Optimal Differentiation Based on Stochastic Signal Models

Bengt Carlsson, Anders Ahlén, *Member, IEEE*, and Mikael Sternad, *Member, IEEE*

***Abstract*—The problem of estimating the time derivative of a signal from sampled measurements is addressed. The measurements may be corrupted by colored noise. A key idea is to use stochastic models of the signal to be differentiated and of the measurement noise. Two approaches are suggested. The first is based on a continuous-time stochastic process as model of the signal. The second approach uses a discrete-time ARMA model of the signal and a discrete-time approximation of the derivative operator. The introduction of this approximation normally causes a small performance degradation, compared to the first approach. There exists an optimal (signal dependent) derivative approximation, for which the performance degradation vanishes.**

Digital differentiators are presented in a shift operator polynomial form. They minimize the mean-square estimation error. In both approaches, they are calculated from a linear polynomial equation and a polynomial spectral factorization. (The first approach also requires sampling of the continuous-time model.) Estimators can be designed for prediction, filtering, and smoothing problems. Unstable signal and noise models can be handled. The three obstacles to perfect differentiation, namely a finite smoothing lag, measurement noise, and aliasing effects due to sampling, are discussed.

I. Introduction

The need to obtain the time derivative of a measured or observed signal arises frequently. Industrial examples include the estimation of heating rates from temperature data [16] and of net flow rates into a tank from measurements of the level. In radar applications, velocity estimation from position data are of interest [44], [45]. Many biomechanical investigations require estimation of second-order derivatives (forces and moments) from position data [2], [20].

Being an important signal processing problem, numerical differentiation has been the subject of extensive investigations, see the survey papers [1]–[3]. A main complication is that differentiators amplify high-frequency noise. This problem grows with the order of the derivative to be estimated and with the required bandwidth of the filter.

For noise-free sampled signals, wide-band or full-band nth order differentiators can be designed [4]–[9]. The transfer function $(i\omega)^n$ should then be approximated by a realizable filter in some frequency band. This band may include all frequencies up to the Nyquist frequency (full-band differentiation).

If the signal is corrupted by noise, this must be taken into consideration. Loosely speaking, the filter design must be a compromise between good differentiation and low noise sensitivity, to achieve a small total error. Some lower bounds on the noise level of the filter output have been derived in [10]–[12].

Manuscript received June 12, 1989; revised March 5, 1990. This work was supported by the Swedish Board for Technical Development under Contract 87-01573.

The authors are with the Automatic Control and Systems Analysis Group, Department of Technology, Uppsala University, Uppsala, Sweden.

IEEE Log Number 9041158.

Differentiation of noisy data can be based on polynomial trend models [13]–[16] and on regularization techniques [1], [17]–[20]. Frequency domain design is also a popular approach. See, for example, [1], [9], [21]–[23], [46]. In an often considered situation, the signal is of low-frequency character, while the measurement noise is white. The filter should then approximate $(i\omega)^n$ at low frequencies and have low gain at high frequencies. Kalman filtering techniques have been applied to derivative estimation from measurements corrupted by white noise. See [3], [15], [24]–[26]. The Kalman filter is a frequently used tool for velocity estimation from radar position data. See, e.g., [44] and references therein.

This paper addresses the problem of estimating nth order derivatives based on stochastic signal models and noise-corrupted discrete-time measurements. Our goal will be to develop a design procedure which is general, yet simple to use. Measurements may be prefiltered and corrupted by colored noise. Nonstationary signals and noises, generated by unstable linear systems, will be handled. The estimator may be designed as a predictor, a filter or a fixed lag smoother. It is designed to minimize the mean-square estimation error. The estimator synthesis is based on a polynomial equations approach to linear quadratic optimization problems. The calculation of the filter basically requires the solution of a spectral factorization and a system of linear equations, corresponding to a linear polynomial equation. The result is a realizable Wiener filter.

We consider two signal models, corresponding to different types of *a priori* information known to the designer.

1) A continuous-time stochastic signal model is first assumed known. (This knowledge would be valuable since the derivative is, basically, a continuous-time concept.) After sampling of the model, the optimal filter can be calculated. An expression for the minimal estimation error variance is also derived. It is affected by the three basic obstacles to perfect estimation: a limit on the number of future data used for estimation, the presence of noise, and aliasing effects due to sampling.

2) If no continuous-time model is available, a discrete-time ARMA model may be obtained from the measured time series. Computation of an optimal differentiating filter then also requires a discrete-time approximation of the derivative operator. It will be proven that there exists an optimal (signal-dependent) approximation. Using this approximation, the filtering performance becomes identical to that based on a continuous-time model.

Equivalent estimators could be designed by state-space methods, using Kalman filtering. Adaptive smoothers for related problems are treated in [27]. We have preferred to use the polynomial equations approach. In contrast to Kalman filtering, it avoids problems in noise-free (singular) situations. It also leads

1053-587X/91/0200-0341$01.00 © 1991 IEEE

to simpler design calculations than for Kalman filters, in particular for smoothing problems with colored noise. The filter coefficients of the optimal differentiator are obtained directly and classical filter concepts, such as frequency responses, poles, and zeros, etc., can be studied directly. For a treatment of related estimation problems using the polynomial equations approach, see [28]–[30], [32], and [36]. The relationship between state-space and transfer function techniques has been thoroughly investigated by Kučera [36].

The paper is organized as follows. The (continuous-time) signal and (discrete-time) noise models are presented in Section II. The optimal estimator based on the continuous-time model and the expression for the minimal estimation error are discussed in Section III. An optimal filter based on a (discrete-time) ARMA model of the signal is derived in Section IV. The optimal discrete-time approximation of the derivative operator is presented in Section V. With it, the filter of Section IV equals the one derived in Section III. An example illustrates the design procedures in Section VI, and a numerical example is discussed in Section VII. Conclusions are presented in Section VIII.

II. Preliminaries

Let a continuous-time scalar signal $s_c(t)$ be characterized as a linear stochastic process

$$s_c(t) = G(p)\, e_c(t) \tag{2.1}$$

where $e_c(t)$ is zero mean white noise with spectral density $\lambda_c/2\pi$, and $G(p)$ is a rational function in the derivative operator $p \triangleq d/dt$. It has order $\delta \geq n + 1$ and pole excess (relative degree) $\geq n + 1$

$$G(p) = \frac{b_0 p^{\delta-n-1} + b_1 p^{\delta-n-2} + \cdots + b_{\delta-n-1}}{p^\delta + a_1 p^{\delta-1} + \cdots + a_\delta}. \tag{2.2}$$

There is no need for $s_c(t)$ to actually be a filtered white noise. The expression (2.1) just represents a model describing the spectral properties of the signal. We assume λ_c and $G(p)$ to be time invariant. The signal $s_c(t)$ may be filtered

$$s(t) = L(p)\, s_c(t)$$

where

$$L(p) = \frac{l_1 p^{\gamma-1} + l_2 p^{\gamma-2} + \cdots + l_\gamma}{p^\gamma + h_1 p^{\gamma-1} + \cdots + h_\gamma} + l_0 \tag{2.3}$$

of order γ, represents an antialiasing filter and/or the dynamics of a transducer. (If no filtering is used, set $\gamma = 0$ and $L(p) = l_0 = 1$.)

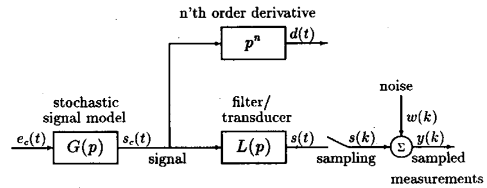


Fig. 1. The differentiation problem. The nth order derivative $d(t)$ of a signal $s_c(t)$ is to be estimated from sampled data. The signal $s_c(t)$ may be filtered by, for example, an antialiasing filter $L(p)$ before sampling. The measurements are corrupted by (possibly colored) noise $w(k)$.

The signal $s(t)$ is sampled with sampling period T. We seek the nth order derivative of the unfiltered signal $s_c(t)$

$$d(t) \triangleq \frac{d^n s_c(t)}{dt^n} = p^n G(p)\, e_c(t) = \frac{b_0 p^{\delta-1} + b_1 p^{\delta-2} + \cdots + b_{\delta-n-1} p^n}{p^\delta + a_1 p^{\delta-1} + \cdots + a_\delta} e_c(t) \tag{2.4}$$

at the time instants $t = kT$; $k = 0, 1, \cdots$. See Fig. 1.

The stochastic model (2.1)–(2.4) can be represented in state space form, [31], as

$$\begin{aligned} dx(t) &= Ax(t)\,dt + B\,dW(t) \\ s(t) &= H_1 x(t) \\ d(t) &= H_2 x(t) \\ s_c(t) &= H_3 x(t) \end{aligned} \tag{2.5}$$

with $dW(t) = e_c(t)\,dt$ being Wiener increments, and with

$$A = \left[\begin{array}{cccc|ccc} -a_1 & \cdots & \cdots & -a_\delta & & & \\ 1 & & & & & 0 & \\ & \ddots & & & & & \\ 0 & & 1 & 0 & & & \\ \hline 0 \cdots 0 & b_0 & \cdots & b_{\delta-n-1} & -h_1 & \cdots & -h_\gamma \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & 0 & & 1 \end{array}\right] \qquad B = \left[\begin{array}{c} 1 \\ 0 \\ \vdots \\ 0 \\ \hline 0 \\ \vdots \\ 0 \end{array}\right]$$

$$\begin{aligned} H_1 &= (\underbrace{0 \cdots 0}_{n} \;\; l_0 b_0 \cdots l_0 b_{\delta-n-1} \,|\, l_1 \cdots l_\gamma) \\ H_2 &= (b_0 \cdots b_{\delta-n-1} \;\; \underbrace{0 \cdots 0}_{n} \,|\, 0 \cdots 0) \\ H_3 &= (\underbrace{0 \cdots 0}_{n} \;\; b_0 \cdots b_{\delta-n-1} \,|\, 0 \cdots 0). \end{aligned} \tag{2.6}$$

Stochastic sampling (see for instance [31]) of (2.5), results in the discrete-time representation

$$\begin{aligned} x(k+1) &= Fx(k) + e_v(k) \\ s(k) &= H_1 x(k) \\ d(k) &\triangleq \left.\frac{d^n s_c(t)}{dt^n}\right|_{t=kT} = H_2 x(k) \end{aligned} \tag{2.7}$$

where $F = e^{AT}$. Note that $d(k)$ is exactly the derivative at the sampling instants. We assume the pair (F, H_1) to be detectable. (Possible unobservable modes must be stable.) The column vector $e_v(k)$ consists of discrete-time stationary white noise elements with zero mean. The covariance matrix equals

$$Ee_v(k)\, e_v(k)' \triangleq \lambda_c R_e = \lambda_c \int_0^T e^{A\tau}BB'e^{A'\tau}\, d\tau \qquad (2.8)$$

where $'$ denotes transpose. Note that while the continuous-time noise process $e_c(t)$ is scalar, $e_v(k)$ will be a vector of dimension $\delta + \gamma$ ($= \dim A$). In general, R_e has full rank.

If the filter $L(p)$ has a stable inverse, a continuous-time filter $p^n L(p)^{-1}$ of infinite bandwidth could, in principle, reconstruct the derivative $d(t)$ perfectly, from noise-free measurements of $s(t)$. With a discrete-time filter, this is impossible due to aliasing effects and the limited bandwidth of $s(k)$. (The signal $s_c(t)$ is not strictly band limited.) The effect of all components of $e_v(k)$ on $d(k) = H_2 x(k)$ cannot, in general, be calculated from their effect on $s(k) = H_1 x(k)$, unless the covariance matrix R_e has rank 1. When the sampling frequency increases, R_e approaches a rank 1 matrix. These points are discussed in more detail in Appendix D.

Measurements of the signal $s(k)$ are assumed to be corrupted by a discrete-time noise $w(k)$ (it normally represents a sampled continuous-time disturbance)

$$y(k) = s(k) + w(k). \qquad (2.9)$$

The sequence $\{w(k)\}$ is modeled as an ARMA process

$$w(k) = \frac{M(q^{-1})}{N(q^{-1})} v(k) \qquad (2.10)$$

where $M(q^{-1})$ and $N(q^{-1})$ are monic polynomials in the backward shift operator ($q^{-1}v(k) = v(k-1)$). They have degrees nm and nn, respectively. The sequence $\{v(k)\}$ is a zero mean and stationary white noise with variance λ_v. It is assumed that $v(k)$ is uncorrelated with all components of $e_v(k)$.

Kalman filtering [35] may be used to obtain an estimate of $d(k)$, based on measurements of $y(k)$ and on the model (2.7)–(2.10). Here, we will present a transfer-function based estimator design. For this reason, introduce the characteristic polynomial $D(q^{-1})$, of degree $nd = \delta + \gamma$, and the polynomial matrix $\mathbf{C}(q^{-1})$ as

$$D(q^{-1}) \triangleq \det(I - q^{-1}F) \qquad (2.11)$$

$$\mathbf{C}(q^{-1}) \triangleq \operatorname{adj}(I - q^{-1}F)q^{-1}. \qquad (2.12)$$

Hence, the sampled system can be expressed as

$$\begin{aligned} x(k) &= \frac{\mathbf{C}(q^{-1})}{D(q^{-1})} e_v(k) \qquad Ee_v(k)\, e_v(k)' = \lambda_c R_e \\ y(k) &= H_1 x(k) + w(k) \\ d(k) &= H_2 x(k) \\ w(k) &= \frac{M(q^{-1})}{N(q^{-1})} v(k) \qquad Ev(k)^2 = \lambda_v. \end{aligned} \qquad (2.13)$$

III. An Estimator Based on the Continuous-Time Signal Model

Assume the parameters of the continuous-time model (2.1)–(2.3) and of the noise description (2.10) to be known *a priori* or correctly estimated in some way. The discrete-time model

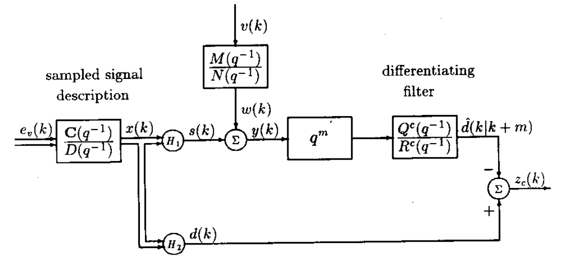


Fig. 2. The differentiation problem, originating from an *a priori* known continuous-time model. The nth order derivative $d(k)$ is to be estimated from the measurements $y(k+m)$.

(2.13) is then obtained by stochastic sampling. From this information, we seek the stable time-invariant linear estimator of the nth derivative

$$\hat{d}(k|k+m) = \frac{Q^c(q^{-1})}{R^c(q^{-1})} y(k+m) \qquad (3.1)$$

which minimizes the stationary mean-square estimation error

$$Ez_c(k)^2 \triangleq E(d(k) - \hat{d}(k|k+m))^2. \qquad (3.2)$$

See Fig. 2. Depending on m, we obtain a fixed lag smoother ($m > 0$), a filter ($m = 0$) or a predictor ($m < 0$).

Let us adopt the following polynomial notation. For any polynomial in the backward shift operator q^{-1}, of degree nu

$$U(q^{-1}) = u_0 + u_1 q^{-1} + \cdots + u_{nu} q^{-nu}$$

let $U_*(q) \triangleq u_0 + u_1 q + \cdots + u_{nu} q^{nu}$ and $\overline{U}(q^{-1}) \triangleq q^{-nu} U_*(q) = u_0 q^{-nu} + u_1 q^{-nu+1} + \cdots + u_{nu}$. In the frequency domain, the complex argument z is substituted for q. The polynomial arguments (q^{-1}, q, z^{-1}, z) will often be omitted. Stable polynominals have all zeros in $|z| < 1$.

Introduce the following polynomials,[1] obtained from the model (2.13)

$$\begin{aligned} P_{ij} &= p^{ij}_{nc} q^{nc} + \cdots p^{ij}_0 + \cdots p^{ij}_{-nc} q^{-nc} \\ &\triangleq H_i \mathbf{C}(q^{-1}) R_e \mathbf{C}'_*(q) H'_j, \qquad i, j = 1, 2. \end{aligned} \qquad (3.3)$$

Also, with $\eta \triangleq \lambda_v / \lambda_c$, introduce the polynomial spectral factorization

$$\tau \beta \beta_* = P_{11} N N_* + \eta D D_* M M_* \qquad (3.4)$$

defining a stable and monic spectral factor $\beta(q^{-1}) = 1 + \beta_1 q^{-1} + \cdots + \beta_{n\beta} q^{-n\beta}$ of degree $n\beta = \max\{nc + nn, nd + nm\}$ and a scalar τ. Reliable iterative algorithms for solving spectral factorizations with respect to τ and β exist [30], [40], [48]. For a stable spectral factor β to exist, it is necessary and sufficient that the two terms on the right-hand side of (3.4) have no common factors with zeros on the unit circle. (If $\eta = 0$, the first term should have no zeros on the unit circle.)

The polynomials P_{ij} and β have specific interpretations. Note, from (2.13), that for stationary signals (stable D and N), the

[1]The elements of the polynomial matrix $\mathbf{C}(q^{-1}) = \operatorname{adj}(I - q^{-1}F)q^{-1}$ are polynomials of degree $\delta + \gamma$, with leading coefficient zero, of type $n_1 q^{-1} + \cdots + n_{\delta+\gamma} q^{-\delta-\gamma}$, where δ and γ are the degrees of $G(p)$ and $L(p)$ in (2.2) and (2.3). The highest power of q^{-1} and q in P_{ij}, nc, will thus be given by $nc = \delta + \gamma - 1 = nd - 1$.

spectral densities of $\{s(k)\}$ and $\{d(k)\}$ will be given by

$$\phi_s(\omega) = \frac{\lambda_c}{2\pi}\frac{P_{11}}{DD_*}, \quad \phi_d(\omega) = \frac{\lambda_c}{2\pi}\frac{P_{22}}{DD_*}. \tag{3.5}$$

The cross spectral density between $\{d(k)\}$ and $\{s(k)\}$ is

$$\phi_{ds}(\omega) = \frac{\lambda_c}{2\pi}\frac{P_{21}}{DD_*} \tag{3.6}$$

where $e^{-i\omega T}$ and $e^{i\omega T}$ have been substituted for q^{-1} and q in all polynomials.

If D and N are stable, the spectral density of the measurement sequence $\{y(k)\}$ is given by

$$\phi_y(\omega) = \phi_s(\omega) + \phi_w(\omega) = \frac{\lambda_c}{2\pi}\frac{P_{11}}{DD_*} + \frac{\lambda_v}{2\pi}\frac{MM_*}{NN_*}$$
$$= \frac{\lambda_c}{2\pi}\frac{\tau\beta\beta_*}{DD_*NN_*}. \tag{3.7}$$

The spectral factor β thus represents the numerator of the innovations model

$$y(k) = \frac{\beta}{DN}\epsilon(k). \tag{3.8}$$

We are now ready to present the following result.

Theorem 1: Consider the sampled signal model described by (2.13). Assume that a stable spectral factor β, defined by (3.4), exists. Possible unstable factors of D are assumed not to be factors of $\bar{\beta}$. A stable linear estimator (3.1) of the derivative then attains the minimum of the estimation error (3.2), if and only if it has the same coprime factors as

$$\frac{Q^c}{R^c} = \frac{Q_1^c N}{\beta}. \tag{3.9}$$

Here $Q_1^c(q^{-1})$, together with a polynomial $L_*^c(q)$, is the unique solution to the linear polynomial equation

$$q^{-m}P_{21}N_* = \tau\beta_* Q_1^c + qDL_*^c \tag{3.10}$$

with polynomial degrees

$$nQ_1^c = \max\{nc + m, nd - 1\}$$
$$nL^c = \max\{nc + nn - m, n\beta\} - 1. \tag{3.11}$$

□

Proof: See Appendix A-2.

A. Remarks and Interpretations

1) Equation (3.10) can be written as a system of linear equations, with equal number of equations and unknowns ($nQ_1^c + 1 + nL^c + 1$). See (A.15) in Appendix A-3. This system has full rank, and precisely one solution $\{Q_1^c, L_*^c\}$.[2] The optimal filter (3.9) may sometimes contain stable common factors. Hence, the remark about coprime factors in Theorem 1.

[2] Linear polynomial equations $AX + BY = C$, also called "diophantine equations," in general have an infinite number of solutions $\{X, Y\}$. (If $\{X_0, Y_0\}$ is a solution, $\{X_0 + ZB, Y_0 - ZA\}$ will also be a solution, for any polynomial Z.) Because Q_1^c must be a polynomial in q^{-1}, and L_*^c a polynomial in q, there will, however, exist at most one solution to (3.10). The system of linear equations corresponding to (3.10) has full rank if and only if the unstable polynomial $q^{-n\beta}\beta_* = \bar{\beta}$ has no factor in common with D. (In particular, this will always be true if D is stable or has zeros on the unit circle.) Equation (3.10) then has a unique solution with respect to Q_1^c and L_*^c, with polynomial degrees (3.11).

2) It is shown in Appendix A-3 that $L^c \to 0$ when $m \to \infty$. Equation (3.10) then reduces to $q^{-m}P_{21}N_* = \tau\beta_* Q_1^c$. The impulse response of $Q_1^c(q^{-1})$ thus approaches that of $q^{-m}P_{21}N_*/\tau\beta_*$. If $d(k)$, $s(k)$, and $w(k)$ are stationary, the use of (3.9), (3.6), and (3.7) gives the limiting frequency domain expression

$$\left.\frac{Q^c(e^{-i\omega T})}{R^c(e^{-i\omega T})}e^{i\omega mT}\right|_{m\to\infty} = \frac{P_{21}N_*N}{\tau\beta_*\beta} = \frac{\phi_{ds}(\omega)}{\phi_s(\omega) + \phi_w(\omega)}. \tag{3.12}$$

This is the well-known unrealizable Wiener filter. When $L(p) = 1$ and no aliasing occurs $\phi_{ds}(\omega) = (i\omega)^n\phi_s(\omega)$, (3.12) then represents the differentiating Wiener filter [1], [3]. An "ideal" differentiator has transfer function $(i\omega)^n$, while the function $\phi_s(\omega)/(\phi_s(\omega) + \phi_w(\omega))$ provides the optimal tradeoff between ideal differentiation and suppression of noise.

3) Note that poles of the noise model (the zeros of N) are canceled by zeros in the estimator. If the noise model has poles close to the unit circle (resonances), the optimal filter will have notches at these frequencies. The filter design might be used also in the presence of nonstationary disturbances (zeros of N on or outside the unit circle). The filtering is, however, nonrobust in such situations. Imperfect cancellation of N would result in a nonstationary estimation error $z_c(k)$.

Corollary 1: With an optimal differentiating filter calculated according to Theorem 1, the minimal variance of the estimation error is finite. It is given by

$$Ez_c(k)^2_{\min} = \frac{\lambda_c}{2\pi i}\oint_{|z|=1} \cdot \left\{ \underbrace{\frac{L^cL_*^c}{\tau\beta\beta_*}}_{\mathrm{I}} + \eta\underbrace{\frac{MM_*P_{22}}{\tau\beta\beta_*}}_{\mathrm{II}} + \underbrace{\frac{NN_*[P_{11}P_{22} - P_{12}P_{21}]}{\tau\beta\beta_*DD_*}}_{\mathrm{III}} \right\} \cdot \frac{dz}{z}. \tag{3.13}$$

□

Proof: See Appendix A-1.

If the D-polynomial is not stable (as in the double integrator model discussed in Section VI), both $d(k)$ and the estimate $\hat{d}(k)$ will, in general, be nonstationary sequences. The estimation error $z_c(k) = d(k) - \hat{d}(k)$ will, however, be a stationary zero mean sequence, with a finite minimal variance given by (3.13). (This implies that unstable factors of D in the denominator of term III in (3.13) are canceled by numerator factors.)

The three terms in (3.13) can be interpreted as follows.

Term I represents the effect of a finite smoothing lag m. As is shown in Appendix A-3, $L^c \to 0$ when $m \to \infty$. Term I then vanishes.

Term II depends on the noise $w(k)$. It represents the unavoidable performance degradation due to noise, which cannot be eliminated even with an aribitrarily large smoothing lag m. The term vanishes in the noise-free case ($\eta = 0$).

Term III remains even when $m \to \infty$ and $\eta = 0$. It represents the performance degradation due to aliasing effects. Asymptotically, when $T \to 0$ and the covariance matrix $\lambda_c R_e$ (defined by

(2.8)) approaches a rank 1-matrix, the term vanishes. This can be shown as follows.

Assume R_e to be a rank 1 matrix, so that $R_e = VV'$, where V is a column vector. Using (2.13), define the polynomials $T(q^{-1}) \triangleq H_1C(q^{-1})V$, $U(q^{-1}) \triangleq H_2C(q^{-1})V$. Then, it is evident from (3.3) that

$$P_{11} = TT_* \qquad P_{22} = UU_* \qquad P_{12} = TU_* \qquad P_{21} = UT_*.$$

Consequently,

$$P_{11}P_{22} - P_{12}P_{21} = TT_*UU_* - TU_*UT_* = 0$$

and term III in (3.13) is eliminated. If R_e has rank >1, so that V is a matrix, T and U will be polynomial row vectors. Then, $TT'_*UU'_* - TU'_*UT'_* \neq 0$ in general, since polynomial vectors, unlike scalar polynomials, do not generally commute.

IV. An Estimator Based on a Discrete-Time Signal Model

Accurate continuous-time models, based on knowledge of the signal-generating process, are often hard to obtain. An obvious alternative is to use the discrete-time data series $\{y(k)\}$ itself, to obtain a model optimized through system identification [41], [42]. An ARMA model of the measured signal $y(k)$ will correspond to the innovations model (3.8)

$$y(k) = \frac{\beta(q^{-1})}{D(q^{-1})\,N(q^{-1})}\,\epsilon(k) \tag{4.1}$$

where $\beta(q^{-1})$ is stable. The white innovations sequence $\{\epsilon(k)\}$ has variance λ_ϵ. From another data series, where $s(k) = 0$, the noise model (2.10) may be obtained

$$w(k) = \frac{M(q^{-1})}{N(q^{-1})}\,v(k). \tag{4.2}$$

With (4.1) and (4.2) known, an ARMA model of the signal $s(k) = y(k) - w(k)$

$$s(k) = \frac{C(q^{-1})}{D(q^{-1})}\,e(k) \tag{4.3}$$

may be calculated by means of a spectral decomposition, see (4.4) below. Here, $e(k)$ is a scalar zero mean white noise with variance λ_e, while C and D are monic. Since $e(k)$ will not be estimated, the phase of $s(k)$, with respect to $e(k)$, will be of no interest. We may consider minimum phase models only. Thus, $z^{nc}C(z^{-1})$ is assumed to have no zeros in $|z| > 1$.

When N and D have zeros on or inside the unit circle, the spectral densities of $y(k)$, $w(k)$ and $s(k)$ are defined, except, possibly, for isolated frequencies. They are related through $\phi_y(\omega) = \phi_s(\omega) + \phi_w(\omega)$, or

$$\frac{\lambda_\epsilon}{2\pi}\frac{\beta\beta_*}{DD_*NN_*} = \frac{\lambda_e}{2\pi}\frac{CC_*}{DD_*} + \frac{\lambda_v}{2\pi}\frac{MM_*}{NN_*}.$$

This relation defines the polynomial spectral factorization equation

$$r\beta\beta_* = CC_*NN_* + \rho DD_*MM_* \tag{4.4}$$

where $r = \lambda_\epsilon/\lambda_e$ and $\rho = \lambda_v/\lambda_e$. With (4.1) and (4.2) known, D is easily determined from N and (DN). With only CC_* unknown, (4.4) represents an overdetermined system of linear equations in the coefficients of CC_*. They can be determined uniquely [33]. (Knowledge of C will not be required for determining an estimator. It will be sufficient to know the product CC_*. The stable polynomial C will be of use in Section V only. It can be determined uniquely by spectral factorization of CC_*.)

Alternatively, if the right-hand side of (4.4) is known, r and β may be calculated. Still another variant is the calculation of both the noise model (4.2) and the signal model (4.3), from knowledge of the innovations model (4.1) only. This is possible, under certain restrictive conditions on the polynomial degrees. See [33], [43].

If the signals defined by (4.1)–(4.3) have the same spectral densities as (3.5)–(3.7), it is obvious that β and D are the same polynomials in the two descriptions, while $\lambda_c P_{11} = \lambda_e CC_*$. The degree nc in (3.3) equals the degree of C in (4.3). Normally, $nc = nd - 1$. The spectral factorizations (3.4) and (4.4) differ by only a constant scaling factor λ_e/λ_c between ρ, r, and η, τ

$$\eta = \nu\rho \qquad \tau = \nu r \qquad \nu \triangleq \lambda_e/\lambda_c. \tag{4.5}$$

By means of the models (4.1)–(4.3), a description of the derivative is sought. The derivative $d(k)$ is, however, not related to the signals described by (4.1)–(4.3) in a simple way. Two procedures for overcoming this problem are conceivable:

1) *Inverse stochastic sampling* [37] may be used to obtain the continuous-time model (2.2) from the discrete-time model (4.3). An optimal estimator is then calculated by applying Theorem 1 on the resampled model (2.7), (2.13).

2) A *discrete-time approximation* of the derivative is introduced

$$\left.\frac{d^n s_c(t)}{dt^n}\right|_{t=kT}^{\text{approx}} \triangleq d_a(k) = q^l\,\frac{B(q^{-1})}{A(q^{-1})}\,s(k) \tag{4.6}$$

where A is stable. The integer $l \geq 0$ is introduced to handle noncausal approximations. Then, a stable linear estimator

$$\hat{d}_a(k|k+m) = \frac{Q^d(q^{-1})}{R^d(q^{-1})}\,y(k+m) \tag{4.7}$$

which minimizes the mean-square estimation error with respect to the approximative derivative $d_a(k)$

$$Ez_d(k)^2 \triangleq E\big(d_a(k) - \hat{d}_a(k|k+m)\big)^2 \tag{4.8}$$

is calculated. See Fig. 3.

We will discuss the second approach. It is much simpler than the use of inverse sampling. The introduction of an approximation (4.6) may degrade the filtering performance. The degradation will, however, be small if the approximation is reasonable and the sampling period is selected properly. (See Section VI.)

In the second approach, the approximation (4.6) (in addition to the smoothing lag m and the sampling period T) is a user choice. Let us assume that the first-order derivative is to be estimated ($n = 1$). Also, assume that $s(t)$ and $s_c(t)$ have the same spectral densities up to the Nyquist frequency.[3] The frequency response of (4.6) should then approximate $i\omega$, up to ω

[3] Either the filter (2.3) is $L(p) = 1$, or it is a good antialiasing filter, with $L(i\omega) \approx 1$ up to the Nyquist frequency ω_N, and $L(i\omega) \approx 0$ above ω_N. If $L(p) \neq 1$ in interesting frequency ranges, one alternative is to treat the problem as a deconvolution, or input estimation problem [32], [47].

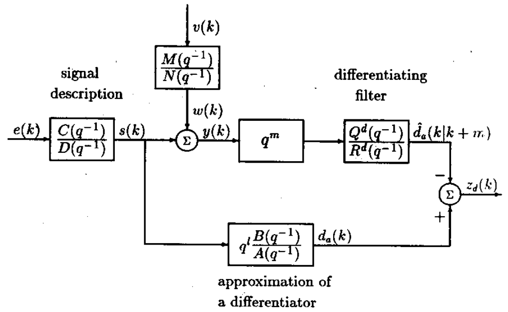


Fig. 3. A problem formulation based on a discrete-time signal model and differentiating approximation. The approximative derivative $d_a(k)$ is estimated.

$= \pi/T$. The simplest choice of approximation is the backward difference: $q^l B/A = (1 - q^{-1})/T$. A higher order alternative is

$$q^l \frac{B}{A} = q^l \sum_{n=1}^{l} \frac{(-1)^{n+1}}{nT} (q^{n-l} - q^{-n-l}). \tag{4.9}$$

This is a truncation of an unrealizable IIR filter which, for $l \to \infty$, has frequency response $i\omega$ for $|\omega| < \pi/T$. In the noise-free case, $q^m(Q^d/R^d)$ should equal the approximation (4.6), giving $z_d(k) \equiv 0$. The design of differentiating estimators for noise-free situations has been discussed in, for example, [4]–[6], [8], and [9]. When noise is taken into consideration, we obtain the following result.

Theorem 2: Let the discrete-time measurements be described by $y(k) = s(k) + w(k)$, using (4.2) and (4.3). Let an approximation (4.6) of the derivative and a smoothing lag m be given. Assume a stable spectral factor β to exist. Unstable factors of D are assumed not to be factors of $\bar{\beta}$. A stable linear estimator (4.7) of the derivative approximation (4.6) then attains the minimum of the estimation error (4.8), if and only if it has the same coprime factors as

$$\frac{Q^d}{R^d} = \frac{Q_1^d N}{\beta A}. \tag{4.10}$$

Here, $Q_1^d(q^{-1})$, together with a polynomial $L_*^d(q)$, is the unique solution to the linear polynomial equation

$$q^{l-m} CC_* N_* B = r\beta_* Q_1^d + qDAL_*^d \tag{4.11}$$

with polynomial degrees

$$nQ_1^d = \max\{nb + nc + m - l, na + nd - 1\}$$
$$nL^d = \max\{nc + nn + l - m, n\beta\} - 1. \tag{4.12}$$

The minimal value of (4.8) is

$$Ez_d(k)^2_{\min} = \frac{\lambda_e}{2\pi i} \oint_{|z|=1} \frac{(L^d L_*^d AA_* + \rho BB_* CC_* MM_*)}{r\beta\beta_* AA_*} \frac{dz}{z}. \tag{4.13}$$

Proof: See Appendix B.

Remarks:

- In the noise-free case ($\rho = 0$), with $m \geq l$, the optimal estimator (4.10) reduces to $q^{l-m}B/A$. This can be seen by inspection of (4.4) and (4.11) or directly from Fig. 3. We thus obtain $z_d(k) = 0$. Perfect estimation of $d_a(k)$ does, however, not imply perfect estimation of the true derivative. With a filter design according to Theorem 2, a new source of error appears. In addition to the three sources discussed after Corollary 1, we now get errors due to an imperfect differentiator approximation. In the next section, a (signal dependent) differentiator approximation which avoids such extra errors will be shown to exist.
- Remarks 1–3 after Theorem 1 apply also to Theorem 2, with small and obvious modifications. It is evident from (4.4) and (4.11) that the optimal estimator does not depend on the phase properties of C or M. Only the factors CC_* and MM_* appear in these expressions.
- For stable B, the differentiation problem, depicted in Fig. 3, may be interpreted as a deconvolution problem, cf. [32]. The signal $d_a(k)$ is then treated as the input to a known system, namely the inverse of the derivative operator approximation. Inversion of (4.6) gives $s(k) = q^{-l}(A/B)\, d_a(k)$. With $d_a(k) = (CB/DA)e(k)$ as an input description, the equations in [32] are directly applicable.

V. The Optimal Derivative Approximation

For a given derivative approximation (4.6), Theorem 2 provides the optimal filter with respect to $d_a(k)$. In general, it differs from the optimal filter with respect to the true derivative $d(k)$, given by Theorem 1. Let the basic assumptions of Theorem 2 hold. We can then prove the following result.

Theorem 3: Assume $z^{nc}C(z^{-1})$ to have no zeros in $|z| \geq 1$. Choose

$$l = m \geq 0 \tag{5.1}$$

$$A(q^{-1}) = C(q^{-1}) \tag{5.2}$$

and $B(q^{-1})$, together with a polynomial $K_*(q)$, as the unique solution to the linear polynomial equation

$$q^{-m} P_{21} = \nu C_* B + qDK_* \tag{5.3}$$

with polynomial degrees

$$nb = \max\{nc + m, nd - 1\} \qquad nk = nc - 1$$

where ν and P_{21} are defined in (4.5) and (3.3). Then, the estimator derived from the discrete-time model in Theorem 2 has the same coprime factors as the one obtained from the continuous-time model in Theorem 1. □

Proof: See Appendix C.

Thus, there exists an optimal approximation of the derivative operator. With it, the use of the discrete-time model (4.1)–(4.3) introduces no extra errors, compared to the use of a sampled continuous-time model. The error variance will be given by (3.13).

The optimal approximation obtained from (5.1)–(5.3) depends on the statistics of $s(k)$. Only a finite number of future values of $s(k)$, equal to the proposed smoothing lag m of the filter, are used in the approximation. The structure of the optimal approximation does not resemble an antisymmetrical FIR-filter like (4.9). Instead, it is an IIR filter, with the signal model numerator C as denominator polynomial. Note that Theorem 3 does not apply when C has zeros on the unit circle.

Theorem 3 is mainly of theoretical interest. It is of limited help for choosing a suitable approximation in practice. Calculation of the optimal polynomial B from (5.3) requires knowledge of P_{21} (the numerator of the cross spectral density (3.6) between $d(k)$ and $s(k)$). This polynomial could be obtained from (4.1)-(4.3) using inverse sampling. However, with knowledge of P_{21}, one might just as well design the estimator from Theorem 1 directly.

VI. Illustrations of the Results

We will in this section illustrate Theorems 1-3, for a simple example.

A. Design Based on Theorem 1

We wish to estimate a velocity, described as an integrated white noise, from position data. Thus, consider the double integrator $G(p) = 1/p^2$, with $L(p) = 1$. Representation in state space form (2.5) and sampling leads to (2.7), where

$$F = \begin{bmatrix} 1 & 0 \\ T & 1 \end{bmatrix} \quad R_e = \begin{bmatrix} T & T^2/2 \\ T^2/2 & T^3/3 \end{bmatrix}$$

$$H_1 = (0 \;\; 1) \quad H_2 = (1 \;\; 0). \tag{6.1}$$

From (2.11) and (2.12), we obtain

$$D = (1 - q^{-1})^2 \quad C = \begin{bmatrix} q^{-1} - q^{-2} & 0 \\ Tq^{-2} & q^{-1} - q^{-2} \end{bmatrix}. \tag{6.2}$$

Thus, the polynomials P_{ij} in (3.3) become

$$P_{11} = \frac{T^3}{6}[q^{-1} + 4 + q]$$

$$P_{21} = P_{12*} = \frac{T^2}{2} q[1 - q^{-1}][1 + q^{-1}]$$

$$P_{22} = T[-q^{-1} + 2 - q]. \tag{6.3}$$

White measurement noise is assumed, i.e., $M = N = 1$. We calculate Q^c/R^c for the filter case $m = 0$. Let the spectral factor β, of order 2, be denoted

$$\beta = 1 + \beta_1 q^{-1} + \beta_2 q^{-2} \tag{6.4}$$

and solve (3.10), with degrees $nL^c = 1$, $nQ_1^c = 1$

$$\frac{T^2}{2} q(1 - q^{-1})(1 + q^{-1}) = \tau\beta_* Q_1^c + q(1 - q^{-1})^2(L_0 + L_1 q). \tag{6.5}$$

Since $(1 - q^{-1})$ is a factor of two terms, it must be a factor of the third, too. Factor out $(1 - q^{-1})$. This gives $Q_1^c = Q_0(1 - q^{-1})$ and

$$\frac{T^2}{2} q(1 + q^{-1}) = \tau(1 + \beta_1 q + \beta_2 q^2)Q_0 + q(1 - q^{-1})(L_0 + L_1 q). \tag{6.6}$$

This polynomial equation corresponds to a set of linear equations in Q_0, L_0, and L_1. We can find Q_0 directly, by evaluating (6.6) for "$q = 1$," where the second term vanishes. With $1 + \beta_1 + \beta_2 = \beta(1)$, this gives

$$Q_0 = \frac{T^2}{\tau\beta(1)} = \frac{T^2\beta(1)}{\tau\beta(1)^2} = \frac{\beta(1)}{T}.$$

In the last equality, $\tau\beta(1)^2 = T^3$ was used. This relation is derived from (3.4) for "$q = 1$": $\tau\beta(1)^2 = P_{11}(1) + D(1)^2 = P_{11}(1) = T^3$, since $D(1) = 0$. Thus, the optimal differentiating filter is found to be

$$\frac{Q^c}{R^c} = \frac{\beta(1)(1 - q^{-1})}{T\beta(q^{-1})} \triangleq H^c. \tag{6.7}$$

B. Illustration of Theorem 3

Consider the optimal approximation B/A, according to Theorem 3. Let $C = 1 + c_1 q^{-1}$. Observe, from (3.4), (4.4), and (4.5), that $\lambda_e CC_* = \lambda_c P_{11}$. Thus

$$\nu(1 + c_1 q^{-1})(1 + c_1 q) = \frac{T^3}{6}(q^{-1} + 4 + q) \tag{6.8}$$

which is satisfied for

$$c_1 = 2 - \sqrt{3} \quad \nu = \lambda_e/\lambda_c = T^3/6c_1 = T^3/(1 + c_1)^2. \tag{6.9}$$

Solution of (5.3) and use of (5.2) leads to the approximation

$$\left(\frac{B}{A}\right)_{\text{opt}} = \frac{(1 + c_1)(1 - q^{-1})}{T(1 + c_1 q^{-1})}. \tag{6.10}$$

See [23] for details. From (4.10) and (4.11), the differentiating filter is then found to be

$$\frac{Q^d}{R^d} = \frac{\beta(1)}{T} \frac{(1 + c_1 q^{-1})(1 - q^{-1})}{\beta(q^{-1})(1 + c_1 q^{-1})}. \tag{6.11}$$

Apart from a stable common factor $C = A$, it coincides with (6.7), as expected.

C. Use of a Simple Derivative Operator Approximation

Let us use the approximation

$$\frac{B}{A} = \frac{1 - q^{-1}}{T} \tag{6.12}$$

instead of the optimal approximation (6.10) in (4.11). The differentiating filter is then found to be

$$\frac{Q^d}{R^d} = \frac{(1 - q^{-1})(Q_0 + Q_1 q^{-1})}{\beta(q^{-1})} \triangleq H^d \tag{6.13}$$

where

$$Q_0 = \frac{1}{T}\left(\beta(1) - \frac{c_1}{r}\right) \quad \text{and} \quad Q_1 = \frac{1}{T}\frac{c_1}{r}.$$

D. Performance Analysis

We will now investigate the estimator performance degradation due to noise, sampling, and the use of the simple approximation (6.12) of the derivative operator. At low frequencies, the estimator (6.13) approximates H^c from (6.7). This is evident from a Taylor expansion of $Q_0 + Q_1 q^{-1}$ in (6.13). Substitute $e^{-i\omega T}$ for q^{-1}

$$Q_0 + Q_1 e^{-i\omega T} \approx Q_0 + Q_1(1 - i\omega T)$$

$$= \frac{\beta(1) - c_1/r}{T} + \frac{c_1/r}{T}(1 - i\omega T)$$

$$\approx \frac{\beta(1)}{T} \quad \text{for } \omega T \ll 1.$$

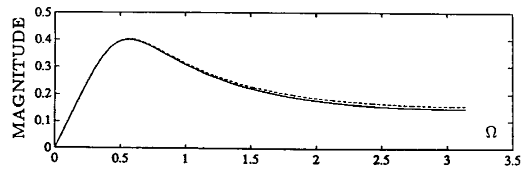


Fig. 4. Magnitude of TH^c (dotted line) and TH^d (solid line) as a function of $\Omega \triangleq \omega T$, for $\alpha = 10$.

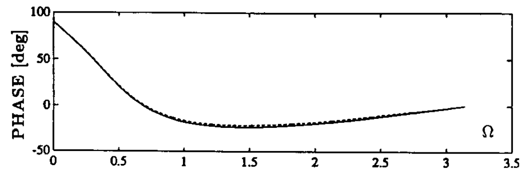


Fig. 5. Phase of TH^c (dotted line) and TH^d (solid line) as a function of $\Omega \triangleq \omega T$, for $\alpha = 10$.

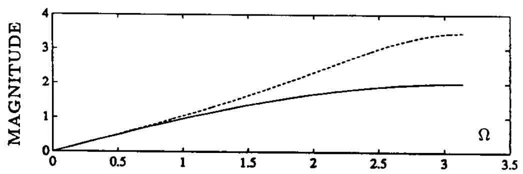


Fig. 6. Magnitude of TH^c (dotted line) and TH^d (solid line) as a function of $\Omega \triangleq \omega T$, for $\alpha = 0$. (Noise-free case.)

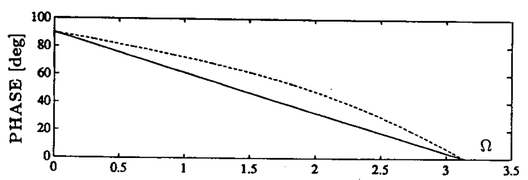


Fig. 7. Phase of TH^c (dotted line) and TH^d (solid line) as a function of $\Omega \triangleq \omega T$, for $\alpha = 0$. (Noise-free case.)

The estimators (6.7) and (6.13) depend on the sampling interval T and on the noise ratio $\eta = \lambda_v/\lambda_c$, which affect the spectral factor β. From the relations $\eta = \nu\rho$ and (6.9), we get $\rho = 6c_1(\lambda_v/\lambda_c T^3)$. For a given C and D, it is evident from (6.8), (3.4) and (4.4) that β, and thus the filters TH^c and TH^d (with normalized gains), are determined uniquely by the factor $\alpha \triangleq \lambda_v/(\lambda_c T^3)$. The ratio α may be large for two reasons: T small, or λ_v large, compared to λ_c. Hence, fast sampling and a low measurement noise may give the same filter as when slow sampling is used and the measurement noise is high. In the latter case, the differentiating filter performance will certainly be worse. See Table I and compare cases 1 and 2.

In the cases 1–3 in Table I, the estimation error is dominated by noise. The performance difference between the filters H^c from (6.7) and H^d from (6.13) is negligible. Hence, a simple approximation of the derivative operator will behave as well as a sophisticated one. The filters (6.7) and (6.13) differ significantly in the noise-free cases (see Fig. 6). Despite this, their performance differs only slightly. The reason is that the signal power at high frequencies, where the transfer functions differ, is insignificant compared to that at low frequencies.

TABLE I

DIFFERENTIATING FILTER PERFORMANCE FOR DIFFERENT SAMPLING INTERVALS T AND NOISE VARIANCES λ_v. $V^c \triangleq E(d(k) - H^c y(k))^2$, $V^d \triangleq E(d(k) - H^d y(k))^2$. THE RATIO $\alpha = \lambda_v/(\lambda_c T^3)$ UNIQUELY SPECIFIES BOTH TH^c AND TH^d

Case	T	λ_c	λ_v	α	V^c	V^d	Frequency Response
1	0.44	1	1	10	0.94	0.94	Figures 4 and 5
2	1	1	10	10	2.10	2.10	Figures 4 and 5
3	1	1	1	1	1.02	1.03	
4	0.1	1	0	0	0.029	0.033	Figures 6 and 7
5	10	1	0	0	2.9	3.3	Figures 6 and 7

Comparing cases 2 and 5 with the others, we see that if the measurement noise and/or the sampling period T is large, the estimation error is large. Case 4 shows that the estimation error is small if the sampling period and the measurement noise is small. The performance degradation due to sampling decreases rapidly with T. (In (3.13), $P_{11}P_{22} - P_{12}P_{21} = O(T^4) \rightarrow 0$ as $T \rightarrow 0$, in this example.)

From this investigation, we may conclude that the use of a simple approximation of the derivative operator is reasonable, if the sampling frequency is chosen sufficiently high, compared to the frequency content in the signal. When the sampling frequency is increased, the number of samples must, of course, be increased in order to cover a prespecified time interval.

VII. A NUMERICAL EXAMPLE

In this section, we will illustrate Theorem 2 for designing digital differentiating filters. The following signal and noise description will be used ($T = 1$):

$$\frac{C(q^{-1})}{D(q^{-1})} = \frac{1 - 0.180q^{-1} - 0.263q^{-2}}{1 - 0.285q^{-1} + 0.036q^{-2} - 0.638q^{-3}}$$

$$\frac{M(q^{-1})}{N(q^{-1})} = \frac{1 - 1.141q^{-1} + 1.082q^{-2} - 0.941q^{-3}}{1 - 1.081q^{-1} + 0.96q^{-2}}$$

$$\lambda_e = 1 \qquad \lambda_v = 0.5.$$

(Compare with Fig. 3.) The spectral densities of the signal $s(k)$ and of the noise $w(k)$ are shown in Fig. 8.

The spectral density of the noise has a resonance at $\omega_r^n = 1.0$ rad/s, with magnitude $M_r^n = 23.8$ dB. The spectral density of the signal has a resonance at $\omega_r^s = 2.0$ rad/s, with magnitude $M_r^s = 14.1$ dB.

The following approximation of the derivative operator is chosen (see [5])

$$\frac{B(q^{-1})}{A(q^{-1})} = \frac{1.150 - 0.378q^{-1} - 0.771q^{-2}}{1 + 0.860q^{-1} + 0.102q^{-2}}. \tag{7.1}$$

The transfer function magnitudes of the filter (7.1) and of the optimal filter, calculated from Theorem 2 with $m = 0$, are shown in Fig. 9.

We see from Fig. 9 that the optimal filter is close to B/A at frequencies where the signal spectral density dominates over the noise spectral density. The optimal filter has a notch at $\omega = 1$ (the resonance frequency of the noise).

Since the signal and/or the noise spectral densities may be incompletely known or time varying, the robustness properties of the filter are of interest. The result of an investigation of some nonideal design cases is presented in Table II.

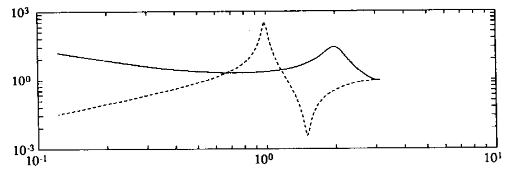


Fig. 8. Spectral densities of signal (solid line) and noise (dashed line).

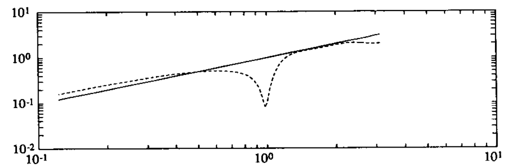


Fig. 9. Magnitude of the derivative approximation (7.1) (solid line). Magnitude of the optimal filter calculated from Theorem 2 (dashed line).

TABLE II
THE VARIANCE OF THE ESTIMATION ERROR FOR SOME ERRONEOUS DESIGN ASSUMPTIONS. THE VARIANCE OF THE ESTIMATED SIGNAL $d_a(k)$ IS 19.6

Case	Filter Designed from Theorem 2 Using:	$Ez_d(k)^2$
1	perfect signal and noise model	2.2
2	$\omega_r^n = 1.3$ rad/s, instead of $\omega_r^n = 1.0$ rad/s	5.4
3	$M_r^n = 16$ dB, -"- $M_r^n = 23.8$ dB	5.4
4	$\omega_r^s = 2.3$ rad/s, -"- $\omega_r^s = 2.0$ rad/s	2.4
5	$M_r^s = 0.5$ dB, -"- $M_r^s = 14.1$ dB	4.4
6	$\lambda_v = 1.0$, -"- $\lambda_v = 0.5$	2.4
7	$\lambda_v = 0$ (giving $Q/R = B/A$)	8.7

Most of the noise energy in this example is concentrated in a narrow peak. It is therefore natural that mismodeling of this peak degrades the performance more than errors in the signal model. None of the cases 1–6 considered in Table II leads to a worse performance than not using any stochastic noise model at all (case 7).

VIII. CONCLUSIONS

The design of digital differentiating filters has been addressed in a stochastic perspective. Signals to be differentiated and measurement noises were described by stochastic models. First, an approach based on a continuous-time signal model was considered. The model was sampled and an estimate of the derivative at the sampling instants was sought. The solution, minimizing the mean-square estimation error, involved a spectral factorization and a linear polynomial equation. The estimation error revealed, in a clear way, three contributing error sources: a finite smoothing lag, measurement noise, and aliasing effects. Even with an infinite smoothing lag and no noise, perfect differentiation is impossible since band-limited signals do not exist. A high sampling rate will, however, alleviate this effect.

A continuous-time model is not usually known *a priori*. Two alternatives are then available: inverse sampling of a discrete-time model or the use of a discrete-time approximation of the derivative operator. The second alternative has been investigated. The best estimator, in a mean-square sense, was found from the solution to a linear polynomial equation. This approach introduced an additional performance degradation due to an imperfect discrete-time approximation of the derivative. The existence of an approximation which eliminates this error was proven. It was found to consist of an IIR filter, having the numerator of the discrete-time signal model as denominator.

Using the optimal approximation of the derivative operator, the discrete-time approach gave a differentiating filter identical to that based on the continuous-time model. A practical problem is the signal dependence of the optimal approximation. However, an example stressed that a simple suboptimal approximation may be sufficient, if the measurement noise is significant and/or the sampling period is small enough.

An advantage with the suggested approaches is that prediction, filtering, and fixed-lag smoothing problems are treated in a unified way. In all three cases, the same design equation is used. The estimators are provided in transfer function form. Some of their frequency-domain properties are immediately obvious. The low gain of the estimators at noise resonances is evident from the numerator polynomials.

A limitation with the discussed differentiators is their time invariance. When derivatives have infrequent but large changes, time-varying estimators, combined with detection of changes, may provide much better performance [23]. A higher gain is then used for a short period of time after a change has been detected. The estimators could be retuned on line in such implementations. The noise ratios η in (3.4) or ρ in (4.4) may then be used as a time-varying tuning parameters.

While the estimator design is based on known and time-invariant models, it could form the basis of adaptive algorithms, as has been demonstrated for a related problem [43].

APPENDIX A
PROOFS OF RESULTS IN SECTION III

Using (2.13) and (3.1), the estimation error $z_c(k) = d(k) - \hat{d}(k \mid k+m)$ can be expressed as

$$z_c(k) = \left(\frac{R^c H_2 C - q^m Q^c H_1 C}{R^c D} \right) e_v(k) - q^m \frac{Q^c M}{R^c N} v(k). \tag{A.1}$$

First, we assume that the estimator is calculated according to Theorem 1. We show that $z_c(k)$ is then stationary even if D and/or N are unstable and that its variance is given by (3.13) (Corollary 1). Then, in subsection 2 below, it is shown that this is the minimal variance. In subsection 3, it is shown that $L^c \to 0$ as $m \to \infty$.

1. Proof of Stationarity and of Corollary 1

If D and N are stable, it is obvious that the estimation error is stationary, since $R^c = \beta$ is stable by construction. Cases where D and/or N contain unstable factors are now investigated. With (3.9), the estimation error (A.1) then becomes

$$z_c(k) = \left(\frac{\beta H_2 C - q^m Q_1^c N H_1 C}{\beta D} \right) e_v(k) - q^m \frac{Q_1^c N M}{\beta N} v(k). \tag{A.2}$$

The second term is stationary, since N in the denominator is cancelled exactly by the corresponding filter numerator factor. It remains to show that the numerator of the first term has D as a factor. Multiply the denominator and numerator in the first term of (A.2) by $\tau\beta_*$. Consider

$$W = \frac{\tau\beta\beta_* H_2 C - q^m Q_1^c \tau\beta_* N H_1 C}{\tau\beta\beta_* D}. \qquad (A.3)$$

We shall show that the numerator of (A.3) has D as a factor. This will then also hold for the first term of (A.2), since $\bar{\beta}$ and D are assumed to be coprime.

The use of (3.4) for $\tau\beta\beta_*$ and (3.10) for $Q_1^c\tau\beta_*$ gives

$$W = \frac{NN_*(P_{11}H_2 - P_{21}H_1)C}{\tau\beta\beta_* D} + \frac{D(\eta D_* MM_* H_2 + q^{m+1}L_*^c NH_1)C}{\tau\beta\beta_* D}. \qquad (A.4)$$

It now remains to show that D is a factor in $(P_{11}H_2 - P_{21}H_1)C$. The use of (3.3) gives

$$(P_{11}H_2 - P_{21}H_1)C = H_1 C R_e C_*^T H_1^T H_2 C - H_2 C R_e C_*^T H_1^T H_1 C = H_1 C_* R_e C^T(H_1^T H_2 - H_2^T H_1)C. \qquad (A.5)$$

Now note that $(H_1^T H_2 - H_2^T H_1)$ is skew-symmetric. Applying Lemma 3.1 in [23] on the factor $C^T(H_1^T H_2 - H_2^T H_1)C$ in (A.5) gives

$$C^T(H_1^T H_2 - H_2^T H_1)C = \mathfrak{N} D \qquad (A.6)$$

where $\mathfrak{N}$ is a polynomial matrix.

Hence, the numerator of (A.3) has D as a factor. Possible unstable factors of D are cancelled in W. As a consequence, $\{z_c(k)\}$ is stationary. Parseval's formula may therefore be used in order to express its variance. Noting that $e_v(k)$ and $v(k)$ are zero mean and mutually uncorrelated, the use of (A.2) and (3.3) gives

$$Ez_c(k)^2 = \frac{\lambda_c}{2\pi i}\oint \frac{(\beta\beta_* P_{22} - z^{-m}\beta P_{21} Q_{1*}^c N_* - z^m Q_1^c N P_{12}\beta_* + P_{11} Q_1^c Q_{1*}^c NN_* + \eta Q_1^c Q_{1*}^c DD_* MM_*)}{\beta\beta_* DD_*}\frac{dz}{z}$$
$$= \frac{\lambda_c}{2\pi i}\oint \frac{(\beta\beta_* P_{22} - z^{-m}\beta P_{21} Q_{1*}^c N_* - z^m Q_1^c N P_{12}\beta_* + Q_1^c Q_{1*}^c \tau\beta\beta_*)}{\beta\beta_* DD_*}\frac{dz}{z} \qquad (A.7)$$

where the integration path is counterclockwise around the unit circle. Completing the square gives for the numerator

$$\beta\beta_* P_{22} - \frac{1}{\tau}P_{12}P_{21}NN_* + \frac{1}{\tau}(z^m P_{12} N - \tau\beta Q_{1*}^c)(z^{-m}P_{21}N_* - \tau\beta_* Q_1^c) \qquad (A.8)$$

which, by using (3.4), (3.10) and noting that $P_{12} = P_{21*}$, becomes

$$\frac{1}{\tau}(P_{11}NN_* + \eta DD_* MM_*)P_{22} - \frac{1}{\tau}P_{12}P_{21}NN_* + \frac{1}{\tau}DD_* L^c L_*^c. \qquad (A.9)$$

Replacing the numerator in (A.7) with (A.9) gives (3.13). □

2. Proof of Theorem 1

Let an arbitrary derivative estimator be written as

$$\hat{d}(k|k+m) = \frac{Q^c}{R^c}y(k+m) + n(k) \qquad (A.10)$$

where Q^c/R^c is the estimator calculated according to Theorem 1 and where $n(k)$ is an arbitrary additional signal, generated from linear combinations of measurements $y(k)$ up to time $k + m$. It will be shown that $n(k) = 0$ is the optimal choice. (This proof technique has been used by Åström and Wittenmark [34].) The estimation error variance, when using (A.10), is given by

$$E(d(k) - \hat{d}(k|k+m))^2 = Ez_c(k)^2 - 2Ez_c(k)n(k) + En(k)^2 \qquad (A.11)$$

where $z_c(k)$ is the estimation error resulting from the use of (3.9). Stationarity of $z_c(k)$ was established in Appendix A-1. The signal $n(k)$ can be expressed as

$$n(k) = \frac{G(q^{-1})}{H(q^{-1})}y(k+m) \qquad (A.12)$$

where H is restricted to be stable. If N or D contain unstable factors, the criterion (A.11) is indefinite unless G cancels possible unstable factors of N and D. Hence, G is restricted to cancel these factors. Using (A.1), (A.12), and (2.13), the mixed term in (A.11) can be expressed as

$$2Ez_c(k)n(k) = 2E\left(\frac{[R^c H_2 C - q^m Q^c H_1 C]}{R^c D}e_v(k) - q^m\frac{Q^c M}{R^c N}v(k)\right)\cdot\frac{G}{H}q^m\left(\frac{H_1 C}{D}e_v(k) + \frac{M}{N}v(k)\right)$$
$$= \frac{\lambda_c}{\pi i}\oint\frac{[R^c H_2 C - z^m Q^c H_1 C]}{R^c D}R_e z^{-m}\frac{C_*' H_1' G_*}{D_* H_*}\frac{dz}{z} - \frac{\lambda_v}{\pi i}\oint\frac{Q^c MM_* G_*}{R^c NN_* H_*}\frac{dz}{z}$$
$$= \frac{\lambda_c}{\pi i}\oint\frac{[z^{-m}\beta P_{21}NN_* - Q_1^c NNN_* P_{11} - \eta Q_1^c NMM_* DD_*]G_*}{\beta DD_* NN_* H_*}\frac{dz}{z}$$
$$= \frac{\lambda_c}{\pi i}\oint\frac{[z^{-m}\beta P_{21}N_* - Q_1^c\tau\beta\beta_*]G_*}{\beta DD_* N_* H_*}\frac{dz}{z}.$$

In the second last equality, (3.3) and (3.9) were used and in the last one (3.4). Finally, by using (3.10), the mixed term reduces to

$$2Ez_c(k)n(k) = \frac{\lambda_c}{\pi i}\oint\frac{L_* G_*}{D_* N_* H_*}dz = 0. \qquad (A.13)$$

The expression (A.13) is zero because H is stable and G is assumed to cancel all unstable factors in D and N. Thus, the integrand has no poles inside the unit circle. Since the mixed term in (A.11) is zero, the criterion is obviously minimized by choosing $n(k) = 0$. Any estimator transfer function with coprime factors different from those of (3.9) would, by definition, correspond to $n(k) \neq 0$. No such estimator attains the minimal error variance.

The degrees of Q_1^c and L_*^c in (3.10) are determined by the requirement that the maximal occurring powers in q^{-1} and q, respectively, in the polynomial equation must be covered. □

3. The Polynomial Equation, in the Limit $m \to \infty$

Multiply (3.10) by q^{-nL^c-1}, to obtain an equation in powers of q^{-1} only

$$\tau\bar{\beta}Q_1^c + \bar{D}\bar{L}^c = q^{-nL^c-1-m}P_{21}N_*. \tag{A.14}$$

Write the right-hand side of (A.14) as

$$q^{-nL^c-1-m}P_{21}N_* \triangleq q^{-f}(s_0 + s_1q^{-1} + \cdots + s_{ns}q^{-ns}).$$

Since the highest powers in q of P_{21} and N_* are nc and nn, respectively, we have $f = nL^c + 1 + m - nc - nn$, which increases with m. By equating equal powers of q^{-1} on the right- and left-hand side of (A.14), the following set of simultaneous equations is obtained:

$$\begin{bmatrix} \tau\beta_{n\beta} & & 0 & 1 & & 0 \\ \vdots & \ddots & & d_1 & \ddots & \\ \tau\beta_1 & & \tau\beta_{n\beta} & \vdots & & 1 \\ \tau & & \vdots & d_{nd} & & d_1 \\ & \ddots & \tau\beta_1 & & \ddots & \vdots \\ 0 & & \tau & 0 & & d_{nd} \end{bmatrix} \begin{bmatrix} Q_0 \\ \vdots \\ Q_{nQ_1^c} \\ L_{nL^c} \\ \vdots \\ L_0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ s_0 \\ \vdots \\ s_{ns} \end{bmatrix} \begin{matrix} \uparrow \\ f \\ \downarrow \\ \\ \\ \\ \end{matrix} \tag{A.15}$$

As $m \to \infty$, nQ_1^c grows, while $nL^c = n\beta - 1$. With an increasing smoothing lag, less and less information about $d(k)$ will be present in the measurement $y(k+m)$ [35]. Thus, the leading coefficients of Q_1^c tend to zero as m increases. This will be true, in particular, for the first nL^c coefficients.

Consider the first nL^c equations of (A.15). Since $\{Q_0, \cdots, Q_{nL^c}\} \to 0$ as $m \to \infty$, it follows that $\{L_{nL^c}, \cdots, L_0\} \to 0$ as $m \to \infty$, since the right-hand side of the equations are zero.

Appendix B
Proof of Theorem 2

This proof follows the proof of Theorem 1, in Appendix A. We therefore present it as an outline. Using (4.6) and (4.10), the estimation error can be expressed as

$$z_d(k) = \frac{(Cq^l\beta B - q^m CQ_1^d N)}{\beta AD} e(k) - q^m \frac{Q_1^d M}{\beta A} v(k). \tag{B.1}$$

It is simple to show that it is stationary, even for unstable signal models C/D. Using Parseval's formula, (4.4), (4.10), and (4.11), it is straightforward to derive the expression (4.13) for the minimal estimation error variance. Proceeding as in Appendix A-2, we propose the use of an arbitrary estimator

$$\hat{d}_a(k|k+m) = \frac{Q^d}{R^d} y(k+m) + n(k) \tag{B.2}$$

where $n(k) = (G/H)y(k+m)$ is stationary and Q^d/R^d is given by (4.10).

Evaluation of the filtering error, using (B.2) and (4.4), gives the mixed term

$$2Ez_d(k)\,n(k) = \frac{\lambda_e}{\pi i} \oint \frac{(z^{l-m}CC_*N_*\beta B - r\beta\beta_*Q_1^d)}{\beta ADD_*N_*} \frac{G_*}{H_*} \frac{dz}{z}$$

which, by using (4.11), is found to be

$$2Ez_d(k)\,n(k) = \frac{\lambda_e}{\pi i} \oint \frac{zDAL_*^d G_*}{ADD_*N_*} \frac{dz}{z}. \tag{B.3}$$

Since G_* is required to cancel possible unstable factors of D and N, (B.3) will have no poles inside the unit circle. The integral will thus be zero. As in Appendix A-2, the optimal choice must be $n(k) = 0$. □

Appendix C
Proof of Theorem 3

With a given model (2.1), (2.2), the estimator Q^c/R^c is calculated from Theorem 1. We seek $\{l, A, \text{and } B\}$, such that Q^d/R^d from Theorem 2 equals Q^c/R^c. Introduce $K_*(q)$ and $L_*^d(q)$ such that

$$L_*^c - \nu L_*^d \triangleq N_*K_*. \tag{C.1}$$

Here L_*^c is determined, while L_*^d and K_* are undetermined.

By solving (5.3), we get B and K_*. Once K_* is determined, L_*^d is also determined from (C.1). It is easily verified that the degrees of L_*^d obtained from (C.1) and (4.12) are consistent. Multiply both sides of (4.11) with ν. Use (5.1) and the expressions (C.1) and (5.3) for νL_*^d and qDK_*. This gives

$$\begin{aligned} \nu r\beta_*Q_1^d &= \nu CC_*N_*B - qDA(L_*^c - N_*K_*) \\ &= \nu CC_*N_*B - qDAL_*^c + AN_*(q^{-m}P_{21} - \nu C_*B) \\ &= \nu CC_*N_*B - \nu C_*AN_*B - qDAL_*^c + AN_*q^{-m}P_{21} \\ &= \tau\beta_*Q_1^cA. \end{aligned}$$

In the last equality, we used (5.2) and (3.10). Since $\nu r = \tau$ (cf. (4.5)), we get $Q_1^d = Q_1^cA$. Canceling the stable factor A, the filter (4.10) becomes

$$\frac{Q^d}{R^d} = \frac{Q_1^cAN}{\beta A} = \frac{Q^c}{R^c}. \qquad \square$$

Appendix D
Exact Estimation of $d(k)$ and the Rank of R_e

Assume that R_e, defined by (2.8), has rank 1. Then the vector $e_v(k)$ in (2.7) can be expressed by a column vector M and a scalar noise source $e(k)$ with variance λ_e as $e_v(k) = Me(k)$, with $R_e = MM'$.

In transfer function form, with $q^{-1}e(k) \triangleq e(k-1)$ and $D(q^{-1}) \triangleq \det(I - Fq^{-1})$, (2.7) then gives

$$s(k) = H_1(I - Fq^{-1})^{-1}q^{-1}Me(k) \triangleq \frac{C_1(q^{-1})}{D(q^{-1})} e(k)$$

$$d(k) = H_2(I - Fq^{-1})^{-1}q^{-1}Me(k) \triangleq \frac{C_2(q^{-1})}{D(q^{-1})} e(k).$$

If the polynomial C_1 is stable, the filter $\hat{d}(k) = (C_2/C_1)s(k)$ is stable, and provides a perfect estimate of the derivative $d(k)$, from noise-free measurements of the signal $s(k)$. This is a simple example of an "unknown-input observer" [38], [39]. Such observers are generically impossible to construct if the number of unknown inputs (rank R_e in our case) is larger than the number of measured signals (one in our case, since $s(k)$ is scalar). However, if the sampling frequency is large, the matrix R_e becomes (approximately) a rank 1-matrix. This can be seen by using a series expansion of $e^{A\tau}$ in (2.8)

$$R_e = \int_0^T \left(I + A\tau + A^2\frac{\tau^2}{2} + \cdots\right) \cdot BB'\left(I + A'\tau + A'^2\frac{\tau^2}{2} + \cdots\right) d\tau \rightarrow \int_0^T BB'\, d\tau = \begin{pmatrix} T & \\ & 0 \end{pmatrix} \quad \text{as } T \rightarrow 0.$$

The performance degradation due to sampling then becomes negligible.

References

[1] R. S. Anderssen and P. Bloomfield, "Numerical differentiation procedures for nonexact data," *Numer. Math.*, vol. 22, pp. 157–182, 1974.
[2] G. A. Wood, "Data smoothing and differentiation procedures in biomechanics," *Exercise Sport Sci. Rev.*, vol. 10, pp. 308–362, 1982.
[3] B. Carlsson, T. Söderström, and A. Ahlén, "Digital differentiating filters," Uppsala University, Sweden, Rep. UPTEC 8792R, Sept. 1987.
[4] J. F. Kaiser, "Digital filters," in *Systems Analysis by Digital Computer*, F. F. Kuo and J. F. Kaiser, Eds. New York: Wiley, 1966, pp. 218–285.
[5] L. R. Rabiner and K. Steiglitz, "The design of wide-band recursive and nonrecursive digital differentiators," *IEEE Trans. Audio Electroacoust.*, vol. AU-18, pp. 204–209, June 1970.
[6] J. H. McClellan, T. W. Parks, and L. R. Rabiner, "A unified approach to the design of optimum FIR linear-phase digital filters," *IEEE Trans. Circuit Theory*, vol. CT-20, pp. 697–701, Nov. 1973.
[7] J. Spriet and J. Bens, "Optimal design and comparison of wide-band digital on-line differentiators," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 46–52, Feb. 1979.
[8] A. Antioniou and C. Charalambous, "Improved design procedure for equiripple digital differentiators," in *Proc. IEEE Int. Symp. Circuits Syst.* (Chicago, IL), Apr. 1981, pp. 540–545.
[9] S.-C. Pei and J.-J. Shyu, "Design of FIR Hilbert transformers and differentiators by eigenfilter," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-37, pp. 505–511, Apr. 1989.
[10] H. Lanshammar, "On precision limits for derivatives numerically calculated from noisy data," *J. Biomechan.*, vol. 15, pp. 459–470, 1982.
[11] D. L. Ragozin, "Error bounds for derivative estimates based on spline smoothing of exact or noisy data," *J. Approximation Theory*, vol. 37, pp. 335–355, 1983.
[12] R. J. Marks, II, "Noise sensitivity of band-limited signal derivative interpolation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 1028–1032, Aug. 1983.
[13] C. Lanczos, *Applied Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1956.
[14] J. H. J. Allum, "A least squares CUBIC algorithms for on-line differential of sampled analog signals," *IEEE Trans. Comput.*, vol. 24, pp. 585–590, June 1975.
[15] S. Fioretti, L. Jetto, and T. Leo, "A discrete-time delay system for efficient simultaneous derivative estimation," in *Signal Processing III: Theories and Applications*, I. T. Young *et al.*, Eds. Elsevier Science Publishers B. V. (North-Holland), 1986, pp. 89–92.
[16] B. Carlsson, "Double exponential smoothing with adaptive forgetting applied to differentiation," Uppsala University, Sweden, Rep. UPTEC 8715R, Oct. 1987.
[17] P. Craven and G. Wahba, "Smoothing noisy data with spline functions," *Numer. Math.*, vol. 31, pp. 377–403, 1979.
[18] H. Hatze, "The use of optimally regularized Fourier series for estimating higher order derivatives of noisy biomechanical data," *J. Biomechan.*, vol. 14, pp. 13–18, 1981.
[19] M. L. Baart, "Computational experience with the spectral smoothing method for differentiating noisy data," *J. Computational Phys.*, vol. 42, pp. 141–151, 1981.
[20] H. J. Woltring, "On optimal smoothing and derivative estimation from noisy displacement data in biomechanics," *Human Movement Sci.*, vol. 4, pp. 229–245, 1985.
[21] S. Usui and I. Amidror, "Digital low-pass differentiation for biological signal processing," *IEEE Trans. Biomed. Eng.*, vol. BME-29, pp. 686–693, Oct. 1982.
[22] B. Kumar and S. C. D. Roy, "Design of digital differentiators for low frequencies," *Proc. IEEE*, vol. 76, pp. 287–289, Mar. 1988.
[23] B. Carlsson, "Digital differentiating filters and model-based fault detection," Ph.D. dissertation, Uppsala University, Sweden, 1989.
[24] T. Söderström, "Estimation of time derivatives from noisy data," Uppsala University, Sweden, Rep. UPTEC 8290R, Sept. 1982.
[25] A. Ahlén, "An input estimation approach to differentiation of noisy data," Uppsala University, Sweden, Rep. UPTEC 8450R, May 1984.
[26] T. Glad and L. Ljung, "Velocity estimation from irregular noisy position measurements," in *Proc. IFAC World Congr.* (Budapest, Hungary), vol. VI, July 1984, pp. 275–279.
[27] J. M. Mendel, *Optimal Seismic Deconvolution. An Estimation-Based Approach*. New York: Academic, 1983.
[28] M. J. Grimble, "Polynomial systems approach to optimal linear filtering and prediction," *Int. J. Contr.*, vol. 41, pp. 1545–1564, June 1985.
[29] A. P. Roberts and M. M. Newmann, "Polynomial approach to Wiener filtering," *Int. J. Contr.*, vol. 47, pp. 681–696, Mar. 1988.
[30] V. Kučera, *Discrete Linear Control. The Polynomial Equation Approach*. Chichester: Wiley, 1979.
[31] K. J. Åström, *Introduction to Stochastic Control Theory*. New York: Academic, 1970.
[32] A. Ahlén and M. Sternad, "Optimal deconvolution based on polynomial methods," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-37, pp. 217–226, Feb. 1989.
[33] A. Ahlén, "Identifiability of the deconvolution problem," *Automatica*, vol. 26, pp. 177–181, Jan. 1990.
[34] K. J. Åström and B. Wittenmark, *Computer Controlled Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
[35] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.
[36] V. Kučera, *Analysis and Design of Discrete Linear Control Systems*. Prague: Academia, and London: Prentice-Hall International, 1991.
[37] T. Söderström, "Algorithms for computing stochastic continuous-time models from ARMA models," Uppsala University, Sweden, Rep. UPTEC 89030R, Apr. 1989.
[38] G. Basile and G. Marro, "On the observability of linear time-invariant systems with unknown inputs," *J. Optimiz. Theory Appl.*, vol. 3, pp. 410–415, 1969.
[39] E. Emre and M. L. Hautus, "A polynomial characterization of (A, B)-invariant and reachability subspaces," *SIAM J. Contr. Optimiz.*, vol. 18, pp. 420–436, 1980.
[40] Z. Vostrý, "New algorithm for polynomial spectral factorization with quadratic convergence," *Kybernetika* (Prague), vol. 11, pp. 415–422, 1975.
[41] T. Söderström and P. Stoica, *System Identification*. Hemel Hempstead: Prentice-Hall International, 1989.
[42] L. Ljung, *System Identification. Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
[43] A. Ahlén and M. Sternad, "Adaptive input estimation," in *Proc. IFAC Symp. ACASP-89: Adaptive Syst. Contr., Signal Processing* (Glasgow, Scotland), Apr. 1989, pp. 631–636.
[44] F. Farina and F. A. Studer, *Radar Data Processing*, vol. I. New York: Research Student Press and Wiley, 1985.
[45] C. C. Schooler, "Optimal $\alpha - \beta$ filters for systems with mod-

eling inaccuracies," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-11, pp. 1300–1306, Nov. 1975.

[46] G. W. Medlin and J. W. Adams, "A new design technique for maximally linear differentiators," in *Preprints ICASSP* (Glasgow, Scotland), vol. 2, May 1989.

[47] B. Carlsson, M. Sternad, and A. Ahlén, "Digital differentiation of noisy data, measured through a dynamic system," submitted for publication.

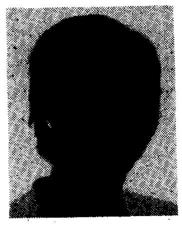
[48] C. J. Demeure and C. T. Mullis, "A Newton–Raphson method for moving-average spectral factorization using the Euclid algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 10, pp. 1697–1709, Oct. 1990.

Bengt Carlsson was born in Uppsala, Sweden, in 1958. He received the M.S. degree in engineering physics in 1984, the licentiate degree in 1987, and the Ph.D. degree in automatic control in 1989, all from the Institute of Technology, Uppsala University, Uppsala, Sweden.

From 1983 to 1989 he held various teaching positions at the Automatic Control Group, Uppsala University. He was a Visiting Research Associate at the Department of Electrical and Computer Engineering, University of Newcastle, NSW, Australia, January to August, 1988. Currently, he is employed as an Assistant Professor at the Automatic Control Group in Uppsala. His research interests include adaptive signal processing, design of differentiating filters, and model based fault detection.

Anders Ahlén (S'80–M'84) was born in Kalmar, Sweden, in 1953. He received the Ph.D. degree in automatic control in 1986, from the Institute of Technology, Uppsala University, Uppsala, Sweden.

He has worked with the Automatic Control Group at Uppsala University, from 1978 to 1984 as a Research and Teaching Assistent; from 1984 to 1989 as an Assistant Professor; and from 1990 as an Associate Professor in Automatic Control. His research interests are adaptive control and estimation, adaptive signal processing; deconvolution algorithms, differentiating filters and problems related to equalization, noise canceling, and feedforward control.



Mikael Sternad (S'83–M'86) was born in Eskilstuna, Sweden in 1957. He received the M.S. degree in engineering physics in 1981 and the Ph.D. degree in automatic control in 1987, both from the Institute of Technology, Uppsala University, Sweden.

He is with the Automatic Control Group at Uppsala University, where he was a Teaching Assistant (1981–1986), an Assistant Professor (1987), and is now an Associate Professor. His research interests have included the study of optimal and adaptive feedforward regulators. It is now centered around the polynomial equations approach to linear quadratic optimization. It also includes robust control and signal processing and adaptive algorithms for control and signal processing.