INTERNATIONAL JOURNAL OF ROBUST AND NONLINEAR CONTROL, VOL. 1, 229–294 (1991)

A FACTORIZATION PRINCIPLE FOR STABILIZATION OF LINEAR CONTROL SYSTEMS

JOSEPH A. BALL
Department of Mathematics, Virginia Tech., Blacksburg, VA 24061, U.S.A.

J. WILLIAM HELTON
Department of Mathematics, University of California, San Diego, La Jolla, CA 92093, U.S.A.

AND

MADANPAL VERMA
Department of Electrical Engineering, McGill University, Montreal, Quebec, Canada H3A 2KC

SUMMARY

By introducing a fictitious signal y^0 if necessary we define a transform

$$\tilde{\mathscr{P}}: \begin{bmatrix} u \\ y \\ y^0 \end{bmatrix} \rightarrow \begin{bmatrix} z \\ w \end{bmatrix}$$

of a given linear control system

$$\mathscr{P}: \begin{bmatrix} w \\ u \end{bmatrix} \rightarrow \begin{bmatrix} z \\ y \end{bmatrix}$$

which generalizes the passage from the scattering to the chain formalism in circuit theory. Given a factorization $\tilde{\mathscr{P}} = \Theta R$ of $\tilde{\mathscr{P}}$ where R is a block matrix function with a certain key block equal to a minimal phase (or outer) matrix function, we show that a given compensator $u = Ky$ is internally stabilizing for the system $\mathscr{P}$ if and only if a related compensator K' is stabilizing for Θ. Factorizations $\tilde{\mathscr{P}} = \Theta R$ with Θ having a certain block upper triangular form lead to an alternative derivation of the Youla parametrization of stabilizing compensators. Factorizations with Θ equal to a J-inner matrix function (in a precise weak sense) lead to a parametrization of all solutions K of the H^∞ problem associated with $\mathscr{P}$. This gives a new solution of the H^∞ problem completely in the transfer function domain. Computation of the needed factorization $\tilde{\mathscr{P}} = \Theta R$ in terms of a state-space realization of $\mathscr{P}$ leads to the state-space formulas for the solution of the H^∞ problem recently obtained in the literature.

KEY WORDS Feedback stabilization H^∞ control J-inner–outer factorization J-spectral factorization

0. INTRODUCTION

A standard general feedback configuration in terms of which many problems of interest in

This paper was recommended for publication by editor M. J. Grimble

1049–8923/91/040229–66$33.00

Received 7 July 1991
Revised September 1991

control theory can be posed is given in Figure 1 (see Reference 1 or 2). Here we assume that

$$\mathscr{P}:\begin{bmatrix} w \\ u \end{bmatrix} \rightarrow \begin{bmatrix} z \\ y \end{bmatrix}$$

is the input–output map for a linear, time-invariant, finite-dimensional system; after transforming to the frequency domain by using the Laplace transform, we may assume that the operator $\mathscr{P}$ is multiplication by a block matrix function

$$\mathscr{P}(s) = \begin{bmatrix} \mathscr{P}_{11}(s) & \mathscr{P}_{12}(s) \\ \mathscr{P}_{21}(s) & \mathscr{P}_{22}(s) \end{bmatrix} \tag{1}$$

The signals w, u, z, y take values in finite dimensional linear spaces W, U, Z, Y; the signal w is variously known as the *command*, *reference* or *disturbance* signal depending on the application, z is the *error* signal, y is the *measurement* signal, and u is the *control* signal. Given $\mathscr{P}$, a first design objective is to build the compensator K so that the closed loop system is *internally stable*, i.e. so that the output z and all internal signals u and y are stable (i.e. have all poles in the left half plane) whenever the input w and disturbances v_1 and v_2 are stable (see Figure 2). This is a formulation of internal stability completely at the input–output transfer function level; other formulations in terms of state-space realizations are also possible.[1,2]

The main result of this paper is the definition of an equivalence relation $\sim$ between two systems $\Sigma = \Sigma(\mathscr{P}, K)$ and $\Sigma' = \Sigma(\mathscr{P}', K')$ such that Σ is internally stable if and only if Σ' is internally stable whenever Σ and Σ' are equivalent. The notion of equivalence is in terms of factorization of a matrix function $\tilde{\mathscr{P}}$ obtained from $\mathscr{P}$ by rearranging the system of equations

$$\begin{aligned} \mathscr{P}_{11}w + \mathscr{P}_{12}u &= z \\ \mathscr{P}_{21}w + \mathscr{P}_{22}u &= y \end{aligned} \tag{2}$$

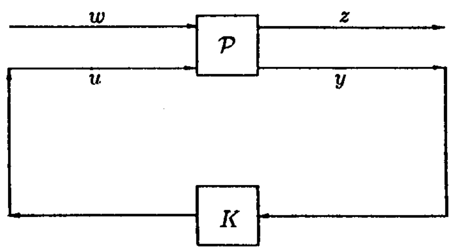

Figure 1

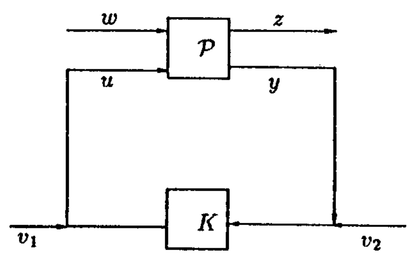

Figure 2

associated with $\mathscr{P}$. A standing assumption in the theory (often referred to as the *regular case*) is that $\mathscr{P}_{12}(s)$ and $\mathscr{P}_{21}(s)$ are injective and surjective respectively on the extended imaginary line. If $\mathscr{P}_{21}(s)$ is in fact square and invertible, one can rearrange the system of equations (2) to be of the form

$$\begin{aligned} \tilde{\mathscr{P}}_{11}u + \tilde{\mathscr{P}}_{12}y &= z \\ \tilde{\mathscr{P}}_{21}u + \tilde{\mathscr{P}}_{22}y &= w \end{aligned} \tag{3}$$

where

$$\begin{bmatrix} \tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{12} \\ \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} \end{bmatrix} = \begin{bmatrix} \mathscr{P}_{12} - \mathscr{P}_{11}\mathscr{P}_{21}^{-1}\mathscr{P}_{22} & \mathscr{P}_{11}\mathscr{P}_{21}^{-1} \\ -\mathscr{P}_{21}^{-1}\mathscr{P}_{22} & \mathscr{P}_{21}^{-1} \end{bmatrix} \tag{4}$$

A new observation and tool here is that, if $\mathscr{P}_{21}$ is not square and invertible but merely surjective, one may add a row $[\mathscr{P}_{21}^0 \quad \mathscr{P}_{22}^0]$ to $\mathscr{P}$ such that the augmented (2, 1)-block

$$\begin{bmatrix} \mathscr{P}_{21} \\ \mathscr{P}_{21}^0 \end{bmatrix}$$

is square and invertible. This induces an augmented system of equations

$$\begin{aligned} \mathscr{P}_{11}w + \mathscr{P}_{12}u &= z \\ \mathscr{P}_{21}w + \mathscr{P}_{22}u &= y \\ \mathscr{P}_{21}^0 w + \mathscr{P}_{22}^0 u &= y^0 \end{aligned} \tag{5}$$

in which y^0 is to be thought of as a fictitious output signal inserted for mathematical convenience. The system (5) can be rearranged in the form

$$\begin{aligned} \tilde{\mathscr{P}}_{11}u + \tilde{\mathscr{P}}_{12}y + \tilde{\mathscr{P}}_{13}y^0 &= z \\ \tilde{\mathscr{P}}_{21}u + \tilde{\mathscr{P}}_{22}y + \tilde{\mathscr{P}}_{23}y^0 &= w \end{aligned} \tag{6}$$

where

$$\begin{bmatrix} \tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{12} & \tilde{\mathscr{P}}_{13} \\ \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} & \tilde{\mathscr{P}}_{23} \end{bmatrix} = \begin{bmatrix} \mathscr{P}_{12} - \mathscr{P}_{11}\begin{bmatrix} \mathscr{P}_{21} \\ \mathscr{P}_{21}^0 \end{bmatrix}^{-1}\begin{bmatrix} \mathscr{P}_{22} \\ \mathscr{P}_{22}^0 \end{bmatrix} & \mathscr{P}_{21}\begin{bmatrix} \mathscr{P}_{21} \\ \mathscr{P}_{21}^0 \end{bmatrix}^{-1} \\ -\begin{bmatrix} \mathscr{P}_{21} \\ \mathscr{P}_{21}^0 \end{bmatrix}^{-1}\begin{bmatrix} \mathscr{P}_{22} \\ \mathscr{P}_{22}^0 \end{bmatrix} & \begin{bmatrix} \mathscr{P}_{21} \\ \mathscr{P}_{21}^0 \end{bmatrix}^{-1} \end{bmatrix} \tag{7}$$

The following is our basic equivalence principle for stability of feedback systems. A special case was given in Reference 3. Recall that a rational matrix function R_0 is said to be *minimal phase* (or *outer* in the mathematical literature) if both R_0 and R_0^{-1} are stable.

Theorem A (see Theorems 2.1.1 and 2.2.1)

Let $\mathscr{P}$ and $\mathscr{P}'$ be two given plants of the form as in Figures 1 and 2. Let $\tilde{\mathscr{P}}$ and $\tilde{\mathscr{P}}'$ be the associated transforms as in (7) and suppose that

$$\tilde{\mathscr{P}} = \tilde{\mathscr{P}}'R \tag{8}$$

where

$$R: \begin{bmatrix} u \\ y \\ y^0 \end{bmatrix} \rightarrow \begin{bmatrix} u' \\ y' \\ y^{0'} \end{bmatrix}$$

has the form

$$R = \begin{bmatrix} R_{11} & R_{12} & 0 \\ R_{21} & R_{22} & 0 \\ R_{31} & R_{32} & R_{33} \end{bmatrix}$$

with

$$R_0 = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

a minimal phase (outer) rational matrix function and R_{31}, R_{32}, R_{33}, R_{33}^{-1} rational matrix functions analytic on the extended imaginary line. Suppose K is a proper rational matrix function such that $(R_{21}K + R_{22})^{-1}$ is also proper. Then the compensator K is stabilizing for $\mathscr{P}$ if and only if the compensator $K' = (R_{11}K + R_{12})(R_{21}K + R_{22})^{-1}$ is stabilizing for $\mathscr{P}'$.

We shall call matrix functions R of the type described in Theorem A *restricted outer* matrix functions.

The practical utility of Theorem A is that a given control problem for a complicated plant $\mathscr{P}$ can be reduced to the same problem for a much simpler plant $\mathscr{P}'$ if $\mathscr{P}$ and $\mathscr{P}'$ are related as in Theorem A. We offer two illustrations of this general principle.

The set of stabilizing compensators K for a given plant $\mathscr{P}$ in general is a complicated nonlinear set; a basic issue therefore is the *stabilization problem.*

(STAB) For a given plant $\mathscr{P}$ describe the set of all stabilizing compensators K

A special case where (STAB) is easy is the case where $\mathscr{P}$ has the form

$$\mathscr{P} = \begin{bmatrix} T_1 & -T_2 \\ T_3 & 0 \end{bmatrix}$$

associated with a model matching problem.[1] A plant $\mathscr{P}$ of this type we shall say is in *model matching form*. In this case one can show that $\mathscr{P}$ is stabilizable if and only if T_1, T_2 and T_3 are all stable, and then K stabilizes if and only if K itself is stable. If $\mathscr{P}$ has the model matching form

$$\mathscr{P} = \begin{bmatrix} T_1 & -T_2 \\ T_3 & 0 \end{bmatrix}$$

then the transform $\tilde{\mathscr{P}}$ has the block upper triangular form

$$\tilde{\mathscr{P}} = \begin{bmatrix} \tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{12} & \tilde{\mathscr{P}}_{13} \\ 0 & \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} \end{bmatrix}$$

Thus we have

Corollary. (Youla parametrization)

Reduction of (STAB) for a general plant $\mathscr{P}$ to the case of a model matching plant

$$\mathscr{P}' = \begin{bmatrix} T_1 & -T_2 \\ T_3 & 0 \end{bmatrix}$$

is equivalent to the computation of an (upper triangular)–(restricted outer) factorization

$$\begin{bmatrix} \tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{22} & \tilde{\mathscr{P}}_{23} \\ \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} & \tilde{\mathscr{P}}_{23} \end{bmatrix} = \begin{bmatrix} \tilde{\mathscr{P}}'_{11} & \tilde{\mathscr{P}}'_{12} & \tilde{\mathscr{P}}'_{13} \\ 0 & \tilde{\mathscr{P}}'_{22} & \tilde{\mathscr{P}}'_{23} \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & 0 \\ R_{21} & R_{22} & 0 \\ R_{31} & R_{32} & R_{33} \end{bmatrix}$$

of $\tilde{\mathscr{P}}$.

In this way we recover the Youla parametrization of stabilizing compensators (see Reference 1 or 2) as a direct application of the general factorization principle Theorem A.

The second illustration of Theorem A to be discussed in this paper is the standard problem of H^∞ control. This problem has received a lot of attention and interest in the literature in the past decade; for engineering motivation and background, we refer to References 4, 1 and 2 and the references therein. We formulate here the strictly suboptimal version of the problem: given the plant $\mathscr{P}$ and a tolerance level γ, one seeks to describe all stabilizing compensators K for which the induced operator norm of the closed-loop transfer function $T_{zw} = \mathscr{P}_{11} + \mathscr{P}_{12}K(I - \mathscr{P}_{22}K)^{-1}\mathscr{P}_{21}$ (as an operator on vector-valued L^2 of the imaginary line) is strictly less than γ.

(HINFγ) Find all stabilizing K such that $\| \mathscr{P}_{11} + \mathscr{P}_{12}K(I - \mathscr{P}_{22}K)^{-1}\mathscr{P}_{21} \|_\infty < \gamma$

Without loss of generality in the following discussion we may set $\gamma = 1$.

As before in the (STAB) problem, there is a special class of plants $\mathscr{P}$ as in (1) for which the solution of the strictly suboptimal H^∞ problem is easy. We say that the rational matrix function

$$\mathscr{P} = \begin{bmatrix} \mathscr{P}_{11} & \mathscr{P}_{12} \\ \mathscr{P}_{21} & \mathscr{P}_{22} \end{bmatrix}$$

as in (1) is *sub-all-pass* if

$$\| \mathscr{P}(s) |_{\left[\begin{smallmatrix} \operatorname{Ker} \mathscr{P}_{21}(s) \\ 0 \end{smallmatrix}\right]} \| < 1 \tag{9}$$

and

$$\mathscr{P}(s) |_{\left\{\left[\begin{smallmatrix} \operatorname{Ker} \mathscr{P}_{21}(s) \\ 0 \end{smallmatrix}\right]\right\}^\perp} \tag{10}$$

is isometric for s on the extended imaginary line. If in addition $\mathscr{P}(s)$ is stable, then we say that $\mathscr{P}$ is *subinner*. For example, if Ker $\mathscr{P}_{21}(s) = \{0\}$ for all s, then $\mathscr{P}$ is subinner if and only if $\mathscr{P}$ is a matrix inner function in the usual sense. It turns out that for a sub-all-pass plant $\mathscr{P}$, the strictly suboptimal H^∞ problem has solutions if and only if $\mathscr{P}$ is subinner, and in this case a compensator K solves the H^∞ problem if and only if K is stable with $\| K \|_\infty < 1$.

The transform $\tilde{\mathscr{P}}$ given by (7) of a sub-all-pass function $\mathscr{P}$ (i.e., a $\mathscr{P}$ satisfying (8) and (9)) can be arranged to have the property

$$\tilde{\mathscr{P}}(s)^* J_2 \tilde{\mathscr{P}}(s) = J_1, \qquad \operatorname{Re} s = 0 \tag{11}$$

and conversely, where $J_2 = I_Z \oplus - I_W$ and $J_1 = I_U \oplus - I_Y \oplus - I_{Y^0}$. Let us call such a matrix function $\tilde{\mathscr{P}}$ a (J_1, J_2)-isometry. The additional stability property required for $\mathscr{P}$ to be subinner then corresponds to

$$\tilde{\mathscr{P}}(s)^* J \tilde{\mathscr{P}}(s) \leqslant \begin{bmatrix} I_U & 0 & 0 \\ 0 & -I_Y & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \operatorname{Re} s > 0 \tag{12}$$

Let us say that a $\tilde{\mathscr{P}}$ satisfying (11) and (12) is a (J_1, J_2)-subinner matrix function. Note that the concept of (J_1, J_2)-subinner requires a specification of the space Y^0 of fictitious signals; when this space is absent, then the concept of (J_1, J_2)-subinner collapses to that of (J_1, J_2)-inner in the usual sense.[5] An application of Theorem A now leads to the following solution of the H^∞ problem.

Corollary

The solution of the H^∞ control problem for a given plant $\mathscr{P}$ reduces to constructing a factorization $\tilde{\mathscr{P}} = \tilde{\mathscr{P}}'R$ of the transform $\tilde{\mathscr{P}}$ of $\mathscr{P}$ (given by (7)) such that $\tilde{\mathscr{P}}'$ is (J_1, J_2)-subinner and R is restricted outer.

This gives a direct analysis of the H^∞ problem completely at the level of transfer functions and input–output operators. When one goes on to compute the $((J_1, J_2)$-subinner)–(restricted outer) factorization $\tilde{\mathscr{P}} = \tilde{\mathscr{P}}'R$ in terms of state-space realizations, one arrives at the elegant state-space formulas for the solution of the H^∞ problem recently obtained in the literature; this is done in Section 6.

Thus the general factorization principle Theorem A provides a unifying general framework from which the Youla parametrization of the set of stabilizing compensators and the parameterization of all solutions of the H^∞ problem both flow as particular illustrations. In particular we obtain a new direct conceptual solution of the H^∞ problem at the level of transfer functions which bypasses completely the Youla parametrization.

The ideas behind Theorem A for the case where $\mathscr{P}_{21}$ is square and invertible can be illustrated quite simply by a picture (see Figure 3). Here the original plant $\mathscr{P}$ is the transfer function

$$\text{from } \begin{bmatrix} w \\ u \end{bmatrix} \text{ to } \begin{bmatrix} z \\ y \end{bmatrix}$$

and it has been factored as in equation (8) of Theorem A where

$$\mathscr{P}' : \begin{bmatrix} w \\ u' \end{bmatrix} \to \begin{bmatrix} z \\ y' \end{bmatrix}$$

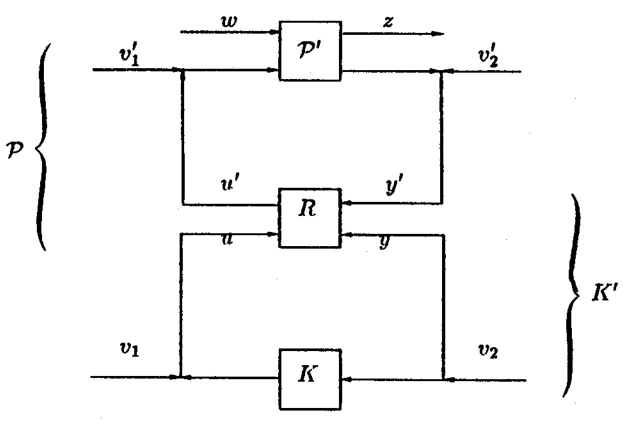

Figure 3

(with $v_1 = 0, v_2' = 0$) and where it is assumed that

$$R: \begin{bmatrix} u \\ y \end{bmatrix} \to \begin{bmatrix} u' \\ y' \end{bmatrix}$$

is outer. Figure 3 depicts this factorization as applied to Figure 2. Internal stability for the original closed loop system (Figure 2) associated with $\mathscr{P}$ and compensator K is defined with respect to the disturbance signals v_1 and v_2 (with v_1' and v_2' taken equal to zero). Internal stability for the closed loop system associated with the plant $\mathscr{P}'$ and compensator $K' = \mathscr{G}_R[K] : y' \to u'$ is defined with respect to the disturbance signals v_1' and v_2' (with $v_1 = 0, v_2 = 0$). The content of Theorem A is that these two notions of internal stability are the same, i.e. one can shift the location of the disturbance signals without changing the problem, in case R is outer. This was one of the basic ideas behind the proof of the Youla parametrization appearing in Reference 3. This idea can be made precise for the general case where $\mathscr{P}_{21}$ is not invertible but is made more complicated by the necessity of introducing an auxiliary space Y^0 of fictitious signals.

Historically the Youla parametrization was used as a tool for the study of the H^∞ problem. Specifically, the Youla parametrization enables one to reduce the H^∞ problem for the general case to the model-matching case. Further reductions and manipulations were required in order to apply the state-space solution of the Nehari problem from Reference 25. After having done all this, one still had to back-solve for the compensator, so the whole process was rather cumbersome. Later Reference 7 showed show to get a more direct parametrization for the set of all performances T_{zw} for an H^∞ problem directly via a J-spectral factorization procedure for a plant $\mathscr{P}$ assumed to be in the model-matching form. The landmark paper (Reference 4) avoided the Youla parametrization completely and was the first to set down a clean state-space solution of the general H^∞ problem in terms of solutions of two Riccati equations. Solutions of parallel problems involving optimization in a 2-norm rather than infinity-norm and strong analogies with LQG theory were also given with all the analysis done in the time domain in terms of a state-space representation of the original plant $\mathscr{P}$. Among the many follow-up accounts we mention Reference 8, which used the bounded real lemma, and Reference 9 which used the Pontrjagin maximum principle. The recent papers (References 6 and 10) returned to the transfer function domain to derive the solution of Reference 4 via factorization of transfer functions and a streamlined application of the Youla parametrization. For the 2-block case and in the context of the H^∞ problem, this paper has considerable overlap with Reference 26.

The formulas we arrive at here for the solution of the H^∞ problem are the same as in these other papers. The contribution of this paper to the H^∞ control theory is conceptual rather than computational. As in the original approach, our solution is set at the level of transfer functions with state-space formulas arrived at as a method of implementing transfer function operations. The novelty of our approach is that all the factorizations and manipulations of the original approach are condensed into a single generalized J-inner–outer factorization for the general regular case. This paper can also be viewed as a more definite improvement of Reference 7. There the solution of the H^∞ problem, under the assumption that the plant had already been brought to model-matching form, was reduced to the same type of single J-inner–outer factorization as we have here. The derivation in Reference 7 was via a more sophisticated Krein space analysis based on ideas from Reference 11 rather than the elementary ideas behind Theorem A as presented here. Also in Reference 7 the state-space analysis did not include the existence criterion; there appeared a superfluous Riccati equation, and the parametrization of the compensators K (as opposed to the closed loop transfer function T_{zw}) was left obscure (see the end of Section 4).

Since this paper was written, we found that the same factorization idea for a solution of the standard H^∞ problem at the frequency-domain level also appears in the recent work of Liu, Mita and Kimura.[12] Also a preliminary form of some of the ideas here appeared in work of Verma and Zames.[27]

This paper deals only with the so-called *regular* H^∞ problem and does not touch the singular case studied in Reference 13; an interesting direction for future research would be to understand the singular case from the factorization point of view of this paper. Also, although not emphasized here, the transfer function analysis in Sections 1–4 makes sense for distributed parameter systems.

We expect that the computation of the generalized J-inner–outer factorization presented in Section 5 will eventually have some technical improvement. Specifically, one should be able to prove that existence of the desired generalized J-inner–outer factorization guarantees that stabilizing solutions of the Riccati equations exist if (A, B) is stabilizable and (C, A) is detectable; these improvements are worked out in a separate report[14] for the 2-block case (corresponding to $\mathscr{P}_{21}(s)$ is square and invertible on the extended imaginary line in the control problem.) In any case the work of Section 5 extends the work of References 15 and 16 on state-space formulas for inner–outer factorization to more complicated factorization problems. The payoff is that one thereby solves completely in one stroke the entire H^∞ problem for a given plant $\mathscr{P}$ rather than merely implementing one of many steps in a solution algorithm, a motivation behind the work of References 15 and 16.

The paper is organized as follows: Section 1 systematically develops basic properties of the general transform $\mathscr{P} \to \tilde{\mathscr{P}}$ and of the associated linear fractional maps. Section 2 presents a proof of Theorem A, with first a separate proof for the simpler case where $\mathscr{P}_{21}$ is square and invertible. In Section 3 we show how the Youla parametrization of stabilizing compensators follows from an (upper triangular)–(restricted outer) factorization of the transformed plant $\tilde{\mathscr{P}}$ as a consequence of Theorem A. Section 4 shows how the H^∞ problem can be solved by a single generalized J-inner–outer factorization of the transformed plant $\tilde{\mathscr{P}}$. Next Section 5 shows how necessary and sufficient conditions for the existence of such a J-inner–outer factorization together with formulas for the factors can be given in terms of a state-space realization of $\tilde{\mathscr{P}}$. Finally in Section 6 we specialize this analysis to a function $\tilde{\mathscr{P}}$ of the form arising from an H^∞ control problem to recover the state-space formulas for the solution of an H^∞ control problem given by Petersen *et al.*[8]

1. LINEAR FRACTIONAL MAPS

In this section we develop some basic principles concerning matrix linear fractional transformations which shall be needed in the sequel.

1.1. Linear fractional maps on matrices

Suppose $\mathscr{P}$ is a block 2×2 matrix

$$\mathscr{P} = \begin{bmatrix} \mathscr{P}_{11} & \mathscr{P}_{12} \\ \mathscr{P}_{21} & \mathscr{P}_{22} \end{bmatrix}$$

where $\mathscr{P}_{ij}$ has size $m_i \times n_j$ for $i, j = 1, 2$. In this section we consider only matrices over the field of complex numbers $\mathbb{C}$; the results of this section will be applied pointwise to the case where the entries of $\mathscr{P}$ are rational matrix functions. The linear fractional map induced by $\mathscr{P}$ is the map

$$K \to \mathscr{P}_{11} + \mathscr{P}_{12}K(I - \mathscr{P}_{22}K)^{-1}\mathscr{P}_{21} =: \mathscr{F}_{\mathscr{P}}[K]$$

transforming $n_2 \times m_2$ matrices K to $m_1 \times n_1$ matrices $\mathscr{F}_{\mathscr{P}}[K]$; $\mathscr{F}_{\mathscr{P}}[K]$ is defined for all $n_2 \times m_2$ matrices K for which the inverse of $K - \mathscr{P}_{22}K$ exists, so in particular $\mathscr{F}_{\mathscr{P}}[0] = \mathscr{P}_{11}$ is always defined. It is easy to see that the map $K \to \mathscr{F}_{\mathscr{P}}[K]$ is one-to-one exactly when

$$\mathscr{P}_{12} \text{ has linearly independent columns (i.e. } \mathscr{P}_{12} \text{ is injective)} \tag{13}$$

and

$$\mathscr{P}_{21} \text{ has linear independent rows (i.e. } \mathscr{P}_{21} \text{ is surjective)} \tag{14}$$

When assumptions (13) and (14) hold, we shall say that the map $\mathscr{F}_{\mathscr{P}}$ is *regular*; in this section we shall deal exclusively with regular maps. The map $\mathscr{F}_{\mathscr{P}}$ can also be viewed as the formula for the input–output map from w to z in the feedback signal flow diagram depicted in Figure 4. Algebraically, if one substitutes $u = Ky$ in the system of equations

$$\begin{bmatrix} \mathscr{P}_{11} & \mathscr{P}_{12} \\ \mathscr{P}_{21} & \mathscr{P}_{22} \end{bmatrix} \begin{bmatrix} w \\ u \end{bmatrix} = \begin{bmatrix} z \\ y \end{bmatrix}$$

and solves for z in terms of w, one obtains

$$z = \mathscr{F}_{\mathscr{P}}[K]\, w$$

The next lemma describes the extent to which a regular map $K \to \mathscr{F}_{\mathscr{P}}[K]$ determines the matrix $\mathscr{P}$ which induces it. We omit the elementary proof.

Lemma 1.1.1

Suppose

$$\mathscr{P}^k = \begin{bmatrix} \mathscr{P}^k_{11} & \mathscr{P}^k_{12} \\ \mathscr{P}^k_{21} & \mathscr{P}^k_{22} \end{bmatrix}$$

is a block 2×2 matrix with $\mathscr{P}^k_{12}$ injective and $\mathscr{P}^k_{21}$ surjective and with $\mathscr{P}^k_{ij}$ of size $m_i \times n_j$ for $k = 1, 2$. Then $\mathscr{F}_{\mathscr{P}^1}[K] = \mathscr{F}_{\mathscr{P}^2}[K]$ for all $n_2 \times m_2$ matrices K in a neighbourhood of 0 if an only if

$$\begin{bmatrix} \mathscr{P}^2_{11} & \mathscr{P}^2_{12} \\ \mathscr{P}^2_{22} & \mathscr{P}^2_{22} \end{bmatrix} = \begin{bmatrix} \mathscr{P}^1_{11} & \alpha\mathscr{P}^1_{12} \\ \alpha^{-1}\mathscr{P}^1_{21} & \mathscr{P}^1_{22} \end{bmatrix}$$

for some non-zero scalar α.

To be consistent with terminology from H^∞ control theory[1] we identify three types of linear

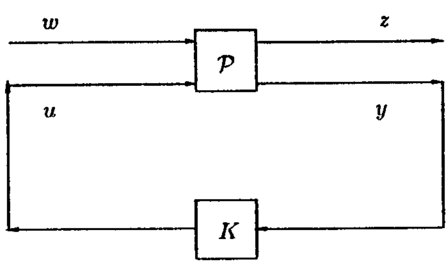

Figure 4

fractional maps $\mathscr{F}_{\mathscr{P}}$ in order of increasing level of generality:

1-block case: both $\mathscr{P}_{12}$ and $\mathscr{P}_{21}$ are square and invertible.
2-block case: $\mathscr{P}_{12}$ is injective and $\mathscr{P}_{21}$ is square and invertible.
4-block case: the general regular case ($\mathscr{P}_{12}$ is injective and $\mathscr{P}_{21}$ is surjective).

The goal of this section is to rearrange the form of the linear fractional map $K \to \mathscr{F}_{\mathscr{P}}[K]$ to an alternative form $K \to \mathscr{G}_{\tilde{\mathscr{P}}}[K]$ having more convenient mathematical properties; in the language of classical circuit theory, the transformation which we are about to define amounts to the conversion from the scattering formalism to the chain formalism.[17] The 1-block case corresponds to the classical case, the 2-block case is a relatively straightforward generalization while the 4-block case is more complicated and involves the introduction of a fictitious channel. We discuss each in turn.

1-block case. We convert to the chain formalism as follows. Rearrange the system of equations

$$\mathscr{P}\begin{bmatrix} w \\ u \end{bmatrix} = \begin{bmatrix} z \\ y \end{bmatrix}$$

to be in the form

$$\tilde{\mathscr{P}}\begin{bmatrix} u \\ y \end{bmatrix} = \begin{bmatrix} z \\ w \end{bmatrix}$$

In the terminology of Reference 18 this amounts to 'partial inversion'; in the context of Figure 4, this amounts to 'reversing the arrows' for the w and y signals to arrive at a system as in Figure 5. The result is

$$\tilde{\mathscr{P}} = \begin{bmatrix} \tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{12} \\ \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} \end{bmatrix} = \begin{bmatrix} \mathscr{P}_{12} - \mathscr{P}_{11}\mathscr{P}_{21}^{-1}\mathscr{P}_{22} & \mathscr{P}_{11}\mathscr{P}_{21}^{-1} \\ -\mathscr{P}_{21}^{-1}\mathscr{P}_{22} & \mathscr{P}_{21}^{-1} \end{bmatrix} \tag{15}$$

When we make the identification $u = Ky$ and solve for z in terms of w, the result is

$$z = (\tilde{\mathscr{P}}_{11}K + \tilde{\mathscr{P}}_{12})(\tilde{\mathscr{P}}_{21}K + \tilde{\mathscr{P}}_{22})^{-1}w$$

If we define a map $\mathscr{G}_{\tilde{\mathscr{P}}}$ by

$$\mathscr{G}_{\tilde{\mathscr{P}}}[K] = (\tilde{\mathscr{P}}_{11}K + \tilde{\mathscr{P}}_{12})(\tilde{\mathscr{P}}_{21}K + \tilde{\mathscr{P}}_{22})^{-1}$$

the above analysis shows that

$$\mathscr{G}_{\tilde{\mathscr{P}}}[K] = \mathscr{F}_{\mathscr{P}}[K]$$

for a generic set of $m_2 \times n_2$ matrices K. From (15) it is clear that $\mathscr{P}$ being in the 1-block case

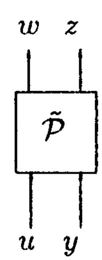

Figure 5

(i.e., $\mathscr{P}_{12}$ and $\mathscr{P}_{21}$ being invertible) corresponds to

$$\tilde{\mathscr{P}}_{22} \quad \text{and} \quad \tilde{\mathscr{P}}_{11} - \tilde{\mathscr{P}}_{12}\tilde{\mathscr{P}}_{22}^{-1}\tilde{\mathscr{P}}_{21} \text{ are square and invertible} \tag{16}$$

Note that invertibility of $\tilde{\mathscr{P}}_{22}$ is equivalent to $K = 0$ being in the domain of definition of $\mathscr{G}_{\tilde{\mathscr{P}}}$ and then, by a Schur complement argument (see, for example, Reference 19), invertibility of $\tilde{\mathscr{P}}_{11} - \tilde{\mathscr{P}}_{12}\tilde{\mathscr{P}}_{22}^{-1}\tilde{\mathscr{P}}_{21}$ amounts to the invertibility of $\tilde{\mathscr{P}}$ itself. A useful property of the chain formalism is the group property

$$\mathscr{G}_{\tilde{\mathscr{P}}^{(1)}} \circ \mathscr{G}_{\tilde{\mathscr{P}}^{(2)}} = \mathscr{G}_{\tilde{\mathscr{P}}^{(1)}\tilde{\mathscr{P}}^{(2)}} \tag{17}$$

where the product on the right is ordinary matrix multiplication.

Conversely, suppose that the block 2×2 matrix

$$\tilde{\mathscr{P}} = \begin{bmatrix} \tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{12} \\ \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} \end{bmatrix}$$

satisfies (16). We then may reverse the process done above, and rearrange the system of equations

$$\tilde{\mathscr{P}}\begin{bmatrix} u \\ y \end{bmatrix} = \begin{bmatrix} z \\ w \end{bmatrix}$$

to arrive at a system of the form

$$\mathscr{P}\begin{bmatrix} w \\ u \end{bmatrix} = \begin{bmatrix} z \\ y \end{bmatrix}$$

The result is

$$\mathscr{P} = \begin{bmatrix} \mathscr{P}_{11} & \mathscr{P}_{12} \\ \mathscr{P}_{21} & \mathscr{P}_{22} \end{bmatrix} = \begin{bmatrix} \tilde{\mathscr{P}}_{12}\tilde{\mathscr{P}}_{22}^{-1} & \tilde{\mathscr{P}}_{11} \tilde{\mathscr{P}}_{12}\tilde{\mathscr{P}}_{22}^{-1}\tilde{\mathscr{P}}_{21} \\ \tilde{\mathscr{P}}_{22}^{-1} & -\tilde{\mathscr{P}}_{22}^{-1}\tilde{\mathscr{P}}_{21} \end{bmatrix} \tag{18}$$

Then assumption (16) is exactly equivalent to $\mathscr{P}_{12}$ and $\mathscr{P}_{21}$ being square and invertible, i.e. to $\mathscr{P}$ being in the 1-block case. Moreover it is easily checked that

$$\mathscr{G}_{\tilde{\mathscr{P}}}[K] = \mathscr{F}_{\mathscr{P}}[K]$$

for all K in the common domain of definition of $\mathscr{G}_{\tilde{\mathscr{P}}}$ and $\mathscr{F}_{\mathscr{P}}$, and that the two transformations $\mathscr{P} \to \tilde{\mathscr{P}}$ and $\tilde{\mathscr{P}} \to \mathscr{P}$ defined above are inverses of each other. We conclude that linear fractional maps of the form $\mathscr{G}_{\tilde{\mathscr{P}}}$ with $\tilde{\mathscr{P}}$ satisfying (16) constitute the chain formalism version of the 1-block case.

2-block case. Now suppose that we are given a block 2×2 matrix

$$\mathscr{P} = \begin{bmatrix} \mathscr{P}_{11} & \mathscr{P}_{12} \\ \mathscr{P}_{21} & \mathscr{P}_{22} \end{bmatrix}$$

with $\mathscr{P}_{21}$ square and invertible but $\mathscr{P}_{12}$ known only to be injective. The derivation of (15) above required only the invertibility of $\mathscr{P}_{21}$; thus we may still define $\tilde{\mathscr{P}}$ by

$$\tilde{\mathscr{P}} = \begin{bmatrix} \tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{12} \\ \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} \end{bmatrix} = \begin{bmatrix} \mathscr{P}_{12} - \mathscr{P}_{1}\mathscr{P}_{21}^{-1}\mathscr{P}_{22} & \mathscr{P}_{11}\mathscr{P}_{21}^{-1} \\ -\mathscr{P}_{21}^{-1}\mathscr{P}_{22} & \mathscr{P}_{21}^{-1} \end{bmatrix}$$

as before. The condition (16) now assumes the weaker form

$$\tilde{\mathscr{P}}_{22} \text{ is square and invertible and } \tilde{\mathscr{P}}_{11} - \tilde{\mathscr{P}}_{12}\tilde{\mathscr{P}}_{22}^{-1}\tilde{\mathscr{P}}_{21} \text{ is injective} \tag{19}$$

We again define the chain formalism version of the linear fractional map

$$K \to \mathcal{G}_{\tilde{\mathcal{P}}}[K] = (\tilde{\mathcal{P}}_{11}K + \tilde{\mathcal{P}}_{12})(\tilde{\mathcal{P}}_{21}K + \tilde{\mathcal{P}}_{22})^{-1}$$

and note that

$$\mathcal{F}_{\mathcal{P}}[K] = \mathcal{G}_{\tilde{\mathcal{P}}}[K]$$

on the common domain of definition. The condition that $\tilde{\mathcal{P}}_{22}$ is invertible guarantees that $K = 0$ is in the domain of definition of $\mathcal{G}_{\tilde{\mathcal{P}}}$ and then by a Schur complement argument the additional condition that $\tilde{\mathcal{P}}_{11} - \tilde{\mathcal{P}}_{12}\tilde{\mathcal{P}}_{22}^{-1}\tilde{\mathcal{P}}_{21}$ is injective amounts to the condition that the matrix $\tilde{\mathcal{P}}$ itself be injective. The group property (17) makes sense only if $\tilde{\mathcal{P}}^{(1)}$ and $\tilde{\mathcal{P}}^{(2)}$ are allowed to have appropriate (different) sizes.

Conversely, if

$$\tilde{\mathcal{P}} = \begin{bmatrix} \tilde{\mathcal{P}}_{11} & \tilde{\mathcal{P}}_{12} \\ \tilde{\mathcal{P}}_{21} & \tilde{\mathcal{P}}_{22} \end{bmatrix}$$

is a block 2×2 matrix satisfying (19), we may again define $\mathcal{P}$ by the inverse transform (18)

$$\mathcal{P} = \begin{bmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{21} & \mathcal{P}_{22} \end{bmatrix} = \begin{bmatrix} \tilde{\mathcal{P}}_{12}\tilde{\mathcal{P}}_{22}^{-1} & \tilde{\mathcal{P}}_{11} - \tilde{\mathcal{P}}_{12}\tilde{\mathcal{P}}_{22}^{-1}\tilde{\mathcal{P}}_{21} \\ \tilde{\mathcal{P}}_{22}^{-1} & -\tilde{\mathcal{P}}_{22}^{-1}\tilde{\mathcal{P}}_{21} \end{bmatrix}$$

Then hypothesis (19) on $\tilde{\mathcal{P}}$ translates exactly to $\mathcal{P}_{21}$ is square and invertible together with $\mathcal{P}_{12}$ is injective, i.e., to $\mathcal{P}$ being in the 2-block case. We conclude that linear fractional maps of the form $\mathcal{G}_{\tilde{\mathcal{P}}}$ with $\tilde{\mathcal{P}}$ satisfying (19) constitute the chain formalism version of the 2-block case.

4-block case. We are now ready to consider the general regular case. Suppose that we are given a block 2×2 matrix

$$\mathcal{P} = \begin{bmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{21} & \mathcal{P}_{22} \end{bmatrix}$$

such that $\mathcal{P}_{12}$ is injective and $\mathcal{P}_{21}$ is surjective. Transformation to the chain formalism appears to require the invertibility of $\mathcal{P}_{21}$. However, since $\mathcal{P}_{21}$ is surjective it is possible to append a second block row $\mathcal{P}_{21}^0$ so that

$$\begin{bmatrix} \mathcal{P}_{21} \\ \mathcal{P}_{21}^0 \end{bmatrix}$$

is square and invertible. Define $\mathcal{P}_{22}^0$ of a compatible size arbitrarily. Then

$$\boldsymbol{\mathcal{P}} = \begin{bmatrix} \mathcal{P}_{11} & \vdots & \mathcal{P}_{12} \\ \cdots & & \cdots \\ \mathcal{P}_{21} & \vdots & \mathcal{P}_{22} \\ \mathcal{P}_{21}^0 & \vdots & \mathcal{P}_{22}^0 \end{bmatrix} : \begin{bmatrix} w \\ u \end{bmatrix} \to \begin{bmatrix} z \\ y \\ y^0 \end{bmatrix}$$

is as in the 2-block case. The associated linear fractional map $\mathcal{F}_{\boldsymbol{\mathcal{P}}}$ acts on compensators of block row type $[K\,K_0]$ and has the form

$$\mathcal{F}_{\boldsymbol{\mathcal{P}}}[\,[K, K_0]\,] = \mathcal{P}_{11} + \mathcal{P}_{12}[K, K_0]\left\{\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} - \begin{bmatrix} \mathcal{P}_{22} \\ \mathcal{P}_{22}^0 \end{bmatrix}[K, K_0]\right\}^{-1}\begin{bmatrix} \mathcal{P}_{21} \\ \mathcal{P}_{21}^0 \end{bmatrix}$$

Note that if we restrict K_0 to $K_0 = 0$ we get

$$\mathcal{F}_{\boldsymbol{\mathcal{P}}}[\,[K, 0]\,] = \mathcal{P}_{11} + \mathcal{P}_{12}[K, 0]\begin{bmatrix} I - \mathcal{P}_{22}K & 0 \\ -\mathcal{P}_{22}^0 K & I \end{bmatrix}^{-1}\begin{bmatrix} \mathcal{P}_{21} \\ \mathcal{P}_{21}^0 \end{bmatrix} = \mathcal{F}_{\mathcal{P}}[K]$$

In words, a general 4-block linear fractional map is equivalent to a 2-block linear fractional map of a larger size restricted to compensators which annihilate a certain subset of the channels feeding into the compensator (see Figure 6). We use formula (15) to convert $\mathcal{P}$ to the chain formalism; the result is

$$\tilde{\mathcal{P}} = \begin{bmatrix} \tilde{\mathcal{P}}_{11} & \vdots & \tilde{\mathcal{P}}_{12} & \tilde{\mathcal{P}}_{13} \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{\mathcal{P}}_{21} & \vdots & \tilde{\mathcal{P}}_{22} & \tilde{\mathcal{P}}_{23} \end{bmatrix} = \begin{bmatrix} \mathcal{P}_{12} - \mathcal{P}_{11}\begin{bmatrix} \mathcal{P}_{21} \\ \mathcal{P}_{21}^0 \end{bmatrix}^{-1}\begin{bmatrix} \mathcal{P}_{22} \\ \mathcal{P}_{22}^0 \end{bmatrix} & \mathcal{P}_{11}\begin{bmatrix} \mathcal{P}_{21} \\ \mathcal{P}_{21}^0 \end{bmatrix}^{-1} \\ -\begin{bmatrix} \mathcal{P}_{21} \\ \mathcal{P}_{21}^0 \end{bmatrix}^{-1}\begin{bmatrix} \mathcal{P}_{22} \\ \mathcal{P}_{22}^0 \end{bmatrix} & \begin{bmatrix} \mathcal{P}_{21} \\ \mathcal{P}_{21}^0 \end{bmatrix}^{-1} \end{bmatrix} \tag{20}$$

where the associated system of equations is

$$\tilde{\mathcal{P}}\begin{bmatrix} u \\ y \\ y^0 \end{bmatrix} = \begin{bmatrix} z \\ w \end{bmatrix}$$

From results derived above for the 2-block case, we know that

$$\mathcal{F}_{\mathcal{P}}[K] = \mathcal{F}_{\boldsymbol{\mathcal{P}}}[\,[K,0]\,] = \mathcal{G}_{\tilde{\boldsymbol{\mathcal{P}}}}[\,[K,0]\,] = (\tilde{\mathcal{P}}_{11}[K,0] + [\tilde{\mathcal{P}}_{12}, \tilde{\mathcal{P}}_{13}]\,)(\,[\tilde{\mathcal{P}}_{21}[K,0] + [\tilde{\mathcal{P}}_{22}, \tilde{\mathcal{P}}_{23}]\,)^{-1}$$

This leads us to define a block 2×3 matrix

$$\tilde{\mathcal{P}} = \begin{bmatrix} \tilde{\mathcal{P}}_{11} & \tilde{\mathcal{P}}_{12} & \tilde{\mathcal{P}}_{13} \\ \tilde{\mathcal{P}}_{21} & \tilde{\mathcal{P}}_{22} & \tilde{\mathcal{P}}_{23} \end{bmatrix}$$

(the same matrix as $\tilde{\boldsymbol{\mathcal{P}}}$ but considered as having a block 2×3 structure rather than block 2×2 structure) and define $\mathcal{G}_{\tilde{\mathcal{P}}}[K]$ by

$$\mathcal{G}_{\tilde{\mathcal{P}}}[K] = \mathcal{G}_{\tilde{\boldsymbol{\mathcal{P}}}}[\,[K,0]\,] = [\tilde{\mathcal{P}}_{11}K + \tilde{\mathcal{P}}_{12}, \tilde{\mathcal{P}}_{13}]\,[\tilde{\mathcal{P}}_{21}K + \tilde{\mathcal{P}}_{22}, \tilde{\mathcal{P}}_{23}]^{-1} \tag{21}$$

In this way we arrive at the chain formalism version $\mathcal{G}_{\tilde{\mathcal{P}}}$ of the linear fractional map $\mathcal{F}_{\mathcal{P}}$ for the general 4-block case. Properties inherited by $\tilde{\mathcal{P}}$ from $\mathcal{P}$ being regular which are read off from (20) are

$$[\tilde{\mathcal{P}}_{22} \;\; \tilde{\mathcal{P}}_{23}] \text{ is invertible and } \tilde{\mathcal{P}}_{11} - [\tilde{\mathcal{P}}_{12} \;\; \tilde{\mathcal{P}}_{13}]\,[\tilde{\mathcal{P}}_{22} \;\; \tilde{\mathcal{P}}_{23}]^{-1}\tilde{\mathcal{P}}_{21} \text{ is injective} \tag{22}$$

Note that invertibility of $[\tilde{\mathcal{P}}_{22} \; \tilde{\mathcal{P}}_{23}]$ guarantees that $K = 0$ is in the domain of definition of $\mathcal{G}_{\mathcal{P}}$ and then by a Schur complement argument the injectivity of $\tilde{\mathcal{P}}_{11} - [\tilde{\mathcal{P}}_{12}\tilde{\mathcal{P}}_{13}]\,[\tilde{\mathcal{P}}_{22}\tilde{\mathcal{P}}_{23}]^{-1}\tilde{\mathcal{P}}_{21}$ is equivalent to the injectivity of the big matrix $\tilde{\mathcal{P}}$.

Conversely, suppose that we are given a block 2×3 matrix

$$\tilde{\mathcal{P}} = \begin{bmatrix} \tilde{\mathcal{P}}_{11} & \tilde{\mathcal{P}}_{12} & \tilde{\mathcal{P}}_{13} \\ \tilde{\mathcal{P}}_{21} & \tilde{\mathcal{P}}_{22} & \tilde{\mathcal{P}}_{23} \end{bmatrix}$$

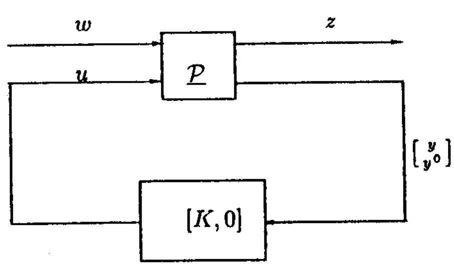

Figure 6

which satisfies condition (22). We merge the last 2-block columns to a 1-block column to generate a block 2×2 matrix

$$\tilde{\mathscr{P}}=\begin{bmatrix}\tilde{\mathscr{P}}_{11} & \vdots & \tilde{\mathscr{P}}_{12} & \tilde{\mathscr{P}}_{13}\\ \tilde{\mathscr{P}}_{21} & \vdots & \tilde{\mathscr{P}}_{22} & \tilde{\mathscr{P}}_{23}\end{bmatrix}$$

which satisfies condition (19) and which has the property

$$\mathscr{G}_{\tilde{\mathscr{P}}}[K]=\mathscr{G}_{\tilde{\mathscr{P}}}[[K,0]]$$

We next use the general formula (18) to convert back to the scattering formalism; the result is

$$\mathscr{P}=\begin{bmatrix}[\tilde{\mathscr{P}}_{12},\tilde{\mathscr{P}}_{13}][\tilde{\mathscr{P}}_{22},\tilde{\mathscr{P}}_{23}]^{-1} & \tilde{\mathscr{P}}_{11}-[\tilde{\mathscr{P}}_{12},\tilde{\mathscr{P}}_{13}][\tilde{\mathscr{P}}_{22},\tilde{\mathscr{P}}_{23}]^{-1}\tilde{\mathscr{P}}_{21}\\ [\tilde{\mathscr{P}}_{22},\tilde{\mathscr{P}}_{23}]^{-1} & -[\tilde{\mathscr{P}}_{22},\tilde{\mathscr{P}}_{23}]^{-1}\tilde{\mathscr{P}}_{21}\end{bmatrix} \quad (23)$$

We then define $\mathscr{P}$ so that

$$\mathscr{F}_{\mathscr{P}}[K]=\mathscr{F}_{\mathscr{P}}[K,0]$$

This amounts to ignoring the last block row of $\mathscr{P}$. The result is

$$\mathscr{P}=\begin{bmatrix}\mathscr{P}_{11} & \mathscr{P}_{12}\\ \mathscr{P}_{21} & \mathscr{P}_{22}\end{bmatrix}$$

where

$$\begin{aligned}\mathscr{P}_{11}&=[\tilde{\mathscr{P}}_{12},\tilde{\mathscr{P}}_{13}][\tilde{\mathscr{P}}_{22},\tilde{\mathscr{P}}_{23}]^{-1}\\ \mathscr{P}_{12}&=\tilde{\mathscr{P}}_{11}-[\tilde{\mathscr{P}}_{22},\tilde{\mathscr{P}}_{13}][\tilde{\mathscr{P}}_{22},\tilde{\mathscr{P}}_{23}]^{-1}\tilde{\mathscr{P}}_{21}\\ \mathscr{P}_{21}&=[I\ \ 0][\tilde{\mathscr{P}}_{22},\tilde{\mathscr{P}}_{23}]^{-1}\\ \mathscr{P}_{22}&=-[I\ \ 0][\tilde{\mathscr{P}}_{22}\ \ \tilde{\mathscr{P}}_{23}]^{-1}\tilde{\mathscr{P}}_{21}\end{aligned} \quad (24)$$

Note that assumption (21) on $\tilde{\mathscr{P}}$ is equivalent to $\mathscr{P}_{12}$ being injective and $\mathscr{P}_{21}$ being surjective, i.e. $\mathscr{P}$ is in the regular case. We conclude that maps of the form (21) with $\tilde{\mathscr{P}}$ a block 2×3 matrix satisfying (22) gives the chain formalism form of linear fractional maps for the general regular case.

We close this section with a discussion of to what extent the map $K\to\mathscr{G}_{\tilde{\mathscr{P}}}[K]$ determines the matrix $\tilde{\mathscr{P}}$. For regular maps in the conventions of the scattering formalism $K\to\mathscr{F}_{\mathscr{P}}[K]$, this is settled by Lemma 1.1.1 given above. For the chain formalism we have the following result.

Lemma 1.1.2

Suppose

$$\tilde{\mathscr{P}}^{k}=\begin{bmatrix}\tilde{\mathscr{P}}^{k}_{11} & \tilde{\mathscr{P}}^{k}_{12} & \tilde{\mathscr{P}}^{k}_{13}\\ \tilde{\mathscr{P}}^{k}_{21} & \tilde{\mathscr{P}}^{k}_{22} & \tilde{\mathscr{P}}^{k}_{23}\end{bmatrix}$$

is a block 2×3 matrix satisfying condition (22) with $\tilde{\mathscr{P}}^{k}_{ij}$ of block size $\tilde{m}_i\times\tilde{n}_j$ for $k=1,2$. In order that

$$\mathscr{G}_{\tilde{\mathscr{P}}^1}[K]=\mathscr{G}_{\tilde{\mathscr{P}}^2}[K]$$

for a generic set of $\tilde{n}_1\times\tilde{n}_2$ matrices K it is necessary and sufficient that there exists a non-zero scalar α and $\tilde{n}_3 \times \tilde{n}_j$ matrices γ_j for $j = 1, 2, 3$ with γ_3 invertible such that

$$\begin{bmatrix} \tilde{\mathscr{P}}^2_{11} & \tilde{\mathscr{P}}^2_{12} & \tilde{\mathscr{P}}^2_{13} \\ \tilde{\mathscr{P}}^2_{21} & \tilde{\mathscr{P}}^2_{22} & \tilde{\mathscr{P}}^2_{23} \end{bmatrix} = \begin{bmatrix} \tilde{\mathscr{P}}^1_{11} & \tilde{\mathscr{P}}^1_{12} & \tilde{\mathscr{P}}^1_{13} \\ \tilde{\mathscr{P}}^1_{21} & \tilde{\mathscr{P}}^1_{22} & \tilde{\mathscr{P}}^1_{23} \end{bmatrix} \begin{bmatrix} \alpha I & 0 & 0 \\ 0 & \alpha I & 0 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix} \tag{25}$$

In particular, if the third column of $\tilde{\mathscr{P}}$ is vacuous (i.e., $\tilde{n}_3 = 0$ and $\tilde{\mathscr{P}}$ is in the 2-block case (19)), then equality of the two maps $\mathscr{G}_{\tilde{\mathscr{P}}^1}$ and $\mathscr{G}_{\tilde{\mathscr{P}}^2}$ is equivalent to

$$\tilde{\mathscr{P}}^2 = \alpha \tilde{\mathscr{P}}^1$$

for a non-zero scalar α.

Proof. This result can be established by tracing through the amount of freedom involved in the transformation from $\mathscr{P}$ to $\boldsymbol{\mathscr{P}}$ to $\tilde{\mathscr{P}}$ and using the characterization of non-uniqueness in the correspondence $\mathscr{P} \to \mathscr{F}_{\mathscr{P}}$ given by Lemma 1.1.1.

1.2. Linear fractional maps of contractions

In this subsection we establish some general facts concerning linear fractional maps which map contraction matrices into contraction matrices. For convenience, with the exception of Proposition 1.2.3, we will state the results only for linear fractional maps presented in the chain formalism.

Our main interest here is to define various classes of block 2×3 matrices

$$U = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \end{bmatrix}$$

(with block sizes $m_i \times n_j$, say, for $i = 1, 2$ and $j = 1, 2, 3$) satisfying the regularity assumption (22) with the property that $\mathscr{G}_U[K]$ is a contraction if and only if K is a contraction. We first consider the 2-block case where $n_3 = 0$ and U_{22} is invertible (so $m_2 = n_2$). We work with the linearization of the equation

$$\mathscr{G}_U[K]\, w = z$$

given by the system of equations

$$\begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} = \begin{bmatrix} z \\ w \end{bmatrix} \tag{26a}$$

$$u = Ky \tag{26b}$$

Let us suppose that U is (j, J)-isometric, i.e., that $U^* J U = j$, where $J = I_{m_1} \oplus -I_{m_2}$ and $j = I_{n_1} \oplus -I_{n_2}$. Then we have explicitly

$$\begin{bmatrix} U_{11}^* U_{11} - U_{21}^* U_{21} & U_{11}^* U_{12} - U_{21}^* U_{22} \\ U_{12}^* U_{11} - U_{22}^* U_{21} & U_{12}^* U_{12} - U_{22}^* U_{22} \end{bmatrix} = \begin{bmatrix} I_{n_1} & 0 \\ 0 & -I_{n_2} \end{bmatrix}$$

In particular $U_{22}^* U_{22} = I_{n_2} + U_{12}^* U_{12}$. Thus U_{22} (which by assumption is square) is invertible and, from

$$(U_{12} U_{22}^{-1})^* U_{12} U_{22}^{-1} = I - (U_{22}^{-1})^* U_{22}^{-1}$$

we conclude that $U_{12} U_{22}^{-1}$ is a strict contraction, that is

$$\| U_{22}^{-1} U_{21} w \| < \| w \| \qquad \text{for all } w \neq 0$$

From this it follows that for any $n_1 \times n_2$ matrix K with $\|K\| \leqslant 1$,

$$U_{21}K + U_{22} = U_{22}(U_{22}^{-1}U_{21}K + I)$$

is invertible, i.e., any $n_1 \times n_2$ K with $\|K\| \leqslant 1$ is in the domain of definition of $\mathscr{G}_K$. Moreover, the (j, J)-isometric property of U implies that

$$\|z\|^2 - \|w\|^2 = \|u\|^2 - \|y\|^2$$

whenever u, y, z, w satisfy (26a). If we now plug in $u = Ky$ from (26b) we get

$$\|\mathscr{G}_U[K]w\|^2 - \|w\|^2 = \|Ky\|^2 - \|y\|^2$$

for all w, where $y = (U_{21}K + U_{22})^{-1}w$. Since for any fixed K in the domain of definition of $\mathscr{G}_U$ the matrix $U_{21}K + U_{22}$ is invertible, we see that y sweeps through all of $\mathbb{C}^{n_2}$ as w sweeps through $\mathbb{C}^{m_2} = \mathbb{C}^{n_2}$. Hence K is contractive, isometric or expansive (i.e. $\|Ky\|^2 \geqslant \|y\|^2$ for all y) if and only if $\mathscr{G}_U[K]$ has the same property. This gives us the following result.

Proposition 1.2.1

Suppose

$$U = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}$$

is (j, J)-isometric, where $j = I_{n_1} \oplus -I_{n_2}$, $J = I_{m_1} \oplus -I_{m_2}$ and $n_2 = m_2$. Then U satisfies (19), and any contractive $n_1 \times n_2$ matrix K is in the domain of definition of $\mathscr{G}_U$. Moreover, for a given K in the domain of definition of $\mathscr{G}_U$, $\mathscr{G}_U[K]$ is contractive (respectively, isometric) (respectively, expansive) if and only if K is contractive (respectively, isometric) (respectively, expansive).

Our next goal is to extend Proposition 1.2.1 to a more general class of block 2×3 matrices U. Therefore consider a block 2×3 matrix

$$U = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \end{bmatrix}$$

where U_{ij} has size $m_i \times n_j$ and U satisfies the regularity assumption (22). We now suppose that U is $(j \oplus -\Psi, J)$-isometric where, as before, $j = I_{n_1} \oplus -I_{n_2}$, $J = I_{m_1} \oplus -I_{n_2}$ and $\Psi = \psi^*\psi$ is a general positive semidefinite $n_3 \times n_3$ matrix. As for the case $n_3 = 0$ done above, one can see that the domain of definition of $\mathscr{G}_U$ includes all $n_1 \times n_2$ contraction matrices K. The linearization of the linear fractional equation

$$\mathscr{G}_U[K]w = z$$

leads us to the system of equations

$$\begin{bmatrix} U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \end{bmatrix} \begin{bmatrix} u \\ y \\ y^0 \end{bmatrix} = \begin{bmatrix} z \\ w \end{bmatrix} \tag{27a}$$

$$u = Ky \tag{27b}$$

The $(j \oplus -\Psi, J)$-isometric property of U implies that

$$\|z\|^2 - \|w\|^2 = \|u\|^2 - \|y\|^2 - \|\psi y^0\|^2 \tag{28}$$

whenever u, y, y^0, z, w satisfy (27a). If we now plug in $u = Ky$ from (27b) we get

$$\| \mathscr{G}_U[K]\, w \|^2 - \| w \|^2 = \| Ky \|^2 - \| y \|^2 - \| \psi y^0 \|^2 \tag{29}$$

for all w, where

$$\begin{bmatrix} y \\ y^0 \end{bmatrix} = [U_{21}K + U_{22}, U_{23}]^{-1} w.$$

In particular, if K is contractive we see from (29) that $\mathscr{G}_U[K]$ is contractive. Conversely, if $\mathscr{G}_U[K]$ is contractive, we have

$$\| \mathscr{G}_U[K]\, w \|^2 - \| w \|^2 \leqslant 0$$

for all w. If we specialize w to be of the form

$$[U_{21}K + U_{22}, U_{23}] \begin{bmatrix} y \\ 0 \end{bmatrix}$$

and use (28), we get

$$\| Ky \|^2 - \| y \|^2 \leqslant 0$$

for all $y \in \mathbb{C}^{n_2}$, i.e., necessarily K is contractive. Also, if $y^0 \in \operatorname{Ker} \Psi$ and we choose

$$w = [U_{21}K + U_{22}, U_{23}] \begin{bmatrix} 0 \\ y^0 \end{bmatrix}$$

from (28) we read off that $\| \mathscr{G}_U[K]\, w \|^2 = \| w \|^2$; hence when Ψ has a non-zero kernel there is no choice of K for which $\| \mathscr{G}_U[K] \| < 1$. On the other hand, if $\Psi > 0$ we read off from (28) that $\| \mathscr{G}_U[K]\, w \| < 1$ for all choices of w if and only if $\| K \| < 1$. We have arrived at the following result.

Proposition 1.2.2

Suppose

$$U = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \end{bmatrix}$$

is $(j \oplus -\Psi, J)$-isometric, where $j = I_{n_1} \oplus -I_{n_2}$, Ψ is an $n_3 \times n_3$ positive semidefinite matrix, $J = I_{m_1} \oplus -I_{m_2}$ and where U satisfies the regularity condition (22) (with U in place of $\tilde{\mathscr{P}}$). Then any contractive $n_1 \times n_2$ matrix K is in the domain of definition of $\mathscr{G}_U$. Moreover, (i) $\mathscr{G}_U[K]$ is contractive if and only if K is contractive. (ii) There exist contractive K for which $\mathscr{G}_U[K]$ is strictly contractive if and only if Ψ is positive definite, and in this case, $\mathscr{G}_U[K]$ is strictly contractive if and only if K is strictly contractive.

In this paper we will be mainly interested in linear fractional transformations having the property (ii) described in Proposition 1.2.2. The following result characterizes the class of such maps in the scattering formalism.

Proposition 1.2.3

Suppose

$$\mathscr{P} = \begin{bmatrix} \mathscr{P}_{11} & \mathscr{P}_{12} \\ \mathscr{P}_{21} & \mathscr{P}_{22} \end{bmatrix}$$

is a block matrix with $\mathscr{P}_{12}$ injective and $\mathscr{P}_{21}$ surjective with $\mathscr{P}_{ij}$ of size $m_i \times n_j$ for $i, j = 1, 2$ such that

(i) $\|\mathscr{P}\| \leqslant 1$,
(ii) $\|\mathscr{P}_{11} \mid \text{Ker } \mathscr{P}_{21}\| < 1$, and
(iii) $\text{rank}(I - \mathscr{P}^*\mathscr{P}) = \dim \text{Ker } \mathscr{P}_{21}$.

Then

(a) $\|\mathscr{F}_{\mathscr{P}}[K]\| \leqslant 1 \Leftrightarrow \|K\| \leqslant 1$, and
(b) $\|\mathscr{F}_{\mathscr{P}}[K]\| < 1 \Leftrightarrow \|K\| < 1$.

Proof. We show that if $\mathscr{P}$ satisfies conditions (i) and (ii) in the statement of Proposition 1.2.3 then there is a choice of augmentation $\mathscr{P}^0 = [\mathscr{P}_{21}^0 \ \ \mathscr{P}_{22}^0]$ for $\mathscr{P}$ so that the resulting transform $U = \tilde{\mathscr{P}}$ defined by (20) is as in part (ii) of Proposition 1.2.2 with $\psi = I$. From (28) and the connection (20) between $\mathscr{P}$ and $\tilde{\mathscr{P}}$, the issue is the construction of $\mathscr{P}^0 = [\mathscr{P}_{21}^0 \ \ \mathscr{P}_{22}^0]$ such that

$$\begin{bmatrix} \mathscr{P} \\ \mathscr{P}^0 \end{bmatrix} \quad \text{is isometric} \tag{30}$$

and

$$\begin{bmatrix} \mathscr{P}_{21} \\ \mathscr{P}_{21}^0 \end{bmatrix} \quad \text{is invertible} \tag{31}$$

By condition (iii) there is a surjective linear transformation $\mathscr{P}^0 = [\mathscr{P}_{21}^0 \ \ \mathscr{P}_{22}^0]$ from $\mathbb{C}^{n_1} \oplus \mathbb{C}^{n_2}$ to $\mathbb{C}^{n_3}$ (where $n_3 = \dim \text{Ker } \mathscr{P}_{21}$) so that

$$I - \mathscr{P}^*\mathscr{P} = (\mathscr{P}^0)^*\mathscr{P}^0 \tag{32}$$

Thus (30) is satisfied with this choice of $\mathscr{P}^0$. To check (31), write $\mathbb{C}^{n_1}$ as a direct sum

$$\mathbb{C}^{n_1} = (\text{Ker } \mathscr{P}_{21})^{\perp} \oplus \text{Ker } \mathscr{P}_{21}$$

and partition $\mathscr{P}_{11}$, $\mathscr{P}_{21}$ and $\mathscr{P}_{21}^0$ conformably:

$$\begin{aligned} \mathscr{P}_{11} &= [\mathscr{P}_{111} \quad \mathscr{P}_{112}] \\ \mathscr{P}_{21} &= [\mathscr{P}_{211} \ \ 0] \\ \mathscr{P}_{21} &= [\mathscr{P}_{211}^0 \ \ \mathscr{P}_{212}^0] \end{aligned}$$

and note that $\mathscr{P}_{211}$ is invertible since $\mathscr{P}_{21}$ is surjective by assumption. From

$$\begin{bmatrix} \mathscr{P}_{21} \\ \mathscr{P}_{21}^0 \end{bmatrix} = \begin{bmatrix} \mathscr{P}_{211} & 0 \\ \mathscr{P}_{211}^0 & \mathscr{P}_{212}^0 \end{bmatrix}$$

we note that (31) is equivalent to

$$\mathscr{P}_{212}^0 \qquad \text{is square and invertible}$$

On the other hand (32) in particular gives

$$I - \mathscr{P}_{112}^*\mathscr{P}_{112} = (\mathscr{P}_{212}^0)^*\mathscr{P}_{212}^0 \tag{33}$$

Assumption (ii) guarantees that

$$\text{rank}(I - \mathscr{P}_{112}^*\mathscr{P}_{112}) = m_3$$

Hence (33) implies that $\mathscr{P}_{212}^0$ is invertible as required. □

Remark 1. One can also prove a converse to the general result proved in the proof of Proposition 1.2.3 is the following matrix extension result. Given a block matrix

$$\mathscr{P}=\begin{bmatrix}\mathscr{P}_{11} & \mathscr{P}_{12}\\ \mathscr{P}_{21} & \mathscr{P}_{22}\end{bmatrix}$$

where $\mathscr{P}_{ij}$ has size $m_i \times n_j$, then there is a choice of $m_3 \times (n_1+n_2)$ matrix $\mathscr{P}^0=[\mathscr{P}^0_{21}\ \ \mathscr{P}^0_{22}]$ such that

(a) $$\begin{bmatrix}\mathscr{P}\\ \mathscr{P}^0\end{bmatrix} \quad \text{is an isometry}$$

and

(b) $$\begin{bmatrix}\mathscr{P}_{21}\\ \mathscr{P}^0_{211}\end{bmatrix} \quad \text{is square and invertible}$$

if and only if

(i) $\|\mathscr{P}\| \leqslant 1$,
(ii) $\|\mathscr{P}_{11} \mid \operatorname{Ker} \mathscr{P}_{21}\| < 1$,
(iii) $\operatorname{rank}(I-\mathscr{P}^*\mathscr{P}) = \dim \operatorname{Ker} \mathscr{P}_{21}$,
(iv) $\mathscr{P}_{21}$ is surjective.

Remark 2. The proof of Proposition 1.2.3 showed that if

$$\mathscr{P}=\begin{bmatrix}\mathscr{P}_{11} & \mathscr{P}_{12}\\ \mathscr{P}_{21} & \mathscr{P}_{22}\end{bmatrix}$$

is a block matrix satisfying (i), (ii) and (iii) in Proposition 1.2.3 in addition to (13) and (14), then $\mathscr{P}$ has a chain formalism transform $U=\tilde{\mathscr{P}}$ as in (1.1.8) which is

$$\left(\begin{bmatrix} j & 0\\ 0 & -I\end{bmatrix}, J\right)\text{-isometric}$$

Conversely one can reverse the steps in the argument to show that an

$$\left(\begin{bmatrix} j & 0\\ 0 & -I\end{bmatrix}, J\right)\text{-isometric } U$$

is in turn the chain formalism transform $U=\tilde{\mathscr{P}}$ of a block matrix $\mathscr{P}$ satisfying the conditions of Proposition 1.2.3. We shall use this remark in Section 4.

2. A FACTORIZATION PRINCIPLE FOR STABILITY OF FEEDBACK SYSTEMS

Consider the feedback configuration $\Sigma(\mathscr{P}, K)$ depicted in Figure 7. Here

$$\mathscr{P}\colon \begin{bmatrix} w\\ u\end{bmatrix} \to \begin{bmatrix} z\\ y_1\end{bmatrix}$$

and $K\colon y \to u-v$ represent input–output maps (called the *plant* and the *compensator*) for linear time-invariant finite dimensional systems, w is the disturbance or generalized reference signal, u is the control signal, z is the error signal, y is the measurement signal, v_1 and v_2 are auxiliary disturbance signals introduced to define internal stability in an input–output setting. Figure 7 represents a general paradigm into which a variety of control problems can be cast;

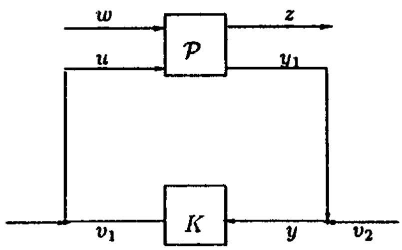

Figure 7

for more details, see References 1 and 2. After applying the Laplace transformation, we may assume that the signals w, u, z, y, are rational vector functions with values in finite dimensional spaces W, U, Z, Y respectively, and that $\mathscr{P}$ and K are given as multiplication by rational matrix functions, also denoted as $\mathscr{P} = \mathscr{P}(z)$ and $K = K(z)$. The configuration in Figure 7 stands for the system of algebraic equations

$$\mathscr{P}\begin{bmatrix} w \\ u \end{bmatrix} = \begin{bmatrix} z \\ y - v_2 \end{bmatrix} \qquad (34)$$
$$u = Ky + v_1$$

The system $\Sigma(\mathscr{P}, K)$ in Figure 7 is said to be *well-posed* if one can solve the system (34) uniquely for z, u, y in terms of w, v_1, v_2 and the resulting map

$$\begin{bmatrix} w \\ v_1 \\ v_2 \end{bmatrix} \to \begin{bmatrix} z \\ u \\ y \end{bmatrix}$$

is given by multiplication by a proper rational matrix function $\mathscr{H}(z) = [\mathscr{H}_{ij}(z)]_{1 \leqslant i, j \leqslant 3}$:

$$\begin{bmatrix} z \\ u \\ y \end{bmatrix} = \mathscr{H}\begin{bmatrix} w \\ v_1 \\ v_2 \end{bmatrix} \qquad (35)$$

If in addition $\mathscr{H}$ is stable (i.e., all poles are in the open left half plane) then the system $\Sigma(\mathscr{P}, K)$ is said to be *internally stable*. Explicitly, in terms of

$$\tilde{\mathscr{P}} = \begin{bmatrix} \mathscr{P}_{11} & \mathscr{P}_{12} \\ \mathscr{P}_{21} & \mathscr{P}_{22} \end{bmatrix} : \begin{bmatrix} w \\ u \end{bmatrix} \to \begin{bmatrix} z \\ y - v_2 \end{bmatrix}$$

and $K: y \to u - v_1$, well-posedness works out to be equivalent to

The rational matrix functions $\Delta = K - \mathscr{P}_{22}K$ is invertible (36)

together with associated rational matrix function $\mathscr{H}$ as in (35) being proper. Explicitly $\mathscr{H}$ is given by

$$\mathscr{H} = \begin{bmatrix} \mathscr{P}_{11} + \mathscr{P}_{12}K\Delta^{-1}\mathscr{P}_{21} & \mathscr{P}_{12} + \mathscr{P}_{12}K\Delta^{-1}\mathscr{P}_{22} & \mathscr{P}_{12}K\Delta^{-1} \\ K\Delta^{-1}\mathscr{P}_{21} & I + K\Delta^{-1}\mathscr{P}_{22} & K\Delta^{-1} \\ \Delta^{-1}\mathscr{P}_{21} & \Delta^{-1}\mathscr{P}_{22} & \Delta^{-1} \end{bmatrix} \qquad (37)$$

In this section we establish a notion of equivalence between two feedback systems $\Sigma(\mathscr{P}, K)$ and $\Sigma(\mathscr{P}', K')$; the useful property of a pair of equivalent systems $\Sigma(\mathscr{P}, K)$ and $\Sigma(\mathscr{P}', K')$ is that one is internally stable if and only if the other is. In this way a complicated system $\Sigma(\mathscr{P}, K)$ can be analysed by exhibiting its equivalence to a simpler system $\Sigma(\mathscr{P}', K')$; illustrations of this general principle will be given in Sections 3 and 4 to follow.

The formulation of this notion of equivalence requires conversion of $\mathscr{P}$ to its representation $\tilde{\mathscr{P}}$ in the chain formalism; this was done in detail in Section 1.1 for constant matrices. We now use the extension of this analysis to matrices of rational functions; equivalently, we may think in terms of the analysis in Section 1.1 being applied pointwise. It is convenient to do the simpler 2-block case separately first.

2.1. *The 2-block case*

We consider first the simpler 2-block case where $\mathscr{P}_{21}$ is invertible (as a rational matrix function). Then, as in Section 1.1, the system of equations

$$\begin{bmatrix} \mathscr{P}_{11} & \mathscr{P}_{12} \\ \mathscr{P}_{21} & \mathscr{P}_{22} \end{bmatrix} \begin{bmatrix} w \\ u \end{bmatrix} = \begin{bmatrix} z \\ y_1 \end{bmatrix}$$

can be rearranged in the form

$$\begin{bmatrix} \tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{12} \\ \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} \end{bmatrix} \begin{bmatrix} u \\ y_1 \end{bmatrix} = \begin{bmatrix} z \\ w \end{bmatrix}$$

where

$$\tilde{\mathscr{P}} = \begin{bmatrix} \tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{12} \\ \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} \end{bmatrix} = \begin{bmatrix} \mathscr{P}_{12} - \mathscr{P}_{11}\mathscr{P}_{21}^{-1}\mathscr{P}_{22} & \mathscr{P}_{11}\mathscr{P}_{21}^{-1} \\ -\mathscr{P}_{21}^{-1}\mathscr{P}_{22} & \mathscr{P}_{21}^{-1} \end{bmatrix} \tag{38}$$

We rewrite Figure 7 in the suggestive form of Figure 8 where it is to be understood that $u_1 = Ky$. Now suppose that $\tilde{\mathscr{P}}$ has a factorization

$$\tilde{\mathscr{P}} = \Theta \circ R$$

where

$$\Theta = \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{bmatrix} : \begin{bmatrix} u' \\ y_1' \end{bmatrix} \to \begin{bmatrix} z \\ w \end{bmatrix}$$

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} : \begin{bmatrix} u \\ y_1 \end{bmatrix} \to \begin{bmatrix} u' \\ y_1' \end{bmatrix}$$

(Here u and u' both have values in U, y_1 and y_1' both have have values in Y.) With this substitution for $\tilde{\mathscr{P}}$, Figure 8 takes the form of Figure 9. Recall that internal stability for this configuration means that the output and internal signals z, u, y are stable whenever the external disturbances w, v_1, v_2 are stable. Now consider the modified configuration where the

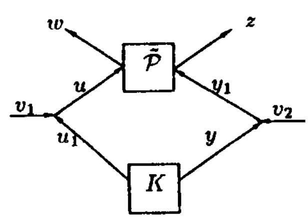

Figure 8

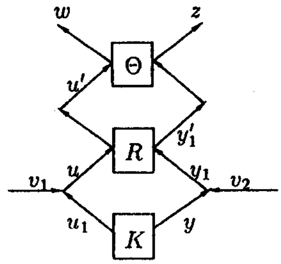

Figure 9

disturbances are shifted to the other side of R as in Figure 10. The compensator K' is considered to be the map from y_1' to u_1' and the plant $\mathscr{P}'$ the map from

$$\begin{bmatrix} w \\ u'' \end{bmatrix} \text{ to } \begin{bmatrix} z \\ y_1'' \end{bmatrix}$$

internal stability for this modified system $\Sigma(\mathscr{P}', K')$ means: the output and internal signals z, u'', y_1' are stable whenever the external disturbances w, v_1', v_2' are stable. It is intuitively plausible that internal stability of $\Sigma(\mathscr{P}, K)$ is equivalent to internal stability of $\Sigma(\mathscr{P}', K')$ if it is the case that R is outer. (Here we say that the rational matrix function R is *outer* if both R and R^{-1} are stable, or, in engineering terminology, R is stable and minimum phase.) One can almost do the proof with pictures; however, a formal proof does require some algebra.

A special case of the following result, the main result of this section for the 2-block case, appears as Lemma 1 in Reference 3.

Theorem 2.1.1

Suppose that the rational matrix function $\tilde{\mathscr{P}}$ has a factorization $\tilde{\mathscr{P}} = \Theta \circ R$ with the rational matrix function R outer. Here we assume that $\tilde{\mathscr{P}}$ arises from a rational matrix function

$$\mathscr{P} = \begin{bmatrix} \mathscr{P}_{11} & \mathscr{P}_{12} \\ \mathscr{P}_{21} & \mathscr{P}_{22} \end{bmatrix}$$

as in (38) where $\mathscr{P}_{21}$ is invertible. Then the system $\Sigma(\mathscr{P}, K)$ depicted in Figure 9 is internally stable if and only if the system $\Sigma(\mathscr{P}', K')$ depicted in Figure 10 is internally stable.

The proof of Theorem 2.1.1 requires a couple of lemmas.

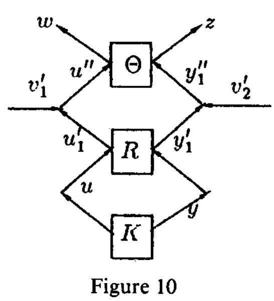

Figure 10

Lemma 2.1.2

Consider the system of equations

$$\begin{bmatrix} \tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{12} \\ \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} \end{bmatrix} \begin{bmatrix} u \\ y_1 \end{bmatrix} = \begin{bmatrix} z \\ w \end{bmatrix} \tag{39a}$$

$$u = K(y_1 + v_2) + v_1 \tag{39b}$$

Then one can solve for z, u, y_1 in terms of w, v_1, v_2 if and only if $\tilde{\mathscr{P}}_{21}K + \tilde{\mathscr{P}}_{22}$ is invertible, in which case

$$\tilde{\mathscr{H}} \begin{bmatrix} w \\ v_1 \\ -v_1 \end{bmatrix} = \begin{bmatrix} z \\ u \\ y_1 \end{bmatrix}$$

where

$$\tilde{\mathscr{H}} = \begin{bmatrix} I & -\tilde{\mathscr{P}}_{11} & -\tilde{\mathscr{P}}_{12} \\ 0 & -\tilde{\mathscr{P}}_{21} & -\tilde{\mathscr{P}}_{22} \\ 0 & I & -K \end{bmatrix}^{-1} \begin{bmatrix} 0 & 0 & 0 \\ -I & 0 & 0 \\ 0 & I & -K \end{bmatrix} \tag{40}$$

Thus, if $\mathscr{P}_{21}$ is invertible, then the system $\Sigma(\mathscr{P}, K)$ in Figure 7 is internally stable if and only if $\tilde{\mathscr{H}}$ given by (40) exists and is stable, where $\tilde{\mathscr{P}}$ is given by (38).

Proof. Once the formula (40) for $\tilde{\mathscr{H}}$ is verified, the remaining assertions follow from the definitions. To verify (40), rearrange the system of equations (39a) and (39b) in matrix form

$$\begin{bmatrix} 0 & 0 & 0 \\ -I & 0 & 0 \\ 0 & I & -K \end{bmatrix} \begin{bmatrix} w \\ v_1 \\ -v_2 \end{bmatrix} + \begin{bmatrix} -I & \tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{12} \\ 0 & \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} \\ 0 & -I & K \end{bmatrix} \begin{bmatrix} z \\ u \\ y_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Well-posedness thus requires invertibility of

$$\begin{bmatrix} -I & \tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{12} \\ 0 & \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} \\ 0 & -I & K \end{bmatrix}$$

or equivalently, of

$$\begin{bmatrix} \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} \\ -I & K \end{bmatrix}$$

By a Schur complement test, this is easily seen to be equivalent to invertibility of $\tilde{\mathscr{P}}_{21}K + \tilde{\mathscr{P}}_{22}$. Solving for z, u, y_1 in terms of $w, v_1, -v_2$ now leads to formula (40) for $\tilde{\mathscr{H}}$. □

To emphasize the dependence of $\tilde{\mathscr{H}}$ in (40) on $\tilde{\mathscr{P}}$ and K, we shall often write $\tilde{\mathscr{H}}(\tilde{\mathscr{P}}, K)$ when more than one pair $(\tilde{\mathscr{P}}, K)$ is under consideration.

Lemma 2.1.3

(See Theorem 3.5 of Reference 5 for a symmetric version.) Suppose

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

is invertible (with R_{11} and R_{22} square) and

$$R^{-1} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}$$

Then $R_{21}K + R_{22}$ is invertible if and only if $r_{11} - Kr_{21}$ is invertible, and then

$$(R_{11}K + R_{12})(R_{21}K + R_{22})^{-1} = -(r_{11} - Kr_{21})^{-1}(r_{12} - Kr_{22}) \tag{41}$$

Proof. Suppose K is such that $R_{21}K + R_{22}$ is invertible. Then define

$$\hat{R} = \begin{bmatrix} \hat{R}_{11} & \hat{R}_{12} \\ \hat{R}_{21} & \hat{R}_{22} \end{bmatrix} \overset{\text{def}}{=} \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} I & K \\ 0 & I \end{bmatrix}$$

Then $\hat{R}$ is invertible with $\hat{R}^{-1}$ given by

$$\hat{R}^{-1} = \begin{bmatrix} \hat{r}_{11} & \hat{r}_{12} \\ \hat{r}_{21} & \hat{r}_{22} \end{bmatrix} \overset{\text{def}}{=} \begin{bmatrix} I & -K \\ 0 & I \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}$$

Then we see that $\hat{R}_{22} = R_{21}K + R_{22}$ is invertible. Then by a well-known Schur complement argument (see, for example, Reference 19) $\hat{r}_{11} = r_{11} - Kr_{21}$ is invertible. The converse follows by replacing the roles of R and R^{-1}. To prove (41) we compute

$$\begin{aligned} 0 &= [I, \quad -K] \begin{bmatrix} K \\ I \end{bmatrix} \\ &= [I, \quad K] R^{-1} \circ R \begin{bmatrix} K \\ I \end{bmatrix} \\ &= [r_{11} - Kr_{21}, r_{12} - Kr_{22}] \begin{bmatrix} R_{11}K + R_{12} \\ R_{21}K + R_{22} \end{bmatrix} \\ &= (r_{11} - Kr_{21}) [I, (r_{11} - Kr_{21})^{-1}(r_{12} - Kr_{22})] \circ \begin{bmatrix} (R_{11}K + R_{12})(R_{21}K + R_{22})^{-1} \\ I \end{bmatrix} (R_{21}K + R_{22}) \end{aligned}$$

Since $r_{11} - Kr_{21}$ and $R_{21}K + R_{22}$ are non-singular, we conclude that

$$[I, (r_{11} - Kr_{21})^{-1}(r_{12} - Kr_{22})] \begin{bmatrix} (R_{11}K + R_{12})(R_{21}K + R_{22})^{-1} \\ I \end{bmatrix} = 0 \qquad \square$$

Proof of Theorem 2.1.1. By Lemma 2.1.2 the content of Theorem 2.1.1 is that, given that R is outer, then $\tilde{\mathscr{H}}(\Theta \circ R, K)$ is stable if and only if $\tilde{\mathscr{H}}(\Theta, \mathscr{G}_R[K])$ is stable. From (40) we have

$$\begin{aligned} \tilde{\mathscr{H}}(\Theta \circ R, K) &= \left[\begin{array}{c|c} I & -\Theta \circ R \\ 0 & \\ \hline 0 & [I - K] \end{array}\right]^{-1} \begin{bmatrix} 0 & 0 & 0 \\ -I & 0 & 0 \\ 0 & I & -K \end{bmatrix} \\ &= \left\{ \left[\begin{array}{c|c} I & -\Theta \\ 0 & \\ \hline 0 & [I, \ -K] R^{-1} \end{array}\right] \left[\begin{array}{c|c} I & 0 \quad 0 \\ \hline 0 & R \\ 0 & \end{array}\right] \right\}^{-1} \begin{bmatrix} 0 & 0 & 0 \\ -I & 0 & 0 \\ 0 & I & -K \end{bmatrix} \end{aligned}$$

Note that

$$\begin{aligned}[I,\ -K]R^{-1} &= [r_{11} - Kr_{21}, r_{12} - Kr_{22}] \\ &= (r_{11} - Kr_{21})[I, (r_{11} - Kr_{21})^{-1}(r_{12} - Kr_{22})] \\ &= (r_{11} - Kr_{21})[I,\ -(R_{11}K + R_{12})(R_{21}K + R_{22})^{-1}] \end{aligned} \tag{42}$$

where we used Lemma 2.1.3 for the last step. We have thus verified

$$[I,\ -K]R^{-1} = (r_{11} - Kr_{21})[I,\ -\mathcal{G}_R[K]]$$

Hence

$$\tilde{\mathcal{H}}(\Theta \circ R, K) = \left\{ \begin{bmatrix} I & \vdots & 0 \quad 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & \\ 0 & \vdots & R^{-1} \end{bmatrix} \begin{bmatrix} I & \vdots & \\ 0 & \vdots & -\Theta \\ \cdots & \cdots & \cdots \\ 0 & \vdots & [I,\ -\mathcal{G}_R[K]] \end{bmatrix}^{-1} \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & (r_{11} - Kr_{21})^{-1} \end{bmatrix} \right\} \begin{bmatrix} 0 & 0 & 0 \\ -I & 0 & 0 \\ 0 & I & -K \end{bmatrix} \tag{43}$$

Now use (42) in the form

$$(r_{11} - Kr_{21})^{-1}[I - K] = [I,\ -\mathcal{G}_R[K]]R$$

When the dust settles, this combined with (43) leads to the fundamental identity

$$\tilde{\mathcal{H}}(\Theta \circ R, K) = \begin{bmatrix} I & \vdots & 0 \quad 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & \\ 0 & \vdots & R^{-1} \end{bmatrix} \tilde{\mathcal{H}}(\Theta,\ \mathcal{G}_R[K]) \begin{bmatrix} I & \vdots & 0 \quad 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & \\ 0 & \vdots & R \end{bmatrix} \tag{44}$$

From this we read off that $\tilde{\mathcal{H}}(\Theta \circ R, K)$ is stable if and only if $\tilde{\mathcal{H}}(\theta,\ \mathcal{G}_R[K])$ is stable whenever R is outer. □

2.2. *The general 4-block case*

The goal of this subsection is to obtain the analogue of Theorem 2.1.1 for the general (regular) 4-block case. Thus assume that we are given a rational matrix function (the plant)

$$\mathcal{P} = \begin{bmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{21} & \mathcal{P}_{22} \end{bmatrix}$$

with $\mathcal{P}_{21}$ surjective and $\mathcal{P}_{12}$ injective (as matrices over the field of scalar rational functions). Consider any convenient augmentation of $\mathcal{P}$ to

$$\boldsymbol{\mathcal{P}} = \begin{bmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \cdots & \cdots \\ \mathcal{P}_{21} & \mathcal{P}_{22} \\ \mathcal{P}_{21}^{0} & \mathcal{P}_{22}^{0} \end{bmatrix}$$

such that the augmented (2, 1)-corner $\begin{bmatrix} \mathscr{P}_{21} \\ \mathscr{P}_{21}^0 \end{bmatrix}$ is square and invertible. We then transform to

$$\tilde{\mathscr{P}} = \begin{bmatrix} \tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{12} & \tilde{\mathscr{P}}_{13} \\ \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} & \tilde{\mathscr{P}}_{23} \end{bmatrix}$$

via formula (20) so that the system of equations

$$\mathscr{P}\begin{bmatrix} w \\ u \end{bmatrix} = \begin{bmatrix} z \\ y \\ y^0 \end{bmatrix}$$

is equivalent to the system

$$\tilde{\mathscr{P}}\begin{bmatrix} z \\ y \\ y^0 \end{bmatrix} = \begin{bmatrix} z \\ w \end{bmatrix}$$

The system configuration $\Sigma(\mathscr{P}, K)$ depicted in Figure 7 can alternatively be expressed in the form depicted in Figure 11 with associated system of algebraic equations.

$$\tilde{\mathscr{P}}\begin{bmatrix} u \\ y \\ y^0 \end{bmatrix} = \begin{bmatrix} z \\ w \end{bmatrix} \tag{45a}$$

$$[K \;\; 0]\begin{bmatrix} y + v_2 \\ y^0 \end{bmatrix} + v_1 = u \tag{45b}$$

Now let us suppose that the size of block $\tilde{\mathscr{P}}_{ij}$ is $m_i \times n_j$ for $i = 1, 2$ and $j = 1, 2, 3$ (so $m_2 = n_2 + n_3$ and $m_1 \geqslant n_1$) and that $\tilde{\mathscr{P}}$ factors as

$$\tilde{\mathscr{P}} = \Theta R \tag{46}$$

where

$$\Theta = \begin{bmatrix} \Theta_{11} & \Theta_{12} & \Theta_{13} \\ \Theta_{21} & \Theta_{22} & \Theta_{23} \end{bmatrix}$$

also has blocks Θ_{ij} of the same size as the corresponding block $\tilde{\mathscr{P}}_{ij}$ of $\tilde{\mathscr{P}}$ and where $R = [R_{ij}]_{1 \leqslant i, j \leqslant 3}$ has 3×3 block matrix structure with the size of R_{ij} equal to $n_i \times n_j$.

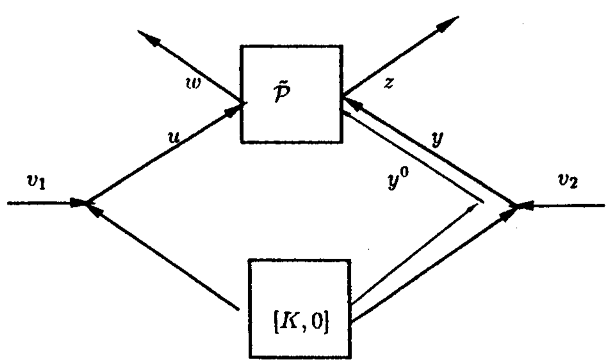

Figure 11

Substituting (46) for $\tilde{\mathscr{P}}$ in Figure 11 leads to the system depicted in Figure 12 with associated set of algebraic equations

$$\Theta\begin{bmatrix} u \\ y \\ y_0 \end{bmatrix} = \begin{bmatrix} z \\ w \end{bmatrix}$$

$$R\begin{bmatrix} u_1 \\ y_1 \\ y_1^0 \end{bmatrix} = \begin{bmatrix} u \\ y \\ y^0 \end{bmatrix}$$

$$[K\ 0]\begin{bmatrix} y_1 + v_2 \\ y_1^0 \end{bmatrix} + v_1 = u_1$$

We wish to analyse internal stability, not for the system in Figure 12, but rather for the system $\Sigma(\mathscr{P}, K)$ from which it came, namely, that the output and internal signals z, u_1, y_1 should be stable whenever the disturbance signals w, v_1, v_2 are stable. The signal y_1^0 is to be considered as physically fictitious, added merely for mathematical convenience. To avoid confusion, let us refer to this criterion as *restricted internal stability* for the system in Figure 12. A seemingly minor modification of Figure 12 is obtained by shifting the disturbances v_1 and v_2 to the other side of the box R; the result is depicted in Figure 13. We consider the bottom two boxes R and $[K\ 0]$ as lumped to define a compensator

$$K': \begin{bmatrix} y - v_2^1 \\ y^0 \end{bmatrix} \to u - v_1'$$

and the plant $\mathscr{P}'$ to be

$$\mathscr{P}': \begin{bmatrix} w \\ u \end{bmatrix} \to \begin{bmatrix} z \\ y \end{bmatrix}$$

We define *restricted internal stability* for the system Figure 13 to mean that the output and internal signals z, u, y are stable whenever the disturbance signals w, v_1', v_2' are stable. The idea

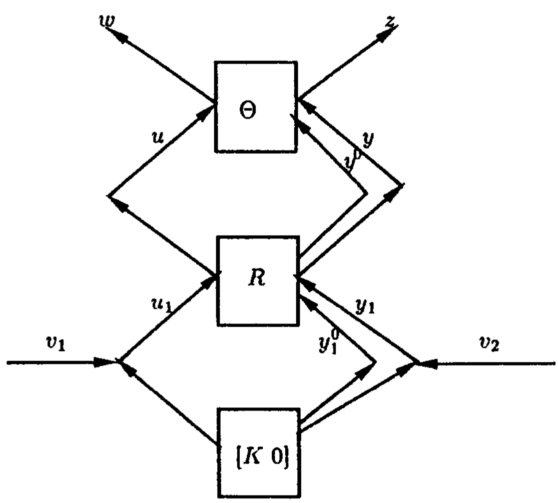

Figure 12

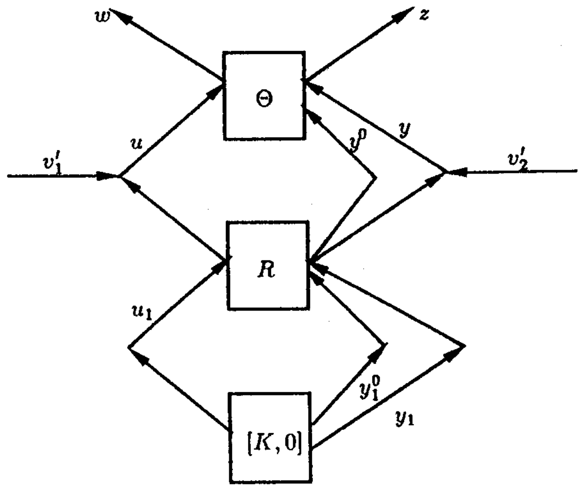

Figure 13

again is that y^0 is physically fictitious added merely for mathematical convenience. The analogue of Theorem 2.1.1 is as follows.

Theorem 2.2.1

Suppose that the rational matrix function

$$\tilde{\mathscr{P}} = \begin{bmatrix} \tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{12} & \tilde{\mathscr{P}}_{13} \\ \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} & \tilde{\mathscr{P}}_{23} \end{bmatrix},$$

arising from the plant

$$\mathscr{P} = \begin{bmatrix} \mathscr{P}_{11} & \mathscr{P}_{12} \\ \mathscr{P}_{21} & \mathscr{P}_{22} \end{bmatrix}$$

as in formula (20) has a factorization

$$\tilde{\mathscr{P}} = \Theta R$$

where R is an invertible block 3×3 rational matrix function of the form

$$R = \begin{bmatrix} R_{11} & R_{12} & 0 \\ R_{21} & R_{22} & 0 \\ R_{31} & R_{32} & R_{33} \end{bmatrix}$$

with

$$\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

outer. Then the system $\Sigma(\mathscr{P}, K)$ in Figure 7 is internally stable if and only if the system $\Sigma(\mathscr{P}', K')$ described by Figure 13 has the property of restricted internal stability.

The proof is most easily presented by first isolating some preliminary lemmas.

Lemma 2.2.2

Suppose that the configuration depicted in Figure 13 is well-posed and R has the block upper triangular form

$$R = \begin{bmatrix} R_{11} & R_{12} & 0 \\ R_{21} & R_{22} & 0 \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \tag{47}$$

Then

$$K' : \begin{bmatrix} y - v_2' \\ y^0 \end{bmatrix} \rightarrow u - v_1'$$

is given by

$$K' = [\mathcal{G}_{r_0}[K], 0]$$

where

$$R_0 = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

and

$$\mathcal{G}_{R_0}[K] = (R_{11}K + R_{12})(R_{21}K + R_{22})^{-1}$$

Proof. In Figure 13 set

$$u_2 = u - v_1', \qquad y_2 = y - v_2'$$

so we have that K' is defined by

$$K' \begin{bmatrix} y_2 \\ y^0 \end{bmatrix} = u_2$$

where

$$R \begin{bmatrix} u_1 \\ y_1 \\ y_1^0 \end{bmatrix} = \begin{bmatrix} u_2 \\ y_2 \\ y^0 \end{bmatrix}$$

and

$$[K \;\; 0] \begin{bmatrix} y_1 \\ y_1^0 \end{bmatrix} = u_1$$

Writing out in detail, we have

$$\begin{aligned} R_{11}u_1 + R_{12}y_1 &= u_2 \\ R_{21}u_1 + R_{22}y_1 &= y_2 \\ R_{31}u_1 + R_{32}y_1 &= y^0 \\ Ky_1 &= u_1 \end{aligned}$$

To solve for K' we must solve for u_2 in terms of y_1 and y_1^0. From the form of these equations we see that the last row of R is irrelevant and we have

$$u_1 = (R_{11}K + R_{12})y_1$$

where

$$y_2 = (R_{21}K + R_{22})y_1$$

This leads to

$$u_2 = [\mathscr{G}_{R_0}[K], 0]\begin{bmatrix} y_2 \\ y_1^0 \end{bmatrix}$$

as asserted. □

The significance of Lemma 2.2.2 we now see is that the compensator K' in Figure 13 again has the form $[\tilde{K}_0' \; 0]$ (making the role of $\tilde{y}$ again that of a fictitious signal) while the compensator $\tilde{K}$ at the bottom of the figure has the form $[K \; 0]$ (making $\tilde{y}$ a fictitious signal), under the assumption that R has the block upper triangular structure (47).

Lemma 2.2.3

Consider the system of equations

$$\begin{bmatrix} \tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{12} & \tilde{\mathscr{P}}_{13} \\ \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} & \tilde{\mathscr{P}}_{23} \end{bmatrix}\begin{bmatrix} u \\ y \\ y_0 \end{bmatrix} = \begin{bmatrix} z \\ w \end{bmatrix} \tag{48a}$$

$$u - v_1 = [K \; 0]\begin{bmatrix} y + v_2 \\ y_0 \end{bmatrix} \tag{48b}$$

Then one can solve uniquely for z, u, y in terms of w, v_1, v_2 if and only if $[\tilde{\mathscr{P}}_{21}K + \tilde{\mathscr{P}}_{22}, \tilde{\mathscr{P}}_{23}]$ is invertible. In this case the mapping

$$\begin{bmatrix} w \\ v_1 \\ v_2 \end{bmatrix} \to \begin{bmatrix} z \\ u \\ y \end{bmatrix} \quad \text{is given by} \quad \begin{bmatrix} z \\ u \\ y \end{bmatrix} = \tilde{\mathscr{H}}\begin{bmatrix} w \\ v_1 \\ v_2 \end{bmatrix}$$

where $\tilde{\mathscr{H}} = \tilde{\mathscr{H}}(\tilde{\mathscr{P}}, K)$ is given by

$$\tilde{\mathscr{H}} = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \end{bmatrix}\left[\begin{array}{c:c} \begin{matrix} I \\ 0 \end{matrix} & -\tilde{\mathscr{P}} \\ \hdashline 0 & I - K \quad 0 \end{array}\right]^{-1}\begin{bmatrix} 0 & 0 & 0 \\ -I & 0 & 0 \\ 0 & I & -K \end{bmatrix} \tag{49}$$

Thus, if $\tilde{\mathscr{P}}$ arises from the regular plant $\mathscr{P}$ as in (20), then the system $\Sigma(\mathscr{P}, K)$ in Figure 7 is internally stable if and only if $\tilde{\mathscr{H}}(\tilde{\mathscr{P}}, K)$ is stable.

Proof. As in the proof of Lemma 2.1.2, all assertions follow from the definitions once the formula (49) for $\tilde{\mathscr{H}}$ is verified.

To verify (49) let use view $\tilde{\mathscr{P}}$ as a block 2×2 matrix function $\tilde{\mathscr{P}}$ by merging the last two block columns to a single block column and then apply Lemma 2.1.2 with $[K, 0]$ in place of K (see Figure 14).

Then by Lemma 2.1.2 the mapping

$$\tilde{\mathscr{H}}: \begin{bmatrix} w \\ v_1 \\ v_2 \\ v_2^0 \end{bmatrix} \to \begin{bmatrix} z \\ u \\ y \\ y^0 \end{bmatrix}$$

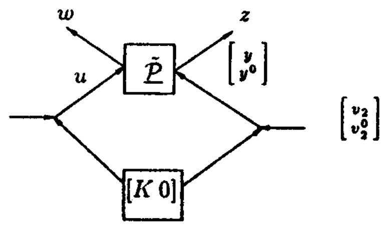

Figure 14

is given by

$$\tilde{\mathcal{H}} = \tilde{\mathcal{H}}(\tilde{\mathcal{P}}, [K\ \ 0]) = \begin{bmatrix} I & -\tilde{\mathcal{P}}_{11} & \vdots & \tilde{\mathcal{P}}_{22} & -\tilde{\mathcal{P}}_{13} \\ 0 & -\tilde{\mathcal{P}}_{21} & \vdots & -\tilde{\mathcal{P}}_{22} & -\tilde{\mathcal{P}}_{23} \\ 0 & I & \vdots & -K & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 0 & \vdots & 0 & 0 \\ -I & 0 & \vdots & 0 & 0 \\ 0 & I & \vdots & -K & 0 \end{bmatrix}$$

However, in the context of the 4-block problem the signal u^0 is factitious and the disturbance v_2^0 is irrelevant. The desired $\tilde{\mathcal{H}} = \tilde{\mathcal{H}}(\tilde{\mathcal{P}}, K)$ is had by simply discarding the last row and last column of $\tilde{\mathcal{H}}$; this results in the formula (49) for $\tilde{\mathcal{H}}$. □

We are now ready for the proof of Theorem 2.2.1.

Proof of Theorem 2.2.1. By Lemma 2.2.2, restricted internal stability of the configuration in Figure 12, which by the discussion above is equivalent to internal stability of $\Sigma(\mathcal{P}, K)$ in Figure 7, amounts to stability of the rational matrix function

$$\tilde{\mathcal{H}}(\Theta \circ R, K) = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \end{bmatrix} \left[\begin{array}{c:ccc} I & & -\Theta \circ R & \\ 0 & & & \\ \hdashline 0 & I & -K & 0 \end{array}\right]^{-1} \begin{bmatrix} 0 & 0 & 0 \\ -I & 0 & 0 \\ 0 & I & -K \end{bmatrix}$$

When R has the upper triangular form (47), by Lemma 2.2.2 combined with Lemma 2.2.3 we have that restricted internal stability of the configuration in Figure 13 is equivalent to stability of the rational matrix function

$$\tilde{\mathcal{H}}(\Theta, \mathcal{G}_{R_0}[K]) = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \end{bmatrix} \left[\begin{array}{c:ccc} I & & -\Theta & \\ 0 & & & \\ \hdashline 0 & I & -\mathcal{G}_{R_0}[K] & 0 \end{array}\right]^{-1} \begin{bmatrix} 0 & 0 & 0 \\ -I & 0 & 0 \\ 0 & I & -\mathcal{G}_{R_0}[K] \end{bmatrix}$$

The content of Theorem 2.2.1 therefore is: if R has the special properties in the hypotheses of Theorem 2.2.1, then $\tilde{\mathcal{H}}(\Theta \circ R, K)$ is stable if and only if $\tilde{\mathcal{H}}(\Theta, \mathcal{G}_{R_0}[K])$ is stable.

The idea for verifying this latter fact is the same as in the proof of Theorem 2.1.1 but the details are a shade more subtle. We begin with

$$\left[\begin{array}{c:ccc} I & & -\Theta \circ R & \\ 0 & & & \\ \hdashline 0 & I & -K & 0 \end{array}\right] = \left[\begin{array}{c:c} I & -\Theta \\ 0 & \\ \hdashline 0 & [I\ \ -K\ \ 0]R^{-1} \end{array}\right] \left[\begin{array}{c:ccc} I & 0 & 0 & 0 \\ \hdashline 0 & & & \\ 0 & & R & \\ 0 & & & \end{array}\right] \tag{50}$$

Decompose R^{-1} as

$$R^{-1} = \begin{bmatrix} r_{11} & r_{12} & 0 \\ r_{21} & r_{22} & 0 \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = \left[\begin{array}{cc:c} R_0^{-1} & & 0 \\ & & 0 \\ \hdashline r_{31} & r_{32} & r_{33} \end{array}\right]$$

Thus

$$[I-K\ \ 0]R^{-1}=[r_{11}-Kr_{21}\ \vdots\ r_{12}-Kr_{22}\ \vdots\ 0]$$
$$=(r_{11}-Kr_{21})[I-\mathscr{G}_{r_0}[K]\ \ 0] \tag{51}$$

where we used Lemma 2.1.3 for the last step. Next observe that

$$\begin{bmatrix} I & 0 & 0 & 0\\ 0 & I & 0 & 0\\ 0 & 0 & I & 0\end{bmatrix}\begin{bmatrix} I & 0 & 0 & 0\\ \hdashline 0 & r_{11} & r_{12} & 0\\ 0 & r_{21} & r_{22} & 0\\ 0 & r_{31} & r_{32} & r_{33}\end{bmatrix}=\begin{bmatrix} I & \vdots & 0 \quad 0\\ \hdashline 0 & \vdots & \\ 0 & \vdots & R^{-1}\end{bmatrix}\begin{bmatrix} I & 0 & 0 & 0\\ 0 & I & 0 & 0\\ 0 & 0 & I & 0\end{bmatrix} \tag{52}$$

When (51) and (52) are combined with (50) we get

$$\begin{bmatrix} I & 0 & 0 & 0\\ 0 & I & 0 & 0\\ 0 & 0 & I & 0\end{bmatrix}\begin{bmatrix} I & \vdots & -\Theta\circ R & \\ 0 & \vdots & & \\ \hdashline 0 & \vdots & I \quad -K & 0\end{bmatrix}^{-1}=\begin{bmatrix} I & \vdots & 0 \quad 0\\ \hdashline 0 & \vdots & \\ 0 & \vdots & R_0^{-1}\end{bmatrix}$$
$$\begin{bmatrix} I & 0 & 0 & 0\\ 0 & I & 0 & 0\\ 0 & 0 & I & 0\end{bmatrix}\begin{bmatrix} I & \vdots & -\Theta & \\ 0 & \vdots & & \\ \hdashline 0 & \vdots & I \quad -\mathscr{G}_{R_0}[K] & 0\end{bmatrix}^{-1}\begin{bmatrix} I & 0 & 0\\ 0 & I & 0\\ 0 & 0 & (r_{11}-Kr_{21})^{-1}\end{bmatrix} \tag{53}$$

Another consequence of Lemma 2.1.3 is

$$(r_{11}-Kr_{21})^{-1}[I,\ -K]=[I,\ -\mathscr{G}_{r_0}[K]]R_0 \tag{54}$$

Finally combine (53) and (54) to get

$$\tilde{\mathscr{H}}(\Theta\circ R,K)=\begin{bmatrix} I & \vdots & 0 \quad 0\\ \hdashline 0 & \vdots & \\ 0 & \vdots & R_0^{-1}\end{bmatrix}\tilde{\mathscr{H}}(\Theta,\ \mathscr{G}_{R_0}[K])\begin{bmatrix} I & \vdots & 0 \quad 0\\ \hdashline 0 & \vdots & \\ 0 & \vdots & R_0\end{bmatrix}$$

From this fundamental identity we read off that, given that R_0 is outer, $\tilde{\mathscr{H}}(\Theta\circ R,K)$ is stable if and only if $\tilde{\mathscr{H}}(\Theta,\ \mathscr{G}_{R_0}[K])$ is stable as needed. □

Note that Theorem 2.1.1 is just the special case of Theorem 2.2.1 where the third block column of $\tilde{\mathscr{P}}$ is trivial.

3. PARAMETRIZATION OF STABILIZING COMPENSATORS

We consider again the feedback system $\Sigma(\mathscr{P},K)$ depicted in Figure 7. A common philosophy in control engineering is to consider the plant $\mathscr{P}$ as given and to use the compensator K as a design parameter to achieve some desired performance characteristics (quantitative and/or qualitative) of the closed-loop system. Since internal stability (as defined in Section 2) is always one such characteristic, it is particularly useful to have a parametrization of the set of all stabilizing compensators for a given plant $\mathscr{P}$, i.e. the set of all compensators K for which the closed-loop system $\Sigma(\mathscr{P},K)$ is internally stable. In addition it is usually also demanded that K itself be proper. A solution to this problem, including many refinements and generalizations, has been known now for some time in the engineering community (see References 1 and 2), and is usually referred to as the Youla parametrization of stabilizing compensators. Here we would like to derive such a parametrization in a different way as a simple application of the general factorization principle (Theorems 2.1.1 and 2.2.1) in Section 2. This section is independent of the other sections of the paper.

In general we say that a given plant

$$\mathscr{P}=\begin{bmatrix}\mathscr{P}_{11} & \mathscr{P}_{12}\\ \mathscr{P}_{21} & \mathscr{P}_{22}\end{bmatrix}$$

is *stabilizable* if and only if there exists a compensator K for which the system $\Sigma(\mathscr{P},K)$ in Figure 7 is stable. We first describe a particular class of plants $\mathscr{P}$ (referred to as the *model-matching* case in the literature) for which the identification of stabilizing compensators is particularly simple.

Theorem 3.1

Suppose $\mathscr{P}$ has the form

$$\mathscr{P}=\begin{bmatrix}\mathscr{P}_{11} & \mathscr{P}_{12}\\ \mathscr{P}_{21} & 0\end{bmatrix}$$

(i.e. $\mathscr{P}_{22}=0$). Then $\mathscr{P}$ is stabilizable if and only if $\mathscr{P}_{11}, \mathscr{P}_{12}, \mathscr{P}_{22}$ are all stable. In this case K is stabilizing for $\mathscr{P}$ if and only if K itself is stable.

Proof. The compensator K is stabilizing for $\mathscr{P}$ if and only if the rational matrix function $\mathscr{H}$ given by (37) is stable. For the case where $\mathscr{P}_{22}=0$, $\Delta=I$ and $\mathscr{H}$ collapses to

$$\mathscr{H}=\begin{bmatrix}\mathscr{P}_{11}+\mathscr{P}_{12}K\mathscr{P}_{21} & \mathscr{P}_{12} & \mathscr{P}_{12}K\\ K\mathscr{P}_{22} & I & K\\ \mathscr{P}_{21} & 0 & I\end{bmatrix}$$

We now read off that $\mathscr{H}$ is stable exactly when each of $K, \mathscr{P}_{11}, \mathscr{P}_{12}, \mathscr{P}_{21}$ is stable, as asserted. □

Combining Theorem 3.1 with Theorem 2.2.1 leads to the following.

Theorem 3.2

Let

$$\mathscr{P}=\begin{bmatrix}\mathscr{P}_{11} & \mathscr{P}_{12}\\ \mathscr{P}_{21} & \mathscr{P}_{22}\end{bmatrix}$$

be a given plant with $\mathscr{P}_{12}$ injective and $\mathscr{P}_{21}$ surjective and form

$$\tilde{\mathscr{P}}=\begin{bmatrix}\tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{12} & \tilde{\mathscr{P}}_{13}\\ \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} & \tilde{\mathscr{P}}_{23}\end{bmatrix}$$

with $\tilde{\mathscr{P}}$ given by (20). Suppose $\tilde{\mathscr{P}}$ has a factorization

$$\tilde{\mathscr{P}}=\Theta R$$

where

(i) R has the block upper-triangular form

$$R=\begin{bmatrix}R_{11} & R_{12} & 0\\ R_{21} & R_{22} & 0\\ R_{31} & R_{32} & R_{33}\end{bmatrix}$$

with $R_0 := \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$ outer

and

(ii) Θ has the triangular form

$$\Theta = \begin{bmatrix} \Theta_{11} & \Theta_{12} & \Theta_{13} \\ 0 & \Theta_{22} & \Theta_{23} \end{bmatrix}$$

Set

$$\begin{bmatrix} T_3 \\ T_3^0 \end{bmatrix} = [\Theta_{22} \ \Theta_{23}]^{-1} \tag{55}$$

$$T_2 = -\Theta_{11} \tag{56}$$

$$T_1 = [\Theta_{12} \ \Theta_{13}][\Theta_{22} \ \Theta_{23}]^{-1} \tag{57}$$

The $\mathscr{P}$ is stabilizable if and only if T_1, T_2, T_3 are stable. In this case K stabilizes $\mathscr{P}$ if and only if

$$K = \mathscr{G}_{R_0^{-1}}[Q]$$

for a stable Q such that $Q(\infty)$ is in the domain of definition $\mathscr{G}_{R_0^{-1}(\infty)}$.

Proof. By Theorem 2.2.1 internal stability of $\Sigma(\mathscr{P}, K)$ is equivalent to restricted internal stability for the system depicted in Figure 13. But this is the same as internal stability for the system $\Sigma(\mathscr{P}', \mathscr{G}_{R_0}[K])$, where $\mathscr{P}'$ is determined from Θ via formula (23) (with Θ in place of $\tilde{\mathscr{P}}$ and $\mathscr{P}'$ in place of $\mathscr{P}$). These formulas work out to give

$$\mathscr{P}' = \begin{bmatrix} T_1 & T_2 \\ T_3 & 0 \end{bmatrix}$$

where T_1, T_2, T_3 are given by (55)–(57) since by assumption $\Theta_{21} = 0$. Thus $\mathscr{P}'$ has the model-matching form so Theorem 3.1 applies. We conclude that the original system $\Sigma(\mathscr{P}, K)$ is stabilizable if and only if the three rational matrix functions T_1, T_2, T_3 are stable, and in this case K stabilizes if and only if $Q := \mathscr{G}_{R_0}[K]$ is stable. Back-solving for K gives that $K = \mathscr{G}_{R_0^{-1}}[Q]$ for some stable rational matrix function Q. The condition that $Q(\infty)$ be in the domain of definition of the (constant matrix) linear fractional map $\mathscr{G}_{R_0^{-1}(\infty)}$ arises from the restriction that K be proper. □

Remark 1. If $\tilde{\mathscr{P}}$ is any injective rational matrix function with $[\tilde{\mathscr{P}}_{22} \ \tilde{\mathscr{P}}_{23}]$ square and invertible, a factorization $\tilde{\mathscr{P}} = \Theta R$ with R and Θ satisfying conditions (i) and (ii) in Theorem 3.2 is always possible. Indeed, given that $\tilde{\mathscr{P}}$ and Θ are both injective, a factorization of the form $\tilde{\mathscr{P}} = \Theta R$ as described in the theorem is equivalent to $\tilde{\mathscr{P}}$ and Θ generating the same subspace $\mathscr{M}$ (in fact modules over the ring $\mathscr{S}$ of stable scalar rational functions) of $\mathscr{R}^{m_1+m_2}$ (column vector rational functions with $m_1 + m_2$ components) given by

$$\mathscr{S} := \tilde{\mathscr{P}}(\mathscr{S}^{n_1} \oplus \mathscr{S}^{n_2} \oplus \mathscr{R}^{n_3}) = \Theta(\mathscr{S}^{n_1} \oplus \mathscr{S}^{n_2} \oplus \mathscr{R}^{n_3}) \tag{58}$$

(where $\mathscr{S}^n$ denotes the space of stable rational column vector functions with n components); this fact can be seen from the easily verified characterization of the class of factors R described by (i) in Theorem 3.2 as exactly those $(n_1 + n_2 + n_3) \times (n_1 + n_2 + n_3)$ rational matrix functions R such that

$$R(\mathscr{S}^{n_1} \oplus \mathscr{S}^{n_2} \oplus \mathscr{R}^{n_3}) = \mathscr{S}^{n_1} \oplus \mathscr{S}^{n_2} \oplus \mathscr{R}^{n_3}$$

We now sketch how to construct $(m_1+m_2)\times(n_1+n_2+n_3)$ rational matrix function Θ which satisfies (58) for a given $\tilde{\mathscr{P}}$.

Set $\mathscr{S}_- = [\tilde{\mathscr{P}}_{21}\ \tilde{\mathscr{P}}_{22}\ \tilde{\mathscr{P}}_{23}](\mathscr{S}^{n_1}\oplus\mathscr{S}^{n_2}\oplus\mathscr{R}^{n_3})$. Since $[\tilde{\mathscr{P}}_{22}\ \tilde{\mathscr{P}}_{23}]$ is invertible one can see that $\mathscr{S}_-$ is the direct sum of a free $\mathscr{S}$-module with n_2 generators and an $\mathscr{R}$-vector subspace of dimension n_3. By choosing bases, it follows that $\mathscr{S}_-$ can be represented as

$$\mathscr{S}_- = [\Theta_{22}\ \Theta_{23}](\mathscr{S}^{n_2}\oplus\mathscr{R}^{n_3})$$

where Θ_{22} and Θ_{23} are rational matrix functions of sizes $m_2\times n_2$ and $m_2\times n_3$ respectively, and where $[\Theta_{22}\ \Theta_{23}]$ is invertible; construction of such a basis can be carried out by an adaptation of the algorithm in Forney's well-known paper[20] involving a sequence of column operations. From the definition of $\mathscr{S}_-$ there then exist rational matrix functions Θ_{12} and Θ_{23} of sizes $m_1\times n_2$ and $m_1\times n_3$ so that

$$\begin{bmatrix}\Theta_{12}\\ \Theta_{22}\end{bmatrix}\mathscr{S}^{n_2}\subset\mathscr{M},\qquad \begin{bmatrix}\Theta_{13}\\ \Theta_{23}\end{bmatrix}\mathscr{R}^{n_3}\subset\mathscr{M}$$

Now $\mathscr{M}$ itself is the direct sum of a free $\mathscr{S}$-module with n_1+n_2 generators and an $\mathscr{R}$-subspace of dimension n_3. By dimension count,

$$\begin{bmatrix}\Theta_{13}\\ \Theta_{23}\end{bmatrix}\mathscr{R}^{n_3}$$

exhausts the latter subspace, and

$$\begin{bmatrix}\Theta_{22} & \Theta_{13}\\ \Theta_{22} & \Theta_{23}\end{bmatrix}(\mathscr{S}^{n_2}\oplus\mathscr{R}^{n_3})$$

has codimension over $\mathscr{S}$ equal to n_1 in $\mathscr{M}$. To pick up the rest of $\mathscr{M}$, we represent $\mathscr{M}\cap(\mathscr{R}^{m_1}\oplus 0)$ in the form

$$\mathscr{M}\cap(\mathscr{R}^{m_1}\oplus 0)=\Theta_{11}\mathscr{S}^{n_1}\oplus 0$$

where Θ_{11} is a $m_1\times n_1$ rational matrix function. Then

$$\Theta=\begin{bmatrix}\Theta_{11} & \Theta_{12} & \Theta_{13}\\ 0 & \Theta_{22} & \Theta_{23}\end{bmatrix}$$

is the desired first factor for the factorization $\tilde{\mathscr{P}}=\Theta R$ satisfying conditions (i) and (ii) in Theorem 3.2. The factor R is then uniquely determined since Θ is injective.

In principle the construction sketched above could be turned into an explicit algorithm involving column operations on the matrix $\tilde{\mathscr{P}}$ similar to the algorithm in the paper of Forney;[20] keeping track of the column operations would also lead to computation of the factor R as well. However, we do not enter into these details here.

Remark 2. The result of Theorem 3.2 is the same principle as in the Youla parametrization of stabilizing compensators, namely, that there is a block 2×2 outer function R_0^{-1} such that $\mathscr{G}_{R_0^{-1}}$ acting on an essentially free stable parameter matrix function Q yields the set of all stabilizing compensators K for a given plant $\mathscr{P}$. The usual derivation of the outer function R_0^{-1} is through a double coprime factorization of $\tilde{\mathscr{P}}$ (see Reference 1 or 2); our derivation through a (block, upper-triangular)–(restricted outer) factorization of $\tilde{\mathscr{P}}$ appears to be new. The construction in Remark 1 represents a new input–output approach to the construction of the Youla parametrization.

Remark 3. The model matching scheme in Theorem 3.1 is closely associated with the divisor–remainder formulation of interpolation conditions.[21] In particular, in view of Theorem 3.2, internal stability of the system $\Sigma(\mathscr{P}, K)$ can be characterized in terms of the transfer function $\mathscr{F}_{\mathscr{P}}[K]: w \rightarrow z$ being stable and satisfying certain interpolation conditions. This phenomenon has been well-known in principle for some time and is usually derived by using the Youla parametrization. Only recently have the interpolation conditions on $\mathscr{F}_{\mathscr{P}}[K]$ been given explicitly in terms of $\mathscr{P}$ for some special instances of the matrix 1-block case (see Chapters 23 and 25 of Reference 21). In Reference 22 a more general interpolation theory (involving both the usual discrete (or lumped) as well as continuous (or generic) interpolation conditions) is developed and used to obtain a characterization of internal stability of $\Sigma(\mathscr{P}, K)$ in terms of interpolation conditions on $\mathscr{F}_{\mathscr{P}}[K]$ for the general case.

4. STABILIZING COMPENSATORS WITH INFINITY NORM PERFORMANCE MEASURE (THE H_∞ PROBLEM)

We return again to the feedback system $\Sigma(\mathscr{P}, K)$ depicted in Figure 7, and in this section consider the standard problem of H^∞ control theory.[1] In this problem, as in Section 3, one considers the plant $\mathscr{P}$ as given and the compensator K as a design parameter to be determined. Now one demands not only that K be proper and that the closed loop system $\Sigma(\mathscr{P}, K)$ be internally stable as in Section 3, but also that the transfer function $T_{zw} = \mathscr{F}_{\mathscr{P}}[K]$ from the reference signal w to the error signal z have infinity norm (along the imaginary axis) less than some prescribed tolerance γ:

$$\| \mathscr{F}_{\mathscr{P}}[K] \|_\infty < \gamma$$

When such Ks exist, one would then also like a parametrization of all such Ks. In this section we present a solution to this problem based on the general factorization principle Theorems 2.1.1 and 2.1.1. In Section 5 we shall implement the recipe prescribed here in terms of a state-space representation of the plant $\mathscr{P}$ to recover the elegant state-space solution of the problem recently obtained in Reference 4.

We first remark that analysis of a general performance level γ for the plant $\mathscr{P}$ is equivalent to analysis of the performance level $\gamma = 1$ for the plant

$$\mathscr{P}_\gamma = \begin{bmatrix} \gamma^{-1}\mathscr{P}_{11} & \gamma^{-1}\mathscr{P}_{12} \\ \mathscr{P}_{21} & \mathscr{P}_{22} \end{bmatrix}.$$

Indeed, it is easily seen from the form of $\mathscr{H}(\mathscr{P}, K)$ in (37) that $\Sigma(\mathscr{P}, K)$ is internally stable if and only if $\Sigma(\mathscr{P}_\gamma, K)$ is internally stable. Equally apparent is that $\mathscr{F}_{\mathscr{P}}[K] = \gamma \mathscr{F}_{\mathscr{P}_\gamma}[K]$. Thus the H^∞ problem associated with $\mathscr{P}$ and performance level γ has the same set of compensators K as solutions as the H^∞ problem associated with $\mathscr{P}_\gamma$ and performance level 1. Hence in this section we deal explicitly only with performance level $\gamma = 1$.

In parallel with the organization in Section 3, we begin with the special class of sub-all-pass plants for which the solution of the H^∞ problem is easy. Recall from the introduction that we say that a rational matrix function

$$\mathscr{P} = \begin{bmatrix} \mathscr{P}_{11} & \mathscr{P}_{12} \\ \mathscr{P}_{21} & \mathscr{P}_{22} \end{bmatrix}$$

is *sub-all-pass* if

(i) $\| \mathscr{P}(s) \| \leqslant 1$,
(ii) $\| \mathscr{P}_{11}(s) \mid \operatorname{Ker} \mathscr{P}_{21}(s) \| < 1$,

and

(iii) $\operatorname{rank}(I - \mathscr{P}(s)^* \mathscr{P}(s)) = \dim \operatorname{Ker} \mathscr{P}_{21}(s)$.

Theorem 4.1

Suppose that for all $s = \mathrm{i}w$ on the imaginary line (including at infinity) the rational matrix function

$$\mathscr{P} = \begin{bmatrix} \mathscr{P}_{11} & \mathscr{P}_{12} \\ \mathscr{P}_{21} & \mathscr{P}_{22} \end{bmatrix}$$

is analytic with $\mathscr{P}_{12}(s)$ injective and $\mathscr{P}_{21}(s)$ surjective and is sub-all-pass. If in addition $\mathscr{P}(s)$ is stable, we say that $\mathscr{P}$ is *subinner*. Then the H^∞ problem associated with $\mathscr{P}$ and tolerance level $\gamma = 1$ has a solution $K = K(s)$ if and only if $\mathscr{P}$ is subinner. In this case the compensator K is stabilizing and meets the H^∞ performance criterion (i.e. $\| \mathscr{F}_{\mathscr{P}}[K] \|_\infty < 1$) if and only if K is stable with $\| K \|_\infty < 1$.

Proof. By Proposition 1.2.3, it is clear that a necessary condition for a compensator K to solve the H^∞ problem with performance level $\gamma = 1$ is that $\| K(\mathrm{i}w) \| < 1$ for all real w (including $w = \infty$). Assuming that we have such a K, we observe that then, by a standard homotopy argument, wno det $\Delta = 0$ (where $\Delta = I - \mathscr{P}_{22}K$ and 'wno' = winding number or change of argument along the imaginary line). We now consider the matrix $\mathscr{H} = \mathscr{H}(\mathscr{P}, K)$ in (37) and recall that K stabilizes $\mathscr{P}$ if and only if $\mathscr{H}$ is stable. From the (3, 3)-entry of $\mathscr{H}$, we see that internal stability requires that Δ^{-1} be stable. But since wno det $\Delta = 0$, this forces Δ itself to be stable. But then $\mathscr{P}_{21} = \Delta \mathscr{H}_{31}$, $\mathscr{P}_{22} = \Delta \mathscr{H}_{32}$, $K = \mathscr{H}_{23}\Delta$, $\mathscr{P}_{12} = \mathscr{H}_{12} - \mathscr{H}_{13}\Delta\mathscr{H}_{32}$ and $\mathscr{P}_{11} = \mathscr{H}_{11} - \mathscr{H}_{13}\Delta\mathscr{H}_{32}$ must all be stable. Conversely, given that Δ^{-1} is stable, it is clear that $\mathscr{H}$ is stable if all of K, $\mathscr{P}_{11}$, $\mathscr{P}_{12}$, $\mathscr{P}_{21}$ and $\mathscr{P}_{22}$ are stable. Finally, since $\mathscr{P}$ is contractive on the imaginary line, stability of $\mathscr{P}$ is equivalent to $\mathscr{P}$ being contractive on the right half plane by the maximum modulus theorem. □

Combining Theorems 4.1 and 2.2.1 leads to the following result.

Theorem 4.2

Let

$$\mathscr{P} = \begin{bmatrix} \mathscr{P}_{11} & \mathscr{P}_{12} \\ \mathscr{P}_{21} & \mathscr{P}_{22} \end{bmatrix}$$

be a given plant. Assume that $\mathscr{P}_{12}(\mathrm{i}w)$ is injective and $\mathscr{P}_{21}(\mathrm{i}w)$ is surjective for all real w (including $w = \infty$). Define a rational matrix function

$$\tilde{\mathscr{P}} = \begin{bmatrix} \tilde{\mathscr{P}}_{11} & \tilde{\mathscr{P}}_{12} & \tilde{\mathscr{P}}_{13} \\ \tilde{\mathscr{P}}_{21} & \tilde{\mathscr{P}}_{22} & \tilde{\mathscr{P}}_{23} \end{bmatrix}$$

as in (20) (applied over the field $\mathscr{R}$ of scalar rational functions rather than over $\mathbb{C}$) and suppose that a factorization

$$\tilde{\mathscr{P}} = \Theta R$$

is known such that

(i) R has the block upper triangular form

$$R = \begin{bmatrix} R_{11} & R_{12} & 0 \\ R_{21} & R_{22} & 0 \\ R_{31} & R_{32} & R_{33} \end{bmatrix}$$

with

$$R_0 := \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

outer and $R_{33}(z)$ analytic and invertible on the extended imaginary line.
(ii) For all $s = iw$ on the extended imaginary line (including $w = \infty$),

$$\Theta(s)^* J \Theta(s) = j \oplus -I_{n_3}$$

where $J = I_{m_1} \oplus -I_{m_2}$, $j = I_{n_1} \oplus -I_{n_2}$. Then the H^∞ problem associated with the plant $\mathscr{P}$ and performance level $\gamma = 1$ has a solution if and only if

$$\Theta(s)^* J \Theta(s) \leqslant \begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix}$$

for all s in the right half plane. When this is the case, then K stabilizes $\mathscr{P}$ and meets the H^∞ performance criterion $\| \mathscr{F}_{\mathscr{P}}[K] \|_\infty < 1$ if and only if

$$K = \mathscr{G}_{R_0^{-1}}[H]$$

where H is a stable rational matrix function with $\| H \|_\infty < 1$ such that $H(\infty)$ is in the domain of definition of $\mathscr{G}_{R_0^{-1}(\infty)}$. The set of all closed-loop transfer functions $T_{zw} = \mathscr{F}_{\mathscr{P}}[K]$ associated with such performing compensators K is given by

$$\begin{aligned} T_{zw} &= \mathscr{G}_\Theta[H] \\ &= [\Theta_{11}H + \Theta_{12}, \Theta_{13}]\,[\Theta_{21}H + \Theta_{22}, \Theta_{23}]^{-1} \end{aligned}$$

where the free parameter H is as described above.

Remark. In Theorem 4.2 it is possible to strengthen the converse direction. Specifically, if solutions of the strictly suboptimal H^∞ problem exist, then necessarily also a factorization $\tilde{\mathscr{P}} = \Theta R$ as in the statement of the theorem also exists. One can give a proof based on ideas from Reference 7; as we do not have a short proof based on the ideas of this paper we do not go into details on this point here. In any case, the existence of the factorization $\tilde{\mathscr{P}} = \Theta R$ is generic with respect to the tolerance level γ in the control problem.

Proof. Suppose a factorization $\tilde{\mathscr{P}} = \Theta R$ is known with R satisfying (i) and Θ satisfying (ii). By Theorem 2.2.1, internal stability of $\Sigma(\mathscr{P}, K)$ is equivalent to restricted internal stability of the system $\Sigma(\mathscr{P}', [\mathscr{G}_{R_0}[K], 0])$ depicted in Figure 13, where $\mathscr{P}'$ is obtained from Θ via (23) (with $\mathscr{P}'$ in place of $\mathscr{P}$ and Θ in place of $\tilde{\mathscr{P}}$). This in turn is equivalent to internal stability of a system $\Sigma(\mathscr{P}', \mathscr{G}_{R_0}[K])$ where $\mathscr{P}'$ is obtained from Θ via formula (24) (with $\mathscr{P}'$ in place of $\mathscr{P}$ and Θ in place of $\tilde{\mathscr{P}}$). Now the hypothesis (ii) on Θ combined with Remark 2 after

Proposition 1.2.3 guarantees that $\mathscr{P}'$ meets the hypotheses of Theorem 4.1. Hence the H^∞ problem for $\mathscr{P}'$ has solutions if and only if $\| \mathscr{P}'(s) \| \leqslant 1$ for all s in the right half plane. We now show that $\| \mathscr{P}'(s) \| \leqslant 1$ for all s in the right half plane if and only if

$$\Theta(s)^* J\Theta(s) \leqslant \begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix}.$$

Recall the connection between $\mathscr{P}'$ and Θ: $\mathscr{P}'$ has an augmentation

$$\mathscr{P}' = \begin{bmatrix} \mathscr{P}'_{11} & \mathscr{P}'_{12} \\ \mathscr{P}'_{21} & \mathscr{P}'_{22} \\ \mathscr{P}'^{0}_{21} & \mathscr{P}'^{0}_{22} \end{bmatrix}$$

so that the system of equations

$$\begin{bmatrix} \mathscr{P}'_{11} & \mathscr{P}'_{12} \\ \mathscr{P}'_{21} & \mathscr{P}'_{22} \\ \mathscr{P}'^{0}_{21} & \mathscr{P}'^{0}_{22} \end{bmatrix} \begin{bmatrix} w \\ u \end{bmatrix} = \begin{bmatrix} z \\ y \\ y^0 \end{bmatrix} \tag{59}$$

is equivalent to the system

$$\begin{bmatrix} \Theta_{11} & \Theta_{12} & \Theta_{13} \\ \Theta_{21} & \Theta_{22} & \Theta_{23} \end{bmatrix} \begin{bmatrix} u \\ y \\ y^0 \end{bmatrix} = \begin{bmatrix} z \\ w \end{bmatrix} \tag{60}$$

Now consider $\mathscr{P}'$ and Θ evaluated at some fixed point s. Then $\| \mathscr{P}(s) \| \leqslant 1$ means that

$$\| z \|^2 + \| y \|^2 \leqslant \| w \|^2 + \| u \|^2 \tag{61}$$

whenever w, u, z, y satisfy

$$\begin{bmatrix} \mathscr{P}''_{11} & \mathscr{P}'_{12} \\ \mathscr{P}'_{21} & \mathscr{P}'_{22} \end{bmatrix} \begin{bmatrix} w \\ u \end{bmatrix} = \begin{bmatrix} z \\ y \end{bmatrix}$$

Then we see that (61) holds also whenever w, u, z, y, y^0 satisfy (59) (irrespective of the value of y^0). Rearrange (61) in the form

$$\| z \|^2 - \| w \|^2 \leqslant \| u \|^2 - \| y \|^2 \tag{62}$$

Using that (59) and (60) are equivalent, we see from (62) that $\Theta(s)$ is $\left(\begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix}, J\right)$-contractive as asserted. The converse, that $\mathscr{P}(s)$ is contractive if $\Theta(s)$ is $\left(\begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix}, j\right)$-contractive, can be established in a similar vein. We conclude that the H^∞ problem for $\mathscr{P}'$ has solutions if and only if Θ is $\left(\begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix}, J\right)$-contractive in the right half plane. Assume now that this condition holds.

By Theorem 2.2.1 we know that K stabilizes $\mathscr{P}$ if and only if $H = \mathscr{G}_{R_0}[K]$ stabilizes $\mathscr{P}'$. By Theorem 3.1, H stabilizes $\mathscr{P}'$ and meets the performance criterion $\| T_{zw} \| = \| \mathscr{F}_{\mathscr{P}'}\ [H] \| < 1$ if and only if H is stable with $\| H \|_\infty < 1$. Back-solving $H = \mathscr{G}_{r_0}[K]$ for K gives $K = \mathscr{G}_{R_0^{-1}}[H]$; the constraint that K be proper limits H to those stable contractions such that $H(\infty)$ is in the domain of definition of $\mathscr{G}_{R_0^{-1}(\infty)}$. Finally note that

$$\begin{aligned}\mathscr{F}_{\mathscr{P}}[K] &= \mathscr{G}_{\tilde{\mathscr{P}}}[K] \\ &= \mathscr{G}_{\Theta R}[K] \\ &= \mathscr{G}_{\Theta}[\mathscr{G}_{R_0}[K]] \\ &= \mathscr{G}_{\Theta}[H]\end{aligned}$$

and hence $\mathscr{G}_{\Theta}$ acting on stable contractions provides a parametrization of the associated performances. This completes the proof of Theorem 4.2. □

To make Theorem 4.2 useful we need a systematic way of computing a factorization $\tilde{\mathscr{P}} = \Theta R$ as in Theorem 4.2. Note that the parametrization of all the compensators solving the H^∞ problem requires computation only of R_0 and not all of R. Also, by Lemma 1.1.2, to parametrize the set of performances T_{zw}, it is sufficient to compute any $\tilde{\Theta}$ of the form

$$\tilde{\Theta} = \Theta \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{bmatrix}$$

where γ_1, γ_2, and γ_3 are any rational matrix functions (with γ_3 invertible) rather than Θ itself if it is more convenient, since then $\mathscr{G}_{\Theta} = \mathscr{G}_{\tilde{\Theta}}$ as linear fractional maps. As we shall see it is more convenient to compute a certain $\tilde{\Theta}$ rather than Θ itself. These observations lead to the following result which will be the basis for the state space calculations in the next section.

Theorem 4.3

Let $\mathscr{P}$ and $\tilde{\mathscr{P}}$ be as in Theorem 4.2, and say $\tilde{\mathscr{P}}_{ij}$ has size $m_i \times n_j$ for $i = 1, 2$ and $j = 1, 2, 3$. Define a block 3×3 matrix function $W(z)$ on the extended imaginary line by

$$W(z) = [W_{ij}(z)]_{1 \leqslant i, j \leqslant 3}$$

where

$$W_{ij}(z) = \tilde{\mathscr{P}}_{1i}(z)^* \tilde{\mathscr{P}}_{1j}(z) - \tilde{\mathscr{P}}_{2i}(z)^* \tilde{\mathscr{P}}_{2j}(z)$$

has size $n_i \times n_j$. In addition, define rational matrix functions

$$W_0 = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}, \qquad W_{30} = [W_{31} \;\; W_{32}]$$

so that we may also consider the block 2×2 decomposition of W,

$$W = \begin{bmatrix} W_0 & W_{30}^* \\ W_{30} & W_{33} \end{bmatrix}$$

and set $J = I_{m_1} \oplus -I_{m_2}$, $j = I_{n_1} \oplus -I_{n_2}$. Then the H^∞ problem associated with the plant $\mathscr{P}$ and performance level $\gamma = 1$ has a solution if and only if

(i) $W_{33}(z) = \tilde{\mathscr{P}}_{13}(z)^* \tilde{\mathscr{P}}_{13}(z) - \tilde{\mathscr{P}}_{23}(z)^* \tilde{\mathscr{P}}_{23}(z)$ is negative definite for z on the extended imaginary line,

(ii) there is an $(n_1 + n_2) \times (n_1 + n_2)$ outer rational matrix function

$$R_0(z) = \begin{bmatrix} R_{11}(z) & R_{12}(z) \\ R_{21}(z) & R_{22}(z) \end{bmatrix}$$

such that

$$V(z) = R_0(z)^* jR_0(z)$$

for z on the extended imaginary line, where $V = W_0 - W_{30}^* W_{33}^{-1} W_{30}$ is the Schur complement of W with respect to W_{33}.

(iii) The rational matrix function

$$\tilde{\Theta} = \tilde{\mathscr{P}} \begin{bmatrix} R_0^{-1} & 0 \\ 0 & I \end{bmatrix}$$

where R_0 is as in (ii), satisfies

$$\tilde{\Theta}(s)^* J \tilde{\Theta}(s) \leqslant \begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix}$$

at all points of analyticity s in the right half plane.

Moreover, when conditions (i), (ii), and (iii) hold, the set of all compensators K satisfying the H^∞ problem is given by

$$K = \mathscr{G}_{R_0^{-1}}[H]$$

and the set of all associated performances T_{zw} is given by

$$T_{zw} = \mathscr{G}_{\tilde{\Theta}}[H] = [\tilde{\Theta}_{11} H + \tilde{\Theta}_{12}, \tilde{\Theta}_{13}]\,[\tilde{\Theta}_{21} H + \tilde{\Theta}_{22}, \tilde{\Theta}_{23}]^{-1}$$

where

$$\tilde{\Theta} = \begin{bmatrix} \tilde{\Theta}_{11} & \tilde{\Theta}_{12} & \tilde{\Theta}_{13} \\ \tilde{\Theta}_{21} & \tilde{\Theta}_{22} & \tilde{\Theta}_{23} \end{bmatrix}$$

is as in (iii) and H is any stable rational matrix function with $\| H \|_\infty < 1$ such that $H(\infty)$ is in the domain of definition of $\mathscr{G}_{R_0^{-1}(\infty)}$.

Proof. By Theorem 4.2 we need to find a factorization

$$\tilde{\mathscr{P}} = \Theta R$$

where R and Θ satisfy properties (i) and (ii) in Theorem 4.2. If Θ and R provide such a factorization then

$$\begin{aligned} W &= \tilde{\mathscr{P}}^* J \tilde{\mathscr{P}} \\ &= (\Theta R)^* J (\Theta R) \\ &= R^*(j \oplus -I_{n_3}) R \end{aligned}$$

by the $(j \oplus -I_{n_3}, J)$-isometric property (ii) of Θ. Conversely, if R satisfies (i) in Theorem 4.2 and provides a factorization of W of the form

$$W = R^*(j \oplus -I_{n_3}) R \tag{63}$$

then $\tilde{\mathscr{P}} = \Theta R$ with $\tilde{\Theta} = \tilde{\mathscr{P}} R^{-1}$ is the desired factorization satisfying (i) and (ii) in Theorem 4.2. Therefore a first step is to analyse the factorization problem (63) where W is given and the unknown R is to have the form (i) in Theorem 4.2.

First suppose that

$$R=\begin{bmatrix} R_{11} & R_{12} & 0 \\ R_{21} & R_{22} & 0 \\ R_{31} & R_{32} & R_{33} \end{bmatrix}$$

provides such a factorization. Thus

$$R_0=\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

is an $(n_1+n_2)\times(n_1+n_2)$ outer matrix function, R_{33} is an $n_3\times n_3$ rational matrix function with invertible values on the imaginary line, and if we set $R_{30}=[R_{31}\ \ R_{32}]$, we may write R in the block 2×2 form

$$R=\begin{bmatrix} R_0 & 0 \\ R_{30} & R_{33} \end{bmatrix}.$$

Then the factorization (63) in more detail is the same as

$$\begin{bmatrix} W_0 & W_{30}^* \\ W_{30} & W_{33} \end{bmatrix}=\begin{bmatrix} R_0^*jR_0-R_{30}^*R_{30} & -R_{30}^*R_{33} \\ -R_{33}^*R_{30} & -R_{33}^*R_{33} \end{bmatrix}$$

Then necessarily $W_{33}=-R_{33}^*R_{33}$ is negative definite on the extended imaginary line and the Schur complement V of W_{33} is given by

$$\begin{aligned} V &:= W_0-W_{30}^*W^{-1}W_{30} \\ &= [R_0^*jR_0-R_{30}^*R_{30}]+(R_{30}^*R_{33})(R_{33}^*R_{33})^{-1}(R_{33}^*R_{30})=R_0^*jR_0 \end{aligned}$$

Thus R_0 arises via a conventional j-spectral factorization of V. Then R_{30} is determined from R_{33} and W_{30} via $R_{30}=-R_{33}^{*-1}W_{30}$. Conversely, suppose W_{33} is negative definite on the extended line and and that V has a j-spectral factorization. Then we can find an $n_3\times n_3$ rational matrix function R_{33} invertible on the extended imaginary line and an $(n_1+n_2)\times(n_1+n_2)$ outer matrix function R_0 such that

$$W_{33}=-R_{33}^*R_{33}, \qquad V=R_0^*jR_0$$

Then

$$R=\begin{bmatrix} R_0 & 0 \\ R_{33}^{*-1}W_{30} & R_{33} \end{bmatrix}$$

provides the desired factorization $W=R^*(j\oplus -I_{n_3})R$ of W.

Let us now assume that $W_{33}=-R_{33}^*R_{33}$ is negative definite and $V=R_0^*jR_0$ has a j-spectral factorization. Set

$$R_{30}=-R_{33}^{*-1}W_{30}R=\begin{bmatrix} R_0 & 0 \\ R_{30} & R_{33} \end{bmatrix}$$

and $\Theta=\tilde{\mathscr{P}}R^{-1}$. Then $\tilde{\mathscr{P}}=\Theta R$ is a factorization with factors Θ and R having the properties (i) and (ii) in Theorem 4.2. Hence by Theorem 4.2 the H^∞ problem (with performance level $\gamma=1$) has a solution K if and only if

$$\Theta \text{ is } \left(\begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix}, J\right)\text{-contractive}$$

in the right half plane, in which case $\mathscr{G}_{R_0^{-1}}$ parametrizes the compensators K which solve the problem and $\mathscr{G}_\Theta$ parametrizes the associated closed loop transfer functions $T_{zw} = \mathscr{F}_{\mathscr{P}}[K]$.

We need to argue that $\Theta = \tilde{\mathscr{P}}R^{-1}$ can be replaced by

$$\tilde{\Theta} = \tilde{\mathscr{P}}\begin{bmatrix} R_0^{-1} & 0 \\ 0 & I \end{bmatrix}$$

in the above analysis. To do this, note that

$$\tilde{\Theta} = \Theta R\begin{bmatrix} R_0^{-1} & 0 \\ 0 & I \end{bmatrix}$$

$$= \Theta\begin{bmatrix} I & 0 \\ R_{30}R_0^{-1} & R_{33} \end{bmatrix}$$

Hence by Lemma 1.1.2, $\mathscr{G}_\Theta$ is the same linear fractional map as $\mathscr{G}_{\tilde{\Theta}}$. Also the identity

$$\begin{bmatrix} I & R_0^{-1*}R_{30}^* \\ 0 & R_{33}^* \end{bmatrix}\begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} I & 0 \\ R_{30}R_0^{-1} & R_{33} \end{bmatrix} = \begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix}$$

shows that Θ is $\left(\begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix}, J\right)$-contractive on the right half plane if and only if $\tilde{\Theta}$ is. This completes the proof of Theorem 4.2. □

Remark. The idea for much of the analysis of this section comes from Reference 7. There the H^∞ problem was assumed to be in the model-matching form and proofs were couched in the less elementary language of projective geometry of Krein spaces using the techniques of Reference 11. Points missed there and contributing to the eventual obscurity of the paper were: (1) the outer factor R_0^{-1} can be used to parametrize the compensators K which solve the H^∞ problem, and (2) the matrix function Θ appearing in Theorem 4.3, which is not $(j \oplus -I, J)$ isometric on the real line, can be used as a parametrizer of the performances T_{zw} equally effectively as the $(j \oplus -I, J)$-isometry Θ appearing in Theorem 4.2. This latter point accounts for the appearance of an additional unnecessary Riccati equation in the state-space formulas derived in Reference 7.

5. GENERALIZED J-INNER–OUTER FACTORIZATION: STATE-SPACE COMPUTATIONS

Theorem 4.3 reduced the solution of the H^∞ control problem to a certain type of factorization problem. In this section we analyse this factorization problem in state-space terms; specialization to the factorization problem arising from the H^∞ control context enables us to recover the state-space formulas and results of Reference 4 (see also References 8, 10, and 6), for the H^∞ problem; we do this in Section 6.

In the following we consider a rational 1×2 block matrix function $G(z) = [G_\alpha(z) \;\; G_\beta(z)]$ where G_α has size $p \times q_\alpha$ and G_β has size $p \times q_\beta$. We assume that we are also given signature matrices J and j of respective sizes $p \times p$ and $q_\alpha \times q_\alpha$. The problem we wish to study is that of factoring G as

$$G = \Theta R$$

where $\Theta = [\Theta_\alpha \ \ \Theta_\beta]$ has size $p \times (q_\alpha + q_\beta)$ and

$$R = \begin{bmatrix} R_\alpha & 0 \\ R_{\beta\alpha} & R_\beta \end{bmatrix}$$

has size $(q_\alpha + q_\beta) \times (q_\alpha + q_\beta)$ such that

$$\Theta(s)^* J(\Theta(s) = \begin{bmatrix} j & 0 \\ 0 & -I \end{bmatrix} \quad \text{for Re } s = 0 \tag{64}$$

$$\Theta(s)^* J(\Theta(s) \leqslant \begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix} \quad \text{for Re } s > 0 \tag{65}$$

and

$$R = \begin{bmatrix} R_\alpha & 0 \\ R_{\beta\alpha} & R_\beta \end{bmatrix} \tag{66}$$

where R_α is an outer rational matrix function and $R_{\beta\alpha}$, R_β and R_β^{-1} are rational matrix functions with no poles on Re $s = 0$ (including infinity). By the argument in the proof of Theorem 4.3 we see that, once (64) and (66) are achieved, then (65) can be replaced by

$$\begin{bmatrix} \Theta_\alpha(s)^* \\ G_\beta(s)^* \end{bmatrix} J[\Theta_\alpha(s) \ \ G_\beta(s)] \leqslant \begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix} \quad \text{for Re } s > 0 \tag{67}$$

We assume that we have a state-space realization for G

$$\begin{aligned} G(s) &= D + C(s - A)^{-1}B \\ &= [D_\alpha \ \ D_\beta] + C(s - A)^{-1}[B_\alpha \ \ B_\beta] \end{aligned} \tag{68}$$

and wish to find an existence criterion for the factorization and then formulas for the factors (in particular R_α and Θ_α) in terms of the matrices A, B_α, B_β, C, D_α, D_β. In the sequel we also assume

$D_\alpha^* J D_\alpha$ is invertible with a factorization of the form $D_\alpha^* J D_\alpha$
$= d_\alpha^* j d_\alpha$ for an invertible $q_\alpha \times q_\alpha$ matrix d_α (69)

$D_\beta^* J D_\beta$ is invertible with a factorization of the form $D_\beta^* J D_\beta$
$= - d_\beta^* d_\beta$ for an invertible $q_\beta \times q_\beta$ matrix d_β (70)

$$D_\beta^* J D_\alpha = 0, \qquad D_\beta^* J C = 0 \tag{71}$$

To condense notation we write

$$E_\alpha = (D_\alpha^* J D_\alpha)^{-1} = d_\alpha^{-1} j d_\alpha^{*-1}, \qquad E_\beta = (D_\beta^* J D_\beta)^{-1} = -d_\beta^{-1} d_\beta^{*-1} \qquad E = \begin{bmatrix} E_\alpha & 0 \\ 0 & E_\beta \end{bmatrix} \tag{72}$$

$$\Delta_\alpha = J - J D_\alpha E_\alpha D_\alpha^* J \tag{73}$$

$$A^\times = A - B_\alpha E_\alpha D_\alpha^* J C \tag{74}$$

$$K = \begin{bmatrix} A & B_\beta E_\beta B_\beta^* \\ C^* J C & -A^* \end{bmatrix} \tag{75}$$

$$K^\times = \begin{bmatrix} A^\times & BEB^* \\ C^* \Delta_\alpha C & -(A^\times)^* \end{bmatrix} \tag{76}$$

For M any square matrix, $\rho_\infty(M)$ denotes the spectral radius of M and $\sigma(M)$ denotes the spectrum of M.

Theorem 5.1

Let $G(s) = [D_\alpha \;\; D_\beta] + C(s-A)^{-1}[B_\alpha \;\; B_\beta]$ be a rational matrix function as above satisfying (69)–(71). Assume that (A, B_β) is controllable, that the matrices A, K and $K^\times$ have no eigenvalues on the imaginary line, and that the Riccati equations

$$XA^\times + (A^\times)^*X - XBEB^*X + C^*\Delta_\alpha C = 0 \tag{77}$$

and

$$YA^* + AY + YC^*JCY - B_\beta E_\beta B_\beta^* = 0 \tag{78}$$

have Hermitian solutions $X = X^*$ and $Y = Y^*$ meeting the stability side conditions

$$\sigma(A^* - BEB^*X) \subset \{\mathrm{Re}\, s < 0\} \tag{79}$$

$$\sigma(A + YC^*JC) \subset \{\mathrm{Re}\, s < 0\} \tag{80}$$

Then G has a factorization $G = \Theta R$ with Θ satisfying (64) and (65) and R satisfying (66) if

$$X \geqslant 0, \qquad Y \geqslant 0, \qquad \rho_\infty(XY) < 1 \tag{81}$$

In this case, $Z = (I - XY)^{-1}$ exists and if $\Theta = [\Theta_\alpha \;\; \Theta_\beta]$ and $R = \begin{bmatrix} R_\alpha & 0 \\ R_{\beta\alpha} & R_\beta \end{bmatrix}$, one may take

$$R_\alpha(s) = d_\alpha + d_\alpha E_\alpha(D_\alpha^*JC + B_\alpha^*X)Z^*(s - (A + YC^*JC))^{-1}(B_\alpha + YC^*JD_\alpha) \tag{82}$$

$$R_\alpha(s)^{-1} = d_\alpha^{-1} - E_\alpha(D_\alpha^*JC + B_\alpha^*X)(s - (A^\times - BEB^*X))^{-1}Z^*(B_\alpha + YC^*JD_\alpha)d_\alpha^{-1} \tag{83}$$

$$\begin{aligned}\tilde{\Theta}_\alpha(s) = G_\alpha(s)R_\alpha(s)^{-1} &= D_\alpha d_\alpha^{-1} + [J\Delta_\alpha C - D_\alpha E_\alpha B_\alpha^* X \;\; C] \\ &\times \left\{ s - \begin{bmatrix} A^\times - BEB^*X & 0 \\ B_\beta E_\beta B_\beta^* X & A \end{bmatrix} \right\}^{-1} \begin{bmatrix} Z^* & Z^*Y \\ -YZX & -YZ \end{bmatrix} \begin{bmatrix} B_\alpha \\ C^*JD_\alpha \end{bmatrix} d_\alpha^{-1}\end{aligned} \tag{84}$$

Conversely, if G has a factorization $G = \Theta R$ satisfying (64)–(66) and the pair $(\hat{A}, \hat{B})$ is controllable, where

$$\hat{A} = \begin{bmatrix} A^\times - BEB^*X & 0 \\ B_\beta E_\beta B_\beta^* X & A \end{bmatrix} \tag{85}$$

and

$$\hat{B} = \begin{bmatrix} Z^*(B_\alpha + YC^*JD_\alpha) \\ -YZ(XB_\alpha + C^*JD_\alpha) \end{bmatrix} d_\alpha^{-1} \tag{86}$$

then necessarily (81) holds.

Remark. The point of this theorem is to exhibit the connection between the

$$\left(\begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix}, J \right)\text{-contractive}$$

property for Θ holding on $\mathrm{Re}\, s > 0$ and the set of conditions (81) holding on the stabilizing solutions X, Y of the Riccati equations (77) and (78). We expect that a much stronger result holds; namely, *under the assumptions that* (C, A) *is detectable and* (A, B) *is stabilizable, then the factorization* $G = \Theta R$ *exists satisfying* (64)–(66) *if and only if stabilizing solutions* X, Y *of* (77)–(80) *exists which also satisfy* (81). By using the recent results on the H^∞ control problem 4, 8, 10, 6, and the connection between H^∞ control and factorization given by Theorem 4.2 (see Section 6), we expect that this result holds in complete generality, but of course a direct proof would be desirable.

The proof of Theorem 5.1 requires the construction of a j-spectral factorization for a given square (say $N \times N$) rational matrix function $V(s)$. Here j is an $N \times N$ signature matrix ($j = j^* = j^{-1}$) and V is analytic with invertible Hermitian values on the extended real line (so $V(s)$ $V(s) = V^{\sim}(s)$ where $V^{\sim}(s) = V(-\bar{s})^*$). We say that V admits a j-spectral factorization (with respect to the right half plane) if there is a square rational matrix function $V_-(s)$ with no poles and no zeros in Re $s \geqslant 0$ (including at infinity) such that

$$V(s) = V_-^{\sim}(s) j V_-(s)$$

Note that we do not need to assume that the realization $V(s) = \mathbf{D} + \mathbf{C}(s - \mathbf{A})^{-1}\mathbf{B}$ is minimal, as is done in Reference 23; this was obtained in Reference 24 as a corollary of an elaborate machinery set up to handle non-canonical factorization. When $\mathbf{A}$ is an $n \times n$ matrix, we define the *modal subspace* $X_+(\mathbf{A})$ relative to the right half plane to be the span of all the eigenvectors and generalized eigenvectors associated with eigenvalues in the left half plane. Similarly the model subspace $X_-(\mathbf{A})$ relative to the left half plane is the analogous object with the left half plane in place of the right half plane. In general, if $\mathcal{M}$ and $\mathcal{N}$ are two subspaces of $\mathbb{C}^n$ we write

$$\mathbb{C}^n = \mathcal{M} \dot{+} \mathcal{N}$$

if the pair $(\mathcal{M}, \mathcal{N})$ form an (internal) direct sum decomposition of $\mathbb{C}^n$.

Theorem 5.2

Let $V(s) = V^{\sim}(s) = \mathbf{D} + \mathbf{C}(s - \mathbf{A})^{-1}\mathbf{B}$ be an $N \times N$ rational matrix function with no poles and no zeros on the imaginary line, where $\mathbf{A}$ is an $n \times n$ matrix with no eigenvalues on the imaginary line and the Hermitian matrix $\mathbf{D}$ factors as

$$\mathbf{D} = d^* j d$$

for an $N \times N$ signature matrix j. Put $A^\times = \mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}$. Then $V(s)$ admits a j-spectral factorization $V = V_-^{\sim} j V_-$ if and only if

$$\mathbb{C}^n = X_-(\mathbf{A}^\times) \dot{+} X_+(\mathbf{A})$$

Furthermore, in this case a j-spectral factorization of V is constructed as follows:

(1) Choose bases $x_1, \ldots, x_r$ of $X_-(\mathbf{A}^\times)$ and $x_{r+1}, \ldots, x_n$ of $X_+(\mathbf{A})$ and put

$$T = [x_1 \ldots x_r \,\vdots\, x_{r+1} \ldots x_n] \tag{87}$$

Then T is non-singular.

(2) Make the following partitionings

$$T^{-1}\mathbf{A}T = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}, \qquad T^{-1}\mathbf{B} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \qquad \mathbf{C}T = [C_1 \;\; C_2] \tag{88}$$

and

$$T^{-1}\mathbf{A}^{\times}T = \begin{bmatrix} A_{11}^{\times} & A_{12}^{\times} \\ 0 & A_{22}^{\times} \end{bmatrix}$$

according to the partitioning of T in (84).

(3) Put

$$V_{-}(s) = d + d\mathbf{D}^{-1}C_1(s - A_{11})^{-1}B_1 \tag{89}$$

Then $V(s) = V_{-}^{\sim}(s)jV_{-}(s)$ is a j-spectral factorization and $V_{-}(s)^{-1}$ is given by

$$V_{-}(s)^{-1} = d^{-1} - \mathbf{D}^{-1}C_1(s - A_{11}^{\times})^{-1}B_1 d^{-1} \tag{90}$$

In addition $V_{-}(s)$ is real-rational if $V(s)$ is real-rational and d is real.

Before proving Theorem 5.1 we analyse in state space terms the existence of a

$$\left(\begin{bmatrix} j & 0 \\ 0 & -I \end{bmatrix}, J\right)\text{-isometric-(restricted outer) factorization } G = \Theta R;$$

i.e. we demand that Θ and R satisfy only (64) and (66). We have the following result.

Theorem 5.3

Let $G(s) = [D_\alpha \;\; D_\beta] + C(s - A)^{-1}[B_\alpha \;\; B_\beta]$ be a rational matrix function such that (69)–(72) are satisfied. Assume that the matrices A, K and $K^{\times}$ (where K and $K^{\times}$ are defined in (75) and (76)) have eigenvalues on the imaginary line and that X and Y are Hermitian matrices satisfying (77)–(80). Then G admists a $\left(\left(\begin{bmatrix} j & 0 \\ 0 & I \end{bmatrix} - J\right)\text{-isometric}\right)$-(restricted outer) factorization

$$G = \Theta R = [\Theta_\alpha \;\; \Theta_\beta]\begin{bmatrix} R_\alpha & 0 \\ R_{\beta\alpha} & R_\beta \end{bmatrix}$$

if and only if $I - XY$ is invertible. In this case, R_α and R_α^{-1} and $\tilde{\Theta}_\alpha = G_\alpha R_\alpha^{-1}$ can be taken to be given by (82)–(84).

Proof. We analyse the factorization $G = \Theta R$ ((64) and (66)) by following the strategy given by Theorem 4.3. Thus, starting with the state-space realization (68) for G, we need to compute state-space realizations for

$$W(s) = \begin{bmatrix} W_{\alpha\alpha}(s) & W_{\alpha\beta}(s) \\ W_{\beta\alpha}(s) & W_{\beta\beta}(s) \end{bmatrix} = \begin{bmatrix} G_\alpha^{\sim}(s) \\ G_\beta^{\sim}(s) \end{bmatrix} J[G_\alpha(s) \quad G_\beta(s)] \tag{91}$$

and for $V(s)$ and $V^{-1}(s)$, where

$$V(s) = W_{\alpha\alpha}(s) - W_{\alpha\beta}(s)[W_{\beta\beta}(s)]^{-1}W_{\beta\alpha}(s) \tag{92}$$

In general we shall indicate that a rational matrix function $g(s)$ has the realization $g(s) = d + c(s - a)^{-1}b$ by

$$g(s) \sim \left[\begin{array}{c:c} a & b \\ \hdashline c & d \end{array}\right]$$

From

$$G(s) \sim \left[\begin{array}{c:cc} A & B_\alpha & B_\beta \\ \hdashline C & D_\alpha & D_\beta \end{array}\right]$$

we get

$$G^{\sim}(s) \sim \left[\begin{array}{c:c} -A^* & C^* \\ \hdashline -B_\alpha^* & D_\alpha^* \\ -B_\beta^* & D_\beta^* \end{array}\right]$$

Multiplication of realizations (see, for example, Reference 23 or 19) then gives

$$W(s) = G^{\sim}(s)JG(s) \sim \left[\begin{array}{cc:cc} A & 0 & B_\alpha & B_\beta \\ C^*JC & -A^* & C^*JD_\alpha & 0 \\ \hdashline D_\alpha^*JC & -B_\alpha^* & D_\alpha^*JD_\alpha & 0 \\ 0 & -B_\beta^* & 0 & D_\beta^*JD_\beta \end{array}\right] \tag{93}$$

where we used (71). In particular

$$W_{\beta\beta}(s) = G_\beta^{\sim}(s)JG_\beta(s) \sim \left[\begin{array}{cc:c} A & 0 & B_\beta \\ C^*JC & -A^* & 0 \\ \hdashline 0 & -B_\beta^* & D_\beta^*JD_\beta \end{array}\right] \tag{94}$$

From the general rule

$$w^{-1} \sim \left[\begin{array}{c:c} a - bd^{-1}c & bd^{-1} \\ \hdashline -d^{-1}c & d \end{array}\right] \quad \text{if } w \sim \left[\begin{array}{c:c} a & b \\ \hdashline c & d \end{array}\right] \tag{95}$$

whenever d is invertible[23] we deduce that

$$[W_{\beta\beta}(s)]^{-1} \sim \left[\begin{array}{cc:c} A & B_\beta E_\beta B_\beta & B_\beta E_\beta \\ C^*JC & -A^* & 0 \\ \hdashline 0 & E_\beta B_\beta^* & E_\beta \end{array}\right]$$

where

$$\begin{bmatrix} A & B_\beta E_\beta B_\beta \\ C^*JC & -A^* \end{bmatrix} = K$$

has no spectrum on the imaginary line by hypothesis. Since $[W_{\beta\beta}(s)]^{-1}$ has Hermitian values on the imaginary line and $[W_{\beta\beta}(\infty)]^{-1} = E_\beta = -d_\beta^{-1}d_\beta^{*-1} < 0$ (see (70)) we conclude that $W_{\beta\beta}(s) < 0$ on the whole extended real line. To show that G has a

$$\left(\left(\begin{bmatrix} j & 0 \\ 0 & -I \end{bmatrix}, J\right)\text{-isometric}\right)\text{-(restricted outer) factorization,}$$

by an argument as in the proof of Theorem 4.3 it suffices to show that

$$V := W_{\alpha\alpha} - W_{\alpha\beta}(W_{\beta\beta})^{-1}W_{\beta\alpha}$$

admits a j-spectral factorization. By a Schur complement argument (see, for example, Reference 19), $V=(w_{\alpha\alpha})^{-1}$ where

$$W^{-1}=\begin{bmatrix} w_{\alpha\alpha} & w_{\alpha\beta} \\ w_{\beta\alpha} & w_{\beta\beta} \end{bmatrix}$$

Using again the general principle (95) in combination with (93) we get

$$W^{-1}(s)=\begin{bmatrix} w_{\alpha\alpha}(s) & w_{\alpha\beta}(s) \\ w_{\beta\alpha}(s) & w_{\beta\beta}(s) \end{bmatrix} \sim \left[\begin{array}{cc|cc} A^{\times} & BEB^{*} & B_{\alpha}E_{\alpha} & B_{\beta}E_{\beta} \\ C^{*}\Delta_{\alpha}C & -(A^{\times})^{*} & C^{*}JD_{\alpha}E_{\alpha} & 0 \\ \hline -E_{\alpha}D_{\alpha}^{*}JC & E_{\alpha}B_{\alpha}^{*} & E_{\alpha} & 0 \\ 0 & -E_{\beta}B_{\beta}^{*} & 0 & E_{\beta} \end{array}\right]$$

where we use the notation (72)–(74). Hence the (1, 1)-block $w_{\alpha\alpha}(s)$ is given by

$$V^{-1}(s)=w_{\alpha\alpha}(s)\sim \left[\begin{array}{cc|c} A^{*} & BEB^{*} & B_{\alpha}E_{\alpha} \\ C^{*}\Delta_{\alpha}C & -(A^{\times})^{*} & C^{*}JD_{\alpha}E_{\alpha} \\ \hline -E_{\alpha}D_{\alpha}^{*}JC & E_{\alpha}B_{\alpha}^{*} & E_{\alpha} \end{array}\right] \tag{96}$$

Thus we get a realization of V by using (94) to compute a realization of $(w_{\alpha\alpha})^{-1}$ from (96); the result is

$$V(s)\sim \left[\begin{array}{cc|c} A & B_{\beta}E_{\beta}B_{\beta}^{*} & B_{\alpha} \\ C^{*}JC & -A^{*} & C^{*}JD_{\alpha} \\ \hline D_{\alpha}^{*}JC & -B_{\alpha}^{*} & D_{\alpha}^{*}JD_{\alpha} \end{array}\right] \tag{97}$$

Thus

$$V\sim \left[\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right]$$

where

$$\mathbf{A}=K \qquad \text{(see 75)}$$

$$\mathbf{B}=\begin{bmatrix} B_{\alpha} \\ C^{*}JD \end{bmatrix}$$

$$\mathbf{C}=[D_{\alpha}^{*}JC \quad -B_{\alpha}^{*}]$$

$$\mathbf{D}=D_{\alpha}^{*}JD_{\alpha}=d_{\alpha}^{*}jd_{\alpha} \qquad \text{(see (69))}$$

and

$$\mathbf{A}^{\times}:=\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}=K^{\times} \qquad \text{(see (76)).}$$

By assumption K and $K^{\times}$ have no imaginary axis eigenvalues. Then by Theorem 5.2, we conclude that V has a j-spectral factorization

$$V(s)=R_{\alpha}^{\sim}(s)jR_{\alpha}(s)$$

if and only if

$$\mathbb{C}^{2n}=X_{-}(K^{\times})+X_{+}(K) \tag{98}$$

Now we assume that X and Y are solutions of (77)–(80). Note that (77) can be written as

$$\begin{bmatrix} A^{\times} & BEB^* \\ C^*\Delta_\alpha C & -(A^{\times})^* \end{bmatrix}\begin{bmatrix} I \\ -X \end{bmatrix} = \begin{bmatrix} I \\ -X \end{bmatrix}(A^{\times} - BEB^*X)$$

Hence $\operatorname{Im}\begin{bmatrix} I \\ -X \end{bmatrix}$ is invariant for $K^{\times}$ and $K^{\times} \mid \operatorname{Im}\begin{bmatrix} I \\ -X \end{bmatrix}$ is similar to $A^{\times} - BEB^*X$ where $\sigma(A^* - BEB^*X)$ is stable by (79). Moreover, since K is a Hamiltonian matrix

$$\left(eK = -K^*e \text{ where } e = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}\right)$$

the eigenvector–eigenvalue structure of K is symmetric about the imaginary line. This implies that $X = X^*$ and

$$X_-(K^{\times}) = \operatorname{Im}\begin{bmatrix} I \\ -X \end{bmatrix}$$

by dimension count. Similarly, if Y satisfies (78) and (80) then necessarily

$$\operatorname{Im}\begin{bmatrix} -Y \\ I \end{bmatrix}$$

is invariant for K,

$$K \mid \operatorname{Im}\begin{bmatrix} -Y \\ I \end{bmatrix}$$

is similar to $-A^* - C^*JCY$ which has all eigenvalues in the right half plane by (80) and

$$X_+(K) = \operatorname{Im}\begin{bmatrix} -Y \\ I \end{bmatrix}$$

by dimension count. Now condition (80) translates to the invertibility of the matrix

$$\begin{bmatrix} I & -Y \\ -X & I \end{bmatrix}$$

or equivalently to

$$I - XY \text{ is invertible} \tag{99}$$

Let us now assume that (98) (and hence (99)) holds. Set

$$Z = (I - XY)^{-1}$$

We may compute the j-spectral factor $R_\alpha(s)$ of $V(s)$ by setting

$$T = \begin{bmatrix} I & -Y \\ -X & I \end{bmatrix}$$

$$T^{-1} = \begin{bmatrix} Z^* & Z^*Y \\ ZX & Z \end{bmatrix}$$

$$\begin{bmatrix} K_{11} & 0 \\ K_{21} & K_{22} \end{bmatrix} = \begin{bmatrix} Z^* & Z^*Y \\ ZX & Z \end{bmatrix}\begin{bmatrix} A & B_\beta E_\beta B_\beta^* \\ C^*JC & -A^* \end{bmatrix}\begin{bmatrix} I & -Y \\ -X & I \end{bmatrix}$$

$$\begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} Z^* & Z^*Y \\ ZX & Z \end{bmatrix} \begin{bmatrix} B_\alpha \\ C^*JD_\alpha \end{bmatrix}$$

$$[C_1 \;\; C_2] = [D_\alpha^*JC \quad -B_\alpha^*] \begin{bmatrix} I & -Y \\ -X & I \end{bmatrix}$$

and

$$\begin{bmatrix} K_{11}^\times & K_{12}^\times \\ 0 & K_{22}^\times \end{bmatrix} = \begin{bmatrix} Z^* & Z^*Y \\ ZX & Z \end{bmatrix} \begin{bmatrix} A^\times & BEB^* \\ C^*\Delta_\alpha C & -(A^\times)^* \end{bmatrix} \begin{bmatrix} I & -Y \\ -X & I \end{bmatrix}$$

In particular

$$\begin{aligned} C_1 &= D_\alpha JC + B_\alpha^* X \\ K_{11} &= Z^*(A - B_\beta E_\beta B_\beta^* X + YC^*JC + YA^*X) \end{aligned}$$

Using (78) to replace YA^* by $-AY - YC^*JCY + B_\beta E_\beta B_\beta^* X$ then gives

$$\begin{aligned} K_{11} &= Z^*(A + YC^*JC - AYX - YC^*JCYX) \\ &= Z^*(A + YC^*JC)(I - YX) \\ &= Z^*(A + YC^*JC)Z^{*-1} \end{aligned} \tag{100}$$

Next

$$B_1 = Z^*B_\alpha + Z^*YC^*JD_\alpha \tag{101}$$

and

$$K_{11} = Z^*(A^* - BEB^*X + YC^*\Delta_\alpha C + Y(A^\times)^*X)$$

Using (77) to replace $(A^\times)^*X$ with $(-XA^\times + XBEB^*X - C^*\Delta_\alpha C)$ leads to

$$\begin{aligned} K_{11}^\times &= Z^*(A^\times - BEB^*X - YXA^\times + YXBEB^*X) \\ &= Z^*(I - YX)(A^\times - BEB^*X) \\ &= A^\times - BEB^*X \end{aligned} \tag{102}$$

Substituting these expressions into the formulas (89) and (90) (with K_{11} in place of A_{11} and $K_{11}^\times$ in place of $A_{11}^\times$) leads to the formulas (82) and (83) for R_α and R_α^{-1}.

We next verify formula (84) for $\tilde{\Theta}_\alpha = G_\alpha R_\alpha^{-1}$. From $G_\alpha(s) = D_\alpha + C(s - A)^{-1}B_\alpha$ and formula (83) for $R_\alpha^{-1}(s)$, we get

$$\begin{aligned} \tilde{\Theta}_\alpha(s) = D_\alpha d_\alpha^{-1} + C(s - A)^{-1}B_\alpha d_\alpha^{-1} - D_\alpha E_\alpha C_1(s - K_{11}^\times)^{-1}B_1 D_\alpha^{-1} \\ - C(s - A)^{-1}M(s - K_{11}^\times)^{-1}B_1 d_\alpha^{-1} \end{aligned} \tag{103}$$

where

$$M = B_\alpha E_\alpha C_1 = B_\alpha E_\alpha(D_\alpha^*JC + B_\alpha^*X)$$

Recall from (74)

$$A - A^\times = B_\alpha E_\alpha D_\alpha^* JC$$

and hence

$$M = A - A^\times + B_\alpha E_\alpha B_\alpha^* X$$

From (101)

$$B_\alpha E_\alpha B_\alpha^* X = -B_\beta E_\beta B_\beta^* X + A^\times - K_{11}^\times$$

and hence

$$\begin{aligned} M &= A - K_{11}^\times - B_\beta E_\beta B_\beta^* X \\ &= (A - s) - (K_{11}^\times - s) - B_\beta E_\beta B_\beta^* X \end{aligned} \tag{104}$$

Now (104) combined with (103) gives

$$\begin{aligned} \tilde{\Theta}_\alpha(s) &= D_\alpha d_\alpha^{-1} + C(s-A)^{-1} B_\alpha d_\alpha^{-1} - D_\alpha E_\alpha C_1 (s - K_{11}^\times)^{-1} B_1 d_\alpha^{-1} \\ &\quad + C(s-A)^{-1}[(s-A) - (s - K_{11}^\times) + B_\beta E_\beta B_\beta^* X](s - K_{11}^\times)^{-1} B_1 d_\alpha^{-1} \\ &= D_\alpha d_\alpha^{-1} + C(s-A)^{-1}(B_\alpha d_\alpha^{-1} - B_1 d_\alpha^{-1}) \\ &\quad + (-D_\alpha E_\alpha C_1 + C)(s - K_{11}^\times)^{-1} B_1 d_\alpha^{-1} \\ &\quad + C(s+A)^{-1} B_\beta E_\beta B_\beta^* X (s - K_{11}^\times)^{-1} B_1 d_\alpha^{-1} \end{aligned} \tag{105}$$

From (73) we have

$$C - D_\alpha E_\alpha D_\alpha^* JC = J\Delta_\alpha C$$

and hence from (98)

$$\begin{aligned} C - D_\alpha E_\alpha C_1 &= C - D_\alpha E_\alpha (D_\alpha^* JC + B_\alpha^* X) \\ &= J\Delta_\alpha C - D_\alpha E_\alpha B_\alpha^* X \end{aligned}$$

Similarly

$$\begin{aligned} B_\alpha - B_1 &= (I - Z^*) B_\alpha - Z^* Y C^* J D_\alpha \\ &= -Z^*(YXB_\alpha + YC^* J D_\alpha) \end{aligned}$$

Now (105) collapses to the formula (84) for $\tilde{\Theta}_\alpha(s)$. Theorem 5.3 follows. □

Proof of sufficiency in Theorem 5.1. We suppose that $G(s) = [D_\alpha \ D_\beta] + C(s-A)^{-1}[B_\alpha \ B_\beta]$ is as in Theorem 5.1 and there exist matrices X, Y satisfying (79)–(81). In particular, since $\rho_\infty(XY) < 1$, $I - XY$ is invertible and we may set $Z = (I - XY)^{-1}$. Now by Theorem 5.2, G admits a

$$\left(\left(\begin{bmatrix} j & 0 \\ 0 & -I \end{bmatrix}, J\right)\text{-isometric}\right)\text{-(restricted outer) factorization } G = \Theta R$$

with the formulas (82)–(84).

It remains to show that in the setting of Theorem 5.2, given that (A, B_β) is controllable, if (81) holds, then Θ satisfies the additional constraint (65) on the right half plane.

We first observe that the controllability of (A, B_β) implies that Y is invertible. Indeed, if $y \in \operatorname{Ker} Y$, then from (78) and (70) we see that

$$0 = -y^* B_\beta E_\beta B_\beta^* y = y^* B_\beta d_\beta^{-1} d_\beta^{*-1} B_\beta^* y$$

and hence $B_\beta^* y = 0$. But then again from (78) we have $YA^* y = 0$. Thus Ker Y is an A^*-invariant subspace contained in the kernel of B_β^*. The controllability of the pair (A, B_β) now forces Ker $Y = (0)$, i.e. Y is invertible. This is the only role of the controllability assumption on (A, B_β); one could assume instead that the solution Y of (78) and (80) is invertible.

By Theorem 4.3 the validity of (65) is equivalent to $\tilde{\Theta} = [\Theta_\alpha G_\beta]$ satisfying

$$\begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix} - \tilde{\Theta}(s)^* J\tilde{\Theta}(s) = \begin{bmatrix} j - \tilde{\Theta}_\alpha(s)^* J\tilde{\Theta}_\alpha(s) & -\tilde{\Theta}_\alpha(s)^* G_\beta(s) \\ -G_\beta(s)^* J\tilde{\Theta}_\alpha(s) & -G_\beta(s)^* JG_\beta(s) \end{bmatrix} \geqslant 0 \tag{106}$$

for Re $s \geqslant 0$, where $\tilde{\Theta}_\alpha(s) = G_\alpha(s)R_\alpha(s)^{-1}$.

As a first step we wish to calculate $j - \tilde{\Theta}_\alpha(s)^* J\tilde{\Theta}_\alpha(s)$. From (84) we have

$$\tilde{\Theta}_\alpha(s) = \hat{D} + \hat{C}(s - \hat{A})^{-1}\hat{B}$$

where $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{D}$ are given by

$$\hat{A} = \begin{bmatrix} A^\times - BEB^*X & 0 \\ B_\beta E_\beta B_\beta^* X & A \end{bmatrix} \tag{107}$$

$$\hat{B} = \begin{bmatrix} Z^* & Z^*Y \\ -YZX & -YZ \end{bmatrix} \begin{bmatrix} B_\alpha \\ C^*JD_\alpha \end{bmatrix} d^{-1} \tag{108}$$

$$\hat{C} = [J\Delta_\alpha C - D_{\alpha\alpha} B_\alpha^* X \ \vdots \ C] \tag{109}$$

and

$$\hat{D} = D_\alpha d_\alpha^{-1} \tag{110}$$

while $G_\beta(s) = D_\beta + C(s - A)^{-1}B_\beta$.

We observed at the beginning of the proof that Y is invertible. Let us set

$$\hat{H} = \begin{bmatrix} X & X \\ X & Y^{-1} \end{bmatrix} \tag{111}$$

Then $\hat{H}$ is Hermitian.

The following lemma gives a formula for the quantity in (106); we postpone the proof of the lemma to the appendix at the end of the section.

Lemma 5.4

With notation as in (106),

$$\begin{aligned} \begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix} - \tilde{\Theta}(s)^* J\tilde{\Theta}(s) &= (s + \bar{s}) \begin{bmatrix} \hat{B}^*(\bar{s} - \hat{A}^*)^{-1} & 0 \\ 0 & B_\beta(\bar{s} - A^*)^{-1} \end{bmatrix} \\ &\times \begin{bmatrix} \hat{H} & \begin{bmatrix} X \\ Y^{-1} \end{bmatrix} \\ [X \ Y^{-1}] & Y^{-1} \end{bmatrix} \begin{bmatrix} (s - \hat{A})^{-1}\hat{B} & 0 \\ 0 & (s - A)^{-1}B_\beta \end{bmatrix} \\ &+ \left\{ \left(\begin{bmatrix} 0 \\ d_\beta^* \end{bmatrix} - \begin{bmatrix} \hat{B}^*(\bar{s} - \hat{A}^*)^{-1} & 0 \\ 0 & B_\beta^*(\bar{s} - A^*)^{-1} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} X \\ Y^{-1} \end{bmatrix} \\ Y^{-1} \end{bmatrix} B_\beta d_\beta^{-1} \right) \right\} \\ &\times \left\{ [0 \ d_\beta] - d_\beta^{*-1} B_\beta^* [[X \ Y^{-1}] \, Y^{-1}] \begin{bmatrix} (s - \hat{A})^{-1}\hat{B} & 0 \\ 0 & (s - A)^{-1}B_\beta \end{bmatrix} \right\} \end{aligned} \tag{112}$$

The next lemma gives the connection between condition (81) in the statement of Theorem 5.1

and the matrices $\hat{H}$ and

$$\begin{bmatrix} \hat{H} & \begin{bmatrix} X \\ Y^{-1} \end{bmatrix} \\ [X \;\; Y^{-1}] & Y^{-1} \end{bmatrix}$$

appearing in (112).

Lemma 5.5.

Let X and Y be Hermitian matrices with Y invertible and set

$$\hat{H} = \begin{bmatrix} X & X \\ X & Y^{-1} \end{bmatrix}$$

Then the following conditions are equivalent.

(i) $X \geqslant 0$, $Y \geqslant 0$ and $\rho_\infty(XY) \leqslant 1$
(ii) $\hat{H} \geqslant 0$
(iii) $\begin{bmatrix} \hat{H} & \begin{bmatrix} X \\ Y^{-1} \end{bmatrix} \\ [X \;\; Y^{-1}] & Y^{-1} \end{bmatrix} \geqslant 0$

Moreover, if any one of these three conditions holds then $\rho_\infty(XY) < 1$ if and only if in addition

$$T = \begin{bmatrix} I & -Y \\ -X & I \end{bmatrix}$$

is invertible.

Proof. Note

$$\hat{H} = \begin{bmatrix} X & X \\ X & Y^{-1} \end{bmatrix}$$

is positive semidefinite if and only if $Y^{-1} > 0$ and $X - XYX = X^{1/2}(I - X^{1/2}YX^{1/2})X^{1/2} \geqslant 0$. Note that if (i) holds then $\rho_\infty(X^{1/2}YX^{1/2}) = \rho_\infty(XY) \leqslant 1$ so $I - X^{1/2}YX^{1/2} \geqslant 0$. This shows that (i) $\Rightarrow$ (ii). Conversely, if (ii) holds that $Y^{-1} > 0$ $X \geqslant XYX \geqslant 0$ and $X^{1/2}(I - X^{1/2}YX^{1/2})X^{1/2} \geqslant 0$. Hence, if $x \in \operatorname{Im} X$, $x^*(I - X^{1/2}YX^{1/2})x \geqslant 0$ while if $x \in \operatorname{Ker} X$, $x^*(I - X^{1/2}YX^{1/2})x = x^*x \geqslant 0$. We conclude that $\rho_\infty(XY) = \rho_\infty(X^{1/2}YX^{1/2}) \leqslant 1$. Thus (ii) $\Rightarrow$ (i).

Clearly (iii) $\Rightarrow$ (ii). Conversely, if (ii) holds, then certainly $Y^{-1} > 0$. By the Schur complement test,

$$\begin{bmatrix} \hat{H} & \begin{bmatrix} X \\ Y^{-1} \end{bmatrix} \\ [X \;\; Y^{-1}] & Y^{-1} \end{bmatrix} \geqslant 0$$

if and only if

$$0 \leqslant \hat{H} - \begin{bmatrix} X \\ Y^{-1} \end{bmatrix} Y [X \;\; Y^{-1}] = \begin{bmatrix} X - XYX & 0 \\ 0 & 0 \end{bmatrix}$$

But we have already noted that (ii) implies that $X - XYX \geqslant 0$. Hence (ii) $\Rightarrow$ (iii).

Finally, if $\rho_\infty(XY) < 1$ then $I - XY$ is invertible so

$$\begin{bmatrix} I & -Y \\ -X & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ -X & I \end{bmatrix} \begin{bmatrix} I & -Y \\ 0 & I - XY \end{bmatrix}$$

is invertible. Conversely, if

$$T = \begin{bmatrix} I & -Y \\ -X & I \end{bmatrix}$$

is invertible, then $I - XY$ is invertible. If also $\rho_\infty(XY) \leqslant 1$, then (since the non-zero spectrum of XY coincides with the non-zero spectrum of the Hermitian matrix $X^{1/2}YX^{1/2}$) necessarily $\rho_\infty(XY) < 1$. The lemma follows. □

Completion of the proof of sufficiency in Theorem 5.1. By assumption, condition (i) in Lemma 5.5 holds. Then by Lemma 5.5,

$$\begin{bmatrix} \hat{H} & \begin{bmatrix} X \\ Y^{-1} \end{bmatrix} \\ [X \;\; Y^{-1}] & Y^{-1} \end{bmatrix} \geqslant 0$$

Now by formula (112), it is clear that

$$\begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix} - \tilde{\Theta}(s)^* J \tilde{\Theta}(s) \geqslant 0$$

i.e., (67) is satisfied, so G has the desired factorization.

Proof of necessity in Theorem 5.1. We now suppose that there are Hermitian matrices X, Y satisfying (77)–(80) and that $G = [D_\alpha \;\; D_\beta] + C(s - A)^{-1}[B_\alpha \;\; B_\beta]$ has a factorization $G = \Theta R$ as in (64)–(66). Then in particular $G = \Theta R$ is a

$$\left(\left(\begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix}, J \right) \text{-isometric} \right) \text{-(restricted outer) factorization for } G$$

so by Theorem 5.2, $I - XY$ is invertible and without loss of generality the formulas (82)–(84) apply. Since Θ satisfies (65), $\tilde{\Theta}$ satisfies (67), i.e.

$$\begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix} - \tilde{\Theta}(s)^* J \tilde{\Theta}(s) \geqslant 0$$

for $s + \bar{s} \geqslant 0$. In particular

$$[I \;\; 0] \left\{ \begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix} - \tilde{\Theta}(s)^* J \tilde{\Theta}(s) \right\} \begin{bmatrix} I \\ 0 \end{bmatrix} \geqslant 0$$

Using the formula (112) from Lemma 5.4, this becomes

$$(s + \bar{s})\hat{B}(\bar{s} - \hat{A}^*)^{-1}\hat{H}(s - \hat{A})^{-1}\hat{B} +$$

$$+ \hat{B}^*(\bar{s} - \hat{A}^*)^{-1} \begin{bmatrix} X \\ Y^{-1} \end{bmatrix} B_\beta d_\beta^{-1} d_\beta^{*-1} B_\beta^* [X \;\; Y^{-1}] (s - \hat{A})^{-1}\hat{B} \geqslant 0 \quad (113)$$

By letting s tend to infinity through real values we see from (113) that

$$\hat{B}(\bar{s} - \hat{A}^*)^{-1}\hat{H}(s - \hat{A})^{-1}\hat{B} \geqslant 0$$

for all real s sufficiently large. The assumption that $(\hat{A}, \hat{B})$ is controllable now implies that $\hat{H} \geq 0$. An application of (ii) ⇒ (i) in Lemma 5.5 now recovers condition (5.18) as desired.

Alternatively, one may work with the formula (112) for the full block 2×2 matrix function

$$\Lambda(s) := \begin{bmatrix} j & 0 \\ 0 & 0 \end{bmatrix} - \tilde{\Theta}(s)^* J \tilde{\Theta}(s)$$

by an asymptotic argument analogous to that just given, $\Lambda(s) \geq 0$ on the right half plane forces

$$\begin{bmatrix} \hat{B}^*(\bar{s} - \hat{A}^*)^{-1} & 0 \\ 0 & B_\beta^*(\bar{s} - A^*)^{-1} \end{bmatrix} \begin{bmatrix} \hat{H} & \begin{bmatrix} X \\ Y^{-1} \end{bmatrix} \\ [XY^{-1}] & Y^{-1} \end{bmatrix} \begin{bmatrix} (s - \hat{A})^{-1}\hat{B} & 0 \\ 0 & (s - A)^{-1}B_\beta \end{bmatrix} \geq 0$$

for all s with $s + \bar{s}$ sufficiently large. Then the controllability of the pair

$$\left(\begin{bmatrix} \hat{A} & 0 \\ 0 & A \end{bmatrix}, \begin{bmatrix} \hat{B} & 0 \\ 0 & B_\beta \end{bmatrix} \right) \quad \text{implies that} \quad \begin{bmatrix} \hat{H} & \begin{bmatrix} X \\ Y^{-1} \end{bmatrix} \\ [X \;\; Y^{-1}] & Y^{-1} \end{bmatrix} \geq 0$$

Now use (iii) ⇒ (i) in Lemma 5.5 to recover condition (81). □

Appendix: Proof of Lemma 5.4

Verification of the formula (107) amounts to checking the validity of the following three identities:

$$j - \hat{\Theta}_\alpha(s)^* J \Theta_\alpha(s) = \hat{B}^*(\bar{s} - \hat{A}^*)^{-1} \left\{ (s + \bar{s})\hat{H} - \begin{bmatrix} X \\ Y^{-1} \end{bmatrix} B_\beta E_\beta B_\beta [X \;\; Y^{-1}] \right\} (s - \hat{A})^{-1} \hat{B} \tag{A.1}$$

$$\begin{aligned} - \tilde{\Theta}_\alpha(x)^* J G_\beta(s) = {} & (s + \bar{s}) \hat{B}^*(\bar{s} - \hat{A}^*)^{-1} \begin{bmatrix} X \\ Y^{-1} \end{bmatrix} (s - A)^{-1} B_\beta - \hat{B}^*(\bar{s} - \hat{A}^*)^{-1} \begin{bmatrix} X \\ Y^{-1} \end{bmatrix} B_\beta \\ & - \hat{B}^*(\bar{s} - \hat{A}^*)^{-1} \begin{bmatrix} X \\ Y^{-1} \end{bmatrix} B_\beta E_\beta B_\beta^* Y^{-1} (s - A)^{-1} B_\beta \end{aligned} \tag{A.2}$$

and

$$\begin{aligned} - G_\beta(s)^* J G_\beta(s) = {} & (\bar{s} + s) B_\beta^*(\bar{s} - A^*)^{-1} Y^{-1} (s - A)^{-1} B_\beta \\ & + d_\beta^* d_\beta - B_\beta^* Y^{-1}(s - A)^{-1} B_\beta - B_\beta^*(\bar{s} - A^*)^{-1} Y^{-1} B_\beta \\ & - B_\beta^*(\bar{s} - A^*)^{-1} Y^{-1} B_\beta E_\beta B_\beta^* Y^{-1} (s - A)^{-1} B_\beta \end{aligned} \tag{A.3}$$

(where we use throughout that $E_\beta = - d_\beta^{-1} d_\beta^{*-1}$ (see (70)).

A first step toward checking (A.1) is to verify the identities

$$\hat{H}\hat{B} = - \hat{C}^* J \hat{D} \tag{A.4}$$

and

$$\hat{A}^* \hat{H} + \hat{H} \hat{A} + \hat{C}^* J \hat{C} = \begin{bmatrix} X \\ Y^{-1} \end{bmatrix} B_\beta E_\beta B_\beta^* [X \;\; Y^{-1}] \tag{A.5}$$

To verify (A.4), we calculate first

$$\hat{H}\hat{B} = \begin{bmatrix} X & X \\ X & Y^{-1} \end{bmatrix} \begin{bmatrix} Z^* & Z^*Y \\ -YZX & -YZ \end{bmatrix} \begin{bmatrix} B_\alpha \\ C^* J D_\alpha \end{bmatrix} d_\alpha^{-1}$$

$$= \begin{bmatrix} X & -XY \\ X & -I \end{bmatrix} \begin{bmatrix} Z^* & Z^*Y \\ ZX & Z \end{bmatrix} \begin{bmatrix} B_\alpha \\ C^*JD_\alpha \end{bmatrix} d_\alpha^{-1}$$

$$= \begin{bmatrix} X & 0 \\ 0 & -I \end{bmatrix} TT^{-1} \begin{bmatrix} B_\alpha \\ C^*JD_\alpha \end{bmatrix} d_\alpha^{-1}$$

$$= \begin{bmatrix} XB_\alpha d_\alpha^{-1} \\ -C^*JD_\alpha d_\alpha^{-1} \end{bmatrix}$$

On the other hand,

$$-\hat{C}^*J\hat{D} = -\begin{bmatrix} C^*\Delta_\alpha J - XB_\alpha E_\alpha D_\alpha^* \\ C^* \end{bmatrix} JD_\alpha d_\alpha^{-1}$$

$$= -\begin{bmatrix} C^*\Delta_\alpha D_\alpha d_\alpha^{-1} - XB_\alpha E_\alpha (D_\alpha^* JD_\alpha) d_\alpha^{-1} \\ C^*JD_\alpha d_\alpha^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} XB_\alpha d_\alpha^{-1} \\ -C^*JD_\alpha d_\alpha^{-1} \end{bmatrix}$$

where we used

$$\begin{aligned} \Delta_\alpha D_\alpha &= (J - JD_\alpha E_\alpha D_\alpha^* J)D_\alpha \\ &= JD_\alpha - JD_\alpha E_\alpha E_\alpha^{-1} = 0 \end{aligned} \tag{A.6}$$

As a first step to verification of (A.5) we analyse the term $\hat{C}^*J\hat{C}$:

$$\hat{C}J\hat{C} = \begin{bmatrix} C^*\Delta_\alpha J - XB_\alpha E_\alpha D_\alpha^* \\ C^* \end{bmatrix} J[J\Delta_\alpha C - D_\alpha E_\alpha B_\alpha^* X \;\vdots\; C]$$

From (A.6) we see that the cross terms in the (1, 1)-block vanish, and the (1, 1)-block is

$$[\hat{C}^*J\hat{C}]_{11} = C^*\Delta_\alpha J\Delta_\alpha C + XB_\alpha E_\alpha (E_\alpha)^{-1} E_\alpha B_\alpha^* X$$

Note

$$\begin{aligned} \Delta_\alpha J\Delta_\alpha &= (J - JD_\alpha E_\alpha D_\alpha^* J)J(J - JD_\alpha E_\alpha D_\alpha^* J) \\ &= J - JD_\alpha E_\alpha D_\alpha^* J - JD_\alpha E_\alpha D_\alpha^* J + JD_\alpha E_\alpha (E_\alpha)^{-1} E_\alpha D_\alpha^* J \\ &= J - JD_\alpha E_\alpha D_\alpha^* J \end{aligned}$$

and hence

$$[\hat{C}^*J\hat{C}]_{11} = C^*\Delta_\alpha C + XB_\alpha E_\alpha B_\alpha^* X \tag{A.7}$$

Therefore verification of (A.5) amounts to checking the validity of the three identities

$$\begin{aligned} ((A^\times)^* - XBEB^*)X + XB_\beta E_\beta B_\beta^* X + X(A^\times - BEB^*X) \\ + XB_\beta E_\beta B_\beta^* X + C^*\Delta_\alpha C + XB_\alpha E_\alpha B_\alpha^* X \\ = XB_\beta E_\beta B_\beta^* X \end{aligned} \tag{A.8a}$$

$$((A^\times)^* - XBEB^*)X + XB_\beta E_\beta B_\beta^* Y^{-1} + XA + C^*\Delta_\alpha C - XB_\alpha E_\alpha D_\alpha^* JC = XB_\beta E_\beta B_\beta^* Y^{-1} \tag{A.8b}$$

$$A^*Y^{-1} + Y^{-1}A + C^*JC = Y^{-1}B_\beta E_\beta B_\beta^* Y^{-1} \tag{A.8c}$$

To verify (A.8a), use the Riccati equation (77) together with the identity

$$XBEB^*X = XB_\alpha E_\alpha B_\alpha^* X + XB_\beta E_\beta B_\beta^* X$$

As for (A.8b), recall

$$B_\alpha E_\alpha D_\alpha^* JC = A - A^\times$$

and use the consequence of the Riccati equation (77)

$$-XBEB^*X + C^*\Delta_\alpha C = -XA^\times + (A^\times)^* X$$

Hence the left-hand side of (A.8b) is

$$\begin{aligned}(A^\times)^*X - XBEB^*X + XB_\beta E_\beta B_\beta^* Y^{-1} + XA + C^*\Delta_\alpha C - XB_\alpha E_\alpha D_\alpha^* JC \\ = (A^\times)^*X + (-XA^\times - (A^\times)^*X) + XB_\beta E_\beta B_\beta^* Y^{-1} + XA - X(A - A^\times) \\ = XB_\beta E_\beta B_\beta^* Y^{-1}\end{aligned}$$

and (A.8b) follows. Finally (A.8c) is an immediate consequence of the Riccati equation (78).

Given that $\tilde{\Theta}_\alpha(s)$ has the form $\tilde{\Theta}_\alpha(s) = \hat{D} + \hat{C}(s-\hat{A})^{-1}\hat{B}$ (see (107)–(110) and (84)) and that $\hat{C}^*J\hat{D} = \hat{H}\hat{B}$ (from (A.4)), we compute

$$\begin{aligned}j - \tilde{\Theta}_\alpha(s)^* J\tilde{\Theta}_\alpha(s) &= j - [\hat{D}^* + \hat{B}^*(\bar{s} - \hat{A}^*)^{-1}\hat{C}^*]\, J[\hat{D} + \hat{C}(s-\hat{A})^{-1}\hat{B}] \\ &= j - \hat{D}^*J\hat{D} - \hat{B}^*(\bar{s}-\hat{A}^*)^{-1}\hat{C}^*J\hat{D} - \hat{D}^*J\hat{C}(s-\hat{A})^{-1}\hat{B} \\ &\quad - \hat{B}^*(\bar{s}-\hat{A}^*)^{-1}\hat{C}^*J\hat{C}(s-\hat{A})^{-1}\hat{B} \\ &= \hat{B}^*(\bar{s}-\hat{A}^*)^{-1}\hat{H}\hat{B} + \hat{B}^*\hat{H}(s-\hat{A})^{-1}\hat{B} \\ &\quad - \hat{B}^*(\bar{s}-\hat{A}^*)^{-1}\hat{C}^*J\hat{C}(s-\hat{A})^{-1}\hat{B} \\ &= \hat{B}^*(\bar{s}-\hat{A}^*)^{-1}[\hat{H}(s-\hat{A}) + (\bar{s}-\hat{A})\hat{H} - \hat{C}^*J\hat{C}]\,(s-\hat{A})^{-1}\hat{B}\end{aligned}$$

Finally, if we use (A.5), we arrive at (A.1) as required.

To verify (A.2) we begin with

$$-\tilde{\Theta}_\alpha(s)^* JG_\beta(s) = -\{\hat{D}^* + \hat{B}^*(\bar{s}-\hat{A}^*)^{-1}\hat{C}^*\}\, J\{D_\beta + C(s-A)^{-1}B_\beta\} \tag{A.9}$$

From assumption (71) and the definitions (107)–(110) of $\hat{A}, \hat{B}, \hat{C}, \hat{D}$, we see that

$$\hat{D}^*JD_\beta = 0 \tag{A.10}$$

and

$$\hat{C}^*JD_\beta = \begin{bmatrix} C^*\Delta_\alpha D_\beta - XB_\alpha^* E_\alpha D_\alpha^* JD_\beta \\ 0 \end{bmatrix}$$

where

$$\begin{aligned}C^*\Delta_\alpha D_\beta &= C^*(J - JD_\alpha E_\alpha D_\alpha^* J)D_\beta \\ &= C^*JD_\beta - C^*JD_\alpha E_\alpha (D_\alpha^* JD_\beta) = 0\end{aligned}$$

and where

$$XB_\alpha^* E_\alpha (D_\alpha^* JD_\beta) = 0$$

Hence

$$\hat{C}^*JD_\beta = 0 \tag{A.11}$$

Next compute

$$\hat{C}^*JC = \begin{bmatrix} C^*\Delta_\alpha C - XB_\alpha E_\alpha D_\alpha^* JC \\ C^*JC \end{bmatrix} \tag{A.12}$$

Thus substitution of (A.10)–(A.12) in (A.9) gives

$$-\tilde{\Theta}_\alpha(s)^*JG_\beta(s) = -\hat{D}^*JC(s-A)^{-1}B_\beta - \hat{B}^*(\bar{s}-\hat{A}^*)^{-1}\begin{bmatrix} C^*\Delta_\alpha C - XB_\alpha E_\alpha D_\alpha^* JC \\ C^*JC \end{bmatrix}(s-A)^{-1}B_\beta \tag{A.12}$$

Now use (A.4) to get

$$\begin{aligned} -\hat{D}^*JC(s-A)^{-1}B_\beta &= -\hat{D}^*J\hat{C}\begin{bmatrix} 0 \\ I \end{bmatrix}(s-A)^{-1}B_\beta \\ &= \hat{B}^*\hat{H}\begin{bmatrix} 0 \\ I \end{bmatrix}(s-A)^{-1}B_\beta \\ &= \hat{B}^*(\bar{s}-\hat{A}^*)^{-1}\cdot(\bar{s}-\hat{A}^*)\hat{H}\begin{bmatrix} 0 \\ I \end{bmatrix}(s-A)^{-1}B_\beta \end{aligned} \tag{A.13}$$

In more detail,

$$\begin{aligned} (\bar{s}-\hat{A}^*)\hat{H}\begin{bmatrix} 0 \\ I \end{bmatrix} &= \begin{bmatrix} \bar{s}-(A^\times)^* + XBEB^* & -XB_\beta E_\beta B_\beta^* \\ 0 & \bar{s}-A^* \end{bmatrix}\begin{bmatrix} X \\ Y^{-1} \end{bmatrix} \\ &= \begin{bmatrix} \bar{s}X-(A^\times)^*X + XBEB^*X - XB_\beta E_\beta B_\beta^* Y^{-1} \\ \bar{s}Y^{-1} - A^*Y^{-1} \end{bmatrix} \end{aligned} \tag{A.14}$$

and hence, making use of the identity $B_\alpha E_\alpha D_\alpha^* JC = A - A^\times$ (see (74)) we get

$$\begin{aligned} &(\bar{s}-\hat{A}^*)\hat{H}\begin{bmatrix} 0 \\ I \end{bmatrix} - \begin{bmatrix} C^*\Delta_\alpha C - XB_\alpha E_\alpha D_\alpha^* JC \\ C^*JC \end{bmatrix} \\ &= \begin{bmatrix} \bar{s}X-(A^\times)^*X + XBEB^*X - XB_\beta E_\beta B_\beta^* Y^{-1} - C^*\Delta_\alpha C + XA - XA^\times \\ \bar{s}Y^{-1} - A^*Y^{-1} - C^*JC \end{bmatrix} \end{aligned} \tag{A.15}$$

From the Riccati equation (77), the first row of the right-hand side of (A.15) becomes

$$\bar{s}X - XB_\beta E_\beta B_\beta^* Y^{-1} + XA = (s+\bar{s})X - X(s-A) - XB_\beta E_\beta B_\beta^* Y^{-1}$$

From the Riccati equation (78) the second row in (A.15) is

$$\begin{aligned} \bar{s}Y^{-1} - A^*Y^{-1} - C^*JC &= \bar{s}Y^{-1} - Y^{-1}B_\beta E_\beta B_\beta^* Y^{-1} + Y^{-1}A \\ &= (s+\bar{s})Y^{-1} - Y^{-1}(s-A) - Y^{-1}B_\beta E_\beta B_\beta^* Y^{-1} \end{aligned}$$

Putting the pieces (A.13)–(A.15) together with (A.12) we get

$$\begin{aligned} &-\tilde{\Theta}_\alpha(s)^*JG_\beta(s) \\ &= \hat{B}^*(\bar{s}-\hat{A}^*)\left\{(s+\bar{s})\begin{bmatrix} X \\ Y^{-1} \end{bmatrix} - \begin{bmatrix} X \\ Y^{-1} \end{bmatrix}(s-A) - \begin{bmatrix} X \\ Y^{-1} \end{bmatrix}B_\beta E_\beta B_\beta^* Y^{-1}\right\}(s-A)^{-1}B_\beta \end{aligned}$$

This yields the claimed identity (A.2).

As for (A.3) begin with

$$-G_\beta(s)^*JG_\beta(s) = -\{D_\beta^* + B_\beta^*(s-A^*)^{-1}C^*\}J\{D_\beta + C(s-A)^{-1}B_\beta\}$$

From (71) we see that the cross terms vanish, while $D_\beta^* J D_\beta = -d_\beta^* d_\beta$ by (70). Hence

$$-G_\beta(s)^* J G_\beta(s) = d_\beta^* d_\beta - B_\beta^*(\bar{s} - A^*)^{-1} C^* J C (s - A)^{-1} B_\beta$$

From the Riccati equation (78) we get

$$\begin{aligned} -G_\beta(s)^* J G_\beta(s) &= d_\beta^* d_\beta - B_\beta^*(\bar{s} - A^*)^{-1}[(\bar{s} - A^*)Y^{-1} + Y^{-1}(s - A) \\ &\quad + Y^{-1} B_\beta E_\beta B_\beta^* Y^{-1} - (\bar{s} + s) Y^{-1}](s - A)^{-1} B_\beta \\ &= d_\beta^* d_\beta - B_\beta^* Y^{-1}(s - A)^{-1} B_\beta - B_\beta^*(\bar{s} - A^*)^{-1} Y^{-1} B_\beta \\ &\quad - B_\beta^*(\bar{s} - A^*)^{-1} Y^{-1} B_\beta E_\beta B_\beta^* Y^{-1}(s - A)^{-1} B_\beta \\ &\quad + (\bar{s} + s) B_\beta^*(\bar{s} - A^*)^{-1} Y^{-1}(s - A)^{-1} B_\beta \end{aligned}$$

This yields (A.3) as required. □

6. THE H^∞ CONTROL PROBLEM: STATE-SPACE COMPUTATIONS

In this section we apply the results on state space formulas for generalized J-inner–outer factorization in Section 5 and the connection of such factorization with H^∞ control as explained in Section 4 to recover the recent elegant state space formulas in Reference 8 for the solution of the H^∞ problem. We suppose that the plant

$$\mathscr{P}: \begin{bmatrix} w \\ u \end{bmatrix} \to \begin{bmatrix} z \\ y \end{bmatrix}$$

for the H^∞ control problem (as explained in Section 4) is given in terms of a state-space presentation

$$\begin{aligned} \dot{x} &= Ax + B_1 w + B_2 u \\ z &= C_1 x + D_{12} u \\ y &= C_2 x + D_{21} w \end{aligned} \tag{114}$$

Since we are interested in studying the problem with some tolerance level γ, we consider instead the plant $\mathscr{P}_\gamma$ given by

$$\begin{aligned} \dot{x} &= Ax + B_1 w + B_2 u \\ z &= \gamma^{-1} C_1 x + \gamma^{-1} D_{12} u \\ y &= C_2 x + D_{21} w \end{aligned} \tag{115}$$

in order to normalize the tolerance level to 1. Here we discuss only the so-called *regular case*; hence we assume:

$$D_{12}^* D_{12} > 0, \qquad D_{21} D_{21}^* > 0 \tag{116}$$

as well as

$$\begin{bmatrix} A - sI & B_1 \\ C_2 & D_{21} \end{bmatrix} \text{ is surjective on the imaginary line} \tag{117}$$

and

$$\begin{bmatrix} A - sI & B_2 \\ C_1 & D_{12} \end{bmatrix} \text{ is injective on the imaginary line} \tag{118}$$

Note that (117) and (118) have the consequence that the transfer function $\mathscr{P}_{21}(s) = D_{21} + C_2(s-A)^{-1}B_1$ is surjective and $\mathscr{P}_{12}(s) = D_{12} + C_1(s-A)^{-1}B_2$ is injective on the imaginary line. A consequence of (116) is that, by choosing appropriate decompositions $Z = Z_1 \oplus Z_2$ and $W = W_1 \oplus W_2$ of the output space Z where z takes values and the input space W where w takes values, we may assume that D_{12} and D_{21} have the form

$$D_{12} = \begin{bmatrix} D_{121} \\ 0 \end{bmatrix}, \qquad D_{21} = [D_{211} \ \ 0] \tag{119}$$

where D_{121} and D_{211} are square and invertible. We then form the partitionings of B_1 and C_1

$$B_1 = [B_{11} \ \ B_{12}], \qquad C_1 = \begin{bmatrix} C_{11} \\ C_{12} \end{bmatrix} \tag{120}$$

consistent with the above decompositions of W and Z. Now the state-space equations (114) for $\mathscr{P}_\gamma$ have the form

$$\begin{aligned} \dot{x} &= Ax + B_{11}w_1 + B_{12}w_2 + B_2u \\ z_1 &= \gamma^{-1}C_{11}x + \gamma^{-1}D_{121}u, D_{121} \quad \text{invertible} \\ z_2 &= \gamma^{-1}C_{12}x \\ y &= C_2x + D_{211}w_1, D_{211} \quad \text{invertible} \end{aligned} \tag{121}$$

A choice of augmentation $[\mathscr{P}^0_{21}(s) \ \ \mathscr{P}^0_{22}(s)]$ to $\mathscr{P}_\gamma$ so that $\begin{bmatrix} \mathscr{P}_{21}(s) \\ \mathscr{P}^0_{21}(s) \end{bmatrix}$ is invertible at infinity is to include the component w_2 of the input signal w as the additional fictitious output signal; this amounts to adding the equation

$$y^0 = w_2 \tag{122}$$

to those already listed in (120). Reversing the arrows (as explained in Section 1) to get

$$\tilde{\mathscr{P}}_\gamma : \begin{bmatrix} u \\ y \\ y^0 \end{bmatrix} \to \begin{bmatrix} z_1 \\ z_2 \\ w_2 \\ w_2 \end{bmatrix}$$

then leads to the state-space equations

$$\begin{aligned} \dot{x} &= (A - B_{11}D_{211}^{-1}C_2)x + B_2u + B_{11}D_{211}^{-1}y + B_{12}y^0 \\ z_1 &= \gamma^{-1}C_{11}x + \gamma^{-1}D_{121}u \\ z_2 &= \gamma^{-1}C_{12}x \\ w_1 &= -D_{211}^{-1}C_2x + D_{211}^{-1}y \\ w_2 &= y^0 \end{aligned} \tag{123}$$

for $\tilde{\mathscr{P}}_\gamma$. By Theorem 4.2, solving the H^∞ control problem is intimately related to a generalized J-inner–outer factorization of $\tilde{\mathscr{P}}$. We can get a state-space analysis and formulas for such a factorization by applying Theorem 5.1 with the substitution

$A - B_{11}D_{212}^{-1}C_2$ in place of A

$[B_2 \ \ B_{11}D_{211}^{-1}]$ in place of B_α

B_{12} in place of B_β

$$\begin{bmatrix} \gamma^{-2}C_{11} \\ \gamma^{-1}C_{12} \\ -D_{211}^{-1}C_2 \\ 0 \end{bmatrix} \quad \text{in place of } C$$

$$\begin{bmatrix} \gamma^{-1}D_{121} & 0 \\ 0 & 0 \\ \cdots & \cdots \\ 0 & D_{211} \\ 0 & 0 \end{bmatrix} \quad \text{in place of } D_\alpha$$

$$\begin{bmatrix} 0 \\ 0 \\ \cdot \\ 0 \\ I \end{bmatrix} \quad \text{in place of } D_\beta$$

$$\begin{bmatrix} I_{Z_1} & 0 & 0 & 0 \\ 0 & I_{Z_2} & 0 & 0 \\ 0 & 0 & -I_{W_1} & 0 \\ 0 & 0 & 0 & -I_{W_2} \end{bmatrix} \quad \text{in place of } J$$

and

$$\begin{bmatrix} I_U & 0 \\ 0 & -I_Y \end{bmatrix} \quad \text{in place of } j \tag{124}$$

with

$$G_\alpha(s) \sim \left[\begin{array}{c:cc} A - B_{11}D_{211}^{-1}C_2 & B_2 & B_{11}D_{211}^{-1} \\ \hdashline \gamma^{-1}C_{11} & & \gamma^{-1}D_{121} \\ \gamma^{-1}C_{12} & 0 & 0 \\ -D_{211}^{-1}C_2 & 0 & D_{211}^{-1} \\ 0 & 0 & 0 \end{array}\right]$$

and

$$G_\beta(s) \sim \left[\begin{array}{c:c} A - B_{11}D_{211}^{-1}C_2 & B_{12} \\ \hdashline \gamma^{-1}C_{11} & 0 \\ \gamma^{-1}C_{12} & 0 \\ -D_{211}^{-1}C_2 & 0 \\ 0 & I \end{array}\right]$$

One easily checks that the hypotheses (69)–(71) are fulfilled; indeed

$$D_\alpha^* J D_\alpha = d_\alpha^* j d_\alpha \qquad \text{with } d_\alpha = \begin{bmatrix} \gamma^{-1}D_{121} & 0 \\ 0 & D_{211} \end{bmatrix} \tag{125}$$

and

$$D_\beta^* J D_\beta = -d_\beta^* d_\beta \qquad \text{with } d_\beta = I_{Y^0} \tag{126}$$

Further notation from (73)–(76) becomes

$$\Delta_\alpha = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & I_{Z_2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -I_{W_2} \end{bmatrix} \tag{127}$$

$$A^\times = A - B_2(D_{121})^{-1}C_{11} \tag{128}$$

The Riccati equations (77) and (78) become

$$X(A - B_2(D_{121})^{-1}C_{11}) + (A - B_2(D_{121})^{-1}C_{11})^* X - X(B_2E_1B_2^* - B_{11}B_{11}^* - B_{12}B_{12}^*)X + \gamma^{-2}C_{12}^*C_{12} = 0 \tag{129}$$

and

$$Y(A - B_{11}(D_{211})^{-1}C_2)^* + (A - B_{11}(D_{211})^{-1}C_2)Y + Y(\gamma^{-2}C_{11}^*C_{11} - \gamma^{-2}C_{12}^*C_{12} - C_2^*E_2C_2)Y + B_{12}B_{12}^* = 0 \tag{130}$$

where we have set

$$E_1 = \{(D_{121})^*D_{121}\}^{-1}, \qquad E_2 = \{D_{211}(D_{211})^*\}^{-1} \tag{131}$$

with the stability side conditions (79) and (80) assuming the form

$$\sigma(A - B_2(D_{121})^{-1}C_{11} - (B_2E_1B_2^T - B_{11}B_{11}^* - B_{12}B_{12}^*)X) \subset \{\text{Re } s < 0\} \tag{132}$$

and

$$\sigma(A - B_{11}(D_{211})^{-1}C_2 - Y(\gamma^{-2}C_{11}^*C_{11} + \gamma^{-2}C_{12}^*C_{12} - C_2^*E_2C_2)) \subset \{\text{Re } s < 0\} \tag{133}$$

The reader will note that, after a change in notation, (129)–(133) are exactly the same as the equations appearing in Reference 8 in connection with the H^∞ control problem. Combining Theorems 4.2, 4.3 and 5.1 now gives us the following result for the H^∞ control problem.

Theorem 6.1

Let

$$\mathscr{P}(s) \sim \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right]$$

be a plant satisfying the conditions (116)–(118) and partition D_{12}, D_{21}, B_1 and C_1 as in (119)–(120). Assume that the Riccati equations (129) and (130) have stabilizing solutions X, Y (satisfying (132) and (133)) and assume that $(A - B_{11}D_{211}^{-1}C_2, B_{12})$ is controllable. Then the H^∞ problem with tolerance level γ (HINFγ) has solutions K if

$$X \geqslant 0, \qquad Y \geqslant 0 \qquad \text{and} \qquad \rho_\infty(XY) < 1 \tag{134}$$

Conversely, if (134) holds and in addition if the pair

$$\left(\begin{bmatrix} A - B_2(D_{121})^{-1}C_{11} & 0 \\ -B_{12}B_{12}^*X & A - B_{11}(D_{211})^{-1}C_2 \end{bmatrix}, \begin{bmatrix} Z^* & ZXY \\ -YZX & -YZ \end{bmatrix} \begin{bmatrix} B_2 & B_{11}(D_{211})^{-1} \\ \gamma^{-2}C_{11}^*D_{121} & C_2^*E_2 \end{bmatrix} \right)$$

is controllable, then the H^∞ problem (HINFγ) has solutions. In this case the set of compensators $K(s)$ solving the problem (HINFγ) is given by

$$K(s) = (r_{11}(s)H(s) + r_{12}(s))(r_{21}(s)H(s) + r_{22}(s))^{-1} \tag{135}$$

where $H(s)$ is any stable rational matrix function with $\|H\|_\infty < 1$ for which $(r_{21}H + r_{22})^{-1}$ is proper and where

$$r(s) = \begin{bmatrix} r_{11}(s) & r_{12}(s) \\ r_{21}(s) & r_{22}(s) \end{bmatrix}$$

is given by

$$r(s) = d + c(s-a)^{-1}b$$

with

$$d = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} = \begin{bmatrix} \gamma D_{121}^{-1} & 0 \\ 0 & D_{211} \end{bmatrix} \tag{136a}$$

$$c = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -D_{121}^{-1}C_{11} - \gamma^2 D_{121}^{-1}(D_{121}^*)^{-1}B_2^*X \\ C_2 + D_{211}B_{11}^*X \end{bmatrix} \tag{136b}$$

$$b = [b_1 \ b_2] = [\gamma B_2 D_{121}^{-1} + \gamma^{-1}YC_{11}^* \ \ B_{11} + YC_2^*(D_{211}^{-1})^*] \tag{136c}$$

and

$$a = A - B_2(D_{121})^{-1}C_{11} + (-\gamma^2 B_2 D_{121}^{-1}(D_{121}^{-1})^* B_2^* + B_1 B_1^*)X \tag{136d}$$

One particular compensator is given by

$$K_0(s) = C_{K_0}(s - A_{K_0})^{-1}B_{K_0}$$

where

$$A_{K_0} = A - B_2(D_{121})^{-1}C_{11} + [-\gamma^2 B_2 D_{121}^{-1}(D_{121}^{-1})^* B_2^* + B_1 B_1^*]X \\ - [B_{11} + YC_2^*(D_{211}^{-1})^*]D_{211}^{-1}[C_2 + D_{211}B_{11}^*X] \tag{137a}$$

$$B_{K_0} = B_{11}D_{211}^{-1} + YC_2^*(D_{211}^{-1})^* D_{211}^{-1} \tag{137b}$$

and

$$C_{K_0} - D_{121}^{-1}C_{11} - \gamma^2 D_{121}^{-1}(D_{121}^{-1})^* B_2^* X$$

Proof. We apply Theorem 5.3 to $\tilde{\mathscr{P}}_\gamma$ given by (123). The formula (136a–d) for $r(s)$ comes from the formula (83) for $R_\alpha(s)^{-1}$. By Theorem 4.3, the formula (135) provides a parametrization of the compensators solving the H^∞ problem. The particular compensators K given by (137) results from plugging $H = 0$ into (135). To see this note that a realization for

$$r(s) = \begin{bmatrix} r_{11}(s) & r_{12}(s) \\ r_{21}(s) & r_{22}(s) \end{bmatrix} : \begin{bmatrix} u \\ y \end{bmatrix} \to \begin{bmatrix} z \\ w \end{bmatrix}$$

is

$$\dot{x} = ax + b_1 u + b_2 y$$
$$z = c_1 x + d_1 u$$
$$w = c_2 x + d_2 y$$

Solving for $\begin{bmatrix} \dot{x} \\ z \\ y \end{bmatrix}$ in terms of $\begin{bmatrix} x \\ w \\ u \end{bmatrix}$ results in

$$\begin{aligned} \dot{x} &= (a - b_2 d_2^{-1} c_2)x + b_2 d_2^{-1} w + b_1 u \\ z &= c_1 x + d_1 u \\ y &= -d_2^{-1} c_2 x + d_2^{-1} w \end{aligned}$$

The associated transfer function from w to z (with input u taken equal to zero and output y ignored) therefore has the realization

$$\begin{aligned} \dot{x} &= (a - b_2 d_2^{-1} c_2)x + b_2 d_2^{-1} w \\ z &= c_1 x \end{aligned}$$

But this transfer function coincides with $r_{12} r_{22}^{-1} = \mathscr{G}_r[0] = K_0$. In this way we get a realization (137a–d) for $r_{12} r_{22}^{-1}$ of the same order as the realization (136a–d) for r. In particular, we see that there is a compensator solving the H^∞ problem with McMillan degree no more than the McMillan degree of the original plant $\mathscr{P}$.

ACKNOWLEDGEMENTS

The first two authors were partially supported by the Air Force Office of Scientific Research and the National Science Foundation. The third author was partially supported by the NSF and the Canadian NSERC.

REFERENCES

1. Francis, B. A., *A Course in H^∞ Control*, Springer-Verlag, New York, 1987.
2. Vidyasagar, M., *Control System Synthesis: A Factorization Approach*, The MIT Press, Cambridge, Mass., 1985.
3. Ball, J. A., and J. W. Helton, 'Linear fractional parametrization of matrix-function spaces and a new proof of the Youla–Jabr–Bongiorno parametrization for stabilizing compensators', in *Mathematical Theory of Networks and Systems*, P. Fuhrmann (ed.), Springer-Verlag, New York, 1984, pp. 16–23.
4. Doyle, J. C., K. Glover, P. P. Khargonekar and B. A. Francis, 'State-space solutions to standard H^2 and H^∞ control problems', *IEEE Trans. Auto. Control*, **AC-34**, 831–847 (1989).
5. Dym, H., *J-Contractive Matrix Functions, Reproducing Kernel Hilbert Spaces and Interpolation*, CBMS #71, American Mathematical Society, Providence, RI, 1989.
6. Green, M., 'H^∞ controller synthesis by J-lossless coprime factorization', *SIAM J. Control and Opt.*, to appear.
7. Ball, J. A., and N. Cohen, 'Sensitivity minimization in H^∞ norm: parametrization of all suboptimal solutions', *Int. J. Control*, **46**, 785–816 (1987).
8. Petersen, I. R., B. D. O. Anderson and E. A. Jonkheere, 'A first principles solution to the nonsingular H^∞ control problem'. *International Journal of Robust and Nonlinear Control*, **1**(3), 171–185 (1991).
9. Tadmor, G., 'The standard H_∞ problem and the maximum principle: the general linear case', preprint.
10. Green, M., K. Glover, D. J. N. Limebeer and J. Doyle, 'A J-spectral factorization approach to H^∞ control', *SIAM J. Control and Opt.*, **28**, 1350–1371 (1990).
11. Ball, J. A., and J. W. Helton, 'A Beurling–Lax theorem for the Lie group $U(m, n)$ which contains most classical interpolation', *J. Operator Theory*, **9**, 107–142 (1983).
12. Liu, K.-Z., T. Mita and H. Kimura, 'Complete solution to the standard H^∞ control problem of discrete time systems', in *Proceedings of the 29th Conference on Decision and Control*, Honolulu, 1990, pp. 1786–1793.
13. Stoorvogel, A. A., 'The singular H^∞ control problem with dynamic measurement feedback', *SIAM J. Control and Opt.*, **29**, 160–184 (1991).
14. Ball, J. A., and M. A. Kaashoek, 'Indefinite inner–outer factorization of rational matrix functions', in preparation.
15. Chen, T., and B. A. Francis, 'Spectral and inner-outer factorizations of rational matrices', *SIAM J. Matrix Anal. Appl.*, **10**, 1–17 (1989).
16. Green, M., 'On inner–outer factorization', *Systems and Control Letters*, **11**, 93–97 (1988).

17. Belevitch, V., *Classical Network Theory*, Holden Day, San Francisco, 1968.
18. Gohberg, I. and S. Rubinstein, 'Cascade decomposition of rational matrix functions and their stability, *International Journal of Control*, **46**, 603–629 (1987).
19. Kailath, T., *Linear Systems*, Prentice Hall, Englewood Cliffs, NJ, 1980.
20. Forney, G. D., 'Minimal bases of rational vector spaces with applications to multivariable linear systems', *SIAM J. Control*, **13**, 493–520 (1975).
21. Ball, J. A., I. Gohberg and L. Rodman, *Interpolation of Rational Matrix Functions*, Birkhäuser Verlag, Basel, 1990.
22. Ball, J. A., and M. Rakowski, 'Interpolation by rational matrix functions and stability of feedback systems: the 4-block case', to appear in *Proc. Workshop on Operator Theory and Complex Analysis*, T. Ando (ed.), Birkhäuser-Verlag, Basel–Berlin–Boston.
23. Bart, H., I. Gohberg and M. A. Kaashoek, 'Minimal factorization of matrix and operator functions', Birkhäuser Basel, 1979.
24. Bart, H., I. Gohberg and M. A. Kaashoek, 'Explicit Wiener–Hopf factorization and realization', in *Constructive Methods of Wiener–Hopf Factorization*, I. Gohberg and M. A. Kaashoek (eds), Birkhäuser Basel, 1986, pp. 235–316.
25. Glover, K., 'All optimal Hankel-norm approximations of linear multivariable systems and their L^∞ error bounds', *Int. J. Control*, **39**, 1115–1193 (1984).
26. Hara, S., and R. Kondo, 'Characterization and computation of H^∞-optimal controllers in the state space', *Proc. 27th IEEE Conf. on Decision and Control*, Austin, IEEE, Piscataway, N.J., 1988, pp. 20–25.
27. Verma, M. and G. Zames, 'An energy and two part framework for H^∞ control', in *Control of Uncertain Dynamic Systems*, S. P. Bhattacharyya and L. H. Keel (eds), CRC Press, 1991, pp. 39–51.