[3] D. W. Jorgensen, "Capital theory and investment behavior," *Amer. Econ. Rev.*, vol. 53, pp. 247–259, 1963.
[4] R. Muth, "A vintage model of the housing stock," Papers of the Regional Science Ass., 1972.
[5] C. P. Rydell, "Review of factors affecting maintenance and operating costs in public housing," Papers of the Regional Science Ass., 1971.
[6] J. L. Sweeney, "A commodity hierarchy model of the rental housing market," *J. Urban Econ.*, vol. 3 pp. 288–323, 1974.
[7] C. F. Curry and D. H. Gensch, "Feasibility of rent and tax incentives for renovation in older residential neighborhoods," *Management Sci.*, vol. 21, pp. 883–896, 1975.

Pravin Varaiya (M'68) was born in Bombay, India, in October 1940. He received the Ph.D. degree from the University of California, Berkeley, in 1966.

Since 1970 he has been a Professor of Electrical Engineering and Computer Sciences at the University of California, Berkeley. He is currently an Associate Editor of the *Journal of Optimal Theory and Applications.* His current interests are in the fields of stochastic control, large-scale systems, and urban economics.

Modern Wiener–Hopf Design of Optimal Controllers—Part II: The Multivariable Case

DANTE C. YOULA, FELLOW, IEEE, HAMID A. JABR, MEMBER, IEEE, AND JOSEPH J. BONGIORNO, JR., MEMBER, IEEE

***Abstract*—In many modern-day control problems encountered in the fluid, petroleum, power, gas and paper industries, cross coupling (interaction) between controlled and manipulated variables can be so severe that any attempt to employ single-loop controllers results in unacceptable performance. In all these situations, any workable control strategy must take into account the true multivariable nature of the plant and address itself directly to the design of a compatible multivariable controller. Any practical design technique must be able to cope with load disturbance, plant saturation, measurement noise, process lag, sensitivity and also incorporate suitable criteria delimiting transient behavior and steady-state performance. These difficulties, when compounded by the fact that many plants (such as chemical reactors) are inherently open-loop unstable have hindered the development of an inclusive frequency-domain analytic design methodology. However, a solution based on a least-square Wiener–Hopf minimization of an appropriately chosen cost functional is now available. The optimal controller obtained by this method guarantees an asymptotically stable and dynamical closed-loop configuration irrespective of whether or not the plant is proper, stable, or minimum-phase and also permits the stability margin of the optimal design to be ascertained in advance. The main purpose of this paper is to lay bare the physical assumptions underlying the choice of model and to present an explicit formula for the optimal controller.**

I. Introduction

In many modern-day control problems encountered in the fluid, petroleum, power, gas, and paper industries, cross coupling (interaction) between controlled and manipulated variables can be so severe that any attempt to employ single-loop controllers results in unacceptable performance. In all these situations, any workable control strategy must take into account the true multivariable nature of the plant and address itself directly to the design of a compatible multivariable controller. Any practical design technique must be able to cope with load disturbance, plant saturation, measurement noise, process lag, sensitivity and also incorporate suitable criteria delimiting transient behavior and steady-state performance. These difficulties, when compounded by the fact that many plants (such as chemical reactors) are inherently open-loop unstable have hindered the development of an inclusive frequency-domain analytic design methodology. However, these obstacles have been overcome and a solution based on a least-squares Wiener–Hopf minimization of an appropriately chosen cost functional E is now available. This solution, which is the natural culmination of earlier work [1]–[4], offers the following concrete accomplishments:

1) There are no restrictions on the plant transfer matrix. It can be rectangular, unstable, improper,[1] and nonminimum phase.

2) The design incorporates input noise, load disturbance, measurement noise, and feedforward compensation. The noise can be colored.

3) The optimal controller minimizing E is proper and guarantees a dynamical asymptotically stable closed-loop design possessing proper sensitivity matrices equal to the identity matrix at $s=\infty$.

Manuscript received June 5, 1975; revised February 11, 1976. Paper recommended by J. B. Pearson, Chairman of the IEEE S-CS Linear Systems Committee. This work was supported by the National Science Foundation under Grant ENG 74-13054 and is taken in part from a Ph.D. dissertation submitted by H. A. Jabr to the Faculty of the Polytechnic Institute of New York.

D. C. Youla and J. J. Bongiorno, Jr., are with the Department of Electrical Engineering and Electrophysics, Polytechnic Institute of New York, Long Island Center, Farmingdale, NY 11735.
H. A. Jabr was with the Department of Electrical Engineering and Electrophysics, Polytechnic Institute of New York, Long Island Center, Farmingdale, NY 11735. He is now with the University of Jordan, Amman, Jordan.

[1] A transfer matrix $A(s)$ is proper if $A(\infty)$ is finite and strictly proper if $A(\infty)=O$, the zero matrix. Otherwise it is improper.

4) The loop can track ramp-type inputs and recover from step-type disturbances of the correct order with zero steady-state error.

5) Transient response (system accuracy) can be traded off against linear operation.

6) The stability margin of the optimal design is ascertainable in advance.

7) The sensor transfer matrices are absorbed directly into the cost and various delays can be simulated by suitable preequalization.

The primary purpose of this paper is to lay bare the physical assumptions underlying the choice of model and to derive an explicit expression for the optimal controller. To achieve this objective it is first necessary to solve several difficult intermediate problems of the "model matching" variety and, for the sake of continuity and clarity, some of the more involved details have been relegated to two Appendixes. Finally, to help place the contributions of the present work in perspective we offer a comparison with the linear quadratic Gaussian (LQG) approach [17]. It is pointed out that the problem addressed and solved by LQG is quite different from the one considered in this paper. Nevertheless, there is a common class of problems that can be treated by both methods and for any such problem the optimal controller is the same. However, we show by actual example that forcing the optimal controller to be realized via the Kalman structure is not always possible. This limitation is inherent in LQG, but not in ours.

II. The Model

We focus our attention exclusively on the design of optimal controllers for multi-input–output finite-dimensional linear time-invariant plants imbedded in a multivariable single-loop configuration of generic type shown in Fig. 1.[2] Suppose $y_d(s)$, the *desired* closed-loop output is related to $\boldsymbol{u}_i(s)$, the *actual* input set-point signal in the linear fashion

$$y_d(s)=T_d(s)\boldsymbol{u}_i(s) \tag{1}$$

via the *ideal* transfer matrix $T_d(s)$. The *prefilter* $W(s)$ is selected in advance, but once chosen,[3]

$$\boldsymbol{u}=W(\boldsymbol{u}_i+\boldsymbol{n}) \tag{2}$$

[2]To avoid proliferating symbols, all quantities are Laplace transforms, deterministic or otherwise. All stochastic processes are either zero-mean second-order stationary or shape-deterministic or a sum of both with rational spectral densities. For example, η/s, η a random variable, is the transform of a random step with spectral density

$$\langle(\eta/s)(\eta/s)_*\rangle=-\frac{\sigma^2}{s^2}$$

where $\sigma^2=\langle|\eta|^2\rangle$ and $\langle,\rangle$ denotes ensemble average.

[3]Function arguments are omitted wherever convenient and for any matrix A, A', $\bar{A}$, $A^*(\equiv\bar{A}')$ and det A denote the transpose, complex-conjugate, ajoint and determinant of A, respectively. Column vectors are written $\boldsymbol{a}$, $\boldsymbol{x}$, etc., or as $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)'$ to exhibit the components explicitly. Last, for any real rational matrix $A(s)$ of the complex frequency variable $s=\sigma+j\omega$, $A_*(s)\equiv A'(-s)$. Note that for $s=j\omega$, ω real, $A_*(j\omega)=A^*(j\omega)$.

must be considered the best available linear version of $y_d(s)$. Any reasonable performance measure should be based on the vector error difference

$$\boldsymbol{e}(s)=\boldsymbol{u}(s)-\boldsymbol{y}(s) \tag{3}$$

between the actual smoothed input $\boldsymbol{u}(s)$ driving the loop and the plant output $\boldsymbol{y}(s)$. If plant delays are excessive the suppression of load disturbance by means of feedback alone may not suffice and it is usually advisable to incorporate feedforward compensation $L(s)$ as an integral part of the design. For given choices of overall sensors $F(s)$ and $L(s)$, the design of the controller $C(s)$ evolves from an appropriate minimization procedure subject to a power-like constraint on $\boldsymbol{r}(s)$ to avoid plant saturation and to extend the linear range. (In nonlinear applications the constraint on $\boldsymbol{r}(s)$ is imposed to avoid permanent departures from the neighborhood of a desired equilibrium state.) Plant disturbance $\boldsymbol{d}(s)$ and instrument noise $\boldsymbol{m}(s)$, $\boldsymbol{l}(s)$ are modeled in a perfectly general way by assuming that

$$\boldsymbol{y}(s)=P(s)\boldsymbol{r}(s)+P_o(s)\boldsymbol{d}(s), \tag{4}$$

$$\boldsymbol{v}(s)=F(s)\boldsymbol{y}(s)+F_o(s)\boldsymbol{m}(s) \tag{5}$$

and

$$\boldsymbol{z}(s)=L(s)\boldsymbol{d}(s)+L_o(s)\boldsymbol{l}(s) \tag{6}$$

where $P(s)$, $P_o(s)$, $F(s)$, $F_o(s)$, $L(s)$, and $L_o(s)$ are real rational matrices.

In Fig. 1, $P(s)$ is $n\times m$, $F(s)$ is $n\times n$, and $C(s)$ is $m\times n$. Hence FPC is $n\times n$ and it is assumed of course that all other matrices are dimensioned compatible.[4] Straightforward analysis yields

$$\boldsymbol{y}=PR(\boldsymbol{u}-F_o\boldsymbol{m}-L_o\boldsymbol{l})+(P_o-PRP_d)\boldsymbol{d}, \tag{7}$$

$$\boldsymbol{r}=R(\boldsymbol{u}-F_o\boldsymbol{m}-L_o\boldsymbol{l}-P_d\boldsymbol{d}), \tag{8}$$

$$\boldsymbol{e}=(1_n-PR)\boldsymbol{u}+PR(F_o\boldsymbol{m}+L_o\boldsymbol{l})-(P_o-PRP_d)\boldsymbol{d} \tag{9}$$

where

$$R=CS, \tag{10}$$

$$S=(1_n+FPC)^{-1}, \tag{11}$$

$$P_d=FP_o+L. \tag{12}$$

In the absence of measurement noise and load disturbance, $\boldsymbol{y}=(PR)\boldsymbol{u}$. Thus,

$$T(s)=P(s)R(s) \tag{13}$$

is the closed-loop transfer matrix and $S(s)$ is the sensitivity matrix. In most industrial applications the available choices of physical sensing devices $L_t(s)$ and $F_t(s)$ are severely restricted and more or less dictated by the problem at hand. However, as explained later, low-power preequalizers $L_e(s)$ and $F_e(s)$ can and in many cases should be employed to improve stability margin, to assure

[4]1_n is the $n\times n$ identity matrix and $\boldsymbol{O}_n$, $O_{n,m}$, O_n denote the n-dimensional zero vector, the $n\times m$ and $n\times n$ zero matrices, respectively.

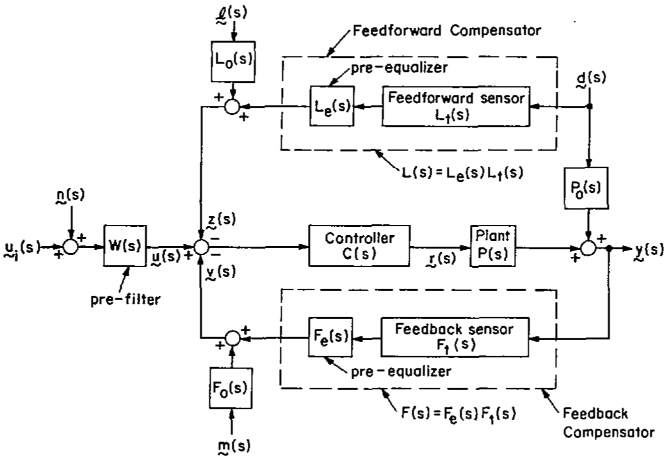

Fig. 1. Multivariable single-loop configuration.

zero steady-state error and to simulate delay in the feedback link. From this point on it is assumed that the data $P_o(s)$, $P(s)$, $L_o(s)$, $F_t(s)$, $L_t(s)$,

$$F(s)=F_e(s)F_t(s) \tag{14}$$

and

$$L(s)=L_e(s)L_t(s) \tag{15}$$

are prescribed in advance.

An interpretation of (9) with $F=1_n$ (unity feedback) and $\boldsymbol{z}=\boldsymbol{O}_n$ (no feedforward compensation) reveals most clearly the potential for tradeoff in the various frequency bands. In fact, for this case $1_n-PR=S$ and

$$\boldsymbol{e}=S(\boldsymbol{u}-P_o\boldsymbol{d})+(1_n-S)F_o\boldsymbol{m} \tag{16}$$

is composed of two contributions $\boldsymbol{e}_1(s)$ and $\boldsymbol{e}_2(s)$. The first,

$$\boldsymbol{e}_1=S(\boldsymbol{u}-P_o\boldsymbol{d}), \tag{17}$$

subsumes steady-state error, accuracy, and load disturbance while the second,

$$\boldsymbol{e}_2=(1_n-S)F_o\boldsymbol{m}, \tag{18}$$

is the error produced by measurement noise. The impossibility of making both $S(j\omega)$ and $1_n-S(j\omega)$ arbitrarily "small" over any frequency band is partly intrinsic and partly conditioned by the plant restrictions [5], [6]. This fundamental conflict is inevitable and responsible for a great deal of the difficulty surrounding practical analytic feedback design. According to (8), the spectral amplification from $\boldsymbol{m}(j\omega)$ to $\boldsymbol{r}(j\omega)$ must depend on the "size" of the matrix $R(j\omega)$. Since (with $F=1_n$)

$$P(j\omega)R(j\omega)=1_n-S(j\omega) \tag{19}$$

and plant bandwidth is usually confined to some low-frequency interval $0\leqslant\omega\leqslant\omega_o$, good transient response and sensitivity require

$$S(j\omega)\approx O_n, \qquad 0\leqslant\omega\leqslant\omega_o; \tag{20}$$

or, qualitatively,

$$C(j\omega)\approx\infty, \qquad 0\leqslant\omega\leqslant\omega_o. \tag{21}$$

Hence, even if $1_n-S(j\omega)$ approaches O_n as $\omega\to\infty$, there usually exists an intermediate high-frequency band $\omega_1\leqslant\omega\leqslant\omega_2$, $\omega_2-\omega_1\gg\omega_o$, over which $1_n-S(j\omega)\approx 1_n$ and $P(j\omega)\approx O_{n,m}$. In view of (19), this implies[5]

$$R(j\omega)\approx\infty, \qquad \omega_1\leqslant\omega\leqslant\omega_2, \tag{22}$$

and the rms value of $\boldsymbol{r}(j\omega)$ can easily exceed the saturation level of the plant because of the extremely wide-band nature of $\boldsymbol{m}(j\omega)$. If we assume the feedback sensors to be as noise free as possible the only remaining remedy is to concentrate on a "best" choice for $C(s)$ subject to the constraint that the closed loop be asymptotically stable. A successful quantitative reformulation of this latter requirement has enabled us to achieve a least-squares solution for $C(s)$ which takes into account all the pertinent performance criteria and is applicable to arbitrary open-loop unstable nonminimum-phase plants.

The McMillan degree, $\delta(A;s_o)$, of $s=s_o$ (finite or infinite) as a pole of the rational matrix $A(s)$ is the largest multiplicity it possesses as a pole of any minor of $A(s)$. The McMillan degree, $\delta(A)$, of $A(s)$ is the sum of the McMillan degrees of its distinct poles. Let the distinct *finite* poles of $A(s)$ be denoted by s_i and their associated McMillan degrees by δ_i, $i=1\to\mu$. The monic polynomial

$$\psi_A(s)=\prod_{i=1}^{\mu}(s-s_i)^{\delta_i} \tag{23}$$

is the *characteristic* denominator of $A(s)$. It is easily shown that $\psi_A(s)$ is also the monic least common multiple of all denominators of all minors of $A(s)$, each minor assumed expressed as the ratio of two relatively prime polynomials.

Lemma 1 (Appendix A): If the plant, feedback compensator, and controller (dynamical or otherwise) are free of unstable hidden modes, the closed loop of Fig. 1 is asymptotically stable iff

$$\varphi(s)\equiv\frac{\psi_P(s)\psi_C(s)\psi_F(s)}{\det S(s)} \tag{24}$$

is a strict Hurwitz polynomial.[6]

By assumption, $P(s)$ and $F(s)$ are given, but even granting that the plant and feedback compensator are free of unstable hidden modes, it may still be impossible to find a controller $C(s)$ which stabilizes the closed loop. Such is the case iff for some finite $s=s_o$, $\mathrm{Re}\,s_o\geqslant 0$,

$$\delta(FP;s_o)<\delta(P;s_o)+\delta(F;s_o). \tag{25}$$

This means that there exists at the output of the plant or feedback compensator an exponential ramp-modulated

[5]At this stage these arguments are intentionally informal. Incidentally, for nonminimum-phase unstable plants it is also possible to have $1_n-S(j\omega)\approx\infty$ over $\omega_1\leqslant\omega\leqslant\omega_2$. See the example in [6, p. 12] and the accompanying table in Fig. 5.

[6]A polynomial free of zeros in $\mathrm{Re}\,s\geqslant 0$ is said to be strict Hurwitz and "iff" abbreviates "if and only if." A system with transfer matrix $A(s)$ is dynamical if $A(s)$ is proper, i.e., if $A(\infty)$ is finite.

sinusoid whose growth in time exceeds that of the signal at the corresponding input and the necessary corrective action to effect stabilization is lacking.

Definition 1: The plant and feedback compensator form an admissible pair if each is individually free of unstable hidden modes and

$$\psi_{FP}^{+}(s)=\phi_{F}^{+}(s)\psi_{P}^{+}(s). \tag{26}$$

(The monic polynomials $\psi^{+}(s)$ and $\psi^{-}(s)$ absorb all the zeros of $\psi(s)$ in $\mathrm{Re}\,s\geqslant 0$ and $\mathrm{Re}\,s<0$, respectively. Thus, up to a multiplicative constant, $\psi=\psi^{+}\psi^{-}$.)

Lemma 2 (Appendix A): There exists a controller stabilizing the given plant and feedback compensator in the closed-loop configuration of Fig. 1 iff the pair is admissible.

Let the spectral densities of $\boldsymbol{u}(s)$, $\boldsymbol{d}(s)$, $\boldsymbol{l}(s)$, and $\boldsymbol{m}(s)$ be denoted by $G_u(s)$, $G_d(s)$, $G_l(s)$, and $G_m(s)$, respectively. Setting aside for the moment all questions of convergence,[7]

$$2\pi jE_t=\mathrm{Tr}\int_{-j\infty}^{j\infty}\langle \boldsymbol{e}(s)\boldsymbol{e}_*(s)\rangle ds \tag{27}$$

is the usual quadratic measure of steady-state response. Similarly, if $P_s(s)$ represents the transfer matrix coupling the plant input $\boldsymbol{r}(s)$ to those "sensitive" plant modes which must be especially guarded against excessive dynamic excursions,

$$2\pi jE_s=\mathrm{Tr}\int_{-j\infty}^{j\infty}\langle P_s(s)\boldsymbol{r}(s)\boldsymbol{r}_*(s)P_{s*}(s)\rangle ds \tag{28}$$

is a proven useful penalty function for saturation [7]. Hence,

$$E=E_t+kE_s, \tag{29}$$

k an adjustable positive constant, serves as a weighted cost combining both factors and the optimal controller is chosen to minimize E. Now referring to (8) and (9) it is seen that $R(s)$ determines $\boldsymbol{r}(s)$, $\boldsymbol{e}(s)$, and E. Consequently $R(s)$ embodies all the design freedom and the next lemma plays an obvious and indispensable role.

Lemma 3 (Appendix A): Let the given plant and feedback compensator form an admissible pair with transfer matrix descriptions $P(s)$, $F(s)$. Let

$$F(s)P(s)=A^{-1}(s)B(s)=B_1(s)A_1^{-1}(s) \tag{30}$$

where the pairs $A(s)$, $B(s)$ and $B_1(s)$, $A_1(s)$ constitute any left–right coprime polynomial decompositions of $F(s)P(s)$, respectively. Select polynomial matrices $X(s)$ and $Y(s)$ such that[8]

$$A(s)X(s)+B(s)Y(s)=1_n. \tag{31}$$

[7] $\mathrm{Tr}A=\mathrm{trace}A$. In (27) and (28), $s=j\omega$, ω real.

[8] The existence of real polynomial matrices $X(s)$, $Y(s)$ satisfying (31) is guaranteed by the left-coprimeness of $A(s)$ and $B(s)$ (Appendix A). These polynomials need not be unique.

Then, 1) the closed-loop of Fig. 1 is asymptotically stable iff

$$R(s)=H(s)A(s) \tag{32}$$

where

$$H(s)=Y(s)+A_1(s)K(s)$$

and $K(s)$ is any $m\times n$ real rational matrix analytic in $\mathrm{Re}\,s\geqslant 0$ which satisfies the constraint

$$\det(X(s)-B_1(s)K(s))\not\equiv 0. \tag{33}$$

2) The stabilizing controller associated with a particular choice of admissible $K(s)$ possesses the transfer matrix

$$C=(Y+A_1K)(X-B_1K)^{-1}. \tag{34}$$

(From $AB_1=BA_1$ and (31) we deduce that

$$A(X-B_1K)+B(Y+A_1K)=1_n, \tag{35}$$

a useful relationship.)

In view of this lemma, the natural way to attack the problem of minimizing E is to vary over all $m\times n$ real rational matrices $K(s)$ analytic in $\mathrm{Re}\,s\geqslant 0$ which satisfy restriction (33). We are now ready to discuss in detail the assumptions which justify the entire optimization scheme.

1) Rate gyros and tachometers are examples of practical sensing devices which are not modeled as dynamical systems. Yet almost invariably sensors are stable and their transfer matrices are analytic in $\mathrm{Re}\,s\geqslant 0$. However, for our purposes it suffices to assume that the feedforward cascade is stable and that $F(s)$ is analytic on the finite $s=j\omega$-axis. In particular, $L(s)$ is analytic in $\mathrm{Re}\,s\geqslant 0$. If $P_o(s)$ and $F_o(s)$ represent distinct physical blocks, these blocks must be stable and both $P_o(s)$ and $F_o(s)$ are analytic in $\mathrm{Re}\,s\geqslant 0$. On the other hand, if $P_o(s)$ and $F_o(s)$ are merely part of the paper modeling, it is possible to relax the analyticity requirements.

2) A pole of $F(s)P(s)$ in $\mathrm{Re}\,s>0$ reveals true open-loop instability whereas a finite pole on the $s=j\omega$-axis is usually present because of intentional high-gain preconditioning. Recall that in the absence of load disturbance and measurement noise, a unity-feedback single-input–output loop enclosing a plant whose transfer function possesses a pole of order ν at $s=j\omega_o$ will track any causal linear combination of $e^{j\omega_o t}\cdot t^{k-1}$, $k=1\rightarrow\nu$, with zero steady-state error. The correct generalization to the multivariable case is easy to find. Setting $\boldsymbol{d}=\boldsymbol{O}$, $\boldsymbol{l}=\boldsymbol{O}$, and $\boldsymbol{m}=\boldsymbol{O}$ in (7) and (9) we obtain

$$\boldsymbol{y}=T\boldsymbol{u} \tag{36}$$

and

$$\boldsymbol{e}=(1_n-T)\boldsymbol{u}. \tag{37}$$

For a stable configuration $T(s)$ is analytic in $\mathrm{Re}\,s\geqslant 0$, but not necessarily proper (Appendix A). Nevertheless, there are cogent reasons for insisting on a dynamical closed-loop design. Consider the conditions that must prevail if the loop is to track any one-sided input of the form

$$e^{j\omega_o t}\cdot \sum_{k=1}^{\nu} t^{k-1}\boldsymbol{d}_{\nu-k}, \tag{38}$$

$\boldsymbol{d}_k$ a constant vector, $k=1\to\nu$, with zero steady-state error. Clearly,[9]

$$\boldsymbol{e}(s)=\sum_{k=1}^{\nu}\Gamma(k)\cdot\frac{1_n-T(s)}{(s-j\omega_o)^k}\cdot\boldsymbol{d}_{\nu-k} \tag{39}$$

is the transform of a bounded time function which vanishes as $t\to\infty$ iff it is analytic in $\operatorname{Re}s\geq 0$ and $\boldsymbol{e}(\infty)=\boldsymbol{O}$. As necessary consequences, $T(s)$ must be proper and

$$(1_n-T(j\omega_o))\boldsymbol{d}_o=\boldsymbol{O}_n. \tag{40}$$

As is well known, (40) possesses a nontrivial solution $\boldsymbol{d}_o$ iff

$$\det(1_n-T(j\omega_o))=0. \tag{41}$$

Conversely, any ω_o satisfying (41) generates a generalized ramp-modulated sinusoid (38) capable of being tracked with zero steady-state error. These inputs and their finite linear combinations constitute a most important class of shape-deterministic information-bearing signals and play a key role in industrial applications. The set of all such possible "infinite gain" frequencies ω_o coincides with the totality of real solutions of (41).

In view of the arguments presented in the initial paragraph of 2), poles of $P(s)$ on the $s=j\omega$-axis enable the loop to track certain inputs with zero steady-state error. With unity feedback, $1_n-T=S$ and all plant poles in $\operatorname{Re}s\geq 0$, counted according to their McMillan degrees, are indeed zeros of $\det S$ (Appendix A). However, if $F\neq 1_n$, this perfect tracking capability is lost unless $F(s)$ is also conditioned suitably. Employing the easily derived formulas

$$S=(X-B_1K)A \tag{42}$$

and

$$FPR=1_n-S, \tag{43}$$

it is seen that

$$\begin{aligned}1_n-T&=S+(F-1_n)PR\\&=(X-B_1K+(F-1_n)P(Y+A_1K))A.\end{aligned} \tag{44}$$

Since $P(s)$ and $F(s)$ form an admissible pair,

$$\det{}^+A=\psi_{FP}^+=\psi_F^+\psi_P^+ \tag{45}$$

and (44) shows that the purely imaginary zeros of $\psi_P^+(s)$ will surely be zeros of $\det(1_n-T)$ provided

$$Z=X-B_1K+(F-1_n)P(Y+A_1K) \tag{46}$$

is designed to be analytic on the finite $s=j\omega$-axis.[10] This analyticity precludes any possibility of cancellation and is achieved iff $(F-1_n)P$ is analytic on the finite $s=j\omega$-axis. The proof of this assertion is somewhat tedious but because of its great importance we supply it in detail.

In 4) it is shown that $P(s)A_1(s)$ is automatically analytic for all finite $s=j\omega$ and it follows from (46) that the same is true of $Z(s)$ iff $(F-1_n)PY$ is analytic for all finite $s=j\omega$. Multiplication of both sides of (31) on the left by $A^{-1}(s)$ yields

$$X+FPY=A^{-1} \tag{47}$$

since $FP=A^{-1}B$. By assumption, $F(s)$ is analytic for $s=j\omega$ whence, by admissibility and (47), any finite purely imaginary pole $s_o=j\omega_o$ of $P(s)$ of McMillan degree ν_o must also be a pole of $P(s)Y(s)$ of the same degree. Write

$$PY=A_2^{-1}B_2 \tag{48}$$

and

$$P=A_p^{-1}B_p \tag{49}$$

where the polynomial pairs (B_2,A_2) and (B_p,A_p) are both left-coprime. By hypothesis,

$$(F-1_n)PY=(F-1_n)A_2^{-1}B_2 \tag{50}$$

is analytic for finite $s=j\omega$. But (X_2 and Y_2 are polynomial)

$$A_2X_2+B_2Y_2=1_n$$

implies

$$(F-1_n)X_2+(F-1_n)A_2^{-1}B_2Y_2=(F-1_n)A_2^{-1} \tag{51}$$

and $(F-1_n)A_2^{-1}$ is also analytic on the finite $s=j\omega$-axis. Substituting (49) into (48) we get

$$B_pY=A_pA_2^{-1}B_2 \tag{52}$$

and by an argument similar to the above, the analyticity of B_pY for finite $s=j\omega$ forces that of $A_pA_2^{-1}$. In other words,

$$A_2^{-1}=A_p^{-1}\mathcal{P} \tag{53}$$

where $\mathcal{P}(s)$ is analytic on the finite $s=j\omega$-axis. However, because the finite $j\omega$-axis poles of PY agree with those of P, McMillan degrees included, $\det A_2(s)$ and $\det A_p(s)$ possess the same $j\omega$-axis zeros, multiplicities included. Thus,

$$\det\mathcal{P}(s)\neq 0, \qquad s=j\omega \tag{54}$$

and $\mathcal{P}^{-1}(s)$ and $(F-1_n)A_2^{-1}$ are, therefore, both analytic on the finite $s=j\omega$-axis. It is now clear that

$$(F-1_n)P=(F-1_n)A_p^{-1}B_p=(F-1_n)A_2^{-1}\mathcal{P}^{-1}B_p \tag{55}$$

is analytic for all finite $s=j\omega$. Q.E.D.

This constraint is of decisive importance and replaces the usual unity-feedback desideratum $F=1_n$ which due to

[9] $\Gamma(k)$ is the Gamma function of argument k.

[10] The multiplicity of any zero of $\psi_P(s)$ equals its McMillan degree as a pole of $P(s)$. Hence, the purely imaginary zeros of $\psi_P^+(s)$ constitute the totality of finite $j\omega$-axis poles of $P(s)$.

ever-present delays and transducer inertia is never realizable. In the actual design the constraint is met by a correct choice of preequalizer $F_e(s)$ and the two degrees of freedom inherent in the problem are exploited to maximum advantage.

3) In process control the recovery of steady state under load disturbance $\boldsymbol{d}(s)$ is a requirement of paramount importance. From (9), with $\boldsymbol{u}$, $\boldsymbol{m}$, and $\boldsymbol{l}$ set equal to $\boldsymbol{O}$,

$$e=(P_0-PRP_d)\boldsymbol{d} \tag{56}$$

$$=(1_n-PRF)P_o\boldsymbol{d}-(PRL)\boldsymbol{d}$$
$$=S_1P_o\boldsymbol{d}-TL\boldsymbol{d}=(S_1P_o-TL)\boldsymbol{d} \tag{57}$$

where

$$S_1=(1_n+PCF)^{-1}. \tag{58}$$

Again as in 2), the shape-deterministic component of $\boldsymbol{d}(s)$ is envisaged to be the transform of a sum of generalized ramp-modulated sinusoids and for bounded zero steady-state error, $\boldsymbol{e}(s)$ must vanish at infinity and be analytic in $\text{Re}\, s\geqslant 0$. Assuming $S_1P_o—TL$ proper and $S_1P_o\boldsymbol{d}$ and $L\boldsymbol{d}$ analytic in $\text{Re}\, s\geqslant 0$ is evidently sufficient. Invoking closed-loop stability it can be shown (Appendix A) that

$$S_1(s)=\mathcal{P}_1(s)A_p(s), \tag{59}$$

$\mathcal{P}_1(s)$ analytic in $\text{Re}\, s\geqslant 0$. Hence, $S_1P_o\boldsymbol{d}$ analytic in $\text{Re}\, s \geqslant 0$ can be replaced by $A_pP_o\boldsymbol{d}$ analytic in $\text{Re}\, s\geqslant 0$ and once again the $j\omega$-axis poles of the plant are brought into evidence through $A_p(s)$. Observe that the $j\omega$-axis analyticity of $A_p(P_oG_dP_{o*})A_{p*}$, LG_dL_*, and $LG_dP_{o*}A_{p*}$ is a corollary.

4) Let

$$P=B_{p1}A_{p1}^{-1} \tag{60}$$

be any right-coprime decomposition of $P(s)$. Then,

$$B_1A_1^{-1}=FP=FB_{p1}A_{p1}^{-1} \tag{61}$$

and it follows from the assumed analyticity of $F(s)$ on the finite $j\omega$-axis that

$$B_1A_1^{-1}A_{p1}=FB_{p1} \tag{62}$$

is also analytic for all finite $s=j\omega$. Hence, reasoning as in 2),

$$A_{p1}(s)=A_1(s)\mathcal{P}_2(s), \tag{63}$$

$\mathcal{P}_2(s)$ analytic and nonsingular for all finite $s=j\omega$. Consequently,

$$PA_1=B_{p1}A_{p1}^{-1}A_1=B_{p1}\mathcal{P}_2^{-1} \tag{64}$$

is analytic for all finite $s=j\omega$. Q.E.D.

5) From (8), (9) and the definitions (28), (27),[11]

$$2\pi jE_s=\text{Tr}\int_{-j\infty}^{j\infty}QR\left(G_u+G_{ml}+P_dG_dP_{d*}\right)R_*\,ds \tag{65}$$

and

$$2\pi jE_t=\text{Tr}\int_{-j\infty}^{j\infty}\big((1_n-PR)G_u(1_n-PR)_*+(PR)G_{ml}(PR)_* + (P_o-PRP_d)G_d(P_o-PRP_d)_*\big)ds \tag{66}$$

where

$$Q(s)=P_{s*}(s)P_s(s) \tag{67}$$

and

$$G_{ml}=F_oG_mF_{o*}+L_oG_lL_{o*}. \tag{68}$$

In terms of

$$G=G_u+G_{ml}+P_dG_dP_{d*} \tag{69}$$

and H,

$$2\pi jE_s=\text{Tr}\int_{-j\infty}^{j\infty}QH\left(AGA_*\right)H_*\,ds. \tag{70}$$

The nonnegative parahermetian matrices $G_{ml}(s)$ and $Q(s)$ are assumed to be free of finite $j\omega$-axis poles. (There is no physical reason for doing otherwise.) Since a stable closed-loop design forces $H(s)$ to be analytic in $\text{Re}\, s\geqslant 0$ (Lemma 3), the integrand of (70) will be devoid of finite $j\omega$-axis poles if AGA_* is analytic on the $j\omega$-axis. Consider first AG_uA_* and the equality

$$P+(F-1_n)P=FP=A^{-1}B. \tag{71}$$

Write, as before, $P=A_p^{-1}B_p$ where A_p, B_p is a left-coprime pair and substitute into (71). Bearing in mind that $(F-1_n)P$ is assumed to be analytic on the finite $j\omega$-axis, familiar reasoning[12] permits us to conclude that $A(s)=\mathcal{P}_3(s)A_p(s)$, $\mathcal{P}_3(s)$ analytic and nonsingular for all finite $s=j\omega$. Hence, the analyticity of $A_pG_uA_{p*}$ on the finite $s=j\omega$-axis guarantees that of AG_uA_*. This $j\omega$-axis analyticity is in accord with our previous reasoning. Namely, the deterministic part of $\boldsymbol{u}(s)$ is the transform of a sum of generalized ramp-modulated sinusoids whose resonant frequencies coincide with the $j\omega$-axis poles of $P(s)$. These poles *and only these poles* should appear as $j\omega$-axis poles of $G_u(s)$. But these poles are also imbedded in the Smith canonic structure of $A_p(s)$ and the $j\omega$-axis analyticity of $A_pG_uA_{p*}$ is merely a succinct formulation of one design objective. Regarding $A(P_dG_dP_{d*})A_*$, its $j\omega$-axis analyticity follows from the assumptions in 3) and the readily deduced relation

$$AF=\mathcal{P}_4A_p, \tag{72}$$

$\mathcal{P}_4(s)$ analytic for $s=j\omega$.

Let us now examine the $j\omega$-axis analyticity of the individual terms making up the integrand of (66). First,

[11] All random processes are assumed to be independent. Note $\text{Tr}(L_1L_2)=\text{Tr}(L_2L_1)$.

[12] Both the admissibility of the pair $F(s)$, $P(s)$ and the $j\omega$-axis analyticity of $F(s)$ must be invoked.

$$(1_n - PR)G_u(1_n - PR)_* = (1_n - PR)A^{-1}(AG_uA_*)A_*^{-1}(1_n - PR)_* \quad (73)$$
$$= (A^{-1} - PH)(AG_uA_*)(A^{-1} - PH)_* \quad (74)$$

and it suffices to prove that $A^{-1} - PH$ is analytic on the $j\omega$-axis. This is clear because K, PA_1, and $(F-1_n)P$ are analytic on $j\omega$, $A^{-1} = X + FPY$ and

$$A^{-1} - PH = A^{-1} - P(Y + A_1K) = A^{-1} - PY - PA_1K$$
$$= X + (F - 1_n)PY - (PA_1)K. \quad (75)$$

The $j\omega$-axis analyticity of the second and third terms in (66) follows from that of $G_{ml}(s)$ and 3). In order to exclude meaningless, but mathematically allowed physical degeneracies, we must also impose the restriction

$$\det(AGA_*)\cdot\det\big(A_{1*}(P_*P + kQ)A_1\big) \neq 0, \qquad s = j\omega. \quad (76)$$

This inequality is essential (Appendix B). It is also shown in Appendix B[13] that

$$G_u(j\omega) \leqslant O(1/\omega^2) \quad (77)$$

and

$$P_oG_dP_{o*} \leqslant O(1/\omega^2) \quad (78)$$

are suggested naturally by the requirement of finite cost. Furthermore, if

$$(P_*P + kQ)G \approx \omega^{2\mu}1_m, \quad (79)$$
$$G_d(j\omega) \approx \omega^{-2i}1 \quad (80)$$

and

$$P(s) = O(s^\nu), \quad (81)$$

the inequalities

$$\mu \geqslant \nu - 1 \quad (82)$$

and

$$i \leqslant 1 \quad (83)$$

assure the properness of T and $S_1P_o - TL$, respectively. In most applications load disturbance contains a step-component and (80) is satisfied with $i = 1$. (The integers μ and ν can be negative.)

6) In general, the effects of parameter uncertainty on P and F are more pronounced as ω increases and closed-loop sensitivity is an important consideration. Let F, PC, S, S_1, and T undergo changes from $(F)_a$, $(PC)_a$, $(S)_a$, $(S_1)_a$, and $(T)_a$ to $(F)_b$, $(PC)_b$, $(S)_b$, $(S_1)_b$, and $(T)_b$ at a fixed ω. Noting that $T = PCS = S_1PC$,

$$(T)_b - (T)_a = (S_1)_b(PC)_b - (PC)_a(S)_a$$
$$= (S_1)_b\big[(PC)_b(1_n + (F)_a(PC)_a) - (1_n + (PC)_b(F)_b)(PC)_a\big](S_a)$$
$$= (S_1)_b\big[(PC)_b - (PC)_a - (PC)_b\cdot((F)_b - (F)_a)(PC)_a\big](S)_a.$$

Thus

$$\Delta T = (S_1)_b\cdot\Delta(PC)\cdot(S)_a - (T)_b\cdot(\Delta F)\cdot(T)_a, \quad (84)$$

an exact formula valid for arbitrary increments ΔF, $\Delta(PC)$. To first order,

$$\delta T = S_1\cdot\delta(PC)\cdot S - T(\delta F)T \quad (85)$$

and we recover the classical differential version of (84). If $\det(FPC) \not\equiv 0$, (85) may also be rewritten as

$$T^{-1}\cdot\delta T = (PC)^{-1}\cdot\delta(PC)\cdot S - (\delta F)\cdot F^{-1}\cdot(1_n - S). \quad (86)$$

In words, at frequency ω,

left percent change in T = (left percent change in PC) $\cdot S(j\omega)$ − (right percent change in F) $\cdot (1_n - S(j\omega))$ (87)

and again $S(j\omega)$ and $1_n - S(j\omega)$ emerge as the pertinent matrix gain functions for the forward and return links, respectively. Clearly then, to combat the adverse effects of high-frequency uncertainty in the modeling of $F(j\omega)$ and the plant matrix $P(j\omega)$, it is sound engineering practice to insist on a design with $S(j\omega)$ proper and equal to 1_n at $\omega = \infty$. This feature is easily introduced into the analytic framework by means of the constraint

$$O(P) + O(F) \leqslant \mu \quad (88)$$

which simultaneously ensures that the Wiener–Hopf controller defined by (34) makes sense and is proper if $\mu \geqslant -1$ (Appendix B). Furthermore, (88) also forces $S_1 = (1_n + PCF)^{-1} \to 1_n$ as $\omega \to \infty$ which is consistent with the engineering interpretation of the right-percentage formula for T,

$$(\delta T)\cdot T^{-1} = S_1\cdot\delta(PC)\cdot(PC)^{-1} - (1_n - S_1)\cdot F^{-1}\cdot\delta F. \quad (89)$$

We should like to emphasize that the cost E already imposes a weighted penalty on the choice of forward and return-link sensitivities through the (somewhat disguised) presence of S and $1_n - S$ in the error $\boldsymbol{e}$. (Equation (16) for $F = 1_n$ illustrates the point.) All this is in accord with a basic tenet of the classical theory which states that good immunity to load disturbance and good forward-link sensitivity usually go hand in hand.

III. The Optimal Controller

By way of recapitulation we shall collate the major working assumptions.

[13] $A(s) \leqslant O(s^r)$ means that no entry in $A(s)$ grows faster than s^r as $s \to \infty$. The order of $A(s)$ equals r, i.e., $A(s) = O(s^r)$ if 1) $A(s) \leqslant O(s^r)$ and 2) *at least one* entry grows exactly like s^r. For $A(s)$ square, $A(s) \approx s^r 1$ abbreviates

$$\operatorname*{limit}_{s\to\infty} s^{-r}A(s) = A_\infty,$$

A_∞ *a constant nonsingular matrix*. Note $A(s) \approx s^r 1$ implies $A(s) = O(s^r)$, but not conversely.

Assumption 1: The plant and feedback compensator form an admissible pair (Definition 1), the feedforward compensator is asymptotically stable and the respective transfer matrices, $P(s)$, $F(s)$, $L(s)$ are prescribed in advance. (Note, in particular, that $L(s)$ is analytic in Re $s \geq 0$.)

Assumption 2: $P_o(s)$, $F_o(s)$, $L_o(s)$, $Q(s) = P_{s*}(s)P_s(s)$ and the spectral densities $G_u(s)$, $G_d(s)$, $G_m(s)$, $G_l(s)$ are given. Any block outside the loop which represents an actual physical component must be asymptotically stable and its transfer matrix is therefore analytic in Re $s \geq 0$. (On the other hand, if any such block is merely part of the paper modeling the analyticity requirement can be relaxed.) The input signal, load disturbance and measurement noises are stochastically independent.

Assumption 3: Let $P = A_p^{-1}B_p$ be any left-coprime factorization of $P(s)$ and let

$$G_{ml} = F_o G_m F_{o*} + L_o G_l L_{o*}. \tag{90}$$

The matrices Q, F, $(F - 1_n)P$, $A_p G_u A_{p*}$, $A_p(P_o G_d P_{o*})A_{p*}$, $LG_d L_*$, and G_{ml} are analytic on the finite $s = j\omega$-axis.

Assumption 4: Let k be any positive constant,

$$G = G_u + P_d G_d P_{d*} + G_{ml} \tag{91}$$

and

$$P_d = FP_o + L. \tag{92}$$

The matrices AGA_* and $A_{1*}(P_*P + kQ)A_1$ are nonsingular on the finite $s = j\omega$-axis. (Their $j\omega$-analyticity is ensured by the above assumptions.)

Assumption 5: The data satisfy the order relations[14]

$$G_u \leq O(1/s^2); \qquad P_o G_d P_{o*} \leq O(1/s^2)$$

$$G_d \approx s^{-2i} 1; \qquad P = O(s^{\nu}) \tag{93}$$

$$O(P) + O(F) \leq \mu \tag{94}$$

and

$$(P_*P + kQ)G \approx s^{2\mu} 1_m \tag{95}$$

where

$$i \leq 1; \qquad \mu \geq \max(\nu - 1, -1). \tag{96}$$

We are now in a position to state the master result.

Theorem 1 (Appendix B): Under Assumptions 1–5 the optimal design is carried out in the following manner.

1) Construct two square real rational matrices $\Lambda(s)$, $\Omega(s)$ *analytic together with their inverses in* $Res \geq 0$ such that

$$A_{1*}(P_*P + kQ)A_1 = \Lambda_* \Lambda \tag{97}$$

and

$$AGA_* = \Omega\Omega_*. \tag{98}$$

2) Let

$$I = A_{1*}P_*(G_u + P_o G_d P_{d*})A_* \tag{99}$$

and choose any two real polynomial matrices $X(s)$, $Y(s)$ such that

$$A(s)X(s) + B(s)Y(s) = 1_n. \tag{100}$$

3) The transfer matrix of the optimal controller is given by

$$C = (Y + A_1 K)(X - B_1 K)^{-1} \tag{101}$$

where[15]

$$K = \Lambda^{-1}\left(\left\{\Lambda_*^{-1} I \Omega_*^{-1}\right\}_+ + \left\{\Lambda A_1^{-1} Y \Omega\right\}_-\right)\Omega^{-1} - A_1^{-1}Y; \tag{102}$$

or, in a form more suitable for numerical implementation,

$$C = H_o\left(A^{-1}\Omega - FPH_o\right)^{-1}, \tag{103}$$

$$H_o = A_1\Lambda^{-1}\left(\left\{\Lambda_*^{-1} I \Omega_*^{-1}\right\}_+ + \left\{\Lambda A_1^{-1} Y \Omega\right\}_-\right). \tag{104}$$

The (nonhidden) poles of the optimally compensated loop are *precisely* the *zeros* of the strict Hurwitz polynomial

$$\theta(s) = \frac{\psi_F^-(s)\psi_P^-(s)}{\psi_{FP}^-(s)} \tag{105}$$

plus the finite *poles* of $K(s)$, each of these poles counted according to its McMillan degree. Both $H_o(s)$ and $K(s)$ are analytic in Re $s \geq 0$ and the *distinct* finite poles of $K(s)$ are included in those of the primary data

$$FP, \left(A_{1*}(P_*P + kQ)A_1\right)^{-1}, \left(AGA_*\right)^{-1},$$
$$A_{1*}P_*\left(G_u + P_o G_d P_{d*}\right)A_* \tag{106}$$

located in $\mathrm{Re}\, s < 0$. Thus stability margin is ascertainable in advance.

Several comments are in order. First, there exist effective computer algorithms for the realization of the canonic factors $\Lambda(s)$, $\Omega(s)$ [8], [9]. Second, the combination of plant and feedback compensator is said to be nonminimum-phase if the polynomial matrix $B(s)$ appearing in the left-coprime decomposition $F(s)P(s) = A^{-1}(s)B(s)$ has rank *less* than row-rank for some finite $s = s_o$ in $\mathrm{Re}\, s > 0$.[16] As is shown in Appendix A, any such s_o is also a zero of $\det(1_n - S)$. Now for any choice of nonzero constant $\eta \neq 1$ the zeros of the stability polynomial

$$\varphi(s, \eta) = \psi_F \psi_P \psi_C \cdot \det(1_n + \eta FPC)$$
$$= \psi_F \psi_P \psi_C \cdot \det(1_n + FPC + (\eta - 1)FPC) \tag{107}$$

[14]Refer to footnote 13 for an explanation of the notation.

[15]In the partial fraction expansion $\{\ \}_\infty + \{\ \}_+ + \{\ \}_-$ of any rational matrix, $\{\ \}_\infty$ is the part associated with the pole at infinity and $\{\ \}_+$, $\{\ \}_-$ the parts associated with all the finite poles in $\mathrm{Re}\, s < 0$ and $\mathrm{Re}\, s \geq 0$, respectively. Clearly, $\{\ \}_+$ is analytic in $\mathrm{Re}\, s \geq 0$, $\{\ \}_-$ in $\mathrm{Re}\, s < 0$ and both vanish at infinity.

[16]Although this definition is the most natural generalization of the one accepted in the scalar case, other definitions also make physical sense when examined in the context of the standard control problem [3].

coincide with those of[17]

$$\varphi(s,1)\cdot\det\left(\frac{1}{\eta-1}1_n+1_n-S\right). \tag{108}$$

Thus, by continuity, at least one of these zeros tends to $s=s_o$ as $|\eta|\to\infty$ and all attempts to decrease transient error to zero by a simple constant-gain modification of some already predetermined controller $C(s)$ must, therefore, fail.

Corollary 1: Suppose $F(s)P(s)$ is analytic in $\mathrm{Re}\,s\geqslant 0$. Then

$$C=H_o(\Omega_r-FPH_o)^{-1} \tag{109}$$

where

$$H_o=\Lambda_r^{-1}\{\Lambda_{r*}^{-1}I_r\Omega_{r*}^{-1}\}_+, \tag{110}$$

$$(P_*P+kQ)=\Lambda_{r*}\Lambda_r, \tag{111}$$

$$G=\Omega_r\Omega_{r*}, \tag{112}$$

$$I_r=P_*(G_u+P_oG_dP_{d*}) \tag{113}$$

and $\Lambda_r(s)$, $\Omega_r(s)$ are square, real rational matrices analytic together with their inverses in $\mathrm{Re}\,s\geqslant 0$. ($\Lambda_r,\Omega_r,I_r$ are "reduced" quantities.)

Proof: The analyticity of $F(s)P(s)$ in $\mathrm{Re}\,s\geqslant 0$ implies $\det^+A(s)=\det^+A_1(s)=\psi_{FP}^+(s)=1$. Thus $A\Omega_r=\Omega$, $\Lambda_rA_1=\Lambda$, $\{\Lambda A_1^{-1}Y\Omega\}_-=O$ and the rest follows by direct substitution. Q.E.D.

Under the conditions of the corollary, the feedback sensor and plant are asymptotically stable and the resulting simplification, as evidenced in (109)–(113), is striking. Note in particular that the polynomial factors $A(s)$, $A_1(s)$, and $Y(s)$ are no longer needed!

The general formula (103) for the optimal controller transfer matrix is excellently conditioned. In fact, it is easily shown (Appendix B) that in exact arithmetic

$$H_o=(Y+A_1K)\Omega \tag{114}$$

and

$$A^{-1}\Omega-FPH_o=(X-B_1K)\Omega. \tag{115}$$

Thus, in exact arithmetic both H_o and $A^{-1}\Omega-FPH$ are analytic in $\mathrm{Re}\,s\geqslant 0$. Consequently, if the numerical scheme employed to compute (103) automatically ensures the closed right half-plane analyticity of H_o and $A^{-1}\Omega-FPH_o$, the corresponding exact arithmetic K is such that A_1K and B_1K are also analytic in $\mathrm{Re}\,s\geqslant 0$.[18] But then,

$$X_1A_1K+Y_1B_1K=K \tag{116}$$

reveals that $K(s)$ is analytic in $\mathrm{Re}\,s\geqslant 0$ *and the closed-loop structure realized with the computed C(s) is asymptotically stable.*

[17] $1_n-S=(1_n+FPC)^{-1}FPC$.

[18] $\Omega^{-1}(s)$ is analytic in $\mathrm{Re}\,s\geqslant 0$.

Corollary 2: Let[19]

$$a=\Lambda_*^{-1}I\Omega_*^{-1}, \tag{117}$$

$$b=\Lambda A_1^{-1}Y\Omega, \tag{118}$$

$$c=\{a-b\}_- \tag{119}$$

and

$$\rho=G_u+P_oG_dP_{o*}-a_*a+c_*c. \tag{120}$$

Then, under the assumption[20]

$$a_\infty(s)=O, \tag{121}$$

the *minimum* cost $E_{\min}$ is given by

$$2\pi jE_{\min}=\mathrm{Tr}\int_{-j\infty}^{j\infty}\rho(s)\,ds. \tag{122}$$

In particular, if $F(s)P(s)$ is analytic in $\mathrm{Re}\,s\geqslant 0$ (the stable case) we can choose

$$\rho=G_u+P_oG_dP_{o*}-a_{+*}a_+. \tag{123}$$

Proof: From (B2), Appendix B,

$$a-b-\Lambda K\Omega=\Lambda_*^{-1}\Delta_*\Omega_*^{-1} \tag{124}$$

and it follows immediately that $c=\{a-b\}_-$ is analytic in $\mathrm{Re}\,s\leqslant 0$ (which includes the $j\omega$-axis). Now by combining (B19), (B22), and (B24) and exploiting the closed left half-plane analyticity of $c(s)$ with the aid of Cauchy's theorem, we easily reach (120)+(122). If $F(s)P(s)$ is analytic in $\mathrm{Re}\,s\geqslant 0$, $b_-(s)\equiv O$ and

$$\rho=G_u+P_oG_dP_{o*}-a_{+*}a_+-a_{+*}a_--a_{-*}a_+.$$

However, since $a_{+*}a_-$ is $O(1/\omega^2)$ and analytic in $\mathrm{Re}\,s\leqslant 0$, contour integration yields

$$\mathrm{Tr}\int_{-j\infty}^{j\infty}a_{+*}a_-\,ds=\mathrm{Tr}\int_{-j\infty}^{j\infty}a_{-*}a_+\,ds=0.$$

Thus

$$2\pi jE_{\min}=\mathrm{Tr}\int_{-j\infty}^{j\infty}(G_u+P_oG_dP_{o*}-a_{+*}a_+)\,ds.$$

Q.E.D.

In the stable case $G_u+P_oG_dP_{o*}$ and a_*a are both individually $j\omega$-analytic, but, in general, it is only the combination $G_u+P_oG_dP_{o*}-a_*a$ which is devoid of purely imaginary poles.

Corollary 3:[21] Let $P(s)$ be square and analytic *together with its inverse* in $\mathrm{Re}\,s\geqslant 0$, let $F=1$ (unity feedback), let $k=0$ (no saturation constraint) and assume feedforward compensation is not employed (L and G_l are zero). Then, if G and $G_u+P_oG_dP_{o*}$ are diagonal matrices, the optimal controller $C(s)$ satisfies the noninteraction condition

[19] Here $a_+\equiv\{a\}_+$, $b_-\equiv\{b\}_-$, $a_\infty\equiv\{a\}_\infty$, etc.

[20] Quite usual.

[21] Suggested some years ago by I. M. Horowitz [19].

$$P(s)C(s) = \text{diagonal matrix}. \tag{125}$$

Proof: Clearly, from (109)–(113) and the stipulated assumptions, it follows that $\Lambda_r = P$, Ω_r is diagonal and

$$I_r = \Lambda_{r*}(G_u + P_o G_d P_{o*}). \tag{126}$$

Thus

$$\Lambda_{r*}^{-1} I_r \Omega_{r*}^{-1} = (G_u + P_o G_d P_{o*})\Omega_{r*}^{-1}, \tag{127}$$

$$PH_o = \Lambda_r H_o = \{(G_u + P_o G_d P_{o*})\Omega_{r*}^{-1}\}_+ \tag{128}$$

and

$$PC = PH_o(\Omega_r - PH_o)^{-1} \tag{129}$$

are also diagonal. Q.E.D.

Let us mention some obvious generalizations. First, suppose the integrand in (27) is also weighted so that

$$2\pi j E_t = \int_{-j\infty}^{j\infty} \langle e_*(s) Q_t e(s)\rangle \, ds = \operatorname{Tr}\int_{-j\infty}^{j\infty} Q_t \langle e(s) e_*(s)\rangle \, ds, \tag{130}$$

Q_t an arbitrary real, constant, symmetric nonnegative-definite matrix. Then we simply make the substitutions

$$(P_* P + kQ) \rightarrow P_* Q_t P + kQ, \tag{130a}$$

$$I \rightarrow A_{1*} P_* Q_t (G_u + P_o G_d P_{d*}) A_* \tag{130b}$$

and

$$\rho \rightarrow Q_t (G_u + P_o G_d P_{o*}) - a_* a + c_* c \tag{130c}$$

and continue to use the same formulas as before. In particular, the canonic factor $\Lambda(s)$ is found from the decomposition

$$\Lambda_* \Lambda = A_{1*}(P_* Q_t P + kQ) A_1 \tag{130d}$$

and (76) is altered to read

$$\det(A_{1*}(P_* Q_t P + kQ) A_1) \cdot \det(A G A_*) \neq 0, \qquad s = j\omega. \tag{130e}$$

Second, we have assumed all processes to be zero-mean. This is always true for the measurement noises $\boldsymbol{m}$ and $\boldsymbol{l}$ and almost invariably true for $\boldsymbol{u}$. In any case, if at least one of the means $\langle \boldsymbol{u}\rangle$ or $\langle \boldsymbol{d}\rangle$ vanishes, all formulas remain intact. Otherwise, $\boldsymbol{e}$ in the integrand of E_t and $\boldsymbol{r}$ in the integrand of E_s must be replaced by $\boldsymbol{e} - \langle \boldsymbol{e}\rangle$ and $\boldsymbol{r} - \langle \boldsymbol{r}\rangle$, respectively. This then entails identifying G_u with the spectral density of $\boldsymbol{u} - \langle \boldsymbol{u}\rangle$ and G_d with that of $\boldsymbol{d} - \langle \boldsymbol{d}\rangle$. The optimal controller now minimizes the steady-state rms error fluctuation subject to a steady-state rms constraint on the fluctuation of the plant input.

IV. Discussion and Conclusions

It would be superfluous to list the numerical problems which beset algorithms involving the factorization and manipulation of rational matrices. Nevertheless, work now in progress leads us to believe that a feasible computer implementation of the optimal controller is within reach. The availability of such an algorithm will undoubtedly suggest related simpler suboptimal strategies.

Although many of the physical ideas propounded in this paper have already been touched upon in the literature by several authors [14], [15], the various attempts to evolve an inclusive frequency-domain least-squares approach to multivariable controller design have by and large been unsuccessful because of an imprecise grasp of the full implications of closed-loop stability and a failure to recognize at the outset the need to condition the cost functional in a manner compatible with the physical constraints introduced by the given data $F(s)$, $P(s)$. As this paper clearly shows, within a linear framework a correct treatment depends essentially on an in-depth analytic characterization of all those engineering factors which figure meaningfully in any practical design scheme and a successful parametric solution of the concomitant "model-matching" problems associated with the matrix Wiener–Hopf equation.[22]

In LQG [17], [18] the objective is to optimally reset the state of the plant to a fixed known equalibrium state in the face of both Gaussian white background noise and Gaussian white measurement noise given the measured output. However, in the problem solved in this paper the "set point" excursion $\boldsymbol{u}(t)$ is both variable and stochastic and the functions of time to be reset are subsumed as components of an output vector $\boldsymbol{y}(t)$. The task of the optimal controller is to optimally reset $y(t)$ to a new level dictated by a shape-deterministic or second-order stationary $\boldsymbol{u}(t)$. This must be accomplished despite the presence of shape deterministic or second-order stationary load disturbance, measurement noise, nonideal sensor dynamics, a zero steady-state error requirement, etc. Accordingly, that part of the cost E_t reflecting loop accuracy has been imposed directly on $\boldsymbol{u}(t) - \boldsymbol{y}(t)$ and does not necessarily involve all the state variables. Nevertheless, these other variables are kept within bounds by an appropriate weighting of the saturation constraint E_s. It appears, therefore, that underlying our design philosophy is the assumption that any variable which is to be reset to a time-varying stochastic set-point must be available as a measured output. This attitude is of course consistent with the classical viewpoint and its true merit can only be judged after sufficient experience with applications of the optimal controller formula (103) to problems of industrial importance has been obtained.[23]

In LQG the optimal regulator structure is prescribed in advance in a manner which identifies the separate roles played by state estimation, Kalman-weighting (via the

[22] The Ph.D. dissertation of H. A. Jabr [16] contains some nontrivial fully worked numerical examples and also includes a transfer-matrix description of the stirred-tank chemical reactor linearized about its unstable equilibrium state.

[23] One possible way to enlarge LQG to encompass a special variable set-point problem is described in [18] but it appears to us that the proposed method can be very sensitive to parameter variations and we prefer the infinite-gain plant preconditioning scheme discussed in 2), Section II.

innovative input) and noise-free optimal deterministic state feedback [17]. However, it is easily shown that the entire configuration is simply a special case of Fig. 1 in which $\boldsymbol{u}=\boldsymbol{O}$, feedforward compensation is absent, $F(s)=1$, $F_o(s)=1$, $G_d(s)=\Sigma$ and $G_m(s)=\Theta$; Σ is real, constant, symmetric nonnegative-definite and Θ is real, constant, symmetric positive-definite. In the frequency-domain approach the objective is to find the optimal controller transfer matrix $C(s)$ and its mode of realization is based on other considerations. This shifting of the emphasis to $C(s)$ is really a restatement of the problem in invariantive fashion and has some distinct advantages.

For example, consider a time-invariant single-input, single-output plant with the constant-coefficient state-variable description

$$\dot{\boldsymbol{x}}=\begin{bmatrix}0 & 1\\ 0 & 2\end{bmatrix}\boldsymbol{x}+\begin{bmatrix}0\\ 1\end{bmatrix}u_o+\xi, \tag{131}$$

$$y=\begin{bmatrix}-1 & 1\end{bmatrix}\boldsymbol{x}; \qquad z=y+\theta. \tag{132}$$

Here $u_o(t)$ is the plant input, $y(t)$ the output, and $z(t)$ the measured output. The noise processes $\xi(t)$ and $\theta(t)$ are both white Gaussian with respective covariance matrices

$$\langle \xi(t)\xi(\tau)\rangle=\begin{bmatrix}\sigma^2 & 0\\ 0 & 0\end{bmatrix}\cdot\delta(t-\tau)=\Sigma\cdot\delta(t-\tau) \tag{133}$$

and

$$\langle \theta(t)\theta(\tau)\rangle=\sigma_m^2\cdot\delta(t-\tau)=\Theta\cdot\delta(t-\tau). \tag{134}$$

Clearly,

$$P(s)=\frac{s-1}{s(s-2)}; \qquad P_o(s)=\left[-\frac{1}{s}\,\middle|\,\frac{s-1}{s(s-2)}\right] \tag{135}$$

$$G_d(s)=\Sigma; \qquad G_m(s)=\sigma_m^2. \tag{136}$$

Subject to the choices $k>0$ and

$$Q_o=\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}, \tag{137}$$

it is a consequence of the LQG solution that the cost functional

$$J=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\left(\boldsymbol{x}'Q_o\boldsymbol{x}+ku_o^2\right)dt \tag{138}$$

is minimized by choosing the control law

$$u_o=-\frac{1}{k}\left[\sqrt{k}\ \middle|\ 2k+\sqrt{4k^2+2k\sqrt{k}+k}\ \right]\hat{\boldsymbol{x}} \tag{139}$$

and placing the Kalman column-vector gain

$$W=\begin{bmatrix}\dfrac{3\sigma}{\sigma_m}+4\\[2mm] \hline \dfrac{4\sigma}{\sigma_m}+8\end{bmatrix} \tag{140}$$

at the innovative input.[24] The observer of course provides the state estimate $\hat{\boldsymbol{x}}$. Since $\boldsymbol{x}'Q_o\boldsymbol{x}=y^2$, this problem also falls within the scope of our solution (with $\boldsymbol{u}=\boldsymbol{O}$) and we obtain[25]

$$C(s)=\frac{\left[(3+4c_1+4c_2+8c_1c_2)\sigma+4(1+2c_1+2c_2+4c_1c_2)\sigma_m\right]s-2\sigma}{\sigma_m c_1c_2s^2+\left[c_1c_2\sigma+(c_1+c_2+4c_1c_2)\sigma_m\right]s-\left[(3+3c_1+3c_2+4c_1c_2)\sigma+(3+4c_1+4c_2+8c_1c_2)\sigma_m\right]} \tag{141}$$

where

$$\sqrt{2}\,c_{1,2}=\sqrt{(4k+1)\pm\sqrt{(4k+1)^2-4k}}\,. \tag{142}$$

Now in view of (139) and (140), the LQG design fails if either $k=0$ or $\sigma_m=0$. Nevertheless, $C(s)$ is perfectly well-defined and, in fact, setting $k=0$ in (141) yields

$$C(s)=\frac{\left(\dfrac{7\sigma}{\sigma_m}+12\right)s-\dfrac{2\sigma}{\sigma_m}}{s-\left(\dfrac{6\sigma}{\sigma_m}+7\right)}; \qquad k=0. \tag{143}$$

If, in addition, σ_m also equals zero,

$$C(s)=\frac{1}{3}-\frac{7}{6}s; \qquad k=\sigma_m=0. \tag{144}$$

Thus, the LQG technique misses these extremely simple practical controllers. On the other hand, if $k\sigma_m\neq 0$, it is straightforward to show that (139) and (140) lead to an overall controller with transfer function given precisely by (141). It is interesting to note that in this limiting case $(k=\sigma_m=0)$,

$$O(P)+O(F)=-1>-2=\mu$$

and since inequality (94) is violated it is not surprising to find that $C(s)$ is improper and $S(\infty)=-6\neq 1$.

The other Kalman gain alluded to in the previous footnote is given by

$$\hat{W}=\begin{bmatrix}-\dfrac{\sigma}{\sigma_m}\\[2mm] \hline 0\end{bmatrix} \tag{145}$$

and its associated cost

$$\hat{J}=\sigma^2+\sigma\sigma_m+\sigma^2\sqrt{4k+2\sqrt{k}+1} \tag{146}$$

is obviously *less* than the cost

[24] Actually, there is a second Kalman column-vector gain which yields a smaller cost than (139) but leads to an unstable closed-loop design. We address this point later on.

[25] All details are omitted.

$$J=\big(9+9\sqrt{4k+2\sqrt{k}+1}+24\sqrt{k}+32k+16\sqrt{4k^2+2k\sqrt{k}+k}\big)\sigma^2+$$
$$+\big(25+24\sqrt{4k+2\sqrt{k}+1}+80\sqrt{k}+128k+64\sqrt{4k^2+2k\sqrt{k}+k}\big)\sigma\sigma_m+$$
$$+\big(20+16\sqrt{4k+2\sqrt{k}+1}+64\sqrt{k}+128k+64\sqrt{4k^2+2k\sqrt{k}+k}\big)\sigma_m^2 \tag{147}$$

induced by W, (140). The two gains W and $\hat{W}$ are generated by the two distinct solutions

$$Z_R=\left[\begin{array}{c|c} 4\sigma^2+9\sigma\sigma_m+4\sigma_m^2 & 4\sigma^2+12\sigma\sigma_m+8\sigma_m^2 \\ \hline 4\sigma^2+12\sigma\sigma_m+8\sigma_m^2 & 4\sigma^2+16\sigma\sigma_m+16\sigma_m^2 \end{array}\right] \tag{148}$$

and

$$\hat{Z}_R=\left[\begin{array}{c|c} \sigma\sigma_m & 0 \\ \hline 0 & 0 \end{array}\right], \tag{149}$$

respectively, of the pertinent matrix Riccati equation. Although Z_R is positive-definite, $\hat{Z}_R$ is only semipositive-definite. Observe, that with our choice of Σ in (133), the plant described by (131) is not controllable from the equivalent scalar disturbance input and the various theorems relating to the uniqueness of the solution of the Riccati equation do not apply [18, p. 36].[26]

The controller[27]

$$\hat{C}(s)=\frac{\frac{\sigma}{\sigma_m}(2-s)}{ks^2+\left(\frac{k\sigma}{\sigma_m}+\sqrt{4k^2+2k\sqrt{k}+k}\right)s+\sqrt{k}\left(1+\frac{\sigma}{\sigma_m}\right)+\frac{\sigma}{\sigma_m}\sqrt{4k^2+2k\sqrt{k}+k}} \tag{150}$$

paired with the choice $\hat{Z}_R$ possesses a zero at $s=2$ coincident with a pole of $P(s)$, and this fact makes the instability of the LQG design immediately apparent. It appears, therefore, that the optimal stabilizing controller (141) yields a relative and not an absolute minimum for the cost functional J. This observation suggests the following question. Since LQG prejudges the structure of the controller and does not invoke closed-loop stability as an *a priori* constraint, is it really clear that the optimal stabilizing Z_R is always included in its several solutions? In any reasonable topology the collection of stabilizing controllers for a given plant-feedback sensor combination should form an open set and the answer is probably yes, but in our opinion the conjecture is in need of strict proof. On the other hand, the frequency-domain solution advanced in this paper not only absorbs many important practical factors easily and naturally, but also succeeds in completely circumventing the above difficulty.

[26] The existence of this second solution $\hat{Z}_R$ was kindly brought to the authors' attention by Dr. J. Boyd Pearson of Rice University, Houston, TX, who also supplied some interesting insights regarding its implications for LQG. We gratefully acknowledge his comments and helpful editorial suggestions.

[27] All details are omitted.

Appendix A

For sound practical reasons the components in the loop of Fig. 1 must not be restricted to be dynamical and a stability criterion must be general enough to encompass this case. Let the zero-state Laplace transform descriptions of the feedback compensator, controller, and plant be given by

$$F_i(s)\boldsymbol{x}_i(s)=G_i(s)\boldsymbol{u}_i(s), \tag{A1}$$
$$\boldsymbol{y}_i(s)=J_i(s)\boldsymbol{u}_i(s)+H_i(s)\boldsymbol{x}_i(s), \tag{A2}$$

$i=2\to 4$, respectively. All coefficient matrices are real and polynomial, all F_i's are square and as usual $\boldsymbol{x}_i$, $\boldsymbol{u}_i$, $\boldsymbol{y}_i$ denote, in the same order, the internal state, the input, and the output. Physical degeneracies are excluded by imposing the determinantal condition,

$$\prod_{i=2}^{4}\det F_i(s)\not\equiv 0. \tag{A3}$$

Clearly then

$$P_i(s)=J_i(s)+H_i(s)F_i^{-1}(s)G_i(s) \tag{A4}$$

is the transfer matrix of system no. i, $i=2\to 4$. As is well known [10], [11], system number i is asymptotically stable iff the scalar polynomial

$$\Delta_i(s)\equiv\det F_i(s) \tag{A5}$$

has all its zeros in $\operatorname{Re}s<0$, $i=2\to 4$. In the present notation,

$$\Delta_F(s)=\Delta_2(s); \qquad F(s)=P_2(s) \tag{A6}$$
$$\Delta_C(s)=\Delta_3(s); \qquad C(s)=P_3(s) \tag{A7}$$
$$\Delta_P(s)=\Delta_4(s); \qquad P(s)=P_4(s). \tag{A8}$$

In Fig. 2 the three systems are shown interconnected through a linear, time-invariant frequency-insensitive grid, and it is assumed that the inputs to this grid uniquely determine its outputs. Hence, there exist real constant matrices M_a and M_b such that

$$\boldsymbol{u}_t(s)=M_a\boldsymbol{y}_t(s)+M_b\boldsymbol{u}(s) \tag{A9}$$

where

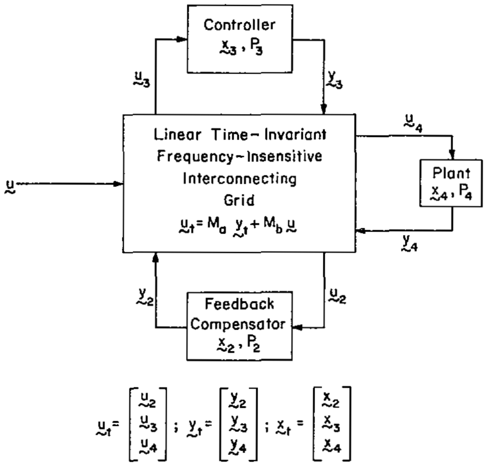

Fig. 2. Generalized interconnection scheme.

$$\boldsymbol{u}_t(s) = \begin{bmatrix} \boldsymbol{u}_2 \\ \boldsymbol{u}_3 \\ \boldsymbol{u}_4 \end{bmatrix}; \; \boldsymbol{y}_t(s) = \begin{bmatrix} y_2 \\ y_3 \\ y_4 \end{bmatrix}; \quad \boldsymbol{x}_t = \begin{bmatrix} x_2 \\ x_3 \\ x_4 \end{bmatrix}. \tag{A10}$$

Eliminating $\boldsymbol{u}_t$ in (A1) and (A2) with the help of (A9) and (A10) we obtain

$$\left[\begin{array}{c|c} F & -GM_a \\ \hline -H & 1-JM_a \end{array}\right] \left[\begin{array}{c} x_t \\ \hline y_t \end{array}\right] = \left[\begin{array}{c} G \\ \hline J \end{array}\right] M_b \boldsymbol{u} \tag{A11}$$

in which[28]

$$\begin{aligned} F &= F_2 \dotplus F_3 \dotplus F_4, \\ G &= G_2 \dotplus G_3 \dotplus G_4, \\ H &= H_2 \dotplus H_3 \dotplus H_4, \\ J &= J_2 \dotplus J_3 \dotplus J_4. \end{aligned} \tag{A12}$$

Consequently [10], [11], the interconnected system is asymptotically stable iff the determinant $\Delta(s)$ of the coefficient matrix on the left-hand side of (A11) has all its roots in $\mathrm{Re}\, s < 0$. A straightforward row operation yields

$$\Delta(s) = \det(1 - P_t(s)M_a) \cdot \prod_{i=2}^{4} \Delta_i(s), \tag{A13}$$

$$P_t(s) = P_2(s) \dotplus P_3(s) \dotplus P_4(s). \tag{A14}$$

The interconnection is nondegenerate iff

$$\det(1 - P_t(s)M_a) \not\equiv 0 \tag{A15}$$

which is exactly the necessary and sufficient condition for the existence of an overall transfer matrix description $T_t(s)$. In fact if $\boldsymbol{y}_t = T_t \boldsymbol{u}$,

$$T_t(s) = (1 - P_t(s)M_a)^{-1} P_t(s) M_b. \tag{A16}$$

(The easy derivation is left to the reader.)

[28] $A \dotplus B$ is the "direct sum" of matrices A and B.

For the topology depicted in Fig. 1,

$$\begin{bmatrix} \boldsymbol{u}_2 \\ \boldsymbol{u}_3 \\ \boldsymbol{u}_4 \end{bmatrix} = \begin{bmatrix} O & O & 1 \\ -1 & O & O \\ O & 1 & O \end{bmatrix} \begin{bmatrix} y_2 \\ y_3 \\ y_4 \end{bmatrix} + \begin{bmatrix} O \\ 1 \\ O \end{bmatrix} \boldsymbol{u}. \tag{A17}$$

Thus,

$$1 - P_t M_a = 1 - (F \dotplus C \dotplus P) \begin{bmatrix} O & O & 1 \\ -1 & O & O \\ O & 1 & O \end{bmatrix} \tag{A18}$$

$$= \begin{bmatrix} 1 & O & -F \\ C & 1 & O \\ O & -P & 1 \end{bmatrix} \tag{A19}$$

and

$$\det(1 - P_t M_a) = \det(1_n + FPC) = 1/\det S. \tag{A20}$$

Expression (A13) for $\Delta(s)$ now assumes the form

$$\Delta(s) = \frac{\Delta_P(s)\Delta_C(s)\Delta_F(s)}{\det S(s)}. \tag{A21}$$

If all components have asymptotically stable hidden modes,

$$\begin{aligned} \Delta_P(s) &= h_P(s)\psi_P(s), \\ \Delta_C(s) &= h_C(s)\psi_C(s), \\ \Delta_F(s) &= h_F(s)\psi_F(s) \end{aligned} \tag{A22}$$

where the h's are strict Hurwitz and ψ_P, ψ_C, ψ_F are the characteristic denominators of plant, controller, and feedback compensator, respectively [10], [11].[29] Thus, the loop is asymptotically stable iff

$$\varphi(s) = \frac{\psi_P(s)\psi_C(s)\psi_F(s)}{\det S(s)} \tag{A23}$$

is strict Hurwitz which is precisely the assertion of Lemma 1. Q.E.D.

Let

$$F(s)P(s) = A^{-1}(s)B(s) \tag{A24}$$

be any left-coprime polynomial decomposition of $F(s)P(s)$ and

$$C(s) = B_C(s)A_C^{-1}(s) \tag{A25}$$

any right-coprime decomposition of $C(s)$.

Then [10], [12],

$$\psi_{FP}(s) = \det A(s) \tag{A26}$$

and

$$\psi_C(s) = \det A_C(s). \tag{A27}$$

[29] Equation (A22) is also obvious from (A4).

Evidently,

$$S=(1_n+FPC)^{-1}=A_C(AA_C+BB_C)^{-1}A \quad (A28)$$

and

$$\det S=\frac{\psi_C\psi_{FP}}{g(s)}\not\equiv 0 \quad (A29)$$

where

$$g(s)=\det(AA_C+BB_C) \quad (A30)$$

is a polynomial. Substituting into (A23),

$$\varphi=g\cdot\frac{\psi_F\psi_P}{\psi_{FP}}. \quad (A31)$$

Since the McMillan degree of any pole of $F(s)P(s)$ cannot exceed the sum of its degrees as a pole of $F(s)$ and $P(s)$, ψ_{FP} must divide $\psi_F\psi_P$ without remainder and the quotient $\psi_F\psi_P/\psi_{FP}$ is polynomial. Thus, if the loop is asymptotically stable $g(s)$ is necessarily strict Hurwitz. Furthermore, any zero of the product $\psi_F^+\psi_P^+$ must be cancelled by a zero of ψ_{FP}^+ whence

$$\psi_{FP}^+(s)=\psi_F^+(s)\psi_P^+(s) \quad (A32)$$

is also necessary for closed-loop stability.

Suppose now that (A32) is satisfied and let the real polynomial matrices $X(s)$ and $Y(s)$ be chosen so that[30]

$$A(s)X(s)+B(s)Y(s)=1_n \quad (A33)$$

and

$$\det X(s)\not\equiv 0. \quad (A34)$$

Select any controller with asymptotically stable hidden modes and with transfer matrix

$$C(s)=Y(s)X^{-1}(s). \quad (A35)$$

According to (A33), the pair (Y,X) is right-coprime and

$$AA_C+BB_C=AX+BY=1_n. \quad (A36)$$

Thus $g(s)=1$ and the associated stability polynomial $\varphi(s)$ is given by

$$\varphi=\frac{\psi_F\psi_P}{\psi_{FP}}=\frac{\psi_F^-\psi_P^-}{\psi_{FP}^-} \quad (A37)$$

which is devoid of zeros in $\mathrm{Re}\, s\geqslant 0$. Consequently, the closed-loop structure is asymptotically stable and Lemma 2 is established. Q.E.D.

Recall from (10) that $R=CS$ or, in terms of the polynomial factors A, B, B_C, and A_C,

$$R=B_C(AA_C+BB_C)^{-1}A=HA, \quad (A38)$$

$$H=B_C(AA_C+BB_C)^{-1}. \quad (A39)$$

[30]The left-coprimeness of the pair $A(s)$, $B(s)$ guarantees that such a choice is always possible [10], [12].

For a stable loop, $g=\det(AA_C+BB_C)$ is strict Hurwitz and it follows immediately that $H(s)$ is analytic in $\mathrm{Re}\, s\geqslant 0$. Let

$$F(s)P(s)=B_1(s)A_1^{-1}(s) \quad (A40)$$

be any right-coprime polynomial factorization of $F(s)\ P(s)$ and define $K(s)$ via the equation

$$Y+A_1K=B_C(AA_C+BB_C)^{-1}. \quad (A41)$$

With this choice of K,

$$\begin{aligned}X-B_1K&=X-B_1A_1^{-1}\big(B_C(AA_C+BB_C)^{-1}-Y\big)\\&=X+FPY-FPB_C(AA_C+BB_C)^{-1}\\&=A^{-1}\big(1_n-BB_C(AA_C+BB_C)^{-1}\big)\\&=A_C(AA_C+BB_C)^{-1}=SA^{-1}.\end{aligned} \quad (A42)$$

Hence, if the polynomial matrices $X_1(s)$ and $Y_1(s)$ are constructed to satisfy

$$X_1A_1+Y_1B_1=1_m, \quad (A43)$$

(A41) and (A42) combine to give

$$K=(X_1B_C-Y_1A_C)(AA_C+BB_C)^{-1}+Y_1X-X_1Y \quad (A44)$$

which is obviously analytic in $\mathrm{Re}\, s\geqslant 0$. Conversely, let $K(s)$ be any real rational matrix analytic in $\mathrm{Re}\, s\geqslant 0$ such that $\det(X-B_1K)\not\equiv 0$ and select any controller with stable hidden modes and transfer matrix $C(s)$ given by

$$C=(Y+A_1K)(X-B_1K)^{-1}. \quad (A45)$$

Reasoning exactly as in Appendix B in the derivation of (B64) we find that

$$\varphi(s)=\frac{\psi_F^-(s)\psi_P^-(s)}{\psi_{FP}^-(s)}\cdot\psi_K(s). \quad (A46)$$

Since $K(s)$ is analytic in $\mathrm{Re}\, s\geqslant 0$, $\psi_K(s)$ and therefore $\varphi(s)$ are both strict Hurwitz and the closed-loop is asymptotically stable. This completes the proof of Lemma 3. Q.E.D.

The closed-loop transfer matrix

$$T=PR=PCS=(1_n+PCF)^{-1}PC. \quad (A47)$$

Let $PC=A_5^{-1}B_5$ and $F=B_FA_F^{-1}$ be left-and right-coprime polynomial factorizations. Evidently,

$$S_1=(1_n+PCF)^{-1}=A_F(A_5A_F+B_5B_F)^{-1}A_5 \quad (A48)$$

and

$$\det S_1=\det S=\frac{\psi_F\psi_{PC}}{g_1(s)}. \quad (A49)$$

$$g_1=\det(A_5A_F+B_5B_F). \quad (A50)$$

Hence

$$\varphi(s)=g_1(s)\cdot\frac{\psi_P(s)\psi_C(s)}{\psi_{PC}(s)} \quad (A51)$$

and it follows as before that the conditions $g_1(s)$ strict Hurwitz and $\psi_P^+\psi_C^+ = \psi_{PC}^+$ are both necessary for closed-loop stability. In particular,

$$T = PR = PCS \doteq S_1 PC = A_F(A_5 A_F + B_5 B_F)^{-1} B_5 \quad (A52)$$

is analytic in $\mathrm{Re}\, s \geq 0$.

From (A42),

$$\det S = \det(X - B_1 K) \cdot \det A = \det(X - B_1 K)\psi_{FP} \quad (A53)$$

and because of the analyticity of $K(s)$ in $\mathrm{Re}\, s \geq 0$, all zeros of $\psi_{FP}^+(s)$ are zeros of det $S(s)$, multiplicities included.

Introducing the left–right coprime polynomial decompositions $P = A_p^{-1} B_p$ and $CF = B_6 A_6^{-1}$ into (A48) gives

$$S_1 = A_6(A_p A_6 + B_p B_6)^{-1} A_p = \mathcal{P}_1 A_p \quad (A54)$$

and closed-loop stability forces $\mathcal{P}_1(s)$ to be analytic in $\mathrm{Re}\, s \geq 0$. We have now justified the three comments preceeding (38), (42) and (59). Q.E.D.

Also, let us remark that for given polynomial matrices $A_7(s)$, $B_7(s)$, the existence of polynomial matrices $X_7(s)$, $Y_7(s)$ such that

$$X_7(s) A_7(s) + Y_7(s) B_7(s) = 1 \quad (A55)$$

is possible iff for every fixed finite s the homogeneous pair

$$A_7(s)\boldsymbol{a} = \boldsymbol{O}, \quad (A56)$$

$$B_7(s)\boldsymbol{a} = \boldsymbol{O} \quad (A57)$$

admits only the trivial solution $\boldsymbol{a} = \boldsymbol{O}$. Necessity is trivial and sufficiency is easily established by actually constructing a solution pair $X_7(s)$, $Y_7(s)$ with the help of the Smith–McMillan theorem. The idea underlying the construction is very simple to grasp. Let

$$\Omega_c(s) = \mathrm{diag}\left[\frac{e_1(s)}{\psi_1(s)}, \frac{e_2(s)}{\psi_2(s)}, \cdots, \frac{e_k(s)}{\psi_k(s)}\right] \quad (A58)$$

be the canonic form of $F(s)P(s)$. Then [8], 1) k = normal rank $F(s)P(s)$; 2) the e's and ψ's are real monic polynomials uniquely determined by $F(s)P(s)$; 3) each $e_i(s)$ is relatively prime to its mate $\psi_i(s)$, $i = 1 \to k$; 4) $e_i(s)$ divides $e_{i+1}(s)$ and $\psi_{i+1}(s)$ divides $\psi_i(s)$, $i = 1 \to k-1$; 5) the distinct finite zeros and poles of $F(s)P(s)$ are identical, respectively, with the distinct zeros of $e_1(s)$ and $\psi_1(s)$; 6) the McMillan degree of any finite pole of $F(s)P(s)$ equals its multiplicity as a root of the characteristic denominator

$$\psi_{FP}(s) = \prod_{i=1}^{k} \psi_i(s). \quad (A59)$$

From the Smith–McMillan theorem [8],

$$FP = U(\Omega_c \dotplus O_{n-k,m-k})V \quad (A60)$$

where $U(s)$ and $V(s)$ are square, real elementary polynomial matrices.[31] Since $e_i(s)$ is relatively prime to $\psi_i(s)$ there exist [13] two real polynomials $\alpha_i(s)$, $\beta_i(s)$ such that $\beta_i(s) \not\equiv 0$ and

$$\alpha_i(s) e_i(s) + \beta_i(s)\psi_i(s) = 1, \qquad i = 1 \to k. \quad (A61)$$

Let[32]

$$\epsilon = \mathrm{diag}[e_1, e_2, \cdots, e_k], \quad (A62)$$

$$\chi = \mathrm{diag}[\psi_1, \psi_2, \cdots, \psi_k], \quad (A63)$$

$$\alpha = \mathrm{diag}[\alpha_1, \alpha_2, \cdots, \alpha_k], \quad (A64)$$

$$\beta = \mathrm{diag}[\beta_1, \beta_2, \cdots, \beta_k]. \quad (A65)$$

Then, putting

$$A = U(\chi \dotplus 1_{n-k})U^{-1}, \quad (A66)$$

$$B_1 = B = U(\epsilon \dotplus O_{n-k,m-k})V, \quad (A67)$$

$$A_1 = V^{-1}(\chi \dotplus 1_{m-k})V, \quad (A68)$$

$$X = U(\beta \dotplus 1_{n-k})U^{-1}, \quad (A69)$$

$$Y_1 = Y = V^{-1}(\alpha \dotplus O_{m-k,n-k})U^{-1} \quad (A70)$$

and

$$X_1 = V^{-1}(\beta \dotplus 1_{m-k})V, \quad (A71)$$

we verify by inspection that

$$\alpha\epsilon + \beta\chi = 1_k, \quad (A72)$$

$$AX + BY = 1_n, \quad (A73)$$

$$X_1 A_1 + Y_1 B_1 = 1_m, \quad (A74)$$

$$XA = AX = U(\chi\beta \dotplus 1_{n-k})U^{-1}, \quad (A75)$$

$$A^{-1}B = BA_1^{-1} = FP, \quad (A76)$$

$$XB = BX_1 = U(\beta\epsilon \dotplus O_{n-k,m-k})V, \quad (A77)$$

$$Y_1 X = X_1 Y = V^{-1}(\alpha\beta \dotplus O_{m-k,n-k})U^{-1}. \quad (A78)$$

Of course, other decompositions may not possess all the symmetry properties enumerated in (A66)–(A78).

According to (A29), any zero of the characteristic denominator $\psi_{FP}(s)$ in $\mathrm{Re}\, s \geq 0$ of multiplicity μ is a zero of det S of at least the same multiplicity. Define

$$e_{FP}(s) = \prod_{i=1}^{k} e_i(s) \quad (A79)$$

to be the characteristic *numerator* of $F(s)P(s)$. Suppose $\det(1_n - S) \not\equiv 0$. Then, any zero of $e_{FP}(s)$ in $\mathrm{Re}\, s \geq 0$ of multiplicity μ is a zero of $\det(1_n - S)$ of multiplicity at least μ. For the proof, note that

$$1_n - S = 1_n - (1_n + FPC)^{-1} = (1_n + FPC)^{-1} FPC;$$

or, using (A24), (A25), and (A30),

$$1_n - S = A_C(AA_C + BB_C)^{-1} \cdot (BB_C)A_C^{-1}. \quad (A80)$$

[31] Det $U(s)$ and det $V(s)$ equal nonzero constants.

[32] A square matrix A whose only nonzero elements are its main diagonal elements $a_1, a_2, \cdots, a_k$ is written $A = \mathrm{diag}[a_1, a_2, \cdots, a_k]$.

$$\therefore \det(1_n - S) = \frac{\det(BB_C)}{g(s)} \tag{A81}$$

and it is clear from (A62) and (A67) with $n=k$ that $\det(BB_C)$ is divisible by $e_{FP}(s)$. Since $g(s)$ is a strict Hurwitz polynomial, the assertion follows. Q.E.D.

Appendix B

Adding (66) to $k\times(65)$ we obtain $2\pi j(E_t + kE_s) = 2\pi jE$. Since $R=(Y+A_1K)A$,

$$\delta R = A_1(\delta K)A \tag{B1}$$

with $\delta K(s)$ analytic in $\mathrm{Re}\, s \geq 0$. Use of the standard variational argument [1] to examine the increment in E produced by the perturbation (B1) leads directly to the Wiener–Hopf equation

$$\Phi - A_{1*}(P_*P + kQ)A_1K(AGA_*) = \Delta_* \tag{B2}$$

where

$$\Phi = A_{1*}P_*(G_u + P_oG_dP_{d*})A_* - A_{1*}(P_*P+kQ)Y(AGA_*) \tag{B3}$$

and $\Delta(s)$ is analytic in $\mathrm{Re}\, s \geq 0$. If (B2) possesses a real rational matrix solution $K(s)$ analytic in $\mathrm{Re}\, s \geq 0$ which satisfies (33) and has a finite associated cost E, then this $K(s)$ is optimal. According to 5) and (76), $A_{1*}(P_*P + kQ)A_1$ and AGA_* are analytic for all $s=j\omega$ and the existence of a $K(s)$ with the desired properties implies the $j\omega$-axis analyticity of $\Phi(s)$. Since the latter is a unique construct from the prescribed data it is important to verify at the outset that this is indeed the case.

Using (69),

$$\Phi = A_{1*}P_*(G + (P_o - P_d)G_dP_{d*} - G_{ml})A_* - A_{1*}(P_*P + kQ)Y(AGA_*) \tag{B4}$$

and its $j\omega$-analyticity follows from that of[33]

$$A_{1*}P_*(P_o - P_d)G_dP_{d*}A_* + \left(A_{1*}P_*A^{-1} - A_{1*}(P_*P+kQ)Y\right)AGA_*$$

which in turn follows from that of

$$A_{1*}P_*(P_o - P_d)G_dP_{d*}A_* \tag{B5}$$

and

$$A_{1*}P_*A^{-1} - A_{1*}(P_*P+kQ)Y \equiv \beta. \tag{B6}$$

Expanding (B5),

$$\begin{aligned}(P_o - P_d)G_dP_{d*}A_* &= ((1_n - F)P_o - L)G_d(FP_o + L)_*A_* \\ &= (1_n - F)P_oG_dP_{o*}(AF)_* - LG_dL_*A_* \\ &\quad + (1_n - F)P_oG_dL_*A_* - LG_dP_{o*}(AF)_*.\end{aligned} \tag{B7}$$

[33] PA_1 and G_{ml} are $j\omega$-analytic.

All four terms are $j\omega$-analytic. First, $AF = \mathcal{P}_4A_p$, $\mathcal{P}_4(s)$ $j\omega$-analytic. Second, the analyticity of

$$(1_n - F)P_oG_dP_{o*}A_{p*} = (1_n - F)A_p^{-1}A_pP_oG_dP_{o*}A_{p*}, \tag{B8}$$

$$(1_n - F)P_oG_dL_* = (1_n - F)A_p^{-1}A_pP_oG_dL_*, \tag{B9}$$

LG_dL_* and $LG_dP_{o*}A_{p*}$ is implied by that of $(F-1_n)P$ and the assumptions introduced in 3).

With regard to (B6), replacing A^{-1} by $X+FPY$ transforms it into

$$A_{1*}P_*X + A_{1*}P_*(F-1_n)PY - kA_{1*}QY \tag{B10}$$

which is eveidently analytic on the $j\omega$-axis since Q, PA_1 and $(F-1_n)P$ are $j\omega$-analytic.

The solution of (B2) is now routine. Construct[34] two square real rational matrices $\Lambda(s)$, $\Omega(s)$ *analytic together with their inverses in* $\mathrm{Re}\, s \geq 0$ such that

$$A_{1*}(P_*P + kQ)A_1 = \Lambda_*\Lambda \tag{B11}$$

and

$$AGA_* = \Omega\Omega_*. \tag{B12}$$

From (B2),

$$\Lambda_*^{-1}\Phi\Omega_*^{-1} - \Lambda K\Omega = \Lambda_*^{-1}\Delta_*\Omega_*^{-1}. \tag{B13}$$

Effect the partial fraction decomposition

$$\Lambda_*^{-1}\Phi\Omega_*^{-1} = \left\{\Lambda_*^{-1}\Phi\Omega_*^{-1}\right\}_\infty + \left\{\Lambda_*^{-1}\Phi\Omega_*^{-1}\right\}_+ + \left\{\Lambda_*^{-1}\Phi\Omega_*^{-1}\right\}_- \tag{B14}$$

where $\{\ \}_\infty$ is the polynomial part of the Laurent expansion of $\Lambda_*^{-1}\Phi\Omega_*^{-1}$ associated with the pole at infinity and $\{\ \}_+$, $\{\ \}_-$ the parts associated with all the poles in $\mathrm{Re}\, s<0$ and $\mathrm{Re}\, s \geq 0$, respectively. Clearly, since Φ is analytic on $j\omega$, $\{\ \}_+$ is analytic in $\mathrm{Re}\, s\geq 0$, $\{\ \}_-$ in $\mathrm{Re}\, s \leq 0$ and both vanish for $s=\infty$. The substitution of (B14) into (B13) yields

$$\left\{\Lambda_*^{-1}\Phi\Omega_*^{-1}\right\}_+ - \Lambda K\Omega = \Lambda_*^{-1}\Delta_*\Omega_*^{-1} - \left\{\Lambda_*^{-1}\Phi\Omega_*^{-1}\right\}_- - \left\{\Lambda_*^{-1}\Phi\Omega_*^{-1}\right\}_\infty. \tag{B15}$$

However, with $K(s)$ forced to be analytic in $\mathrm{Re}\, s \geq 0$, the left-hand side of (B15) is also analytic in $\mathrm{Re}\, s\geq 0$ and equals the right-hand side which is analytic in $\mathrm{Re}\, s \leq 0$. Thus (B15) is polynomial and we obtain

$$K = \Lambda^{-1}J\Omega^{-1} + \Lambda^{-1}\left\{\Lambda_*^{-1}\Phi\Omega_*^{-1}\right\}_+\Omega^{-1}, \tag{B16}$$

$J(s)$ a real polynomial matrix to be determined by the requirement of finite cost. Observe that $K(s)$, as defined by (B16) is actually analytic in $\mathrm{Re}\, s\geq 0$ while

[34] Inequality (76) guarantees the analyticity of the factors $\Lambda^{-1}(s)$, $\Omega^{-1}(s)$ in $\mathrm{Re}\, s \geq 0$. Without (76) analyticity is assured only in $\mathrm{Re}\, s > 0$. It can be shown that the factors are unique up to real constant orthogonal multipliers [8], [9].

$$\Delta_* = \Lambda_*\left(\left\{\Lambda_*^{-1}\Phi\Omega_*^{-1}\right\}_- + \left\{\Lambda_*^{-1}\Phi\Omega_*^{-1}\right\}_\infty - J\right)\Omega_* \quad \text{(B17)}$$

is analytic in $\mathrm{Re}\, s \leqslant 0$ (as it should be). In 5) we imposed conditions guaranteeing the $j\omega$-analyticity of all integrands in E_s and E_t and to study the convergence of the cost under the choice (B16) for K it suffices to examine the behavior of the integrand of E as $\omega \to \infty$. Denote this integrand by $\rho(s)$. Noting that $R=(Y+A_1K)A=HA$ and

$$\begin{aligned}\mathrm{Tr}\left(H_*(P_*P+kQ)H(AGA_*)\right) &= \mathrm{Tr}\left[kQRGR_* + (PR)G(PR)_*\right] \\ &= \mathrm{Tr}\left(\Omega_* H_*(P_*P+kQ)H\Omega\right), \quad \text{(B18)}\end{aligned}$$

simple algebra yields

$$\begin{aligned}\rho = \mathrm{Tr}\left(\Omega_* H_*(P_*P+kQ)H\Omega\right) + \mathrm{Tr}\, G_u + \mathrm{Tr}\left(P_o G_d P_{o*}\right) \\ -2\,\mathrm{Tr}(PRG_u) - 2\,\mathrm{Tr}\left(PRP_d G_d P_{o*}\right). \quad \text{(B19)}\end{aligned}$$

To evaluate the first term in ρ we need H. From (B16) and (B3),

$$\begin{aligned}K = \Lambda^{-1}J\Omega^{-1} + \Lambda^{-1}\left\{\Lambda_*^{-1}I\Omega_*^{-1}\right\}_+\Omega^{-1} \\ -\Lambda^{-1}\left\{\Lambda A_1^{-1}Y\Omega\right\}_+\Omega^{-1} \quad \text{(B20)}\end{aligned}$$

where

$$I = A_{1*}P_*\left(G_u + P_o G_d P_{d*}\right)A_*. \quad \text{(B21)}$$

Multiplying (B20) on the left by A_1 and combining,

$$\Lambda A_1^{-1}H\Omega = J_1 + \left\{\Lambda_*^{-1}I\Omega_*^{-1}\right\}_+ + \left\{\Lambda A_1^{-1}Y\Omega\right\}_- \quad \text{(B22)}$$

where

$$J_1 = J + \left\{\Lambda A_1^{-1}Y\Omega\right\}_\infty \quad \text{(B23)}$$

is also polynomial. Since

$$\Omega_* H_*(P_*P+kQ)H\Omega = \left(\Lambda A_1^{-1}H\Omega\right)_*\left(\Lambda A_1^{-1}H\Omega\right), \quad \text{(B24)}$$

the integral of the first term in (B19) converges iff

$$\Lambda A_1^{-1}H\Omega \leqslant O(1/\omega), \qquad \omega\to\infty. \quad \text{(B25)}$$

Now both curly brackets in (B22) are already $\leqslant O(1/\omega)$ and, therefore, $J_1 \leqslant O(1/\omega)$. But being polynomial J_1 can only be $\leqslant O(1/\omega)$ for $\omega\to\infty$ if it is identically zero, whence

$$J = -\left\{\Lambda A_1^{-1}Y\Omega\right\}_\infty \quad \text{(B26)}$$

is identified. According to (B18) this convergence entails that of

$$\mathrm{Tr}\int_{-j\infty}^{j\infty}(PR)G(PR)_*\, ds \quad \text{(B27)}$$

which entails that of

$$\mathrm{Tr}\int_{-j\infty}^{j\infty}(PR)G_u(PR)_*\, ds, \quad \text{(B28)}$$

$$\mathrm{Tr}\int_{-j\infty}^{j\infty}(PR)G_{ml}(PR)_*\, ds \quad \text{(B29)}$$

and

$$\mathrm{Tr}\int_{-j\infty}^{j\infty}(PR)\left(P_d G_d P_{d*}\right)(PR)_*\, ds. \quad \text{(B30)}$$

We can exploit the integrability of

$$\mathrm{Tr}\left[H_*(P_*P+kQ)H(AGA_*)\right] = \mathrm{Tr}\left((P_*P+kQ)RGR_*\right) \quad \text{(B31)}$$

to derive a sharp sufficient condition for $T(s)=PR$ to be proper. Let

$$G(j\omega)\approx\omega^{2l}1_n, \quad \text{(B32)}$$

$$\left(P_*P+kQ\right)\approx\omega^{2q}1_m \quad \text{(B33)}$$

and

$$R(j\omega)=O(\omega^r) \quad \text{(B34)}$$

for $\omega\to\infty$. Then[35]

$$\left(P_*P+kQ\right)RGR_* = O\left(\omega^{2l+2q+2r}\right) \quad \text{(B35)}$$

and invoking integrability, $l+q+r\leqslant -1$. Thus $r\leqslant -(1+l+q)$ and if $P(s)=O(s^\nu)$, order $T=$ order $(PR)\leqslant \nu-(1+l+q)$. It follows that the constraint

$$\nu-1\leqslant l+q \quad \text{(B36)}$$

guarantees $T(s)$ proper. Stated differently, if

$$\left(P_*P+kQ\right)G\approx\omega^{2\mu}1_m, \quad \text{(B37)}$$

$$P(s)=O(s^\nu) \quad \text{(B38)}$$

and

$$\mu\geqslant\nu-1, \quad \text{(B39)}$$

then $T(s)$ is proper. Irrespective of (B39), the assumptions

$$G_u(j\omega)\leqslant O(1/\omega^2) \quad \text{(B40)}$$

and

$$P_o G_d P_{o*}\leqslant O(1/\omega^2) \quad \text{(B41)}$$

plus the finiteness of (B28) and (B30) imply $E<\infty$. For, using Schwartz's inequality,[36]

$$\begin{aligned}\left|\mathrm{Tr}\int(PRG_u)\,d\omega\right|^2 \leqslant \mathrm{Tr}\int(PR)G_u(PR)\,d\omega \\ \cdot\mathrm{Tr}\int G_u\,d\omega<\infty \quad \text{(B42)}\end{aligned}$$

[35]This conclusion is reached by making use of some properties of positive-definite matrices.

[36]$\left|\mathrm{Tr}\int F_1F_2\,dx\right|^2 \leqslant \mathrm{Tr}\int F_1F_1^*\,dx\cdot\mathrm{Tr}\int F_2F_2^*\,dx$.

and

$$|\mathrm{Tr}\int PRP_dG_dP_{o*}\,d\omega|^2 \leqslant \mathrm{Tr}\int PR\,(P_dG_dP_{d*})(PR)_*\,d\omega \cdot \mathrm{Tr}\int P_oG_dP_{o*}\,d\omega < \infty. \quad \text{(B43)}$$

(The range of integration is over $|\omega| > \omega_o$, ω_o sufficiently large.) Writing $G_d = K_dK_{d*}$, it is seen that (B30) is finite iff

$$PRP_dK_d = ((1_n - S_1)P_o + TL)K_d \leqslant O(1/\omega). \quad \text{(B44)}$$

From (B41), $P_oK_d \leqslant O(1/\omega)$ and substituting into (B44) we obtain

$$(S_1P_o - TL)K_d \leqslant O(1/\omega). \quad \text{(B45)}$$

Obviously, if

$$G_d(j\omega) \approx \omega^{-2i}1, \qquad i \leqslant 1, \quad \text{(B46)}$$

then $K_d(j\omega) \approx \omega^{-i}1$ and (B45) forces $S_1P_o - TL$ to be proper.

From $FPR = 1_n - S$ and $PRF = 1_n - S_1$ it is clear that

$$O(P) + O(R) + O(F) \leqslant -1 \quad \text{(B47)}$$

is a sufficient condition for limit $S(j\omega) =$ limit $S_1(j\omega) = 1_n$ as $\omega \to \infty$. Since $O(R) \leqslant -(1+\mu)$, (B47) is certainly valid if

$$O(P) + O(F) \leqslant \mu. \quad \text{(B48)}$$

It now follows from $R = CS$ that $O(C) = O(R)$ and, therefore, $\mu \geqslant -1$ guarantees $C(s)$ proper. Note that $S(j\omega) \to 1_n$ as $\omega \to \infty$ implies det $S(j\omega) \not\equiv 0$ and in particular det $(X - B_1K) \not\equiv 0$ because $S = (X - B_1K)A$. This means that $C(s)$, as defined by (34), makes sense.

Employing the formulas $R = CS = HA$, $S = 1_n$-FPR and (B22) with $J_1 = O$ we obtain

$$C = R(1_n - FPR)^{-1} = H(A^{-1} - FPH)^{-1} = H(1_n - BH)^{-1}A \quad \text{(B49)}$$

where

$$H = A_1\Lambda^{-1}\left(\left\{\Lambda_*^{-1}I\Omega_*^{-1}\right\}_+ + \left\{\Lambda A_1^{-1}Y\Omega\right\}_-\right)\Omega^{-1}. \quad \text{(B50)}$$

The product $A_1\Lambda^{-1}\{\Lambda A_1^{-1}Y\Omega\}_-\Omega^{-1}$ is obviously analytic in $\mathrm{Re}\,s \geqslant 0$ and the closed right-half-plane analyticity of $H(s)$ is, therefore, apparent.[37]

According to Appendix A, to study the stability margin of the optimally compensated loop it is necessary to find the zeros of the associated polynomial[38]

$$\Delta(s) = \frac{\Delta_F(s)\Delta_P(s)\Delta_C(s)}{\det S(s)}. \quad \text{(B51)}$$

Granting that any hidden modes of the plant and feedback compensator are known or at least localizable and that $C(s)$ shall be realized minimally,[39] it suffices instead to locate the zeros of the polynomial

$$\varphi(s) = \frac{\psi_F(s)\psi_P(s)\psi_C(s)}{\det S(s)} \quad \text{(B52)}$$

where $\psi_F(s)$, $\psi_P(s)$ and $\psi_C(s)$ are the characteristic denominators of $F(s)$, $P(s)$, and $C(s)$, respectively. Since

$$\det S = \det(X - B_1K)\cdot\det A \quad \text{(B53)}$$

and $\psi_F^+\psi_P^+ = \psi_{FP}^+ = \det{}^+A$ (by admissibility),

$$\varphi = \frac{\psi_F^-\psi_P^-\psi_C}{\det(X - B_1K)\cdot\psi_{FP}^-}. \quad \text{(B54)}$$

Evidently, ψ_{FP}^- divides $\psi_F^-\psi_P^-$ and

$$\varphi = \frac{\psi_C}{\det(X - B_1K)}\cdot\theta \quad \text{(B55)}$$

where

$$\theta(s) = \frac{\psi_F^-(s)\psi_P^-(s)}{\psi_{FP}^-(s)} \quad \text{(B56)}$$

is a strict Hurwitz polynomial. To make further progress we must relate ψ_C to $\det(X - B_1K)$. Let $K = NM^{-1}$ be a right-coprime factorization of $K(s)$. Then

$$C = (Y + A_1K)(X - B_1K)^{-1} = (YM + A_1N)(XM - B_1N)^{-1} \quad \text{(B57)}$$

and the pair $(YM + A_1N,\ XM - B_1N)$ is right-coprime. For the proof it is necessary to show (Appendix A) that the equations

$$(YM + A_1N)\boldsymbol{a} = \boldsymbol{O}, \quad \text{(B58)}$$

$$(XM - B_1N)\boldsymbol{a} = \boldsymbol{O} \quad \text{(B59)}$$

possess only the trivial solution $\boldsymbol{a} = \boldsymbol{O}$ which we accomplish by using the identity

$$A(XM - B_1N) + B(YM + A_1N) = M. \quad \text{(B60)}$$

Clearly, in view of (B60) any $\boldsymbol{a}$ satisfying (B58) and (B59) must also satisfy

$$M\boldsymbol{a} = \boldsymbol{O}, \quad \text{(B61)}$$

$$A_1N\boldsymbol{a} = \boldsymbol{O}; \quad B_1N\boldsymbol{a} = \boldsymbol{O}. \quad \text{(B62)}$$

Since the pair (A_1, B_1) is right-coprime, (B62) implies $N\boldsymbol{a} = \boldsymbol{O}$ and invoking (B61) and the right-coprimeness of (N, M), $\boldsymbol{a} = \boldsymbol{O}$. Q.E.D.

Hence, up to a multiplicative constant,

$$\psi_C = \det(XM - B_1N) \quad \text{(B63)}$$

and substituting into (B55),

$$\varphi(s) = \theta(s)\cdot\det M(s). \quad \text{(B64)}$$

[37] Unfortunately, the best numerical scheme for carrying out the computation (B50) is not so apparent.

[38] $\Delta(s)$ in (B51) has no connection with the $\Delta(s)$ appearing in (B2).

[39] If $C(s)$ is not realized minimally its hidden modes must also be localizable.

It is seen therefore that the (nonhidden) poles of the optimally compensated loop are *precisely* the zeros of $\theta(s)$ plus the finite poles of $K(s)$, each of the latter counted according to its McMillan degree.

An examination of the formula

$$K=\Lambda^{-1}\left(\left\{\Lambda_*^{-1}I\Omega_*^{-1}\right\}_+ + \left\{\Lambda A_1^{-1}Y\Omega\right\}_-\right)\Omega^{-1}-A_1^{-1}Y \tag{B65}$$

reveals immediately that the distinct finite poles of $K(s)$ are *included* in those of $A_1^{-1}(s)$, $\Lambda^{-1}(s)$, $\Omega^{-1}(s)$, and $I(s)$ in $\mathrm{Re}\, s<0$; or, in terms of primary data, in those of

$$FP,\ \left(A_{1*}(P_*P+kQ)A_1\right)^{-1},\ (AGA_*)^{-1},\ A_{1*}P_*\left(G_u+P_oG_dP_{d*}\right)A_*. \tag{B66}$$

Finally, instead of (B49), experience indicates that the formula

$$C=H_o\left(A^{-1}\Omega-FPH_o\right)^{-1} \tag{B67}$$

where

$$H_o=H\Omega=A_1\Lambda^{-1}\left(\left\{\Lambda_*^{-1}I\Omega_*^{-1}\right\}_+ + \left\{\Lambda A_1^{-1}Y\Omega\right\}_-\right) \tag{B68}$$

is more suitable for computer implementation. To complete the proof of Theorem 1, Section III, it is finally necessary to prove that the controller defined by (B67) and (B68) provides a global minimum for the cost E from among the class of all admissible controllers.

Combining (B16) and (B26), it is seen that

$$K_o(s)=\Lambda^{-1}\left(\left\{\Lambda_*^{-1}\Phi\Omega_*^{-1}\right\}_+ - \left\{\Lambda A_1^{-1}Y\Omega\right\}_\infty\right)\Omega^{-1} \tag{B69}$$

is the Wiener–Hopf solution for $K(s)$. Clearly, $K_o(s)$ is analytic in $\mathrm{Re}\, s\geqslant 0$ and as we have already shown in great detail, the associated cost

$$E(K_o)=E_s(K_o)+kE_t(K_o) \tag{B70}$$

obtained by substituting $R_o=(Y+A_1K_o)A$ into (65) and (66) is finite. According to Lemma 3, any $R(s)$ corresponding to a stable closed-loop design must be of the form $R=(Y+A_1K)A$ where $K(s)$ is analytic in $\mathrm{Re}\, s\geqslant 0$. Hence, for our present purposes we say that $K(s)$ is admissible if it is analytic in $\mathrm{Re}\, s\geqslant 0$ and the associated cost $E(K)<\infty$.[40] Our objective is to prove that $E(K)\geqslant E(K_o)$ for any choice of admissible $K(s)$.

Let

$$(R_1,R_2)_s\equiv \mathrm{Tr}\int_{-\infty}^{\infty} QR_1GR_{2*}\,d\omega, \tag{B71}$$

$$(R_1,R_2)_u\equiv \mathrm{Tr}\int_{-\infty}^{\infty} (1_n-PR_1)G_u(1_n-PR_2)_*\,d\omega, \tag{B72}$$

$$(R_1,R_2)_{ml}\equiv \mathrm{Tr}\int_{-\infty}^{\infty} (PR_1)G_{ml}(PR_2)_*\,d\omega \tag{B73}$$

and

$$(R_1,R_2)_d\equiv \mathrm{Tr}\int_{-\infty}^{\infty} (P_o-PR_1P_d)G_d(P_o-PR_2P_d)_*\,d\omega. \tag{B74}$$

In view of (66), (69), and (70),

$$2\pi E(K)=k(R,R)_s+(R,R)_u+(R,R)_{ml}+(R,R)_d. \tag{B75}$$

Moreover, since each of the four terms on the right-hand side of (B75) is nonnegative, $E(K)<\infty$ iff these terms are all finite.

Suppose R_1 and R_2 correspond to admissible choices K_1 and K_2, respectively. Then, $E(K_1)<\infty$, $E(K_2)<\infty$ and using the version of Schwartz's inequality given,[36] it is easily shown that

$$\begin{aligned}|(R_1,R_2)_s|^2&\leqslant (R_1,R_1)_s\cdot(R_2,R_2)_s<\infty,\\ |(R_1,R_2)_u|^2&\leqslant (R_1,R_1)_u\cdot(R_2,R_2)_u<\infty,\\ |(R_1,R_2)_{ml}|^2&\leqslant (R_1,R_1)_{ml}\cdot(R_2,R_2)_{ml}<\infty,\\ |(R_1,R_2)_d|^2&\leqslant (R_1,R_1)_d\cdot(R_2,R_2)_d<\infty.\end{aligned} \tag{B76}$$

For example, recalling that $Q=P_{s*}P_s$ and $G=\Omega\Omega_*$,

$$\begin{aligned}|(R_1,R_2)_s|^2&=\left|\mathrm{Tr}\int_{-\infty}^{\infty}(P_sR_1\Omega)(P_sR_2\Omega)_*\,d\omega\right|^2\\ &\leqslant \mathrm{Tr}\int_{-\infty}^{\infty}(P_sR_1\Omega)(P_sR_1\Omega)_*\,d\omega\\ &\quad\cdot\mathrm{Tr}\int_{-\infty}^{\infty}(P_sR_2\Omega)(P_sR_2\Omega)_*\,d\omega=\\ &=\mathrm{Tr}\int_{-\infty}^{\infty}QR_1GR_{1*}\,d\omega\cdot\mathrm{Tr}\int_{-\infty}^{\infty}QR_2GR_{2*}\,d\omega\\ &=(R_1,R_1)_s\cdot(R_2,R_2)_s<\infty.\end{aligned}$$

The other three inequalities are established in exactly the same way. (The result $\mathrm{Tr}AB=\mathrm{Tr}BA$ is used repeatedly.)

Identify R_1 with R_o and R_2 with any R defined by an admissible $K=K_o+\delta K$. Of course, $\delta K(s)$ is analytic in $\mathrm{Re}\, s\geqslant 0$ and $R=R_o+\delta R$ where $\delta R=A_1(\delta K)A$. Since

$$(R_o,R)_s=(R_o,R_o)_s+(R_o,\delta R)_s$$

is finite and $(R_o,R_o)_s<\infty$, it is also true that $\alpha_o=(R_o,\delta R)_s<\infty$. Similarly,

$$\begin{aligned}\alpha_1&=(R_o,\delta R)_{ml}<\infty,\\ \alpha_2&=\mathrm{Tr}\int_{-\infty}^{\infty}(1_n-PR_o)G_u(P\delta R)_*\,d\omega<\infty\end{aligned} \tag{B77}$$

and

$$\alpha_3=\mathrm{Tr}\int_{-\infty}^{\infty}(P_o-PR_oP_d)G_dP_{d*}(P\delta R)_*\,d\omega<\infty.$$

[40]The argument that follows is independent of the assumption $\det(X-B_1K)\not\equiv 0$ in (33).

From

$$(R,R)_s=(R_o,R_o)_s+2(R_o,\delta R)_s+(\delta R,\delta R)_s$$

it now follows that $\beta_o=(\delta R,\delta R)_s<\infty$. In the same manner, exploiting the remaining inequalities in (B77), we get

$$\beta_1=\operatorname{Tr}\int_{-\infty}^{\infty}(P\delta R)G_{ml}(P\delta R)_*\,d\omega<\infty,$$

$$\beta_2=\operatorname{Tr}\int_{-\infty}^{\infty}(P\delta R)G_u(P\delta R)_*\,d\omega<\infty \qquad \text{(B78)}$$

and

$$\beta_3=\operatorname{Tr}\int_{-\infty}^{\infty}(P\delta R)P_dG_dP_{d*}(P\delta R)_*\,d\omega<\infty.$$

Clearly, all four β's are nonnegative.

Let $E(K)=E(K_o)+\delta E$. By a straightforward expansion of (B75),

$$2\pi(\delta E)=2(k\alpha_o+\alpha_1-\alpha_2-\alpha_3)+(k\beta_o+\beta_1+\beta_2+\beta_3). \qquad \text{(B79)}$$

However, it is readily verified by grouping terms that

$$-j(k\alpha_o+\alpha_1-\alpha_2-\alpha_3)=\operatorname{Tr}\int_{-j\infty}^{j\infty}\Delta_*(\delta K)_*\,ds \qquad \text{(B80)}$$

where $\Delta_*(s)$ is as defined in (B2) and (B3) and $K(s)$ replaced by $K_o(s)$. Now the Wiener–Hopf solution $K_o(s)$ guarantees the analyticity of $\Delta_*(s)$ in Re $s\leqslant 0$ and the finiteness of the α's implies that of the integral. The integrand

$$\operatorname{Tr}\left(\Delta_*(\delta K)_*\right)$$

is therefore analytic in Re $s\leqslant 0$ and $O(1/\omega^2)$ for large ω^2. By Cauchy's theorem the integral equals zero whence, $k\alpha_o+\alpha_1-\alpha_2-\alpha_3=0$ and

$$2\pi(\delta E)=k\beta_o+\beta_1+\beta_2+\beta_3\geqslant 0.$$

Consequently, $E(K)\geqslant E(K_o)$ for every admissible $K(s)$. Q.E.D.

References

[1] J. E. Weston and J. J. Bongiorno, Jr., "Extension of analytical design techniques to multivariable feedback control systems," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 613–620, Oct. 1972.

[2] D. C. Youla, "Modern classical multivariable feedback control theory: Part I," Rome Air Development Center, Griffiss Air Force Base, NY, Tech Rep. RADC-TR-70-98, June 1970.

[3] D. C. Youla, J. J. Bongiorno, Jr., and C. N. Lu, "Single-loop feedback stabilization of linear multivariable dynamical plants," *Automatica*, vol. 10, pp. 159–173, 1974.

[4] D. C. Youla, J. J. Bongiorno, Jr., and H. A. Jabr, "Modern Wiener–Hopf design of optimal controllers—Part I: The single-input-output case," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 3–13, Feb. 1976.

[5] H. W. Bode, *Network Analysis and Feedback Design*. New York: Van Nostrand, 1945.

[6] D. C. Youla, "Network ideas applied to the design of optimal multivariable feedback controllers," Polytechnic Inst. of New York, Farmingdale, Memo POLY-EE/EP-75-146, Dec. 1973.

[7] G. C. Newton, Jr., L.A. Gould, and J. F. Kaiser, *Analytical Design of Linear Feedback Controls*. New York: Wiley, 1957.

[8] D. C. Youla, "On the factorization of rational matrices," *IRE Trans. Inform. Theory*, vol. IT-7, pp. 172–189, July 1961.

[9] W. G. Tuel, Jr., "Computer algorithm for spectral factorization of rational matrices," *IBM J. Res. Develop.*, vol. 12, pp. 163–170, Mar. 1968.

[10] H. H. Rosenbrock, *State-Space and Multivariable Theory*. New York: Wiley-Interscience, 1970.

[11] R. E. Kalman, "Irreducible realizations and the degree of a rational matrix," *J. Soc. Industrial and Applied Mathematics*, vol. 13, pp. 520–544, June 1965.

[12] H. H. Rosenbrock, "Relatively prime polynomial matrices," *Electron. Lett.*, no. 4, pp. 227–228, 1968.

[13] K. Hoffman and R. Kunze, *Linear Algebra*. Englewood Cliffs, NJ: Prentice-Hall, 1961.

[14] R. T. Yanushevskii, "Synthesis of closed-loop multivariable control systems for a certain class of plants with delay," *Automation and Remote Control*, pp. 1394–1402, Feb. 1975; Russian original, vol. 35, no. 9, pt.1, Sept. 1974.

[15] I. M. Horowitz and U. Shaked, "Superiority of transfer function over state-variable methods in linear time-invariant feedback system design," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 84–97, Feb. 1975.

[16] H. A. Jabr, "Modern analytical design of optimal multivariable control systems," Ph.D. dissertation, Polytechnic Institute of New York, Farmingdale, 1975.

[17] *IEEE Trans. Automat. Contr. (Special Issue on Linear-Quadratic-Gaussian Problem)*, vol. AC-16, Dec. 1971.

[18] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. New York: Wiley-Interscience, 1972.

[19] I. M. Horowitz, *Synthesis of Feedback Systems*. New York: Academic, 1963.

Dante C. Youla (SM'59–F'75), for a photograph and biography, see page 13 of the February 1976 issue of this TRANSACTIONS.

Hamid A. Jabr (S'73–M'76), for a photograph and biography, see page 13 of the February 1976 issue of this TRANSACTIONS.

Joseph J. Bongiorno, Jr. (S'56–M'60), for a photograph and biography, see page 13 of the February 1976 issue of this TRANSACTIONS.