Rational and Polynomial Matrix Factorizations via Recursive Pole-Zero Cancellation

P. Van Dooren
Philips Research Laboratory Belgium
1348 Louvain-la-Neuve
B-1170 Brussels, Belgium

Submitted by Leiba Rodman

ABSTRACT

We develop a recursive algorithm for obtaining factorizations of the type

$$R(\lambda) = R_1(\lambda) R_2(\lambda)$$

where all three matrices are rational and $R_1(\lambda)$ is nonsingular. Moreover the factors $R_1(\lambda)$ and $R_2(\lambda)$ are such that either the *poles* of $[R_1(\lambda)]^{-1}$ and $R_2(\lambda)$ are in a prescribed region Γ of the complex plane, or their *zeros*. Such factorizations cover the specific cases of coprime factorization, inner-outer factorization, GCD extraction, and many more. The algorithm works on the state-space (or generalized state-space) realization of $R(\lambda)$ and derives in a recursive fashion the corresponding realizations of the factors.

1. INTRODUCTION

Several problems occurring in the literature of linear systems theory can be rephrased as a factorization problem for some rational matrix $R(\lambda)$. In this paper we consider a certain class of such factorizations, namely where the $p \times m$ rational matrix $R(\lambda)$ is factored into a product of two rational matrices:

$$R(\lambda) = R_1(\lambda) R_2(\lambda) \tag{1.1}$$

655 Avenue of the Americas, New York, NY 10010 0024-3795/90/$3.50

where:

(i) $R_1(\lambda)$ is $p \times p$ nonsingular, i.e., $R_1^{-1}(\lambda)$ exists.

(ii) Given a region Γ of the complex plane, one of the following two conditions is satisfied:

(a) the *poles* of $R_1^{-1}(\lambda)$ and $R_2(\lambda)$ lie in Γ;
(b) the *zeros* of $R_1^{-1}(\lambda)$ and $R_2(\lambda)$ lie in Γ.

(iii) The (McMillan) degree δ of $R_1(\lambda)$ is minimal, which is shown to imply that, according to the choice (ii)(a) or (ii)(b), either

$$\delta(R_1) = \#\ \textit{poles} \text{ of } R(\lambda) \text{ outside } \Gamma \tag{1.2a}$$

or

$$\delta(R_1) = \#\ \textit{zeros} \text{ of } R(\lambda) \text{ outside } \Gamma. \tag{1.2b}$$

We will show in the sequel that there always exist such factorizations for any region Γ and that they are in fact far from unique unless some additional conditions are imposed. Notice also that to each factorization of the above type there corresponds a *dual* factorization where the roles of $R_1(\lambda)$ and $R_2(\lambda)$ are interchanged. These are easily obtained by working as above on the transpose of $R(\lambda)$, and transposing the obtained factors. It is common to call the above factorization a "left" one and the dual type a "right" one, but in the sequel we will only work with left factorizations and therefore drop this adjective. The above class of factorizations is rather general and covers the following special cases, which are well known in different application areas:

(1) *Coprime factorization.* Given a $p \times m$ rational matrix $R(\lambda)$, one wants to find polynomial matrices $D(\lambda)$ and $N(\lambda)$, with $D(\lambda)$ nonsingular, such that

$$R(\lambda) = D^{-1}(\lambda)N(\lambda). \tag{1.3}$$

This fits into the above formulation with condition (ii)(a), where $\Gamma = \{\infty\}$. The condition (1.2a) is usually replaced here by the less restrictive condition that (1.2a) should only hold for the determinantal degree (see e.g. [11]):

$$\delta(\det D(\lambda)) = \#\text{ finite poles of } R(\lambda). \tag{1.4}$$

We will see in the sequel that the stronger constraint

$$\delta(D(\lambda)) = \#\text{ finite poles of } R(\lambda) \tag{1.5}$$

is always satisfied by our constructive algorithm.

(2) *All-pass extraction.* Given a $p \times m$ rational matrix $R(z)$, one wants to find an *all-pass* transfer function $R_1(z)$, i.e.,

$$[R_1(z^{-*})]^* R_1(z) = I_p \tag{1.6}$$

[where $\cdot^*$ denotes the conjugate transposed and $\cdot^{-*}$ means $(\cdot^*)^{-1}$], such that $R_2(z) = R_1^{-1}(z)R(z)$ is stable in the sense of discrete-time systems (i.e., all poles are inside the unit circle). This can always be obtained by a stable factor $R_1^{-1}(z)$, and the problem thus clearly fits into the above formulation with condition (ii)(a) where Γ is the closed unit disc: $\Gamma = \{z \mid |z| \leqslant 1\}$. Equation (1.6) is an *additional* condition to the general ones, but can easily be incorporated in the method presented in this paper, as will be shown in the sequel.

There is also a continuous-time analogue to the previous problem, where now "all-pass" means a transfer function $R_1(s)$ satisfying

$$[R_1(-s^*)]^* R_1(s) = I_p, \tag{1.7}$$

and where the stable region Γ is now the closed left half plane: $\Gamma = \{s \mid \mathrm{Re}(s) \leqslant 0\}$.

(3) *Inner-outer factorization.* Here a stable (in the discrete-time sense) transfer matrix $R(z)$ is given, and one seeks to extract a factor $R_1(z)$ which is all-pass—i.e. satisfies (1.6)—and such that the resulting factor $R_2(z)$ has all its zeros inside the unit disc. The "inner factor" $R_1(z)$ can always be chosen to have zeros inside the unit disc as well, which then clearly fits into the above formulation with condition (ii)(b) where Γ is the closed unit disc. Again, the additional condition (1.6) is easily incorporated in our method. This is the rational matrix version of what is usually called the "inner-outer factorization" [8] in functional analysis. In systems theory the factor $R_2(z)$ is also often called the "minimum-phase" factor of $R(z)$ [21]. Here also there is a continuous-time analogue to the problem of inner-outer factorization, which should perhaps be termed "left-right factorization," since now the stability region Γ for the zeros of $R(s)$ is the closed left half plane. The factor $R_1(s)$ now satisfies (1.7), and $R_2(s)$ has all its zeros in Γ.

(4) *GCD extraction.* Let $P_i(\lambda)$, $i=1,\dots,k$, be a set of polynomial matrices of dimensions $p \times m_i$. Then their greatest common (left) divisor (GCD) is defined as the *nonsingular* polynomial matrix $D(\lambda)$ such that

$$P(\lambda) = D(\lambda)Q(\lambda) \tag{1.8a}$$

where

$$P(\lambda) = [P_1(\lambda),\dots,P_k(\lambda)], \tag{1.8b}$$

$$Q(\lambda) = [Q_1(\lambda),\dots,Q_k(\lambda)], \tag{1.8c}$$

and where the *quotients* $Q_i(\lambda)$ form together a polynomial matrix $Q(\lambda)$ with no Smith zeros [11]. Hence, both $D^{-1}(\lambda)$ and $Q(\lambda)$ have all their McMillan zeros at ∞, which is thus condition (ii)(b) with $\Gamma = \{\infty\}$. The additional degree condition usually imposed is

$$\delta(\det D(\lambda)) = \#\text{ finite zeros of } P(\lambda), \tag{1.9}$$

which is less restrictive than (1.2b). When $P(\lambda)$ has normal rank $r < p$, this degree condition is crucial for restricting the class of possible solutions.

(5) *Polynomial factor extraction.* Given a $p \times m$ polynomial matrix $P(\lambda)$, one wants to find a $p \times p$ *nonsingular divisor* $P_1(\lambda)$ which contains all the zeros of $P(\lambda)$ inside a region Γ_f *not* containing the point at infinity:

$$P(\lambda) = P_1(\lambda)P_2(\lambda). \tag{1.10}$$

The factors $P_1^{-1}(\lambda)$ and $P_2(\lambda)$ have thus their zeros in the complement $\Gamma = \Gamma_f^c$, and this factorization thus satisfies (ii)(b). The additional degree condition

$$\delta(\det P_1(\lambda)) = \#\text{ zeros of } P(\lambda) \text{ in } \Gamma_f \tag{1.11}$$

again makes the factorization essentially unique when the normal rank r of $P(\lambda)$ is smaller than p. The determinantal conditions (1.5), (1.9), (1.11), which are less restrictive than (1.2b), are easy to fulfill by our algorithm, as is shown later on.

2. A RECURSIVE APPROACH

In the development of our algorithm we were strongly inspired by ideas found in Belevitch [2] and more elaborately in the work of Dewilde and Vandewalle [5, 14]. There it is shown that it is always possible to find a nonsingular $p \times p$ rational transfer function $C(\lambda)$ of degree 1, i.e. with one pole γ and one zero δ, such that in the product $R_2(\lambda) = C(\lambda)R(\lambda)$, *either* δ cancels with a pole (say α_1) of $R(\lambda)$, *or* γ cancels with a zero (say β_1) of $R(\lambda)$. For such a cancellation to occur, one of course needs $\delta = \alpha_1$ (respectively, $\gamma = \beta_1$), but in the matrix case some additional vector conditions are required and can always be satisfied as shown in [5]. In some cases, when both $\delta = \alpha_1$ and $\gamma = \beta_1$, it can happen that *both* γ and δ cancel, but this is inessential for our present discussion.

Let us represent the rational matrix $R(\lambda)$ via its poles $\{\alpha_i \,|\, i = 1,\dots,l\}$ and zeros $\{\beta_j \,|\, j = 1,\dots,k\}$. Then the product $R_2(\lambda) = C(\lambda)R(\lambda)$ can be represented as either of the following two:

$$\left[\frac{\cancel{\delta}}{\gamma}\right]\left(\frac{\beta_1,\dots,\beta_k}{\cancel{\alpha}_1,\dots,\alpha_l}\right) = \left(\frac{\beta_1,\beta_2,\dots,\beta_k}{\gamma,\alpha_2,\dots,\alpha_l}\right), \tag{2.1a}$$

$$\left[\frac{\delta}{\cancel{\gamma}}\right]\left(\frac{\cancel{\beta}_1,\dots,\beta_k}{\alpha_1,\dots,\alpha_l}\right) = \left(\frac{\delta,\beta_2,\dots,\beta_k}{\alpha_1,\alpha_2,\dots,\alpha_l}\right). \tag{2.1b}$$

Here we have chosen to use square brackets for *nonsingular* rational matrices and round ones for (possibly) *singular* rational matrices. The number of poles is by definition the McMillan degree of the (nonsingular or singular) rational matrix. Notice that for nonsingular rational matrices, this also equals the number of zeros, while for the singular case, the number of zeros can be less than the McMillan degree [20]. Moreover, the inverse $C^{-1}(\lambda)$ of a nonsingular rational matrix has its poles and zeros interchanged. Therefore, from (2.1) one obtains for $R(\lambda) = C^{-1}(\lambda)R_2(\lambda) = R_1(\lambda)R_2(\lambda)$

$$\left(\frac{\beta_1,\dots,\beta_k}{\alpha_1,\dots,\alpha_l}\right) = \left[\frac{\gamma}{\alpha_1}\right]\left(\frac{\beta_1,\dots,\beta_k}{\gamma,\alpha_2,\dots,\alpha_l}\right) \tag{2.2a}$$

or

$$\left(\frac{\beta_1,\dots,\beta_k}{\alpha_1,\dots,\alpha_l}\right) = \left[\frac{\beta_1}{\delta}\right]\left(\frac{\delta,\beta_2,\dots,\beta_k}{\alpha_1,\dots,\alpha_l}\right). \tag{2.2b}$$

It is shown in [5] that in the construction of $C(\lambda)$ in (2.1a) the pole γ can be chosen arbitrarily, while δ is fixed, since it must cancel with α_1. Now if α_1 had been the only *pole* of $R(\lambda)$ outside a given region Γ, then the factorization (2.2a) with $R_1(\lambda)=C^{-1}(\lambda)$ would correspond to a factorization as described in Section 1. All conditions (i), (ii), and (iii)(a) would indeed be satisfied iff γ were chosen inside Γ as well, which is always possible as indicated above. The same holds for (2.1b), where now $\gamma=\beta_1$ cancels a zero β_1 of $R(\lambda)$. If β_1 were the only *zero* outside Γ, then (2.2b) would satisfy all conditions (i), (ii) and (iii)(b) of the factorization described in Section 1, provided δ was chosen inside Γ. The factor $C(\lambda)$ in (2.1) has thus "dislocated" one pole $\alpha_1=\delta$ (respectively, zero $\beta_1=\gamma$) outside Γ to a pole γ (zero δ) inside Γ, using a factor $R_1^{-1}(\lambda)=C(\lambda)$ which itself has a pole γ (zero δ) inside Γ.

When several poles $\{\alpha_i \mid i=1,\dots,l_0\}$ (zeros $\{\beta_j \mid j=1,\dots,k_0\}$) are outside Γ, they can be dislocated recursively one after the other by such first-degree sections $C_i(\lambda)$ as are described in (2.1), (2.2). This then yields

$$\left[\frac{\cancel{\delta}_{l_0}}{\gamma_{l_0}}\right]\cdots\left[\frac{\cancel{\delta}_2}{\gamma_2}\right]\left[\frac{\cancel{\delta}_1}{\gamma_1}\right]\left(\frac{\beta_1,\beta_2,\dots,\beta_k}{\cancel{\alpha}_1,\dots,\cancel{\alpha}_{l_0},\alpha_{l_0+1},\dots,\alpha_l}\right)$$
$$=\left(\frac{\beta_1,\beta_2,\dots,\beta_k}{\gamma_1,\dots,\gamma_{l_0},\alpha_{l_0+1},\dots,\alpha_l}\right), \tag{2.3a}$$

$$\left[\frac{\delta_{k_0}}{\cancel{\gamma}_{k_0}}\right]\cdots\left[\frac{\delta_2}{\cancel{\gamma}_2}\right]\left[\frac{\delta_1}{\cancel{\gamma}_1}\right]\left(\frac{\cancel{\beta}_1,\dots,\cancel{\beta}_{k_0},\beta_{k_0+1},\dots,\beta_k}{\alpha_1,\alpha_2,\dots,\alpha_l}\right)$$
$$=\left(\frac{\delta_1,\dots,\delta_{k_0},\beta_{k_0+1},\dots,\beta_k}{\alpha_1,\alpha_2,\dots,\alpha_l}\right). \tag{2.3b}$$

Considering the product of the nonsingular factors $C_i(\lambda)$ as $R_1^{-1}(\lambda)$ and the right-hand side as $R_2(\lambda)$, this certainly satisfies the imposed conditions, since (i) $R_1(\lambda)$ is constructed to be nonsingular, (ii) the poles (zeros) of $R_2(\lambda)$ and $R_1^{-1}(\lambda)$ are in Γ by construction, and (iii) the degree of $R_l(\lambda)$ equals the number of poles (zeros) to be moved inside Γ.

REMARK 2.1. Notice that the latter also sheds some light on condition (iii). This condition restricts the possible solutions $R_1(\lambda)$ to those having *minimal* degree. Indeed, any solution $R_1(\lambda)$ to conditions (i) and (ii) will have degree at least equal to (1.2), as is easily seen from the above discussion.

The above recursive scheme thus allows us to generate factorizations of the type described in Section 1, for any given region Γ. Things can be different, though, when additional conditions are imposed on any of the factors, such as the condition of $R_1(\lambda)$ being all-pass. As shown e.g. in [2], this implies that the poles γ_i and zeros δ_i are each other's mirror image with respect to the unit circle ($\gamma_i = \delta_i^{-*}$) or to the $j\omega$-axis ($\gamma_i = -\delta_i^*$). Since either γ or δ has to cancel with a point outside Γ, the other will be inside Γ (for both the discrete-time and continuous-time cases), and thus this does not contradict our requirements. Moreover, it is shown in [2] that these special conditions can be incorporated in factors that satisfy the cancellation property of (2.3). The other types of factorizations discussed in Section 1 imply a relaxation of some of the standard conditions and thus cause no trouble, except that our algorithm will only generate a restricted (but nonempty) set of all possible solutions.

On the other hand when $R(\lambda)$ is real and one imposes the factors $R_1(\lambda)$ and $R_2(\lambda)$ to be real as well, problems will occur when $R(\lambda)$ has complex poles or zeros. When the region Γ is symmetric with respect to the real axis, such real factorizations do exist, but are not necessarily obtained by the above recursive scheme using first-degree factors, since some of these factors will be complex. On the other hand, when using real *second*-degree factors $C_i(\lambda)$ with complex conjugate pairs of poles and/or zeros, one can still completely follow the above reasoning and thus derive real factorizations recursively, but then using second-degree factors whenever needed. In [14] the general formulas are given for constructing such real second-degree canceling factors.

The major disadvantage of the above "transfer-function" approach is its complexity. First one has to compute the poles (or zeros) of $R(\lambda)$ in order to know which ones have to be canceled. Then one has to compute a partial-fraction expansion of $R(\lambda)$, and from the coefficient matrices of this expansion one derives certain vectors needed for the construction of the factors $C_i(\lambda)$ [a process that can be very complicated when $R(\lambda)$ has coinciding poles and zeros]. Moreover, after each pole/zero cancellation with a factor $C_i(\lambda)$, the expansion has to be updated. Finally, the process becomes even more involved when extracting real second-degree sections with complex conjugate poles and/or zeros, as is easily seen from the formulas given in [14]. This disadvantage suggests the use of another approach. Apparently, the most appealing methods for calculating poles and zeros use state-space models [9, 15]. Since one has a compact description of the system, one could as well try to solve the problem using this parametrization. The necessary and sufficient conditions of pole/zero cancellations using state-space models are derived in the next two sections and indeed turn out to be much simpler.

3. DISLOCATING POLES IN STATE SPACE

Here we first assume for simplicity that the matrices $R(\lambda)$, $R_1^{-1}(\lambda)$, and $R_2(\lambda)$ only have finite poles. These conditions will be removed later on, but can always be obtained via a first-degree conformal mapping $\lambda = (a\mu + b)/(c\mu + d)$ which does not affect the degree of rational matrices. The region $\Gamma(\lambda)$ of course has to be transformed accordingly to $\Gamma(\mu)$. Since we assume now (without loss of generality) that $R(\lambda)$ has no infinite poles, it has a realization quadruple $\{A, B, C, D\}$, i.e.,

$$R(\lambda) = C(\lambda I - A)^{-1}B + D, \tag{3.1}$$

which we will denote as

$$R(\lambda) \sim \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]. \tag{3.2}$$

Let also

$$R_1^{-1}(\lambda) \sim \left[\begin{array}{c|c} F & G \\ \hline H & J \end{array}\right]. \tag{3.3}$$

In order to construct a quadruple for the product $R_2(\lambda) = R_1^{-1}(\lambda)R(\lambda)$ we make use of the following lemma.

LEMMA 3.1. *Let*

$$\left[\begin{array}{c|c} F & G \\ \hline H & J \end{array}\right] \quad \textit{and} \quad \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

be realizations of two transfer matrices. Then a realization of the product of the two corresponding transfer matrices (in that order) is given by the constant matrix product

$$\left[\begin{array}{cc|c} F & 0 & G \\ 0 & I & 0 \\ \hline H & 0 & J \end{array}\right]\left[\begin{array}{cc|c} I & 0 & 0 \\ 0 & A & B \\ \hline 0 & C & D \end{array}\right] = \left[\begin{array}{cc|c} F & GC & GD \\ 0 & A & B \\ \hline H & JC & JD \end{array}\right]. \tag{3.4}$$

Proof. See Appendix. ∎

The right-hand side of (3.4) is often found in the literature [3], but it is rarely mentioned that it can be written as a product of two matrices as well (see also [6]). This is an elegant feature of working with realizations and will be exploited in the sequel. Notice also that these realizations need not be minimal.

Using this lemma, we now derive necessary and sufficient conditions for canceling the poles of a transfer function $R(\lambda)$. Let us choose a *minimal* realization (3.2) for $R(\lambda)$ where A is in (upper) Schur form and where the eigenvalues of A outside Γ are all grouped in the top left corner A_{11}. Let us assume that the blocks A_{ii} have dimensions $n_i \times n_i$ for $i = 1, 2$ (where $n = n_1 + n_2$):

$$R(\lambda) \sim \left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ 0 & A_{22} & B_2 \\ \hline C_1 & C_2 & D \end{array}\right] \tag{3.5}$$

(such a realization can always be obtained by updating (3.2) with a *unitary* state-space transformation [13, 19]). Then the canceling factor $R_1(\lambda)$ will have degree n_1 according to condition (iii). Let its minimal realization be given by (3.3), where F has order n_1. Following Lemma 3.1, a (nonminimal) realization for $R_2(\lambda)$ is then given by

$$\left[\begin{array}{ccc|c} F & & & G \\ & I_{n_1} & 0 & \\ & 0 & I_{n_2} & \\ \hline H & & & J \end{array}\right] \left[\begin{array}{ccc|c} I_{n_1} & & & \\ & A_{11} & A_{12} & B_1 \\ & 0 & A_{22} & B_2 \\ \hline & C_1 & C_2 & D \end{array}\right] = \left[\begin{array}{ccc|c} F & GC_1 & GC_2 & GD \\ & A_{11} & A_{12} & B_1 \\ & & A_{22} & B_2 \\ \hline H & JC_1 & JC_2 & JD \end{array}\right]. \tag{3.6}$$

This now has all its poles inside Γ iff the eigenvalues of A_{11} are either unobservable or uncontrollable, since A_{22} is Γ-stable by assumption and F is chosen to be Γ-stable (see Section 2). Since we assumed (3.5) to be a minimal realization, the eigenvalues of A_{11} are clearly controllable in (3.5) and hence also in (3.6). Let $\mathscr{X}$ be the invariant subspace of the state space of (3.6) corresponding to the eigenvalues of A_{11}. This space is uniquely defined, since the spectrum of A_{11} is disjoint from the rest of the poles of (3.6). A basis for $\mathscr{X}$ in the coordinate system of (3.6) is easily seen to be

$$\mathscr{X} = \left\langle \begin{bmatrix} X \\ I_{n_1} \\ 0 \end{bmatrix} \right\rangle, \tag{3.7}$$

where X is the (unique) solution of the Sylvester equation

$$XA_{11} - FX = GC_1. \tag{3.8}$$

Since this space must be unobservable, one has $[H, JC_1, JC_2]\mathscr{X} = 0$, or

$$HX + JC_1 = 0. \tag{3.9}$$

A reduced realization for $R_2(\lambda)$ is then obtained using the state-space transformation

$$T^{-1} = \begin{bmatrix} I_{n_1} & -X & 0 \\ & I_{n_1} & 0 \\ & & I_{n_2} \end{bmatrix}, \quad T = \begin{bmatrix} I_{n_1} & X & 0 \\ & I_{n_1} & 0 \\ & & I_{n_2} \end{bmatrix}. \tag{3.10}$$

Performed on (3.6) this yields, because of (3.8), (3.9),

$$R_2(\lambda) \sim \left[\begin{array}{ccc|c} F & 0 & GC_2 - XA_{12} & GD - XB_1 \\ 0 & A_{11} & A_{12} & B_1 \\ 0 & 0 & A_{22} & B_2 \\ \hline H & 0 & JC_2 & JD \end{array}\right]$$

$$\sim \left[\begin{array}{cc|c} F & GC_2 - XA_{12} & GD - XB_1 \\ 0 & A_{22} & B_2 \\ \hline H & JC_2 & JD \end{array}\right], \tag{3.11}$$

which again has state-space dimension $n_1 + n_2$. Summarizing, we have thus proved the following theorem.

THEOREM 3.1. *Let* $R(\lambda)$ *be a* $p\times m$ *rational matrix with a minimal realization of order* $n=n_1+n_2$ *as given in* (3.5), *where* $\Lambda(A_{22})\subset\Gamma$ *and* $\Lambda(A_{11})\subset\Gamma_c$, *the complement of* Γ. *Then* $R_1(\lambda)$ *and* $R_2(\lambda)$ *realized by*

$$R_1^{-1}(\lambda)\sim\left[\begin{array}{c|c} F & G \\ \hline H & J \end{array}\right], \tag{3.12a}$$

$$R_2(\lambda)\sim\left[\begin{array}{cc|c} F & GC_2-XA_{12} & GD-XB_1 \\ 0 & A_{22} & B_2 \\ \hline H & JC_2 & JD \end{array}\right] \tag{3.12b}$$

satisfy the required conditions of the factorization (1.1) *iff*:

(i) $\{F,G,H,J\}$ *represents a nonsingular transfer matrix*;
(ii) *the following equation is satisfied*:

$$\left[\begin{array}{c|c} F & G \\ \hline H & J \end{array}\right]\left[\begin{array}{c} X \\ \hline C_1 \end{array}\right]=\left[\begin{array}{c} X \\ \hline 0 \end{array}\right]A_{11}; \tag{3.13}$$

(iii) $\Lambda(F)\subset\Gamma$.

Proof. That (i) is equivalent to condition (i) of the factorization (1.1) is obvious. Then, assuming (i) holds, it is clear that (iii) is equivalent to the poles of (3.12a) lying in Γ. Condition (ii) is a rewriting of (3.8), (3.9) and is thus equivalent to the poles of (3.12b) lying in Γ also. Finally, the degree condition (1.2a) is automatically satisfied, since $R_1^{-1}(\lambda)$ has a realization of order n_1. ∎

The question now of course remains how to find matrices F, G, H, J, and X satisfying the conditions of Theorem 3.1. Although we know from the previous section that such a solution must exist, we would like to derive here a *constructive* proof which does not rely on the material of Section 2. For this, we first show that the nonsingularity of the system $\{F,G,H,J\}$ implies X is invertible.

LEMMA 3.2. *Let* (A_{11},C_1) *be observable, and let* $\{F,G,H,J\}$ *satisfy* (3.13). *Then* $\{F,G,H,J\}$ *represents a nonsingular system only if* X *in* (3.13) *is nonsingular*.

Proof. We assume that $\{F,G,H,J\}$ represents a nonsingular system. Using (3.8) and (3.9) one easily derives

$$\left[\begin{array}{c|c} F-\lambda I & G \\ \hline H & J \end{array}\right]\left[\begin{array}{c} X \\ \hline C_1 \end{array}\right]=\left[\begin{array}{c} X \\ \hline 0 \end{array}\right](A_{11}-\lambda I) \tag{3.14}$$

for any value of λ. Choosing a point λ which is neither an eigenvalue of A_{11} nor a zero of $\{F,G,H,J\}$, we have that

$$A_{11}-\lambda I \quad \text{and} \quad S(\lambda)=\left[\begin{array}{c|c} F-\lambda I & G \\ \hline H & J \end{array}\right] \tag{3.15}$$

are both nonsingular [20]. Hence from (3.14),

$$X \quad \text{and} \quad \left[\begin{array}{c} X \\ \hline C_1 \end{array}\right]$$

must have the same rank and also the same null space, which we represent by $\mathscr{N}$. Thus we have

$$\mathscr{N}=\operatorname{Ker} X, \qquad C_1\mathscr{N}=0. \tag{3.16}$$

Applying this null space to (3.14) for $\lambda=0$, we find

$$XA_{11}\mathscr{N}=0, \qquad \text{or} \quad A_{11}\mathscr{N}\subset\mathscr{N}. \tag{3.17}$$

But then $\mathscr{N}$ must be an unobservable subspace of (A_{11},C_1), which contradicts the assumptions. ■

From this one also easily obtains the following result, which is to be expected from the discussion of Section 2.

COROLLARY 3.1. *The eigenvalues of A_{11}—i.e. the poles of $R(\lambda)$—to be canceled are the zeros of the nonsingular system $\{F,G,H,J\}$.*

Proof. Since X is nonsingular, it follows from (3.14) that $S(\lambda)$ and $A_{11}-\lambda I$ lose rank together at the eigenvalues of A_{11}, and the zeros of $S(\lambda)$ are indeed those of its associated transfer function $R_1^{-1}(\lambda)$ [20]. ■

Since X is now invertible, it can be "absorbed" into the quadruple $\{F,G,H,J\}$ as a state-space transformation. Putting

$$\{\hat{F},\hat{G},\hat{H},\hat{J}\}=\{X^{-1}FX,X^{-1}G,HX,J\},\tag{3.18}$$

it follows indeed from (3.13) that we are looking for a system $\{\hat{F},\hat{G},\hat{H},\hat{J}\}$ satisfying

$$\left[\begin{array}{c|c}\hat{F} & \hat{G}\\ \hline \hat{H} & \hat{J}\end{array}\right]\left[\begin{array}{c} I\\ \hline C_1\end{array}\right]=\left[\begin{array}{c} A_{11}\\ \hline 0\end{array}\right],\tag{3.19}$$

and this is easily solved via the following procedure:

ALGORITHM 3.1.

Step 1. Determine $\hat{F}$, $\hat{G}$, with $\Lambda(\hat{F})\subset\Gamma$ by solving the pole placement problem

$$\hat{F}=A_{11}-\hat{G}C_1.\tag{3.20}$$

This has always a solution, since (A_{11},C_1) is observable.

Step 2. Determine $[\hat{H}\,|\,\hat{J}]^T$ as any basis for the null space of $[I_{n_1}\,|\,C_1^T]$, i.e.

$$[\,\hat{H}\mid\hat{J}\,]=M[\,-C_1\mid I\,]\tag{3.21}$$

for an arbitrary *invertible* M.

Since the non-singularity of the system $\{\hat{F},\hat{G},\hat{H},\hat{J}\}$ immediately follows from the invertibility of $\hat{J}=M$, all conditions of Theorem 3.1 are clearly satisfied. We have thus derived here a constructive proof that the undesired poles of a transfer function $R(\lambda)$ can be canceled by a nonsingular transfer function $R_1^{-1}(\lambda)$ whose degree equals the number of poles to be canceled (here n_1), whose zeros will indeed be those unwanted poles, and whose poles can be chosen arbitrarily in Γ. Moreover, it is interesting that here we did not need a recursive argument, as was the case in the work of Belevitch [2] and of Dewilde and Vandewalle [5]. The recursiveness, though, when also used in state space, will yield a very simple and elegant algorithm, as is shown in Section 5.

4. DISLOCATING ZEROS IN STATE SPACE

Here we consider the case of canceling the undesired zeros of $R(\lambda)$ that are outside a specified region Γ by the poles of $R_1^{-1}(\lambda)$. As before, let

$$R_1^{-1}(\lambda) \sim \left[\begin{array}{c|c} F & G \\ \hline H & J \end{array}\right] \tag{4.1}$$

be a minimal realization of the factor $R_1^{-1}(\lambda)$, and let n_1 be its state-space dimension. Then according to Lemma 3.1, the product $R_2(\lambda) = R_1^{-1}(\lambda)R(\lambda)$ is realized by

$$R_2(\lambda) \sim \left[\begin{array}{cc|c} F & GC & GD \\ 0 & A & B \\ \hline H & JC & JD \end{array}\right]. \tag{4.2}$$

We would like now the poles of $R_1^{-1}(\lambda)$—i.e. the spectrum of F—to cancel with the undesired zeros of $R(\lambda)$, which are assumed to be n_1 in number. Since the pair (F, H) is observable in (4.1), it is also observable in (4.2), and the spectrum of F in (4.2) must thus be uncontrollable. Since (A, B) is controllable in (4.2), the controllable subspace $\mathscr{X}$ of the realization (4.2) must be of the form [4]

$$\mathscr{X} = \left\langle \begin{bmatrix} X \\ I_n \end{bmatrix} \right\rangle, \tag{4.3}$$

where X satisfies [4]

$$XA - FX = GC, \qquad XB = GD. \tag{4.4}$$

The state-space transformation

$$T^{-1} = \begin{bmatrix} I_{n_1} & -X \\ 0 & I_n \end{bmatrix}, \qquad T = \begin{bmatrix} I_{n_1} & X \\ 0 & I_n \end{bmatrix}, \tag{4.5}$$

applied to (4.2), then yields

$$R_2(\lambda) \sim \left[\begin{array}{cc|c} F & 0 & 0 \\ 0 & A & B \\ \hline H & JC + HX & JD \end{array}\right], \tag{4.6}$$

where we have used (4.4) for the zero blocks in the top row of (4.6).

Unfortunately, the condition (4.4) is only a *necessary* condition for the pole-zero cancellation to occur. This is illustrated by the following example. Let

$$R(\lambda)=\begin{bmatrix}\dfrac{\lambda-2}{\lambda-1}\\ 0\end{bmatrix}\sim\left[\begin{array}{c|c}1&-1\\ \hline 1&1\\ 0&0\end{array}\right]=\left[\begin{array}{c|c}A&B\\ \hline C&D\end{array}\right]\tag{4.7}$$

and

$$R_1^{-1}(\lambda)=\begin{bmatrix}1&0\\ 0&\dfrac{\lambda}{\lambda-2}\end{bmatrix}\sim\left[\begin{array}{c|cc}2&0&1\\ \hline 0&1&0\\ 2&0&1\end{array}\right]=\left[\begin{array}{c|c}F&G\\ \hline H&J\end{array}\right].\tag{4.8}$$

Clearly $R(\lambda)$ has a zero at $\lambda=2$, which we hope to dislocate by the factor $R_1^{-1}(\lambda)$, which has a pole at $\lambda=2$ and a zero at $\lambda=0$. Now following (4.2), we have

$$R_2(\lambda)=R_1^{-1}(\lambda)\cdot R(\lambda)\sim\left[\begin{array}{cc|c}2&0&0\\ 0&1&-1\\ \hline 0&1&1\\ 2&0&0\end{array}\right],\tag{4.9}$$

and the condition (4.4) is clearly satisfied for $X=[0\ \ 0]$. The eigenvalue $\lambda=2$ of F indeed is uncontrollable in (4.9), but it has not canceled the zero of $R(\lambda)$, since $R_2(\lambda)$ is again equal to $R(\lambda)$. No zero dislocation from $\lambda=2$ to $\lambda=0$ has taken place here.

Apparently some additional conditions have to be imposed on X in order to enforce a zero dislocation. In order to do so, one has to explicitly specify first which zeros of $R(\lambda)$ have to be dislocated, as was done in (3.5) for the poles. A similar decomposition for the zeros is now derived. For this we first recall the following lemmas, proved in [18]:

LEMMA 4.1. *Let U be any invertible transformation such that*

$$\left[\begin{array}{c|c}\lambda I-A&B\\ \hline -C&D\end{array}\right]U=\left[\begin{array}{c|c}\lambda\hat{E}-\hat{A}&\lambda\hat{B}-\hat{F}\\ \hline 0&\hat{D}\end{array}\right],\tag{4.10}$$

where $\hat{D}$ has linearly independent columns. Then the generalized eigenvalues of $\lambda\hat{E}-\hat{A}$ are the zeros of $R(\lambda)=C(\lambda I-A)^{-1}B+D$. ■

Notice that, since $[-C\,|\,D]U=[0\,|\,\hat{D}]$, rank $\hat{D}$ equals rank$[-C\,|\,D]$ and the numbers of columns $\hat{m}$ of $\hat{D}$ and m of D may thus differ [which is why we used a dashed line in (4.10)]. We also remark that in practice one uses unitary matrices U, which yields a numerically stable construction of what was called the *zero pencil* $\lambda\hat{E}-\hat{A}$ in [18]. The use of unitary matrices is also maintained in the following lemma, leading to the separation between two parts of the spectrum of $\lambda\hat{E}-\hat{A}$.

LEMMA 4.2. *Let $\lambda\hat{E}-\hat{A}$ be an arbitrary singular pencil with spectrum $\Lambda(\hat{E},\hat{A})$—i.e., $\Lambda(\hat{E},\hat{A})$ is the set of generalized eigenvalues of $\lambda\hat{E}-\hat{A}$. Then for any complementary regions Γ and Γ_c of the complex plane separating $\Lambda(\hat{E},\hat{A})$ into two disjoint parts Λ_Γ and Λ_{Γ_c}, there exist unitary transformations Q and Z such that*

$$Q^*(\lambda\hat{E}-\hat{A})Z=\begin{bmatrix}\lambda\hat{E}_{11}-\hat{A}_{11} & \lambda\hat{E}_{12}-\hat{A}_{12} & \lambda\hat{E}_{13}-\hat{A}_{13}\\ 0 & \lambda\hat{E}_{22}-\hat{A}_{22} & \lambda\hat{E}_{23}-\hat{A}_{23}\\ 0 & 0 & \lambda\hat{E}_{33}-\hat{A}_{33}\end{bmatrix} \quad (4.11)$$

whereby

(i) $\Lambda(\hat{E}_{11},\hat{A}_{11})=\Lambda_\Gamma$, $\Lambda(\hat{E}_{22},\hat{A}_{22})=\Lambda_{\Gamma_c}$, $\Lambda(\hat{E}_{33},\hat{A}_{33})=\varnothing$;
(ii) $\lambda\hat{E}_{11}-\hat{A}_{11}$ *is right invertible for* $\lambda\in\Gamma_c$;
(iii) $\lambda\hat{E}_{22}-\hat{A}_{22}$ *is invertible for* $\lambda\in\Gamma$;
(iv) $\lambda\hat{E}_{33}-\hat{A}_{33}$ *is left invertible.*

Lemma 4.2 essentially says that there exists a (generalized) block Schur decomposition (4.11) with the generalized eigenvalues of $\lambda\hat{E}-\hat{A}$ inside Γ gathered in $\lambda\hat{E}_{11}-\hat{A}_{11}$, and the remaining ones gathered in $\lambda\hat{E}_{22}-\hat{A}_{22}$ (see [16, 1] for algorithms). The additional conditions (ii), (iii), (iv) guarantee that such a decomposition is essentially unique, and they will play an important role in the sequel. These two lemmas now lead to the following theorem.

THEOREM 4.1. *One can always update a minimal realization* (3.2) *by a unitary state-space transformation Q such that its zero pencil* (4.10) *is automatically in generalized Schur form* (4.11), *i.e.*,

$$\left[\begin{array}{c|c} Q^*(\lambda I-A)Q & Q^*B\\ \hline -CQ & D\end{array}\right]=\left[\begin{array}{ccc|c}\lambda I_{n_1}-A_{11} & -A_{12} & -A_{13} & B_1\\ -A_{21} & \lambda I_{n_2}-A_{22} & -A_{23} & B_2\\ -A_{31} & -A_{32} & \lambda I_{n_3}-A_{33} & B_3\\ \hline -C_1 & -C_2 & -C_3 & D\end{array}\right]$$

(4.12)

and

$$\left[\begin{array}{ccc|c} \lambda_{I_{n1}} - A_{11} & -A_{12} & -A_{13} & B_1 \\ -A_{21} & \lambda I_{n_2} - A_{22} & -A_{23} & B_2 \\ -A_{31} & -A_{32} & \lambda I_{n_3} - A_{33} & B_3 \\ \hline -C_1 & -C_2 & -C_3 & D \end{array}\right] U$$

$$= \begin{array}{c} \left[\begin{array}{ccc|c} \lambda \hat{E}_{11} - \hat{A}_{11} & \lambda \hat{E}_{12} - \hat{A}_{12} & \lambda \hat{E}_{13} - \hat{A}_{13} & \lambda \hat{F}_1 - \hat{B}_1 \\ 0 & \lambda \hat{E}_{22} - \hat{A}_{22} & \lambda \hat{E}_{23} - \hat{A}_{23} & \lambda \hat{F}_2 - \hat{B}_2 \\ 0 & 0 & \lambda \hat{E}_{33} - \hat{A}_{33} & \lambda \hat{F}_3 - \hat{B}_3 \\ \hline \underbrace{\quad 0 \quad}_{\hat{n}_1} & \underbrace{\quad 0 \quad}_{\hat{n}_2} & \underbrace{\quad 0 \quad}_{\hat{n}_3} & \underbrace{\quad \hat{D} \quad}_{\hat{m}} \end{array}\right] \end{array} \begin{array}{l} \}n_1 \\ \}n_2 \\ \}n_3 \\ \}p \\ \end{array} \tag{4.13}$$

where $\lambda \hat{E}_{11} - \hat{A}_{11}$, $\lambda \hat{E}_{22} - \hat{A}_{22}$, *and* $\lambda \hat{E}_{33} - \hat{A}_{33}$ *satisfy conditions* (i), (ii), (iii), (iv) *of Lemma* 3.2 (*hence* $\hat{n}_1 \geqslant n_1$, $\hat{n}_2 = n_2$, $\hat{n}_3 < n_3$).

Proof. Let the zero pencil $\lambda \hat{E} - \hat{A}$ of a given realization (3.2) be as in (4.10), and let (4.11) be its required Schur form. Then we have

$$\left[\begin{array}{c|c} Q^* & \\ \hline & I_p \end{array}\right] \left[\begin{array}{c|c} \lambda I - A & B \\ \hline -C & D \end{array}\right] U \left[\begin{array}{c|c} Z & \\ \hline & I_{\hat{m}} \end{array}\right]$$

$$= \left[\begin{array}{ccc|c} \lambda \hat{E}_{11} - \hat{A}_{11} & \lambda \hat{E}_{12} - \hat{A}_{12} & \lambda \hat{E}_{13} - \hat{A}_{13} & \lambda \hat{F}_1 - \hat{B}_1 \\ 0 & \lambda \hat{E}_{22} - \hat{A}_{22} & \lambda \hat{E}_{23} - A_{23} & \lambda \hat{F}_2 - \hat{B}_2 \\ 0 & 0 & \lambda \hat{E}_{33} - \hat{A}_{33} & \lambda \hat{F}_3 - \hat{B}_3 \\ \hline 0 & 0 & 0 & \hat{D} \end{array}\right], \tag{4.14}$$

where $\hat{m}$ is the number of columns of $\hat{D}$. From this we then obtain (4.12), (4.13) by putting

$$U \stackrel{\text{def}}{=} \left[\begin{array}{c|c} Q^* & \\ \hline & I_p \end{array}\right] U \left[\begin{array}{c|c} Z & \\ \hline & I_{\hat{m}} \end{array}\right]. \tag{4.15}$$

This proves that Q as given in (4.11) is the required unitary state-space transformation. ∎

Without loss of generality we can thus assume that our state-space realization is in a form satisfying (4.13) where $\lambda \hat{E}_{22} - \hat{A}_{22}$ contains the zeros of $R(\lambda)$ to be dislocated. This form is to be considered an analogue to (3.5), now isolating the *zeros* to be cancelled in a separate block $\lambda \hat{E}_{22} - \hat{A}_{22}$. The fact that now there is a third block $\lambda \hat{E}_{33} - \hat{A}_{33}$ is due to the possible singularity of the pencil (4.12)—and hence of $R(\lambda)$ (see [18]). Indeed, when $R(\lambda)$ happens to be right invertible, so will (4.12), and the block $\lambda \hat{E}_{33} - \hat{A}_{33}$ will vanish [18].

We are now ready to formulate a theorem on zero dislocation analogously to Theorem 3.1.

THEOREM 4.2. *Let $R(\lambda)$ be a $p \times m$ rational matrix with a minimal realization of order $n = n_1 + n_2 + n_3$ as given in* (4.12), (4.13), *where $\Lambda(\hat{E}_{11}, \hat{A}_{11}) \subset \Gamma$ and $\Lambda(\hat{E}_{22}, \hat{A}_{22}) \subset \Gamma_c$, the complement of Γ, and $\Lambda(\hat{E}_{33}, \hat{A}_{33}) = \varnothing$. Then $R_1(\lambda)$ and $R_2(\lambda)$ realized by*

$$R_1^{-1}(\lambda) \sim \left[\begin{array}{c|c} F & G \\ \hline H & J \end{array}\right], \qquad R_2(\lambda) \sim \left[\begin{array}{c|c} \lambda I - A & B \\ \hline -JC - HX & JD \end{array}\right] \quad (4.16\text{a,b})$$

satisfy the required conditions of the factorization (1.1) *iff*:

(i) *$\{F, G, H, J\}$ represents a nonsingular transfer matrix*;

(ii) *the following equation is satisfied for the $n_2 \times (n_1 + n_2 + n_3)$ matrix $X = [0 \;\; X_2 \;\; X_3]$*:

$$[X \quad G]\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = F[X \quad 0], \qquad (4.17)$$

with X_2 invertible;

(iii) *the zeros of $\{F, G, H, J\}$ lie in Γ.*

Proof. That (i) is equivalent to condition (i) of the factorization (1.1) is obvious. Then, assuming (i) holds, (iii) is necessary and sufficient for (4.16a) to have its zeros in Γ. If we can also prove that, under the assumption that (i) and (iii) hold, (ii) is necessary and sufficient for (4.16b) to have its zeros in Γ, then the theorem is proved, since the degree condition (1.2b) is satisfied by construction.

We now first prove the *necessity* of (ii). Since $R_1(\lambda)$ is invertible, the zeros of $R_2(\lambda)$ can only be dislocated by a cancellation with the poles of

$R_1^{-1}(\lambda)$. Hence (4.4), (4.6) must hold, or equivalently, condition (ii), without any specification for X, must hold. But—as mentioned in (4.7), (4.8)—this is not enough, since for some X the zeros of $R_2(\lambda)$ *may* not be dislocated despite the cancellation of the poles of $R_1^{-1}(\lambda)$. For deriving the additional condition on X, we partition X, U, and the product $[X\,|\,0]U$ conformably with (4.12), (4.13). Using (4.13) for the structure of U, we thus have

$$X \stackrel{\text{def}}{=} \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}, \tag{4.18}$$

$$U = \begin{bmatrix} \hat{E}_{11} & \hat{E}_{12} & \hat{E}_{13} & \hat{F}_1 \\ 0 & \hat{E}_{22} & \hat{E}_{23} & \hat{F}_2 \\ 0 & 0 & \hat{E}_{33} & \hat{F}_3 \\ * & * & * & * \end{bmatrix}, \tag{4.19}$$

$$[X|0]U \stackrel{\text{def}}{=} \begin{bmatrix} \hat{X}_1 & \hat{X}_2 & \hat{X}_3 & | & \hat{X}_4 \end{bmatrix}. \tag{4.20}$$

Multiplying the columns of (4.16b) by U and using (4.20) yields

$$\left[\begin{array}{c|c} \lambda I - A & B \\ \hline -JC - HX & JD \end{array}\right] U$$

$$= \left[\begin{array}{ccc|c} \lambda\hat{E}_{11} - \hat{A}_{11} & \lambda\hat{E}_{12} - \hat{A}_{12} & \lambda\hat{E}_{13} - \hat{A}_{13} & \lambda\hat{F}_1 - \hat{B}_1 \\ 0 & \lambda\hat{E}_{22} - \hat{A}_{22} & \lambda\hat{E}_{23} - \hat{A}_{23} & \lambda\hat{F}_2 - \hat{B}_2 \\ 0 & 0 & \lambda\hat{E}_{33} - \hat{A}_{33} & \lambda\hat{F}_3 - \hat{B}_3 \\ \hline -H\hat{X}_1 & -H\hat{X}_2 & -H\hat{X}_3 & J\hat{D} - H\hat{X}_4 \end{array}\right]. \tag{4.21}$$

Similarly, the condition (4.17) becomes, after multiplying both sides by U and using (4.18), (4.20),

$$\begin{bmatrix} X_1 & X_2 & X_3 & G \end{bmatrix} \left[\begin{array}{ccc|c} \lambda\hat{E}_{11} - \hat{A}_{11} & \lambda\hat{E}_{12} - \hat{A}_{12} & \lambda\hat{E}_{13} - \hat{A}_{13} & \lambda\hat{F}_1 - \hat{B}_1 \\ 0 & \lambda\hat{E}_{22} - \hat{A}_{22} & \lambda\hat{E}_{23} - \hat{A}_{23} & \lambda\hat{F}_2 - \hat{B}_2 \\ 0 & 0 & \lambda\hat{E}_{33} - \hat{A}_{33} & \lambda\hat{F}_3 - \hat{B}_3 \\ \hline 0 & 0 & 0 & \hat{D} \end{array}\right]$$

$$= (\lambda I_{n_2} - F)\begin{bmatrix} \hat{X}_1 & \hat{X}_2 & \hat{X}_3 & \hat{X}_4 \end{bmatrix}. \tag{4.22}$$

From this we extract $X_1(\lambda \hat{E}_{11} - \hat{A}_{11}) = (\lambda I - F)\hat{X}_1$. Now $\lambda I - F$ and $\lambda \hat{E}_{11} - \hat{A}_{11}$ are left and right invertible, respectively, and have no common spectrum by assumption. Therefore we have $X_1 = 0$, $\hat{X}_1 = 0$ because of Lemma A.2 of the Appendix. Now consider $\hat{X}_2$, which equals $X_2 \hat{E}_{22}$. First note that $\hat{E}_{22}$ must be invertible, since we assumed that the zeros to be dislocated were *finite*. We now prove that X_2 (and hence $\hat{X}_2$) must be invertible if $R_2(\lambda)$ has no zeros outside Γ. For this, consider the pencil (4.21), which, after multiplying the second block column by $\hat{E}_{22}^{-1}$, looks like

$$\left[\begin{array}{ccc|c} \lambda \hat{E}_{11} - \hat{A}_{11} & * & * & * \\ 0 & \lambda I - \hat{A}_{22}\hat{E}_{22}^{-1} & \lambda \hat{E}_{23} - \hat{A}_{23} & \lambda \hat{F}_2 - \hat{B}_2 \\ 0 & 0 & \lambda \hat{E}_{33} - \hat{A}_{33} & \lambda \hat{F}_3 - \hat{B}_3 \\ \hline 0 & -HX_2 & -H\hat{X}_3 & J\hat{D} - H\hat{X}_4 \end{array}\right]. \tag{4.23}$$

Because of Lemma 4.1, the finite zeros of $R_2(\lambda)$ are the finite eigenvalues of $\lambda \hat{E} - \hat{A}$ and hence also of (4.23). But this pencil will clearly drop rank at any unobservable mode of the pair $(\hat{A}_{22}\hat{E}_{22}^{-1}, -HX_2)$, which contradicts the singularity of X_2. Indeed, let $\mathcal{N}$ be the kernel of X_2; then from $X_2(\lambda \hat{E}_{22} - \hat{A}_{22}) = (\lambda I - F)\hat{X}_2$ [extracted from (4.22)] one also finds that $\mathcal{N}$ is an unobservable space of $(\hat{A}_{22}\hat{E}_{22}^{-1}, -HX_2)$.

Next, we prove the *sufficiency* of (ii). We thus have to prove that $R_2(\lambda)$ has no spectrum left outside Γ. Using (4.23) and the invertibility of X_2, we define

$$Y_3 \overset{\text{def}}{=} X_2^{-1}X_3, \qquad \hat{Y}_3 \overset{\text{def}}{=} X_2^{-1}\hat{X}_3 = \hat{E}_{23} + Y_3\hat{E}_{33}, \quad \text{and}$$

$$\hat{Y}_4 \overset{\text{def}}{=} X_2^{-1}\hat{X}_4 = \hat{F}_2 + Y_3\hat{F}_3.$$

From (4.22) one then obtains the identity

$$\begin{bmatrix} I & Y_3 & X_2^{-1}G \end{bmatrix} \left[\begin{array}{cc|c} \lambda \hat{E}_{22} - \hat{A}_{22} & \lambda \hat{E}_{23} - \hat{A}_{23} & \lambda \hat{F}_2 - \hat{B}_2 \\ 0 & \lambda \hat{E}_{33} - \hat{A}_{33} & \lambda \hat{F}_3 - \hat{B}_3 \\ \hline 0 & 0 & \hat{D} \end{array}\right]$$

$$= (\lambda I - X_2^{-1}FX_2)\begin{bmatrix} \hat{E}_{22} & \hat{Y}_3 & \hat{Y}_4 \end{bmatrix}, \tag{4.24}$$

which can be used to transform (4.23) to

$$\left[\begin{array}{ccc|c} \lambda\hat{E}_{11}-\hat{A}_{11} & * & * & * \\ 0 & X_2^{-1}(\lambda I - F)X_2 & 0 & X_2^{-1}G\hat{D} \\ 0 & 0 & \lambda\hat{E}_{33}-\hat{A}_{33} & * \\ \hline 0 & -HX_2 & 0 & J\hat{D} \end{array}\right]. \tag{4.25}$$

From this one now clearly sees that the zeros of $R_2(\lambda)$ are the union of $\Lambda(\hat{E}_{11},\hat{A}_{11})$ ($\subset\Gamma$), of $\Lambda(\hat{E}_{33},\hat{A}_{33})$ ($=\varnothing$), and of the *zeros* of $\{F,G\hat{D},H,J\hat{D}\}$ ($\subset\Gamma$). This thus completes the proof. ■

For the construction of a solution to the above theorem, we can again "absorb" the invertible factor X_2 into the quadruple $\{F,G,H,J\}$ as a state-space transformation. Putting

$$\{\hat{F},\hat{G},\hat{H},\hat{J}\} = \{X_2^{-1}FX_2, X_2^{-1}G, HX_2, J\}, \tag{4.26}$$

it then follows that we are looking for a system $\{\hat{F},\hat{G},\hat{H},\hat{J}\}$ satisfying

$$\begin{bmatrix} I & Y_3 & \hat{G} \end{bmatrix}\left[\begin{array}{cc|c} \lambda\hat{E}_{22}-\hat{A}_{22} & \lambda\hat{E}_{23}-\hat{A}_{23} & \lambda\hat{F}_2-\hat{B}_2 \\ 0 & \lambda\hat{E}_{33}-\hat{A}_{33} & \lambda\hat{F}_3-\hat{B}_3 \\ \hline 0 & 0 & \hat{D} \end{array}\right]$$

$$= (\lambda I - \hat{F})\begin{bmatrix} \hat{E}_{22} & \hat{Y}_3 & \hat{Y}_4 \end{bmatrix}, \tag{4.27}$$

and this is easily solved via the following procedure:

ALGORITHM 4.1.

Step 1. Put $\hat{F} = \hat{A}_{22}\hat{E}_{22}^{-1}$.

Step 2. Solve

$$Y_3(\lambda\hat{E}_{33}-\hat{A}_{33}) - (\lambda I - \hat{F})\hat{Y}_3 = -(\lambda\hat{E}_{23}-\hat{A}_{23}) \tag{4.28}$$

for Y_3 and $\hat{Y}_3$. This has always a solution according to Lemma A.3 of the Appendix.

Step 3. Put $\hat{Y}_4 = \hat{F}_2 + Y_3\hat{F}_3$, and solve for $\hat{G}$ from

$$\hat{G}\hat{D} = \hat{B}_2 + Y_3\hat{B}_3 - \hat{F}\hat{Y}_4. \tag{4.29}$$

Step 4. Choose $\hat{H}$, $\hat{J}$ such that the zeros of $\{\hat{F},\hat{G},\hat{H},\hat{J}\}$ lie in Γ. For *finite* zeros this is a pole placement problem, since then $\hat{J}$ is invertible. Choose then K such that $\Lambda(\hat{F}+\hat{G}K)\subset\Gamma$, and then solve for $[\hat{H}\ \hat{J}]^T$ as any basis for the null space of $[I_{n_2}\,|\,K^T]$, i.e.

$$[\hat{H}\mid\hat{J}] = M[-K\mid I] \tag{4.30}$$

for an arbitrary *invertible* M. (For *infinite* zeros the solution is more involved but can be constructed also—see concluding remarks.)

5. RECURSIVE POLE DISLOCATIONS IN STATE SPACE

The method for dislocating poles in state space described in Section 3 has the advantage of dealing with constant matrices only and of solving the dislocation problem for arbitrary degree n_1 at once. But the method has also a number of drawbacks:

(1) the additional conditions on the factor $R_1^{-1}(\lambda)$ mentioned in the introduction are hard to add on to condition (i) of Theorem 3.1;

(2) the choice of the state-space coordinate system for the system $\{F,G,H,J\}$ is not apparent;

(3) the state-space transformation (3.10), depending on the matrix X, ought to be avoided when it is badly conditioned (i.e. when X has a large norm);

(4) the system matrix dimensions may be significantly enlarged while constructing $\{F,G,H,J\}$ via (3.6).

In this section we show that by returning to the idea of recursiveness presented in Section 2, one can easily include the special conditions on the factor $R_1(\lambda)$ in Theorem 3.1 and at the same time avoid the above drawbacks. For a recursive solution of the problem in state space, we want to dislocate one "bad" eigenvalue of $R(\lambda)$ at a time. For this we thus consider a realization (3.5) where $n_1 = 1$. The "updating" transformation for obtaining such a realization (3.5) with $n_1 = 1$ is known to require only $O(n^2)$ operations [13]. We show that one can then give explicit formulas for the construction in Algorithm 3.1, while at the same time satisfying some additional constraints on $R_1(\lambda)$.

Notice that $R_1(\lambda)$ is now of first degree and thus (3.20) is a scalar condition. This suggests that one can exploit the degrees of freedom still present in Algorithm 3.1. Indeed, we show below that the realization of $R_1^{-1}(\lambda) \sim \{F, G, H, J\}$ can be chosen in the following form:

$$\left[\begin{array}{c|c} F & G \\ \hline H & J \end{array}\right] = \left[\begin{array}{c|cccc} f & g & 0 & \cdots & 0 \\ \hline h & j & & & \\ 0 & & & & \\ \vdots & & & I_{p-1} & \\ 0 & & & & \end{array}\right] \left[\begin{array}{c|c} 1 & \\ \hline & V \end{array}\right], \tag{5.1}$$

where V is a unitary transformation (i.e., $VV^* = I_m$) satisfying

$$V^* C_1 = \begin{bmatrix} c \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{5.2}$$

where c is nonzero (otherwise a_{11} would be unobservable) and can be chosen real. Using this form, let us now see how to satisfy all conditions of Theorem 3.1. The transfer function $R_1^{-1}(\lambda)$ of (5.1) is given by

$$R_1^{-1}(\lambda) = \left[\begin{array}{c|c} r(\lambda) & \\ \hline & I_{p-1} \end{array}\right] V, \qquad \text{with} \qquad \begin{aligned} r(\lambda) &= j + h(\lambda - f)^{-1} g \\ &= \frac{j\lambda + hg - jf}{\lambda - f}, \end{aligned} \tag{5.3}$$

which is nonsingular and of degree 1 iff $hg \neq 0$. Because of (5.1), (5.2), condition (ii) simplifies to

$$\left[\begin{array}{c|c} f-a_{11} & g \\ \hline h & j \end{array}\right]\left[\begin{array}{c} x \\ \hline c \end{array}\right]=0, \tag{5.4}$$

where x and a_{11} are now also scalar. Finally, condition (iii) requires that $f \in \Gamma$, which is again a scalar condition. Other special conditions to be imposed on $R_1^{-1}(\lambda)$ for the applications of Section 1 are obtained by concentrating on the *scalar* function $r(\lambda)$ only:

(1) For coprime factorizations and polynomial factorizations one has to work in a transformed variable, say $\lambda = 1/\mu$. Then $R_1^{-1}(1/\mu)$ is polynomial in μ if $f=0$, while $R_1(\mu)$ is polynomial in μ if $hg = jf$. In both cases one may take $|g| = |h|$ and either one of them real. The corresponding scalar function $r(1/\mu)$ is then equal to $j - gh\mu$ or $j/(1\text{-}\mu f)$, respectively.

(2) For an all-pass factor $R_1^{-1}(\lambda)$ in the variable $s = \lambda$ one takes $j=1$ and $g = h = \sqrt{\bar{f}+f}$. The corresponding scalar function $r(s)$ is then equal to $(s+\bar{f})/(s-f)$.

(3) For an inner factor $R_1^{-1}(\lambda)$ in the variable $z=\lambda$ one takes $j=\bar{f}$ and $g=-h=\sqrt{1-f\bar{f}}$. The corresponding scalar function $r(z)$ is then equal to $(\bar{f}z-1)/(z-f)$.

REMARK 5.1. Notice that the above realizations $\{f,g,h,j\}$ of the scalar functions $r(\mu)$, $r(s)$, and $r(z)$ are in balanced form, since for each of them $g^2 = h^2$. This choice of realization (which for a first-degree factor only consists of a scaling of g and h) is recommended for numerical reasons. For the case of an inner factor $r(z)$, it results e.g. in a realization where the compound matrix

$$\begin{bmatrix} f & g \\ h & j \end{bmatrix} = \begin{bmatrix} f & g \\ -g & \bar{f} \end{bmatrix}$$

is unitary, since $|f|^2 + g^2 = 1$. The overall state-space model (5.1) is then also unitary, which has good numerical consequences for the construction given in (3.4).

We now show that these conditions are easy to satisfy for each of the above restrictions. Without loss of generality, we illustrate this for the case of inner factors only. Let a_{11} and c be given; then one proceeds as follows. The pole a_{11} must be cancelled by a zero in $r(z)$, which completely fixes the factor $(f = 1/\bar{a}_{11})$. Since by assumption $a_{11} \in \Gamma_c$, we automatically have $f \in \Gamma$. Finally, the scalar x is computed from (5.4) and is then used to construct the updating state-space transformation (3.10) and then compute (3.11). Together with (5.1), (5.2), this then satisfies all required conditions. It is easily seen that for other types of first-degree factors the same reasoning goes through. This thus circumvents the two first drawbacks mentioned above.

Let us now turn to the last two drawbacks. For this we first define

$$V^*\begin{bmatrix} C_2 & D \end{bmatrix} \stackrel{\text{def}}{=} \begin{bmatrix} C_2^+ & D^+ \\ C_2^- & D^- \end{bmatrix} \begin{matrix} \}1 \\ \}p-1 \end{matrix} \tag{5.5}$$

and then rewrite (3.6) for the particular choice (5.1), (5.2) made here:

$$\left[\begin{array}{cccc} f & gc & gC_2^+ & gD^+ \\ & a_{11} & A_{12} & B_1 \\ & & A_{22} & B_2 \\ \hline h & jc & jC_2^+ & jD^+ \\ 0 & 0 & & \\ \vdots & \vdots & C_2^- & D^- \\ 0 & 0 & & \end{array}\right]. \tag{5.6}$$

Since x may be large, one may be forced to apply a nonunitary state-space transformation that is poorly conditioned and likely to induce rounding errors of unacceptable size. This can be avoided by replacing (3.10) with the unitary state-space transformation

$$T = \begin{bmatrix} \gamma & \sigma & 0 \\ -\sigma & \bar{\gamma} & 0 \\ & & I_{n_2} \end{bmatrix}, \tag{5.7}$$

which has the effect of "swapping" f and a_{11} in the triangular form (5.6).

Such a Givens transformation always exists, and the coefficients γ and σ are determined via

$$\gamma = \frac{jc}{\sqrt{g^2 + j^2c^2}},$$

$$\sigma = \frac{h}{\sqrt{g^2 + j^2c^2}}. \tag{5.8}$$

One easily checks that in fact $x = \gamma/\sigma$ and that (5.7) is the unitary equivalent of (3.10) for isolating the unobservable space (3.7). Each time that an unobservable eigenvalue a_{11} is isolated, we can also reduce the transformed state-space model (5.6) again to one of order n such that there is never an increase of state dimension in the overall process. This thus overcomes the last two drawbacks as well.

The overall algorithm now becomes very similar to the recursive cancellation algorithm of Section 2. It can in fact be viewed as a state-space equivalent of this method, whereby unitary transformations are used as much as possible.

ALGORITHM 5.1.

Step 1. Construct a state-space realization $\{A, B, C, D\}$ of $R(\lambda)$ (possibly by using some transformation of variable).

Step 2. Perform a unitary transformation to put $\{A, B, C, D\}$ in Schur form (3.5).

Step 3. WHILE there is still an undesired pole DO

Transform an undesired eigenvalue to position a_{11}.
Transform the (new) corresponding first column of C according to (5.2).
Choose $\{f, g, h, j\}$ satisfying the conditions (5.3), (5.4).
Construct the expanded state space model (5.6).
Deflate the uncontrollable eigenvalue a_{11} using (5.7).

END

6. RECURSIVE ZERO DISLOCATIONS IN STATE SPACE

We now turn to the recursive dislocation of zeros in state space. Since the zeros of the transfer function are described by a (possibly singular) generalized eigenvalue problem, one may now expect a slightly more involved procedure. Yet, the techniques used here are still very similar to those of the previous section.

We dislocate one zero at a time ($n_2 = 1$), and we assume that a state-space model is used along the lines of Theorem 4.2. More precisely, we assume that

$$\left[\begin{array}{ccc|c} \lambda I_{n_1} - A_{11} & -A_{12} & -A_{13} & B_1 \\ -A_{21} & \lambda - a_{22} & -A_{23} & B_2 \\ -A_{31} & -A_{32} & \lambda I_{n_3} - A_{33} & B_3 \\ \hline -C_1 & -C_2 & -C_3 & D \end{array}\right] U$$

$$= \begin{array}{l} \left[\begin{array}{ccc|c} \lambda \hat{E}_{11} - \hat{A}_{11} & \lambda \hat{E}_{12} - \hat{A}_{12} & \lambda \hat{E}_{13} - \hat{A}_{13} & \lambda \hat{F}_1 - \hat{B}_1 \\ 0 & \lambda \hat{e}_{22} - \hat{a}_{22} & \lambda \hat{E}_{23} - \hat{A}_{23} & \lambda \hat{F}_2 - \hat{B}_2 \\ 0 & 0 & \lambda \hat{E}_{33} - \hat{A}_{33} & \lambda \hat{F}_3 - \hat{B}_3 \\ \hline 0 & 0 & 0 & \hat{D} \end{array}\right] \begin{array}{l} \}n_1 \\ \}1 \\ \}n_3 \\ \}p \end{array} \\ \begin{array}{cccc} \qquad\underbrace{\qquad\qquad}_{\hat{n}_1} & \underbrace{\qquad\qquad}_{1} & \underbrace{\qquad\qquad}_{\hat{n}_3} & \underbrace{\qquad\qquad}_{\hat{m}} \end{array} \end{array} \tag{6.1}$$

where $\lambda \hat{E}_{33} - \hat{A}_{33}$ has full column rank for all λ and $\hat{D}$ also has full column rank, and where $\lambda \hat{E}_{11} - \hat{A}_{11}$ has full row rank for *almost* all λ. The matrices X_2 and F of Theorem 4.2 are now just scalars x_2 and f. For the realization (4.26) one obtains

$$\{f, G, H, J\} \overset{\text{def}}{=} \{\hat{f}, \hat{G}x_2, \hat{H}/x_2, \hat{J}\}, \tag{6.2}$$

and one can e.g. choose x_2 such that $\|G\| = \|H\|$. Moreover (4.27) simplifies to

$$\begin{bmatrix} I & Y_3 & \hat{G} \end{bmatrix} \begin{bmatrix} \hat{a}_{22}\hat{E}_{23} - \hat{e}_{22}\hat{A}_{23} & \hat{a}_{22}\hat{F}_2 - \hat{e}_{22}\hat{B}_2 \\ \hat{a}_{22}\hat{E}_{33} - \hat{e}_{22}\hat{A}_{33} & \hat{a}_{22}\hat{F}_3 - \hat{e}_{22}\hat{B}_3 \\ 0 & \hat{D} \end{bmatrix} = 0, \tag{6.3}$$

which is obtained by filling in $\lambda = \hat{f} = \hat{a}_{22} / \hat{e}_{22}$ in (4.27) and multiplying by $\hat{e}_{22}$. Because of the above rank conditions, this always has a solution. Moreover, if a special staircase form is used for $\lambda \hat{E}_{33} - \hat{A}_{33}$ [1], then (6.3) can be solved in $O(n^2)$ operations, since the matrix is already in column echelon form. This thus yields both Y_3 and $\hat{G}$ at the same time.

As in the previous section, we try to construct a realization (6.2) of the form

$$\left[\begin{array}{c|c} f & G \\ \hline H & J \end{array}\right] = \left[\begin{array}{c|cccc} f & g & 0 & \cdots & 0 \\ \hline h & j & & & \\ 0 & & & & \\ \vdots & & & I_{p-1} & \\ 0 & & & & \end{array}\right] \left[\begin{array}{c|c} 1 & \\ \hline & V \end{array}\right]$$

$$= \left[\begin{array}{c|cccc} \hat{f} & \hat{g}x_2 & 0 & \cdots & 0 \\ \hline \hat{h}/x_2 & \hat{j} & & & \\ 0 & & & & \\ \vdots & & & I_{p-1} & \\ 0 & & & & \end{array}\right] \left[\begin{array}{c|c} 1 & \\ \hline & V \end{array}\right], \tag{6.4}$$

with V unitary. Let $\hat{G}$ be the solution of (6.3). Then it can always be written as a product $[\hat{g}\ 0\ \ldots\ 0]V$ for some unitary V. If we choose V to be just a sequence of $p-1$ Givens rotations, this essentially determines V. The transfer function of (6.4) is clearly equal to

$$R_1^{-1}(\lambda) = \left[\begin{array}{c|c} r(\lambda) & \\ \hline & I_{p-1} \end{array}\right] V, \qquad \text{with} \qquad r(\lambda) = j - \frac{hg}{\lambda - f} = \hat{j} - \frac{\hat{h}\hat{g}}{\lambda - \hat{f}}, \tag{6.5}$$

and it is thus easy to choose the remaining $\hat{h}$ and $\hat{j}$ (or h and j) such that (6.5) has a zero $hg/j - f$ in the right location ($\in \Gamma$). Any additional conditions such as those described in the previous section are also easy to impose. This is due to the remaining degree of freedom x_2, which is typically used to balance the realization (6.2).

REMARK 6.1. Notice that the conditions of Theorem 4.2 are no longer all satisfied here. Indeed, since we only isolated *one* undesired zero in (6.1), $\Lambda(\hat{E}_{11}, \hat{A}_{11})$ may still contain elements outside Γ. Yet, in Theorem 4.2 we only used this condition to show that X_1 then had to be zero. In this section, we *choose* $X_1 = 0$ and gave a *simplified* construction for the case $n_2 = 1$, which does yield a zero dislocation.

Once the transfer function (6.5) has been constructed, we update (6.1) using (4.16b) with H, J having the special form (6.4). This is done by first multiplying C, D, and $\hat{D}$ from the left by V. This first step is cheap, since V is just a sequence of $p-1$ Givens rotations. Next, we update only one row of (6.1) using the scalars h and j, which is again cheap. The result is

$$\left[\begin{array}{ccc|c} \lambda I_{n_1} - A_{11} & -A_{12} & -A_{13} & B_1 \\ -A_{21} & \lambda - a_{22} & -A_{23} & B_2 \\ -A_{31} & -A_{32} & \lambda I_{n_3} - A_{33} & B_3 \\ \hline -JC_1 & -JC_2 - Hx_2 & -JC_3 - HX_3 & JD \end{array}\right] U$$

$$= \begin{array}{c} \left[\begin{array}{ccc|c} \lambda \hat{E}_{11} - \hat{A}_{11} & \lambda \hat{E}_{12} - \hat{A}_{12} & \lambda \hat{E}_{13} - \hat{A}_{13} & \lambda \hat{F}_1 - \hat{B}_1 \\ 0 & \lambda \hat{e}_{22} - \hat{a}_{22} & \lambda \hat{E}_{23} - \hat{A}_{23} & \lambda \hat{F}_2 - \hat{B}_2 \\ 0 & 0 & \lambda \hat{E}_{33} - \hat{A}_{33} & \lambda \hat{F}_3 - \hat{B}_3 \\ \hline 0 & -H\hat{x}_2 & -H\hat{X}_3 & J\hat{D} - H\hat{X}_4 \end{array}\right] \begin{array}{l} \}n_1 \\ \}1 \\ \}n_3 \\ \}p \end{array} \\ \begin{array}{cccc} \underbrace{\qquad\qquad}_{\hat{n}_1} & \underbrace{\qquad\qquad}_{1} & \underbrace{\qquad\qquad}_{\hat{n}_3} & \underbrace{\qquad\qquad}_{\hat{m}} \end{array} \end{array} \tag{6.6}$$

Here U has now to be updated in order to put the right-hand side of (6.6) again in the form (6.1). Since there is only one row to be annihilated in (6.1), this can be done with a sequence of $\hat{n}_3$ Givens rotations, as e.g. indicated in [16, 1]. The same techniques can also be used to "reorder" (6.1) so that a new undesired zero appears in the middle position of (6.1). It is also shown there that these updating transformations require only $O(n^2)$ operations.

The overall algorithm now becomes:

ALGORITHM 6.1.

Step 1: Construct a state-space realization $\{A, B, C, D\}$ of $R(\lambda)$ (possibly by using some transformation of variable).

Step 2. Perform a unitary transformation to put $\{A, B, C, D\}$ in the special form (6.1).

Step 3. WHILE there is still an undesired zero DO

Transform an undesired eigenvalue to position $\lambda \hat{e}_{22} - \hat{a}_{22}$ as in (6.1).

Solve for $\hat{f}$ and $\hat{G}$ according to (6.3) and construct V.
Choose $\{f, g, h, j\}$ according to (6.4) and such that (6.5) has the required zero (and other properties).
Construct the updated state-space model (6.6).
Update U in (6.6) to yield again a model in the form (6.1).
END

7. CONCLUDING REMARKS

In this paper we have presented a state-space approach to pole-zero cancellation, which can be used for constructing a wide variety of polynomial and rational matrix factorizations. The major advantage of the state-space approach over the transfer-function approach of Dewilde and Vandewalle [5, 14] is its simplicity. This is reflected e.g. in the fact that in state space a solution can be given to cancel all the poles or zeros at once, using the solution of a Sylvester equation (Sections 3 and 4). (Note that a similar "block" approach was also recently used in [10] for a related problem.) In Sections 5 and 6 we also presented recursive versions of this state-space approach which are inspired by the transfer-function methods and cancel only one pole or zero at a time. The equations to be solved then are scalar and can in fact be viewed as the scalar systems encountered when solving Sylvester equations recursively via a Schur approach. This results in a simple recursive method that can be implemented using basic building blocks of linear-algebra software.

One should point out here that once the Schur forms corresponding to the zeros or poles of the transfer function are known, the cancellation problem then boils down to the solution of a system of linear equations (namely a pole placement problem tackled via a Sylvester equation). This simple observation is not easy to deduce from the transfer-function approach. The factorization problems described in the introduction are thus all solved in a two-stage fashion: first compute the Schur forms of the zeros or poles (this is an *iterative* process), then solve the system of linear equations to construct the factors (this is a *finite* process). One could argue that it is a roundabout approach to use an iterative process to solve problems such as coprime factorization which can also be solved using a finite recursion [11]. Yet, it is e.g. recommended to solve Lyapunov and Sylvester equations via Schur methods as well [19].

The overall complexity of the method using state-space models can be evaluated, since the complexity of all its subproblems is well known. The Schur decompositions require $O(n^3)$ operations, and the recursive cancellations $O(n_{\gamma_c} n^2)$, where n_{γ_c} is the number of pole-zero cancellations to be performed. When the rational matrix is not given in state-space form, one still has to construct a realization, and the complexity of this initial step depends on the original representation of the transfer function [15]. In general, though, one would even expect that passing through state space would yield an economical way to construct the factorization. The other advantage of the state-space approach is the controlled numerical behavior of all intermediate steps. All linear-algebra problems encountered in this approach are well studied for their numerical behavior. Proving the stability of the concatenation of these subtasks into one algorithm requires a separate analysis, of course.

Not all issues were tackled here for reasons of simplicity. A full-scale algorithmic implementation of our method would also require the ability to deal with complex conjugate pairs using real arithmetic only. This would follow the lines of Sections 5 and 6, where now 2×2 blocks are being dealt with. This of course complicates the formulas, but all algorithmic details can be dealt with appropriately. This is very similar again to the solution of 2×2 blocks in the Schur method for Sylvester equations. Another problem is that of using generalized state-space models instead of state-space models. This can e.g. be needed if one wants to avoid a conformal mapping $\lambda = (a\mu + b)/(c\mu + d)$ when constructing a realization for $R(\lambda)$. Such techniques are required in order to deal with infinite poles or zeros efficiently, rather than using conformal mapping. All linear-algebra tools for such generalizations are in fact available, and this should cause no difficulties.

An unsolved problem in this paper, is the dislocation of the null-space structure of a transfer function to certain zero locations. This would lead to factorizations (1.1) of a $p \times m$ transfer function $R(\lambda)$ with $p > m$, where e.g. $R_2(\lambda)$ has a zero block at the bottom [i.e., the null-space structure of $R_2(\lambda)$ is now trivially displayed]. Such factorization can e.g. be found in [12]. It is currently being investigated whether our approach can also handle such cases. Other possible application areas could be sought in the minimal-design problem [22] or in state-space formulations of certain $\mathscr{H}_2$ and $\mathscr{H}_\infty$ control problems [7].

APPENDIX A

LEMMA A.1. *Let*

$$S_i(\lambda) = \left[\begin{array}{c|c} T_i(\lambda) & U_i(\lambda) \\ \hline -V_i(\lambda) & W_i(\lambda) \end{array}\right] \qquad \textit{for} \quad i = 1, 2$$

be the system matrices of the rational matrices $R_i(\lambda)$ *of dimensions* $p_i \times m_i$. *Then, provided* $m_1 = p_2$, *one obtains a system matrix* $S(\lambda)$ *for the product*

$$R(\lambda) = R_1(\lambda) R_2(\lambda) \tag{A.1}$$

by the product of the "embedded" system matrices:

$$S(\lambda) = \left[\begin{array}{cc|c} T_1(\lambda) & & U_1(\lambda) \\ & I_{n_2} & \\ \hline -V_1(\lambda) & & W_1(\lambda) \end{array}\right] \left[\begin{array}{cc|c} I_{n_1} & T_2(\lambda) & U_2(\lambda) \\ \hline & -V_2(\lambda) & W_2(\lambda) \\ & & \end{array}\right]$$

$$= \left[\begin{array}{cc|c} T_1(\lambda) & -U_1(\lambda)V_2(\lambda) & U_1(\lambda)W_2(\lambda) \\ 0 & T_2(\lambda) & U_2(\lambda) \\ \hline -V_1(\lambda) & -W_1(\lambda)V_2(\lambda) & W_1(\lambda)W_2(\lambda) \end{array}\right]. \tag{A.2}$$

Proof. The transfer function of a system matrix is its Schur complement with respect to it bottom right corner. It is easily seen that the transformations

$$\left[\begin{array}{cc|c} I_{n_1} & & \\ & I_{n_2} & \\ \hline V_1T_1^{-1} & & I_{p_1} \end{array}\right] S(\lambda) \left[\begin{array}{cc|c} I_{n_1} & & \\ & I_{n_2} & -T_2^{-1}U^2 \\ \hline & & I_{m2} \end{array}\right]$$

$$= \left[\begin{array}{cc|c} T_1(\lambda) & * & * \\ 0 & T_2(\lambda) & 0 \\ \hline 0 & * & R_1(\lambda)R_2(\lambda) \end{array}\right] \tag{A.3}$$

do not affect this Schur complement. It is now clear from the structure of the right-hand side of (A.3) that this Schur complement is exactly $R_1(\lambda)R_2(\lambda)$, which completes the proof. ■

COROLLARY A.1. *Let*

$$S_i(\lambda) = \left[\begin{array}{c|c} \lambda E_i - A_i & B_i \\ \hline -C_i & D_i \end{array}\right] \qquad \textit{for} \quad i = 1,2$$

be system matrices in the generalized state-space form of $R_i(\lambda)$. Then the above formula automatically yields a generalized state-space model for the product

$$S(\lambda)=\left[\begin{array}{cc|c}\lambda E_1-A_1 & -B_1C_2 & B_1D_2\\ 0 & \lambda E_2-A_2 & B_2\\ \hline -C_1 & -D_1C_2 & D_1D_2\end{array}\right]. \tag{A.4}$$

COROLLARY A.2. *Let $\{A_i, B_i, C_i, D_i\}$ for $i=1,2$ be realizations of $R_i(\lambda)$. Then the above formula automatically yields a state-space quadruple $\{A, B, C, D\}$ for the product*

$$\left[\begin{array}{c|c}A & B\\ \hline C & D\end{array}\right]=\left[\begin{array}{cc|c}A_1 & B_1C_2 & B_1D_2\\ 0 & A_2 & B_2\\ \hline C_1 & D_1C_2 & D_1D_2\end{array}\right]$$

$$=\left[\begin{array}{cc|c}A_1 & & B_1\\ & I_{n_2} & \\ \hline C_1 & & D_1\end{array}\right]\left[\begin{array}{cc|c}I_{n_1} & & \\ & A_2 & B_2\\ \hline & C_2 & D_2\end{array}\right]. \tag{A.5}$$

It is interesting to notice that the embedded matrices in (A.2), (A.5) are in fact nonminimal system matrices and realizations, respectively, of $R_1(\lambda)$ and $R_2(\lambda)$. System matrices (and realizations) of a product of two rational matrices can thus be obtained as a product of their respective system matrices (and realizations), provided these are appropriately chosen. We remark that similar observations were also made in [6].

The following lemmas are proved in [17].

LEMMA A.2. *If the pencils $\lambda E_l - A_l$ and $\lambda E_r - A_r$ are left and right invertible (for some λ), respectively, and if they have no common spectrum, then the equation*

$$(\lambda E_l - A_l)P - Q(\lambda E_r - A_r) = 0 \tag{A.6}$$

has the unique solution $P=0$, $Q=0$.

LEMMA A.3. *If the pencils $\lambda_l E - A_l$ and $\lambda E_r - A_r$ are left and right invertible (for some λ), respectively, and if they have no common spectrum, then the equation*

$$P(\lambda E_l - A_l) - (\lambda E_r - A_r)Q = \lambda S - T \tag{A.7}$$

always has a solution for P, Q.

This research was partly done while the author was visiting the Department of Systems Engineering of the Australian National University (winter 1985).

REFERENCES

1 T. Beelen and P. Van Dooren, An improved algorithm for the computation of Kronecker's canonical form of a singular pencil, *Linear Algebra Appl.* 105:9–65 (1988).
2 V. Belevitch, *Classical Network Theory*, Holden Day, San Francisco, 1968.
3 C. T. Chen, *Introduction to Linear System Theory*, Holt, Rinehart & Winston, New York, 1970.
4 C. Desoer, *Notes for a Second Course on Linear Systems*, Van Nostrand, New York, 1970.
5 P. Dewilde and J. Vandewalle, On the factorization of a nonsingular rational matrix, *IEEE Trans. Circuits and Systems* CAS-22:637–645 (1975).
6 A. Dijksma, H. Langer, and H. de Snoo, Characteristic functions of unitary operator colligations in Π_k-spaces, in *Operator Theory and Systems*, Vol. OT 19, Birkhäuser, 1986, pp. 125–194.
7 J. Doyle, K. Glover, P. Khargonekar, and B. Francis, State-space solutions to standard $\mathscr{H}_2$ and $\mathscr{H}_\infty$ control problems, *IEEE Trans. Automat. Control* AC-34:831–847 (1989).
8 F. Duren, *Theory of $\mathscr{H}^p$ spaces*, Academic, New York, 1970.
9 A. Emami and P. Van Dooren, Computation of zeros of linear multivariable systems, *Automatica* 18:415–430 (1982).
10 I. Gohberg, M. Kaashoek, and A. Ran, Partial pole and zero displacement by cascade connection, *SIAM J. Matrix Anal. Appl.* 10:316–325 (1989).
11 T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, N.J., 1980.
12 L. Pernebo, An algebraic theory for the design of controllers for linear multivariable systems I, II, *IEEE Trans. Automat. Control* AC-26:171–193 (1981).
13 G. W. Stewart, HQR3 and EXCHNG. Fortran subroutines for calculating and ordering the eigenvalues of a real upper Hessenberg matrix, *ACM Trans. Math. Software* 2:275–280 (1976).

14 J. Vandewalle and P. Dewilde, On the irreducible cascade synthesis of a system with a real rational transfer matrix, *IEEE Trans. Circuits and Systems* CAS-24:481–494 (1977).

15 P. Van Dooren, The generalized eigenstructure problem in linear system theory, *IEEE Trans. Automat. Control* AC-26:111–129 (1981).

16 P. Van Dooren, Algorithm 590. DSUBSP and EXCHQZ: Fortran subroutines for computing deflating subspaces with specified spectrum, *ACM Trans. Math. Software* 8:376–382 (1982).

17 P. Van Dooren, Reducing subspaces: Definitions, properties, and algorithms, in *Matrix Pencils* (A. Ruhe and B. Kågström, Eds.), Lecture Notes in Math. 973, Springer-Verlag, New York, 1983, pp. 58–73.

18 P. Van Dooren, Factorization of a rational matrix: The singular case, *Integral Equations Operator Theory* 7:704–741 (1984).

19 P. Van Dooren and M. Verhaegen, On the use of unitary state space transformations, in *Contemp. Math.*, Vol. 47, Amer. Math. Soc., 1985.

20 G. Verghese, P. Van Dooren, and T. Kailath, Properties of the system matrix of a generalized state space system, *Internat. J. Control* 30:235–243 (1979).

21 M. Vidyasagar, *Control System Synthesis: A Factorization Approach*, MIT Press, Cambridge, Mass., 1985.

22 S.-H. Wang and E. Davison, A minimization algorithm for the design of linear multivariable systems, *IEEE Trans. Automat. Control* AC-18:220–225 (1973).

Received 26 September 1989; final manuscript accepted 11 March 1990