vector is $\theta \in \Theta \triangleq \{H_i\}$, and subscripts are required in (10). Since specific Θ_i parameterizations may better represent various classes, improved classification may be possible, but at the cost of additional complexity.

Based on the representation of f by θ, it is evident that relations exist between the various feature sets given above, as well as between additional feature sets that may be constructed. The complete analytical development of these inherently nonlinear relations is often quite difficult and they are only indicated here. Using the Volterra representation in (5) and white noise or pseudorandom signals, crosscorrelation relations may be developed to obtain the Volterra kernels, and the truncated model (6) may similarly be used to relate the $\{a_{r_1 \cdots r_{N_v}}\}$ to crosscorrelation terms. Using the expansion in Hermite polynomials, the coefficients $\{a'_{r_1 \cdots r_n}\}$ may be determined by crosscorrelation of $y(j)$ with the outputs of the various polynomial $\phi_{r_k}(\cdot)$ similar to the determination of the Wiener coefficients in the continuous-time case [5]. Such crosscorrelations were used in [1] and [2] for classification of nonlinearity.

VI. Additional Considerations

Classification Error

The classification error relations presented in [1] and [2] will be briefly reviewed here due to space limitation. If $(\Omega, \beta(\Omega), \mu_\Omega)$ represents the probability space, the probability of error in classification depends on how the probability mass is distributed over $\beta(\Omega)$ by the measure μ_Ω. However, since most applications involve consideration of density functions related to the pattern vector, it is more convenient to consider the related space $(\mathcal{V}, \beta(\mathcal{V}), \mu_V)$ with $\mathcal{V} = R^n$ where n is the dimension of v. With appropriate assumptions on μ, the mixture density function description $p(v) = \Sigma^M p(v|\omega_i) \Pr[\omega_i]$ is possible and may be used to write the well-known Bayes error expressions [11]. It should be noted that error expressions based on μ_Ω depend only on the nonlinear classes ω_i, while those based on μ must also include effects of truncation and estimation error. Further discussion of error rates and experimental studies to approximate Bayes error rates may be found in [2], [1], respectively.

Other Approaches to Classification

Feature selection using system representations discussed in [5] are based on the information that is required to implicitly characterize an unknown nonlinear system and lead to features that are easily estimated for use in statistical classification methods. However, in some applications alternate approaches to classification may be more suitable such as behavior exhibited in phase-plane portraits, for which different features may be found useful. These and other possibilities remain to be considered.

If sufficient *a priori* information is available about the $\{\omega_i\}$ to be considered, parameterizations more suitable to a specified class may be employed as the basis for classification. In particular, if the form of the $f(\cdot,\cdot,\omega_i)$ is suitably restricted, more specialized means of parameter estimation may be used such as maximum *a posteriori* estimation [12]. Results of these hypothesis conditional estimates are subsequently used as a basis for classification.

Conclusions

For systems with nonlinear structure, a formulation of the classification problem has been given a decision-theoretic format which leads to a pattern recognition solution [2]. While it is difficult to implement the theoretical solutions in [4] without making simplifying assumptions, the formulation provides for systematically considering various aspects of the overall problem. Some analogies have been noted with other problems, both from a mathematical and applications point of view, and other examples involving waveform classification may be cited.

Attention has been restricted to a single input-single output formulation for simplicity, and it is evident that as the system order n_x and the allowed nonlinear forms increase in number, classification becomes more difficult. A more realistic approach is to assume that all the states are measureable, possibly with additive noise, and formulate a pattern vector making use of the additional information.

The problem of feature selection has been considered from two main viewpoints, both of which of necessity involve *implicit* characterizations of systems. Experimental results in [1] suggest that theoretical Bayes error rates for classification based on truncated moments are quite low. Classification based on crosscorrelations in [1], [2] also gave low error rates, and although the crosscorrelations can be related to the θ parameterizations of Section III, experimental studies using actual estimates $\hat{\theta} = v$ or $\hat{\theta}(H_i) = v$ may also be considered.

References

[1] G. N. Saridis and R. F. Hofstadter, "A pattern recognition approach to the classification of nonlinear systems," *IEEE Trans. Syst., Man, Cybern.*, vol. 4, pp. 362–371, July 1974.

[2] ——, "A pattern recognition approach to the classification of nonlinear stochastic systems," Purdue Univ., West Lafayette, IN., Tech. Rep. TR-EE 74-13, Apr. 1974.

[3] T. T. Soong, *Random Differential Equations in Science and Engineering*. New York: Academic, 1973.

[4] P. Alper, "A consideration of the discrete volterra series," *IEEE Trans. Automat. Contr.*, vol. 10, pp. 322–327, 1965.

[5] N. Wiener, *Nonlinear Problems in Random Theory*. Cambridge, MA: Mass. Inst. Technol. Press, 1958.

[6] D. Middleton, *Statistical Communication Theory*. New York: McGraw-Hill, 1960.

[7] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, vol. 1. New York: Wiley Interscience, 1953.

[8] P. Kuznetsov, R. L. Stratonovich, and V. I. Tikhonov, "Quasi-moment functions in the theory of random processes," *Theory Prob. Appl.*, vol. 5, pp. 80–97, Jan. 1960.

[9] J. R. Fisher, "Optimal nonlinear filtering," in *Advances in Control Systems*, vol. 5, C. T. Leondes, Ed. New York: Academic, 1967, pp. 197–300.

[10] R. Stratonovich, *Topics in the Theory of Random Noise*. New York: Gordon and Breach, 1963.

[11] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley Interscience, 1973.

[12] A. P. Sage and J. L. Melsa, *System Identification*. New York: Academic, 1971.

Asymptotic Root Loci of Multivariable Linear Optimal Regulators

HUIBERT KWAKERNAAK

***Abstract*—The asymptotic loci of the closed-loop poles of the multivariable time-invariant linear regulator are considered as the weight on the input in the criterion approaches zero. It is proved that those poles that go to infinity group into several Butterworth configurations of different orders and with different radii. It is furthermore shown that the first-order patterns are easily explicitly determined. Some examples illustrate the results.**

I. Introduction

The purpose of this short paper is to settle a question about the asymptotic loci of the closed-loop poles of multivariable optimal linear regulators with quadratic criteria, as the weight on the input goes to zero. This problem has been discussed by Chang [1] and Kalman [2] for the single-input case, and for the multiinput case by Tyler and Tuteur [3], Kwakernaak and Sivan [4], [5], and, in a recent book, by Wonham [6]. The asymptotic behavior of the closed-loop poles is of interest because it gives an indication of the type of dynamic response that the optimal regulator may be anticipated to exhibit.

In [5, theorem 3.12, pp. 288–289] it is claimed, but not proved, that the far-away closed-loop poles group into several Butterworth configurations of different orders and with different radii. An example is offered in support of this claim. Wonham [6, theorem 13.2, p. 317] considers the problem under certain restrictive conditions, which make it a special case of the problem of [5], and states that under these conditions no Butterworth patterns, or combinations of Butterworth patterns, occur.

Manuscript received February 14, 1975; revised January 14, 1976. Paper recommended by D. L. Kleinman, Past Chairman of the IEEE S-CS Optimal Systems Committee.

The author is with the Department of Applied Mathematics, Twente University of Technology, Enschede, The Netherlands.

It will be proved in this short paper that the claim that the far-away closed-loop poles group into several Butterworth patterns is correct. To this end, in Section II first some results are stated concerning higher order root loci. In Section III the principal results are obtained, while Section IV deals with the special case considered by Wonham [6].

For a discussion concerning the consequences of the asymptotic behavior of the closed-loop poles for the expected regulator response, and in particular concerning the fact that some of the closed-loop poles approach the system zeros or their mirror images, we refer to [5, sec. 3.8].

II. Asymptotic Properties of Higher Order Root Loci

The results detailed in this section may to a large extent be inferred from Rosenau [7]. We shall consider the asymptotic behavior of the roots of a polynomial, whose coefficients are themselves polynomials in a complex parameter σ, as $|\sigma|\rightarrow\infty$. The polynomial is represented in the form

$$p(\lambda;\sigma)=\sum_{j=0}^{n}\left(\sum_{k=0}^{k_j}\alpha_{jk}\sigma^{k}\right)\lambda^{j} \tag{1}$$

where n and k_j, $j=0,1,\cdots,n$ are nonnegative integers, and where the complex coefficients α_{jk} are such that $\alpha_{jk_n}\neq 0$, and $\alpha_{jk_j}\neq 0$ for all $j\in\{0,1,\cdots,n-1\}$ such that $k_j>0$.

To determine the asymptotic behavior of the roots of p as $|\sigma|\rightarrow\infty$, we first define positive real numbers κ_p and k_p^*, both for $p=0,1,\cdots,r$, with r the largest integer that is found, such that 1) $0=\kappa_0<\kappa_1<\cdots<\kappa_r$; 2) $k_0^*=\max(k_0,k_1,\cdots,k_n)$; 3) for $p\in\{1,2,\cdots,r\}$, κ_p and k_p^* satisfy the inequality $j\kappa_p+k_j\leqslant k_p^*$ for all $j\in\{0,1,\cdots,n\}$, while $j\kappa_p+k_j=k_p^*$ for at least two different values of j in the set $\{0,1,\cdots,n\}$. Thus, the numbers r, κ_p and k_p^*, $p=0,1,\cdots,r$, are entirely determined by the powers k_j, $j\in\{0,1,\cdots,n\}$. Fig. 1 illustrates how these numbers may be obtained. For each $p\in\{1,2,\cdots,r\}$, κ_p and k_p^* are chosen such that $k_j\leqslant k_p^*-j\kappa_p$ for all j and such that $k_j=k_p^*-j\kappa_p$ for at least two different values of j.

Lemma 1: The numbers κ_p, $p\in\{0,1,\cdots,r\}$ assume rational values only, and $0\leqslant r\leqslant n$.

This lemma is easily proved. Now, in order to determine the asymptotic behavior of the roots of (1) as $|\sigma|\rightarrow\infty$ we consider the polynomial obtained from (1) by only retaining of each coefficient the term with the highest power in σ:

$$q(\lambda;\sigma)=\sum_{j=0}^{n}\alpha_{jk_j}\sigma^{k_j}\lambda^{j}. \tag{2}$$

Substitution of $\lambda=z\sigma^{\kappa_p}$ yields

$$\sum_{j=0}^{n}\alpha_{jk_j}\sigma^{k_j+j\kappa_p}z^{j}, \tag{3}$$

which is a polynomial in z. Since $k_j+j\kappa_p\leqslant k_p^*$ for each $p\in\{0,1,\cdots,r\}$, division of (3) by $\sigma^{k_j+j\kappa_p}$ shows that those roots (in z) of (3) that remain finite as $|\sigma|\rightarrow\infty$ are the roots of the polynomial $\phi_p(z)$, where

$$\phi_p(z)=\sum_{j\in K_p}\alpha_{jk_j}z^{j} \tag{4}$$

with K_p the integer set $K_p=\{j: k_j+j\kappa_p=k_p^*\}$.

In the sequel we shall retain all roots of ϕ_0, which will be denoted as z_{0k}, $k\in\{1,2,\cdots,n_0\}$, and the *nonzero* roots of ϕ_p, $p\in\{1,2,\cdots,r\}$, which will be denoted as z_{pk}, $k\in\{1,2,\cdots,n_p\}$. If ϕ_0 has no roots we shall set $n_0=0$. For $p\neq 0$, ϕ_p has at least one nonzero root, so that $n_p\geqslant 1$ for $p\neq 0$. The main result of this section is the following.

Lemma 2: As $|\sigma|\rightarrow\infty$, the n roots of (1) asymptotically behave as $z_{pk}\sigma^{\kappa_p}$, $k\in\{1,2,\cdots,n_p\}$, $p\in\{0,1,\cdots,r\}$.

Since $\kappa_0=0$, the lemma implies that n_0 of the roots asymptotically approach fixed positions in the complex plane. The remaining $n-n_0$ roots approach infinity according to different positive rational powers of σ.

The arguments leading up to Lemma 2 make the result more or less

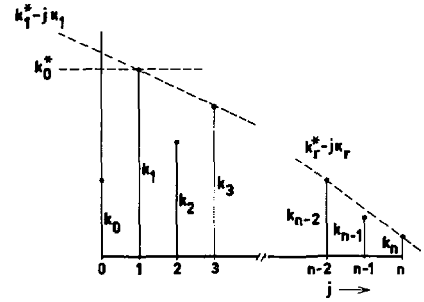

Fig. 1. Determination of the numbers κ_p and k_p^*, $p=0,1,\cdots,r$.

plausible. A more rigorous proof uses Rouché's theorem (compare [6]). Here we shall only verify that the asymptotic behavior of all n roots of (1) has been identified. The degree of ϕ_0 is of course n_0. Inspection of Fig. 1 shows that the lowest power of z occurring in ϕ_1 is n_0; since ϕ_1 has n_1 nonzero roots, the highest power of z occurring in ϕ_1 is n_0+n_1. Accordingly, the lowest power of z occurring in ϕ_2 is n_0+n_1, and as a result the highest power is $n_0+n_1+n_2$. Continuing like this it follows that the highest power of z occuring in ϕ_r is $n_0+n_1+\cdots+n_r$, which equals n. This proves that in Lemma 2 the asymptotic behavior of all n roots is identified.

Example 1: As an application, we consider the longitudinal motion of an airplane, as discussed in [4, example 3.21, pp. 293–297]. The system is described by the state differential equation

$$\dot{x}(t)=\begin{bmatrix}-0.1580 & 0.02633 & -9.810 & 0\\ -0.1571 & -1.030 & 0 & 120.5\\ 0 & 0 & 0 & 1\\ 0.0005274 & -0.01652 & 0 & -1.466\end{bmatrix}x(t)+\begin{bmatrix}0.0006056 & 0\\ 0 & -9.496\\ 0 & 0\\ 0 & -5.565\end{bmatrix}u(t) \tag{5}$$

while the output is given by

$$z(t)=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}x(t). \tag{6}$$

For a discussion of the physical significance of state, input and output variables we refer to [4]. Assume that linear output feedback of the form $u(t)=-\sigma Fz(t)$ is applied, with F the matrix

$$F=\begin{bmatrix}50 & 0\\ 0 & 1\end{bmatrix}. \tag{7}$$

We shall study the asymptotic pole locations of the closed-loop system as the scalar gain factor σ goes to infinity. The closed-loop characteristic polynomial is given by $p(s,\sigma)=\det\,(sI-A+\sigma BFC)$. Evaluation of this characteristic polynomial yields, retaining the leading terms only,

$$q(s,\sigma)=s^4-0.03028\sigma s^3+5.489\sigma s^2-0.1685\sigma^2 s-0.1688\sigma^2. \tag{8}$$

Clearly, the closed-loop system is not asymptotically stable for large σ. It is easily found that $r=2$, and that $\kappa_0=0,\kappa_1=\frac{1}{2}$, and $\kappa_2=1$. Furthermore, the integer sets K_0, K_1, and K_2 are given by $K_0=\{0,1\}$, $K_1=\{1,3\}$, and $K_2=\{3,4\}$. It follows that the polynomials ϕ_0, ϕ_1, and ϕ_2 are $\phi_0(z)=-0.1688-0.1685z$, $\phi_1(z)=-0.1685z-0.03028z^3$, $\phi_2(z)=-0.03028z^3+z^4$, with the result that $z_{01}=-1.002$, $z_{11}=2.359i$, $z_{12}=-2.359i$, and $z_{21}=0.03028$. This means that one closed-loop pole asymptotically approaches the fixed location -1.002 (this is in fact the zero of the system [4]), while a pair of roots asymptotically behaves as $\pm 2.359i\sigma^{1/2}$, and the remaining root asymptotically behaves as 0.03028σ.

III. Asymptotic Root Loci of Linear Optimal Regulators

In this section we consider the stabilizable and detectable time-invariant linear system

$$\dot{x}(t)=Ax(t)+Bu(t) \tag{9}$$

$$z(t)=Dx(t) \tag{10}$$

and the quadratic criterion

$$\int_0^\infty [z^T(t)Qz(t)+u^T(t)Ru(t)]\,dt. \tag{11}$$

Q and R are positive-definite symmetric matrices. In the sequel we shall denote $\dim(x)=n$ and $\dim(u)=m$. It is well known (see, e.g., [5] or [6]) that the criterion (11) is minimized if the input is chosen as $u(t)=-Fx(t)$, where the gain matrix F is given by $F=R^{-1}B^TP$, with P the unique nonnegative-definite solution of the algebraic Riccati equation

$$0=D^TQD-PBR^{-1}B^TP+A^TP+PA. \tag{12}$$

It is moreover known (see, e.g., [5, sec. 3.8.1]) that the closed-loop characteristic polynomial $\phi_c(s)=\det(sI-A+BF)$ satisfies the relation

$$\phi_c(s)\phi_c(-s)=\phi(s)\phi(-s)\det[I+R^{-1}H^T(-s)QH(s)] \tag{13}$$

where $\phi(s)=\det(sI-A)$ is the open-loop characteristic polynomial, and $H(s)=D(sI-A)^{-1}B$ the open-loop transfer matrix of the system.

We shall study the asymptotic behavior of the closed-loop poles of the optimal regulator, i.e., the roots of $\phi_c(s)$, when R in (11) is replaced with ρR, and we let $\rho\downarrow 0$. In this case we have to replace (13) with

$$\phi_c(s)\phi_c(-s)=\phi(s)\phi(-s)\det\left[I+\frac{1}{\rho}R^{-1}H^T(-s)QH(s)\right]. \tag{14}$$

We consider the right-hand side of (14) as $\rho\downarrow 0$. The roots of $\phi_c(s)$ that remain finite are easily found; they are the left-half plane roots of the polynomial

$$\phi(s)\phi(-s)\det[H^T(-s)QH(s)]. \tag{15}$$

Here we use the fact that the roots of $\phi_c(s)$ are always in the left-half complex plane since the closed-loop optimal regulator is asymptotically stable. It will be assumed throughout that (15) is not identically zero, and has degree $2q\geq 0$. As a result, q of the closed-loop poles remain finite as $\rho\downarrow 0$.

To study the behavior of the roots that do not remain finite, we first introduce a few new quantities. By replacing R^{-1} in (14) with $R^{-\frac{1}{2}}R^{-\frac{1}{2}}$, and using the fact that for any two matrices P and Q of compatible dimensions $\det(I+PQ)=\det(I+QP)$, we rewrite (14) as

$$\phi_c(s)\phi_c(-s)=\phi(s)\phi(-s)\det\left[I+\frac{1}{\rho}R^{-\frac{1}{2}}H^T(-s)QH(s)R^{-\frac{1}{2}}\right] \tag{16}$$

and define

$$M(s)=\phi(s)\phi(-s)R^{-\frac{1}{2}}H^T(-s)QH(s)R^{-\frac{1}{2}}. \tag{17}$$

It is noted that the polynomial matrix M is para-Hermitian, i.e., $M(-s)=M^T(s)$. Now we write

$$\begin{aligned}\phi_c(s)\phi_c(-s)&=\phi(s)\phi(-s)\det\left[I+\frac{1}{\rho}\frac{M(s)}{\phi(s)\phi(-s)}\right]\\&=\phi(s)\phi(-s)\prod_{j=1}^{m}\left[1+\frac{1}{\rho}\frac{\lambda_j(s)}{\phi(s)\phi(-s)}\right]\end{aligned} \tag{18}$$

where the $\lambda_j(s), j=1,2,\cdots,m$, are the eigenvalues of $M(s)$ (regarded as functions of s).

Lemma 3: As $|s|\to\infty$, the eigenvalues of $M(s)$ asymptotically behave as $u_{pk}(-1)^{\kappa_p}s^{2\kappa_p}$, with $k\in\{1,2,\cdots,n_p\}$ and $p\in\{0,1,\cdots,r\}$. Here u_{pk} is a real number with $u_{pk}>0$ for $k\in\{1,2,\cdots,n_p\}$ and $p\in\{1,2,\cdots,r\}$, and $u_{0k}\geq 0$ for $k\in\{1,2,\cdots,n_0\}$. For each $p\in\{0,1,\cdots,r\}$ the number κ_p is an integer satisfying $0=\kappa_0<\kappa_1<\cdots<\kappa_r\leq n-1$. Furthermore, r is a nonnegative integer, while $n_0,n_1,\cdots,n_r$ are nonnegative integers with $n_0\geq 0, n_p>0$ for $p\in\{1,2,\cdots,r\}$, and $n_0+n_1+\cdots+n_r=m$.

The proof of this lemma is given in Appendix A. It heavily relies on Section II of this short paper. The lemma may be applied in identifying the asymptotic behavior of the roots of the right-hand side of (18). We replace each polynomial or other function in (18) with its leading term or asymptotic behavior as $|s|\to\infty$. Thus, substitution of $u_{pk}(-1)^{\kappa_p}s^{2\kappa_p}$ for $\lambda_j(s)$, and of $(-1)^n s^{2n}$ for $\phi(s)\phi(-s)$ shows that there exists a group of roots of (18) asymptotically satisfying

$$s^{2(n-\kappa_p)}+(-1)^{\kappa_p}\frac{u_{pk}}{\rho}=0 \tag{19}$$

as $\rho\downarrow 0$. Those roots of (19) that lie in the left-half complex plane together form a Butterworth configuration (see, e.g., [8]) of order $n-\kappa_p$, with radius $(u_{pk}/\rho)^{1/2(n-\kappa_p)}$. We thus have the following result.

Theorem 1: As $\rho\downarrow 0$, q of the closed-loop regulator poles approach finite locations in the complex plane. The remaining $n-q$ closed-loop poles approach infinity and asymptotically group into a number of Butterworth configurations, of different orders and different radii.

In Appendix B it is verified that in the argument preceding the theorem all $n-q$ closed-loop poles that approach infinity are identified.

Example 2: We again consider the airplane of Example 1, and suppose that we wish to control the system such that a criterion of the form (11) minimized, where

$$Q=\begin{bmatrix}0.02 & 0\\ 0 & 50\end{bmatrix},\quad R=\rho\begin{bmatrix}0.0004 & 0\\ 0 & 2500\end{bmatrix}. \tag{20}$$

The asymptotic behavior of the closed-loop poles of the resulting regulator as $\rho\downarrow 0$ has been determined numerically in [5, example 3.21]. We shall verify these results here. Let us write $\phi(s)$ for the open-loop characteristic polynomial, $Q=\operatorname{diag}(q_1,q_2)$, $R=\rho\operatorname{diag}(r_1,r_2)$, and $H(s)=N(s)/\phi(s)$, where the polynomial matrix N has entries $h_{ij}(s), i,j=1,2$. Finally, if $p(s)$ is an arbitrary polynomial, we shall write $\bar{p}(s)=p(-s)$. With this notation is not difficult to derive from (13) that

$$\begin{aligned}\phi_c\bar{\phi}_c=\phi\bar{\phi}+\frac{1}{\rho}\Bigg[\frac{1}{r_1}&\left(q_1h_{11}\bar{h}_{11}+q_2h_{21}\bar{h}_{21}\right)\\&+\frac{1}{r_2}\left(q_1h_{12}\bar{h}_{12}+q_2h_{22}\bar{h}_{22}\right)\Bigg]+\left(\frac{1}{\rho}\right)^2\frac{q_1q_2}{r_1r_2}\psi\bar{\psi}.\end{aligned} \tag{21}$$

For brevity we have omitted the argument s. Furthermore, $\psi(s)$ is the numerator polynomial of the transfer matrix H, i.e., $\psi(s)/\phi(s)=\det[H(s)]$. After calculation of the polynomials h_{ij} and ψ, the right-hand side of (21) can be evaluated. Including the leading terms only, we find that the asymptotic behavior of the roots of $\phi_c(s)\phi_c(-s)$ is determined by the roots of the polynomial $q(\lambda,1/\rho)$, with $\lambda=s^2$, and

$$\begin{aligned}q\left(\lambda,\frac{1}{\rho}\right)=\lambda^4&-1.834\times10^{-5}\left(\frac{1}{\rho}\right)\lambda^3+0.6192\left(\frac{1}{\rho}\right)\lambda^2\\&-1.1357\times10^{-5}\left(\frac{1}{\rho}\right)^2\lambda+1.1402\times10^{-5}\left(\frac{1}{\rho}\right)^2.\end{aligned} \tag{22}$$

Using the method of analysis of Section II, it is easily found that as $\rho\downarrow 0$ the roots of q asymptotically behave as 1.004, $\pm 0.7869 i\rho^{-1/2}$, and $1.834\times10^{-5}\rho^{-1}$, respectively. Taking the square roots, and selecting the left-half plane values, it follows that one of the closed-loop poles approaches the fixed location -1.002, that furthermore a pair of closed-loop poles asymptotically behaves according to the second-order Butterworth pattern $0.6273(-1\pm i)\rho^{-1/4}$, while the remaining closed-loop pole traces the first-order Butterworth pattern $-0.004283\rho^{-1/2}$. These are exactly the patterns that were obtained in [5, example 3.21].

IV. Determination of the First-Order Butterworth Patterns

In this section it is shown how to determine the asymptotic first-order Butterworth patterns. We write

$$M(s)=\sum_{i=0}^{2(n-1)}N_is^i \tag{23}$$

where $N_i, i=0,1,\cdots,2(n-1)$ are constant matrices. Inspection of

$$\det\left[\lambda I - \sum_{i=0}^{2(n-1)} N_i s^i\right] = 0 \tag{24}$$

shows that the eigenvalues of $M(s)$ that go fastest to ∞ as $|s| \to \infty$ are obtained by determining the nonzero roots (in λ) of

$$\det[\lambda - N_{2(n-1)} s^{2(n-1)}]. \tag{25}$$

It is easily verified (e.g., by using Leverrier's expansion) that $N_{2(n-1)} = (-1)^{n-1} R^{-\frac{1}{2}} (DB)^T Q DBR^{-\frac{1}{2}}$. Let the nonzero (positive) eigenvalues of the nonnegative-definite symmetric matrix $R^{-\frac{1}{2}}(DB)^T QDBR^{-\frac{1}{2}}$ be given by μ_i, $i = 1,2,\cdots,m'$. Then (25) has as nonzero roots $(-1)^{n-1}\mu_i s^{2(n-1)}$, $i = 1,2,\cdots,m'$, which are the asymptotic roots of $M(s)$ corresponding to $\kappa_r = n-1$. The resulting faraway regulator poles are from (19) the left-half plane roots of

$$s^2 - \frac{\mu_i}{\rho} = 0, \qquad i = 1,2,\cdots,m'. \tag{26}$$

For each i a first-order Butterworth pattern is obtained, consisting of a single pole $-(\mu_i/\rho)^{\frac{1}{2}}$ on the negative real axis. The number of first-order patterns equals rank $(DBQ^{\frac{1}{2}})$.

If rank $(DBQ^{\frac{1}{2}}) = m$, which is the situation considered by Wonham [6], $m' = m$, and the first-order patterns are the only patterns found, since it follows from (15) that in this case $q = n - m$ and hence there are exactly m faraway poles.

Example 3: We use the result of this section to determine the first-order Butterworth pattern of Example 2 directly. It is easily found that

$$R^{-\frac{1}{2}}(DB)^T QDBR^{-\frac{1}{2}} = \begin{bmatrix} 1.834\times 10^{-5} & 0 \\ 0 & 0 \end{bmatrix} \tag{27}$$

which has a single nonzero eigenvalue $\mu_1 = 1.834\times 10^{-5}$. Correspondingly, the closed-loop system asymptotically has a single first-order Butterworth pattern, with the pole location $-0.004283\rho^{-1/2}$. This agrees with what was found in Example 2.

V. Conclusion

It has been proved that those closed-loop poles of the time-invariant multivariable optimal linear regulator that go to infinity as the weight on the input decreases asymptotically group into several Butterworth configurations. It has also been shown that the number and asymptotic radii of the first-order Butterworth patterns may be determined relatively easily. Similar methods to determine the higher order patterns seem not to be available at present.

Appendix A
Proof of Lemma 3

We study the asymptotic behavior of the eigenvalues of the matrix $M(s)$, given by (17). Since M is para-Hermitian, the characteristic polynomial $d(\lambda,s) = \det[\lambda I - M(s)]$ satisfies $d(\lambda,-s) = d(\lambda,s)$, so that we may write

$$d(\lambda,s) = \sum_{j=0}^{m}\left(\sum_{k=0}^{k_j} \alpha_{jk} s^{2k}\right)\lambda^j \tag{A1}$$

where the coefficients α_{jk} are real. From Section II we conclude (taking $\sigma = s^2$) that as $|s| \to \infty$ the roots of $d(\lambda,s)$ asymptotically behave as $z_{pk} s^{2\kappa_p}$, $k \in \{1,2,\cdots,n_p\}$, $p \in \{0,1,\cdots,r\}$. The κ_p are rational numbers satisfying $0 = \kappa_0 < \kappa_1 < \cdots < \kappa_r$. For each $p > 0$ and k the real number z_{pk} is nonzero. Furthermore, $0 \leq r \leq m$, while $n_p > 0$ for $p > 0$.

Let us consider the case that s is purely imaginary. Setting $s = i\omega$, with ω real, the asymptotic behavior of the roots as $\omega \to \pm\infty$ is given by $z_{pk}(-1)^{\kappa_p}\omega^{2\kappa_p}$. Now, for $s = i\omega$ the matrix $M(s)$ is nonnegative-definite Hermitian for each ω, which implies that its eigenvalues are real and nonnegative for each ω. Consequently, we cannot but conclude that for each p the number κ_p must be an integer, and $z_{pk}(-1)^{\kappa_p}$ nonnegative real. Returning to the case that s is not purely imaginary, and substituting $(-1)^{\kappa_p} z_{pk} = u_{pk}$, we conclude that the asymptotic behavior of the eigenvalues of $M(s)$ is $(-1)^{\kappa_p} u_{pk} s^{2\kappa_p}$, where the κ_p are integers satisfying $0 = \kappa_0 < \kappa_1 < \cdots < \kappa_r$, and where $u_{pk} > 0$ for $p > 0$, and $u_{0k} \geq 0$.

To complete the proof of Lemma 3 it remains to demonstrate that $\kappa_r \leq n-1$. This is trivially true if $r = 0$. Suppose that $r > 0$. It follows from Section II that the inequality $k_j + j\kappa_p \leq k_p^*$ is satisfied with equality for $j = n_0 + n_1 + \cdots + n_{p-1}$ and for $j = n_0 + n_1 + \cdots + n_p$. Substitution of these two values of j in $k_j + j\kappa_p$ and equating the results it follows that $n_p\kappa_p = k_{n_0+n_1+\cdots+n_{p-1}} - k_{n_0+n_1+\cdots+n_p}$. Taking $p = r$, and using the fact that $k_{n_0+n_1+\cdots+n_r} = k_m = 0$, it is seen that

$$\kappa_r = \frac{k_{n_0+n_1+\cdots+n_{r-1}}}{n_r}. \tag{A2}$$

Now, since the coefficient of λ^j in $d(\lambda;s) = \det[\lambda I - M(s)]$ is the sum of all $(m-j)\times(m-j)$ principal minors of $M(s)$ (Gantmacher [9, p. 70]), and $2k_j$ is the degree of this coefficient it follows from the fact that $M(s)$ has degree $2(n-1)$ or less, that $2k_j \leq 2(n-1)(m-j)$. Applying this to (A2) we obtain

$$\kappa_r = \frac{k_{n_0+n_1+\cdots+n_{r-1}}}{n_r} = \frac{k_{m-n_r}}{n_r} \leq \frac{(n-1)n_r}{n_r} = n-1 \tag{A3}$$

which terminates the proof of Lemma 3.

Appendix B
Identification of Closed-Loop Poles

In this Appendix it will be verified that in the argument preceding Theorem 1 all $n-q$ faraway closed-loop poles are identified. Since $u_{pk} > 0$ for $k \in \{1,2,\cdots,n_p\}$ and $p \in \{1,2,\cdots,r\}$, corresponding to each such u_{pk} we obtain n-κ_p asymptotic locations of closed-loop poles. Altogether,

$$\sum_{p=1}^{r} (n-\kappa_p) n_p \tag{B1}$$

closed-loop poles may thus be identified. In Appendix A we found that $n_p\kappa_p = k_{n_0+n_1+\cdots+n_{p-1}} - k_{n_0+n_1+\cdots+n_p}$ for $p \in \{1,2,\cdots,r\}$. It follows that

$$\sum_{p=1}^{r} n_p\kappa_p = k_{n_0} - k_{n_0+n_1+\cdots+n_r} = k_{n_0} - k_m = k_{n_0}. \tag{B2}$$

In addition,

$$\sum_{p=1}^{r} n_p = m - n_0. \tag{B3}$$

It follows from (B1), (B2) and (B3) that

$$\sum_{p=1}^{r} (n-\kappa_p) n_p = n(m-n_0) - k_{n_0}. \tag{B4}$$

The possibility that $u_{0k} = 0$ for one or several values of k has to be investigated. We consider the following cases.

1) $n_0 = 0$. Then $2k_{n_0} = 2k_0$ is the degree of $\det[M(s)]$, which is easily determined to be $2(nm-n+q)$. Consequently, $k_0 = nm-n+q$ and

$$\sum_{p=1}^{n} (n-\kappa_p) n_p = nm - (nm-n+q) = n-q \tag{B5}$$

which means that all closed-loop poles are found. There are no closed-loop poles corresponding to $\kappa_0 = 0$.

2) $n_0 \geq 1$, and $u_{0k} = 0$ for all $k \in \{1,2,\cdots,n_0\}$. In this case we have $k_{n_0} > k_j$ for $j \in \{1,2,\cdots,n_0-1\}$, and hence also $k_{n_0} > k_0 = nm-n-q$. This would imply

$$\sum_{p=1}^{r} (n-\kappa_p) n_p < 0 \tag{B6}$$

from which it follows that $r=0$. As a consequence, $\det[\lambda I - M(s)]$ has degree n_0 in λ, which means that $n_0 = m$, and hence $k_{n_0} = k_m = 0$. This contradicts the conclusion that $k_{n_0} > k_0$, which means that the case cannot occur.

3) $n_0 \geq 1$, and $u_{0k} > 0$ for one value of $k \in \{1,2,\cdots,n_0\}$. Corresponding to this u_{0k} we obtain an nth-order Butterworth configuration, yielding n closed-loop poles at once. Consequently $q=0$ and $r=0$. Since $r=0$, $\det[\lambda I - M(s)]$ is of degree n_0 in λ, so that $n_0 = m$, which in turn implies $k_{n_0} = k_m = 0$. Since u_{0k} can be nonzero for one value of k only (otherwise several nth-order patterns of closed-loop poles would result, which is impossible), we evidently have $\det[\lambda I - M(s)] = -a\lambda^{m-1} + \lambda^m$, with $a > 0$. Unless $m=1$, this implies $\det[M(s)] = 0$, which is contrary to assumption. The case $m=1$, in which a single nth-order Butterworth pattern is obtained, corresponds to the single-input case where $H^T(-s)QH(s) = c/\phi(s)\phi(-s)$, with c a constant.

Summarizing, we have demonstrated that if $n_0 = 0$, one or several Butterworth patterns are obtained. The case $n_0 = 1$, which results in a single nth-order Butterworth pattern, only occurs when in the single-input case $H^T(-s)QH(s) = c/\phi(s)\phi(-s)$, with c a constant. The case $n_0 > 1$ does not occur.

References

[1] S. S. L. Chang, *Synthesis of Optimum Control Systems*. New York: McGraw-Hill, 1961.
[2] R. E. Kalman, "When is a linear control system optimal?" *Trans. ASME* (J. Basic Eng.), ser. D, vol. 86, pp. 51–60, 1964.
[3] J. S. Tyler and F. B. Tuteur, "The use of a quadratic performance index to design multivariable control systems," *IEEE Trans. Automat. Contr.*, vol. AC-11, pp. 84–92, Jan. 1966.
[4] H. Kwakernaak and R. Sivan, "Asymptotic pole locations of time-invariant linear optimal regulators," in *Proc. 3rd Annu. Southeastern Symp. System Theory*, Atlanta, GA, Apr. 5–6, 1971.
[5] —, *Linear Optimal Control Systems*. New York: Wiley-Interscience, 1972.
[6] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach* (lecture notes in economics and mathematical systems), vol. 101. Berlin: Springer, 1974.
[7] G. Rosenau, "Höhere Wurzelortskurven bei Mehrgrössensystemen," *IFAC Symp. Multivariable Systems, Preprints*, Oct. 7–8, 1968.
[8] L. Weinberg, *Network Analysis and Synthesis*. New York: McGraw-Hill, 1962.
[9] F. R. Gantmacher, *The Theory of Matrices*, vol. I. New York: Chelsea, 1959.

On Optimal and Suboptimal Actuator Selection Strategies

Y. VANBEVEREN AND M. R. GEVERS, MEMBER, IEEE

***Abstract*—This short paper studies a particular class of optimization problems dealing with the selection, at each instant of time, of one out of many actuators in order to obtain a determined result. A cost is associated with each actuator. The cost function is the integral of a weighted combination of the achieved accuracy on the state of the system and the control energy. The control energy term depends upon both the selected actuator and the magnitude of the applied control. The problem is to design an optimal actuator selection strategy. The analysis is limited to the class of linear deterministic systems with measurable states. A discrete approach is considered. The analytic solution to this optimization problem is given first. When the number of actuators and the number of stages in the time interval become large the optimal analytic solution requires a considerable combinatorial work; a suboptimal algorithm is then proposed to alleviate this defect.**

I. Introduction

The problem of selecting, at each instant of time, one out of many available actuators is presently untreated in the literature. There are, however, applications in which several different or incompatible actions can be applied on a process. Classes of examples are: problems with a bottleneck (such as hierarchical systems in which a single line is to transmit different effects having the same potentialities to the various subsystems), or problems with different zones for the control (e.g., a gearbox). In this last example the problem is both to select the best gear and to determine the pressure on the accelerator.

Some aspects of the dual problem on the optimal selection of sensors have been solved by Athans [1], Herring and Melsa [2], and Bensoussan [3].

Athans [1] has considered the determination of optimal costly measurement strategies in the case of finite-dimensional systems. At each instant during a time interval, one out of a finite number of sensors must be selected to minimize a payoff that depends on two terms: the accumulated observation cost and the prediction accuracy at final time. The accumulated prediction error cost is not considered.

Herring and Melsa [2] have generalized these results to allow the selection at each instant of time of the best combination of a finite number of sensors. The payoff depends on the observation cost as before, but also on the accuracy of prediction at each instant of the time interval considered.

Bensoussan [3] has extended Athans' results (but with different methods) to infinite-dimensional spaces in order to optimize the location of sensors in a distributed parameter system. He uses the same payoff as Athans. Aidarous, Gevers, and Installé have derived a numerically implementable algorithm for the optimal allocation of sensors [7] and actuators [8] in a distributed parameter system.

In this short paper the problem of designing an optimal actuator selection strategy is solved using the optimality principle. The cost function is not the dual of any of the measurement strategy problems mentioned above, since it includes an instantaneous cost depending upon both the chosen actuator and the control energy. The problem is stated in Section II, and the N-stage optimization problem is solved in Section III. Two criteria are presented for the *a priori* elimination of certain "bad" sequences. For the remaining sequences the solution depends on the initial state. For a long time interval (N large) or a large amount of actuators, the computational effort required to find the optimal actuator policy can become prohibitive. Therefore, a suboptimal algorithm has been developed that drastically reduces the computation time. This "forward-backward" algorithm is presented in Section IV. All the simulations performed so far show that the "forward-backward" algorithm is near optimal; some numerical results are given in Section V.

II. Problem Statement

Consider a time-invariant linear dynamic system

$$X(i+1) = AX(i) + BU(i) \tag{1}$$

where X is an $n \times 1$ state vector and U is an $mq \times 1$ control vector. A and B are $n \times n$ and $n \times mq$ matrices. B will be represented as follows:

$$B = [b_1 \; b_2 \; \cdots \; b_m]$$

where b_j is an $n \times q$ matrix corresponding to the jth actuator. m actuators are available, but only one actuator can be used at any given time. Therefore,

$$U(i) = \begin{bmatrix} u_1(i) \\ u_2(i) \\ \\ u_m(i) \end{bmatrix} \in U, \quad U \triangleq \left\{ \begin{bmatrix} u \\ 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ u \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \cdots, \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ u \end{bmatrix} \right\}, \qquad u \in R^q. \tag{2}$$

Hence, if the jth actuator is chosen at time i, $u_j(i)$ may take any real value, while $u_k(i) = 0$, $k \neq j$.

It is assumed that each pair $[A, b_j]$ is completely controllable, and that the state $X(i)$ is exactly measurable. The cost function to be minimized for a N-stage problem is

$$J_N = \sum_{i=0}^{N-1} [X'(i+1)QX(i+1) + U'(i)RU(i)] \tag{3}$$

Manuscript received March 13, 1975; revised January 14, 1976. Paper recommended by D. L. Kleinman, Chairman of the IEEE S-CS Optimal Systems Committee.

Y. Vanbeveren was with Universite Catholique de Louvain, Louvain-la-Neuve, Belgium. He is now with the Centre d'Automatique, Fontainebleau, France.
M. R. Gevers is with the Universite Catholique de Louvain, Louvain-la-Neuve, Belgium.