W. Wolf

Digital Computer Laboratory
Massachusetts Institute of Technology
Cambridge 39, Massachusetts

SUBJECT: **THE APPLICATION OF STATISTICAL FILTER DESIGN TO A PARTICULAR PROBLEM**

To: W.K. Linvill

From: W.I. Wells

Date: July 21, 1953

Abstract: A particular solution of the optimum filter design is presented for the class of problems where the noise is normally distributed and the data is described by a normal distribution. This is an application of the general results obtained in previous memos.[1,2] This particular problem lends itself to an analytic solution and in fact turns out to yield a linear filter. This linear filter has long been used in tracking problems but it is believed that this is the first rigourous derivation. Three examples are given which represent the complete range of characteristics of the filter and one is able from them to make statements about the necessary "memory" of the filter. One is also able to draw conclusions about how often one should sample a function based upon the irregularity of the function and on the noise that contaminates the measurements.

I. INTRODUCTION

As a direct application of the general concepts set forth in previous reports[1,2], a problem is presented which not only illustrates many of the important general points, but fortunately, leads to an exact analytical solution. Under certain circumstances this problem may have important practical applications, but for the present it will be treated as a purely hypothetical case.

We will describe the problem in the statistical sense and then present three particular solutions. As a preparation for this it is necessary to describe the problem in terms that are readily interpreted so as to define the two functions of the filter, detection and selection.

The problem to be treated is one which occurs in systems that are designed to track, in one dimension, on the basis of sampled data. The characteristics of the track and of the noise that contaminate the data are given. From these characteristics we are able to compute on the basis of the received data the probability distribution functions for

---

[1] M-1812, "The Philosophy of Statistical Filter Design," W.I. Wells, Jan. 27, 1953.

[2] M-1886, "The Specification of an Ideal Detector as a First Step in Filter Design," W.I. Wells, March 6, 1953.

the expected position and velocity of the track after the reception of data. This is the detection problem.

The selection problem comes in choosing, from this probability density distribution function, the value of position and velocity that represents the "best" guess of position and velocity under the constraints imposed by the desired overall function of the filter. This action of selection will be discussed first since its action, in this case, is such that we simplify the calculations if it is known.

We suppose that the probability distribution function has been calculated. We now desire to find the spot at which we next expect to find the track when the next data arrives. If we are going to examine the following data in such a way that only data that is extremely close to the predicted position will be used, we will be obliged to look at the most probable value. If, however, we are going to look for data that may be anywhere in the general neighborhood, we may choose another criterion. The mathematical statement of this problem is that we wish to center our area of search on the probability distribution curve in such a way as to cause the product of these two curves to include the most area. Under these conditions we will have the greatest probability of "seeing" the data that comes in next time. In order to do this process in a mathematical way we note that this is just the convolution integral representation. Thus we construct a function which represents our effective area of search and convolve this with the probability distribution function. Then the maximum point on the resulting curve is the correct place to put the center of area of search in order to maximize the probability of "seeing" the new data.

We note that if the probability distribution curve is symmetrical with its highest point as the center and that the curve of search area is symmetrical, then the convolution of these two leads to the fact that the highest point is where the search area curve is centered on the highest, center point, of the distribution curve. If this is the case, one may just forget about this convolution operation and look for the center or maximum of the probability distribution function. This is the case in the problem being reported in this paper. It will be shown that the probability distribution functions come out as normal distributions, and since the search area is assumed symmetrical, we need only ask for the highest point on the normal curve as the best place for us to center our search area.

The problem from here on is to do the job of detection, that is, to calculate the probability distribution curves. It will be understood that the process of selection merely picks out the peak of the distribution curve so we will give formulae that do this as part of the following calculations. Our function that we expect to filter is $X(t)$.

II. THE PROBLEM

First we shall discuss the characteristics of $X(t)$. We choose to say that the second derivative of $X(t)$ is a constant in the intervals between samples. This is an arbitrary representation but as the samples become closer together the representation becomes exact. This means that the curve $X(t)$ is to be approximated with a sequence of parabolas, each connecting two adjacent sample times in such a way that the curve of $X(t)$ is smooth, i.e., the curve is continuous at the sample times, with the second derivative assuming new values for each interval. We will call the first derivative the velocity and the second derivative the acceleration. We have not yet completely described the process. For any set of samples one may fill in several smooth curves with the above characteristics. See Fig. 1.

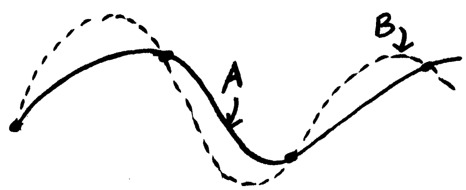


Fig. 1

The remaining restriction must tell which curve will be used for the approximation. We do this by requiring that the probability distribution density of the acceleration be a normal distribution centered at zero. That is, the smaller accelerations are more likely than the larger ones, hence, the approximating curve will be the one with the least violent turns. This means in the above figure, (Fig. 1), the continuous curve, (A), is much more likely than the dashed curve, (B). We write this distribution as

$$W(a) = \frac{1}{\sqrt{2\pi}\,\alpha} e^{-\frac{a^2}{2\alpha^2}} \qquad (1)$$

where (a) is the acceleration, α^2 is the variance of the normal distribution.

Next we will discuss the type of noise to be considered. When we sampled the curve we make a measurement of the value of X at a certain time. If this measurement is exact, there is no noise. We wish to consider the particular case where there is some doubt as to the exact value of X. In fact we assume that the exact value of X is normally distributed about the measurement value of the sample. Let the n'th sample be called S_n. Then

$$W(X) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(X-S_n)^2}{2\sigma^2}} \qquad (2)$$

where σ^2 is the variance of the normal distribution.

III. SOLUTION

This completes the characterization of the problem. Next we shall show how these two distributions may be used to calculate the probability distribution function of predicted X and V. (Detection)

We begin by assuming that we have received the first sample (S_1) at t=1. Now we may ask for the probability distribution of X. This is obviously

$$W_1(X) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(X-S_1)^2}{2\sigma^2}} \tag{3}$$

Now, there is a time elapse before the next sample and two effects become apparent. First, if there is a first derivative, or velocity (V), at t=1, then we can say, if the acceleration (a) is zero, that at t=2 the distribution will be the same as for t=1 except that X will be increased to X+V. We measure (V) and (a) in terms of the sample interval so that $V(t_2-t_1)= V$. In other words, if a=0 then at t=2

$$W_2(X) = e^{\frac{-(X-V-S_1)^2}{2\sigma^2}} \tag{4}$$

where V is the velocity at t=1. We can imagine this as a distribution curve which is a function of both (X) and (V). We plot this on an (X,V) plane where the height above the plane is the height of the probability density distribution. When the first sample S_1 was received, we had a plot as follows:

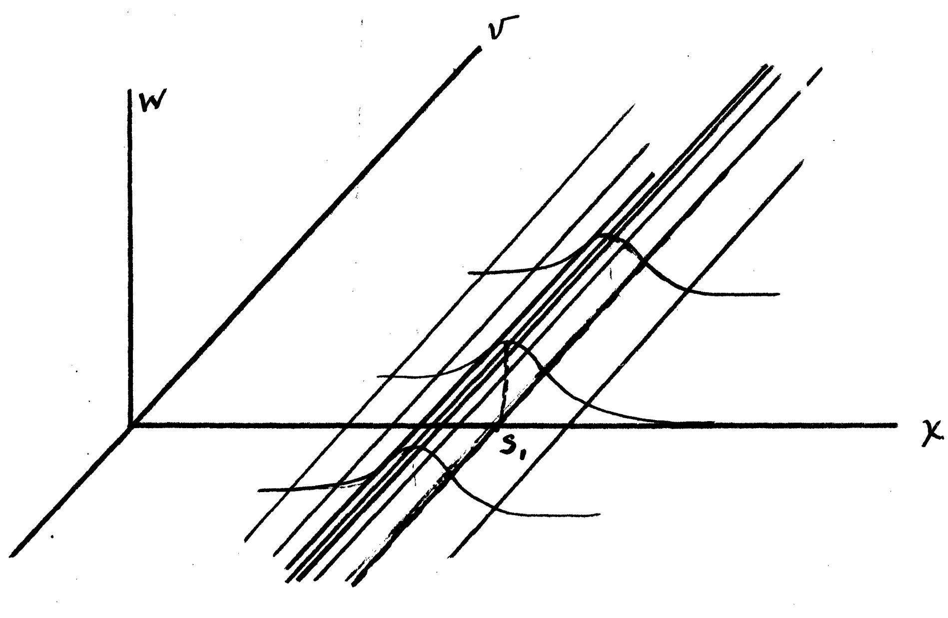


Fig. 2

We see here that the "mound" is uniform in the V coordinate. This means that after the first sample is received, one knows nothing about derivatives and all velocities are equilikely. In addition, no matter where we cut through, the X distribution is normal about S_1.

Now, if a=0, we can draw the distribution at t=2. We note that <u>if</u> the initial velocity is zero, the distribution curve is the same. Thus, on the X axis, where V=0, the "mound" remains fixed. <u>If</u> the velocity is some other value, the mound shifts to the right a distance V. One sees that the whole mound is twisted as follows.

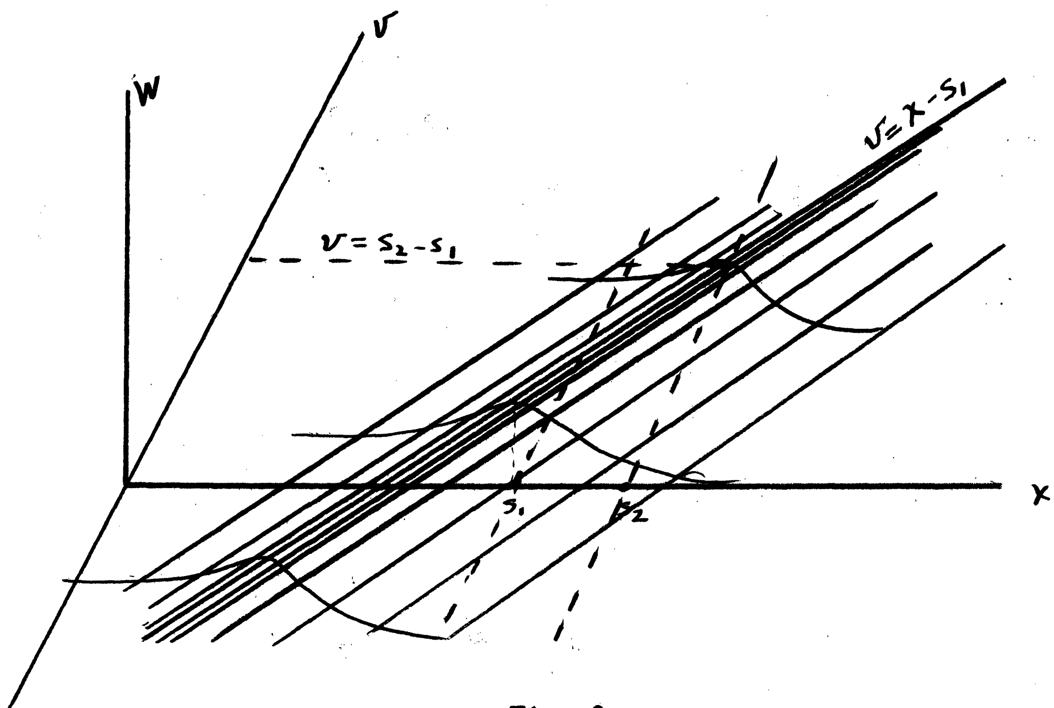


Fig. 3

It is seen to be skewed with respect to the V axis. This means that <u>if</u> the position at t=2 is found to be S_2, the most likely value for V <u>will</u> be found at $V=S_2-S_1$, where this mound has its highest point, which intersects the line $X = S_2$.

Now we examine what happens to this distribution if the acceleration is not zero. Suppose it is exactly (a). Then we know that in the time t_2-t_1, X will increase to

$$X_2 = X_1 + V_1 + \frac{a}{2} \qquad (5)$$

and V to

$$V_2 = V_1 + a \qquad (6)$$

since (a) is constant in the interval. We ask what this has done to the "mound" that we started with in Fig. 2. Obviously if a=0, we get Fig. 3. If (a) is some number, the mound retains its same cross-sectional shape but is now centered at $X=S_1+a/2$ on the X axis and is parallel to the line

$$2V = X-S_1 \tag{7}$$

rather than

$$V = X-S_1 \tag{8}$$

as in Fig. 3.

The formal procedure for accomplishing transition from Fig. 3 to the desired curve is the convolution. When the acceleration is fixed at a known value (a_o) its distribution function is an impulse. If this is thought of as a function of V, we have $U_o(V-a_o)$. $U_o(X)$ is the unit impulse function. It is zero everywhere except at X=o, and its integral over X is one. Convoling this with a distribution containing V, merely replaces V by V-a, which is the desired transformation. Likewise, if the impulse is written $U_o(2X-a)$, and convolved with a distribution containing X, we replace X by X-a/2. So we have for the case above

$$W_2(X_1V) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_1(\sigma-V,\tau)U_o(\sigma-X-\tfrac{a}{2})U_o(\sigma-V-a)\,d\sigma\,d\tau. \tag{9}$$

Both steps have been combined here. We note $W_1(\sigma-V, \tau)$ has had X replaced by X-V to account for the initial velocity at t=1. (This is the transition from Fig. 2 to Fig. 3.) The convolution shifts the curves further to allow for the acceleration during the interval.

It is helpful to think of this process in terms of filters. We note that as W(a) is an impulse, we pass the function $W_1(X-V,V)$ through a "two-dimensional" filter whose impulse response is an impulse, delayed (a_o) units in the V direction, $\frac{a_o}{2}$ units in the X direction. Since the operations of integration and differentiation involved are linear, it is proper that when W(a) is different from an impulse, the impulse response of the filter will also be different. In fact it will be W(a), also shifted (a) units in the V direction and a/2 units in the X direction. The needed filter has the impulse response related to W(a) as follows

$$W_f(a) \longrightarrow W_f(V)W_f(\tfrac{X}{2}) \tag{10}$$

We note that in the X,V plane these convolutions all take place along a straight line V-2X = const. For convenience we change variables from

X,V to W,Z as follows .

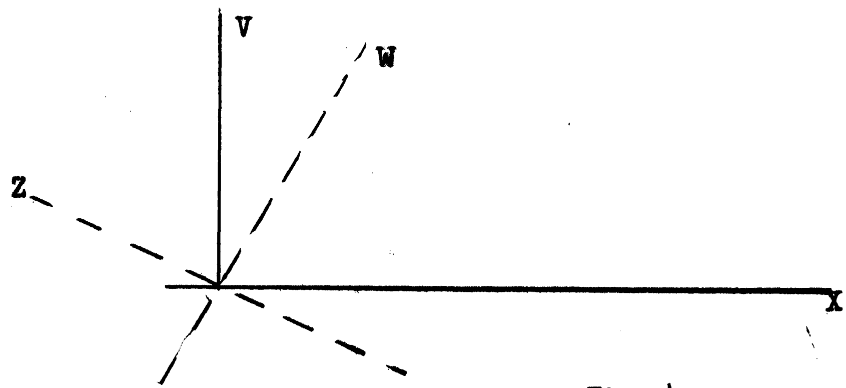


Fig. 4

where Z=const. is the equation of V-2X=const. In terms of W,Z we should expect that the impulse response of our filter contain only W. The equations for the transformed coordinates are:

$$\begin{aligned} 5\ Z &= V-2X \\ 5\ W &= 2V+X \\ 5\ V &= Z+2W \\ 5\ X &= -2Z+W \end{aligned} \qquad (11)$$

We note that the filter corresponding to a fixed known acceleration

$U_0(X-\frac{a_o}{2})U_o(V-a_o)$ becomes

$$U_o\left(\frac{-4Z+2W}{\sqrt{5}} - a_o\right) U_o\left(\frac{Z+2W}{\sqrt{5}} - a_o\right) \qquad (12)$$

This has a value only where both brackets are zero, thus:

$$-\frac{4Z+2W}{\sqrt{5}} - a_o = 0 = \frac{Z+2W}{\sqrt{5}} - a_o \qquad (13)$$

which is satisfied only for Z = 0 and the impulse response becomes

$$U_o\left(\frac{2W}{\sqrt{5}} - a_o\right)$$

or in general

$$W(a) \longrightarrow W_1\left(\frac{2W}{\sqrt{5}}\right) \qquad (14)$$

Now we return to the problem laid out. The distribution of acceleration was to be normal.

$$W(a) = \frac{1}{\sqrt{2\pi}\alpha} e^{-\frac{a^2}{2\alpha^2}} \qquad (15)$$

Dropping the constant multiplier, we have from Eq. 14,

$$^{*}W(W) \doteq e^{-\frac{4/5W^2}{2\alpha^2}}$$

$$\text{Let} \quad \delta = \frac{4}{10\alpha^2} \qquad (16)$$

$$W(W) \doteq e^{-\delta W^2} \qquad (17)$$

The normal distribution of time position about observed position may be written

$$W(W_1 Z) \doteq e^{-\beta(W-2Z-b_1)^2} \qquad (18)$$

$$\text{where} \quad \beta = \frac{1}{10\sigma^2} \qquad b_1 = \sqrt{5}\; S_1 \qquad (19)$$

Now with the notations we may begin the process of constructing the probability density curves for $W(W,Z)$ after receiving several pieces of data.

We receive the first sample S_1 and at $t=1$ we have

$$W(W,Z,t_1) \doteq e^{-\beta(W-2Z-b_1)^2} \qquad (20)$$

* $\doteq$ stands for "is proportional to."

We replace X by X-V to allow for initial velocities. This is done in W,Z by replacing

$$\left.\begin{array}{l} W \text{ by } 1/5(3W-Z) \\ Z \text{ by } 1/5(7Z+4W) \end{array}\right\} \qquad (21)$$

giving:

$$e^{-\beta(W+3Z+b_1)^2} \qquad (22)$$

We now convolve this with $e^{-\delta W^2}$ which gives:

$$W(W,Z,t_2) \doteq e^{-\frac{\beta\delta}{\beta+\delta}(W+3Z+b_1)^2} \qquad (23)$$

This is the distribution just before the reception of the second sample. When the second sample is received we have:

$$W(W,Z,t_2) \doteq e^{-\frac{\beta\delta}{\beta+\delta}(W+3Z+b_1)^2} e^{-\beta(W-2Z-b_2)^2} \qquad (24)$$

Again we change variables to allow for initial velocities and convolve with $e^{-\delta W^2}$, getting

$$W(W,Z,t_3) \doteq e^{-\left[\frac{1}{\frac{9\beta\delta}{\beta+\delta}+\beta+\delta}\left\{\frac{\beta\delta^2}{\beta+\delta}(3W+4Z+b_1)^2 + \beta\delta(W+3Z+b_2)^2 + \frac{\beta^2\delta}{\beta+\delta}(5Z+3b_2-b_1)^2\right\}\right]} \qquad (25)$$

This is the distribution just before the reception of the third sample. The values of(W and Z)(X,V) which maximize this expression are the best guess at the predicted position and velocity at the third sample.

From the length of this expression one notes that the algebra is becoming very long. Thus we shall attempt to find a recursion formula to be used in numerical examples. We note that by further manipulations each of the previous expressions can be put in the form:

$$e^{-A_1(W+BZ+C_1)^2 - A_2(Z+C_2)^2} \qquad (26)$$

Since the expressions retain form we merely choose an appropriate beginning and calculate the A,B,C of one time in terms of those at one preceding time.

A convenient spot to begin is where we have just received a sample and then changed variables to allow for initial velocities. Then we have

$$W(W,Z,t_{n-1}) = e^{-A_{1n-1}(W+B_{n-1}+C_{1n-1})^2 - A_{2n-1}(Z+C_{2n-1})^2} \tag{27}$$

where this gives the distribution of (V) and (X+V) after the n-1 sample. Now we convolve this with $e^{-\delta W^2}$ to get the distribution of the predicted X and V. Then we receive the n'th sample and change variables to allow for initial velocities. Putting this back in the above form, we have

$$A_{1n} = XY^2 + A_{2n-1}\frac{16}{25} + \beta \tag{28}$$

$$A_{2n} = \frac{X\left[A_{2n-1} + \beta U^2\right] + \beta A_{2n-1}}{A_{1n}} \tag{29}$$

$$B_n = \frac{XYQ + A_{2n-1}\frac{28}{25} + 3\beta}{A_{1n}} \tag{30}$$

$$C_{1n} = \frac{XYC_{1n-1} + 4/5A_{2n-1}C_2 + \beta b_n}{A_{1n}} \tag{31}$$

$$C_{2n} = \frac{\left\{ XA_{2n-1}\left[YC_{2n-1} - 4/5C_{1n-1}\right] + X\beta U\left[Yb_n - C_{1n-1}\right] + A_{2n-1}\beta\left[4/5b_n - C_{2n-1}\right] \right\}}{A_{1n}A_{2n}} \tag{32}$$

where

$$
\left.
\begin{aligned}
X &= \frac{A_{1n-1}\,\delta}{A_{1n-1}+\delta} & Y &= \frac{4B_{n-1}+3}{5} \\
Q &= \frac{7B_{n-1}-1}{5} & U &= B_{n-1}+2 \\
b_n &= \sqrt{5}\; S_n & S_n &= \text{n'th sample.}
\end{aligned}
\right\} \qquad (33)
$$

These are the desired recursion relationships. We may now ask for the most probable (W,Z) or (X,V). These are the coordinates of the peak of the "mound" that we have. These are found to be.

$$V_n = (2B_{n-1})\,\frac{C_{2n}}{\sqrt{5}} - \frac{2C_{1n}}{\sqrt{5}} \qquad (34)$$

$$\bar{X}_{n+1} = (B_n+2)\,\frac{C_{2n}}{\sqrt{5}} - C_{1n} \qquad (35)$$

$\bar{X}_{n+1} = X_n + V$ is the predicted position for the following (n+1) sample.

$$X_n = (3-B_{1n})\,\frac{C_{2n}}{\sqrt{5}} + \frac{C_{1n}}{\sqrt{5}} \qquad (36)$$

Somewhat more useful formulae are obtained when one notes that these equations may be written as follows:

$$V_n = V_{n-1} + M(S_n - \bar{X}_n) \qquad (37)$$

$$X_n = \bar{X}_n + N(S_n - \bar{X}_n) \qquad (38)$$

where:

$$M = \frac{(2B_{1n}-1)\,(XYU\beta + 4/5A_{2n-1}\beta)}{A_{1n}\,A_{2n}} - \frac{2\beta}{A_{1n}} \qquad (39)$$

$$N = \frac{(3-B_{1n})\,(XYU\beta + 4/5A_{2n-1}\beta)}{A_{1n}\,A_{2n}} + \frac{\beta}{A_{1n}} \tag{40}$$

The size of the distributions are of interest, especially in the X direction, since the samples are received in X. To get the unconditional distribution of X, we merely integrate $W(X,V)$ over all V. This leaves a normal distribution in X of the form:

$$e^{-\frac{(X-m)^2}{2\gamma^2}} \tag{41}$$

If we ask for the distribution of X_n, we have

$$\gamma_n^2 = \frac{A_{1n}(3-B_{1n})^2 + A_{2n}}{10\,A_{1n}\,A_{2n}} \tag{42}$$

and for X_{n+1} we have

$$\gamma_{n+1}^2 = \frac{\frac{A_{1n}\,\delta}{A_{1n}+\delta}\,(2+B_{1n})^2 + A_{2n}}{10\,\frac{A_{1n}\,\delta}{A_{1n}+\delta}\,A_{2n}} \tag{43}$$

These are all of the relations that are needed to solve any particular problem. The particular problems are described by giving α and σ.

IV. DISCUSSION

First, let us discuss the outstanding characteristics of the solution and then present some examples. To actually construct the optimum filter we need know Eqs. 37 and 38, and the appropriate values of M and N for the particular data. This filter has an important characteristic which is not usually so evident. M and N are functions of the number of pieces of data received. In other words, when one first begins to filter data, the filter uses one set of values of M and N, then as more and more data are received, these values of M and N change, eventually reaching a fixed value asymptotically. In many cases the "transient" behavior is over very quickly and one looses very little by using the fixed filter straight through. On the other hand, there are cases where the final form of the filter is useless unless the "transient" is used first. The most striking example of this behavior is that where the acceleration is zero at all times. In this case $X(t)$ is a straight line. The final values of M and N are zero. This occurs only after an infinite number of samples have been received. Thus, during the "transient" the filter determines the straight line and then after a long time no more data is accepted. In this case the "transient" behavior is the important part. One, therefore, sees that we cannot ask for the final values of M and N and automatically expect to have a good filter.

In many cases, however, the transient is relatively short and new data is weighted rather heavily and thus only small errors are encountered by using a fixed filter with the final values of M and N. Whether this procedure is possible depends upon the relative values of α and σ. In any case, the procedure for calculating the optimum filter is given above.

The next thing to be considered is the "goodness" of the filter. Even though the filter may be "optimum" it may not do enough good to warrant its use. Therefore we ask, how much sharpening of the distribution curves do we get by using this filter with several pieces of data. Since W(X) is a normal curve, its variance γ^2 gives its sharpness. Eqs. 42,43 give γ^2 for the whole family of filters after any number of samples. As the number of samples increase, γ^2 decreases asymptotically to a fixed value which represents the "best" that the filter can do with this particular type of data. For a good filter γ^2 should be considerably smaller than σ^2.

We will now give three simple examples. The first leads to the straight line smoothing equations.

V. EXAMPLE 1.

Let $\alpha = 0$. This says a = 0, thus X(t) is a straight line. We receive our first sample and have

$$W(X,V,t_1) = e^{-\frac{(X-S_1)^2}{2\sigma^2}}$$

which after a shift due to velocity and change of variable becomes:

$$e^{-\beta(W+3Z+b_1)^2}$$

From Eq. 27. $A_{11} = \beta$ $A_{21} = 0$ $B_1 = 3$ $C_{11} = b_1$ $C_{21} = 0$

From Eq. 16 $\delta = \frac{4}{10\;{}^{2}} = \infty$

From Eqs. 33 $X = \beta$ $Y = 3$ $Q = 4$ $U = 5$

Eq. 28 $A_{12} = 10\beta$

Eq. 29 $A_{22} = 5/2\beta$

Eq. 30 $B_2 = 3/2$

Eq. 39,40 $M_2 = 1$ $N_2 = 1$

Eq. 42 $\gamma_2^2 = \sigma^2$ $\gamma_{\overline{3}}^2 = 5\sigma^3$

Now repeating these steps one finds in general

$$M_n = \frac{\sigma}{n(n+1)} \qquad N_n = \left(\frac{4n-2}{n(n+2)}\right)$$

$$\gamma_n^2 = \frac{4n-2}{n(n+1)}\sigma^2$$

$$\gamma_{\overline{n+1}}^2 = \frac{2(2n+1)}{n(n-1)}\sigma^2$$

where n is the number of samples.

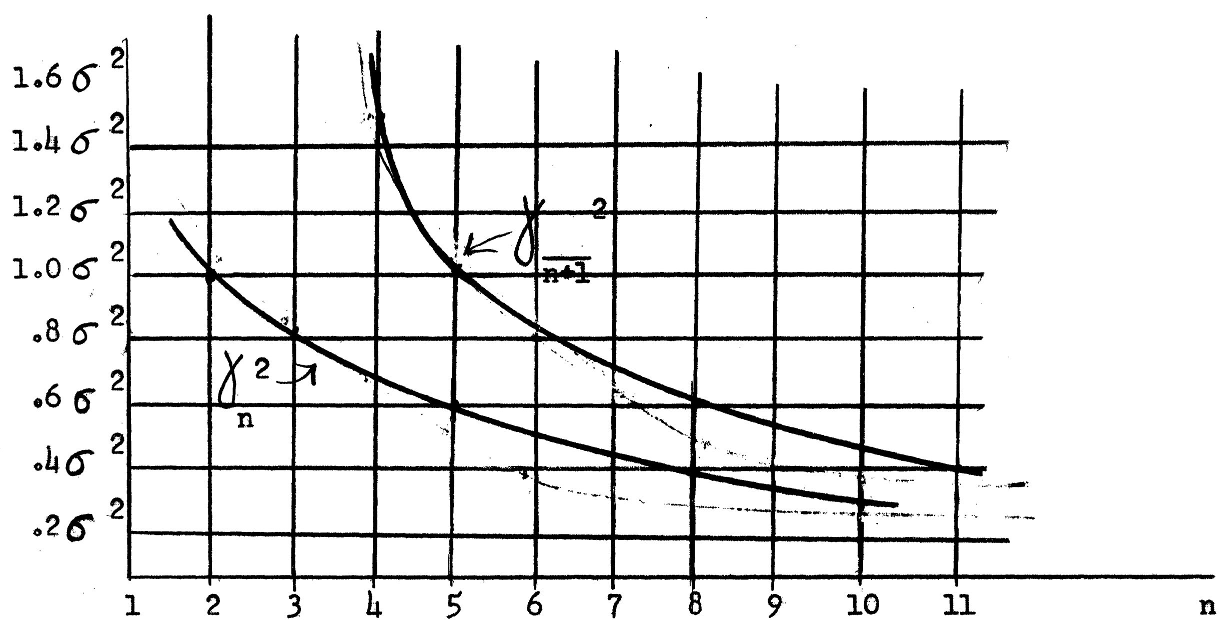


Fig. 5

VI. EXAMPLE 2.

Let $\sigma = 0$, which says there is no noise, hence we should expect very little filtering.

We being again by receiving the first sample, converting to (W,Z) and allowing for initial velocities; this gives as before:

$$e^{-\beta(W+3Z+b_1)^2}$$

$A_{11} = \beta = \infty \quad A_{21} = 0 \quad B_1 = 3 \quad C_{11} = b_1 \quad C_{21} = 0$

$\beta = \frac{1}{10\sigma^2} = \infty$

$X = \delta \quad Y = 3 \quad Q = 4 \quad U = 5$

$A_{12} = \beta = \infty \quad A_{22} = 25\delta \quad B_2 = 3$

$M_2 = 1 \qquad N_2 = 1$

$\gamma_2^2 = 0 \qquad \gamma_{\overline{3}}^2 = 1/2\,\alpha^2$

Repeating.

$X = \delta \quad Y = 3 \quad Q = 4 \quad U = 5$

$A_{13} = \beta = \infty \quad A_{23} = 50\delta \quad B_3 = 3$

$M_3 = 3/2 \qquad N_2 = 1$

$\gamma_3^2 = 0 \qquad \gamma_{\overline{4}}^2 = 3/8\,\alpha^2$

Again:

$A_{14} = \infty \quad A_{24} = 75\delta \quad B_4 = 3$

$M_4 = 5/3 \qquad N_2 = 1$

$\gamma_4^2 = 0 \qquad \gamma_{\overline{4}}^2 = 1/3\,\alpha^2$

or in general:

$$M_n = \frac{2n-3}{n-1} \qquad N_n = 1$$

$$\gamma_n^2 = 0 \qquad \gamma_{n+1}^2 = \frac{n}{4(n-1)}\alpha^2$$

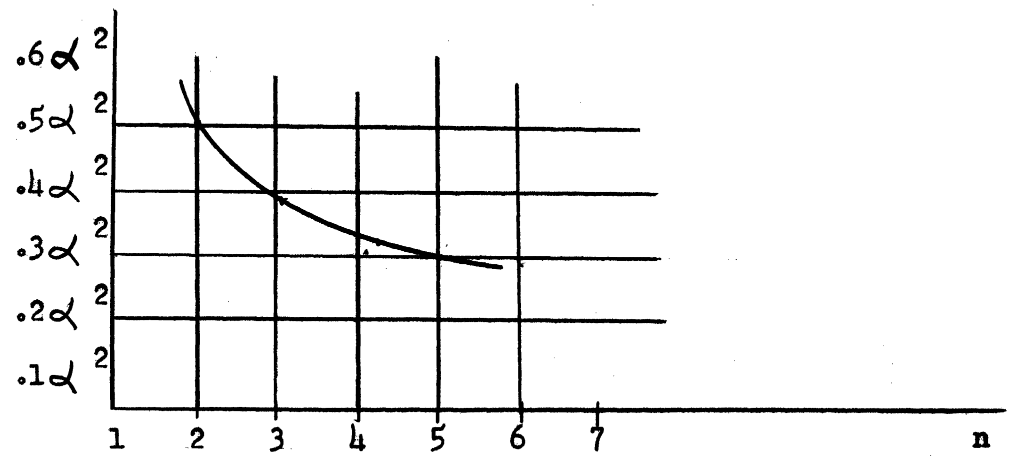


Fig. 6

We see that the filter reaches a nearly fixed value after about 6 pieces of data. In this case we conclude that the "transient" is fairly short and reasonably good results will be obtained by using the filter M=2, N=1, from the beginning. Since there is no noise in the data, the smoothed positions are obviously not in error. The predicted positions are such that there is an improvement of $\gamma^2 = .5\alpha^2$ to $\gamma^2 = .25\alpha^2$, for the final filter.

Both of the above examples could have been worked by simpler means. The third example is a more realistic problem and also the results do not come out so nicely; however, by continuation of the above processes we can work this example for as many steps as we desire.

VII. EXAMPLE 3.

Let $\beta = \delta = 1$, thus $\sigma^2 = 1/10 \quad \alpha^2 = 4/10$.

This is the combined case of noise plus non-straight line signals. Again we find:

$$A_{11} = \beta = 1 \quad A_{21} = 0 \quad B_1 = 3 \quad C_{11} = b_1 \quad C_{21} = 0$$

$$X = \frac{1}{2} \quad Y = 3 \quad Q = 4 \quad U = 5$$

$$A_{12} = 11/2 \quad A_{22} = \frac{25}{11} \quad B_2 = \frac{18}{11}$$

$$M_2 = 1 \quad N_2 = 1$$

$$\gamma_2^2 = 1/10 = \sigma^2 \quad \gamma_3^2 = 7/10 = 7\sigma^2$$

repeating $X = 11/13 \quad Y = \frac{21}{11} \quad Q = \frac{23}{11} \quad U = \frac{40}{11}$

$$A_{13} = 72/13 \quad A_{23} = \frac{25}{9} \quad B_3 = 29/18$$

$$M_3 = 3/4 \quad N_2 = 7/8$$

$$\gamma_3^2 = 7/80 = 7/8\,\sigma^2 \quad \gamma_4^2 = \frac{47}{80} = \frac{47}{8}\sigma^2$$

again:

$$X = \frac{72}{85} \quad Y = 17/9 \quad Q = \frac{37}{18} \quad U = 65/18$$

$$A_{14} = \frac{29}{5} \quad A_{24} = \frac{1375}{493} \quad B_4 = 47/29$$

$$M_4 = 42/55 \quad N_4 = 47/55$$

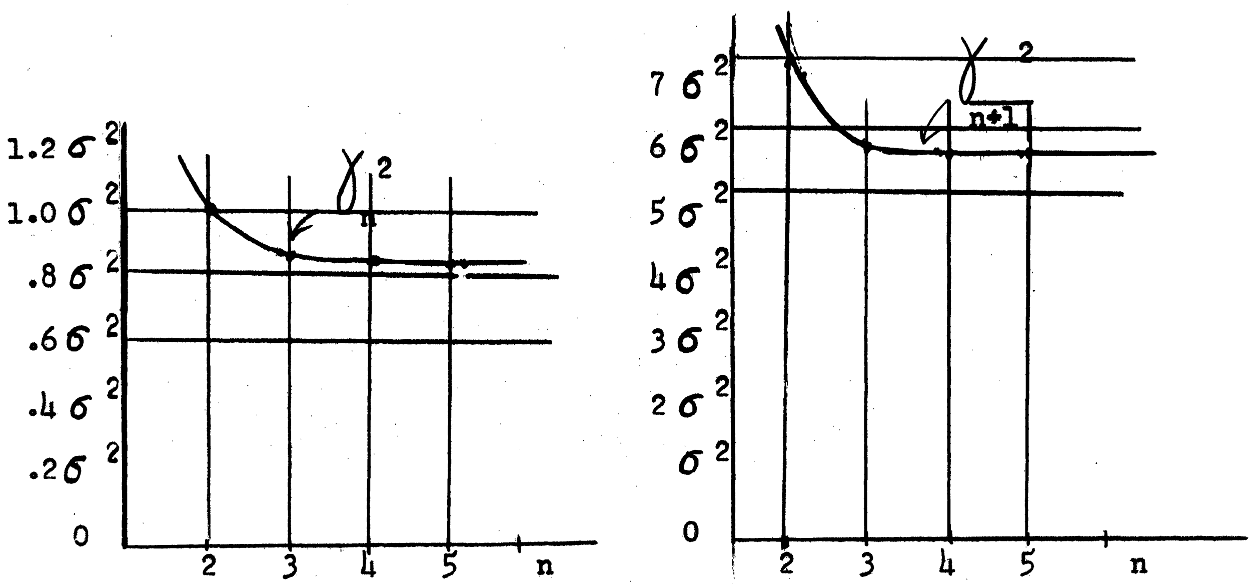
$$\gamma_4^2 = 47/550 = 47/55\,\sigma^2 \quad \gamma_5^2 = \frac{323}{550} = \frac{323}{55}\sigma^2$$

Once more:

$$A_{15} = \frac{199}{34} \qquad A_{25} = \frac{274{,}050}{98{,}107} \qquad B_5 = \frac{322}{199}$$

$$M_5 = \frac{289}{378} \qquad N_5 = \frac{323}{378}$$

$$\gamma_5^{\,2} = \frac{323}{378}\,\sigma^2 \qquad \gamma_5^{\,2} = \frac{1189}{203}\,\sigma^2$$



n is the number of samples.

Fig. 7

We see for this filter, the transient is over within about 4 samples. Also, we see no great improvement in the smoothed data, i.e., from $\gamma_n^2 = 1$ to .85. The predicted position has a distribution roughly $\sqrt{5.8}$ or 2.4 times as wide as the received data. This is to be expected since with $\alpha^2 = 4/10$, there is a fairly large possibility of variation from a straight line.

These three examples illustrate the two extreme problems and one intermediate one. Any other problem of this type can be worked out using the same recursion formulae.

It should be repeated that for every combination of α, σ one can determine M and N (actually these depend only on α / σ) which completely define the linear filter; then one also may calculate the γ^2's which describe the quality of the filter. There is one further criterion which could be easily calculated. This is the variance of the unconditional probability distribution of velocity. It was not included since it is of secondary importance, when the data involves X only.

VIII. CONCLUSION

The complete solution of the design of the desired filter is now given. This involves the two processes, detection and selection. In this particular case the optimum filter turned out to be linear. It is conceivable that for the same problem, but for a different search area criterion, one would find a filter that is not linear. The detection or calculation of the distribution curves would be the same, only the selection problem would differ.

For other problems, in which the distribution curves of the acceleration and noise are not normal, one would in theory follow the same process to calculate the distribution functions of predicted position and velocity. Then an appropriate criterion would be used in the selection process. Since the calculation of the distribution curves is not always possible by analytic means for the general distribution, it is hoped that these problems may be solved by composing the actual distributions from sums of appropriately chosen normal functions. Then the distribution curves of the predicted quantities will be combinations of sums of the type of solution presented here.

There are certain extensions that permit further constraints on certain of the variables, such as requiring the acceleration distribution to be a function of the present velocity, etc. These concepts are very clear; the only question is whether solutions of these more general forms can be done analytically. It does not seem economically feasible to do these calculations with some special analog device, although it is perfectly possible conceptually.

Signed W. I. Wells
W.I. Wells

Approved William K. Linvill
W.K. Linvill W.I.W.

WIW/mrs