General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Produced by the NASA Center for Aerospace Information (CASI)

Final Report

SUBOPTIMAL FILTERING

Part 3:

LIMITED MEMORY OPTIMAL FILTERING

Andrew H. Jazwinski
Stanley F. Schmidt
Ann E. Bailie
Norman Levine

Report No. 68-12
Contract NAS 5-11048

GPO PRICE $ ______
CSFTI PRICE(S) $ ______

Hard copy (HC) 3.00
Microfiche (MF) .65

ff 653 July 65

May 1968

FACILITY FORM 602
N68-33829 (ACCESSION NUMBER) (THRU)
13 (PAGES) 1 (CODE)
CR 96579 (NASA CR OR TMX OR AD NUMBER) 21 (CATEGORY)

Washington Office
Analytical Mechanics Associates, Inc.
9301 Annapolis Road
Lanham, Maryland 20801

FOREWORD

This final report is in four parts:

Part 1: ADAPTIVE FILTERING

Part 2: COMPENSATION FOR MODELING ERRORS IN ORBIT DETERMINATION PROBLEMS

Part 3: LIMITED MEMORY OPTIMAL FILTERING

Part 4: TEST-BED COMPUTER PROGRAM

The first three parts describe several suboptimal filter concepts developed under this Contract. A number of these filters have been simulated in the rectilinear orbit problem. These simulations are described therein. In order to provide a more realistic environment for testing these suboptimal filters, a more general test-bed computer program is under development. This program enables the simulation of real observation schedules and combined effects of dynamical model errors in three-dimensional satellite motion. This program is briefly described in Part 4.

The authors wish to express their appreciation for the active interest and support of this work by Mr. R. K. Squires of Goddard Space Flight Center. The contributions of Dr. H. Wolf and Mr. S. Pines are also gratefully acknowledged.

LIMITED MEMORY OPTIMAL FILTERING*

Andrew H. Jazwinski
Manager, Guidance and Control
Analytical Mechanics Associates, Inc.
Lanham, Maryland

Abstract

Linear and nonlinear optimal filters with limited memory length are developed. The filter output is the conditional probability density function and, in the linear-Gaussian case, is the conditional mean and covariance matrix where the conditioning is only on a fixed amount of most recent data. This is related to maximum likelihood (least squares) estimation. These filters have application in problems where standard filters diverge due to dynamical model errors. This is demonstrated via numerical simulations.

1. Introduction and Summary

The filtering theory developed to date (e.g. [1]-[4]) assumes that system dynamics are completely known and are precisely modeled in the filter. Clearly, this is never true in practice, and furthermore, finite arithmetic precludes the exact computation of the filter state. The modeling and computational errors which are invariably present may not present particular difficulties when the noise inputs to the system are large. When these are small, however; when model errors (such as dynamic biases) exist; and when the filter operates over long time intervals (over much data), its operation is sometimes rendered totally unacceptable.

This is often the case in the determination of space-vehicle trajectories ([5],[6]) via a 'modified' Kalman filter. The observed phenomenon is a 'divergence' of the errors in the estimates to values totally inconsistent with the rms values predicted by theory. The covariance matrix becomes unrealistically small (optimistic); the filter gain thus becomes small, and subsequent measurements are ignored. The state and its estimate then diverge, due to model errors in the filter.

An analysis of error 'divergence' may be found in [6]. Some techniques found useful in controlling divergence are outlined in [5] and [6]. These range from arbitrary incrementation of the covariance matrix, to keep its elements above an a priori lower bound, to a computational error noise model [7]; experimentally determined system noise input levels [6]; modeling additional state variables (biases) and including their uncertainties in the filter, with or without actually estimating such biases [7],[8]. This last technique requires state augmentation and may not be practical from the computational point of view. More recently, Schmidt [9] proposed two new schemes. One computes an estimate which is a linear combination of the estimate given all prior data with the estimate given no prior data. Past information (data) is thus degraded. The other scheme imposes a priori lower bounds on certain projections of the covariance matrix. A method which 'covers' modeling errors with noise and adaptively estimates the noise variance was recently proposed by this author [10]. This adaptive filter maintains small residuals (estimation errors) by automatically degrading the covariance matrix when residuals become large.

This paper presents a new optimal filter which has direct application to problems in which the standard filters diverge because of dynamical model errors. This new filter, called the Limited Memory Filter, computes the conditional probability density function (in the nonlinear case) and its parameters, namely the mean and covariance matrix (in the linear case), where the conditioning is on a pre-specified amount of most recent data (not the data and the initial condition). This is related to maximum likelihood estimation. In terms of discrete measurements, the conditioning is on the most recent N measurements, where N is fixed a priori. The Limited Memory filter is not a 'batch' processor.* It is truly a filter in that discrete observations are processed one by one and,

---

* Research supported by NASA Goddard Space Flight Center under Contracts NAS 5-9085 and NAS 5-11048.

* Similar results can clearly be obtained by making independent least squares fits to each batch of N measurements, at least in the linear case.

in certain modes of operation, need not be stored in memory. This filter can, however, be used to process batches of data which have been pre-processed and 'compressed' for easy storage.

That the Limited Memory filter represents a reasonable approach to problems where standard filters diverge seems obvious. If the dynamical model is sufficiently accurate over only a limited time arc, then data given at earlier times can be worse than useless and should be discarded. This will be demonstrated in some simulations. The idea of limiting the filter memory has been utilized by others [12].

The next section defines the problem and gives the results for the nonlinear case. This is specialized to the linear-Gaussian case in Section 3. Simulations which demonstrate the performance of the Limited Memory filter in a situation where the standard (Kalman) filter diverges are presented in Section 4.

2. Nonlinear Results

Let the dynamical system be described by the vector differential equation

$$\dot{x}_t = f(x_t, t) \quad , \qquad t \geq 0 \ , \tag{1}$$

where x_o is random with specified probability density function (pdf) $p(x_o)$. Note the absence of a random forcing term in (1). This restriction is essential in the following development. Discrete, nonlinear, noise-corrupted observations are taken on system (1)

$$y_k = h(x_k, t_k) + v_k \quad , \qquad k = 1, 2, \ldots , \tag{2}$$

where $\{v_k\}$ is a zero-mean, white, Gaussian vector noise sequence with

$$\mathcal{E}\left\{v_k v_\ell^T\right\} = R_k \, \delta_{k\ell} \quad , \qquad R_k > 0 \ . \tag{3}$$

Consider the sequence of data

$$y_1, \ldots, y_m, y_{m+1}, \ldots, y_n, \ldots ; \ n-m = N > 0 \tag{4}$$

and define

$$\begin{aligned} \mathcal{Y}_m &\equiv \{y_1, \ldots, y_m\} \\ \mathcal{Y}_n &\equiv \{y_1, \ldots, y_n\} \\ \mathcal{Y}_{n-m} &\equiv \{y_{m+1}, \ldots, y_n\} \end{aligned} \tag{5}$$

<u>Lemma</u>. <u>Assume dynamical system (1) with observations (2). Suppose that the required pdf's exist. Then</u>

$$p(\mathcal{Y}_{n-m} | x_n) = c \, \frac{p(x_n | \mathcal{Y}_n)}{p(x_n | \mathcal{Y}_m)} \tag{6}$$

<u>where</u>

$$c = \frac{p(\mathcal{Y}_n)}{p(\mathcal{Y}_m)}$$

<u>is a constant, independent of</u> x_n. <u>Also,</u>

$$p(x_n | \mathcal{Y}_{n-m}) = c_1 \, \frac{p(x_n | \mathcal{Y}_n)}{p(x_n | \mathcal{Y}_m)} \tag{7}$$

<u>where</u>

$$c_1 = \frac{p(\mathcal{Y}_n)}{p(\mathcal{Y}_m) \, p(\mathcal{Y}_{n-m})} \, p(x_n) \ .$$

Now $p(x_n | \mathcal{Y}_n)$ is the output of the nonlinear filter consisting of Kolmogorov's forward equation and Bayes' rule ([4],[11]). $p(x_n | \mathcal{Y}_m)$ is a 'prediction' computed via Kolmogorov's forward equation with initial condition $p(x_m | \mathcal{Y}_m)$ at time m. Then (6) gives the likelihood function which can be used to compute the maximum likelihood estimate of x_n based on $\mathcal{Y}_{n-m}$.

<u>Proof of Lemma</u>. By Bayes' rule

$$p(x_n | \mathcal{Y}_n) = \frac{p(\mathcal{Y}_n | x_n) \, p(x_n)}{p(\mathcal{Y}_n)}$$

$$= \frac{p(\mathcal{Y}_m, \mathcal{Y}_{n-m} | x_n) \, p(x_n)}{p(\mathcal{Y}_n)}$$

But

$$p(\mathcal{Y}_m, \mathcal{Y}_{n-m} | x_n) = p(\mathcal{Y}_m | \mathcal{Y}_{n-m}, x_n) \, p(\mathcal{Y}_{n-m} | x_n)$$

and since system (1) is noise-free

$$p(\mathcal{Y}_m | \mathcal{Y}_{n-m}, x_n) = p(\mathcal{Y}_m | x_n) \ .$$

Therefore

$$p(x_n | \mathcal{Y}_n) = \frac{p(\mathcal{Y}_m | x_n) \, p(\mathcal{Y}_{n-m} | x_n) \, p(x_n)}{p(\mathcal{Y}_n)}$$

Now applying Bayes' rule to $p(\mathcal{Y}_m|x_n)$

$$p(\mathcal{Y}_m|x_n) = \frac{p(x_n|\mathcal{Y}_m)\,p(\mathcal{Y}_m)}{p(x_n)}$$

and rearranging terms

$$p(x_n|\mathcal{Y}_n) = \frac{p(\mathcal{Y}_m)}{p(\mathcal{Y}_n)}\,p(x_n|\mathcal{Y}_m)\,p(\mathcal{Y}_{n-m}|x_n)$$

which is (6). Equation (7) follows after application of Bayes' rule to $p(\mathcal{Y}_{n-m}|x_n)$

$$p(\mathcal{Y}_{n-m}|x_n) = \frac{p(x_n|\mathcal{Y}_{n-m})\,p(\mathcal{Y}_{n-m})}{p(x_n)} \quad . \qquad (8)$$

The conditional density $p(x_n|\mathcal{Y}_{n-m})$ in (7) is the density of x_n given $\mathcal{Y}_{n-m}$ and the density of x_o. It is defined in terms of the joint density $p(x_n, \mathcal{Y}_{n-m})$ and the marginal density $p(\mathcal{Y}_{n-m})$. We wish to define the conditional density of x_n given only $\mathcal{Y}_{n-m}$, excluding any prior information about x_o. To do this, we suppose formally that, for each n, $p(x_n)$ is Gaussian with zero-mean and covariance matrix

$$(1/\epsilon)I$$

and define $p(x_n|\mathcal{Y}_{n-m}|)$, the density of x_n given only $\mathcal{Y}_{n-m}$, by

$$p(x_n|\mathcal{Y}_{n-m}|) = \lim_{\epsilon\to 0} p(x_n|\mathcal{Y}_{n-m}) \quad , \qquad (9)$$

assuming the limit exists.

Theorem 1. Hypotheses of the Lemma. Then

$$p(x_n|\mathcal{Y}_{n-m}|) = c_2\,\frac{p(x_n|\mathcal{Y}_n)}{p(x_n|\mathcal{Y}_m)} \quad , \qquad (10)$$

where c_2 is a constant, independent of x_n. [The normalizing constant can obviously be determined by

$$\frac{1}{c_2} = \int \frac{p(x_n|\mathcal{Y}_n)}{p(x_n|\mathcal{Y}_m)}\,dx_n \quad .]$$

Proof. From (8)

$$\frac{p(\mathcal{Y}_{n-m}|x_n)}{p(x_n|\mathcal{Y}_{n-m})} = \frac{p(\mathcal{Y}_{n-m})}{p(x_n)} \quad . \qquad (11)$$

Now $p(\mathcal{Y}_{n-m}|x_n)$ does not depend on the statistics of x_n. Assuming the limit in (9) exists and taking the limit in (11), we get

$$\frac{p(\mathcal{Y}_{n-m}|x_n)}{p(x_n|\mathcal{Y}_{n-m}|)} = \lim_{\epsilon\to 0} \frac{p(\mathcal{Y}_{n-m})}{p(x_n)} = \bar{c} \qquad (12)$$

which proves that the limit in (12) (namely $\bar{c}$) exists. $\bar{c}$ is clearly independent of x_n. Equation (12), together with (6) of the Lemma, proves the theorem.

As was already noted, $p(x_n|\mathcal{Y}_n)$ and $p(x_n|\mathcal{Y}_m)$ are outputs of the nonlinear filter and predictor. Equation (10) then produces the density of x_n given only $\mathcal{Y}_{n-m}$. In view of (12), $p(\mathcal{Y}_{n-m}|x_n)$ and $p(x_n|\mathcal{Y}_{n-m}|)$ are equal up to a multiplicative constant. Thus any estimate of x_n obtained from $p(\mathcal{Y}_{n-m}|x_n)$ can also be obtained from $p(x_n|\mathcal{Y}_{n-m}|)$ by the same operation, and vice versa.

3. Linear Limited Memory Filter

The above result is now specialized to the linear case. A discrete dynamical system is considered

$$x_{k+1} = \Phi_{k+1,k}\,x_k \quad , \qquad k = 0, 1, \ldots , \qquad (13)$$

$x_o \sim N(\hat{x}_o, P_o)$, with observations

$$y_k = M_k x_k + v_k \quad , \qquad k = 1, 2, \ldots , \qquad (14)$$

where the noise sequence $\{v_k\}$ is as previously defined (3). Let

$$\hat{x}_{n|n} = \mathcal{E}\{x_n|\mathcal{Y}_n\}$$

$$P_{n|n} = \mathcal{E}\{(x_n - \hat{x}_{n|n})(x_n - \hat{x}_{n|n})^T|\mathcal{Y}_n\} \qquad (15a)$$

$$\hat{x}_{n|m} = \mathcal{E}\{x_n | \mathcal{J}_m\} = \Phi_{n,m}\, \hat{x}_{m|m}$$

$$P_{n|m} = \mathcal{E}\{(x_n - \hat{x}_{n|m})(x_n - \hat{x}_{n|m})^T | \mathcal{J}_m\} \qquad (15b)$$

$$= \Phi_{n,m}\, P_{m|m}\, \Phi^T_{n,m} \quad .$$

The assumption of existence of the pdf's in the Lemma can now be replaced by an explicit observability condition and the following holds.

Theorem 2. Assume dynamical system (13) with observations (14). Suppose x_n is observable with respect to $\mathcal{J}_{n-m}$. Then the maximum likelihood estimate of x_n based on $\mathcal{J}_{n-m}$ is given by

$$\hat{x}_{n|N} = P_{n|N}\,(P^{-1}_{n|n}\,\hat{x}_{n|n} - P^{-1}_{n|m}\,\hat{x}_{n|m}) \quad ,$$

$$P^{-1}_{n|N} = P^{-1}_{n|n} - P^{-1}_{n|m} \quad , \qquad (16)$$

where $P_{n|N}$ is the covariance of the errors in the estimate.

Remark 1. $\hat{x}_{n|n}, P_{n|n}$ are outputs of the Kalman filter [1]. $\hat{x}_{n|m}, P_{n|m}$ are simply predictions [see (15b)].

Remark 2. The observability condition guarantees the existence of the inverse of $P^{-1}_{n|N}$, which is required in the computation of $\hat{x}_{n|N}$.

Remark 3. Solved for $\hat{x}_{n|n}, P_{n|n}$, (16) can be used to process 'batches' of data characterized by $\hat{x}_{n|N}, P_{n|N}$.

Remark 4. Equation (16) may be used to eliminate or replace initial conditions $\hat{x}_o, P_o$ which are frequently arbitrarily set to start the Kalman filter.

Remark 5. Equations (16) have the intuitive limiting properties. With $n=m$ $(N=0)$, $P^{-1}_{n|N} = 0$ (no information) and $\hat{x}_{n|N}$ is arbitrary. For N large, $P^{-1}_{n|N} \approx P^{-1}_{n|n}$ and $\hat{x}_{n|N} \approx \hat{x}_{n|n}$.

Proof. In view of the hypotheses, the pdf's in (6) exist and are Gaussian, with the parameters listed in (15a) and (15b). Therefore, from (6)

$$p(\mathcal{J}_{n-m}|x_n) = c_3 \exp\left\{-\frac{1}{2}\Big[(x_n - \hat{x}_{n|n})^T P^{-1}_{n|n}(x_n - \hat{x}_{n|n}) - (x_n - \hat{x}_{n|m})^T P^{-1}_{n|m}(x_n - \hat{x}_{n|m})\Big]\right\} . \qquad (17)$$

To maximize (17) (with respect to x_n), minimize

$$\frac{1}{2}\Big[(x_n - \hat{x}_{n|n})^T P^{-1}_{n|n}(x_n - \hat{x}_{n|n}) - (x_n - \hat{x}_{n|m})^T P^{-1}_{n|m}(x_n - \hat{x}_{n|m})\Big]. \qquad (18)$$

The gradient of (18) is set equal to zero,

$$P^{-1}_{n|n}(x_n - \hat{x}_{n|n}) - P^{-1}_{n|m}(x_n - \hat{x}_{n|m}) = 0 \quad . \qquad (19)$$

$\hat{x}_{n|N}$ is the solution of (19). $P^{-1}_{n|N}$ is the matrix of second derivatives

$$P^{-1}_{n|n} - P^{-1}_{n|m} \equiv P^{-1}_{n|N} \quad , \qquad (20)$$

and (16) follows.

Theorem 3. Hypotheses of Theorem 2. Then $\hat{x}_{n|N}$ given by (16) and $P_{n|N}$ are, respectively, the mean and covariance matrix of $p(x_n|\mathcal{J}_{n-m}|)$.

Proof. This follows immediately from Theorem 1 and the fact that, in this case, the densities in (10) are Gaussian (the mean of a Gaussian density coincides with its mode). In this linear-Gaussian case, we can prove the existence of $p(x_n|\mathcal{J}_{n-m}|)$. $\mathcal{J}_{n-m}$ can be regarded as a single (vector) measurement

$$\mathcal{J}_{n-m} = A\,x_n + w \qquad (21)$$

with

$$\mathcal{E}\{w\} = 0 \quad , \quad \mathcal{E}\{w\,w^T\} = Q > 0 \quad ,$$

where the matrices A and Q are appropriately defined. $\mathcal{J}_{n-m}$ can be processed via the Kalman filter

$$\hat{x}_{n|N,p(x_n)} = (P^{-1}_n + A^T Q^{-1} A)^{-1}(A^T Q^{-1}\mathcal{J}_{n-m} + P^{-1}_n \hat{x}_n) \quad ,$$

$$P^{-1}_{n|N,p(x_n)} = P^{-1}_n + A^T Q^{-1} A \quad . \qquad (22)$$

$\hat{x}_n, P_n$ are the parameters of $p(x_n)$, while $\hat{x}_{n|N,p(x_n)}$, $P_{n|N,p(x_n)}$ are the parameters of $p(x_n|\mathcal{Y}_{n-m})$. Now

$$\hat{x}_{n|N} = \lim_{\epsilon \to 0} \hat{x}_{n|N,p(x_n)} = (A^T Q^{-1} A)^{-1} A^T Q^{-1} \mathcal{Y}_{n-m} \quad , \tag{23}$$

$$P^{-1}_{n|N} = \lim_{\epsilon \to 0} P^{-1}_{n|N,p(x_n)} = A^T Q^{-1} A \quad ,$$

exist since $A^T Q^{-1} A$ is non-singular (observability condition). $\hat{x}_{n|N}$, $P_{n|N}$ are the parameters of $p(x_n|\mathcal{Y}_{n-m}|)$.

The linear Limited Memory filter (16), applied directly, generates an estimate based on a 'moving window' of the most recent N observations. It essentially requires two Kalman filters and a predictor. N observations have to be stored in memory, and three matrix inversions are required at each step.

Aside from the problems of storage and matrix inversions, this type operation is clearly unacceptable in problems where the Kalman filter diverges.* An acceptable procedure, which will at the same time eliminate the storage requirement and involve infrequent matrix inversions, is to discard the 'conditioning' on old data in batches of N according to the following procedure. Run the Kalman filter to N and then repeat the following cycle:

(0) $k = 1$,

(i) Run Kalman filter from kN to $(k+1)N$,

(ii) Predict from kN to $(k+1)N$,

(iii) Discard the 'conditioning' on the observations in $[t_{kN} - t_{(k-1)N}]$ (and initial conditions for $k=1$) via (16), thus generating new initial conditions for filter and predictor in (i) and (ii) ,

(iv) Set $k=k+1$ and return to (i) .

This procedure produces limited memory estimates, with memory varying between N and 2N. Clearly, variations on this are possible, and N itself may be varied in the process. N can be chosen so that the dynamical model represents an adequate approximation over time arcs of length 2N. The matrix inversion required in (16) every N observations provides a check on the observability of the state with respect to the data retained. This mode of operation is utilized in the simulations of the following section.

* The two required Kalman filters would diverge.

4. Simulations

The dynamical system chosen for numerical study is the 'rectilinear' orbit problem with dynamics

$$\ddot{x} = -\frac{\mu}{x^2} \qquad (x \text{ scalar}) \tag{24}$$

μ is the Gravitational Constant times earth mass, 19.9094165 er^3/hr^2 (er - earth radii , hr - hour). In first order form

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = -\frac{\mu}{x_1^2} \quad , \tag{25}$$

where x_1 is position and x_2 velocity.

The observations** consist of discrete positional (range) data

$$y_k = (x_1)_k + v_k \tag{26}$$

with $\mathcal{E}\{v_k\} = 0$, $\mathcal{E}\{v_k v_\ell\} = r\,\delta_{k\ell}$, $r = 1.\times 10^{-7}$ er^2.

The Kalman filter is applied to this problem by assuming (an approximation) that the estimate satisfies (25) between observations. A closed form (although implicit) solution of (25) is available. This solution is re-linearized about the most current estimate for the purpose of computing the state transition matrix (for the propagation of the covariance matrix only). Thus the problem is completely discretized, and no numerical integration is required. To summarize, between observations the estimate evolves according to (25) but via a closed form solution, and the covariance matrix according to

** Simulations were also performed using velocity (range rate) data, and simultaneous range and range rate data. Conclusions for these data types are similar, and only range data simulations will be presented here.

$$P_{k+1|k} = \Phi_{k+1,k} P_{k|k} \Phi^T_{k+1,k} \quad , \qquad (27)$$

where $\Phi_{k+1,k}$ is computed as described above.

A number of trajectories of (25) were used in simulations. The one presented here is a rectilinear 'ellipse' starting at approximately 8 er (zero time), going to apogee at 34 er (42 hrs), and terminating at 26 er (70 hrs). A bias in μ of +0.015 was introduced in (25) for the purpose of generating observations, while the filter used the unbiased value. This represents a 100 sigma bias, which is quite large, but the interest here is to observe, over short time, long term effects in real orbits. Range data (26) was generated at the rate of 1 every 0.1 hr, and a 20 hr gap in the data was left between 40 and 60 hrs. The data gap will demonstrate how the limited memory filter performs when faced with prediction errors caused by system nonlinearity.

Figures 1 and 2 show how the Kalman filter behaves in terms of estimation error time histories. The position estimation error, in particular, is seen to grow very rapidly. The velocity estimation error does not grow as much, but is definitely not random. Figures 3 and 4 show the histories of the root sum square/root mean square error (RSS/RMS) ratios* for position and velocity. These are seen to grow almost exponentially in the Kalman filter. Thus, in the presence of the bias, the Kalman error uncertainties quickly become small, since much accurate data has been processed, and then the filter fails to track the orbit.

The linear Limited Memory filter performance is shown in the same figures. This filter is used according to the procedure outlined in the preceding section. That is, old data is discarded in batches of 10, and the filter memory varies between 10 and 20. Reference to Figures 1 and 2 shows that the estimation errors are rather random, with appropriately small sums. The estimation errors remain within approximately 2 sigma measurement noise level, and the filter is tracking the orbit. More accurate determination of the orbit is not possible in view of the μ bias. More accurate determination would require the estimation of the bias itself.

Figures 3 and 4 contain several interesting features. First, it is seen that the filter covariance matrix is consistent with the estimation errors (RSS/RMS$<$3), except at 60 hrs in Figure 4, which will be discussed later. Note that the RSS/RMS ratios always grow and then drop the instant old data is discarded. Apparently the filter begins to diverge; the model is not quite accurate enough over time-arcs of 2 hrs. Note the unusual RSS/RMS ratios at 60 hrs. Apparently the computation of the covariance matrix over the data gap via (27) was inaccurate due to system nonlinearity. Thus at 60 hrs, the covariance matrix is in error. This error is eliminated, together with old data, at 61 hrs (see Remark 4 of Section 3).

5. Conclusions

Linear and nonlinear optimal filters with limited memory length have been developed. These filters have application in problems where the standard filters 'diverge' due to modeling and computational errors. This has been demonstrated by numerical simulations.

6. Acknowledgments

The author gratefully acknowledges the stimulating discussions with Mr. R. K. Squires which led to the formulation of this problem, and his continuing support throughout this work. The author is also indebted to Mrs. A. Bailie for performing the analysis and computer program development associated with the simulations.

References

[1] Kalman, R. E.; "A New Approach to Linear Filtering and Prediction Problems," Trans. ASME, J. Basic Engrg., Vol. 82, pp. 35-44, March 1960.

---

*
$$RSS(n) = \frac{1}{n}\sum_{i=1}^{n} [x_k(i) - \hat{x}_k(i)]^2 \quad ;$$

$$RMS(n) = [(P_{n+})_{kk}]^{1/2} \quad ; \quad k=1,2 \text{ for position}$$

and velocity, respectively.

[2] Kalman, R. E. and Bucy, R. S.; "New Results in Linear Filtering and Prediction Theory," Trans. ASME, J.Basic Engrg., Vol. 83, pp. 95-108, March 1961.

[3] Kushner, H. J.; "On the Differential Equations Satisfied by Conditional Probability Densities of Markov Processes," J. SIAM on Control, Vol. 2, No. 1, pp. 106-119, 1964.

[4] Jazwinski, A. H.; "Filtering for Nonlinear Dynamical Systems," IEEE Trans. on Automatic Control, Vol. AC-11, No. 4, pp. 765-766, October 1966.

[5] Schlee, F. H., Standish, C. J. and Toda, N. F.; "Divergence in the Kalman Filter," AIAA Journal, Vol. 5, No. 6, pp. 1114-1120, June 1967.

[6] Fitzgerald, R. J.; "Error Divergence in Optimal Filtering Problems," Second IFAC Symposium on Automatic Control in Space, Vienna, Austria, September 1967.

[7] Pines, S., Wolf, H., Bailie, A. and Mohan, J.; "Modifications of the Goddard Minimum Variance Program for the Processing of Real Data," Analytical Mechanics Associates, Inc. Tech. Rept. Contract NAS 5-2535, October 1964.

[8] Schmidt, S. F.; "The Application of State Space Methods to Navigation Problems," Philco Western Development Labs., Rept. TR4, October 1963.

[9] Schmidt, S. F.; "Estimation of State with Acceptable Accuracy Constraints," Interim Rept. 67-4, Contract NAS 5-9085, Analytical Mechanics Associates, Inc., January 1967.

[10] Jazwinski, A. H. and Bailie, A. E.; "Adaptive Filtering," Interim Rept. 67-6, Contract NAS 5-9085, Analytical Mechanics Associates, Inc., March 1967.

[11] Jazwinski, A. H.; "Nonlinear Filtering with Discrete Observations," Paper 66-38, AIAA Aerospace Sciences Meeting, New York, January 1966.

[12] Cosaert, R. and Gottzein, E.; "A Decoupled Shifting Memory Filter Method for Radio Tracking of Space Vehicles," Paper presented at the XVIII International Astronautical Congress, Belgrade, Yugoslavia, 1967.

Fig. 1. Position Estimation Errors

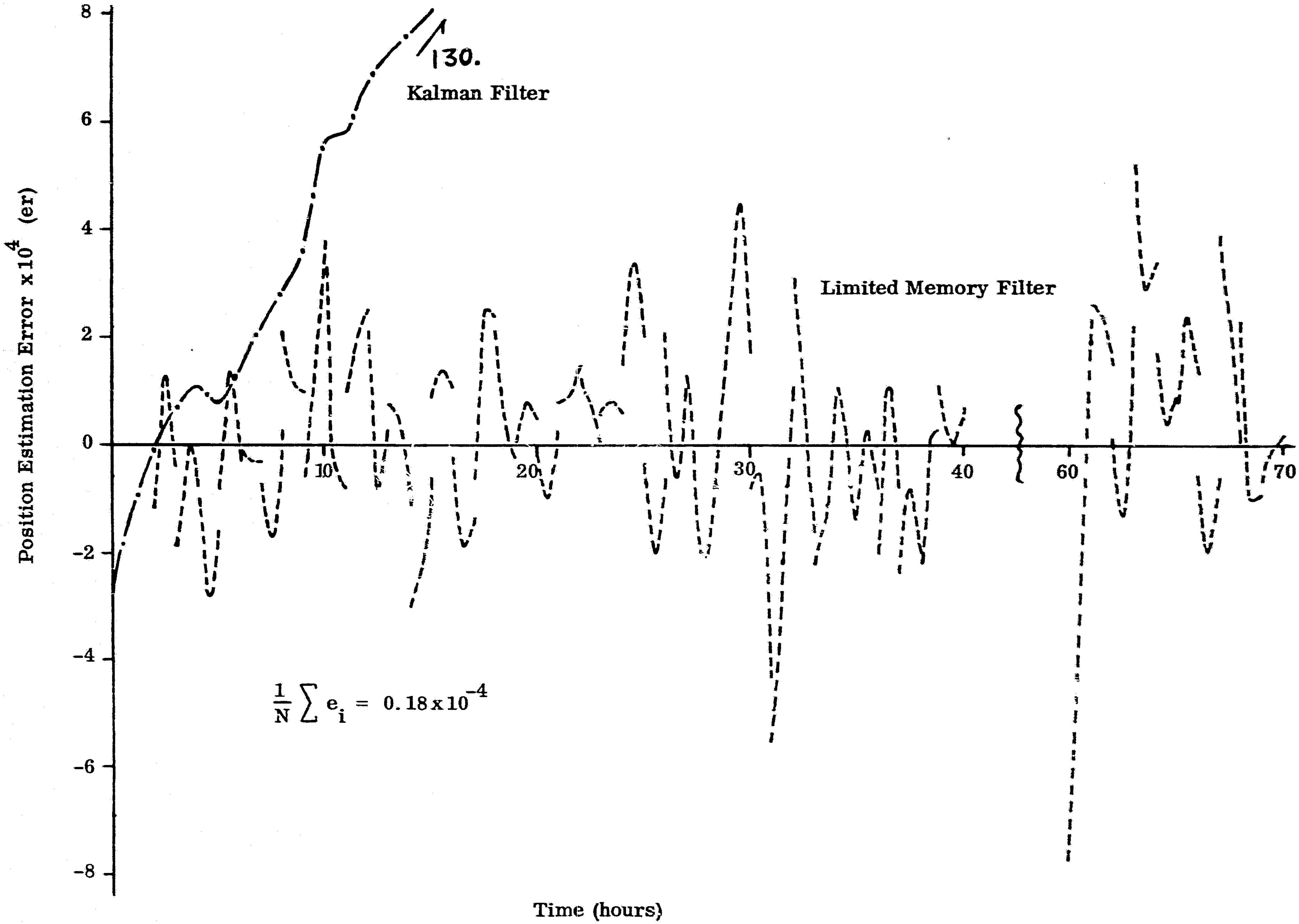

Fig. 4. Velocity RSS/RMS History

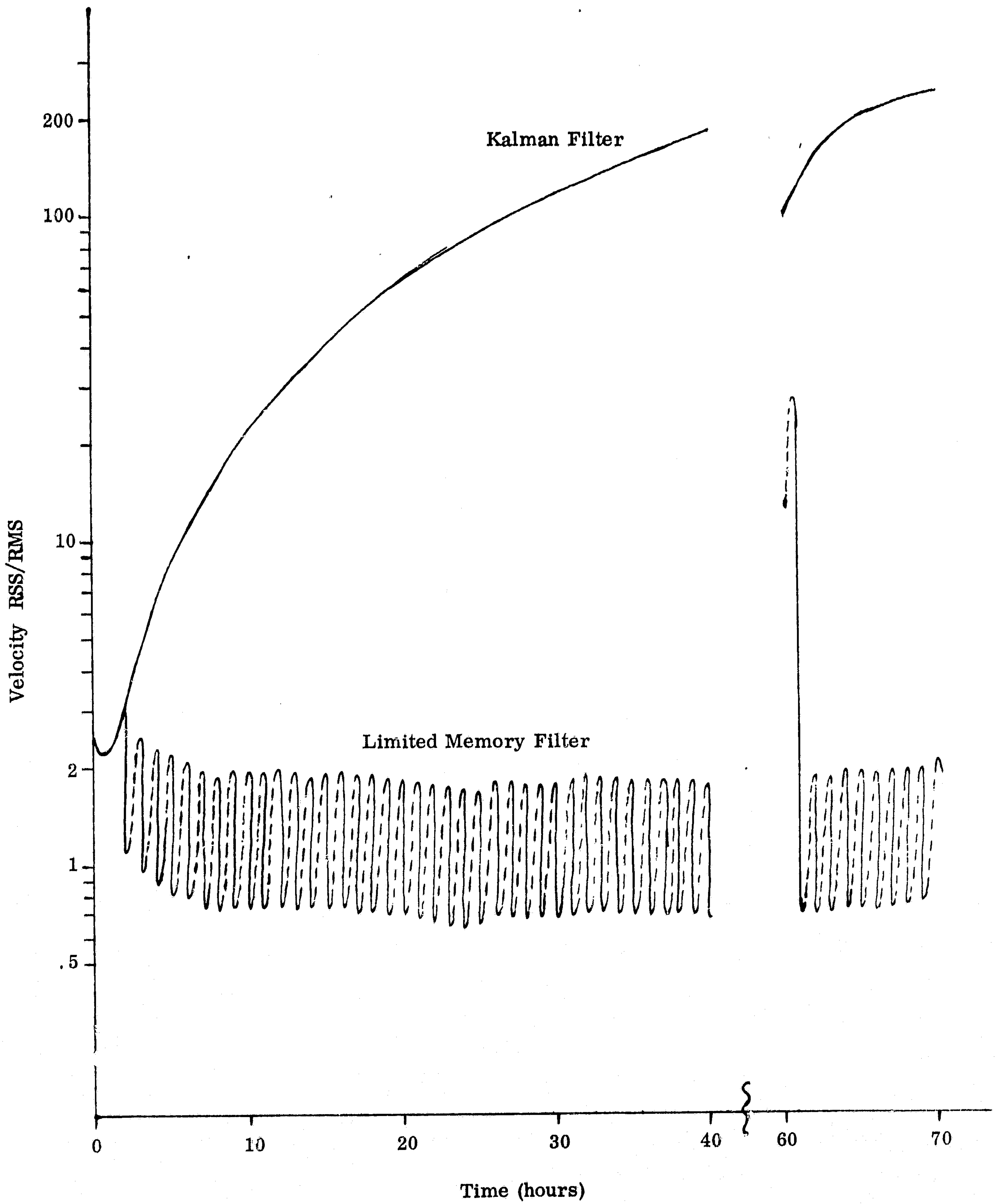

Fig. 3. Position RSS/RMS History

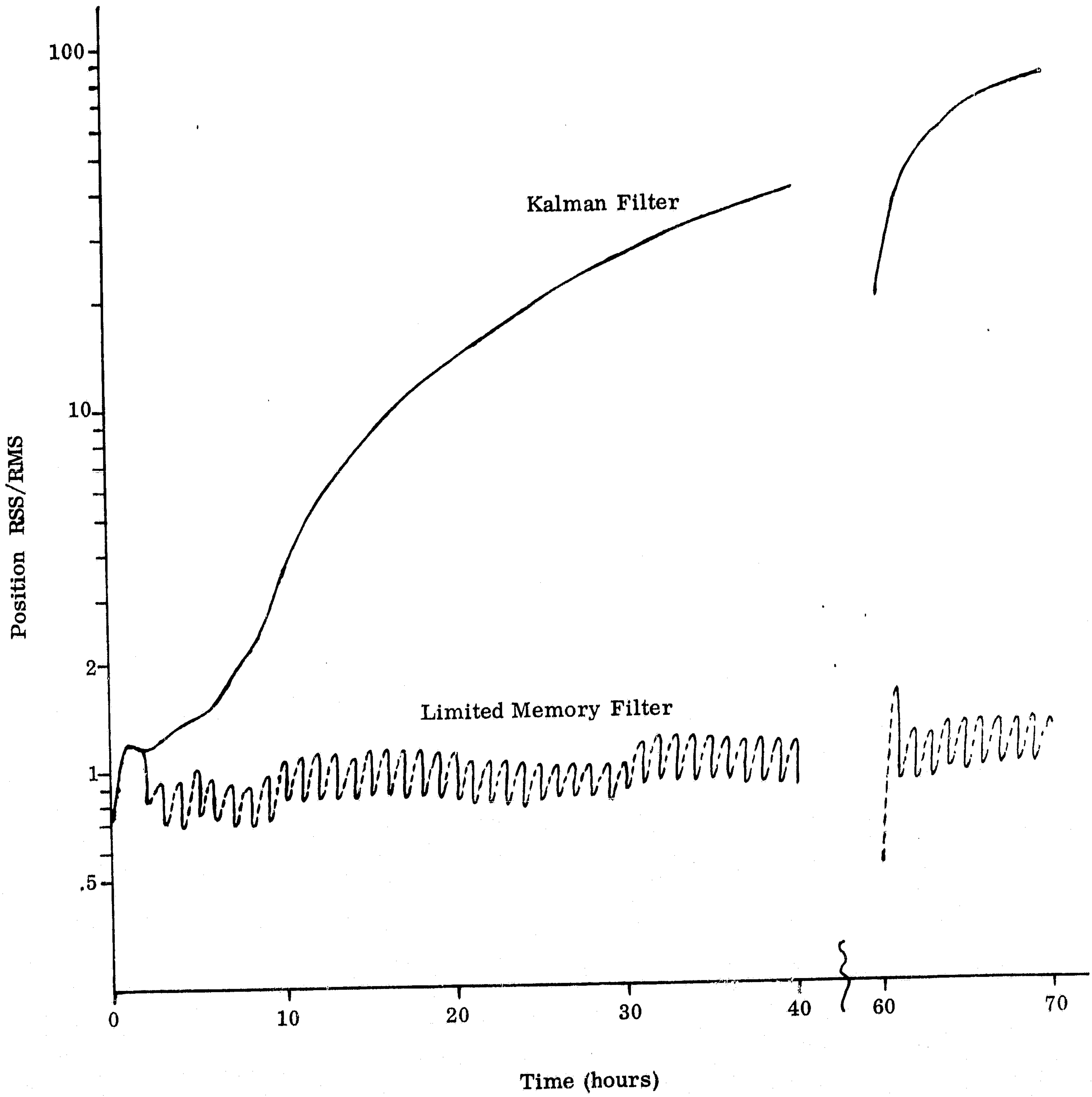

Fig. 2. Velocity Estimation Errors