Hence we choose $B=N$, where N is the $n\times n$ matrix such that $d^Te^{Ah}N=0$. Unknown elements of N are obtained from (10)–(12).

Construction of B can also be obtained from the following equivalent equations. Let $L(t)=-Z(t)e^{A(t-h)}$ be a solution of (6). Then it follows from (6) and (9) that

$$\dot{Z}(t)=(AZ-ZA)-B \tag{14}$$

$$d^T\left(e^{Ah}Z(h)-Z(2h)e^{Ah}+e^{2Ah}\right)=0. \tag{15}$$

Let us assume a series solution of $Z(t)$ of the form

$$Z(t)=\sum_{n=0}^{\infty}\frac{Z_nt^n}{n!}.$$

From (14), it follows that

$$B=(AZ_0-Z_0A)-Z_1 \tag{16}$$

$$Z_i=(AZ_{i-1}-Z_{i-1}A), \qquad i=2,3,4,\cdots,\infty. \tag{17}$$

From (10), (11), and (16), it follows that

$$d^T\left(A^nZ_0-Z_0A^n-Z^{(n)}+e^{Ah}A^n\right)B=0 \tag{18}$$

$$d^Te^{Ah}\left(A^nZ_0-Z_0A^n-Z^{(n)}\right)B=0, \qquad n=0,1,2,3,\cdots,\infty \tag{19}$$

where $Z^{(n)}$ are given by the following recurrence relations:

$$Z^{(n)}=A^{n-1}Z_1+Z^{(n-1)}A, \quad n=1,2,3,\cdots,\infty$$

$$Z^{(n)}=0, \qquad n\leqslant 0.$$

Hence, to find B we choose Z_0, Z_1, satisfying (15) and (17) containing unknown parameters, and these unknown parameters are determined from (18) and (19), and finally B is determined from (16).

Example

Let $h=1$, $d^T=(1,-2,-1,0)$,

$$A=\begin{bmatrix}0&2&0&0\\0&0&-1&0\\0&0&0&0\\0&0&0&0\end{bmatrix}.$$

Then from (15)–(19) we find

$$B=AZ_0-Z_0A-Z_1=\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&2&0&0\\b_1&b_2&b_3&b_4\end{bmatrix}.$$

Popov's construction in this case gives B as

$$B=\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&2&0&0\\0&b&0&0\end{bmatrix}$$

where b,b_1,b_2,b_3,b_4 are arbitrary.

V. Conclusions

The above method of construction of degenerate systems is more general than earlier constructions and gives all possible solutions. It can be shown that if the system is regular, then the foregoing method of construction reduces to Popov's method.

References

[1] B. A. Asner, Jr. and A. Halanay, "Algebraic theory of pointwise degenerate delay-differential systems," *J. Differential Equations*, vol. 14, pp. 293–306, 1973.
[2] A. K. Choudhury, "Necessary and sufficient conditions of pointwise completeness of linear time-invariant delay-differential systems," *Int. J. Contr.*, vol. 16, pp. 1083–1100, 1972.
[3] V. M. Popov, "Pointwise degeneracy of linear time-invariant delay-differential equations," *J. Differential Equations*, vol. 11, pp. 541–561, 1972.

Minimum Error Dispersion Linear Filtering of Scalar Symmetric Stable Processes

B. W. STUCK

***Abstract*—The well-known Kalman–Bucy linear-filtering theory for Gaussian Markov processes is generalized to cover a particular class of non-Gaussian Markov processes, the scalar symmetric stable Markov processes. Results are presented only for discrete time because of certain pathologies that arise in the continuous-time analog (except in the Gaussian case). Attention is confined to the scalar case because of technical problems arising in characterizing multivariate stable distributions (except in the Gaussian case).**

I. Introduction

Since the original work by Wiener [1], [2], using spectral factorization techniques, and Kalman and Bucy [3], [4], using time domain methods, linear filtering of Gaussian Markov random processes with rational spectra has been extensively studied in both the engineering [5] and information-theoretic [6] literature.

Here we deal with a natural extension of the Kalman–Bucy technique to linear filtering of a particular class of scalar Markov random processes that include the Gaussian as a special case, the symmetric stable random processes in discrete time. The motivation for this work is drawn from problems associated with noise on telephone lines [7], extremely low-frequency electromagnetic communication [8], and the random-walk model of stock-market prices [9], [10].

One reason for the popularity of Gaussian random processes in modeling uncertainty is that often a large number of independent disturbances contribute to the uncertainty in a measurement, and at that point the Central Limit Theorem of probability theory is invoked to claim that the noise is Gaussian. It is often overlooked that the Gaussian is only one member of a family of distributions, the stable distributions, that arise from examining the limiting distributions of suitable normalized and translated sums of random variables. In an engineering situation, one would like to be able to combat uncertainty in a manner independent of the choice of a particular stable distribution.

Two main reasons have contributed to attention being confined to the Gaussian rather than non-Gaussian stable case. First, only in the Gaussian case does there exist a closed-form analytic expression for the probability density function; in the non-Gaussian stable case one must resort to power series and asymptotic series. However, with the availability of digital computers that are inexpensive to use, numerical calculations of error probabilities are straight-forward to carry out. The second main reason that non-Gaussian stable distributions have not been studied or used as widely as Gaussian stable distributions is that only the Gaussian distribution has a finite second moment, and hence a finite variance, and infinite variance is felt to be physically inappropriate. While superficially appealing, this reason is inadequate because the Gaussian distribution is unbounded, which is clearly physically inappropriate. The question here is over what range the model should fit observed data: it may well be that a finite-variance Gaussian model is

Manuscript received July 22, 1977. This work was originally presented at the 1974 IEEE International Information Theory Symposium, Notre Dame, IN.
The author is with Bell Laboratories, Murray Hill, NJ 07974.

0018-9286/78/0600-0507$00.75 ©1978 IEEE

adequate over a limited range of data, while an infinite-variance stable-distribution model is adequate over a larger range of data, but both models are inappropriate outside of a given range of observations, although the infinite-variance model is better than the finite-variance model in terms of matching observed data over a wider range. For sufficiently large deviations from the model nonlinearities and limiting operations would become significant and make either model inappropriate.

Three items are particularly interesting about these results.

1) At the present time, no analogous Wiener–Hopf spectral factorization theory exists for the problems considered here. Hopefully, they will provide stimulus for clarifying the role played by Hilbert space orthogonal projection techniques [1], [2], [6].

2) The generalization to the multivariate case is nontrivial because of the complicated nature of multivariate stable distributions [12]; an Appendix gives the reader who is not familiar with stable distributions a brief summary of their properties in the scalar case. Only in the Gaussian case is it seemingly straightforward to extend these results to multivariate distributions. Because of this we deal only with the scalar case.

3) The extension of these results to continuous time results in a degenerate problem, because of certain pathologies in the sample paths of stable processes [11]. In particular, we can discriminate perfectly between two stable processes with differing parameters in any finite time interval because the associated path space probability measures are orthogonal [11]. The intuitive reason for this is that continuous-time stable processes have sample paths that are left continuous with right-hand limits everywhere defined and are constant except for jumps of random amplitudes at random time epochs; however, the jump time epochs, while countable in any finite time interval, are dense, and thus there is a countably infinite set of jumps available for processing. It may not be not surprising that with an infinite amount of data, i.e., observations of the jump epochs and jump amplitudes, perfect filtering and discrimination is possible. In the Gaussian case the sample paths consist of jumps of zero amplitude, and perfect filtering and discrimination is not possible in all cases. This reasoning is only meant as a plausibility argument to suggest why attention is confined here solely to the discrete time case; the diligent reader should consult the references for the detailed technical arguments substantiating these claims [11].

II. The Result

The reader is referred to the Appendix for a brief discussion of properties of stable distributions. The Appendix also fixes the notation and nomenclature used throughout this note.

Let $\{u_k\}$, $\{w_k\}$, $k=0,1,\cdots$, be sequences of mutually independent stable random variables with the same characteristic exponent $\alpha, 0<\alpha\leqslant 2$, but differing dispersion parameters, denoted by γ_u and γ_w, respectively, where k is the discrete time parameter.

$$E(e^{iwu_k})=\exp(-\gamma_u|\omega|^\alpha)\quad E(e^{i\omega w_k})=\exp(-\gamma_w|\omega|^\alpha),\qquad k=0,1,\ldots.$$

Here we have explicitly assumed $\{u_k\}$, $\{w_k\}$ are sequences of symmetric stable random variables. The dispersion plays the same role as the variance in the Gaussian case, and thus $\gamma^{1/\alpha}$ plays the same role as the standard deviation in the Gaussian case. For example, if x is a stable random variable with dispersion γ, then if we scale x by $\gamma^{1/\alpha}$, we see that with location parameter zero,

$$E[\exp(i\omega x\gamma^{-1/2})]=\exp(-|\omega|^\alpha)$$

and hence the natural units of x are in terms of $\gamma^{1/\alpha}$. For simplicity of exposition, we omit the case of asymmetric or skewed stable distributions.

The model of the signal process $\{x_k\}$ and observation process $\{y_k\}$ is assumed to be of the form

$$x_{k+1}=Ax_k+Bu_k\quad E(e^{i\omega x_0})=\exp(-\gamma_{x_0}|\omega|^\alpha)\quad y_k=Cx_k+w_k.$$

This encompasses both autoregressive and moving average models, which are currently quite popular in applications.

We now closely follow the original Kalman–Bucy [3], [4] work. We seek a linear signal estimate using a recursive algorithm such that the error sequence has minimum dispersion for each value of time. Thus, we denote by $\{\hat{x}_k\}$ the signal estimate sequence, and $\{e_k\equiv x_k-\hat{x}_k\}$ denotes the error sequence.

Proposition

Given the preceding assumptions, the minimum error-dispersion linear signal-estimate recursive sequence and error sequence satisfy the following relations:

$$\hat{x}_{x+1}=A\hat{x}_k+G_k(y_k-C\hat{x}_k)\quad E(e^{i\omega\hat{x}_0})=\exp(-\gamma_{\hat{x}_0}|\omega|^\alpha)$$

$$e_{k+1}=(A-G_kC)e_k+Bu_k-G_kw_k$$

$$E(e^{i\omega e_0})=\exp(-\gamma_{e_0}|\omega|^\alpha),\qquad \gamma_{e_0}=\gamma_{x_0}+\gamma_{\hat{x}_0}$$

where for $1<\alpha\leqslant 2$

$$G_k=\frac{A}{C}\frac{\left[|C|^\alpha\gamma_{e_k}\right]^{1/(\alpha-1)}}{\left[|C|^\alpha\gamma_{e_k}\right]^{1/(\alpha-1)}+\gamma_w^{1/(\alpha-1)}}$$

$$\gamma_{e_{k+1}}=|B|^\alpha\gamma_u+\left|\frac{A}{C}\right|^\alpha\frac{\gamma_w|C|^\alpha\gamma_{e_k}}{\left[\gamma_w^{1/(\alpha-1)}+\left\{|C|^\alpha\gamma_{e_k}\right\}^{1/(\alpha-1)}\right]^{(\alpha-1)}}$$

while for $0<\alpha\leqslant 1$

$$G_k=\begin{cases}A/C & \gamma_w<|C|\gamma_{e_k}\\ 0 & \gamma_w>|C|\gamma_{e_k}\end{cases}$$

$$\gamma_{e_{k+1}}=|B|^\alpha\gamma_u+\begin{cases}|A/C|^\alpha\gamma_w & \gamma_w<|C|\gamma_{e_k}\\ |A/C|^\alpha|C|^\alpha\gamma_{e_k}, & \gamma_w>|C|\gamma_{e_k}.\end{cases}$$

Proof. The method of proof is straightforward. The basic idea is to minimize the error dispersion for each successive datum or observation as much as possible and thus minimize the pointwise error dispersion. The question of whether or not there exists a solution to this problem, and whether it is unique, can be handled by exhibiting the unique solution explicitly. The error dispersion is simply given by

$$\gamma_{e_{k+1}}=|A-G_kC|^\alpha\gamma_{e_k}+|B|^\alpha\gamma_u+|G_k|^\alpha\gamma_w.$$

Two cases arise. For $1<\alpha\leqslant 2$, this is a convex function of G_k; thus it is necessary to compute the derivative of $\gamma_{e_{k+1}}$ with respect to G_k, set it equal to zero, and solve for the resulting recursive gain. For $0<\alpha\leqslant 1$, the error dispersion is not a strictly convex function of G_k, and we must examine the extreme points of the interval $(0,A/C)$. Depending on the values of γ_w,γ_u, and $|C|$, one or the other of these extrema will minimize the error dispersion.

Having solved for the recursive gain for one particular value of k, the result is substituted into the recursion for the error dispersion and the process is repeated. The results are summarized in the above formulas. Q.E.D.

Note that the recursive gain sequence $\{G_k\}$ is a continuous function of the characteristic exponent α in the neighborhood of $\alpha=2$, the Gaussian case. Preliminary evidence strongly indicates that *nonlinear* filtering offers performance superior to *linear* filtering, as seen by [13] and the references therein. We leave that subject for future work.

Appendix
A Brief Summary of Properties of Stable Distributions

Suitable references on stable distributions are Feller [14] or Gnedenko and Kolmogorov [15] from which this material is taken.

Definition.

A probability distribution $P(x)$ is said to be stable if for all $a_1>0, b_1$, $a_2>0, b_2$ there exists constants $a>0, b$ such that

$$P(a_1x+b_1)*P(a_2x+b_2)=P(ax+b).$$

Theorem

The characteristic function of a stable distribution can be put in the following canonical form

$$E(e^{i\omega x})=\int_0^\infty e^{i\omega x}dP(x)\equiv\exp(i\omega\delta)$$

$$\cdot\begin{cases}\exp(-\gamma|\omega|^\alpha[1+i\beta\,\mathrm{sign}(\omega)]), & 0<\alpha\leqslant 2,\ \alpha\neq 1\\ \exp(-\gamma|\omega|[1+i\beta\,\mathrm{sign}(\omega)\ln(\gamma|\omega|)]), & \alpha=1\end{cases}$$

where $-1\leqslant\beta\leqslant 1$, $\gamma>0$, $-\infty<\delta<\infty$.

The parameter α is called the characteristic index of the stable distribution $P(x)$: for $-1<\beta<1$

$$\lim_{x\to\infty}|x|^\alpha P(-x)=k_-\geq 0\qquad \lim_{x\to\infty}|x|^\alpha[1-P(x)]=k_+>0$$

From this it is easy to see that $E(|x|^p)$ is finite for $p<\alpha$, $-1<\beta<1$, or, in particular, that for $0<\alpha\leqslant 1$ stable distributions possess no mean, while for $1<\alpha<2$ stable distributions possess a mean but no variance. The parameter β is associated with the skewness of the distribution: when $\beta=0$, the distribution is symmetric about $x=\delta$, while for $\beta\neq 0$

$$\lim_{x\to\infty}=\frac{P(-x)}{1-P(x)}=\frac{1+\beta}{1-\beta}.$$

γ is associated with the dispersion of the distribution; when $\alpha=2$ we see that γ equals twice the variance of the distribution. δ is a centering or location parameter; for $\alpha>1$ the location parameter is also the mean.

Only three analytic closed-form expressions for stable probability-density functions are known at present; the Gaussian ($\alpha=2$), the Cauchy ($\alpha=1, \beta=0$), and the Pearson $V(\alpha=\frac{1}{2}, \beta=-1)$. Otherwise, one must resort to series expansions [14].

References

[1] N. Wiener, *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. New York: Wiley, 1949.

[2] A. M. Yaglom, *An Introduction to the Theory of Stationary Random Functions*. New York: Dover, 1973.

[3] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, pp. 35–45, 1960.

[4] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *J. Basic Eng.*, vol. 83, pp. 95–108, 1961.

[5] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic, 1970.

[6] E. Wong, *Stochastic Processes in Information and Dynamical Systems*. New York: McGraw-Hill, 1971.

[7] B. W. Stuck and B. Kleiner, "A statistical analysis of telephone noise," *Bell Syst. Tech. J.*, vol. 53, pp. 1263–1320, 1974.

[8] J. E. Evans, "Preliminary analysis of ELF noise," M.I.T. Lincoln Laboratory Tech. Note 1969-18, DDC AD-691814, Mar. 26, 1969.

[9] P. Cootner, *The Random Character of Stock Market Prices*. Cambridge, MA: M.I.T. Press, 1964.

[10] E. Fama, "The behavior of stock market prices," *J. Business*, vol. 38, pp. 34–105, 1965.

[11] A. V. Skorokhod, "On the differentiability of measures which correspond to stochastic processes. I. Processes with independent increments," *Theory Prob. Appl.*, vol. 2, pp. 407–432, 1957.

[12] V. J. Paulauskas, "Some remarks on multivariate stable distributions," *J. Multivariate Anal.*, vol. 6, pp. 356–369, 1976.

[13] C. J. Masreliez and R. D. Martin, "Robust Bayesian estimation for the linear model and robustifying the Kalman filter, *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 361–372, 1977.

[14] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. II, 2nd ed. New York: Wiley, 1971, pp. 169–176 and pp. 574–583.

[15] B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*, rev. ed. Reading, MA: Addison-Wesley, 1968, ch. 3, 4, 6, 7.

Comments on "The Multivariable Servomechanism Problem from the Input–Output Viewpoint"

F. M. CALLIER

***Abstract*—The purpose of this correspondence is to point out that some results of the above paper[1] are not new and can be viewed as special cases of more general results.**

In the above paper[1] some general results were presented, and B. Francis deserves congratulations for their simplicity and generality, as well as the elegance of his derivations. Unfortunately, his Theorems 1A and 1B are special cases of results submitted in 1974 and published in 1975 [1] and 1976 [2]. Furthermore Francis' expression (6) is one of four possible expressions [1], [2]. Theorem 1B was also covered in a review paper devoted to lumped systems [3]. For the more general case of arbitrary interconnections, corresponding stability results are also known [4], [5, theorem II and corollary II.1].

References

[1] F. M. Callier and C. A. Desoer, "A stability theorem concerning open-loop unstable feedback systems with dynamical feedbacks," presented at the 6th IFAC World Congr. 1975, Paper 44.2.

[2] ——, "Open-loop unstable convolution feedback systems with dynamical feedbacks," *Automatica*, vol. 12, pp. 507–518, Sept. 1976.

[3] C. A. Desoer and W. S. Chan, "The feedback interconnection of lumped linear time-invariant systems," *J. Franklin Inst.*, vol. 300, pp. 335–351, Nov.–Dec. 1975.

[4] F. M. Callier, W. S. Chan, and C. A. Desoer, "Input-output stability theory of interconnected systems using decomposition techniques," *IEEE Trans. Circuits Syst.*, vol. CAS-23, pp. 714–729, Dec. 1976.

[5] ——, "Input-output stability of interconnected systems using decompositions: An improved formulation," Electron. Res. Lab., Univ. Calif., Berkeley, CA, Memo. M77/29, May 3, 1977. Also *IEEE Trans. Automat. Contr.*, vol. AC-23, Apr. 1978.

Manuscript received July 24, 1977.

The author is with the Department of Mathematics, University of Namur, Namur, Belgium.

[1] B. Francis, *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 322–328, June 1977.

Comments on Finding the Generic Rank of a Structural Matrix

MANFRED MORARI AND GEORGE STEPHANOPOULOS

The appendix of Shields and Pearson's paper[1] suggests an algorithm to determine the generic rank of a structural matrix. Consider the following example:

$$\begin{bmatrix}x & x & 0 & 0 & 0\\ 0 & 0 & x & x & x\\ 0 & 0 & x & x & x\\ 0 & 0 & x & x & x\\ 0 & 0 & x & x & x\end{bmatrix}.$$

Manuscript received September 27, 1977.

M. Morari was with the Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455. He is now with the Department of Chemical Engineering, University of Wisconsin, Madison, WI 53706.

G. Stephanopoulos is with the Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455.

[1] R. W. Shields and J. B. Pearson, *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 203–212, Apr. 1976.

0018-9286/78/0600-0509$00.75