Fixed-Lag Smoothing Results for Linear Dynamical Systems*

J. B. MOORE,
Department of Electrical Engineering
University of Newcastle, N.S.W. 2308

Recent results in the smoothing of signals from linear systems with gaussian input and measurement noise have provided new perspectives on receiver design. The particular new results and perspectives discussed in this paper are those concerning the important class of "on line" filters known as filters with delay, lag-filters, or fixed-lag smoothers.

Fixed-lag smoothers are now sufficiently developed to be strong competitors of the widely used Wiener and Kalman filters for some applications, and the paper discusses the various trade-offs between filter performance, permissible delay, filter complexity, and design difficulty.

1. Introduction

In control and communication systems, noisy signals are frequently processed to achieve estimates of signals, states, or parameters. Usually an effort is made to construct estimators which achieve an optimal estimate for some convenient criterion of optimality such as the minimum mean square error criterion. Of the three basic types of estimators, namely filters, predictors, and smoothers, we focus attention in this paper on those estimators which have the best inherent performance characteristics, namely smoothers. The performance improvement possible from smoothers relative to filters is achieved at the expense of estimator complexity and at the expense of true on-line estimation. Of the three basic types of smoothers, namely fixed-point smoothers, fixed-lag smoothers, and fixed-interval smoothers, the fixed-lag smoother is the most useful since it yields an estimate which, apart from a small fixed delay in processing the incoming signal data, is "on-line". In this paper, we study "on-line" estimation via fixed-lag smoothing and in particular the trade-offs between smoother complexity, delay in estimation, and improvement in performance.

To be more precise in our various definitions, consider that for discrete-time systems, any estimate of $x(k)$, $k = 0, 1, \ldots$ based on a sequence of noisy measurements $\{z(0), \ldots, z(j)\}$ is denoted by $\hat{x}(k/j)$ and similarly for continuous-time systems, any estimate of $x(t)$ derived from the set of measurements $\{z(\tau), t_0 \leq \tau \leq s\}$ is denoted by $\hat{x}(t/s)$. The relationship between k and j (or t and s) determines the type of estimator. Thus, we have for

filtering $j = k$ (or $t = s$) and j (or t) is variable;
smoothing $j > k$ (or $s > t$);
fixed-point smoothing
k (or t) is fixed and $j > k$ (or $s > t$) is variable;
fixed-lag smoothing
$(j - k)$ [or $(s - t)$] is fixed and k [or t] is variable;
fixed-interval smoothing
j (or s) is fixed and $k < j$ (or $t < s$) is variable;
prediction $j < k$ (or $s < t$).

* Work supported by the Australian Research Grants Committee and Radio Research Board.
This is an invited paper.

1.1 History

The history of the development of fixed-lag smoothing results is very much bound up with the history of filtering theory and the broader field of estimation theory. In 1601, Kepler, using rather awkward methods, sought to estimate the orbit of Mars from twelve observations, but it was not until 1795 that Gauss with his more elegant least squares methods was able to predict the planetary orbits with reasonable accuracy. In the late 1930's and early 1940's, Kolmogorov and Wiener working with stationery random processes developed the least squares approach for the discrete-time and continuous-time prediction problems respectively. It appears that Kolmogorov's work was motivated from purely mathematical consideration while Wiener was interested in applications to a wide variety of problems including weather forecasting, economics and communication systems. This latter interest led him to study filtering and fixed-lag smoothing problems (Ref. 1). Wiener's work was all in the frequency domain with signals characterized by their power spectral densities and filters derived in terms of their transfer functions. Further developments (Refs. 2, 3) made their appearance when the signal was modelled by a dynamic system driven from a white noise source. This enabled nonstationary processes to be modelled conveniently, and subsequently such estimation problems were solved in the time domain. In this respect, the work of Kalman (Refs. 4, 5) and Kalman and Bucy (Ref. 6) was the most significant. With the assumption of finite-dimensionality and the use of state space ideas, these workers developed highly efficient computational filtering and prediction algorithms which have found numerous applications, the most spectacular of which are in the aerospace field.

The application of the state-space ideas to smoothing problems followed soon after (Refs. 7, 12), but inherent in the early realizations of the fixed-lag smoothing equations are instability problems which have only recently been recognised. Stable realizations of the fixed-lag smoothing equations have since been studied in (Refs. 13-16). A more complete survey of the literature on smoothing is given in Ref. 19.

In summary, we have inherited some well developed fixed-lag smoothing results using frequency domain ideas and following in the tradition of Wiener. We also have a number of more general and more recent fixed-lag smoothing results using state space ideas and following in the tradition of Kalman. The former set of results are restricted in application to stationary processes while the latter set of results are more generally applicable to nonstationary processes but require a finite dimensionality assumption. Both sets of results require the assumption of gaussian signals and gaussian noise. There are also nonlinear estimation results available but these are usually restricted in application to limited classes of nonlinear problems and the resultant estimators are usually considerably more complex and less well understood than those for linear-gaussian problems. Further mention of these will not be included here.

In the next section we review some of the most significant fixed-lag smoothing results in the Wiener tradition and in the following sections introduce the more recent results in the Kalman tradition. Some overall perspectives are discussed in the final section.

2. Fixed-lag Smoothing via Frequency Domain Techniques

The application of Wiener's fixed-lag smoothing results and some of the ideas involved is perhaps best illustrated by reference to a simple example which we now study in some detail.

Consider the scalar valued stationary measurement process

$$z(t) = y(t) + v(t) \quad (1)$$

where the power spectral densities of the signal $y(t)$ and additive noise $v(t)$, denoted $\Phi_{yy}(\omega)$ and $\Phi_{vv}(\omega)$ respectively, are specified. For our example, we will consider the specifications

$$\Phi_{yy}(\omega) = \frac{a^2}{\omega^2 + b^2}, \quad \Phi_{vv}(\omega) = c^2 \quad (2)$$

where a^2, b^2 and c^2 are positive real numbers. Suppose also that the random processes $v(t)$ and $y(t)$ are independent and thus $\Phi_{yv}(\omega) = \Phi_{vy}(\omega) = 0$. We have immediately that

$$\Phi_{zz}(\omega) = \Phi_{yy}(\omega) + \Phi_{vv}(\omega)$$

and thus for our example

$$\Phi_{zz}(\omega) = \frac{(a^2 + c^2b^2) + c^2\omega^2}{\omega^2 + b^2} \quad (3)$$

The optimal *unrealizable* estimator yielding an estimate $\hat{y}(t|\infty)$ which minimizes the mean square error $E[\hat{y}(t|\infty) - y(t)]^2$ has the transfer function

$$H_u(\omega) = \frac{\Phi_{yy}(\omega)}{\Phi_{zz}(\omega)}$$

For our example,

$$H_u(\omega) = \frac{a^2}{c^2(\omega^2 + k^2)}$$

where $k^2 = b^2 + (a^2/c^2)$. This transfer function is easily inverted to yield the impulse response

$$H_u(t) = \frac{a^2}{2kc^2} e^{-k|t|}$$

The optimal *realizable* estimator yielding an estimate $\hat{y}(t - T|t)$ (where t is variable and T is fixed) is more difficult to calculate. The relevant formula for the transfer function involves a spectral factorisation of $\Phi_{zz}(\omega)$ as follows:

$$\Phi_{zz}(\omega) = \Phi^{+}_{zz}(\omega)\Phi^{-}_{zz}(\omega) \quad (4)$$

where $\Phi^{+}_{zz}(\omega)$ $[\Phi^{-}_{zz}(\omega)]$ has no lower [upper] half plane singularities or zeros. The transfer function of the realizable estimator is

$$H(\omega) = \frac{1}{2\pi\Phi^{+}_{zz}(\omega)} \int_{0}^{\infty} e^{-j\omega T} \int_{-\infty}^{\infty} \frac{\Phi_{zy}(\nu)}{\Phi^{-}_{zz}(\nu)} . e^{j\nu(t-T)} d\nu dt \quad (5)$$

where $\Phi_{zy}(\omega)$ is the cross spectral density. Of course, the spectral factorization of $\Phi_{zz}(\omega)$ is frequently a nontrivial problem, although for our example it is straightforward. Evaluation of the integrals in (5) for our example yields

$$H(\omega) = (k - b)\left[\frac{(k + b)e^{-j\omega T} - (b + j\omega)e^{-kT}}{k^2 + \omega^2}\right] \quad (6)$$

Two points should be noted concerning $H(\omega)$. First, $H(\omega)$ is not a rational function for $T > 0$ (the fixed-lag smoother case) and thus cannot be realized by a finite dimensional network—in contrast to the filtering case when $T = 0$ and $H(\omega)$ is rational. Secondly, there is a lower half plane and therefore unstable pole at $\omega = -jk$ and an identical lower half plane zero of $H(\omega)$. These cancel with one another in the filtering case when $T = 0$, but as can be seen, no cancellation is computationally possible in the fixed-lag smoothing case when $T > 0$, for it is not possible to present in closed-form the quotient resulting from dividing the numerator by $(j\omega + k)$.

The above two points have been made for the particular transfer function (6) but they are also applicable to the more general transfer functions (5) and thus help to explain the cause of instability problems in the earlier realizations of the fixed-lag smoother, which can be regarded as realization of (6), or more generally (5), without removal of the unstable poles. Filters on the other hand, normally have stable realizations, since the unstable poles are removable.

The impulse response of the transfer function (b) is

$$h(t) = \begin{cases} 0 & \text{for } t < 0, \\ \frac{k - b}{k} e^{-kt} (k \cosh kt + b \sinh kt) & \text{for } 0 \le t \le T, \\ \frac{k - b}{k} (k \cosh kT + b \sinh kT)e^{-kt} & \text{for } t \ge T \end{cases} \quad (7)$$

Observe that there is a cusp at $t = T$ and that this impulse response has the general shape of the unrealizable response $h_u(t - T)$ for $T \ge t \ge 0$, see Fig. 1.

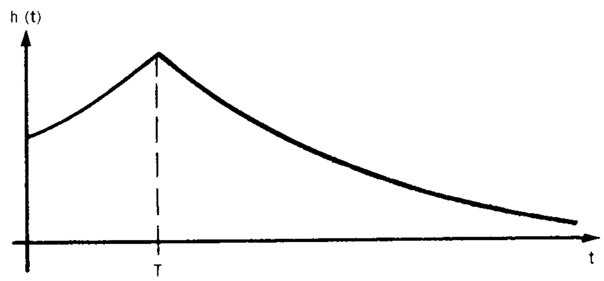

Fig. 1—Fixed-lag smoother impulse response

The realization of a system which achieves this impulse response for the filtering problem when $T = 0$ is straightforward, but for the fixed-lag smoothing case when $T > 0$, the realization of the desired impulse response $h(t)$, even for our simple example, is a non-trivial problem. Evidentally, approximation with a stable finite dimensional system is required, and a number of these may have to be tested for their suitability.

So far, we have touched on the realization of fixed-lag smoothers using Wiener's results, but have yet to examine their performance.

The *minimum mean-square error* for the optimal *unrealizable* estimator above is, by stationarity, independent of t, being

$$\mathrm{MMSE_u} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\Phi_{yy}(\omega)\Phi_{vv}(\omega)}{\Phi_{zz}(\omega)}\,d\omega = \frac{a^2}{2k} \tag{8}$$

The *minimum mean-square error* for the optimal *realizable* fixed-lag smoother is more complicated. However, for our example the MMSE reduces to

$$\mathrm{MMSE} = c^2(k-b)\left[1 - \frac{k-b}{2k}(1 - e^{-2kT})\right]$$

which clearly illustrates a general and intuitively reasonable result, namely, that *the MMSE is monotone decreasing as T increases and in the limit as T approaches infinity is the unrealizable error* (8).

The trade-off between the fixed-lag T and the MMSE as illustrated above suggests that in practice a fixed-lag need not be chosen which achieves more than say 95% of the improvement that is possible from smoothing. It seems reasonable in the case of our simple example to select T to be say 2 or 3 times the time constant of the optimal filter, namely $\tau = 1/k$. It turns out that more generally it is reasonable to *select T to be say 2 or 3 times the dominant time constant of any optimal filter* to achieve essentially all the improvement in performance that can be achieved by smoothing.

Before going on to a study of fixed-lag smoothing using Kalman filtering type ideas, we may note that Wiener's approach yields straightforward results for the design of optimal realizable filters, and very simple results for the design of unrealizable (or equivalently infinite lag) smoothers, but when it comes to designing stable realizable fixed-lag smoothers, the going is considerably more difficult. This is because it is difficult to carry out the manipulations to yield the optimal realizable estimator impulse response in the first place, and then it is tedious to design a stable finite-dimensional realization to approximate this response. The difficulties are clearly compounded if one seeks the best trade-off for a given application between estimator performance, estimator complexity, and the delay required.

3. Fixed-lag Smoothing of Discrete-time Signals via Kalman Filtering Techniques

Kalman filtering techniques can be applied when the signal model is described by state space equations such as the following equations for $k = 0, 1, 2, \ldots$.

$$\mathbf{x}(k+1) = \boldsymbol{\phi}\mathbf{x}(k) + \mathbf{G}\mathbf{w}(k) \tag{9}$$

$$y(k+1) = \mathbf{h}'\mathbf{x}(k+1) \tag{10}$$

Here, the state $\mathbf{x}(\cdot)$ is an n-vector, the disturbance $\mathbf{w}(\cdot)$ is a p-vector, and the signal $y(\cdot)$ is a scalar. The signal measurement $z(\cdot)$ is a scalar quantity given from

$$z(k+1) = y(k+1) + v(k+1) \tag{11}$$

Also the disturbances $\mathbf{w}(\cdot)$ and $v(\cdot)$ are independent zero mean, white gaussian sequences with covariance matrices

$$E[\mathbf{w}(k)\mathbf{w}'(k)] = Q, \quad E[v^2(k)] = R$$

with R positive. This signal model is more restrictive than the corresponding Wiener model in that a finite dimensionality constraint is involved, but we hasten to point out that, in contrast to the Wiener results, the results discussed in this section are readily generalized for time-varying, multiple output signal models.

The Kalman filter yields the filtered state estimates*

$$\hat{\mathbf{x}}(k|k) = E\{\mathbf{x}(k)|z(0), z(1), \ldots z(k)\}$$

from which signal estimates $y(k|k)$ can also be achieved since $\hat{y}(k|k) = \mathbf{h}'\mathbf{x}(k|k)$.

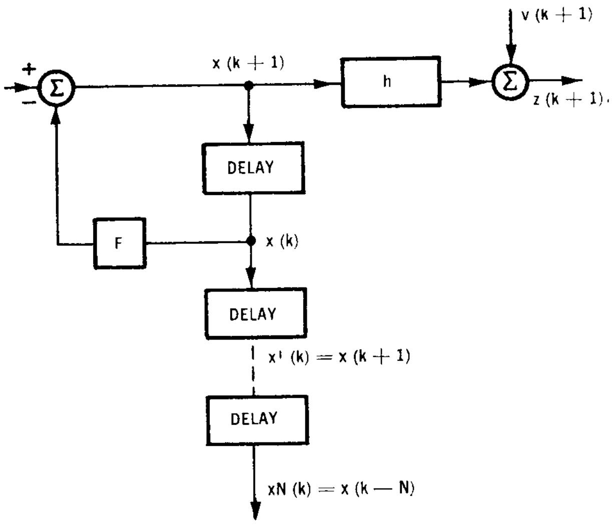

Fig. 2—Augmented signal model

In order to outline just how Kalman filter results can be applied to yield fixed-lag smoother results, suppose that the signal model (9)-(11) is augmented with delay elements as depicted in Fig. 2 and is described by the equations

$$\begin{bmatrix} x(k+1) \\ x^1(k+1) \\ x^2(k+1) \\ \cdot \\ \cdot \\ x^N(k+1) \end{bmatrix} = \begin{bmatrix} \emptyset & 0 & \cdot & \cdot & 0 \\ & & & & 0 \\ & & & & 0 \\ & & \mathbf{I}_N & & \cdot \\ & & & & \cdot \\ & & & & 0 \end{bmatrix} \begin{bmatrix} x(k) \\ x^1(k) \\ x^2(k) \\ \cdot \\ \cdot \\ x^N(k) \end{bmatrix} + \begin{bmatrix} G \\ 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} w(k) \tag{12}$$

$$y(k+1) = [1 \quad 0 \quad \cdot \quad \cdot \quad \cdot \quad 0] \begin{bmatrix} x(k) \\ x^1(k) \\ x^2(k) \\ \cdot \\ \cdot \\ x^N(k) \end{bmatrix} \tag{13}$$

* The conditioned mean estimate is also a minimum mean square error estimate in the scalar case.

where manipulations yield the identities $x^i(k) = x(k-i)$. Application of the Kalman filter equations to the signal model (12), (13) and (11) yields optimal estimates $\hat{\mathbf{x}}(k|k), \hat{\mathbf{x}}^1(k|k), \ldots \hat{\mathbf{x}}^N(k|k)$. The filtered estimates $\hat{\mathbf{x}}^i(k|k)$ are in fact the desired fixed-lag estimates $\hat{\mathbf{x}}(k-i|k)$ since from (12) and the definitions of filtered estimates and fixed-lag estimates, $\hat{\mathbf{x}}(k-i|k) = E\{\mathbf{x}(k-i)|z(0), z(1), \ldots z(k)\} = E\{\mathbf{x}^i(k)|z(0), z(1), \ldots z(k)\} = \hat{\mathbf{x}}^i(k|k)$.

After appropriate simplifications, Kalman filter equations (Refs. 4, 5) for the augmented signal model yield the fixed-lag smoothing equations together with the filtering equations as follows for $i = 1, 2, \ldots N$.

$$\begin{aligned}\hat{\mathbf{x}}(k+1|k+1) &= \mathbf{F}\hat{\mathbf{x}}(k|k) + \mathbf{k}\tilde{z}(k+1)\\ \hat{\mathbf{x}}(k+1-i|k+1) &= \hat{\mathbf{x}}(k+1-i|k) + \mathbf{k}_i\tilde{z}(k+1)\end{aligned} \tag{14}$$

Here $\tilde{z}(k+1)$ is a white noise innovations process given from

$$\tilde{z}(k+1) = z(k+1) - \mathbf{h}'\boldsymbol{\phi}\hat{\mathbf{x}}(k|k)$$

and $\mathbf{F}$ and $\mathbf{k}_i$ are given from

$$\begin{aligned}\mathbf{F} &= \boldsymbol{\phi} - \mathbf{kh}'\\ \mathbf{k} &= \bar{\mathbf{P}}\mathbf{h}(\mathbf{h}'\bar{\mathbf{P}}\mathbf{h} + R)^{-1}\\ \mathbf{P} &= \bar{\mathbf{P}}(\mathbf{I} - \mathbf{hk}')\\ \mathbf{P}_i &= \bar{\mathbf{P}}(\mathbf{F}')^i\\ \mathbf{k}_i &= \mathbf{P}_i\mathbf{P}^{-1}\mathbf{k}\end{aligned} \tag{15}$$

where $\bar{\mathbf{P}}$ is calculated from the equations

$$\begin{aligned}\bar{\mathbf{P}} &= \lim_{k\to\infty} \bar{\mathbf{P}}_k\\ \bar{\mathbf{P}}_{k+1} &= \boldsymbol{\phi}[\bar{\mathbf{P}}_k - \bar{\mathbf{P}}_k\mathbf{h}(\mathbf{h}'\bar{\mathbf{P}}_k\mathbf{h} + R)\mathbf{h}'\bar{\mathbf{P}}_k]\boldsymbol{\phi}' + \mathbf{GQG}'\\ \bar{\mathbf{P}}_0 &= 0\end{aligned} \tag{16}$$

In these equations, $\mathbf{P}$ is the covariance of the filtering error $[\mathbf{x}(k) - \hat{\mathbf{x}}(k|k)]$ and $\bar{\mathbf{P}}$ is the covariance of the one-stage prediction error $[\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k)]$. Both these covariances are independent of k. The steady state error covariance of the fixed-lag smoother is given from

$$\begin{aligned}\mathbf{P}_S &= \lim_{k\to\infty} E\{[\mathbf{x}(k) - \hat{\mathbf{x}}(k-N|k)][\mathbf{x}(k) - \hat{\mathbf{x}}(k-N|k)]'\}\\ &= \mathbf{P} - \sum_{i=1}^{N} \mathbf{D}^{-1}(\bar{\mathbf{P}} - \mathbf{P})(\mathbf{D}^{-1})'\end{aligned} \tag{17}$$

where $\mathbf{D} = \mathbf{P}(\mathbf{F}')^{-1}\mathbf{P}^{-1}$

Observe once again that as the lag N increases there is an improvement in smoother performance. Further manipulations yield that the error covariance reduction implied by (17) effectively obtains its maximum value when the fixed-lag N is of the order of two or three times the value of the dominant filter time constant. For any specified situation there appears to be no simple way to predict how much the improvement is going to be, other than by calculation via (17). In many cases the improvement may be negligible, but in other cases the improvement may be greater than 50%.

There are numerous alternative discrete-time fixed-lag smoothing structures including those driven directly from the filtered states (Refs. 13, 14) and various *reduced order* structures (Ref. 13). It should be noted that the abovementioned smoothers are all inherently stable, (in the sense of Lyapunov) in contrast to structures earlier suggested in the literature. Another feature which all these discrete-time fixed-lag smoothers share is that the trade-offs between the delay in estimation N, the estimator performance given by (17), and estimator dimension (dimension of estimator (14) is (n + 1)N) are immediately available to the designer.

A further point worthy of note is that for the case where the measurement set is finite, as for fixed-interval smoothing, it is not difficult to adapt the fixed-lag smoother to yield the fixed-lag smoothed estimate for the first part of the interval (equal in length to the interval itself minus the fixed lag), and the fixed-interval smoothed estimate for the latter part of the interval (equal in length to the fixed lag). All that is required is apply the recurrence relationships (14) but set $z(k+1) = 0$ and $\hat{\mathbf{x}}(k|k) = 0$ and thus $\tilde{z}(k+1) = 0$ for k greater than the fixed-interval magnitude. This result is interesting in the special case when the fixed lag is chosen as identical to the fixed interval, for then the adapted fixed-lag smoother is in fact an "on-line" fixed-interval smoother, which incidentally is more efficient computationally than the various well known "off-line" fixed-interval smoothers in the literature (Ref. 20); as these involve both a forward pass of the data through a filter and then a reverse pass of relevant data through a second filter.

4. Fixed-lag Smoothing of Continuous-time Signals via Kalman-Bucy Filtering Techniques

For continuous-time signals $y(\cdot)$ it is not difficult to express the fixed-lag smoothed estimate $\hat{y}(t-T|t)$ (of $y(t-T)$ given measurements to time t) in terms of a filtered estimate $y_T(t|t)$ of a signal $y_T(t)$ defined as $y_T(t) = y(t-T)$. In fact, this definition leads immediately to the identity

$$\hat{y}(t-T|t) = \hat{y}_T(t|t) \tag{18}$$

Pressing the analogy with the discrete-time results of the earlier section, we would expect that for the cases where $\hat{y}(t|t)$ can be found via Kalman-Bucy filtering techniques, then $y_T(t|t)$ could be found by these same techniques. This is not the case since the signal model for $y_T(t)$ involves a time delay and unless this is approximated by a finite dimensional system, and thereby an approximate value for $y_T(t|t)$ obtained, Kalman-Bucy filtering cannot be applied as such. It turns out that in practice, the filtering of an approximate value for $y_T(t)$ achieves an approximate value for $\hat{y}_T(t|t)$ and thus of $\hat{y}(t-T|t)$. To be more precise, let us consider the linear, time-invariant signal model equations as

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{Fx} + \mathbf{Gu}\\ y &= \mathbf{h}'\mathbf{x}\\ z &= y + v\end{aligned} \tag{19}$$

where $\mathbf{u}(\cdot)$ and $v(\cdot)$ are zero mean white gaussian processes with covariances $E[\mathbf{u}_T(t)\mathbf{u}'(\tau)] = \mathbf{Q}\delta(t-\tau)$ and $E[v(t)v(\tau)] = R\delta(t-\tau)$ with R positive definite. Also, $E[\mathbf{u}(t)\,v(\tau)] = 0$. (Once again, we have used a signal model which is more restrictive than the Wiener one in that a finite-dimensional constraint is involved, but we also note that results for time-varying multiple output systems are readily obtained.) In addition, we consider the following augmentation to the signal model (19),

$$\begin{aligned}\dot{\mathbf{x}}_a &= \mathbf{F}_a\mathbf{x}_a + \mathbf{G}_a y\\ y_a' &= \mathbf{h}_a\mathbf{x}_a + \mathbf{J}_a y\end{aligned} \tag{20}$$

where $\mathbf{F}_a$, $\mathbf{G}_a$, $\mathbf{h}_a$, and $\mathbf{J}_a$ are selected as the system matrices of an approximate delay network (where the delay is T) such as are studied in the networks literature. Thus, we have that $y_a(t) \simeq y_T(t) = y(t - T)$. Observe that the above augmentation does not affect the signal y or the measurements z. It is purely a device for solving the smoothing problem.

For the stationary smoothing problem, application of Kalman-Bucy filtering equations (Refs. 5, 6) to the augmented signal model, after some simplifications, yields the following estimator for $\hat{y}_a(t|t)$ and $\hat{y}(t|t)$.

$$\dot{\hat{\mathbf{x}}} = \mathbf{F}\hat{\mathbf{x}} + \mathbf{k}(z - \mathbf{h}'\hat{\mathbf{x}})$$

$$\mathbf{k} = \overline{\mathbf{P}}\mathbf{h}R^{-1}$$

$$\mathbf{F}\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}' - \mathbf{P}(t)\mathbf{h}R^{-1}\mathbf{h}'\mathbf{P}(t) + \mathbf{G}\mathbf{Q}\mathbf{G}' = \dot{\mathbf{P}}(t),\quad \mathbf{P}(0) = 0$$

$$\overline{\mathbf{P}} = \lim_{t\to\infty} \mathbf{P}(t) \tag{21}$$

and

$$\dot{\hat{\mathbf{x}}}_a = \mathbf{G}_a\mathbf{h}'\hat{\mathbf{x}} + \mathbf{F}_a\hat{\mathbf{x}}_a + \mathbf{k}_a(z - \mathbf{h}'\hat{\mathbf{x}})$$

$$\hat{y}_a = \mathbf{h}'_a\mathbf{x}_a$$

$$\mathbf{k}_a = \mathbf{P}_a\mathbf{h}R^{-1}$$

$$\mathbf{F}_a\mathbf{P}_a' + \mathbf{P}_a(\mathbf{F}' - \mathbf{h}\mathbf{k}') + \mathbf{G}_a\mathbf{h}'\mathbf{P} = 0 \tag{22}$$

The equations (21) are the Kalman-Bucy filter equations associated with the signal model (19), and equations (22) are augmentations of these associated with the augmentation of the signal model. The more general nonstationary equations can also be obtained.

The above filtering equations yield approximate values for the fixed-lag smoothed estimates $\hat{y}(t - T|t)$ since $y(t - T|t) = y_T(t) \simeq y_a(t)$ and thus $\hat{y}(t - T|t) = y_T(t|t) \simeq \hat{y}_a(t|t)$. Of course, the study of the approximations is a study in its own right, but the performance measure that is relevant is

$$E\{[\hat{y}_a(t|t) - y(t - T)]^2\}$$

which we denote as P_e. The value of P_e can be determined by solving linear equations (Ref. 16), but the details are perhaps too lengthy to summarize here. No firm results are available which offer help in the selection of a suitable value for the dimension of x_a in a first trial design, but experience indicates that if this is chosen to be of the order of that for x, a smoother design, in most cases, can be achieved with P_e virtually the same as the optimal fixed-lag smoothed signal error estimate which in turn may be considerably less than the filtered signal error.

In spite of the inherent tedium in determining a suitable trade-off between delay, estimator performance, and estimator complexity, such a study can be carried out systematically. The results can be readily extended to multiple output systems or to achieving fixed-lag smoothed *state* estimates but since more than one approximate delay is usually involved in solving such problems, any optimization procedure adopted has many more parameters to contend with.

5. Further Fixed-lag Smoothing Results

So far, we have presented Wiener fixed-lag smoothing results (chiefly because of their historical significance) and the application of Kalman filtering results to achieve fixed-lag smoothers (chiefly because this approach has both simplicity and power). Other results are, however, available and are very important for an understanding of the subject.

Ref. 17 derives simple formulas for the continuous-time fixed-lag smoothing problem as

$$\mathbf{x}(t - T|t) = \hat{\mathbf{x}}(t - T|t - T) + \mathbf{P}\int_{t-T}^{t} \boldsymbol{\phi}'_f(\tau - t)\mathbf{h}R^{-1}[z - \mathbf{h}'\hat{\mathbf{x}}(\tau|\tau)]d\tau \tag{23}$$

$$\mathbf{P}_S = \mathbf{P} - \mathbf{P}\int_{t-T}^{t} \boldsymbol{\phi}'_f(\tau - t)\mathbf{h}R^{-1}\mathbf{h}'\boldsymbol{\phi}_f(\tau - t)\,d\tau.\mathbf{P} \tag{24}$$

where $\boldsymbol{\phi}_f(\cdot)$ is the transition matrix associated with the system matrix $\mathbf{F} - \mathbf{k}\mathbf{h}'$ of the optimal filter. An alternative expression for $\hat{\mathbf{x}}(t - T|t)$ is given in Ref. 18 by

$$\hat{\mathbf{x}}(t - T|t) = \mathbf{P}\boldsymbol{\phi}'_f(T)\mathbf{P}^{-1}\hat{\mathbf{x}}(t|t) + \mathbf{P}\int_{t-T}^{t} \boldsymbol{\phi}'_f(\tau - t)\mathbf{P}^{-1}\mathbf{G}\mathbf{Q}\mathbf{G}'\mathbf{P}^{-1}\hat{\mathbf{x}}(\tau|\tau)d\tau \tag{25}$$

A relatively simple realization of (23) is depicted in Fig. 3 where $\mathbf{F}_f$ denotes the filter closed loop system matrix. Unhappily, in the usual case when $\mathbf{F}_f$ has negative eigenvalues, the matrix $(-\mathbf{F}_f')$ has positive eigenvalues, and as a consequence the realization of Fig. 3 contains instabilities.

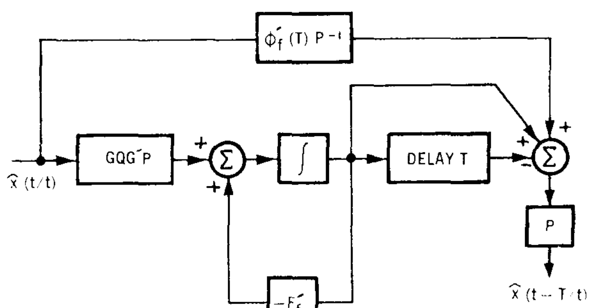

Fig. 3—Unstable fixed-lag smoother realization

The instability problem has been solved in Ref. 15 by permitting a resetting to zero of the states of the integrator in Fig. 3 at time instants 2nT for n = 1, 2, ... It turns out that using this resetting scheme solves the instability problems but the modified smoother output yields a smoothed estimate only over the time intervals [2nT, 2(n + 1)T]. However, switching between two such stable smoothers (the second reset at instants 2(n + 1)T) achieves a stable realization of the fixed-lag smoothing equations. Stability is achieved at the expense of resetting, switching, and complexity, rather than at the expense of optimality as in the previous section and as discussed in the next paragraph.

The instability problem associated with (23) and (25) can also be avoided by approximate but stable realizations of the impulse response associated with the smoothing equations (23) and (25). This approach also avoids the necessity for pure time-delay elements. For further details, see Ref. 16. A second approach to approximate realizations is to replace the integral in (23) or (25) by an approximating sum. Then, in the case of (25) for example, the smoothed estimate becomes a linear combination of weighted filtered estimates, all with various delays.

6. Concluding Remarks

The various fixed-lag smoothing results briefly discussed in the previous sections suggest that the possibility of fixed-lag smoothing for solving linear estimation problems should not be overlooked in favour of the more straightforward task of filtering, as appears to be the current practice. The improvement in estimator performance in many cases would justify the increase in estimator complexity required for fixed-lag smoothing.

A s o suggested by the results mentioned in the paper is the conc lusion that the "on-line" fixed-lag smoothing approach t o solving fixed-interval smoothing problems should be used in the interests of efficiency rather than the various "off-line" fixed-interval algorithms currently in use, even if this means the selection of a fixed lag equal to the period of the fixed interval.

7. References

1. Wiener, N., *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*, John Wiley and Sons, New York, 1949.
2. Bode H. W. and Shannon, C. E., "A simplified derivation of linear least-squares smoothing and prediction", *Proc. IRE*, 1950, vol. 38, pp. 417–425.
3. Zadeh, L. A. and Ragazzini, J. R., "An extension of Wiener's theory of prediction", *J. Appl. Phys.*, 1950, vol. 21, pp. 645–655.
4. Kalman, R. E., "A new approach to linear filtering and prediction problems", *Trans. ASME, Ser. D., J. Basic Engineering*, 1960, vol. 82, pp. 35–45.
5. Kalman, R. E., "New methods and results in linear prediction and filtering theory", *Tech. Rept.* no. 61–1, RIAS, Martin Co., Baltimore, Md., 1961.
6. Kalman, R. E. and Bucy, R. S., "New results in linear filtering and prediction theory", *Trans. ASME, Ser. D., J. Basic Engineering*, 1961, vol. 83, pp. 95–108.
7. Carlton, A. G., "Linear estimation in stochastic processes", *Bumblebee Series Rept.* No. 311, Applied Physics Lab., Johns Hopkins University, Silver Springs, Md., March 1962.
8. Rauch, H. E., "Linear estimation of sampled stochastic processes with random parameters", *Tech. Rept.* 2108–1, Stanford Electronics Lab., Stanford Univ., Stanford, Calif., April 1962.
9. Rauch, H. E., "Solutions to the linear smoothing problem", *IEEE Trans. Autom. Contr.*, 1963, vol. AC–8, pp. 371–372.
10. Bryson, A. E., Jr., and Frazier, M., "Smoothing for linear and nonlinear dynamic systems", *TDR* 63–119, Aero. Sys. Div., Wright-Patterson Air Force Base, Ohio, Feb. 1963, pp. 353–364.
11. Rauch, H. E., Tung, F. and Streibel, C. T., "Maximum likelihood estimates of linear dynamic systems", *AIAA J.*, 1965, vol. 3, pp. 1445–1450.
12. Weaver, C. S., "Estimating the output of a linear discrete system with gaussian inputs", *IEEE Trans. Autom. Contr.*, 1963, vol. AC–8, pp. 372–374.
13. Moore, J. B., "Discrete-time fixed-lag smoothing algorithms", *Automatica*, March 1973, vol. 9, no. 2, pp. 175–184.
14. Chirarattananon, S. and Anderson, B. D. O., "The fixed-lag smoother as a finite-dimensional linear system", *Automatica*, 1971, vol. 7, pp. 657–665.
15. Tam, P. K. S., Moore, J. B. and Anderson, B. D. O., "Stable realizations of the optimal continuous-time fixed-lag smoothing equations", *Proceedings of Third Symposium on Nonlinear Estimation Theory and its Applications*, to appear.
16. Chirarattananon, S. and Anderson, B. D. O., "Stable fixed-lag smoothing of continuous-time processes", *IEEE Trans Inf. Theory* (to appear).
17. Kailath, T. and Frost, P. A., "An innovations approach to least-squares estimation: Part II—Linear smoothing in additive white noise", *IEEE Trans. Autom. Contr.*, Dec. 1968, vol. AC–13, no. 6, pp. 646–655.
18. Anderson, B. D. O. and Chirarattananon, S., "New linear smoothing formulas", *IEEE Trans. Autom. Contr.*, Feb. 1972, vol. AC–17, no. 1, pp. 160–161.
19. Meditch, J. S., "A survey of data smoothing for linear and nonlinear dynamic systems", *Automatica*, March 1973, vol. 9, no. 2, pp. 151–162.
20. Fraser, D. C. and Potter, J. E., "The optimum linear smoothers as a combination of two optimum filters", *IEEE Trans. Autom. Contr.*, 1969, vol. AC–14, pp. 387–390.