation Equation

$$\underline{M}_{k+1} = \frac{\partial \tilde{\underline{f}}}{\partial \underline{x}_k} \underline{P}_k \left(\frac{\partial \tilde{\underline{f}}}{\partial \underline{x}_k}\right)^T + \tilde{\underline{g}}(\hat{\underline{x}}_k)\,\tilde{\underline{Q}}\,\tilde{\underline{g}}^T(\hat{\underline{x}}_k) \tag{4-50}$$

Remark:

1) We can evaluate at each stage the Jacobians $\frac{\partial \tilde{\underline{f}}}{\partial \underline{x}_k}$ at $\underline{x}_k = \hat{\underline{x}}_k$ and that is exactly what we mean by the notation $\frac{\partial \tilde{\underline{f}}}{\partial \hat{\underline{x}}_k}$. The Jacobians $\underline{H}_k$ can be evaluated at the predicted estimate.

2) Because at each stage we relinearize the system and observation functions about the previous estimate we cannot compute off-line the covariances, and hence the Kalman gain has to be computed on-line since it is coupled to the current estimates. This on-line computation of the propagated covariance values increases the computational burden of the filter and a desirable property is to see if the gain reaches a steady state. We will see that for some cases this is in fact true and that these steady state values are intimately related to the assumed parameters (array geometry, signal to noise ratio, driving noise level).

A close look at equations (4-46) to (4-50) shows that the main computational effort relates to the matrix inversion in (4-48), which is a 2N x 2N matrix. By standard manipulations starting with (4-48) we have:

$$\begin{aligned} \underline{P} &= \underline{M} - \underline{M}\underline{H}^T [\underline{H}\underline{M}\underline{H}^T + \underline{R}]^{-1} \underline{H}\underline{M} \\ \underline{P}\underline{H}^T\underline{R}^{-1} &= \underline{M}\underline{H}^T\underline{R}^{-1} [\underline{I} - \underline{R}[\underline{H}\underline{M}\underline{H}^T + \underline{R}]^{-1} \underline{H}\underline{M}\underline{H}^T\underline{R}^{-1}] \\ &= \underline{M}\underline{H}^T\underline{R}^{-1} [\underline{H}\underline{M}\underline{H}^T\underline{R}^{-1} + \underline{I}]^{-1} [\underline{H}\underline{M}\underline{H}^T\underline{R}^{-1} + \underline{I} - \underline{H}\underline{M}\underline{H}^T\underline{R}^{-1}] \\ &= \underline{M}\underline{H}^T [\underline{H}\underline{M}\underline{H}^T + \underline{R}]^{-1} \end{aligned} \tag{4-51}$$

Substitution of (4-51) in (4-48) gives:

$$\underline{P} = \underline{M} - \underline{P}\underline{H}^T\underline{R}^{-1} \underline{H}\underline{M}$$

or

$$\underline{P} = [\underline{M}^{-1} + \underline{H}^T\underline{R}^{-1} \underline{H}]^{-1} \tag{4-52}$$

or equivalently:

$$\underline{P} = \underline{M}[\underline{I} + \underline{H}^T\underline{R}^{-1}\underline{H}\underline{M}]^{-1} \tag{4-53}$$

Looking at (4-53) it seems that we have increased the computational effort since we need the inversion of a 2N x 2N matrix $\underline{R}$ plus the inversion of the n x n bracketed matrix.

But in the case of stationary noise, $\underline{R}$ is a constant matrix and so the 2N x 2N matrix inversion takes place only once and by considering the coding (4-53) instead of (4-48) we do in fact decrease at each stage the computational burden.

Equation (4-52) will be used for theoretical discussions.

MAP Discrete Filter

This algorithm is essentially similar to the discrete EKF described above. The only difference relates to the measurement update covariance equation.

This equation becomes, instead of (4-53):

$$\underline{P}_k = [\underline{M}_k^{-1} - \frac{\partial}{\partial \underline{\bar{x}}_k} \{\underline{H}_k^T \underline{\tilde{R}}_k^{-1} (\underline{r}_k - \underline{h}(\underline{\bar{x}}_k, k))\}]^{-1} \tag{4-54}$$

where

$$\frac{\partial}{\partial \underline{\bar{x}}_k} (.) = \frac{\partial}{\partial \underline{x}_k} (.) \Bigg|_{\underline{x}_k = \underline{\bar{x}}_k} \tag{4-55}$$

We observe that $\underline{P}_k$ as given by (4-54) is essentially a stochastic matrix (in the sense that its elements are random numbers) and so a priori, we suspect that for some cases (essentially depending on the signal to noise ratio and on the geometry tracking) it can lose its positive semidefiniteness.

If
$$- \frac{\partial}{\partial \underline{\bar{x}}_k} \{\underline{H}_k^T \underline{R}_k^{-1} (\underline{r}_k - \underline{h}(\underline{\bar{x}}_k, k))\} \geq 0 \tag{4-56}$$

then we can find a matrix $\underline{\hat{H}}_k$ such that

$$\underline{\hat{H}}_k \underline{R}_k^{-1} \underline{\hat{H}}_k = - \frac{\partial}{\partial \underline{\bar{x}}_k} \{.\} \tag{4-57}$$

Bearing in mind (4-57) (and (4-56)) we can formally reduce equation (4-54) to equation (4-52) or (4-53).

For the sake of easy reference we summarize the whole reformulated discrete nonlinear estimation problem.

I. Discrete System:

$$\begin{cases} \underline{x}_{k+1} = \underset{\sim}{\underline{f}}(\underline{x}_k) + \underline{g}(\underline{x}_k)\,\underline{u}_k \\ \underline{x}_o = \underline{x}_o \end{cases} \tag{4-58}$$

A. Rectangular coordinates

$$\underset{\sim}{\underline{f}}(\underline{x}_k) = \underset{\sim}{\underline{F}}\,\underline{x}_k = \begin{bmatrix} 1 & T & \underline{o} & \\ o & 1 & & \\ & & 1 & T \\ \underline{o} & & o & 1 \end{bmatrix} \underline{x}_k \tag{4-59}$$

$$\underset{\sim}{\underline{g}}(\underline{x}_k) = \underset{\sim}{\underline{G}} = \begin{bmatrix} T^2/2 & o \\ T & o \\ o & T^2/2 \\ o & T \end{bmatrix} \tag{4-60}$$

B. Polar coordinates

$$\underset{\sim}{\underline{f}}(\underline{x}_k) = [T\,\underline{f}(\underline{x}(t)) + \underline{x}(t)]\Big|_{t=kT} = T \begin{bmatrix} x_2(k) \\ x_1 x_4^2(k) \\ x_4(k) \\ \dfrac{-2x_2x_4(k)}{x_1(k)} \end{bmatrix} + \underline{x}_k \tag{4-61}$$

$$\underset{\sim}{\underline{g}}(\underline{x}_k) = T\,\underline{g}(\underline{x}(t))\Big|_{t=kT} = T \begin{bmatrix} o & o \\ \sin x_3(k) & \cos x_3(k) \\ o & o \\ \dfrac{\cos x_3(k)}{x_1(k)} & \dfrac{-\sin x_3(k)}{x_1(k)} \end{bmatrix} \tag{4-62}$$

II. Received Signal:

$$\underline{r}_k = \underline{r}(t)\Big|_{t = kT} = \underline{h}(\underline{x}_k, k) + \underline{w}_k \tag{4-63}$$

with

$$\begin{aligned} h_{c_i}(\underline{x}_k, k) &= \sqrt{P}\ \sin \frac{w_c}{c}\,[p_i \sin\Theta(k) - R(k)] \\ h_{s_i}(\underline{x}_k, k) &= \sqrt{P}\ \cos \frac{w_c}{c}\,[p_i \sin\Theta(k) - R(k)] \end{aligned} \tag{4-64}$$

and $\Theta(k)$ and $R(k)$ were previously defined in terms of the state variables, either in polar or rectangular coordinates.

III. Statistics

$$\begin{cases} E\,\underline{u}(k) = o \\ E\,\underline{u}(k)\,\underline{u}^T(j) = \underline{Q}_k\,\delta_{kj} \quad \text{with} \quad \underline{Q}_k = \frac{Q}{T}\,\underline{I} \end{cases} \tag{4-65}$$

$$\begin{cases} E\,\underline{x}_o = \hat{\underline{x}}_o \\ \operatorname{cov}(\underline{x}_o, \underline{x}_o) = \underline{P}_o \end{cases} \tag{4-66}$$

$$\begin{cases} E\,\underline{w}_k = \underline{o} \\ E\,\underline{w}_k \underline{w}_j^T = \underline{R}_k\,\delta_{kj} \text{ with } \underline{R}_k = \frac{1}{T}\,\frac{N_o}{2}\,\underline{I} \end{cases} \tag{4-67}$$

$$\begin{cases} E\,x_{o_i}\,u_\ell(k) = o \ , \quad \text{every } i = 1, \ldots, 4,\ \ell = 1,2,\ k \geq o \\ E\,x_{o_i}\,w_\ell(k) = o \ , \quad \text{every } i = 1, \ldots, 4, \text{ every } \ell, \text{ any } k \\ E\,u_i(k)\,w_\ell(j) = o, \quad \text{every } i, \ell, k \text{ and } j \end{cases} \tag{4-68}$$

IV. Extended Kalman Filter

A. Starting Values

$$\begin{cases} \hat{\underline{x}}(o) = E\,\underline{x}_o \\ \underline{P}(o) = \mathrm{cov}(\underline{x}_o, \underline{x}_o) = \underline{P}_o \end{cases} \tag{4-69}$$

B. Time Update (prediction)

1. Estimator Equation

$$\bar{\underline{x}}_{k+1} = \underset{\sim}{f}(\hat{\underline{x}}_k) \tag{4-70}$$

2. Covariance Equation

$$\underline{M}_{k+1} = \frac{\partial \underset{\sim}{f}}{\partial \hat{\underline{x}}_k}\,\underline{P}_k\,\frac{\partial \underset{\sim}{f}^T}{\partial \hat{\underline{x}}_k} + \underset{\sim}{g}(\hat{\underline{x}}_k)\,\underline{Q}_k\,\underset{\sim}{g}^T(\hat{\underline{x}}_k) \tag{4-71}$$

C. Measurement Update Equation (correction)

1. Estimator Equation

$$\hat{\underline{x}}_k = \bar{\underline{x}}_k + \underline{K}_k(\underline{r}_k - \underline{h}(\bar{\underline{x}}_k, k)) \tag{4-72}$$

2. Kalman Gain

$$\underline{K}_k = \underline{P}_k\,\underline{H}_k^T\,\underline{R}_k^{-1} \tag{4-73}$$

3. Covariance Equation

$$\underline{P}_k = \underline{M}_k[\underline{I} + \underline{H}_k^T\,\underline{R}_k^{-1}\underline{H}_k\,\underline{M}_k]^{-1} \tag{4-74}$$

V. MAP Filter

Framework of equations (4-69) - (4-73). The covariance equation can still formally be written as (4-74) with the foregoing definition (4-57).

$$\hat{\underline{H}}_k^T\,\underline{R}_k^{-1}\,\hat{\underline{H}}_k \triangleq -\frac{\partial}{\partial \bar{\underline{x}}_k}\{\underline{H}_k^T\,\underline{R}_k^{-1}[\underline{r}_k - \underline{h}(\bar{\underline{x}}_k, k)]\} \tag{4-75}$$

4.5 Square Root Algorithm

As it will become apparent from the simulation results the incorporation of new accurate available data reduces drastically the covariance matrix to a small value.

Due to the truncation and roundoff errors introduced by the computer, particularly relevant for example when subtraction of small numbers of almost the same size occurs (e.g. covariance measurement update equation) the covariance matrix may lose its positive semidefiniteness becoming meaningless.

The numerical computation of the covariance matrix becomes particularly sensitive when the matrix condition number, defined as

$$\mu = \frac{|\text{maximum eigenvalue}|}{|\text{minimum eigenvalue}|}$$

is of the order of b^n, where

b is the arithmetic basis used

n is the number of significant digits.

As we will see, particularly with the polar frame where in the diagonal of the covariance matrix appears the variance of the range and the variance of the bearing rate, μ is in fact of the order of b^n.

The numerical improvement of the covariance matrix reflects directly in the overall performance of the filter since the Kalman gain is directly proportional to the covariance matrix.

There are several remedies to improve the numerical ability of the filter and to obviate the inherently theoretical impossibility of the propagated covariance losing its positive semidefiniteness. We consider the square root implementation of the filter equations.

For a symmetric positive semidefinite matrix $\underline{P}$ we can always find a (not unique) matrix $\underline{W}$ such that

$$\underline{P} = \underline{W}\underline{W}^T \tag{4-77}$$

The nonuniqueness can be used to impose additional requirements on the form of $\underline{W}$, in order to expedite the numerical computations (e.g. the Cholesky algorithm [E-4], computes $\underline{W}$ in lower Δ form with positive diagonal elements). We will call $\underline{W}$ the square root form of $\underline{P}$.

The square root filter algorithm propagates the square root of $\underline{P}$ and $\underline{M}$ instead of $\underline{P}$ and $\underline{M}$ themselves. This improves the filter's ability of keeping $\underline{P} \geq o$ as well as improves its numerical efficiency since

$$\mu[\underline{P}^{1/2}] = [\mu[\underline{P}]]^{1/2} \tag{4-78}$$

There are several available methods. Suitable survey references include Schmidt [E-5] and Kaminski et al. [E-4].

After a brief discussion of the pertinent methods, we construct the square root algorithm (SQRT) as it applies to our problem.

Let

$$\underline{P} = \underline{W}\underline{W}^T \tag{4-79}$$

$$\underline{M} = \underline{S}\underline{S}^T \tag{4-80}$$

$$\underline{G}\underline{Q}\underline{G}^T = \underline{V}\underline{V}^T \tag{4-81}$$

The SQRT includes algorithms for:

Step 1) Initializing the filter, i.e., to find from a given covariance $\underline{P}_o$ its square root.

Step 2) Correcting the square root covariance matrix when incorporating new data, i.e. for computing $\underline{W}$.

Step 3) Propagating forward in time the square root covariance matrix, i.e. an equation for $\underline{S}$.

We consider separately each one of the above steps.

Step 1) Initialization of the filter

Cholesky algorithm, [E-4], is an efficient algorithm constructing the square root of a given matrix in triangular form. We present the routine in Appendix B.

Step 2) Equation for propagating $\underline{W}$

The first algorithm was developed by Potter, e.g. Battin [A-1], problem 9.11, for the case of a no driving noise plant and scalar measurement. This method is referred to as Potter's algorithm and is essentially the method discussed below. Bellantoni and Dodge [E-2] generalized Potter's method to the case of r-vector valued measurements by essentially considering a spectral decomposition procedure. Andrews [E-1] gave another (simpler) generalization of Potter's method for the case of r-vector measurements and also gave an alternate formulation for the SQRT.

Schmidt [E-5] discusses the different algorithms as they apply to the nonzero driving noise case.

We discuss very briefly Potter's method and some of its generalizations in order to have the necessary background to construct the SQRT.

Potter's Method

The essential step in Potter's method consists of the factorization of

$$\underline{P} = \underline{M} - \underline{M}\underline{h}^T \left[\underline{h}\underline{M}\underline{h}^T + \frac{N}{2}o\right]^{-1} \underline{h}\underline{M} \tag{4-82}$$

We remark that in (4-82) $\underline{h}^T$ is a column vector and $\underline{h}\underline{M}\underline{h}^T + \frac{N}{2}o$ is a scalar.

We present the main argument in the proof of the following

Lemma : Let $\underline{P} = \underline{W}\underline{W}^T$

The measurement correction of $\underline{W}$ for the no plant noise and acalar measurement case can be implemented by the recurrent equation

$$\underline{W} = \underline{S}[\underline{I} - \frac{\underline{b}\underline{b}^T}{a(1 + c)}] \tag{4-83}$$

where

$$a = \underline{h}\underline{M}\underline{h}^T + \frac{N}{2}o \tag{4-84}$$

$$c = \frac{N_o/2}{\underline{b}\underline{b}^T + N_o/2} \qquad \underline{b} = \underline{S}^T\underline{h}^T \tag{4-85}$$

Proof

Using definitions (4-79) to (4-81) we can write equation (4-82) in the form

$$\underline{W}\underline{W}^T = \underline{S}(\underline{I} - a^{-1}\,\underline{b}\underline{b}^T)\underline{S}^T \tag{4-86}$$

We now show by direct verification that

$$\underline{I} - a^{-1}\,\underline{b}\underline{b}^T = \left(\underline{I} - \frac{\underline{b}\underline{b}^T}{a(1 + c)}\right)\left(\underline{I} - \frac{\underline{b}\underline{b}^T}{a(1 + c)}\right)^T \tag{4-87}$$

where a and c are given as in the statement of the Lemma.

The expansion of the right hand side leads to:

$$= I + \frac{\underline{b}\underline{b}^T\underline{b}\underline{b}^T}{a^2(1 + c)^2} - 2\,\frac{\underline{b}\underline{b}^T}{a(1 + c)} \tag{4-88}$$

$$= \underline{I} - a^{-1}\,\underline{b}\left\{\frac{2}{1 + c} - \frac{\underline{b}^T\underline{b}}{a(1 + c)^2}\right\}\underline{b}^T \tag{4-89}$$

But the scalar

$$\{.\} = \frac{2}{1+c} - \frac{\underline{b}^T\underline{b}}{a(1+c)^2} = \frac{2a(1+c) - \underline{b}^T\underline{b}}{a(1+c)^2} = \frac{a + N_o/2 + 2ac}{a(1+c)^2}$$

$$= \frac{a(1 + c^2 + 2c)}{a(1+c)^2} = 1$$

and so equation (4-83) follows.

QED.

Andrews generalized Potter's method to the r-measurement case by first diagonalizing the covariance matrix $\underline{R}$ via a similarity transformation (recall that $\underline{R}$ being covariance matrix of a real vector stochastic process is $\underline{R} = \underline{R}^T \geq o$). Let $\underline{T}$ be an r x r orthogonal matrix s.t. $\underline{\Lambda}$ defined by

$$\underline{\Lambda} = \underline{T}^T\underline{R}\underline{T} \tag{4-90}$$

is a diagonal matrix.

Then equation (4-48) can be written as

$$\underline{P} = \underline{M} - \underline{M}(\underline{HT})\ [(\underline{HT})^T\ \underline{M}(\underline{HT}) + \underline{\Lambda}]^{-1}\ (\underline{HT})^T\ \underline{M} \tag{4-91}$$

where the matrix $\underline{H}$ is replaced by $\underline{HT}$ with the equivalent measurement noises uncorrelated. Potter's method can then be applied to an r-vector measurement by processing sequentially one at a time.

Before we apply essentially these arguments for implementing the measurement update of the square root covariance matrix $\underline{W}$, we recall trivial facts from matrix theory.

Any real symmetric matrix $\underline{A}$ with eigenvalues x_i and eigenvectors $\underline{a}_i$ can be written as

$$\underline{A} = \Sigma\, x_i\, \underline{a}_i\, \underline{a}_i^T \tag{4-92}$$

It immediately then follows the

Property: If $\underline{A} = \underline{A}^T \geq o$ then

$$\underline{A} = \underline{B}\underline{B}^T \tag{4-93}$$

with

$$\underline{B} = [\underline{b}_1 \cdots \underline{b}_n\ ]\ ,\ \underline{b}_i = x_i^{1/2}\ \underline{a}_i\ ,\quad i = 1, \ldots, n \tag{4-94}$$

If any $x_i = o$, some of the $\underline{b}_i$ are $\underline{o}$ and the remaining are linearly independent (Standard Linear Algebra result).

Remark: In the above fact the matrix $\underline{B}$ was constructed to be unique. Obviously it's not the unique solution to equation (4-93).

We have chosen to implement the covariance measurement update equation via

$$\underline{P} = [\underline{M}^{-1} + \underline{H}^T\underline{R}\underline{H}]^{-1} \tag{4-95}$$

Since $\underline{R} > o$, $\underline{H}^T\underline{R}^{-1}\underline{H} \geq o$ and so

$$\underline{B} = \underline{H}^T\underline{R}^{-1}\underline{H} = \sum_{i=1}^{\ell} \underline{b}_i\underline{b}_i^T \tag{4-96}$$

where ℓ is the rank of $\underline{B}$, $\ell \leq n$ $(n = 4)$ and the $\underline{b}_i$ are linearly independent.

Recalling from matrix calculus the

Fact: $$[\underline{A}^{-1} + \underline{b}\underline{b}^T]^{-1} = \underline{A} - \underline{A}\underline{b}(\underline{I} + \underline{b}^T\ \underline{A}\underline{b})^{-1}\ \underline{b}^T\ \underline{A} \tag{4-97}$$

Proof: Algebra manipulation.

Then the square root algorithm follows by applying sequentially ℓ times Potter's algorithm to the equation

$$\underline{W}_j \underline{W}_j^T = [(\underline{S}_{j-1} \underline{S}_{j-1}^T)^{-1} + \underline{b}_j \underline{b}_j^T]^{-1} \qquad j = 1, \ldots, \ell \qquad (4\text{-}98)$$

and where

$$\begin{cases} \underline{S}_o = \underline{S} & (4\text{-}99) \\ \underline{S}_j = \underline{W}j \qquad j = 1, \ldots, \ell - 1 & (4\text{-}100) \\ \underline{W} = \underline{W}_\ell & (4\text{-}101) \end{cases}$$

We remark that if we had used directly Andrews' argument we would have to do 2N iterations at each stage to update the covariance square root of $\underline{P}$.

In our particular problem ℓ is always = 2, while N (number of array elements) may be very large (e.g. 20).

Step 3) Time Update

We can apply directly the method as developed by Schmidt, e.g. Kaminski et alt. [E-4].

In Appendix B we present the complete routine.

Chapter V

Mathematical Analysis of the Optimum Receiver

In this chapter we analyse the receiver structure derived in 4.4 and Appendix A.

Our goal is to have a certain understanding of how the filter behaves in order to be able to interpret the simulation results.

This analysis is carried in two steps. First the estimator equation is considered and several block diagrams are drawn. The polar coordinates case is explored more in depth and we conclude that in the receiver we can distinguish essentially two "phase locked loops", one trying to lock on the waveform

$$\phi_1(t) = \frac{w_c}{c} R(t) \tag{5-1}$$

and the second one on

$$\phi_2(t) = \frac{w_c}{c} \sin\Theta(t) \tag{5-2}$$

We remark that this analogy gives hindsight on the expected behavior of the several receivers.

Secondly, we analyse the covariance propagation and conclude that a steady state regime is reached (polar coordinates case). As pointed out earlier, this steady state regime is a desirable characteristic of the filter since it avoids the computational burden of propa-

gating the covariance equation. But even in the real world where the simplified hypothesis of the analysis do not necessarily hold this shows that at least after a transient the filter will essentially stabilize its covariance matrix and so in each application we only do need to propagate the covariance matrix for the initial transient period.

5.1 Estimator Equation

We summarize the results from section 4.4 and Appendix A:

$$\underline{\hat{x}}_k = \underline{\bar{x}}_k + \underline{P}_k \underline{H}_k^T \underline{\tilde{R}}_k^{-1} [\underline{r}_k - \underline{h}(\underline{\bar{x}}_k, k)] \tag{5-3}$$

$$\underline{\bar{x}}_{k+1} = \underline{\tilde{f}}(\underline{\hat{x}}_k) \tag{5-4}$$

where

$$\underline{B} = \underline{H}^T \underline{R}^{-1} [\underline{r} - \underline{h}] = \begin{bmatrix} Z1 \\ o \\ Z3 \\ o \end{bmatrix} \tag{5-5}$$

$$Z1 = \frac{2\sqrt{P}\,T}{N_o} \sum_{i=1}^{N} d_{1i}\{\cos \bar{a}_i(r_{c_i} - h_{c_i}) - \sin \bar{a}_i(r_{s_i} - h_{s_i})\} \tag{5-6}$$

$$Z3 = \frac{2\sqrt{P}\,T}{N_o} \sum_{i=1}^{N} d_{3i}\{\cos \bar{a}_i(r_{c_i} - h_{c_i}) - \sin \bar{a}_i(r_{s_i} - h_{s_i})\} \tag{5-7}$$

d_{1i}, d_{3i}, are given in Table A-1,

$$\bar{a}_i = \frac{w_c}{c} [p_i \sin\bar{\Theta} - \bar{R}] \tag{5-8}$$

In figure (5-1) we present the general block diagram of the estimator processor in the general setting of equations (5-6) and (5-7). Figures (5-2) and (5-3) particularize for the polar and rectangular coordinate systems, respectively.

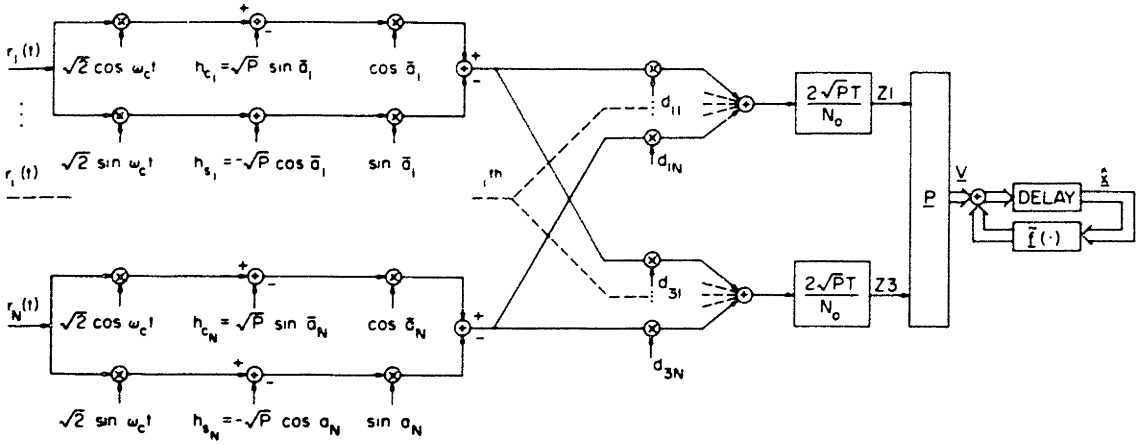


Fig. 5-1

Estimator Processor: General Case

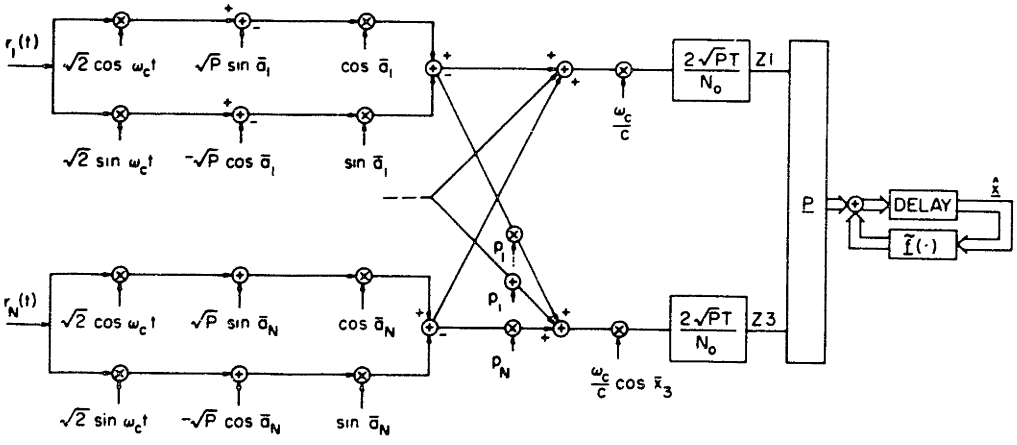


Fig. 5-2

Estimator Processor: Polar System

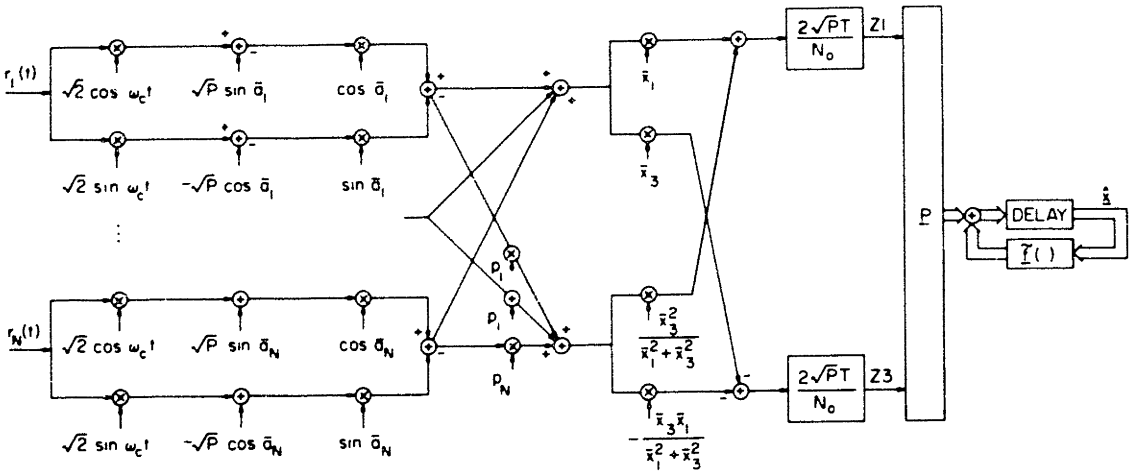
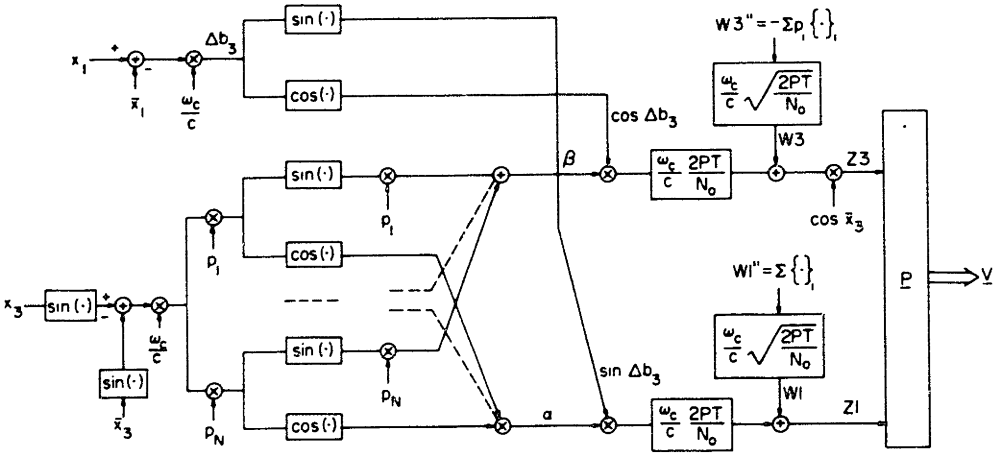


Fig. 5-3

Estimator Processor: Rectangular Coordinate



$\alpha=\Sigma\cos(p_i\Delta b_i)$

$\beta=\Sigma p_i\sin(p_i\Delta b_i)$

$\Delta b_i=\frac{\omega_c}{c}(\sin x_3-\sin\bar{x}_3)$

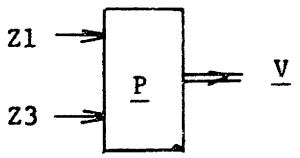
$\Delta b_3=\frac{\omega_c}{c}(x_1-\bar{x}_1)$

$\{\cdot\}_i=\cos\bar{\alpha}_i\omega_{ci}-\sin\bar{\alpha}_i\omega_{si}$

Fig. 5-4

Mathematical Model: Polar System

Figure (5-1) (general case) shows that the filter performs two beams Z1 and Z3 each one being a given weighted sum of the individual sensor outputs. These two beams are input to a beam combiner which essentially consists of the matrix operation $\underline{P}$



generating the input (control) $\underline{V}$ to the receiver's copy of the dynamical system.

Performing the $\underline{P}$ matrix multiplication

$$\begin{cases} V_1 = P_{11}Z1 + P_{13}Z3 \\ V_2 = P_{12}Z1 + P_{23}Z3 \\ V_3 = P_{13}Z1 + P_{33}Z3 \\ V_4 = P_{14}Z1 + P_{34}Z3 \end{cases} \tag{5-9}$$

we see that the input V_i to each state variable processor equation is a weighted sum of the two beams Z1 and Z3, the weights being the elements of the first and third column of the covariance matrix.

Figure (5-2) and figure (5-3) show the two beam performing operations in each one of the coordinate systems.

In the polar case each beam is essentially performed in an "independent" channel. Z1 corresponds to a conventional delay and sum beam (i.e. Van Trees [A-11]) and Z3 to a "difference" beam. This "independence" is not complete at this stage since a coupling exists through $\bar{a}_i$, $i = 1, \ldots, N$ and the factor $\cos \bar{x}_3$. In the rectangular case both channels intervene in the construction of both beams.

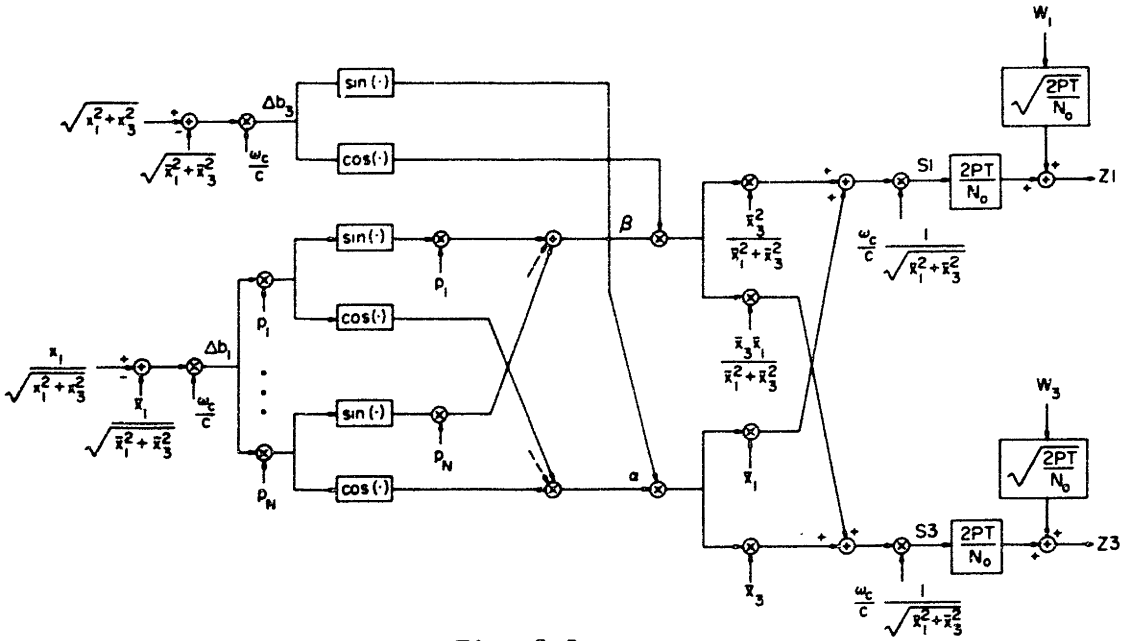


Fig. 5-5

Mathematical Model I: Rectangular System

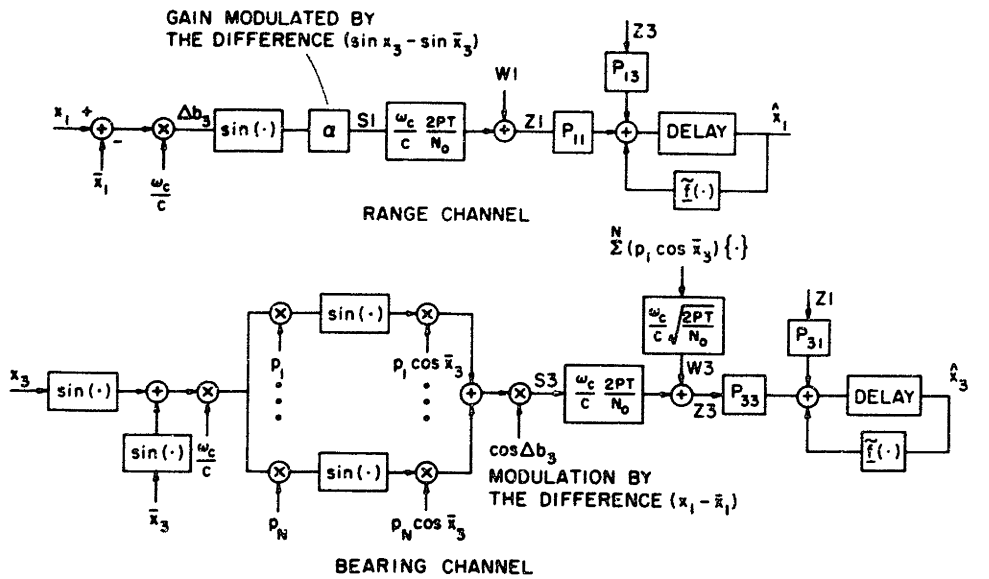


Fig. 5-6

Range and Bearing Channel in Polar Coordinates

In Appendix A we derived a mathematical model for the filter which is more suitable to pursue the analysis. Essentially we considered that

$$Z1 = \frac{2PT}{N_o} S1 + \sqrt{\frac{2PT}{N_o}} W1$$
$$Z3 = \frac{2PT}{N_o} S3 + \sqrt{\frac{2PT}{N_o}} W3 \qquad (5\text{-}10)$$

where S1, S3, W1, W3 are given in Table A-3 and equation (A-20).

Figures (5-4) and (5-5) show the block diagrams in polar and rectangular coordinates respectively.

In (5-10) each beam is decomposed as the weighted sum of a "signal term" and an additive "noise disturbance", the weights being proportional to the signal to noise ratio (SNR). We see that if

$$\frac{2PT}{N_o} > 1$$

the coefficient of the signal term is larger than the coefficient of the noise term. But if

$$\frac{2PT}{N_o} < 1$$

exactly the opposite occurs. Hence a degradation of performance is expected for values of $\frac{2PT}{N_o}$ around 1.

5.1.1 Polar Coordinates

We now concentrate in the polar case. From its mathematical model we see that the two beam forming operations can be decomposed in the generation of two "signals" S1 and S3 and two "additive noise disturbances" W1 and W3.

If we assume that the linearized assumption is valid, the errors will be small and

$$\cos \Delta b_3 \sim 1 \tag{5-11}$$

$$\cos p_1 \Delta b_1 \sim 1 \tag{5-12}$$

from which it follows that

$$\alpha \sim N = \text{number of sensors of the array} \tag{5-13}$$

and the real active elements in both channels are the sin(.)elements, respectively $\sin(p_i \Delta b_1)$, $i = 1, \ldots, N$ which perform β and $\sin\Delta b_3$.

This argument shows that S1 contains essentially information about the difference $x_1 - \hat{x}_1$ while S3 has essentially the available information about the difference between the true bearing and the bearing estimate. So we see that the Z3 channel is mainly a bearing channel while the Z1 is a range channel with a weak coupling, e.g. modulation of β by $\cos \Delta b_3$ in the bearing channel and through the modulated gain α in the range channel.

Z1 and Z3 are as noted before combined by the matrix operation $\underline{P}$ to give a weighted input to each state variable equation, the weights being the covariance elements which represent the coupling between the several channels. At this stage we are not able to see how strong or weak this coupling is but if the cross covariances are small when compared with the variances (e.g. $P_{13} << P_{11}$) and if the covariance elements reach a steady state then a decoupling of the range channel and of the bearing channel effectively takes place.

To see this, we reorganize figure (5-4) in figure (5-6). In figure (5-6) it becomes apparent that in the linear regime the modulation

effects by cos Δb_3 in the bearing channel and the modulation of the range channel gain α are negligible. Also it is now obvious that the main coupling between the two channels, if it exists, is through the P_{11}, P_{13} and P_{33} gains.

For the two remaining state variables, figure (5-7) shows the respective processors.

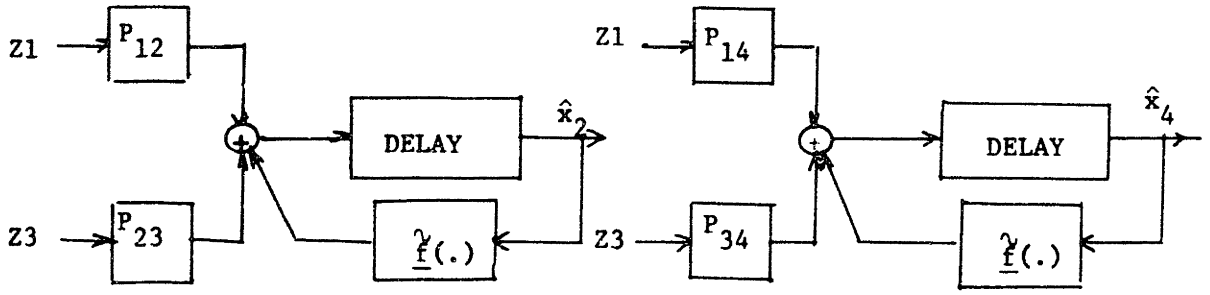


Fig. 5-7

In figure (5-8) we show in detail the last section of figure (5-6):

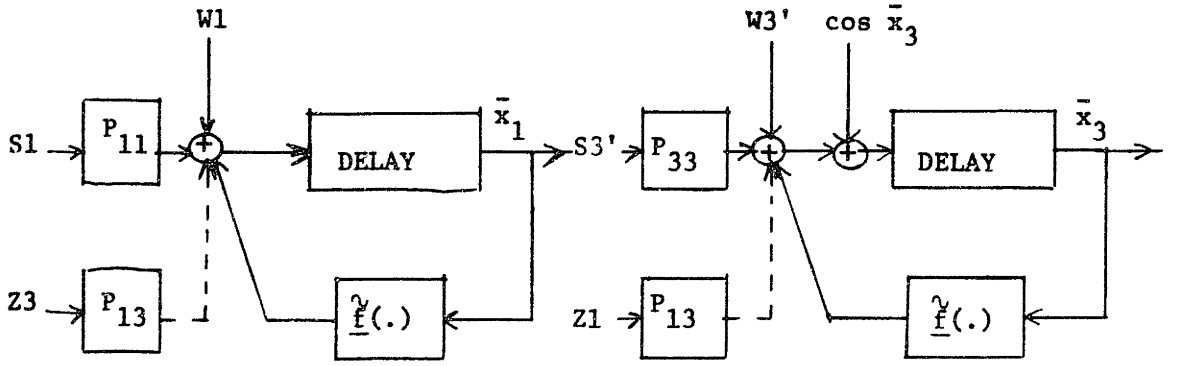


Fig. 5-8

We observe first that in dashed lines we represented the coupling effect (assuming the gains P_{11}, P_{33} reach a steady state value) between the two channels.

If we can show that

$$P_{11} >> P_{13}$$
$$P_{33} >> P_{13} \qquad (5\text{-}14)$$

than the two channels are really decoupled.

We see that when the relations (5-14) hold, figure (5-8) shows that the range channel and the bearing channel perform like phase locked loops tracking the waveforms

$$\phi_1(t) = \frac{w_c}{c} R(t) \tag{5-15}$$

and

$$\phi_3(t) = \frac{w_c}{c} \sin\theta(t) \tag{5-16}$$

as previously indicated.

We remark that in figure (5-8) we have explicitly indicated the cos $\bar{x}_3$ multiplying factor to show that the performance of the bearing "phase lock loop" is degraded for increasing bearing angles. In figure (5-6) we have associated the cos $\bar{x}_3$ factor with the position elements p_i, showing that the net effect of increasing the bearing angle is the reduction of the array length.

The rate channels of figure (5-7) are essentially coupled to the previous ones. This is obvious since no direct information is available at the received signal about the rates. The only way the filter can estimate the range rate is looking at the modulation of the received waveform induced by the target dynamics. At each rate channel we have present both beams Z1 and Z3 weighted by the respective crosscovariances between the error of x_1 and x_j and the errors of x_3 and x_j respectively, where j = 2, 4.

One suspects that in each case, in the linear analysis:

$$|P_{12}| >> |P_{23}| \tag{5-17}$$

and

$$|P_{34}| >> |P_{14}| \tag{5-18}$$

but at this stage this can only be verified by the simulation results.

Remark: We should point out that because the whole filter corresponds to a linearized analysis, the above arguments are only valid when we are indeed in a linear region of performance. We also see that since the available copy of the source at the receiver is an "estimated" copy, this means that the performance of the different channels is degraded by the errors on the estimates. But we suspect that in the linear region these will be small.

To finish the analysis of the polar coordinates receiver, we make a brief error analysis.

From figure (5-6) we can see that if $\Delta b_3 > o$ ($x_1 > \hat{x}_1$) then $S1 > o$ and for sufficiently high signal to noise ratio $Z1 > o$ the estimate moves up in the range channel to track the right value.

Conversely, if $\Delta b_3 < o$ the same argument shows $Z1 < o$ and the estimate will be reduced and so once again the filter moves in the right direction.

This whole argument is true whenever

$$|\Delta b_3| < \pi \tag{5-19}$$

In fact if we consider the estimate range parameter space we see that there are equilibrium points in the range space where the filter will lock in. These equilibrium points are half of a wavelength apart of the true range.

This is essentially due to the fact that the net effect of Δb_3 is through the sin (.) function.

In figure (5-9) we show diametrically this effect.

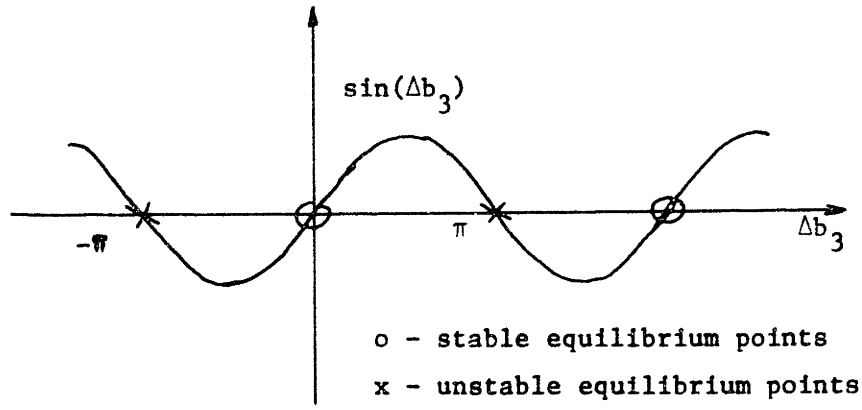


Fig. 5-9

Of these equilibrium points the ones apart $2k\lambda^*$ of the true range are stable and the ones apart $(2k + 1)\frac{\lambda}{2}$ are unstable equilibrium points in the sense that a small perturbation will make the filter to lock on the nearest stable equilibrium point.

These phenomena just parallel the similar behaviour of a phase lock loop (see Van Trees [A-9]).

From this analysis we conclude that the primary lockin range is limited to an interval centered on the true range and width of one wavelength.

5.1.2 Rectangular Coordinates Receiver

We can rearrange figure (5-5), i.e. the derived mathematical model of the rectangular coordinates case , in a similar way. This is done in figure (5-10).

In the rectangular coordinates case for small bearing angles, i.e. $x_1 << x_3$, we can neglect, in the linear analysis, in the upper path of figure (5-10) the summand S_{31} compared to S_{33}. So the bottom path

---

* λ is the wavelength, i.e. $\lambda w_c = 2\pi c$

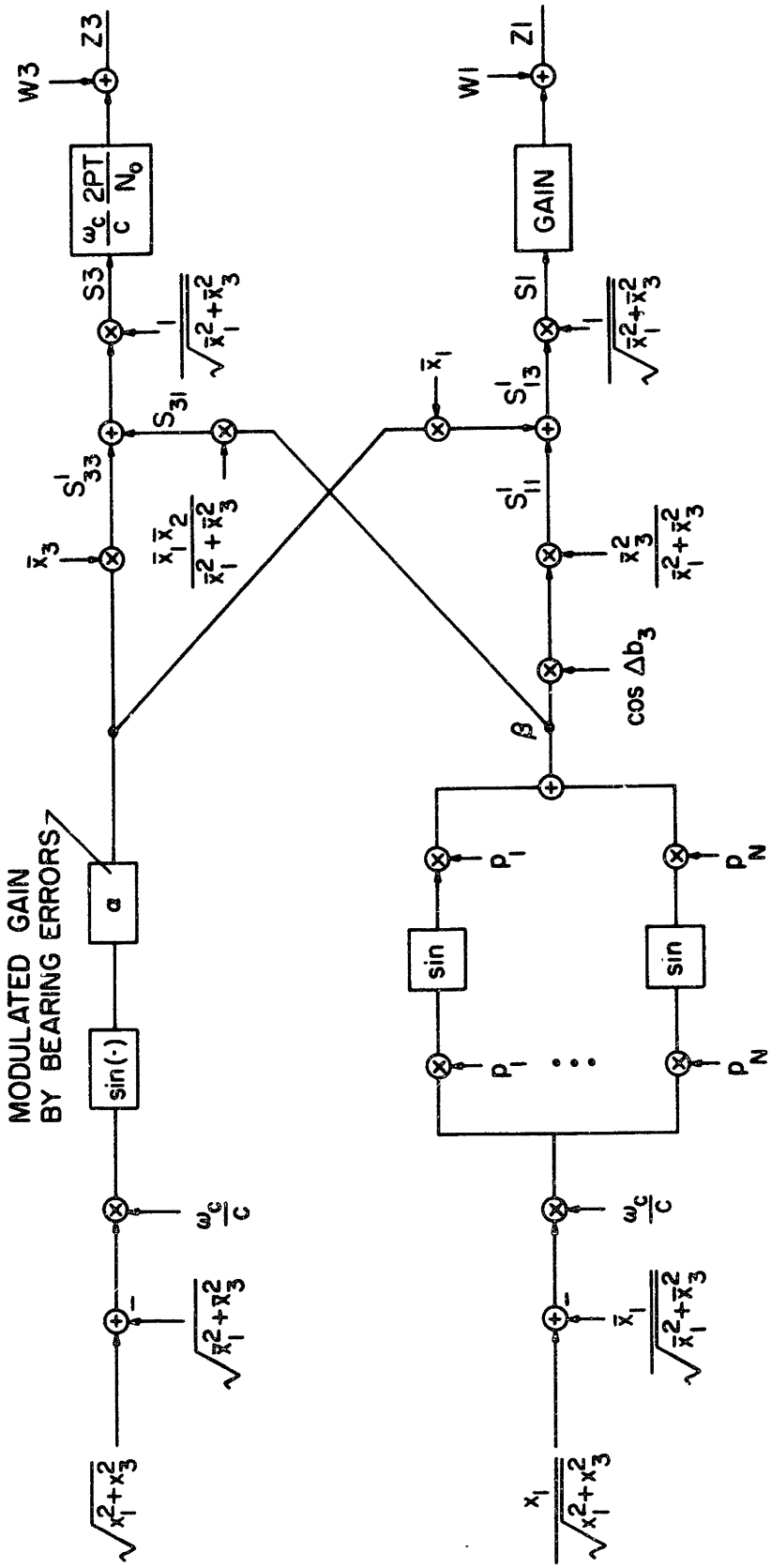


Fig. 5-10

Mathematical Model II: Rectangular Coordinate

only influences the upper one through the coupling between the covariance gains and through the dependence on the estimates (which in a linear analysis can be translated qualitatively in terms of a degraded performance).

We see that now the essential dynamics in each path are governed by Δb_1 and Δb_3 with both depending on x_1 and x_3.

But we still can argue that Z3 is essentially controlled by Δb_3 (if $x_1 << x_3$) and that

$$\sqrt{x_1^2 + x_3^2} - \sqrt{\bar{x}_1^2 + \bar{x}_3^2} \sim |x_3| - |\bar{x}_3| \tag{5-20}$$

So the upper path represents essentially a phase lock loop similar to the range channel in polar coordinates, trying to lock on the waveform

$$\psi_1(t) = \frac{w_c}{c} x_3 \tag{5-21}$$

We call it the Y-channel. We remark again that the Y-channel will have a degraded performance with respect to the range channel, and that this analysis is essentially valid only for $x_1 << x_3$.

On the bottom path Z1 is controlled by both Δb_3 and Δb_1, being difficult to compare the two quantities S_{11} and S_{13}.

It is easily seen that S1 << S3 for the geometry under consideration and so even if

$$P_{13} \sim P_{33} \tag{5-22}$$

we can neglect the coupling to the Y-channel induced by Z1. The only coupling remaining is then through the covariance gain P_{33}.

For the bottom path which we will call X-channel, both beams are processed and it is hard a priori to pursue the analytical argument.

5.2 Covariance Equation

Recalling the covariance measurement update equation in the form

$$\underline{P} = [\underline{M}^{-1} + \underline{H}^T \underline{R}^{-1} \underline{H}]^{-1} \tag{5-23}$$

from Appendix A we have

$$\underline{B} = \underline{H}^T \underline{R}^{-1} \underline{H} = \begin{bmatrix} H1 & o & H2 & o \\ o & o & o & o \\ H2 & o & H3 & o \\ o & o & o & o \end{bmatrix} \tag{5-24}$$

with H1, H2, H3 given in the Table A-4.

The special form of the matrix $\underline{B}$ just illustrated the fact that on the received waveform the receiver has only present direct information about x_1 and x_3. The fact that the filter can effectively estimate x_2 and x_4 is a result, as pointed out many times before, of the modulation induced in the signal by the target's dynamics.

We have seen before that the filter's processing of the received waveform consisted in filtering two beams Z1 and Z3 where essentially $Z1(\Delta b_1)$ and $Z3(\Delta b_3)$ and that these two beams were the inputs of 4 dynamical systems. We argued that excluding the coupling through the covariance gains no additional coupling existed between the range channel and the bearing channel (in the linearized analysis the fact that the copy of the dynamical system at the receiver depends on the actual estimates can be thought of as a degradation in the performance).*

---

* As a final general remark on the covariance equation, we observe that in the EKF (polar or rectangular), as shown by Table A-4, the covariance equation (5-23) only depends on the estimated values of the state vector and not on its actual values. In the MAP filter the measurement update equation explicitly depends on the actual values of the state vector.

To pursue the analysis, we concentrate first on the:

5.2.1 Extended Kalman Filter, Polar Coordinates Case

The covariance measurement update equation in this case has

$$H2 = o \tag{5-25}$$

in the $\underline{B}$ matrix.

To get an upper bound on the values of $\underline{P}$, we imagine the worst situation where we assume that we rely on nothing in the previous estimate, i.e. make

$$\underline{M}^{-1} = \underline{o} \tag{5-26}$$

(This could, for example, be the starting attitude).

Because $\underline{B}$ is singular, this says nothing of interest about $\underline{P}$. But if we recall our previous discussion about the way the filter processes the data (beams Z1 and Z3) and that the important dynamics of the filter relate essentially to Δb_1 and Δb_3, i.e. to the range and bearing estimate errors, this suggests that we think of our problem as a 2-state variable problem

$$\begin{bmatrix} x_1 \\ x_3 \end{bmatrix}$$

Then we get from equation (5-23) and (5-24) in the "worst" situation

$$\begin{bmatrix} P_{11} & P_{13} \\ P_{13} & P_{33} \end{bmatrix} \simeq \begin{bmatrix} H1 & o \\ o & H3 \end{bmatrix}^{-1} \tag{5-27}$$

or

$$\begin{cases} P_{13} \simeq o & (5\text{-}28) \\ P_{11} \simeq \frac{1}{H1} \triangleq B1 & (5\text{-}29) \\ P_{33} \simeq \frac{1}{H3} \triangleq B3 & (5\text{-}30) \end{cases}$$

So, if the argument holds, we should expect that in the linearized analysis

$$P_{11} \sim < \frac{1}{H1} \tag{5-31}$$

$$P_{33} \sim < \frac{1}{H3} \tag{5-32}$$

and that

$$P_{13} \sim o \tag{5-33}$$

Recalling from Table A-4

$$H1 = \frac{2PT}{N_o} \left(\frac{w_c}{c}\right)^2 N \tag{5-34}$$

$$H3 = \frac{2PT}{N_o} \left(\frac{w_c}{c}\right)^2 \cos^2 \bar{x}_3 \sum_1^N p_i^2 \tag{5-35}$$

we see that essentially the bounds B1 and B3 depend on the signal to noise ratio, the array height (or number of elements) and in the bearing channel also on the bearing angle.

We show below in figure (5-11) graphs of B1 and B3 when the several parameters vary.

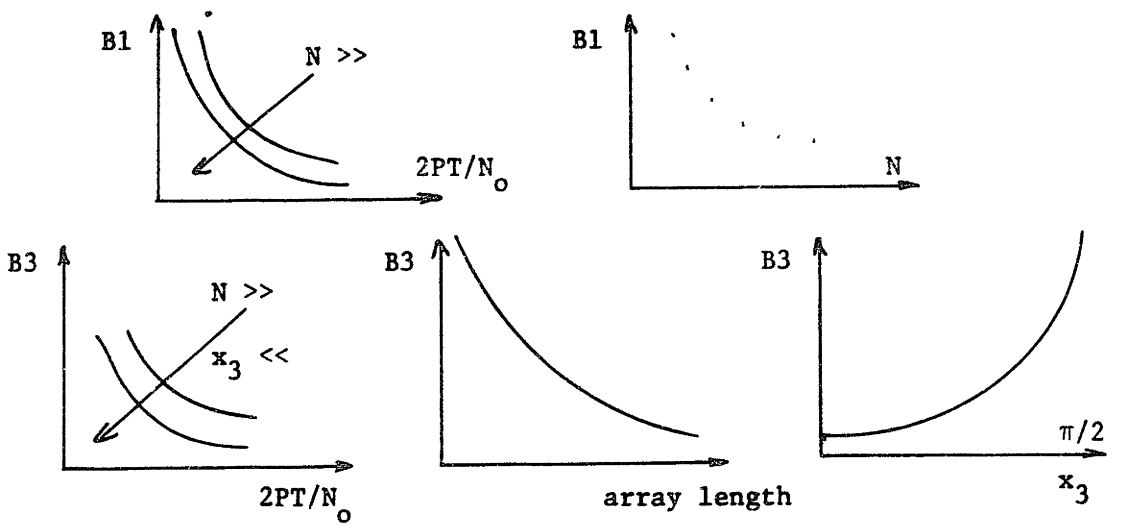


Fig. 5-11

Obviously we expect that the agreement between the actual behaviour of P_{11} and P_{33} and the sketched bounds B_{11} and B_{33} is better when the values of the parameters imply a good filter performance (e.g. high signal to noise ratio (SNR), long arrays, small bearing angles, etc.).

The above heuristic argument can be made more rigorous by considering the expanded form of the equations of the covariance elements P_{11}, ..., P_{44} and by taking limits of these values when for example $M_{11} \to \infty$.

We present now a

Numerical example

$$\frac{2P}{N_o} = 10$$

$T = .5$ sec

$c = 5000$ feet/sec

$w_c = 2\pi f_c$

$f_c = 100$ cycles/sec

$N = 20$ elements

array length = 500 feet

$x_3 = o$

Then

$$B1 \simeq .625$$
$$B3 \simeq 2.5 \times 10^{-5}$$

The numerical values of B1 and B3 and physical consideration lead one to suspect that B1 and B3 are tight bounds of P_{11} and P_{33} respectively. In fact the simulation results agree with the above quantities, i.e.

$$P_{11} < B1 \tag{5-36}$$

$$P_{33} < B3 \tag{5-37}$$

and they show that the covariances remain practically unchanged after a finite number of iterations.

Nevertheless the strict inequalities in (5-36), (5-37) justify in terms of a gain in performance the actual work of propagating the covariance matrix by the Kalman filter in order to evaluate the final values of the covariance elements.

5.2.2 Extended Kalman Filter, Rectangular Coordinates

By an argument similar to the one used in the previous paragraph, we get in the "worst" situation

$$\begin{bmatrix} P_{11} & P_{13} \\ P_{13} & P_{33} \end{bmatrix} = \begin{bmatrix} H1 & H2 \\ H2 & H3 \end{bmatrix}^{-1} = \frac{1}{H4} \begin{bmatrix} H3 & -H2 \\ -H2 & H1 \end{bmatrix} \tag{5-38}$$

where

$$H4 = H1\ H3 - H2^2 \tag{5-39}$$

From Table A-4 we get

$$H4 = \left[\frac{2PT}{N_o} \left(\frac{w_c}{c} \right)^2 \frac{1}{(.)} \right]^2 N \sigma \bar{x}_3^{-2} \tag{5-40}$$

$$B1 \triangleq \frac{H3}{H4} = \frac{\frac{\bar{x}_1^{-2}\bar{x}_3^{-2}}{(.)^2} \sigma + \bar{x}_3^{-2} N}{DEN} \tag{5-41}$$

$$B3 \triangleq \frac{H3}{H4} = \frac{\frac{\bar{x}_3^{-4}}{(.)^2} \sigma + \bar{x}_1^{-2} N}{DEN} \tag{5-42}$$

$$B13 \triangleq -\frac{H2}{H4} = \frac{\left[\frac{\bar{x}_3^2}{(.)^2}\sigma - N\right]\bar{x}_1\bar{x}_3}{DEN} \tag{5-43}$$

where

$$(.) = \bar{x}_1^2 + \bar{x}_3^2 \quad , \qquad \sigma = \sum_1^N p_i^2$$

$$DEN = \frac{2PT}{N_o}\left(\frac{w_c}{c}\right)^2 \frac{\bar{x}_3^2}{\bar{x}_1^2 + \bar{x}_3^2} N\sigma$$

For a geometry where

$$x_1 << x_3 \tag{5-44}$$

we can approximate the above bounds by:

$$B1 \simeq \frac{\bar{x}_3^2}{\frac{2PT}{N_o}\left(\frac{w_c}{c}\right)^2 \sigma} \tag{5-45}$$

$$B3 \simeq \frac{1}{\frac{2PT}{N_o}\left(\frac{w_c}{c}\right)^2 N} \tag{5-46}$$

$$B13 \simeq \frac{-\bar{x}_1\bar{x}_3}{\frac{2PT}{No}\left(\frac{w_c}{c}\right)^2 \sigma} \tag{5-47}$$

Once again we present a

Numerical example

For a well sized array

N = 20

length = 500 feet

$\sigma \doteq 5 \times 10^5$

and for

$$\frac{2P}{N_o} = 100$$

$$T = .5$$

$$x_3 = 10^4$$

$$x_1 = 10^2$$

we get

$$B1 \simeq 250$$

$$B3 \simeq .0625$$

$$B13 \simeq 2.5$$

We remark that for the broad side array the Y-channel has a performance similar to the range channel in polar coordinates as measured by the bounds (5-46) and (5-29) which are the same.

From the numerical example we suspect that the performance of the X-channel will be worst than the performance of the Y-channel. The simulation results will show that

$$P_{11} << B1 \tag{5-48}$$

and in fact

$$P_{11} > P_{33} \tag{5-49}$$

We see that the general expressions (5-41) to (5-43) give the bounds as functions of the geometry. As a final comment, we observe that for an endfire array the values of B1, B3, B13 are in modulo much larger than the ones corresponding to the broadside array.

5.2.3 MAP Filter

By a similar argument we obtain

$$\begin{cases} B1 = \dfrac{H3}{H4} & (5\text{-}50) \\ B3 = \dfrac{H1}{H4} & (5\text{-}51) \\ B13 = -\dfrac{H2}{H4} & (5\text{-}52) \end{cases}$$

where as before

$$H4 = H1\ H3 - H2^2 \qquad (5\text{-}53)$$

and H1, H2, H3 are given by expressions (A-47) to (A-59). The complexity of these equations make almost useless the heuristic analysis pursued in the previous paragraphs. But two remarks are in order in this case:

1) The dependence of the covariances on the actual errors of the estimates which show that the propagated covariance and so the Kalman gain will tend to follow the actual errors. Physically and in other words, the filter bandwidth is directly modeled by the errors in the estimate.

2) In the expressions of H1, H2, H3 we distinguish essentially two terms, one related to the signal and the other related to the noise, the "signal term" being weighted by the signal to noise ratio while the "noise term" is multiplied by the square root of the signal to noise ratio.

The presence of the additive "noise terms" in H1, H2, H3 show that the filter bandwidth is also directly modeled by the noise. For large SNR this modulation is small. But when the SNR decreases, it may become very significant and cause the filter to diverge. One expects the threshold to be about

$$\frac{2PT}{N_o} \sim 1$$

This discussion ends the mathematical analysis of the Kalman filter. The next chapter presents the results of the simulation runs which support the main conclusions developed in this chapter.

Chapter VI

Simulation Results

In this chapter we present the results of the simulation runs in a digital computer where the receiver structures discussed before were implemented for several geometric configurations and different sets of parameters and initial conditions.

The practical experience with the Kalman filter confirmed the main arguments developed in Chapter 5 when a linearized assumption holds.

Before pursuing we briefly discuss the organization of the chapter. After we explain the criteria used to compare the several structures, we present representative runs showing that the linearized assumption holds in the sense that the actual ensemble mean square errors are small and converge to the corresponding propagated covariances. This justifies that the comparative study of the different structures be carried out by simply comparing the propagated values of the covariance matrix.

Then we concentrate mainly on the polar, EKF, SQRT implementation and study the effect of the several parameters and different initial conditions on the propagated covariance elements. Finally, we discuss in summary the main conclusions drawn from the simulation work.

Performance criteria

The performance is measured by the bias

$$\underline{b}(i) = \frac{1}{M} \sum_{j=i-(M-1)}^{i} [\underline{x}(j) - \underline{\hat{x}}(j)] \tag{7-1}$$

and by the standard deviation as computed by

$$\underline{s}(i) = \left[\left\{ \frac{1}{M} \sum_{j=i-(M-1)}^{i} [x_k(j) - \hat{x}_k(j)]^2 \right\}^{1/2} \right] \quad k = 1, \ldots, 4 \tag{7-2}$$

Our studies are based in Monte Carlo simulations. The processors are compared in terms of their performance and actual regions of convergence.

Representative Runs

Figures 6-1 to 6-9 show several runs of the four processors for the indicated conditions. In all the cases the filters converged.

In fig. 6-1 we summarize a typical situation for the four filters. The bias is effectively removed and they all work in a linear region, i.e.

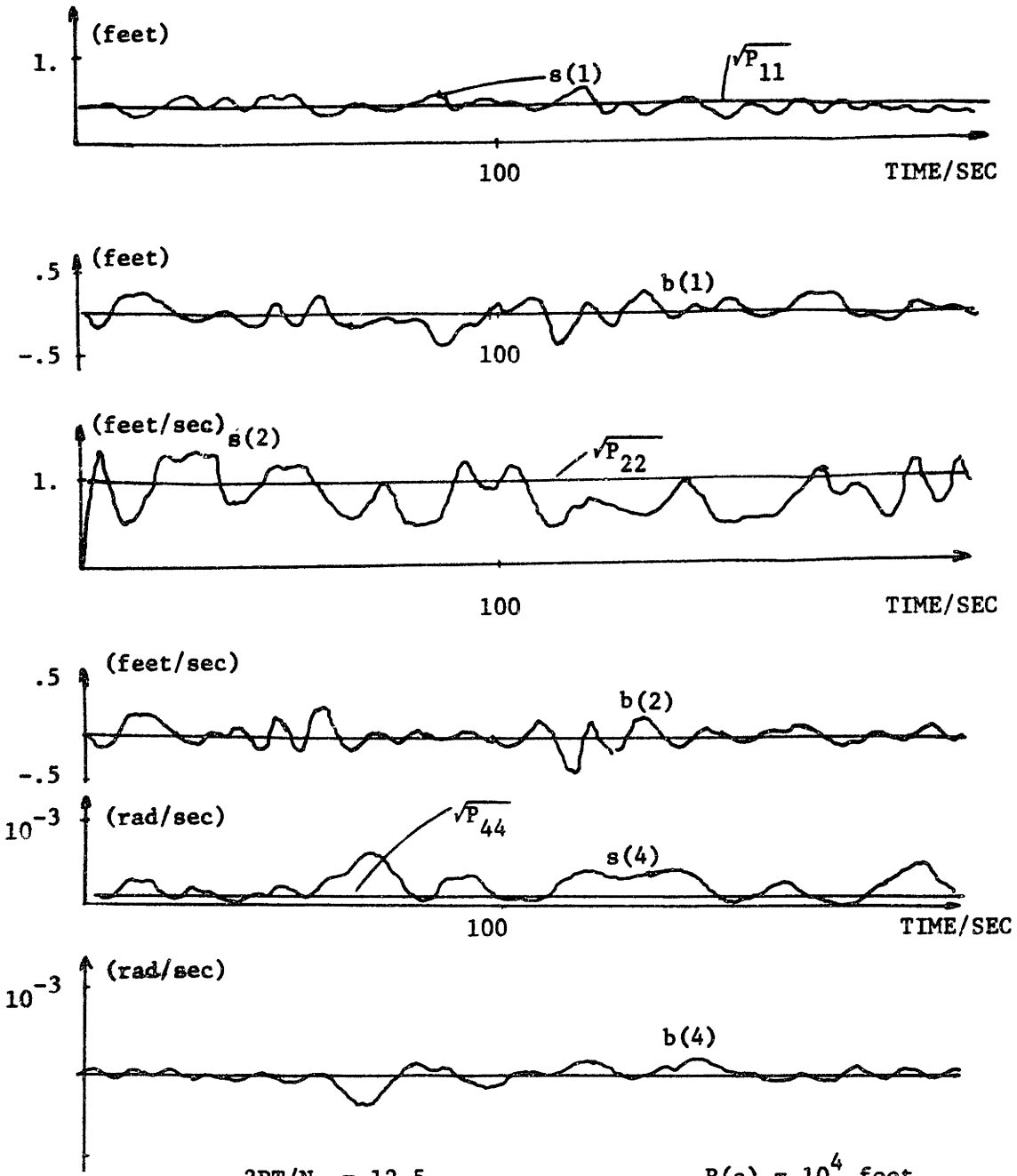
$$s(i) \rightarrow \sqrt{P_{ii}}$$

Figure 6-2 shows in expanded scale the bearing curves for the four receivers. This was done in order to appreciate the spread of the different curves. The small magnitude of $\sqrt{P_{33}}$ coupled with possible computational numerical errors and the relatively small sized sample over which we averaged the results explain the relatively large spread of the curves in fig. 6-2. But they show the errors are small and that their expected magnitude is related to the (optimistic) propagated standard deviations.

Figure 6-3 and fig. 6-4 show in particular for the polar, EKF, SQRT algorithm the bearing channel performance for different initial bearing rates. We remark as an interesting fact the inertia exhibited by the filter.

In figures 6-5 to 6-7 we considered a sizable nonzero initial error on the range and range rate as well as on the X and Y coordinates. It is seen that as long as the range lies inside the primary lock in range the filters converge. As for the range rate the filter is not very sensitive to the initial value as we note from fig. 6-5.

Figure 6-8 compares the X-channel and the Y-channel in the rectangular coordinates case for a broadside geometric configuration.



$2PT/N_o = 12.5$ $\quad$ $R(o) = 10^4$ feet

$Q/T = 2$ $\quad$ $\dot{R}(o) = 30$ feet/sec

$N = 20$

Fig. 6-1

Typical Statistical History for the 4 Nonlinear Filters

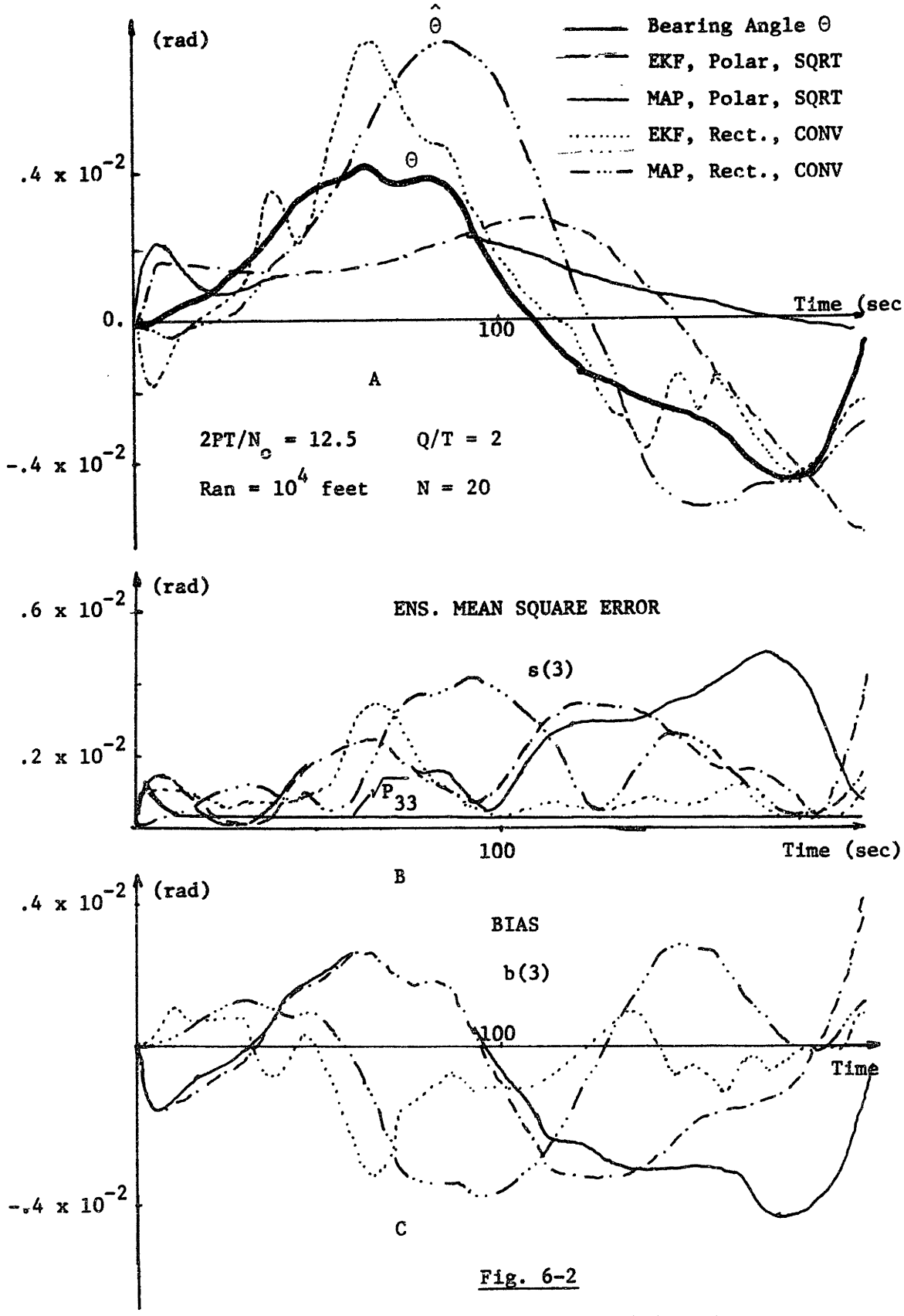
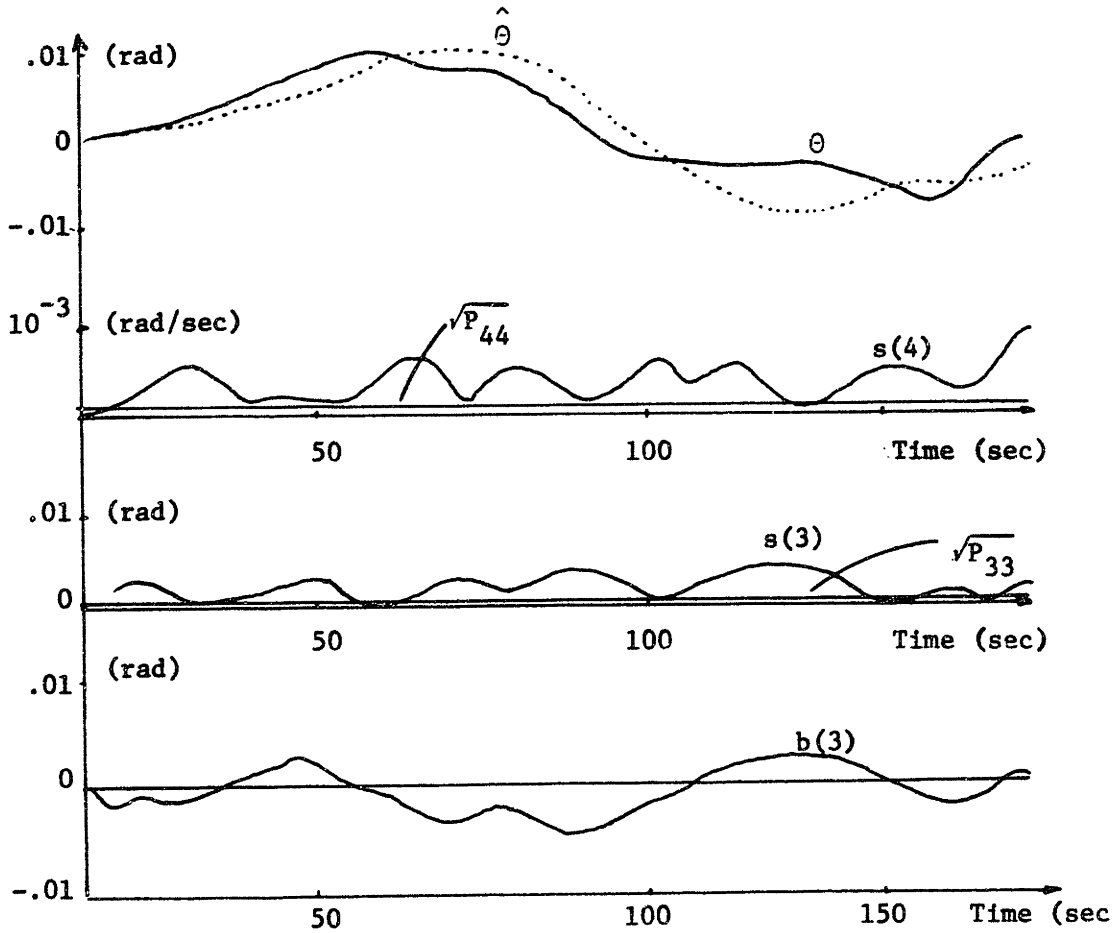


Fig. 6-2

Bearing History in an Expanded Scale



Polar, EKF, SQRT

Fig. 6-3

Bearing Angle History of the EKF, Polar, SQRT

for Small Initial Bearing Rate

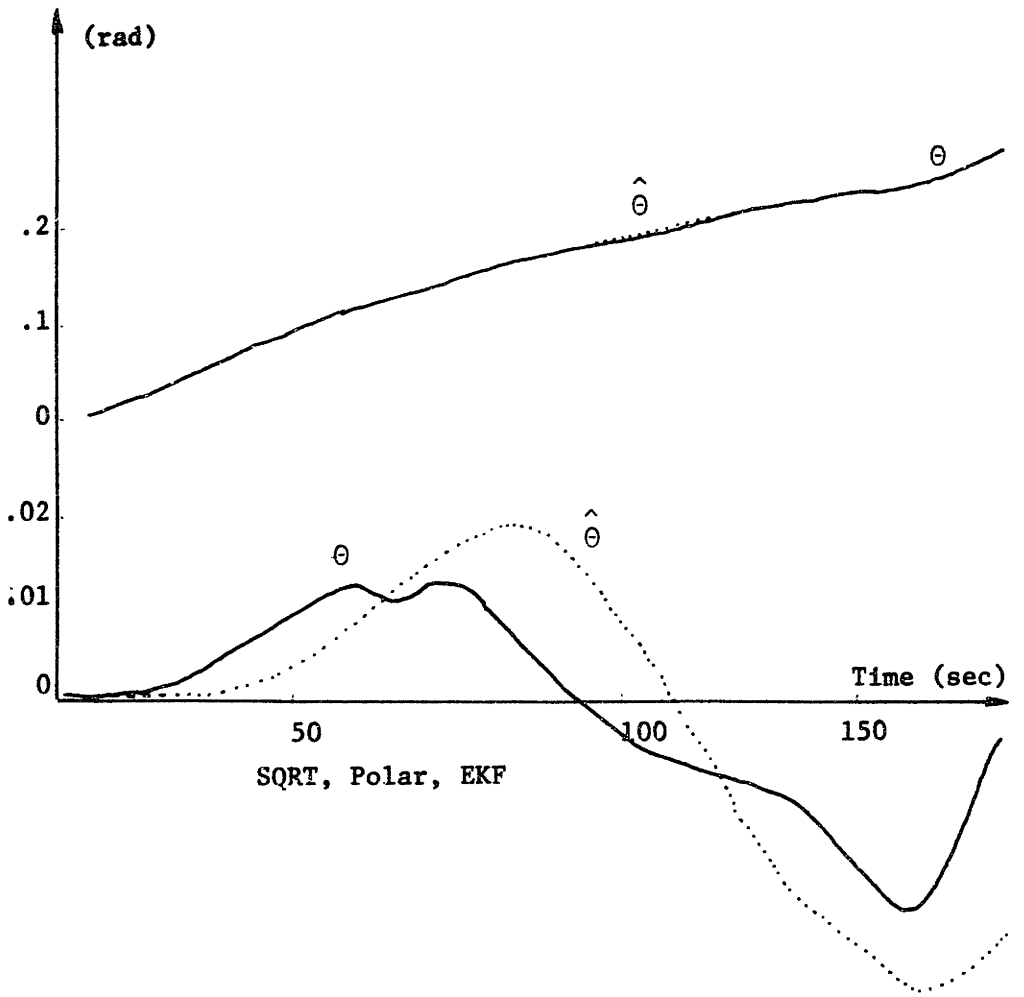


Fig. 6-4

Bearing Angle History for Two Different

Initial Bearing Rates

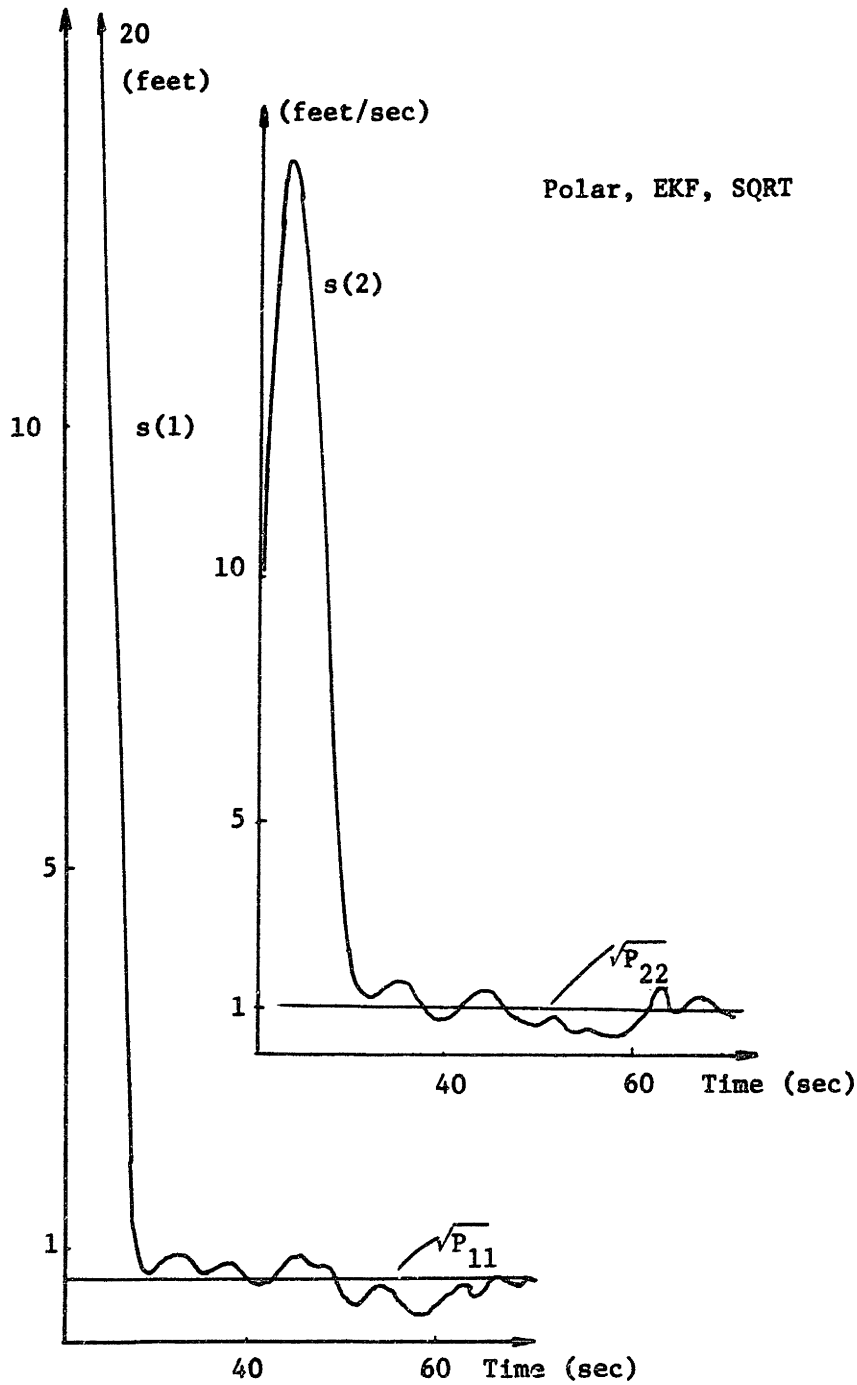


Fig. 6-5

Study of an Initial Error of the Range and Range Rate Estimates on the Standard Deviations

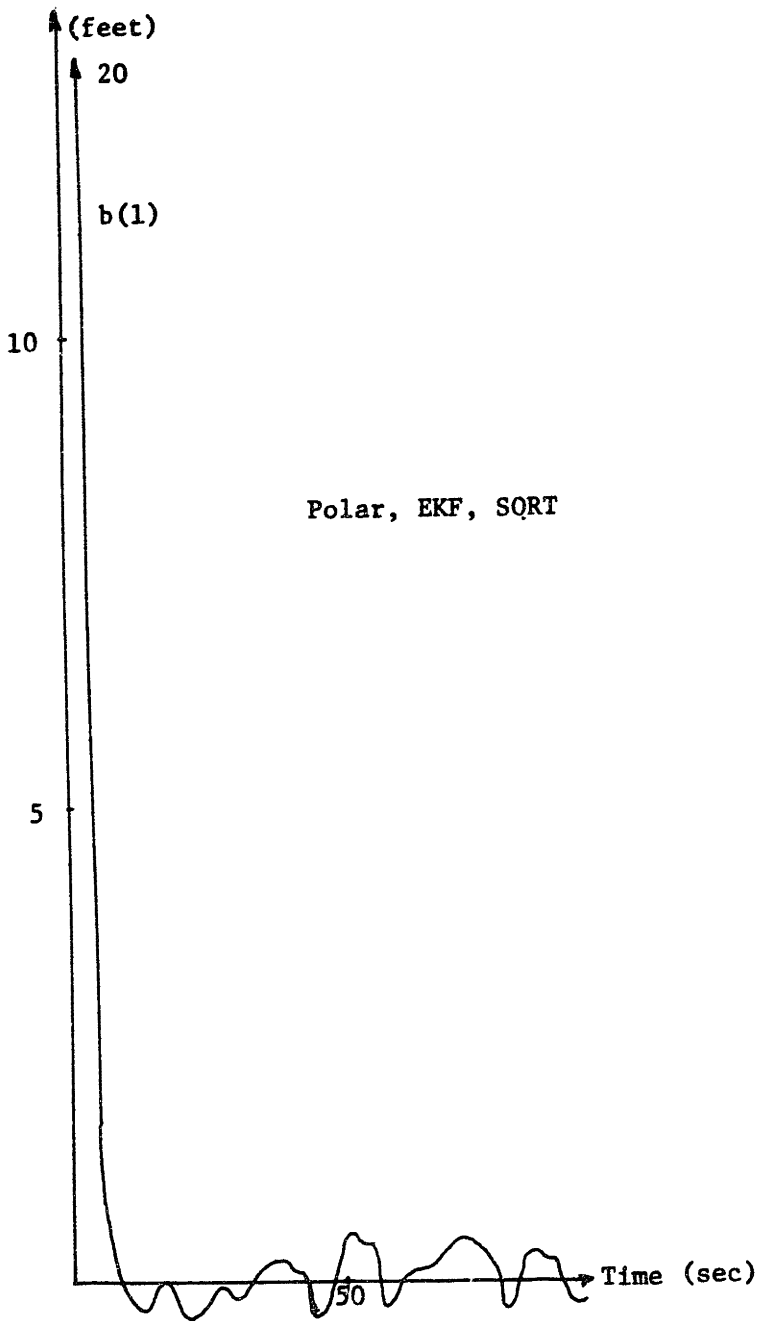


Fig. 6-6

Effect of an Initial Error of the Range Estimate on the Range Bias

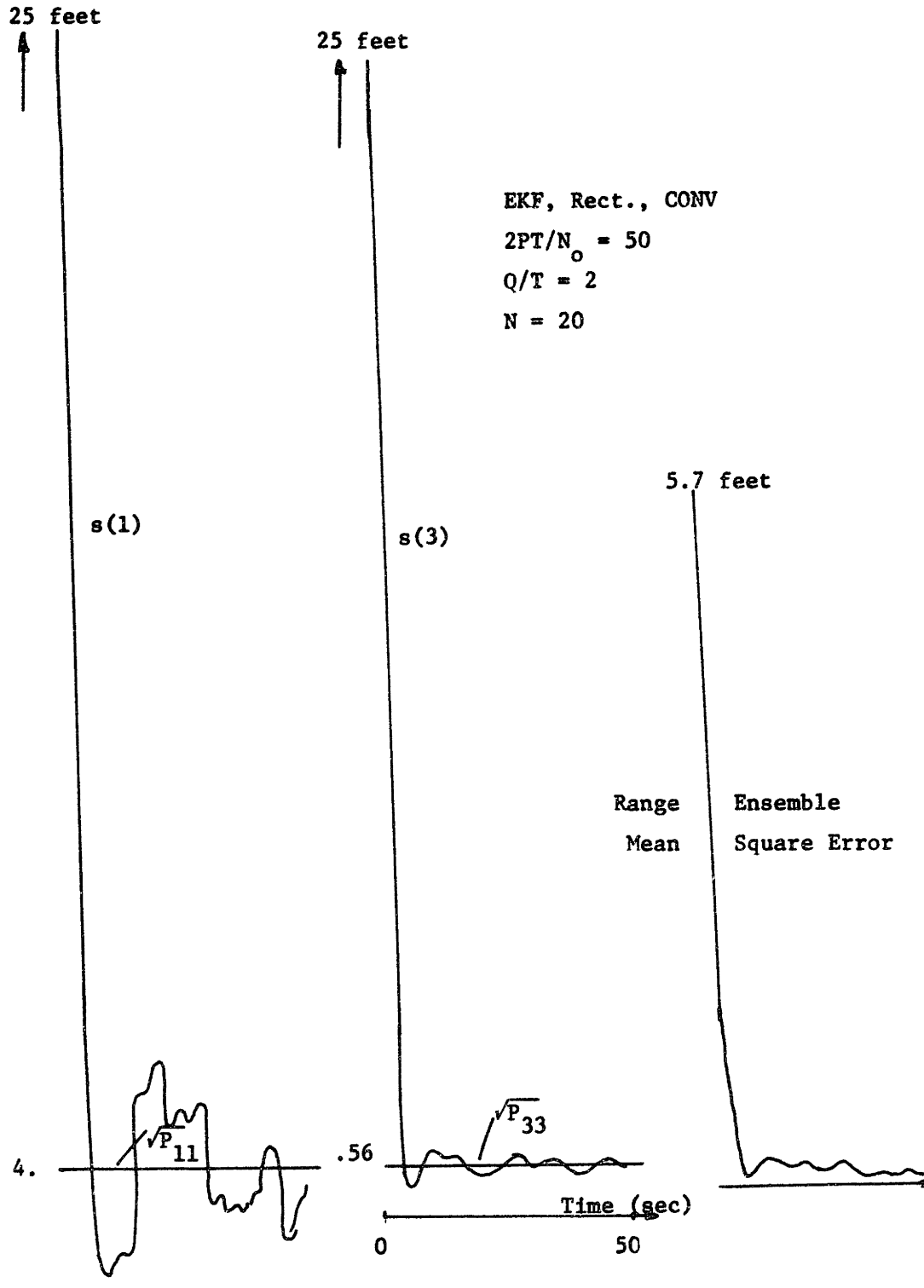


Fig. 6-7

Effect of Initial Error on

Rectangular Coordinate Receiver

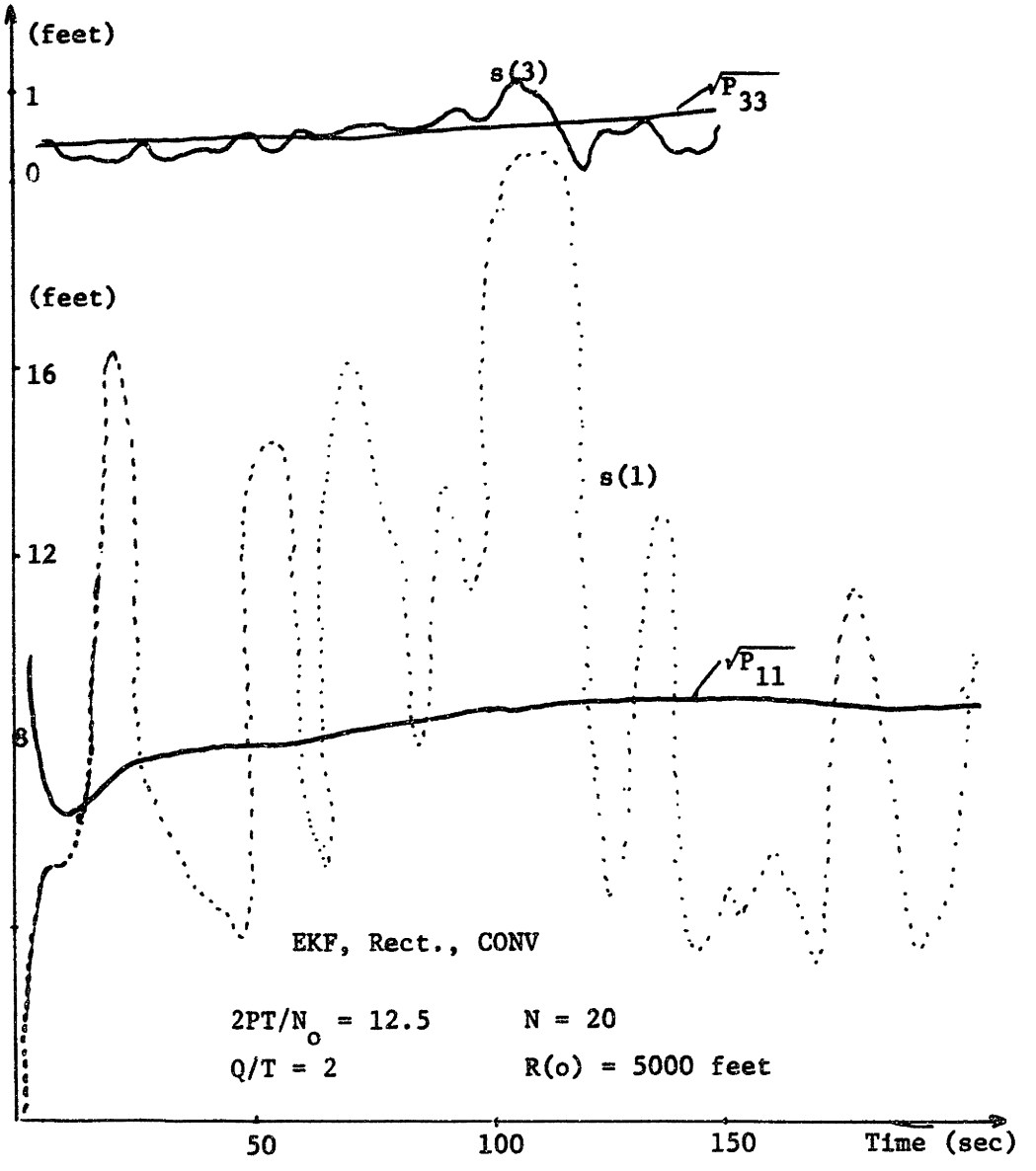
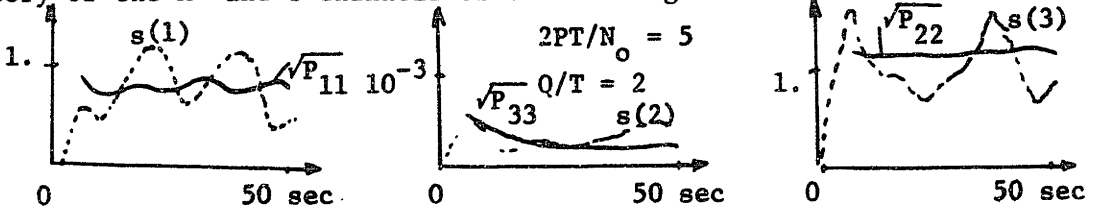


Fig. 6-8

History of the X- and Y-channels of the Rectangular Coordinates Receiver



MAP, Polar, SQRT

Fig. 6-9

Modulation of the Covariances in the

MAP, Polar, SQRT Filter

Figure 6-9 shows in particular the modulated covariances for the MAP filter.

Driving Noise Level 2

Figures 6-10 and 6-11B show the effect on the performance when Q varies.

As Q increases the error covariance and so the Kalman gain also increases. This accounts for the increase of the filter's bandwidth* necessary for the filter to follow the larger source dynamics. As a net effect more noise is allowed in the filter with a correspondent performance degradation.

SNR

Figure 6-11A and 6-12 show the effect of the SNR on the several diagonal elements of the covariance matrix.

The filters exhibited a sharp threshold which, as argued in Chapter 5, is located in the neighborhood of

$$\frac{2PT}{N_o} \sim < 1$$

For several runs the MAP filter diverged. This occurred for

$$\frac{2PT}{N_o} \sim \geq 1$$

this being caused by the fact that the covariance matrix lost its positive semidefinite character.

We remark that this divergence is a true divergence in the sense that once

$$\underline{P} \not\geq \underline{o}$$

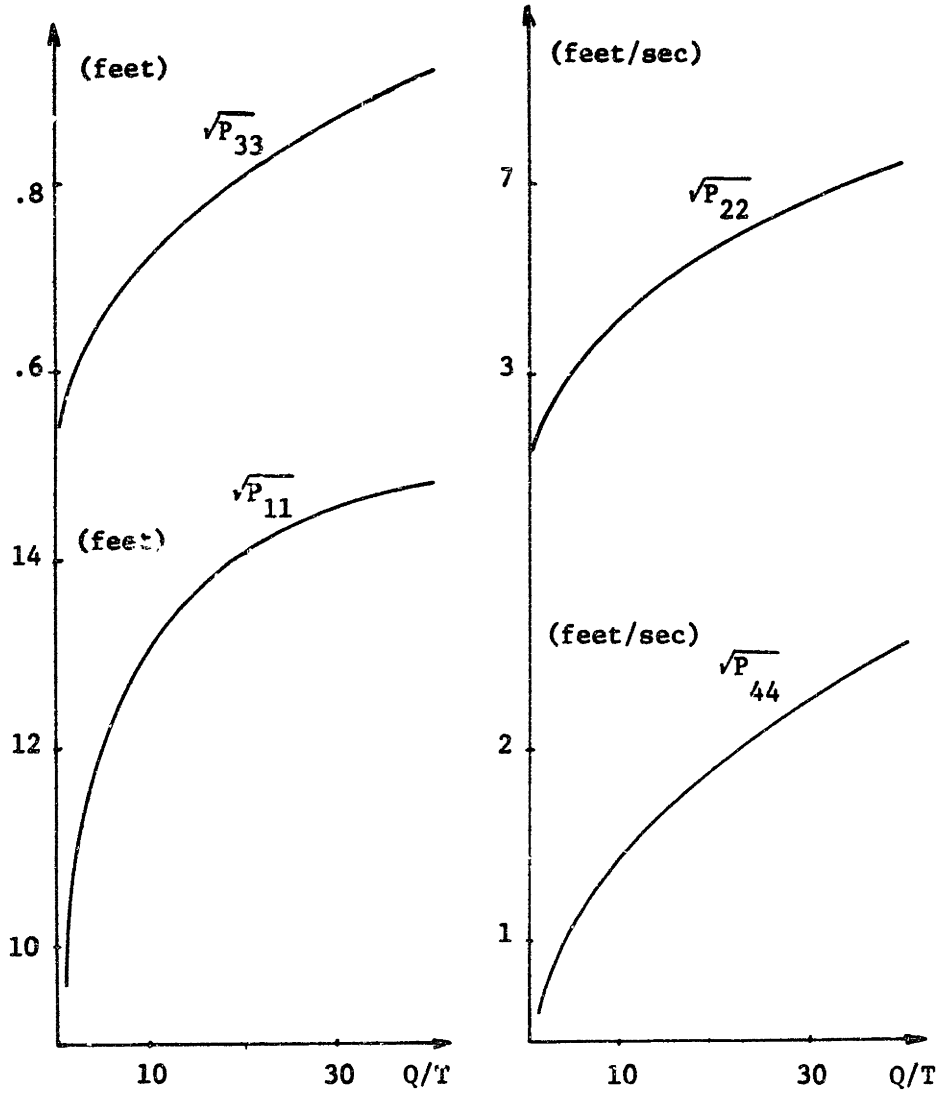
---

* This concept is here used in a heuristic sense.

then the errors become actually unbounded as compared with the divergence observed with the EKF below threshold where because of the large values assumed by the covariance elements the spread of the actual errors is so large that the estimate loses its practical significance. In particular for the range channel when the standard deviation is near

$$\frac{\lambda}{4}$$

the filter may lose the lock, jumping to other stable equilibrium points of the range space.



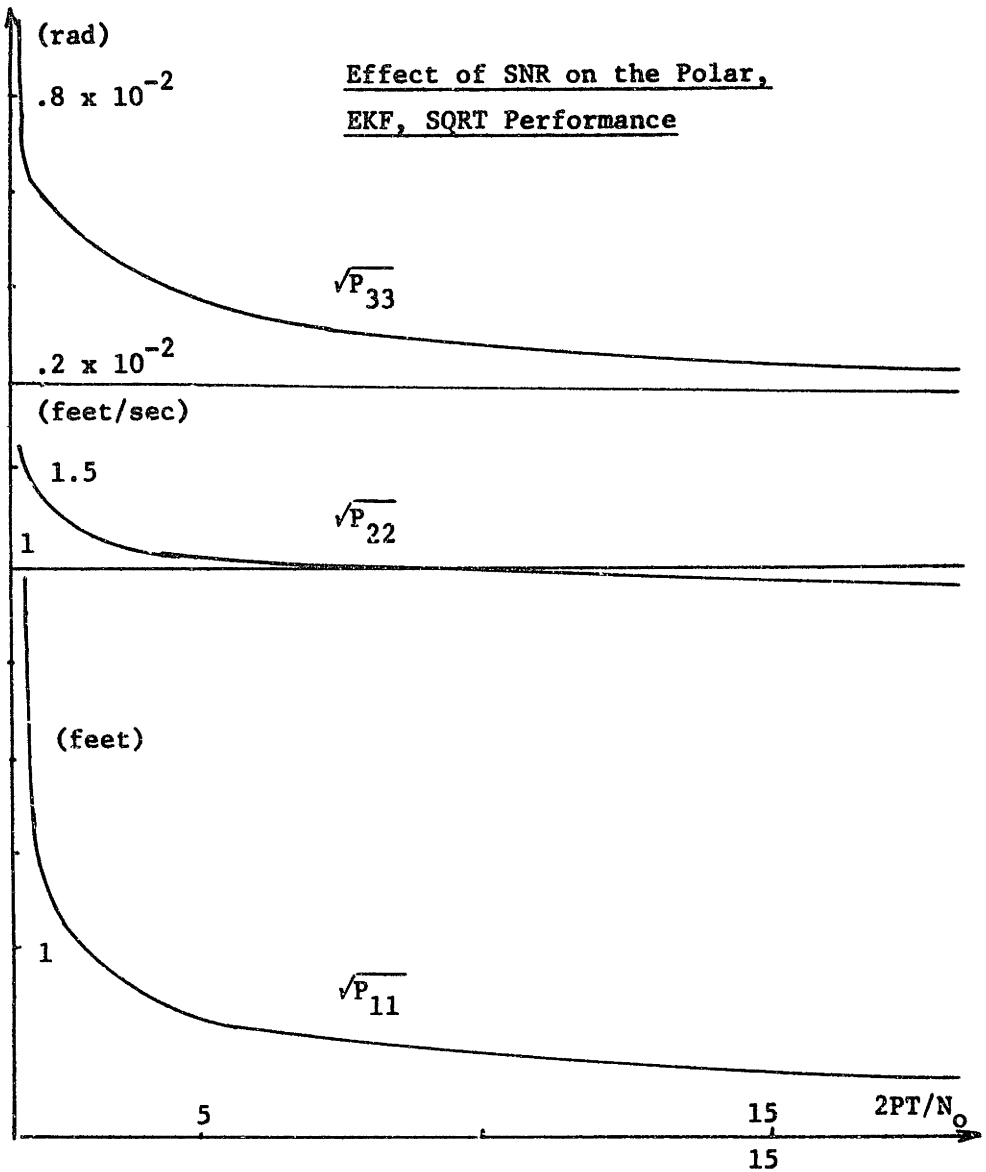
EKF, Rect., CONV

Fig. 6-10

Effect of Q on the Rect., EKF, SQRT Performance

Fig. 6-11

A EKF, Polar, SQRT B



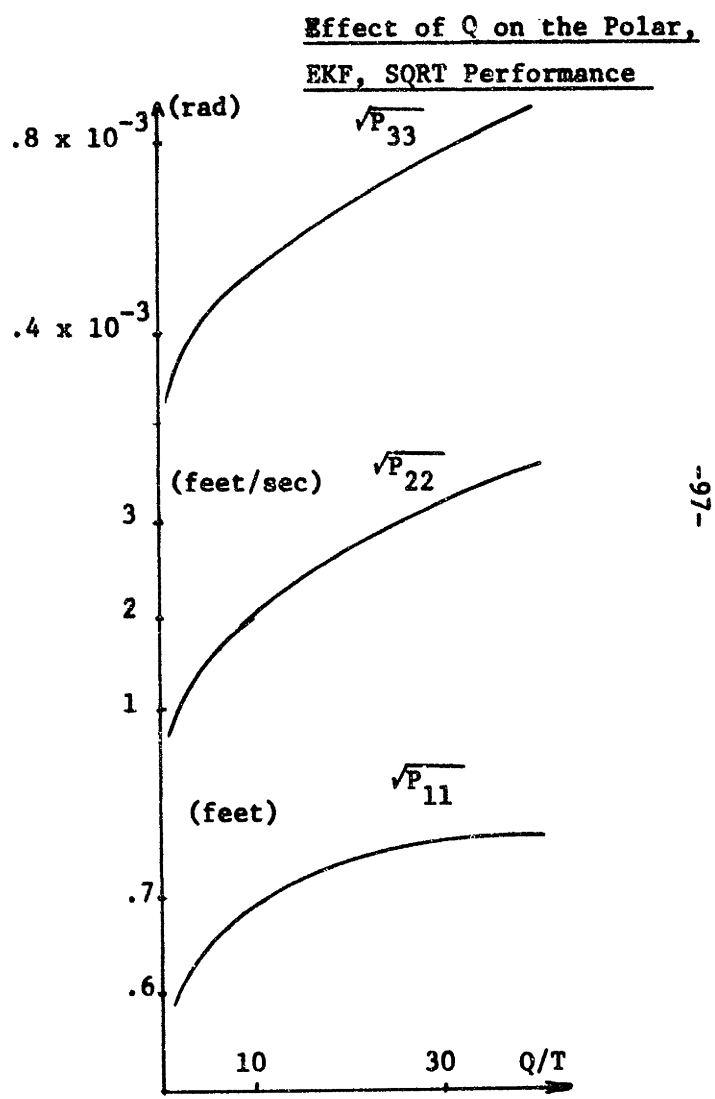
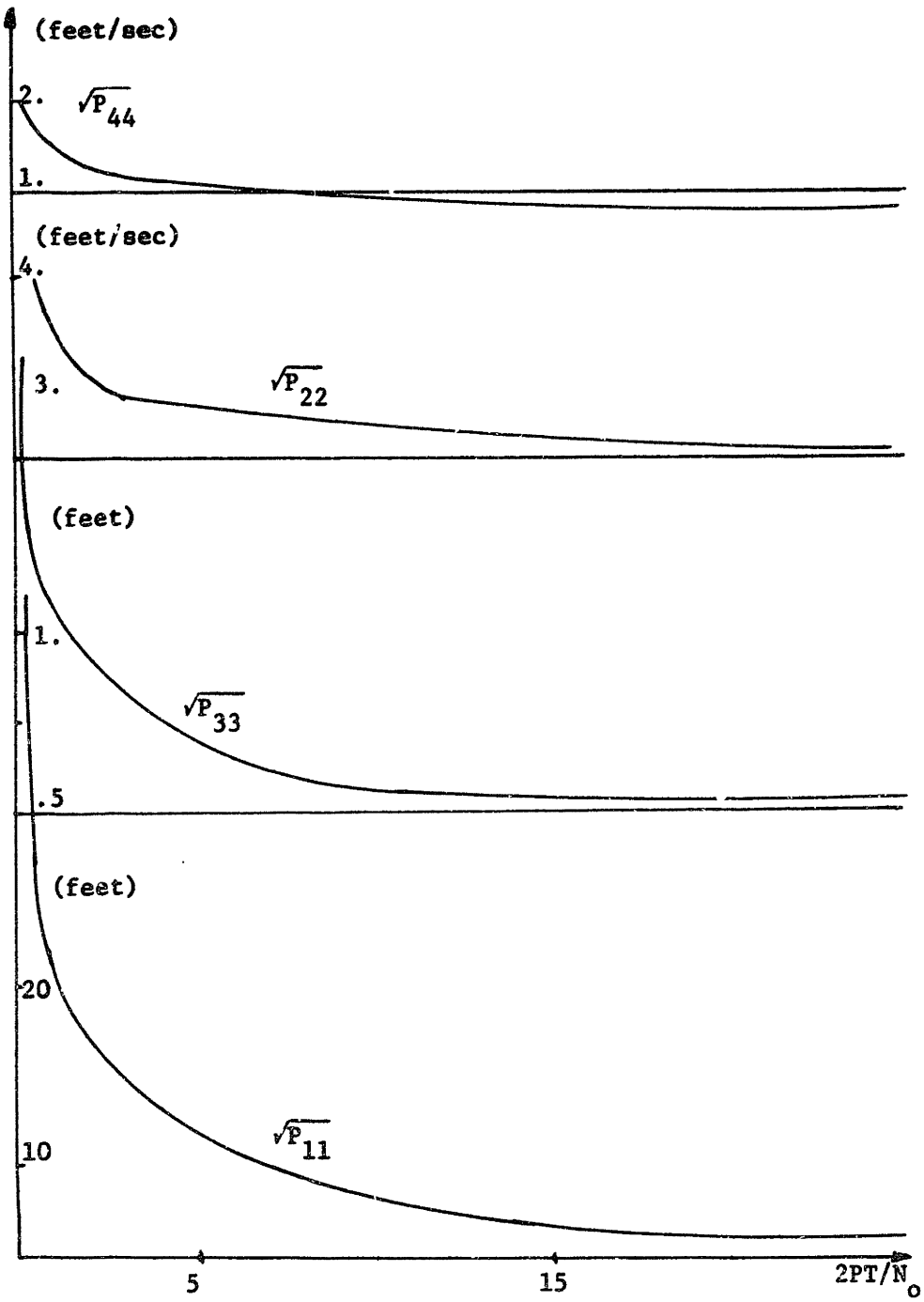


Fig. 6-12

EKF, Rect., CONV

Effect of SNR on the Rect., EKF, SQRT Performance



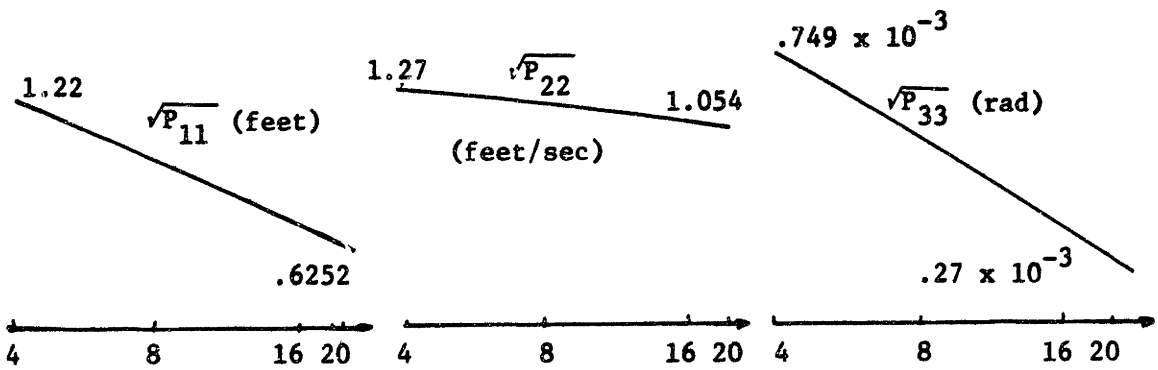
Array Length

Keeping the space between sensors equal to half of a wavelength, i.e.

$$\Delta p = \frac{\lambda}{2}$$

decreasing the array length also means reducing the number of sensors N.

Figure 6-13 shows the effect in the performance when we vary N.



Polar, EKF, SQRT

Fig. 6-13

Study of the Effect of the Array Length on the Filter Performance

Geometry

The geometry mainly affects the bearing angle channel as pointed out in Chapter 5.

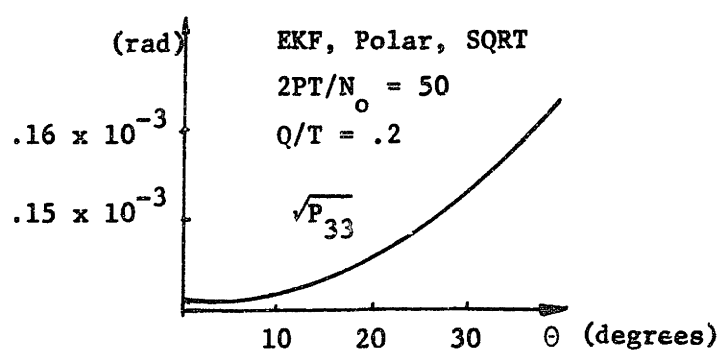


Fig. 6-14

Effect of the Range on the Bearing Channel Performance

Fig. 6-15

Effect of the Bearing on the Bearing Channel Performance

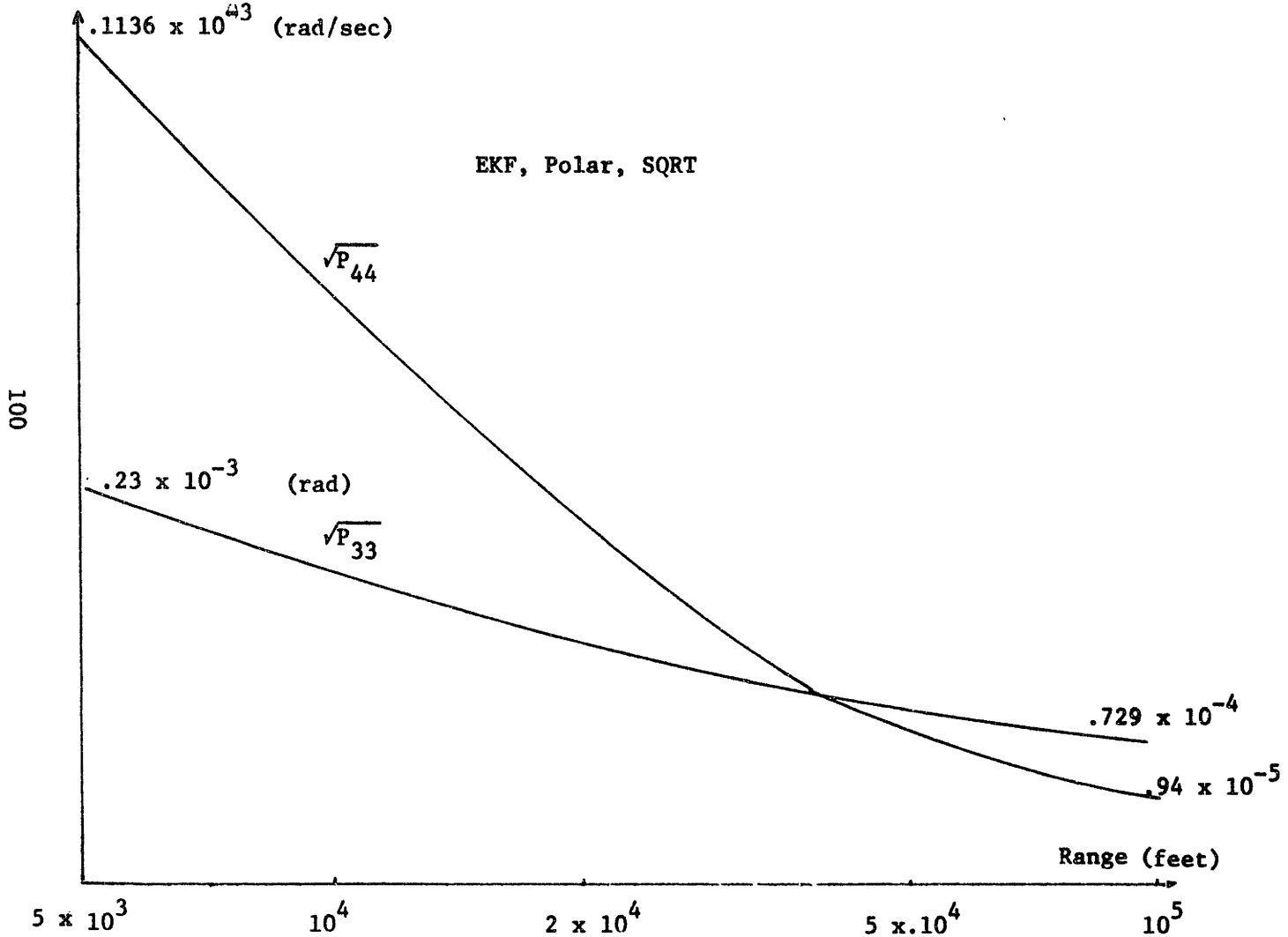


Figure 6-14 shows the effect on the bearing covariance when the bearing angle increases. We observe that the broadside array performs better. The net effect of increasing Θ is comparable to a reduction on the effective array length as discussed in Chapter 5.

Figure 6-15 shows the performance gain (reduction of $\sqrt{P_{33}}$, $\sqrt{P_{44}}$) of the bearing channel when the range increases. We observe that for ranges of the order of 10^5 feet a saturation seems to occur while for small ranges the performance degrades almost exponentially.

This can be explained by the fact that small errors with large range values induce small bearing errors, while the same errors for small ranges may induce large bearing errors.

This completes the more interesting cases studied. Many more variants could be simulated but the above examples consider the important theoretical questions. We next recap the main issues explored in this chapter.

Summary

1) Choice of the Coordinate System

We picked two different coordinate frameworks to describe the target's and observer's geometry. In the natural system (polar frame) both the system function and the observation function exhibit strong nonlinearities; the cartesian dynamical system is linear but the observation function is nonlinear. From the mathematical analysis it was difficult to predict how the systems would perform.

The simulation results show that in both frames for specific configurations of the nonlinear estimation problem the filters converge and that for a broadside array, as analysed in Chapter 5, the Y-channel

of the rectangular coordinate processor has a performance almost equivalent to the performance of the range channel of the polar frame.

2) Choice of the Filter: EKF versus MAP

Both filters are obtained by a first order linearization of the problem about the previous estimate with the two filters differing only in the covariance measurement update equation as pointed out in Chapter 4.

For each filter and on both frames one can find a set of parameters, a geometric configuration and source dynamics such that the processors will converge and work in a linear region.

In the MAP filter the measurements are explicitly present in the covariance equation and so the errors in the estimate modulate directly the covariance propagation as is exhibited by the graphs of Fig. 6-9.

In the EKF only through the estimates is the covariance propagation coupled to the measurements. As it is seen from Fig. 6-3 and Fig. 6-5 the covariance matrix reaches a steady state value which depends mainly on the assummed parameters and the geometric configuration.

The simulation runs showed that the MAP filter is more unstable than the EKF. This unstability of the filter can be traced to the explicit presence of the measurement noise in the covariance equation, from which the covariance matrix may lose its positive semidefiniteness character, thus becoming meaningless (we remark that the SQRT algorithm cannot circumvent this fact since the problem now is not of a numerical difficulty). Obviously this fact is more important for small SNR and for geometric configurations different of the optimal one.

3) Effect of the Different Parameters

Signal to noise ratio: All the filters show a sharp threshold with respect to the SNR below which the performance deteriorated quickly.

This threshold is not uniform for the class of filters studied. For the EKF it is about

$$\frac{2PT}{N_o} \sim .5$$

while for the MAP filter it is

$$\frac{2PT}{N_o} \sim 2.$$

With the MAP filter the danger of unstability is much more pronounced for SNR near the threshold.

Driving noise power: Q also directly influences the filter's performance and convergence ability. As we increase Q the filter tracks with more and more difficulty as one would expect.

Array length: The array length is an important parameter of the simulation with a better overall performance for larger arrays as it is physically intuitive and was predicted by the bounds of Chapter 5.

4) Tracking Geometry

The filter's performance is dependent of the tracking geometry, range and bearing as well as their assumed rates of change.

As we increase the bearing angle the performance deteriorates. But essentially the bearing estimate is much more affected than the range estimate.

With respect to the range we observe that as the target gets further away the performance gets better.

5) The Conventional Implementation Versus the SQRT Algorithm

The conventional implementation does not converge in the case of polar coordinates because of numerical difficulties as it was explained in Chapter 4.5. With the polar frame we have to resort to the SQRT algorithm.

With the rectangular frame the conventional implementation is able to converge because on the one hand the condition number of the covariance matrix is smaller than the b^n factor of the computer and on the other hand the elements of the covariance matrix are sufficiently far away from zero. The SQRT algorithm does not give a significant improvement in this case.

6) Initialization

In some of the runs we have started the polar filter with an initial guess for which

$$|R - \hat{R}| > \frac{\lambda}{2}$$

It was assessed that the Kalman filter would converge to the nearest stable equilibrium point (see Chapter 5). This confirmed the previous theoretical analysis that asserted that because of the periodic structure of the signal the filter can only solve for the range within a half of a wavelength interval of the true value. So we have to initialize the filter within the primary lock in range (section 5.1.1).

As for the initial value of the remaining state variables the filter was particularly sensitive to the initial bearing error and practically insensitive to the initial guess of the range rate.

A maximum likelihood approach is taken in Chapter 7 to find the starting "guess" of the Kalman filter.

We observe from the simulation results that the final values of the propagated covariances are intimately related to the parameters and geometry, and are almost independent of the assumed a priori covariance thus confirming the analysis of Chapter 5.2.

Chapter VII

Initialization of the Kalman Filter

As discussed earlier in this report, the convergence of the filters depends intimately on the starting point. This is theoretically obvious given that the nonlinear estimation problem was solved by a first order approximation, i.e., by linearizing the system about the previous estimate.

If the "a priori" knowledge, i.e., the initial estimate, is not good enough, the neglected terms in the Taylor series expansions contribute with important errors and the linearized assumption fails.

In this chapter we return to the conventional "bearings only" problem. Our main concern is to find a good estimate of the bearing angle. Then a triangularization procedure solves for the range in order to get in the primary lock in interval of the range channel.

The solution is carried in two steps. In the first one an M-ary detection problem is solved to find the correct subinterval in the bearing parameter space, i.e., in the interval $[o, 2\pi]$ where the true bearing angle is. In the second step the bearing maximum likelihood (ML) estimate is found by carrying a local maximization.

The global accuracy is measured in terms of the probability of large errors (i.e. of choosing the wrong cell) given by equation (7-44). The local accuracy is measured by the mean square error of the ML-estimate given by (7-56).

In section 7.1 we present the maximum likelihood estimation problem, the detailed model and the optimum receiver. In 7.2 we derive the expression of the probability of large errors, in 7.3 the variance of the ML-estimate, in 7.4 the low pass model and finally in 7.5 we present the simulation results.

7.1 Maximum Likelihood Estimation

The Model

We recall from section 2.3 that at the i-th sensor the received signal is modeled by equation (2-41):

$$r_i(t) = s_i(t) + w_i(t)$$

$$= \sqrt{2P}\ \sin w_i\left(t - \frac{R(t)}{c} + \frac{P_i}{c}\ \sin\Theta(t)\right) + w_i(t),\ i = 1, \ldots, N \tag{7-1}$$

We will assume that during the processing interval [o,T], $\Theta(t)$ and $R(t)$ do not vary significantly so that they can be treated as unknown, non varying parameters.

Furthermore we will model

$$\psi = -\frac{R}{c} \tag{7-2}$$

as an unwanted random parameter with a prespecified probability distribution.

We want to find the maximum likelihood estimate of the bearing angle. This problem falls in the category of the detection and estimation of signals with unwanted parameters (see Van Trees [A-8], section 4-4).

The maximum likelihood function follows by a simple generalization to the vector case of equation 4-359 in Van Trees [A-8]. We get

$$\Lambda[\underline{r}(t)] = \int_{\pi}^{\pi} p_{\psi}(\psi)d\psi \exp[+\frac{2}{N_o} \int_o^T \underline{r}^T(t) - \underline{s}(t, \psi, \hat{\Theta})dt - \frac{1}{N_o} \int_o^T \underline{s}^T(t, \psi, \hat{\Theta}).$$

$$\underline{s}(t, \psi, \hat{\Theta})dt] \qquad (7\text{-}3)$$

where

$$\underline{r}(t) = \begin{bmatrix} r_1(t) \\ \cdot \cdot \cdot \\ r_N(t) \end{bmatrix} \qquad (7\text{-}4)$$

$$\underline{s}(t, \psi, \hat{\Theta}) = \begin{bmatrix} s_1(t, \psi, \hat{\Theta}) \\ \cdot \cdot \cdot \\ s_N(t, \psi, \hat{\Theta}) \end{bmatrix} \qquad (7\text{-}5)$$

and

$$r_i(t) = s_i(t, \psi, \Theta) + w_i(t) \qquad (7\text{-}6)$$

with

$$s_i(t, \psi, \Theta) = \sqrt{2P} \sin w_c(t + \frac{p_i}{c} \sin\Theta + \psi) \qquad (7\text{-}7)$$

We remark that

$$\int_o^T \underline{s}^T(t, \Theta)\ \underline{s}(t, \Theta)dt \qquad (7\text{-}8)$$

gives the received energy which is independent of the phase and so can be ignored in the maximization procedure.

Following the rationale in Van Trees (for the details refer to chapter 4-4 of [A-8]) we define

$$L_{c_i} = \int_o^T \sqrt{\frac{2}{T}}\, r_i(t) \cos\left[w_c t + \frac{w_c}{c} p_i \sin\hat{\theta}\right]dt \tag{7-9}$$

$$L_{s_i} = \int_o^T \sqrt{\frac{2}{T}}\, r_i(t) \sin\left[w_c t + \frac{w_c}{c} p_i \sin\hat{\theta}\right]dt \tag{7-10}$$

and

$$\Lambda^1[\underline{r}(t)] = \int_{-\pi}^{\pi} p(\psi)d\psi \exp\left\{ \frac{2\sqrt{PT}}{N_o} \sum_{i=1}^{N} [L_{c_i} \cos\psi - L_{s_i} \sin\psi] \right\} \tag{7-11}$$

Choosing

$$p(\psi: \Lambda_m) = \frac{\exp[\Lambda_m \cos\psi]}{2\pi\, I_o(\Lambda_m)} \qquad -\pi \leq \Theta \leq \pi \tag{7-12}$$

For $\Lambda_m = o$

$$p(\psi: \Lambda_m) = \frac{1}{2\pi} \qquad -\pi \leq \Theta \leq \pi \tag{7-13}$$

i.e. ψ is uniformly distributed in $[-\pi, \pi]$.

Define

$$L_c = \sum_{i=1}^{N} L_{c_i} \tag{7-14}$$

$$L_s = \sum_{i=1}^{N} L_{s_i} \tag{7-15}$$

Substitution in (7-11) leads to

$$\Lambda^1[\underline{r}(t)] = I_o\left(\frac{2\sqrt{PT}}{N_o} [L_c^2 + L_s^2]^{1/2} \right) \tag{7-16}$$

where $I_o(.)$ is a modified Bessel function of the first kind.

The log-maximum-likelihood function becomes

$$\ell n\, \Lambda^1[\underline{r}(t)] = \ell n\, I_o\left(\frac{2\sqrt{PT}}{N_o} [L_c^2 + L_s^2]^{1/2} \right) \tag{7-17}$$

Because of the monotonicity of the $\ln I_o(.)$ we need only to maximize the argument in (7-17). So we want to maximize

$$\ell(\underline{r}) = \left(\frac{2\sqrt{PT}}{N_o}\right)^2 [L_c^2 + L_s^2] \tag{7-18}$$

The maximization of $\ell(\underline{r})$ is carried out in two steps. First we do a parallel processing dividing the parameter space in M equal cells of length $\frac{2\pi}{M}$

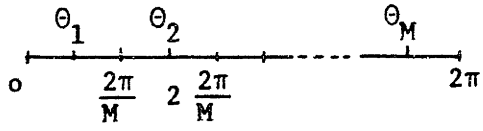


Fig. 7-1

and constructing the function $\ell(\underline{r})$ for the central points Θ_i in each cell as indicated in Fig. 7-2.

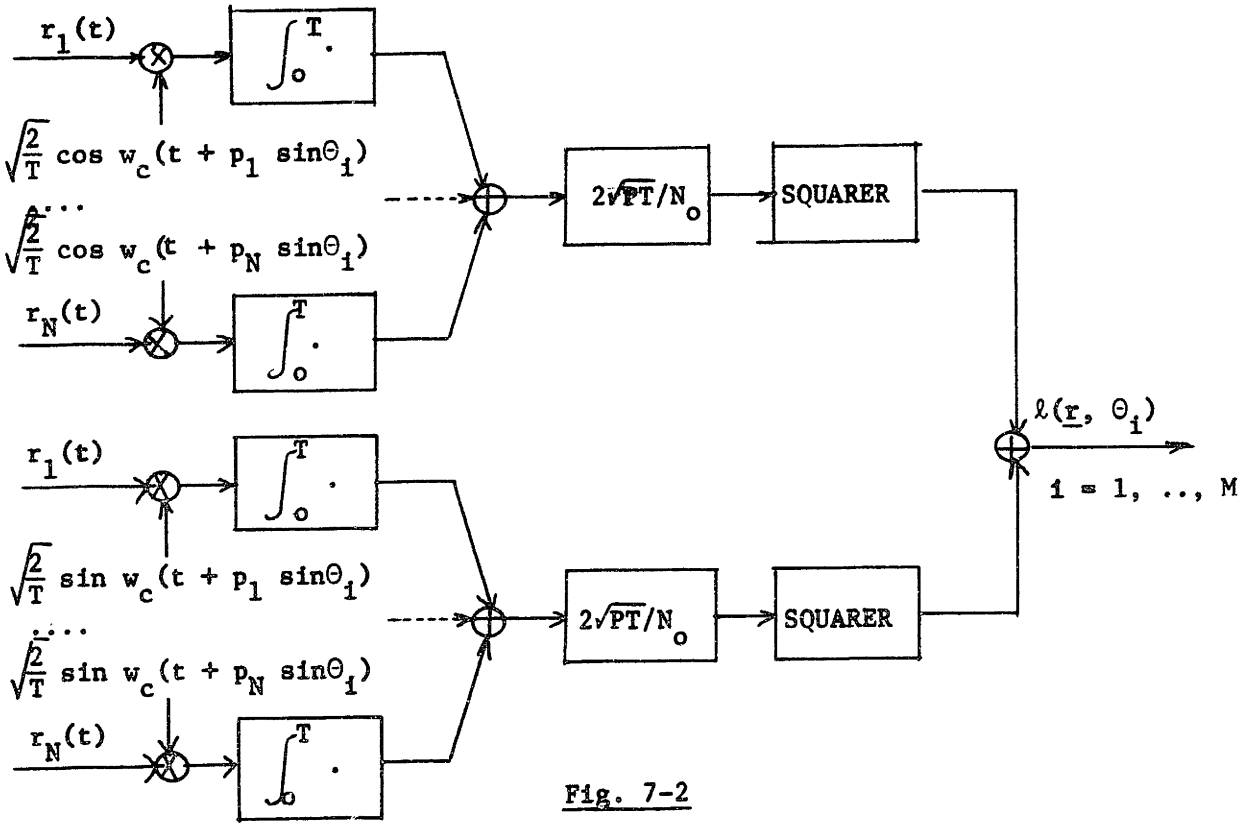


Fig. 7-2

The largest of $\ell(\underline{r}, \Theta_i)$, $i = 1, \ldots, M$, indicates in which cell the true bearing angle is. Then we carry out a local maximization procedure to find $\hat{\Theta}_{m\ell}$, given that the maximum is in the chosen cell.

Equivalently the first step corresponds to an M-ary detection problem:

$$\begin{aligned} H_1: &\quad \underline{r}(t) = \underline{s}(t, \psi, \Theta_1) + \underline{w}(t) \qquad o \leq t \leq T \\ H_2: &\quad \underline{r}(t) = \underline{s}(t, \psi, \Theta_2) + \underline{w}(t) \qquad o \leq t \leq T \\ &\cdots \qquad\qquad\qquad\qquad\qquad\qquad \cdots \\ H_M: &\quad \underline{r}(t) = \underline{s}(t, \psi, \Theta_M) + \underline{w}(t) \qquad o \leq t \leq T \end{aligned} \tag{7-19}$$

In each hypothesis the same waveform is present corrupted by an additive zero mean Gaussian, white noise disturbance and it is desired to decide what the value assumed by the parameter Θ is. We remark further the presence of the unwanted parameter ψ modeled, as pointed out earlier, as a random variable with a priori probability density specified by equation (7-18).

The optimum receiver for the M-ary detection problem outlined above, given the equal energy on all hypothesis, assumming equal a priori probabilities

$$P(H_1) = \ldots = P(H_M) = \frac{1}{M} \tag{7-20}$$

and choosing as criterion the minimum probability of error is the "largest of" receiver implemented in Fig. 7-2.

The final estimate is obtained by conducting a local maximization, assumming the correct value of the bearing is inside the chosen region.

To study the performance of the designed receiver, we consider two kinds of errors: the large errors and the small errors (when a high signal to noise ratio assumption holds).

We want to minimize the occurrence of large errors and so chose as criterion the minimum probability of error $Pr(\varepsilon)$ in the M-ary detection problem

$$Pr(\varepsilon) = \sum_{i=1}^{M} Pr(\varepsilon|H_i)\, Pr(H_i) \tag{7-21}$$

where $Pr(\varepsilon|H_i)$ is the probability of choosing $\Theta = \Theta_j$, $j = 1, \ldots, M$, $j \neq i$ given that hypothesis H_i is true, i.e. $\Theta = \Theta_i$.

Under the hypothesis of small errors we want the mean square error

$$E\,[(\Theta - \hat{\Theta})^2] \tag{7-22}$$

to be small.

We will compute the probability of error $Pr(\varepsilon)$ in section 7.2 and the variance of the M-L estimate in section 7.3.

7.2 Global Accuracy (Ambiguity): Probability of Error

When maximizing

$$\ell(\underline{r}) = L_c^2 + L_s^2$$

an M-ary detection problem is solved by assumming that the parameter space of Θ is discretized to the set

$$\{\Theta_i,\ i = 1, \ldots, M\}$$

Under each hypothesis the received signal is:

$$H_i:\ \underline{r}(t) = \underline{s}(t, \psi, \Theta_i) + \underline{w}(t) \qquad o \leq t \leq T \qquad i = 1, \ldots, M \tag{7-23}$$

where

$$\underline{s}(t, \psi, \Theta_i) = \sqrt{2P} \begin{bmatrix} \sin w_c\left(t + \frac{P_1}{c} \sin\Theta_i + \psi\right) \\ \cdots \\ \sin w_c\left(t + \frac{P_N}{c} \sin\Theta_i + \psi\right) \end{bmatrix} \tag{7-24}$$

Defining:

$$L_c^i = \int_o^T \underline{r}^T(t)\, \underline{s}_c(t, \Theta_i)dt \tag{7-25}$$

$$L_s^i = \int_o^T \underline{r}^T(t)\, \underline{s}_s(t, \Theta_i)dt \tag{7-26}$$

with

$$\underline{s}_c(t, \Theta) = \sqrt{\frac{2}{T}} \begin{bmatrix} \cos(w_c t + \frac{w_c}{c} p_1 \sin\Theta) \\ \cdots \\ \cos(w_c t + \frac{w_c}{c} p_N \sin\Theta) \end{bmatrix} \tag{7-27}$$

and

$$\underline{s}_s(t, \Theta) = \sqrt{\frac{2}{T}} \begin{bmatrix} \sin w_c\left(t + \frac{p_1}{c} \sin\Theta\right) \\ \cdots \\ \sin w_c\left(t + \frac{p_N}{c} \sin\Theta\right) \end{bmatrix} \tag{7-28}$$

The likelihood ratio test (LRT) computes

$$\ell_i(\underline{r}) = L_c^{i\,2} + L_s^{i\,2} \qquad i = 1, \ldots, M \tag{7-29}$$

and chooses the largest.

Up to this point we did not question what the correlation of the signals is under the different hypothesis. We know that a considerable simplification occurs in the computation of the probability of error if the signals are orthogonal.

We investigate now this point by computing the time-frequency correlation function, or equivalently its modulo squared, the ambiguity

function (for a detailed discussion of the ambiguity function and its properties see Van Trees [A-10]. Also references [F]).

The ambiguity function for the problem under consideration is defined by:

$$\phi_N = \hat{L}_c^2 + \hat{L}_s^2 \tag{7-30}$$

where

$$\hat{L}_c = \int_o^T \underline{s}^T(t, \Theta)\, \underline{s}_c(t, \hat{\Theta})dt \tag{7-31}$$

$$\hat{L}_s = \int_o^T \underline{s}^T(t, \Theta)\, \underline{s}_s(t, \hat{\Theta})dt \tag{7-32}$$

Neglecting double order frequency terms and assumming the equal spaced linear array of Chapter 2, we get

$$\phi_N = \left| \frac{\sin(N \frac{1}{2} \frac{w_c}{c} \Delta p(\sin\Theta - \sin\hat{\Theta}))}{\sin(\frac{1}{2} \frac{w_c}{c} \Delta p(\sin\Theta - \sin\hat{\Theta}))} \right|^2 \tag{7-33}$$

For $\Theta = o$ and $\Delta p = \frac{\lambda}{2}$, where λ is the wavelength, i.e.,

$$\lambda f_c = c \tag{7-34}$$

we get

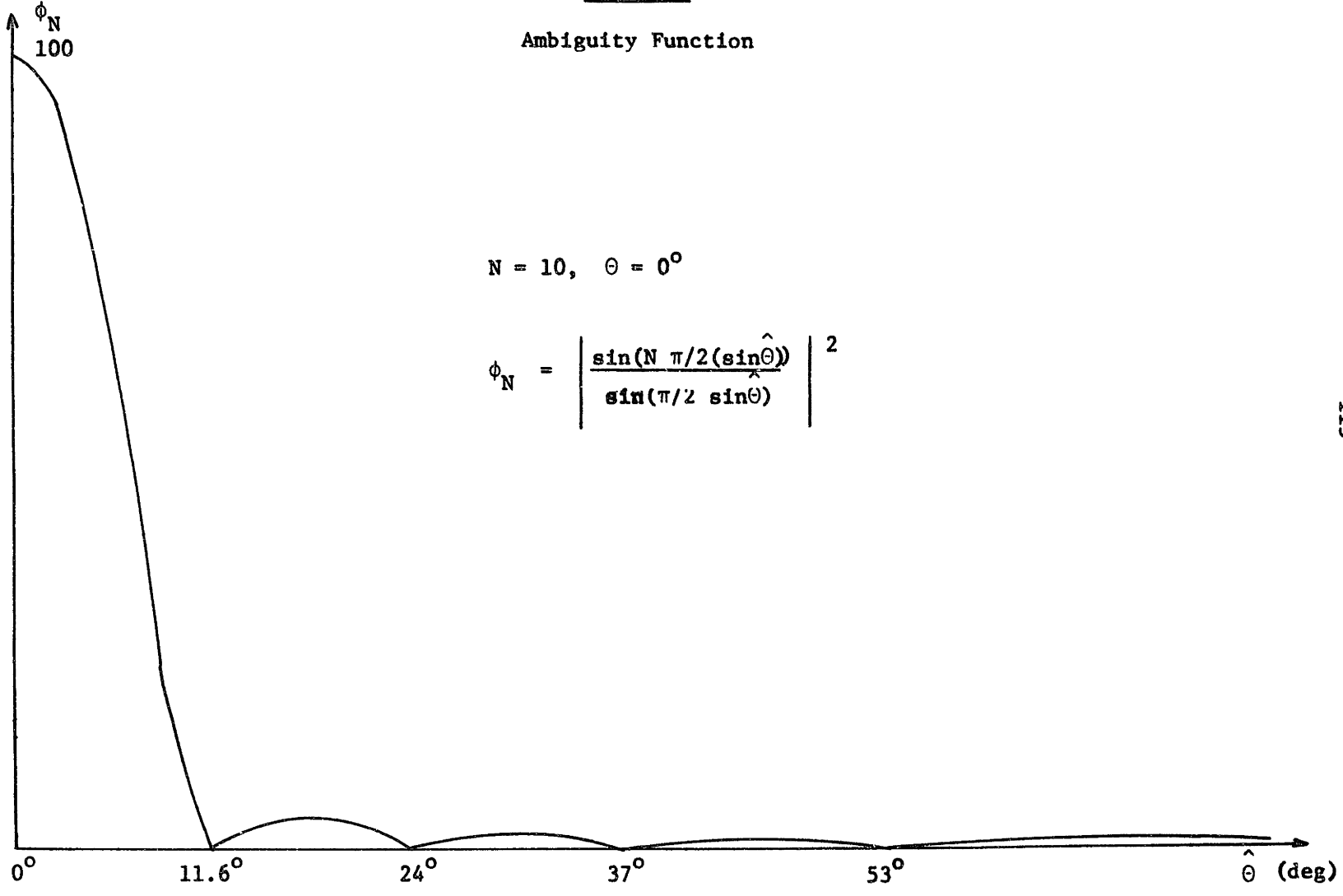
$$\phi_N = \left| \frac{\sin(N \frac{\pi}{2} \sin\hat{\Theta})}{\sin(\frac{\pi}{2} \sin\hat{\Theta})} \right|^2 \tag{7-35}$$

A sketch of the ambiguity function ϕ_N is presented in Fig. 7-3 where $N = 10$, $\Theta = o^o$, $\Delta p = \frac{\lambda}{2} = 25$ feet.

The graph shows the existence of several lobes (ambiguity). But we observe that ϕ_N is sharply reduced for $\hat{\Theta}$ outside the first lobe and so choosing M such that $\frac{2\pi}{M} \sim 10^o$ it is plausible to assume that

Fig. 7-3

Ambiguity Function



the signals are orthogonal under the different hypothesis.

In summary we have an M-ary detection problem with M orthogonal signals, with an unknown unwanted phase ψ.

We proceed to compute the

Probability of error

$$P_r(\varepsilon) = \int_{-\pi}^{\pi} p_r(\varepsilon|\psi)\ p(\psi)d\psi \tag{7-36}$$

But

$$P_r(\varepsilon|\psi) = \sum_{i=1}^{M} P_r(\varepsilon|\psi, H_i)\ P_r(H_i) = \frac{1}{M} \sum_{i=1}^{M} P_r(\varepsilon|\psi, H_i) \tag{7-37}$$

Now

$$P_r(\varepsilon|\psi, H_i) = 1 - P_r(\ell_i > \ell_j, \text{ all } j \neq i\ |\psi, H_i) \tag{7-38}$$

But

$$P_r(\ell_i > \ell_j \quad \text{all } j \neq i\ |\psi, H_i) = \int_0^\infty p(\ell_i|\psi, H_i) \prod_{\substack{j=1 \\ j\neq i}}^{M} \{\int_0^\infty p(\ell_j\ |H_i, \psi)d\ell_j\}d\ell_i \tag{7-39}$$

and we need

$$p(\ell_i\ |\ \psi, H_i)$$

$$p(\ell_j\ |\ \psi, H_i),\ j \neq i$$

These probability densities follow immediately from the statistics of L_c^i , L_c^j , L_s^i , L_s^j conditioned on ψ and H_i.

Under the orthogonality assumption of the signals in the different hypothesis we summarize the results on Table 7-1.

	L_c^i	L_s^i	L_c^j	L_s^j
$E[. \mid H_i, \psi]$	$N\sqrt{PT}\sin\psi$	$N\sqrt{PT}\cos\psi$	o	o
$\text{var}[. \mid H_i, \psi]$	$N\frac{N_o}{2}$			

Table 7-1

Remark: In Table 7-1 $j \neq i$

Then

$$p(\ell_i \mid \psi, H_i) = (2\pi N\frac{N_o}{2})^{-1} \exp\left\{-\frac{(L_c^i - N\sqrt{PT}\sin\psi)^2 + (L_s^i - N\sqrt{PT}\cos\psi)^2}{2N\frac{N_o}{2}}\right\} \tag{7-40}$$

and

$$p(\ell_j \mid \psi, H_i) = (2\pi N\frac{N_o}{2})^{-1} \exp\left\{-\frac{L_c^{j\,2} + L_s^{j\,2}}{2N\frac{N_o}{2}}\right\} \quad j = 1, \ldots, M,\ j \neq i \tag{7-41}$$

We remark that equations (7-40) and (7-41) are symmetric on H_i and ψ and so

$$P_r(\varepsilon|\psi) = P_r(\varepsilon|\psi, H_i) = 1 - \int_o^\infty p(\ell_i|\psi, H_i)\left[\int_o^{\ell_i} p(\ell_j|H_i, \psi)d\ell_j\right]^{M-1} d\ell_i \tag{7-42}$$

This integral is easily computed (see for example Van Trees [A-8])

$$P_r(\varepsilon|\psi) = 1 - E\left\{\left[1 - \exp(-\frac{x^2 + y^2}{2})\right]^{M-1}\right\} \tag{7-43}$$

and for ψ uniformly distributed in $[-\pi, \pi]$ we obtain

$$P_r(\varepsilon) = \sum_{k=1}^{M-1} \binom{M-1}{k} (-1)^{k+1} \frac{\exp\left\{ -\sqrt{\frac{2NPT}{N_o}} \frac{k}{k+1} \right\}}{k+1} \tag{7-44}$$

7.3 Local Accuracy: Variance of the Maximum Likelihood Estimate

The received signal is

$$\underline{r}(t) = \underline{s}(t, \psi, \Theta) + \underline{w}(t) \qquad o \leq t \leq T \tag{7-45}$$

with

$$s_i(t, \psi, \Theta) = \sqrt{2P}\, a(t) \sin w_c\left(t + \frac{p_i}{c} \sin\Theta + \psi\right) \tag{7-46}$$

where in the case under consideration we took

$$a(t) = 1 \tag{7-47}$$

We assumed $\Theta(t)$ to be slowly varying so that it can be considered constant in $[o, T]$, i.e. we modeled Θ not as a random process (as before, Chapter 2) but as an unknown quantity.

The variance of the maximum likelihood estimate of

$$u = \frac{\sin\Theta}{c} \tag{7-48}$$

has been computed by H. Urkowitz [F-5] with the result

$$\frac{1}{\text{Var } \hat{u}} = \frac{2E_T w_c^2}{N_o} \left\{ \left(1 + \frac{B_r^2}{w_c^2}\right) L_r^2 + B_r^2 X_c^2 / w_c^2 \right\} \tag{7-49}$$

where:

Mean square radian bandwidth

$$B_r^2 = \frac{\int w^2 |A(f)|^2 df}{\int |A(f)|^2 df} \tag{7-50}$$

Horizontal centroid

$$X_c = \frac{\sum_{i=1}^{N} p_i |G_i|^2}{\sum_{i=1}^{N} |G_i|^2} \tag{7-51}$$

Mean square aperture length

$$L_r^2 = \frac{\sum_{i=1}^{N} p_i^2 |G_i|^2}{\sum_{i=1}^{N} |G_i|^2} - X_c^2 \tag{7-52}$$

Beam pattern of i'th element of discrete antenna

$$G_i = G_i(w_o u)\Big|_{u = o} \tag{7-53}$$

Total energy absorbed by antenna in [o, T]

$$E_T = PT \tag{7-54}$$

Carrier frequency

$$f_o$$

Fourier transform of a(t)

$$A(f)$$

Applying the Urkowitz result to our problem we get, assumming all array elements are isotropic:

$$G_i = 1$$

$$X_c = o$$

$$L_r^2 = \frac{\Sigma p_i^2}{N} \tag{7-55}$$

$$B_r^2 = o$$

$$E_T = PT$$

$$\text{var } \sin\hat{\Theta} = \frac{N}{\frac{2PT}{N_o}\left(\frac{w_c}{c}\right)^2 \sum_{i=1}^{N} p_i^2} \tag{7-56}$$

For Θ near zero $\sin\Theta \sim \Theta$ and the variance of $\hat{\Theta}$ will be approximately given by the expression above.

7.4 Low Pass Model

For simulation purposes we present the low pass version of the model discussed in section 7.1.

Following the rationale of section 4.1, we have at the ith sensor

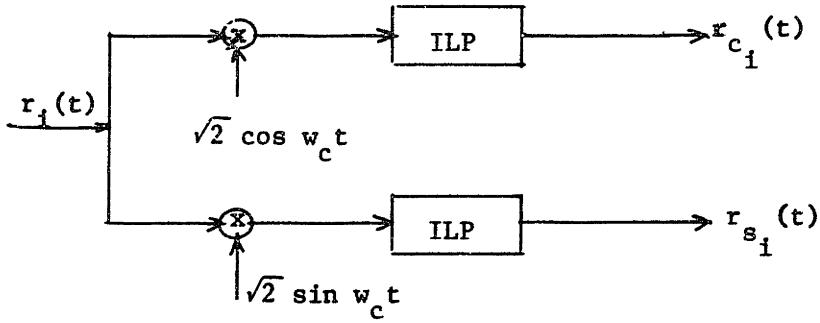


Fig. 7-4

where ILP stands for Ideal Low Pass Filter.

$$r_i(t) = \sqrt{2P}\sin(w_c t + \frac{w_c}{c} p_i \sin\Theta + \psi) + w_i(t) \tag{7-57}$$

$$r_{c_i}(t) = \sqrt{P}\sin(\frac{w_c}{c} p_i \sin\Theta + \psi) + w_{c_i}(t) \tag{7-58}$$

$$r_{s_i}(t) = \sqrt{P}\cos(\frac{w_c}{c} p_i \sin\Theta + \psi) + w_{s_i}(t) \tag{7-59}$$

In vector notation

$$\underline{r}(t) = \begin{bmatrix} r_{c_i}(t) \\ \cdots \\ r_{c_N}(t) \\ r_{s_i}(t) \\ \cdots \\ r_{s_N}(t) \end{bmatrix} = \underline{s}(t, \psi, \Theta) + \underline{w}(t) \tag{7-60}$$

$$\underline{s}(t, \psi, \Theta) = \begin{bmatrix} s_{c_1}(t, \psi, \Theta) \\ \cdots \\ s_{c_N}(t, \psi, \Theta) \\ s_{s_1}(t, \psi, \Theta) \\ \cdots \\ s_{s_N}(t, \psi, \Theta) \end{bmatrix} \tag{7-61}$$

with

$$s_{c_i}(t, \psi, \Theta) = \sqrt{P} \quad \sin(\frac{w_c}{c} p_i \sin\Theta + \psi) \tag{7-62}$$

$$s_{s_i}(t, \psi, \Theta) = \sqrt{P} \quad \cos(\frac{w_c}{c} p_i \sin\Theta + \psi) \tag{7-63}$$

and

$$\underline{w}(t) = \begin{bmatrix} w_{c_i}(t) \\ \cdot \cdot \cdot \\ w_{c_N}(t) \\ w_{s_i}(t) \\ \cdot \cdot \cdot \\ w_{s_N}(t) \end{bmatrix} \tag{7-64}$$

where, as argued in section 4.1, $w_{c_i}(t)$, $w_{s_i}(t)$ are sample functions of zero mean white Gaussian processes and ψ is a random variable uniformly distributed in the interval $[-\pi, \pi]$.

The optimum receiver follows by the same argument as in 7.1:

$$\Lambda^1[\underline{r}(t)] = \int_{-\pi}^{\pi} p_\psi(\psi)d\psi \exp\{ + \frac{2}{N_o} \int_o^T \underline{r}^T(t)\, \underline{s}(t, \psi, \hat{\Theta})dt\} \tag{7-65}$$

and so:

$$L_{c_i} = \int_o^T \sqrt{\frac{2}{T}} \{r_{c_i}(t) \sin\hat{\phi}_i + r_{s_i}(t) \cos\hat{\phi}_i\}dt \tag{7-66}$$

$$L_{s_i} = \int_o^T \sqrt{\frac{2}{T}} \{- r_{c_i}(t) \cos\hat{\phi}_i + r_{s_i}(t) \sin\hat{\phi}_i\}dt \tag{7-67}$$

with

$$\hat{\phi}_i = \frac{w_c}{c} p_i \sin\hat{\Theta} \tag{7-68}$$

Defining

$$L_c = \sum_{i=1}^{N} L_{c_i} \tag{7-69}$$

$$L_s = \sum_{i=1}^{N} L_{s_i} \tag{7-70}$$

the test is

$$\ell(\underline{r}) = L_c^2 + L_s^2 \tag{7-71}$$

Mathematical model

Substitution of (7-58), (7-59) and (7-68) in (7-69) and (7-70) leads to, after some algebraic computations and for the equally spaced linear array previously considered:

$$L_c = \sqrt{PT}\left\{ \frac{\sin Nx}{\sin x}\cos\psi + \sqrt{\frac{N_o}{PT}} \sum_{i=1}^{N} (w_{n_{ci}} \sin\hat{\phi}_i + w_{n_{si}} \cos\hat{\phi}_i) \right\} \quad (7\text{-}72)$$

$$L_s = \sqrt{PT}\left\{ -\frac{\sin Nx}{\sin x}\sin\psi + \sqrt{\frac{N_o}{PT}} \sum_{i=1}^{N} (w_{n_{si}} \sin\hat{\phi}_i - w_{n_{ci}} \cos\hat{\phi}_i) \right\} \quad (7\text{-}73)$$

where

$$x = \frac{\Delta P}{2} \frac{w_c}{c} (\sin\theta - \sin\hat{\theta}) \quad (7\text{-}74)$$

$$\hat{\phi}_i = \frac{w_c}{c} p_i \sin\hat{\theta} \quad (7\text{-}75)$$

and $w_{n_{ci}}$, $w_{n_{si}}$ are two Gaussian random variables N(o, 1).

The "largest of" receiver is then mathematically equivalent to the block diagram of Fig. 7-5.

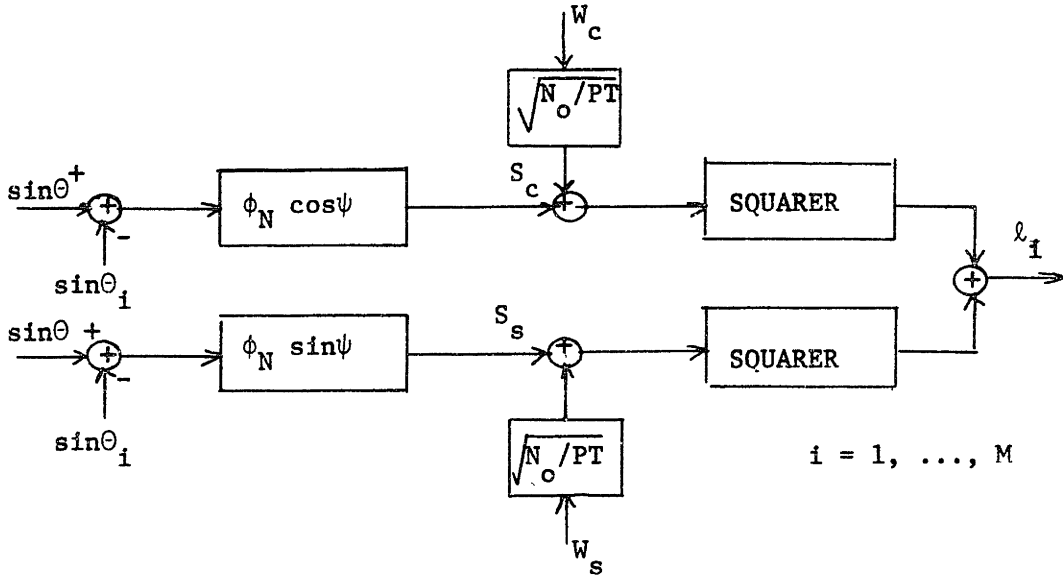


Fig. 7-5

7.5 Simulation Results*

We present now the results of the maximum likelihood (ML) simulation work.

These results were averaged over the observation noise and over the uniformly distributed unwanted parameter ψ.

Values of parameters:

$N = 10$

$SNR = 10$

$T = 1$ sec

The probability of error as given by (7-44) is

$$P_r = .082$$

The variance of the ML estimate as computed from (7-56) is

$$\text{Var} = .1102 \times 10^{-2} \quad \text{rad}^2$$

with the corresponding standard deviation

$$\text{STD} = .033 \quad \text{rad}$$

The experimental results are summarized in Table 7-1.

---

* Because the ambiguity function as given by equation (7-33) has two main peaks (e.g. at $\hat{\Theta} = o$ and $\hat{\Theta} = \bar{\pi}$, for $\Theta = o$) the complete solution of the ML estimate of the bearing angle in $(o, 2\pi)$ requires that the process be repeated with two different geometries which can be done by mechanically steering the array. But this puts no conceptual problems.

θ (rad)	$\hat{\theta}$	var (rad^2)	STD
0	-.0062	$.298 \times 10^{-3}$	.0172
.1396	.1386	$.488 \times 10^{-3}$	.022
.2792	.2768	$.284 \times 10^{-3}$	.0168
.4189	.4185	$.345 \times 10^{-3}$	.0185
.5585	.5585	$.3164 \times 10^{-3}$	.0177
.9774	.9913	$.529 \times 10^{-3}$	.023
1.256	1.178	$.533 \times 10^{-2}$	.073

Table 7-2

Table 7-2 summarizes the case when the bearing is in a neighborhood of zero.

θ (rad)	$\hat{\theta}$	var (rad^2)	STD
.0349	.03298	$.151 \times 10^{-3}$	.0123
.0698	.0688	$.196 \times 10^{-3}$	.0140
.1047	.1021	$.373 \times 10^{-3}$	.0193
.1396	.1385	$.488 \times 10^{-3}$	.0221
.1745	.1732	$.578 \times 10^{-3}$	.0240

Table 7-3

We observe from these results that the bearing estimate is close to the true value and that the standard deviation as computed from the simulation runs agrees with the theoretical result given by equation (7-56).

Initialization of the Kalman Filter:

Solved the "bearings only" problem several strategies can be implemented in order to start the Kalman filter (KF). As an example we can start the KF with

$$\underline{\hat{x}}(o) = \begin{bmatrix} o \\ o \\ \hat{\Theta}_{m\ell} \\ o \end{bmatrix} \tag{7-76}$$

From Chapter 6 we know that the KF is then able to solve for the other state variables, except for an integer multiple of wavelengths in the range estimate.

This can finally be solved by a triangularization procedure as sketched in Fig. 7-6.

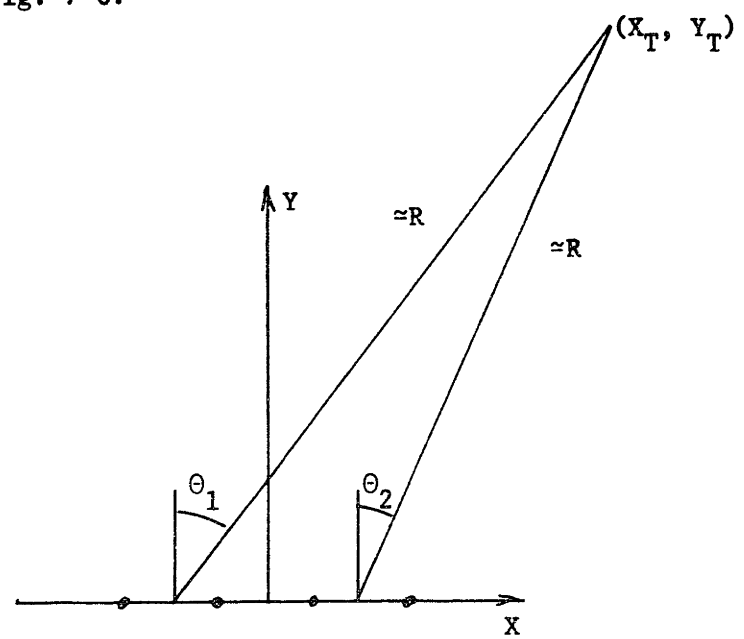


Fig. 7-6

Summary

In this chapter we solved the problem of initializing the Kalman filter(KF).

We looked into the nonlinear estimation problem from a different standpoint reducing it to a "bearings only" problem.

The bearing's maximum likelihood estimate was successfully solved via a composite hypothesis test. Then a simple triangularization provides the remaining required information to start the KF.

Chapter VIII

Conclusion

A new model, developed in [B-1], for the study of the motion of a source generating a narrow band signal, was presented.

The source's motion induces a modulation on the structure of the signal and by conveniently exploiting the changes in its temporal and spatial structure an integrated spatially-temporal tracking is possible which allows the simultaneous estimation of the range, range rate, bearing and bearing rate by a stationary observer (linear array).

The moving source was modeled by a stochastic dynamical system. The nonlinear estimation theory for lumpen state systems (Kalman filtering) was then applied. Simplicity considerations led to the implementation of first order approximations of the optimum (∞-dimensional) receiver. We implemented the discrete version of the Extended Kalman filter (EKF) and of the Maximum a Posteriori filter (MAP).

One of the issues explored in this work related to the choice of the coordinate system. We studied the source's motion in the natural frame (polar) and in the rectangular system. Because of the nonlinear nature of the problem and of the nonlinear transformation between the two coordinate systems and because of the approximations involved in the actual receivers only through comparative simulation studies could one assess their relative performance.

When implementing the polar EKF we encountered the problem of filter divergence which was related to roundoff errors in the computer and made the covariance matrix lose its positive semidefiniteness character. To prevent this the square root algorithm was implemented which improved the numerical accuracy and the overall filter's performance.

The results of Chapter 6 showed that all filters presented a threshold on the signal to noise ratio and on the driving noise level above which a good performance is to be expected and below which the performance degrades significantly and the filter diverges. The filters performance also depended on the array length (number of elements of the array) and on the geometry, being better for the broad side array as one would expect (maximum effective array length).

The polar frame is to be chosen for simplicity reasons. The EKF is more reliable than the MAP for a larger range of the parameters.

With the polar EKF filter the error covariance matrix reached practically a steady state solution. This can be used to further simplify the processor avoiding the propagation of the covariance equation.

In Chapter 5 a mathematical analysis of the filter was made with special incidence on the polar EKF. Essentially it was concluded that the filter performs two beams Z1 and Z3 which are inputs to two phase lock loops, one locking on the waveform

$$\psi_1(t) = \frac{w_c}{c} R(t) \tag{8-1}$$

and called the range channel and the other locking on

$$\psi_2(t) = \Theta(t) \tag{8-2}$$

and designated the bearing channel.

The coupling between the two channels, when a linearized analysis holds (small errors) was mainly introduced via the cross- covariance of the range and bearing. In the linear region of performance we have seen that this cross-covariance is small so that the two channels may be thought decoupled (except for the presence of the whole state estimate in the receiver's copy of the dynamical source).

With rectangular coordinates the analysis was more involved but an X-channel and a Y-channel could also be distinguished. For a broad side configuration the Y-channel presented a similar behavior to the range channel of the polar frame.

The starting of the Kalman filter required some a priori information. In Chapter 7 we remodeled the received waveform at the ith sensor as

$$r_i(t) = \sqrt{2P}\sin(w_c t + \Theta + \psi) \tag{8-3}$$

where Θ is an unknown quantity and

$$\psi = -\frac{w_c}{c} R \tag{8-4}$$

is a uniformly distributed random variable.

We then applied the maximum likelihood estimation techniques to this "bearings only" type problem which led to a composite hypothesis test.

This approach successfully solved the initialization of the Kalman filter as discussed in Chapter 7.

In Chapter 5 bounds for the range and bearing covariances were derived which allowed important theoretical conclusions about the expected performance.

Comparing the bound (5-35) of the bearings covariance with the variance (7-56) of the ML-estimate on the "bearings only" problem of Chapter 7, for the single signal case and conventional array receiver (delay and sum) (e.g. Van Trees [A-11]) under consideration we get

$$\frac{\text{Var } \hat{\Theta}_{m\ell}}{P_{33}} = N \tag{8-5}$$

We see that the Kalman filter increases the performance by a factor which is exactly the array gain of the linear array for the single signal case.

BIBLIOGRAPHY

A - Books

[A-1] - Battin, Richard H., "Astronautical Guidance", McGraw Hill Co., 1964.

[A-2] - Brockett, Roger W., "Finite Dimensional Linear Systems", John Wiley and Sons, 1970.

[A-3] - Bryson and Ho, "Applied Optimal Control", Ginn and Co., 1969.

[A-4] - Jazwinski, Andrew H., "Stochastic Processes and Filtering Theory", Academic Press, 1970.

[A-5] - Sage and Melsa, "Estimation Theory With Applications to Communications and Control", McGraw Hill, 1971.

[A-6] - Snyder, Donald L., "The State Variable Approach to Continuous Estimation With Applications to Analog Communication Theory", MIT Press, Cambridge, MA, 1969.

[A-7] - Stratonovitch, R.L., "Conditional Markov Processes and Their Application to the Theory of Optimal Control", American Elsevier Publishing Co., New York, 1968.

[A-8] - Van Trees, Harry L., "Detection, Estimation and Modulation Theory", Part I, J. Wiley, 1968.

[A-9] - Van Trees, Harry L., "Detection, Estimation and Modulation Theory", Part II, J. Wiley, 1971.

[A-10] - Van Trees, Harry L., "Detection, Estimation and Modulation Theory", Part III, J. Wiley, 1971.

[A-11] - Van Trees, Harry L., "Array Processing", MIT notes, 1969.

[A-12] - Viterbi, A.J., "Principles of Coherent Communications", McGraw Hill Co., 1966.

Papers

B - Model

[B-1] - Van Trees, Harry L., unpublished memorandum, MIT, 1971.

[B-2] - Van Trees, Harry L., internal memorandum, MIT, 1971.

C - Nonlinear Estimation Theory

[C-1] - Kushner, H.J., "Dynamical Equations for Nonlinear Filtering", J. Differential Equations, vol. 3, pp 179-190, 1967.

[C-2] - Kushner, H.J., "On the Differential Equations Satisfied by Conditional Probability Densities of Markov Processes, With Applications", J. Siam Control, vol. 2, pp 106-119, 1964.

[C-3] - Stratonovich, R.L., "Conditional Markov Processess", Th. Prob. and Its Applications, 3, 2, pp 156-178, 1960.

[C-4] - Wong, E. and Zakai, M., "On the Convergence of Ordinary Integrals to Stochastic Integrals", Ann. Math. Stat., vol. 36, pp 1500-1504, 1965.

[C-5] - Wong, E. and Zakai M., "On the Relation Between Ordinary and Stochastic Differential Equations and Applications to Problems in Control Theory", 3rd International Fed. Automatic Control Cong., London, June, 1966.

D - Nonlinear Estimation Algorithms

[D-1] - Athans, M., Wishner, R.P., and Bertolini, A., "Suboptimal State Estimation for Continuous Time Nonlinear Systems from Discrete Noisy Measurements", IEEE, Transactions on Automatic Control, vol. AC-13, No. 5, October 1968, pp 504-514.

[D-2] - Kushner, H.J., "Approximations to Optimal Nonlinear Filters", IEEE, Transactions on Automatic Control, vol. AC-12, No. 5, October 1967, pp 546-556.

[D-3] - Schwartz, Lawrence and Stear, Edwin B., "A Computational Comparison of Several Nonlinear Filters", IEEE, Transactions on Automatic Control, vol. AC-13, February 1968, pp 83-86.

E - Square Root Algorithms

[E-1] - Andrews, A., "A Square Root Formulation of the Kalman Covariance Equations", AIAA J., vol. 6, pp 1165-1166, June 1968.

[E-2] - Bellantoni, J.F. and Dodge, K.W., "A Square Root Formulation of the Kalman-Schmidt Filter", AIAA J., vol. 5, pp 1309-1314, July 1967.

[E-3] - Dyer, P. and McReynolds, S., "Extension of Square-Root Filtering to Include Process Noise", J. Optimiz Theory Appl., vol. 3, No. 6, pp 444-459, 1969.

[E-4] - Kaminski, P.G., Bryson, Arthur E., and Schmidt, Stanley F., "Discrete Square Root Filtering: A Survey of Current Techniques", IEEE, Transactions on Automatic Control, vol. AC-16, No. 6, pp 727-736, December 1971.

[E-5] - Schmidt, S.F., "Computational Techniques in Kalman Filtering", Theory and Applications of Kalman Filtering, NATO Advisory Group for Aerospace and Development, AGAR Dograph 139, February 1970.

F - Maximum Likelihood Estimation

[F-1] - Kelly, E.J., Reed, I.S., and Root, W.L., "The Detection of Radar Echoes in Noise", Part I, J. Siam, vol. 8, No. 2, pp 309-341, June 1960.

[F-2] - Kelly, E.J., Reed, I.S., and Root, W.L., "The Detection of Radar Echoes in Noise", Part II, vol. 8, pp481-510, September 1960.

[F-3] - Urkowitz, H., Haver, C.A., and Koval, J.F., "Generalized Solution in Radar Systems", Proc. IRE, vol.50, pp 2093-2105, October 1962.

[F-4] - Urkowitz, H., "The Angular Ambiguity Function of a Discrete-Continuous Array", Proc. IEEE, vol. 51, pp 1775, December 1963.

[F-5] - Urkowitz, H., "The Accuracy of Maximum Likelihood Angle Estimates in Radar and Sonar", IEEE, Transactions on Military Electronics, pp 39-45, January 1964.

Appendix A

Derivation of the Processor Equations

In this appendix* we derive the actual form of the EKF and MAP as given by equations (4-69) - (4-75). To avoid unnecessary repetitions we derive the actual mathematical models for the receivers in a general setting, presenting at each stage the final result in both coordinate frames.

Step 1) Measurement Update Equations

A. Estimator Equation

$$\underline{\hat{x}}_k = \underline{\bar{x}}_k + \underline{P}_k \underline{H}_k^T \underline{R}_k^{-1}[\underline{r}_k - \underline{h}(\underline{\bar{x}}_k, k)] \tag{A-1}$$

We recall the Jacobian matrix

$$\underline{H}_k = \begin{bmatrix} \frac{\partial \underline{h}_c}{\partial \underline{x}_k} \\ \vdots \\ \frac{\partial \underline{h}_s}{\partial \underline{x}_k} \end{bmatrix} = \begin{bmatrix} \frac{\partial h_{c_1}}{\partial \underline{x}_k} \\ \vdots \\ \frac{\partial h_{s_N}}{\partial \underline{x}_k} \end{bmatrix} \tag{A-2}$$

where $\frac{\partial h_{c_i}}{\partial \underline{x}_k}$, $\frac{\partial h_{s_i}}{\partial \underline{x}_k}$, i = 1, ..., N stand for the covector gradients:

---

* To keep the notation as simple as possible we will often delect the explicit time dependence. Also, for avoiding confusion, throughout we reserve the letter k for time dependence parameter. When explicitly indicated, either as subindex or as parameter, the letter i, i = 1, ..., 4, is used to refer to the i'th element of the vector, e.g. x_i stands for the i'th component of $\underline{x}_k$.

$$\frac{\partial h_{c_i}}{\partial \underline{x}_k} = \left[\frac{\partial h_{c_i}}{\partial x_1} \; . \; . \; . \; . \; \frac{\partial h_{c_i}}{\partial x_4} \right] \tag{A-3}$$

and similarly for $\dfrac{\partial h_{s_i}}{\partial \underline{x}_k}$.

Defining:

$$a_i = \frac{w_c}{c} \left[p_i \sin\Theta(k) - R(k) \right] \tag{A-4}$$

we have:

$$\begin{aligned} \frac{\partial h_{c_i}}{\partial \underline{x}_k} &= \sqrt{P} \cos a_i \, \nabla_{\underline{x}_k}\left(a_i\right) \\ & \qquad\qquad\qquad\qquad i = 1, \ldots, N \\ \frac{\partial h_{s_i}}{\partial \underline{x}_k} &= -\sqrt{P} \sin a_i \, \nabla_{\underline{x}_k}\left(a_i\right) \end{aligned} \tag{A-5}$$

where $\nabla_{\underline{x}_k}(.)$ is the covector gradient. Specializing to

Rectangular coordinates:

$$a_i = \frac{w_c}{c} \left[p_i \frac{x_1}{\sqrt{x_1^2 + x_3^2}} - \sqrt{x_1^2 + x_3^2} \right] \tag{A-6}$$

and so

$$\nabla^{\mathrm{r}}_{\underline{x}_k}\left(a_i\right) = \frac{w_c}{c} \frac{1}{\sqrt{x_1^2 + x_3^2}} \left[\left(p_i \frac{x_3^2}{x_1^2 + x_3^2} - x_1\right) \vdots \; 0 \; \vdots \left(-p_i \frac{x_3 x_1}{x_1^2 + x_3^2} - x_3\right) \vdots \; 0 \right] \tag{A-7}$$

Polar coordinates:

$$a_i = \frac{w_c}{c} [\, p_i \sin x_3 - x_1] \tag{A-8}$$

$$\nabla_{\underline{x}_k}^{p}(a_i) = \frac{w_c}{c} [\, -1 \;\vdots\; o \;\vdots\; p_i \cos x_3 \;\vdots\; o] \tag{A-9}$$

In a general setting

$$\nabla_{\underline{x}_k}(a_i) = [d_{1j} \quad o \quad d_{3j} \quad o] \qquad i = 1, \ldots, N \tag{A-10}$$

where

	Rectangular Coordinates	Polar Coordinates
d_{1i}	$\frac{w_c}{c} \frac{1}{\sqrt{x_1^2 + x_3^2}} \left(p_i \frac{x_3^2}{x_1^2 + x_3^2} - x_1\right)$	$-\frac{w_c}{c}$
d_{3i}	$\frac{w_c}{c} \frac{1}{\sqrt{x_1^2 + x_3^2}} \left(-p_i \frac{x_3 x_1}{x_1^2 + x_3^2} - x_3\right)$	$\frac{w_c}{c} p_i \cos x_3$

Table A-1

The Jacobian $\underline{H}_k$ is:

$$\underline{H}_k = \sqrt{P} \begin{bmatrix} \cos a_1 \, \nabla_{\underline{x}_k} a_1 \\ \vdots \\ \cos a_N \, \nabla_{\underline{x}_k} a_N \\ \\ -\sin a_1 \, \nabla_{\underline{x}_k} a_1 \\ \vdots \\ -\sin a_N \, \nabla_{\underline{x}_k} a_N \end{bmatrix} \Bigg|_{\underline{x}_k = \bar{\underline{x}}_k} \tag{A-11}$$

$$\underline{H}_k = \sqrt{P} \left[\begin{array}{cccc} d_{11}\cos a_1 & o & d_{31}\cos a_1 & o \\ \vdots & \vdots & \vdots & \vdots \\ d_{1N}\cos a_N & o & d_{3N}\cos a_N & o \\ -d_{11}\sin a_1 & o & -d_{31}\sin a_1 & o \\ \vdots & \vdots & \vdots & \vdots \\ -d_{1N}\sin a_N & o & -d_{3N}\sin a_N & o \end{array} \right] \Bigg|_{\underline{x}_k = \bar{\underline{x}}_k} \tag{A-11}$$

Let: $$\underline{B}_k = \underline{H}_k^T R_k^{-1}[\, \underline{r}_k - \underline{h}(\bar{\underline{x}}_k, k)] \tag{A-12}$$

We get:

$$\underline{B}_k = \begin{bmatrix} Z1 \\ o \\ Z3 \\ o \end{bmatrix} \tag{A-13}$$

where

$$Z1 = \frac{2T}{N_o}\sqrt{P} \sum_{i=1}^{N} d_{1i}\{\cos \bar{a}_i(r_{c_i} - h_{c_i}) - \sin \bar{a}_i(r_{s_i} - h_{s_i})\} \tag{A-14}$$

$$Z3 = \frac{2T}{N_o}\sqrt{P} \sum_{i=1}^{N} d_{3i}\{\cos \bar{a}_i(r_{c_i} - h_{c_i}) - \sin \bar{a}_i(r_{s_i} - h_{s_i})\}$$

Remark: $\bar{a}_i = a_i\Big|_{\underline{x}_k = \bar{\underline{x}}_k}$; h_{c_i}, h_{s_i} are always assumed to be evaluated at $\bar{\underline{x}}_k$, e.g. $h_{c_i} = h_{c_i}(\bar{\underline{x}}_k, k)$.

In order to get a mathematically equivalent model more suitable to analytical and simulation analysis we proceed by substituting the mathematical expressions of r_{c_i}, r_{s_i}, h_{c_i}, h_{s_i}, etc. in (A-14). By algebraic manipulation:

$$\cos \bar{a}_i (r_{c_i} - h_{c_i}) - \sin \bar{a}_i (r_{s_i} - h_{s_i}) = \sqrt{P} \sin(a_i - \bar{a}_i) + \cos \bar{a}_i \, w_{c_i} - \sin \bar{a}_i \, w_{s_i} \tag{A-15}$$

but:

$$a_i - \bar{a}_i = \frac{w_c}{c} [p_i(\sin\theta - \sin\bar{\theta}) - (R - \bar{R})] \tag{A-16}$$

or, defining:

$$\Delta b_1 = \frac{w_c}{c} (\sin\theta - \sin\bar{\theta})$$
$$\Delta b_3 = \frac{w_c}{c} (R - \bar{R}) \tag{A-17}$$

By expanding the $\sin(a_i - \bar{a}_i)$ and $\cos(a_i - \bar{a}_i)$ terms in (A-14) and collecting terms, we get:

$$Z1 = \frac{2PT}{N_o} \; S1 + \sqrt{\frac{2PT}{N_o}} \; W1$$
$$Z3 = \frac{2PT}{N_o} \; S3 + \sqrt{\frac{2PT}{N_o}} \; W3 \tag{A-18}$$

where:

$$S1 = \sum_{i=1}^{N} d_{1i} \{\cos \bar{\Delta b}_3 \sin(p_i \Delta b_1) - \sin \Delta b_3 \cos(p_i \Delta b_1)\}$$
$$S3 = \sum_{i=1}^{N} d_{3i} \{\cos \Delta b_3 \sin(p_i \Delta b_1) - \sin \Delta b_3 \cos(p_i \Delta b_1)\} \tag{A-19}$$

$$W1 = \sum_{i=1}^{N} d_{1i} \{\cos \bar{a}_i \, w_{c_i} - \sin \bar{a}_i \, w_{s_i}\}$$
$$W3 = \sum_{i=1}^{N} d_{3i} \{\cos \bar{a}_i w_{c_i} - \sin \bar{a}_i \, w_{s_i}\} \tag{A-20}$$

where in (A-20) we normalized the noise samples to unit variance.

To proceed, we write

$$d_{1i} = q_{1i} + q_1$$
$$d_{3i} = q_{3i} + q_3 \qquad \text{(A-21)}$$

where

	Rectangular Coordinates	Polar Coordinates
q_1	$-\frac{w_c}{c}\ \frac{1}{\sqrt{.}}\ x_1$	$-\frac{w_c}{c}$
q_{1i}	$\frac{w_c}{c}\ \frac{1}{\sqrt{.}}\ p_i\ \frac{x_3^2}{(.)}$	o
q_3	$-\frac{w_c}{c}\ \frac{1}{\sqrt{.}}\ x_3$	o
q_{3i}	$-\frac{w_c}{c}\ \frac{1}{\sqrt{.}}\ p_i\ \frac{x_3 x_1}{(.)}$	$\frac{w_c}{c}\ p_i \cos x_3$

where $\sqrt{.} = \sqrt{x_1^2 + x_3^2}$ and $(.) = (x_1^2 + x_3^2)$

Table A-2

We can then rewrite S1 and S3 as

$$S1 = \sum_{i=1}^{N} (q_1 + q_{1i})\ (.)$$
$$S3 = \sum_{i=1}^{N} (q_3 + q_{3i})\ (.) \qquad \text{(A-22)}$$

where as before:

$$(.) = \cos \Delta b_3 \sin(p_i\ \Delta b_1) - \sin \Delta b_3 \cos(p_i \Delta b_1) \qquad \text{(A-23)}$$

Choosing for center of coordinates the geometric center of the array, we get:

$$\begin{aligned} \Sigma q_1 \sin(p_i \Delta b_1) &= o \\ \Sigma q_3 \sin(p_i \Delta b_1) &= o \\ \Sigma q_{1i} \cos(p_i \Delta b_1) &= o \\ \Sigma q_{3i} \cos(p_i \Delta b_1) &= o \end{aligned} \qquad \text{(A-24)}$$

Calling

$$\begin{aligned} \alpha &= \sum_{i=1}^{N} \cos(p_i \Delta b_1) \\ \beta &= -\frac{\partial \alpha}{\partial(\Delta b_1)} = \sum_{i=1}^{N} p_i \sin(p_i \Delta b_1) \end{aligned} \qquad \text{(A-25)}$$

we can rewrite S1 and S3

	Rectangular Coordinates	Polar Coordinates
S1	$\frac{w_c}{c} \frac{1}{\sqrt{\bar{x}_1^2 + \bar{x}_3^2}} \left\{ \frac{\bar{x}_3^2}{\bar{x}_1^2 + \bar{x}_3^2} \cos \Delta b_3 \beta + \bar{x}_1 \sin \Delta b_3 \alpha \right\}$	$\frac{w_c}{c} \sin \Delta b_3 \alpha$
S3	$\frac{w_c}{c} \frac{1}{\sqrt{\bar{x}_1^2 + \bar{x}_3^2}} \left\{ - \frac{\bar{x}_1 \bar{x}_3}{\bar{x}_1^2 + \bar{x}_3^2} \cos \Delta b_3 \beta + \bar{x}_3 \sin \Delta b_3 \alpha \right\}$	$\frac{w_c}{c} \cos \bar{x}_3 \cos \Delta b_3 \beta$

Table A-3

Substitution in equation (A-1) gives the estimator equation

$$\hat{x}_i(k) = \bar{x}_i(k) + \frac{2PT}{N_o}\,[P_{1i}(k)\ S1 + P_{i3}(k)\ S3] + \sqrt{\frac{2PT}{N_o}}\,[P_{1i}(k)\ W1 + P_{i3}(k)\ W3]$$

$$\text{for } i = 1, \ldots, 4 \qquad \text{(A-26)}$$

Step 2) Measurement Update Covariance Equation

$$\underline{P}_k = \underline{M}_k[\underline{I} + \underline{H}_k^T \underline{R}_k^{-1} \underline{H}_k \underline{M}]^{-1} \qquad \text{(A-27)}$$

Remark: We omit thereoff the explicit time dependence unless ambiguity arises.

Recalling the form of the Jacobian $\underline{H}$, from (A-11), one gets:

$$\underline{H}^T\underline{R}^{-1}\underline{H} = \begin{bmatrix} H1 & o & H2 & o \\ o & o & o & o \\ H2 & o & H3 & o \\ o & o & o & o \end{bmatrix} \qquad \text{(A-28)}$$

where:

$$H1 = \frac{2PT}{N_o} \sum_{i=1}^{N} d_{1i}^2$$

$$H2 = \frac{2PT}{N_o} \sum_{i=1}^{N} d_{1i}d_{3i} \qquad \text{(A-29)}$$

$$H3 = \frac{2PT}{N_o} \sum_{i=1}^{N} d_{3i}^2$$

Substitution of the values of d_{1i}, d_{3i}, from Table A-1, and recalling that for the particular choice of the coordinate center, the first moment

$$\sum_{i}^{N} p_i = o \qquad \text{(A-30)}$$

Calling the second geometric moment:

$$\sigma = \sum_{i=1}^{N} p_i^2 \tag{A-31}$$

we are led to the Table A-4:

	Rectangular Coordinates	Polar Coordinates
H1	$\frac{2PT}{N_o}\left(\frac{w_c}{c}\right)^2 \frac{1}{(.)} \left[\frac{\bar{x}_3^{-4}}{(.)^2}\sigma + \bar{x}_1^{-2}N\right]$	$\frac{2PT}{N_o}\left(\frac{w_c}{c}\right)^2 N$
H2	$\frac{2PT}{N_o}\left(\frac{w_c}{c}\right)^2 \frac{\bar{x}_1\bar{x}_3}{(.)}\left[N - \frac{\bar{x}_3^{-2}}{(.)^2}\sigma\right]$	o
H3	$\frac{2PT}{N_o}\left(\frac{w_c}{c}\right)^2 \frac{1}{(.)}\left[\frac{\bar{x}_1^{-2}\bar{x}_3^{-2}}{(.)^2}\sigma + \bar{x}_3^{-2}N\right]$	$\frac{2PT}{N_o}\left(\frac{w_c}{c}\right)^2 \cos^2\bar{x}_3\,\sigma$

where $(.) = \bar{x}_1^{-2} + \bar{x}_3^{-2}$

Table A-4

Step 3) Time Update Estimator Equation (prediction)

$$\underline{\bar{x}}_{k+1} = \underset{\sim}{f}(\hat{\underline{x}}_k) \tag{A-32}$$

with the foregoing definitions of $\underset{\sim}{f}(.)$, equations (4-59) and (4-61).

Step 4) Time Update Covariance Equation (prediction)

$$\underline{M}_{k+1} = \frac{\partial \underline{f}_k}{\partial \hat{\underline{x}}} \underline{P}_k \frac{\partial \underline{f}_k^T}{\partial \hat{\underline{x}}} + \underset{\sim}{g}_k \underline{Q}_k \underset{\sim}{g}_k^T \tag{A-33}$$

where

Rectangular coordinates:

$$\frac{\partial \underset{\sim}{f}_k}{\partial \underset{\sim}{\hat{x}}} = \begin{bmatrix} 1 & T & & \\ 0 & 1 & & \\ & & 1 & T \\ & & 0 & 1 \end{bmatrix} \qquad \text{(A-34)}$$

$$\underset{\sim}{g}_k \; \underset{\sim}{Q}_k \; \underset{\sim}{g}_k^T = QT \begin{bmatrix} \frac{T^2}{4} & \frac{T}{2} & & \\ \frac{T}{2} & 1 & & \\ & & \frac{T^2}{4} & \frac{T}{2} \\ & & \frac{T}{2} & 1 \end{bmatrix} \qquad \text{(A-35)}$$

Polar coordinates:

$$\frac{\partial \underset{\sim}{f}_k}{\partial \underset{\sim}{\hat{x}}} = \begin{bmatrix} 1 & T & 0 & 0 \\ T\hat{x}_4^2 & 1 & 0 & 2T\,\hat{x}_1\hat{x}_4 \\ 0 & 0 & 1 & T \\ \dfrac{2T\,\hat{x}_2\hat{x}_4}{\hat{x}_1^2} & -\dfrac{2T\,\hat{x}_4}{\hat{x}_1} & 0 & 1 - T\dfrac{2\hat{x}_2}{\hat{x}_1} \end{bmatrix} \qquad \text{(A-36)}$$

$$\underset{\sim}{g}_k \; \underset{\sim}{Q}_k \; \underset{\sim}{g}_k = QT \begin{bmatrix} 0 & 0 & & \\ 0 & 1 & & \\ & & 0 & 0 \\ & & 0 & \frac{1}{\hat{x}_1^2} \end{bmatrix} \qquad \text{(A-37)}$$

Step 5) MAP Measurement Update Covariance Equation

$$\underline{P}_k = [\underline{M}_k^{-1} - \frac{\partial}{\partial \underline{\hat{x}}_k} \{\underline{H}_k^T \underline{R}^{-1}[\underline{r}_k - \underline{h}(\underline{\bar{x}}_k, k)]\}]^{-1} \tag{A-38}$$

We recall from equations (A-12) and (A-13):

$$\underline{B} = \underline{H}^T \underline{R}^{-1}[\underline{r} - \underline{h}] = \begin{bmatrix} Z1 \\ o \\ Z3 \\ o \end{bmatrix} \tag{A-39}$$

and so the Jacobian

$$-\frac{\partial}{\partial \underline{\hat{x}}} \underline{B} = \begin{bmatrix} -\nabla_{\underline{\hat{x}}} Z1 \\ o \\ -\nabla_{\underline{\hat{x}}} Z3 \\ o \end{bmatrix} \tag{A-40}$$

where $\nabla_{\underline{\hat{x}}}$. is the previously defined covector operator, see equation (A-3).

Because of the linearity

$$\nabla_{\underline{\hat{x}}_k} Z1 = \frac{2PT}{N_o} \nabla_{\underline{\hat{x}}} S1 + \sqrt{\frac{2PT}{N_o}} \nabla_{\underline{\hat{x}}} W1 \tag{A-41}$$

$$\nabla_{\underline{\hat{x}}} Z3 = \frac{2PT}{N_o} \nabla_{\underline{\hat{x}}} S3 + \sqrt{\frac{2PT}{N_o}} \nabla_{\underline{\hat{x}}} W3 \tag{A-42}$$

S1, S3, W1, W3, are only functions of $\hat{x}_1$, $\hat{x}_3$ and given their definitions

$$-\frac{\partial}{\partial \underline{\hat{x}}}\underline{B} = \begin{bmatrix} H1 & o & H2 & o \\ o & o & o & o \\ H2 & o & H3 & o \\ o & o & o & o \end{bmatrix} \tag{A-43}$$

which is formally equivalent to (A-28).

Recalling (A-25) and further defining:

$$\gamma = \frac{\partial \beta}{\partial (\Delta b_1)} = -\frac{\partial^2 \alpha}{\partial (\Delta b_1)^2} = \sum_{i=1}^{N} p_i^2 \cos(p_i \Delta b_1) \tag{A-44}$$

$$\hat{\gamma}W1_i = \sin \bar{a}_i \, w_{c_i} + \cos \bar{a}_i \, w_{s_i} \tag{A-45}$$

$$\hat{\gamma}W2_i = \cos \bar{a}_i \, w_{c_i} - \sin \bar{a}_i \, w_{s_i} \tag{A-46}$$

we get

Polar coordinates

$$H1 = \frac{2PT}{N_o}\left(\frac{w_c}{c}\right)^2 \cos \Delta b_3 \, \alpha + \sqrt{\frac{2PT}{N_o}}\left(\frac{w_c}{c}\right)^2 \sum_{i=1}^{N} \hat{\gamma}W_i \tag{A-47}$$

$$H2 = -\frac{2PT}{N_o}\left(\frac{w_c}{c}\right)^2 \cos \bar{x}_3 \sin \Delta b_3 \beta + \sqrt{\frac{2PT}{N_o}}\left(\frac{w_c}{c}\right)^2 \cos \bar{x}_3 \sum p_i \hat{\gamma}W_i \tag{A-48}$$

$$H3 = \frac{2PT}{N_o}\left\{\left(\frac{w_c}{c}\right)^2 \cos^2 \bar{x}_3 \cos \Delta b_3 \gamma - \frac{w_c}{c} \sin \bar{x}_3 \cos \Delta b_3 \beta\right\} +$$

$$+\sqrt{\frac{2PT}{N_o}}\left(\frac{w_c}{c}\right)^2 \cos^2 \bar{x}_3 \sum_{i=1}^{N} p_i^2 \hat{\gamma}W_i \tag{A-49}$$

Rectangular coordinates

$$-\frac{2PT}{N_o}\frac{\partial S1}{\partial \bar{x}_1} = G1\left\{\left[\frac{2}{\sqrt{.}}\cos 3 + 2\frac{w_c}{c}\sin 3\right]\frac{\bar{x}_1\bar{x}_3^{-2}}{(.)}\beta + \frac{\bar{x}_3^{-4}}{(.)^2}\cos 3\gamma - \right.$$

$$\left. - \left[\frac{\bar{x}_3^{-2}}{\sqrt{.}}\sin 3 + \frac{w_c}{c}\bar{x}_1^{-2}\cos 3\right]\alpha\right\} \qquad \text{(A-50)}$$

$$-\frac{2PT}{N_o}\frac{\partial S3}{\partial \bar{x}_3} = G1\left\{\left[\frac{\bar{x}_1^{-2} - \bar{x}_3^{-2}}{\sqrt{.}}\cos 3 - 2\bar{x}_3^{-2}\frac{w_c}{c}\sin 3\right]\frac{\bar{x}_1}{(.)}\beta + \frac{(\bar{x}_1\bar{x}_3)^2}{(.)^2}\frac{w_c}{c}\cos 3\gamma + \right.$$

$$\left. + \left[\frac{\bar{x}_1^{-2}}{\sqrt{.}}\sin 3 + \bar{x}_3^{-2}\frac{w_c}{c}\cos 3\right]\alpha\right\} \qquad \text{(A-51)}$$

$$-\frac{2PT}{N_o}\frac{\partial S_j}{\partial \bar{x}_1} = G1\left\{\left[\frac{\bar{x}_3^{-2} - 2\bar{x}_1^{-2}}{\sqrt{.}}\cos 3 + (\bar{x}_3^{-2} - \bar{x}_1^{-2})\frac{w_c}{c}\sin 3\right]\frac{\bar{x}_3}{(.)}\beta + \right.$$

$$\left. + \frac{\bar{x}_1\bar{x}_3^{-3}}{(.)^2}\frac{w_c}{c}\cos 3\gamma - \left[\frac{1}{\sqrt{.}}\sin 3 - \frac{w_c}{c}\cos 3\right]\bar{x}_1\bar{x}_3\alpha\right\} \qquad \text{(A-52)}$$

$$-\sqrt{\frac{2PT}{N_o}}\frac{\partial W1}{\partial \bar{x}_1} = G2\left\{\sum_1^N \left[\frac{3p_i\bar{x}_1}{(.)} + 1\right]\frac{\bar{x}_3^{-2}}{\sqrt{.}}\hat{\gamma}W2_i - \left[\frac{p_i\bar{x}_3^{-2}}{\bar{x}_1^{-2} + \bar{x}_3^{-2}} - \bar{x}_1\right]^2\frac{w_c}{c}\hat{\gamma}W1_i\right\} \qquad \text{(A-53)}$$

$$-\sqrt{\frac{2PT}{N_o}}\frac{\partial W3}{\partial \bar{x}_3} = G2\left\{\sum_1^N \left[p_i\frac{\bar{x}_1^{-2} - 2\bar{x}_3^{-2}}{(.)} - \bar{x}_1\right]\frac{\bar{x}_1}{\sqrt{.}}\hat{\gamma}W2_i + \left[p_i\frac{\bar{x}_1}{(.)} + 1\right]^2\bar{x}_3^{-2}\frac{w_c}{c}\hat{\gamma}W1_i\right\}$$

(A-54)

$$-\sqrt{\frac{2PT}{N_o}}\frac{\partial W_i}{\partial \bar{x}_i} = G2\left\{\sum_1^N \left[p_i\frac{\bar{x}_3^{-2} - 2\bar{x}_1^{-2}}{(.)} + \bar{x}_1\right]\frac{\bar{x}_3}{\sqrt{.}}\hat{\gamma}W2_i + \right.$$

$$\left. + \left[p_i\frac{\bar{x}_1}{(.)} + 1\right](-p_i\frac{\bar{x}_3^{-2}}{(.)} + \bar{x}_1)\bar{x}_3\frac{w_c}{c}\hat{\gamma}W1_i\right\} \qquad \text{(A-55)}$$

where:

$$G1 = \frac{w_c}{c} \frac{2PT}{N_o} \frac{1}{\bar{x}_1^{-2} + \bar{x}_3^{-2}}$$

$$G2 = \frac{w_c}{c} \sqrt{\frac{2PT}{N_o}} \frac{1}{\bar{x}_1^{-2} + \bar{x}_3^{-2}}$$

$$i \neq j, \; i = 1, 3; \quad j = 1, 3$$

$$(.) = (\bar{x}_1^{-2} + \bar{x}_3^{-2}) \tag{A-56}$$

$$\sqrt{.} = \sqrt{\bar{x}_1^{-2} + \bar{x}_3^{-2}}$$

$$\sin 3 = \sin(\Delta b_3)$$

$$\cos 3 = \cos(\Delta b_3)$$

Finally:

$$H1 = - \frac{2PT}{N_o} \frac{\partial S1}{\partial \bar{x}_1} - \sqrt{\frac{2PT}{N_o}} \frac{\partial W1}{\partial \bar{x}_1} \tag{A-57}$$

$$H2 = - \frac{2PT}{N_o} \frac{\partial S_i}{\partial \bar{x}_i} - \sqrt{\frac{2PT}{N_o}} \frac{\partial W_i}{\partial \bar{x}_i} \tag{A-58}$$

$$H3 = - \frac{2PT}{N_o} \frac{\partial S3}{\partial \bar{x}_3} - \sqrt{\frac{2PT}{N_o}} \frac{\partial W3}{\partial \bar{x}_3} \tag{A-59}$$

Appendix B

In this appendix we present the SQRT algorithm used in the implementation of the optimum processors as discussed in section 4-5. The following coding has been considered in the literature, e.g. [E-4].

Step 1) Matrix Decomposition: Cholesky Algorithm

The square root form of a symmetric positive semidefinite n x n matrix $\underline{P}$ can be written as

$$\underline{P} = \underline{W}\,\underline{W}^T \tag{B-1}$$

with

$$\underline{W} = \begin{bmatrix} W_{11} & & \\ \cdot & & \underline{o} \\ W_{n1} & \cdots & W_{nn} \end{bmatrix} \tag{B-2}$$

where the W_{ij} are given by the Cholesky recursive algorithm

$$W_{ii} = \sqrt{P_{ii} - \sum_{j=1}^{i-1} W_{ij}^2} \tag{B-3}$$

$$W_{ji} = \begin{cases} o & j < i \\ \dfrac{1}{W_{ii}}\left(P_{ji} - \sum_{k=1}^{i-1} S_{jk}\, S_{ik}\right) & j = i+1, n \end{cases} \tag{B-4}$$

Step 2) Measurement Update

I. Estimator Equation: same as equation (4-46)

II. Covariance Equation:

Recall equation (4-96)

$$\underline{B} = \underline{H}^T \underline{R}^{-1} \underline{H} = \sum_{j=1}^{\ell} \underline{b}_j \underline{b}_j^T \tag{B-5}$$

Apply ℓ times the following routine

$$\underline{W}_j = \underline{S}_{j-1} \left[\underline{I} - \frac{\underline{b}_j \underline{S}_{j-1}^T \underline{S}_{j-1} \underline{b}_j^T}{a(1 + c)}\right] \tag{B-6}$$

with

$$a = \underline{b}_j^T \underline{S} \underline{S}^T \underline{b}_j + 1 \tag{B-7}$$

$$c = \sqrt{1/(1 + \underline{b}_j^T \underline{S} \underline{S}^T \underline{b}_j)} \tag{B-8}$$

starting with

$$\begin{aligned} \underline{S}_o &= \underline{S} \\ \underline{S}_j &= \underline{W}_j \qquad j = 1, \ldots, \ell - 1 \\ \underline{W} &= \underline{W}_\ell \end{aligned} \tag{B-9}$$

Step 3) Time Propagation

I. Estimator Equation: same as equation (4-49)

II. Covariance Equation:

Recall equations (4-80) and (4-81)

$$\underline{M} = \underline{S} \underline{S}^T \tag{B-10}$$

$$\underline{G} \underline{Q} \underline{G}^T = \underline{V} \underline{V}^T \tag{B-11}$$

Then:

$$\begin{bmatrix} \underline{S}^T(k+1) \\ \cdots\cdots \\ \underline{o} \end{bmatrix} = \underline{T} \begin{bmatrix} \underline{W}^T(k) \quad \frac{\partial \underline{f}^T}{\partial \underline{x}}(k) \\ \cdots\cdots\cdots\cdots \\ \underline{V}^T \end{bmatrix} \tag{B-12}$$

where $\underline{T}$ is an orthogonal transformation such that equation (B-12) is satisfied with $\underline{S}$ in lower triangular form.

There are several routines for the construction of $\underline{T}$. We use the

Modified Gram-Schmidt (MGS) Algorithm

Notation: Let

$$\begin{bmatrix} \underline{S}^T \\ \cdots \\ \underline{o} \end{bmatrix} = \underline{T}\,\underline{A} \tag{B-13}$$

where $\underline{S}$ is n x n lower triangular matrix.

Define

$$\begin{aligned} &\underline{A}^{(1)} \triangleq \underline{A} \\ &\underline{A}_i^{(j)} \triangleq \text{ith column of } \underline{A}^{(j)} \\ &\underline{A}_{i\ell}^{(j)} \triangleq \text{element in ith row and } \ell\text{th column of } \underline{A}^{(j)}. \end{aligned} \tag{B-14}$$

Then the MGS algorithm follows:

For j = 1, n

$$\sigma_j = \sqrt{\underline{A}_j^{(j)T}\,\underline{A}_j^{(j)}} \tag{B-15}$$

$$S_{ij} = \begin{cases} o & i = 1, j-1 \\ \sigma_j & i = j \\ \frac{1}{\sigma_j}\,\underline{A}_j^{(j)T}\,\underline{A}_i^{(j)} & i = j+1, n \end{cases} \tag{B-16}$$

$$\underline{A}_i^{(j+1)} = \underline{A}_i^{(j)} - \frac{S_{ij}}{\sigma_j} \underline{A}_j^{(j)} \qquad i = j + 1, \ldots, n \qquad \text{(B-17)}$$