SIAM J. CONTROL AND OPTIMIZATION
Vol. 18, No 6, November 1980

0363-0129/80/1806-0008/$01.00/0

FACTORIZATIONS OF TRANSFER FUNCTIONS*

H. BART†, I. GOHBERG‡, M. A. KAASHOEK† AND P. VAN DOOREN§

Abstract. This paper is concerned with minimal factorizations of rational matrix functions. The treatment is based on a new geometrical principle. In fact, it is shown that there is a one-to-one correspondence between minimal factorizations on the one hand and certain projections on the other. Considerable attention is given to the problem of stability of a minimal factorization. Also the numerical aspects are discussed. Along the way, a stability theorem for solutions of the matrix Riccati equation is obtained.

Introduction. The problem of factorizing a rational matrix-valued function $W(\lambda)$ into "simpler" rational factors has network theory as one of its origins. In this theory $W(\lambda)$ appears as a transfer function of a network. Its minimal factorizations (see Chapter II) are of particular interest because it allows one to obtain the network by a cascade connection of elementary sections which have the simplest synthesis [6], [22].

In the present paper the treatment of the factorization problem is based on a new geometrical principle. This principle has been observed independently by the first three authors and by the fourth (and has been communicated at a miniconference on Operators and System Theory held at Amsterdam and Delft, February, 1978). For the fourth author network theory [22], [23] has been the main motivation, while the first three authors were inspired by [3], [7], [20].

The new geometrical principle referred to allows for a unifying approach to seemingly disjoint topics such as the network problems mentioned above, the matrix Riccati equation [19], the factorization theory of characteristic functions for linear operators [7], the theory of Wiener-Hopf (or spectral) factorization [10], [11] and the divisibility theory of operator polynomials [3], [12], [13]. Here we treat only the first two topics; the other connections will be investigated in detail in a forthcoming publication [5].

The problem of computing numerically the minimal factors of a transfer function led us to investigate the stability of divisors under small perturbations. We pay considerable attention to the measure of stability.

The matrix functions studied here are viewed as transfer functions of systems. A system consists of three matrices A, B, and C, of appropriate sizes, and the corresponding transfer functions are of the form

$$W(\lambda)=I+C(\lambda I-A)^{-1}B,$$

where λ is the complex variable and I the identity matrix. In the first chapter multiplication and division of transfer functions are described in terms of systems. Applications to matrix Riccati equations are also considered here. The special type of minimal factorization and its properties are studied in Chapter II. In geometrical terms an explicit description of all minimal factors is given. Stability and numerical aspects are studied in the last two chapters. Throughout the paper we confine ourselves to the finite dimensional case, but with minor modifications the results of Chapters I and III are also valid in the infinite dimensional situation (see [5]).

* Received by the editors June 22, 1979, and in final revised form March 10, 1980.

† Wiskundig Seminarium, Vrije Universiteit, Amsterdam, The Netherlands.

‡ Department of Mathematical Sciences, Tel-Aviv University, Ramat Aviv and Department of Pure Mathematics, Weizmann Institute of Science, Rehovot, Israel.

§ Division of Applied Mathematics and Computer Science, Katholieke Universiteit, Leuven, Belgium.

As far as notation and terminology is concerned we stipulate the following. The term *linear space* stands for a complex vector space. All linear spaces appearing below are assumed to be finite dimensional. In Chapters III and IV it is also assumed that they are endowed with a norm, which is always denoted by $\|\cdot\|$. By an *operator* we mean a linear transformation between two linear spaces. The null space and range of an operator T are denoted by Ker T and Im T, respectively. The identity operator on a linear space X is always denoted by I. The symbol I_n is used for the $n\times n$ identity matrix. Whenever this is convenient, an $m\times n$ matrix A will be identified with the operator from $\mathbb{C}^n$ into $\mathbb{C}^m$ given by the canonical action of A with respect to the standard bases in $\mathbb{C}^n$ and $\mathbb{C}^m$. In particular a rational $n\times n$ matrix function may be viewed as a rational function whose values are operators acting on $\mathbb{C}^n$.

I. Divisibility of transfer functions and the Riccati equation. In this chapter multiplication and division of transfer functions are described in terms of systems. The main result on factorization is presented in § 1.1. A slightly more sophisticated factorization theorem, involving the notion of an angular operator, is given in § 1.2. In § 1.3 we discuss the operator Riccati equation.

1.1. Multiplication and divisibility of systems. A *system* is a quintet $\theta=(A, B, C; X, Y)$ of two linear spaces X, Y and three operators $A: X\to X$, $B: Y\to X$ and $C: X\to Y$. The space X is called the *state space*; the space Y is called the *input/output space*. The operator A is referred to as the *state space* or *main operator*. A common way to give systems is to specify three matrices of appropriate sizes. To be more specific, if A is a $\delta\times\delta$ matrix, B is a $\delta\times n$ matrix and C is an $n\times\delta$ matrix, then (identifying A, B, and C in the usual way with operators) the quintet $(A, B, C; \mathbb{C}^\delta, \mathbb{C}^n)$ is a system.

Two systems $\theta_1=(A_1, B_1, C_1; X_1, Y)$ and $\theta_2=(A_2, B_2, C_2; X_2, Y)$ are said to be *similar*, written $\theta_1\simeq\theta_2$, if there exists an invertible operator $S: X_1\to X_2$, called a *system similarity*, between θ_1 and θ_2 such that

$$A_1=S^{-1}A_2S,\quad B_1=S^{-1}B_2,\quad C_1=C_2S.$$

The relation $\simeq$ is reflexive, symmetric and transitive.

Let $\theta=(A, B, C; X, Y)$ be a system, and put

$$W(\lambda)=I+C(\lambda I-A)^{-1}B. \tag{1.1}$$

Then $W(\lambda)$ is a rational operator function and $W(\infty)=I$. This function is called the *transfer function* of θ, and is denoted by W_θ. Obviously, similar systems have the same transfer function.

If $W(\lambda)$ is any rational function whose values are operators acting on Y and $W(\infty)=I$, then it is known from system theory (cf. [2]) that $W(\lambda)$ can be represented in the form (1.1). Such a representation is called a *realization* for $W(\lambda)$; we also use this term for the system $(A, B, C; X, Y)$.

Our terminology is taken from system theory, where the transfer function (1.1) is used to describe the input/output behavior of the linear dynamical system

$$\dot{x}(t)=Ax(t)+Bu(t),\qquad y(t)=Cx(t)+u(t).$$

In the theory of characteristic operator functions, certain systems with special properties are called *nodes* (see, for instance, [7]). The connections with this theory are further developed in [5]. In the next paragraph we shall define the product of two systems. The definition is motivated by the notion of a series connection of two linear dynamical systems. For details, the reader is referred to [18] (cf. also [7]).

Let $\theta_1 = (A_1, B_1, C_1; X_1, Y)$ and $\theta_2 = (A_2, B_2, C_2; X_2, Y)$ be systems. Put $X = X_1 \oplus X_2$, and

$$A = \begin{bmatrix} A_1 & B_1C_2 \\ 0 & A_2 \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = [C_1 \quad C_2].$$

Then $(A, B, C; X, Y)$ is a system. It is called the *product* of θ_1 and θ_2 and denoted by $\theta_1\theta_2$. A straightforward calculation shows that

$$W_{\theta_1\theta_2}(\lambda) = W_{\theta_1}(\lambda) W_{\theta_2}(\lambda). \tag{1.2}$$

So if θ_1 and θ_2 are realizations for $W_1(\lambda)$ and $W_2(\lambda)$, respectively, then $\theta_1\theta_2$ is a realization for $W_1(\lambda)W_2(\lambda)$.

If $\theta = (A, B, C; X, Y)$ is a realization for the rational operator function $W(\lambda)$, then

$$\theta^\times = (A - BC, B, -C; X, Y)$$

is a realization for $W(\lambda)^{-1}$. We call $\theta^\times$ the *associate system* of θ. The operator $A - BC$ is called the *associate* (*main*) *operator* of θ. By abuse of notation, we write $A^\times = A - BC$. Note that $A^\times$ depends not only on A, but also on the other operators appearing in the system θ. One checks without difficulty that $(\theta^\times)^\times = \theta$ (so in particular $(A^\times)^\times = A$) and $(\theta_1\theta_2)^\times = \theta_2^\times\theta_1^\times$, the natural identification of $X_1 \oplus X_2$ and $X_2 \oplus X_1$ being a system similarity.

Consider the system $\theta = (A, B, C; X, Y)$ and let Π be a projection of X. So Π is an idempotent operator on X. With respect to the decomposition $X = \operatorname{Ker} \Pi \oplus \operatorname{Im} \Pi$, we write

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = [C_1 \quad C_2]. \tag{1.3}$$

The system $\mathrm{pr}_\Pi(\theta) = (A_{22}, B_2, C_2; \operatorname{Im} \Pi, Y)$ is called the *projection* of θ associated with Π (cf. [7]). Observe that $\mathrm{pr}_{I-\Pi}(\theta) = (A_{11}, B_1, C_1; \operatorname{Ker} \Pi, Y)$. One easily verifies that $\mathrm{pr}_\Pi(\theta)^\times = \mathrm{pr}_\Pi(\theta^\times)$. The projection Π is said to be a *supporting projection* for θ if

$$A[\operatorname{Ker} \Pi] \subset \operatorname{Ker} \Pi, \qquad A^\times[\operatorname{Im} \Pi] \subset \operatorname{Im} \Pi. \tag{1.4}$$

If Π is a supporting projection for θ, then $I - \Pi$ is one for $\theta^\times$. The second part of (1.4) is equivalent to the rank condition.

$$\operatorname{rank} \begin{bmatrix} A_{12} & B_1 \\ C_2 & I \end{bmatrix} = \dim Y. \tag{1.5}$$

This is immediate from the fact that the left-hand side of (1.5) is equal to rank $(A_{12} - B_1C_2) + \dim Y$.

The following theorem admits a very simple proof. Nevertheless it is one of the cornerstones for the rest of the present paper. A somewhat more sophisticated factorization theorem will be presented in § 1.2.

THEOREM 1.1. *Let Π be a supporting projection for the system $\theta = (A, B, C; X, Y)$. Then*

$$\theta = \mathrm{pr}_{I-\Pi}(\theta) \cdot \mathrm{pr}_\Pi(\theta). \tag{1.6}$$

If $W(\lambda)$, $W_1(\lambda)$ and $W_2(\lambda)$ are the transfer functions of θ, $\mathrm{pr}_{I-\Pi}(\theta)$ and $\mathrm{pr}_\Pi(\theta)$, respectively, then $W(\lambda) = W_1(\lambda)W_2(\lambda)$. In other words,

$$I + C(\lambda I - A)^{-1}B = [I + C(\lambda I - A)^{-1}(I - \Pi)B][I + C\Pi(\lambda I - A)^{-1}B].$$

Proof. With respect to the decomposition $X = \operatorname{Ker} \Pi \oplus \operatorname{Im} \Pi$, we write the operators A, B, and C as in (1.3). Then $A^\times$ may be written as

$$A^\times = \begin{bmatrix} A_{11} - B_1C_1 & A_{12} - B_1C_2 \\ A_{21} - B_2C_1 & A_{22} - B_2C_2 \end{bmatrix}.$$

Hence (1.4) is equivalent to $A_{21} = 0$ and $A_{12} - B_1C_2 = 0$. It follows that

$$A = \begin{bmatrix} A_{11} & B_1C_2 \\ 0 & A_{22} \end{bmatrix}.$$

But then (1.6) is clear from the definition of the product of two systems. The second part of the theorem is now an immediate consequence of formula (1.2).

In a certain sense Theorem 1.1 gives a complete description of all possible factorizations of the system θ. Indeed, if $\theta \simeq \theta_1\theta_2$ for some systems θ_1 and θ_2, then there exists a supporting projection Π for θ such that $\theta_1 \simeq \operatorname{pr}_{I-\Pi}(\theta)$ and $\theta_2 \simeq \operatorname{pr}_\Pi(\theta)$.

1.2. Angular operators and the division theorem. Throughout this section, X is a linear space and Π is a projection of X onto X_2 along X_1. (Block) matrix representations of operators acting on X will always be taken with respect to the decomposition $X = X_1 \oplus X_2$.

A subspace N of X is called *angular* with respect to Π if $X = \operatorname{Ker} \Pi \oplus N$. If R is an operator from X_2 into X_1, then the space

$$N_R = \{Rx + x \mid x \in X_2\}$$

is angular with respect to Π. The next proposition shows that every angular subspace is of this form.

PROPOSITION 1.2. *Let N be a subspace of X. Then N is angular with respect to Π if and only if $N = N_R$ for some operator R from X_2 into X_1.*

Proof. We have already observed that if $N = N_R$, then N is angular with respect to Π. To prove the converse, assume that N is angular with respect to Π, and let Q be the projection of X onto N along X_1. Put $Rx = (Q - \Pi)x$ for $x \in X_2$. Then $N = N_R$.

Given an angular subspace N, the operator R for which $N = N_R$ is uniquely determined. It is called the *angular operator* for N with respect to Π. This notion was introduced by M. G. Krein in [17]. We are now in a position to bring the division theorem for systems into a slightly more general form.

THEOREM 1.3. *Let $\theta = (A, B, C; X, Y)$ be a system, let Π be a projection of X onto X_2 along X_1, and let N be an angular subspace of X with respect to Π. Assume that*

$$A[X_1] \subset X_1, \qquad A^\times[N] \subset N. \tag{1.7}$$

Further, let

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = [C_1 \quad C_2],$$

be the matrix representations of A, B and C with respect to the decomposition $X = X_1 \oplus X_2$, let R be the angular operator for N with respect to Π, and put

$$\theta_1 = (A_{11}, B_1 - RB_2, C_1; X_1, Y), \tag{1.8}$$

$$\theta_2 = (A_{22}, B_2, C_1R + C_2; X_2, Y). \tag{1.9}$$

Then $\theta \simeq \theta_1\theta_2$. More precisely,

$$\theta_1\theta_2 = (E^{-1}AE, E^{-1}B, CE; X, Y),$$

where E is the invertible operator

$$E=\begin{bmatrix} I & R \\ 0 & I \end{bmatrix}.$$

Proof. For convenience, put $\hat{A}=E^{-1}AE$, $\hat{B}=E^{-1}B$, $\hat{C}=CE$ and $\hat{\theta}=(\hat{A},\hat{B},\hat{C};X,Y)$. Observe that $\hat{A}^{x}=E^{-1}A^{x}E$. Now E maps X_1 onto X_1 and X_2 onto N. Thus (1.7) implies that

$$\hat{A}[X_1]\subset X_1,\qquad \hat{A}^{x}[X_2]\subset X_2.$$

Apply now Theorem 1.1 to show that

$$\hat{\theta}=\mathrm{pr}_{I-\Pi}(\hat{\theta})\cdot \mathrm{pr}_{\Pi}(\hat{\theta}).$$

But $\mathrm{pr}_{I-\Pi}(\hat{\theta})=\theta_1$ and $\mathrm{pr}_{\Pi}(\hat{\theta})=\theta_2$, and the proof is complete.

Suppose that the angular subspace N in Theorem 1.3 is the image of X_2 under some invertible operator

$$S=\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}: X_1\oplus X_2\rightarrow X_1\oplus X_2.$$

Then it is not difficult to prove that S_{22} is invertible. Moreover the angular operator R for N is given by $R=S_{12}S_{22}^{-1}$. By substituting this in (1.8) and (1.9), we get

$$\theta_1=(A_{11},B_1-S_{12}S_{22}^{-1}B_2,C_1;X_1,Y),$$

$$\theta_2=(A_{22},B_1,C_1S_{12}S_{22}^{-1}+C_2;X_2,Y).$$

This together with formula (1.2), can be used to give a quick proof of Theorem 4 in Sahnovič's paper [20].

1.3. The Riccati equation. As in the previous section, X is a linear space and Π is a projection of X onto X_2 along X_1. In view of Theorem 1.3 the following question is of interest. Given an angular subspace N of X and an operator T on X, when is N invariant under T? The next proposition shows that the answer involves an operator Riccati equation.

PROPOSITION 1.4. *Let N be an angular subspace of X with respect to Π, and let*

$$T=\begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}: X_1\oplus X_2\rightarrow X_1\oplus X_2$$

be an operator on X. Then N is invariant under T if and only if the angular operator R for N satisfies the Riccati equation

$$RT_{21}R+RT_{22}-T_{11}R-T_{12}=0. \tag{1.10}$$

Proof. The operator

$$E=\begin{bmatrix} I & R \\ 0 & I \end{bmatrix}: X_1\oplus X_2\rightarrow X_1\oplus X_2$$

is invertible and maps X_2 onto N. So

$$E^{-1}TE=\begin{bmatrix} T_{11}-RT_{21} & -RT_{21}R-RT_{22}+T_{11}R+T_{12} \\ T_{21} & T_{22}+T_{21}R \end{bmatrix}$$

leaves invariant X_2 if and only if T leaves invariant N. But $E^{-1}TE$ leaves invariant X_2 if and only if (1.10) is satisfied, and the proof is complete.

In view of formula (1.2) and Theorem 1.3, the problem of finding factorizations for transfer functions of systems is related to that of solving a certain Riccati operator equation. As a matter of fact, the condition $A^{\times}[N]\subset N$ is equivalent to the requirement

$$RB_2C_1R+R(B_2C_2-A_{22})+(A_{11}-B_1C_1)R+A_{12}-B_1C_2=0.$$

Here we use the notation of Theorem 1.3.

Now let us introduce some more notation and terminology. Let T be an operator on X and let μ be an eigenvalue of T. The subspace $\operatorname{Ker}(\mu I-T)^m$, where m is the dimension of X, is called the *generalized eigenspace* of T corresponding to μ. If $\lambda_1,\cdots,\lambda_r$ are eigenvalues of T, the space

$$\operatorname{Ker}(\lambda_1 I-T)^m\oplus\cdots\oplus\operatorname{Ker}(\lambda_r I-T)^m \tag{1.11}$$

is called the *spectral subspace* for T corresponding to the eigenvalues $\lambda_1,\cdots,\lambda_r$. This spectral subspace can also be described as follows. Let Γ be a contour in $\mathbb{C}$ such that $\lambda_1,\cdots,\lambda_r$ are inside and the remaining eigenvalues of T are outside Γ. Put

$$P(T;\Gamma)=\frac{1}{2\pi i}\int_\Gamma(\lambda I-T)^{-1}\,d\lambda.$$

Then the spectral subspace (1.11) coincides with the image of $P(T;\Gamma)$. In view of this, (1.11) is also called the spectral subspace for T corresponding to Γ. The operator $P(T;\Gamma)$ is a projection of X, called the *Riesz projection* corresponding to T and Γ (or $\lambda_1,\cdots,\lambda_r$).

PROPOSITION 1.5. *Let N be an angular subspace of X with respect to Π, and let*

$$T=\begin{bmatrix}T_{11} & T_{12}\\ T_{21} & T_{22}\end{bmatrix}:X_1\oplus X_2\rightarrow X_1\oplus X_2,$$

be an operator on X. Then N is a spectral subspace for T if and only if the angular operator R for N satisfies the Riccati equation (1.10) *and the operators $T_{11}-RT_{21}$ and $T_{22}+T_{21}R$ have no common eigenvalues.*

It will appear from the proof that if N is the spectral subspace for T corresponding to the contour Γ, then the eigenvalues of $T_{22}+T_{21}R$ are inside Γ and the eigenvalues of $T_{11}-RT_{21}$ are outside Γ.

Proof. Define E as in the proof of Proposition 1.4. It is clear that N is a spectral subspace for T (corresponding to a contour Γ) if and only if X_2 is a spectral subspace (corresponding to the same contour Γ) for $S=E^{-1}TE$. With respect to the decomposition $X=X_1\oplus X_2$, we write

$$S=\begin{bmatrix}S_{11} & S_{12}\\ S_{21} & S_{22}\end{bmatrix}.$$

Recall that $S_{11}=T_{11}-RT_{21}$, $S_{12}=-RT_{21}R-RT_{22}+T_{11}R+T_{12}$, $S_{21}=T_{21}$, and $S_{22}=T_{22}+T_{21}R$.

Now suppose that $S_{12}=0$ and that S_{11} and S_{22} have no common eigenvalues. Let Γ be a Cauchy contour such that the eigenvalues of S_{11} are outside and the eigenvalues of S_{22} are inside Γ. Then $P(S;\Gamma)$ has the form

$$P(S;\Gamma)=\begin{bmatrix}0 & 0\\ * & I\end{bmatrix},$$

and it follows that $X_2=\operatorname{Im}P(S;\Gamma)$.

Next assume that X_2 is the spectral subspace for S corresponding to the contour Γ.

Then in particular X_2 is S-invariant and so $S_{12}=0$. Write $P=P(S;\Gamma)$. The operator S_{22} is the restriction of S to Im P. Thus the eigenvalues of S_{22} are precisely the eigenvalues of S lying inside Γ. Let S_0 be the restriction of S to Ker P. Then the eigenvalues of S_0 are precisely the eigenvalues of S lying outside Γ. In particular S_{22} and S_0 have no common eigenvalues. It remains to prove that S_0 and S_{11} have the same eigenvalues.

Since $\operatorname{Im} P = X_2 = \operatorname{Im} \Pi$, we have $I-P=(I-P)(I-\Pi)$ and the map

$$F=(I-P)|X_1\colon X_1 \to \operatorname{Ker} P$$

is an invertible operator. One easily verifies that $S_0F=FS_{11}$. So S_0 and S_{11} are similar, and the proof is complete.

II. Minimality and minimal factorizations. In this chapter we discuss minimal systems and minimal factorizations of rational matrix functions. The main result is Theorem 2.2, which shows that there is a one-to-one correspondence between minimal factorizations and supporting projections of minimal systems.

2.1. Minimal nodes. Let X and Y be linear spaces. A pair of operators (A, B),

$$A\colon X\to X, \qquad B\colon Y\to X,$$

is called *controllable* if, for k sufficiently large,

$$\operatorname{Im} B+\operatorname{Im} AB+\cdots+\operatorname{Im} A^{k-1}B=X. \tag{2.1}$$

Similarly, a pair (A, C)

$$A\colon X\to X, \qquad C\colon X\to Y,$$

is said to be *observable* if, for k sufficiently large,

$$\operatorname{Ker} C\cap \operatorname{Ker} CA\cap\cdots\cap \operatorname{Ker} CA^{k-1}=(0). \tag{2.2}$$

Observe that the left-hand sides of (2.1) and (2.2) are independent of k, provided k is larger than or equal to the degree of the minimal polynomial of A.

A system $\theta=(A,B,C;X,Y)$ is called *minimal* if (A,B) is controllable and (A,C) is observable. Such systems play an important role in the sequel. Below we collect together a number of facts concerning minimal systems that are either wellknown or easy to prove (cf. [2], [16] and the references given there).

If θ is minimal, then so is $\theta^{\times}$. Similarity of systems implies that their transfer functions coincide. The converse of this is not true in general. However, if θ and Δ are minimal systems for which $W_\theta=W_\Delta$, then θ and Δ are similar. This result is known as the state space isomorphism theorem. If S is a system similarity between two minimal systems, then S is uniquely determined. In other words, the only system similarity between a minimal system and itself is the identity operator. Given a system θ, there exists a minimal system (unique up to similarity) whose transfer function coincides with that of θ. The product of two minimal systems need not be minimal. However, if the product of two systems is minimal, then so are the factors. In particular, if Π is a supporting projection for the minimal system θ, then $\operatorname{pr}_\Pi(\theta)$ and $\operatorname{pr}_{I-\Pi}(\theta)$ are both minimal.

2.2. Minimality and McMillan degree. Let $W(\lambda)$ be a rational $n\times n$ matrix function, and let λ_0 be a complex number. Then λ_0 is at worst a pole of $W(\lambda)$. So, taking p sufficiently large, we may write

$$W(\lambda)=\sum_{j=-p}^{\infty}(\lambda-\lambda_0)^jW_j,$$

the expansion being valid in some deleted neighborhood of λ_0. The rank of the block Hankel matrix

$$\begin{bmatrix} W_{-1} & W_{-2} & \cdots & W_{-p} \\ W_{-2} & \cdots & W_{-p} & 0 \\ \vdots & & & \vdots \\ W_{-p} & 0 & \cdots & 0 \end{bmatrix},$$

is called the *degree* of W at λ_0. It is denoted by $\delta(W;\lambda_0)$. Observe that $\delta(W;\lambda_0)$ does not depend on the choice of p. For equivalent definitions and generalizations, see [15]. We also define the degree $\delta(W;\infty)$ of W at ∞ to be the degree of $W(\lambda^{-1})$ at 0.

It is clear that $\delta(W;\mu)=0$ if and only if $W(\lambda)$ is analytic at μ. Therefore it makes sense to put

$$\delta(W)=\sum_{\mu\in\mathbb{C}_\infty}\delta(W;\mu).$$

Here $\mathbb{C}_\infty=\mathbb{C}\cup\{\infty\}$. The number $\delta(W)$ is called the *McMillan degree* of W.

Assume that $W(\infty)=I_n$. Then $W(\lambda)$ admits a realization of the form

(2.3) $$W(\lambda)=I_n+C(\lambda I_\delta-A)^{-1}B.$$

The system $\theta=(A,B,C;\mathbb{C}^\delta,\mathbb{C}^n)$ has $W(\lambda)$ as its transfer function. We call the realization (2.3) *minimal* if θ is a minimal system. In fact (2.3) is minimal if and only if δ is equal to the McMillan degree $\delta(W)$ of W. If (2.3) is not minimal, then $\delta>\delta(W)$. From (2.3) it is clear that each pole of $W(\lambda)$ is an eigenvalue of A. In general the converse is not true, but if the realization is minimal, each eigenvalue of A is a pole of $W(\lambda)$. So in that case the set of poles of $W(\lambda)$ coincides with the set of eigenvalues of A. Similarly, if (2.3) is minimal, the set of poles of $W(\lambda)^{-1}$ coincides with the set of eigenvalues of $A^\times=A-BC$. Poles of $W(\lambda)^{-1}$ are usually called *zeros* of $W(\lambda)$.

2.3. Minimal factorizations. Let $W(\lambda)$, $W_1(\lambda)$ and $W_2(\lambda)$ be rational $n\times n$ matrix functions, and assume that

(2.4) $$W(\lambda)=W_1(\lambda)W_2(\lambda).$$

Then it is known (cf., e.g., [26]) that $\delta(W)\leqq\delta(W_1)+\delta(W_2)$. In fact this inequality holds pointwise in the following sense:

(2.5) $$\delta(W;\mu)\leqq\delta(W_1;\mu)+\delta(W_2;\mu),\qquad \mu\in\mathbb{C}_\infty.$$

The factorization (2.4) is called *minimal* if $\delta(W)=\delta(W_1)+\delta(W_2)$. An equivalent requirement is that in (2.5) we have equality for all $\mu\in\mathbb{C}_\infty$.

In dealing with minimal factorizations, we shall always suppose that det $W(\lambda)\not\equiv 0$. This implies the existence of $a\in\mathbb{C}$ such that $W(a)$ is invertible. Put $\tilde W(\lambda)=W(a)^{-1}W(\lambda^{-1}+a)$. Then clearly $\tilde W(\infty)=I_n$. There is a one-to-one correspondence between the (minimal) factorizations of $W(\lambda)$ and those of $\tilde W(\lambda)$. So (from a theoretical point of view) there is no loss of generality in assuming that $W(\infty)=I_n$.

Suppose $W(\infty)=I_n$. We are interested in the minimal factorizations of $W(\lambda)$. We claim that it suffices to consider only those factorizations (2.4) of $W(\lambda)$ for which $W_1(\infty)=W_2(\infty)=I_n$. To make this claim more precise, assume that (2.4) is a minimal factorization of $W(\lambda)$. Then $\delta(W_1;\infty)+\delta(W_2;\infty)=\delta(W;\infty)=0$, because W is analytic at ∞. Hence $\delta(W_1;\infty)=\delta(W_2;\infty)=0$, or, in other words, W_1 and W_2 are analytic at ∞. Moreover $I_n=W(\infty)=W_1(\infty)W_2(\infty)$, and so $W_1(\infty)$ and $W_2(\infty)$ are each other's inverse. Put $U=W_1(\infty)^{-1}$. By multiplying $W_1(\lambda)$ from the right with U and

$W_2(\lambda)$ from the left with U^{-1}, we obtain a minimal factorization of $W(\lambda)$ whose factors have the value I_n at ∞.

These considerations justify the fact that, *from now on, without further mentioning, we only deal with rational matrix functions that are analytic at ∞ with value the identity matrix.* In other words, the rational matrix functions considered below will be viewed as transfer functions of systems.

PROPOSITION 2.1. *Let $W(\lambda)=W_1(\lambda)W_2(\lambda)$ be a factorization of the rational matrix function $W(\lambda)$, let θ_1 be a minimal realization for $W_1(\lambda)$ and let θ_2 be a minimal realization for $W_2(\lambda)$. Then the factorization is minimal if and only if the product $\theta_1\theta_2$ is a minimal system.*

Proof. Let $\theta=(A, B, C; \mathbb{C}^\delta, \mathbb{C}^n)$ be a realization for the rational matrix function $W(\lambda)$, i.e., formula (2.3) is satisfied. Then θ is a minimal system if and only if $\delta=\delta(W)$. From this and the definition of the product of two systems, the desired result is clear.

We now come to the main result of this chapter.

THEOREM 2.2. *Let θ be a minimal realization of the rational matrix function $W(\lambda)$.*

(i) *If Π is a supporting projection for θ, $W_1(\lambda)$ is the transfer function of* $\mathrm{pr}_{I-\Pi}(\theta)$ *and $W_2(\lambda)$ is the transfer function of* $\mathrm{pr}_\Pi(\theta)$, *then $W(\lambda)=W_1(\lambda)W_2(\lambda)$ is a minimal factorization of $W(\lambda)$.*

(ii) *If $W(\lambda)=W_1(\lambda)W_2(\lambda)$ is a minimal factorization, then there exists a unique supporting projection Π for the system θ such that $W_1(\lambda)$ and $W_2(\lambda)$ are the transfer functions of* $\mathrm{pr}_{I-\Pi}(\theta)$ *and* $\mathrm{pr}_\Pi(\theta)$, *respectively.*

Proof. Statement (i) is an immediate consequence of Proposition 2.1. Therefore we concentrate on (ii). Assume that $W(\lambda)=W_1(\lambda)W_2(\lambda)$ is a minimal factorization. For $i=1,2$, let θ_i be a minimal realization of $W_i(\lambda)$ with state space $\mathbb{C}^{\delta_i}$. Here δ_i is the McMillan degree $\delta(W_i)$ of W_i. By Proposition 2.1, the product $\theta_1\theta_2$ is a minimal realization for $W(\lambda)$. Hence $\theta_1\theta_2$ and θ are similar, say with system similarity $S\colon \mathbb{C}^{\delta_1}\oplus\mathbb{C}^{\delta_2}\to\mathbb{C}^\delta$, where $\delta=\delta_1+\delta_2=\delta(W)$. Let Π be the projection of $\mathbb{C}^\delta$ along $S[\mathbb{C}^{\delta_1}]$ onto $S[\mathbb{C}^{\delta_2}]$. Then Π is a supporting projection for θ. Moreover $\mathrm{pr}_{I-\Pi}(\theta)$ is similar to θ_1 and $\mathrm{pr}_\Pi(\theta)$ is similar to θ_2. It remains to prove the unicity of Π.

Suppose P is another supporting projection of θ such that $\mathrm{pr}_P(\theta)$ and $\mathrm{pr}_{I-P}(\theta)$ are realizations of $W_2(\lambda)$ and $W_1(\lambda)$ respectively. Then $\mathrm{pr}_P(\theta)$ and $\mathrm{pr}_{I-P}(\theta)$ are minimal again. Hence $\mathrm{pr}_{I-\Pi}(\theta)$ and $\mathrm{pr}_{I-P}(\theta)$ are similar, say with system similarity $U\colon \operatorname{Ker}\Pi\to\operatorname{Ker}P$, and $\mathrm{pr}_\Pi(\theta)$ and $\mathrm{pr}_P(\theta)$ are similar, say with system similarity $V\colon \operatorname{Im}\Pi\to\operatorname{Im}P$. Define T on $\mathbb{C}^\delta$ by

$$T=\begin{bmatrix} U & 0\\ 0 & V\end{bmatrix}\colon \operatorname{Ker}\Pi\oplus\operatorname{Im}\Pi\to\operatorname{Ker}P\oplus\operatorname{Im}P.$$

Then T is a system similarity between θ and itself. Since θ is minimal, it follows that T is the identity operator on $\mathbb{C}^\delta$. But then $\Pi=P$, and the proof is complete.

III. Stability of spectral divisors. The problem of computing numerically the minimal factorizations of a given transfer function leads in a natural way to questions concerning the stability of divisors of a system. These and related questions form the main topic of this chapter.

3.1. Examples and first results. The property of having nontrivial minimal factorizations may be ill-conditioned. To see this, consider the following example. Let

$$W_\varepsilon(\lambda)=\begin{bmatrix} 1+\dfrac{1}{\lambda} & \dfrac{\varepsilon}{\lambda^2}\\ 0 & 1+\dfrac{1}{\lambda}\end{bmatrix}. \tag{3.1}$$

For each ε this is the transfer function of the minimal system $\theta_\varepsilon = (A, I, I; \mathbb{C}^2, \mathbb{C}^2)$, where $I = I_2$ and

$$A_\varepsilon = \begin{bmatrix} 0 & \varepsilon \\ 0 & 0 \end{bmatrix}.$$

To find a nontrivial minimal factorization of the function (3.1), we have to find nontrivial supporting projections of the system θ_ε (cf. Theorem 2.2); i.e., we must look for nontrivial subspaces M and $M^\times$ of $\mathbb{C}^2$, invariant under A_ε and $A_\varepsilon^\times = A_\varepsilon - I$, respectively, such that $M \oplus M^\times = \mathbb{C}^2$. The operators A_ε and $A_\varepsilon - I$ have the same invariant subspaces, and for $\varepsilon \neq 0$ there is only one such subspace of dimension one, namely the first coordinate space. It follows that for $\varepsilon \neq 0$, the rational matrix function (3.1) has no nontrivial minimal factorizations. For $\varepsilon = 0$, we have

$$W_0(\lambda) = \begin{bmatrix} 1 + \dfrac{1}{\lambda} & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 + \dfrac{1}{\lambda} \end{bmatrix},$$

and this factorization is minimal, because the McMillan degree of $W_0(\lambda)$ is equal to 2 and the McMillan degree of each factor is 1.

The next theorem shows that under certain conditions the existence of a minimal factorization is a stable property.

THEOREM 3.1. *Consider the rational matrix function*

$$W_0(\lambda) = I_n + C_0(\lambda I_\delta - A_0)^{-1} B_0. \tag{3.2}$$

Assume that the realization (3.2) *is minimal and that* $W_0(\lambda)$ *admits a minimal factorization* $W_0(\lambda) = W_{01}(\lambda) W_{02}(\lambda)$,

$$W_{0i}(\lambda) = I_n + C_{0i}(\lambda I_{\delta_i} - A_{0i})^{-1} B_{0i}, \qquad i = 1, 2, \tag{3.3}$$

where $\delta = \delta_1 + \delta_2$ *and the factors* $W_{01}(\lambda)$ *and* $W_{02}(\lambda)$ *have no common zeros and no common poles. Then there exist positive constants* ω *and* ε *such that the following holds. If* A, B, *and* C *are matrices* (*of the appropriate sizes*) *such that*

$$\|A - A_0\| + \|B - B_0\| + \|C - C_0\| < \omega, \tag{3.4}$$

then the realization $W(\lambda) = I_n + C(\lambda I_\delta - A)^{-1} B$ *is minimal and* $W(\lambda)$ *admits a minimal factorization* $W(\lambda) = W_1(\lambda) W_2(\lambda)$,

$$W_i(\lambda) = I_n + C_i(\lambda I_{\delta_i} - A_i)^{-1} B_i, \qquad i = 1, 2, \tag{3.5}$$

such that the factors $W_1(\lambda)$ *and* $W_2(\lambda)$ *have no common zeros and no common poles, and*

$$\|A_i - A_{0i}\| + \|B_i - B_{0i}\| + \|C_i - C_{0i}\| \leqq \varepsilon (\|A - A_0\| + \|B - B_0\| + \|C - C_0\|),$$

for $i = 1, 2$.

The above theorem deals with "spectral" minimal factorizations. The stability of nonspectral minimal factorizations is investigated in [5]. The theorem is proved in § 3.3. The proof provides explicit estimates for ω and ε.

3.2. Opening between subspaces and angular operators. Let M and N be subspaces of the linear space X, and let $\|\cdot\|$ be a norm on X. The number

$$\eta(M, N) = \inf \{\|x + y\| \mid x \in M, y \in N, \max \{\|x\|, \|y\|\} = 1\}$$

is called the *minimal opening* between M and N. Note that always $0 \leqq \eta(M, N) \leqq 1$, except when both M and N are trivial, in which case $\eta(M, N) = \infty$. It is well known (see [14, Lemma 1]) that $\eta(M, N) > 0$ if and only if $M \cap N = (0)$. If Π is a projection of the space X, then

$$\max\{\|\Pi\|, \|I - \Pi\|\} \leqq \frac{1}{\eta(\operatorname{Ker}\Pi, \operatorname{Im}\Pi)}. \tag{3.6}$$

Sometimes it will be convenient to describe $\eta(M, N)$ in terms of the *minimal angle* $\phi_{\min}$ between M and N. By definition this quantity is given by the following formulas:

$$0 \leqq \phi_{\min} \leqq \frac{\Pi}{2}, \qquad \sin \phi_{\min} = \eta(M, N).$$

(cf. [14]). Put

$$\rho(M, N) = \sup_{0 \neq x \in M} \inf_{y \in N} \frac{\|x - y\|}{\|x\|}.$$

The number

$$\operatorname{gap}(M, N) = \max\{\rho(M, N), \rho(N, M)\}$$

is called the *gap* or *maximal opening* between M and N. There is an extensive literature on this concept[1].

Now let us assume that $X = \mathbb{C}^n$, endowed with the usual Euclidean norm. Let P and Q be the orthogonal projections of X onto M and N, respectively. It can be shown that $\operatorname{gap}(M, N) = \|P - Q\|$ (cf. [1]). Also

$$1 - \eta(M, N)^2 = \sup_{0 \neq x \in M} \frac{\|Qx\|^2}{\|x\|^2} = \sup_{0 \neq y \in N} \frac{\|Py\|^2}{\|y\|^2}, \tag{3.7}$$

provided both M and N are nontrivial.

LEMMA 3.2. *Let Π_0, Π and Π_1 be projections of the linear space X, and assume that* $\operatorname{Ker}\Pi_0 = \operatorname{Ker}\Pi = \operatorname{Ker}\Pi_1$. *Let R be the angular operator of* $\operatorname{Im}\Pi$ *with respect to Π_0 and let R_1 be that of* $\operatorname{Im}\Pi_1$. *The following assertions hold*:

(i) $\eta(\operatorname{Ker}\Pi_0, \operatorname{Im}\Pi_0) \cdot \rho(\operatorname{Im}\Pi_1, \operatorname{Im}\Pi) \leqq \|R_1 - R\|$.

(ii) *If* $\rho(\operatorname{Im}\Pi_1, \operatorname{Im}\Pi) < \eta(\operatorname{Ker}\Pi, \operatorname{Im}\Pi)$, *then*

$$\|R_1 - R\| \leqq \frac{\rho(\operatorname{Im}\Pi_1, \operatorname{Im}\Pi) \cdot (1 + \|R\|)}{\eta(\operatorname{Ker}\Pi, \operatorname{Im}\Pi) - \rho(\operatorname{Im}\Pi_1, \operatorname{Im}\Pi)}.$$

In particular, if $\rho(\operatorname{Im}\Pi_1, \operatorname{Im}\Pi_0) < \eta(\operatorname{Ker}\Pi_0, \operatorname{Im}\Pi_0)$, *then*

$$\|R_1\| \leqq \frac{\rho(\operatorname{Im}\Pi_1, \operatorname{Im}\Pi_0)}{\eta(\operatorname{Ker}\Pi_0, \operatorname{Im}\Pi_0) - \rho(\operatorname{Im}\Pi_1, \operatorname{Im}\Pi_0)}.$$

Results of this type seem to be well known. Therefore we omit the proof. We proceed with a lemma which will be most useful in the next section.

LEMMA 3.3. *Let P, $P^\times$, Q and $Q^\times$ be projections of the linear space X, and put* $\alpha_0 = \frac{1}{6}\eta(\operatorname{Im}P, \operatorname{Im}P^\times) \cdot (1 + \|P^\times\|)^{-1}$. *Assume* $X = \operatorname{Im}P \oplus \operatorname{Im}P^\times$ *and*

$$\|P - Q\| + \|P^\times - Q^\times\| < \alpha_0. \tag{3.8}$$

[1] For details, see T. Kato: *Perturbation Theory For Linear Operators*, Springer, Berlin–Heidelberg–New York, 1966, and the references given there.

Then $X = \operatorname{Im} Q \oplus \operatorname{Im} Q^\times$ *and there exists an invertible operator* $S: X \to X$ *such that*

(i) $S[\operatorname{Im} Q] = \operatorname{Im} P, \quad S[\operatorname{Im} Q^\times] = \operatorname{Im} P^\times$,

(ii) $\max\{\|S - I\|, \|S^{-1} - I\|\} \leqq \beta(\|P - Q\| + \|P^\times - Q^\times\|)$,

where $\beta = 2[\alpha_0 \eta(\operatorname{Im} P, \operatorname{Im} P^\times)]^{-1}$.

Proof. For simplicity we put $d = \|P - Q\| + \|P^\times - Q^\times\|$ and $\eta = \eta(\operatorname{Im} P, \operatorname{Im} P^\times)$. Since $\operatorname{gap}(\operatorname{Im} P, \operatorname{Im} Q) \leqq \|P - Q\|$ and $\operatorname{gap}(\operatorname{Im} P^\times, \operatorname{Im} Q^\times) \leqq \|P^\times - Q^\times\|$, condition (3.8) implies that

$$2 \operatorname{gap}(\operatorname{Im} P, \operatorname{Im} Q) + 2 \operatorname{gap}(\operatorname{Im} P^\times, \operatorname{Im} Q^\times) < \eta.$$

But then we may apply Theorem 2 from [14] to show that $X = \operatorname{Im} Q \oplus \operatorname{Im} Q^\times$.

Note that (3.8) also implies that $\|P - Q\| < \frac{1}{4}$. Hence $S_1 = I + P - Q$ is invertible and we can write $S_1^{-1} = I + V$ with $\|V\| \leqq \frac{4}{3}\|P - Q\| < \frac{1}{3}$. As $I - P + Q$ is invertible too, we have $S_1[\operatorname{Im} Q] = \operatorname{Im} P$. By direct calculation, it can be shown that

$$\|S_1 Q^\times S_1^{-1} - P^\times\| \leqq 3\|P^\times - Q^\times\| + 3 \cdot \|P^\times\| \cdot \|P - Q\|,$$

and hence

$$\rho(\operatorname{Im} S_1 Q^\times S_1^{-1}, \operatorname{Im} P^\times) \leqq \|S_1 Q^\times S_1^{-1} - P^\times\| \leqq 3d(1 + \|P^\times\|) < \frac{\eta}{2}.$$

Let $\Pi_0(\Pi)$ be the projection of X along $\operatorname{Im} P(\operatorname{Im} Q)$ onto $\operatorname{Im} P^\times(\operatorname{Im} Q^\times)$, and put $\tilde{\Pi} = S_1 \Pi S_1^{-1}$. Then $\tilde{\Pi}$ is a projection of X and $\operatorname{Ker} \tilde{\Pi} = \operatorname{Ker} \Pi_0$. Further, $\operatorname{Im} \tilde{\Pi} = \operatorname{Im} S_1 Q^\times S_1^{-1}$, and so we have

$$\rho(\operatorname{Im} \tilde{\Pi}, \operatorname{Im} \Pi_0) < \frac{\eta}{2} = \frac{1}{2}\eta(\operatorname{Ker} \Pi_0, \operatorname{Im} \Pi_0).$$

Hence, if R denotes the angular operator of $\operatorname{Im} \tilde{\Pi}$ with respect to Π_0, then because of Lemma 3.2,

$$\|R\| \leqq \frac{2}{\eta} \rho(\operatorname{Im} \tilde{\Pi}, \operatorname{Im} \Pi_0).$$

Since $\rho(\operatorname{Im} \tilde{\Pi}, \operatorname{Im} \Pi_0) \leqq 3d(1 + \|P^\times\|)$, this implies that $\|R\| \leqq d\alpha_0^{-1}$.

Next, put $S_2 = I - R\Pi_0$, and take $S = S_2 S_1$. Clearly, S_2 is invertible; in fact $S_2^{-1} = I + R\Pi_0$. It follows that S is invertible too. From the properties of the angular operator one easily sees that with this choice of S statement (i) holds true. It remains to prove (ii).

From $S = (I - R\Pi_0)(I + P - Q)$ and the fact that $\|P - Q\| < \frac{1}{4}$, one deduces that $\|S - I\| \leqq \|P - Q\| + \frac{5}{4}\|R\| \cdot \|\Pi_0\|$. Moreover $\|R\| \leqq d\alpha_0^{-1}$, and from (3.6) we know that $\|\Pi_0\| \leqq \eta^{-1}$. It follows that

$$\|S - I\| \leqq d + \frac{5d}{4\alpha_0 \eta}. \tag{3.9}$$

Recall that $S_1^{-1} = I + V$ with $\|V\| \leqq \frac{4}{3}\|P - Q\| < \frac{1}{3}$. This can be used to show that

$$\|S^{-1} - I\| \leqq \frac{4d}{3} + \frac{4d}{3\alpha_0 \eta}. \tag{3.10}$$

Statement (ii) is now an easy consequence of (3.9), (3.10), and the fact that $6\alpha_0 \eta \leqq 1$.

3.3. Stability of spectral divisors. Let $\theta = (A, B, C; X, Y)$ and $\theta_0 = (A_0, B_0, C_0; X, Y)$ be systems. The *distance* between θ and θ_0 is defined to be the number

$$\|\theta-\theta_0\|=\|A-A_0\|+\|B-B_0\|+\|C-C_0\|.$$

We also put $\|\theta\|=\|A\|+\|B\|+\|C\|$. If $W(\lambda)$ and $W_0(\lambda)$ are the transfer functions of θ and θ_0, respectively, then

$$\|W(\lambda)-W_0(\lambda)\|\leqq\frac{\|\theta-\theta_0\|\cdot\|\theta\|\cdot\|\theta_0\|}{\|A\|\cdot\|A_0\|},$$

provided that $|\lambda|>2\max\{\|A\|,\|A_0\|\}$.

THEOREM 3.4. *Let Π_0 be a supporting projection for the system $\theta_0=(A_0,B_0,C_0;X,Y)$, and assume that*

$$\operatorname{Ker}\Pi_0=\operatorname{Im}P,\qquad \operatorname{Im}\Pi_0=\operatorname{Im}P^{\times},$$

where P and $P^{\times}$ are projections of X. Put

$$\alpha_0=\tfrac{1}{6}\eta(\operatorname{Ker}\Pi_0,\operatorname{Im}\Pi_0)\cdot(1+\|P^{\times}\|)^{-1}.$$

Let $\theta=(A,B,C;X,Y)$ be another system, and let Q and $Q^{\times}$ be projections of X such that

$$A[\operatorname{Im}Q]\subset\operatorname{Im}Q,\qquad A^{\times}[\operatorname{Im}Q^{\times}]\subset\operatorname{Im}Q^{\times},\tag{3.11}$$
$$\|P-Q\|+\|P^{\times}-Q^{\times}\|<\alpha_0.$$

Then $X=\operatorname{Im}Q\oplus\operatorname{Im}Q^{\times}$, there exists an invertible operator $S:X\to X$ such that $S^{-1}\Pi_0S$ is the projection Π of X onto $\operatorname{Im}Q^{\times}$ along $\operatorname{Im}Q$, and the projection Π_0 is a supporting projection for the system $\hat\theta=(SAS^{-1},SB,CS^{-1};X,Y)$, while for the corresponding factors we have

$$\max\{\|\mathrm{pr}_{I-\Pi_0}(\theta_0)-\mathrm{pr}_{I-\Pi_0}(\hat\theta)\|,\|\mathrm{pr}_{\Pi_0}(\theta_0)-\mathrm{pr}_{\Pi_0}(\hat\theta)\|\}$$
$$\leqq\frac{9}{\eta(\operatorname{Im}P,\operatorname{Im}P^{\times})^3}\left[\|\theta-\theta_0\|+\frac{\|\theta_0\|}{\alpha_0}(\|P-Q\|+\|P^{\times}-Q^{\times}\|)\right].$$

Proof. From Lemma 3.3 we know that $X=\operatorname{Im}Q\oplus\operatorname{Im}Q^{\times}$. Take S as in Lemma 3.3. Then $S^{-1}\Pi_0S$ is the projection Π of X onto $\operatorname{Im}Q^{\times}$ along $\operatorname{Im}Q$. It is now clear from formula (3.11) that $S^{-1}\Pi_0S$ is a supporting projection for the system $\theta=(A,B,C;X,Y)$. But then Π_0 is a supporting projection for $\hat\theta=(SAS^{-1},SB,CS^{-1};X,Y)$.

Let θ_{01} and $\hat\theta_1$ be the left factors of θ_0 and $\hat\theta$ associated with Π_0, and let θ_{02} and $\hat\theta_2$ be the right factors. From the definition of the factors (see § 1.1) it is clear that $\|\theta_{01}-\hat\theta_1\|\leqq\|I-\Pi_0\|\cdot\|\theta_0-\hat\theta\|$ and $\|\theta_{02}-\hat\theta_2\|\leqq\|\Pi_0\|\cdot\|\theta_0-\hat\theta\|$. Using (3.6), we obtain

$$\max\|\theta_{0i}-\hat\theta_i\|\leqq\frac{1}{\eta}\|\theta_0-\hat\theta\|,\qquad i=1,2,\tag{3.12}$$

where $\eta=\eta(\operatorname{Im}P,\operatorname{Im}P^{\times})$. As $\|\theta_0-\hat\theta\|\leqq\|\theta_0-\theta\|+\|\theta-\hat\theta\|$, it remains to compute a suitable upper bound for $\|\theta-\hat\theta\|$.

Put $S=I+V$ and $S^{-1}=I+W$. Note that $\|\theta-\hat\theta\|\leqq\|A\|(\|V\|+\|W\|+\|V\|\cdot\|W\|)+\|B\|\cdot\|V\|+\|C\|\cdot\|W\|$. By Lemma 3.3, we have $\max\{\|V\|,\|W\|\}\leqq 2d(\alpha_0\eta)^{-1}$, where $d=\|P-Q\|+\|P^{\times}-Q^{\times}\|$. It follows that

$$\|\theta-\hat\theta\|\leqq\frac{4d}{\alpha_0\eta}\left(1+\frac{d}{\alpha_0\eta}\right)\|\theta\|.\tag{3.13}$$

Since $d\alpha_0^{-1} < 1$ and $\eta \leqq 1$, one can use formula (3.13) to show that

$$\|\theta_0 - \hat{\theta}\| \leqq \frac{9}{\eta^3}\left[\|\theta - \theta_0\| + \frac{d}{\alpha_0}\|\theta_0\|\right].$$

This, together with formula (3.12), gives the desired result.

THEOREM 3.5. *Let Π_0 be a supporting projection for the system $\theta_0 = (A_0, B_0, C_0; X, Y)$. Assume that*

$$\operatorname{Ker} \Pi_0 = \operatorname{Im} P(A_0; \Gamma), \qquad \operatorname{Im} \Pi_0 = \operatorname{Im} P(A_0^\times; \Gamma^\times),$$

where Γ and $\Gamma^\times$ are contours such that A_0 and $A_0^\times$ have no eigenvalues on Γ and $\Gamma^\times$, respectively. Then there exist positive constants α, β_1 and β_2 such that the following holds. If $\theta = (A, B, C; X, Y)$ is a system such that $\|\theta - \theta_0\| < \alpha$, then A has no eigenvalues on Γ, $A^\times$ has no eigenvalues on $\Gamma^\times$,

$$X = \operatorname{Im} P(A; \Gamma) \oplus \operatorname{Im} P(A^\times; \Gamma^\times),$$

the projection Π of X along $\operatorname{Im} P(A; \Gamma)$ onto $\operatorname{Im} P(A^\times; \Gamma^\times)$ is a supporting projection for θ, and there exists a similarity transformation S such that

$$\|S - I\| \leqq \beta_1 \|\theta - \theta_0\|,$$

$\Pi_0 = S\Pi S^{-1}$, Π_0 is a supporting projection for the system $\hat{\theta} = (SAS^{-1}, SB, CS^{-1}; X, Y)$ and for the corresponding divisors we have

$$\|\mathrm{pr}_{I-\Pi_0}(\theta_0) - \mathrm{pr}_{I-\Pi_0}(\hat{\theta})\| \leqq \beta_2 \|\theta - \theta_0\|,$$

$$\|\mathrm{pr}_{\Pi_0}(\theta_0) - \mathrm{pr}_{\Pi_0}(\hat{\theta})\| \leqq \beta_2 \|\theta - \theta_0\|.$$

Furthermore, if θ_0 is minimal, then α can be chosen such that θ is minimal whenever $\|\theta - \theta_0\| < \alpha$.

Proof. Let ℓ be the maximum of the lengths of the curves Γ and $\Gamma^\times$,

$$\gamma = \max\left\{\max_{\lambda \in \Gamma} \|(\lambda I - A_0)^{-1}\|, \max_{\lambda \in \Gamma^\times} \|(\lambda I - A_0^\times)^{-1}\|\right\},$$

and

$$\alpha_0 = \tfrac{1}{6}\eta(\operatorname{Ker} \Pi_0, \operatorname{Im} \Pi_0) \cdot (1 + \|P(A_0^\times; \Gamma^\times)\|)^{-1}.$$

Put

$$\alpha = (1 + \|\theta_0\|)^{-1} \min\left\{1, \frac{1}{2\gamma}, \frac{\alpha_0 \pi}{2\gamma^2 \ell}\right\},$$

$$\beta_1 = 4(1 + \|\theta_0\|)\gamma^2 \ell [\pi \alpha_0 \eta(\operatorname{Ker} \Pi_0, \operatorname{Im} \Pi_0)]^{-1}, \tag{3.14}$$

$$\beta_2 = \frac{9}{\eta(\operatorname{Ker} \Pi_0, \operatorname{Im} \Pi_0)^3}\left[1 + \frac{2\gamma^2 \ell}{\pi \alpha_0}\|\theta_0\|(1 + \|\theta_0\|)\right].$$

We shall prove that α, β_1 and β_2 have the properties mentioned in the first part of the theorem. For convenience we write $\eta = \eta(\operatorname{Ker} \Pi_0, \operatorname{Im} \Pi_0)$, $P = P(A_0; \Gamma)$ and $P^\times = P(A_0^\times; \Gamma^\times)$.

Suppose $\theta = (A, B, C; X, Y)$ is a system such that $\|\theta - \theta_0\| < \alpha$. Then, in particular, $\|\theta - \theta_0\| < 1$. Since $\|A^\times - A_0^\times\| \leqq \|\theta - \theta_0\| \cdot (1 + \|\theta_0\|)$, it follows that

$$\max\{\|A - A_0\|, \|A^\times - A_0^\times\|\} < \frac{1}{2\gamma}.$$

Using elementary spectral theory, we may conclude that A has no eigenvalues on Γ, $A^\times$ has no eigenvalues on $\Gamma^\times$, while further,

$$\|(\lambda I-A)^{-1}-(\lambda I-A_0)^{-1}\|\leqq 2\gamma^2\|\theta-\theta_0\|\cdot(1+\|\theta_0\|), \qquad \lambda\in\Gamma,$$
$$\|(\lambda I-A^\times)^{-1}-(\lambda I-A_0^\times)^{-1}\|\leqq 2\gamma^2\|\theta-\theta_0\|\cdot(1+\|\theta_0\|), \qquad \lambda\in\Gamma^\times.$$

Hence for the corresponding Riesz projections $Q=P(A;\Gamma)$ and $Q^\times=P(A^\times;\Gamma^\times)$, we have

$$\|P-P^\times\|+\|Q-Q^\times\|\leqq\frac{2\gamma^2\ell}{\pi}\|\theta-\theta_0\|\cdot(1+\|\theta_0\|)<\alpha_0.$$

The fact that α, β_1, and β_2 meet the requirements of the first part of the theorem is now an easy consequence of Lemma 3.3 and Theorem 3.4.

Assume that θ_0 is minimal. We want to define the constant α in such a way that it also has the property that θ is minimal whenever $\|\theta-\theta_0\|<\alpha$. Let n be the dimension of X. Since θ_0 is minimal, the operator $\operatorname{col}(C_0A_0^j)_{j=0}^{n-1}$ is left invertible, say with left inverse L, and the operator $\operatorname{row}(A_0^jB_0)_{j=0}^{n-1}$ is right invertible, say with right inverse R. If $E:X\to Y^n$ is an operator satisfying

$$\|E-\operatorname{col}(C_0A_0^j)_{j=0}^{n-1}\|<\|L\|^{-1},$$

then E is also left invertible. A similar remark can be made involving $\operatorname{row}(A_0^jB_0)_{j=0}^{n-1}$. Hence there exists a positive number $\omega(\theta_0)$ such that $\|\theta-\theta_0\|<\omega(\theta_0)$ implies that θ is minimal. The new α may now be defined as

$$\alpha=\min\left[\omega(\theta_0),(1+\|\theta_0\|)^{-1}\min\left\{1,\frac{1}{2\gamma},\frac{\alpha_0\pi}{2\gamma^2\ell}\right\}\right]. \tag{3.15}$$

This completes the proof of the theorem.

We now come to the proof of Theorem 3.1.

Proof of Theorem 3.1. The matrices appearing in Theorem 3.1, will be identified with their canonical action as operators. Put $\theta_0=(A_0,B_0,C_0;\mathbb{C}^\delta,\mathbb{C}^n)$ and

$$\theta_{0i}=(A_{0i},B_{0i},C_{0i};\mathbb{C}^{\delta_i},\mathbb{C}^n), \qquad i=1,2.$$

Since the factorization $W_0(\lambda)=W_{01}(\lambda)W_{02}(\lambda)$ is minimal and $\delta_1+\delta_2=\delta$, the realizations (3.3) are minimal. Hence θ_0 is similar to the product $\bar\theta_0=\theta_{01}\theta_{02}$, say with system similarity $T:\mathbb{C}^\delta\to\mathbb{C}^{\delta_1}\oplus\mathbb{C}^{\delta_2}$.

With respect to the direct sum $\mathbb{C}^{\delta_1}\oplus\mathbb{C}^{\delta_2}$, the main operator $\bar A_0$ and the associate main operator $\bar A_0^\times$ of the system $\bar\theta_0=\theta_{01}\theta_{02}$ have the form

$$\bar A_0=\begin{bmatrix}A_{01} & B_{01}C_{02}\\ 0 & A_{02}\end{bmatrix}, \qquad \bar A_0^\times=\begin{bmatrix}A_{01}^\times & 0\\ -B_{02}C_{01} & A_{02}^\times\end{bmatrix}.$$

The hypothesis of Theorem 3.1 concerning the poles and zeros of $W_{01}(\lambda)$ and $W_{02}(\lambda)$ just means that A_{01} and A_{02} have no common eigenvalues and that $A_{01}^\times$ and $A_{02}^\times$ have no common eigenvalues. Let Γ be a contour that separates the eigenvalues of A_{01} from those of A_{02}. Similarly, let $\Gamma^\times$ be a contour that separates the eigenvalues of $A_{01}^\times$ from those of $A_{02}^\times$. Then

$$\operatorname{Im}P(\bar A_0;\Gamma)=\mathbb{C}^{\delta_1}\oplus(0), \qquad \operatorname{Im}P(\bar A_0^\times;\Gamma^\times)=(0)\oplus\mathbb{C}^{\delta_2}.$$

It follows that we may apply Theorem 3.5 to the system $\bar\theta_0$.

Let α and β_2 be the positive numbers that according to Theorem 3.5 correspond to the system $\bar{\theta}_0$ (cf. (3.14) and (3.15)), and put

$$\omega = \alpha[\|T\| \cdot \|T^{-1}\| + \|T\| + \|T^{-1}\|]^{-1},$$
$$\varepsilon = \beta_2[\|T\| \cdot \|T^{-1}\| + \|T\| + \|T^{-1}\|]^{-1}.$$

Suppose (3.4) holds and write $\bar{\theta} = (TAT^{-1}, TB, CT^{-1}; \mathbb{C}^{\delta_1} \oplus \mathbb{C}^{\delta_2}, \mathbb{C}^n)$. Then

$$\|\bar{\theta} - \bar{\theta}_0\| \leqq \|\theta - \theta_0\| \cdot (\|T\| \cdot \|T^{-1}\| + \|T\| + \|T^{-1}\|) < \alpha.$$

Hence $\bar{\theta}$ is minimal. This means that the realization $W(\lambda) = I_n + C(\lambda I_\delta - A)^{-1}B$ is minimal. Moreover, there exists a similarity transformation S such that for the system

$$\tilde{\theta} = (STAT^{-1}S^{-1}, STB, CT^{-1}S^{-1}; \mathbb{C}^{\delta_1} \oplus \mathbb{C}^{\delta_2}, \mathbb{C}^n),$$

the projection Π_0 of $\mathbb{C}^{\delta_1} \oplus \mathbb{C}^{\delta_2}$ along $\mathbb{C}^{\delta_1} \oplus (0)$ onto $(0) \oplus \mathbb{C}^{\delta_2}$ is a supporting projection. This shows that $W(\lambda)$ admits a minimal factorization $W(\lambda) = W_1(\lambda) W_2(\lambda)$, with $W_1(\lambda)$ and $W_2(\lambda)$ of the form (3.5). We also know that

$$\|\mathrm{pr}_{I-\Pi_0}(\bar{\theta}_0) - \mathrm{pr}_{I-\Pi_0}(\tilde{\theta})\| \leqq \beta_2 \|\bar{\theta}_0 - \bar{\theta}\|,$$
$$\|\mathrm{pr}_{\Pi_0}(\bar{\theta}_0) - \mathrm{pr}_{\Pi_0}(\tilde{\theta})\| \leqq \beta_2 \|\bar{\theta}_0 - \bar{\theta}\|.$$

But this is the same as

$$\|A_i - A_{0i}\| + \|B_i - B_{0i}\| + \|C_i - C_{0i}\| \leqq \beta_2 \|\bar{\theta}_0 - \bar{\theta}\|$$
$$\leqq \varepsilon \|\theta - \theta_0\| = \varepsilon(\|A - A_0\| + \|B - B_0\| + \|C - C_0\|).$$

Let $\bar{A}$ be the main operator of $\bar{\theta}$, and let $\bar{A}^\times$ be the main operator of the associate system $\bar{\theta}^\times$. As $\|\bar{\theta} - \bar{\theta}_0\| < \alpha$, we can apply Theorem 3.5 to show that $\bar{A}$ has no eigenvalues on Γ, $\bar{A}^\times$ has no eigenvalues on $\Gamma^\times$, and

$$\mathbb{C}^{\delta_1} \oplus \mathbb{C}^{\delta_2} = \mathrm{Im}\, P(\bar{A}; \Gamma) \oplus \mathrm{Im}\, P(\bar{A}^\times; \Gamma^\times).$$

Let Π be the projection of $\mathbb{C}^{\delta_1} \oplus \mathbb{C}^{\delta_2}$ along $\mathrm{Im}\, P(\bar{A}; \Gamma)$ onto $\mathrm{Im}\, P(\bar{A}^\times; \Gamma^\times)$. Then $\Pi_0 = S\Pi S^{-1}$. It follows that the eigenvalues of A_1 are inside and those of A_2 are outside the contour Γ. Similarly the eigenvalues of $A_2^\times$ are inside and those of $A_1^\times$ are outside $\Gamma^\times$. Thus $W_1(\lambda)$ and $W_2(\lambda)$ have no common zeros and no common poles. This completes the proof.

3.4. Application to the Riccati equation. In this section we show that the method of § 3.3 also can be used to prove stability theorems for solutions of the operator Riccati equation. Here we restrict ourselves to "spectral" solutions (cf. Theorem 3.6 below). The general case has been investigated in [4], [8].

Throughout this section, X_1 and X_2 are linear spaces. We use the symbol $\mathcal{L}(X_j, X_i)$ to indicate the space of all linear operators from X_j into X_i.

THEOREM 3.6. *Let $T_{ij} \in \mathcal{L}(X_j, X_i)$, $i, j = 1, 2$, and let $R \in \mathcal{L}(X_2, X_1)$ be a solution of the Riccati equation*

$$ZT_{21}Z + ZT_{22} - T_{11}Z - T_{12} = 0.$$

Assume that $T_{11} - RT_{21}$ and $T_{22} + T_{21}R$ have no common eigenvalues, and let Γ be a contour whose interior domain contains all eigenvalues of $T_{22} + T_{21}R$ and whose exterior domain contains all eigenvalues of $T_{11} - RT_{21}$. Then there exist positive constants ω and ε such that the following holds. If $S_{ij} \in \mathcal{L}(X_j, X_i)$ and

$$\|S_{ij} - T_{ij}\| < \omega, \qquad i, j = 1, 2, \tag{3.16}$$

then the equation

$$ZS_{21}Z + ZS_{22} - S_{11}Z - S_{12} = 0 \tag{3.17}$$

has a solution $Q \in \mathcal{L}(X_2, X_1)$ *such that all eigenvalues of* $S_{22} + S_{21}Q$ *are inside* Γ, *all eigenvalues of* $S_{11} - QS_{21}$ *are outside* Γ *and*

$$\|R - Q\| \leqq \varepsilon \max_{i,j=1,2} \|T_{ij} - S_{ij}\|. \tag{3.18}$$

Proof. Consider the operators

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}, \qquad S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{12} \end{bmatrix},$$

on $X = X_1 \oplus X_2$. Assume that X is endowed with the norm $\|(x_1, x_2)\| = \|x_1\| + \|x_2\|$. Then

$$\|T - S\| \leqq 2 \max_{i,j=1,2} \|T_{ij} - S_{ij}\|. \tag{3.19}$$

From Proposition 1.5 we know that $N_R = \{(Rx, x) | x \in X_2\}$ is a spectral subspace for T. In fact, if Γ is as in the statement of the theorem, then T has no eigenvalues on Γ and $N_R = \operatorname{Im} P(T; \Gamma)$.

Let ℓ be the length of Γ, and put $\gamma = \max_{\lambda \in \Gamma} \|(\lambda I - T)^{-1}\|$. Take $\|T - S\| < (2\gamma)^{-1}$. By elementary spectral theory this implies that S has no eigenvalues on Γ and

$$\|(\lambda I - T)^{-1} - (\lambda I - S)^{-1}\| \leqq 2\gamma^2 \|S - T\|, \qquad \lambda \in \Gamma.$$

But then $\|P(T; \Gamma) - P(S; \Gamma)\| \leqq \pi^{-1}\gamma^2 \ell \|S - T\|$.

As $X = X_1 \oplus N_R$, the number $\eta(X_1, N_R)$ is positive. Put

$$\omega = \min\left\{\frac{1}{4\gamma}, \frac{\pi}{4\gamma^2 \ell}\eta(X_1, N_R)\right\},$$

and assume that (3.16) holds true. By (3.19) this implies that $\|T - S\| < 2\omega \leqq (2\gamma)^{-1}$, and we can apply the result of the previous paragraph to show that

$$\|P(T; \Gamma) - P(S; \Gamma)\| < \tfrac{1}{2}\eta(X_1, N_R).$$

In particular we see that

$$\operatorname{gap}(N_R, \operatorname{Im} P(S; \Gamma)) < \tfrac{1}{2}\eta(X_1, N_R). \tag{3.20}$$

By Theorem 2 in [14] this implies that

$$X = X_1 \oplus \operatorname{Im} P(S; \Gamma).$$

It follows that there exists $Q \in \mathcal{L}(X_2, X_1)$ such that

$$N_Q = \{(Qz, z) | z \in X_2\} = \operatorname{Im} P(S; \Gamma).$$

By the results of § 1.3, this operator Q is a solution of (3.17), the eigenvalues of $S_{22} + S_{21}Q$ are inside Γ, and the eigenvalues of $S_{11} - QS_{21}$ are outside Γ.

By (3.20), we have $\operatorname{gap}(N_R, N_Q) < \frac{1}{2}\eta(X_1, N_R)$. So we can apply Lemma 3.2 to show that

$$\|R - Q\| \leqq \frac{2(1 + \|R\|)}{\eta(X_1, N_R)} \operatorname{gap}(N_R, N_Q). \tag{3.21}$$

But

(3.22) $$\operatorname{gap}(N_R, N_Q) \leqq \|P(T, \Gamma) - P(S; \Gamma)\| \leqq \frac{\gamma^2 \ell}{\pi} \|T - S\| \leqq 2 \frac{\gamma^2 \ell}{\pi} \max_{i,j=1,2} \|T_{ij} - S_{ij}\|.$$

Put

$$\varepsilon = 4(1 + \|R\|) \frac{\gamma^2 \ell}{\pi \eta(X_1, N_R)}.$$

Then we see from (3.21) and (3.22) that (3.18) holds true. This completes the proof of the theorem.

IV. Numerical and computational aspects. In this chapter we shall discuss some of the practical numerical aspects of minimal factorization of rational matrix functions. In contrast with the results obtained in earlier sections, the coordinate system becomes here of crucial importance. Indeed, for computational problems, the matrices A, B, and C determining the transfer function

(4.1) $$W(\lambda) = I_n + C(\lambda I_\delta - A)^{-1} B,$$

are known with a certain relative accuracy. Any coordinate transformation T, required to construct a factorization, causes a loss of accuracy which is proportional to

$$\operatorname{cond}(T) = \|T\| \cdot \|T^{-1}\|,$$

(cf. [24]). The number cond (T) is called the *condition number* of T.

THEOREM 3.7. *Let $\mathbb{C}^\delta = \mathbb{C}^{\delta_1} \oplus \mathbb{C}^{\delta_2}$ be the (Euclidean) direct sum of $\mathbb{C}^{\delta_1}$ and $\mathbb{C}^{\delta_2}$, and let $T_1: \mathbb{C}^{\delta_1} \to \mathbb{C}^\delta$ and $T_2: \mathbb{C}^{\delta_2} \to \mathbb{C}^\delta$ be operators. If T_1 and T_2 are isometries, then*

(4.2) $$\|T_1^* T_2\| = \cos \phi_{\min},$$

where $\phi_{\min}$ is the minimal angle between Im T_1 *and* Im T_2. *Moreover, if*

$$T = [T_1 T_2]: \mathbb{C}^{\delta_1} \oplus \mathbb{C}^{\delta_2} \to \mathbb{C}^\delta$$

is invertible, then

$$\operatorname{cond}(T) \geqq \frac{1 + \cos \phi_{\min}}{\sin \phi_{\min}},$$

equality occurring when T_1 and T_2 are isometries.

Proof. First assume that T_1 and T_2 are isometries. Put $Q_1 = T_1 T_1^*$. Then Q_1 is the orthogonal projection of $\mathbb{C}^\delta$ onto $M_1 = \operatorname{Im} T_1$. It is not difficult to prove that

$$\|T_1^* T_2\| = \sup_{0 \neq x \in M_2} \frac{\|Q_1 x\|}{\|x\|},$$

where $M_2 = \operatorname{Im} T_2$. Hence, by formula (3.7),

$$\|T_1^* T_2\| = [1 - \eta(M_1, M_2)^2]^{1/2}.$$

The equality (4.2) is now immediate from the definition of $\phi_{\min}$.

Suppose that T is invertible and that T_1 and T_2 are isometries. In order to determine $\|T\|$ and $\|T^{-1}\|$, we compute the spectrum of T^*T. With respect to the decomposition $\mathbb{C}^\delta = \mathbb{C}^{\delta_1} \oplus \mathbb{C}^{\delta_2}$, we have

$$\lambda I - T^* T = \begin{bmatrix} (\lambda - 1) I & -T_1^* T_2 \\ -T_2^* T_1 & (\lambda - 1) I \end{bmatrix}.$$

For $\lambda \neq 1$, one can write the right-hand side as a product of three operator matrices as follows:

$$(4.3)\quad \begin{bmatrix} \frac{1}{\lambda-1}I & -\frac{1}{(\lambda-1)^2}T_1^*T_2 \\ 0 & \frac{1}{\lambda-1}I \end{bmatrix} \cdot \begin{bmatrix} (\lambda-1)^2I - T_1^*T_2T_2^*T_1 & 0 \\ 0 & (\lambda-1)^2I \end{bmatrix} \cdot \begin{bmatrix} I & 0 \\ -\frac{1}{\lambda-1}T_2^*T_1 & I \end{bmatrix}.$$

In this way one sees that for $\lambda \neq 1$, the operator $\lambda I - T^*T$ is invertible if and only if $(\lambda-1)^2I - T_1^*T_2T_2^*T_1$ is invertible. It follows that

$$\|T\|^2 = 1 + \|T_1^*T_2\|, \qquad \|T^{-1}\|^{-2} = 1 - \|T_1^*T_2\|.$$

But $\|T_1^*T_2\| = \cos\phi_{\min}$, and hence

$$\operatorname{cond}(T)^2 = \frac{1+\cos\phi_{\min}}{1-\cos\phi_{\min}} = \frac{(1+\cos\phi_{\min})^2}{\sin^2\phi_{\min}}.$$

This proves the theorem for the case when T_1 and T_2 are isometries.

Finally we consider the general case, where T_1 and T_2 are arbitrary operators such that $T = [T_1 T_2]$ is invertible. Using polar decomposition, we may write $T_1 = U_1R_1$ and $T_2 = U_2R_2$, where U_1 and U_2 are isometries and R_1 and R_2 are strictly positive selfadjoint operators acting on $\mathbb{C}^{\delta_1}$ and $\mathbb{C}^{\delta_2}$, respectively. Put $S = [U_1U_2]$, and $R = \operatorname{diag}(R_1, R_2)$. Then R is invertible and $T^*T = RS^*SR$.

Set $\alpha = \cos\phi_{\min}$. Then $\alpha = \|U_1^*U_2\|$, and there exists $x \in \mathbb{C}^{\delta_1}$ such that $\|x\| = 1$ and $U_1^*U_2U_2^*U_1x = \alpha^2x$. Put

$$z_i = \begin{bmatrix} x \\ \frac{(-1)^i}{\alpha}U_2^*U_1x \end{bmatrix}, \qquad \lambda_i = 1 + (-1)^i\alpha, \quad i = 1, 2.$$

For $\lambda \neq 1$, we know that $\lambda I - S^*S$ is equal to the product (4.3), provided the operators T_1 and T_2 are replaced by U_1 and U_2, respectively. It follows that

$$S^*Sz_i = \lambda_iz_i, \qquad i = 1, 2.$$

Note that $\|R^{-1}z_1\| = \|R^{-1}z_2\| \geqq \|R\|^{-1} > 0$. So

$$\operatorname{cond}(T)^2 \geqq \frac{\|T^*TR^{-1}z_2\|}{\|T^*TR^{-1}z_1\|} = \frac{\|RS^*Sz_2\|}{\|RS^*Sz_1\|}$$
$$\geqq \frac{\lambda_2\|Rz_2\|}{\lambda_1\|Rz_1\|} = \frac{\lambda_2}{\lambda_1} = \frac{1+\cos\phi_{\min}}{1-\cos\phi_{\min}} = \frac{(1+\cos\phi_{\min})^2}{\sin^2\phi_{\min}},$$

and the proof is complete.

The preceding theorem sheds some light on the numerical problem of computing minimal factorizations of rational matrix functions. Consider the realization (4.1). We assume that the realization is minimal. From Theorem 2.2 we know that there is a one-to-one correspondence between the minimal factorizations of $W(\lambda)$ and the supporting projections of the (minimal) system $\theta = (A, B, C; \mathbb{C}^\delta, \mathbb{C}^n)$. In turn these supporting projections are completely determined by pairs of subspaces M, $M^\times$ satisfying

$$AM \subset M, \quad A^\times M^\times \subset M^\times, \quad \mathbb{C}^\delta = M \oplus M^\times. \tag{4.4}$$

Here, as usual, $A^\times = A - BC$.

For the computation of invariant subspaces of a matrix, reliable algorithms are available in the literature [25]. A common way to proceed is to construct a unitary matrix Q_1 such that

$$A_S = Q_1^* A Q_1 = \begin{bmatrix} \alpha_1 & * & \cdots & * \\ 0 & \ddots & & \vdots \\ \vdots & & & * \\ 0 & \cdots & 0 & \alpha_\delta \end{bmatrix} \tag{4.5}$$

is in upper Schur form. The diagonal elements $\alpha_1, \cdots, \alpha_\delta$ of A_S are the poles of $W(\lambda)$. Similarly, one can construct a unitary matrix Q_2 which transforms $A^\times$ to lower Schur form:

$$Q_2^* A^\times Q_2 = A_S^\times = \begin{bmatrix} \beta_1 & 0 & \cdots & 0 \\ * & \ddots & & \vdots \\ \vdots & & & 0 \\ * & \cdots & * & \beta_\delta \end{bmatrix}. \tag{4.6}$$

Here $\beta_1, \cdots, \beta_\delta$ are the zeros of $W(\lambda)$. Algorithms that perform these decompositions are known as the QR and QL algorithms [25].

Given natural numbers δ_1 and δ_2 for which $\delta = \delta_1 + \delta_2$, we partition Q_1 and Q_2 as follows,

$$Q_1 = [\underbrace{V_1}_{\delta_1} \vdots \underbrace{W_1}_{\delta_2}], \quad Q_2 = [\underbrace{V_2}_{\delta_1} \vdots \underbrace{W_2}_{\delta_2}].$$

From (4.5) and (4.6) it is clear that the columns of V_1 and W_2 form orthonormal bases for invariant subspaces M and $M^\times$ of A and $A^\times$, respectively. Now $\mathbb{C}^\delta = M \oplus M^\times$ if and only if the minimal angle $\phi_{\min}$ between M and $M^\times$ is nonzero. Thus M and $M^\times$ satisfy (4.4) if and only if $\phi_{\min} > 0$. By Theorem 3.7 we have $\cos \phi_{\min} = \|V_1^* W_2\|$. Therefore, defining the matrix Q by $Q = Q_1^* Q_2$ and partitioning it as follows

$$Q = \begin{array}{c} \left[\begin{array}{c:c} Q_{11} & Q_{12} \\ \hdashline Q_{21} & Q_{22} \end{array}\right] \begin{array}{l} \}\delta_1 \\ \}\delta_2, \end{array} \\ \begin{array}{cc} \delta_1 & \delta_2 \end{array} \end{array}$$

one can measure $\phi_{\min}$ from the block $Q_{12} = V_1^* W_2$. Indeed, whenever the norm of Q_{12} is smaller than one, the spaces M and $M^\times$ yield a supporting projection of the system θ, and, consequently, a minimal factorization $W(\lambda) = W_1(\lambda) W_2(\lambda)$ of $W(\lambda)$. Observe that δ_1 and δ_2 are the McMillan degrees of W_1 and W_2, respectively.

In order to determine the factors W_1 and W_2 we put $T = [V_1 \vdots W_2]$. If $\mathbb{C}^\delta = M \oplus M^\times$, the matrix T is invertible. But then the system $(T^{-1}AT, T^{-1}B, CT; \mathbb{C}^\delta, \mathbb{C}^n)$ is similar to the system θ and has $W(\lambda)$ as its transfer function. One easily verifies that the matrices $T^{-1}AT$, $T^{-1}B$, and CT admit a partitioning of the following type

$$T^{-1}AT = \begin{array}{c} \left[\begin{array}{c:c} A_1 & B_1C_2 \\ \hdashline 0 & A_2 \end{array}\right] \begin{array}{l} \}\delta_1 \\ \}\delta_2, \end{array} \\ \begin{array}{cc} \delta_1 & \delta_2 \end{array} \end{array} \quad T^{-1}B = \left[\begin{array}{c} B_1 \\ \hdashline B_2 \end{array}\right] \begin{array}{l} \}\delta_1 \\ \}\delta_2, \end{array} \quad CT = [\underbrace{C_1}_{\delta_1} \vdots \underbrace{C_2}_{\delta_2}].$$

Put $\theta_1 = (A_1, B_1, C_1; \mathbb{C}^{\delta_1}, \mathbb{C}^n)$ and $\theta_2 = (A_2, B_2, C_2; \mathbb{C}^{\delta_2}, \mathbb{C}^n)$. Then $\theta = \theta_1\theta_2$. The factors W_1 and W_2 are now the transfer functions of θ_1 and θ_2, respectively. The poles of W_1 are the first δ_1 diagonal entries in A_S; the zeros of W_1 are the first δ_1 diagonal entries in $A_S^\times$. A similar remark can be made about W_2.

The smaller cond (T) is, the more accurate the constructed factors W_1 and W_2 will be. This shows the significance of Theorem 3.7. Indeed, the similarity transformation T constructed above maps $\mathbb{C}^{\delta_1}\oplus(0)$ onto M and $(0)\oplus\mathbb{C}^{\delta_2}$ onto $M^\times$. By Theorem 3.7, a lower bound for the condition number of a transformation having this property is given by the number

$$\frac{1+\cos\phi_{\min}}{\sin\phi_{\min}}.$$

In the present situation cond (T) is actually equal to this bound, for V_1 and W_2 are isometries. So in this respect T is optimal. On the other hand, for a very small angle $\phi_{\min}$, the condition number of T will be very large. In that case one can expect a very bad relative accuracy that may even exceed 1. This will occur whenever $\phi_{\min}$ is smaller than a certain threshold ϕ_0 which depends on the accuracy of the data. Therefore the spaces M and $M^\times$ cannot be used when their minimal angle is too small. If that happens, one can try different choices of δ_1 and δ_2, while using the same matrices Q_1 and Q_2. Also one can try other Schur decompositions of A and $A^\times$.

For the amount of computations involved in the construction of a factorization of the transfer function (4.1), we can give the following rough estimates, where 1 operation stands for 1 multiplication plus 1 addition:

$\delta^2 n$	operations for constructing $A^\times$,
$20\delta^3$	operations for each Schur decomposition,
δ^3	operations for the product $Q = Q_1^* Q_2$,
$\frac{10}{3}\delta_1^2\delta_2$	operations (if $\delta_1 < \delta_2$) for calculating $\cos\phi_{\min}$,
$2\delta^2(n+\delta)$	operations for computing A_1, B_1, C_1, A_2, B_2 and C_2 if $\phi_{\min} > \phi_0$.

In general, the Schur decompositions constitute the most time-consuming step, but for $\delta = 100$, e.g., experiments have yielded run times that are still within the orders of seconds [9].

As we have seen, the determination of minimal factorizations is closely related to that of pairs of "matching" invariant subspaces. The number of invariant subspaces involved may be very large or even infinite. In practice this may lead to very cumbersome combinatorial problems. For more details, the reader is referred to [21].

REFERENCES

[1] N. I. AHIEZER AND I. M. GLAZMAN, *Theorie der Linearen Operatoren im Hilbertraum*, Akademie Verlag, Berlin, 1960.

[2] S. BARNETT, *Introduction to Mathematical Control Theory*, Clarendon Press, Oxford, 1975.

[3] H. BART, I. GOHBERG AND M. A. KAASHOEK, *Operator polynomials as inverses of characteristic functions*, Integral Equations Operator Theory, 1 (1978), pp. 1–12.

[4] ———, *Stable factorizations of monic matrix polynomials and stable invariant subspaces*, Integral Equations Operator Theory, 1 (1978), pp. 496–517.

[5] ———, *Minimal factorizations of matrix and operator functions*, to appear.

[6] V. BELEVITCH, *Classical Network Theory*, Holden-Day, San Francisco–Cambridge–Amsterdam, 1968.

[7] M. S. BRODSKII, *Triangular and Jordan representations of linear operators*, Transl. Math. Monographs, Vol. 32, American Mathematical Society, Providence RI, 1970.

[8] S. CAMPBELL AND J. DAUGHTRY, *The stable solutions of quadratic matrix equations*, Proc. Amer. Math. Soc., 74 (1979), pp. 19–23.

[9] EISPACK (Guide), *Matrix eigensystem routines*, 2nd Ed., Springer-Verlag, Berlin, 1976.

[10] I. C. GOHBERG AND A. I. FELDMAN, *Convolution equations and projection methods for their solution*, Transl. Math. Monographs, Vol. 41, American Mathematical Society, Providence RI, 1974.

[11] I. C. GOHBERG AND M. G. KREIN, *Systems of integral equations on a half line with kernels depending on the difference of arguments*, Uspehi Mat. Nauk, 13 (1958), pp. 3–72 (in Russian) translated as Amer. Math. Soc. Transl., (2) 14 (1960), pp. 217–287.

[12] I. GOHBERG, P. LANCASTER AND L. RODMAN, *Spectral analysis of matrix polynomials I. Canonical forms and divisors*, Linear Algebra Appl., 20 (1978), pp. 1–44.

[13] ——, *Spectral analysis of matrix polynomials II. The resolvent form and spectral divisors*, Linear Algebra Appl., 21 (1978), pp. 65–88.

[14] I. C. GOHBERG AND A. S. MARKUS, *Two theorems on the gap between subspaces of a Banach space*, Upsehi Mat. Nauk, 14 (1959), pp. 135–140 (in Russian).

[15] I. C. GOHBERG AND E. I. SIGAL, *An operator generalization of the logarithmic residue theorem and the theorem of Rouché*, Mat. Sbornik, 84 (1971), pp. 607–629 (in Russian), transl. as Math. USSR Sbornik, 13 (1971), pp. 603–625.

[16] R. E. KALMAN, P. L. FALB AND M. A. ARBIB, *Topics in Mathematical System Theory*, McGraw-Hill, New York, 1966.

[17] M. G. KREIN, *Introduction to the geometry of indefinite J-spaces and to the theory of operators in these spaces*, Amer. Math. Soc. Transl., (2) Vol. 93, 1970.

[18] H. KWAKERNAAK AND R. SIVAN, *Linear Optimal Control Systems*, Wiley-Interscience, New York, 1972.

[19] B. P. MOLINARI, *Equivalence relations for the algebraic Riccati equation*, SIAM J. Control, 11 (1973), pp. 272–285.

[20] L. A. SAHNOVIČ, *On the factorization of an operator-valued transfer function*, Soviet Math. Dokl., 17 (1976), pp. 203–207.

[21] P. VAN DOOREN AND P. DEWILDE, *Minimal factorization of rational matrices*, Proceedings of the 17th IEEE Conference on Decision Control, 1979, pp. 170–172.

[22] P. DEWILDE, *Cascade Scattering Matrix Synthesis*, Tech. Rep. 6560-21. Information System Lab., Stanford University, Stanford CA, 1970.

[23] P. DEWILDE AND J. VANDEWALLE, *On the factorization of a non-singular rational matrix*, IEEE Trans. Circuits and Systems, CAS-22 (1975), pp. 637–645.

[24] J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.

[25] J. H. WILKINSON AND C. REINSCH, *Handbook for Automatic Computation, Vol. II: Linear Algebra*, Springer-Verlag, New York, 1971.

[26] D. C. YOULA AND P. TISSI, *An explicit formula for the degree of a rational matrix*, Electrophysics Memo. PIBM RI-1273-65, Polytechnic Institute of Brooklyn, Electrophysics Department, Brooklyn NY, 1965.